



# 计算机科学

COMPUTER SCIENCE

## 基于改进DeepLabv3+的遥感影像道路提取算法

王谦, 何朗, 王展青, 黄坤

引用本文

王谦, 何朗, 王展青, 黄坤. 基于改进DeepLabv3+的遥感影像道路提取算法[J]. 计算机科学, 2024, 51(8): 168-175.

WANG Qian, HE Lang, WANG Zhanqing, HUANG Kun. Road Extraction Algorithm for Remote Sensing Images Based on Improved DeepLabv3+ [J]. Computer Science, 2024, 51(8): 168-175.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

### [基于注意力机制的CNN和BiGRU的加密流量分类](#)

Encrypted Traffic Classification of CNN and BiGRU Based on Self-attention  
计算机科学, 2024, 51(8): 396-402. <https://doi.org/10.11896/jsjcx.230500032>

### [基于知识图谱与邻域感知注意力机制的推荐算法研究](#)

Study on Recommendation Algorithms Based on Knowledge Graph and Neighbor PerceptionAttention Mechanism  
计算机科学, 2024, 51(8): 313-323. <https://doi.org/10.11896/jsjcx.230500143>

### [基于RoBERTa和加权图卷积网络的中文地质实体关系抽取](#)

Chinese Geological Entity Relation Extraction Based on RoBERTa and Weighted Graph Convolutional Networks  
计算机科学, 2024, 51(8): 297-303. <https://doi.org/10.11896/jsjcx.230600231>

### [基于多模态注意力网络的红外人体行为识别方法](#)

Infrared Human Action Recognition Method Based on Multimodal Attention Network  
计算机科学, 2024, 51(8): 232-241. <https://doi.org/10.11896/jsjcx.230600143>

### [基于多样化标签矩阵的医学影像报告生成](#)

Diversified Label Matrix Based Medical Image Report Generation  
计算机科学, 2024, 51(8): 200-208. <https://doi.org/10.11896/jsjcx.230600018>

# 基于改进 DeepLabv3+ 的遥感影像道路提取算法

王 谦 何 朗 王展青 黄 坤

武汉理工大学理学院 武汉 430070

(chanwj@whut.edu.cn)

**摘 要** 道路提取可以帮助人们更好地理解城市环境,是城市交通和城市规划等方面的重要部分,随着深度学习与计算机视觉的发展,利用基于深度学习的语义分割算法从遥感影像中提取道路的技术趋于成熟。针对现有的深度学习道路提取算法存在的提取速度慢和容易受背景环境因素干扰而产生漏分割、不连续等问题,提出了一种基于 ECANet 注意力机制和级联空洞空间金字塔池化模块的轻量化算法 CE-DeepLabv3+。首先,将主干特征提取网络更换为轻量级的 MobileNetv2,减少参数量,提高模型的执行速度;其次,通过增加空洞空间金字塔池化模块的卷积层进一步扩大感受野,再级联不同特征层来增强语义信息的复用性,从而加强对细节特征的提取能力;再次,加入 ECANet 注意力机制,抑制背景环境中的干扰因素,聚焦道路信息;最后,采用改进的损失函数进行训练,消除了道路与背景样本不均衡对模型性能产生的影响。实验结果表明,改进算法的性能优良,与原始 DeepLabv3+ 算法相比,在分割效率、分割精度上有较大的提升。

**关键词:** 语义分割;遥感影像;道路提取;注意力机制;DeepLabv3+;级联空洞空间金字塔池化

**中图分类号** TP319

## Road Extraction Algorithm for Remote Sensing Images Based on Improved DeepLabv3+

WANG Qian, HE Lang, WANG Zhanqing and HUANG Kun

School of Science, Wuhan University of Technology, Wuhan 430070, China

**Abstract** Road extraction can help us better understand the urban environment and is an important part of urban transportation and planning. With the development of deep learning and computer vision, the use of deep learning-based semantic segmentation algorithm to extract roads from remote sensing images has become increasingly mature. However, existing deep learning road extraction algorithms suffer from slow extraction speed and susceptibility to background environmental factors, resulting in missed segmentation and discontinuity. To address these issues, a lightweight algorithm called CE-DeepLabv3+ based on ECANet attention mechanism and cascade atrous spatial pyramid pooling module is proposed. Firstly, the main feature extraction network is replaced with the lightweight MobileNetv2 to reduce parameter volume and improve model execution speed. Secondly, the convolution layers of the atrous spatial pyramid pooling module are further expanded to increase the receptive field, and different feature layers are cascaded to enhance semantic information reuse, thereby improving the ability to extract fine-grained features. Thirdly, the ECANet attention mechanism is added to suppress environmental interference and focus on road information. Finally, an improved loss function is used for training to address the impact of road and background sample imbalance on model performance. Experimental results show that the improved algorithm achieves excellent performance, with significant improvements in segmentation efficiency and accuracy compared to the original DeepLabv3+ algorithm.

**Keywords** Semantic segmentation, Remote sensing images, Road extraction, Attention mechanism, DeepLabv3+, Cascade atrous spatial pyramid pooling

## 1 引言

道路提取在城市规划、环境检测、城市管理和导航地图制作等应用领域意义重大<sup>[1]</sup>。随着世界城市化的快速发展,道路信息更新频率也越来越快,因此对道路信息进行实时和准确提取具有关键的作用。早期传统的人工道路提取方法包括

人工数字化绘图法、视觉解译法以及特征提取法,这些方法都需要耗费大量的人力和物力,而且很容易受到主观因素的干扰,导致提取结果不准确,并且提取速度不能跟上道路更新的频率。遥感影像中具有很强的道路特征,目前获取遥感影像的方式越来越方便,遥感影像受地面条件限制因素小,并且实时传输,探测范围大,非常适合于进行道路提取任务。遥感

到稿日期:2023-06-14 返修日期:2023-08-31

基金项目:国家自然科学基金(62176194)

This work was supported by the National Natural Science Foundation of China(62176194).

通信作者:何朗(helang@whut.edu.cn)

影像还具有多光谱、高分辨率等特点,可以提供更丰富、准确的信息,有助于提高道路信息提取的精度和效率,利用图像分割方法可以快速提取遥感影像中大范围区域内的道路信息,而且不需要实地勘测,可以节约大量的时间和人力成本<sup>[2]</sup>。

传统的图像分割方法包括基于区域的方法<sup>[3]</sup>、基于对象的方法<sup>[4]</sup>、基于图切割的方法<sup>[5]</sup>和基于聚类的方法<sup>[6]</sup>等。Lopez等<sup>[7]</sup>提出了一种基于行的遥感影像分割方法,即利用数据驱动自动分割;Wan等<sup>[8]</sup>提出了一种基于轨迹数据提取道路的方法,该方法采用决策树并结合移动开窗式的轨迹线分割模型进行道路交叉口提取;Cheng等<sup>[9]</sup>提出了一种基于图切割的概率传播方法,用于从复杂的遥感图像中提取道路,该方法在收集的两个数据集上的定性和定量比较中都取得了较好的表现;Maurya等<sup>[10]</sup>提出了一种聚类方法来进行道路分割,该方法可以方便地提取道路,并且在少量简单图像的情况下给出了令人满意的结果。但是,以上方法普遍存在耗时长、需要大量人工干预、算法精度差及适用性有限等缺点。

随着深度学习的快速发展,基于卷积神经网络的图像语义分割技术得到了广泛的应用。深度学习模型可以自动地从大量的遥感影像数据中学习道路的特征,并且可以实现端到端的道路提取,从而弥补传统方法中需要手动调整参数和特征的缺点。此外,深度学习模型还具有较高的准确性和效率,在对大规模遥感影像进行道路提取时表现出了明显的优势。2016年 Zhong等<sup>[11]</sup>利用全卷积网络(FCN)从高分辨率图像中提取道路,在 Massachusetts 道路数据集上提取精度超过了 78%。Chen等<sup>[12]</sup>提出了一种针对遥感图像中道路提取任务的重建偏差 U-Net,该方法增加了解码分支,从不同的上采样中获取多个语义信息,在 3 个不同数据集上进行测试,获得了比其他 6 种先进分割模型更好的效果。Badrinarayanan等<sup>[13]</sup>提出了一种基于编码器-解码器的 SegNet 网络,并且在道路场景中获得了良好的推理时间和分割精度。Li等<sup>[14]</sup>针对道路狭长结构导致提取困难的特点,提出了一种改进的 ResNet 道路提取方法,通过构建条形池化模块和混合池化模块,提升了道路提取的效果。Poudel等<sup>[15]</sup>提出了一种快速分割卷积神经网络 Fast-SCNN,能对高分辨率图像数据进行实时语义分割,适用于嵌入式设备的高效计算。虽然以上方法解决了人工干预的问题,为道路的自动提取做出了一定贡献,但是它们仍然存在漏分割、不连续等问题,在提取精度上仍然有进步的空间。

近年来,由谷歌提出的 DeepLab 系列算法得到了广泛的关注,其提出了空洞空间金字塔池化(Atrous Spatial Pyramid Pooling, ASPP)模块,通过空洞卷积扩大了感受野,提高了语义分割的准确性。DeepLabv3+算法<sup>[16]</sup>重新引入了编码器-解码器结构,将浅层特征与深层特征融合,提高了分割精度。虽然 DeepLabv3+ 在一些公共数据集上表现良好,但是对于遥感影像的道路提取任务,DeepLabv3+ 中主干特征提取网络 Xception 的模型参数量较大,需要花费大量的训练时间和计算量,并不适用于道路提取这种目标种类少的提取任务,并且由于道路环境中背景复杂,存在很多干扰因素,也容易出现细节部分提取效果较差的问题。

针对以上问题,本文提出了一种改进的 DeepLabv3+ 算法,

用于遥感图像中的道路提取任务,称为 CE-DeepLabv3+。首先,采用轻量级的 MobileNetv2<sup>[17]</sup> 网络作为主干特征提取网络,降低模型的参数量,加快分割速度。其次,对 ASPP 模块进行改进,增加一层卷积层并级联相邻的卷积层,使网络能够获得更密集的像素采样,从而增强网络提取细节特征的能力,加强了特征复用和表达能力。此外,再在改进的 ASPP 模块中的每层卷积和浅层特征后加入 ECANet 注意力机制<sup>[18]</sup> 来抑制无用信息,使模型更关注道路信息,从而提升分割精度。最后,为解决道路场景中道路与背景类别不均衡的问题,采用改进的损失函数进行训练,通过增加道路信息的权重来引导模型更加关注难学习到的道路特征,进一步提升了模型性能。

## 2 DeepLabv3+

DeepLabv3+ 算法<sup>[16]</sup> 是 Chen 等于 2018 年提出的一种先进的语义分割算法,具有出色的图像分割能力,已经被广泛应用于遥感图像的分割任务中。DeepLabv3+ 算法通过 ASPP 模块中的空洞卷积有效地扩大了网络感受野并捕获了广泛的上下文信息,提高了图像分割的精度。空洞卷积的等效卷积核  $K$  的计算式为:

$$K = k + (k-1)(r-1) \quad (1)$$

其中,  $k$  为原始卷积核大小,  $r$  为空洞率。

由主干特征提取网络提取到一个低阶特征和一个高阶特征,高阶特征通过 ASPP 模块捕获不同尺度的信息,其中包含 1 层  $1 \times 1$  卷积、3 层不同空洞率的  $3 \times 3$  卷积和 1 层池化层,最后将并行提取的特征进行融合,融和后的特征图包含了丰富的语义信息,再通过  $1 \times 1$  卷积调整通道数,在解码器中通过上采样操作与低阶特征进行融合,再通过  $3 \times 3$  卷积和上采样操作恢复原来的大小,得到分割结果。整体的网络结构如图 1 所示。

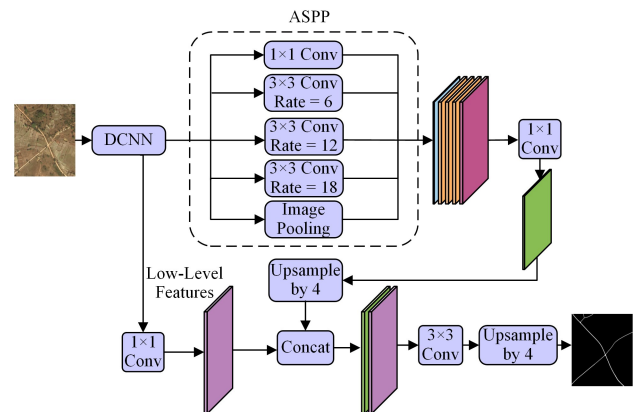


图 1 DeepLabv3+ 算法的结构

Fig. 1 Structure of DeepLabv3+ algorithm

## 3 CE-DeepLabv3+

由于道路信息在遥感图像中所占像素比例较小,且容易受到背景环境因素的干扰,因此对算法的细节提取能力的要求很高。为解决 DeepLabv3+ 模型参数量大,且对道路易产生漏分割和不连续的问题,本文对 DeepLabv3+ 模型的主干特征提取网络进行轻量化,再对 ASPP 模块进行改进,最后在改进的 ASPP 模块中的每层卷积和浅层特征后加入 ECANet

注意力机制,使分割速度和分割精度得到有效提升。改进后的 CE-DeepLabv3+算法的网络结构如图 2 所示。

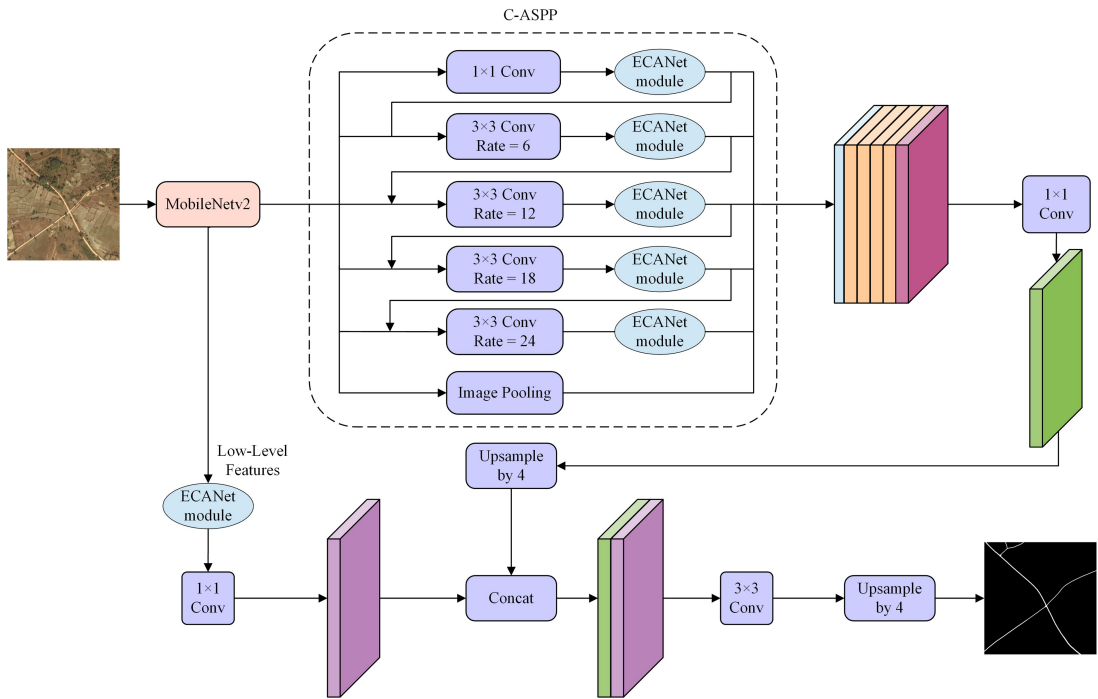


图 2 CE-DeepLabv3+算法的结构

Fig. 2 Structure of CE-DeepLabv3+ algorithm

### 3.1 轻量化的主干特征提取网络

由于原 DeepLabv3+算法中的 Xception 网络结构非常复杂,这对于道路提取这种小目标的提取任务来说,可能会导致网络在提取细节和纹理信息时存在一些模糊和混淆的问题,从而出现不连续现象,并且 Xception 庞大的参数量需要消耗大量的计算资源与时间。本文采用轻量级的 MobileNetv2<sup>[17]</sup> 结构作为主干特征提取网络,相对于 Xception 参数量更少,训练速度更快,MobileNetv2 可以更快地提取道路图像中的特征,使模型更容易捕捉到道路的细节和纹理,更适用于道路提取任务。

MobileNetv2<sup>[17]</sup> 是一种轻量级的卷积神经网络,其主要是通过深度可分离卷积来减少参数量,从而能够在保持模型精度的同时下减少计算量。深度可分离卷积是一种卷积神经网络中的优化方法,可以减少计算量并提升性能。与传统的卷积层不同,深度可分离卷积将卷积层拆分为两个独立的操作,即深度卷积和逐点卷积,深度卷积通过对每个输入通道应用单个卷积滤波器来执行轻量级滤波,逐点卷积负责通过计算输入通道的线性组合来构建新特征。当输入特征图大小为  $h_i \times w_i \times d_i$ , 卷积核尺寸为  $k \times k \times d_i \times d_j$  时,标准卷积的计算量为  $h_i \cdot w_i \cdot d_i \cdot d_j \cdot k \cdot k$ , 而深度可分离卷积的计算成本仅为  $h_i \cdot w_i \cdot d_i \cdot (k^2 + d_j)$ 。MobileNetv2 使用了  $k=3$  ( $3 \times 3$  深度可分离卷积), 因此计算成本为标准卷积的  $1/9 \sim 1/8$ , 而精度仅略有降低。深度可分离卷积的过程如图 3 所示。

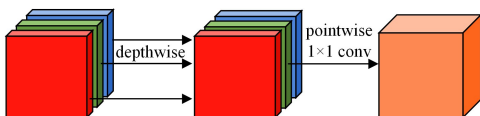


图 3 深度可分离卷积过程示意图

Fig. 3 Schematic diagram of deep separable convolution process

### 3.2 改进的 ASPP 模块

DeepLabv3+算法中的 ASPP 模块利用空洞卷积来获取更大的感受野,但是对于道路这种小目标提取任务,容易出现细节特征提取不足、特征丢失等现象,从而造成漏分割等问题。因此,本文对 ASPP 模块进行改进,提出了一种级联的空洞空间金字塔池化(Cascade Atrous Spatial Pyramid Pooling, C-ASPP)模块,增加一层卷积层,从而进一步提取更多不同尺度的特征信息,提高模型对小目标的感知能力,再通过级联不同的分支,将前一层的特征输出与后一层的特征输入融合并将其作为后一层的特征输入,来融合多级跨尺度特征,从而获取更全面的语义信息,增强语义信息的复用性,有利于加强对道路细节特征的提取。C-ASPP 的结构如图 4 所示。

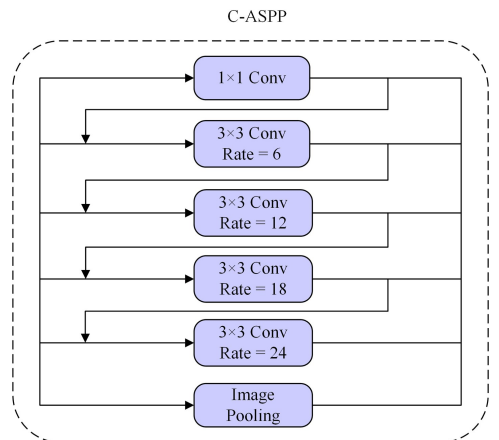


图 4 C-ASPP 的结构

Fig. 4 Structure of C-ASPP

C-ASPP 结构增加了不同卷积层之间的特征相关性。在一维空洞卷积中,卷积运算的计算式为:

$$f[i] = \sum_{k=1}^k x[i+r \cdot k]w[k] \quad (2)$$

其中,定义一维输入为  $x[i]$ ,  $k$  为卷积核大小,  $r$  为空洞率,  $w[k]$  表示卷积核中第  $k$  个参数,  $f[i]$  为一维空洞卷积的输出。

空洞卷积的感受野大小的计算式为:

$$F = (k-1)(r-1) + k \quad (3)$$

当空洞率为 6、卷积核大小为 3 时,一维空洞卷积的感受野大小为 13,但是每次只有 3 个像素参与卷积运算,这种情况在二维卷积中会更局限。而在本文提出的 C-ASPP 模块中,将前一层的特征输出与后一层的特征输入融合,后一层的空洞卷积中将有 7 个像素参与卷积运算,可以获得更加密集的像素采样,从而提高像素利用率,更利于提取小目标道路信息。图 5 给出了一维空间下普通空洞卷积和级联的空洞卷积的像素采样差异。

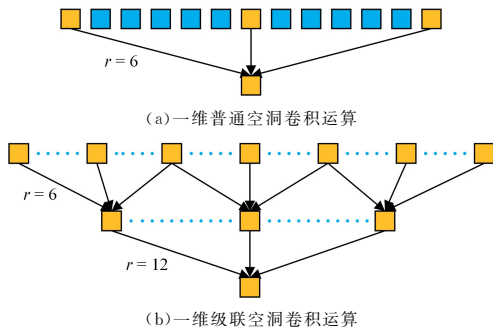


图 5 空洞卷积示意图

Fig. 5 Schematic diagram of atrous convolution

### 3.3 注意力模块 ECANet

遥感影像中的道路环境特征往往非常复杂,存在建筑物、树木等物体遮挡道路的情况,导致模型难以正确识别道路,因此道路提取任务变得非常具有挑战性。ECANet<sup>[18]</sup> (Efficient Channel Attention Module) 是一种高效通道注意力模块,它可以让网络更加关注重要的特征通道,从而更好地捕捉图像中的特征信息,并且它还提出了一种不降维的局部跨通道交互策略,有效避免了降维对通道注意力学习效果的影响。如图 6 所示,ECANet 主要由 3 部分组成。首先,通过全局平均池化获得大小为  $1 \times 1 \times C$  的特征图,然后自适应确定内核大小  $k$ 。

$$k = \psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (4)$$

其中,  $\lfloor t \rfloor_{\text{odd}}$  表示最接近  $t$  的奇数,  $\gamma$  和  $b$  分别设置为 2 和 1。

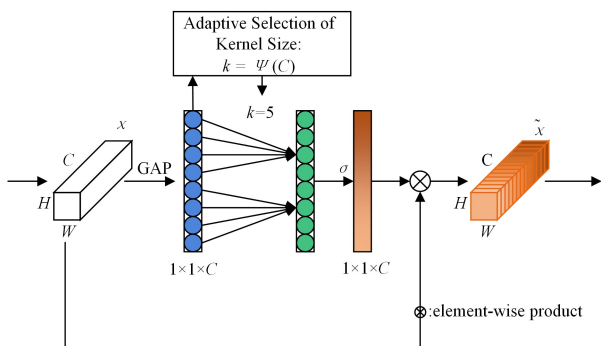


图 6 ECANet 注意力模块的结构

Fig. 6 Structure of ECANet attention module

最后执行一维卷积,通过 Sigmoid 函数学习通道注意力。其表达式如下:

$$wx = \sigma(C1D_k(y))x \quad (5)$$

其中,  $\sigma$  代表 Sigmoid 函数,  $C1D$  代表一维卷积,  $y$  代表  $x$  的池化输出。

本文在低阶特征和 C-ASPP 中每层卷积之后加入 ECA-Net 注意力机制,使得模型更好地聚焦于遥感图像中的道路信息,并抑制无效且复杂多样的背景特征,从而减少复杂的背景因素对道路提取的干扰,提高模型的准确性。

### 3.4 改进的损失函数

Dice loss 是图像分割任务中广泛使用的损失函数,常被用于二分类任务。对于遥感影像的道路提取任务,由于道路占整幅图像的像素比例较小,Dice loss 的优点在于它对类别不平衡的问题比较敏感,并且还能够减轻边界模糊现象,因此它适用于道路提取任务。Dice loss 定义为:

$$L_d = 1 - \frac{2|X \cap Y| + 1}{|X| + |Y| + 1} \quad (6)$$

其中,  $X$  为预测结果的分割区域,  $Y$  为实际标签区域,  $X \cap Y$  为预测结果与真实标签区域的重合部分。

但是,在 Dice Loss 计算中,如果预测值和真实值相差很大,则梯度会变得非常小,这使得模型难以学习。Focal loss 也是一个针对类别不平衡问题而设计的一种损失函数,它可以通过降低易分类样本的权重来引导模型更加关注难学习到的道路信息,还可以引入 Focal Loss 中的调节因子来缓解 Dice Loss 的梯度消失问题。Focal loss 定义为:

$$L_f = -\alpha(1-p_t)^\gamma \log(p_t) \quad (7)$$

其中,  $p_t$  为分类结果的置信度,越接近 1,则越容易区分,其权重越小;  $\alpha$  和  $\gamma$  为权重的控制程度参数,  $\gamma=0$  时,即为交叉熵损失。本文中取  $\gamma=0, \gamma=2$ 。

因此,本文结合了 Dice loss 和 Focal loss 两者的优势,可以有效解决遥感影像中道路提取任务的类别不平衡和梯度消失问题,同时提高模型的精度和鲁棒性。改进的损失函数 Loss 定义为:

$$L = L_d + \lambda L_f \quad (8)$$

其中,  $L_d$  为 Dice loss,  $L_f$  为 Focal loss,  $\lambda$  表示权重系数。

## 4 实验结果与分析

### 4.1 实验环境与数据集

本文实验是在 Linux 系统下搭建的 Pytorch 深度学习框架上进行的,其软硬件详细环境配置如表 1 所列。

表 1 实验环境配置

实验环境	配置说明
硬件环境	CPU: Intel(R) Xeon(R) Platinum 8225C@2.50 GHz GPU: NVIDIA RTX 2080 Ti(11GB) 内存: 40 GB
操作系统	Linux 操作系统
开发环境	Python 3.8
开发框架	Pytorch 深度学习框架

训练过程中的参数设置如下:学习率为  $5 \times 10^{-4}$ ,每批次输入大小为 8,训练迭代次数为 100,使用 cos 学习率下降方式选择了 Adam 优化器,相比其他优化器,Adam 优化器能够

同时考虑一阶动量和二阶动量,并动态调整每个参数的学习率,从而使模型更快收敛至性能最优。

实验所用数据集来源于 DeepGlobe 卫星图像理解挑战赛<sup>[19]</sup>中的道路数据集。该数据集包含了来自不同地区、不同季节和不同天气条件下的高分辨率卫星图像,包括街区、农村、高速公路、山区等多个场景,是道路提取任务中广泛应用的数据集之一。其中包含了 6226 张大小为  $1024 \times 1024$  像素的带标签的图像,空间分辨率为 0.5m。按照 7:2:1 的比例随机划分数据集,得到了 4314 张训练图像、1289 张验证图像和 623 张测试集图像进行实验。

## 4.2 评价指标

评价指标用于评价语义分割模型的检测效果。在遥感影像目标提取领域,常用的评价指标包括准确率(Accuracy)、平均像素精度(mPA)、平均交并比(mIoU)等。其计算式分别如式(9)一式(11)所示:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (9)$$

其中,TP 代表被正确识别的道路的像素点数量,TN 代表被正确识别的背景的像素点数量,FP 代表被错误识别为道路的背景像素点数量,FN 代表被错误识别为背景的道路像素点数量。

$$mPA = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}} \quad (10)$$

其中, $k+1$  为类别数, $p_{ii}$  为正确分类的像素数, $p_{ij}$  为  $i$  类被预测为  $j$  类的像素数, $p_{ji}$  为  $j$  类被预测为  $i$  类的像素数。PA 表示计算每个类中正确分类的像素数量与像素总数之间的比率,mPA 则表示计算每个类的 PA,然后再累加求平均。

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (11)$$

## 4.3 消融实验

### 4.3.1 改进的损失函数的有效性

首先探究改进的损失函数中不同权重系数  $\lambda$  对 DeepLabv3+算法分割精度的影响,从而确定  $\lambda$  的最优值并验证改进的损失函数的有效性,在 DeepLabv3+算法的基础上进行 5 组实验,逐渐增加 Focal loss 的权重。其中,当  $\lambda$  为 0 时,表示单独使用 Dice Loss 损失函数进行训练。实验结果显示,将 Dice Loss 与 Focal loss 结合后分割精度均高于单独使用 Dice Loss 损失函数,且当  $\lambda$  取 0.6 时,得到的 mPA, Accuracy 和 mIoU 的值最高,分割效果最好,因此本文取  $\lambda$  为 0.6。根据上述内容可以得出结论,使用改进的损失函数来处理样本不平衡问题的效果比单独使用 Dice Loss 损失函数的效果更佳。

表 2 DeepGlobe 道路数据集上不同比例 Focal loss 的性能结果

Table 2 Performance results of Focal loss with different proportions on DeepGlobe road dataset

Loss Function	$\lambda$	mPA/%	Accuracy/%	mIoU/%
$L = L_d + \lambda L_f$	0.0	87.01	97.27	76.81
	0.2	87.13	97.31	76.90
	0.4	87.35	97.53	77.35
	0.6	<b>87.54</b>	<b>97.68</b>	<b>77.60</b>
	0.8	87.46	97.60	77.48

### 4.3.2 主干特征提取网络的有效性

为验证 MobileNetv2 主干特征提取网络对提升模型分割效率的有效性,本节在相同的实验环境下设计了以下的对比实验,这些实验均采用改进的损失函数进行,在测试集上进行验证,实验结果如表 3 所列。

表 3 DeepGlobe 道路数据集上不同主干特征提取网络的性能结果

Table 3 Performance results of different backbone feature extraction networks on DeepGlobe road dataset

backbone	mPA/%	Accuracy/%	mIoU/%	推理时间/ (毫秒每幅)
Xception	<b>88.81</b>	<b>97.81</b>	<b>78.77</b>	104.4
Mobile netv2	87.54	97.68	77.60	<b>40.7</b>

从表 3 中可以看出,当把 DeepLabv3+算法的主干特征提取网络从原来的 Xception 替换为轻量级的 MobileNetv2 时,虽然分割精度略有下降,但是单幅图像推理时间显著缩短,分割速度达到了原始 DeepLabv3+算法的两倍多,说明使用轻量级的 MobileNetv2 作为 DeepLabv3+的主干特征提取网络来提取特征,可以有效减少参数量并提升分割速度,可见 MobileNetv2 更简单的模型结构相比 Xception 更适合于道路提取的任务,满足实时性的条件。

### 4.3.3 C-ASPP 和 ECANet 的有效性

本节验证所提 C-ASPP 模块与加入 ECANet 对模型分割精度的影响,使用 MobileNetv2 作为主干特征提取网络,并使用改进的损失函数进行以下实验。结果如表 4 所列,“√”表示所标注的模块用于实验,无“√”表示未添加该模块,其中 C-ASPP(add conv)表示仅在 ASPP 中增加卷积层,C-ASPP(cascade conv)表示仅在 ASPP 中串联相邻的卷积层。

表 4 DeepGlobe 道路数据集上不同改进模块的性能结果

Table 4 Performance results of different improved modules on DeepGlobe road dataset

C-ASPP (add conv)	C-ASPP (cascade conv)	ECANet	mPA	Accuracy	mIoU
			87.54	97.68	77.60
√			87.78	97.75	78.01
	√		87.97	97.82	78.43
√	√		88.20	97.90	79.09
		√	88.53	97.98	79.71
√	√	√	<b>89.54</b>	<b>98.12</b>	<b>80.91</b>

通过对比前 4 组的数据可以看出,在原模型的 ASPP 模块中分别增加卷积层和级联相邻的卷积层后,各项评价指标均有不同程度的提升,同时进行两项操作(即替换成 C-ASPP 模块)后,mPA, Accuracy 和 mIoU 分别提升了 0.66%, 0.22% 和 1.49%,C-ASPP 通过增加卷积层进一步扩大了感受野,提取到更多的不同尺度的道路特征,再通过不同特征层共享,使得模型能够学习到更加丰富的特征信息,从而提高对道路的识别精度。通过对比第 1 组和第 5 组可以看出,在加入 ECANet 注意力机制后,mPA, Accuracy 和 mIoU 分别提升了 0.99%, 0.3% 和 2.11%,说明 ECANet 可以使模型更加关注更重要的特征,从而更好地对道路信息进行提取,提高分割精度。从第 6 组的数据可以看出,在同时加入 C-ASPP 模块

和 ECANet 后,模型取得了最高的分割精度,相比未改进前, $mPA$ ,  $Accuracy$ ,  $mIoU$  分别提高了 2%, 0.44%, 3.31%, 有较大幅度的提升,证明了本文引入的不同改进模块的有效性。

为了进一步验证不同注意力机制模块和不同 ASPP 改进模块对算法分割精度的提升情况,进行以下实验,如表 5 所列。实验结果表明,在使用了 MobileNetv2 和改进的损失函数的 DeepLabv3+ 算法的基础上,引入 ECANet 注意力机制<sup>[18]</sup>相比引入 SENet 注意力机制<sup>[20]</sup>和 CBAM 注意力机制<sup>[21]</sup>的  $mPA$ ,  $Accuracy$  和  $mIoU$  更高,分别为 88.53%, 97.98% 和 79.71%;引入 C-ASPP 模块相比 DenseASPP<sup>[22]</sup>模块取得了更好的分割精度结果, $mPA$ ,  $Accuracy$  和  $mIoU$  分别为 88.2%, 97.9% 和 79.09%。总体而言,本文提出的改进模块在同类型改进方案中有更好的提升效果。

表 5 不同注意力机制模块和不同 ASPP 改进模块的性能结果

Table 5 Performance results of different attention mechanism modules and different ASPP improvement modules

Algorithm	$mPA$	$Accuracy$	$mIoU$
Deeplabv3+(MobileNetv2)	87.54	97.68	77.60
Deeplabv3+(MobileNetv2)+SENet <sup>[20]</sup>	88.05	97.79	78.82
Deeplabv3+(MobileNetv2)+CBAM <sup>[21]</sup>	87.99	97.76	78.66
Deeplabv3+(MobileNetv2)+ECANet <sup>[18]</sup>	<b>88.53</b>	<b>97.98</b>	<b>79.71</b>
Deeplabv3+(MobileNetv2)+DenseASPP <sup>[22]</sup>	87.85	97.71	78.74
Deeplabv3+(MobileNetv2)+C-ASPP	<b>88.20</b>	<b>97.90</b>	<b>79.09</b>

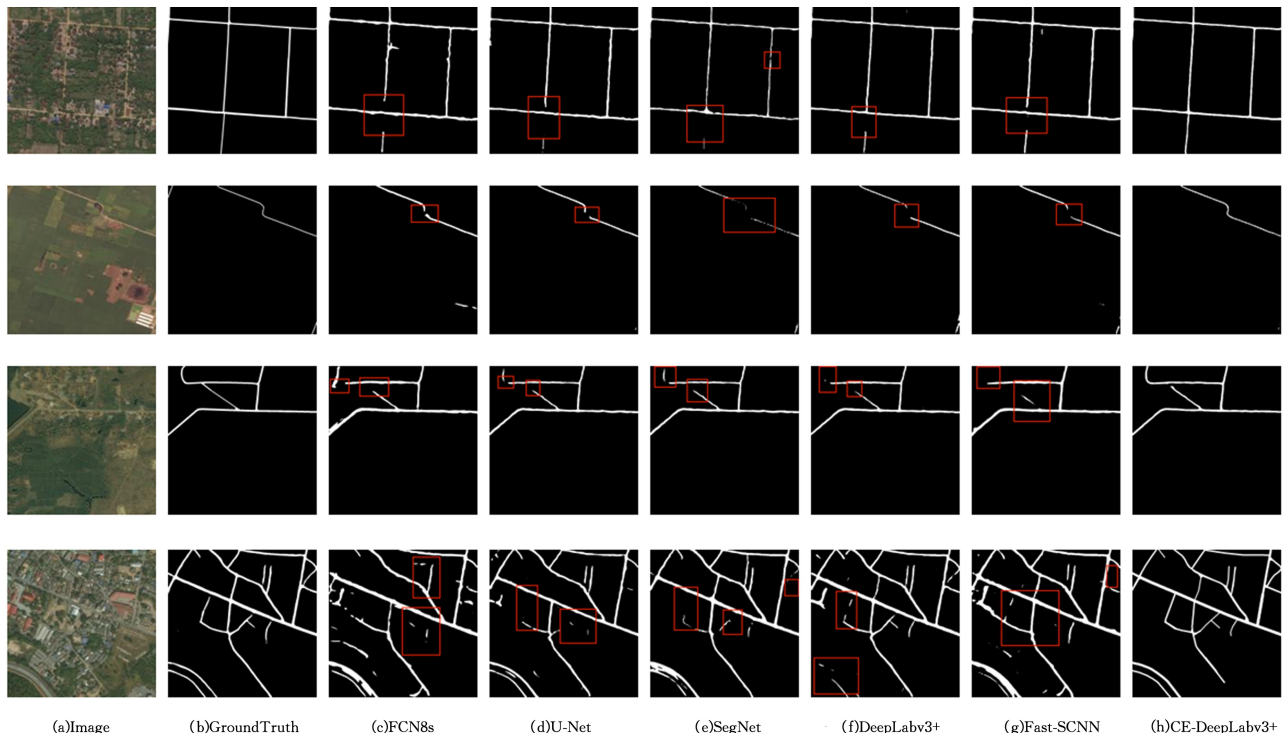


图 7 DeepGlobe 道路测试集上不同算法的分割效果图(电子版为彩图)

Fig. 7 Segmentation effect of different algorithms on DeepGlobe road test set

#### 4.4 与其他模型的分割效果对比

为进一步验证本文算法在遥感影像道路提取任务中的分割优势以及有效性,选取 FCN8s<sup>[11]</sup>, U-Net<sup>[12]</sup>, SegNet<sup>[13]</sup>, Fast-SCNN<sup>[15]</sup>等经典语义分割算法和本文算法进行对比测试分析,所有实验均在相同的实验环境和数据集上进行。不同算法的分割精度评价指标以及单幅图片的推理时间结果如表 6 所列。

表 6 DeepGlobe 道路数据集上不同算法的对比结果

Table 6 Comparison results of different algorithms on DeepGlobe road dataset

Algorithm	$mPA$ /%	$Accuracy$ /%	$mIoU$ /%	推理时间/ (毫秒每幅)
FCN8s <sup>[11]</sup>	84.56	96.34	72.68	253.4
U-Net <sup>[12]</sup>	85.23	96.43	76.81	63.3
SegNet <sup>[13]</sup>	83.76	95.98	74.89	156.5
DeepLabv3+ <sup>[16]</sup>	88.81	97.81	78.77	104.4
Fast-SCNN <sup>[15]</sup>	86.17	97.12	75.02	<b>15.8</b>
CE-DeepLabv3+	<b>89.54</b>	<b>98.12</b>	<b>80.91</b>	47.5

根据实验结果可知,本文算法相比其他算法获得了最高的分割精度指标结果, $mPA$ ,  $Accuracy$ ,  $mIoU$  分别为 89.54%, 98.12%, 80.91%。并且,虽然本文算法的推理时间比轻量级网络 Fast-SCNN 略长,但总体而言,其在分割精度和分割速度上取得了平衡,且都具有良好的表现。

#### 4.5 可视化结果分析

为了更好地描述 4.4 节的实验结果,利用表 6 所列的实验数据完成可视化结果分析,在 DeepGlobe 道路测试集上的部分提取结果如图 7 所示。

对于背景环境因素简单的道路影像图,不同语义分割

算法都有较好的表现,但在细节部分,如树木遮挡的部分

道路,本文算法有更好的提取效果,漏提现象较少,提取结果的连续性更好。由于 DeepGlobe 道路数据集中环境复杂,存在很多建筑、植被等遮挡物,同时还包含大量乡村、荒郊和裸地样本,因此道路和背景颜色、纹理相似,从而导致提取存在一定难度。但即使是在复杂的环境中,本文算法也能体现出它的优越性能表现。由图 7 中红色框所标记的区域可以看出,FCN8s、U-Net、SegNet、原始 DeepLabv3+ 和 Fast-SCNN 所分割的图像丢失了很多细节的部分,都产生了由部分道路漏分割导致的不连续的问题,而本文所提出的 CE-DeepLabv3+ 算法的分割结果最贴近标签图像。这归功于本文所提出的级联空洞空间金字塔池化模块,它可以捕获更多尺度的特征图信息进行融合,更好地提取了道路细节和轮廓特征,减少了不完整和不连续的现象。同时,由于引入了 ECANet 注意力机制和使用改进的损失函数进行训练,减少了背景复杂环境因素的干扰,并且降低了道路与背景类别不平衡造成的影响,使得对道路特征提取更加准确。综上所述,与经典语义分割算法相比,本文算法对道路信息提取结果更加完整,保证了道路的连通性,优于其他经典语义分割算法。

**结束语** 本文中,针对遥感影像道路提取任务中提取速度慢、漏分割和不连续的问题,提出了一种改进 DeepLabv3+ 算法。首先使用轻量级的主干特征提取网络 MobileNetv2 来减少参数量,加快图像分割速度;再利用级联的 ASPP 模块来获取更大的感受野并融合多级特征,学习更丰富的特征信息;再加入 ECANet 注意力机制,增加道路特征信息的权重,同时降低背景环境特征信息的权重,使模型更加关注像素占比较小的道路信息,从而有效降低背景复杂环境因素的干扰,提升道路提取的准确率;最后,采用改进的损失函数进行训练,以解决道路与背景样本不均衡的问题。消融实验研究表明,所提出的改进模块是有效的,与其他算法的对比实验表明,所提算法优于其他的经典语义分割算法。该算法基于深度学习方法提取遥感影像中的道路特征,能快速并准确地提取出不同环境下的道路信息,这项工作为在城市道路规划等应用中利用深度学习提供了真实案例研究。未来,还打算在更多的地理环境条件下验证该算法的泛化性。

## 参考文献

- [1] WANG W, YANG N, ZHANG Y, et al. A review of road extraction from remote sensing images[J]. Journal of Traffic and Transportation Engineering(English Edition), 2016, 3(3): 271-282.
- [2] LIAN R, WANG W, MUSTAFA N, et al. Road extraction methods in high-resolution remote sensing images: A comprehensive review[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 5489-5507.
- [3] WANG Z, JENSEN J R, IM J. An automatic region-based image segmentation algorithm for remote sensing applications[J]. Environmental Modelling & Software, 2010, 25(10): 1149-1165.
- [4] BLASCHKE T, BURNETT C, PEKKARINEN A. Image segmentation methods for object-based analysis and classification [M] // Remote Sensing Image Analysis: Including the Spatial Domain. Dordrecht: Springer Netherlands, 2004: 211-236.
- [5] YI F, MOON I. Image Segmentation: A Survey of Graph-Cut Methods[C] // 2012 International Conference on Systems And Informatics(ICSAI2012). IEEE, 2012: 1936-1941.
- [6] SHAN P. Image segmentation method based on K-mean algorithm[J]. EURASIP Journal on Image and Video Processing, 2018, 2018(1): 1-9.
- [7] LOPEZ J, BRANCH J W, CHEN G. Line-based image segmentation method: A new approach to segment VHSR remote sensing images automatically[J]. European Journal of Remote Sensing, 2019, 52(1): 613-631.
- [8] WAN Z J, LI L Y, YANG M, et al. Decision tree Model for Extracting Road Intersection Features from Vehicle Trajectory Data[J]. Acta Geodaetica et Cartographica Sinica, 2019, 48(11): 1391-1403.
- [9] CHENG G, WANG Y, GONG Y, et al. Urban road extraction via graph cuts based probability propagation[C] // 2014 IEEE International Conference on Image Processing (ICIP). IEEE, 2014: 5072-5076.
- [10] MAURYA R, GUPTA P R, SHUKLA A S. Road extraction using k-means clustering and morphological operations[C] // 2011 International Conference on Image Information Processing. New York: IEEE, 2011: 1-6.
- [11] ZHONG Z, LI J, CUI W, et al. Fully convolutional networks for building and road extraction: Preliminary results[C] // 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2016: 1591-1594.
- [12] CHEN Z, WANG C, LI J, et al. Reconstruction bias U-Net for road extraction from optical remote sensing images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 2284-2294.
- [13] 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.
- [14] LI D D, HE X H, LI P L, et al. Remote Sensing Image Road Extraction Network Based On SPUD-ResNet[J]. Computer Engineering and Applications, 2021, 57(23): 240-247.
- [15] POUDEL R P K, LIWICKI S, CIPOLLA R. Fast-scnn: Fast semantic segmentation network[J]. arXiv: 1902. 04502, 2019.
- [16] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C] // Proceedings of the European Conference on Computer Vision(ECCV). 2018: 801-818.
- [17] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 4510-4520.
- [18] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and

Pattern Recognition, 2020:11534-11542.

- [19] DEMIR I, KOPERSKI K, LINDENBAUM D, et al. A challenge to parse the earth through satellite images. 2018 IEEE[C]// CVF Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). 2018.
- [20] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:7132-7141.
- [21] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]// Proceedings of the European Conference on Computer Vision(ECCV). 2018:3-19.
- [22] YANG M, YU K, ZHANG C, et al. Densnaspp for semantic segmentation in street scenes[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:3684-3692.



**WANG Qian**, born in 1999, postgraduate. His main research interest is image processing.



**HE Lang**, born in 1974, professor, Ph.D. His main research interests include intelligent calculation and image processing.

(责任编辑:喻黎)

## 由 CCF 承办的科协年会分论坛“大数据关键技术与数字经济发展论坛”在京举办

7月12日由 CCF 承办的科协年会分论坛“大数据关键技术与数字经济发展论坛”于北京中国科技馆成功举办。该论坛由中国科协主办,CCF 承办,CCF 中国数字经济 50 人论坛(以下简称“50 人论坛”)、CCF 大数据专家委员会(以下简称“大专委”)、CCF 青年工作委员会(以下简称“青工委”)、CCF 数据治理发展委员会(以下简称“数发委”)联合协办,于 7 月 12 日成功举办。本场论坛设置了“数据要素乘数效应释放路径”、“数据资产的价值评估与入表”、“数据高质量供给与有效利用”、“数据治理的科学问题与关键技术”、“我国推进数据治理的可行路径”、“数据治理行业落地的”小切口等 6 项议题,并邀请大数据和数字经济相关领域的院士、专家,从技术创新、应用实践、政策导向、未来趋势等多个维度展开深入交流和研讨,解析在当前的数字经济大环境下,数据如何“供得出”、“流得动”、“用得好”。

论坛由 CCF 会士、中国人民大学教授、50 人论坛委员杜小勇,CCF 会士、华中科技大学教授、CCF 副理事长、50 人论坛委员金海,清华大学副教授、大专委副秘书长、青工委主任崔鹏,北京交通大学计算机学院院长、大专委副秘书长李滢东,北京邮电大学教授、大专委执行委员、数发委秘书长高志鹏依次主持。CCF 会士、前理事长、中国科学院院士、CCF 咨询委员会主任、50 人论坛主任梅宏作了开场讲话。梅宏院士强调了大数据关键技术和数字经济的重要性。首先,他指出如今数据的重要性毋庸置疑,只有数据才是信息时代最有价值、最值得留存的东西,其他都是工具。目前在 AI 的热潮中,数据起到了非常重要的作用,AI 越来越依赖于数据,导致现在的大模型最后变成了数据工程,这突显了数据的重要性。其次,他认为国家是第一个把数据明确作为生产要素来对待的国家,推动数字经济的发展,是我们未来弯道超车的重要机会,是开启数字文明新机遇。最后,梅院士呼吁,我们要抓住这个时代机遇,实现数字经济的发展,推动数字化转型,需要建立数据要素市场,构建数字的治理体系和数据技术体系。并鼓励与会专家畅所欲言,希望年轻学者共同探索大数据与数字经济的新路径,推动社会向数字文明迈进。

北京大学光华管理学院教授/原副校长、50 人论坛委员张国有,阿里研究院副院长、50 人论坛委员安筱鹏,中国社科院信息化研究中心主任、中国科学院《互联网周刊》主编、50 人论坛委员姜奇平,蚂蚁集团研究院院长、50 人论坛委员李振华,中国社会科学院经济研究所人工智能经济研究室主任、研究员张磊,中国人民大学教授、50 人论坛委员杜小勇,伏羲智库创始人、清华大学互联网治理研究中心主任、50 人论坛委员李晓东,CCF 数字治理发展委员会主任、中国电子党组成员/副总经理、50 人论坛委员陆志鹏,北京大学信息科学技术学院教授黄罡,北京市大数据中心数据部部长、数发委常委贾晓丰,安徽飞数信息常务副总、大专委常委谭昶在论坛上作专题发言,20 余名专家学者参与论坛讨论。

随后,梅宏院士对讨论的主题进行了补充与深化,以其独到的视角为讨论内容增添了新的维度与深度。他认为,经济专家和技术专家关注的角度可能不太一样,的确需要不同视角的碰撞以开阔视野。就数据资产入表而言,目前尚有难度,存在很多问题有待解决,如果各方面的利益没有照顾到,有的部门一定会成为数据共享的障碍。最后,他着重强调技术发展应务实,避免过度超前,强调国家层面应慎重规划研究路径。

针对此次论坛的议题,各单位已于 6 月 20 日及 7 月 4 日分别召开了两场相关调研及闭门研讨,就数据流通与交易及管理、数据治理的科学问题与研究路径等议题形成了成果输出。未来,还将不定期举办多场系列专题研讨会,针对数字经济领域的热点、重点和难点问题进行深入研讨,鼓励跨学科群议群策,形成并发布独立、客观且交融的智库思想建议,以实际行动推动数字经济的健康快速发展。