



一种用于地震断层图像识别的SPD-UNet模型

席英杰¹, 李克文¹, 徐延辉¹, 朱剑兵²

(1. 中国石油大学(华东) 计算机科学与技术学院, 山东 青岛 266580;

2. 中国石化胜利油田分公司物探研究院, 山东 东营 257022)

摘要: 断层是控制油气田形成和分布的主要因素, 断层检测和识别对于油气勘探具有重要作用。基于 Attention-UNet 神经网络模型, 构建一种面向地震断层图像识别的 SPD-UNet 模型。引入空洞卷积, 在保证卷积核感受野大小且不损失原始图像分辨率的情况下, 增强 SPD-UNet 模型的断层图像特征提取能力。将金字塔结构的空洞卷积组合成 SPD 模块, 解决空洞卷积的局部信息丢失问题, 提高断层信息关联性及图像识别精度。实验结果表明, SPD-UNet 模型对于地震断层图像的识别精度优于 SegNet 与 ResUNet 模型, 并且识别结果与实际标注的地震断层形状及位置更接近。

关键词: 地震断层识别; 图像分割; 神经网络; UNet 模型; 空洞卷积; 金字塔结构

开放科学(资源服务)标志码(OSID):



中文引用格式: 席英杰, 李克文, 徐延辉, 等. 一种用于地震断层图像识别的 SPD-UNet 模型[J]. 计算机工程, 2021, 47(12): 249-255.

英文引用格式: XI Y J, LI K W, XU Y H, et al. A SPD-UNet model for seismic fault image identification[J]. Computer Engineering, 2021, 47(12): 249-255.

A SPD-UNet Model for Seismic Fault Image Identification

XI Yingjie¹, LI Kewen¹, XU Yanhui¹, ZHU Jianbing²

(1. College of Computer Science and Technology, China University of Petroleum (East China), Qingdao, Shandong 266580, China;

2. Geophysical Research Institute of Sinopec Shengli Oilfield Branch, Dongying, Shandong 257022, China)

[Abstract] Fault is the main factor that controls the formation and distribution of oil and gas fields, so the detection and identification of fault plays an important role in the exploration oil and gas fields. Based on the Attention-UNet model, this paper proposes an improved SPD-UNet model for fault identification in earthquake images. SPD-UNet introduces dilated convolution, which can effectively enhance image feature extraction while expanding the receptive field and preventing resolution loss. At the same time, the dilated convolutions in the pyramid structure are stacked to form the SPD module, which avoids the local information loss of dialted convolutions, and improves the correlation between fault information and image identification accuracy. Experimental results show that SPD-UNet exhibits a higher identification accuracy than SegNet and ResUNet. The fault position and shape identified by SPD-UNet are closer to actual information.

[Key words] seismic fault identification; image segmentation; neural network; UNet model; dilated convolution; pyramid structure

DOI: 10.19678/j.issn.1000-3428.0059327

0 概述

断层是地下岩层沿一个破裂面或破裂带两侧发生相对位错的现象。地震往往是由断层活动引起的, 是断层活动的一种表现, 因此地震与断层的关系

十分密切。在常规地震剖面上, 断层可以通过反射波同相轴错动、分叉、合并、扭曲、形状和数目突变、相邻层位的错动等表象特征进行直接识别。在油气勘探领域, 勘探阶段的区域断裂研究对于地质构造、沉积环境解释以及隐蔽性油气田开发具有重要作

基金项目: 国家自然科学基金重大项目(51991361); 国家科技重大专项(2016ZX05021-002)。

作者简介: 席英杰(1995—), 男, 硕士研究生, 主研方向为计算机视觉、深度学习; 李克文, 教授、博士、博士生导师; 徐延辉, 硕士研究生; 朱剑兵, 博士。

收稿日期: 2020-08-24 修回日期: 2020-12-01 E-mail: xitutu2019@163.com

用。常规的断层解释方法是工作人员在三维地震数据的垂直剖面 and 水平切片上手动解释断层,通过视觉识别反射层的不连续性来实现断层解释,难度大、周期长、主观性强,在很大程度上依赖于解释人员的经验和相关区域的前期调研。长期以来,学者们围绕提高断层解释的精度和速度进行了大量研究,通过相干体属性^[1]、方差体属性、断层切片、边缘增强属性等技术来提取三维地震数据精确描述断层,并为计算机自动识别断层打下了良好的基础。

近年来,随着计算机运算速度的不断提高,深度学习技术在多个领域均取得重大突破。深度学习的概念来源于人工神经网络,神经网络结构可以很好地抓取对象的特征属性,发现数据的隐藏特征,并具有良好的自我学习能力。因此,越来越多的学者开始尝试将深度学习技术应用到地震领域,与地震数据相结合解释地震断层。TINGDAHL等^[2]将深度学习与断层识别相结合用于地震目标检测。HUANG等^[3-4]通过引入卷积神经网络(Convolutional Neural Networks, CNN)来识别地震属性检测断层。GUITTON等^[5-7]将CNN与地震断层属性相结合来检测断层。WU等^[8]利用基于CNN的像素级图像分割方法检测断层,并获得了显著效果。图像分割技术是计算机视觉领域的重要研究方向且已有许多CNN模型被用于图像分割任务,例如LONG等^[9]提出的全卷积神经网络(Fully Convolutional Networks, FCN)模型、RONNEBERGER等^[10-12]提出的UNet、SegNet网络以及残差网络模型,可利用图像分割技

术实现地震断层的自动化识别。

本文基于 Attention-UNet 神经网络模型和空洞卷积,提出一种用于地震断层图像识别的 SPD-UNet 模型。基于神经网络的 Encoder-Decoder 结构,在 Encoder 阶段输入地震断层图像,采用 CNN 与 SPD 模块进行特征提取,在 Decoder 阶段进行图像上采样恢复与 Skip Connection 操作,完成对地震断层图像的训练与识别,同时引入 Focal Loss 函数,解决地震图像的正负样本极度不平衡问题。

1 卷积神经网络模型

1.1 UNet神经网络模型

UNet网络是U对称结构,由一个收缩路径(左边,也称特征提取)和一个扩张路径(右边,也称上采样)组成。收缩路径遵循典型的卷积网络结构,包含4个Convolutional layer,每层layer包含2个3×3卷积层以及1个最大池化层。扩张路径包括4个Upsampling layer,每个layer与对应的feature map通过Skip Connection在通道维度进行拼接,形成更全面的特征信息,再进行下一步处理。

利用收缩路径部分对原始图像进行特征提取,之后采用扩张路径将feature map恢复到输入图片大小,并恢复每一个像素对应的空间位置。由于经过上采样存在部分信息丢失的问题,通过Skip Connection将对应的输出层(具有更好的全局信息)与浅层(具有更丰富的局部细节)相结合,从而进行更有效的预测。UNet模型结构如图1所示。

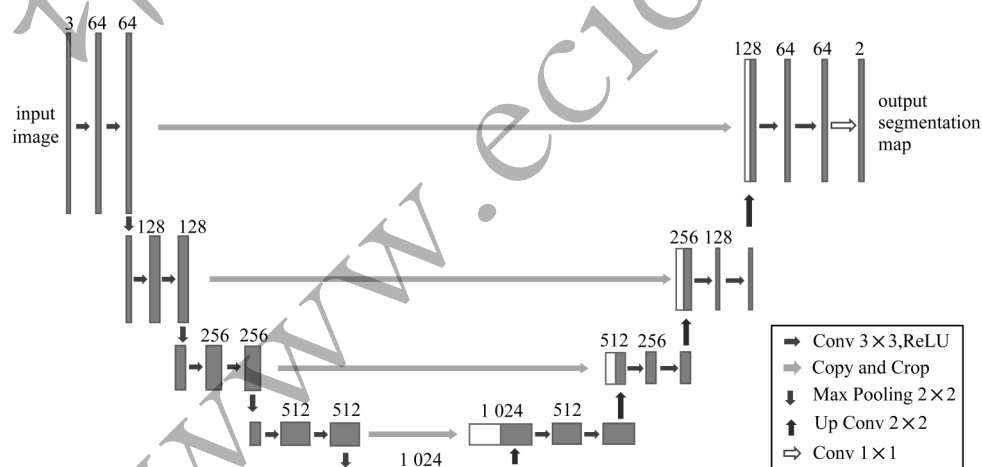


图1 UNet模型结构

Fig.1 Structure of UNet model

1.2 Attention-UNet神经网络模型

在UNet的基础上,在每个跳跃链接的末端使用Attention Gate^[13]结构(以下简称AG),对需要提取的feature map实现Attention机制。Attention机制是模仿人类注意力提出的一种解决方案,可以从大量信息中筛选出高价值信息,主流Attention机制有通道

注意力^[14]、空间注意力^[15]以及上下文注意力^[16],应用于图像领域时可以抑制不相关背景区域的特征响应,使模型关注更有价值的特征。Attention-UNet模型结构与UNet一致,如图2所示,网络接收输入图像,通过与UNet类似的卷积和下采样操作进行特征提取,不同之处在于,在Skip Connection阶段添加了

AG注意力模块,用于提升对断层区域的关注度。在图2中, H_i 、 W_i 、 D_i 分别代表不同阶段输入输出的高、宽以及通道数, F 表示每批次输入的数量。

增加的AG模块结构如图3所示。AG模块接收左侧上一层 feature map 和右侧下一层特征,然后拼接右侧上一层得到最后输出。其中, g 信号来自

右侧下一层的输入, x_i 信号来自左侧上一层的输入。左侧输入分为两部分:一部分与 g 信号相加,然后共同进行 ReLU 和 Sigmoid 激活,得到的结果进行重采样,获得最终输出为 x_i 的注意力系数;另一部分与 x_i 进行相乘,以突出通过 Skip Connection 传递的显著特征。

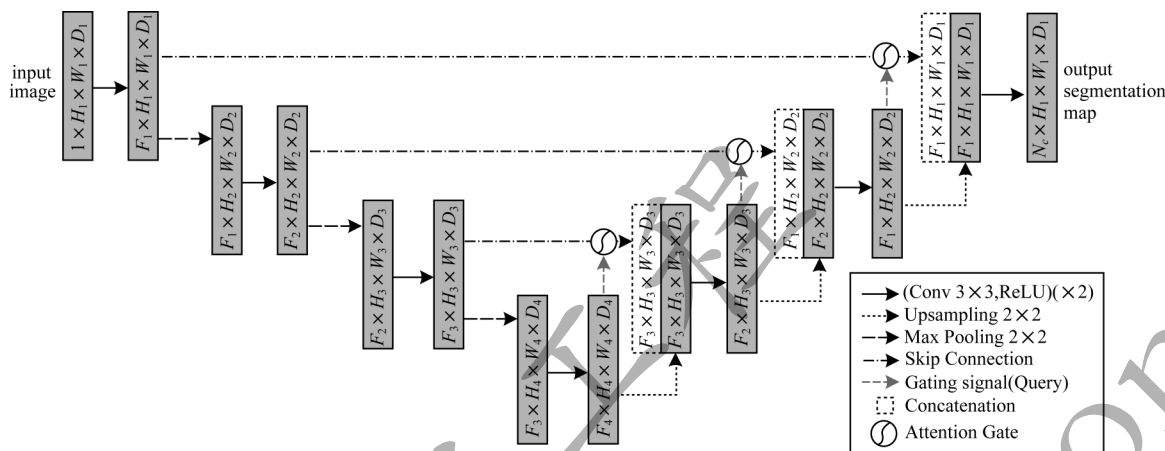


图2 Attention-UNet模型结构

Fig.2 Structure of Attention-UNet model

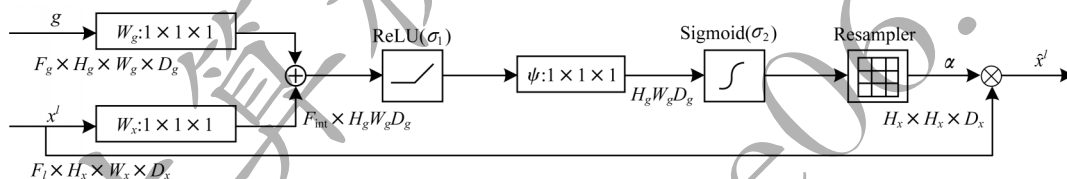


图3 AG模块结构

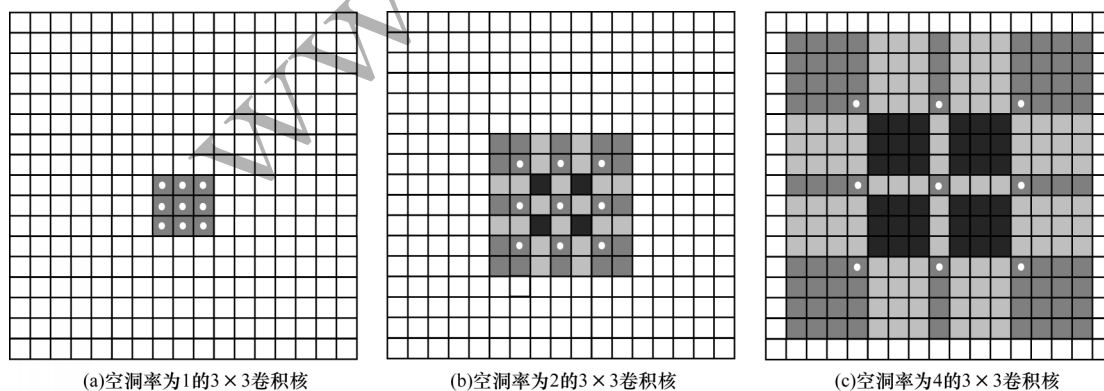
Fig.3 Structure of AG module

1.3 空洞卷积

在图像分割领域,增大卷积核的感受野十分重要,有利于捕获更大区域的特征信息,在传统神经网络中,一般采用池化操作降低图像尺寸以增大感受野,同时减少计算参数量,但池化存在一定的缺陷,由于降低了输入图像的尺寸,会造成信息的丢失,尤其在地震断层识别中,地震断层属于小目标,在池化操作下极易造成

识别不连续,甚至无法重建目标断层。

空洞卷积^[17]最初来源于DeepLab系列v1^[18]和v2^[19]网络,经过PSPNet^[20]、DeepLab v3+^[21]等网络的不断完善,具有更大感受野的同时保持原有的权重个数不变,并且可通过空洞卷积的不同空洞率更好地捕获多尺度^[22]的上下文信息。空洞卷积示意图如图4所示。



(a)空洞率为1的 3×3 卷积核

(b)空洞率为2的 3×3 卷积核

(c)空洞率为4的 3×3 卷积核

图4 空洞卷积示意图

Fig.4 Schematic diagram of dilated convolution

空洞卷积的实际卷积核计算公式如下:

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

其中: k 为原始卷积核大小; r 为空洞卷积的空洞率; K 为空洞卷积大小。

2 SPD-UNet模型

对于地震断层识别,可以将地震剖面切片视为原始输入图片,将断层识别问题视为在断层图片上通过标记断层位置进而识别图像的二分割问题。利用UNet网络并引入Attention机制和空洞卷积来实现图像分割,使得神经网络能进行相应的学习与识别。

SPD-UNet模型相比UNet模型在特征提取部分的前3层结构保持不变,输入图像大小为256像素×256像素,每层包含2个3×3的卷积层,然后加入Batch Normalization对数据进行归一化处理,使得数据在进行ReLU之前不会因为过大而导致网络性能不稳定,ReLU激活之后接一个2×2的Max Pooling池化操作,用来增大感受野并且降低参数量。由于断层识别属于小目标识别,为避免图像信息损失过多,在特征提取的后两层,取消池化操作,采用空洞卷积来增大感受野,提高特征提取精度。此外,由于空洞卷积是稀疏的,存在空洞率越大获得的长距离信息关联性越低的问题,会导致输出的结果连续性较差,空间关联性较低,并影响识别精度。为此,在本文模型中将金字塔结构^[23]的空洞卷积进行组合,整合成SPD模块,如图5所示。SPD模块共包括4个卷积,1个常规3×3卷积核与3个空洞卷积,空洞率 r 分别为2、3和4。4个卷积核接收来自上一层的同一个输出,分别与其进行卷积操作后输出同样尺寸的结果。将4个输出进行sum后的值作为该layer最后

的输出,SPD模块使得模型可以同时兼顾小范围和大范围内的特征信息,在具有更大感受野的同时,利用不同空洞率的金字塔卷积组合,避免空洞卷积产生的网格效应。

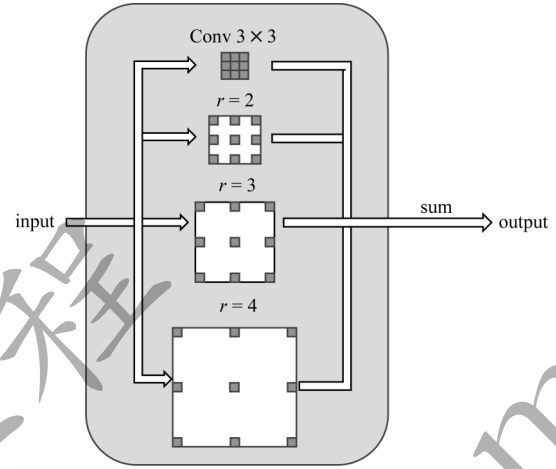


图5 SPD模块结构

Fig.5 Structure of SPD module

在上采样部分,输入部分改进了传统UNet网络的Skip Connection部分,加入一个Attention Gate来添加注意力机制,AG接收左侧上一层feature map和右侧下一层特征,然后拼接右侧上一层得到最后输出。右侧包括3层,每层采用Upsample进行扩展,然后采用2×2的卷积及Batch Normalization层进行归一化,之后进行ReLU激活,传递到右边上一层。模型采用Sigmoid函数对最终输出进行处理,使得最终输出图像中每个像素点的值对应于0~1,保证网络模型的输出大小与输入大小一致。SPD-UNet模型结构如图6所示。

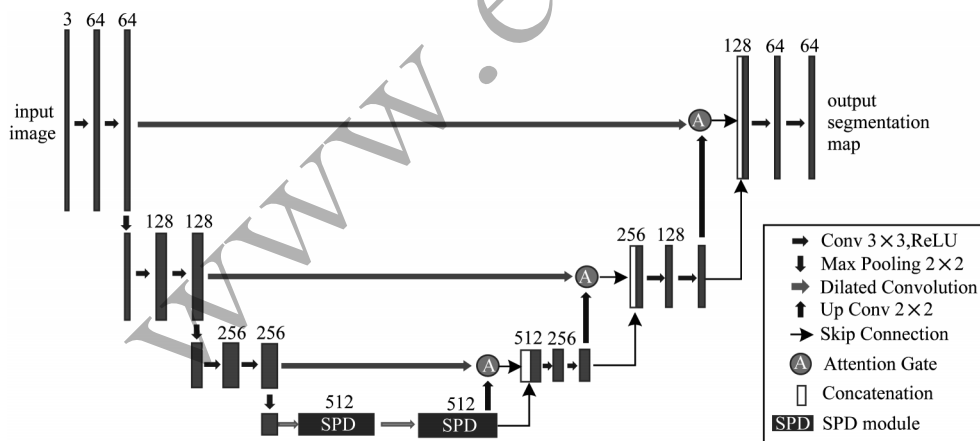


图6 SPD-UNet模型结构

Fig.6 Structure of SPD-UNet model

3 模型训练

模型训练实验选取胜利油田某区块地震断层数据,地震数据为高精度地震采集数据。提取SEG-Y

地震数据体,将SEG-Y数据体通过提取识别获得断层区块剖面图集,选取图片进行断层标注组建样本集、训练集与测试集。

在传统的图像分割领域,最常用的二分类损失函数为BCE Loss,但因为地震图像正负样本比例极度不平衡,负样本占比达到90%以上,在神经网络学习的过程中,过多的无用负样本会使得模型整体学习方向发生偏差,产生无效学习,导致预测效果差,识别效果不明显。因此,本文中所有网络模型均采用Focal Loss作为损失函数。

Focal Loss^[24]是在标准交叉熵基础上修改得到的。标准交叉熵是将各个训练样本交叉熵直接求和,即各个样本权重一致,计算公式如下:

$$CE(p, y) = \begin{cases} \ln p, y = 1 \\ -\ln(1-p), \text{其他} \end{cases} \quad (2)$$

为表示简便,用 p_i 表示 true class 概率,将式(2)改写为:

$$CE(p, y) = CE(p_i) = -\ln p_i \quad (3)$$

通过添加系数 α 来控制正负样本对总体 Loss 的共享权重, α 取值较小时可以降低负样本的权重,如式(4)所示:

$$CE(p_i) = -\alpha_i \ln p_i \quad (4)$$

此外,需要解决控制易分类和难分类样本的权重问题,如式(5)所示:

$$FL(p_i) = -(1-p_i)^\gamma \ln p_i \quad (5)$$

通过调制系数 γ 减少易分类样本的权重,使得模型更关注于难分类样本,因此Focal Loss最终表达式如式(6)所示,既能调整正负样本的权重,又能控制难易分类样本的权重。

$$FL(p_i) = -\alpha_i (1-p_i)^\gamma \ln p_i \quad (6)$$

对于Focal Loss函数,将系数 α 设置为0.25,参数对比结果如图7所示,可以看出,当 γ 取2时,模型损失函数最小,拟合精度最好,因此选取 $\alpha=0.25$ 、 $\gamma=2$ 来共同调节损失函数。

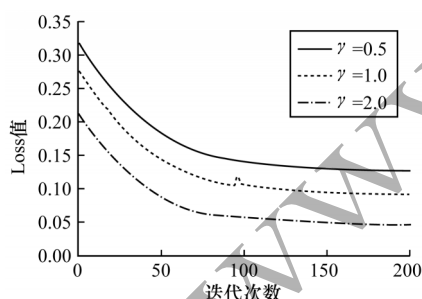


图7 Focal Loss函数参数对比结果

Fig.7 Comparison results of Focal Loss function parameters

深度学习网络有多种训练算法,常用的有随机梯度下降法(SGD)、自适应梯度下降法(Adgrad)、自适应学习率调整法(Adadelta)、自适应动量估计法(Adam)等。本文实验采用收敛速度较快的Adam算法。

4 实验验证

在图像分割领域,最常用的评价指标为交并比(Intersection over Union, IOU)和DICE系数,本文实验选取这两种指标进行性能评测。

4.1 评价指标

DICE系数和IOU是衡量两个集合之间相似性的度量指标,在图像分割领域用来衡量网络分割结果与实际结果间的相似性,数值越大,图像的相似性越高。DICE系数和IOU虽然在表达方式上有所差别,但均是关于图像分割精度的指标,计算公式分别如式(7)和式(8)所示。为更好地应用于断层识别任务,将其转换为如式(9)和式(10)所示。

$$D_{DICE}(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (7)$$

$$I_{IOU}(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (8)$$

其中: A 、 B 分别表示两个集合,即两张图片的像素集合。

$$D_{DICE} = \frac{2 \times T_{TP}}{(T_{TP} + F_{FN}) + (T_{TP} + F_{FP})} \quad (9)$$

$$I_{IOU} = \frac{T_{TP}}{T_{TP} + F_{FN} + F_{FP}} \quad (10)$$

其中: T_{TP} 是样本目标和预测目标的交集部分; F_{FP} 是将背景误认为断层的部分; F_{FN} 是将断层误认为背景的部分。鉴于地震断层图像正负样本极不平衡,实验在进行IOU及DICE系数计算时选择去除背景元素的图像。

4.2 实验结果与分析

不同UNet模型的识别结果如图8所示,其中,图8(a)表示原始图像,图8(b)表示实际标注,图8(c)表示用UNet模型进行断层识别的结果,图8(d)表示用Attention-UNet模型进行断层识别的结果,图8(e)表示加入单层空洞卷积的Attention-UNet模型进行断层识别的结果,图8(f)表示用本文SPD-UNet模型进行断层识别的结果。不同UNet模型在断层图像识别中识别性能比较如表1所示。从实验结果可以看出:UNet模型可以识别部分断层,但边缘较为模糊,且识别精度较低,缺失较严重,断层识别连续性较差,在测试集上的DICE值为0.703 6, IOU值为0.542 8;加入Attention机制后,断层识别的连续性和精度均有了明显提升,但整体识别效果仍然不够好,在测试集上的DICE值为0.769 2, IOU值为0.625 0;加入单层空洞卷积后结果差别不大,并且单层空洞卷积存在长距离信息残缺问题,导致相比Attention-UNet模型预测断层的连续性更差一些,在测试集上的DICE值为0.670 3, IOU值为0.504 1;本文SPD-UNet模型较其他模型有较大改善,识别精度明显提升。

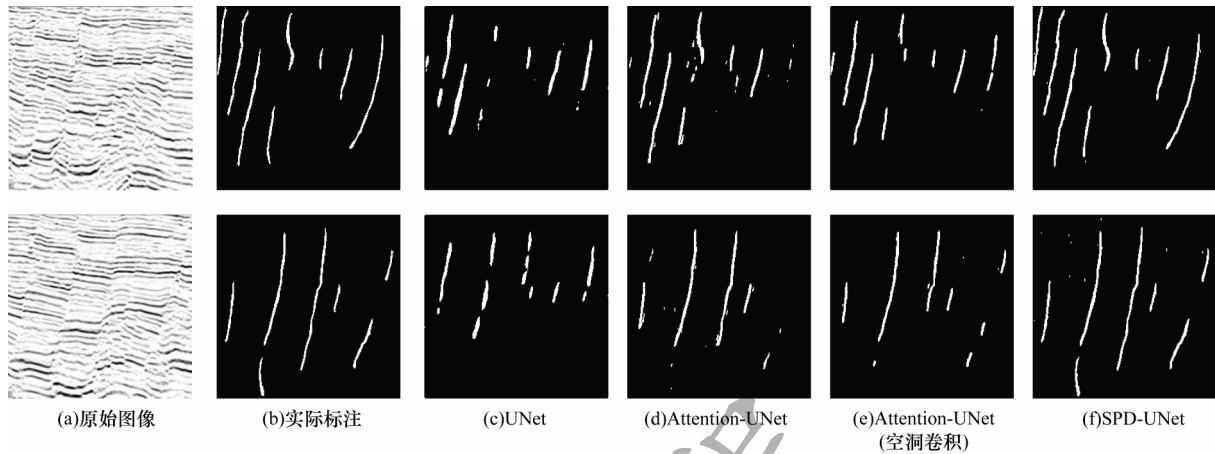


图8 4种UNet模型的识别结果比较

Fig.8 Comparison of identification results of four UNet models

表1 4种UNet模型在断层图像识别中的识别性能比较

Table 1 Comparison of identification performance of four UNet models in fault image identification

模型	IOU		DICE 系数		参数量/ 10^6
	训练集	测试集	训练集	测试集	
UNet	0.611 0	0.542 8	0.758 5	0.703 6	13.4
Attention-UNet	0.755 1	0.625 0	0.860 4	0.769 2	34.9
Attention-UNet(空洞卷积)	0.781 1	0.504 1	0.877 1	0.670 3	20.3
SPD-UNet	0.852 9	0.720 9	0.920 6	0.837 8	39.5

图9给出了用常见的语义分割领域的神经网络模型进行断层识别,其中,图9(a)表示FCN模型的识别结果,图9(b)表示基于34层网络的ResUNet模型识别结果,图9(c)表示SegNet模型的识别结果,图9(d)表示本文SPD-UNet模型的识别结果。不同

语义分割模型在断层图像识别中的识别性能比较如表2所示。从实验结果可以看出,本文SPD-UNet模型识别效果优于SegNet与ResUNet模型,与FCN模型接近,识别连续性更好,与实际标注的地震断层形状及位置更接近。

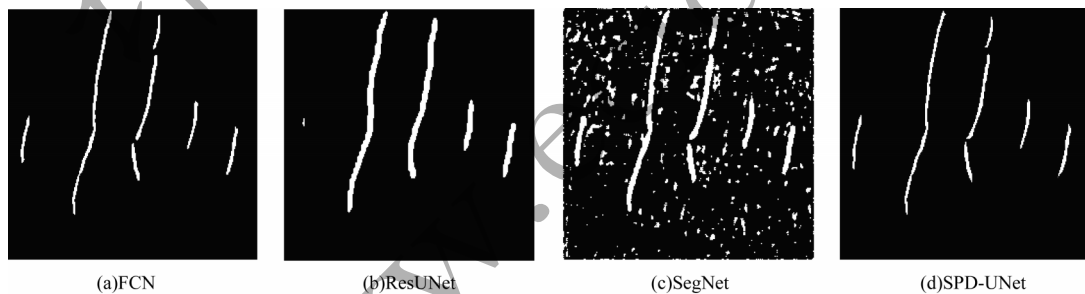


图9 4种语义分割模型的识别结果比较

Fig.9 Comparison of identification results of four semantic segmentation models

表2 4种语义分割模型在断层图像识别中的识别性能比较

Table 2 Comparison of identification performance of four semantic segmentation models in fault image identification

模型	IOU		DICE 系数		参数量/ 10^6
	训练集	测试集	训练集	测试集	
FCN	0.881 5	0.773 0	0.937 0	0.914 2	35.9
ResUNet	0.817 8	0.690 2	0.899 7	0.816 7	21.8
SegNet	0.799 0	0.671 1	0.888 3	0.803 1	29.4
SPD-UNet	0.852 9	0.720 9	0.920 6	0.837 8	39.5

综上所述,SPD-UNet模型在断层识别方面,相比UNet、Attention-UNet以及单层空洞卷积模型,IOU值、DICE系数和参数量指标均获得较大性能提

升,得到了更好的识别结果。相比其他语义分割模型,SPD-UNet模型在测试集上,IOU值相比SegNet与ResUNet模型约有0.05和0.02的提升,与FCN模

型的IOU值也较接近,DICE系数达到0.837 8,IOU值达到0.720 9,说明了SPD-UNet模型的有效性。SPD-UNet模型的参数量为 39.5×10^6 ,大于SegNet、ResUNet以及其他UNet系列模型的参数量,但IOU也高于上述模型。与此同时,在识别精度略低于FCN的情况下,参数量略大于FCN模型的 35.9×10^6 ,说明SPD-UNet模型存在优化空间。

5 结束语

本文构建面向断层图像识别的改进SPD-UNet模型,利用神经网络强大的编码和解码能力,通过特征提取与上采样,并引入注意力机制、空洞卷积和Focal Loss损失函数,增强SPD-UNet模型的断层图像特征提取能力,在保证感受野大小且不损失分辨率的情况下,更全面地捕捉断层的整体分布信息。同时,利用SPD模块解决空洞卷积的局部信息丢失问题,提高断层信息关联性 & 图像识别精度。实验结果表明,应用SPD-UNet模型进行地震断层识别效果较好。后续将通过优化Encoder-Decoder结构、Attention机制以及门控模块进一步提升SPD-UNet模型的地震断层图像识别精度,并利用GAN网络降低地震断层图像噪声对识别结果的影响。

参考文献

- [1] FINN C J, BACKUS M M. Estimation of three-dimensional dip and curvature from reflection seismic data[EB/OL]. [2020-07-11]. <https://library.seg.org/doi/abs/10.1190/1.1893089>.
- [2] TINGDAHL K M, DE ROOIJ M. Semi-automatic detection of faults in 3D seismic data[J]. *Geophysical Prospecting*, 2005, 53(4): 533-542.
- [3] HUANG L, DONG X S, CLEE T E. A scalable deep learning platform for identifying geologic features from seismic attributes[J]. *The Leading Edge*, 2017, 36(3): 249-256.
- [4] 李鹏松,李俊达,吴良武,等. 基于阈值分割法和卷积神经网络的图像识别算法[J]. *吉林大学学报(理学版)*, 2020, 58(6): 1436-1442.
LI P S, LI J D, WU L W, et al. Image recognition algorithm based on threshold segmentation method and convolutional neural network[J]. *Journal of Jilin University (Science Edition)*, 2020, 58(6): 1436-1442. (in Chinese)
- [5] GUITTON A. 3D convolutional neural networks for fault interpretation[C]//*Proceedings of the 80th EAGE Conference and Exhibition*. Copenhagen, Denmark: EAGE Publications Press, 2018: 1-17.
- [6] GUO B W, LIU L, LUO Y. Automatic seismic fault detection with convolutional neural network[C]//*Proceedings of International Geophysical Conference*. Beijing, China: Chinese Petroleum Society, 2018: 24-36.
- [7] ZHAO T, MUKHOPADHYAY P. A fault-detection workflow using deep learning and image processing[EB/OL]. [2020-07-11]. <https://library.seg.org/doi/10.1190/segam2018-2997005.1>.
- [8] WU X M, LIANG L M, SHI Y Z, et al. FaultSeg3D: using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation[J]. *Geophysics*, 2019, 84(3): 35-45.
- [9] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE Press, 2015: 3431-3440.
- [10] RONNEBERGER O, FISCHER P, BROX T. UNet: convolutional networks for biomedical image segmentation[C]//*Proceedings of MICCAI'15*. Washington D. C., USA: IEEE Press, 2015: 4-28.
- [11] 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.
- [12] HE K M, ZHANG X Y, REN S Q, et al. Identity mappings in deep residual networks[C]//*Proceedings of ECCV'16*. Amsterdam, the Netherlands: [s. n.], 2016: 630-645.
- [13] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-Net: learning where to look for the pancreas[EB/OL]. [2020-07-11]. <https://arxiv.org/abs/1804.03999v2>.
- [14] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011-2023.
- [15] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[EB/OL]. [2020-07-11]. <https://arxiv.org/abs/1807.06521>.
- [16] FU J, LIU J, TIAN H J, et al. Dual attention network for scene segmentation[C]//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE Press, 2019: 3141-3149.
- [17] ANTHIMOPOULOS M, CHRISTODOULIDIS S, EBNER L, et al. Semantic segmentation of pathological lung tissue with dilated fully convolutional networks[J]. *IEEE Journal of Biomedical and Health Informatics*, 2019, 23(2): 714-722.
- [18] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[EB/OL]. [2020-07-11]. <https://arxiv.org/abs/1412.7062>.
- [19] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.
- [20] ZHAO H S, SHI J P, QI X J, et al. Pyramid scene parsing network[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Washington D. C., USA: IEEE Press, 2017: 6230-6239.
- [21] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[M]//*Proceedings of ECCV'18*. Berlin, Germany: Springer, 2018: 833-851.
- [22] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[EB/OL]. [2020-07-11]. <https://arxiv.org/abs/1511.07122>.
- [23] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916.
- [24] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal Loss for dense object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 318-327.