

基于多尺度网络的苗绣绣片纹样分割算法研究

陈世婕,王卫星*,彭莉
(贵州大学 机械工程学院,贵州 贵阳 550025)

摘要:苗绣作为苗族人民传承历史文化的主要形式,图案内容丰富。但纹样设计周期相对较长且大多为手工劳动,因此生产效率较低。与此同时,在互联网平台上的相关资源质量不佳且数量较少,大多纹样及实体获取往往需要到实地考察。为解决传统手工苗绣绣片难以保存的问题,采用深度学习的图像语义分割方法进行相关苗绣绣片纹样提取用以保存和分类。该文构建了一组小型苗绣绣片纹样数据库并进行相关类别标注,同时提出一种多尺度图像语义分割网络架构(Multi-Scale Network)。该方法包含四个模块,分别为多尺度模块(Multi-Scale Block, MSB)、多尺度编码器(Multi-Scale Encoder, MSE)、多尺度解码器(Multi-Scale Decoder, MSD)以及语义分割头(Semantic-Head, SH)。每个模块均采用高效设计的多尺度提取机制和残差结构,可以提取更多全局纹样信息,并显著提升模型在不同任务下的泛化能力。所提方法在PASCAL VOC 2012数据集以及自建苗绣绣片纹样数据库中进行测试,实验结果相比全卷积神经网络的交并比(MIoU)提高了14.84%和6.7%。在相同实验条件下也验证了算法在应对错误分割、遗漏分割的有效性。

关键词:苗绣;数据库;深度学习;语义分割;数字化保护

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2023)11-0149-07

doi:10.3969/j.issn.1673-629X.2023.11.022

Research on Miao Embroidery Pattern Segmentation Algorithm Based on Multi-scale Network

CHEN Shi-jie, WANG Wei-xing*, PENG Li

(School of Mechanical Engineering, Guizhou University, Guiyang 550025, China)

Abstract: The Miao Embroidery in Qiongzhusi, Guizhou Province, China is a precious intangible cultural heritage, as well as a national costume handicraft and textile. Its exquisite pattern design requires exquisite workmanship. However, the design cycle is relatively long and most of them are manual labor, so the production efficiency is low. At the same time, the quality and quantity of relevant resources on the Internet platform are poor, and most of the patterns and entities need to be obtained on the spot. In order to solve the problem that traditional manual Miao Embroidery pieces are difficult to save, the image semantic segmentation method based on deep learning is used for saving and classification. We have constructed a group of small Miao Embroidery pattern databases and proposed a multi-scale image semantic segmentation network architecture (Multi Scale Network). The method includes four modules, namely, Multi Scale Block (MSB), Multi Scale Encoder (MSE), Multi Scale Decoder (MSD) and Semantic Head (SH). Each module adopts an efficient multi-scale extraction mechanism and residual structure, which can extract more global information and significantly improve the generalization of the model under different tasks. The proposed method was tested in PASCAL VOC 2012 dataset and self built Miao Embroidery Piece Pattern Database. The experimental results showed that the mean intersection over union (MIoU) of the full convolution neural network was improved by 14.84% and 6.7% compared with the full convolution neural network. Under the same experimental conditions, the proposed method in dealing with error segmentation and omission segmentation is also verified.

Key words: Miao Embroidery; database; deep learning; semantic segmentation; digital protection

0 引言

苗族是中国人口最多的少数民族之一,分布在贵

州省不同地区。苗族的民族历史大多记录在刺绣图案中,不仅可以作为装饰,同时还具备记录历史的功能。

收稿日期:2022-12-15

修回日期:2023-04-19

基金项目:贵州省科学技术基金(黔科合基础[2020]1Y262);贵州省教育厅青年科技人才成长项目(黔教合KY字[2018]112);贵州省科技计划项目(黔科合支撑[2021]一般396);贵州省科技计划重点项目(黔科合基础-ZK[2023]重点015)

作者简介:陈世婕(1998-),女,硕士研究生,研究方向为工业设计、产品创新设计;通讯作者:王卫星(1982-),男,博士,副教授,研究方向为数字化设计及理论、人机交互、智能穿戴设计。

苗族人民现在主要聚居在贵州省东南部,在这里可以找到丰富且具有特色的苗绣,因此对黔东南苗绣的研究具有代表性。同时,苗绣作为非物质文化遗产,是中

国少数民族服饰文化的重要组成部分,也是一种民族工艺品和纺织品^[1],如图 1 所示。

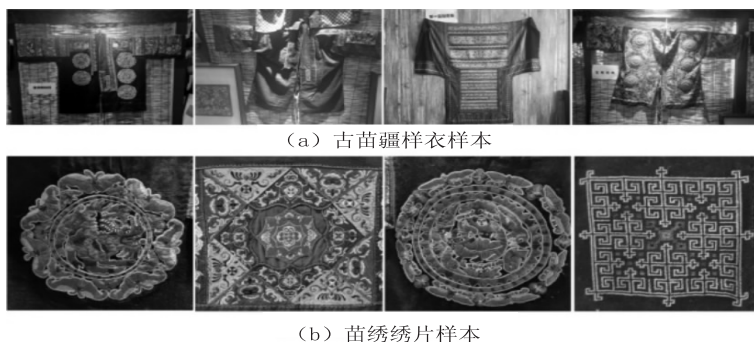


图 1 苗绣实例

但随着经济社会的发展,人们愈发追求快节奏生活,外来文化对少数民族传统文化产生了巨大影响。因此,苗绣的保存及发展也面临挑战:一方面,传统苗绣几乎都是纯手工制作,然而具备刺绣生产技能的一辈逐渐老去,同时年轻一代没有充分的动机学习苗绣技艺。另一方面,苗绣的制作流程相对复杂,从绣片选取到纹样设计几乎是纯手工劳动。相反,流水线生产制作简单,性价比高,挤压了传统苗绣的售卖市场,因此对刺绣数字化保护进行研究是必要的。受计算机视觉技术的启发,该文拟结合现代信息技术对苗绣资源进行设计、存储与传承,既体现了苗绣的历史特色,又结合了现代纹样的创新。

目前,苗绣绣片纹样的设计、挑选与分类主要由人工处理,这些任务较为繁琐且需要耗费大量人力、物力。随着计算机性能的提升,凡是涉及图像调整均可使用计算机进行处理,但传统计算机图像处理有其不足之处,如边缘检测或线条检测对噪声敏感且难以掌握最佳阈值。为改善传统方法的弊端,计算机也逐渐开始学习图像特征。随着深度学习技术,特别是卷积神经网络(Convolutional Neural Network, CNN)在视觉应用中取得了巨大成功,在医疗、教育、文化等领域得到了广泛应用。在医疗领域,Olaf 等人提出 U-net^[2]用于解决生物医学图像分割的问题,同时也是首次提出将图像语义分割技术应用于医学领域。在文化领域,贾小军等人^[3]提出一种基于神经网络的分类方法,用于构建蓝色花布纹理图案的图像元素。在纺织和服装领域,吴欢等人^[4]提出一种基于 CNN 模型的服装廓形分类方法,该方法用于对女性裤子的轮廓进行分类。

图像深度学习在刺绣纹样研究中也取得了相关成就。彭宇等人^[5]综述了实时语义分割的优势和限制,同时也提出面向快速语义分割方案。郑锐等人^[6]为解决刺绣风格数字化缺少立体感和线条方向等问题,提出一种基于深度学习的风格迁移方法,采用语义分割

网络对图像轮廓进行提取。Gatys 等人^[7]提出一种基于卷积神经网络的自然纹理参数化模型,纹理合成的质量和方法较以往有了实质性改进。Luan 等人^[8]提出在风格迁移过程中添加引导图以避免图像内容难以匹配等问题。王禹君等人^[9]提出基于空间领域的模糊均值算法用于处理少数民族刺绣纹样分割任务。通过以上分析,提取丰富图像信息是高质量语义分割的关键。为此,有关学者提出了扩张卷积,该方法可以在不降低空间分辨率下扩大图像感受野。DeepLab^[10]和 PSPNet^[11]通过捕获多尺度空间信息实现高效图像分割。RefineNet^[12]提出多路径细化模块,增强了图像特征信息的提取。Deeplab v3+^[13]将扩张卷积和空间金字塔池化集成到网络中进一步提高分割效率。

综上所述,将图像语义分割方法用于苗绣绣片纹样的提取,让网络理解图像中的语义信息并提炼出别具一格的苗绣纹样值得研究。因此,该文提出一种基于深度学习语义分割框架用于提取苗绣绣片纹样的轮廓,主要贡献有:(1)将苗绣文化与图形图像识别技术有机结合;(2)为弥补相关苗绣绣片纹样数据的稀少,提出并构建了苗绣绣片数据库,同时对其进行分类和分割的标注^[14];(3)针对传统图像处理提取刺绣纹样轮廓不连续、易丢失以及误分割等现象,提出一种多尺度深度学习网络。据笔者了解,很少有将深度学习模型应用于苗绣绣片纹样提取的研究,因此,文中工作在苗绣数字化保护中具有明显意义。

1 苗绣绣片数据库构建

为了深入研究苗绣绣片纹样,进一步保护非物质文化遗产,通过实地调查对收集到的图像进行数据库的构建。

在收集图像样本的过程中,分别从当地博物馆、手工作坊以及纹样设计公司获取苗绣纹样样本,基于相关数据和知识共收集 3 600 幅苗绣绣片纹样建立数据

库,用于开展不同类型研究。据笔者了解,这是第一个仅以苗绣绣片纹样为主的图像数据库。分类标准主要以线条类型、纹样大小、物体类型为主,每幅图像均

进行了类别标注以及用于语义分割的轮廓标注。已建立苗绣数据库的一些示例,如图 2 所示。

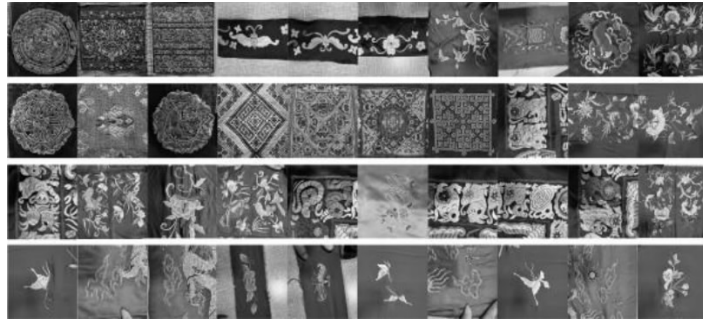


图 2 自建苗绣绣片图像数据库部分示例

2 基于多尺度网络的绣片纹样语义分割

本节主要介绍图像语义分割的多尺度网络 (MSNet)。首先,阐述网络的主要模块并分析相对于其他网络框架的优势;其次,描述 MSNet 的体系结构;最后,表明该网络目的在于改善语义分割任务中出现的“丢失”、错误分割以及遗漏分割等问题。

通常,带有尺度特征提取的卷积神经网络是一个分层结构。它由堆叠的卷积层组成,可选项是归一化层、池化层以及损失函数。卷积层通过线性卷积滤波器和非线性激活函数生成特征图 (Feature Map),然后将特征图馈送至损失函数中进行优化。卷积神经网络的训练包括两个阶段:前向阶段和反向阶段。前一阶

段用每层中的参数(权重和偏置)表示输入图像,最后一层的输出用于计算具有真值标签 (Ground-Truth Label) 的损失函数。同时采用链式法则用于反向传播更新梯度,并为下一次正向计算做好准备。

2.1 MSNet 架构

网络架构如图 3(a) 所示,采用类似 U-net^[2] 网络的编码器-解码器架构且图像尺寸的缩放倍率为 16。编码器部分包含特征金字塔下的采样操作,主要用于提取图像特征;解码器部分采用上采样双线性插值用于恢复图像的尺寸。其中每个特征提取层均包含多尺度特征提取步骤,同时采用跳跃连接^[15]维持编码器和解码器之间的特征关系。

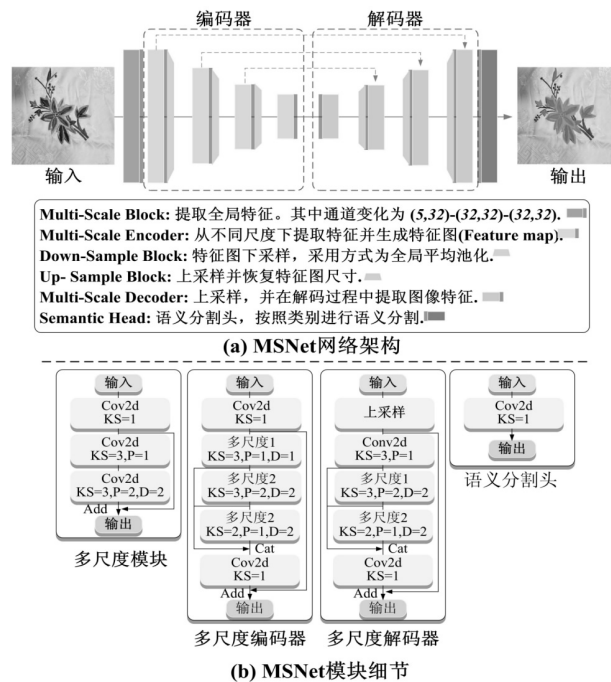


图 3 MSNet 整体流程

2.2 MSNet 网络细节

网络架构由四个模块组成:多尺度模块 (Multi-Scale Block, MSB)、多尺度编码器 (Multi-Scale

Encoder, MSE)、多尺度解码器 (Multi-Scale Decoder, MSD) 以及语义分割头 (Semantic-Head, SH)。同时, 为了进一步利用空间线索, 在编码器和解码器之间叠

加一系列跳跃连接来进行特征融合。模块详细分析如下:

多尺度模块 (Multi-Scale Block, MSB): 如图 3 (b) 所示, 图像语义分割的一个关键问题是如何提取上下文信息。因此有必要利用不同尺寸的卷积核获取全局信息。在该模块中将不同卷积核的输出依次叠加并通过 1×1 , 3×3 以及扩张比率分别为 1 和 2 的卷积核, 最后采用跳跃连接进行融合, 其中通道数变化为 $(5, 32) \rightarrow (32, 64) \rightarrow (64, 64)$ 。

多尺度编码器 (Multi-Scale Encoder, MSE): 该模块用于提取不同尺度的特征, 采用顺序堆叠+跳跃连接的形式, 下采样阶段采用全局平均池化用来避免卷积操作带来的额外开销。消融实验表明, 该步骤对最终的 MIoU 几乎无影响。因此在该类条件下, 特征图像的通道数变化为 $(64 \rightarrow 128 \rightarrow 256 \rightarrow 512)$, 尺寸变化为 $(224, 224) \rightarrow (14, 14)$ 。

多尺度解码器 (Multi-Scale Decoder, MSE): 与上采样操作和转置卷积相比, 采用多尺度特征+双线性插值构建解码器部分。主要原因有: (1) 上采样忽略了“参数”可学习因素, 最终结果可能会降低分割精度。相反, 文中方法具有无参数学习以及训练速度快的优点; (2) 转置卷积包含可学习因素, 但难以保证与编码器信息的对齐。因此, 该文采用多尺度特征+双线性插值替换常规转置卷积步骤, 并在消融实验中验证该步骤的合理性。

3 损失函数

从实际情况和数据集分析可知, 采用损失函数的标准是提高神经网络的泛化能力, 其中需要解决的问题包括: (1) 图像物体边界分割模糊; (2) 如何缓解场景中的类别失衡; (3) 如何更好地优化交并比 (MIoU)。对于上述问题, 使用损失函数的组合来监督模型训练, 主要包括交叉熵损失 L_{wce} , Lova 'sz - Softmax 损失以及边界损失 L_{bd} 。在语义分割任务中, 二维图像主要存在语义分割时物体边界模糊的问题。因此, 定义边界损失 L_{bd} 为:

$$L_{\text{bd}}(\hat{y}, y) = 1 - \frac{2P^c \cdot R^c}{P^c + R^c} \quad (1)$$

P^c, R^c 对于类别 C , 预测的边界特征 y_{bd} 相对于真实标签 y_{gt} 的精度和召回率。因此, 边界可定义为:

$$\begin{aligned} y_{\text{bd}}^b &= \text{pool}(1 - y_{\text{bd}}, \theta_0) - (1 - y_{\text{bd}}) \\ y_{\text{gt}}^b &= \text{pool}(1 - y_{\text{gt}}, \theta_0) - (1 - y_{\text{gt}}) \end{aligned} \quad (2)$$

$\text{pool}(\cdot)$: 大小为 θ_0 的滑动窗口上的最大池化操作。

为了解决类别标签不平衡问题, 通过类别出现频率开方的倒数来制定权重 Softmax 交叉熵损失函数:

$$L_{\text{wce}}(\hat{y}, y) = - \sum_i \alpha_i p(y_i) \log(p(\hat{y}_i)) \quad (3)$$

$$\alpha_i = \frac{1}{\sqrt{f_i}}$$

y_i, \hat{y}_i : 表示类别标签的真实值与预测值; f_i : 表示 C 类别标签出现的频率。同时使用 Lova 'sz - Softmax 损失函数来最大化交并比 (MIoU):

$$L_{\text{ls}} = \frac{1}{|C|} \sum_{c \in C} \Delta J_c(m(c)), m_i(c) = \begin{cases} 1 - x_i(c), & c = y_i(c) \\ x_i(c), & \text{otherwise} \end{cases} \quad (4)$$

$|C|$: 表示类别标签的数目; ΔJ_c : 雅可比矩阵索引的扩展项; $x_i(c) \in [0, 1]$, $y_i(c) \in \{-1, 1\}$: 表示类别标签 C 的第 i 个像素的真实标签概率和预测标签概率。最后, 损失函数的组合如下:

$$L = w_1 \cdot L_{\text{ls}} + w_2 \cdot L_{\text{wce}} + w_3 \cdot L_{\text{bd}} \quad (5)$$

w_1, w_2, w_3 : 表示不同损失函数的权重, 这些参数的组合形式将会在消融实验中展示。

4 实验

4.1 实验配置

为保证实验的有效性和公平性, 分别在公共数据集 PASCAL VOC 2012 数据集 (<http://host.robots.ox.ac.uk/pascal/VOC/voc2012/>) 和苗绣绣片自建数据库中进行实验, 主要用于验证模型的鲁棒性和泛化能力。

数据集: 实验选用数据集 I 是 PASCAL VOC 2012 的语义分割部分。该数据集作为公开的基准数据集, 主要用于验证图像分类、检测与语义分割等主流计算机视觉任务, 同时提供完善的指标对比方式。该数据集包含 1 个背景类别和 20 个物体对象类别, 其中包含 1 456 张测试用例、1 464 组训练用例以及 1 449 组验证用例。为了获取更多可用训练数据, 本实验依据数据增强方法额外生成 8 000 组训练数据; 数据集 II 是自建的苗绣绣片数据库, 共包含 3 600 张苗绣绣片纹样且采用数据增强进行扩充。但相较 PASCAL VOC 2012 数据集来说, 自建基础数据体量不充分, 因此仅按照 9:1 划分训练用例和测试用例, 不采用验证用例, 该策略主要针对小型数据集, 其目的是为了保证充分的训练用例。

数据增强: 数据增强作为一种重要的机器学习方法, 在基于已有的训练样本数据下来生成更多的训练数据, 目的是为了使扩增的训练数据尽可能接近真实分布的数据, 从而提高检测精度。实验中对输入图像做随机水平翻转、随机旋转以及随机缩放等数据增强操作。其中, 以 DeiT^[16] 为基准, 设置随机水平翻转的概率为 0.5, 随机缩放比例的范围为 $[0.5, 3]$, 随机旋

转范围为 $[-15^\circ, 15^\circ]$ 。该数据增强准则目前被广泛用于图像深度学习中,实验中对对比模型均以此为前置条件,以保证预测结果的公平性。

实验配置:本实验选用设备 CPU 为 Intel(R) Core(TM) i5-9300H CPU @ 2.40 GHz, GPU NVIDIA GeForce GTX 2080Ti;所用操作系统为 UBUNTU,网络框架为 Pytorch 1.7。

评价指标:本次实验采用的评估指标 I 为 MIoU,该指标主要是先计算每种类别真实标签和预测标签的两个对应像素点集合的交并比,然后在所有类别上取平均值。该指标的取值为概率形式,因此设定范围为 $[0,1]$ 。其中交并比越大说明图像的分割越精准,该指标的计算公式如下:

$$MIoU = \frac{1}{k+1} \cdot \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (6)$$

k :表示数据集中的类别标签总数; p_{ii} :表示第 i 类标签被分配到第 i 类类别的像素数量; p_{ij} :表示第 i 类标签被分配到第 j 类类别的像素数量,即将 i 类别预测为 j 类别; p_{ji} :表示将所有的 j 类别预测为 i 类标签的像素数量。评价指标 II 为模型推理速度和模型参数量。主要用于评估模型的迁移特性和泛化能力。

训练策略:训练策略即正则化方法,由于神经网络较为复杂且拟合能力强,很容易在训练数据上产生过拟合。因此在训练神经网络时需要通过一定的正则化方法来改进网络的泛化能力。该实验以全景分割方法 Panoptic segformer^[17] 为基准设置训练策略,使用随机梯度下降(Stochastic Gradient Descent, SGD)算法,权重衰减为 0.000 1,基础学习率为 0.1,动量为 0.9,训练迭代次数为 300,实验涉及的模型均采用该训练策略。

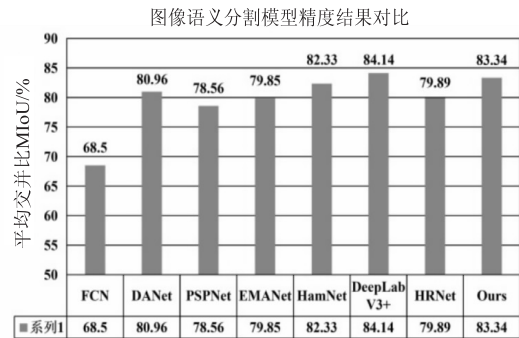
4.2 不同模型对比分析

在 PASCAL VOC 2012 数据集上进行相关实验,主要从两方面阐述:(1)与不同的语义分割框架对比,包括 PSPNet^[11]、FCN^[18]、DeepLab^[10,13] 系列框架等,如表 1 所示;(2)从模型的推理速度和模型参数上进行对比。

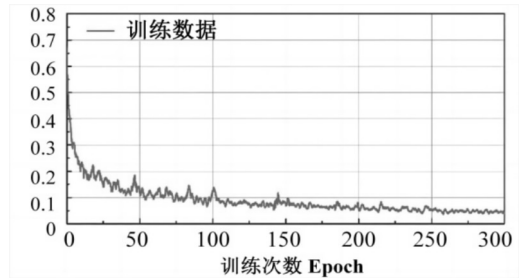
表 1 图像语义分割模型精度对比

模型	交并比(MIoU)
FCN ^[18]	68.50
DANet ^[19]	80.96
PSPNet ^[11]	78.56
EMANet ^[20]	79.85
HamNet ^[21]	82.33
DeepLab V3+ ^[13]	84.14
HRNet	79.89
Ours	83.34

由表 1 和图 4 可知,文中方法(MSNet)的 MIoU 相比于 FCN 提高了 14.84 百分点,提升幅度相对较大的原因是 FCN^[18] 作为首个语义分割框架,在目前看来仍有许多可以优化的技术,由于 FCN 解码器中特征融合的方式较为单一,所以相比现代语义分割方法来说精度最低。相比于近期主流框架 DANet^[19]、PSPNet^[11] 以及 DeepLab V3+^[10,13] 分别提高了 2.38 百分点、4.78 百分点以及 2.2 百分点。同时,MSNet 在训练数据上的损失值变化曲线如图 4(b)所示,在前期损失值快速下降,最终稳定在 0.03 区间范围,从而体现了网络可以准确、快速地寻找到梯度下降的方向。



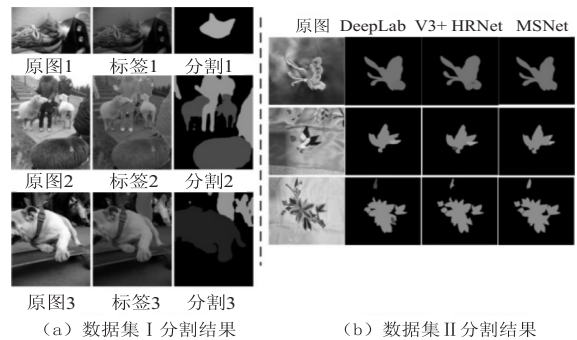
(a) 模型分割精度结果对比



(b) MSNet 网络损失值变化

图 4 模型训练及精度指标对比

文中方法的优势主要体现在提出的多尺度特征提取模块,由于感受野在每一阶段都会做出相应的扩大,从而可以更好地提取特征。这同时也表明了在同一阶段,局部信息与全局信息的融合对模型性能的提升有很大的帮助,相关实验展示如图 5 所示。



(a) 数据集 I 分割结果

(b) 数据集 II 分割结果

图 5 模型训练及精度指标对比

如图 5(b)所示,在自建苗绣数据库上进行纹样语义分割测试,对比框架为 DeepLab V3+ 和 HRNet。实

验结果表明, MSNet 在提取纹样边界上明显优于前者, 同时误匹配、误分类现象也相对降低。实验结果可知, MSNet 在面对颜色相近纹样时可以更好地提取边界特征。

表 2 展示了相同实验条件下不同模型推理速度和参数量的对比。MSNet 在模型参数量上相较 FCN 增加了 32.5 百分点, 但精度提升了 14.84 百分点。DeepLab V3+ 的 MIoU 指标与 MSNet 基本一致, 但参数量增加了 39.4 百分点, 在后期推理速度上明显低于文中方法。由此可知, MSNet 可以更好地迁移至嵌入式平台。

表 2 推理速度和模型参数分割结果对比

模型	参数量 m ($\times 10^6$)	FLOPs/G
FCN	24.2	193.7
DeepLab V3+	59.3	240.3
EMANet	61.1	246.1
HamNet	69.1	277.9
PSPNet	51.4	208.3
Ours	35.9	197.1

浮点运算 (Floating Point Operation Per Second, FLOPs) 是衡量神经网络效率的度量标准之一, 表 2 比较了 MSNet 与不同模型的 FLOPs 运算量。对于相同运算量的 FLOPs, MSNet 在模型参数量以及运算速度上有着明显优势。但同样需要注意的是, 浮点运算并非完全代表模型的效率, 因为该指标并未考虑与网络推理相关的因素, 如内存访问、并行特性和平台特性。

4.3 消融实验

为了验证不同模块的作用, 进行了一系列消融实验, 包括模型参数、FLOPs (浮点运算)、损失函数的分配比、解码器选择以及训练策略的对比, 同时从常用的双线性插值上采样方式逐步扩展到当前模型, 以增量式消融实验来验证模型性能。

表 3 图像语义分割方法参数量对比
(Decoders: 解码器)

方法	模型堆叠方式	交并比 (MIoU)	参数量 m ($\times 10^6$)	时间/s
MSNet	5C+BI	79.53	31.8	0.069
	4C+M+BI	79.61	33.1	0.073
	3C+2M+BI	80.29	33.9	0.076
	2C+3M+BI	80.35	34.7	0.076
	C+4M+BI	81.02	35.2	0.077
	5M+BI	81.19	35.3	0.081
	5C+Decoders	81.68	32.4	0.077
	5M+Decoders	82.31	35.9	0.089

上采样模型堆叠: 为验证网络特性, 将解码器部分

替换为双线性插值上采样步骤, 并逐步替换该文所提的多尺度方法, 主要阶段有: 5C+BI (5C: 卷积下采样, BI: 双线性插值上采样); 增量模式同时包括: (4C+M+BI), (3C+2M+BI), (2C+3M+BI), (C+4M+BI) 和 (5M+BI), 其中 M 表示多尺度模块, 结果如表 3 所示。

损失函数组合比例: 损失函数的选择同样会影响网络的性能, 该文选择的损失函数组合为: $w_1 \cdot L_{ls} + w_2 \cdot L_{wce} + w_3 \cdot L_{lbd}$ 。首先, 仅采用 L_{ls} 和 L_{wce} 损失函数组合, 但是在进行大量实验之后发现边界损失的比例也会明显影响最终的平均交并比, 如表 4 所示。

表 4 损失函数系数比率对比

损失函数组合	w_1	w_2	w_3	交并比 (MIoU)
$w_1 \cdot L_{ls} + w_2 \cdot L_{wce} +$	1	1	0	82.31
$w_3 \cdot L_{lbd}$	1	1	1	82.95
	1	1	1.5	83.34

5 结束语

图像语义分割任务是计算机视觉领域的研究热点, 相较于传统图形学, 语义分割在提取物体轮廓上有着更大优势。该文首先建立了一个苗绣纹样片纹样的图像数据库, 并为每个图像手动分配一个准确的苗绣类别标签。然后, 在基于公共数据库和自建数据库的基础上提出了一种基于深度学习的多尺度语义分割方法, 可以有效弥补通用框架在此类问题上的缺陷, 例如绣片纹样轮廓错误分割、遗漏分割等。实验结果表明, 模型不仅能对边界特征进行高效提取, 还在精度与速度之间取得平衡。此外, 从文中方法的有效性上表明, 将深度学习与传统绣片特色相结合是一种比较有意义的方式。因此, 在未来的工作中, 会继续探索与多尺度特征融合相关的网络及触发点, 同时结合风格迁移对苗绣纹样片的再设计做出更多研究。

参考文献:

- [1] 谢蕊. 贵州苗族女性传统服饰技艺及其文化的变迁与发展研究[D]. 北京: 中央民族大学, 2022.
- [2] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[C]// Medical image computing and computer-assisted intervention - MICCAI 2015: 18th international conference. Munich: [s. n.], 2015: 234-241.
- [3] 贾小军, 叶利华, 邓洪涛, 等. 基于卷积神经网络的蓝印花布纹样基元分类[J]. 纺织学报, 2020, 41(1): 111-117.
- [4] 吴欢, 丁笑君, 李秦曼, 等. 采用卷积神经网络 CaffeNet 模型的女裤廓形分类[J]. 纺织学报, 2019, 40(4): 118-121.
- [5] 彭宇, 姬森展, 于希明, 等. 语义分割网络的 FPGA 加速

- 计算方法综述[J]. 仪器仪表学报, 2021, 42(9): 1–12.
- [6] 郑 锐, 钱文华, 徐 丹, 等. 基于卷积神经网络的刺绣风格数字合成[J]. 浙江大学学报: 理学版, 2019, 46(3): 270–278.
- [7] GATYS L A, ECKER A S, BETHGE M. Texture synthesis and the controlled generation of natural stimuli using convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2015, 28: 1–9.
- [8] LUAN F J, PARIS S, SHECHTMAN E, et al. Deep photo style transfer[C]//2017 IEEE conference on computer vision and pattern recognition (CVPR). Honolulu: IEEE, 2017: 4990–4998.
- [9] 王禹君, 周菊香, 徐天伟. 改进模糊 C 均值算法在民族服饰图像分割中的应用[J]. 计算机工程, 2017, 43(5): 261–267.
- [10] 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.
- [11] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]//2017 IEEE conference on computer vision and pattern recognition (CVPR). Maryland: IEEE, 2017: 2881–2890.
- [12] LIN G, MILAN A, SHEN C, et al. Refinenet: multi-path refinement networks for high-resolution semantic segmentation [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Maryland: IEEE, 2017: 1925–1934.
- [13] 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). Munich: [s. n.], 2018: 801–818.
- [14] 韦 婷, 李馨蕾, 刘 慧. 小样本困境下的图像语义分割综述[J]. 计算机工程与应用, 2023, 59(2): 1–11.
- [15] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas: IEEE, 2016: 770–778.
- [16] TOUVRON H, CORD M, DOUZE M, et al. Training data-efficient image transformers & distillation through attention [C]//International conference on machine learning. Vienna: [s. n.], 2021: 10347–10357.
- [17] LI Z, WANG W, XIE E, et al. Panoptic segformer: delving deeper into panoptic segmentation with transformers [C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. New Orleans: IEEE, 2022: 1280–1289.
- [18] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). Boston: IEEE, 2015: 3431–3440.
- [19] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation [C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Long Beach: IEEE, 2016: 3146–3154.
- [20] LI X, ZHONG Z, WU J, et al. Expectation-maximization attention networks for semantic segmentation [C]//Proceedings of the IEEE/CVF international conference on computer vision. Seoul: IEEE, 2019: 9167–9176.
- [21] GENG Z, GUO M H, CHEN H, et al. Is attention better than matrix decomposition? [C]//International conference on learning representations. Barcelona: [s. n.], 2021: 1–20.