

摘 要

目的：基于语义分割算法在膀胱镜图像中建立深度学习模型，对膀胱肿瘤的膀胱镜图像进行诊断分类，区分病理分级，为膀胱肿瘤的治疗方式和预后随访提供依据。

方法：回顾性收集我院 2010 年 1 月-2021 年 8 月收治的 261 例膀胱肿瘤患者的膀胱镜图像。根据不同病理分级的膀胱肿瘤在膀胱镜下的图像特征将图像分为 3 类，其中乳头状瘤或低度恶性潜能的乳头状瘤 116 例，185 张图像；低级别尿路上皮癌 63 例，97 张图像；高级别尿路上皮癌 42 例，62 张图像。经数据扩增，最终得到 2408 张图片，并以 8:1:1 的比例进行随机划分得到训练集，测试集和验证集。采用多层次特征语义融合分割算法（multilevel feature semantic fusion network, MFSFNet）对图像进行分割。并将其与 4 种经典语义分割模型在我们的数据集上进行对比实验，最后将外部数据集按临床信息（是否浸润肌层、性别、年龄、是否吸烟、有无膀胱癌家族史）进行分组探索模型的临床应用价值，预测膀胱镜图片的分类，采用平均像素精度（Mean Pixel Accuracy, MPA）、平均交并比（Mean Intersection over Union, MIoU）、交叉熵损失函数、模型参数量（Parameters）、画面每秒传输帧数（Frame Per Second,FPS）作为模型实验分割的评价指标，并计算模型图像分类的准确率。

结果：经过图像识别训练的 MFSFNet 区分乳头状瘤或低度恶性潜能乳头状瘤、低级别尿路上皮癌、高级别尿路上皮癌的 MPA 为 95.93%和 MIoU 为 92.54%，其图像分类的准确率 96%，相较于其他 4 类经典的模型具有一定的优势。在分组实验中，整体图像分类准确率为 51%，非浸润组、男性组、<60 岁、吸烟组及有膀胱癌家族史组中图像分类准确率具有更高的准确率，分别为 56%、56%、54%、53%、71%。

结论：MFSFNet 网络模型在对膀胱镜图像上的膀胱肿瘤分割的效果良好，同时能较好的对膀胱肿瘤图像进行分类，证明了深度学习语义分割模型在膀胱肿瘤分割任务中的应用前景，具有一定的临床应用价值，能够为膀胱肿瘤患者的早期诊断分类和个性化治疗提供有价值的信息，促进人类健康事业的发展。

关键词：人工智能；深度学习；语义分割；膀胱肿瘤；膀胱镜图像

ABSTRACT

Objective To establish a deep learning model in cystoscopy images based on semantic segmentation algorithm, classify cystoscopy images of bladder tumors, distinguish pathological grades, and provide basis for the treatment of bladder tumors and prognostic follow-up.

Methods Cystoscopy images of 261 patients with bladder tumor admitted to our hospital from January 2010 to August 2021 were retrospectively collected. According to the characteristics of cystoscopic images of bladder tumors with different pathological grades, the images were divided into 3 categories, including 116 cases of papilloma or papilloma with low malignant potential, 185 images. 63 cases with low-grade urothelial carcinoma, 97 images; 42 cases of high-grade urothelial carcinoma, 62 images. After data amplification, 2408 images were obtained and randomly divided at a ratio of 8:1:1 to get the training set, test set and verification set. The multilevel feature semantic fusion network (MFSFNet) was used to segment the image. It was compared with four classical semantic segmentation models on our data set. Finally, external data sets were grouped according to clinical information (whether there is infiltration of muscle layer, gender, age, smoking, and family history of bladder cancer) to explore the clinical application value of the model and predict the classification of cystoscopy pictures. Mean Pixel Accuracy (MPA), Mean Intersection over Union, mean pixel accuracy (mpa), mean intersection over Union, MIoU), cross entropy loss function, model Parameters and Frame Per Second (FPS) were used as evaluation indexes of model experiment segmentation, and the accuracy of model image classification was calculated.

Results After image recognition training, MFSFNet distinguished papilloma or low-malignant potential papilloma, low-grade urothelial carcinoma, and high-grade urothelial carcinoma with 95.93% MPA and 92.54% MIoU, and its image classification accuracy was 96%, which had certain advantages compared with the other 4 types of classical models. In the grouping experiment, the overall image classification accuracy was 51%, and the non-invasive group, male group, <60 years

old group, smoking group and the group with a family history of bladder cancer had higher accuracy, which were 56%, 56%, 54%, 53% and 71%, respectively.

Conclusion MFSFNet network model has a good effect on the segmentation of bladder tumors on cystoscopy images, and can better classify bladder tumor images, which proves the application prospect of deep learning semantic segmentation model in the task of bladder tumor segmentation and has certain clinical application value. It can provide valuable information for early diagnosis classification and personalized treatment of patients with bladder tumor, and promote the development of human health.

Key words : Artificial intelligence;Deep learning;Semantic segmentation;Bladder tumor;Cystoscopic image

目 录

第1章 前言	1
第2章 临床资料与方法	4
2.1 膀胱镜设备与操作	4
2.2 数据收集	4
2.2.1 纳入标准与排除标准	5
2.2.2 临床数据收集	5
2.2.3 膀胱镜图像数据收集	5
2.3 研究思路与算法构架	6
2.4 数据处理与实验分析	7
2.4 算法对比实验与分组实验	10
2.6 评价指标	11
2.7 损失函数	12
第3章 结果	13
3.1 算法对比实验	13
3.2 分割结果可视化	14
3.3 分组实验	14
第4章 分析与讨论	18
第5章 总结与展望	20
致 谢	21
参考文献	22
综 述	25

中英文缩略词表

英文缩写词	英文全称	中文全称
WHO	World Health Organization	世界卫生组织
NMIBC	non-muscle-invasive bladder cancer	非肌层浸润性膀胱肿瘤
MIBC	muscle-invasive bladder cancer	肌层浸润性膀胱肿瘤
TURBT	transurethral resection of bladder tumor	经尿道膀胱肿瘤电切术
AI	Artificial intelligence	人工智能
ML	Machine Learning	机器学习
DL	Deep Learning	深度学习
NLP	Natural Language Processing	自然语言处理
CNN	Convolutional Neural Network	卷积神经网络
MRI	magnetic resonance imaging	核磁共振成像
MFSFNet	multi-level feature semantic fusion network	多层次特征语义融合分割算法
PSPNet	pyramid scene parsing net	金字塔场景解析网络
HRNet	high-resolution net	高分辨率网络
CASPP	Compound Atrous Spatial Pyramid Pooling	复合空洞空间金字塔池化
UICC	Union for International Cancer control	国际抗癌联盟
SGD	Stochastic Gradient Descent	随机梯度下降
MPA	Mean Pixel Accuracy	平均像素精度
MIoU	Mean Intersection over Union	平均交并比
RC	radical cystectomy	根治性膀胱切除术
CT	computed tomography	电子计算机断层扫描

第1章 前言

膀胱肿瘤是泌尿系统常见的肿瘤，好发部位为膀胱侧壁和后壁。膀胱肿瘤的发病机制尚未明确，高龄、男性和吸烟以及有膀胱癌家族史等因素都会导致膀胱癌的发生，其最常见的临床表现为血尿^[1]。据世界卫生组织（World Health Organization, WHO）对全球 185 个国家的 36 种癌症的研究数据显示，膀胱癌是全球第 10 大最常诊断的癌症，2020 年约有 57.3 万例新发病例和 21.3 万例死亡病例，并且全球男性的发病率和死亡率约为全球女性的 4 倍，相同分期的膀胱癌，女性的预后比男性差^[2]。

膀胱肿瘤根据是否浸润肌层可分为非肌层浸润性膀胱肿瘤（non-muscle-invasive bladder cancer, NMIBC）和肌层浸润性膀胱肿瘤（muscle-invasive bladder cancer, MIBC），大多数患者(75%)表现为 NMIBC^[3]。目前 NMIBC 的治疗方式是手术治疗辅以放化疗，手术治疗常规标准术式是经尿道膀胱肿瘤电切术（transurethral resection of bladder tumor, TURBT）^[4]。NMIBC 患者微创手术治疗后 2 年内有 50%-70% 的病例复发，其中约 40% 的病例 5 年后进展至 MIBC^[5,6]。有研究表明膀胱肿瘤患者的早期复发可能是膀胱镜检查 and 经尿道膀胱肿瘤切除术(TURBT)未发现病变的结果^[7]。

根据 WHO2016 年病理分级标准，膀胱肿瘤可分为乳头状瘤、低度恶性潜能乳头状瘤、低级别尿路上皮癌、高级别尿路上皮癌^[8]。根据不同病理分级，乳头状瘤、低度恶性潜能的乳头状尿路上皮肿瘤及低级别尿路上皮癌一般推荐行经尿道膀胱肿瘤电切术，而高级尿路上皮癌一般推荐行膀胱癌根治术+淋巴清扫。膀胱镜是诊断膀胱肿瘤的重要检查，是临床上诊治膀胱肿瘤必不可少的一环^[9]。根据膀胱镜下膀胱肿瘤形状、瘤体有无血流、有无蒂以及蒂的宽窄可将不同病理分级的膀胱肿瘤分为 3 类（乳头状瘤或低度恶性潜能的乳头状瘤；低级别尿路上皮癌；高级别尿路上皮癌）^[10]。然而对于膀胱肿瘤的结果的识别仍具有挑战性，这结果因检查人员技能、工作环境以及劳动强度而有所不同，因此具有一定的误差和主观性^[11]。在这种情况下，使用特征提取和深度学习的计算机辅助诊断工具有望作为进行诊断分类的工具，并实现可重复的膀胱镜检查结果，从而减少不必要的活检，并且能有效地缓解医生的工作强度、减少误差，有利于提升诊疗水平。

人工智能 (Artificial intelligence, AI) 是基于复杂的非线性数学计算机模拟系统对已知的数据进行处理, 执行从对象识别等小型可管理任务到预测等复杂任务, 并估计有关未来的未知数据, 涉及各种人工智能子领域, 如机器学习 (Machine Learning, ML)、深度学习 (Deep Learning, DL)、自然语言处理 (Natural Language Processing, NLP)、计算机视觉、预测分析、视觉识别等领域, 其中大多数用于当今的医学和医疗保健^[12]。其中医学图像分析 (Medical image analysis, MIA) 领域衍生于计算机视觉领域, 而计算机视觉领域又是人工智能或计算机科学的一个子学科。医学图像分析在临床诊断、图像引导手术和模式发现中发挥着核心作用。机器学习是在算法的帮助下教会计算机做出准确预测的过程, 这些算法经过训练, 并从模型中的过去经验中学习, 将特征映射到相应的结果变量, 是人工智能不可或缺的一部分。机器学习可以根据不同的原理组织分成不同的类别, 其中根据训练样本标签的利用率, 可以分为监督学习、半监督学习和无监督学习。在医学图像分析领域中, 基本上采用的是监督学习。深度学习是 Hinton 在 2006 年提出的, 其利用多个计算机处理层来发现高维数据中复杂结构的方法^[13]。与传统机器学习方法相比, 深度学习最大的优势是能够自动提取图像的特征, 无需手动。随着深度学习在计算机视觉以及自然语言处理方面取得进展, 以卷积神经网络(Convolutional Neural Network, CNN) 为基础的深度学习开始应用于医学影像分析领域^[14]。Zunair^[15]等人对深度卷积神经网络的研究证实其能成功应用于医学图像的分析任务。Eminaga^[16]等人对 479 例患者的膀胱镜图像进行处理, 得到 18681 张图像, 再对这些图像预处理后开发了 5 种深度卷积网络模型, 这些模型经验证都能正确的区分正常膀胱粘膜和膀胱肿瘤。国内也有学者通过构建卷积神经网络模型对膀胱镜图像数据进行处理, 该模型显示出一定的辅助诊断病理分级价值^[10]。

医学图像分割是医学影像分析中的重要内容, 是通过各类方法和算法分析所涉及的医学视频或图片内容, 进而使得人工智能系统理解周围环境信息的重要环节, 而基于深度学习的医学图像语义分割是当前的热门研究^[17]。Renard^[18]等人展示了使用深度学习算法进行分割的重要性。目前有很多学者基于深度学习来构建医学图像语义分割模型, Liu^[19]等人展示了语义分割在乳腺癌诊断中的应用。Meraj^[20]等人对肺部 CT 影像分割的研究, 用于肺结节检测。也有学者利用一种迁移学习的方法对 CT 影像中的膀胱肿瘤进行语义分割, 并取得较高的性能^[21]。Shkolyar^[22]等人在 2019 年报道了使用深度学习在膀胱镜图像上准确分

割膀胱肿瘤，并且开发了一个图像分析平台 CystoNet。国内也有学者将改进的语义分割算法应用于肾脏及肿瘤图像^[23]。邬硕^[24]等人基于混合神经网络进行改进，将改进好的语义分割算法应用于脑部核磁共振成像(magnetic resonance imaging, MRI)的图像，获得了较好的结果。张娜^[25]等人设计了一种多尺度特征融合的轻量化膀胱癌分割算法对膀胱癌 MRI 图像进行分割，结果显示该算法分割速度快，具有一定的辅助诊断价值。

综上所述，目前国内关于膀胱肿瘤的膀胱镜图像分割的研究较少，本文利用膀胱镜图像结合语义分割算法，采用一种多层次特征语义融合分割算法(multi-level feature semantic fusion network,MFSFNet)，对三种类别的膀胱肿瘤图像进行分割，判断其像素级别对肿瘤特征提取的能力以及对膀胱镜图像分类的准确性，将其与目前 4 种经典的语义分割算法(金字塔场景解析网络(pyramid scene parsing net,PSPNet)^[26]、U-Net^[27]、DeepLabv3+^[28]、高分辨率网络(high-resolution net,HRNet)^[29])相比较，同时也根据病人的临床数据(是否浸润肌层、性别、年龄、是否吸烟、有无膀胱癌家族史)对外部数据集进行再分类，探索模型在不同情况下的临床应用价值，为膀胱肿瘤患者提供早期诊断及个性化治疗提供有价值的信息。

第2章 临床资料与方法

2.1 膀胱镜设备与操作

本研究所有患者的膀胱镜检查均在我院膀胱镜室进行，膀胱镜规格型号为奥林巴斯 CLV-S40，其是由内窥镜、鞘套及闭孔器、操作器、窥镜桥清洗杆及配套手术器械（软性活检钳、软性锯齿钳、软性异物钳和软性剪刀）组成。在进行膀胱镜检查前，嘱咐患者注意事项，使用盐酸奥布卡因凝胶进行局部麻醉。之后取已经装好闭孔器的膀胱尿道镜镜鞘，从尿道外口进入，女性病人应仔细辨认尿道口，将镜鞘缓慢送入尿道，进入膀胱后有尿液自冲水开关流出。男性病人应左手提起阴茎，右手执镜鞘，将镜鞘自尿道外口送入到尿道膜部处，后将镜鞘缓慢往下压，使镜鞘缓慢地进入膀胱，镜鞘进入过程切勿暴力进镜，以免产生假道。如若进入阻力过大，可采用尿道扩张器扩张后再尝试进镜。待镜鞘进入膀胱后，抽出闭孔器，安放内窥镜，视尿液的清晰度进行放水和灌水。检查膀胱时应系统、全面的观察，不能遗漏。应利用膀胱尿道镜的旋转、角度变化以及进退进行观察，一般先将膀胱镜尿道镜推向三角区近端，然后沿着镜的轴心进行旋转观察，观察一周，之后便边退镜边旋转观察，最后检查膀胱颈部。检查过程可根据输尿管口判断位置，注意观察膀胱表面的血管纹理、颜色以及是否有肿物，正常膀胱黏膜平整，为粉红色，并且可以清晰地看到血管纹理，左右输尿管口之间有输尿管间襞。如有肿物，将镜头对准肿物并拉近镜头，周围尿液应尽量保持清晰，拍取清晰完整以及不同角度的肿物照片，如有需要则夹取肿物进行病理活检，检查完毕，退出观察镜，放出膀胱内液体，安装闭孔器，退出镜鞘，检查完毕，嘱咐患者注意观察尿液情况并多喝水。行膀胱镜检查的医生均具有3年以上临床经验，能熟练的操作膀胱镜。

2.2 数据收集

选取2011年01月至2021年10月在我院泌尿外科行经尿道膀胱肿瘤电切术（TURBT）且术后病理均确诊为尿路上皮肿瘤的共261名患者，其中外部测试集患者人数为40人。

2.2.1 纳入标准与排除标准

纳入标准：（1）病例筛选：临床病理信息完整，分级准确，无病理诊断争议的病例；（2）图片筛选：图片质量优良，无阴影；（3）手术方式：经尿道膀胱肿瘤电切术。

排除标准：（1）病理诊断不明确的病例；（2）有明显干扰扫描因素的图片，如模糊以及肿瘤显露不佳等；（3）未行手术或者行膀胱癌根治术。

2.2.2 临床数据收集

通过我院电子病历系统收集患者的基本临床资料，包括性别、年龄、是否吸烟、有无膀胱癌家族史等。查询术后记录，收集患者的病理资料，包括肿瘤的病理分级与分期。病理报告先是由一名中级职称医师初步书写，再由一名高级职称医师严格审阅。其中病理分期参照 2014 版中国泌尿外科指南，采用国际抗癌联盟（Union for International Cancer control，UICC）的 2009 年第 7 版 TNM 分期标准。病理分级根据 WHO2016 年病理分级标准，可分为乳头状瘤、低度恶性潜能乳头状瘤、低级别尿路上皮癌、高级别尿路上皮癌。

2.2.3 膀胱镜图像数据收集

根据临床数据资料完整且符合纳入标准的患者信息检索膀胱镜系统，收集对应患者的膀胱镜图像，图像均保存为 JPG 格式，再由专业技术人员协助导出图像。

2.3 研究思路与算法构架

本研究技术路线图如下

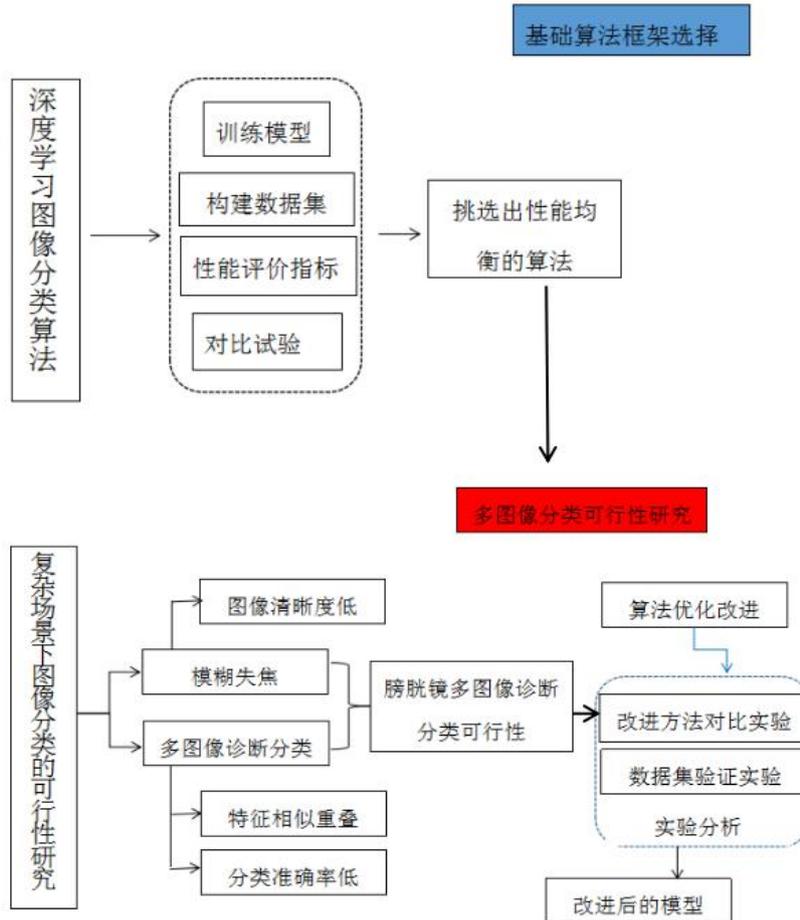


图 1. 技术路线图

如图 2 所示本文整体网络结构由编码器和解码器两部分组成。编码器-解码器结构特别适合图像分割任务，因为它们能够在潜在空间中对输入图像的复杂底层特征进行建模，然后生成与输入大小相同的输出图像，即包含每个像素分类的分割图^[30]。编码器部分包括 MobileNetV2^[31]和复合空洞空间金字塔池化（Compound Atrous Spatial Pyramid Pooling, CASPP）模块。该算法模型采用轻量的 MobileNetV2 作为网络主干进行特征提取，生成高级语义特征图。再利用 CASPP 模块用于获取输入特征图的多尺度信息，该模块将级联空洞卷积与条状

池化模块并行设计，利用级联空洞卷积在控制感受野和分辨率情况下获取多尺度上下文信息，通过条状池化模块对上下文信息补充和完善，加强全局特征信息的捕获^[32]。同时增加小膨胀率 $r=3$ 的分支帮助小范围局部特征的提取，总的采用互质膨胀率 3、5、9、17 作为 CASPP 模块的膨胀率，有助于提高了像素的利用率，将目标细节特征更充分的传递到更深的网络层，进而提高网络的分割能力。解码器部分采用特征融合模块进行多尺度特征融合，构建不同层次特征像素之间的位置对应关系，将深浅层次特征自底向上逐级融合，提升模型的分割性能。

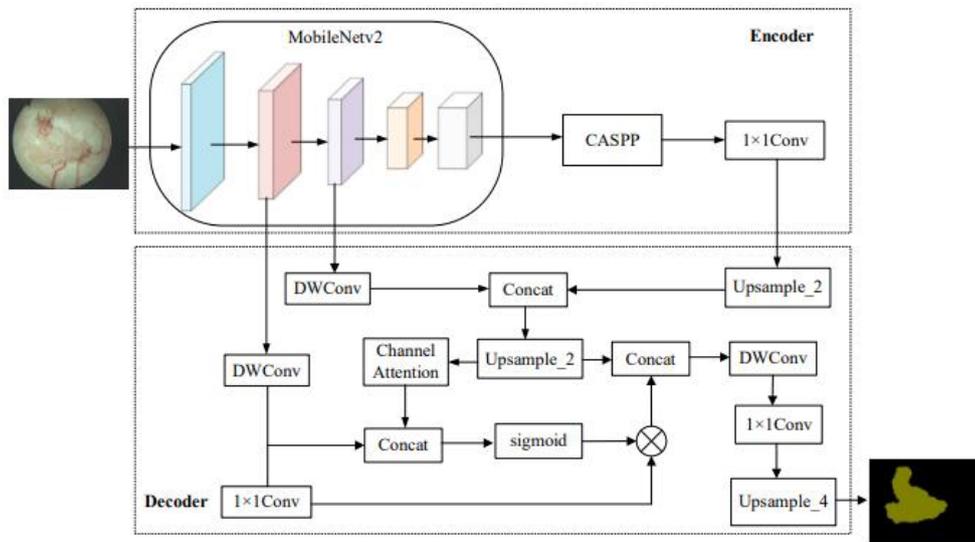


图 2. 整体网络结构

2.4 数据处理与实验分析

所有病例中男性 159 人，女性 62 人，年龄 24-81 岁，按不同病理分级的膀胱肿瘤在膀胱镜下的图像特征分为 3 类，如表 2 所示其中乳头状瘤或低度恶性潜能的乳头状瘤 116 例，185 张图像；低级别尿路上皮癌 63 例，97 张图像；高级别尿路上皮癌 42 例，62 张图像。本研究将 221 个膀胱肿瘤患者的 344 张膀胱镜图像分为 3 个类别，分别为 High bladder tumor（高级别尿路上皮癌），Low bladder tumor（低级别尿路上皮癌），Papilloma（乳头状瘤或低度恶性潜能的乳头状瘤）。一名具有多年临床工作经验的医生使用 Labelme 软件进行语义分割标签的制作。由于本研究样本较少，为了防止过拟合以及数据类别不均衡，通过翻转、旋转、镜像、缩放、剪切等方式进行数据增强（如图 3 示），

扩大训练样本，以增强模型的泛化能力，最终得到 2408 张图片，并以 8:1:1 的比例进行随机划分得到训练集，测试集和验证集（训练集 1950 张，测试集 241 张，验证集 217 张），其中训练集、验证集用于建立模型，测试集用于测试模型。为区分图像，测试前设定图像具有乳头状瘤或低度恶性潜能的乳头状瘤图像特征的输出结果为黄色，具有低级别尿路上皮癌图像特征的输出结果为绿色，具有高级别尿路上皮癌图像特征的输出结果为红色，分类正确的图像为 1 种颜色并分类正确，如有 2 种或 3 种颜色则为分类不正确，表 1 为 221 名患者的临床信息表。本研究所有的训练和测试图像尺寸均设置为 512x512 大小，在训练期间，使用随机梯度下降（Stochastic Gradient Descent, SGD）优化器，动量参数为 0.9，采用余弦退火对学习率进行更新，初始学习率设置为 0.007，随着训练次数增加，学习率逐渐降低，BatchSize 设置为 2，总共迭代 500epochs。实验引入迁移学习的策略，将在 ImageNet 数据库上训练好的 MobileNetV2 网络权重作为实验的初始权重，使用迁移参数初始化网络提升泛化性能。表 3 为实验操作环境。

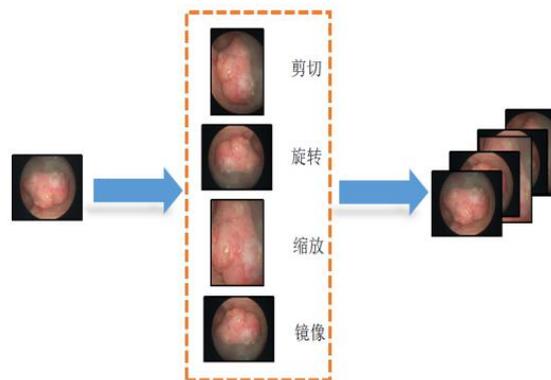


图 3. 图像扩增

第 2 章 临床资料与方法

表 1. 临床资料分类

临床特征	病人数量
性别	
男	159 人
女	42 人
年龄	
<60	77 人
≥60	144 人
T 分期	
<T2	178 人
≥T2	43 人
吸烟	
是	97 人
否	124 人
膀胱癌家族史	
是	29 人
否	192 人

表 2. 图像分类信息表

图像分类	病人数量	原始数据
乳头状瘤或低度恶性潜能的乳头状瘤	116 人	185 张
低级别尿路上皮癌	63 人	97 张
高级别尿路上皮癌	42 人	62 张
	221 人	344 张

表 3. 实验操作环境

名称	环境参数
操作系统	Windows11 64 位
CPU	Intel(R) Core(TM) i9-10850K CPU @ 3.60GHz
GPU	GeForce RTX 3070Ti(8.0 GB)
编程语言	Python 3.8
深度学习框架	Pytorch1.7.1、CUDA11.1

2.5 算法对比实验与分组实验

将本研究的模型与目前4种经典的语义分割算法（金字塔场景解析网络（pyramid scene parsing net,PSPNet）、U-Net、DeepLabv3+、高分辨率网络（high-resolution net,HRNet））相比较，探索算法的创新性。为保证具有可比性，5种算法使用的均是本研中的测试集。分组实验中采用的是外部数据集，其中有40名患者，男性有32人，女性有8人，共49张膀胱镜图像，图像不采用数据增强的方式扩增数量，年龄范围为24-77岁，该数据同样按不同病理分级的膀胱肿瘤在膀胱镜下的图像特征分为3类。再根据是否浸润肌层、性别、年龄、是否吸烟、有无膀胱癌家族史）进行亚组探索模型的临床应用价值。表4和表5为外部数据集的患者基本信息表和膀胱镜图像分组信息表。

表4. 患者基本信息表

患者基本情况	低级别尿路	高级别尿路	乳头状瘤或低
	上皮癌	上皮癌	度恶性潜能的乳头状瘤
患者数量 (n)	9	14	17
年龄 n			
≥60	8	9	12
<60	1	5	5
性别 n			
男	7	12	13
女	2	2	4
吸烟史 n			
吸烟	6	8	10
不吸烟	3	6	7
肌层浸润 n			
浸润	3	11	0
非浸润	6	3	7
膀胱癌家族史 n			
有	2	2	2
无	7	12	15

表 5. 膀胱镜图像分组信息表

图像分组	低级别尿路	高级别尿路	乳头状瘤或低度恶性
	上皮癌	上皮癌	潜能的乳头状瘤
图片数量 (n)	10	18	21
年龄 n			
≥60	9	11	16
<60	1	7	5
性别 n			
男	8	15	16
女	2	3	5
吸烟史 n			
吸烟	7	10	13
不吸烟	3	8	8
肌层浸润 n			
浸润	3	14	0
非浸润	7	4	21
膀胱癌家族史 n			
有	3	2	2
无	7	16	19

2.6 评价指标

本文采用平均像素精度 (Mean Pixel Accuracy, MPA)、平均交并比 (Mean Intersection over Union, MIoU) 作为实验评价指标。MIoU 计算每个类别的真实值和预测值及两个集合交集与并集的比值, 再对所有类别的交并比求均值, 如式 (5) 所示。MIoU 的值越大, 代表分割精度越高。

$$MIoU = \frac{1}{n+1} \sum_{i=0}^n \frac{P_{ii}}{\sum_{j=0}^n P_{ij} + \sum_{j=0}^n P_{ji} - P_{ii}}$$

平均像素精度 (MPA) 分别对每个标注类别计算像素准确率, 然后再对所有类别的像素准确率求均值, 如式 (6) 所示。MPA 的值越大, 图像分割性能越好。

$$MPA = \frac{1}{n+1} \sum_{i=0}^n \frac{P_{ii}}{\sum_{j=0}^n P_{ij}}$$

其中， n 表示标签标记的类别， $n+1$ 表示包含背景的总类别， p_{ii} 表示实际为 i 类预测为 i 类的像素数量， p_{ij} 表示实际为 i 类但预测为 j 类的像素数量， p_{ji} 表示实际为 j 类但预测为 i 类的像素数量。模型参数量（Parameters）计算了模型总的可训练参数量，用来评价模型的空间复杂度。FPS 表示每秒处理图片的数目，用来衡量网络的分割速度。

2.7 损失函数

针对不同的训练模型需要选择合适的损失函数，这样有助于优化网络中的参数，提高训练过程中的收敛速度。在图像语义分割任务中，常用的损失函数为交叉熵损失（CE loss），其公式如式（6）所示：

$$Loss_{CE} = -\frac{1}{N} \sum_i [y_i \ln \hat{y}_i + (1 - y_i) \ln(1 - \hat{y}_i)]$$

其中， N 表示样本总数， i 表示样本， y_i 表示像素 i 的实际标签， \hat{y}_i 表示像素 i 的预测标签。但交叉熵损失函数在对每个像素进行类别预测时，对每个像素的损失赋予了相同的权重，因此，当正负样本数差距过大时，损失函数的效果会变差。使用 Dice loss 可以避免这一问题，Dice loss 是一种衡量交并比指标的损失函数，其公式如式（7）所示：

$$Loss_{dice} = 1 - \frac{2|X \cap Y|}{|X| + |Y|}$$

其中： X 表示目标类预测集合， Y 表示真实标签类集合。

Dice loss 本质上度量了预测值与真实值的重叠程度，侧重于标签与预测的一致性，这对于正负不平衡样本表现更好。然而，Dice loss 对噪声很敏感，可能会忽略边界信息，导致边界分割不佳。因此，本文将 CE loss 与 Dice loss 结合作为损失函数，不仅解决样本分布不平衡的问题，也提高了肿瘤边界提取效果。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/558131037014006034>