

文章编号: 2095-2163(2022)07-0025-10

中图分类号: TP391.41

文献标志码: A

基于多特征融合的中医症候舌象分类方法研究

徐雍钦¹, 杨晶东¹, 姜泉², 韩曼², 宋梦歌²

(1 上海理工大学 光电信息与计算机工程学院 自主机器人实验室, 上海 200093;

2 中国中医科学院广安门医院 风湿病科, 北京 100053)

摘要: 中医舌象与脏器病理特征具有一定相关性,通常用来分析人体脏器病理变化。采用深度学习方法可以提取舌象深层特征,反映脏器病理变化,但现有的深度学习模型结构单一,无法有效提取舌象深层的局部和全局特征,也未结合临床舌象浅层特征,如边缘特征、纹理特征等进行综合分析,降低了模型分类精度和泛化性能。本文提出了一种基于多特征融合的迁移学习分类网络,采用基于自适应注意力机制迁移学习框架,提取舌象深层特征并融合舌象边缘特征、纹理特征,提高舌象分类精度和泛化性能。实验结果表明,本文方法针对6分类中医证候舌象样本,模型分类精度为 0.953 ± 0.031 ,灵敏度 0.952 ± 0.032 , $F1$ 值 0.952 ± 0.032 ,与典型分类模型相比,具有较高的分类精度和泛化性能。

关键词: 舌象分类;深度学习;注意力机制;特征融合

Research on classification of tongue image of traditional Chinese medicine syndrome based on multiple feature fusion

XU Yongqin¹, YANG Jingdong¹, JIANG Quan², HAN Man², SONG Mengge²

(1 Autonomous Robotics Laboratory, School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China; 2 Division of Rheumatology, Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing 100053, China)

[Abstract] Tongue image of traditional Chinese medicine has relationship with viscera pathological features, and are usually used to judge the pathological variants of the human organs in the clinical diagnosis of TCM. Deep learning methods can extract in-depth features of tongue image to reflect the pathological variants of the human organs. However, current deep learning model has simple and single functions, and is unable to extract in-depth local and global features of tongue image and to combine the superficial features such as edge, texture, resulting reduction of classification accuracy and generalization. In this article we proposes a novel transfer learning network based on multi-feature fusion, which can effectively extract in-depth features of tongue image using a transfer learning network with adaptive attention mechanism combined with superficial features such as edge, texture feature etc. The experimental results show that the proposed model can achieve precision of 0.953 ± 0.031 , sensitivity of 0.952 ± 0.032 , and $F1$ -score of 0.952 ± 0.032 on the 6-class tongue images of TCM syndrome, which is significantly higher than the state-of-the-art networks in classification accuracy and generalization performance.

[Key words] tongue images; deep learning; attention mechanism; feature fusion

0 引言

近年来,深度学习已被广泛运用到医学图像分类领域。张晓彤^[1]等将 ResNet50 应用到更年期泌尿生殖系统综合症的分类,高效地实现病灶识别;增思涛^[2]等则将 ResNet 应用到肺灌注图像分类,并有效地检测了肺栓塞病变;刘张^[3]等人将 EfficientNet 用于消化道常见病变,平均 AUC 值达 0.94;赵晓平^[4]等,将 DenseNet 网络应用于乳腺癌病理的八分类研

究,在亚型八分类表现上准确率达 95.5%,且在数据不均衡的条件下取得了较好的分类效果。

近年来,许多学者已将深度学习应用于舌象分类。汤一平^[5]等在舌象分类中引入多任务结构,使用多个二分类网络来消除多标签的歧义性;孙萌^[6]等基于 Inception-ResNet-V1 网络改进了损失函数,对提取特征形成了有效的映射空间,并取得了较好的分类效果;李渊彤^[7]等采用 VGG16 网络对舌象进行多分类,并使用迁移学习降低网络训练成本;宋超^[8]等采

基金项目: 国家自然科学基金(81973749);中国中医科学院科技创新工程项目(CI2021A01503)。

作者简介: 徐雍钦(2000-),男,本科生,主要研究方向:人工智能、机器学习与大数据分析等;杨晶东(1973-),男,博士,副教授,主要研究方向:人工智能、机器学习与大数据分析、机器视觉等;姜泉(1961-),女,博士,教授,主任医师,主要研究方向:风湿免疫病的中医、中西医结合临床及基础研究;韩曼(1984-),女,博士,副主任医师,主要研究方向:风湿免疫病的中医、中西医结合临床及基础研究;宋梦歌(1993-),女,博士研究生,主要研究方向:风湿免疫疾病的临床与基础研究。

收稿日期: 2022-02-22

用多分支网络结构进行舌象分类,并且在每个分支网络中采用不同策略的迁移学习参数,提高网络分类性能;翟鹏博^[9]等将注意力机制引入舌象分类,降低样本噪声,并针对每种特征建立单独分支网络,构建一种简单网络结构;邱童^[10]将浅层特征和深层语义特征融合后,采用传统机器学习方法进行舌象分类,但传统机器学习分类器无法充分学习融合特征,无法充分利用特征间的差异性提高分类精度。

本文针对临床中医证候六分类舌象样本,构建一种自适应注意力机制的迁移学习舌象分类网络,提取舌象深层特征,并融合舌象边缘特征、纹理特征,实现多特征融合舌象分类。

1 基于多特征融合的舌象自主分类模型

1.1 基于自适应注意力机制的迁移学习网络 FusionNet

注意力机制被广泛应用于深度学习网络,提高网络特征提取能力,Hu^[11]等将通道注意力机制 SE 引入计算机视觉领域,提出 SE (Squeeze-and-Excitation) 网络,该结构基于通道维度构建特征相关性;Woo^[12]等提出 CBAM (Convolutional Block Attention Module) 注意力机制,在通道注意力机制基础上引入了空间注意力机制,通过通道和空间两个维度提取特征,保证了图像通道域和空间域上的信息关联;Li^[13]等参考 Inception 结构提出 SK (Selective Kernel) 网络,通过多分支并行结构,学习图像不同尺度下的多层次特征;Yang^[14]等受脑科学中神经元放电模式启发,提出 SimAM (Simple, Parameter-Free Attention Module) 注意力机制,能够通过一种轻量无参方法直接得出三维权重。

传统神经网络为了提升分类精度,会选择对网络深度或宽度进行一定比例的缩放,即通过增加网络深度、特征层数来获得更好效果。而迁移学习

EfficientNet 网络通过对网络深度、宽度以及输入图像分辨率的统一缩放,在等量资源下取得更好的分类性能^[15]。EfficientNetV2 则在 EfficientNet 的基础上引入 Fused-MBConv 结构,逐步增加特征图尺寸,以提高模型分类性能^[16]。

本文在 EfficientNetV2 的基础上引入 SimAm 注意力机制,提出一种自适应迁移学习网络 FusionNet,采用 SimAm-MBConv 结构替换原网络中的 MBConv 结构,改进后的网络结构如图 1 所示。SimAm-MBConv 注意力机制是一种 SE+SimAm 串联结构,原始注意力机制 SE 可以提取图像通道域上的特征,但难以提取空间域特征,即图像通道内平面位置上的特征,而 SimAM 注意力机制的轻量无参特性,虽能提取舌象三维特征,但空间域特征关联性不足。因此,本文串联 SE 和 SimAM 两种注意力机制,在不额外引入参数的情况下将混合注意力机制应用到本文网络,使得 FusionNet 具有更好的分类效果和更高的精度。

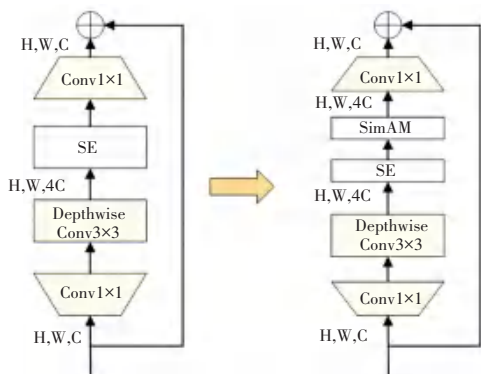


图1 MBConv 模块与改进后的 SimAM-MBConv 模块

Fig. 1 MBConv module and improved SimAM-MBConv module

本文采用 FusionNet 网络提取舌象多维度深层语义特征,并利用迁移学习预训练网络,加快网络收敛速度,增强网络分类精度,自适应的迁移学习 FusionNet 网络结构如图 2 所示。

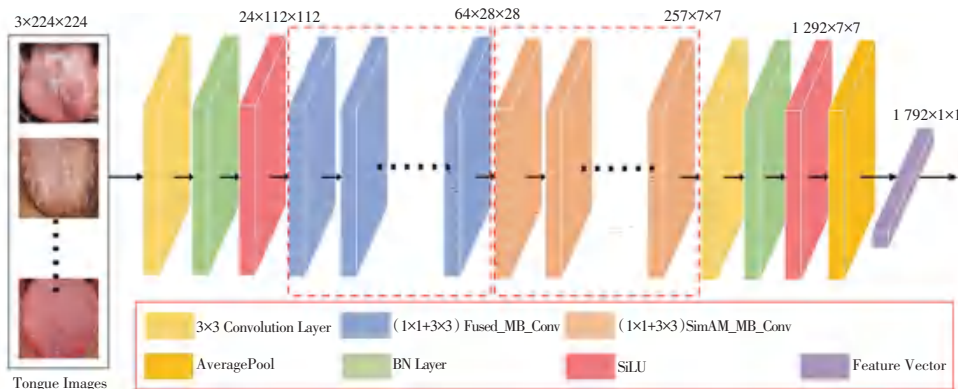


图2 自适应的迁移学习 FusionNet 网络结构

Fig. 2 Adaptive transfer learning network structure of FusionNet

1.2 基于多特征融合的舌象分类网络

中医症候舌象样本具有多样性和浅层特征差异性,传统网络难以感知多维度深层特征,且训练过程中的浅层噪声难以剔除,本文分别采用 Laplacian 算

子、LBP 算子提取舌象边缘特征、纹理特征,并与 FusionNet 网络提取深层特征融合,实现基于多特征融合的舌象自主分类。基于多特征融合的舌象分类网络结构如图 3 所示。

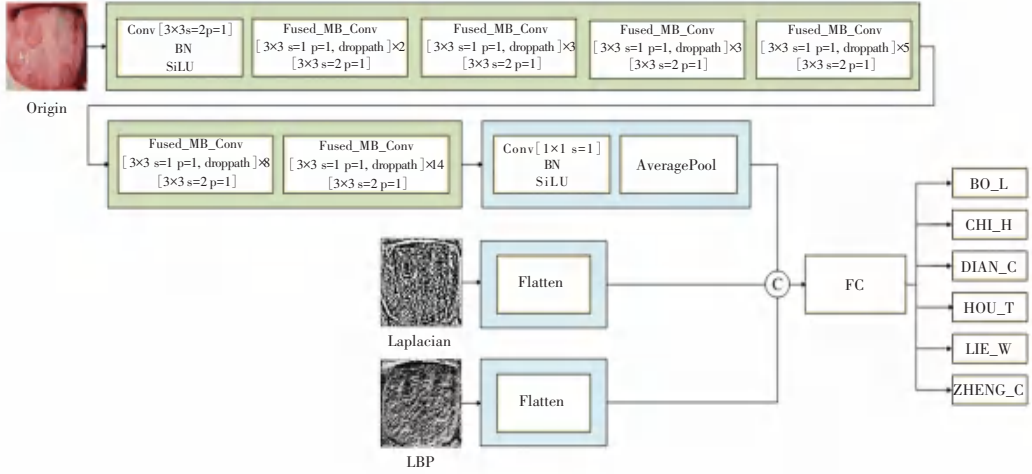


图 3 基于多特征融合的舌象分类网络结构

Fig. 3 Tongue image classification network structure based on multi-feature fusion

本文根据全连接层输出权重,采用 Softmax 函数进行分类,交叉熵损失函数如式(1):

$$Loss = -\frac{1}{N} \sum_i \sum_{j=1}^k y_j \log(p_j) \quad (1)$$

其中, N 为样本数; k 为图像数据类的个数; y_j 为图像分类结果的符号函数,正确分类取 1, 否则取 0; p_j 为对应图像被预测为 j 类的概率。

本文采用 Adam 优化器,总训练 $epoch$ 为 100 (ε),前 50 $epoch$ 学习率为 $1e - 3(\mu)$,后 50 $epoch$ 学习率为 0.000 1, batch-size 为 16(δ_1 / β)。在每个 $epoch$ 中, δ_1 / β 为训练样本数, δ_2 / β 为验证样本数,当前正确率大于 λ 时则更新 λ ,并保存当前网络权重。训练和验证算法流程见表 1。

表 1 模型训练和验证流程

Tab. 1 Model training and validation process

Input	舌象训练集 δ_1 , 测试集 δ_2
	μ → 初始学习率
	ξ → 当前 $epoch$
	ε → 最大 $epoch$
	α → 当前 $epoch$ 中迭代次数
	β → 一个 $epoch$ 中最大迭代次数
	λ → 当前最大正确率
Output	ω → 网络权重
Start	
1.	初始化模型参数: $\mu, \varepsilon, \beta, \lambda$
2.	预处理图像为 $3 \times 224 \times 224$
3.	for $\xi = 1$ to ε do
4.	for $\alpha = 1$ to δ_1 / β do
5.	从 δ_1 获取当前批次的图像
6.	网络前馈并计算训练损失
7.	反向传播并使用 Adam 优化器更新网络权重 ω
8.	end
9.	for $\alpha = 1$ to δ_2 / β do
10.	从 δ_2 获取当前批次的图像
11.	网络前馈并获取预测结果
12.	end
13.	if $accuracy > \lambda$ do
14.	保存网络权重 ω
15.	$\lambda \leftarrow$ 当前 $epoch$ 的正确率
16.	end
17.	end
End	

2 实验结果与分析

2.1 数据预处理和数据增强

本文采用的数据集来自于北京广安门医院中医证候临床舌象样本,该数据集共分为6个类别:正常、裂纹、厚苔、点刺、齿痕、剥落苔,每个类别30张,共180张。由于样本数量较少,本文对该数据集进行了数据增强处理。预处理原始图像像素为 224×224 ,其对应的纹理特征和边缘特征图像像素为 1×448 ,并将数据集划分为训练集和测试集,其比例为4:1。数据增强过程如图4所示。



图4 数据增强

Fig. 4 Data augmentation

将舌象旋转 30° ,进行0值填充处理,使得数据集每类样本数量均扩展12倍,即每类360张图像样本;将原始图像样本水平和垂直翻转,每类得到60张样本,共得到每类420张舌象样本,6类共2520张。本文将 256×256 原始图像通过双线性插值变换为 $224 \times$

224。

2.2 评估方法与超参数

本文使用正确率 *Accuracy*;精确率 *Precision*;灵敏度 *Sensitivity*; F_1 -score, *AUC* 值(ROC 曲线与坐标轴围成的面积)对模型进行评估,其公式如式(2)~(4)。

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (4)$$

$$F_1 - score = \frac{2 \times Precision \times Sensitivity}{Precision + Sensitivity} \quad (5)$$

本文 FusionNet 网络由若干 Fused_MB_Conv 和 SimAM_MBConv 组成,见表2,网络输入为三通道 224×224 尺寸图像,分别重复2、3、3次的 Fused_MB_Conv 结构,输出格式为 $64 \times 28 \times 28$,其中通道加深发生在每次重复前一个未经 DropPath 失活的 Fused_MB_Conv 结构^[16]。同理,SimAM_MB_Conv 结构重复5、8、14次,最终输出格式为 $256 \times 7 \times 7$,分别再经过 1×1 卷积, BN 层, SiLU 激活函数,下采样层后与舌象边缘特征和纹理特征进行通道维度拼接,再经过全连接层和 SiLU 激活后,通过最后全连接层输出分类结果。

表2 FusionNet 模型超参数

Tab. 2 Hyperparameters of FusionNet model

Name	Layers	Size of Filters	Repeat	DropPath	Output shape(C×H×W)	
Features	Conv	3×3, s=2, p=1	-	-	3×224×224	
		BN			24×112×112	
		SiLU				
		Fused_MB_Conv	3×3, s=1, p=1	2	True	24×112×112
		Fused_MB_Conv	3×3, s=2, p=1	-	False	48×56×56
		Fused_MB_Conv	3×3, s=1, p=1	3	True	48×56×56
		Fused_MB_Conv	3×3, s=2, p=1	-	False	64×28×28
		Fused_MB_Conv	3×3, s=1, p=1	3	True	64×28×28
		SimAM_MB_Conv	3×3, s=2, p=1	-	False	128×14×14
		SimAM_MB_Conv	3×3, s=1, p=1	5	True	128×14×14
		SimAM_MB_Conv	3×3, s=1, p=1	-	False	160×14×14
		SimAM_MB_Conv	3×3, s=1, p=1	8	True	160×14×14
Classifiers	Conv	3×3, s=2, p=1	-	False	256×7×7	
		3×3, s=1, p=1	14	True	256×7×7	
		1×1, s=1	-	-	1 792×7×7	
		BN				
		SiLU				
		AveragePool			1 792×1×1	
		Concat			2 688×1×1	
		FC			1 792×1×1	
SiLU						
FC			6×1×1			

2.3 消融实验

为分析各种注意力机制对于舌象分类精度影响,本文对 EfficientNetv2 网络原有注意力机制模块进行改进,分别采用 SE、CBAM、SK、SimAM 4 种典型注意力机制组合形式,共设计了 7 种注意力机制进行消融实验,并比较测试集上的 ROC 曲线和 PR 曲线,如图 5 所示。可以看出无论对于 ROC 曲线,还是 PR 曲线,SE+SimAM 和 CBAM 具有最好的分类性能,SimAM 和 SE+CBAM 次之,SE、SE+SK、SK 的分类效果较差。说明混合注意力机制分类性能优于单一注意力机制,这是由于舌象样本深层特征具有

多样性和空域特性,单一注意力机制难以提取舌象多维度空域特征,降低了模型分类精度。

本文在网络浅层加入 SimAM 模块,并对不同注意力机制组合进行比较,如图 6 所示。可以看出,SimAM+SE+SK、SimAM+SE+SimAM 的组合取得了较好的分类结果,分类性能均优于没有在浅层添加 SimAM 的网络结构。这是由于 SimAM 作为一种轻量无参的注意力机制可以提取舌象浅层多维度的特征,且不会增加运算负担,在浅层添加注意力机制可以更好地消除浅层噪声,减小其在后续卷积中的比重,从而优化分类结果。

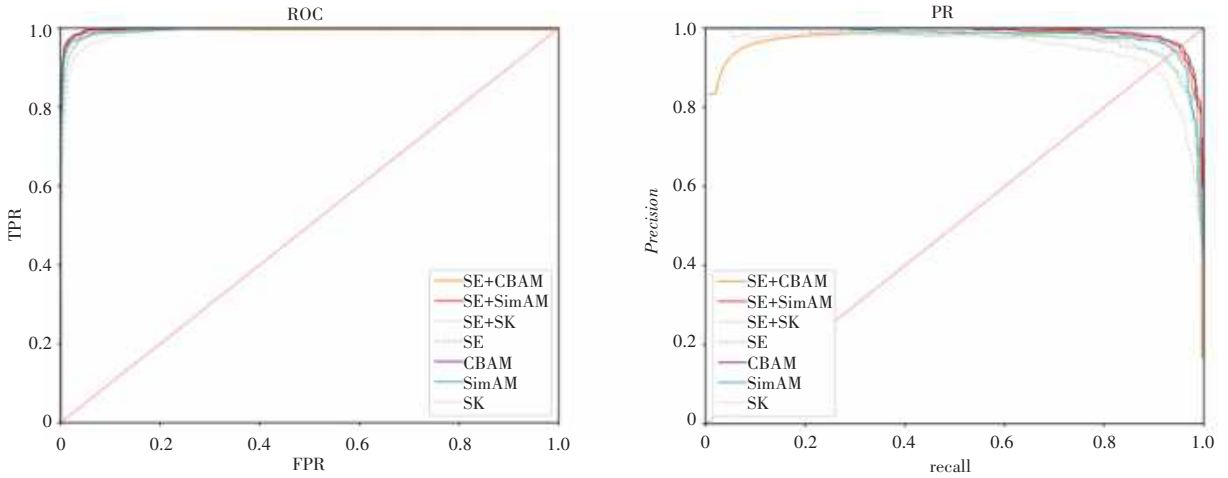


图 5 各种注意力机制的 ROC、PR 曲线

Fig. 5 ROC and PR curve of various model based on different attention mechanism

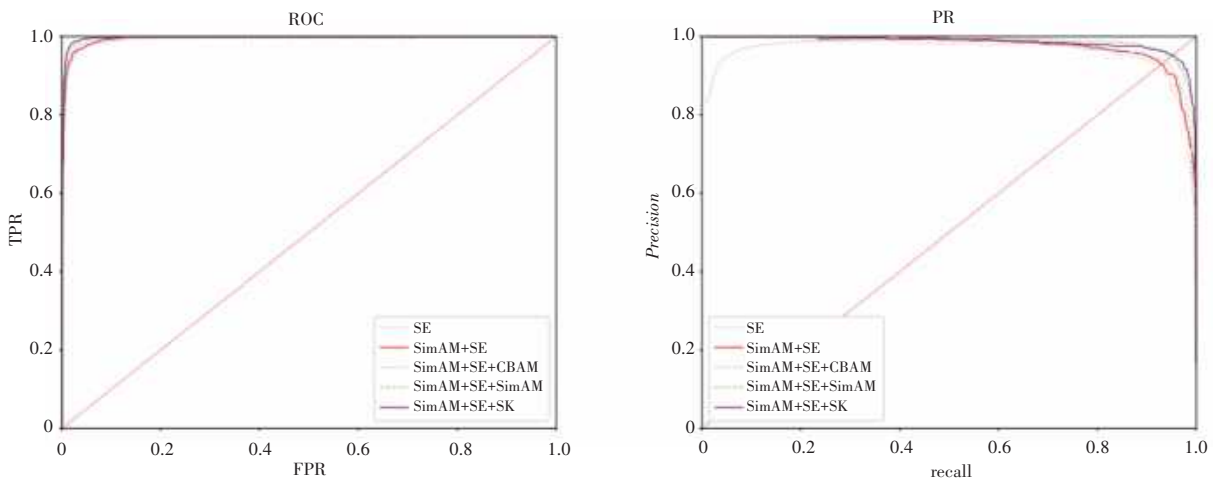


图 6 在网络浅层加入 SimAM,不同注意力机制的比较

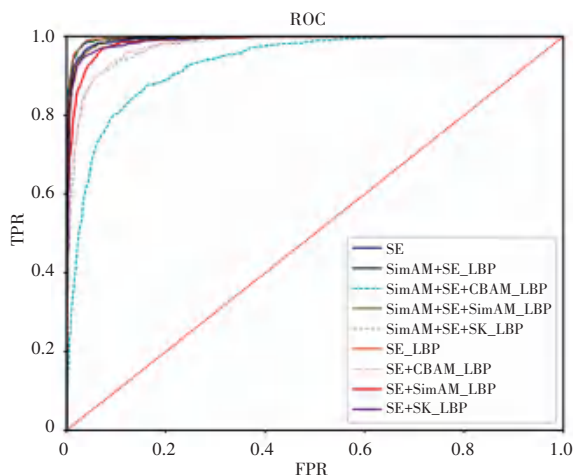
Fig. 6 Comparison of different attention mechanisms after adding SimAM in shallow layers

本文在网络浅层加入 SimAM 模块,并对不同注意力机制组合进行比较,如图 6 所示。可以看出,SimAM+SE+SK、SimAM+SE+SimAM 的组合取得了较好的分类结果,优于没有在浅层添加 SimAM 的网

络结构。这是由于 SimAM 作为一种轻量无参的注意力机制可以提取舌象浅层多维度的特征,且不会增加运算负担,在浅层添加注意力机制可以更好地消除浅层噪声,减小其在后续卷积中的比重,从而优

化分类结果。

上述注意力机制一定程度上提高了舌象深层特征提取能力,但仍然缺乏对浅层特征提取能力,无法有效提取舌面或舌尖等纹理特征、边缘特征。因此,本文分别采用 LBP 算子和拉普拉斯算子来提取舌象浅层的纹理特征与边缘特征,并与 FusionNet 网络提取的深层特征进行多特征融合分类,提高分类性能。



本文设计了多种组合形式,以测试多特征融合对分类性能影响。本文首先以 LBP 算子提取的纹理特征结合 FusionNet 网络提取的深层特征进行融合,并添加各种注意力机制进行消融实验。同样对比各模型测试集 ROC 曲线、PR 曲线,分析舌象分类模型性能。如图 7 所示。可以看出,SE+LBP 和 SimAM+SE+LBP 的组合具有最好的分类性能。

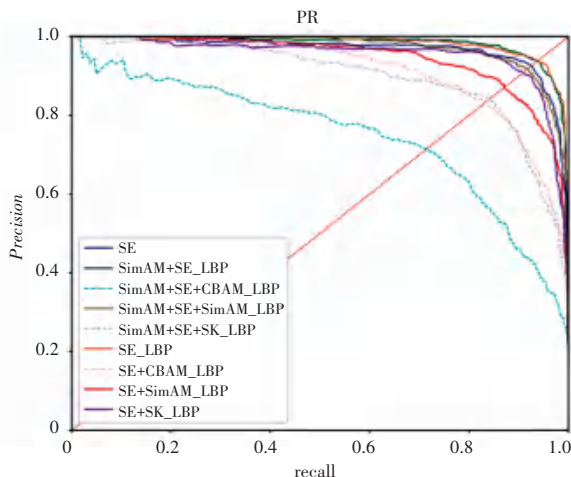


图 7 双特征融合下不同注意力机制的比较

Fig. 7 Comparison of different attention mechanisms with two-feature fusion

这是由于注意力机制 SimAM 模块可以提取舌象深层空域特征,消除部分浅层噪声,与 LBP 算子提取的浅层纹理特征进行融合,获得了较好分类精度。而 SimAM+SE+CBAM+LBP、SimAM+SE+SK+LBP、SE+CBAM+LBP 的组合分类性能较差,这是由于过多注意力机制组合影响了舌象深层特征提取,导致分类精度下降。分析可知,浅层舌象纹理特征有利于提高网络分类精度。

此外,本文在双特征基础上融合 Laplacian 算子

提取的边缘特征,采用三特征融合模式实现舌象分类。各模型测试集 ROC 曲线、PR 曲线分类精度对比曲线如图 8 所示。可以看出,SE+SimAM+LBP+LAPLACIAN 的组合获得了最好的分类性能。这是由于诸如剥裂苔、齿痕类数据的边缘特征较为明显,因此,多特征融合提升了网络分类性能;而 SE+SimAM+CBAM+LBP+LAPLACIAN 等组合由于过多的注意力机制影响了深层特征提取,导致网络整体分类性能降低。

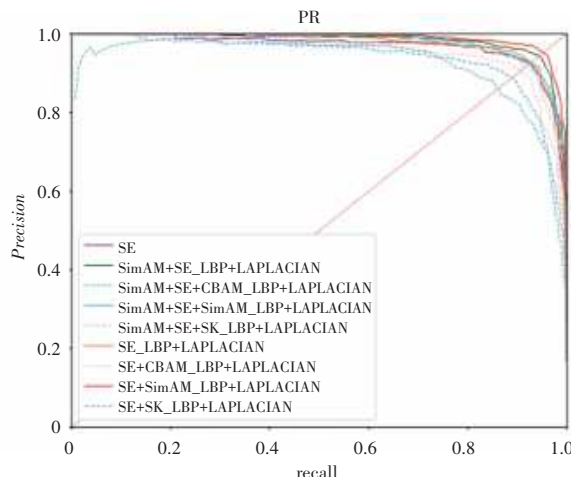
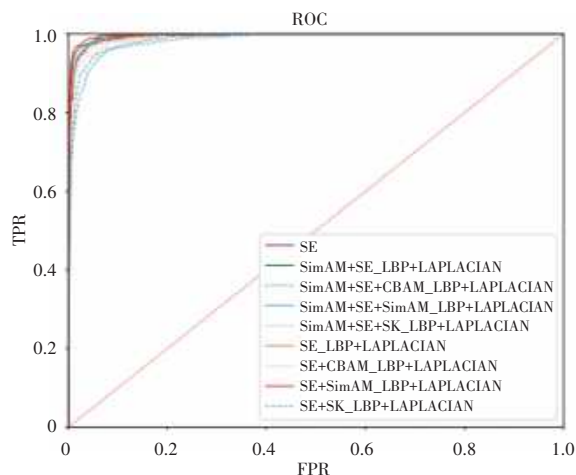


图 8 三特征融合下不同注意力机制的比较

Fig. 8 Comparison of different attention mechanisms with three-feature fusion

综上所述,本文选取 SE+SimAM 的注意力机制,并结合 LBP 算子提取纹理特征和拉普拉斯算子提取的边缘特征,实现舌象自主分类。

2.4 分类性能

2.4.1 交叉验证

将本文方法与现有模型对比,进一步分析本文网络的分类性能。采用 Inception_v4、ConvNeXt_t、ConvNeXt_s 和 RegNetY_200MF 作为对照模型,并进

行 5-2 折交叉验证,每折 506 张舌象样本共 6 类均匀分布,对比 Accuracy、Precision、Sensitivity、F1 - score 和 AUC 值,ROC 曲线和 PR 曲线以及混淆矩阵评估模型分类性能,不同模型在测试集下分类性能,见表 3。与 Inception_v4 网络相比,本文方法准确率增加 0.018 52,精度增加 0.017 15,灵敏度增加 0.018 52, F1 - Score 增加 0.018 76,在 AUC 减少 0.000 69。

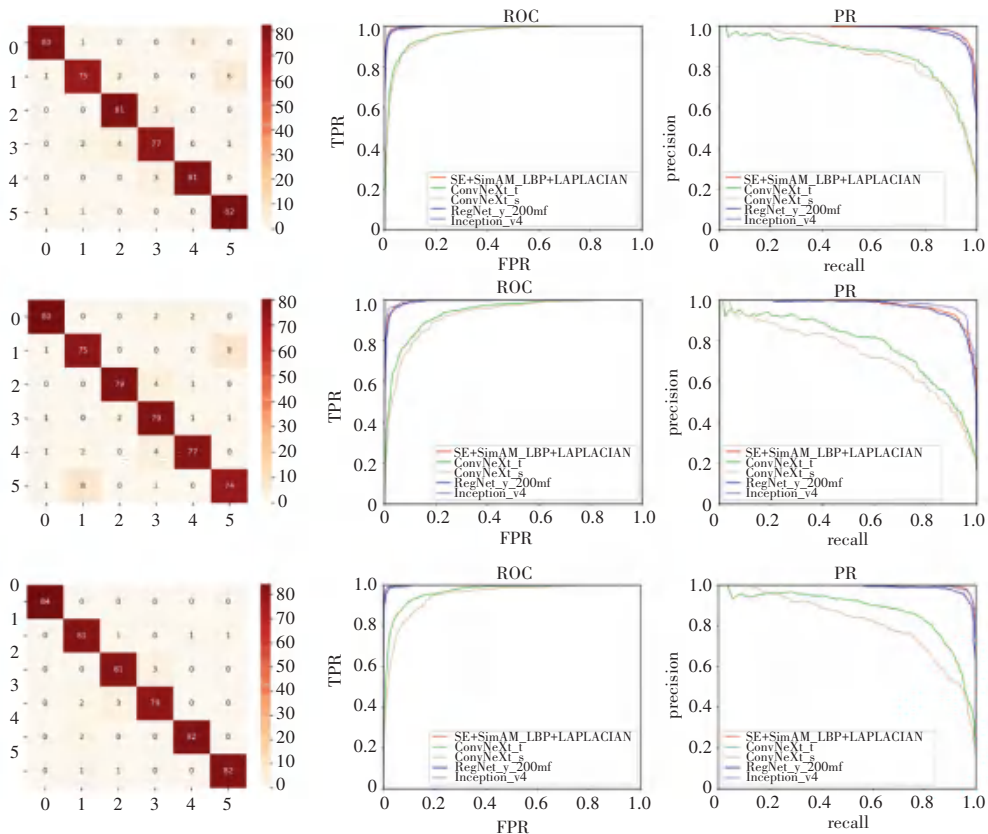
表 3 不同模型的性能对比

Tab. 3 Comparison of classification performance of different models

Model	Accuracy	Precision	Sensitivity	F1 - score	AUC
Ours	0.956 35	0.957 08	0.956 35	0.956 46	0.997 42
Inception_v4	0.937 83	0.939 93	0.937 83	0.937 7	0.998 11
ConvNeXt_t	0.757 94	0.758 62	0.757 94	0.752 89	0.946 63
ConvNeXt_s	0.718 25	0.715 31	0.718 25	0.713 43	0.926 9
RegNetY_200MF	0.908 73	0.909 47	0.908 73	0.908 4	0.990 63

5 折交叉验证实验下 5 种模型的 ROC 曲线和 PR 曲线,以及本文模型分类结果的混淆矩阵如图 9

所示。与 5 种典型模型相比,本文模型具有最好的分类精度和泛化性能。



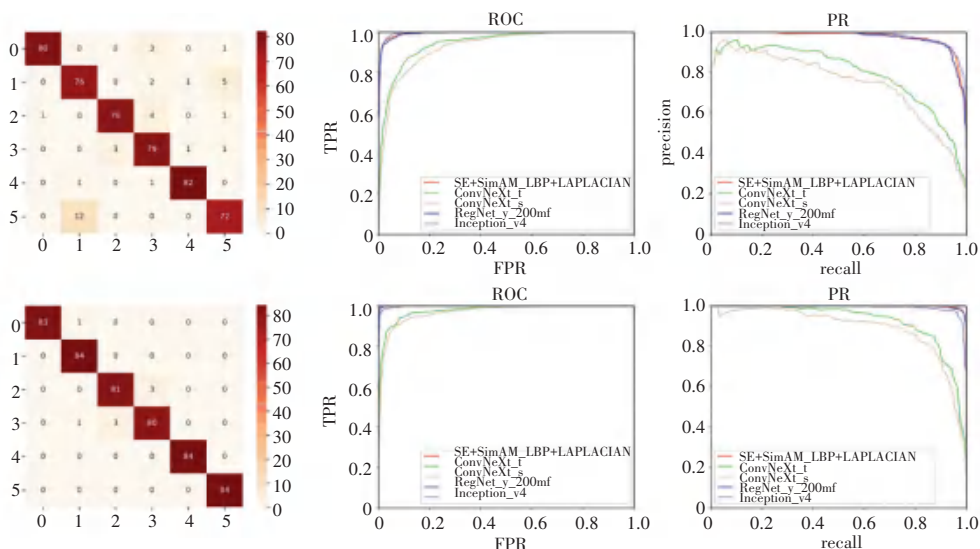


图9 5折交叉验证下不同模型的混淆矩阵、ROC曲线和PR曲线

Fig. 9 Confusion matrix, ROC curve and PR curve under 5-fold cross validation

2.4.2 性能分析

为进一步验证本文模型分类性能,首先选取了与本文模型主干网同系列的其他网络进行对比,共选取了 EfficientNet 系列 b0 ~ b7 共 8 种模型,

EfficientNet v2 系列 medium 和 large 两种模型,并对测试集分类模型 ROC 曲线和 PR 曲线,实验结果如图 10 所示。

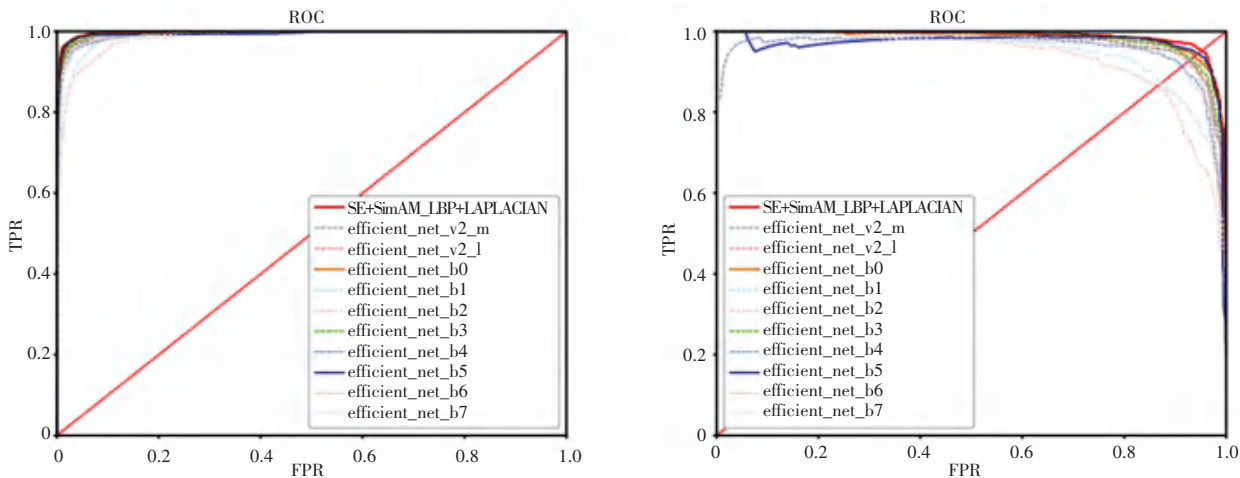


图10 EfficientNet 不同结构网络比较

Fig. 10 Comparison of different EfficientNet

可以看出,本文模型具有最好的分类精度。这是因为 EfficientNet 系列网络结构简单,单 SE 通道注意力机制难以提取舌象深层特征。而 EfficientNet v2 系列两种网络结构 EfficientNet v2 medium 和 EfficientNet v2 large 网络结构过于复杂,无法提取舌象浅层特征,降低模型泛化性能。本文采用混合注意力机制提取舌象多维度的特征,并融合浅层纹理和边缘特征进行综合分类决策,具有更高分类精度和泛化性。

5 折交叉验证各模型评估参数箱线图如图 11

所示,本文方法具有最好分类精度。本文还统计了模型分类精度、计算复杂度和卷积层和全连接层参数量,见表 4。可以看出本文模型虽然具有较高计算复杂度和参数量,但具有最高的分类精度和泛化性能。与 Inception_v4、ConvNeXt_t、ConvNeXt_s、RegNetY_200MF 相比,本文方法的 Precision 增加 0.015、0.162、0.182、0.013,灵敏度增加 0.017、0.172、0.190、0.013, F1 - score 增加 0.019、0.169、0.192、0.013。

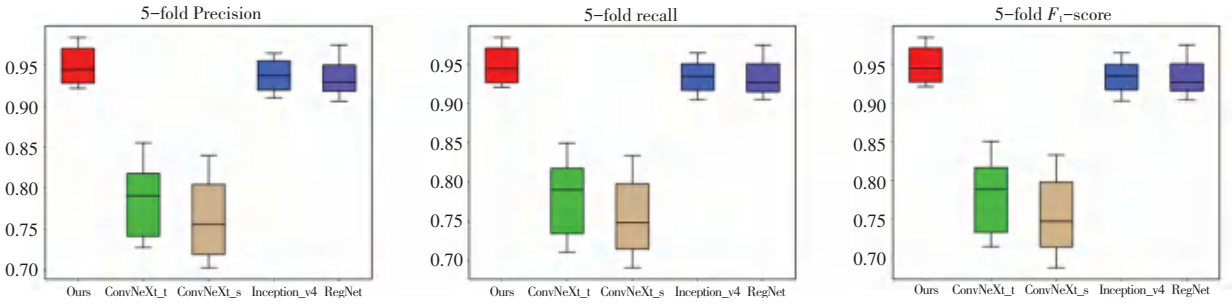


图 11 各模型分类性能对比

Fig. 11 Comparison of classification performance on different models

表 4 5 折交叉验证不同模型的分类结果和计算复杂度

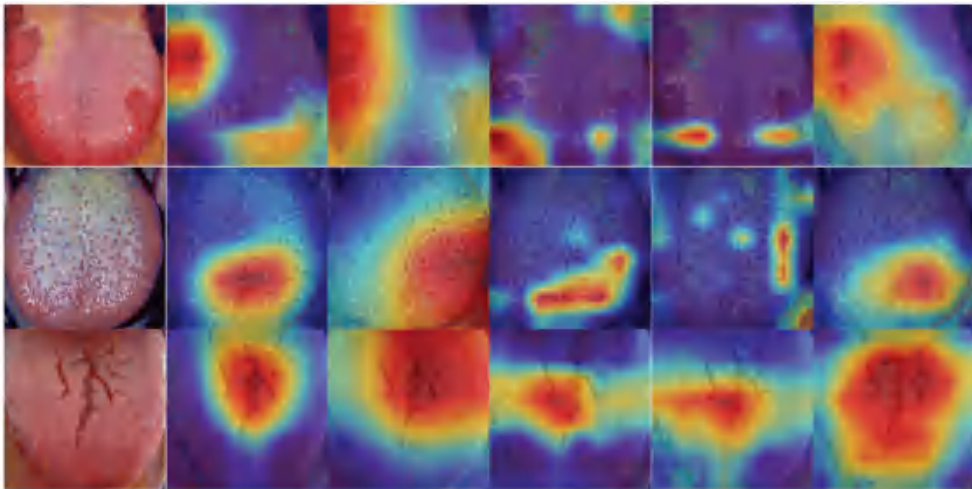
Tab. 4 Classification result and computational complexity of different models under 5-fold cross-validation

Model	Precision	Sensitivity	F1 - score	Time/GFLOPS	Parameters/n
Ours	0.953±0.031	0.952±0.032	0.952±0.032	5.82	43 127 662
Inception_v4	0.938±0.027	0.935±0.030	0.933±0.031	12.28	41 152 038
ConvNeXt_t	0.791±0.064	0.780±0.069	0.783±0.068	0.59	27 824 742
ConvNeXt_s	0.771±0.069	0.762±0.071	0.760±0.072	0.74	49 459 302
RegNetY_200MF	0.940±0.034	0.939±0.035	0.939±0.035	0.42	2 904 112

3 讨论

为可视化各模型对舌象特征提取能力,本文引入了 grad-CAM^[17]方法可视化了网络关注的特征区域。各种模型特征关注度热力图如图 12 所示,图 12(a)分别为剥裂苔、点刺、裂纹舌象原图,图(b)~图(f)分别为本文方法、Inception_v4、ConvNeXt_t、

ConvNeXt_s、RegNetY_200MF 在 grad-CAM 下的特征热力图。对于剥裂苔类舌象,本文模型提取特征均处在舌象边缘明显剥裂的位置;对于点刺类舌象,本文方法提取特征集中于舌象中下部点刺区域;对于裂纹类舌象,本文方法所提取的特征范围能精确覆盖裂纹部分。可见,本文模型特征提取更精确,更适合于中医证候舌象样本分类。



(a) 原始图像 (b) Ours (c) Inception_v4 (d) ConvNeXt_t (e) ConvNeXt_s (f) RegNetY_200MF

图 12 不同模型对舌象样本的特征关注度

Fig. 12 The feature attention of different models to tongue samples

4 结束语

本文提出了一种基于多特征融合和改进注意力机制的中医证候舌象迁移学习分类网络。与当前主流分类网络相比,具有更好的分类精度和泛化性能。由于本文舌象样本属于宏观类医学样本,舌象浅层

特征(纹理、边缘特征)对舌象分类具有较大影响,为此,本文采用自适应注意力机制的迁移学习网络提取舌象深层特征;采用 LBP 算子、Laplacian 算子分别提取纹理特征、边缘特征;最后,将深层特征与浅层特征融合,确定舌象的最终分类。实验结果证明,本文模型可以有效提取中医证候舌象多维度深

层、浅层特征,有利于加速网络收敛。今后将针对更多舌象样本进行研究,改进模型结构和泛化性能。

参考文献

- [1] 张晓彤,孙铭阳,王浩然,等. ResNet 对更年期泌尿生殖系统综合征光学相干断层成像图像的识别[J]. 激光杂志, 2021, 42(4):96-99.
- [2] 增思涛,曹永春,林强,等. 基于 ResNet 深度模型的 SPECT 肺灌注图像分类[J]. 西北民族大学学报(自然科学版), 2021, 42(2):27-35.
- [3] 刘张,郭旭东,张璐璐. 基于 EfficientNet 的胶囊内镜图像多病灶协同检测[J]. 智能计算机与应用, 2021, 11(11):157-161.
- [4] 赵晓平,王荣发,孙中波,等. 改进 DenseNet 的乳腺癌病理图像八分类研究[J/OL]. 计算机工程与应用:1-10[2022-02-18].
- [5] 汤一平,王丽冉,何霞,等. 基于多任务卷积神经网络的舌象分类研究[J]. 计算机科学, 2018, 45(12):255-261.
- [6] 孙萌,张新峰. 基于 TripletLoss 损失函数的舌象分类方法研究[J]. 北京生物医学工程, 2020, 39(2):131-137.
- [7] 李渊彬,罗裕升,朱珍民. 基于深度学习的舌象特征分析[J]. 计算机科学, 2020, 47(11):148-158.
- [8] 宋超,王斌,许家佗. 基于深度迁移学习的舌象特征分类方法研究[J]. 计算机工程与科学, 2021, 43(8):1488-1496.
- [9] 翟鹏博,杨浩,宋婷婷,等. 融合注意力机制的多阶段舌象分类算法[J]. 计算机工程与设计, 2021, 42(6):1606-1613.

- [10] 邱童. 基于深度学习与多特征融合的舌象诊断算法[J]. 现代信息技术, 2020, 4(1):63-65, 68.
- [11] 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.
- [12] 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.
- [13] LI X, WANG W, HU X, et al. Selective kernel networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 510-519.
- [14] YANG L, ZHANG R Y, LI L, et al. Simam: A simple, parameter-free attention module for convolutional neural networks [C]//International Conference on Machine Learning. PMLR, 2021: 11863-11874.
- [15] TAN M, LE Q. Efficientnet: Rethinking model scaling for convolutional neural networks [C]//International conference on machine learning. PMLR, 2019: 6105-6114.
- [16] TAN M, LE Q. Efficientnetv2: Smaller models and faster training [C]//International Conference on Machine Learning. PMLR, 2021: 10096-10106.
- [17] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization [C]//Proceedings of the IEEE international conference on computer vision. 2017: 618-626.

(上接第24页)

4 结束语

针对两幅同一地区、不同时间的 SAR 图像变化检测问题,本文提出了一种基于邻域信息优化以及差异图融合的方法,该方法根据像素点邻域的异质性 θ , 自适应地提取邻域信息,极大程度降低了噪声对变化检测结果的影响,同时生成的差值差异图像可以很好地找到弱变化区域,将两幅差异图结合,便可以实现抑制噪声和保留细节之间的平衡。实验结果证明了本文提出的方法在 SAR 图像变化检测上的有效性,通过与其他变化检测方法对比,本文算法在 PCC 和 Kappa 系数上均优于其他算法,能够更加准确地获取变化信息,有效提高了 SAR 图像变化检测精度。下一步将针对 SAR 图像边界区域思考,构造更加适合边界部分的邻域窗口,以减小图像局部区域信息损失,进一步提高变化检测精度。

参考文献

- [1] 张晓红,李茂堂,王桂琴. 航迹显示在实时 SAR 图像传输中的实现与应用[J]. 计算机应用与软件, 2007(1):112-114.
- [2] GONG M, YU C, WU Q. A Neighborhood-Based Ratio Approach for Change Detection in SAR Images [J]. IEEE Geoscience & Remote Sensing Letters, 2012, 9(2):307-311.

- [3] TSAI D M, LIN C C. Fuzzy C-means based clustering for linearly and nonlinearly separable data[J]. Pattern Recognition, 2011, 44(8):1750-1760.
- [4] 刘若辰. 基于邻域比值和自步学习的 SAR 图像变化区域检测方法. 201710739761.2[P]. 2019-11-01.
- [5] MA J, GONG M, ZHOU I. Wavelet fusion on radio images for change detection in SAR images[J]. IEEE Geoscience and Remote Sensing letters, 2012, 9(6):1122-1126.
- [6] Majidi Milad, Ahmadi Salman, Shah Hosseini Reza. A saliency-guided neighbourhood ratio model for automatic change detection of SAR images [J]. International Journal of Remote Sensing, 2020, 41(24):9606-9627.
- [7] Yin Deshuai, Hou Rui, Du Junchao, et al. SAR image change detection method based on intuitionistic fuzzy C-means clustering algorithm[J]. Journal of Intelligent & Fuzzy Systems, 2020, 38(4):3595-3604.
- [8] LU J, MING L, PENG Z, et al. SAR Image Change Detection Based on Multiple Kernel K-Means Clustering With Local-Neighborhood Information [J]. IEEE Geoscience & Remote Sensing Letters, 2017, 13(6):856-860.
- [9] ZHENG Y, ZHANG X, HOU B, et al. Using Combined Difference Image and k-Means Clustering for SAR Image Change Detection[J]. IEEE Geoscience & Remote Sensing Letters, 2013, 11(3):691-695.
- [10] ZHU W P. Color image segmentation by using clustering of spatial patterns [J]. Computer Engineering and Applications, 2009, 45(34):161-163.
- [11] 公茂果,苏临之,李豪,等. 合成孔径雷达影像变化检测研究进展[J]. 计算机研究与发展, 2016, 53(1):123.