

基于小波池化 Gabor 卷积神经网络的皮肤癌分类

刘瑞欣 刘海华^{通讯作者}

中南民族大学 生物医学工程学院, 武汉, 430074;

摘要: 皮肤癌是目前人类最频发的癌症之一, 严重影响人类的生命安全, 为了降低皮肤癌的影响, 对皮肤病的早期精确诊断和分类具有非常重要的意义。迄今为止, 深度学习已被广泛应用于疾病的临床辅助诊断中, 如卷积神经网络 (CNN) 在基于医学图像分类的疾病辅助诊断中表现出优越的性能。针对皮肤疾病图像的特点, 本文采用 Gabor 滤波器对图像中纹理信息, 边缘信息的敏感性, 以及离散小波变换对噪声的鲁棒性, 构建由 Gabor 卷积核和小波下采样层组成的 Gabor 卷积网络模型 (WAPL-Gabor CNN)。该网络利用 Gabor 的实部和虚部建立卷积层, 而利用注意力小波下采样构建下采样层。为了检验 Gabor 卷积网络的性能, 采用 HAM10000 和 Small-ISIC2018 等公开数据集进行实验, 其结果表明, 与其它方法相比该网络模型非常有效, 其分类准确率分别达到 99.76% 和 91.6%。

关键词: 皮肤癌; Gabor 滤波器; 注意力机制; 离散小波变换

Skin cancer classification based on wavelet pooling Gabor convolutional neural network

Liu Ruixin Liu Haihuacorresponding author

Institute of Biomedical Engineering, South-Central University for Nationalities, Wuhan, 430074;

Abstract: Skin cancer ranks among the most prevalent malignancies worldwide, posing significant threats to human health. Early and accurate diagnosis and classification of skin lesions are crucial for mitigating its impact. To date, deep learning has been extensively applied in clinical decision support systems, with convolutional neural networks (CNNs) demonstrating superior performance in medical image-based disease diagnosis. To address the unique characteristics of dermatological images, this study leverages the sensitivity of Gabor filters to texture and edge information, combined with the noise robustness of discrete wavelet transform (DWT), to construct an Attention-Wavelet Gabor CNN (WAPL-Gabor CNN). The proposed architecture integrates Gabor-based convolutional layers (utilizing both real and imaginary components of Gabor filters) with attention-guided wavelet downsampling layers. Experimental validation on publicly available datasets (HAM10000 and Small-ISIC2018) demonstrates the model's superior performance, achieving classification accuracies of 99.76% and 91.6%, respectively, outperforming existing approaches. These results underscore the effectiveness of the proposed framework in dermatological image analysis.

Keywords: skin cancer; Gabor filters; attention mechanism; discrete wavelet transform

DOI:10. 69979/3029-2808. 25. 06. 054

引言

人类活动工业化产生的有害化学物质可直接接触皮肤, 或间接导致到达地面的紫外线增多令皮肤细胞的 DNA 受损, 致使人类患皮肤癌几率增加。值得引起注意的是, 一些良性皮肤病变在后期可能转变为可危及生命的恶性皮肤癌, 因而在早期正确分类皮肤病变对于降低皮肤癌致死率意义重大。在涉及微妙的恶性特征或非典型表现的情况下, 病变之间的相似性导致医生的主观评价不一, 难以推广。而深度学习强大的特征学习能力和优越的计算效率可以显著提高诊断准确率, 缓解医生临

床诊断压力。

皮肤病变的诊断主要有两大难点。首先, 皮肤病变存在很大的类内差异性和类间相似性, 同一种类型的皮肤病变在质地、颜色、形状等方面会存在较大的差异, 而不同的皮肤癌也可能具有很高的视觉相似性。其次, 边界特征是确定皮肤癌种类的重要依据, 但皮肤病变区域和周围正常皮肤之间的界限往往不明显、不清晰。随着对深度学习研究的深入与实践应用的拓展, 传统卷积神经网络 (CNN) 也逐渐暴露出一些有待优化的问题。针对皮肤癌分类任务, 普通卷积的卷积核是固定的, 在

处理具有复杂方向和尺度变化的纹理信息时,缺乏对方向和尺度的自适应调整能力。而 Gabor 滤波器的方向尺度选择特性在纹理特征识别上更有优势,且参数量少。Deng 等人^[1]采用 Gabor 滤波器的实部高效地提取局部纹理特征,在 ISIC 2017 和 ISIC 2018 数据集上的分类结果优于当时的所有方法。此外传统的池化作为下采样会导致图像信息的丢失,使得边缘分界更加模糊,如果噪声与有用特征在空间上分布较为接近,池化在减少数据量的同时可能会把有用特征和噪声一起进行处理,使得噪声在某种程度上被放大。DWT 凭借其特有的多分辨率分析特性,能够在保留重要细节的前提下进行下采样,有效降低混叠现象出现的概率。Wu 等人^[2]采用离散小波变换作为下采样,并加入多通道注意机制,将小波变换后的高频分量与低频分量有效结合。在增强 HAM10000 数据集上准确率显著提高,达到 95.84%,并且加快了网络分类速度,增强了抗噪性。

为攻克传统卷积神经网络在皮肤癌病变分类中面临的难题,提升计算机辅助诊断系统性能,本文凭借 Gabor 滤波器特性重新设计卷积,并利用离散小波变换多分量分解优势设计小波池化层替代传统下采样层集成于同一卷积神经网络中,提出了小波池化 Gabor 卷积神经网络(WAPL-GaborCNN),并与现有的 ResNet101、Eff-CTM 等模型进行比较,WAPL-GaborCNN 在准确性方面表现优异。

1 小波池化 Gabor 卷积神经网络

1.1 网络模型的基本结构

小波池化 Gabor 卷积神经网络(WAPL-Gabor CNN)是将 VGG16 网络结构中第一个普通卷积块替换为 Gabor Block,用小波池化层(WAPL)取代网络中最大池化层构建而成,包括四个卷积块,共十个卷积层、四个下采样层和三个全连接层,其结构如图 1 所示。

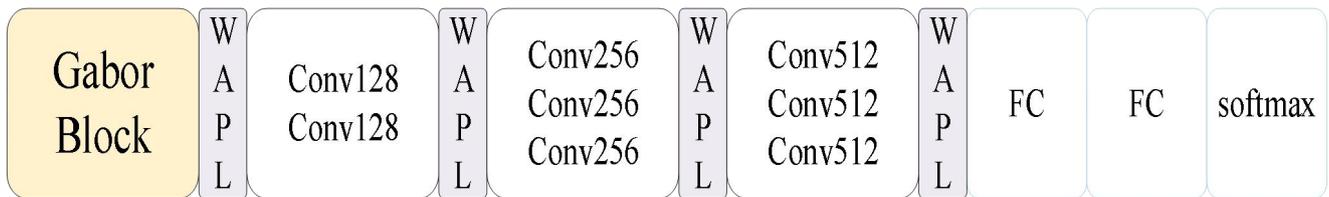


图 1 WAPL-Gabor CNN

1.2 Gabor 卷积层

Gabor 滤波器通过调整尺度和方向参数捕捉不同大小和方向的图像特征,适用于分析多种不同类型的纹理。Gabor 函数是由表示正交方向的实部和虚部组成的复数形式,事实上大多数与 Gabor 相关的 CNN 只利用了 Gabor 滤波器的实部,这意味着它们只收集数据中的低频信

息,而忽略了可能有用的高频信息。Jiang 和 Su^[3]在 CNN 中使用了实部和虚部的直接连接相连的方式。然而,这种方法在提取 Gabor 特征时实部和虚部两个部分之间的关系无法调整。Liu 等人^[4]设计了由相位诱导的实部虚部相关联的 Gabor 滤波器对高光谱图像进行分类,使得神经网络在分类任务上性能有显著提升,以下为相位诱导 Gabor 滤波器的推导过程:

$$g(x, y; \lambda, \phi, \psi, \sigma, \gamma) = \exp\left(-\frac{\bar{x}^2 + \gamma^2 \bar{y}^2}{2\sigma^2}\right) \exp\left(i\left(2\pi\frac{\bar{x}}{\lambda} + \psi\right)\right) \quad (1)$$

其中 $\bar{x} = x \cos \phi + y \sin \phi$ 和 $\bar{y} = y \cos \phi - x \sin \phi$ 是 x 和 y 在给定角 ϕ 下旋转坐标, ϕ 表示滤波器并行条纹的方向,令 $\phi = 0$, $\bar{x} = x$, $\bar{y} = y$; σ 为高斯函数窗宽度,为确保高斯包络部分在整个二维平面上的积分值为 1,

$$G(X, Y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \exp\left(i(x\omega_x + y\omega_y)\right) \quad (2)$$

其中 $\frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$ 是高斯包络,令其为 K 。 $\omega_x = |\omega| \cos \theta$, $\omega_y = |\omega| \sin \theta$ 是给定角频率 ω 在 x 、 y 方向

对高斯函数使用归一化常数 $1/2\pi\sigma_x\sigma_y$,令 $\sigma_x = \sigma_y = \sigma$; γ 表示空间纵横比,令 $\gamma = 1$,使 Gabor 函数的形状为圆形。

上的投影, θ 是 ω 和 x 方向间的夹角。根据欧拉关系,令 $M = x\omega_x + y\omega_y$,并引入表示正弦谐波的相位偏移 P ,

可以将二维 Gabor 滤波器改写为以下复形式：

$$\begin{aligned} G(X, Y) &= K \times \exp(i(M + P)) \\ &= K \cos(M + P) + iK \sin(M + P) \\ &= \Re\{G_p(X, Y)\} + i\Im\{G_p(X, Y)\} \end{aligned} \quad (3)$$

由于正弦谐波和余弦谐波的相位特性：

$$K \sin(M + P) = K \cos\left(M + \left(P - \frac{\pi}{2}\right)\right) \quad (4)$$

可以看出，具有特定相位偏移 P 的虚部正好是 $P - \pi/2$ 的相应的真实的虚部：

$$\Im\{G_p(X, Y)\} = \Re\left\{G_{\left(P - \frac{\pi}{2}\right)}(X, Y)\right\} \quad (5)$$

不同尺寸的 Gabor 卷积核在感受野上具有显著差异：小尺寸卷积核能精细捕捉局部纹理和边缘细节；大尺寸卷积核则整合宏观结构特征。多尺度融合机制弥补了单一尺度的局限性，使得微观到宏观的各类特征都能被充分挖掘，本文设计了融合不同尺寸 Gabor 卷积特征的 Gabor Block。如图 2 所示，Gabor Block 由一个 Gabor 卷积层和一个普通卷积层组成。首先将 3×3 、 5×5 和 7×7 Gabor 卷积核提取的特征进行通道拼接得到多通道的融合特征，为保证不同尺寸 Gabor 核输出特征图大小保持一致，为其分别做适当填充。再通过 $\text{stride}=2$ 的 1×1 卷积进行降维，该操作不仅减少计算量，减轻后续网络层负担，还通过特征重组优化信息表达提升特征的代表性，提高网络效率和性能。

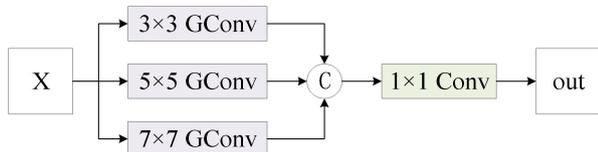


图 2 Gabor Block

1.3 小波池化层

小波基函数是离散小波变换的核心，它由分别对应低频滤波器和高频滤波器的尺度函数 $\phi(x)$ 和小波函数 $\psi(x)$ 组成，与信号卷积可得到反映信号整体特征的近似分量和反映细节特征的高频分量。二维离散小波变换可以看作对输入图像行滤波，对得到的行低频分量 L 和行高频分量 H 上再进行列滤波，两次滤波后得到低频分量 LL 、水平高频分量 LH 、垂直高频分量 HL 和对角高频分量 HH 。令输入为 $X(x, y)$ ， x 、 y 分别为输入 X 的水平分量和垂直分量。其表达如公式 (6) 和 (7) 所示。

$$\{LL, LH, HL, HH\} = 2DDWT(X) \quad (6)$$

$$\begin{cases} LL = \varphi(X) = \varphi(x)\varphi(y) \\ LH = \psi^{LH}(X) = \varphi(x)\psi(y) \\ HL = \psi^{HL}(X) = \psi(x)\varphi(y) \\ HH = \psi^{HH}(X) = \psi(y)\psi(x) \end{cases} \quad (7)$$

WaveCNet 只利用了保留主体结构信息的低频分量 LL ^[5]，丢弃了包含丰富细节信息的所有高频分量，而 F WFC^[6]认为有助于识别小物体的图像域微小差异可以在某些小波带中突出显示。因此，我们选择保留低频分量 LL ，和低频分量 LH 、 HL ，舍弃噪声最多的高频分量 HH ，通过混合注意力机制进一步提取特征，构建注意力小波下采样层。

注意力小波下采样 (WAPL) 结构如图 3 所示。离散小波变换 (DWT) 将输入图像分解成四个子带分量，各分量的空间分辨率均降为原图的 $1/2$ 。对经空间注意力 SA 模块得到 LH_S 和 HL_S 元素相加得到高频特征 $high_W$ ，并与 LL_S 进行通道拼接得到混合特征 Att 。为优化特征表示， Att 进一步由自适应通道权重的通道注意力 CA 模块处理征通道维度压缩。这种设计实现了多分量特征的自适应融合，为后续图像处理任务提供了更具判别性的特征表示，在空间维度实现降采样效果的同时确保了特征表达的紧凑性和有效性。

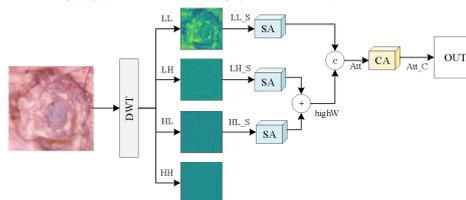


图 3 注意力小波下采样 (WAPL) 示意图

空间注意力模块 SA 通过卷积和激活函数学习特征图上的每个像素点的重要性，第一个卷积以缩放比 R 减少通道数，第二个卷积将通道数恢复为输入图像的大小，然后与输入特征逐元素相乘进行空间调制。通道注意力模块 CA 与 SA 结构相似，CA 先通过全局池化捕获通道的全局信息再通过卷积层和激活函数学习通道间的关系。

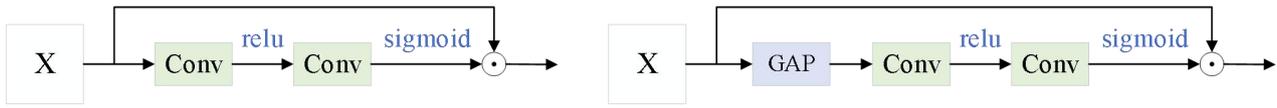


图4 注意力模型框图. (a) 空间注意力 (SA) ; (b) 通道注意力 (CA)

1.4 网络实现

相位诱导 Gabor 滤波器设置 $\{\theta, \omega, \sigma, P\}$ 来调整滤波器的形状和状态。本文将 θ_0 's 预定义为 $[0, \pi)$ 内的均匀间隔序列; ω_0 's 设置为初始值为 $(\pi/2)$ 的几何级数为 $(1/2)$ 的等比序列; σ 设为核尺寸的四分之一; 在 $[0, 2\pi)$ 内随机初始化核相位 P_0 's。小波基函数选择不当可能导致对医学图像执行 DWT 操作时所生成的小波系数中大量有效信息流失, 影响后续对特征的处理。除了 Haar 小波之外具有对称性的正交小波并不存在, 为维持对

称性, 常规选择小波基函数时可选用双正交小波基函数。本文选择双正交性小波 bior2.2 作为实验的小波基函数。

本文使用 Adam 优化器训练网络, 训练次数设置为 100 次, batch size 设置为 64。初始学习率设置为 1×10^{-4} , 并使用余弦退火策略进行调整, 超参数 T_{max} 设置为 10。

2 实验结果与分析

2.1 数据集介绍

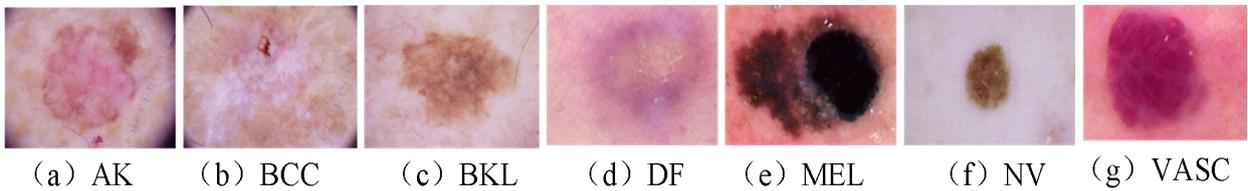


图5 HAM10000 数据库 7 种皮肤疾病示例

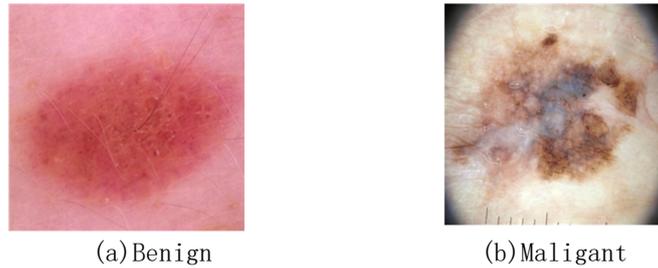


图6 Small-ISIC2018 数据库: (a) Benign; (b) Malignant

(1) HAM10000 (Human Against Machine with 10 000 training images) 数据集^[7]包含 10,015 张图像, 涵盖 7 种病变类型。该数据集数量最多的黑色素细胞痣 (NV) 多达 6705 张, 而血管病变 (VASC) 只有 142 张, 只占总数量的约 1%, 直接训练将严重影响模型评估的准确性。使用数据增强平衡类间数量的同时扩充数据集,

使得增强后数量约为原来的四倍, 类别分布如下表 1。

(2) Small-ISIC2018 数据集是 ISIC 2018^[8] 挑战数据集的一个子集。从原始数据中删除了除黑色素瘤以外的其他疾病的皮肤镜图像, 只保留了两个样本数量更平衡的类别: 良性皮肤痣和恶性皮肤痣。Small-ISIC2018 数据集中病变分布如表 2 所示。

表 1 HAM10000 数据集增强前后数据分布

Dataset split	AK	BCC	BKL	DF	MEL	NV	VASC	Total
Original number	327	514	1099	115	1113	6705	142	10015
Increased number	6213	6168	5495	6555	5565	6705	6532	43233

表 2 Small-ISIC2018 数据集中病变分布

Dataset split	Benign	Malignant	Total
Train	1440	1197	2637
Test	360	300	660

2.2 评价指标

在医学图像分类方向上,单一的评价指标不能全面反映模型的性能。为了对模型的性能进行综合评价,本文采用了多种评价指标,以保证评价结果的准确性和可

靠性。本文选取准确率(Accuracy, Acc)、精密度(Precision)、召回率(Recall)和F1分数四个指标全面准确地评价模型在医学图像分类任务中的表现。这些评价指标的计算方法如下:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$Pre = \frac{TP}{TP + FP} \quad (9)$$

$$Rec = \frac{TP}{TP + FN} \quad (10)$$

$$F1 = \frac{2Pre \times Rec}{Pre + Rec} = \frac{2TP}{2TP + FP + FN} \quad (11)$$

2.3 实验结果

本文实验环境为Centos系统,配置了Intel(R)Xeon(R)Gold 5118 CPU@2.30GH和NVIDIA Tesla V100-SXM2显卡的大型浪潮服务器,软件环境为anaconda 4.5.4, Python 3.6.9, Pytorch 1.1.0。

为了对模型的性能进行综合评价,本文采用了多种评价指标,以保证评价结果的准确性和可靠性。本文选

取准确率(Accuracy, Acc)、精密度(Precision)、召回率(Recall)和F1分数四个指标全面准确地评价模型在医学图像分类任务中的表现。这些评价指标的计算方法如下:表3,表4为不同深度学习模型在HAM1000和Small-ISIC 2018分类任务上表现结果。

表3 HAM10000数据集不同方法的实验结果

Model	Acc(%)	F1(%)	Pre(%)	Rec(%)
FixCaps(2022) ^[9]	96.49			
DCENSnet(2024) ^[10]	99.53	96.9	97.3	96.9
SMobile-Net(2024) ^[11]	99.00	98.85	99.00	98.85
SADASNet(2025) ^[12]	99.31	97.58	97.64	97.57
WAPL_GaborCNN	99.77	99.76	99.76	99.75

表4 Small-ISIC 2018数据集不同方法的实验结果

Model	Acc(%)	F1(%)	Pre(%)	Rec(%)
ResNet101	88.64	87.73	86.17	89.33
VIT-B-16(2021) ^[16]	89.09	88.50	84.97	92.33
Eff-CTM(2023) ^[17]	90.45	89.62	88.60	90.67
Conv-SdMLPMixer (2025) ^[18]	90.15	89.11	89.56	88.67
WAPL-Gabor CNN	91.06	90.21	89.4	90

2.4 结论

本文在VGG结构的基础上提出了WAPL_GaborCNN模型进行皮肤病变分类:利用Gabor滤波器的实部和虚部建立卷积层,在特征提取阶段融合不同尺度的Gabor特征信息;引入小波池化层作为下采样,保留更多信息并提高网络鲁棒性。在HAM1000和Small-ISIC 2018数据集上进行实验,并与其他模型进行对比,结果表明本文提出的WAPL_GaborCNN的皮肤病变分类准确率在几种模型中最高,分别达到99.77%和90.15%,提高了医学图像分类的准确性和可靠性。

[1] Deng X, Yin Q, Guo P. Efficient structural pseudoinverse learning-based hierarchical representation learning for skin lesion classification[J]. Complex & Intelligent Systems, 2022, 8(2): 1445-1457.

[2] Wu Q, Yu Y, Zhang X. A skin cancer classification method based on discrete wavelet down-sampling feature reconstruction[J]. Electronics, 2023, 12(9): 2103.

[3] Ji C, Su J. Gabor binary layer in convolutional neural networks[C]//2018 25th IEEE ICIP. 2018: 3408-3412.

[4] Liu C, Li J, He L, et al. Naive Gabor netw

参考文献

- orks for hyperspectral image classification[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(1): 376-390.
- [5] Li Q, Shen L, Guo S, et al. WaveCNet: Wavelet Integrated CNNs to Suppress Aliasing Effect for Noise-Robust Image Classification[J]. IEEE Transactions on Image Processing, 2021, 30: 7074-7089.
- [6] Li S, Florencio D, Li W, et al. A Fusion Framework for Camouflaged Moving Foreground Detection in the Wavelet Domain[J]. IEEE Trans. Image Process., 2018, 27(8):3918-3930
- [7] Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions[J]. Scientific data, 2018, 5(1): 1-9
- [8] Codella N, Rotemberg V, Tschandl P, et al. Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic)[J]. arXiv preprint arXiv:1902.03368, 2019.
- [9] Lan Z, Cai S, He X, et al. FixCaps: An improved capsules network for diagnosis of skin cancer[J]. IEEE Access, 2022, 10: 76261-76267.
- [10] Chanda D, Onim M S H, Nyeem H, et al. DCE NSnet: A new deep convolutional ensemble network for skin cancer classification[J]. Biomedical Signal Processing and Control, 2024, 89: 105757.
- [11] Sulthana R, Chamola V, Hussain Z, et al. A novel end to end deep convolutional neural network based skin lesion classification framework[J]. Expert Systems with Applications, 2024, 246: 123056.
- [12] İlker G, Özkan İ. SADASNet: A Selective and Adaptive Deep Architecture Search Network with Hyperparameter Optimization for Robust Skin Cancer Classification[J]. Diagnostics, 2025, 15(5): 541.
- [13] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.
- [14] Liu S, Wang L, Yue W. An efficient medical image classification network based on multi-branch CNN, token grouping Transformer and mixer MLP[J]. Applied Soft Computing, 2024, 153: 111323.
- [15] Ren Z, Liu S, Wang L, et al. Conv SdMLPMixer: A hybrid medical image classification network based on multi-branch CNN and multi-scale multi-dimensional MLP[J]. Information Fusion, 2025: 102937.

作者简介: 刘瑞欣(1999-)女,硕士,研究方向: 医学图像处理。

通信作者: 刘海华(1967-)男,教授,博士,研究方向: 视觉认知计算。