异构多分支网络超声图像自动诊断方法

李昕昕[1]，师　恩[2*]

(1. 四川大学锦城学院计算机与软件学院　成都　611731；2. 西南交通大学信息科学与技术学院　成都　610031)

【摘要】超声 (US) 是乳腺结节的主要影像学检查和术前评估手段之一，然而由于其在良性和恶性乳腺结节图像上的相似表现形式，使得诊断结果在很大程度上依赖于医生的经验判断，以计算机辅助手段帮助医生提高诊断准确率逐渐成为当前的热点。该文提出了一种用于乳腺超声图像的良性和恶性分类的异构多分支网络 (HMBN)，该网络同时使用了图像信息 (超声图像及造影图像) 和非图像信息 (包括患者年龄和其他 6 个病理特征)。还提出了一种适合该异构多分支网络的混合型损失函数，在附加角边距损失的基础上应用了最小超球面能量来提升分类精度。实验结果表明，在本文收集的包含 1 303 个病例的乳腺超声数据集上，该文提出的异构多分支网络的分类精度为 92.41%，相较于具有 5 年资历的医师的平均诊断准确性提升了 7.11%，相较于其他最新研究成果诊断准确性名列前茅。说明通过将医学知识纳入优化过程，以及将超声造影和非图像信息添加到网络中，可以在很大程度上提高乳腺超声诊断的准确性和鲁棒性。

关　键　词　乳腺超声；　造影图像；　异构数据；　病理；　超声影像识别

中图分类号　TP391.41　　**文献标志码**　A　　doi: 10.12178/1001-0548.2020246

Automatic Diagnosis Method of Ultrasound Image Based on Heterogeneous Multi-Branch Network

LI Xin-xin[1] and SHI En[2*]

(1. School of Computer and Software, Jincheng College of Sichuan University　Chengdu　611731;
2. School of Information Science and Technology, Southwest Jiaotong University　Chengdu　610031)

Abstract　Objective Ultrasound (US) is one of a primary imageological examination and preoperative assessment for breast nodules. However, in the field of ultrasound diagnosis, it relies heavily on the experience of physicians due to the overlapping image expression of benign and malignant breast nodules. Computer-aided medical diagnosis has gradually become a hot spot of current research. In this paper, a heterogeneous multi-branch network (HMBN) is presented for benign and malignant classification of the breast ultrasound images. In HMBN, the image information includes ultrasound images and contrast-enhanced ultrasound (CEUS) images while non-image information includes patients' age and other six pathological features. On the other hand, a fusion loss function suitable for this heterogeneous multi-branch network is also proposed. This loss function uses the minimum hyperspherical energy (MHE) based on additive angular margin loss to improve the classification accuracy. Experimental results show that on the breast ultrasound data set of 1 303 cases collected, the classification accuracy of the proposed heterogeneous multi-branch network is 92.41%, which is 7.11% higher than the average diagnostic accuracy of doctors with five years of experience, and ranks among the best in diagnostic accuracy in comparison with other latest research results. It is proved that the accuracy and robustness of breast diagnosis are greatly improved by incorporating medical knowledge into the optimization process and adding contrast-enhanced ultrasound images and non-image information to the network.

Key words　breast ultrasound; contrast-enhanced ultrasound; heterogeneous data; pathology; ultrasound classification

乳腺癌是目前女性发病率最高的恶性肿瘤，尽早发现可以提供更好的治疗机会[1]。当前常用的乳腺癌筛查手段包括：粗针穿刺、乳腺 X 线摄影、乳腺磁共振检查和超声影像。而超声影像因其无创

收稿日期：2020－06－02；修回日期：2020－10－22
基金项目：四川省重点研发项目 (2019YFS0432)
作者简介：李昕昕 (1981－)，女，副教授，主要从事人工智能、模式识别方面的研究.
通信作者：师恩，E-mail：1404916243@qq.com

伤、便捷、价廉，对致密性乳腺诊断敏感性和准确性高，成为首选的乳腺检查方法。但由于超声影像中良恶性结节的图像表现存在重叠，严重依赖医生经验，不同资历的医生误诊率达到 10%～30%[2]。因此，如何借助计算机视觉和大数据技术辅助临床医生提高对超声影像进行良恶性肿瘤判断的准确率，逐渐成为目前的研究热点。

针对上述问题，本文提出了一种端到端的自动结节分类方法。该方法主要采用基于乳腺超声图像的良性和恶性分类的异构多分支网络 (heterogeneous multi-brach network, HMBN)。包括如下 4 个模块：

1) 图像预处理：为提高超声图片的质量，本文采用基于 mean shift 聚类算法[3] 对原始视频进行关键帧选取。然后对超声及造影图像进行基于增强的 Wiener 滤波[4] 去噪和基于多重分形和对比度自适应的直方图均衡技术[5] 的增强处理。

2) 超声病理特征信息提取：病理信息包括原始超声图像提取的病理信息和包括年龄以及其他 6 个与乳腺良恶性相关度最高的乳腺病理特征在内的非图像信息。通过将病理信息做归一化处理后融合到分类的神经网络中。

3) 异构多分支网络模型：该模型的输入数据包括图像信息和非图像信息。对于图像信息，使用了由 ImageNet 数据集迁移的 Resnet50 网络进行特征提取，对于非图像信息，使用了一个全连接层将非图像的信息向量与其他分支结果融合进行分类。

4) 优化的损失函数：在附加角边距损失 (additive angular margin loss, AAM loss)[6] 的基础上应用了最小超球面能量 (minimum hyperspherical energy, MHE)[7]来改善分类效果。

1　相关工作

在乳腺超声的自动诊断分类任务中，文献 [8] 在针对乳腺病灶区进行筛选和分类的任务中，将自适应的去卷积网络模型作为一种产生式的无监督分层深度学习框架，用于诊断乳腺肿瘤的超声图像特征学习以及特征图谱生成，从而实现乳腺良恶性肿瘤的分类。实验结果表明该卷积网络的性能已达到传统手工特征的 CADx 系统的水平。文献 [9] 研究了一种将从预先训练的 CNN 中提取的低级和中级特征与从传统计算机辅助系统中获得的手工设计特征相结合的方法，并得到了显著的性能改善。文献 [10] 设计了 3 种能够分别接受横向平面图像，冠状面图像和图像注释信息的网络模型。实验结果表明，文中设计的模型可以同时处理图像和注释，且与单输入模型相比，多种信息融合模型将分类准确率提高了 2.91%，达到了 75.11%。

本文结合临床医学知识，将传统卷积神经网络扩展为异构的多分支网络进行乳腺超声的良恶性分类识别，最终获得了较好的结果。

2　本文算法

2.1　数据预处理

超声图像对于乳腺癌等疾病的诊断有着很大的优势。但是超声图像中可能包含大量噪声，尤其是斑点伪影和高斯噪声，因此需要大量的图像预处理工作来提升图像质量。

1) 超声图像去噪

为了探索针对乳腺超声图像数据集最佳的去噪方法，本文选择基于散斑统计的记忆各向异性扩散去噪方法 (anisotropic diffusion with memory based on speckle statistics, ADMSS)[11]、小波阈值算法、以及基于增强的 Wiener 滤波的超声图像去噪算法在团队自建的乳腺超声数据集上进行定量和定性评估。实验结果如表 1 所示。

表 1　不同算法的去噪效果对比

方法	ENL	USDSAI
原始图片	3.15	1
ADMSS	21.6	1.32
小波阈值	19.5	2.25
增强的Wiener滤波	31.5	6.41

实验中采用了两个应用最普遍的去噪效果评价指标：等效外观指数 (equivalent number of looks, ENL) 和超声去斑评估指数 (ultrasound despeckling assessment index, USDSAI)。其中，ENL 用于测量降噪效果，USDSAI 则用于测量去噪后图像中不同类别 (均质区域) 的可分离性，值越大表示该算法的性能越好。基于上述实验结果，本文选择了增强的 Wiener 滤波算法对乳腺超声图像进行去噪，基于增强的 Wiener 滤波方法实现过程如下。

首先考虑以下采集模型：

$$z(x,r)=y(x,r)n(x,r) \tag{1}$$

式中，$z()$ 是采集的数据；$y()$ 是无噪声信号；$n()$ 是斑点噪声；(x,r) 是根据采集几何定义的空间索引。然后通过应用对数变换，将乘法模型转换为加法模型，从而得到：

$$s'(x,r)=\log[z(x,r)]=\log[y(x,r)]+\log[n(x,r)]=y'(x,r)+n'(x,r) \quad (2)$$

最后在新的获取模型上，应用 Wiener 滤波器 (wiener filter, WF)。WF 是线性不变滤波器，其二维频率响应函数定义为：

$$\mathrm{HW}(\zeta,\eta)=\frac{\mathrm{Py}'(\zeta,\eta)}{\mathrm{Py}'(\zeta,\eta)+\alpha\mathrm{Wn}'(\zeta,\eta)} \quad (3)$$

式中，$\mathrm{Py}'(\zeta,\eta)=|\mathrm{Y}'(\zeta,\eta)|^2$是无噪声信号$y'(x,r)$的功率谱；$(\zeta,\eta)$为二维频谱中的空间频率。根据超声图片采集系统的不同，噪声随机过程表现出特定的自相关函数，其特征在于功率密度函数$\mathrm{Wn}'(\zeta,\eta)$；参数α用于调整滤波器强度的标量值，在标准 WF 下，将其设置为单位值 1。为了提高滤波效果，采用马尔科夫随机场 (Markov random field, MRF) 理论对图像进行建模。

2) 超声图像增强

为了更好地提升图像质量，还需要对乳腺超声图像进行数据增强。通过对多峰广义直方图均衡化算法 (multi-peak generalized histogram equalization, Multi-peak GHE)[12]，基于多重分形和对比度自适应直方图均衡的图像增强算法 (Clahe+Mfrac)[5] 以及一种基于模糊逻辑的增强算法[13] 进行对比实验，最终选择基于多重分形和对比度自适应直方图均衡的图像增强算法进行超声影像数据集的增强。

多重分形技术是增强医学图像和提取纹理特征的有效工具。多重分形技术中，最常用的 4 种强度度量分别是：求和度量、最大度量、逆最小度量和 iso 度量。与其他 3 个度量相比，逆最小度量在对斑点滤波过后的超声图片进行增强操作时，其处理后的图片在纹理特征等方面表现更好，因此本文在多重分形中使用逆最小度量。

将多重分形度量表示为$\mu_w(\rho)$，其中ρ是大小为w的方形窗口内的中心像素。令$g(k,l)$表示窗口内(k,l)处像素的强度值，且Ω表示窗口的中心像素ρ的所有邻近像素的集合。

计算多重分形特征的第一步是 Holder 指数α的估计。最小强度度量遵循逆幂定律，并给出α的负值。然后将计算出的最小值相对于最大强度值求反，以获得具有所需缩放属性的最小逆度量值：

$$\mu_{\mathrm{inv}}-\min(\rho)=1.0-\min_{(k,l)\in\Omega} g(k,l) \quad (4)$$

令图像中每个像素点的强度代表原始图像中相应点的 Holder 指数，得到与原始图像大小相同的α图像。为了生成α切片，选择图像中属于α值特定范围$(\alpha_{\min},\alpha_{\max})$的像素，并缩放其强度值。通过多重分形获得的输出特征增强图像在乳腺超声扫描图像中得到了改善。从图 1 为多重分形图像增强效果比较图。图 1a 为斑点滤波过后的乳腺超声图像，图 1b 为多重分形 (中心) 的α图像，图 1c 为使用多重分形测量值的增强型乳腺超声图像。图 1b 可以看出，将图像多重分解为α图像可用于表征超声图像中的解剖结构及各种形状和纹理特征。

a. 斑点滤波增强　b. 多重分形 α 图　c. 多重分形测量增强

图 1　多重分形图像增强效果图

表 2 为定量增强效果评价表。从表 2 可以看出，本文采用的基于多重分形和对比度自适应直方图均衡的图像增强算法 (Clahe+Mfrac)，在结构相似性指标 (structural similarity index metric, SSIM)、边缘保持系数 (edge preservation index, EPI) 和通用质量指标 (universal quality index, UQI) 这 3 个指标上较其他两种方法效果更好。图 2 为应用多重分形和对比度自适应直方图均衡增强算法后的超声造影效果图。

表 2　定量增强效果评价表

方法	SSIM	EPI	UQI
Clahe+Mfrac	0.761	0.874	0.913
Multi-peak GHE	0.675	0.783	0.785
基于模糊逻辑的增强算法	0.683	0.752	0.902

a. 原始超声造影图像

b. 增强效果图

图 2　多重分形和对比度自适应直方图均衡增强效果图

3) 图像关键帧提取

本文采用 mean shift 聚类算法对原始视频进行关键帧选取。将包含良性 (150 例) 和恶性 (150 例) 的多张乳腺超声视频进行均值化处理后，对每 5 帧图片进行距离可视化，得到图 3 所示的结果。

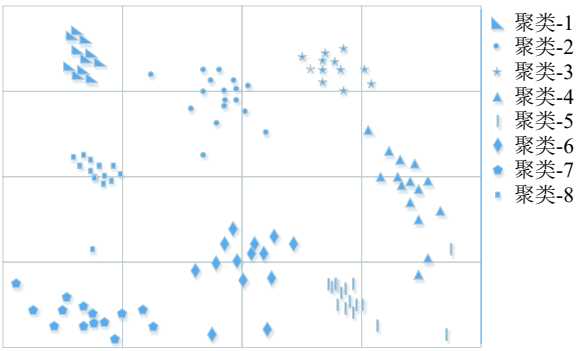


图 3　乳腺超声图像关键帧提取算法聚类结果

算法步骤：

1) 从视频帧序列中选取中间一帧作为起始关键帧；

2) 计算出当前关键帧与其前面所有图像和后面所有图像的相似度，得到前后平均相似度的值；

3) 将当前关键帧沿着平均相似度值大的方向移动，移动的距离为移动方向所有图像数量的一半，修改当前关键帧；

4) 重复步骤 2) 和步骤 3)，直到当前关键帧与其前后图像平均相似度值差距很小，即目前的关键帧为最后算法选取的关键帧。

为了量化评价 mean shift 聚类算法对乳腺超声视频关键帧提取的有效性，引入两个量化指标：视频保真度和视频压缩率。其中，保真度是提取的关键帧中所能表达的视频包含的有效信息。压缩率则是提取的关键帧总数与原始视频总帧数之比。将所有视频分类送入该算法并计算结果的平均值，得到 mean shift 聚类算法的关键帧提取结果，如表 3 所示。

表 3　乳腺超声图像关键帧提取算法聚类结果

视频类型	总帧数	关键帧数	平均保真度	压缩率
良性	2342	223	0.821	0.095
恶性	2653	284	0.845	0.107
合计	4995	507	0.833	0.101

2.2　损失函数设计

本文通过将最小超球形能量 MHE[7] 与 AAM loss[6] 相结合，提出一种新的混合损失函数。下面分别介绍附加角边距损失、MHE 与混合损失。

1) 附加角边距损失

附加角边距损失 AAM 函数是在传统 softmax 损失函数[14] 的基础上改进得到的。传统的 softmax 损失函数可表示为：

$$L_1=-\frac{1}{N}\sum_{i-1}^{N}\log\frac{\mathrm{e}^{\boldsymbol{W}_{y_i}^{\mathrm{T}}\boldsymbol{x}_i+b_{y_i}}}{\sum_{j=1}^{n}\mathrm{e}^{\boldsymbol{W}_j^{\mathrm{T}}\boldsymbol{x}_i+b_j}}\tag{5}$$

式中，$x_i\in R^d$表示第 i 个样本的深度特征；$\boldsymbol{W}_j\in R^d$表示权重$\boldsymbol{W}\in R^{d\times n}$的第 j 列；$b_j\in R^n$ 是偏置项。

但是，Softmax 损失函数由于缺少明确的优化嵌入式特征，无法为类内样本提供更高的相似度，为类间样本提供更高的多样性，导致当类内样本的外观差距较大时，图像识别的性能降低。而 AAM 函数在x_i和W_{y_i}之间增加了附加的角度损失 m，可以同时增强类内部的紧凑性和类间的差异：

$$L_{\mathrm{arc}}=-\frac{1}{N}\sum_{i=1}^{N}\log\frac{\mathrm{e}^{s(\cos(\theta_{y_i}+m))}}{\mathrm{e}^{s(\cos(\theta_{y_i}+m))}+\sum_{j=1,j\neq y_i}^{n}\mathrm{e}^{s\cos\theta_{y_i}}}\tag{6}$$

2) MHE

分类器神经元是每个类特征簇的近似中心。输出层的全空间 MHE 使得分类器神经元在超球体上分布更均匀，可以有效地改善类间特征的可分离性。MHE 将原始卷积视为角函数 $g(\theta)=\cos(\theta)$ 和幅度函数的乘积：

$$h(\|w\|,\|x\|)=\|w\|\|x\|:f(x,x)=h(\|w\|,\|x\|)g(\theta)\tag{7}$$

式中，θ是输入 x 和核 w 之间的角度。可以看出，核的方向和核的范数对内积相似性有着不同的影响。通常，可以通过最小化 l_2 范数来规范化权重衰减，但是内核的方向并没有被规范化。因此，使用 MHE 来规范化内核角度：

$$\begin{aligned}\mathrm{MHE}=&\underbrace{\frac{1}{m}\sum_{j=1}^{m}\frac{1}{m}l(\langle w_i^{\mathrm{out}},x_j\rangle_{i=1}^{c},y_{y_j})}_{\text{training data fitting}}+\\&\underbrace{\lambda_h\sum_{j=1}^{L-1}\frac{1}{N_j(N_j-1)}\{E_s\}_j}_{\text{Th:hyperspherical energy for hidden layers}}+\\&\underbrace{\lambda_o\frac{1}{N_L(N_L-1)}E_s(\widehat{w_i^{\mathrm{out}}}|_{i=1}^{c})}_{\text{To:hyperspherical energy for output layer}}\end{aligned}\tag{8}$$

式中，x_i 是进入输出层的第 i 个样本的特征；w_i^{out} 是输出全连接层中第 i 个类的分类器神经元，且

$\widehat{w}_i^{\text{out}}$表示其规范化形式；$\{E_s\}_i$代表第 i 层中神经元的超球面能量；m 是 batch 的大小；c 是 X 的类别数；L 表示神经网络的层数；N_i 是神经网络第 i 层的核数量；$E_s(\widehat{w}_i^{\text{out}}|_{i=1}^{c})$表示神经元的超球面能量$\left\{\widehat{w}_1^{\text{out}},\widehat{w}_2^{\text{out}},\cdots,\widehat{w}_c^{\text{out}}\right\}$；$\lambda_o$和$\lambda_h$是这 3 个正则化项的加权超参数。

从解耦的角度而言，MHE 在规范化神经网络方面具有许多意义，因为它可以充当权重衰减的补充和正交角色：

$$\text{MHE}=\lambda_M\frac{1}{m(M-1)}\sum_{i=1}^{m}\sum_{j=1,j\neq y_j}^{M}f_s(\|\widehat{w}_i^{\text{out}}w_j^{\text{out}}\|) \tag{9}$$

式中，m 是最小 batch 的大小；c 是类别数目；M 是分类器神经元的数量；x_i 表示第 i 个图片的深度特征 (是其真实的标签)；w_i^{out} 是第 i 个分类器的神经元。

3) 融合损失函数

通过以上分析和大量实验验证，AAM loss 的主要作用是减小类内距离，而 MHE 可以显著增加类间的间距。因此，本文将 MHE 和 AAM 进行融合，形成新的损失函数。

正则化项定义为：

$$L_{\text{Fus}}=L_{\text{arc}}+\text{MHE} \tag{10}$$

式 (10) 展开可表示为：

$$L_{F\text{us}}=\underbrace{-\frac{1}{N}\sum_{i=1}^{N}\log\frac{\mathrm{e}^{s(\cos(\theta_{y_i}+m))}}{\mathrm{e}^{s(\cos(\theta_{y_i}+m))}+\sum_{j=1,j\neq y_i}^{n}\mathrm{e}^{s\cos\theta_{y_i}}}}_{\text{ArcFace: promoting intra-class compactness}}+\underbrace{\lambda_M\frac{1}{m(M-1)}\sum_{i=1}^{m}\sum_{j=1,j\neq y_j}^{M}f_s(\|\widehat{w}_i^{\text{out}}w_j^{\text{out}}\|)}_{\text{MHE:promoting inter-class separability}} \tag{11}$$

上述融合损失函数可以进一步增加学习特征的角余量，减少类内距离，分类效果显著改善。

2.3 病理信息提取

从超声及造影图片或视频中提取结节病理信息，对于医学诊断具有重要意义。并且在本文所研究的乳腺超声自动诊断课题中，结节病理信息可以作为乳腺超声图片良恶性判断的依据。有时原始数据中的病理信息可能存在缺失的情况，且很多目前公开的医学超声图像数据集中基本都未包含与其相应的病理信息。因此，探究对乳腺超声诊断有意义的病理信息并从超声图片中将其提取出来对乳腺超声图片的诊断具有重要的意义。

对于从图片中提取的病理信息如表 4 所示，包括 3 个方面的工作。

表 4 病理信息提取技术表

病理特征	提取方法	所需数据
增强均匀性	CNN-RNN多标签分类技术	超声图片
蟹足征	CNN-RNN多标签分类技术	超声图片
滋养血管	CNN-RNN多标签分类技术	超声图片
增强后边界	CNN-RNN多标签分类技术	超声图片
增强后形态	CNN-RNN多标签分类技术	超声图片
增强后病灶范围	病灶定位及分割技术	超声和造影图片
增强强度	峰值及斜率计算法	造影视频
增强时相	峰值及斜率计算法	造影视频

1) 基于多标签分类的病理提取

基于多标签的乳腺造影图片病理信息的提取主要包含以下内容：增强均匀性 (类别：均匀、不均匀)、蟹足征 (类型：有、无)、滋养血管 (类型：无、有)、增强后边界 (清楚、不清楚、难以分辨) 和增强后形态 (规则、不规则、难以分辨)。这些基于乳腺造影图片的病理信息被认为是对良恶性判断非常有帮助的病理信息[15]。首先通过已有的文本格式的病理信息对数据集进行多标签标记，然后将训练数据集作为多标签分类的训练数据集进行分类并用测试数据集对分类结果进行定量分析，分析结果如表 5 所示。

表 5 良性和恶性乳腺病变形态特征的比较

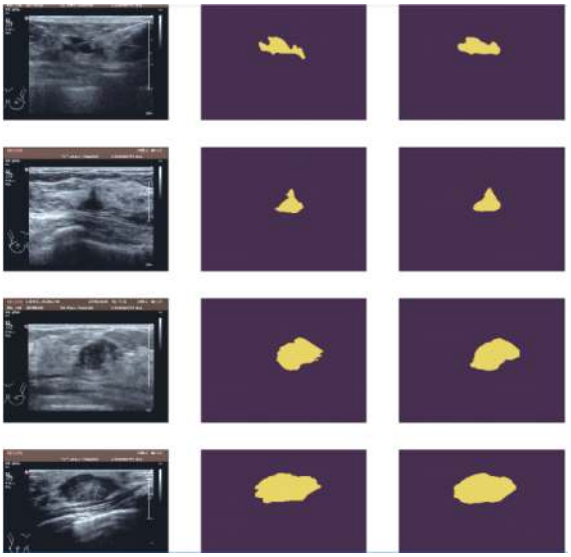
造影病理信息		良性		恶性	
		数目	比例/%	数目	比例/%
增强后病灶范围	扩大	203	25	590	75
	不变	599	87	84	13
	缩小	49	75	17	25
	难以分辨	169	64	39	36
增强后均匀性	均匀	459	72	184	28
	不均匀	539	51	540	49
蟹足征	有	868	70	380	30
	无	122	26	337	74
滋养血管	无	830	69	373	31
	有	121	26	340	74
增强后边界	清晰	551	72	219	28
	不清晰	258	36	468	64
	难以分辨	186	15	34	85
增强后形状	规则	489	83	100	17
	不规则	207	26	567	74
	难以分辨	294	87	44	13

2) 增强后病灶范围提取

由于增强后病灶范围的特征描述的是通过造影 (超声增强) 后的图像病灶范围的大小相比于普通超声图像中的病灶范围的变化情况。因此单独使

用超声图像或造影图像都不能得出针对该病理特征的正确结论。同时，有相关研究发现，增强后病灶范围扩大特征与恶性病灶的相关性更高[15]。因此，本文同时使用到了病灶定位和病灶区分割技术。病灶区分割技术采用 Mask-RCNN[16]。首先分别输入超声图片和其对应的造影图片，使用病灶区定位及分割技术，确定病灶范围；然后分别通过分割结果(结节形状掩码) 进行计算；最后对计算出的超声病灶范围和其对应的造影病灶范围进行判断，确定病灶区范围 (扩大/缩小/不变)。

病灶区分割结果如图 4 所示，其中第一列为原始图像，第二列为实际病灶区掩码，即标签，第三列为分割结果。



a. 原图　　b. 标签　　c. 分割结果

图 4　病灶区分割结果示例

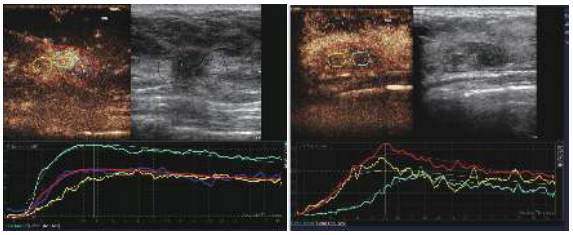
3) 增强强度及增强时相后的病理特征提取

增强强度和增强时相这两个病灶特征是通过分析造影视频得出的。增强强度是指病灶区相对于其他正常区域达到最亮的峰值高低，而增强时相则表征病灶区相对于其他正常区域增强的快慢。如图 5 所示，该图为某医院的设备中，将造影视频的亮暗信息进行分析后所绘制的波形图。其中图 5a 中的恶性病灶显示出增强时相快进、高增强的特征，而图 5b 中的良性病灶显示出增强时相慢进、低增强的特征。

针对这两个病灶特征的提取流程为：首先，分析整个视频的病灶区域并记录下病灶区平均亮度的最大值 p 和到达该值时的时间值 t，然后通过斜率 k 和峰值 p 的经验阈值对其进行判断分析。

k 值的计算方式为：

$$k = \frac{p}{t} \tag{12}$$



a. 恶性病灶增强时相和增强强度　　b. 良性病灶增强时相和增强强度

图 5　增强时相和增强强度示例图

图 5a 中的病灶区的平均亮度最大值为 15.63，到达该值的时间为 12.33 s，通过式 (12) 得到 k 值为 1.27[17]。由于不同的医生可能存在不同的经验阈值，针对本文实验的数据，设定经验阈值为 0.74。

对于增强强度及增强时相后的病理特征，采用医生的经验值作为算法的经验阈值。该经验值即为：将病灶区域与病灶周围的其他正常区域 (两个区域的组织结构尽可能保持一致，且区域范围由医生标注) 相比较。根据良恶性肿瘤的一般特性，如果病灶区域的增强时相比正常区域更快，且其增强强度更高则其病灶为恶性的概率则更大，反之亦然。另外，为了尽量减少非相关变量对结果的影响，算法还对该区域进行了归一化操作 (结果除以该区域的面积)。根据该经验阈值进行计算后的准确率结果分析显示在表 5 中。

2.4　异构多分支超声图像识别网络

调研发现，目前已有的针对超声图片良恶性诊断的研究仍然存在以下局限性：

1) 现有的大多数研究仅基于乳腺超声图像，因此获得的结果很大程度上依赖于输入图像的质量。本文提出的方法将超声图像与相应的造影图像相结合，增加了乳腺病变的信息量，为神经网络的特征提取提供了更丰富的信息。

2) 目前所做研究大多数都没有使用医生的经验信息，例如病人或病灶的描述信息等。本文提出的方法以医学专业知识为基础，将临床上针对良恶性判别最有效的病理特征提取到网络中，进一步提高了分类的准确性。

针对目前研究工作的不足，本文提出了一种端到端的乳腺超声自动结节分类模型。该模型用于乳腺超声图像的良性和恶性分类的异构多分支网络。

该网络通过将医学知识作为非图像信息，和经过数据预处理的超声图像信息一起输入异构多分支网络，提高了乳腺超声诊断的准确率。

本文实验所采取的技术路线如图 6 所示。

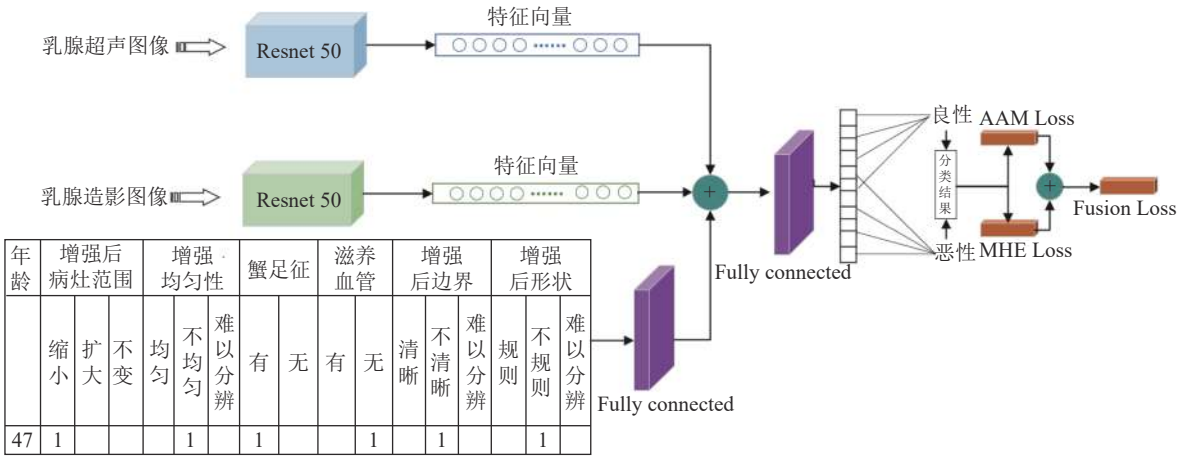


图 6 异构多分支超声良恶性分类实验技术路线图

分类步骤分为 3 步：1) HMBN 网络采用 ResNet 作为主干网，其中，两个分支采用 Rsenet50 用于提取超声图像和造影图像的特征。即将大小为 224×224 的超声图片和造影图片通过网络中步长为 2、大小为 3×3 的卷积核和全局平均池化层操作，得到 7×7 的特征图以及具有 Softmax 的全连接的分类层。为了避免瓶颈，当特征图的空间大小变化时，特征的数量将增加一倍。每次卷积后，应用 batch 归一化和 ReLU 函数。

2) 对于单独的用于处理非图像信息的分支，使用一个全连接层将非图像信息 (包含患者的年龄和病灶的病理特征 (如表 5 中 6 项造影病理信息所示)) 所构成的特征矢量的长度从 7 转换到 2048。对于用于提取图像特征的分支，使用在 ImageNet 数据集上预训练得到的模型，在模型输出时去掉最后一个卷积层，得到长度为 2048 的特征向量。

3) 通过相加运算和一个全连接层对 3 个异构网络的分支结果进行融合，使用本文提出的 MHE 和 AMM 的融合损失函数得到最终的诊断分类结果。

3 实验结果分析

3.1 数据集介绍

本文实验所用的数据集来源于四川省的若干个医疗机构，由各医疗机构的超声科采集的乳腺超声科室提供。原始的乳腺超声数据集中包含 3 种不同格式的视频文件，其中每个视频文件分别包含超声图像和相应的造影文件，以及与超声造影视频文件相对应的乳腺详细病理报告。不同医院所用设备不同，因此所采集的图像的分辨率和噪声情况等都略有不同。

综上所述，本实验所用乳腺超声数据集包含 791 例良性病例和 572 例恶性病例，每个病例通过视频关键帧提取技术抽取了包含 100～200 张超声图像和相应的造影图像和与其对应的病理报告。然后将数据集分为 3 个独立的数据集，分别是包含 1 166 个样本集的训练集、包含 102 个样本集的验证集和包含 95 个样本集的测试集。其分配比例基本按照 10∶1∶1 的关系进行随机抽取。表 6 中显示了该数据集的详细信息。

表 6 乳腺超声病例数据集

类别	训练集	验证集	测试集	总数
良性	680	58	53	791
恶性	486	44	42	572
总数	1 166	102	95	1 363

其中，该数据集的病理报告组成有：研究中心基本情况、病灶基本资料 (如为多发病灶等)——灰阶超声、患者基本资料 (包含患者年龄、性别、民族、身高及体重等)、CEUS、超声造影增强模式定性数据、乳腺钼靶、MRI、穿刺活检病理、术后病理、随访内容、其他情况。

3.2 病理信息提取结果

为选取最适合乳腺超声数据病理信息提取的多标签分类的网络模型，本文基于乳腺超声数据集分别对 CNN-RNN 多标签分类技术[18]、GCN 多标签分类技术[19] 及 VGG16 多标签分类技术[20] 进行多标签提取的实验。为了验证本文所设计的基

于造影视频的增强强度和增强时相特征的提取，以及基于乳腺超声图片和其对应的造影图片的增强后病灶范围特征提取的有效性，分别做了准确率的验证实验。本将主要针对基于多标签分类的病理提取技术的实验结果和其他病理信息的提取实验结果进行比较和分析。

1) 基于多标签分类的病理提取技术实验

为了提取乳腺超声图片中的病理特征 (增强均匀性、蟹足征、滋养血管、增强后边界和增强后形态)，综合考虑算法时间复杂度、准确率、发表时间等因素分别筛选出 3 个网络进行多标签分类的训练和测试实验：CNN-RNN 网络，GCN 网络和 VGG16 网络。这 3 种网络下的分类实验结果如表 7 所示。表中显示了多个标签累计分类结果的准确率、敏感性及特异性。

表 7　多标签分类网络病理信息提取实验结果

网络	准确率/%	敏感性/%	特异性/%
VGG16	58.3	83.76	61.64
GCN	73.33	69.54	75.42
CNN-RNN	77.54	73.53	79.82

由表 7 可知，CNN-RNN 网络、基于 GCN 的网络及 VGG16 这 3 种网络的准确率都有待提高，其中表现最佳的是 CNN-RNN 网络。为了验证从超声图片中提取的病理信息可以有效提升分类的准确率，将通过上述 3 种网络提取的结果输入到只包含超声图片和造影图片的 HMBN 二分支网络中，进行分类结果的对比，其实验结果如表 8 所示。

表 8　不同多标签分类网络结果对 HMBN 良恶性分类结果的影响

网络	准确率/%	敏感性/%	特异性/%
VGG16+二分支网络	80.4	83.76	61.64
基于GCN+二分支网络	82.8	69.54	75.42
CNN-RNN+二分支网络	83.24	85.43	86.52

由表 8 可知，由于 VGG16 网络所提取的病理特征在良恶性分类时的准确率较低，无法对分类效果进行改善，甚至对网络有反向的抑制作用。而基于 GCN 的网络和基于 CNN-RNN 的方法在加入 HMBN 的二分支网络后，准确率有了明显提升，而 CNN-RNN+二分支网络的分类准确率最高，达到 83.24%。因此，本文后续实验选用 CNN-RNN 网络对超声图片内的病灶特征进行提取。

2) 其他病理信息提取技术实验

其他病理信息提取涉及增强后病灶范围 (enhanced lesion range, ELR)、增强时相 (enhanced phase, EP) 及增强强度 (enhanced strength, ES) 这 3 个病理特征的提取。考虑到在医院的数据分析仪器中，增强强度和增强时相的定义与造影图片亮度值变化曲线的峰值和斜率相关，因此本文认为利用该定义进行特征提取可以获得很好的效果，如表 9 所示。

表 9　其他病理信息提取结果

病理特征	准确率/%	敏感性/%	特异性/%
增强后病灶范围	92.1	93.6	95.3
增强时相	96.4	97.7	98.6
增强强度	97.8	98.3	98.3

为了验证该方法提取到的增强后病灶范围、增强时相及增强强度这 3 个特征对多分支网络进行乳腺良恶性分类结果的影响，分别在训练集和测试集上做了训练和测试，其中包含两个方面的实验：1) 提取得到的增强后病灶范围、增强时相及增强强度这 3 个特征相对于二分支网络的作用；2) 结合 CNN-RNN 网络所提取的其他病理特征和这 3 个特征共同作用下相对于二分支网络的分类作用，实验结果如表 10 所示。

表 10　提取的病理特征对二分支网络良恶性分类的影响

病理特征	准确率/%	敏感性/%	特异性/%
二分支网络	82.1	83.6	85.3
二分支网络+（ELR+EP+ES）	83.3	86.2	87.5
二分支网络+（ELR+EP+ES）+CNN-RNN	84.7	88.2	89.4

从表 10 可以看出，利用本文设计的其他病理信息提取技术提取的增强后病灶范围、增强时相及增强强度这 3 个特征对网络具有积极改善的作用，且将其与利用 CNN-RNN 网络所提取的超声图片中的特征相结合后，对异构多分支网络中良恶性肿瘤的的识别性能有明显改善。但是其 84.7% 的分类准确率相较于表 1 中采用 HMBN 网络得到的 92.41% 的准确率仍有很大的改进空间。

3.3　实验分析

3.3.1　HMBN 实验过程

针对超声图像分支，输入图像大小为 224×224，当 Resnet50 网络在 con2_x 提取图像特征时，卷积核为 3×3，输入的特征图大小为 56×56×64，输出的特征图大小为 28×28×128，其过程如图 7 所示。

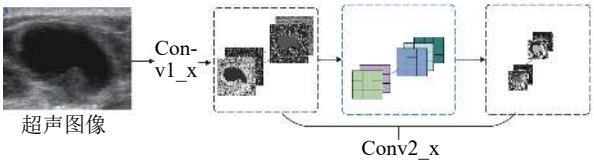
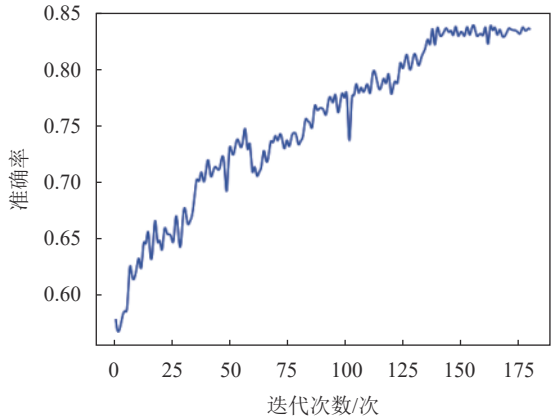
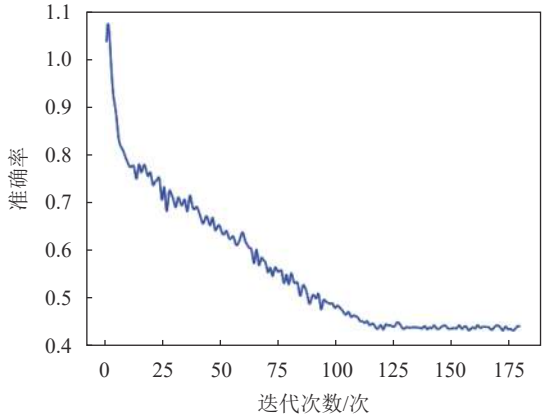


图 7 超声分支网络特征提取图

在训练过程中，每一个 epoch 中取出其中 3 个阶段的精确值和损失函数的值，得到训练精度变化曲线和 loos 变化曲线，如图 8 所示。从图中可以看出在第 125 个 epoch 左右收敛。



a. 训练精度变化



b. loss 变化

图 8 训练精确度和 loss 变化情况

3.3.2 不同数目分支分类效果对比实验

为了验证本文提出的异构多分支网络的有效性，分别对以下 4 个不同结构的网络进行实验并进行比较和分析：

1) Single-Stream net_1，即仅使用一个 ResNet-50 网络。它的输入图像是通过超声图像及其对应的对比度增强超声图像的拼接形成的；

2) Single-Stream net_2，其主干网络结构与 Single-stream net_1 的网络相同，但该网络的输入图像仅是超声图像；

3) Two-Stream network，该网络的主干网络是 ResNet50，它是具有两个单独的图片特征提取网络的双流网络，可同时处理超声图像和其相应的造影图像，但是没有添加处理非图像的病理特征的分支网络。

4) HMBN 网络，即本文提出的网络，是同时使用图像特征和非图像特征的异构多分支网络。

这 4 个网络的对比实验结果如表 11 所示。

表 11 不同网络结构准确率比较

网络	准确率/%
HMBN	92.41
Two-Stream network	89.45
Single-Stream net_1	88.39
Single-Stream net_2	85.68

从实验结果可以看出，同时使用了超声图像、造影图像和非图像信息的病理特征的 HMBN 网络的识别准确率最高，达到 92.41%。证明造影图像和非图像的病理特征对提高乳腺超声诊断的准确性具有积极的作用。

3.3.3 损失函数对比实验

为了验证本文设计的融合损失函数对乳腺超声诊断的异构多分支网络的有效性，进行了多组比较实验。首先，将广泛使用的原始 Softmax 损失函数用作参考实验，然后将附加角边距损失函数作为第一组对比实验，将 SphereFace +损失函数作为第二组对比实验，而将本文设计的融合损失函数作为第三组对比实验。最终的实验结果如表 12 所示。

表 12 不同损失函数分类效果比较

损失函数	准确率/%	错误样本数(总数：14 250)
Softmax损失	87.46	1 787
附加角边距损失	90.63	1 335
SphereFace+损失	91.7	1 144
融合损失	92.41	1 081

从表 12 可以看出，本文所提出的融合损失函数具有最佳的表现，其测试集的准确率达到 92.41%。

3.3.4 与其他方法的对比实验

目前针对乳腺超声图像良恶性分类效果最好的方法有两种：1) 基于加权 BI-RADS 的 B-wBI[21]，该方法利用减少病理分级的方法解决分类中常见的分级失衡问题，进而提高诊断性能。其中经过加权训练的 RF 分类器用于对提取的特征进行分类；2) 基于 3D-CNN 的方法[22]。为了验证本文提出的 HMBN 方法与其他方法的对比实验效果，将上述 3 种方法在数据集上进行了实验，由于 B-wBI 和 3D-CNN

两种方法均未使用造影图像、年龄以及病理信息等非图像信息，因此实验中仅将超声图像构造为以上两种方法的训练和测试数据集。实验结果如表 13 所示。

表 13　HMBN 与最新技术分类准确性的比较

方法	准确率/%	错误样本数 (总数：14 250)
B-wBI	91.89	1 155
3D-CNN	92.82	1 026
HMBN	92.41	1 081

在准确性方面，HMBN 的准确性比 B-wBI 方法提高了 0.52%，但略低于 3D-CNN。原因在于 3D-CNN 直接处理视频，考虑了超声视频中时间维度信息，而 HMBN 方法在融合病理信息时虽然也引入了与时间相关的增强强度和增强时相等病理特征，但是并不能将时间维度上的信息完全利用，所以导致 HMBN 的分类准确率略低于 3D-CNN。但是 3D-CNN 对图像质量要求较高，目前符合其要求的医学超声图像并不多，而 HMBN 由于采取了前期的数据预处理，更适合现有的数据集，具有更好的普遍适用性。

4　结 束 语

本文的工作主要分为 3 点：1) 通过对乳腺超声图像去噪和图像增强操作在一定程度上改善了超声图像的质量；2) 提出了一种异构多分支网络，将 3 个针对良恶性诊断相关度较高的信息 (即超声图像、造影图像和包括患者年龄的非图像信息) 和其他 6 个病理信息结合在一起进行乳腺超声影像的良恶性分类的网络模型；3) 基于最小超球面能量和附加角边距损失的融合提出了一种新的损失函数。实验结果表明，上述工作对于提升乳腺良恶性分类的性能和准确率都有积极的效果，其最终识别准确率在目前最新的分类算法中名列前茅。

参　考　文　献

[1] PISANI P, PARKIN D M, FERLAY J. Estimates of the worldwide mortality from eighteen major cancers in 1985. Implications for prevention and projections of future burden[J]. Int J Cancer, 1993, 55(6): 891-903.

[2] GIGER M L, CHAN H P, BOONE J. Anniversary paper: History and status of CAD and quantitative image analysis: The role of medical physics and AAPM[J]. Medical Physics, 2008, 35(12): 5799-5820.

[3] HANJALIC A, ZHANG H J. An integrated scheme for automated video abstraction based on unsupervised cluster validity analysis[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(8): 1280-1289.

[4] BASELICE F, FERRAIOLI G, PASCAZIO V, et al. Enhanced wiener filter for ultrasound image denoising[C]// European Medical and Biological Engineering Conference Nordic-Baltic Conference on Biomedical Engineering and Medical Physics. Singapore: Springer, 2017: 65-68.

[5] SINGH P, MUKUNDAN R, RYKE R D. Feature enhancement in medical ultrasound videos using multifractal and contrast adaptive histogram equalization techniques[C]//The 2nd International Conference on Multimedia Information Processing and Retrieval (MIPR). [S.l.]: IEEE, 2019: 240-245.

[6] DENG Jian-kang, GUO Jia, XUE Nian-nan, et al. ArcFace: Additive angular margin loss for deep face recognition[C]// Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.]: IEEE, 2019: 4690-4699.

[7] LIU W, LIN R, LIU Z, et al. Learning towards minimum hyperspherical energy[C]//Proceedings of the 32nd International Conference on Neural Information Processing SystemsDecember (NeurIPS). [S.l.]: ACM, 2018: 6225-6236

[8] JAMIESON A, DRUKKER K, GIGER M. Breast image feature learning with adaptive deconvolutional networks[C]// Proceedings of the SPIE Medical Imaging. San Diego: SPIE, 2012: 64-76.

[9] ANTROPOVA N, HUYNH B Q, GIGER M L. A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets[J]. Medical Physics, 2017, 44(10): 5162-5171.

[10] KONG X, TAN T, BAO L, et al. Classification of breast mass in 3D ultrasound images with annotations based on convolutional neural networks[J]. Chinese Journal of Biomedical Engineering, 2018, 37(4): 414-422.

[11] RAMOS-LLORDN G, VEGAS-SNCHEZ-FERRERO G, MARTIN-FERNANDEZ M, et al. Anisotropic diffusion filter with memory based on speckle statistics for ultrasound images[J]. IEEE Transactions on Image Processing, 2015, 24: 345-358.

[12] SHI X, CHENG H D, HU L, et al. Detection and classification of masses in breast ultrasound images[J]. Digital Signal Processing, 2010, 20(3): 824-836.

[13] GUO Y, CHENG H D, HUANG J, et al. Breast ultrasound image enhancement using fuzzy logic[J]. Ultrasound in Medicine & Biology, 2006, 32(2): 237-247.

[14] OMKAR M P, ANDREA V, ANDREW Z. Deep face recognition[C]//Proceedings of the British Machine Vision Conference (BMVC). UK: BMVA, 2015: 41.1-41.12.

[15] SHEN Ruo-xia, YANG Li-chun, LUO Xiao-mao, et al. Retrospective study of qualitative features of contrast-enhanced ultrasound in benign and malignant breast lesions based on data from Chinese multicenter studies[J]. Chinese Journal of Medical Imaging, 2018, 26(12): 11-15.

[16] 杜章锦, 龚勋, 罗俊, 等. 乳腺超声图像中易混淆困难样本的分类方法[J]. 中国图象图形学报, 2020, 25(7): 1490-1500.
DU Zhang-jin, GONG Xun, LUO Jun, et al. Classification

method for samples that are easy to be confused in breast ultrasound images[J]. Journal of Image and Graphics, 2020, 25(7): 1490-1500.

[17] 沈若霞, 杨丽春, 罗晓茂, 等. 基于中国多中心研究数据的乳腺良恶性病灶超声造影定性特征的回顾性研究[J]. 中国医学影像学杂志, 2018, 26(12): 885-889.
SHEN Ruo-xia, YANG Li-chun, LUO Xiao-mao, et al. Qualitative features of contrast-enhanced ultrasonography in benign and malignant breast lesions based on data of a multi-center study in China: A retrospective study[J]. Chinese Journal of Medical Imaging, 2018, 26(12): 885-889.

[18] WANG J, YANG Y, MAO J, et al. CNN-RNN: A unified framework for multi-label image classification[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 2285-2294.

[19] CHEN Z, WEI X, WANG P, et al. Multi-label image recognition with graph convolutional networks[C]// Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2019: 5177-5186.

[20] KAREN S, ANDREW Z. Very deep convolutional networks for large-scale image recognition[C]// International Conference on Learning Representations. Ithaca, NY: [s.n.], 2015: 1-14.

[21] RODRÍGUEZ-CRISTERNA A, GÓMEZ-FLORES W, DE ALBUQUERQUE P W C, et al. A computer-aided diagnosis system for breast ultrasound based on weighted BI-RADS classes[J]. Computer Methods and Programs in Biomedicine, 2018, 153: 33-40.

[22] CHIANG T, HUANG Y, CHEN R, et al. Tumor detection in automated breast ultrasound using 3-D CNN and prioritized candidate aggregation[J]. IEEE Transactions on Medical Imaging, 2019, 38(1): 240-249.

编 辑 叶 芳