激光与光电子学进展 先进成像

基于自适应聚合权重联邦学习的肺结节CT图像分类

侍江烽[1]，冯宝[2*]，陈业航[2]，陈相猛[3]

[1]桂林电子科技大学电子工程与自动化学院，广西 桂林 541004；

[2]桂林航天工业学院生物医学与人工智能实验室，广西 桂林 541004；

[3]江门市中心医院医学影像智能计算及应用实验室，广东 江门 529030

摘要 针对目前医学影像面临多中心数据存在数据孤岛以及非独立同分布的问题(Non-IID)，提出了一种基于自适应聚合权重的联邦学习算法(FedAaw)。在全局模型聚合过程中，提出准确率阈值来筛选出本地模型，并由中心服务器采用筛选后模型的准确率计算相应的聚合权重，从而对全局模型进行聚合，使得分类性能较佳的模型参与全局模型的构建，以达到缓解多中心数据Non-IID的问题。同时，为提高模型挖掘图像长短距离信息之间的能力，在本地和全局模型中引入多头自注意力(MHSA)机制。此外，为缓解端对端的冗余特征造成的模型过拟合问题，提取全局模型中卷积核的特征，并采用基于L_1范数的稀疏贝叶斯极限学习机(SBELML$_1$)的集成学习方法完成各中心数据的特征分类。最后，通过多次打乱不同中心的数据分布来验证FedAaw算法的抗干扰能力。5个中心的测试集AUC变化范围为中心1(0.7947～0.8037)、中心2(0.8105～0.8405)、中心3(0.6768～0.7758)、中心4(0.8496～0.9063)、中心5(0.8913～0.9348)，该结果表明：FedAaw在多中心数据上具有良好的分类性能且抗干扰能力较强。

关键词 自适应聚合权重；联邦学习；多头自注意力；L_1范数的极限学习机；对抗验证

中图分类号 TN911.73-34；TP391.41 **文献标志码** A **DOI**：10.3788/LOP223027

Lung Nodule CT Image Classification Based on Adaptive Aggregate Weight Federated Learning

Shi Jiangfeng[1], Feng Bao[2*], Chen Yehang[2], Chen Xiangmeng[3]

[1]*School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin* 541004, *Guangxi, China*;

[2]*Laboratory of Artificial Intelligence of Biomedicine, Guilin University of Aerospace Technology, Guilin* 541004, *Guangxi, China*;

[3]*Laboratory of Intelligent Computing and Application of Medical Imaging, Jiangmen Central Hospital, Jiangmen* 529030, *Guangdong, China*

Abstract The field of medical imaging currently faces the problems of data island and non-independent and independently distributed (Non-IID) variables in multi-center data. In this study, a federated learning algorithm based on adaptive aggregate weight (FedAaw) is proposed. Using a global model polymerization process, this study utilized the accuracy threshold to filter out the local model; the model accuracy is calculated by the center server. The corresponding weights of polymerization, which are updated in the global model, yielded models with better classification performances that are used to construct a global model, which helps address the problems associated with Non-IID multicenter data. Furthermore, to improve the applicability of the model to mining the information between the long and short distance of the image, the multi head self-attention mechanism is introduced to the local and global models. In addition, to address the problem of model overfitting caused by end-to-end redundant features, the convolution kernel features in the global model are extracted. The learning of sparse Bayesian extreme learning machine based on L_1 norm (SBELML$_1$) framework is used for the feature classification of the data obtained from each center. Finally, the anti-interference ability of the FedAaw algorithm is verified by shuffling the data distribution of different centers several times. The AUC ranges of the test sets

收稿日期：2022-11-11；**修回日期**：2022-12-24；**录用日期**：2023-02-22；**网络首发日期**：2023-03-09

基金项目：国家自然科学基金(81960324,62176104)、广西自然科学基金(粤桂联合基金)(2021GXNSFAA075037)、广东省医学科学技术研究基金(A2021138)、桂林航天工业学院校级科研基金(XJ21KT24)

通信作者：*fengbao1986.love@163.com

used in the five centers are as follows: center 1: (0.7947-0.8037), center 2: (0.8105-0.8405), center 3: (0.6768-0.7758), center 4: (0.8496-0.9063), and center 5: (0.8913-0.9348). These results indicate that FedAaw has good classification performance on multi-center data and a strong anti-interference ability.

Key words adaptive aggregate weight; federated learning; multi-head self-attention; L_1-norm extreme learning machine; against the validation

1 引 言

随着计算机断层扫描(Computed tomography, CT)技术的发展,孤立型肺实性结(Solitary pulmonary solid nodule, SPSN)的检出率逐渐提高。SPSN是一种肺部的病变组织,具有良恶性之分,且均在CT图像上表现为实性结节。肺腺癌(Lung adenocarcinoma, LAC)属于恶性实性结节的一种,在生物学行为上属于侵袭性生长,从而表现为向周围组织浸润侵犯而出现分叶征、毛刺征、胸膜牵拉征等影像学征象;肺结核(Lung tuberculosis, LTB)属于良性结节的一种其大部分呈现慢性生长过程,但部分结核性肉芽肿结节内部也会发生机化和纤维化,结节形态收缩,在CT图像上亦出现分叶征等征象。同时,在临床干预处理上,LAC需要尽早手术切除处理,改善患者的预后和生活质量;LTB应该避免侵入性诊疗方案如穿刺活检或手术干预等。临床实践中,主要通过影像学征象和医生的主观经验对LAC和LTB进行分类诊断,但由于影像学征象难以定量评估,同时临床医生的主观诊断经验一致性不高,导致了实性肺结节的精准诊断和个性化治疗较难[1]。因此,通过先进和无创的方法,在术前准确鉴别诊断LAC和LTB,具有非常重要的临床价值。

近年来,深度学习方法具有较强学习特征的能力,能从影像数据中挖掘出与任务相关的深度特征,从而获得更高的疾病诊断准确率,因此其应用于医学辅助诊断的各个领域[2-5]。然而,深度学习方法需要大样本数据才能训练出一个鲁棒模型。在临床数据分析中,因医疗数据隐私规定,各机构之间不能收集和共享患者的数据,同时多中心机构的数据存在采集设备、跨地域、图像质量等差异,从而导致了单中心医学影像数据样本量较少和多中心数据分布差异较大,进而影响深度模型的性能[6]。

针对以上问题,近年来联邦学习(Federated learning, FL)被一些学者提出[7]。FL是指多中心机构在不交换本地数据的基础上,共同参与训练模型的一种算法。Li等基于Fedavg[7]的基础上在本地模型优化中引入正则化项,以约束本地模型与全局模型间的参数差异,从而抑制数据的漂移[8]。Li等[9]通过在本地模型训练期间,计算本地模型与全局模型的余弦相似度以缩小本地模型与全局模型之间的差异,以确保本地模型训练期间不会过于偏离中心最优解。Jiang等[10]提出使用图像低频特征和幅度归一化的方式来减少模型受异质性数据的干扰,同时为本地模型训练加入扰动项以扩展模型收敛区间,从而减少模型参数的漂移。上述方法,虽然能够提高全局模型的收敛速度和稳定性,但存在两个问题:1) 多中心的数据存在Non-IID的问题[11-12],导致多中心数据参与全局模型聚合时,因部分本地模型分类性能不佳,从而影响全局模型分类性能下降;2) 本地和全局模型均采用卷积神经网络(Convolutional neural network, CNN)来对图像局部进行特征提取,无法建模图像长短距离的依赖关系,从而影响模型的分类效果[13-15]。

基于上述问题,本文提出了一种基于自适应聚合权重的联邦学习(Federated learning of adaptive aggregate weight, FedAaw)算法,来缓解多中心数据存在数据孤岛以及Non-IID的问题。首先,在FL中,提出FedAaw策略对全局模型进行聚合。FedAaw采用各本地模型的准确率筛选出符合要求的本地模型,并通过筛选后本地模型的准确率计算聚合权重,以使全局模型进行自适应聚合,从而提高全局模型的泛化性能。然后,在本地模型以及中心服务器模型构建中,考虑到不同CT图像的肺结节形状、大小存在差异,若采用卷积层提取肺结节特征容易丢失图像长短距离的依赖关系,从而影响模型的分类性能。因此,采用多头自注意力(Multi head self-attention, MHSA)机制来构建肺结节CT图像中的长距离之间的依赖关系,提取多尺度的肺结节空间特征。此外,为缓解端对端的冗余特征造成的模型过拟合问题,通过提取全局模型中的卷积核特征,并采用集成学习的方法来完成特征分类,实现LAC与LTB的分类。

2 FedAaw的算法框架

FedAaw的整体算法结构如图1所示。该算法包括FedAaw的特征提取和基于L_1范数的稀疏贝叶斯极限学习机(Sparse Bayesian extreme learning machine based on L_1 norm, SBELML$_1$)的特征分类两个部分。特征提取部分:首先,使用各个中心的数据训练一轮各自的本地模型(基于MHSA的MsaNet);然后,各个本地模型上传该轮训练得到的模型参数和准确率至中心服务器;最后,中心服务器通过自适应聚合权重(Adaptive weight, Aaw)机制来完成全局模型(基于MHSA的MsaNet)的构建,并将全局模型参数下发至本地服务器用以更新本地模型参数。直到迭代次数达到预先设置后,将全局模型的卷积核作为特征提取器对各个中心的数据进行特征提取。分类器构建部分:当提取特征后,各个本地服务器分别使用SBELML$_1$

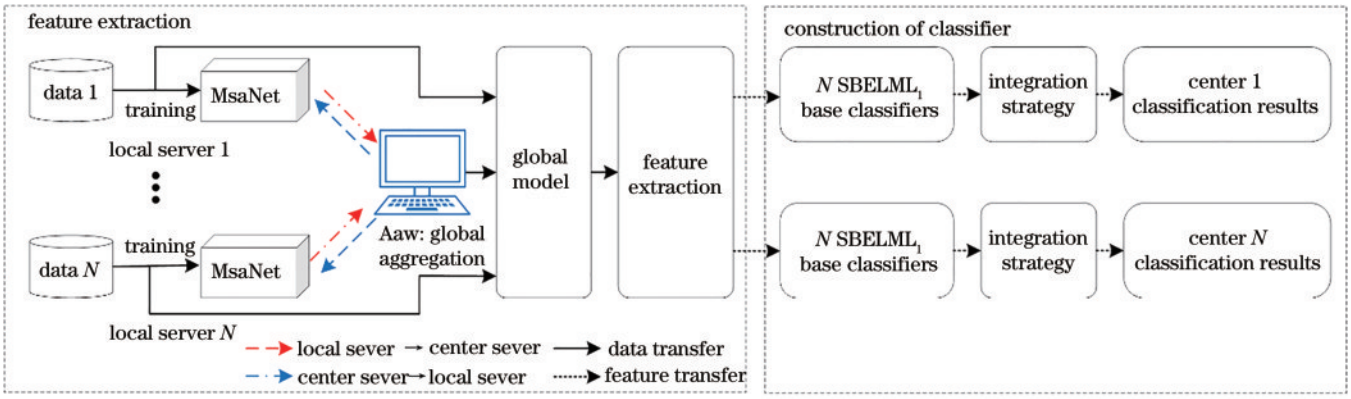

图 1　算法整体框图

Fig. 1　Overall algorithm block diagram

构建的集成分类器对各个中心的特征进行分类，并得到最终的预测结果。

2.1　基于 FedAaw 的特征提取

在中心服务器进行全局模型聚合时，因各个中心的数据量以及数据分布不同，所以各本地模型的训练速度和性能具有差异。在每轮训练时，一些性能较差的本地模型上传参数参与全局模型的聚合容易造成全局模型准确率下降，从而导致全局模型在训练过程中产生震荡。因此，在本地模型上传参数时引入限制阈值 Q，来选择在该轮训练中性能较好的一些本地模型，以缓解因部分本地模型性能不佳而造成全局模型泛化性能不佳的问题。

2.1.1　基于 Aaw 机制的全局模型聚合

Aaw 机制主要任务为决定本地模型参与全局模型的聚合、计算筛选后的本地模型参与全局模型聚合的权重。具体训练过程如图 2 所示。

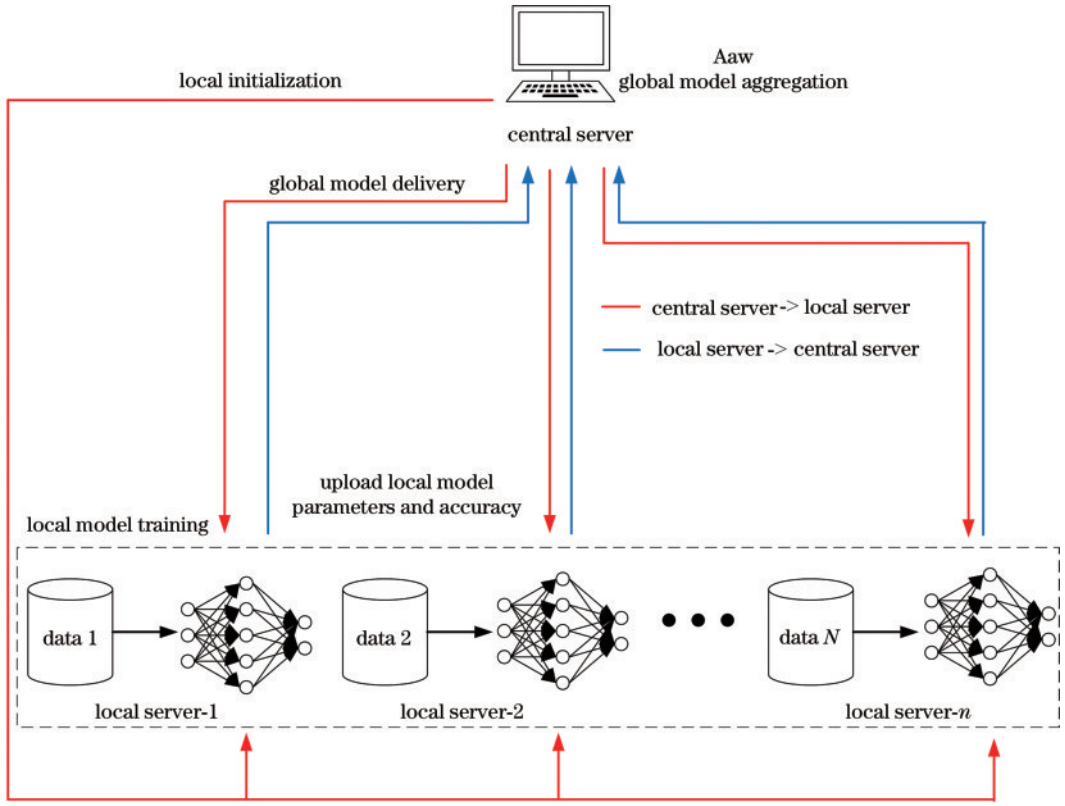

图 2　准确值阈值调度的联邦学习训练过程

Fig. 2　Federated learning training process of accurate value threshold scheduling

设本地服务器的集合为 $D_{\mathrm{All}}=\{D_1, D_2, \cdots, D_n\}$，$n$ 为本地服务器的个数。各本地服务器 D_i 通过中心服务器获取模型参数完成本地模型初始化，然后各本地服务器采用本地数据集进行训练。当第 t 轮训练完成后，各个本地服务器将该轮本地模型参数和准确率 $(\omega_i^t, f_{\mathrm{Acc}\,i}{}^t)$ 上传至中心服务器，并由中心服务器通过 Aaw 机制中预先设定好的准确率阈值 Q 来筛选出符合要求的本地服务器集合 D_{Accord}，之后对筛选后的本地

服务器集合 D_{Accord} 按照如下公式计算出集合中各本地服务器参与全局模型聚合的权重。

$$\phi_i^t=\frac{A_i^t}{\sum_{i=1}^{s}A_i^t},\tag{1}$$

式中：s 为 D_{Accord} 中模型的数量；A_i^t、ϕ_i^t 分别为第 i 个本地服务器第 t 轮的准确率和权重。由 Aaw 机制筛选出的本地模型的参数 ω_i^t 以及式(1)计算得到的对应权重 ϕ_i^t 通过加权平均的方式完成全局模型的聚合。

$$\omega^{t+1}=\omega^t-\sum_{i=1}^{s}\sum_{i=1}^{s}(\omega_i^t\times\phi_i^t),\tag{2}$$

式中：ω^t 为当前第 t 轮的全局模型参数。当本轮全局模型完成聚合后，中心服务器将全局模型参数下发至参与聚合的各个本地模型，其他未参与本轮聚合的本地模型使用该轮自身参数 ω_i^t，从而完成所有本地模型的更新，该过程可表示为

$$\begin{cases}\omega_i^{t+1}=\omega^t, D_i\in D_{\text{Accord}}\\ \omega_i^{t+1}=\omega_i^t, D_i\notin D_{\text{Accord}}\end{cases},\tag{3}$$

直到迭代次数达到预先要求，中心服务器完成全局模型的构建，并将最后全局模型参数下发至全部的本地服务器 D_{All}。此外，考虑到中央服务器和本地服务器之间的通信成本，采用本地训练 5 轮和中心服务器完成一次数据的交互。

2.2　基于 MHSA 机制的 CNN

基于 MHSA 机制的 CNN 本地模型和全局模型[16]，网络结构如图 3 所示。该网络将 MHSA 机制引入 CNN 中，通过浅层卷积层获取肺结节图像的低级语义特征，在深层采用 MHSA 机制关注与肺结节分类任务高度相关的抽象特征。MHSA 机制可建立肺结节的长距离特征之间的依赖关系，从而提高模型对病灶整体部分的关注度，以获取不同病灶部位的多尺度特征。

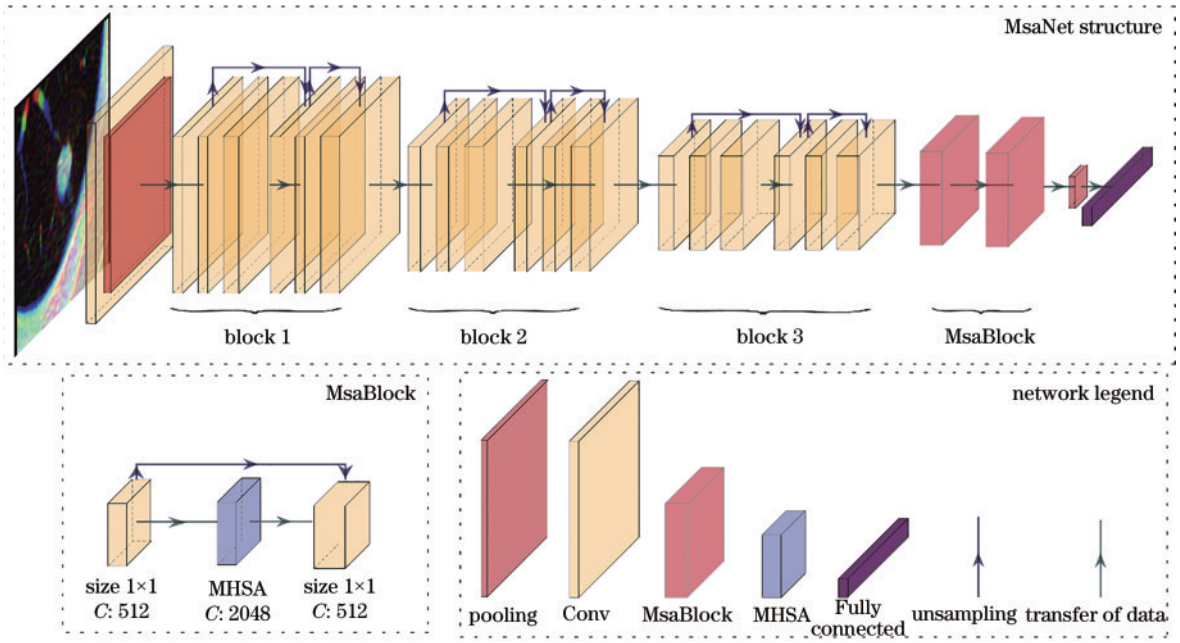

图 3　MsaNet 网络结构图

Fig. 3　Framework of MsaNet

2.2.1　基于 MHSA 的 Msa 块结构

Msa 块由两个 1×1 的卷积层和一个 MHSA 模块构成。1×1 的卷积层用来融合不同通道之间的信息和降低参数的计算量。MHSA 模块通过获取全局上下文信息、位置上下文信息、多尺度信息，以获取更稳定的肺结节特征。MHSA 结构如图 4 所示。

设 $X\in R^{H\times W\times D}$ 为输入图像，并初始化参数矩阵 $\boldsymbol{W}_h^{\mathrm{Q}}$、$\boldsymbol{W}_h^{\mathrm{K}}$、$\boldsymbol{W}_h^{\mathrm{V}}$。为建立 $\boldsymbol{X}$ 图像矩阵中较为孤立的各行向量之间的联系，通过 $\boldsymbol{X}$ 图像矩阵与 $\boldsymbol{W}_h^{\mathrm{Q}}$、$\boldsymbol{W}_h^{\mathrm{K}}$、$\boldsymbol{W}_h^{\mathrm{V}}$ 分别相乘获得索引矩阵 $\boldsymbol{q}\in R^{H\times W\times d}$、键矩阵 $\boldsymbol{k}\in R^{H\times W\times d}$、值矩阵 $\boldsymbol{v}\in R^{H\times W\times d}$ 等 3 个矩阵。

$$\begin{cases}\boldsymbol{q}=\boldsymbol{W}_h^{\mathrm{Q}}\\ \boldsymbol{k}=\boldsymbol{W}_h^{\mathrm{K}}\\ \boldsymbol{v}=\boldsymbol{W}_h^{\mathrm{V}}\end{cases},\tag{4}$$

式中：$\boldsymbol{W}_h^{\mathrm{Q}}$、$\boldsymbol{W}_h^{\mathrm{K}}$、$\boldsymbol{W}_h^{\mathrm{V}}$ 分别为 $\boldsymbol{X}$ 图像矩阵的第 h 个子空间的线性映射权重矩阵。

通过将 $\boldsymbol{q}\in R^{H\times W\times d}$ 与 $\boldsymbol{k}\in R^{H\times W\times d}$ 矩阵相乘，得到 $\boldsymbol{X}$ 图像矩阵中任意两点像素之间的长短距离依赖关系，从而捕捉到图像的全局信息。考虑到肺结节病灶部位与整体之间的关系，本文加入 2D 相对位置编码。通过将输入图像的长和宽进行对应通道的归一化，得到高为 $\boldsymbol{R}_{\mathrm{h}}\in R^{H\times 1\times d}$、宽为 $\boldsymbol{R}_{\mathrm{w}}\in R^{1\times W\times d}$，二者进行矩阵叠加融合得到图像的位置空间先验 $\boldsymbol{r}=\boldsymbol{R}_{\mathrm{h}}\oplus\boldsymbol{R}_{\mathrm{w}}$，

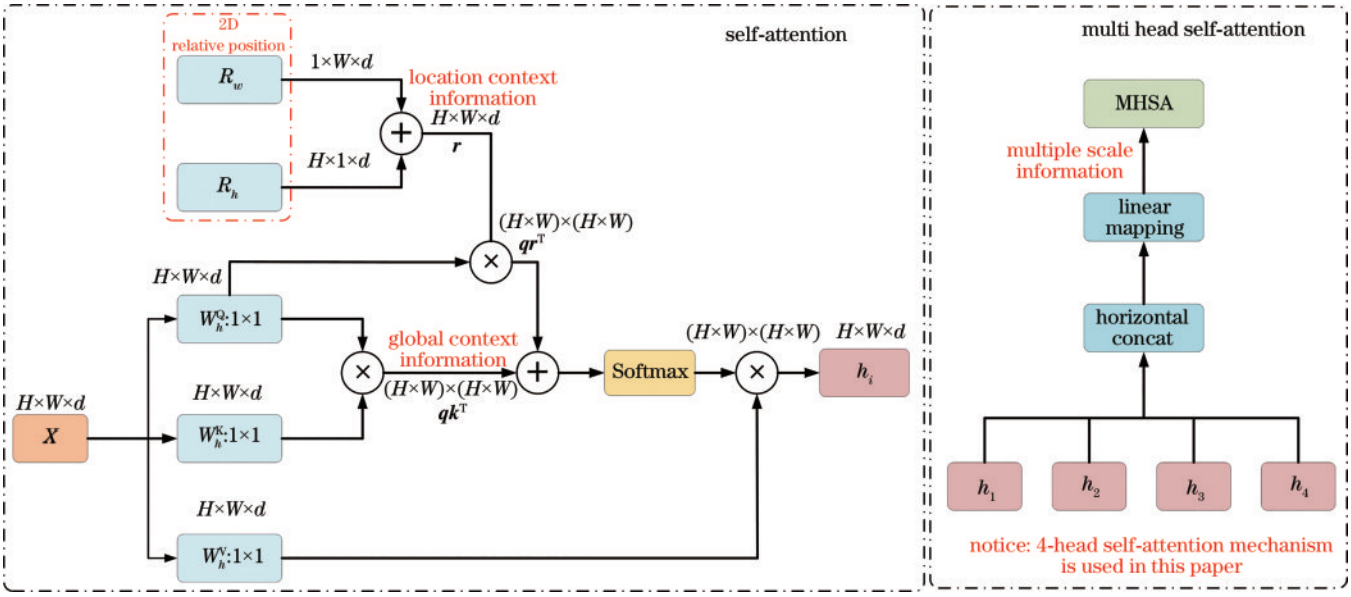

图 4　MHSA 结构图

Fig. 4　Framework of MHSA

$r \in R^{H\times W\times d}$。之后通过计算 $\boldsymbol{q}\otimes\boldsymbol{k}+\boldsymbol{q}\otimes\boldsymbol{r}$ 得到图像的空间敏感的相似性特征，从而提高模型关注肺结节的合适区域。当前图像的整体注意力得分 Attention 由空间敏感的相似性特征通过 Softmax 函数运算得到，Attention 反映当前图像中每个像素与整体图像之间的相关程度。将 Attention 矩阵与值矩阵 $\boldsymbol{v}$ 相乘，得到图像的单头自注意力特征 $\boldsymbol{h}$，$\boldsymbol{h}$ 可表示为

$$\boldsymbol{h}_i=\mathrm{Softmax}\left[\boldsymbol{q}\otimes\boldsymbol{k}^{\mathrm{T}}\oplus\boldsymbol{q}\otimes(\boldsymbol{R}_{\mathrm{w}}\oplus\boldsymbol{R}_{\mathrm{h}})^{\mathrm{T}}\right]\otimes\boldsymbol{v},\quad(5)$$

式中：⊗ 为矩阵相乘；⊕ 为矩阵相加。

因肺结节病灶部分特征表现形式具有多样性，采用 MHSA 机制能够提高模型提取特征多样性的能力。将其多个单头自注意力矩阵进行横向拼接进而获取图像的多尺度信息。MHSA 公式可表示为

$$\mathrm{MHSA}(\boldsymbol{X})=\mathrm{Concat}(h_1,h_2,\cdots,h_h)。\quad(6)$$

2.2.2　基于 CNN 的特征提取

为了充分利用 CNN 的特征，读取网络中所有卷积层滤波器生成的特征图，并对每张特征图求均值以得到对应滤波器的深度特征。然后，对所有的深度特征进行横向拼接，从而得到该图像在整个网络中的深度特征，在 MsaNet18 中共有 14400 个（ResNet18 中为 3904 个）卷积滤波器。肺结节的特征提取过程如图 5 所示。

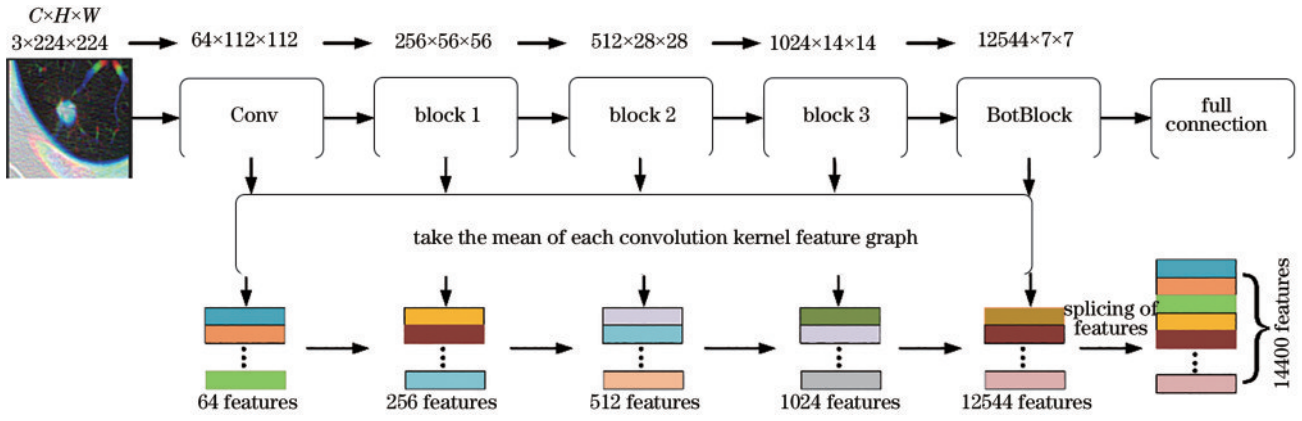

图 5　特征提取过程

Fig. 5　Feature extraction process

为了提高模型的预测性能以及获得更低的运算复杂度，需要筛选出对任务具有高度相关性的特征。本研究采用统计学的 Mann-Whitney U test（U 检验）方法对提取的深度特征进行筛选；使用最大相关最小冗余算法对 U 检验后的特征进行降维，以提高特征与任务之间的相关性和去除特征之间的冗余性，之后采用筛选后的深度特征进行分类。

2.3　基于 SBELM 集成分类器的构建

与单一分类器相比，采用集成学习方法构建分类器泛化能力更强，单一分类器的泛化能力有限，容易造

成训练集过拟合[17]。通过集成策略对多个弱分类器进行集成，构建一个强分类器提高模型的泛化能力，使用的基分类器为 $SBELML_1$。集成学习框图如图 6 所示。

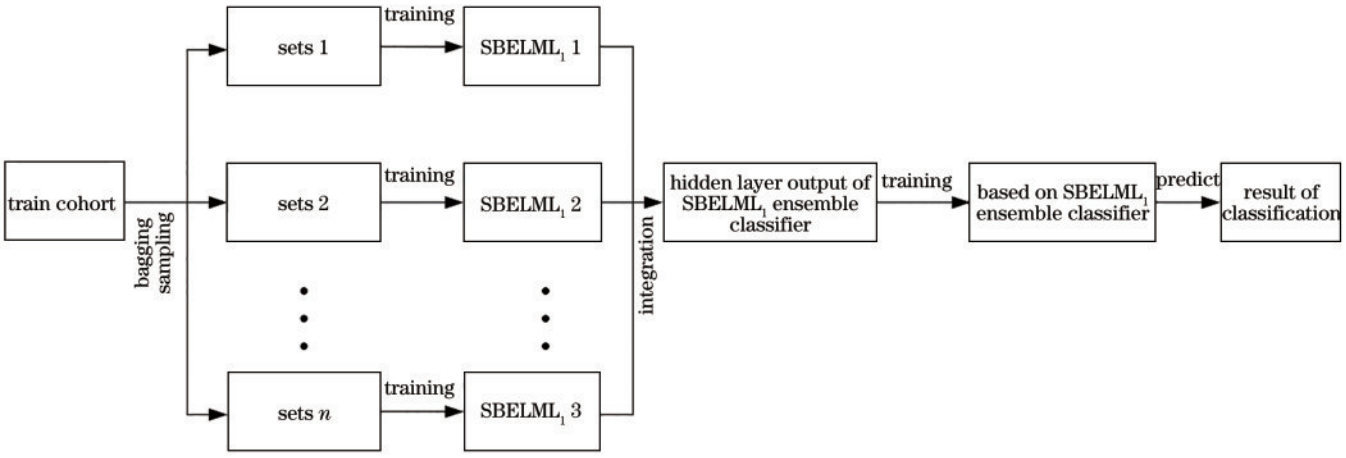

图 6　集成 $SBELML_1$ 分类器构建

Fig. 6　Integrated $SBELML_1$ classifier construction

2.3.1　$SBELML_1$ 算法

极限学习机(Extreme learning machine，ELM)是一种新型求解的单隐层前向传播算法，其主要特点是随机给定输入层和隐含层之间的连接权重 ω 和偏置 b，并且只需要设置网络结构就能得出相应的输出矩阵，具有较快的学习速度和更好的泛化能力[18-19]。利用 L_1 范数自动筛选最有价值特征的特性来防止基分类器过于复杂，导致模型过拟合问题，因此在 ELM 的优化求解中引入 L_1 范数对模型进行约束，使得模型具有稀疏解。

$$\arg\min E(\omega, b)=\|y-X\omega\|^2+\lambda\sum_{i-1}^{L}\|\omega_i\|_1, \quad (7)$$

式中：y 为样本的真实标签；λ 为 L_1 范数约束项系数，$\lambda>0$；ω 为隐藏层与输出层之间的权重；L 为隐藏层神经元个数；X 为隐藏层的输出。

式(7)的求解过程如下：设噪声符合逆方差为 β 的零均值高斯随机变量，y 为高斯噪声的线性组合。因此针对样本数据 (X, t)，ω 的似然函数可表示为

$$P(y|X, \omega, \beta)=N(y|X^{\mathrm{T}}\omega, \beta^{-1}), \quad (8)$$

为获得 ω 的后验概率，需引入关于 ω 的稀疏先验，稀疏先验可表示为

$$P(\omega|\alpha)=\prod_{k=1}^{L}N(\omega_i|0, \alpha_i^{-1}), \quad (9)$$

式中：$\alpha=[\alpha_1, \alpha_2, \cdots, \alpha_L]^{\mathrm{T}}$ 为每一个输出参数 ω_i 的先验独立参数。根据贝叶斯定理，可以求得 ω 的后验概率为

$$P(\omega|y, X, \alpha, \beta)=N(\omega|m, \Sigma), \quad (10)$$

且后验均值 m 及其协方差可表示为

$$\begin{cases} m=\beta\Sigma X^{\mathrm{T}}y \\ \Sigma^{-1}=\mathrm{diag}(\alpha)+\beta X^{\mathrm{T}}X \end{cases}, \quad (11)$$

式中：diag 为对角矩阵表达形式。α、β 利用边际极大似然估计方法求解，具体迭代公式可表示为

$$\begin{cases} a_i^{j+1}=\dfrac{1-a_i^j\Sigma_{ii}}{m_i^{\mathrm{T}}m_i} \\ (\beta^{j+1})^{-1}=\dfrac{\|y-Xm\|^2}{N-\sum\limits_{i=1}^{L}(1-a_k^j\Sigma_{kk})} \end{cases}, \quad (12)$$

式中：m_i 为后验均值 m 的第 i 个分量；Σ_{ii} 为后验分布协方差 Σ 的第 i 个对角线分量。

为使对数似然最大化，采用以下方式进行优化求解：给定 a 与 σ^2 的初始值，通过式(11)依次迭代求解得到高斯分布均值 m 和协方差 Σ。并判断 ω 的收敛性，若不收敛则根据式(12)更新参数 α 以及 β，并重新计算均值 m 和协方差 Σ；若收敛，则 $\omega=m$。因在求解过程中，α 的一些元素取向无穷导致对应的 ω 具有零后验分布，之后利用非 0 向量的输出 ω 构建模型并进行预测。

3　数据处理

采用随机划分数据集的方式，易造成训练集和测试集数据分布差异过大，从而影响模型真实泛化性能的评估[20-21]。因此，采用对抗验证对数据集进行划分，使训练集测试集模拟数据的真实分布并缓解两个集合之间分布的差异性，从而更加精确的评估模型的泛化性能。对抗算法划分数据集流程，如图 7 所示。

当前中心数据有 M 个病人数据样本记为 $\{X_1, X_2, \cdots, X_M\}$，取前 N 个病人数据作为训练集，剩下病人数据作为测试集。初步划分后的训练集为 $\{X_1, X_2, \cdots, X_N\}$，测试集为 $\{X_{N+1}, \cdots, X_M\}$，其中记训练集数据标签为正类，测试集标签极为负类。训练一个二分类判别器 $D(x)$ 可表示为

$$E_{x\sim p(x)}\{\lg[1-D(x)]\}+E_{x\sim q(x)}[\lg D(x)], \quad (13)$$

式中：$x\sim p(x)$ 为训练集的数据分布；$x\sim q(x)$ 为测试

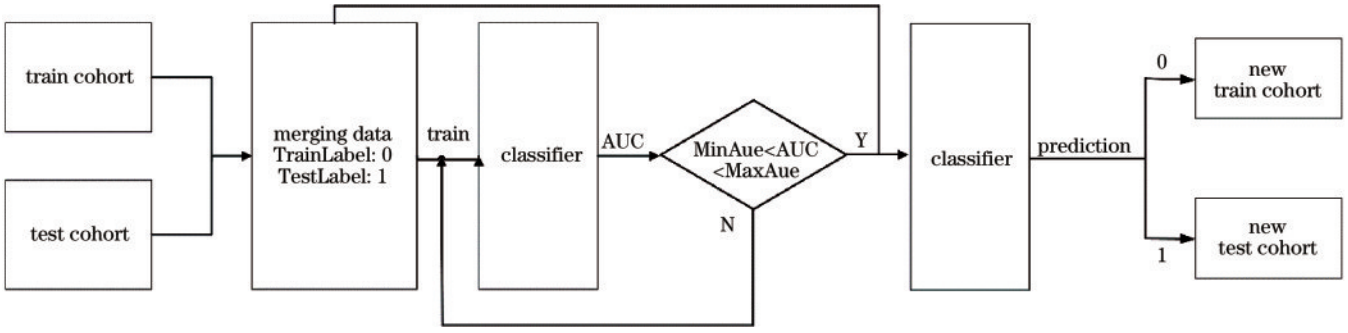

图 7　对抗验证算法流程图

Fig. 7　Flowchart of adversarial verification algorithm

集的数据分布。

因此，该优化方程的理论最优解为

$$D(x)=\frac{q(x)}{p(x)+q(x)}。\tag{14}$$

只有当$x\sim p(x)$与$x\sim q(x)$较为接近时，$D(x)$难以判断该病人数据来源，从而确认两个数据集的分布较为接近。

4　实验结果

4.1　实验数据

本研究使用的肺结节CT图像数据分别来自5个医院中心。对抗验证划分数据集的算法参数：分类器为支持向量机，优化算法使用带动量的随机梯度下降算法（Stochastic gradient descent algorithm，SGDM），训练轮次为120，学习率为0.001，动量为0.9，衰减因子为0.00001，受试者工作特征曲线（ROC）下的面积（Area under curve，AUC）划分范围为0.7～0.82。划分后数据集的结果如表1所示。（ROC是反映敏感性和特异性连续变量的综合指标）

表 1　对抗验证划分结果

Table 1　Division results of adversarial verification

Center	Train cohort		Test cohort	
	LTB case	LAC case	LTB case	LAC case
Center 1	105	155	86	156
Center 2	14	26	10	37
Center 3	32	120	25	76
Center 4	26	33	12	8
Center 5	2	37	3	46

4.2　FedAaw训练参数设置

本地模型在训练过程中，网络初始化学习率l_0=0.1，模型每轮输入的样本量为50张224×224×3的图片，训练轮次为120，优化器使用带动量的随机梯度下降（SGD），动量为0.9，衰减因子为0.0001，调整学习率使用余弦退火算法，余弦周期T_{max}=10，最小学习率l_{min}=0.001，周期调整倍数T_{mult}=2。损失函数采用交叉熵损失函数来反映目标标签和响应期望之间的损失。余弦退火算法可表示为

$$l_{new}=l_{min}+0.5\times(l_0-eta_{min})\times\left\{1+\cos\left\{\frac{e}{T_{max}\times\left[\left(e/T_{max}\right)\times T_{mult}\right]}\times\pi\right\}\right\},\tag{15}$$

式中：l_{new}为新的学习率；l_0为初始学习率；l_{min}为最小学习率；T_{max}为余弦周期的1/2；e为训练轮次。各个模型初始网络参数为在ImageNet数据集（1400万张自然图像包含1000个常见类的开放数据集）上训练的网络参数。此外，为保证中心服务器运算的同态性，各本地服务器采用相同模型和优化方法进行训练。

4.3　分类性能评价指标

采用的评价指标可以通过混淆矩阵计算得到，真阳性（True positive，T_P）表示实际为肺腺癌且被检测为肺腺癌的样本数量；假阴性（False negative，F_N）表示实际为肺腺癌且被检测为肺结核的样本数量（漏报）；假阳性（False positive，F_P）表示为实际为肺结核检测为肺腺癌的样本数量（误报）；真阴性（True negative，T_N）表示实际为肺结核检测为肺结核的样本数量。

AUC、准确率（$f_{Accuracy}$）、敏感度（$f_{Sensitivity}$）、特异度（$f_{Specificity}$）等评价指标评估分类结果可表示为

$$\begin{cases}f_{Sensitivity}=\dfrac{T_P}{T_P+F_N}\\ f_{Specificity}=\dfrac{T_N}{T_N+F_P}\\ f_{Accuracy}=\dfrac{T_P+T_N}{T_P+F_N+T_N+F_P}\end{cases}。\tag{16}$$

4.4　不同模型和Aaw机制参数的对比结果

在实验中，针对5个中心的数据，采用2种不同模型ResNet18和MsaNet18进行对比，并将联邦客户端准确率阈值Q分别设置为0、0.5、0.8。FL训练的损失曲线如图8所示。由图8可知，6种对比算法的模型均完成收敛，最终损失值均趋于0。在同一准确率阈值Q下，与MsaNet18相比，ResNet18的损失值均较大且收敛速度较慢。

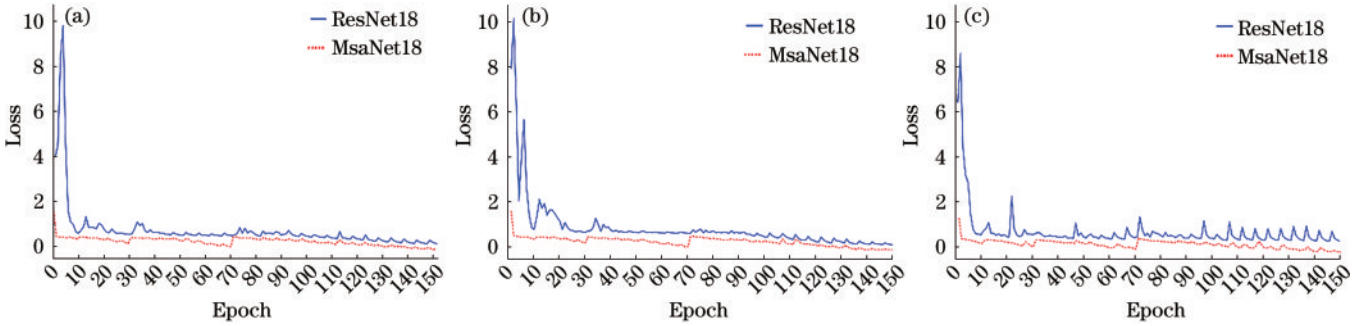

图 8　联邦损失曲线

Fig. 8　Federal loss curve

各个中心最后训练集和测试集的AUC性能指标如表2所示，测试集性能指标如表3～表5所示，AUC曲线如图9所示。与端对端方法相比，本文方法在测试集上性能均取得更好的结果；与ResNet18相比，本文使用的MsaNet18也在测试集上取得较大的提升。本文方法在5个中心测试集上的AUC分别为0.8037、0.8405、0.7758、0.9063、0.9348。

表2　不同模型和Aaw阈值AUC结果对比

Table 2　Comparison of AUC results of different models and Aaw threshold parameters

Center	Q=0				Q=0.5				Q=0.8			
	ResNet18		MsaNet18		ResNet18		MsaNet18		ResNet18		MsaNet18	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
Center 1	0.8987	0.7467	0.9393	0.7519	0.9030	0.7373	0.9459	0.7732	0.8704	0.7431	0.8938	0.8037
Center 2	0.9368	0.7622	0.9423	0.8135	0.9093	0.7919	0.9423	0.8351	0.9258	0.7784	0.9313	0.8405
Center 3	0.8341	0.6726	0.9154	0.7047	0.9276	0.7317	0.9464	0.7126	0.8638	0.7089	0.9526	0.7758
Center 4	0.8345	0.8229	0.8928	0.8542	0.9091	0.8229	0.8916	0.8542	0.8823	0.8333	0.9138	0.9063
Center 5	0.9054	0.8043	0.9459	0.8768	0.9189	0.8188	0.9189	0.8986	0.9324	0.8333	0.9459	0.9348

表3　Q=0时，MsaNet18的测试集性能指标

Table 3　When Q=0, MsaNet18 test set performance indicator

Parameter	Center 1	Center 2	Center 3	Center 4	Center 5	Sum
Sensitivity	0.7821(122/156)	0.6486(24/37)	0.4737(36/76)	0.7500(6/8)	0.8261(38/46)	0.6996(226/323)
Specificity	0.6628(57/86)	0.9000(9/10)	0.8800(22/25)	0.8333(10/12)	1.0000(3/3)	0.7426(101/136)
Accuracy	0.7397(179/242)	0.7021(33/47)	0.5743(58/101)	0.8000(16/20)	0.8367(41/49)	0.7124(327/459)

表4　Q=0.5时，MsaNet18的测试集性能指标

Table 4　When Q=0.5, MsaNet18 test set performance indicator

Parameter	Center 1	Center 2	Center 3	Center 4	Center 5	Sum
Sensitivity	0.8077(126/156)	0.8108(30/37)	0.6711(51/76)	0.8750(7/8)	0.8261(38/46)	0.7802(252/323)
Specificity	0.6279(54/86)	0.9000(9/10)	0.8400(21/25)	0.7500(9/12)	1.0000(3/3)	0.7059(96/136)
Accuracy	0.7438(180/242)	0.8298(39/47)	0.7129(72/101)	0.8000(16/20)	0.8367(41/49)	0.7581(348/459)

表5　Q=0.8时，MsaNet18的测试集性能指标

Table 5　When Q=0.8, MsaNet18 test set performance indicator

Parameter	Center 1	Center 2	Center 3	Center 4	Center 5	Sum
Sensitivity	0.8333(130/156)	0.7568(28/37)	0.6579(50/76)	1.0000(8/8)	0.9348(43/46)	0.8019(259/323)
Specificity	0.6395(55/86)	0.9000(9/10)	0.8000(20/25)	0.6667(8/12)	1.0000(3/3)	0.6985(95/136)
Accuracy	0.7645(185/242)	0.7872(37/47)	0.6931(70/101)	0.8000(16/20)	0.9388(46/49)	0.7712(354/459)

由表2可知，在同一阈值Q下，MsaNet18网络的分类效果AUC均好于ResNet18网络，同时随着准确率阈值Q的提高，5个中心测试集的AUC也随之提高；由表3～表5可知，随着阈值Q的增加，模型的分类准确率和灵敏度的整体数值逐步增加，特异度的整体数值则出现了一定程度的下滑。其中，灵敏度的提高

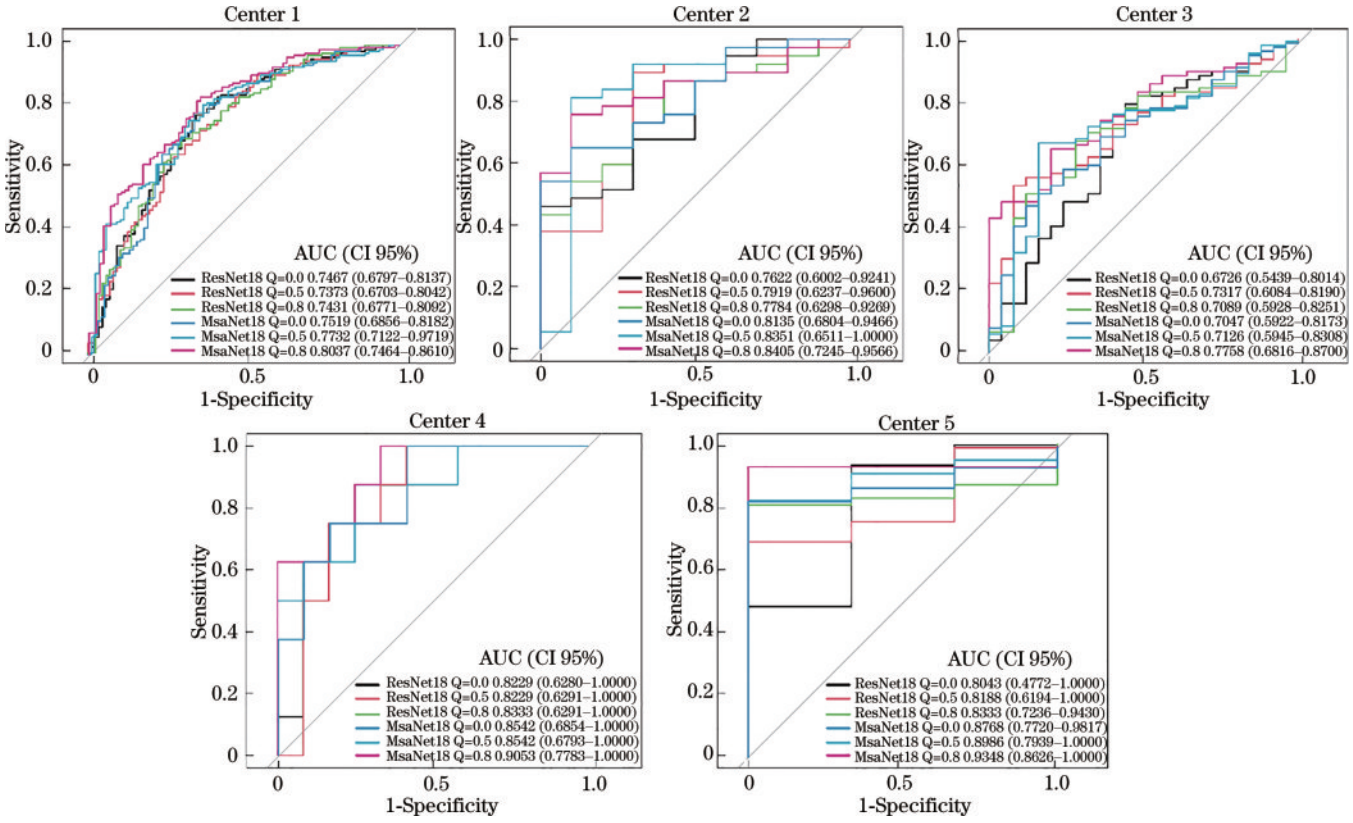

图 9　各中心测试集 AUC 曲线

Fig. 9　AUC curve of each center test cohort

代表模型针对肺腺癌的鉴别能力提高，准确率的提高代表模型对疾病的检出率提高，特异度的下降代表模型对肺结核的鉴别能力下降。然而，在临床中，肺腺癌相较于肺结核而言，肺腺癌形成的恶性肿瘤会出现侵犯周边组织的情况，出现误诊更容易导致患者错失最佳治疗时间，所以肺腺癌和肺结核出现的漏诊造成损失是不同的。因此，实际应用中更加倾向于选择肺腺癌鉴别能力较强的模型。

4.5　对比试验

表 6、表 7 将所提的 FL 算法与其他学者近期研究提出的 FL 算法进行性能比较，其中包括了 FedAvg[7]、FedProx[8]、Moon[9]、HarmoFL[10]算法。为确保实验的一致性，在进行联邦实验对比时，统一采用端对端的实验结果，本地模型及全局模型均为 MsaNet18。因为 5 个中心数据存在着异质性，导致 FedAvg 算法在全局模型聚合后，较易出现多中心数据欠拟合的现象。在 FedProx 算法中，虽然加快了模型的拟合速度，但因为不同中心数据样本的差距，导致全局模型对不同中心数据产生了偏移。本文算法与 Moon 和 HarmoFL 算法相比，在 5 个中心性能上的表现均有明显提升。

表 6　不同联邦算法在肺结节分类任务上训练集的 AUC 结果

Table 6　AUC results of different federal algorithms on lung nodule classification tasks

Algorithm	Center 1	Center 2	Center 3	Center 4	Center 5
FedAvg	0.6077	0.8159	0.5849	0.5303	0.6622
FedProx	0.8228	0.9890	0.8232	0.9930	0.9189
Moon	0.6371	0.8709	0.6789	0.7587	0.7027
HarmoFL	0.7799	0.9780	0.5727	0.9405	0.6216
FedAaw	0.9963	0.9988	0.9375	0.6395	0.7605

表 7　不同联邦算法在肺结节分类任务上测试集的 AUC 结果

Table 7　AUC results of different federal algorithms on lung nodule classification tasks

Algorithm	Center 1	Center 2	Center 3	Center 4	Center 5
FedAvg	0.6473	0.7351	0.5932	0.4479	0.5870
FedProx	0.6506	0.7297	0.3758	0.3333	0.2899
Moon	0.6491	0.6568	0.5142	0.5652	0.7101
HarmoFL	0.6520	0.7703	0.2937	0.5313	0.3551
FedAaw	0.6943	0.7013	0.5591	0.6617	0.8478

4.6　联邦抗干扰实验

为验证联邦学习框架与分布式学习两种方式对数据分布的敏感程度。随机调换 3 次 5 个中心数据分布来验证两个学习框架的抗干扰性能。数据调换后正负样本的分布如图 10(a)所示。联邦学习和分布式学习对 3 次数据分布调换后的抗干扰实验结果如图 10(b)所示。

由图 10(b)可知，在两种方法的对比和测试中，联邦学习在各个中心测试集上的 AUC 均大于分布式学习。同时根据箱线图的上下四分位点和均值，发现联

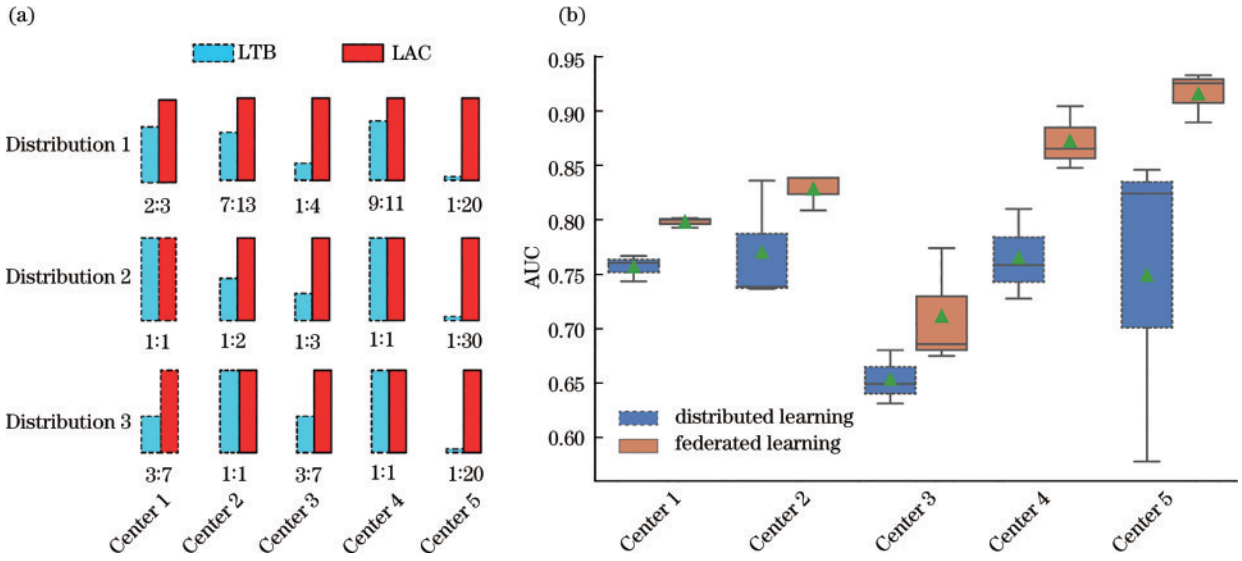

图 10　抗干扰实验数据分布。(a)实验数据比例;(b)Fed-Aaw抗干扰实验

Fig. 10　Distribution of anti-interference experimental data. (a) Proportion of positive and negative sample data; (b) FedAaw anti-interference experiment

邦学习对各个中心正负样本的分布敏感程度低于本地训练,进而反应出本文联邦学习的框架抗干扰能力强于本地训练。

采用的联邦学习经过3次数据调换后各中心测试集的AUC均值为0.8001、0.8305、0.7134、0.8744、0.9179,相对于传统的分布式学习的结果为0.7588、0.7723、0.6554、0.7672、0.7519,各中心具有较大的提升,具体实验结果如表8、表9所示。通过实验表明,FedAaw受到各个中心数据中正负样本的比例差异的干扰较小,同时也进一步证明使用的FedAaw相对传统的分布式学习具有更强的鲁棒性和模型泛化能力。

表8　联邦抗干扰训练集结果

Table 8　Results of federal anti-interference experimental training set

Distribution	Method	Center 1	Center 2	Center 3	Center 4	Center 5
Distribution 1	Distributed learning	0.8963	0.9148	0.9268	0.8858	0.8919
	Federated learning	0.8938	0.9313	0.9526	0.9138	0.9459
Distribution 2	Distributed learning	0.8682	0.9286	0.8770	0.8789	0.8917
	Federated learning	0.8816	0.9405	0.9200	0.9368	0.9417
Distribution 3	Distributed learning	0.8323	0.9306	0.9286	0.8632	0.9000
	Federated learning	0.9098	0.8264	0.8262	0.9053	0.9000

表9　联邦抗干扰测试集结果

Table 9　Results of federal anti-interference experimental test set

Distribution	Method	Center 1	Center 2	Center 3	Center 4	Center 5
Distribution 1	Distributed learning	0.7686	0.8378	0.6821	0.7604	0.8261
	Federated learning	0.8037	0.8405	0.7758	0.9063	0.9348
Distribution 2	Distributed learning	0.7453	0.7405	0.6331	0.7293	0.5797
	Federated learning	0.7947	0.8405	0.6768	0.8672	0.9275
Distribution 3	Distributed learning	0.7624	0.7386	0.6511	0.8120	0.8478
	Federated learning	0.8020	0.8105	0.6875	0.8496	0.8913

5　结　　论

针对目前医学影像面临多中心数据存在数据孤岛以及非独立同分布的问题,提出一种基于自适应聚合权重的联邦学习算法FedAaw。首先,中心服务器通过阈值Q筛选出每轮参与聚合的本地模型;其次,通过准确率计算相应的聚合权重来完成全局模型的自适应聚合;然后,为丰富模型提取多尺度信息和长短距离依赖的能力,在ResNet18网络中加入Transformer结构;最后,为缓解端对端模型因冗余特征导致的过拟合问

题，通过提取全局模型卷积核的特征，并使用基于L_1范数稀疏贝叶斯的集成学习完成多中心数据特征的分类。3项实验表明，所提方法具有较好的预测精度、泛化能力以及抗干扰能力。

参 考 文 献

[1] Travis W D, Brambilla E, Noguchi M, et al. International association for the study of lung cancer/American thoracic society/European respiratory society international multidisciplinary classification of lung adenocarcinoma: executive summary[J]. Journal of Thoracic Oncology, 2011, 8(5): 381-385.

[2] Müller M, Gromicho M, de Carvalho M, et al. Explainable models of disease progression in ALS: learning from longitudinal clinical data with recurrent neural networks and deep model explanation[J]. Computer Methods and Programs in Biomedicine Update, 2021, 1: 100018.

[3] 邸拴虎, 杨文瀚, 廖苗, 等. 基于RA-Unet的CT图像肝脏肿瘤分割[J]. 仪器仪表学报, 2022, 43(8): 65-72.
Di S H, Yang W H, Liao M, et al. Liver tumor segmentation from CT images based on RA-Unet[J]. Chinese Journal of Scientific Instrument, 2022, 43(8): 65-72.

[4] 何峰峰, 张强, 杨超, 等. 基于深度学习的图像分割方法在主动脉疾病中的应用[J]. 临床心血管病杂志, 2022, 38(6): 449-454.
He F F, Zhang Q, Yang C, et al. Application of image segmentation methods based on deep learning in aortic diseases[J]. Journal of Clinical Cardiology, 2022, 38(6): 449-454.

[5] Liu Q D, Chen C, Qin J, et al. FedDG: federated domain generalization on medical image segmentation via episodic learning in continuous frequency space[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 1013-1023.

[6] Sheller M J, Reina G A, Edwards B, et al. Multi-institutional deep learning modeling without sharing patient data: a feasibility study on brain tumor segmentation[M]// Crimi A, Bakas S, Kuijf H, et al. Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries. Lecture notes in computer science. Cham: Springer, 2019, 11383: 92-104.

[7] McMahan H B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[EB/OL]. (2016-02-17) [2022-11-12]. https://arxiv.org/abs/1602.05629.

[8] Sahu A K, Li T, Sanjabi M, et al. Federated optimization in heterogeneous networks[EB/OL]. [2022-11-12]. https://arxiv.org/pdf/1812.06127.pdf.

[9] Li Q B, He B S, Song D. Model-contrastive federated learning[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA. New York: IEEE Press, 2021: 10708-10717.

[10] Jiang M R, Wang Z R, Dou Q. HarmoFL: harmonizing local and global drifts in federated learning on heterogeneous medical images[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(1): 1087-1095.

[11] Hsieh K, Phanishayee A, Mutlu O, et al. The non-IID data quagmire of decentralized machine learning[EB/OL]. (2019-10-01) [2022-11-12]. https://arxiv. org/abs/1910.00189.

[12] 张泽辉, 李庆丹, 富瑶, 等. 面向非独立同分布数据的自适应联邦深度学习算法[J/OL]. 自动化学报: 1-13 [2022-11-12]. https://doi.org/10.16383/j.aas.c201018.
Zhang Z H, Li Q D, Fu Y, et al. Adaptive federated deep learning algorithm for Non-IID data[J/OL]. Acta Automatica Sinica: 1-13[2022-11-12]. https://doi. org/10.16383/j.aas.c201018.

[13] 陈辉, 张甜, 陈润斌. 基于轻量级卷积Transformer的图像分类方法及在遥感图像分类中的应用[J/OL]. 电子与信息学报, 2022: 1-9[2022-07-07]. https://kns.cnki.net/kcms/detail/11.4494.TN.20220705.1638.014.html.
Chen H, Zhang T, Chen R B. Image classification method based on lightweight convolutional transformer and its application in remote sensing image classification [J/OL]. Journal of Electronics & Information Technology, 2022: 1-9[2022-07-07]. https://kns. cnki. net/kcms/detail/11.4494.TN.20220705.1638.014.html.

[14] McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[EB/OL]. (2016-02-17) [2022-11-12]. https://arxiv.org/abs/1602.05629.

[15] Schlemper J, Oktay O, Schaap M, et al. Attention gated networks: learning to leverage salient regions in medical images[J]. Medical Image Analysis, 2019, 53: 197-207.

[16] Srinivas A, Lin T Y, Parmar N, et al. Bottleneck transformers for visual recognition[EB/OL]. (2021-01-27) [2022-11-08]. https://arxiv.org/abs/2101.11605.

[17] Yang M, Zhang Y X, Wang X Z, et al. Multi-instance ensemble learning with discriminative bags[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(9): 5456-5467.

[18] Khellal A, Ma H B, Fei Q. Convolutional neural network based on extreme learning machine for maritime ships recognition in infrared images[J]. Sensors, 2018, 18 (5): 1490.

[19] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: theory and applications[J]. Neurocomputing, 2006, 70(1/2/3): 489-501.

[20] Shafahi A, Najibi M, Ghiasi A, et al. Adversarial training for free! [EB/OL]. (2019-04-29) [2022-11-09]. https://arxiv.org/abs/1904.12843.

[21] Pan J, Pham V, Dorairaj M, et al. Adversarial validation approach to concept drift problem in user targeting automation systems at uber[EB/OL]. (2020-04-07)[2022-11-09]. https://arxiv.org/abs/2004.03045.