

文章编号: 2095-4980(2024)02-0142-10

## 基于动态权重模型的数据不平衡SEI方法

段可欣<sup>1,2</sup>, 闫文君<sup>\*1</sup>, 刘凯<sup>1</sup>, 张建廷<sup>3</sup>, 李春雷<sup>4</sup>, 王艺卉<sup>1,5</sup>

(1. 海军航空大学 信息融合研究所, 山东 烟台 264001; 2. 91422 部队, 山东 烟台 265200; 3. 海军研究院, 北京 100071; 4. 92038 部队, 山东 青岛 266109; 5. 31401 部队, 山东 烟台 264099)

**摘要:** 针对辐射源个体识别(SEI)中个体数据分布不平衡导致的识别准确率下降的问题, 提出一种基于动态权重模型的SEI方法。通过设计一个动态类权重(DCW)模型, 先利用元学习算法使用少量样本数据通过2层计算得到一个适中的权重初始值; 再设计一种新的代价敏感损失函数计算预测值与真实值之间的距离, 反向调整赋予少数类的学习权重, 适度增加对少数类数据的重视程度。对少数类更加友好, 对高度不平衡数据的处理有明显优势, 缓解多数类样本对整个识别过程的计算误导, 提高整体的识别正确率。

**关键词:** 辐射源个体识别; 不平衡数据; 动态类权重; 元学习; 代价敏感损失

中图分类号: TN974

文献标志码: A

doi: 10.11805/TKYDA2023181

## Data imbalance SEI method based on dynamic weight model

DUAN Kexin<sup>1,2</sup>, YAN Wenjun<sup>\*1</sup>, LIU Kai<sup>1</sup>, ZHANG Jianting<sup>3</sup>, LI Chunlei<sup>4</sup>, WANG Yihui<sup>1,5</sup>

(1. Institute of Information Fusion, Naval Aviation University, Yantai Shandong 264001; 2. Unit 91422 of Chinese People's Liberation Army, Yantai Shandong 265200, China; 3. Navy Research Institute of People's Liberation Army, Beijing 100071, China; 4. Unit 92038 of Chinese People's Liberation Army, Qingdao Shandong 266109, China; 5. Unit 31401 of Chinese People's Liberation Army, Yantai Shandong 264099, China)

**Abstract:** To tackle with the problem of decreased recognition accuracy caused by imbalanced individual data distribution in Specific Emitter Identification(SEI), a dynamic weight model based method is proposed for individual identification of radiation sources. A Dynamic Class Weight(DCW) model is built. A moderate initial weight value is obtained by using a meta learning algorithm through two-layer calculation with a small amount of sample data. Then, a new cost sensitive loss function is designed to calculate the backward adjustment of the distance between the predicted value and the true value, which gives the minority learning weight, and moderately increases the attention to the minority data. It is more friendly to the minority. It has obvious advantages in the processing of highly unbalanced data, which alleviates the calculation misleading of the majority of samples in the whole recognition process, thus improving the overall recognition accuracy.

**Keywords:** Specific Emitter Identification; unbalanced data; Dynamic Class Weights; meta learning; cost sensitive losses

科技的快速发展带动了数据爆炸性增长, 带来了大量的数据资源, 这是巨大的机遇; 同时越来越复杂的数据结构也带来了许多挑战, 其中一个很值得注意的方面是数据不平衡现象。在雷达辐射源识别中同样存在这样的问题, 标准分类器在处理不平衡数据时, 往往会将少数类样本错误地划分为多数类样本, 导致分类性能非常差。大多数分类器都是在假设数据平衡且均匀分布在每个类上的情况下设计的, 这些分类器从不平衡的数据集中学习显著损害标准学习算法的性能, 当遇到复杂的不平衡数据集时, 标准算法往往无法正确地表示数据的正确分布特征。

目前已有不少专家学者对此进行了研究, 提出的数据不平衡处理办法大致可分为数据级、算法级和混合形

收稿日期: 2023-06-28; 修回日期: 2023-09-25

基金项目: 国家自然科学基金面上资助项目(62271499; 62371465); 电磁空间安全全国重点实验室开放基金资助项目

\*通信作者: 闫文君 email:1346859039@qq.com

式等3类<sup>[1-2]</sup>。在数据级层面，通过采样技术<sup>[3-6]</sup>，将不平衡数据转换为平衡的数据以提供平衡的分布，然后使用分类算法获得结果。在算法级层面，通过设计分类算法处理数据不平衡问题，采用成本敏感方法<sup>[7-9]</sup>为每类数据分配不同的权重，以不同的错误分类成本降低总体成本；或用集成的方法同时从多个分类器中学习，每个分类器学习一类数据的特征，减少多数类数据对少数类数据的误导。混合形式是将采样技术与分类算法相结合，产生更好的方案解决不平衡问题。在现实世界收集的数据中，少数类具有丰富的实例，而大多数类具有很少的实例。对于这个问题，一个简单的解决方案是删除少数类，使类分布平衡，但这种方式会削弱训练数据的部分重要特征。为保证数据集特征的完整度，Nitesh等<sup>[10]</sup>选择在训练期间使用基于类的采样方法，以相同的概率对每个类进行采样，但这个方法仅降低了计算量，对于少数类没有产生偏斜，依然会出现大量少数类被误判为多数类的情况，对识别准确率贡献不大；Zhong等<sup>[11]</sup>提出一种两阶段训练机制，先用多数类进行训练，再用多数类和少数类对已有的模型进行微调，这对识别效果有一定的提升，但第2阶段需要较为复杂的方法对参数进行调整才能获得较好的识别性能，整个工作过程较为繁琐且效果具有一定的随机性。

在数据不平衡问题中，多数类的数据量是丰富<sup>[12-13]</sup>的，而少数类只有有限的数据量。当样本有限时，神经网络对挖掘到的隐式特征所形成的描述，可能会变得过于具体，导致过拟合。目前的一些采样方法对于这些不平衡数据有一定的平衡作用，但对于高度不平衡的数据集，平衡作用有限，且当采样的规则与数据集本身的特点融合度不够时，可能会造成采样策略失效。针对这些问题，本文采用不平衡比(Imbalanced Rate, IR)<sup>[14]</sup>标准定义高度不平衡数据集，提出一种动态类权重(DCW)方法，根据数据集的不平衡分布情况灵活分配权重，使训练效果得到较好的提升；同时还提出一种包含数据集不平衡度的损失函数，弥补由于数据不平衡带来的识别效果变差问题。

## 1 相关工作

### 1.1 辐射源信号双谱分析

雷达辐射源识别不可避免地要对信号进行特征提取，高阶谱<sup>[15]</sup>以其独特的抑制高斯噪声的优势广泛用于信号分析和处理。其中双谱因其计算量和复杂度较为均衡，应用较为广泛，双谱分析对信号的幅度保持性和相位保持性显现出信号的非线性特性。双谱的概念如下所述。

假定高阶累积量 $c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1})$ 是绝对可和的，即

$$\sum_{\tau_1}^{\infty} \dots \sum_{\tau_{k-1}}^{\infty} |c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1})| < \infty \quad (1)$$

则 $k$ 阶谱定义为 $k$ 阶累积量的 $(k-1)$ 阶 Fourier 变换，即有：

$$S_{kx}(\omega_1, \omega_2, \dots, \omega_{k-1}) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} \dots \sum_{\tau_{k-1}=-\infty}^{\infty} c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1}) e^{-j(\omega_1 \tau_1 + \omega_{k-1} \tau_{k-1})} \quad (2)$$

则双谱可定义为：

$$B_x(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} c_{3x}(\tau_1, \tau_2) e^{-j(\omega_1 \tau_1 + \omega_2 \tau_2)} \quad (3)$$

可以得出接收机接收到的离散含相噪信号为：

$$x(n) = s(n) + w(n) \quad (4)$$

式中： $w(n)$ 为高斯噪声信号； $s(n)$ 为发射机输出的含有非高斯噪声的信号。 $w(n)$ 和 $s(n)$ 相互独立。

对 $x(n)$ 求三阶累积量：

$$c_{3x}(\tau_1, \tau_2) = E \left\{ [s(n) + w(n)][s(n + \tau_1) + w(n + \tau_1)][s(n + \tau_2) + w(n + \tau_2)] \right\} \quad (5)$$

式中 $E$ 为期望。将式(5)展开，然后合并为：

$$c_{3x}(\tau_1, \tau_2) = c_{3s}(\tau_1, \tau_2) + c_{3w}(\tau_1, \tau_2) + E[w(n)] [c_{2x}(\tau_1) + c_{2x}(\tau_2) + c_{2x}(\tau_2 - \tau_1)] E[s(n)] [c_{2w}(\tau_1) + c_{2w}(\tau_2) + c_{2w}(\tau_2 - \tau_1)] \quad (6)$$

因此只要信号和噪声的均值为零，则有

$$c_{3x}(\tau_1, \tau_2) = c_{3s}(\tau_1, \tau_2) + c_{3w}(\tau_1, \tau_2) \quad (7)$$

若  $w(n)$  为高斯有色噪声, 则  $c_{3\omega}(\tau_1, \tau_2)$  可以忽略不计。可见信号的双谱可以消除高斯有色噪声对信号的影响, 其双谱由  $c_{3s}(\tau_1, \tau_2)$  确定, 即

$$c_{3s}(\tau_1, \tau_2) = c_{3s}(\tau_1, \tau_2) = E[s(n)s(n+\tau_1)s(n+\tau_2)] \quad (8)$$

由上述分析可知双谱分析可滤除高斯噪声的影响, 从而最大化保留辐射源信号个体特征。由于双谱具有对称性和周期性, 考虑到计算量, 选择双谱对角线作为显式特征参与后面的识别过程。

### 1.2 元学习生成权重初始值

参数初始化是模型训练前比较重要的一步, 主要是因为其可以加快梯度下降收敛的速度, 并尽量使其收敛于全局最优。权重参数初始化的好坏会影响模型最后的识别性能, 初始化参数不能太大或太小, 参数太小, 会导致特征在每层间逐渐缩小而难以产生作用; 参数太大, 会导致数据在逐层间传递时逐渐放大, 导致梯度消失发散。元学习<sup>[16-17]</sup>专注于提升模型整体的学习能力, 训练时, 不停地在不同的任务上切换, 从而达到初始化网络参数的目的, 最终得到的模型在面对新任务时可以表现出较好的性能。

元学习又称为学会学习, 通过学出一个较好的权重初始化参数, 帮助模型在新的任务上进行学习。随着对每一项历史任务的学习和经验积累, 元学习在学习新任务时变得更加容易, 所需的训练样本更少, 同时能保证一定的算法精确度。区别于以映射方式为主的深度学习和以试错方式为主的强化学习, 元学习旨在学会如何学习。每当模型尝试学习某项任务时, 无论成功与否, 模型都会获得有用的经验, 整合这些经验形成智能体的“价值观”, 代表一种会学习的能力, 即抽象成函数  $F(x)$ 。若出现新的任务, 在“价值观”的协助下, 模型继续学习新任务的极少量样本, 即可快速适应和掌握新任务, 也即抽象出一个个对应新任务  $i$  的函数  $f_i(x)$ 。该算法可直接用于任何梯度下降过程的网络学习训练模型, 具有良好的泛化性, 适用于各种不同的学习问题, 能够使用少量训练数据解决新的学习任务, 这与想通过学习少量样本获取少数类更多特征需求正相符。根据对元学习的理解, 建立适合本文数据结构的学习模型, 其算法结构如图 1 所示。

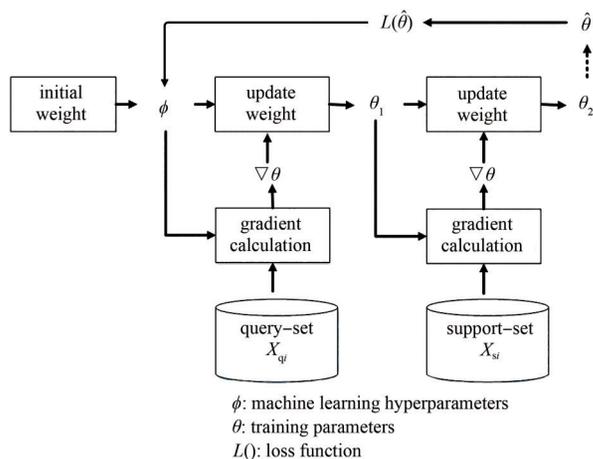


Fig.1 Diagram of meta learning algorithm  
图1 元学习算法结构图

表 1 MAML 算法流程

Table 1 Process of MAML algorithm

algorithm: Model-Agnostic Meta-Learning
require: $p(T)$ : distribution over tasks
require: $\alpha, \beta$ : step size hyperparameters
1: randomly initialize $\theta$
2: while not done do
3: sample batch of tasks $T_i \sim p(T)$
4: for all $T_i$ do
5: evaluate $\nabla_{\theta} L_{T_i}(f_{\theta})$ with respect to $K$ examples
6: compute adapted parameters with gradient descent: $\theta'_i = \theta - \alpha \nabla_{\theta} L_{T_i}(f_{\theta})$
7: end for
8: Update $\theta \leftarrow \theta - \beta \nabla_{\theta} \sum_{T_i \sim p(T)} L_{T_i}(f_{\theta})$
9: end while

元学习比迁移学习对未知新任务的适应有着更好的稳健性, 元学习中比较经典是 Finn<sup>[18]</sup>等提出的模型不可知元学习算法(Model-Agnostic Meta-Learning, MAML), 算法流程如表 1 所示。

在元学习算法中, 通过选取少量样本送进权重生成器, 在少量的学习中获取合适的初始化权重  $w_i$ , 再将  $w_i$  送入处理样本数据的神经网络中赋予不同的样本。

### 1.3 动态权重模型

在高度不平衡数据集中, 多数类在进行数据学习分析中的主导作用更强, 识别过程会更偏向多数类。在深度神经网络训练过程中, 对多数类的训练包含了大部分损失函数和主导梯度, 标准交叉熵损失函数已不足以弥补由于巨大的数据量差异带来的梯度差。因此引入一种动态的思想, 通过动态地分配权重, 提高对少数类的重视程度, 同时不过分减小多数类的存在率, 使神经网络能兼顾少数类与多数类的特点, 获取各自的隐式特征并保存在模型中, 提高对不平衡数据集的整体识别正确率。动态获取权重的过程如图 2 所示。

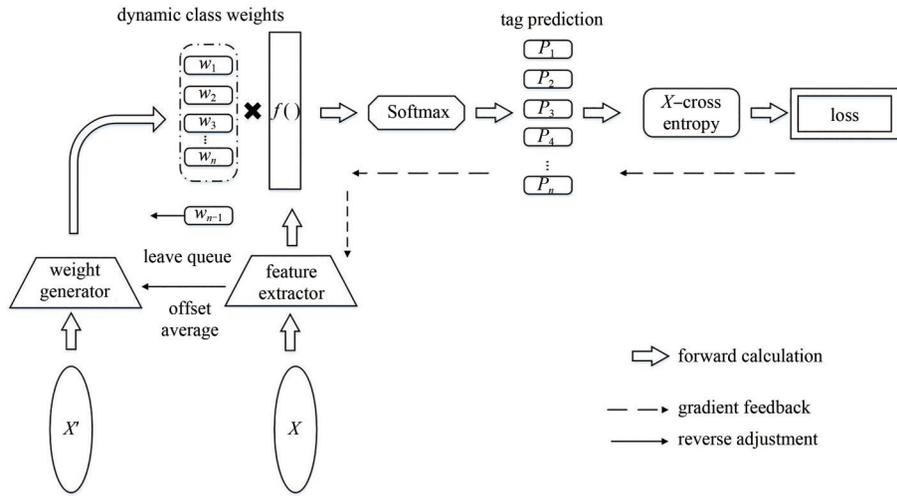


Fig.2 Overall structure of dynamic class weights  
图2 动态类权重整体结构

在权重生成器部分，随机选取少数类样本  $X$  中的一小部分  $X'$  作为元学习的训练样本。对于  $N$  个少数类数据集  $X' = \{X'_1, X'_2, \dots, X'_N\}$ ，取每个  $X'_i$  的子集  $X'_{si}$ 、 $X'_{qi}$ ，其中  $X'_{si}$  作为支持集，每个支持集有  $K$  条数据； $X'_{qi}$  作为查询集，包含  $q$  条未分类数据。在第一阶段，用  $X'_{si}$  中的数据训练神经网络，进行 1 次梯度下降，每个  $X'_{si}$  通过反馈机制对初始参数  $\theta_0$  进行 1 次更新得到  $\theta'_i$ ，即

$$\theta'_i \leftarrow \phi - \alpha \partial l(\phi) / \partial \phi \tag{9}$$

用 1 次更新后得到的  $\theta'_i$  测试在查询集  $X'_{qi}$  上的效果，获得损失函数  $l(\theta'_i)$ ，对  $X'_{qi}$  上的  $l(\theta'_i)$  求均值，得到子集  $X'_i$  的损失函数均值为：

$$l_i(\phi) = \frac{1}{Q} \sum_{i=1}^Q l_i(\theta'_i) \tag{10}$$

式中  $Q$  为  $Q$  条数据。对每个子集的查询集计算损失函数均值求和，得到总损失函数：

$$L(\phi) = \sum_{i=1}^N l_i(\theta'_i) \tag{11}$$

进行第二阶段梯度下降，即更新初始化参数  $\phi$ ，更新过程为：

$$\phi \leftarrow \phi - \beta \partial L(\phi) / \partial \phi \tag{12}$$

由此得到权重初始值  $w_i$ ，赋予预处理后的样本数据，进入下一步的 Softmax 分类计算。通过神经网络的计算以及设计损失函数，记录预测标签与真实标签的差值，并反馈给特征提取阶段和权重生成阶段。根据真实值与预测值的偏移量相应地调整当前权重，由特征提取器的移动平均值对权重生成器进行更新。

本文采用元学习利用少量数据计算出一个较适合的初始权重  $w_i$ ，再通过 DCW 模型在  $w_i$  的基础上，根据每一类的的数据量灵活分配神经网络对该类的重视程度，并利用损失函数反向更新，使少数类的学习率得到相应的补偿，提高对少数类的识别正确率。整个算法的结构流程如图 3 所示。系统输出为：

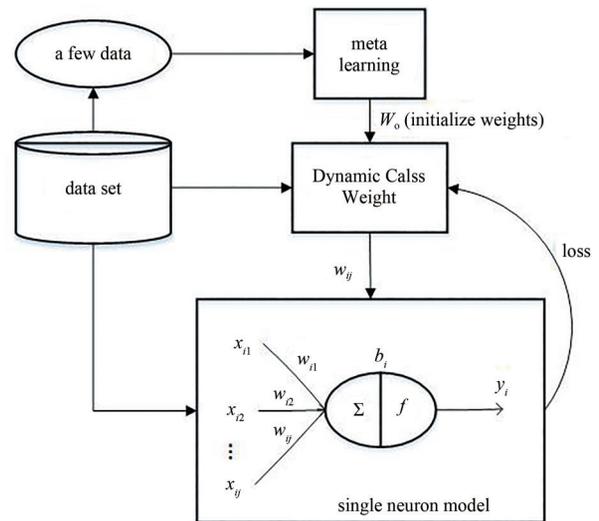


Fig.3 Block diagram of overall process  
图3 整体流程框图

$$y_i = f\left(b_i + \sum_j (x_{ij} * w_{ij})\right) \quad (13)$$

式中： $x_{ij}$ 为输入信号； $y_i$ 为输出； $w_{ij}$ 为赋予输入信号的权重； $b_i$ 为神经元内部状态即偏置； $f(\cdot)$ 为激活函数。

## 2 网络模型

### 2.1 卷积神经网络结构

卷积神经网络(Convolutional Neural Network, CNN)<sup>[19-21]</sup>是深度学习中一个非常重要的神经网络结构，多用于图像、音视频等信息的处理。由于具有强大的前向反馈功能，CNN能够自动学习并提取数据中的特征，进而利用所获取的特征信息进行识别。CNN主要包括卷积层、池化层、全连接层，卷积层目的是提取输入数据的特征，利用卷积核作为“特征过滤器”，采用不同的卷积核提取不同的特征，并在迭代过程中提取越来越复杂的特征，卷积层数和卷积核的个数决定了提取特征的数量和复杂度，网络设计时需平衡网络结构的宽度和深度；池化层是对卷积层提取到的特征进行亚采样处理，以减少网络中参数个数，避免过拟合，提高神经网络识别的鲁棒性；全连接层把卷积神经网络学到的特征映射到样本空间，完成最终的识别任务。本文使用的CNN网络模型如表2所示。

卷积核大小设为 $3 \times 1$ ，卷积层1(Conv 1)由32个长度为3的一维卷积核构成，卷积运算后做最大池化。卷积层2(Conv 2)由64个长度为3的一维卷积核构成，卷积运算后做最大池化。全连接层1(Dense 1)1 024个结点，全连接层2(Dense 2(Softmax输出))4个结点，输出的4个结点代表4个不同的辐射源。

### 2.2 损失函数设计

针对数据集的不平衡现状，提出一种改进的交叉熵损失函数，用以弥补由于样本类别不平衡带来的梯度失衡问题，提高少数类样本生成的损失对于整体损失的重要程度。

标准交叉熵损失(Cross Entropy loss, CE)表示为：

$$L_{CE} = -\sum_{i=1}^N y_i \log(\hat{y}_i) \quad (14)$$

式中： $y_i$ 为第*i*个元素的真实标签； $\hat{y}_i$ 为模型预测该元素属于第*i*个类别的概率。交叉熵损失的本质是衡量2个概率分布之间的距离，其中一个概率分布是真实标签的分布 $y$ ，另一个是模型预测的概率分布 $\hat{y}$ 。当2个概率分布越接近时，交叉熵损失越小，模型预测结果越准确。CE没有考虑样本数据不平衡时多数类对梯度的误导，少数类对整体损失的贡献更小，交叉熵损失最终偏向对多数类的判断，对少数类的误判概率增加。对此，Lin等<sup>[22]</sup>提出焦点损失(Focal Loss, FL)：

$$L_{FL} = -\alpha_i (1 - p_i)^\gamma \log(p_i) \quad (15)$$

式中： $\alpha$ 为控制正负样本的比例参数，取值范围为 $[0, 1]$ ， $\alpha$ 的取值一般可通过交叉验证选择合适的值； $\gamma$ 为聚焦参数，取值范围为 $[0, +\infty)$ ，目的是通过减少易分类样本的权重，从而使模型在训练时更专注于困难样本，当 $\gamma=0$ 时，FL退化为交叉熵损失， $\gamma$ 越大，对易分类样本的惩罚力度就越大。Zhang等<sup>[14]</sup>提出一种类不平衡损失(Class Imbalance loss, CI)，采用三角函数因子调节各类数所占的比重：

$$L_{CI} = -v \sum_{i=1}^C y_i \left(1 - \sin\left(\frac{\pi}{2} \hat{y}_i\right)\right) \cos\left(\frac{\pi}{2} \hat{y}_i\right) \log(\hat{y}_i) \quad (16)$$

式中 $v$ 为超参数。这些改进都调整了损失函数的下降速率，置信度较高数据的损失函数梯度下降更快，但对于本文提出的高度不平衡数据集，不平衡比(IR)小于0.1(数据集中最小类和最大类样本数量的比值小于0.1)，低置信概率部分获得的重视程度还不够充足，高置信概率部分的损失值尽管处在较低水平，但在数量的优势加成下，在整体损失中所占的比重仍有相当的优势，对高度不平衡数据的梯度弥补效果不够，并没有使低置信概率区间的不占优势的少数类获得足够高的损失值。基于此，本文提出另一种改进的交叉熵损失函数，引入IR，将少数类占总体的比重作为一个系数因子加进损失函数中，提高少数类对总体损失的贡献率。设不平衡数据集包含*c*个

表2 神经网络结构

Table2 Neural network structure

layer name	parameter settings
input layer	300×1
convolutional 1 ReLU	filters=32, kernel_size=3×1, strides=1, padding=same
pooling	pool_size=3, stride=3
convolutional 2 ReLU	filters=64, kernel_size=3×1, strides=1, padding=same
pooling	pool_size=3, stride=3
Dense 1 ReLU	expand 1 024 nodes
Dense 2	4 nodes
Softmax output	correspond to 4 types of labels

不同类的数据，第  $i$  类的数据量为  $N_i$ ，占整个数据集总数的比例为： $P_i = N_i / \sum_{i=1}^c N_i$ ，定义不平衡损失(Imbalanced Loss, IL)为：

$$L_{IL} = -\mu \sum_{i=1}^N y_i e^{-(\lambda P_i + \chi)} \log(\hat{y}_i) \quad (17)$$

式中： $P_i$ 为各类占总数的比例，即不平衡比； $\mu$ 、 $\lambda$ 和 $\chi$ 为超参数， $\mu$ 为权重因子，当少数类数据信息极少，即 $P_i$ 过于小时，可以通过增大权重因子提高少数类样本损失在总体损失中的比重； $\lambda$ 为缩放系数，多数类数据占总体比例过高，即 $P_i$ 过大时， $\lambda$ 可调整对 $P_i$ 的限制程度，即对多数类损失的抑制程度； $\chi$ 为调制系数，用于调节IL损失曲线的下降速度，其实质是对各类样本的损耗分离度进行分别设置。

图4为包含本文所提损失函数在内的几种损失函数损失值随置信度变化的曲线。从图中可以看出，四类损失函数均能随着置信度的变化自动地调整下降速度：对于置信概率较高的样本，损失值接近于0；对于置信概率较低的样本，几类损失的增益都处在较高水平，以此来提高少数类样本在计算中的权重。相比于FL中的幂指数调制因子和CI损失中的三角函数因子，本文提出的IL损失函数采用指数函数，在低置信区间有更高的损失值，在高置信区间有较低损失值，加入不平衡比弥补不平衡因子并赋值给自然指数部分，超参数调整损失函数的下降速率。图中IL损失函数曲线表现为下降更快，使易分类样本的损失速度下降更快，边界分离性能也更好。

### 3 仿真实验

#### 3.1 仿真条件

本文实验平台为Windows 10，双谱算法平台为Matlab R2018b，深度学习模型架构为Tensorflow2.2.0。采用基于GNU Radio软件的模拟辐射源，通用软件无线电外设(Universal Software Radio Peripheral, USRP)作为信号收发器。实验过程中，搭建由装有GNU Radio软件的计算机和8台USRP-B210组成的软件无线电(Software Defined Radio, SDR)平台作为信号发射装置，另一台USRP-B210作为信号接收装置，信号采集过程如图5所示。

在GNU Radio的GUI界面中，发射信号参数设置如下：信号调制方式为16QAM，载波频率为1.09 GHz，带宽为2 MHz，发射增益为50 dB，USRP采样率为2 MHz，通道选择TX/RX2:0。在接收端配置B210接收机设备信息，工作频率设为20 MHz，采样频率为2 MHz，通道选择RX2:0，接收增益为60 dB。在发射机流图与接收机装置均连接好之后，进行真实环境下的数据收集，实验场景如图5所示。

#### 3.2 数据集设置

由于所采集的数据是实测数据，从图5中可以看出发射机和接收机之间距离较近，可近似认为信号传播过程中没有噪声干扰。现实情况下，信道噪声无法避免，因此在信号处理过程中按能量加入高斯白噪声，模拟不同信噪比(Signal-to-Noise Ratio, SNR)条件下的数据集，对DCW的性能进行验证和分析。设置SNR的取值范围为0~24 dB，每台发射机单一信噪比值下采集样本数为2 000，每条样本的采样点数为2 048，训练集和测试集的比例为3:1。为得到不平衡数据集，在训练集中按一定的比例随机采样，选择奇偶交叉采样法构造不平衡训练集，训练集1和训练集3的奇数类分别以标准训练集的大约20%和大约5%作为少数类，训练集2和训练集4的偶数类以标准训练集的大约20%和大约5%作为少数类，得到的训练集和测试集如表3所示，其分布情况如图6所示。

#### 3.3 网络收敛性分析

由于在神经网络中加入DCW模块，在计算时偏向少数类，因此本节对神经网络的收敛性进行分析。在各不平衡训练集上运行DCW-CNN，记录迭代后的损失函数如图7所示，batch-size设为256，优化器为Adam，学习率设为0.001。曲线表明，在DCW模块加持下，随着迭代次数增加，各不平衡数据集条件下网络的损失函数在训练65个epoch后均能在低水平达到稳定状态。

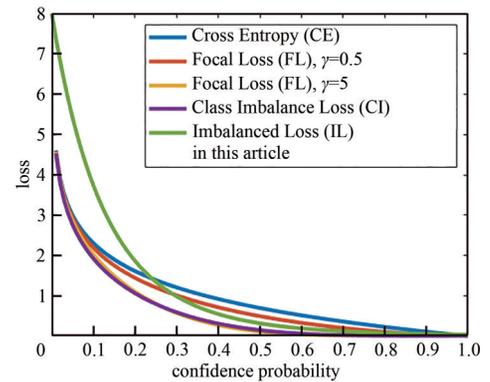


Fig.4 Four types of cost sensitive loss curves

图4 四类代价敏感损失曲线



Fig.5 Radiation source signal acquisition system

图5 辐射源信号采集系统实物图

表3 训练集和测试集设置  
Table3 Training and test set settings for simulation data

dataset	category							
	E1	E2	E3	E4	E5	E6	E7	E8
standard training set	1 500	1 500	1 500	1 500	1 500	1 500	1 500	1 500
standard test set	500	500	500	500	500	500	500	500
training set1	200	1 500	300	1 500	200	1 500	400	1 500
training set2	1 500	80	1 500	100	1 500	100	1 500	90
training set3	100	1 500	50	1 500	90	1 500	60	1 500
training set4	1 500	200	1 500	300	1 500	350	1 500	200

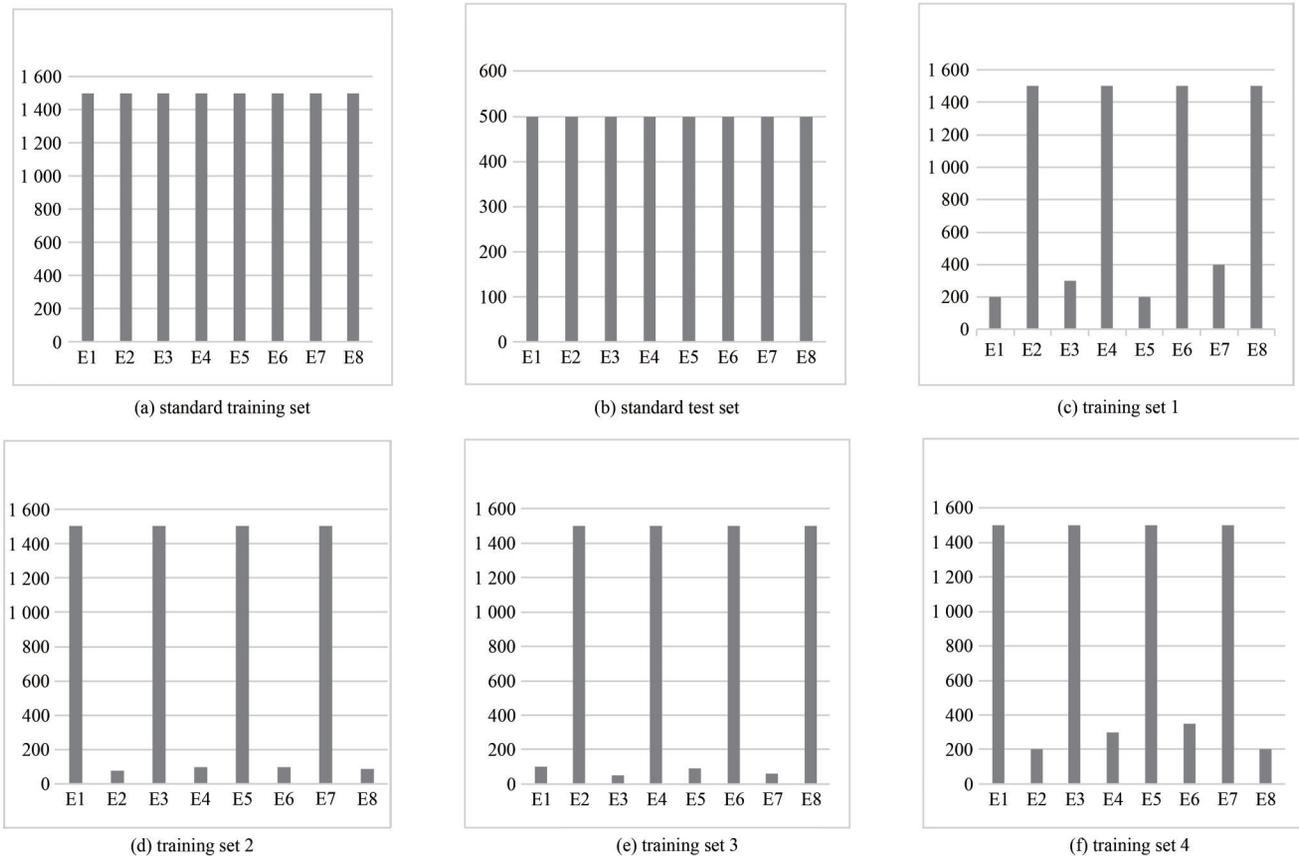


Fig.6 Dataset column charts  
图6 数据集设置柱形图

### 3.4 DCW 的有效性验证

为证明 DCW 模块的有效性，选取 DCW-CNN 和原始 CNN 进行识别性能对比。2 个网络除了有无 DCW 模块不同，其余网络结构和网络参数均相同。分别在 2 个网络上训练标准训练集、训练集 1、训练集 2、训练集 3 和训练集 4，保存训练好的模型，使用标准测试集检验 2 个网络在不同训练集上的学习训练效果，2 种网络的识别准确率如图 8 所示。

2 种网络在标准训练集上的训练效果相差不多，但在训练不平衡数据集时，添加了 DCW 模块的 DCW-CNN 相比于原始 CNN 有明显的优势。说明提出的 DCW 在处理不平衡数据集时，确实可以通过分配偏向于少数类的权重来平衡多数类对整体识别的误导，有控制地加强对少数类特征的学习，在测试集中无论数据集是否不平衡，都能达到较高的识别率。

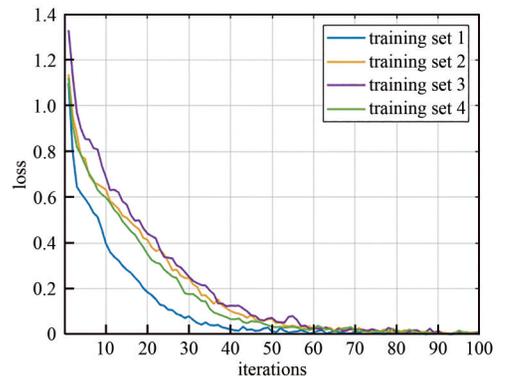


Fig.7 Loss value change curves of unbalanced training set  
图7 不平衡训练集损失值变化曲线

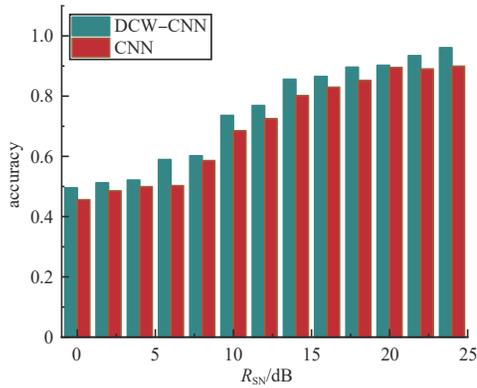


Fig.8 Comparison of recognition accuracy between DCW-CNN and CNN  
图 8 DCW-CNN 与 CNN 的识别准确率对比

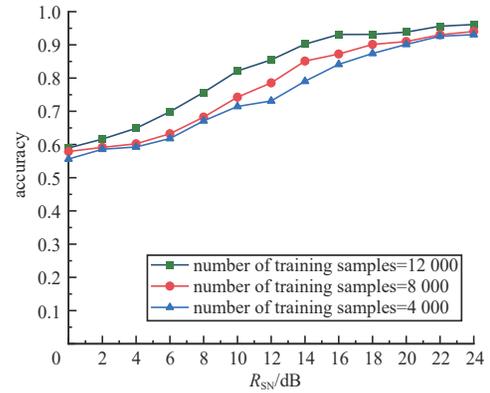


Fig.9 Comparison of recognition performance under different training sample numbers  
图 9 不同训练样本数下识别性能对比

### 3.5 网络鲁棒性分析

为验证不同样本数量对网络性能的影响，对标准训练集进行均匀随机采样，按能量加入高斯白噪声模拟不同的信噪比，损失函数设置为标准交叉熵损失，实验结果如图 9 所示。在不同的信噪比条件下，识别准确率随着训练样本数量的增大而增大。低信噪比条件下，几个训练集的识别准确率相差不大；6 dB 条件下，训练集数量为 12 000 时的识别率高于数量为 4 000 时约 11%，这是由于样本较少时，神经网络在较少的样本中产生过于密集的计算，产生过拟合现象，可以通过修改网络参数调整。

### 3.6 类不平衡损失性能评估

通过在 1 个均衡数据集和 4 个不均衡数据集上训练 3 类损失函数(CE 损失，FL 损失和 IL 损失)，分析 IL 损失与其他 2 类损失的性能，并在标准测试集上比较这 3 类损失函数的测试准确率，结果如表 4 所示。

表 4 不同损失函数在类不平衡数据集上的性能评估(%)  
Table4 Accuracy of different loss functions on class-imbalanced datasets(%)

		$R_{SN}/dB$												
		0	2	4	6	8	10	12	14	16	18	20	22	24
training set 0	CE	53.46	54.34	58.94	62.87	67.42	73.81	84.98	90.17	92.88	94.25	95.43	96.51	97.12
	FL	50.15	51.23	54.01	58.64	63.97	70.54	80.16	87.49	91.69	93.55	94.07	94.19	94.35
	IL	50.10	52.11	52.84	60.51	63.15	68.51	79.37	88.09	90.87	91.88	92.35	92.61	92.94
training set 1	CE	29.36	34.47	37.51	45.64	48.91	53.17	60.25	66.18	68.20	70.23	75.62	76.76	78.45
	FL	29.65	34.89	38.14	47.43	50.68	57.91	67.29	71.58	75.66	80.44	81.62	82.34	83.56
	IL	30.16	35.08	39.26	49.24	55.35	61.24	70.19	74.63	82.14	83.51	83.96	84.05	85.73
training set 2	CE	27.24	31.76	33.24	39.91	43.15	49.64	55.11	60.58	64.99	65.19	71.45	71.05	72.64
	FL	28.39	32.62	35.08	42.58	47.95	56.28	62.31	70.06	76.25	78.24	81.35	82.28	83.96
	IL	29.09	33.24	38.41	46.28	53.65	60.19	66.79	72.18	79.26	81.54	82.29	83.36	84.64
training set 3	CE	27.09	29.77	31.02	37.15	41.25	47.61	53.59	57.62	62.15	63.32	70.42	71.02	71.96
	FL	29.28	31.92	36.74	43.46	48.68	57.02	62.15	69.89	74.56	76.25	77.65	78.97	81.23
	IL	30.05	33.99	40.21	45.69	54.03	59.65	65.02	71.65	80.06	80.26	81.65	82.68	83.96
training set 4	CE	28.55	33.24	35.24	42.12	45.05	50.53	57.26	64.13	66.05	67.13	72.96	73.47	76.68
	FL	29.86	33.68	36.49	46.85	49.78	57.64	67.79	70.26	76.26	79.39	81.26	82.99	83.25
	IL	31.25	34.58	38.94	47.65	55.06	61.57	69.08	73.28	81.07	82.58	83.14	84.68	85.58

在同一训练集中不同的信噪比条件下，可以看出识别准确率随着信噪比的增大而稳定提高。在信噪比小于 12 dB 时，信噪比的增加对于识别准确率的提高有明显增进作用，这是因为在一定阈值内，噪声能量由大变小，使信号中细微的个体差异能够极大程度地释放出来并被捕捉到，使识别准确率增大的速度较快；当信噪比达到这个阈值时，再往上，对于细微特征相对来说的“放大”效果就不再那么明显，此时信噪比的提高对于提高识别准确率的贡献度降低。

比较不同训练集的识别效果，可以看出，在标准训练集上三类损失的识别准确率差别不大，这是因为在均衡数据集上，FL 损失和 IL 损失的加权模块作用不大，无法根据数据集不平衡的误差反馈对权重进行重新分配，因此三类损失的工作量也相差不大。对于四类不均衡的训练集，FL 损失和 IL 损失的识别准确率均高于 CE 损失，且在多数情况下 IL 损失的识别效果要优于 FL 损失。这是因为 IL 损失在调整权重的基础上引入了不平衡度因子，

对分配权重再次修正,使不平衡数据集中的少数类凭借在数据集总量中占的比例得到相应程度的权重弥补。由此验证了IL损失的性能确实较CEL损失和FL损失有不同程度的提升,平均识别准确率相比CEL损失和FL损失分别提高了9.14%和2.71%。

### 3.7 算法识别性能分析

为验证DCW-CNN配合IL损失在真实场景下的识别效果,将收集到的8台USRP-B210信号进行双谱预处理后,选取表3中不平衡训练集(training set 3),研究待测辐射源个数 $k(k=4、5、6、7、8)$ 对本文识别框架的影响,结果如图10所示。由图10可知,辐射源数量 $k=4$ 时,平均识别准确率能达到87.37%;辐射源数量 $k=8$ 时,平均识别准确率为83.91%,识别性能随着辐射源数量的增加而逐渐下降。使用混淆矩阵中的错识别概率对应训练集3(training set 3)的不平衡数据分布情况可以看出,一个待测样本的标签预测概率与其在整个数据集中所占的份额成一定比例。这也侧面印证了通过调整学习权重以及将数据集的类不平衡比(IR)加入到不平衡损失函数中的方法可以有效提升不平衡数据的识别正确率。

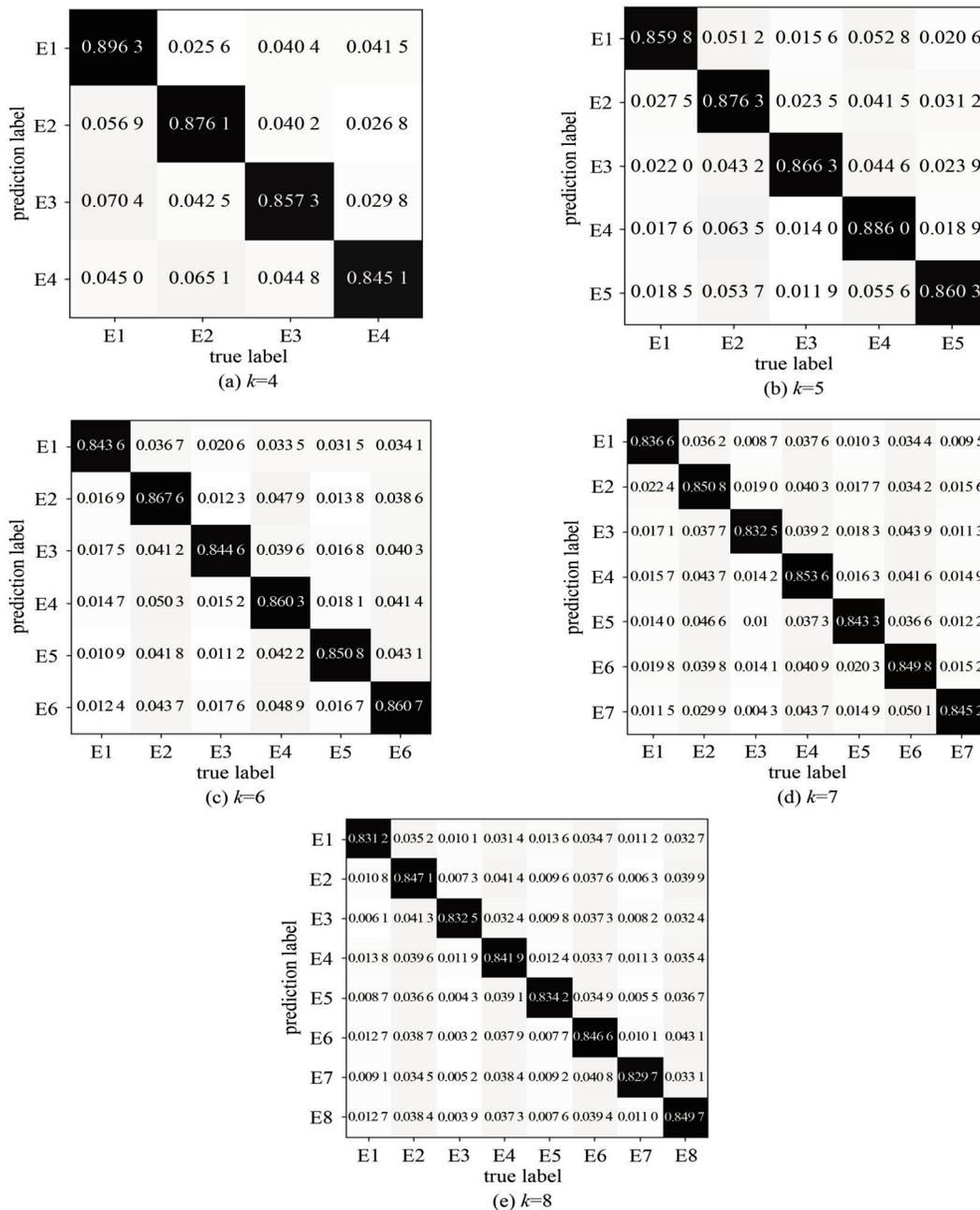


Fig.10 Recognition results for different numbers of emitters  
图 10 模型对于不同辐射源数量的识别结果

#### 4 结论

针对数据不均衡导致识别准确率下降的问题，本文提出使用元学习生成权重初始值并用两层反馈动态调整类权重的方法，给不同数量级的样本数据分配不同的权重，给少数类样本一定的重视程度，纠正多数类对神经网络判断的误导；同时提出一种改进的交叉熵损失函数，克服多数类带来的模型性能退化问题，为不平衡识别问题提供新的解决方案，对于改善识别处理高度不平衡数据的效果有一定的应用意义，但本文对于损失函数的研究还存在一定遗憾，对处于置信概率中段的数据没有给予足够的重视，损失函数从低置信概率到高置信概率下降极快，对于中段数据的考虑有所欠缺，希望在以后的研究中对这一情况加以重视。

#### 参考文献：

- [ 1 ] LI Bi,XI Teng,ZHANG Gang,et al. Dynamic class queue for large scale face recognition in the wild[C]// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Nashville,TN,USA:IEEE, 2021:3763–3772.
- [ 2 ] LIU Tianyu,CHEN Bojian,HUANG Weiguang,et al. Assessment of tool wear with insufficient and unbalanced data using improved conditional generative adversarial net and high-quality optimization algorithm[J]. IEEE Transactions on Industrial Electronics, 2023,70(11):11670–11680. doi:10.1109/TIE.2022.3231300.
- [ 3 ] 李成,晋玉猛,田文彪. 多尺度自适应直接信息采样与重构[J]. 海军航空大学学报, 2019,34(2):187–192. (LI Cheng,JIN Yumeng, TIAN Wenbiao. Multi scale adaptive direct information sampling and reconstruction[J]. Journal of Naval Aviation University, 2019,34(2):187–192.) doi:10.7682/j.issn.1673–1522.2019.02.002.
- [ 4 ] KHANDOKAR I A, ABDULLAH A T, KHONDOKAR T, et al. A clustering based priority driven sampling technique for imbalance data classification[C]// 2022 the 14th International Conference on Software, Knowledge, Information Management and Applications(SKIMA). Phnom Penh,Cambodia:IEEE, 2022:176–180. doi:10.1109/SKIMA57145.2022.10029565.
- [ 5 ] MALHOTRA R,KASHYAP V,SHARMA V. Comparative study of sampling techniques for software defect prediction[C]// 2022 the 8th International Conference on Advanced Computing and Communication Systems(ICACCS). Coimbatore,India:IEEE, 2022: 760–766. doi:10.1109/ICACCS54159.2022.9785238.
- [ 6 ] DISHA N,CHADAGA N B,SINGH H,et al. Empirical analysis of sampling methods on imbalanced data[C]// 2022 IEEE North Karnataka Subsection Flagship International Conference(NKCon). Vijaypur, India: IEEE, 2022: 1–6. doi:10.1109/NKCon56289. 2022.10126733.
- [ 7 ] 陈白强,盛静文,江开忠. 基于损失函数的代价敏感集成算法[J]. 计算机应用, 2020,40(S2):60–65. (CHEN Baiqiang,SHENG Jingwen,JIANG Kaizhong. Cost sensitive integration algorithm based on loss function[J]. Computer Applications, 2020,40(S2): 60–65.) doi:10.11772/j.issn.1001–9081.2020020213.
- [ 8 ] LI Kai,WANG Bo,TIAN Yingjie,et al. Fast and accurate road crack detection based on adaptive cost-sensitive loss function[J]. IEEE Transactions on Cybernetics, 2023,53(2):1051–1062. doi:10.1109/TCYB.2021.3103885.
- [ 9 ] WANG Chao,JI Liqiang,SHI Fan. Collapsed building detection in high-resolution remote sensing images based on mutual attention and cost sensitive loss[J]. IEEE Geoscience and Remote Sensing Letters, 2023(20): 1–5. doi:10.1109/LGRS. 2023.3268701.
- [ 10 ] CHAWLA Nitesh V,BOWYER Kevin W,HALL Lawrence O,et al. Smote:synthetic minority over-sampling technique[J]. Journal of Artificial Intelligence Research, 2002,16(1):321–357. doi:10.1613/jair.953.
- [ 11 ] ZHONG Yaoyao,DENG Weihong,WANG Mei,et al. Unequal-training for deep face recognition with long-tailed noisy data[C]// CVF Conf Comput Vis Pattern Recognit(CVPR). Long Beach,CA,USA:IEEE, 2019:7804–7813. doi:10.1109/CVPR.2019.00800.
- [ 12 ] ZHENG Jian,QU Honchun,LI Zhoni,et al. Conformal transformation twin-hyperspheres for highly imbalanced data to binary classification[C]// 2022 IEEE the 9th International Conference on Data Science and Advanced Analytics(DSAA). Shenzhen, China:IEEE, 2022:1–6. doi:10.1109/DSAA54385.2022.10032448.
- [ 13 ] SADREDDIN A,SADAOU S. Training and testing cascades for imbalanced data classification[C]// 2022 IEEE Symposium Series on Computational Intelligence(SSCI). Singapore,Singapore:IEEE, 2022:261–268. doi:10.1109/SSCI51031.2022.10022190.
- [ 14 ] ZHANG Linbin,ZHANG Caiguang,QUAN Sinong,et al. A class imbalance loss for imbalanced object recognition[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020,13(1): 2778–2792. doi:10.1109/JSTARS. 2020.2995703.
- [ 15 ] 张贤达. 现代信号处理[M]. 2版. 北京:清华大学出版社, 2002:274–281. (ZHANG Xianda. Modern signal processing[M]. 2nd ed. Beijing:Tsinghua University Press, 2002:274–281.)