

引用格式: 夏文博, 范威, 高莉. 基于卷积神经网络的水下多目标方位估计方法[J]. 声学技术, 2023, 42(3): 290-296. [XIA Wenbo, FAN Wei, GAO Li. Underwater multi-target azimuth estimation method based on convolutional neural network[J]. Technical Acoustics, 2023, 42(3): 290-296.] DOI: 10.16300/j.cnki.1000-3630.2023.03.004

基于卷积神经网络的水下多目标方位估计方法

夏文博¹, 范威¹, 高莉²

(1. 上海船舶电子设备研究所, 上海 201108;
2. 中科长城海洋信息系统有限公司, 北京 100085)

摘要: 针对水下多目标方位估计问题, 提出了一种利用卷积神经网络模型对目标声源进行方位估计的方法。该方法使用不等强度的声源数据进行训练并使用焦点损失函数作为训练损失函数。通过对阵列接收到的信号进行特征提取, 使用焦点损失函数指导卷积神经网络训练, 最终利用训练好的卷积神经网络模型进行目标方位估计。对不同模型参数的训练进行对比, 结果表明所训练的卷积神经网络模型在较低信噪比条件下也能正确估计弱目标的方位。试验结果表明, 与采用二元交叉熵损失函数的卷积神经网络模型相比, 该方法对弱目标的方位估计能力更强, 提高了方位估计的准确率。

关键词: 方位估计; 特征提取; 卷积神经网络; 深度学习; 焦点损失函数

中图分类号: TB566

文献标志码: A

文章编号: 1000-3630(2023)-03-0290-07

Underwater multi-target azimuth estimation method based on convolutional neural network

XIA Wenbo¹, FAN Wei¹, GAO Li²

(1. Shanghai Marine Electronic Equipment Research Institute, Shanghai 201108, China;
2. Zhongke Great Wall Ocean Information System Co., Ltd., Beijing 100190, China)

Abstract: To solve the problem of underwater multi-target azimuth estimation, a method to estimate the target sound source azimuth using the convolutional neural network (CNN) model is proposed. In this method, unequal intensity sound source data are used for training and the focal loss function is taken as the training loss function. Through the feature extraction of the signals received by the array, the focal loss function is used to guide the convolutional neural network training, and finally the trained convolutional neural network model is used to estimate the target azimuth. By comparison with the training results of different model parameters, it is shown that the trained convolutional neural network model can correctly estimate the azimuth of weak targets under the condition of low SNR. And, the experimental results show that in contrast with the convolutional neural network model using the binary cross-entropy loss function, the method in this paper has a stronger ability to estimate weak target azimuth and improves the estimation accuracy.

Key words: azimuth estimation; feature extraction; convolutional neural network(CNN); deep learning; focal loss function

0 引言

信号的目标方位(Direction of Arrival, DOA)估计是水声阵列信号处理中的一个基本问题, 常用的水下目标方位估计方法包括最小方差无失真响应(Minimum Variance Distortionless Response, MVDR)法、多重信号分类(Multiple Signal Classification, MUSIC)法等^[1]。这些方法都基于的是声呐阵列流形

模型。随着人工智能的不断发展, 基于深度学习的水下目标方位估计技术受到广泛的关注。

Zhu 等^[2]采用深度学习的方法对单目标方位估计进行了研究。多目标方位估计问题相对于单目标方位估计更为复杂, 目标数量的增多使得神经网络的训练开支与模型的复杂性都相应增加, 不同声源之间的强弱差异性也会影响估计结果的准确性。文献[3-4]采用遍历所有目标方位组合的方式训练网络模型, 研究了两目标的方位估计。对于声源数量未知、声源方位分布更多的情况, 遍历声源方位的所有分布是不现实的, 一些学者采用均匀随机抽样的方法^[5-9], 实现了对多目标情况的方位估计。在水声领域中, Ozanich 等^[17]将目标方位估计问题转化为深

收稿日期: 2022-01-16; 修回日期: 2022-03-01

作者简介: 夏文博(1996—), 男, 陕西咸阳人, 硕士研究生, 研究方向为水声信号处理。

通信作者: 夏文博, E-mail: fu315184qiu82@163.com

度学习中的分类问题，提出了一种用于DOA估计的前馈神经网络(Feed-forward Neural Network, FNN)。借鉴图像处理领域中卷积神经网络(Convolutional Neural Network, CNN)的应用，在全连接层前加入卷积层可以提取阵列信号中的方位特征^[5-6, 8, 10]，减少了冗余特征对模型估计结果的不利影响，提高了方位估计能力。在这些研究的基础上，通过对信号自协方差矩阵和相位矩阵的联合检测，以及不同信噪比环境下的训练，可以提高深度学习方法在低信噪比条件下的目标方位估计能力^[9]。此外，深度学习方法在矢量水听器的目标方位估计中也同样适用^[11-12]。

以上研究是在多声源等强度或强度相差不大的情况下进行的。但在实际中，由于大型船舶等强干扰声源的存在，不同水下目标信号到达声呐阵列的强度差异性较大，对神经网络模型来说，同时对强目标和弱目标进行方位估计是一项新的挑战。文献[7]中的试验结果也体现了这一现状，当多目标源的强度不同时，神经网络模型对强目标、弱目标与背景噪声的区分效果变差。此时神经网络模型会倾向于将更弱的目标信号估计为背景噪声，在时间方位历程图上表现为有间断产生。

针对在强目标源存在的情况下深度学习方法对弱目标检测困难的问题，本文借鉴卷积神经网络对图像的分类识别方法，对信号的协方差矩阵进行特征提取后送入全连接层中进行方位估计。训练时采用不等强度的声源信号，与传统深度学习采用均方误差和交叉熵作为训练时的损失函数不同，本文采用焦点损失作为损失函数指导卷积神经网络模型的收敛与参数调整，增强对弱目标与噪声的区分能力。仿真及试验结果表明，本文所提方法收敛速度更快，检测概率更高，可以解决深度学习网络模型在多目标方位估计中弱目标方位估计困难的问题。

1 声呐阵列信号协方差模型

1.1 声呐阵列信号平面波模型

假设声呐阵列接收到的信号为远场平面波，如图1所示。在到达 M 元阵列的信号中，第 n 个阵元接收到的信号的声程差 D_n 为

$$D_n = (n-1)d \sin \theta_k \quad (1)$$

式中： d 为阵元间距， θ_k 为第 k 个声源的入射角度。则时延 τ_n 为

$$\tau_n = \frac{D_n}{c} \quad (2)$$

式中： c 为水中声速。则阵元间接收信号的频域相

移 β 可表示为

$$\beta = e^{-j \frac{2\pi(n-1)d \sin \theta_k}{\lambda}} \quad (3)$$

式中： λ 为波长。线列阵接收到的信号 $y_n(t)$ 可表示为

$$y_n(t) = \sum_{k=1}^K x_k(t) e^{-j \frac{2\pi(n-1)d \sin \theta_k}{\lambda}} + b_n(t) \quad (4)$$

式中： K 为声源数， $x_k(t)$ 为第 k 个目标信号， $b_n(t)$ 为加性高斯白噪声。将式(4)以矩阵形式表示为

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{B} \quad (5)$$

$$\mathbf{Y} = [y_1(t) \ y_2(t) \ \cdots \ y_M(t)]^T \quad (6)$$

$$\mathbf{X} = [x_1(t) \ x_2(t) \ \cdots \ x_K(t)]^T \quad (7)$$

$$\mathbf{B} = [b_1(t) \ b_2(t) \ \cdots \ b_M(t)]^T \quad (8)$$

$$\mathbf{A} = [a(\theta_1) \ a(\theta_2) \ \cdots \ a(\theta_K)] \quad (9)$$

$$a(\theta_k) = \left[1 \ e^{-j \frac{2\pi d \sin \theta_k}{\lambda}} \ \cdots \ e^{-j \frac{2\pi(M-1)d \sin \theta_k}{\lambda}} \right] \quad (10)$$

式中： $\mathbf{Y}$ 为阵列接收信号矩阵， $\mathbf{X}$ 为 t 时的传输信号， $\mathbf{B}$ 为阵列接收到的噪声， $\mathbf{A}$ 为阵列流形矩阵， T 代表矩阵转置。

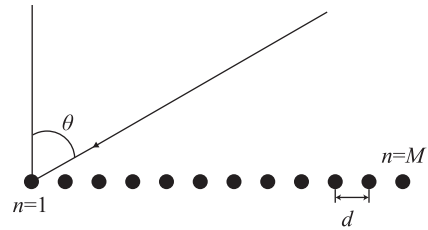


图1 远场平面波模型示意

Fig.1 Schematic diagram of the far-field plane wave model

1.2 训练数据集构造

本文将水下多目标的DOA估计问题转化为一个多标签多分类问题。为此，将信号的到达方位均匀划分为180个扇角，设有目标源的扇角标签为“1”，无目标源的扇角标签为“0”，如图2所示。所得的长度为180的一维矩阵即为训练数据的标签。

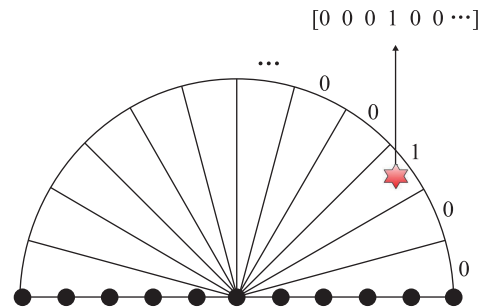


图2 训练数据的标签构建

Fig.2 Construction of training data labels

将阵列接收到的阵元域信号的自协方差矩阵 $\mathbf{R}_y$ 作为模型的输入, 有:

$$\mathbf{R}_y = \mathbf{E}[\mathbf{Y}\mathbf{Y}^H] \quad (11)$$

式中: 符号 $\mathbf{E}[\cdot]$ 代表期望, H 代表矩阵的共轭转置。

2 卷积神经网络模型

2.1 网络结构

本文的多目标 DOA 估计卷积神经网络模型的参数如表 1 所示。首先, 卷积层提取输入数据中目标方位的特征信息, 各层网络学习目标方位特征后输出空间方位的分类结果。卷积神经网络模型的输出即为到达角空间上各方位是否存在目标的概率。

表 1 卷积神经网络的详细参数
Table 1 Detailed parameters of CNN

层类型	核大小	通道数	神经元数量	激活函数
卷积层 1	7×7	64	—	Relu
卷积层 2	5×5	64	—	Relu
卷积层 3	3×3	64	—	Relu
卷积层 4	2×2	64	—	Relu
全连接层 1	—	—	1 024	Relu
全连接层 2	—	—	512	Relu
全连接层 3	—	—	256	Relu
全连接层 4	—	—	180	Sigmoid

对卷积层 1 接收到的输入数据进行 L2 正则化, L2 正则化对每个元素进行不同比例的放缩, 可以减小过拟合现象, 提升模型对真实数据的适应能力。各卷积层中卷积运算的步长为 1 且边缘不补 0。层与层之间采用批归一化的方法来加快模型的收敛速度, 减少梯度弥散现象, 使训练更加稳定。除输出层采用 Sigmoid 激活函数外, 各层均采用 Relu 激活函数。Sigmoid 激活函数的表达式为

$$S(z) = (\mathbf{H}e^{-z})^{-1} \quad (12)$$

式中: z 为任意实数。Sigmoid 激活函数的输出是 0 到 1 之间的值, 对应目标源是否存在的概率。Relu 激活函数的表达式为

$$f(x) = \max(0, x) \quad (13)$$

2.2 焦点损失函数

在神经网络训练中, 损失函数用于指导网络的收敛和网络中参数与偏置的调整。在文献[4-8]和[10]中常常采用交叉熵(Cross Entropy, CE)与二元交叉熵(Binary Cross Entropy, BCE)作为损失函数。交叉熵损失函数的表达式为

$$E_{Cn} = t_n \lg y_n \quad (14)$$

式中: t_n 为第 n 个样本的期望输出, 即真实值; y_n

为第 n 个样本的输出值, 即估计值。二元交叉熵损失函数的表达式为

$$E_{Cn} = -[t_n \lg y_n + (1 - t_n) \lg(1 - y_n)] \quad (15)$$

在式(14)、(15)的两个损失函数中, 若估计值与真实值完全一致, 则损失值降到最低。在文献[6-8]中, 阵列接收到的各目标源强度被认为是相等或相近的, 而实际中各声源到达阵列时的强度往往是不相等的。当输入数据中弱目标与强目标同时存在时, 由于声源之间的强度差异在数值上是呈指数变化的, 此时弱目标的参数值远小于强目标, Sigmoid 激活函数输出的弱目标的估计值会偏向于 0。且由于弱目标的参数值远小于强目标, 其对神经网络权重和偏置更新的影响也远小于强目标的影响。

从式(15)中可以看出, 当 $t_n=1$ 时, 在弱目标存在的方位, y_n 更接近 0, 在强目标存在的方位 y_n 更接近 1, 使得损失函数的第一项变动很大, 整个模型的收敛减缓, 且容易收敛到局部极小值。在本研究中, 强目标的方位估计是一个易分类样本, 弱目标的方位估计是一个难分类样本。当强目标、弱目标与背景噪声同时存在时, 此时神经网络模型会倾向于将弱目标与背景噪声划为一类。

对于难分类样本的估计问题, 深度学习图像处理领域的 Lin 等^[13]提出了焦点损失(Focal Loss, FL)函数, 其表达式为

$$E_{FLn} = -\alpha(1 - y_n)^\gamma \lg y_n \quad (16)$$

式中: α 为平衡因子, $\alpha \in [0, 1]$ 。 α 用于平衡正负样本比例不均的问题, 调节 α 可以使神经网络在训练时着重于正样本或负样本。在实际情况中, 有目标存在的方位是正样本, 其数量比负样本即无目标存在的方位的数量少, 因此可以调节 α 的值来增加正样本在训练中的权重。 γ 为焦点参数, 用于调节易分类样本与难分类样本对损失值的影响权重, $\gamma > 0$ 时可以减少易分类样本对损失值的贡献, 使神经网络在训练时更多地关注难分类样本。在训练过程中易分类样本的预测值与真实值差距较小, 难分类的样本预测值与真实值差距较大, 经过 γ 次方运算后较小的值会更小, 此时对损失值起更大影响的是与真实值差距较大的难分类的样本。

在本文中多源方位估计问题是一个多目标多分类问题, 也可以看作多个二分类问题, 即估计每一个方位是否有目标的概率。因此本文采用如下形式的损失函数, 其表达式为

$$E_n = -\left[t_n \alpha (1 - y_n)^\gamma \lg y_n + (1 - t_n)(1 - \alpha) y_n^\gamma \lg(1 - y_n)\right] \quad (17)$$

其中：当 $\alpha=0.5$ ， $\gamma=0$ 时该损失函数退化为二元交叉熵损失函数。

2.3 参数设定

神经网络各层的初始超参数经正态分布初始化随机生成，减少模型陷入局部极小值的可能。数据输入后其估计结果经焦点损失函数计算损失值，之后将误差反向传播更新各层的权值与偏置，完成一次训练学习。

使用 M 元线列阵生成训练数据，阵元间距为 $\lambda/2$ ， λ 为波长。仿真数据训练集共有 2×10^5 组数据，数据集为给定频率下随机生成的6个声源的阵列自协方差矩阵信号。各声源信号到达方位不同，到达阵列的强度也各不相同，训练集中加入强度随机的加性高斯白噪声。在强度不等的声源训练数据集中，各声源强度相差过大会使噪声被错误地估计为目标，降低神经网络模型对目标与噪声的区分能力。各声源强度相差过小会降低神经网络对弱目标的检测能力，结果接近于使用等强度声源训练集的检测结果。

经过实际训练与验证后发现，声源强度差与实际情况接近时效果较好，同时为了保证模型对不同信噪比数据的适用性，本文适当增加了训练数据集的源强度差。同一组训练数据中，2/3的声源信噪比在 $-20 \sim -5$ dB之间，另外1/3的声源信噪比在 $-5 \sim 5$ dB之间。

为了在训练时验证模型的泛化能力，验证集为与训练集不同方位的 2×10^4 组数据。在同一组验证数据中，2/3的声源信噪比在 $-25 \sim -5$ dB之间，另外1/3的声源信噪比在 $-5 \sim 5$ dB之间。

由于训练集中正负标签数量不均衡，在一组数据中有目标存在的方位(正样本)为6个，无目标的方位(负样本)为174个，负样本数量远大于正样本数量，此时无目标的方位对损失值影响更大，会主导梯度的更新方向。因此需要人为增加正样本的权重，通过调整使焦点损失函数中的 $\alpha > 0.5$ 。但 α 过大也会使得神经网络对负样本信息的学习不充分，表现为对弱目标与背景噪声的区分能力变差。合适的 α 需要通过调整测试得到。本文经测试对比，得到合适的 α 的取值为0.9，此时正负样本对损失值的影响比例约为1:3。同样，经测试对比后得到本文的焦点参数 $\gamma=2$ 。

模型训练采用Tensorflow 2.3深度学习库^[14]。训练参数为：训练的总迭代次数为100；初始学习率为0.001，若经5个迭代后损失值仍不下降，学习率降低为原学习率的1/10；采用自适应矩估计优化

器；每批次输入数据量为2 000。

3 仿真分析及试验数据验证

由于Tensorflow在处理复数时会丢失其虚部的信息，因此将 R_y 的实部与虚部分别存储可得一个 $M \times M \times 2$ 型的矩阵。则训练数据的特征集为 $\{\text{Re}(R_y), \text{Im}(R_y)\}$ ，与目标方位的标签集构成神经网络模型训练的输入与输出对。

3.1 仿真分析

在模型的训练学习过程中发现，采用不等强度声源数据集进行训练时相比采用等强声源训练的结果，在对真实数据的目标方位估计中检测效果更好。这是因为真实数据中多目标源到达阵列的强度是不相等的，为了验证在更改损失函数后模型的优越性，对采用二元交叉熵损失函数的模型与采用焦点损失函数的模型，两者的训练结果进行对比。在对比试验中，除损失函数外模型的其余参数不变。

由于无目标的方位远多于有目标的方位，因此标签中0的数量远大于1，此时的识别准确率并不适合作为判别标准，因为若模型只给出0这一估计值，模型的准确率也会较高。因此采用召回率结合精确率一起作为模型判别标准。模型训练过程中的精确率与召回率如图3所示。召回率是对模型估计

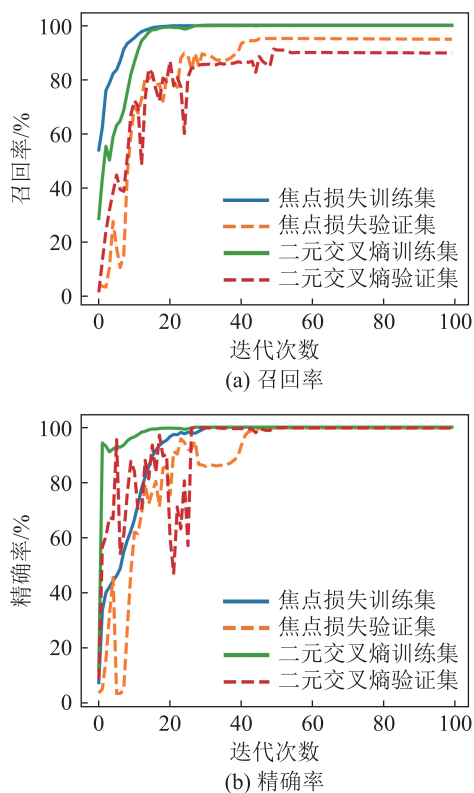


图3 两种模型召回率与精确率的对比

Fig.3 Comparison of recall rates and precisions of the two models

结果覆盖面的度量，衡量了模型对正例的识别能力，或称查全率。召回率可以判断出模型漏报了多少目标，召回率越高，模型对弱目标的方位估计效果就越好。精确率表示对正样本结果中的预测准确程度。

在训练初期召回率较低的时候，模型倾向于只估计易估计目标，随着迭代次数增多，召回率增加，精确率也随之波动。由图3可以看出，改进后的网络收敛速度更快，相同迭代次数下召回率更高。

两种训练模型的召回率与精确率的对比结果如表2所示。焦点损失函数模型验证集的召回率和精确率分别为94.8%和99.7%，召回率比二元交叉熵模型提高了5.0%，说明改进后的模型提高了对弱目标的方位估计能力。

表2 两种模型验证集召回率与精确率的比较
Table 2 Comparison of recall rates and precision rates on validation sets of the two models

损失函数类型	模型召回率%	模型精确率%
二元交叉熵	89.8	99.7
焦点损失	94.8	99.7

仿真测试数据的时延求和波束形成的时间方位历程图如图4所示。图4中共有6个目标源，在0时刻0°~180°的6个目标的信噪比分别为0、-9、-9、-7、-7和-14 dB。

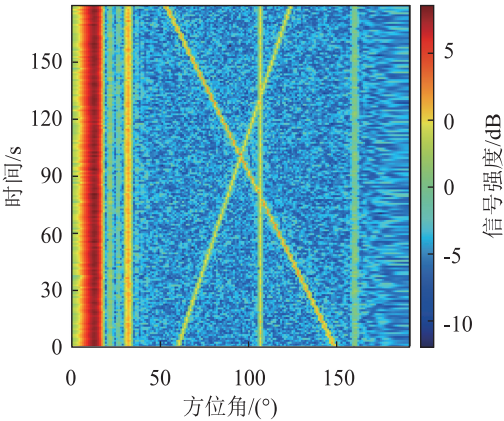


图4 测试数据常规波束形成结果
Fig.4 Conventional beamforming results of test data

三种深度学习方法二元交叉熵损失函数等强度声源训练模型、二元交叉熵损失函数不等强度声源训练模型、焦点损失函数不等强度声源训练模型，对仿真数据的方位估计结果如图5所示。

图5中，色棒的灰度代表目标检出的概率，范围为(0~1)。从仿真测试数据的方位估计结果中可以看出，图5(a)中等强度声源训练的模型对较弱目标方位估计效果不佳，在时间方位历程上存在不连

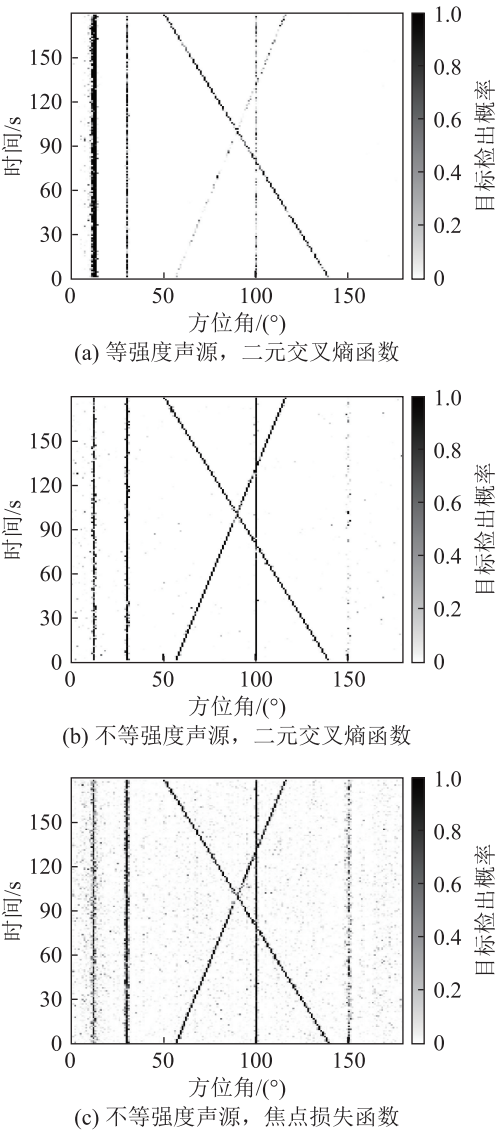


图5 三种深度学习训练结果对比
Fig.5 Comparison of training results of three deep learning methods

续，有间断的现象，且无法估计出声强与背景噪声接近的目标的方位；图5(b)中不等强度声源下训练的模型可以估计出较弱的目标，但当目标强度与噪声强度接近时，模型会将弱目标错判为噪声，对弱目标估计效果变差；采用焦点损失函数训练的模型，即使在低信噪比条件下，也能正确估计出目标的方位，其估计结果优于二元交叉熵损失函数训练的模型。

三种深度学习训练方法得到的模型对仿真测试数据的目标方位估计结果如表3所示。在同一检测阈值下，基于焦点损失模型的目标方位估计准确率比基于二元交叉熵的等强声源模型提升了31.2个百分点，比基于二元交叉熵的不等强度声源模型提升了9.6个百分点，证明改进后的模型对目标方位的估计效果更好。

表3 三种深度学习模型估计结果对比
Table 3 Comparison of estimation results of the three deep learning models

训练模型	方位估计准确率/%
等强度声源、二元交叉熵函数	61.9
不等强度声源、二元交叉熵函数	83.5
不等强度声源、焦点损失函数	93.1

3.2 试验验证

采用被动声呐阵列接收到的湖上实测数据对训练后的模型进行验证,结果如图6所示。试验数据中在78°处存在一个方位固定的强干扰目标,在100°处有一个弱目标。根据式(11)构建给定频率下的协方差矩阵,作为试验验证数据。

图6(a)为经时延求和的常规波束形成得到的方位估计结果,将其作为深度学习方法的对照。将阵

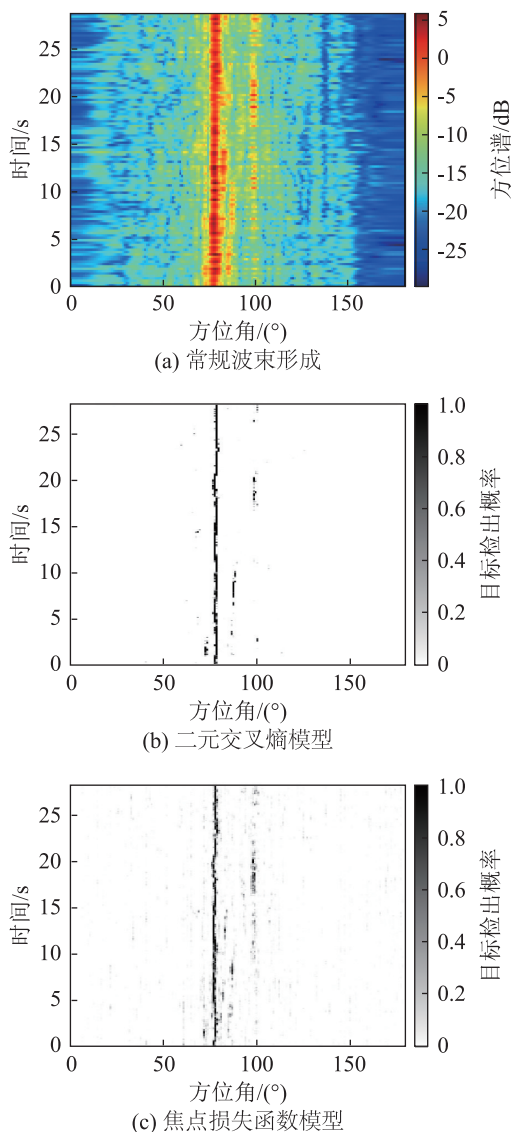


图6 三种方法的目标方位估计的试验结果对比
Fig.6 Comparison of experimental results of target azimuth estimation by three different methods

列自协方差矩阵送入两种已经训练好的卷积神经网络中进行估计,结果如图6(b)、6(c)所示。结合图6(b)、6(c)可以看出,两种深度学习模型对强目标都有着可靠的方位估计结果。与基于二元交叉熵的模型相比,图6(c)中基于焦点损失训练的模型对弱目标的检出率更高,方位估计结果更为连贯。

4 结论

本文对多目标方位估计问题中各声源强度的差异对方位估计结果的影响进行了分析研究。解决了强目标、弱目标与背景噪声同时存在时对弱目标方位估计困难的问题。本文提出了采用不等强度声源作为训练集的训练方式和一种采用焦点损失作为损失函数的卷积神经网络模型,将目标方位估计问题转化为多标签多分类问题来解决。通过卷积层提取阵列信号自协方差矩阵中的方位特征信息,最终由全连接层输出目标方位的估计结果。

本文对比了不同训练方法与损失函数对模型训练过程与估计结果的影响,研究发现:

(1) 采用声源强度不等的数据集进行训练可有效提高卷积神经网络模型对弱目标的检测能力。

(2) 采用焦点损失作为损失函数能增加模型对弱目标与背景噪声之间的区分能力。焦点损失函数训练得到的模型对弱目标的方位估计能力优于采用二元交叉熵损失函数训练出的模型。

本文所提方法收敛速度更快,检测概率更高,可以解决深度学习网络模型在多目标方位估计中弱目标方位估计困难的问题。试验数据处理结果也验证了本文所提方法的优越性。

参考文献

- [1] [美]第斯. 最优阵列处理技术[M]. 北京: 清华大学出版社, 2008: 863-872.
- [2] ZHU W L, ZHANG M. A deep learning architecture for broadband DOA estimation[C]//2019 IEEE 19th International Conference on Communication Technology (ICCT). Xi'an, China. IEEE, 2020: 244-247.
- [3] KASE Y Y, NISHIMURA T, OHGANE T, et al. DOA estimation of two targets with deep learning[C]//2018 15th Workshop on Positioning, Navigation and Communications (WPNC). Bremen, Germany. IEEE, 2018: 1-5.
- [4] KASE Y Y, NISHIMURA T, OHGANE T, et al. Performance analysis of DOA estimation of two targets using deep learning [C]//2019 22nd International Symposium on Wireless Personal Multimedia Communications (WPMC). Lisbon, Portugal. IEEE, 2020: 1-6.
- [5] ZHANG M, PAN X, SHEN Y N, et al. Deep learning-based direction-of-arrival estimation for multiple speech sources using a small scale array[J]. The Journal of the Acoustical Soci-

- ety of America, 2021, **149**(6): 3841.
- [6] LIU C L, FENG W K, LI H B, et al. Single snapshot DOA estimation based on spatial smoothing MUSIC and CNN[C]//2021 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). Xi'an, China. IEEE, 2021: 1-5.
- [7] OZANICH E, GERSTOFT P, NIU H Q. A feedforward neural network for direction-of-arrival estimation[J]. The Journal of the Acoustical Society of America, 2020, **147**(3): 2035.
- [8] LIU Y J, CHEN H X, WANG B. DOA estimation based on CNN for underwater acoustic array[J]. Applied Acoustics, 2021, **172**: 107594.
- [9] PAPAGEORGIOU G K, SELLATHURAI M, ELDAR Y C. Deep networks for direction-of-arrival estimation in low SNR [J]. IEEE Transactions on Signal Processing, 2021, **69**: 3714-3729.
- [10] 刘文哲, 柯雨璇, 郑成诗, 等. 基于卷积神经网络和改进聚类算法的多声源定位[C]//2019年全国声学大会论文集. 深圳, 2019: 485-486.
- [11] 曹怀刚, 任群言, 郭圣明, 等. 卷积神经网络单矢量水听器方位估计[J]. 哈尔滨工程大学学报, 2020, **41**(10): 1524-1529.
- CAO Huaigang, REN Qunyan, GUO Shengming, et al. Source azimuth estimation with single vector sensor based on convolutional neural network[J]. Journal of Harbin Engineering University, 2020, **41**(10): 1524-1529.
- [12] WHITAKER S, BARNARD A, ANDERSON G D, et al. Recurrent networks for direction-of-arrival identification of an acoustic source in a shallow water channel using a vector sensor[J]. The Journal of the Acoustical Society of America, 2021, **150**(1): 111.
- [13] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, **42**(2): 318-327.
- [14] (印)纳温·库马尔·马纳西(Navin Kumar Manaswi), 著. Python深度学习实战: 基于TensorFlow和Keras的聊天机器人以及人脸、物体和语音识别[M]. 刘毅冰, 薛明, 译. 北京: 机械工业出版社, 2019: 1-2.