



云南民族大学学报(自然科学版)

Journal of Yunnan Minzu University(Natural Sciences Edition)

ISSN 1672-8513,CN 53-1192/N

## 《云南民族大学学报(自然科学版)》网络首发论文

题目: CV-CNN 与稀疏贝叶斯学习结合的声源定位方法研究  
作者: 邢传玺, 崔晶, 魏光春, 董赛蒙  
网络首发日期: 2024-06-06  
引用格式: 邢传玺, 崔晶, 魏光春, 董赛蒙. CV-CNN 与稀疏贝叶斯学习结合的声源定位方法研究[J/OL]. 云南民族大学学报(自然科学版).  
<https://link.cnki.net/urlid/53.1192.n.20240605.1752.007>



**网络首发:** 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

**出版确认:** 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

# CV-CNN 与稀疏贝叶斯学习结合的声源定位方法研究

邢传玺，崔晶，魏光春，董赛蒙  
(云南民族大学 电气信息工程学院，昆明 650500)

**摘要：**针对现有水下目标定位算法多数是在声源数目已知的情况下对目标进行定位，而当声源数目未知或估计错误时会出现目标定位不准确甚至无法定位的问题。本文采用复数卷积神经网络（Complex-Valued Convolutional Neural Networks, CV-CNN）与稀疏贝叶斯学习相结合的声源定位方法。首先在声源数目预测阶段，利用神经网络学习传感器接收数据与声源数目之间的关系，估计未知声源的数目。然后在声源定位阶段，在估计出声源数目的基础上采用离格稀疏贝叶斯学习方法对目标进行定位。仿真结果表明，采用 CV-CNN 的网络模型对传感器在不同信噪比下混合数据集的声源数量估计准确率达到 99.16%，在估计出声源数目的基础上，本文方法在信噪比为-5dB 时的均方根误差在  $1^\circ$  以内，在快拍数为 100 时的均方根误差也保持在  $1^\circ$  以内。

**关键词：**阵列信号处理；深度学习；离格稀疏贝叶斯学习；DOA 估计

**中图分类号：**TB567

## CV-CNN combined with Sparse Bayesian Learning for Sound Source Localization

XING Chuan-xi, CUI Jing, WEI Guang-chun, DONG Sai-meng

(School of Electrical and Information Technology, Yunnan Minzu University, Kunming 650500, China)

**Abstract:** Most of the existing algorithms for underwater target localization focus on scenarios where the number of sound sources is known, which can lead to issues of inaccurate or even unattainable target localization when the number of sound sources is unknown or inaccurately estimated. In this study, we propose a sound source localization approach that integrates Complex-Valued Convolutional Neural Networks (CV-CNN) with Bayesian learning. In the initial stage of sound source prediction, a neural network is employed to establish the correlation between the received data from the sensor array and the number of sound sources, enabling an estimation of the unknown sound source count. Subsequently, in the sound source localization phase, the off-grid sparse Bayesian learning technique is employed to pinpoint the target based on the previously estimated sound source count. Simulation results underscore the effectiveness of the proposed method. The network model employing CV-CNN achieves a remarkable accuracy of 99.16% in estimating the number of sound sources within a mixed dataset acquired from the receiving array under varying signal-to-noise ratios. Moreover, building upon the estimated sound source count, the method presented in this paper maintains a root-mean-square error below  $1^\circ$  at a signal-to-noise ratio of -5 dB. This level of accuracy is also maintained with a snap count of 100.

**基金项目：**国家自然科学基金（61761048）；云南省基础研究专项面上项目（202101AT070132）。

**作者简介：**邢传玺（1982-），男，博士，教授。主要研究方向：水声信号处理。

**通讯作者：**崔晶（1998-），女，硕士研究生。主要研究方向：水声定位。

**Key words:** Array signal processing; Deep learning; Off-grid sparse Bayesian learning; DOA estimation

## 0 引言

水下定位<sup>[1]</sup>是指使用水听器接收来自声源的信息,以预测声源的位置。水下声源定位可分为主动定位和被动定位两大类,被动定位隐蔽性好、安全性高,应用领域更加广泛。然而水下声源由于具有种类繁多且不确定的特点,给声源识别和定位带来了挑战。因此,对水下声源源数目的准确估计就显得至关重要,准确知道水下声源源数目不仅可以提高定位精度,而且对于避免潜艇碰撞、保障安全都具有特殊的意义。

最初进行数目估计的基本思想为假设检验,这类方法需要预先设置,受主观影响很大,并且随着需要估计的数据量增大,其准确度急剧下降,估计效率极低。大多数经典的信源数估计方法都侧重估计信号的空间参数,由阵列信号处理的空间谱估计的相关理论可知,信号成分用空间协方差矩阵的较大特征值表示,噪声成分用较小特征值表示<sup>[2]</sup>。经过多年的发展,形成了基于信息论、基于盖氏圆盘和基于奇异值分解的源信号数目估计理论体系。机器学习在分类问题上的突出表现使得研究学者试图将源信号数目的估计也看作一个分类问题,因此引入机器学习的手段来解决。Lee Y 探索基于机器学习的聚类算法生成归一化特征值间隙的数据集,推导出信源数目估计的最佳阈值<sup>[3]</sup>,并通过实验证明,如果能提供适当的学习数据集,可以提高源数目估计的性能。

DOA 技术最早提出于 20 世纪 60 年代,在水下声学中常用的 DOA 估计方法主要有常规波束形成法<sup>[4]</sup>、多重信号分类(MUSIC)法<sup>[5]</sup>等和旋转不变技术算法<sup>[6]</sup>。近年来,稀疏贝叶斯学习<sup>[7]</sup>(SBL)由于具有强噪声鲁棒性且能实现高精度的方位估计,也逐渐被应用于水声领域。然而,海洋环境中存在着各种未知噪声,如风成噪声、生物发声,面对这种有色噪声,传统 DOA 估计方法的性能会严重下降。为了解决这一问题,李秀坤等人确定了阵列增益与阵元数量、噪声和二阶统计特性等因素之间的关系,发现使用四阶累积量波束形成的方法可以获得更高的阵增益和分辨率<sup>[8]</sup>。Fu 等人在 2019 年针对声学信号在时频域的稀疏特性,提出了一种基于深度学习和卷积非负矩阵分解相结合的盲 DOA 估计方法<sup>[9]</sup>。Kase 等人在 2022 年提出一种基于深度学习的方法来提高 DOA 估计的准确性<sup>[10]</sup>。然而,这些方法只适用于有固定信号源数目的情景,应用场景较为单一且具有较大局限性,无法在声源数目未知的情况下实现对水下目标的定位。

针对上述问题,本文提出了一种结合深度学习与离格稀疏贝叶斯学习方法的两阶段声源定位方法,实现在未知声源数目的情况下对水下目标的高精度 DOA 估计。本文主要内容如下:

(1) 考虑到水下声源种类繁多和数量的不确定,提出了基于 CV-CNN 的声源数目预测模型,通过学习阵列在不同时间接收到的水声信号的特征,预测声源数目。

(2) 考虑在高斯噪声条件下,传统的 DOA 估计方法受噪声干扰明显,在低信噪比情况下目标方位估计不准确的问题,在估计出声源数目的基础上,选择离格稀疏贝叶斯学习方法进行水下目标定位。

## 1 理论基础

### 1.1 信号模型

水声信号处理中，水听器阵列是由一组水听器按照某种规则或空间排列，或特定数量排列形成的。图 1 展示了一个均匀水听器阵列的示意图，该阵列由  $M$  个水听器组成，这些水听器按照等间距方式排列。在平面波假设条件下， $K$  个远场窄带不相关声源信号  $s(t)=[s_1(t),s_2(t),\cdots,s_K(t)]^T$  入射到该阵列上，入射方向与各个阵元的法线方向的夹角为  $\theta$ ， $\theta=[\theta_1,\theta_2,\cdots,\theta_K]^T$ ，其中  $\theta_k$  表示第  $K$  个信号的入射角。在时刻  $t$ ，阵列的输出矢量  $x(t)=[x_1(t),x_2(t),\cdots,x_M(t)]^T$  表示为：

$$x(t) = A(\theta)s(t) + n(t) \quad (1)$$

式中， $A(\theta)=[a(\theta_1),a(\theta_2),\cdots,a(\theta_K)]$  为  $K$  个入射信号对应的阵列流形矩阵， $a(\theta_k)$  表示角度  $\theta_k$  对应的阵列响应矢量，且

$$A(\theta) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-j\frac{2\pi d}{\lambda}\sin\theta_1} & e^{-j\frac{2\pi d}{\lambda}\sin\theta_2} & \cdots & e^{-j\frac{2\pi d}{\lambda}\sin\theta_K} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j\frac{2\pi(M-1)d}{\lambda}\sin\theta_1} & e^{-j\frac{2\pi(M-1)d}{\lambda}\sin\theta_2} & \cdots & e^{-j\frac{2\pi(M-1)d}{\lambda}\sin\theta_K} \end{bmatrix} \quad (2)$$

$n(t)=[n_1(t),n_2(t),\cdots,n_M(t)]^T$  表示水听器阵列接收到的噪声。当每个信号源的快拍数为  $T$  时，式(1)的单快拍模型变为多快拍模型可写为：

$$X = AS + N \quad (3)$$

式中， $X=[x(1),x(2),\cdots,x(T)]$  表示水听器阵列接收到的信号矩阵， $S=[s(1),s(2),\cdots,s(T)]$  表示入射信号矢量矩阵， $N=[n(1),n(2),\cdots,n(T)]$  表示噪声信号矢量矩阵。

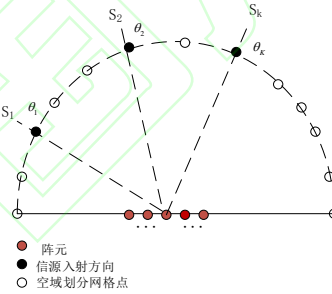


图 1 信号入射阵列模型图

## 1.2 复数卷积神经网络

复数卷积神经网络主要由复数卷积、复数激活、复数批量归一化和池化运算构成。其中池化运算与实数卷积神经网络相同。

### 1.2.1 复数卷积

定义被卷积张量  $h = x + iy$ ，复数卷积核  $W = A + iB$ ，卷积运算的结果可表示为：

$$W * h = (A * x - B * y) + i(B * x + A * y) \quad (4)$$

图 2 为进行复卷积运算示意图，使用矩阵来表示卷积运算的实部和虚部，则式（4）表示为：

$$\begin{bmatrix} R(W * h) \\ I(W * h) \end{bmatrix} = \begin{bmatrix} A & -B \\ B & A \end{bmatrix} * \begin{bmatrix} x \\ y \end{bmatrix} \quad (5)$$

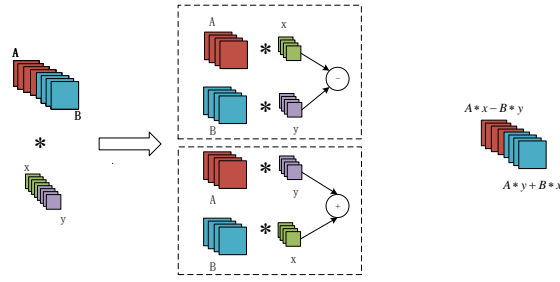


图2 复卷积运算示意图

### 1.2.2 复数激活函数

激活函数应用于特征向量上增强其非线性表达能力，在复数域中选择 CReLU 函数分别对复数的实部和虚部进行激活，CReLU 表达式如下：

$$CReLU(x) = ReLU(Re(x)) + iReLU(Im(x)) \quad (6)$$

### 1.2.3 复数批量归一化

批量归一化能将每层网络输入值的分布拉回到标准正态分布，复数批量归一化表达式如下式(7)所示：

$$BN(x) = \gamma x + \beta \quad (7)$$

式中  $\gamma = \begin{bmatrix} \gamma_{rr} & \gamma_{ri} \\ \gamma_{ir} & \gamma_{ii} \end{bmatrix}$ ，包含 3 个可学习参数  $\gamma_{rr}$ 、 $\gamma_{ri} = \gamma_{ir}$ 、 $\gamma_{ii}$ ； $\beta = \beta^R + i\beta^I$  是一个复数，包含

$\beta^R$ 、 $\beta^I$  2 个可学习参数； $x = V^{-1/2}(x - E[x])$  且  $V = \begin{bmatrix} Cov(x^R, x^R) & Cov(x^R, x^I) \\ Cov(x^I, x^R) & Cov(x^I, x^I) \end{bmatrix}$ 。

## 2 基于复数卷积神经网络的水下声源数量估计

由于水听器阵列接收到的信号是一种时间序列数据，卷积神经网络能够捕捉信号的时序特征和变化，从而有效地提取与声源数目相关的重要信息。针对均匀线性阵列 (Uniform Linear Array, ULA) 接收数据的声源数目预测问题构造了基于 CV-CNN 的网络模型，该模型可以预测出声源数目，为接下来目标定位提供关键性信息。

### 2.1 构造数据集

将水听器阵列接收到的数据  $X$  作为网络模型的输入，将数据集按照 7: 2: 1 的比例分为训练集、验证集和测试集，数据标签(1, 2, 3)代表声源个数，首先在训练集上进行不断学习保存最优网络模型，然后在测试集上进行测试。

### 2.2 基于 CV-CNN 的声源预测的网络模型

基于 CV-CNN 的水下声源数量估计的网络模型如图 3 所示。其中，模型的输入为水下传感器接收到的复数矩阵；卷积层和池化层用来对复数矩阵的深层提取；Flatten 层、全连接层和 Softmax 分类器组成输出层，实现对声源数目的结果输出。具体的模型参数如表 1 所示。

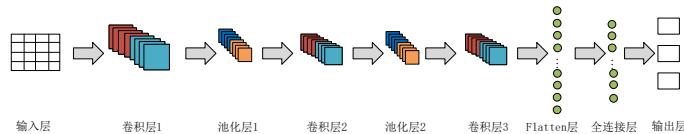


图3 CV-CNN 网络结构

表 1 CV-CNN 网络结构模型参数

网络层数	类型	输入	卷积核/神经元 量	步长	输出
1	输入层	8×100×1	/	/	/
2	卷积层 1	8×100×1	3×3	1×1	6×98×32
3	池化层 1	6×98×32	2×2	2×2	6×49×32
4	卷积层 2	6×49×32	3×3	1×1	4×47×64
5	池化层 2	4×47×64	2×2	2×2	2×23×64
6	卷积层 3	2×23×64	3×3	1×1	2×21×64
7	Flatten 层	2×21×64	2688	/	1×2688
8	全连接层	2688	64	/	1×64
9	Softmax	64	3	/	1×3

网络输出层选择 Softmax 激活函数，表达式如式(12)。通过 Softmax 函数，会返回一个由 3 个概率值组成的数组，这三个概率值和为 1。每个概率值用来表示当前水听器接收阵列接收到的数据属于三个声源数量中的某一个声源数目的概率。

$$\text{Soft max}(x_i) = e^{x_i} / \sum_{c=1}^C e^{x_c} \quad (8)$$

其中， $x_i$  是输出层中第  $i$  个神经元的值， $C$  是输出层的神经元数量。Softmax 函数输出声源数目的预测概率，网络将具有最大预测概率的源的数量视为预测结果，并据此确定声源数目  $K$ ， $K$  在接下来的离格稀疏贝叶斯学习中，用于控制目标定位结果的稀疏程度，即决定估计结果中的非零元素数量。

### 3 基于离格稀疏贝叶斯学习的测向算法

水下传感器在  $t$  时刻接收到信号的四阶累积量可以表示为<sup>[11-12]</sup>：

$$\begin{aligned}
C(i, j, n, l) &= \text{cum}\{x_i(t), x_j(t), x_n^*(t), x_l^*(t)\} \\
&= \text{cum}\left\{\sum_{k=1}^K a_i(\theta_k) s_k(t), \sum_{k=1}^K a_j(\theta_k) s_k(t), \sum_{k=1}^K a_n^*(\theta_k) s_k^*(t), \right. \\
&\quad \left. \sum_{k=1}^K a_l^*(\theta_k) s_k^*(t)\right\} + \text{cum}\{n_i(t), n_j(t), n_n^*(t), n_l^*(t)\} \\
&= \sum_{k=1}^K a_i(\theta_k) a_j(\theta_k) a_n^*(\theta_k) a_l^*(\theta_k) \text{cum}(s_k(t), s_k(t), s_k^*(t), s_k^*(t)),
\end{aligned} \quad (9)$$

式中， $(\cdot)^*$  表示共轭， $a_i(\theta_k)$  表示阵列流形矩阵  $A(\theta)$  的第  $i$  行、第  $k$  列，cum 表示累积量，变量  $i, j, n, l \in \{1, 2, \dots, M\}$ 。根据高斯噪声三阶累积量以上均为零的理论可以进一步简化模型。进而，从稀疏重构的角度出发，由于四阶累积量维度高且存在很多零值，所以两次构造选择矩阵对其进行降维，从而去除冗余元素，只保留带有信息的非零元素，得到模型的稀疏表示：

$$\mathbf{Z} = \mathbf{D}(\theta)\Gamma \quad (10)$$

式中， $\mathbf{D}(\theta) = [d(\theta_1), d(\theta_2), \dots, d(\theta_N)]$  为过完备的阵列流形矩阵， $\Gamma$  中列向量大部分元素为零，非零元素对应估计角度的网格点位置。而在实际情况中，存在目标离格情况，因此采用一阶泰勒公式来近似真实的离格点，构造新的阵列流型过完备矩阵，其表示为：



$$\Phi = \tilde{D} + \tilde{G} \text{diag}(\beta) \quad (11)$$

式中,  $\Phi = [\varphi(\tilde{\theta}_1), \varphi(\tilde{\theta}_2), \dots, \varphi(\tilde{\theta}_N)]$  为线性化处理后的阵列导向矢量,  $\tilde{\theta}_{n_k}$  表示距离目标方位  $\theta_k$  最近的网格点。  $\tilde{G} = [\tilde{g}(\tilde{\theta}_1), \tilde{g}(\tilde{\theta}_2), \dots, \tilde{g}(\tilde{\theta}_N)]$ , 其中  $\tilde{g}(\tilde{\theta}_{n_k})$  为  $\tilde{d}(\tilde{\theta}_{n_k})$  的一阶求导项, 即  $\tilde{g}(\tilde{\theta}_{n_k}) = \tilde{d}'(\tilde{\theta}_{n_k})$ 。  $\beta$  为网格修正向量, 定义  $\beta = [\beta_1, \beta_2, \dots, \beta_N]^T \in [-0.5r, 0.5r]^N$  表示目标位置与网格点之间的失配误差,  $r$  表示相邻网格的划分间距, 则可以得到网格修正向量为:

$$\beta_n = \begin{cases} \theta_k - \tilde{\theta}_{n_k}, & \tilde{I}_n \neq 0, n = n_k, \forall k \in \{1, 2, \dots, K\} \\ 0, & \tilde{I}_n = 0, n \neq n_k \end{cases} \quad (12)$$

由此, 得到离格稀疏 DOA 估计模型为:

$$Z = \Phi \Gamma \quad (13)$$

为了解决离格稀疏模型的 DOA 估计问题, 将稀疏贝叶斯学习应用于稀疏信号恢复中。将水下载感器接收到的信号视为是随机观测的样本数据, 并将待恢复的稀疏信号看作贝叶斯估计中的位置参数, 运用贝叶斯学习来求解 DOA 估计的全局最优解。

假设噪声服从复高斯分布, 则  $Z$  的似然函数的表达式为:

$$p(Z | \Gamma; \alpha_0) = CN(Z | \Phi \Gamma, \alpha_0^{-1}) \quad (14)$$

式中,  $\alpha_0 = \sigma^{-2}$ , 噪声的方差为  $\sigma^2$ 。CN 代表复高斯分布, 假设  $\alpha_0$  是未知参量, 并且  $\alpha_0$  满足伽马 (Gamma) 分布。

假设稀疏信号  $\Gamma$  的先验概率为:

$$p(\Gamma | \alpha) = CN(\Gamma | 0, \Lambda) \quad (15)$$

式中,  $\alpha$  是一个超参数集合, 代表水听器阵列接收到的不同空域角度的信号源功率,  $\Lambda = \text{diag}(\alpha)$ 。

结合先验概率和似然函数, 可以得到联合概率密度函数为

$$p(\Gamma, Z, \alpha_0, \alpha, \beta) = p(Z | \Gamma, \alpha_0, \beta) p(\Gamma | \alpha) p(\alpha) p(\alpha_0) p(\beta) \quad (16)$$

得到  $\Gamma$  的后验概率密度函数为:

$$p(\Gamma | Z, \alpha_0, \alpha, \beta) = CN(\Gamma | \mu, \Sigma) \quad (17)$$

其中, 根据推导可得到后验方差、均值为:

$$\begin{aligned} \mu &= \alpha_0 \Sigma \Phi^H Z \\ \Sigma &= (\alpha_0 \Phi^H \Phi + \Lambda^{-1})^{-1} \end{aligned} \quad (18)$$

$\Sigma$  和  $\mu$  是关于超参数  $\alpha_0$ ,  $\alpha$  和  $\beta$  的函数, 采用最大化后验概率准则最大化  $p(\alpha_0, \alpha, \beta | Z)$  来估计  $\alpha_0, \alpha$  和  $\beta$  三个参数, 则根据 EM 算法, 可以得到  $\alpha$  和  $\alpha_0$  的更新公式为:

$$\alpha_n^{\text{new}} = \frac{\sqrt{1 + 4\rho E\{\|\Gamma^n\|_2^2\}} - 1}{2\rho}, n = 1, 2, \dots, N \quad (19)$$

$$\alpha_0^{new} = \frac{M + c - 1}{E\{\|Z - \Phi\Gamma\|_F^2\}} \quad (20)$$

且  $E\{\|\Gamma^n\|_2^2\} = \|\mu^n\|_2^2 + \Sigma_m$ ,  $E\{\|Z - \Phi\Gamma\|_F^2\} = \|Z - \Phi\mu\|_F^2 + \alpha_0^{-1} \sum_{n=1}^N (1 - \alpha_n^{-1} \Sigma_m)$ ,  $\rho$  为

取值很小的正约束。

网格误差  $\beta$  由最大化  $E\{\lg \rho(Z|\Gamma, \alpha_0, \alpha, \beta) p(\beta)\}$  得到，相当于最小化：

$$E\{\|Z - (D + G \text{diag}(\beta))\Gamma\|_2^2\} = \beta^T P \beta - 2q^T \beta + C \quad (21)$$

式中，C 为常量，其值取决于  $\beta$ ，P 为半正定矩阵，P 和 q 的表达示为：

$$P = \text{Re}\{G^H G \odot (\mu\mu^H + \Sigma)\} \quad (22)$$

从而得到采样点的离格误差更新矩阵为：

$$\beta^{new} = \arg \min_{\beta \in [-0.5r, 0.5r]} \{\beta^T P \beta - 2q^T \beta\} \quad (23)$$

由式 (25) 对  $\beta$  求导并令其导数为零，且由  $\beta_n \in [-0.5r, 0.5r]$  可以得到离格误差的表达式为：

$$\beta_n^{new} = \begin{cases} \beta_n, & \beta_n \in [-0.5r, 0.5r]; \\ -0.5r, & \beta_n < -0.5r \\ 0.5r, & \text{else} \end{cases} \quad (24)$$

最后，初始化参数  $\alpha_0$ ， $\alpha$  和  $\beta$ ，不断迭代计算式并更新参数值，直到收敛，得到 K 个信源的 DOA 估计值，即  $\hat{\theta}_k = \hat{\theta}_{nk} + \beta_{nk}, k=1, 2, \dots, K$ 。

综上所述，得到水下目标方位估计的总体流程图，如图 4 所示：

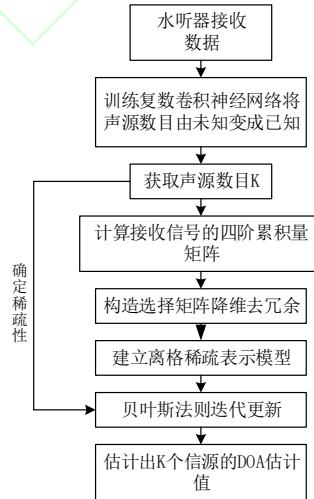


图 4 总体算法流程图

#### 4 数值仿真分析

为了验证 CV-CNN 与离格稀疏贝叶斯学习相结合的水下目标定位方法的性能。本节将进行仿真分析，仿真实验中，若无特殊说明，以下参数和初始值的设置不变：假设均匀线性



阵列由 8 个完全一致的传感器组成，相邻传感器之间的距离为 0.5 米，空间角度划分为  $[-\pi/2, \pi/2]$ ，网格间距  $r=1^\circ$ ，声源是中心频率为 1000Hz 的线性调频信号，定义信噪比为  $SNR=10\lg(P_s/P_n)$ ，其中  $P_s$  表示信号功率、 $P_n$  表示噪声功率。通过均方根误差（RMSE）衡量所使用的 DOA 估计方法的定位结果与真实目标方位之间的差距来衡量算法的性能，均方根误差的数学表达式为  $RMSE = \sqrt{\frac{1}{KS} \sum_{s=1}^S \sum_{k=1}^K (\hat{\theta}_{k_s} - \theta_k)^2}$  其中，S 为蒙特卡罗实验的次数，K 为信源的个数， $\hat{\theta}_{k_s}$  为第 s 次实验中第 k 个信源的方位估计值。

#### 4.1 基于 CV-CNN 的声源数目预测模型的仿真与性能分析

##### 4.1.1 结果分析

基于 CV-CNN 的水下声源数量的估计在测试集上的估计准确率的混淆矩阵如图 5 所示。

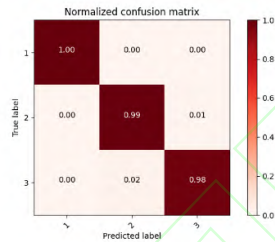


图 5 CV-CNN 在测试集上的准确率的混淆矩阵

混淆矩阵中方格上的数值代表估计准确率，对角线上的数值即代表正确估计的概率，从图 5 中可以看出随着声源数目的增多，声源数目的估计准确率有所降低。当仅有 1 个声源时，此时模型估计结果最为准确，准确率达到 100%。，与单目标相比，多目标的声源估计准确率有所降低，但整体上该网络模型对于声源数目估计的准确率达到 98%及以上，可信度较高。

##### 4.1.2 性能分析

为了更好地验证 CV-CNN 模型在不同声源数目(1, 2, 3)和不同信噪比(-10、-5、0、5、10、15)的性能，将分别对其进行测试。将传感器接收到的数据作为网络模型的输入，标签为 1、2、3 代表声源数量。通过准确率衡量算法的输出结果与预期结果之间的差距来衡量算法的性能，准确率的数学表达式定义为  $acc = num_{K_{pre}=K_{real}} / num_{\text{试验次数}}$ ，即正确估计的数量占总估计数量的比例。其中， $K_{pre}$  代表该网络模型预测的声源数量， $K_{real}$  代表真实的声源数量，当  $K_{pre} = K_{real}$  时代表该网络模型估计的声源数量正确。

验证结果如表 2 所示，从表中可以看出，随着信噪比的增加，网络模型对声源数目的预测精度增大，在较低信噪比-10dB 的情况下，网络模型对于单目标的声源数目的预测精度为 100%，对于多目标的声源数目预测也达到 95%及以上，结果较为可信。由此可以看出，该网络模型具有较高的精度，可以准确地估计出声源数目 K。

表 2 不同信噪比情况下 CV-CNN 的预测精度

声源数目	SNR/dB					
	-10	-5	0	5	10	15
1	100	100	100	100	100	100
2	97	100	100	100	100	100

## 4.2 基于离格稀疏贝叶斯学习测向算法的仿真与性能分析

本节通过数值仿真分析,在使用 CV-CNN 网络模型来估计出水下声源数目的基础上,将本文算法与较为经典的算法(MUSIC-like 算法、OGSBI-SVD 算法)进行比较分析。

### 4.2.1 空间谱及均方根误差分析

图 6 给出了在信源入射方向为 $[-30.3^\circ \ 20.6^\circ]$ ,输入信噪比  $\text{SNR}=3\text{dB}$  以及快拍数  $T=100$  的条件下,三种算法(MUSIC-like 算法、OGSBI-SVD 算法和本文算法)在相同条件下的 DOA 空间谱估计图。由理论分析可知,在声源数目未知时,MUSIC-like 和 OGSBI-SVD 算法无法通过特征值分解和奇异值分解的方法来实现对水下目标的方位估计。而在估计出水下声源数目的基础上,这三种算法均能估计出水下目标的位置,但本文所采用的算法具有两个较为明显的谱峰,且主瓣宽度较窄,与声源数目的估计结果吻合;OGSBI-SVD 算法对目标方位的估计结果主瓣宽度较宽且存在伪峰,估计结果可能不准确;MUSIC-like 算法对目标方位的估计主瓣宽度较宽且波峰不明显。

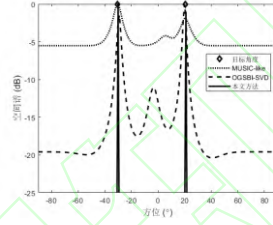


图 6 空间谱估计图

图 7 为三种算法在阵元输入信噪比  $\text{SNR}=-5:1:10\text{dB}$ ,快拍数  $T=100$  的条件下对于入射方向为 $[-40.6^\circ \ 10.3^\circ]$ 的 DOA 估计的均方根误差。为保证算法的可靠性,在对水下声源数目进行估计的基础上,进行 200 次蒙特卡罗实验。从图 7 中可以看出,MUSIC-like 算法的 RMSE 值高于另外两种算法;随着信噪比的增加,OGSBI-SVD 算法的 RMSE 值逐渐降低,显示出较好的性能;而本文提出的算法在 DOA 估计中表现出较低的 RMSE 值,显示出更高的准确性,对目标方位的估计精度误差在  $1^\circ$  以内,估计精度高于其他两种算法。

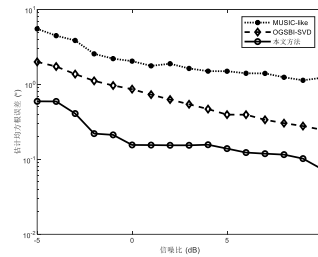


图 7 RMSE 随信噪比变化图

图 8 为三种算法在信噪比  $\text{SNR}=2\text{dB}$ ,快拍数  $T=100:100:1000$  的条件下对于入射方向为 $[-10.1^\circ \ 22.4^\circ]$ 的 DOA 估计的均方根误差。为保证算法的可靠性,在对水下声源数目进行估计的基础上,进行 200 次蒙特卡罗实验。从图 8 中可以看出,随着快拍数目的增大,这三种算法的 RMSE 整体上逐渐降低。而本文算法具有更低的 RMSE 值,其估计误差在  $1^\circ$  以内,能更准确地估计出目标的角度信息。

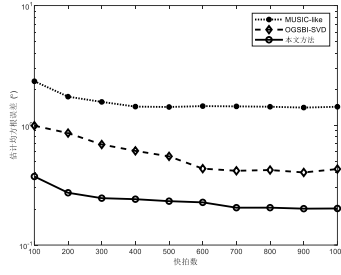


图 8 RMSE 随快拍数变化图

为验证采样角度间隔对 DOA 估计精度的影响, 设置入射角度为 $[-15.3^{\circ} \ 35.6^{\circ}]$ , 输入信噪比  $\text{SNR}=-5:1:10\text{dB}$ , 快拍数  $T=100$ , 在估计出声源数目的基础上, 进行 200 次蒙特卡罗实验, 图 9 展示了在角度间隔  $r=[0.5^{\circ} \ 1^{\circ} \ 2^{\circ}]$  时的均方根误差变化曲线。从图 9 可以观察到, 网格间距越小, RMSE 值越低, 同时随着信噪比增加, 每个角度间隔的 RMSE 值也呈现下降的趋势, 当角度间隔为  $2^{\circ}$  时, 本文方法仍然能够保持在  $1^{\circ}$  以内的 RMSE 值, 其估计结果相对准确。

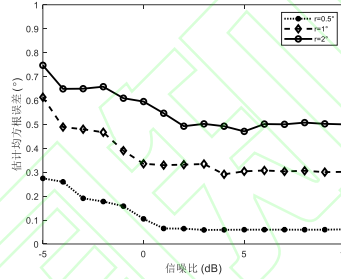


图 9 RMSE 在不同网格间距下随信噪比的变化图

#### 4.2.2 相邻声源分辨力

考虑当多个声源紧密分布时三种算法的性能, 在其他条件保持不变的情况下, 目标信号入射角度为 $[-3.9^{\circ} \ 4.3^{\circ}]$ , 角度间隔  $r=1^{\circ}$ , 考虑在信噪比  $\text{SNR}=0$  以及快拍数  $T=200$  的情况下三种算法的 DOA 空间谱估计结果。从图 10 中可以看出, 本文方法能够清晰地分辨出相近的两个声源, 主瓣宽度较窄, 与声源估计数目的结果吻合; MUSIC-like 算法仅有一个谱峰, 与实际的水下声源数目不符, 目标定位不准确; OGSBI-SVD 算法在无法分辨位置距离较近的声源。

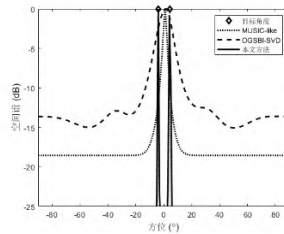


图 10 相邻声源空间谱估计图

## 5 结论

考虑到水下环境复杂, 噪声干扰过大和水下声源数目的不确定性给水下目标定位带来的困难, 本文在声源数目估计和声源定位两个方面进行了研究。该方法首先采用 CV-CNN 网络不断学习水下传感器接收到的信号和声源数目之间的关系, 先估计出水下声源的数目, 然

后在此基础上采用基于离格稀疏贝叶斯学习方法实现对水下目标的角度估计。该方法的关键在于可以实现在未知声源数目的情况下,对水下目标高精度的 DOA 估计。数值仿真分析表明,信噪比、快拍数、声源方位和网格间距划分等因素对本文方法的影响较小,可实现对多目标的方位估计,估计精度更高。这不仅为实现水下目标高精度快速稳定定位提供了新的思路,也为解决在水下目标定位中当声源数目未知或估计不准确时所带来的问题提供了一个有效的解决途径。

最后,本文感谢昆明人工智能计算中心提供支持。

#### 参考文献

- [1] 黄海宁,李宇.水声目标探测技术研究现状与展望[J].中国科学院院刊, 2019,34(03):264-271
- [2] 梅超,潘晴,常桂林等.一种信源数目估计的改进算法[J].信息通信,2018,182(02):26-29.
- [3] Lee Y, Park C, Kim T. Source Enumeration Approaches Using Eigenvalue Gaps and Machine Learning Based Threshold for Direction-of-Arrival Estimation[J]. Applied Sciences, 2021, 11(04):1942.
- [4] Li Qihu, Wei Chonghua, Xue Shanhua. Iterative inverse beamforming algorithm and its application in multiple targets detection of passive sonar[J].Chinese Journal of Acoustics,2017,36(02):208-216.
- [5] 马思强,李雄风,许衍.多重信号分类算法用于通信信号测向的仿真分析[J].航天电子对抗,2021,37(04):22-28.
- [6] 杨德森,朱中锐,时胜国,莫世奇.声矢量圆阵相位模态域目标方位估计[J].声学学报,2014,39(01):19-26.
- [7] 游康勇,杨立山,刘玥良,郭文彬,王文博.基于稀疏贝叶斯学习的网格自适应多源定位[J].电子与信息学报,2018,40(09):2150-2157.
- [8] Li Xiukun, Li Tingting, Gu Xinyu, et al.Array gain of fourth-order cumulants beamforming under typical probability density background[J]. Chinese Journal of Acoustics,2015,34(01):15-26.
- [9] Fu Q, Jing B, He P. Blind DOA Estimation in a Reverberant Environment Based on Hybrid Initialized Multichannel Deep 2-D Convolutional NMF With Feedback Mechanism[J].IEEE Access, 2019, 7:179679-179689.
- [10] Kase Y, Nishimura T, Ohgane T. Accuracy Improvement in DOA Estimation with Deep Learning[J]. IEICE Transactions on communications, 2022, 105(05): 588-599.
- [11] 刘庆华,周秀清,晋良念.相关色噪声下无冗余累积量稀疏表示 DOA 估计[J]. 航空学报,2017,38(4):10.
- [12] Sheng Liu, Jing Zhao, Yu Zhang. Array manifold matching algorithm based on fourth-order cumulant for 2D DOA estimation with two parallel nested arrays[J]. International Journal of Computational Science and Engineering,2021,24(2).丁文祥. 数字革命与竞争国际化[N]. 中国青年报, 2000-11-20(15).