

DOI:10.20189/j.cnki.CN/61-1527/E.202604008

基于自适应特征增强的时域卷积网络-门控循环单元 单通道盲源分离网络

彭瑞琰, 郭文普*, 高绍原, 牟宸琛

(火箭军工程大学, 西安 710025)

摘要:针对单通道盲源分离在通信信号极端同频混叠和高阶调制场景中的长时依赖建模不足及特征捕捉能力弱的问题,提出一种基于自适应特征增强的时域卷积网络-门控循环单元(temporal convolutional network-gated recurrent unit, TCN-GRU)单通道盲源分离网络。该架构由4个核心模块构成:多核深度卷积编码器、自适应多尺度特征注意力(adaptive multi-scale feature attention, AMFA)模块、堆叠式自适应 TCN 与双向 GRU,以及多尺度解码器。首先,通过多核深度卷积编码器实现可靠的场景适配性;其次,利用 AMFA 模块动态调整通道权重,聚焦关键差异特征;随后,采用 TCN-GRU 融合方式捕捉信号的多尺度依赖关系;最后,多尺度解码器保证信号重构完整性。在包含同/异调制与符号率的4类典型混叠场景下开展仿真实验,结果表明:在信噪比为 $-10\sim 25$ dB 范围内,分离信号波形与原始信号高度吻合,信噪比超过 10 dB 后皮尔逊相关系数均接近 1;针对同频混叠程度的鲁棒性测试显示,当载频间隔低至 250 Hz 时,各场景尺度不变信噪比(scale invariant signal-to-noise ratio, SI-SNR)仍保持在 22 dB 以上,误码率(bit error rate, BER)处于可用范围,有效工作边界可延伸至 250 Hz;与 Conv-TasNet、DPRNN 等先进模型相比,高信噪比环境下 SI-SNR 分别提升 3.01 dB、9.42 dB,最优场景达 34.55 dB;消融实验进一步表明,所提 GRU 模块使 BER 显著降低;AMFA 模块在中高 SNR 区间可为模型带来 1~2 dB 的 SI-SNR 增益;充分验证了各模块对通信解调性能的实质性贡献。

关键词:单通道盲源分离;门控循环单元;同频混叠;时域卷积

中图分类号: TN911.7

文献标志码: A

开放科学标识码(OSID):

文章编号: 2097-2741(2026)04-0111-15



听语音
聊科研
与作者互动

随着无线通信技术迅猛发展,电磁频谱资源的拥挤导致信号混叠已成为常见问题。在军事侦察、频谱监测及认知无线电等领域,接收端需在未知信道参数及无源信号先验信息条件下,从单一观测通道提取多个通信信号源,这构成了单通道盲源分离(single-channel blind signal separation, SCBSS)的核心挑战。SCBSS 的性能直接关系到频谱资源的高效利用与通信安全,尤其在载频间隔极小、调制方式复杂的同频混叠场景下,其技术瓶颈更为突出。

早期的单通道盲源分离方法主要依赖于经典的信号处理理论,如稀疏分量分析(sparse component analysis, SCA)^[1],以及从多通道分离框架衍生出的独立分量分析(independent component analysis, ICA)^[2]变体等。其中,SCA 通常假设信号在时频域等变换域中具有稀疏性,但在同频时两信号时频结构高度重叠,稀疏性优势消失。经典 ICA 适用于多通道观测场景,为将其应用于单通道问题,学者们提出了诸如基于时间延迟嵌入的 ICA (time-

收稿日期: 2026-01-05

修回日期: 2026-01-29

录用日期: 2026-03-12

第一作者: 彭瑞琰(1998-),女,硕士研究生,主要从事单通道盲源分离研究。

* 通信作者: 郭文普(1976-),男,副教授,主要从事网络与信息安全以及通信信号分析与识别研究。

引用格式: 彭瑞琰,郭文普,高绍原,等. 基于自适应特征增强的时域卷积网络-门控循环单元单通道盲源分离网络[J]. 火箭军工程大学学报,2026,40(4):111-125. PENG Ruiyan, GUO Wenpu, GAO Shaoyuan, et al. Temporal convolutional network-gated recurrent unit single channel based on adaptive feature enhancement blind source separation network[J]. Journal of Rocket Force University of Engineering, 2026, 40(3):111-125.

delay embedding independent component analysis, TDE-ICA)、基于经验模态分解的ICA等多种变体^[3-4],其核心是通过信号重构构造虚拟多通道以满足ICA的输入条件,但同频混叠下统计特征更为接近,且高阶调制和成型滤波进一步弱化了信号独立性,导致其分离效果大幅降低。

近年来,深度学习的兴起为解决SCBSS问题开辟了新的途径,特别是在语音分离、通信信号的调制识别上展现出强大的特征提取能力和时序信号处理能力。Luo等^[5]提出时域音频分离网络(time-domain audio separation network, TasNet),采用“编码器-解码器”端到端架构,通过线性编码器将语音波形转换为高维特征表示,并结合时域掩码生成模块实现信号分离。但该框架仅针对语音信号进行优化,其特征提取与分离机制未考虑通信信号特有的载频、相位、符号率等关键参数差异。Luo等^[6]提出全卷积时域音频分离网络(convolutional time-domain audio separation network, Conv-TasNet),引入时序卷积网络替代循环结构,强化了语音信号的长时相关性建模,但仍未突破语音任务域的局限。Luo等^[7]提出双路径循环神经网络(dual-path recurrent neural network, DPRNN),通过分割-块处理-重叠相加实现长序列高效建模,在长时序信号分离中兼具轻量化与高性能,但难以捕捉微小载频/码率差异及相位累积偏移。Hou等^[8]提出跳跃连接增强型混合神经网络,通过GRU-based跳跃连接实现跨层级低维特征传输,解决了深度网络的特征退化问题,但模型的复杂度较高且仅针对自动相关监视广播信号,未扩展到雷达、通信等其他类型信号。Dai^[9]提出信号数量估计与分离网络,仅需输入“混合信号”即可完成分离,无需知道源数量、调制类型等信息,适配真实场景,但仅基于模拟信号(正弦、脉冲)进行了训练与验证。Ma等^[10]提出注意力卷积单元深度分离网络,基于“编码器-分离器-解码器”架构,核心为“大核卷积+改进全局上下文块”的注意力单元,实现了BPSK/BPSK、BPSK/QPSK等同/异调制混合信号的分离,但其对8PSK、8QAM等高阶调制信号的分离性能显著下降,且验证条件相对理想,缺乏对极小载频差、码率差等微弱可分特征的鲁棒性验证。Guo等^[11]提出复杂时域扩张卷积循环网络,创新性地直接处理复信号的I/Q分量而非拆分实虚部,避免了相位信息丢失,但其实验设置过于理想,限定混合信号载频一致,无初始相位偏移且无码率差异。这种设定虽

然能验证复数网络的基础有效性,但本质上回避了“极端同频混叠”场景中的最核心挑战。Fu等^[12]提出堆叠式时域盲分离卷积网络,通过融入高效通道注意力(efficient channel attention, ECA)机制,实现了PSK、PAM、QAM等多调制类型信号的分离,且在隐蔽通信、非合作通信场景中验证了该网络的有效性;在载频间隔为0.1 MHz、信噪比(signal-to-noise ratio, SNR)为10 dB时QPSK的符号错误率为 10^{-5} 量级,但该载频间隔与kHz级间隔相比混叠程度尚不极端。赵孟晨^[13]提出堆叠式时域卷积网络和堆叠式长短期记忆网络2种网络,通过空洞卷积、注意力机制强化特征提取,可处理10种混合调制信号,但未对信号相位差进行进一步研究。

综上,现有单通道盲源分离方法在处理同频混叠通信信号时,尤其是在载频间隔极小、存在码率差异及高阶调制的复杂场景下,仍面临两大核心挑战。1) 关键差异特征的提取与增强。混叠信号在时频域高度重叠,其可分性依赖于载波频偏、符号速率、初始相位等细微差异。如何设计一种机制,能自适应地从强噪声和另一路信号的干扰中聚焦并放大这些微弱但关键的区分性特征,是提升分离性能的瓶颈。2) 信号内在多尺度时序依赖的建模。一方面,码元内部及码间过渡体现了局部、平稳的短期依赖;另一方面,由载频偏移累积的相位旋转、不同符号速率导致的异步性则呈现出全局、非平稳的长期依赖。现有单一结构的网络(如RNN)难以同时高效捕捉这两种不同尺度的依赖关系^[14-18]。

针对上述挑战,本文提出一种基于自适应特征增强的时域卷积网络-门控循环单元(temporal convolutional network-gated recurrent unit, TCN-GRU)单通道盲源分离网络。该网络架构的核心思想在于:首先通过多核深度卷积捕获信号的复合特征;继而利用一种自适应注意力机制动态增强关键差异信息;最后,通过融合TCN与GRU的混合结构,协同建模信号的短时局部平坦性和长时非平稳依赖性,实现源信号精准重构。

1 盲源分离问题建模与方法理论分析

1.1 盲源分离问题建模

发射端通过星座映射与成型滤波完成基带信号构建。符号序列经调制映射为星座点,再经根升余弦滤波实现脉冲成型;基带信号通过载波调制上变频至射频频段,不同场景分别配置不同载波频

率、符号速率与时延参数。信道环节将两路射频信号线性叠加,并引入加性高斯白噪声模拟信道衰减^[19]。生成的信号模型可表示为

$$\mathbf{x}(t) = \sum_{i=1}^2 \alpha_i \mathbf{s}_i(t - \tau_i) e^{j(2\pi f_{c,i}t + \theta_i)} + \mathbf{n}(t) \quad (1)$$

式中: α_i 为第 i 路信号幅度; $\mathbf{s}_i(t - \tau_i)$ 为第 i 路带时延的基带成型滤波信号; t 为时间; τ_i 为时延; $f_{c,i}$ 为第 i 路信号的载波频率; θ_i 为第 i 路信号的初始相位; $\mathbf{n}(t)$ 为高斯噪声。接收端采用频率为 f_{L0} 的本振载波进行下变频,将射频混合信号与本振载波共轭相乘,此时第 i 路信号相对于本振载波的残余频偏定义为 $\Delta f_i = f_{c,i} - f_{L0}$ 。下变频后各信号的载波项由原射频载波转换为残余频偏对应的相位项,最终形成的多维度基带数据输出为

$$\mathbf{y}(t) = \mathbf{x}(t) \cdot e^{-j2\pi f_{L0}t} \quad (2)$$

将式(1)代入式(2),基带混合信号可进一步展开为

$$\mathbf{y}(t) = \sum_{i=1}^2 \alpha_i \mathbf{s}_i(t - \tau_i) e^{j(2\pi \Delta f_i t + \theta_i)} + \mathbf{n}_b(t) \quad (3)$$

式中: $\mathbf{n}_b(t) = \mathbf{n}(t) \cdot e^{-j2\pi f_{L0}t}$ 为下变频后的基带噪声。

1.2 方法理论分析

式(3)所描述的基带混合信号模型中,2路源信号 $\mathbf{s}_1(t)$ 和 $\mathbf{s}_2(t)$ 因非线性和线性运算深度耦合。盲源分离的目标是设计一个逆向映射函数,以从观测信号 $\mathbf{y}(t)$ 中恢复出原始信号的估计。该过程面临以下严峻挑战,这些挑战直接驱动了本文网络架构的设计。

1.2.1 特征的非线性深度耦合与解码需求

首先,将混合信号模型 $\mathbf{y}(t)$ 分解为 $\mathbf{y}_I(t)$ 、 $\mathbf{y}_Q(t)$ 。令源信号 $\mathbf{s}_i(t) = \mathbf{s}_{iI}(t) + j\mathbf{s}_{iQ}(t)$, 相位项 $\varphi_i = 2\pi \Delta f_i t + \theta_i$, $\mathbf{n}_b(t) = \mathbf{n}_{bI}(t) + j\mathbf{n}_{bQ}(t)$, 则 $\mathbf{y}(t)$ 可另表示为

$$\mathbf{y}(t) = \sum_{i=1}^2 \alpha_i (\mathbf{s}_{iI}(t) + j\mathbf{s}_{iQ}(t)) (\cos(\varphi_i(t)) + j\sin(\varphi_i(t))) + (\mathbf{n}_{bI}(t) + j\mathbf{n}_{bQ}(t)) \quad (4)$$

展开后,观测信号的 I/Q 分量可表示为

$$\begin{cases} \mathbf{y}_I(t) = \sum_{i=1}^2 \alpha_i [\mathbf{s}_{iI}(t)\cos(\varphi_i(t)) - \mathbf{s}_{iQ}(t)\sin(\varphi_i(t))] \\ \quad + \mathbf{n}_{bI}(t) \\ \mathbf{y}_Q(t) = \sum_{i=1}^2 \alpha_i [\mathbf{s}_{iI}(t)\sin(\varphi_i(t)) + \mathbf{s}_{iQ}(t)\cos(\varphi_i(t))] \\ \quad + \mathbf{n}_{bQ}(t) \end{cases} \quad (5)$$

从式(5)可以看出,观测信号的每一个采样点都是4个未知源信号分量 $\{\mathbf{s}_{1I}, \mathbf{s}_{1Q}, \mathbf{s}_{2I}, \mathbf{s}_{2Q}\}$ 经过时变、非线性的三角函数加权求和的结果。这是一个典型的欠定问题,其复杂的深度耦合和非线性特性使得传统线性方法失效。为了解耦这种复杂关系,分离网络必须能近似一个强大的非线性函数。深度神经网络通过多层非线性激活函数的堆叠,具备拟合任意复杂函数的能力。编码器的作用就是将输入 $\mathbf{y}(t)$ 映射到一个高维隐空间,其数学目标是找到一个特征变换,使得源信号的特征表示更易于分离。为实现这一目标,采用多核深度卷积网络,以学习不同尺度上的信号特征。例如,大卷积核 k_1 可以捕捉由符号速率决定的包络信息(低频特征),而小卷积核 k_2 则能更好地捕捉由载波频偏引起的快速相位变化(高频特征)。

1.2.2 时序依赖的双尺度特性

通信基带信号 $\mathbf{s}_i(t)$ 本身由脉冲成型滤波器 $p(t)$ 对离散符号序列 $\mathbf{c}_{ik}$ 进行卷积生成,其表达式为

$$\mathbf{s}_i(t) = \sum_k \mathbf{c}_{i,k} \cdot p(t - kT_{s,i}) \quad (6)$$

式中: $T_{s,i}$ 为第 i 路信号的符号周期; $\mathbf{c}_{i,k}$ 为第 i 路信号在第 k 个符号周期内的调制符号; k 为符号序号索引。

基带信号生成原理导致了信号时序依赖具有双尺度特性。

局部短期依赖主要体现在单个码元内部的波形平滑性以及码元之间的过渡特性,由成型滤波器决定,这种依赖关系具有局部性和相对平稳的特点。时域卷积网络利用其因果卷积和空洞卷积结构,能够以固定的卷积核高效地学习这种局部时间结构,并通过扩大感受野捕获一定范围内的上下文信息,非常适合对此类短期依赖进行建模。对于一个 L 层的 TCN,第 l 层的输出特征可表示为

$$\mathbf{Z}^{(l)}(t) = f_r(\sum_{m=0}^{M-1} \mathbf{w}_m^{(l)} \cdot \mathbf{Z}^{(l-1)}(t - d_l \cdot m) + \mathbf{b}^{(l)}) \quad (7)$$

式中: $\mathbf{Z}^{(l)}(t)$ 为第 l 层在 t 时刻的输出特征向量; $f_r(\cdot)$ 为线性整流函数,将其定义为: $f_r(x) = \max(0, x)$; m 为卷积核内权重索引; $\mathbf{w}_m^{(l)}$ 为第 l 层卷积核的第 m 个权重矩阵; $\mathbf{b}^{(l)}$ 为第 l 层的偏置向量; M 为卷积核大小; d_l 为第 l 层的膨胀率,通常按指数增长。

全局长期依赖主要源于以下两个方面。首先,载波频率的微小失配 ($\Delta f_i \neq 0$) 会导致相位随时间 t 线性累积旋转 ($2\pi \Delta f_i t$),这是一个贯穿整个信号

序列的全局非平稳过程;其次,当两路信号符号速率不同时,它们的时序结构会产生异步性,这种异步模式也构成长时依赖。CRU 作为一种循环神经网络,其内部的门控机制和隐藏状态能够有效地记忆和传递长距离的时序信息。针对上述长时依赖,GRU 通过隐藏状态更新实现精准建模。

$$\begin{aligned} h_n &= (1 - z_n) \odot h_{n-1} + \\ & z_n \odot \tanh(W_h \cdot [s(n), r_n \odot h_{n-1}] + b_h) \end{aligned} \quad (8)$$

式中: h_n 为第 n 个采样点的 GRU 隐藏状态; $s(n)$ 为 GRU 在时刻 n 的输入向量; z_n 为第 n 个采样点的更新门,用于控制“保留历史长时信息”与“更新当前信息”的比例; r_n 为第 n 个采样点的重置门,用于筛选历史隐藏状态中的有效信息; $\tanh(\cdot)$ 为激活函数; W_h 、 b_h 分别为可学习的权重矩阵与偏置项; $\odot$ 为逐元素乘法。

CRU 弥补了 TCN 在处理全局非平稳依赖时的不足。因此,将 TCN 与双向 GRU 进行融合,可以形成一个优势互补的分离核心,TCN 负责精细刻画局部波形特征,GRU 负责把握全局相位演变和时序异步性,两者共同实现对信号双尺度时序依赖的全面建模。

1.2.3 关键差异特征的动态性与聚焦需求

在不同的混叠场景下,起主导作用的差异特征是动态变化的。例如,在同符号率、异调制场景(如 QPSK+16QAM),星座点的幅相差异是关键;而在同调制、异符号率场景(如 8PSK+8PSK, 5 MBd+2.5 MBd),时序结构的节拍差异则更为重要。但混

叠场景类型复杂多变,仅依靠静态的、一成不变的特征提取网络无法最优地适应所有情况。这就要求模型具备一种注意力机制,能够根据当前输入信号的特性,动态地为不同的特征通道分配权重,即“聚焦”于当前场景下最具区分度的信息,同时抑制噪声和无关特征的干扰。本文提出的自适应多尺度特征注意力(adaptive multi-scale feature attention module, AMFA)模块正是为此设计,通过引入动态系数调制分支,使注意力权重能够根据信号的信噪比、调制类型等内在统计特性进行自适应调整,从而实现对关键差异特征的精准增强。

基于上述理论分析,本文构建了一个由多核深度卷积编码器、AMFA 模块、堆叠式自适应 TCN 与双向 GRU 以及多尺度解码器组成的端到端网络。该架构的每一部分都针对同频混叠信号分离的核心物理挑战而设计,形成了理论驱动与数据驱动相结合的解决方案。

2 模型架构设计

本文设计并实现了一种深度学习框架,以解决单通道同频混叠信号的盲源分离任务。该框架包括 4 个主要组成部分:多核深度卷积编码器、AMFA 模块、堆叠式自适应 TCN 与双向 GRU 以及多尺度解码器。具体架构如图 1 所示,AMFA 模块如图 2 所示,图中 GAP 为全局平均池化(global average pooling),FC 为全连接层(full connection),又称线性层。

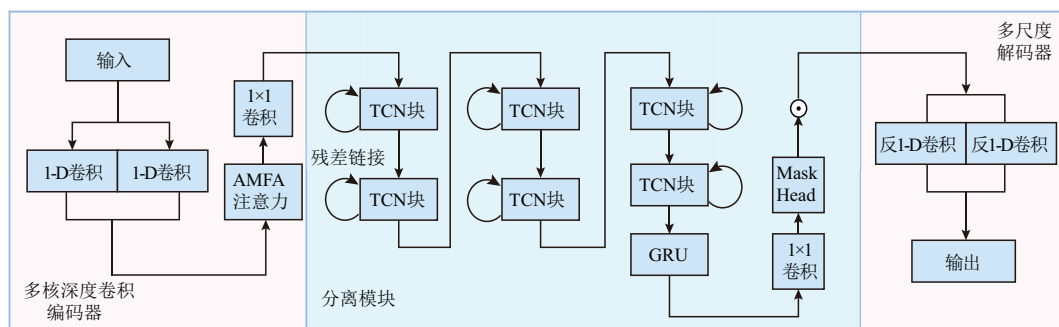


图 1 TCN-GRU 模型架构

Fig. 1 architecture of TCN-GRU model

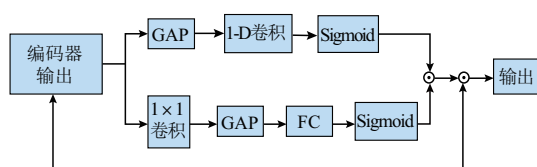


图 2 AMFA 模块

Fig. 2 AMFA module

模型以混合通信信号为输入(维度为 $B \times 2 \times L$, 其中 B 为批量大小, 2 对应 I/Q 双通道, L 为信号长度)。首先,通过编码器将其时域信号映射为高维特征表示;随后利用 AMFA 模块增强关键特征的表达能力;分离模块通过堆叠 TCN 与 GRU 捕捉信号的长时依赖并生成时域掩码;最后多尺度解

码器将掩码加权后的特征映射回时域,得到分离后的源信号(维度为 $B \times 2 \times 2 \times L$)。

2.1 多核深度卷积编码器

多核深度卷积编码器通过多分支深度可分离卷积(由深度卷积和逐点卷积组成)对混合信号的高维特征进行编码,以实现不同尺度信号特征的捕获。对于输入的混合信号 $\mathbf{y} \in \mathbf{R}^{2 \times L}$,首先通过两组并行的“深度卷积+逐点卷积”分支提取不同尺度特征,卷积核尺寸分别设为 $k_1=8$ 和 $k_2=16$,步长为 1,填充设为 $k_i/2$ 以保持序列长度不变。编码后的混合信号特征可表示为

$$\mathbf{f}_i = \mathbf{f}_s(N(\mathbf{C}_{pw}(\mathbf{C}_{dw}(\mathbf{y}, k_i)))) , i=1,2 \quad (9)$$

式中: $\mathbf{C}_{dw}(\cdot, \cdot)$ 为深度卷积; $\mathbf{C}_{pw}(\cdot)$ 为逐点卷积; $N(\cdot)$ 为实例归一化; $\mathbf{f}_s(\cdot)$ 为 SiLU 激活函数; k_i 为第 i 个分支的深度卷积核大小。

逐点卷积通过 1×1 的卷积核将通道数从 2 映射至 256,两支路输出拼接后形成 512 维基础特征。其表达式为

$$\mathbf{f}_{cat} = \mathbf{f}_1 \oplus \mathbf{f}_2 \in \mathbf{R}^{B \times 512 \times L} \quad (10)$$

式中: $\oplus$ 为拼接操作。拼接后的特征再经实例归一化与 SiLU 非线性单元激活处理,可有效避免梯度消失并增强非线性表达能力。编码器的最终输出特征可表示为

$$\mathbf{f}_{enc} = \mathbf{f}_s(N(\text{Conv1D}(\mathbf{f}_{cat}))) \quad (11)$$

式中: Conv1D 为一维卷积层。

2.2 自适应多尺度特征注意力模块

注意力机制本质是对输入的信息进行重要性加权,让模型自动聚焦于当前输入任务中更加关键的信息,同时抑制次要或者无关信息。ECA 机制旨在通过局部跨通道交互实现通道注意力的高效计算,且其参数量少,适合轻量级模型。然而,通信调制信号的特征具有非线性、多尺度的特点,ECA 无法有效地提取这些复杂特征的高阶统计量,导致对不同调制信号的区分能力有限。ECA 的注意力权重完全由静态的一维卷积生成,但是通信场景下混合信号的调制类型组合、信噪比是动态变化的,ECA 无法根据实时信号分布调整注意力策略,导致信号分离的鲁棒性降低。针对以上问题,本研究提出了融合全局 ECA 动态系数调制机制的 AMFA 模块,核心流程包括全局 ECA 计算、动态系数生成及最终特征加权 3 部分,具体过程如下。

首先,对编码特征 $\mathbf{f}_{enc} \in \mathbf{R}^{B \times 512 \times L}$ 进行全局

ECA 计算。通过全局平均池化将特征压缩为 $\mathbf{R}^{B \times 512 \times 1}$,转置后经核尺寸为 5、填充为 2 的一维卷积捕捉通道间依赖,再经 Sigmoid 激活生成全局通道权重 $\mathbf{w}_{eca}$ 。

$$\mathbf{w}_{eca} = \sigma[\text{Conv1D}([\mathbf{P}_g(\mathbf{f}_{enc})]^T)]^T \quad (12)$$

式中: $\mathbf{P}_g(\cdot)$ 为全局平均池化操作; $\sigma(\cdot)$ 为 Sigmoid 函数。

其次,为解决传统通道注意力权重固定、对不同信号特征适配性不足的问题,引入动态系数调制分支对全局通道权重进行修正。该分支首先通过卷积核尺寸为 3 的一维卷积将特征通道数从 C 压缩至 $C/4$,随后经实例归一化与 SiLU 激活后再次执行全局平均池化,得到特征的全局统计量;最后通过全连接层完成维度压缩及动态系数 α 生成(映射至 $[0.5, 1]$ 区间),实现权重随调制类型、信噪比动态调整。

$$\mathbf{f}_{stat} = \mathbf{P}_g(\mathbf{f}_s(N(\text{Conv1D}(\mathbf{f}_{enc})))) \quad (13)$$

$$\alpha = \sigma(\mathbf{f}_c(\cdot)(\mathbf{f}_{stat})) \times 0.5 + 0.5 \in \mathbf{R}^{B \times 1 \times 1} \quad (14)$$

式中: $\mathbf{f}_{stat}$ 为特征的全局统计量; $\mathbf{f}_c(\cdot)$ 为线性变换(全连接层)。

最后,将全局 ECA 权重与动态系数进行逐元素相乘,得到最终自适应特征权重。

$$\mathbf{w}_{final} = \mathbf{w}_{eca} \odot \alpha \quad (15)$$

式中: $\odot$ 为逐元素乘法。将 $\mathbf{w}_{final}$ 扩展至与编码特征 $\mathbf{f}_{enc}$ 相同的维度后,进行逐元素加权操作,实现关键特征的强化与次要信息的抑制。输出加权后的特征为

$$\mathbf{f}_{out} = \mathbf{f}_{enc} \odot \mathbf{w}_{final} \quad (16)$$

该过程结合了全局通道依赖与信号调制类型、信噪比的动态特性,使注意力权重能够自适应地匹配不同信号的非线性、多尺度特征^[20-23]。

2.3 堆叠式自适应 TCN 与双向 GRU

同频混叠信号(载频间隔小于 1 kHz)的时域特征存在双尺度依赖特性,短期依赖呈现局部平稳性,而长期依赖则表现为全局非平稳性。基于此,单一网络结构难以同时优化“局部平稳特征的精细捕捉”与“非平稳长时依赖的全局建模”。因此,通过堆叠 TCN 与双向 GRU 捕捉信号的长时依赖关系,并生成时域掩码实现源信号分离,具体步骤如下。

1) 对输入的信号进行全局归一化,再通过 1×1 卷积将 512 通道的特征降维至 128 通道,减少计算量。通过瓶颈层之后,全局多尺度特征变成了核心

压缩特征,但 128 通道的核心特征仍可能存在局部波动,因此需再经过归一化以解决局部波动问题,让激活函数和后续的模块能更有效地学习信号差异。该过程可表示为

$$\mathbf{f}_b = N_g(\mathbf{C}_{pw}(\mathbf{f}_{in})) \quad (17)$$

式中: $N_g(\cdot)$ 为分组归一化; $\mathbf{f}_{in}$ 为输入特征; $\mathbf{f}_b$ 为瓶颈层输出特征。

2) 堆叠 TCN-GRU 模块提取长时特征,用于捕捉信号的多尺度局部特征和长时依赖关系。TCN 块基于深度可分离卷积设计,通过“分组卷积+逐点卷积”的组合,在减少参数的同时扩大感受野。TCN 块采用固定填充配置:前两组 TCN 的卷积核大小为 3×1 ,填充为 1,2,第三组 TCN 的卷积核大小为 5×1 ,填充为 4,8,所有 TCN 块输出通过残差求和聚合,保留多尺度局部特征^[24]。该过程可表示为

$$\mathbf{f}_{tcn} = \mathbf{f}_1(N_g(\mathbf{C}_{pw}(\mathbf{C}_{dw}(\mathbf{f}_b, \delta)))) + \mathbf{f}_b \quad (18)$$

式中: $\mathbf{f}_{tcn}$ 为 TCN 块输出的时序特征; δ 为填充; $f_1(\cdot)$ 为 LeakyReLU 激活函数。

针对通信信号的长时依赖(如相位连续性、突发干扰持续影响),在第三组 TCN 后引入双向 GRU(隐藏层维度为 384),将 TCN 特征转置为序列维度在前的格式 $(B, L, 512)$ 。经双向 GRU 输出的特征格式为 $(B, L, 2 \times 384)$,再通过 1×1 卷积投影回 512 通道,确保特征维度一致性。

$$\mathbf{f}_{gru} = \text{Conv1D}(G(\mathbf{f}_{tcn})) \quad (19)$$

式中: $\mathbf{f}_{gru}$ 是双向 GRU 模块输出并经 1×1 卷积投影回原通道数后的时序特征; $G(\cdot)$ 为双向 GRU 网络作为一个整体函数。

3) 为隐式刻画调制信号的高阶统计关联(如四阶矩阵反映的是星座对称性、符号间互相关特性),引入高阶增强分支。首先,通过深度可分离卷积提取精细特征,再经逐点卷积调整通道后进行归一化激活,最后与 TCN-GRU 融合特征加权融合。

4) 对特征执行实例归一化与 LeakyReLU 激活^[25-26]。首先通过 1×1 卷积生成与源数量匹配的特征 $(512 \rightarrow 2 \times 512)$,再经 Sigmoid 激活后重塑为掩码 $\mathbf{M} \in \mathbf{R}^{B \times 2 \times 512 \times L}$ (对应 2 路源信号);随后将瓶颈层特征扩展维度后与掩码逐元素相乘,得到各源信号的分离特征 $\mathbf{f}_{sep} \in \mathbf{R}^{B \times 2 \times 512 \times L}$ 。

$$\mathbf{f}_{sep} = \mathbf{f}_{bottleneck} \odot \mathbf{M} \quad (20)$$

式中: $\mathbf{f}_{bottleneck}$ 为扩展维度后的瓶颈层特征。

2.4 多尺度解码器

多尺度解码器采用多“分支转置卷积+特征融

合+相位校正”的级联操作,从分离后的高维编码特征中重构出各源信号。特征融合模块通过对拼接后的多分支输出进行 1×1 卷积压缩,实现不同尺度时域特征(如符号周期、包络趋势)的互补融合;相位校正模块由 3×3 卷积、实例归一化与 tanh 激活组成,一方面将输出幅度限制在合理范围,另一方面补偿分离过程中引入的相位偏移,保证星座图等相位敏感特征的还原精度。

3 网络训练和实验设置

3.1 网络训练

本实验采用的硬件平台配置为:显卡为 NVIDIA RTX 5070,处理器为 Intel(R)Core(TM) i9-14900HX;基于 Python3.10.18 的深度学习框架。

通信信号的盲源分离对分离精度要求较高,因此训练时采用自适应组合损失函数,由尺度不变信噪比(scale invariant signal-to-noise ratio, SI-SNR)和星座损失构成。SI-SNR 损失,可通过计算 2 种源置换组合尺度不变信噪比之和的最小值求得。其数学表达式为

$$\begin{cases} \hat{\mathbf{s}}_{\text{aim}} = \frac{\langle \hat{\mathbf{s}}, \mathbf{s} \rangle}{\|\mathbf{s}\|^2} \mathbf{s} \\ \mathbf{e}_{\text{noise}} = \hat{\mathbf{s}} - \hat{\mathbf{s}}_{\text{aim}} \\ L_{\text{SI-SNR}} = 10 \lg \frac{\|\hat{\mathbf{s}}_{\text{aim}}\|^2}{\|\mathbf{e}_{\text{noise}}\|^2} \end{cases} \quad (21)$$

式中: $\hat{\mathbf{s}} \in \mathbf{R}^{2 \times L}$ 为模型输出的分离信号; $\mathbf{s} \in \mathbf{R}^{2 \times L}$ 为原始源信号; $\hat{\mathbf{s}}_{\text{aim}}$ 为估计信号的目标对齐分量; $\mathbf{e}_{\text{noise}}$ 为估计信号的噪声残差分量; $L_{\text{SI-SNR}}$ 为尺度不变信噪比损失; $\|\cdot\|^2$ 为信号功率。

自适应星座损失,无论信号采用何种调制方式,其星座图均呈现出显著的幅度聚类与相位聚类特征,这是通信调制信号的共性规律。因此,针对此通信信号星座图特性,需动态计算幅度损失、相位损失与聚类损失。首先,将模型输出的分离信号与原始源信号转换为复数基带信号,再进行能量归一化以消除幅度缩放的影响。该过程可表示为

$$\begin{cases} \hat{\mathbf{s}} = \hat{\mathbf{x}}_1 + j \cdot \hat{\mathbf{x}}_2 \\ \mathbf{s} = \mathbf{x}_1 + j \cdot \mathbf{x}_2 \\ \hat{\mathbf{s}}_{\text{norm}} = \frac{\hat{\mathbf{s}}}{\sqrt{\mathbb{E}[\|\hat{\mathbf{s}}\|^2]} + \epsilon} \\ \mathbf{s}_{\text{norm}} = \frac{\mathbf{s}}{\sqrt{\mathbb{E}[\|\mathbf{s}\|^2]} + \epsilon} \end{cases} \quad (22)$$

式中: $\hat{\mathbf{x}}_1$ 、 $\hat{\mathbf{x}}_2$ 分别为分离信号的实部与虚部; $\mathbf{x}_1$ 、 $\mathbf{x}_2$ 分别为原始源信号的实部与虚部; $\mathbb{E}[\cdot]$ 表示序列维度的均值; $\epsilon = 10^{-8}$ 为防止分母为 0 的微小量; $\hat{\mathbf{s}}_{\text{norm}}$ 为分离信号的能量归一化复数形式; $\mathbf{s}_{\text{norm}}$ 为纯净标签信号的能量归一化复数形式。基于归一化后的复数信号计算幅度损失以约束星座点的幅度聚类特性,其表达式为

$$L_{\text{amp}} = \mathbb{E}[\|\hat{\mathbf{s}}_{\text{norm}} - \mathbf{s}_{\text{norm}}\|] \quad (23)$$

式中: L_{amp} 为自适应幅度损失。计算相位损失时,需适配通信信号的圆周相位特性(将相位差约束在 $[-\pi, \pi]$),再约束星座点的相位聚类特性。相位损失可表示为

$$\begin{cases} \phi_s = \arg(\hat{\mathbf{s}}_{\text{norm}}) \\ \phi_s = \arg(\mathbf{s}_{\text{norm}}) \\ \Delta\phi = \text{rem}(\phi_s - \phi_s + \pi, 2\pi) - \pi \\ L_{\text{phase}} = \mathbb{E}[\Delta\phi^2] \end{cases} \quad (24)$$

式中: $\arg(\cdot)$ 为复数的相位角; $\text{rem}(\cdot, 2\pi)$ 表示对 2π 取余运算; ϕ_s 为分离信号能量归一化后的相位角; ϕ_s 为原始纯净信号能量归一化后的相位角; $\Delta\phi$ 为经圆周归一化后的相位偏差; L_{phase} 为自适应相位损失。

最终,自适应星座损失为幅度损失与相位损失的加权融合,可表示为

$$L_c = L_{\text{amp}} + 0.9 \cdot L_{\text{phase}} \quad (25)$$

式中: L_c 为自适应星座总损失。在自适应组合损失中,还会根据原始源信号的幅度方差动态调整 L_c 的权重;对幅度方差较大的调制类型采用更高权重

(0.9),对幅度方差较小的调制类型采用较低权重(0.5)。

优化器选择自适应矩估计器(adaptive moment estimation, Adam),初始学习率 $1e^{-3}$,若验证损失连续 5 次未下降,学习率乘以 0.5;同时引入早停机机制,若验证损失连续 15 次未下降,则终止训练,避免模型过拟合。考虑到神经网络的非凸特性,训练过程中会保存“最新模型”与“验证损失最优模型”,便于后续选择泛化性能最佳的模型。

模型参数通过反向传播更新,涉及编码器、分离模块、解码器 3 个组件的参数集,分别记为 θ_e 、 θ_s 、 θ_d ,更新方式为

$$\theta_e \leftarrow \theta_e - \eta \frac{\partial L_{\text{combined}}}{\partial \theta_e} \quad (26)$$

$$\theta_s \leftarrow \theta_s - \eta \frac{\partial L_{\text{combined}}}{\partial \theta_s} \quad (27)$$

$$\theta_d \leftarrow \theta_d - \eta \frac{\partial L_{\text{combined}}}{\partial \theta_d} \quad (28)$$

式中: η 为当前学习率, $\partial L_{\text{combined}}$ 为自适应组合损失。

3.2 实验设置及评价指标

实验数据选取 QPSK、8PSK、16QAM 共 3 种调制类型的通信基带信号,基于同调制、异调制、相同符号速率、不同符号速率的维度分成 4 个场景进行训练分离。混合信号由软件定义无线电平台 GNU Radio 生成,信噪比范围为 $-10 \sim 25$ dB,步长 2 dB。每个场景下信号具体参数如表 1 所示。

表 1 不同场景信号参数
Table 1 Signal parameters at different scenarios

场景	调制类型	符号速率	初始相偏	采样率/MHz	相对时延/ μs	载频间隔/Hz
场景 1	8PSK+8PSK	5 MBd	$\pi/5$	100	0.06	1 000
场景 2	QPSK+16QAM	5 MBd	$\pi/5$	100	0.06	1 000
场景 3	8PSK+8PSK	5 MBd+2.5 MBd	$\pi/5$	50	4.8×10^{-8}	750
场景 4	QPSK+8PSK	5 MBd+2.5 MBd	$\pi/5$	50	4.8×10^{-8}	750

数据预处理阶段,将复数基带信号拆解为实部与虚部的双通道形式输入网络,同时对 2 通道混合信号进行序列长度归一化;若原始信号长度不足则补零,过长则截断,确保输入维度统一。

为全面衡量分离信号与目标信号的匹配程度,客观评估分离效果,本次实验采用 SI-SNR、皮尔逊相关系数(Pearson correlation coefficient, ρ)、误码

率(bit error rate, BER)3 类指标构建评价体系。

SI-SNR:量化分离信号与目标信号的失真程度,其核心计算逻辑与 3.1 节训练所用 SI-SNR 损失一致。作为评价指标时,其值越高表明分离信号与原始信号的波形一致性越好。

皮尔逊相关系数 ρ :量化分离信号与目标信号之间的线性相关程度,反映两者的整体趋势匹配

性,其公式为

$$\rho(\hat{s}, s) = \frac{\text{cov}(s, \hat{s})}{\sigma_s \sigma_{\hat{s}}} \quad (29)$$

式中: $\text{cov}(s, \hat{s})$ 为原始信号 s 与分离信号 $\hat{s}$ 的协方差; σ_s 和 $\sigma_{\hat{s}}$ 分别为两者的标准差。 ρ 的取值范围为 $[-1, 1]$, 越接近 1, 表明分离过程中信号的核心特征保留越完整。

误码率: 量化数字通信场景下分离信号的解调质量, 反映码元判决的错误比例, 其计算公式可表示为

$$L_{\text{BER}} = \frac{N_{\text{error}}}{N_{\text{total}}} \quad (30)$$

式中: N_{error} 为解调后判定错误的比特数, N_{total} 为信号的总比特数。 L_{BER} 值越低, 表明分离信号的幅相调制特征还原越精准, 解调性能越好。

4 实验与结果分析

4.1 分离结果及可视化

图 3、4 分别为模型在不同混合场景下, SI-SNR 和 ρ 随输入信噪比变化的特性曲线。如图 3 所示, 低 SNR 区间 ($-10 \sim 5$ dB) 内, 各场景 SI-SNR 的平均增幅约为 25.2 dB, 增长速率最快; 中 SNR 区间 ($5 \sim 15$ dB) 内, 增长明显放缓, 平均增幅约 7.6 dB; 高 SNR 区间 ($15 \sim 25$ dB) 内, SI-SNR 趋于饱和, 平均增幅仅为 3.1 dB。同时, 在 SNR 到达 5 dB 后各场景的性能排序保持稳定分层: 场景 4 > 场景 1 > 场景 2 > 场景 3, 这一排序直观反映了模型对不同“调制-符号率组合”场景的适配差异。

在低 SNR 区间, 各场景的 SI-SNR 曲线近乎重叠, 场景间差距不足 1 dB。这是因为低 SNR 下高斯噪声的强干扰主导了信号特征, 源信号的调制类型差异、符号率差异均被噪声掩盖, 模型的核心作用聚焦于“噪声压制与信号初步提取”, 场景固有特性暂未凸显, 故性能表现趋同。在中高 SNR 区间, 场景 4 的异调制组合 (QPSK + 8PSK) 具有清晰的相位间隔差异, 可被 AMFA 模块高效聚焦; 同时其异符号率 (5 MBd + 2.5 MBd) 带来的时序差异, 恰好通过 GRU 的长时依赖建模实现了关联捕捉, 模块协同作用使其性能领先; 场景 1 虽无调制差异, 但统一的时序节奏让 TCN 的局部卷积建模充分发挥作用, 故性能次之; 场景 2 的异调制 (QPSK + 16QAM) 虽能被 AMFA 捕捉, 但 16QAM 的高密度星座点增加了特征区分难度, 导致其性能弱于

场景 1; 场景 3 (8PSK + 8PSK) 既无调制差异供 AMFA 聚焦, 又因异符号率 (5 MBd + 2.5 MBd) 加剧了时序建模复杂度, 故始终效果略低于其他几个场景。

如图 4 所示, 平均皮尔逊相关系数 ρ 的变化趋势与 SI-SNR 高度契合, 在信噪比提升至 10 dB 后, 各场景的 ρ 值均趋近于 1, 且后续随信噪比继续增大而保持稳定, 表明分离信号与原始纯净信号的线性相关性已达到极高水平, 信号的核心时序特征与整体趋势被完整保留。

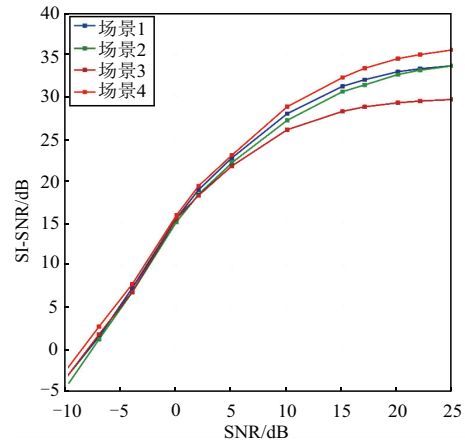


图 3 分离后 4 个场景信号的平均尺度不变信噪比

Fig. 3 Average SI-SNR of the signals from four scenarios after separation

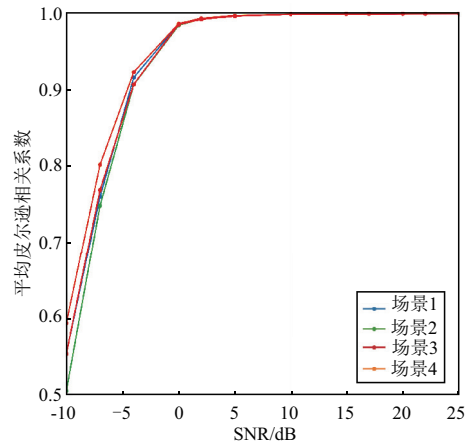


图 4 分离后 4 个场景信号的平均皮尔逊相关系数

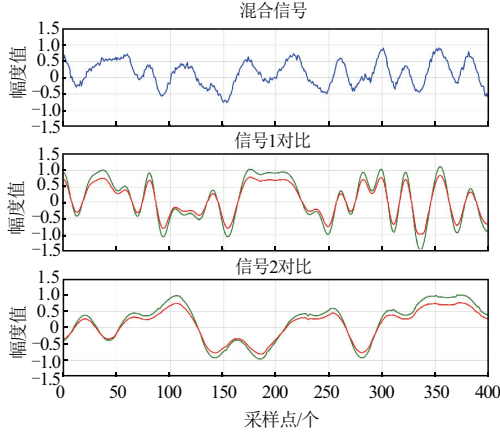
Fig. 4 Average Pearson correlation coefficient of the signals from four scenarios after separation

图 5 分别为 4 个场景在信噪比为 20 dB 时的信号分离结果。从图中可以看出, 分离波形和原始波形呈高度匹配, 这也证明了本文模型的架构设计可以有效分离 4 种不同场景下的混合信号。

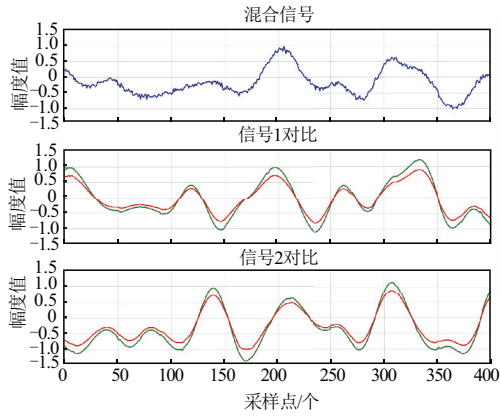
4.2 模型对同频混叠程度的鲁棒性分析

同频混叠指两路或多路信号的中心频率极

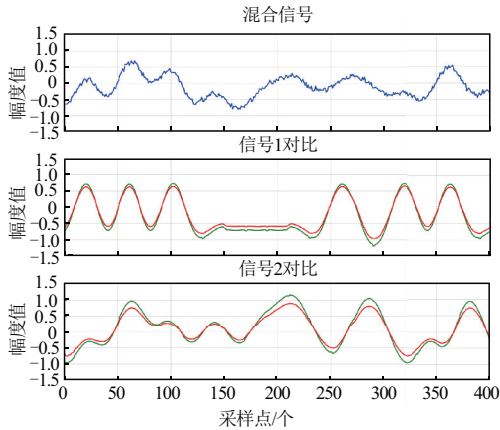
为接近,导致其频谱在接收端发生严重重叠,无法通过传统线性滤波器进行有效分离的场景。在实际的频谱监测与非合作通信中,这种“小载频间隔”(Δf)并没有一个绝对的标准,但通常当 Δf 小于信号自身的带宽时,即可视为混叠。当 Δf 进一步缩小至 1 kHz 甚至数百 Hz 量级时,信号在时域上因载频差异而产生的相位旋转变得非常缓慢,这对依赖时序特征的分离算法构成了严峻挑战。



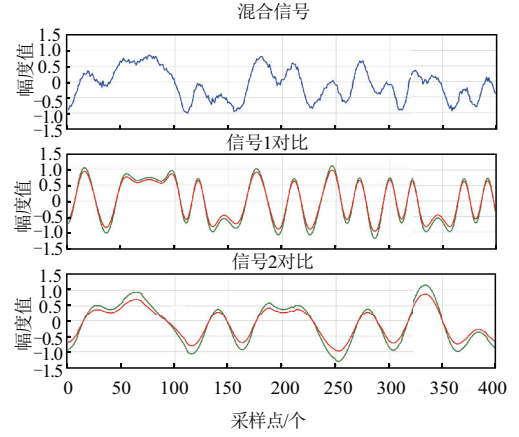
(a) 场景 1 信号的分离效果波形可视化



(b) 场景 2 信号的分离效果波形可视化



(c) 场景 3 信号的分离效果波形可视化



(d) 场景 4 信号的分离效果波形可视化

— 纯净信号 — 分离信号

图 5 不同场景信号分离结果的波形可视化

Fig. 5 Waveform visualization of signal separation results for different scenarios

因此设计变载频间隔的仿真实验来验证模型对同频混叠程度的鲁棒性。实验条件设置固定输入信噪比为 15 dB,这是一个能代表中等通信质量的典型值。针对表 1 中定义的 4 种场景,将载频间隔 Δf 作为唯一变量,在 200 Hz 至 10 kHz 范围内进行扫描测试。该范围覆盖了从极端混叠($\Delta f < 250$ Hz),到核心同频混叠($250 \text{ Hz} \leq \Delta f < 1 \text{ kHz}$),再到较大频偏($\Delta f \geq 1 \text{ kHz}$)的完整区间,旨在全面考察模型的性能变化曲线与工作边界。实验采用尺度不变信噪比和误码率作为核心评价指标,结果分别如图 6 和图 7 所示。

综合分析图 6 和图 7,可以观察到所有场景随着载频间隔 Δf 的增大,SI-SNR 显著提升,BER 则相应降低。其原因在于, Δf 越大,混合信号在时域上相位旋转差异更明显,可为模型的 TCN-GRU 混合结构捕捉和解耦信号的长时依赖特征提供更丰富、更具区分度的信息。具体性能表现可分为以下 3 个区间进行剖析。

1) 在较大频偏区($\Delta f \geq 1 \text{ kHz}$),模型性能表现出色且趋于稳定。此时各场景的 SI-SNR 均达到 30 dB 以上的较高水平,而 BER 也基本达到其性能上限(对于大部分场景,BER 优于 0.03)。这表明当频偏大于 1 kHz 时,信号间的差异已足够明显,模型能够稳定、高效地实现混合信号的高质量分离,且可以进一步分析得出,影响该区域信号分离效果的不再是载频间隔这一因素,而是模型自身拟合能力等其他因素。

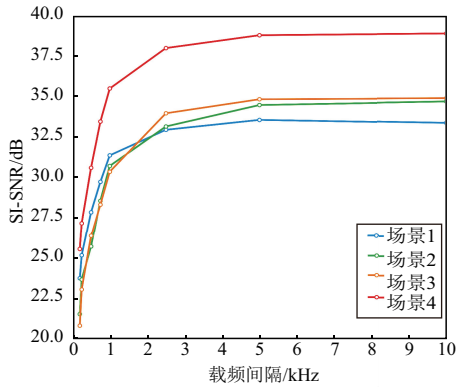


图 6 不同载频间隔下的 SI-SNR 性能

Fig. 6 SI-SNR performance under different carrier frequency intervals

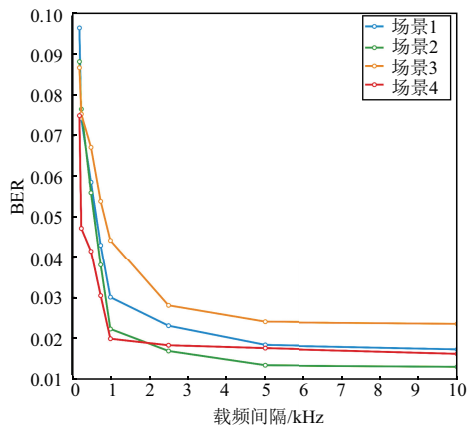


图 7 不同载频间隔下的 BER 性能

Fig. 7 BER performance under different carrier frequency intervals

2) 核心同频混叠区 ($250 \text{ Hz} \leq \Delta f < 1 \text{ kHz}$)。这是最具挑战性也最能体现模型鲁棒性的区域。当 Δf 从 1 kHz 逐步下降至 500 Hz 时,各场景 SI-SNR 出现 3~4 dB 的小幅下降,但仍维持在 25 dB 以上,整体保持良好水平;相应地,BER 呈不同程度上升,但除场景 3 外均保持在 0.06 以下。当 Δf 进一步压缩至 250 Hz 时,性能开始出现更明显的下降,但即便是性能相对最差的场景 3,其 SI-SNR 仍维持 23.05 dB,BER 为 0.075 4,处于基本可用范围。这证明模型在典型的同频混叠场景下并未失效,而是表现出一种渐进式、可预测的性能下降,具有较强的鲁棒性。

3) 极端混叠区 ($\Delta f < 250 \text{ Hz}$)。当载频间隔压缩至 200 Hz 时,所有场景的性能均急剧下降。此时,4 类场景的 SI-SNR 分别降至 23.73 dB(场景 1)、21.56 dB(场景 2)、20.86 dB(场景 3)和 25.50 dB(场景 4),与载频间隔为 1 kHz 时的对应值 (31.25、30.61、30.26、35.30 dB)相比,平均降幅达 9

dB 左右;与此同时,BER 分别升至 0.096 6、0.088 3、0.086 8 和 0.075 0,较 1 kHz 时的 0.030 0、0.022 2、0.044 2、0.019 8 显著恶化,部分场景的 BER 已超出可接受通信范围。两路信号的载波频率极为接近,时域特征的区分度被大幅削弱,已接近本模型特征提取能力的极限。

综合 SI-SNR 和 BER 两项指标的分析,可以明确本文所提 TCN-GRU 模型对同频混叠问题具有优良的鲁棒性。其有效工作边界可以界定在 250 Hz。当载频间隔大于 500 Hz 时,模型能提供稳定且高质量的分离效果;当间隔在 250 Hz 到 500 Hz 之间时,性能会随间隔减小而平滑下降,但仍处于完全可用的水平;而当间隔小于 250 Hz 时,则进入了模型的性能极限区。

4.3 对比实验

表 2 不同算法的参数

Table 2 Parameters of different algorithms

算法	参数	值
TDE-ICA	迭代次数	800
	迭代次数	500
NMF	基信号数	16
	编码器核数	8、16
TCN-GRU	瓶颈层通道数	128
	TCN 扩张率	1、2、4、8
GRU 隐层大小	GRU 隐层大小	384
	输出源数量	2
DPRNN	隐藏层维度	512
	双路径块数量	8
Conv-TasNet	注意力头数	8
	滤波器数量	64
滤波器长度	滤波器长度	16
	瓶颈层通道数	128
输出源数量	TCN 隐层通道数	256
	输出源数量	2

为验证本文所提 TCN-GRU 模型的有效性,选取了 TDE-ICA^[3-4]、非负矩阵分解 (non-negative matrix factorization, NMF)^[14] 两种经典算法以及 Conv-TasNet^[6]、DPRNN^[7] 两种深度学习模型进行对比。其中,TDE-ICA 通过时间延迟嵌入这一技巧构建虚拟多通道,将单通道问题转化为多通道问题,以满足 ICA 的输入要求;NMF 则假设信号在某

个变换域中可由一组非负基信号线性表示。以上两种是盲源分离领域的经典方法,将其作为基线对比,可以衡量深度学习方法带来的性能飞跃。Conv-TasNet 开创性地在时域使用 TCN 进行端到端分离,有效建模了长时依赖;DPRNN 则通过创新的块内与块间循环处理,进一步提升了对超长序列的建模效率和性能。选择上述模型作为对比,是因为二者是时域分离任务中最具代表性的 SOTA 架构,能够客观评估本文针对通信信号同频混叠特性所做改进的实际增益。各算法参数按合理配置统一设定,如表 2 所示。以 SI-SNR 与皮尔逊相关系数为评价指标,在 SNR 为 25 dB 通信场景下进行性能测试,结果如表 3 所示。

从表 3 的分离对比结果可见:传统算法 TDE-ICA、NMF 的分离性能显著弱于深度学习类算法,其 SI-SNR 均低于 10 dB、皮尔逊相关系数均值低于 0.8;DPRNN 与 Conv-TasNet 性能有所提升,SI-SNR 均值分别为 22.95 dB、29.36 dB,皮尔逊相关系数均值分别为 0.93、0.96;而所提 TCN-GRU 在所有场景中均表现最优,SI-SNR 均值为 32.37 dB、皮尔逊相关系数均值为 0.97,充分验证了其在信号分离任务中的性能优势。Conv-TasNet 和 DPRNN 源于语音分离领域,其网络结构是为语音信号的短时平稳特性优化的。然而,通信信号的分离,尤其是在同频混叠下,其核心挑战在于捕捉由载频差和符号速率差引起的长时、非平稳的相位和时序依赖。本文的 TCN-GRU 融合结构正是针对这一通信信号特有的物理机理而设计,因此本文模型的分离效果比 Conv-TasNet 和 DPRNN 在 SI-SNR 指标上分别提升 3.01 dB、9.42 dB。

表 3 不同算法在 25 dB 混合信号下的分离结果

Table 3 Separation results of different algorithms under 25 dB mixed signals

算法	指标	场景 1	场景 2	场景 3	场景 4
TDE-ICA	SI-SNR/dB	8.52	8.25	7.83	8.06
	ρ	0.763	0.752	0.768	0.746
NMF	SI-SNR/dB	9.81	9.56	9.03	9.36
	ρ	0.794	0.816	0.776	0.783
DPRNN	SI-SNR/dB	23.55	24.02	17.65	26.58
	ρ	0.923	0.934	0.915	0.943
Conv-TasNet	SI-SNR/dB	28.52	30.06	27.25	31.59
	ρ	0.953	0.963	0.961	0.953
TCN-GRU	SI-SNR/dB	32.99	32.65	29.28	34.55
	ρ	0.968	0.971	0.962	0.979

表 4 分离混合信号所用时间

Table 4 Time consumed for separating mixed signals

算法	单样本用时/ms
TDE-ICA	0.6
NMF	0.9
DPRNN	1.5
Conv-TasNet	0.9
TCN-GRU	1.3

由表 4 可知本文模型的单样本分离用时较长。TDE-ICA、NMF 作为模型驱动的传统算法,本质是浅层结构,其建模逻辑依赖“信号独立/非负”等先验假设,难以刻画单通道盲源分离中“同频混叠、异构调制”的复杂特征。因此,即便其单样本用时仅 0.6~0.9 ms,也因精度无法满足通信需求而难以落地。

4.4 消融实验

4.4.1 GRU 消融实验

为验证 GRU 模块对同频混叠基带信号长时序依赖建模的贡献,开展了消融模型对比实验,实验设置如下:移除本文模型中的 GRU 模块,保持其他结构与训练超参数完全一致,对比“带 GRU”(基础模型)与“无 GRU”(消融模型)在 4 类混叠场景下的性能差异,结果如图 8、图 9 所示。

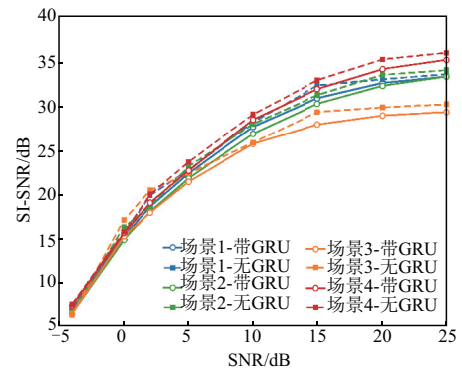


图 8 GRU 消融实验的 SI-SNR 对比

Fig. 8 SI-SNR comparison in the GRU ablation experiment

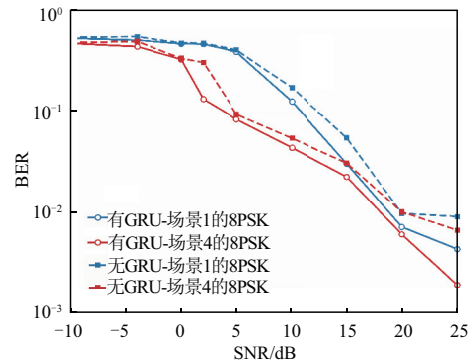


图 9 GRU 消融实验的 BER 对比

Fig. 9 BER comparison in the GRU ablation experiment

图 8 中的 SI-SNR 曲线显示,所有场景下无 GRU 模型的 SI-SNR 均高于对应带 GRU 模型;图 9 中的 BER 曲线却显示,带 GRU 模型的 BER 低于无 GRU 模型。这一现象揭示了波形拟合与信息解调之间的不一致性:SI-SNR 主要通过最小化估计信号与真实信号的时域波形误差实现优化,其以能量为核心的计算逻辑使得模型倾向于拟合时域中高能量的波峰波谷区域;而 BER 取决于星座点判决,对信号的相位连续性和微小幅相失真更为敏感。无 GRU 模型虽然通过 TCN 的大感受野较好地还原了波形包络(高 SI-SNR),但缺乏 GRU 对长时序相位演变的记忆能力,导致解调特征(星座点)发散,从而在通信评价指标(BER)上表现不佳。这证明了引入 GRU 挖掘长时依赖的必要性。

图 9 仅选取场景 1 与场景 4 进行 BER 对比分析:场景 1 为同调制同符号率,GRU 增益最小(下限);场景 4 为异调制异符号率,GRU 增益最大(上限)。场景 2、3 的 GRU 增益介于二者之间,因此通过上下限对比即可有效展示 GRU 模块的作用边界。中高 SNR 区间带 GRU 模型的 BER 下降速度快于无 GRU 模型,且场景 4 的分离效果明显优于场景 1。这一结果本质是 GRU 模块核心作用的体现:GRU 能有效捕捉信号的长时序依赖,精准建模载频偏移累积的相位旋转与符号率差异引发的时序异步特征,让与解调相关的特征更具判别性。而场景 4 为异调制+异符号率组合,时序依赖特征更复杂,GRU 对长时特征的捕捉与判别性增强作用表现得更为突出,相比之下场景 1 同调制同符号率的时序特征更简单,GRU 的增益表现相对平缓,因此呈现出场景间的性能差异。

本研究的 GRU 消融实验结果显示,移除 GRU 后,尽管对 SI-SNR 的影响甚微,但 BER 却显著恶化。这一看似矛盾的现象,深刻揭示了单通道盲源分离任务的双重目标:既要完成信号处理层面的波形重构,又要实现通信层面的信息解调。当前已有研究^[27-30]针对多种信号场景及评估维度进行了探索,在此基础上本文针对同频混叠通信信号所提出的网络架构,进一步验证了在通信解调任务中,兼顾波形重构与信息恢复的联合优化策略的必要性,且本文的所设计的多尺度相位校正解码器设计正是该研究思路的具体实验。

4.4.2 AMFA 消融实验

为验证所提出的 AMFA 模块对模型分离性能的增益,本节构建了 AMFA 消融实验,实验设置如

下。保持模型其余结构与超参数完全一致,对比“集成 AMFA(基准模型)”与“移除 AMFA(消融模型)”在 4 类典型通信场景下的性能差异。

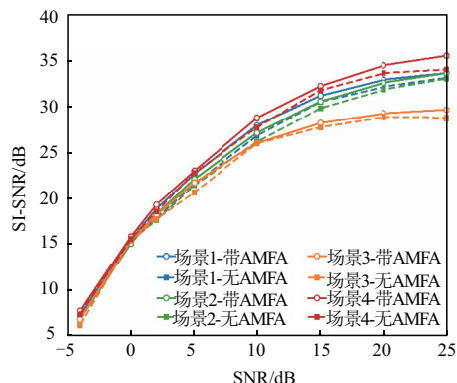


图 10 AMFA 消融实验的 SI-SNR 对比

Fig. 10 SI-SNR comparison in the AMFA ablation experiment

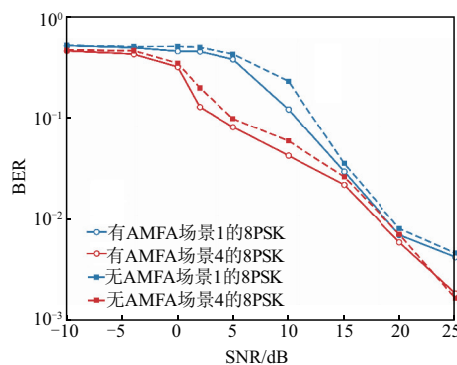


图 11 AMFA 消融实验的 BER 对比

Fig. 11 BER comparison in the AMFA ablation experiment

图 10 展示了 4 类场景下有无 AMFA 模块的模型在不同 SNR 下的 SI-SNR 曲线。所有场景中,集成 AMFA 的基准模型 SI-SNR 均略高于对应消融模型,在中高 SNR 区间(10~20 dB)可为模型带来 1~2 dB 的 SI-SNR 增益,验证了 AMFA 模块对特征质量的提升作用。图 11 BER 曲线表明,两类场景中基准模型的 BER 整体低于消融模型,且中高 SNR 区间(10~20 dB)内性能差距更为显著。这一现象主要源于噪声强度与特征提取效率的耦合关系。低 SNR 下,噪声的能量占比掩盖了信号的细微特征,AMFA 的自适应尺度调整作用被噪声压制,因此基准模型与消融模型的 BER 差距不大;中 SNR 下,噪声能量占比下降,信号的多尺度特征可被有效提取,AMFA 通过动态系数调制精准匹配不同尺度的信号特征,因此基准模型和消融模型的 BER 差异可以明显区分;高 SNR 下,信号本身的干扰与特征模糊度已极低,AMFA 模块的增益空间被压缩,即便是移除 AMFA 的消融模型,也可依托基

础网络的建模能力,较好地还原信号的调制特征。

5 结论

为解决极端同频混叠场景下单通道盲源分离问题,提出了基于自适应特征增强的 TCN-GRU 单通道盲源分离网络,有效工作边界延伸至 250 Hz,与 Conv-TasNet、DPRNN 等模型相比,在高信噪比场景下 SI-SNR 分别提升 3 dB、9.4 dB;针对性弥补了现有研究存在的缺口,并得到以下结论。

1) 多核深度卷积编码器采用多尺度场景感知模块,通过多核深度可分离卷积提取信号多尺度特征,并融入场景特征分支,有效解决了不同通信场景下信号特征的适配性问题。实验表明,在-10~25 dB 全 SNR 区间内,4 类场景的分离信号与原始信号波形高度吻合,10 dB 以上皮尔逊相关系数均接近 1,验证了编码器在多调制、多符号速率场景下的可靠适配能力。

2) 自适应注意力模块通过通道注意力的动态调控,可精准聚焦微小载波频偏、相近符号速率等细微差异特征,显著提升了关键差异信息的区分度。消融实验表明,AMFA 在中高 SNR 区间(10~20 dB)可为模型带来 1~2 dB 的 SI-SNR 增益,并有效降低误码率,验证了其调制差异特征的有效聚焦能力。

3) 分离模块核心为堆叠式自适应 TCN 与双向 GRU 的融合结构。TCN 通过自适应调整扩张率,可灵活捕捉不同尺度的时序关联,适配符号速率差导致的非平稳长时异步场景;双向 GRU 强化了对信号长时依赖的建模能力,弥补了纯 TCN 在长时关联建模中的稳定性不足。消融实验表明,移除 GRU 后虽对 SI-SNR 影响轻微,但 BER 在中高 SNR 区间显著恶化,证明 GRU 对保障解调性能具有不可替代的作用。在 25 dB 下,4 场景 SI-SNR 分别高达 32.99、32.65、29.28、34.55 dB,BER 均处于极低水平,充分支撑了通信解调的可靠性。

4) 多尺度解码器采用多尺度转置卷积与相位校正模块,专门解决信号重构过程中的相位失真问题,保障了分离信号的相位一致性与完整性。

该模型可有效处理微小载频间隔、符号速率异构的同频混叠信号,为非合作通信、频谱监测等领域提供技术方案。但模型在极端低 SNR 下性能仍有提升空间,且 GRU 模块的循环结构对 GPU 并行效率有一定限制;未来可探索基于结构化状态空间

的长时建模方案,并扩展多源分离场景适配能力。

参考文献

- [1] Bofill P, Zibulevsky M. Underdetermined blind source separation using sparse representations[J]. Signal Processing, 2001, 81(11): 2353-2362.
- [2] Davies M E, James C J. Source separation using single channel ICA [J]. Signal Processing, 2007, 87(8): 1819-1832.
- [3] 邓文,黄知涛,王翔.单通道通信信号盲分离方法的研究进展综述[J].通信学报,2023,44(8):179-194.
Deng Wen, Huang Zhitao, Wang Xiang. Overview of research progress on blind separation methods for single channel communication signal[J]. Journal of Communications, 2023, 44(8): 179-194.
- [4] 刘荣,余志勇,韩佳,等.欠定盲源分离技术的研究进展综述[J].电子技术与软件工程,2022(22):84-89.
- [5] Luo Yi, Mesgarani N. TasNet: time-domain audio separation network for real-time single-channel speech separation[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ, USA; IEEE, 2018:696-700.
- [6] Luo Yi, Mesgarani N. Conv-TasNet: surpassing ideal time-frequency magnitude masking for speech separation[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2019, 27(8): 1256-1266.
- [7] Luo Yi, Chen Zhuo, Yoshioka T. Dual-path RNN: efficient long sequence modeling for time-domain single-channel speech separation[C]//ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ, USA; IEEE, 2020: 46-50.
- [8] Hou Xiaoqi, Gao Yong. Skip connection-enhanced hybrid neural network for single-channel blind separation[J]. Signal, Image and Video Processing, 2025, 19(8): 660.
- [9] Dai Zean. NES2Net: a number estimation and signal separation network for single-channel blind signal separation [C]//2024 7th International Conference on Advanced Algorithms and Control Engineering (ICAACE). Piscataway, NJ, USA; IEEE, 2024: 469-474.
- [10] Ma Hao, Zheng Xiang, Yu Lu, et al. A novel end-to-end deep separation network based on attention mechanism for single channel blind separation in wireless communication [J]. IET Signal Processing, 2023, 17(2): e12173.
- [11] Guo Pengcheng, Yu Miao, Shen Lei, et al. Single-channel blind source separation in wireless communications: a complex-domain deep learning approach [J]. IEEE Wireless Communications Letters, 2024,

- 13(6): 1645-1649.
- [12] Fu Weihong, Zhao Wensheng, Zhang Xinyu. A single-channel blind separation convolutional network combined with attention mechanism for communication signals[J]. *Circuits, Systems, and Signal Processing*, 2024, 43(2): 1240-1269.
- [13] 赵孟晨. 基于深度学习的空间混叠信号单通道盲源分离技术研究 [D]. 北京: 中国科学院大学(中国科学院国家空间科学中心), 2021: 47-82.
- Zhao Mengchen. Single channel blind source separation of spatial aliasing signal based on deep learning[D]. Beijing: University of Chinese Academy of Sciences, 2021: 47-82.
- [14] Févotte C, Idier J. Algorithms for nonnegative matrix factorization with the β -divergence[J]. *Neural Computation*, 2011, 23(9): 2421-2456.
- [15] 陈柄旭, 高勇. 基于多编码器及双路径 Transformer 网络的单通道盲分离 [J]. *通信技术*, 2025, 58(9): 950-954.
- Chen Bingxu, Gao Yong. Single-channel blind separation based on multi-encoder and dual-path Transformer network[J]. *Communications Technology*, 2025, 58(9): 950-954.
- [16] Nie Wei. A single-channel blind source separation method based on modulation parameter estimation and signal reconstruction[C]//2025 IEEE 8th Information Technology and Mechatronics Engineering Conference (ITOEC). Piscataway, NJ, USA: IEEE, 2025: 1118-1123.
- [17] 颜子彦, 刘宇腾, 周建二, 等. 单通道盲源分离的多尺度模态融合复数卷积神经网络算法 [J]. *西安邮电大学学报*, 2026, 31(1): 1-8.
- Yan Ziyang, Liu Yuteng, Zhou Jianer, et al. A multi-scale modal-fusion complex-valued neural convolutional network algorithm for single-channel blind source separation [J]. *Journal of Xi'an University of Posts and Telecommunications*, 2026, 31(1): 1-8.
- [18] 于宏毅, 查仁鹏, 沈智翔, 等. 单通道同频混合信号的最佳接收及盲解调性能界 [J]. *通信学报*, 2024, 45(12): 142-152.
- Yu Hongyi, Zha Renpeng, Shen Zhixiang, et al. Optimal receiver and blind demodulation performance bounds for single channel co-frequency mixed signals[J]. *Journal of Communications*, 2024, 45(12): 142-152.
- [19] Tu Shilong, Zheng Hui, Gu Na. Single-channel blind separation of two QPSK signals using per-survivor processing[C]//APCCAS 2008-2008 IEEE Asia Pacific Conference on Circuits and Systems. Piscataway, NJ, USA: IEEE, 2008: 473-476.
- [20] 付卫红, 周亚飞, 张晓燕, 等. 基于参数估计和 Kalman 滤波的单通道盲源分离算法 [J]. *系统工程与电子技术*, 2024, 46(8): 2850-2856.
- Fu Weihong, Zhou Yafei, Zhang Xiaoyan, et al. Single-channel blind source separation algorithm based on parameter estimation and Kalman filter[J]. *Systems Engineering and Electronics*, 2024, 46(8): 2850-2856.
- [21] Zhang Liyuan, Liu Yusi, He Fei, et al. ASCNet: attention mechanism and self-calibration convolution fusion network for X-ray femoral fracture classification [J]. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 2023, 27(6): 1192-1199.
- [22] 付卫红, 张鑫钰, 刘乃安. 基于多尺度融合神经网络的同频同调制单通道盲源分离算法 [J]. *系统工程与电子技术*, 2025, 47(2): 641-649.
- Fu Weihong, Zhang Xinyu, Liu Naian. Single-channel blind source separation algorithm for co-frequency and co-modulation based on multi-scale fusion neural network[J]. *Systems Engineering and Electronics*, 2025, 47(2): 641-649.
- [23] Liu Xiaobei, Guan Yongliang, Koh S N, et al. Low-complexity single-channel blind separation of co-frequency coded signals[J]. *IEEE Communications Letters*, 2018, 22(5): 990-993.
- [24] 施昊, 郭文普, 陈昊. 基于深度学习的生成式文本隐写研究综述 [J]. *火箭军工程大学学报*, 2025, 39(2): 124-136.
- Shi Hao, Guo Wenpu, Chen Hao. Review on deep learning-based generative text steganography[J]. *Journal of Rocket Force University of Engineering*, 2025, 39(2): 124-136.
- [25] Zha Renpeng, Yu Hongyi, Shen Zhixiang. A fast demodulation algorithm for co-frequency mixed signals based on pre-decision[C]//2024 IEEE 24th International Conference on Communication Technology (ICCT). Piscataway, NJ, USA: IEEE, 2024: 1869-1873.
- [26] 陈昊, 郭文普, 康凯. 基于相位变换和 GhostNet-门控循环单元的自动调制识别方法 [J]. *火箭军工程大学学报*, 2024, 38(4): 86-92.
- Chen Hao, Guo Wenpu, Kang Kai. Automatic modulation recognition method based on phase transformation and GhostNet-GRU[J]. *Journal of Rocket Force University of Engineering*, 2024, 38(4): 86-92.
- [27] Ma Baoze, Zhang Tianqi. Single-channel blind source separation for vibration signals based on TVF-EMD and improved SCA[J]. *IET Signal Processing*, 2020, 14(4): 259-268.
- [28] Deng Wen, Wang Xiang, Huang Zhitao. Co-channel

- multiuser modulation classification using data-driven blind signal separation[J]. IEEE Internet of Things Journal, 2024, 11(8): 14829-14843.
- [29] 付卫红,周雨菲,赵文胜. 面向能量差异混合信号的单通道盲源分离算法[J]. 系统工程与电子技术, 2023, 45(10):3302-3311.
- Fu Weihong, Zhou Yufei, Zhao Wensheng. Single-channel blind source separation algorithm for energy differences mixed signals[J]. Systems Engineering and Electronics, 2023, 45(10):3302-3311.
- [30] 解元,邹涛,孙为军,等. 面向卷积混叠环境下的盲源分离新方法[J]. 自动化学报, 2023, 49(5):1062-1072.
- Xie Yuan, Zou Tao, Sun Weijun, et al. Novel blind source separation method for convolutive mixed environment[J]. Acta Automatica Sinica, 2023, 49(5): 1062-1072.

Temporal Convolutional Network-Gated Recurrent Unit Single Channel Based on Adaptive Feature Enhancement Blind Source Separation Network

PENG Ruiyan, GUO Wenpu*, GAO Shaoyuan, MU Chenchen

(Rocket Force University of Engineering, Xi'an 710025, China)

ABSTRACT: To address the issues of insufficient long-term dependency modeling and weak feature-capture capability in single-channel blind source separation for communication signals under extreme co-frequency aliasing and high-order modulation scenarios, a single-channel blind separation network based on an adaptive feature-enhanced temporal convolutional network-gated recurrent unit (TCN-GRU) was proposed. The architecture comprised four core modules: a multi-kernel depthwise convolutional encoder, an adaptive multi-scale feature attention module, a stacked adaptive TCN combined with a bidirectional GRU, and a multi-scale decoder. First, the multi-kernel depthwise convolutional encoder ensured reliable scenario adaptability. Second, the adaptive attention module dynamically adjusted channel weights to focus on critical differential features. Subsequently, the TCN-GRU fusion method captured the multi-scale dependencies of the signals. Finally, the multi-scale decoder guaranteed the integrity of signal reconstruction. Simulation experiments conducted on four typical aliasing scenarios involving homogeneous/heterogeneous modulation and symbol rates demonstrate that the proposed method achieves a stable performance across a signal-to-noise ratio (SNR) range of -10 dB to 25 dB, with the separated signal waveforms closely matching the original signals and Pearson correlation coefficients approaching one above 10 dB. Robustness tests with respect to the degree of co-frequency aliasing reveal that when the carrier frequency interval is as low as 250 Hz, the scale-invariant signal-to-noise ratios (SI-SNRs) for all the scenarios remain above 22 dB, and the bit error rate (BER) stays within a usable range, with the effective operating boundary extending to 250 Hz. Compared with advanced models such as Conv-TasNet and DPRNN, the proposed method achieves average SI-SNR improvements of 3.01 dB and 9.42 dB under a high SNR environment, respectively, reaching 34.55 dB in the optimal scenario. Furthermore, ablation experiments indicate that the GRU module significantly reduces the BER and the adaptive multi-scale feature attention module brings SI-SNR gain of approximately $1-2$ dB in the mid-to-high SNR range, verifying the substantial contribution of each module to the communication demodulation performance.

KEYWORDS: single-channel blind source separation; gated recurrent unit; co-frequency aliasing; temporal convolution

(编辑:吴文冬)