



系统工程与电子技术  
Systems Engineering and Electronics  
ISSN 1001-506X, CN 11-2422/TN

## 《系统工程与电子技术》网络首发论文

题目: 基于 Stacked-TCN 的空间混叠信号单通道盲源分离方法  
作者: 赵孟晨, 姚秀娟, 王静, 董苏惠  
收稿日期: 2020-10-12  
网络首发日期: 2021-03-02  
引用格式: 赵孟晨, 姚秀娟, 王静, 董苏惠. 基于 Stacked-TCN 的空间混叠信号单通道盲源分离方法[J/OL]. 系统工程与电子技术.  
<https://kns.cnki.net/kcms/detail/11.2422.tn.20210227.2208.006.html>



**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。

# 基于 Stacked-TCN 的空间混叠信号单通道盲源分离方法

赵孟晨<sup>1, 2</sup>, 姚秀娟<sup>2</sup>, 王 静<sup>2</sup>, 董苏惠<sup>1, 2</sup>

(1. 中国科学院大学电子电气与通信工程学院, 北京 100049;

2. 中国科学院国家空间科学中心, 北京 100190; )

**摘 要:** 针对空间互联网星地通信场景中的混叠信号分离精度不足问题, 提出了基于深度学习的堆叠时域卷积网络(stacked time domain convolutional network, Stacked-TCN)分离方法。首先, 对混合信号提取编码特征表示。然后, 通过时域卷积网络训练得到源信号的深层特征掩模, 将每个信号源的掩模与混合信号编码特征做 Hadamard 乘积, 得到源信号的编码特征表示。最后, 使用 1-D 卷积, 对源信号特征进行解码, 得到原始波形。实验采用负的比例不变信噪比(scale-invariant source to noise ratio, SISNR)作为网络训练的损失函数, 即单通道盲源分离性能的评价指标。结果表明, Stacked-TCN 方法与独立成分分析(independent component analysis, ICA)、非负矩阵分解(non-negative matrix factorization, NMF)两种经典分离算法和时域音频分离网络(time-domain audio separation network, TasNet)、Wave-U-Net 两种深度学习分离方法相比, 损失函数降低了 9.11~36.63dB。其中, 8PSK\_16QAM 分离损失为 -3.23dB@SNR=5dB, 比其他 4 种方法降低了 12.77dB。Stacked-TCN 方法具有更好的分离精度和噪声鲁棒性。

**关键词:** 欠定盲源分离; 同频干扰; 单通道; 时域卷积网络(TCN)

**中图分类号:** TN 911.7

**文献标志码:** A

## Single-channel blind source separation of spatial aliasing signal based on Stacked-TCN

ZHAO Mengchen<sup>1, 2</sup>, YAO Xiujuan<sup>2</sup>, WANG Jing<sup>2</sup>, DONG Suhui<sup>1, 2</sup>

(1. School of Electronic and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China; 2. National Space Science Center; CAS, Beijing 100190, China)

**Abstract:** Aiming at the problem of insufficient separation accuracy of the aliased signals in the satellite-to-ground communication scene, a Stacked-TCN method based on deep learning is proposed. First, the coding feature of the mixture is extracted. Then, the mask of the source is obtained through TCN. The Hadamard products is performed to the mask of source and the mixture feature to obtain the source feature. Finally, 1-D convolution is used to decode the source feature to obtain the waveform. SISNR is used as the loss function, that is, the evaluation index of single-channel blind source separation performance. The results show that the Stacked-TCN method reduce the loss by 9.11dB~ 36.63dB compared with algorithms of ICA, NMF, TasNet and Wave-U-Net, and reduce the loss by 12.77dB in 8PSK\_16QAM mixture when SNR is 5dB in which the loss is -3.23dB. This method has better separation accuracy and noise robustness.

**Keywords:** underdetermined blind source separation; co-frequency interference; single channel; time-domain convolutional network (TCN)

## 0 引言

近年来,国内外的卫星星座组网计划发展迅速,星地通信、星间通信使得频谱拥挤进一步加剧,同频信号混叠干扰难以避免<sup>[1]</sup>,空间通信电磁环境呈现时频混叠、空间交错,在现实空间通信环境中,鲁棒的信号处理通常需要自动信号分离<sup>[2, 3]</sup>。在密集的星座互联网通信场景下,达到接收端的源信号可能是多个通道的混叠信号,而地面站观测通道数目有限,因此欠定盲源分离的应用场景极为常见。常规的二步法、稀疏特征表示法等欠定分离方法<sup>[4-6]</sup>在单通道条件下分离效果欠佳,单通道通信信号盲源分离问题亟待解决。

研究者提出了许多解决分离问题的方法。独立成分分析 ICA<sup>[7]</sup>在超定分离方面表现良好,但在欠定盲源分离场景下效果欠佳。基于聚类法的稀疏分量分析(sparse component analysis, SCA)<sup>[8]</sup>需要已知混合信号的数目。时-频掩模方法<sup>[9-12]</sup>提高了分离精确度,但存在幅度和相位解耦,短时傅立叶变换需要较高分辨率的频率分解窗,限制在低延迟系统中的适用性。NMF<sup>[13-14]</sup>可以在时域进行分解,但在泛化为大数据的能力方面较弱。基于 Kalman 滤波<sup>[15]</sup>, LCL-FRESH 滤波<sup>[16]</sup>, 循环维纳滤波的单通道分离<sup>[17]</sup>计算复杂度高,实用效果有待提高。

随着大数据的发展和算力的提高,深度学习在时序信号处理如语音识别、语音分离<sup>[18-38]</sup>、通信信号调制识别<sup>[39]</sup>上取得了巨大成功,展示出了深度学习强大的特征提取和时序信号处理能力。然而,深度学习在通信信号处理上的应用多见于常规的调制识别与分类任务,在单通道通信信号分离等复杂任务上尚未涉足。

本文提出的通信信号单通道盲源分离方法,引入了数据驱动的深度学习方法,采用基于 stacked-TCN 的全卷积时域分离网络(convolution time-domain audio separation network, Conv-TasNet)分离框架。提出的方法首先使用 1-D 卷积对混合通信信号提取编码特征表示,将其作为输入。然后,使用时域卷积网络,训练得到纯净信号源的深层特征掩模,将每个信号源的掩模与得到的混合信号编码特征做 Hadamard 乘积,可以得到源信号的编码特征表示。最后,再次使用 1-D 卷积,对得到的源信号特征进行解码,得到原始波形。训练的损失函数采用比例不变信

噪比,以此度量掩模估计准确度和信号分离准确度。实验结果表明,该方法有效提高了单通道通信信号盲源分离精度,具有较好的噪声鲁棒性。

## 1 盲源分离数学模型

盲源分离(blind source separation, BSS)根据观测信号  $\mathbf{x}(t)$ , 利用信号源之间的独立性,使得估计信号  $\mathbf{s}^*(t)$  逼近于信号源  $\mathbf{s}(t)$ 。信号源为  $\mathbf{s}(t) = \{s_1(t), s_2(t), \dots, s_n(t)\}^T$ , 接收到的观测混合信号为  $\mathbf{x}(t) = \{x_1(t), x_2(t), \dots, x_m(t)\}^T$ , 估计信号为  $\mathbf{s}^*(t) = \{s_1^*(t), s_2^*(t), \dots, s_n^*(t)\}^T$ , 盲源分离数学模型<sup>[12, 18-21, 23-31]</sup>采用线性瞬时混合模型:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \quad (1)$$

式中,  $\mathbf{A}$  为混合矩阵;  $m$  代表源信号的个数;  $n$  代表接收天线阵元个数。当  $n < m$  时, 定义为欠定盲源分离。其中, 当  $n=1$  时, 定义为欠定条件下的单通道盲源分离。

单通道欠定盲源分离瞬时混合模型为

$$x(t) = \sum_{i=1}^N a_i s_i(t) \quad (2)$$

## 2 全卷积时域信号分离网络

为了对通信信号的深层特征进行有效提取, 以提高分离精度, 本文采用以 Stacked-TCN 为基础的 Conv-TasNet 全卷积时域信号分离网络框架进行训练。序列信号的准确分离需要较长的时间窗口信息, 即长时依赖性, 而 TCN 在序列信号建模和处理上的优越性能, 可以满足语音信号、通信信号等序列信号的长时依赖性建模。每个 TCN 为一个 1-D 单元块, 如图 1 所示。全卷积时域信号分离网络结构如图 2 所示。

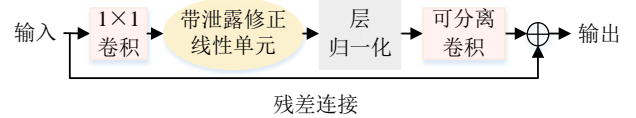


图 1 1-D 单元块结构图

Fig.1 Structure of 1-D block

### 2.1 混合信号编码特征表示

对一维混合通信源信号使用 1-D 卷积进行线性编码特征表示, 其中, 使用 512 组卷积核, 生成混合信号的多维编码特征, 将多维编码特征表示作为分离网络的输入:

$$\mathbf{x}_{encoder} = h_{encoder}(\mathbf{x}) = \mathbf{w}_1 * \mathbf{x} + \mathbf{b}_1 \quad (3)$$

式中,  $h_{encoder}(\cdot)$  为卷积操作;  $\mathbf{w}_1$  和  $\mathbf{b}_1$  分别为卷积核的权重和偏置。

## 2.2 源信号掩模生成

**步骤 1** 首先, 进行层归一化 (layer normalization, LN)。层归一化<sup>[40]</sup>是对一个中间层的所有神经元进行归一化, 使其输入分布在训练过程保持一致, 解决每次的梯度下降更新参数导致的内部协变量偏移问题, 从而增强模型的泛化能力, 同时避免了梯度消失和梯度爆炸。再使用  $1 \times 1$  卷积作为瓶颈层, 以减少特征通道的数目。

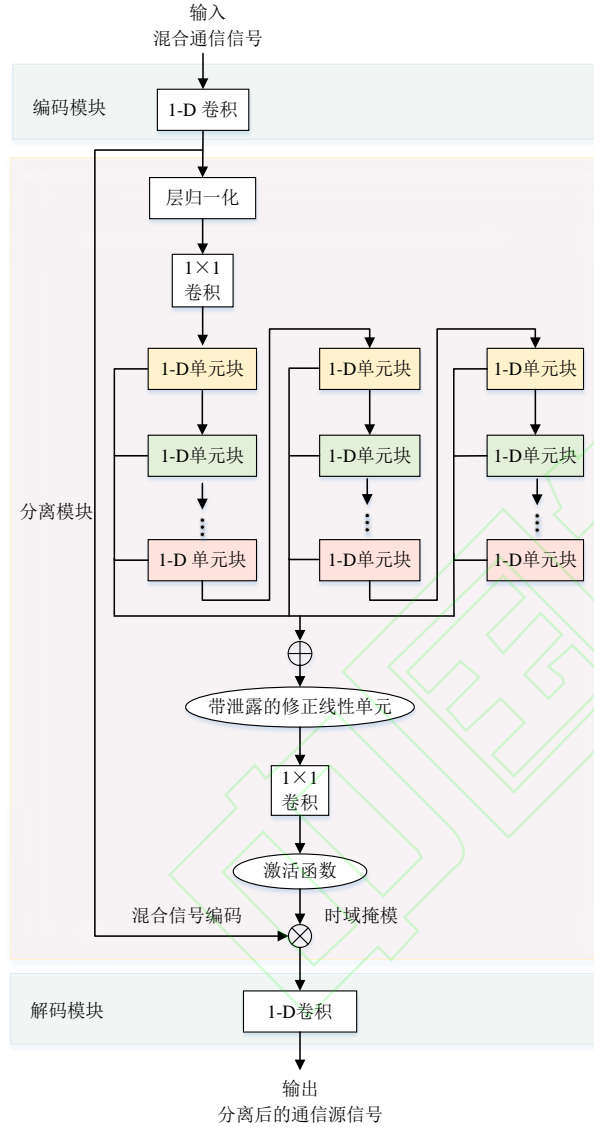


图 2 分离模块网络结构

Fig.2 Structure of separation

**步骤 2** 随后, 特征进入堆叠的 TCN, 每个堆叠块包含 4 个 TCN, 其卷积核为空洞卷积核<sup>[18]</sup>膨胀率依次增加, 分别为 1, 2, 4, 8。不同膨胀率的卷积核, 可以有效提取并融合多尺度的时序通信信号特征, 充分对信号分离的长时依赖性进行建模。其中, 每个 1-D 单元块采用深度可分离卷积 (depthwise separable convolution)。深

度可分离卷积<sup>[41]</sup>是普通卷积在通道上的分解, 将一个卷积核分裂为独立的深度卷积 (depthwise, DW) 和点向卷积 (pointwise, PW), 可以有效减少参数数量, 运算成本和模型大小, 用于构建轻量级网络<sup>[19]</sup>。1-D 单元块的残差网络连接方式, 可以防止随着网络层数加深导致的训练退化问题。此堆叠块进行 3 次重复连接, 将每一次堆叠输出作为下一次的输入, 进一步加深网络层数, 提取源信号的深层本质特征。对通信信号进行不同窗口不同尺度和深度的特征提取, 有效融合多层特征。

**步骤 3** 通过带泄露的修正线性单元 (Leaky ReLU) 作为激活函数, 由于其导数不存在零点, 可以防止负区间中神经元不学习问题。然后再次使用  $1 \times 1$  卷积瓶颈层, 恢复特征通道数量, 经过 sigmoid 激活函数, 训练获得两个源信号的时域掩模:

$$s_{mask} = f_{mask}^{stacked-TCN}(x_{encoder}) \quad (4)$$

式中,  $f_{mask}^{stacked-TCN}$  为包含 stacked-TCN 的分离模块, 用于生成时域掩模  $s_{mask}$ 。

**步骤 4** 混合信号的编码特征实际包含了两个源信号编码特征, 每个源信号存在潜在的时域掩模<sup>[18]</sup>, 通过时域掩模可以提取每个源信号的编码特征。将得到的每个源信号的时域掩模分别与混合信号的编码特征表示相乘, 得到两个通信源信号的特征编码:

$$s_{sep} = x_{encoder} \circ s_{mask} \quad (5)$$

式中,  $\circ$  是 Hadamard 乘积, 即两个操作数的按元素乘积。

## 2.3 源信号波形恢复

对分离得到的通信源信号特征编码, 采用 1-D 卷积反变换, 进行 512 维特征解码, 得到一维时域波形:

$$s^* = h_{decoder}(x_{sep}) \quad (6)$$

式中,  $h_{decoder}(\cdot)$  为解码器。

## 3 网络训练

衡量盲源分离性能的基本分离评价指标通常采用比例不变信噪比 (scale-invariant source-to-noise ratio, SISNR)<sup>[42]</sup>。SISNR 衡量了信号与分离误差的比值关系, SISNR 越高, 表明分离误差越低, 分离性能越好。计算之前将源



信号和分离后的源信号标准化为零均值,确保尺度不变。

网络训练时一般采用梯度下降法,需要最小化损失函数,因此采用负的 SISNR 作为损失函数,如式(7)所示,保证端到端的训练使得损失最小,比例不变信噪比最大,以确保模型训练的准确性。

$$\min_{\theta} \ell_{SISNR}(s, s^*) = -10 \log_{10} \left( \frac{\|s_{aim}\|^2}{\|s^* - s_{aim}\|^2} \right) \quad (7)$$

$$s_{aim} = \frac{\langle s^*, s \rangle s}{\|s\|^2}$$

通过反向传播梯度下降算法,进行参数更新:

$$\theta_{encoder} \leftarrow \theta_{encoder} - \eta \frac{\partial \ell_{SISNR}}{\partial \theta_{encoder}} \quad (8)$$

$$\theta_{mask} \leftarrow \theta_{mask} - \eta \frac{\partial \ell_{SISNR}}{\partial \theta_{mask}} \quad (9)$$

$$\theta_{decoder} \leftarrow \theta_{decoder} - \eta \frac{\partial \ell_{SISNR}}{\partial \theta_{decoder}} \quad (10)$$

反向传播梯度下降算法对编码部分、掩膜部分和解码部分的参数  $\theta = \{\theta_{encoder}, \theta_{mask}, \theta_{decoder}\}$  更新,其中,编码器参数集,掩膜参数集,解码器参数集分别为  $\theta_{encoder} = \{w_1, b_1\}$ ,  $\theta_{mask} = \{w_2, b_2\}$ ,  $\theta_{decoder} = \{w_3, b_3\}$ 。

## 4 实验与结果分析

实验包括两部分。4.1 为混合信号在 20dB 高信噪比下的分离实验,根据 SISNR 损失函数值,与 NMF、ICA 算法对比验证 Stacked-TCN 方法的分离性能。4.2 为不同信噪比的混合信号的分离实验,验证 Stacked-TCN 方法的泛化性能与噪声鲁棒性。

### 4.1 实验条件

实验采用的硬件资源为 Tesla k80 GPU, Intel Xeon E5 2.60GHz CPU, 深度学习框架为 PyTorch1.4。所有实验均采用五折交叉验证。

如图 3 所示,实验流程分为两部分:数据生成与混合部分,信号分离部分。

在数据生成与混合部分,通过软件定义无线电平台 GNUradio<sup>[39]</sup>生成 5 种调制方式 BPSK, 8PSK, QAM16, QAM64, PAM4 的通信数据,采样率为 1MHz,码速率为 125K symbol/s。参照表 1 中已有研究成果<sup>[43-46]</sup>,结合仿真平台的实际运行效率,选择的信噪比范围为 5~20dB,步长为 2.5dB。在仿真中,假定不同来源的混叠信号具有相同的频偏和定时偏差,本文重点讨论不

同算法中信噪比对分离效果的影响。

每类信号在每种信噪比下生成 1000 条纯净数据信号,每条数据含  $L=32768$  个采样点,即为 32.768ms,满足分离任务的长时依赖性需求。首先对纯净信号的幅值进行标准化处理。然后采用如式(1)所示的线性瞬时混合模型,进行信号混合。将 5 种不同调制方式的信号两两混合,得到 10 种混合方式的数据,在每种混合方式中,将同噪比的信号进行混合。10 种混合方式数据分别为 BPSK\_16QAM、8PSK\_64QAM、8PSK\_PAM4、64QAM\_PAM4、BPSK\_8PSK、BPSK\_64QAM、BPSK\_PAM4、8PSK\_16QAM、16QAM\_64QAM、16QAM\_PAM4 作为混合数据集,共 70000 条样本。

在分离部分,分为 3 个模块:1-D 卷积编码模块,分离模块,1-D 卷积解码模块。

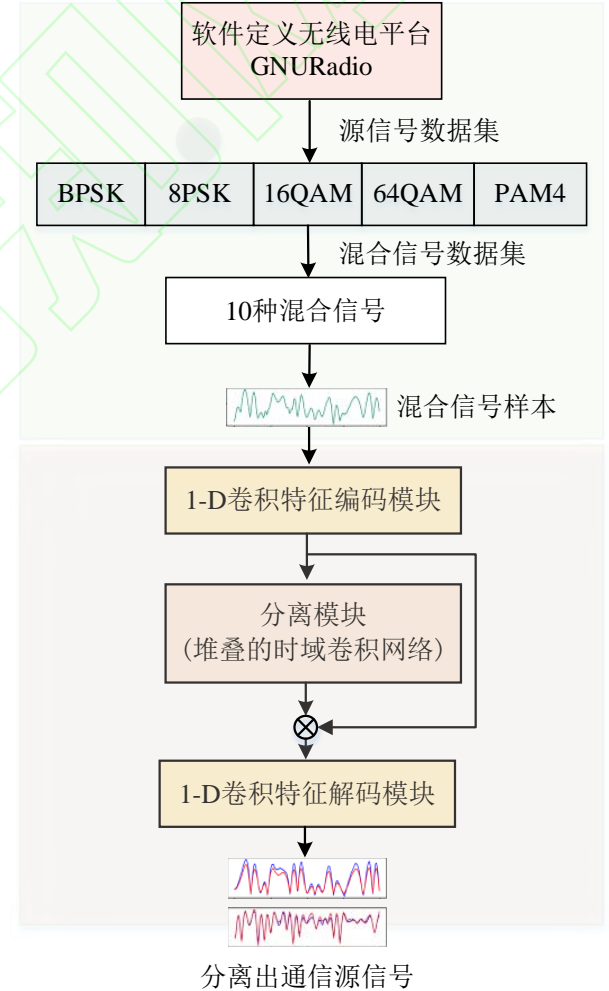


图 3 实验流程图

Fig. 3 Experimental flow chart

表 1 单通道盲分离算法比较  
Table 1 single-channel BSS algorithms

序号	方法	信号种类	SNR 范围(dB)	步长(dB)
1	[43]	BPSK	5-30	5
2	[44]	BASK 、BPSK	5-25	5
3	[45]	QPSK	12-22	2
4	[46]	雷达信号	10-30	0.5
5	[7]	EEG、ECG	/	/
6	[47]	语音信号	-6-9	3
7	[48]	语音信号	/	/
8	[49]	语音信号	/	/
9	Stacked-TCN	BPSK、8PSK、16QAM、64QAM、PAM4	5-20	2.5

#### 4.2 实验 1: SNR=20dB 下的算法性能对比

实验目的: Stacked-TCN 算法与 ICA<sup>[7]</sup>、NMF<sup>[47]</sup> 两种经典机器学习算法, TasNet<sup>[48]</sup>、Wave-U-Net<sup>[49]</sup> 两种深度学习方法在高信噪比

SNR=20dB 下进行对比分离试验, 验证在理想的低噪声的信号传输情况下, 不同算法的分离性能。4 种对比方法的基本参数配置见表 2。

表 2 不同算法的参数配置  
Table 2 Parameter configuration of 5 methods

算法\参数配置	参数	值
ICA	迭代次数	100
	迭代阈值	$1 \times 10^{-8}$
NMF	迭代次数	100
	基信号数, N	128
TasNet	帧长, L	64
	LSTM 隐层单元数	128
	LSTM-block 数, X	2
	卷积核大小, P	5
Wave-U-Net	block 数, X	5
	通道数	16-32-64-128-256
Stacked-TCN	编码器 filter 数, N	512
	帧长, L	16
	瓶颈层通道数, B	128
	卷积核大小, P	3
	1D-block 通道数, H	512
	1D-block 数, X	8
	Repeat 数, R	3

ICA<sup>[7]</sup>和 NMF<sup>[47]</sup>作为传统机器学习方法,是盲源分离领域的两种经典的算法。ICA 利用动态嵌入将单通道观测数据转化为多通道数据进行分离,在正定和超定分离上性能优越,在单通道欠定分离方面的分离精度欠佳。NMF 根据 Kullback-Leibler 散度最小化计算信源的基本矩阵和系数矩阵,实现信号分离。此类传统算法相当于浅层模型,不提取信号深层特征。TasNet<sup>[48]</sup>和 Wave-U-Net<sup>[49]</sup>作为两种深度学习分离方法,可以实现信号的单通道盲源分离。TasNet 网络采

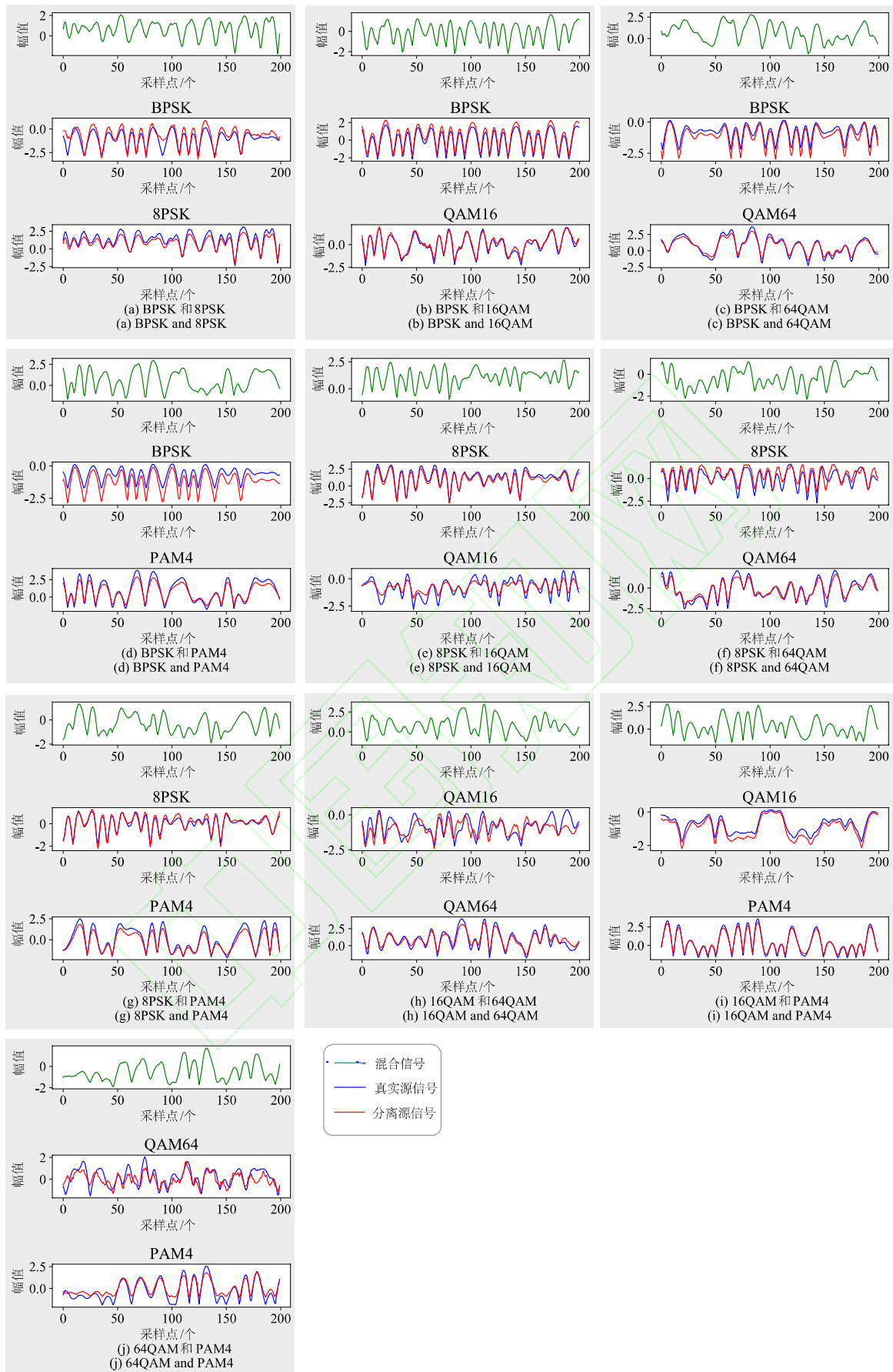
用编解码与分离模块的三级结构,分离模块中每个块采用长短时记忆网络 (long short-term memory, LSTM), 其大量参数显著增加了其计算成本。Wave-U-Net 网络通过重复进行特征图的下采样和卷积来处理时间序列上下文,在不同时间尺度上组合高级和低级特征,卷积产生的每个特征映射以原始信号的采样率作为分辨率,内存消耗较高。Stacked-TCN 方法在分离模块采用空洞卷积,参数量减少,与 LSTM 相比,下一步计算无需等待上一步输出。

表 3 20dB 混合信号下不同算法的损失值 dB  
Table 3 Loss of different algorithms on 20dB mixture dB

混合信号\算法	Stacked-TCN	TasNet	Wave-U-Net	ICA	NMF
BPSK_8PSK	-22.51	-2.94	-17.22	5.36	9.92
BPSK_16QAM	-18.71	-4.06	-17.02	1.64	6.32
BPSK_64QAM	-22.83	-1.44	-19.08	6.02	8.72
BPSK_PAM4	-28.44	-2.35	-12.09	5.92	8.19
8PSK_16QAM	-13.00	-3.63	-13.55	4.87	9.82
8PSK_64QAM	-5.96	-1.98	-11.31	1.07	7.67
8PSK_PAM4	-15.64	-1.95	-18.67	2.68	6.13
16QAM_64QAM	-6.32	-2.51	-1.51	5.48	8.73
16QAM_PAM4	-21.61	-2.25	-17.31	5.32	7.79
64QAM_PAM4	-5.53	-2.36	-11.92	2.50	5.98
平均值	<b>-16.05</b>	-2.55	-13.97	4.09	7.93

实验结果如表 3 所示, 可以看出 Stacked-TCN 方法损失最低, 为-16.05 dB, 其次是 Wave-U-Net 网络, 损失为-13.97dB, 第三是 TasNet 网络, 损失为-2.55dB, 最后是 ICA 和 NMF 算法, 损失分别为 4.09 dB 和 7.93dB。在单通道通信信号盲源分离任务中, 3 种深度学习方法在 SISNR 上的性能明显优于传统的 ICA、NMF 盲源分离算法, 其中, Stack-TCN 方法的分离性能最佳。该算法的损失与其他 4 种算法相比平均降低了 14.93 dB。

图 4 给出了 Stacked-TCN 算法在 SNR=15dB 情况下的 10 种混合信号的波形部分结果图, 可以看出, 分离结果与源信号具有高度吻合性。这是由于多层的堆叠时域卷积网络充分利用了通信源信号的深层次特征, 深度神经网络的优秀的拟合能力在学习时域掩模上发挥了重要作用。同时, 空洞卷积的不同时间窗口随层数递增, 对通信时序信号的长时依赖性进行了建模, 融合不同窗口长度特征。





#### 4.3 实验 2: SNR=5~20dB 时的对比实验

实验目的: 在信噪比 SNR=5~20dB 范围内, 进行对比分离试验, 验证在不同噪声的传输情况下, Stacked-TCN 算法的泛化性能与噪声鲁棒性。

图 5 给出了随着混合信号信噪比增加, 不同算法的损失结果变化。第一, 随着信噪比的增加, Stacked-TCN 方法的分离性能随之呈现增加的趋势, 分离准确度增加。第二, 3 种深度学习算法在不同信噪比下, 性能均优于两种机器学习算

法 ICA、NMF。即便在较低信噪比下, 分离精度亦占优势, 其中, Stacked-TCN 与 Wave-U-Net 效果最为突出, 对于噪声干扰, 鲁棒性较强。第三, Wave-U-Net 方法在 8PSK\_64QAM、64QAM\_PAM4 两种混合信号的分离中, 分离效果优于 Stacked-TCN, 但在其他情况下, 存在波动与不稳定性。而 Stacked-TCN 方法整体较为平稳, 呈现较好的下降趋势。

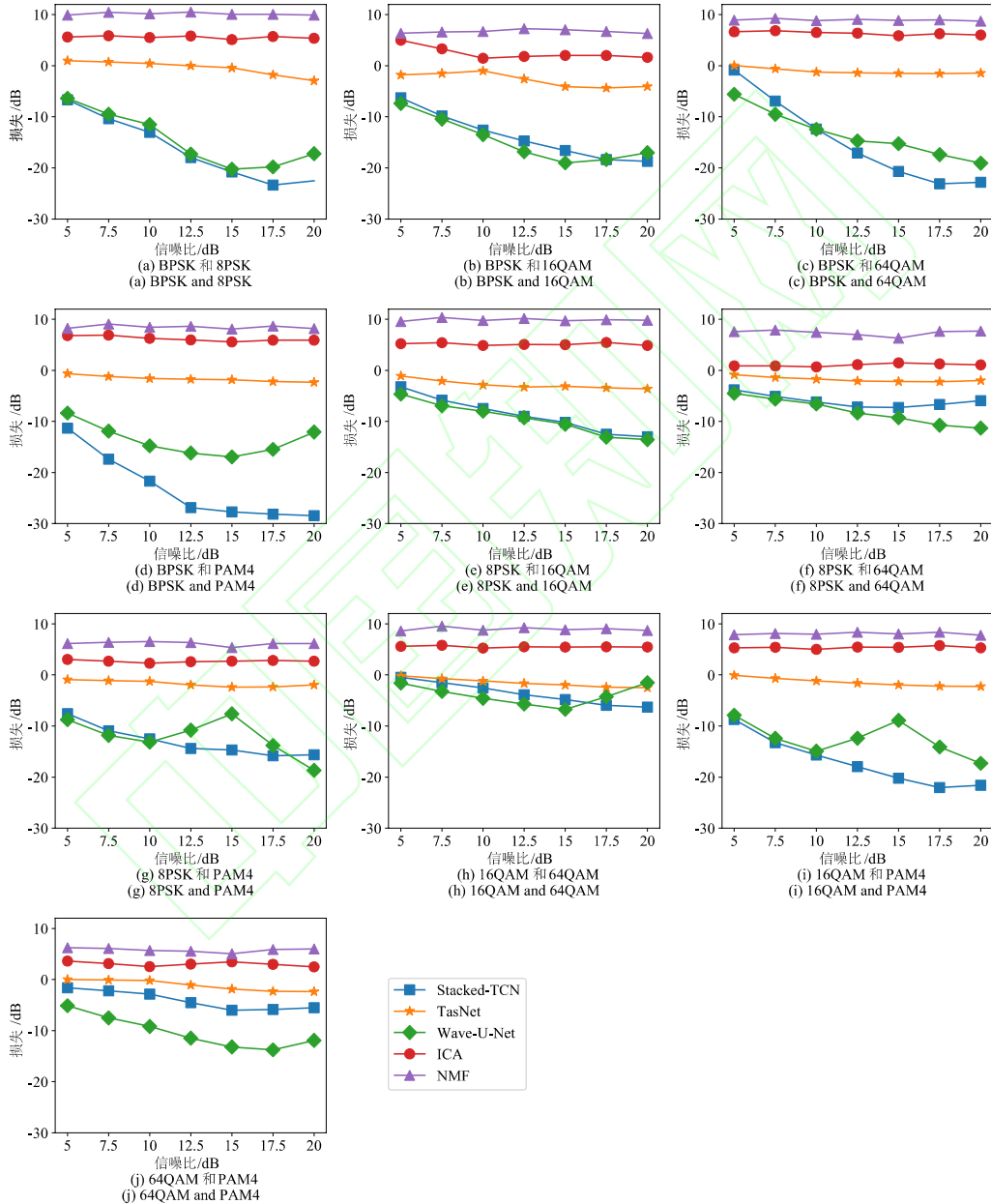


图 5 不同信噪比下的损失变化  
Fig.5 Loss under different SNR

这是由于 NMF、ICA 作为模型驱动的机器学习方法, 本质上是浅层模型, 适合样本小、模型精的任务。然而单通道盲源分离作为先验

知识极少的难点问题, 浅层模型不足以刻画其本质特征, 未能充分利用信号的深层特征与信息, 在泛化为大数据能力上较弱, 因此在单通

道极端病态条件下, 分离性能不佳。当任务复杂, 无法用模型驱动的方法精确刻画时, 数据驱动的深度学习方法可以弥补模型上的困难, 从大量样本中学习到深层本质特征, 具有强大的拟合能力, 满足信号分离等任务的需求。因此, Stacked-TCN、Wave-U-Net、TasNet 的分离效果优于 ICA、NMF。其中, TasNet 网络采用 LSTM 作为 block, 大量参数增加了计算成本, 为保证合理的网络模型和计算速度, 需控制 block 的数量, 对时序信号的建模准确性不够, 因此效果较差。Wave-U-Net 网络通过重复进行特征图的下采样和卷积来处理时间序列上下文, 效果较好, 但存在波动。Stacked-TCN 方法在分离模块采用空洞卷积和, 参数量减少, 同等网络模型大小, 可以增加 block 的重复次数, 同时深度可分离卷积保证了不同窗口尺度信息的融合, 因此在 3 种深度学习方法中, 其准确性和稳定性上最优。

表 4 处理每帧所用时间

Table 4 Times per frame ms		
算法	单帧时长	CPU 所用时间
Stacked-TCN 时域掩模法	0.032	$7.90 \times 10^{-3}$
时-频域掩模法	0.256	1.24

此外, 计算了 Stacked-TCN 时域掩模法与时-频域掩模法每帧的计算时间, 如表 4, 大多数时-频域掩模方法的信号分离研究中, STFT 的窗口长度至少为 256 点<sup>[9-12]</sup>, 因此本实验单帧时长取为 0.256 ms, 其计算时间为 1.24 ms, 较长的时间窗和计算时间, 使得系统最低时延增加。而 Stacked-TCN 时域掩模法在不降低分离准确度的前提下, 单帧时长可降至 0.032 ms, 其计算时间仅为  $7.90 \times 10^{-3}$  ms。具有较低的系统最小延迟。时-频域掩模法要成功地从时-频表示中分离源信号, 需要混合信号的高分辨率频率分解, 这需要较长的 STFT 时间窗, 该要求增加了系统的最小延迟, 这限制了其在实时及低延迟应用中的适用性, 因此, 越来越多的研究开始转向时域方法<sup>[18-31]</sup>。

表 5 分离混合信号所用时间

Table 5 Times of separation s	
算法	所用时间
NMF	$0.38 \times 10^{-4}$
ICA	0.15727
Wave-U-Net	0.20259
TasNet	0.23369
Stacked-TCN	0.33449

表 5 给出了不同算法分离混合信号的速度, 在单样本的分离速度方面, 时域分离方法中, NMF>ICA>Wave-U-Net>TasNet>Stacked-TCN, ICA、NMF 算法运行较快, 但未对信号充分建模, 牺牲了分离准确度。Wave-U-Net、TasNet、Stacked-TCN 运行速度相当。由于样本的长时采样点在 Stacked-TCN 中得到了充分利用, 使得分离性能提高。

## 5 结 论

基于深度学习框架 Stacked-TCN 的单通道通信信号盲源分离, 免去了繁杂的特征工程, 具有良好的泛化性能, 克服了 ICA、NMF 等传统机器学习方法的泛化大数据能力弱的缺陷, 解决了单通道分离中可行解爆炸问题。分离模块采用深度可分离卷积, 卷积核为空洞卷积, 减少了参数量, 在分离之前提取信号的多尺度窗口深层特征以学习源信号的掩模, 与 TasNet 和 Wave-U-Net 网络相比, 具有较小的网络规模和更优的分离性能。

实验结果表明, 该方法能够用于常见调制方式的混合信号的分离, 其损失函数比其他 4 种方法降低了 9.11~36.63 dB。其意义体现在以下 3 方面: 第一, 在 5dB 低信噪比下, 具备良好的分离性能和噪声鲁棒性。第二, 在保证分离精度的同时, 具有较短的单帧处理时长和较低的系统最小延迟。第三, 通过与信号处理系统的结合, 提高前端处理模块的混叠信号分离精度。

## 参考文献

- [1] 肖楠, 梁俊, 王伟, 等. 卫星认知无线网络中频谱感知与分配策略设计[J]. 中南大学学报(自然科学版), 2015, 46(12): 4520-4528.  
XIAO N, LIANG J, WANG W, et al. Design of spectrum sensing and allocating strategy for satellite cognitive radio networks[J]. Journal of Central South University (Science and Technology), 2015, 46(12): 4520-4528.
- [2] 蒋丽丽, 陈国彬, 张广泉. 认知无线网络中基于盲源分离的频谱检测算法[J]. 中国科技论文, 2017, 12(8): 868-872.

- JIANG L L, CHEN G B, ZHANG G Q. Blind source separation-based spectrum sensing in cognitive radio networks[J]. China Science Paper, 2017, 12(8): 868-872.
- [3] 杨柳, 张杭. 通信中的盲源分离问题及解决方案探讨[J]. 通信技术, 2014, 47(1): 1-6.
- YANG L, ZHANG H. Discussion of blind source separation problem and its solution in communication[J]. Communications Technology, 2014, 47(1): 1-6.
- [4] LI C, ZHU L, LUO Z. Underdetermined blind source separation of adjacent satellite interference based on sparseness[J]. China Communications, 2017, 14(4): 140-149.
- [5] LI Y, WANG Y, DONG Q. A novel mixing matrix estimation algorithm in instantaneous under-determined blind source separation[J]. Signal Image and Video Processing, 2020, 14(5): 1-8.
- [6] XIE Y, XIE K, XIE S. Underdetermined blind source separation for heart sound using higher-order statistics and sparse representation[J]. IEEE Access, 2019, 7: 87606-87616.
- [7] DAVIES M E, JAMES C J. Source separation using single channel ICA[J]. Signal Processing, 2007, 87(8): 1819-1832.
- [8] BOFILL P, ZIBULEVSKY M. Underdetermined blind source separation using sparse representations[J]. Signal Processing, 2001, 81(11): 2353-2362.
- [9] ISIK Y, ROUX J L, CHEN Z, et al. Single-channel multi-speaker separation using deep clustering[J]. arXiv preprint arXiv:1607.02173, 2016.
- [10] YU D, KOLBÆK M, TAN Z H, et al. Permutation invariant training of deep models for speaker-independent multi-talker speech separation[C]// Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2017: 241-245.
- [11] KOLBÆK M, YU D, TAN Z H, et al. Multi-talker speech separation with utterance-level permutation invariant training of deep recurrent neural networks[J]. IEEE Trans. on Audio Speech and Language Processing, 2017, 25(10): 1901-1913.
- [12] DO H D, TRAN S T, CHAU D T. Speech separation in the frequency domain with autoencoder[J]. Journal of Communications, 2020, 15(11): 841-848.
- [13] YOSHII K, TOMIOKA R, MOCHIIHASHI D, et al. Beyond NMF: time-domain audio source separation without phase reconstruction[C]// Proc. of the International Society for Music Information Retrieval, 2013: 369-374.
- [14] FÉVOTTE C, IDIER J. Algorithms for nonnegative matrix factorization with the  $\beta$ -divergence[J]. Neural computation, 2011, 23(9): 2421-2456.
- [15] 何继爱, 宋宇霄. Kalman 滤波下的多信号单通道盲源分离[J]. 信号处理, 2018, 34(7): 843-851.
- HE J A, SONG Y X. Blind source separation of the multi-signal single channel based on Kalman filtering[J]. Signal Processing, 2018, 34(7): 843-851.
- [16] 何继爱, 刘琳芝, 李英堂. LCL-FRESH 滤波器实现单通道盲源分离[J]. 信号处理, 2014, 30(2): 236-242.
- HE J A, LIU LZ, LI YT. Single-channel blind source separation achieved by the LCL-FRESH filter[J]. Signal Processing, 2014, 30(2): 236-242.
- [17] 宋阳, 王翔, 马忠正, 等. CFE 条件下基于循环维纳滤波的单通道分离算法[J]. 中国电子科学研究院学报, 2014, 9(2): 186-193.
- SONG Y, WANG X, MA Z Z, et al. Single channel signals separation algorithm based on cyclostationary filtering under CFE correction[J]. Journal of China Academy of Electronics and Information Technology, 2014, 9(2): 186-193.
- [18] LUO Y, MESGARANI N. Conv-TasNet:

- surpassing ideal time-frequency magnitude masking for speech separation[J]. IEEE Trans. on Audio, Speech, and Language Processing, 2019, 27(8): 1256-66.
- [19] CHOLLET F. Xception: Deep learning with depthwise separable convolutions[C]// Proc. of the IEEE conference on computer vision and pattern recognition, 2017: 1251-1258.
- [20] LUO Y, CHEN Z, YOSHIOKA T. Dual-path rnn: efficient long sequence modeling for time-domain single-channel speech separation[C]// Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2020: 46-50.
- [21] SHI Z, LIU R, HAN J. La furca: Iterative context-aware end-to-end monaural speech separation based on dual-path deep parallel inter-intra bil-STM with attention[J]. arXiv preprint arXiv:2001.08998, 2020.
- [22] SUBAKAN C, RAVANELLI M, CORNELL S, et al. Attention is all you need in speech separation[J]. arXiv preprint arXiv:2010.13154, 2020.
- [23] HENNEQUIN R, KHLIF A, VOITURET F, et al. Spleeter: a fast and efficient music source separation tool with pre-trained models[J]. Journal of Open Source Software, 2020, 5(50): 2154.
- [24] HAN C, LUO Y, LI C, et al. Continuous speech separation using speaker inventory for long multi-talker recording[J]. arXiv preprint arXiv: 2012.09727, 2020.
- [25] FAN C, TAO J, LIU B, et al. Deep attention fusion feature for speech separation with end-to-end post-filter method[J]. arXiv preprint arXiv:2003.07544, 2020.
- [26] LIU Y, DELFARAH M, WANG D L. Deep CASA for talker-independent monaural speech separation[C]// Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2020: 6354-6358.
- [27] NGUYEN V N, SADEGHI M, RICCI E, et al. Deep variational generative models for audio-visual speech separation[J]. arXiv preprint arXiv:2008.07191, 2020.
- [28] SHI J, XU J, FUJITA Y, et al. Speaker-conditional chain model for speech separation and extraction[J]. arXiv preprint arXiv:2006.14149, 2020.
- [29] WANG D L, CHEN J. Supervised speech separation based on deep learning: an overview[J]. IEEE Trans. on Audio, Speech, and Language Processing, 2018, 26(10): 1702-1726.
- [30] ZHAO J, GAO S, SHINOZAKI T. Time-domain target-speaker speech separation with waveform-based speaker embedding[C]// Proc. of the Interspeech, 2020: 1436-1440.
- [31] LUO Y, HAN C, MESGARANI N. Ultra-lightweight speech separation via group communication[J]. arXiv preprint arXiv:2011.08397, 2020.
- [32] ZHANG L, SHI Z, HAN J, et al. FurcaNeXt: End-to-end monaural speech separation with dynamic gated dilated temporal convolutional networks[C]// Proc. of the International Conference on Multimedia Modeling, 2020: 653-665.
- [33] KAVALEROV I, WISDOM S, ERDOGAN H, et al. Universal sound separation[C]// Proc. of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2019: 175-179.
- [34] PRÉTET L, HENNEQUIN R, ROYO-LETELIER J, et al. Singing voice separation: A study on training data[C]// Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2019: 506-510.
- [35] LUO Y, HAN C, MESGARANI N, et al. FaSNet: Low-latency adaptive beam forming for multi-microphone audio processing[C]// Proc. of the IEEE Automatic Speech Recognition and Understanding Workshop, 2019: 260-267.
- [36] LLUÍS F, PONS J, SERRA X. End-to-end



- music source separation: is it possible in the waveform domain?[J]. arXiv preprint arXiv:1810.12187, 2018.
- [37] BAHMANINEZHAD F, WU J, GU R, et al. A comprehensive study of speech separation: spectrogram vs waveform separation[J]. arXiv preprint arXiv:1905.07497, 2019.
- [38] XU C, RAO W, XIAO X, et al. Single channel speech separation with constrained utterance level permutation invariant training using grid lstm[C]// Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2018: 6-10.
- [39] O'SHEA T J, CORGAN J, CLANCY T C. Convolutional radio modulation recognition networks[C]// Proc. of the International Conference on Engineering Applications of Neural Networks, 2016: 213-226.
- [40] XIONG R, YANG Y, HE D, et al. On layer normalization in the transformer architecture[C]//Proc. of the International Conference on Machine Learning, 2020: 10524-10533.
- [41] KHAN Z Y, NIU Z. CNN with depth wise separable convolutions and combined kernels for rating prediction[J]. Expert Systems with Applications, 2020, 170: 114528.
- [42] ISIK Y, ROUX J L, CHEN Z, et al. Single-channel multi-speaker separation using deep clustering[J]. arXiv preprint arXiv:1607.02173, 2016.
- [43] 王少波, 郭英, 睦萍, 等. 欠定条件下同步组网跳频信号盲源分离方法[J]. 计算机工程, 2020, 46(10): 166-72+81.
- WANG S B, GUO Y, SUI P, et al. Under-determined blind source separation method for synchronous network frequency-hopping signals[J]. Computer Engineering, 2020, 46(10): 166-72+81.
- [44] 赵知劲, 黄艳波. 基于小波包分解的单通道盲源分离算法[J]. 通信技术, 2017, 50(3): 425-429.
- ZHAO Z J, HUANG Y B. Single-channel blind-source separation algorithm based on wavelet packet decomposition[J]. Communications Technology, 2017, 50(3): 425-429.
- [45] 郭一鸣, 杨勇, 张冬玲, 等. 基于 SIC 的单通道同频混合信号低复杂度盲分离算法[J]. 信号处理, 2015, 31(7): 860-866.
- GUO Y M, YANG Y, ZHANG D L, et al. Blind separation algorithm of single-channel co-frequency signals with low complexity based on SIC[J]. Signal Processing, 2015, 31(7): 860-866.
- [46] 张建中, 文树梁, 谭澄, 等. 盲源分离联合阻塞矩阵抗雷达主瓣干扰研究[J]. 现代防御技术, 2018, 46(1): 135-140.
- ZHANG J Z, WEN S L, TAN C, et al. Radar main-lobe jamming suppression based on the method of BSS-BM union[J]. Modern Defense Technology, 2018, 46(1): 135-140.
- [47] WENINGER F, ROUX J L, HERSHEY J R, et al. Discriminative NMF and its application to single-channel source separation[C]// Proc. of the 15th Annual Conference of the International Speech Communication Association, 2014: 865-869.
- [48] LUO Y, MESGARANI N. Tasnet: time-domain audio separation network for real-time, single-channel speech separation[C]// Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2018: 696-700.
- [49] STOLLER D, EWERT S, DIXON S. Wave-u-net: a multi-scale neural network for end-to-end audio source separation[J]. arXiv preprint arXiv:1806.03185, 2018.

## 作者简介

赵孟晨 (1996-), 女, 硕士研究生, 主要研究方向为盲源分离、信号处理、机器学习与深度学习。

E-mail: zhaomengchen0484@163.com

姚秀娟 (1977-), 女, 研究员, 博士研究生导



师，主要研究方向为空间频谱感知、空间频率规划与干扰仿真技术。

E-mail: yaoxj@nssc.ac.cn

**王 静** (1990-), 女, 博士后, 主要研究方向为空间频谱感知、高动态时空干扰规避技术。

E-mail: wangj@nssc.ac.cn

**董苏惠** (1994-), 女, 博士研究生, 主要研究方向为空间互联网干扰仿真技术。

E-mail: dongsuhui17@mails.ucas.ac.cn

