

## 基于 Swin-Transformer 的短波协议信号识别

朱政宇<sup>1,2,3</sup>, 陈鹏飞<sup>1</sup>, 王梓暄<sup>1</sup>, 巩克现<sup>1</sup>, 吴迪<sup>4</sup>, 王忠勇<sup>1</sup>

(1. 郑州大学电气与信息工程学院, 河南 郑州 450001; 2. 郑州大学河南省智能网络和数据分析国际联合实验室, 河南 郑州 450001; 3. 郑州大学电子材料与系统国际联合研究中心, 河南 郑州 450001; 4. 信息工程大学数据与目标工程学院, 河南 郑州 450001)

**摘 要:** 针对短波复杂信道环境下信号所属协议识别困难的问题, 提出一种基于 Swin-Transformer 神经网络模型的短波协议信号识别算法。首先使用时频分析方法得到信号的灰度时频图作为神经网络的输入; 其次设计一种基于 Swin-Transformer 的神经网络模型, 对信号时频图进行特征提取; 最后将特征与协议建立映射关系, 从而实现信号协议的识别。仿真实验结果表明, 在信噪比大于  $-4$  dB 的高斯信道下, 所提算法的识别准确率接近 100%, 高于现有算法。此外, 在强干扰以及多径时延衰落的信道条件下, 所提算法仍具有较高的短波协议信号识别率。

**关键词:** 短波协议信号识别; 神经网络; 时频分析; 多径时延衰落; Swin-Transformer

中图分类号: TN92

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2022209

## Short wave protocol signals recognition based on Swin-Transformer

ZHU Zhengyu<sup>1,2,3</sup>, CHEN Pengfei<sup>1</sup>, WANG Zixuan<sup>1</sup>, GONG Kexian<sup>1</sup>, WU Di<sup>4</sup>, WANG Zhongyong<sup>1</sup>

1. School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou 450001, China

2. Joint International Laboratory of Intelligent Network and Data Analysis in Henan Province, Zhengzhou University, Zhengzhou 450001, China

3. National Center for International Joint Research of Electronic Materials and Systems, Zhengzhou University, Zhengzhou 450001, China

4. College of Data Target Engineering, Information Engineering University, Zhengzhou 450001, China

**Abstract:** Aiming at the problem that it is difficult to identify the protocol to which the signal belongs in the complex SW channel environment, a SW protocol signal recognition algorithm based on Swin-Transformer neural network model was proposed. Firstly, the gray-scale time-frequency map of the signal was obtained by using the time-frequency analysis method as the input of the neural network. Secondly, a neural network model based on swing transformer was designed to extract the features of the signal time-frequency map. Finally, the mapping relationship between the features and the protocol was established to realize the recognition of the signal protocol. The simulation results show that the recognition accuracy of the proposed algorithm is close to 100% in the Gaussian channel with SNR greater than  $-4$  dB, which is higher than the existing algorithms. In addition, under the channel conditions of strong interference and multipath delay fading, the proposed algorithm still has a high SW protocol signals recognition rate.

**Keywords:** SW protocol signals recognition, neural network, time frequency analysis, multipath delay fading, Swin-Transformer

收稿日期: 2022-07-04; 修回日期: 2022-10-19

通信作者: 巩克现, ggkx@163.com

基金项目: 国家重点研发计划基金资助项目 (No.2019QY0302); 中国博士后科学基金资助项目 (No.2020M682345); 河南省高校科技创新人才支持计划资助项目 (No.23HASTIT019); 河南省博士后经费资助项目 (No.202001015)

**Foundation Items:** The National Key Research and Development Program of China (No.2019QY0302), China Postdoctoral Science Foundation Funded Project (No.2020M682345), Program for Science and Technology Innovation Talents in Universities of Henan Province (No.23HASTIT019), Henan Postdoctoral Foundation Program (No.202001015)

## 0 引言

短波通信凭借其设备简便、通信距离远等优点,广泛应用于多种通信场景中。为保证在不同通信系统下信息的有效传输,短波频段有种类繁多的通信协议<sup>[1]</sup>。由于短波信道存在多径衰落、多普勒频移以及人为干扰等特点,识别短波信号所属协议具有一定的难度。

目前,用于通信协议识别的传统算法主要分为基于特征提取的识别算法<sup>[2-3]</sup>和基于模板匹配的识别算法<sup>[4-6]</sup>。基于特征提取的识别算法通过对提取的信号特征进行分类实现信号协议识别,文献[2]使用调制识别中的熵距离参数分离频移键控(FSK, frequency shift keying)模式的110A<sup>[7]</sup>与STANAG4285<sup>[8]</sup>这2种信号,但该算法仅适用于识别调制方式不同的信号;文献[3]通过识别信号中心频率、调制方式以及编码方式等属性形成信号特征向量,使用支持向量机对特征向量进行分类实现信号识别,但无法解决在低信噪比(SNR, signal to noise ratio)下信号特征难以提取的问题。文献[4]提出使用时域子序列匹配算法识别STANAG4285等相移键控(PSK, phase shift keying)信号,并提出一种频域波峰匹配算法识别2GALE<sup>[9]</sup>等FSK信号;文献[5]通过对待识别信号与构造的信号频谱模板进行相关运算,实现对未知短波协议信号的识别,但算法性能受频偏影响较大;文献[6]在频域模板匹配算法的基础上使用相位差分抑制频偏,对特定短波信号的识别具有较好的效果;但基于模板匹配的识别算法存在判决门限难以确定、对先验信息需求高等缺点。

近年来,随着深度学习的快速发展,神经网络算法广泛应用于信号处理领域,给通信信号识别技术带来了新思路。在信号调制识别领域,文献[10]使用原始同向正交信号作为输入,设计一维卷积神经网络(CNN, convolutional neural network),能够有效识别5种调制方式信号。文献[11]关注调制信号的时频特性,使用时频图作为CNN的输入,对多种信号取得了较好的识别效果。针对闭集识别存在易把干扰信号识别为有效信号的问题,文献[12]对神经网络中常用的Softmax分类器加以改进,并采用改进的GE2E(generalized end-to-end)<sup>[13]</sup>损失函数,达到开集识别的效果。在短波信号识别领域,文献[14]使用信号时频图作为ConvNet输入,对具有2、4、8这3种载波数以及窄、中、宽这3种载

波间隔的9种模式进行区分,但所识别的信号模式有限;文献[15-16]利用特定信号在时频图上呈现的视觉特征,使用卷积神经网络对灰度时频图进行特征提取和映射,实现信号所属协议的识别,识别准确率与传统方法相比有明显提升。CNN感受野大小对捕捉时频图中的特征差异有重要作用,文献[17]将ResNet<sup>[18]</sup>中的卷积核改进为扩张卷积核,在不增加参数数量的情况下扩大感受野,进一步提升短波协议信号的识别率,但仍未解决CNN中卷积层存在的感受野相对有限的问题。

Swin-Transformer<sup>[19]</sup>是一种基于自注意力机制<sup>[20]</sup>的神经网络模型,有对全局特征进行建模的能力,可以捕捉不同短波协议信号在时频图上呈现出的视觉特征差异。基于此,本文提出一种基于Swin-Transformer的短波信号协议识别算法。在多种复杂信道环境下,与CNN识别算法相比,所提算法的信号识别准确率明显提升。本文具体研究工作如下。

1) 分析了短波协议信号在时频图上视觉特征的成因,建立一个用于训练神经网络的短波信号时频图数据集。数据集包含9种常见的短波协议,为了模拟真实通信环境并提高数据多样性,在高斯白噪声环境下进行仿真,并对信号进行随机截断,模拟非合作通信无法确保从起始时刻接收信号的场景。

2) 研究了引入自注意力机制的Transformer模型结构以及Swin-Transformer的特点,并分析了Swin-Transformer可以用于短波协议信号识别的原因。提出一种基于Swin-Transformer的短波协议信号识别算法。

3) 仿真结果验证了所提算法的有效性。在短波多径时延、瑞利衰落以及强混叠的信道环境下,基于Swin-Transformer的短波信号协议识别算法均具有较高的识别率。与现有算法相比,所提算法在识别率上明显提升。

## 1 特定短波信号的时频特性分析

不同短波通信协议规定的帧结构、调制方式、载波中心频率以及信号带宽等差异较大,而这些属性的差异导致信号在时频图呈现出不同的视觉特性。其中,CLOVER2000<sup>[21]</sup>是一种典型的采用多载波调制的短波协议,其采取了多种技术来适应短波复杂信道环境,在短波通信中有广泛应用;LINK11<sup>[22]</sup>作为一种活跃时间较长的短波协议,是LINK16和LINK22协议的研发基础,对后续协议

的研究有重要参考意义。因此，本节以 CLOVER2000 和 LINK11 协议为例，对短波协议信号的时频特性做出具体分析。

时频分析结合时域和频域的特点，同时将信号的时间信息和频率信息展示在一幅时频图中，对时变非平稳信号的分析具有独特优势。常用的时频分析方法主要有短时傅里叶变换（STFT, short time Fourier transform）、Wigner-Ville 分布和小波变换等。STFT 由于计算量小、不存在交叉项等优点被广泛应用<sup>[23]</sup>，本文采取基于短时傅里叶变换的时频分析方法。

短时傅里叶变换的定义为

$$\text{STFT}(t, f) = \int_{-\infty}^{+\infty} s(\tau)w^*(\tau - t)e^{-j2\pi f\tau} d\tau \quad (1)$$

其中， $t$  为时间， $f$  为频率， $s(t)$  为信号， $w^*(\cdot)$  为窗函数。

CLOVER2000 信号的波形由 8 个音频并行组成，其中，最低频率为 625 Hz，最高频率为 2 375 Hz，相邻音频发送间隔为 250 Hz。时域上，每个音频的发送时序按照音频序号从小到大间隔 2 ms 依次发送。CLOVER2000 脉冲序号与频率的对应关系如表 1 所示。

表 1 CLOVER2000 脉冲序号与频率的对应关系

脉冲序号	频率/Hz
0	625
1	1 625
2	875
3	1 875
4	1 125
5	2 125
6	1 375
7	2 375

CLOVER2000 信号  $s(t)$  可以表示为

$$s_c(t) = \sum_{i=0}^7 \sum_n P c_{i,n} e^{j2\pi f_i t} g(t - nT_s - i\Delta t) \quad (2)$$

其中， $P$  为信号的平均功率， $c_{i,n}$  为第  $i$  个音频的第  $n$  个码元， $f_i$  为第  $i$  个音频的频率， $g(t)$  为成形脉冲， $T_s$  为脉冲持续时间， $\Delta t = 2 \text{ ms}$  为相邻音频的发送间隔。

对 CLOVER2000 信号进行短时傅里叶变换得到其灰度时频图，如图 1 所示。从图 1 可以看出，CLOVER2000 信号在时频上既呈现出了频域的 8 个音频，也体现出了每个音频发送时序不同的特点。

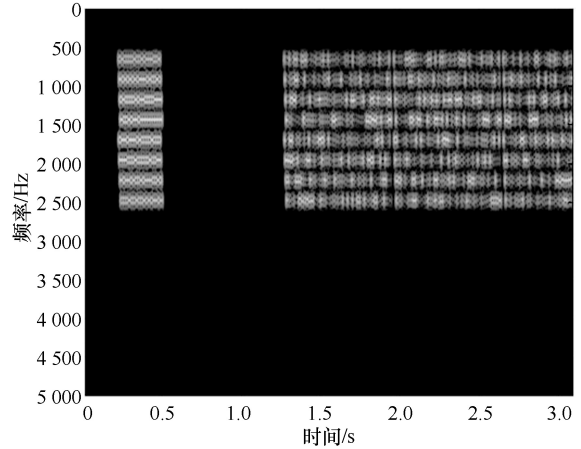


图 1 CLOVER2000 信号灰度时频图

LINK11 信号共有 6 种工作模式，本文以其中的网络控制站报告信号为例分析 LINK11 的时频特性<sup>[22]</sup>，其帧结构如图 2 所示。

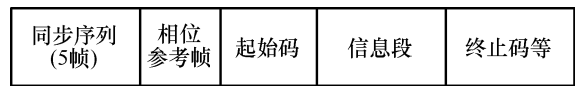


图 2 网络控制站报告信号帧结构

图 2 中，同步序列由 605 Hz 和 2 915 Hz 音频组成，605 Hz 音频为多普勒校正音频，其功率比 2 915 Hz 的同步音频大 5~7 dB，此外，相邻帧的同步音频以  $\pi$  进行相移，设初始相位为 0，同步序列信号可表示为

$$s'(t) = \sum_{n=1}^5 A_1 \cos(2\pi f_1 t) - (-1)^n A_{16} \cos(2\pi f_{16} t) \quad (3)$$

其中， $A_1$  为多普勒校正音频幅度， $A_{16}$  为同步音频幅度， $f_1 = 605 \text{ Hz}$ ， $f_{16} = 2 915 \text{ Hz}$ 。

位于同步序列之后的相位参考帧、起始码、信息段以及终止码等均由 16 个音频分量组成。除了 605 Hz 和 2 915 Hz 之外的 14 个音频频率位于 935~2 365 Hz 之间，相邻音频频率间隔为 110 Hz。多普勒校正音频不携带信息，采用四相移相键控（QPSK, quaternary phase shift keying）制方式，除多普勒校正音频外的 15 个音频，每帧携带 2 bit 信息，采用  $\frac{\pi}{4}$  四相相对相移键控（QDPSK, differential QPSK）的调制方式。其信号可表示为

$$s(t) = \sum_{i=2}^{16} \sum_n (A_i \cos(f_i t) + A_i \cos(f_i t + \varphi_{i,n})) g(t - nT_s) \quad (4)$$

其中， $A_i$  为第  $i$  个音频的幅度， $\varphi_{i,n}$  为第  $n$  帧第  $i$  个音频的相位。图 3 给出了 LINK11 信号频谱。

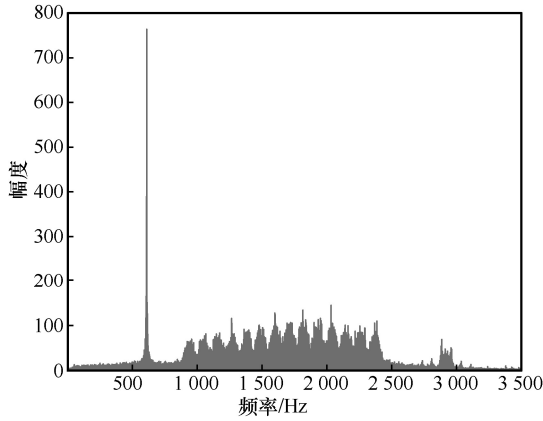


图 3 LINK11 信号频谱

对整段 LINK11 信号进行 STFT，得到如图 4 所示的时频图。

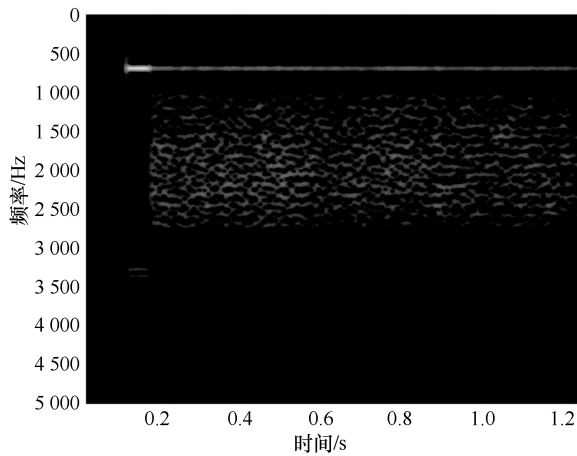


图 4 LINK11 信号时频图

由图 1 和图 4 可知，CLOVER2000 和 LINK11 信号时频图呈现出截然不同的视觉特征，这是由于不同短波协议在制定过程中，对信号调制方式、传输模式以及帧结构等做出了不同设计。因此，通过对短波信号的时频进行视觉特征提取可以实现对不同短波协议信号的识别。

### 2 Swin-Transformer 网络模型结构

近年来，神经网络凭借其强大的特征提取能力在图像处理等领域取得了巨大成功，其中，Vaswani 等<sup>[20]</sup>提出基于编码器解码器框架的 Transformer 网络模型，该模型引入多头自注意力（MSA, multi-head self-attention）机制，具备学习全局特征的能力，其编码器结构如图 5 所示。

假设输入序列长度为  $L$ ，batch size 大小为  $B$ ，则 Transformer 编码器的输入张量可以表示为  $M \in \mathbb{R}^{B \times L}$ 。M 首先经过 Input Embedding 映射成一个  $d_x$  维的 Embedding，再与位置编码相加得到 Transformer Block 的输入  $I \in \mathbb{R}^{B \times d_x \times L}$ ，这里的位置编码既可以采用正弦形式，也可以是通过学习得到的参数。每个 Transformer 模块都由多头自注意力模块和两层的前馈网络（FFN, feed forward network）组成。其中 MSA 和 FFN 的输入、输出都采用残差连接的方式，输出还需要进行层归一化，整个 Transformer Block 过程可以表示为

$$O_A = I + \text{LayerNormal}(\text{MSA}(I)) \tag{5}$$

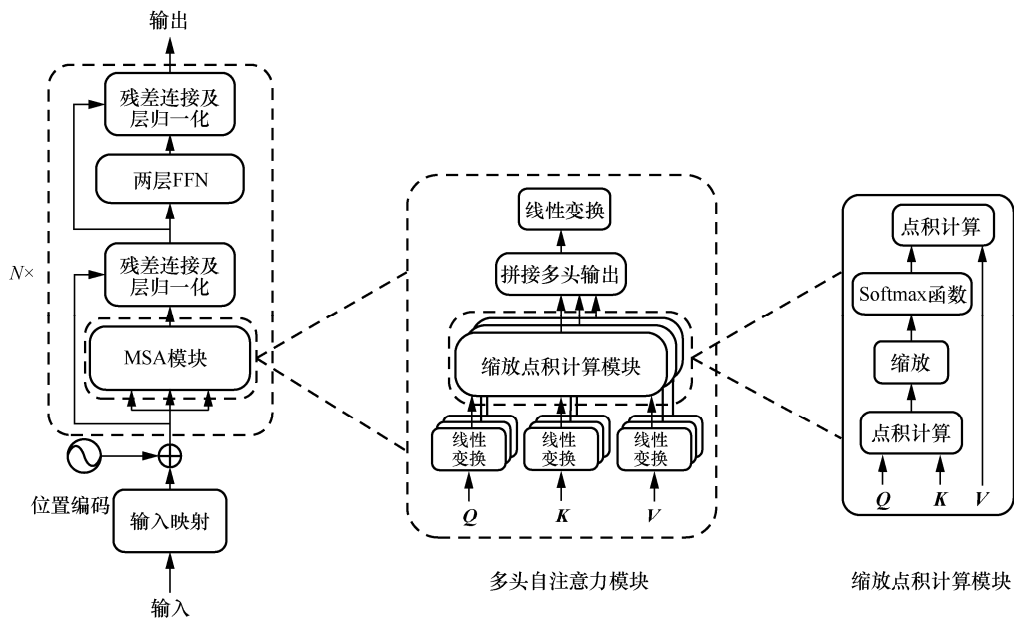


图 5 编码器结构

$$\mathbf{O} = \mathbf{O}_A + \text{LayerNormal}(F_2(\text{ReLu}(F_1(\mathbf{O}_A)))) \quad (6)$$

其中,  $\mathbf{O}_A$  为 MSA 的输出,  $\mathbf{O}$  为 Transformer 模块的输出,  $F_1(\cdot)$  和  $F_2(\cdot)$  分别为前馈网络的第一层和第二层, 其形式为  $F(x) = Wx + b$ 。

MSA 是 Transformer 的核心机制, 对于输入  $\mathbf{X} \in \mathbb{R}^{L \times d_x}$ , 其提供 3 个大小分别为  $\mathbb{R}^{d_x \times d_q}$ 、 $\mathbb{R}^{d_x \times d_k}$ 、 $\mathbb{R}^{d_x \times d_v}$  的权重矩阵分别与  $\mathbf{X}$  进行矩阵乘法得到矩阵  $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ , 即

$$\begin{cases} \mathbf{Q} = \mathbf{XW}^q \\ \mathbf{K} = \mathbf{XW}^k \\ \mathbf{V} = \mathbf{XW}^v \end{cases} \quad (7)$$

假定 MSA 有  $h$  个“头”, 则每一个“头”的输出  $\mathbf{A}_i$  定义为

$$\mathbf{A}_i = \text{Softmax}\left(\frac{\mathbf{Q}_i \mathbf{K}_i^T}{\sqrt{d}}\right) \mathbf{V}_i \quad (8)$$

其中,  $\mathbf{Q}_i$ 、 $\mathbf{K}_i$  和  $\mathbf{V}_i$  分别为

$$\begin{cases} \mathbf{Q}_i = \mathbf{W}_q^i \mathbf{Q} \\ \mathbf{K}_i = \mathbf{W}_k^i \mathbf{K} \\ \mathbf{V}_i = \mathbf{W}_v^i \mathbf{V} \end{cases} \quad (9)$$

这里,  $\mathbf{W}_q^i$ 、 $\mathbf{W}_k^i$  和  $\mathbf{W}_v^i$  分别是对  $\mathbf{Q}$ 、 $\mathbf{K}$  和  $\mathbf{V}$  进行线性拆分的矩阵。将每个“头”的输出拼接后进行线性变换, 使每个“头”的输出进一步融合, 即可得到最终的输出  $\mathbf{Y}$ , 此过程可以表示为

$$\mathbf{Y} = \text{Concat}(\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_h) \mathbf{W}^O \quad (10)$$

其中,  $\mathbf{W}^O \in \mathbb{R}^{d_x \times d_x}$  是一个线性变换矩阵。

鉴于 Transformer 在自然语言处理领域取得的成功, 文献[24]提出 Vision Transformer, 将 Transformer 架构成功用于图像处理领域。具体地, 首先将一个 224 像素  $\times$  224 像素大小的图像均分为 196 个 16 像素  $\times$  16 像素大小的区域, 其次将每个区域视为一个长度为 256 的向量输入 Transformer 编码器, 并增加

一个分类向量来表示全局特征, 最后将分类向量经过 Transformer 编码器的输出用于 Softmax 进行分类。

Swin-Transformer 在 Vision Transformer 的基础上引入 CNN 中常用的层次化构建方式, 在不同的层次对不同大小的窗口在内部进行 MSA 操作, 大幅降低了计算复杂度。同时通过滑动窗口操作增加窗口与窗口的信息交互, 确保模型性能不会降低。具体来说, 每一个 Swin-Transformer 模块均采用与 Transformer 编码器类似的架构, 但与 Transformer 编码器不同的是, Swin-Transformer 模块将普通的 MSA 改进为 Window-MSA (W-MSA) 以及 Shift Window-MSA (SW-MSA)。SW-MSA 工作原理如图 6 所示, 本文采用的 Swin-Transformer 网络模型具体结构如图 7 所示。

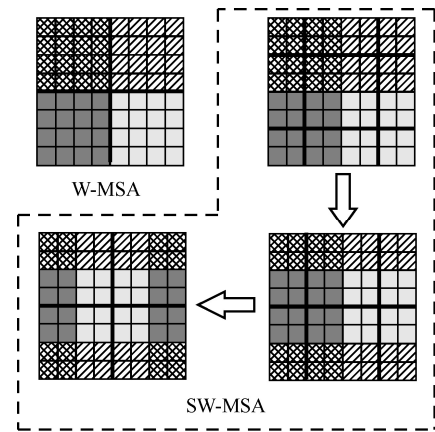


图 6 SW-MSA 工作原理示意

W-MSA 模块将输入特征图划分为多个窗口, 在每个窗口内部进行 MSA 操作从而大大减少了计算量。由于浅层网络的特征图尺寸相对较大, 因此浅层的 W-MSA 模块将特征图划分为更多的窗口, 使每个窗口尽可能小, 随着网络的加深, 特征图划分的窗口相应减少。

为了使不同的窗口间进行信息交互, 每个 W-MSA 模块后面会紧接一个 SW-MSA 模块。

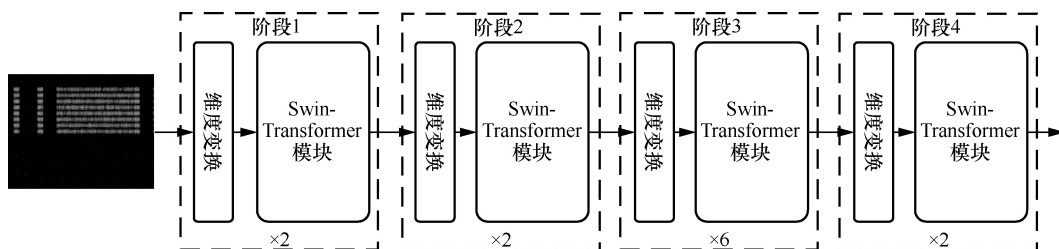


图 7 Swin-Transformer 网络模型具体结构

SW-MSA 模块首先对特征图重新划分窗口, 其次根据特定的规则对窗口进行移位, 对移位后的特征图进行窗口内的 MSA 操作, 此时的一个窗口内同时包含了多个移位前的其他窗口的特征, 从而发挥出 Transformer 长距离感知的优势。

Transformer 的全局感知能力对短波协议识别具有重要意义。CLOVER2000 与 2GALE 时频图局部和整体对比如图 8 所示。从图 8 可以看出, 调制方式为 8FSK 的 2GALE 信号的时频图与具有 8 个音频的 CLOVER2000 信号的时频图在局部具有类似的特征, 但由于二者帧结构、带宽以及中心频率等存在差异, 因此可以凭借全局特征对 2 种信号进行分辨。综上, 对图像具有长距离感知能力的 Swin-Transformer 可以通过对信号在时频图上呈现出的视觉特性进行特征提取, 从而实现短波信号的协议识别。

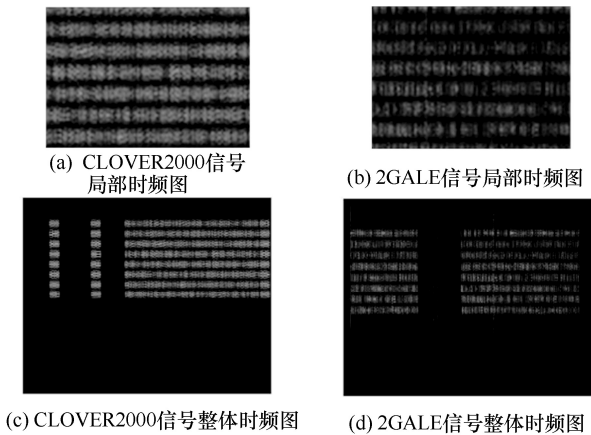


图 8 CLOVER2000 与 2GALE 时频图局部和整体对比

### 3 仿真测试与分析

短波协议数量众多, 短波协议识别暂时没有一个公开的标准数据集, 因此, 本文从目前常见的短波协议中选取了具有一定代表性的 110A、110B<sup>[25]</sup>、2GALE、3GALE<sup>[26]</sup>、CLOVER2000、CIS-45、LINK11、PRC4+4 和 STANAG4285 这 9 种短波协议信号的时频图制作数据集。其中, CLOVER2000 和 LINK11 前文已有介绍; 2GALE、3GALE 分别为短波第二代、第三代自动链路建立协议, 分别使用 FSK 和 PSK 波形对信道进行探测实现信道评估; 110A、110B、STANAG4285 为短波通信军用标准, 采用了多种纠错编码、交织、加扰等技术, 并且信号按照严格的帧结构进行发送, 包含了多数短波协议的特点; CIS-45 及 PRC4+4 分别为典型的采用

OFDM 和多载波调制的短波协议, 而正交频分复用 (OFDM, orthogonal frequency division multiplexing) 和多载波调制是短波协议使用较多的调制模式。综上, 9 种信号的调制模式包含多种短波信号常用的调制模式, 在短波协议中具有一定代表性。因此, 本文选择这 9 种短波协议信号制作数据集, 通过对这 9 种信号的识别, 可以说明本文算法对多种类型的短波协议信号识别具有一定的普适性。此外, 为了增加数据集的可靠性和多样性, 对仿真信号及数据集做以下处理。

1) 短波信道环境恶劣, 信号受噪声影响严重。本文采用高斯白噪声, 且 SNR 范围为  $-10 \sim 10$  dB。

2) 在非合作通信场景下, 接收方无法确保接收信号的完整性, 同时考虑到实时识别的要求, 信号的长度不宜过长, 因此对仿真信号随机截取时长为  $0.5 \sim 5$  s 的子序列。

3) 考虑到闭集识别的局限性, 本文使用 9 种信号之外的部分短波信号以及调制方式为 PSK、FSK 的普通信号组成单独的噪声类, 来提高算法对数据集中 9 种信号之外的其他部分短波信号的区分能力。

基于以上 3 种处理措施, 训练集样本由每类信号产生 500 个灰度时频图样本得到, 每个样本大小压缩为  $224$  像素  $\times$   $224$  像素。数据集中 9 种短波协议信号的时频图如图 9 所示。从图 9 中可以看出, 由于不同信号在帧结构、带宽、中心频率及调制模式等方面存在差异, 各信号在时频图上呈现出不同的视觉特性。

本文仿真实验环境的硬件与软件配置信息如表 2 所示。训练过程的学习率设为 0.000 1, batch size 设置为 32, epoch 最大设为 20, 采用 AdamW 优化器。

表 2 仿真实验环境配置信息

硬件与软件	配置信息
GPU	NVIDIA GeForce RTX 2080Ti
CPU	Intel(R) Xeon(R) CPU E5-2678 v3 @ 2.50 GHz
内存	64 GB
运行系统	Ubuntu 18.04 64 bit
Python	3.8
CUDA	11.0
PyTorch	1.7.1

图 10 给出了所提算法在高斯信道下对 9 种信号的识别率, 每种信号在各信噪比下的样本数为 500。从图 10 可以看出, 低信噪比下各信号识别率差异较大, 但当  $\text{SNR} > -4$  dB 时, 9 种信号的识别

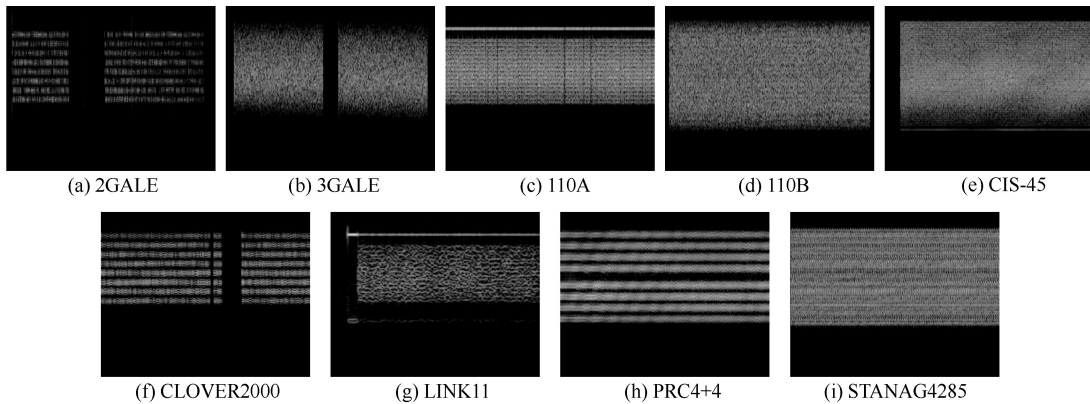


图 9 数据集中的 9 种短波协议信号时频图

率均已接近 100%，说明所提算法能够对 9 种信号进行有效识别。

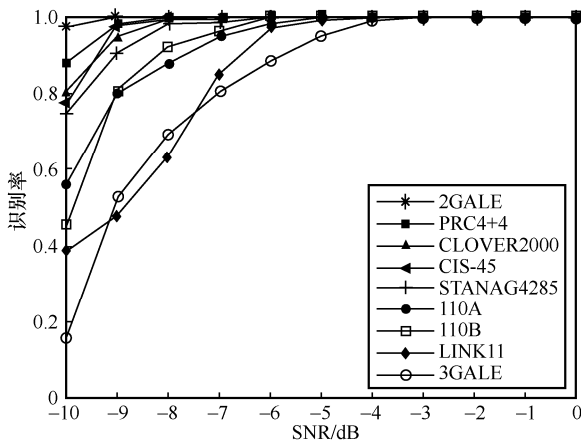


图 10 所提算法在高斯信道下对 9 种信号的识别率

理论上，基于 Swin-Transformer 和基于 ResNet 的神经网络算法均能够通过增加网络深度实现性能的有限提升，但网络层数的增加会导致算法计算量增加以及过拟合风险加大，因此本文选择了与所提算法计算量接近的 ResNet50 和 ResNeXt50<sup>[27]</sup> 这 2 种 CNN 算法。此外，在非神经网络算法中，时域模板匹配算法作为一种通用的短波协议信号识别算法，事先对待识别协议信号建立模板库，通过计算待识别信号与模板信号的相关系数来确定信号所属类别。图 11 给出了高斯信道下 4 种算法的性能对比。

由图 11 可知，神经网络算法凭借强大的特征提取能力，其对短波协议信号的识别性能要优于传统的时域模板匹配算法，且神经网络算法在识别过程中不需要将待识别信号与信号库一一对比，识别流程相对更直接。此外，由于 Swin-Transformer 引

入了全局感知能力更强的自注意力机制，相比于传统的 CNN 具有更大的感受野，而短波协议间的差异在时频图呈现出的视觉特性差异同样是全局性的，因此 Swin-Transformer 算法在短波协议识别上比感受野相对有限的 CNN 性能更好。

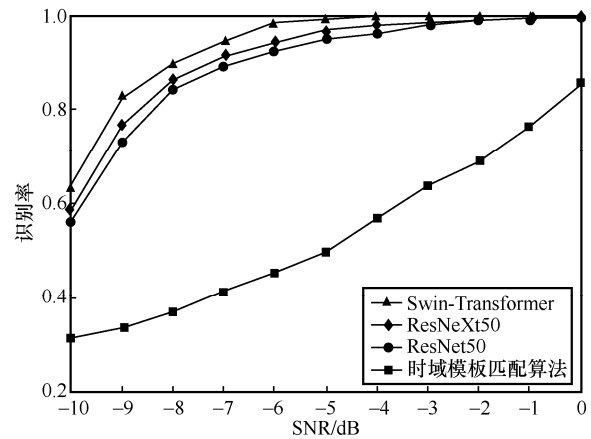


图 11 高斯信道下 4 种算法的性能对比

另外，短波频段频谱资源紧张，实际接收到信号可能存在叠加干扰，因此需验证所提算法对存在干扰的信号识别性能。本文使用的干扰信号调制方式为 8PSK，带宽为  $f_1$ ，被混叠信号带宽为  $f_2$ ，两者重叠带宽为  $f_{12}$ ，2 个信号在时间上完全重叠，文献[16]对频域上干扰混叠度  $D$  的定义为

$$D = \frac{f_{12}}{f_1 + f_2 - f_{12}} \quad (11)$$

在存在叠加干扰的情形下，此时用信号干扰比 (SIR, signal to interfere ratio) 表示信号与干扰功率之比，单位为 dB。图 12 给出了干扰混叠度及干扰信号功率对所提算法识别性能的影响。从图 12 中

可以看出, 在干扰信号功率较大 ( $SIR=0$ ) 且时频混叠程度较大的情况下 ( $D=0.5$  或  $D=0.7$ ), 本文所提算法性能受影响较大, 因为此时信号在时频图上呈现的视觉特征大部分被干扰信号所遮挡, 所提算法无法提取出有效特征进行识别。但当  $SIR>10$  dB, 即信号功率明显大于干扰功率时, 本文所提算法对 4 种混叠程度下的短波协议信号均有超过 90% 的识别率, 说明所提算法具有较好的抗干扰能力。

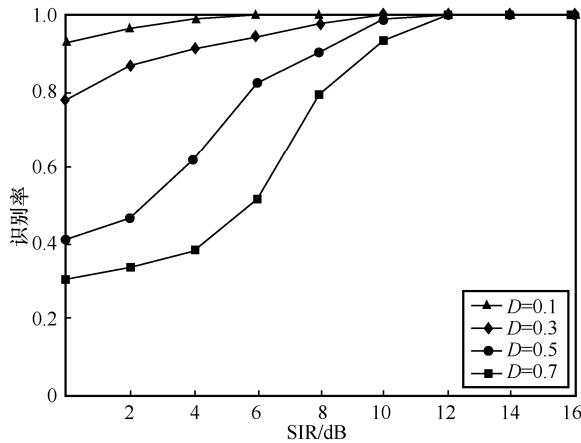


图 12 叠加干扰信号下所提算法识别率

此外, Watterson 信道模型是一种经典的短波信道模型, 本文采用文献[28]建议的信道参数。图 13 给出了所提算法在 Watterson 信道下算法的识别性能。由图 13 可知, 在 Watterson 信道下, 当  $SNR>13$  dB 时, 本文所提算法在 8 种信道条件下的识别率均接近 100%。

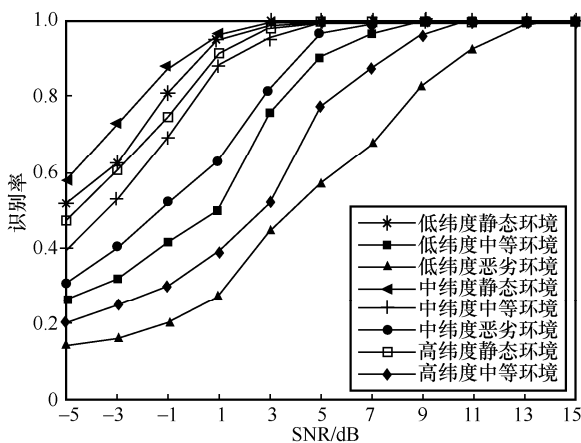


图 13 Watterson 信道下所提算法识别率

## 4 结束语

本文针对传统短波协议信号识别算法低信噪

比下识别率低、识别信号单一以及对先验信息需求高等问题, 研究了短波协议信号在时频图上视觉特征的原因, 提出一种基于 Swin-Transformer 的神经网络短波协议信号识别算法, 通过对信号的灰度时频图进行特征提取并映射实现对信号所属协议的识别。实验结果表明, 基于 Swin-Transformer 的短波协议信号识别算法在高斯白噪声信道、存在强混叠信号和 Watterson 信道环境下均具有较高的识别率。

## 参考文献:

- [1] 张有光, 江琥, 熊鹰. 通用短波数据传输协议发展状况综述[J]. 电讯技术, 2002, 42(2): 140-144.  
ZHANG Y G, JIANG H, XIONG Y. Introduction to and comparison of commercial HF digital protocols[J]. Telecommunication Engineering, 2002, 42(2): 140-144.
- [2] GIESBRECHT J E, CLARKE R, ABBOTT D. Modulation recognition for real HF signals[C]//Proceedings of Microelectronics: Design, Technology, and Packaging II. Bellingham: SPIE Press, 2006: 458-469.
- [3] 聂东举, 叶进, 闫坤, 等. 基于 SVM 算法的短波通信协议识别技术[J]. 系统工程与电子技术, 2013, 35(6): 1307-1311.  
NIE D J, YE J, YAN K, et al. Recognition technology for high frequency communication protocol based on SVM algorithm[J]. Systems Engineering and Electronics, 2013, 35(6): 1307-1311.
- [4] DEARLOVE C M, WRIGHT P G. Rapid HF waveform identification[C]//Proceedings of Frequency Selection and Management Techniques for HF Communications. London: IET, 1999: 1-5.
- [5] 林肖辉, 张润生. 基于模板匹配的 2G-ALE 信号识别技术[J]. 无线电通信技术, 2016, 42(3): 46-48, 69.  
LIN X H, ZHANG R S. Technology of 2G-ALE signal recognition based on spectrum template matching[J]. Radio Communications Technology, 2016, 42(3): 46-48, 69.
- [6] 谷双春, 施昉, 张润生. 短波 110A 信号检测分析方法研究[J]. 无线电通信技术, 2018, 44(3): 277-281.  
GU S C, SHI F, ZHANG R S. Study on HF 110A signal detection and analysis method[J]. Radio Communications Technology, 2018, 44(3): 277-281.
- [7] US Department of Defense. Interoperability and performance standards for data modems: MIL-STD-188-110A[S]. 1991.
- [8] North Atlantic Treaty Organization. Characteristics of 1200/2400/3600 bits per second single tone modulators/ demodulators for HF radio links: STANAG 4285[S]. 1989.
- [9] US Department of Defense. Interoperability and performance standards for medium and high frequency radio systems: MIL-STD-188-141A[S]. 1988.
- [10] 安泽亮, 张天骐, 马宝泽, 等. 基于一维 CNN 的多入多出 OSTBC 信号协作调制识别[J]. 通信学报, 2021, 42(7): 84-94.  
AN Z L, ZHANG T Q, MA B Z, et al. Cooperative modulation recognition based on one-dimensional convolutional neural network for MIMO-OSTBC signal[J]. Journal on Communications, 2021, 42(7): 84-94.
- [11] LIN S G, ZENG Y, GONG Y. Learning of time-frequency attention mechanism for automatic modulation recognition[J]. IEEE Wireless

- Communications Letters, 2022, 11(4): 707-711.
- [12] HUANG Z H, LI S, YANG X H, et al. OAE-EEKNN: an accurate and efficient automatic modulation recognition method for underwater acoustic signals[J]. IEEE Signal Processing Letters, 2022, 29: 518-522.
- [13] WAN L, WANG Q, PAPIR A, et al. Generalized end-to-end loss for speaker verification[C]//Proceedings of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2018: 4879-4883.
- [14] XU Z J, SAVVARIS A, TSOURDOS A, et al. High-frequency band automatic mode recognition using deep learning[C]//Proceedings of 2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC). Piscataway: IEEE Press, 2018: 1-5.
- [15] 杨司韩, 彭华, 许漫坤, 等. 基于卷积神经网络的超短波特定信号谱图识别[J]. 系统工程与电子技术, 2019, 41(4): 744-751.  
YANG S H, PENG H, XU M K, et al. Ultra-short wave specific signal spectrogram recognition based on convolution neural network[J]. Systems Engineering and Electronics, 2019, 41(4): 744-751.
- [16] 查雄, 许漫坤, 彭华, 等. 基于深度残差网络的特定协议信号识别[J]. 电子学报, 2019, 47(7): 1532-1537.  
ZHA X, XU M K, PENG H, et al. Specific protocol signal recognition based on deep residual network[J]. Acta Electronica Sinica, 2019, 47(7): 1532-1537.
- [17] 查雄. 基于神经网络的数字调制信号检测、识别与解调算法[D]. 郑州: 信息工程大学, 2020.  
ZHA X. Research on digital modulation signal detection, recognition and demodulation based on neural network[D]. Zhengzhou: Information Engineering University, 2020.
- [18] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 770-778.
- [19] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2021: 9992-10002.
- [20] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017). Massachusetts: MIT Press, 2017: 5998-6008.
- [21] HAL Communications Corporation. CLOVER2000 waveform & protocol[S]. 1999.
- [22] US Department of Defense. Interoperability and performance standards for tactical digital information link(TADIL)A: MIL-STD-118-203-1A[S]. 2013.
- [23] XIAO Z L, YAN Z Y. Radar emitter identification based on novel time-frequency spectrum and convolutional neural network[J]. IEEE Communications Letters, 2021, 25(8): 2634-2638.
- [24] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth  $16 \times 16$  words: transformers for image recognition at scale[C]//International Conference on Learning Representations (ICLR). [S.l.:s.n.], 2021: 1-21.
- [25] US Department of Defense. Interoperability and performance standards for data modems: MIL-STD-188-110B[S]. 2000.
- [26] US Department of Defense. Interoperability and performance standards for medium and high frequency radio systems: MIL-STD-188-141B[S]. 1989.
- [27] XIE S N, GIRSHICK R, DOLLÁR P, et al. Aggregated residual transformations for deep neural networks[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 5987-5995.
- [28] WATTERSON C, JUROSHEK J, BENSEMA W. Experimental confirmation of an HF channel model[J]. IEEE Transactions on Communication Technology, 1970, 18(6): 792-803.

### [作者简介]



朱政宇 (1988- ), 男, 河南周口人, 博士, 郑州大学副教授、硕士生导师, 主要研究方向为无线通信与信号处理、智能反射表面技术、物理层安全技术等。

陈鹏飞 (1998- ), 男, 河南南阳人, 郑州大学硕士生, 主要研究方向为短波通信、智能信号处理。

王梓暄 (1998- ), 男, 河南周口人, 郑州大学博士生, 主要研究方向为通信信号处理、多源信息融合等。

巩克现 (1976- ), 男, 山东泰安人, 博士, 郑州大学教授、博士生导师, 主要研究方向为无线通信信号分析与处理、信道编码、无线接入、目标监测及电子对抗等。

吴迪 (1984- ), 男, 福建建阳人, 博士, 信息工程大学讲师, 主要研究方向为通信信号分析与智能处理、电子对抗等。

王忠勇 (1965- ), 男, 江西遂川人, 博士, 郑州大学教授、博士生导师, 主要研究方向为通信信号处理等。