

基于脉冲神经网络的轻量化SAR图像 舰船识别算法

谢洪途¹, 陈佳兴¹, 张琳², 朱楠楠³

(1. 中山大学·深圳 电子与通信工程学院, 广东 深圳 518107; 2. 空军预警学院, 湖北 武汉 430019;

3. 中山大学 系统科学与工程学院, 广东 广州 510275)

摘要: 针对传统方法进行合成孔径雷达(SAR)图像目标识别存在参数多、能耗高等问题,提出了一种基于脉冲神经网络(SNN)的轻量化SAR图像舰船识别算法。首先,利用视觉注意力机制提取SAR图像视觉显著图,采用泊松编码器进行脉冲序列编码,能抑制背景噪声干扰。然后,结合泄漏整合发射(LIF)脉冲神经元和卷积神经网络,构建融合时序信息的SNN模型,能实现SAR图像舰船识别。最后,采用反正切函数作为反向传播时脉冲发射函数的梯度替代函数对SNN模型进行优化,能解决模型难以训练的问题。实验结果表明所提算法具有高精度、少参数、高效率 and 低能耗等优势,能实现SAR图像高效准确舰船识别。

关键词: 合成孔径雷达图像;舰船识别;脉冲神经网络;轻量化

中图分类号: TN 957.52; TP 391.4 文献标志码: A 文章编号: 1005-3026(2024)04-0474-09

Lightweight Ship Recognition Algorithm Based on SNN in SAR Images

XIE Hong-tu¹, CHEN Jia-xing¹, ZHANG Lin², ZHU Nan-nan³

(1. School of Electronics and Communication Engineering, Shenzhen Campus of Sun Yat-sen University, Shenzhen 518107, China; 2. Air Force Early Warning Academy, Wuhan 430019, China; 3. School of Systems Science and Engineering, Sun Yat-sen University, Guangzhou 510275, China. Corresponding author: ZHU Nan-nan, E-mail: zhunn25@mail.sysu.edu.cn)

Abstract: Due to the more parameters and higher energy-consumption in the traditional methods for the synthetic aperture radar (SAR) image target recognition, this paper proposes a lightweight ship recognition algorithm based on the spiking neural network (SNN) in SAR images. Firstly, the visual attention mechanism is adopted to extract the visual saliency map from SAR images, and the Poisson encoder is adopted for the spike train encode, which can suppress the background noise interference. Then, combined with the leaky integrate-and-fire (LIF) spiking neuron and convolutional neural network, the SNN model integrating the time series information is constructed, which can realize the ship recognition in SAR images. Finally, the SNN model is optimized by using the arctangent function as the surrogate gradient function of the spiking emission function during the backpropagation, which can solve the problem that the SNN model is difficult to train. The experiment results show that the proposed algorithm has higher accuracy, fewer parameters, higher efficiency, and lower energy-consumption, which can achieve efficient and accurate ship recognition in SAR images.

Key words: synthetic aperture radar (SAR) image; ship recognition; spiking neural network (SNN); lightweight

收稿日期: 2022-12-09

基金项目: 广东省基础与应用基础研究基金资助项目(2023A1515011588, 2021A1515010768); 深圳市科技计划资助项目(202206193000001, 20220815171723002); 国家自然科学基金资助项目(62203465, 62001523, 62201614, 6210593)。

作者简介: 谢洪途(1986-),男,湖南邵阳人,中山大学副教授。

近年来,利用合成孔径雷达(synthetic aperture radar, SAR)^[1-2]图像进行舰船检测识别的研究在海洋遥感领域得到高度重视^[3-4].然而,传统SAR图像识别方法主要依赖人工设计特征,易受到复杂背景干扰,存在识别精度差、效率低、泛化能力弱等缺点.以人工神经网络(artificial neural network, ANN)为代表的深度学习方法能自主学习图像特征,具有自动化程度高、识别能力强、抗干扰强等优点,在SAR图像目标识别任务中取得较大发展^[5].

随着深度学习方法的发展,ANN模型越来越复杂,导致模型训练中计算成本和能量消耗越来越大.2018年发布的自然语言处理模型BERT(bidirectional encoder representation from transformers)达到3亿的参数量^[6],2020年发布的GPT-3模型参数量为1750亿^[7],2021年的语言模型Switch Transformer更达到1.6万亿参数量^[8].因此,效率和能耗问题极大限制了ANN在轻量化图像目标识别中的应用.脉冲神经网络(spiking neural network, SNN)模型能够更精确地模仿人类大脑的信息表达及处理过程,是一种具有高生物可塑性、高事件驱动性和低功耗的类脑计算模型^[9].Hong等^[10]通过在传统的残

差单元之间加入SNN神经元层,提出了基于脉冲神经元的残差网络以解决图像分类问题.Kim等^[11]提出了一种基于SNN的检测模型Spiking-YOLO,获得了比ANN模型更好的检测效果和更低的能耗.

为解决传统ANN进行SAR图像目标识别存在参数较多、能耗较高等问题,本文提出了基于SNN的轻量化SAR图像舰船识别算法,即SpikingSAR.

1 SAR图像舰船识别

基于SNN的轻量化SAR图像舰船识别算法的模型框架见图1.首先利用视觉注意力机制提取SAR图像视觉显著图,采用泊松编码器进行脉冲序列编码.然后,结合泄漏整合发射(leaky integrate-and-fire, LIF)脉冲神经元和卷积神经网络(convolutional neural network, CNN),构建融合时序信息的SNN模型,利用脉冲神经元发射频率实现SAR图像舰船识别.最后,采用反正切函数作为反向传播时脉冲发射函数的梯度替代函数,解决SNN模型难以训练的问题.

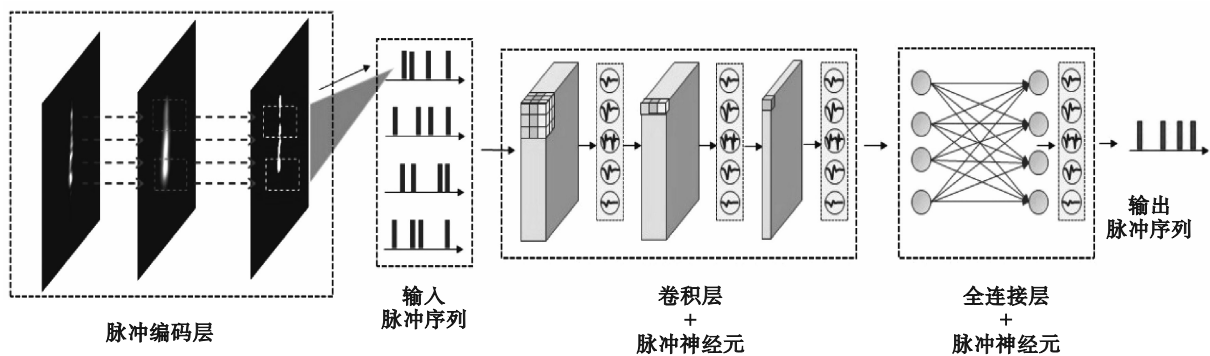


图1 基于SNN的轻量化SAR图像舰船识别算法的模型框架

Fig. 1 Model framework of the lightweight ship recognition algorithm based on SNN in SAR image

1.1 基于视觉注意力机制的脉冲编码器

SAR图像为连续的值,难以表示离散的时间相关脉冲序列,因此需对输入图像进行脉冲编码.基于速率的编码是将输入图像转换为脉冲在各时间步长上的映射.给定仿真时长,根据合适分布从每个像素采样一个泊松脉冲序列,其脉冲发射概率与像素灰度强度成正比.该编码方式既能保留输入图像的完整性,又能对数据进行二值化编码^[8].

SAR图像像素值分布与光学图像存在较大差异,直接进行基于速率的编码难以获得有效脉

冲序列.由于海洋环境及SAR图像存在各类噪声,需采用预处理对大量干扰信息进行甄别和剔除,才能精确地获得SAR图像中的舰船目标.传统方法一般对SAR图像进行滤波以降低噪声,然而滤波会导致图像边缘变宽,损失边缘定位信息,易造成较小舰船漏检.考虑所关注的舰船仅占据SAR图像小部分区域,若能预先选取SAR图像中较关注的区域,并根据各区域重要性将其重新编码,有利于提升识别模型的准确率,加速模型的收敛速度.为此,本文提出基于视觉注意力机制^[12]的脉冲编码器,优先聚焦在SAR图像中

视觉特征明显的显著性区域,忽略背景杂波噪声,从而缩小舰船识别范围.

1.1.1 视觉注意力机制

本文将视觉注意力机制应用于 SAR 图像舰船识别,重点关注 SAR 图像中亮度信息和方向信息.本文采用的视觉注意力机制模型结构如图 2 所示.

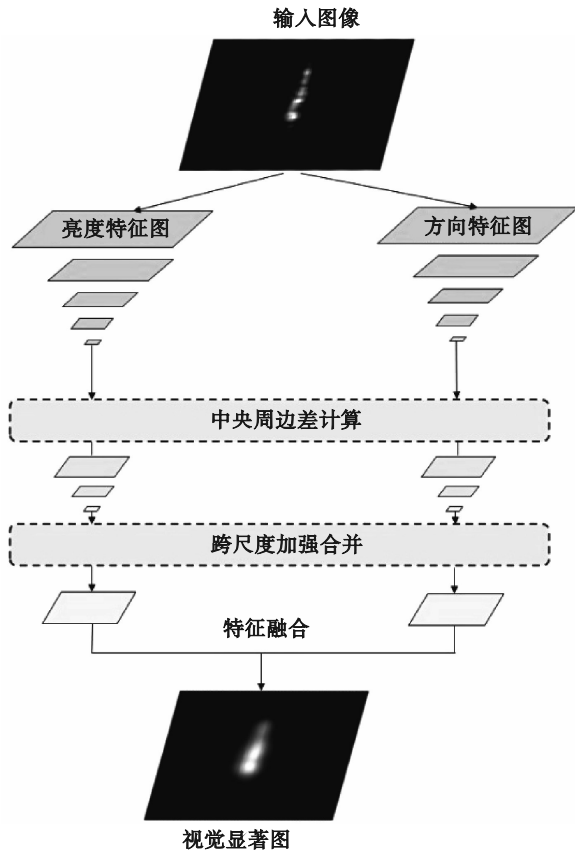


图2 视觉注意力机制模型结构

Fig.2 Model architecture of the visual attention mechanism

1) 亮度信息提取.给定输入图像 J ,利用高斯金字塔进行分辨率为2的亮度特征图提取,实现 SAR 图像在水平和垂直向尺度递减的八级下采样.每层图像的亮度特征图为 $I(k)$, $k \in [0, 8]$ 代表图像高斯金字塔结构中不同层级.图像中相邻像素间在纹理和灰度上有很强相关性,某个像素点与周围像素点差异越大,越易引起视觉注意,成为视觉显著点.因此,引入中央周边差法对不同尺度分辨率的特征图进行处理获得注意力信息. \ominus 为中央周边差算子,亮度特征图的提取是将不同层级的特征图缩放到同一尺度后进行逐像素点相减.

$$I(c, s) = |I(c) \ominus I(s)|. \quad (1)$$

其中: $c \in \{2, 3, 4\}$; $s = c + \delta$; $\delta \in \{3, 4\}$.

2) 方向信息提取.与亮度信息提取类似,利用高斯金字塔对输入图像进行八级下采样得到多尺度特征图,再进行方向特征提取.本文采用二维 Gabor 滤波器提取图像方向通道特征信息

$$G(x, y) = \frac{1}{2\pi\sigma\beta} \exp \left\{ -\pi \left[\frac{(x-x_0)^2}{\sigma^2} + \frac{(y-y_0)^2}{\beta^2} \right] \right\} \times \exp \{ i[\zeta_0 x + \nu_0 y] \}. \quad (2)$$

其中: (x_0, y_0) 为目标中心在图像的坐标位置; (ζ_0, ν_0) 为滤波器在频域的最优空间频率; σ^2 和 β^2 分别为高斯函数在 x 轴和 y 轴的方差.通过方向特征提取,获得不同尺度的方向特征图 $O(k)$, $k \in [0, 8]$ 代表图像高斯金字塔结构中的不同层级.进一步,可以利用中央周边差法提取方向特征的注意力信息

$$O(c, s, \theta) = |O(c, \theta) \ominus O(s, \theta)|. \quad (3)$$

其中: O 为 Gabor 滤波器的 4 个方向; $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$.

3) 全局加强合并.由于 SAR 图像目标的识别通常缺乏自上而下的先验信息,且视觉注意力机制为无监督方法,因此本文采用全局加强合并策略融合亮度和方向特征图的注意力信息,以得到视觉显著图.包含三个步骤:1) 将特征图归一化到 $[0, N]$, $N \in [0, 255]$ 限制了归一化范围;2) 计算全局最大值 M 和局部平均值 \bar{m} ;3) 对每1幅特征图,乘以权值 $(M - \bar{m})^2$.假设 N 表示全局加强合并,则融合了中央周边差算子的亮度通道显著图的合并过程为

$$I = \sum_{c=2}^4 \sum_{s=c+3}^{c+4} N(I(c, s)). \quad (4)$$

类似地,方向通道显著图的合并过程为

$$O = \sum_{\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} N \left(\sum_{c=2}^4 \sum_{s=c+3}^{c+4} N(O(c, s, \theta)) \right). \quad (5)$$

最后,利用全局加强合并策略得到视觉显著图 S ,

$$S = \frac{1}{2} (N(I) + N(O)). \quad (6)$$

通过融合,若所得特征图像之间差别较大,其显著性也被放大,特征则在最终融合图像中被突出;同时能减少微弱差别所带来的干扰,图像中需注意的目标被凸显,贡献较小的特征被抑制.由于该策略采用了归一化操作,最终得到的视觉显著图 $S \in [0, 1]$ 符合本文脉冲编码的前提条件.此外,由于视觉特征较显著的区域灰度值较大,该区域对应编码后的脉冲序列有更大的概率发射脉冲信息.

1.1.2 泊松编码器

SNN 模型中,编码器主要将输入的连续值信号转换为带有时间信息的离散脉冲序列,并尽可能保留数据的大部分信息.基于速率的泊松编码器,对于输入 $x \in [0, 1]$,在一个仿真时长 T 内,设脉冲发射的概率 $p=x$,则泊松编码器的采样过程为

$$\hat{x} = \begin{cases} 1, & x \geq p, \\ 0, & \text{其他}, \end{cases} \quad \text{s.t. } p \in U(0,1). \quad (7)$$

其中 $U(0,1)$ 为均匀分布.如图 3 所示, t 表示仿真时间,每个像素在仿真时长 T 内产生 n 个脉冲序列,则脉冲发射速率 $v=n/T$.此外,脉冲的发射概率与 SAR 图像的像素值的大小成正比.

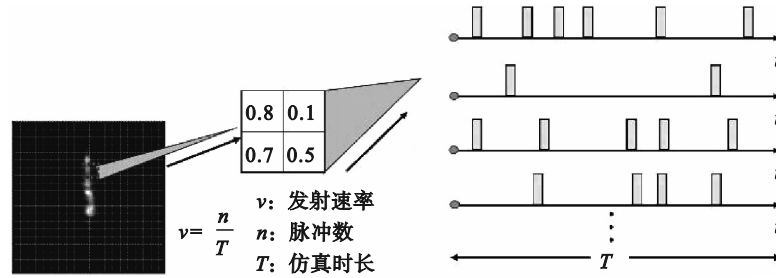


图 3 泊松编码器的编码过程

Fig. 3 Encoding process of the Poisson encoder

1.2 脉冲神经网络模型

1.2.1 脉冲神经元状态方程

SNN 的输入是当前时刻的电压增量,表示对神经元电位的充电过程;输出值为 0 或 1 的离散值,表示脉冲发射与否,为神经元电位的放电过程.神经元在放电后短暂时间内存在不应期,不响应外界输入信号.为模拟该过程,在神经元放电后额外引入重置过程,将电压重置使神经元处于静息状态.因此,典型的 SNN 神经元可以用充电、放电和重置三个过程描述,对应离散方程为

$$\left. \begin{aligned} H(t-1) &= f(V(t-1), X(t)), \\ S(t) &= \psi(H(t) - V_{th}), \\ V(t) &= g(H(t), S(t), V_{reset}). \end{aligned} \right\} \quad (8)$$

其中: $V(t)$ 是神经元在 t 时刻的膜电压; $X(t)$ 是 t 时刻的输入,通常为来自突触前神经元的脉冲信号; $H(t)$ 是神经元在 t 时刻的隐藏状态,通常为充电后的瞬时电压; f 是神经元充电时刻的状态更新函数; $S(t)$ 是神经元的脉冲发射状态; V_{th} 为脉冲发射的阈值电压; ψ 为脉冲发射的门控函数; g 是脉冲发射后的重置函数; V_{reset} 为神经元的静息电压或平衡电压.

1.2.2 LIF 脉冲神经元模型

LIF 脉冲神经元模型^[13] 是对整合发射 (integrate-and-fire, IF) 脉冲神经元模型的改进. LIF 模型将细胞膜的电特性可以等效为电阻和电容并联,相较于 IF 模型具有更高生物可信度和生

物精确性,能准确描述神经元的状态变化.对于 LIF 神经元模型,用微分方程进行描述:

$$C \frac{dV(t)}{dt} = -\frac{1}{R}(V(t-1) - V_{reset}) + I(t). \quad (9)$$

其中: R 为膜电阻; C 为电容量; $I(t)$ 为 t 时刻的输入电流.为保证 SNN 模型计算的适用性,通常采用离散形式来近似微分表达式,即

$$\tau(V(t) - V(t-1)) = -(V(t-1) - V_{reset}) + X(t). \quad (10)$$

其中, $\tau = RC$ 为膜时间常数. LIF 模型的充电过程为

$$f(V(t-1), X(t)) = V(t-1) + \frac{1}{\tau}(-V(t-1) - V_{reset}) + X(t). \quad (11)$$

其中, $X(t) = I(t)R$ 为 t 时刻的外源输入. LIF 模型的放电过程与 IF 模型放电类似.当电荷累积到一定程度,即膜电压达到阈值 V_{th} , 神经元发射脉冲,即为放电过程.脉冲发射可以采用 Heaviside Step 阶跃函数进行描述:

$$\Psi(x) = \begin{cases} 1, & x \geq 0, \\ 0, & \text{其他}. \end{cases} \quad (12)$$

LIF 模型的重置发生在脉冲发射后,发射脉冲消耗神经元之前积累的电荷,则膜电位会瞬间降低.考虑到膜电位未达到阈值时,神经元不发射脉冲,电位保持不变,因此 LIF 模型的脉冲重置过程为

$$g(H(t), S(t), V_{reset}) = H(t)(1 - S(t) + V_{reset}S(t)). \quad (13)$$

1.2.3 主干网络设计

本文将 ANN 与脉冲神经元模型相结合,作

为 SAR 图像目标识别的主干网络,见图 4.利用 ANN 对神经元 j 的输入电压进行模拟,其过程可表示为

$$X^l(t) = \sum_i w'_{ij} s_i^{l-1} + b'_j. \quad (14)$$

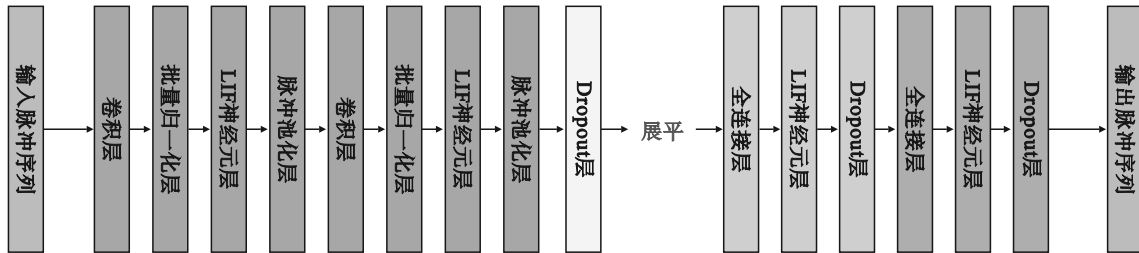


图 4 主干网络结构图

Fig. 4 Structure diagram of the backbone network

ANN 包含 CNN 和全连接网络 (fully connected network, FCN), 两者都能较好地模拟 SNN 输入电压. 鉴于输入的脉冲序列由 SAR 图像的视觉显著图编码而得, 使用 CNN 更能捕获输入的局部空间信息, 且兼具参数较少的优势, 再在 CNN 末端引入全连接层, 可整合卷积层提取的特征信息, 对图像目标进行识别. 因此, 本文采用 CNN, FCN 分别与 LIF 脉冲神经元交替堆叠作为 SNN 模型的主干网络.

ANN 通常与批量归一化层、Dropout 层、池化层和激活函数层结合使用. 批量归一化层能在尽可能不丢失重要信息的同时对每层输入数据进行规范化处理, 使得数据分布相对稳定, 加速模型学习速度. Dropout 层能以一定概率随机丢弃神经元, 减少模型隐层节点间相互作用, 避免模型过拟合. 池化层的作用是进行特征图下采样, 逐步减少特征空间大小, 避免网络的高计算复杂度, 增强网络的空间不变性. Hong 等^[10]研究表明最大池化与 SNN 的时间信息处理能力一致, 可提高 SNN 对时序数据的拟合能力, 同时降低下一层网络的计算成本. 因此, 本文采用脉冲最大池化作为网络的池化层. 最大池化层能保持脉冲序列的离散性, 且可使用逻辑运算代替矩阵乘法运算, 从而提高 SNN 的运算速度和减少能耗. ANN 最重要特性是可任意逼近非线性函数, 主要是得益于非线性激活函数的存在. 由于 Heaviside Step 脉冲发射函数的存在, 引入脉冲神经元的 SNN 本身就具备非线性的特点, 因此可直接利用脉冲神经元替代激活层, 激活函数则替代为脉冲发射函数 $\psi(x)$, 从而简化模型运算复杂度.

1.2.4 模型训练

SNN 的脉冲发射函数 $\psi(x)$ 通常是不可微分

其中: s_i^{l-1} 为第 $l-1$ 层输入的脉冲序列; w'_{ij} 和 b'_j 分别为第 l 层的权重和偏置; $X^l(t)$ 为第 l 层的输入电压.

的, 导致 SNN 不能使用梯度下降和反向传播算法优化训练. 本文提出利用替代梯度进行训练的方法, 即利用一个形式与 $\psi(x)$ 十分相似, 但可微分的门控函数 $\sigma(x)$ 逼近脉冲函数, 再计算近似的替代梯度对 SNN 进行更新. 该方法能对 SNN 进行有效训练, 训练后的模型在诸多任务方面的表现可与 ANN 相比拟^[10]. 其核心思想: 1) 在前向传播时, 仍然使用脉冲发射函数 $\psi(x)$ 作为门控函数, 激活脉冲神经元并在电压达到阈值时发射脉冲; 2) 在反向传播时, 使用替代函数 $\sigma(x)$ 计算反向传播时的近似梯度, 并利用替代函数梯度对网络进行优化更新. 常见的替代函数是 Sigmoid 函数 $\sigma(\alpha x) = (1 + \exp(-\alpha x))^{-1}$, α 越大会使得替代函数越平滑, 从而更逼近脉冲发射函数 $\psi(x)$, 但越容易在靠近原点时引起梯度爆炸, 远离原点时引起梯度消失, 导致网络变得越难以训练. 反向传播时, 计算 Sigmoid 函数的梯度为 $\nabla \sigma(\alpha x) = \alpha \cdot \sigma(\alpha x) \cdot (1 - \sigma(\alpha x))$.

然而, Sigmoid 函数难以拟合 Heaviside Step 脉冲发射函数, 并且当 α 较大时 Sigmoid 函数容易在两端出现饱和现象, 从而造成梯度消失, 增加了网络训练的难度. 为解决上述问题, 本文提出使用平滑化反正切函数 (arctan) 作为模型训练时的梯度替代函数. 在前向传播时, 平滑化 arctan 函数为

$$\sigma(x) = \frac{1}{\pi} \arctan\left(\frac{\pi}{2} \alpha x\right) + \frac{1}{2}; \quad (15)$$

在反向传播时, 平滑化 arctan 函数的梯度为

$$\nabla \sigma(x) = \frac{\alpha}{2(1 + (\pi \alpha x / 2)^2)}. \quad (16)$$

为利用反向传播算法训练 SNN, 需要设计一个合适的损失函数. 假设 SAR 图像的类别数

为 K ,输出的脉冲序列为 $O=[o_{i,t}] \in \mathbf{R}^{K \times T}$,对应的类别 $Y=[y_{i,t}] \in \mathbf{R}^{K \times T}$.如果1张SAR图像的真实标签为 l ,那么最终输出表示类别 l 的神经元应当具有最高的兴奋水平(即发射最多脉冲),其余神经元应当保持抑制状态(即发射较少脉冲).因此,本文采用均方差误差函数作为模型训练的损失函数,即

$$L(O, Y) = \frac{1}{T} \sum_{t=0}^{T-1} \frac{1}{K} \sum_{i=0}^{K-1} (o_{i,t} - y_{i,t})^2. \quad (17)$$

其中, $y_{i,t}$ 当且仅当 $i=l$ 时为1,其余为0.当模型训练收敛后,对输入图像的识别预测类别 l_p 为发射脉冲数最多的神经元所代表的类别,即

$$l_p = \arg \max_i \frac{1}{T} \sum_{t=0}^{T-1} o_{i,t}. \quad (18)$$

2 实验与分析

2.1 实验设置

为了验证所提算法的正确性和适用性,本文使用SAR舰船数据集SAR-Dataset^[14]进行实验.该数据集是2017年发布用于舰船识别和检测的数据集,由22张哨兵一号和3张RADARSAT-2卫星采集的图像构建,包含42幅双极化和4幅单极化辐射校准图像.数据集涵盖南非专属经济区约80%区域,包含多个高船舶密度港口场景,有三类图像样本,分别为舰船图像(正样本)、类船图像(负样本)和海洋背景图像(负样本).每类样本数量相同,都包含1596张75像素×75像素的SAR图像.该数据集规模较小,但样本分布比较均衡,有利于模型训练和评估.该数据集的75%为训练集,25%为测试集,并从训练集中划分20%作为验证集.

为验证所提算法的有效性,本文采用传统识别方法与深度学习识别方法进行对比.传统识别方法主要包括逻辑回归(logistic regression, LR)、K-近邻(K-nearest neighbors, KNN)、支持向量机(support vector machine, SVM)和决策树(decision tree, DT).深度学习识别方法主要包括本文所提SpikingSAR, AlexNet, GoogleNet, MobileNet, DenseNet和ResNet, ResNet分别采用了ResNet-18, ResNet-34和ResNet-50不同层数模型.对于传统方法,使用主成分分析法将输入SAR图像进行降维,提取图像特征作为输入以预测类别,设置特征维度为80.对于神经网络模型,使用Adam作为优化器进行训练,训练批次为32,学习率为0.001,权重衰减为0.0001,采用交

叉熵损失函数,共训练100次,根据验证集上的结果保存模型的最优权重.对于SpikingSAR模型,膜时间常数 τ 为2s,仿真时长 T 为12s,损失函数为均方差函数,其余设置与对比方法中神经网络模型一致.采用CPU为Intel i9-9900X, GPU为NVIDIA RTX 2080Ti,操作系统为Ubuntu18.04的计算机.传统方法使用Scikit-learn框架实现,而深度学习方法使用PyTorch框架实现,且利用CUDA10.1实现训练和测试过程中的加速.

本文采用准确度(Precision),召回率(Recall)和F1分数(F1-score)作为评估指标,F1分数由准确度和召回率定义.针对目标类别 i 的识别结果为

$$P_i = \frac{TP}{TP + FP}, \quad (19)$$

$$R_i = \frac{TP}{TP + FN}, \quad (20)$$

$$F1_i = 2 \times \frac{P_i \times R_i}{P_i + R_i}. \quad (21)$$

其中: P 表示准确度; R 表示召回率; TP 表示正确识别舰船的数量; FN 表示未识别或未被发现的舰船数量; FP 表示错误识别的舰船数量.考虑数据集中存在各类目标数量可能不均衡问题,为更准确地评估模型的性能,本文使用各个类别目标识别结果的加权平均作为最终的性能指标:

$$P = \sum_i \frac{N_i}{N} P_i, \quad (22)$$

$$R = \sum_i \frac{N_i}{N} R_i, \quad (23)$$

$$F1 = \sum_i \frac{N_i}{N} F1_i. \quad (24)$$

其中: N_i 为类别 i 的目标样本数; N 为整个数据集的样本数.

2.2 实验结果

2.2.1 舰船识别结果

表1为不同算法舰船目标识别结果,总体上传统识别方法表现都不如深度学习识别方法,表现最好的LR算法的F1分数仅为92.58%. SAR图像通常存在较大噪声干扰,传统识别方法仅利用SAR图像的表层特征不能充分提取图像蕴含的信息,容易受到背景杂波或相干斑噪声的影响.而深度学习识别方法通过卷积层进行图像特征提取,利用梯度下降法进行迭代训练优化,能更准确识别SAR图像中舰船,且图像中噪声可通过

卷积核权重的自适应学习在一定程度上进行避免. AlexNet表现较差, $F1$ 分数仅为 92.68%, 由于其模型层数较浅, 不能准确提取图像中高层特征, 容易受图像噪声影响. 相比 AlexNet, GoogleNet 加深了网络深度, 识别性能得到了提高. 然而, CNN 在模型层数加深情况下容易存在梯度消失的问题, 导致模型难以训练. 因此, DenseNet, ResNet 和 MobileNet 通过引入残差结构来缓解梯度消失的问题, 取得了较好识别性能. 此外, 随着 ResNet 的层数加深, 模型识别效果反而出现一定程度的下降. 这是由于 SAR 图像相比于光学图像不具有丰富的特征, 更深的网络结构反而导致模型性能变差.

表 1 不同方法的舰船目标识别结果
Table 1 Ship recognition results of different methods

方法	模型	%		
		P	R	$F1$
传统识别方法	LR	92.67	92.65	92.58
	KNN	92.50	92.49	92.50
	SVM	88.13	87.37	87.75
	DT	89.34	89.39	89.36
深度学习识别方法	AlexNet	92.67	92.73	92.68
	GoogleNet	94.44	94.49	94.46
	MobileNet	94.81	94.82	94.80
	DenseNet	95.14	95.15	95.15
	ResNet-18	95.05	95.07	95.06
	ResNet-34	94.92	94.90	94.96
	ResNet-50	94.10	94.06	94.05
本文 SpikingSAR	95.57	95.59	95.58	

本文 SpikingSAR 采用 CNN 与 SNN 相结合架构, 设计了较浅的结构能减少模型复杂度和避免深度网络带来的问题, 取得最好的识别性能, $F1$ 分数达到 95.58%. 实验结果验证了本文所提算法的有效性, 各项评价指标表现都优于传统方法, 且在大多数情况下取得与 CNN 相当甚至更优性能. 表 2 为本文 SpikingSAR 在 SAR-Dataset 上舰船目标识别的混淆矩阵. 混淆矩阵常用于监督学习中的分类识别任务, 能可视化展示模型对不同类别的预测结果. 由表 2 可知, 本文 SpikingSAR 能较准确地识别舰船样本, 仅对于少数其他样本出现虚警现象; 同时由于类船样本与背景样本相似度较高, 因此本文 SpikingSAR 只在少数情况下将二者错误识别.

表 2 舰船目标识别结果的混淆矩阵

样本集	舰船样本	类船样本	背景样本	$F1/\%$
舰船样本	396	2	1	98.55
类船样本	2	377	20	93.31
背景样本	1	24	374	94.21
$F1/\%$				95.58

2.2.2 模型参数量对比

图 5 为不同神经网络模型的参数量比较. AlexNet 模型参数量达到 228.03 MB, 由于其浅层结构设计依赖后续全连接层对图像特征的提取, 过多的全连接层导致较大模型参数量. GoogleNet 与 ResNet 采用堆叠卷积层的结构设计, 仅在最后一层使用全连接层, 大大减少了模型参数量. MobileNet 引入深度可分离卷积操作, 以更少参数和计算量获得同样输出, 进一步降低模型参数量. DenseNet 通过将网络中所有层进行相互连接, 提高了网络的信息和梯度流动, 且由于大大缩减卷积层的通道数, 因此能使用较少参数取得较优性能. 本文 SpikingSAR 的模型参数量是最少 (仅 3.11 MB) 的, 远少于 AlexNet, GoogleNet 和 ResNet 的模型参数量, 也比优化后的 MobileNet 与 DenseNet 的模型参数量更少. 其主要原因是: 1) 使用了卷积层与 LIF 脉冲神经元层相结合的结构设计, LIF 层主要作用是模拟脉冲神经元充放电过程, 不包含可训练的参数; 2) 在输入层对 SAR 图像进行基于视觉注意力机制的脉冲编码, 增强了图像特征信息, 因此不需要较深的网络结构就能有效提取到图像的高阶特征信息.

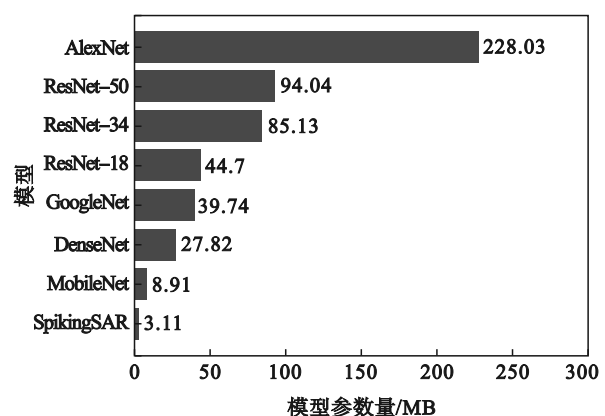


图 5 不同神经网络模型的参数量

Fig. 5 Parameters of the different neural network models

2.2.3 模型能耗对比

为了验证所提算法的能耗性能,本文分别对SpikingSAR与对比的神经网络使用合适指标衡量其操作数,计算各自的能耗.能耗的计算与硬件中操作数息息相关,不同的神经网络操作数计算方法也不同.对于传统ANN,衡量操作数的指标主要是浮点运算次数(floating-point operations, FLOP),认为计算量是衡量模型计算性能的1个指标.对于SNN,衡量操作数的指标主要是突触运算次数(synaptic operations, SOP),即神经元膜电位的变化.传统ANN运行于Intel Stratix 10 TX FPGA,能量消耗为12.5 pJ/FLOP. SNN运行于神经形态硬件ROLLS上,通过在神经元间传输脉冲信号来提供触发事件的高效计算,仅在神经元发射脉冲过程才消耗计算资源,能量消耗为77 fJ/SOP.

表3为不同神经网络操作数与能耗对比.由表3可以发现,基于脉冲信号传递的SpikingSAR具有较少操作数(SOP为17.97 MB),仅为其他ANN操作数的1/30~1/2,说明利用脉冲传输信号过程的高效性.由表3可以发现,在神经形态硬件ROLLS上使用SpikingSAR消耗的能量远小于传统的ANN.此外,SpikingSAR获得将近3个数量级的能耗优势,说明本文所提算法具有显著的能源利用效率.

表3 不同神经网络模型操作数(FLOP/SOP)与能耗对比
Table 3 Comparison of the operations (FLOP/SOP) and energy-consumption for the different models

模型	(FLOP/SOP)/MB	能耗/J
AlexNet	98.59	1.23×10^{-3}
GoogleNet	143.07	1.79×10^{-3}
MobileNet	43.54	0.54×10^{-3}
DenseNet	291.63	3.65×10^{-3}
ResNet-18	243.53	3.04×10^{-3}
ResNet-34	489.96	6.12×10^{-3}
ResNet-50	549.02	6.86×10^{-3}
本文 SpikingSAR	17.97	1.38×10^{-6}

2.2.4 仿真时长分析

由于实验中脉冲序列长度固定为输入图像编码后的像素值个数,因此脉冲序列传播的计算主要受到脉冲数量影响,而脉冲数量则由选定的仿真时长 T 决定.当 T 过小时可减少计算复杂度和能耗,但可能无法精确描述神经元状态的变化过程;当 T 过大会造成网络计算量与能耗增加,

也可能破坏网络对脉冲信号的模拟精度.脉冲神经元模型的选取对SNN性能也至关重要,合适的神经元模型能利用合理数学模型建模生物系统智能运算过程,并融合生物启发的、高效精确的神经信息处理机制到SNN中,赋予SNN更好生物可信度与强大拟合能力.

针对不同的仿真时长 T 与脉冲神经元模型在SAR-Dataset数据集进行实验分析,见图6.可发现:1)不同的脉冲神经元模型具有不同的性能表现.LIF模型在 $F1$ 分数的表现优于IF模型,这是因为LIF模型在简化神经元状态变化过程中,考虑了生物神经元膜电位变化过程中存在泄漏的关键特征,能更准确对神经元活动细节进行描述,增强SNN的生物精确度和性能表现;2)SpikingSAR模型对不同仿真时长 T 的敏感性.仿真时长 T 较小时,使用IF与LIF模型的SpikingSAR表现较差;随着 T 逐渐增大,模拟的时间步长逐渐精细,SpikingSAR的识别准确率逐渐上升,且在 $T=12$ s时达到峰值,说明 $T=12$ s是一个较为合理的仿真时长.当 $T>12$ s时,SpikingSAR的识别准确率开始下降,说明 T 较大时反而会影响模型的识别性能.该现象可能原因是较大 T 会导致输入的脉冲序列过多,使用替代梯度训练方法造成的误差累积逐渐增大,从而影响模型的学习能力.

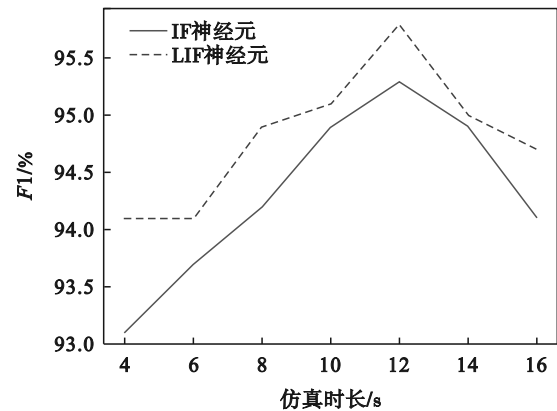


图6 不同神经元模型仿真时长与 $F1$ 关系

Fig. 6 Relation between simulation time and $F1$ for different neuron models

3 结 语

本文提出了一种基于SNN的轻量化SAR图像舰船目标识别算法,基于实际场景下SAR图像数据集SAR-Dataset进行了实验验证,得出以下结论:

1) 结合视觉注意力机制和泊松编码器的方

法能有效对输入图像进行脉冲编码,保留图像的重要信息并消除噪声干扰;

2) 使用替代梯度方法对 SNN 进行训练是有效的, SpikingSAR 获得了与传统方法和 ANN 模型相当甚至更优的舰船识别性能;

3) SpikingSAR 使用稀疏脉冲在神经网络中传递消息,利用神经形态硬件进行高效计算,具有更低能耗。

然而,由于 SAR 舰船目标细节特征不明显、类别分布不平衡,使得基于深度学习的舰船目标识别同样面临较大挑战,直接应用经典的图像分类网络往往效果不佳。而 SAR 图像具有不同的极化方式、不同域的数据形式,因而研究人员逐渐将视角转向融合不同来源的数据以期提升舰船目标识别的表现,如融合不同极化方式的 SAR 图像数据、融合复数域数据、融合手工特征和神经网络特征等。因此,基于多极化数据融合的 SAR 图像舰船目标识别技术研究将是课题组后续开展研究工作的重要方向之一。

参考文献:

- [1] Hu X, Xie H T, Zhang L, et al. Fast factorized backprojection algorithm in orthogonal elliptical coordinate system for ocean scenes imaging using geosynchronous spaceborne - airborne VHF UWB bistatic SAR [J]. *Remote Sensing*, 2023, 15(8):2215-2241.
- [2] Chen J X, Xie H T, Zhang L, et al. SAR and optical image registration based on deep learning with co-attention matching module [J]. *Remote Sensing*, 2023, 15(15):3879-3896.
- [3] Jiang X Q, Xie H T, Chen J X, et al. Arbitrary-oriented ship detection method based on long-edge decomposition rotated bounding box encoding in SAR images [J]. *Remote Sensing*, 2023, 15(3):673-694.
- [4] Jiang X Q, Xie H T, Lu Z, et al. Energy-efficient and high-performance ship classification strategy based on siamese spiking neural network in dual-polarized SAR images [J]. *Remote Sensing*, 2023, 15(20):4966-4983.
- [5] Xie H T, He J F, Lu Z, et al. Two-level feature-fusion ship recognition strategy combining HOG features with dual-polarized data in SAR images [J]. *Remote Sensing*, 2023, 15(18):4393-4405.
- [6] Devlin J, Chang M W, Lee K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [C]//Conference of the North American Chapter of the Association for Computational Linguistics. Minneapolis, 2019:4171-4186.
- [7] Brown T B, Mann B, Ryder N, et al. Language models are few-shot learners [C]//International Conference on Neural Information Processing Systems. Vancouver, 2020: 1877-1901.
- [8] Fedus W, Zoph B, Shazeer N. Switch transformers: scaling to trillion parameter models with simple and efficient sparsity [J]. *Journal of Machine Learning Research*, 2022, 23:1-40.
- [9] Roy K, Jaiswal A, Panda P. Towards spike-based machine intelligence with neuromorphic computing [J]. *Nature*, 2019, 575(7784):607-617.
- [10] Hong C F, Yuan M W, Zhang M X, et al. SPAIC: a spike-based artificial intelligence computing framework [J]. *IEEE Computational Intelligence Magazine*, 2024, 19(1):51-65.
- [11] Kim S, Park S, Na B, et al. Spiking - YOLO: spiking neural network for energy - efficient object detection [C]//AAAI Conference on Artificial Intelligence. New York, 2020: 11270-11277.
- [12] 李超, 李悦丽, 安道祥, 等. 基于视觉注意机制的 UWB SAR 叶簇隐蔽目标变化检测 [J]. *电子学报*, 2016, 44(1): 39-45.
(Li Chao, Li Yue-li, An Dao-xiang, et al. UWB SAR change detection of foliage - concealed targets based on visual attention [J]. *Acta Electronica Sinica*, 2016, 44(1): 39-45.)
- [13] Gerstner W, Kistler W, Naud R, et al. *Neuronal dynamics: from single neurons to networks and models of cognition* [M]. Cambridge: Cambridge University Press, 2014.
- [14] Schwegmann C P, Kleynhans W, Salmon B P, et al. Very deep learning for ship discrimination in synthetic aperture radar imagery [C]//IEEE International Geoscience and Remote Sensing Symposium. Beijing, 2016: 104-107.