

## 一种基于EfficientNet与BiGRU的多角度SAR图像目标识别方法

赵鹏菲<sup>①②③</sup> 黄丽佳<sup>\*①②</sup>

<sup>①</sup>(中国科学院空天信息创新研究院 北京 100094)

<sup>②</sup>(中国科学院空间信息处理与应用系统技术重点实验室 北京 100190)

<sup>③</sup>(中国科学院大学 北京 100049)

**摘 要:** 合成孔径雷达(SAR)的自动目标识别(ATR)技术目前已广泛应用于军事和民用领域。SAR图像对成像的方位角极其敏感,同一目标在不同方位角下的SAR图像存在一定差异,而多方位角的SAR图像序列蕴含着更加丰富的分类识别信息。因此,该文提出一种基于EfficientNet和BiGRU的多角度SAR目标识别模型,并使用孤岛损失来训练模型。该方法在MSTAR数据集10类目标识别任务中可以达到100%的识别准确率,对大俯仰角(擦地角)下成像、存在版本变体、存在配置变体的3种特殊情况下的SAR目标分别达到了99.68%, 99.95%, 99.91%的识别准确率。此外,该方法在小规模的数据集上也能达到令人满意的识别准确率。实验结果表明,该方法在MSTAR的大部分数据集上识别准确率均优于其他多角度SAR目标识别方法,且具有一定的鲁棒性。

**关键词:** 合成孔径雷达; 自动目标识别; 多角度识别; EfficientNet

中图分类号: TP753

文献标识码: A

文章编号: 2095-283X(2021)06-0895-10

DOI: [10.12000/JR20133](https://doi.org/10.12000/JR20133)

**引用格式:** 赵鹏菲, 黄丽佳. 一种基于EfficientNet与BiGRU的多角度SAR图像目标识别方法[J]. 雷达学报, 2021, 10(6): 895–904. doi: 10.12000/JR20133.

**Reference format:** ZHAO Pengfei and HUANG Lijia. Target recognition method for multi-aspect synthetic aperture radar images based on EfficientNet and BiGRU[J]. *Journal of Radars*, 2021, 10(6): 895–904. doi: 10.12000/JR20133.

## Target Recognition Method for Multi-aspect Synthetic Aperture Radar Images Based on EfficientNet and BiGRU

ZHAO Pengfei<sup>①②③</sup> HUANG Lijia<sup>\*①②</sup>

<sup>①</sup>(*Aerospace Information Research Institutue, Chinese Academy of Sciences, Beijing 100094, China*)

<sup>②</sup>(*Key Laboratory of Technology in Geo-spatial Information Processing and Application System, Chinese Academy of Sciences, Beijing 100190, China*)

<sup>③</sup>(*University of Chinese Academy of Sciences, Beijing 100049, China*)

**Abstract:** Automatic Target Recognition (ATR) in Synthetic Aperture Radar (SAR) has been extensively applied in military and civilian fields. However, SAR images are very sensitive to the azimuth of the images, as the same target can differ greatly from different aspects. This means that more reliable and robust multiaspect ATR recognition is required. In this paper, we propose a multiaspect ATR model based on EfficientNet and BiGRU. To train this model, we use island loss, which is more suitable for SAR ATR. Experimental results have revealed that our proposed method can achieve 100% accuracy for 10-class recognition on the Moving and Stationary Target Acquisition and Recognition (MSTAR) database. The SAR targets in three special imaging cases with large depression angles, version variants, and configuration variants reached recognition accuracies of 99.68%, 99.95%, and 99.91%, respectively. In addition, the proposed method achieves satisfactory accuracy even

收稿日期: 2020-10-26; 改回日期: 2020-12-21; 网络出版: 2021-01-07

\*通信作者: 黄丽佳 [iecasshuanglijia@163.com](mailto:iecasshuanglijia@163.com) \*Corresponding Author: HUANG Lijia, [iecasshuanglijia@163.com](mailto:iecasshuanglijia@163.com)

基金项目: 国家自然科学基金(61991420, 62022082), 中科院青促会专项支持

Foundation Items: The National Natural Science Foundation of China (61991420, 62022082), Special Support of Youth Innovation Promotion Association Chinese Academy of Sciences

责任编辑: 林赞 Corresponding Editor: LIN Yun

with smaller datasets. Our experimental results show that our proposed method outperforms other state-of-the-art ATR methods on most MSTAR datasets and exhibits a certain degree of robustness.

**Key words:** Synthetic Aperture Radar (SAR); Automatic Target Recognition (ATR); Multi-aspect SAR; EfficientNet

## 1 引言

合成孔径雷达(Synthetic Aperture Radar, SAR)是一种主动式微波遥感设备,其通过天线主动地向目标发射微波,并接收目标反射的回波实现对目标的成像<sup>[1]</sup>。SAR不仅可以在光照不足和恶劣天气下工作,而且可以实现对隐蔽目标的穿透探测,具有全天时、全天候工作的特点<sup>[2]</sup>。经过30余年的发展,SAR技术越来越成熟,目前已经成为遥感信息获取的重要途径,在军用以及民用领域都有广泛的应用<sup>[3]</sup>。

与光学图像不同,由于电磁散射和相干成像机理,SAR图像是目标三维几何和辐射信息在二维图像中的映射,存在叠掩、阴影等特点,且含有大量相干斑噪声。这使SAR图像在视觉上较光学图像更难解释与理解,对目标的检测和识别有较大的影响<sup>[4]</sup>。因此,从SAR图像中识别目标的真实属性,即SAR自动目标识别技术(SAR Automatic Target Recognition, SAR ATR)成为一直以来的研究热点。

近年来,随着深度学习技术的发展,基于深度学习的SAR ATR技术得到了广泛的应用<sup>[5-8]</sup>。都设计了不同的卷积神经网络来对SAR图像进行目标识别任务,并取得了不错的效果,但这些方法都是基于SAR的单角度图像进行的目标识别。

由于SAR采用侧视成像原理,SAR对目标观测俯仰角和方位角的变化都十分敏感。对于同一观测目标,即便观测的方位角仅有较少变化,所得目标的SAR图像也会有较大的差别。而对于不同目标在某一个观测角度下又呈现较大的相似性,这些情况对基于单角度SAR图像的目标识别造成相当的困难。此外,同一观测目标的多角度SAR图像序列中蕴含着多个角度的散射回波,较单角度SAR图像更加全面精细地刻画了目标散射特性<sup>[9]</sup>。

多角度SAR目标识别技术基于SAR平台获取目标不同方位角的多幅图像组成的图像序列,利用不同视角的散射特性对目标进行类别属性判别,其拥有更高的准确度和鲁棒性<sup>[10,11]</sup>。

目前,多角度SAR目标识别技术也有了一定的发展。2017年,Zhang等人<sup>[12]</sup>提出多角度感知双向长短期记忆网络(Multi-Aspect-aware Bidirectional Long-Short Term Memory, MA-BLSTM)用于

多角度SAR目标识别,其首先提取单张图像的Gabor特征和3片局部二值模式(Three Patches Local Binary Pattern, TPLBP)特征,并进一步利用长短期记忆网络(Long-Short Term Memory, LSTM)提取了图像特征序列的时序特征,最后通过Softmax层进行目标的分类,取得了不错的效果。2018年,Pei等人<sup>[13]</sup>提出多视角深度卷积神经网络(Multi-View Deep Convolutional Neural Network, MVDCNN)方法,利用并行的卷积神经网络(Convolutional Neural Network, CNN)来提取不同角度图像的深度特征,并逐步将这些特征进行融合,从而进行识别。邹浩等人<sup>[14]</sup>将3张不同角度的SAR图像合并作为1张3通道的图像送入CNN网络进行目标识别,取得了不错的效果。Zhao等人<sup>[15]</sup>提出了多流卷积神经网络(Multi-Stream Convolutional Neural Network, MS-CNN)方法,设计了傅里叶特征融合层,来对由卷积神经网络提取的深度特征进行了融合。2019年,Zhang等人<sup>[16]</sup>在文献<sup>[12]</sup>的基础上,将Gabor特征和TPLBP特征换为由ResNet提取的深度特征,并提出基于投票的后处理方法,取得了目前最高的准确率。

尽管上述方法都取得了不错的效果,但仍然存在一些尚未解决的问题,如在训练集和测试集图像成像条件不同时,识别准确率较低;基于深度学习的方法通常需要大量数据进行训练等。

为解决以上问题,本文提出基于EfficientNet<sup>[17]</sup>和BiGRU<sup>[18]</sup>的多角度SAR目标识别方法,主要贡献如下:

(1) 本文采用EfficientNet提取单张SAR图像的空间特征,并进一步利用BiGRU提取图像序列的时序特征。这使模型保持高性能的同时也降低了模型参数量;

(2) 本文采用孤岛损失(Island Loss)<sup>[19]</sup>改进分类模型的训练损失。在SAR图像目标识别任务中,Island Loss通过扩大各个目标类别间的类间距离及缩小类内距离来有效提升识别准确率。

实验证明,本方法不仅在大部分数据集上取得了当前最好的识别效果,还在小规模数据集上取得了不错的识别性能。

本文的第2节详细介绍了本文所提多角度SAR目标识别方法;第3节介绍实验所采用的图像数据以及

多角度图像数据集的构造；第4节介绍实验细节与结果；第5节总结本方法的优越性并安排了后续的工作。

## 2 多角度SAR图像目标识别方法

### 2.1 多角度SAR图像序列识别网络结构

本文所提多角度SAR图像目标识别网络的结构如图1所示，其主要由提取单张图像特征的EfficientNet网络与进一步提取多角度SAR图像序列时序特征的BiGRU网络构成。

该网络的输入数据是一组同一目标的多角度SAR图像序列 $(B, L, C, H, W)$ 。其中， $B$ 为模型训练时的Batch Size， $L$ 为图像序列的图像数量， $C, H, W$ 分别为图像通道数、图像高度与宽度。

该方法首先使用一组共享权重的EfficientNet网络提取图像序列中单张图像的空间特征 $(B \times L, U)$ ，其中 $U$ 为单张图像空间特征的维度。对图像序列所提取的图像特征序列进行维度变换，得到特征 $(B, L, U)$ ，再送入BiGRU网络继续提取多角度图像序列时序特征 $(B, V)$ ，其中 $V$ 为特征的维度。最后，将经过BiGRU网络提取的特征送入全连接层获取最终的输出，即该组多角度图像的目标类别概率 $(B, k)$ ，其中 $k$ 为待识别目标的类别数目。

### 2.2 EfficientNet

EfficientNet系列模型是目前性能最好的深度学习分类模型之一，该模型的基础网络架构是通过神经网络架构搜索(Neural Architecture Search, NAS)技术得到的<sup>[17]</sup>。本文选取了EfficientNet系列模型中的EfficientNet-B0网络，该网络在EfficientNet系列模型中尺寸最小，速度最快，更适用于SAR领域这类小规模数据集。

本文所使用的EfficientNet-B0网络结构组成如表1所示，其由16个移动倒置瓶颈卷积模块(Mobile inverted Bottleneck Convolution, MB-Conv)<sup>[17]</sup>、2个卷积层、1个全局平均池化层和1个全连接层构成。

深度学习任务中常用的残差网络ResNet分类网络<sup>[20]</sup>，其也应用于文献<sup>[16]</sup>进行多角度SAR目标图像的识别。表2对比了EfficientNet-B0和ResNet50网络的参数量、计算量与性能，top1/top5准确率是在ImageNet数据集中得出的<sup>[17,21]</sup>。top1/top5准确率如式(1)所示， $n$ 为样本总数， $g_k$ 表示第 $k$ 个样本的真实标签， $l_j$ 表示模型对当前样本所得概率向量中第 $j$ 大的概率所对应的标签，函数 $d$ 如式(2)所示

$$\left. \begin{aligned} \text{Acc}_{\text{top1}} &= \frac{1}{n} \sum_{k=1}^n d(g_k, l_1) \\ \text{Acc}_{\text{top5}} &= \frac{1}{n} \sum_{k=1}^n \sum_{j=1}^5 d(g_k, l_j) \end{aligned} \right\} \quad (1)$$

$$d(x, y) = \begin{cases} 1, & x = y \\ 0, & x \neq y \end{cases} \quad (2)$$

如表2所示，对比ResNet50网络，EfficientNet-B0网络的速度快10倍、参数仅为其1/5且性能更好，这使其更适用于较难获取、数量较少的SAR遥感目标数据集。此外，由于EfficientNet-B0网络参数较少，模型部署更加容易。

在本文所提多角度SAR图像识别框架中，EfficientNet-B0网络的输入是单张SAR的目标图像，输出是其空间特征。模型将多角度图像序列的特征进行维度变换后送入后续的BiGRU网络。

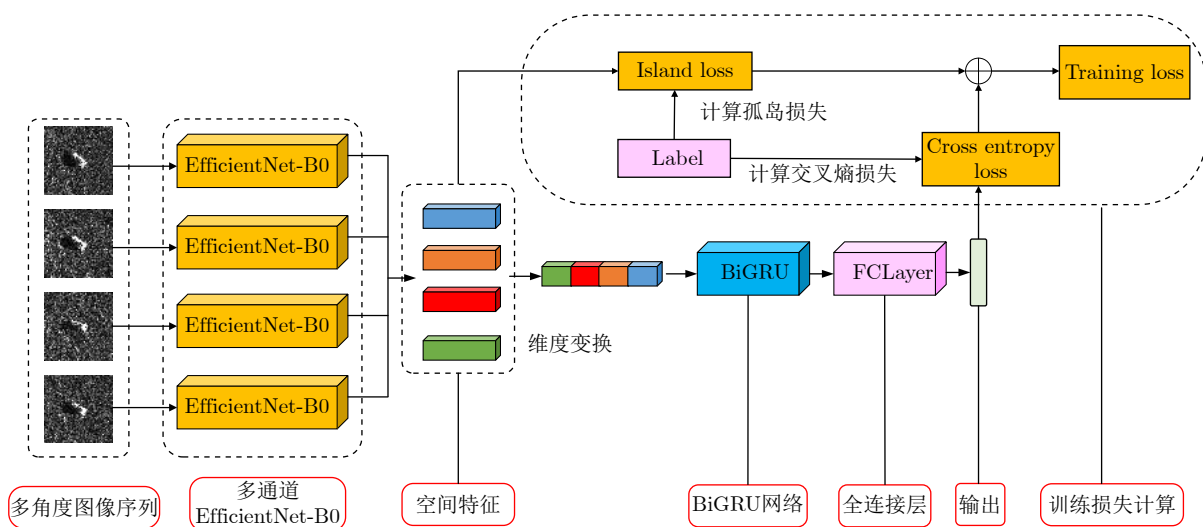


图1 多角度SAR目标识别网络结构图

Fig. 1 Multi-aspect SAR ATR framework

表 1 EfficientNet-B0网络结构

Tab. 1 EfficientNet-B0 network structure

阶段	模块	输出尺寸	层数
1	Conv3×3	16×32×32	1
2	MBCConv1, k3×3	24×32×32	1
3	MBCConv6, k3×3	40×16×16	2
4	MBCConv6, k5×5	80×8×8	2
5	MBCConv6, k3×3	112×8×8	3
6	MBCConv6, k5×5	192×4×4	3
7	MBCConv6, k5×5	320×2×2	4
8	MBCConv6, k3×3	1280×2×2	1
9	Conv1×1 & Pooling & FC	k	1

表 2 EfficientNet-B0与ResNet50网络对比

Tab. 2 Comparison of EfficientNet-B0 and ResNet50 networks

模型	参数量 (M)	FLOPS (B)	top1/top5准确率 (%)
EfficientNet-B0	5.3	0.39	77.3/93.5
ResNet50	26.0	4.10	76.0/93.0

### 2.3 BiGRU

门控循环单元(Gate Recurrent Unit, GRU)<sup>[18]</sup>是一种特殊的循环神经网络(Recurrent Neural Network, RNN), 其与长短期记忆网络(LSTM)<sup>[21]</sup>相似, 是为了解决长期记忆和反向传播中的梯度等问题而提出的。

GRU适宜于处理时间序列数据, 本文使用GRU来提取多角度SAR图像序列的空间变化特征。对比LSTM, GRU在性能相当的同时参数量更少、更易于训练。相比之下更适宜于小规模数据集, 因此本文选取了GRU网络用于多角度图像序列的处理。

GRU由更新门和重置门两个门控单元组成, 如图2所示, 其利用门控单元控制输入、记忆等数据在当前时刻做出预测, 其表达式为式(3)–式(6), 其中 $z_t$ 为更新门,  $r_t$ 为重置门,  $\mathbf{W}_z, \mathbf{W}_r, \mathbf{W}$ 分别为更新门、重置门和候选隐层状态的权值矩阵,  $x_t$ 为输入信息,  $h_t$ 为 $t$ 时刻的隐层状态,  $\tilde{h}_t$ 为 $t$ 时刻的候选隐层状态,  $h_{t-1}$ 为 $t-1$ 时刻的隐层状态,  $\sigma$ 是Sigmoid函数。重置门将新的输入信息与前面的记忆相结合, 输出一个0~1的值, 其决定将 $h_{t-1}$ 向下一状态传递的多少。更新门定义了先前记忆保存至当前时刻的量, 即先前记忆对当前状态的影响程度

$$z_t = \sigma(\mathbf{W}_z \cdot [h_{t-1}, x_t]) \quad (3)$$

$$r_t = \sigma(\mathbf{W}_r \cdot [h_{t-1}, x_t]) \quad (4)$$

$$\tilde{h}_t = \tanh(\mathbf{W} \cdot [r_t \times h_{t-1}, x_t]) \quad (5)$$

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \tilde{h}_t \quad (6)$$

然而, 普通的GRU网络仅能编码从前向后的序列信息, 无法编码从后往前的序列信息。因此, 本文采用了双向的GRU网络BiGRU来获取了图像特征序列双向信息。

本文所采用的BiGRU结构如图3所示。

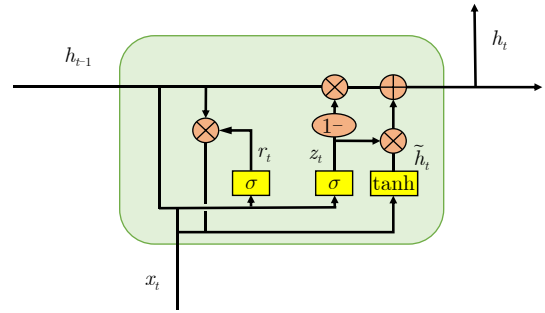


图 2 GRU结构示意图

Fig. 2 The structure of GRU

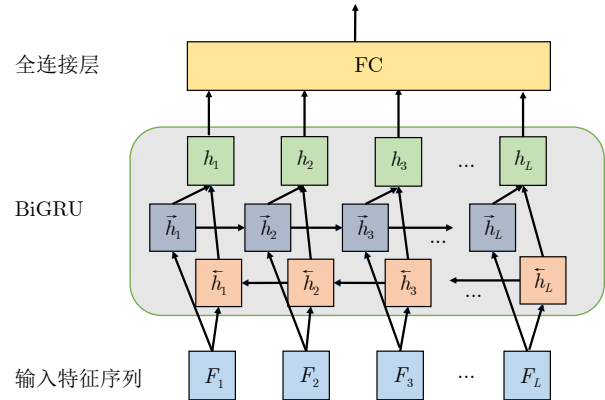


图 3 BiGRU结构示意图

Fig. 3 The structure of BiGRU

BiGRU的计算公式如式(7)–式(9)所示, 其中 $\vec{h}_t$ 与 $\overleftarrow{h}_t$ 分别表示前向GRU与后向GRU输出的隐层状态,  $w_t$ 与 $v_t$ 分别表示 $\vec{h}_t$ 与 $\overleftarrow{h}_t$ 所对应的权重,  $b_t$ 则为偏置参数。可以看出, 在BiGRU中, 当前隐层状态 $h_t$ 由前向GRU与后向GRU输出的隐层状态加权所得

$$\vec{h}_t = \text{GRU}(x_t, \vec{h}_{t-1}) \quad (7)$$

$$\overleftarrow{h}_t = \text{GRU}(x_t, \overleftarrow{h}_{t-1}) \quad (8)$$

$$h_t = w_t \vec{h}_t + v_t \overleftarrow{h}_t + b_t \quad (9)$$

在本文的实验中, BiGRU网络的隐层节点数为128, 网络层数为4。

### 2.4 训练损失

在基于深度学习的目标识别任务中, 交叉熵损

失是最常使用的。然而，由于SAR图像自身成像机理的复杂性和成像角度的敏感性，SAR图像通常存在同一目标之间类内距离大，不同目标之间类间距离小的问题。因此，简单地使用交叉熵不适用于SAR图像目标任务。如图1所示，本文使用孤岛损失(Island Loss)<sup>[19]</sup>和交叉熵损失共同作为训练损失来解决以上问题，其中交叉熵损失公式为

$$L_{CE} = - \sum_{i=1}^m y_i \log(p_i) \quad (10)$$

其中， $y_i$ 是第*i*个样本的类别， $p_i$ 是模型输出的第*i*个值。

Island Loss可以在训练中扩大类间距离、缩小类内距离，从而提高模型的性能。

Island Loss基于中心损失(Center Loss)<sup>[22]</sup>提出，Center Loss的公式为

$$L_C = \frac{1}{2} \sum_{i=1}^m \|x_i - c_{y_i}\|^2 \quad (11)$$

其中， $y_i$ 是第*i*个样本的类别， $x_i$ 是第*i*个样本由EfficientNet-B0输出的特征向量， $c_{y_i}$ 第*i*个样本所在类别的所有样本计算的中心点。根据式(11)，Center Loss为每个样本特征距离其所在类中心的欧氏距离之和。因此，通过最小化Center Loss，可以将同一类别的不同样本的特征聚向类别中心，以此来缩小类内距离。

Center Loss仅仅可以缩小类内距离，而Island Loss在其基础上进一步优化了不同目标的类间距离。其公式为

$$L_{IL} = L_C + \lambda_1 \sum_{c_j \in N} \sum_{\substack{c_k \in N \\ c_j \neq c_k}} \left( \frac{c_j \cdot c_k}{\|c_j\|_2 \|c_k\|_2} + 1 \right) \quad (12)$$

其中， $c_j$ 和 $c_k$ 分别是第*j*个类别和第*k*个类别的中心点， $\lambda_1$ 为超参数。由式(12)可知，Island Loss的前半部分是Center Loss，后半部分是训练样本所属各个类别的特征中心之间的余弦相似度加1后的总和。当第*j*个类别和第*k*个类别的中心点 $c_j$ 和 $c_k$ 相差较远时，其余弦相似度接近于-1，此时的两个类别之间所产生的损失为0；当 $c_j$ 和 $c_k$ 相近时，其余弦相似度接近于1，此时两个类别之间所产生的损失接近于最大值2。由此可见，可以通过最小化Island Loss，将同一类别的不同样本的特征聚向类别中心的同时，并将类别中心互相推离彼此，在高维空间中形成一座座“孤岛”，因此Island Loss可以扩大类间距离，缩小类内距离。

在联合交叉熵损失后，本文所使用的训练损失为

$$L = L_{CE} + \lambda L_{IL} \quad (13)$$

其中， $\lambda$ 为超参数。

### 3 数据集

本文使用了由桑迪亚国家实验室(Sandia national laboratory)提供的SAR图像数据集——移动与静止目标获取识别(Moving and Stationary Target Acquisition and Recognition, MSTAR)数据集，针对SAR图像目标识别的研究大多都基于该数据集而展开。在该数据集中，图像的分辨率为0.3 m × 0.3 m，工作波段为X波段，所用的极化方式为HH极化方式。数据集中包含了10种不同类型的地面车辆目标在不同俯仰角、不同方位角下的图像数据。如图4所示，其包含了目标车辆在0°~360°间各个方位角的图像数据，因此该数据集适宜于建立多角度图像目标识别数据集。

本文采用了文献[16]中多角度SAR图像数据集的构造方式。目前，宽角SAR一次可以对目标实现90°以下方位角的连续观察，为了使方法更实用，本文设置连续观测角 $\theta$ 为45°，多角度图像序列包含的图像个数*L*分别设置为2, 3, 4，数据集的构造方式如图5所示。

构建完成后的多角度SAR图像序列数据集分为标准工作条件(Standard Operating Condition, SOC)和扩展工作条件(Extended Operating Condition, EOC)两类。其中SOC指的是训练集和测试集数据的SAR传感器的成像条件相似，在实验中，训练集的成像俯仰角为17°，测试集的成像俯仰角为15°；EOC指的是训练集和测试集数据的SAR传感器的成像条件存在一定的差异。图像序列*L* = 4的SOC数据集目标种类及数量如表3所示，其中训练集合计10592组，测试集合计8809组；当图像序列*L* = 3时，训练集合计8016组，测试集合计6771组；当图像序列*L* = 2时，训练集合计5347组，测试集合计4669组。

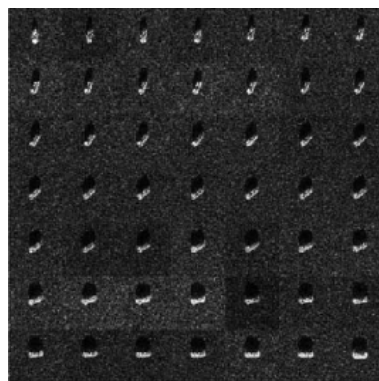


图4 不同方位角、同一目标的SAR图像

Fig. 4 SAR images of the same target with different azimuth

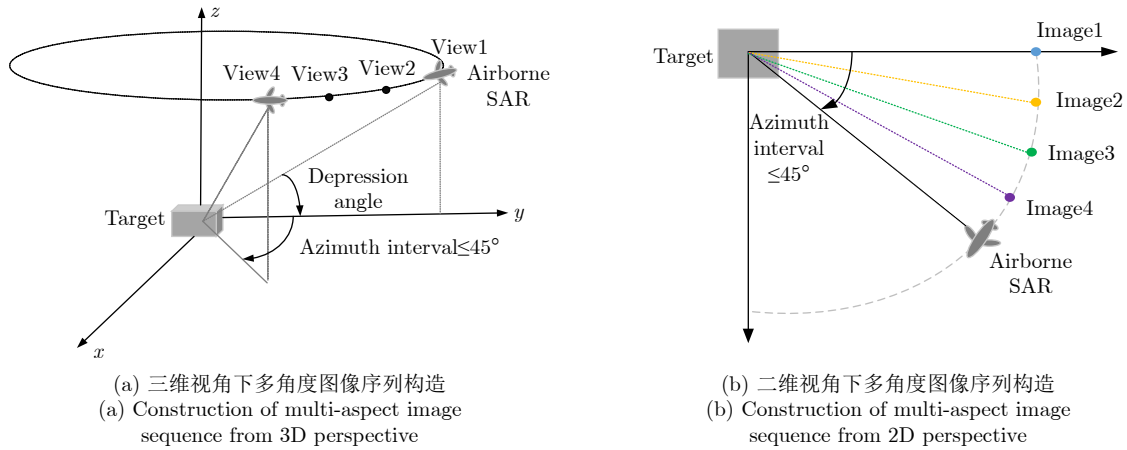


图 5 多角度图像序列构造示意图

Fig. 5 Schematic diagram of multi-angle image sequence structure

表 3 图像序列L为4时, SOC数据集大小

Tab. 3 SOC dataset size when L=4

目标名称	训练集数量	测试集数量
2S1	1162	1034
BMP2	883	634
BRDM_2	1158	1040
BTR70	889	649
BTR60	978	667
D7	1162	1037
T62	1162	1032
T72	874	642
ZIL131	1162	1034
ZSU_234	1162	1040
合计	10592	8809

表 4 图像序列L为4时, EOC-1数据集大小

Tab. 4 EOC-1 dataset size when L=4

目标名称	训练集数量	测试集数量
2S1	1166	1088
BRDM_2	1162	1084
T72	913	1088
ZSU_234	1166	1088
合计	4407	4348

在MSTAR数据集中, EOC通常有3种, 第1类为大俯仰角情况, 本文称为EOC-1, EOC-1的训练集成像俯仰角为17°, 测试集成像俯仰角为30°, 其中图像序列L = 4时, EOC-1数据集目标种类及数量如表4所示。

此外, 另外两种EOC分别指车辆版本变化与车辆配置变化。版本变化是指车辆的功能改变, 即原始车辆变化为救护车、运输车、侦察车等; 配置变化是指车辆上一些部件的增加或移除, 如T72移除车后的油罐等<sup>[15]</sup>。在本文称这两种变化分别为EOC-2和EOC-3。

实验中选取了SOC训练集中的BMP2, BRDM\_2, BTR70与T72作为EOC-2, EOC-3的训练集, 选择T72的5类版本变体S7, A32, A62, A63, A64作为EOC-2的测试集, 选择BMP2的两类配置变体9566, C21及T72的5类配置变体812, A04, A05, A07, A10作为EOC-3的测试集<sup>[15]</sup>。

构建完成后的EOC-1, EOC-2与EOC-3训练集与测试集总数量如表5所示。

在图像预处理中, 本文对所有的图像均中心裁剪成64×64。此外, 本文还对部分训练集进行了数据增广, 具体方法采用了由文献<sup>[8]</sup>提出的SAR图像增广方法。首先将原始图片中心裁剪为68×68, 之后再随机裁剪多张64×64的图像, 进行数据增广的训练集图像序列数量如表6所示。

表 5 EOC-1, EOC-2与EOC-3数据集大小

Tab. 5 EOC-1, EOC-2 and EOC-3 dataset size

L	数据集	训练集总数	测试集总数
4	EOC-1	4407	4384
4	EOC-2	4473	9996
4	EOC-3	4473	12969
3	EOC-1	3307	3310
3	EOC-2	2889	7773
3	EOC-3	2889	10199
2	EOC-1	2202	2312
2	EOC-2	1934	5258
2	EOC-3	1934	6911

## 4 实验

### 4.1 训练与测试

在SOC实验中，模型的各项参数设置如表7所示。

在EOC-1实验中，数据增广前 $\lambda$ 为0.01，数据增广后的 $\lambda$ 为0.0001，训练的Epoch增加为300。对于Island Loss的超参数 $\lambda$ ，其根据数据集的大小而变化，数据集图像序列越大，其数值应越少。在EOC-2与EOC-3实验中，其各项参数与SOC保持一致。

在训练过程中，每80个Epoch，Adam优化器的学习率会缩减到1/10。此外，实验中使用了EfficientNet-B0在ImageNet数据集上的预训练模型，而BiGRU部分，则从头开始训练。

表 6 部分进行数据增广的数据集增广后大小

Tab. 6 The size of some data sets for data augmentation

$L$	数据集类型	训练集总数
4	EOC-1	17392
3	SOC	16032
3	EOC-1	13228
3	EOC-2&EOC-3	11544
2	SOC	16041
2	EOC-1	8808
2	EOC-2&EOC-3	7736

表 7 SOC实验中各参数设置

Tab. 7 Parameter in SOC experiment

名称	设置参数
Batch Size	32
优化器	Adam
Adam的学习率	0.001
Island Loss的优化器	SGD
SGD的学习率	0.5
Island Loss参数 $\lambda$	0.001
Island Loss参数 $\lambda_1$	10
Epochs	260

本文所有训练与测试实验均在Ubuntu18.04系统下进行。在硬件方面，实验采用了Intel i9-9900 CPU和NVIDIA RTX 2080Ti GPU。

### 4.2 SOC与EOC实验

在SOC实验中，其中图像序列个数为4，未进行数据集的增广，在正常数据集大小就达到了100%的识别准确率。

在EOC-1实验中，当图像序列个数为4时，且数据未增广时，识别准确率已达到99.08%，进行数据增广后，识别准确率提高为99.68%，其混淆矩阵如表8所示。

如表9所示，本文对比了当图像序列数为4时各主要方法的识别准确率。本文所提方法不仅在SOC数据集上取得了100%的识别准确率，在EOC-1数据集上取得的识别准确率也优于其他方法。

当图像序列数 $L$ 为3, 2时，SOC与EOC-1两种类型的测试集准确率如表10、表11，其中当图像序列数 $L$ 为3时，SOC的准确率为99.94%，EOC-1的准确率为98.58%；图像序列数 $L$ 为2时，SOC的准确率为99.87%，EOC-1的准确率为97.60%。与主要多角度识别方法进行对比如表10与表11所示。

此外，本文还测试了EOC-2, EOC-3两类数据集上的识别准确率，测试结果如表12、表13。由实验结果，本文所提方法在EOC-2与EOC-3数据集上均取得了不错的识别效果，当 $L$ 为4时，本方法在EOC-2的准确率99.95%，在EOC-3上的准确率为99.91%。

表 8 图像序列数 $L$ 为4时，EOC-1混淆矩阵

Tab. 8 The EOC-1 confusion matrix when  $L=4$

类型	S1	BRDM_2	T72	ZSU_234	Acc (%)
2S1	1076	2	10	0	98.90
BRDM_2	0	1084	0	0	100.00
T72	0	0	1088	0	100.00
ZSU_234	2	0	0	1086	99.82
平均值					99.68

表 9 图像序列数 $L$ 为4时，各方法识别准确率在SOC与EOC-1数据集上对比

Tab. 9 Comparison of the recognition accuracy on SOC and EOC-1 dataset when  $L$  is 4

序号	方法	SOC			EOC-1		
		准确率 (%)	图像样本数量	图像序列样本数量	准确率 (%)	图像样本数量	图像序列样本数量
1	MVDCNN <sup>[13]</sup>	98.52	690	43533	94.61	283	19705
2	MS-CNN <sup>[15]</sup>	99.92	2747	2747	98.61	1128	1128
3	ResNet-LSTM <sup>[16]</sup>	<b>100.00</b>	2000	7720	98.97	928	3614
4	本文方法	<b>100.00</b>	2747	10592	<b>99.08</b>	1128	4407
5	经过图像增广的本文方法	-	-	-	<b>99.68</b>	1128	17628

表 10 图像序列数 $L$ 为3时, 各方法准确率对比(%)  
Tab. 10 Comparison of test accuracy when  $L=3$  (%)

方法	SOC准确率	EOC-1准确率
MVDCNN <sup>[13]</sup>	98.17	94.34
MS-CNN <sup>[15]</sup>	99.88	97.48
本文方法	<b>99.94</b>	<b>98.58</b>

表 11 图像序列数 $L$ 为2时, 各方法准确率对比(%)  
Tab. 11 Comparison of test accuracy when  $L=2$  (%)

方法	SOC准确率	EOC-1准确率
MVDCNN <sup>[13]</sup>	97.81	93.29
MS-CNN <sup>[15]</sup>	99.84	96.69
本文方法	<b>99.87</b>	<b>97.60</b>

表 12 EOC-2数据集识别准确率对比(%)  
Tab. 12 Comparison of accuracy on EOC-2 (%)

方法	$L=4$	$L=3$	$L=2$
MVDCNN <sup>[13]</sup>	95.46	95.08	93.75
MS-CNN <sup>[15]</sup>	<b>100.00</b>	<b>100</b>	<b>99.67</b>
本文方法	99.95	99.82	99.39

表 13 EOC-3数据集识别准确率对比(%)  
Tab. 13 Comparison of accuracy on EOC-3 (%)

方法	$L=4$	$L=3$	$L=2$
MVDCNN <sup>[13]</sup>	95.45	95.25	94.98
MS-CNN <sup>[15]</sup>	99.58	99.08	98.71
本文方法	<b>99.91</b>	<b>99.57</b>	<b>99.13</b>

对比其他方法, 在SOC, EOC-1及EOC-3数据集上, 本方法取得了相对更优的识别结果, 在EOC-2数据集上, 本文方法仅略低于MS-CNN<sup>[15]</sup>。

#### 4.3 数据集缩减实验

由于SAR图像较难获取, SAR目标识别通常面临着数据样本较少的共性问题, 因此, 本文对图像序列 $L$ 为4时的训练集进行了缩减, 分别测试了本文所提方法和文献[13]所提方法在5%, 15%, 50%训练集规模上的识别准确率, 测试结果如表14所示, 可以看出, 本文所提方法在数据集缩减的情况下同样保持了较高的准确率。

#### 4.4 消融实验

为验证不同模块对模型性能的影响, 本文在图像序列数为4且未进行图像增广的EOC-1数据集上进行了消融实验。

消融实验以ResNet与GRU组成的基线模型为基础, 依次进行添加Center Loss、改变Center Loss为Island Loss、改变骨干网络为EfficientNet-

B0、改变GRU为BiGRU, 操作对性能的影响如表15所示, 其中序号1为基线模型的测试结果。

表 14 在缩减数据集上的识别准确率(%)  
Tab. 14 Recognition accuracy on the reduced dataset (%)

数据集规模	5%	15%	50%
本文方法	<b>95.98</b>	<b>99.72</b>	<b>99.93</b>
ResNet-LSTM <sup>[16]</sup>	93.97	99.37	99.58

表 15 消融实验结果  
Tab. 15 Results of ablation experiments

序号	Center Loss	Island Loss	EfficientNet	BiGRU	准确率 (%)	提升 (%)
1					94.08	-
2	✓				95.81	1.73
3		✓			97.03	1.22
4		✓	✓		98.46	1.43
5		✓	✓	✓	99.08	0.62

由表15的实验结果, 本文所采用的不同模块对模型的识别准确率均有较大的提升作用。在训练损失的改进上, Island Loss和Center Loss的加入对模型的识别准确率均有所提升, 其中更改训练损失为Island Loss的效果更好, 对比Center Loss提升了1.22%, 对比基线模型提升了2.95%; 更改骨干网络为EfficientNet-B0网络后, 对比ResNet模型, 识别准确率提升了1.43%; 将GRU网络更改为BiGRU后, 模型具有获取序列中双向信息的能力, 进一步使识别准确率提升了0.62%。

消融实验展示了本文所采用的EfficientNet, BiGRU和Island Loss对识别性能的影响, 三者的组合对比基线模型提升了5.00%的识别准确率, 从而使本文所提方法取得了相当的目标识别性能。

## 5 结束语

本文提出一种基于EfficientNet, BiGRU与Island Loss的多角度SAR图像目标识别模型, 其在SOC, EOC-1, EOC-3 3类数据集下的识别准确率均优于其他主流方法, 在EOC-2数据集下的识别准确率也接近目前的最优准确率。

同时此模型也在训练样本减小的情况下维持了较高的识别准确率, 有较好的鲁棒性, 是多角度SAR目标识别任务中一种行之有效的方法。

然而, 本文方法虽提高了EOC数据集的识别准确率, 但仍没有达到SOC数据集的识别准确率水平, 因此后续可以继续研究测试集和训练集存在一定差距情况下的模型改进和准确率提升问题。



## 参 考 文 献

- [1] 盖旭刚, 陈晋汶, 韩俊, 等. 合成孔径雷达的现状与发展趋势[J]. 飞航导弹, 2011(3): 82–86, 95.  
GAI Xugang, CHEN Jinwen, HAN Jun, *et al.* Development status and trend of synthetic aperture radar[J]. *Aerodynamic Missile Journal*, 2011(3): 82–86, 95.
- [2] 张红, 王超, 张波, 等. 高分辨率SAR图像目标识别[M]. 北京: 科学出版社, 2009.  
ZHANG Hong, WANG Chao, ZHANG Bo, *et al.* Target Recognition in High Resolution SAR Images[M]. Beijing: Science Press, 2009.
- [3] MOREIRA A, PRATS-IRAOLA P, YOUNIS M, *et al.* A tutorial on synthetic aperture radar[J]. *IEEE Geoscience and Remote Sensing Magazine*, 2013, 1(1): 6–43. doi: [10.1109/MGRS.2013.2248301](https://doi.org/10.1109/MGRS.2013.2248301).
- [4] 王瑞霞, 林伟, 毛军. 基于小波变换和PCA的SAR图像相干斑抑制[J]. 计算机工程, 2008, 34(20): 235–237. doi: [10.3969/j.issn.1000-3428.2008.20.086](https://doi.org/10.3969/j.issn.1000-3428.2008.20.086).  
WANG Ruixia, LIN Wei, and MAO Jun. Speckle suppression for SAR image based on wavelet transform and PCA[J]. *Computer Engineering*, 2008, 34(20): 235–237. doi: [10.3969/j.issn.1000-3428.2008.20.086](https://doi.org/10.3969/j.issn.1000-3428.2008.20.086).
- [5] CHEN Sizhe and WANG Haipeng. SAR target recognition based on deep learning[C]. 2014 International Conference on Data Science and Advanced Analytics, Shanghai, China, 2015.
- [6] 田壮壮, 占荣辉, 胡杰民, 等. 基于卷积神经网络的SAR图像目标识别研究[J]. 雷达学报, 2016, 5(3): 320–325. doi: [10.12000/JR16037](https://doi.org/10.12000/JR16037).  
TIAN Zhuangzhuang, ZHAN Ronghui, HU Jiemin, *et al.* SAR ATR based on convolutional neural network[J]. *Journal of Radars*, 2016, 5(3): 320–325. doi: [10.12000/JR16037](https://doi.org/10.12000/JR16037).
- [7] CHEN Sizhe, WANG Haipeng, XU Feng, *et al.* Target classification using the deep convolutional networks for SAR images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(8): 4806–4817. doi: [10.1109/TGRS.2016.2551720](https://doi.org/10.1109/TGRS.2016.2551720).
- [8] FURUKAWA H. Deep learning for target classification from SAR imagery: Data augmentation and translation invariance[R]. SANE2017-30, 2017.
- [9] 袁媛, 袁昊, 雷玲, 等. 一种同步轨道星机双基SAR成像方法[J]. 雷达科学与技术, 2007, 5(2): 128–132. doi: [10.3969/j.issn.1672-2337.2007.02.011](https://doi.org/10.3969/j.issn.1672-2337.2007.02.011).  
YUAN Yuan, YUAN Hao, LEI Ling, *et al.* An imaging method of GEO Spaceborne-Airborne Bistatic SAR[J]. *Radar Science and Technology*, 2007, 5(2): 128–132. doi: [10.3969/j.issn.1672-2337.2007.02.011](https://doi.org/10.3969/j.issn.1672-2337.2007.02.011).
- [10] 史洪印, 周荫清, 陈杰. 同步轨道星机双基地三通道SAR地面运动目标指示算法[J]. 电子与信息学报, 2009, 31(8): 1881–1885.  
SHI Hongyin, ZHOU Yinqing, and CHEN Jie. An algorithm of GEO spaceborne-airborne bistatic three-channel SAR ground moving target indication[J]. *Journal of Electronics & Information Technology*, 2009, 31(8): 1881–1885.
- [11] LI Zhuo, LI Chunsheng, YU Ze, *et al.* Back projection algorithm for high resolution GEO-SAR image formation[C]. 2011 IEEE International Geoscience and Remote Sensing Symposium, Vancouver, Canada, 2011: 336–339.
- [12] ZHANG Fan, HU Chen, YIN Qiang, *et al.* Multi-aspect-aware bidirectional LSTM networks for synthetic aperture radar target recognition[J]. *IEEE Access*, 2017, 5: 26880–26891. doi: [10.1109/ACCESS.2017.2773363](https://doi.org/10.1109/ACCESS.2017.2773363).
- [13] PEI Jifang, HUANG Yulin, HUO Weibo, *et al.* SAR automatic target recognition based on Multiview deep learning framework[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(4): 2196–2210. doi: [10.1109/TGRS.2017.2776357](https://doi.org/10.1109/TGRS.2017.2776357).
- [14] 邹浩, 林赞, 洪文. 采用深度学习的多方位角SAR图像目标识别研究[J]. 信号处理, 2018, 34(5): 513–522. doi: [10.16798/j.issn.1003-0530.2018.05.002](https://doi.org/10.16798/j.issn.1003-0530.2018.05.002).  
ZOU Hao, LIN Yun, and HONG Wen. Research on multi-aspect SAR images target recognition using deep learning[J]. *Journal of Signal Processing*, 2018, 34(5): 513–522. doi: [10.16798/j.issn.1003-0530.2018.05.002](https://doi.org/10.16798/j.issn.1003-0530.2018.05.002).
- [15] ZHAO Pengfei, LIU Kai, ZOU Hao, *et al.* Multi-stream convolutional neural network for SAR automatic target recognition[J]. *Remote Sensing*, 2018, 10(9): 1473. doi: [10.3390/rs10091473](https://doi.org/10.3390/rs10091473).
- [16] ZHANG Fan, FU Zhenzhen, ZHOU Yongsheng, *et al.* Multi-aspect SAR target recognition based on space-fixed and space-varying scattering feature joint learning[J]. *Remote Sensing Letters*, 2019, 10(10): 998–1007. doi: [10.1080/2150704X.2019.1635287](https://doi.org/10.1080/2150704X.2019.1635287).
- [17] TAN Mingxing and LE Q V. EfficientNet: Rethinking model scaling for convolutional neural networks[J]. arXiv: 1905.11946, 2019.
- [18] CHO K, VAN MERRIENBOER B, GULCEHRE C, *et al.* Learning phrase representations using RNN encoder-decoder for statistical machine translation[J]. arXiv: 1406.1078, 2014.
- [19] CAI Jie, MENG Zibo, KHAN A S, *et al.* Island loss for learning discriminative features in facial expression recognition[C]. The 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 2018: 302–309.
- [20] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Deep

residual learning for image recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 2016.

- [21] HOCHREITER S and SCHMIDHUBER J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735–1780. doi:

[10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735).

- [22] WEN Yandong, ZHANG Kaipeng, LI Zhifeng, *et al.* A discriminative feature learning approach for deep face recognition[C]. The 14th European Conference on Computer Vision – ECCV 2016, Amsterdam, The Netherlands, 2016.

### 作 者 简 介



赵鹏飞(1996–), 男, 硕士生, 研究方向为合成孔径雷达图像分析。



黄丽佳(1984–), 女, 博士, 研究员, 硕士生导师, 研究方向为合成孔径雷达信号处理与图像分析。