

# 基于轻量级残差网络的红外遥感船只检测

朱天佑<sup>1,2,3</sup>, 黄凌锋<sup>1,2,3</sup>, 董峰<sup>1,2</sup>, 龚惠兴<sup>1,2\*</sup>

<sup>1</sup>中国科学院红外探测与成像技术重点实验室, 上海 200083;

<sup>2</sup>中国科学院上海技术物理研究所, 上海 200083;

<sup>3</sup>中国科学院大学, 北京 100049

**摘要** 针对红外遥感船只检测领域存在的硬件存储资源和功耗的限制,以及目标检测输出边界矩形框形式结果不够精细的问题,提出了一种轻量化且具有像素级输出的分割网络 TRS-Net。将图像分割的编码-解码结构用于船只检测,以获得像素级的输出;将 32 bit 的浮点型参数二值化(目的是压缩网络模型的大小),提出了 BS-Net;针对 BS-Net 带来的检测精度低的问题,引入残差连接,提出了 BRS-Net;根据神经网络稀疏性特点引入参数三元化,提出了 TS-Net;为进一步提升检测效果,将 TS-Net 改进成 TRS-Net。采用实验室自主研制的长波红外相机进行成像实验,获取红外船只图片并制作数据集,对 4 种网络的结果进行对比分析。结果表明:TRS-Net 检测的精确率为 88.73%,召回率为 83.34%, $F_{1\text{-score}}$  为 85.95%,交并比为 75.36%,模型大小压缩为原先的 1/16。TRS-Net 对红外船只的实时检测具有一定的工程应用价值。

**关键词** 成像系统; 红外船只检测; 二值网络; 编码-解码结构; 三元残差网络

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/AOS202040.0111018

## Infrared-Remote-Sensing Ship Detection Based on Lightweight Residual Network

Zhu Tianyou<sup>1,2,3</sup>, Huang Lingfeng<sup>1,2,3</sup>, Dong Feng<sup>1,2</sup>, Gong Huixing<sup>1,2\*</sup>

<sup>1</sup>Key Laboratory of Infrared System Detection and Imaging Technology, Chinese Academy of Sciences, Shanghai 200083, China;

<sup>2</sup>Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;

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

**Abstract** To address the limitations of hardware storage resource and power consumption in infrared-remote-sensing ship detection and the inadequate precision of the output boundary rectangular box form of target detection, a lightweight and pixel-level-output segmentation network TRS-Net (ternary residual segmentation network) is proposed. We apply the encoder-decoder structure of image segmentation to ship detection to obtain the pixel-level output. Further, we binarize the 32-bit floating-point parameters to compress the size of the network model and propose a binary segmentation network (BS-net). Then, to solve the problem of poor detection accuracy caused by BS-Net, we introduce residual connection and propose a binary residual segmentation network (BRS-Net). Furthermore, owing to the sparsity of the neural network, we introduce ternary parameters and propose a ternary segmentation network (TS-Net); therefore, we propose a ternary residual segmentation network (TRS-Net) to further improve the detection effect. Using a long-wave infrared camera independently developed by the laboratory for imaging experiments, we obtain infrared images of ships, make the datasets, and compare and analyze the results of four kinds of networks. The results demonstrate that the detection precision, recall rate,  $F_{1\text{-score}}$ , and intersection-over-union of TRS-Net are 88.73%, 83.34%, 85.95%, and 75.36%, respectively. Furthermore, the model size is reduced to one-sixteenth of its original size. Therefore, the proposed TRS-Net has practical engineering value for real-time infrared ship detection.

**Key words** imaging systems; infrared ship detection; binary network; encoder-decoder structure; ternary residual network

**OCIS codes** 110.2960; 120.0280; 150.1135; 100.3008

收稿日期: 2019-07-26; 修回日期: 2019-08-27; 录用日期: 2019-09-09

基金项目: 十三五装备预研中科院联合基金(61XXA011XXXX)

\* E-mail: hxgong@mail.sitp.ac.cn

# 1 引 言

随着经济社会的发展,人类的海上活动越来越频繁,对海上大量船只的监控成为难题,尤其是完全被动条件下的监控。因此,利用天基遥感手段对船只进行监控在海上运输、海战态势感知<sup>[1]</sup>等领域具有重要意义<sup>[2]</sup>,其中,对船只进行高效检测是关键<sup>[3]</sup>。为加快检测速度和提高检测精度,本文对深度神经网络方法在红外遥感船只检测方面的应用进行了研究。

深度神经网络是近年来的研究热点,在遥感舰船检测中具有独特的应用价值,它通过多层网络获取不同层次的特征信息,具有超高的特征描述能力,可以提取目标更复杂的特征<sup>[4]</sup>。Li 等<sup>[5]</sup>提出的 HSF-Net(hierarchical selective filtering network)由 fast R-CNN (region-convolutional neural network)网络改进而来,可以实现端到端的可见光遥感船只检测,并以边界矩形框(bounding box)的形式输出结果。Cheng 等<sup>[6]</sup>针对遥感船只等检测的多方向问题,提出了一种 RICNN(rotation-invariant CNN)网络,有效提升了识别精度,但对计算资源的需求也更高。虽然现有的深度学习船只检测方法不断向前发展,但仍存在一定问题:1)网络结果均以边界矩形框的形式输出,不够精细,船只区域的表达方式有待改进;2)这些网络使用 32 bit 精度的参数,给低功耗平台带来巨大压力。

为了提高对船只检测结果的表达能力,本文采用了编码(encoder)-解码(decoder)网络结构。编码-解码结构的网络在目标识别方面发展迅速,且主要应用在语义分割中,如 U-Net<sup>[7]</sup>、SegNet<sup>[8]</sup>等,输出像素级的检测结果。相比于边界矩形框,编码-解码结构不仅可以识别目标,还能分割出目标所在像素的位置,更符合视觉要求。编码-解码结构在遥感船只目标识别中的相关研究较少,是值得探索的重点方向。为了降低硬件设备存储资源、功耗方面的压力,本文对编码-解码结构网络的轻量化(二值化、三元化)方法进行了探索。Courbariaux 等<sup>[9]</sup>在 2015 年提出的二值连接将传统的卷积神经网络中的权重进行二值化处理,Hubara 等<sup>[10]</sup>在二值连接的基础上,将权重和特征图全部二值化,实现了网络参数的高效压缩,但这样会导致检测精度下降。Li 等<sup>[11]</sup>实现了网络权重的三元化,以在压缩的同时改善检测效果。除了权重量化,基于改变卷积核的轻量化方法同样有很多应用,如分组卷积

( ShuffleNet<sup>[12]</sup>、 MobileNet<sup>[13]</sup> )、分解卷积(ErfNet<sup>[14]</sup>)等。

基于此,针对船只目标检测中的问题,本文结合编码-解码结构与参数轻量化优势,提出了一种基于语义分割的轻量级网络方法 BS-Net (binary segmentation network),输出的是像素级的检测结果;为了提高检测精度,引入残差连接,提出 BRS-Net(binary residual segmentation network);依据神经网络稀疏性特点引入参数三元化,提出 TS-Net (ternary segmentation network),并进一步提出了 TRS-Net(ternary residual segmentation network)。

鉴于红外可全天时工作的特点及其在复杂海况下的优良适用性,本文对红外遥感舰船目标检测进行了重点探索。采用实验室自主研制的红外采集设备进行成像实验,获取红外图像,制作数据集,并对本文网络算法进行对比。

## 2 基本原理

本节主要包括三部分:1)轻量级网络结构设计;2)根据层级结构与运算方法对二值网络实现过程中的前向传输、反向传输及损失函数进行详细说明;3)二值网络的改进措施,包括引入残差连接和三元化。

### 2.1 船只检测分割网络的搭建

图 1 中,encoder 部分实现船只特征的提取,decoder 部分实现像素级船只识别结果的表达。在复制和裁剪过程中,将 encoder 中的位置引入 decoder,改善船只分割的分辨率。这样的结构为本文船只检测带来两大优势:一是相比传统的神经网络,本文的船只检测网络结构无全连接层(全连接层一般占据整个网络约 80%资源),可有效实现参数量的压缩;二是编码-解码结构的输出可以直接输出像素级的分割结果,实现船只检测结果的清晰表达。网络中的 encoder 部分包括残差卷积模块(Res-Conv2D Block)、池化(pooling)操作,Res-Conv2D Block 涉及卷积、Batch Normalization(BN,批量标准化)、ReLU(激活)、残差连接等,如图 2 所示。Decoder 过程与 encoder 部分的运算一致,upooling 部分也继承 encoder 部分 pooling 的位置信息,可以补充更多的特征信息,输出更精确的识别结果。

### 2.2 轻量级网络中函数的选取

#### 1) 前向传输

轻量级网络在精度上与 32 位浮点数的精度越来越接近,而在硬件资源消耗上却占据着巨大优势,

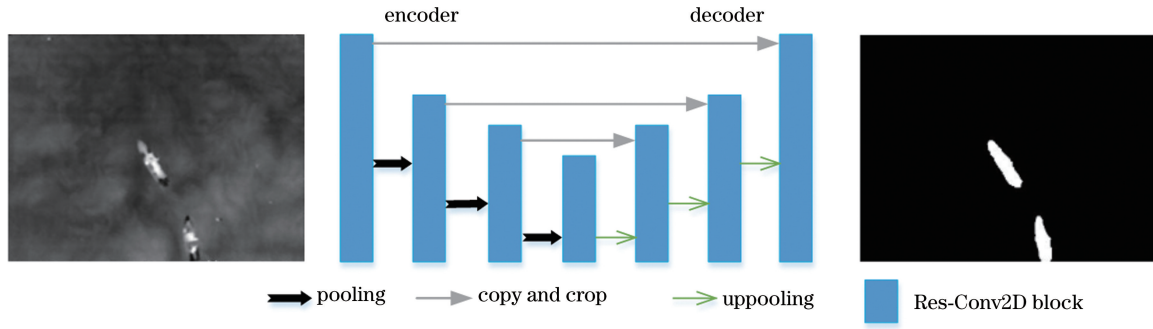


图 1 船只检测网络结构图

Fig. 1 Architecture of ship detection network

具有强大的吸引力及不可估量的经济价值。对网络参数存储的压缩存在极限,即 1 bit,可表征两种状态(二值化)。

在实现网络参数的二值化方面,考虑到天基遥感应用场景的实时性要求,本文采用决定式方法,即采用 sign 函数判别运算过程的输出结果,权重  $W$  的二值化过程可表示为

$$W_b = \text{sign}(W) = \begin{cases} +1, & \text{if } W \geq 0 \\ -1, & \text{otherwise} \end{cases} \quad (1)$$

二值网络运算过程中的 feature map 及权重均为二值结果,但图像数据不是二值的,需要进行特殊的处理,将原始图像转为 8 bit 定点数,权重参数是 1 bit。为了实现 8 bit 的还原,像素值用  $x^n$  表示。运算过程实现  $n$  位与权重的单独相乘后相加,累加输出结果为  $s$ ,具体为

$$s = \sum_{n=1}^8 2^{n-1} (x^n \cdot W_b) \quad (2)$$

二值化网络中的 BN 可以解决协方差偏移,采用 2 的次方和移位实现乘除操作,即

$$r = \frac{x - \mu}{\sqrt{\sigma^2 + \epsilon}} \gamma + b, \quad (3)$$

式中: $r$  为输出结果; $x$  为输入数据; $\mu$  为均值; $\sigma^2$  为方差; $\gamma$  为尺度因子; $b$  为偏置; $\epsilon$  是为防止出现分母为零而设置的一个很小的值。BN 在实现网络加速的同时,引入的噪声也可以起到正则化的作用,可在一定程度上防止模型过拟合。本实验在编码及解码过程中均保留 BN 操作。

在网络的池化层采用最大池化的方法对特征图  $u_j^l$  进行池化操作得到  $p_j^l$ ,如此可在实现特征降维的同时在一定程度上保持一定的平移不变性,表示为

$$p_j^l(x, y) = \max\{u_j^l(p, q) \mid xM \leq p \leq xM + K - 1, yM \leq q \leq yM + K - 1\}, \quad (4)$$

式中: $l$  为特征图的序号; $j$  为池化块的序号; $K$  表示设置的邻域大小; $M$  为池化窗口的步长,通常小

于  $K$ 。

### 2) 误差反向传递

为了减小反向传播过程二值化引起的误差,除了二值化权重值外,还保留了浮点型权重值。浮点值可以更精确地用于更新二值权重。针对二值化网络反向传播的不可导问题(反向传播需链式求导,而 sign 函数不可导),采用放缩松弛的方法,即

$$g_w = \begin{cases} g_{w_b}, & |W| \leq 1 \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

式中: $g_w$  为浮点权重的梯度; $g_{w_b}$  为二值权重的梯度。当权重  $W$  的绝对值小于 1 时, $W$  的梯度与  $W_b$  的梯度相同;当权重  $W$  的绝对值大于 1 时,梯度为 0,可以实现将  $[-1, 1]$  之外的数据裁减掉的功能。接下来,计算浮点型权重梯度与激活函数的残差,采用梯度下降法实现权重的更新。

### 3) 损失函数

损失函数对神经网络性能的影响较大,网络训练过程的本质是对损失函数进行优化。深度网络中应用最广的是 MSE 损失函数。MSE 表示预测结果与标准结果间的均方误差,MSE 越小,表明预测结果越精确,越接近理想的结果。

$$C = \frac{(y - a)^2}{2}, \quad (6)$$

式中: $y$  为期望的输出结果; $a$  为实际输出; $C$  为损失值。设  $w$  为单层权重, $b$  为偏置, $z$  为中间层激活值,用  $\sigma(\cdot)$  表示激活函数,依据损失值和激活函数,基于链式求导,实现  $w$  和  $b$  的更新。

实验网络最后预测的输出层采用 Sigmoid 激活函数  $f(x) = 1/(1 + e^{-x})$ ,以对两类结果进行判别。在  $z$  值较大时会出现  $\sigma'(z)$  很小的情况,这使得  $w$  和  $b$  的更新非常慢。在变化相同的  $\Delta x$  的情况下,针对 Sigmoid 函数随着  $x$  不断增大而变化越来越小的缺陷,采用交叉熵损失函数实现损失函数与

Sigmoid 函数的有效结合。

$$C = -y \cdot \log_2(a) - (1 - y) \cdot \log_2(1 - a). \quad (7)$$

由(7)式可以看出,在交叉熵损失函数中,当  $y=1$  时,  $a$  越大,  $a$  与  $y$  就越接近,预测就更加准确,损失函数就越小。

优化方法选取随机梯度下降(SGD)法。SGD 法随机选取一批数据样本完成迭代,所在批次样本均参与一次迭代,即实现一个迭代周期。数据批较大时,可以得到更为精确的梯度估计,小批量数据有助于更快收敛。数据批的大小可以根据需要设置合理的数值:

$$w_i^{k+1} = w_i^k - \alpha \cdot G(w_i^k), \quad (8)$$

式中:  $\alpha$  为学习率,取值范围为  $0 \sim 1$ ;  $i$  为权重参数的序号;  $w_i^k$  为第  $k$  次迭代第  $i$  个权重参数;  $G(w_i^k)$  为权重参数  $w_i^k$  对应的梯度值。

### 2.3 轻量级网络的改进方法

#### 1) 三元化方法

神经网络的权重参数具有稀疏性<sup>[15]</sup>,其数值大部分为 0 或接近 0,服从正态分布,因此对权重强制二值化时带来的误差较大,导致精度损失严重。参

数三元化就是将绝对值小于阈值的权重归为 0,其余权重依然归为 +1 和 -1,即

$$W_T = \begin{cases} +1, & \text{if } W > \Delta \\ 0, & \text{if } |W| \leq \Delta \\ -1, & \text{if } W < -\Delta \end{cases}, \quad (9)$$

式中:  $\Delta$  是阈值。通常取所有权重绝对值和的平均值  $E(|W|)$  的 0.7,即

$$\Delta = 0.7 \cdot E(|W|) = \frac{0.7}{n} \sum_{i=1}^{n'} |W_i|, \quad (10)$$

式中  $n'$  为权重参数的个数。

三元化参数的存储比二值化多 1 bit,其中 1 bit 用以表征是否稀疏,其余 1 bit 仍表征 +1 和 -1,因此在运行效率上与二值化相当,并且能够使神经网络权重稀疏化,在减小计算复杂度的同时也能得到更高的准确率和泛化能力。

#### 2) 残差连接模块

随着网络层的加深,会出现梯度消失或爆炸现象,导致深层网络的准确度相较于浅层网络反而下降。残差连接<sup>[16]</sup>可以把特征图以及所包含的信息无损地传递到下一层,保证网络性能随着深度加深不会变差,同时能加快网络训练,结构如图 2 所示。

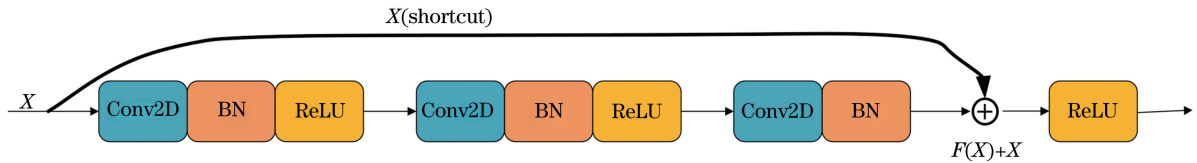


图 2 残差连接模块(Res-Conv2D)

Fig. 2 Block of residual connection (Res-Conv2D)

图 2 中,  $X$  表示网络层的输入,通向两部分:一是直连的恒等映射,跳过两层把值送到后层的激活(ReLU)前;二是正常经过几层权重。假设经几层网络的输出映射是  $H(X)$ ,则  $F(X)$  相当于残差函数,即

$$F(X) = H(X) - X. \quad (11)$$

因此,新层的权重只需学习残差,这样就减小了网络学习的难度。增加残差连接能够在无需新增权重参数的基础上提高精度,且残差网络更容易优化,能够通过增加相当的深度来提高准确率。

## 3 实验与结果

### 3.1 实验场景设计与数据采集

采用本实验室自主研制的长波红外相机作为采集设备,利用水槽和船只模型进行海上船只和海上环境的模拟,设计实验中的场景如图 3 所示。

采集 35 次成像实验数据,形成原始船只图片万余张。在数据量满足训练要求的情况下,数据集标签的制作是必不可少的内容。在数据集制作过程中,要实现编码-解码结构的像素级输出,制备标签会耗费大量的时间。在人工标记过程中,红外图像中的目标边缘不够清晰,船只目标偏小,图片均需放大足够的倍数以获取足够的边界点信息。每张图片耗费 1 min 左右的时间,1 万张将耗费约 167 h,人力成本消耗相当大。在确保标记精度的前提下实现快速标注也是重点探索的方向。

针对上述问题,本实验在标注过程探索出了一种新的像素级标签制作方法。首先,选取 1000 张图片制作初始样本数据集,用于进行网络训练集的测试任务,将剩余的近 9000 张数据用于测试,得到输出结果;然后,对输出结果中满足人眼要求的高精度

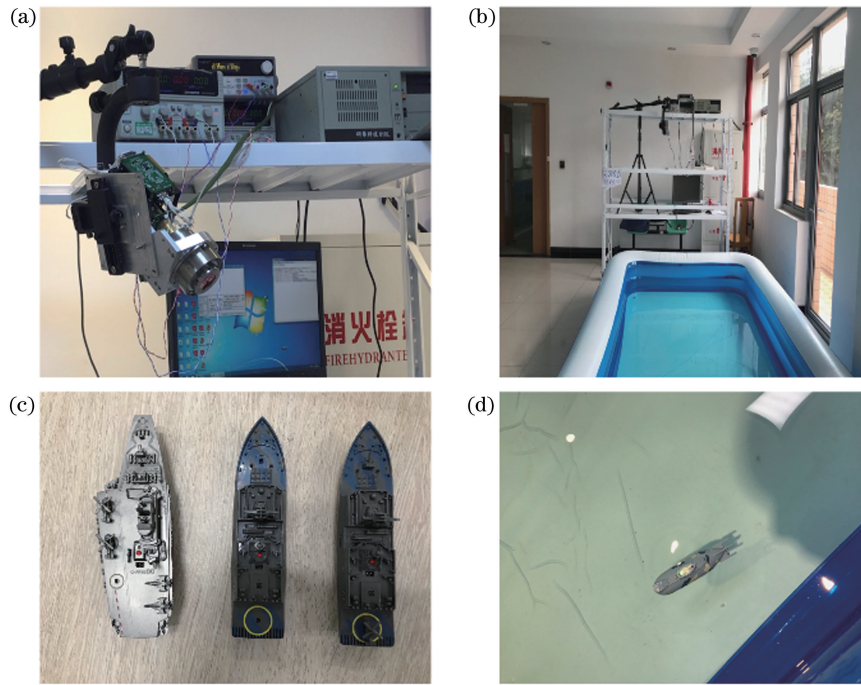


图 3 船只模型和红外成像系统。(a)红外成像系统;(b)水槽;(c)船模;(d)水面船只  
Fig. 3 Ship models and infrared imaging system. (a) Infrared imaging system; (b) flume;  
(c) ship models; (d) surface ship

识别结果的边界进行手动调整,以扩充数据集。采用这样的操作后可以获取近千余张满足标签要求的样本,经过十余次网络训练与优化,以及最终的手工标注调整,就可有效实现数据集的扩充,得到完整的

目标分割像素级标注结果。

在初始样本数据集的制作中,采用 LabelMe 对区域进行标记,利用其多点选取的优势实现像素级的分割。标签制作前后的示例如图 4 所示。

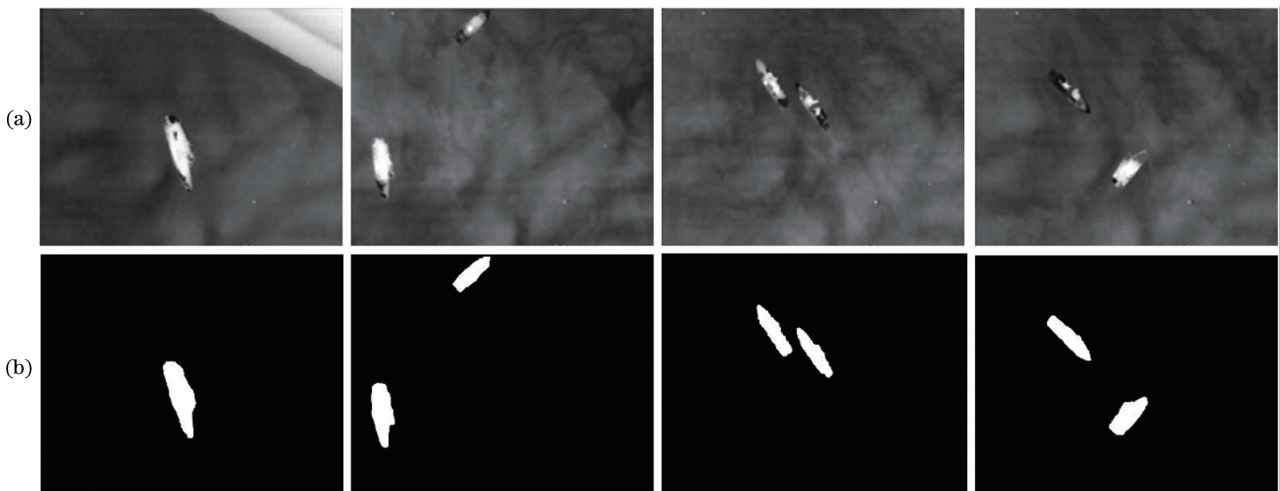


图 4 数据集制作示例。(a)原始数据;(b)用 LabelMe 标记过的数据  
Fig. 4 Samples of ship dataset production. (a) Original data; (b) data labeled by LabelMe

在网络训练过程中,参考经验按照 8:1:1 的规则拆分<sup>[17]</sup>数据集,即:将 80% 的图片作为训练集,10% 作为测试集,10% 作为交叉验证集。为了减少运算量,将数据集中原始图像和标签统一处理为  $256 \times 256$  的图像。

### 3.2 训练与平台

在完成网络结构设计及优化、数据获取与数据集制作之后,进行网络训练实验。根据现有条件,网络训练与测试平台基于 NVIDIA GTX1080 GPU, 64 位 Ubuntu16.04 LTS 系统。训练代码框架基于

TensorFlow, 其余的图像数据处理方法采用 OpenCV3.2.0 实现。

训练是获取网络各项参数的有效途径,在训练过程中,通过多次迭代更新网络参数,使损失函数取得全局极小值。其中,网络初始参数采用随机生成的方式,实验中,每层滤波器大小均设为  $3 \times 3$ ,池化尺寸为  $2 \times 2$ ,滑窗步长均为 1。梯度下降法的初始学习率设置为  $\alpha=0.001$ ,每 10 个迭代周期对学习率进行一次减半操作,训练过程中交叉熵损失值的变化、像素准确率的变化分别如图 5、6 所示。

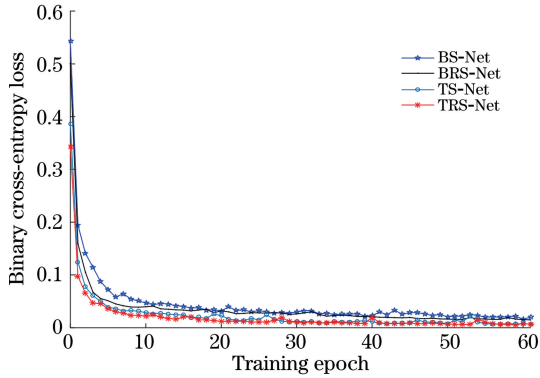


图 5 4 种网络在训练过程中交叉熵损失值的变化  
Fig. 5 Variation of binary cross-entropy loss in training processing for four networks

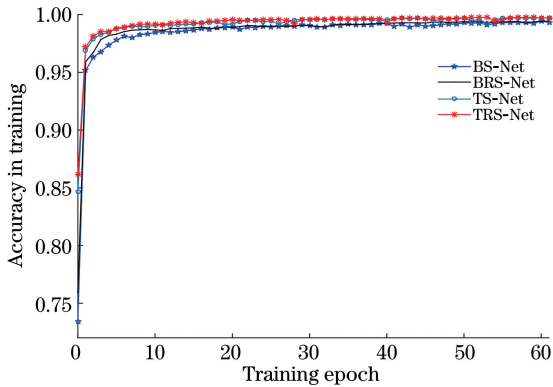


图 6 4 种网络在训练过程中像素准确率的变化  
Fig. 6 Variation of pixel accuracy in training processing for four networks

### 3.3 实验结果评价

为了更好地对实验结果进行评价,分别对分割图像视觉结果及定量的分割指标进行详细分析。实验选取目标识别及语义分割中的经典评价指标对识别结果进行分析,所用评价指标包括像素准确率  $\beta_{PA}$ 、精度  $P$ 、召回率  $R$ ,以及综合评价指标  $F_{1-score}$  (balanced  $F$  score)、 $\beta_{IoU}$  (Intersection over union)。评价指标的计算公式为

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (12)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (13)$$

$$\beta_{PA} = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}}, \quad (14)$$

$$F_{1-score} = \left( \frac{R^{-1} + P^{-1}}{2} \right)^{-1} = 2 \cdot \frac{P \cdot R}{P + R}, \quad (15)$$

$$\beta_{IoU} = (A \cap B) / (A \cup B), \quad (16)$$

式中:  $N_{TP}$  为正确识别出船只目标像素的个数;  $N_{TN}$  为正确识别出背景像素的个数;  $N_{FP}$  为将背景误识为船只像素的个数;  $N_{FN}$  为将船只误识成背景像素的个数;  $A$  为预测船只目标像素集合;  $B$  为真实船只目标像素所在区域。

$\beta_{PA}$  即像素准确率,表示识别正确的像素(包括船只目标和背景)占有所有像素的比例;  $P$  即精确率,表示所有被分类为船只目标的像素中真实的船只像素所占比例;  $R$  即召回率,表示在所有的真实船只像素中被正确分类的比例。通常希望同时获得较高的精确率和召回率,但这两者在某些情况下是互相矛盾的,因此需要综合考虑它们,常见的方法是取两者的调和平均数,即  $F_{1-score}$ 。

测试集图片上的视觉检测效果如图 7 所示,可知:二值网络 BS-Net 的检测精度较差;引入残差连接的 BRS-Net 和三元化 TS-Net 网络的检测结果得到了明显改进;三元残差网络 TRS-Net 的视觉效果更好,船只像素级的区域得到了有效检测,单帧中的非单一目标也得到了有效检测,并且红外图像右下角人工漏标注的船只也被有效地检测出来,可用来有效地辅助新数据的标注。

表 1 测试集中 4 种网络对船只检测结果的对比

Table 1 Comparison of ship detection results of four networks on testing datasets %

Network	$\beta_{PA}$	$P$	$R$	$F_{1-score}$	$\beta_{IoU}$
BS-Net	95.76	34.46	75.27	47.27	30.95
BRS-Net	98.70	84.78	59.31	69.79	53.60
TS-Net	99.18	93.46	72.60	81.72	69.09
TRS-Net	99.31	88.73	83.34	85.95	75.36

由表 1 可知:虽然二值分割网络 BS-Net 的模型大小是浮点型的  $1/32$ ,但其检测效果较差,  $F_{1-score}$  仅为 47.27%,  $\beta_{IoU}$  仅为 30.95%;引入残差连接后的 BRS-Net 的  $F_{1-score}$  较 BS-Net 提高了 22.52 个百分点,  $\beta_{IoU}$  提高了 22.65 个百分点,表明残差结构能有效提高网络性能;引入三元化后的 TS-Net 比 BS-

Net 的  $F_{1-score}$  提高了 34.45 个百分点,  $\beta_{IoU}$  提高了 38.14 个百分点, 说明三元化更符合神经网络权重的正态分布特性和稀疏性, 而二值化带来的量化误差过大; 三元残差网络 TRS-Net 的检测精度更高,

$F_{1-score}$  为 85.95%,  $\beta_{IoU}$  为 75.36%, 模型大小压缩为原来的 1/16。综合模型的压缩尺寸和检测效果可以看出, 与其他网络相比, TRS-Net 具有更重要的实际工程应用价值。

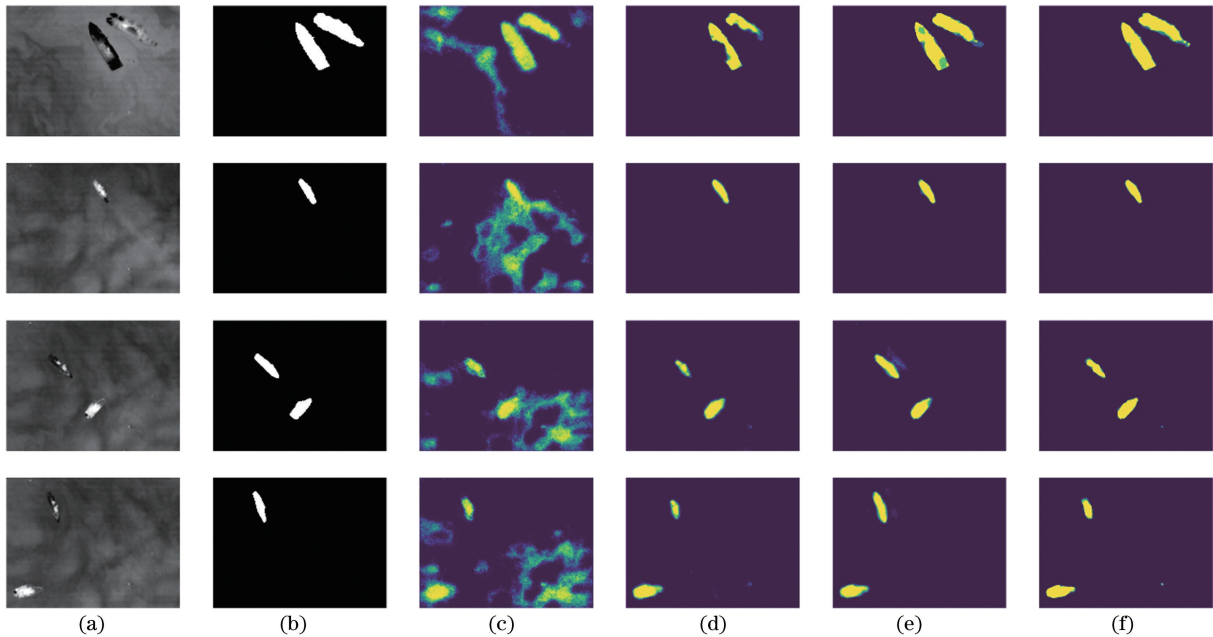


图 7 输出结果对比。(a)输入图像;(b)人工标注结果;(c) BS-Net 输出结果;(d) BRS-Net 输出结果;(e) TS-Net 输出结果;(f) TRS-Net 输出结果

Fig. 7 Output results. (a) Input images; (b) output results of manual annotation; (c) output results of BS-Net; (d) output results of BRS-Net; (e) output results of TS-Net; (f) output results of TRS-Net

## 4 结 论

相比可见光遥感船只检测, 红外船只检测受大气、光照、云雾和海岛等海面不确定条件的影响较小, 且具有昼夜观测不受限的优点, 因此成为重点研究方向。本实验室针对红外遥感船只公开数据集缺少的现状, 采用自主研发的长波红外相机进行成像实验, 并获取、制作数据集; 针对目标检测矩形框形式输出结果不够精细的问题, 采用分割方法输出像素级的结果; 针对硬件资源和功耗的限制, 将 32 bit 的浮点型参数二值化, 以减少存储空间, 提出二值分割网络 BS-Net, 使得模型压缩为原来的 1/32。针对 BS-Net 检测输出效果差的问题 ( $F_{1-score}$  为 47.27%,  $\beta_{IoU}$  为 30.95%), 引入残差连接, 提出了 BRS-Net。与 BS-Net 相比, BRS-Net 可有效提高检测精度,  $F_{1-score}$  提高了 22.52 个百分点,  $\beta_{IoU}$  提高了 22.65% 个百分点。针对神经网络参数稀疏性的特点, 引入三元化, 提出了 TS-Net。最后结合残差和三元化两种手段将 TS-Net 改进成 TRS-Net 网络, 进一步提升了检测效果。TRS-Net 的检测精确率为 88.73%,

准确率为 83.34%,  $F_{1-score}$  为 85.95%, 交并比 ( $\beta_{IoU}$ ) 为 75.36%, 模型大小压缩为原来的 1/16。

## 参 考 文 献

- [1] Bao S Z, Zhong X, Zhu R F, et al. Automatic detection method of ships based on shortwave infrared remote sensing images[J]. Acta Optica Sinica, 2018, 38(5): 0528001.  
鲍松泽, 钟兴, 朱瑞飞, 等. 基于短波红外遥感影像的船只自动检测方法[J]. 光学学报, 2018, 38(5): 0528001.
- [2] Wang W X, Fu Y T, Dong F, et al. Infrared ship target detection method based on deep convolution neural network [J]. Acta Optica Sinica, 2018, 38(7): 0712006.  
王文秀, 傅雨田, 董峰, 等. 基于深度卷积神经网络的红外船只目标检测方法[J]. 光学学报, 2018, 38(7): 0712006.
- [3] Jiang B T, Ma X F, Lu Y, et al. Ship detection in spaceborne infrared images based on convolutional neural networks and synthetic targets [J]. Infrared Physics & Technology, 2019, 97: 229-234.
- [4] Liu F, Shen T S, Ma X X. Convolutional neural

- network based multi-band ship target recognition with feature fusion[J]. *Acta Optica Sinica*, 2017, 37(10): 1015002.
- 刘峰, 沈同圣, 马新星. 特征融合的卷积神经网络多波段舰船目标识别[J]. *光学学报*, 2017, 37(10): 1015002.
- [5] Li Q P, Mou L C, Liu Q J, et al. HSF-Net: multiscale deep feature embedding for ship detection in optical remote sensing imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(12): 7147-7161.
- [6] Cheng G, Zhou P C, Han J W. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(12): 7405-7415.
- [7] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W, et al. *Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science*. Cham: Springer, 2015, 9351: 234-241.
- [8] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [9] Courbariaux M, Bengio Y, David J P. BinaryConnect: training deep neural networks with binary weights during propagations[C]//*Proceedings of the 28th International Conference on Neural Information Processing Systems*, December 7-12, 2015, Montreal, Canada. Canada: NIPS, 2015: 3123-3131.
- [10] Hubara I, Courbariaux M, Soudry D, et al. Binarized neural networks [C]//*Advances in Neural Information Processing Systems 29 (NIPS 2016)*, December 5-10, 2016, Barcelona, Spain. Canada: NIPS, 2016: 4107-4115.
- [11] Li F F, Zhang B, Liu B. Ternary weight networks [J/OL]. (2016-11-19) [2019-07-25]. <https://arxiv.org/abs/1605.04711>.
- [12] Zhang X Y, Zhou X Y, Lin M X, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C]//*2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 6848-6856.
- [13] Howard A G, Zhu M L, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [J]. (2017-04-17) [2019-07-25]. <https://arxiv.org/abs/1704.04861>.
- [14] Romera E, Alvarez J M, Bergasa L M, et al. ERFNet: efficient residual factorized ConvNet for real-time semantic segmentation [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 19(1): 263-272.
- [15] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks [C]//*Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, April 11-13, 2011, Fort Lauderdale, USA. USA: MIT Press, 2011: 315-323.
- [16] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C]//*2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [17] Brownlee J. Tactics to combat imbalanced classes in your machine learning dataset [DB/OL]//*Machine learning mastery*, 2015-08-19 [2019-07-25]. <https://machinelearningmastery.com/tactics-to-combat-imbalanced-classes-in-your-machine-learning-dataset/>.