

## 一种改进的 Capsule 及其在 SAR 图像目标识别中的应用

张盼盼<sup>1,2,3,4,5</sup>, 罗海波<sup>1,2,4,5\*</sup>, 鞠默然<sup>1,2,3,4,5</sup>, 惠斌<sup>1,2,4,5</sup>, 常铮<sup>1,2,4,5</sup>

- (1. 中国科学院沈阳自动化研究所, 辽宁 沈阳 110016;
2. 中国科学院机器人与智能制造创新研究院, 辽宁 沈阳 110169;
3. 中国科学院大学, 北京 100049;
4. 中国科学院光电信息处理重点实验室, 辽宁 沈阳 110016;
5. 辽宁省图像理解与视觉计算重点实验室, 辽宁 沈阳 110016)

**摘要:** 为了解决 Capsule 网络随着输入图像增大计算量和参数数量急剧增加的问题, 对 Capsule 网络进行了改进并将其用于 SAR 自动目标识别 (SAR-ATR) 中。基于大脑视觉皮层以层级结构以及柱状形式处理信息的机制, 提出了完全实例化的思想, 并运用类脑计算对 Capsule 网络进行了改进。具体方法是: 使用多个卷积层实现层级处理, 同时使用了较少的卷积核, 但每一层使用的卷积核数量随着层级加深逐渐增加, 使得提取的特征更加趋于抽象化; 在 PrimaryCaps 层中, Capsule 向量由最后一层卷积层输出的所有特征图构成, 使得 Capsule 单元包含目标局部或整体的全部特征, 以实现目标的完全实例化。在 SAR-ATR 上, 将改进的 Capsule 网络与原 Capsule 网络、传统目标识别算法和基于经典卷积神经网络的目标识别算法进行对比实验。实验结果表明, 改进的 Capsule 网络训练参数和计算量大大减少, 并且训练速度得到很大提升, 在 SAR 图像数据集上的识别准确率较 Capsule 网络和前两类方法分别提高了 0.37 和 1.96~8.96 个百分点。

**关键词:** 目标识别; Capsule 网络; 完全实例化; 类脑计算; 卷积神经网络

**中图分类号:** TP391 **文献标志码:** A **DOI:** 10.3788/IRLA20201010

## An improved Capsule and its application in target recognition of SAR images

Zhang Panpan<sup>1,2,3,4,5</sup>, Luo Haibo<sup>1,2,4,5\*</sup>, Ju Moran<sup>1,2,3,4,5</sup>, Hui Bin<sup>1,2,4,5</sup>, Chang Zheng<sup>1,2,4,5</sup>

- (1. Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China;
2. Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China;
3. University of Chinese Academy of Sciences, Beijing 100049, China;
4. Key Laboratory of Opto-Electronic Information Processing, Chinese Academy of Sciences, Shenyang 110016, China;
5. The Key Lab of Image Understanding and Computer Vision, Liaoning Province, Shenyang 110016, China)

**Abstract:** In order to solve the problem that the Capsule network increases the amount of calculation and the number of parameters increases sharply with the input picture, the Capsule network is improved and the improved Capsule network is used in SAR automatic target recognition (SAR-ATR). In this paper, based on the mechanism of brain visual cortex processing information in hierarchical structure and column form, the idea of complete instantiation was proposed, and the brain-like calculation was used to improve the Capsule network. The specific method was to use multiple convolution layers to achieve hierarchical processing. The number of convolution

收稿日期: 2020-01-05; 修订日期: 2020-02-08

作者简介: 张盼盼 (1992-), 博士生, 主要从事模式识别与智能系统方面的研究。Email: zhangpanpan@sia.cn

导师简介: 罗海波 (1967-), 男, 研究员, 博士, 主要从事图像处理、模式识别与智能系统方面的研究。Email: luohb@sia.cn

kernels used in each layer increases with the depth of the hierarchy, which made the extracted abstract features gradually increase. In the PrimaryCaps layer, the Capsule vector consisted of all the feature maps output by the last layer of the convolutional layer, so that the Capsule unit contained all the features of the target part or the whole to achieve full instantiation of the target. On the SAR-ATR, a comparison experiment was performed with the Capsule network, the traditional target recognition algorithm and the target recognition algorithm based on the classical convolutional neural network. The experimental results show that the improved Capsule network training parameters and calculations are greatly reduced, and the training speed is greatly improved, and the recognition accuracy on the SAR image data set is increased by 0.37 and 1.96-8.96 percentage points compared with the Capsule network and the first two methods respectively.

**Key words:** target recognition; Capsule network; complete instantiation; brain-like calculation; convolutional neural networks

## 0 引言

作为深度学习网络模型,卷积神经网络在图像分类、计算机视觉等领域取得了长足的进展<sup>[1]</sup>。但是,卷积神经网络缺乏对目标的相对位置和空间关系的处理能力。因此,为了识别不同视角的目标,需要更多的训练样本或者更多的网络结构<sup>[2-3]</sup>。针对经典卷积神经网络存在的问题,一种新的深度学习网络 Capsule<sup>[4-5]</sup>被提出。Capsule 网络包含很多 Capsule 单元,而 Capsule 单元是由多种特征组合而成的多维向量,代表一个物体的整体或者一部分。Capsule 输出既有本身的激活概率也有描述它们属性的实例化参数,其中属性包括姿态、形变、方向和纹理等<sup>[6]</sup>。在传输过程中,耦合过滤原则被用来激活更高层的 Capsules 和在 Capsules 之间建立局部与整体的空间关系。Capsule 网络通过动态路由机制<sup>[4-5]</sup>调整局部与整体的空间关系,这种空间关系的建立使得 Capsule 网络能更好地处理目标的相对位置关系和空间关系问题。但是,由于 Capsule 网络需要消耗大量的计算和存储资源,因此处理图片大小一般不超过 32×32。这给 Capsule 网络在军事领域的应用带来了极大的限制,如在 SAR-ATR 应用中,合成孔径雷达(SAR)图像的目标切片大小一般大于 128×128,可见光和红外图像具有更大的图像尺寸,采用 Capsule 网络进行目标识别,对计算和存储资源的需求成为了一个挑战。

文中基于脑机制提出了一种针对 Capsule 网络的改进算法,大大减小了计算量和参数数量。在 SAR-ATR 上,与 Capsule 网络、SVM<sup>[7]</sup>、AdaBoost<sup>[7]</sup>、IGT<sup>[7]</sup>、DCNN<sup>[8-9]</sup>、CGM<sup>[10]</sup>和 2-VDCNN<sup>[11]</sup>识别算法相比,准确率也得到明显的提升。这为深度学习技术

的发展借鉴生物脑机制提供了有力的证据。

## 1 相关工作

Capsule 概念的提出借鉴了初级大脑视觉皮层中具有相同功能的神经元以柱状的形式组织的机制<sup>[12]</sup>,以多个神经元组合表征实体的局部或整体。在初级大脑视觉皮层中,神经元以层级结构组织,而每一层中具有相同或相似功能的神经元又以柱状的形式被组织在一起,比如处理颜色属性和处理纹理属性的神经元分别组成处理颜色属性的柱状结构和处理纹理属性的柱状结构。

### 1.1 Capsule 单元

一个 Capsule 单元是一个多维向量,它的长度代表了一个实体存在的可能性。向量的每个维度代表了实体的属性,包括位置、方向、色度、纹理等。Capsule 单元的结构如图 1 所示,用特征图中不同的特征作为 Capsule 单元的各个维度,图中不同的颜色代表了不同的特征,虚线框表示一个 Capsule 单元。

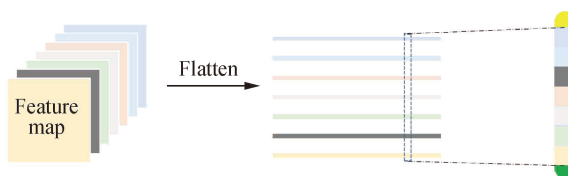


图 1 Capsule 单元结构

Fig.1 Structure of Capsule unit

### 1.2 Capsule 网络结构

Capsule 网络包含两个卷积层,卷积核的大小都是 9×9 像素,输出通道皆为 256 通道。以 Capsule 网络处理手写数据集图像为例,手写数据集的图像大小

为 28×28。在 PrimaryCaps 层中,经过第二次卷积,得到 256 个像素大小为 6×6 的特征图;然后进行矩阵维度变换,将 256 个特征图分成 32 组,每一组包含 8 个特征图,形成 1 152 个向量维度为 8 的 Capsule 单元。最后一层为 DigitCaps 层,包含 10 个向量维度为 16 的 Capsule,表示 10 个实体类别。除了这些基本的结构外,原论文中还添加了重构层,将识别出来的目标还原成图片,用来计算重构损失以及可视化目标识别类别。重构层中包含 3 个全连接层,每一层神经元的数量分别为 512、1 024 和 784。

Capsule 网络用向量的长度表征实体存在的可能性。通过非线性变换,将向量映射到 (0, 1) 之间。映射函数为非线性函数,定义如下:

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \frac{s_j}{\|s_j\|} \quad (1)$$

式中:  $v_j$  是 Capsule  $j$  的向量输出;  $s_j$  是 Capsule  $j$  的总输入。

实例化向量  $v_j$  的长度代表一个 Capsule 实体存在的可能性。对于手写数字数据集,如果数字出现在图像中,希望最后一层的 Capsules 对于每一个类别  $k_i (i = 0, 1, 2, \dots, 9)$ , 都有一个长的输出向量。为了在一张图像中检测多个类别,网络中使用一个分离的边缘损失,  $L_k$  代表其中一类目标的损失,公式定义如下:

$$L_k = T_k \cdot \max(0, m^+ - \|V_k\|^2) + \lambda(1 - T_k) \max(0, \|V_k\| - m^-)^2 \quad (2)$$

式中: 如果数字类别  $k$  出现,  $T_k = 1$ ;  $m^+ = 0.9$ ;  $m^- = 0.1$ ;  $\lambda$  是为了阻止缺失的类别缩短所有激活向量的长度。

### 1.3 Capsule 网络的动态路由机制

Capsule 网络通过动态路由机制<sup>[4-5]</sup>调整输出层的输出。输出层中的一个 Capsule 单元的总输入  $s_j$  是所有预测向量  $\hat{u}_{ji}$  的加权和。 $\hat{u}_{ji}$  是前一层 Capsules 和权重矩阵  $W_{ij}$  的乘积。计算公式如下:

$$s_j = \sum_i c_{ij} \hat{u}_{ji}, \hat{u}_{ji} = W_{ij} u_i \quad (3)$$

式中:  $c_{ij}$  是耦合系数,通过动态路由迭代过程调整其大小。耦合系数  $c_{ij}$  之和为 1,由 softmax 计算得到:

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})} \quad (4)$$

式中:  $b_{ij}$  由  $\hat{u}_{ji}$  和  $v_j$  的乘积加上其本身计算得到;  $v_j$  是 Capsule  $j$  的向量输出。下面是动态路由算法的实现过程:

#### Procedure 1 Routing algorithm.

- 1: procedure Routing ( $\hat{u}_{ji}, r, l$ )
- 2: for all Capsule  $i$  in layer  $l$  and Capsule  $j$  in layer  $(l + 1)$ :  $b_{ij} \leftarrow 0$
- 3: for  $r$  iterations do
- 4: for all Capsule  $i$  in layer  $l$ :  $lc \leftarrow \text{softmax}(b_{ij})$
- 5: for all Capsule  $j$  in layer  $(l + 1)$ :  $s_j \leftarrow \sum_i c_{ij} \hat{u}_{ji}$
- 6: for all Capsule  $j$  in layer  $(l + 1)$ :  $v_j \leftarrow \text{squash}(s_j)$
- 7: for all Capsule  $i$  in layer  $l$  and Capsule  $j$  in layer  $(l + 1)$ :  $b_{ij} \leftarrow b_{ij} + \hat{u}_{ji} \cdot v_j$
- 8: return  $v_j$

### 1.4 Capsule 网络的重构

Capsule 网络使用重构损失作为总损失的一部分,促使输出 Capsules 编码输入图像的实例化参数。重构结构如图 2 所示,激活的 Capsule 的输出被送到一个包含 3 个全连接层的解码器。重构损失的计算公式如下:

$$L_{rec} = \sum_0^{783} (\hat{X} - X)^2 \quad (5)$$

式中:  $\hat{X}$  表示重构图像;  $X$  表示目标图像。为了防止训练过程中重构损失主导边缘损失,在总损失计算时,  $L_{rec}$  需要乘以系数  $\alpha = 0.000 5$ 。

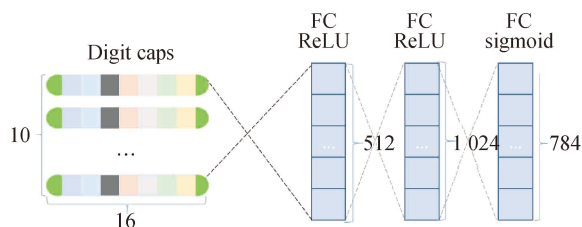


图 2 Capsule 网络的重构层

Fig.2 Layers of reconstruction of Capsule network

### 1.5 Capsule 网络的优势

- (1) 提取的特征具有局部与整体空间关系的不变性。这一特性有助于识别或分割较小的目标。
- (2) 路由机制代替池化操作。这一机制保全了目标的信息,在识别邻近或重叠目标方面具有更好的优势。
- (3) 以向量输出代替标量输出。向量包含了目标的各个属性,实例化目标,能更好地解译目标。
- (4) 在相同识别效果下,与经典卷积神经网络相比,需要更少的训练数据、更少的层和参数。

## 2 改进的 Capsule 网络

文中基于脑机制提出了完全实例化的思想,并对原 Capsule 网络进行了改进,不仅保留了原网络的优势,而且减少了计算量和参数数量,同时也提升了训练的速度及识别的准确率。

初级大脑视觉皮层是以层级结构以及柱状组织的形式处理信息,并且随着层级的加深,提取的特征也更加抽象<sup>[12]</sup>。在初级视觉皮层中,获取到的视觉信号被逐级的处理,高层信息特征是对低层信息特征的整合<sup>[12]</sup>。基于此,文中对原 Capsule 网络结构进行了改进。一是特征提取结构的改进:原 Capsule 网络仅采用两层卷积层,而且卷积核的数量固定且较多,没能更好地体现初级大脑视觉皮层以层级处理信息的机制,以及提取的特征没有更加趋于抽象化,如图 3 中 CapsNet 所示。针对此问题,文中采用多层卷积结

构实现层级处理,以及采用较少且逐渐增加的卷积核用于特征提取,不仅减小了特征图的尺寸,而且使得特征图所包含的信息更加趋向于抽象化,结构如图 3 中 Improved CapsNet 所示(绿色部分为添加的卷积层,红色部分为卷积核的数量);二是 Capsule 单元的改进:原 Capsule 网络借鉴大脑初级视觉皮层柱状组织的机制采用 Capsule 单元实例化目标,但在原 Capsule 网络的 PrimaryCaps 层中,通过将特征图分组的方式构成 Capsule 单元,如图 3 中 CapsNet 所示(橙色部分),因此 Capsule 单元只包含了部分特征信息,导致信息的缺失,不能更好地实例化目标。针对该问题,在改进后的 PrimaryCaps 层中,采用全体特征图的形式构成 Capsule 单元,使得 Capsule 向量包含所有特征图的特征信息,避免了信息的缺失,实现了目标的完全实例化,结构如图 3 中 Improved CapsNet 所示(橙色部分)。

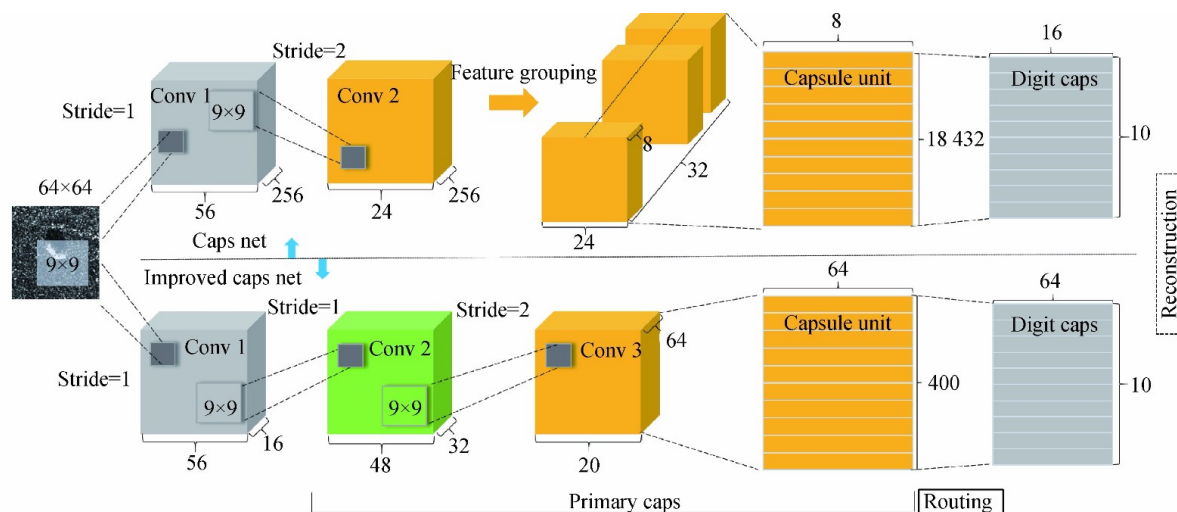


图 3 原 Capsule 网络结构与改进的 Capsule 网络结构

Fig.3 Structure of original Capsule network and improved Capsule network

改进的 Capsule 网络卷积层的卷积核数量依次为 16, 32 和 64, 卷积核大小都是 9×9 像素,如图 3 中 Improved CapsNet 所示;同时采用网络重构作为正则化方式。改进的 Capsule 网络与原 Capsule 网络性能对比如表 1 所示(用于训练和测试实验的原始 SAR 数据集见表 2),参数数量和计算量分别降低了 1.6 倍和 31 倍,训练速度提高了一倍。虽然改进的 Capsule 网络采用了较少的卷积核,但是在识别的准确度上却得到了提升,不仅得益于更加抽象化的特征信息,还在于改进的 Capsule 单元较原 Capsule 单元能更好地实例化目标,避免了信息的缺失,从而提高了

识别的准确度,识别结果如表 3 和表 4 所示。表 1 中对比结果在批次为 9、输入图像像素大小为 64×64 条件下进行测试,服务器系统为 Ubuntu 16.04.6 LTS,

表 1 改进的 Capsule 网络与 Capsule 网络性能对比

Tab.1 Performance comparison of improved Capsule network and Capsule network

	Model size (parameters)	Training_time/epoch	BFLOPs
Capsule	33.73 M	2 min 8 s	33.519
Improved Capsule	21.65 M	1 min 3 s	1.078

GPU 为 GeForce GTX 1080 Ti。

### 3 实验与分析

文中采用了标准工作条件 (SOC) 下采集的 10 类军用车辆的 SAR 图像数据集进行训练和测试, 并与 Capsule 网络在 SAR 图像数据集上的识别结果进行对比。对比算法除了原 Capsule 网络之外, 还有其他 7 种算法, 主要分为传统 SAR 目标识别算法和基于经典卷积神经网络的 SAR 目标识别算法。

#### 3.1 数据描述

获取该数据集的传感器为高分辨率的聚束式合成孔径雷达, 该雷达的分辨率为  $0.3\text{ m}\times 0.3\text{ m}$ 。MSTAR

数据集的采集条件分为两类, 分别为标准工作条件 (SOC) 和扩展工作条件 (EOC)。文中主要采用标准工作条件 (SOC) 下采集的数据集进行实验。图 4 中展示了 SOC 数据集 10 类车辆目标 (坦克: T62 和 T72; 自行榴弹炮: 2S1; 货运卡车: ZIL131; 装甲运输车: BMP2、BTR70、BTR60 和 BRDM2; 自行高炮: ZSU23/4; 推土机: D7) 的可见光图像和对应的 SAR 图像。训练数据集和测试数据集分别在  $17^\circ$  俯仰角下和  $15^\circ$  俯仰角下进行采集, 各类目标的切片图像大小为  $128\times 128$ 。由于实验设备的资源有限, 将原图像转换成  $64\times 64$  的图像作为输入图像进行对比仿真实验, 所用数据详细信息如表 2 所示。



图 4 (a) 和 (b) 分别为 BMP2、BTR70、T72、BTR60 和 2S1 的光学图像和相对应的 SAR 图像; (c) 和 (d) 分别为 BRDM2、D7、T62、ZIL131 和 ZSU23/4 的光学图像和相对应得 SAR 图像

Fig.4 Optical images and their corresponding MSTAR SAR images for (a) and (b) BMP2, BTR70, T72, BTR60, and 2S1; (c) and (d) BRDM2, D7, T62, ZIL131, and ZSU23/4

表 2 用于训练和测试实验的原始 SAR 数据集

Tab.2 Raw SAR dataset for training and testing in experiment

Class	BMP2sn-9563	BTR70	T72sn-132	BTR60	2S1	BRDM2	D7	T62	ZIL131	ZSU23/4
Train samples( $17^\circ$ )	117	117	116	128	150	149	150	150	150	150
Test samples( $15^\circ$ )	195	196	196	195	274	274	274	273	274	274

#### 3.2 实验结果

实验中采用图像重构作为正则化方法, 并将重构损失计算到总损失之中, 改进的 Capsule 网络训练过程中图像的重构结果如图 5 所示。在迭代将近 6 000 步时, 损失曲线开始趋于平缓, 训练过程中重构的错误曲线和训练损失曲线如图 6 所示。

表 3 和表 4 以混淆矩阵的形式分别显示了原 Capsule 网络和改进的 Capsule 网络的识别结果, 输入图像大小皆为  $64\times 64$ 。混淆矩阵中的每一个元素代表了每一种类别被识别的可能性。表中的行代表实际的目标类别, 列代表被预测的类别。从表 3 和表 4 中

可以看出, 改进的 Capsule 网络的识别率达到了 98.85%, 较原 Capsule 网络 98.48% 的识别率提高

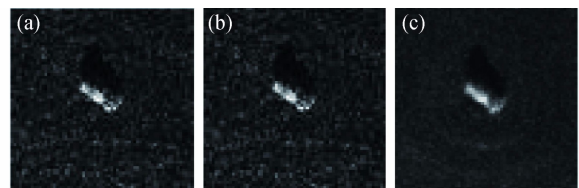


图 5 改进的 Capsule 网络的重构结果。(a) 原始图像, (b) 目标图像, (c) 重构图像

Fig.5 Reconstruction result of improved Capsule. (a) Original image, (b) Target image and (c) Reconstruction image

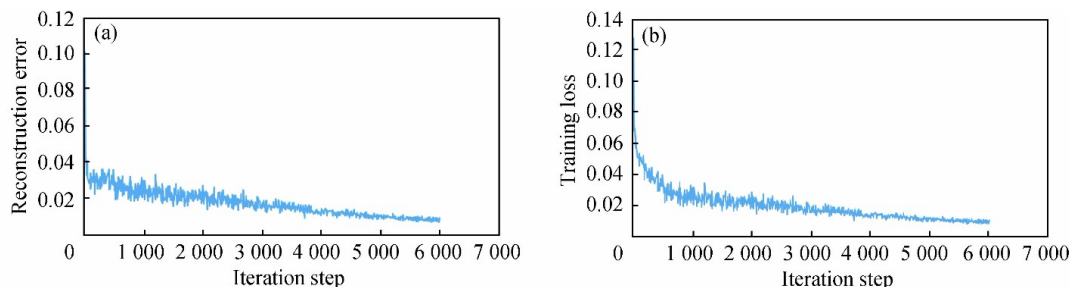


图 6 改进的 Capsule 网络训练中重构错误曲线和训练损失曲线。(a) 重构错误曲线, (b) 训练损失曲线

Fig.6 Reconstruction error curve and training loss curve of improved Capsule network. (a) Reconstruction error curve and (b) training loss curve

表 3 原 Capsule 网络 10 类目标识别结果的混淆矩阵 (识别率: 98.48%)

Tab.3 Confusion matrix of 10-class target recognition results of Capsule network(recognition rate: 98.48%)

Class	BMP2sn-9563	BTR70	T72sn-132	BTR60	2S1	BRDM2	D7	T62	ZIL131	ZSU23/4
BMP2sn-9563	96.92	0.51	2.57	0	0	0	0	0	0	0
BTR70	0	100.00	0	0	0	0	0	0	0	0
T72sn-132	0	0	100.00	0	0	0	0	0	0	0
BTR60	0	0	0	98.46	0	0.51	0	0	0	1.03
2S1	0	0	0	2.92	94.16	1.46	0	0.73	0.73	0
BRDM2	0	0.365	0	0.73	0	97.45	0	0	1.09	0.365
D7	0	0.73	0	0	0	0	99.27	0	0	0
T62	0	0	0	0	0.73	0	0	98.90	0	0.37
ZIL131	0	0	0	0	0	0	0	0	100.00	0
ZSU23/4	0	0	0	0	0	0	0.36	0	0	99.64

表 4 改进的 Capsule 网络 10 类目标识别结果的混淆矩阵 (识别率: 98.85%)

Tab.4 Confusion matrix of 10-class recognition results of improved Capsule network (recognition rate: 98.85%)

Class	BMP2sn-9563	BTR70	T72sn-132	BTR60	2S1	BRDM2	D7	T62	ZIL131	ZSU23/4
BMP2sn-9563	96.41	0	3.59	0	0	0	0	0	0	0
BTR70	0	100.00	0	0	0	0	0	0	0	0
T72sn-132	0	0	100.00	0	0	0	0	0	0	0
BTR60	0	0	0	98.97	0	0	0	0	0	1.03
2S1	0	0	0	2.555	96.35	1.095	0	0	0	0
BRDM2	0	0	0	0.73	0	98.54	0.365	0	0.365	0
D7	0.365	0	0	0	0	0.365	99.27	0	0	0
T62	0	0	0	0	0	0	0	99.63	0	0.37
ZIL131	0	0	0	0	0	0	0.36	0	99.64	0
ZSU23/4	0	0	0	0	0	0	0.36	0	0	99.64

0.37 个百分点。

### 3.3 识别效果对比

在标准工作条件 (SOC) 下, 除与原 Capsule 网络进行了对比外, 还与基于经典卷积神经网络的目标识别算法和传统目标识别算法进行了对比。它们分别是 SVM<sup>[13]</sup>、AdaBoost<sup>[14]</sup>、CGM<sup>[10]</sup>、IGT<sup>[7]</sup>、DCNNs<sup>[8-9]</sup>

和 2-VDCNN<sup>[11]</sup>。DCNNs 和 2-VDCNN 都是经典卷积神经网络模型, 其中 2-VDCNN 是多视角目标识别算法, 相较于单一视角目标识别更容易取得好的识别效果。对比结果如表 5 所示, 其中被引用的结果分别来自参考文献 [7]、[9-11]、Capsule 网络的识别结果以及文中改进的 Capsule 网络的识别结果, 对比实验使

用图片大小在 64×64 和 128×128 之间。

表 5 不同方法的识别效果

Tab.5 Recognition performance of different methods

Methods	SOC	
	Rates	Training images
SVM <sup>[7]</sup>	90.10%	3 670
AdaBoost <sup>[7]</sup>	92.70%	3 670
DCNN <sup>[9]</sup>	92.30%	3 671
DCNN <sup>[8]</sup>	94.56%	2 747
IGT <sup>[7]</sup>	95.00%	3 670
CGM <sup>[10]</sup>	97.18%	3 670
2-VDCNN <sup>[11]</sup>	97.81%	1 377
CapsNet	98.48%	1 377
Improved CapsNet	98.85%	1 377

从表 5 中可以看出,原 Capsule 网络和改进的 Capsule 网络的识别率均超过了基于卷积神经网络的目标识别算法,证明了原 Capsule 网络相较于卷积神经网络的优越性。而改进的 Capsule 网络的识别率相较于原 Capsule 网络提高了 0.37 个百分点,证明改进后的 Capsule 网络具有更好的优势,和其他算法相比,不仅使用了更少的训练数据,识别率也有明显的提升。这些数据表明,在 SAR-ATR 上,改进的 Capsule 网络具有更好的识别效果。

#### 4 结 论

文中首先分析了 Capsule 网络的结构、原理及特性,并结合脑机制对原 Capsule 网络存在的不足进行了改进。然后,将原 Capsule 网络和改进后的 Capsule 网络应用于 SAR 图像目标识别,并与其他 SAR 图像目标识别算法进行了对比。实验结果表明,在性能上,改进的 Capsule 网络较原 Capsule 网络的参数数量和计算量分别降低了 1.6 倍和 31 倍,训练速度提高了一倍;在 SAR-ATR 上,改进的 Capsule 网络的识别率较原 Capsule 网络和多视角 2-VDCNN 分别提高了 0.37 和 1.04 个百分点。与其他 6 种算法相比,在训练样本分别减小 62.5% 和 49.9% 的条件下,改进的 Capsule 网络的识别率分别提高了 1.96~8.96 个百分点。从而验证了基于脑机制改进的 Capsule 网络较原 Capsule 网络在性能上有较大的提升,并且在 SAR-ATR 上具有更好的识别效果。尽管改进的 Capsule

网络在一定程度上减少了计算量和参数数量,但是针对更大图片的处理仍然存在计算消耗和存储需求较大的问题。因此,在后续的工作中,将考虑去除没有或包含较少信息量的特征,进一步减少 Capsule 单元的维度,从而使得对计算资源和存储空间的需求降低。

#### 参考文献:

- [1] Yang Nan, Nan Lin, Zhang Dingyi, et al. Research on image interpretation based on deep learning [J]. *Infrared and Laser Engineering*, 2018, 47(2): 0203002. (in Chinese)  
杨楠, 南琳, 张丁一, 等. 基于深度学习的图像描述研究[J]. *红外与激光工程*, 2018, 47(2): 0203002.
- [2] Cohen T, Welling M. Group equivariant convolutional networks[C] //International Conference on Machine Learning. 2016: 2990-2999.
- [3] Cohen T S, Geiger M, Köhler J, et al. Spherical cnns[J]. arXiv preprint arXiv: 1801.10130, 2018.
- [4] Sabour S, Frosst N, Hinton G E. Dynamic routing between Capsules [J]. *Computer Vision and Pattern Recognition*, 2017, arXiv: 1710.09829: 1-11.
- [5] Sabour S, Frosst N, Hinton G. Matrix capsules with EM routing[C]//6th International Conference on Learning Representations, ICLR. 2018: 1-15.
- [6] Hinton G E, Krizhevsky A, Wang S D. Transforming auto-encoders[C]//International Conference on Artificial Neural Networks. Berlin, Heidelberg: Springer, 2011: 44-51.
- [7] Srinivas U, Monga V, Raj R G. SAR automatic target recognition using discriminative graphical models [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2014, 50(1): 591-606.
- [8] Ding J, Chen B, Liu H, et al. Convolutional neural network with data augmentation for SAR target recognition [J]. *IEEE Geoscience and Remote Sensing Letters*, 2016, 13(3): 364-368.
- [9] Morgan D A E. Deep convolutional neural networks for ATR from SAR imagery[C]//Algorithms for Synthetic Aperture Radar Imagery XXII. International Society for Optics and Photonics, 2015, 9475: 94750F.
- [10] O'Sullivan J A, DeVore M D, Kedia V, et al. SAR ATR performance using a conditionally Gaussian model [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2001, 37(1): 91-108.
- [11] Pei J, Huang Y, Huo W, 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.
- [12] Eric R K, James H S, Thomas M J, et al. Principles of Neural Science [M]. 5th ed. Beijing: China Machine Press, 2013.
- [13] Zhao Q, Principe J C. Support vector machines for SAR automatic target recognition [J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2001, 37(2): 643–654.
- [14] Sun Y, Liu Z, Todorovic S, et al. Adaptive boosting for SAR automatic target recognition [J]. *IEEE Transactions on Aerospace & Electronic Systems*, 2007, 43(1): 112–125.