

基于卷积神经网络的SAR图像目标识别研究

田壮壮* 占荣辉 胡杰民 张军

(国防科学技术大学自动目标识别重点实验室 长沙 410073)

摘要: 针对合成孔径雷达(Synthetic Aperture Radar, SAR)的图像目标识别应用, 该文提出了一种基于卷积神经网络(Convolutional Neural Network, CNN)的SAR图像目标识别方法。首先通过在误差代价函数中引入类别可分性度量, 提高了卷积神经网络的类别区分能力; 然后利用改进后的卷积神经网络对SAR图像进行特征提取; 最后利用支持向量机(Support Vector Machine, SVM)对特征进行分类。使用美国运动和静止目标获取与识别(Moving and Stationary Target Acquisition and Recognition, MSTAR)SAR图像数据进行实验, 识别结果证明了所提方法的有效性。

关键词: 合成孔径雷达; 自动目标识别; 卷积神经网络; 支持向量机; BP算法

中图分类号: TN957.52

文献标识码: A

文章编号: 2095-283X(2016)03-0320-06

DOI: 10.12000/JR16037

引用格式: 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.

Reference format: 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.

SAR ATR Based on Convolutional Neural Network

Tian Zhuangzhuang Zhan Ronghui Hu Jiemin Zhang Jun

(ATR Key Laboratory, National University of Defense Technology, Changsha 410073, China)

Abstract: This study presents a new method of Synthetic Aperture Radar (SAR) image target recognition based on a convolutional neural network. First, we introduce a class separability measure into the cost function to improve this network's ability to distinguish between categories. Then, we extract SAR image features using the improved convolutional neural network and classify these features using a support vector machine. Experimental results using moving and stationary target acquisition and recognition SAR datasets prove the validity of this method.

Key words: Synthetic Aperture Radar (SAR); Automatic Target Recognition (ATR); Convolutional Neural Network (CNN); Support Vector Machine (SVM); Back Propagation (BP)

1 引言

合成孔径雷达(Synthetic Aperture Radar, SAR)是当前对地观测的重要手段之一。由于具有全天候、全天时和一定的穿透成像能力, SAR已经被广泛应用于战场侦查和情报获取等领域。SAR图像自动目标识别(Automatic Target Recognition, ATR)技术可以有效获取目标信息, 提高军队自动

化水平。目前识别技术上常用的方法包括模板匹配法^[1], 支持向量机(Support Vector Machine, SVM)法^[2], 文献[3]中提出的基于Gabor滤波器和局部纹理特征提取的方法, 文献[4]中提出的多信息字典学习和稀疏表示相结合的方法等等。但是这些方法都需要人工凭借经验进行特征提取, 具有一定的盲目性和不确定性。

近年来, 深度学习(Deep Learning)理论在模式识别领域掀起了一股浪潮, 相关的理论和实验成果不断出现。2006年, Hinton等人^[5]提出了非监督的逐层贪婪训练方法, 解决了深度增加所带来的“梯度耗散”问题, 使神经网络能够朝着更深的方

收稿日期: 2016-02-03; 改回日期: 2016-03-15; 网络出版: 2016-05-09

*通信作者: 田壮壮 tzz14@nudt.edu.cn

基金项目: 国家自然科学基金(61471370)

Foundation Item: The National Natural Science Foundation of China (61471370)

向进行发展。随后许多学者根据不同的应用背景提出了多种DL模型,如深度置信网络(Deep Belief Network, DBN)^[6]、栈式消噪自动编码器(Stacked Denoising Autoencoders, SDA)^[7]和卷积神经网络(Convolutional Neural Network, CNN)^[8]等。由于可以直接处理2维图像,卷积神经网络在图像识别领域取得了较多的应用,如人脸识别^[9],人体行为识别^[10],医学图像识别^[11]等。

CNN主要有以下特点:采用原始图像作为网络输入,避免了复杂的特征提取过程;局部感受域获取的特征与图像的平移、缩放和旋转无关;权值共享结构降低了网络模型的复杂度;下采样利用局部相关性原理,在有效减少数据处理量的同时保留了结构信息^[12]。CNN作为一种特征学习方法,通过非线性的简单模型将原始数据转化为更为抽象的表达,在特征学习过程中不需要过多的人工参与^[13]。

本文主要研究将CNN应用于SAR图像目标识别领域,减少特征选取中的人工操作。为了进一步增强CNN所提取特征的可分性,本文尝试将类别可分性度量引入代价函数之中。为了提高CNN对特征的分类能力,利用支持向量机对CNN模型中的Softmax分类器进行替换。实验结果表明,该方法具有良好的识别分类效果,证明了其有效性。

2 卷积神经网络

卷积神经网络是一种多层神经网络结构,主要由输入层、卷积层、下采样层、输出层等组成,其中卷积层和下采样层为隐含层。在卷积神经网络中,输入层用于接收原始图像;卷积层用于提取图像特征,并降低噪声的影响;下采样层利用图像局部相关性原理,减少需要处理的数据量;输出层将提取特征映射为最终所得到的标签。

图像在CNN中的映射过程是一个前向传播过程,上一层的输出作为当前层的输入。为了避免线性模型表达能力不够的问题,各层在传递时需要加入非线性的激活函数。第 l 层的输出可以表示为:

$$\left. \begin{aligned} \mathbf{u}^l &= \mathbf{W}^l \mathbf{x}^{l-1} + \mathbf{b}^l \\ \mathbf{x}^l &= f(\mathbf{u}^l) \end{aligned} \right\} \quad (1)$$

式中, l 表示层数, \mathbf{x} 为特征图, \mathbf{W} 为当前网络层的映射权值矩阵, \mathbf{b} 为当前网络的加性偏置项, f 为激活函数。

CNN模型常用后向传播(Back Propagation, BP)算法^[14]来对参数进行迭代更新。BP算法是一种常见的监督式学习方法。该算法首先利用样本的实际输出与期望输出求得代价函数,然后将代价函数

反向传播至各层,最后利用梯度下降法(Gradient Descent, GD)沿着代价函数的负梯度方向调整网络中的权值和偏置等参数。

首先假设输出的目标类别维数为 B ,即共有 B 类样本,网络以均方误差作为代价函数。第 n 个样本的代价函数 E_0 为:

$$E_0 = \frac{1}{2} \sum_{k=1}^B (t_k^n - y_k^n)^2 = \frac{1}{2} \|\mathbf{t}^n - \mathbf{y}^n\|_2^2 \quad (2)$$

式中, t_k^n 表示第 n 个样本中第 k 个神经元所对应的期望值, y_k^n 表示第 k 个神经元的实际输出值。

然后定义 $\delta^l = \frac{\partial E_0}{\partial \mathbf{u}^l}$ 为第 l 层的灵敏度,设第 L 层为输出层,根据式(1)、式(2)可得输出层的灵敏度为:

$$\delta^L = f'(\mathbf{u}^L) \circ (\mathbf{y}^n - \mathbf{t}^n) \quad (3)$$

式中,“ \circ ”表示点乘。

在反向传播过程中,可由第 $l+1$ 层的灵敏度 δ^{l+1} 和式(1)得到第 l 层的灵敏度为:

$$\delta^l = \mathbf{W}^{l+1} \delta^{l+1} \circ f'(\mathbf{u}^l) \quad (4)$$

完成了代价函数的反向传播之后,需要利用梯度下降法根据代价函数对网络中参数的偏导数对参数进行更新,最终使代价函数达到极小值。偏导数在此处就是 $\frac{\partial E_0}{\partial \mathbf{W}^l}$ 和 $\frac{\partial E_0}{\partial \mathbf{b}^l}$ 。

分别求 E_0 对 \mathbf{W}^l 和 \mathbf{b}^l 的偏导,由式(1)可得:

$$\left. \begin{aligned} \frac{\partial E_0}{\partial \mathbf{W}^l} &= \frac{\partial E_0}{\partial \mathbf{u}^l} \circ \frac{\partial \mathbf{u}^l}{\partial \mathbf{W}^l} = \delta^l \circ \mathbf{x}^{l-1} \\ \frac{\partial E_0}{\partial \mathbf{b}^l} &= \frac{\partial E_0}{\partial \mathbf{u}^l} \circ \frac{\partial \mathbf{u}^l}{\partial \mathbf{b}^l} = \delta^l \end{aligned} \right\} \quad (5)$$

根据梯度下降法,可以得出第 l 层神经元权值的更新变化量为:

$$\left. \begin{aligned} \Delta \mathbf{W}^l &= -\eta \frac{\partial E_0}{\partial \mathbf{W}^l} \\ \Delta \mathbf{b}^l &= -\eta \frac{\partial E_0}{\partial \mathbf{b}^l} \end{aligned} \right\} \quad (6)$$

式中, η 为学习率。

3 卷积神经网络的改进

3.1 引入类别可分性度量

为了使CNN模型所提取的特征具有更好的类别可分性,代价函数在式(2)的基础上,增加了样本的类别可分性度量^[15],即类内距离与类间距离:

$$E_1 = \frac{1}{2} \|\mathbf{y}_c^n - \mathbf{M}_c\|_2^2 \quad (7)$$

$$E_2 = \frac{1}{2} \sum_{c=1}^B \|\mathbf{M}_c - \mathbf{M}_c\|_2^2 \quad (8)$$

式中, \mathbf{y}_c^n 表示属于第 c 类的第 n 个样本的实际输出值, $\mathbf{M}_c, \mathbf{M}_{c'}$ 分别为第 c 类和第 c' 类样本输出的均值。 E_1 和 E_2 分别表示样本经过卷积神经网络后输出特征的类内距离和类间距离。为了使得网络的输出特征具有更强的可分性, 在迭代中应使类内距离减小, 类间距离增大。引入类内及类间距离后, 代价函数修正如下:

$$E = E_0 + \alpha E_1 - \beta E_2 \quad (9)$$

式中, α 与 β 为权值参数。

对代价函数进行修正的目的是在调整权值参数时, 使网络向着有利于分类的方向进行调整, 提高最终训练结果的可分性。为了使代价函数能够反向传播, 需要求得修正后代价函数的灵敏度。

对于类内距离函数 E_1 , 其输出层的灵敏度为:

$$\begin{aligned} \delta_1^L &= \frac{\partial E_1}{\partial \mathbf{u}^L} = \frac{\partial}{\partial \mathbf{u}^L} \frac{1}{2} \|\mathbf{y}_c^n - \mathbf{M}_c\|_2^2 \\ &= \left(1 - \frac{1}{n_c}\right) f'(\mathbf{u}^L) \circ (\mathbf{y}_c^n - \mathbf{M}_c) \end{aligned} \quad (10)$$

式中, n_c 表示属于 c 类的样本个数。

对于类间距离函数 E_2 , 其输出层的灵敏度为:

$$\begin{aligned} \delta_2^L &= \frac{\partial E_2}{\partial \mathbf{u}^L} = \frac{\partial}{\partial \mathbf{u}^L} \frac{1}{2} \sum_{c'=1}^B \|\mathbf{M}_c - \mathbf{M}_{c'}\|_2^2 \\ &= \frac{1}{n_c} f'(\mathbf{u}^L) \circ \sum_{c'=1}^B (\mathbf{M}_c - \mathbf{M}_{c'}) \end{aligned} \quad (11)$$

根据式(3)、式(9)、式(10)和式(11)可以得到修正后的输出层灵敏度为:

$$\begin{aligned} \delta' &= \delta^L + \alpha \delta_1^L - \beta \delta_2^L \\ &= f'(\mathbf{u}^L) \circ (\mathbf{y}^n - \mathbf{t}^n) + \alpha \left(1 - \frac{1}{n_c}\right) \\ &\quad \cdot f'(\mathbf{u}^L) \circ (\mathbf{y}_c^n - \mathbf{M}_c) \\ &\quad - \beta \frac{1}{n_c} f'(\mathbf{u}^L) \circ \sum_{c'=1}^B (\mathbf{M}_c - \mathbf{M}_{c'}) \end{aligned} \quad (12)$$

得到了输出层的灵敏度之后, 根据式(4)即可迭代得到各层的灵敏度, 再根据式(5)和式(6), 可以得到各层参数的更新量。

3.2 支持向量机的使用

在CNN模型中, 通常利用Softmax回归模型作为分类器。该分类器通过全连接方式将CNN模型所提取特征映射为最终输出标签。这种分类方法在训练时可以有效地与反向传播过程结合, 方便模型中参数的更新, 但是Softmax单层全连接的方式并不能很好地解决非线性分类问题。

支持向量机^[16]是Vapnik等人于1995年首先提出的, 主要利用核函数将原始特征空间中的非线性分

类界面隐式地映射到更高维的变换特征空间中产生线性分类界面, 达到更好的分类效果^[17]。SVM在解决非线性分类问题上表现优良。为解决Softmax对非线性样本特征的支持问题, 本文利用SVM代替Softmax作为CNN模型的分器, 对样本通过所有隐含层后所得特征进行分类。

4 实验结果

4.1 实验数据集

为了验证本文方法的有效性, 采用美国国防研究规划局(DARPA)和空军研究实验室(AFRL)联合资助的运动和静止目标获取与识别(Moving and Stationary Target Acquisition and Recognition, MSTAR)数据库进行实验。数据库中的SAR图像分辨率为0.3 m×0.3 m, 目标图像大小为128×128像素。MSTAR数据库中包含了BMP2、BRDM2、BTR60、BTR70、D7、T62、T72、ZIL131、ZSU234、2S1等10类目标的SAR图像数据。实验采用17°俯仰角下的图像作为训练样本, 15°俯仰角下的图像作为测试样本。

4.2 实验模型

实验中用于SAR图像识别的CNN模型如图1所示。该模型包括输入层, 输出层, 2个卷积层和1个下采样层。由于目标在图像中仅占据较小的部分, 为了减少输入维数, 我们选择截取图像中央64×64大小的部分作为模型的输入, 除裁剪外, 未进行其他预处理操作。卷积层采用的卷积核大小为13×13, C1卷积层神经元个数为20个, 输出特征图大小为52×52。C3卷积层神经元有120个, 输出特征图大小为1×1。S2下采样层局部感受域的大小为4×4, 神经元个数为20个, 输出特征图大小为

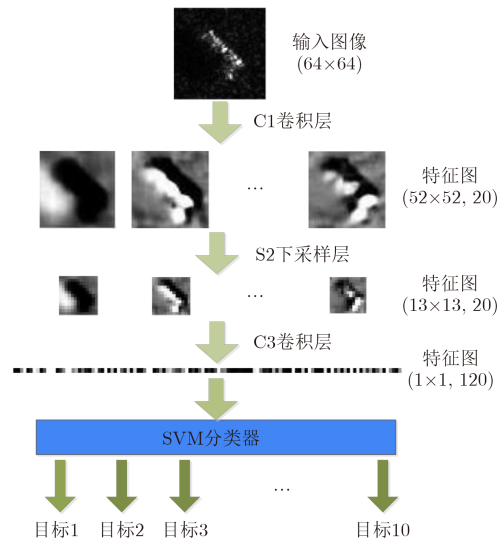


图 1 本文卷积神经网络结构示意图

Fig. 1 CNN structural plan of the method

13×13。激活函数采用的是sigmoid函数，偏置初始值设为0，学习率为1。各层权值的初始值在区间 $\left[-4\sqrt{\frac{6}{fan_{in}+fan_{out}}}, 4\sqrt{\frac{6}{fan_{in}+fan_{out}}}\right]$ 内随机生成^[8]，其中 fan_{in} 为各层的输入特征图个数， fan_{out} 为输出特征图个数。式(9)中的参数设置为 $\alpha=0.03, \beta=2$ 。实验平台为Windows 764bit+Matlab 2010。

4.3 3类目标的识别结果与分析

实验中，训练样本为17°俯仰角下BMP2_SN9563、BTR70_C71、T72_SN132的SAR图像数据，测试数据为15°俯仰角下BMP2_SNC21、BMP2_SN9563、BMP2_SN9566、BTR70、T72_SN132、T72_SN812和T72_SNS7的SAR图像数据。3类目标的识别结果如表1所示。

从表1可以看出，对于相同型号目标的SAR图像，本文方法的正确识别率达到了98.97%以上。对

表 1 3类MSTAR目标识别结果

Tab. 1 The recognition results of three types of MSTAR targets

测试目标	识别结果			正确识别率(%)
	BMP2	BTR70	T72	
BMP2(9563)	193	0	2	98.97
BMP2(9566)	183	9	4	93.37
BMP2(C21)	187	4	5	95.41
BTR70	0	196	0	100
T72(132)	1	0	195	99.49
T72(812)	9	6	180	92.31
T72(S7)	12	4	175	91.62
整体识别率(%)	95.90			

于相同类别不同型号的变体，正确识别率也都在91.62%以上。本文方法整体的平均识别率达到了95.90%。表2中可以看出本文方法识别率优于SVM、ELM以及稀疏表示等方法。实验中，在相同条件下利用相同结构但未经改进的CNN模型进行了识别测试，得到的平均识别率为93.63%，说明本方法相较于改进前CNN模型也有一定提高。

表 2 与其他方法的效果对比

Tab. 2 The recognition performance compare with other methods

方法	平均识别率(%)
SVM(文献[2])	93.54
Gabor+LPM+ELM(文献[3])	94.80
字典学习+稀疏表示(文献[4])	92.90
CNN方法	93.63
本文方法	95.90

4.4 10类目标的识别结果与分析

在10类目标的识别实验中，目标包括BRDM2、BMP2(BMP2_SNC21、BMP2_SN9563、BMP2_SN9566)、BTR60、BTR70、D7、ZIL131、ZSU234、T62、T72(T72_SN132、T72_SN812、T72_SNS7)、2S1等10类。训练样本为17°俯仰角下不包括变体(即不包括BMP2_SNC21、BMP2_SN9566、T72_SN812、T72_SNS7等型号)的SAR图像数据，测试样本为15°俯仰角下的包含变体的各目标SAR图像数据。识别结果如表3。

表 3 10类MSTAR目标识别结果

Tab. 3 The recognition results of ten types of MSTAR targets

测试目标	识别结果										正确识别率(%)
	BMP2	T72	2S1	BRDM2	BTR60	BTR70	D7	T62	ZIL	ZSU	
BMP2(9563)	193	1	0	0	1	0	0	0	0	0	98.97
BMP2(9566)	173	9	1	2	4	7	0	0	0	0	88.27
BMP2(C21)	168	6	3	2	9	3	1	1	3	0	85.71
T72(132)	0	194	0	0	1	0	0	1	0	0	98.98
T72(812)	0	154	6	0	0	5	0	25	5	0	78.97
T72(S7)	4	164	2	0	1	2	0	17	1	0	85.86
2S1	0	4	255	3	2	4	0	5	1	0	93.07
BRDM2	1	1	1	257	5	2	0	0	6	1	93.80
BTR60	0	0	1	0	190	2	0	0	0	2	97.44
BTR70	0	0	1	0	0	195	0	0	0	0	99.49
D7	0	0	0	0	0	0	256	0	4	14	93.43
T62	0	2	0	0	0	0	0	259	8	4	94.87
ZIL	0	0	0	0	0	0	0	0	273	1	99.64
ZSU	0	0	0	0	0	0	2	0	0	272	99.27
整体识别率(%)	93.76										

从表3的分类结果可以看出, 本文方法10类目标分类结果仍然较好, 平均识别率仍能达到93.76%。

5 结 论

针对SAR图像目标识别问题, 本文提出了一种基于卷积神经网络的目标识别方法。通过在代价函数中引入类别可分性度量以及使用支持向量机代替Softmax分类器, 对卷积神经网络进行了改进。利用MSTAR图像数据对该方法进行验证。结果表明, 在不需要人工进行特征提取的情况下, 该方法在3类目标分类实验中, 达到了95.90%的识别率, 在10类目标分类实验中, 平均识别率仍然能达到93.76%。实验结果说明本文方法的有效性以及卷积神经网络在SAR图像目标识别上具有一定的应用潜力。

参 考 文 献

- [1] Ross T D, Worrell S W, Velten V J, *et al.*. Standard SAR ATR evaluation experiments using the MSTAR public release data set[C]. Aerospace/Defense Sensing and Controls, International Society for Optics and Photonics, 1998: 566–573.
- [2] Tao W, Xi C, Xiangwei R, *et al.*. Study on SAR target recognition based on support vector machine[C]. 2009 2nd Asian-Pacific Conference on Synthetic Aperture Radar, 2009: 856–859.
- [3] 王璐, 张帆, 李伟, 等. 基于 Gabor 滤波器和局部纹理特征提取的 SAR 目标识别算法[J]. 雷达学报, 2015, 4(6): 658–665. DOI: 10.12000/JR15076.
Wang Lu, Zhang Fan, Li Wei, *et al.*. A method of SAR target recognition based on Gabor filter and local texture feature extraction[J]. *Journal of Radars*, 2015, 4(6): 658–665. DOI: 10.12000/JR15076.
- [4] 齐会娇, 王英华, 丁军, 等. 基于多信息字典学习及稀疏表示的 SAR 目标识别[J]. 系统工程与电子技术, 2015, 37(6): 1280–1287.
Qi Huijiao, Wang Yinghua, Ding Jun, *et al.*. SAR target recognition based on multi-information dictionary learning and sparse representation[J]. *Systems Engineering and Electronics*, 2015, 37(6): 1280–1287.
- [5] Hinton G E and Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504–507.
- [6] Hinton G E, Osindero S, and Teh Y W. A fast learning algorithm for deep belief nets[J]. *Neural Computation*, 2006, 18(7): 1527–1554.
- [7] Vincent P, Larochelle H, Lajoie I, *et al.*. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion[J]. *The Journal of Machine Learning Research*, 2010, 11: 3371–3408.
- [8] Lecun Y, Bottou L, Bengio Y, *et al.*. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278–2324.
- [9] Ma Y, He J, Wu L, *et al.*. An effective face verification algorithm to fuse complete features in convolutional neural network[C]. MultiMedia Modeling. Springer International Publishing, 2016: 39–46.
- [10] Ijjina E P and Mohan C K. Human action recognition based on motion capture information using fuzzy convolution neural networks[C]. 2015 Eighth International Conference on Advances in Pattern Recognition (ICAPR), Kolkata, 2015: 1–6.
- [11] Ciompi F, de Hoop B, van Riel S J, *et al.*. Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box[J]. *Medical Image Analysis*, 2015, 26(1): 195–202.
- [12] 尹宝才, 王文通, 王立春. 深度学习研究综述[J]. 北京工业大学学报, 2015, 41(1): 48–59.
Yin Bao-cai, Wang Wen-tong, and Wang Li-chun. Review of deep learning[J]. *Journal of Beijing University of Technology*, 2015, 41(1): 48–59.
- [13] Lecun Y, Bengio Y, and Hinton G. Deep learning[J]. *Nature*, 2015, 521(7553): 436–444.
- [14] Rumelhart D E, Hinton G E, and Williams R J. Learning representations by back-propagating errors[J]. *Nature*, 1986, 323(6088): 533–536.
- [15] 孙艳丰, 齐光磊, 胡永利, 等. 基于改进 Fisher 准则的深度卷积神经网络识别算法[J]. 北京工业大学学报, 2015, 41(6): 835–841.
Sun Yanfeng, Qi Guanglei, Hu Yongli, *et al.*. Deep convolution neural network recognition algorithm based on improved fisher criterion[J]. *Journal of Beijing University of Technology*, 2015, 41(6): 835–841.
- [16] Cortes C and Vapnik V. Support-vector networks[J]. *Machine Learning*, 1995, 20(3): 273–297.
- [17] 孙即祥, 等. 现代模式识别[M]. 北京: 高等教育出版社, 2008: 624–625.
Sun Jixiang, *et al.*. Pattern Recognition[M]. Beijing: Higher Education Press, 2008: 624–625.
- [18] Glorot X and Bengio Y. Understanding the difficulty of training deep feedforward neural networks[J]. *Journal of Machine Learning Research*, 2010, 9: 249–256.

作者简介

田壮壮(1993-), 男, 硕士生, 研究方向为雷达目标识别。

E-mail: tzz14@nudt.edu.cn

张 军(1973-), 男, 研究员, 博士, 研究方向为雷达智能信号处理、制导雷达应用技术。

占荣辉(1978-), 男, 讲师, 博士, 研究方向为雷达目标识别、雷达信息处理。

胡杰民(1983-), 男, 讲师, 博士, 研究方向为空间目标识别、雷达成像。