

《电子学报》稿件加工样本

尚书林^{1,2} 左年明² 张哲^{3,4}

摘要 辐射源的快速、准确和鲁棒的自动目标识别在现代军事中的作用十分重要,自动目标识别算法需要可以准确区分出已知目标和未知目标,同时可以正确的对于已知目标进行分类。我们需要在未收集大量数据的前提下,可以迅速的识别出新的目标。与传统的利用已知类别的样本进行训练测试的机器学习算法不同,我们这个问题是在 Open Set 的情况下,将需要考虑将输入识别为未知的情况。本文利用深度卷积神经网络与支持向量机进行结合,以雷达信号的模糊函数作为训练样本,构建了一个可以对未知分类进行识别的分类器。我们利用实际数据进行验证,证明我们的分类器具有很强的准确性。

关键词 关键词1,关键词2,关键词3,关键词4,关键词5

引用格式 尚书林,左年明,张哲.《自动化学报》稿件加工样本.自动化学报,201X,XX(X):X-X

DOI 10.3724/SP.J.1004.201x.xxxxx

Preparation of Papers for Acta Automatica Sinica

SHANG Shu-Lin^{1,2} ZUO Nian-Ming² ZHANG Zhe^{3,4}

Abstract An abstract should be a concise summary of the significant items in the paper, including the results and conclusions. It should be about 5% of the length of the article, but not more than 200 words. Define all nonstandard symbols, abbreviations and acronyms used in the abstract. Do not cite references in the abstract.

Key words Open Set, 深度学习, 支持向量机, 辐射源识别, 模糊函数

Citation Shang Shu-Lin, Zuo Nian-Ming, Zhang Zhe. Preparation of papers for Acta Automatica Sinica. *Acta Automatica Sinica*, 201X, XX(X): X-X

随着科学技术的进步,现代战场形势瞬息万变,信息对抗在现代军事中的作用越来越重要。纵观整个20世纪所爆发的两次世界大战和数次局部战争、21世纪初的美阿、美伊之战以及最近闹得沸沸扬扬的韩国的萨德事件,无一不昭示着现代战争已成为电子战的“天下”,电子战技术也在历次实战演练中逐渐成熟。电子战也称电子对抗,包括电子侦察、电子攻击和电子防护三个方面。电子侦察主要

指从敌方雷达及其武器系统获取有用信息,通过雷达辐射源个体识别,可以对战场环境中敌我双方雷达辐射源的分布情况实施侦察,提供更加全面的、精确的电磁斗争与武器的态势,进行有效的战场指挥与决策,雷达辐射源个体识别已成为当前电子战特别是电子侦察领域的研究热点和难点。然而由于辐射源的特征未知、信号波形日趋复杂、战时电磁环境恶劣,给辐射源的精确识别带来了越来越严峻的挑战。

在雷达辐射源信号特征挖掘方面,已有很多学者作了大量研究工作,在上世纪70年代国外相关研究人员就开始了该部分的研究,该部分研究可以分为两个阶段:第一阶段为辐射源基本参数特征研究。对于原始信号特征直接求取其载波频率、脉冲宽度、脉冲幅度、到达角度和到达时间等信息,利用其中一个或多个作为特征向量。这种情况主要是应用于电磁环境相对单一、辐射源类别较少、信号形式单一、雷达参数固定的早期。

第二阶段自20世纪90年代以来,西方的军事强国开始研究雷达辐射信号的脉内特征,相继提出了多种分析雷达信号脉内特征的方法。有代表性的工作有:时域波形分析法、谱相关法、基于专家知识信号处理法、时频综合法、小波分析法、信息理论准则与聚类技术综合法、脉内瞬时频率特征与累积法、

收稿日期 XXXX-XX-XX 录用日期 XXXX-XX-XX
Manuscript received Month Date, Year; accepted Month Date, Year

国家重点基础研究发展计划(973计划)(XXXXXX),国家高技术研究发展计划(863计划)(XXXXXX),国家自然科学基金(XXXXXX)资助
Supported by National Basic Research Program of China (973 Program)(XXXXXX), National High Technology Research and Development Program of China (863 Program)(XXXXXX), National Natural Science Foundation of China (XXXXXX)

本文责任编辑 XXX

Recommended by Associate Editor BIAN Wei

1. 中国科学院自动化研究所高技术创新中心 北京 100190 2. 中国科学院自动化研究所模式识别国家重点实验室 北京 100190 3. 中国科学院自动化研究所《国际自动化与计算杂志》编辑部 北京 100190 4. 中国科学院自动化研究所《自动化学报》编辑部 北京 100190

1. Hi-Tech Innovation Center, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 2. National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 3. Editorial Office of *International Journal of Automation and Computing*, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 4. Editorial Office of *Acta Automatica Sinica*, Institute of Automation, Chinese Academy of Sciences, Beijing 100190

信号的分型特征等。

国内对雷达辐射源个体识别技术的研究始于上世纪80年代初,虽然起步较晚,但受到了军方的高度重视,在“九五”、“十五”和“十一五”国防预研中给予了大力资助。在脉内特征挖掘方面,毕大平提出易于工程实现的脉内瞬时频率提取技术;张葛祥提出了雷达辐射源信号的小波包特征、相像系数特征、熵特征、粗集理论、信息维数和分形盒维数;朱明提出了基于原子分解的特征、基于Chirplet原子的特征、时频原子特征;普运伟提出了瞬时频率派生特征、模糊函数主脊切面特征;陈稻伟提出了符号化脉内特征、围线积分双谱特征等;余志斌提出的局域波分解、小波脊频级联特征。

另一方面,雷达辐射源识别是一个典型的分类问题,其主要思路为在得到辐射源信号的特征表示之后,借助有效的分类算法来实现特征空间到决策空间的转换,从而确定信号的所属类别。大量的分类算法被成功运用于雷达辐射源识别中,如模板匹配、神经网络、支持向量机等。一般被应用于该领域的有三种分类方法,一种是判别型分类器,其需要在学习过程中最优化某种目标函数;另一种为生成模型分类器,其主要是基于先验概率和类别条件概率密度进行估计,如线性判别分类器、K最近邻等;第三种是决策树分类算法,通过人类专家的先验知识进行分类,如ID3、C4.5算法。

在跟踪发达国家研究进展的基础上,国内也开始加紧研发可装备于军方的SEI系统,以逐步缩小与发达国家在技术实用方面的差距。但总体来说,我国在该方面的研究还赶不上美英等发达国家,除了缺乏深层次且系统化的理论分析研究外,更缺乏成熟的工程化系统实践和实际装备,因而迫切需要加快在该领域的理论研究和应用步伐。本文综合雷达信号处理、深度学习等多学科理论,重点围绕在复杂电磁环境下不同辐射源的个体识别所面临的识别能力差等问题与挑战,提出合理的雷达脉内细微特征模型,结合深度学习的理论与方法,解决传统辐射源识别方法的局限性,为雷达辐射源识别提供理论支撑于技术指导。

Open Set目标识别系统必须可以准确的处理下面三种类型的数据类:

- 已知的(目标)类,被标记为正面训练样本的数据。
- 已知的未知(非目标)类,被标记为负面训练样本的数据。
- 未知的未知(非目标)类,在训练样本中不存在的类别的数据。

国内外对于Open Set的识别也早有研究,Simonson提出了一种称作probabilistic fusion(PF)的利用统计的方法来进行Open Set识别,其主要通过合并来自不同数据源的证据得到一个统计测试模型,根据此模型的分布来对于类别进行判断。Scheirer等人提出了一种通过分析后验数据得分来进行类型判断的方法。

1 辐射源信号分析

对于辐射源信号的处理,本文主要考虑两方面:信号预处理、特征提取优化。

在信号预处理方面,首先需要剔除无用和错误的数据。然后将信号进行分选,其主要是从随机交叠的脉冲信号流中分离出各个雷达的脉冲信号并选出有用信号。其实质是去交叠、去交错,所利用的是同一部雷达信号参数的相关性和不同雷达信号参数的差异性。

在特征提取优化方面,合理的特征是分类识别的基础。本文以模糊函数为平台,分析提取其切片特征。

1.1 模糊函数

模糊函数不仅能描述雷达信号的分辨特性和模糊度,还能描述由雷达信号所决定的测量精度和杂波抑制特性等,并根据这种雷达无意调制产生的信号脉内细微特征来进行分类所需的特征。对于信号 $x(t)$,其瞬时自相关函数为

$$R_x(t, \tau) = x(t + \tau/2)x^*(t - \tau/2) \quad (1)$$

,其中 τ 为时延,模糊函数的定义为,

$$A(\tau, \nu) = \int_{-\infty}^{+\infty} R_x(t, \tau) e^{j2\pi\nu t} dt \quad (2)$$

即 $R_x(t, \tau)$ 关于时间 t 的傅里叶反变换。由于在实际中发射机自身存在相位噪声以及各类杂散输出,故可以区分出型号、参数均相同的辐射源,通过模糊函数在时延和频偏这二维上的变换,可以多角度的刻画出无意调制对于发射信号的影响。

为了方便在数字信号中使用,上式可以经过变换等价于下面的形式:

$$A(\tau, \nu) = \int_0^\tau x(t)x^*(t + \tau)e^{j2\pi\nu t} dt \quad (3)$$

对信号均匀采样,即对接收信号和参考信号离散化后,上式可以表示为:

$$A(\tau_l, \nu_m) = A(l, m) = \sum_{n=0}^{N-1} x(n)x^*(n + l)e^{\frac{j2\pi mn}{N}} \quad (4)$$

其中, $\tau_l = l/f_s, \nu_m = mf_s/N$ 。

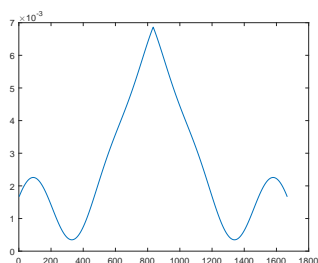


图1 模糊函数切片图

Fig. 1 Ambiguity Function

2 Open Set 分类器设计

此部分主要解决的问题是当得到一个新的测试样本, 如果该样本不属于已经经过训练的分类, 那么传统的神经网络模型会将该样本指派给与其最相似的一个类别, 此种情况对于一个Open Set识别系统, 也即类似于辐射源识别系统这种具有较多尚未经过训练的样本的一个数据集, 首先这会导致其识别率下降, 另一方面是由于对于未知的辐射源无法很好的确定, 无法很好的完成预警等任务。国内外学者对于该问题的研究主要分为下面两个思路:

- 在训练集中添加一个“未知”类别, 利用不同的来自非已知类别的数据作为训练样本对该类别进行训练, 然后对于所有的输入数据进行类别的识别, 对于识别结果为该类别的数据作为未知分类。然而该思路最大的问题是我们无法得到所有可能的未知类别的样本来进行训练, 具有一定的局限性。
- 针对于多分类使用的softmax函数, 可以设立阈值或者对于该识别结果进行一个评价(例如与已知类别数据的一个“距离”)等进行分辨出未知的分类。

针对于该问题, 我们基于思路2的想法设计了一个基于meta-recognition的可以识别未知辐射源

的深度神经网络。首先是创建一个深度卷积神经网络分类器, 该分类器的输出为该训练样本属于各个类别的概率, 我们然后将此类别作为一个输入, 输入到我们的meta-recognition中, 这里我们设计一个支持向量机分类器作为meta-recognition, 然后从该meta-recognition会进行判断该输入是否为一个未知分类。



图2 Open Set 分类器设计结构图

Fig. 2 Open Set Recognition Framework

2.1 深度卷积神经网络分类器设计

在分类器设计方面, 本文设计利用卷积神经网络的分类器。典型的卷积神经网络由深层结构堆叠在一起的多个不同的层组成: 输入层, 多组卷积和池化层, 有限数量的完全连接的隐藏层, 以及输出层。其中最主要的部分为卷积层。其利用输入数据中的局部结构, 将整个输入空间划分成很小的隐藏单元。将各个隐藏单元的权重构建得到的卷积核作用于整个输入空间, 从而得到特征向量。利用这种机制, 我们不仅大大减少了参数数量同时提高了数据的平移不变性。据辐射源信号的实际数据以及其反映出来的特性, 构建了基本的具有13层的卷积神经网络, 每层具有多个特征向量, 每个特征向量具有多个神经元, 并且每个特征向量来自于一种卷积核所提取输入的一种特征。主要过程为对于输入的辐射源信号, 进行多次的卷积、池化操作, 进行再次特征提取, 然后通过BP网络进行训练。

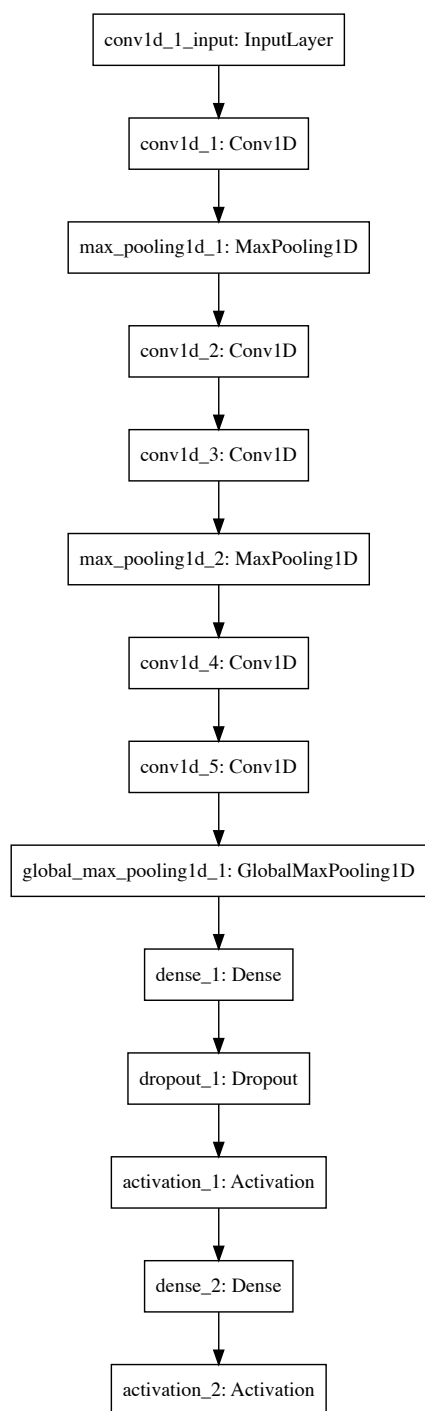


图3 深度卷积神经网络框架图

Fig. 3 Deep Learning Recognition Framework

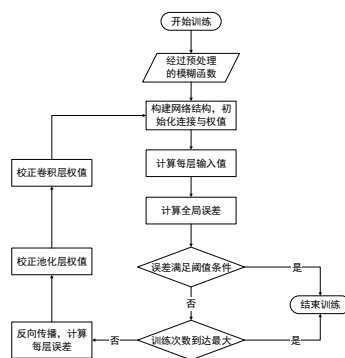


图4 深度卷积神经网络算法流程图

Fig. 4 Flow of Deep Learning Recognition

3 SVM meta-recognition 设计

支持向量机是一种流行的分类方法,因为它们可以在不需要大量数据的情况下产生良好的结果。我们可以利用所有的目标数据和未知目标的数据来作为训练样本对该SVM分类器进行训练,本部分我们以深度卷积神经网络的输出作为该分类器的输入,利用各类别的概率作为其特征进行训练识别。由于在类别的识别过程中,存在一定的波动性,这个会影响对于是否属于未知类别的分类判断,我们这里选取对于来自同一个辐射源的连续10拍的识别结果进行一个平均作为最终的输入。

下面是对于SVM分类器的设计,首先是核函数的选择。核函数将输入空间映射到高维特征空间,最终在高维特征空间中构造出最优分离超平面,从而把平面上本身不好分的非线性数据分开。常用的核函数为线性核函数和高斯核函数。对于核函数的选择,一般分为三种情况:

- Feature的数量很大,跟样本数量差不多,这时候选用LR或者是Linear Kernel的SVM
- Feature的数量比较小,样本数量一般,不算大也不算小,选用SVM+Gaussian Kernel
- Feature的数量比较小,而样本数量很多,需要手工添加一些feature变成第一种情况

由于我们的问题符合情况2,故选择高斯核函数。

支持向量机具有两个关键参数,惩罚参数 C 的和核参数 σ ,这两个参数的取值在很大程度上决定了SVM的性能的优劣。核函数的参数主要影响样本数据在高维特征空间中分布的复杂程度,即维数。特征子空间的维数越高,那么得到的最优分类超平面就会越复杂。反之亦然。因此只有选择合适的核参数得到合适的特征子空间,才能得到推广能

力良好的SVM分类器。本文中用到的是高斯核参数。大量的实验数据表明,如果与样本点之间的距离很小, $\sigma \rightarrow 0$; 如果与样本点之间的距离很大时, $\sigma \rightarrow \infty$; 当 σ 很小, 高斯核函数支持向量机得到的判别函数差不多是一个常数, 出现过拟合现象。当 σ 很大时, 样本的正确分类率也会比较低。

惩罚参数是影响SVM算法性能的另一个重要因素。它的作用主要是调节特征子空间中SVM模型的置信范围与经验风险的比例, 使支持向量机的泛化能力达到最好。特征子空间不同时, 最优参数值取值也会不同。惩罚参数与经验误差的惩罚和SVM的复杂度成正比, 与经验风险值成反比, 反之亦然。因此, 选择合适的惩罚参数也是非常重要的。

从上面的分析可以看出, 核参数影响着映射函数、进而影响样本子空间的复杂度。最后会影响分类机性能的好坏。惩罚参数作用是在数据子空间中调节学习机器自信区间的范围。这些都说明了惩罚参数和核参数的选择非常重要。

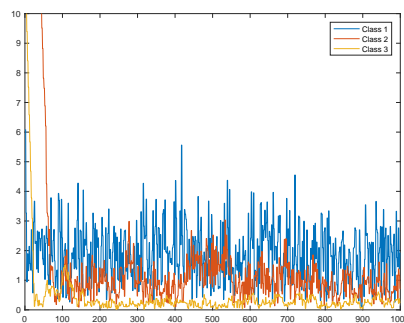


图5 不同类别样本特征图

Fig. 5 Comparison of Different Classes Data

4 仿真实验与分析

4.1 实验环境

由于我们原始获得的数据为IQ两路数据, 为了更好的捕捉到回波的特征信息, 我们对数据进行了一个变换, 求取其模糊函数, 并做偏移为0附件的一个切片。由于深度学习需要大量的数据进行训练学习, 而本身数据量偏少, 故我们在已有数据的基础上在一定信噪比的前提下, 生成部分仿真数据。

对于数据的选择方面, 我们从数据中选择出2至8个类别分别进行实验, 对于每一个类别, 我们均选择大约10000组数据, 其中70%作为训练样本, 20%作为交叉验证样本, 10%作为测试样本, 同时在测试样本中又添加了与已知分类等数量的未知分类的数据进行测试。

由于该变换后, 数据之间的差距比较大, 我们对数据进行了归一化。我们采用的归一化方法为min-max标准化 (Min-Max Normalization), 也称为离差标准化, 是对原始数据的线性变换, 使结果值映射到[0, 1]之间。转换函数如下:

$$x^* = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (5)$$

其中 x_{max} 为样本数据的最大值, x_{min} 为样本数据的最小值。

我们采用网格搜索法对SVM参数进行调优, 最终选择参数惩罚参数为32, 核参数 σ 为0.0312。

4.2 实验结果分析

在上述样本的情形下, 我们通过选取不同的类别数目, 进行训练和测试得到下表的识别结果。从表中数据我们可以看出, 随着样本类别数的增加, 对于未知分类的识别准确率也随之有了大幅度的增加, 而另一方面随着类别的增加, 对于每个类别的识别准确率有一定的降低, 但是仍然维持在比较高的水平。

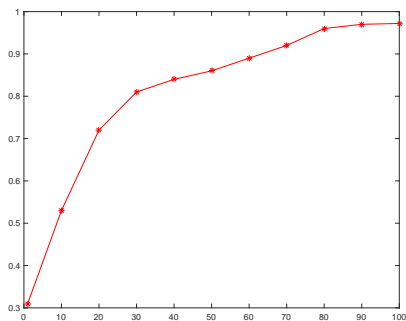


图6 迭代次数与识别准确率曲线图

Fig. 6 Comparasion of epoches and recogrization rate.

表1 不同样本类别下的识别正确率

Table 1 The recognition rate in different classes.

已知类别识别正确率	未知类别识别正确率
99.55%	84.32%
98.50%	93.10%
98.56%	97.81%
98.48%	98.42%
96.32%	98.85%
96.26%	99.22%
96.08%	99.14%

5 结束语

本文针对复杂电磁环境下辐射源的识别面临的电磁信号干扰大、雷达信号参数相近等问题与挑战，利用深度学习的思想与方法，深入研究辐射源脉内细微特征，设计合适的神经网络结构，并基于实际机载气象雷达数据进行初步验证。主要特色与创新点如下：

- (1) 利用深度学习方法进行辐射源识别前沿。通过对现有辐射源信号进行分析，利用其脉内细微特征作为训练样本，使得识别准确率有了较大的进步。虽然已有研究利用神经网络、支持向量机等机器学习算法进行识别，但是仍然需要基于雷达信号的基本参数，没有考虑信号的内部特征参数。
- (2) 本文采用方法具有较强的抗噪声、抗干扰能力。传统方法进行辐射源个体识别前均需进行降

噪、多径抑制和分选等复杂的信号预处理工作，这些操作会在一定程度上削弱雷达的个体特征。深度学习方法可以通过大量的样本，智能地判断各特征的权重，通过赋予不同的权重在保留雷达个体特征的情况下，避免干扰的影响。由此可见，本文所运用的方法具有较好的鲁棒性。

我们希望未来在目前的基础上，在下面几方面做进一步的研究：(1) 利用更多的数据对算法进行进一步的验证，目前类别较少的情况下，部分结果会对选择的类别具有一定的依赖性，另一方面是选取更多更合适的特征进行训练学习；(2) 利用集成学习提高学习精度；(3) 尝试新方法对于未知分类识别；(4) 对于未知分类，尝试利用无监督或者半监督学习进行训练学习。

致谢

文章作者可以在此处对支持和帮助其研究工作或本文写作的其他组织或人员表示感谢。注意,支持此研究的基金项目请直接标注在首页页脚处,不用在此再次致谢。

References

1 Ran B, Boyce D E. *Modeling Dynamic Transportation Network*. Berlin: Springer-Verlag, 1996. 69–83

2 Payton D, Estkowski R, Howard M. Compound behaviors in pheromone robotics. *Robotics and Autonomous Systems*, 2003, **44**(3): 229–240

3 Su Lian-Cheng, Zhu Feng. Design of a novel omnidirectional stereo vision system. *Acta Automatica Sinica*, 2006, **32**(1): 67–72
(苏连成, 朱枫. 一种新的全向立体视觉系统的设计. *自动化学报*, 2006, **32**(1): 62–72)

4 Roychoudhury R, Bandyopadhyay S, Paul K. Adistributed mechanism for topology discovery in ad hoc wireless networks using mobile agents. In: *Proceedings of IEEE First Annual Workshop on Mobile and Ad hoc Networking and Computing*. New York, USA: IEEE, 2000. 145–146

5 Hryniewicz O. An evaluation of the reliability of complex systems using shadowed sets and fuzzy lifetime data. *International Journal of Automation and Computing*, to be published

6 Zhang W. Reinforcement Learning for Job-Shop Scheduling [Ph. D. dissertation], Peking University, 1996

7 The Math Works. Image Processing Toolbox for Use with Matlab: User's Guide [Online], available: <http://www.mathworks.com>, November 3, 2006

8 Reily R C, Mack J L. The Self-organization of Spatially Invariant Representations, Technical Report PDP.CNS.92.5, Department of Psychology, Mellon University, USA, 1993

9 Wilkinson J P. Nonlinear Resonant Circuit, U.S. Patent 362412, July 1990

10 IEEE Criteria for Class IE Electric System, IEEE Standard 308, 1969