

一种深度强化学习的雷达辐射源个体识别方法

冷鹏飞, 徐朝阳

(中国船舶重工集团有限公司 第 723 研究所, 江苏 扬州 225001)

摘要: 针对传统依赖于人工经验提取辐射源个体特征的不足, 提出一种基于深度强化学习的雷达辐射源个体识别方法。利用发射机非理想信道造成的辐射源信号包络在信号变化时呈现的不同瞬态信息, 以信号包络前沿作为深度神经网络的输入状态, 以辐射源类别作为当前输入状态的可选动作, 通过卷积神经网络自动提取辐射源包络个体特征, 并拟合当前状态动作对的 Q 值, 进而以强化学习模型完成雷达辐射源个体识别任务。讨论了深度 Q 网络模型、深度双 Q 网络模型以及 Dueling Network 模型 3 种深度强化学习模型在辐射源识别任务中的应用。实测数据仿真实验表明: 传统机器学习算法的识别率不足 80%, 而深度强化学习网络的识别率高达 98.42%。

关键词: 雷达; 辐射源个体识别; 深度神经网络; 强化学习

中图分类号: TN971⁺.1 **文献标志码:** A **文章编号:** 1000-1093(2018)12-2420-07

DOI: 10.3969/j.issn.1000-1093.2018.12.016

Specific Emitter Identification Based on Deep Reinforcement Learning

LENG Peng-fei, XU Chao-yang

(No. 723 Institute, China Shipbuilding Industry Corporation, Yangzhou 225001, Jiangsu, China)

Abstract: A specific emitter identification (SEI) method based on deep reinforcement learning is proposed on account of the deficiency of emitter individual feature extraction depending on artificial experience. Due to the differences of the transient information of signal envelope, which results from the change of the signal owing to a nonideal transmitter channel, an envelope rising edge is used as the input state of deep neural network, and the emitter classifications are used as the optional actions of the current input state. The envelope features are extracted automatically through the convolutional neural network (CNN), and Q values of the current state action pairs are fitted, thus completing the specific emitter identification task based on the reinforcement learning model. The applications of deep Q network (DQN), deep double Q network (DDQN) and Dueling network in the specific emitter identification are discussed. The measured results show that the recognition rate of traditional machine learning algorithm is less than 80%, but the deep reinforcement learning model can achieve the high recognition rate of 98.42%.

Key words: radar; specific emitter identification; deep neural network; reinforcement learning

收稿日期: 2018-05-07

基金项目: 装备预研船舶重工联合基金项目(2016 年)

作者简介: 冷鹏飞(1994—), 男, 硕士研究生。E-mail: ttl_eye@163.com

通信作者: 徐朝阳(1968—), 男, 研究员, 硕士生导师。E-mail: edaxcy@vip.sina.com

0 引言

在雷达对抗领域中,雷达辐射源识别技术是电子对抗情报分析领域中的重要研究内容,其水平是衡量电子侦察系统和侦察设备信息处理技术先进程度的重要标志^[1]。雷达辐射源个体识别技术通过脉内无意调制特征区分雷达的不同个体^[2],当前雷达辐射源识别问题的一个研究热点为提取辐射源的有效特征^[3],通常可将信号包络特征^[4]、高阶矩^[5]、时频特征^[6]以及模糊函数特征^[7]等参数作为辐射源的个体特征。这些特征均需基于先验知识人工提取,因而所提取特征未必能描述辐射源的本质特点,且人工提取特征过程繁琐,使得数据库更新缓慢。

近年来,深度神经网络的兴起使得机器自动提取特征成为现实,深度神经网络可逼近复杂的非线性函数,具有较强的泛化能力,能够刻画数据本质信息^[8],且现场可编程门阵列(FPGA)、图形处理单元(GPU)的加速处理方法使得运算的实时性得以保证。基于此,本文提出了一种深度强化学习的辐射源个体识别方法,利用不同雷达个体发射信号包络的差异实现辐射源识别,在实际电子对抗环境中,雷达辐射源脉冲宽度从微秒级到毫秒级变化,这种大范围的脉宽变化使得处理整个辐射源包络较为困难,考虑到工程应用中系统的实时性要求,本文将辐射源包络前沿(包络上升沿及其前后部分数据)作为个体特征进行辐射源识别。

1 辐射源包络

1.1 辐射源侦收

图 1 给出了一种电子侦察设备个体识别框图,为了兼顾宽带接收机侦察范围大、窄带接收机灵敏度高的优点,本文采用一种宽带、窄带接收机并行工作的方案。宽带接收机对信道带宽进行子信道划分,进而完成各频段辐射源信号的检测与参数估计以获取宽带脉冲描述字(PDW)样本集;窄带接收机分时侦收各频段辐射源信号,对射频前端中频输出进行采样、信号检测及参数估计以获取窄带 PDW 样本集;宽带、窄带 PDW 样本集经数据融合后再对其融合结果进行信号分选。信号分选后的辐射源描述字(EDW)将作为个体识别模块的控制信号,依据窄带 PDW 样本集选择感兴趣的辐射源脉冲 $s(n)$ 对其进行个体识别,其中 n 为离散时间。

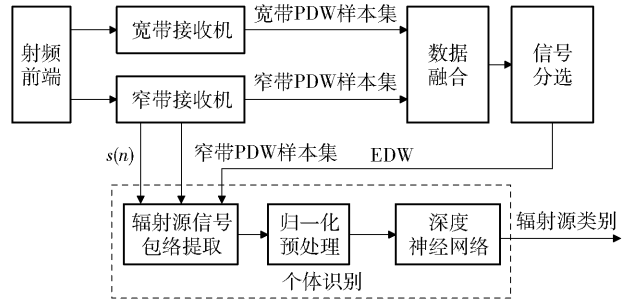


图 1 一种电子侦察设备个体识别框图

Fig. 1 Block diagram of specific emitter identification of electronic reconnaissance equipment

1.2 辐射源包络提取

设窄带接收机侦收辐射源基带信号为 $s(n)$, 则其包络可描述为

$$a(n) = \sqrt{|s(n)|^2 + |\text{Hilbert}[s(n)]|^2}, \quad (1)$$

式中: $\text{Hilbert}[\cdot]$ 为希尔伯特变换。对 $a(n)$ 进行滑动平均处理可获得平滑的包络曲线,以上升沿中点对齐信号包络,可截取一段包络前沿作为深度神经网络的输入。图 2 给出了两类雷达个体辐射源包络前沿,图 2 中每类辐射源有 10 个样本,接收机带宽为 200 MHz,采样频率为 1 GHz,信噪比 $SNR \geq 15$ dB。从图 2 中可以看出,两类辐射源包络上升沿形状、过冲点均存在差异,表明不同辐射源个体可通过包络前沿区分开来。

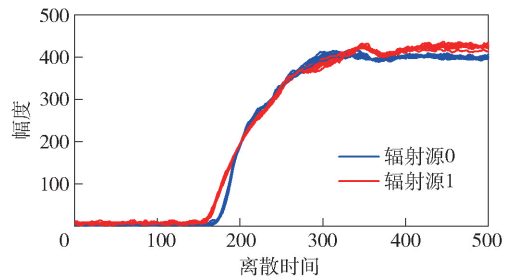


图 2 两类雷达个体辐射源包络前沿

Fig. 2 Envelope rising edges of different radar emitter individuals

2 深度强化学习

2.1 马尔可夫决策过程

讨论强化学习模型在 C 类已知辐射源识别问题中的应用。强化学习任务通常用马尔可夫决策过程(MDP)来描述,MDP 对应四元组 $\langle S, A, p, r \rangle$, 其中: S 为输入状态集合, $s \in S$ 为某个特定状态; A 为可选动作集合, $a \in A$ 为某个特定动作; p 为当前

状态 s 在机器执行动作 a 后转化为状态 s' 的概率; r 为当前状态 s 在机器执行动作 a 后转化为状态 s' 时所获得的奖励。MDP 具有马尔可夫性质; 系统下一时刻状态只与当前时刻状态有关, 与过往时刻状态无关。在所述辐射源识别任务中, 状态 s 为电子侦察设备从外部电磁环境侦收提取的辐射源包络前沿; a 为辐射源类别编号 $0, 1, 2, \dots, C-1$; 奖励 r 依据模型识别结果与雷达数据库中样本真实类别的异同而定, 相同时 r 取 $+1$, 不同时 r 取 -1 ; 对于状态转移概率 p 而言, 尽管下一时刻辐射源包络前沿 s' 与当前时刻状态动作对 (s, a) 无关, 但是在网络训练过程中, 为了防止过拟合, 通常会随机乱序训练样本且各训练样本数目分布均匀, 因此可以认为状态转移概率 $p = 1/C$, 且 s' 仅在状态 s 下执行动作 a 时以概率 p 转移而得, 与过往时刻的状态无关, 即满足马尔可夫性质。

2.2 策略

在强化学习中, 对于当前输入状态 s , 机器在某种策略 π 的指引下选择动作 $a = \pi(s)$, 机器学习的目的即为求得最优策略 π^* 。通常可用状态-动作值函数 (即 Q 值函数) 评估策略, 其数学模型为

$$Q(s, a) = E \left[\sum_{t'=t}^{+\infty} \gamma^{t'-t} r_{t'+1} \mid s, a, \pi \right], \quad (2)$$

式中: γ 为折扣系数, 其值在 $0 \sim 1$ 之间, γ 的存在是为了确保长期累积奖励能够收敛。

(2) 式描述了从状态 s 出发, 执行动作 a 后使用策略 π 所带来的期望累积奖励, 对其进一步推导可得 Bellman 方程为

$$Q(s, a) = E_{s'} [r + \gamma E_{a' \sim \pi(s')} [Q(s', a')] \mid s, a, \pi], \quad (3)$$

式中: s' 为状态 s 下执行动作 a 后转移到的新状态; r 为状态 s 执行动作 a 后转移至 s' 时获取的奖励。强化学习任务中策略的改进与值函数的改进是一致的^[9], 根据 (3) 式求得 $Q(s, a)$ 后应以策略 $\pi' =$

$\operatorname{argmax}_a Q(s, a)$ 替换策略 π 。机器学习得最优策略时满足最优 Bellman 方程为

$$Q^*(s, a) = E_{s'} [r + \gamma \max_{a'} Q^*(s', a') \mid s, a], \quad (4)$$

式中: $Q^*(s, a)$ 为使用最优策略 π^* 时所产生的期望累积奖励。(4) 式明确了机器学习的方向, 即使得 Q 值函数收敛于最优 Q 值函数 $Q^*(s, a)$ 的方向。

在雷达辐射源识别任务中, 状态空间 S 中的状态数目无限, 无法用有限空间存储所有状态动作对的 Q 值, 深度神经网络具有很强的函数泛化能力, 可用于 Q 值函数的拟合, 将其记为 $Q(s, a; \theta)$, 其中 θ 为神经网络的未知参数向量。当前状态 s 在执行动作 a 后以一定概率转移到状态 s' , 令 $r + \gamma \max_{a'} Q(s', a'; \theta)$ 为目标 Q 值函数, 采用 ε -贪心策略, 以概率 ε 选择当前状态 s 下使 Q 值函数最大的动作, 以概率 $1 - \varepsilon$ 随机选择动作得到 $Q(s, a)$, 于是原机器学习问题转换为如下无约束最优化问题:

$$\theta = \operatorname{argmin}_{\theta} E_{\pi} [|r + \gamma \max_{a'} Q(s', a'; \theta) - Q(s, a; \theta)|^2]. \quad (5)$$

(5) 式使得 Q 值函数依据 ε -贪心策略在最小均方差意义下收敛于最优 Q 值函数 $Q^*(s, a)$ 。

在深度学习任务中, 输入向量的维数一般较高, 即使最优化问题中损失函数在某点处的梯度为 0, 其在该点所有分量上呈现出相同凹凸性的概率亦几乎为 0, 因此, 深度学习中损失函数梯度为 0 的点更多地考虑为鞍点而非局部最优点, 由于梯度噪声的存在使得网络能够跳出鞍点, 因而 (5) 式可用随机梯度下降类算法求解。

3 具体网络模型

3.1 深度 Q 网络模型

本文讨论 10 类已知雷达辐射源的个体识别问题, 输入辐射源包络前沿长度为 1 024, 以卷积神经网络^[10]拟合 Q 值函数, 图 3 给出了其模型, 该模型

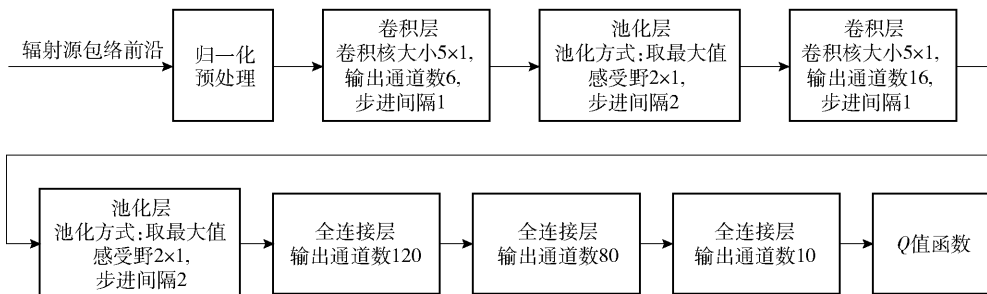


图 3 Q 值函数拟合 DQN 模型

Fig. 3 DQN model of Q value function fitting

为深度 Q 网络 (DQN) 模型^[11], 由 2 层卷积池化层、3 层全连接层组成。第 1 层卷积层卷积核大小为 5×1 , 提取辐射源包络前沿 6 个底层特征, 第 2 层卷积层进一步提取包络前沿高层特征, 池化层使用某一位置相邻输出的总体统计特征代替网络在该位置的输出^[12], 通过对卷积层输出特征图中相邻像素点取最大值的方式实现池化运算, 使网络在保留辐射源本质特征信息的同时可以减少数据量。全连接层用于将这些特征连接起来, 多层全连接层级联能够增强网络的学习能力。

对于辐射源包络前沿 s , 经过网络正向传播后即可得到当前状态 s 下执行各动作时的 Q 值, 即 $Q(s, a; \theta)$, 其中 a 为整数 $0, 1, \dots, 9$ 。

经过训练后, 网络将收敛至最优 Q 值函数。由于目标 Q 值函数与 Q 值函数之间存在相关性, 深度强化学习模型在训练时难以收敛, 为了缓解这个问题, 可在 Q 值函数拟合网络多次迭代后更新一次目标 Q 值函数以降低其相关性, 即将 (5) 式变为

$$\theta = \underset{\theta}{\operatorname{argmin}} E_{\pi} [|r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta) |^2], \quad (6)$$

式中: 参数向量 θ^- 在多次迭代后得到参数向量 θ 。

3.2 深度双 Q 网络模型

DQN 模型以卷积神经网络拟合 Q 值函数, 进而通过梯度下降类算法求解 (6) 式最优化问题, 分析 (6) 式可知, 其目标 Q 值函数采用同样的拟合函数选择最优动作并对其进行评估, 即有

$$y = r + \gamma \max_{a'} Q(s', a'; \theta^-) = r + \gamma Q(s', \operatorname{argmax}_{a'} Q(s', a'; \theta^-); \theta^-). \quad (7)$$

在这个过程中, θ^- 并没有更新, 这种方式会导致 Q 值的高估, 若这种过高的估计对于所有潜在决策是不一致的, 则可能导致策略选择次优解^[13]。对于状态空间连续取值的强化学习任务来说, 有限的学习样本不能使神经网络拟合出适用于所有状态动作对的 Q 值函数, 因此所拟合的 Q 值函数曲线会在真实 Q 值曲线周围波动。由于目标 Q 值函数值求解过程中需要选择, 使得 Q 值达到最大的动作, 网络估计的 Q 值可能比真实 Q 值高。深度双 Q 网络 (DDQN) 模型采用两个卷积神经网络, 分别用于选择最优动作和策略评估, 缓解了这个问题, 其目标 Q 值数学模型为

$$y = r + \gamma \max_{a'} Q(s', a'; \theta^-) =$$

$$r + \gamma Q(s', \operatorname{argmax}_{a'} Q(s', a'; \theta); \theta^-). \quad (8)$$

在实际应用时, 与 (6) 式中降低目标 Q 值函数与 Q 值函数相关性的方法类似, 可以用一个网络实现, 即参数向量 θ^- 为参数向量 θ 多次迭代前的向量。

3.3 Dueling Network 模型

与 DQN 模型不同, Dueling Network 模型^[14] 将 Q 值函数分解为状态值函数 $V(s)$ 与动作优势函数 $A(s, a)$, 其数学模型为

$$Q(s, a; \theta, \alpha, \beta) = V(s; \theta, \alpha) + A(s, a; \theta, \beta) - \max_{a'} A(s, a'; \theta, \beta), \quad (9)$$

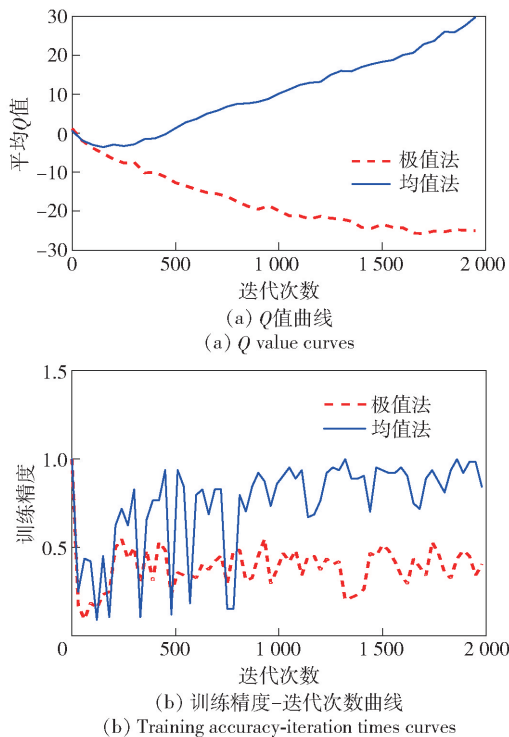
$$Q(s, a; \theta, \alpha, \beta) = V(s; \theta, \alpha) + A(s, a; \theta, \beta) - \frac{1}{C} \sum_{a' \in A} A(s, a'; \theta, \beta), \quad (10)$$

式中: α, β 分别为状态值函数与动作优势函数全连接层的参数向量。Dueling Network 模型与 DQN 模型的区别在于最后一层全连接层, DQN 模型直接拟合 $Q(s, a)$, 而 Dueling Network 最后一层由并联的两个全连接层组成, 拟合 $A(s, a)$ 的全连接层与 DQN 模型相同, 输出 C 个通道, 而拟合 $V(s)$ 的全连接层只有一个输出通道。为了便于论述, 记 (9) 式为“极值法”、(10) 式为“均值法”。网络收敛后, 极值法中 $V(s)$ 、 $A(s, a)$ 能够给出状态值函数及动作优势函数的估计^[14], 但是 (9) 式使得网络难以收敛; 均值法中 $V(s)$ 、 $A(s, a)$ 在网络收敛后与实际估计存在一个常数偏差, 但 (10) 式能够保证网络稳定收敛。图 4 给出了网络训练过程中两种 Q 值拟合方法的 Q 值曲线以及训练精度-迭代次数曲线 (具体仿真实验参数见第 4.1 节、4.2 节)。为了清楚地对比两种方法的性能, 图 4 中曲线为实际曲线均匀降采样的结果, 图 4 中平均 Q 值为 64 个辐射源包络前沿 Q 值的平均 (动作为使 Q 值最大的动作)。从图 4 中可以看出, 均值法 Q 值、训练精度稳定上升, 且训练精度逐渐收敛。 Q 值曲线仍呈现上升趋势的原因是 (2) 式中 γ 为 0.99, 理想情况下网络收敛时奖励 r 恒等于 1, 因而理论上最优 Q 值为 100, 在分类识别任务中, 长时间的迭代会导致模型过拟合, 因此当模型训练精度满足工程需求时即可停止训练。

4 实测数据仿真实验

4.1 实验环境

实验平台为 Ubuntu Linux16.04, 深度学习模

图4 两种 Q 值拟合方法性能对比图Fig. 4 Performances of two Q value fitting methods

型架构为 TensorFlow, 程序接口为 Python 2.7, 图形处理器为 NVIDIA GeForce GTX 1080Ti. 为了确保深度学习模型可靠泛化, 主要考虑 3 个方面问题:

- 1) 模型对不同型号、不同雷达个体辐射源信号的识别能力;
- 2) 模型对相同型号、不同雷达个体辐射源的识别能力;
- 3) 模型对不同接收系统所侦收辐射源的识别能力。

为此, 样本集中含同型号导引头辐射源信号 4 类、同型号机械扫描雷达辐射源信号 4 类、同型号雷达模拟器辐射源信号 2 类, 共 10 类辐射源样本, 每类样本数据量为 6 000, 其中训练集样本数量为 54 000, 测试集样本数量为 6 000, 训练集与测试集中各辐射源类别数量分布均匀。雷达辐射源经不同侦收系统采集, 利用抽取、插值等处理方法将信号采样率统一为 1 GHz, 对齐包络后取 1 024 个离散样本点作为辐射源包络前沿数据。图 5 为某雷达个体辐射源包络前沿 (10 个样本)。从图 5 中可以看出, 样本集中相同雷达个体辐射源包络前沿间亦存在差异。因此, 可用模型的测试精度评估其泛化能力。

4.2 模型训练

模型训练步骤如下:

- 1) 采用 Xavier 方法^[15]初始化拟合 Q 值函数的

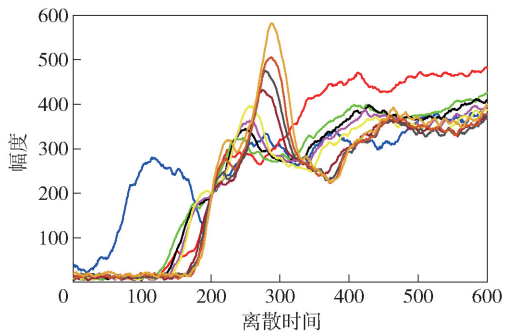


图5 某雷达个体辐射源包络前沿

Fig. 5 Envelope rising edges of a specific emitter

网络参数向量 θ ;

- 2) 初始化拟合目标 Q 值函数的网络参数向量 θ^- ;

- 3) 对于当前输入状态 (辐射源包络前沿) s , 利用 ε -贪心算法选择动作 a (某个类别);

- 4) 若 a 与雷达数据库中的样本真实类别相同, 则获取奖励 $r = +1$, 否则 $r = -1$;

- 5) 载入下一时刻状态 s' , 利用 (8) 式计算目标 Q 值函数 y (DQN 模式为 (7) 式);

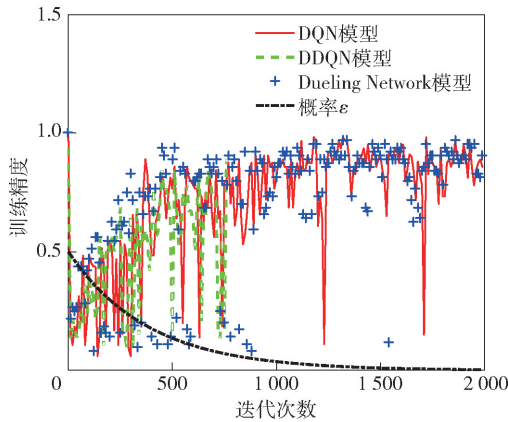
- 6) 对 $|y - Q(s, a; \theta)|^2$ 使用梯度下降类算法更新网络参数 θ ;

- 7) 每 D 次迭代使得 $\theta^- = \theta$;

- 8) 重复步骤 3 ~ 步骤 7, 直至网络收敛。

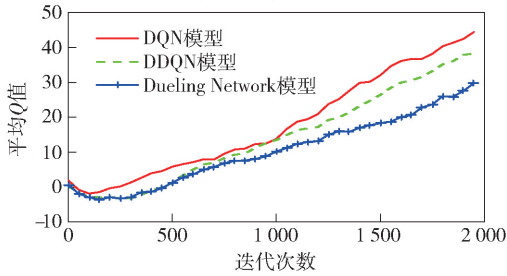
模型训练参数如下: 学习率为 0.000 25, mini-batch 大小为 64, 迭代次数为 2 000, 参数更新周期 D 为 15, 概率 $\varepsilon = 0.5 \times 0.02^{i/1500}$ (i 为当前迭代次数), 折扣系数 γ 为 0.99, 采用 Adam 算法训练网络, 其中梯度动量 β_1 为 0.9, 梯度平方动量 β_2 为 0.999, 引入误差 e 为 10^{-8} 以确保计算过程数值稳定。

图 6 为 3 种模型的训练结果, 为了清楚地对比 3 种模型的性能, 图 6 中曲线为实际曲线均匀降采样所得。从图 6 中可以看出, 3 种模型训练精度均已收敛, 且 DDQN 模型、Dueling Network 模型的 Q 值估计低于 DQN 模型。图 6 (a) 中存在许多训练精度瞬间下降的毛刺, 这是因为迭代初期机器不断选择非最优动作导致的, 图 6 (a) 中 ε -贪心策略的概率 ε 随着迭代次数的增加而降低, 在迭代初期 ε 较大, 机器会以较大概率选择非最优动作训练网络, 因而毛刺较为密集; 随着 ε 的减小, 这种现象得以缓解。事实上, 在测试时机器总是选择最优动作执行, 因此这些毛刺不会对网络测试精度造成影响。



(a) 训练精度-迭代次数曲线

(a) Training accuracy-iteration times curve



(b) Q值曲线

(b) Q value curve

图 6 3 种模型的训练结果

Fig. 6 Training results of three models

4.3 测试结果

测试实验对比了传统机器学习算法与深度神经网络的识别性能,其中, k 近邻(k -NN)算法的近邻数 k 取 20,以 L1 范数与 L2 范数两种方法度量待识别辐射源包络与已识别样本的距离,支持向量机(SVM)采用线性函数(linear function)与径向基函数(RBF)作为核函数,对辐射源包络进行识别。表 1 为各模型运算时间对比,其中训练平均时间为模型训练时 1 个 minibatch(使用梯度下降类算法求解最优化问题过程中单步迭代参与运算的样本数据,即 64 个辐射源包络前沿)平均每次迭代所消耗的时间,测试平均时间为测试模型识别能力时 1 个 minibatch 数据正向传播过程所消耗的时间,其值为多次实验的平均结果。分析表 1 中数据可知,3 种模型可实时实现且运算时间大致相当,这是因为 3 种模型中运算复杂度较高的深层网络结构相同,其差别仅在于 Q 值函数、目标 Q 值函数的拟合方式。在实际工程应用时,可将一台服务器专门用于模型的训练,并将训练完成的网络参数即时传输给 FPGA(或 PowerPC 等实时处理设备),最终的辐射源识别功能则交由 FPGA(或 PowerPC)完成。表 2 为测试精度

对比,从表 2 中可以看出,传统 k -NN 算法、SVM 算法识别率不足 80%,而本文所述 3 种神经网络模型识别率均在 96% 以上,且 Dueling Network 模式识别率高达 98.42%。

表 1 3 种模型的运算时间对比

Tab. 1 Operation times of different models

算法模型	训练平均时间/ (ms·minbatch ⁻¹)	测试平均时间/ (μs·minbatch ⁻¹)
DQN	7.26	324
DDQN	7.26	332
Dueling Network	7.28	358

表 2 7 种模型的测试精度对比

Tab. 2 Testing accuracies of different models

算法模型	识别率/%
k -NN(L1)	61.80
k -NN(L2)	62.20
SVM(RBF)	68.36
SVM(linear function)	76.81
DQN	96.53
DDQN	97.36
Dueling Network	98.42

5 结论

本文讨论了深度强化学习在已知雷达辐射源个体识别中的应用。从用于辐射源个体识别的电子侦察设备方案出发,分析了深度强化学习理论在辐射源识别中的具体表现形式,给出了 DQN 模型、DDQN 模型以及 Dueling Network 模型 3 种具体网络模型,实测数据仿真实验结果表明,相较于传统机器学习算法而言,深度强化学习算法能够在保证一定运算实时性的同时取得更高的识别率。

参考文献 (References)

- [1] 陈昌孝,何明浩,徐璟,等. 雷达辐射源识别技术研究进展[J]. 空军预警学院学报, 2014, 28(1):1-5,9.
CHEN Chang-xiao, HE Ming-hao, XU Jing, et al. Progress of study on recognition technology of radar emitter [J]. Journal of Air Force Early Warning Academy, 2014, 28(1):1-5,9. (in Chinese)
- [2] 周志文,黄高明,陈海洋,等. 雷达辐射源识别算法综述[J]. 电讯技术, 2017, 57(8):973-980.
ZHOU Zhi-wen, HUANG Gao-ming, CHEN Hai-yang, et al. A

- review of radar emitter recognition algorithm[J]. Telecommunication Engineering, 2017, 57(8):973-980. (in Chinese)
- [3] 周志文, 黄高明, 高俊, 等. 一种深度学习的雷达辐射源识别算法[J]. 西安电子科技大学学报, 2017, 44(3):77-82.
ZHOU Zhi-wen, HUANG Gao-ming, GAO Jun, et al. Radar emitter identification algorithm based on deep learning[J]. Journal of Xidian University, 2017, 44(3):77-82. (in Chinese)
- [4] Kawalec A, Owczarek R. Specific emitter identification using intrapulse data[C]//Proceedings of European Radar Conference. Amsterdam, the Netherlands; IEEE, 2004:249-252.
- [5] 王宏伟, 赵国庆, 王玉军. 基于脉冲包络前沿高阶矩特征的辐射源个体识别[J]. 现代雷达, 2010, 32(10):42-45.
WANG Hong-wei, ZHAO Guo-qing, WANG Yu-jun. Specific emitter identification based on higher order moment of the envelope's front edge[J]. Modern Radar, 2010, 32(10):42-45. (in Chinese)
- [6] 程吉祥, 张葛祥, 李志丹. 基于时频原子方法的雷达辐射源个体识别[J]. 航天电子对抗, 2011, 27(1):54-57.
CHENG Ji-xiang, ZHANG Ge-xiang, LI Zhi-dan. A novel specific emitter identification method based on time-frequency atom approach[J]. Aerospace Electronic Warfare, 2011, 27(1):54-57. (in Chinese)
- [7] Wang L, Ji H B, Shi Y. Feature extraction and optimization of representative-slice in ambiguity function for moving radar emitter recognition[C]//Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing. Dallas, TX, US; IEEE, 2010:2246-2249.
- [8] Schmidhuber J. Deep learning in neural networks: an overview [J]. Neural Network, 2014, 61(1):85-117.
- [9] 周志华. 机器学习[M]. 北京:清华大学出版社, 2016:380-381.
ZHOU Zhi-hua. Machine learning[M]. Beijing: Tsinghua University Press, 2016:380-381. (in Chinese)
- [10] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//Proceedings of International Conference on Neural Information Processing Systems. Spain: Neural Information Processing Systems Foundation, 2012:1097-1105.
- [11] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning [J]. Nature, 2015, 518(7540):529-541.
- [12] Bengio Y I, Goodfellow J, Courville A. Deep learning [M]. Cambridge, MA, US; MIT Press, 2016:340-341.
- [13] van Hasselt H, Guez A, Silver D. Deep reinforcement learning with double Q-learning[C]//Proceedings of the 30th AAAI Conference on Artificial Intelligence. Phoenix, AZ, US; AAAI, 2016:2094-2100.
- [14] Wang Z, Schaul T, Hessel M, et al. Dueling network architectures for deep reinforcement learning[C]//Proceedings of International Conference on Machine Learning. New York, NY, US: PMLR, 2016:1995-2003.
- [15] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks [J]. Journal of Machine Learning Research, 2010, 9: 249-256.