



DOI: 10.12086/oe.2019.180307

面向军事目标识别的 DRFCN 深度网络设计及实现

刘俊^{1*}, 孟伟秀¹, 余杰², 李亚辉¹, 孙乔¹¹杭州电子科技大学通信信息传输与融合技术国防重点学科实验室, 浙江 杭州 310018;²中国船舶重工集团公司第七一五研究院, 浙江 杭州 310023

摘要: 自动目标识别(ATR)技术一直是军事领域中急需解决的重点和难点。本文设计并实现了一种新的面向军事目标识别应用的 DRFCN 深度网络。首先,在 DRPN 部分通过卷积模块稠密连接的方式,复用深度网络模型中每一层的特征,实现高质量的目标采样区域的提取;其次,在 DFCN 部分通过融合高低层次特征图语义特征信息,实现采样区域目标类别和位置信息的预测;最后,给出了 DRFCN 深度网络模型结构以及参数训练方法。与此同时,进一步对 DRFCN 算法开展了实验分析与讨论: 1) 基于 PASCAL VOC 数据集进行对比实验,结果表明,由于采用卷积模块稠密连接的方法,在目标识别平均准确率、实时性和深度网络模型大小方面,DRFCN 算法均明显优于已有基于深度学习的目标识别算法;同时,验证了 DRFCN 算法可以有效解决梯度弥散和梯度膨胀问题。2) 利用自建军事目标数据集进行实验,结果表明,DRFCN 算法在准确率和实时性上满足军事目标识别任务。

关键词: 深度学习; 目标识别; PASCAL VOC 数据集; 军事目标

中图分类号: TP391.41; TB872

文献标志码: A

引用格式: 刘俊, 孟伟秀, 余杰, 等. 面向军事目标识别的 DRFCN 深度网络设计及实现[J]. 光电工程, 2019, 46(4): 180307

Design and implementation of DRFCN in-depth network for military target identification

Liu Jun^{1*}, Meng Weixiu¹, Yu Jie², Li Yahui¹, Sun Qiao¹¹Fundamental Science on Communication Information Transmission and Fusion Technology Laboratory, Hangzhou Dianzi University, Hangzhou, Zhejiang 310018, China;²China Shipbuilding Industry Corporation 715 Research Institute, Hangzhou, Zhejiang 310023, China

Abstract: Automatic target recognition (ATR) technology has always been the key and difficult point in the military field. This paper designs and implements a new DRFCN in-depth network for military target identification. Firstly, the part of DRPN is densely connected by the convolution module to reuse the features of each layer in the deep network model to extract the high quality goals of sampling area; Secondly, in the DFCN part, we fuse the information of the semantic features of the high and low level feature maps to realize the prediction of target area and location information in the sampling area; Finally, the deep network model structure and the parameter training method of DRFCN are given. Further, we conduct experimental analysis and discussion on the DRFCN algorithm: 1) Based on

收稿日期: 2018-06-04; 收到修改稿日期: 2018-08-07

基金项目: 海军装备预研创新项目; 国家自然科学基金重点项目(61333009, 61427808)

作者简介: 刘俊(1971-), 男, 博士, 教授, 主要从事信息融合、模式识别与智能系统、计算机视觉的研究。E-mail: gutlj@163.com

the PASCAL VOC dataset for comparison experiments, the results show that DRFCN algorithm is obviously superior to the existing algorithm in terms of average accuracy, real-time and model size because of the convolution module dense connection method. At the same time, it is verified that the DRFCN algorithm can effectively solve the problem of gradient dispersion and gradient expansion. 2) Using the self-built military target dataset for experiments, the results show that the DRFCN algorithm implements the military target recognition task in terms of accuracy and real-time.

Keywords: deep learning; target recognition; PASCAL VOC dataset; military target

Citation: Liu J, Meng W X, Yu J, *et al.* Design and implementation of DRFCN in-depth network for military target identification[J]. *Opto-Electronic Engineering*, 2019, **46**(4): 180307

1 引言

自动目标识别(Automatic target recognition, ATR)技术一直是军事应用领域中急需解决的重点和难点,其中光电探测是现代预警探测情报网中的关键探测途径之一。在实际作战中,通过部署在不同武器装备平台上的光电设备,可以获得海量图像、视频数据,如何在这些采集的数据中实现快速、精准的军事目标识别,提高作战系统的军事目标识别能力,具有十分重要的意义。得益于大数据、云计算和人工智能技术的发展,尤其是以深度学习算法为基础的智能目标识别技术的进步,利用机器智能对军事目标进行识别,越来越受到各国学术界和应用部门的高度重视。

从 20 世纪 50 年代开始,美国率先开展了基于图像的军事目标识别技术研究工作。从国外军事目标识别研究的技术路线及发展历史来看,目标识别的技术路线及发展历史可大致划分为六个阶段:经典的统计模式识别方法^[1]→基于知识的自动目标识别方法^[2]→基于模型的自动目标识别方法^[3]→基于多传感器信息融合的自动目标识别方法^[4]→基于人工神经网络和专家系统混合应用的自动目标识别方法^[5]→基于人工智能和深度学习的自动目标识别方法^[6]。其中,基于前 5 个发展阶段的技术路线,其实现的军事目标识别技术水平一般难以实现超越人的识别水平。对于最后一个发展阶段,即基于人工智能和深度学习的自动目标识别方法,近十年来,由于超大规模集成电路、超高速集成电路、神经网络、量子计算机技术、多传感器集成技术、人工智能技术、深度学习、大数据技术等的高速发展,国外军事目标识别研究已由理论探索、实验室仿真逐渐走向实际应用。我国基于深度学习的军事目标识别研究主要处于理论探索和实验室仿真阶段,有待逐渐突破并走向实际应用。

在理论研究方面,基于深度学习的目标识别算法

大致可以分为三类方法:1) 基于区域采样、目标分类、目标位置修正的方法,可称为基于区域采样的深度卷积目标识别算法,该类方法的优点是识别准确率高,缺点是算法模型庞大,训练耗时长,很多高准确率的算法模型需要多 GPU 的联合训练,算法识别速度慢,对硬件要求高,如全卷积目标识别(region-based fully convolutional networks, RFCN)^[7]算法模型;2) 基于回归的深度卷积目标识别算法,该类方法的优点是目标识别速度快,但识别准确率总体低于基于区域采样策略的目标识别方法,如 YOLO(you only look once)算法模型^[8];3) 前两类方法的组合方法,典型代表是 SSD(single shot MultiBox detector)算法模型^[9]。

ALEXNet^[6]验证了基于 CNN(convolutional-neural-networks)的深度学习方法能够大幅度提升分类的准确率。2012 年 Cirshick 在 ALEXNet 基础上进一步提出 R-CNN(regions with CNN features)^[10]方法,第一次实现基于 CNN 方法的自动目标识别,但其缺点是由于采用分段级联模型不能共享计算;特征向量需要海量存储空间,存在时空冗余;输入图片需固定尺寸。进一步文献[11]提出 SPP(spatial pyramid pooling)算法,在模型底层采用空间金字塔池化方式^[12]将不同大小的特征图聚合归一化到特定维度,在特征图上获取采样区域的方法共享了卷积层的计算量,大幅提高网络的训练和测试速度,然而由于分类和坐标回归需要两个代价函数,不能实现端到端的训练和测试。随后, Cirshick 在 SPP 算法的基础上提出了 Fast R-CNN(fast regions with CNN features)算法模型^[13],采用联合代价函数实现了端到端的训练与测试,并且特征不需要额外存储空间,但是区域采样步骤耗费的时间,相比检测算法也要高出一个数量级。文献[14]针对上述问题,提出一种全卷积区域采样算法 RPN(region proposal network)^[15]的 CNN 网络,能够快速产生高质量采样区域。

通过适当堆叠深度学习算法卷积模块可以提高网

络的学习能力,提升网络性能^[16],但在算法训练过程中将面临梯度弥散和梯度膨胀^[17]的问题,过多堆叠卷积模块反而造成准确率下降。为解决此类问题,提出了残差网络 ResNet^[18]和 Highway Networks^[19]等,采用岔路连接的策略解决了卷积模块堆叠过多而难以训练的问题。基于残差结构的目标识别算法模型 RFCN 仍然保持着 2017 年 PASCAL VOC 数据集中最高的目标识别准确率,但该算法每一层学习到的特征相互独立,没有共享特征信息。那么,有没有一种算法能突破这种单向连接的建模方式?在 ILSVRC2017 的比赛中,由 Facebook 人工智能团队提出的 DenseNet^[20]分类算法通过共享每一层特征图的方式,获得了 ILSVRC2017 最高的目标分类准确率。DenseNet 通过特征重用而不是拓展算法模型深度^[21]抑或是拓展算法宽度(如 GoogLeNet^[22])或变体的 ResNet^[23]来提高算法性能,DenseNet 在分类准确率和模型大小上存在很大的优势。

因此,本文拟通过卷积模块稠密连接的方式,各个卷积模块都和上层的所有卷积模块稠密连接,复用深度网络模型中每一层的特征,底层的卷积模块学习到了上面所有卷积模块的特征,必将提高深度网络模型的特征表达能力,不仅提高识别平均准确率和减小深度网络模型,而且有效解决梯度弥散、梯度膨胀问题。本文将此模型应用在军事目标识别上在准确率、实时性等方面都取得了比较好的识别效果。

2 面向军事目标识别应用的 DRFCN 深度网络模型设计

面向军事目标识别应用的 DRFCN 深度网络模型总体结构图如图 1 所示,主要包括:1) 基于稠密卷积

神经网络的区域采样算法(dense connected region proposal network, DRPN); 2) 基于稠密卷积神经网络的快速区域分类算法(dense connected fast classification network, DFCN); 3) DRFCN 算法模型结构; 4) RFCN 参数训练方法等。

2.1 基于稠密卷积神经网络的区域采样算法(DRPN)

基于稠密卷积神经网络的 DRPN 算法模型结构如图 2 所示,该模块主要完成采样区域的提取任务,通过将采样区域进行前景和背景类别的预测。该部分的类别预测输出仅为 0 或者 1,其中 1 表示前景(包含目标),0 表示背景(不包含目标)。

1) 采样区域获取:输入通道的特征图片经过稠密卷积网络层变换后,生成了多个 $W \times H$ 大小的特征图,特征图的每一个像素(神经元)具有很宽泛的感受野,对于一个 $W \times H$ 的特征图将产生 $W \times H \times k$ 个锚点框(k 为样本的数量),这些采样区域包含了大量的前景区域(包含目标)和背景区域(不包含目标),在 $W \times H \times k$ 个样本中利用交并比(intersection over union, IoU)指标对产生的采样区域进行筛选,使得锚点框正负样本比例数量维持在 3:1 左右。正样本:对于任意一个锚点框,它和前景区域有最大的重叠区域,或者它和前景区域有大于 70%的重叠区域;负样本:对于任意一个锚点框,它满足不是正样本的前提下,并且和所有前景区域有低于 30%的重叠区域;剩下的既不是正样本也不是负样本,不参与最终训练。

2) 联合代价函数:为了实现目标采样的提取任务,在此采用联合代价函数训练 DRPN。通过分别设定分类和区域采样算法的代价函数,利用边框回归与逻辑回归算法计算采样区域的类别损失 LOSS 和位置损失 LOSS,设计如式(1)所示的联合代价函数:

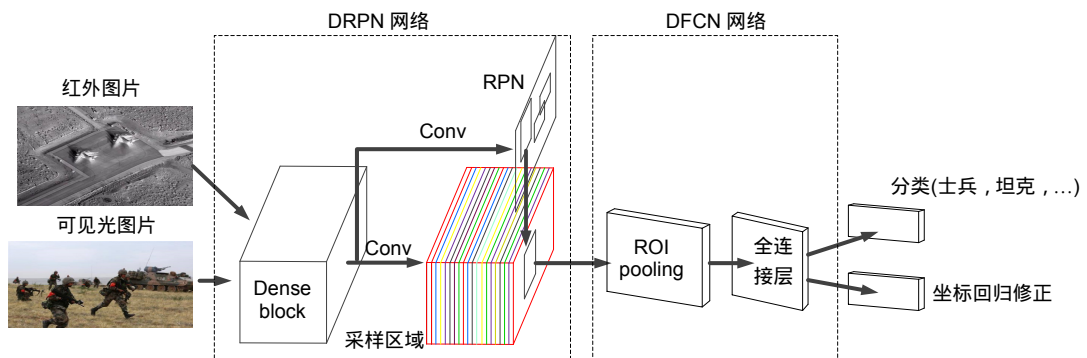


图 1 DRFCN 深度网络模型总体结构框图

Fig. 1 General structure diagram of DRFCN depth network model

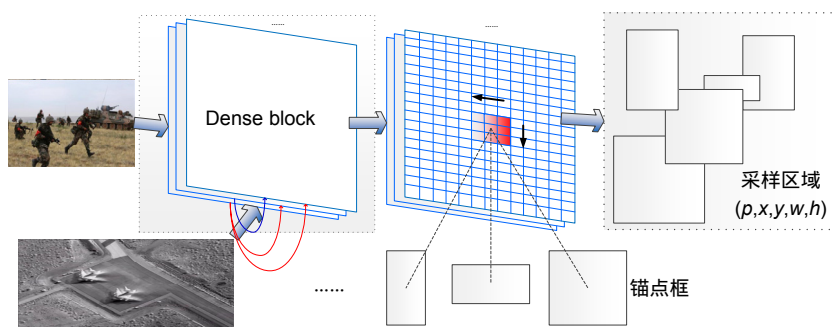


图 2 DRPN 的总体结构图

Fig. 2 The overall structure diagram of DRPN

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} L_{cls} \sum_i (p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (1)$$

其中： L_{cls} 和 L_{reg} 分别表示分类前景区域(包含目标)和背景区域(不包含目标)和锚点框回归的代价函数， $L_{cls}(p_i, p_i^*)$ 中 p_i 表示预测的锚点框置信得分，如果一个锚点框为负样本， p_i^* 为 0；如果一个锚点框为正样本， p_i^* 为 1。 N_{cls} 表示一次训练选取的锚点框数量， N_{reg} 表示选取锚点框的特征图大小。

2.2 基于稠密卷积神经网络的快速区域分类算法

基于稠密卷积神经网络的快速区域分类算法(DFCN)的总体结构图如图 3 所示。本文采取了区域池化层(region of interest pooling layer, ROI pooling)和全连接层组合的方式构建 DFCN 算法模型。利用 DRPN 快速、高效地获得高质量的采样区域后，需要对获得的采样区域进行细分类。在深度网络中，特征的质量直接决定了目标识别算法的最终准确率。DFCN 算法相较已有深度网络可以实现：1) 融合高低层次语义特征信息进行采样区域分类；2) 利用特征图信息同时输

出了军事目标的类别和位置信息。基于深度学习的 DFCN 分类算法主要有以下几个部分：

1) 特征转换：对于每一层特征图，算法模型需要对特征图做卷积、池化、归一化、线性或非线性转换等操作。

2) 采样区域池化层：将不同维度的特征矩阵归一化到相同维度的变换。

3) 联合代价函数：DFCN 有两个同级的输出：输出 $k+1$ 个目标的类别信息，其中 k 表示类别；输出每一类的坐标位置。训练 DFCN 的联合代价函数如下式：

$$L(p, u, t^u, v) = L_{cls}(p, u) + \lambda [u \geq 1] L_{loc}(t^u, v) \quad (2)$$

式中： $L_{cls}(p, u) = -\log(p_u)$ 是真实类别 u 概率分布的对数损失。边框回归的损失 L_{loc} 的定义基于两组参数：类别 u 输出的边框偏移量 $v = (v_x, v_y, v_w, v_h)$ ；类别 u 的预测边框偏移量 $t^u = (t_x^u, t_y^u, t_w^u, t_h^u)$ 。只有正样本才参与边框回归，通过 u 对 y 进行控制。具体计算式：

$$y = [u \geq 1] \begin{cases} 1 & u \geq 1 \\ 0 & u = 0 \end{cases} \quad (3)$$

λ 用来控制分类损失和边框回归损失的正则项值，平衡这两个损失 L_{loc} 和 L_{cls} 的比重。

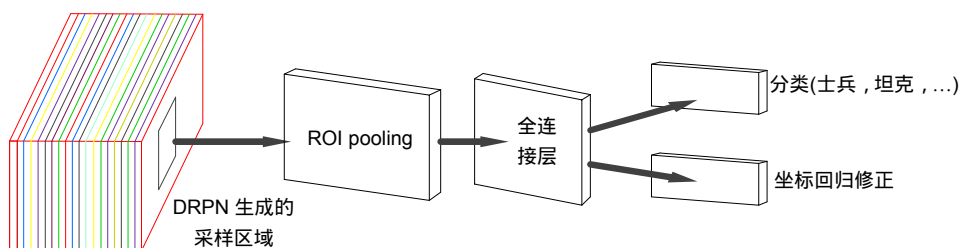


图 3 DFCN 的总体结构图

Fig. 3 General structure diagram of DFCN

2.3 DRFCN 算法模型结构

密集连通区域快速分类网络(Dense connected region fast classification network, DRFCN)算法主要有以下几个部分,具体模型结构的如表 1 所示。

表 1 DRFCN16 目标识别算法模型结构
Table 1 DRFCN16 object recognition algorithm model structure

DRFCN 层	输出尺寸(w×h)	DRFCN16
卷积	250×500	7×7 卷积, 步长 2, 填充 3
池化	126×251	3×3 池化, 步长 2, 填充 1
稠密连接卷积层	126×251	多层卷积
卷积	63×26	5×5 卷积, 步长 2, 填充 2
池化	32×64	3×3 池化, 步长 2, 填充 1
稠密连接卷积层	32×64	多层卷积
卷积	32×64	3×3 卷积, 步长 1, 填充 1
池化	32×64	3×3 池化, 步长 1, 填充 1
稠密连接卷积层	32×64	多层卷积
卷积	32×64	3×3 卷积, 步长 1, 填充 1
ROI	7×7	ROI 池化层
FC	21	连接分类器输出置信度

1) 特征提取基础网络部分:由 3 个稠密卷积块和多层卷积池化层组成,稠密卷积块由 6 层卷积层稠密连接而成;

2) 特征降维部分:在每个稠密卷积块之间分布着卷积池化层,每个稠密卷积块都是级联了上层所有特征图,每一层表现为综合了前面所有层的特征信息,为了避免算法底层权值维度过高带来的内存溢出问题,在每一个稠密卷积块之间添加了卷积池化层降低特征图矩阵的维度;

3) 分类器和位置坐标回归预测部分:在提取了目标特征信息后,采用 DRPN 独有的层获得采样区域,采用 ROI 归一化特征图大小和全连接层加分类器^[24]的方式获得目标的置信得分和最终的位置信息。

2.4 DRFCN 参数训练方法

DRPN 算法用于产生高质量的采样区域,DFCN 算法用于计算关于采样区域类别信息的离散概率分布和位置坐标。假如要使得 DRPN 和 DFCN 共享稠密卷积层,则 DRFCN 需要一种训练机制,使得 DRPN 和 DFCN 共享权重。本文实验采样区域生成网络和采样区域分类网络公用共享卷积层的模型训练方法,采用 DRPN 和 DFCN 联合分布参数训练,将 DRPN 和 DFCN 看做一个统一的整体,在每一次采用批量随机梯度下降算法(mini-batch stochastic gradient descent, MSGD)

的过程中,前向传递阶段用 DRPN 产生的采样区域直接训练 DFCN,反向传递阶段 DRPN 和 DFCN 梯度依次反传,多次迭代后直到算法收敛。大致训练步骤如下:

步骤 1:使用 MSGD 算法训练 DRPN,稠密连接的卷积模块用超前训练的权重初始化;

步骤 2:使用步骤 1 中 DRPN 产生的采样区域训练 DFCN,稠密连接的卷积模块用超前训练的权重初始化;

步骤 3:使用步骤 2 中 DFCN 稠密连接的卷积层初始化 DRPN,保持稠密连接部分权重不变,只微调 DRPN 独有的层,到此为止,DRPN 和 DFCN 已经共享了卷积层;

步骤 4:保持稠密连接的卷积层块权值不变,使用步骤 3 中 DRPN 生成的采样区域训练 DFCN,这一步只微调 DFCN 独有的层;

步骤 5:算法收敛,训练结束。

3 DRFCN 方法实验分析及讨论

本节中将对 DRFCN 深度网络开展实验验证和性能测评。首先,利用 PASCAL VOC 标准数据集,开展 DRFCN 方法和已有典型深度学习方法的性能对比实验和分析;其次,利用实际作战环境中采集得到的图像、视频资源,仿照 PASCAL VOC 数据集的格式,自建军事目标数据集,并利用该数据集对 DRFCN 算法开展实验验证及分析。具体分析如下:

3.1 实验环境及数据集

1) 实验软硬件环境:Caffe 框架;CPU: Xeon E5-2450@2.00 GHz;内存 16 G;GPU:TITAN X(12 G)。

2) 数据集:

PASCAL VOC^[25]数据集:VOC2007(9963 张已标注的图片)和 VOC2012(17125 张已标注的图片)。

自建军事目标数据集:包含陆、海、空三个军种的军事目标,具体分为 9 类:战斗机、坦克、直升飞机、军舰、枪、导弹、加农炮、潜艇、士兵。数据集中包含了各种分辨率的图片,图片中包含了各个角度的军事目标,不仅如此,数据集中还包含了大量近景、远景军事目标(10000 张已标注的图片)。

3.2 性能评价指标

1) 目标识别方法的准确性指标:

首先,给出目标识别中精确率 P ,召回率 R 的计算式分别为

$$P = \frac{TP}{TP + FP}, \quad (4)$$

$$R = \frac{TP}{TP + FN}, \quad (5)$$

式中： TP 表示在一次测试中正确预测的正样本数量， FP 表示在一次测试中被错分为正样本的数量， FN 表示所有的被错误预测为负样本的正样本数量。

本文采用目前基于深度学习的目标识别方法中广泛使用的平均准确率 mAP 作为识别准确性指标。利用 11 点轴法分别计算某一类别目标的平均准确率 AP ，如式(6)所示，再计算平均准确率 mAP ，如式(7)所示，式中： N 为目标识别框架的类别总数。

$$AP = \frac{\sum_{k=0,0.1,\dots,1} \max(P[R > k])}{11}, \quad (6)$$

$$mAP = \frac{1}{N} \sum_{i=0}^N AP_i. \quad (7)$$

2) 目标识别方法的实时性指标：

本文采用测试时间和网络模型大小两个指标来评价目标识别方法的实时性指标。具体如下：

测试时间 t_s ：是指算法从任意图像帧数据读入到检测结果输出所需的时间，通常用毫秒(ms)作为其度量单位。在给定的软硬件条件下，当 t_s 取值越小，说明目标识别方法的实时性能越好。

网络模型大小 N_p ：是指整个网络可训练参数

的大小，其主要包含卷积层参数的数量和全连接层数。其计算式：

$$N_p = \sum_{i=1}^m N_{c_i} + \sum_{j=1}^n N_{f_j} + N_a, \quad (8)$$

式中： m 为卷积层数量； n 为全连接层数； N_{c_i} 为第 i 层卷积层参数量； N_{f_j} 为第 j 层全连接层参数量； N_a 为其他参数。

3.3 实验结果及分析

1) DRFCN 方法的收敛性实验

本文首先在 100 万张的 ImageNet^[26]数据集上训练了针对 1000 个不同类别的分类模型；随后用预训练的分类模型初始化 DRFCN 基础网络部分和全连接部分，DRPN 独有层的权值采用均值为 0，方差为 0.01 的高斯分布随机初始化；最后对算法模型进行微调。根据稠密连接的卷积块数量差异，本文训练了两种极深的 DRFCN 模型，分别命名为 DRFCN5、DRFCN16。DRFCN16 算法收敛过程如图 4 所示，横坐标表示迭代次数，纵坐标表示对应的类别 LOSS，总体 LOSS(代价函数损失值)，DRPN 下采样区域分类 LOSS 和 DRPN 下采样算法坐标回归 LOSS，算法总共迭代了 15 万次，DRFCN 方法很快得到收敛。

为了进一步描绘 DRFCN 算法的收敛性，通过召回率和对应的精度之间的函数关系可以画出对应的

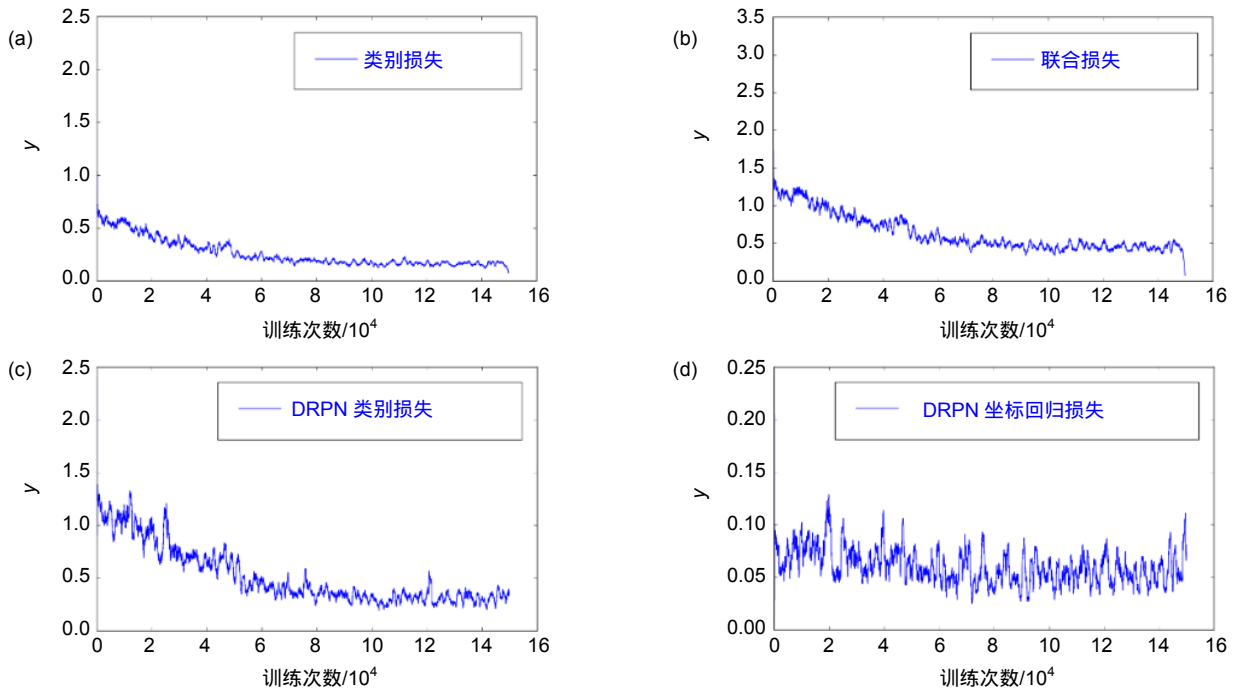


图 4 DRFC16 迭代收敛过程示意图

Fig. 4 Schematic diagram of the DRFC16 iterative convergence process

PR 曲线, PR 曲线很好地表现一个分类器的性能, 平均准确率 mAP 就是 PR 曲线下方的区域面积, 所以要求曲线包围的面积越大越好。DRFCN 算法的 P-R 曲线如图 5 所示, 从图中可以看出: 通过 150000 次迭代, 算法很快收敛且分类器的性能表现也越来越佳。

2) 预训练模型分类准确性对比实验

DRFCN 方法需要通过预训练模型对基础网络部分和全连接部分进行初始化。在 100 万张 ImageNet 数据集上针对 1000 个不同类别的分类模型, 分别使用 DRFCN 方法和已有典型深度学习模型方法开展预训练, 得到实验结果如表 2 所示。

网络模型大小分析:

从表 2 中可以看出, DRFCN16、VGG16、ResNet-101 的 Top-5 准确率差距小于 1%, 但 DRFCN16 网络模型大小分别是 VGG16 和 ResNet-101 的 20.5% 和 55.2%。表明当分类准确率相同, DRFCN 网络模型小于已有典型深度学习网络模型。即 DRFCN 方法可实现在大幅度压缩模型大小的情况下, 获得和前沿算法相近的准确率。

分类准确率分析:

从表 2 中可以看出, DRFCN5 和 ResNet-18 的网

络模型大小分别为 50.8 MB 和 44.6 MB ;但是 DRFCN5 的 Top-1 和 Top-5 分类准确率对比 ResNet-18, 分别提高了 4.5% 和 0.5%。表明: 当模型大小相同, DRFCN 可以取得更好的分类准确率。

3) DRFCN 和已有方法的对比实验

基于 VOC 数据集的训练时间、测试时间和识别准确性比较实验结果, 如表 3 所示。

从表 3 中可以看出, 在 VOC2007 数据集上, DRFCN16 平均准确率 75.3%、每幅图像测试时间为 0.12 s, VGG16 的平均准确率 69.9%、每幅图像测试时间 0.42 s, 相对 VGG16, DRFCN16 在平均准确率上提高了 5.4%, 在测试时间上减少了 30 ms。结果表明, 从识别结果的准确性和实时性分析, DRFCN 方法整体性能优于 VGG 方法。

DRFCN16 平均准确率 75.3%、每幅图像测试时间为 0.12 s, RFCN-101 的平均准确率 76.6%、每幅图像测试时间 0.17 s, 对于 DRFCN16 和 RFCN-101, 在平均准确率相当的情况下, DRFCN 的测试时间减少了 50 ms。结果表明, DRFCN16 与 RFCN-101 的识别结果的准确性相当, 但是 DRFCN 方法的实时性优于 RFCN-101。这里, 对 DRFCN 算法的实时性提高极具

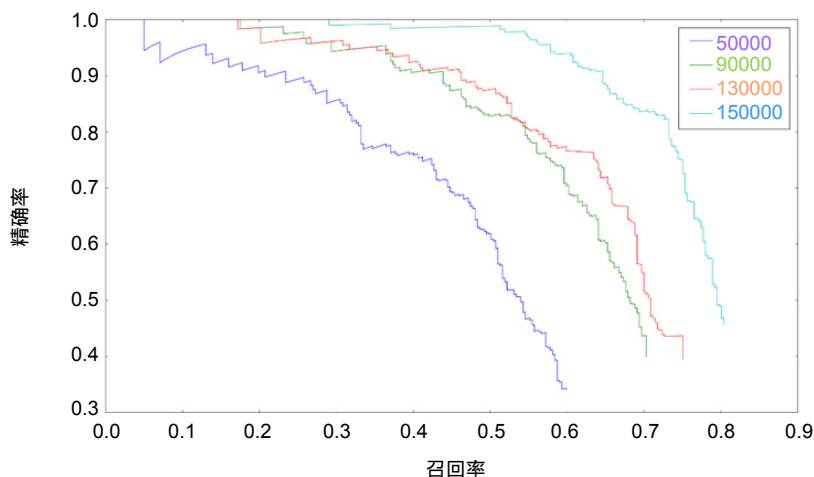


图 5 精确率-召回率曲线
Fig. 5 Precision-recall curve

表 2 DRFCN 在 ImageNet 数据集上预训练模型大小及分类准确率比较
Table 2 Comparison of pre-training model sizes for DRFCN on ImageNet datasets

算法模型	Top-1/%	Top-5/%	模型大小/MB
DRFCN5	74.9	90.1	50.8
DRFCN16	76.1	93.1	112.3
VGG16	76.0	93.2	548.3
ResNet-18	70.4	89.6	44.6
ResNet-101	80.1	94.0	203.5

体的分析：在目标识别任务中，通过堆叠卷积层可以得到表达能力更强的特征，然而深度学习算法模型会面临梯度弥散和梯度膨胀的问题，何凯明等人提出的残差算法模型^[17]解决了这一类问题。在 RFCN-101 残差网络中分别采用了如式(9)所示的前向传递表达式和如式(10)所示的反向传递表达式：

$$X_n = X_l + \sum_{i=1}^{n-1} F_i(X_i) \quad (9)$$

$$\frac{\delta J}{\delta X_l} = \frac{\delta J}{\delta X_n} \left(1 + \frac{\delta F(X_l)}{\delta X_l}\right) \quad (10)$$

式中： J 表示代价函数， X_n 表示算法模型的第 n 层特征图， F_i 表示第 i 层特征图的变换。而在本文提出的 DRFCN 算法中，前向传递过程和反向传递过程如式(11)、式(12)所示：

$$X_n = \sum_{i=0}^{n-1} X_i + F_{n-1}(X_{n-1}) \quad (11)$$

$$\frac{\delta J}{\delta X_l} = \frac{\delta J}{\delta X_n} \left(1 + \frac{\delta F(X_l)}{\delta X_l}\right) \quad (12)$$

通过对比残差网络可以发现，DFCN 算法和残差算法模型反向传递表达式相同，但是前向传递表达式

不同，利用稠密连接卷积层方式，梯度同样是通过和的方式在算法模型中反向传递。所以，DFCN 算法加强了深度学习算法模型的特征表达能力，从而在较少的卷积操作计算的条件下，实现了模型特征提取能力的增强，使得模型在保证准确性的条件下实时性得到进一步提升。

4) 基于自建军事目标数据集的识别实验

在自建的军事目标数据集上测试了 DRFCN 算法，得到实验结果如表 4 所示，部分识别结果如图 6 所示。

从表 4 中可以看出，在具体的场景下，DRFCN 算法在准确率上获得了 77.5%，每幅图像测试时间为 0.20 s，相比该算法在 VOC 数据集上，准确率提升了 2.2%，测试时间减少了 80 ms。该实验结果表明，DRFCN 算法在准确率和实时性上满足军事目标识别任务。

4 结 论

本文提出了一个面向军事目标识别应用的 DRFCN 深度网络模型算法，该算法通过卷积模块稠密连接的方式，复用深度网络模型中每一层的特征，可

表 3 DRFCN 算法和前沿目标识别算法模型在 VOC2007 数据集上的比较

Table 3 Comparison of DRFCN algorithm and frontier target recognition algorithm model on VOC2007 dataset

算法模型	每幅图像训练时间/s	每幅图像测试时间/s	mAP/%
DRFCN5	0.21	0.09	72.1
DRFCN16	0.28	0.12	75.3
VGG16	1.20	0.42	69.9
RFCN-101	0.45	0.17	76.6

表 4 DRFCN16 算法在自建的军事目标数据集上的平均准确率和测试时间

Table 4 The average accuracy and test time of DRFCN16 algorithm on the data set of the self-built military target

每幅图像测试时间/s	mAP/%	战斗机	坦克	直升机	军舰	枪	导弹	加农炮	潜艇	士兵
0.20	77.5	90.7	77.2	91.6	78.7	69.1	74.2	68.8	67.7	79.5



图 6 DRFCN 检测结果部分展示

Fig. 6 DRFCN test results display in part

以提高目标识别的平均准确率并且减小深度网络模型的大小。利用 PASCAL VOC 数据集,从目标识别准确率、实时性和网络模型大小等评价指标,对 DRFCN 算法和典型深度目标识别网络进行对比实验和分析,结果表明,DRFCN 算法综合性能指标更优,并可有效的解决梯度弥散、梯度膨胀问题。进一步,利用自建的军事目标数据集,结合军事目标识别应用,使用 DRFCN 算法进行学习训练、测试、评价,验证了基于 DRFCN 的军事目标识别方法可以基本满足军事目标识别需求。一方面,DRFCN 算法采取的稠密连接卷积模块的方式提高了目标识别准确率,缩减了模型大小,但是该算法还存在冗余,因此,如何在保持准确率的前提下,对 DRFCN 算法模型做进一步的压缩,是后续研究的一个方向;另一方面,如何将大数据处理技术和云计算技术与 DRFCN 算法相结合,提升军事目标识别的准确性,也是后续研究的重点。

参考文献

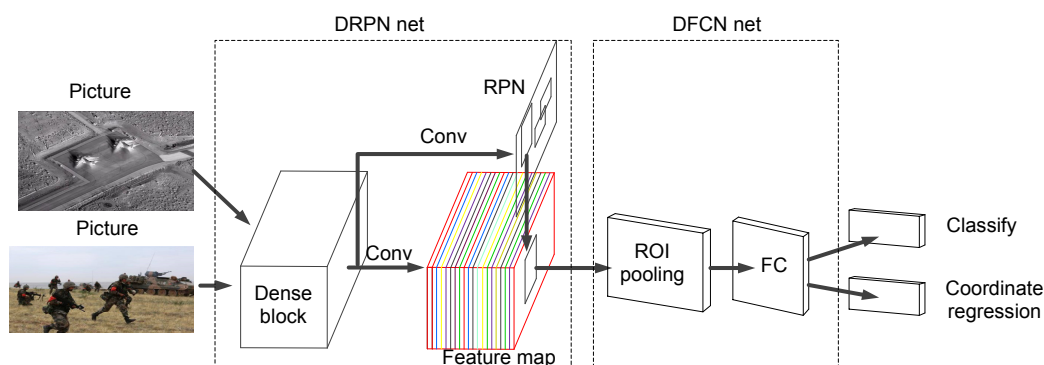
- [1] Nair D, Aggarwal J K. Bayesian recognition of targets by parts in second generation forward looking infrared images[J]. *Image and Vision Computing*, 2000, **18**(10): 849–864.
- [2] Crevier D, Lepage R. Knowledge-Based image understanding systems: A survey[J]. *Computer Vision and Image Understanding*, 1997, **67**(2): 161–185.
- [3] Bartlett T A. Simplified IR signature prediction for model-based ATR[J]. *Proceedings of SPIE*, 1993, **1957**: 111–121.
- [4] Watkins W R, CuQlock-Knopp V G, Jordan J B, et al. Sensor fusion: a preattentive vision approach[J]. *Proceedings of SPIE*, 2000, **4029**: 59–67.
- [5] Rogers S K, Colombi J M, Martin C E, et al. Neural networks for automatic target recognition[J]. *Neural Networks*, 1995, **8**(7–8): 1153–1184.
- [6] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//*Proceedings of the 25th International Conference on Neural Information Processing Systems*, 2012: 1097–1105.
- [7] Dai J F, Li Y, He K M, et al. R-FCN: object detection via region-based fully convolutional networks[C]//*30th Conference on Neural Information Processing Systems*, 2016: 379–387.
- [8] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//*2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779–788.
- [9] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[C]//*14th European Conference on Computer Vision*, 2016: 21–37.
- [10] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//*2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014: 580–587.
- [11] He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(9): 1904–1916.
- [12] Grauman K, Darrell T. The pyramid match kernel: discriminative classification with sets of image features[C]//*Tenth IEEE International Conference on Computer Vision*, 2005, **2**: 1458–1465.
- [13] Girshick R. Fast R-CNN[C]//*Proceedings of the 2015 IEEE International Conference on Computer Vision*, 2015: 1440–1448.
- [14] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(6): 1137–1149.
- [15] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//*IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 3431–3440.
- [16] Wang Z L, Huang M, Zhu Q B, et al. The optical flow detection method of moving target using deep convolution neural network[J]. *Opto-Electronic Engineering*, 2018, **45**(8): 180027.
王正来, 黄敏, 朱启兵, 等. 基于深度卷积神经网络的运动目标光流检测方法[J]. *光电工程*, 2018, **45**(8): 180027.
- [17] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult[J]. *IEEE Transactions on Neural Networks*, 1994, **5**(2): 157–166.
- [18] 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*, 2016: 770–778.
- [19] Srivastava R K, Greff K, Schmidhuber J. Training very deep networks[C]//*Neural Information Processing Systems*, 2015: 2377–2385.
- [20] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//*2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 2261–2269.
- [21] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C]//*2015 IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 1–9.
- [22] Zagoruyko S, Komodakis N. Wide residual networks[C]//*Proceedings of the British Machine Vision Conference*, 2016, **87**(7): 1–12.
- [23] Xie S N, Girshick R, Dollár P, et al. Aggregated residual transformations for deep neural networks[C]//*2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 5987–5995.
- [24] Gu Y, Xu Y. Fast SAR target recognition based on random convolution features and ensemble extreme learning machines[J]. *Opto-Electronic Engineering*, 2018, **45**(1): 170432.
谷雨, 徐英. 基于随机卷积特征和集成超限学习机的快速 SAR 目标识别[J]. *光电工程*, 2018, **45**(1): 170432.
- [25] Everingham M, Eslami S M A, Van Gool L, et al. The pascal visual object classes challenge: a retrospective[J]. *International Journal of Computer Vision*, 2015, **111**(1): 98–136.
- [26] Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database[C]//*2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009: 248–255.

Design and implementation of DRFCN in-depth network for military target identification

Liu Jun^{1*}, Meng Weixiu¹, Yu Jie², Li Yahui¹, Sun Qiao¹

¹Fundamental Science on Communication Information Transmission and Fusion Technology Laboratory, Hangzhou Dianzi University, Hangzhou, Zhejiang 310018, China;

²China Shipbuilding Industry Corporation 715 Research Institute, Hangzhou, Zhejiang 310023, China



General structure diagram of DRFCN depth network model

Overview: Automatic target recognition (ATR) technology has always been the key and difficult point in the military field. Photoelectric detection is one of the key detection methods in modern early warning and detection information network. In actual combat, massive images and video data of different types, timings and resolutions can be obtained by optoelectronic devices. For these massive infrared images or visible light images, this paper designs and implements a DRFCN in-depth network for military target identification applications. Firstly, the DRFCN algorithm inputs images and the part of DRPN is densely connected by the convolution module to reuse the features of each layer in the deep network model to extract the high quality goals of sampling region; Secondly, in the DFCN part, we fuse the information of the semantic features of the high and low level feature maps to realize the prediction of target area and location information in the sampling area; Finally, the deep network model structure and the parameter training method of DRFCN are given. In the experimental analysis and discussion part: 1) Through a large number of experiments, we draw various types of LOSS curves and P-R curves to prove the convergence of the DRFCN algorithm. 2) On the pre-training classification model based on the ImageNet dataset, the DRFCN algorithm achieved 93.1% Top-5 accuracy, 76.1% Top-1 accuracy and the model size was 112.3 MB. 3) Based on the PASCAL VOC dataset, the accuracy of DRFCN algorithm is 75.3%, which is 5.4% higher than that of VGG16 network. The test time of the DRFCN algorithm is 0.12 s. Compared to VGG16, the test time was reduced by 0.3 s. The DRFCN algorithm has advantages over the existing algorithm. Therefore, it is superior to the existing depth learning based target recognition algorithm. At the same time, it is verified that the DRFCN algorithm can effectively solve the vanishing gradient and exploding gradient. 4) Using the self-made military target data set for experiments, the DRFCN algorithm has an accuracy rate of 77.5% and a test time of 0.20 s. Compared to the PASCAL VOC2007 dataset algorithm, the accuracy is increased by 2.2%. The time is reduced by 80 milliseconds. The results show that the DRFCN algorithm achieves the military target recognition task in accuracy and real-time. In summary, compared with the existing deep learning network, the comprehensive performance of the DRFCN algorithm is better. The DRFCN algorithm improves the recognition average accuracy, reduces the depth network model and effectively solves the vanishing gradient and exploding gradient.

Citation: Liu J, Meng W X, Yu J, *et al.* Design and implementation of DRFCN in-depth network for military target identification[J]. *Opto-Electronic Engineering*, 2019, 46(4): 180307

Supported by Naval Equipment Pre-research Innovation Project and National Natural Science Foundation of China (61333009, 61427808)

* E-mail: gutlj@163.com