

DOI: 10.3969/j.issn.1672-2337.2018.02.009

基于分块 CNN 的多尺度 SAR 图像目标分类算法

曲长文¹, 刘晨¹, 周强², 李智¹, 李健伟¹

(1. 海军航空大学电子信息工程系, 山东烟台 264001;
2. 海军航空大学科研部, 山东烟台 264001)

摘要: 针对合成孔径雷达(Synthetic Aperture Radar, SAR)图像目标分辨率差异大, 多尺度 SAR 图像目标分类准确率不高的问题, 提出了一种基于迁移学习和分块卷积神经网络(Convolutional Neural Network, CNN)的 SAR 图像目标分类算法。首先通过大量与目标域相近的源域数据对分块 CNN 的参数进行训练, 得到不同尺度下的 CNN 特征提取网络; 其次将 CNN 的卷积和池化层迁移到新的网络结构中, 实现目标特征的提取; 最后用超限学习机(Extreme Learning Machine, ELM)网络对提取的特征进行分类。实验数据采用美国 MSTAR 数据库以及多尺度 SAR 图像舰船目标数据集, 实验结果表明, 该方法对多尺度 SAR 图像的分类效果优于传统 CNN。

关键词: 卷积神经网络; 迁移学习; 多尺度 SAR 图像; 目标分类

中图分类号: TN957.5 文献标志码: A 文章编号: 1672-2337(2018)02-0169-05

Multi-Scale SAR Images Target Classification Algorithm Based on Block Convolutional Neural Network

QU Changwen¹, LIU Chen¹, ZHOU Qiang², LI Zhi¹, LI Jianwei¹

(1. Department of Electronic and Information Engineering, Naval Aeronautical University, Yantai 264001, China;
2. Department of Scientific Research, Naval Aeronautical University, Yantai 264001, China)

Abstract: Because of big difference of SAR image resolutions and low accuracy rate of multi-scale SAR image classification, an algorithm based on transfer learning and block convolutional neural network (CNN) is proposed. Firstly, the parameters of block CNN are trained by large number of source domain data which is similar to the target domain, and the CNN feature extraction networks under different scales are obtained. Secondly, the convolution and the pool layer are transferred to the new network structure to achieve the feature extraction. Finally, the extracted features are classified by extreme learning machine(ELM). The MSTAR database and multi-scale SAR ship image data sets are used as experimental data. The experimental results show the proposed method is superior to traditional CNN in the classification of multi-scale SAR images.

Key words: convolutional neural network (CNN); transfer learning; multi-scale SAR image; target classification

0 引言

合成孔径雷达(SAR)最早由美国科学家 Carl Wiley 于 1951 年提出, 相比于传统光学、高光谱和红外传感器, 具有不受天气、光照等外界条件的限制, 可以大范围、全天候、实时对目标进行检测, 并且有一定的穿透能力^[1]。随着 SAR 技术的日益成熟, 在民用和军事领域都得了较为广泛的应用,

具有较高的研究价值。但是由于合成孔径雷达类型众多, 图像分辨率差异较大, 同一区域在不同雷达中的成像效果不同, 对 SAR 图像目标分类的准确率产生较大的影响。

卷积神经网络(CNN)可以有效地从大量样本中学习到相应的特征, 提取出优于人工设计的特征, 并且样本数量越大, 提取的特征越有利于分类识别^[2]。CNN 虽然具有空间位置不变性, 对于图像中任意位置的目标都可以提取出相应的特征,

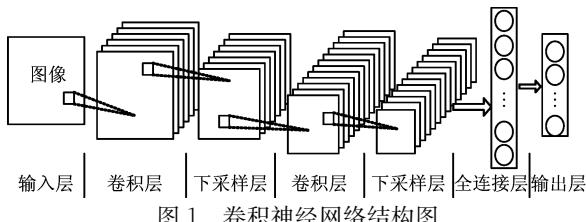
但是对于输入图像的空间变化适应性较差,要求训练图像与测试图像在空间尺度上一致,造成网络对与训练集尺度相似的测试数据有较高的激活值,进而导致网络对多尺度输入图像的分类准确率下降^[3]。

本文主要研究了分块 CNN 在多尺度 SAR 图像上的应用,改善了传统 CNN 模型的尺度不变性,提升了多尺度 SAR 图像目标分类识别的准确率。同时,为降低网络训练过程中对训练样本数量的需求,通过大量与目标域相似的源域图像对分块 CNN 进行训练,得到较优的网络参数;然后将网络参数迁移到新的分类网络结构中,用于特征的提取;最后将提取的特征用超限学习机(ELM)网络结构进行分类,ELM 的结构参数通过少量带标签的目标域数据训练得到。实验结果表明,本文方法可以有效提升网络对多尺度 SAR 图像目标的分类识别效果。

1 分块 CNN 的基本原理

1.1 卷积神经网络基本结构

典型的卷积神经网络由输入层、卷积层、下采样层(池化层)、全连接层和输出层五部分组成,如图 1 所示。



卷积层主要用于提取特征和降低噪声。假设第 l 层为卷积层,则第 l 层第 j 个特征图的计算公式如下所示:

$$\mathbf{x}_j^l = f \left(\sum_{i \in M_j} \mathbf{x}_i^{l-1} * \mathbf{w}_{ij}^l + \mathbf{b}_j^l \right) \quad (1)$$

式中, \mathbf{x}_i^{l-1} 表示第 $l-1$ 层的第 i 个特征图, \mathbf{w}_{ij}^l 表示第 $l-1$ 层的第 i 个特征图与第 l 层的第 j 个特征图之间的卷积核, $*$ 表示卷积操作, \mathbf{b}_j^l 表示第 $l-1$ 层的偏置, $f(\cdot)$ 表示非线性的激活函数。

下采样层主要利用图像局部之间相关性的原理,降低需要处理的数据量。将邻域 4 个像素变为

一个像素,常用的方法有均匀下采样、最大下采样和概率最大下采样,最终产生一个大概缩小 4 倍的特征映射图。

1.2 分块卷积神经网络结构

从 CNN 的结构可以看出,整个网络的结构和参数只能单一设定,对于不同尺度的输入图像,不能有效地提取出全局和局部特征,进而限制了网络参数对不同尺度输入图像特征提取的灵活性。本文对同一幅图像构造 3 个不同尺寸的 CNN,不同网络采用不同的卷积核与池化大小,最终输出维数相近的特征图。分块 CNN 的网络结构如图 2 所示。

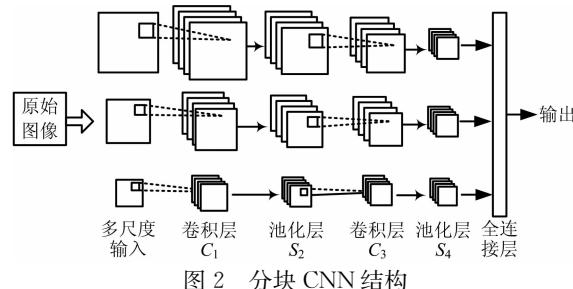


图 2 分块 CNN 结构

图 2 中,对每幅输入图像进行尺度变换,得到 3 种不同尺度的输入图像,然后分别对 3 种 CNN 进行训练,使最终的网络输出维数相近。网络中不同 CNN 的设计如图 3 所示。

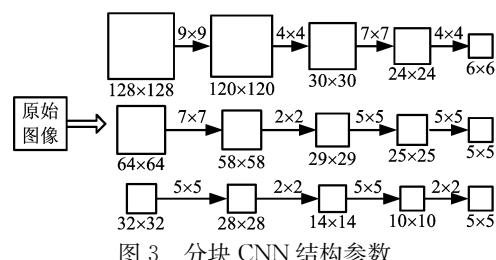


图 3 分块 CNN 结构参数

图 3 中,方框下面的数字表示该层图像的大小,箭头上的数字分别表示卷积核和池化大小,最终将 3 个网络提取出的特征图排列成一个一维向量,用 ELM 网络完成分类识别。

2 多尺度目标分类网络

2.1 迁移学习的基本理论

近年来,迁移学习(Transfer Learning)成为了卷积神经网络的发展趋势之一,它降低了 CNN

学习过程中对样本数量的要求,解决了训练样本和测试样本必须独立同分布,以及训练样本数目需求过大的问题。迁移学习可以通过已有的知识来解决小样本数据的学习问题,从而提升了卷积神经网络在小样本数据集上分类识别的准确率^[4-5]。

根据文献[6]对域和任务进行定义,域 D 由特征空间 \mathbf{X} 及其边缘概率分布 $P(\mathbf{X})$ 构成,其中 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \in \mathcal{X}$ 。任务 T 由标签空间 \mathbf{y} 与目标预测函数 $f(\cdot)$ 构成。

对于给定的源域 D_s 及其对应的源任务 T_s 、目标域 D_t 及其对应的目标任务 T_t ,迁移学习即是通过从给定的 D_s, T_s 中学到相应的知识,从而得到 D_t 的目标预测函数 $f(\cdot)$ 。迁移学习与传统方法的对比如图 4 所示。

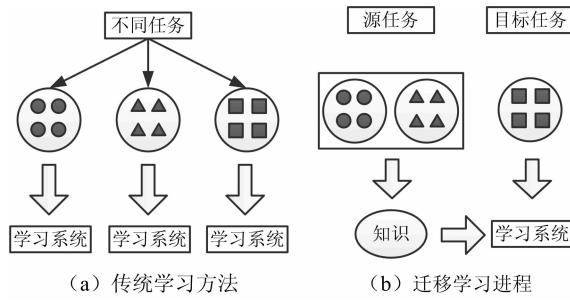


图 4 迁移学习与传统方法对比

从图 4 中可以看出,传统学习方法对于每个不同的训练任务需要训练出不同的学习系统,而通过迁移学习,可以从不同的任务中学习到相应的知识,进而用于新目标任务的预测。但是,并不是所有的知识都可以用来进行迁移学习,迁移学习过程需要源和目标具有相关性。

2.2 超限学习机模型

超限学习机因其训练参数少、训练速度快得到了广泛关注。相比于传统人工神经网络的复杂调参过程,ELM 在网络训练过程中,隐藏层节点参数随机确定,无需调节,仅需求解隐藏层与输出层之间权值的最小二乘范数解。因为网络参数的训练过程中不需要进行迭代过程,从而大大降低了训练过程中对样本数量的需求并提升了网络的训练速度。目前超限学习机在特征学习、聚类、回归和分类等方面得到了广泛的应用^[7-8]。ELM 的结构如图 5 所示。

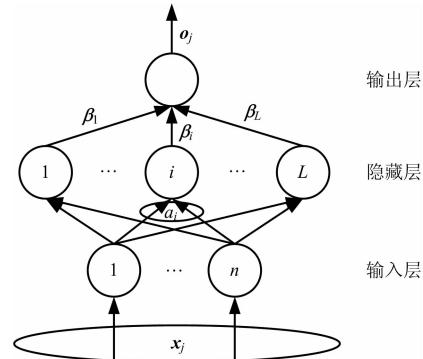


图 5 ELM 结构图

x_j 表示第 j 个输入样本, n 个节点表示输入的样本维度, 隐含层的节点数为 L , L 越大, 表达能力越强, a_i 表示输入样本与节点 i 的连接权值, β_i 表示隐藏层与输出之间的权值连接, o_j 表示第 j 个输入样本的输出。

输出权值的最小二乘范数解 β 的求解过程如下:

假设样本集的输入为 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, a_i 表示输入与隐藏层连接的权值, 则网络隐藏层的输出公式为

$$z^i = f(a_i x_j + b_i) \quad (2)$$

式中, z^i 表示第 i 个隐藏节点的输出, x_j 表示第 j 个输入样本, b_i 表示第 i 个隐藏节点的偏置, $f(\cdot)$ 表示激活函数。

网络的最终输出结果为

$$o_j = \sum_i^L \beta_i f(a_i x_j + b_i) \quad (3)$$

式中, o_j 为第 j 个输入样本的输出结果, β_i 为第 i 个隐藏节点与输出之间的权重, $f(a_i x_j + b_i)$ 为第 i 个隐藏节点的输出。

假设输入的样本为 (x_j, t_j) , 其中 x_j 表示 n 维向量, t_j 表示样本所对应的标签。令前馈网络输出与标签相等, 即 $o_j = t_j$, 得到

$$\left. \begin{array}{l} \beta_1 f(a_1 x_1 + b_1) + \dots + \beta_L f(a_L x_1 + b_L) = t_1 \\ \beta_1 f(a_1 x_2 + b_1) + \dots + \beta_L f(a_L x_2 + b_L) = t_2 \\ \vdots \\ \beta_1 f(a_1 x_N + b_1) + \dots + \beta_L f(a_L x_N + b_L) = t_N \end{array} \right\} \quad (4)$$

令 $H = \begin{bmatrix} f(a_1 x_1 + b_1) & \dots & f(a_L x_1 + b_L) \\ \dots & \dots & \dots \\ f(a_1 x_N + b_1) & \dots & f(a_L x_N + b_L) \end{bmatrix}_{N \times L}$, $\beta = [\beta_1, \dots, \beta_L]^T$, $T = [t_1, \dots, t_N]^T$, 则式(4)可以改写为

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T} \quad (5)$$

通过式(5)可求得 $\boldsymbol{\beta} = \mathbf{H}^{-1}\mathbf{T}^{[9]}$ 。

2.3 混合模型分类网络

本文通过迁移学习的方法降低网络结构对带标签样本数量的需求,并且与 ELM 结合形成最终的混合分类网络,网络结构如图 6 所示。

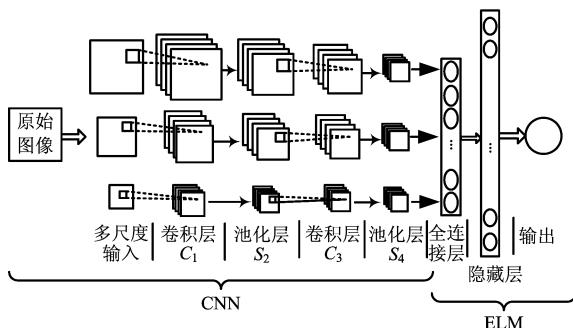


图 6 混合模型分类网络

算法步骤如下:

- 1) 将输入图像通过尺度变换为 $128 \times 128, 64 \times 64, 32 \times 32$ 三种不同尺度的图像,并用变换后的图像分别对 3 种 CNN 进行训练,得到相应的网络参数;
- 2) 通过迁移学习将 CNN 与 ELM 结构结合,进而得到适用于带标签样本数量不足数据的分类识别,并用带标签训练集样本对 ELM 参数进行训练,得到最终的混合模型分类网络;
- 3) 输入测试集数据,用最终训练好的混合分类模型网络对测试数据进行分类识别,得出相应分类结果。

3 实验仿真

3.1 实验数据集

为对本文提出的方法进行验证,数据集采用美国 MSTAR 数据库与 SAR 舰船目标数据集进行实验。MSTAR 数据库中图像的分辨率为 $0.3 \text{ m} \times 0.3 \text{ m}$,大小为 128 像素 \times 128 像素,训练样本为 17° 俯视角的 SAR 图像目标,共 2 320 幅,测试样本为 15° 俯视角的 SAR 图像目标,共 2 100 幅。为加快网络的训练速度,截取图像中心包含目标的 42×42 大小的图像块进行实验,如图 7 所示。

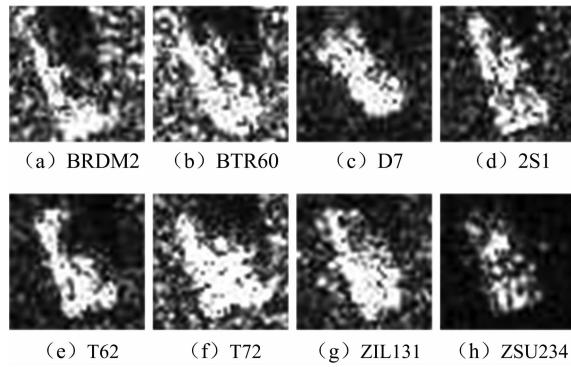


图 7 MSTAR 数据实验目标

SAR 舰船目标数据集对舰船和杂波进行了分类,一共有 999 幅图片,图片来自多幅不同分辨率的 SAR 图像,经过目标检测后获得船只的图像切片,通过专家解译对所有船只进行人工标注,数据集分布如表 1 所示。

表 1 SAR 舰船数据分布

类别	训练	测试	总计
船只	336	200	536
非船	333	130	463
总计	669	330	999

3.2 MSTAR 数据分类结果与分析

首先用 T72, ZIL131, ZSU234 三类目标数据集(训练样本 870 个,测试样本 810 个)对分块 CNN 的网络参数进行训练,然后通过迁移学习得到最终分类识别网络。最后对 BRDM2, BTR60, D7, 2S1 和 T62 五类目标(训练样本 1 450 个,测试样本 1 290 个)对混合分类网络进行训练,并对测试样本进行分类识别,结果如表 2 所示。

表 2 五类目标的 SAR 图像分类结果

测试目标	分类结果					正确分类 准确率/%
	BRDM2	BTR60	D7	2S1	T62	
BRDM2	272	0	1	0	1	99.27
BTR60	1	194	0	0	0	99.49
D7	0	0	273	0	1	99.64
S1	0	7	0	267	0	97.45
T62	0	0	0	0	273	100.00
整体分类 准确率/%						99.15

将本文方法与其他方法进行对比,结果如表 3 所示。

表 3 与其他方法效果对比

方法	识别率/%
CNN 方法	93.86
改进 CNN 方法	94.44
本文方法	99.15

从上述结果可以看出,本文方法相比于其他方法有效地提升了目标分类识别的准确率。为验证本方法对小样本数据集的有效性,将 MSTAR 目标训练集每次以递增 100 个样本的方式进行测试实验,测试样本数量保持不变,实验结果如图 8 所示。

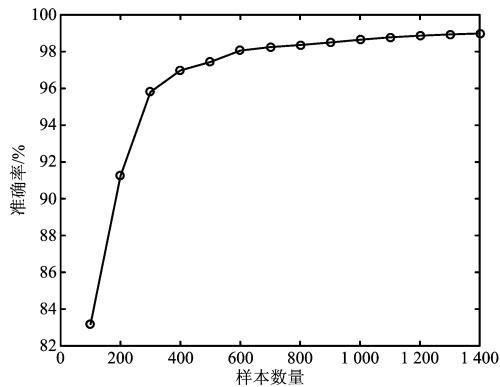


图 8 不同训练样本数实验结果

从实验结果可以看出,当训练样本数量达到 300 时,测试结果就已经超过了改进卷积神经网络,实验证明该方法有助于提升小样本数据的分类识别准确率。

3.3 SAR 舰船数据分类结果与分析

实验仍保持分块 CNN 的网络参数不变,用 SAR 舰船目标数据集对混合网络进行训练,并对测试集进行分类识别,得到每幅图像是舰船的可能性,对大于等于 0.5 的判定为舰船,否则为非舰船目标,部分识别效果如图 9 所示,最终识别结果如表 4 所示。

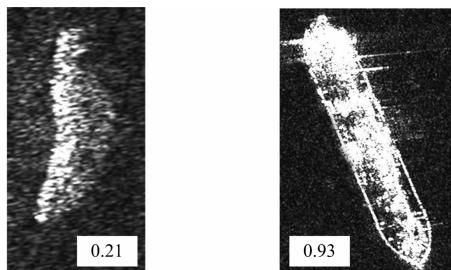


图 9 部分识别效果图

表 4 SAR 舰船目标分类识别结果

测试目标	识别结果		正确识别率/%
	舰船	杂波	
舰船	190	10	96.00
杂波	3	127	97.69
整体识别率/%			96.06

将本文方法与其他方法进行对比,在训练过程中,训练集每次递增 100 个样本的方式进行实验,测试样本数量保持不变,实验结果如图 10 所示,最终对比结果如表 5 所示。

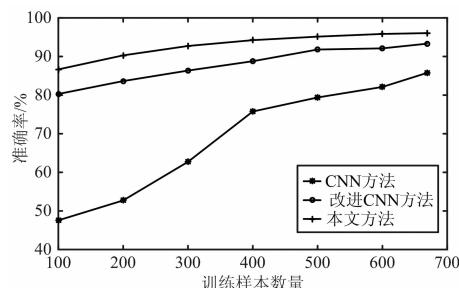


图 10 不同方法的实验结果

表 5 与其他方法对比结果

方法	识别率/%
CNN 方法	85.76
改进 CNN 方法	93.33
本文方法	96.06

从上述结果可以看出,本文方法提升了多尺度目标的分类识别准确率,为之后多尺度 SAR 图像进行目标检测识别提供了可靠保障。

4 结束语

针对 SAR 图像分辨率差异大,多尺度 SAR 图像目标的分类效果不佳的问题,提出了一种迁移学习与分块 CNN 相结合的混合分类网络,通过分块 CNN 有利于提升多尺度 SAR 图像的分类准确率;同时由于分块 CNN 网络参数多,对训练样本需求量大,采用迁移学习方法降低网络对带标签训练样本数量的需求。实验结果表明,两种方法结合后对带标签样本数量不足的多尺度 SAR 图像的分类效果得到了很大的提升。

参考文献:

- [1] CURLANDER J C, McDONOUGH R N. Synthetic Aperture Radar: Systems and Signal Processing[J]. New York: Wiley, 1991.

(下转第 180 页)