

基于注意力机制的偏振成像目标分类方法

孙瑞^{1,2,3}, 孙晓兵^{1,3*}, 刘晓^{1,3}, 宋强^{1,2,3}¹中国科学院合肥物质科学研究院安徽光学精密机械研究所, 安徽 合肥 230031;²中国科学技术大学, 安徽 合肥 230026;³中国科学院通用光学定标与表征技术重点实验室, 安徽 合肥 230031

摘要 基于注意力机制的神经网络可重点提取样本中关键区域的特征信息, 将此特点应用于偏振图像目标分类任务中有助于充分获取不同偏振图像之间的相互关系, 并从关键区域中提取出较多特征信息。针对杂乱自然背景中目标识别难度大的问题, 提出一种基于注意力机制的偏振图像目标分类方法。首先, 将注意力机制和卷积神经网络相结合, 构建出适用于数量有限样本的偏振特征提取模型; 然后, 选择合适的偏振图像作为模型输入并对其进行训练, 使得注意力模块中易于分类的通道域特征信息和空间域特征信息具有较大的权重, 以实现分类准确率的进一步提高。实验结果表明, 所提方法在不同自然背景中可有效提升目标分类的准确率, 在自建偏振目标数据库中准确率达到 95% 以上。相较于传统的深度学习分类方法, 所提方法的准确率得到明显的提高, 因此更适用于杂乱背景中目标的分类。

关键词 成像系统; 偏振成像; 卷积神经网络; 注意力机制; 目标分类

中图分类号 TP751

文献标志码 A

doi: 10.3788/AOS202141.1611004

Polarimetric Imaging Target Classification Method Based on Attention Mechanism

Sun Rui^{1,2,3}, Sun Xiaobing^{1,3*}, Liu Xiao^{1,3}, Song Qiang^{1,2,3}¹Anhui Institute of Optics and Fine Mechanics, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei, Anhui 230031, China;²University of Science and Technology of China, Hefei, Anhui 230026, China;³Key Laboratory of Optical Calibration and Characterization, Chinese Academy of Sciences, Hefei, Anhui 230031, China

Abstract The neural network based on attention mechanism can focus on extracting the feature information from the key areas. The application of this characteristic in the polarimetric imaging target classification can help us to obtain the relationships among different polarimetric images and to extract more feature information from critical areas. To solve the difficulty of target recognition in cluttered natural backgrounds, this paper presents a polarimetric imaging target classification method based on an attention mechanism. Firstly, the attention mechanism and the convolutional neural network are combined to construct a polarimetric feature extraction model suitable for limited samples. Then, proper polarimetric images are selected as the input model for training so that the attention module can give more weights to the channel domain and spatial domain feature information that is easily classified to obtain higher classification accuracy. The experimental results show that the classification accuracy of the proposed method can be further improved in different natural backgrounds and reach more than 95% in the self-built polarimetric target database, which is obviously improved to compare with that of the traditional

收稿日期: 2021-01-19; 修回日期: 2021-03-01; 录用日期: 2021-03-18

基金项目: 国家重点研发计划(2016YEF0201400)、中国科学院合肥物质科学研究院重点项目(Y73H9P1801)、卫星应用共性关键技术项目(30-Y20A010-9007-17/18)

通信作者: *xbsun@aiofm.ac.cn

deep learning classification method. Thus, our method is more suitable for target classification in cluttered backgrounds.

Key words imaging systems; polarimetric imaging; convolutional neural network; attention mechanism; target classification

OCIS codes 110.5405; 100.4996

1 引言

光学偏振成像检测技术在过去的 20 年里一直保持着快速发展,该技术不仅可以获取目标光的强度信息,同时还可以获取目标多个维度的偏振信息。在复杂的自然环境中,光学偏振成像系统获得的偏振图像含有丰富的细节信息,经特征提取后其目标分类的准确率得到提升。偏振图像目标检测研究始于 2000 年,Chenault 等^[1]实验验证了偏振图片中目标与背景之间具有较高的对比度。Sadjadi 等^[2]将偏振成像应用于目标检测领域,利用目标的偏振特性降低了目标检测的误警率。在后续的研究中,研究人员通过建立偏振双向反射模型来分析目标表面的偏振信息分布规律^[3-6]。在偏振图像算法上,张晶晶等^[7-8]提出了将偏振图像和普通图像融合的算法,该算法在保留图像频谱信息的同时,可以提取出目标的偏振信息。2015 年,研究人员提出了适用于偏振图像的特征提取算子,该算子通过图像二值化来筛选目标像素,有效提升了分类的效果^[9]。

随着深度学习技术的快速发展和电子芯片算力的大幅提升,基于卷积神经网络(CNN)的目标分类算法在性能上取得了巨大的突破,并成为当前计算机视觉领域目标分类任务的主流解决方案。研究人员提出了大量的经典神经网络模型(如 AlexNet^[10]、VGG^[11]、Resnet^[12]等),这些模型在 ImageNet 大规模视觉识别挑战赛中不断刷新任务记录,人类的识别精度得到了突破。近几年,研究人员在偏振图像中采用 CNN 算法来实现车辆检测,相较于传统图像,取得了更高的识别准确率^[13-14]。然而,神经网络技术在光学偏振成像检测领域的应用还处于初级阶段。

注意力机制是目前深度学习的研究热点之一,神经网络在提取完整图像特征信息的过程中,通过模拟人类大脑对环境的感知机理,对输入数据中的关键信息赋予较大的权重,并减小输入数据中无关信息的权重。2014 年,Mnih 等^[15]将注意力机制应用于计算机视觉领域中,采用循环神经网络与注意力机制相结合的方法实现图像分类。之后,研究人员利用残差注意力网络来完成图像分类任务,通过

堆叠注意力模块构建了一个端到端的训练网络^[16]。2020 年,关世豪等^[17]和赵斌等^[18]分别将注意力机制应用于红外图像和高光谱数据中,崔涓涓等^[19]和鞠默然等^[20]分别将注意力机制应用于无人机目标跟踪与多尺度目标检测等领域,取得了不错的效果。

在光学偏振相机拍摄的图像中,目标表面的偏振信息受到目标材质、观测角度、环境变量等多个因素的影响,目标表面和背景的偏振信息的辐射强度会随着这些因素的改变而发生较大幅度的变化。不同偏振图像采用相同的特征提取权重会导致优质通道的偏振信息无法被充分挖掘。此外,CNN 中含有较多的下采样运算,这些运算会进一步压缩提取的目标特征信息,因此偏振成像目标分类的难点在于如何在复杂多变的背景中高效提取有限目标区域的特征信息。因此,将注意力机制与 CNN 相结合是提升偏振成像目标识别效果的有效解决方案之一。为了实现自然背景下偏振成像人造目标的分类,本文利用深度学习技术提出了一种基于注意力机制的偏振成像目标分类方法。该方法利用不同偏振图像的通道信息和空间信息,设计出适用于数量有限偏振样本的神经网络,该网络聚焦于关键且易于分类的特征信息,提升了目标分类的准确率。

2 光学偏振图像特性

光的偏振状态可通过斯托克斯矢量来描述。斯托克斯矢量由四个参数组成,分别为 S_0 、 S_1 、 S_2 和 S_3 ,这四个参数的理论定义为

$$\begin{cases} S_0 = I_{0^\circ} + I_{90^\circ} \\ S_1 = I_{0^\circ} - I_{90^\circ} \\ S_2 = I_{45^\circ} - I_{-45^\circ} \\ S_3 = I_{lh} - I_{rh} \end{cases}, \quad (1)$$

式中: I_α ($\alpha = 0^\circ, 45^\circ, 90^\circ, 135^\circ$)表示不同偏振方向 α 所对应的偏振图; I_{lh} 和 I_{rh} 分别表示左旋偏振图和右旋偏振图; S_0 表示普通强度图像,常用于传统的目标分类任务中; S_1 和 S_2 表示偏振方向相互正交的两个线偏振图; S_3 表示圆偏振图,由于人造目标的圆偏振特性并不明显,其可忽略^[21]。因此,目标分类算法的输入数据由 S_0 、 S_1 和 S_2 组成。

假设入射光的斯托克斯矢量为 \mathbf{S}^i ,照射至物体表

面后反射光的斯托克斯矢量为 \mathbf{S} 。入射光与目标相互作用后,光束偏振状态的改变可采用穆勒矩阵 \mathbf{M} 来表示。入射光、反射光和穆勒矩阵之间的关系为

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \end{bmatrix} = \begin{bmatrix} M_{00} & M_{01} & M_{02} \\ M_{10} & M_{11} & M_{12} \\ M_{20} & M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} S_0^i \\ S_1^i \\ S_2^i \end{bmatrix}. \quad (2)$$

由(1)式和(2)式可知,目标表面的偏振信息受到多个环境参数的影响,不同的偏振图像存在较大差异,且自身的偏振信息也会因环境不同而发生明显的变化。

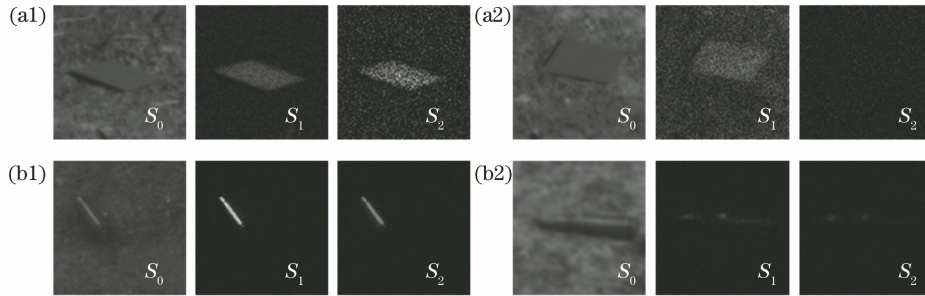


图 1 偏振成像目标特性。(a1)(a2)不同铁片的 S_0 、 S_1 和 S_2 样本;(b1)(b2)不同子弹的 S_0 、 S_1 和 S_2 样本
Fig. 1 Characteristics of polarimetric imaging targets. (a1)(a2) S_0 , S_1 and S_2 samples of different iron flakes;
(b1)(b2) S_0 , S_1 and S_2 samples of different bullets

3 基于注意力机制的偏振成像目标分类方法

3.1 网络中的基本模块

3.1.1 卷积块注意力模块

卷积块注意力模块(CBAM)是一种可无缝集成到神经网络中的轻量级注意力模块^[22],能够在CNN中完成端到端的训练且不会占用过多的计算资源。其核心思想就是依次采用通道注意力模块和空间注意力模块计算出输入特征图的注意力权重,再将其与原特征图相乘以实现特征权重的自适应调整,CBAM数据处理流程图如图2(a)所示。

在CBAM中,数据的处理主要分为两个独立的模块,分别为通道注意力模块和空间注意力模块。通道注意力模块的输入是一个尺寸为 $H \times W \times C$ 的特征图 \mathbf{F}_1 ,依次通过全局平均池化(AvgPool, f_{Avg})和全局最大池化(MaxPool, f_{Max})操作实现维度压缩,得到两个尺寸为 $1 \times 1 \times C$ 的特征向量,即 $\mathbf{F}_{\text{Avg}}^C$ 和 $\mathbf{F}_{\text{Max}}^C$ 。池化输出的结果先经过卷积运算实现数据的降维,输出尺寸为 $1 \times 1 \times (C/r)$,其中 r 为压缩比,激活函数为 Relu。再利用卷积运算实现数据的增维,输出尺寸为 $1 \times 1 \times C$,此过程在多层感知机(MLP, f_{MLP})中进行。最后,将得到的两个特征

图1展示了距离镜头4~6m的人造目标。在相同的拍摄环境下,有的图像中目标与背景之间区别明显,有的图像中目标与背景相近且难以区分。因此,增加易于分类图像的权重可以提升目标识别的效果。除此之外,不同人造目标占据图像的像素比例也不同,传统的神经网络是通过读取图像中的每个像素来提取特征信息的。在目标尺寸较小的情况下,背景占据的像素比例较大,提取的特征信息中含有大量无用的背景特征,这会影响偏振成像目标分类的准确率。

值相加并利用 Sigmoid 激活函数,得到通道域的权重系数,将权重系数与输入的特征图相乘就可得到新的特征图。通道注意力模块数据处理流程图如图2(b)所示,通道注意力的计算公式^[22]为

$$\begin{aligned} \mathbf{M}_C(\mathbf{F}_1) = & \sigma\{f_{\text{MLP}}[f_{\text{Avg}}(\mathbf{F}_1)] + f_{\text{MLP}}[f_{\text{Max}}(\mathbf{F}_1)]\} = \\ & \sigma\{\mathbf{W}_1[\mathbf{W}_0(\mathbf{F}_{\text{Avg}}^C)] + \mathbf{W}_1[\mathbf{W}_0(\mathbf{F}_{\text{Max}}^C)]\}, \\ & \mathbf{W}_0 \in \mathbb{R}^{C/r \times C}, \mathbf{W}_1 \in \mathbb{R}^{C \times C/r}, \end{aligned} \quad (3)$$

式中: \mathbf{M}_C 为通道注意力; $\sigma(\cdot)$ 为 Sigmoid 函数; \mathbf{W}_0 和 \mathbf{W}_1 为 MLP 的权重参数。

空间注意力模块的输入是一个尺寸为 $H \times W \times C$ 的特征图 \mathbf{F}_2 ,先对输入数据分别进行平均池化和最大池化操作,得到两个尺寸为 $H \times W \times 1$ 的通道特征矩阵,即 $\mathbf{F}_{\text{Avg}}^S$ 和 $\mathbf{F}_{\text{Max}}^S$,并将这两个矩阵进行拼接。然后,利用卷积核与 Sigmoid 函数得到空间域的权重系数。最后,新的权重系数和原特征图相乘即可得到新的特征图。空间注意力模块数据处理流程图如图2(c)所示,空间注意力的计算公式^[22]为

$$\begin{aligned} \mathbf{M}_S(\mathbf{F}_2) = & \sigma\{f^{n \times n}[f_{\text{Avg}}(\mathbf{F}_2); f_{\text{Max}}(\mathbf{F}_2)]\} = \\ & \sigma[f^{n \times n}(\mathbf{F}_{\text{Avg}}^S; \mathbf{F}_{\text{Max}}^S)], \end{aligned} \quad (4)$$

式中: \mathbf{M}_S 为空间注意力; $f^{n \times n}$ 为卷积核大小为 $n \times n$ 的卷积运算。

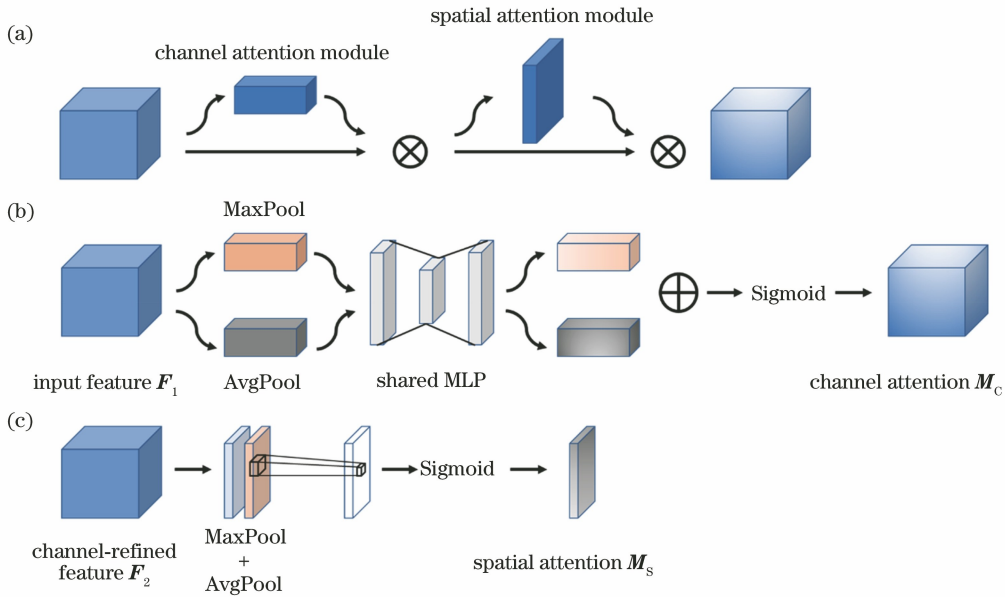


图 2 数据处理流程图。(a) CBAM 中的数据处理流程；(b) 通道注意力模块中的数据处理流程；(c) 空间注意力模块中的数据处理流程

Fig. 2 Flow charts of data processing. (a) Flow chart of data processing in CBAM; (b) flow chart of data processing in channel attention module; (c) flow chart of data processing in spatial attention module

3.1.2 卷积模块

卷积模块是由卷积核、批归一化、非线性激活函数和 CBAM 组成。首先,利用 CBAM 模块计算输入多通道偏振数据的通道注意力和空间注意力,并得到新的特征图,再采用卷积核通过滑窗操作提取新特征图像的区域信息,并生成多个通道的特征映

射图。卷积运算结果经过批归一化处理,数据呈均值为 0、方差为 1 的正态分布^[23],采用非线性激活函数 Relu 来提升网络的非线性描述能力。卷积模块数据处理流程图如图 3(a)所示,其中 CBAM_3 表示 CBAM 的通道注意力模块中参数 r 设置为 3,Conv 表示卷积操作,Batch Norm 表示批归一化处理。

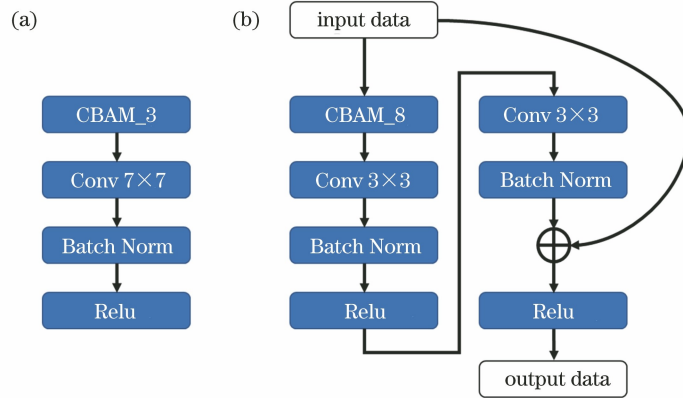


图 3 不同模块的数据处理流程。(a) 卷积模块中的数据处理流程；(b) 残差模块中的数据处理流程

Fig. 3 Flow charts of data processing in different modules. (a) Flow chart of data processing in convolution module; (b) flow chart of data processing in residual module

3.1.3 残差模块

残差模块是基于跳层连接思路构建的天然恒等映射,可有效解决深度卷积神经网络的退化问题,有利于神经网络采用更深的网络设计方案^[12]。该模块的计算公式为

$$x_{l+1} = x_l + F(x_l, W_l) \tag{5}$$

残差模块由直接映射部分和残差部分组成, x_l 为第 l 层的直接映射,残差部分由一个 CBAM、两个卷积操作、两个批标准化处理和一个非线性激活函数组成,对应 $F(x_l, W_l)$ 部分。其中, W_l 为第 l 层

的参数, x_{l+1} 为第 $l+1$ 层的直接映射。残差模块的数据处理流程图如图 3(b) 所示, CBAM_8 表示 CBAM 的通道注意力模块中参数 r 设置为 8。

3.2 基于注意力机制的偏振成像目标分类方法

深度卷积神经网络完成偏振成像目标分类任务需要解决两个问题: 1) 不同偏振图像的目标特性差异明显, 需要采用注意力机制加大利于分类的特征信息的权重; 2) 偏振图像本身存在训练样本数量不足与特征信息有限的问题, 设计合理的网络架构可

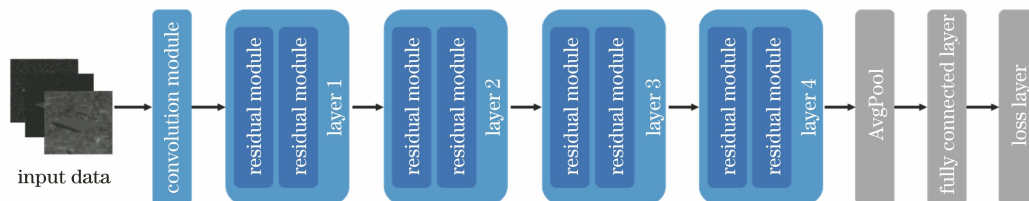


图 4 基于注意力机制的偏振成像目标分类网络结构图

Fig. 4 Structural diagram of polarimetric imaging target classification network based on attention mechanism

传统的 Resnet18 网络通过多层端到端的方式集成了特征提取器和分类器, 所采用的残差网络使得深度网络易于优化^[12], 也是目前应用最为广泛的 CNN 特征提取网络之一。虽然该网络可以完成复杂自然环境中的目标分类任务, 但是无法为偏振图像中的优质信息赋予较大的权重。基于以上考虑, 需要重新设计目标分类网络, 通过嵌入多个注意力模块来提升网络关键信息的筛选能力。该网络要求输入数据的大小为 $224 \times 224 \times 3$ 。首先, 采用一个 CBAM 学习 S_0 、 S_1 和 S_2 图像之间的通道权重与输入图像的空间区域权重, 并通过卷积模块完成偏振数据特征的初步提取。再采用多个残差模块完成后续的特征提取工作, 每个残差模块的头部都含有一个 CBAM 以学习输入数据的权重。最终, 提取的结果经池化层与全连接层映射到损失函数中, 完成本次的模型训练。

所提神经网络将随机梯度下降作为优化器算法, 初始学习率设置为 0.01, 并随着迭代次数的增加而不断降低, 动量大小设置为 0.9, 权重衰减设置

保证目标特征信息的充分提取且可有效避免过拟合问题的发生。所提出的目标分类神经网络是在 Resnet18 网络的基础上添加多个 CBAM, 第一个 CBAM 为高质量的偏振图像和目标区域赋予了更大的权重, 其他 CBAM 用于各残差模块特征图的权重学习。将 CBAM 和残差模块结合可有效解决样本数量受限和网络深度加深所带来的过拟合、梯度爆炸及网络退化等问题。整个卷积神经网络的结构如图 4 所示。

为 0.0001。模型每次训练的样本数量 (batch size) 设为 10, 网络训练的迭代次数 epoch 设为 40, 采用 CrossEntropyLoss 作为损失函数。

偏振成像目标分类方法的具体实现步骤为: 1) 数据预处理。采用 Matlab 软件读取偏振图像, 将每个原始图像的大小改为 $224 \text{ pixel} \times 224 \text{ pixel}$, 并按照比例将其随机分配给训练集与测试集; 2) 模型训练。将偏振数据 ($224 \times 224 \times 3$) 转化为张量, 将其作为神经网络的输入进行训练, 直到损失收敛。

4 实验数据与结果分析

实验设备与实验环境如图 5 所示。实验场地选为中国科学院合肥物质科学研究院安徽光学精密机械研究所, 采用美国 Bossa Nova 公司生产的 SALSAs 被动偏振成像相机完成室外实验的拍摄, 通过电压快速改变铁电液晶的状态以实现偏振参数信息的实时获取。该相机可拍出 $1040 \text{ pixel} \times 1040 \text{ pixel}$ 的单色偏振照片, 观测距离为 $4 \sim 6 \text{ m}$ 。识别的人造样本为铁片和子弹壳, 自然背景选为沙



图 5 实验设备与实验环境。(a) SALSAs 偏振相机系统; (b) 室外拍摄; (c) 目标放置于草地; (d) 裁剪的目标样本

Fig. 5 Experimental equipment and experimental environment. (a) SALSAs polarization camera system; (b) outdoor shooting; (c) targets placed in grassland; (d) cropped target sample

地、裸土和草地,目标随机摆放于不同自然环境中,以获取不同观测角度下的训练样本。铁片和子弹壳与自然背景具有相似的光谱特征,仅依靠人眼与常规光学成像难以对其进行准确分类。

基于注意力机制的偏振成像目标分类方法的训练样本由 S_0 、 S_1 和 S_2 三种偏振图像组成,且所有的图像由 SALSA 同时拍摄。在获取拍摄原图后,需要对原图进行裁剪。将原图裁剪为多个固定大小的训练样本,并对样本的标签进行标记,再通过 Matlab 2020b 软件将其分别存储为训练数据和验

证数据。程序运行环境为 64 位 Windows 操作系统,处理器型号为 AMD Ryzen7-4800H,显卡型号为 NVIDIA Geforce RTX2060,训练平台采用基于 Python3.6 的 Pytorch 1.7.0 机器学习库。

在沙地的自然背景中,裁剪的铁片样本数目为 160,子弹样本数目为 150,沙地背景的样本数目为 172。在裁剪的样本中随机选取 238 个训练样本,选取 244 个测试样本并进行归一化处理。在沙地的自然背景中,不同偏振成像目标分类方法的效果如图 6 所示,其中 L 为损失值。

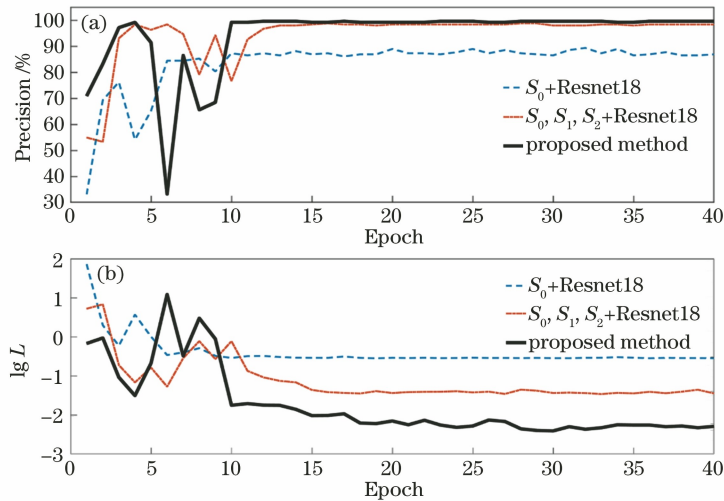


图 6 沙地区域中目标分类实验结果。(a)所提方法与传统方法的分类准确度;(b)所提方法与传统方法的训练损失

Fig. 6 Experimental results for target classification in sandy area. (a) Classification accuracies of proposed method and traditional methods; (b) training losses of proposed method and traditional methods

随着迭代次数的增加,不同神经网络的分类准确率不断提升。传统的 Resnet18 网络在沙地环境中只采用 S_0 图像进行训练,分类准确率达到 86.89%,并保持稳定。随着偏振信息的加入,目标分类的准确率得到大幅提升,并稳定在 98.36%。然而,基于注意力机制的偏振成像目标分类方法能够达到更加出色的识别效果,其分类准确率可以达到 99.59%,高于传统方法。在 1~19 epoch 内,所提分类方法的训练损失下降较快,从最高的 12.1699 降至 0.0060,之后训练损失下降速度变慢并趋于稳定。在沙地环境中,偏振信息的加入可以进一步提升目标分类准确率,且基于注意力机制的偏振成像目标分类方法具有更好的优化效果。

在裸土的自然背景中,裁剪的铁片样本数目为 161,子弹样本数目为 100,裸土背景的样本数目为 102。在裁剪的样本中随机选取 234 个训练样本和 129 个测试样本。在裸土的自然背景中,不同偏振成像目标分类方法的效果如图 7 所示。

随着迭代次数的增加,不同神经网络的分类准

确率不断提升。传统的 Resnet18 网络只采用普通光强图像,分类准确率达到 89.92%。随着偏振信息的加入,目标分类的准确率可升至 93.02%。然而,基于注意力机制的偏振成像目标分类方法的目标分类准确率在多次迭代后可达到 97.67%。因此在裸土环境下,通过在卷积神经网络中加入注意力机制可以大幅提升神经网络的分类性能。同时,在 1~26 epoch 内,所提分类方法的训练损失下降较快,从 1.2850 降至 0.1116,之后训练损失下降速度变慢并趋于稳定,损失值低于传统方法。在裸土环境中,自然背景信息杂乱,存在如石块、落叶和树根等自然物体,这些自然背景信息会对人造目标的正确分类造成一定的干扰,注意力机制可以更加有效地减小无用特征信息的权重。

在草地的自然背景中,裁剪的铁片样本数目为 185,草地背景的样本数目为 225。在裁剪的样本中,随机选取 286 个训练样本和 124 个测试样本。在草地的自然背景中,不同偏振成像目标分类方法效果如图 8 所示。

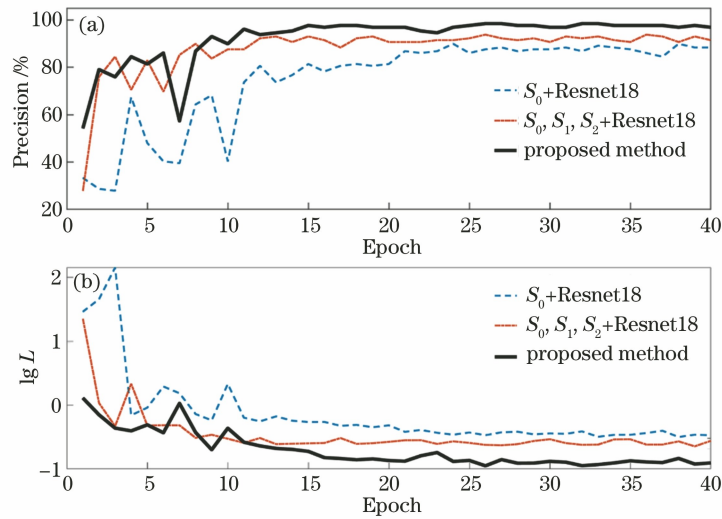


图 7 裸土区域中目标分类实验结果。(a)所提方法与传统方法的分类准确度;(b)提出的方法与传统方法的训练损失
 Fig. 7 Experimental results for target classification in bare soil area. (a) Classification accuracies of proposed method and traditional methods; (b) training losses of proposed method and traditional methods

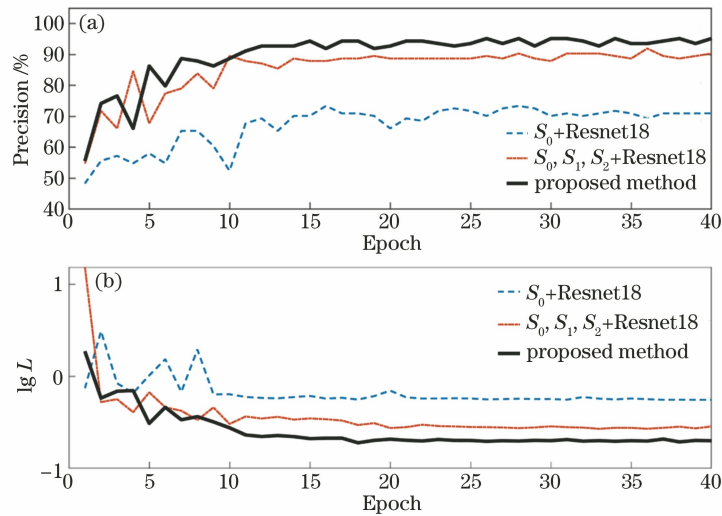


图 8 草地区域中目标分类实验结果。(a)提出的方法与传统方法的分类准确度;(b)本文方法与传统方法的训练损失
 Fig. 8 Experimental results for target classification in grassland area. (a) Classification accuracie of proposed method and traditional methods; (b) training losses of proposed method and traditional methods

随着迭代次数的增加,不同神经网络的分类准确率不断提升。传统的 Resnet18 网络只采用 S_0 图像,分类准确率达到 70.97%,并保持稳定。偏振信息的加入使得目标分类的准确率升至 90.32%。然而,基于注意力机制的偏振成像目标分类方法的目标分类准确率在多次迭代后可达到 95.16%。因此在草地环境中,通过在卷积神经网络中加入注意力机制也可有效提升网络的分类性能。同时,在 1~18 epoch 内,所提分类方法的训练损失下降较快,从 1.8674 降至 0.1896,之后训练损失下降速度变慢并趋于稳定,损失值低于传统方法。在草地环境中,自

然背景信息更为杂乱,除了存在如落叶和杂草等自然物体,还会出现人造目标被杂草遮挡从而目标形状不规则的情况,注意力机制和卷积神经网络的结合可有效降低以上因素的影响,有利于神经网络充分利用目标偏振成像特性,进而提取出更深层次的信息。

在样本数量有限时,虽然多通道的偏振图像可带来更多维度信息,但是复杂的神经网络会发生过拟合且传统的神经网络无法提取更多有效的特征信息。因此,采用注意力机制不仅可以提取不同通道偏振图像的权重,还可以给予重点空间区域更多的关注,从

而进一步提取出更多利于分类的特征信息。基于注意力机制的偏振成像目标分类方法更适用于复杂环境下的人造目标分类。

表 1 为网络参数数量与运行时间。从表 1 可

表 1 网络参数数量与运行时间

Table 1 Numbers of model parameters and running time

Network	Number of model parameters / 10^7	Running time in sandy area /s	Running time in bare soil area /s	Running time in grassland area /s
Resnet18	1.1181891	192	177	197
CBAM-Resnet18	1.1292276	241	221	250

5 结 论

为了提升复杂自然背景中光学偏振图像的目标识别效果,提出了一种基于注意力机制的偏振成像目标分类算法。该方法的主要思路就是将 CBAM 与 CNN 相结合,通过嵌有多个注意力模块的神经网络提取偏振图像之间的非线性关系与目标区域特征,注意力机制为特征明显、利于分类的通道与空间区域赋予较大的权重,最终实现了分类准确率的提升。实验结果表明,基于注意力机制的偏振成像目标分类算法的性能优于传统的深度学习分类算法,对不同自然背景中的偏振数据集的分类准确率可达到 95% 以上,相较于传统的深度学习分类方法得到明显的提升,且不会占用过多的计算资源,适用于复杂背景中光学偏振成像的目标分类。

仅针对铁片与子弹两种人造样本进行了实验验证,所提方法在其他领域光学偏振图像中的目标分类性能还需要进一步验证。后续将针对注意力机制与光学偏振成像原理的深度结合、网络架构的进一步优化及网络学习资源消耗的进一步降低等展开更深入的研究。

参 考 文 献

- [1] Chenault D B, Pezzaniti J L. Polarization imaging through scattering media[J]. *Proceedings of SPIE*, 2000, 4133: 124-133.
- [2] Sadjadi F A, Chun C S L. Automatic detection of small objects from their infrared state-of-polarization vectors[J]. *Optics Letters*, 2003, 28(7): 531-533.
- [3] Thilak V, Creusere C D, Voelz D G. Material classification using passive polarimetric imagery[C]// 2007 IEEE International Conference on Image Processing, September 16-October 19, 2007, San Antonio, TX, USA. New York: IEEE Press, 2007: IV-121-IV-124.
- [4] Hyde IV M W, Schmidt J D, Havrilla M J. A

知,注意力机制的加入并不需要过多的计算资源,也不会消耗大量的运算时间。因此,基于注意力机制的偏振成像目标分类方法适用于偏振样本数量有限的目标分类任务。

geometrical optics polarimetric bidirectional reflectance distribution function for dielectric and metallic surfaces[J]. *Optics Express*, 2009, 17(24): 22138-22153.

- [5] Feng W W, Wei Q N, Wang S M, et al. A hybrid model of polarized BRDF for painted surfaces [J]. *Acta Photonica Sinica*, 2009, 38(4): 962-966.
冯巍巍,魏庆农,汪世美,等. 涂层表面偏振双向反射分布函数的多参量混合模型[J]. *光子学报*, 2009, 38(4): 962-966.
- [6] Chen C, Zhao Y Q, Luo L, et al. Model and analysis of spectropolarimetric BRDF of painted target based on GA-LM method [J]. *Spectroscopy and Spectral Analysis*, 2010, 30(3): 729-734.
陈超,赵永强,罗丽,等. 基于遗传 LM 算法的涂层目标光谱偏振 BRDF 建模分析[J]. *光谱学与光谱分析*, 2010, 30(3): 729-734.
- [7] Zhang J J, Fang Y H. Novel image fusion algorithm for multi-band polarimetric image based on visible light[J]. *Acta Optica Sinica*, 2008, 28(6): 1067-1072.
张晶晶,方勇华. 基于可见光的多波段偏振图像融合新算法[J]. *光学学报*, 2008, 28(6): 1067-1072.
- [8] Zhang J J, Fang Y H. Algorithm and evaluation for polarization image fusion based on edge information [J]. *Opto-Electronic Engineering*, 2007, 34(11): 78-81, 87.
张晶晶,方勇华. 基于边缘信息的偏振图像融合算法及评价[J]. *光电工程*, 2007, 34(11): 78-81, 87.
- [9] Romano J M, Rosario D S. Morphological operators for enhanced polarimetric image target detection[J]. *Proceedings of SPIE*, 2015, 9613: 96130R.
- [10] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [11] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04) [2020-01-15]. <https://arxiv.org/abs/1409.1556>.

- [12] 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 (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [13] Köse S K, Ergünay S, Ott B, et al. Target detection with deep learning in polarimetric imaging [J]. Proceedings of SPIE, 2018, 10794: 107940O.
- [14] Blin R, Ainouz S, Canu S, et al. Adapted learning for polarization-based car detection [J]. Proceedings of SPIE, 2019, 11172: 1117218.
- [15] Mnih V, Heess N, Graves A, et al. Recurrent models of visual attention [EB/OL]. (2014-06-24) [2020-01-15]. <https://arxiv.org/abs/1406.6247>.
- [16] Wang F, Jiang M Q, Qian C, et al. Residual attention network for image classification [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6450-6458.
- [17] Guan S H, Yang G, Lu S, et al. Multi-objective optimization of hyperspectral band selection based on attention mechanism [J]. Acta Optica Sinica, 2020, 40(21): 2128002.
关世豪, 杨桃, 卢珊, 等. 基于注意力机制的多目标优化高光谱波段选择 [J]. 光学学报, 2020, 40(21): 2128002.
- [18] Zhao B, Wang C P, Fu Q, et al. Multi-scale infrared pedestrian detection based on deep attention mechanism [J]. Acta Optica Sinica, 2020, 40(5): 0504001.
赵斌, 王春平, 付强, 等. 基于深度注意力机制的多尺度红外行人检测 [J]. 光学学报, 2020, 40(5): 0504001.
- [19] Cui Z J, An J S, Zhang Y F, et al. Light-weight Siamese attention network object tracking for unmanned aerial vehicle [J]. Acta Optica Sinica, 2020, 40(19): 1915001.
崔洲涓, 安军社, 张羽丰, 等. 面向无人机的轻量级 Siamese 注意力网络目标跟踪 [J]. 光学学报, 2020, 40(19): 1915001.
- [20] Ju M R, Luo J N, Wang Z B, et al. Multi-scale target detection algorithm based on attention mechanism [J]. Acta Optica Sinica, 2020, 40(13): 1315002.
鞠默然, 罗江宁, 王仲博, 等. 融合注意力机制的多尺度目标检测算法 [J]. 光学学报, 2020, 40(13): 1315002.
- [21] Zeng H L, Zhang M W, Liu Z, et al. Experiment of target detection based on polarization imaging using polarization parameter [J]. Opto-Electronic Engineering, 2016, 43(2): 22-26.
曾恒亮, 张孟伟, 刘征, 等. 利用偏振特性因子的偏振成像目标探测实验 [J]. 光电工程, 2016, 43(2): 22-26.
- [22] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module [M] // Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [23] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [EB/OL]. (2015-02-11) [2020-01-15]. <https://arxiv.org/abs/1502.03167>.