

基于神经网络的典型地物的偏振多光谱图像分类 (特邀)

张 颖, 李河申, 王 昊, 孙军华*, 张 晞, 刘惠兰, 吕妍红

(北京航空航天大学 仪器科学与光电工程学院, 北京 100191)

摘要: 相比传统的多光谱成像探测, 偏振多光谱成像探测方法可以探测目标表面的粗糙度、含水量等更多信息, 给目标检测带来了很大便利, 但目前主要用于目标探测, 尚未广泛应用于目标分类。BP 神经网络是目前常用的一种典型神经网络, 可以建立从端到端的映射, 在训练样本集足够大的前提下, 训练完毕且效果良好的神经网络是一种高效、精确、快速的工具。首先, 利用基于旋转偏振片和滤波片的偏振光谱成像探测系统获取了典型地物的偏振多光谱图像, 对图像进行了预处理, 建立了数据集; 其次, 在该数据集上进行了神经网络的训练, 训练后的神经网络可以处理未知的偏振多光谱图像, 并实现了对几种典型地物的分类; 最后, 对神经网络分类的效果进行了评价, 并与其他几种典型分类方法的效果进行了对比, 发现神经网络方法具有更好的分类精度和效果, 相比典型的最大似然分类算法, 其总体分类精度可从 91.7% 提升至 94.2%, Kappa 系数可从 0.851 提升至 0.898。研究结果表明: 基于神经网络的偏振光谱图像分类方法对于改进和优化现有的偏振多光谱图像数据处理方法具有一定的研究意义。

关键词: 偏振多光谱图像; 图像分类; BP 神经网络

中图分类号: TP751 **文献标志码:** A **DOI:** 10.3788/IRLA20220249

Polarized multispectral image classification of typical ground objects based on neural network (*Invited*)

Zhang Ying, Li Heshen, Wang Hao, Sun Junhua*, Zhang Xi, Liu Huilan, Lv Yanhong

(School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100191, China)

Abstract: Compared with the traditional multispectral imaging detection, polarized multispectral imaging detection can detect more information of the detected object surface such as roughness and moisture content, which brings great convenience to target detection. However, at present, it is mainly used for target detection and not widely used in target classification. BP neural network is a typical neural network commonly used at present. Neural network can establish the start-to-end mapping. On the premise that the training sample set is large enough, the trained neural network with good consequences is an efficient, accurate and high-speed tool. Firstly, the polarized multispectral images of typical ground objects were obtained by using the polarized multispectral imaging detection system based on rotating polarizer and filter, and after the images were preprocessed, the data set could be established; Secondly, the neural network was trained on this data set. The trained neural network could process the unknown polarized spectrum images and realize the classification of several typical ground objects; Finally, the effect of neural network classification was evaluated and compared with several other typical

收稿日期: 2022-04-10; 修订日期: 2022-05-10

基金项目: 国家自然科学基金面上项目 (62071024)

作者简介: 张颖, 女, 副教授, 博士, 主要研究方向为智能光电成像探测技术。

通讯作者: 孙军华, 男, 教授, 博士, 主要研究方向为先进光学测量技术。

classification methods. It was found that the neural network method has better classification accuracy and effect. Compared with the typical maximum likelihood classification algorithm, its overall classification accuracy could be improved from 91.7% to 94.2%, and the Kappa coefficient could be improved from 0.851 to 0.898. The results show that the polarized multispectral image classification method based on neural network has certain research significance for improving and optimizing the existing data processing methods of polarized multispectral images.

Key words: polarized multispectral image; image classification; BP neural network

0 引言

随着遥感科学与技术的不断突破和进步,人们越来越重视偏振多光谱成像探测。在传统的图像中,人们所获取的仅仅是电磁波的强度分量,但它只是更完备的偏振信息的一个子集。对于地球和大气中的大多数目标,在与电磁波的相互作用中,如反射、透射和散射,都会产生由其表面性质决定的偏振特征。为了充分利用电磁波中携带的物体信息,就必须研究电磁波的偏振成像。正因为偏振光谱成像能提供更多的信息,包括目标的粗糙度、含水量等,在某些场景下,采用传统成像方式很难被区分的物体,在偏振多光谱图像中可以很容易地被区分出来,对人工目标材料的识别作用更为明显,且空间目标偏振特性不受大气消光影响。因此偏振光谱成像对于物体探测有着先天的优势,它可以用于映射感兴趣场景的偏振状态,并已经在遥感^[1]、目标检测^[2]、材料识别^[3]、金属探伤^[4]和伪装目标探测^[5-6]等领域得到应用。因此,研究物体的偏振光谱图像数据特征,并依据这些特征对图像进行分类,对于物体的探测和识别有着重要意义^[7]。

人工神经网络,简称神经网络,是利用仿生学,根据生物神经元连接方式和原理,以拓扑理论为基础,建立的一种数学模型^[8]。其中,多层前馈网络及反向传播算法,简称 BP 神经网络,是一种典型的神经网络算法模型。它灵活性好,能很好地拟合数学模型中的非线性关系。在偏振多光谱图像分类过程中,利用神经网络输入端到输出端的数据处理过程,省去了中间步骤,直接获得结果,在具备便捷性、快速性的同时兼具准确性。而偏振多光谱探测可以获取更多的信息,这正是神经网络适合处理的。将神经网络应用到偏振多光谱探测中去,将二者的优势相结合,可以极大提高数据处理的效率和探测系统的探测与识别能力。将神经网络应用到偏振多光谱图像数据的处理中,是近年来的发展趋势。

2018 年, Ergunay S 等人^[9]采用神经网络处理采集到的汽车样本的偏振图像,实现了目标检测任务。2019 年, Blin R 等人^[10]将偏振探测和神经网络相结合,完成了车辆的目标探测,并通过实验证明偏振图像在目标分类中的优越性。2020 年,中国科学技术大学的孙瑞^[11]在偏振目标分类中应用了卷积神经网络,取得了良好的效果。但以上研究中仅采用了单波段或全色波段下的偏振图像,而没有采用多光谱信息。而偏振多光谱探测同时探测了目标表面的偏振信息和多光谱信息,再与神经网络相结合,更利于实现目标分类。

该研究使用了一种可附加偏振片和滤波片的偏振多光谱成像系统采集典型地物场景下的偏振多光谱图像,并建立一个应用于典型地物分类的偏振多光谱图像分类的 BP 神经网络。然后,利用训练好的神经网络,对数据进行处理,实现自动分类。最后,实现其他几种典型的分类算法,并比较它们的效果,评估基于神经网络的典型地物的偏振多光谱图像分类方法的有效性,证明了神经网络分类方法在处理偏振多光谱图像中的优越性。

1 偏振多光谱图像数据获取与预处理

该研究采用的图像数据获取系统是传统的偏振多光谱成像系统,其结构如图 1 所示,该系统由光学滤波片、偏振片、镜头、工业 CCD 相机等组成。其原理是入射光通过滤波片,成为中心波长为特定值的窄带光;通过旋转滤波轮更换光学通道中的滤波片,可以获得目标在不同谱段内的光强度;入射光到达偏振片,只有振动方向与偏振片透光方向相同的光分量可以完全透过,其他光分量将被部分或全部阻挡,在四次成像的过程中,通过光机结构控制偏振片的透光方向依次旋转 45°,获取四幅透射光分量互成 45°夹角的偏振光谱强度图像,利用斯托克斯矢量法逐个像元地

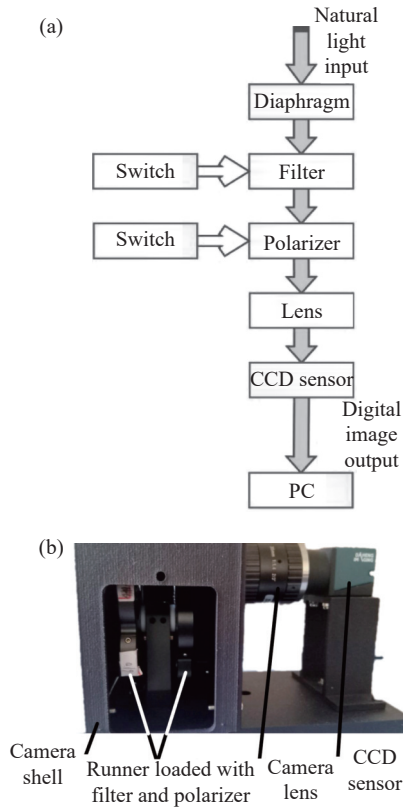


图 1 (a) 偏振多光谱成像系统结构示意图; (b) 系统实物图

Fig.1 (a) Structure diagram of polarized multispectral imaging system; (b) Physical system image

解算偏振度与偏振角,合成偏振度图像与偏振角图像,从而实现对静态目标的偏振多光谱成像探测。

斯托克斯矢量可以由通过上述偏振片的光强分

量计算得到,如公式 (1) 所示:

$$\begin{bmatrix} I \\ Q \\ U \end{bmatrix} = \begin{bmatrix} I(0^\circ) + I(90^\circ) \\ I(0^\circ) - I(90^\circ) \\ I(45^\circ) - I(135^\circ) \end{bmatrix} \quad (1)$$

由公式 (2)、(3) 可计算线偏振度与偏振角:

$$DOLP = \frac{\sqrt{Q^2 + U^2}}{I} \quad (2)$$

$$AOP = \frac{1}{2} \arctan \frac{U}{Q} \quad (3)$$

偏振多光谱图像数据获取系统的 6 片通带带宽为 10 nm 的滤波片安装在转轮上,中心波长分别为: 441、488、532、610、676、808 nm。偏振片透偏角度分别为 0°、45°、90°、135°,安装在另一个转轮上。系统的视场光阑位于外壳上且大小固定。这样设计的原因是转轮切换滤波片和偏振片时,保证系统的视场不会发生变化。CCD 传感器选用大恒图像水星二代 MER2-502-79 U3 M/C 型号相机,可以实现实时刷新,通过 USB3.0 数据接口向 PC 端进行图像数据传输。系统主要参数如表 1 所示。实验期间,采用该系统采集了多组偏振多光谱图像数据。

在获得原始的图像数据之后,需要对其进行预处理。由于采用了逐次采集的准静止成像方式,实际采集过程中会因相机微小振动、风力等原因使得被摄目标在图像中出现位移,这是设备正常使用过程中产生的随机误差。通过测量发现,位移一般在 10 个像素

表 1 偏振多光谱成像系统主要器件参数

Tab.1 Main device parameters of polarized multispectral imaging system

Device	Performance	Parameter
Specifications of CCD sensor	Type	MER2-502-79 U3 M/C
	Port	USB3.0
	Resolution	2 448(H) × 2 048(V)
	Frame rate	79.1 fps
	Sensor	2/3", Sony IMX250Global shutter CMOS
Specifications of lens	Lens mount	Port C
	Type	GSP-25
Specifications of polarizer	Extinction ratio	>1 000:1
	Half bandwidth	10 nm
Specifications of filter	Peak transmittance	50%
	Weight	1.5 kg
Mechanical specifications	Size	18(L) cm×10(W) cm×13(H) cm

以内,因此需要对偏振多光谱图像进行配准。

设 r 和 s 是两幅只有位移差距的图像,位移大小为 (x_0, y_0) , r 与 s 的关系可用公式 (4) 表示:

$$s(x, y) = r(x - x_0, y - y_0) \quad (4)$$

设 R 和 S 分别为 r 和 s 的傅里叶变换,由位移定理,有:

$$S(u, v) = \exp(-j2\pi(ux_0 + vy_0))R(u, v) \quad (5)$$

其互功率谱有:

$$\frac{R(u, v)S^*(u, v)}{|R(u, v)S(u, v)|} = \exp(j2\pi(ux_0 + vy_0)) \quad (6)$$

式中: S^* 为 S 的复共轭。

对公式 (6) 进行傅里叶反变换,会得到一个冲激函数,这个函数在匹配点,即 (x_0, y_0) 处取得最大值,其他位置的值很小。利用这个原理,可以在频率域中对图像的平移进行检测。此算法的配准精度能够达到 0.1 pixel^[12-13], 满足了偏振多光谱图像解析目标偏振信息的要求,进行亚像素级别配准的算法流程如图 2 所示。

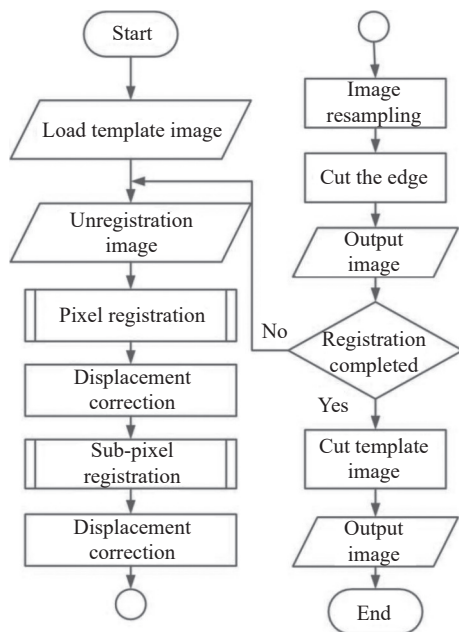


图 2 配准算法流程图

Fig.2 Algorithm flow chart for registration

该研究使用 6 个滤波片和 4 个偏振片,每组偏振多光谱图像数据包括 24 幅灰度图像,对这 24 幅图像进行互功率谱配准和中值滤波以后,得到了包含目标区域多光谱信息和偏振信息的数据立方体,使之可以

在程序脚本中以元胞数组格式来进行储存和显示,每一个元胞中存储一幅灰度图像,经过配准的图像消除了彼此间的微小位移,使得每幅图像相同坐标处的像点对应同一个物点。

2 神经网络实现

一个完整的神经网络包括输入层、隐含层、输出层。理论上,没有隐含层的网络仅能拟合线性函数;隐含层数量为 1 的网络可以拟合任意从输入空间到输出空间的连续映射函数;隐含层数量大于等于 2 时,可以拟合任意精度的任意平滑映射,同时可以学习复杂的描述^[8]。对于数据集较为简单的项目而言,单隐含层网络通常已经足够使用,如图 3 所示。

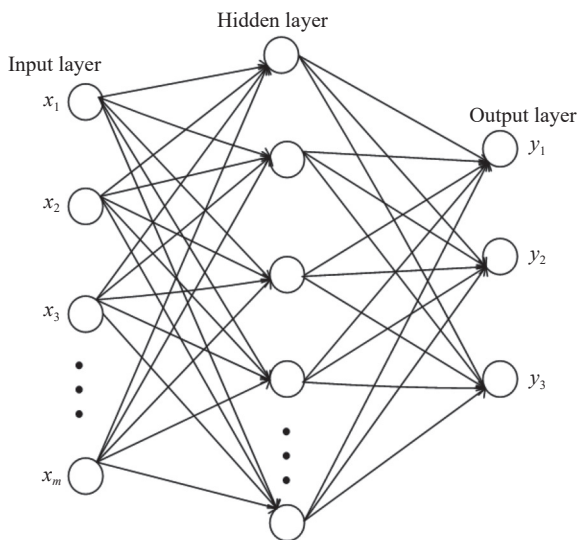


图 3 网络结构示意图

Fig.3 Network structure diagram

输入层的神经元数目要与输入特征向量维数相对应,因此输入层神经元数目为 24。输出层的神经元数目要与输出相对应。该研究将对被摄目标像素进行分类,因此输出层神经元数目为 3,激活函数使用 softmax。

隐含层的神经元数目确定则依赖于一些经验化原则,过少的神经元数目会导致欠拟合,而过多的神经元数目则可能导致过拟合,同时延长网络的训练时间。过拟合和欠拟合都会降低网络的分类效果^[11]。这些经验化的原则包括:隐含层神经元数量应不超过输入层神经元数量的 2 倍;一般介于输入层神经元数

量和输出层神经元数量之间;可以选择输入层和输出层神经元数量之和的 2/3 等。因此可以大致确定隐含层神经元数量在 9~24 之间。

2.1 训练样本选取

在训练 BP 神经网络前需要依据所采集的偏振多光谱图像数据搭建训练集和验证集。神经网络在训练集上进行训练,学习不同地物类别在偏振光谱图像上的灰度分布特征和规律。因此,训练集应当选取被摄目标中最典型、最重要、最有代表性的区域,训练样本应涵盖所有被摄的地物类别,并选择研究的重点、典型地物进行采集。验证集用以验证网络训练后的效果。网络在训练集上取得不错的准确率后,要在验证集上进行验证。

样本的选取方法采用人机交互方法,利用计算机处理图像并显示,操作者利用鼠标点选出目标地物的

区域,建立训练样本集。这种方法需要有格式化的数据作为输入,速度较快,适合处理大量数据。该研究搭建的地物场景共包括“土壤”、“植被”、“伪装植被”三类,其中伪装植被的材质是塑料。在點選数据时首先选择某个波段的图象进行显示,然后,利用目视解译的方法,同时对照拍摄时搭建的场景彩色照片,对获取到的偏振多光谱图像中的像素进行分类,并记录其灰度值的特征向量和类别标签。将获取的数据集进行随机划分,该研究划分后的训练集共 5 450 个样本,验证集共 1 850 个样本。图 4(a) 所示为其中一组图像中點選的样本。

在获得样本之后,还可以根据样本数据,利用公式 (1)~(3),分别计算每一个样本在各个波段下的 I 、 Q 、 U 分量以及偏振度和偏振角。其中一个“伪装植被”样本的 I 、 Q 、 U 分量图像及其偏振度、偏振角图

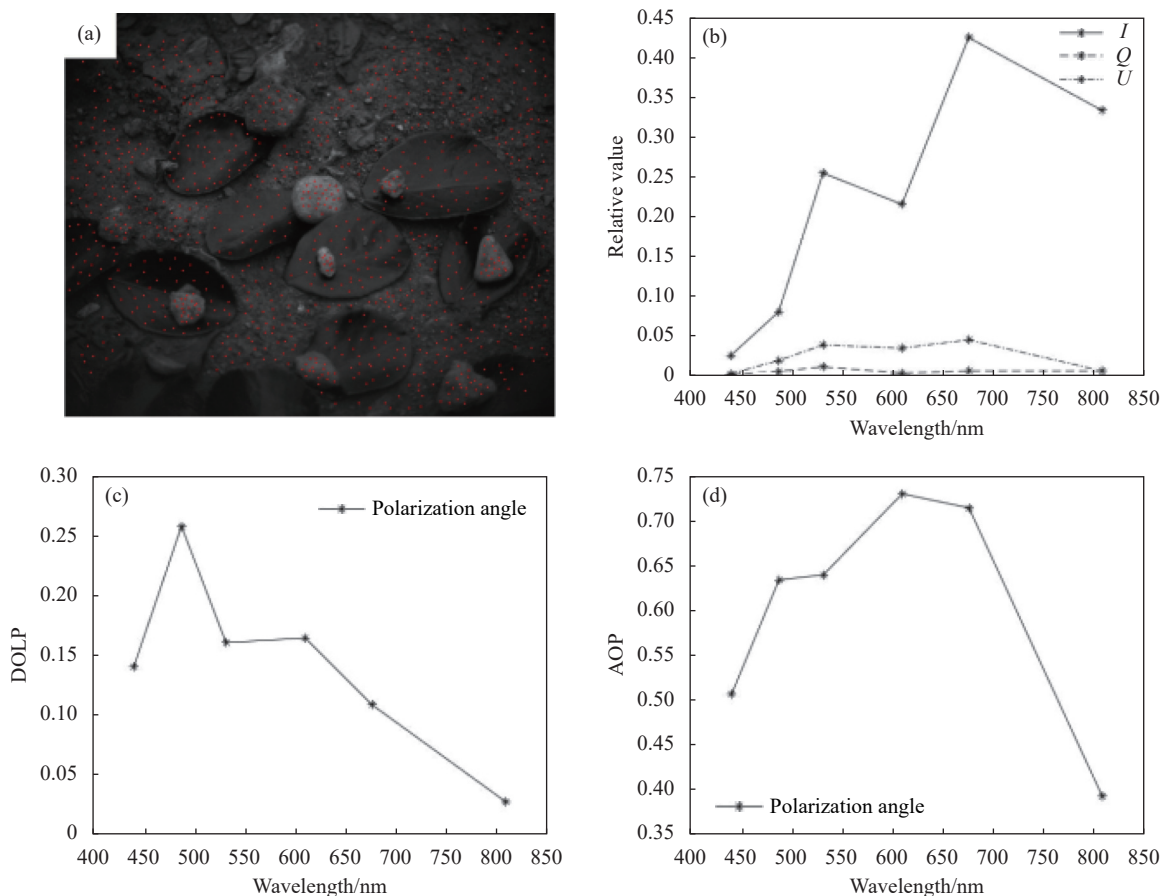


图 4 (a) 从一组图像中點選的样本; (b) 一个样本的 I 、 Q 、 U 分量; (c) 一个样本的线偏振度; (d) 一个样本的偏振角

Fig.4 (a) Samples selected from a set of images; (b) I , Q , U components of a sample; (c) Degree of linear polarization of a sample; (d) Polarization angle of a sample

像如图 4(b)~(d) 所示。

2.2 神经网络搭建和训练

该研究在 Python3.7 环境下, 基于谷歌 Tensor Flow2.3-gpu 开源框架, 采用 Adam 优化器以及稀疏交叉熵损失函数 (keras.Sparse Categorical Crossentropy), 在多次试验以后, 项目训练出了效果较为良好的神经网络模型。谷歌 TensorFlow2.3-gpu 框架可以调用 NVIDIA 系列 GPU 参与神经网络的训练和预测, 能够显著提高神经网络分类方法的运行速度。Adam 优化器^[14] 由于其性能优秀、实现简单、适用于不稳定目标函数, 在很多神经网络中得到应用。

项目建立的网络模型的隐含层有 12 个神经元, 在经过整个数据集上 150 次迭代训练后, 其在训练集上准确率达到 97.3%, 损失函数值低于 0.58, 在验证集上准确率达到 94.2%, 损失函数值低于 0.61。图 5(a) 和图 5(b) 展示了其准确率和损失函数随训练次数的变化情况。

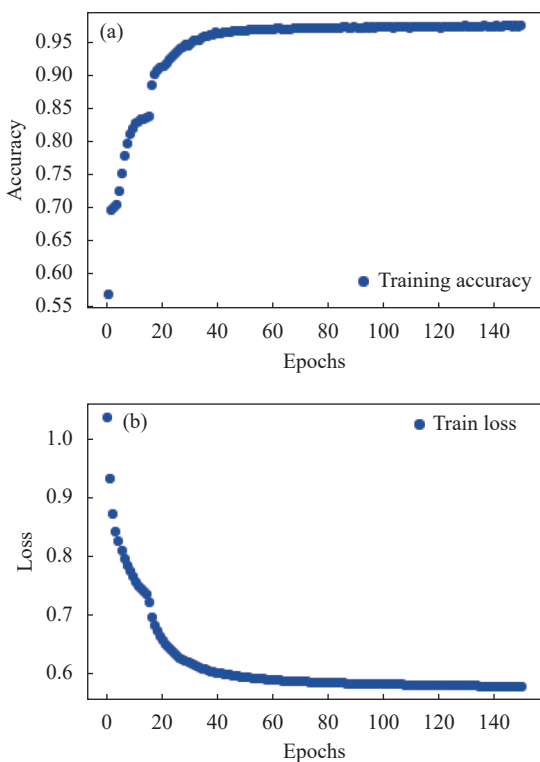


图 5 (a) 网络的训练集准确率迭代图像; (b) 网络的损失函数迭代图像

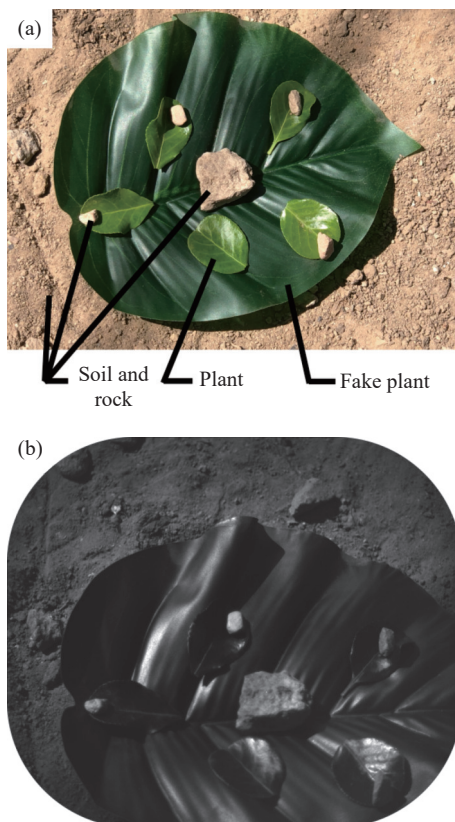
Fig.5 (a) Training accuracy iterative image of network; (b) Training loss iterative image of network

3 神经网络分类效果的评价

3.1 神经网络分类结果

网络训练完毕后, 需要对网络的分类效果进行评价。首先可以通过目视解译的方法大致判断分类的效果。获得一组图像数据的预测结果比较简单。首先将经过预处理的数据导入, 然后调用已保存的神经网络模型并进行编译, 再对导入的数据逐像素进行预测。神经网络会自动预测该像素归类为每一类的概率, 并输出这个预测的结果。将预测结果恢复至原输入图片大小, 就获得了输出图像。图 6(a) 展示了获取一组偏振多光谱图像时的实验场景, 图 6(a) 为这个场景下的一幅原始图像, 图 6(c) 展示了神经网络对该组图像的分类结果。

由于引用的偏振多光谱成像系统中添加了光阑, 导致获取到的图像四角有较为明显的渐晕, 因此通过圆角截去了这部分像素。从图 6 可以看出, 总体的分类效果较为良好, 基本完成了土壤、塑料假叶片、植被叶片的分类。



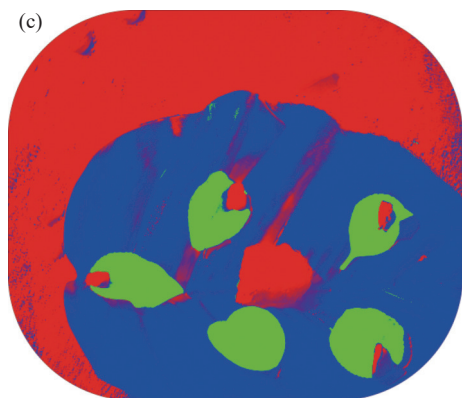


图 6 (a) 实验场景; (b) 一幅原始图像; (c) 神经网络对该组图像分类结果

Fig.6 (a) Experimental scene; (b) One original image; (c) Classification results of this set of images by neural network

3.2 分类效果的精度评价

精度评价是图像分类后必须进行的重要任务。通过对分类结果的定量化评价,可以评估神经网络模型分类的效果,指示神经网络结构或参数进行优化的方向,进而对网络结构或参数进行针对性的改善。

评价某种分类方法精度时,可以采用的主要指标包括以下几种^[15]:

(1) 总体分类精度。总体分类精度即分类的总体分类精度,计算方法为正确分类的像素数量与总像素数量之比。

(2) 混淆矩阵。混淆矩阵又称误差矩阵,主要用于比较分类结果与真实情况之间的差异性,可以用混淆矩阵的格式的形式表示分类精度。在混淆矩阵中,错分类的样本数目分布在非对角线上,而正确分类的样本分布在对角线上,因此可以借助混淆矩阵方便地

查看多种类别的混淆情况。

(3) Kappa 系数。Kappa 系数是一个更加客观描述分类精度的量。其计算方法为:

$$Kappa = \frac{N \sum_{i=1}^n x_{ii} - \sum_{i=1}^n x_{i-x_i}}{N^2 - \sum_{i=1}^n x_{i-x_i}} \quad (7)$$

式中: N 表示总像素数量, x_{ii} 表示第 i 类像素归类为第 i 类地物的数量, x_{i-x_i} 表示第 i 类像素归类为非第 i 类地物的数量和非第 i 类像素归类为第 i 类地物的数量。

综上所述,在不同分类方法之间定量化评价分类精度时,选择总体分类精度、混淆矩阵、Kappa 系数这三项指标。

在神经网络方法实现后,项目基于 MATLAB 平台实现了最小距离分类器、最大似然分类器、支持向量机三种典型分类方法,并对比神经网络分类方法与这些典型分类方法在处理偏振多光谱图像中的效果,如表 2 所示。可以看出,神经网络分类方法效果最好,其次是最大似然分类方法。使用 RBF 内核的支持向量机方法和最小距离分类方法的分类效果最差。总体上看,神经网络分类方法相比其他分类方法,虽然需要较长时间进行训练,但训练完毕的神经网络可以快速调用,训练好的网络处理一组 24 幅 2448×2 048 分辨率的图像数据用时约 3~5 s,使用的计算机搭载一片 NVIDIA GTX 1050 显卡和一片 Intel i5-7300 HQ 处理器,与最大似然分类方法和最小距离分类方法相比耗时差距不大,二者较为接近。训练完毕的神经网络可以在保证快速性的同时,取得更佳分类效果。

表 2 不同分类方法效果对比

Tab.2 Comparison of different classification methods

Classification method	Training set accuracy	Validation set accuracy	Accuracy ranking	Kappa coefficient	Kappa coefficient ranking
Network	97.3%	94.2%	1	0.898	1
MLC	95.3%	91.7%	3	0.851	3
MDC	76.5%	76.5%	5	0.584	5
SVM-RBF	71.9%	73.6%	6	0.456	6
SVM-linear	88.9%	85.5%	4	0.725	4

4 结 论

将神经网络应用到偏振光谱探测中去,将二者的

优势相结合,可以极大提高数据处理的效率和探测系统的探测与识别能力。将神经网络分类方法应用到偏振光谱图像数据的处理中,是近年来的发展趋势。

该研究尝试将神经网络分类方法应用在偏振光谱图像分类中,取得了较好的效果,并将其分类效果同其他典型分类方法进行了对比。

结果表明,虽然神经网络需要更长的训练时间,但它给出的分类效果是最好的,其总体分类精度可达 94%, Kappa 系数可达 0.898。除此之外,同样是神经网络分类,采用不同参数的神经网络也会取得不同的训练和分类效果,表明了神经网络结构参数对神经网络模型有着很大的影响。

该研究仍然有一些不足,需要加以改进和完善。尽管神经网络方法在该研究中具有较好地效果,但鉴于实验场景搭建的单一性和人机交互方法选取样本过程中的局限性,其普适性和泛化性仍然有待进行更深入地研究。此外,将目标的视觉特征与目标的偏振多光谱数据相结合,将图像分割、语义分割等计算机视觉处理手段应用到偏振多光谱图像的处理中,是下一步的研究和改进方向。

参考文献:

- [1] Tyo J S, Goldstein D L, Chenault D B, et al. Review of passive imaging polarimetry for remote sensing applications [J]. *Applied Optics*, 2006, 45(22): 5453-5469.
- [2] Tuo H, Shi G, Luo X. Infrared polarization image fusion and detection system for ship target [J]. *Journal of Physics: Conference Series*, 2021, 1802(2): 022070.
- [3] Krylov V A, Moser G, Sebastiano Bruno Serpico, et al. Supervised high-resolution dual-polarization SAR image classification by finite mixtures and copulas [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2011, 5: 554-566.
- [4] Wang Fangbin, Sun Fan, Zhu Darong, et al. Metal fatigue damage assessment based on polarized thermograph [J]. *Acta Optica Sinica*, 2020, 40(14): 124-134. (in Chinese)
- [5] Wang Xia, Liang Jian'an, Long Huabao, et al. Experimental study on long wave infrared polarization imaging of typical background and objectives [J]. *Infrared and Laser Engineering*, 2016, 45(7): 0704002. (in Chinese)
- [6] Zhou Qiangguo, Huang Zhiming, Zhou Wei. Research progress and application of polarization imaging technology [J]. *Infrared Technology*, 2021, 43(9): 817-828. (in Chinese)
- [7] Li Yanan, Sun Xiaobing, Mao Yongna, et al. Spectral polarization characteristic of space target [J]. *Infrared and Laser Engineering*, 2012, 41(1): 205-210. (in Chinese)
- [8] Ian Goodfellow, Yoshua Benjio, Aaron Courville. Deep Learning[M]. UK: MIT Press, 2017.
- [9] Ergunay S, Wellig P, Leblebici Y, et al. Target detection with deep learning in polarimetric imaging[C]//Target and Background Signatures, 2018.
- [10] Blin R, Ainouz S, Canu S, et al. Adapted learning for polarization-based car detection[C]//Fifteenth International Conference on Quality Control by Artificial Vision, 2019.
- [11] Sun Rui. Polarimetric imaging target classification methods and experiments based on convolutional neural network[D]. Hefei: University of Science and Technology of China, 2020. (in Chinese)
- [12] Yang Changjiu, Li Shuang, Qiu Zhenwei, et al. Study on image registration of simultaneous imaging polarization system [J]. *Infrared and Laser Engineering*, 2013, 42(1): 262-267. (in Chinese)
- [13] Yu Xuelian, Chen Qian, Sui Xiubao, et al. Registration method of infrared images based on the optimal peak of phase-correlation criterion [J]. *Infrared and Laser Engineering*, 2013, 42(9): 2589-2596. (in Chinese)
- [14] Kingma D, Adam B J. A method for stochastic optimization [J]. *arXiv preprint arXiv*, 2014, 1412: 6980.
- [15] Zhang Dongyan, Dai Zhen, Xu Xingang, et al. Crop classification of modern agricultural park based on time-series Sentinel-2 images [J]. *Infrared and Laser Engineering*, 2021, 50(5): 20200318. (in Chinese)