

深度置信网络模型的机载多光谱数据罂粟识别

陆永帅, 李元祥, 彭希帅

(上海交通大学 航空航天学院, 上海 200240)

摘要:针对传统识别算法在罂粟植物地块识别上精度不足的问题,提出了基于深度置信网络模型的机载多光谱数据罂粟识别算法。采用模拟人脑多层结构的方式,可以对数据自动地进行特征提取,挖掘内在联系,建立更准确的识别模型;同时将随机隐退过程引入到罂粟识别的深度网络中,避免了传统神经网络因为随机初始化而陷入局部最优解的情况。无人机航拍数据的实验表明,在小样本罂粟训练集的情况下,与支持向量机和传统神经网络方法相比,基于随机隐退的深度置信网络模型可取得更好的识别结果。

关键词:罂粟识别;多光谱数据;深度置信网络;深度学习;随机隐退

doi: 10.3969/j.issn.1000-3177.2017.04.015

中图分类号: TP75 文献标志码: A 文章编号: 1000-3177(2017)152-0098-06

Poppy Detection in Airborne Multispectral Data Based on Deep Belief Network

LU Yongshuai, LI Yuanxiang, PENG Xishuai

(School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: To solve the problem of low detection precision on poppy planting blocks, we propose a poppy detection algorithm based on DBN with dropout, which avoids the situation that traditional neural networks caught in the local optimal solution because of random initialization. This algorithm allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. Aerial multispectral data from unmanned aerial vehicles was used for evaluation. The experimental results show the effectiveness of the proposed method with small training sample set in comparison with support vector machine and traditional neural network methods.

Key words: poppy detection; multispectral data; deep belief network; deep learning; dropout

0 引言

我国面临的禁毒形势严峻复杂,毒品犯罪多发高发,禁毒工作任务十分艰巨^[1]。精确识别毒品原植物是禁毒工作的关键,而罂粟作为一种主要毒品原植物,是需要识别的重点。非法罂粟的种植一般具有隐蔽性的特点,传统的人工地面识别排查的手段需要耗费大量的人力物力。而利用无人机航拍技术来进行非法罂粟的识别具有覆盖范围广、信息源可靠、节约资源等优点^[2]。因为罂粟的样本非常难

以得到,在实际过程中,很难得到大量带有标签的罂粟训练数据。因此,研究针对小样本训练集的识别模型十分有必要。

与全色相机相比,机载多光谱相机能得到地物多个通道的光谱反射信息^[3],通过比较不同地物在不同通道上反射率的差异,可以准确识别出特定的地物^[4]。高光谱数据^[5]虽然可以得到更高光谱分辨率的图像,但因为高光谱相机的质量已经超出了一般无人机的载荷要求,因此不做考虑。

在利用光谱技术识别目标的研究中,系统建模、

收稿日期: 2016-04-26 修订日期: 2016-11-28

作者简介: 陆永帅(1991—),男,硕士研究生,主要研究方向为图像处理与机器学习。

E-mail: lys_2008@sjtu.edu.cn

通信作者: 李元祥(1967—),男,副教授,主要研究方向为遥感图像解译、图像识别、图像重构与评估。

E-mail: yuanxli@sjtu.edu.cn

辨识和预测中有很多算法。但这些方法需要已知数据的内在规律并且人工选择出特征才能很好地进行识别^[6]。而传统机器学习的方法,如人工神经网络、支持向量机等往往不能很好地归纳出目标的特征,建立精准的识别模型。

2006年,多伦多大学机器学习领域的著名教授Hinton和他的学生Salakhutdinov在国际顶尖学术期刊《Science》上提出利用深度置信网络(deep belief network, DBN)实现数据的降维及分类^[7],深度学习开始备受关注。

深度神经网络采用模拟人脑深层结构的方式,对输入数据从低层到高层渐进地进行特征提取,最终形成适合的适合目标识别的模型,从而提升识别的准确性^[8]。目前,深度学习在语音识别^[9]、视觉识别^[10]、目标检测、遥感图像分类^[11]等领域都有着广泛的应用。

本文提出一种基于随机隐退 DBN 模型的多光谱数据罂粟识别方法,在真实航拍的罂粟试验田数据进行实验,以提高罂粟作物的识别精度。

1 DBN 原理简介

结构上,深度置信网络由多个限制玻尔兹曼机(restricted Boltzmann machine, RBM)层以及一层反向传播网络(back-propagation, BP)组成,如图1所示。

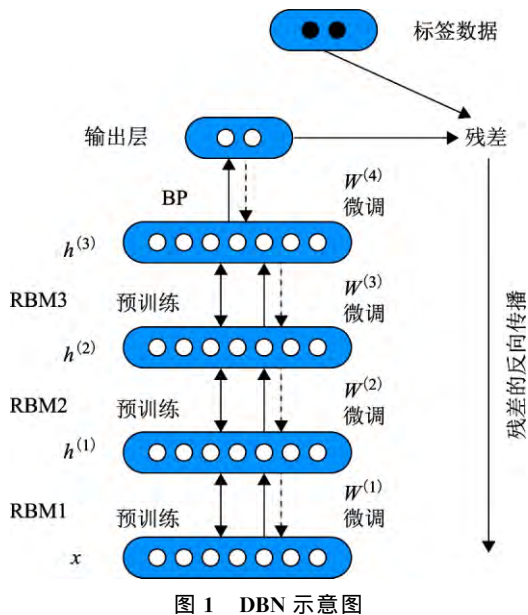


图1 DBN 示意图

DBN 中主要的训练方式可以分为“预训练(pre-training)”和“微调(fine-tuning)”2个步骤。首先进行的是无监督的预训练阶段,采用逐层训练的方式,对各层 RBM 进行训练,低一层的 RBM 输出

层(隐含层)作为高一层的 RBM 输入层(可见层)。在预训练结束后,对最后一层 BP 网络进行有监督方法的训练,把实际输出与真值之间的误差逐层向后传播,对整个 DBN 网络的权值进行微调。

预训练过程的加入,使得深层 BP 神经网络的权值初始化过程更为合理,克服了传统网络因为随机初值而造成的训练困难和局部最优的问题。

1.1 限制玻尔兹曼机

限制玻尔兹曼机由一个可见层(visible layer)与一个隐含层(hidden layer)组成^[12],如图2所示。其中, v 和 h 分别表示可见层和隐含层, W 表示它们之间的连接权值。各层的内部是两两独立的,而相邻层之间是两两双向互通的。

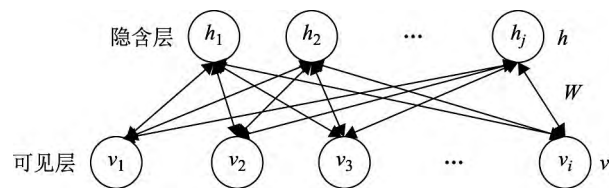


图2 限制玻尔兹曼机

设可见层节点数目为 I ,第 i 个节点状态为 v_i 。同样,隐含层节点数目为 J ,第 j 个节点的状态为 h_j 。假设每个节点均为二值变量,对于一组状态 (v, h) ,RBM 的能量定义为

$$E(v, h | \theta) = - \sum_{i=1}^I a_i v_i - \sum_{j=1}^J b_j h_j - \sum_{i=1}^I \sum_{j=1}^J w_{ij} v_i h_j \quad (1)$$

式中: $\theta = (a_i, b_j, w_{ij})$ 是模型的参数; a_i 为 v_i 的偏置量; b_j 为 h_j 的偏置量; w_{ij} 为 v_i 和 h_j 之间的连接权重。根据上面的能量函数可以得到 (v, h) 的联合概率分布为

$$p(v, h | \theta) = \frac{e^{-E(v, h | \theta)}}{Z(\theta)} \quad (2)$$

式中: $Z(\theta) = \sum_v \sum_h e^{-E(v, h | \theta)}$ 为归一化项。在实际应用中,更关心联合概率分布 $p(v, h | \theta)$ 的边缘分布,即 $p(v | \theta)$ 。

由于 RBM 每层节点相互独立,有

$$p(h_j = 1 | v, \theta) = \sigma(b_j + \sum_{i=1}^I v_i w_{ji}) \quad (3)$$

式中: $\sigma(x) = 1/(1 + e^{-x})$ 为 sigmoid 函数。因为 2 层之间为两两双向相连的,用隐藏层的状态也可以推出可见层的状态

$$p(v_i = 1 | h, \theta) = \sigma(a_i + \sum_{j=1}^J h_j w_{ji}) \quad (4)$$

RBM 采用迭代的方式训练,训练的目的在于得

到模型参数 $\theta = (a_i, b_j, w_{ij})$, 以用于拟合给定的训练数据。参数 θ 可以通过求得训练集上的极大对数似然函数得到。

对比散度算法(contrastive divergence, CD)是 Hinton 提出的用于计算 RBM 参数更新值的一种迭代算法^[13]。他提出了各个参数的更新规则如下:

$$\Delta w_{ij} = \epsilon(\langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{recon}}) \quad (5)$$

$$\Delta a_i = \epsilon(\langle v_i \rangle_{\text{data}} - \langle v_i \rangle_{\text{recon}}) \quad (6)$$

$$\Delta b_j = \epsilon(\langle h_j \rangle_{\text{data}} - \langle h_j \rangle_{\text{recon}}) \quad (7)$$

式中: ϵ 为与训练的学习率, $\langle h_j \rangle_{\text{data}}$ 为训练数据集所定义的分布之上的数学期望, $\langle h_j \rangle_{\text{recon}}$ 为重构后的模型所定义的分布上的数学期望。

1.2 BP 网络

BP 神经网络是一种可以反向传递误差并不断修正网络权值的神经网络模型。它分为正向传播和反向传播 2 个过程。

1) 正向传播。输入样本从输入层经过隐含单元一层层进行处理, 传向输出层, 这种逐层处理的过程, 每层神经元的状态只对下一层神经元的状态产生影响。在输出层将当前的输出与期望的输出进行比较, 如果当前的输出不等于期望输出, 则进行反向传播。

2) 反向传播。将误差信号按正向传播的通路反向传回, 对隐含层的各个神经元的连接权进行修正, 以使误差信号最小。

深度置信网络其实就是一个利用预训练事先得到各层网络权值初值的 BP 神经网络, 在预训练结束后, 后面的训练步骤与传统 BP 网络一致。

1.3 基于随机隐退的 DBN 方法

传统的 DBN 方法就是在深层 BP 网络的前方加入了预训练过程。这样的网络, 在对于小样本低维度的数据进行训练时, 常常会出现过拟合的问题, 即训练集结果达到非常高的逼近精度, 但将模型运用到测试集时, 效果却不理想。

针对这个问题, Hinton 在 2012 年提出了随机隐退(dropout)的 DBN 方法^[14], 文献^[15]也说明了随机隐退的方法对于提高小样本 DBN 模型的泛化能力很有帮助。随机隐退的思想就是在微调过程的正向过程, 随机使某些节点的输出值变为 0, 反向过程也不对这个节点的权值进行调整, 相当于这个节点在这次迭代中不作为网络的一部分^[16]。

具体的过程就是设定一个隐退比例 α , 训练的每次迭代过程中为每个节点产生一个 $[0, 1]$ 的随机数 r , 当 $r < \alpha$ 将此节点作隐退操作。在测试阶段, 所有节点都发挥作用, 但每个节点的输出值都要乘

以 $(1 - \alpha)$ 。

随机隐退其实就相当于同时训练很多模型, 最后将这些模型的结果加权平均进行融合, 提高了模型的泛化能力, 提高了小样本 DBN 模型的识别精确度。

2 基于 DBN 的罂粟识别

2.1 光谱特性分析

实验数据(见 3.1 节)是通过搭载三通道多光谱相机的无人机对某处罂粟试验田进行航拍得到的。在数据集 1 中对每一类地物随机挑选 50 个数据点, 可以得到图 3 的光谱特性图。

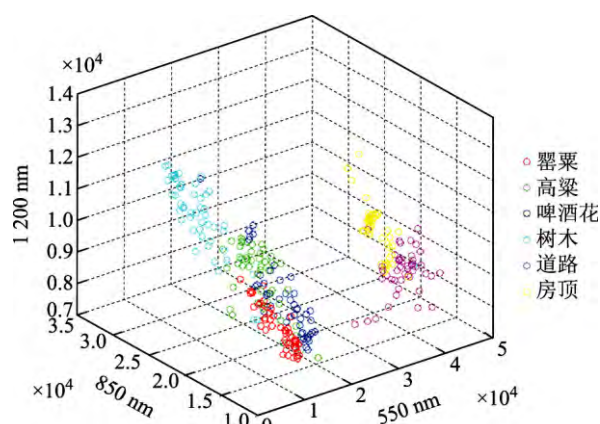


图 3 实验数据集中典型地物的三通道光谱特性

通过对典型地物的三通道光谱曲线进行分析, 可以看出, 植物(高粱、罂粟)与人造建筑(道路、房顶)的光谱曲线存在着明显的差异。其中, 人造建筑在 550 nm 波段的反射强度明显强于植物, 而在 850 nm 处又弱于植物。利用这个特征可以很好地将建筑与目标区分开。而乔木(高大的树木)与灌木(高粱、罂粟)之间光谱特性也有着明显的差异, 乔木在各波段的光谱反射强度都高于灌木, 在 850 nm 波长处则最为明显。而同是灌木的罂粟与高粱、啤酒花, 光谱特性极为相似, 区别在 550 nm 波段, 罂粟的反射强度略弱, 而 1 200 nm 波段罂粟的反射强度也略弱。若要更好地区分罂粟与其他灌木, 还需要更多波段的支持或者通过纹理特性来辅助识别。

2.2 特征输入

DBN 的样本输入是一个向量。识别罂粟, 最主要的特征就是其光谱特征。因为数据波段数较少, 为了更好地进行识别, 本文对于每一个要识别的像素, 需要考虑包括其周围相邻像素的一片区域。邻域窗口的大小设为 3×3 , 而每个像素位置实际包括 3 个通道的信息, 可以将其展开维度为 $3 \times 3 \times 3 = 27$ 的向量作为 DBN 的初始输入, 如图 4 所示。

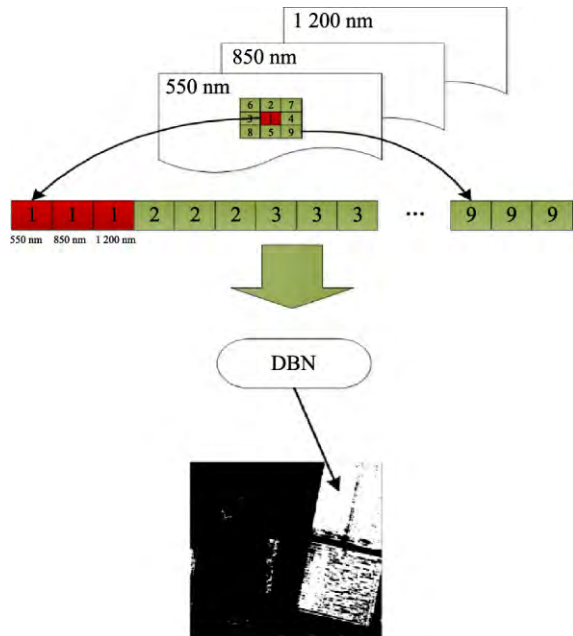


图4 基于DBN模型的罂粟识别方法流程图

在进行DBN训练的时候,几个RBM层会逐层训练,自动对目标输入进行降维和特征提取,学习罂粟作物内在的特有特征,并最后利用标签数据进行微调,从而得到精确的识别模型。

预测阶段,将样本以同样的数据结构输入模型,通过向前传播计算可以得到识别的结果。

因为罂粟的样本非常难以得到,在实际过程中,很难得到大量带有标签的罂粟训练数据。因此,本文采用基于随机隐退的DBN方法,通过少量训练数据即可得到较精确的罂粟识别模型。

3 实验与分析

3.1 实验数据

实验数据集是2张某地罂粟试验田的航拍数据,拍摄时间为7月,三通道波长分别为550 nm、850 nm和1200 nm,空间分辨率为0.1 m,经过了几何校正三通道配准等预处理操作。将三通道的数据分别作为R、G、B通道,可以得到图5的伪彩色图像。

该地区除了罂粟目标以外,还包括了高粱、啤酒花等多种作物和道路房子等人造建筑。数据集1和数据集2均有80万像素点,罂粟区域分别占275 538像素和246 578像素。

实验中,罂粟作为需要识别的目标,而其他非罂粟区域统一标示为背景,内部类别不做区分。

3.2 识别结果评价

为了表示识别的准确性,需要事先定义性能指标。

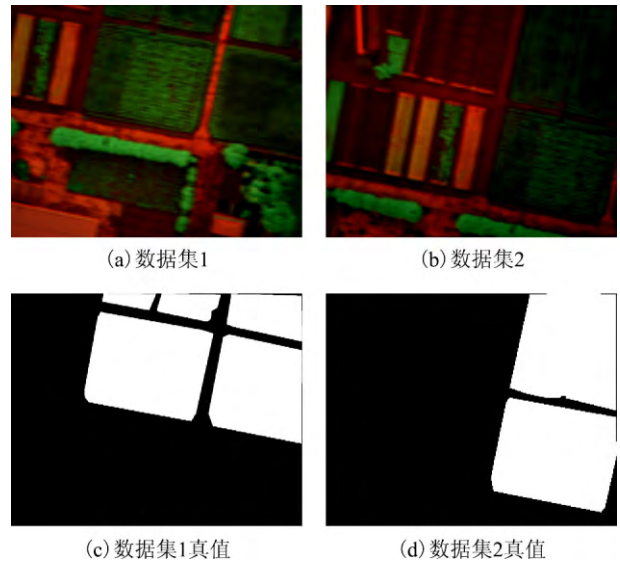


图5 罂粟实验田伪彩色图像与真值图像

以 N_T 表示场景内目标的像素个数; N_D 表示正确识别的目标数(真实是罂粟检测为罂粟); N_{FA} 表示虚警数(真实是背景检测为罂粟)。准确率表示为 $P_D = N_D/N_T$,虚警率表示为 $P_{FA} = N_{FA}/(N_{FA} + N_D)$ 。在比较的时候发现,准确率和虚警率是对立的关系,常常准确率高的算法同时也有着较高的虚警率。本文采用品质因子(figure of merit, FoM)作为衡量罂粟识别性能的评价指标^[17]。

$$\text{FoM} = \frac{N_D}{N_T + N_{FA}} \quad (8)$$

FoM大小能够很有效地反映识别方法的性能,FoM越大,识别方法的性能越好,识别率越高,虚警率越低。

3.3 实验

1)DBN的结构分析实验。目前,DBN的隐层数和隐层节点数的选择大多取决于经验。因此本文先通过实验来确定一个较优的模型参数。假设各隐层的节点数相同,隐层数在(1,2,3,4)中选择,节点数在(5,10,15,20,25,30,35)中选择,每层预训练迭代次数设为50,微调次数为5 000。通过网格搜索之后得到的结果如图6所示。

可以看出,在同样节点下,随着隐层数的增加,识别结果是呈现先变优后变差的趋势。变优是因为随着网络深度的增加,DBN所提取的特征更抽象,建模能力也随之增加。但网络深度增加的同时也需要更多的迭代次数用于调整权值参数,若训练达不到反而会使模型识别的精度下降。同样,隐层节点数太少无法满足特征表示的需要,太多会出现过拟合等问题。因此,取图中的极值点3层15个节点作为下文DBN实验的隐层参数。

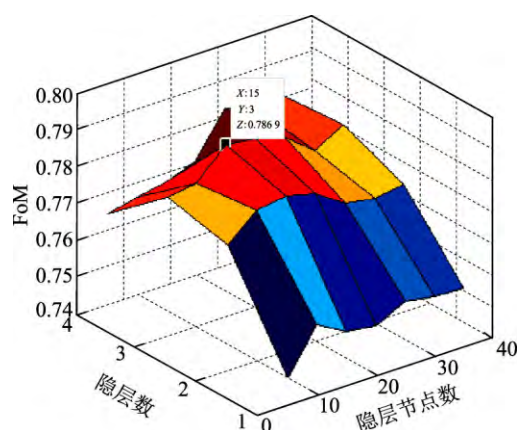


图 6 DBN 结构分析

2) 小样本训练集下不同识别方法结果比较。用传统的 BP 神经网络方法和支持向量机(SVM)方法得到的结果与本文 DBN 方法的结果进行对比。DBN 模型包含 3 个节点数为 15 的隐藏层, 预处理和微调学习率都设为 0.3, 预训练迭代次数设为 50, 微调次数为 5 000。BP 方法采用与 DBN 模型相同的参数, 只是省去了 RBM 预训练的过程。SVM 采

用径向基函数, 参数 g 和惩罚因子 C 采用五折交叉验证的方法进行寻优得到, 对于数据集 1, $C=128$, $g=32$; 对于数据集 2, $C=2\ 048$, $g=4$ 。随机隐退 DBN 模型中, 隐退系数 α 设置为 0.1。

每种方法每幅图随机选取 1 000 样本作为训练, 剩下的 799 000 样本作为测试, 以品质因子 FoM 作为评价标准。重复 10 次, 取平均值作为该算法的识别结果。

从表 1 结果可以看出, 高识别率的算法一般都对应的较高的虚警率, 因此通过这 2 个指标比较难判别一个算法的好坏。因此, 取 3.2 中提到的 FoM 作为评价标准。DBN 方法得到的 FoM 指标相对于 SVM 和 BP 的都更高一些。相比于 NN 的随机初值, DBN 多了一个预训练的过程, 使网络初值一定程度上分配得更合理。每幅图训练样本只有 1 000 个, 为全体数据的 0.13%, 相比经典的 DBN 方法, 随机隐退的 DBN 方法, 有更强的泛化能力, 在小样本识别上有着更高的准确度。

表 1 不同方法识别结果比较

算法	数据集 1			数据集 2		
	识别率	虚警率	FoM	识别率	虚警率	FoM
SVM	0.896	0.138	0.784	0.932	0.066	0.876
BP	0.840	0.101	0.768	0.946	0.085	0.870
DBN	0.907	0.144	0.787	0.952	0.079	0.880
DBN(dropout)	0.933	0.147	0.803	0.976	0.088	0.892

为了更直观地体现识别的结果, 图 7 为数据集 2 某次实验的识别结果图(1 000 训练样本)。其中白色区域为罂粟目标, 黑色区域表示背景。可以看出, 随机隐退的 DBN 方法在可以在各个区域更精准的识别出罂粟作物, 而 SVM 方法在内部会有大

量的点状暗斑(漏检), BP 和 DBD 方法在内部也会出现少量的块状暗斑。同时也可以看出 4 种算法在树木和其他农作物的边缘位置会出现少量的虚警, 毕竟三通道的数据信息量十分有限, 会出现少量的不可分情况, 更多通道的数据支持可以解决这个问题。

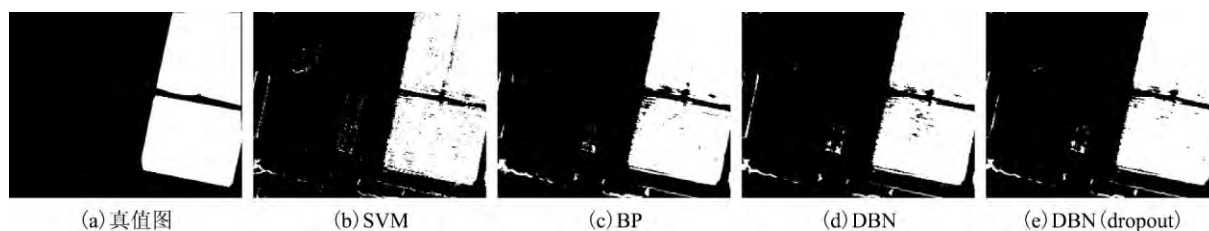


图 7 数据集 2 使用 1 000 训练样本各个算法的罂粟识别结果

3) 不同训练集比例下识别性能比较。为比较训练集大小对识别效果的影响, 数据集 1 的训练比例分别为 0.025、0.1、0.2、0.3、0.4、0.5 等 6 种情况。每种识别方法在相同训练集比例下重复实验 3 次,

取平均值作为实验结果; 其余实验参数的设置与上文相同。

从图 8 可以看出, 4 种方法的 FoM 指标均随着训练集比例的增加而增加, 其中神经网络和深度学

习的方法的增长速度相对于 SVM 方法更为迅速。SVM 方法针对小样本训练集更具有优势。当训练集比例取 0.4 及 0.4 以下的时候,各方法的识别性能为 $\text{DBN}(\text{Dropout}) > \text{DBN} > \text{SVM} > \text{BP}$ 。但当训练集比例上升到 0.5 的时候,识别性能为 $\text{DBN} > \text{DBN}(\text{Dropout}) > \text{SVM} \approx \text{BP}$ 。随着训练集比例的增

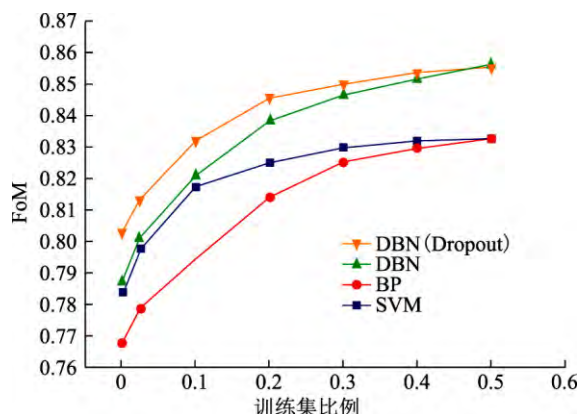


图 8 数据集 1 不同训练集比例下的 FoM 指标

加,随机隐退带来的优势在不断减少。随机隐退的目的是为了防止模型的过拟合,当训练样本足够大并且可以有效表现出罂粟作物特征的时候,一般是不存在过拟合问题的。此时,随机隐退的加入反而减慢网络的收敛速度,难以得到精确的模型。

4 结束语

本文提出了一种基于随机隐退的深度置信网络模型的机载多光谱数据罂粟识别算法。深度置信网络模型自动提取罂粟的光谱和区域特征,发掘其内在规律,提高了识别准确度。在深度置信模型的基础上加入随机隐退,针对少量罂粟样本可以得到泛化能力极强的识别模型。实验表明,基于随机隐退的深度置信网络模型用于多光谱罂粟识别是可行的。

下一步,将尝试使用其他深度学习模型,对比各模型的特点以进一步提高罂粟的识别精度。

参考文献

- [1] 胡训珉. 中国禁毒形势分析评估[J]. 云南警官学院学报, 2010(3): 6-15.
- [2] 罗颖, 张敏, 刘军. 鹰击长空织就“天网”: 无人机及其警务应用现状[J]. 警察技术, 2014(5): 54-62.
- [3] 许洪, 王向军. 多光谱超光谱成像技术在军事上的应用[J]. 红外与激光工程, 2007, 36(1): 13-17.
- [4] 张中山, 燕琴, 余洁, 等. 基于粒子群训练的人工神经网络应用于多光谱遥感影像分类研究[J]. 遥感信息, 2008, 23(5): 3-7.
- [5] SUN K, GENG X, JI L. A new sparsity-based band selection method for target detection of hyperspectral image[J]. Geoscience and Remote Sensing Letters, IEEE, 2015, 12(2): 329-333.
- [6] 李民赞, 韩东海, 王秀. 光谱分析技术及其应用[M]. 北京: 科学出版社, 2006: 112-162.
- [7] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [8] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [9] DENG L, YU D, DAHL G E. Deep belief network for large vocabulary continuous speech recognition; U. S. Patent 8,972, 253[P]. 2015-3-3.
- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Image net classification with deep convolutional neural networks [C]//Advances in Neural Information Processing Systems, 2012: 1097-1105.
- [11] 吕启, 窦勇, 牛新, 等. 基于 DBN 模型的遥感图像分类[J]. 计算机研究与发展, 2014, 51(9): 1911-1918.
- [12] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [13] HINTON G E. Training products of experts by minimizing contrastive divergence[J]. Neural Computation, 2002, 14 (8): 1771-1800.
- [14] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. Computer Science, 2012, 3(4): 212-223.
- [15] 李卫. 深度学习在图像识别中的研究及应用[D]. 武汉: 武汉理工大学, 2014.
- [16] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. The Journal of Machine Learning Research, 2014, 15(1): 1929-1958.
- [17] WANG Z, LI Y, YU F, et al. Object detection capability evaluation for SAR image[C]//International Geoscience and Remote Sensing Symposium 2016 (GARSS 2016), IEEE, 2016: 1548-1551.