

基于神经网络的气象卫星影像森林火点 自动识别的试验研究*

梁益同¹⁾ 胡江林²⁾ 谢萍¹⁾ 刘良明³⁾

¹⁾ (武汉中心气象台, 武汉 430074)

²⁾ (湖北省气象科学研究所, 武汉 430074)

³⁾ (测绘遥感信息工程国家重点实验室, 武汉 430079)

摘要

森林火点的识别是利用气象卫星资料监测森林火灾的基础。传统的目视解译火点识别法难以实现计算机的自动识别,神经网络技术为解决这一问题提供了新的工具。作者探讨了应用神经网络实现气象卫星影像森林火点自动识别的技术方法,并在湖北省地理范围内进行了试验。试验结果显示,经过训练的神经网络能够记忆火点的特征,具备将森林火点从气象卫星影像中识别出来的能力。与目视解译法相比,神经网络方法的精度接近目视解译法,最重要的是实现了森林火点的自动识别。

关键词: 森林火点 自动识别 神经网络 气象卫星影像

引言

森林火灾是一种世界性的重大自然灾害,越来越受到各国政府的重视。卫星遥感技术的出现为森林火灾的及早发现提供了便利的条件。由于时间分辨率高、覆盖范围广、成本低,气象卫星在森林火灾的监测领域里大有用武之地。将森林火灾在气象卫星遥感图像上所反映的信息从其它错综复杂的信息中识别出来,即进行火点识别,是利用气象卫星资料监测森林火灾的基础。传统的火点识别方法主要是目视解译法(即人工识别)。在实践中,这种方法因具有较高的精度而获得广泛的应用,但其有两个不足之处:一是需要有丰富经验的判读人员;二是难以满足计算机自动化的要求,从而影响火点信息分布状态的描述和时效的提高。许多遥感工作者在探索在不降低精度的前提下实现森林火点自动识别的方法,这对于森林火灾及时、准确的发现和迅速的扑灭,具有特别重要的意义,这也是今后森林火灾卫星监测业务的主要方向。美国的 Harris A J L 等^[1]进行了半自动化监测森林火灾的研究。袁飞等^[2]用阈值法来识别火点,但需要人工方法来确定阈值,仍未能真正达到自动识别的目的。覃先林等^[3]采用多因子方法建立小火点自动识别模型,但精度有待进一步提高。

近年来迅速发展起来的人工神经网络理论(简称神经网络)为森林火点的自动识别提供了一条新的途径。由于神经网络是模拟生物神经网络的人工智能技术,具有并行处理、

* 测绘遥感信息工程国家重点实验室开放研究基金资助项目(WKL(99)0102)。

2002-06-24 收到,2003-01-20 收到修改稿。

非线性、容错性和自适应、自学习的特点,因而广泛应用于模式识别、信号处理和自动控制以及遥感图像的分类^[4-5]。到目前为止,已经出现许多种神经网络模型及相应的学习方法,其中反向传播神经网络模型(简称 BP 模型)是应用较广泛的一种。本文探讨了利用 BP 神经网络实现气象卫星影像的森林火点自动识别的基本原理和技术方法,并在湖北省地理范围内进行了试验。

1 BP 神经网络的原理简介

20 世纪 80 年代 Rumelhart 和 Webb 等人建立了 BP(Back Propagation Learning Algorithm)反向学习算法^[6],该算法针对前馈型神经网络,其典型代表是多层感知机。常用的 BP 模型由 3 个神经网络层组成,其最下层为输入层,中间层为隐含层,最上层为输出层,相邻层次之间的神经元之间用连接权系数作相互连接,而各层内的神经元之间没有连接,如图 1 所示。

BP 神经网络的数学模型是:

$$C = f(WB + \theta_1)$$

$$B = f(VA + \theta_2)$$

这里 A, B, C 分别为输入层、隐含层和输出层矢量(节点向量), W 和 θ_1 表示隐含层与输出层之间的连接权和阈值, V 和 θ_2 表示输入层与隐含层之间的连接权和阈值, $f(x)$ 为网络激活函数,这里采用 S 型函数,即

$$f(x) = \frac{1}{1 + e^{-x}}$$

通过对样本的学习和训练,不断调整各节点的连接权和阈值,进而得到样本的输入输出的非线性关系。在 BP 神经网络中,引入代价函数(也称能量函数)

$$E = \frac{1}{2} \sum (C - C_0)^2$$

这里 C_0 是期望输出矢量。BP 神经网络通过代价函数的最小化过程来完成输入到输出的映射。它的基本思想是如果利用已有权重和阈值正向传播得不到期望的输出,则反向传播学习,反复修改(迭代)各节点的权重和阈值,逐步减少代价函数,直到代价函数达到预先设定的要求,一般以代价函数小于某一相当小的正数或迭代时不再减少而反复振荡为止,此时完成 BP 网络的训练,输入与输出映射的确立。

2 火点自动识别模型的建立

2.1 气象卫星监测森林火灾原理简介

目前,用于监测森林火灾的气象卫星主要是极轨气象卫星,如美国的 NOAA 卫星和

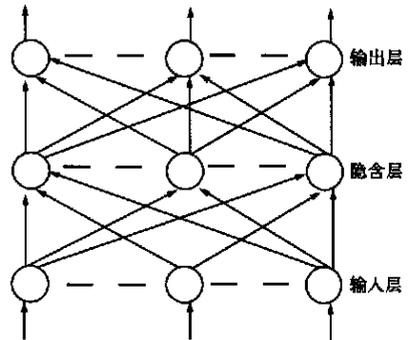


图 1 BP 神经网络的拓扑结构

我国的 FY-1 卫星。以 NOAA-14 卫星为例,其星下点分辨率为 1.1 km,它的甚高分辨辐射仪(AVHRR)有 5 个通道,各通道编号、波长范围及其主要用途如表 1。其中通道 3、4 和 5 是遥感目标温度的,而对地面温度的变化非常敏感的是通道 3,当地面温度高于 800 K (森林火灾一般高于此值)时,通道 3 测到的辐射值远高于通道 4 和 5,人们正是利用通道 3 对森林火灾极为敏感的特性来进行火点识别的。

表 1 NOAA 14 通道编号、波长范围及其主要用途

通道号	波长(μm)	主要用途
1	0.58 ~ 0.68	白天云层、冰、雪、植被
2	0.725 ~ 1.10	白天云层、植被、水体
3	3.55 ~ 3.93	热源、夜间云层
4	10.5 ~ 11.5	洋面温度、白天/夜间云层
5	11.5 ~ 12.5	洋面温度、白天/夜间云层

2.2 数据预处理和样本的提取

本文所用的资料是 1994 ~ 1999 年武汉中心气象台接收的 NOAA 卫星资料,此外还有 1:100 万的湖北省地表类型分布的数据。对卫星资料进行麦卡托投影变换处理,水平和垂直分辨率均为 1.1 km,得到多幅覆盖湖北省范围的区域 NOAA 影像。查看同期的森林火灾遥感监测服务值班记录,选取覆盖湖北省范围,已确认有森林火点的 NOAA 影像 43 幅,将其中 31 幅作为训练集,其余 12 幅作为测试集。

这里定义一个变量叫邻近亮温差,其意义是影像中某像元通道 3 的亮温与周围背景点的平均亮温之差,其表达式为

$$\Delta T = T - \bar{T}$$

式中 T 为该像元第三波段的亮温, \bar{T} 为以该像元为中心一个 5×5 像元窗口通道 3 的平均亮温。

这样,对于影像中的任一像元,我们都可以获得 7 类数据,即 AVHRR 的 5 个通道值、邻近亮温差和地表类型。

为了加快神经网络收敛,需要将 7 类数据进行归一化处理,表达式如下:

$$S = (X - X_{\min}) / (X_{\max} - X_{\min})$$

这里 X_{\max} 和 X_{\min} 分别是某数据的最大值和最小值; X 是该数据的原值; S 是归一化后该数据的新值。

在训练集中,选取已确认森林火灾发生的像元作为火点的训练样本,无森林火灾发生的一些有代表性像元(注:有代表性是指包括各种不同类型的像元,如:云、水体、城区、裸地、耕地、山区、平原等)作为非火点的训练样本。本文选取火点样本 287 个,非火点样本 305 个。

2.3 BP 模型设计

按照 BP 网络设计方法^[6],若 BP 网络中隐含节点可根据需要自由设定,那么三层 BP 神经网络可实现任意连续函数的任意精度逼近。为了简单起见,我们在研究中只使用三层神经网络。我们把能够反映火点特征的信息叫做火点的特征因子,火点的特征因子从上述 7 类数据中确定。输入层神经元数目应与所确定的特征因子数一致。隐含层神经元数目的选择方法目前尚无定论,一般的经验是应略大于输入层神经元个数。从减少误差积累和缩短训练时间考虑,输出层神经元个数应尽量少。火点识别的实质就是将 NOAA

影像的像元划分为火点和非火点两种类型,实际上就是对火点“真假”的判别,因此可以用一个输出端,其理想输出值为 0 或 1,分别表示火点的有或无。BP 模型的学习率以及各节点的连接权和阈值都在 $(-1, +1)$ 区间内随机取值。

2.4 火点特征因子的确定

火点特征因子的确定是影响识别精度的关键。一般地,用于识别的因子越多,识别结果越好。但是,这必须是建立在所选因子确实能反映出一定的火点特征。如果所选因子火点信息量少反而容易产生模糊识别。本文主要目的是在不降低精度的前提下研究森林火点的自动识别,因此所选因子应与识别精度较高的人工识别法所涉及到的因子相同。为了验证这一观点,本文选取了多种方案,相应地构造了多种神经网络进行试验。这里列举 3 个较有代表的模型来分析比较。BP₅ 表示选择 5 类数据(5 个通道值)作为火点特征因子的模型,BP₆ 表示选择 6 类数据(5 个通道值和邻近亮温差)作为火点特征因子的模型,BP₇ 表示选择全部 7 类数据作为火点特征因子的模型。基于上述 BP 模型设计原则,BP₅、BP₆ 和 BP₇ 的拓扑结构分别设计为(5,7,1)、(6,8,1)和(7,11,1)。3 个模型的学习率均取 0.01。

将训练样本的特征因子作为输入,对相应的样本类型(0 或 1)作为输出,对各网络模型进行反复训练,可以发现,随着训练的进行,代价函数稳定减少,当代价函数基本稳定后,训练结束。表 2 列出各模型的训

表 2 各模型的训练次数和稳定后的
代价函数

	BP ₅	BP ₆	BP ₇
训练次数	4960	4200	3380
代价函数	48.5	36.8	17.8

练次数和稳定后的代价函数。表 3 给出了 20 个训练样本的期望输出与各模型稳定后的

表 3 训练样本对各模型的特性输出

样本号	BP ₅ 输出	BP ₆ 输出	BP ₇ 输出	期望输出
1	0.028	0.102	0.018	0
2	0.304	0.234	0.033	0
3	0.188	0.224	0.103	0
4	0.453	0.336	0.097	0
5	0.090	0.159	0.050	0
6	0.010	0.221	0.014	0
7	0.200	0.293	0.306	0
8	0.280	0.382	0.270	0
9	0.430	0.473	0.201	0
10	0.307	0.273	0.309	0
11	0.751	0.601	0.652	1
12	0.231	0.321	0.401	1
13	0.402	0.360	0.883	1
14	0.732	0.940	0.946	1
15	0.369	0.261	0.900	1
16	0.901	0.818	0.902	1
17	0.445	0.916	0.932	1
18	0.732	0.823	0.902	1
19	0.670	0.800	0.781	1
20	0.892	0.832	0.915	1

实际输出。由表 2 可见, BP_7 比 BP_5 和 BP_6 具有更小的代价函数。由表 3 可见, 虽然 BP_5 和 BP_6 对火点也有一定的识别能力, 但它们很容易将一些非火点(如噪声点)判为火点, 只有 BP_7 能更好地记忆火点和非火点的基本特征, 具备了将火点从影像中识别出来的能力。试验还发现, 即使增加各模型的隐含层神经元数目, 试验结果也没太大改变。

由上面试验可见, 选择 7 类数据为火点的特征因子才能充分反映火点的特征。这实际上也和我们人工判别火点的经验相吻合。可以回想一下, 人工判别某像元是否为森林火点, 我们通常至少要进行如下 5 方面分析: (1) 分析该像元通道 3 的亮温; (2) 比较该像元和周围像元通道 3 的亮温; (3) 分析该像元所处的地表类型以及植被状况; (4) 分析 5 个通道值以判断该像元是否为噪声点(由于接收设备原因偶尔产生); (5) 分析通道 1、2 的反射率以及通道 4、5 的亮温以判断该像元是否为云像元。

2.5 测试结果和精度分析

将测试集图像的每个像元输入已训练好的 BP_7 模型, 根据其输出值判断该像元是否火点。至此, 完成了 NOAA 影像森林火点识别的试验过程。图 2 是一幅经过 BP_7 模型识别后的湖北省东南部火点分布图(1999 年 4 月 7 日, NOAA-14)。

在测试集的 12 幅图像中, 总共有 31 457 28 ($512 \times 512 \times 12$) 个像元。为了检验用神经网络法火点识别的精度, 需将目视解译法的识别结果与神经网络方法的识别结果相比, 得到两种方法识别结果的精度检验表, 如表 4 所示。表中括弧里的数值为目视解译法的判别结果。

由表 4 可算出火点和非火点的识别精度(注: 算法为判定正确的火点像元数除以实际火点像元总数)。神经网络的精度为 92.4%, 目视解译法的精度为 93.7%, 两者非常接近。同样可算出神经网络和目视解译法的误判率(注: 算法为判定错误的非火点像元数除以实际非火点像元总数)均在百万分之二以内。

从表 4 还可以看出, 神经网络共将 18 个火点误判为非火点。分析图像并查看值班记录可以发现, 这些火点有的是过火面积小, 有的为薄云覆盖, 火点特征不太明显, 即使是目视解译法也很难分辨出来。事实上表 4 中, 目视解译法将火点判为非火点的个数为 15, 其中大部分包括在这 18 个之内。同样, 有极少数非火点被神经网络误判为火点, 分析图像发现这些点具备一定的火点特征, 即使用目视解译法也极容易误认为是火点。



图 2 经过 BP_7 模型识别后的卫星图像
(1999 年 4 月 7 日, NOAA-14, 图中箭头所指黑点为火点)

表 4 火点识别结果的精度检验

	总数	判定正确个数	判定错误个数
火点	237	219 (222)	18 (15)
非火点	31 454 91	31 454 87 (31 454 85)	4 (6)

注: 括号内数值为目视解译法结果

3 结 论

本文研究表明,神经网络可用来识别气象卫星影像的森林火点信息,并能得到和目视解译法非常接近的精度,更为重要的是,神经网络方法可以实现森林火点的自动识别。因此,当收到新的 NOAA 影像后,只要将各像元所对应的特征因子输入已训练好的神经网络模型,就能快速准确地识别出森林火点,这必然大大地提高森林火灾遥感监测的时效。

参 考 文 献

- 1 Harris A J L. Towards automated fire monitoring from space: semi-automated mapping of the January 1994 New South Wales wildfires using AVHRR data. *Wild Land Fire.*, 1996, 6(3): 107~116.
- 2 袁飞,李亚军.气象卫星遥感火灾自动识别业务系统的研究.见:中国气象局气象服务与气候司、国家卫星气象中心、中国气象科学研究院主编.气象卫星遥感技术为农业服务应用研讨会文集,北京:1996.11.63~68.
- 3 覃先林,易浩若,纪平. AVHRR 数据小火点自动识别方法的研究. *遥感技术与应用*, 2000, 15(1): 36~40.
- 4 Bernard A C, Wilkinson G G, Kanellopoulos I. Training strategies for neural network soft classification of remotely-sensed imagery. *INT. J. Remote Sensing*, 1997, 18(8): 1851~1856.
- 5 梁益同,胡江林. NOAA 卫星图像神经网络分类方法的探讨. *武汉测绘科技大学学报*, 2000, 25(2): 148~151.
- 6 李学桥,马莉.神经网络·工程应用.重庆:重庆大学出版社,1996.37~44.

AUTO-IDENTIFYING FOREST FIRE POINTS IN METEOROLOGICAL SATELLITE IMAGES BASED ON NEURAL NETWORK

Liang Yitong¹⁾ Hu Jianglin²⁾ Xie Ping¹⁾ Liu Liangming³⁾

¹⁾ (Wuhan Central Meteorological Service, Wuhan 430074)

²⁾ (Hubei Research Institute of Meteorology, Wuhan 430074)

³⁾ (National Lab for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan 430079)

Abstract

Identification of forest fire-points in meteorological satellite images is the basis of monitoring forest fire using meteorological satellite data. The traditional approach of identification is visual interpretation. In the visual interpretation, better precision is obtained in practice, but experienced interpreter is needed and it is difficult to settle for auto-identification with computer, so that there is always unfavorable in the description of information distribution and the improvement of efficiency. In order to find out the forest fire in time and eradicate it quickly, it is very significant to make the fire-points auto-identified without reducing the precision. The artificial neural network theory(short for NN) developing rapidly in recent

years provides a new means for solving this problem. NN has the basal characters of human brains, such as learning, recollecting and generalizing. The peculiarity of NN is massive parallel computing, distributive memory of information, nonlinear dynamics of consecutive time, global behaviour, great fault-tolerance and robust, self-organization, self-learning and real time processing. BP (Back Propagation Learning Algorithm) model is a NN used widely. The method using BP model to automatically identify fire-points in meteorological satellite images are discussed and the test in the range of Hubei province is presented. Besides the NOAA/AVHRR data from 1994 to 1999, the earth's surface distribution is used. 43 meteorological satellite images with fire-points are selected, and 31 of them are regarded as trained collection, the others as tested collection. The result of the test shows that a fire-point is correlative with 7 character factors, which are the radiation-values of the 5 channels of NOAA/AVHRR, the earth's surface distribution and the difference of the contiguous temperature. The disciplined NN has recollected the characters of fire-points and non-fire-points and has ability to identify fire-points in images. Comparing NN with visual interpretation, the conclusion is drawn that NN can auto-identify fire-points in meteorological satellite images with the almost same precision.

Key words: Forest fire-point Auto-identification Neural network Meteorological satellite images