

基于遗传算法优化的 BP 神经网络遥感数据土地覆盖分类

骆成凤^{1,2}, 刘正军¹, 王长耀², 牛 铮²

(1. 中国测绘科学研究院摄影测量与遥感所, 北京 100039; 2. 中国科学院遥感应用研究所遥感科学国家重点实验室, 北京 100101)

摘 要: 提出了一种基于遗传算法优化的 BP 神经网络遥感分类方法。该方法兼顾了遗传算法和梯度下降优化算法分别在 全局和局部搜索极小点的优势; 避免了在 BP 网络训练过程中过早收敛于局部极小点的风险; 与 BP 算法相比, 该算法多次 重复过程所得网络的均方差比较稳定。在算法验证中, 用中巴地球资源一号卫星数据作为试验数据, 详细描述了网络优化 过程中的参数设置和关键参数变化过程, 比较了该算法与 BP 算法、最大似然法的分类精度。分类试验表明: 该算法不但有 较高的执行效率, 也能达到很高的分类精度。

关键词: BP 算法; 遗传算法; 人工神经网络; 土地覆盖分类; 遥感数据

中图分类号: S127; TP183; TP274

文献标识码: A

文章编号: 1002-6819(2006)12-0133-05

骆成凤, 刘正军, 王长耀, 等. 基于遗传算法优化的 BP 神经网络遥感数据土地覆盖分类 [J]. 农业工程学报, 2006, 22(12): 133-137.

Luo Chengfeng, Liu Zhengjun, Wang Changyao, et al. Optimized BP neural network classifier based on genetic algorithm for land cover classification using remotely-sensed data [J]. Transactions of the CSAE, 2006, 22(12): 133-137. (in Chinese with English abstract)

0 引 言

在过去的近 20 年中, 人工神经网络 (Artificial Neural Network, ANN) 算法已经被广泛应用到遥感信息提取过程中。与基于统计的分类算法不同, ANN 分类算法依赖的是数据本身, 不需要数据中的统计分布假设。对遥感数据而言, 由于地物波谱辐射的复杂性以及获取过程中干扰因素的多样性等原因, 获取的地物信息具有模糊性、不确定性等特征, 而且数据分布也复杂, 蕴涵在遥感数据中的特征很难用常规的统计方法来获取。许多研究者用 ANN 分类器对不同尺度不同数据源的遥感图像进行了分类研究^[1-4], 几乎所有的研究案例均证明 ANN 分类器优于传统分类器。

目前在 ANN 的实际应用中, 绝大部分的 ANN 模型是采用反向传播神经网络 (Back Propagation, BP) 和它的变化形式。BP 算法使用了优化中最普通的梯度下降法, 经过多年试验研究, 现通常采用 Sigmoid 函数作为激发函数^[5]。梯度下降优化算法和 Sigmoid 函数自身的特点可能导致 BP 网络最终收敛到局部极小点, 而不是全局最优极小点。而且 BP 算法多次训练结果不稳定, 网络的执行和收敛速率低。

本文提出了基于 GA 优化的 BP 神经网络遥感数据分类算法, 以弥补 BP 算法在遥感影像分类过程中的不足。文章先对基于 GA 优化的 BP 算法进行了算法描述; 然后以 CBERS-1 卫星数据为信息源, 对该算法进行了试验验证, 并与 BP 分类算法和最大似然法进行了

分类精度比较。试验证明该算法不但有较高的运行效率, 而且能得到很高的分类精度。

1 研究方法

1.1 反向传播神经网络算法

反向传播神经网络算法是由 Rumelhart 等在 1985 年提出的, 即 EBP (Error Back Propagation) 算法^[5], 简称 BP 算法。BP 网络的学习过程由正向和反向两部分组成, 输入信息经过正向传播得到相应的输出, 输出结果与目标输出比较, 如果误差超出预定值, 则转入反向传播, 将误差信号由输出层反馈给隐含层和输入层, 根据误差调整各层节点 (神经元) 间的连接权重, 反复该过程直到信号误差达到允许的误差范围。BP 算法使用了优化中最普通的梯度下降法, 经过多年实验研究, 现通常采用 Sigmoid 函数作为激发函数^[5]。

Sigmoid 函数 $f(x)$:

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

其中 $x \in (-\infty, +\infty)$ 。

网络训练过程中各节点权值和阈值调整函数为:

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji} + T[w_{ji}(t) - w_{ji}(t-1)] \quad (2)$$

$$\theta_j(t+1) = \theta_j(t) + \Delta \theta_j + T[\theta_j(t) - \theta_j(t-1)] \quad (3)$$

式中 w_{ji} ——输入节点和输出节点间的连接权重; θ_j ——节点的阈值; t ——训练迭代次数, 在实际的学习过程中, Z 是按梯度搜索的步长, 对学习过程的影响很大, 步长越大, 网络权重变化越剧烈。实际应用中, 通常以不导致振荡为前提提取尽量大的值, T 是一个常数, 它决定过去网络权重变化对目前网络权重值变化的影响程度。

训练过程中整个网络的误差函数用均方差 (mean square error, MSE) 表示, 具体公式如下:

收稿日期: 2005-03-24 修订日期: 2006-01-18

项目来源: 科技部科研院所社会公益研究专项 (2004DIB3J107); 中国科学院知识创新工程重大项目 (KZCX1-SW-01-02)

作者简介: 骆成凤 (1976-), 女, 陕西人, 博士, 主要从事土地覆盖 / 土地利用分类以及变化监测研究。北京 中国测绘科学研究院摄影测量与遥感所, 100039 Email: chf_luo@163.com

$$X_{net} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^k (T_i^j - f_i^j)^2 \tag{4}$$

式中 X — 均方差; net — 神经网络; f_i^j — 输入向量在输出层节点的实际输出; T_i^j — 期望输出值。

BP算法中的梯度下降优化算法和 Sigmoid函数自身的特点可能导致网络最终收敛到局部极小点。梯度下降依赖误差面的形状、网络连接权重的初始值,以及一些复杂参数。一般误差面上存在多个局部极小值,当沿着误差面移动时它就大多会收敛到一些局部极小点,而不是收敛到全局极小点,这使网络不能很好地收敛到期望值。Sigmoid函数将初始无限输入域压缩到有限的输出域,通常也叫“挤压函数”(squashing function),它有一个特征,当输入变得很大时,其误差面斜率趋向零。因此,当用快速梯度下降算法来训练具有 Sigmoid函数的多层网络时就出现一个问题:当梯度下降一个很小的量级时,权重和偏差也随之发生很小的变化,使权重和偏差值偏离它们的最优值。受网络随机初始化的影响,BP算法多次训练所得结果不稳定,导致分类结果有差异。再者,微分计算的低效率大大降低了网络的执行和收敛速率。

1.2 基于 GA优化的 BP算法

遗传算法 (Genetic Algorithm, GA)是一种根据生物学中所谓自然选择和遗传机理随机搜索的优化算法,其本质特征在于群体策略和简单的遗传算子^[6]。它突破邻域搜索的限制实现整个解空间分布上的信息搜索、采集和继承。GA通过将待寻优参数空间进行编码,不依赖梯度信息,而用随机选择作为工具来引导搜索过程向着更高效的方向发展。尤其适合于处理传统搜索方法难于解决的复杂和非线性问题。

用 GA进行全局最优 BP网络权重搜索包括 3个主要步骤: A)建立 GA染色体的位串结构与 BP网络权重之间的关系。 B)确定适配函数,它是评价染色体对网络权重适配程度的基准。 C)确定遗传策略,包括选择群体大小,选择、交叉、变异方法等。典型的 BP网络权重进化过程包括:

- 1) 将当前每个染色体个体解码成 BP网络权重集。本文采用实数编码方案,直接将每个节点的连接权重和偏差依次与相应的基因段关联。
- 2) 用整个网络的误差函数作为适配函数,通过解码构建相应的神经网络结构,然后计算实际输出与目标输出总的均方差来评估网络权重集。个体的适配性取决于网络总的均方差,均方差越高,适配性越低。
- 3) 基于适配值来选择可被复制到下一代的父本。这里应用轮盘赌的选择方法^[7,8]来选择父本。这种选择方法将当前的群体影射到轮盘上,并且认为每个染色体在轮盘上的空间与其适配值成正比。
- 4) 应用交叉、变异等搜索算子从父本产生子代。为了增加进化的多样性,本文用了单点随机变异算子、多点随机变异算子和三种交叉算子。

整个算法流程如图 1所示

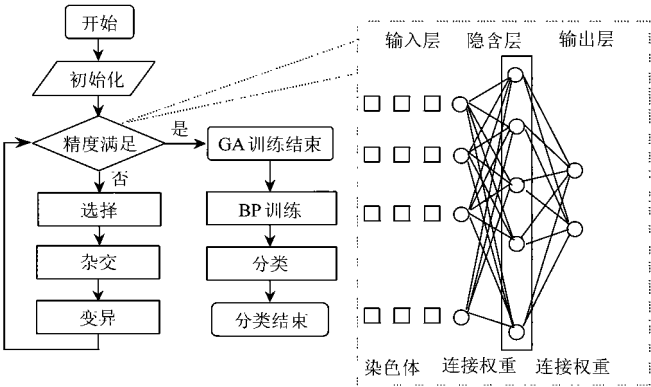


图 1 基于 GA优化的 BP算法流程图
Fig. 1 Flow chart of optimized BP algorithm based on genetic algorithm

算法先利用 GA全局寻优的优势,在迭代搜索早期用各种遗传算子在整个解空间进行查找,有效地搜索全局近最优网络连接权重,这样既不用计算梯度信息,也不用初始化网络连接权重,可大大提高搜索执行效率,但有时得不到最优的解。研究的目的是找出最优的网络连接权重。因此在全局近最优解的基础上,在迭代搜索后期继续用梯度下降算法训练网络,在全局近最优解附近区域进行局部协调,直到找到局部最优网络权重集。

BP与 GA结合后的算法可以解决 BP局部收敛问题。同时只在局部寻优过程中使用梯度下降法,大大提高了整个训练过程中的执行和收敛速度。经过 GA全局寻优后,BP局部寻优的初始化网络比较稳定,这样多次训练所得网络结果和分类结果都比较稳定。

2 分类算法验证

2.1 试验数据

用中巴地球资源一号卫星 (CBERS-1)数据作为试验数据,选取 4个波段作为输入波段,波长范围分别为 0.45~ 0.52 0.52~ 0.59 0.63~ 0.69和 0.77~ 0.89 μm ,图像空间分辨率为 19.5 m,用于试验的图像大小为 800 \times 1000像素。选择 240个样本,分属于 6种土地利用类型:小麦、棉花、盐碱地、裸地、城镇用地、水体。在分类之前,引入 Jeffries-Matusita距离 (J-M距离)^[9]来度量类别间的统计可分性。该度量值范围在 0~ 2之间,一般情况下,该值大于 1.8就表明类别间分离性较好。由 J-M距离表示的 6类样本间分离度见表 1。从表 1中可以看出,除了裸地和棉花之间的分离度为 1.841外,其它各类之间的分离度都在 1.98以上,说明所选取的样本点有较强的代表性。

表 1 各类别样本之间的分离度

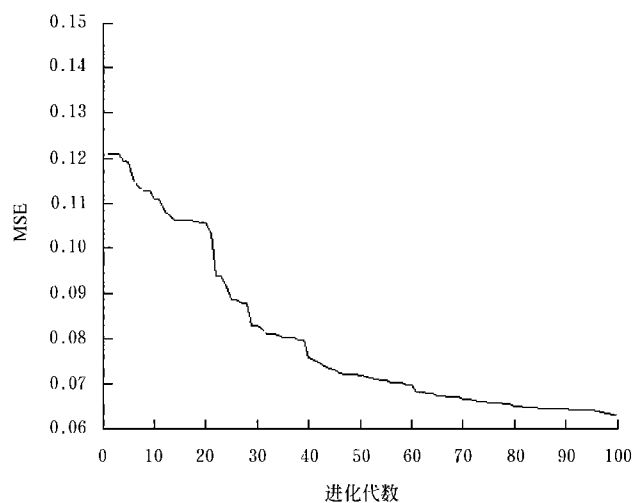
Table 1 Separability of samples of different classifications					
	小麦	棉花	盐碱地	裸地	城镇用地
棉花	1.982				
盐碱地	1.999	1.997			
裸地	1.998	1.841	1.999		
城镇用地	1.999	1.997	1.998	1.986	
水体	1.999	1.999	2.000	1.981	1.999

根据先验知识,对各波段输入数据进行归一化,使所有的输入数据在 0~1 之间。最后分类时同样先对整个影像数据进行归一化。

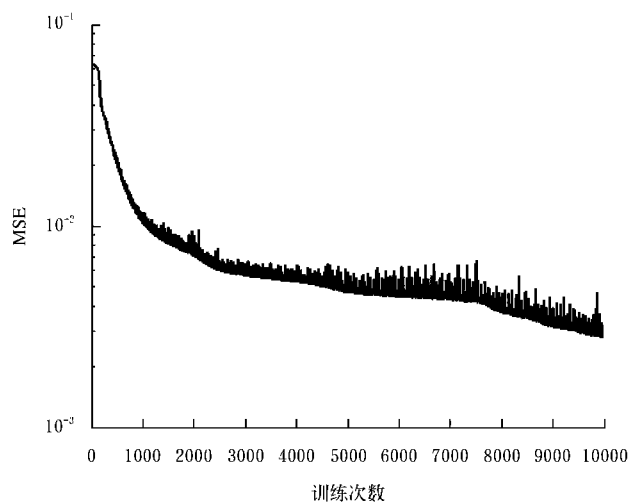
2.2 参数设置

设置三层 BP 网络结构,参照 Paola 和 Schowengert 关于中间层节点设置的论述^[10],将中间层节点数设置为 8,输入层节点为 4(对应 4 个波段),输出层节点数为 6(对应 6 个类别)。将 BP 网络的权重和阈值按照一定顺序排列成染色体,每个染色体基因数目为 94,基因值范围为 $[-5, 5]$ 。对于遗传过程,初始种群大小为 160,交叉率 0.09,变异率为 0.11,三种交叉算子各占 0.03,多点变异占变异数目 1/3,单点变异共占变异数目的 2/3。遗传迭代 100 次后,继续进行 BP 传播算法,学习步长为 0.01,步长增长为 1.05,训练目标误

差为 0.0001。图 2a 是对应遗传进化过程中每一代的最佳网络均方差,从图中可以看出,进化了 40 代后,最佳网络均方差的变化因局部调整而变慢,70 代后最佳网络均方差的变化非常小,虽然此时最小误差达到 0.063,但是用适配性最好的染色体解码后构建 BP 网络,对所有训练样本进行网络仿真,其实际输出与目标输出还是相差非常大;图 2b 是对应 BP 网络训练过程每次循环的网络均方差,训练次数为 10000,尽管在训练过程中,网络训练函数均方差 (Mean squared error, 简称 MSN) 出现动荡反复,但在训练全过程中, MSE 总趋势还是呈明显降低。由图 2 可以看出,尽管遗传进化已经大大降低了网络总的均方差 (0.121~0.063),但运用 BP 网络训练过程进行权重调整仍可改进网络,网络总的均方差达到 0.0031,提高近 0.06。



a. 进化训练的最佳 MSE



b. BP 网络训练的 MSE

图 2 网络训练均方差 (MSE) 在网络训练中的变化过程

Fig. 2 Changing process of mean squared error (MSE) in the net training

2.3 分类结果及分析

用于进行分类精度比较的有 3 种监督分类算法: 最大似然法 (Maximum likelihood classifier, MLC), BP 网络分类法, 基于遗传优化算法的 BP 网络分类法。其中, MLC 是遥感影像监督分类经典的算法之一, 它通过统计方法计算各个像元属于每个类的概率, 将该像元归属于概率最大的类。BP 网络是一种多层感知器 (Multi-layer perception) 模型, 在此为了表述方便, 用 BP-MLP 代表 BP 网络分类器, 用 GA-MLP 表示基于 GA 优化的 BP 网络分类器。其中 BP-MLP 分类过程中训练参数设置与 GA-MLP 过程中 BP 网络训练参数相同。

对 3 个分类器使用同样的训练样本, 3 种方法的分类结果分别见图 3b、3c、3d 所示。这 3 种不同的分类算法对水体、小麦、裸地 3 种类别的分类差异比较大。如图上右上角部分, MLC 分类结果的水体明显多于其它两个分类结果, BP-MLP 分类结果将其分为以裸地为主的类别。对比原始彩色合成图像, 水体呈黑色调, 但稍暗的非黑色调区域不是水体, MLC 将其误分为水体; 另外

小麦呈红色调和浅红色调, 暗红色调为裸地, BP-MLP 将部分小麦误分为裸地, 分类结果夸大了裸地所占比例。3 种不同的分类算法分类结果中各个类别所占面积百分比比较见图 4。

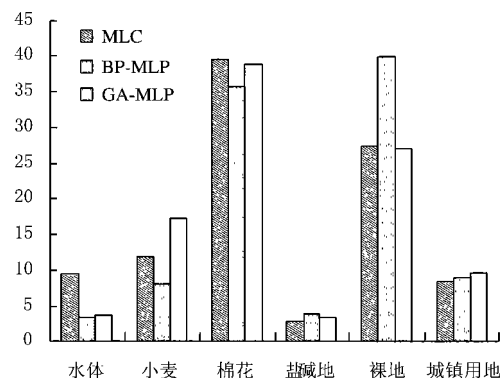


图 4 各种分类算法分类结果对比

Fig. 4 Comparison of different classification results using different classifiers

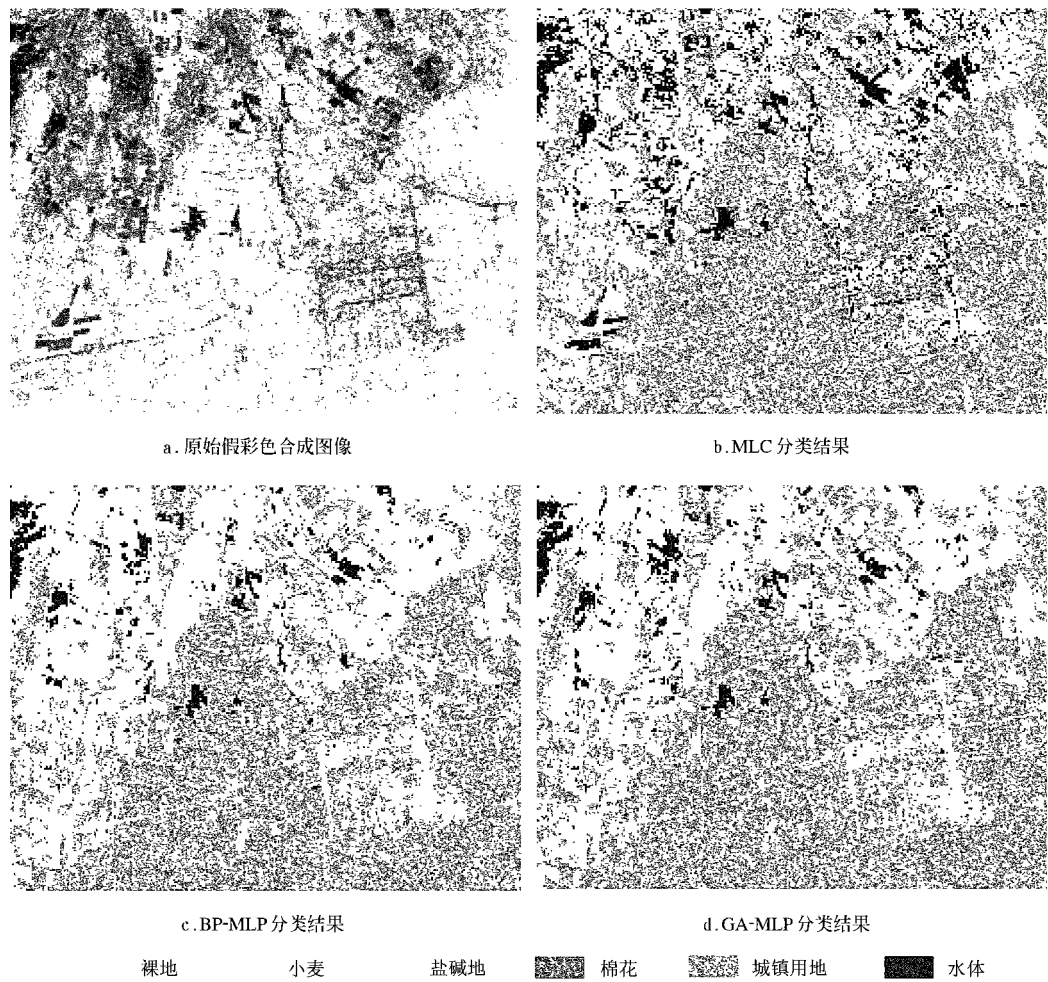


图 3 试验区图像及不同分类法对相同样本的分类结果

Fig. 3 Images and results of land cover classifications using different classifiers experimental areas

随机抽取 400个样点来评估分类精度,对单个类别精度、总精度、总 Kappa系数^[11]进行比较,如表 2所示

表 2 各分类算法精度对比

Table 2 Comparison of classification accuracy of each algorithms

土地覆盖类型	MLC	BP-MLP	GA-MLP
水体	85.71%	91.43%	98.57%
小麦	87.14%	85.71%	100.00%
棉花	88.89%	95.81%	95.83%
盐碱地	88.57%	94.29%	97.14%
裸地	87.14%	98.57%	94.29%
城镇用地	87.14%	92.86%	98.57%
总精度	87.44%	93.13%	97.39%
Kappa系数	0.849	0.918	0.968

从表 2可以看出,ANN分类算法的单类分类精度、总精度、Kappa系数几乎全部高于基于统计的最大似然法。MLC、BP-MLP、GA-MLP 的总精度分别为: 87.44%、93.13%、97.39%,GA-MLP的总精度比 BP-MLP提高 4.26%,这是因为 GA-MLP训练得到了比 BP-MLP训练更低的网络 MSE

3 结论与讨论

- 1) 本文提出了基于 GA优化的 BP算法,它兼顾了 GA在全局搜索和 BP在局部搜索的优势。BP算法在搜索最优解时,需要微分计算,在整个解空间求解时会大大降低执行和收敛速率,而且可能过早收敛于局部最小;一般的 BP网络受随机初始化网络的影响,多次训练所得结果网络权重不稳定,据估计初始化网络的随机性对最终网络 MSE的影响达 3% 或者更高^[12],而用经过 GA优化所得全局近最优解作为 BP继续训练的初始值,不但能提高算法执行效率,避免网络过早收敛于局部最小值,而且得到的最终网络权重相对比较稳定。
- 2) 用 CBERS-1数据(1.5 m分辨率)进行的试验表明:基于 GA优化的 BP网络分类器,执行效率高于基于梯度下降的 BP分类器;在训练过程中能得到更低的网络均方差,其分类精度达到 97%,比 BP分类器高 4%,与传统的最大似然分类法相比,总分类精度提高近 10%。试验证明,文中提出的基于 GA优化的 BP网络分类器用于中高分辨率的遥感影像分类,能得到比较高的分类精度。
- 3) 用基于 GA优化的 BP网络对遥感数据进行分

类时,怎样决定最优神经网络拓扑结构是个问题,例如对不同分辨率、不同数据源、不同样本数目,怎样设计才能使得网络结构达到最佳。本文试验中 BP网络结构的设计是在借鉴 Paola 等关于中间层节点设置方法的基础上,经过多次前期试验设置的。一个可能的替代方法是应用 GA来优化神经网络结构使其达到最佳值

[参 考 文 献]

- [1] Benediktsson J A, Swain P H, Ersoy O K. Neural network approaches versus statistical methods in classification of multisource remote sensing data [J]. IEEE Transactions on Geoscience and Remote Sensing, 1990, 28 (4): 540– 552.
- [2] Hara Y, Atkins R G, Yueh S H, et al. Application of neural networks to radar image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 1994, 32 100– 111.
- [3] Tzeng Y C, Chen K S, Kao W L, et al. A dynamic learning neural network for remote sensing applications [J]. IEEE Transactions on Geoscience and Remote Sensing, 1994, 32 1096– 1103.
- [4] Benediktsson J A, Swain P H, Ersoy O K. Conjugate-gradient neural networks in classification of multisource and very-high-dimensional remote sensing data [J]. International Journal of Remote Sensing, 1993, 14 2883– 2903.
- [5] 杨建刚. 人工神经网络实用教程 [M]. 杭州: 浙江大学出版社, 2001.
- [6] 李敏强, 寇纪松, 等. 遗传算法的基本理论有应用 [M]. 北京: 科学出版社, 2002.
- [7] Goldberg D E. Genetic algorithms in search, optimization, and machine learning [M]. Addison-Wesley New York, 1989.
- [8] Holland J H. Adaptation in natural and artificial systems [M]. The University of Michigan Press, 1975.
- [9] 赵英时, 等. 遥感应用分析原理与方法 [M]. 北京: 科学出版社, 2003.
- [10] Paola J D, Schowengerdt R A. A review and analysis of backpropagation neural networks for classification of remote-sensed multi-spectral imagery [J]. International Journal of Remote Sensing, 1995, 16(16): 3033– 3058.
- [11] Fitzgerald R W, Lee B G. Assessing the classification Accuracy of multisource remote sensing data [J]. Remote Sensing Environment, 1994, 47 362– 368.
- [12] Paola J D, Schowengerdt R A. The effect of neural network structure on a multispectral land-use/land-cover classification [J]. Photogrammetric Engineering and Remote Sensing, 1997, 5 535– 544.

Optimized BP neural network classifier based on genetic algorithm for land cover classification using remotely-sensed data

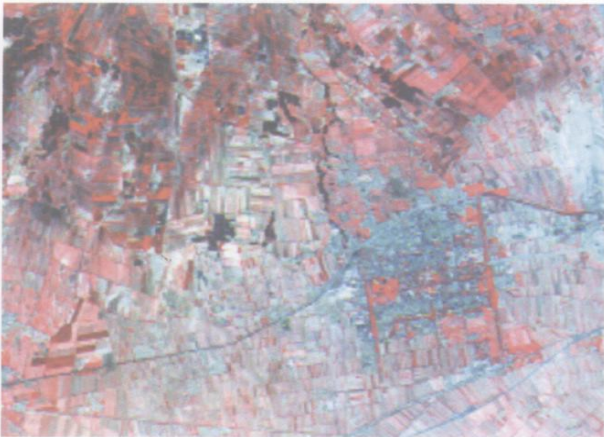
Luo Chengfeng^{1,2}, Liu Zhengjun², Wang Changyao², Niu Zheng¹

(1. Institute of Photogrammetry and Remote Sensing, Chinese Academy of Surveying and Mapping, Beijing 100039, China; 2. State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing 100101, China)

Abstract A new Error Back Propagation algorithm based Genetic Algorithm was proposed in the article, and the key steps and framework were described in detail. The algorithm gives attention to two optimization algorithms, genetic algorithm and back propagation algorithm, which have the advantage during searching the infinitesimal point in local space and global space respectively, avoiding the risk of premature convergence while BP network was training; compared with BP algorithm, the end total mean square error of the network is more stable even if people redo the whole course several times. The data from China-Brazil Earth Resources Satellite were used to validate the algorithm; meanwhile the setting parameters and change processing of parameter were depicted carefully. Maximum likelihood classifier, back propagation neural network classifier were involved for a comparison purpose. The experiment results show that the new algorithm cannot only run with better efficiency, but also achieve the best classification accuracy.

Key words BP algorithm; genetic algorithm; Artificial Neural Network; land cover classification; remotely-sensed data

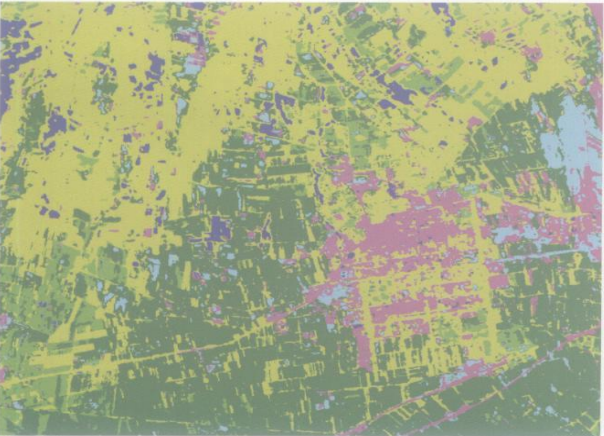
骆成凤, 等. 基于遗传算法优化的BP神经网络遥感数据土地覆盖分类 (图4)
Luo Chengfeng, et al. Optimized BP neural network classifier based on genetic algorithm for laand cover classification using remotely-sensed data (Figure 4)



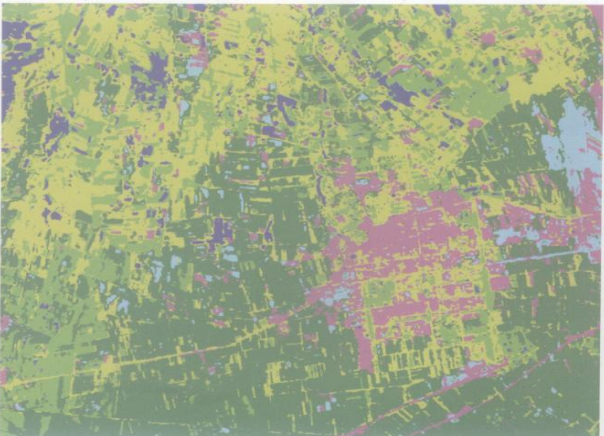
a. 原始假彩色合成图像



b. MLC分类结果



c. BP-MLP分类结果



d. GA-MLP分类结果

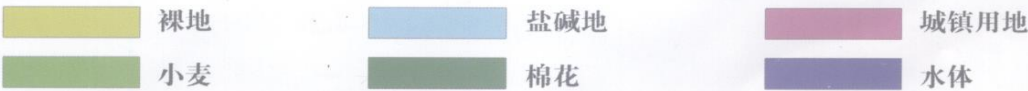


图 4 试验区图像及不同分类法对同样本的分类结果
Fig.4 Images and results of land cover classification using different classifiers in experimental areas