

## 基于近似域划分的可变离散精度粗逻辑网络及其遥感图像分类应用

张东波<sup>①②</sup> 王耀南<sup>②</sup>

<sup>①</sup>(湘潭大学信息工程学院 湘潭 411105)

<sup>②</sup>(湖南大学电气与信息工程学院 长沙 410082)

**摘要:** 为解决粗逻辑神经网络精度与网络规模复杂性和推广泛化能力之间的矛盾, 该文提出了一种具有可变离散精度的粗逻辑神经网络设计方法。该方法通过近似域划分, 将论域空间划分为确定性区域和可能性区域, 由于可能性区域信息粒度过大是造成误分类的重要原因, 只需对可能性区域离散区间进一步细化, 即可达到提高粗逻辑网络的精度, 同时抑制网络规模增长过快的目的。在长白山地区的遥感图像分类实验中, 常规方法在离散等级为 7 时有最好性能, 而该文方法以较小的网络代价和训练时间获得了逼近的分类结果。

**关键词:** 遥感图像分类; 粗糙集; 粗逻辑网络

中图分类号: TP751

文献标识码: A

文章编号: 1009-5896(2007)11-2720-05

## Variable Discretization Precision Rough Logic Neural Network Based on Approximation Area Partition and Its Application to Remote Sensing Image Classification

Zhang Dong-bo<sup>①②</sup> Wang Yao-nan<sup>②</sup>

<sup>①</sup>(*Institute of Information Engineering, Xiangtan University, Xiangtan 411105, China*)

<sup>②</sup>(*College of Electrical and Information Engineering, Hunan University, Changsha 410082, China*)

**Abstract:** A variable discretization precision rough logic neural network is proposed to solve contradiction between network precision and the size of network as well as generalization ability. Based on the approximation area partition, the universe discussed can be partitioned into certain area and possibility area. The important reason of misclassification is the granularity of the possibility area is too coarse. In this work, only possibility area is refined and the precision of the rough logic neural network is improved while the size of network is restrained. In the experiment of the remote sensing image classification about Changbai mountain area, the performance of conventional method is best when the discretization level is 7. The most approximated result is acquired, while less network cost and training time are expended, when this method is used.

**Key words:** Remote sensing image classification; Rough set; Rough logic neural network

### 1 引言

粗糙集理论<sup>[1-3]</sup>是一种对不确定、不完整、不精确知识进行处理的数学理论, 由于具备强大的定性分析能力, 目前已成为除模糊集、神经网络之外的一种新的不确定性计算工具。神经网络作为不确定性计算工具, 具有较强的抗干扰能力、推广泛化能力和自学习能力, 但在输入空间维数过高时, 网络结构比较复杂, 含有较多的结构冗余, 而且神经网络的内在机理和决策过程不透明, 而粗糙集具有强大的定性分析和知识约简能力, 决策过程具有语义解释性, 但粗糙集本身抗噪声性能差, 推广能力弱。正因为粗糙集和神经网络具有很强的互补关系, 将两者进行集成已成为一个引人注目的研究领域。近十几年来, 多种粗集神经网络集成技术<sup>[4]</sup>得到了

研究, 其中, 基于粗逻辑决策规则的“强耦合”集成方式的粗逻辑神经网络更是得到了广泛的应用<sup>[5-8]</sup>。建立在通过粗糙集数据分析所获得的输入输出映射关系基础上的粗逻辑神经网络具有结构精简、冗余较少、学习速度快、决策过程透明、具备语义解释性的优势。文献[9]中提到了基于“与”、“或”逻辑元的粗逻辑网络, 该网络完全依据粗逻辑决策规则进行构造, 隐含层分别由进行“与”、“或”逻辑运算的逻辑元组成, 同时连接权值仅表示联接关系而不进行学习训练, 因此, 其决策过程是一种简单的规则匹配过程, 没有冲突解决机制, 也不具备推广、容错能力。鉴于此, 文献[9]中提出将“与”、“或”逻辑元分别用神经元取代, 同时相关权值和偏置阈值可以进行学习, 从而使网络具有自学习能力和推广、容错能力。文献[6, 9]中的粗逻辑网络第 2 层作为离散化层, 由相应离散量化函数进行区间划分, 由于实行的是简单的边界划分, 当输入变量为连续数据时, 不利于输入数据中模糊信息的描述, 因此, 有的文献中<sup>[5, 8]</sup>, 将离散化层

2006-05-08 收到, 2006-10-08 改回

国家自然科学基金(60375001), 高等学校博士点基金(20030532004)和湖南省教育厅科研项目(05C093)资助课题

改造为模糊量化层, 通过隶属函数进行输入变量的模糊描述, 而且隶属函数的参数(如高斯型隶属函数的中心、标准差参数)可以通过学习进行调整, 使其更符合模式分类的动态自适应过程。由于粗逻辑网络的精度在很大程度上取决于离散化时的不可分辨区间划分的精细程度, 一般来讲, 划分的知识粒度越小, 精度越高, 但相应网络的规模复杂程度、训练时间和计算量也越大, 同时可能造成网络的推广能力减弱, 极端情况下, 如果一个样本代表一个模式类时, 网络仅具有记忆能力而丧失了推广泛化能力。因此如何平衡网络复杂性、推广泛化能力和精度之间的矛盾是一个亟待解决的问题, 在以往的文献中, 都还未进行过深入的研究和提出有效的解决方法。本文通过粗糙集理论中近似域的思想将输入空间论域划分为确定性区域和可能性区域, 对确定性区域和可能性区域分别采用不同精度等级的离散量化函数, 从而在一定程度上为解决上述问题提供了一种新的思路。本文方法在长白山地区遥感图像地物分类的实验中, 取得了良好的应用效果。

## 2 粗糙集基本概念

一个信息系统可以用四元组来表示  $S = (U, A, V, f)$ , 其中  $U$  表示非空有限对象集, 即论域;  $A$  为所有属性的集合;  $V = \bigcup_{a \in A} V_a$ ,  $V_a$  是属性  $a$  的值域;  $f: U \times A \rightarrow V$  是一个信息函数, 为每一个对象的属性赋予属性值。信息系统通常可以简写为  $S = (U, A)$ 。

粗糙集从模式分类的观点来描述知识, 通过不可分辨的等价关系, 将知识空间划分为不同的模式等价类, 从而将知识空间表示为粒状结构。正是因为知识的这种粒状结构, 造成知识表示的粗糙性。令子集  $X \subseteq U$ ,  $R$  是对论域  $U$  进行划分的一个等价关系, 则一个划分  $U/R$  定义为:

**定义** 划分  $U/R = \{X_1, X_2, \dots, X_n\}$ ;  $X_i \subseteq U$ ,  $X_i \neq \phi$ ,  $X_i \cap X_j = \phi$ , 对于  $i \neq j$ ,  $i, j = 1, 2, \dots, n$ ;  $\bigcup_{i=1}^n X_i = U$ , 可以证明, 在  $R$  已知的情况下, 这种划分具有唯一性。

$U/R$  表示基于  $R$  对  $U$  划分的所有等价类集合,  $\forall x, y \in X_i (i = 1, \dots, n)$ , 有  $(x, y) \in R$ , 记作  $xRy$ , 其中包含元素  $x$  的  $R$  等价类, 用  $[x]_R$  来表示。如果  $X$  能表达成  $U/R$  中某些等价类的并时, 则称集合  $X$  是  $R$  可定义集或  $R$  精确集, 否则集合  $X$  是  $R$  不可定义集或  $R$  粗糙集。粗糙集  $X$  可以用  $R$  上、下近似集  $\langle \bar{R}(X), \underline{R}(X) \rangle$  来近似描述:

$$\underline{R}(X) = \bigcup \{ [x]_R \mid [x]_R \subseteq X, x \in U \} \quad (1a)$$

$$\bar{R}(X) = \bigcup \{ [x]_R \mid [x]_R \cap X \neq \phi, x \in U \} \quad (1b)$$

$\text{BN}_R(X) = \bar{R}(X) - \underline{R}(X)$  称为  $X$  的  $R$  边界域,  $\text{POS}_R(X) = \underline{R}(X)$  称为  $X$  的  $R$  正域,  $\text{NEG}_R(X) = U - \bar{R}(X)$  称为  $X$  的  $R$  负域。显然:  $\bar{R}(X) = \text{POS}_R(X) \cup \text{BN}_R(X)$ 。

当一个信息系统的属性集  $A$  由条件属性  $C$  与决策属性  $D$  组成时, 此时的信息系统也称之为决策系统, 记作

$S = (U, C \cup D)$ 。如果决策系统中存在条件相同而决策不同的对象, 则决策系统称为不相容的或不一致的, 否则, 称为相容的或一致的。条件属性  $C$  和决策属性  $D$  之间往往存在着某些程度的依赖和关联, 这种依赖性可以用依赖度来定义:

$$\gamma_C(D) = |\text{POS}_C(D)|/|U| \quad (2)$$

其中  $| \cdot |$  表示求集合的基数,  $\text{POS}_C(D) = \bigcup_{x \in U/D} \underline{C}(X)$ ,

称为  $D$  的  $C$  正域,  $\gamma_C(D)$  衡量了根据属性  $C$  描述的知识, 论域对象能够被正确划分到相应决策类的比例。

由于条件属性之间的依赖与关联, 对决策属性来说, 并非所有的条件属性都是必要的, 从而引出属性的约简问题, 约简可以理解为在保证分类能力不变的前提下, 最简单地表示决策系统的决策属性对条件属性集的关联和依赖。

如果条件属性  $C$  的非空子集  $C'$ , 满足条件: (1)  $\gamma_C(D) = \gamma_{C'}(D)$ ; (2) 不存在  $C'' \subset C'$ , 使得  $\gamma_{C''}(D) = \gamma_{C'}(D)$ , 则称  $C'$  为  $C$  的一个关于  $D$  的约简。

## 3 论域空间的近似域分割

对于决策系统而言, 通常所要获取的知识是条件属性与决策属性之间的因果关系, 因此条件属性分类相对于决策属性的分类关系非常重要, 也就引出了相对近似域的概念, 针对决策系统  $S = (U, C, D)$ , 定义:

$$\text{POS}_C(D) = \bigcup_{x \in U/D} \text{POS}_C(X), \text{ 称为 } D \text{ 的 } C \text{ 正域};$$

$$\text{BN}_C(D) = \bigcup_{x \in U/D} \text{BN}_C(X), \text{ 称为 } D \text{ 的 } C \text{ 边界域};$$

**定理 1** 对于给定的有限集论域  $U$ , 决策系统  $S = (U, C, D)$ , 相对近似域  $\text{POS}_C(D)$  和  $\text{BN}_C(D)$  互为补集。

定理 1 不难证明, 根据定理 1, 相对近似域有如下性质:  $\text{POS}_C(D) \cup \text{BN}_C(D) = U$ ,  $\text{POS}_C(D) \cap \text{BN}_C(D) = \phi$ 。

根据相对近似域, 给出如下决策系统分解的定理:

**定理 2** 任何一个决策系统  $S = (U, C, D)$  都可以唯一分解为两个不相交的子决策系统  $S_1 = (U_1, C, D)$  和  $S_2 = (U_2, C, D)$ , 其中  $U_1 = \text{POS}_C(D)$ ,  $U_2 = \text{BN}_C(D)$ 。对于  $S_1$ ,  $\gamma_C^1(D) = 1$ , 而对于  $S_2$ ,  $\gamma_C^2(D) = 0$ 。

定理 2 的成立是易于理解的, 限于篇幅, 在此不给出详细证明。特殊情况下, 即  $S = (U, C, D)$  是完全一致的决策系统时, 有  $S_1 = S$ ,  $S_2 = \phi$ , 或者当  $S = (U, C, D)$  是完全不一致的决策系统时, 有  $S_1 = \phi$ ,  $S_2 = S$ 。

根据上述定理, 可以将一个决策系统的论域空间  $U$  划分为两个不相交的子空间  $U_1$  和  $U_2$ 。 $U_1 = \text{POS}_C(D)$  是所有根据条件属性  $C$  可以被唯一分类的对象集合, 在论域空间中, 对应决策表中由确定性规则描述的区域。 $U_2 = \text{BN}_C(D)$  是所有根据条件属性  $C$  不能被唯一分类的对象集合, 在论域空间中, 对应决策表中由可能性规则描述的区域。本文分别将其定义为确定性区域和可能性区域。

## 4 粗逻辑神经网络分类器的设计

粗逻辑神经网络的精度很大程度上取决于输入变量的离

离散化精度, 而离散化精度和网络的规模复杂性、推广泛化能力又存在一定的矛盾。从定理 2 可知, 有限论域的决策系统可以分解为对应确定性区域和可能性区域的两个决策子系统。显然, 在当前的离散精度情况下, 确定性区域的对象已经可以被唯一确定分类, 而可能性区域中的对象有相同的条件却可能属于不同的类别, 实际上, 可能性区域对应的是样本的混合区域, 而对混合区域的样本划分的粒度过大正是造成系统误分类、影响分类精度的重要原因。鉴于此, 本文只针对可能性区域进行进一步的细分, 细化其信息粒度, 从而在一定程度上既可提高分类精度, 又可限制网络规模增长过快。

#### 4.1 粗逻辑网络结构

综合以往各种粗逻辑网络结构的优缺点, 本文采用具有模糊量化层、权值可调、聚类中心和标准差参数可学习的粗逻辑网络结构, 如图 1。网络总共由 4 层组成。

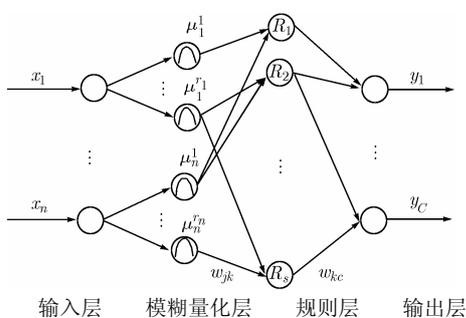


图 1

第 1 层为输入层, 输入  $x = (x_1, x_2, \dots, x_n)^T$ ,  $x_i (i = 1, 2, \dots, n)$  是输入特征变量。

第 2 层为模糊量化层, 依据某种离散化方法, 分别将输入变量  $x_i$  离散化为  $r_i$  个不同值。为有利于对连续数据的描述, 该层神经元激励函数可取为高斯型隶属函数:

$$\mu_i^j = \exp\left(-\frac{x_i - c_{ij}}{\delta_{ij}}\right)^2 \quad (3)$$

式中  $c_{ij}$ ,  $\delta_{ij}$  分别是被离散化到第  $j (j = 1, 2, \dots, r_i)$  档的特征变量  $x_i$  的平均值和标准差, 其初始值可由训练样本进行估算。

第 3 层为规则层, 每一个神经元代表一条规则, 规则的获取和描述依据的是粗糙集数据分析理论。该层神经元和第 2 层、第 4 层神经元的连接关系, 由规则的前件和后件来决定, 该层第  $k$  个神经元  $R_k$  的输入、输出:

$$I_k^3 = \sum_j w_{jk} \mu_i^j \quad (4)$$

$$O_k^3 = f_3(I_k^3 + \beta_k) \quad (5)$$

$k = 1, 2, \dots, s$ ,  $s$  是规则数,  $\beta_k$  为第  $k$  个神经元阈值,  $f_3$  是规则层的激励函数, 取为正切 S 函数 tansig:

$$f_3(X) = \frac{2}{1 + \exp(-2X)} - 1 \quad (6)$$

第 4 层为输出决策层, 每一个神经元代表一个决策类, 第  $c$  个神经元的输入、输出:

$$I_c^4 = \sum_k w_{kc} O_k^3 \quad (7)$$

$$y_c = O_c^4 = f_4(I_c^4 + \theta_c) \quad (8)$$

$c = 1, 2, \dots, C$ ,  $C$  是决策类数,  $\theta_c$  为第  $c$  个输出神经元阈值,  $f_4$  是输出层的激励函数, 取为对数 S 函数 logsig。

$$f_4(X) = \frac{1}{1 + \exp(-X)} \quad (9)$$

网络可调节的权值参数有  $w_{jk}$ ,  $w_{kc}$ , 及神经元阈值  $\beta_k$ ,  $\theta_c$ , 高斯隶属函数的中心和标准差参数  $c_{ij}$ ,  $\delta_{ij}$ , 均可由反向传播 BP 算法来进行训练。

#### 4.2 粗逻辑网络设计步骤

整个粗逻辑推理决策网络由两步设计来实现:

第 1 步 对训练样本进行粗略的离散区间划分, 进行属性约简及规则提取, 提取的规则集由确定性规则和可能性规则构成, 训练样本中与确定性规则匹配的训练样本被划分到确定性样本集, 与可能性规则匹配的样本被划分到可能性样本集。针对确定性样本集设计确定性粗逻辑推理网络;

第 2 步 对可能性样本集进行进一步的区间细化, 在新的离散精度区间重新进行属性约简及规则提取, 并依此构造可能性粗逻辑推理网络。最后, 分别通过确定性样本集和可能性样本集对网络进行训练。整个网络设计过程如图 2。

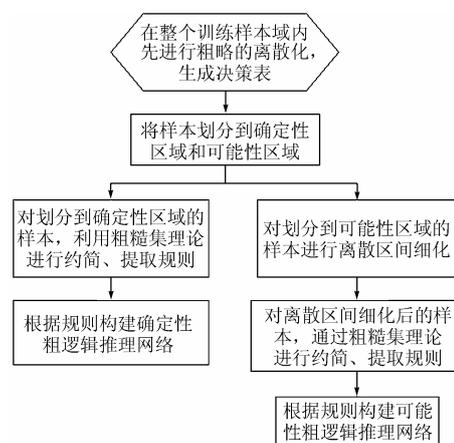


图 2 整个粗逻辑神经网络结构设计过程

#### 4.3 分类决策

网络经过训练后, 首先对测试样本或未知样本进行区域划分, 然后分别由相应的粗逻辑推理网络进行决策分类。对待分类的测试样本或未知样本进行区域划分时, 绝大部分样本可以与从训练集中提取的确定性规则或可能性规则匹配, 相应被分别划分到确定性区域和可能性区域。但由于训练样本有限, 或者有噪声等其它原因, 也可能造成极少数的测试样本或未知样本不能与从训练集中提取的确定性规则或可

能性规则匹配, 对于这小部分奇异样本的处理可以根据基于距离的度量原则, 将其归入到确定性区域或可能性区域进行分类。

假设在第 1 步的离散区间粗略划分时, 确定性区域由  $s$  个小区域组成、而可能性区域由  $k$  个小区域组成, 这里的每一个小区域均可以看作一个模式类。则

$$d_b(x) = \min_i(d_{b_i}(x)), i = 1, \dots, s, d_{b_i}(x) \text{ 是待分类样本 } x$$

到第  $i$  个确定性模式类中心的欧氏距离;

$$d_p(x) = \min_j(d_{p_j}(x)), j = 1, \dots, k, d_{p_j}(x) \text{ 是待分类样本 } x$$

到第  $j$  个非确定的可能性模式类中心的欧氏距离;

如果  $d_b(x) < d_p(x)$ , 则将  $x$  归入到确定性区域; 否则, 将  $x$  归入到可能性区域。

### 5 遥感图像分类实验

本文选择了长白山天池地区的 landsat TM 遥感图像进行实验, 虽然 TM 遥感图像共有 7 个波段, 但如果采用全部 7 个波段的信息, 在设计粗逻辑网络时, 会造成组合爆炸, 使提取的规则数过多, 导致网络规模过大, 因此, 本文只用了 Band5,4,3(分别对应 R, G, B 颜色分量)3 个波段的信息, 由这 3 个波段合成的图像, 各类地物界限清晰, 有利于目视解译和分类。

通过目视解译, 将地物分成 5 类, 湖水( $d_1$ )、裸岩( $d_2$ )、裸土( $d_3$ )、稀疏林草地( $d_4$ )、密林( $d_5$ )。从每一类地物中抽取部分样本, 提取了总共 1149 个样本, 其中 477 个样本作为训练样本, 514 个样本作为测试样本。另外选取了 158 个样本作为校验样本, 在训练过程中, 如果校验样本的分类性能出现恶化, 则训练结束。

在进行样本离散化时, 第 1 步的离散化可以比较粗略, 通常可以采用等频法或等区间法<sup>[10]</sup>, 实验中采用了等频法, 而在第 2 步对可能性区域进一步细化时, 我们采用了基于布尔推理的离散化方法<sup>[11]</sup>, 该方法可以在保证分类性能的基础上, 获得足够的离散精度。除第一步等频法的离散区间个数需人为给定外, 第 2 步对可能性区域离散化时采用的布尔推理方法可以自动确定离散区间的个数, 相应各离散区间的范围都可以由相应离散算法自动确定。实验中, 在第 1 步的等频法区间离散时, 每一波段分为 3 级, 相应离散区间见表 1。离散化后, 通过对由训练集构成的决策表进行粗糙集数据分析及规则约简后, 提取了 19 条规则, 其中规则 1~14 为确定性规则, 其余 5 条为可能性规则, 见表 2。根据定理 2, 训练样本集可以被划分为确定性样本集和可能性样本集, 在实际应用中, 与确定性规则匹配的训练样本被划分为确定性样本集, 与可能性规则匹配的训练样本被划分为可能性样本集。第 2 步, 通过布尔推理的离散化方法对可能性规则描述的区域进行了进一步的离散细化, 离散区间的划分情况见表 3。表 3 中的 0, 1, 2 表示采用的仍是原来的离散区间编码, 不需再进行细化。区间细化后, 原来的不一致决策规则集变

成了一致的决策规则集。  $a_2b_1c_1 \rightarrow d_3 \vee d_4 \vee d_5$

表 1 离散区间的粗略划分

波段编码	波段	离散区间编码		
		0	1	2
a	Band 5	[*, 66)	[66, 124)	[124, *)
b	Band 4	[*, 59)	[59, 133)	[159, *)
c	Band 3	[*, 39)	[39, 82)	[82, *)

表 2 第 1 步所提取的规则集

规则号	规则	规则号	规则
1	$a_0c_1 \rightarrow d_1$	11	$a_2b_1c_2 \rightarrow d_4$
2	$b_0c_2 \rightarrow d_1$	12	$a_1b_1c_0 \rightarrow d_5$
3	$a_0b_1c_2 \rightarrow d_2$	13	$a_0b_1c_0 \rightarrow d_5$
4	$a_1c_2 \rightarrow d_2$	14	$a_1b_2c_1 \rightarrow d_5$
5	$a_0b_2 \rightarrow d_2$	15	$a_0b_0c_0 \rightarrow d_1 \vee d_5$
6	$a_2b_2c_2 \rightarrow d_3$	16	$a_2b_2c_1 \rightarrow d_3 \vee d_5$
7	$b_2c_0 \rightarrow d_3$	17	$a_2b_1c_1 \rightarrow d_3 \vee d_4 \vee d_5$
8	$a_2b_0 \rightarrow d_4$	18	$a_1b_0c_0 \rightarrow d_4 \vee d_5$
9	$a_1b_0c_1 \rightarrow d_4$	19	$a_1b_1c_1 \rightarrow d_4 \vee d_5$
10	$a_2b_1c_0 \rightarrow d_4$		

表 3 对可能性规则描述区域的细化

可能性规则	Band 5	Band 4	Band 3	决策类
$a_0b_0c_0 \rightarrow d_1 \vee d_5$	[*, 30)	0	0	$d_1$
	[30, 66)	0	0	$d_5$
$a_2b_2c_1 \rightarrow d_3 \vee d_5$	[124, 152)	2	1	$d_5$
	[152, *)	2	1	$d_3$
$a_2b_1c_1 \rightarrow d_3 \vee d_4 \vee d_5$	[127, *)	[111, 133)	1	$d_3$
	[127, *)	[59, 111)	1	$d_4$
$a_1b_0c_0 \rightarrow d_4 \vee d_5$	[124, 127)	[111, 133)	1	$d_5$
	[80, 124)	0	0	$d_4$
$a_1b_1c_1 \rightarrow d_4 \vee d_5$	[66, 80)	0	0	$d_5$
	1	[59, 82)	1	$d_4$
	1	[82, 133)	1	$d_5$

根据上述步骤, 分别设计出确定性粗逻辑推理网络和可能性粗逻辑推理网络, 分别由相应的被划分到确定性区域和可能性区域的训练样本进行训练, 最后分别对确定性测试样本和可能性测试样本进行分类决策。为进行比较和验证, 实验中也采用常规的离散精度固定的方法设计了粗逻辑推理网络。离散化方法为等频离散法, 离散等级分别取 3, 5, 7, 9 这 4 种情况。表 4 给出了各种情况下的网络规模、训练时

间。由于输入层和输出层神经元数是相同的,分别是 3 个波段数据输入和 5 个决策类的输出,因此只需比较隐层的神经元数和权值数。初始规则是直接从决策表提取的还未进行规则约简的规则集,通过规则约简,可以获得精简的约简规则集。实验中采用的计算机为 PIII999, 256M RAM。从表 4 可以得知,常规方法中,随着离散等级的增加,隐层神经元数和需调节的权值数也在增加,同时训练时间延长。虽然对不同区域设计了两个粗逻辑推理网络,但两个网络整体的规模仍然是适中的,而且训练时间是最短的,其中确定性网络训练时间为 1.2min,可能性网络训练时间为 0.8min。

表 4 网络规模和训练时间

离散等级	隐层神经元数	权值数	初始规则数	约简规则数	训练时间 (min)	
常	3	27	66	21	19	2.0
规	5	50	112	49	35	5.6
方	7	62	125	77	41	6.2
法	9	78	150	112	51	7.4
本文方法	23/28	44/46	16/11	14/11		1.2/0.8

注:表示项“a/b”,前项和后项分别表示的是确定性网络和可能性网络的数据。

表 5 给出了各种情况下的测试样本的分类精度。在常规方法中,随着离散精度的提高,分类精度也越高,在离散等级为 7 时,分类性能最好,正确分类率达到 96.1%,但当离散等级进一步增加到 9 时,分类性能反而下降,说明离散等级并非越多越好,当超过某一阈值时,会造成信息颗粒划分过细,反而影响网络的推广泛化能力,导致分类性能下降。本文方法以较小的网络代价和训练时间,使分类精度也达到了 95.9%,非常逼近常规方法中的最好结果,说明本文方法在平衡网络复杂性、推广泛化能力和分类精度之间的矛盾方面确实有很好的效果。限于篇幅,分类结果图略。

表 5 测试样本分类精度(%)

离散等级	湖水	裸岩	裸土	稀疏林草地	密林	全部	
常	3	90.6	90.2	95.0	93.2	90.8	92.0
规	5	96.4	93.8	96.0	93.8	92.5	94.3
方	7	96.4	97.2	98.0	92.9	96.2	96.1
法	9	91.8	90.5	97.0	94.6	93.4	93.5
本文方法		96.4	98.0	98.0	92.4	95.3	95.9

## 6 结束语

(1)根据粗糙集不可分辨关系和近似域的概念,任何有限论域均可分解为确定性区域和可能性区域,为针对不同区域进行相应处理提供了理论依据和可能;

(2)常规的粗逻辑神经网络设计中,离散精度对网络的性

能影响很大。通常情况下,离散精度越高,信息颗粒越细,分类精度越高,但网络的规模越复杂、训练时间越长,而且当离散精度达到一定阈值时,会因为信息颗粒划分过细,造成网络的推广泛化能力下降,影响分类性能;

(3)通过对样本域的分割,本文只针对混合样本区(可能性区域)进行离散区间的细化,避免了对整个论域空间的细化处理,从而达到了提高分类精度,同时有效抑制网络规模增长过快的目的。

## 参考文献

- [1] Pawlak Z. Rough set theory and its application to data analysis. *Cybernetics and Systems*, 1998, 29(9): 661-688.
- [2] Pawlak Z. Rough set theory for intelligent industrial applications. *Intelligent Processing and Manufacturing of Materials*, 1999, IPMM'99, Honolulu, 10-15 July, 1999, 1: 37-44.
- [3] Pawlak Z. Rough logic. *Bulletin of the Polish Academy of Sciences, Technical Science*, 1987, 35: 253-258.
- [4] 张东波, 王耀南, 易灵芝. 粗集神经网络及其在智能信息处理领域的应用. *控制与决策*, 2005, 20(2): 121-126.  
Zhang Dong-bo, Wang Yao-nan and Yi Ling-zhi. Rough neural network and its application to intelligent information processing. *Control and Decision*, 2005, 20(2): 121-126.
- [5] Mohua B, Sushmita M, and Sankar K P. Rough fuzzy MLP: knowledge encoding and classification. *IEEE Trans. on Neural Networks*, 1998, 9(6): 203-216.
- [6] Yu C Y, Wu M H, and Wu M. Combining rough set theory with neural network theory for pattern recognition. *International Conference on Robotics, Intelligent Systems and Signal Processing*, Changsha, 2003: 880-885.
- [7] Chen S Y and Yi J K. A fuzzy neural network based on rough sets and its applications to chemical production process. *Info-tech and Info-net*, 2001. Proceedings, Beijing Oct 29-Nov 1, 2001, 4: 405-410.
- [8] Wu Z C. Research on remote sensing image classification using neural network based on rough sets. *Info-tech and Info-net*, 2001. Proceedings, Beijing Oct 29-Nov 1, 2001, 4: 79-284.
- [9] Liu H J, Tuo H Y, and Liu Y C. Rough neural network of variable precision. *Neural Processing Letters*, 2004, 19(1): 73-87.
- [10] Dougherty J, Kohavi R, and Shami M. Supervised and unsupervised discretization of continuous features. *Proceedings of 12<sup>th</sup> International Conference on Machine Learning*, CA: Morgan Kaufmann, 1995: 194-202.
- [11] Nguyen H S and Skowron A. Quantization of real-valued attributes, rough set and Boolean reasoning approaches. *2nd International Joint Conference on Information Sciences*, Wrightsville Beach, NC, USA, 1995: 34-37.

张东波: 男, 1973 年生, 讲师, 博士生, 从事粗糙集、神经网络、图像处理和模式识别等的研究。

王耀南: 男, 1957 年生, 教授, 博士生导师, 从事人工智能、智能控制、图像处理和模式识别等的研究.