

多分类问题的 RBF 二叉神经树网络方法

岳喜才¹, 叶大田¹, 管桦²

¹(清华大学 电机系, 北京 100084)

²(空军工程大学 科研部, 陕西 西安 710068)

摘要: 神经网络是一种普遍使用的分类方法。当类别数目较大时,神经网络结构复杂、训练时间激增、分类性能下降。基于两类问题的树网络多分类方法将两分类方法和判决树相结合,利用两分类方法来减少神经网络的训练时间,利用树型分类器来提高识别率。提出了一种多分类问题的二叉神经树网络结构和训练算法。利用两分类网络的训练结果对类别进行排序处理,并应用排序后的类别序号构成树型分类器,使可分性最差的类别的识别率提高最大,从而提高了整体分类性能。使用径向基函数(RBF)网络作为节点网络,使节点网络结构适应两类间的可分性,从而最终优化了神经树网络的结构。仿真实验表明该方法的分类性能优于现有方法。

关键词: 径向基函数网络;模式识别

分类号: TP391 **文献标识码:** A **文章编号:** 1009-3516(2000)01-0034-06

人工神经网络是由大量神经元互连形成的非线性网络。它具有学习能力、记忆能力、计算能力以及各种智能处理功能,它在不同程度和层次上模仿人脑神经系统的信号处理、存储及检索功能,是对人脑的简化与模拟,反映了人脑的基本特性。人工神经网络具有任意复杂的模式分类能力,这使其成为一种普遍使用的模式分类器。然而当需要增加新的类别时,神经网络需重新训练。网络训练通常需要较长的时间,在网络规模较大时尤其如此,这给系统的扩充、修改和维护带来很大的不便,所以使用单一网络进行分类在系统结构上是很不灵活的。除了结构上的不灵活外,单个大网络还存在难以胜任大类识别这一致命的缺点。这是因为当 N 增大时,神经网络的训练时间以指数增长^[1],当 N 很大时,将很难完成神经网络的训练。

一个多分类(N 类)问题可转化成 $N-1$ 个两分类问题,其中第 i ($1 \leq i \leq N$) 个问题是把 i 类和其他 $N-1$ 类相区分。我们可分别建立 $N-1$ 个网络,每个网络仅需确认一个类别^[2]。在训练过程中,对类别 i ,网络仅有两种输出;是 i 类和非 i 类,所以训练数据除包括类别 i 本身的数据外,训练数据还应包括类别集合中所有其他类别的数据,这样一方面存在训练数据的均衡问题,另一方面随着 N 数目的增大,网络的训练数据量也随之增大,进而使得网络的训练时间随着类别数的增大而增加。就训练数据量而言,该方法并未触及大类识别中的大数据量训练问题,故而该方法仅部分解决了网络的结构问题(至少使输出层的神经元的数目减为 1),未从根本上解决大类别分类问题。

多分类问题的另一转化方法是将 N 分类问题两两组合作为子类,这样用于 N 类识别的一个大网络被分成了 $C_N^2 = N(N-1)/2$ 个子网络。神经网络的大类别分类问题比较复杂,但其对小类别分类则性能较好,因此这种转换方法特别适合用神经网络来实现大类别分类。文献[3]基于两分类方法,提出了一种可用于说话人识别的神经阵列网络结构和搜索方法,说话人识别实验证明该方法可用很短的训练语音和识别语音获得较高的识别率。文献[4]证明了神经阵列网络方法可降低分类错误率,并由仿真计算说明神经阵列网络更适合于解决大类别分类问题和复杂分类问题。神经阵列网络方法可从 N 分类问题的全部 $N(N-1)/2$ 个子网络中快速搜索到其中的 $N-1$ 个子网络,经 $N-1$ 次判决后得出分类结果,从而将复杂的 N 分类问题转化为简单的两分类问题。本文在文献[3、4]的基础上,提出了一种二叉神经树网络结构,它从两分类子网络的训练

结果中得到各类别的可分性信息,并以此作为神经树网络结构设计的依据。由于神经树网络可利用类别可分性这一分类先验信息,故而可望进一步提高识别率。

1 神经树网络

1.1 神经网络阵列的结构及搜索方法

将大类别分类问题归结为两分类问题后,一个 N 分类问题被分成了 $N(N-1)/2$ 个两分类问题。为有效地从 $N(N-1)/2$ 个两分类问题中得到 N 分类结果,文献[3]提出的两分类神经网络阵列,其结构及搜索算法如下:

将子网络按类别序号规律进行排列,就形成神经网络阵列。对 N 类问题,其子网络排列如图 1 所示,图中所有的子网络排成 $N-1$ 行和 $N-1$ 列,行号和列号分别对应类别序号,其中最大的行号还对应类别的个数,图中行和列的交叉点为该行和列所对应的两个类别的子网络。从图中可清楚地看到,对 N 类问题的网络阵列,再增加一个类别后,原有的子网络不变,仅需增加 N 个子网络,也就是说网络结构具有可扩充性。

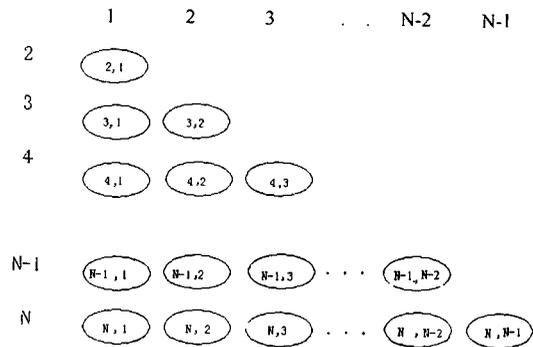


图 1 神经网络阵列的结构

在神经网络阵列中,子网络数 $N(N-1)/2$ 远大于分类数 N 。根据网络阵列的结构,文献[3]确定了一种识别搜索方法。该方法可完成对全部子网络集的搜索并最终得出分类结果。

网络阵列中每个子网络仅由其所在行和所在列分别对应的两类模式训练,在识别时对某一子网络作出的是模式 A 而不是模式 B 的判决结果,我们仅可理解为象模式 A 而非模式 B。这样每个子网络虽不能确定一模式,但可排除另一模式。神经网络阵列的搜索算法可描述如下:对图 1 的网络结构,从网络的顶层开始搜索,采用自上而下的搜索顺序,每排除一模式,则放弃与该模式有关的全部子网络,即放弃图 1 中该模式所对应的全部子网络。由于网络中每一层比上一层仅增加了一个子网络,而上一层的判决结果又放弃了网络的一个列,所以当搜索到某一层时,该层实际上已仅剩唯一一个子网络。经全部的 $N-1$ 层搜索后就得出分类结果。

在实际的 N 分类问题中,由于类别号未知,无法确定与待识类别相关的 $N-1$ 个网络,按搜索算法我们仍可选出 $N-1$ 个网络,然而却不能保证选出的 $N-1$ 个网络都与待识类别相关。

假定待识类别号为 $i(2 \leq i \leq N)$,其所选的 $N-1$ 个网络有两种:

(1)当从行号 2 搜索到行号 i 前,所选出的 $i-2$ 个网络都与该类别不相关,也就是说待识数据需进入仅与另外两个类别相关的网络,由于子网络不具备识别第三种模式的能力,子网络的输出结果将很难预料。但在搜索算法中,每一子网络只需作出一个排除决定,就上述情形而言,无论排除该两类别中的任何一个都不是错误的决定。

(2)从行号 i 开始,选出的子网络是与待识类别相关的。由于每一子网络仅需区分两个类别,经充分训练后一般不易发生分类错误。

1.2 神经二叉树网络

在上述的神经网络阵列中,每一类别的识别过程都是经过 $N-1$ 个网络。这初看起来似乎很象将每一类别与其它 $N-1$ 个类别相区分,但由于网络阵列的结构和搜索方法的特点,使得在神经网络阵列中某一类别并非需同其他的 $N-1$ 个类别相区分。搜索算法所选出的 $N-1$ 个子网络并非全部与待识类别相关,当不相关时,子网络的输出并无错误发生。 $N-1$ 个子网络中仅有 $N-1-(i-2)=N-i+1$ 个网络与类别 i 相关,此时类别 i 仅需与其他 $N-1$ 个类别中的 $N-i+1$ 个类别进行区分。换句话说,在去除无关子网络后,搜索路径实质上是一非等长路径,其长度取决于类别 i ,且 i 序号越大则相关路径越短,该类的识别错误率也越低。基于这一考虑,我们提出一种利用类别序号信息的神经阵列网络结构,并根据网络结构的特点将其改造成神经树网络以期进一步提高识别率。

假定类别 i, j 的子网络为 $S_{i,j}$,其对应的识别率为 $P_{i,j}(0 < P_{i,j} \leq 1)$ 。对于 N 分类的单个大网络,由于每一类别需同时与其它的 $N-1$ 个类别相区分,类别 i 的分类正确率为

$$P_{Ai} = \prod_{\substack{j=1 \\ j \neq i}}^N P_{i,j} \quad (P_{i,j} = P_{j,i}) \quad (1)$$

式中 P_{Ai} 是在 i 和其他任何其他类别之间不产生任何错误的概率期望值。 N 分类器的分类正确率为^[5]

$$P = E[P_{Ai}] \quad (2)$$

在图 1 中,搜索算法所选出的 $N-1$ 个子网络与类别 i 相关的子网络为 $\{S_{i,1}, S_{i,2}, \dots, S_{i,i-1}\}$ 中的一个和全部的 $\{S_{i+1,i}, S_{i+2,i}, \dots, S_{N,i}\}$ 。在识别中,从 $\{S_{i,1}, S_{i,2}, \dots, S_{i,i-1}\}$ 中随机选取过程应服从均匀分布,故而类别 i 的分类正确率为

$$P'_{Ai} = \frac{1}{i-1} \sum_{j=1}^{i-1} P_{i,j} \prod_{k=i+1}^N P_{k,i} \quad (3)$$

由于 $0 < P_{i,j} \leq 1$, 易证

$$\frac{1}{i-1} \sum_{j=1}^{i-1} P_{i,j} \geq \prod_{j=1}^{i-1} P_{i,j} \quad (4)$$

故而

$$P_{Ai} = \prod_{\substack{j=1 \\ j \neq i}}^N P_{i,j} = \prod_{j=1}^{i-1} P_{i,j} \prod_{k=i+1}^N P_{k,i} \leq \frac{1}{i-1} \sum_{j=1}^{i-1} P_{i,j} \prod_{k=i+1}^N P_{k,i} = P'_{Ai} \quad (5)$$

从以上分析可知,单个网络时类别 i 的识别过程相当于在图 1 中选择了与 i 相关的全部的 $N-1$ 个子网络 $\{S_{i,1}, S_{i,2}, \dots, S_{i,i-1}\}$ 和 $\{S_{i+1,i}, S_{i+2,i}, \dots, S_{N,i}\}$, 而神经阵列网络将 $\{S_{i,1}, S_{i,2}, \dots, S_{i,i-1}\}$ 从级联关系(相乘)变为了随机选一,对均匀分布时为取均值(相加平均)。对类别 i ($2 < i \leq N$), 使用神经阵列网络后该类识别率的改进率为

$$\Delta = \frac{P'_{Ai} - P_{Ai}}{P_{Ai}} = \frac{\frac{1}{i-1} \sum_{j=1}^{i-1} P_{i,j}}{\prod_{j=1}^{i-1} P_{i,j}} - 1 \quad (6)$$

上式中,对不同 i 值识别率的改进率显然不同。由于分式的分子项为一取均值的计算,故而该项变化较小,而分母为一累积,由于 $0 < P_{i,j} \leq 1$, i 值越大分母越小,使得识别率的改进率越大。

在神经阵列网络中,类别序号是随机的,这使得改进率最高的类别并非是最需要改进的类别,即类别可分性最差的类别(类别可分性由(1)式计算)。由于一个 N 分类器的分类正确率为 $P = E[P_{Ai}]$, 故而要提高 N 分类器的识别率则应使可分性最差的类别的识别率改进最大。由(6)式知将可分性最差的类的序号改为最大即可。

基于上述分析,对于 N 分类问题,我们提出一种二叉神经树网络结构,其构成步骤如下:

1) 任意给定类别序号,将 N 类两两组合,分别完成对 $N(N-1)$ 个子网络的训练。

2) 根据(1)式分别计算各类别的可分性。对类别可分性进行排序。

3) 按可分性从好到差的次序,从小到大依次分配类别次序。

4) 对新类别次序,按图 1 构造神经阵列网络。

5) 从第 2 行到第 $N-1$ 行,将神经阵列网络中的每一子网络与它下一行的两个子网络相连:下一行同列的子网络和下一行最右边列的子网络。

6) 在最后一行增加叶接点。以子网络对应的两类别分别作为该子网络的两个叶接点。

对于树型分类器,其类别识别搜索算法为:

1) 从根接点(子网络 $S_{2,1}$)开始搜索。

2) 对接点 $S_{i,j}$,如果网络的输出值大于 0.5,则剪去右枝,否则剪去左枝。

3) 沿路径搜索直到到达叶接点。

1.3 径向基函数(RBF)二叉树网络

采用两分类方法后,各个子网络独立进行训练,这使得子网络的结构具有灵活性。也就是说,各子网络结构可随两类别间的可分性而改变,即较难区分的两类,其对应的识别网络隐层神经元数目较大,网络结构复杂,反之亦然。当选用径向基函数(RBF)网络^[6]时,隐层神经元的个数可由训练数据自动确定^[7],故而我们将子网络的型式选为 RBF 网络,形成神经树网络。

RBF 网络具有很强的生物学背景,在人的大脑皮层区域中,局部调节及交叠的感受野是人脑反应的特

点。基于感受野这一特性,Moody 和 Darken 提出了 RBF 网络。RBF 网络是一种两层前向网络,其结构示于图 2,它的输出是隐层基函数的线性组合。基函数一般选高斯形式:

$$\mu_{1,j} = \exp \left[- \frac{(X - W_{1,j})^T (X - W_{1,j})}{2\sigma_j^2} \right] \quad j = 1, 2, \dots, N_1 \quad (7)$$

式中 $\mu_{1,j}$ 是网络隐层(第一层)第 j 个节点的输出, $W_{1,j}$ 是该节点的权矢量, σ_j^2 是该节点的归一化参数, N_1 是该层的节点个数, X 为输入矢量。各隐层节点的输出在 0—1 的范围,输入信号越靠近高斯分布的中心,节点的输出越大。RBF 网络输出层节点完成隐层输出的线性组合:

$$y_j = W_{2,j}^T \mu_1 \quad j = 1, 2, \dots, N_2 \quad (8)$$

式中 y_j 为第 j 个节点的输出, $w_{2,j}$ 为该节点的权矢量, μ_1 为隐层的输出矢量, N_2 为该层的节点个数,故而 RBF 网络实现高斯基函数的加权和。

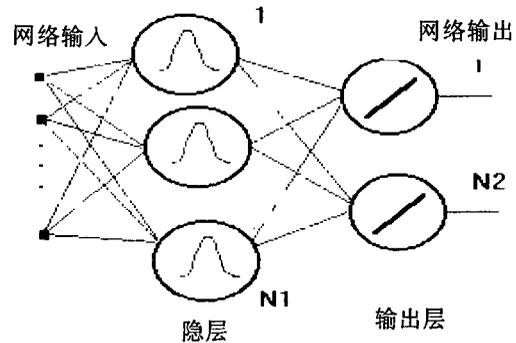


图 2 RBF 网络结构

RBF 网络有多种训练方法。正交最小二乘法^[6]是目前 RBF 网络训练中应用较多的一种方法。在前向网络的应用问题中,网络隐层神经元个数的选择问题是应用者需解决的一个基本问题,正交最小二乘法可在 RBF 网络的训练过程中自动确定隐层神经元的数目,从而解决了这一问题。

RBF 二叉树神经网络中使用的 RBF 网络结构很简单,其输入层由类别特征决定,输出层神经元数目为 1,隐层神经元数由训练过程自动确定。这一特点一方面使得两分类子网络的训练一般都比较充分,从而保证了二叉神经网络的识别效果,另一方面该特点为我们进行类别可分性的评估提供了一个简易的方法。在二叉神经网络的构成步骤 2 中,要使用(1)式计算类别可分性。由于网络的正确识别率指标并不能从训练过程中得到,计算这一指标常需另附加数据,这就增大了构成二叉神经网络的数据需求,给某些受数据量限制的应用带来困难。使用 RBF 网络,相同训练数据量和相同误差指标所得到的不同网络结构(即不同的隐层神经元数目)可作为衡量类别可分性的指标。故而(1)式中的 $P_{i,j}$ 可用 $N_{i,j}$ 来代替。这里 $N_{i,j}$ 代表识别类别 i, j 的 RBF 网络的隐层神经元个数。

2 神经二叉树网络识别性能仿真计算

对类别数为 5,任意给定的两类间识别率 $P_{i,j}$,我们分别仿真计算了单个大网络、神经阵列网络和神经树的分类正确率,示于表 1。其中大网络的识别率根据(1)式和(2)式计算,神经阵列网络的正确分类率根据(3)式计算。根据(1)式的计算结果调整类别序号,根据(3)式可计算神经树网络的识别率。表 1 中神经树网络的类别号是按单个大网络的类别可分性重新排定的,即单个大网络可分性最差的类别 1 为神经树网络的类别 5,由(6)式可算得其识别率的提高率为 17.88%,从而使得神经树网络的识别率较神经阵列网络进一步提高(见表 1)。

表 1 不同网络的识别率比较

类别序号	1	2	3	4	5	平均识别率
单个大网络	0.8145	0.8867	0.9039	0.8784	0.8625	0.8692
神经阵列网络	0.8145	0.8867	0.9437	0.9148	0.9639	0.9047
神经树网络	0.9475	0.8811	0.9022	0.9190	0.9601	0.9220

对于大类别分类问题,我们分别仿真计算了类别数为 50、60、70、80、90、100 的单个大网络、神经阵列网络和神经树网络的分类正确率。为便于对比,我们同时计算了与神经树网络类别反序排列的神经阵列网络的识别率,一并示于图 3。图中的曲线总体上呈下降趋势,说明随着类别的增加,识别问题的难度增加,导致识别率的减小。从图中我们可看出,神经树网络、神经阵列网络和与神经树网络反序排列(即类别可分性越差,类别序号越小)的神经阵列网络的三条识别率曲线形成一族,其值远大于单个大网络,这说明使用基于两分类方法的神经网络方法的识别性能远好于单个大网络。而就三条一组的曲线而言,神经树网络识别率曲线位

于最上方,神经阵列网络识别率曲线位于其他两条曲线的中间,最下边的一条是与神经树网络类别反序排列的神经阵列网络的识别率。与神经树网络类别反序排列的神经阵列网络的识别率小于神经阵列网络,这从另一侧面说明神经树网络对神经阵列网络中类别序列重排方法正确。相对神经树网络和与神经树网络反序排列的神经阵列网络而言,神经阵列网络中类别的序号是随机排列的,所以其识别率曲线位于中间,其最差值为其下方紧邻的曲线,而神经树网络则是其最佳识别率曲线,故而使用神经树网络可改进神经阵列网络的性能。

为便于比较,表2对神经树网络和单个大网络的识别性能进行了比较。表中识别率改进率是用单个大网络和神经树网络识别率之差除以单个大网络的识别率得到的。比较结果显示,使用神经树网络可在很大程度上提高识别率,且随着类别的增长,改进率越来越高。

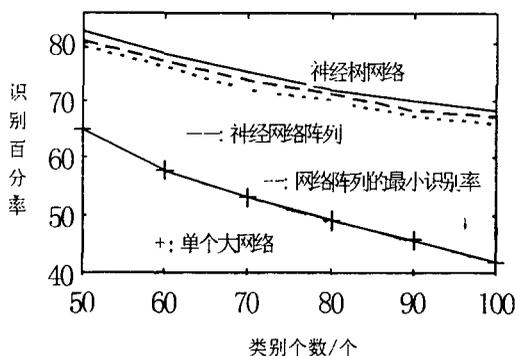


图3 神经树网络与神经阵列网络的识别性能比较

表2 神经树网络和单个大网络的识别率

类别个数	50	60	70	80	90	100
单个大网络(%)	64.9058	58.7821	52.8892	49.0109	45.0895	41.2447
神经树网络(%)	81.8154	78.2884	74.9348	72.3730	69.8588	67.4232
识别率改进率(%)	26.0526	33.1842	41.6826	47.6672	54.9335	63.4714

3 结论及讨论

神经树网络是对神经阵列网络的改进和优化,故而它具有神经阵列网络的结构灵活、便于修改已有类别和增加新的类别、缩短训练时间、提高识别率等优点。本文提出的二叉树结构使得神经树网络的结构简单,分类直观,搜索算法易于描述,同时由于附加了类别可分性这一先验信息,改进了分类效果。仿真计算表明神经树网络达到了优化网络结构和提高识别率的目标。本文将神经树网络的节点网络选为RBF网络,充分利用了该网络的最小二乘训练算法,一方面使得节点网络的训练更加充分,另一方面为神经树网络中类别序列的重新排列提供了简单可行的方法。

虽然神经树网络可用于解决任何复杂的分类问题,但对于一般分类问题,我们仍应注意分类策略的应用,即可先使用简单分类方法进行类别间的简单粗分类,对粗分类后的复杂分类问题再使用神经树网络进行分类。总之只有灵活应用分类策略和分类方法,才能更好地解决分类问题。

参 考 文 献

- [1] Rudasi L, Zahorian S A. Text-Independent Talker Identification with Neural Networks[J]. IEEE Proc. Inter. Conf. Acoustic Speech and Signal Processing, 1991: 389~392.
- [2] Farral K R, Mammone R J, Assaleh K T. Speaker Recognition Using Neural Networks and Conventional Classifiers[J]. IEEE Trans on Speech and Audio Processing. 1994, 2(1): 194~205.
- [3] 岳喜才, 伍晓宇, 郑崇勋. 用神经阵列网络进行文本无关的说话人识别研究[J]. 声学学报, 2000, 25(1).
- [4] 岳喜才, 伍晓宇, 郑崇勋等. 一种大类别数分类的神经网络方法[J]. 计算机研究与发展, 2000, 37(1).
- [5] Furui S. Digital Speech Processing, Synthesis and Recognition[M]. New York: Marcel Dekker, 1989.
- [6] Chen S, Cowan C F N, Grant P M. Orthogonal Least Squares Algorithm for Radial Basis Function Networks [J]. IEEE Trans on Neural Networks, 1991, 2(2): 302~309.
- [7] Mulgrew B. Applying Radial Basis Functions[J]. IEEE Signal Processing Magazine, 1996, (3): 50~65.

A Classifier of Binary Radial Basis Function Neural Tree Networks

YUE Xi-cai¹, YE Da-tian¹, GUAN Hua²

¹(Dept. of Electrical Machinery, Tsinghua University, Beijing 100084, China)

²(Dept. of Science Research, AFEU., Xi'an 710068, China)

Abstract: Neural network has been used for pattern recognition popularly. The training time of neural network for N catalogs classification increases exponentially with N, so it is difficult to deal with large number of catalogs by normal neural networks. Based on binary partition method and decision tree, a binary neural tree network (BNTN) classifier is proposed. Each node in BNTN is a simple neural network which only processes 2 catalog classification. Thus the architecture of BNTN is flexible and expansible, and the training time is reduced largely. The key to construct a BNTN is to sort the classes by separation of each class. We proposed a simple way to calculate separability of each class after radial basis function (RBF) neural networks have been selected as a type of node. Simulation shows that BNTN is better than other classifying methods.

Key words: RBF networks; Pattern recognition