

一种自动生成神经网络结构的新方法¹⁾

王继成 蔡义发 吕维雪

(浙江大学生物医学工程研究所 杭州 310027)

摘 要

神经网络已广泛应用于模式识别、优化计算等领域。但是，人们为了寻求某一问题的神经网络结构，往往采用穷举法，从而使得选择合适的神经网络结构随着问题规模的变大而变得越来越困难。本文根据神经元状态的变化导致人脑的空间结构和状态变化的研究，在神经网络中引入神经元的兴奋、抑制和突触修改机制、退化机制、死亡机制、自修复机制等，通过神经网络的学习，自动生成解决某一具体问题的合适的神经网络结构。实验结果表明，该方法是可行的、有效的，为神经网络结构的设计提供了一种新方法。

关键词：神经网络，认知科学，网络结构设计，自适应理论。

1 引言

目前，神经网络越来越广泛地应用于模式识别、优化计算和信号处理等领域。但是，对于某一具体问题，如何设计合适的神经网络结构，大多采用试凑法，主要靠主观经验判断，其中有许多弊病。例如，对于一个前馈神经网络用于模式分类，若隐含层或隐含神经元过少，则不能对某一具体问题完成分类的任务，不能进行充分的学习；反之，若隐含神经元过多，虽然可以完成训练样本的分类，但由于神经网络参数过多，产生了过剩学习，使神经网络的参数初值、训练样本的特异性、外界干扰对联接权的影响增大，使得输入模式和训练样本在有小的畸变情况下不能得到正确的分类结果。同时，由于隐含神经元的增加，大大增加了网络的训练时间。由此看来，对于某一具体问题，选择合适的神经网络结构，尤为重要。

Gutierrez 等人^[1]研究了一种包含有输入层、隐含层和输出层，层间神经元采用全联接方式，层内神经元无联接的前馈神经网络所需隐含神经元数的估计方法。一般来说，估计的隐含神经元数比实际需要的隐含神经元数多 10% 以上。Kung 和 Hwang^[2]采用了用代数理论来确定隐含神经元数的方法。这种方法要求神经网络在学习的初始阶段，联接权值已接近最佳分布。实际上，这是很难实现的。

本文根据神经元状态的变化导致人脑的空间结构和状态变化的研究，在神经网络中引入神经元的兴奋、抑制和突触修改机制、退化机制、死亡机制和自修复机制等，以及神经元的正常状态、睡眠状态、临界状态、病态和死亡状态及其神经元各状态的转换。从而，

1) 本课题得到国家自然科学基金资助。

本文于 1993 年 10 月 13 日收到

通过神经网络的学习,自动生成解决某一具体问题的合适的神经网络结构.为了验证该方法的可行性,作者用该方法设计了一个用于心电图分类的神经网络结构.实验结果表明,用该方法设计的神经网络,既有好的分类能力,同时又可以减少神经网络的训练时间.更为重要的是,该方法可以根据学习任务的变化,而自动调整神经网络的结构,适应外部环境的变化.

2 神经元状态及其转换理论

根据神经科学的研究^[3],人脑是由大约 10^{11} 个神经元组成的巨系统,但人在处理具体问题时,并不是所有的神经元都参加人的思维活动,有相当一部分神经元处于休眠状态,同时人脑每天都有神经细胞的死亡.因此,一般的神经网络系统是由许多神经元联接而成的,并且各神经元可以具有不同的状态:正常状态、睡眠状态、病态、临界状态和死亡状态.神经元的状态可以随着外界环境的变化而发生转移.

所谓睡眠状态,即神经元处于“活”的状态,但此时没有加入思维过程(神经计算过程);所谓正常状态,即神经元加入思维过程;所谓病态,即神经元处于一种不正常的工作状态;所谓临界状态,即神经元处于正常和病态之间的一种状态;所谓死亡状态,

即神经元停止工作.图 1 示出了神经元状态的变化.从图 1 可以看出,神经元状态的变化是由于其本身的退化和来自于其它神经元的激活信息等引起的.为此,研究神经网络联接权重和偏置修改规则,神经元退化规则,病态神经元修复规则和睡眠神经元唤醒规则.

联接权重和偏置修改规则为

$$\Delta W_{ij}(t) = \Delta W'_{ij}(t) + \Delta W''_{ij}(t) \quad (1)$$

$$\Delta \theta_i(t) = \Delta \theta'_i(t) + \Delta \theta''_i(t) \quad (2)$$

其中 $\Delta W'_{ij}(t)$ 为接收神经元 i 与发送神经元 j 在时刻 t 由于权重退化引起的联接权变化; $\Delta W''_{ij}(t)$ 为接收神经元 i 与发送神经元 j 在时刻 t 由于学习引起的联接权变化; $\Delta \theta'_i(t)$ 为神经元 i 在时刻 t 由于偏置退化引起的偏置值变化; $\Delta \theta''_i(t)$ 为神经元 i 在时刻 t 由于学习引起的偏置值变化.

神经元退化规则包括联接权重退化规则,偏置值退化规则,激活值退化规则.它们分别为

$$W'_{ij}(t) = W_{ij}(t) * e^{-ct}, \quad (3)$$

其中: c 为大于零的常数;

$$\theta'_i(t) = \theta_i(t) * e^{-at}, \quad (4)$$

其中: a 为大于零的常数;

$$a'_i(t) = a_i(t) * e^{-bt}, \quad (5)$$

其中 b 为大于零的常数.

当 $W'_{ij}(t) \in [-W_{\delta 2}, -W_{\delta 1})$ 或 $W'_{ij}(t) \in (W_{\delta 2}, W_{\delta 1}]$ 时(其中: $W_{\delta 2}$, $W_{\delta 1}$ 为预先设

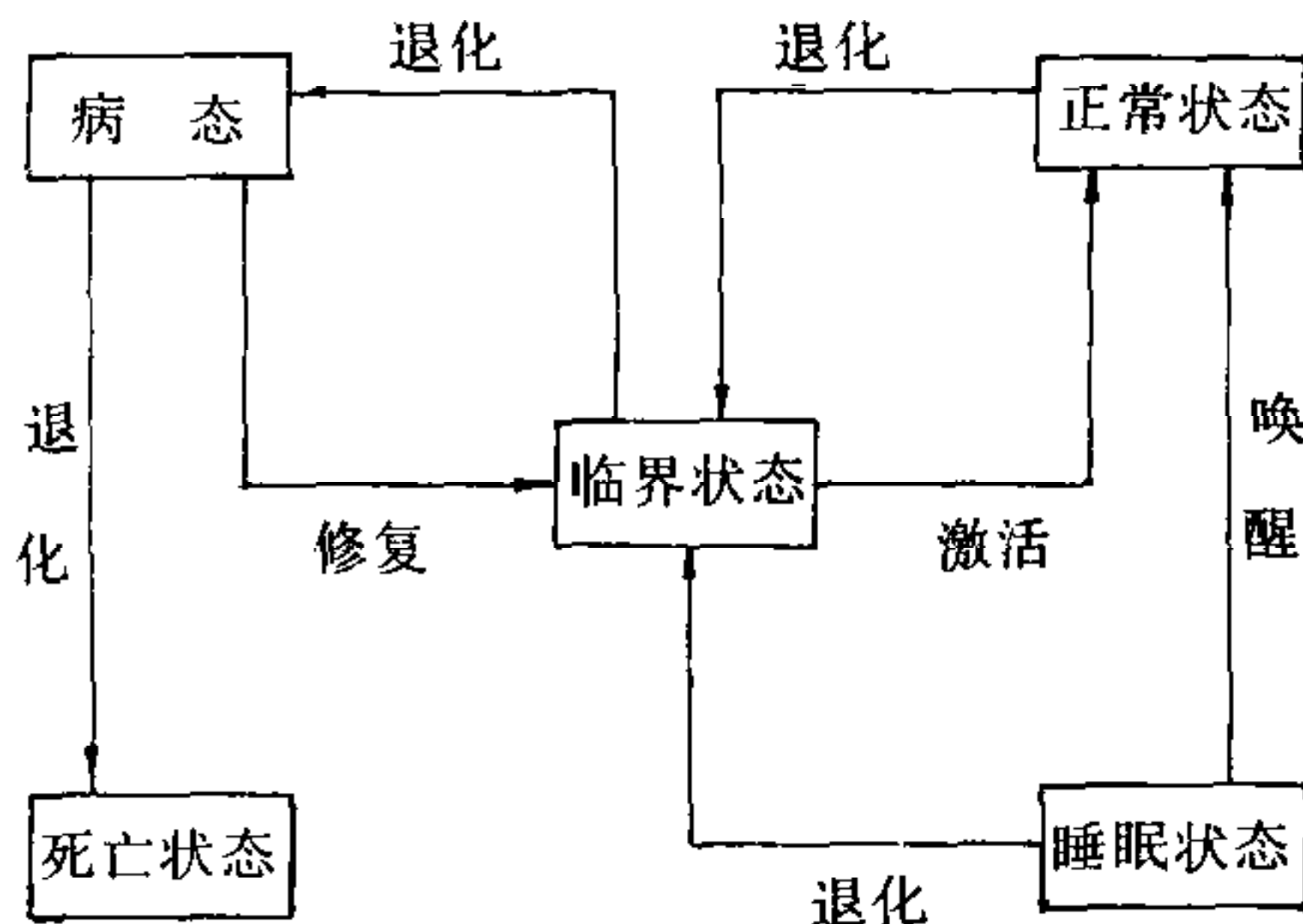


图 1 神经元状态转移图

置的联接权限值), 则称联接权 $W'_{ij}(t)$ 处于临界状态. 当 $\theta'_i(t) \in [\theta_{\max} - \theta_{\delta 2}, \theta_{\max} - \theta_{\delta 1})$ 时 (其中 θ_{\max} , $\theta_{\delta 1}$, $\theta_{\delta 2}$ 为预先设置的偏置限值), 则称神经 i 处于临界状态. 如果处于临界状态的联接权 $W'_{ij}(t)$ 继续退化, 当

$$|W'_{ij}(t)| < W_{\delta 1} \text{ 时,}$$

则称联接权 $W'_{ij}(t)$ 处于病态. 同样, 如果处于临界状态的神经元 i 的偏置值 $\theta'_i(t)$ 继续退化, 当 $\theta'_i(t) > \theta_{\max} - \theta_{\delta 1}$ 时, 则称神经元处于病态.

病态神经元修复规则: 如果病态神经元只是由于病态联接权引起的, 则删去病态的联接, 产生新的联接替换之; 如果病态神经元由于本身偏置退化引起的, 则删去该神经元, 唤醒睡眠神经元替换之; 如果此时无睡眠神经元可用, 则通过学习将该神经元的功能分布到神经网络中的其它神经元中去.

睡眠神经元唤醒规则: 当神经网络中某个神经元的所有联接权都处于病态或神经元由于偏置值退化而处于病态, 删除该神经元后, 就唤醒处于睡眠状态中的某个神经元, 以代替被删除神经元的功能.

3 自动生成神经网络结构的算法

本文所提出的自动生成神经网络结构的算法, 是通过神经网络学习过程中引入神经元的状态及其状态变化和神经元的自适应规则 (退化规则、病态神经元修复规则、联接权重和偏置修改规则、睡眠神经元唤醒规则等), 使得神经网络在学习某一具体问题, 可以根据问题的特点, 自动生成一个可以解决该问题的合适的神经网络结构. 为了说明的方便, 不妨假设需解决的问题为: 已知有 n 个输入变量, 要求根据输入变量的值, 进行分类, 且类别为 m . 采用前馈式层次型神经网络解决该分类问题, 其中输入层神经元数为 n , 负责接收 n 个输入变量的值, 输出层神经元数为 m , 输出最后的类别信息, 隐含层数和隐含神经元数以及各层次间神经元的联接通过学习自动产生. 具体的学习算法如下:

1) 从睡眠神经元中唤醒 l 个神经元作为隐含层神经元.

2) 隐含层神经元与输入层神经元之间、输出层神经元与隐含层神经元之间联接权修改规则为:

$$\Delta W_{ij}(t) = \Delta W_{ij}^{(BP)}(t) + W_{ij}(t-1) * (e^{-at} - 1) \quad (6)$$

其中: $\Delta W_{ij}^{(BP)}(t)$ 是采用 BP 学习算法时权值修改规则, a 是大于零的常数, 表示神经元联接权退化.

3) 当某个联接权值 $|W_{ij}(t)| < W_{\delta 1}$ (其中 $W_{\delta 1}$ 是预先设置的联接权限值), 则删除联接 W_{ij} , 无任何联接的神经元也被删除.

4) 重复步骤 2)、3), 若网络训练周期小于 N (N 为预先设定的训练周期限值), 且满足分类精度要求, 则此时所得到的神经网络结构为解决该分类问题的合适的神经网络结构.

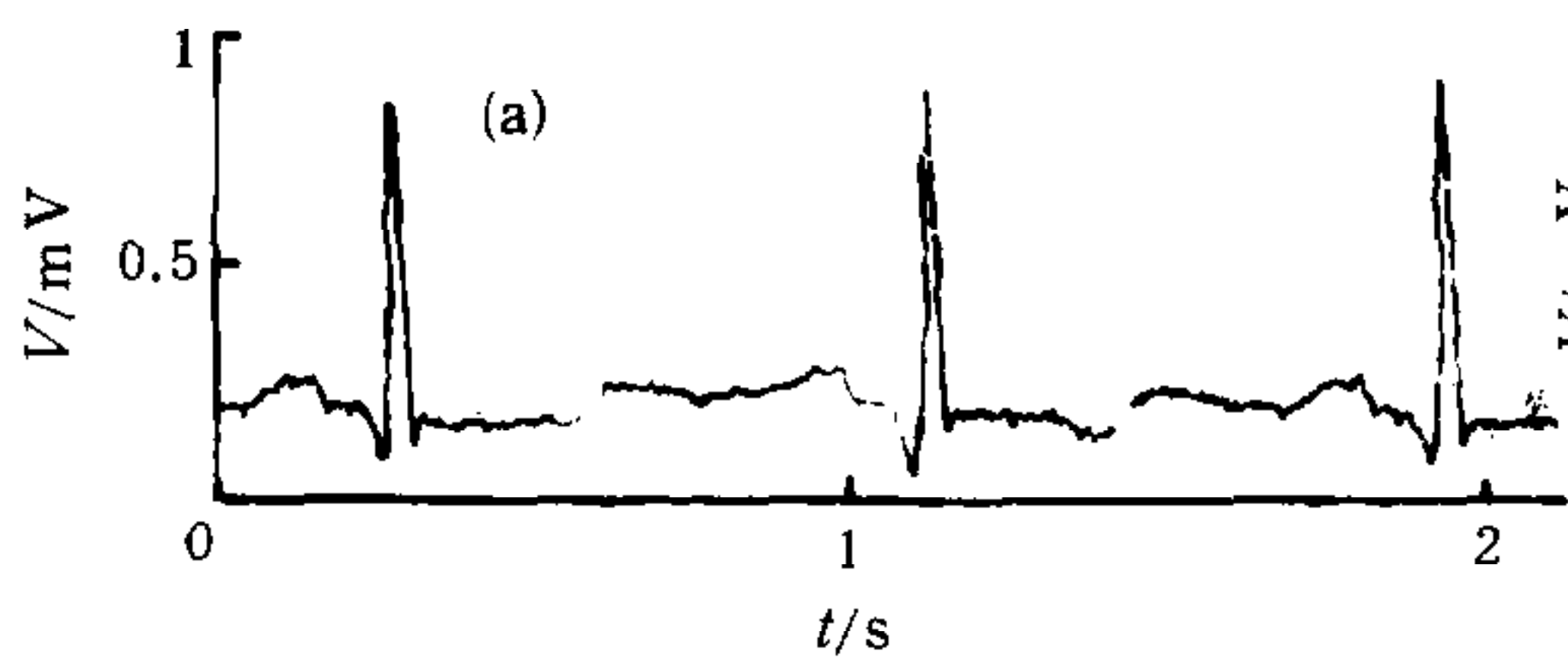
5) 若网络在训练周期的限值内不能满足分类精度要求, 则从睡眠神经元中唤醒神经元, 转步骤 2) (唤醒的神经元作为隐含神经元使用).

从上述的学习算法可以看出, 神经网络的结构 (隐含神经元数和神经元间的联接) 并

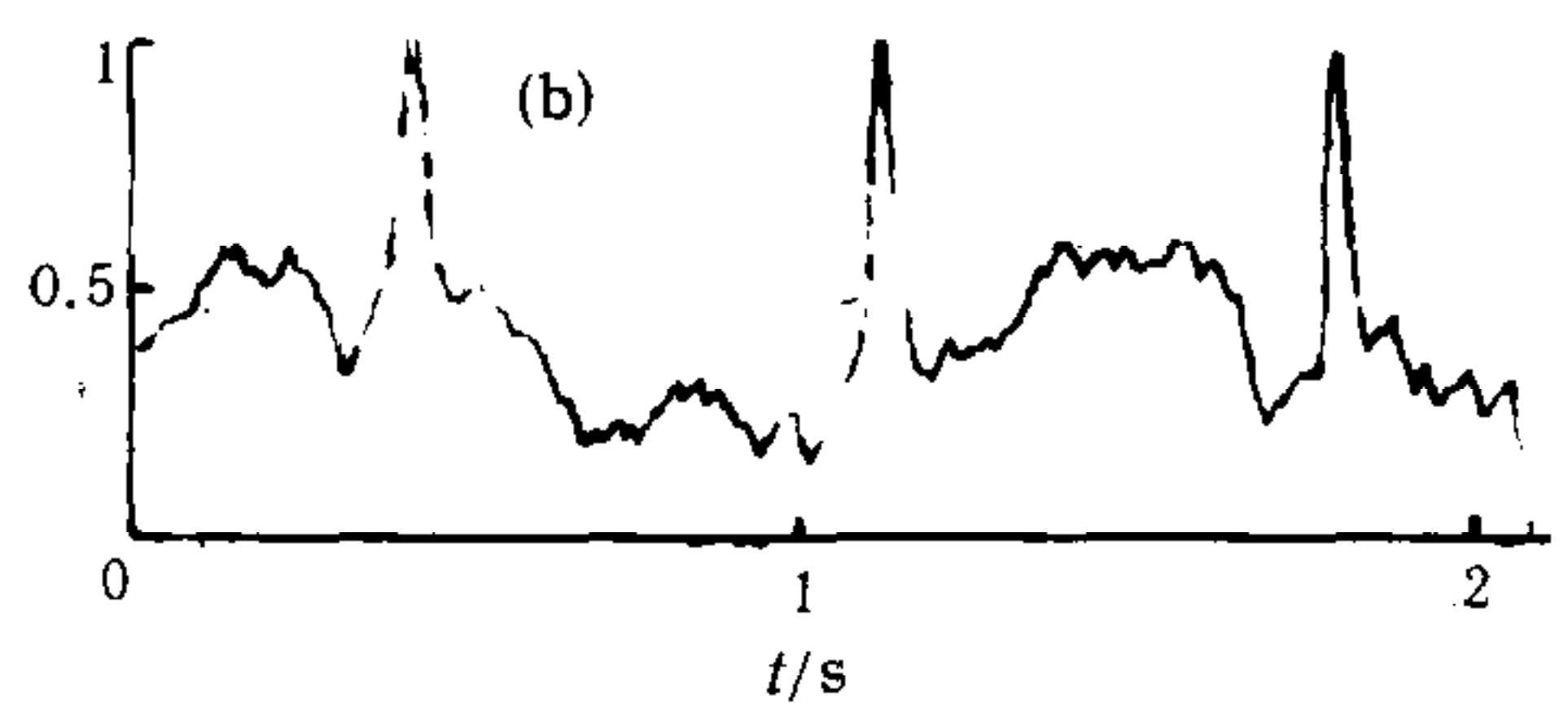
不是一成不变的,而是根据任务的要求,神经网络通过学习,自动产生实现任务所需的神经网络结构.并且,可以随着任务变化或训练样本的变化,通过学习产生新的神经网络结构来适应这种变化.事实上,人脑神经网络系统也是通过不断学习,使其解决问题的能力越来越强,功能越来越复杂.

4 实验与讨论

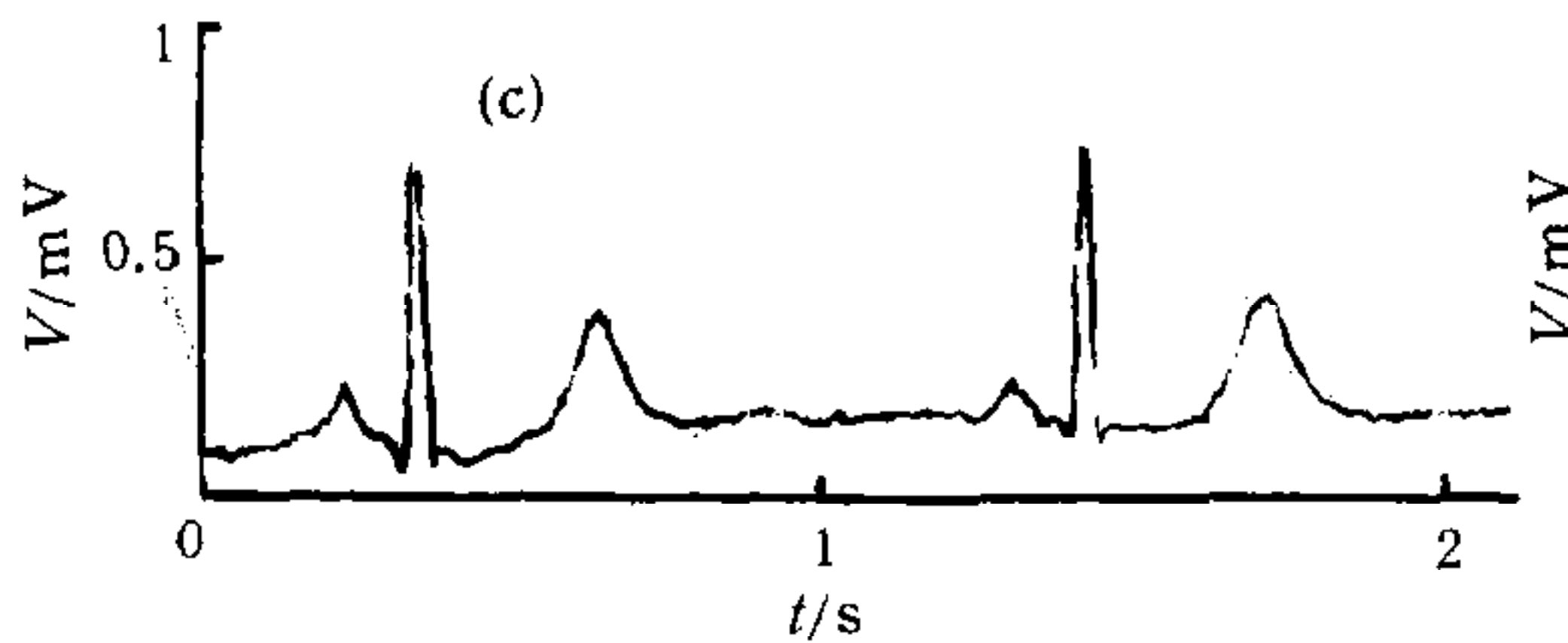
为了验证本文提出的自动生成神经网络结构的算法的有效性,设计了用于进行分类心电图的神经网络结构.该神经网络对美国 MIT BME 中心提出的 MIT/BIH 心电图数据库中的部分心电图波形 T100、T105、T106、T108、T111、T112、T217、T219、T220、



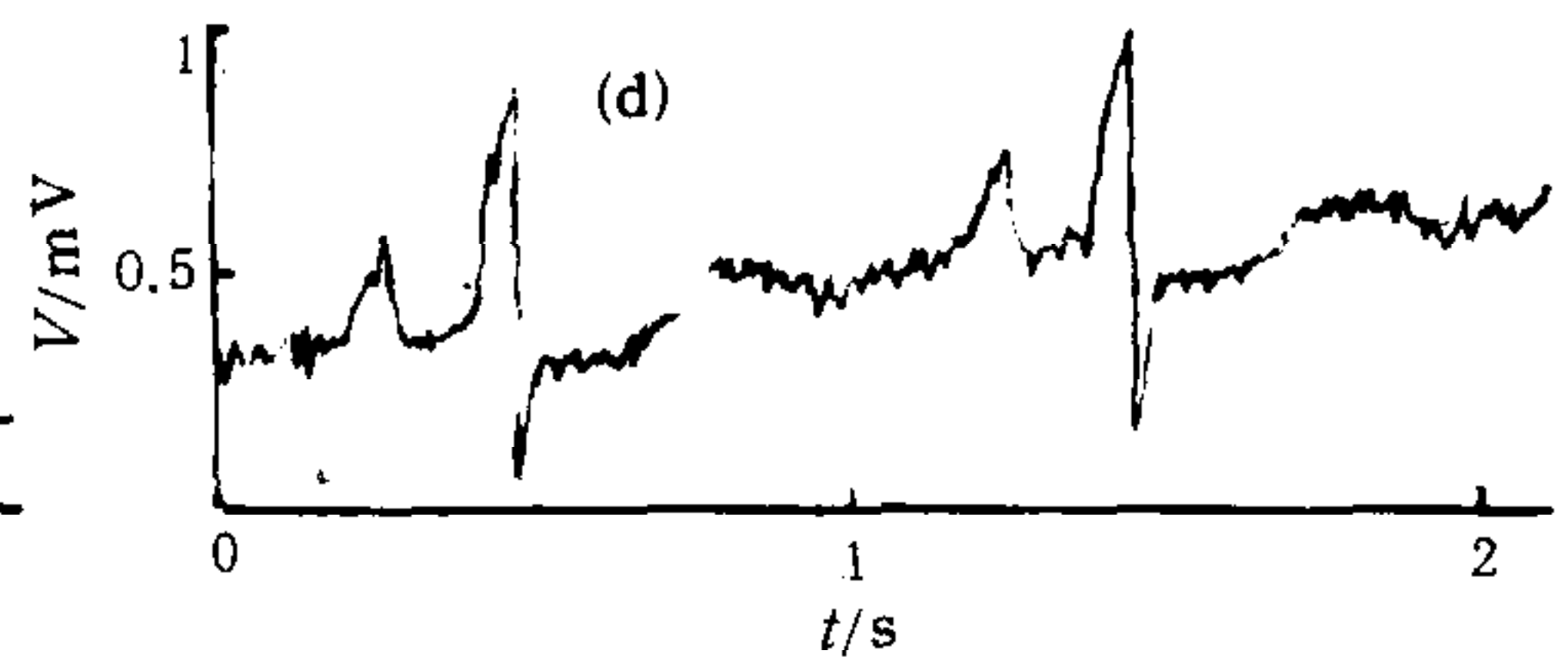
(a) T100



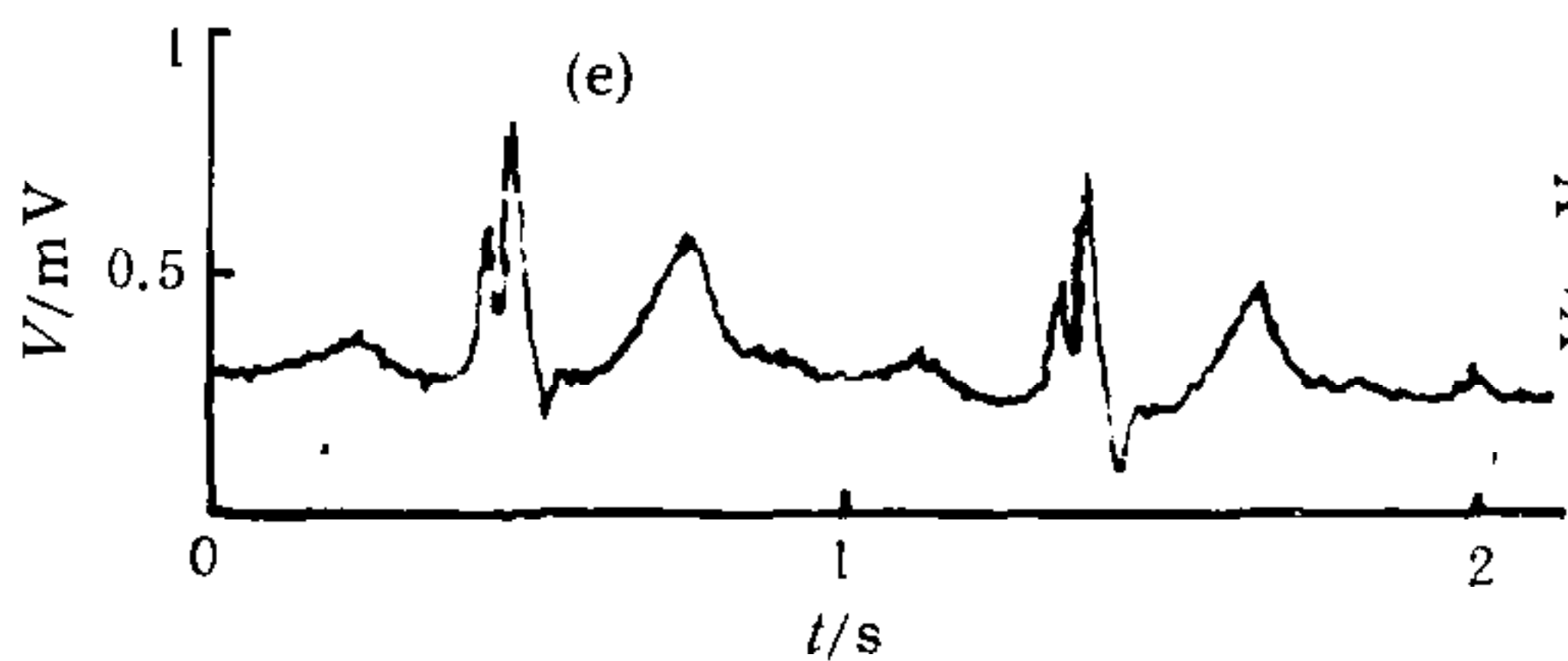
(b) T105



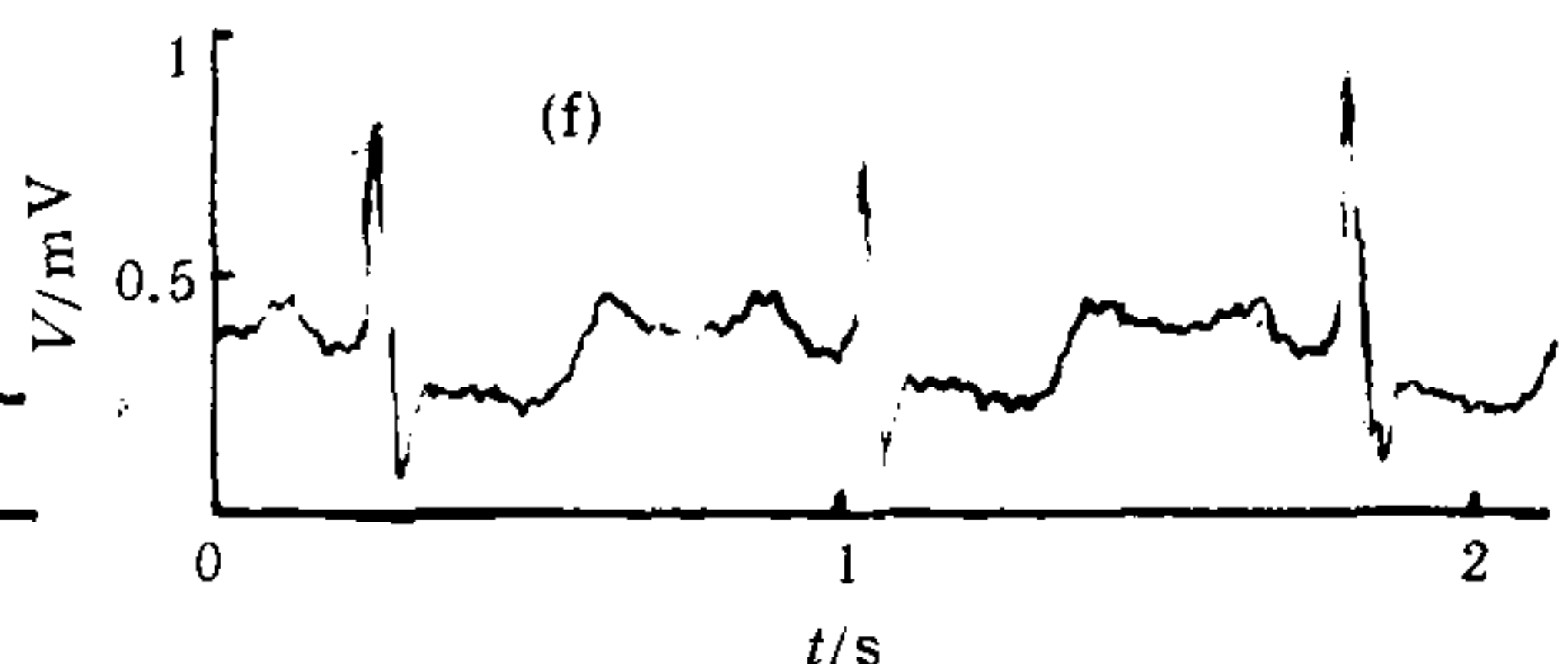
(c) T106



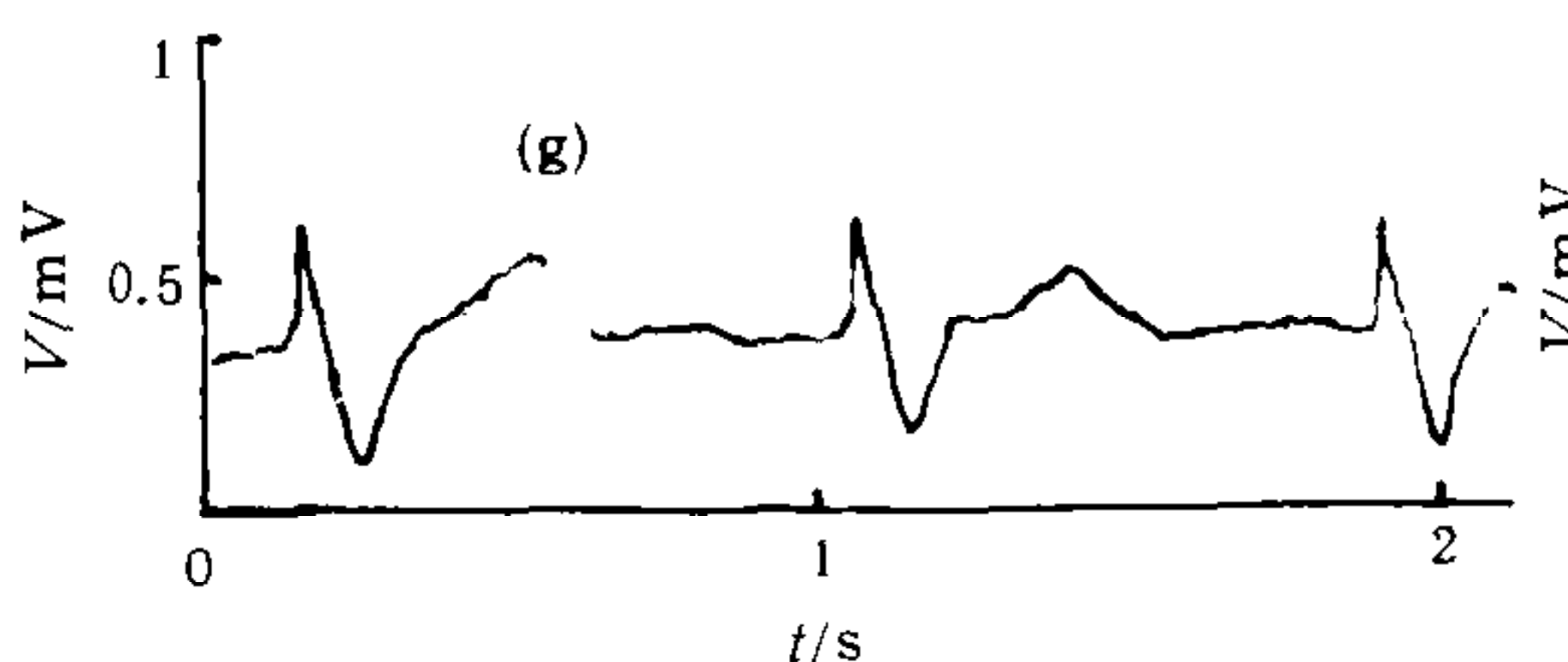
(d) T108



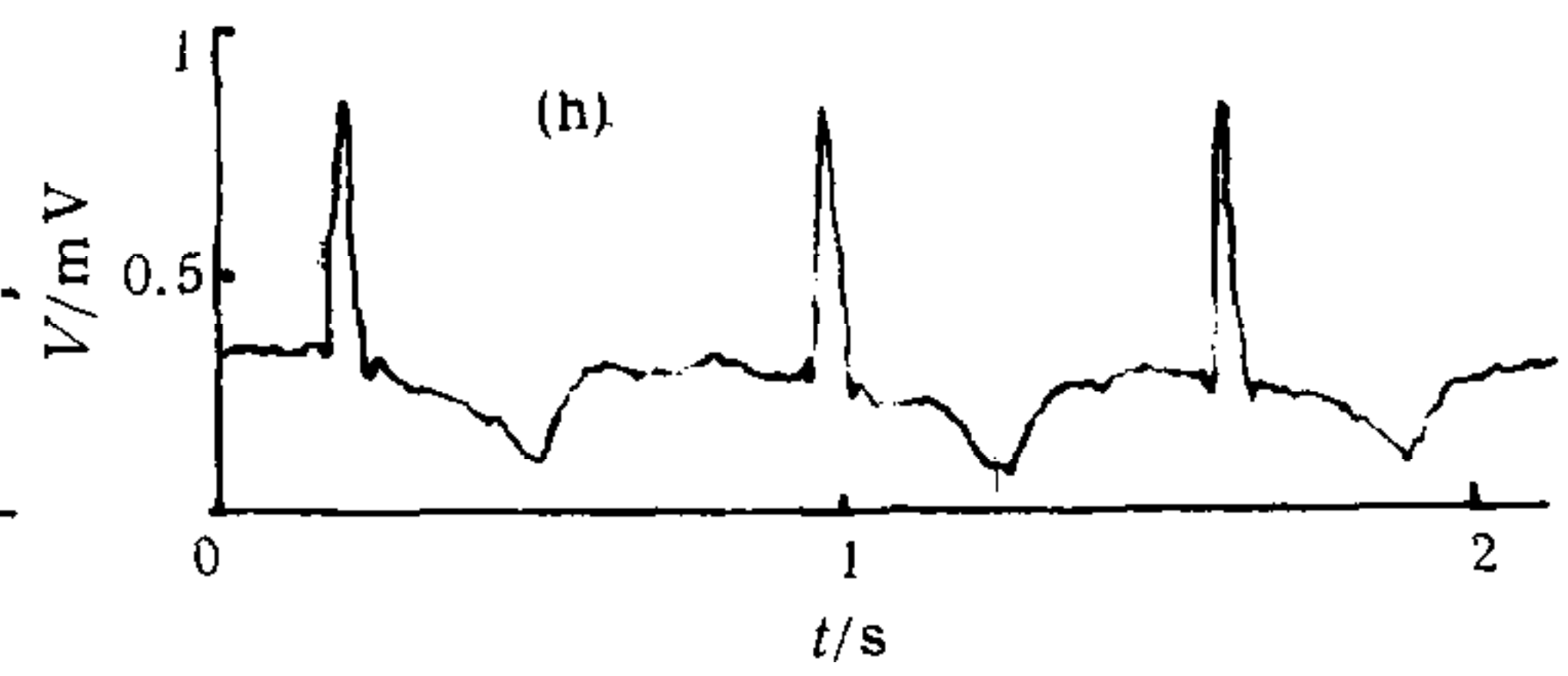
(e) T111



(f) T112



(g) T217



(h) T219

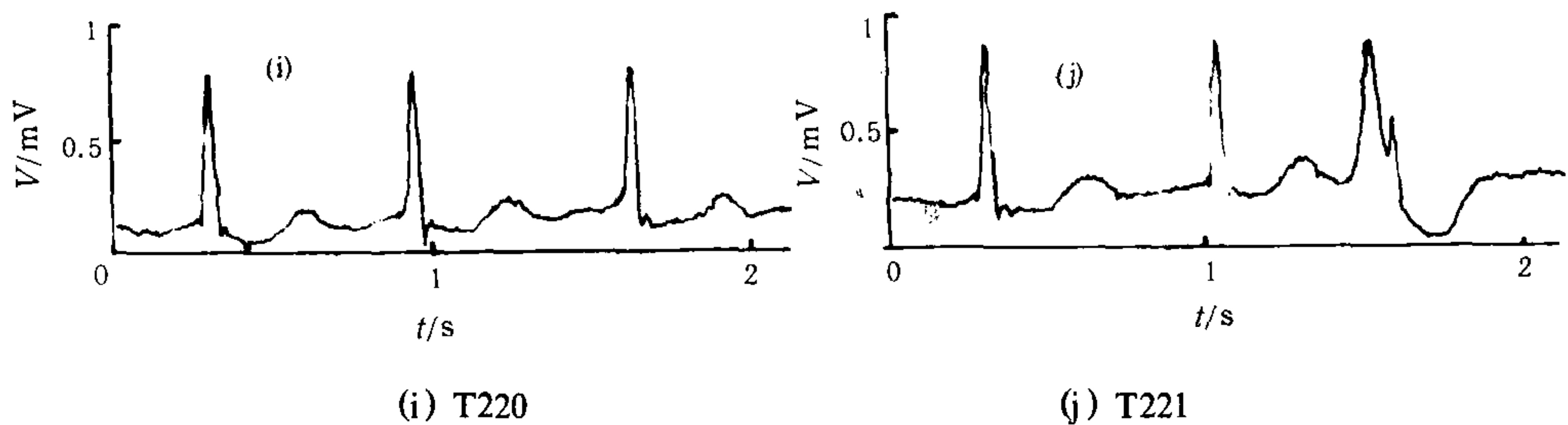


图 2 (a) — (j) 部分心电图波形。

T221 进行分类。图 2 示出了各心电图的部分波形。

作为模式分类对象的心电图波形，各心搏取 R 波顶点之前 105 点，R 波顶点之后 165 点。对这 270 点数据进行抽用，即在 R 波前后 15 点全部使用，其它部分每隔 6 点采样，其结构为每一心搏 70 个数据。因此，取神经网络的输入层神经元为 70 个，负责输入心电图上每一心搏的 70 个数据；取输出层神经元为 10 个，输出心电图上每一心搏的分类信息。

采用本文提出的算法（其中 $a=10^{-3}$ ，学习率 $b=0.1$ ，联接权限值 $W_{\delta 1}=0.0001$ ），首先从睡眠神经元中唤醒两个神经元，使之处于正常状态，用作隐含神经元，开始采用全联接方式，并用随机数生成函数初始化神经网络联接权和偏置值；然后用 10 类心电图 T100 — T221 各 4 个波形训练该神经网络，当训练周期数为 10^5 时，神经网络对于已学习过的心电图波形的正确识别率为 85%，对未学习过的心电图波形的正确识别率为 42%。如果继续训练该神经网络，由于心电图波形的正确分类率没有达到预先的要求且神经网络训练的周期数已经达到限值，所以神经网络学习算法继续从睡眠神经元中唤醒神经元投入运行，这时把神经网络的学习样本也改变为 10 类心电图 T100 — T221 各 20 个波形，神经网络经过 100 000 个训练周期的学习后，这时隐含层神经元数为 8，神经元间为部分联接。隐含层的神经元数和神经元间的联接随着学习样本的变化而动态改变。

表 1 基于三种方法生成的神经网络对 MIT 数据库中 10 类心电图的分类结果

心电数据	学习 波形数	训练 周期数	测试 波形数	正 确 识 别 率					
				本文的方法		Gutierrez 方法		基于遗忘的结构化方法	
				学习过的	未学习过的	学习过的	未学习过的	学习过的	未学习过的
T100	20	10^5	40	100%	69%	100%	60%	100%	68%
T105	20	10^5	40	100%	64%	100%	55%	100%	65%
T106	20	10^5	40	100%	66%	100%	61%	98%	45%
T108	20	10^5	40	100%	61%	100%	58%	97%	54%
T111	20	10^5	40	100%	53%	100%	47%	91%	48%
T112	20	10^5	40	100%	56%	100%	44%	93%	51%
T217	20	10^5	40	95%	43%	95%	45%	78%	46%
T219	20	10^5	40	93.4%	48%	93%	38%	89%	44%
T220	20	10^5	40	88.6%	37%	89%	35%	60%	36%
T221	20	10^5	40	87.8%	34%	87%	32%	55%	35%

为了分析比较,采用 Gutierrez^[1] 等人提出的方法,神经网络的结构为:输入神经元数为70,隐含神经元数为10,输出神经元数为10;采用基于遗忘机制的结构化算法^[4],其隐含神经元数为6,输入、输出神经元数同上.表1分别列出了基于上述三种方法设计的神经网络对 MIT 数据库中10类心电图的分类结果.其中,训练样本集为200个心电图波形;检测样本集为400个心电图波形(包括200个已学习过的波形).

从表1的实验结果可以看出,由本文提出的方法设计出的神经网络结构,无论对学习过的心电图波形,还是对未学习过的心电图波形,都有更好的分类能力.同时,神经网络的结构还可以随着外界环境的变化和神经元本身的退化以一种最佳的形式适应外界环境的变化.

5 结束语

本文在神经网络学习过程中引入了神经元兴奋、抑制和突触修改机制、退化机制、死亡机制、自修复机制等.提出了神经网络通过学习,自动生成解决某一问题的合适的神经网络结构.虽然实验中只讨论了前馈网络,事实上,其思想同样适合于其它类型的神经网络,这将是今后进一步研究的课题.

参 考 文 献

- [1] Gutierrez M, Wang J. Estimating hidden unit number for two-layer perceptrons. Proc. Int. Joint Conf. Neural Networks (Washington, DC). 1989. 677—681.
- [2] Kung S Y, Hwang J N. An algebraic projection analysis for optimal hidden units size and learning rate in back-propagation learning, Proc. IEEE Int. Conf. Neural Networks (San Diego., CA). 1988. 363—370.
- [3] Sejnowski T J, Koch C, Churchland P S. Computational neuroscience. Science. 1988. 241: 1299—1306.
- [4] Haddon S J. Comparing biases for minimal network construction with back-propagation. Advances in Neural Information Processing Systems. 1989. 1: 177—185.

A NEW METHOD ON AUTOMATICALLY BUILDING NEURAL NETWORK STRUCTURE

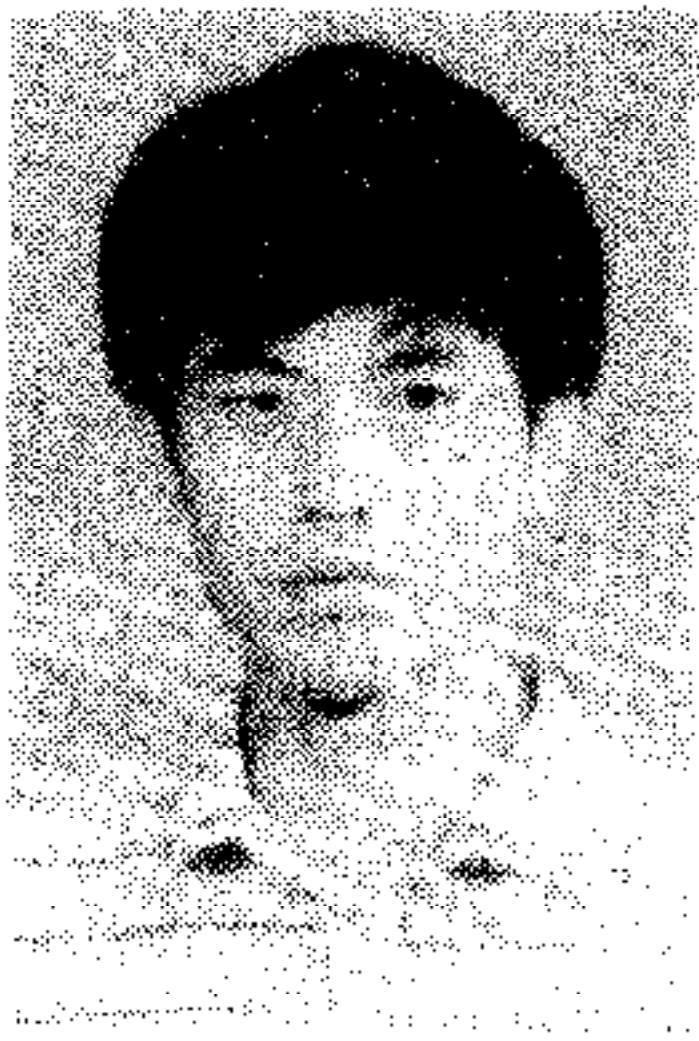
WANG JICHENG CAI YIFA LU WEIXUE

(*Biomedical Engineering Institute., Zhejiang University Hangzhou 310027*)

ABSTRACT

Neural networks are widely used in pattern recognition, optimization computation, etc. But, for a given problem, in order to get a neural network structure, designers usually use the method of trial and error. Thus the selection of neural network structure is more and more difficult with the problem more complex. In this paper, according to the research of neuron state change resulting in brain space structure and brain state change, the excitation, inhibition, synaptic change mechanism, degradation mechanism, death mechanism and auto-repairing mechanism of neuron are introduced in neural network. The suitable neural network structure for a given problem can be built through neural network learning. The experiment results have shown that the method is available and provides a new way of designing neural network structure.

Key words: Neural network., cognition science, network structure design, adaptive theory.



王继成 浙江大学生物医学工程研究所；人工智能研究所联合培养博士后。1994年6月浙江大学生物医学工程专业，获博士学位。主要研究方向包括神经网络、人工智能、认知科学、生物医学信息处理。发表论文20余篇。