



位移不变的视觉模式在线学习和识别¹⁾

沈 实 王正志 胡德文 周宗潭

(国防科技大学自动控制系 长沙 410073)

摘 要 提出了一种二维模式识别的神经网络模型 ART-Cognizer, 该模型采用新认知机的结构和自适应共振网络的自顶向下的注意和匹配机制, 因此对模式在视野中的缩放和平移有容忍能力, 并对自组织学习过程有自稳定能力, 它可以在线学习新的模式而不损伤其原有记忆.

关键词 视觉模式识别, 新认知机, 自适应共振.

1 引言

神经网络模型在模式学习与识别方面具有自组织、自适应等优点. 然而, 在基于神经网络的视觉模式学习和识别系统中, 以下两个问题尤为突出和重要: 1) 对模式在视野中的缩放、平移的容忍能力; 2) 对自组织学习的自稳定能力, 或稳定性、可塑性难题. 其中可塑性和稳定性是密切相关的特性, 如果系统具有可塑性以适应环境的变化, 则会因不断学习重复出现的类似模式而无法在熟悉的环境中达到稳定状态, 或由于学习新的模式而忘却了已学的模式; 反之, 如果系统保持稳定性, 则在设定的学习过程结束后, 系统拒绝学习任何新知识, 无法适应环境的变化.

Neocognitron(新认知机)神经网络^[1,2]能对具有较大位移、形变、缩放的模式进行正确识别, 但它的学习过程是人为结束的, 没有自稳定能力. 而基于自适应共振理论的 ART^[3]神经网络能在线地对输入模式进行自组织和自稳定的学习, 较好地解决了系统可塑性和稳定性之间的矛盾. 但 ART 模型本身不具备从原始输入模式中提取不变性特征的能力, 不能容忍模式的缩放和平移. 本文利用新认知机和 ART1模型研制了一种视觉模式学习和识别系统 ATR-Cognizer, 它能够容忍平移和缩放并具有自组织和自稳定能力.

2 ATR-Cognizer 的基本原理

ART-Cognizer 是一种结构上类似新认知机的多层神经网络, 如图1所示.

模型中 L_0 — L_4 层采用新认知机中的前五层, L_4 到 L_5 是由许多平行排列的 ART1网

1) 本文工作得到国家自然科学基金资助.

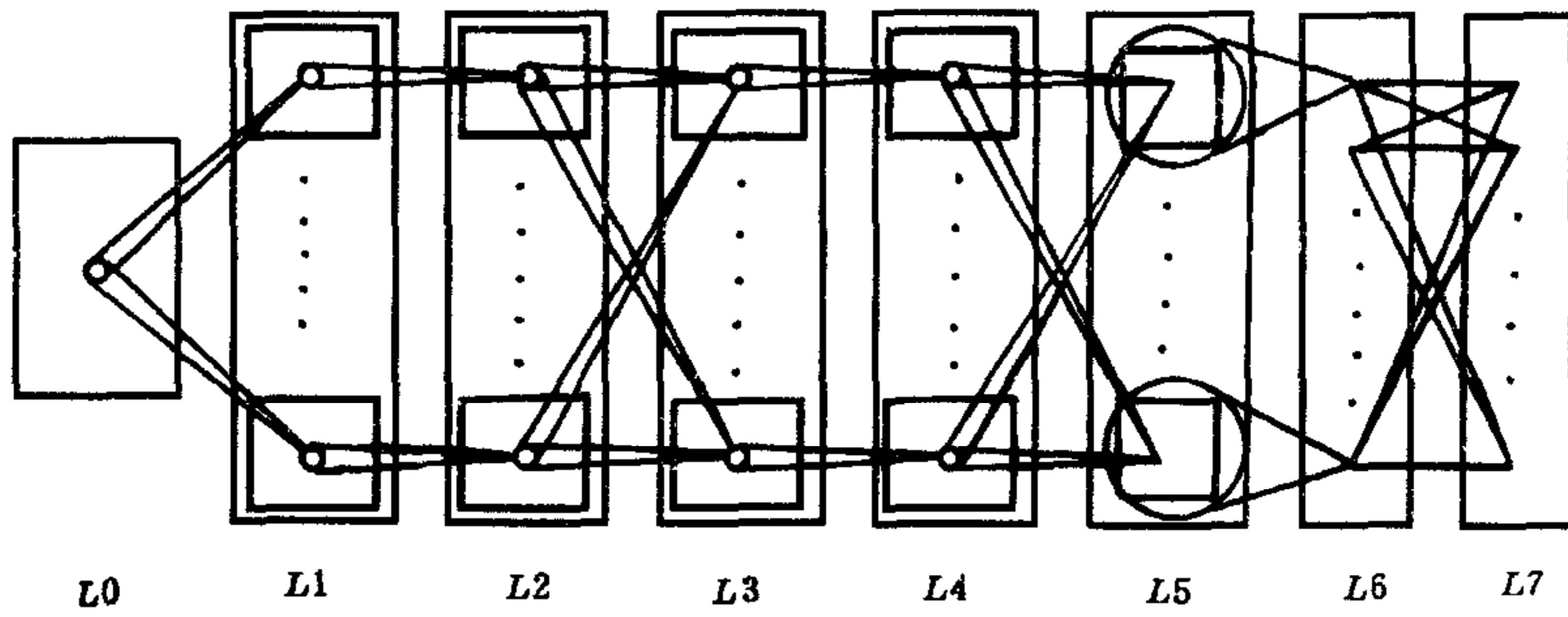


图1 ART-Cognizer 结构示意图

络阵列组成的, $L6$ 到 $L7$ 层是一个 ART1 网络. 其中 $L4$ 和 $L6$ 层分别对应于 ART1 模型中的 $F1$ 层, 而 $L5$ 和 $L7$ 层则分别对应于 ART1 中的 $F2$ 层. $L4$ 到 $L5$ 层之间的多个 ART1 网络是这样排列的: 每个 ART1 网络的输出节点是由 $L5$ 层的各细胞平面中位置相同的细胞所组成, 其输入节点是由这些细胞在 $L4$ 层的感受野中所包含的全部细胞构成的. 图 2 给出了对应于 $L5$ 层各细胞平面中第一个细胞的 ART1 网络的具体连接关系, 其中细胞间的每条连线包含着相反方向的两条连接权.

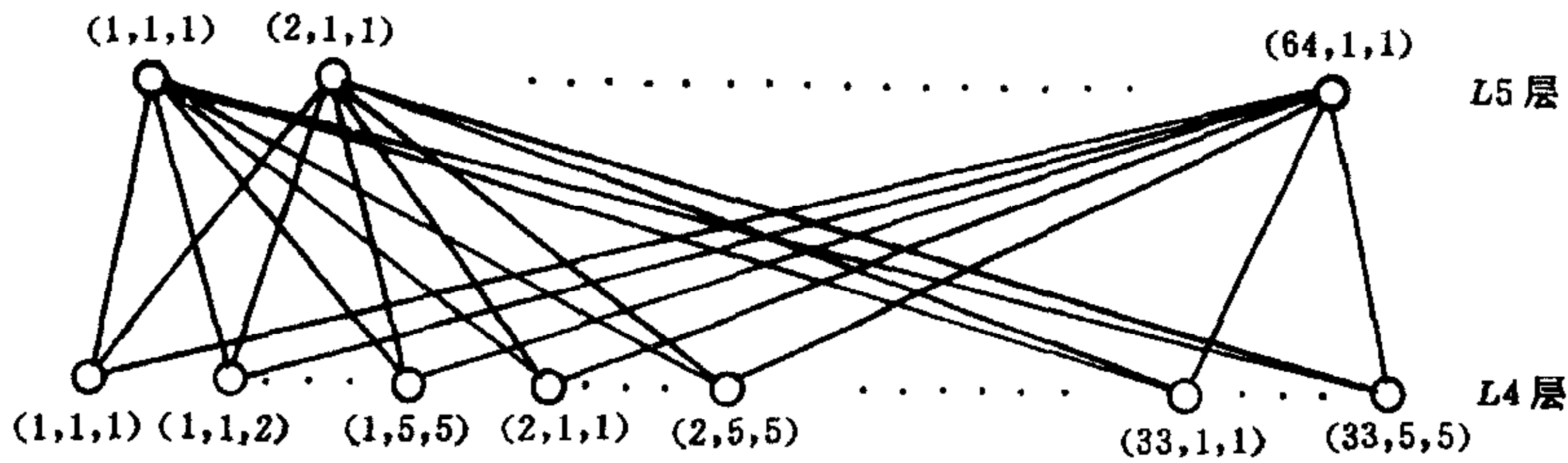


图2 $L5$ 层的一个 ART1网络的组成

在 ART-Cognizer 中, $L1$ — $L4$ 层具有基本特征抽取能力, 由于不同模式只是一些基本特征的不同空间组合, 所以 $L1$ — $L4$ 层使用一组事先确定的特征集从输入模式中抽取后继层所要用的所有必需的信息. 训练特征是一般模式中所共有的各种线段端点、拐角、交叉点等基本特征, 而复杂特征由 $L5$ 层抽取, 因此 $L1$ — $L4$ 层的训练不影响系统的在线学习能力. $L1$ 和 $L3$ 层中的 S 细胞和 V 细胞以及 $L2$ 和 $L4$ 层中的 C 细胞的输出由以下各式决定:

$$u_{sl}(k_l, n) = r_l \cdot \varphi \left[\frac{1 + \sum_{k_{l-1}=1}^{K_{l-1}} \sum_{v \in S_l} a_l(k_{l-1}, v, k_l) \cdot u_{cl-1}(k_{l-1}, n + v)}{1 + \frac{r_l}{1 + r_l} \cdot b_l(k_l) \cdot v_{cl-1}(n)} + 1 \right], \quad (1)$$

$$V_{cl-1}(n) = \sqrt{\sum_{k_{l-1}=1}^{K_{l-1}} \sum_{v \in S_l} C_{l-1}(v) \cdot u_{cl-1}^2(k_{l-1}, n + v)}, \quad (2)$$

$$u_{cl}(k_l, n) = \psi \left[\frac{1 + \sum_{v \in D_l} d_l(v) \cdot u_{sl}(k_l, n + v)}{1 + v_{sl}(n)} - 1 \right], \quad (3)$$

$$V_{sl}(n) = \frac{1}{K_l} \sum_{k_l=1}^{K_l} \sum_{v \in D_l} d_l(v) \cdot u_{sl}(k_l, n + v), \quad (4)$$

其中

$$\varphi[x] = \begin{cases} x, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0, \end{cases} \quad \psi[x] = \frac{\varphi[x]}{1 + \varphi[x]} \quad (5), (6)$$

$a_l(k_{l-1}v, k_l)$ 和 $b_l(k_l)$ 通过训练来改变, $C_{l-1}(v)$ 和 $d_l(v)$ 是固定的. 在模式学习开始前先对 $L1$ 和 $L3$ 层进行训练, $a_l(k_{l-1}, v, k_l)$ 和 $b_l(k_l)$ 按如下规则改变:

$$\Delta a_l(k_{l-1}, v, \hat{k}_l) = q_l \cdot C_{l-1}(v) \cdot u_{cl-1}(k_{l-1}, \hat{n} + v), \quad (7)$$

$$\Delta b_l(\hat{k}_l) = q_l \cdot V_{cl-1}(\hat{n}). \quad (8)$$

其中 \hat{k}_l 表示当前训练的细胞平面, \hat{n} 代表该平面中种子细胞的坐标, q_l 是常量. $a_l(k_{l-1}, v, k_l)$ 和 $b_l(k_l)$ 的初值均取为零.

$L5$ 层把 $L1$ — $L4$ 层抽取出来的基本特征组合成更加复杂的综合特征集. 系统在这一级将确定一个特征组合(对应于一个输入模式)是否已被编码到 $L5$ 层的特征集中. 如果这个特征组合未被识别, 就补充 $L5$ 层的特征集以包含这个新的特征组合. $L5$ 层中同一细胞平面里的所有细胞具有相同空间分布的输入连接, 因而同一平面中的细胞抽取不同位置的同一特征. 在 $L4$ 到 $L5$ 层采用的 ART1网络中, 每次只允许每个网络中有一个输出节点以胜者全得的方式发生共振. $L5$ 层的所有细胞为成为一个特征组合的代表而竞争, 直到每个有特征信息输入的 ART1网络都达到共振状态. 另外, 若一个输入模式中有两个或多个相同的基本特征出现在不同位置上, 每个特征必须分别地被识别和编码, 以保证得到正确的分类结果. 因此在 $L5$ 层中, 对任一输入模式, 每个平面中只允许一个细胞发生共振.

位移不变层 $L6$ 是由二态神经元组成的一维阵列, 细胞数等于 $L5$ 层中的细胞平面数. $L6$ 层的每个细胞对应于 $L5$ 层中的一个平面, 对于 $L5$ 层中包含有共振细胞的平面, $L6$ 层中对应的细胞输出为“1”, 而其它细胞输出为“0”. 同时, $L6$ 和 $L7$ 层也被作为一个 ART1网络的输入层 $F1$ 和输出层 $F2$. 不管一个输入模式是否有位移, 它对该 ART1网络的输入都是同样的, 使在输出层得到正确分类, 或在出现一个新模式时进行学习, 即更新其从底向上和从顶向下的长记忆权重. 由于在 $L4$ 到 $L5$ 层和 $L6$ 到 $L7$ 层均采用了 ART1网络, 而 $L1$ — $L4$ 层只抽取基本特征, 系统的总体行为表现出在可塑性和稳定性之间自主转换的能力.

3 仿真结果

ART-Cognizer 中所有的 ART 网络均采用快速学习方式. 由于 $L4$ 层的 C 细胞输出的是模拟量, 而随后的 ART1网络只能处理二值输入信号, 所以对 $L4$ 层的输出做如下处理:

$$f(x) = \begin{cases} 1, & \text{if } x \geq 0.58, \\ 0, & \text{if } x < 0.58. \end{cases} \quad (9)$$

在对模型的 $L1$ 和 $L3$ 层进行训练后, 让系统以学习和识别交叉进行的方式, 在线地学习了阿拉伯数字 1—9 和英文大写字母 A—Z (数字0和字母 O 视为一样). 结果表明,

该模型所学的新模式不被以后所学的模式破坏,直到输出层细胞被用完.适当选择模型中各 ART1网络的警戒参数,可使系统识别有收缩、平移、噪声和形变的模式.实验表明,系统对于学习过的模式可达97.1%的识别率,对于测试模式若其形变程度不超过 L3 层的变形训练模式,则识别率为 77.3%.当模式的缩小、噪声或形变增大时,识别率会有所下降;对无限制的手写体数字字母的识别率为 68%.所有中间层 ART1网络的警戒参数选为0.82,而输出层 ART1网络的警戒参数选为 0.92.

参 考 文 献

- [1] Fukushima K. Neocognitron: A self-organizing neural network model. *Biological Cybernetics*. 1980, **36**(4):193—202.
- [2] Fukushima K *et al.* Neocognitron: A new algorithm for pattern recognition. *Pattern Recognition*. 1982, **15**(6):455—469.
- [3] Carpenter GA *et al.* A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics and Image Processing*. 1987, **37**(1):54—115.

SHIFT INVARIANT VISUAL PATTERN LEARNING AND RECOGNITION

SHEN SHI WANG ZHENGZHI HU DEWEN ZHOU ZONGTAN

(Dept. of Automatic Control, National Univ. of Defence Technology, Changsha 410073)

Abstract This paper outlines a self-organizing neural network model, called ART-Cognizer, that automatically learns to classify visual patterns into recognition categories. It has the ability of tolerance to shift and scale variance of patterns, and the ability of self-stabilizing the self-organizing learning process, hence it can learn new patterns on-line without degrading its corpus of memories.

Key words Visual pattern recognition, neocognitron, adaptive resonance theory.