Cascade Fuzzy Adaptive Hamming Net: A Coarse-to-Fine Representation Scheme for Object Recognition*

Hong-Yuan Mark Liao[††], Hai-Lung Hung[†], Chwen-Jye Sze[‡],

Shing-Jong Lin[‡], Wei-Chung Lin[†] and Kuo-Chin Fan[‡]

† Department of Electrical Engineering and Computer Science,

Northwestern University, IL 60208

‡ Institute of Computer Science and Information Engineering,

National Central University, Chung-Li, Taiwan

†† Institute of Information Science, Academia Sinica, Taiwan

Email: liao@iis.sinica.edu.tw

Tel:+886(2)788-3799 ext. 1811, Fax:+886(2)782-4814

Keywords: neural network, unsupervised learning, self organization, hierarchical representation

⋆ A preliminary version of this technical report was presented at the **International Symposium on Artificial Neural Networks**, Hsinchu, Taiwan, Dec., 1995.

*This work was supported by the National Science Council of Taiwan under grant no. NSC 85-2213-E-001-017.

ABSTRACT

In this report, we propose a cascade fuzzy adaptive Hamming net (CFAHN) which can function as an extensible database in a model-based object recognition system. The proposed CFAHN can accept both binary and analog inputs. The architecture of a CFAHN not only preserves the prominent characteristics of the FAHN (i.e., parallel pattern matching, fast learning and stable categorization), but also extends its capability to the hierarchical class representation of input patterns. The developed CFAHN is an unsupervised learning neural network, which can be used to store new object categories in an extensible manner. Moreover, every path in the database reveals a coarse-to-fine representation of an input pattern.

1 INTRODUCTION

Unsupervised pattern classification can be defined as the categorization of input patterns into clusters or groups without any *a priori* information. Patterns categorized into a cluster are more similar to each other than to those belonging to other clusters. In recent years, neural network clustering has become a hot area of research because of its adaptivity and intrinsic parallelism. Among different unsupervised neural network models, Adaptive Resonance Theory(ART)-type networks [1–4] have several salient characteristics; i.e., (1) they can self-organize data set on line; (2) they can create new output nodes (categories) incrementally; and (3) they do not suffer from the problem of forgetting previously learned categories if the environment changes.

Usually, an unknown input scene is very complex and may exhibit a hierarchical structure [5–7]. If input data contain structural relationships, a single layer of output nodes in traditional self-organizing neural networks [1, 2, 4, 8] is definitely not sufficient to reveal their hierarchical characteristics. In order to make the scheme more flexible and powerful, Bartfai proposed a Hierarchical ART (HART) [6] that is capable of developing hierarchical class representation through self-organization of input patterns. However, his architecture learns to categorize only binary input patterns. Inspired by his work, we propose a cascade fuzzy adaptive Hamming net (CFAHN), which is able to accept both binary and analog inputs. In order to improve the learning process, we incorporate a fuzzy adaptive Hamming net into the structure of the proposed CFAHN. In the architecture of a traditional ART-type network, the learning process incorporates a search-hypothesis test cycle so that the winning node can be located and updated. If the hypothesis test fails, a reset will ensue to start another searching cycle. It is definitely true that a series of mismatched resets in response to a single input will increase the computational load and lower the efficiency of the networks. The adaptive Hamming net (AHN) [9, 10] is, thus, proposed to solve this problem. An AHN is functionally equivalent to an ART-1 network [1] except that its comparison scheme is a parallel search. Extended from an AHN, a fuzzy adaptive Hamming net (FAHN) [10] is proposed which allows analog input and is functionally equivalent to a fuzzy ART network [4]. In a CFAHN, the adaptive thresholding rule of an FAHN module is modified so that it can allow complement coding (which is an essential process for solving the proliferation problem [4]). This modification can be applied to an analog input data when it propagates between layers.

In a CFAHN, each module of an FAHN will learn the input data by either updating the matched category or by creating a new category if no existing categories can represent it. Through the connections of categories between consecutive layers, the class hierarchy of input data is formed from bottom to top (fine to coarse). The physical meaning of the hierarchical relationship is that categories in higher layers represent a more general view than do those in lower layers. Based on this kind of arrangement, a coarse-to-fine hierarchical representation scheme that can reveal the intrinsic structure of input patterns is presented. In order to make things more clear, an example is given as follows. An input set contains three events {man, woman, cat}. A coarse-to-fine representation can be learned by a 3-layered CFAHN, and the result after learning is shown in Fig. 1. In layer 1, an individual event occupies a single category. In layer 2, because of the coarser representation, a human being category and an animal category are generated. The human being category in layer 2 represents a more general view which includes both the man category and woman category in layer 1. In layer 3, only one category, the living creature, is generated to express the most general view of all events.

The organization of the rest of this report is as follows. In Section 2, the architecture and learning strategies of the proposed CFAHN are presented. In Section 3, a simulation example is illustrated. Concluding remarks for this work are given in Section 4.

2 CASCADE FUZZY ADAPTIVE HAMMING NET

In this section, we shall describe the details of a CFAHN.

2.1 Architecture

Fig. 2 shows the basic architecture of a CFAHN. In this architecture, every layer l contains a module of $FAHN_l$. A module of $FAHN_l$ (Fig. 3) is divided into two parts, $F1_l$ and $F2_l$. The process of pattern matching and category selection is performed in $F1_l$ of each layer, which is able to filter out impossible categories and pick up a best matched category from the candidates in one run. After the winning category, J^l, is generated in $F1_l$, its corresponding category in $F2_l$ is then activated to learn the input data. The learning process either updates the existing category or creates a new category in each layer. In the meantime, learning between layers (i.e., $FAHN$s) is triggered by a compression-based data-driven process [6]. That is, except for the first layer which receives original input data, every other layer receives prototypes generated from its previous layer. For example, the input to $F1_l$ with $l \geq 2$ is the output of $F1_{l-1}$. Let I^1 represent the input data to the first layer; then we have

$$I^l = I^{l-1} \wedge W^{l-1}, \quad l = 2, \cdots, L, \tag{1}$$

where L is the number of layers in a CFAHN, and the fuzzy AND operator $\wedge$ is defined by

$$(x \wedge y)_i = min(x_i, y_i). \tag{2}$$

2.2 Determination of thresholds

To extend an FAHN to the proposed multiple-layered CFAHN, one important consideration is whether complement coding is applied to the learning process or not. Complement coding is a process used to solve the proliferation problem [4] (i.e., to avoid generating too many small categories). In our design, only the first layer needs complement coding. As to the higher layers, since their inputs become hyperbox-shaped prototypes, it is unnecessary to apply complement coding again. In a CFAHN, the learning process with or without complement coding results in two different interpretations in terms of a threshold value. Basically, this threshold value plays an important role because it can control the number of categories generated in every layer. The larger the threshold is set, the greater is the number of categories generated. If complement coding is not applied, the threshold is used to determine the compression ratio of a selected category in response to an input data pattern. A category is selected as a candidate from the category database if its weights are very close to those of an input data pattern. By nature, the generated category is a maximal subset of an input data pattern. If complement coding is applied, and the input data is in analog format, then by Eq. (1) the input to higher layers (i.e. $l \geq 2$) becomes a hyperbox-shaped set. Under these circumstances, whenever a category is selected as a candidate from the category database, this means its weights can maximize a hyperbox that represents the input data pattern. The selected category in higher layers is, therefore, considered as the superset of an input data pattern. To retain the learning properties of ART-type networks, the adaptive thresholding value in multiple layers is redefined as

$$\theta_l = \begin{cases} 0 & \text{if } |I^l \wedge W^l| = |W^l|; \\ |I^l|\rho_l & \text{otherwise,} \end{cases} \tag{3}$$

where the norm $|\cdot|$ is defined by

$$|\mathbf{x}| \equiv \sum_{i=1}^{m} |x_i|.$$

ρ_l is the vigilance parameter in layer l, and its value is set to [0, 1].

2.3 Learning Algorithm

The learning algorithm of a CFAHN is described as follows:

1. **Initialization:** Determine the number of layers, L, and the vigilance parameter, ρ_l, for layer l.

2. **Input:** Present a binary or analog pattern $X = [x_1, ..., x_n]$ to the input nodes of $FAHN_1$, where x_i is in [0, 1], and n is the number of input nodes (i.e., the dimension of the input vector).

3. **Input coding:** If complement coding is applied, the input to $FAHN_1$ is extended from dimension n to m by

$$I^1 = [x_1, \cdots, x_n, x_1^c, \cdots, x_n^c], \quad \text{where } x_i^c = 1 - x_i \text{ and } m = 2n;$$

otherwise, $I^1 = X$ and $m = n$. After the number of input nodes in the first layer is determined, the number of input nodes to every upper layer is set to the same value (i.e., m).

4. **For every** $FAHN_l$ $(l : 1 \to L)$ **:**

4.1 **Pattern matching:**

4.1.1 If the category database in $FAHN_l$ is empty, Goto (4.4.1).

4.1.2 For every input I^l and category W_j^l, compute the matching score by

$$s_j^l = \sum_{i=1}^{m} min(I_i^l, w_{ji}^l).$$

Here, a linear function f_θ^l is used to filter out impossible categories if their matching scores cannot pass the vigilance test. That is,

$$f_\theta^l(s_j) = \begin{cases} 0 & \text{if } s_j^l < \theta_l, \\ s_j^l & \text{if } s_j^l \geq \theta_l. \end{cases}$$

The threshold value θ_l is determined by Eq. (3).

4.1.3 If all W_j^l are filtered out, Goto (4.4.1).

4.2 **Category selection:** Choose a pattern category according to the selection function

$$u_j^l = \frac{f_\theta^l(s_j)}{\alpha + |W^l|},$$

where $0 < \alpha \ll 1$. The category selection J^l is, thus, obtained by

$$J^l = \arg_j \max_j u_j^l, \quad j = 1, \cdots, N^l,$$

where N^l is the number of categories generated at layer l. If more than one u_j^l is maximal, the output node with the smallest index is chosen to break the tie.

Perform between-layer and within-layer learning at the same time.

4.3 **Between-layer learning:** For the higher layers in $FAHN_l$ ($l \geq 2$),

$$I^l = I^{l-1} \wedge W_J^{l-1},$$

where $l \geq 2$, and W_J^{l-1} is the weights of the activated category located in $FAHN_{l-1}$.

4.4 **Within-layer learning:**

4.4.1 If the category database is empty, or if there is no category match, a new category is generated and initialized by

$$W_J^l = I^l.$$

4.4.2 If the J^lth category in $FAHN_l$ is chosen, the weight values are updated by

$$W_J^{(new)} = W_J^{(old)} + \gamma(I^l \wedge W_J^{(old)} - W_J^{(old)}),$$

where γ is set to [0,1]. If γ is equal to 1, the learning is considered to be fast learning; otherwise, the winning weight vector is updated with a slower rate of forgetting.

5. **Goto** (2) until the network is stable; i.e., no new category in $FAHN_l$ is created, and the weights are not further changed.

2.4 Coarse-to-fine representation

Basically, the output of a CFAHN is a selected category represented in a hierarchical form. This hierarchical representation scheme represents a category with a linked path, from fine to coarse (bottom up). In what follows, we shall give an example of this category representation scheme. Given an input pattern, the pattern matching and category selection module ($F1$) of a CFAHN will try to select a best matched category located in $F2$ (Fig. 2). Since the counterparts of the input image in every layer of $F2$ will be activated, the category (or class) hierarchy will be generated. To explain this using an example, a simple data set with five patterns (10111 00101 10010 00010 01000) was fed into a 2-layered CFAHN. The network was trained until each data pattern could directly access its category in each layer. As shown in Fig. 4, in layer $FAHN_1$, five categories were generated to represent five input patterns. However, in layer $FAHN_2$, only three categories are generated. Categories in these two layers thus form a hierarchical structure by activation of input patterns. Category $C0$ in layer $FAHN_2$ represents category $C0$ and $C4$ in layer $FAHN_1$ due to the similarity of the pattern (_ 0 1 _ 1). Category $C1$ in layer $FAHN_2$ represents categories $C1$ and $C3$ in layer $FAHN_1$ due to the similarity of the pattern (_ 0 0 1 0). In addition to exhibiting the hierarchical relationships of categories, the prototypes of these categories can also be used to locate features that are not inherited from more general classes located at higher levels. For example, the pattern (_ 1 _ _ _) represented by Category $C2$ in $FAHN_1$ cannot be represented by Category $C0$ or $C1$ in $FAHN_2$.

3 NUMERICAL EXAMPLE

In the experiments, we used 340 two-dimensional feature points as a training data set to test the effectiveness of the proposed CFAHN. These data were spread into seven Gaussian distributions, and among these distributions, four overlapped to form two bigger clusters. The test data were related to object recognition by considering that features of objects were extracted by a preprocessing process and then transformed into an appropriate feature vector form. In this example, each two-dimensional data point represented a feature vector with two feature components. Fig. 5(a) shows the scatter plot of the data used in this example. The number of layers of CFAHN required for a CFAHN used in the experiment was set to be three. The vigilance parameters of CFAHN were set to be $\rho_1 = 0.8$, $\rho_2 = 0.70$ and $\rho_3 = 0.35$. Further, we applied complement coding in the learning process of the experiment. After the learning process of the network became stable (i.e., every input could access a matching category in each layer, and the weights of a selected category did not change), the cascading relationships of different categories represented by multi-layers of an $F2$ network were as shown in Fig. 5(b). Figs. 5(c), (d) and (e) show a series of generated categories in different layers of $F2$. It is obvious that categories in higher layers formed a superset of categories in lower layers (Fig. 5(f)). Therefore, categories in higher layers could express a more general view than could those in lower layers. In this way, the representation of categories was constructed through input data propagation from the bottom layer to the top layer. In layer $FAHN_2$, Category $C3$ was initialized but never activated and expanded. An additional garbage collection process is, thus, needed at the end of every learning epoch. This process is designed to delete empty categories which are never activated by any input data.

4 CONCLUDING REMARKS

We have successfully built a cascade fuzzy adaptive Hamming net. The cascade networks are able to create new categories in every layer to learn input data. The hierarchical relationships of input data are then constructed through the connections of categories in layers. The cascade fuzzy adaptive Hamming net is able to learn complex and structured input data. Thus, it can act as an extensible database in a object recognition system and provide a coarse-to-fine representation scheme of input data.

References

[1] G. A. Carpenter and S. Grossberg, "A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine", *Computer Vision, Graphics, and Image Processing*, vol. 37, pp. 54–115, 1987.

[2] G. A. Carpenter and S. Grossberg, "ART2: Self-Organizing of Stable Category Recognition Codes for Analog Input Patterns", *Applied Optics*, vol. 26, no. 23, pp. 4919–4930, Dec. 1987.

[3] G. A. Carpenter and S. Grossberg, "ART 3: Hierarchical Search Using Chemical Transmitters in Self-Organizing Pattern Recognition Architectures", *Neural Networks*, vol. 3, pp. 129–152, 1990.

[4] G. A. Carpenter, S. Grossberg, and D. B. Rosen, "Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System", *Neural Networks*, vol. 4, pp. 759–771, 1991.

[5] P. K. Simpson, "Fuzzy Min-Max Neural Networks-Part 2: Clustering", *IEEE Transactions on Fuzzy Systems*, vol. 1, no. 1, pp. 32–45, Feb. 1993.

[6] G. Bartfai, "An ART-based Modular Architecture for Learning Hierarchical Clusterings", Tech. Rep. CS-TR-95/3, Victoria University of Wellington, Feb. 1995, To appear on *Neurocomputing*.

[7] H. L. Hung and W. C. Lin, "Dynamic Hierarchical Self-organizing Neural Networks", in *Proceedings of IEEE International Conference on Neural Networks*, Orlando, FL., Jun. 28 - Jul. 2 1994, vol. II, pp. 627–632.

[8] T. Kohonen, *Self-Organization and Associative Memory*, Springer-Verlag, Berlin, 3 edition, 1989.

[9] Y. H. Pao, *Adaptive Pattern Recognition and Neural Networks*, Addison-Wesley Publishing Company, Inc., 1989.

[10] C. A. Hung and S. F. Lin, "Adaptive Hamming Net: A Fast-Learning ART 1 Model Without Searching", *Neural Networks*, vol. 8, no. 4, pp. 605–618, 1995.

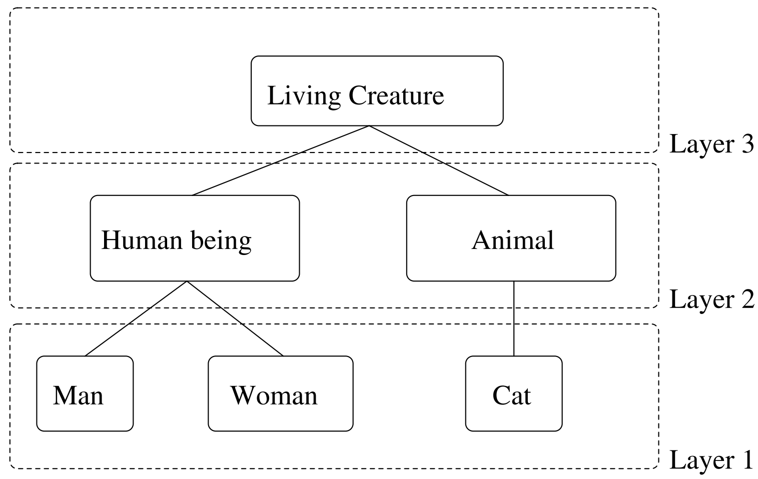


Figure 1: A coarse-to-fine representation example

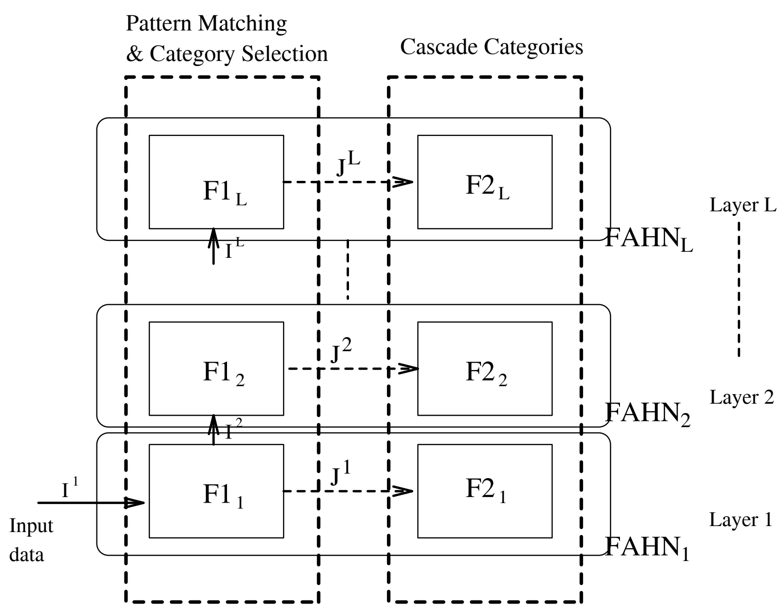


Figure 2: The architecture of a Cascade Fuzzy Adaptive Hamming Net is composed of multiple modules of a Fuzzy Adaptive Hamming Net (FAHN).

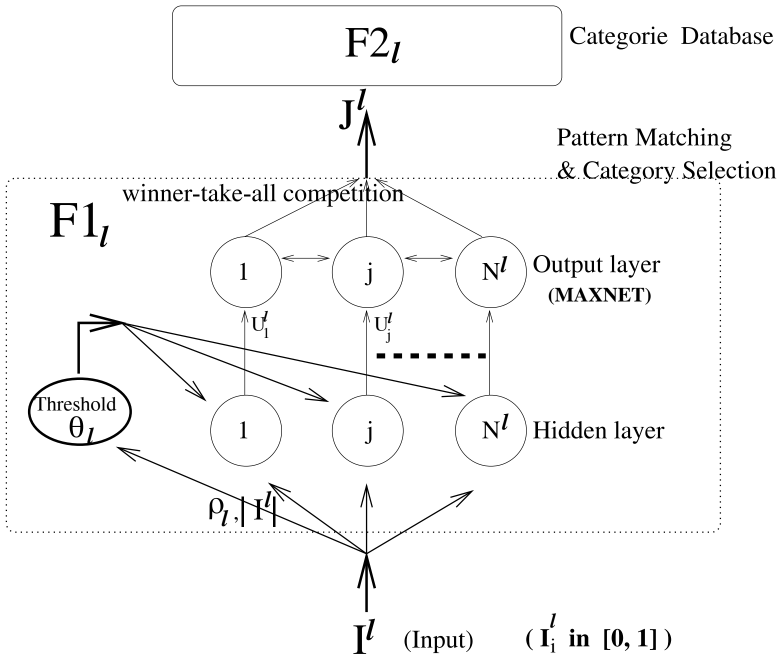


Figure 3: The architecture of a 2-layered fuzzy adaptive Hamming net.

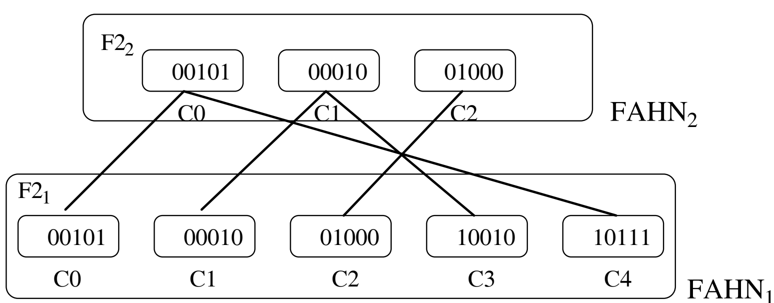
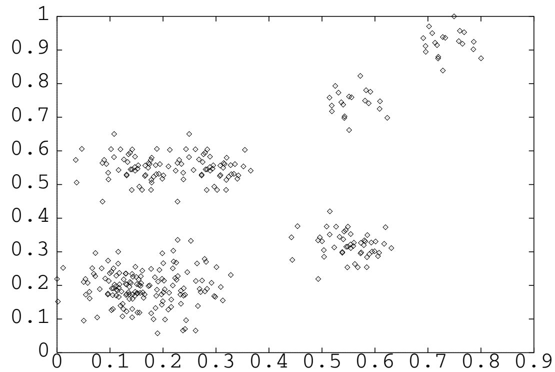
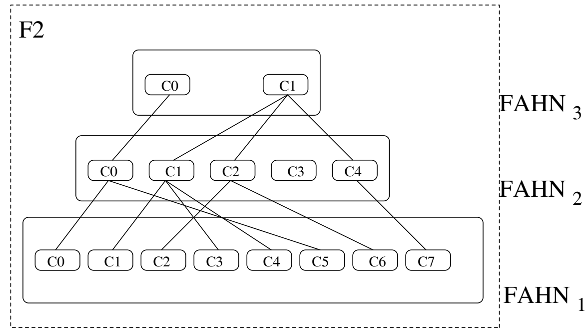


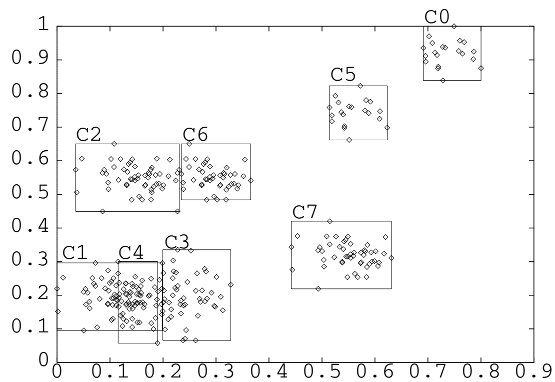
Figure 4: Hierarchical categories of a 2-layered CFAHN with five input patterns.



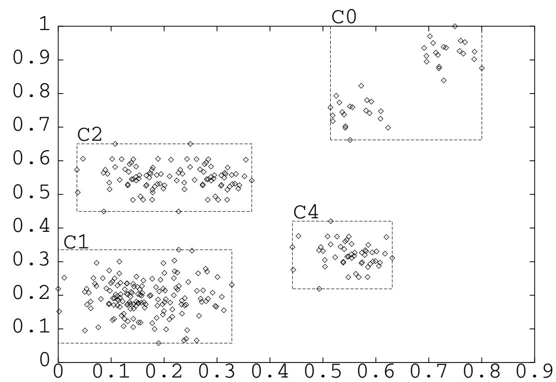
(a) 2-dimensional test data



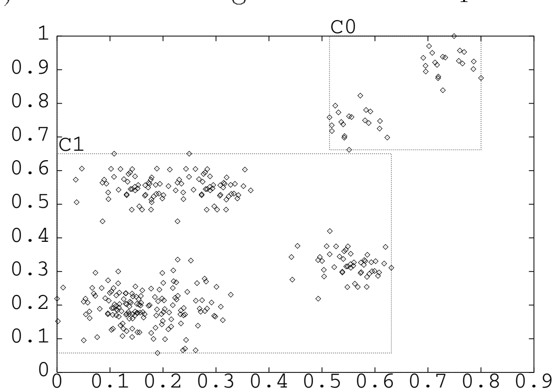
(b) Cascade relationships of variant categories represented by multi-layers of an $F2$ network.



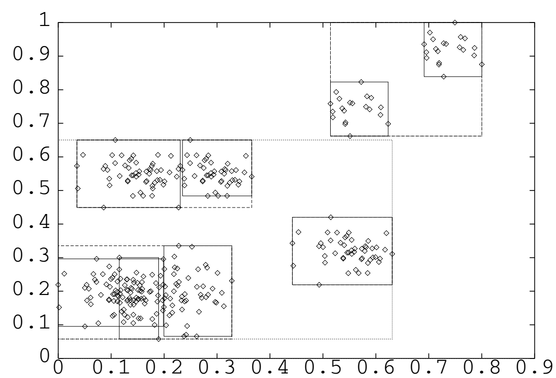
(c) Generated categories in $FAHN_1$.



(d) Generated categories in $FAHN_2$.



(e) Generated categories in $FAHN_3$.



(f) Categories in higher layers form supersets of categories in lower layers.

Figure 5: A 2-dimensional case with analog input fed into a 3-layered CFAHN ($\rho_1 = .8$, $\rho_2 = .75$, $\rho_3 = .35$).