

# 퍼지 RBF 네트워크의 학습 성능 개선

## Learning Performance Improvement of Fuzzy RBF Network

김재용\*, 김광백\*  
\*신라대학교 컴퓨터공학과

Jae-Yong Kim\* and Kwang-Baek Kim\*  
\*Dept. of Computer Engineering, Silla University  
E-mail : gbkim@silla.ac.kr

### 요 약

본 논문에서는 퍼지 RBF 네트워크의 학습 성능을 개선하기 위하여 Delta-bar-Delta 알고리즘을 적용하여 학습률을 동적으로 조정하는 개선된 퍼지 RBF 네트워크를 제안한다. 제안된 학습 알고리즘은 일반화된 델타 학습 방법에 퍼지 C-Means 알고리즘을 결합한 방법으로, 중간층의 노드를 자가 생성하고 중간층과 출력층의 학습에는 일반화된 델타 학습 방법에 Delta-bar-Delta 알고리즘을 적용하여 학습률을 동적으로 조정하여 학습 성능을 개선한다.

제안된 RBF 네트워크의 학습 성능을 평가하기 위하여 컨테이너 영상에서 추출한 40개의 식별자를 학습 데이터로 적용한 결과, 기존의 ART2 기반 RBF 네트워크와 기존의 퍼지 RBF 네트워크 보다 학습 시간이 적게 소요되고, 학습의 수렴성이 개선된 것을 확인하였다.

### 1. 서론

RBF 네트워크는 빠른 학습 시간, 일반화(generality), 단순화(simplicity)의 특징으로 학습 데이터를 분류하는 작업과 비선형 시스템 모델링 등에 적용되고 있다[1]. 기존의 오류 역전파 학습 알고리즘과 같은 다층 퍼셉트론(multilayer perceptron)의 연결 가중치들이 지역 최소화에 위치하여 학습 패턴들이 잘 분류되지 않는 경우를 방지할 수 있으며 학습 시간을 단축시킬 수 있다[2]. RBF 네트워크의 중간층은 클러스터링 하는 층이다. 즉, 이 층의 목적은 주어진 자료 집합을 유사한 클러스터들(homogeneous clusters)로 분류하는 것이다. 여기서 유사하다는 것은 입력 데이터들에 대한 특징 벡터 공간사이에서 한 클러스터 내의 벡터들 간에 거리를 측정하여 정해진 반경 내에 존재하면 같은 클러스터로 분류하고 정해진 반경 내에 존재하지 않으면 다른 클러스터로 분류한다[3]. 그러나 정해진 반경 내에서 클러스터링 하는 것은 잘못된 클러스터를 선택하는 단점을 가지게 된다. 그러므로 중간층을 결정하는 것은 RBF 네트워크의 전반적인 효율성에 큰 영향을 준다. RBF 네트워크는 학습이 완료

된 상태, 즉 연결 강도가 특정 값으로 고정된 상태에서 새로운 패턴을 학습시키게 되면 이미 설정된 모든 연결 강도에 영향을 주므로 RBF 네트워크를 다시 학습 시키는데 상당한 학습 시간이 소요되는 단점과 RBF 네트워크에 학습되지 않은 전혀 새로운 형태의 패턴들이 들어오는 경우 이전에 학습된 유사한 패턴으로 분류하여 인식률을 저하시킨다[4]. 따라서 본 논문에서는 RBF 네트워크의 학습 구조를 개선한다. 제안된 학습 구조에서 입력층과 중간층 간의 연결구조는 퍼지 C-Means의 구조와 같다. 제안된 학습 구조는 전체적으로 완전 연결 구조이나 출력층에서 목표 벡터와 출력 벡터를 비교하여 역전파시킬 때는 대표 클래스와 연결된 가중치만을 역전파시켜 연결 가중치를 조정하므로 새로운 패턴들이 이전에 학습 패턴으로 분류되는 부분을 개선한다. 그리고 중간층의 대표 클래스와 연결된 출력층의 노드만이 일반화된 델타 방법을 적용하여 지도 학습으로 수행한다. 이때에 학습 시간을 단축하기 위하여 Delta-bar-Delta 방법을 적용하여 학습률을 동적으로 조정한다. 따라서 본 논문에서 제안된 학습 방법의 학습 성능을 분석하기 위하여 컨

테이너 영상에서 추출한 식별자에 적용하여 기존의 ART2 기반 RBF 네트워크와 퍼지 RBF 네트워크 그리고 제안된 학습 방법 간의 학습 성능을 비교 분석한다.

## 2. 관련 연구

### 2.1 ART2 기반 RBF 네트워크

ART2 기반 RBF 네트워크는 중간층을 생성하는 과정에서 경계 변수의 설정에 따라 중간층의 노드 수가 결정된다. 경계 변수는 클러스터들을 분류하는 반지름 값이다. 경계 변수를 적게 설정하면 입력 패턴과 저장 패턴 사이에 약간의 차이만 있어도 서로 다른 패턴으로 분류하여 새로운 클러스터를 생성하고 경계 변수를 크게 설정하면 입력 패턴과 저장 패턴 사이에 많은 차이가 있어도 같은 패턴으로 분류한다. 따라서 경계 변수의 설정에 따라 학습과 인식 성능이 달라지는 문제점이 있다[5].

### 2.2 퍼지 C-Means 기반 RBF 네트워크

퍼지 C-Means 기반 RBF 네트워크는 중간층을 생성하기 위하여 퍼지 C-Means 알고리즘을 적용한다. 문자 인식 분야에 이 알고리즘을 적용할 경우에는 학습 시간이 많이 소요되는 단점이 있다. 일반적으로 문자 인식에 있어서 입력 패턴은 이진 패턴을 사용한다. 따라서 0과 1로 구성된 학습 패턴을 퍼지 C-Means 알고리즘을 적용할 경우에는 정확히 입력 패턴들을 분류하기 어렵고 다른 클러스터링 알고리즘에 비해 학습 시간이 많이 소요된다[6]. 따라서 본 논문에서는 퍼지 C-Means 기반 RBF 네트워크의 학습 성능을 개선하기 위하여 중간층과 출력층 사이에 연결 강도를 조정하는 과정에서 학습률을 동적으로 조정하여 학습 시간을 단축하기 위하여 Delta-Bar-Delta 알고리즘을 적용한다.

### 2.3 Delta-Bar-Delta 알고리즘

오류 역전파 알고리즘의 학습 성능 개선을 위한 방법 중에는 Jacobs 등에 의해 제안된 delta-bar-delta 알고리즘이 있다[7]. 다층 신경 회로망이 입력층  $i$ , 은닉층  $j$ , 출력층  $k$ 로 구성될 때, delta-bar-delta 알고리즘은 식(1)과 같은 연결강도의 변화를 갖는다.

$$w_{kj}(t+1) = w_{kj}(t) - \alpha_{kj}(t+1) \frac{\partial E}{\partial w_{kj}} \quad (1)$$

$$= w_{kj}(t) + \alpha_{kj}(t+1) \delta_k z_j$$

식(1)에서  $\alpha_{kj}$ 는 학습률,  $\delta_k$ 는 출력층 오차이

며,  $z_j$ 는 은닉층 뉴런의 출력이다. 각 출력층 뉴런에서 delta는 식(2)와 같으며, 각 은닉층 뉴런에서 delta는 식(3)과 같다.

$$\Delta_{kj} = \frac{\partial E}{\partial w_{kj}} = -\delta_k z_j \quad (2)$$

$$\Delta_{ji} = \frac{\partial E}{\partial w_{ji}} = -\delta_j x_i \quad (3)$$

delta-bar-delta 규칙은 각 출력 뉴런에 대하여 현재의 정보를 식(4)와 같이 delta-bar를 형성하기 위하여 파생된 과거의 정보를 조합하여 이용한다.

$$\overline{\Delta}_{kj}(t) = (1-\beta)\Delta_{kj}(t) + \beta\overline{\Delta}_{kj}(t-1) \quad (4)$$

식(4)에서 파라미터  $\beta$ 의 값은  $0 < \beta < 1$  사이의 고정된 상수이다. delta와 delta-bar 변화에 대한 학습률의 변화는 다음과 같다. 만약 연결강도의 변화가 연속된 학습과정에서 같은 방향으로 이루어진다면 학습률은 증가한다.  $\overline{\Delta}_{kj}(t-1)$ 과  $\Delta_{kj}(t)$ 가 같은 부호에서 발생하면, 이때 학습률은 증가해야 한다. 또한 만약  $\overline{\Delta}_{kj}(t-1)$ 과  $\Delta_{kj}(t)$ 의 부호가 반대 부호를 가지면 학습률은 현재값의  $(1-\gamma)$ 의 비율로 감소한다. 이러한 가변 학습률은 각 층에 대하여 다음과 같은 식으로 표현할 수 있다.

$$\alpha_{kj}(t+1) = \alpha_{kj}(t) + k \quad \text{if } \overline{\Delta}_{kj}(t-1) \cdot \Delta_{kj}(t) > 0$$

$$= (1-\gamma)\alpha_{kj}(t) \quad \text{if } \overline{\Delta}_{kj}(t-1) \cdot \Delta_{kj}(t) < 0$$

$$= \alpha_{kj}(t) \quad \text{if } \overline{\Delta}_{kj}(t-1) \cdot \Delta_{kj}(t) = 0$$

(5)

$$\alpha_{ji}(t+1) = \alpha_{ji}(t) + k \quad \text{if } \overline{\Delta}_{ji}(t-1) \cdot \Delta_{ji}(t) > 0$$

$$= (1-\gamma)\alpha_{ji}(t) \quad \text{if } \overline{\Delta}_{ji}(t-1) \cdot \Delta_{ji}(t) < 0$$

$$= \alpha_{ji}(t) \quad \text{if } \overline{\Delta}_{ji}(t-1) \cdot \Delta_{ji}(t) = 0 \quad (6)$$

## 3. 개선된 퍼지 RBF 네트워크

RBF 네트워크의 중간층은 학습 패턴들을 클러스터링하는 층이다. 이 중간층의 목적은 주어진 학습 패턴들을 유사한 클러스터들로 분류하는 것이다. 여기서 유사하다는 것은 학습 패턴들에 대한 특징 벡터 공간 사이에서 한 클러스터내의 벡터들 간에 거리를 측정하여 정해진 반경 내에 존재하면 같은 클러스터에 속하고 정해진 반경 내에 존재하지 않으면 다른 클러스터에 속한다. 정해진 반경 내에서 클러스터링하는 것은 잘못된 클러스터를 선택하여 출력층의 입력값으로 구성하므로 RBF 네트워크의 학습 성능을 저하시킨다. 또한 학습 과정에서 중간층의 노드는 그 노드의 목표 벡터를 알 수 없기 때문에 출력층 노드의 오차를 역으로 중간층 노드들로 책임 분담(credit assignment)하는

방법을 사용하여 정체현상이 일어나므로 학습 시간이 많이 소요된다. 따라서 본 논문에서는 중간층에 퍼지 C-Means 알고리즘을 적용하여 소속도가 가장 큰 노드를 승자 노드로 선택하여 출력층으로 전달하는 학습 구조를 제안한다. 그리고 중간층과 출력층 사이의 학습으로는 일반화된 델타 학습 방법을 적용하고 학습 시간을 개선하기 위하여 Delta-Bar-Delta 알고리즘을 적용한다. 제안된 학습 모형은 그림 1과 같다.

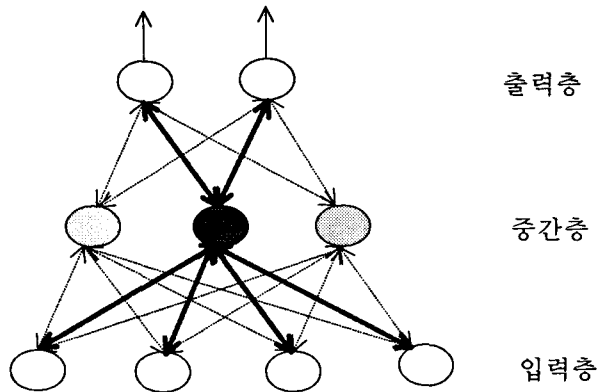


그림 1. 제안된 학습 모형

제안된 퍼지 RBF 네트워크의 학습 방법은 다음과 같이 요약할 수 있다.

1. 입력층과 중간층의 연결 구조는 퍼지 C-Means 알고리즘과 같고, 퍼지 C-Means 알고리즘의 출력층이 제안된 학습 구조의 중간층이다.
2. 중간층의 노드는 각 클래스를 의미한다. 따라서 전체적으로는 완전 연결 구조이나 목표 벡터와 실제 출력 벡터와 비교하여 역전파시킬 때는 대표 클래스와 연결된 가중치만 역전파하는 승자 노드 방식을 채택한다.
3. 중간층의 승자노드는 퍼지 C-Means 알고리즘에서 소속도가 가장 큰 중간층의 노드를 승자 노드로 선정한다.
4. 중간층과 출력층 사이의 학습 구조로는 일반화된 델타 학습 방법을 적용하여 지도 학습으로 동작한다.

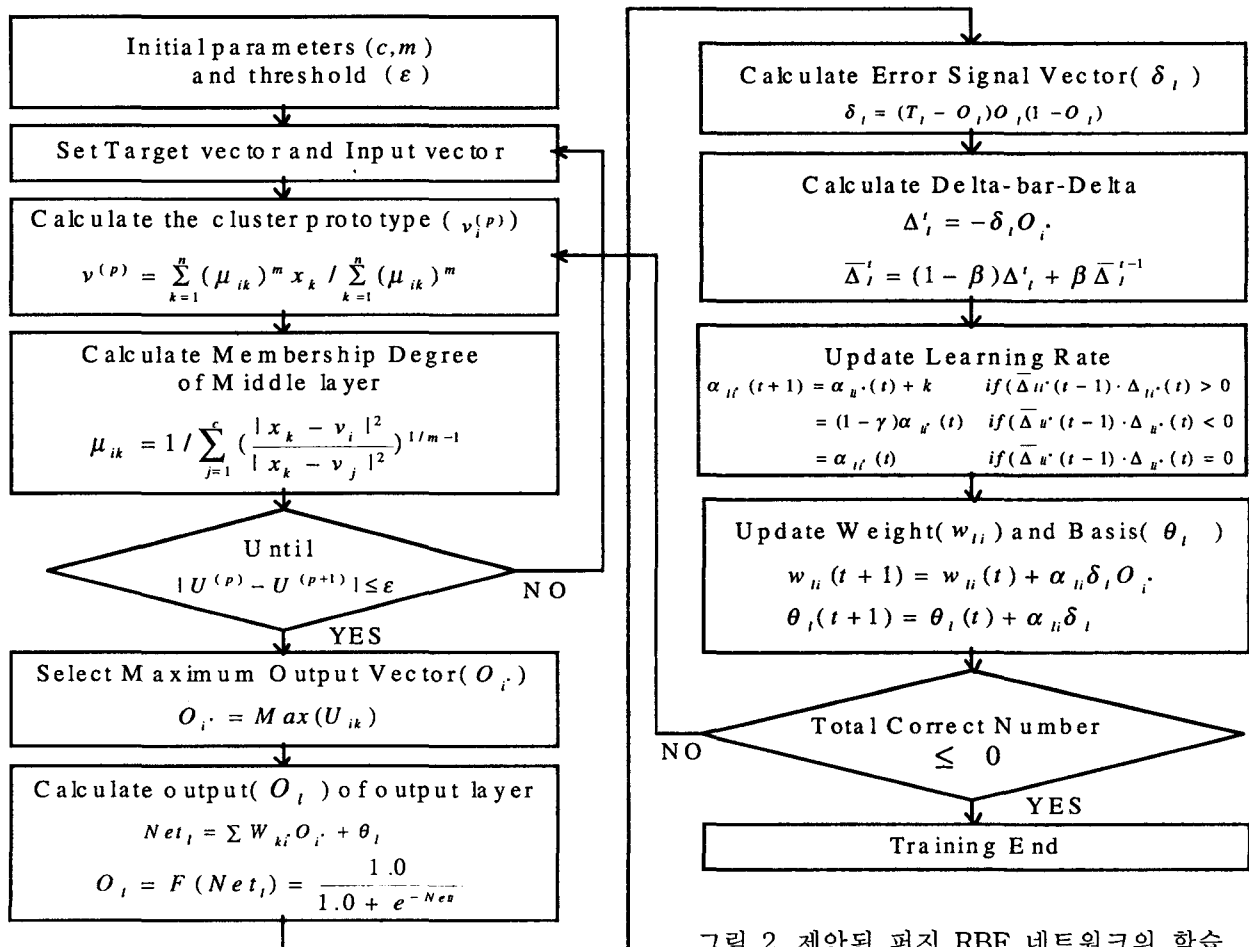


그림 2. 제안된 퍼지 RBF 네트워크의 학습 알고리즘

5. 학습 성능을 개선하기 위하여 일반화된 델타 학습에 Delta-bar-Delta 알고리즘을 적용하여 학습률을 동적으로 조정하는데 목표 벡터와 출력 벡터의 차이가 0.1 보다 적으면 정확성으로 정의하고 크면 부정확성으로 정의하여 총 패턴의 정확성의 수가 부정확성의 수보다 크거나 같은 경우에만 Delta-bar-Delta 알고리즘을 적용한다. 그 이유는 학습 과정에서 경쟁 단계로 인하여 조기 포화 현상(early saturation)이 발생하여 오류가 거의 일정하게 유지되면서 학습이 진전되지 않거나 진동하는 부분을 방지하기 위해서이다. 따라서 제안된 퍼지 RBF 네트워크의 학습 알고리즘은 그림 2와 같다.

4. 실험 및 결과 분석

제안된 학습 알고리즘의 학습 성능을 분석하기 위하여 Intel Pentium-IV 2GHz CPU와 512MB RAM이 장착된 IBM호환 PC상에서 비주얼 C++ 6.0으로 구현하였다. 제안된 학습 알고리즘의 성능을 평가하기 위하여 컨테이너 영상에서 추출한 40개의 숫자 식별자 분류 문제에 적용하여 학습 반복 횟수와 인식 개수를 ART2 기반 RBF 네트워크, 퍼지 RBF 네트워크와 비교 분석하였다.

컨테이너 식별자를 추출하기 위해 [8]의 방법을 적용하여 개별 숫자 식별자를 추출하였다. 먼저 컨테이너 영상에서 Canny 마스크를 적용하여 에지를 검출하였다. 검출된 에지 정보에서 영상 획득 시 외부 광원에 의해 수직으로 길게 발생하는 잡음들을 제거하기 위하여 퍼지 추론을 적용하여 잡음을 제거한 후에 컨테이너 식별자 영역을 추출 및 이진화 하였다. 이진화된 식별자 영역을 윤곽선 추적 알고리즘을 적용하여 개별 식별자를 추출하였다.

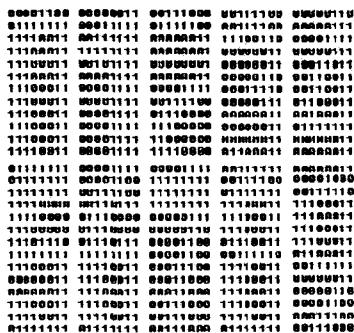


그림 3. 컨테이너 숫자 식별자 학습 패턴의 예

컨테이너 영상에서 추출 및 정규화된 패턴들 중에서 40개의 숫자 패턴을 선택하여 학습 패

턴으로 구성하였고 학습 패턴의 예는 그림 3과 같다.

숫자 식별자 분류 실험에 사용된 제안된 학습방법의 파라미터는 표 1과 같다.

표 1. 퍼지 RBF 네트워크의 파라미터

학습방법	$\alpha$	$\mu$	$k$	$\gamma$	$\beta$
제안된 퍼지 RBF 네트워크	0.9	0.5	0.05	0.2	0.7

여기서  $\epsilon$ 는 0.1로 설정하였다. 표 1에서  $\alpha$ 는 학습률,  $\mu$ 는 모멘텀 계수,  $k$ ,  $\gamma$ ,  $\beta$ 는 delta-bar-상수이다.

표 2. 각 학습 방법에 따른 학습 결과

평가방법	중간층 노드수	Epoch 수	TSS	인식수
ART2 기반 RBF 네트워크	13	950	0.067065	40
퍼지 RBF 네트워크	10	822	0.072591	40
제안된 퍼지 RBF 네트워크	10	526	0.065005	40

표 2에서 Epoch수는 학습 반복 횟수이고, TSS는 총 오차자승 합이다. 그리고 ART2기반 RBF 네트워크와 퍼지 RBF 네트워크에서 학습률과 모멘텀을 각각 0.5와 0.6으로 설정하였고 제안된 퍼지 RBF 네트워크에서는 초기 학습률을 0.5로 설정하여 총 패턴의 정확성의 수가 부정확성의 수보다 크거나 같으면 Delta-bar-Delta 알고리즘을 적용하였다. 그리고 모멘텀은 0.6으로 설정하였다. 표 2에서 학습 판정은 총 패턴의 부정확성의 수가 0보다 적거나 같은 경우에만 학습을 종료하였다. 표 2에서 알 수 있듯이 제안된 퍼지 RBF 네트워크가 ART2 기반 RBF 네트워크와 기존의 퍼지 RBF 네트워크 보다 학습 속도가 개선되었다. ART2 기반 RBF 네트워크는 ART2의 단점인 경계변수 설정에 따라 중간층의 노드의 수가 증가하거나 감소하였다. 따라서 표 2에서는 경계 변수를 0.15로 설정하는 것이 최적으로 나타났다. 퍼지 RBF 네트워크와 제안된 퍼지 RBF 네트워크에서는 퍼지 C-Means 알고리즘을 중간층으로 적용하므로 중간층의 노드의 소속도에 따라서 클

러스터를 생성하므로 ART2 기반 RBF 네트워크보다 중간층의 노드가 적게 생성되었다. 그러나 학습 패턴이 이진 패턴으로 구성되므로 퍼지 RBF 네트워크에서는 중간층의 노드를 클러스터링 하는 과정에서 학습 시간이 비교적 많이 필요하였다. 제안된 퍼지 RBF 네트워크에서는 학습 과정에서 경쟁 단계에 해당하는 조기 포화 현상을 줄이기 위하여 학습률을 동적으로 조정하므로 기존의 퍼지 RBF 네트워크 보다 학습 시간이 적게 소요되었다. 그림 4는 기존의 방법들과 제안된 방법들 간의 오차 제곱 합 곡선을 나타내었다.

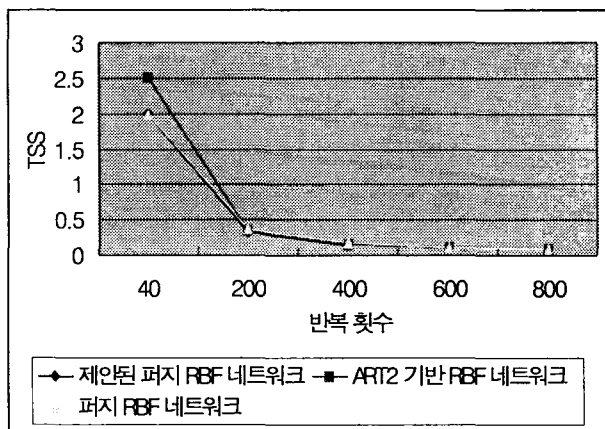


그림 4. 학습 방법에 따른 TSS의 변화율

그림 4에서와 같이 제안된 방법이 기존의 방법들보다 초기의 수렴 속도가 빠르고 학습 시간이 적게 소요되는 것을 확인할 수 있다.

### 5. 결론

본 논문에서는 퍼지 RBF 네트워크의 학습 성능을 개선하기 위하여 Delta-bar-Delta 알고리즘을 적용하여 학습률을 동적으로 조정하는 개선된 퍼지 RBF 네트워크를 제안하였다. 또한 RBF 네트워크의 학습 구조도 개선하였다. 제안된 학습 구조에서 입력층과 중간층 간의 연결 구조는 퍼지 C-Means 알고리즘을 적용하였다. 제안된 학습 구조는 전체적으로 완전 연결 구조이나 출력층에서 목표 벡터와 출력 벡터를 비교하여 역전파시킬때는 대표 클래스와 연결된 가중치만을 역전파시켜 연결 가중치를 조정하므로 새로운 패턴들이 이전에 학습된 패턴으로 분류되는 부분이 개선되었고 출력층 노드들의 오차들 역으로 중간층 노드들로 책임 분담하는 방법을 사용하여 정체현상이 일어나는 부분을 개선하였다. 그리고 중간층의 대표 클래스

와 연결된 출력층의 노드만을 일반화된 델타 방법에 적용하였다. 또한 학습 시간을 단축하기 위하여 Delta-bar-Delta 방법을 적용하여 학습률을 동적으로 조정하였다. 본 논문에서 제안된 학습 방법의 학습 성능을 분석하기 위하여 컨테이너 영상에서 추출한 식별자 분류에 적용한 결과, 기존의 ART2 기반 RBF 네트워크와 퍼지 RBF 네트워크보다 학습 성능이 개선됨을 확인하였다.

### 참고문헌

- [1] M. L. Kothari, S. Madnani and R. Segal, "Orthogonal Least Square Learning Algorithm Based Radial Basis Function Network Adaptive Power System Stabilizer." Proceeding of IEEE SMC, Vol.1, pp.542-547, 1997.
- [2] 황영섭, 방승양, "RBF 신경망 분류기의 효율적 구성 방법." 한국정보과학회논문지, Vol. 21, No. 5, pp.451-460, 1997.
- [3] C. Panchapakesan, D. Ralph and M. Palaniswami, "Effects of Moving the Centers in an RBF Network." Proceedings of IJCNN, Vol. 2, pp.1256-1260, 1998.
- [4] K. B. Kim, S. W. Jang and C. K. Kim, "Recognition of Car License Plate by Using Dynamical Thresholding Method and Enhanced Neural Networks," Lecture Notes in Computer Science, LNCS 2756, pp. 309-319, 2003.
- [5] K. B. Kim, Y. H. Joo, J. H. Cho, "An Enhanced Fuzzy Neural Network," Lecture Notes in Computer Science, LNCS 3320, pp.176-179, 2004.
- [6] 김광백, 윤홍원, 노영욱, "퀄러 정보와 퍼지 C-Means 알고리즘을 이용한 주차관리시스템 개발." 한국지능정보시스템학회 논문지, 제 8권, 1호, pp.87-102, 200.
- [7] R. A. Jacobs, "Increased rates of convergence through learning rate adaptation," IEEE Transactions on Neural Networks, Vol. 1, no. 4, pp.295-308, 1988.
- [8] K. B. Kim, "Recognition of Identifiers from Shipping Container Images using Fuzzy Binarization and Neural Network with Enhanced Learning Algorithm," Applied Computational Intelligence, World Scientific, pp.215-221, 2004.