

은닉층 뉴우런 추가에 의한 역전파 학습 알고리즘

(A Modified Error Back Propagation Algorithm Adding Neurons to Hidden Layer)

白 竣 豪*, 金 有 信**, 孫 慶 植**

(Jun Ho Baek, Yoo Sin Kim, and Kyung Sik Son)

要 約

본 논문에서는 은닉층 뉴우런 추가에 의한 새로운 역전파 알고리즘을 제안한다. 이 제안한 알고리즘에는 학습 패턴 감소 알고리즘을 이용하였다. 본 논문에서 제안된 알고리즘을 필기체 숫자 인식에 적용하여 기존 역전파 알고리즘 및 학습 패턴 감소 알고리즘과 수렴속도 및 인식률을 비교하였다. 학습 속도에 있어서는 본 논문에서 제안한 알고리즘이 학습 패턴 감소 알고리즘보다 약 2배 이상이 향상되었고, 기존의 학습 알고리즘에 비해서는 약 4배까지 향상되었다. 인식률은 기존의 학습 알고리즘의 경우 96.2%, 학습 패턴 감소 알고리즘의 경우 96.5%이며 본 논문에서 제안한 은닉층 뉴우런 추가 학습 알고리즘의 경우 97.4%이었다.

Abstract

In this paper new back propagation algorithm which adds neurons to hidden layer is proposed. this proposed algorithm is applied to the pattern recognition of written number coupled with back propagation algorithm through omitting redundant learning. Learning rate and recognition rate of the proposed algorithm are compared with those of the conventional back propagation algorithm and the back propagation through omitting redundant learning. The learning rate of proposed algorithm is 4 times as fast as the conventional back propagation algorithm and 2 times as fast as the back propagation through omitting redundant learning. The recognition rate is 96.2% in case of the conventional back propagation algorithm, 96.5% in case of the back propagation through omitting redundant learning and 97.4% in the proposed algorithm.

I. 서 론

역전파 알고리즘은 Rumelhart^[1]가 제안한 이래 여러 분야에서 사용해 왔다. 이 역전파 알고리즘은 수렴속도가 느리고, 최적의 은닉층 뉴우런수를 결정하

는 방법이 결정되지 않아 응용에 많은 한계를 지니고 있다. 이것을 해결하기 위해 학습 파라메타 변화,^[2] 오차함수의 수정을 통한 알고리즘 개선^[3] 또는 학습 데이터 중복 처리과정 생략^[4] 등을 사용하여 학습속도를 개선하고 있다. 그러나 이러한 방법은 다층 퍼셉트론의 구조는 변경시키지 않는다. 최근 Haigawara^[5]는 다층 퍼셉트론의 은닉층 뉴우런을 축소시키는 알고리즘과 그것을 통해 학습속도를 개선하는 방

*準會員, **正會員, 釜山大學校 電子工學科

(Dept. of Elec. Eng., Pusan Nat'l Univ.)

接受日字 : 1991年 12月 3日

법을 제안했다. 최적의 은닉층 뉴우런 수를 결정하기 위해 Kung et al⁽⁶⁾은 projection analysis를 통해서 학습 패턴간의 규칙성 여부에 의해 은닉층의 뉴우런 수를 결정하는 연구를 했다. 이들은 모두 처음에 많은 수의 뉴우런으로 시작했다가, 점차 잉여 뉴우런 수를 축소시켜 최적화 뉴우런의 수를 결정했다. 송현경등⁽⁷⁾은 중간층 노드의 출력값을 분석하여 중간층과 출력층 사이의 가중치 크기와 전체 학습 패턴에 대한 각 중간층 노드의 출력값 평균을 가지고 불필요한 노드를 찾아 제거에 의해 중간층 노드를 줄이는 방법을 제안하였다. 또한 최근 Hirose⁽⁸⁾에 의해 제안된 은닉층 가변을 통해 학습 속도를 개선하면서 최적의 은닉층 뉴우런 수를 줄이려는 시도가 있었다. Hirose는 이 논문에서 학습의 정체 현상을 local minima에 의한 것으로 간주하고 local minima를 제거하기 위해 학습의 정체시마다 은닉층에 뉴우런을 추가시켜 갔다. 학습이 다 이루어진 후에 은닉층의 뉴우런 수를 축소시키면서 수렴에 도달하게 하여 은닉층의 뉴우런 수를 결정하고 local minima를 벗어나게 하는 방법을 제안했다. 그러나 이 알고리즘은 단지 인쇄체 8×8 영문 알파벳과 숫자의 학습에 적용하여 학습속도를 개선하였다.

본 논문에서는 역전파 알고리즘이 갖는 중복학습을 생략하는 알고리즘을 이용하여 은닉층 뉴우런을 추가시켜 수렴에 도달하게 하는 새로운 학습 알고리즘을 제안한다. 그리고 이 알고리즘을 필기체 숫자 인식에 적용하여 기존의 역전파 알고리즘과 학습 패턴 감소 알고리즘과 비교한다. 본 논문에서 제안한 알고리즘을 은닉층의 뉴우런 수가 여러 종류일때 모두 비교하여 학습 속도가 어떻게 개선되었는지 확인했다.

II. 학습 패턴 감소 역전파 알고리즘

다층 퍼셉트론의 구조는 입력층과 출력층 그리고 그 사이에 은닉층으로 구성된다. Rumelhart가 제안한 역전파 알고리즘(backpropagation)은 가중치에 대한 오차함수의 급하강법(steepest descent method)을 사용하여 반복적 계산으로 오차를 최소화 하는 것이다. 이의 학습을 위한 역전파 알고리즘은 다음과 같다.

$$E = \sum_p E_p = \frac{1}{2} \sum_p \sum_k (t_{pk} - O_{pk})^2 \quad (1)$$

$$\Delta W_{kj} = -\eta \frac{\partial E}{\partial W_{kj}} \quad (2)$$

$$\delta_{pk} = (t_{pk} - O_{pk}) f'(net_{pk}) \quad (3)$$

$$\delta_{pj} = f'(net_{pj}) \sum_k \delta_{pk} W_{kj} \quad (4)$$

$$\Delta_p W_{kj}(n) = \eta \delta_{pk} O_{pj} + \alpha \Delta_p W_{kj}(n-1) \quad (5)$$

k:출력층의 k번째 뉴우런
j:은닉층의 j번째 뉴우런
p:p번째 패턴

각 층은 뉴우런으로 구성되고 각 층에 있는 뉴우런은 인접층과 서로 연결되어 각 연결은 가중치를 가진다. 그 구조는 다음 그림1과 같다. 여기서 t_{pk} 는 p번째 패턴에 대한 k번째 출력 뉴우런의 원하는 출력이고 o_{pk} 는 실제 나오는 출력이다.

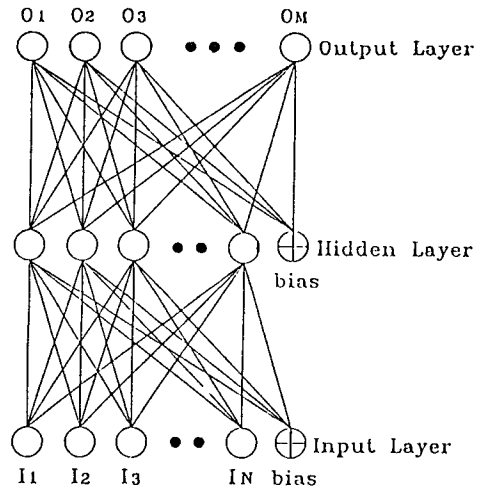


그림 1. 다층 퍼셉트론의 구조
Fig. 1. The structure of multi-layer perceptron.

학습시켜야 하는 총 패턴수를 n_p 라고 하자. 종래의 방법은 모든 학습 패턴의 오차 합계가 어떤 ϵ 보다 작을 때까지 모든 학습패턴을 매 학습때 마다 학습시키게 된다. 따라서 이미 학습이 되었다고 생각되는 것까지 중복해서 학습과정을 거치게 된다. 학습시켜야 할 모든 패턴을 한번 신경회로망에 입력시켜 출력을 계산하여 가중치를 수정하는 과정을 반복 횟수 1회라고 정의한다. 식(2)에서 계산된 각 패턴의 오차의 합을 패턴수로 나눈 것을 평균오차라 정의한다. 그 식은 다음과 같다.

$$E = \frac{1}{2} \sum_k^{n_p} (t_{pk} - o_{pk})^2 \quad (6)$$

$$\bar{E} = \frac{1}{n_p} \sum_1^{n_p} E_i \quad (7)$$

여기서 \bar{E} 는 평균오차이고 E_i 는 각 패턴의 오차이다.

학습은 윗층에서 아래층으로 오차가 δ_{pk} 에 의해 역전파하여 식(5)에 의해 가중치를 수정함으로써 이루어진다. 각 패턴을 학습시킬 때 평균오차는 주어진 기준치(ϵ)보다 작지 않지만 각각의 오차를 E_p 라 했을 때 어떤 패턴들은 ϵ 보다 작을 경우가 생긴다. 이 패턴들을 p_1, p_2, \dots, p_i 이라 하고 패턴들의 오차를 E_{p_1}, \dots, E_{p_i} 이라고 하자.

$$E_{p_1}, \dots, E_{p_i} < \epsilon_1 \quad (8)$$

이 패턴 p_1, p_2, \dots, p_i 에 대해서는 식(4)와 (5)에 따라 δ_{pk} 값이 무척 작기 때문에 식(1)에 의해 가중치 변화에 영향을 거의 미치지 않는다. 그러나 종래의 역전파 알고리즘의 경우에는 이러한 패턴들도 매번 학습시키는 불필요한 과정을 수행하여 학습속도를 느리게 한다. 학습 패턴 감소 알고리즘에서는 이러한 패턴 p_1, p_2, \dots, p_i 의 학습 과정을 생략하여 수렴속도를 개선한다. 그 알고리즘은 다음과 같다.

먼저 순방향 과정을 통하여 입력된 각 패턴의 출력을 발생시키고 각 패턴의 오차를 구한다. 그리고 각 패턴의 오차가 주어진 ϵ_1 보다 적게되면 그 패턴은 이미 학습된 것으로 생각하고 그 패턴에 대해서는 역전파를 통한 가중치 수정 과정을 생략한다. 그리고 이 패턴들에 의한 출력층의 값은 0 또는 1로 수렴한 상태이므로 식(4)와 (5)에 의해 가중치 변화에 거의 영향을 주지 않는다. 따라서 위의 생략 과정은 타당하다. 학습이 진행함에 따라 이러한 과정을 각 반복횟수마다 수행할 경우 역방향을 거치는 패턴 수는 점점 줄어들 것이다. 따라서 학습시간은 단축된다. 여기에 대한 반론이 제기될 수도 있다. 즉 패턴 p_1, p_2, \dots, p_i 이 특정한 n 번째의 반복횟수를 거칠 때 가중치의 집합 $\{W_{k1}^n, \dots, W_{kn}^n\}$ 에 의해 수렴되었다고 하더라도 다른 패턴들의 학습에 의해 수정된 가중치들에 의해서도 여전히 수렴되는지는 문제가 된다. 이 문제를 해결하기 위해서 이미 학습되었다고 여겨지는 패턴들은 반복 횟수마다 순방향 과정을 거치면서 오차를 다시 구하고 ϵ_1 보다 큰 패턴들이 생기면 이들은 다시 역방향 과정을 거치면서 가중치 수정을 행한다. 따라서 위에서 제기된 문제는 해결된다. 이를 요약하면 학습 패턴 감소 역전파 알고리즘은 T. P. Vogl et al^[2] 및 A. Krzyzak et al^[9]과는 달리 전체 오차를 이미 설명한 수렴 기준치와 비교하지 않고 평균 오차를 기준치와 비교하도록 하는 것 외에는 순방향 과정은 종래의 역전파 알고리즘과 같은 방식으로 수행되 이미 학습된 패턴은 역방향 과정을 거쳐 가중치 수정을 행하는 불필요한 과정을 생략하는 것이다. 이 학습 패턴감소 알고리즘을 학습 알고리즘 I이라 한다. 그 순서도는 그림 2와 같다.

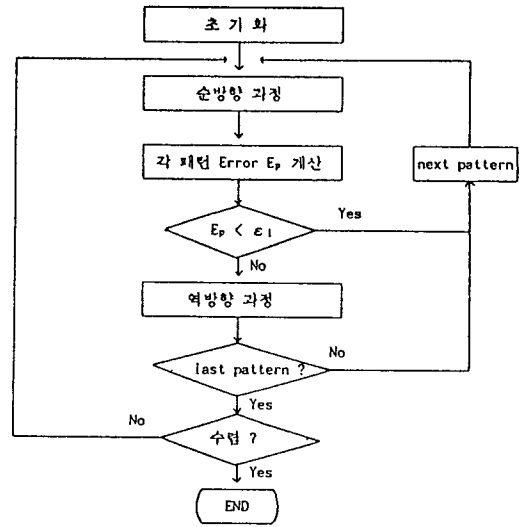


그림 2. 학습 패턴 감소 알고리즘(학습 알고리즘 1)의 순서도

Fig. 2. Flow chart of back propagation algorithm through omitting redundant learning (learning algorithm 1).

III. 은닉층 뉴우런 추가 학습 알고리즘의 제안

기존의 역전파 알고리즘에서 은닉층 뉴우런의 갯수를 보통 미리 주어지고 학습과정에서는 은닉층의 뉴우런 갯수를 변화하지 않는다. 본 논문에서 제안한 알고리즘은 2장에서 설명한 학습 패턴 감소 알고리즘을 이용하여 처음 학습을 시작할 때 은닉층의 뉴우런 수를 1개부터 시작한다. 이때 패턴들의 평균 오차가 은닉층의 뉴우런 수로는 원하는 수준보다 더 떨어지지 않아 학습을 멈추기에 충분한 평균오차 수준에 이르지 못한다. 그런데 이 평균오차가 어느 수준까지 떨어지면 오차의 변화율 $[\Delta E = E(i) - E(i-1)]$ 이 매우 작아져 주어진 기준치 ϵ_2 보다 적을 때 수렴한 것으로 볼 수 있다. 여기서 i 는 반복횟수를 나타낸다. 오차의 변화율이 미리 정한 기준값보다는 적으나 평균오차가 기준치보다 적어 학습은 되어 있지만 않았지만 학습이 더 이상 이루어지지 않을 때 은닉층의 뉴우런을 하나 더 추가시킨다. 그리고 추가된 은닉층의 뉴우런과 연결되는 모든 가중치는 적은 값으로 초기화 시킨다. 이는 새로운 은닉층 뉴우런을 추가시킴으로서 생기는 가중치를 작은 값으로 초기화하여 지금까지 해온 학습에 영향을 줄이기 위한 것이다. 그리고 같은 방법으로 전체 학습을 멈추기 위하여 오차가 미리 설정해 놓은 기준값 이하가 될 때까지 수렴시켜 나간다.

학습된 오차가 기준값 이하가 되고, 오차의 변화율 ΔE 가 매우 적어 일정한 값 이하가 되었을 때 끝나게 된다. 이러한 방법을 사용하여 학습을 시켜나가면 학습이 진행됨에 따라 은닉층의 뉴우런 수는 증가하게 되고 은닉층 뉴우런 수가 어떤 값에 이르면 학습이 끝난다. 비록 여기서 결정된 은닉층의 뉴우런 수가 최적의 뉴우런 수인가는 알 수는 없고 앞으로 이것은 해결해야 할 과제이다. 그러나 학습을 시작하기 전에 외부로부터 은닉층의 뉴우런 수를 결정해야 하는 어려움을 해결할 수 있다. 그런데 뉴우런을 추가할 때마다 연결되는 시냅스의 가중치의 초기값에 따라 학습이 끝나는데 필요한 뉴우런 수는 달라진다. 어느 것이 최적화 뉴우런 수인지 알기가 힘들다. 실험에서는 심지어 은닉층 뉴우런 수가 많은 경우에 은닉층 뉴우런 수가 적은 경우보다 더 빠른 학습속도와 높은 인식률을 보인 경우도 있다. 이때 학습률 η 는 고정하였다. 따라서 최적화 뉴우런 수를 이야기할 때 최적화의 기준도 문제가 되리라고 생각한다. 최적화 은닉층 뉴우런 수 결정은 은닉층의 뉴우런 수와 학습률 및 학습속도간에 발생하는 trade-off를 잘 처리하여야 한다. 은닉층 뉴우런의 증가를 이용해 학습을 마쳤으나 은닉층 뉴우런 축소 방법은 적용하지 못했다. 이것은 앞으로 연구해야 할 과제이다. 은닉층 뉴우런 추가 학습 알고리즘을 학습 알고리즘 2라 하고, 순서도는 그림3과 같다.

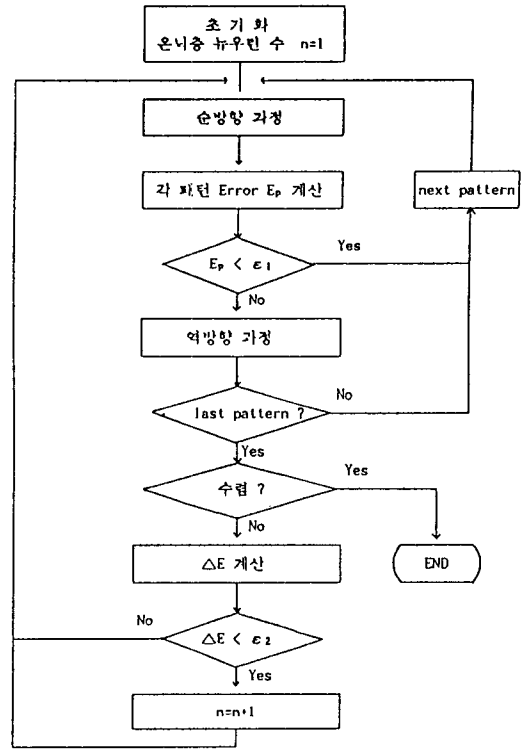


그림 3. 은닉층 뉴우런 추가 학습 알고리즘(학습 알고리즘 2)의 순서도

Fig. 3. Flow chart of back propagation algorithm adding neurons to hidden layer (learning algorithm 2).

IV. 실험결과 및 검토

본 논문에서 제안한 알고리즘의 우수성을 입증하기 위하여 0에서 9까지의 필기체 숫자 인식에 적용하여 보았다. 전산 시뮬레이션을 통해서 제안한 알고리즘과 기존의 역전파 알고리즘 및 학습 패턴감소 알고리즘을 학습속도 및 인식률 등을 비교 검토하였다.

1. 기존의 역전파 알고리즘에 의한 실험결과

기존의 은닉층 뉴우런 수를 결정하는 방법은 뉴우런 수가 서로 다른 여러가지 경우에 실험하여 그 중간가장 좋은 특성을 나타내는 것으로 결정하는 방법이다. 이러한 방법에 의한 실험의 결과가 그림 4이다. 이 실험에서 뉴우런 수는 5개에서 10까지 변화시켜 보았다. 그림4에서 은닉층의 뉴우런 수가 증가함에 따라 학습시간이 늘어나고 있다. 인식률은 은닉층 뉴우런 수가 8개일 때 가장 높다는 것을 알 수 있다. 따라서 기존의 은닉층 뉴우런 수를 결정방법으로 결정한 은닉층 뉴우런 수는 8개이다.

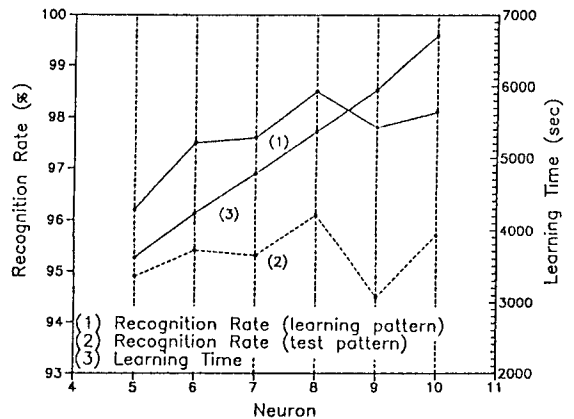


그림 4. 기존의 역전파 알고리즘의 학습 결과
Fig. 4. Results of learning of conventional back-propagation.

2. 학습패턴 감소 알고리즘에 의한 실험결과 및 검토

학습패턴 감소 알고리즘에서 각 패턴의 수렴여부를 결정하는 ϵ_1 값은 적당히 작은 값으로 선택하였다. 이는 여러번 실험을 실시한 결과 ϵ_1 이 너무 큰 경우 수렴하지 않은 패턴도 수렴한 패턴으로 판정하게 되고 ϵ_1 이 너무 작은 경우에는 수렴한 패턴도 가중치 수정을 거치게 된다. 그림5는 학습패턴 감소 알고리즘에 의한 학습 결과이다. 그림5에서 학습 패턴 감소 알고리즘은 은닉층 뉴우런 수가 8개일 때 가장 좋은 특성을 나타낸다.

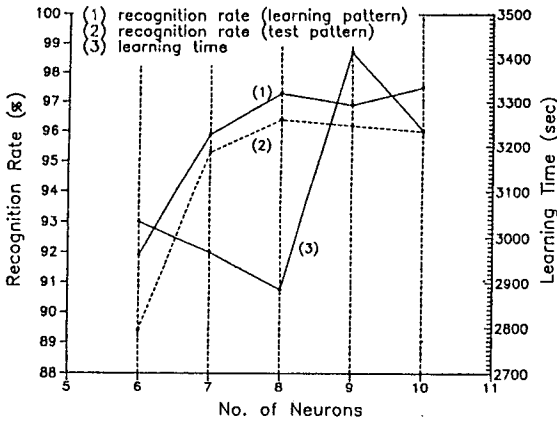


그림 5. 학습 알고리즘 1의 학습 결과
Fig. 5. Results of learning of learning algorithm 1.

그림6은 학습패턴 감소 알고리즘에 의하여 학습이 진행됨에 따라 역방향 과정을 거치는 학습 패턴의 수를 나타낸 것이다. 그림6에서 가로축은 반복횟수를 나타내며 세로축은 학습이 진행됨에 따른 학습패턴 수의 변화를 나타낸다. 그림6에서 볼 수 있는 것처럼 학습 패턴은 학습이 진행됨에 따라 감소한다는 것을 알 수 있다. 그리고 그림6에서 학습 패턴의 수가 진동하는 것을 볼 수 있다. 이는 이전 학습에는 수렴한 것으로 판정된 패턴이 다음 학습에는 수렴되지 않는 것으로 판정되는 경우이다. 이러한 현상은 이전에 학습되지 않은 패턴에 의하여 가중치의 변화가 발생하게 되고 이러한 가중치의 변화로 이미 수렴한 것으로 판정된 패턴에 대한 수렴 상태가 바뀐다. 그러나 이러한 과정이 반복됨에 의하여 모든 패턴을 동시에 만족할 수 있는 가중치를 구하게 된다.

그림7은 기존의 역전파 알고리즘과 학습 패턴감소 알고리즘의 학습시간 비교이다. 은닉층 뉴우런 수가

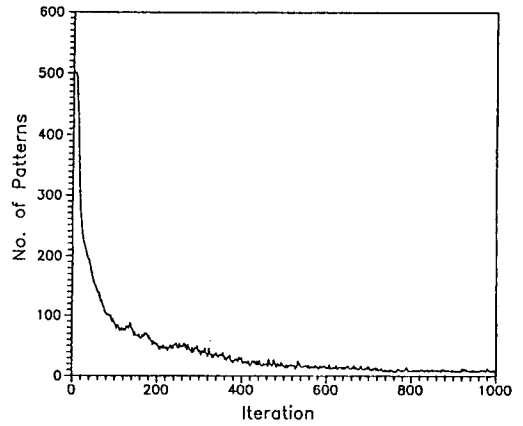


그림 6. 학습 패턴수와 반복횟수의 관계 (알고리즘 1)
Fig. 6. The relation between number of learning patterns and iteration (algorithm 1).

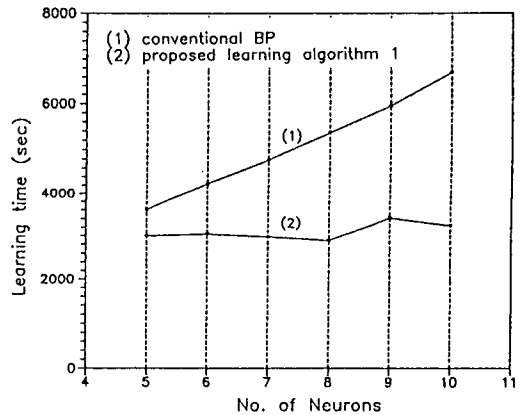


그림 7. 수렴시간 비교
Fig. 7. Comparison of learning time.

증가함에 따라 학습시간은 더욱 큰 차이를 보인다. 따라서 학습 패턴 감소 알고리즘은 학습 패턴이 많을수록 그리고 은닉층 뉴우런 수가 많을수록 더 큰 효과가 있다.

3. 제안된 학습 알고리즘에 의한 실험결과 및 검토

은닉층 뉴우런 추가 학습 알고리즘에서 ϵ_2 값은 학습이 진행되는 현재의 은닉층 뉴우런 수에서 수렴했는가를 결정하는 기준이다. 특정 은닉층 뉴우런 수에서 수렴했는가를 판정하기 위해서는 충분히 적은 값으로 선택해야 한다. 그러나 너무 작은 값을 선택했을 경우에는 특정 뉴우런 갯수에서 새로운 뉴우런

을 추가시켜야 하는 시점에 뉴우런을 추가시키지 못하고 지나치게 오랜 학습을 하는 경우가 발생한다. 본 논문에서는 실험을 통하여 적당히 선택하였다. 그리고 은닉층의 뉴우런이 새롭게 추가될 때마다 생기는 새로운 가중치는 -0.1 에서 0.1 사이엔 난수를 발생시켜 적은 값으로 초기화 시켰다. 표1은 제안한 은닉층 뉴우런 추가 학습 알고리즘에 의한 학습 결과이다. 7회의 실험결과 은닉층 뉴우런 수는 14개에서 20개까지 각각 다른 결과가 나타났다.

표 1. 제안한 학습 알고리즘 2의 학습결과
Table 1. Results of learning of proposed learning algorithm 2.

실험#	뉴우런 개수	반복횟수	평균오차	학습률(%)	수렴시간(sec)
1	18	624	0.002445	100	1819
2	14	878	0.003656	100	2799
3	17	976	0.003161	100	3589
4	20	786	0.003385	100	2198
5	17	740	0.003284	100	2339
6	19	584	0.003248	100	1917
7	18	613	0.002605	100	1778

그림8은 제안한 은닉층 뉴우런 추가 학습 알고리즘에 의한 학습과정을 나타낸 것이다. 가로축은 반복횟수를 나타내며 세로축은 평균 오차와 은닉층 뉴우런 수를 나타낸다. 실선은 오차의 변화를 나타내고 점선은 은닉층 뉴우런의 변화를 나타낸다. 여기서 학습 초기에 평균 오차가 반복횟수에 대하여 큰 변화가 없다가 갑자기 감소하는 지점이 있다. 이 때가 새로운 은닉층의 뉴우런이 추가되는 지점이다. 이것이 반복되면서 뉴우런은 첨가되고 오차의 크기는 줄어든다. 그러나 학습 중반(반복횟수 300번 정도)부터 진동이 심하게 발생되고 뉴우런 추가에 의한 효과는 무척 적어 오차가 거의 영이 될때까지 많은 수의 뉴우런이 첨가된다. 이것은 개선되어야 할 것이라고 생각된다. 앞으로 뉴우런 추가 기준을 단지 오차 변화율에만 둘 것이 아니라 보다 세련된 기준을 마련하여 추가 뉴우런 수를 줄이는 방법을 검토해야 할 것이다.

그림9는 제안한 은닉층 뉴우런 추가 학습 알고리즘의 학습이 진행됨에 따른 역방향 과정을 거치는 학습 패턴수의 변화를 나타낸 것이다. 그림9의 오차 곡선과 비슷한 형태를 가지는 것을 알 수 있다. 이는 새로운 뉴우런이 추가됨에 따라 이전까지 학습이 이루어지지 않은 패턴에 대하여 학습이 이루어진다는 것을 의미한다. 그리고 제안한 은닉층 뉴우런 추

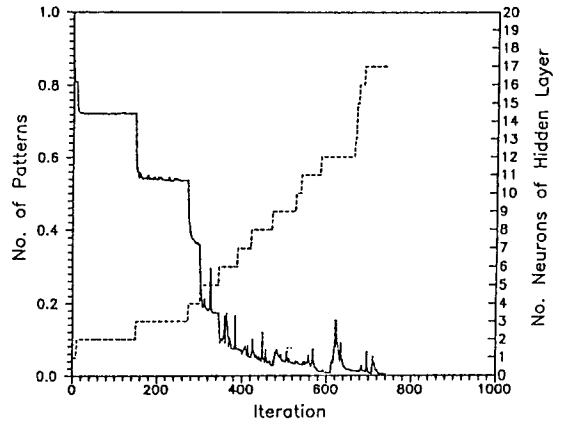


그림 8. 오차와 반복횟수와의 관계 (제안한 학습 알고리즘 2)
Fig. 8. The relation between error and iteration (proposed learning algorithm 2).

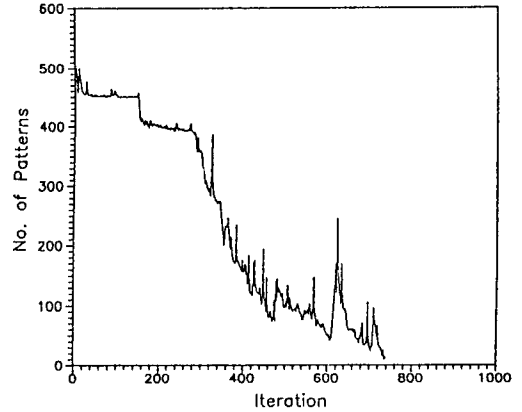


그림 9. 학습 패턴수와 반복횟수와의 관계 (제안한 학습 알고리즘 2)
Fig. 9. The relation between number of learning patterns and iteration (proposed learning algorithm 2).

가 학습 알고리즘의 경우 은닉층의 뉴우런 수가 많이 증가했음에도 불구하고 수렴속도가 빨라지는 것은 학습 초기에는 적은 은닉층의 뉴우런에 의하여 학습되고 학습이 진행됨에 따라 학습되는 패턴 수가 감소함으로써 비록 은닉층의 뉴우런이 증가하더라도 학습시간은 증가하지 않은 것으로 생각된다. 그것은 뉴우런 추가에 의해서 새로운 학습이 이루어졌을 때는 모든 시냅스의 가중치 변화가 있음을 확인한다. 그러나 이 변화는 이미 이루어진 학습이 파괴되지 않는 범위에서 가중치 변화가 일어날 것이기 때문이

다. 학습 중반부터는 그전과는 달리 진동이 심하여 뉴우런 추가 효과가 어디서 일어나는지 알기 힘들다. 그러나 곧 수렴이 되어 학습이 완료되는 것으로 보아 진동은 일시적으로 학습된 것이 파괴될 정도로 가중치의 변화가 있지만 그 변화 범위는 재학습이 단번에 이루어질 수 있는 가중치 영역에 포함된다.

그림10은 학습시간의 비교결과이다. 은닉층 뉴우런 추가 학습 알고리즘과 기존의 학습 알고리즘의 결과를 비교하면 은닉층 뉴우런 추가 학습 알고리즘의 수렴속도가 4배정도 향상되었다. 그리고 학습 패턴 감소 알고리즘의 결과와 은닉층 뉴우런 추가 학습 알고리즘의 결과를 비교해 보면 은닉층 뉴우런 추가 학습 알고리즘의 수렴속도가 2배정도 향상되었다. 이는 학습이 진행됨에 따라 은닉층의 뉴우런을 추가시켜 나가는 방법이 학습 전에 은닉층의 뉴우런 수를 고정시켜 놓고 학습을 시키는 것보다 수렴속도를 향상시킬 수 있다는 것을 나타낸다. 그리고 은닉층의 뉴우런 수를 여러가지 변화시켜 실험할 필요가 없이 학습이 진행됨에 따라 자동적으로 은닉층의 뉴우런 수를 결정하게 된다. 따라서 역전파 알고리즘의 은닉층의 뉴우런 수를 결정해야 하는 문제점을 해결할 수 있다.

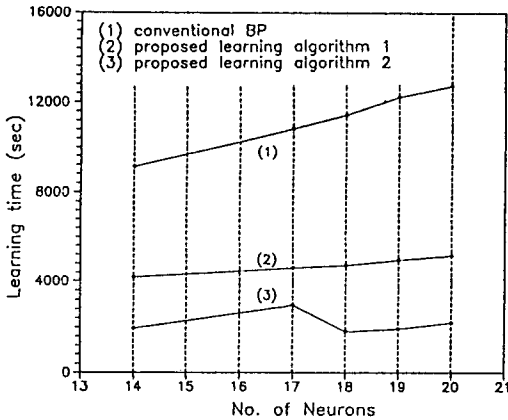


그림 10. 수렴시간의 비교
Fig. 10. Comparison of learning time.

V. 결 론

본 논문에서는 은닉층 뉴우런 추가에 의한 새로운 역전파 알고리즘을 제안한다. 이 제안한 알고리즘에는 학습 패턴 감소 알고리즘을 이용하였다. 본 논문에서 제안된 알고리즘을 기존 역전파 알고리즘 및

학습 패턴 감소 알고리즘과 수렴속도와 인식률을 비교하였다.

본 논문에서 제안한 알고리즘을 학습 패턴 감소 알고리즘과 수렴속도를 비교한 결과 약 2배이상 향상되었다. 또한 기존의 학습 알고리즘에 비해서 수렴속도는 약 4배까지 향상되었다. 인식률은 기존의 학습 알고리즘의 경우 96.2%, 학습 패턴 감소 알고리즘의 경우 96.5%이며 본 논문에서 제안한 은닉층 뉴우런 추가 학습 알고리즘의 경우 97.4%이었다. 따라서 은닉층 뉴우런 추가 학습 알고리즘의 결과가 수렴속도와 인식률면에서 가장 좋다는 것을 확인할 수 있었다. 연구결과 은닉층 뉴우런 수가 많음에도 인식률과 수렴속도가 높아지는 것이 보이는데 이것이 우연인지 아니면 수렴속도와 인식률과 은닉층 뉴우런 수 사이에 trade-off가 있는지를 알 필요가 있다. 따라서 은닉층 뉴우런 수의 최적화 기준이 무엇인지를 연구해야 한다. 앞으로 할 연구과제는 다음과 같다. 첫째 제안된 은닉층 뉴우런 수 축소방법을 적용해서 잉여 뉴우런 수를 없애는 방법을 통해 최적 뉴우런 수에 도달하는 방법을 연구한다. 둘째 뉴우런 추가 기준을 보다 세련되게 만들어 뉴우런 수를 축소할 필요없이 바로 최적의 뉴우런 수에 도달하는가를 연구해야 한다. 셋째 은닉층 뉴우런의 역할에 대한 이론적 분석을 통해 본 논문에서 제안한 알고리즘의 보편성을 입증해야 한다.

參 考 文 獻

- [1] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, "Learning internal representations by error propagation," in D.E. Rumelhart and J.L. McClelland(Eds.), *Parallel Distributed processing: Explorations in the Microstructure of Cognition*, vol. 1: Foundations, MIT Press, pp. 318-362, 1986.
- [2] T.P. Vogl, J.K. Mangis, A.K. Rigler, W.T. Zink, and D.L. Alkon, "Acceleration the convergence of the back-propagation method," *Biol. cybern.* 59, pp. 257-263, 1988.
- [3] A. Krzyzak, W. Dai and C.Y. Suen, "Classification of large set of handwritten charaters using modified back propagation model." *IJCNN'90*, vol. 3, pp. 225-232, 1990.
- [4] H. Sawai, a. Waibel, P. Haffner, M. Miyatake, K. shikano, "Parallelism, hierarchy, scaling in time-delay neural networks for spotting japanese phonemes/CV-syllables," *IEEE 3rd Int'l Conf. on Neural Network*, vol. 2, pp. 81-88, 1989.

[5] S.Y. Kung and J.N. Hwang, "An algebraic projection analysis for optimal hidden units size and learning rates in back propagation learning," *Proc. IJCNN'88*, vol. 1, pp. 363-370, 1988.

[6] Masafumi Hagiwara, "Novel back propagation algorithm for reduction of Hidden units and acceleration of convergence using artificial selection." *Proc. IJCNN'90*, vol. 1, pp. 625-

630, 1990.

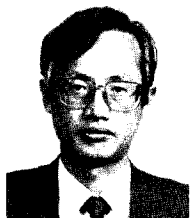
[7] 송현경, 이영지, "역전파 학습시 중간층의 역할 분석과 노드 수 축소 방법", *대한전자공학회 논문지*, 28-B, pp. 65-78, 1991, 3.

[8] Yoshio Hirose, Koichi Yamashita and Shimpei Hijiya, "Back propagation algorithm which varies the number of hidden units," *Neural Network*, vol. 4, pp. 61-66, 1991.

著 者 紹 介

白 垞 豪 (準會員)

1967年 8月 7日生. 1990年 2月 부산대학교 공과대학 전자공학과 졸업. 1992年 2月 부산대학교 전자공학과 석사과정 졸업. 1992年 2月 현재 삼성전관 근무



金 有 信 (正會員)

1951年 9月 21日生. 1974年 2月 서울대학교 공과대학 전자공학과 졸업. 1974年 3月~1978年 8月 한국원자력 연구소 근무. 1980年 6月 U. C. Berkely 전자공학과 석사과정 졸업. 1983年 6月 Stanford대학 박사과정 수료. 1983年 9월부터 부산대학교 전자공학과 재직, 현재 부산대학교 전자공학과 부교수. 주관심분야는 신경회로망 알고리즘 개발 및 집적회로 등임.



孫 慶 植 (正會員)

1950年 3月 25日生. 1973年 2月 부산대학교 전자공학과 졸업. 1975年 3月~1977年 8月 부산대학교 대학원 전자공학과 공학 석사학위 취득. 1977年 11月~1979年 2月 부산대학교 전자공학과 조교. 1979年 3月~1982年 12月, 1985年 1月~1985年 9月 부산대학교 전자공학과 전임강사. 1983年 1月~1985年 1月 미국 알라바마 주립대학 박사과정 수료. 1985年 10月~1991年 9月 부산대학교 전자공학과 조교수. 1991年 10月~현재 부산대학교 전자공학과 부교수. 1987年 3月~1991年 8月 경북대학교 대학원 전자공학과 공학박사 학위 취득. 주관심분야는 디지털 신호처리 및 신경회로망 등임.