

유전자 알고리즘을 이용한 신경 회로망 성능향상에 관한 연구

임정은*, 김해진*, 장병찬*, 서보혁**
 경북대학교 전기공학과*, 경북대학교**

A study on Performance Improvement of Neural Networks Using Genetic algorithms

Jung-Eun Lim*, Hae-Jin Kim*, Byung-chan Chang*, Bo-Hyeok Seo**
 Kyungpook Nat. Univ. Graduation School*,
 Kyungpook Nat. Univ.**

Abstract - In this paper, we propose a new architecture of Genetic Algorithms(GAs)-based Backpropagation(BP). The conventional BP does not guarantee that the BP generated through learning has the optimal network architecture. But the proposed GA-based BP enable the architecture to be a structurally more optimized network, and to be much more flexible and preferable neural network than the conventional BP. The experimental results in BP neural network optimization show that this algorithm can effectively avoid BP network converging to local optimum. It is found by comparison that the improved genetic algorithm can almost avoid the trap of local optimum and effectively improve the convergent speed.

1. 서 론

최근 인공신경망의 설계에서 많은 시도가 이루어지고 있다. 유전자 알고리즘에 의한 진화된 신경망이 연구가 되고 있는데, 유전자 알고리즘이 신경망 알고리즘에 끼치는 영향은 아주 크다.

역전파 알고리즘은 피드포워드 뉴럴 네트워크에서 잘 알려진 학습 알고리즘이다. 이 알고리즘은 기울기를 따라가는 (gradient descent)방법의 하나인데 지역 최소값(local minima)에 빠질 염려가 있는 단점이 있다. 또한 역전파 알고리즘은 학습 과정이 수렴되기까지 아주 많은 양의 학습 데이터가 필요하고 기억된 패턴의 수정이나 추가적인 학습이 불가능하다. 그렇지만 비교적 구현이 쉽고 학습이 어느 정도 가능하다는 이유로 현재 가장 많이 사용되고 있다.

한편, 유전자 알고리즘[1]은 전역해를 찾는 데 적합한 탐색 전략이다. 전역 탐색을 수행함으로써 부분극소점에 빠질 가능성이 적다. 이러한 장점 때문에 근래 주목받는 최적화 기법의 하나가 되었으며 신경망, 퍼지이론과 결합하여 적용된 다양한 사례들이 발표되고 있다. Lin[2]은 유전 알고리즘을 신경망의 연결강도를 최적화하는데 이용하였고, Homaifar[3]은 유전자 알고리즘을 이용하여 퍼지 제어기의 소속 함수와 제어규칙을 동시에 최적화 하였다.

역전파 알고리즘이 보다 강력하고 빠른 탐색 성능을 갖기 위해서는 연결강도, 오픈셋 값을 처음 학습 단계부터 적합한 값을 가지고서 신경망 학습을 시작하는 방법이 필요하다. 따라서 본 논문에서는 필기체 숫자 인식에서 수렴속도를 줄이는 방법으로서 역전파 알고리즘과 유전자 알고리즘을 이용하여 병행처리 하는 방법을 제안하였다.

2. 본 론

2.1 유전자 알고리즘과 역전파 알고리즘

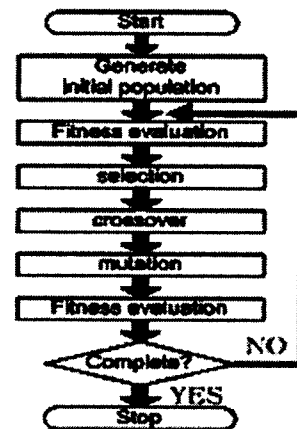
2.1.1 유전자 알고리즘

최근까지 여러 분야에서 일어나는 다양한 최적화 문제들을 해결하기 위한 많은 기법들 가운데 유전자 알고리즘은 자연 생태계의 적자생존의 원리를 모방한 이론으로서 1975년 Holland에 의해 개발되었다[1]. 유전자 알고리즘은 평가 함수(evaluation function), 유전 연산자, 문제의 해를 표현하는 염색체(chromosome)로 구성되며, 기본 연산자로는 선택(selection), 교배(crossover)와 돌연변이(mutation)등이 있다.[4]

유전자 알고리즘의 수행 과정은 다음과 같다. 우선 찾고자 하는 문제의 해를 하나의 개체(chromosome)라 하고 여러 개의 초기해로 해집단(population)을 구성한다. 해집단 내의 각 개체에 대하여 평가치를 계산하고 평가치 평가치에 따라 이들 중 일부를 다음 세대를 구성하기 위한 개체로 선택한다. 선택된 개체들에 여러 가지 유전자 연산을 적용하여 새로운 세대(generation)의 해집단을 구성한다. 이러한 일련의 과정을 반복함으로써 점진적으로 최적해에 가까운 해를 찾아가는 것이다.

다음 세대를 구성하기 위해 개체를 선택할 때에는 평가치가 높은 개체들이 높은 확률로 선택되도록 하며, 새로운 개체를 만들어 내기 위한 유전 연산으로는 복제(reproduction), 교차(crossover) 및 돌연변이(mutation)를 사용한다. 복제는 선택된 개체를 변형 없이 그대로 다음세대로 전달하는 것이며, 교차는 두 개체의 일부를 임의로 선택된 교차점을 중심으로 교환하는 연산이고, 변이는 한 개체내의 임의의 유전자를 변

경시켜서 새로운 유전자를 만들어 내는 연산이다. 이러한 연산자들을 반복 시행하면서 평가치가 높은 개체들을 선택해나간다면 세대가 진행될수록 평가치가 높은 해집단을 얻을 수 있게 된다. 해집단 내에 일정한 점수 이상의 평가치를 갖는 개체가 생성되거나, 또는 일정한 세대가 진행되어도 더 나은 개체가 생성되지 않으면 알고리즘을 종료한다



<그림 1> 유전자 알고리즘의 처리

2.1.2 역전파 알고리즘

역전파 알고리즘[5][6]은 다층 구조의 신경망을 학습시키는데 가장 널리 사용되고 있는 기법이다. 역전파 알고리즘은 교사학습(supervised learning) 방법에 의해 학습 자료들이 주어지면 원하는 값과 실제 출력된 값 사이의 차이를 줄여나가는 데 학습 규칙은 학습을 진행하는 다계층 전방향(multilayer feed-forward)의 네트워크로서 일반화된 delta rule의 법칙으로 학습을 진행한다. 특히 계층 구조를 갖는 신경망의 학습에 이용되고 있다.

역전파 알고리즘은 두 가지의 단계로 나뉘는데, 첫 번째 단계는 전향 단계(Forward phase)이다. 전향 단계에서는 각 뉴런에 입력이 들어와서 망으로 전파되고, 출력 값이 계산된다. 각 뉴런에 대해서 모든 현재의 출력들은 첫 번째 단계와 비교되고, 그 차이인 오차 값이 계산된다. 두 번째 단계는 역전파 단계에서 얻은 오차 값을 초기의 입력 값으로 역전파한다. 이 두 단계는 시스템이 안정되어 오류합이 오류 기준치를 넘지 않을 때까지 계속 반복 실행에 의해 최적의 해를 탐색하게 된다.

목적치 T_j 와 계산된 출력 결과 O_j 로 오차가 계산된다.

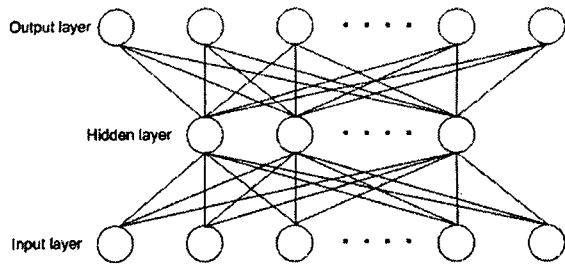
$$E = \frac{1}{2} \sum_j (T_j - O_j)^2 \quad (1)$$

후에 오차가 후방으로 전파되고 결과에 따라 가중치가 갱신된다.

$$\Delta w_i = [2K | X |^2] \frac{E_{xi}}{|X|^2} \quad (2)$$

$$\eta = 2K | X |^2 \quad (3)$$

Δw_i 은 가중치 변화량이고 η 는 학습률이다.



〈그림 2〉 역전파 알고리즘의 구조

2.2 제안된 방법의 적용

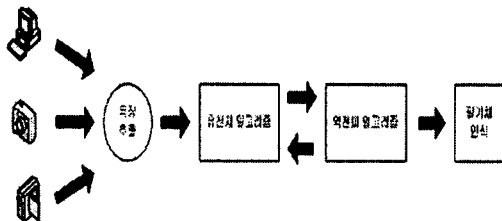
유전자 알고리즘과 역전파 알고리즘의 단점을 보완하기 위해서 두 알고리즘을 병합한 모델이 제안되었다. Kitano[7]는 다층 신경망의 전체 자승 에러(total square error)가 임계값 이하로 감소하면 신경망 학습이 전역적 최적해에 근접하게 수렴한 것으로 가정하여 에러가 임계값에 도달하는 전역적 최적해 근처까지는 유전자 알고리즘을 사용하여 수렴하고, 그 결과인 가장 높은 적합도의 해를 역전파 알고리즘 파라미터의 초기값으로 한 학습으로 전역적 최적해에 수렴하는 유전자와 역전파 알고리즘을 병합하는 방법을 제안하였다. 본 논문에서는 유전자 알고리즘으로 역전파 알고리즘의 초기 입력값들을 설정하고 학습하여 최적해에 수렴하는 방법을 이용하였다.

유전자 알고리즘의 탐색 결과로 숫자 인식에 적합한 연결강도, 오프셋 값을 역전파 학습 알고리즘의 입력으로 하고 기울기 강하 기법(gradient descent)의 알고리즘으로 오차를 줄여나가는 학습을 실행하여 인식한다.

```

단계 1 : Initial setup
단계 2 : GA execution
procedure GA (population)
begin
  for individual i∈P do i=Local-search(i);
  repeat;
    for i=1 to #crossovers do
      select two parents ia, ib ∈ P randomly;
      ic = Crossover(ia, ib);
      id = Crossover(ia, ib);
      add individual ic, id to P;
    endfor;
    for j = 1 to #mutations do
      select an individual i∈P randomly;
      im=Mutate(i);
      add individual im to P;
    endfor;
  end;
단계 3 : 연결강도와 오프셋 값 결정
단계 4 : BP Neural Network
단계 5 : BP adaptation
  
```

2.3 실험결과



〈그림 2〉 제안된 구조

필기체 숫자의 인식에 대해서 제안된 방법의 향상된 수렴속도의 성능을 알아보기 위해서 먼저 역전파 알고리즘만으로 실행한 결과를 표2에서 알아보고 다음으로 유전자 알고리즘과 역전파 알고리즘을 병합하여 실행한 경우를 표3에 나타내었다. 유전자 알고리즘과 역전파 알고리즘을 하나만 실행한 경우보다 제안한 방법이 빠른 수렴시간의 결과를 나타내어 탐색시간의 성능에 효과가 있음을 알 수 있었다. 첫번째는 역전파 알고리즘 단독으로 실행한 결과이고 두 번째는 유전자 알고리즘과 역전파 알고리즘을 병합한 경우의 수렴시간을 보았다.

〈표 1〉 유전자 알고리즘의 초기조건값

		유전자 알고리즘의 초기조건 값
신경망 설계	은닉노드수의 범위	1~255
	허용오차 범위	0.005~0.015
	학습률 값의 범위	0.1~1.0
	모멘트값의 범위	0.1~1.0
신경망 정보	목표오차	0.01
	실험 Cycle	10
	전환함수	시그모이드 함수
GAs정보	종료조건	지정된 적합도 수렴률
	Population 수	30
	Crossover 확률	0.1
	Mutation 확률	0.05
	Elitism 확률	0.01

〈표 2〉 역전파 알고리즘 실행 학습결과

학습횟수	20	50	100	250	400
학습시간	20s	53s	98s	243s	463s
수렴률(%)	80.6	92.7	97.65	99.03	99.78

〈표 3〉 유전자 알고리즘을 적용한 역전파 알고리즘 실행 학습결과

학습횟수	20	50	100	250	400
학습시간	18s	52s	99s	241s	456s
수렴률(%)	81.79	93.5	98.45	99.1	99.87

3. 결 론

본 논문에서는 필기체 숫자 인식에서 유전자 알고리즘을 적용한 역전파 알고리즘을 적용하였다. 제안한 방법은 수렴 시간과 학습시간의 성능 비교에서 역전파 알고리즘을 단독으로 실행했을 경우보다 수렴 시간의 성능에서 더 우수함을 나타내었다. 또한 제안한 방법에서 가장 수렴시간이 빠른 경우는 그 결과를 역전파 알고리즘의 파라미터로 하여 학습을 실행하여 최적해에 수렴한 경우로 유전자 알고리즘 단독 실행보다 더 빠른 수렴시간을 나타내었다.

앞으로의 연구방향은 초기 입력값에 대하여 더 나은 학습 방법이 이루어질 수 있도록 하는 연구가 계속 되어야 할 것이고, 유전자 알고리즘의 기존의 연산자 외에 탐색 기능을 향상시킬 수 있는 새로운 연산자에 대한 연구가 있어야 하겠다.

[참 고 문 헌]

- [1] J.H.Holland, "Adaptation in natural and artificial systems." The Univ. of Michigan Press, 1975.
- [2] C.T. Lin, C.P. Jou, and C.J. Lin, "GA-based reinforcement learning for neural networks", International Journal of System Science, Vol. 29, No.3, pp.233-247, 1998
- [3] A. Homaifar and E. McCormick, "Simultaneous design of membership functions and rule sets for fuzzy controllers using genetic algorithms", IEEE Trans. Fuzzy Syst., Vol. 3, pp. 129-139, 1995.
- [4] D.E. Goldberg, Genetic Algorithms in search Optimization & Machine Learning, Addison. Wesley, 1989
- [5] D.E. Rumelhart, R. Durbin, R. Golden. and Y. Chauvin, "Backpropagation : The basic theory." in Backpropagation - Theory, Architectures, and Applications, Y. Chauvin and D.E.Rumelhart(Eds.), pp.1~34. Lawrence Erlbaum Associates, Publishers, 1995.
- [6] D.E. Rumelhart, G.E. Hinton, and R.J.Williams,"Learning Internal Representation by Error Propagation," in Parallel Distributed Processing, Vol.1. pp.318~362, 1986
- [7] Kitano. H. "Empirical studies on the speed of convergence of neural network training using genetic algorithms " In proc. 8th JMIT National Conf in Artificial Intelligence, Vol. 2, pp.789~796, 1990.