

다중 신경망의 계층 결합에 의한 필기체 숫자 인식에 관한 연구

김 두 식(金斗植), 임 길 택(林吉澤), 남 윤 석(南潤奭)
한국전자통신연구원 정보화기술연구본부 우정자동화팀
전화 : (042) 860-5809 / 팩스 : (042) 860-6508

A Study on Handwritten Digit Recognition by Layer Combination of Multiple Neural Network

Doo Sik Kim, Kil Taek Lim, Yun Seok Nam
Postal Automation Technology Development Team
Electronics and Telecommunications Research Institute
E-mail : doosik@etri.re.kr

Abstract

In this paper, we present a solution for combining multiple neural networks. Each neural network is trained with different features. And the neural networks are combined by four methods. The recognition rates by four combination methods are compared. The experimental results for handwritten digit recognition shows that the combination at hidden layers by single layer neural network is superior to any other methods. The reasons of the results are explained.

I. 서론

패턴 인식의 문제가 복잡해질수록, 신경망의 크기가 커지고, 학습 시간과 계산량이 증가한다. 이러한 문제를 극복하기 위하여 여러 개의 작은 신경망들을 결합하는 방법이 제시되고 있다[1-5].

소규모의 다중 신경망을 결합하는 접근 방법은 대규모의 단일 신경망에서 발생하는 문제들을 해결한다. 즉, 다중 인식기의 결합에 의하여 인식률과 정확도의 향상, 신뢰도와 결합 허용도의 개선, 상이한 성질을 갖는 각 다중 인식기들의 단점 보완, 복잡한 문제의 분할 정복 등과 같은 효과를 기대할 수 있다.

다중 인식기의 결과를 결합하는 방법은 선형 결합(linear combination), 단순 평균(simple averaging), 다수결 투표(majority voting) 방법 등이 있으며[1], 패턴 인식 방법을 적용하기도 한다. 이는 다중 인식기의 각 결과를 종합하여 최종 결과를 출력하는 문제가 패턴 인식의 문제와 근본적으로 유사하다는 점에 기인하며, 이에 따라 다중 신경망(multilayer neural network), LVQ(learning vector quantization), 퍼지 이론 등을 이용하여 다중 신경망을 결합하는 방법에 관한 연구 결과가 소개되고 있다[6, 7]. 다중 인식기의 결합에 의한 문제 해결 방법은 패턴 인식뿐만 아니라, 영상 압축, 통계 분석, 예측 등과 같이 다양한 분야에서 적용되고 있으며, 기존의 단일 인식기로는 해결하기 어려운 복잡한 문제를 다중 인식기의 결합 방법으로 효과적으로 해결할 수 있음을 보여주고 있다[2].

본 논문은 다수의 다중 신경망의 각 은닉층(또는 출력층)의 결합 결과를 하나의 다중(또는 단층) 신경망의 입력층으로 재구성하여 필기체 숫자 인식률을 향상시키는 방법을 소개한다.

본 논문의 구성은 다음과 같다. 2 절에서 기존의 신경망을 이용한 인식기 결합 방법을 간단히 소개하고, 3 절에서는 은닉층 또는 출력층을 결합한 결합 신경망의 구조와 학습 방법을 제안한다. 4 절에서는 제안된 방법을 실험 결과를 이용하여 비교 및 분석한다. 마지막으로 5 절에서는 간단한 결론과 추후 연구 방향을 제시한다.

II. 신경망을 이용한 다중 인식기의 결합

그림 1은 다중 신경망의 일반적인 구조를 보여주고 있다. 이와 같은 전통적인 신경망은 추가 학습을 하는 경우에 대하여 기존의 학습된 정보를 점진적으로 상실하며, 문제가 복잡하고 입력 패턴의 차원이 커질수록 신경망의 크기가 커져야 하고, 이로 인하여 학습 속도 및 학습 능력의 저하와 메모리의 증가를 초래한다.

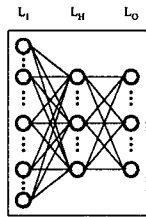


그림 1 다중 신경망의 구조
Fig. 1 Structure of Multilayer Neural Network

인식 대상 집합의 분포가 복잡하면, 하나의 인식기로 해결하려는 방법보다 다수의 작은 인식기로 나누어 해결하는 방법이 효과적이며[2], 인식률이 낮은 다수의 인식기들을 결합하여 인식률을 향상시킬 수 있다[3].

Waibel[2]은 각 다중 신경망들의 은닉층을 입력층으로 하는 다중 신경망을 이용하여 다중 신경망들을 결합함으로써, 복잡한 음성 인식의 문제를 분할하여 해결하였으며, 은닉층에서의 결합이 효과적임을 보였다.

Wesolkowski[3]는 신경망의 출력층을 단층 신경망과 다중 신경망으로 각각 결합하여 비교 실험한 결과, 단층 신경망에 의한 결합 방법이 상대적으로 우수한 인식률을 보이고, 다수결 투표법과 같은 비적응적(non-adaptive) 결합 방법들에서뿐만 아니라, 유전자 알고리즘 등을 이용한 적응적(adaptive) 결합 방법들 중에서도 가장 높은 인식 성능을 보임을 확인했다.

III. 은닉층 결합과 출력층 결합,

단층 결합 신경망과 다중 결합 신경망

지금까지 신경망을 이용한 다양한 결합 방법 및 형태가 발표되어 왔으나, 본 논문에서는 두가지 기준을 이용하여 네 유형으로 구분하고자 한다. 즉, 다중 신경망의 어느 층을 결합했는가에 따라 은닉층 결합 및 출력층 결합으로 구분하고, 어떤 결합 신경망을 사용했는가에 따라 단층 결합 신경망과 다중 결합 신경망으로 구분한다. 그림 2는 다중 신경망을 결합 신경망으로 사용한 경우를 보여주며, 그림 3은 단층 신경망을

결합 신경망으로 사용한 경우를 보여준다. 각각의 경우에 대하여 다중 신경망의 입력층을 결합했는가, 출력층을 결합했는가에 따라 다시 구분된다. 이러한 구분에 따르면, TCNN[1]과 Waibel[2] 방법은 은닉층의 다중 신경망 결합(그림 2b)에 해당하며, DCNN[4]은 출력층의 단층 신경망 결합(그림 3a)에 해당한다. 2절에서 언급한 바와 같이 Wesolkowski[3]는 3a 유형의 신경망 결합 방법이 2a 유형보다 다소 우수한 인식률을 보임을 확인하였다. 그러나, 그림 3b와 같은 은닉층의 단층 신경망 결합 형태에 대한 선행 연구는 없다.

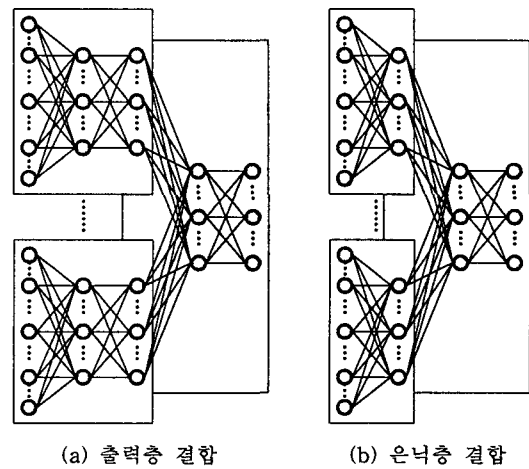


그림 2 다중 신경망에 의한 다중 신경망 결합
Fig. 2 Combination of Multiple Neural Networks by Multi-Layer Neural Network

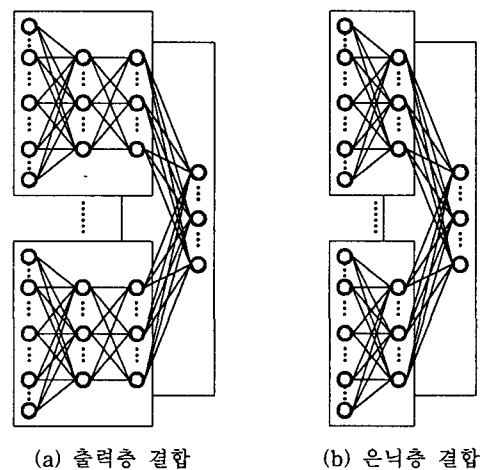


그림 3 단층 신경망에 의한 다중 신경망 결합
Fig. 3 Combination of Multiple Neural Networks by Single Layer Neural Network

IV. 실험 결과

실험 데이터는 Concordia 대학에서 수집된 필기체 숫자 데이터로 전체 6,000자를 학습용 2,000자, 검증용 2,000자, 평가용 2,000자로 각각 구분하였다. 인식기에 입력되는 특징은 표 1의 네 가지 특징을 사용하였다.

표 1 실험에 사용된 특징들
Tab. 1 Features Used for Experiment

기호	특징 이름	차원수
DC	방향 기여도 (directional contributivity)	144
DT	거리 변환 (distance transform)	144
GD	기울기 (gradient)	144
PC	투영 및 교차 거리 (projection + cross distance)	108

오류 역전과 신경망의 학습을 위한 매개변수로 관성항은 0.5, 학습률은 0.05를 사용하였고, 바이어스항을 사용하였으며, 다중 신경망에 대한 은닉층에서의 유닛 수는 10부터 50까지 다양하게 실험하였다.

학습용 데이터에 대한 학습을 무작위 순서로 300회를 수행하는 동안, 매 회마다 검증용 데이터로 인식률을 측정하여 인식률이 가장 높았을 때의 신경망 상태를 저장하였고, 학습이 끝나면 저장된 신경망 상태를 최종 학습 결과로 사용하였다.

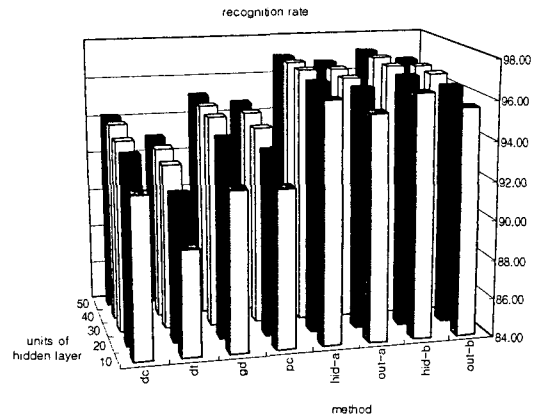
이렇게 학습된 네 종류의 요소(component) 신경망에 대한 결합 방법은 3절에서 언급한 바와 같고, 결합 신경망의 학습 단계에서는 요소 신경망의 연결 강도는 고정한 채, 결합 신경망의 연결 강도만 변화시킨다.

다중 결합 신경망에서의 은닉층 유닛 수는 10개로 고정하였으며, 학습을 위한 매개변수 값과 학습 방법은 요소 신경망의 학습과 동일하게 진행하였다.

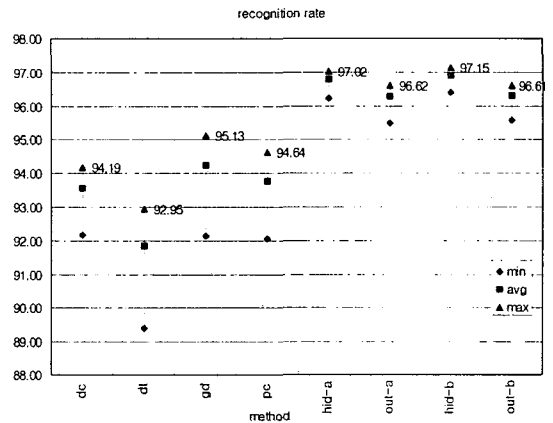
각 신경망의 학습이 끝나면 평가용 데이터로 각 방법에 대한 인식 성능을 평가하였고, 이와 같은 학습 및 평가 과정을 20회 반복하여 각각에 대한 평균 인식률을 구하였다. 그 결과는 그림 4에서 보여주고 있다.

그림에서 dc, dt, gd, pc는 각 요소 신경망을 나타내며, hid-a, out-a, hid-b, out-b는 요소 신경망들에 대한 은닉층 다중 결합, 출력층 다중 결합, 은닉층 단층 결합, 출력층 단층 결합 유형에 의한 다중 결합 신경망을 각각 나타낸다. 그림 4a는 요소 신경망의 은닉층에서의 유닛 수에 따른 인식률의 변화를 보여주며, 그림 4b는 이에 대한 평균과 최소, 최대값을 나타낸다.

실험 결과를 살펴보면 다음과 같은 몇 가지 사실을 확인할 수 있다.



(a) 은닉층의 유닛 수에 따른 인식률 변화



(b) 각 방법에 대한 평균, 최소, 최대 인식률

그림 4 각 방법에 대한 인식률
Fig. 4 Recognition Rate of Each Method

첫째, 각 개별 신경망보다 각 신경망을 결합한 다중 결합 신경망의 인식률이 상대적으로 우수했다. 이는 다중 인식기를 결합함으로써 단일 인식기보다 우수한 인식률을 갖는 인식 시스템을 구성할 수 있음을 재확인 한 것이다.

둘째, 다중 결합 신경망의 인식률의 변화는 개별 요소 신경망의 인식률의 변화보다 작다. 이는 다중 결합 신경망에 대한 인식률이 보다 안정적이며, 또한 단일 신경망의 크기를 증가시키는 것보다 소규모의 단일 신경망들을 다중 결합하는 방법이 인식률과 인식 시스템의 크기 측면에서 효과적임을 나타낸다.

셋째, 출력층 결합보다 은닉층 결합이 보다 우수한 결과를 보여준다. 이는 은닉층에서의 특징 분포가 출력층에서의 특징 분포보다 분별이 용이함을 나타낸다.

넷째, 다중 결합 신경망보다 단층 결합 신경망이 보다 우수한 결합 결과를 보여준다. 이는 선형 구분이 가능한 은닉층 또는 출력층 상의 특징 분포에 대한 인식의 문제에서 다중 신경망보다 단층 신경망이 효과적임을 의미한다. 특히, 단층 결합 신경망의 학습 곡선은 다중 결합 신경망보다 빨리 수렴할 수 있었다.

다섯째, 은닉층의 단층 결합 신경망이 가장 우수한 인식 성능을 보였다. 이는 은닉층 상에서의 특징 분포가 선형 구분이 가능함을 의미하며, 다른 세 가지 유형의 결합 방법에 비하여 신경망의 크기가 상대적으로 작기 때문에, 빠른 속도의 수렴과 학습의 안정이 비교적 용이했던 것으로 추정된다.

V. 결론

본 논문에서는 인식 성능의 향상을 위해 신경망을 이용하여 다수 개의 신경망 인식기를 병렬 결합하는 기존의 방법들을 정리하였으며, 기존의 다양한 결합 방법에 비하여 다수 신경망 인식기의 각 은닉층들을 단층 신경망으로 결합하는 새로운 방법이 인식 성능 측면에서 가장 우수하였음을 실험에 의해 확인할 수 있었다. 특히, 제안된 방법은 기존의 방법에 비하여 신경망 구조의 크기가 가장 작기 때문에 학습 및 인식 속도가 가장 우수하였다. 즉, 제안된 방법은 인식률의 향상뿐만 아니라, 인식 시스템의 크기 축소, 계산량 감소, 속도 향상 효과를 함께 얻을 수 있었으며, 다양한 입력 특징에 대하여 개별적으로 미리 학습된 인식기를 통합함으로써 점진적인 학습에 대한 빠른 학습 효과를 기대할 수 있었다.

실험 결과, 다중 신경망의 은닉층에 대한 특징 분포가 선형 분포를 이루고 있음을 재확인할 수 있었으며, 선형 특징 분포 공간상의 분류 문제에서 다중 신경망보다 단층 신경망이 보다 효과적임을 알 수 있었다.

앞으로의 연구 방향은 다음과 같다. 여러 신경망이 결합되어 있는 다중 결합 신경망은 하나의 부분 연결 신경망(partially connected neural network)과 구조가 유사하다는 점에서 연관성을 찾을 수 있지만, 다중 결합 신경망은 두 단계의 학습 과정이 필요한 반면에, 부분 연결 신경망은 단일 단계의 학습이 가능하다는 점에서 차이가 있다. 다중 결합 신경망의 두 단계 학습 과정에서 요소 신경망들을 각각 독립적으로 학습시킬 수 있다는 점은 장점이 될 수도 있으나, 실시간 추가 학습과 같은 몇몇 요구 사항에는 적합하지 않다. 따라서, 다중 결합 신경망에 대한 단일 단계의 추가 학습 방법을 개발하고, 이를 통하여 전체적인 학습 속도가 개선될 수 있는지를 연구할 계획이다.

참고문헌

- [1] 류강수, 진성일, "모듈화된 신경회로망 중간층 출력의 재학습에 의한 필기체 숫자 인식", 정보과학회 논문지(B), 제 23권 제 9호, pp. 931-940, 1996.
- [2] A. Waibel, H. Sawai, and K. Shikano, "Consonant recognition by modular construction of large phonemic time-delay neural networks", Int. Conf. on Acoustics, Speech and Signal Processing, Vol. 1, pp. 112-115, 1989.
- [3] S. Wesolkowki and K. Hassanein, "A comparative study of combination schemes for an ensemble of digit recognition neural networks", Int. Conf on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation, Vol. 4, pp. 3534-3439, 1997.
- [4] D.-S. Lee and S. N. Srihari, "A theory of classifier combination: the neural network approach", Proc. of the 3rd Int. Conf. on Document Analysis and Recognition, Vol. 1, pp. 42-45, 1995.
- [5] S. Hashem, "Optimal linear combinations of neural networks", Neural Networks, Vol. 10, No. 4, pp. 599-614, 1997.
- [6] S.-B. Cho and J.-H. Kim, "Combining multiple neural networks by fuzzy integral for robust classification", IEEE Trans. on Systems, Man, and Cybernetics, Vol. 25, No. 2, pp. 380-384, Feb. 1995.
- [7] A. Weingessel, H. Bischof, K. Hornik, and F. Leisch, "Adaptive Combination of PCA and VQ networks", IEEE Trans. on Neural Networks, Vol. 8, No. 5, pp. 1208-1211, Feb. 1997.