

한글의 구조적 특징을 이용한 인쇄체 한글인식을 위한 신경망 설계

서원택^o, 조범준

조선대학교 컴퓨터공학과, 조선대학교 컴퓨터공학부

Neural Network design for Printed Hangeul recognition using structural characteristic of Hangeul

Wontaek Seo, Beomjoon Cho

Dept. of Computer Science Engineering, School of Computer Science Engineering

요 약

본 논문에서는 한글의 구조적인 특징을 이용하여 이를 효과적으로 인식할 수 있는 신경망을 설계해보았고, 이를 이용하여 주민등록증에 있는 이름을 인식하는 시스템을 구성해 보았다. 본 시스템은 한글의 6형식에 따른 구조적인 특징을 효과적으로 구분하기 위해 형식을 구분하는 신경망을 먼저 구성하여 형식별로 분류한 뒤, 형식에 따라 자모음을 분리하여 각 형식에 따라 구성된 2차 신경망으로 입력을 하여 인식하는 구조로 설계되었다. 훈련용 데이터는 각 형식 별로 자소를 분리해서 얻은 영상들을 자소별 평균이미지로 만들어서 이를 조합하여 만든 글자로 사용하였다. 그래서 같은 형식의 같은 자음이라도 글자의 모양과 위치가 조금 다른것에 대해서 강인한 훈련을 할 수 있었다. 또한 히스토그램의 국부 평균을 적용함으로써 잡음에 효과적으로 대응하였다. 100명의 주민등록증을 컴퓨터 카메라를 이용하여 입력 받아서 테스트한 결과 98.1%의 높은 인식률을 얻을 수 있었다.

1. 서 론

컴퓨터를 이용하여 문서에 인쇄되어 있는 정보를 자동으로 저장하고자 하는 문제에 대해서 많은 연구가 이루어지고 있다. 문서에 있는 정보를 효과적으로 얻어내기 위해서는 문서의 구조를 잘 파악하고 글자부분을 잘 추출해야 하는데, 이것에 대한 연구가

현재까지 활발히 연구되고 있다[1]. 추출된 글자부분에서 인쇄되어 있는 문자이미지를 효과적으로 인식하는 연구도 많이 연구되고 있다. 그 중에서 신경망은 매우 효과적인 방법으로 알려져 있는데, 그 이유는 기존의 방법들은 이미지의 잡음에 대해서 민감해서 이를 해결하기 위해서 전처리 부분이 복잡한 반면, 신경망은 자체적으로 잡음에 대해 강인한 성질을

가지고 있어서 전처리 부분이 훨씬 간단해진다. 또한, 기존의 방법들은 글자의 형태에 따라 모두 수학적 모델을 찾는데 어려움을 갖게 되지만, 신경망은 인간의 두뇌 신경조직을 모델로 만들었기 때문에 그 자체의 적응능력으로 적절한 알고리즘을 생성해 내서 적용할 수 있다는 것에서 문자인식이나 음성인식에서 각광을 받고 있다. 특히, 문자인식 분야에서는 오류역전파(Backpropagation) 학습법을 이용한 신경망을 이용하는 연구가 활발히 진행되고 있다[2].

본 논문에서는 문자인식 중에서도 인쇄체 한글인식을 효과적으로 수행하는 계층적 신경망에 대해서 연구하였다. 본 논문에서 구현된 인식기는 주민등록증을 이용한 신원확인 시스템에서 앞면의 성명 부분을 인식하는 모듈단위의 프로그램으로 사용하도록 만들어졌다.

2 장에서는 본 연구에서 사용된 오류역전파 신경망에 대해서 소개하고, 3 장에서는 인쇄체 한글을 인식하기 위해 계층적으로 신경망을 구성한 방법에 대해서 소개한다. 4 장에서는 실제 실험에서 사용된 데이터세트와 종류에 대해서 소개하고 인식을 구현한다. 5 장에서는 결과 및 향후 연구 방향에 대해서 기술한다.

2. 신경망 알고리즘

2.1 오류 역전파 알고리즘

오류역전파 알고리즘은 최대경사도법(steepest descent rule)을 사용하여 에러평균의 제곱근이 최소가 되도록 연결강도를 조정하는 학습법이다. 이 알고리즘은 크게 두 과정으로 나눌 수 있는데, 연산과정과 학습과정이다. 먼저 연산과정 동안에는 하나의 입력패턴에 대해 현재의 연결강도로 은닉층의 출력과 출력층의 출력을 구한다.

$$NET_z = x_k v_k^T, \quad z = f(NE_T_z) = \frac{1 - \exp(-NET_z)}{1 + \exp(-NET_z)} \quad (1)$$

$$NET_y = zw^T, \quad y = f(NE_T_y) = \frac{1 - \exp(-NET_y)}{1 + \exp(-NET_y)} \quad (2)$$

위식에서 전달함수로는 양극성 시그모이드 함수를

사용하였고 v 는 입력층과 은닉층간의 연결강도이고 w 는 은닉층과 출력층 간의 연결강도이다. z 는 은닉층의 출력이고 y 는 출력층의 출력이다.

그 다음 학습과정에서는 실제 출력과 원하는 출력간의 차이를 계산하여 오차(E)의 값이 최소가 되도록 하는 연결강도의 변화량을 역으로 전파 시켜 연결강도를 변경시킨다.

$$E = \frac{1}{2}(d - y)^2 + E \quad (3)$$

$$\delta_y = (d - y)(1 - y^2) \quad (4)$$

$$\delta_z = \frac{1}{2}(1 - z^2) \sum_{j=1}^m \delta_j w_j \quad (5)$$

$$w^{k+1} = w^k + \Delta w^k = w^k + \alpha \delta_y z^k \quad (6)$$

$$v^{k+1} = v^k + \Delta v^k = v^k + \alpha \delta_z x^k \quad (7)$$

위의 과정들이 오차가 허용할 수 있는 최소값에 도달할 때까지 계속 반복하게 된다.

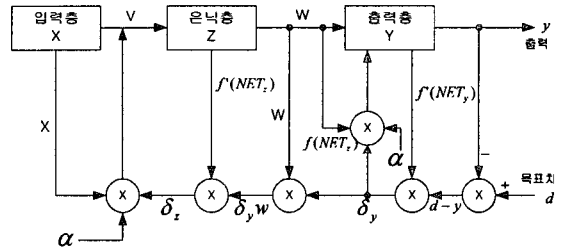


그림 1. 오류역전파 학습 알고리즘의 블록다이어그램

2.2 모멘텀 알고리즘

본 논문에서는 신경망의 학습을 좀더 효과적으로 하기 위해서 모멘텀 알고리즘을 적용하였다. 모멘텀 BP알고리즘은 연결강도를 조정할 때 그 전단계의 연결강도 변화량을 보조적으로 사용하는 방법이다. 알고리즘을 아래의 식에 나타내었다.

$$\Delta v^k = \alpha \delta_z x^k + \beta \Delta v^{k-1} \quad (8)$$

$$\Delta w^k = \alpha \delta_y z^k + \beta \Delta w^{k-1} \quad (9)$$

위식에서 α 와 β 는 각각 학습률과 모멘텀 상수

이며, δ_x 와 δ_y 는 각각 은닉층과 출력층의 오차신호이다. 따라서 $k+1$ 학습단계의 연결강도 v^{k+1} 과 w^{k+1} 은 다음과 같다.

$$v^{k+1} = v^k + \alpha \delta_x x^k + \beta \Delta v^{k-1} \tag{10}$$

$$w^{k+1} = w^k + \alpha \delta_y z^k + \beta \Delta w^{k-1} \tag{11}$$

모멘텀 BP알고리즘에서 모멘텀 상수 β 를 일반적으로 0~0.8 사이의 값으로 설정하며, BP알고리즘에 비하여 학습시간이 단축된다. 다만, 이전 학습 단계의 연결강도를 저장하여야 하는 등의 절차가 요구된다.

3. 인쇄체 한글 인식을 위한 계층적 신경망

3.1 유형분류 신경망

본 연구에서 한글을 인식하기 위해 구성한 신경망은 먼저 문자 이미지를 입력받아 한글의 6형식중의 하나로 분류하는 신경망으로부터 시작된다. 한글은 기본자소가 그림 2와 같이 일정한 위치에 존재하며, 각 위치에 올 수 있는 자소의 종류가 표 1와 같이 정해져 있다.

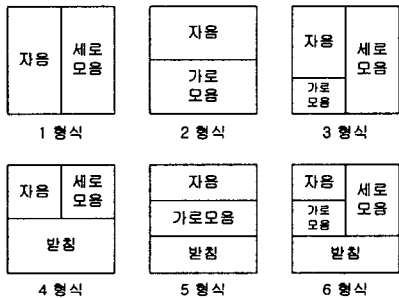


그림 2. 한글 6형식의 구조

표 1. 각 부분에 올 수 있는 자소 그룹 분류

자음	ㄱ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅅ, ㅇ, ㅈ, ㅊ, ㅋ, ㆁ, ㆁ, ㆁ, ㆁ
세로모음	ㅏ, ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ, ㅡ

가로모음	ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ, ㅡ
받침	ㄱ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅅ, ㅇ, ㅈ, ㅊ, ㅋ, ㆁ, ㆁ, ㆁ, ㆁ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅅ, ㅇ, ㅈ, ㅊ, ㅋ, ㆁ, ㆁ, ㆁ, ㆁ

3.2 자소별 문자 인식 신경망

일단 유형이 분류된 문자 이미지는 형식에 따라서 자소 이미지로 분리가 되어 각각의 형식에 맞는 신경망으로 입력이 된다. 이에 본 논문에서는 각형식별 자소별로 총 15개의 개별 신경망을 구성하였다. 3형식과 6형식의 이중모음은 하나의 신경망으로 구성하였다. 신경망의 구성은 입력층과 은닉층, 출력층으로 구성하였고 노드간의 연결은 모두 연결(Full Connected)로 하였다.

3.3 전체 신경망의 구조

본 연구에서 사용한 신경망의 전체 구조를 그림 3에 나타내었다. '냄'이라는 글자 이미지가 유형분류 신경망을 거치면서 4형식으로 판정되고, 그 이미지가 4형식의 구조에 맞춰 분리가 되어서 4형식을 담당하는 신경망의 각 자소부분으로 입력이 되었다.

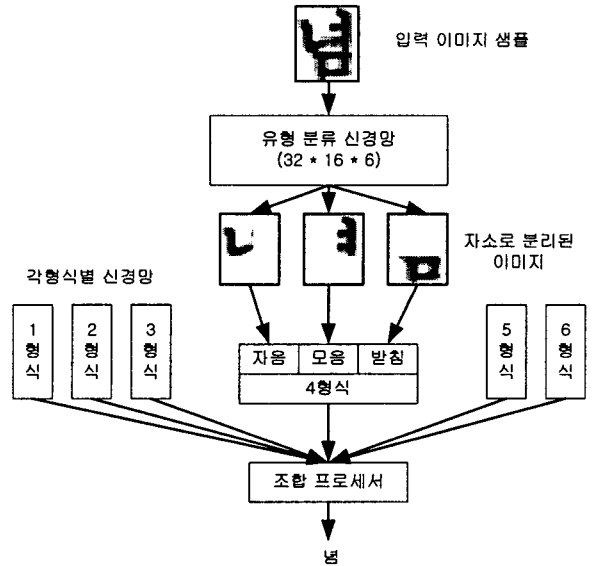


그림 3. 인쇄체 한글인식을 위한 계층적 신경망 구조

4. 실험 및 분석

주민등록증상에서 성명 부분의 위치는 고정적이기 때문에 간단하게 스펙트럼을 이용하여 글자부분을 추출하였고, 각 글자의 이미지를 32*27의 이미지로 분리했다. 이진화 과정을 거치고, 신경망의 입력으로는 히스토그램을 사용하였는데 (2W+1)의 범위의 국부평균을 이용하였다.

$$H(x) = \frac{1}{2w+1} \sum_{i=x-w}^{x+w} h(i) \quad (12)$$

단, w는 평균을 구하는 범위이고, 본 실험에서는 3을 사용하였다.

본 실험에서 신경망을 학습 시키기 위한 글자세트로는 본교의 연구실에서 연구의 목적으로 정보과학회지 논문 100편의 초록의 글자세트를 분류해 놓은 데이터베이스를 사용하였다. 실험데이터는 그 중에서 각 형식별, 자소별로 평균이미지를 구하고 이를 합성한 글자 이미지 1800자를 이용하였다. 평균이미지를 사용한 이유는 같은 형식의 같은 자음이라 할지라도 모양에 따라 그 모양과 위치, 크기가 다르기 때문에 실제 신경망에 적용을 시키면 애러가 될 가망성이 많아 진다. 이에 평균의 이미지를 이용하여 학습을 하게 되면 그러니 위치의 변화에 강인해 지는 것을 확인하였다.

주민등록증 100 장에 대해서 실험한 결과 98.1%의 성공률을 나타내었다.

5. 결론 및 향후 방향

본 연구에서는 인쇄체 한글의 인식기를 계층적 신경망을 이용하여 구현하여 보았고, 이를 주민등록증에 있는 성명을 인식하는데 적용해 보았다. 본 연구에서는 인식률을 높이기 위해 훈련 데이터를 실제 데이터에서 얻어낸 데이터의 평균 이미지를 이용해서 인위적으로 잡음을 넣는 방법보다 훨씬 효과적이라는 것을 알 수 있었다. 단 신경망의 개수가 15개가 된다는 점은 단점이라 할 수 있겠다.

신경망의 개수를 줄일 수 있는 방법을 찾아보는 것이 향후 과제라 할 수 있겠다.

참고문헌

- [1] Yi Xiao, Hong Yan "Text region extraction in a document image based on the Delaunay tessellation" Pattern Recognition Vol.36, pp.799-809, 2003
- [2] 조성배, 김진형 "인쇄체 한글문자의 인식을 위한 계층적 신경망" 한국정보과학회 논문지, Vol.17, No.3 pp.306-316, 1990