

역전파 학습 신경망을 이용한 한글 숫자 인식에 관한 연구

박 창민, 박 귀순, 김 대원, 이 동춘, 김 명원, 배 현주, 차 의영
부산 대학교 계산통계학과

A Study on the Handwritten Korean Numeric Recognition using a Backpropagation Learning Neural Network

Changmin Park, Kwisoon Park, Daewon Kim, Dongchoon Lee,
Myengwon Kim, Hyunjoo Bae, Euiyoung Cha
Dept. of Computer Science and Statistics, PNU

요 약

본 논문에서는 신경망 구조의 한 모델인 *feed-forward multi-layered network*에 역전파 학습(back-propagation learning) 기법을 이용하여 필기체 한글 숫자를 인식하고 그 가능성을 보였다.

문자 인식에 있어 입력 대상의 모양이 왜곡 되거나, 대상의 크기 혹은 위치의 변화 등과 같은 잡음(noise)에 대해서 정확히 대상을 인식하는 데는 대상의 구조 추출에 크게 관여 되므로 한글의 구조 추출에 적합하다고 생각되는 bar mask 부사법을 제안하였다. 모델의 학습은 필기체 한글 숫자 16자의 입력 패턴과 타겟(target) 입력의 쌍을 이용해 학습시켰다. 또한, 모델의 인식 정도를 측정해 보기위해 시험패턴을 적용하여 훈련된 패턴과 훈련 되지 않은 패턴간의 인식률을 비교 하여 보았다.

I. 서 론

디지털 컴퓨터가 0,1의 이진신호를 이용하여 yes/no의 판단을 하는 반면 신경망 회로는 아날로그 신호를 이용해 불일치하거나 모순된 데이터에 대해서도 가중판단(weighted decision)을 내림으로써 문제를 해결한다. 또한, 데이터를 분류함에 있어서는 완전정합(exact matching)보다는 근접정합(closest matching)을 행함으로써 정보의 일부가 파손된 데이터에 대해서도 근접된 응답을 보이기 때문에 fault tolerance가 강하다. 이러한 이유로 인하여 여러 분야에 걸쳐서 신경망 연구가 활발히 진행되고 있으며, 특히 형태 인식이나 형태 분류, 영상 압축 등에서의 응용이 발표되었다.

문자인식에 있어서는 문자로부터 자획을 추출해내는

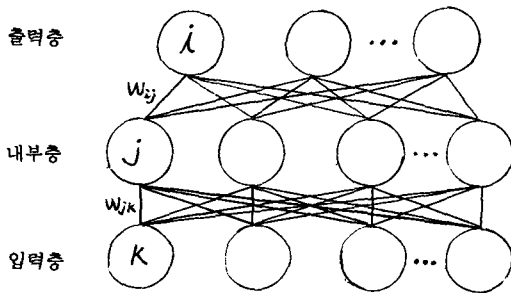
원리를 이용해 문자를 인식하는 neocognitron과 영문, 숫자의 구조에 적합한 mask를 만들어 특징을 추출함으로써 문자를 인식하는 shadow projection 방법이 대표적인데 neocognitron 은 복잡한 한글의 구조추출에 적합하지 않으며 구조가 복잡하여 계산량이 많아 시간이 오래 걸리기 때문에 실시간(real-time)처리가 어렵다. 그러나 shadow projection 방법은 입력 패턴의 필체를 추적하면서 각 화소값들을 가장 가까운 bar에 직선 이동시켜 각 bar에 나타난 화소값들의 밀도에 의하여 특징을 추출한다.

본 논문에서는 shadow projection 방법을 이용하여 문자의 특징을 추출하여 적은 수의 입력패턴을 정확하게 추출하였으며 이를 신경망 모델 중의 하나인 역전파 학습기법을 이용하여 필기체 한글 숫자 인식에 적용하고자 한다.

II. 역전파 모델(Backpropagation Model)

1. 구 조

지도학습(supervised learning) 알고리즘 중의 하나인 역전파 학습은 학습시켜야 할 모든패턴에 대해서 각각에 대한 기대되는 출력(desired output)과 실제로 신경망이 산출해 낸 출력(actual output)과의 차이를 최소화 하기 위하여 회로망 가중치(network weight)를 구해내는 알고리즘 으로서 일반화 된 delta rule을 사용하며 내부층(hidden layer)을 포함하는 feed-forward multi-layered neural network 구조를 갖는다.



(그림 1) 역전파 학습 모형의 topology

학습은 두 단계로 진행되는데 첫번째 단계는 입력 패턴과 타겟 입력을 제시하여 각 unit에 대해서 출력(즉 o_i)을 산출하는 전향 단계이고 두번째 단계는 desired output 과 actual output 과의 차이를 계산해서 이 차이를 역전파 시키면서 각 weight 를 조절하는 후향단계이다.

학습은 시스템이 안정될때 까지 즉, 총오류합(total error sum)이 오류기준치 (error criterion)를 만족할때 까지 계속 반복 실행한다. 오류 측정을 위해서는 Mean Square Error인

$$E = \sum_i (t_i - o_i)^2$$

{ 단, t_i : unit 의 desired output,
 o_i : unit 의 actual output }

을 이용하여 이 오류 E 를 최소로 하는 weights 를 구하는

gradient descent 방법을 이용한다. actual output을 산출하기 위해서는 net-input 과 활성화함수를 선택해야 한다

. net-input

$$\rightarrow net_i = \sum_j w_{ij} o_j + \theta_i \quad \{ \text{단 } \theta_i \text{는 unit}_i \text{의 bias} \}$$

. 활성화함수(semilinear activation function)

$$\rightarrow o_i = f(net_i) = 1 / (1 + \exp(-net_i))$$

각 unit의 오차를 구하기 위해서는 활성화함수의 미분치인 $f'(net_i)$ 를 구해보면 $o_i(1 - o_i)$ 가 산출된다.

이로써 출력층과 내부층의 unit에 대한 각 δ 를 구하는 식은

. 출력층의 unit $\rightarrow \delta_i = (t_i - o_i) o_i (1 - o_i)$.

. 내부층의 unit $\rightarrow \delta_j = (o_j(1 - o_j)) \sum_k \delta_k w_{kj}$

이 되어 내부층의 unit에 대해서는 재귀적(recursive) 연산이 이루어 지게 된다.

결과적으로 weight의 변화량은

$$\Delta w_{ij}(n+1) = \eta \delta_i o_j + \alpha \Delta w_{ij}(n)$$

{ 단, n은 time sequence , η 는 learning rate ,

α 는 momentum term }

이 된다.

2. 알고리즘

제 1단계 : . network 구성

- 입력, 내부, 출력 unit 수 지정

. 오류 한계치와 상수값 지정

. biases, weights 초기화

- 입력 unit을 제외한 모든 units와 모든

weights 에 -0.5 0.5 사이의

random number 를 할당

제 2단계 : 입력 패턴과 타겟 입력 제시

제 3단계 : . 각 unit 에 대해서 actual output 산출

- net-input

$$\Rightarrow net_i = \sum_j w_{ij} o_j + \theta_i$$

- logistic activation function

$$\Rightarrow o_i = f(net_i) = 1 / (1 + \exp(-net_i))$$

제 4단계 : . biases 와 weights 조절

- 모든 연결 가중치 w_{ij} 와 입력 units를

제외한 모든 units의 biases(θ_i)를 조절

한다.

- 조절규칙

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n)$$

$$\theta_i(n+1) = \theta_i(n) + \Delta \theta_i(n)$$

=> 변화량

$$\Delta w_{ij}(n) = \eta \delta_i \theta_j + \alpha \Delta w_{ij}(n-1)$$

$$\Delta \theta_i(n) = \eta \delta_i + \alpha \Delta \theta_i(n-1)$$

제 5단계 : 학습 상태 점검

- 총 오류합(tss)가 기준치 보다 크면

제 2단계로 가서 계속 수행하고 그렇지

않으면 학습을 끝내고 weights 와 biases

를 저장

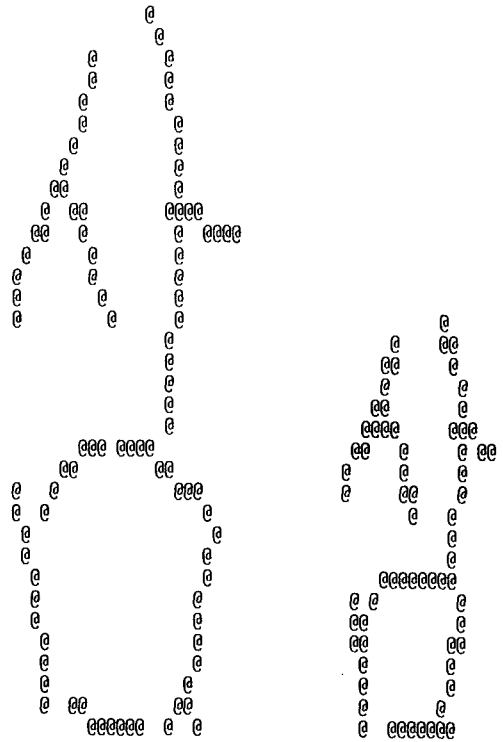
III. 한글 숫자 인식 시스템 구현

본 장에서는 입력장치를 통해 얻어진 패턴정보를 전처리(pre-processing) 하는 과정과 효율적인 입력을 구성하기 위한 bar mask projection 방법을 제안하고 학습과 인식에 대해 기술한다.

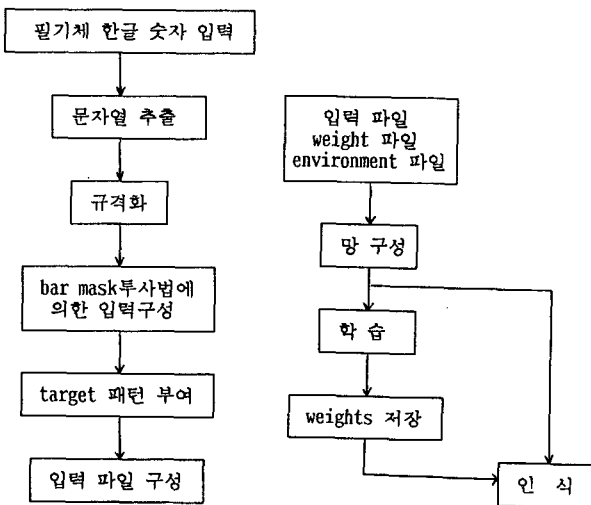
1. 전처리(Preprocessing)

전처리 과정은 그림2 에서 문자열 추출, 규격화 단계까지를 말한다. 입력 장치에 있는 문자를 이루는 화소(pixel)들의 좌표열을 추출하여 각 좌표의 최대값과 최소값을 구하여 입력된 문자의 크기를 결정한다. 이 과정이 끝나면 bar mask 투사법에 적용시키기 위해서 패턴 배열을 bar mask 크기인 20*16 로 규격화 한다.

그 방법은 패턴의 X축 크기와 Y축 크기에 대한 mask 크기의 비율에 따라 축소시키거나 확대시킨다.



(그림 4) scanner 를 통한 패턴 (그림 5) 규격화된 패턴



(그림 2) 전처리 과정

(그림 3) 인식 시스템의 구성

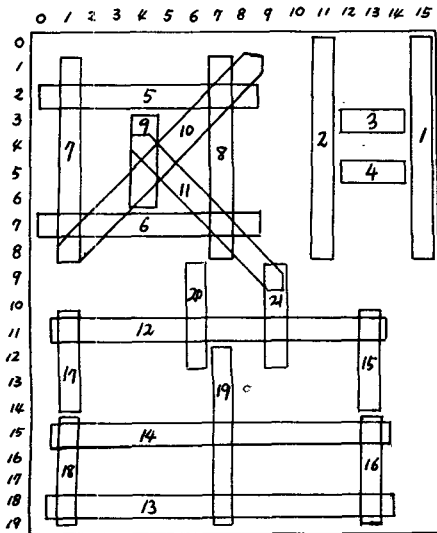
2. 입력 구상을 위한 bar mask 투사법

신경망에 입력을 주는 방법에는 한 화소(pixel)당 하나의 입력 unit 를 담당하는 방법이 있으나 이렇게 되면 200*200 배열에 있어서 40000개의 입력 units이 필요하게 되고 이로 인해 학습, 인식시간, 기억장소 면에서 상당한 낭비를 가져온다. 이러한 낭비를 줄이기 위해서는 패턴의

특징을 잃지 않는 범위 내에서 가급적 input units 의 수를 줄이는 것이 바람직하다.

이러한 필요성 때문에 영문자와 숫자의 특징 추출을 위한 shadow projection 방법이 발표 되었는데, 이 방법은 문자를 추적하면서 각 화소들을 가장 가까운 수평 bar, 가장 가까운 수직 bar, 가장 가까운 대각 bar 에 각각 직교 이동시켜 OR 연산을 이루는 방법이다.

본 논문에서는 이 shadow projection 방법에서의 bar 구조를 한글 숫자 인식에 맞도록 재구성 하였다.



(그림 6) 제안된 bar mask

그림5의 '삼' 패턴을 그림6의 bar mask에 투사 시켜면 각 bar 에는 그 입력 패턴에 대한 정보가 기록되고, 구해진 이 21 개의 bar 정보를 신경망의 입력으로 제시한다. bar mask 투사법을 통과 한 후의 결과는 그림7과 같다.

X_value[1]= 0.111111 X_value[2]= 1.000000
 X_value[3]= 0.250000 X_value[4]= 1.000000
 X_value[5]= 0.222222 X_value[6]= 0.555556
 X_value[7]= 0.500000 X_value[8]= 0.375000
 X_value[9]= 0.750000 X_value[10]= 0.875000
 X_value[11]= 1.000000 X_value[12]= 0.615385
 X_value[13]= 0.615385 X_value[14]= 0.384615
 X_value[15]= 1.000000 X_value[16]= 0.750000
 X_value[17]= 0.750000 X_value[18]= 1.000000
 X_value[19]= 0.125000 X_value[20]= 0.250000
 X_value[21]= 0.500000

그림 7) bar mask 투사법에 의해 만들어진 입력패턴

3. 학습 및 인식

본 실험에서 적용한 문자는 한글숫자 16자이다.

영, 일, 이, 삼, 사, 오, 육,
 칠, 팔, 구, 십, 백, 천, 만,
 억, 조

(그림 8) 실험에 적용한 한글숫자

위 16자를 구분하기 위해서 output unit 수를 4개(즉, $\log_2 16 = 4$) 로 해서 이들 unit에 2진수로 0에서 15까지의 숫자를 할당하는 방법인데 각 문자에 대한 target output은 (표 1)과 같다.

영	0000	팔	1000
일	0001	구	1001
이	0010	십	1010
삼	0011	백	1011
사	0100	천	1100
오	0101	만	1101
육	0110	억	1110
칠	0111	조	1111

(표 1) 각 패턴에 대한 target output

학습 패턴은 한 문자에 대한 필체가 모두 다르므로 한 문자당 5번씩 수집했다. 즉, 총 80자(16자*5번)를 학습에 적용 했다. 80 자의 학습 패턴에 대해서 target output할당이 끝나면 바로학습에 들어 가는데 학습이 끝나면 신경망의 weights 를 저장해서 인식에 사용한다.

IV. 실험결과

1. 내부층 의 unit 수에 따른 실험

주어진 문제를 해결하기 위한 domain knowledge를

저장하는 내부층은 그 unit의 수에 따라 신경망의 성능에 상당한 영향을 주는것으로 알려져있다.

본 논문에서 설계된 한글 숫자 인식 시스템이 내부갯수의 변화에 따라 어느 정도의 성능 변화 특성을 나타내는지 알아 보기 위해서 내부층의 unit갯수를 30, 40, 50, 60, 70개로 달리 하면서 각각에 대한 학습 속도와 인식률에 대해서 분석했다. 실험결과 내부층의 unit 수를 증가시킬수록 학습 속도가 빨랐으며 인식률 또한 증가함을 알 수 있었다.

V. 결 론

본 논문에서는 신경망 구조중의 하나인 feed-forward layered neural network 에 backpropagation 학습 알고리즘을 이용하여 한글 숫자를 인식하는 시스템을 소프트웨어 시뮬레이션 하였다. 학습시 기억 장소와 시간을 절약하고, 또한 변형된 필기체에도 영향을 받지 않도록 하기 위해서 한글 구조 추출에 적합한 bar mask 투사법을 제안 하였으며 입력 장치를 이용하여 수집된 문자 패턴에 이 bar mask를 투사시켜 입력 파일을 형성한 후 이것을 신경망에 제시했다. 실험은 16자의 한글 숫자에 대하여 각각 10개씩 입력 패턴을 얻어 낸 후 그 중 다섯자는 그 문자에 대한 학습 패턴으로 사용하였고 나머지 다섯 자는 인식 시험 패턴으로 사용하였다.

실험 결과, 학습 패턴에 대해서는 90% 이상의 인식률을 보였으나 시험 패턴에 대해서는 조금 떨어지는 인식률을 보였다.

[참고 문헌]

- [1] D.J.Burr, "Experiments on Neural Net Recognition of Spoken and Written Text," IEEE trans. on ASSP. Vol 36, No. 7, pp.1162-1168, July, 1988.
- [2] K.Fukushima, S.Miyake and Takayukiito, "A Neural Network for Visual Pattern Recognition," IEEE COMPUTER, pp. 65-74, March 1988.
- [3] K.Fukushima, "Neural network model for selective attention in visual pattern recognition and associative recall," APPLIED OPTICS, Vol.26, No.23, pp.4985-4992, December, 1988.
- [4] D.E.Rumelhart, G.E. Hinton, and R.J.Williams, "Learning Internal Representations by Error Propagation," PDP, Vol. 1, pp. 318-361, Cambridge, MA:MIT Press, 1986.
- [5] D.E.Rumelhart and J.L.McClelland, eds., "Explorations in PDP. MIT Press/Bradford Books, 1988.
- [6] "신경회로 컴퓨터: 이론, 응용 및 구현," 한국과학기술원 산학협동 공개강좌 교재, 1988.
- [7] 이일병, "한글문자 인식을 위한 Neocognitron 연구," 88 국내외 한국 과학 기술자 학술회의 추계 Workshop 정보 산업 분과 pre-prints, pp. 71-76, 10, 1988.