

필기체 숫자와 비숫자의 인식을 위한 MLP 인식기의 구현 방법에 관한 연구

임길택*

*경주대학교 컴퓨터멀티미디어공학부

A Study on the Implementation Methods of the MLP Recognizer for Handwritten Numerals and Non-Numerals

Kil-Taek Lim*

*School of Computer and Multimedia Engineering, Gyeongju University

E-mail : ktlim@ju.ac.kr

요 약

본 논문은 MLP (multilayered perceptrons) 를 이용하여 필기체 숫자와 비숫자를 인식할 경우의 MLP 인식기의 구현 방법에 관한 것이다. MLP는 인식률 및 속도, 메모리 요구량 등에 있어서 필기체 숫자인식에 매우 효율적인 인식기로 알려져 있다. 그러나 기존 연구에서의 MLP는 숫자 입력에 대해서만 인식을 하고 있으며 비숫자 입력에 대해서는 인식률, 기각률 및 동작 특성에 대해서는 연구된 바가 거의 없다. 본 논문에서는 숫자와 비숫자가 혼재하는 경우의 MLP 인식기의 구현방법에 대해서 논한다. MLP 인식기는 세 가지 방법으로 구현되며, 세 가지의 오류 유형을 정의하여 각 인식 방법의 인식 특성을 분석하였다. 인식 실험은 약 63,000여자의 필기체 숫자와 비숫자를 이용하여 이루어지며, 세가지 오류 유형의 측면에서 숫자와 비숫자에 대한 가장 적절한 인식 방법이 논의된다.

ABSTRACT

This paper describes the implementation methods of the MLP (multilayer perceptrons) recognizers for numerals and non-numerals. The MLP has known to be a very efficient classifier to recognize handwritten numerals in terms of recognition accuracy, speed, and memory requirements. The MLP in the previous researches, however, focuses on the only numeral inputs and does not pay attention to non-numeral inputs with respect to recognition accuracy, rejection rates, and other characteristics. In this paper, we present some implementation methods of the MLP in the environments that numeral and non-numerals are mixed. The MLP had been developed by three methods, and investigated with three error types introduced. The experiments had been conducted on a total of about 63,000 numerals and non-numerals. The promising method to recognize numeral and non-numerals is described in terms of the three error types.

키워드

숫자인식, MLP, 인식률, 기각률

I. 서 론

필기체 숫자인식은 우편물의 자동처리, 전장표 자동처리 등 각종 문서의 자동처리에 있어서 필수적으로 요구되는 기반 기술로 이에 대한 많은 연구가 진행되어왔다[1-4]. 숫자인식은 인식하고자 하는 대상 숫자를 추출한 후 숫자인식기에 입력을 하는 과정을 거친다. 이러한 숫자의 추출 과정에서 숫자는 다른 숫자 또는 문자 등과 접촉되어 있지 않거나, 다른 숫자와 약간만 접촉된 상태로 존재하는 경우가 많다. 따라서 인식 대상 숫자의 추출과정이 매우 용이하다고 가정하여 인식기는 온전히 분리된 숫자를 인식하게 된다. 기존의 대

부분의 연구는 이렇게 온전히 분리된 고립 숫자의 연구 [1,2,3]에 초점이 맞추어져 있으며 분리가 제대로 되지 않거나 숫자가 아닌 비숫자 입력에 대한 연구는 거의 없다.

본 논문은 숫자와 비숫자가 혼재하는 환경에서 MLP [5]로 숫자인식기를 구현하는 방법에 관한 것이다. MLP를 이용한 숫자인식 방법을 제시하며 각 방법의 특성을 분석하여 적절한 방법을 제안하였다. 각 방법의 체계적 분석을 위해 세 가지 종류의 오류 유형을 정의하며 이를 이용함으로써 기존의 분석 방법에 비해 인식기의 인식 특성을 보다 명확하게 분석하였다. 제시한 방법에 따른 숫자인식기의 구현 및 인식 실험은 약 53,000여자

의 숫자 영상 데이터와 14,000여자의 비슷자 영상 데이터를 이용하여 이루어졌다. 실험 결과, 제한한 MLP 숫자인식기가 숫자와 비슷자가 혼재하는 경우에 적절하게 사용될 수 있음을 알 수 있었다.

II. 인식기 구조 및 오류 유형

1. 인식기 구조

본 연구에서는 MLP를 이용하여 숫자뿐만 아니라 비슷자도 처리하고자 한다. 입력이 숫자라면 0~9까지의 숫자 중 하나를 결과로 출력하고 비슷자이면 비슷자로 인식하여야 한다. 비슷자는 숫자가 아닌 것이므로 기각을 하는 것을 제대로 인식한 것으로 한다. 이를 그림 1에 나타낸 것과 같은 두 가지의 방법으로 구현한다. 그림 1의 각 인식기의 구조를 그림 2에 나타내었다.

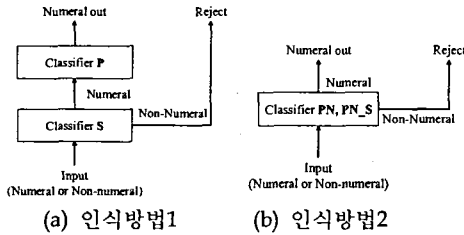


그림 1. 숫자와 비슷자의 인식 방법

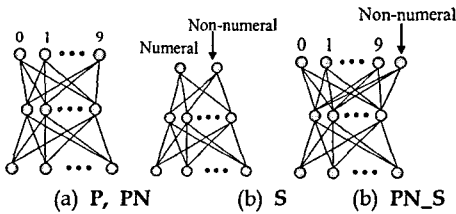


그림 2. 각 인식기의 구조

인식방법1에서 S는 입력이 숫자 클래스인지 비슷자 클래스인지를 구분하는 이진분류기의 역할을 한다. 만약 S가 입력을 숫자 클래스로 분류한다면 입력은 P로 전달되고 P는 0~9의 숫자 중 하나를 출력하게 된다. 여기서 P는 숫자 클래스들만을 대상으로 학습하며, S는 숫자 클래스와 비슷자 클래스들을 모두 이용하여 구현된다. 인식방법2에서 PN과 PN_S도 S와 마찬가지로 숫자 클래스와 비슷자 클래스들을 모두 이용하여 학습을 한다. PN은 최고 출력값이 임계치를 넘으면 숫자로 받아들이고 그렇지 않다면 비슷자 클래스로 인정하여 기각을 한다. PN_S의 경우는 출력층에 기각노드를 별도로 두어 기각노드에 해당하는 출력값이 가장 클 경우 비슷자 클래스로 인식을 한다. S, P, PN, PN_S 모두 하나의 중간층을 가

지는 MLP로 구현되는데, P와 PN은 10개의 출력노드, PN_S는 11개의 출력노드, 그리고 S는 두 개의 출력노드로 구성되어 있다.

2. 세 가지 오류 유형

일반적으로 인식기의 인식성능은 오류율과 기각률의 관점에서만 논의되는데 오류율과 기각률의 관점에서만 측정할 경우 인식대상 숫자의 추출이 완벽하지 않아서 발생하는 잘못된 입력에 대한 정확도의 측정이 어렵다. 따라서 이러한 오류율과 기각률을 포함하여 본 논문에서는 표 1과 같이 인식기의 오류 유형을 세 가지로 정의한다. Type1 오류는 입력이 숫자 클래스이고 출력도 숫자 클래스이지만 잘못된 클래스로 인식을 하는 경우에 해당한다. Type2 오류는 입력이 숫자 클래스이지만 인식점수가 낮아 인식을 거절하는 것을 의미한다. Type3 오류는 비슷자 클래스 입력인데 숫자 클래스로 결정하는 경우의 오류를 의미한다. 이와 같이 오류 유형을 정의하고 분석함으로써 숫자인식기의 정확도를 판단할 수 있다.

표 1. 숫자인식기의 세 가지 오류 유형

오류 유형	입력	출력	비고
Type1 오류	숫자	숫자	인식오류
Type2 오류	숫자	기각	기각오류
Type3 오류	비숫자	숫자	

III. 실험 및 분석

1. 실험 데이터 세트

본 논문에서는 실험을 위하여 4가지 종류의 숫자 영상 데이터 세트를 사용하였다. 학습을 위해서 11,690자의 숫자 영상 데이터 세트와 9,113자의 비슷자 영상 데이터 세트를 사용하였다. 학습된 인식기의 테스트를 위해서 41,193자의 숫자 영상 데이터 세트와 4,709자의 비슷자 영상 데이터 세트를 사용하였다. 숫자 영상과 비슷자 영상의 예를 그림 3에 나타내었다. 비슷자 영상 데이터 세트는 필기체로 작성된 우리나라의 우편봉투에 존재하는 주소 문자열에서 숫자간의 접촉에 의해 숫자 분리에 실패한 영상과 한글문자 및 잘못 분리된 한글문자들로 구성되어 있다.

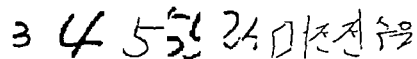


그림 3. 숫자 영상 및 비슷자 영상 예

일반적인 인식방법에서는 인식기에 인식 대상 영상을 그대로 입력하지 않고 인식에 보다 용이한 특징을 추출하여 입력함으로써 영상을 인식한다. 본 논문에서는 영상의 특징으로 [6]에서 제시된 것과 같은 방향각도 특징을 이용한다.

2. 오류 유형별 실험

그림 1과 그림 2에 나타난 방법에서의 P, PN, PN_S, S를 오류역전과 알고리즘[5]을 이용하여 학습하였다. 각 MLP 인식기는 하나의 중간층을 가지며 중간층 노드의 개수를 30, 50, 70, 90개로 달리하여 구현하였다. 학습시의 학습율은 0.1, 관성항은 0.7로 설정하였으며 학습 반복회수는 100회로 제한하였다.

인식기들의 학습을 완료한 후 우선 Type1 오류 유형에 대한 인식기별 성능을 분석하였다. Type1 오류의 분석에 사용한 인식 규칙은

recognition rule : If $\text{argmax}_j O_j = \text{true class}$,
 $j=0,1,\dots,9$, then the recognition result is correct.

와 같다. 인식 규칙에서 O_j 는 j번째 클래스에 해당하는 MLP 출력층 노드의 출력값을 의미한다.

표 2는 숫자 테스트 데이터 세트에 대한 Type1 오류율을 나타낸다. 모든 인식기에서 중간층 노드의 수가 많아질수록 Type1 오류가 감소한다. 각 인식기별로 성능을 비교해보면, 동일한 중간층 노드의 수를 가지는 인식기에서 숫자만 학습한 P의 오류율이 가장 낮으며, PN의 오류율이 PN_S보다 낮다. 그러나 PN과 PN_S의 중간층 노드의 수를 증가시키면 작은 규모의 P의 성능과 유사하게 나타남을 알 수 있다. 약 6.4%의 오류율을 기준으로 보면 P의 경우 30개, PN은 50개 PN_S의 경우 90개의 중간층 노드가 있음을 확인할 수 있다.

표 2. 숫자 테스트 데이터 세트에 대한 Type1 오류율(%)

인식기 \ 중간층 노드 수	30	50	70	90
P	6.44	5.48	5.04	4.73
PN	8.05	6.34	5.75	5.43
PN_S	8.77	8.02	6.88	6.41

다음 실험으로 Type2의 오류율을 분석하였으며 그 결과를 표 3에 나타내었다. Type2 오류는 기각 오류로써 숫자가 입력되었음에도 인식을 하지 못하고 기각을 하는 경우이며 다음과 같은 규칙으로 오류율이 계산된다.

rejection rule1 : If $\max O_j < 0.5$, $j=0,1,\dots,9$, then reject the input.

rejection rule2 : If $\max O_j < O_{10}$, $j=0,1,\dots,9$, then reject the input.

일반적으로 인식기는 인식기의 출력값을 기준으로 인식결정에 대한 확신도를 나타내며 이에 따라 기각 결정을 한다. 본 논문에서 최대 출력값이 0.5보다 낮으면 입력을 기각한다. 이를 기각 규칙1이라 한다. 기각 규칙2는 PN_S에 대해 적용된다. PN_S에서 마지막 열 번째 노드가 비슷자 클래스에 해당하는 노드이다. 따라서 숫자에 해당 하는 처음 10개의 출력층 노드의 출력값의 최고

값이 기각 노드의 출력값 보다 작으면 기각한다. Type2 오류율은 전체 입력된 숫자 중에 기각된 숫자의 비율을 의미한다. 표 3에서 PN_S2에 대한 오류율은 기각 규칙2를 적용했을 경우의 오류율이다. Type2 오류의 측면에서 P가 가장 우수한 성능을 보이고, PN이 PN_S보다 대체적으로 좋은 성능을 보이며, PN_S2가 가장 나쁜 성능을 보인다. P보다 나머지 세 개의 인식기가 Type2 오류율이 높은 것은 단순한 기각 규칙을 적용했을 경우 P보다 나머지 세 개의 인식기가 보다 신중하게 인식을 시도한다는 것을 의미한다.

표 3. 숫자 테스트 데이터 세트에 대한 Type2 오류율(%)

인식기 \ 중간층 노드 수	30	50	70	90
P	3.87	4.32	4.09	4.33
PN	22.77	20.11	21.48	14.62
PN_S	27.27	18.37	18.19	17.58
PN_S2	31.24	21.71	16.71	16.26

표 3의 Type2 오류율에 포함되지 않은 즉, 기각되지 않은 숫자에 대한 인식을 살펴보는 것도 의미있을 것이다. 본 논문에서는 다음과 같이 Type1*오류율을 정의하여 인식기별로 오류율을 분석하였으며, 그 결과를 표 4에 나타내었다. 이 측정값들은 인식기의 인식결정에 대한 신뢰도를 반영한다.

Type1* error (%) = the number of misrecognized numerals × 100 / the total number of not-rejected numeral inputs

Type1* 오류의 측면에서 PN이 가장 우수한 성능을 보였으며, P가 가장 좋지 않은 성능을 보였다. 이는 Type1 오류의 측면에서는 P가 가장 좋은 성능을 보인 것과 다른 특성이다. 즉, 기각을 하지 않고 인식 결정을 하였을 경우의 신뢰도는 PN이 가장 좋다는 것을 알 수 있다. 인식결과와 신뢰도가 중요한 전장표 처리, 우편물 구분 등의 경우에는 기각이 없을 때 최고 인식성능을 나타내는 인식기보다는 기각을 적정 수준으로 하더라도 신뢰도 높은 인식기가 보다 더 효과적이다. 논문에 나타내지 않았지만 기각 임계치를 0.5로 하지 않고 변경하여 실험하였을 때도 동일한 양상이 나타났다.

표 4. 숫자 테스트 데이터 세트에 대한 Type1* 오류율(%)

인식기 \ 중간층 노드 수	30	50	70	90
P	4.51	3.53	3.19	2.74
PN	2.24	1.74	1.32	1.71
PN_S	2.10	2.35	1.96	1.73

Type3 오류의 측면에서 실험한 결과를 표 5에 나타내었다. 이 오류는 비슷자 입력을 숫자인 것으로 잘

못 받아들여 인식한 오류이며 다음의 인정 규칙에 의해 구해진다. 인정 규칙2는 PN_S의 마지막 노드의 출력값을 이용한 것이다.

acceptance rule 1 : $If \max O_j > 0.5, j=0,1,\dots,9,$
then accept the input.

acceptance rule 2 : $If \max O_j > O_{10}, j=0,1,\dots,9,$
then accept the input.

표 5. 비슷자 테스트 데이터 세트에 대한 Type3 오류율(%)

인식기 \ 중간층 노드 수	30	50	70	90
P	74.47	77.15	75.60	76.32
PN	1.85	2.55	1.85	2.78
PN_S	1.32	2.91	1.83	1.89
PN_S2	0.87	1.57	1.81	1.57

P의 Type3 오류율은 최소 약 74%이상의 오류를 보인다. 이는 비슷자 입력에 대해 거의 학습하지 못한 인식기는 비슷자 입력에 대한 기각 신뢰도가 현저히 낮아진다는 것을 의미한다. 반면 PN과 PN_S는 1%~3%내외의 낮은 오류율을 보여 안정적으로 기각을 잘 하는 것으로 나타났다.

Type1 오류가 최적인 P의 장점을 취하고자 그림 2의 방법 1에서와 같이 숫자와 비슷자만을 구분하는 S를 도입하는 경우는 어떠한 양상이 나타나는지 검토해 볼 가치가 있다. 그림 2에 나타낸 S의 Type2, Type3 오류율에 대한 실험 결과를 표 6에 나타내었다. 실험 결과 비슷자에 대한 오류는 적지만 숫자에 대한 오류는 매우 많다. 숫자인 것으로 올바르게 분류한 경우에 정확한 숫자 클래스를 인식할 수 있는 지에 대해서도 실험해 보았다. S에서 바르게 분류된 숫자 입력을 P에 입력하여 인식한 결과 성능을 표 7에 나타내었다. 이 오류율과 표 4에 나타낸 P의 Type1* 오류율을 비교해보면 거의 유사하며, PN과 PN_S보다는 나쁜 결과임을 알 수 있다.

표 6. S의 Type2, Type3 오류율(%) - Type2 오류율은 숫자 데이터세트, Type3 오류율은 비슷자 데이터세트에 대한 오류율임.

오류유형 \ 중간층 노드 수	30	50	70	90
Type2	30.77	28.31	30.29	31.44
Type3	1.76	1.74	1.57	1.53

표 7. 숫자 데이터세트에 대한 S+P의 Type1* 오류율(%) - S는 50개의 중간층을 노드를 가짐.

오류유형 \ 중간층 노드 수	30	50	70	90
Type1*	4.42	3.62	3.19	3.03

VI. 결 론

본 논문에서는 숫자와 비슷자가 혼재하는 경우의 MLP 인식기의 구현방법에 대해서 논하였다. MLP 인식기는 세 가지 방법으로 구현하였으며, 세 가지의 오류 유형을 정의하여 각 인식 방법의 인식 특성을 분석하였다. 인식 실험은 약 63,000 여자의 필기체 숫자와 비슷자를 이용하여 이루어 졌다.

본 논문의 실험 결과를 토대로 다음과 같은 결론을 내릴 수 있다. 비슷자 환경에서도 Type1 오류만을 최소화시키고자 하는 경우에는 기존에 숫자만으로 학습된 인식기 P를 사용하는 것이 더 낫다. P의 장점을 취하고자 숫자와 비슷자, 두 가지를 구분하는 새로운 인식기 S를 도입하더라도 S의 오류율이 높아 P 단독으로 사용하는 것보다 나쁜 결과를 초래할 가능성이 높다. 입력을 기각하지 않고 인식결정을 내린 경우에는 숫자와 비슷자를 동시에 학습한 PN의 인식결정이 신뢰도가 가장 높으며, 출력값 특성도 후처리에 유리하게 나타난다. PN의 출력층에 비슷자를 담당하는 기각 노드를 새로이 추가한 PN_S도 PN과 동작 특성이 유사하며 특별히 우수한 장점이 없다.

참고문헌

- [1] S. Cho, "Neural-network classifiers for recognizing totally unconstrained handwritten numerals," *IEEE Transactions on Neural Networks*, vol. 8, pp. 43-53, 1997.
- [2] F. Kimura and M. Shridhar, "Handwritten numeral recognition based on multiple algorithms," *Pattern Recognition*, vol. 24, no. 10, pp. 976-983, 1991.
- [3] C. Y. Suen, C. Nadal, R. Legault, T. A. Mai, and L. Lam, "Computer recognition of unconstrained handwritten numerals," *Proc. of IEEE*, vol. 80, no. 7, pp. 1162-1180, 1992.
- [4] 임길택, 김호연, 이상호, 송재관, 남윤석, "우편물 자동구분을 위한 문자인식 시스템", *대한전자공학회 컴퓨터/반도체 소사이터 추계학술대회 논문집*, 제 25권 제 2호, pp. 103-106, 2002.
- [5] D. E. Rumelhart, G.E. Hinton, and R.J. Williams, "Learning Internal Representations by Error Propagation," *Parallel Distributed Processing*, vol. 1, Eds. D. E. Rumelhart and J. L. McClelland, pp. 319-362, 1986.
- [6] 김호연, 임길택, 김두식, 남윤석, "서장 우편물 자동처리를 위한 우편영상 인식 시스템," *정보처리학회논문지*, Vol.10 No.4, pp.429-442, 2003.