

# 규칙적인 잡음을 이용한 인쇄체 한글 자소인식 개선

이진수, 권오준, 방승양  
포항공과대학교 전자계산학과 컴퓨터시스템 연구실

## Improvement of The Printed Korean Grapheme Recognition using Meaningful Noises

Jin-Soo Lee, Oh-Jun Kwon, Sung-Yang Bang  
Dept. of Computer Science & Engineering, Postech

### 요 약

한글은 문자수가 많고 초성, 중성, 종성의 조합으로 이루어진 2차원적인 특성 때문에, 신경망을 이용한 한글 인식의 경우에는 자소를 분리한 후 자소별로 인식하는 방법이 많이 사용된다. 이러한 방법의 경우 분리된 자소영역에 원하는 자소 이외의 부분이 첨가되면 학습이 어려워 오인식의 주된 원인이 되기 때문에, 정확한 자소분리 알고리즘이나 전처리들을 통하여 그러한 잡음을 없애려는 시도가 많이 있었으나 아직도 원하는 자소부분만을 정확히 분리하는 것은 어려운 문제로 남아있다. 본 논문에서는 그러한 잡음이 규칙적임을 이용하여, 필요한 자소영역만을 추출하려하기보다는 오히려 필요한 자소영역 외의 부분을 포함시킴으로써, 잡음이라고만 생각했던 부분을 하나의 정보로 역이용하여 이로 인한 여러 오인식 경우를 해결하였다. 또한 자소의 위치가 불규칙적인 부분에 있어서, 그 위치를 고정시키는 알고리즘을 사용하여 인식률을 더욱 높였다.

### 1. 서론

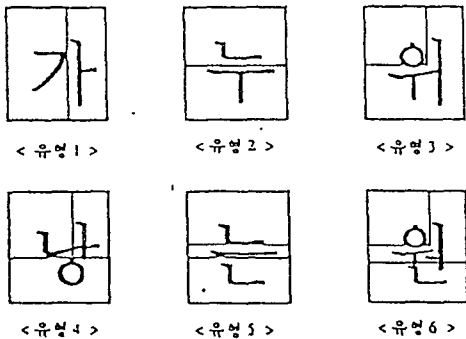
문자인식은 오래 전부터 연구되어온 분야로서 그 동안 많은 연구가 이루어져 왔다. 특히 근래에는 인쇄체 문자 인식을 위한 상용화된 제품도 많이 나와 있는 등 사실상 실용화되고 있는 실정이며, 많은 연구자들이 인쇄체 문자보다는 필기체 문자인식에 더 많은 관심을 갖고 있다. 하지만 그러한 실정에도 불구하고 한글의 경우 대부분의 문자인식 제품들이 선전과는 달리 다양한 폰트와 크기에 대해서 사용자가 만족할 만큼의 높은 인식률을 보여주지 못하기 때문에 아직 실생활에 많이 사용되지는 못하고 있다. 따라서 실질적인 한글 문자 인식의 실용화를 위해서는 다양한 폰트와 크기의 한글 문자에 대해서 매우 높은 인식률을 보장하는 인식 모델이 요구된다. 본 연구에서는 이러한 목적으로 비교적 좋은 결과를 보여준 기존 방법[4]을 기반으로 하여, 높은 인식률을 갖는 다중폰트 다중크기 인쇄체 한글인식 모델을 제안한다.

문자인식은 크게 통계적 방법, 구조적 방법, 신경망 이용 방법으로 나눌 수 있다. 통계적 방법은 같은 문자를 모아 통계적으로 공통된 특징을 추출한 후 인식할 새로운 문자의 특징과 비교하는 방법이다. 이러한 방법은 글자 모양이 일정한 경우에는 좋지만 다양한 변형을 가진 문자를 인식하거나 비교대상이 많은 경우에는 그 특

징의 분포를 알아내기 어려워 적용하기 어려운 단점이 있다. 구조적 방법은 문자의 선분, 획 등의 구조적 특성을 이용하여 인식하는 방법으로서, 구조적 특징을 이루는 기본 특징만 정확하게 추출할 수 있다면 좋은 방법이 될 수 있다. 하지만 이 방법은 폰트가 다양해지고 특징을 추출하기까지의 과정에서 생기는 잡음 등으로 인해 정확한 특징을 추출하기가 힘들다. 이러한 이유로 근래에는 잡음에 강한 특성을 갖는 신경망을 이용한 방법이 많이 사용되고 있다[2][3][4]. 하지만 신경망을 이용하는 방법은 입력 특징소나 신경망 모델에 따라 그 성능이 매우 달라지기 때문에 주어진 문제에 가장 적합한 입력 특징소와 신경망 모델을 선택해야 좋은 결과를 얻을 수 있다.

한글은 문자수가 많고 초성, 중성, 종성이 조합되어 하나의 문자를 이루는 2차원적인 특성을 갖고 있기 때문에, 신경망을 이용한 한글 인식의 경우에는 자소를 분리한 후 자소별로 인식하는 방법이 많이 사용된다 [1][3][4]. 이와 같이 한글의 자소를 분리한 후 각 자소별로 인식할 경우 문제의 크기가 축소되어 인식에 유리하다. 하지만 각 자소간의 위치가 약간씩 다르고, 자소간의 연결 또한 매우 불규칙적이어서, 구조적 특징들을 사용하여 원하는 자소부분만을 분리하는 것이 힘들다 [1]. 그래서 한글을 일정한 자소영역을 갖는 6가지 유형

[그림 1]로 분리한 후 각 자소 영역별로 신경망을 이용하는 방법을 사용하여 비교적 좋은 결과를 얻은 연구가 보고된 바 있다[4]. 이 방법의 경우 대부분의 오인식이 모음에서 나타났는데 이것은 자소영역 안에 다른 자소의 일부가 잡음으로 들어왔기 때문이다. 모음의 경우에 작은 모음 영역에 비해 그 잡음의 양이 많아 신경망이 제대로 잡음으로 처리하기가 어렵다. 이러한 잡음의 일부는 전처리 등으로 처리할 수 있으나, 대부분 일그러짐에 의해 원래의 자소와 잡음간의 연결 형태가 불규칙적이어서 전처리를 통해 원하는 자소만을 추출하는 것은 사실상 힘들다. 만일 이러한 문제를 해결하여 대부분의 모음의 오인식 경우를 없앨 수 있다면 높은 인식률을 얻을 수 있을 것이다. 따라서 본 연구에서는 유형별 인식 모델을 기반으로 하여, 그러한 잡음의 영향을 없애므로써 보다 높은 인식률을 갖는 인식 모델을 구현하였다. 잡음의 영향을 없애기 위해서, 필요한 자소영역만을 추출하려고 하기 보다는 오히려 필요한 자소영역 외의 부분을 포함시킴으로써, 잡음이라고만 생각했던 부분을 하나의 정보로 역이용하여 위의 문제를 해결하였다.

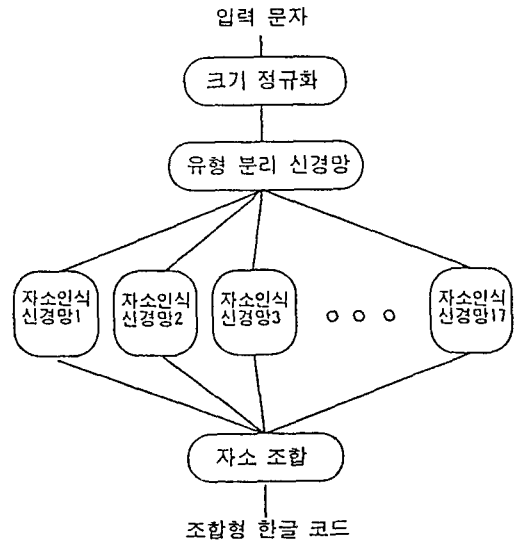


[그림 1] 한글의 6가지 유형

## 2. 신경망을 이용한 인체체 한글 인식기의 구현

유형 분류 신경망과 자소 인식 신경망으로 이루어진 계층적 신경망[4]을 기본으로 구현한 신경망 모델의 기본 구조는 [그림 2]와 같다. 인식과정은 먼저 입력 노드수에 맞도록 크기 정규화를 거친 후 특별한 전처리는 하지 않고 바로 입력 피쳐를 추출한다. 입력 피쳐는 망피쳐(mesh)[5]를 사용하였으며 망피쳐의 윈도우 크기는 각 신경망의 종류마다 다른데 글자의 조밀도와 문제의 복잡도에 따라 3\*3, 3\*2, 또는 2\*2로 정하였다. 추출된 입력 피쳐는 한글의 6가지 유형중 하나를 분류하는 유형분류 신경망을 거치게 되고, 여기에서의 결과에 따라 각 유형별 자소영역을 나누어 해당 영역에서 다시 입력 피쳐를 추출한다. 이들은 각각 자소 신경망의 입력이 되고 해당되는 자소인식 신경망은 자소를 인식한 결과를 출력하게 된다. 위와 같이 계층적인 신경망을 사용할 경우 앞 단계, 즉 유형분류 신경망에서의 작은 오인식이 뒤의 자소

인식에 큰 영향을 주어 전체적으로 많은 인식을 저하를 가져올 수 있다. 이를 막기 위하여 본 모델에서는 [그림 3]과 같이 인식된 자소의 출력치와 유형의 출력치를 이용하여 후보 유형도 고려함으로써, 앞서 유형분류가 잘못되었다고 하더라도 다음 자소인식 단계에서 인식 결과의 확실정도에 따라 다시 올바른 유형으로 인식할 수 있도록 하였다. 마지막으로 각 자소의 인식 결과는 2바이트 조합형 코드로 조합되어 한글로 출력하게 된다.



[그림 2] 신경망 모델의 기본구조

```

Do {
  if(jaso_output < threshold1
    .or. type_output < threshold2)
  {
    new_type = next candidate;
    recognition (new type);
  }
  else
    break;
}while(iteration_number < type_number)

jaso_output : 자소인식 신경망의 해당 출력치
type_output : 유형 분류 신경망의 해당 출력치
  
```

[그림 3] 후보 유형을 이용한 재인식 알고리즘

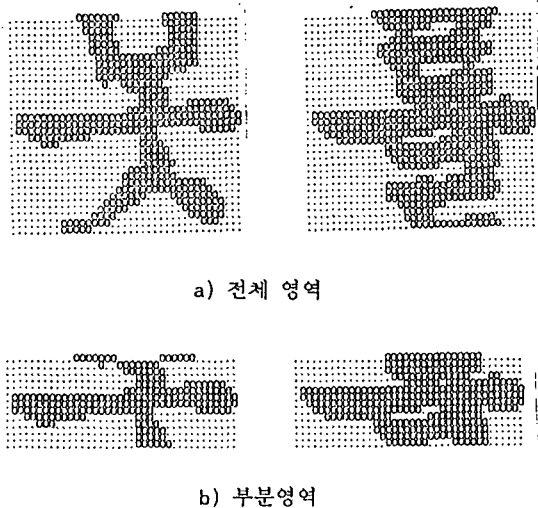
## 3. 규칙적인 잡음을 이용한 개선된 한글 자소 인식 방법

### 3.1 규칙적인 잡음정보의 활용

인체체 문자인식에서 잡음이란 일반적으로 스캐너로 이미지를 받는 과정이나 또는 크기 정규화 과정에서 생기는 원하지 않는 픽셀들이나 이미지의 일그러짐을 말한다. 이러한 잡음은 구조적 방법이나 통계적 방법을 사용

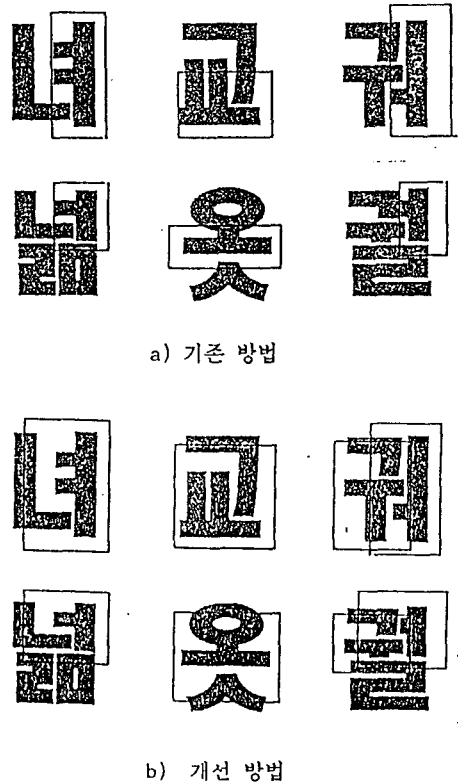
할 때 오인식의 주된 원인이 되며, 또 구조적 특성을 이용하여 자소를 분리할 때도 어려움을 주는 요소이다. 또 다른 잡음으로는 원래의 문자 이미지의 일부분인데도 불구하고 원하지 않는 부분이기에 때문에 잡음이 되는 경우가 있다. 앞에서도 언급하였듯이 고정된 자소영역을 분리할 경우 원하는 자소 이외의 부분이 같이 들어오게 된다. 이러한 잡음은 일반적으로 말하는 잡음과는 다른 것으로 위의 방법을 사용할 경우 오인식의 주된 원인이 된다. 이것은 특히 모음의 경우, 잡음이 차지하는 비율이 인식에 필요한 특징요소 보다도 크기 때문에 더욱 큰 영향을 미친다. 이러한 잡음은 다음과 같은 특성을 갖는데, 같은 폰트 내에서 같은 글자라면 같은 잡음이 들어오게 된다. 즉 학습 데이터의 한 문자에서 자소 분리 시에 생기는 해당 자소 이외의 부분은, 새로운 데이터의 같은 문자를 자소 분리할 때도 같은 형태로 생기게 된다. 하지만 이러한 잡음 역시 일반적인 의미의 잡음으로 인하여 일그러지게 되어 규칙성을 잃게 되고 결국에는 커다란 잡음의 역할을 하게 된다. 본 연구에서는 잡음의 이러한 특성을 이용하여 잡음의 규칙성을 살리어 정보로 이용하였다.

사람의 경우 원하는 자소영역만을 보았을 경우 위의 잡음들로 인하여 인식하기가 어려우나 문자 전체를 보면 쉽게 원하는 자소를 알 수 있는 경우가 있다. [그림 4]의 a)에서는 모음이 'ㅛ'와 'ㅡ'라는 것을 바로 알 수 있으나 b)에서는 'ㅛ'나 'ㅑ', 'ㅡ'의 구분이 a)에서처럼 쉽지 않다. a)에서 '옷'의 경우 중성이 'ㅅ'임을 알고 있으므로 수평성분 아래에 붙어있는 선분은 'ㅑ'가 아닌 'ㅅ'에서 온 잡음임을 짐작할 수 있으며, '를'의 경우에는 모음 부분에 잡음이 심하더라도 전체적인 형태를 볼 때 모음이 'ㅡ'임을 알 수 있다. 하지만 b)의 경우에는 수평 성분 위아래, 모두가 비슷한 모양을 하고 있기 때문에 모음을 인식하기가 힘들다. 이것은 신경망을 사용한 문자인식의 경우에도 마찬가지로 적용된다.



[그림 4] 문자의 부분 영역과 전체 영역

기존 방법[4]에서는 [그림 5]의 a)와 같이 필요한 자소영역만을 입력부분으로 사용하였다. 이 경우 앞에서 설명한 바와 같이, 원하는 자소 이외의 부분이 입력으로 들어오고, 같은 문자라 해도 폰트마다 적합한 자소영역이 조금씩 다르기 때문에, 그러한 잡음의 모양과 크기는 폰트마다 다양하여 인식률이 저하된다. 이것은 주로 모음의 경우에 나타나게 되는데, 이전에 보고되었던 결과[4]에 의하면 전체 오인식 문자의 65% 이상이 모음에서 나타났고, 본 실험에서도 [표 1]과 같이 전체 오인식 문자 중 65.68%가 모음에서 나타났다. [그림 5]의 b)는 모음의 인식을 위하여 필요한 영역 이외의 영역까지 입력으로 한 경우이다. 자소를 정확히 분리할 수 있거나 일부분이 잡음으로 들어오더라도 모음의 인식에 영향을 주지 않는 부분은 포함시키지 않는다. 예를 들면 유형 4에서 중성 부분은 폰트와 관계없이 어느 정도 정확히 분리할 수 있으며, 유형 3에서 세로모음은 가로 모음을 분류하기 위한 특징에 영향을 주지 않으므로 포함시키지 않는다. 하지만 각 유형의 초성과 같이, 다른 자소의 일부분이 모음의 일부처럼 보일 가능성이 있는 자소에 대해서는 그 자소의 전체를 모음 영역 안에 포함시킨다. 이렇게 함으로써 신경망이 그러한 자소의 일부분을 모음의 일부로 잘못 학습하는 경우를 막을 수 있다. 즉 특정한 자소가 어떤 모양의 잡음으로 나타나는 지를 학습함으로써 모음의 일부처럼 보일 수 있는 잡음을 신경망 스스로가 처리할 수 있는 것이다.



[그림 5] 기존 방법과 개선 방법에서의 모음영역 분리

	자음	모음
고딕체	5	11
명조체	2	11
신명조체	11	29
중고딕체	13	21
궁서체	27	39
전체	58	111

[표 1] 기존 방법에서의 모음과 자음의 오인식 자소수

5번째 유형 문자는 모음 중에서도 가장 인식하기 어려운 문자이다. 이것은 초성과 중성에 따라 수평모음의 위치가 매우 다양하게 변화되고 또한 모음의 영역이 문자에서 차지하는 비율이 상대적으로 작기 때문이다. 따라서 모음부분에서 수평성분의 위치를 고정시켜줄 필요가 있다. 이를 위해 5번 유형의 문자일 경우 수평성분의 세로위치를 찾아내어 그 위치를 기준으로 자소영역을 분리하는 알고리즘을 사용한다[그림 6].

```
for(y = 0; y < Y_MAX; y++)
  if(horizontal line exists at total_image[y])
  {
    초성 하단좌표 = y + Δy1;
    중성 상단좌표 = y + Δy2;
    중성 상단좌표 = y + Δy1';
    중성 하단좌표 = y + Δy2';
  }
  초성 영역 size normalization;
  중성 영역 size normalization;
```

[그림 6] 수평모음 위치 고정 알고리즘

#### 4. 실험 결과 및 분석

기존 방법과 개선된 방법의 인식률 비교가 [표 2]에 나타나 있다. 실험 대상은 많이 사용되면서도 모양이 상이하게 다른 명조체, 신명조체, 고딕체, 중고딕체, 궁서체에 대해서 사용 빈도수가 큰 순서로 1405 글자를 사용하였다[10]. 유형 오류는 단순한 유형분류에서의 오류가 아닌, 자소인식을 모두 거친 후 최종 인식 결과에서의 유형 오류를 말한다. 따라서 자소 인식률이 높으면 유형 분류가 잘못되었다고 하더라도 다시 바로잡을 수 있으므로 유형 오류도 줄어들게 된다. 결과의 객관성을 위해서 오인식의 대부분을 차지하는 4번과 5번 유형에 대해서는 은닉노드 수등의 변수를 달리하면서 여러 번 학습한 후 각각 테스트한 결과를 평균하였다.

	인식률	1유형	2유형	3유형	4유형	5유형	6유형	유형오류
고딕체	98.86%	2	1	0	8	3	0	2
명조체	99.07%	0	1	0	9	2	0	1
신명조체	97.15%	1	0	4	18	9	2	6
중고딕체	97.58%	0	2	2	13	11	0	6
궁서체	95.30%	4	0	3	21	29	2	7
전체	97.59%	7	4	9	69	54	4	22

a) 기존 방법에서의 인식률 및 유형별 오인식 자소수

	인식률	1유형	2유형	3유형	4유형	5유형	6유형	유형오류
고딕체	99.50%	1	1	0	2	2	0	1
명조체	99.72%	0	1	0	1	1	0	1
신명조체	98.65%	0	0	1	8	3	2	5
중고딕체	98.86%	0	2	0	6	2	0	6
궁서체	97.44%	3	0	2	12	10	2	7
전체	98.83%	4	4	3	29	18	4	20

b) 개선 방법에서의 인식률 및 유형별 오인식 자소수

	기존 방법	개선 방법
고딕체	11	3
명조체	11	2
신명조체	29	9
중고딕체	21	3
궁서체	39	9
전체	111	26

c) 폰트별 모음의 오인식 수 비교

[표 2] 기존 방법과 개선 방법의 인식률 비교

ㅏ->ㅑ 1	ㅓ->ㅕ 1	ㅗ->ㅛ 2
ㅓ->ㅕ 7	ㅓ->ㅕ 2	ㅓ->ㅕ 1
ㅓ->ㅕ 1	ㅓ->ㅕ 2	ㅓ->ㅕ 1
ㅓ->ㅕ 1	ㅓ->ㅕ 1	ㅓ->ㅕ 1
ㅓ->ㅕ 1	ㅓ->ㅕ 1	
ㅓ->ㅕ 2	ㅓ->ㅕ 1	

[표 3] 오인식된 모음

[표 2]에서 알 수 있듯이 전체 오인식 중에서 평균 50.30%가 개선되었고, 특히 모음의 경우에는 약 76.58%가 개선되었다. 폰트별로는 명조체와 고딕체가 비교적 높은 인식률을 보이기는 하지만, 기존연구와는 달리 각 폰트들이 비교적 균일하게 높은 인식률을 보였다. 특히 궁서체의 경우 많은 인식률 향상을 보였는데, 이것은 궁서체의 모음이 곡선 모양이라 자소의 위치와 모양이, 다른 폰트에 비해 많이 틀리기 때문에 기존 방법에서는 인식이 어려웠으나, 개선된 방법에서는 새로운 자소의 위치와 모양에 비교적 덜 민감하기 때문으로 보인다. 하지만 아직도 오인식된 모음이 남아있는데 [표 3]은 오인식된 모음의 종류와 수를 나타낸다.

인식해야 할 폰트가 다양할수록 각 자소가 차지하는 영역 또한 다양해진다. 따라서 기존의 방법의 경우에는 인식 대상 폰트가 증가할 경우 자소 영역을 다시 잡아주어야 하고 인식률도 저하된다. 하지만 본 방법은 자소영역을 따로 잡을 필요 없이 기존 신경망을 사용해 학습만 시키면 되고, 또한 궁서체등과 같이 자소의 위치가 많이 다른 문자를 학습하더라도 인식률이 영향을 덜 받는다.

#### 5. 결론

영어 등에 비해 한글은 초성, 중성, 종성의 조합으로 한 문자를 이루는 2차원적인 특성 때문에, 문자의 수가 많고 문자 인식이 어렵다. 특히 폰트가 다양해지면서 인식률은 매우 저하된다. 본 연구에서는 기존의 방법들 중

신경망을 이용하여 좋은 결과를 보여주었던 계층적 신경망 모델[4]의 오인식 원인을 분석하고 이를 개선하는 알고리즘을 제안하고, 제안된 알고리즘을 이용하여 높은 인식률을 갖는 다중폰트 다중크기 인쇄체 한글 인식기를 구현하였다. 자소를 분리한 후 신경망을 이용하여 자소 별로 인식하는 알고리즘을 사용할 경우 지금까지는 되도록 원하는 자소 부분만을 분리하려고 노력하였으나 완벽하게 해당 자소 부분만을 분리하는 것이 힘들므로 오인식의 주된 원인이 되어왔다. 본 연구에서는 이러한 잡음의 영향을 막기 위해 해당 자소의 주변 이미지가 학습시킴으로써 잡음의 규칙적인 특성을 정보로 이용하였다. 이 방법으로 모음의 오인식을 대부분 없앨 수 있었고, 다양한 폰트가 같이 학습되었어도 높은 인식률을 유지할 수 있었다.

[10] 이강섭, 이기성, "한글 출력코드 체계에 관한 기초 연구" 제 4회 한글 및 한국어 정보처리 학술발표 논문집, pp.285-291, 1992.

< 감사의 글 >

본 연구는 한국통신 장기기초 연구 과제 95-55의 지원아래 수행되었습니다.

참 고 문 헌

[1] 도정인, "인쇄체 한글 문자의 인식을 위한 자소 분리에 관한 연구" 한국 정보과학회 가을 학술발표 논문지, pp.175-178, 제 17권 제 2호, 1990.  
 [2] 김상우, 김윤호, 최종호, "신경회로망을 이용한 인쇄체 한글 문자의 인식" 전자공학회 논문지, pp.228-234, 제 27권 제 2호, Feb. 1990.  
 [3] 장석진, 강선미, 김혁구, 노우식, 김덕진, "자소 인식 신경망을 이용한 한글 문자 인식에 관한 연구" 전자공학회 논문지, pp.81-87, 제 31권 제 1호, Jan. 1994.  
 [4] 권재욱, 조성배, 김진형, "계층적 신경망을 이용한 다중 크기의 다중 활자체 한글문서 인식". 한국정보과학회 논문지, pp.69-79, 제 19권 제 1호, Jan. 1992.  
 [5] 김우태, 윤병식, 박인규, 진성일, "인쇄체 한글 문자인식을 위한 특징성능의 비교" 한국정보과학회 논문지, pp.1103-1110, 제 20권 제 8호, Aug. 1993.  
 [6] Y.H.Yu, R.F.Simmons, "Descending Epsilon in Back-Propagation : A Technique for Better Generation," in Proceedings of the IEEE IJCNN, Vol. III, pp.167-172. 1990.  
 [7] Rumelhart, D.E., and J.L.McClelland, eds., Parallel Distributed Processing : Explorations in the Microstructure of Cognition, Vol. 1. Cambridge, MA : MIT Press, 1989.  
 [8] Rangachari Anand, Kishan Mehrotra, Chilukuri K. Mohan, and Sanjay Ranka, "Efficient Classification for Multiclass Problems Using Modular Neural Networks," IEEE Transactions on Neural Networks, pp. 117-124, Vol.6.No.1.JANUARY, 1995.  
 [9] LeCun, Y., "Generalization and network design strategies," In Connectionism in Perspective (R. Pfeifer, Z. Schreter, F. Fogelman -Soulie, and L. Steels, eds.), pp. 144-155. Amsterdam: North-Holland.