

# 계층적 신경망을 이용한 다중 크기의 다중활자체 한글문서 인식

## (Multi-size/multi-font Hangul Document Recognition with Hierarchical Neural Networks)

권재욱\* 조성배\* 김진형\*\*  
(Jae-Ook Kwon) (Sung-Bae Cho) (Jin H. Kim)

### 요 약

본 논문에서는 신경망 기법을 이용하여 다중 크기 및 다중활자체로 구성된 한글문서를 인식하였다. 전체 시스템은 3종류의 신경망 — 문자추출 신경망, 유형분류 신경망, 그리고 문자인식 신경망 — 으로 구성된다. 즉 문자결합 신경망과 문자분리 신경망으로 이루어진 문자추출 신경망을 이용하여 문자열로부터 각 문자 영상을 추출한 후, 유형분류 신경망에서 문자 영상을 6가지의 유형으로 분류하고, 해당 유형을 처리하는 문자인식 신경망에서 실제 문자를 인식하도록 구성되었다. 또한 문자인식 신경망을 학습시키기 위하여 descending epsilon 방법을 이용한 backpropagation 알고리즘을 사용하였는데, 이 방법은 신경망을 모든 입력 영상의 모든 출력 노드에 대해 고르게 학습시키기 때문에 다중활자체의 변형을 흡수할 수 있다. 그 결과 8가지 활자체의 사용빈도수가 높은 한글 520자를 94% 이상, 그리고 다중활자체로 구성된 한글문서를 94.5% 인식함으로써 신경망이 다중 크기의 다중활자체 한글문서를 효과적으로 처리할 수 있음을 보였다.

### ABSTRACT

This paper describes an attempt to develop a multi-font and multi-size Hangul recognition system. The system consists of three kinds of networks: a character extraction network, a type classification network, and six character recognition networks. At first the system extracts each character image from a line image using the character extraction network which is, in turn, composed of a character splitting network and a character merging network. Then, the type classification network classifies the character image into one of the six types. Finally, the corresponding recognition network classifies it into a Hangul character. In order to train the recognition networks,

\*준 회원 한국과학기술원 전산학과

\*\*종신회원 한국과학기술원 전산학과 교수  
접수일자 : 1991년 7월 26일

the backpropagation algorithm is used with descending epsilon. The descending epsilon approach is able to absorb shape variations in various font types because it helps train neural networks tolerable. The proposed system is evaluated by two sets of data, a collection of 520 most frequently used characters in eight font types, and images of a multi-font magazine. The recognition rates are about 94% for the first experiment, and 94.5% for the second one. Judging from the recognition results, we may conclude neural network approach is quite effective in handling multi-font and multi-size Hangul documents

## I. 서 론

대부분의 정보가 전산화되고 있는 현재의 시점에서 문서 자동인력 시스템의 필요성은 날로 증가하고 있다. 이에 따라 인쇄체 문서인식에 관한 연구가 활발히 진행되었고, 그 결과로 단일활자체 한글문서를 96% 이상 인식하는 시스템이 개발된 바 있다[9] 그러나 실제 문서는 여러 크기의 다양한 활자체로 구성되므로 이를 인식하기 위한 다중활자체 한글문서 인식에 관한 연구가 필요하다.

지금까지의 연구에서는 인쇄체 한글을 인식하기 위하여 원형 정합(template matching) 방법, 통계적(statistical) 방법, 그리고 구조 분석적(structural analysis) 방법 등을 사용하였다. 그러나 이러한 방법은 다중활자체의 많은 변형을 처리하기 어렵기 때문에[13, 16], 본 논문에서는 최근에 패턴 인식 분야에서 각광을 받고 있는 신경망을 사용하였다. 신경망은 프로그래머가 문제 해결을 위한 정확한 지식을 갖고 있지 못한 경우에도 그 자체의 적응 능력으로 적절한 해결책을 제공하며, 일반화 능력을 가지므로 여러 형태의 입력에 대해 일관성있는 답을 낼 수 있게 한다[11]. 또한 약간의 잡음이나 손실 등으로 변형된 입력을 처리할 수 있는 결합 극복 능력이 있으므로 다중활자체 한글문서 인식이라는 문제를 해결할 수 있는 하나의 적절한 방법론이라 할 수 있다.

따라서 본 논문에서는 단일활자체 한글 문자를 인식하기 위하여 제안되었던 계층적 신경망[11]을 기본으로 하여 신경망을 구성하고, 이를 다중활자체 한글문서 인식에 적합하도록 Descending Epsilon 알고리즘[17]을 사용하여 학습시킨 다중활자체 한글문서 인식 시스템을 구현하였다. 또한 문서인식시 가장 어려운 과정인 문자 추출부를 신경망을 이용하여 구현함으로써 다중 크기의 다중활자체 한글문서 인식 시스템을 구성하였다.

## II. 배 경

### 2.1 경험적 지식에 의한 문자 추출

문서영상으로부터 문자를 추출하는 것은 문서인식의 전체 과정 중 가장 어려운 부분이다. 그 이유는 문자의 크기가 일정하지 않고, 두 영역이 한 문자를 이룰 수 있으며, 글자체나 인쇄 매체에 따라 문자의 가로와 세로 비율이 다르고, 문자의 영역이 붙거나 겹치기 때문이다[4]. 따라서 그 동안 신문[3, 8, 9], 잡지[10], 프린터로 출력된 문서[4]를 대상으로 많은 연구가 진행되었다.

그러나 이러한 연구는 그 입력 문서가 한글로만 구성되었거나[10], 단일활자체이거나[3, 8, 9], 또는 문자간의 접촉이 발생하지 않는다고 가정하는[4] 등 입력 문서에 대한 많은 제약을 포함하고 있었다. 또한 현재까지의 모든 연구가 기본적으로 수평 투영(horizontal projection)과 수직 투영(vertical projection)에 의하여 분리된(run)에 경험적 지식(heuristic)을 적용하여 문자를 추출하는 방법을 사용하고 있다. 즉 미리 계산된 문자의 가로와 세로 비율을 이용하여 문자의 너비를 추정하거나[3, 9], 분리된 run을 단어 단위로 재구성한 후 각각의 문자를 추출하거나[8], run을 몇 개의 유형으로 분류한 후 각 유형별로 처리하는 방법[4, 10]을 사용하고 있다.

그러나 실제 문서의 경우 각 활자체 별로 가로와 세로의 비율이 다르고 숫자나 기호 등이 섞여 있으므로 모든 활자체에 적합한 임계치를 구하기 어렵다. 실제로 경험적 지식을 사용하여 다중활자체 한글문서에서 문자를 추출하려고 시도해 본 결과 활자체 별 임계치의 차이로 인하여 좋은 결과를 얻지 못하였다.

### 2.2 다중활자체 문자인식에 관한 기존의 연구

본 절에서는 다중활자체 문자를 인식하기 위하여 제

안되었던 기존의 연구에 대하여 영문자, 한자, 그리고 한글의 경우로 나누어 설명한다 먼저 다중활자체 영문자 인식의 경우, 가장 널리 알려진 것은 Pavlidis[15]의 연구와 Shilen[16]의 연구인데 두 방법 모두 입력된 문자영상으로부터 기하학적인 특성(geometrical feature) 및 위상학적인 특성(topological feature)을 추출한 후 이것으로 모델을 만들어 인식하는 방법이다.

Shilen의 연구에서는 하나의 입력 문자영상을  $3 \times 3$ 의 부분영역으로 나눈 후, 각 영역별로 8방향에 대해 직선(straight line), 요철(convex and concave), 그리고 홀이 존재하는가를 조사하여 만든 189개의 이진 특성과 문자의 길이 등 전체적인 정보를 반영하는 8개의 이진 특성을 이용하여 모델을 만든다. 학습시킨 6가지 활자체가 섞여 있는 문서에 대하여 실험한 결과 90%에서 99% 사이의 인식률을 나타냈다. Pavlidis의 연구에서는 입력 영상으로부터 직선, 홀, 호(arc) 등의 특성을 추출한 후, 이 특성을 특성의 위치를 나타내는 패러미터 공간과 특성의 방향 및 크기를 나타내는 패러미터 공간 상의 점으로 변환한다. 변환된 점이 패러미터 공간 상에서 나타나는 위치를 여러 개의 영역으로 분류하여 모델을 만들고 Bayesian 분류자(classifier)를 사용하여 인식한 결과 97% 이상의 인식률을 나타냈다.

한자의 경우는 입력 문자영상에서 기본 자획을 추출하고, 이 자획을 특징점(feature point)의 수와 자획의 길이에 따라 순서대로 나열한 후, 전후 자획간의 위치관계를 이용한 결정 트리 형태의 모델을 만들어 다중활자체 한자를 인식한 Hsu[14]의 연구 등 많은 연구가 있었다. Hsu의 연구에서는 6가지 활자체를 사용하여 실험한 결과 92.42%의 인식률을 얻었다.

한글의 경우에는 다중활자체 문자인식에 관한 연구가 극히 미약한 실정인데 입력 문자영상으로부터 윤곽선을 추출한 후, 추출된 윤곽선을 분석하여 모음을 검출하고 검출된 모음에 따라 한글을 3가지 유형으로 분류한 다음 입력된 문자를 자소별로 분리하여, 분리된 자소를 인식한 이광호[5]의 연구가 자소 분리율 95%, 분리된 자소에 대한 문자 인식률 96%를 얻었다. 그러나 이 방법은 각 활자체에 대한 경험적 지식을 너무 많이 포함하고 있다는 문제점을 가지고 있다. 이외에 다중활자체 한글 문자의 정보량에 관한 연구[6]나 트리분류기를 사용한 인식에 관한 연구[7]가 최근에 발표된 바 있다.

### 2.3 신경망을 이용한 한글문자 인식

본 절에서는 신경망으로 단일활자체 한글 문자를 인식하기 위하여 제안되었던 연구에 대하여 알아 보겠다. 스캐너를 통하여 입력받은 실제적인 문자 영상을 신경망에 입력하여 한글 문자를 인식한 연구로는 경쟁하는 신경망 조직을 이용하여 597자의 한글 문자를 학습시키고 이를 인식한 고병기의 연구[1]가 부분 인식률 96.0%, 전체 인식률 83.6%의 결과를 얻었고,  $20 \times 20$ 의 크기로 정규화(normalize)된 입력 영상을 신경망의 입력으로 사용하여 한글의 유형을 15가지로 분류한 후 각 유형별 문자인식 신경망에서 한글을 인식하도록 구성한 김병기의 연구[2]가 510자의 한글을 대상으로 실험한 결과 97.61%의 유형 분류율과 96.31%의 문자 인식률을 나타냈다. 또한 한글을 6가지 유형으로 분류한 후 각 유형별 문자인식 신경망에서 자소별로 인식하도록 신경망을 계층적으로 구성하고, 변형을 첨가한 학습과 선택적 재학습 등으로 Backpropagation 알고리즘을 개선하여 신경망을 학습시킨 조성배의 연구[11]가 한글 짚기 우선 순위별 1500자[12] 중 상위 990자에 대하여 98.28%의 인식률을 얻은 바 있다.

### III. 문자추출 신경망

본 장에서는 기존의 경험적 지식에 의한 문자 추출 방법 대신 신경망의 학습에 의한 문자 추출 방법에 대하여 설명하겠다. 문서로부터 문자를 추출하는 일반적인 방법은 수평 투영에 의하여 하나의 문자열을 검출하고, 검출된 문자열을 수직 방향으로 투영하여 다수의 련으로 나눈 후 분리된 련에 여러가지 방법을 적용하여 각각의 문자를 추출하는 것이다. 영문자 추출시에는 접촉된 문자의 분리만으로 문자를 추출할 수 있으나, 한글 문서의 문자추출 과정에서는 분리된 두 개의 련이 결합되어 문자를 형성하는 경우가 많으므로 결합 과정과 분리 과정이 병행되어야 한다.

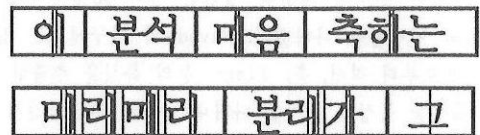
현재까지 신경망을 사용하여 한글 문서로부터 문자를 추출하는 연구가 수행된 바 없다. 즉, 하나의 신경망을 사용하여 문자 추출의 모든 과정을 처리하고자 하였기 때문에 신경망을 문자 추출에 적용할 수 없었다. 본 논문에서는 수평 투영에 의하여 검출된 문자열을 (그림 3.1)과 같이 WR(White Run)과 BR(Black Run)으로 나눈

후 각 런에 대하여 좌측에서 우측의 순서로 진행하며 현재의 런이 BR일 경우는 문자분리 신경망에서 연결된 문자를 분리시킬 것인가를 결정하고, WR일 경우는 문자결합 신경망에서 분리된 문자를 결합시킬 것인가를 결정하여 그 결과에 따라 분리와 결합 과정을 수행하는 문자 추출 방법을 제안한다(그림 3.2).

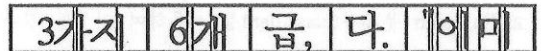
한글 문서의 경우 수직 투영에 의하여 분리된 런은 일반적으로 (그림 3.3)과 같은 형태로 나타난다. 예를 들면, 'ㅎ', 'ㅏ', 'ㅓ', 'ㅣ'로 분리된 런으로부터 두 문자 '하', '지'를 추출하여야 한다. 따라서 이러한 형태를 모두 처리하기 위하여 문자결합 신경망은 좌우의 BR이 결합되어야 문자가 형성되는 경우에는 좌우의 런을 결합시키라는 결정을 하도록 학습된다. 또한 문자 분리 신경망은 원래의 런이나 문자 결합 신경망에 의해 생성된 런에 대하여 BR의 크기가 문자 크기의 짝수 배일 경우 중심에서 분리하고, 문자 크기의 홀수 배이거나 특수 기호

가 문자의 끝 부분과 접촉되었을 경우는 앞에서 분리하도록 학습되며, 숫자 등이 문자의 앞 부분과 접촉된 경우는 뒤에서 분리하도록 학습된다. 따라서 좌우에 분리되어야 할 런이 위치하고 이 두 런이 결합되어야 할 경우나 기타 다른 여러 복잡한 경우에도 결합 과정과 분리 과정이 반복되어 결국은 각각의 문자로 추출된다. 신경망이 현재의 런을 분리할 것인가 또는 좌우의 런을 결합할 것인가를 결정하면 이 결정에 따라 분리와 결합을 수행하는 과정은 단순하다. 다만 한글의 경우 두 문자가 접촉되어 있으면 분리해야 할 위치의 결정 방법이 모호하므로 분리 과정에서 문자 크기에 대한 정보를 이용해야 한다.

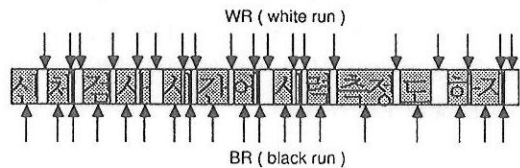
< 한글의 경우 >



< 숫자나 기호와 혼합된 경우 >

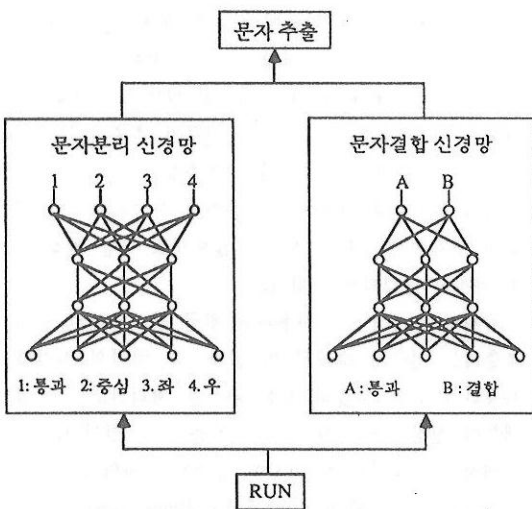


(그림 3.3) 수직 투영에 의해 얻어진 한글 문자 영상의 유형

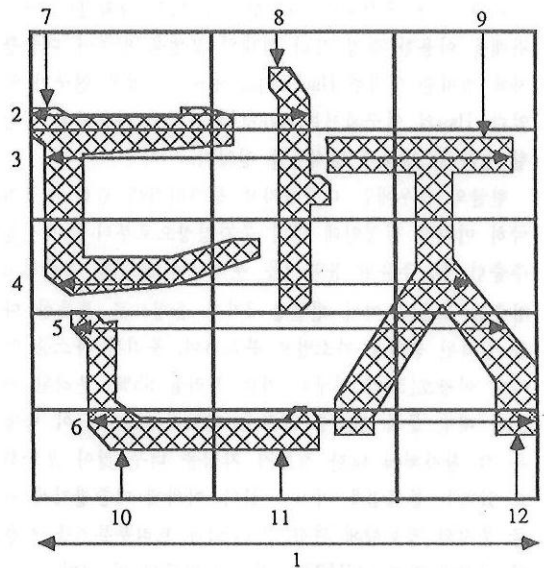


추출되는 문자: 신 체 검 사 시 간 에 시 력 측 정 도 하 지

(그림 3.1) 수직 투영에 의해 WR과 BR로 나뉘어진 문자열



(그림 3.2) 문자추출 신경망의 구성



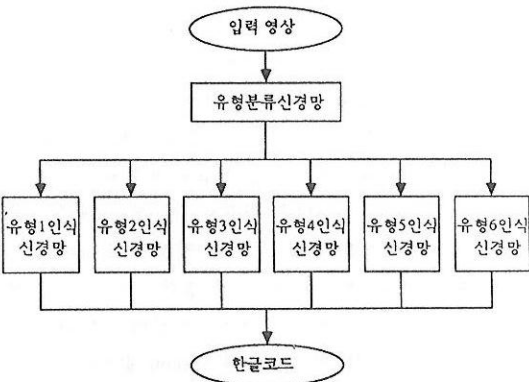
(그림 3.4) 문자 추출 신경망의 입력

문서로부터 문자를 추출하기 위하여 앞에서 설명한 것과 같이 구성된 신경망의 입력으로는 현재의 련과 좌우 두개의 BR, 즉 5개의 련에 대해 (그림 3.4)에 나타난 값을 사용한다. 이 입력은 박진규[4]의 논문에서 사용한 특성을 기초로 하여 (그림 3.3)의 유형을 구분하는데 중요한 역할을 하는 문자의 가로 및 세로 길이 정보(1-12)와, 특수 기호 및 수직 모음을 구분할 수 있는 정보(7-12)를 포함하도록 실험적으로 결정하였다.

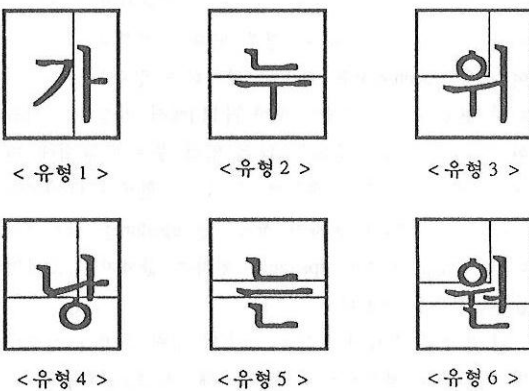
#### IV. 다중활자체 한글 인식을 위한 신경망

##### 4.1 신경망의 구조

본 논문에서는 조성배[11]가 제안한 계층적 신경망을 기본으로 하여 신경망을 구성하였다(그림 4.1). 이 방법은 신경망을 두 단계로 나누어, 먼저 유형 분류 신경망



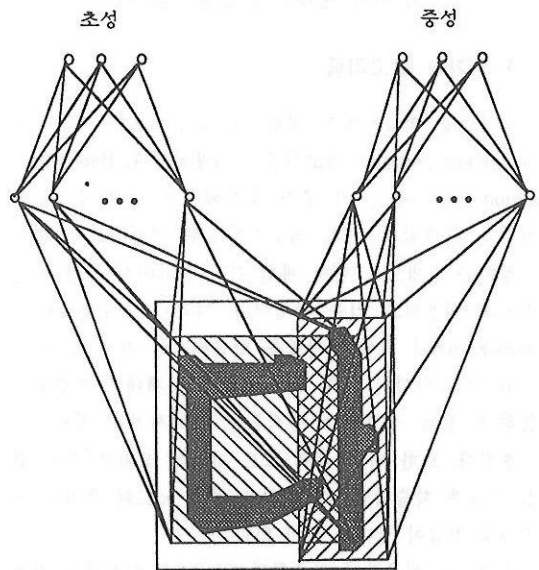
(그림 4.1) 전체 시스템의 구성



(그림 4.2) 한글의 형태와 유형

이 입력된 문자를 (그림 4.2)와 같은 6가지 유형 중의 하나로 분류하고 이에 따라 유형별 문자인식 신경망이 문자를 자소별로 인식하는 것이다. 유형 분류 신경망은 한글의 구조적 특성인 모음의 형태와 받침의 존재 유무에 따라 한글의 유형을 6가지로 분류한다. 이와 같이 유형을 분류하면 같은 출력 노드에는 같은 유형의 입력 문자만 대응되므로 하나의 출력 노드에 대응하는 입력의 형태를 제한할 수 있고 따라서 신경망의 학습 속도 및 인식률을 높일 수 있다. 또한 같은 유형의 문자에서는 각 자소가 유사한 위치에 놓이게 되므로 한글을 자소별로 인식하기 위해 필요한 부분만을 신경망에 입력시킬 수 있고, 따라서 불필요한 입력이 제거되므로 연결 강도의 수를 줄일 수 있다.

유형 분류 신경망에서 분류된 문자는 그에 해당하는 유형별 문자인식 신경망에서 자소별로 인식된다. 이 신경망에서는 각 유형별로 자소가 공통적으로 위치하는 영역을 결정하여 그 영역만을 신경망의 중간층과 연결시키게 된다. 자소의 영역을 결정하기 위하여 본 논문에서는 각 활자체의 모든 경우를 포함할 수 있도록 충분히 큰 영역을 설정하여 중간층 노드와 연결시켰다(그림 4.3). 물론 이러한 방법으로 자소의 영역을 결정하면 원하지 않는 자소의 일부가 입력에 포함될 수 있지만 신경망은 단순히 결합 극복 능력을 가진 원형 정합 방법

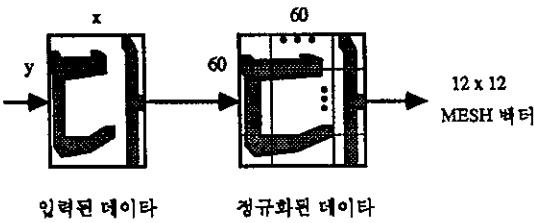


(그림 4.3) 유형 1 문자인식 신경망의 구성

이 아니라 자체의 일반화 능력을 가지고 있으므로 이러한 변형의 문제를 해결할 수 있다.

4.2 신경망의 입력

신경망을 사용하여 문자를 인식하는 대부분의 연구에서 신경망의 입력으로 이진 영상을 그대로 이용하거나 전처리 과정을 통해 추출된 특징을 사용한다. 그러나 전자의 경우는 입력의 크기가 너무 커져서 학습 시간이 많이 소요되며 약간의 위치 변화가 인식 과정에서 큰 영향을 미칠 수 있다는 단점을 가지고 있고, 후자의 경우에는 신경망 전체의 성능이 특징 추출 결과에 크게 의존하므로 신경망의 장점을 최대한으로 활용할 수 없다는 단점을 갖는다. 이와 같은 두가지 면을 고려하여 본 논문에서는 입력의 수를 줄이면서 약간의 위치 변화를 흡수할 수 있고 또한 신경망의 성능이 입력의 표현에 의해 결정되지 않도록 하기 위해 메쉬(mesh) 벡터를 입력으로 사용하였다(그림 4.4)



(그림 4.4) 문자인식 신경망의 입력

4.3 학습 알고리즘

신경망을 학습시키기 위한 알고리즘으로는 기본적으로 Backpropagation 알고리즘을 사용하였다. Backpropagation 알고리즘은 각 출력 노드에서의 실제 출력값과 원하는 출력값의 차가 최소화되도록 연결강도의 값을 조정하는 방법으로 현재 패턴 인식 분야에서 가장 일반적으로 사용되고 있는 방법이다. 그러나 순수한 Backpropagation에 의해 신경망을 학습시키면 전체 입력 문자와 학습 방향이 다른 소수의 입력이 제대로 학습되지 못하여 전체 신경망의 인식률을 저하시키는 원인으로 작용한다. 또한 비록 출력 노드 전체의 오차는 적은 값을 갖도록 학습되었다고 일부 출력 노드의 오차는 큰 상태로 학습이 끝나는 수가 있다.

신경망을 이용하여 다중활자체 한글을 인식하는 경우와 같이 하나의 출력 노드에 대응될 입력이 상호 큰 차

이를 보이는 경우에는 유사한 다른 자소를 해당 자소의 변형으로부터 확실히 구분할 수 있어야 한다 또한 이러한 많은 변형 중의 일부가 학습되지 않은 입력으로 존재할 가능성이 높으므로 이를 처리할 수 있는 방법이 고안되어야 한다. 이러한 문제의 해결을 위하여 본 논문에서는 Descending Epsilon 방법[17]을 사용하였다. 이 방법의 핵심은 활성화되는 출력 노드의 값을 다른 노드의 값보다 확실히 크게 하기 위하여 학습되지 않은 출력 노드의 오차만으로 신경망을 학습시키는 것인데 신경망이 학습된 정도를 나타내는 척도로

$$E = \sum_{p=1}^P \sum_{i=1}^N (t_{pi} - o_{pi})^2$$

의 식으로 계산되는 오차를 사용하는 대신

$$E = \sum_{p=1}^P \sum_{i=1}^N (\text{number of } \{t_{pi} - o_{pi} > \epsilon\})$$

을 사용한다. 즉 실제 출력값을 이전 출력값으로 바꾸었을 때, 원하는 출력값과 다른 값을 갖는 출력 노드의 수를 사용한다.

Epsilon Descending 방법의 전체 알고리즘은 다음과 같다. 먼저 어떤 epsilon 값을 설정하여 각 출력 노드에서의 오차가 이 값보다 작을 경우 오차를 0으로 만든다. 따라서 이 출력 노드의 오차는 학습에 반영되지 않는다. 이 방법에 의하여 신경망을 학습시키면 모든 활성화 출력 노드와 비활성화 출력 노드가 일정 범위외의 값을 가지므로 비활성화되어야 할 어떤 출력 노드가 높은 출력값을 내는 것을 방지한다. 또한 각 입력에 대하여 모든 출력 노드에서의 오차가 현재 epsilon 값보다 작으면 그 입력은 학습에 참여시키지 않는다. 따라서 오차가 큰 입력에 대한 집중적인 학습이 이루어지므로 학습되지 않은 소수의 입력이 전체 인식률을 저하시키는 문제를 해결할 수 있다. 특히 이 경우 변화된 연결강도로 인하여 다음 epoch에서는 학습에 참여하는 입력 문자의 집합이 변화하므로 선택적 재학습[11]에서 발생하는 학습된 입력 문자와 학습되지 않은 입력 문자의 혼합에 따른 문제를 자동으로 해결할 수 있다. 현재의 epsilon에 대하여 모든 입력 문자가 학습되면 epsilon을 감소시켜 다시 학습을 시키며 epsilon이 원하는 값까지 감소되었을 때 학습을 마친다.

이 방법은 학습되지 않은 소수의 입력 문자가 신경망의 인식률을 저하시키는 문제를 해결하였고, 학습된 입력 문자는 현재 epoch의 학습에 참여시키지 않음으로써

학습 시간을 줄일 수 있다. 또한 활성화되는 출력 노드와 비활성화되는 출력 노드 간의 출력값의 차가 크도록 학습되므로 인식 과정에서 다소의 변형을 흡수할 수 있고 비교적 큰 오차에서도 높은 인식률을 얻을 수 있다. <표 4.1>은 Backpropagation과 Descending Epsilon 방법을 사용하여 학습된 신경망으로 다중활자체 한글 문자를 인식한 결과이다. 그러나 이 방법은 다른 답을 요구하는, 같은 형태의 입력이 학습 자료에 포함되어 있을 경우 신경망을 학습시킬 수 없다는 단점을 가지고 있다.

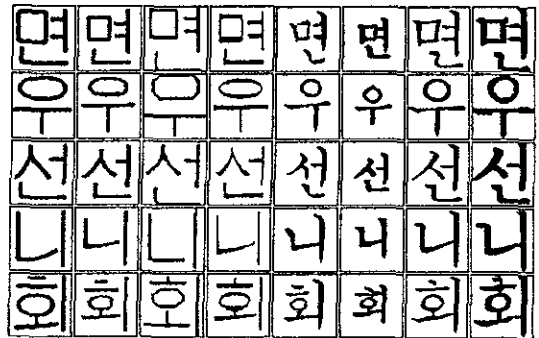
<표 4.1> BP와 DE에 의한 다중활자체 한글 문자 인식

	BP	DE
디너루체	93.27	97.78
고딕체(중)	94.23	97.98
고딕체(대)	94.62	97.74
공작체	93.65	94.61
궁서체	86.73	94.41
명조체(소)	91.73	95.76
명조체(중)	91.35	98.36
명조체(대)	92.31	97.64

BP, Backpropagation  
오차가 0.005보다 작아질때까지 학습

DE Descending Epsilon  
epsilon이 0.1보다 작아질때까지 학습

NEWS image reader를 이용하여 이진 영상으로 입력 받았다. 실험에 사용한 문자 추출 신경망은 2개의 중간층을 가지며 Backpropagation 알고리즘으로 전체 오차가 0.01보다 작아질 때까지 학습시켰고, 문자 인식 신경망은 하나의 중간층을 가지며 Descending Epsilon 알고리즘으로 epsilon이 0.1이 될 때까지 학습시켰다. 문자 추출 신경망에서는 같은 형태의 입력에 대하여 다른 답을 요구하는 경우가 발생하므로 신경망을 학습시키기 위하여 Descending Epsilon 알고리즘을 사용하지 못하고 Backpropagation 알고리즘을 사용하였다. (그림 5.1)과 (그림 5.2)는 실제로 실험에 사용한 입력의 예로서 스캐너로부터 입력된 이진 영상을 레이지 프린터를 이용하여 출력시킨 것이다.



디너루체 고딕체(중) 고딕체(대) 공작체 궁서체 명조체(소) 명조체(중) 명조체(대)  
(그림 5.1) 각 활자체별 입력 문자영상의 예

## V. 실험 결과 및 분석

### 5.1 실험 환경

본 논문에서는 ELXSI 컴퓨터를 사용하여 신경망을 학습시키고, SONY NEWS workstation에서 C 언어를 사용하여 전체 시스템을 구현하였다. 실험은 크게 2가지로 나누어지는데 첫번째 실험에서는 2가지 크기의 고딕체와 3가지 크기의 명조체, 그리고 단일 크기의 공작체, 궁서체, 디너루체 각각에 대하여 한글 기계화 연구소에서 조사한 한글 찾기 순위별 상위 1,500자[12] 중에서 상위 520자를 대상으로 하여 인식하였다. 이 실험에서는 각 문자 사이에 일정한 공간이 있는 문서를 사용하였으므로 추출된 문자에 대한 인식만을 수행하였다.

두번째 실험은 실제적인 문서를 가지고 수행하였는데, 크기 또는 형태가 다른 6가지의 활자체로 구성된 일반 대중 잡지를 선택하여 인식하였다. 실험에 사용한 입력 문서는 400 DPI의 해상도를 갖는 광학 문서입력 장치인

**근시는 악물이나 근육체로  
치유될수 없다**

각막절개술은 각막의 굴절 정도를 낮추어주는 수술인데 성공율은 1백 퍼센트이지만 합병증이 생길 가능성이 있다는 점이 문제로 남아 있다.  
**김용규(안과전문의)**

**근시원인은 유전**

(그림 5.2) 실험에 사용한 실제 문서의 예

5.2 활자체별 실험

5.2.1 각 자소별 인식률

앞 절에서 언급한 모든 활자체를 입력으로 하여 신경망을 학습시킨 후 각 자소별로 인식한 결과 (표 5.1)과 같은 인식률을 얻었다. 이 결과를 분석해 보면 초성과 중성의 인식률은 99% 이상인데 반하여 중성, 즉 모음의 인식률이 상대적으로 낮은 것이 문자 오인식의 가장 큰 요인이라는 것을 알 수 있다. 실제로 잘못 인식된 문자의 50.8%가 모음의 오인식으로 인한 것이었다. 특히 모음의 오인식 중 89.1%가 유형 2 또는 유형 5의 모음, 즉 ‘-’, ‘ㄱ’, ‘ㄷ’, ‘ㅇ’, ‘ㅍ’에서 발생했고 그 중 95% 이상이 ‘-’, ‘ㄱ’, ‘ㄷ’의 혼동에서 발생한 것이다. 이러한 결과는 신경망의 학습과 인식 과정에서 자소를 분리하지 않고 문자의 일정 영역을 입력으로 하여 처리하므로, ‘그’와 ‘오’에서와 같이 ‘-’와 ‘ㄱ’가 문자내에서 유사한 형태를 가지고 있기 때문에 발생하는 것이다. 이러한 문제가 본 논문과 유사한 방식으로 한글을 인식할 경우 가장 시급히 해결해야 할 문제라고 생각하며 참고로 이 경우의 오인식을 제외하면 전체의 인식률이 3% 정도 높아진다.

이 밖에 입력으로 12×12의 배쉬 벡터를 선택하였기 때문에 발생한 문제로 중성이나, 유형 5 또는 유형 6의 초성에서 ‘ㄷ’, ‘ㅁ’, ‘ㅂ’, ‘ㅋ’, ‘ㅌ’, ‘ㅍ’, ‘ㅎ’의 구분이 어려워진다는 문제점을 가지고 있다. 실제로 입력을 이진화하여 사람이 판단하더라도 구분이 어려울 정도이므로 이러한 문제점과 입력의 크기에 따른 학습 및 인식 속도, 그리고 위치 이동에 따른 변화 흡수 등을 고려하여 적절한 입력의 크기를 결정하는 것이 중요하다고 하겠다. 이러한 오인식이 중성 오인식의 90% 정도를 차지한

〈표 5.1〉 자소별 인식률

	유형	초성	중성	중성	전체
디나루체	99.14	99.23	98.27	99.02	97.78
고딕체(중)	99.71	99.62	98.36	99.58	97.98
고딕체(대)	99.76	99.37	98.80	98.87	97.74
공작체	99.52	99.23	96.15	98.16	94.61
궁서체	99.42	98.65	96.82	97.04	94.41
명조체(소)	99.33	98.27	97.50	98.04	95.76
명조체(중)	99.81	99.27	98.94	99.16	98.36
명조체(대)	99.66	99.42	98.27	99.09	97.64

다. 각 유형별로 자주 오인식되는 문자는 (표 5.2)와 같다.

5.2.2 각 유형별 인식률

앞 절의 결과를 유형별로 분석해 보면 (표 5.3)과 같다. 실험에 사용한 520자를 유형별로 나누면 유형 1이 93자, 유형 2가 50자, 유형 3이 21자, 유형 4가 218자, 유형 5가 123자, 그리고 유형 6이 15자이다. 따라서 인식률에 큰 영향을 미치는 것은 유형 1과 유형 4, 그리고 유형 5인데 유형 1과 유형 4는 활자체별로 고른 인식률을 보이므로 결국 유형 5의 인식률이 전체 인식률을 결정한다고 볼 수 있다. 실제로 인식률이 낮은 공작체와 궁서체의 경우 유형 5의 인식률이 83.61%와 90.61%에 머물러 있어 다른 유형과 많은 차이를 나타냄을 알 수

〈표 5.2〉 많은 오인식을 발생하는 자소

	유형1	유형2	유형3	유형4	유형5	유형6
초성	ㄱ 2	ㄷ 2	ㅎ 1	ㅁ 4 ㅌ 3 ㅍ 7	ㄷ 4 ㅌ 2 ㅎ 2	
중성	ㅣ 2 ㅣ 2 ㅌ 2	ㅡ 3	ㅣ 3	ㅣ 4 ㅣ 9 ㅣ 3 ㅣ 4 ㅣ 6	ㅡ 57 ㅡ 21 ㅡ 21	
중성	X			ㅁ 25 ㅌ 6 ㅌ 4 ㅌ 4	ㅍ 2 ㅌ 2 ㅌ 3	

〈표 5.3〉 유형별 인식률

	유형1	유형2	유형3	유형4	유형5	유형6	전체
디나루체	98.91	100	90.70	98.40	96.73	93.75	97.78
고딕체(중)	98.93	99.00	100	97.95	96.31	100	97.98
고딕체(대)	99.46	99.00	100	97.59	95.50	100	97.74
공작체	99.46	97.03	90.91	98.15	83.61	100	94.61
궁서체	95.63	96.04	86.36	96.09	90.61	100	94.41
명조체(소)	97.80	97.96	95.35	95.20	94.33	96.77	95.76
명조체(중)	99.47	100	100	97.93	97.14	100	98.36
명조체(대)	98.92	99.50	97.59	97.49	95.91	100	97.64



있다. 이것은 앞 절에서 유형 5의 모음 ‘ㅡ’, ‘ㅓ’, ‘ㅜ’의 혼동이 전체 인식률에 큰 영향을 미친다는 것을 더욱 잘 나타내주고 있다.

### 5.2.3 활자체별 인식률

앞 절의 결과에서 볼 수 있듯이 공작체와 궁서체의 인식률이 다른 활자체에 비해 매우 낮다. 그 원인은 고딕체와 명조체에 대해서는 여러 크기의 입력 문자를 학습시켰으므로 신경망이 이 두 활자체의 입력 문자를 주로 학습하였기 때문이다. 즉 학습 기회가 적었던 디나루체, 공작체, 궁서체 중 문자의 형태와 자소의 위치가 고딕체나 명조체와 상이한 공작체와 궁서체의 인식률이 상대적으로 낮아진 것이다(그림 5.1) 따라서 보다 효율적으로 다중활자체 한글을 인식하는 신경망을 학습시키기 위해서는 각 활자체에 대하여 입력의 빈도수를 고르게 유지하는 것이 필요하다.

### 5.3 실제 문서에의 적용 결과

본 논문에서 구현한 다중활자체 한글문서 인식 시스템이 실제 문서를 어느 정도 인식할 수 있는가를 알아보기 위하여 ‘마리안느’라는 여성 잡지로 실험하여 보았다. 이 잡지를 인식 대상으로 선택한 이유는 한 장의 문서에 비교적 다양한 활자체가 섞여 있으므로 본 시스템의 학습 및 인식 성능을 평가하기에 적절하다고 생각하였기 때문이다.

실험은 크기나 형태가 다른 6가지의 활자체를 대상으로 행하여졌으며, 7장의 문서에 포함된 2,631자를 학습시켰으나 같은 문자가 여러 번 중복되어 나타나므로 실제로 학습된 문자의 종류는 403자이었다 이와 같이 학습된 시스템으로 학습 과정에서 사용하지 않은 4장의 문서에 포함된 964자를 인식한 결과 문자 추출률 97.5%, 추출된 문자에 대한 인식률 96.7%를 나타내어 총 94.5%의 인식률을 얻을 수 있었다. 이 결과는 좀더 많은 학습을 시킬 경우 향상될 수 있다.

이 실험에서 문자 오인식의 주요 원인은 문자 추출의 실패이다. 특히 볼드체의 경우 주로 3개 이상의 문자가 접착되어 나타나므로 오인식의 80% 이상이 문자 추출 단계에서 발생하였다. 문자가 제대로 추출되었을 경우에는 앞 절에서 설명한 바와 같이 수평 모음의 오인식이 대부분을 차지한다. 본 시스템의 속도는 스케너로 문서를 입력받는 시간과 문자 추출 및 인식 시간, 그리고 X

window 상에 출력시키는 시간을 포함하여 SONY workstation에서 637자가 포함된 문서를 468초에 인식하였으므로 한 자를 인식하는데 평균 0.73초 정도를 소요한다. 이것은 기존의 방법을 사용한 문서인식 시스템에 비교하면 다소 느린 속도이지만, 신경망 부분을 VLSI로 구현한다면 실시간(real time)에 가까운 인식 속도를 얻을 수 있을 것이다.

## VI. 결 론

본 논문에서는 단일활자체 한글 문자를 효과적으로 인식하기 위하여 제안된 계층적 신경망을 기본으로 하여 실제 문서에서 나타나는 여러 크기의 다양한 활자체 한글문자를 인식하였다. 또한 신경망을 이용하여 실제 문서로부터 문자를 추출하는 방법을 제안하고, 이를 문자인식 부분과 결합하여 다중 크기의 다중활자체 한글 문서 인식 시스템을 구현하였다.

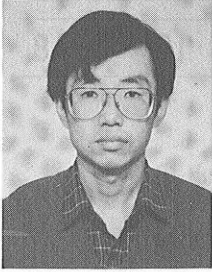
문자 추출 과정에서는 입력에 여러 제약을 두었던 경험적 지식에 의한 문자 추출 방법 대신 신경망을 이용함으로써 다양한 활자체의 한글 문자가 숫자 및 특수 기호와 섞여 있는 실제 대중 잡지로부터 한글 문자를 성공적으로 추출할 수 있었다. 또한 추출된 문자를 인식하는 단계에서는 신경망의 학습 능력, 일반화 능력, 그리고 결합 극복 능력을 이용하여 다중활자체 한글의 많은 변형을 흡수할 수 있었으며, 특히 모든 입력영상의 모든 출력 노드를 고르게 학습시키는 Descending Epsilon 알고리즘을 신경망의 학습에 사용함으로써 학습되지 않은 소수의 입력에 의한 학습 속도 및 인식률의 저하라는 문제를 해결하였다. 그 결과 각 활자체에 대해 94% 이상의 인식률을 나타냈으며, 이를 실제 문서에 적용하여 94.5%의 인식률을 얻을 수 있었다. 한글사전 및 문법적 지식을 이용하여 오인식된 문자를 교정하는 후처리부를 설계하여, 본 논문의 문서 인식 시스템과 결합한다면 기존의 연구보다는 한 단계 더 실용적인 한글문서 인식 시스템에 접근할 수 있을 것이다.

앞으로 더 연구해야 할 과제로는 다음의 3가지를 생각할 수 있다. 첫째는 문자추출 신경망의 성능을 향상시킬 수 있는 입력에 관한 연구이다. 본 논문에서는 신경망의 입력 방식에 대한 충분한 실험을 하지 못하였으므로 입력을 개선할 경우 더 높은 문자 추출률을 얻을 수 있을 것이다. 다음으로 ‘ㅡ’, ‘ㅓ’, ‘ㅜ’의 혼동을 감소시킬

수 있는 방법의 개발이 필요하다. 위의 세가지 수평 모음에 대한 혼동이 오인식의 34% 이상을 차지하므로 이에 관한 연구가 필수적이다. 셋째로 신경망의 일반화 능력을 높일 수 있는 학습 방법이 개발되어야 하겠다. 보다 실용적인 문서 인식 시스템의 개발을 위해서는 현재의 학습되지 않은 활자체의 입력 문자에 대한 인식률 80%에서 85%를 90% 이상으로 높일 수 있는 신경망의 일반화 방법에 대한 연구가 필요하다. 마지막으로, 한 단계 더 발전한 문서인식 시스템 즉 한글뿐만 아니라 영어와 한문 등이 섞여 있는 문서를 인식할 수 있는 시스템 개발에 관한 연구가 진행되어야 한다고 생각한다

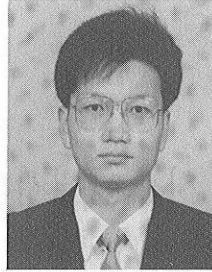
### 참 고 문 헌

1. 고병기, 김진형, 양현승, "경쟁하는 신경망 조직을 이용한 인체체 한글문자의 인식", 제1회 영상처리 및 이해에 관한 워크숍 발표논문집, 포항, 1989년 1월.
2. 김병기, 김항준, "신경망 모델을 이용한 한글 문자의 형태 분류와 인식", 한국정보과학회 추계 학술 발표 논문집, 서울, pp. 303-306, 1989년 10월.
3. 김형훈, 이성환, 김진형, "신문의 구조적 분석을 통한 한국 신문 기사의 추출", 한국정보과학회 학술 논문지, 제15권, 제5호, pp. 392-404, 1988.
4. 박진규, 최기석, 안영경, "다중 문자로 이루어진 문서의 문자 추출 및 유형 분류에 관한 연구", 한국정보과학회 추계 학술발표 논문집, 서울, pp. 319-322, 1989년 10월
5. 이광호, "다중 활자체 한글 인식을 위한 자모의 분리", 한국과학기술원 전산학과 석사학위논문, 1989.
6. 이성환, "한글 문자 영상에서의 정보량 및 엔트로피의 분포", 제2회 한글 및 한국어정보처리 학술발표 논문집, 서울, pp. 208-215, 1990년 11월.
7. 이성환, 김홍섭, "다양한 활자체 및 크기의 한글 문자 인식을 위한 최적화된 계층적 트리 분류기", 제3회 영상처리 및 이해에 관한 워크숍 발표논문집, 수안보, pp. 151-161, 1991년 2월.
8. 이승형, 전종익, 조용주, 남궁재찬, "신문 자동인식 시스템을 위한 문자의 분류에 관한 연구", 제1회 한글 및 한국어정보처리 학술발표 논문집, 서울, pp. 209-215, 1989년 10월.
9. 이승호, 조창제, 김일영, 조성배, 양현승, 김진형, "한글 문서 인식시스템 SILNOON의 개발", 한국정보과학회 춘계 학술발표 논문집, 부산, pp. 211-214, 1989년 4월.
10. 이현표, 양순성, 황교철, 이균하, "한글 문서에서의 날자 분리 알고리즘", 제1회 한글 및 한국어정보처리 학술발표 논문집, 서울, pp. 203-208, 1989년 10월.
11. 조성배, "신경망 기법의 현실적 적용을 위한 개선 전략: 인체체 한글문자 인식을 중심으로", 석사학위논문, 한국과학기술원 전산학과, 1990.
12. 한글 기계화 연구소. 한글 기계화 연구, 1975.
13. K. S. Fu and A. Rosenfeld, "Pattern Recognition and Computer Vision," *IEEE Computer*, pp. 274-282, October 1984.
14. P. N. Chen, Y. S. Chen and W. H. Hsu, "Stroke Relation Coding - A New Approach to the Recognition of Multi-font Printed Chinese Characters," *Int Jour. of PRAI*, Vol. 1, No. 1, pp. 149-160, 1988.
15. S. Kahan, T. Pavlidis and H. S. Baird, "On the Recognition of Printed Characters of Any Font and Any Size," *IEEE Trans. on PAMI*, Vol. PAMI-9, No. 2, pp. 274-288, 1987.
16. S. Shilen, "Multi Font Character Recognition for Typeset Documents," *Int. Jour. of PRAI*, Vol. 2, No. 4, pp. 603-620, 1988.
17. Y. H. Yu, R. F. Simmons, "Descending Epsilon in Back-Propagation: A Technique for Better Generalization," in *Proceedings of the IEEE IJCNN*, Vol. III, pp. 167-172, 1990.



권재욱

1989년 연세대학교 이과대학  
전산과학과 졸업  
1991년 한국과학기술원 전산  
학과 석사학위 취득  
1991년~현재 한국과학기술원  
전산학과 박사과정 재학중  
관심분야: 문서의 구조 분석,  
문자인식, 신경망, 지능형 Man-Machine Interface 등



조성배

1988년 연세대학교 이과대학  
전산과학과 졸업  
1990년 한국과학기술원 전산  
학과 석사학위 취득  
1991년~현재 한국과학기술원  
전산학과 박사과정 재학중  
관심분야: 신경망, 패턴인식,  
지능형 Man-Machine Interface 등



김진형

1971년 서울대학교 공과대학  
졸업  
1979년 UCLA 전산학 석사학  
위 취득  
1983년 UCLA 전산학 박사학  
위 취득  
1973년~1976년 KIST 전산실

연구원

1976년~1977년 미국 Californiaia 도로국 연구원  
1981년~1985년 미국 Hughes 인공지능센터 선임연구원  
1986년~1988년 한국정보과학회 산하 인공지능연구회  
위원장  
1985년~현재 한국과학기술원 부교수  
1991년~현재 한국과학기술원 인공지능연구센터 부소  
장, 국제로봇자동차시스템 논문지의 편집자  
관심분야: 컴퓨터 시각, 전문가 시스템, 지능형 Man-  
Machine Interface 등