

碩士學位論文

神經回路網을 利用한 英文字 認識
알고리즘에 關한 研究

濟州大學校 大學院

通信工學科



金 德 周

1994年 12月

神經回路網을 利用한 英文字 認識
알고리즘에 關한 研究

指導教授 林 載 允

金 德 周

이 論文을 工學 碩士學位 論文으로 提出함

1994年 12月

金德周의 工學 碩士學位 論文을 認准함

審査委員長

李 鎔 鶴

委 員

康 鎭 植

委 員

林 載 允

濟州大學校 大學院

1994年 12月

**A Study on Algorithm for English Character
Recognition by Neural Networks**

Deog-Ju Kim

(Supervised by professor Jae-Yun Lim)

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF MASTER
OF SCIENCE IN ENGINEERING**

**DEPARTMENT OF TELECOMMUNICATION ENGINEERING
GRADUATE SCHOOL
CHEJU NATIONAL UNIVERSITY**

1994. 12

목 차

Abstract	1
I. 서 론	2
II. 문자 인식 시스템	4
1. 패턴 인식과 문자 인식	4
2. 종래의 문자 인식 방법 및 문제점	6
III. 문자 인식 시스템으로서의 신경회로망	10
1. 신경회로망	10
2. 문자 인식 시스템으로서의 신경회로망	14
IV. 영문자 인식용 신경회로망 모델 구현	15
V. 알고리즘 흐름도 및 시뮬레이션 결과	21
1. 영문자 인식 알고리즘	21
2. 시뮬레이션 순서도	22
VI. 결 론	31
참 고 문 헌	32

Abstract

An efficient algorithm for English character recognition by neural networks is proposed. After normalizing input patterns, input neurons detect local distribution of pixels and determine input vectors. Neural networks were constructed in three layers and trained by Error Back Propagation method. During the process of training, neural network's weight were determined. English characters were recognized by neural networks after training. These algorithms are programmed by C-language on IBM PC 486-DX. The learning is repeated until the difference between object output and real output is less than 0.00001. After setting the hidden neuron's number optimally, learning velocity and recognition rate are analyzed as momentum α and learning rate η vary from 0.1 to 0.5. Through the trained neural networks, it is shown that recognition ratio is 96.2% for learning patterns and 81.5% for no learning patterns.



I. 서 론

현대사회의 방대한 정보량의 증가는 활자를 매체로 하는 수용 한계를 넘었고, 이미 컴퓨터를 활용하지 않으면 안되게 됐다. 특히, 기존에 문서화되어 있는 많은 데이터와 정보를 데이터베이스화하여 원하는 정보를 신속히 찾아내기 위해서는 수작업으로 입력하는 것보다 신속하고 정확하게 데이터를 입력할 수 있는 장치의 개발이 요구된다. 이러한 데이터 입력장치의 개발이 선행되지 않는다면 컴퓨터의 정보처리 속도와 연산 속도가 아무리 향상되어도 효율적인 정보처리를 할 수 없으므로 패턴 인식 분야의 개발이 요구되고 있다(Hussain 과 Kabuka, 1994), (오, 1993), (이, 1993).

대부분의 병렬 처리 문제들은 순차적인 해결 알고리즘으로는 모든 경우의 수에 대한 처리 시간이 시스템의 처리 시간 중 대부분을 차지하게 되어 전체 처리 시간이 길어진다. 또한 순차형 컴퓨터를 이용한 문자 인식 방법은 처리 데이터를 순차적으로 처리하므로 처리 속도에 한계가 있다. 이에 비해 분산된 뉴런들의 처리를 이용한 신경회로망에서의 병렬 처리에 의한 문제 해결 방법은, 각 뉴런의 처리 과정이 병렬적으로 이루어져서 처리 시간이 효율적이므로 많은 관심을 모으고 있다(Koso, 1993). 특히 신경회로망은 순차형 컴퓨터의 계산 방식으로는 해결하기 어려운 여러가지 문제 특히, 패턴 인식 문제에 성공적으로 응용되고 있으며, 필기체 영문자와 숫자(남, 1990), (Fukus, 1988) 및 필기체 한글 인식(이, 1993)에도 가능성을 제시하고 있다.

본 논문에서는 신경회로망을 이용하여 영문자 패턴을 분류 인식하는 시스템을 제안하여 패턴 인식 문제에 신경회로망의 적용이 유용함을 보인다. 본 시스템에서는 마우스를 통하여 20x20개의 화소로 구성된 입력 패턴에 대해 화소의 국소

분포를 이용해 패턴의 특징 성분을 추출하며, 추출된 특징 성분은 신경회로망의 입력 벡터로 주어 영문자를 학습시킨다. 영문자 인식 시스템은 3층 전방향 전파 신경회로망으로 구성되며 학습에는 오류 역방향 전파(EBP : Error Back Propagation) 방법 (Rumelhart,1986)을 사용한다. 영문자 인식 시스템의 특성과 학습 속도와 인식 능력을 분석하기 위해 중간층 뉴런의 수와 학습률 그리고 변화율을 변화시켜 가면서 시뮬레이션 한다.

본 논문은 2 장에서 문자 인식 시스템을 패턴 인식 부분과 문자 인식 부분으로 나누어 서술하고, 3 장에서는 문자 인식 시스템으로서의 신경회로망에 대하여 서술한다. 그리고 4 장에서 영문자 인식용 신경회로망 모델 구현에 관하여 논의하며, 5 장에서 알고리즘 흐름도 및 시뮬레이션 결과를 서술하며 6 장에서 결론을 서술한다.



II. 문자 인식 시스템

1. 패턴 인식과 문자 인식

1) 패턴 인식

패턴이란 물리적으로 표현되어 사람이 인지할 수 있는 정보의 외형을 말하며, 인식이란 패턴의 분류를 통하여 궁극적이면서 구체적인 부류로 확정하는 일을 말한다. 물체를 보고 단지 대상물의 존재를 아는 과정은 인지라 한다. 따라서 패턴 인식이란 단순히 패턴의 성질이나 특징을 정량적으로 관측, 측정하는 것뿐만 아니라, 패턴과는 차원이 다른 의미나 정보의 내용을 한정하는 것으로 외형으로서의 패턴에 내용으로서의 개념을 대응시키는 작업이 패턴 인식이다(오,1993).

패턴 인식은 외부의 정보 모델을 자기 내부에 형성하는 과정이라고 할 수 있으며, 공학적 의미로서의 패턴 인식은 외부의 관측과 해석에 의해 이미 자기 내부에 형성된 모델에 따라 주어진 패턴에 분류 표시를 하여 그것이 속하는 데이터의 명칭을 출력하는 과정이라고 할 수 있다.

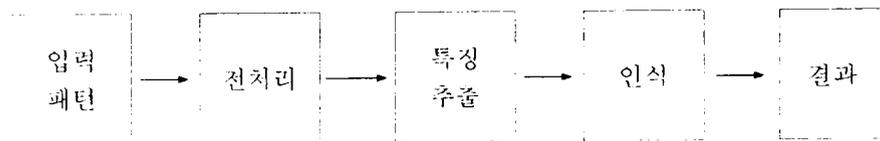


Fig.1 Construction of pattern recognition

일반적인 패턴 인식 시스템은 Fig.1과 같이 구성된다. 입력 자료는 잡음을 가지고 있으므로, 시스템에서 인식을 위해 필요한 자료만을 얻어내는 전처리를 하

게 된다. 다음은 입력된 패턴을 인식하기 용이한 작은 단위로 분할하는 과정을 거친다. 특징 추출 단계에서는 인식에 중요한 실마리가 되는 특징을 얻어내고, 이 특징을 비교하여 패턴을 식별하게 된다. Fig.1의 각 단계는 고정된 것이 아니며 순서를 일부 바꾼다든지, 일부를 생략할 수 있다.

패턴 인식의 문제점은 질적인 문제와 양적인 문제로 나눌 수 있다. 질적인 문제는 해는 알고 있으나 해법이 불분명한 경우를 말한다. 인간의 패턴 인식 기능에 의해 해의 존재는 알 수 있지만, 인간의 인식 과정은 알기 어렵다. 즉, 인간은 뛰어난 패턴 인식 기능을 가지고 있으나, 그 과정은 무의식적이면서 직관과 경험에 바탕을 두고 있다. 이러한, 인간의 패턴 인식 과정을 객관화하고 정량화시켜 알고리즘의 형태로 정식화하는 방법은 아직 완성되어 있지 않았다. 공학적 관점에서는 부분적이거나 체계적으로 접근하여, 일부 기능의 실용화를 목표로 하고 있다.

2) 문자 인식 시스템

문자는 음성과 같이 인간 상호간에 언어로 표현되는 정보를 전달하기 위해 발명된 대표적인 부호 체계이다. 정보처리 장치로서의 컴퓨터의 출현과 발달은 급속도로 이루어졌으나, 프로그램과 데이터의 입력 수단은 인간이 키보드를 통하여 입력하거나 인쇄한 문자를 천공하는 수작업에 의해서 이루어져 왔다. 따라서 컴퓨터를 이용하여 처리할 정보량이 많아질수록 사람에 의한 수작업의 기계화가 더욱 요망되어 왔다. 문자 인식은 이와 같은 요구에 부응하는 기술로서 발달하여 왔다.

기계로 문자를 인식할 경우에, 그 문자가 숫자인가 혹은 문자인가에 따라 어려움이 달라진다. 문자를 판독하는 난이도에는 일반적으로 양적인 어려움과 질적인 어려움의 두 가지 측면이 있다. 양적인 난이도는 예를 들면 인식 대상이 숫자에서 영문자 혹은 한글, 더 나아가 한자가 됨에 따라 증가하게 된다. 질적인 난이

도는, 단일 글자체의 인쇄 문자로부터 제한 필기 문자, 자유 필기 문자로 인식 대상이 바뀔에 따라 인식이 곤란하게 되는 것과 같다.

2. 종래의 문자 인식 방법 및 문제점

문자 인식을 인간의 패턴 인식 기능의 모의 실험이라는 입장에서 본다면, 문자 인식 방법도 본질적으로는 한 종류라야 한다. 그러나, 현재의 문자 인식 방법은 실용을 목적으로 하는 기술적 필요성에서 개발된 것이기 때문에 개별적이며 다양하다. 따라서, 한정된 대상에 대해서는 효과적이지만 필기체 한자나 한글 등을 포함하는 임의의 문자를 인식할 수 있는 통일적인 방법은 아직 알려져 있지 않다. 그러므로, 현재 실용화되고 있는 여러 가지 문자 인식 방법은 대부분 일장 일단이 있기 때문에 인식 대상에 따라 바꾸어 사용하는 것이 보통이다.



1) 패턴 정합법

패턴 정합법은 중첩법 또는 패턴 매칭법이라 하며 가장 오래 전부터 사용되어 온 방법이다. 이 방법은 인식 장치에 문자의 종류에 따라 각각을 대표하는 표준 패턴을 기억시키고, 인식하고자 하는 미지 문자와 각각의 표준 패턴과 일치하는 정도를 비교해서 가장 잘 일치하는 표준 패턴이 미지 문자로 판정하는 방법이다.

유사도법은 패턴 정합법을 구체화한 것으로, 문자 인식 연구 초기에 제안되어 현재까지도 널리 이용되고 있는 기본적인 문자 인식 방법이다. 이 방법은 인식 처리가 지극히 단순하고, 표준 패턴도 인식 대상 문자의 표본으로부터 일정한 절차에 의해 만들 수 있기 때문에, 일단 장치만 만들면, 특정한 문자를 인식할 수

있도록 장치를 교육하는 일도 자동적으로 수행된다. 유사도법은 인식에 유사도를 이용하지만, 문자 인식 방식으로서의 성능, 특히 여러 가지 품질의 문자를 인식할 때의 인식 정밀도의 좋고 나쁨은 전처리나 상관법의 내용에 의존하게 된다.

유사도법은 미지 문자 및 표준 패턴 데이터를 비교적 원형을 지니게 하여 직관과 대응되기 쉬운 단계에서 양자의 일치도를 계산하게 된다. 이에 비해, 특징 정합법은 문자가 복수 개의 부분 패턴으로 구성되어 있다 생각하고, 부분 패턴의 유무를 적당한 방법으로 추출하여 같은 부분 패턴의 유무 및 조합으로 표현되는 표준 패턴과 비교해서 인식한다. 그러나, 전형적인 유사도법을 포함하는 좁은 의미의 패턴 정합법과 특징 정합법 사이에 확실한 경계가 있는 것이 아니다.

2) 획분석법

이 방법은 문자 인식 장치의 실용화 초기에 가장 널리 이용된 방법이다. 획분석법이 추출하는 특징과 인식 논리를 "+" 및 "-"로 표시하여 대응하는 문자에 그 특징이 존재 또는 존재하지 않는 것을 나타낸다. 획분석법에서 특징 추출이 행해지는 미지 문자는 항상 일부분밖에 보이지 않는다. 이 때문에 확실하게 추출할 수 있는 특징에도 제약이 있으며 반대로 추출하기 쉬운 특징으로 구성된 전용 문자형이 사용되었다. 획분석법은 간단한 처리에 의해 문자를 인식하지만, 그 하드웨어 구성상의 제약 때문에 곡선 부분을 포함하는 자연스러운 형태의 문자를 인식하기는 어렵다. 이러한 이유와 하드웨어의 크기가 그다지 문제가 되지 않는 면도 있어 현재는 거의 사용되지 않는다. 그러나 획분석법의 접근 방식 자체는 특징 정합법과 동일하며, 다른 영역간의 특징 조합을 고려하는 점에서는 보다 복잡한 처리를 행할 수 있는 가능성을 내포하고 있다.

3) 구조 해석법

실용적으로 인식 가능한 글자 종류는 필기 영문자 및 숫자로까지 확장되었다.

구조 해석법은 인식 처리의 추상적 개념으로는 특징 정합법이나 획분석법과 동일하지만, 실제의 인식 처리는 이들 방법이 내포하고 있는 하드웨어 구성상의 제약을 벗어나, 자유로운 방법을 실현한 것으로 볼 수 있다.

구조 해석법의 순서는, 브라운관의 빔 주사를 원형 주사라고 부르는 특수한 방식으로 제어하고 미지 문자의 윤곽 및 기울기를 검출한다. 원형 주사 방식은 비점을 원호 모양으로 회전하면서, 종이로부터의 반사광을 검출하여, 문자의 검은 부분에 들어가면 그 회전 반경을 줄여 비점은 문자의 윤곽을 따라 이동하도록 제어된다. 이 때에 윤곽의 경사와 위치도 검출하여 특징으로 한다. 두드러진 점은 추출하는 특징이 획분석법이나 특징 정합법에 비해 자연스러운 이점이 있으나, 이를 구체화하기 위해 사용하는 하드웨어는 획분석법 등과 같고, 표준 패턴은 인식 논리의 형태로 하드웨어의 연결에 의해 구현된다.

필기 숫자 인식 처리는 크기의 정규화, 획 방향 추출, 수평 특징 추출 및 인식의 4단계로 행해진다. 크기의 정규화는 일정한 값보다 큰 문자를 가로, 세로 모두 1/2로 축소한다. 축소는 원 도형에서의 $2 \times 2 = 4$ 비트 영역을 축소 후의 1비트에 대응시켜 실현한다. 이때, 원 도형의 4비트의 영역에 1비트라는 검은 점이 있으면 축소 후 대응하는 비트는 검은 점으로 한다. 선소의 방향 추출은 3×3 비트의 데이터를 검사하여 부분 특징 중의 하나로 추출한다. 수평 방향 특징 추출은, 미지 문자로부터 선소 방향을 추출한 후 미지 문자를 수평 방향으로 14-16개의 부분으로 자르고, 이 잘린 각각의 부분을 먼저 추출한 획 특징의 조합에 의해 수평 특징으로 표현한다. 미지 문자를 처리해서 얻어진 수평 특징의 계열이 상태 천이도에 입력되며, 만일 마지막 상태에 도달하여 통과하면, 미지 문자는 이 표준 패턴과 일치한다고 본다. 여기서, 미지 문자가 두개 이상의 서로 다른 카테고리에 속하는 상태 천이도를 통과한 경우에는 기각한다.

4) 종래 문자 인식 방법의 문제점

유사도법의 한계는 변형에 약하기 때문에 필기 문자나 저품질의 인쇄 문자에는 적용할 수 없다. 특징 정합법의 결점은 개념으로서의 특징과 실제로 추출한 결과가 일치하지 않는다. 또 특징의 개념 자체가 애매하고 특징간의 관련성을 평가할 타당한 척도를 찾아내기 어렵다. 대부분의 경우, 특징 추출 때 정보 손실을 동반하며, 특징 추출 처리의 내용에 필연성이 없고 시행 착오적이다.

구조 해석법의 한계로는 문자 구성 요소의 접속 관계를 너무 많이 이용하기 때문에, 2차원 도형 상에서 보면 문제가 되지 않을 정도의 사소한 접속 관계의 혼란에도 큰 영향을 받는다.

종래 방식의 문제점은 이상과 같으나, 이들 방법에 의존하는 한, 앞으로의 과제인 필기체 한글 등의 실용적인 인식은 불가능할 것이고, 또 현재 실용화되어 있다고는 하나, 필기 영문자 등의 인식 장치의 능력도 충분하다고 말하기는 어렵다.

그러나, 인쇄 한자의 인식에서 실증된 바와 같이, 적절한 인식 처리만 해주면 글자 종류가 많은 것은 문자 인식을 수행할 때에 본질적인 방해 요인은 되지 않는다. 따라서, 유사도법을 발전시켜서 단일 글자체 인쇄 문자를 인식하는 것과 같은 정도로 필기 문자를 쉽게 인식할 수 있는 방법으로 구성할 수 있다면, 그 인식 방법은 필기 한글이나 필기 한자를, 현재 인쇄 한자를 인식하는 것과 거의 동등한 정도로 처리할 가능성이 있다.

III. 문자 인식 시스템으로서의 신경회로망

1. 신경회로망

신경회로망은 간단한 계산 능력을 가진 처리기들이 서로 복잡하게 연결되어 구성된 컴퓨터 시스템으로서, 외부에서 주어진 입력에 대하여 동적인 상태 반응을 함으로써 정보를 처리하는 시스템이다. 신경회로망에서 중요한 구성 요소는 뉴런 또는 처리 소자라 불리는 처리기와 이들 상호간의 연결이다. 뉴런간의 연결은 방향성 연결선이고 가중치를 가지고 있다. 가중치는 뉴런간에 미칠 수 있는 영향력의 크기를 표시한다. 그러므로 양의 가중치는 흥분시키는 작용을 하고 음의 가중치는 억제하는 작용을 한다. 가중치의 값에 의해 회로망의 동작이 결정되므로 가중치가 전통적인 소프트웨어에서의 프로그램의 역할을 한다(이 와 김, 1990).

여기서 인간의 뇌세포 활동과 신경회로망의 동작 원리를 비교하면, 인간의 뇌세포는 어떠한 상황이 발생하면 이를 인간의 오감을 통하여 뇌세포에 전달되고, 뇌세포는 이를 감지하고 일반적인 활동을 주시한다. 반응을 감시하고 자기 스스로 체계를 갖추게 되며 일련의 사건을 생각한다. 이를 일반화하여 대응 동작을 한다. 이에 대한 신경회로망의 동작은 다수의 입력 신호를 시냅스를 통하여 가중치를 실어 뉴런에게 전달한다. 전달된 신호들은 합산을 하여 문턱치와 비교하여, 문턱치를 넘을 경우 높은 출력을 발생하고 넘지 못할 경우 낮은 출력을 발생하여 상위 뉴런으로 그 값을 전달한다. 이러한 과정은 누적되어 각 뉴런에 전반적으로 저장되며, 학습 규칙을 만들어 가면서 가중치가 저장된다. 이 저장된 가중치를 통하여 차후 입력 데이터를 받아들여도 원하는 출력치를 얻을 수 있게 된다.

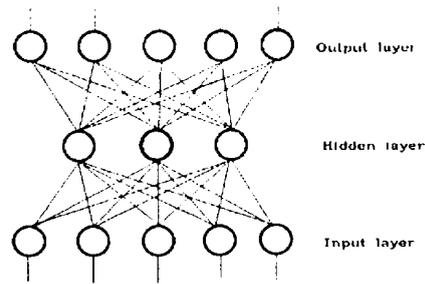


Fig.2 Construction of neural networks

인공 신경회로망의 수학적 모델은 뉴런과 시냅스로 구성된다. 한 뉴런으로부터 전달된 출력 신호는 시냅스 연결을 통하여 다른 뉴런으로 전달된다. 한 뉴런의 입력 신호의 상태는 각각의 가중치가 시냅스 연결의 세기가 되는 다른 뉴런으로부터 전달되는 가중된 입력 신호의 선형적인 합에 의해 결정된다. Fig.2의 회로망은 전방향 회로망이다.

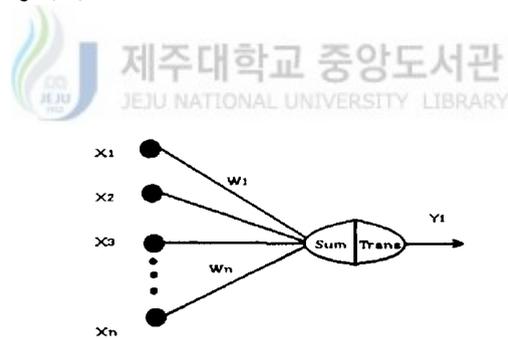


Fig.3 Construction of processing element

신경회로망에 존재하는 뉴런은 뇌에서의 신경세포에 해당하며, 활성화 값이라 불리는 하나의 출력값을 계산하는 작용을 한다. 뉴런은 Fig.3과 같은 구조로 되어 있으며 한 뉴런의 활성화 값은 전단에서 그 뉴런에 직접 연결되어 있는 뉴런들의 출력값과 해당 연결선의 가중치를 이용하여 계산된다. 전형적인 신경회로망 모델에서는 다른 뉴런들로부터의 입력에 연결선의 가중치를 곱하여 그 합을 구하고 그 값이 임계치를 초과하는가를 응답 함수로 계산하여 초과하면 그 값을 그 뉴런의 출력값으로 한다.

$$U_i = \sum_k W_{ki} V_k \quad (1)$$

식(1)에서 V_k 는 k 번째 뉴런의 출력이고, U_i 는 i 번째 뉴런의 입력이다. W_{ki} 는 k 번째 뉴런으로부터 i 번째 뉴런으로 전달되는 시냅스 연결의 세기이다.

McCulloch-Pitts의 뉴런 입/출력 함수는 식(2)와 같다.

$$V_i = f(U_i) = \begin{cases} 1 & \text{if } U_i > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

여기서, V_i 와 U_i 는 각각 i 번째 뉴런의 출력과 입력이다.

이러한 응답 함수는 그 특성에 따라 Fig.4와 같이 구분할 수 있다.

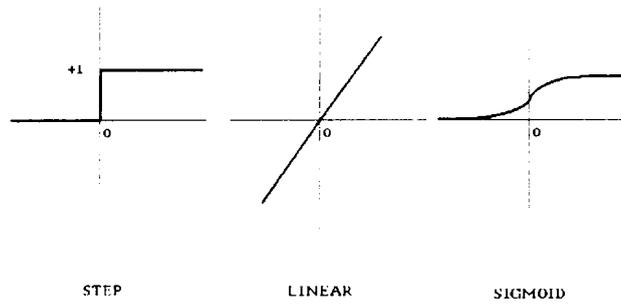


Fig.4 Response functions of processing element

신경회로망에서 지식은 뉴런들이 어떻게 상호 연결되어 있고, 각 연결선의 가중치가 어떤 값을 갖고 있는가에 따라 결정되어 저장된다. 즉 지식이란 특정 장소의 내용이 아니라 회로망 구조에 의해 나타내어진다. 뉴런은 네트워크 상의 역할에 따라 입력 뉴런, 중간 뉴런, 출력 뉴런의 세 계층으로 나뉘어진다. 입력 뉴런은 회로망의 외부로부터 입력을 받아들여 활성화되어 입력 값을 그대로 출력한다. 출력 뉴런은 회로망 전체의 출력으로 사용된다. 입력 뉴런이나 출력 뉴런에 속하지 않는 뉴런을 중간 뉴런이라 하며 선형 비분리 함수와 같은 복잡한 기능을 수행하는데 필요하다. 신경회로망을 구조에 따라 방향성 사이클이 존재하지 않는 전방향 회로망과 방향성 사이클이 포함되어 있는 역방향 회로망으로 나눌 수 있다. 신경회로망에서 학습이란 주어진 신경회로망이 우리가 원하는 동작을 수행할 수 있는 가중치를 찾아내는 작업을 의미한다. 이러한 학습은 입력 값과 우리가 기대하는 출력값을 가지고 가중치를 조정해 나감으로써 이루어진다.

2. 문자 인식 시스템으로서의 신경회로망

신경회로망이란 신경조직에서 착안하여 모델화 한 정보처리 시스템으로서 단순한 소자들이 병렬, 분산 연결 구조를 가지고 있다. 외부로부터 입력을 받아들여 동적인 반응을 일으킴으로써 필요한 출력을 생성시킨다. 따라서 신경회로망은 기존의 순차형 컴퓨터에서 비효율적으로 처리되던 패턴 인식 문제를 효율적으로 해결하는데 매우 적합하다. 즉, 신경회로망은 패턴 인식에 필요한 많은 양의 데이터를 병렬 처리 능력을 이용하여 인식 대상 문자 수에 상관없이 인식에 걸리는 시간을 일정하게 할 수 있다. 디지털 컴퓨터의 순차적인 정보처리 방식과는 달리 수많은 뉴런들에 분산 저장되어 있는 정보가 대규모 병렬 연산에 의해 즉각적으로 처리된다.

그리고 패턴 인식과 같은 수학적 알고리즘의 적용이 곤란한 문제도 학습에 의하여 효과적으로 처리할 수 있다. 신경회로망에서 가장 중요한 특징 중의 하나가 학습 능력이다. 신경회로망에서의 학습이란 주변으로부터의 자극과 응답에 대한 경험적 정보를 일반화하고 그 데이터로부터 신경회로망 내부에 지식 기반을 형성, 입력 정보에 대해 적절하고 융통성 있는 출력을 생성하도록 하는 것이다. 학습의 궁극적인 목표는 원하는 출력과 가중치를, 목표치와 현 출력과의 오차를 최소화하도록 바꿔 나가는 것이다.

신경회로망은 패턴 인식에 있어 자주 발생하는 잡음을 포함하거나 애매한 데이터를 효과적으로 처리할 수 있다. 신경회로망을 이용한 시스템은 학습과 기억을 통하여 문자를 인식하는 동적인 시스템이다. 이러한 동적인 특성으로 인하여 훈련되지 않은 다른 문자 체에도 약간의 학습 과정을 통하여 쉽게 적응할 수 있다. 신경회로망 시스템은 간단한 학습에 의해 개발이 가능하므로 다른 방식에 비해 개발 기간이 훨씬 단축된다.

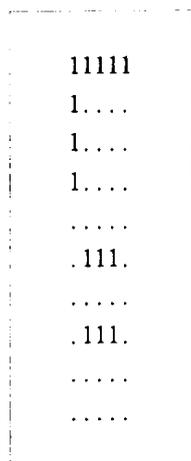
IV. 영문자 인식용 신경회로망 모델 구현

본 논문에서 영문자 인식을 위해 영문자 26자에 대한 입력 패턴을 준비하고 이를 신경회로망에 입력으로 주어 학습시킨다. 문자 인식은, 우선 입력 패턴의 정형화, 잡음 제거, 세션화 과정을 거치고 특징점 추출을 하여야 한다. 이렇게 추출된 특징 정보를 신경회로망을 통하여 분류시킴으로서 문자 인식이 가능하게 된다. 정형화 및 세션화 과정을 거친 패턴은 20x20의 비트 행렬이 되는데 이를 특징점 추출과 분류를 위해 특징점 추출 알고리즘을 이용한다. 영문자는 획을 분류



Fig.5 Detection of local distribution

해 보면 가로, 세로 혹은 원 등의 선분들의 집합이다. 이러한 선분들의 특징 및 국소적 분포 성분을 그 문자의 특징 벡터로 추출하기 위하여 국소 분포 성분 행렬을 Fig. 5와 같이 추출한다. 4비트의 검색창이 입력 패턴을 50개의 국소 분포 성분으로 특징 행렬을 정의할 수 있다. 검색창에 '1'인 화소가 검출되면 그 지역의 특징 행렬은 '1'의 값을 갖는다. Fig. 6는 국소 분포로부터 얻어진 50개의 특징 행렬이다.




제주대학교 중앙도서관
 JEJU NATIONAL UNIVERSITY LIBRARY
 Fig. 6 Detection feature on local distribution

본 논문에서 사용한 신경회로망의 구조는 입력층과 출력층 그리고 그 사이에 중간 처리층을 두어 3개의 층으로 구성되어진다. 입력층은 50개의 입력단을 갖고 있고 이 입력단을 통하여 특징 추출 행렬을 입력으로 받아들여진다. 그리고 중간 층은 30개에서 80개의 뉴런으로 변화시키면서 구성하였고 출력단은 26개의 뉴런으로 영문자 갯수와 일치 시켰다. 각 뉴런간의 연결은 전방향으로 모두 연결되어 있다. 각 뉴런은 연결되어 있는 전단의 출력과 가중치를 곱한 값의 총 합을 식 (3)과 같이 구하여 이 값을 입력으로 하여 문턱치와 비교 처리한 후 이를 출력하

여 상위층 뉴런으로 전달된다.

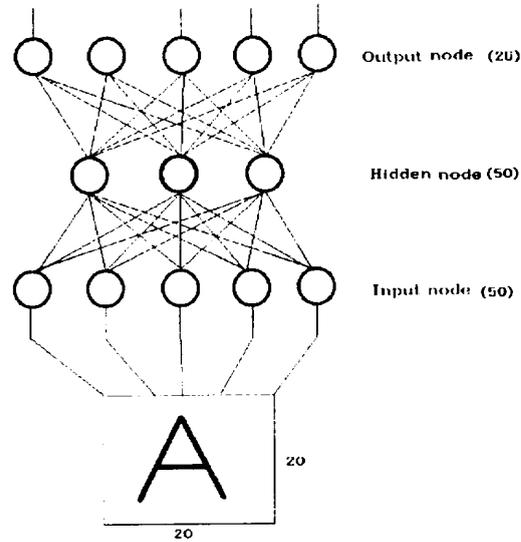
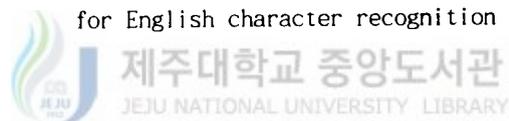


Fig.7 Construction of neural networks



$$net_j = \sum_i o_i \cdot w_{ji} + \theta_j \quad (3)$$

$$o_j = \frac{1}{(1 + e^{-net_j})} \quad (4)$$

여기서 첨자 i 와 j 는 뉴런의 번호이고 o_i 는 i 번째 뉴런의 출력이다. w_{ji} 는 j 번째 뉴런과 i 번째 뉴런간의 가중치이며, θ_j 는 j 번째 뉴런의 문턱치이다. 따라서 net_j 는 j 번째 뉴런으로 들어오는 입력의 합이며 이 값은 식(4)와 같은 응답 함수를 통과하여 o_j 라는 j 번째의 출력값이 결정되어 진다. 식(4)의 응답 함수는 시그

모이드 함수로서 미분 가능하고 단조 증가하는 특징을 가지고 있다.

신경회로망의 학습 방법은 Rumelhart(1986)가 제안한 오류 역방향 전파(EBP : Error Back Propagation)를 사용하였다. 입력층과 중간층 그리고 출력층으로 구성된 신경회로망을 오류 역방향 전파 알고리즘을 통하여 필요한 정보를 저장하도록 학습시킬 수 있다. 역방향 전파는 출력의 기대치와 실제 값의 오차를 감소하는 방향으로 연결 강도를 조절하고 상위층의 오차를 현재의 연결 강도를 가중치로 하여 다음 하위층에 역전파하며 하위층에서는 이를 근거로 하여 자기층의 연결 강도를 조정해 나간다. 이 신경회로망을 3계층으로 구성할 경우 이론상으로는 어떠한 형태의 패턴도 형성할 수 있으나, 학습시 국소극소에 빠질 우려가 있고, 학습 과정이 너무 오래 걸리고, 또 기억된 패턴의 수정, 추가 학습 등이 불가능하다는 단점이 있다. 하지만 이 신경회로망은 구현이 쉽고 학습 방식이 용이하기 때문에 현재 가장 폭넓게 응용되고 있다. 어떤 입력을 주었을 때 의도 출력값 (t_j)와 실제의 출력값 (o_j)와의 차를 식(5)와 같이 정의한다.

$$E = \frac{1}{2} (t_j - o_j)^2 \quad (5)$$

식(5)에서 E는 j번째 출력단의 오차 값을 나타낸다. 신경회로망의 학습은 이 오차를 줄이도록 접합 가중치를 조정하는 것이다. 여기서 입력을 주었을 때의 j번째와 i번째의 가중치 w_{ji} 의 변화량은 식(6)과 같다.

$$\Delta w_{ji} = - \frac{\partial E}{\partial w_{ji}} \quad (6)$$

식(6)의 우변은 식(7)과 같이 다시 쓸 수 있다.

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}} \quad (7)$$

식(3)에서 i대신 k를 변수로 쓴다면 $net_j = \sum_k o_k \cdot w_{jk}$ 이므로

식(7)의 우변의 우측은 $\frac{\partial net_j}{\partial w_{ji}} = \frac{\partial}{\partial w_{ji}} \sum_k o_k \cdot w_{jk} = o_j$ 가된다.

또한 좌변 식은 δ_j 를 j번째 항의 입력 합인 변화량에 대한 오차의 변화량이라 두면 $\delta_j = -\frac{\partial E}{\partial net_j}$ 이고, $\frac{\partial net_j}{\partial w_{ji}} = o_j$ 로부터 $-\frac{\partial E}{\partial w_{ji}} = \delta_j \cdot o_j$ 가되어 간략화된다.

그러므로 $\Delta w_{ji} = \eta \cdot \delta_j \cdot o_j$ 가되며 여기서 η 는 상수로서 학습률이다. δ_j 는 뉴런이 중간층이나 출력층이나에 따라 다르다. 출력층 뉴런일 경우 $E = \frac{1}{2}(t_j - o_j)^2$, $o_j = f_j(net_j)$ 이므로 식 (8)과 같다.

$$\begin{aligned} \delta_j &= -\frac{\partial E}{\partial net_k} \\ &= -\frac{\partial E}{\partial o_j} \cdot \frac{\partial o_j}{\partial net_k} \\ &= (t_j - o_j) \cdot f'_j(net_j) \end{aligned} \quad (8)$$

중간층일 경우, $(t_j - o_j)$ 에 대응하는 항은 식(9)와 같다.

$$\begin{aligned} -\sum_k \frac{\partial E}{\partial net_k} \cdot \frac{\partial net_k}{\partial o_j} &= -\sum_k \frac{\partial E}{\partial net_k} \cdot \frac{\partial}{\partial o_j} \sum_i w_{ki} \cdot o_i \\ &= -\sum_k \frac{\partial E}{\partial net_k} \cdot w_{kj} \\ &= \sum_k \delta_k \cdot w_{kj} \\ \delta_j &= f'_j(net_j) \sum_k \delta_k \cdot w_{kj} \end{aligned} \quad (9)$$

Δw 의 계산은 출력단으로부터 시작하여 중간층 뉴런에 전달된다. 중간층 뉴런의 Δw 를 계산하기 위해서는 앞단의 Δw 가 계산되지 않으면 안된다. 그러므로 최후의 입력단까지 거슬러 올라가야만 처음으로 계산이 가능하게 된다. 이러한

역방향 전파가 이루어져야 학습을 할 수 있으므로 역방향 전파 학습법이라 한다.

일반적인 Δw 를 정식화한 것이 식 (10)과 같다.

$$\Delta_{ji}(n+1) = \eta \cdot \delta_j \cdot o_i + \alpha \cdot \Delta w_{ji}(n) \quad (10)$$

여기서 n 은 학습 횟수, η 는 학습률(learning ratio), α 는 변화율(momentum)이며, $\alpha \cdot \Delta w_{ji}(n)$ 은 오차의 진동을 적게 하여 수렴 속도를 빨리 하기 위하여 첨가한 것이다.

본 논문에서는 학습 횟수는 오차가 0.00001 이하 일 때까지로 정하였으며 중간층 뉴런의 갯수와 학습 속도 및 인식률에 따른 관계를 살펴보고 최적의 중간층 뉴런 수를 선정하여 다시, 변화율 α 를 0.1에서 0.5까지 변화시키면서 학습 속도와 인식률을 분석하였다. 설정된 중간층 뉴런 수에서의 학습률 η 의 영향을 분석하기 위하여 0.1 - 0.5를 주어 학습 및 인식을 시켰다.



V. 알고리즘 흐름도 및 시뮬레이션 결과

1. 영문자 인식 알고리즘

본 논문에서 제시된 영문자 인식을 위한 알고리즘을 나타내면 다음과 같다.

[과정 1] 20×20비트 행렬로부터 국소 분포 특성을 찾아내기 위하여 정규화를 거친 후 20×20행렬을 5×5의 국소 지역으로 분할하고 5×5의 가로 검색창과 5×5의 세로 검색창을 통하여 50개의 특성 벡터를 추출한다.

[과정 2] 특성 벡터를 신경망의 입력 뉴런에 입력으로 주어 전방향 전파한다. 출력은 임의의 가중치로부터 임의의 출력값을 갖으므로 원하는 출력값과 비교하여 오차를 계산한다.

[과정 3] 오차 값을 이용하여 역방향 전파를 하면서 가중치를 조정해 나간다.

[과정 4] 조정된 가중치로 다시 전방향 전파를 하고 오차를 계산한다.

[과정 5] 계산된 오차 값이 허용 범위를 넘을 경우 과정 3과 과정 4를 반복한다.

[과정 6] 계산된 오차 값이 허용 범위를 넘지 않을 경우 가중치를 저장하고 그 가중치를 이용하여 문자 인식을 한다.

2. 시뮬레이션 순서도

본 논문에서 제시한 영문자 인식 알고리즘을 구현해 보기 위하여 C언어로 시뮬레이션 하였으며 Fig.8은 알고리즘 흐름도 이다.

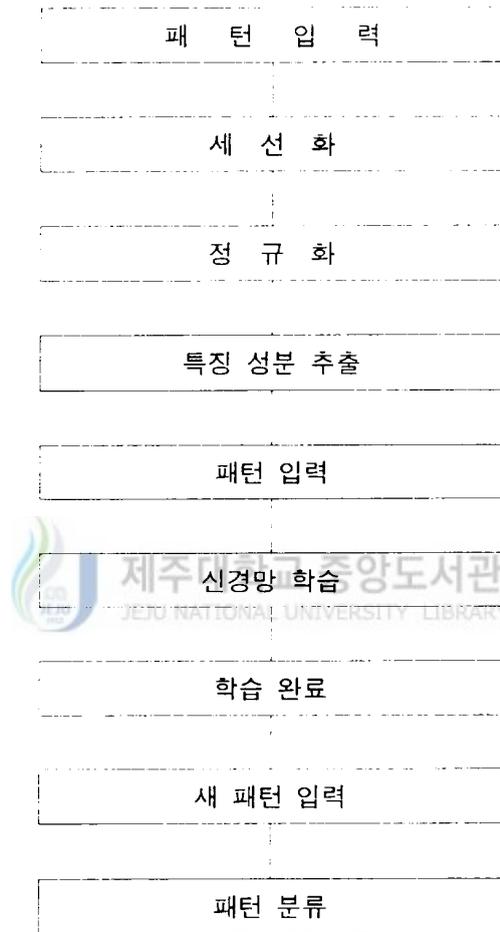


Fig.8 System flowchart

Fig.9 은 영문자 "E"와 "F"에 대한 입력 패턴이다. "E"와 "F"를 예를 든 것은 먼저 "E"패턴이 하단 부의 존재 유무가 "E"와 "F"를 구분하는 특징이 되기 때문인데 유사도가 있음에도 이러한 특징은 "E"와 "F"패턴 구분을 명확히 해주었다.

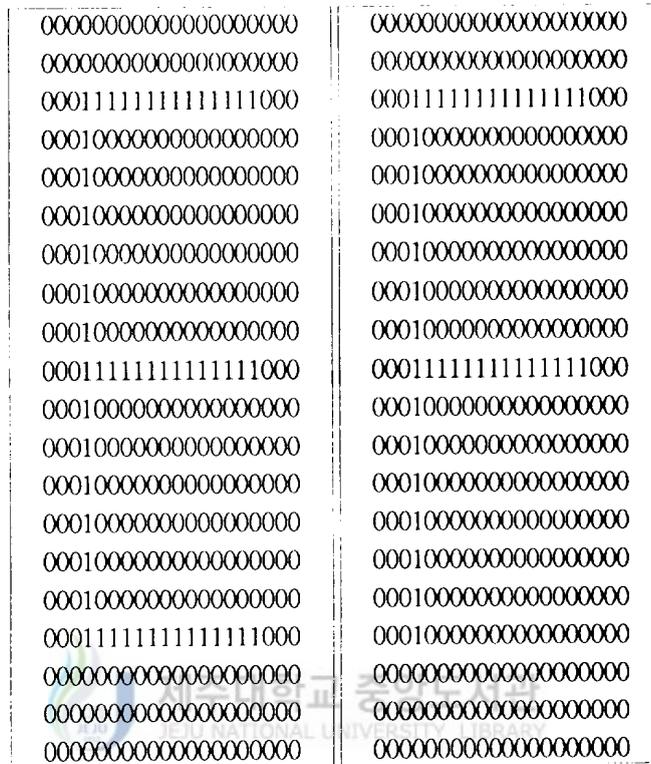


Fig.9-a Input pattern of "E" Fig.9-b Input pattern of "F"

Fig.9 Input patterns of "E" and "F"

"E"와 "F"입력 패턴으로부터 국소 분포를 검출한 특징 행렬을 Fig.10에 나타내었다. Fig.10의 국소 분포 행렬로 신경망을 학습시켜 학습이 완료된 그림이 Fig.11에 보여진다.

1111
 1....
 1....
 1....

 .111.

 .111.

 .111.

Fig.10-a Distribution characteristic of "E"



1111
 1....
 1....
 1....

 .111.
 .111.

제주대학교 중앙도서관
 JEJU NATIONAL UNIVERSITY LIBRARY

Fig.10-b Distribution characteristic of "F"

Fig.10 Detection of the characteristic on local area

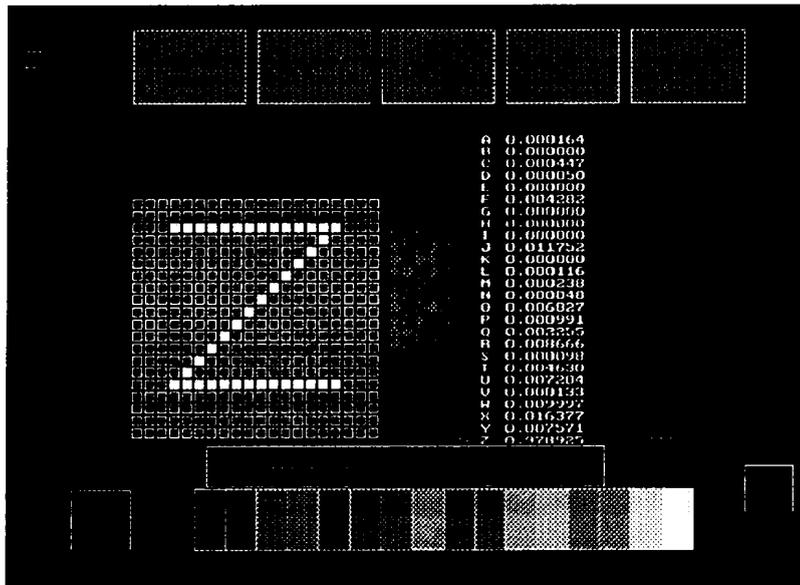


Fig.11 Learning pattern

학습이 완료된 후 다시 신경망에 입력 패턴 "E"와 "F"를 주었을 때의 결과는 Fig.12와 Fig.13에 보여진다. Fig.12에서 "E"로의 인지도는 0.976298 이며 "F"로의 인지도는 0.02149 이며, Fig.13에서는 "E"로의 인지도는 0.016929 이며 "F"로의 인지도는 0.959140 으로 나타나 확실히 구분함을 보여준다. Fig.14은 이렇게 학습된 신경회로망에 영문자 "A"가 예러가 있을 때의 인식 정도인데 인지도는 0.895935 이다.

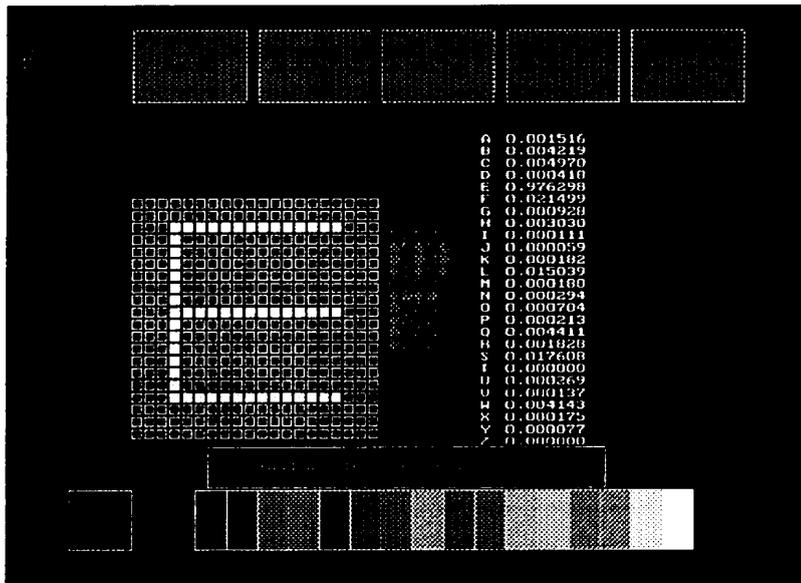


Fig.12 The result of "E"

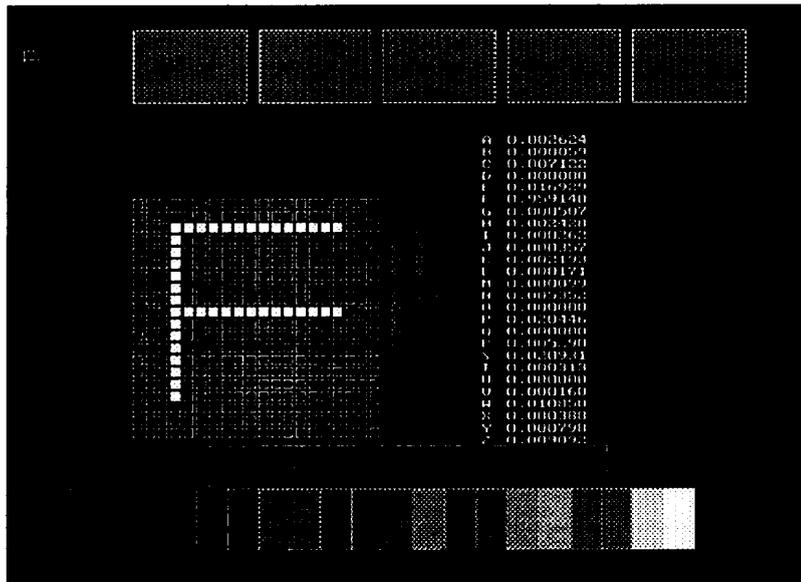


Fig.13 The result of "F"

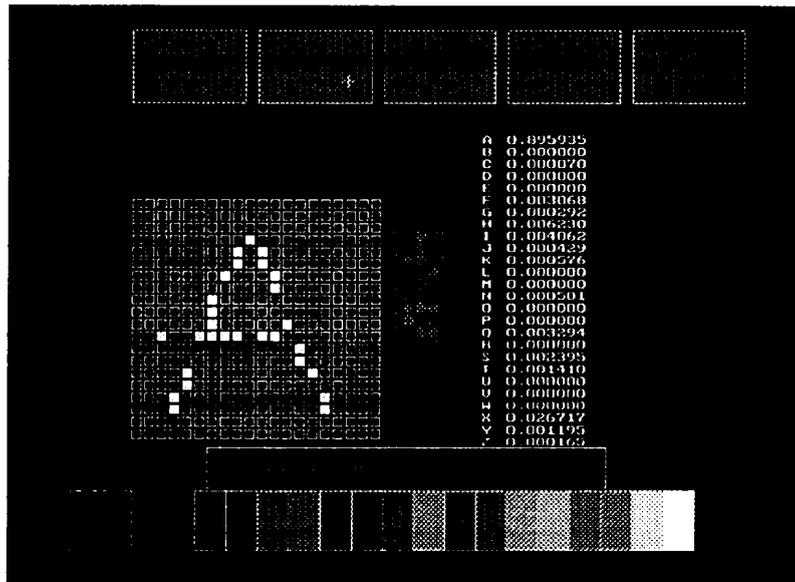


Fig.14 The result of "A" with error

Table 1 The number of learning and recognition ratio according to hidden neurons (momentum=0.1)

중간 뉴런 수	학습 종료시 학습 횟수	인식률 % (학습 패턴)	인식률 % (미학습 패턴)
30	2290	96.2	76.9
40	1588	92.3	80.3
50	1476	96.2	81.5
60	1171	76.9	70.76
70	980	80.8	73.1
80	911	76.9	65.4

Table 1은 영문자 인식 시스템의 중간층 뉴런의 수에 따른 학습 속도와 인식률의 변화를 알아보기 위하여 중간층 뉴런의 수를 30개에서 80개까지 변화시키면서 학습시키고 학습이 완료된 후 인식률을 알아보았다. 이때의 학습률은 0.1이고 변화율도 0.1로 정하였다. Table 1에서 알 수 있듯이 중간층 뉴런의 수가 많아지면 학습 속도는 빨라지지만 인식률에 있어서는 50개 이상일 경우 오히려 떨어지고 있다. 이는 적당한 수까지의 증가는 정보의 분산 효과가 커서 각 뉴런이 학습할 부분이 적어지므로 효율적이지만 필요 이상으로 많아질 경우 필요 이상의 정보 분산으로 학습률이 떨어져 인식률도 떨어지고 있음을 알 수 있다.

Table 2 The number of learning and recognition ratio
according to momentum (learning ratio = 0.5)

변화율	학습 종료시 학습 횟수	인식률 % (학습 패턴)	인식률 % (미학습 패턴)
0.1	1476	96.2	81.5
0.2	1303	96.2	80.0
0.3	1174	96.2	81.5
0.4	939	96.2	77.7
0.5	745	96.2	86.2

Table 2는 영문자 인식 시스템의 변화율 α 값의 변화에 따른 학습 속도와 인식률의 변화를 알아보기 위하여 변화율을 0.1에서 0.5까지 변화시키면서 학습시키고 학습이 완료된 후 인식률을 알아보았다. 이때의 중간층 뉴런의 수는 인식률을 높게 나타낸 50개로 정하였고 학습률은 0.1로 정하였다. Table 2에서 보면 변화

율의 증가는 학습 속도를 빠르게 하지만 이렇게 빠른 종료로 인해 충분한 학습이 이루어지지 못하여 인식률이 저하되었다. 그러나 필요 이상의 학습은 필요치 않으므로 변화율의 조정으로 적정의 학습 횟수를 정하는 것이 좋다.

Table 3 The number of learning and recognition ratio according to learning ratio (learning ratio = 0.5)

학습률	학습 종료시 학습 횟수	인식률 % (학습 패턴)	인식률 % (미학습 패턴)
0.5	735	96.2	86.2
0.6	730	96.2	83.5
0.7	519	96.2	78.8
0.8	457	96.2	84.2
0.9	398	96.2	86.2
1.0	346	92.3	82.7
1.5	230	92.3	84.2
2.0	153	96.2	84.6

Table 3는 영문자 인식 시스템의 학습률 0.1값의 변화에 따른 학습 속도와 인식률의 변화를 알아보기 위하여 학습률을 0.1에서 0.5까지 변화시키면서 학습시키고 학습이 완료된 후 인식률을 알아보았다. 이때의 중간층 뉴런의 수는 인식률을 높게 나타낸 50개로 정하였고, 변화율은 0.5로 정하였다. Table 3에서 보면 학습률의 증가는 학습 속도를 빠르게 하지만 인식률에 있어서는 많은 변화를 보이지

않아 적정한 학습률은 학습 속도를 빠르게 하는데 중요한 요인이 됨을 알 수 있다. 그러나 학습률 η 값이 2.0이 이상에서는 학습 속도는 빨라지나 인식률이 떨어져 더 이상 영문자 인식 시스템에 좋은 요인으로 작용하지 못했다.



VI. 결 론

본 논문에서는 영문자 인식을 위해 전처리 과정 중 입력 패턴에 대해 특징을 추출하였다. 입력 패턴의 화소의 국소적 분포를 특성 벡터로 표현하여 이를 신경 회로망 입력으로 사용하였다. 이 입력 벡터로부터 역방향 전파 학습법으로 학습된 신경회로망은 새로운 입력 패턴에 대해서도 학습된 가중치와 문턱값, 그리고 특성 행렬을 통하여 영문자 인식을 수행함으로써 예러가 섞인 문자에 대해서도 인식할 수 있음을 보였다. 본 논문에서는 학습 횟수는 오차가 0.00001 이하 일 때까지로 정하였으며, 중간층 뉴런의 갯수와 학습 속도 및 인식률에 따른 관계를 살펴보고 최적의 중간층 뉴런 수를 선정하였다. 그리고, 변화율 α 값을 0.1에서 0.5까지 변화시키면서 학습 속도와 인식률을 분석하였다. 설정된 중간층 뉴런 수에서의 학습률 η 의 영향을 분석하기 위하여 η 값을 0.1 - 0.5를 주어 학습 및 인식을 시켰다. 학습 결과 중간층 뉴런의 수가 30일 경우 2290회에서 학습을 마쳤고, 50일 경우는 1476회 및 80일 경우 911회에 학습을 마쳤다. 학습된 신경망을 통하여 학습 데이터의 인식률은 96.2%이고 미학습 패턴에 대해서는 81.5%의 인식률을 보였다. 추후 연구 과제로는 필기체 한글의 인식과 다양한 패턴 분류이고 이들의 하드웨어 구현이다.

참 고 문 헌

- . Basit Hussain and M. R. Kabuka, "A Novel Recognition Neural Network and its Application to Character Recognition", IEEE transactions on pattern analysis and machine intelligence. Vol.16. NO.1. Jan. 1994., pp.98-106
- . Bart Kosko, "Adaptive bidirectional associative memories", APPLIED OPTICS Vol.26, No.23, Dec.1987, pp4947-4959
- . Bart Kosko, "Constructing an Associative Memory", BYTE Sep.1987, pp137-144
- . D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning Internal Representations by Error Propagation", Parallel Distributed Processing, D.E. Rumelhart and J. L. McClelland, eds., Ch. 8, Cambridge, MA : MIT Press, 1986.
- . K. Fukushima, "A Neural Network for Visual Pattern Recognition", IEEE Computer, Vol. 21, no 3, pp 65-75, 1988.
- . 남호원, 정호선, "영문자 인식 및 전처리용 신경칩의 설계", 한국전자통신학회 논문지, Vol.15 No6, 1990, pp455-466
- . 오영환, "패턴 인식의 개관", 정보과학지 제 11권 5호, 1993.10, pp11-20
- . 이광노, 김명원, "Neural network을 이용한 필기체 한글 자소 인식", 전자 통신 제12권 3호, 1990.10, pp77-91
- . 이광노, 장명욱, 박치선, 이훈복, "패턴 인식을 위한 신경망- 지식 기반 융합 모델", 전자 통신 제 14권 4호, 1993.1, pp125-136

본 논문과 관련된 저자의 발표 논문

1) 현해진, 김덕주, 이기태, 임재운, 1993, “신경회로망에서 효과적인 조정 정보 전달에 의한 4색 배치 문제 해결 알고리즘“, 한국통신학회 추계 종합 학술 발표회 논문집, pp. 295-298.

2) 현해진, 김덕주, 강부식, 이기태, 임재운, 1993, “신경회로망에서 뉴런간의 연결도를 이용한 4색 배치 문제 해결 알고리즘“, 산업 기술 연구보고 제4권, pp. 51-57.

3) 김덕주, 현해진, 이기태, 임재운, 1994, “신경회로망을 이용한 효과적인 영문자 인식 알고리즘“, 대한전자공학회, 한국통신학회 제주 지부 합동 학술 발표 논문집, pp. 28-33.



感謝의 글

본 論文이 完成되기까지 未洽한 저를 指導와 鞭撻로 이끌어 주신 指導教授 林載允 教授님께 感謝드립니다. 未備한 論文을 細心하게 살펴 주시고 忠告해 주신 李鎔鶴 教授님, 張宰成 教授님, 文建 教授님, 金興洙 教授님께 感謝드리며, 先輩로서 論文이 完成되기까지 많은 도움을 주신 梁斗榮 教授님, 康鎭植 教授님께도 感謝드립니다.

研究室의 中心으로 恒常 念慮하여 주신 李起兌 先輩님, 玄海辰 先輩님과 마이크로파 研究室의 康副植 先輩님께 感謝드리며 原稿 整理와 어려울 때도 웃음으로 많은 도움을 준 後輩 景培, 映希, 昌旻, 祺凡, 香辰에게도 고마운 마음을 傳합니다. 그리고, 바쁜 가운데도 時間을 配慮해 주고 많은 도움을 주신 洪成旭 助教님께도 感謝드립니다.

恒常 精誠으로 보살펴 주시고 念慮하여 주신 어머님과 곁에서 힘이 되어 준 兄님들과 누님들께 感謝드리며, 어려움 속에서도 웃음을 잃지 않고 사랑으로 도움이 되어 준 아내와 귀여운 딸 慧璘과 함께 기쁨을 나누고 싶습니다. 끝으로 故人이 되셨지만 제일 기뻐하실 아버님께 이 論文을 바칩니다.