Journal of The Korea Society of Computer and Information
Vol. 24 No. 2, pp. 193-200, February 2019
www.ksci.re.kr
https://doi.org/10.9708/jksci.2019.24.02.193

Sign Language Image Recognition System Using Artificial Neural Network

Hyung-Hoon Kim*, Jeong-Ran Cho**

Abstract

Hearing impaired people are living in a voice culture area, but due to the difficulty of communicating with normal people using sign language, many people experience discomfort in daily life and social life and various disadvantages unlike their desires. Therefore, in this paper, we study a sign language translation system for communication between a normal person and a hearing impaired person using sign language and implement a prototype system for this. Previous studies on sign language translation systems for communication between normal people and hearing impaired people using sign language are classified into two types using video image system and shape input device. However, existing sign language translation systems have some problems that they do not recognize various sign language expressions of sign language users and require special devices. In this paper, we use machine learning method of artificial neural network to recognize various sign language expressions of sign language users. By using generalized smart phone and various video equipment for sign language image recognition, we intend to improve the usability of sign language translation system.

▸Keyword: Sign Language Translation, Artificial Neural Network, Machine Learning, Image Recognition, Hearing Impaired Person

I. Introduction

과학기술의 발전으로 사람들 사이의 소통이 다양한 형태로 발전되어 왔으며, 더욱더 효율적이고 편리한 환경으로 변화되어 가고 있다. 그러나 대부분의 상호간 소통을 위한 환경은 비장애인들을 대상으로 하고 있어 장애인의 경우에 대해서는 많은 어려움이 존재하고 있다. 청각장애인은 선천적 또는 후천적인 요인으로 청신경의 기능에 이상이 생겨 소리를 듣지 못함으로 인해 야기되는 이차적인 장애로 언어에 대한 장애를 겪게 된다. 청각 장애인들은 수화라는 소통 수단을 사용하여 상호간의 소통을 이루어가고 있지만, 비장애인들과의 소통에는 많은 어려움이 존재한다. 청각 장애인은 음성 문화권에서 생활하고 있지만 수화를 사용하여 정상인과 소통해야 하는 어려움으로 인하여 일상생활 및 사회생활에 있어서 많은 불편함과 본인의 욕구와 다르게 여러 가지 불이익을 받고 있는 상황이다.[1]

조사된 통계 자료에 의하면 청각장애인은 수화라는 소통 수단을 사용하는 특별함을 가지고 있으나 독립적으로 외출이 가능하고 사회 각 분야에 진출하여 역할을 수행하고 있는 것으로 파악되고 있다(표 1). 보건복지부 한국보건사회연구원에서 2015년 조사한 자료에 의하면 본인이 병의원에 가고 싶을 때 가지 못한 경험이 있는지 알아본 결과 19.1%가 있다고 응답하여 진료가 필요하나 받지 못한 경우가 있는 것으로 나타나고 있다. 이 가운데 뇌병변장애에 이어 청각장애인이 19.5%의 비율로 다른 장애 유형에 비해 높게 나타나 있다(표 2).[2]

상호간에 원활한 소통이 이루어지지 않았을 때 발생되는 여러 가지 문제적 상황 가운데 특히 병원진료와 같은 경우에는 정상인과의 최소한의 소통이라도 반드시 이루어져야 하는 환경이라 할 수 있다. 이에 본 논문에서는 청각 장애인이 독립적 사회 진출이 가능하지만 수화를 이용한 정상인과의 소통에 대한 어려움으로 인하여 겪는 문제점을 해소하고자 수화 번역 시스템을 검토하였다.

• First Author: Hyung-Hoon Kim, Corresponding Author: Jeong-Ran Cho
*Hyung-Hoon Kim (hhkim@kwu.ac.kr), Dept. of Cosmetic Science, Kwangju Womens University
**Jeong-Ran Cho (jrcho@kwu.ac.kr), Dept. of Health Administration, Kwangju Womens University
• Received: 2019. 01. 11, Revised: 2019. 02. 12, Accepted: 2019. 02. 12.

Table 1. Going out status of persons with disabilities - by year, by type of disability

(단위 : 명, %)

구 분	전국추정수	혼자 외출 여부	외출 빈도					
			전 체	거의 매일	주1~3회	월1~3회	연 10회 이내	전혀 외출 않음
2008년	2,137,226	83.6	100.0	66.6	21.1	7.1	3.9	1.2
2011년	2,611,126	85.7	100.0	66.6	20.0	8.5	-	4.9
2014년	2,646,064	84.3	100.0	67.3	20.3	7.2	-	5.2
지 체	1,339,569	92.9	100.0	73.2	19.4	5.1	-	2.4
뇌병변	289,257	55.7	100.0	47.0	21.3	14.3	-	17.4
시 각	274,484	83.6	100.0	66.9	21.1	8.4	-	3.5
청 각	277,432	94.0	100.0	69.2	22.1	6.1	-	2.6
언 어	21,399	85.1	100.0	66.9	14.7	8.6	-	9.9
지 적	182,333	59.5	100.0	69.8	16.7	7.0	-	6.5
자폐성	18,951	33.6	100.0	86.4	10.1	1.7	-	1.8
정 신	105,002	79.3	100.0	40.3	25.4	17.8	-	16.5
신 장	71,876	88.8	100.0	63.7	30.7	2.1	-	3.5
심 장	8,456	88.9	100.0	56.1	27.1	16.8	-	0.0
호흡기	15,985	71.6	100.0	52.8	14.7	15.4	-	17.1
간	11,839	96.9	100.0	78.5	12.7	4.4	-	4.4
안 면	3,012	100.0	100.0	73.8	26.2	0.0	-	0.0
장루·요루	16,705	91.0	100.0	72.5	18.3	9.2	-	0.0
뇌전증	9,764	75.7	100.0	67.8	16.1	9.2	-	6.9

주 : 지난 1년간 평균 외출 횟수에 대한 응답 값임
자료 : 보건복지부·한국보건사회연구원, 「2008년 장애인 실태조사」, 2009.
보건복지부·한국보건사회연구원, 「2011년 장애인 실태조사」, 2012.
보건복지부·한국보건사회연구원, 「2014년 장애인 실태조사」, 2015.

Table 2. Experiences that disabled people can not get to the clinic - Types of disability

(단위 : 명, %)

구 분	전국추정수	전 체	예	아니오
전 체	2,646,064	100.0	19.1	80.9
지 체	1,339,569	100.0	19.5	80.5
뇌병변	289,257	100.0	25.1	74.9
시 각	274,484	100.0	17.8	82.2
청 각	277,432	100.0	19.5	80.5
언 어	21,399	100.0	19.2	80.8
지 적	182,333	100.0	15.2	84.8
자폐성	18,951	100.0	8.9	91.1
정 신	105,002	100.0	16.6	83.4
신 장	71,876	100.0	13.3	86.7
심 장	8,456	100.0	14.1	86.0
호흡기	15,985	100.0	18.9	81.1
간	11,839	100.0	9.3	90.7
안 면	3,012	100.0	18.3	81.7
장루·요루	16,705	100.0	14.1	86.0
뇌전증	9,764	100.0	11.4	88.6

주 : 1) 최근 1년간 본인이 병의원에 가고 싶을 때 가지 못한 경험 유무 (치과 제외)
2) 2014년 기준임
자료 : 보건복지부·한국보건사회연구원, 「2014년 장애인 실태조사」, 2015.

수화번역시스템에 대한 일반적인 접근 방법은 크게 수화영상에 대한 영상처리 및 인식기법에 기반 한 수화번역시스템과 형상입력장치인 데이터 글로브와 같은 물리적센서를 사용하여 손가락의 구부림과 손의 방향에 대한 정보 입력장치를 기반 한 수화 번역시스템이 연구되고 있다. 형상입력장치인 데이터 글로브를 착용한 수화번역시스템은 손가락의 움직임을 감지하는 방식으로 수화형상을 인식하고 번역하는 시스템이다. 또한 기존의 수화영상에 대한 영상처리 및 인식기법은 표준 수화영상에 대한 결정된 특징정보를 담고 있는 수화영상데이터베이스를 근거로 수화영상을 인식하고 번역하는 시스템이다.[1][7][8][14]

데이터 글로브와 같은 형상입력장치를 사용한 수화번역시스템은 수화를 인식하기에 필요한 상세한 움직임 감지가 어렵다는 기술적인 문제점과 함께 수화번역시스템 사용성 측면에 있어서 특별한 장치 등 이용 환경을 요구함으로써 불편함 등 현실화의 어려움이 존재한다. 또한 기존의 수화영상에 대한 영상처리 및 인식기법은 수화영상에 대한 몇 가지 결정된 특징정보를 담고 있는 수화영상데이터베이스를 근거로 하고 있기 때문에 수화 사용자의 다양한 수화 표현을 인식하기 어렵고, 변화되는 수화영상 및 기호에 대해 적응성이 떨어진다는 문제점이 있다.

기존 수화번역시스템이 공통적으로 가지고 있는 문제점은 수화 사용자의 다양한 수화 표현을 인식하면서 사용의 편리성을 좋게 하는 수화번역시스템으로 만들기가 어렵다는 것이다. 기존 수화번역시스템은 수화에 대한 데이터베이스를 구성하여 비디오 영상 또는 형상입력장치로부터 인식된 수화를 저장된 데이터베이스의 내용을 근거로 번역하는 체계를 가지고 있다. 따라서 수화 사용자의 다양한 수화의 특징을 모두 반영한 시스템을 구현하기란 쉽지 않은 근본적인 문제점을 가지고 있다.

본 논문에서는 기존 수화 번역시스템의 한계점과 단점을 보완할 방법으로 인공신경망을 이용한 수화 번역시스템을 구상하고 이에 대한 프로토타입 시스템을 숫자 수화를 대상으로 구현하여 이의 가능성을 실험하였다. 인공신경망은 일반적으로 선형적이고 함수적인 관계성에 대한 문제뿐만 아니라 다양한 비선형적이며 비함수적인 관계성에 대해서도 뛰어난 패턴인식 능력을 가지고 있다. 이러한 인공신경망의 장점을 사용하여 수화영상으로부터 수화 기호를 인식하여 번역하는 시스템을 위한 프로토타입시스템을 설계하고 구현하여 그 결과를 분석하였다.

Ⅱ. Proposed System

2.1 Design of Artificial Neural Network

지금까지 인공신경망은 과학자 및 공학자에게 큰 관심을 받아왔고, 최고의 계산도구중 하나로 평가 받고 있다. 이와 같이 기대하는 큰 이유 가운데 하나는 인공신경망이 사람처럼 학습할 수 있는 능력을 모방할 수 있기 때문이다. 인공신경망은 완벽하지 않은 학습 정보가 제시되었을 때에도 이를 통하여 학습된 인공신경망이 새로운 환경에 대해 어느 정도 결정하고 결론을 유추할 수 있다. 이처럼 인공신경망이 새로운 상황을 받아들여 처리할 수 있다는 것은 사람의 창의적 처리를 초보적 수준에서 모방하고 있다고 평가할 수 있는 부분이다. 이 인공신경망은 비수치적 계산뿐만 아니라 많은 수치적 계산에 대한 매우 좋은 통계적 도구로 사용 되고 있다. 특히 인공신경망은 다양한 비선형 시스템을 시뮬레이션하기 위한 강력한 도구로 알려져 있고, 많은 영역에서 많은 복잡한 문제에 적용되어 왔다.[3]

인공신경망은 복잡한 다차원적 문제에 대한 데이터 모델링 및

패턴인식 능력을 통하여 사람의 학습과 일반화 능력을 시뮬레이션 한다. 인공신경망이 기존의 일반적 데이터처리 기법과 다른 특징은 특별한 수학적 함수 관계 없이 입력과 출력 사이의 관계성을 일반화 할 수 있다는 것이다. 수화번역은 수화 사용자의 수화 표현에 있어서 다양성을 가지고 있으며, 기본적으로 언어간 번역이 가지는 비선형성으로 인하여 기존의 데이터처리 기법을 사용한 수화번역시스템은 많은 어려움이 존재한다. 따라서 본 논문에서는 인공신경망을 사용하여 다양성과 비선형성의 특징을 갖는 수화번역시스템을 구현하는데 적합한 도구로 평가하고 이를 연구하게 되었다.

새로운 인공신경망이 계속적으로 개발되면서 많은 유형이 존재하지만, 모든 인공신경망은 전이함수, 학습 규칙, 연결 형식에 의하여 종류가 분류될 수 있다. 가장 일반적으로 사용되는 인공신경망의 종류는 오차역전파로 훈련된 전방향 전파 인공신경망이다. 그림 1에 보인 것과 같이 전방향 전파 인공신경망은 입력층, 하나 이상의 은닉층 그리고 출력층으로 구성된다.

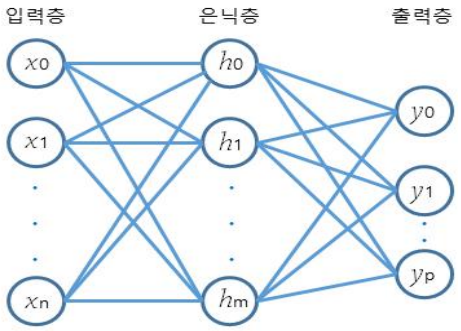


Fig. 1. Structure of artificial neural network

입력층은 외부 소스로부터 데이터를 공급 받으며, 은닉층은 입력층에서 입력된 데이터를 연결된 링크의 가중치에 근거하여 다음 계층에 전달한다. 최종 결과 값은 출력층에 생성된다. 각 계층의 노드는 다음 계층에 대한 출력을 계산하기 위해 이전 계층으로부터 입력을 받아 합계를 계산하고 시그모이드 전이함수에 의하여 출력을 식(1), (2)와 같이 계산한다.[4]

$$y_j = b + \sum w_{ij} x_i \tag{1}$$

$$f(y_j) = \frac{1}{1+\exp(-\alpha y_j)} \tag{2}$$

식 (1)에서, w_{ij}는 이전 계층에 있는 노드 I와 현재 계층에 있는 노드 j 사이의 링크에 대한 가중치이고, x_i는 이전 계층으로부터의 출력 값이고 b는 편향 값이다. 식 (2)는 전이함수로 사용된 시그모이드함수로서 $f(y_j)$는 출력 값으로 다음 계층에 전달되는 값이고, α는 시그모이드 함수의 형태와 관계된 파라미터이다. 이 시그모이드 함수의 장점은 과도한 감쇠 없이 작은 신호를 통과시킬 수 있고, 포화 없이 큰 신호를 수용할 수 있다는 것이다. 시그모이드 함수의 비선형성은 α가 증가함에 따라 강화된다.

인공신경망에서 은닉층의 개수와 은닉층에 포함되는 노드의 개수는 인공신경망의 기능과 성능에 영향을 준다. 최적의 은닉층의 개수와 노드의 개수에 대한 선택은 일반적으로 시행착오의 방법으로 결정되어 진다. 인공신경망에 대한 은닉층 노드의 개수에 대한 상한선에 대해 식 (3)을 참고하여 결정할 수 있다.

$$N_{hidden} = \frac{N_{trn}}{[R + (N_{inp} + N_{out})]} \tag{3}$$

식 (3)에서, N_{hidden}은 은닉층에 있는 노드의 개수이고, N_{trn}은 훈련 데이타의 개수, R은 5에서 10까지 범위 안에 있는 상수 값, N_{inp}는 입력층 노드의 개수 그리고 N_{out}는 출력층 노드의 개수이다.

또한 훈련 데이터의 크기 및 훈련의 양은 인공신경망의 사용에 있어서 매우 중요하다. 덜 훈련된 인공신경망은 큰 오류를 가지게 되며, 너무 과하게 훈련된 인공신경망은 패턴인식 및 일반화에 대한 능력을 오히려 떨어뜨리는 현상을 만들기 때문이다. 식 (4)는 입력층 노드의 개수, 은닉층 노드의 개수, 출력층 노드의 개수와 관련하여 훈련 데이터의 양을 합리적으로 예측하는데 사용된다.

$$n_s = \beta(n_h(n_i + 1) + n_o(n_h + 1)) \tag{4}$$

식 (4)에서, n_h은 은닉층 노드의 개수, n_i은 입력층 노드의 개수, n_o는 출력층 노드의 개수, 그리고 n_s는 훈련 데이터 양의 개수이다. 상수 β는 과적합의 정도에 관련된 파라미터이다.[5-6]

본 논문에서는 수화번역시스템에 대한 프로토타입 시스템으로 숫자 0부터 9까지의 수화에 대해서만 실험하였다. 실험에 사용된 수화 영상은 40x40(높이x너비) 크기에 회색톤 영상을 사용하였다. 숫자 수화영상을 인식하여 번역하기 위해 설계된 인공 신경망의 구조는 그림 2와 같다.

그림 2에서 입력층은 수화 영상 40x40을 1차원 배열로 받아들이기 위해 1600개 입력 노드로 구성되어 있다. 은닉층은 1개 층으로 구성하였고 은닉층 노드의 개수는 10개 노드와, 30개 노드로 구성하여 학습 및 평가를 비교 분석하였다. 출력층 노드는 10개 출력층 노드로 구성하여 인식된 수화 숫자에 대한 출력 값을 평가하도록 하였다.

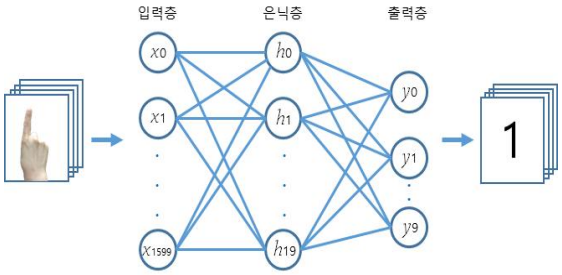


Fig. 2. Artificial neural network concept of sign language video recognition system

2.2 Implementation and Learning of Artificial Neural Network

인공신경망은 입력층 1600개 노드, 하나의 은닉층과 10개 및 30개의 은닉층 노드, 그리고 출력층 10개 노드로 구성하였

다. 인공신경망의 구현은 파이썬(Python) 버전 3.6.5와 OpenCV 3.0 버전을 사용하였다.[9][10][12][13] 구현된 인공신경망에 대한 파이썬 코드의 일부 내용은 그림 3과 같다.

```
# x 입력 수화영상 40x40
# w1, w2 가중치
# b1, b2 편향 값
# y 출력 값
h1 = np.dot(x, w1) + b1
h11 = sigmoid(h1)
h2 = np.dot(h11, w2) + b2
y = softmax(h2)
```

Fig. 3. Part implementation code of artificial neural network

훈련데이터는 그림 4와 같은 숫자 수화영상 40x40 회색톤 영상으로 원본 영상 50개의 수화 영상을 반복시켜 500개 훈련 데이터로 사용하였다.

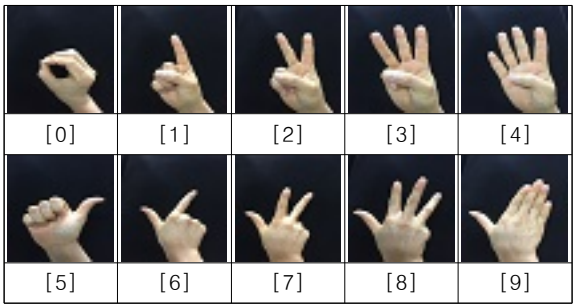


Fig. 4. Example of numeric sign language image

훈련 데이터는 영상 데이터와 함께 각 영상의 숫자가 레이블된 값을 쌍으로 제공하여 인공신경망을 훈련하는데 사용하였다. 인공신경망의 학습은 인공신경망의 각 노드 사이의 연결에 대한 가중치 값을 조정하여 훈련 데이터를 기반으로 하였을 때 오차가 최소화가 되도록 하는 과정이다.

일반적으로 인공신경망의 학습과정에서는 지역적 최소 오차에 빠지게 되는 위험성이 존재한다. 본 논문에서는 이를 위하여 가중치의 초기 값에 난수를 사용하고 학습률에 대해 몇 개의 적정한 값을 사용하여 훈련하였다. 본 논문에서는 학습률을 각각 0.1과 0.05의 값 등 몇 개의 값으로 변경하여 적용하고 비교하였다.

본 논문에서 사용된 훈련 데이터는 데이터 관리의 편의성과 향후 발전된 수화 번역시스템의 확장성을 고려하여 동영상 파일(*.avi)로 압축하여 사용하였다. 훈련 데이터의 수집과정은 숫자 수화에 대한 정지영상을 준비된 배경에서 칼라 영상으로 촬영하여 수집하였다. 수화 동작의 일반성을 확보하기 위하여 여러 사용자들로부터 최대한 다양한 수화 동작에 대한 정지영상을 수집하도록 하였다.

훈련에 사용된 수화영상은 칼라영상에서 256단계의 회색톤영상으로 변환하여 사용하였다. 인공신경망의 학습과정은 준비된 훈련 데이터를 사용하여 그림 5와 같은 순방향전파와 역방향전파를 훈련 횟수만큼 반복 수행하여 진행하였다. 이와 같은 학습과정을 통하여 초기 인공신경망에 부여된 가중치의 값을 미세하게 조정하여 준비된 훈련 데이터의 정답이 일치해 가도록 학습을 진행하였다.

```
# feed forward
h1 = np.dot(x, w1) + b1
h11 = sigmoid(h1)
h2 = np.dot(h11, w2) + b2
y = softmax(h2)

# backward propagation
dy = (y - t) / data_num
grads['w2'] = np.dot(h1.T, dy)
grads['b2'] = np.sum(dy, axis=0)

dh1 = np.dot(dy, w2.T)
dh11 = sigmoid_grad(h1) * dh1
grads['w1'] = np.dot(x.T, dh11)
grads['b1'] = np.sum(dh11, axis=0)
```

Fig. 5. Part implementation code of feed forward and backward propagation of artificial neural network

순방향전파(feed forward)는 입력층 노드로부터 시작하여 출력층 노드 방향으로 진행하는 계산과정을 이르는 과정을 나타낸다. 역방향전파(backward propagation)는 출력층 노드로부터 시작하여 입력층 노드 방향으로 진행하며 잘못된 인공신경망의 평가 결과에 대하여 오차 미분 값을 사용하여 가중치를 미세하게 조정하는 과정을 나타낸다. 이와 같이 순방향전파와 역방향전파를 반복하여 오차 값이 0에 가까워지기 위한 조정 작업으로 인공신경망의 훈련 또는 학습 과정이 진행되었다.

III. Analysis on Proposed System

그림 6은 인공신경망의 구성을 입력층 노드 개수 1600개, 은닉층 노드 개수 10개, 출력층 노드 개수 10개로 구성(인공신경망 1600-10-10)하고 훈련 진행과정에서의 인공신경망의 훈련 데이터 및 평가 데이터에 대한 정확도 결과를 나타내는 그래프이다. 인공신경망 훈련에 사용된 훈련 데이터는 수화 이미지와 수화 라벨이 세트로 구성되어 있으며 본 논문의 실험을 위하여 500 세트를 사용하였다. 인공신경망을 평가하기 위해 사용된 평가 데이터는 수화 이미지와 수화 라벨이 세트로 구성되어 있으며 실험을 위하여 200 세트를 사용하였다. 훈련 과정은 훈련 데이터 500 세트로부터 20세트의 데이터를 임의로 선택하여 3000번의 훈련을 반복 수행하였다. 그림 6에 표현된 각 epoche는 25번의 반복 훈련으로 구성되어 있다.

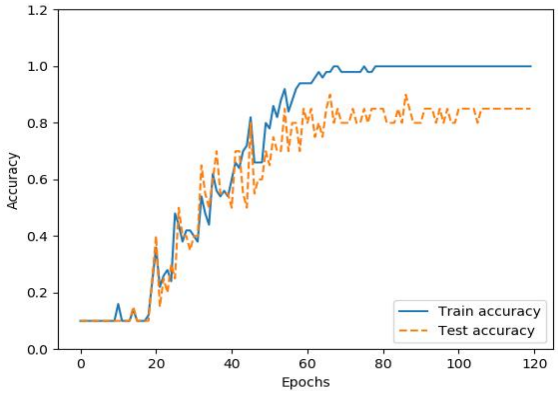


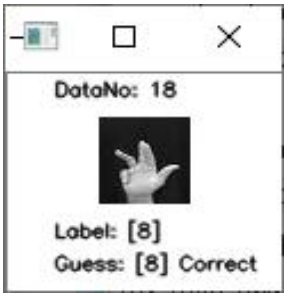
Fig. 6. Learning and evaluation accuracy of artificial neural network 1600-10-10

그림 6에서 실선으로 표현된 Train accuracy는 훈련 데이터에 대한 정확도를 나타내고 있다. Train accuracy 즉, 훈련데이터에 대해서는 약 50 epochs까지는 정확도가 점진적으로 증가하고 있는 모습을 나타내고 있으며, 약 60 epochs 이후부터는 100%의 정확도를 보여주고 있다. 그림 6에서 점선으로 표현된 Test accuracy는 평가 데이터에 대한 정확도를 나타내고 있다. 평가 데이터는 훈련 데이터에 포함되어 있지 않은 데이터로써 인공신경망의 일반화에 대한 능력을 나타내고 있다. Test accuracy 즉, 평가 데이터에 대해서는 60 epochs까지 점진적으로 정확도가 증가하는 모습을 나타내고 있으며, 65 epochs 이후부터는 약 85%의 정확도를 유지하고 있음을 보여 주고 있다.

그림 7에서는 학습된 인공신경망 1600-10-10에 의해서 올바로 인식된 수화의 예를 보여 주고 있고, 그림 8에서는 잘못 인식된 수화의 예를 보여주고 있다.

(a) 수화 [0]이미지

(b) 수화 [3]이미지



(c) 수화 [8]이미지

(d) 수화 [6]이미지

Fig. 7. Correct recognition by artificial neural network 1600-10-10

(a) 수화 [6]이미지

(b) 수화 [7]이미지

Fig. 8. Incorrect recognition by artificial neural network 1600-10-10

그림 8에서 인공신경망 1600-10-10이 잘못 인식하고 있는 (a)는 수화 이미지가 숫자 6을 나타내고 있으나, 인공신경망은 숫자 7로, (b)는 수화 이미지가 숫자 7을 나타내고 있으나 인공신경망은 숫자 8로 잘못 인식하고 있는 예를 보여주고 있다.

그림 9는 인공신경망의 구성을 입력층 노드 개수 1600개, 은닉층 노드 개수 30개, 출력층 노드 개수 10개로 변경하여 구성(인공신경망 1600-30-10) 하고 훈련 진행과정에서의 변경된 인공신경망의 훈련 데이터 및 평가 데이터에 대한 정확도 결과를 나타내는 그래프이다. 즉, 앞에서 구성한 인공신경망과 다르게 은닉층의 노드의 개수를 30개로 확대하여 인공신경망을 구성하고 달라진 학습 정확도의 내용을 그래프로 보여주고 있다. 실험에 사용된 훈련 데이터와 평가 데이터는 인공신경망 1600-10-10과 동일한 훈련 데이터 500세트와 평가 데이터 200세트를 사용하였다.

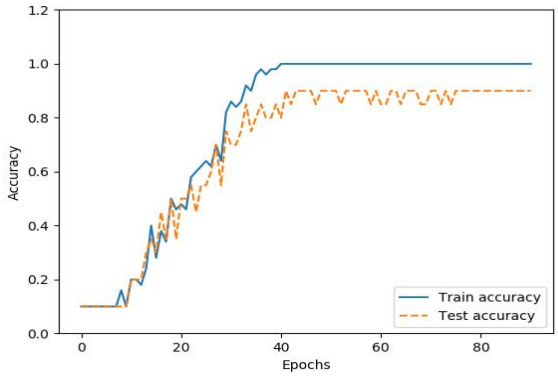


Fig. 9. Learning and evaluation accuracy of artificial neural network 1600-30-10

그림 9에서 실선으로 표현된 Train accuracy는 훈련 데이터에 대한 정확도를 나타내고 있으며 그림 6의 인공신경망 1600-10-10에 대한 것과 다른 점은 학습의 정확도에 대한 상승 속도가 더욱더 빨라졌다는 것이다. Train accuracy 즉, 훈련데이터에 대해서는 약 40 epochs까지는 정확도가 점진적으로 증가하고 있으며, 인공신경망 1600-10-10 보다 빨라진 속도로 증가하고 있는 모습을 나타내고 있다. 훈련 반복 횟수가 늘어남에 따라 훈련 데이터에 대한 정확도가 점진적으로 상승

하다가 약 50 epochs 이후부터는 100%의 안정된 정확도를 보여주고 있다. 은닉층의 노드의 개수를 확대함으로써 정확도에 대한 상승 속도가 더욱 빨라졌고, 안정된 100%의 정확도에 이르는 시점도 더욱 빨라진 것을 보여 줌으로써 은닉층의 노드의 개수가 학습의 정확도 및 안정성에 영향을 주고 있는 것으로 분석할 수 있다.

그림 9에서 점선으로 표현된 Test accuracy는 평가 데이터에 대한 정확도를 나타내고 있으며, 그림 6에서 보인 인공신경망 1600-10-10의 결과 보다 높은 정확성을 나타내고 있다. 즉 인공신경망의 일반화에 대한 능력이 은닉층 확대와 함께 더욱 더 향상되었음을 나타내고 있다. 그림 9에서 Test accuracy 즉, 평가 데이터에 대해서는 50 epochs까지 점진적으로 정확도가 증가하는 모습을 나타내고 있으며, 70 epochs 이후부터는 약 90%의 정확도를 유지하고 있음을 보여 주고 있다.

그림 10에서는 학습된 인공신경망 1600-30-10에 의해서 올바로 인식된 수화의 예를 보여주고 있고 그림 11에서는 잘못 인식된 수화의 예를 보여주고 있다.

(a) 수화 [0]이미지

(b) 수화 [3]이미지

(c) 수화 [8]이미지

(d) 수화 [6]이미지

(e) 수화 [4]이미지

Fig. 10. Correct recognition by artificial neural network 1600-30-10

(a) 수화 [6]이미지

Fig. 11. Incorrect recognition by artificial neural network 1600-30-10

그림 11에서 인공신경망 1600-30-10이 잘못 인식하고 있는 (a)은 수화 이미지는 숫자 6을 나타내고 있으나, 인공신경망 은 숫자 7로 잘못 인식하고 있는 예를 보여주고 있다. 그러나 그림 10의 (e)는 앞의 인공신경망 1600-10-10에서 잘못 인식하였던 수화 이미지를 인공신경망 1600-30-10에서는 정확하게 인식하고 있어 인식 정확도가 향상되었음을 보여주고 있다.

Ⅳ. Conclusion

과학기술의 발전과 함께 사람들 사이의 소통이 다양한 형태로 발전하였고 더욱더 편리하고 효율적인 환경으로 변화해 가고 있다. 그러나 장애인을 위한 소통 환경은 아직도 불편하고 비효율적인 부분이 많이 존재한다. 청각장애인이 사용하고 있는 수화의 경우에 장애인과의 대화 수단으로는 잘 이용되고 있으나 정상인과의 대화에는 많은 어려움이 존재한다. 이와 같은 어려움은 청각장애인들의 일상의 생활뿐만 아니라 의료 환경과 같이 중대한 상황에서 심각한 문제를 발생시키고 있어 이를 해소하기 위한 방안이 필요한 상황이다. 이와 관련하여 수화번역시스템에 대한 연구가 진행되고 있으나 아직 부족한 점이 많이 있어 여전히 어려운 상황에 놓여 있어, 본 논문에서는 이와 같은 문제점을 보완할 수 있는 수화번역시스템에 대해 연구하였다.

지금까지 연구되어 온 수화번역시스템은 크게 영상인식기반 수화번역시스템과 특별한 장치를 사용한 형상입력시스템기반 수화번역시스템이 주를 이루고 있다. 그러나 이와 같은 기존 수화번역시스템은 수화 사용자의 다양한 수화 표현을 인식하지 못하고, 현실적 사용에 있어서 불편함을 요구하는 특별한 장치를 필요로 함으로써 실용적이 못한 문제점이 존재한다. 따라서 본 논문에서는 이러한 단점을 보완한 수화번역시스템을 개발하기 위하여 수화 사용자의 다양한 수화 표현을 인식할 수 있는 인공신경망기술과 특별한 장치의 요구 없이 일상에서 쉽게 접근할 수 있는 스마트폰 영상 및 카메라영상을 사용한 수화번역시스템을 연구하게 되었다.

본 논문에서는 인공신경망을 통한 수화영상인식시스템의 프로토타입을 개발하고 이의 가능성을 실험하기 위하여 숫자 수화영상으로 제한하여 개발하였다. 숫자 수화 이미지는 40x40 크기의 회색톤 영상으로 실험의 편리성과 향후 확장성을 위하여 비디오영상파일로 압축하여 사용하였다. 본 논문에서 설계된 인공신경망은 입력층 1600개 노드, 출력층 10개 노드로 구성하고, 은닉층은 10개 노드와 30개 노드로 각각 구성하여 그 성능을 비교 검토하였다. 은닉층 노드를 10개로 구성한 인공신경망 1600-10-10과 은닉층 노드를 30개로 구성한 인공신경망 1600-30-10에 대한 숫자 수화영상인식 성능은 훈련데이터와 평가데이터에서 각각 100%, 85%와 100%, 90%의 정확도를 나타내었음을 확인하였다. 이로써 인공신경망을 통한 수화영상번역시스템이 다양성과 현실적 실용성을 갖춘 수화번역시스템의 방법으로 적합하다고 분석할 수 있었다.

본 논문에서 사용한 훈련데이터 500 세트와 평가데이터 200 세트의 데이터양은 앞의 인공신경망의 적정한 훈련데이터양에 대한 기준에 비추어 볼 때 데이터양이 많이 부족한 것으로 분석된다. 훈련 데이터의 양과 훈련 반복 크기는 인공신경망의 정확도와 일반화 능력에 많은 영향을 주는 요인 가운데 하나이므로 이의 실용화를 위하여 적정한 수준의 훈련 데이터의 양을 확보하는 것이 필요하다. 또한 학습된 인공신경망 수화번역시스템의 다양성을 더욱더 높이기 위하여 다양한 수화사용자의 수화영상을 훈련 데이터와 평가 데이터로 준비하는 것이 요구된다.

앞으로 수화영상번역시스템의 정확도와 다양성을 높이기 위해서는 인공신경망의 설계 및 훈련방법에 대한 연구, 적정한 훈련데이터의 확보 방안 연구, 수화 이미지에 대한 다양한 영상처리기법의 사용에 대한 연구를 추가적으로 진행하고자 한다.

REFERENCES

[1] Ryu M. W., “Design and Implementation of Bi-directional Sign Language System for Emergency Medical Situation”, Pusan National University M.S, Feb. 2015.

[2] Korea Employment Agency for the Disabled, “2017 Disability statistics”, Survey statistics 2017-01, Nov. 2017.

[3] Vijaykumar Sutariya, Anastasia Groshev, Prabodh Sadana, Deepak Bhatia, Yashwant Pathak, “Artificial Neural Network in Drug Delivery and Pharmaceutical Research”, The Open Bioinformatics Journal, Jul. 2013.

[4] S. Ferrari and R. F. Stengel, "Smooth function approximation using neural networks", IEEE Trans Neural Network, Vol. 16, pp. 24-38, Jan. 2005.

[5] M. N. Jadid and D. R. Fairbairn, "Neural-network applications in predicting moment-curvature parameters from experimental data", Engineering applications of artificial intelligence, Vol. 9, pp. 309-319, Jun. 1996.

[6] H. M. Carpenter WC, "Understanding neural network approximations and polynomial approximations helps neural network performance," AI Expert. Vol. 2, pp. 31-33, Mar. 1995.

[7] Lee Seung Seok, Heo Jeong Hyun, No Seung Woo, Yoon Hyeon Jin, Park So Hyun, Kim Chan Kyu, “Sign language translation system based on deep learning for speech disorders”, Proceedings of Symposium of the Korean Institute of communications and Information Sciences, Jun. 2018.

[8] Kyung-Hyuk Kwon, Yo-Seop Woo, Hong-Ki Min, “Design and Implementation of a Korean Text to Sign Language Translation System”, The transactions of the Korea Information Processing Society, Mar. 2000.

[9] Saito G.. “Deep running starting from the bottom”, O’REILLY, Hanbit Media Inc, Jan. 2018.

[10] Hope, Tom Resheff, YehezkelS. Lieder, Itay, “Learning TensorFlow”, Hanbit Media, May. 2018.

[11] F. Liu, M. J. Er, “A novel efficient learning algorithm for self-generating fuzzy neural network with applications”, International Journal of Neural Systems, Vol. 22, pp21-35, Feb. 2012.

[12] D. K. Kim, “C++ API OpenCV Programming”, The Publish Company of KAME, May 2015.

[13] AmmarAnuar, KhairulMuzzammilSaipullah, NurulAtiqah Ismail, and Soo Yew Guan, “OpenCV Based Real-Time Video Processing Using Android Smartphone”, International Journal of Computer Technology and Electronics Engineering, Vol. 1, No. 3, pp.58-63, 2011.

[14] Lee Hyun-Suk, Kim Seung-Pil, Chung Wan-Young, “Development of Sign Language Translation System using Motion Recognition of Kinect”, Journal of the institute of signal processing and systems, Vol. 14, No. 4, pp.235-242, Oct. 2013.

Authors

Hyung-Hoon Kim received the B.S. degrees in Computer Science from Chonnam National University, Korea, in 1986. He received the M.S. degrees in Computer Science from KAIST, Korea, in 1988. He received the Ph.D. degrees in Computer Science from Hanyang University, Korea, in 2007. Dr. Kim joined the faculty of Kwangju Womens University, Gwangju, Korea, in 1994. He is currently a Professor in the Department of Cosmetic Science at Kwangju Womens University. He is interested in web programming, medical information system, information security, and AI.

Jeong-Ran Cho received the B.S., M.S. and Ph.D. degrees in Computer Science from Chonnam National University, Korea, in 1987, 1989 and 1999, respectively Dr. Cho joined the faculty of Kwangju Womens University, Gwangju, Korea, in 1994. She is currently a Professor in the Department of Health Administration at Kwangju Womens University. She is interested in database, parallel computing, internet and mobile computing, and multimedia contents service.