

립모션 센서 기반 증강현실 인지재활 훈련시스템을 위한 합성곱신경망 손동작 인식

송근산^{1‡} · 이현주^{2‡} · 태기식^{1,3*}

¹건양대학교 대학원 의료공학과, ²건양대학교 물리치료학과
³건양대학교 의공학부

Hand Gesture Recognition with Convolution Neural Networks for Augmented Reality Cognitive Rehabilitation System Based on Leap Motion Controller

Keun San Song^{1‡}, Hyun Ju Lee^{2‡} and Ki Sik Tae^{1,3*}

¹Department of Biomedical Engineering & Science, Graduate School of Konyang University
²Department of Physical Therapy, Konyang University, Daejeon, Republic of Korea
³Department of Biomedical Engineering, Konyang University, Daejeon, Republic of Korea
(Manuscript received 10 August 2021 ; revised 18 August 2021 ; accepted 24 August 2021)

186

Abstract: In this paper, we evaluated prediction accuracy of Euler angle spectrograph classification method using a convolutional neural networks (CNN) for hand gesture recognition in augmented reality (AR) cognitive rehabilitation system based on Leap Motion Controller (LMC). Hand gesture recognition methods using a conventional support vector machine (SVM) show 91.3% accuracy in multiple motions. In this paper, five hand gestures (“Promise”, “Bunny”, “Close”, “Victory”, and “Thumb”) are selected and measured 100 times for testing the utility of spectral classification techniques. Validation results for the five hand gestures were able to be correctly predicted 100% of the time, indicating superior recognition accuracy than those of conventional SVM methods. The hand motion recognition using CNN meant to be applied more useful to AR cognitive rehabilitation training systems based on LMC than sign language recognition using SVM.

Key words: Euler angle spectrograph, Convolutional neural network (CNN), Hand gesture recognition, Augmented reality (AR), Cognitive rehabilitation system, Leap motion controller (LMC)

1. 서 론

인지재활 분야에서의 다양한 연구에서 뇌는 반복적이고 집중적이며 작업 지향적 훈련을 통해 회복이 가능하다는 것이 밝혀졌다[1,2]. 하지만 뇌 손상 환자는 전통적으로 재활에 필요한 운동이 환자에게 특정행동을 반복적으로 유도하여 지루하게 느낀다고 보고하였다[3]. 환자에게 동기를 부여하는 것은

재활 성공의 중요한 요소로 작용하여 재활 결과에 영향을 끼치는 경우가 많다[4]. 재활의 효과를 최대화하기 위한 주요 문제는 환자가 재활 프로그램에 적극적으로 참여하도록 장려하는 방법을 찾는 것이다[5]. 인지 장애에 대한 증거 기반 심리치료에 대한 접근은 저렴하고 시장 확장이 가능해야 하며 치료 개입이 분명하게 필요하다[6]. 이러한 인지 치료 방법들이 시장에 진출하여 보급 시 겪는 가장 큰 어려움은 재활 치료 전문가의 지도를 필연적으로 필요로 하거나 높은 치료 비용을 요구한다는 점이다[7].

오늘날 4차 산업혁명 유망기술에 대한 관심이 높아지고 그와 관련된 연구가 활발해짐에 따라 메타버스(Metaverse) 기반의 디지털 전산 콘텐츠를 이용한 인지재활의 중재는 이러

[‡]Contributed equally to this work

*Corresponding Author : Ki-Sik Tae

Department of Biomedical Engineering, Konyang University,
158 Gwanjeodong-ro, Seogu, Daejeon, 35365, Republic of Korea
Tel: *** - **** - **** +82-42-600-8518
E-mail: tae@konyang.ac.kr

한 비용이나 전문가의 부재로 겪는 어려움에 대한 해결책을 제공 할 수 있다. 그러나 현재 개발되어지고 있는 메타버스 기반 재활 콘텐츠의 입력 메커니즘은 터치스크린, 조이스틱 등의 기성 컨트롤러를 사용하거나 그 사용법이 실제 일상생활 동작과는 거리가 멀고 직관적이지 못하다. 이를 해결하기 위해 입력 메커니즘(input mechanism)의 개선을 통해 콘텐츠 참여를 위한 동기부여가 필요하다. 또한 각각의 재활목적의 전산 콘텐츠를 수행하려고 할 때마다 각 제품만의 컨트롤러 사용법을 학습해야 한다는 것은 환자와 치료사 양쪽 모두에게 부담으로 작용할 수밖에 없고[8] 인지재활을 주로 이용하는 고령자 환자의 경우 이러한 터치스크린과 조이스틱 등의 입력 메커니즘의 동작방식 자체가 익숙하지 않아 금방 흥미를 잃게 될 수 있다. 콘텐츠의 중재방법을 임상에서 적용하는데 필요한 요건은 접근성, 간편성이고 대상자들에게 흥미 있는 콘텐츠로 재활에 적극적으로 참여할 수 있도록 개선하는 과정은 환자가 콘텐츠를 통한 재활치료에 있어서 동기부여에 큰 도움을 줄 수 있다[9].

기존의 가상현실 콘텐츠에서의 손동작 인식은 영상처리를 통해 손끝 추적을 하는 것을 중심으로 발달되어 깊이 카메라(depth camera) 이용해 실시간으로 손 모양을 추적하여 3D 콘텐츠 내부에 사용자의 손을 객체로 구현하는 기술이 사용되어져 왔다. 최근 립모션(leap motion) 장비를 이용하여 이미지나 객체를 추출해내는 것이 매우 간단해졌기 때문에 단순히 손 모양을 구현하는 것을 넘어 손동작이 의미하는 바를 추적하는 알고리즘이 인공지능을 통한 손동작 인식의 개념으로써 자리 잡고 관련 연구들이 제안되고 있다. 현재 립모션을 사용했을 때 가장 많이 이용되는 딥 러닝 방식은 서포트 벡터 머신(support vector machine: SVM)이며[10-13] SVM은 선형예측 도구로써 분류의 기준선을 결정함을 통해 현재 손동작이 미리 사전에 라벨링된 동작인지 아닌지 여부를 검토해서 분류해내는 이진화 알고리즘이다. 그러나 이는 어디까지나 미리 사전에 정의된 라벨만을 학습시키고 분류할 수 있다는 명확한 한계점을 지니고 있으며 새로운 동작을 추가하려고 할 때마다 다시 데이터를 모아 학습시켜야한다는 부담감을 연구자로 하여금 가져올 수 있다. 이는 곧 사용성과 정확도와 직결되는 문제이기에 사용자에게까지 그 부담이 느껴지는 결과를 초래할 수 있다. 반면 Côté-Allard 등[14]은 Myo사의 암밴드를 사용하여 얻은 근전도 데이터로부터 스펙트로그램(spectrogram)으로 변환하여 합성곱신경망(convolutional neural network: CNN)을 통해 이미지 분류하여 주파수 성분을 컨볼루션하여 일반적으로 인지 가능한 영역보다 많은 영역을 스캔할 수 있으며 근전도를 대상으로 분류를 실행하였을 때 보다 적은 편차를 갖는 각도 값에서 무리 없이 손동작 분류 및 식별이 가능하다고 하였다. 이러한 원리를 이용하여 립모션을 통해 추출한 각도로부터 스펙트로그램을 추출

하고 이용할 수 있다. 서포트 벡터 머신을 이용한 이진화 선형예측 알고리즘은 수화 인식, 숫자의 손가락 셈 인식 등 주어진 동작을 인식하는데 그쳤기 때문에 메타버스 콘텐츠 등에 직접적인 활용은 연구된 바가 상대적으로 적다. 때문에 아직 메타버스 콘텐츠에서의 손동작 인식은 각도 상수를 비교하여 특정 상수 이상일 경우 굽힘으로 감지하고 특정상수 이하일 경우 펴므로 감지하는 방식인 상수 대입법(local transform Euler angle)을 상용 어플리케이션에 이용하고 있으며 SVM은 그 특성상 라벨링 손동작을 분류하는 수화인식 등의 콘텐츠에서 사용되어야 하며 실제 상호인식이 필요한 메타버스 콘텐츠에서 사용하기에는 제한이 있다.

따라서 본 논문에서는 스펙트로그램 변환과 합성곱신경망을 이용한 손동작 인식 예측 정확도를 측정하여 기존의 서포트 벡터 머신 방법과의 비교를 통해 립모션 기반 증강현실 인지재활 훈련시스템의 적용 가능성을 검증하고자 하였다.

II. 연구 방법

1. 모션 기반 증강현실 인지재활 훈련시스템

본 연구에서 개발한 립모션 카메라를 이용한 손동작 인식 재활콘텐츠와 구글 음성-문자 변환 인터페이스(Google speech to text API)를 이용한 음성인식 인지재활콘텐츠는 손의 각도와 손가락 굽힘에 관한 손동작 인식, 한국어에 대한 음성인식을 통한 재활콘텐츠 구현을 위한 시스템이다. 영상촬영을 위한 적외선 카메라와 마이크가 포함된 입력부, 신호처리부, 구현 및 출력부로 나뉜다. 그림 1은 본 연구에서의 시스템 설계를 위한 블록도를 나타낸다. 립모션 카메라의 기본 어플리케이션은 모델 구현만을 제공할 뿐 손동작 인식을 위해서는 별도의 라벨링 및 정의함수가 추가적으로 필요하다. 예를 들어 물건을 집기 위해서 손을 오므리거나 피는 동작을 3D 모델 구현 한다고 했을 때, 집는 모델을 렌더링하는 것은 립모션을 통해 관절의 위치를 찾아 매핑하는 것으로 해결 될 수

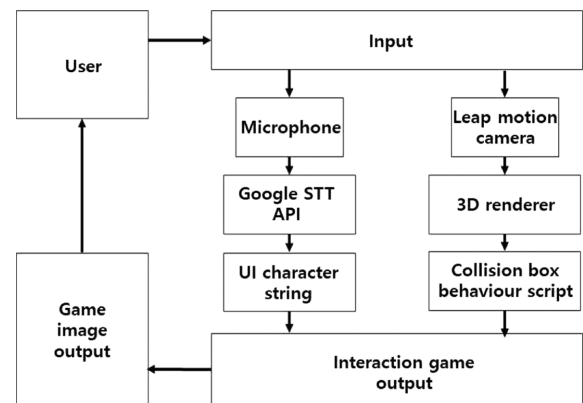


그림 1. 전체 시스템 블록도
Fig. 1. Block diagram of system

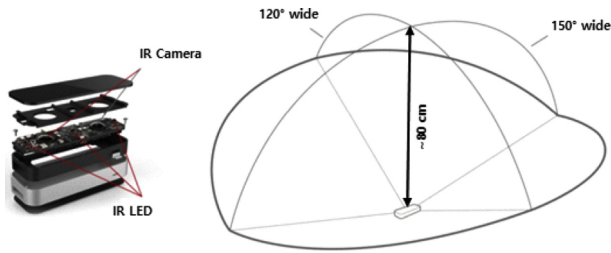


그림 2. 립모션 센서와 동작범위
Fig. 2. Leap motion controller and operating range

있지만 콘텐츠 내부에 존재하는 오브젝트와 상호작용을 하려면 충돌체크와 충돌연산에 동반되는 동작을 구현해서 지정해 주어야한다. 이 때문에 손바닥 부분과 각 손 끝 관절은 충돌의 매개체가 되는 물리량을 갖는 충돌체를 추가해주게 되는데 이를 충돌 판정 박스라고 한다. 충돌 판정 박스를 추가해줌으로써 상호작용이 가능한 물체 혹은 가상현실의 구성 오브젝트가 가지고 있는 충돌 판정 박스와 충돌체크를 진행할 수 있도록 구성하였다.

2. 기구부 및 인지재활 콘텐츠 개발

본 연구에서 콘텐츠를 진행하기 위한 기구부 설계를 진행하였으며 기구부는 영상출력을 위한 디스플레이부, 립모션 센서(LM-010)를 이용해 이미지를 입력받는 영상 감지부, 외장을 구성하는 외부프레임 등으로 나뉜다. 사용된 립모션 카메라는 크기가 76 mm × 30 mm × 12.7 mm, 무게 45.3 g 이며 동작범위는 그림 2와 같다.

디스플레이부는 외부프레임의 디스플레이 거치대에 고정되며 영상 감지부는 외부프레임 하단에 위치하는 상자 형태의 구조(박스 프레임) 내부의 부품 고정부에 위치하여 좌우 수직 프레임으로부터 구성되는 작업스테이션 내에서의 손 움직임을 촬영하도록 구성된다(그림 3).

현행 인지재활 훈련 현황에 대한 사전조사결과 요양원에서 학습지로 실시하고 있는 평가, 한국치매평가검사 평가항목, 우리소프트의 뉴로월드의 재활콘텐츠의 평가 및 재활 항목은 목적을 공유하며 실시하고 있는 것을 확인할 수 있다. 분류

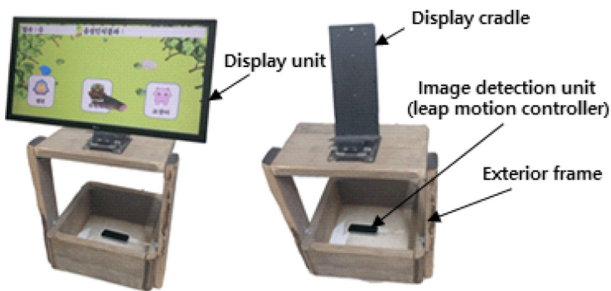


그림 3. 실제 기구부 사진
Fig. 3. Appearance of instrument part

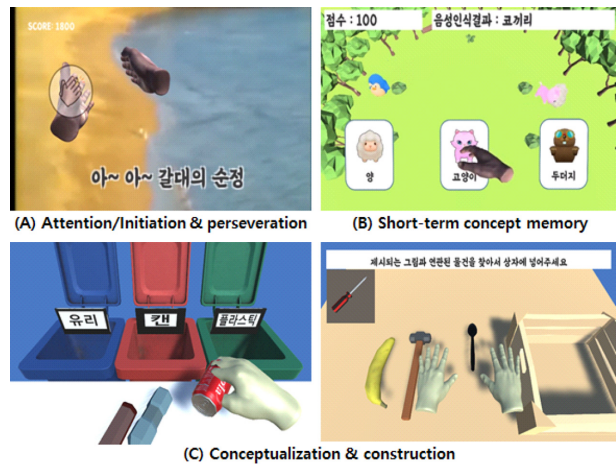


그림 4. 증강현실 기반 훈련 콘텐츠
Fig. 4. Training contents based on augmented reality (AR)

하기와 주의, 주의집중 훈련은 정보를 수동적으로 접수하고 이를 처리하는 용량 및 간섭 없이 주의를 집중시킬 수 있는 능력훈련이 목적이며 서열 짓기, 구성, 개념화, 시지각 훈련은 시공간적인 능력의 하나로 시지각 기능과 운동반응 요구 혹은 구체적 사상이나 사건을 추상적으로 처리하는 사고과정을 훈련을 목적으로 하고 있다. 암기하기, 기억, 기억력훈련은 언어적 회상, 언어적 재인을 훈련을 목적으로 하고 있다.

메타버스를 이용한 재활 콘텐츠의 설계는 위 항목을 훈련하기 위하여 메타버스 콘텐츠 안에서 가상 오브젝트와 사용자가 어떤 방식으로 상호작용을 할지를 구성하는 내용을 포함하여야 한다[15]. 본 연구에서 제작된 손 동작 인식 기반 증강현실 콘텐츠는 (1) 주의훈련 및 관리기능 훈련 콘텐츠 1종, (2) 단기 개념 기억 훈련 콘텐츠 1종, (3) 개념화 및 구성 훈련 콘텐츠 2종으로 총 4종의 콘텐츠가 제작되었다.

3. 손동작 인식 성능평가

(1) 상수대입법 정확도 평가

기존 상수대입법을 이용한 손동작 인식 방법에서 5가지 손동작(그림 5)에 대하여 미리 라벨링한 사용자의 손동작에 대해 증강현실 콘텐츠 내부에서 립모션이 인식한 손 동작 라벨과 실제 의도한 손동작과 비교하여 얼마나 정확하게 인식하였는지를 각 손동작별로 30회씩 반복하여 손동작 인식 정확도를 평가하였다.

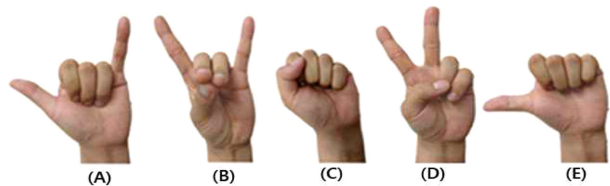


그림 5. 5가지 손동작 정의
Fig. 5. Definition of 5 hand gestures

본 연구에서 사용된 5가지 손동작 중 A는 “Promise” 동작으로 엄지와 소지 인식, 가장 변화가 뚜렷한 엄지와 이미 지상으로 가장 변화가 미미한 소지를 동시에 인식하기 위한 동작이다. B는 “Bunny” 동작으로 손가락과 손바닥 겹침의 경우 손가락의 인식 정확도를 확인하기 위한 동작이다. C는 “Close” 동작으로 손 모델 인식의 기본이 되는 동작이며 모든 손가락 굽힘인 상태를 인식하는지 확인하기 위한 동작이다. D는 “Victory” 동작으로 가장 많이 측정지표로 활용되는 동작 중의 하나로서 인식이 높은 검지와 중지의 상태를 올바르게 검출하는지 확인하기 위한 동작이다. E는 “Thumb” 동작으로 가장 변화가 뚜렷한 엄지만 펴 상태로 두고 나머지를 굽힘 상태로 두었을 때 인식하기 위한 동작이다.

(2) 모델인식 검출 속도 평가

손뺌 동작의 검출 후 비활성화된 손 모델이 복구하는데 걸리는 시간을 50회 측정하여 그 평균치를 측정하였다.

(3) 합성곱신경망 적용 및 평가

학습에 사용된 손동작 데이터는 5가지 손동작에 대하여 각 동작별로 100개의 데이터를 측정한 실험 1과 1,000개의 데이터를 측정한 실험 2로 나뉘며 이러한 데이터를 수집하기 위하여 시뮬레이션 프로그램을 구성하여 동작시켰다. 실험은 나누어서 진행되었으며 실험 1, 2 모두 500개의 개별 데이터를 훈련 세트(training set) 90%(450개)와 평가 세트(test set) 10%(50개)로 나누어 학습 및 테스트를 진행하였다.

기본적으로 가상현실(virtual reality: VR) 콘텐츠 내부의 최소 시간단위는 사용 플랫폼의 성능에 의해 좌우되며 0.04 초를 기본 최소 시간단위로 잡고 있다. 필요에 따라 그 이상을 성능을 내는 것도 가능하지만 시스템의 안정성을 떨어트릴 수 있어 0.04초마다 동작하는 기본 최소시간 단위로 고정하고 이를 고려하여 설계하도록 하였다. 지도학습을 위한 데이터 수집을 위해 데이터 수집버튼을 누른 후 3초 뒤 Unity VR 콘텐츠 내부에서 한가지 손동작을 행할 것을 디스플레이를 통해 지시하고 0.1초 마다 1초 동안 데이터가 수집되어 내부변수에 기록한 뒤 수집한 데이터를 csv 파일로 기록하였다. 파일을 저장한 다음 디스플레이상에서 데이터 수집이 끝났음을 알리고 다음 버튼이 눌릴 때까지 사용자가 대기하는 방식으로 데이터 수집이 이루어졌다.

이후 데이터를 python OpenCV 모듈을 이용하여 그림 6과 같이 스펙트로그램으로 변환 하고 CNN 모델은 미리 동일한 조건으로 분류 기법을 적용하고 학습시키기 위한 라벨링된 데이터셋을 구성하여 지도학습의 형태로 진행하였다.

기존 상수 대입법 및 SVM 기법과 비교하여 그 정확도의 향상을 확인하기 위하여 그림 7과 같은 모델을 설계하여 학습을 진행하도록 한다.

모델은 가장 널리 통용되는 CNN 모델인 mnist를 참조하여 디자인하였으며 3개의 변환 레이어(convolution layer)와 3개의 max pooling 레이어, 2개의 dense 레이어를 사용하며 각 레이어의 상세구조는 표 1과 같이 구성되어있다.

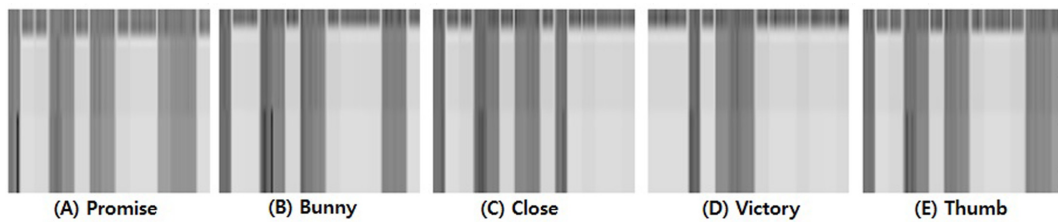


그림 6. 스펙트로그램 변환
Fig. 6. Spectrogram transformation

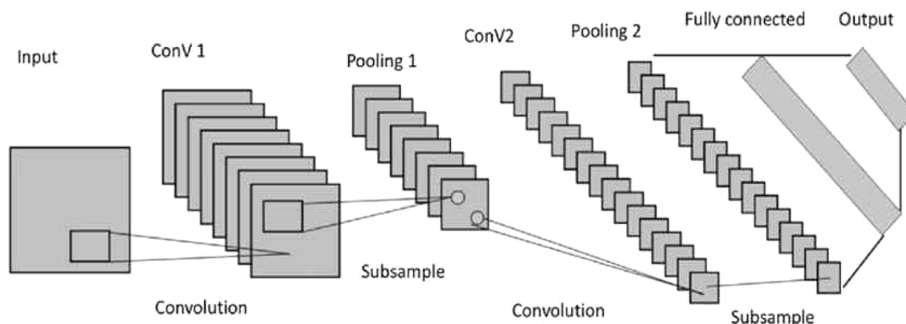


그림 7. CNN 모델 블록 개요도
Fig. 7. CNN model structure for this study

표 1. 모델 레이어 정의 및 구조

Table 1. Definition and structure of model layer

Layer (type)	Output Shape
conv2d (Conv2D)	(28, 28, 16)
max_pooling2d (MaxPooling2D)	(14, 14, 16)
dropout (Dropout)	(14, 14, 16)
conv2d_1 (Conv2D)	(12, 12, 64)
max_pooling2d_1 (MaxPooling2)	(6, 6, 64)
dropout_1 (Dropout)	(6, 6, 64)
conv2d_2 (Conv2D)	(4, 4, 64)
max_pooling2d_2 (MaxPooling2)	(2, 2, 64)
dropout_2 (Dropout)	(2, 2, 64)
flatten (Flatten)	(256)
dense (Dense)	(256)
dropout_3 (Dropout)	(256)
dense_1 (Dense)	(5)

	A	B	C	D	E
A	10				
B		10			
C			10		
D				10	
E					10

그림 8. 컨퓨전 매트릭스

Fig 8. Confusion matrix (X: actual, Y: predicted)

II. 연구 결과

본 연구에서는 상수 대입법을 사용하는 전통적인 손동작 예측 방법과 CNN 기반 스펙트로그램 분석 기법을 이용한 손동작 예측 방법을 통해 실험을 진행하고 예측 정확도를 비교하였다.

1. 상수 대입법을 이용한 예측 정확도

표 2는 상수 대입법을 이용한 각도에 따른 손동작별 손동작 인식 예측 정확도를 보여준다. 기존의 전통적인 손동작 예측방법인 상수 대입법을 사용한 정확도는 모션에 따라 A 모션은 87.5%, 모션은 87.5%, C 모션은 100%, D 모션은 97.9%, E 모션은 85.0%의 정확도를 나타내었다. 또한 손바닥이 카메라를 바라보고 있을 경우 가장 인식률이 높았으며 이때 각도를 0도라고 했을 때, 90도에서 89.3%, 270도에서 86.7%로 낮았으며 높은 정확도를 보인 것은 0도일 때 98.0%, 315도일 때 99.3%로 손바닥이 카메라 정면인 경우 모든 동작에 대하여 정확도가 상대적으로 높았다.

표 2. 상수 대입법을 이용한 손동작별 손동작 인식 예측 정확도

Table 2. Accuracy of gesture predictions of 5 hand gestures cognition using the local transform Euler angles

Motion \ Angle	0	45	90	135	180	225	270	315	Average (number)	Accuracy (%)
A	30	23	27	27	23	30	21	29	26.25±3.49	87.5
B	27	30	21	22	26	30	24	30	26.25±3.65	87.5
C	30	30	30	30	30	30	30	30	30.00±0.00	100
D	30	29	30	28	30	28	30	30	29.37±0.91	97.9
E	30	30	26	30	30	27	25	30	28.50±2.13	95.0

2. 손 모델 인식 반응시간

개발한 립모션 기반 증강현실 인지 재활훈련 시스템의 최대 소요시간은 0.9초였으며 최소 소요시간은 0.512초로 평균 0.727초 내로 손 모델이 소실된 후 인식하는 결과 값을 보였다. 이때 소실되는 것을 방지하기 위해 손 모델 렌더링 활성화에 관한 SDK 기능을 비활성화한 경우 오히려 인식률이 감소하는 경향을 보였다. 이 경우 최대 소요시간은 1.395초이었으며 최소 소요시간은 0.704초로 전체적으로 크게 재인식에 걸리는 시간이 증가하였다. 손 모델 렌더링 활성화 기능을 사용하지 않고 손뺌 동작의 검출을 수행하였을 때 평균 0.727초 이내에 손 모델이 복구되었다.

3. CNN을 이용한 예측 정확도

각 동작 별 10회 예측 실험 결과 CNN 기반의 스펙트로그램 분석기법을 통한 손동작 예측 정확도는 5개 모든 손동작에서 100%로 측정되었다. 그림 8은 실험 결과와 정답이 맞았던 항목에 대하여 컨퓨전 매트릭스(confusion matrix)로 나타낸 것이다.

III. 고 찰

본 논문에서 실험한 CNN 기반의 스펙트로그램 분석 기법을 이용한 손동작 인식은 5가지 손동작에 대하여 100%의 정확도를 보였으며 이러한 결과는 Khalif 등[11]이 손동작 인식 정확률을 전체적으로 91.3%의 정확도를 보였던 것과

Stinghen 등[10]의 SVM을 적용한 손동작 인식 모델에서 “Close” 동작에서 97%의 정확도와 “Thumb” 동작에서 약 98%의 정확도를 냈던 결과보다 높은 인식 정확도를 나타내었다.

현재의 기술은 현재 활용 가능한 하드웨어와 기본적인 소프트웨어 상에서는 매우 고도화되어있어 향후 연구과제에서는 다른 기법이나 하드웨어인 데스크 스테이션, 인공지능을 통한 사람이 사람의 손을 인식하는 수준의 결과물 개발이 기대되고 있는 실정이라고 말할 수 있다. 본 논문에서의 손동작 인식 정확도는 기존 SVM 방식에 비해 높게 나왔으며 이는 첫 번째로 실험 중 각도를 제한하여 손바닥을 바닥면에 대하여 45도 이상의 각도로 회전하지 않도록 하였으며 두 번째로 정확도가 높았던 손동작만을 선별하여 분류를 진행하였다는 점이다. 그러나 참고하였던 SVM을 사용한 실험들 또한 각도를 제한하여 실험하였다[12,13]. 본 실험은 손가락의 각도 값만을 추출하여 주파수 변환을 통해 이미지 분류 모델에 통과시켰다는 점에서 높은 결과를 얻어낸 것으로 보인다. 본 실험의 결과만 가지고 논하였을 때에는 기존의 SVM 모델을 사용한 다른 논문들보다 더 높은 수준의 인식률을 보였다고 결론지을 수 있다. 인식률이 높은 모델을 사용하였고 각도 또한 인식률이 높은 각도에서만 측정하도록 설정한 제한 조건이 있으나 이러한 제한조건에 대해서도 다른 논문들 또한 특정 손동작에 대해서만 정확도가 높게 나타나거나 각도를 카메라의 정면에 고정된 0도 각도에서만 측정한 결과를 실험 결과로 제시하였기에 결과를 비교하기에 문제가 없다고 판단된다. 또한 앞서 SVM의 한계점과 CNN의 특징이 갖는 활용 가능성에 대해 설명한 대로 SVM을 사용한 이진화 모델은 사전 기록한 특정 라벨만을 추천한다는 단점이 있으나[16] CNN과 스펙트로그램을 이용한 경우 손가락 굽힘 정도를 주파수 형태로 변환하여 이미지화 시키는 것으로 각 손가락별 특성에 대응될 수 있다는 점에서 차별성을 갖는다[17]. 손가락 마다 하나씩의 스펙트럼을 분석하도록 알고리즘을 구성할 수 있으며 5개의 모델로 모든 손동작을 검출하고 이를 인식할 수 있다는 장점이 있어 손동작 인식의 정확도가 상승할 수 있을 것이라고 가정을 세웠고 실험결과를 통해 이를 입증할 수 있었다. 이는 수화 인식 등과 같이 손가락의 굽힘 정도를 이진화 판별하여 표시하고자 하는 것이 목적인 SVM 기법보다 메타버스 콘텐츠에 적용하기에 용이하다[16,17]. 손가락의 굽힘 정도를 가지고 손동작을 판별할 때 기존에 메타버스 콘텐츠에서 사용하던 특정 상수를 입력하여 그 이상의 각도일 경우 굽힘으로 분류하는 상수 대입법은 사람에 따른 편차에 따라 그 결과 값이 달라질 수 있고 작성자의 실수로 인해 잘못된 상수가 기입되었을 경우에는 해당 상수를 수정하기 전까지 원하는 결과를 얻을 수 없다는 치명적인 단점을 가지고 있다[18]. 이러한 문제를 SVM은 어떠한 변환 없이 이미지를 가지고 특정 라벨

링에 대한 참 거짓을 구분하는 것만으로 해결하였다. 그러나 각 라벨마다 학습을 다시 시도해야 하며 실시간으로 동작을 인식해야 하는 메타버스 콘텐츠 내부에서 사용할 경우 인식 시킨 라벨 수만큼 모델을 적용시켜 라벨 목록 안에서 일치하는 라벨을 일일이 대조하여야만 한다. 본 연구에서는 촬영된 이미지가 사전인식된 라벨에 속하는지 아닌지를 참 또는 거짓으로 판별하는 SVM 보다 촬영된 손동작을 여러 가지 카테고리 중 어디에 속하는지 구분하는 CNN 모델이 높은 손동작 인식 정확도를 보였다.

IV. 결 론

본 연구에서 사용된 CNN 기법은 SVM을 활용한 손동작 인식 방법보다 예측 정확도가 높고 주변 배경변화 등에 유려하여 증강현실 기반 인지재활 훈련시스템에 유용하게 적용될 수 있을 것으로 기대된다. 향후 연구에서는 립모션과 같은 상용 적외선 카메라가 내장 SDK를 분석하는데 한계가 있음을 분명히 인지하고 오픈소스와 RGB 카메라 3대 이상으로 구성된 데스크 스테이션 구축을 통해 정해진 범위 내에서 손의 움직임을 3차원으로 하거나 단일 카메라에서 딥러닝 모델을 통해 손의 움직임을 추측해 좌표 값으로 변환해줄 수 있는 다양한 연구가 필요할 것으로 생각된다.

References

- [1] Panic AS. Addressing patient motivation in virtual reality based neurocognitive rehabilitation. Delft University of Technology, 2010.
- [2] Doidge N. The Brain That Changes Itself: Stories of Personal Triumph from the Frontiers of Brain Science. Penguin Group. 2007.
- [3] Burdea CG. Virtual Rehabilitation: benefits and Challenges. Methods of Information in Medicine. 2003;42(5):519-523.
- [4] Alankus G, Lazar A, May M, Kelleher C. Towards customizable games for stroke rehabilitation. Proceedings of the 28th Annual CHI Conference on Human Factors in Computing Systems (CHI '10). 2010;2113-2122.
- [5] Burke JW, McNeill M, Charles D, Morrow P, Crosbie J, McDonough S. Designing engaging, playable games for rehabilitation. Proceedings of the 8th International Conference On Disability, 2010.
- [6] Donker T, Cornelisz I, van Klaveren C, van Straten A, Carlbring P, Cuijpers P, van Gelder JL. Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: A randomized clinical trial. JAMA Psychiatry. 2019; 76:682-690.
- [7] Tsao YC, Shu CC, Lan TS. Development of a Reminiscence Therapy System for the Elderly Using the Integration of Virtual Reality and Augmented Reality. Sustainability. 2019;11:2-10.
- [8] Annema JH, Verstraete M, Abeele VV, Desmet S. Video games in therapy: a therapist's perspective. International Journal of Arts and Technology. 2010;6(1):94-98.

- [9] Andersson G, Carlbring P, Titov N, Linderfors N. Internet interventions for adults with anxiety and mood disorders: A narrative umbrella review of recent meta-analyses. *Can. J. Psychiatry*. 2019;64(7):465-470.
- [10] Stinghen I, Chen E, da Silva JM, da Silva R. Gesture recognition using leap motion: a comparison between machine learning algorithms. *ACM SIGGRAPH 2018 Posters*, 2018.
- [11] Khelil B, Amiri H, Hand Gesture Recognition Using leap motion controller for recognition of arabic sign language. *Computer Science*. 2016.
- [12] Tran DS, Ho NH, Yang HJ, Baek ET. Real-Time Hand Gesture Spotting and Recognition Using RGB-D Camera and 3D Convolutional Neural Network. *Applied Sciences*. 2020;10(2): 722.
- [13] Huang DY, Hu WC, Chang SH. Vision-Based Hand Gesture Recognition Using PCA+Gabor Filters and SVM. *15th International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2009.
- [14] Côté-Allard U, Fall CL, Drouin A, Campeau-Lecours A, Gosselin C, Glette K, Laviolette F, Gosselin B. Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2019;27(4):760-771.
- [15] Radianti J, Majchrzak T, Fromm J, Wohlgenannt I. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 2020;147:103778.
- [16] Huang DY, Hu WC, Chang SH. Gabor filter-based hand-pose angle estimation for hand gesture recognition under varying illumination. *Expert Systems with Applications*, 2011;38(5): 6031-6042.
- [17] Sagayam KM, Hemanth DJ. Hand posture and gesture recognition techniques for virtual reality applications: a survey. *Virtual Reality*. 2017;21:91-107.
- [18] Badi H. Recent methods in vision-based hand gesture recognition. *International Journal of Data Science and Analytics*. 2016;1:77-87.