

인공신경망 기반 손영상 인식기술을 이용한 가위바위보 게임

장연수^{0*}, 김다예^{*}, 박동진^{*}, 한윤성^{**}, 전수빈^{*}, 서동만^{00*}

^{*}대구가톨릭대학교 컴퓨터소프트웨어학부

^{**}대구가톨릭대학교 컴퓨터소프트웨어학과

dlakddustn@cu.ac.kr, daye0529@cu.ac.kr, qufehdwls@cu.ac.kr, sa01106@cu.ac.kr,
marsberry@cu.ac.kr, sarum@cu.ac.kr

A Rock-paper-scissors Game Using Hand Image Recognition Technology based on Artificial Neural Network

Yeon-Su Jang^{0*}, Da-ye Kim^{*}, Dong-Jin Park^{*}, YunSung Han^{**},
Soobin Jeon^{*}, Dongmahn Seo^{00*}

^{*}School of Computer Software, Daegu Catholic University

^{**}Department of Computer Software, Daegu Catholic University

요 약

최근 코로나 19로 인한 사회적 거리 두기 확산에 따라 언택트 문화가 새로운 패러다임으로 등장해 사회 전반으로 확산되고 있다. 언택트 문화의 확산으로 컴퓨터를 사용할 때 직접적인 접촉이 있는 키보드나 마우스 같은 입력장치는 공공장소에서 여러 사람이 접촉할 경우 문제가 생길 수 있다. 본 논문에서는 웹캠을 통해 입력된 영상에서 손동작을 인식하는 합성곱 신경망을 학습하고 결과로 나온 추론 모델을 이용하여 비접촉 가위바위보 게임을 구현하였다.

1. 서론

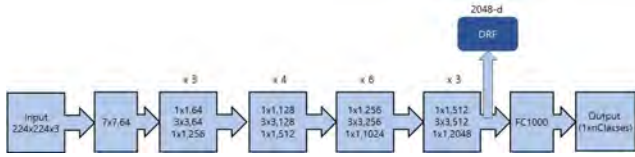
컴퓨터는 일반적으로 사람에게 직접 키보드나 터치 패드, 마우스의 입력을 통해 명령을 받게 된다. 하지만 최근 코로나 19로 인해 ‘사회적 거리 두기’ 확산에 따라 언택트 문화가 새로운 패러다임으로 등장해 사회 전반으로 확산되고 있다. 이러한 사회적 분위기에 따라서 입력장치를 사용하지 않고도 컴퓨터와 사용자 간의 상호작용을 할 수 있도록 하는 연구가 진행되고 있다[1].

하지만 키보드와 같은 입력장치를 사용하지 않고는 사용자와 컴퓨터 간의 원활한 상호작용이 힘들 수 있다. 그렇기에 여러 인체 중에서 큰 불편함 없이 다양한 동작을 만들어내는 손동작을 인식하는 것이 가장 효율적인 방법이다. 손동작 영상을 이용한 물체 인식 방법은 크게 두 가지 인공지능과 영상 처리로 나뉘어진다. 영상처리의 경우 손을 인식하기 위해 피부색을 이용하여 손 모양을 추출하는 방법,

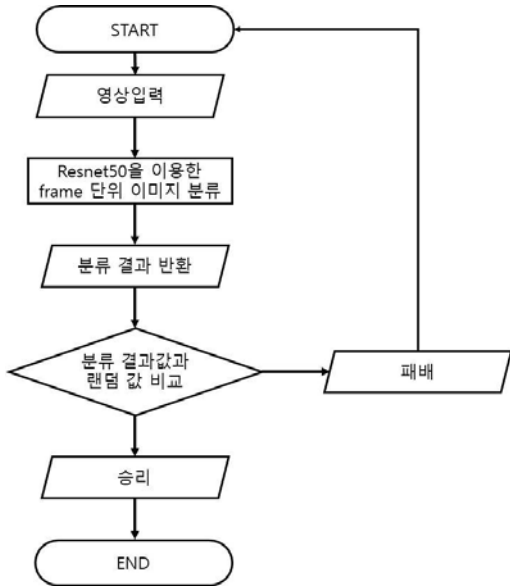
저니키(Zernike) 알고리즘을 이용하여 손 모양을 추출하는 방법, 등이 있다. 피부색을 이용하여 손 모양을 추출하는 방법은 배경색이 피부색과 비슷할 경우 문제가 생기는 단점이 있으며 영상 처리의 특성상 빛에 취약하다는 단점이 있다. 또한, 저니키 알고리즘을 이용하는 방법은 디지털 이미지에서 극 좌표계를 기준으로 하여 저니키 모멘트를 구하기란 쉽지 않은 단점을 지닌다[2-3]. 따라서, 인공지능을 이용한 방법 중 이미지 분류에 성능이 뛰어난 합성곱 신경망(Convolutional Neural Network)알고리즘을 활용하는 방법을 제안한다. 이는 이미지 분류학습으로 사람이 클래스에 따라 이미지들을 분류하고 이를 인공신경망이 학습한 뒤 학습에서 사용하지 않은 새로운 이미지가 어떤 클래스의 요소인지 자동으로 분류하는 인공신경망을 말한다. 이미지 분류학습은 이미지 내에 해당 사물이 있는지 없는지에 대하여 판별하고 해당 이미지를 분류하는 기술로서 이미지 내의 손 모양을 판별하고 분류하기 위해 선행되어야 하는 기술이다. 합성곱 신경망을 이용한 분류기 학습 단계는 다음과 같다. 첫째, 분류하고자 하는 이미지 데이터를 모아

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학지원사업의 연구결과로 수행되었음(2019-0-01056).

⁰⁰교신저자



(그림 1) Resnet50 모델망



(그림 2) 시스템 흐름도

각 이미지 데이터에 라벨을 부여한다. 둘째, 합성곱 신경망 모델을 통해 픽셀 단위 특징 추출 후 분류 학습을 진행한다. 셋째, 학습이 완료되면 입력 이미지에 대한 추론을 진행한다.

본 논문에서는 합성곱 신경망 중 이미지 분류 모델인 Resnet 50을 이용한 전이학습을 통해 손영상에서 손 모양을 인식하고 해당 인식결과를 이용하여 가위바위보 게임을 구현한다.[4-5]

2. 관련 연구

인공지능을 이용한 물체 탐지 기술은 합성곱 신경망을 이용하여 영상을 분류 방법과 물체 탐지 딥러닝을 이용하는 방법, 객체 분할이 있다. 이미지 분류 딥러닝은 입력된 영상을 학습된 모델망을 통해 클래스 중 하나의 결과값으로 반환하는 것이다. 물체 탐지 딥러닝은 입력된 영상 내의 물체를 학습된 모델망을 통해 이미지 내의 위치를 경계상자(Bounding Box)의 형태로 결과를 반환한다. 객체 분할 딥러닝은 영상 내의 물체를 학습된 모델망을 통해 분할하고 물체의 형태를 마스크(Mask)값으로 반환한다.

본 논문에서는 손의 위치나 형태보다는 가위바위



(그림 3) UI 구현

보의 인식을 목적으로 하여 손의 위치와 형태에 대한 데이터는 고려하지 않는다. 추출된 손 영상은 가위와 바위, 보 만을 분류한다. 합성곱 신경망을 이용하여 영상을 가위, 바위, 보로 분류한다. 손동작을 인식하기 위해 합성곱 신경망 중 Resnet50을 이용하였다. 그림1은 대표적인 합성곱 신경망 중 Resnet50으로써 다량의 이미지를 모델망을 통해 학습하고 해당 이미지를 추론하는 분류 딥러닝 알고리즘이다[4].

3. 제안 시스템

3.1 시스템 설계

전체적인 시스템의 흐름은 웹캠을 통해 입력된 영상을 이용하여 Resnet50 모델망을 통해 손동작을 인식한다. 가위바위보 게임에서 컴퓨터의 경우 랜덤 함수를 이용해 가위와 바위, 보를 출력한다. 그리고 인식결과와 컴퓨터 출력 결과를 비교하여 가위, 바위, 보의 승, 패 결과를 사용자 인터페이스(User Interface)의 형태로 출력한다. 그림2는 해당 시스템의 흐름도이다. 시스템은 크게 3가지의 모듈로 나뉜다. 3가지의 모듈은 영상인식을 위한 Resnet50기반 인공지능 모델, 컴퓨터가 랜덤 함수를 이용하여 가위, 바위, 보를 랜덤하게 반환하는 모듈, 사용자의 편의성을 위한 사용자 인터페이스로 구성된다.

3.2 Resnet50기반 인공지능 모델

입력된 영상이 어떤 손동작인지 분류하기 위하여 Resnet50기반 전이학습을 통해 가위, 바위, 보를 분류할 수 있는 분류기를 만들었다. 총 390장의 가위, 바위, 보 이미지 데이터셋을 만들어 분류기를 학습하였다.

3.3 사용자 인터페이스

사용자의 편의성을 위해 사용자 인터페이스를 설계하였다. 프로그램의 전체적인 사용자 인터페이스 구성은 그림3과 같다. 시작과 종료 버튼을 이용하여

<표 1> 실험 환경

Hardware		Software Version	
OS	Window10	Python	3.8.3
CPU	Intel Core i5-8265U		
RAM	8195MB	keras	2.43
Webcam	Microsoft 1080pHD sensor	PyQt	5.9.2



(그림 4) 가위와 바위, 보 인식 결과

<표 2> 각 도에 따른 손동작 인식횟수 비교표

	가위	바위	보
30도	19/20	15/20	16/20
60도	20/20	14/20	10/20
90도	17/20	2/20	18/20

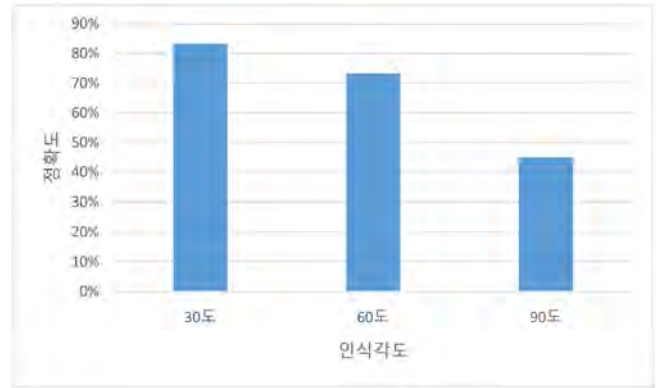
게임을 시작 및 종료시킬 수 있도록 하였고, 사용자의 게임 기록을 남기기 위해 닉네임 입력 기능 및 사용자별 승률 확인 기능을 구현하였다. 또한, 직접 승, 패를 확인할 수 있도록 결과를 사용자 승률, 전체 승률로 나누어 시각화하였다.

4. 구현 및 결과

구현된 Resnet50의 성능을 테스트하기 위한 시스템의 구현 및 실험 환경은 표 1에 정리되어 있다.

테스트는 웹캠을 이용하여 약 30도, 약 60도, 약 90도 일 때 가위, 바위, 보 입력 영상을 받았다. 그림4는 가위, 바위, 보의 인식결과 이미지이다. 표 2는 각 손동작을 30°, 60°, 90° 마다 20회씩 인식했을 때의 결과이다. 가위, 바위, 보 중 아무것도 인식하지 않았을 때, 기본적인 인식 값이 가위이기 때문에, 바위나 보를 내도 가위로 인식되는 경우가 자주 있어 가위의 정확도만 높은 것을 볼 수 있다. 따라서 바위나 보를 냈을 경우 정확도가 낮으나 가위의 경우 정확도 높다는 것을 알 수 있다.

그림 5는 인식 각도 별 정확도 측정 그래프이다. 각도 별 손동작 인식 정확도는 약 30도 일 때 83.33%, 약 60도 일 때 73.33% 그리고 90도 일 때 61.66%이다. 학습 데이터가 30도 각도에서 찍은 데



(그림 5) 인식 각도 별 정확도 측정 그래프

이터가 많아 각도 중에서 30도의 각도로 손동작을 인식했을 때, 다른 각도들에 비해 정확도가 높은 것을 확인할 수 있다.

5. 결론 및 향후 계획

최근 코로나 19로 인해 언택트라는 단어가 사회적으로 강조되고 있다. 그에 따라 컴퓨터에 값을 입력하기 위해 키보드나 마우스가 아닌 비접촉방식의 입력을 위해 많은 연구가 진행되고 있다.

본 논문에서는 입력받은 손영상에서 손 모양을 인식하는 인공지능을 활용하여 가위바위보 게임을 구현하였으며 각 손동작의 인식 각도를 비교하여 실험을 진행하였다. 비교한 인식 각도 30도, 60도, 90도 중 30도가 83.33%의 정확도로 가장 좋은 인식 정확도를 보였다. 향후 계획으로는 Resnet50 알고리즘이 손의 각도에 따라 인식율이 떨어지는 사례를 보여주므로 좀 더 다양한 손의 패턴을 인식할 수 있도록 학습모델을 개선할 것이다.

또한, 현재 다양한 손의 각도와 모양을 고려한 이미지 데이터 셋이 부족하여 인공지능 성능이 낮은 것을 보여 추가적인 데이터 셋을 구축할 예정이다. 위의 개선사항을 통해 손동작 인식률이 높아질 경우 응용하여 수어 인식 인공지능 알고리즘 개발에 도움이 될 것으로 기대된다.

참고문헌

[1] 배영임, 신혜리, “코로나19, 언택트 사회를 가속화하다.” 이슈&진단, 1-26, 2020

[2] 안호석, 사인규, 백영민, 안윤석, 최진영, “영상 처리 기반의 가위 바위 보 게임 로봇”, 대한전기학회 대한전기학회 학술대회 논문집, 2008, 327-328 2pages

[3] 서정민, 김병주, 문형만, 박창선, 황정환, “손영상

인식기술을 이용한 가위바위보 게임”, 한국멀티미디어학회 학술발표논문집, 2010, pp.510 ~ 512, 3pages

[4] HE, Kaiming, et al. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. p. 770-778.

[5] 김진성, 송재열, 김하얀, 이진국, “딥러닝 기반 이미지 자동인식 기술을 활용한 사무집기 자동인식과 정보관리 시스템과의 연동방안”, 한국퍼실리티매니지먼트학회 논문집, 2017, p. 75-80