

# 제스처 인식을 이용한 비접촉식 미디어 아트 키오스크

김유라 서솜이 박구만

서울과학기술대학교 전자IT미디어공학과

somvest@naver.com

## Touchless Media Art Kiosks using Gesture Recognition

Kim, Yura Seo, Somyi Park, Gooman

Seoul National University Of Science And Technology

### 요 약

본 논문은 딥러닝과 핸드 제스처, 미디어 파이프를 이용하여 비접촉식 키오스크를 만들어 (코로나19로 인한) 감염병 예방 및 대형마트에서 원하는 물건을 검색 후 편리한 구매를 위해 구현한 내용을 다룬다. 단순한 핸드 제스처 인식만을 이용한 것이 아니라 그림을 그리는 비접촉 제스처 인식을 더해 하나의 예술로써 제스처 인식을 더 많이 사용하고자 했다. 또한 하나의 손가락으로 그림을 그리는 것이 아닌 주먹을 쥐었을 때 그 중심을 인식해 그림을 그리는 방법을 이용해 기존의 방법보다 더 안정감 있게 그릴 수 있게 구현하였다. 현실에서 사용하기 위해 세부적인 기능들은 학습을 통해 기존보다 정확도가 향상된 미디어 파이프를 이용하였다. 빠른 처리 속도와 정확성에 초점을 두는 것이 아닌 하나의 미디어 아트로써 키오스크를 설계하였다.

## 1. 서론

### 1.1 동기

현재 IT 기술이 발전하면서 키오스크를 사용하지 않는 매장을 보기 힘들 정도로 키오스크를 사용하는 매장의 수가 급격히 늘고 있다. 하지만 코로나19 확산으로 감염에 대한 두려움이 커지면서 불특정 다수가 이용하는 키오스크에 불안함을 느끼는 사람들도 많아졌다. 영국에서는 패스트푸드 매장 상당수에 설치된 키오스크에서 인체에 치명적인 박테라가 검출됐다고 보도했다. 한 전문가는 키오스크에서 검출된 포도상 구균은 패혈증 뿐만 아니라 심하면 사망으로 이어질 수도 있다며 “매우 전염성이 강한 이 박테라가 기기에서 발견된 것이 우려스럽다”라고 덧붙였다. 키오스크에 세균이 많은 것은 사용자들이 손으로 코를 만지는 것에서 시작된다며 “손가락이 닿은 화면을 통해 다른 사용자들에게 옮겨지고 상처 부위로 감염될 경우 더 위험해진다고 말했다.

리가 나올 수 있게 연결시킴으로써 하나의 미디어 아트로 키오스크를 설계하였다.

### 1.2 차별성

본 시스템은 대형마트에서 핸드 제스처로 키오스크를 조작해 상품-찾기 프로세스를 수행하는 것이 목적이다. 그 이유는 많은 종류의 상품을 파는 대형마트에는 아직 키오스크가 사용되지 않아 직원에게 상품의 위치를 물어보는 상황이 빈번한데, 사람과의 접촉을 줄여 코로나19 감염을 예방할 수 있으며 기존보다 편리한 쇼핑을 기대할 수 있기 때문이다.

또한 검지로 그림을 그리는 방식이 아닌 손의 중심을 인식해 카테고리 그림을 그리는 방식을 선택했다. Google 사의 MediaPipe에서 제공하는 핸드 제스처를 사용하지 않고 직접 손의 모양을 약 200번 학습시켜 기존 Google 사의 MediaPipe 보다 빠르고 정확한 핸드 트래킹을 사용하였다.

## 2. 관련 연구

Google 에서 모바일, 엣지, 클라우드에 머신러닝을 실행할 수 있는 솔루션 및 애플리케이션을 제공하는 MediaPipe를 공개했다. 이는 멀티 모달 머신러닝 파이프라인을 구축하기 위한 그래프 기반 프레임워크이다. 이는 모바일 장치, 워크 스테이션 및 서버에서 실행되는 크로스 플랫폼이며 모바일 GPU 가속을 지원한다. MediaPipe를 사용하면 추론 모델 및 미디어 처리 기능과 같은 모듈 식 구성 요소의 그래프로 적용된 머신러닝 파이프라인을 구축할 수 있으며 오디오 및 비디오 스트림과 같은 데이터들이 그래프에 입력되고, object localization 및 face landmark 와 같은 정보를 출력한다.

하지만 복잡한 배경일수록 속도가 느리고 정확도가 떨어지는 것을 볼 수 있는데 이를 해결하고자 데이터 학습을 실행시켰다. 핸드 제스처는 MediaPipe's Hand Landmark의 위치값, 점들 사이의 각도, 점들 사이의 거리를 정의해 손의 정보를 인식하여 정의내리도록 한다. 이 점을 이용해 직접 손가락의 각도와 손의 위치를 계속 변경해가며 한 동작에 대해 약 200-300번의 데이터를 학습 시켰다.



Fig. 1 국내 코로나19 확진자 그래프

이에 본 논문은 영상 처리 및 인식 기술을 이용해 키오스크의 스크린을 접촉하지 않고도 화면을 제어할 수 있는 시스템을 구현하여 코로나19의 확산을 예방하고자 했다. 더 나아가 메뉴판의 카테고리를 직접 손으로 선택하는 것이 아니라 누구나 쉽게 그림을 그리고 그에 맞는 카테고리

### 3. 키오스크 시스템 구조

#### 3.1 H/W 구조

본 시스템의 구현은 QuickDraw 기반의 핸드 제스처와 MediaPipe Framework을 활용하여 사용자가 화면을 접촉하지 않고 제어할 수 있게 함으로써 이루어진다. 이 시스템의 H/W는 그림 (2)와 같이 디바이스 (노트북), 웹캠 그리고 모니터로 구성되어 있다. 디바이스가 제스처 인식과 그림 인식을 통해 웹캠의 이미지를 각 변수로 받은 후 UI의 컴포넌트를 제어한다.

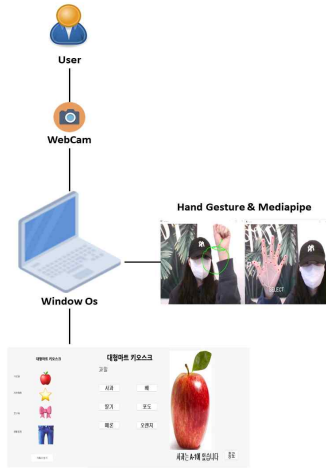


Fig. 2 제스처 인식 키오스크 H/W 구조

#### 3.2 S/W 구조

본 시스템의 S/W 동작을 그림 (3) 시퀀스 다이어그램으로 구현하였다. 키오스크 사용을 Python 기반 전처리 후처리로 OpenCV 라이브러리를 통해 웹캠에서 실시간 영상을 트래킹 시스템에 전송하여 사용자의 손동작을 인식한다. 손동작에 따라 Mediapipe를 통해 손의 정보들을 인식하고 이를 변수로 받으며 PyQt 기반에 키오스크 UI를 제어하고 컴포넌트에 수행 명령을 전달한다. 이러한 동작의 반복으로 키오스크의 S/W를 구성하였다. 또한 RGB 카메라가 아닌 웹캠만을 사용하여 접근성을 높였다.

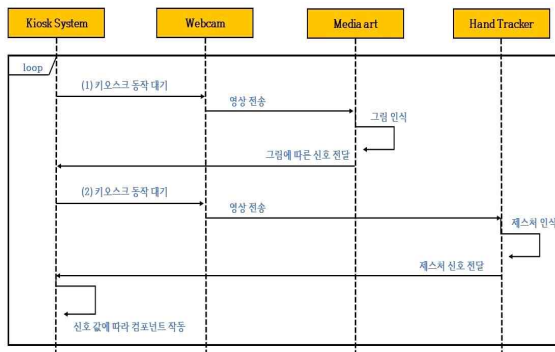


Fig. 3 제스처 인식 키오스크 S/W 시퀀스 다이어그램

### 4. Media art

간단한 제스처 인식 응용 프로그램인 AirGesture 와 Google에서 개발한 온라인 게임인 QuickDraw를 결합했다. Quick Draw는 구글에서 만든 머신러닝 모델을 이용한 낙서게임이다. 플레이어가 낙서(drawing)을 그리면 제한시간내에 머신 러닝 모델이 낙서의 클래스를 맞추는 게임이다. 이 프로그램을 사용하는 이유는 image detection을 사용하면 검

출이 느리고, 명암과 색깔로 구분해서 정확도가 떨어지는 반면, rmn을 사용하면 정확도가 올라가고 더 편리하기 때문이다. 또한 Quick Draw는 마우스를 통해 화면에 그림을 그리는데, 여기서 input 값만 바꾸어서 넣어주면, 손으로 그림을 그리는 효과를 보여 줄 수 있다. 따라서 최종적으로 구성하는 시스템에서 손으로 그림 그리기에 적합한 프로그램이라고 할 수 있다.

웹캠과 손만을 이용하여 응용 프로그램을 쉽게 이용할 수 있도록 제작했다. 우선 웹캠 앞에서 손을 들고 주먹을 쥐면 손의 중간 지점을 감지하여 빨간색 점으로 강조 표시가 된다. 이는 키오스크를 이용할 때 검지 하나를 이용하는 것보다 무게감이 있어 현실적으로 안정감을 준다. 키보드의 스페이스 바를 누르면 그림을 그리는 것을 시작하고 다시 스페이스 바를 누르면 멈출 수 있다.

위 소프트웨어를 활용한 응용프로그램을 그림 (4)에서처럼 비접촉 키오스크의 카테고리를 선택하는데 이용하고자 한다. 그림 (4)는 첫 번째 Base Window 초기 시작화면이다. 식료품과 가전제품, 문구류 그리고 생활용품의 해당되는 그림으로 구성되어 있다.



Fig. 4 대형마트 키오스크 메뉴판

각 카테고리에 맞는 그림을 누구나 쉽게 그릴 수 있고 모양이 겹치지 않는 이미지 4개를 사용하여 오류를 줄였다. apple, bowtie, star, pants 와 유사한 그림 그렸을 때 apple, bowtie, star, pants 가 나오도록 데이터를 훈련시켰다. 사과를 그리면 식료품 카테고리로 연결되고, 나머지도 각 카테고리에 연결하고자 한다. 그림 (5)와 같이 대표 사과를 예시로 보여주었다. 그림 (4)에 있는 카메라 열기 버튼을 누르면 웹캠이 켜지면서 그림을 그릴 수 있는 창이 열린다.

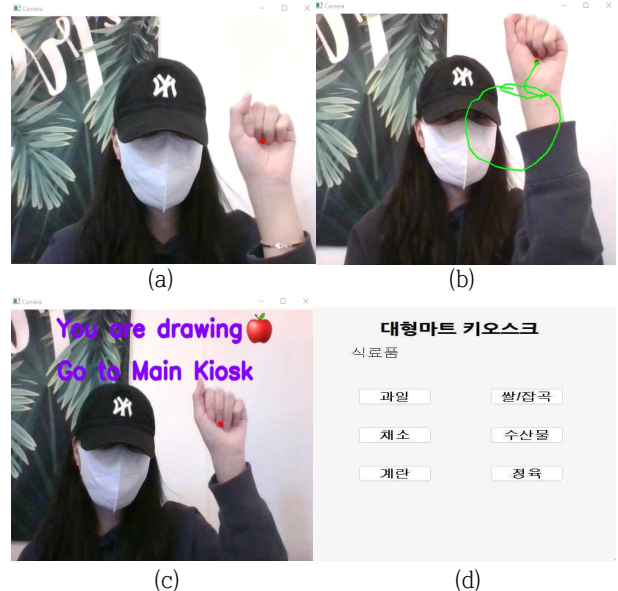


Fig. 5 미디어아트 데모 이미지

그림 (5)에서 (a)를 보면 손의 중심 부분을 그림의 시작점으로 표시하고 눈에 잘 보이게 빨간 점으로 강조한다. (b)와 같이 식품 카테고리 연결되도록 사과를 그린다. 그리는 것을 종료하면 (c)와 같이 그림을 사과로 인식하고 (d)에서 볼 수 있듯이 식품 카테고리를 나타내는 키오스크로 연결하고자 한다.

데이터를 많이 학습시켜서 그림을 그리는 배경이 복잡한 환경에서도 잘 실행되는 것을 볼 수 있다.

### 5. Mediapipe

그림 (6)는 핸드 트래킹의 제스처를 분류한 표이다. 키오스크의 사용자 전 연령층으로 삼았기 때문에 사람이 할 수 있는 가장 쉬운 손동작 6가지로 구성하였다.

기능번호	기능명	핸드 제스처
1	화면 켜기 (ON)	
2	화면 종료하기 (OFF)	
3	선택하기 (SELECT)	
4	취소하기 (CANCEL)	
5	이전으로 가기 (BEFORE)	
6	다음으로 가기 (AFTER)	

Fig. 6 핸드 제스처 분류 표

기존의 MediaPipe 라이브러리는 모델 용량이 작고 속도가 빠른 장점을 가지고 있으나 손의 모양을 오검출이나 미검출하는 경우가 있다. 이를 보완하기 위해 학습 데이터의 개수를 늘려 더 정확하고 빠르게 손의 모양을 검출하게 했다. MediaPipe를 발전시키기 위해 한 동작에 대해 약 200-300개의 학습 데이터를 모았다. 그림 (7)의 (a)와 (b)에서 볼 수 있듯이 손 모양이 빠르고 정확하게 나오며 그림 (8)의 (a)는 처음 MediaPipe의 학습 결과이고 (b)는 데이터를 늘려 훈련시킨 기존보다 발전된 학습 완료 그래프이다.

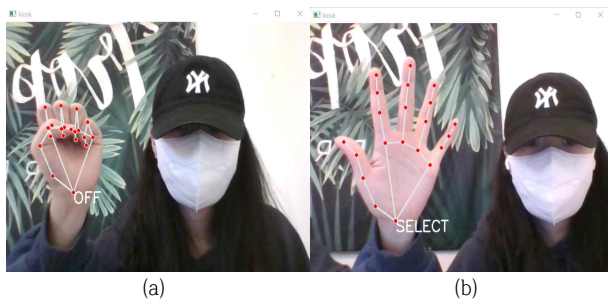


Fig. 7 MediaPipe 실행 결과

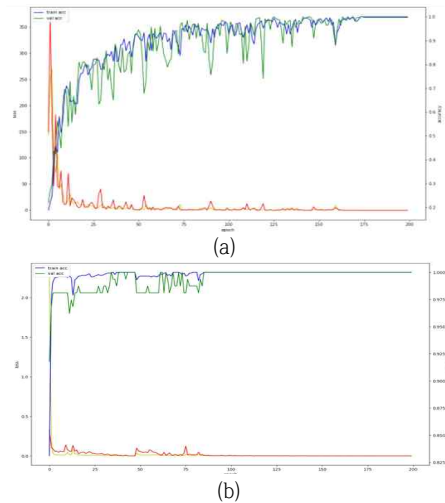


Fig. 8 MediaPipe 학습 완료 그래프 (train accuracy는 파란색, validation accuracy는 초록색, loss는 노란색과 빨간색. accuracy가 100%가 나온 것을 볼 수 있다.)

화면 켜기(On) 손동작을 하면 그림 (9)와 같이 카테고리를 선택할 수 있는 키오스크 메뉴창이 켜진다.

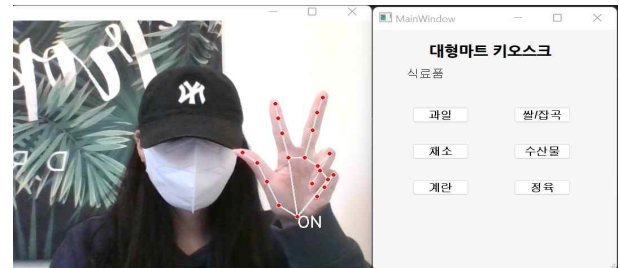


Fig. 9 MediaPipe 실행 결과

핸드 제스처를 이용해 키오스크를 세부적으로 조작하려고 한다. 이전에 만들었던 6가지 MediaPipe를 활용하여 그림(10)과 같이 키오스크를 실행하는 데 화면을 접촉하지 않고 손쉽게 이용할 수 있다.



Fig. 10 키오스크 실행 결과

## 6. 결론

### 6.1 구현결과



Fig. 11 키오스크 구현 결과

본 시스템은 그림 (11)의 (a)-(b)-(c)-(d)-(e)-(f)-(g) 순으로 실행된다. 처음 키오스크 카테고리 선택(a부터 d 과정)은 그림을 그리는 미디어 아트 기술이 적용되었고, 카테고리가 선택되고 나면 자신이 원하는 물건을 선택하는 과정(e 과정 이후)에서는 그림 (12)와 같은 제스처 인식 기술이 적용된다.

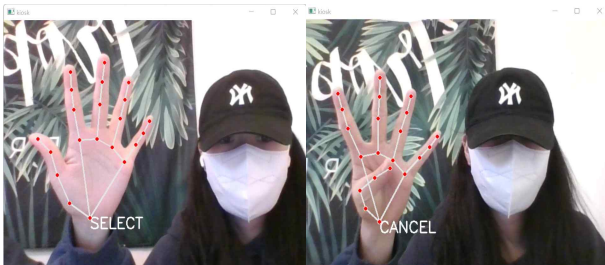


Fig. 12 제스처 인식 예시

## 6.2 기대효과

본 논문에서는 MediaPipe와 핸드 제스처를 이용한 비접촉 키오스크 시스템을 구현해보았다. 비접촉식 키오스크를 제작함으로써 사용자가 화면을 접촉하지 않아 감염자의 비말에 의한 코로나 19 감염병을 예방하고 위생적인 문제를 해결했다. 또한 총 6가지 손동작의 학습 데이터를 늘림으로써 빠르고 높은 인식률을 보였다. 하지만 단순히 빠르고 정확도에 초점을 둔 키오스크가 아니라 비접촉 기술로 제스처 인식이 이어 그림을 그리는 방법을 더해 하나의 미디어 아트 키오스크로써 더 발전하여 다양한 곳에 사용될 수 있을 것으로 예상된다.

첫째, 미디어 아트 기술은 아이들의 교육에도 효과적으로 이용할 정도로 누구나 쉽게 접할 수 있다. 예를 들어, 주어진 제시어를 차례대로 그림을 이어 그려 아이들의 창의력을 증진시킬 수 있다. 제스처 인식과 미디어 아트까지 활용해 화려한 화면 효과를 적용한다면 교육적인 키오스크가 가능할 것이다.

둘째, 대형 마트에 키오스크를 추가하여 마트 이용자들에게 편리함을 제공한다. 현재 대형 마트에서 물건의 위치를 찾을 때 직원을 찾아 직접 물어보는 방식이 만연하다. 이 방식은 매우 비효율적이며 코로나19 감염

병이 재유행하는 시기에는 더 비효율적이고 위생상으로 불안함을 일으킨다. 사람과의 접촉 없이 어린이나 노인분들 모두 편리하게 물건을 찾는 효과를 기대할 수 있다.

## References

- [1] 키오스크를 이용한 장애인 및 약자를 위한 음성인식 주문시스템/ Ordering system for the disabled and the weak using a KIOSK with speech recognition technology, 2021
- [2] Mediapipe를 활용한 핸드트래킹 키오스크 시스템/Hand Tracking Kiosk System Using Mediapipe, 2021
- [3] 가위바위보 기계 만들기 - 손가락 인식 인공지능/ 빵형의 개발도상국 유튜브
- [4] CoronaBoard 사이트- 국내 코로나 확진자 그래프