

AI 를 활용한 대화형 음성인식 키오스크

한지혜¹, 김상현², 김수민¹, 이해원¹, 임길환², 이경용³

¹영남대학교 정보통신공학과 학부생

²영남대학교 통계학과 학부생

³동의대학교 응용소프트웨어 겸임부교수

gkswlgp0331@naver.com, ddanghyni@gmail.com, kimsumin411@gmail.com, hooho931@gmail.com, dtc1128@naver.com, ceo@apptools.co.kr

Interactive Voice Recognition Kiosk Using AI

Ji-Hye Han¹, Sang-Hyun Kim², Su-Min Kim¹, Hye-Won Lee¹, Gil-Hwan Lim², Gyeong-Yong Lee³

¹Dept. of Information & Communication Engineering, Yeung-Nam University

²Dept. of Statistics, Yeung-Nam University

³Dept. of Applied Software Engineering, Dong-Eui University

요 약

고령 사회로 접어들고 디지털 문맹이 대두되면서 기계 조작에 익숙한 정보취약계층이나 고령층들이 느끼는 불편함을 개선하고자 본 연구는 기존의 터치형 키오스크와는 차별된 음성/얼굴인식 기술을 활용한 키오스크 개발을 제안하며, 모든 연령층을 고려한 AI 기술의 활용으로 맞춤형 주문 서비스를 제공하여 고객 만족도를 높이고, 보다 쉽고 빠르게 주문할 수 있도록 한다.

1. 서론

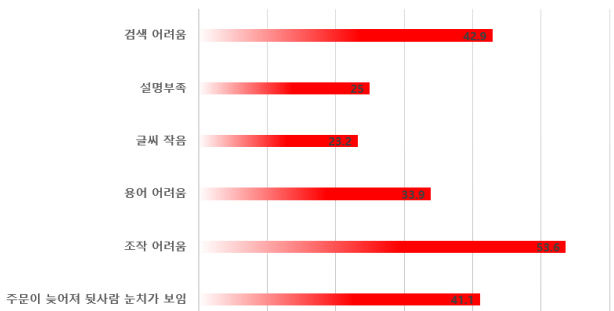
1.1 개발 배경 및 필요성

노인들의 키오스크 포비아가 증가하고 있는 현 시점에서, 고령 사회로 접어들고 디지털 문맹이 대두되면서 이러한 문제를 해결하기 위해 지자체에서 노인 대상 키오스크 활용 교육을 시작하였지만 노인들이 겪는 키오스크 포비아 문제를 해결하지는 못했다[1]. 이러한 상황이 지속됨으로써 디지털 소외 계층이 증가하게 되어 사회에서 배제되는 결과를 초래할 수 있다. 이러한 문제점을 착안해 기존의 키오스크에서 기계 조작에 익숙한 정보취약계층이나 고령층이 느끼는 불편함을 개선하고자 AI 를 활용한 대화형 음성인식 키오스크 프로그램을 개발할 필요성을 느꼈다.

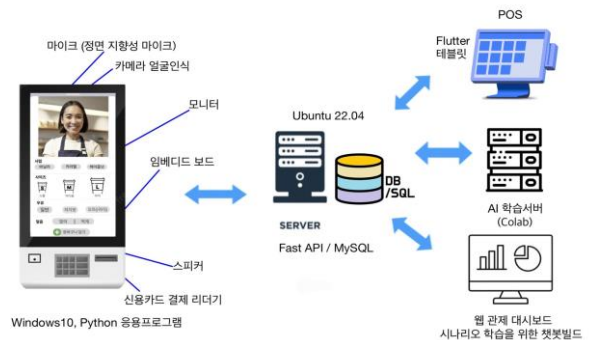
1.2 기존 키오스크의 문제점 보완

STT, TTS 음성인식 기술을 사용한 대화형식의 주문을 통한 AI 키오스크의 사용으로 기계를 이용한 주문으로 느껴지는 피로감을 줄일 수 있으며 키오스크에 대한 거부감을 줄일 수 있다. 또한 실시간 객체감지를 이용한 영상인식을 접목하여 나이, 성별, 감정 분석을 통해 연령대와 성별을 구분하며 고객의 유형을 판단 후 차별화된 메뉴를 추천하고, 감정 분석 결과에 따른 고객 맞춤 대화로 전 연령층에 대해 유연한 대응이 가능하다.

2. 인공지능 기반 키오스크 설계



<표 1> 60대 이상 키오스크 이용 중 경험한 불편 사례



<그림 1> 프로젝트 구성도

주문을 원하는 고객이 키오스크 앞에 섰을 때, 얼굴인식을 통해 객체를 감지한다. 학습된 데이터를 통해 나이와 성별을 인식 후 키오스크가 주문을 시작하는

대화를 진행하며, 상황에 맞는 화면과 대화를 제공한다. DB 에 저장된 데이터(메뉴 정보, 옵션 정보)를 나이, 성별, 감정분석 한 결과에 따라 개인 맞춤형하여 키오스크 화면에 보여주고, 주문을 진행하는 동안 생성되는 데이터를 저장한다. 주문이 완료된 즉시 주문 내역을 Server 를 통해 DB 에 저장하고, POS 에서 DB 에 저장된 주문내역 데이터를 통해 주방에서 주문 내역을 확인할 수 있다. 또한 웹 관제 대시보드를 이용하여 판매내역을 시각화한 데이터를 확인할 수 있다.

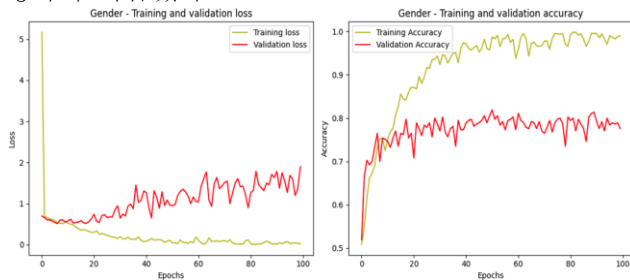
3. 인공지능 기반 키오스크 구현

3.1 얼굴인식

고객이 키오스크 앞에 섰을 때 얼굴이 인식이 되면 3 초 길이의 영상이 촬영된다. 그 영상을 토대로 나이, 성별, 감정 분석을 진행하며 분석한 결과의 평균값을 도출한다. 결과의 평균값을 바탕으로 개인 맞춤형 주문이 시작되며, 메뉴를 추천해준다. 실시간 소통을 위해 고객의 말이 시작될 때마다 위와 같은 과정으로 감정 분석이 이루어진다.

1) 영상에서 얼굴을 검출하기 위해 Haar Cascade 를 이용했다. 얼굴이 있는 이미지(Positive Image)와 얼굴이 없는 이미지(Negative Image)를 학습시켜 사진의 특징을 이용한 Machine Learning 으로 사진의 픽셀들의 밝기 차를 이용하여 얼굴을 검출한다[2].

2) 나이, 성별, 감정분석에 사용되는 데이터는 48x48 크기의 약 3 만장의 Dataset 을 Python 의 Keras 라이브러리를 이용하여 Google Colab 에서 학습시킨 후 .h5 형식의 모델을 만들어 사용했다. [3]정확한 결과를 도출하기 위해 학습 이미지의 순서를 바꾸거나 Accuracy 와 Loss 를 비교하며 epoch 을 조절하는 등 다양한 시도를 했다. 그 결과 나이, 성별, 감정 분석 모델을 만들기 위해 epoch 을 100 으로 설정하여 학습했다.



<그림 2> 성별 모델의 Loss, Accuracy

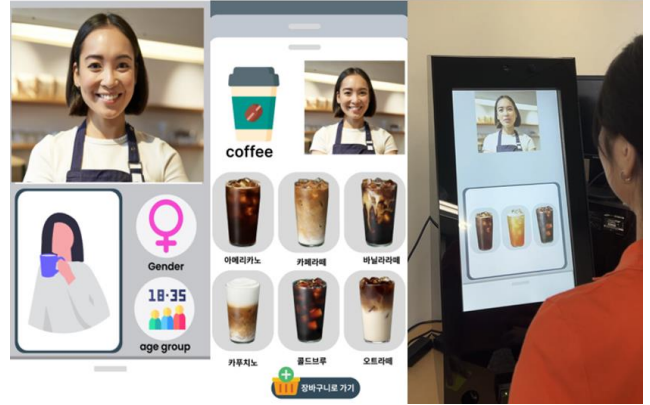
3.2 음성인식

기존 키오스크와 가장 큰 차이점 중의 하나로 터치가 아닌 음성으로 대화를 하며 주문이 가능하다는 점이다. STT 는 Flutter 에서 제공하는 라이브러리를 사용하여 기능을 구현했다. 시끄러운 매장 내부에서 키오스크와의 원활한 소통을 위해 지향성 마이크를 장착하여 주문하는 고객의 소리를 선택적으로 받을 수 있다.

3.3 키오스크 화면

Flutter 와 Dart 언어를 사용하여 키오스크 웹 화

면을 구성했다. 고령층의 편의를 위해 키오스크를 개발하고 있는 만큼 고령자의 시각, 청각, 지각 능력을 인지하여 글자크기, 색 배치, 전환 속도 등을 고려하며 UI/UX 를 구현하였다[4]. 또한 기계와 대화를 하며 느껴지는 거부감을 줄이기 위해 실제 사람의 이미지를 D-id studio 를 이용하여 자연스러운 표정을 구사하는 영상을 만들어 상단에 배치했다.



<그림 3> 키오스크 화면과 외형

4. 결론 및 기대효과

본 논문은 고령자들의 디지털 격차를 줄이기 위해 AI 를 활용한 대화형 음성인식 키오스크를 개발했다. Machine Learning 을 이용한 얼굴인식, 나이, 성별, 감정 분석을 통해 맞춤형 주문, 메뉴 추천이 가능하며 음성인식기능을 사용하여 실제 직원과 대화하듯이 키오스크를 사용하여 주문이 가능하다. 직원을 대체할 수 있는 키오스크를 사용함으로써 인건비 절감 효과를 기대할 수 있으며, 개인의 특성에 맞춘 주문을 통해 키오스크 사용에 대한 거부감을 줄일 수 있다. 또한 은행이나 공공기관과 같은 고객이 직접 처리 가능한 업무가 존재하는 분야에서 활용할 수 있을 것이라 기대한다.

* 본 논문은 과학기술정보통신부 정보통신창의인재 양성사업의 지원을 통해 수행한 ICT 멘토링 프로젝트 결과물입니다.

참고문헌

[1] 한국소비자원. (2022). 키오스크 이용실태조사 보도 자료

[2] J. Whitehill, C.W. Omlin, Haar features for FACES AU recognition, 7th International Conference on Automatic Face and Gesture Recognition, Southampton, UK, 2006

[3]김경태, 최재영, 단일 훈련 샘플만을 활용하는 준지도 학습 심층 도메인 적응 기반 얼굴인식 기술 개발, 한국멀티미디어학회논문지 제 25 권 제 10 호, 1,375-1,385(11p), 2022

[4]이설화, 임희석, 시니어 맞춤형 UI/UX 모델 개발을 위한 특징 요소, 한국컴퓨터교육학회 2017 년도 동계학술발표논문지 제 23 권 제 1 호, 87-89(3p), 2017