

FaceNet 기반의 얼굴인식 로그인을 하는 음성인식 키오스크

유래훈(인하대학교), 양현지(인하대학교), 박대영*(인하대학교)
sapoda55@naver.com, hj6440@naver.com, *dpark@inha.ac.kr

Speech recognition KIOSK using FaceNet based face recognition login

Raehoon Yoo (Inha Univ.), Hyunji Yang(Inha Univ.), *Daeyoung Park(Inha Univ.)

요약

본 논문은 기존 키오스크의 문제점을 개선하는 법을 제안한다. 기계를 잘 사용하지 못하는 사람들을 위해 음성인식을 지원하는 키오스크를 구현하고 얼굴인식을 통한 로그인과 이용자의 편리성을 위한 다양한 기능을 추가한다.

I. 서론

최근 키오스크를 사용하는 음식점, 카페가 점점 늘어나고 있는 추세이다. 하지만 기계를 잘 다루지 못하는 사람들의 경우 키오스크 사용에 어려움을 겪게 되고, 많은 주문 선택사항과 터치스크린의 한계 때문에 직원과 대면하여 주문하는 방법에 비해 주문하는데 더 오랜 시간이 걸린다.

따라서 터치스크린을 통해 주문하는 것이 아닌 직원과 대면하듯 음성으로 주문이 가능한 키오스크를 제안한다. 음성으로 주문함으로써 기존의 키오스크에서 주문하는데 걸리는 시간을 줄이고 기계를 잘 다루지 못하는 사람들도 키오스크를 쉽게 이용할 수 있게 해준다. 또한 얼굴인식을 통한 로그인을 통해 포인트 적립, 지난주문기록 조회 등 사용자 편의를 위한 기능을 사용할 수 있다.

II. 본론

1. 얼굴인식

그림 1은 얼굴인식 과정을 표현한 flow chart이다. 키오스크에는 회원가입 기능이 있는데, 회원가입시 10장의 사진을 찍게된다. 회원가입을 통해 얻은 고객들의 사진에서 MTCNN을 통해 얼굴만을 추출한다. 이후 FaceNet [1]을 통해 얼굴의 임베딩 데이터를 추출하고 이를 npz파일 형태로 저장한다. 이 임베딩 데이터셋을 얼굴인식에 사용한다.

키오스크의 웹캠을 통해 고객이 사진을 찍으면 마찬가지로 MTCNN을 통해 얼굴만을 추출하고 FaceNet을 통해 얼굴 임베딩을 추출한다. 그 후 이미 회원가입된 임베딩 데이터셋을 불러오고 SVM [2]을 통해 고객의 임베딩과 비교한다. 임베딩 데이터셋에 있는 임베딩들 중 고객의 임베딩과 일치 확률이 99% 이상인 임베딩이 있고 이 임베딩이 회원 A의 임베딩이라 가정하면 고객은 회원 A가 맞다고 판별하고 회원 A의 계정으로 로그인해 키오스크를 사용하게 된다. 일치확률이 99% 미만이라면 비회원으로

서 키오스크를 사용하게 된다.

2. 음성인식

Python에서 제공하는 speech recognition 패키지를 이용하여 speech to text를 진행했다. 그 후 KoNLPy [3] 패키지를 이용하여 text를 품사별로 나누어줬다. 그림 2의 첫 번째 줄은 “아이스 아메리카노 두 잔 주세요”라고 말했을 시 speech recognition 패키지를 통해 text로 변환된 것이고 두 번째 줄은 KoNLPy를 통해 품사별로 나눈 것이다.

```
아이스 아메리카노 두 잔 주세요  
['아이스', '아메리카노', '두', '잔', '주세요']
```

그림 2. speech recognition과 KoNLPy의 결과

KoNLPy가 text를 품사별로 나눠주면, 여기서 특정 키워드를 추출해 주문을 분석한다. 예를들어 그림 2를 보면 ‘아이스’라는 단어가 있으므로 아이스 음료를 주문한 것으로 분석한다. ‘아메리카노’라는 단어가 있으므로 아메리카노 음료를 주문한 것으로 분석한다. ‘두’라는 단어가 있으므로 2잔을 주문한 것으로 분석한다. 이렇게 분석한 결과를 GUI의 주문창을 통해 확인시켜주고 주문을 진행한다.

3. GUI

Python의 PyQt5를 통해 GUI를 구현했다. 그림 3은 웹캠을 통한 얼굴인식 후 나오는 기본창이다. 기본화면의 상단에는 회원의 이름이 나와있고 적립된 포인트와 지난 주문기록을 보여준다. ‘지난 주문기록 불러오기’ 버튼을 통해 주문시 음성인식 과정을 생략하고 바로 지난 주문기록 그대로 주문이 가능하도록 해주었는데, 이는 사람들이 주로 선호하는 음료만 마시는 습성을 이용해 주문하는데 편리성을 더하기 위해 넣은 기능이다.

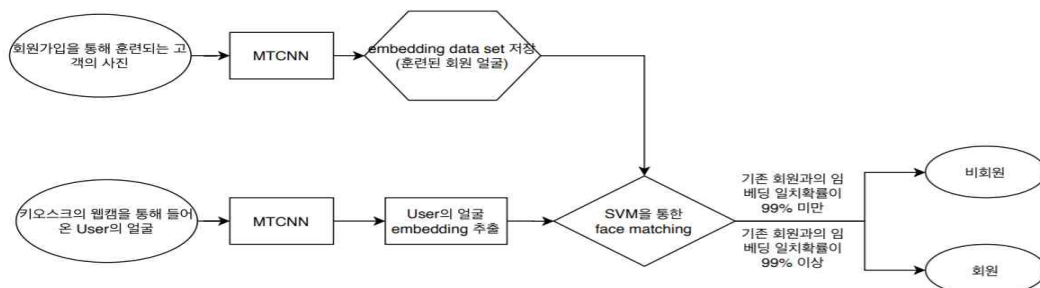


그림 1. 얼굴인식 flow chart

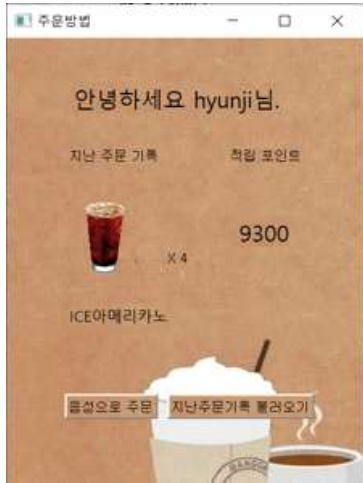


그림 3. 기본화면

그림 4는 음성으로 주문 후 나오는 주문창이다. 상단에는 음성으로 주문한 내용을 나타내준다. 음성으로 주문한 내용을 분석해 주문한 음료를 보여주고, 수량 및 HOT/ICE 변경, 사이즈업이 가능하도록 해주었다. 또한 음료 단가와 수량, 사이즈업을 고려해 총 가격을 도출해주었다. 포인트로 주문하게 되면 보유하고 있는 포인트가 음료의 총 가격보다 클 시에만 보유 포인트에서 총 가격만큼 차감이 된다.

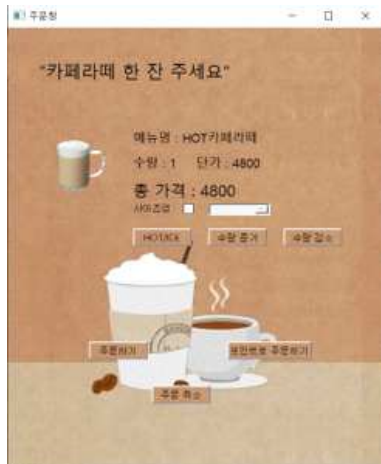


그림 4. 주문창

III. 결론

본 논문에서는 음성인식으로 주문함으로써 주문 시간을 단축시키고 사용하기 간편하도록 키오스크의 개선방안을 제안했다. 뿐만 아니라 얼굴인식을 통한 로그인을 통해 사용자 편의를 위한 여러 기능도 넣을 수 있도록 했다. MTCNN, FaceNet, SVM을 통해 얼굴인식을 진행하고, STT(Speech to Text), KoNLPy를 통한 자연어처리를 통해 음성인식을 진행했다. 이와 같은 기술들을 GUI를 통해 데모 환경을 구축 및 시연함으로써 구현한 프로젝트를 입증하고 상품화할 수 있도록 하였다.

ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.2019R1A2C1005512)

참 고 문 헌

- [1] Florian Schroff, Dmitry Kalenichenko, James Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering", IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp.815-823, 2015
- [2] M Pal, GM Foody, "Feature selection for classification of hyperspectral data by SVM", IEEE Transactions on Geoscience and Remote Sensing, 48, pp.2297-2307, 2010
- [3] 박은정, 조성준. "KoNLPy: 쉽고 간결한 한국어 정보처리 파이썬 패키지", 제26회 한글 및 한국어 정보처리 학술대회 논문집, pp. 133-136, 2014.