

논문 2019-01(수정할 필요 없음)

SORI ; 나만의 만능 비서

(System with Omniwheel & Recognition Interface)

금나연*

(Nayoun Keum)

요 약

5G와 AI의 시대가 도래하며 IoT 산업의 발전 가능성이 새롭게 주목을 받고 있다. 나만의 만능 비서 SORI는 점점 늘어가는 홈 IoT 사용자의 수요와 변화하는 시대에 발맞추고자 개발되었다. SORI는 사용자를 인식하고 그에 맞춤 서비스를 제공한다. 자유로운 이동성과 물리적 도움을 주기 위해 Ohmniwheel 모터를 사용하고 사용자와 물건을 인식하기 위해 딥러닝을 적용하였다. 얼굴인식과 사물인식을 통해 목적지를 설정할 수 있도록 하여 음성인식에 불편함을 느끼는 사람들에게 새로운 선택지를 제공하도록 모델링하였다

Abstract

With the advent of the era of 5G and AI, the potential for development of the IoT industry is receiving renewed attention. My own all-around secretary SORI has been developed to keep pace with the growing demand of home IoT users and changing times. SORI recognizes users and provides customized services. To provide free mobility and physical assistance, Ohmniwheel motors were used and deep learning was applied to recognize users and objects. New options were modeled to provide people who felt uncomfortable with voice recognition by allowing them to set their destination through facial and object recognition.

Keywords: IoT, Robot, Deeplearning, Face Recognition, Object Recognition

I. 서 론

최근 5년간의 통계조사를 통해 IoT 시장이 빠르게 성장하는 사실을 확인할 수 있다. 발전된 형태로 다양한 브랜드에서 출시되는 홈 IoT들의 수요는 높아졌다. 현재 AI speaker의 경우 “1가정 1 AIspeaker” 라는 말이 유행 할만큼 수요가 많아졌지만 TV remote controller 또는 mp3 player와 같은 한정된 서비스 제공

* 금나연, 이화여자대학교

과 물리적인 서비스가 부재하다는 단점이 있다 생각되었다.

이와 같은 문제점을 인식하여 새로운 홈 IoT로서 사용자에게 맞춤형 서비스와 실질적인 물리적인 도움을 줄 수 있는 SORI를 모델링하게 되었다.

만능 비서 SORI는 Ohmniwheel을 이용하여 자유로운 이동성을 Deep Learning을 통한 얼굴 인식 사물인식 기능성을 제공한다. 가정용 비서 외 에도 독거노인 또는 혼자집에 남겨진 반려견/반려묘를 모니터링하는 돌보미의 역할 등 그 활용성이 뛰어난 것이 특징이다.

© Corresponding Author(E-mail: 51209tuna@gmail.com)
접수일자: 2019년 12월 21일

II. 본 론

1. SORI의 Algorithm

SORI ; 나만의 만능 비서는 크게 사용자에게 필요한 물건을 갖다 주거나, 지정한 사람에게 입력한 물건을 갖다 주는 기능을 제공한다. 본 로봇이 제공하는 알고리즘은 다음과 같다.

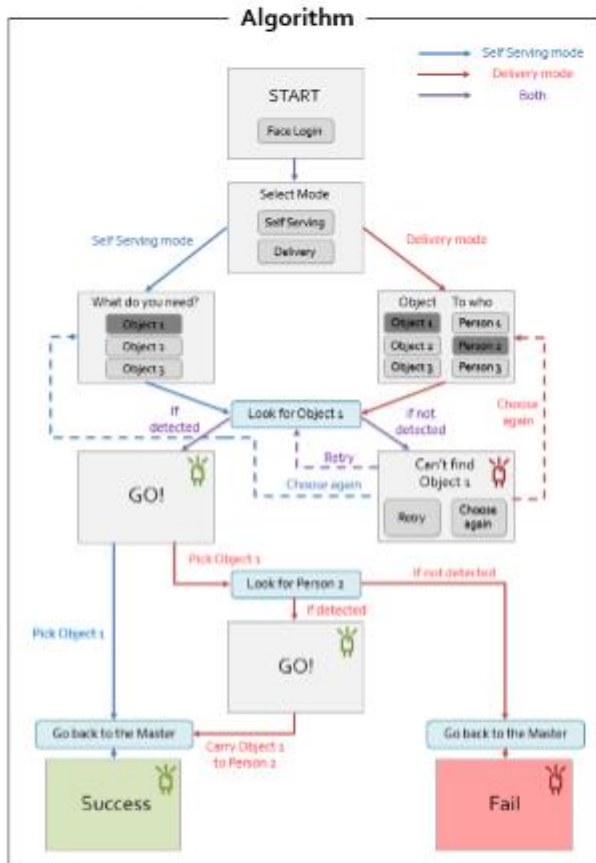


그림 1. SORI 알고리즘

Fig. 1. Algorithm of SORI

가. Face Login 사용자 등록

(1) 사용자 등록 Register the User

SORI를 사용하기에 앞서 사용자 등록을 거친다. 몇 장의 사진을 찍어 사용자 등록을 마치면 SORI는 사용자를 파악한다. 기존의 딥러닝 모델을 사용하기 위해서는 GPU를 통해 학습을 시켜야했다. 이 점은 모델을 사용하기 전, 모든 사용자를 예상하고, 학습을 시켜야 한다는 점에서 딥러닝 기술의 상용화를 방해하는 문제 중 하나였다. 본 모델은 딥러닝 모델을 사용하면서도 사용자가 쉽게 학습을 시킬 수 있는 방법을 사용하며 본문 2장 기술구현 과정단계에서 그 방법을 찾아볼 수 있다.

(2) Face Login

사용자 등록을 마쳤다면 Start 화면에서 이름을 선택하여 Select Mode로 넘어갈 수 있다.

나. Select Mode

(1) Self Serving

사용자가 선택할 수 있는 기능은 2가지이며 첫 번째 기능은 Mode 1인 Self Serving 모드이다. 화면에서 Self-Serving을 클릭하면 사용자는 필요한 물건을 선택하게 된다. SORI는 선택된 물건을 Object detection을 이용해 찾아 사용자에게 돌아온다. 이때 Face detection을 사용하므로 사용자의 위치가 변해도 다른 사람이 존재해도 SORI는 사용자를 정확하게 찾아 Self Serving기능을 수행하여 임무를 맞

(2) Delivery

Mode 2 Delivery 모드에서는 전달 받을 사람과 전달할 물건을 입력한다. SORI는 먼저 전달할 물건을 object detection을 사용해 찾는다. 그 후, Face detection을 사용해 전달 받을 사람을 찾아 물건을 전달하여 임무를 마치고 Start 화면으로 돌아가게 된다.

2. 구현기술 및 방법 IMPLEMENTATION

가. 구성 부품

(1) 라즈베리파이 Raspberry pi

Raspberry pi 3 와 4를 사용하였으며 리눅스 기반의 개발환경을 경험 할 수 있었다. 저가형 개발 보드이기에 가격 부분에 있어 다른 IoT 기기들과의 가격 경쟁력에 있어 우세한 강점을 보인다.



그림 2 Raspberry pi 3 board

Fig. 2. Raspberry pi 3 board

(2) 파이 카메라 V2 (Pi camera module V2, 8MP)

라즈베리파이 보드와 연결되는 pi camera를 통해 사진과 영상 데이터를 다룰 수 있게 하였다. SORI의 주된 기능인 얼굴인식과 사물인식을 가능하게 한 카메라이다.



그림 2. Pi-Camera V2

Fig. 2. Pi-Camera V2

(3) TB6612FnG Motor driver module

TB6612FnG 듀얼 모터 드라이버 모듈은 프로세서의 작은 전압을 제어하여 모터의 속도를 제어하고, 모터의 방향을 바꿀 수 있는 회로 장치이다. DC 모터 2개 또는 스텝모터를 2개를 제어할 수 있으며 Step Down mode, 입력 전압 25V~13.5V, 출력 전류 3A의 사양을 가지고 있다.



그림

3. TB6612FnG 모터 드라이버 모듈

Fig. 3. TB6612FnG motor driver module

(4) OhmniWheel Platform

메카닉 휠과 유사한 옴니 휠은 회전 방향에 수직인 원주 주위에 작은 롤러가 있어 휠을 최대의 힘으로 구동하는 동시에 측면으로도 쉽게 미끄러지는 특징을 가지고 있다. 원하는 방향이 어디든지 자연스럽게 이동할 수 있다는 장점이 있어 이용하게 되었다.

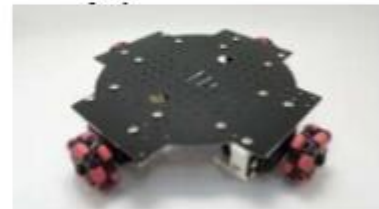


그림 4. TB6612FnG 모터 드라이버 모듈

Fig. 4. TB6612FnG motor driver module

나. Open CV



그림 5. Open CV 로고

Fig. 5. Open CV Logo

(1) 가상환경 Virtual Environment

Python이 버전에 따라 지원되는 서비스가 달라 OpenCV를 설치하는데에 거듭된 실패를 경험한 끝에 Virtual Environment를 사용하게 되었다. Python에서 지원하는 가상환경인 pyenv를 사용하여 임의의 디렉토리 밑에 Python version 3.7 + Open CV version 4.1.0을 설치하여 버전간의 충돌을 해결할 수 있었다.

(2) Deep Learning Frame Work

얼굴인식 또는 사물인식을 위해서는 학습 데이터를 학습시키는 Deep learning 과정이 필요하다. Open CV는 Tensor Flow, Py Torch와 같은 딥러닝 프레임 워크를 지원하기에 실시간 컴퓨터 비전을 목표로하는 프로그래밍 기능 라이브러리인 Open CV를 라즈베리파이 보드에 설치하였다.

다. 얼굴인식 (Face detection & Face Recognition)

(1) 얼굴 탐색 Face Detection

이미지 또는 영상에서 사람의 얼굴을 찾아 네모 칸으로 표시해주는 Face Detection을 시작으로 그 네모칸에 들어온 사람의 얼굴이 어떤 사람인지 또 얼마나 그사람과 일치한지를 퍼센트를 띄우는 Face Recognition을 구현하였다. Face Detection에서는 SSD(Single Shot Detector)모델을 사용하였다.

(2) 얼굴인식 Face Recognition

얼굴을 인식하기 위해서는 얼굴의 위치를 찾는 Face detection 감지한 얼굴을 자르고 정방향으로 바꿔주는 Face Alignment 찾은 얼굴이 누구의 얼굴인지 판별하는 Face Recognition단계를 거치게 된다.

개발 초기에는 Deep Learning에 쓰일 학습 데이터를 인터넷에서 다운받거나 따로 보드에 넣어주어 사용했으나 개발 마무리단계에서는 라즈베리파이 보드에 연결되어있는 Pi camera를 이용하여 찍은 20장 정도의 학습 데이터를 통해 Face Recognition을 가능하게 하였다. Face Recognition 단계에서는 FaceNet 모델을 사용하였다

(3) 딥러닝 학습 모델 (Deep Learning)

SORI가 이렇게 적은 학습 데이터로 실시간 영상에서 Face Recognition이 가능하게 된 이유는 Face Detection 과정에서의 SSD(Single Shot Detector)와 Resnet Backbone의 조합 그리고 Face Recognition단계에서의 FaceNet 모델 덕이었다. SSD 논문¹에서 이용한 VGG16 backbone보다 정확도 측면에서 성능이 더 좋아 Resnet을 이용하게 되었다. Face Net은 embedding을 통해 학습이 된다는 점을 이용하여 face embedding을 추출해 레이블 당 10-20장 정도의 적은 데이터를 단시간 내에 본 모델에서 학습할 수 있도록 구현하였습니다. 이러한 과정을 라즈베리파이 보드에서는 너무 무거워 돌릴 수 없었고 이 과정을 통해 학습된 모델을 SORI에 넣어 사용하여 실시간으로 얼굴 인식이 가능하도록 하였다.

라. 사물인식 Object Detection

(1) 사물인식 (Object Detection)

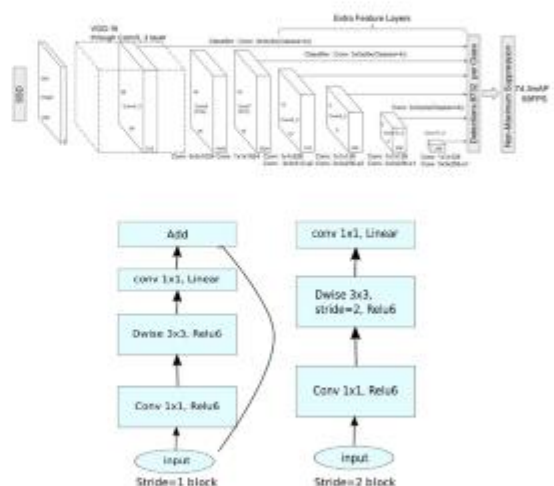
SORI의 또 다른 중요 기술은 문체인식이다. 물체인식에서 물체의 위치를 파악하는 Object Detection 파악된 물체가 무엇인지를 판별하는 Object Recognition 단계를 거치게 된다. 얼굴 인식과 마찬가지로 Object Detection을 위해 사용한 모델은 SSD(Single Shot Detector) model이다.

(2) 딥러닝 학습 모델 (Deep Learning)

SSD 이외에도 Fast R-cnn, YOLO의 선택지도 있었으나 Object Detection에서도 SSD 모델을 사용하였다. 그 이유는 Detection 부분에서 Fater R-cnn은 느리고 YOLO는 빠르지만 정확성이 떨어진다는 점을 감안하여 정확성과 속도 모두를 원하는 기준 이상으로 충족 시키는 SSD모델을 사용하였다.

Object Recognition 과정에서는 MobilNets을 사용하였는데 이는 핸드폰과 같은 디바이스에 커다란 딥러닝 네트워크를 실행시키기 위해 네트워크 크기를 줄인 것과 속도가 빠르다는 이점이 있어 사용하게 되었다.

Dataset은 실생활에서 사용하는 물건들이 대부분 들어있는 COCO dataset을 사용하였다. 이 또한 너무 크기가 크기에 라즈베리 파이에서는 돌릴 수 없었고, pre-trained된 weight만을 들고와서 사용하였다. 또한 조금 더 빠른 속도로 실행 되도록 multiprocessing으로 진행되도록 하여 기대이상으로 빠르게 사물을 인식할 수 있었다.



마. SORI의 하드웨어 Hardware

(1) Omni-Directional Wheels

메카넘 휠과 유사한 옴니 휠은 회전 방향에 수직인 원주 주위에 작은 물러가 있어 휠을 최대의 힘으로 구동하는 동시에 측면으로도 쉽게 미끄러지는 특징을 가지고 있다. 원하는 방향이 어디든지 자연스럽게 이동할 수 있다는 장점이 있어 이용하게 되었다.



그림 7. TB6612FnG 모터 드라이버 모듈
Fig. 7. TB6612FnG motor driver module

(2) Ultrasonic Sensor

초음파 센서는 SORI의 앞면에 부착하여, 얼굴 인식과 사물인식을 바탕으로 대상과 SORI의 방향이 일직선이 된 이후부터는 초음파 센서로 대상과 SORI의 거리 값을 전달받아 적당한 위치에서 멈출 수 있도록 한다.

초기 아이디어는 Pi Camera로 거리 인식을 하려 하였으나 초음파 센서가 너무나도 정확하고 잘 동작하여 거리인식 기능은 초음파 센서에게 맡기게 되었다.



그림 8 초음파 센서
Fig. 8. Ultra Sonic Sensor

3. 아쉬운 점 및 성찰 Reflection

가. Graphic User Interface GUI

사용자가 컴퓨터와 정보를 교환할 때, 그래픽을 통해 작업할 수 있는 환경을 말하는 GUI. 리눅스 기반인 라즈베리파이보다 가시성이 좋기는 Window 처럼 마치 하나의 Application을 실행하듯 버튼을 눌러 모드 변경 입력값을 주고 싶어 개발을 시작하였다.

(1) PyQt5

Python에서 제공하는 PyQt5이지만 Anaconda를 설치하여 아나콘다 가상환경에서 따로 깔아 작업했고 라즈베리파이로 옮겨 실행하였다.

난이도가 높으나 GUI를 이쁘게 만들 수 있고 실제로 Application으로까지 만들 수 있는 도구였기에 사용하였다.

완성하여 알고리즘에서 볼 수 있는 페이지들을 기능에 맞게 만들었지만 원래 작성하고 있던 Face Recognition과정이 순차적으로 코드가 사이에 넣는 것이 어렵다고 인지하여 합치지 못하고 모드를 선택할 수 있는 함수를 추가하여 command창에서 키보드로 1또는 2와 같은 입력값을 주어 모드 변환 부분을 완성시키게 되었다.

(2) Tkinter

이 또한 PyQt5와 같은 GUI를 만들 수 있는 대표 도구이며 난이도는 PyQt5에 비해 낮으나 모양이 이쁘지 않아 사용하지 않으려 하였으나 프로젝트 종료 전에 GUI를 가늠하기 위해 사용하여 GUI를 완성하였으나 코드간의 호응 문제를 해결하지 못하여 완성된 SORI에는 탑재하지 못하였다.



그림 9 PyQ 5 로고

Fig. 9. PyQ 5 Logo

나. IoT 기기 간의 확장성

SORI가 IoT 기기 간의 확장성을 실현하기 위해서는 IoT 기기간의 연결 뿐만 아니라 IoT 기기 내에 Deep Learning Processor가 탑재 되어야 한다는 것을 알게 되었다. 처음부터 같은 브랜드 또는 회사에서 출시된 것이 아닌 IoT기기 간의 소통이 목표이기에 새로운 IoT 기기를 연동시에 IoT 스스로 Firm-Ware를 다운받아 학습 할 수 있는 시스템이 아직 구현되지 않았다.

음성인식의 장점을 가지고 있는 AI Speaker와 물리적인 Service의 장점을 가지고 있는 SORI가 만나 시너지를 만드는 것 처럼 IoT간의 확장성을 갖게 해주고 싶었던 초기 목표는 기술 부족으로 잠시 먼 미래의 목표로 미뤄져야했다.

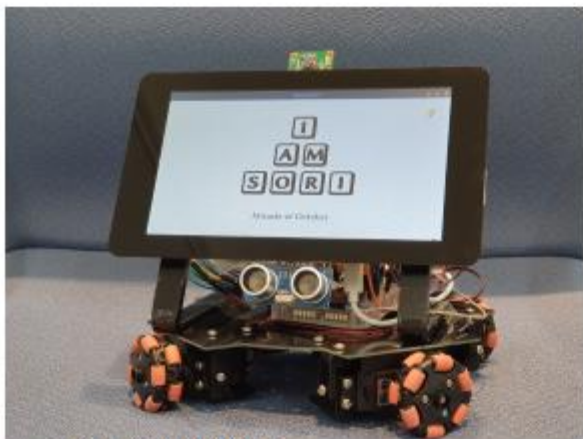


그림 9 SORI의 완성형
Fig. 9. SORI

III. 결 론

최근 5G와 AI의 시대가 도래하며, 편의성에 초점을 맞추어 보급되고 있던 현재의 홈IoT들과는 다르게, SORI는 기존의 IoT 기기들 간의 통신과 학습 가능성 대한 고민 속에 탄생한 로봇이다. 현재 가장 대중화 되고 수요가 급증하고 있는 AI Speaker의 실질적인 불편함들을 개선하고 싶은 생각에 SORI를 개발하게 되었고 그 과정 속에서 IoT 시장의 가능성을 다시한번 엿보게 되었다.

SORI와 같이 얼굴인식과 더불어 물리적인 서비스를 제공하는 로봇은 단순히 물건을 가져다 주는 비서가 아닌 사회 문제를 해결하는 데에도 큰 도움이 될 수 있는 로봇이라 생각한다. 만약 SORI가 노양원에 배치된다면 거동이 불편한 노인계층이나 장애인에게 그들이 각각 필요로 하는 물건 또는 서비스를 맞춤으로 제공이 가능하다. 또한 사회인들의 바쁜 일상 때문에 돌보지 못하는 반려묘 또는 반려견들의 우울증 예방 또는 치료에도 많은 도움이 될것이라 예상된다. 또한 병원에서도 필요 입원한 환자의 맞춤 식사나 약물 제공등의 서비스를 제공하는 등 너무나도 바쁜 간호사들의 손발이 되는 등 SORI는 상상 이상의 다양한 공간에서 다양한 목적으로 활용 될 수 있다.

현재는 기술 부족으로 인해 Deep Learning Processor가 탑재된 IoT가 대중화 되지 않아 SORI ver2가 나온다면 AI Speaker와의 연동이 아닌, 먼저 자연어 처리를 담당하는 기술을 SORI에 구현하게 될 것이라 생각이 든다. 하지만 만약 미래에 Deep learning processor가 탑재된 IoT 기기가 보편화 된다면 서로의 장점만을 극대화 시키며 서로 확장해 나가는 미래도 쉽게 그릴 수가 있다. 그 미래에는 확장성을 갖춘 기기들이 시장을 지배하며 이미 존재하고 있는 IoT기기의 활용도를 200%로 끌어올릴 것으로 예상된다.

REFERENCES

- [1] [Schroff et al. FaceNet: A Unified Embedding for Face Recognition and Clustering. In: CVPR(2015)
- [2] W. Liu, . Anguelov, D. Erhan, C. Szegedy, S. Reed, "SSD: Single shot multibox detector", 2015
- [3] A.Howard. Mobilenets: Efficient convolutional neural networks for mobile vision applications. Forthcoming
- [4] Joseph Redmond, You Only Look Once: Unified, Real-Time Object Detection
- [5] S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. In NIPS,2015
- [6] H.P. Oliveira, A.J. Sousa, A.P. Moreira, P.J.C. Gomes da Costa: "Dynamical Models for OmniDirectional Robots with 3 and 4 Wheels", 2008

저 자 소 개



금나연

2020 이화여대 전자공학과 학사 졸업

업

<주관심분야 : 컴퓨터 아키텍처,
SoC, 반도체, Deep Learning>