

● 2019 국립 한경대학교 컴퓨터공학과 졸업전시회

제 17회 졸업전시회 작품집

올데이아이즈

< 김영남, 박수빈, 홍지수 >



2020학년도 캡스톤디자인 최종보고서

올데이아이즈

수행분야	졸업작품			
지도교수	성 명	김숙연 교수	확 인	
팀 원	학 번	성 명		역할/주요업무
	2014250010	김영남		조원/개발
	2018250061	박수빈		조원/개발
	2019250074	홍지수		조장/개발

한경대학교 컴퓨터응용수학부 소프트웨어&서비스컴퓨팅전공

목 차

1. 서 론

- 1.1. 개발 배경(주제 선정 이유)
- 1.2. 작품 개요
- 1.3. 선행 연구 및 작품관련 자료조사
- 1.4. 작품의 목표
- 1.5. 작품의 특징 및 기대효과
- 1.6. 팀원의 구성 및 역할 분담

2. 작품의 분석/설계

- 2.1. 설계개요(개념설계, 시스템 구성도)
- 2.2. UI 설계
- 2.3. 시스템 상세설계
- 2.4. 시스템 기능별 요약

3. 결론 및 향후과제

- 3.1. 작품보완점 및 목표구현 정도
- 3.2. 개발 환경 및 도구
- 3.3. 향후 발전 방향
- 3.4. 문제점 및 개선 방안

4. 사용자 매뉴얼

부 록

- R-1. 참고문헌 및 참고사이트

1. 서 론

1.1. 개발 배경(주제 선정 이유)

IT 기술을 사용하여 사회적 약자에게 도움을 줄 수 있는 것을 만들어 보는 건 어떨까, 하는 의견에서 시각장애인의 눈이 되어주는 ‘올데이아이즈’ 개발이 시작되었다. 시각장애인이 사용하는 보조기구들은 일부 국가에서만 사용할 수 있다거나 단가가 높아 좋은 기술이 탑재되어 있더라도 구매장벽이 높다는 사실을 알게 되었다. 그래서 국내 시각장애인이 좀 더 가벼운 마음으로 사용할 수 있는 안경 형태의 시각 보조기구를 만들어 보기로 했다.

1.2. 작품 개요

올데이아이즈는 크게 두 가지 기능이 탑재되어 있다. 첫째로 장애물 알림 및 거리안내 기능이다. 사용자 전방에 장애물이 있다면 그것을 감지하여 어떤 장애물인지 종류와 함께 장애물까지의 거리도 함께 알려준다. 두 번째로 글자 읽어주는 기능이다. 특정 손 모양을 취하면 글자 읽어주는 기능으로 전환되고, 인식된 텍스트를 사용자에게 소리로 들려준다.

1.3. 선행 연구 및 작품관련 자료조사

가) 유사시스템 분석

이름	주요기능 및 장점	올데이아이즈와의 차이점
아이라 호라이즌	<ul style="list-style-type: none"> - 120도 광각 카메라가 설치되어 넓은 시야를 카메라에 담을 수 있다. - 상담원을 연결하여 주변 안내를 받는다. - AI 음성비서 서비스 클로이를 사용하여 이미지를 식별 후 글자를 읽어준다. 	<ul style="list-style-type: none"> - 상담원 연결 기능을 이용하려면 높은 가격의 월 정기결제 가 필요하다. - 상담원 서비스는 한국을 지원하지 않는다.
위워크	<ul style="list-style-type: none"> - 초음파센서를 이용해 길거리의 장애물을 감지한다. - 핸들 양쪽에 있는 진동 모터를 통해 사용자에게 위험요소를 알려준다. 	<ul style="list-style-type: none"> - 위험요소만 알려줄 뿐 장애물의 종류는 알려주지 않는다.
서울시 엔젤아이즈	<ul style="list-style-type: none"> - 스마트폰을 통해 실시간 영상을 특정인에게 전송하여 도움 요청 후 안내를 받는다. 	<ul style="list-style-type: none"> - 도움을 요청할 지인이 없거나 상담원 연결이 안 되면 안내를 받기 어렵다.
오어캠	<ul style="list-style-type: none"> - 손으로 특정 대상을 가리키면 관련 정보를 음성으로 들려준다. 	<ul style="list-style-type: none"> - 손으로 가리킨 대상의 정보만 알려준다.

나) 작품에 사용될 기술 조사

1) YOLO(You Only Look Once)

YOLOv3 실시간 물체 감지 : You Only Look Once의 약자로써, 물체인식(Object Detection)을 수행하기 위해 고안된 심층 신경망이다. 테두리상자 조정(Bounding Box Coordinate)과 분류(Classification)를 동일 신경망 구조를 통해 동시에 실행하는 통합인식(Unified Detection)을 구현하는 것이 가장 큰 특징. 이미지를 한 번 보는 것으로 object의 종류와 위치를 추측하고 단일 네트워크를 통해 여러 개의 테두리 상자(Bounding Box)에 대한 객체 확률을 계산한다.

2) 텍스트-음성 변환 기술(TTS)

Google Cloud Text-to-Speech는 텍스트를 30여 개의 언어 및 방언이 지원되는 180여 개의 자연스러운 음성으로 변환해 준다. 음성 합성(WaveNet)의 획기적인 연구 성과와 Google의 강력한 신경망을 적용하여 최상의 음질을 제공한다. 사용이 간편한 API로 사용자와 실제 대화하는 듯한 상호작용을 제공하여 고객 서비스, 기기 상호작용, 기타 애플리케이션에 혁신적인 변화를 줄 수 있다.

3) OCR (optical character recognition)

보통 컴퓨터가 2진법(0/1) 데이터를 폰트를 통해 인간이 인식할 수 있는 형태로 글자를 보여 준다면, OCR은 그 반대로 인간이 종이 위에 써 놓은 글씨를 인지하여 텍스트 데이터로 치환한다. 보통은 스캐너로 읽어 들인 이미지 파일을 분석하여 텍스트나 워드 파일로 결과물을 내놓는다.

1.4. 작품의 목표

올 데이 아이즈의 가장 큰 목표는 ‘위험을 최소화’하는 것이다. 비장애인의 시각이 아닌 시각장애인의 시각에서 바라볼 때 그들은 수많은 위험의 사각지대에 놓여있다.

흐릿하게 보이는 장애물을 판별할 수 있게 되고, 의약품을 잘못 복용하는 위험을 최소화해 그들의 안전성을 보장해주는 것이다.

1.5. 작품의 특징 및 기대효과

가) 작품의 특징

기존의 YOLO에서 제공하는 weights파일이 아닌 시각장애인에게 위험한 장애물에 대한 이미지를 수집해 그들에게 특화된 weights파일을 생성했다.

사물의 자세한 텍스트 정보를 알고 싶을 때, 특정 손가락 모양을 카메라에 가져다대면 가리키는 사물의 텍스트 정보를 음성으로 전달해준다.

나) 기대효과

1. 평소 길을 걸을 때 제대로 알아볼 수 없는 사물의 정보를 알 수 있고, 거리 정보도 함께 제공하여 길을 걸을 때 마주하는 위험에 대한 부담을 줄여준다.

2. 시중에 판매하는 시각장애인 시각보조기구보다 저렴하고, 해외제품의 경우 영어로 전달하지만 한국인에 특화되어 외국 제품에 비해 사용하기 친숙하다.

3. 흐릿하게 보였던 글자를 뚜렷하게 읽을 수 있고, 의약품 등의 오용의 위험을 줄여 줄 수 있다.

1.6. 팀원의 구성 및 역할 분담

학번	이름	역할	참여도
2014250010	김영남	조원 (YOLO 코드 변경 및 추가, weights 파일 생성)	33%
2018250061	박수빈	조원 (작품 제작에 필요한 전반적인 프로그램 설치)	33%
2019250074	홍지수	조장 (학습 파일 수집 및 YOLO 자동 실행 구현)	34%

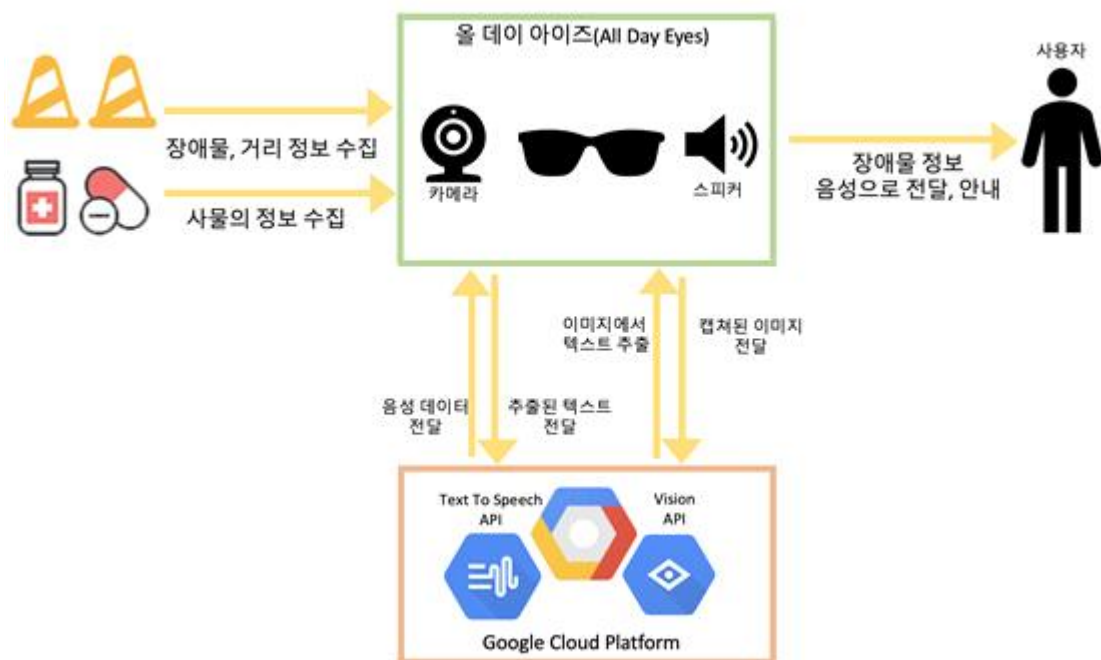
2. 작품의 분석/설계

2.1. 설계개요(개념설계, 시스템 구성도)

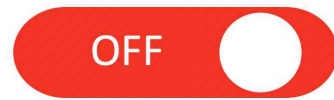
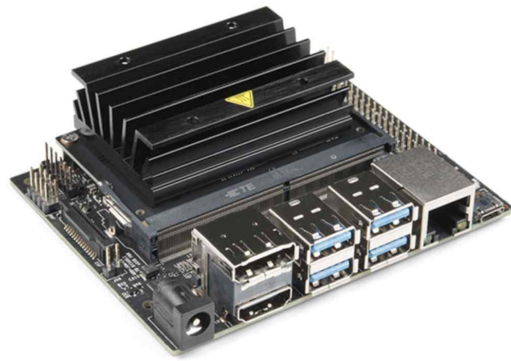
가) 개념 설계

1. 카메라에 들어온 영상 데이터를 활용해서 YOLOv3로 장애물의 종류와 거리 정보를 얻는다.
2. YOLOv3 사이트에서 제공하는 기존의 weights 파일이 아닌 시각장애인에게 특화된 장애물의 정보가 담긴 weights 파일을 직접 생성한다.
3. weights 파일에는 특정 손모양을 학습시켜서해 해당 손모양이 감지될 때 이미지가 캡처된다.
4. Vision API를 통해 이미지에서 텍스트 정보를 추출한다.
5. 장애물의 정보와 사물의 정보를 TTS(Text-To-Speech) API를 통해 사용자에게 음성으로 전달한다.

나) 시스템 구성도



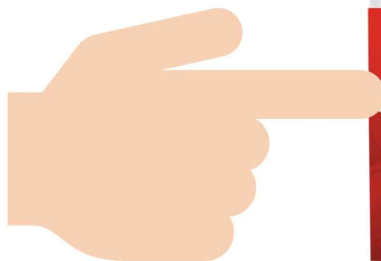
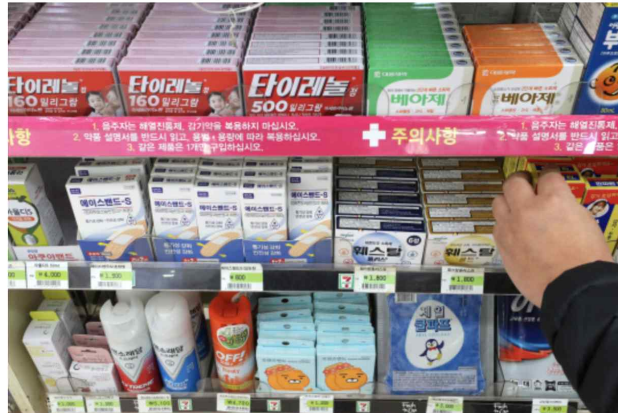
2.2. UI 설계



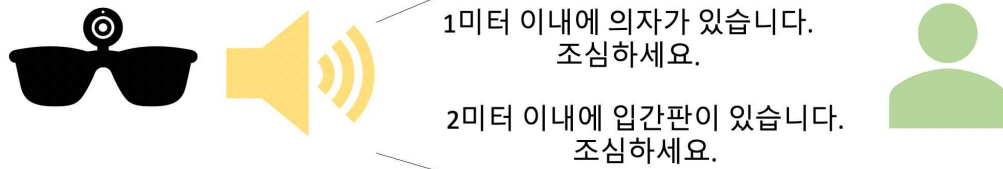
1. 사용자는 젯슨나노의 전원을 on / off 를 제어할 수 있다.



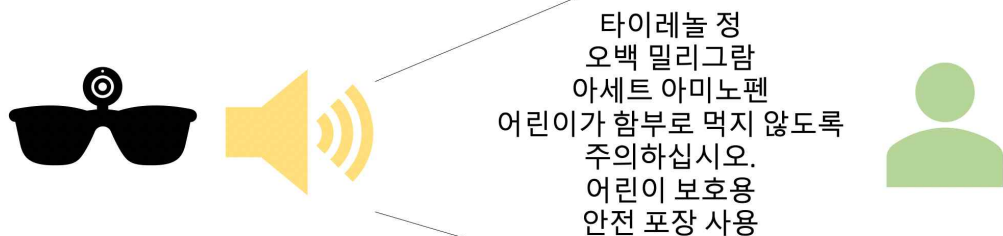
2-1. 안경에 부착된 카메라로 장애물들이 감지된다.



2-2. 안경에 부착된 카메라에 정보를 얻고자 하는 물체에 특정 손가락 모양(총 모양)을 가져온다.



3-1. 장애물의 정보를 스피커를 통해 사용자에게 음성으로 안내한다.



3-2. 특정 손모양(총모양)이 화면에 잡히면 사용자에게 물체에 대한 정보를 음성으로 안내한다

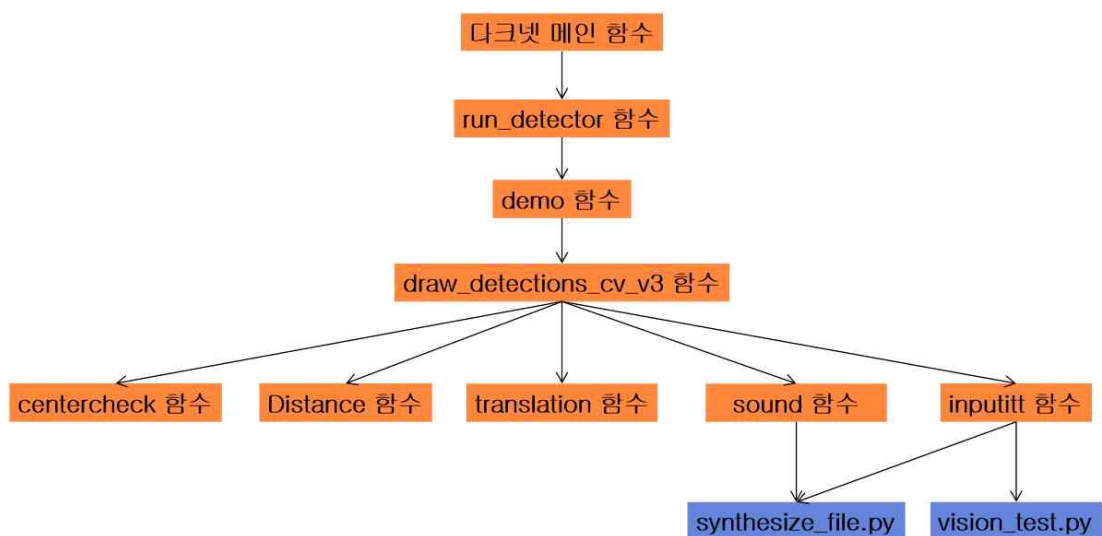
2.3. 시스템 상세설계

2.3.1 하드웨어 상세설계

2.3.1 하드웨어 상세설계

올데이아이즈 안경	사용자의 손발이 자유롭도록 착용하는 안경 형태로 선정
	사용자의 착용 편의성을 위해 부품의 무게를 최소화함
젯슨나노	<ul style="list-style-type: none"> - GPU: 128-코어 엔비디아 맥스웰(NVIDIA Maxwell™) 아키텍처 기반 GPU - CPU: 쿼드 코어(Quad-core) ARM®A57 - 비디오: 4K @ 30 fps (H.264/H.265) / 4K @ 60 fps (H.264/H.265) 인코드 및 디코드 - 카메라: MIPI CSI-2 DPHY 레인(lanes), 12x (모듈) and 1x (개발자 키트) - 메모리: 4 GB 64-비트LPDDR4; 25.6 초당 기가바이트 - 연결: 기가비트 이더넷(Gigabit Ethernet) - 운영체제 지원: 테그라(Tegra®)용 리눅스 - 모듈 크기: 70mm x 45mm - 개발자 키트 크기: 100mm x 80mm
	신속한 실시간 탐지를 위해 AI에 특화된 오픈소스 보드 사용
	발열의 최소화를 위해 별도의 쿨러 설치
	사용자에게 정확한 정보의 전달을 위해 고성능의 TTS 기술 사용
카메라	USB 연결 형태의 카메라 사용
	원활한 장애물 탐지와 거리안내를 위해 고성능의 카메라 사용
스피커	USB 사운드 카드와 귀걸이형 이어폰을 사용.
	전력 사용을 절감하기 위해 무선 대신 유선으로 선정

2.3.1 소프트웨어 상세설계



다크넷 메인 함수	다크넷을 실행할 때 넣어준 메인 함수 인자를 확인하여 잘못된 인자가 들어오면 올바른 사용법을 알림
	다크넷을 사용하는 하드웨어의 환경(CUDA, CUDnn 등)을 체크하고 잘못된 환경이면 잘못된 점을 알리고 다크넷 종료
	메인 함수 인자에 따라서 실행할 기능이 있는 함수를 호출(우리 팀은 run_detector 함수를 사용)
run_detector 함수	동영상 또는 웹캠을 사용한 실시간 분석에 필요한 설정값들을 디폴트 값 또는 사용자의 설정값으로 설정(웹캠 화면의 가로와 세로의 크기, 감지될 물체들의 정확도 하한값 등)
	다크넷을 사용할 때 설정값을 확인하여 연산을 GPU(그래픽카드)로 할지 아니면 CPU로 할지, GPU로 한다면 몇 번째 GPU로 연산할지를 설정
	메인 함수에서 받아온 인자에 따라서 실행할 기능이 있는 함수를 호출(우리 팀은 demo 함수를 사용)
demo 함수	인자값에 따라서 동영상 분석인지 웹캠 실시간 분석인지 판별(우리 팀은 웹캠 실시간 분석을 사용)하고 알맞게 설정, 그리고 웹캠 실시간 분석이면 웹캠에 연결이 되는지 확인
	설정된 DATA 파일, NAMES 파일, CFG 파일, weights 파일이 서로 연동이 되고 연동되는 값들이 일치하는지 확인
	반복문을 사용해서 draw_detections_cv_v3 함수를 연속해서 호출하며, fps(초당 프레임)을 출력하고, 각 프레임의 화면을 출력

draw_detections_ cv_v3 함수	입력된 화면(동영상 또는 웹캠)을 분석해서 학습된 물체를 감지한 후 해당 물체의 이름과 위치(화면상의 픽셀값)를 변수에 저장
	감지된 물체의 위치(픽셀값)를 인자로 넣어서 centercheck 함수 호출 후 centercheck 함수의 조건을 만족하면 이름, 위치, 정확도(학습시킨 파일과 얼마나 유사한지)를 출력
	감지된 물체 중에 사용자와 가장 가까운 물체(세로 픽셀값이 가장 큰 물체)의 픽셀 위치를 넣어서 Distance 함수 호출 후 반환값에 따라서 알맞은 문장을 detection_text.txt 파일에 저장
	감지된 물체 중에 사용자와 가장 가까운 물체(세로 픽셀값이 가장 큰 물체)의 이름을 넣어서 translation 함수 호출
	translation 함수의 반환을 sound 함수의 인자로 넣어서 sound 함수 호출
	학습시킨 특정 손 모양이 감지되면 inputitt 함수 호출
centercheck 함수	감지된 물체의 위치 정보를 인자로 받아서 해당 물체의 왼쪽 좌표값을 전체 화면의 가로 픽셀값 1/4 지점보다 크고 오른쪽 좌표값이 3/4 지점보다 작으면 1(참)을 반환 아니라면 0(거짓)을 반환
Distance 함수	인자로 받은 물체의 아래쪽 좌표값을 미리 설정해둔 화면의 세로 픽셀값과 비교해서 해당 물체와의 거리가 얼마인지 특정(3m 이내, 5m 이내, 5m 이상)
translation 함수	인자로 받은 물체의 이름(영어)에 따라서 알맞은 한글 문장으로 바꿔서 반환("car" -> "자동차가 있습니다" 등)
sound 함수	인자로 받은 문장을 detection_text.txt 파일에 추가로 저장
	감지되는 물체가 없을 때는 "X"를 인자로 받는데 이때는 "전방에 감지되는 물체가 없습니다"라는 문장을 detection_text.txt 파일에 저장
	synthesize_file.py에 detection_text.txt 파일을 인자로 넣어서 실행
inputitt 함수	웹캠의 화면을 캡처 후 캡처한 이미지를 vision_test.py에 인자로 넣어서 실행
	vision_test.py를 사용해서 만들어진 텍스트 파일을 synthesize_file.py의 인자로 넣어서 실행
synthesize_file.py	인자로 받은 텍스트 파일을 음성 파일로 변환 후 해당 음성 파일을 실행
vision_test.py	인자로 받은 이미지에서 텍스트를 추출하여 텍스트 파일을 생성

2.4. 시스템 기능별 요약

가) 장애물 정보 알림 기능

사용자가 전원을 켜면 장애물 알림 기능이 실행되며 사용자와 가장 가까운 장애물의 종류와 장애물과의 거리를 스피커를 통해 소리로 알려준다.

나) 글자 읽어주는 기능

사용자가 특정 손 모양을 취하면 글자 읽어주는 기능이 실행되며 해당 화면을 캡처 후 추출된 텍스트를 스피커를 통해 소리로 알려준다.

3. 결론 및 향후과제

3.1. 작품보완점 및 목표구현 정도

기능	구현도	보완점
장애물 정보 알림 기능	90%	올데이아이즈 전원을 켜면 바로 실행되도록 구상했으나 해당 기능 개발 중 TTS가 실행되지 않는 오류가 발생
글자 읽어주는 기능	100%	카메라에 잡히는 모든 글자를 읽어주기 때문에 주변의 원하지 않은 글자까지 읽어주는 현상이 발생

3.2. 개발 환경 및 도구

개발환경	Ubuntu 18.04, Windows 10
개발도구	Vim, Visual Studio

3.3. 향후 발전 방향

사용자가 직접 소지하고 다녀야 하는 제품이기에 발열, 무게, 부품들의 부품을 더욱 경량화, 최적화하면 사용자의 편의성이 증가할 것이다. 또한, 실시간 감지를 통해 사용자에게 정보를 안내하는 제품인 만큼, 동작 속도 개선에 힘쓰면 사용 효과가 더 상승할 것이다.

3.4. 문제점 및 개선 방안

실시간 감지 속도 개선하여 사용자에게 이전보다 더 빨리 장애물의 정보를 전달하면 좋을 것이다. 또한 weights 파일의 샘플 개수와 장애물의 종류가 더 많아지면 사용자에게 더 많은 장애물을 안내할 수 있다.

4. 사용자 매뉴얼

가. 올 데이 아이즈는 젯슨나노 본체+안경으로 이루어져있다.

1. 초기 세팅이 된 젯슨나노와 안경을 착용한다.
2. 안경에 부착된 이어폰은 직접 착용하거나 귀 뒤에 착용할 수 있다.
3. 젯슨나노와 보조배터리를 연결해 전원을 공급한다.
4. 길을 걸으며 감지되는 장애물이 어떤 것인지, 거리가 몇 m인지에 대한 정보를 이어폰을 통해 음성으로 들을 수 있다.
5. 지팡이 등의 보조기구와 함께 착용하면 더 정확하게 감지할 수 있다.
6. 의약품이나 음식등 사물에 적혀있는 글자를 읽고 싶을 때, 해당 물체를 카메라(안경)앞으로 가까이 가져온 후 특정 손가락 모양(총 모양)을 취한다.
7. 총 모양이 감지되면 주변의 글자를 사용자에게 읽어준다.

부 록

R-1. 참고문헌 및 참고사이트

- [1] YOLO open source site : <https://pjreddie.com/darknet/yolo/>
- [2] Jetson Nano OS install : <https://wendys.tistory.com/141>
- [3] OpenCV camera capture : <https://webnautes.tistory.com/818>
- [4] YOLO mark : <https://reyrei.tistory.com/20>