



국민대학교
전자정보통신대학
컴퓨터공학부

캡스톤 디자인 I

종합설계 프로젝트


프로젝트 명	Be Eyes
팀 명	ViewTopia
문서 제목	계획서

Version	3.1
Date	2019-APR-17

팀원	박 병훈 (조장)
	김 상민
	안 은영
	서 준교
	이 옥걸

CONFIDENTIALITY/SECURITY WARNING

이 문서에 포함되어 있는 정보는 국민대학교 전자정보통신대학 컴퓨터공학부 및 컴퓨터공학부 개설 교과목 캡스톤 디자인 I 수강 학생 중 프로젝트 "Be Eyes"를 수행하는 팀 "ViewTopia"의 팀원들의 자산입니다. 국민대학교 컴퓨터공학부 및 팀 "ViewTopia"의 팀원들의 서면 허락없이 사용되거나, 재가공 될 수 없습니다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

문서 정보 / 수정 내역


Filename	수행계획서 - Be Eyes.doc
원안작성자	박병훈(팀장), 김상민, 안은영, 서준교, 이옥걸
수정작업자	박병훈(팀장), 김상민, 안은영, 서준교

수정날짜	대표수정 자	Revision	추가/수정 항목	내 용
2019-03-09	박병훈	1.0	최초 작성	개요, 개발 목표 및 내용 작성
2019-03-10	김상민	1.1	내용 수정	역할분담, 비용, 개발 일정 작성
2019-03-11	박병훈	2.0	내용 수정	개발 목표 및 내용 수정, 배경 기술 작성
2019-03-13	김상민	2.1	내용 수정	기대효과, 자원관리, 참고문헌 작성
2019-03-14	안은영	2.2	내용 추가	연구 개발 내용 추가 작성
2019-04-15	서준교	3.0	내용 추가	시장 조사 추가 작성
2019-04-17	안은영	3.1	내용 수정	개발 결과 및 내용 수정

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

목 차

1	개요	4
1.1	프로젝트 개요	4
1.2	추진 배경 및 필요성	4
1.2.1	추진 배경 및 사전 조사 내용	4
1.2.2	기술 시장 현황 및 한계점	7
2	“Be Eyes” 개발 목표 및 내용	9
2.1	목표	9
2.2	연구 / 개발 내용	9
2.2.1	데이터 훈련 및 인식	9
2.2.2	OpenCV 를 이용한 객체 추적	11
2.2.3	사용자의 모션인식 및 Image-To-Text, Text-To-Speech 기능	12
2.3	개발 결과	15
2.3.1	시스템 기능 요구사항	15
2.3.2	시스템 비 기능(품질) 요구사항	16
2.3.3	시스템 구조	17
2.3.4	최종 결과물 묘사	18
2.3.5	결과물 목록 및 상세 사양	20
2.4	기대효과 및 활용방안	20
3	배경 기술	21
3.1	기술적 요구사항	21
3.1.1	프로젝트 개발 환경	21
3.2	현실적 제한 요소 및 그 해결 방안	23
3.2.1	하드웨어	23
3.2.2	소프트웨어	23
3.2.3	기타	24
4	프로젝트 팀 구성 및 역할 분담	24
5	프로젝트 비용	25
6	개발 일정 및 자원 관리	26
6.1	개발 일정	26
6.1.1	Gantt Chart Schedule Plan	27
6.2	일정 별 주요 산출물	27
6.3	인력자원 투입계획	29
6.4	비 인적자원 투입계획	30

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

7 참고 문헌

30

1 개요

1.1 프로젝트 개요

본 프로젝트는 **시각장애인을 위한 딥러닝 기반 스마트 안경**을 개발하는 것을 목표로 한다.

안경에 부착된 카메라로 영상처리를 하여 시각장애인 전방의 장애물을 인식하고 데이터 분석을 통해 구체적인 물체의 정보를 시각장애인에게 음성으로 전달해준다. 이를 통해 시각장애인은 안전한 보행이 가능하며 장애물로 인하여 신체에 가해질 위험을 최소화한다.

또한 시각장애인의 모션을 인식하여 손가락이 가리키는 텍스트를 해석하여 음성으로 전달하는 기능을 추가한다. 시각 장애인의 잘못된 약 복용으로 인하여 위험할 수 있는 상황을 예방한다.

위 2 가지 기능을 통하여 시각장애인의 안전한 일상생활이 가능하다.

1.2 추진 배경 및 필요성


1.2.1 추진 배경 및 사전 조사 내용

1) 시각장애인을 위협하는 블라드 [\[1\]](#)

석재 재질에 낮은 블라드 시각장애인 안전 위협

에이블뉴스, 기사작성일 : 2019-01-23 13:43:36

도로의 차량으로부터 사람들을 보호하기 위해 만든 블라드가 시각장애인들에게 오히려 독이 되고 있다. 위 기사에 따르면 시각장애인들은 블라드 때문에 보행에 많은

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

불편을 겪고 있다고 호소한다. 앞이 보이지 않으니 블라드가 오히려 시각 장애인들의 앞길을 가로막는 장애물 같은 역할을 하는 것이다. 양면의 날인 블라드 대신 시각장애인들을 위한 확실한 보행 안전 수단이 필요하다.

2) 점자블록의 현주소 [2]

시각장애인 점자블록, 안전의 또 다른 사각지대


[기획 '안전사회' 경각심 높인다] ⑨ 시각장애인 보행 안전상태 점검해보니

정책기자 박동현 | 2014.06.02



점자블록을 따라 직진하면 차도로 직행하거나, 블록이 부분적으로 아예 사라진 경우도 있다. 또 철봉, 맨홀 등 장애물이 설치돼 있어 보행에 불편을 주고 있다.

보도에 설치된 점자 블록이 차량 진입 등으로 파손되고, 설치한 지 오랜 세월이 흘러 훼손된 곳이 많다. 그로 인해 블록의 요철이 사라져 지팡이로 더듬어 다니기가 쉽지 않고 시각장애인들의 보행 안전을 위협하고 있다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

3) 약조차 먹기 힘든 시각장애인 [3]

HOME > 사회

네 명 중 한명 오남용에도, 시각장애인 배려 없는 의약품

안소영 | 승인 2016.04.01 17:39

댓글 0 | 트위터 | 페이스북



팔에 좁쌀 같은 두드러기가 다닥다닥 올라왔다. 한참을 굶으니 500원짜리 동전만한 붉은 반점이 팔뚝을 뒤덮었다. 병원을 찾아 알레르기약을 받아왔지만, 간지러움에 사흘간은 잠을 설쳤다. 유통기한이 일 년도 더 지난 근육통약을 먹은 뒤 일어난 일이었다. 복용할 당시 약에 대한 정보를 알 수 있는 방법은 전혀 없었다. 시각장애인 이대규 씨의 이야기다.


4) 항상 위험에 노출되어 있는 시각장애인 [4]

"안약인 줄 알았더니 후시딘..." 갈길 먼 시각장애인 '투약권'

입력 2017.11.04 04:40 | 좋아요 20개

로그인 ★ 가 가

시각장애인 손지민(34)씨는 올해 초 저녁부터 열을 동반한 두통을 느꼈던 경험을 얘기하면서 "정말 아찔했다"고 했다. 사연은 이렇다. 병원은 이미 닫았을 시간인데, 점점 심해지는 통증. 급한 대로 책상에 놓여 있던 상비약을 꺼내먹었지만 두통은 나아질 기미가 보이지 않았다. 다음날 아침 그는 약을 다시 복용했지만 두통은 심해져 갔다. 의문은 활동보조인이 와서야 풀렸다. "두통이라면서 잇몸 약은 왜 먹은 거냐"는 그의 질문. 약 포장 형태와 크기가 비슷해 착각으로 구분할 수 없었던 손씨가 잇몸 약을 두통약으로 착각해 두 번이나 복용했던 것이다. 손씨는 "기본 상비약인 두통약조차 점자 표시가 없어 도움 없이는 약 한 알 먹기도 어렵다"고 한숨을 쉬었다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

시급한 필요성에도 불구하고 개선은 여전히 불투명하다. 의약품의 용기, 포장 및 첨부 문서에 점자 또는 음성 변환용 코드를 표시하자는 방안이 나왔지만 약 값 상승 등의 이유로 많은 사람들이 반대하기 때문이다. 시각장애인들의 정보 접근권은 배려가 아닌 의무가 아닐까.

이렇듯 현실에서 시각장애인의 한걸음 한걸음은 살얼음판을 걷는 것과 같다. 이러한 어려움이 반복된다면 시각장애인들은 사회에서 영원히 소외 당할 것이다. 근본적으로 시각장애인이 일상생활 속 겪는 불편함의 원인은 앞을 보지 못하기 때문이다. 따라서 시각장애인의 시야를 대신할 수 있는 무언가의 필요성이 대두된다.

1.2.2 기술 시장 현황 및 한계점

1. 시각장애인용 '호라이즌 스마트 글래스'¹



스타트업 '아이라(Aira)'에서 개발한 호라이즌 스마트 글래스는 일종의 실시간 '비주얼 통역 서비스 안경'이라고 볼 수 있다. 사용자가 이 안경을 쓰면 전면에 부착된 카메라를 통해 보여지는 영상이 실시간 중계되며 연결된 상담원이 통역과 길안내 서비스를 제공하게 된다. 인공지능 음성비서인 클로이(Chloe) 서비스까지 탑재한 것도 호라이즌 안경의 장점이다.

하지만 이용 금액은 한화로 월정액 4 만원수준으로 매우 비싼 편에 속해, 시각장애인에게 매우 부담스러운 가격이다. 그리고 상담원과 직접 연결되기 때문에 서비스를 항상 이용할 수 있지 않다는 한계점이 있다.

¹ 시각장애인용 스마트 글래스 탄생! '아이라', 1boon, 2018 년 04 월 17 일 작성, 2019 년 04 월 17 일 접속, <https://1boon.kakao.com/bizion/5ad53a8b6a8e510001adfeae>

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17


2. 서울시 행정안전부가 시행한 시각장애인을 위한 '엔젤 아이즈' [8]



'엔젤 아이즈'라는 앱을 깔고, 앱에 지인 5~6 명을 등록시킨 뒤, 웨어러블 카메라를 연동시켜 마치 영상통화를 하듯 지인을 연결해 화면을 보여주고 안내를 받는 방식이다.

이용을 위해 앱을 깔고 지인 등록까지 해야 하는데 이는 스마트폰의 사용 자체가 불편한 시각장애인의 입장을 반영하지 않았다고 보여진다. 실제로 시각장애인들이 사용했을 때 접속조차 쉽지 않았다. 또한 웨어러블 카메라도 크기가 너무 크고 10 분만 착용하고 있어도 관자놀이 부분에 통증을 준다는 문제점이 있다.

시중에 출시된 시각장애인을 위한 제품들은 시각장애인들이 홀로 일상생활을 가능하게 하는 것이 아닌 누군가의 도움을 받아야만 되는 기술이다. 따라서 누군가의 도움이 없이 시각장애인의 자립성을 보장해줄 수 없다는 한계가 있다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

2 “Be Eyes” 개발 목표 및 내용

2.1 목표

본 프로젝트는 시각장애인들이 안전하게 일상생활을 할 수 있도록 주변의 물체 및 텍스트를 인식하여 음성으로 전달해주는 스마트 안경을 개발하는 것을 목표로 한다.

- 1) 시각장애인 전방의 장애물의 정보를 알려준다. 예: 불라드
 - 2) 모션 인식을 통해서 시각장애인이 가리키는 문구를 음성으로 출력한다. 예: 약 표시 정보, 식료품 (우유, 과자봉지) 등
- 위 2 가지 기능을 통하여 시각장애인들의 안전한 일상생활에 도움이 되고자 한다.

2.2 연구 / 개발 내용

2.2.1 데이터 훈련 및 인식

목표 : 이 프로젝트는 YOLO(You Only Look Once)를 이용하여 객체 탐지 및 추적을 수행하며 추가적인 데이터 학습을 실시한다. 빠른 객체 인식을 위해 이미지를 한번에 스캔하는 방식으로 객체를 인식하고 뉴럴 네트워크 프레임워크인 다크넷을 기반으로 동작한다. 따라서 YOLO 에서 지원하는 학습 모델들과 추가적인 데이터 학습으로 시각장애인 주변 사물을 실시간으로 인식할 수 있다.

YOLO 는 각 이미지를 $N \times N$ 개의 그리드로 분할하고, 그리드의 신뢰도를 계산한다. 신뢰도는 그리드 내 객체 인식 시 정확성을 반영한다. 아래 그림과 같이 처음에는 객체 인식과는 동떨어진 경계 상자가 설정되지만, 신뢰도를 계산하여 경계 상자의 위치를 조정함으로써, 가장 높은 객체 인식 정확성을 가지는 경계 상자를 얻을 수 있다.

 <div> <p>국민대학교</p> <p>컴퓨터공학부</p> <p>캡스톤 디자인 I</p> </div>	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

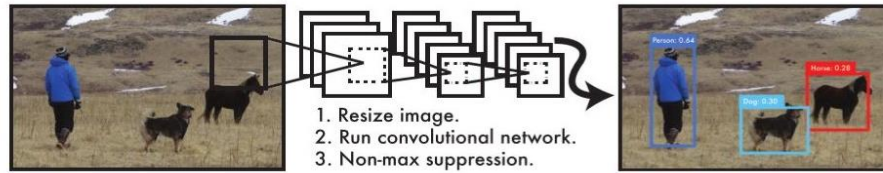
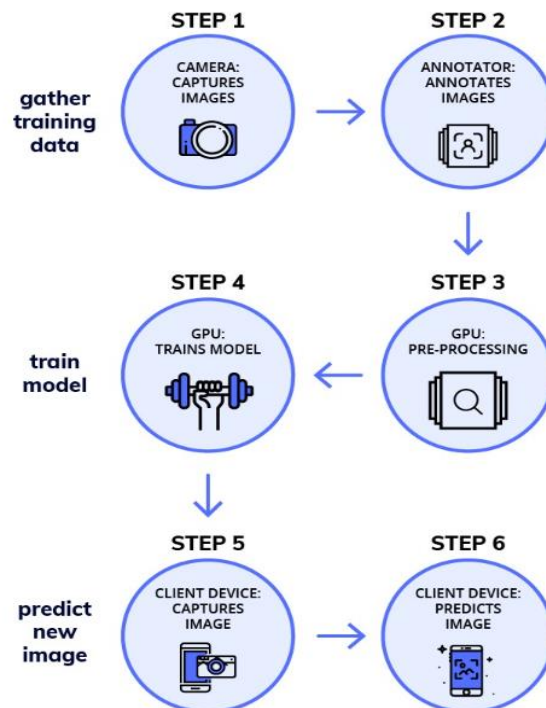



Figure 1: The YOLO Detection System. Processing images with YOLO is simple and straightforward. Our system (1) resizes the input image to 448×448 , (2) runs a single convolutional network on the image, and (3) thresholds the resulting detections by the model's confidence.

<그림> Joseph Redmon By University of Washington,
"You Only Look Once: Unified, Real-Time Object Detection",
Introduction(2016), p1, [\[6\]](#)



<그림> Sarthak Jain, "Object Detection With Deep Learning",

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

How to Build Deep Learning Model(2018) ²

2.2.2 OpenCV 를 이용한 객체 추적

목표 : 관심영역 및 관심객체를 탐지하고 추적한다. 이를 위해 물체 감지 및 추적(Object Detection and Tracking) 방법을 이용한다. 물체 추적의 목적은 비디오 입력 영상으로부터 관심있는 영역을 세분화하는 것과 움직임, 위치에 대한 상황들의 감지를 지속하는 것이다.

1. 관심영역 인식 (The identification of regions of interest)

객체를 탐지하는데 있어서 가장 먼저 이루어져야 하는 것은 관심영역을 인식하는 것이다. 비디오로 입력 받은 이미지는 크게 두 부분으로 나눌 수 있는데 첫번째는 전경객체에 대한 픽셀을 담고 있고, 두번째는 배경 객체에 대한 픽셀 정보를 담고 있다. 처리된 이미지 결과는 이진법으로 변형된 이미지나 마스크 형태로 나타나게 된다.

2. 객체 탐지 (Object Detection)

객체 탐지는 객체의 픽셀들의 집합과 비디오 입력 중에서 관심 있는 객체를 식별하는 과정이다. 객체 탐지의 방법에는 프레임 차이를 이용하는 방법, 광학 흐름을 이용하는 방법, 배경을 빼는 방법 등이 존재한다.


3. 객체 분류 (Object Classification)

객체는 자동차, 새, 구름, 나무 등과 같이 정적인 객체와 움직이는 동적인 객체로 분류할 수 있다. 이러한 객체를 분류하기 위한 방법에는 형상 기반 분류, 동작 기반 분류, 색상 기반 분류, 텍스트처 기반 분류 등이 있다.

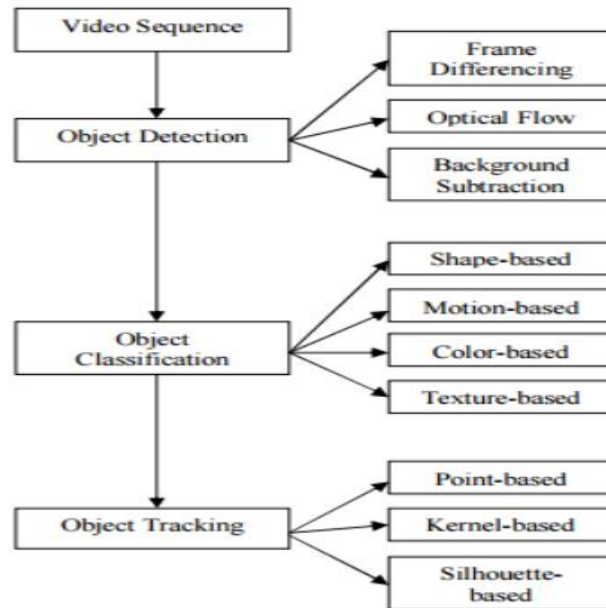
4. 객체 추적 (Object Tracking)

추적은 움직이는 장면에서 이미지면에 있는 물체의 경로의 근사치를 구하는 문제라고 할

² Sarthak Jain, “How to easily Detect Objects with Deep Learning on Raspberry Pi” , 2018 년 3 월 21 일, <https://medium.com/nanonets/how-to-easily-detect-objects-with-deep-learning-on-raspberrypi-225f29635c74>

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17


수 있다. 즉, 영상에서 관심 있는 물체의 경로가 그 전 프레임과 유사한지 알아낸 다음 동일 객체라고 인식하게 되면, 그 물체를 계속 추적하는 것이다. 객체 추적의 방법에는 포인트 추적, 커널 추적, 실루엣 추적 등이 존재한다.

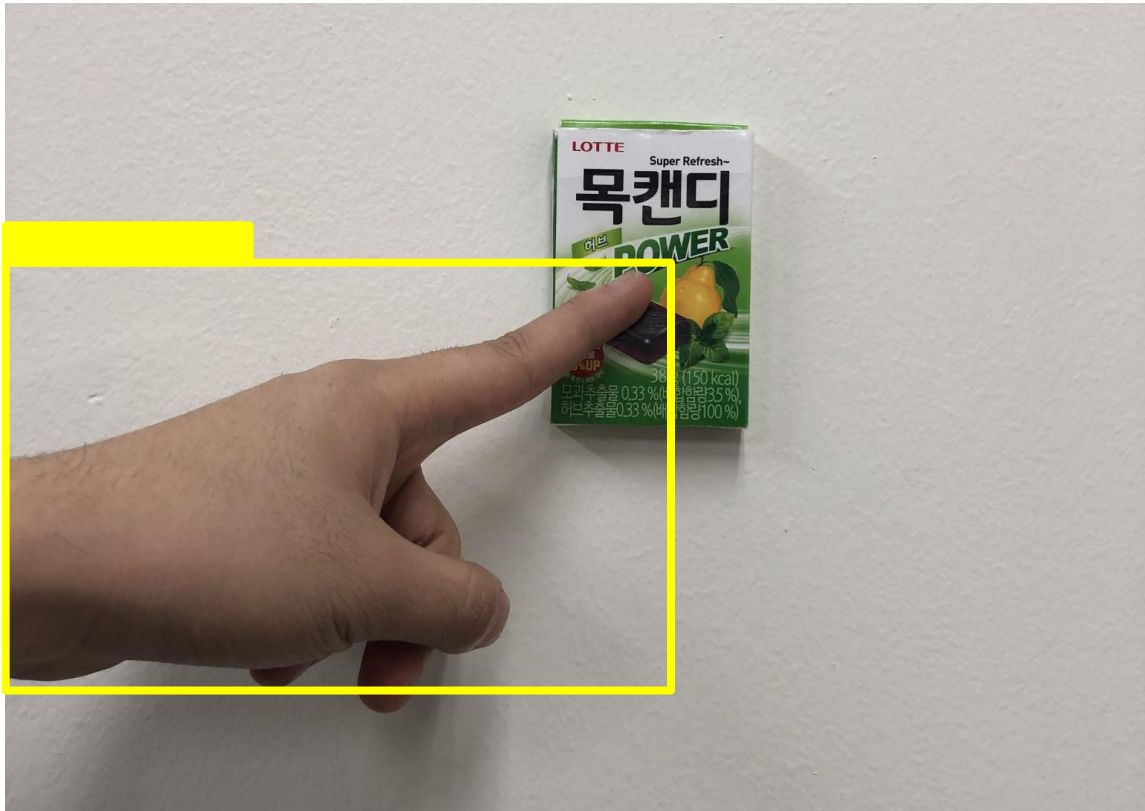


<그림> Himani S. Parekh, "A Survey on Object Detection and Tracking Methods", Basic Steps for tracking an object(2014), [\[7\]](#)

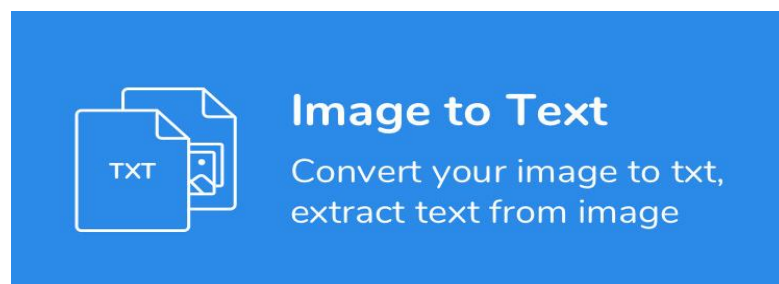
2.2.3 사용자의 모션인식 및 Image-To-Text, Text-To-Speech 기능

목표 : 손가락 모양의 데이터를 미리 학습시킨다. 영상처리를 통해 카메라가 손가락을 인식했을 때, 이미지를 저장한다. 이미지에서 손가락의 끝을 기준으로 바운더리를 잡아 Cloud Vision API 를 이용해 텍스트로 변환한다. 그 후, 텍스트의 내용을 구글에서 지원하는 Text To Speech API 를 이용해 음성으로 말해준다. 모션 인식 기능은 시각장애인들의 주된 정보입력 방식을 '음성'에서 '동작'으로 바꾼 것에 의의가 있다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17




<그림> OpenCV 손동작 인식(2014)



<그림> Hipdf, "Image to Text"(2018) ³

³ Hipdf, "Image to Text", Hipdf(2018), 2019 년 3 월 12 일 접속,
<https://www.hipdf.com/kr/image-to-text>


 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

tts process



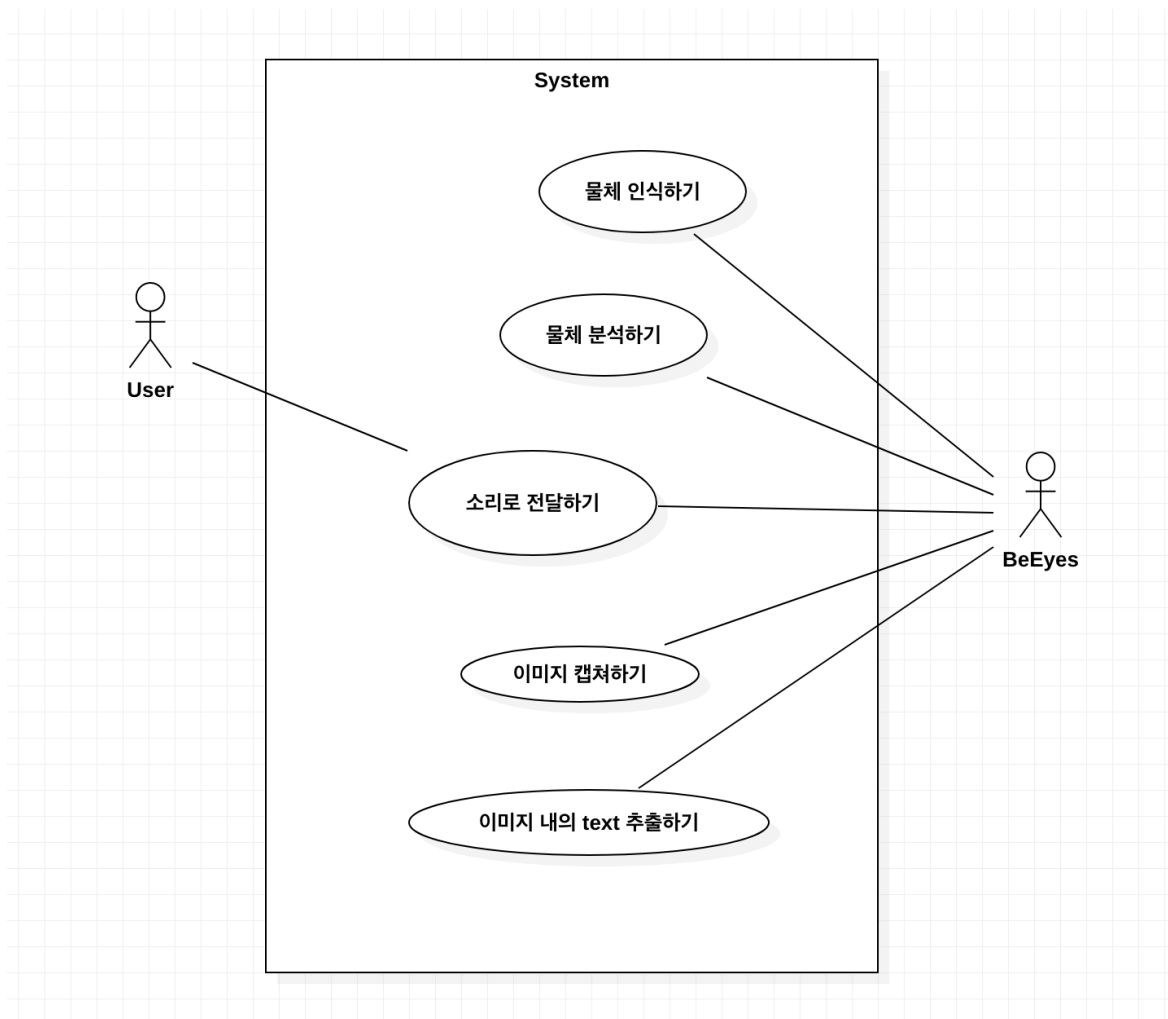
<그림> Ivona Software, "TTS Process", An Amazon Company (2019) ⁴

⁴ IVONA Software, "TTS Process", An Amazon Company (2019), 2019 년 3 월 12 일 접속,
<https://www.ivona.com/us/about-us/text-to-speech/>

 <div> <p>국민대학교</p> <p>컴퓨터공학부</p> <p>캡스톤 디자인 I</p> </div>	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

2.3 개발 결과

2.3.1 시스템 기능 요구사항




<그림> Use case Diagram

- BeEyes의 카메라 모듈을 통해 물체를 인식하기
- 실시간으로 물체 분석하기
- 검지 손가락 객체 인식시
 - 이미지 캡처하기
 - 이미지 내의 text 추출하기
- 분석한 물체 또는 text를 BeEyes의 소리 모듈을 통해 User에게 전달하기

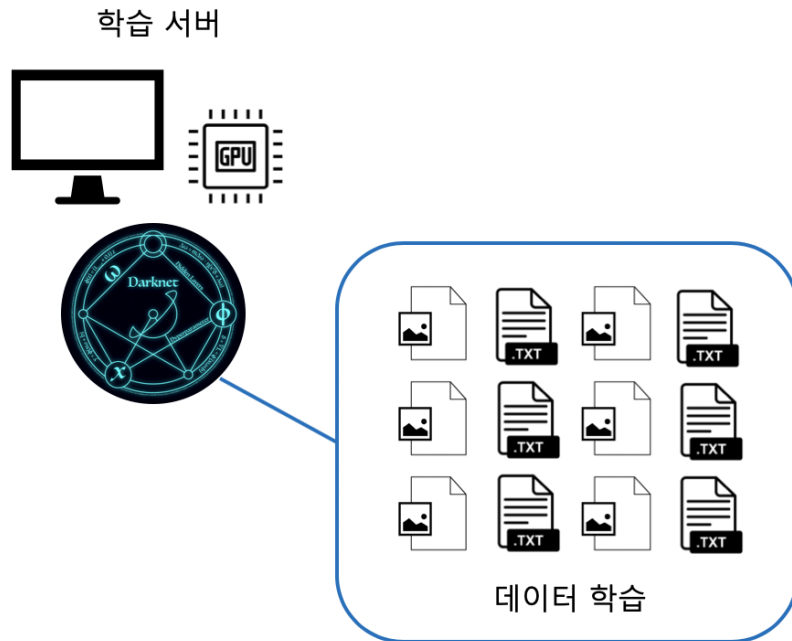
 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

2.3.2 시스템 비 기능(품질) 요구사항

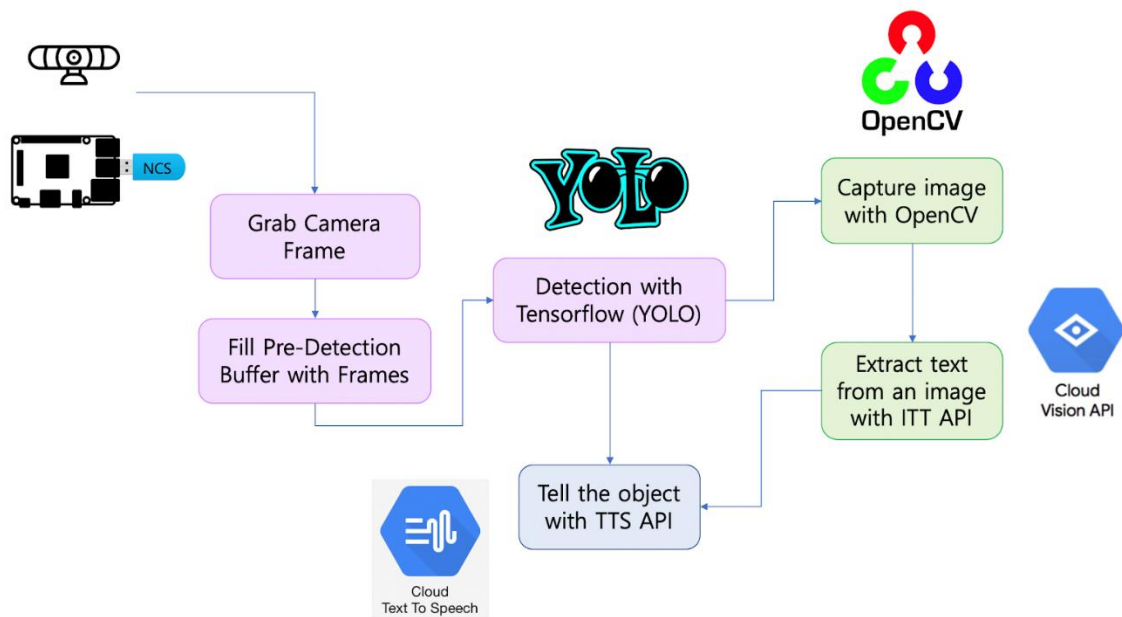
품질 속성	품질 요구사항 이름	설명	측정 요소
Performance(성능) - 1 초 동안 보여주는 화면의 수가 6 프레임 이상으로 탐지한다.	실시간 물체 분석	시각장애인에게 실질적으로 도움이 될 수 있도록 영상에 대한 정보를 분석하고 전달하는 과정이 실시간으로 진행되어야 한다.	Detection 이 6fps 정도 나오는지 확인
Reliability(신뢰성) - 정확도 90% 이상인 결과를 사용자에게 전달한다.	물체 분석 정확성	시각장애인이 사물을 직접 확인 할 수 없으므로 물체 분석의 정확도가 높아야 한다.	정확도 90% 이상 나오는지 확인
Usability(사용성) - 1 초만에 결과를 사용자에게 소리로 전달한다.	정보 전달 용이성	사용자는 소리로 실시간 전달을 받아야한다.	1 초만에 소리 출력되는지 확인
Maintainability(유지보수성) - 디바이스에서 새로 추가 학습한 물체를 인식한다.	새로운 데이터 추가 학습 용이성	사용자가 사용하면서 얻게 되는 새로운 이미지 데이터를 모아 두었다가 서버로 보내주어 학습시킨 뒤 학습 모델을 업데이트 해줄 수 있어야 한다.	디바이스에서 물체 인식 확인

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

2.3.3 시스템 구조



<그림> "Be Eyes" 데이터 훈련



<그림> "Be Eyes" 시스템 구조

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

시스템의 전체적인 구조는 위와 같다. 라즈베리파이에 부착된 카메라가 장애물을 인식한다. 영상처리를 통해 인식된 이미지는 YOLO 에 저장된 모델에 넣어서 분류된 값을 구한다. 그 값을 텍스트로 출력하고 출력 결과를 TTS(Text to Speech) API 를 이용하여 소리로 출력한다. 모션인식 또한 라즈베리파이에 부착된 카메라로 실행되며 이미지를 저장한 뒤 Cloud Vision API 를 이용하여 이미지를 텍스트로 변환하고 그 텍스트 정보를 TTS API 를 이용하여 소리로 출력한다.

사용자의 보행에 방해가 될 만한 장애물들을 추가적으로 서버에서 학습시키고 이를 최신화 하여 분류기 안의 클래스의 개수를 늘린다. 추가적인 분류화로 인해 더 많은 장애물들을 식별할 수 있고, 사용자가 더욱 더 안전한 보행이 가능하도록 한다.


2.3.4 최종 결과물 묘사

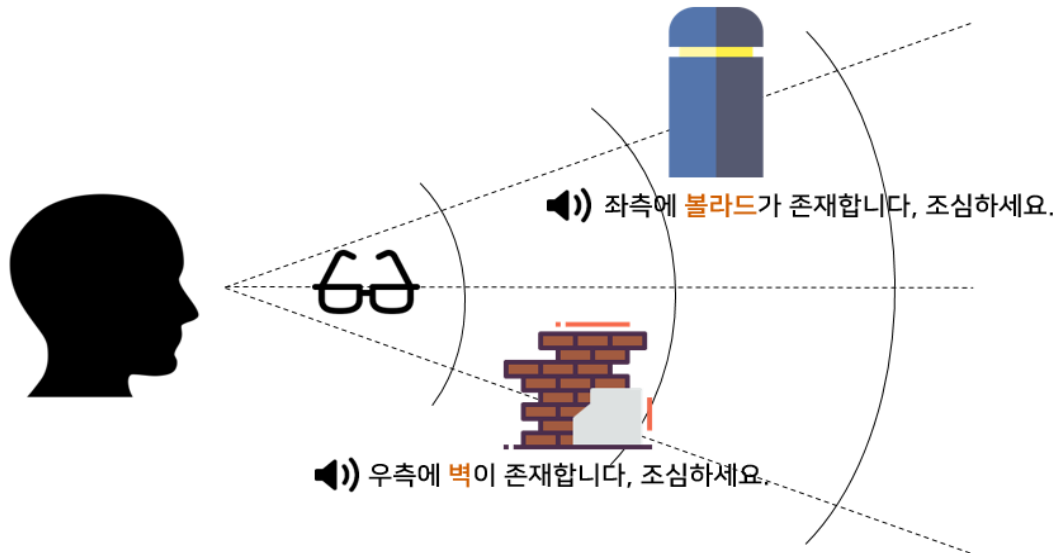


<그림> Orcam Device, Get Orcam ⁵

기능 1) 시각장애인에게 일정한 거리 내에서 인식된 장애물의 정보가 포함된 경고 메시지를 음성으로 전달한다.

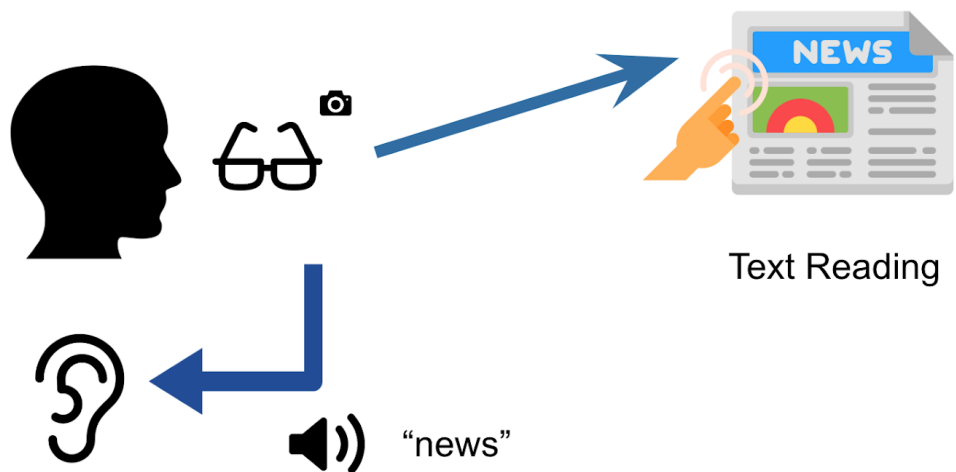
⁵ “What is OrCam?” , Amit Systems, 2019 년 3 월 12 일 접속, <https://www.getorcaml.com/what-is-ocaml>

 <div> <p>국민대학교</p> <p>컴퓨터공학부</p> <p>캡스톤 디자인 I</p> </div>	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17




<그림> “Be Eyes” 객체 탐지 기능 예시

기능 2) 시각장애인이 읽고자 하는 부분을 손가락으로 가리키면, 손가락이 가리키는 영역이 존재하는 텍스트를 음성으로 시각장애인에게 전달해준다.



<그림> “Be Eyes” 모션 인식 및 Text Reading 예시


 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

2.3.5 결과물 목록 및 상세 사양

대분류	소분류	기능	형식	비고
라즈베리 파이	사물 인지	YOLO 를 이용해 만든 모델에 인식해야 할 이미지를 넣어서 분류된 값을 구함	함수	
	모션 인식	특정 손 모양을 학습시킨 모델을 활용하여 모션을 인식	함수	
	이미지 처리	ITT(Image To Text) API 를 이용해 이미지 내의 텍스트를 추출	함수	
	음성 처리	TTS(Text to Speech) API 를 이용해 인식한 물체 및 텍스트를 음성으로 출력	함수	
장치	라즈베리 카메라 모듈 v2	영상 촬영		
	Logitech c920	영상 촬영		
	스피커	소리 출력		
	Intel Movidius NCS	딥러닝용 USB 드라이브		

2.4 기대효과 및 활용방안

세계보건기구(WHO)에 따르면 전 세계 시각장애인은 2 억 3300 만 명에 달하며, 2 억 1700 만 명이 중등도 이상의 시력 손상을 겪고 있다. 시각 장애인을 위한 안경 타입의 스마트 기기는 2017 년 중국(상하이) 기술 박람회에서 '천사의 눈'이라는 이름으로 세계 최초로 등장하였다. 2018 년에는 미국의 한 스타트업에서 AI 를 도입한 음성 비서 역할을 하는 스마트 글라스 아이라(Aira)가 출시되면서 현재에도 꾸준히 발전하고 있으며, 공익 차원에서도 시장성이 풍부한 분야이다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

그러나 인식의 부족과 비싼 가격 때문에 시각장애인들의 인원수에 비해 수요가 적고 시각장애인들은 여전히 수 많은 위험에 노출되어 있다. 따라서 어떠한 돌발적인 상황에서도 기술적인 도움을 통해 시각 장애인들이 스스로 위험을 인지하고 대처할 수 있어야 한다.

“Be Eyes”는 시각장애인들이 장애물의 종류와 위치를 인지하고 스스로 위험을 예방할 수 있도록 제작하여 착용자들의 안전을 보장한다. 또한 식료품이나 의약품 표시 정보를 정확하게 알 수 있도록 하여 시각장애인들의 불안과 걱정을 덜어줄 수 있다.

본 프로젝트에서는 “Be Eyes”를 스마트 안경으로 제작하지만 안경뿐만 아니라 원하는 곳에 기기를 탈부착 할 수 있도록 웨어러블 형태로 제작하는 것을 고려하여 착용자들에게 편리한 사용성을 제공할 것이다. “Be Eyes”를 제작하여 많은 사람들에게 공급할 수 있다면 사용자들에게 신체적 한계를 극복한 기쁨을 주고 삶의 원동력을 제공할 것이다. 더불어 사회적, 상업적인 가치를 창출할 수 있다.

3 배경 기술

3.1 기술적 요구사항


3.1.1 프로젝트 개발 환경

- 프로젝트를 개발하는데 필요한 개발 환경

- 1) 운영체제 : Window 10, Ubuntu 16.04, Mac OS X / Raspbian OS
- 2) 컴파일 : g++
- 3) 개발언어 : Python, C/C++

- 프로젝트 결과물 확인 환경

Raspberry Pi 3 Model B+, Camera Module v2, Audio Module 을 연동하여 설계한 디바이스

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

● 라이브러리

1) TensorFlow Object Detection API

물체를 인식하는 모델을 사용자가 쉽게 학습 / 제작 / 배포 할 수 있는 오픈소스 프레임워크이다. 본 프로젝트에서는 직접 이미지를 학습시키고 모델을 제작해야 하므로 이 API 를 사용한다.

2) YOLO

실시간 사물 탐지 및 인식이 가능하다. 본 프로젝트는 실시간으로 객체를 탐지 및 인지할 수 있어야 하므로 이 라이브러리를 사용한다. 라즈베리파이는 메모리 용량과 전력 소모의 문제 때문에 YOLO 정식 버전 대신 저용량의 Tiny YOLO 를 사용한다.

여러 딥러닝 알고리즘 중 YOLO 를 쓰는 이유는 다음과 같다.


종류	mAp	Fps (Titan X GPU)	우선순위
YOLO V2	76.8	67 Fps	1
SSD 500	75.5	58 Fps	2
SSD 300	72.1	58 Fps	3
Faster R-CNN	73.2	7 Fps	4
Fast R-CNN	68.4	-	5

<그림> 이상현, Titan X GPU 기준의 프레임과 정확도, Tistory (2017) ⁶

3) Google Cloud Platform - Vision API

손 동작을 인식하게 되면 이미지를 가져와서 이미지 안의 텍스트 정보를 알 수 있어야 한다. 따라서 광학 문자 인식(OCR)을 사용하여 자동 언어 식별과 함께 이미지 내의 텍스트를 감지할 수 있는 Vision API 를 사용하여 텍스트를 추출한다.

⁶ 이상현, “Titan X GPU 기준의 프레임과 정확도”, Tistory, 2017 년 12 월 11 일 작성, 2019 년 3 월 12 일 접속, <https://pgmrlsh.tistory.com/2?category=766787>

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

4) Google Cloud Platform - Text To Speech API

객체를 인식한 후 사용자에게 텍스트가 아닌 음성으로 전달을 해주어야 한다.

인식한 객체의 종류가 텍스트로 들어오기 때문에 본 프로젝트에서는 이 텍스트를 입력으로 받아 음성으로 출력한다.

3.2 현실적 제한 요소 및 그 해결 방안

3.2.1 하드웨어

- 1) 라즈베리파이를 통해 실시간으로 영상 처리와 객체 판별을 하는 것은 많은 양의 전력을 소모한다. 따라서 Intel Movidius NCS(Neural Compute Stick)라는 딥러닝용 USB 드라이브를 통해 저전력으로 구동할 수 있도록 고려한다.
- 2) 학습된 모델을 통해 연산을 많이 해야 하기 때문에 만족스러운 속도가 나오지 않는다면 라즈베리파이 이외에 처리 속도가 빠른 보드를 사용하여 동일한 환경으로 구동시킨다.
- 3) 네트워크를 통해 서버에서 학습시킨 데이터 모델을 라즈베리파이 안에 업데이트 하는 것에 제한이 있을 시 물리적 수단(예: USB)을 이용하여 추가적으로 학습된 데이터를 업로드한다.
- 4) 카메라의 시야 안에 물체가 인식되지만 사용자의 보행에 직접적으로 장애가 되지 않는 경우를 대비하여 카메라의 위치를 사용자의 신체에 맞게 조절하여 장애가 되는 물체만 인식하게 한다.

3.2.2 소프트웨어

- 1) 기존의 다른 real-time 비전 기술은 처리 속도가 매우 느렸다. 하지만 YOLO 는 단순한 처리로 속도가 매우 빠르다. 기존의 실시간 물체 탐지 기술과 비교할 때, 2 배 정도 높은 성능을 보인다. 이미지 전체를 한번에 바라보는 방식으로 물체를 판별하기 때문이다. 따라서 데이터 분석 및 처리에 YOLO 모델을 사용한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

- 2) 프로젝트 기간 내에 학습시킬 데이터의 개수에 제한이 생기므로 데이터의 개수를 2개(볼라드, 손가락)로 제한한다.
- 3) 특정 모션을 통해 텍스트를 읽는 기능은 시각장애인이 읽고자 하는 글자가 반대로 또는 거꾸로 되어있으면 글자 인식이 불가능하다. 따라서 디바이스가 글자를 올바르게 인식할 수 있도록 사용자에게 음성으로 텍스트를 인식할 수 없다는 음성 메시지를 출력해준다. 글씨의 크기 또한 제한요소가 될 수 있으므로 본 프로젝트에서는 큰 글씨 (예 : 폰트 72)를 인식하게 한다.

3.2.3 기타

- 1) 안경이나 선글라스에 부착한 디바이스의 크기나 무게가 사용자들이 사용하기에 적합하지 않다. 따라서 안경이 아니더라도 모자나 옷 일부에 탈부착 할 수 있게 소형의 웨어러블 기기 형태로 제작하는 것을 고려한다.
- 2) 본 프로젝트에서 제안한 2 가지 기능을 모두 구현하기에 캡스톤 디자인 프로젝트의 기간이 짧다. 따라서 실시간 객체 인식 기능을 중점적으로 진행한 이후에 모션 인식 기능을 구현하는 것으로 시간을 조율하여 진행한다.

4 프로젝트 팀 구성 및 역할 분담

이름	역할
박병훈	<ul style="list-style-type: none"> - Project Team Leader - Data 처리 및 훈련 - Object Detection & Tracking 기능 구현
김상민	<ul style="list-style-type: none"> - Data 처리 및 훈련 - Object Detection & Tracking 기능 구현

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

안은영	<ul style="list-style-type: none"> - Text Reading 하기 위한 모션 인식 설정 - Text To Speech 기능 구현
서준교	<ul style="list-style-type: none"> - Text Reading 하기 위한 모션 인식 설정 - Text Reading 동작 설정 및 Text To Speech 기능 구현
이옥걸	<ul style="list-style-type: none"> - 라즈베리파이 환경 설정 및 하드웨어 디바이스 제작

5 프로젝트 비용


항목	예산치 (MD)
아이디어 구상	20
개발 환경 구축	5
시스템 설계	20
하드웨어 디바이스 제작	10
Object Detection & Tracking 기능 구현	20
Text Reading 모션 인식 동작 설정	20
Text To Speech 기능 구현	10
시스템 통합 테스트 및 디버깅	25
데모 동영상 및 PPT 제작	10
합	140

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

6 개발 일정 및 자원 관리

6.1 개발 일정

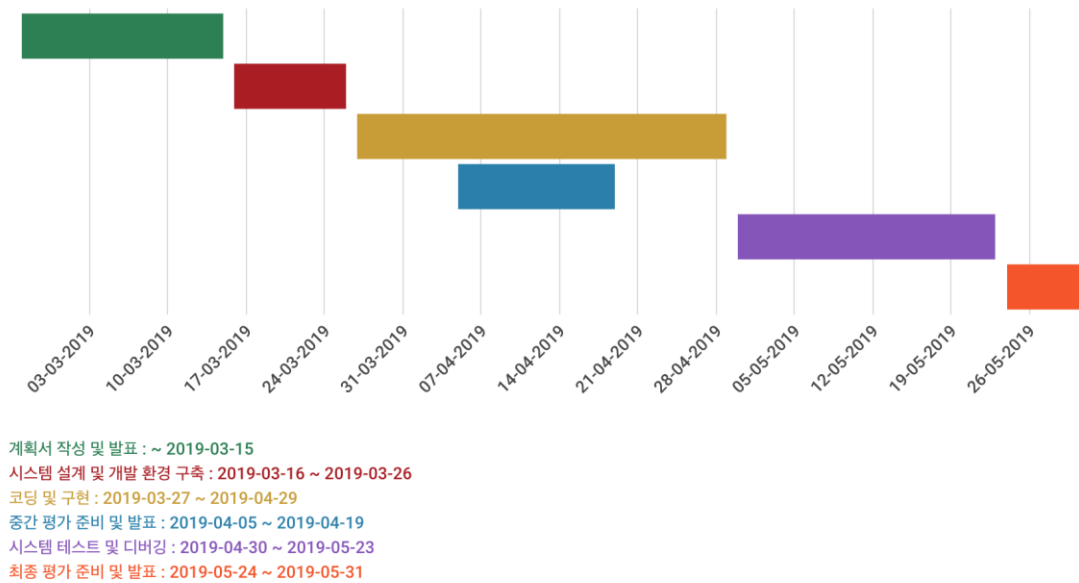
항목	세부내용	2 월	3 월	4 월	5 월	6 월	비고
요구사항분석	요구 분석						
	SRS 작성						
관련분야연구	주요 기술 연구						
	관련 시스템 분석						
설계	시스템 설계						
구현	코딩 및 모듈 테스트						
테스트	시스템 테스트						
평가	전시 및 최종 평가						

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

6.1.1 Gantt Chart Schedule Plan

"Be Eyes" : Team Project Plan

Team BeEyes's Project Plan Of 2019 Capstone Design Class.




<그림> "Be Eyes" Gantt Chart Schedule ⁷

6.2 일정 별 주요 산출물

마일스톤	개요	시작일	종료일
계획서 발표	프로젝트 기획 및 요구 사항 분석 산출물 : 1. 프로젝트 수행 계획서	~	2019-03-15

⁷ "Be Eyes" Gantt Chart, Infogram, 2019 년 3 월 10 일 작성,
https://infogram.com/capstone_design_2019_6-1h0r6rwmlzg34ek?live


 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

	2. 계획서 발표 슬라이드 쇼 3. 계획서 발표 슬라이드 화일		
시스템 설계 및 개발 환경 구축	하드웨어 설계 및 개발 환경 구축 완료 딥러닝 모델 학습 및 객체 탐지 및 추적 기능 설계 산출물: 1. 개발 환경 2. 하드웨어 디바이스	2019-03-16	2019-03-26
중간 평가	Object Detection & Tracking(객체 탐지 및 추적) 구현 완료 Text Reading 모션인식 동작 설정 완료 산출물: 1. 프로젝트 중간 평가 요약 보고서 2. 프로젝트 중간 평가 발표 자료 3. 수행 계획서 수정본	2019-03-27	2019-04-19
구현 완료	TTS(Text To Speech) 기능 구현 완료 시스템 구현 완료 산출물: 1. 프로젝트 소스코드 2. 데모 버전	2019-04-20	2019-05-15
테스트	시스템 통합 테스트 산출물: 1. 최종 버전	2019-05-16	2019-05-23
최종 보고서	최종 발표 산출물: 1. 최종 평가 결과 보고서 2. 전시용 자료(포스터 및 소개 책자) 3. 온라인 평가용 자료 4. 최종 발표 자료 및 시연 동영상	2019-05-24	2019-05-31

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

6.3 인력자원 투입계획

이름	개발항목	시작일	종료일	총개발일(MD)
전원	아이디어 구성 및 요구 분석	2019-02-18	2019-03-15	20
전원	개발 환경 구축	2019-03-16	2019-03-23	5
전원	시스템 설계	2019-03-24	2019-03-31	15
이옥걸	하드웨어 디바이스 설계	2019-03-24	2019-03-31	20
박병훈	- Data 처리 및 훈련 - Object Detection & Tracking 구현	2019-04-01	2019-05-08	30
김상민	- Data 처리 및 훈련 - Object Detection & Tracking 구현	2019-04-01	2019-05-08	30
안은영	-Text Reading 동작 설정 -Text To Speech 기능 구현	2019-04-01	2019-05-08	30
서준교	-Text Reading 동작 설정 -Text To Speech 기능 구현	2019-04-01	2019-05-08	30
전원	시스템 통합 테스트 및 디버깅	2019-05-09	2019-05-23	20
전원	최종 발표 준비	2019-05-24	2019-05-31	10

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 3.1	2019-APR-17

6.4 비 인적자원 투입계획

항목	Provider	시작일	종료일	Required Options
개발용 PC 2 대	Samsung	2019-03-16	2019-05-31	
개발용 PC 1 대	Lenovo	2019-03-16	2019-05-31	
개발용 PC 1 대	Apple	2019-03-16	2019-05-31	
Raspberry Pi 3 Model B+	Raspberry	2019-03-16	2019-05-31	Camera Module V2 Speaker (12S 3W)
Movidius NCS	Intel	2019-03-16	2019-05-31	USB Drive

7 참고 문헌

번호	종류	제목	출처	발행 년도	저자	기타
1	기사	석재 재질에 낮은 볼라드 시각장애인 안전 위협	에이블뉴스	2019	박종태	
2	기사	시각장애인 점자블록, 안전의 또 다른 사각지대	대한민국 정책브리핑	2014	박동현	
3	기사	네 명 중 한 명 오남용에도, 시각장애인 배려 없는 의약품	스토리오브서울	2016	안소영	
4	기사	안약인 줄 알았더니 후시딘...’ 갈길 먼 시각장애인 ‘투약권’	한국일보	2017	강진구	
5	논문	이동 카메라에서 움직이는 물체 검출 및 추적	한국정보기술 응용학회	2007	오운환, 이은주	
6	논문	You only look once: Unified, real-time object detection	cv- foundation	2016	Joseph Redmond	언어 (E NG)
7	논문	A Survey On Moving Object Detection and Tracking Methods	Semantic Scholar	2014	Himani S, Parekh	언어 (EN G)
8	기사	[취재파일] 시각장애인은 안중에도 없는 서울시의 시각장애인 사업	SBS NEWS	2017	한지연	