



국민대학교
전자정보통신대학
컴퓨터공학부

캡스톤 디자인 I

종합설계 프로젝트

프로젝트 명	Be Eyes
팀 명	ViewTopia
문서 제목	결과보고서

Version	1.6
Date	2019-MAY-28

팀원	박 병훈 (조장)
	김 상민
	안 은영
	서 준교
	이 옥걸

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

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.

문서 정보 / 수정 내역

Filename	결과보고서 – Be Eyes.doc
원안작성자	서준교, 안은영
수정작업자	박병훈(팀장), 김상민, 안은영, 서준교

수정날짜	대표수정자	Revision	추가/수정 항목	내 용
2019-05-07	서준교	1.0	최초 작성	
2019-05-17	서준교	1.1	내용 추가	
2019-05-19	안은영	1.2	내용 추가	
2019-05-21	박병훈	1.3	내용 추가	
2019-05-21	김상민	1.4	내용 추가	
2019-05-25	서준교	1.5	내용 수정	
2019-05-28	박병훈	1.6	최종 정리	

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

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	개발 내용 및 결과물	9
2.1	목표	9
2.2	연구 / 개발 내용 및 결과물	9
2.2.1	연구 / 개발 내용	9
2.2.2	시스템 기능 요구사항	17
2.2.3	시스템 비기능(품질) 요구사항	18
2.2.4	시스템 구조 및 설계도	19
2.2.5	활용 / 개발된 기술	21
2.2.6	현실적 제한 요소 및 그 해결 방안	28
2.2.7	결과물 목록	30
2.3	기대효과 및 활용방안	31
3	자기평가	31
4	참고 문헌	33
5	부록	34
5.1	사용자 매뉴얼	34
5.2	운영자 매뉴얼	35
5.3	테스트 케이스	36

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

1 개요

1.1 프로젝트 개요

본 프로젝트는 **시각장애인**을 위한 **스마트 안경**을 개발하는 것을 목표로 한다. 안경에 부착된 카메라로 영상처리를 하여 시각장애인 전방의 장애물을 인식 및 식별을 통해 구체적인 물체의 정보를 시각장애인에게 음성으로 전달해준다. 이를 통해 시각장애인은 안전한 보행이 가능하며 장애물로 인하여 신체에 가해질 위험을 최소화한다. 또한, 시각장애인이 글을 읽으려는 행위(예: 손가락으로 가리킴)를 하면 읽고자 하는 문구를 음성으로 전달하는 기능을 추가한다. 시각 장애인의 잘못된 약 복용 및 음식 섭취로 인하여 위험할 수 있는 상황을 예방한다. 위 두 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 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.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

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	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

3) 약조차 먹기 힘든 시각장애인 [3]

HOME > 사회

네명 중 한명 오남용에도, 시각장애인 배려 없는 의약품

안소영 | 승인 2016.04.01 17:39

댓글 0 | 트위터 | 페이스북



팔에 찰흙 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

🔍 ★ 📄 📄

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	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

의약품에 대한 시각장애인의 접근성의 개선이 시급함에도 불구하고 개선 여부는 여전히 불투명하다. 의약품의 용기, 포장 및 첨부 문서에 점자 또는 음성 변환용 코드를 표시 하자는 방안이 나왔지만 약값의 상승 등의 이유로 많은 사람들이 반대하기 때문이다. 시각장애인들이 정보에 접근할 수 있는 권리를 보장해주는 것은 배려가 아닌 의무가 아닐까.

이렇듯 현실에서 시각장애인의 한걸음 한걸음은 살얼음판을 걷는 것과 같다. 이러한 어려움이 반복된다면 시각장애인들은 사회에서 점점 소외될 것이다. 근본적으로 시각장애인이 일상생활 속 겪는 불편함의 원인은 앞을 보지 못하기 때문이다. 따라서 시각장애인의 시야를 대신할 수 있는 장치의 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 음성 비서인 클로이(Chloe) 서비스까지 탑재한 것도 호라이즌 안경의 장점이다.

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

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

하지만 이용 금액은 한화로 약 월 40 만원 수준으로 매우 비싼 편에 속해, 시각장애인에게 매우 부담스러운 가격이다. 그리고 상담원과 직접 연결되기 때문에 서비스를 항상 이용할 수 있지 않다는 한계점이 있다.

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.

그 밖에 시중에 출시된 제품들 또한 시각장애인들이 홀로 일상생활을 가능하게 하는 것이 아닌 누군가의 도움을 받아야만 되는 기술이 대부분이다. 따라서 타인의 도움 없이는 시각장애인의 자립성을 보장해줄 수 없다는 한계가 있다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

2 개발 내용 및 결과물

2.1 목표

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

- 1) 시각장애인 전방의 장애물의 정보를 알려준다. 장애물의 위치와 정보를 파악함으로써 시각 장애인은 상황에 따른 구체적인 인식 및 대처가 가능하다.
- 2) 시각장애인이 손가락으로 가리키는 영역에 존재하는 문구를 음성으로 전달한다. 문구의 내용을 인지한 시각장애인은 안전한 약 복용 및 음식 섭취가 가능하다.

위 두 가지 기능을 통하여 시각장애인들이 안전한 일상생활을 할 수 있도록 돕는 것을 목표로 한다.


2.2 연구 / 개발 내용 및 결과물

2.2.1 연구 / 개발 내용

1) 딥러닝 기반 객체 탐색 기법

본 프로젝트를 진행함에 있어서 가장 기본이 되는 사항은 장애물로 선별한 물체들을 빠르고 정확하게 탐지하는 것이다. 따라서 효율적인 객체 탐지가 가능한 방법을 선정하는 것이 무엇보다 중요하다. 이에 따라, 딥러닝을 기반으로 한 여러 가지 객체 탐색 기법을 조사하였다. 대표적인 객체 탐색 기법으로는 R-CNN, SSD, YOLO 가 있다.

R-CNN 은 영상에 선택적 탐색(selective search)이라는 방법을 적용하여 객체가 있을 법한 후보 영역을 찾고, 각 후보 영역에 컨볼루션 신경망(CNN)을 적용하여 객체를 분류한 뒤, 객체의 위치를 보정하는 방법이다. R-CNN 은 세 가지 기법 중에서 가장 높은 정확도를 가지고 있지만, 모든 이미지의 바운딩 박스에 대해 CNN 을 돌려야 하기 때문에 시간이 상대적으로 오래 걸린다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

SSD 는 한 번만 입력 이미지에 대해 CNN 을 실행하고 feature map 을 계산한다. 바운딩 박스 및 객체 분류 확률을 예측하기 위해 해당 feature map 에 대해 3 x 3 크기의 필터로 CNN 을 적용한다. CNN 연산 후, 바운딩 박스를 예측하는 방식이다. SSD 는 객체 검출 속도와 정확도가 균형을 잘 이루고 있는 탐색 기법이다.

YOLO 는 이미지를 한번에 스캔하는 방식으로 객체를 인식하고, 뉴럴 네트워크 프레임워크인 다크넷을 기반으로 동작한다. YOLO 는 각 이미지를 N x N 개의 그리드로 분할하고, 그리드의 신뢰도를 계산한다. 신뢰도는 그리드 내 객체 인식 시 정확성을 반영한다. 아래 그림과 같이 처음에는 객체 인식과는 동떨어진 바운딩 박스가 설정되지만, 신뢰도를 계산하여 바운딩 박스의 위치를 조정함으로써, 가장 높은 객체 인식 정확성을 가지는 바운딩 박스를 얻을 수 있다.

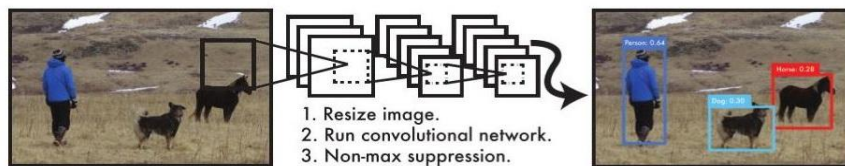
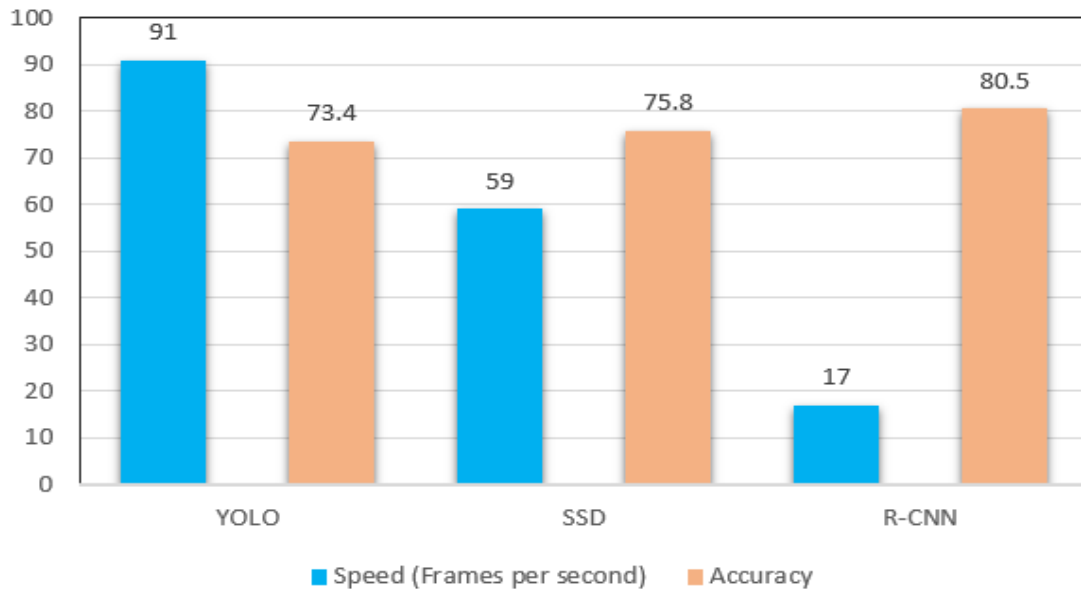


Figure 1: The YOLO Detection System. Processing images with YOLO is simple and straightforward. Our system (1) resizes the input image to 448×448 , (2) runs a single convolutional network on the image, and (3) thresholds the resulting detections by the model's confidence.

<그림> The YOLO Detection System ²

² Joseph Redmon, "You Only Look Once: Unified, Real-Time Object Detection", The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, p.1

아래의 자료는 위의 객체 탐색 기법의 정확도와 속도를 비교한 도표이다.



<그림> 딥러닝 기반 객체 탐색 기법들의 정확도와 속도 비교

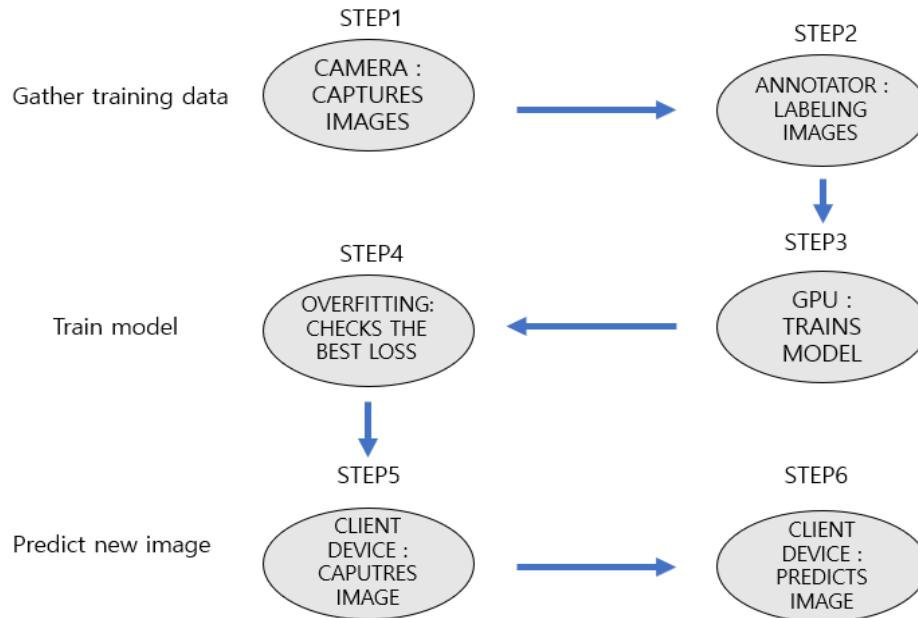
본 프로젝트에서는 실시간으로 사물을 빠르게 인식할 수 있는 모델을 필요로 하기 때문에, 모델 선정에 있어서, 정확도와 속도 모두 중요한 항목이지만 속도를 최우선으로 고려하였다. 따라서 속도가 가장 빠르며 평균이상의 정확도를 확보할 수 있는 YOLO 를 선택하였다.

2)데이터 훈련 및 인식

본 프로젝트에서 사용할 모델은 TINY YOLO V3 로, 일반적인 YOLO 모델이 요구하는 computing power 가 프로젝트 구현에 사용할 라즈베리파이의 성능보다 훨씬 높기 때문에 요구 성능도 낮고 용량도 훨씬 적은 TINY YOLO V3 모델이 적합할 것으로 판단되어 해당 모델을 채택하였다. (YOLO V3 모델은 최소 4GB 의 RAM 을 요구하고, TINY YOLO V3 모델은 1GB 의 RAM 을 요구한다.)

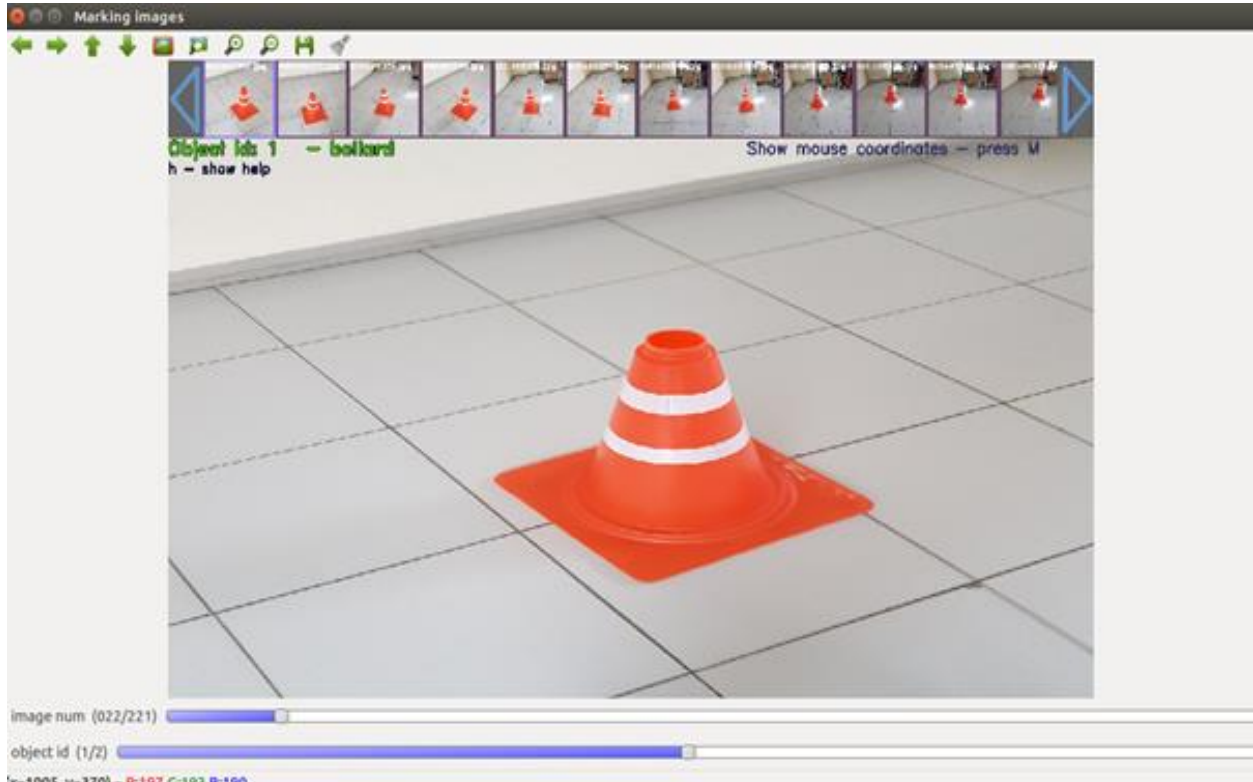
YOLO 모델을 통해 데이터를 인식하는 과정은 다음과 같이 크게 데이터 수집, 학습, 예측으로 구분할 수 있다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28



<그림> 데이터 학습 및 객체 탐지 과정

YOLO 모델에 이미지 데이터를 학습하는 과정에 시각 장애인이 일상생활을 하며 장애가 될 만한 요소들을 선별한다. 이 때, 프로젝트 실제 구현 환경의 제약에 따라 장애물과 유사한 물체들로 대체한다.(예: 볼라드->안전고깔) 또한, 인식률을 높이기 위해 여러 각도, 다양한 배경, 실제 사람의 눈높이 등을 고려하여 이미지를 수집하였다. 이미지 수집 후, 모든 데이터의 확장자를 JPG 로 수정한 후에 Yolo mark 라는 라벨링 프로그램을 이용하여 학습시킬 이미지에서 인식하고자 하는 물체의 범위를 바운딩 박스로 지정하고, 물체의 오브젝트 아이디를 지정하는 라벨링 작업을 진행한다.



<그림> Yolo mark 를 통해 학습시킬 이미지 데이터를 라벨링 하는 과정

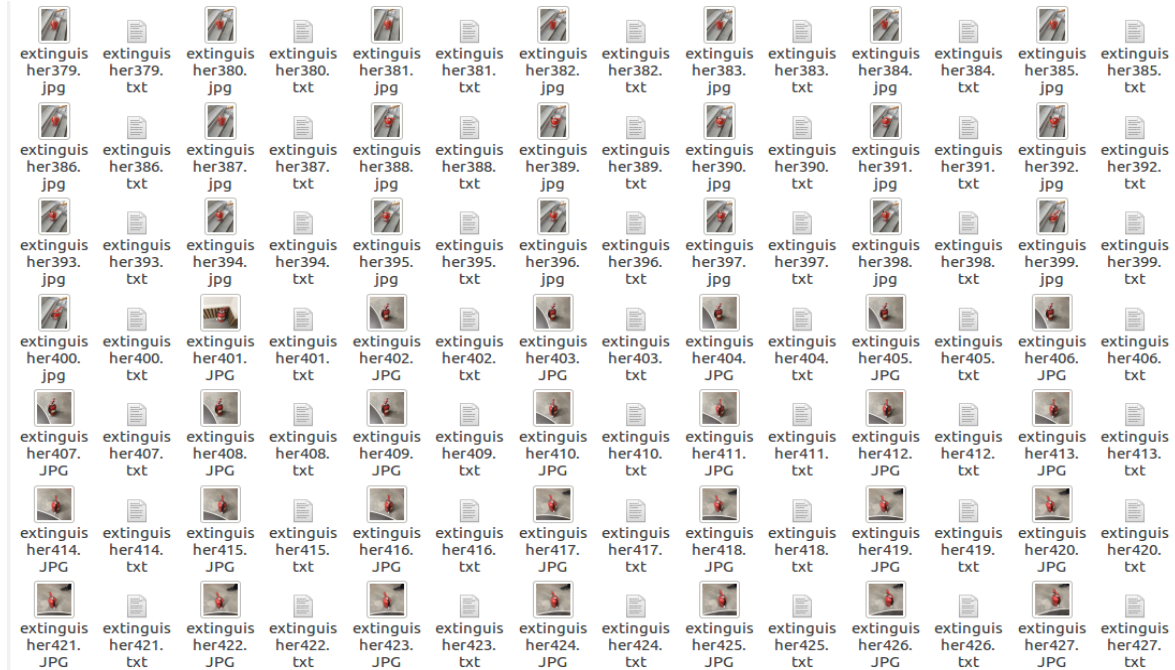
해당 작업이 완료되면 각 이미지 파일에 대해 라벨링 된 id 와 바운딩 박스의 좌표값이 txt 파일 형식으로 생성된다. 생성되는 txt 파일의 형식은 다음과 같다.

```
<object-class> <x_center> <y_center> <width> <height>
```

object-class 에는 id 가 기록되고, x_center, y_center, width, height 에는 각각 바운딩 박스의 중심 좌표와 너비, 높이가 기록된다.

인식률을 높이기 위해 데이터 학습에 약 3400 장의 이미지를 수집하고 (계단 1500 장, 소화기 1000 장, 볼라드 600 장, 손가락 300 장) 해당 데이터에 위와 같은 작업을 실행하였다. GTX 1080 8GB 의 GPU 를 사용하여 overfitting 되기 전까지 학습을 진행하였다. 추가적으로 손가락을 YOLO 모델에 학습시킨 이유는 본 프로젝트에서 손가락 인식에 사용한 OpenCV Haar Cascade 알고리즘과 성능 차이를 비교해보기 위해서이다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28



<그림> 학습에 사용한 이미지 데이터 및 텍스트 파일 중 일부

Loss Function

Localization

$$\lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{obj}} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] \\ + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{obj}} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right]$$

Confidence


$$+ \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 \\ + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2$$

Classification

$$+ \sum_{i=0}^{S^2} \mathbb{1}_i^{\text{obj}} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2$$

<그림> Loss Function Yolo ³

³ Ayoosh Kathuria, "What's new in YOLO v3?", Medium, 2018 년 4 월 23 일 작성, 2018 년 5 월 12 일 접속, <https://towardsdatascience.com/yolo-v3-object-detection-53fb7d3bfe6b>

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

YOLO 에서는 위와 같은 손실 함수를 사용한다. 큰 박스에서의 오차와 작은 박스에서의 오차를 다르게 취급해야 하는데, 이를 다루기 위해 sum-squared error 가 사용된다. Localization Loss 는 실제 바운딩 박스와 예측한 바운딩 박스간의 오차를 의미하고, Confidence Loss 는 해당 바운딩 박스의 신뢰도가 100%에서 얼마나 많은 차이가 발생하는지를 의미한다. 마지막으로 Classification Loss 는 클래스의 실제 확률과 예측된 클래스에 대한 확률 차이를 의미한다. 이 세 가지의 loss 를 모두 합치면 YOLO 에서의 손실 함수 값이 된다. 이 손실 함수의 값을 기준으로 해당 값이 더이상 감소하지 않는 지점, 즉 overfitting 되기 전까지 데이터를 학습시켜 가장 효율적인 모델을 만드는 것이 중요하다.

```
10001: 0.310197, 0.310197 avg, 0.001000 rate, 3.097001 seconds, 640064 images
50001: 0.153419, 0.153419 avg, 0.001000 rate, 2.291412 seconds, 3200064 images
100001: 0.148360, 0.148360 avg, 0.001000 rate, 8.038935 seconds, 6400064 images
110001: 0.122418, 0.122418 avg, 0.001000 rate, 1.412100 seconds, 7040064 images
126860: 0.113058, 0.140726 avg, 0.001000 rate, 0.654941 seconds, 8119040 images
```

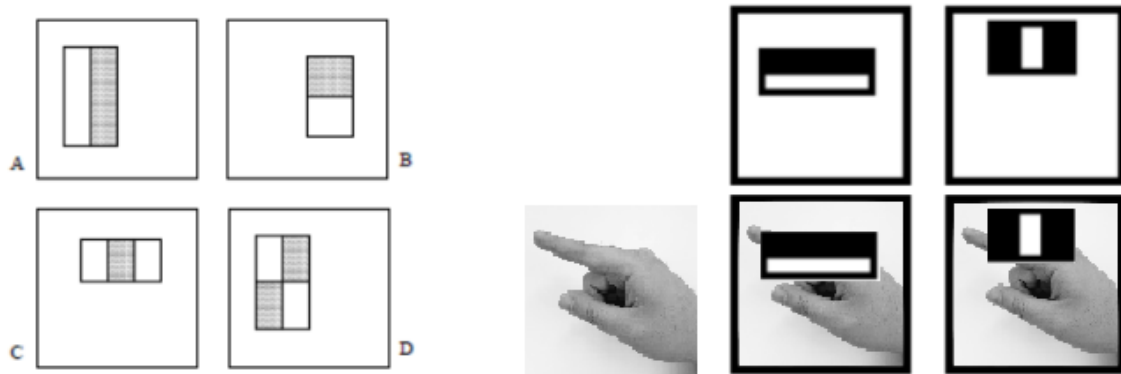
<그림> 학습 횟수에 따른 손실 함수 값의 변화

overfitting 되는 지점을 찾기 위해 5 일동안 약 13 만 번 학습을 진행하였다. 11 만 번 학습시켰을 때 손실 함수의 값이 최저가 되었다가, 12 만 번째부터 점점 손실 함수 값이 증가하는 모습을 볼 수 있었다. 즉, overfitting 현상이 약 12 만 번 학습시켰을 때부터 발생하였다고 판단하여 11 만 번 학습시켰을 때 추출된 가중치 파일을 본 프로젝트에 사용하였다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28


3) OpenCV Haar Cascade

본 프로젝트에서 사용자의 손가락을 가리키는 특정 행동을 인식하기 위해 OpenCV 에서 제공하는 Haar Cascade 알고리즘을 사용한다. Haar Cascade 는 Haar-like feature (유사 하르 특징)로 영상에서의 객체의 특징을 파악하여 탐지하는 기법이다.

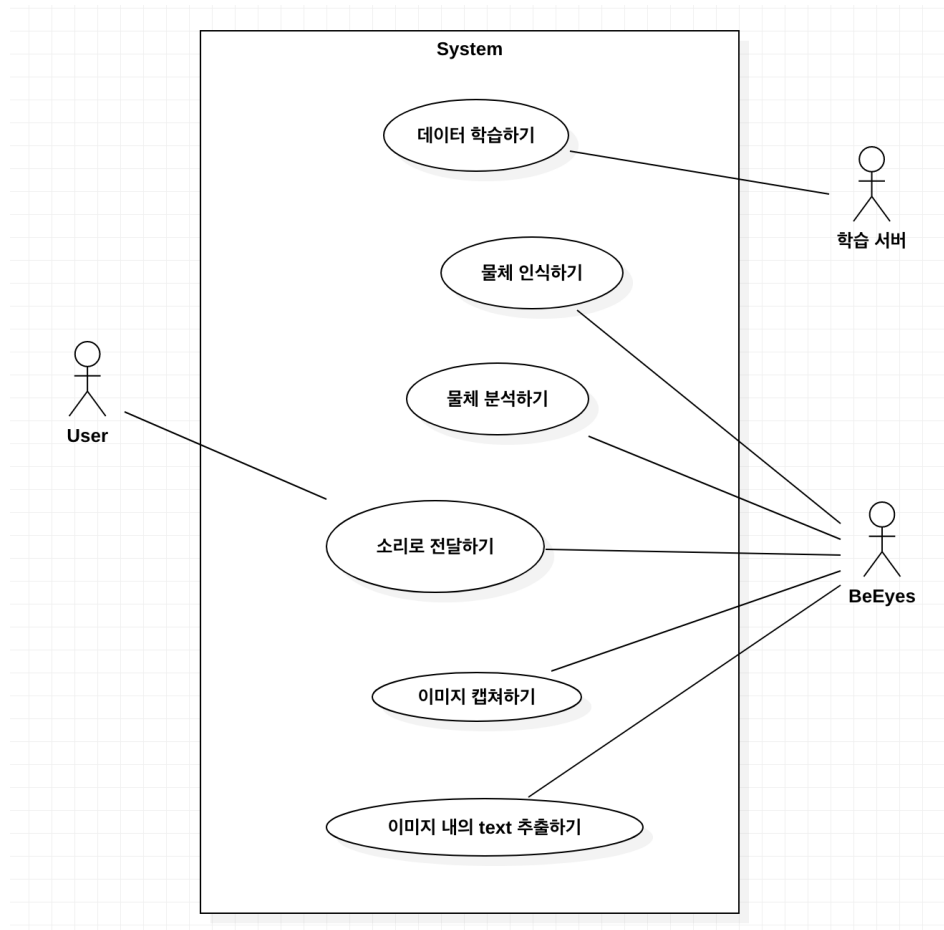


<그림> Haar-like feature (유사 하르 특징)

위 그림과 같이 객체를 검출하기 위한 이미지 위에 흑백의 사각형을 겹쳐 놓은 다음 밝은 영역에 속한 픽셀 값들의 평균에서 어두운 영역에 속한 픽셀 값들의 평균의 차이를 구한 뒤, 그 값이 어떤 임계 값을 넘으면 특정 객체에 대해 유사 하르 특징이 있다고 판단을 내리게 된다. 흑백으로 이루어진 사각형의 크기는 A, B, C, D 와 같이 변하게 되고 그에 따라 밝은 영역에 속한 픽셀 값들의 평균이 바뀌게 된다. 그렇기 때문에 사람의 손은 다양하지만 그 생김새의 패턴은 비슷하므로 손의 특정 위치, 그리고 특정 영역의 픽셀 값들의 분포도 크게 차이가 없을 것이라고 생각할 수 있다. 결론적으로 그 값이 곧 임계 값이 된다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

2.2.2 시스템 기능 요구사항



<그림> Use case Diagram

기능적인 요구사항	내용	상태
데이터 학습하기	인식할 데이터 (볼라드, 소화기, 계단)를 서버에서 학습시킨다. 단, 프로젝트 실제 구현 환경의 제약에 따라 장애물과 유사한 물체들로 대체한다. (예: 볼라드->안전고깔)	완료
물체 인식하기	Be Eyes 의 카메라 모듈을 통해 물체를 인식한다.	완료

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

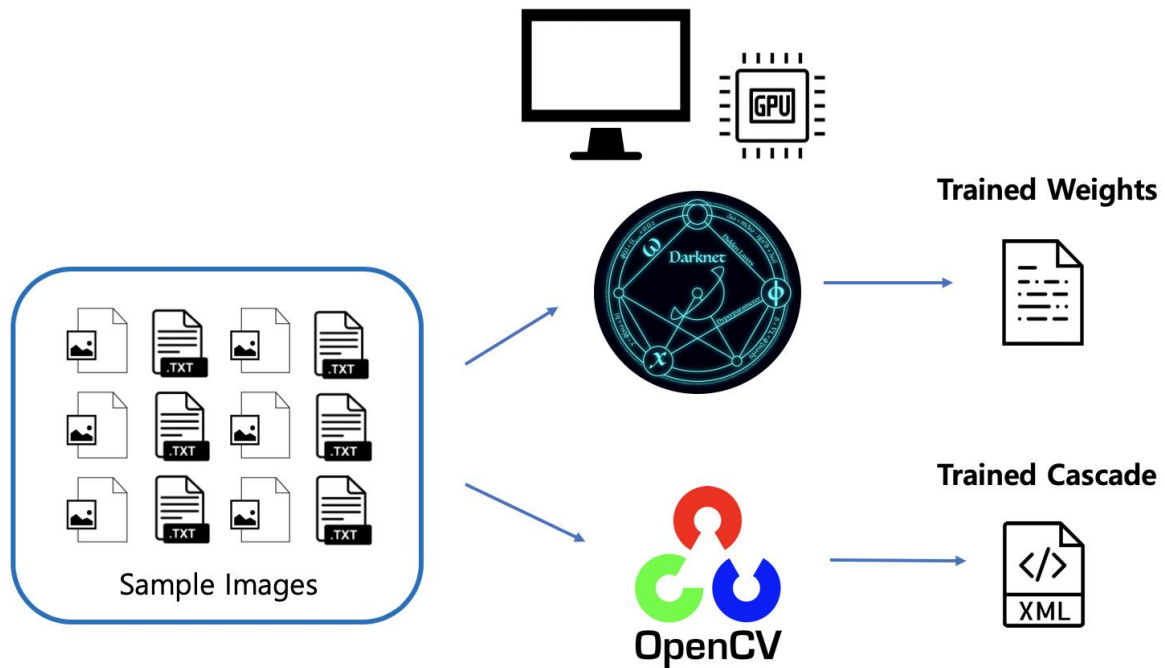
물체 분석하기	실시간으로 물체를 분석한다.	완료
소리로 전달하기	분석한 물체 또는 텍스트를 Be Eyes 의 음성 모듈을 통해 사용자에게 전달한다.	완료
이미지 캡처하기	검지 손가락 인식 후 해당 화면을 캡처한다.	완료
이미지 내의 text 추출하기	캡처된 이미지 내의 text 를 추출한다.	완료

2.2.3 시스템 비기능(품질) 요구사항

품질 속성	품질 요구사항	내용	측정 요소	상태
Performance(성능) - 1 초 동안 보여주는 화면의 수가 6 프레임 이상으로 탐지 가능하다.	실시간 물체 분석	시각장애인에게 실질적으로 도움이 될 수 있도록 영상에 대한 정보를 분석하고 전달하는 과정이 실시간으로 진행되어야 한다.	Detection Speed 가 6FPS 이상 나오는지 확인	달성
Reliability(신뢰성) - 정확도 90% 이상인 결과를 사용자에게 전달한다.	물체 분석 정확성	시각장애인이 사물을 직접 확인 할 수 없으므로 물체 분석의 정확도가 높아야 한다.	장애물 인식에 대한 정확도가 90% 이상 나오는지 확인	달성
Usability(사용성) - 0.5 초만에 결과를 사용자에게 소리로 전달한다.	정보 전달 용이성	사용자는 실시간으로 정보를 소리로 전달받아야 한다.	0.5 초 안에 소리가 출력되는지 확인	달성
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	미달성 - 향후 발전 방향 제시

 <div> <div>국민대학교</div> <div>컴퓨터공학부</div> <div>캡스톤 디자인 I</div> </div>	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

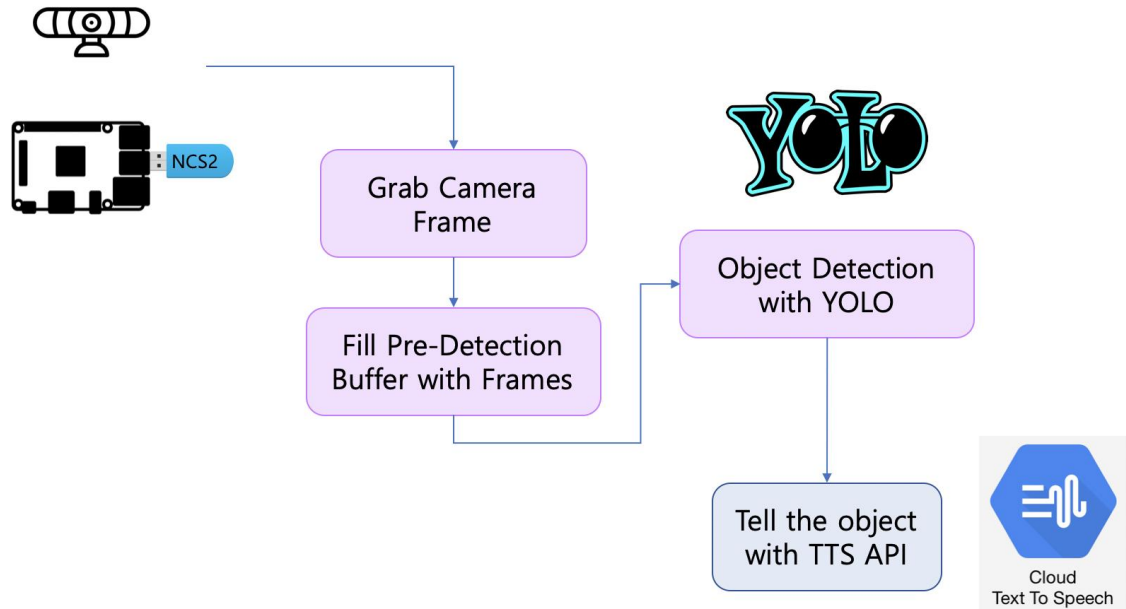
2.2.4 시스템 구조 및 설계도



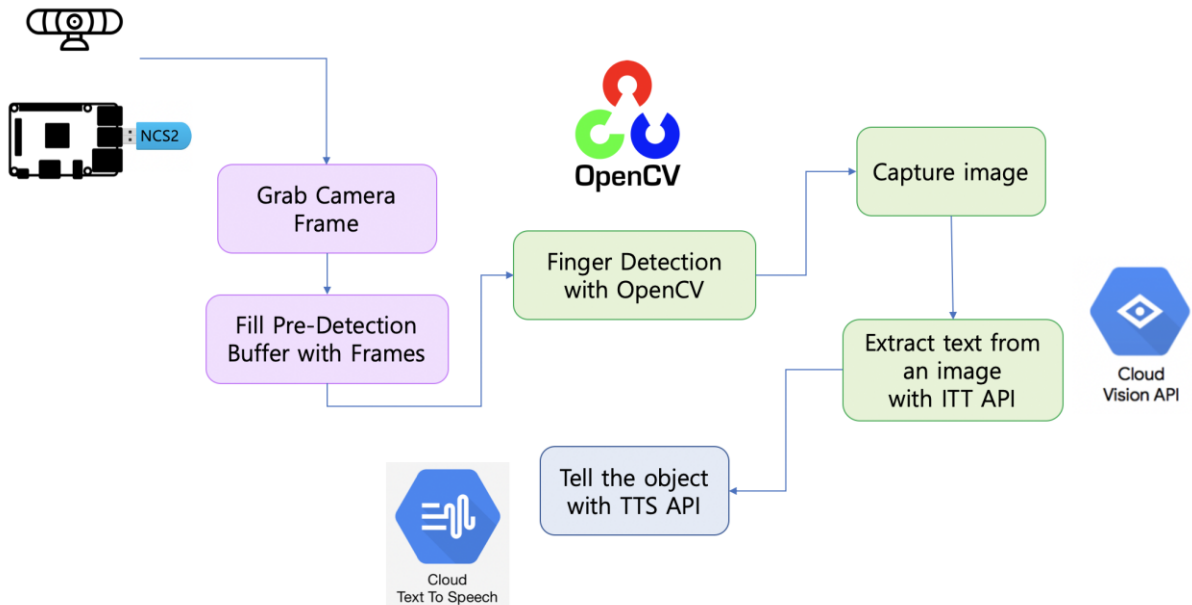
<그림> "Be Eyes" 데이터 훈련

본 프로젝트에서 데이터 훈련은 위와 같이 이루어진다. 장애물에 대한 이미지 데이터는 다크넷 YOLO 를 이용해 학습하여 weight 확장자를 가진 가중치 파일을 추출하는 데에 사용하고, 손가락에 대한 이미지 데이터는 긍정 이미지(손가락)와 부정 이미지(손가락 제외) 데이터 셋을 수집 후 OpenCV 의 haarTraining 을 사용하여 학습시켜 xml 파일을 추출하는 데 사용한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28



<그림> '기능 1'의 시스템 구조



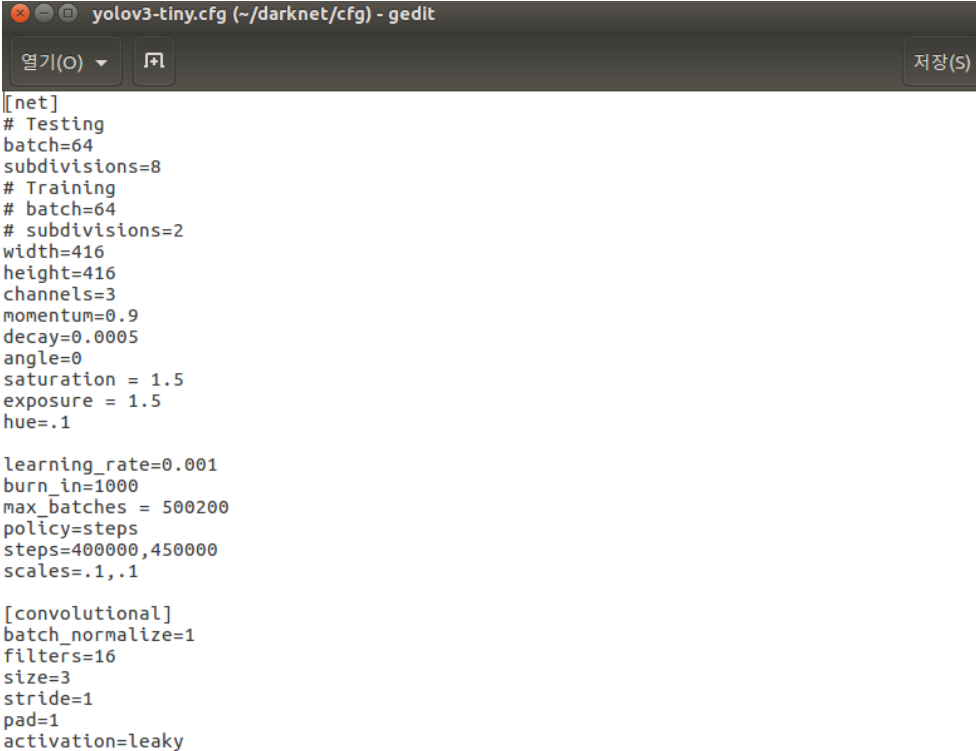
<그림> '기능 2'의 시스템 구조

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

2.2.5 활용/개발된 기술

1) 다크넷 YOLO

본 프로젝트에서는 신속한 장애물 탐지를 위해 YOLO 를 이용하여 객체 탐지 및 추적을 수행한다. YOLO 를 실행시키기 위해서는 다크넷이 필요한데, 여기서 다크넷은 신경망들을 학습시키고 실행시킬 수 있는 신경망 프레임워크를 의미한다. YOLO 를 실행시키기 위해서는 다크넷에서 YOLO 의 cfg 파일과 weight 파일을 불러와야 한다. cfg 파일은 신경망의 구조 (layer 의 개수, 입력 데이터의 구조 등)를 명시한 파일이고, weight 파일은 실제로 학습된 신경망의 가중치 값들을 저장한 파일이다.



```

yolov3-tiny.cfg (~/.darknet/cfg) - gedit
열기(O)  저장(S)

[net]
# Testing
batch=64
subdivisions=8
# Training
# batch=64
# subdivisions=2
width=416
height=416
channels=3
momentum=0.9
decay=0.0005
angle=0
saturation = 1.5
exposure = 1.5
hue=.1

learning_rate=0.001
burn_in=1000
max_batches = 500200
policy=steps
steps=400000,450000
scales=.1,.1

[convolutional]
batch_normalize=1
filters=16
size=3
stride=1
pad=1
activation=leaky

```

<그림> cfg 파일의 내용 중 일부

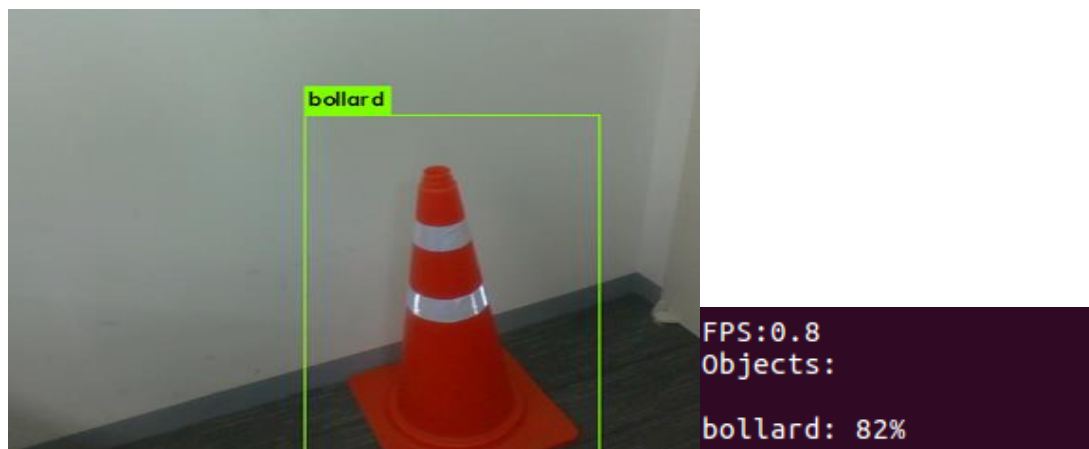
 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

cfg 파일에 기술되어 있는 하이퍼 파라미터를 변경하여 다양한 형태의 학습을 진행할 수 있다. 학습시킨 장애물은 총 3 가지(계단, 소화기, 안전고깔)로 구성하였으며, 학습시킨 이미지 데이터 수는 계단이 1500 장, 소화기 1000 장, 안전고깔 600 장이다. 본 프로젝트에 사용한 YOLO 모델의 구조는 다음과 같다.


layer	filters	size	input	output	
0 conv	16	3 x 3 / 1	416 x 416 x 3	-> 416 x 416 x 16	0.150 BFLOPs
1 max		2 x 2 / 2	416 x 416 x 16	-> 208 x 208 x 16	
2 conv	32	3 x 3 / 1	208 x 208 x 16	-> 208 x 208 x 32	0.399 BFLOPs
3 max		2 x 2 / 2	208 x 208 x 32	-> 104 x 104 x 32	
4 conv	64	3 x 3 / 1	104 x 104 x 32	-> 104 x 104 x 64	0.399 BFLOPs
5 max		2 x 2 / 2	104 x 104 x 64	-> 52 x 52 x 64	
6 conv	128	3 x 3 / 1	52 x 52 x 64	-> 52 x 52 x 128	0.399 BFLOPs
7 max		2 x 2 / 2	52 x 52 x 128	-> 26 x 26 x 128	
8 conv	256	3 x 3 / 1	26 x 26 x 128	-> 26 x 26 x 256	0.399 BFLOPs
9 max		2 x 2 / 2	26 x 26 x 256	-> 13 x 13 x 256	
10 conv	512	3 x 3 / 1	13 x 13 x 256	-> 13 x 13 x 512	0.399 BFLOPs
11 max		2 x 2 / 1	13 x 13 x 512	-> 13 x 13 x 512	
12 conv	1024	3 x 3 / 1	13 x 13 x 512	-> 13 x 13 x 1024	1.595 BFLOPs
13 conv	256	1 x 1 / 1	13 x 13 x 1024	-> 13 x 13 x 256	0.089 BFLOPs
14 conv	512	3 x 3 / 1	13 x 13 x 256	-> 13 x 13 x 512	0.399 BFLOPs
15 conv	27	1 x 1 / 1	13 x 13 x 512	-> 13 x 13 x 27	0.005 BFLOPs
16 yolo					
17 route	13				
18 conv	128	1 x 1 / 1	13 x 13 x 256	-> 13 x 13 x 128	0.011 BFLOPs
19 upsample		2x	13 x 13 x 128	-> 26 x 26 x 128	
20 route	19 8				
21 conv	256	3 x 3 / 1	26 x 26 x 384	-> 26 x 26 x 256	1.196 BFLOPs
22 conv	27	1 x 1 / 1	26 x 26 x 256	-> 26 x 26 x 27	0.009 BFLOPs
23 yolo					

<그림> YOLO 모델의 구조

장애물이 있는 영상을 입력 받으면 학습시킨 모델을 통해 다음과 같은 결과를 얻을 수 있다.



<그림> 학습시킨 장애물에 대한 인식 결과

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28


2) OpenVINO tiny-YOLO v3 with NCS2

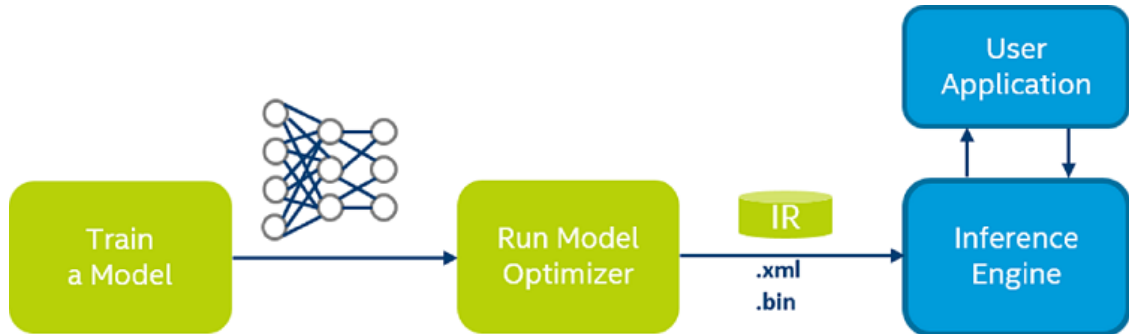
본 프로젝트는 라즈베리파이 환경에서 대표적인 단일 단계 방식의 객체 탐지 모델인 YOLO 를 통해 객체 인식 및 추적을 구현한다. 그러나 다크넷 프레임워크는 라즈베리파이에서 실행 시 전력 소모가 크고 요구 사양이 높아 딥러닝 모델을 정상적으로 실행하지 못하였다.



<그림> USB 형 딥러닝 장치, Neural Compute Stick 2(NCS2)

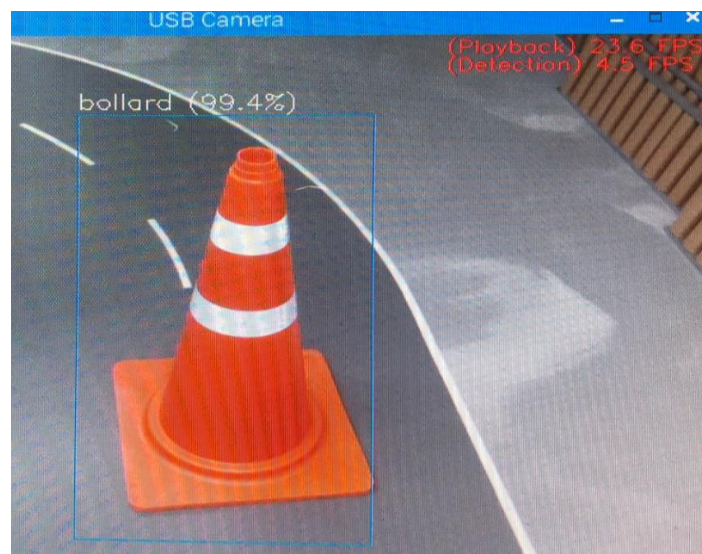
따라서 위 그림과 같이 Movidius™ VPU (비주얼 처리 장치)에 의해 저전력, 고성능으로 구동되는 딥러닝용 USB 드라이브인 Intel® Movidius™ NCS2(Neural Compute Stick 2)를 사용한다. 이 USB 는 개방형 시각 추론 및 신경망 최적화 도구 모음인 OpenVINO 라는 Toolkit 을 제공한다. 그러므로 데이터 학습은 다크넷에서 진행하고 학습된 모델을 OpenVINO 도구를 통해 구동한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28



<그림> 라즈베리파이에서 객체 탐지가 가능한 파일 변환 과정

모델의 학습은 데스크탑(GPU : GTX1080 8GB)에서 별도로 수행한다. 다크넷에서 제공하는 학습 도구 Yolo Mark 를 사용하여 학습 모델 파일(weight)을 생성한다. 여기서 학습된 모델 파일은 NCS2 USB 를 사용할 수 있는 OpenVINO 도구에서 지원하지 않기 때문에 실행되지 않는다. 따라서 위의 그림과 같이 OpenVINO 에서 사용할 수 있도록 최적화된 학습 모델로 변환하는 작업이 필요하다. Model Optimizer 란 최적의 실행이 가능하도록 딥러닝 모델을 자동으로 조정해주는 OpenVINO 의 cross-platform 명령 도구이다. 이 과정을 통해 학습한 모델을 Intermediate Representation(IR)인 xml 과 bin 파일로 변환한다. 이렇게 변환한 모델 파일을 사용하여 라즈베리파이 환경에서 NCS2 를 사용한 실질적인 YOLO 알고리즘 구현이 가능하다.



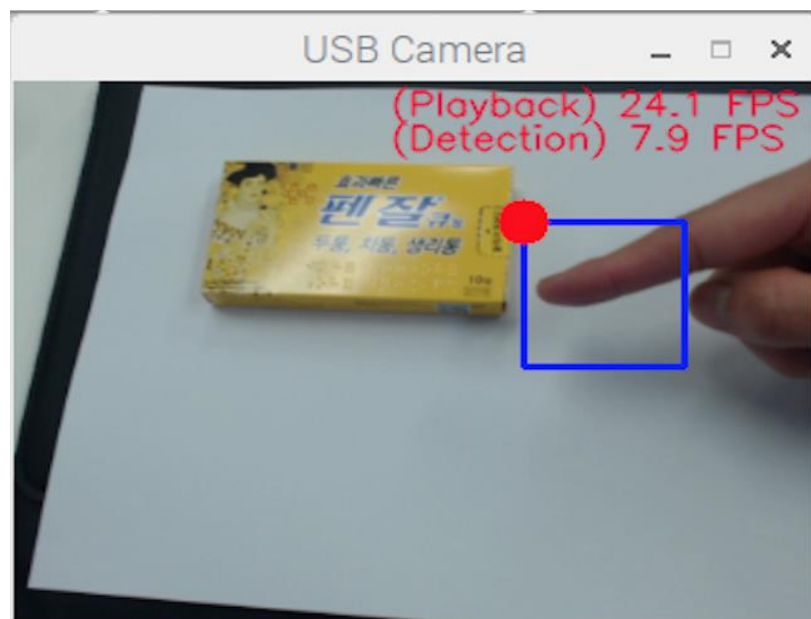
<그림> 라즈베리파이 + Neural Compute Stick 2, 학습 데이터 인식

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28


FPS(Frame Per Second, 초당 프레임)는 초당 처리할 수 있는 프레임 수를 나타내며 이 수치가 높을수록 물체를 인식하는 속도가 향상된다. 위 그림과 같이 학습한 안전고깔 데이터를 약 4~8 FPS 로 인식한다. 라즈베리파이에서 다크넷을 사용했을 때는 화면이 멈추는 오류 현상 및 전력 부족 표시가 나타나면서 전혀 실행이 되지 않았다. 또한 ubuntu 16.04 LTS 환경에서 다크넷을 사용했을 때는 FPS의 범위가 0.6~0.8 이었는데 이와 비교했을 때 약 7~10 배 정도의 성능이 향상된 것을 볼 수 있다.

3) OpenCV의 Haar Cascade 알고리즘

본 프로젝트에서 사용자가 읽고자 하는 의사를 표현하기 위한 목적으로 손가락을 가리키는 행동을 인식하기 위해 Haar Cascade 알고리즘을 사용한다. 손가락이 인식되면, 특정 영역에 존재하는 영역의 글자를 읽어준다.

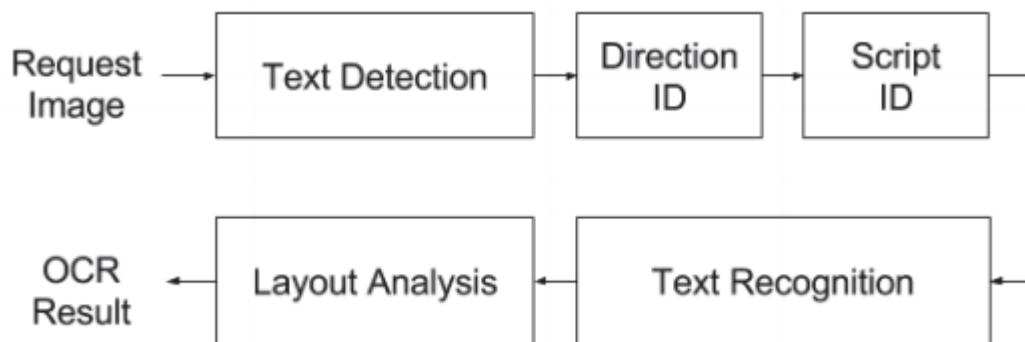


<그림> Haar Cascade 알고리즘을 이용한 손가락 인식

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

4) Google Cloud Platform 의 Vision API

Google Cloud 의 Vision API 는 머신 러닝 기술을 이용하여 이미지를 분석해주는 기능을 제공한다. 이 기능을 활용하여 사용자가 손가락으로 읽고자 하는 부분을 가리키면, 해당 영상을 캡처하여 사진으로 저장 후 Vision API 를 통해 분석하여 사진의 특정 영역에 존재하는 글자를 string 형태로 전달한다.



<그림> Google Cloud Vision API Architecture ⁴

5) Google Cloud Platform 의 Text-To-Speech API

Google Cloud 의 Text-To-Speech API 는 입력 받은 문자를 음성으로 변환하여 출력해주는 기능을 제공한다. Vision API 를 통해 추출된 문자열의 입력을 확인 후 연결된 스피커를 통해 해당 문자열에 대한 음성이 출력된다. 이 기능은 장애물 인식, 글자 인식에 모두 적용되며, 장애물 인식의 경우에는 전방에 장애물이 식별되었을 때 전달되는 문자열 형식의 경고문을 읽어주며, 글자 인식의 경우 Vision API 를 통해 추출된 문자열을 읽어준다.

⁴ Edward Ma, "Secret of Google Web-Based OCR Service", 2019 년 1 월 13 일 작성, 2019 년 5 월 23 일 접속, <https://towardsdatascience.com/secret-of-google-web-based-ocr-service-fe30eeceddd01>

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

tts process



<그림> TTS (Text To Speech) Process ⁵


TTS API 를 통해 다음과 같이 전송된 문자열을 음성으로 사용자에게 전달한다.

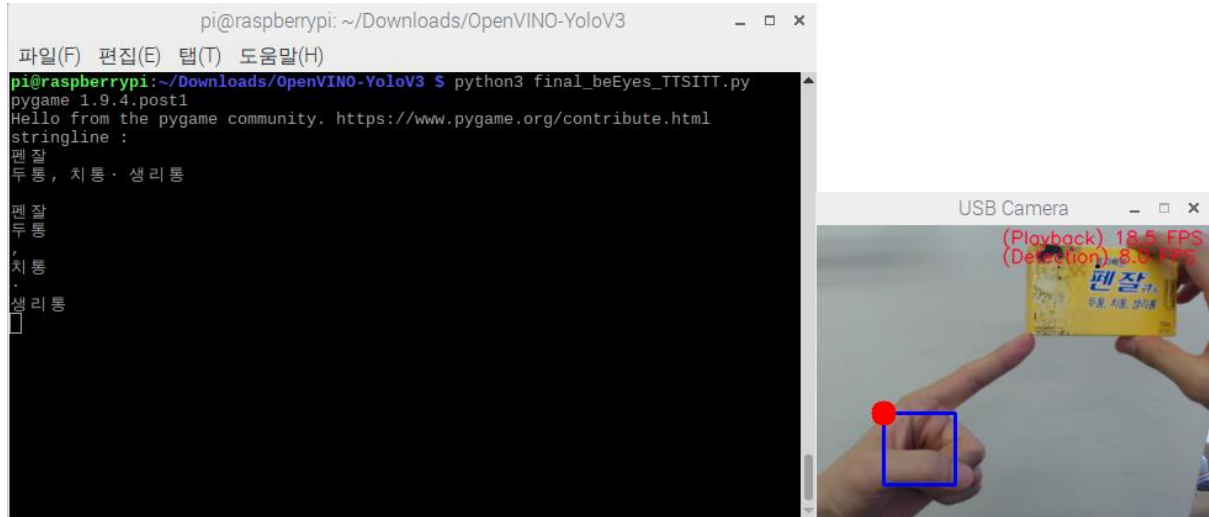
```

pi@raspberrypi: ~/Downloads/OpenVINO-YoloV3
파일(F) 편집(E) 탭(T) 도움말(H)
[setupvars.sh] OpenVINO environment initialized
pi@raspberrypi:~ $ cd Downloads/OpenVINO-YoloV3/
pi@raspberrypi:~/Downloads/OpenVINO-YoloV3 $ python3 ttsitt.py
pygame 1.9.4.post1
Hello from the pygame community. https://www.pygame.org/contribute.html
전방에 안전코갈 1개 있습니다. 조심하세요.
  
```

<그림> 장애물 식별 시 출력되는 문구

⁵ IVONA Software, “TTS Process”, An Amazon Company (2019), 2019 년 3 월 24 일 접속,
<https://www.ivona.com/us/about-us/text-to-speech/>

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28




<그림> 손가락이 가리키는 문구 출력

2.2.6 현실적 제한 요소 및 그 해결 방안

1. 하드웨어

- 1) 라즈베리파이를 통해 실시간으로 영상 처리와 객체 판별을 하는 것은 많은 양의 전력을 소모하고 충분한 연산이 수행되어야 한다. 따라서 Movidius™ VPU(비주얼 처리 장치)에 의해 저전력, 고성능으로 구동되는 딥러닝용 USB 드라이브인 Intel® Movidius™ NCS2(Neural Compute Stick 2)를 사용하여 해결하였다.
- 2) 카메라의 시야 안에 물체가 인식되지만 사용자의 보행에 직접적으로 장애가 되지 않는 경우를 대비하여 카메라의 위치를 사용자의 신체에 맞게 조절이 가능하도록 하여 직접적으로 장애가 되는 물체만 인식하게 하였다.
- 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) 웨어러블 장비로 제작되기 때문에 라즈베리파이에 대한 전원공급과 전력문제를 해결하기 위해 5v3a 보조 배터리를 이용한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

2. 소프트웨어

- 1) 프로젝트 기간 내에 학습시킬 데이터의 개수에 제한이 생기므로 데이터의 개수를 3 개(안전고깔, 소화기, 계단)로 제한한다. 추후에 장애요소를 추가적으로 선별하여 클래스의 개수를 확장할 수 있다.
- 2) 손가락을 인식하기 위해 단순히 피부색을 이용하는 것은 많은 장애요소가 발생한다. 따라서 OpenCV 의 cascade 분류기를 이용하여 손가락 데이터를 모아 학습하였고 성공적으로 손가락을 검출할 수 있었다.
- 3) 문구를 인식할 때, 범위를 지정하였기 때문에 해당 범위에 글자가 포함되지 않으면 인식하지 못하는 경우가 생길 수 있다. 따라서 글자가 인식되지 않은 경우 사용자가 스스로 손가락의 위치를 재조정할 수 있도록 하기 위하여 사용자에게 음성으로 텍스트를 인식할 수 없다는 음성 메시지를 출력해준다.
- 4) 영상이 출력되는 USB 카메라 사이즈를 재조정하여 Detection FPS 를 증가하였다.

3. 기타

- 1) 안경이나 선글라스에 부착한 디바이스의 크기나 무게가 사용자들이 사용하기에 적합하지 않다. 따라서 안경이 아니더라도 모자나 옷 일부에 탈 부착할 수 있게 소형 웨어러블 기기 형태로 제작하는 것을 고려한다.
- 2) 지정된 장애요소 중 하나인 계단은 구현 환경의 제약에 따라 시연영상으로 대체한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

2.2.7 결과물 목록

a) H/W



<그림> 'Be Eyes' 스마트 글라스



<그림> 'Raspberry Pi 3 B+' + 'Neural Compute Stick2'

b) S/W

- 인공지능을 이용한 오브젝트 디텍션 기능 소프트웨어
- TTS & ITT 기능 소프트웨어

c) 기타

- 학습 모델 가중치 파일 (.weights)
- 모델 최적화(Model Optimizer) 과정을 거친 파일(.xml, .bin, .mapping)

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

2.3 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. 이에 따라 시각 장애인을 위한 안경 타입의 스마트 기기는 공익 차원에서도 시장성이 풍부한 분야이다. “Be Eyes”는 시각장애인들이 일상생활에서 마주할 수 있는 장애물에 대한 정보를 인지하고 스스로 위험을 예방할 수 있도록 제작하여 착용자들의 안전을 보장한다. 식료품이나 의약품 표시 정보를 음성으로 전달하여 정확하게 알 수 있으므로 시각장애인들의 불안과 걱정도 덜어줄 수 있다. 특히, 시각장애인의 대부분을 차지하고 있는 저시력 시각장애인들에게 큰 도움을 줄 것이라 예상된다.

본 프로젝트에서는 “Be Eyes”를 스마트 안경으로 제작하지만 안경뿐만 아니라 원하는 곳에 기기를 탈 부착할 수 있도록 웨어러블 형태로 제작하는 것을 향후에 고려하여 착용자들에게 편리한 사용성을 제공할 수 있다. 또한 인공지능을 기반으로 한 제품이기 때문에 본 프로젝트에서 선별한 요소들 외에 장애물들을 추가적으로 학습시켜 확장성을 보장할 수 있다는 큰 장점이 있다. 동시에 연결된 카메라를 스테레오 카메라로 변경하면 물체의 거리에 대한 정보까지 제공할 수 있기 때문에 시각장애인이 더 구체적인 판단을 할 수 있다.


“Be Eyes”를 제작하여 많은 사람들에게 공급할 수 있다면 사용자들에게 신체적 한계를 극복했다는 의미 있는 경험과 삶의 원동력을 제공할 것이다. 더불어 사회적, 상업적인 가치를 창출할 수 있다.

3 자기평가

구성원 이름	평가 내용
박병훈	본 프로젝트에서 목표로 한 시각장애인을 위한 스마트글라스 Be Eyes 제작은 성공적이다. 목표한 두 가지 기능 모두 시연 가능하고 실제 착용 후 데모 환경에서 테스트를 해본 결과 기능이 정상적으로 작동되는 것을 확인할 수 있었다. 프로젝트를 진행하며 나는 프로젝트 리더의 역할을 맡아 팀을 이끌었고 데이터 수집과 전처리, 학습 및 테스트를 담당하였다. 또한 OpenCV haarcascade 를 이용하여 손가락 인식이 가능한 xml

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

	<p>파일을 제작하였다.</p> <p>이 프로젝트는 발전가능성이 무궁무진하다. 만약 시간과 비용이 허락했다면 더 많은 장애요소를 추가할 수 있었고 거리센서를 이용하여 구체적인 시각 장애인의 시야 역할을 할 수 있었을 것이다. 또한 성능이 좋고 가벼운 하드웨어로 대체 가능하여 실제 시각장애인이 안전한 일상생활을 하는데 도움을 줄 수 있을 것이라고 확신한다.</p> <p>프로젝트를 진행하며 경험이 없던 인공지능과 영상처리에 대한 지식을 쌓을 수 있었다. 약 3 개월이라는 시간동안 한 분야에 대한 깊은 연구를 할 순 없었지만 새로운 기술과 경험을 할 수 있었던 좋은 기회였다고 생각한다.</p>
김상민	<p>프로젝트를 효율적으로 관리하고 설정한 목표를 시간 내에 달성하기 위해 To Do List 를 작성하여 우선순위를 정리하였다. 구글 드라이브(Google Drive)와 깃허브를 활용하여 자료 및 문서를 공유하고 슬랙을 통해 담당 교수님과 지속적으로 컨택하고 피드백을 받으면서 업무의 효율을 높였다.</p> <p>본 프로젝트를 진행하면서 나는 라즈베리파이에서 Neural Compute Stick2(NCS2)를 사용하기 위한 환경을 설정하고 USB 규칙을 추가하는 업무를 수행하였다. 또한, OpenVINO 도구를 사용하여 학습된 모델을 최적화하는 작업을 수행하고 Google Cloud Platform 의 TTS 및 ITT API 를 Yolo 알고리즘에 적용하였다. Yolo 알고리즘 코드를 분석하고, 여러 개의 객체를 동시에 인식했을 경우를 고려하여 객체의 이름뿐만 아니라 개수도 각각 출력할 수 있도록 프로그래밍하여 사용자 편의성을 향상시켰다. 최종 평가를 대비하여 프로젝트 영상을 편집하고 팀원들과 스마트 글라스 "Be Eyes" 장치 및 데모 환경을 직접 제작하였다.</p> <p>이번 캡스톤 디자인 프로젝트를 진행하면서 나는 짧은 시간에 분석하고 이해하는 태도에 초점을 맞추었다. 전문 분야에서의 역량, 소통 및 협업 능력을 키울 수 있었고 새로운 기술을 경험할 수 있는 좋은 기회였다.</p>
안은영	<p>시각장애인을 도울 수 있는 프로젝트를 하게 되어 더 의미가 있었던 것 같다. 두가지 기능을 한학기 내에 모두 구현하기에 기간이 짧아 걱정이 되었는데 딥러닝 yolo 모델과 OpenCV 의 harr cascade 알고리즘, 구글에서 제공하는 API 를 이용하여 한 학기 내에 목표한 결과를 다 이룰 수 있었다. 인공지능은 잘 알지 못하는 분야라 어려움도 많이 있었지만 프로젝트를 수행하면서 많이 배울 수 있었다. 학습한 모델 최적화를 조사하고 수행하였고, 손가락 인식을 위한 알고리즘을 분석하고 필요한 부분을 추가하여 구현하였다.</p> <p>초기 하드웨어 선택시 딥러닝에 적합한 젯슨 nano 의 존재를 모르고 라즈베리파이에 딥러닝용 USB 를 결합하는 것을 선택하여 더 좋은 성능을 내지 못한 점이 아쉬운 점으로 남는다. .</p>

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

서준교	<p>이번 프로젝트에서 나는 학습 데이터의 수집 및 전처리 (경로 설정, 라벨링 등)와 데이터 학습, 그리고 회의록 작성을 담당하였다. 최대한 다양한 각도 및 배경을 바탕으로 한 이미지를 수집하여 학습을 시켰고, 가장 인식률이 높게 나온 가중치 모델을 추출 및 적용하였다. 해당 프로젝트를 진행하면서, 전반적인 딥러닝의 이론을 이해할 수 있었고 수많은 시행착오를 겪으며 최적의 데이터를 수집하고 학습시키는 과정에서 인식률에 영향을 주는 요인을 스스로 파악하고 이를 최대한 제어하는 방법을 익힐 수 있었다. 딥러닝에 최적화된 하드웨어에서 해당 프로젝트의 기능을 구현한다면, 더 빠른 성능을 보일 것으로 기대되며, 추가적인 기능도 고려할 수 있다고 생각한다. 하지만 나는 이 프로젝트가 '시각장애인들의 안전한 일상생활' 이라는 목표 달성에 한걸음 내딛는 데에 충분한 역할을 했다고 확신한다. 추후에 기회가 된다면, 데이터 수집 및 학습 과정을 자동화하여 시각장애인이 실제로 사용해도 지장이 없을 정도의 완성도 있는 제품을 만들고 싶다.</p>
-----	---

4 참고 문헌

번호	종류	제목	출처	발행년도	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	오윤환, 이은주	
6	논문	You Only Look Once: Unified, Real-time Object Detection	CVPR 2016 paper	2016	Joseph Redmon	언어 (ENG)

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

7	논문	A Survey On Moving Object Detection and Tracking Methods	Semantic Scholar	2014	Himani S, Parekh	언어 (ENG)
8	기사	[취재파일] 시각장애인은 안중에도 없는 서울시의 시각장애인 사업	SBS NEWS	2017	한지연	

5 부록

5.1 사용자 매뉴얼

5.1.1 디바이스 시작

1. 라즈베리파이에 배터리 팩을 연결한다.
2. 안경을 착용하고, 카메라의 위치를 자신에게 맞게 조정한다.
3. "Be Eyes 실행합니다." 라는 음성과 함께 프로그램이 실행된다.

5.1.2 장애물 인식

1. 사용자의 전방에 특정 장애물이 감지되면 장애물의 이름과 경고 메시지를 출력한다.
 - 인식 가능한 장애물 종류: 안전고깔, 소화기, 계단
2. 장애물이 여러 개가 감지되면, 해당 장애물의 이름과 개수가 출력된다.
3. 사용자는 경고 메시지를 듣고 이후의 행동을 스스로 판단할 수 있다

5.1.3 글자 인식

1. 사용자가 읽기를 원하는 문구가 있는 부분의 하단을 검지손가락으로 가리킨다.
2. 손가락이 인식되면 해당 문구를 두 번 출력해준다.
3. 해당 문구를 모두 읽었으면 "글자를 모두 읽었습니다." 라는 음성이 출력된다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28


5.2 운영자 매뉴얼

5.2.1 H/W 구성

장치	모델명	비고
라즈베리파이 3 B+	Raspberry Pi3 B+	CPU - Broadcom BCM2837B0, Cortex-A53 64-bit SoC @ 1.4GHz Memory - 1GB LPDDR2 SDRAM
딥러닝용 USB 드라이브	Neural Compute Stick 2	X
카메라 모듈	Logitech c920	X
배터리 팩	OWL OSB-F18000	OWL Qualcomm 3.0 18000mAh (출력 : 5V 3.0A)
SD 카드	SanDisk Ultra MicroSD Class10	32GB
스피커 모듈	해당 없음	3.5mm 이어폰

5.2.2 S/W 설정

- 다크넷 데이터 학습 시 필요한 개발 환경
 - Ubuntu 16.04 LTS
- Model Optimizer 작업 시 필요한 개발 환경
 - Ubuntu 16.04 LTS
 - OpenVINO tool 2018 R5 Version
- 라즈베리파이에서 실행 시 필요한 개발 환경
 - Raspbian OS
 - Google Cloud Speech API 를 사용할 KEY 발급

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

- OpenVINO tool 2019 R1 Version
- 파이썬 모듈
 - google text-to-speech
 - google vision
 - pygame

5.2.3 운영자 가이드

- 데이터 학습 매뉴얼:

https://github.com/kookmin-sw/2019-cap1-2019_6/blob/master/DataTrain_guide.md

- openVINO 매뉴얼:

https://github.com/kookmin-sw/2019-cap1-2019_6/blob/master/openVINO_guide.md

- haarcascade 매뉴얼:

https://github.com/kookmin-sw/2019-cap1-2019_6/blob/master/HaarCascade_guide.md

5.3 테스트 케이스

대분류	소분류	기능	테스트 방법	기대 결과	테스트 결과
실행	자동 실행	라즈베리파이에 전원 공급시 프로그램이 자동으로 실행된다.	라즈베리파이에 배터리 팩을 연결한다.	"Be Eyes 실행합니다." 라는 음성 메시지가 나온다.	실패
물체 인식	장애물 인식	학습한 모델을 인식하면 해당	스마트 안경을 쓰고 물체가 있는 쪽으로 간다.	음성으로 해당 장애물을 알려준다.	성공

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	결과보고서		
	프로젝트 명	Be Eyes	
	팀 명	ViewTopia	
	Confidential Restricted	Version 1.6	2019-MAY-28

		장애물의 정보를 음성으로 전달한다.			
글자 읽기	검지손가락이 가리키는 곳의 문자 인식	검지손가락으로 가리키는 글자를 읽어준다.	스마트 안경을 쓰고 검지 손가락으로 약 상자의 아래 쪽을 가리킨다.	약 상자의 이름을 음성으로 출력한다.	성공