

드론 아트쇼 INDOOR 시스템 개발 2

팀원: 강태원, 김규열, 이영균 / 소속: 컴퓨터공학전공 / 지도교수: 안종석 교수(컴퓨터공학전공) / 산업체멘토: (주)군집텍, 서정호 대표이사

과제 개요

- 드론 아트쇼의 핵심기술 3가지 중 하나인 위치 인식 기술을 실외의 RTK-GPS나 기존 실내에서 사용하던 패시브 마커 방식의 모션캡처 기술이 아닌 다른 기술로 대체하여 고정밀 저비용의 실내 드론 아트쇼를 위한 시스템을 구축하는 프로젝트

과제 수행 목표 및 필요성

01 개발 동기

- 최근 무인비행장치를 활용한 기술의 지속적인 발전에 따라 촬영, 농업, 측량, 재난 또는 재해 감시, 건축물 역설계 등 다양한 분야에서 드론의 활용도가 높아지고 있는 추세이다. 그러나 단일 기체 비행의 경우 비행 시간과 탑재 중량의 제약으로 인해 다양한 요구에 대응하기 어렵다. 이에 따라 다수 드론을 동시에 제어하는 군집 비행 기술의 중요성이 대두되고 있다.

02 개발 목적 및 목표

- 실내 드론 군집 비행을 위한 고정밀 저비용의 위치 추적 기술을 적용하여 군집 비행 기술을 쉽게 보일 수 있는 드론 아트쇼를 진행하고자 한다.

03 개발 필요성

- 군집 비행을 위해서는 드론의 정밀한 위치 추정과 그에 따른 제어가 필수적이다. 기존에 실외 군집 비행에서 주로 이용되던 RTK-GPS의 경우, 실내에서 GPS 전파를 수신하는 것이 불가능하므로 실내 군집 비행에 적용하기 어렵다. 실내 측위 기술 중 무선 비콘 방식은 오차 범위가 커 부적합하다. 적외선 카메라와 패시브 마커를 이용한 모션 캡처 방식은 정밀도가 우수하나 시스템 구축 비용이 높다.

진행 과정

1. 구현과정

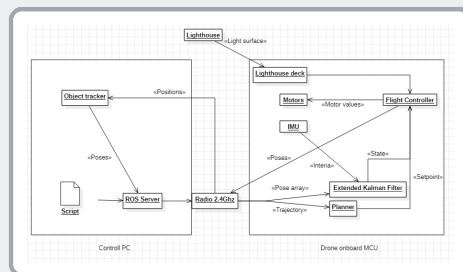
- 하드웨어 플랫폼으로 Crazyflie 2.1을 사용하였다. 해당 플랫폼은 STM32F405 기반의 FC와 가로세로 9cm, 무게 28g의 초소형 드론 플랫폼으로, 하드웨어와 소프트웨어 모두가 오픈 소스로 공개되어 있고 다양한 펌웨어를 이용할 수 있어 연구에 적합하다.
- 이에 SteamVR 베이스 스테이션을 이용한 Lighthouse positioning method를 적용해 기존 군집 비행 프로젝트인 CrazySwarm에서 사용하던 패시브 마커 기반의 모션 캡처 시스템을 대체하는 것을 목표로 하였다. 따라서 기존 CrazySwarm에서 모션 캡처와 관련된 코드 전체를 비활성화하였고, Lighthouse 시스템에 의해 드론에서 계산된 위치 정보만을 받아 이를 통해 드론을 제어하도록 하였다.

2. 기획과정

- 실내 위치 추적 기술의 활용 분야 중 VR/AR에서 사용되는 방식에 주목하였다. 현재 VR/AR 분야에서 사용되는 방식 인사이드-아웃 방식과 아웃사이드-인 방식으로 구분되는데, 인사이드-아웃 방식은 일종의 마커로서 기능하는 외부 하드웨어를 이용하는 Lighthouse 방식과 드론에 내장된 카메라에 의한 영상처리를 이용하는 마커리스 트래킹 방식으로 다시 나뉘고, 그 외 외부 카메라에서 VR 기기의 위치를 추적하는 아웃사이드-인 방식은 현재 소비자용 기기에서는 사장된 상태이다.
- 카메라를 이용하는 마커리스 트래킹 방식은 영상처리를 위한 하드웨어가 필요하여 소형의 드론에서는 적용하기 어려우므로, Lighthouse 방식을 적용하여 드론 위치를 추적하고 이를 바탕으로 군집 비행 시스템을 구현하고자 하였다.

과제 결과

- 과제 결과물은 Crazyflie 2.1 기체와 SteamVR Base Station을 이용한 드론 군집 비행 시스템이다. 이는 해당 하드웨어와 드론의 펌웨어, 제어를 위한 PC 소프트웨어를 포함한다.
- 해당 구성 요소를 전부 포함한 프로토타입 시스템을 구현하여, 드론 4대의 40cm 간격 호버링과 동시 이/착륙을 포함한 기술 데모를 진행하였다. 또한 Lighthouse의 신호로부터 위치를 추정하는 방식 두 가지(crossing beam method, sweep angle method)의 성능 비교를 진행하였다.



[그림 1] 전체 시스템의 구성도



[그림 2] 군집 호버링 비행 시연

활용방안 및 기대효과

- 기존의 실내 군집 비행에 사용되던 고가의 모션 캡처 방식을 사용하지 않아도 되므로 비용을 절감할 수 있음
- 무선 비콘이 아닌 광신호를 이용하므로 높은 수준의 위치정확도를 나타낼 수 있고, 이를 통해 드론 간의 최소이격거리를 줄여 소형 기체의 한계점을 해결할 수 있을 것으로 기대됨.
- 드론에 외부에서의 인식을 위한 별도의 마커를 장착할 필요가 없으므로 추가적인 페이로드 상승이 적고, 이를 통해 드론에 다양한 용도를 기대할 수 있음

산학협력

팀원 역할

주 1회 약 3~4시간씩 오프라인 미팅을 가져 시스템의 기능 구현 및 테스트를 진행하였다.

강태원: 미팅 진행 및 결과 정리, 시스템 구현, 비행 스크립트 작성

김규열: 하드웨어 점검 및 테스트, 시스템 구현

이영균: 시스템 구현, 비행 스크립트 작성

멘토 역할

- 무인동력비행장치 자격 취득 기회 제공
- 과제 결과물이 기술 목표에 부합하는지 확인
- 과제 결과물에 대한 점검 및 보완점 제안
- 테스트에 필요한 하드웨어 및 공간 제공

딥러닝을 이용한 이미지 투 코드

팀원: 김영훈, 이준영 / 소속: 컴퓨터공학전공 / 지도교수: 안종석 교수(컴퓨터공학전공) / 산업체멘토: (주)인셈니아, 이성훈 대표

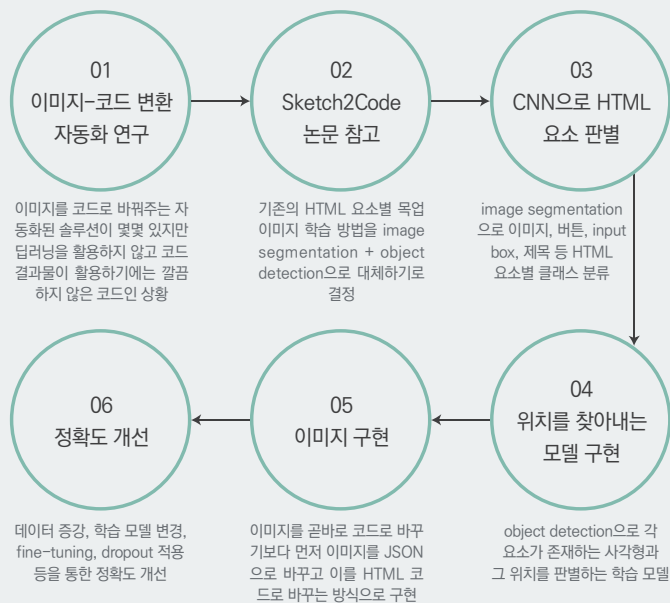
과제 개요

- UX 디자인을 HTML 코드로 변환해주는 딥러닝 기반 솔루션
- 개발 비용의 많은 부분을 차지하는 UI 개발을 자동화하여 생산성 증대 목표

과제 수행 목표 및 필요성

- 디자인 시안을 바탕으로 UI 개발(퍼블리싱)하는 과정은 오래 걸리는 과정
- 그에 비해 복잡한 로직이 필요하기 보다는 단순한 HTML 코딩 과정
- 이를 딥러닝을 통해 자동화하면 개발 비용과 시간을 크게 아낄 수 있음
- 최종적으로는 디자인 시안 이미지를 업로드하면 Tailwind 기반의 HTML로 리턴해주는 API 구축이 목표

진행 과정



활용방안 및 기대효과

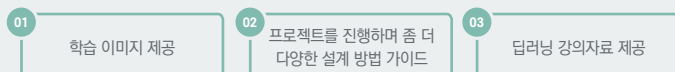
- 개발 프로젝트에서 많은 시간과 비용을 단축시킬 수 있고 개발자는 단순한 코딩이 아니라 비즈니스 로직과 보안 등 더 중요한 업무에 시간을 쓸 수 있음

산학협력

팀원 역할

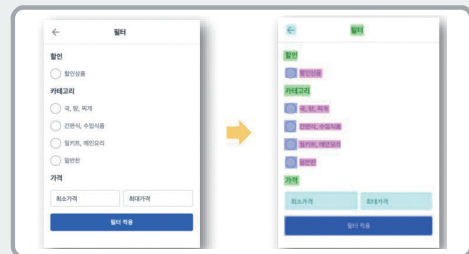
김영훈: 팀장, 보고서 작성, 논문 분석
이준영: 팀원, 아키텍처 설계 및 구현

멘토 역할

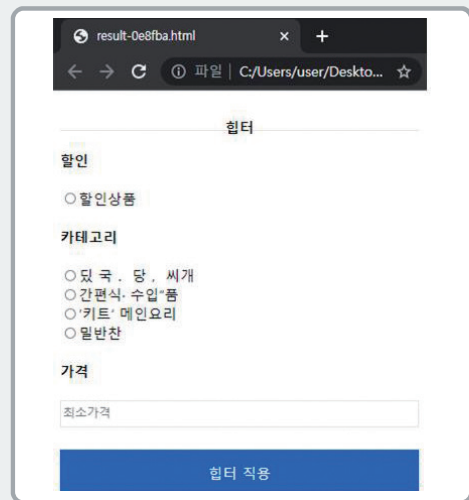


과제 결과

- 먼저 멘토에게서 받은 이미지를 그대로 사용하는 대신 어떤 요소가 제목, 인풋, 버튼 등이 있는지를 나타내는 클래스 맵으로 전처리 (결과물 사진 1이 이미지의 전처리 전후 과정)
- 입력한 이미지를 바탕으로 CNN 모델을 통해 HTML 요소를 판별하고 위치를 찾아 내도록 구현함(결과물 사진 2가 원본 이미지에서 생성한 HTML 결과물)
- 학습 데이터의 종류가 입력 품에 집중되어 있어 상대적으로 입력 요소는 잘 탐지했으나 이미지, 본문 등 입력 품에 없는 요소는 정확도가 상대적으로 낮음



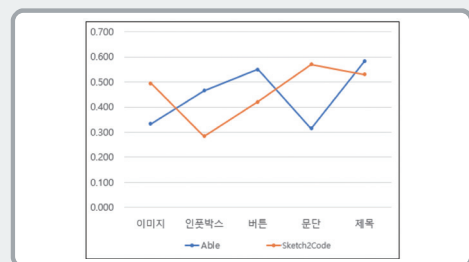
[그림 1] 원본 웹사이트(좌)와 전처리 후 이미지(우)



[그림 2] 생성된 웹사이트의 실행 이미지

	Precision	Recall	F1 Score		Precision	Recall	F1 Score
이미지	1.000	0.200	0.333	이미지	0.843	0.351	0.496
인풋박스	0.400	0.556	0.465	인풋박스	0.190	0.556	0.283
버튼	0.800	0.419	0.550	버튼	0.424	0.419	0.421
문단	0.747	0.200	0.316	문단	0.747	0.461	0.570
제목	0.700	0.500	0.583	제목	0.651	0.447	0.530

[그림 3] 레퍼런스의 정확도 자료와의 비교 표



[그림 4] 그림 3을 도식화한 그래프

딥러닝을 이용한 이미지 투 코드 솔루션

팀원: 하이현, 김병현, 이서영 / 소속: 컴퓨터공학전공 / 지도교수: 안종석 교수(컴퓨터공학전공) / 산업제멘토: (주)인셈니아, 이성훈 대표

과제 개요

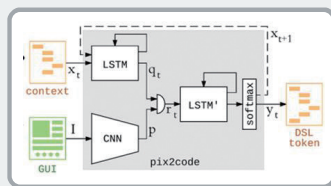
- UI/UX 디자인 이미지를 Tailwind CSS 기반 HTML 문서로 변환해주는 딥러닝 기반 솔루션 설계
- HTML 문서 작성을 자동화하여 UI 개발과정의 효율성 증가

과제 수행 목표 및 필요성

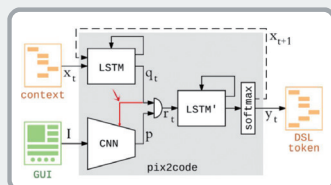
- IT 기술의 발달에 따라 웹 서비스 개발의 수요는 증가하고 있다. 웹 서비스 개발 과정에서 개발자는 디자이너가 생성한 UI/UX 디자인을 HTML로 변환하고, 서비스의 핵심 기능과 로직을 개발하는 업무를 담당한다. 이중 UI/UX 디자인을 코드로 변환하는 업무는 난이도가 어렵지 않지만 양이 많아 작업시간을 많이 요구한다. 또한 웹 서비스 개발은 프로토타입이라는 개발 기법에 따라 진행되는 경우가 많다. 이 경우 개발자는 다양한 프로토타입 디자인에 따른 HTML 문서를 모두 작성하여야 하고, 디자이너는 수정 사항이 반영된 HTML 문서가 완성될 때까지 대기하는 병목 현상이 발생하므로 HTML 코드 변환 과정은 비효율적이라고 판단된다.
- 따라서 HTML 코드 변환 과정을 자동화하여 위와 같은 프로세스의 단점을 보완하는 것이 이번 프로젝트의 목적이다. 프로젝트에서 개발하고자 하는 솔루션은 이미지 형식으로 된 UI/UX 디자인을 입력하면, 해당 디자인에 1:1로 대응하는 HTML 문서를 출력하는 것이다. 해당 솔루션의 목표는 생성된 HTML 디자인은 입력으로 사용한 UI/UX이미지와 90%이상 유사한 렌더링 결과를 출력하는 것이다. 또한, 개발자가 HTML 문서를 기반으로 웹 페이지의 핵심 기능을 구현할 때 수정을 최소화 할 수 있도록 최적화된 HTML 코드가 작성되어야 하며, 이미지 입력에서 HTML 문서의 출력까지 초 단위의 처리 속도를 보장하여야 한다. 이러한 솔루션은 인공지능 딥러닝 모델을 기반으로 설계 되었으므로 비교적 높은 수준의 연산량을 요구하여 서버에서 구동하는 웹 서비스 형태의 형태이며, 사용자가 원활하게 서비스를 이용 가능하도록 API를 제공하는 것이 목표이다.
- 해당 솔루션이 도입될 경우 회사의 관점에서는 기존에 업무 시간의 큰 축을 담당하던 HTML 문서 작성 과정이 초 단위로 단축되어 시간을 더욱 효율적으로 관리할 수 있게 되며, 이는 곧 같은 기간에 많은 웹 서비스를 개발 할 수 있음을 의미한다. 개인의 관점에서 디자이너는 개발자에게 HTML 문서 작성을 의뢰할 필요 없이 직접 솔루션을 이용하여 여러 프로토타입을 제작할 수 있으며, 개발자는 솔루션 사용을 통해 단축된 시간을 핵심 업무에 활용함으로써 업무 효율을 크게 증가시킬 수 있다. 또한 해당 솔루션의 결과는 항상 같은 코딩 스타일로 최적화된 HTML 문서이므로, 실무 인원이 변경되어도 코드의 품질은 변하지 않는다는 장점이 있다. 만약 개발에 사용하는 언어나 기법이 추가되더라도 학습할 데이터만 충분하다면, 컴파일러를 수정하고, 해당 딥러닝 모델을 학습시키는 비교적 간단한 방법으로 대응 가능하다. 결론적으로 해당 솔루션은 웹 서비스 개발에서 많은 시간을 요구하던 과정을 자동화하여 작업 효율을 개선가능하며, 회사와 개인의 관점 모두에서 큰 이점을 가진다.

진행 과정

- 이미지 투 코드를 실현하기 위해서는 이미지를 즉시 소스코드로 변환하기보다는, 쉽게 구성되고 분석될 수 있는 중간 언어 형태로 표현되는 것이 성능 면에서 우수하다. 따라서 기획과정에서는 이미지 인식, 중간 언어(DSL) 생성, 중간 언어를 Compiling 하여 원하는 언어로 작성된 소스코드(본 과제에서는 HTML 코드)로 변환함을 목적으로 하였다.
- 본 과제를 달성하기 위하여 다양한 모델의 적용 가능성을 검토했으며, 기존 연구 중 가장 본 과제에서 목표하는 바와 유사한 성과를 거둔 참고 논문 pix2code(참고문헌 1)의 모델 아키텍처를 기준으로 더 다양한 데이터를 더 좋은 성능으로 학습할 수 있는 개선안을 내놓는 것을 목적으로 하였다. 개선을 위해 기존 모델(pix2code)을 v0(버전 0)으로 명명하였으며, 기존 모델의 아키텍처는 다음과 같다.
- pix2code의 학습 과정에서는 코드로 변환을 원하는 GUI 이미지와 DSL context가 입력으로 사용된다. DSL context는 GUI 이미지와 1:1로 대응하는 정답 DSL을 토큰 별로 분할하고 이를 슬라이딩 윈도우를 사용하여 전체 DSL이 아닌 DSL의 일부분만 입력으로 사용하는 것이다. 입력받은 이미지와 DSL context에 대한 내용을 이해하고 입력으로 주어진 DSL context의 바로 뒤에 등장 할 토큰을 예측한다. 이렇게 예측된 토큰과 정답 토큰을 비교하여 모델의 loss 값을 계산하여 내부 신경망 네트워크를 업데이트한다. 그림 1의 내부에는 총 3개의 신경망 모듈이 사용된다. 이를 크게 분리해보면 좌측의 CNN, LSTM과 우측의 LSTM*으로 분리된다. 좌측의 두 계층은 입력으로부터 정보를 추출하는 인코더 계층이고, 우측의 계층은 추출된 정보가 다음 토큰을 예측하는 디코더 계층의 역할을 수행한다. 좌측의 두 계층은 각각 이미지와 현재 분석 중인 토큰의 정보를 입력으로 받아 이로부터 분석한 내용을 벡터값으로 추출하고, 두 계층의 정보를 연결하여 전달받는 LSTM* 계층은 입력받은 context의 다음 토큰의 확률을 계산하는 역할을 한다.
- 최종적으로는 확률을 기반으로 단 하나의 토큰이 정답 토큰으로 선정되고, 이렇게 도출된 토큰들이 슬라이딩 윈도우에 축적되어 다시 입력으로 제공되는 방식으로 전체 DSL 생성이 진행된다.



[그림 1] pix2code

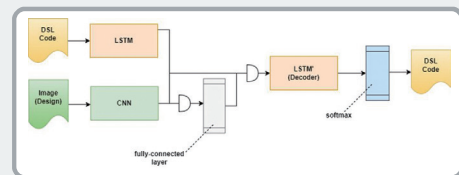


[그림 2] 구조적 개선을 적용한 pix2code v2

- 본 과제에서는 위와 같은 방식으로 동작하는 기존 모델에서 일부 파라미터 값을 변경하는 방식으로 모델을 개선하고자 하였으나, 데이터를 확장해감에 따라 이러한 방식의 개선에는 한계점이 있다는 것을 알게 되었다. 이러한 한계점의 원인을 분석한 결과 이는 Overfitting(과적합, 학습 데이터를 과하게 학습하여 학습 데이터에 대한 예측에서는 높은 정확도를 보이나 테스트 데이터에 대한 예측에서는 낮은 정확도를 보이는 현상)에 의한 것이었고, 해당 문제를 해결하고자 pix2code를 구성하는 CNN, LSTM 계층의 Overfitting 문제 해결법을 찾던 중 모델의 구조적 개선과 관련된 아이디어를 도출하게 되었다.
- 아이디어를 적용한 모델을 v2(버전 2)로 명명하였으며, v2의 아키텍처는 v0과 비교했을 때 다음과 같은 차이가 있다.
- 본 모델에서 context를 입력으로 받아 현재 분석 중인 context에 대한 벡터를 출력하는 LSTM의 출력 값을 CNN 계층의 dense layer의 입력으로 제공한다. 다음과 같이 구조를 개선하면 기존에 전체 이미지를 입력으로 받아 전체 이미지에 대한 특징 벡터값을 제공하던 CNN 계층에 현재 분석 중인 태그에 대한 정보가 제공됨으로써 CNN 계층 또한 특징 벡터를 선정할 때 context를 고려하여 값을 선정할 수 있게 되고, 이는 기존에 Decoder를 담당하던 LSTM*의 책임을 CNN에서 분담하는 것과 같은 효과를 내어 더 적은 학습 데이터로도 높은 정확도를 달성할 수 있게 하고, 유사한 데이터 간 구별 능력이 떨어지던 기존 모델의 문제점 또한 개선할 수 있게 하였다.

과제 결과

- 기존 데이터 보다 좀 더 복잡한 데이터를 만들기 위해 텍스트를 랜덤하게 생성하였고, 버튼태그의 색상을 더욱 다양하게 만들었다. 또한 위치가 다양하게 나올 수 있도록 quadruple태그를 추가하였다.
- 과제의 결과로 기존 pix2code의 모델에서 이미지 해석 부분을 CNN에서 resNet로 수정하였고, 디코더와 context 인코딩 부분을 LSTM에서 GRU로 변경하여 연산과 정확도를 향상시켰다.
- 또한, 모델의 구조를 수정하여 이미지 인코더 뒤쪽에 context 인코딩 결과를 합쳐 fully connected layer를 추가하는 방식으로 한번 더 정보를 처리하여 특징 context에서 주목할 이미지의 특징을 더 잘 찾을 수 있게 되었다.
- 앞서 수정한 모델을 훈련 뒤 원본과 비교해본 결과 테스트한 결과 약 70%의 정확도 향상을 볼 수 있었다. 이를 사용자가 편하게 사용할 수 있도록 서버로 gcp를 사용하여 모델을 serving하였고, 웹 ui를 통해 접근할 수 있도록 하였다.



[그림 3] Model Architecture



[그림 4] Client Program

활용방안 및 기대효과

- 웹 개발시 초반 작성에 활용될 수 있을 것으로 생각된다.
- 웹 UI로 배포하여 웹 개발에 종사하는 사람들의 접근성이 뛰어나 많은 사람이 사용할 수 있을 것으로 예상되며, 저희 결과물을 이용하게 된다면 웹 개발 시 이미지를 이용해 코드를 작성할 수 있게 됨으로써 생산성이 더욱 증대될 것으로 기대된다.

산학협력

팀원 역할

하이현: 팀장, 팀 회의의 스케줄 관리, Tailwind CSS 컴파일러 확장, 클라이언트 구현
김병현: 데이터 셋 확장, 데이터 자동 생성기 구현, 데이터 시각화 모듈 구현
이서영: 모델의 파라미터 개선, 모델의 구조 개선, 개선 모델의 학습 및 평가

멘토 역할

멘토는 프로젝트의 정확한 방향성과 이번 프로젝트에서 중점적으로 다뤄야 할 문제에 대해 제시하였으며, 프로젝트 중, 후반에 걸쳐 프로젝트의 진행 상황을 검토하고 나아가야 할 방향을 제시해 주었다. 프로젝트 중반에 수행된 멘토링에서는 상용화 단계에서 검토해야 할 요소들에 대해 전달받고, 현업 종사자의 시선으로 해당 프로젝트의 개선점을 전달하였다. 또한, 프로젝트 후반에 수행된 멘토링에서는 프로젝트에 대한 총평을 나누고, 추후 개선 방향에 대해 검토하는 시간을 가졌다.

경로 탐색 기반 출발 시간 알림 서비스

팀원: 이정우, 김정제, 이슬기, 한소희 / 소속: 컴퓨터공학전공 / 지도교수: 안중석 교수(컴퓨터공학전공)

과제 개요

- 도보, 버스, 지하철 등의 배차시간 및 소요시간을 예측하여 제시시간에 도착하기 위해 출발해야 하는 시간을 알림으로 제공하는 서비스

과제 수행 목표 및 필요성

01 개발 동기 및 필요성

- 최근 자가 운전자들의 증가 및 대중교통 이용에 불편을 초래하는 시위와 같은 일들로 인해 교통 체증이 증가하고 있다. 이로 인해 같은 경로인 경우에도 소요되는 시간이 일정하지 않다. 예를 들어 임의의 한 출발지로부터 목적지까지 가는 경우, 동일 시각에 출발하더라도 목적지에 도착하는 시각은 날짜 및 시간, 기상 정보, 시위와 같은 임시적 방해 요소 등에 의해 가변적이다. 따라서 목적지에 원하는 시각에 도착하기 위해 개인의 경험을 토대로 출발 시각을 예측하고, 그 예측된 출발시각에 출발해야만 했다. 하지만 이 경우 사고자 하는 경로에 대한 실시간 정보가 부족하여 도착시각을 맞추지 못하는 경우가 적지 않게 발생한다.
- 필요할 때마다 수시로 목적지까지의 경로 및 소요시간을 확인하며 '언제 출발해야만 하는가'를 예측해야 하는 것은 사용자들에게 큰 불편함으로 다가오며 상기와 같이 개인의 경험에 의해 출발 시각을 예측하는 방식은 그 예측 결과가 다소 부정확하다는 단점도 존재한다.
- 이에 본 과제는 버스, 지하철의 실시간 위치와 도착 예정 정보를 바탕으로 사용자가 제시시간에 도착하기 위한 출발 시각을 알려주는 서비스를 제안한다. 이를 통해 필요할 때마다 사용자들이 직접 경로 및 소요시간을 확인하지 않더라도 최적의 경로를 이용할 수 있게끔 하여 시간 관리를 효과적으로 가능하게 한다.

02 개발 목적 및 목표

- 본 과제의 목적은 실시간 교통 상황을 바탕으로 대중교통 경로를 탐색하여 설정한 시간 내 도착을 보장하는 출발시간 알림 서비스를 제공하는 것이다. 과제의 목적을 달성하기 위하여 다음 두 개의 하위 세부 과제를 목표로 한다.
- 첫 번째, 출발지로부터 목적지까지 대중교통 경로를 효율적으로 탐색하기 위한 알고리즘을 개발해야 한다. 이를 위해서 탐색 과정에서 방문하지 않아도 되는 노드를 프루닝하는데, 탐색 범위와 환승 횟수 제약을 기반으로 한다. 탐색 범위 제약은 탐색 범위를 출발지와 목적지를 중심으로 갖는 타원 모양으로 설정하고 단반경을 임의의 거리 계수로 넓히거나 최적의 경로를 최대 4개까지 생성하는 것이다. 환승 횟수 제약은 국내 환승 제도에 따른 환승 가능 횟수가 4회인 점을 고려하여 5회 이상부터는 방문 노드를 고려하지 않는 것이다.
- 두 번째, 실시간 교통 상황을 바탕으로 소요시간을 예측하는 모듈을 개발해야 한다. 이미 탐색이 완료된 경로라고 해도 소요시간을 나타내는 경로에 대한 가중치는 실시간 교통 상황에 따라 변할 수 있다. 이를 고려하여 임의의 갱신 주기를 설정하여 재탐색이 이루어지도록 한다.

진행 과정

1. 구현과정

개념적 설계

1. 사용자에게 출발지와 목적지에 대한 위도와 경도로 이루어진 좌표 정보를 입력 받는다.
2. 좌표 정보를 통해 새로운 노드를 생성하여 대중교통 네트워크에 삽입한다.
3. 삽입된 노드로부터 특정 변경 이내 노드를 탐색하여 연결하며 간선의 가중치는 도보 경로 탐색을 통해 가중치를 도출하여 대입한다.
4. 새롭게 구성된 네트워크에 탐색 범위 제약과 환승 횟수 제약을 반영한 '개선된 다익스트라 알고리즘'을 적용하여 출발지로부터 목적지까지 가중치가 최소가 되는 '최소 비용 경로'를 생성한다.
5. 도출된 최소 비용 경로로부터 Deviation Path 알고리즘을 사용해 '대안 경로'를 생성한다.
6. 최종적으로 대안 경로 4개와 최소 비용 경로 1개가 생성된다.

2. 기획과정

- 사용자가 제시시간에 도착하기 위해 알맞은 시간에 출발 알림을 제공해야 하는데, 이를 위해 경로를 효율적으로 탐색하기 위한 알고리즘을 개발하는 것이 중요했다. 따라서 탐색 과정에서 방문하지 않아도 되는 노드를 프루닝하는 아이디어를 구상하는 데에 큰 노력을 기울였다.

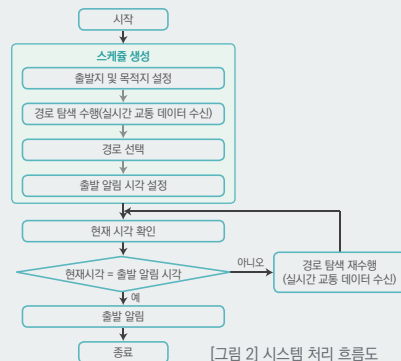


[그림 1] 타원 모형화 예시

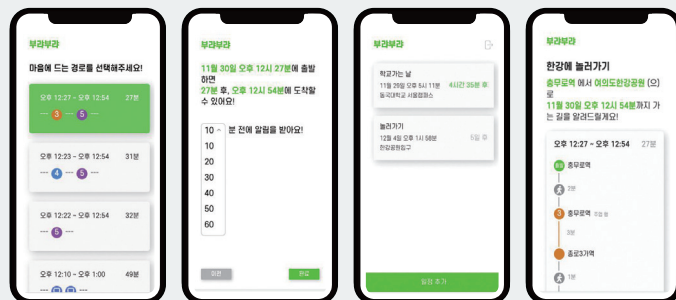
- 그 노력의 예로, 탐색 범위를 모형화하여 노드 탐색 범위를 줄이고자 하였다. 초점과 단축, 장축을 가지는 것을 특징으로 하는 타원을 탐색 모형으로 선정하였고, 이는 초점을 각각 출발지와 목적지로 설정하고, 단축의 길이를 임의의 계수로 늘려가며 점층적으로 타원의 넓이가 커지는 것을 탐색 범위를 넓히는 아이디어로 적용하기 위함이었다.
- 다른 예로, 기존 차용했던 경로 탐색 알고리즘에서 환승 가능한 범위를 측정하기 위해 일정 거리 내의 노드를 집합화했던 집합 노드를 본 과제에서는 제거하는 방식으로 알고리즘을 개선하고자 했다. 이는 각 집합 노드 간의 중복 영역이 다수 발생하여 레이어가 늘어감에 따라 불필요한 추가 연산이 발생한다고 판단했기 때문이다.
- 마지막으로 국가 통계 포털(KOSIS)에 따른 서울 내 대중교통 환승 4회 이상 비율이 단 1.5%에 불과하며 국내 환승 제도에 따른 최대 환승 가능 횟수가 4회인 점을 근거로 환승 5회 이상의 경로 탐색은 수행하지 않는 방식을 구상했다.

과제 결과

- 웹 앱의 형태로 제작되었으며 시스템 처리 흐름도는 아래 '그림 2'와 같다. 회원가입 및 로그인을 완료한 사용자가 스케줄 생성을 진행한다. 출발지 및 목적지를 설정하면 경로 탐색이 이루어진다. 최소 1개에서 최대 4개까지 경로가 생성되고, 사용자가 원하는 경로를 선택한다. 그리고 출발 전 알림을 받고자 하는 시각을 설정하면 스케줄 생성이 완료된다.



[그림 2] 시스템 처리 흐름도



[그림 3] 탐색한 결과를 바탕으로 경로 선택

[그림 4] 출발 알림 시각 설정

[그림 5] 생성한 스케줄 목록을 볼 수 있는 홈

[그림 6] 스케줄 상세 정보 확인

활용방안 및 기대효과

- 개발 동기 및 필요성에서 상술한 바와 같이 사용자들이 필요할 때마다 직접 경로 및 소요시간을 확인하지 않더라도 자동으로 경로를 탐색하고 출발시각을 갱신해주는 출발 알림 서비스로 사용자들의 시간 관리를 극대화한다.
- 또한 기존 사용자 개인의 경험을 바탕으로 출발시각을 예측하는 방식은 부정확할 뿐 아니라, 익숙하지 않은 출발지와 목적지에 대해서는 예측하기가 쉽지 않다는 문제를 해결하여 사용자에게 쉽고 정확한 예측 서비스를 제공한다.
- 추가적으로 실시간 변동 정보를 사용자에게 제공하는 기능, 다른 사용자에게 경로를 공유하는 기능,요일 반복 알림 기능 등을 개발해 함께 제공한다면 스케줄 관리가 필수인 현대인들에게 '맞춤 길찾기 앱'으로서 역할을 기대할 수 있다.

산학협력

팀원 역할

이정우: 팀장, BE, 경로 탐색 알고리즘 API 개발, 웹 서버 구현
한소희: 팀원, BE, 경로 탐색 알고리즘 API 개발, 웹 서버 구현
김정제: 팀원, FE, Application UI 구현 및 기능 개발
이슬기: 팀원, FE, Application UI 구현 및 기능 개발

딥러닝을 이용한 사용자 이슈 분류 시스템

팀원: 이호준, 박서진, 최준호 / 소속: 컴퓨터공학전공 / 지도교수: 안종석 교수(컴퓨터공학전공) / 산업예멘토: 지니뮤직 | 플랫폼개발센터 · 아키텍처그룹, 이호성 차장

과제 개요

- 최근 IT 서비스에서 UX(사용자 경험)을 높이는 것이 중요한 과제이다. 본 과제는 사용자들의 이슈를 기업에 맞는 기준으로 분류하여 피드백을 기존의 방식보다 효율적으로 처리할 수 있도록 하는 시스템을 개발하는 프로젝트이다.

과제 수행 목표 및 필요성

01 개발 동기 및 필요성

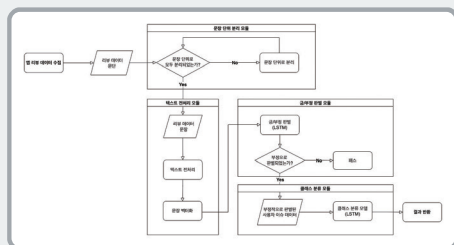
- 최근 IT 서비스에서 UX(사용자 경험)은 굉장히 중요한 가치로 여겨지고 있다. 최근 많은 기업들이 사용자 정보 데이터 및 피드백 데이터를 이용해 UX를 높이려고 하고 있다. 그러나 현재 대부분의 기업은 사용자의 요구사항이나 피드백을 자사의 홈페이지, 구글 스토어, 앱스토어, 고객센터 등을 통해 받고 이를 복잡한 처리 프로세스를 통해 처리하고 있다. 그에 따라 사용자는 자신의 요구사항이나 불편사항이 개선되었다는 것을 느끼지 못하거나 너무 늦게 처리되어 불편함을 겪고 있다.
- 특히 업데이트 이후 버그 발생과 같은 빠르게 처리되어야 하는 개발이슈들도 같은 방식으로 처리되어 사용자 경험이 떨어지는 문제가 존재한다. 즉, 기업들의 기존 사용자 이슈 처리 프로세스에서는 사용자 빠른 피드백 요구와 속도보다는 정확하게 분석해서 해결해야 하는 기업의 입장 사이에 존재하는 속도 차이 문제가 존재한다. 기존의 이슈 처리 프로세스에는 반복적이고 불필요한 과정이 존재하며 이를 분류 시스템을 통해 개선하여 효율적으로 처리할 수 있다.

02 개발 목적 및 목표

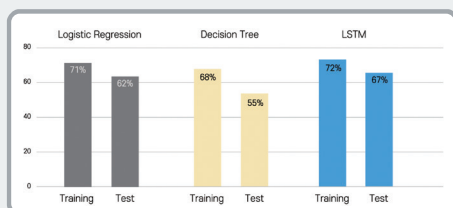
- 본 과제의 목표는 사용자의 이슈를 기업이 처리하는 기준으로 분류하여 해당 이슈를 처리하는 담당 부서로 전달하는 시스템을 개발하는 것이다.
- 이를 위해 사용자 이슈 정보를 얻는 크롤링 모듈, 문장 단위 분리 모듈, 텍스트 전처리 모듈, 긍정/부정 판별 모듈, 클래스 분류 모듈, 결과 전달 모듈 총 6개의 모듈로 개발하여 전체 시스템을 구현하고 이슈를 분류하는 딥러닝 모델을 제안한다.

진행 과정

- 본 프로젝트는 기업들의 사용자 이슈 처리 프로세스에서 사용자 빠른 피드백 요구와 이를 정확하게 분석해서 해결해야 하는 기업의 프로세스 사이의 속도 차이 이슈를 해결하기 위해 딥러닝을 이용한 사용자 이슈 분류 시스템 설계 구현하였다. 또한, 데이터 기반의 분류를 위해 LSTM 기반의 이슈 분류 모델을 제안한다. 사용자 리뷰 데이터를 얻어 이슈 분류 시스템에서 처리되어 이슈 담당 부서로 이슈 분류 결과가 전달되는 것이 전체 시스템 흐름이다. 구글 플레이 스토어, 앱 스토어에서 사용자 리뷰 데이터를 얻는다.
- 이슈 분류 시스템은 크게 6개의 모듈로 구성된다. 크롤링 모듈에서 새로운 사용자 리뷰 데이터를 크롤링하고 문단으로 들어오는 리뷰 데이터를 문장 단위로 나눠 텍스트 전처리 모듈에서 전처리를 진행하고 긍정/부정 판별 모듈에서 판별하여 부정인 문장에 대해서 클래스 분류 모듈에서 앱 개발팀, 서비스 개발팀, 서비스 기획팀, 콘텐츠 운영팀, IT 기획팀, 부정인 문장의 총 6개의 클래스로 이슈를 분류한다. 입력으로 들어온 리뷰에 대한 분류 결과를 결과 전달 모듈에서 처리하여 이슈 담당 부서로 전달한다.



[그림 1] 사용자 이슈 분류 프로세스

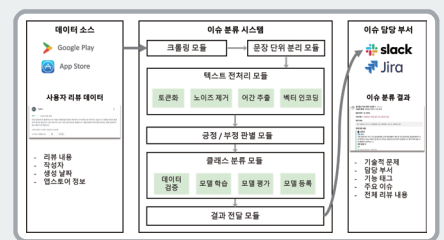


[그림 2] 머신 러닝, 딥러닝 학습 실험 결과

과제 결과

- 하단의 결과물 사진 1은 사용자 이슈 분류 시스템 아키텍처를 그린 사진이다. 이해를 돕기 위해 본 프로젝트의 전체 시스템 아키텍처를 그림으로 표현하였다. 데이터 소스에서 사용자 리뷰 데이터를 얻어 이슈 분류 시스템에서 처리되어 이슈 담당 부서로 이슈 분류 결과가 전달되는 전체 흐름을 표현하였다. 구글 플레이 스토어, 앱 스토어는 사용자 리뷰 데이터를 얻는 데이터 소스이다. 사용자 리뷰 데이터는 리뷰 내용, 작성자, 생성일자, 앱 스토어 정보 등의 정보를 가지고 있다. 이슈 분류 시스템은 크게 6개의 모듈로 구성된다. 크롤링 모듈에서 새로운 사용자 리뷰 데이터를 크롤링하고 문단으로 들어오는 리뷰 데이터를 문장 단위로 나눠 텍스트 전처리 모듈에서 전처리를 진행하고 긍정/부정 판별 모듈에서 판별하여 부정인 문장에 대해서 클래스 분류 모듈에서 앱 개발팀, 서비스 개발팀, 서비스 기획팀, 콘텐츠 운영팀, IT 기획팀, 부정인 문장의 총 6개의 클래스로 이슈를 분류한다. 입력으로 들어온 리뷰에 대한 분류 결과를 결과 전달 모듈에서 처리하여 이슈 담당 부서로 전달한다. 담당 부서로 전달되는 이슈 분류 결과의 내용으로는 기술적 문제, 담당 부서, 기능 태그, 주요 이슈, 전체 리뷰 내용, 리뷰 작성자, 리뷰 정보 등의 정보가 포함되어 있다.
- 이슈 분류 시스템은 하단의 결과물 사진 2는 사용자 이슈 분류 시스템 결과물을 캡처한 사진이다. 실제 시스템을 동작시켜 대표적인 협업 툴인 Slack으로 분류 결과 메시지를 전달한 것을 캡처한 사진이다. "이번 업데이트하고 나서 잠금화면에 노래 재생 이름이 계속 한 곡으로 떠요. 잠금화면에서 노래 재생 되고 있는 부분이 잘 표시가 안되는 것 같아요. 앱 개선이 필요합니다. 빨리해주세요. 너무 불편한 것 같아요.ㅠㅠ" 라는 리뷰에 대한 이슈 분류 시스템의 분류 결과에 대한 내용이다. 앱 개발팀이 해결해야 할 이슈로 정확히 분류하였으며 기술적 문제, 담당 부서, 기능 태그, 주요 이슈, 전체 리뷰 내용, 리뷰 작성자, 리뷰 정보 등의 정보가 포함되어 있다.

[사진 1] 사용자 이슈 분류 시스템 아키텍처



[사진 2] 사용자 이슈 분류 시스템 결과 사진



활용방안 및 기대효과

- 본 프로젝트의 시스템을 통해 앱 서비스를 제공하는 기업은 사용자 이슈를 분류하여 사용자의 피드백을 빠르게 반영할 수 있다. 또한, 앱 서비스를 제공하지 않는 기업이라도 해당 기업의 이슈 분류기준으로 데이터를 생성하고 기존의 모델을 이용하여 전이학습을 진행하여 해당 기업에 맞는 이슈 분류 모델을 생성하여 사용자 이슈를 분류하는 시스템을 이용할 수 있다. 이로 인해 기업은 사용자 피드백을 받아 서비스를 보완, 확장해나갈 수 있고 사용자 경험도 높이고 사용자 수와 기업의 이익을 높일 수 있다. 즉, 본 시스템을 통해 사용자 데이터 기반의 올바른 기업 의사 결정을 도와 이익을 창출하고, 사회적으로는 향상된 고객 경험으로 환원 효과를 기대할 수 있는 이상적인 사이클 구조를 만들 수 있다.

산학협력

팀원 역할

- 이호준: 팀장, 프로젝트 배경지식 및 솔루션을 위한 자료조사, 보고서 작성, 멘토링 주도, 딥러닝 모델 개발, 딥러닝 서버 개발
- 박서진: 팀원, 프로젝트 솔루션 도출을 위한 관련 자료 조사, 크롤링 모듈 개발, 서버 개발 등 실제 프로젝트 테스트 환경 구축
- 최준호: 팀원프로젝트 문제 정의, 솔루션 도출을 위한 논문 및 책 서베이 수행 및 공유, 회의 시 서기 역할 수행, 발표, 보고서 자료 작성 및 삽화 디자인, 전처리 구현, 데이터 분석

멘토 역할

- 지니뮤직 이호성 멘토는 멘토링 미팅을 진행하고 미팅이 아닌 평소에도 질의응답에 답변해주면서 프로젝트를 진행하는데 도움을 주었다. 주로 프로젝트와 관련된 정보와 이슈 그리고 실제 현업에서 발생하는 사례를 설명해주었다. 또한, 질문에 대해서 답변을 해주고 팀이 제안하는 솔루션을 보고 피드백을 주었다. 본 프로젝트의 기초가 되었던 지니뮤직의 애로사항을 설명해주어 본 프로젝트가 시작되었으며 실제 지니뮤직이 사용자 이슈를 처리하는 방식, 각 이슈를 실제로 담당하고 있는 부서의 정보, 시스템의 요구사항 등의 정보를 주며 프로젝트에 대해 조언해주었다. 이외에도 프로젝트와 관련되지 않은 개발자로의 진로, 공부방법, 시스템 아키텍처와 관련된 지식 등에 대한 질문에 대한 멘토링을 했다.

낙상 감지 시스템

팀원: 김범식, 김경래, 류신영, 김정엽 / 소속: 컴퓨터공학전공 / 지도교수: 안종석 교수(컴퓨터공학전공)

과제 개요

- 현대사회에서 문제가 되고 있는 독거노인의 낙상 문제를 해결하기 위해 동영상 기반의 이미지 인식 알고리즘을 통해 낙상을 인식시켜 낙상을 사전에 예방한다.

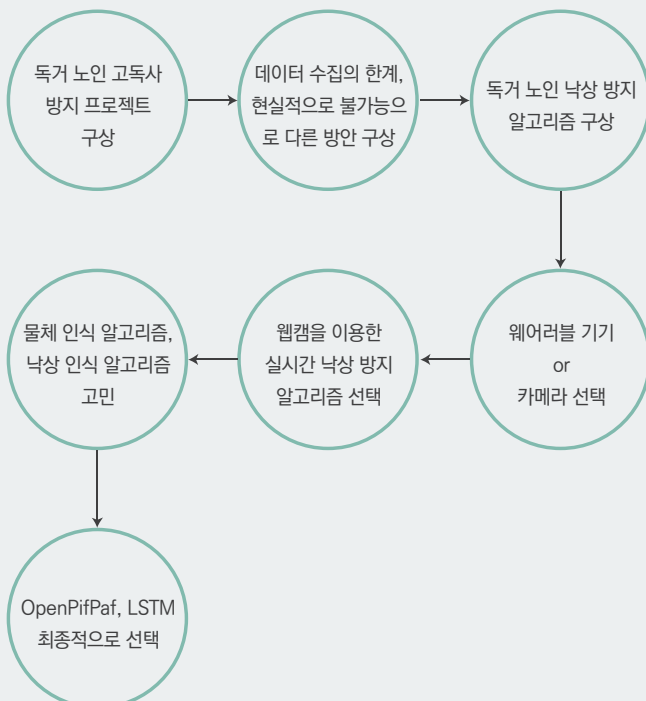
과제 수행 목표 및 필요성

- 현대사회에서 문제가 되고 있는 독거노인의 낙상 문제를 해결하기 위함
- 매년 전 세계적으로 3700만 건의 낙상 사고가 일어나는데, 독거노인들은 속수무책으로 도움을 요청하지 못하는 경우가 생겨 위험하다.
- 카메라 및 웹캠을 이용하여 동영상 정보를 입력받아 LSTM 알고리즘을 이용하여 실시간으로 낙상을 감지하는 시스템을 개발한다.

진행 과정

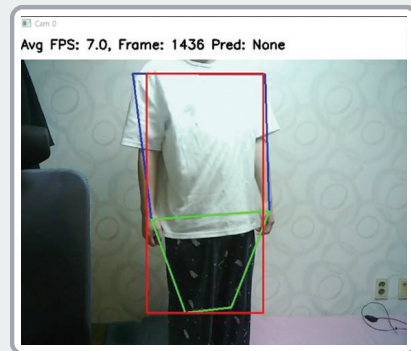
- 구현과정**
 - 카메라 또는 웹캠을 이용하여 동영상 정보를 입력받는다.
 - OpenPifPaf 알고리즘을 이용하여 포즈를 추정한다.
 - 추정한 포즈로부터 5가지 feature를 추출한다.
 - LSTM 알고리즘을 이용하여 각 프레임 간 sequence를 분석하여 낙상을 감지한다.

- 기획과정**
 - 사회공헌적인 프로젝트 진행 계획



과제 결과

- 기존에 가정에서 자주 쓰이는 웹캠을 이용하여 일반적 상황에서 낙상을 감지할 수 있다. 특정 각도에 따라 낙상을 인식하지 못하거나, 낙상과 유사동작을 구분하지 못하는 경우가 있다.
- 각도에 따른 문제는 여러 카메라를 이용하여 사각지대를 없애는 것으로 해결이 가능하다. 유사동작을 구분하지 못하는 것은 더 많은 데이터 수집으로 해결 가능할 것으로 보인다.



[그림 1] 평범히 서있을 때의 사진



[그림 2] 낙상을 감지한 사진

활용방안 및 기대효과

- 정부가 지자체에서 독거노인들에게 저렴한 웹캠을 보급하여 예상치 못한 낙상 사고를 당해 도움을 받지 못하는 것을 예방할 수 있다.
- 나아가 낙상 이외에도 다른 위험요소를 비디오를 통해 판단하여 사전에 문제를 예방할 수 있다.

산학협력

팀원 역할

김경래 & 류신영

: OpenPifPaf 알고리즘을 이용하여 각 프레임에서 인간 자세의 핵심 포인트를 추출하는 역할이다. 각 키 포인트 세트를 이용하여 스켈레톤 모형을 생성한다. 또한 해당 알고리즘을 개선하여 인체, 관절 인식의 정확도를 개선하는 역할을 맡았다.

김범식 & 김정엽

: LSTM 알고리즘을 이용하여 인식된 인간의 행동을 인식하여 예측한다. "Fall" 클래스와 "No Fall" 클래스의 2가지로 구분되며 Pytorch의 Adam optimizer를 사용한다. 또한 해당 알고리즘을 개선하여 낙상 인식의 정확도를 개선하는 역할을 맡았다.

강의 요약 서비스

팀원: 한영빈, 전취준, 이현호, 황세연 / 소속: 컴퓨터공학전공 / 지도교수: 안종석 교수(컴퓨터공학전공) / 산업체멘토: LG CNS, 박정규

과제 개요

- 강의 스크립트를 입력받아 요약 모델을 통해 요약본을 제공한다.

과제 수행 목표 및 필요성

- 코로나19의 영향으로 비대면환경에서 진행되는 강의들이 증가하였다. 이에 따른 이러닝 공급시장 규모가 매년 늘었다. 청각장애학생들은 온라인 강의 수강으로 인해 학습권이 보장되지 않은 상황이 많았다. 교수자의 입모양을 볼 수 없고 기본적 자막 또한 보장되지 않는 강의가 대부분이었기 때문이다.
- 온라인 강의 수강으로 인해 청각 장애 학생들의 학습권이 침해받지 않도록, 그리고 모든 온라인 강의 수강하는 학생들의 효율적인 학습에 대해 큰 도움을 줄 수 있는 서비스이다.
- 한국어 위주의 강의 영상에서 음성을 추출한다. 이를 STT를 통해 텍스트 파일로 변환한다. 이 텍스트 파일을 500자 이하로 분할하여 강의 내용을 요약하여 출력한다.

진행 과정

1. 구현과정



- 요약이 진행되는 파이프라인이다. STT에서 출력된 텍스트를 500자 이내로 끊어서 KoBART Summarization을 진행한다. 이렇게 요약된 텍스트를 출력한다.
- 이때 STT에서 텍스트화된 텍스트 중 사용자가 수정할 수 있으며, clova 기능 중 요약 구간이 지정된다. 이때 500자 이내로 끊을 수 있으며 그 구간별로 요약이 된다.

2. 기획과정

- 한국어 음성추출엔 네이버 클로버 스피치가 제일 높은 성능을 보인다는 것을 알게 되었다. 그래서 강의 요약 서비스의 성능을 위해 클로버 API를 사용하기로 했다. 그리고 요약 부분에는 추상적 요약과 추출적 요약이 있다.
- 기존에 있는 한국어 요약 모델이 추상적 요약을 하는 모델이기도 하고, 전체적인 향상된 요약 모델을 위해 추상적 요약을 하는 모델을 택하였다.

산학협력

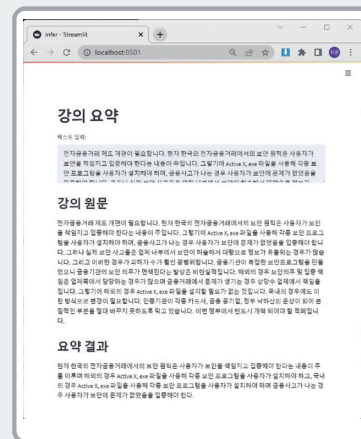
팀원 역할

팀원들 모두 요약 평가 알고리즘과 이미 존재하는 요약 모델인 KoBart summarization의 성능 확인 후 데이터의 종류를 확인하는 등 선행 기술에 대한 조사를 하였다. 격주에 걸쳐 발표자료를 만들었으며 격주 팀원들 끼리 돌아가며 수업 중 교수님과 학생들 앞에서 발표하였다. 발표 중 교수님의 피드백을 받으며 차주에 할 일을 정하고 그에 대한 방안이나 자료를 연구하여 솔루션을 마련하는 등 실제 구현을 위한 계획과 시스템을 이해하고 정하였다.

한영빈, 황세연: koBart 모델의 fine-tuning을 진행
이현호, 전취준: 프론트엔드와 백엔드를 맡아 개발을 진행

과제 결과

- koBart Summarization을 이용해 기존 모델의 성능을 확인해보았다. 그리고 Rouge F1 스코어를 확인해 kobart모델의 정확도를 확인했고, 형태소 분석을 한 본문과 그냥 본문과의 rouge f1 스코어를 비교해 형태소 분석이 요약 모델에 어떤 영향을 끼치는지, 형태소 분석이 필요한 작업이라는 것을 확인하였다.
- 그래서 이번 과목에서는 STT는 Clova API를 통해 정확도가 높은 텍스트화를 진행할 것이며 이를 통해 추출된 텍스트로 딥러닝을 이용한 추상적 요약을 해 요약문을 제공할 것이다. 사용되는 요약 모델은 한국어 데이터베이스로 학습된 KoBart Summarization이며 웹으로 사용자에게 이 서비스를 제공할 예정이다. 그리고 이 요약 모델이 만들어낸 요약문에 대한 평가는 사람의 요약문과 기계 요약문을 비교해서 평가하는 Rouge 알고리즘의 F1 스코어를 이용하여 성능을 측정한다.



[그림 1] 유튜브의 강의영상을 요약한 결과

	ROUGE-1	ROUGE-2	ROUGE-1
F1	0.459	0.150	0.270

[그림 2] 도서 자료 데이터셋으로 학습한 모델의 성능 측정 결과
DACon에서 주최한 한국어 문서 생성요약 시 경진대회에 사용된
모델 성능 측정 코드를 사용하여 측정하였음

활용방안 및 기대효과

- 온라인 강의로 인해 강의자의 입 모양을 볼 수 없는 상황에서 변환된 텍스트를 통해 내용 파악 가능해 청각장애인 학생의 학습권이 보장된다.
- 요약본이 제공됨으로써 학생들이 강의에서 원하는 부분 찾기에 유용해 학습시간이 단축되는 효과가 있다.
- 스크립트와 요약문이 제공되어 복습 시 유용하다.

멘토 역할

- 초반에 진행 방향에 대해 대략적으로 흐름을 알려주신 후 팀플레이를 위한 회의자료 관리, 코드 공유, 개발 방향 진행, restful api 구현에 대해 알려주셨다. 그 후 교수님의 피드백 받은 것 중 모르겠는 내용에 대해 알려주셨다. 가령 모르는 내용을 요약할 수 있는가라는 교수님의 물음에 멘토님께서 오히려 기계나 요약할 것 할 수 있는것이고 시는 요약할 경우에 문장간의 관계를 고려하여 요약하므로 요약이 가능하다고 알려주셨다.
- 그리고 fine tuning을 통해 요약을 정확성을 높이려는 과제는 꽤나 의미있고 도움이 되는 공부라고 알려주셨다. 그 후 GPU서버 대여는 시허브라는 곳에서 진행하고 있다는 정보를 주셨고, 요약 정확도를 어떻게 확인해볼 것인가에 대해 직접 몇 가지를 요약해보고 스코어를 비교하는 방법을 제안 주셨고, STT는 클로버가 제일 정확하니 이 클로버 API로 사용하는 것이 정확도 면에서 좋을 것이라고 추천해주셨다.

YOLO 모델을 이용한 물체 인식 보조 어플리케이션

팀원: 아미나, 유천일, 정여준, 최홍규 / 소속: 컴퓨터공학전공 / 지도교수: 안종석 교수(컴퓨터공학전공)

과제 개요

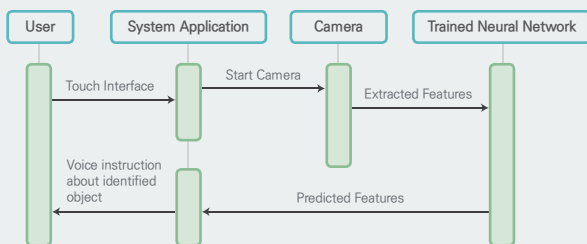
- 2001년과 비해 현재 등록 시각장애인 수가 약 2배로 늘어난 추세에서 시각장애인은 보행 시 점자블록, 볼라드, 및 횡단보도의 인식이 불가능하여 어려움을 겪고 있다. 이러한 상황에서 시각장애인을 위해 보행 보조 어플리케이션을 만들고자 한다.

과제 수행 목표 및 필요성

- 시각장애인의 현재의 환경에서 횡단보도, 점자블록, 볼라드의 위치, 거리 등의 인식이 불가능한 물체의 안내를 휴대폰 카메라를 통해서 전방에 배치된 장애물의 위치와 거리를 인식할 수 있도록 한다. 이렇게 감지된 장애물을 음성으로 출력하여 시각장애인으로 하여금 보행에 도움이 되는 장치를 개발한다.

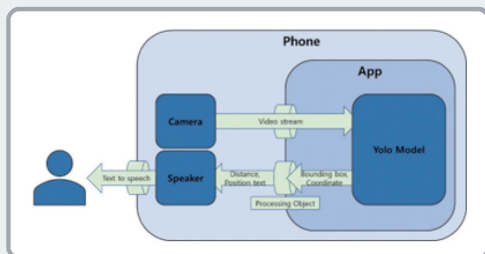
진행 과정

1. 구현과정



[그림 1] 시스템 아키텍처

- 사용자가 인터페이스를 사용하여 어플리케이션을 실행하면 카메라가 작동하고 카메라를 통해 얻은 이미지에서 특징을 추출하여 뉴럴 네트워크로 보낸다. 뉴럴 네트워크에서는 받은 정보를 분석하고 학습하고 학습을 통해 얻은 정보를 바탕으로 객체를 예측하고 예측한 정보를 어플리케이션으로 넘긴다. 이 정보를 가지고 어플리케이션은 인식된 객체를 음성으로 사용자에게 알려준다. 또한 표준거리로 설정한 거리와 비례식을 이용하여 현재 인식한 물체의 거리를 측정하여 음성으로 사용자에게 알려준다.



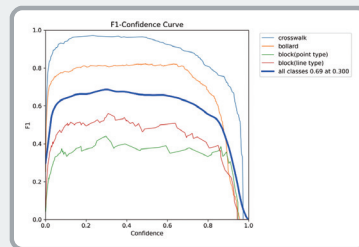
[그림 2] 어플리케이션 설계도

- 본 프로젝트는 앱 내부에 모델을 탑재하는 방식이다. 따라서 인터넷 없는 곳에서도 사용가능하다. 어플리케이션이 작동을 시작하면 카메라로부터 비디오 스트림을 앱 내부의 모델에 실시간으로 입력한다. 모델은 비디오 속 물체를 인식하여 bounding box와 coordinate를 출력으로 내보낸다. 이 값을 통해 앱 내부에서 물체까지의 거리, 위치 등의 정보를 구하여 휴대폰의 스피커를 통해 음성으로 전달한다.

2. 기획과정

- 점자블록, 볼라드, 횡단보도의 인식에 어려움을 겪는 시각장애인들을 보며 본 어플리케이션의 개발을 떠올렸다. 다른 논문에서 나온 설문조사의 결과를 참고하였을 때도, 시각장애인들의 독립적인 보행을 하지 못하고 있는 이유에 위의 이유들이 들어가고 있는 것을 알 수 있었다.
- 이를 통해 글로벌 팀은 딥러닝을 통해서 물체들을 탐지하여 음성으로 알려주는 어플리케이션이라는 아이디어를 떠올리게 되었다.

과제 결과



[그림 3] 현재 모델 정확도

- 현재 구현된 물체 인식 모델의 정확도이다. 2100개의 사진을 학습시킨 결과다. 카메라로 주변을 둘러봤을 때 실시간으로 횡단보도, 점자블록, 볼라드 등을 인식할 수 있다.

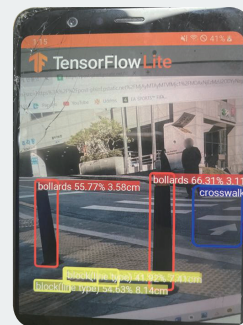


[그림 4] 어플리케이션 구동 화면

- 어플리케이션의 실행하면 화면에 터치 또는 스와이프 기능으로 쉽게 물체 인식 페이지로 이동할 수 있다. 시각장애인을 위한 어플리케이션임으로 모든 안내는 음성으로 나오고, 화면에는 실시간 카메라가 뜬다.



[그림 5] 어플리케이션 메인 화면



[그림 6] 물체 인식 모델이 작동하고 있는 모습

활용방안 및 기대효과

- 시각장애인들이 본 어플리케이션을 사용하게 되면 시각장애인들이 독립적 보행을 할 수 있고, 사고를 방지하고, 예산도 절감할 수 있다.

산학협력

팀원 역할

아미나: 어플리케이션의 기본 디자인 및 구글 Text to Speech 엔진 구현

유천일: Custom Dataset 생성 및 객체탐지 모델 학습

정여준: 객체탐지 모델 학습 및 거리 계산 알고리즘 적용

최홍규: 어플리케이션의 기본 디자인 및 구글 Text to Speech 엔진 구현