

인공지능 데이터 구축·활용 가이드라인

- 버스 승하차 데이터셋 -

인공지능 데이터 구축	사업 총괄	
	데이터 설계	
	원천데이터 수집	
	원천데이터 정제	 
	데이터 가공	 
	데이터 검수	 
	저작도구 개발	
	AI모델 개발	
	응용 서비스 개발	
가이드라인 작성		이 금 석
		정 형 구
가이드라인 버전	가이드라인 문서 rev.8 (2021/03/25)	

<제목 차례>

1. 데이터 명세 정보	1
1.1 데이터 정보 요약	1
1.2 데이터 포맷	1
1.3 어노테이션 포맷 (최종)	2
1.4 데이터 구성	4
1.4.1 일반 데이터 비디오 클립 이름 규칙	4
1.4.2 특수 데이터 비디오 클립 이름 규칙	4
1.4.3 데이터 분류 규칙	5
1.4.4 데이터의 카메라 채널 정의	5
1.5 데이터 통계	8
1.5.1 데이터 구축 규모	8
1.5.2 데이터 분포	8
1.6 원시데이터 특성	10
1.6.1 대상분류	10
1.6.2 제약조건	10
1.6.3 속성	11
1.7 기타정보	11
1.7.1 포괄성	11
1.7.2 독립성	11
1.7.3 유의사항	11
2. 데이터 구축 가이드	12
2.1 데이터 구축 개요	12
2.2 문제정의	12
2.2.1 임무 정의	12
2.2.2 데이터 구축 유의사항	12
2.3 획득·정제	13
2.3.1 원시데이터 선정	13
2.3.2 획득·정제 절차	13
2.3.3 획득·정제 기준	15

2.3.5 획득·정제 조직	15
2.3.6 획득·정제 도구	16
2.4 어노테이션/라벨링	17
2.4.1 어노테이션/라벨링 절차	17
2.4.2 어노테이션/라벨링 기준	19
2.4.3 어노테이션/라벨링 조직	22
2.4.4 어노테이션/라벨링 도구	22
2.5 검수	23
2.5.1 검수 절차	23
2.5.2 검수 기준	28
2.5.3 검수 조직	28
2.5.4 검수 도구	28
2.5.5 기타 품질관리 활동	29
2.6 활용	29
2.6.1 활용 모델	29

1. 데이터 명세 정보

1.1 데이터 정보 요약

데이터 이름	대중교통 승객 승하차 데이터셋	
활용 분야	자율주행 버스, 대중교통 안전, 승객수 측정, 대중교통 보건(마스크 유무)	
데이터 요약	대중교통 탑승객의 안전을 위한 승하차 승객에 대한 스켈레톤, 어노테이션, 마스크착용, 연령 성별 등의 정보	
데이터 출처	서울시 버스 사업자	
데이터 이력	배포버전	1.0.0
	개정이력	-
	작성자/ 배포자	정형구

1.2 데이터 포맷

프레임별 추출된 jpg 파일과 해당 이미지의 데이터 json 쌍으로 구성된다.



그림 1. 데이터 포맷 대표 도면

1.3 어노테이션 포맷 (최종)

No	항목	내용	타입
1	info	이미지 메타 데이터	List
1-1	description	데이터셋 명	String
1-2	year	데이터셋 공개연도	Integer
1-3	version	데이터셋 버전	String
1-4	img_path	이미지파일 경로	String
1-5	width	이미지 Width	Integer
1-6	height	이미지 Height	Integer
1-7	G-Sensor	가속도 센서 측정값	Integer
1-8	Gyro	자이로 센서 측정값	Integer
2	annotations	이미지 어노테이션 정보	List
2-1	Id	해당 객체 ID	Integer
2-2	bbox	승객 Bounding Box 정보 (좌상단 x, y / 우하단 x, y)	List
2-3	occluded	Bounding box 가림/잘림 정보	Boolean
2-4	num_keypoints	어노테이션된 Keypoint 수	Integer
2-5	keypoints	16개 Keypoint 정보	List
2-6	action	현재 동작 상태 정보 (stand, walk, run, sit down, lay down, grab handle, kick, punch)	String
2-7	age	연령 정보 (Children, Adult, Elder)	String
2-8	disabled	신체 부자유 여부	Boolean
2-9	mask	마스크 착용여부	Boolean
2-10	get_on	버스 승차 여부	Boolean
2-11	get_off	버스 하차 여부	Boolean

표 1 어노테이션 구성표

(구축 단계에서의 데이터는 json 포맷에는 date, gnss 정보가 포함되어 있으나 이미지 정보와의 결합 시 개인정보 식별의 가능성 때문에 최종 데이터에서는 제거한다.)

```
{  
  "info": {  
    "description": "AIHub Bus Passenger Dataset",  
    "year": 2020,  
    "version": "1.0",  
    "img_path": "",  
    "width": 1920,  
    "height": 1080,  
    "G-Sensor": "-47,-72,941",  
    "Gyro": "0,0,0"  
  },  
  "annotations": [  
    {  
      "id": 2,  
      "bbox": [  
        710,  
        598,  
        1201,  
        1077  
      ],  
      "occluded": false,  
      "num_keypoints": 16,  
      "keypoints": [  
        "..."  
      ],  
      "action": "stand",  
      "age": "adult",  
      "disabled": false,  
      "mask": false,  
      "get_on": false,  
      "get_off": false  
    }  
  ]  
}
```

그림 2. json 파일 표현 형식

1.4 데이터 구성

최종 데이터 저장소는 다음과 같은 규칙으로 구성한다. (단, 구축 단계에서의 데이터는 관리 효율을 위해 json, jpg, mp4 파일을 나누지 않고 통합하여 관리한다.)

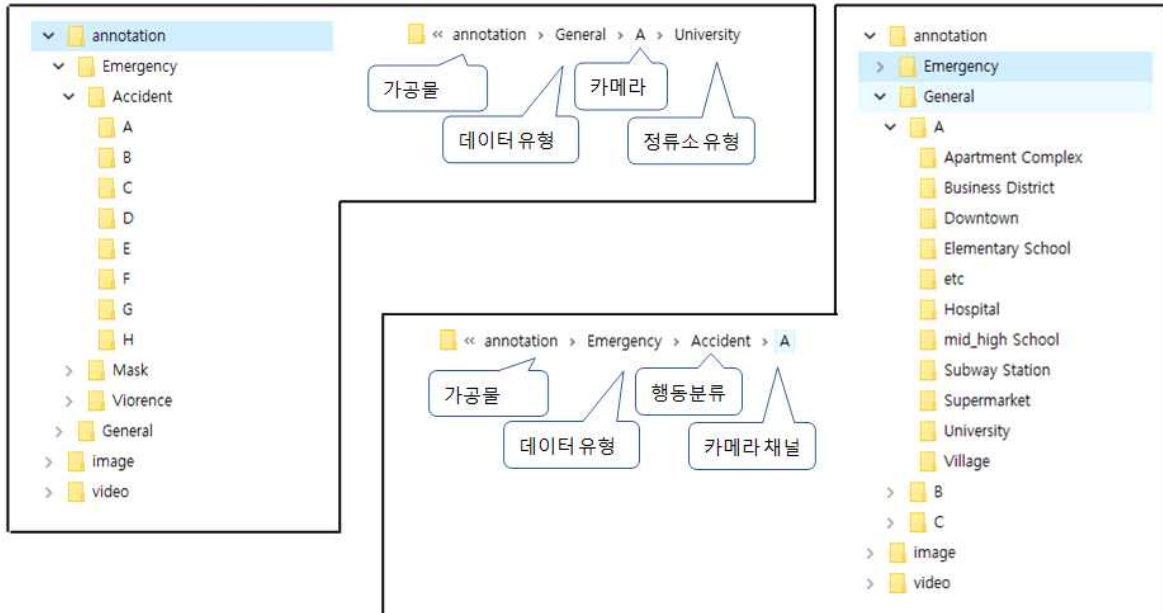


그림 2 데이터 저장 폴더 구조

1.4.1 일반 데이터 비디오 클립 이름 규칙

실제 운용 버스에서 촬영된 영상 데이터의 파일명은 아래 표와 같이 작성한다.

규칙	예시	[장소]시간대_클립번호채널.mp4
완성 예문		[초등학교]심야_001B.mp4
장소	[초등학교]	후술 장소 정의에서 공백과 앞을 떼어 기입
시간구분	심야	{아침, 출근, 낮, 퇴근, 밤}
클립번호	_001	0을 삽입해서 3자리를 맞추고 001부터 증가
채널	B	{A, B, C} (비디오 정의 참조)

1.4.2 특수 데이터 비디오 클립 이름 규칙

특수 목적을 촬영된 영상 데이터의 파일명은 아래 표와 같이 작성한다.

규칙	예시	[분류]시나리오_스크립트채널.mp4
완성 예문		[특수]보건13_023G.avi
분류	[특수]	[특수] 문구 고정 입력 (일반 영상과 구분 목적)
시나리오	01	{사고, 폭력, 보건} + 0을 삽입한 2자리 숫자 시나리오 번호
스크립트 번호	_001	0을 삽입해서 3자리를 맞추고 001부터 증가
채널	G	{A, B, C, D, E, F, G, H} (비디오 정의 참조)

json 및 jpg파일은 파일명에 _프레임번호를 추가 하여 관리한다.

완성 예문 [초등학교]심야_001B_56.jpg / [특수]보건13_023G_22.json

1.4.3 데이터 분류 규칙

정의	시작 시간	종료시간
아침	첫차	07:00
출근	07:00	10:00
낮	10:00	17:00
퇴근	17:00	21:00
밤	21:00	막차

채널 정의	카메라 위치
채널 A	앞문 외측 사이드 카메라
채널 B	운전석 측 내부 카메라
채널 C	후문 앞 내부 카메라

대분류	소분류	설명
General	Village	주거지구로 분류 할 수 있는 곳
	Apartment Complex	주거지구 중 대규모 아파트 단지 앞
	Downtown	상업지구로 분류 할 수 있는 곳
	Business District	업무지구로 분류할 수 있는 곳(상업과 중복 시 상업지구 우선)
	Hospital	대형병원
	University	대학교 앞
	Mid-high school	중고교 앞
	Elementary school	초등학교 앞
	Supermarket	대형마트 앞 (대형 전통시장도 포함)
	Subway station	지하철 환승 정류장
	Etc	분류 불가
Emergency	Violence	버스 안에서 발생하는 폭력 상황에 대한 영상
	Accident	버스 안에서 발생하는 승객의 사고 상황에 대한 영상
	Mask	보건 데이터 훈련을 위한 마스크 영상
본 영상 취득 가이드에 의거 폴더 구조를 작성하고 폴더 내에 비디오 클립 이름 규칙에 따라 비디오를 축적합니다		
본 영상 취득 가이드에 의거 소분류가 고르게 분포하도록 영상을 취득해야 합니다		

나이 구분	영문 정의	정량 기준
유소년	children	신장 약 1미터 미만인 어린이 및 아동
성인	adult	유소년과 고령자를 제외한 모두
고령자	elder	70세 이상으로 평가되거나 고령으로 거동이 느리거나 불편한 사람

행동 구분	영문 정의	설명 및 정의
서 있는 상태	stand	다리를 움직이지 않고 선 채로 고정된 상태
걷는 상태	walk	걸고 있는 상태
뛰는 상태	run	뛰고 있는 상태
앉은 상태	sit down	앉은 상태
쓰러진 상태	lay down	쓰러진 상태
손잡이를 잡은 상태	grab handle	버스 내부의 링 혹은 봉 형태의 손잡이를 잡은 상태 (식별할 수 없을 때는 stand와 비교하여 더 적합한 상태 선택)
발로 차는 상태	kick	발을 이용한 공격
주먹질 상태	punch	주먹을 이용한 공격(물건 휘두르는 경우 포함)

1.4.4 데이터의 카메라 채널 정의



표 9 일반 수집 영상의 취득 채널 정의

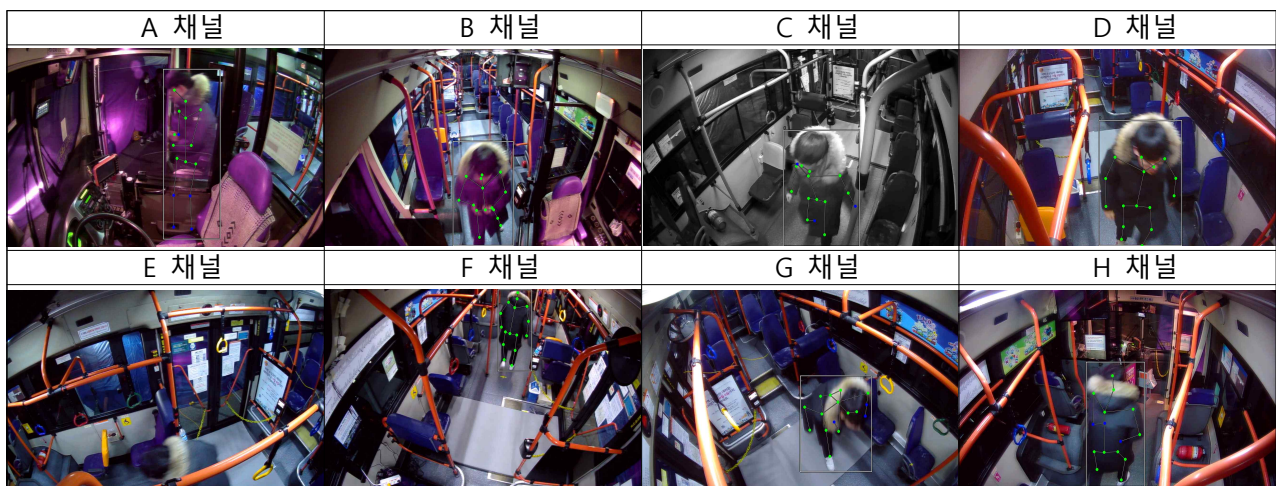


표 10 특수 카메라의 취득 채널 정의

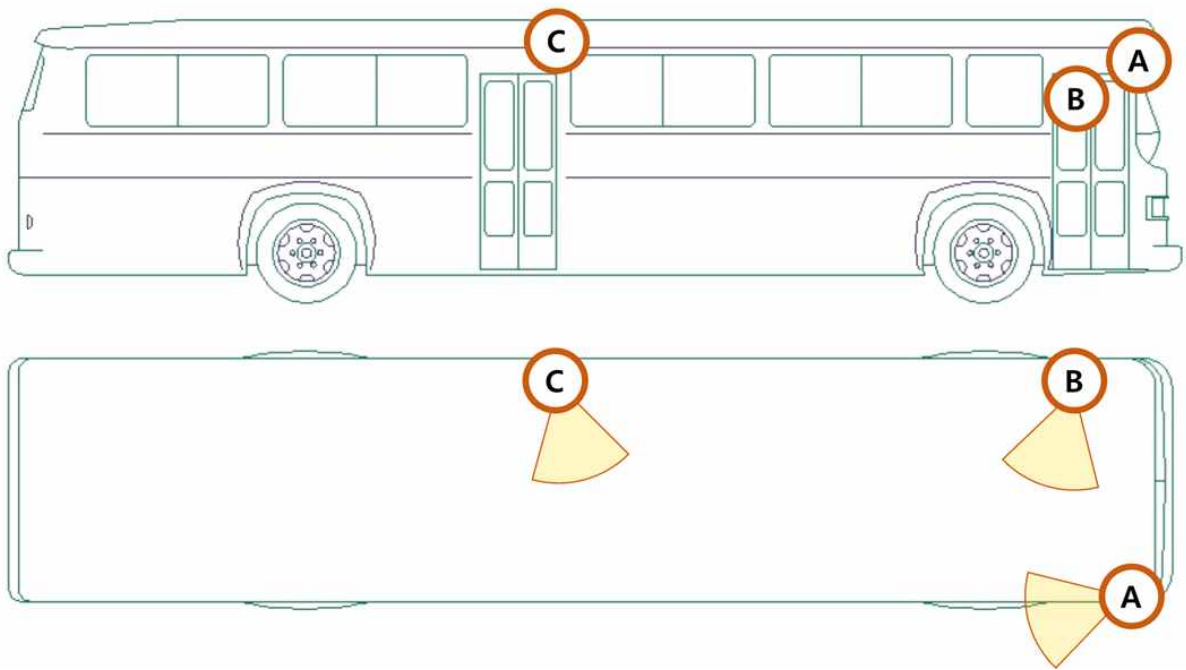


그림 3 일반 수집 영상의 카메라 배치

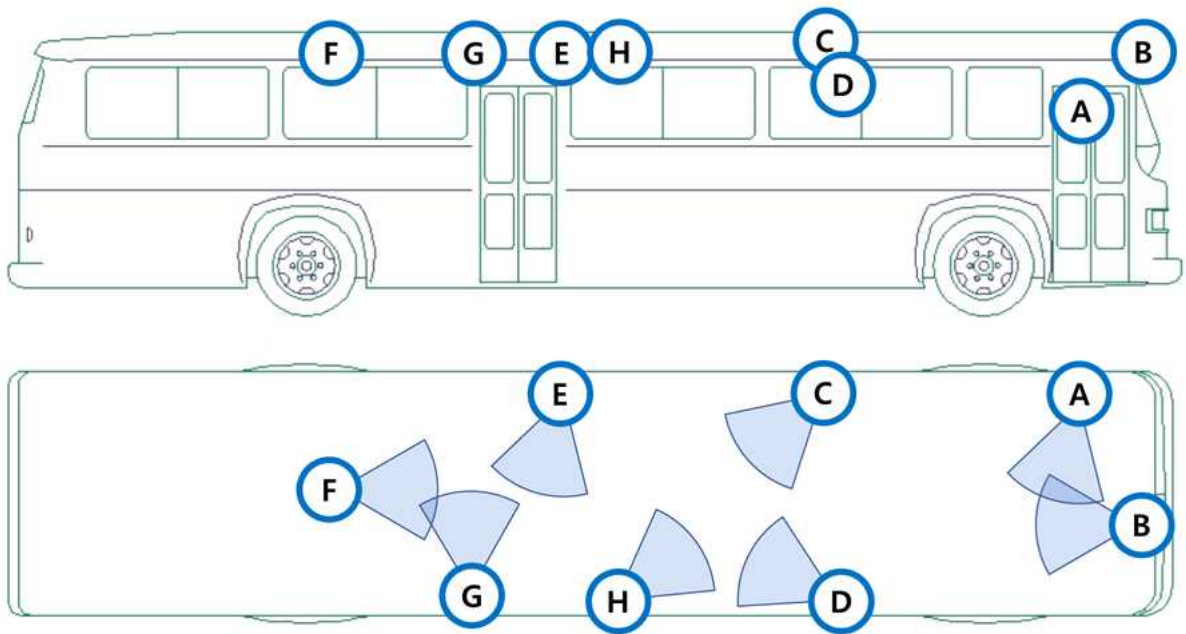


그림 4 특수 수집 영상의 카메라 배치

1.5 데이터 통계

1.5.1 데이터 구축 규모

영상 데이터의 규모 표현

- **정류소 승하차 영상:** 평균 40초 내외의 버스 정류장 정차 영상, 각 채널 2300개 내외, 이미지 80만 장

> HD 이상 해상도 평균 40초 내외 버스 정류장 2300건 내외, 3개 채널 합계 7000여개 내외의 비디오 클립, 각 채널별 10시간 내외 분량

> 3ch / 평균 40초 / 3fps 이미지 추출 / 채널당 평균 2300 클립

$$\begin{aligned}\text{일반 취득 이미지 수} &= 3\text{ch} * 40\text{s} * 3\text{fps} * 2300 \\ &= 828,000\end{aligned}$$

- **특수 행동 영상:** 사고 영상 시나리오 6개, 폭행 영상 시나리오 6개, 와 개인정보 동의를 받은 다수의 배우를 섭외한 보건 데이터셋을 위한 영상. 8개의 채널로 10만장 내외의 이미지 획득

보건 영상의 경우 마스크를 쓰고 벗거나 외투 악세서리를 변경하는 등의 1명의 배우가 3회의 촬영을 주 야간으로 반복한다.

> 폭행영상 규모

8ch / 평균 45초 / 3fps 이미지 추출 / 6개 시나리오 / 시나리오 당 6회 촬영(배우변경)

$$\begin{aligned}\text{폭행 취득 이미지 수} &= 8\text{ch} * 45\text{s} * 3\text{fps} * 6 * 6 \\ &= 38,800\end{aligned}$$

> 사고 영상 규모

8ch / 평균 20초 / 3fps 이미지 추출 / 6개 시나리오 / 시나리오 당 8회 반복 (배우변경)

$$\begin{aligned}\text{사고 취득 이미지 수} &= 8\text{ch} * 20\text{s} * 3\text{fps} * 6 * 8 \\ &= 23,040\end{aligned}$$

> 보건 영상 규모

8ch / 평균 60초 / 3fps 이미지 추출 / 3가지 복장 및 마스크 조합 변경 /

$$\begin{aligned}\text{보건 취득 이미지 수} &= 8\text{ch} * 60\text{s} * 3\text{fps} * 2 * 14 \\ &= 40,320\end{aligned}$$

1.5.2 데이터 분포

- 구축 데이터 이미지 장 수: 900,000 장 (승하차 80만, 특수 행동 10만)

대분류 (%)	
일반	특수
88	12

- 소분류 (취득정류장/특수상황) 클립 수 분포

정류장	아침	출근	낮	퇴근	밤	계
마을 앞	25 이상	60 이상	60 이상	60 이상	60 이상	265 이상
아파트 앞	25 이상	60 이상	60 이상	60 이상	60 이상	265 이상
변화가 앞	25 이상	60 이상	60 이상	60 이상	60 이상	265 이상
업무지구 앞	25 이상	60 이상	60 이상	60 이상	60 이상	265 이상
병원 앞	25 이상	60 이상	60 이상	60 이상	60 이상	265 이상
대학교 앞	25 이상	60 이상	60 이상	60 이상	60 이상	265 이상
중고교 앞	25 이상	60 이상	60 이상	60 이상	60 이상	265 이상
초등학교 앞	25 이상	60 이상	60 이상	60 이상	60 이상	265 이상
대형마트 앞	25 이상	60 이상	60 이상	60 이상	60 이상	265 이상
지하철 앞	25 이상	60 이상	60 이상	60 이상	60 이상	265 이상
기타	25 이상	25 이상	25 이상	25 이상	25 이상	125 이상

- 소분류 (특수상황) 클립 분포 상세

시나리오	폭행 시나리오	시간	횟수
S#01	마스크 제지하는 기사폭행 남성	45s	6 이상
S#02	이유 없이 핸드백으로 기사 폭행	45s	6 이상
S#03	기사와 말싸움 하던 남성이 핸들을 돌림	45s	6 이상
S#04	두 명의 승객이 기사 폭행	45s	6 이상
S#05	1 열에 앉은 여성이 기사에게 오물 투척	45s	6 이상
S#06	운전석 착석 전 기사 폭행	45s	6 이상

시나리오	사고 시나리오	시간	횟수
S#07	뒷좌석에 앉은 여성이 급정거에 버스 통로로 떨어진다.	20s	6 이상
S#08	여행용가방을 놓친 승객으로 앞좌석 승객이 다친다.	20s	6 이상
S#09	급출발로 자리에 앉지 못한 승객이 쓰러진다.	20s	6 이상
S#10	손잡이를 안잡은 승객이 쓰러지며 옆승객도 쓰러진다.	20s	6 이상
S#11	주행중인 버스를 건넌 할머니가 중심을 잃고 쓰러진다.	20s	6 이상
S#12	좌석에서 비져나온 승객의 발에 걸려 넘어지는 학생.	20s	6 이상

시나리오	사고 시나리오	시간	횟수
S#13	주간 마스크 승객 승 하차(개인정보 동의)	60s	14
S#14	야간 마스크 승객 승 하차(개인정보 동의)	60s	14

1.6 원시데이터 특성

1.6.1 대상분류

- 실제 : 현재 운용 중인 버스의 감시카메라로부터 실제 운용 영상을 획득
시뮬레이션 : 사고, 운전사 폭행 등의 상황은 유희 차량을 활용, 실제 시뮬레이션으로 획득
: 보건을 위한 마스크 착용 유무에 대해서는 개인정보 보호 동의를 얻은
사람에 한하여 공개가 가능하므로, 개인정보 이용에 관한 동의를 얻은
촬영자들을 섭외, 시뮬레이션 상황에서 취득함

1.6.2 제약조건

N/A

1.6.3 속성

Video Format : HD 이상 / 6~15 FPS

Image Format : HD 이상 / 비디오 파일로부터 3 FPS 추출 /

1.7 기타정보

1.7.1 포괄성

서울시의 서로 다른 20개의 노선을 선정, 요일, 시간, 정류장의 다양성을 확보

마을버스, 시내버스의 데이터 5:5의 비율로 획득

획득이 불가능한 상황에 대해서는 실제 버스를 이용하여 배우의 연출

개인정보 침해 가능성을 배제하기 위해서 마스크 검출을 위한 마스크 미착용 데이터 취득은 배우를 섭외하여 개인정보 동의 하에 촬영함

1.7.2 독립성

개인정보보호에 관한 법적 문제 있음: 비식별화 처리가 필요함. 상세 규정 준비 중

1.7.3 유의사항

파급효과 : 대중교통 승객에 대한 안전 시스템 구축

유의사항 : 개인정보 보호에 관한 법률 유의

사용처와 목적에 따라 원천 데이터의 소유자와 사전 협의 필요성도 있음

개인정보 취급에 관한 규정, 개인정보 보호 요청에 대한 대응 규정 준비 필요

2. 데이터 구축 가이드

2.1 데이터 구축 개요

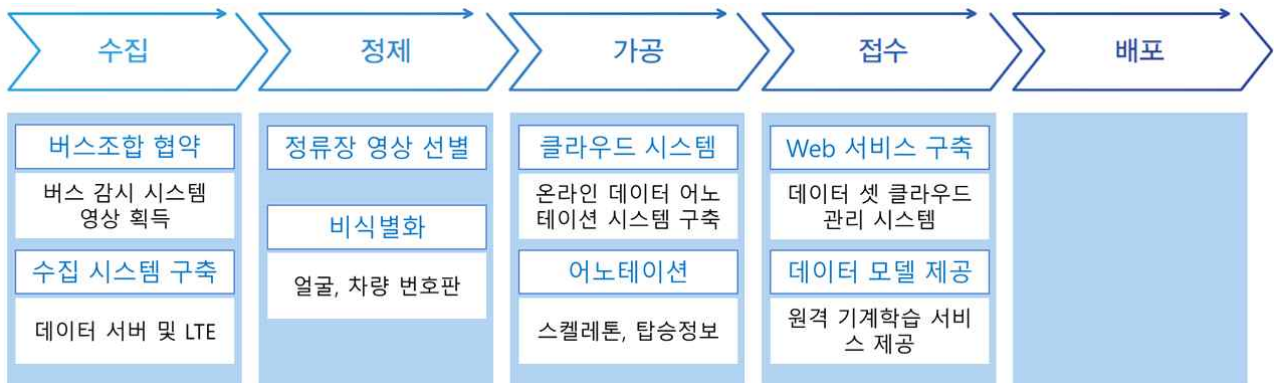


그림 5. 데이터 구축 순서도

2.2 문제정의

2.2.1 임무 정의

핵심 임무 : 버스의 안전한 승하차 시스템 구축

개문발차로 대표되는 승객이 승하차가 완료되지 않은 상황에서의 버스의 출발, 그리고 아직 탑승을 해야하는 승객, 하차를 해야하는 승객이 남아있으나 버스가 먼저 출발하는 경우를 방지하기 위하여 탑승자(미탑승 의사의 보행자 포함)의 위치, 스켈레톤 정보를 취득하고 이를 LSTM과 같은 RNN모델에 적용 하면 정류소에 있는 승객이 탑승의도를 갖고 있는지, 후문 근처 승객이 하차의도를 갖고 있는지 예측 할 수 있음. 또한 앞문 뒷문에서 승차중인 사람, 하차중인 사람을 정확히 구분해 개문발차 사고를 방지하는 것으로 목표로 함

기타 임무1 : 마스크 미착용 승객 검출

현재 호흡기성 전염병과 같은 보건 위기 상황에서 대수의 사용자가 시설을 공유하는 대중교통 시설은 마스크 착용 여부가 전염병의 확산을 막고 국민 건강을 지키는데 매우 중요한 역할을 작데 됨. 이를 위해 시로 마스크 착용 유무를 인식하는 시를 개발하고자 하는 연구자에게 실제 운용중인 대중교통의 감시 카메라를 통해 획득된 마스크 착용, 미착용 어노테이션 영상을 제공하고자 함.

기타 임무2 : 버스 내 사고, 폭력 상황 검출

버스 내에서 사고, 및 폭행 상황의 발생은 향후 차량의 supervisor인 기사가 없는 무인 버스에서 즉각적인 대응이 되지 않아 사건이 악화되는 골든 타임을 놓치는 등의 큰 문제가 될 수 있음. 이에 버스 내에서 발생할 수 있는 폭력이나 사고 상황을 재현하고 이 분야의 국내 연구와 자율주행 버스의 구축을 돕기 위해 스텔레톤 정보를 어노테이션 수행하여 데이터셋으로 구축하고자 함.

2.2.2 데이터 구축 유의사항

본 데이터는 비식별화 처리가 되어있으나 관련 법률에 의거 제삼자로의 공개는 제약됩니다.

본 데이터는 서울시 버스조합과의 협약에 의거 구축되었으며 상업적 이용에 관한 일부 권리는 서울시 버스조합과 협의가 필요 할 수 있습니다.

본 데이터는 실제 운용 중인 대중교통의 영상을 활용합니다. 따라서 본 데이터는 개인정보 보호를 위하여 얼굴 및 차량 번호판에 비식별화 처리가 되어있음에 유의해 주시기 바랍니다.

2.3 획득·정제

2.3.1 원시데이터 선정

원시 데이터

버스의 감시카메라에서 촬영된 비디오 정보는 독자 포맷과 소프트웨어로 요일 및 시간별로 접근 및 추출을 할 수 있다.

원시 데이터는 외장하드에 기록되며 하드 기록은 최소 1주일 분량 이상이 기록할 수 있다. 현재 준비 중인(최종 정정 할 것) 시내버스 10대, 마을버스 10대로부터 취득, 10대는 모두 중복되지 않도록 다양한 지역의 노선을 선택한다. 시간, 환경 변수는 실제 대중교통의 영상인 만큼 충분히 다양성을 확보할 수 있을 것으로 기대되나 돌발상황, 폭력 사태 등의 영상은 본 사업기간내에 충분히 획득하기 어려우므로 유휴 차량에서 배우들의 동작으로 획득 한다.

2.3.2 획득·정제 절차

실제 운용중인 버스 감시 영상 획득 절차

- 가. 감시 시스템으로부터 외장하드 회수
- 나. D-TEG 데이터 서버로 전송
- 다. 대분류, 소분류 기준에 따른 비디오 클립 추출
- 라. 영상의 정합성 판단 (승하차 상황이 아니거나 정류장을 그냥 지나치는 경우 등)
- 마. 소프트웨어를 이용한 영상의 메타 데이터 추출
- 바. 정합성은 충족하다 분류에 적합하지 않은 영상 선별 (이미지 품질, 과편향 등)

배우 섭외를 통한 시뮬레이션 영상 획득 절차

- 가. 촬영 스케줄 및 배우 동작 시나리오 작업
- 나. 현장 섭외(유휴 시내버스) 및 촬영 세트장 구성
- 다. 특수 영상 촬영을 위한 카메라 증설 (6채널)
- 라. 촬영 감독의 통제 하에 take 촬영
- 마. 각 take별 비디오 클립 추출
- 바. 정합성 및 원천 데이터 획득 기준에 근거한 유효 클립 선정





RE-01	화질 열화	영상 취득	A,B,C	비디오 품질 저하 현상 이 발생하는 경우 (깍두기 / 컬러채널불량 / 과도한 잔상 / 모션블러)	
RE-02	승객 과밀	승객 수	C	C 채널에서 한 화면에 서있는 승객 수 7명 이상인 경우 (6인 초과)	
RE-03	승객 없음	승객 수	A,B,C	모든 채널에서 승차객이 한명도 발생하지 않음	
RE-04	역광 번짐	영상 취득	A,B,C	카메라 빛이 물체식별이 어려울 정도로 포화 됨	N/A
RE-05	영역 불량	영상 취득	A	주요 승객의 신체 부위가 영상 내내 20% 미만으로 등장 할 경우	

표 4. 원천데이터 제약조건 평가 규정 rev. 0.1 (2020/10/28)

2.3.3 획득·정제 기준

획득 기준

- 가. 버스가 GPS를 통한 차속이 0이 되고 정류장에 멈춘 경우를 소프트웨어를 통해 식별, 정차 전 15초, 발차 후 15초의 영상을 비디오 클립으로 정제함
- 나. 사전에 정의한 5개의 시간 대역 구분, 11개의 장소 구분이 고르게 분포되도록 영상을 2배수로 선별한다.
- 다. 버스 내 승객 수도 한 다양하고 고르게 분포 가능하도록 데이터를 선별 획득한다.

2.3.4

2.3.5 획득·정제 조직



그림 6. 원천데이터 획득 조직 구성도

2.3.6 획득·정제 도구

2.3.7

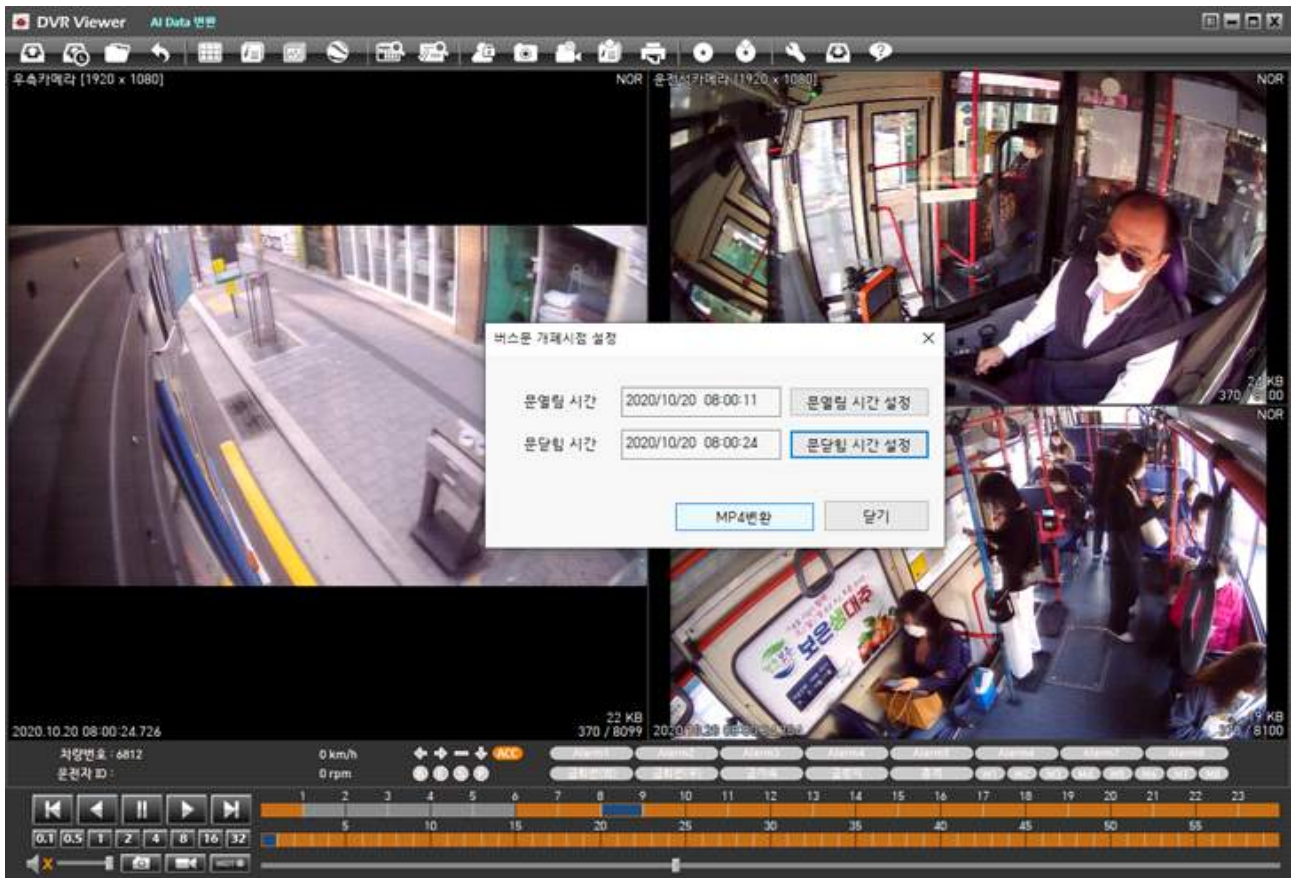


그림 7. 원천데이터 획득 정제용 비디오 추출 소프트웨어

2.4 어노테이션/라벨링

2.4.1 어노테이션/라벨링 절차

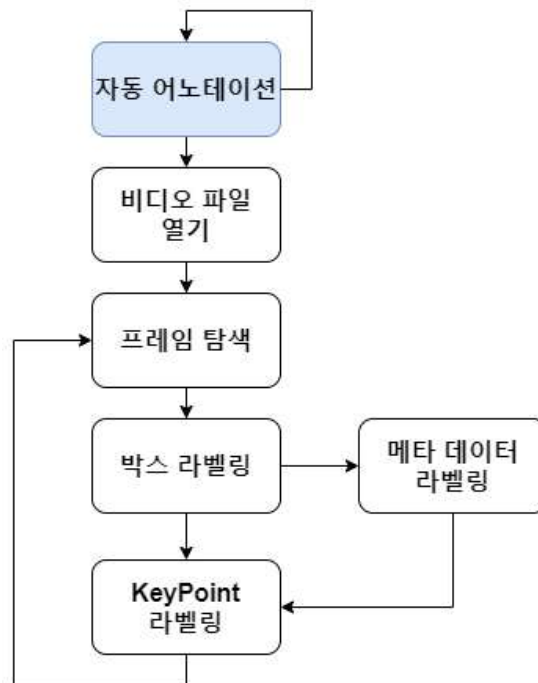


그림 8. 어노테이션 절차 순서도

1. 자동 어노테이션 과정

수집된 원천 영상 데이터에 대해 미리 학습된 인공지능 모델을 사용하여 Bounding Box 및 Skeleton Key Point를 1차 어노테이션 진행한다.

2. 어노테이션 저작도구를 이용한 데이터 오픈

어노테이션 저작도구 TAGGING TOOL 1.9.1을 이용하여 영상폴더를 선택하고 이미지 목록을 오픈한다.

(* 어노테이션 저작도구를 이용한 데이터 오픈은 첨부 비디오 파일 "[채널A 작업가이드] 1. 동영상 전체를 먼저 살펴보기.mp4"을 참조)

3. Bounding Box 어노테이션 과정

선택된 영상에서 승객마다 Bounding Box를 어노테이션 한다. 작업자는 자동 어노테이션된 결과를 확인하며 미검출 및 오검출된 데이터를 수정한다.



그림 9. 박스 어노테이션 과정

4. 메타 데이터 어노테이션 과정

작업자는 어노테이션한 Bounding Box에 해당하는 객체의 기타 메타 데이터를 라벨링 한다.

(* 메타 데이터 어노테이션은 첨부 비디오 파일"[채널A 작업가이드] 3_ 객체 정보 체크하기.mp4"을 참조)

순번

3

연령대

아동

행동

서있는 상태

마스크 착용 여부

☐ 착용

장애 여부

☐ 장애

승차 여부

☐ 승차

하차 여부

☐ 하차

5. Key Point 어노테이션 과정

Bounding Box 어노테이션이 완료된 객체를 대상으로 승객별 Skeleton을 어노테이션 한다. 작업자는 자동 어노테이션 결과에서 오검출된 Skeleton 위치 정보 및 미검출된 Skeleton 정보를 수정한다.

(* 키포인트 어노테이션 과정은 첨부 비디오 파일"[채널A 작업가이드] 2_스켈레톤 박스 정보 작업하기.mp4"을 참조)



6. 비식별화 처리

어노테이션 작업이 완료된 데이터를 대상으로 비식별화 처리를 진행한다. 비식별화 대상은 승객의 안면 이미지, 주변 차량의 번호판이 해당되며, 미리 학습된 인공지능 모델로 비식별화 대상 영역을 검출한다. 작업자는 인공지능 모델이 추출한 비식별화 영역을 검수하여 미검출 및 오검출 된 영역을 수정한다. 검수 및 수정이 완료된 데이터에 Blur 처리를 하여 비식별화 처리를 한다.

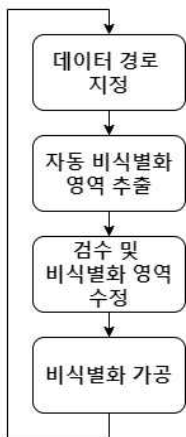



그림 11. 비식별화 과정

2.4.2 어노테이션/라벨링 기준

바운딩 박스 및 승객의 Skeleton Key Point 어노테이션은 MS COCO (Common Objects in Context) 데이터셋의 형식을 따라 구성한다.

바운딩 박스 및 Key Point 어노테이션 시 Ground Truth 세부 기준은 아래 표와 같다.

바운딩 박스 어노테이션 세부 기준

기준	세부 내용
대상	<p>이미지에 보이는 승객 객체를 모두 어노테이션 하되 아래의 경우는 어노테이션 하지 않음.</p> <ul style="list-style-type: none"> 객체 식별이 불가능한 경우(객체나 카메라가 움직여 너무 흐리게 보이는 경우) 객체의 가로 너비가 32 px 이하인 경우 (카메라로부터 원거리에 있는 경우)
박스 그리기	<p>객체의 보이는 부분이 모두 포함되도록 그리며, 사각형의 경계선과 객체의 경계에 빈틈이 없도록 그리는 것을 원칙으로 함.</p> <ul style="list-style-type: none"> 해당 객체가 잘려서 촬영된 경우 이미지 가장자리까지 사각형을 그림. 해당 객체가 다른 객체에 의해 겹쳐 보이는 경우 객체의 온전한 모양을 추정하여 사각형을 그려야 함.  <p>6531 X: -0.016 Y: 0.020 Z: -0.062 2020.02.25 18:38:30.762</p> <p>Left Right 15 rpm 0 km/h C [N] 37°28'12.1"N 127°07'38.39"E</p> <ul style="list-style-type: none"> 객체가 군집을 이루고 있는 경우 객체 하나씩 모두 사각형을 그려야 함.
가려짐	객체끼리 겹치거나 카메라 화각으로 인해 가려진 경우 아래 기준에 따라 어노테이션 함.

	기준	어노테이션
	객체가 95% 보임	가려진 부분까지 추정하여 사각형을 그림
	객체가 20~95% 보임	가려진 부분까지 추정하여 사각형을 그리고 "Occluded" 처리
	객체가 20% 이하로 보임	객체에 사각형을 그리지 않음
이미지	이미지 크기가 1280*720 보다 작은 경우 해당 영상을 사용하지 않음.	
반사 및 광고물	버스 및 정류장에 부착된 거울, 유리, 광고물 등에 반사되거나 인쇄되어있는 객체는 어노테이션 하지 않음.	

Key Point 어노테이션 세부 기준

기준	세부내용	
대상	Bounding Box 데이터가 존재하는 승객 객체의 16개 Keypoint 전부를 어노테이션 한다.	
Keypoint 어노테이션	Keypoint 순서에 맞는 관절 위치에 Keypoint가 위치하도록 어노테이션 하는 것을 원칙으로 한다. <ul style="list-style-type: none"> 일부 관절이 화각에 의해 잘려서 촬영된 경우 보이는 관절만 전부 어노테이션 한다. 일부 관절이 다른 객체에 의해 가려진 경우 관절 위치를 추정하여 어노테이션 한다. 	
가려짐 및 관절 연결	객체끼리 겹치거나 카메라 화각으로 인해 관절이 가려진 경우, 아래 기준에 따라 어노테이션 함.	
	기준	어노테이션
	해당 관절이 보임	관절을 정위치에 어노테이션 함.
	해당 관절이 화각 내에 있으나 가려짐	가려진 부분의 관절 위치를 추정하여 어노테이션 하고 "not visible"처리.
	해당 관절이 화각 밖으로 나감	보이지 않는 관절을 "delete" 처리.
이미지	이미지 크기가 1280*720 보다 작은 경우 해당 영상을 사용하지 않음.	
반사 및 광고물	버스 및 정류장에 부착된 거울, 유리, 광고물 등에 반사되거나 인쇄되어 있는 객체는 어노테이션 하지 않음.	

기타 메타데이터 어노테이션 세부 기준

기준	세부내용	
대상	Bounding Box 데이터가 존재하는 승객 객체의 프레임별 ID 및 승하차 여부, 마스크 착용 여부, 행동 상태 등 메타 데이터를 어노테이션 한다.	
ID 부여	영상의 첫 프레임부터 마지막 프레임까지 Bounding Box 데이터가 존재하는 승객 객체의 ID를 어노테이션 한다. <ul style="list-style-type: none"> ID는 0번부터 부여한다. 연속한 프레임에서의 동일한 객체는 동일한 ID를 부여한다. 프레임에서 여러 객체가 동시에 존재하고 새 ID를 부여하는 경우, 이미지 좌상단부터 우하단 방향으로 ID를 증가시키며 부여한다. 새로운 객체가 나타나는 경우 ID 번호를 증가시킨다. ID가 부여된 객체가 프레임에서 사라지는 경우, 해당 ID는 다시 사용하지 않는다. 	

승하차 여부	<p>영상의 첫 프레임부터 마지막 프레임까지 Bounding Box 데이터가 존재하는 승객 객체의 승하차 여부를 어노테이션 한다.</p> <ul style="list-style-type: none"> 영상에서 객체가 나타나는 프레임부터 사라지는 프레임까지 대상이 승하차 여부를 어노테이션 한다. 채널 A, B 영상에서, 버스가 정류장 접근 시점부터 객체가 카메라 화각 밖으로 나가는 시점까지 버스에 실제 탑승하는 승객 객체를 "get_on"으로 처리한다 채널 C 영상에서 영상 시점부터 승객이 하차할 때까지 실제 하차하는 승객 객체를 "get_off"으로 처리한다. 이외 버스에 승차하지 않거나, 하차하지 않는 승객은 어노테이션하지 않는다.
기타 데이터	영상의 첫 프레임부터 마지막 프레임까지 Bounding Box 데이터가 존재하는 승객 객체의 기타 데이터(연령대, 마스크, 행동 등)을 어노테이션 한다.
이미지	이미지 크기가 1280*720 보다 작은 경우 해당 영상을 사용하지 않음.
반사 및 광고물	버스 및 정류장에 부착된 거울, 유리, 광고물 등에 반사되거나 인쇄되어 있는 객체는 어노테이션 하지 않음.

2.4.3 어노테이션/라벨링 조직

TBD 조직 구성 완료 시 추가

2.4.4 어노테이션/라벨링 도구

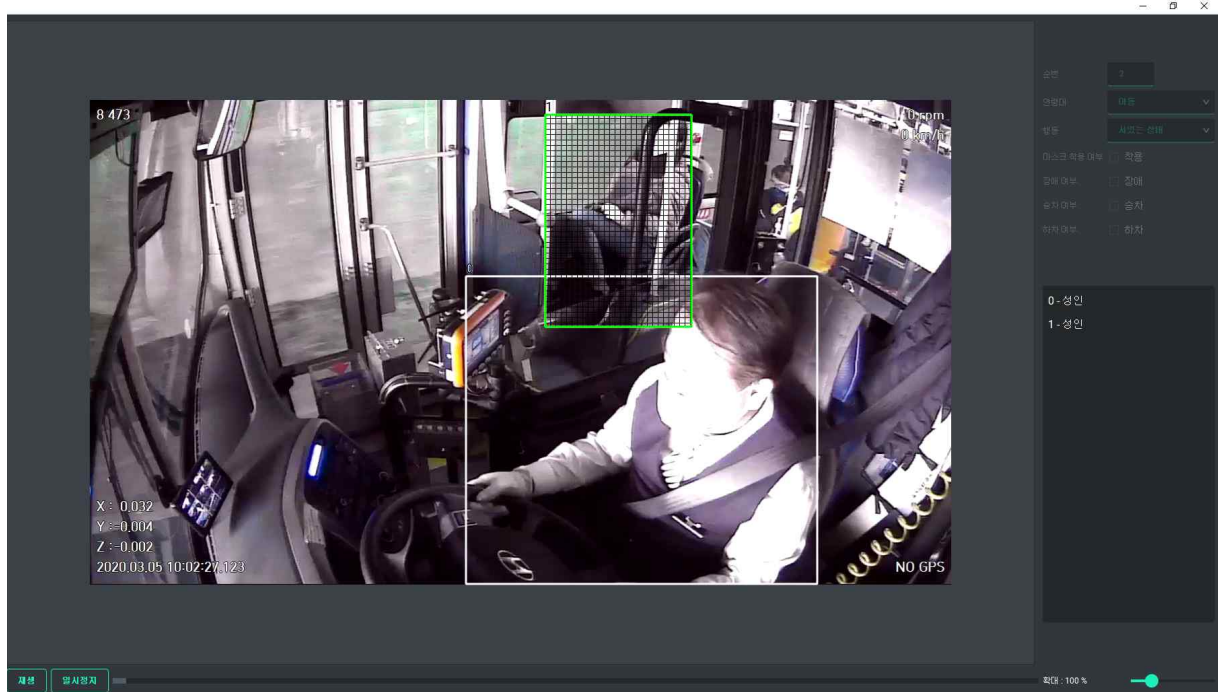


그림 12. 승객 어노테이션 과정

버스 승객 데이터 어노테이션 툴

본 어노테이션 툴은 영상 클립에 대해 버스 승객 승하차 데이터셋 가공을 위한 편집 도구이다. 작업자는 동영상 파일 입력 후 프레임 이미지의 사람 객체의 Bounding Box 및 Key point, 기타 메타 데이터를 가공하여 Json 포맷의 데이터를 생성할 수 있다.

주요 기능

영상 클립에 대해 정해진 간격으로 Frame을 점프하며 어노테이션 작업 .

이미지 또는 영상 클립 내에 원하는 객체의 라벨링 및 Bounding Box 표시 등의 어노테이션 작업.

Bounding Box 객체에 대해 영상 내 ID 부여, 마스크 착용 여부, 연령대, 현재 동작 정보, 승하차 여부에 대한 기타 메타데이터의 어노테이션 작업.

사람 객체에 대한 Key Point 어노테이션 작업

이미지 크기 확대 / 축소 기능

라벨링 된 데이터의 편집 기능

2.5 검수

검수 절차는 크게 4단계의 검증을 거치며 1차 검수는 가공 인력간에 교차 전수 검수, 2차검수부터는 검수 전문 인력이 담당하게 된다.



[단계별 품질보증활동 종합]

2.5.1 검수 절차

전체 학습용 데이터의 품질 관리 및 검증을 모두 외부기관에 의뢰하거나 정밀 검수를 수행 할 경우 그 금전적, 시간적 비용은 천문학적으로 증가할 수 있다. 따라서 다음 5단계 방안을 순차적으로 진행하고자 한다.

방안 a. 작업자의 전수 교차 검사

작업자는 작업 종료전 당일 다른 작업자와 결과를 전수 교환 교차 검증을 수행한다.

검수절차는 검수 전용 소프트웨어를 활용하며 소프트웨어를 통해 검수 결과를 csv로 축적, 검수 이력 및 오류 보고서를 작성한다.

	A	B	C	D	E	F	G	H	I	J	K	L
1		관리		파일	프레임	ID		불량 체크				
2	인덱스	처리일	검수자	비디오 이름	시작	종료	번호	박스	스켈	태깅	규정	기타
3	1	14/01/2021 11:44	이슬비	[mid_high]morning_059A	318	324	0				V	
4	2	14/01/2021 13:32	이슬비	[university]day_005A	355	355	9				V	
5	3	14/01/2021 13:33	이슬비	[university]day_005A	300	315	7			V		
6	4	14/01/2021 13:33	이슬비	[university]day_005A	320	330	8			V		
7	5	14/01/2021 14:09	이슬비	[ele]leave_017B	100	126	6			V		
8	6	14/01/2021 14:09	이슬비	[ele]leave_017B	126	126	6			V		
9	7	14/01/2021 15:00	이슬비	[subway]night_051A	210	216	4			V		
10	8	14/01/2021 15:01	이슬비	[subway]night_051A	282	282	4			V		
11	9	14/01/2021 15:58	이슬비	[mid_high]day_009B	108	110	11	V				
12	10	14/01/2021 16:38	이슬비	[apt]attend_074C	0	114	1			V		
13	11	14/01/2021 16:57	이슬비	[supermarket]leave_013A	66	93	0			V		
14	12	14/01/2021 17:44	이슬비	[supermarket]leave_013A	198	198	2			V		
15	13	14/01/2021 17:47	이슬비	[supermarket]leave_013A	255	336	4			V		
16	14	14/01/2021 17:48	이슬비	[supermarket]leave_013A	345	390	4			V		
17	15	14/01/2021 17:48	이슬비	[supermarket]leave_013A	420	429	4			V		
18	16	14/01/2021 18:36	이슬비	[district]day_025C	48	48	2				V	
19	17	14/01/2021 18:43	이슬비	[district]day_025C	72	74	3				V	
20	18	14/01/2021 18:45	이슬비	[district]day_025C	76	78	4				V	
21	19	14/01/2021 18:48	이슬비	[district]day_025C	132	136	6				V	
22	20	14/01/2021 18:56	이슬비	[district]day_025C	196	200	8				V	
23	21											
24	22											

그림 18 검수용 소프트웨어를 이용한 검수자 검수 이력 및 정정 결과



그림 14 검수 방안 a. 전수 교환 검사 절차

방안 b. 자체 검수 전담 인력 배치

학사 이상의 데이터 분야 전공자를 자체 검수 전담 인력으로 배치하고 매주 추가분 어노테이션 데이터에 대하여 20% 내외의 랜덤 표본 샘플 검수를 진행한다.

검수는 어노테이션의 정확도뿐 아니라 데이터의 중복성 또한 평가 하도록 한다.

2단계의 검수 또한 1단계의 검수와 동일한 소프트웨어를 활용하난.

방안 c. 데이터 품질 검수 시스템 운용

전체 데이터의 통계 모델을 산출하는 소프트웨어를 개발하고 이를 기반으로 데이터의 유효성을 검수 하고 outlier를 찾아낸다. (예시. 전체 신장 길이 대비 팔의 길이 분포모델)



그림 15 검수 방안 b. 전담인력의 표본 검수 절차


```

1 2021-01-21 13:25:36,[apt]attend_1427B_104.json,2,Out Of
2 2021-01-21 13:25:36,[apt]attend_1427B_104.json,2,Out Of
3 2021-01-21 13:25:36,[apt]attend_1452B_124.json,3,Out Of
4 2021-01-21 13:25:36,[apt]attend_1462B_130.json,0,Out Of
5 2021-01-21 13:25:36,[apt]night_1478B_108.json,0,Out Of B
6 2021-01-21 13:40:33,[district]day_629B_122.json,1,Out Of
7 2021-01-21 13:40:33,[district]night_536B_102.json,1,Out
8 2021-01-21 13:40:33,[district]night_536B_104.json,1,Out
9 2021-01-21 13:40:33,[district]night_536B_108.json,1,Out
10 2021-01-21 13:40:33,[district]night_536B_108.json,1,Out
11 2021-01-21 14:06:42,[downtown]day_213B_120.json,2,Out Of
12 2021-01-21 14:06:42,[downtown]day_220B_116.json,1,Out Of
13 2021-01-21 14:06:42,[downtown]leave_157B_110.json,0,Out
14 2021-01-21 14:23:35,[ele]night_334B_260.json,0,Out Of Bo
15 2021-01-21 14:35:07,[etc]attend_037B_275.json,1,Out Of B
16 2021-01-21 14:44:59,[hospital]day_417B_320.json,2,Out Of
17 2021-01-21 14:54:02,[mid_high]attend_252B_98.json,2,Out
18 2021-01-21 14:54:02,[mid_high]night_141B_295.json,2,Out
19 2021-01-21 15:14:19,[subway]morning_1128B_255.json,0,Out
20 2021-01-21 15:23:59,[supermarket]day_131B_178.json,0,Out
21 2021-01-21 15:36:42,[university]attend_124B_245.json,0,0
22 2021-01-21 15:36:42,[university]day_090B_265.json,1,Out
23 2021-01-21 15:36:42,[university]leave_183B_250.json,1,0u
24 2021-01-21 15:36:42,[university]leave_183B_260.json,1,0u
25 2021-01-21 15:36:42,[university]leave_183B_265.json,1,0u
26 2021-01-21 15:36:42,[university]leave_183B_270.json,1,0u
27 2021-01-21 15:36:42,[university]leave_183B_275.json,1,0u
28 2021-01-21 15:36:42,[university]leave_183B_275.json,1,0u
29 2021-01-21 15:36:42,[university]leave_183B_280.json,1,0u
30 2021-01-21 15:36:42,[university]leave_183B_285.json,1,0u
31 2021-01-21 15:36:42,[university]leave_183B_290.json,1,0u
32 2021-01-21 15:36:42,[university]leave_183B_290.json,1,0u
33 2021-01-21 15:36:42,[university]morning_156B_94.json,1,0
34 2021-01-21 15:36:42,[university]morning_184B_485.json,0,
35 2021-01-21 15:51:11,[village]day_205B_290.json,1,Out Of
36 2021-01-21 15:51:11,[village]leave_186B_134.json,3,Out 0
37 2021-01-21 15:51:11,[village]leave_186B_134.json,3,Out 0
38 2021-01-21 15:51:11,[village]leave_189B_96.json,1,Out Of
39 2021-01-21 15:51:11,[village]leave_193B_134.json,4,Out 0
40 2021-01-21 15:51:11,[village]leave_193B_134.json,4,Out 0
41 2021-01-21 15:51:11,[village]morning_202B_2.json,0,Out 0
42 2021-01-21 15:51:11,[village]morning_202B_20.json,0,Out
43 2021-01-21 15:51:11,[village]morning_202B_28.json,0,Out

```

그림 21 3차 검수 로그 화면

방안 d. 유효성 검증 모델을 통한 데이터 정확성, 유효성 산출

k-fold cross validation을 기반으로 최종 데이터의 정확성, 유효성을 증명한다.

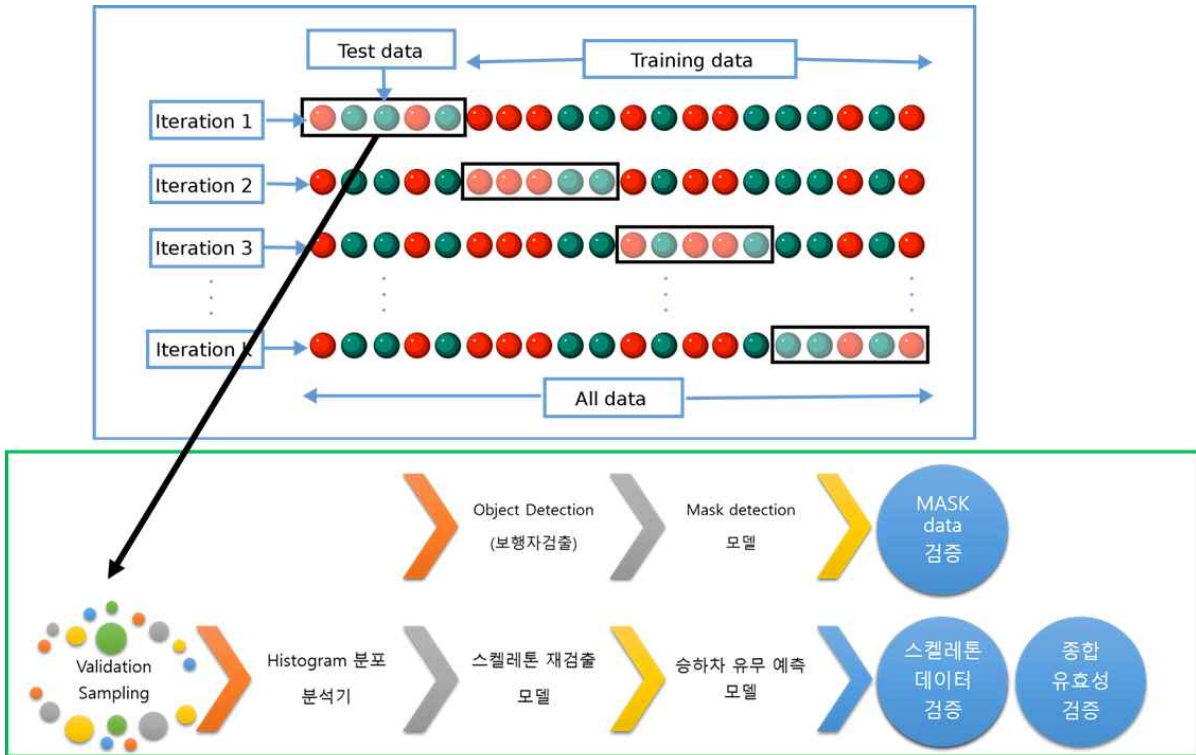


그림 16 검수방안 d. AI모델을 통한 유효성 시험 구성도

검수 작업 절차

- 1) 파일 및 영상 로드
- 2) 탑승객 어노테이션 여부 판단
- 3) 바운딩박스 위치 적합성 검수
- 4) 스켈레톤 적합성 검수
- 5) 탑승객의 식별정보 검수(ID, 성별, 연령, 행동정의, 신체부자유, 마스크착용)
- 6) 종료 및 검수 기록

2.5.2 검수 기준

각 검수자의 검수 기준 및 지침과 교육은 별첨 문서 [버스승하차데이터셋 구축] 검수 지침서.pdf를 참고한다.

2.5.3 검수 조직



그림 23 검수 조직도

2.5.4 검수 도구

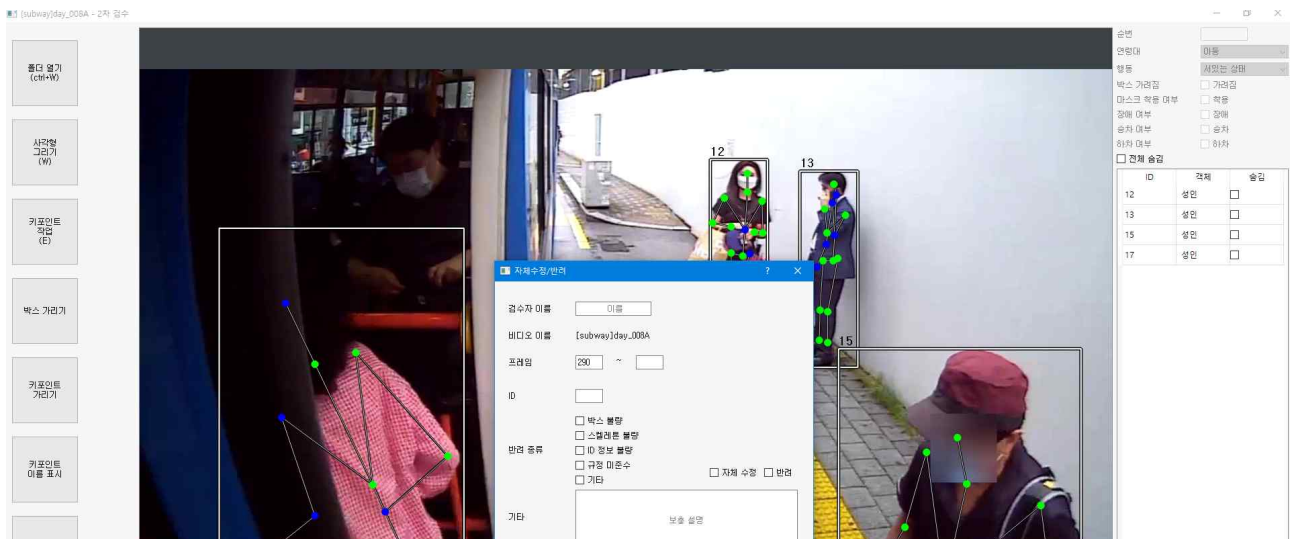


그림 24 검수 전용 소프트웨어

2.5.5 기타 품질관리 활동

검수 참여인력에 대한 전문 교육 인력 배치 및 교육 이수 과정 의무화하고 교육 이력과 교육관리 대장을 관리 한다.

원천데이터 획득 가이드 문서를 작성하고 교육을 수행한다.

어노테이션 작업 가이드북 문서를 작성하고 교육을 수행한다.

검수 지침서를 작성하고 교육을 수행한다.

2.6 활용

2.6.1 활용 모델

2.6.1.1 모델 학습

탐색객 검출 모델

YOLO는 오늘날의 객체탐지 시스템 개발 분야에서 가장 각광받는 모델 중 하나이다. YOLO는 시스템 개발자가 자신의 비전을 구축하고 반영하는데 있어서 표준적이고 이해하기 쉬운 방법으로 할 수 있도록 도와준다.

승객 검출모델

YOLO v4

Backbone

CSPDarknet53

주요 layer 구성

Layer	Filters size	Repeat	Output size
Image			416×416
Conv	$32 \times 3/1$	1	416×416
Conv	$64 \times 3/2$	1	208×208
Conv	$32 \times 1/1$	<div><div>Conv</div><div>Conv</div><div>Residual</div></div> $\times 1$	208×208
Conv	$64 \times 3/1$		208×208
Residual			208×208
Conv	$128 \times 3/2$	1	104×104
Conv	$64 \times 1/1$	<div><div>Conv</div><div>Conv</div><div>Residual</div></div> $\times 2$	104×104
Conv	$128 \times 3/1$		104×104
Residual			104×104
Conv	$256 \times 3/2$	1	52×52
Conv	$128 \times 1/1$	<div><div>Conv</div><div>Conv</div><div>Residual</div></div> $\times 8$	52×52
Conv	$256 \times 3/1$		52×52
Residual			52×52
Conv	$512 \times 3/2$	1	26×26
Conv	$256 \times 1/1$	<div><div>Conv</div><div>Conv</div><div>Residual</div></div> $\times 8$	26×26
Conv	$512 \times 3/1$		26×26
Residual			26×26
Conv	$1024 \times 3/2$	1	13×13
Conv	$512 \times 1/1$	<div><div>Conv</div><div>Conv</div><div>Residual</div></div> $\times 4$	13×13
Conv	$1024 \times 3/1$		13×13
Residual			13×13

Conv

Conv2d Layer

BN Layer

LeakyReLU Layer

Residual

Add

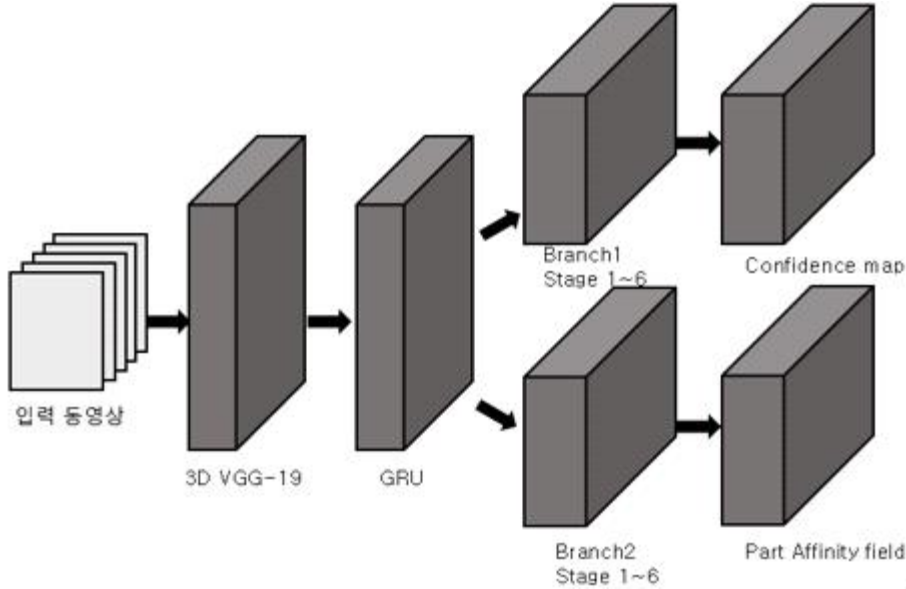
Conv
(1×1)

Conv
(3×3)

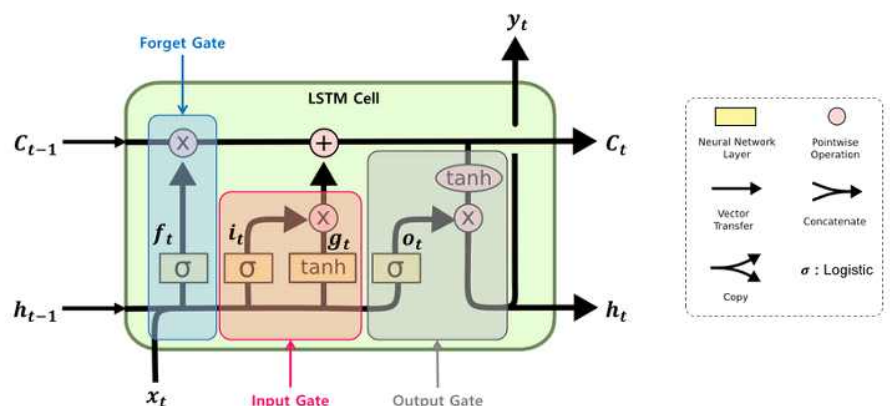
Conv
(3×3)

탐승객 행동 분류 모델

Openpose는 단일 이미지에서 인체의 신체, 손, 얼굴 및 발 키포인트(총 135개의 키포인트)를 공동으로 감지하는 최초의 실시간 다중 사용자 시스템이다. CMU(Carnegie Mellon Univ.)의 딥러닝 기반 오픈소스 프로젝트이다. 2D real-time으로 25개의 keypoint로 detection을 수행하였다. stage 1~6에서 이미지 속 사람 관절의 위치를 파악하고 운동 방향을 파악하여 특징을 추출한다.

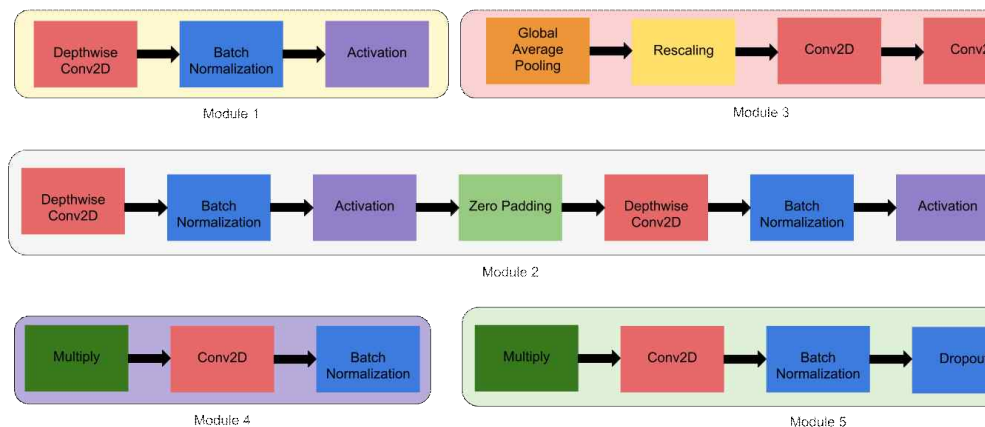
스켈레톤모델	Openpose
Backbone	VGG-19
주요 layer 구성	 <p>The diagram illustrates the Openpose architecture. It starts with '입력 동영상' (Input video) represented by a stack of images. This feeds into a '3D VGG-19' backbone, which then connects to a 'GRU' (Gated Recurrent Unit) block. From the GRU, the flow splits into two parallel branches: 'Branch1 Stage 1~6' and 'Branch2 Stage 1~6'. Branch1 leads to a 'Confidence map', and Branch2 leads to a 'Part Affinity field'.</p>
데이터셋	CMU Panoptic Studio dataset
저작권	CMU Panoptic Studio dataset is shared only for research purposes, and this cannot be used for any commercial purposes.

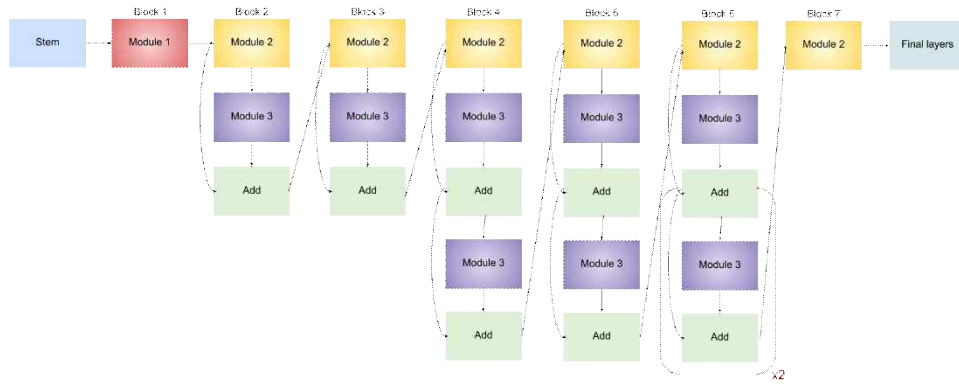
RNN의 주요모델 중 하나로, 순서적인 측면을 고려해서 판단할 수 있는 특성인 Sequence data를 다루는데 도움이 되는 모델이다. 이러한 데이터들을 시계열 데이터라 하는데 직전 데이터 뿐만 아니라 좀 더 거시적으로 과거 데이터를 고려하여 미래의 데이터를 예측하기 위해 활용된다. 하나의 네트워크가 여러 개 복사된 형태로 구성되어 각 네트워크는 다음 단계로 특성을 넘겨주는 형태이다.

승하차 예측모델	LSTM
Backbone	RNN
주요 layer 구성	
데이터셋	FS bus-passenger dataset
저작권	FSsolution - 확인 필요

탑승객 마스크 착용 유무 모델

다크넷 프레임워크와 YOLOv3 Tiny PRN 아키텍처를 기반으로 한 실제 이미지에 대한 빠른 장면 마스크 감지 모델이다. 다크넷은 C와 CUDA로 작성된 오픈소스 신경 네트워크 프레임워크이며 CPU와 GPU 연산을 지원한다. YOLO는 한 번에 한 이미지의 모든 객체를 감지한다.

보건 모델	YOLO Mask Detection
Backbone	EfficientNet-B0
주요 layer 구성	



데이터셋	VictorLin000
저작권	Free

2.6.1.2 서비스 활용 시나리오

버스 내외부에 설치된 카메라로 촬영된 보행자 및 승하차객들에 대한 데이터셋이다. AI trainset에 유의미하게 Human joint(skeleton)가 잘 나타나는 구역에 한하여 취득하였다. 객체가 사람인지 판단하는 Bounding Box와 각 객체의 skeleton을 활용하여 object tracking을 실행하였다. 앞서 말한 세가지 모델을 활용하여 버스에 탑승할 사람인지 아닌지 여부를 예측하는 버스 승하차 예측 응용서비스(1)를 제공한다.

또한, 각 객체의 meta data로 마스크 착용 여부를 활용하여 보건 응용서비스(2)를 제공한다.

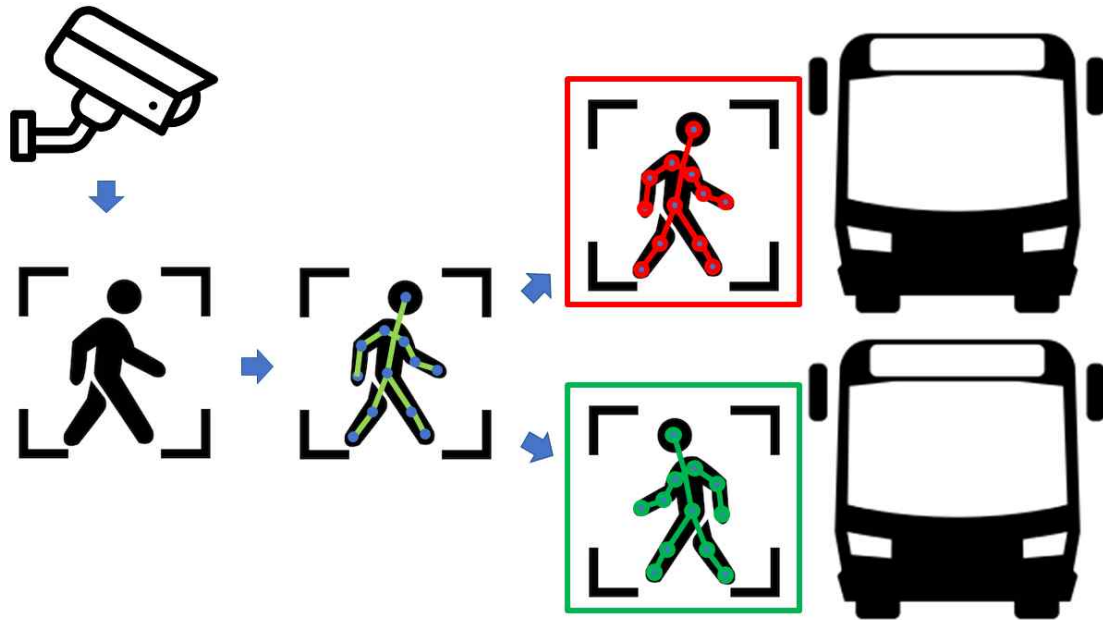


그림 17. 응용 서비스 예시 모식도



카메라 영상으로부터 사람의 객체를 분류

분류된 각 객체를 tracking

CenterNet skeleton, 성별, 연령, 거동 부자유자, 정도 등을 식별

승객의 행동의 시퀀스를 LSTM 모델을 통해 버스 승차 예정, 버스 승차 중, 버스 하차 중, 으로 분류하여

버스의 승하차 승객이 있을 경우 버스가 출발하지 않도록 유도한다.



승객의 승차 중인 경우 Yolo V4를 이용 마스크 착용 유무를 검사 한다.

References

Open Pose

[GitHub - CMU-Perceptual-Computing-Lab/openpose: OpenPose: Real-time multi-person keypoint detection library for body, face, hands, and foot estimation](#)

LSTM

[43905.pdf \(googleusercontent.com\)](#)

YOLO V4 Mask

[GitHub - BogdanMarghescu/Face-Mask-Detection-Using-YOLOv4: Because of the COVID-19 pandemic of 2020, more and more people are concerned with protecting themselves using masks, thus the need of software capable of monitoring whether the people are wearing masks or not. That is why I created a Python application using OpenCV \(with CUDA support\) based on the YOLOv4 algorithm, capable of monitoring the safety level of a space with video surveillance.](#)