

메타데이터 정보 (다중기입가능)	분야	데이터 유형 ¹⁾	구축 데이터량		원천데이터 형식 ²⁾		라벨링 형식 ³⁾		라벨링 유형 ⁴⁾		
	NIA 기입	텍스트	원천 데이터	1,152, 305개	신호 최적화 데이터	-	신호 최적화 데이터	.csv /.xml	신호 최적화 데이터	텍스트	
			라벨링 데이터	651,8 89개	영상 데이터	jpg	영상 데이터	.csv /.xml	영상 데이터	이미지	
	데이터 출처 ⁵⁾	데이터 구축년도	구축기관(총괄)		가공기관		검수기관				
	자체 수집	2022년	(주)유아이네트웍스		(주)유아이네트웍스 (주)테스트웍스		(주)유아이네트웍스 한국전자기술연구원				
	데이터 문의처	기관명	문의담당자명		전화번호 (유선전화번호기입)		메일주소				
		(주)유아이네트웍스	전용현		02-586-0461		kipon7@naver.com				
	데이터 소개	AI를 활용하여 도심지 교통량, 교통밀도, 지체도 등 교통량 지표 데이터를 산출하고, 측 단위 또는 네트워크 단위의 교통흐름을 개선하기 위한 AI(심층강화학습) 기반의 신호 최적화 데이터 구축									
	주요키워드	AI 기반, 신호 최적화, 교통량, 인공지능 학습, 원천데이터, 라벨링데이터, 시뮬레이션									
카테고리 정의서		첨부의 카테고리 정의서 엑셀파일에 데이터카테고리 작성하여 제출(예시참고)									

1) 텍스트, 오디오, 이미지, 비디오,

2) txt, jpg,.....

3) json, csv,.....

4) 내용요약(텍스트), 번역(자연어), 질의응답(자연어), 바운딩박스(이미지/동영상), 키포인트(이미지/동영상), 세그멘테이션(이미지/동영상), 전자(음성)

5) 4대 언론기사, 자체 수집,.....

데이터셋명	국문영문	AI 기반 신호 최적화를 위한 데이터
		Data for Artificial Intelligence-based signal optimization
구축목적	항목	요구사항
	연구분야	<ul style="list-style-type: none"> 완전감응형, 반감응형, TOD 신호 최적화 알고리즘 개발 독립교차로, 축단위, 네트워크 단위 신호최적화 모델 연구 통행시간, 통행속도, 통행량 예측 모델 연구 AI를 활용한 교통신호 제어 모형 개발 방법론 연구
	산업분야	<ul style="list-style-type: none"> 스마트시티 관련 교통 시스템 개발 및 성능개선 AI 기반 감응형 신호제어기 및 관련 서비스 개발 엣지 시스템을 활용한 네트워크 단위 신호제어시스템 개발
활용서비스	<p>○ 다양한 데이터셋의 구성을 통한 신호최적화 연구 활성화</p> <ul style="list-style-type: none"> 교통신호에 관한 연구는 지속적으로 수행되었으며, 앞서, 연구의 배경에도 언급하였듯이 원활한 데이터 수집 및 보급이 이루어지지 않는 한계점을 내포하고 있었음 스마트교차로 보급 및 AI 학습 능력 향상으로 인하여 차량 식별에 대한 정확성이 향상되고 이로 인하여 보다 신호최적화를 위한 데이터셋 보급이 가능해짐 따라서 공간적, 시간적 범위의 데이터셋의 구성이 가능해져 신호최적화에 대한 다양한 연구가 진행될 수 있을 것으로 예상됨 <p>○ 신호 최적화 연구에 변화 및 고도화</p> <ul style="list-style-type: none"> 기존의 회전교통량 기반의 신호최적화 연구가 대부분인 반면 본 과제를 통하여 제공되는 다양한 형태의 데이터셋 (회전교통량 기반, 개별 차량 기반, 기종점 통행량 기반)을 통하여 단일 다중, 축별로의 최적화가 아닌 도시 전체를 포괄할 수 있는 연구로 고도화가 가능 <p>○ 교차로의 통행시간, 통행량, 통행속도의 예측이 가능하여 네비게이션 기초자료로 활용 가능</p> <ul style="list-style-type: none"> 본 사업을 통하여 구축된 데이터셋을 통하여 차량이 이동경로 및 기종점에 대한 정보, 5분 후에 도시 전체의 통행량, 통행속도 등의 예측 및 정체예상 구간도 가능해져 네비게이션의 기초자료로 활용이 가능함 교차로 단위의 예측이 가능함에 따라 예측의 정확성 신뢰성을 확보할 수 있으며 궁극적으로 이용자의 통행시간 단축에 기여할 수 있다고 판단됨 <p>○ 한국형 데이터셋을 통한 표준화 및 적용성 확대</p> <ul style="list-style-type: none"> 교통부문의 AI 학습 데이터셋은 대부분 국외에서 수집된 자료로 국내 실정에 맞는 자료에 구득이 어려움 신호최적화를 위한 시뮬레이션 대부분 국외 상용프로그램 및 오픈소스를 제공해 국내 데이터와 다소 차이가 발생하는 한계점을 가지고 있었음 본 사업에서 구축된 데이터셋의 프로세스 및 분석, 가공 도구는 오픈소스 기반으로 제작되어 향후 사업 후에도 데이터셋 이용자가 다양한 분석에 활용될 수 있는 기반이 되고, 한국형 신호 최적화 데이터셋의 표준체계 정립도 가능할 것으로 예상됨 	
소개	<p>○ 코로나에 의한 사회적 거리두기는 통행행태에 많은 변화를 야기시켰으며, 특히 승용차 통행량의 증가로 인해 교통정체를 일부 가중시키는 변화를 촉진하였다. 이에 도심부의 교통흐름에 영향을 주는 교통신호에 대한 부분이 강조되며, 신호최적화를 위한 AI 데이터는 반드시 필요한 자원</p> <p>○ 전세계적으로 AI 학습에 대한 관심이 높으며 이에 따른 핵심자원인 데이터 구축은 반드시 필요하며, 혁신성장의 시작일 뿐만 아니라 AI 선도국인 대한민국의 초석</p> <p>○ 따라서 아래와 같은 목적으로 데이터 구축</p> <ul style="list-style-type: none"> 연구분야는 4가지 측면에서 데이터 구축의 목적이 있으며, 4가지는 ① 완전 감응형, 반감응형, TOD 신호 최적화 알고리즘 개발, ② 독립교차로, 축단위, 네트워크 단위 신호 최적화 모델 연구, ③ 통행시간, 통행속도, 통행량 예측모델 연구, ④ AI를 활용한 교통신호 제어 모형 개발 방법론 연구 	

1차 경로

2차 경로

3차 경로

4차 경로

5차 경로

파일 포맷

제출 수량

네트워크 데이터

-

-

-

.xml

7개

부천시

교통량데이터
(개별차량
교통량기반)

교차로별 구분

일자별 구분

-

.csv

34,080개

교통량데이터
(기종점
교통량기반)

교차로별 구분

일자별 구분

-

.csv

34,080개

교통량데이터
(회전
교통량기반)

교차로별 구분

일자별 구분

-

.xml

133,559개

신호데이터

교차로별 구분

csv

일자별 구분

.csv

34,080개

xml

일자별 구분

.xml

34,080개

지표데이터

교차로별 구분

일자별 구분

-

.csv

34,080개

네트워크
데이터

-

-

-

.xml

5개

총 합계

648,049개

* 교차로 별 구분은 “데이터셋 구성-1. 데이터구성 및 구분정보-(1) 교차로 별 구분자” 참고

** 일자 별 구분은 “데이터셋 구성-1. 데이터구성 및 구분정보-(2) 날짜 별 구분자” 참고

(2) 영상 데이터

1차 경로

2차 경로

파일 포맷

구축 수량

csv

(방향별_날짜·시간별[5분단위])

.csv

1,920개

xml

(방향별_날짜·시간별[5분단위])

.xml

1,920개

총 수량

3,840개

1. 데이터 구성 및 구분정보

1) 신호최적화 데이터

경로 구분 정보

구분자 정보

1차 경로

지역 구분

안양시/부천시

2차 경로

데이터 종류별 구분

교통량데이터(개별차량기반교통량),
교통량데이터(기종점기반교통량),
교통량데이터(회전교통량),
신호데이터, 지표데이터, 네트워크데이터

3차 경로

교차로별 구분

교차로별 구분자 : 하단 첨부

4차 경로

날짜별 구분

날짜별 구분자 : 하단 첨부

(1) 교차로 별 구분자

경로 구분 정보

구분자 정보(네트워크 교차로 ID)

3차
경로

교차로별
구분

안양시

단일교차로

3 / 13 / 14 / 15 / 16 / 17 / 18 / 19 / 20

다중교차로

13_14 / 15_16 / 13_19_20 / 14_15_16_17 /
13_14_15_16_17_18_3

부천시

단일교차로

26 / 29 / 104 / 105 / 106 / 231 / 233 / 380 / 382

다중교차로

26_233_231 / 106_26_29 / 106_105_104 / all

(2) 날짜 별 구분자

경로 구분 정보

구분자 정보

4차
경로

날짜별
구분

안양시

단일
교차로
(30개)

20220323 / 20220324 / 20220326 / 20220327 / 20220328 /
20220329 / 20220330 / 20220424 / 20220425 / 20220426 /
20220427 / 20220428 / 20220429 / 20220430 / 20220506 /
20220507 / 20220508 / 20220509 / 20220510 / 20220705 /
20220706 / 20220707 / 20220708 / 20220709 / 20220716 /
20220717 / 20220718 / 20220719 / 20220720 / 20220721

다중
교차로
(18개)

20220323 / 20220324 / 20220326 / 20220327 / 20220328 /
20220329 / 20220330 / 20220424 / 20220425 / 20220426 /
20220427 / 20220428 / 20220429 / 20220430 / 20220506 /
20220507 / 20220508 / 20220509

부천시

단일

20220810 / 20220811 / 20220812 / 20220813 / 20220814 /

데이터셋 구성

경로 구분 정보		구분자 정보	
		교차로 (31개)	20220815 / 20220819 / 20220820 / 20220821 / 20220822 / 20220823 / 20220906 / 20220907 / 20220909 / 20220910 / 20220911 / 20220912 / 20220913 / 20220914 / 20220915 / 20220916 / 20220926 / 20220928 / 20220930 / 20221001 / 20221002 / 20221026 / 20221027 / 20221028 / 20221030 / 20221031
		다중 교차로 (19개)	20220810 / 20220811 / 20220812 / 20220813 / 20220814 / 20220819 / 20220820 / 20220821 / 20220822 / 20220823 / 20220906 / 20220907 / 20220909 / 20220910 / 20220911 / 20220926 / 20220928 / 20221028 / 20221030

2) 영상 데이터

경로 구분 정보		구분자 정보
1차 경로	데이터 형식별 구분	csv/xml
2차 경로	방향별_날짜·시간별 구분	(5분단위 구분)

2. 파일명 구성 정보

1) 신호최적화

예시	세부 구성 설명
40003120220202000000b.csv	교차로(묶음)(1)+교차로(개별)(4)+지역구분(1)+연도(4)+날짜(4)+시간(6)+차량타입(2).csv
40001120220202000000total.csv	교차로(묶음)(1)+출교차로(개별)(4)+지역구분(1)+연도(4)+날짜(4)+시간(6)+타입(5)

- 교차로(묶음) : 가장 앞자리 4는 1~18 교차로중 4개의 묶은 개수를 의미함
- 교차로(개별) : 0001은 각각의 개별 교차로를 의미함
- 지역구분 : 안양 1, 부평 2의 지역 코드를 의미함
- 연도 : 이 데이터는 몇 년도에 추출이 되어있는지를 말함
- 날짜 : 데이터 추출 날짜 기입
- 시간 : 이 데이터는 00(시)00(분)00(초)에 End 타입 기준으로 추출을 의미
- 차량 타입 : p(승용차), b(버스), st(소형트럭), lt(대형트럭), s(특수차), m(오토바이)
- 타입 : tra(궤적 trajectory), total(지표)

2) 영상데이터

예시	세부 구성 설명
EW04_20220726071500.xml	방향(2)+교차로번호(2)+연도(4)+날짜(4)+시간(6).xml
EW04_20220726071500.xml	방향(2)+교차로번호(2)+연도(4)+날짜(4)+시간(6).csv

3. 라벨 데이터 파일 내부 구성요소

1) 교통량데이터(기종점기반/개별차량기반)

구분	속성명	타입	필수여부	설명	유효값
1	unix_time	number	Y	이동수집시각	-
2	vhcl_id	string	Y	개별차량OD	-
3	stdr_ymd	number	Y	수집시각 (연월일)	-
4	stdr_hm	number	Y	수집시각(시분)	-
5	stdr_ss	number	Y	수집시각(초)	-
6	lon	number	Y	경도	*하단 참고
7	lat	number	Y	위도	*하단 참고
8	alt	string	Y	고도	-
9	spd	string	N	속도(m/s)	-
10	allowed_spd	string	N	비혼잡속도 (m/s)	-

구분	속성명	타입	필수여부	설명	유효값
11	agl	string	N	진행방향	-
12	slp	string	N	경사	-
13	vhcl_typ	number	Y	차량종류	*하단 참고
14	from_inter_id	string	N	이전통과교차로	-
15	to_inter_id	string	N	목적교차로	-
16	dist2to_inter	string	N	목적교차로까지의 거리	-
17	turn_typ2to_inter	string	N	회전방향	-
18	edge_grp_id	string	N	엣지그룹ID	-
19	edge_id	string	N	엣지ID	-
20	lane_id	string	N	차선ID	-
21	seg_id	string	N	세그먼트ID	-
22	nxt_inter_id	string	N	다음 교차로 ID	-
23	dist2nxt_inter	string	N	다음 교차로까지의 거리	-
24	turn_typ2nxt_inter	string	N	다음교차로 회전방향	-
25	nxt_inter_sig_st	string	N	다음 교차로 현재신호	-
26	nxt_inter_phs_no	string	N	다음 교차로 현시신호	-
27	que_all	string	N	전체 대기행렬관측여부	-
28	que_200	string	N	대기행렬관측여부 (~200m)	-
29	que_200_500	string	N	대기행렬관측여부 (200m~500m)	-
30	que_500	string	N	대기행렬관측여부 (~500m)	-
31	tl	string	N	지체시간	-

* 위도/경도값 범위

지역	구분	lat(위도)**	lon(경도)**
안양시	최소값	37.362	126.943
	최대값	37.401	126.999
부천시	최소값	37.494	126.741
	최대값	37.518	126.780

** 좌표계 : WGS84

* 차량 종류입력 범위

구분구적	차량종류 (carmodel type)	1	2	3	4	5	6
	차종	승용차	버스	트럭	대형트럭	특수차량	오토바이

2) 교통량데이터(회전교통량)

구분	속성명	타입	필수여부	설명	유효값
1	data	object	Y	전체 데이터 묶음	-
2	interval	array	Y	시간 간격	-
3	{}	object	Y	문서	-
4	end	string	Y	종료시간	-
5	id	string	Y	상태값	-
6	begin	string	Y	시작시간	-
7	edgeRelation	object	Y	기종점 노드	-
8	count	number	Y	통과 차량대수(대)	-
9	from	string	Y	시점 노드 ID	-
10	to	string	Y	종점 노드 ID	-

3) 신호데이터(csv)

구분	속성명	타입	필수여부	설명	유효값
1	interid	string	Y	교차로 ID	*하단 참고
2	aringstarttime	number	Y	신호시작시간 (날짜/시분초)	-
3	signalstate	string	Y	SUMO 신호입력값	-
4	unix_time	number	Y	유닉스 타임	-
5	phasepattern	string	N	원본데이터 형식패턴	-

* 교차로 ID 입력 범위

지역명	구문규칙	교차로명	데이터	좌표* (위도, 경도)
	교차로 (개별)		교차로ID (interid)	
안양시	3	별말오거리	cluster_217833_cluster_216283_217694_cluster_215007_216164_216616_217832	37.396608, 126.976359
	13	범계사거리	cluster_216196_216197_217930_217931	37.389058, 126.948116
	14	범계역사거리	215147	37.391117, 126.950987
	15	시청사거리	215173	37.392524, 126.955169
	16	관평사거리	215174	37.394078, 126.959660
	17	평촌역사거리	214964	37.395564, 126.964019
	18	열병합사거리	215024	37.396009, 126.968882
	19	방축사거리	215092	37.382793, 126.952703
부천시	20	신기사거리	cluster_216778_217465	37.378211, 126.955200
	26	석천사거리	204820	37.504764, 126.763169
	29	문예사거리	204822	37.503955, 126.769198
	104	화목사거리	201305	37.499997, 126.756811
	105	중동전화국사거리	204920	37.502754, 126.757570
	106	송내대로사거리	cluster_204796_205202	37.505410, 126.758310
	231	부명사거리	201301	37.499322, 126.762051
	233	포도마을사거리	201292	37.502079, 126.762636
	380	중앙공원사거리	201306	37.498523, 126.768089
	382	구터미널사거리	201293	37.501292, 126.768665
	103**	넙말사거리	201330	37.497218, 126.756068
	379**	복사골아파트사거리	201329	37.495748, 126.767516
	414**	영안아파트사거리	201300	37.496535, 126.761483

* 좌표계 : WGS84

** 신호최적화데이터 수집대상 지점에 포함되지않으나, 교차로 연계성을 위해 추가 설정

4) 신호데이터(xml)

구분	속성명	타입	필수여부	설명	유효값
1	tlLogics	object	Y	전체 신호 데이터	-
2	tlLogic	array	Y	교차로별 신호 데이터	-
3	{ }	object	Y	문서	-
4	phase	array	Y	신호 패턴	-
5	{ }	object	Y	문서	-
6	duration	number	Y	해당 신호가 유지되는 시간(초)	-
7	state	string	Y	신호 상태	-
8	offset	string	Y	연동신호여부	-
9	id	string	Y	교차로 ID	-
10	type	string	Y	형식	-
11	programID	number	N	교차로 순서	-

5) 지표 데이터

구분	속성명	타입	필수여부	설명	유효값
1	edge_grp_id	string	Y	엣지그룹 ID	-
2	avg_spd	number	Y	평균속도(m/s)	-
3	total_time_loss	string	Y	지체시간(초)	-
4	avg_time_loss	string	Y	차량당 평균 지체시간(초)	-
5	avg_travle_time	number	Y	평균 통행시간(초)	-

6) 영상 데이터(csv)

구분	속성명	타입	필수여부	설명	유효값	비고
1	photo_file	String	Y	영상파일 이름 (교차로+시간대)	교차로[4자릿수]_YYYYMMDDhhmmss	
2	photo_id	String	Y	기준선 통과 이미지 파일의 식별값	교차로[4자릿수]_YYYYMMDDhhmmssms_nnnnn.jpg	
3	signal_info.movement	String	Y	통과시간의 신호정보	[s,t,l,y,t,l,y]	
4	departure_time	String	Y	영상기준 차량의 교차로 통과 시간	mm:ss:ms	분:초:밀리초
5	car_type	String	Y	교차로 통과 차량의 종류	[car, bus, small_truck, large_truck, bike, unknown]	
6	lane	String	Y	차량이 지나가는 차로	[1 ~ 10]	
7	car_info.movement	String	Y	차량의 진행 방향	[r,t,l,u]	r: 우회전, t: 직진 l: 좌회전, u: 유훈
8	tracking ID	String	N	추적차량 번호	[0 ~ 10]	

7) 영상 데이터(xml)

구분	속성명	타입	필수여부	설명	유효값	비고
1	annotations	object	Y	어노테이션 정보	-	
2	image	array	Y	이미지 정보	-	
3	{}	object	Y	문서	-	
4	id	string	Y	현재 xml파일 내의 이미지 파일별 고유 식별 숫자(0번부터 시작, 마지막 식별자는 size-1)	-	
5	name	string	Y	이미지 파일명	-	
6	width	string	Y	이미지 파일 가로 해상도	"1920"	
7	height	string	Y	이미지 파일 세로 해상도	"1080"	
8	box	array	N	Bounding-box 객체 정보	-	
9	{}	object	N	문서	-	
10	label	string	N	바운딩박스 가공	"car","bus","sm"	

구분	속성명	타입	필수여부	설명	유효값	비고
				객체의 class 정보. 객체 차종은 총 6종임	all_truck", "large_truck", "bike", "unknown"	
11	occluded	string	N	의미없는 항목 (가려진 차량 판별)	—	
12	z_order	string	N	의미없는 항목 (가공틀에서 산출)	—	
13	xtl	string	N	바운딩박스 가공 객체의 이미지 픽셀을 기준으로 좌측 상단 x좌표	—	
14	ytl	string	N	바운딩박스 가공 객체의 이미지 픽셀을 기준으로 좌측 상단 y좌표	—	
15	xbr	string	N	바운딩박스 가공 객체의 이미지 픽셀을 기준으로 우측 하단 x좌표	—	
16	ybr	string	N	바운딩박스 가공 객체의 이미지 픽셀을 기준으로 우측 하단 y좌표	—	
17	attribute	string	N	객체의 속성값	0,1,2,3,4,5,6,7,8,9,10,"r","l","t","u"	
18	attribute.name	string	N	attribute태그의 이름 속성	"CID","direction","lane"	
19	meta	object	Y	meta 정보	—	
20	task	object	Y	task 정보	—	
21	owner	object	N	의미없는 항목 (가공업체 정보)	—	
22	email	string	N	의미없는 항목 (가공업체 email정보)	—	
23	username	string	N	의미없는 항목 (가공업체 책임자 정보)	—	
24	flipped	string	N	의미없는 항목 (뒤집힌 이미지 판별)	—	
25	created	string	N	의미없는 항목 (작업생성일)	—	
26	project	string	N	의미없는 항목 (과제명)	—	
27	start_frame	string	N	의미없는 항목(시작 프레임)	—	
28	frame_filter	string	N	의미없는 항목 (가공틀상에서 작업화면 페이지 전환 (ex. Step=1 : 1장씩, step=2 : 2장씩))	—	
29	labels	object	Y	라벨링 대상 task 정보	—	
30	label	array	Y	객체 class 정의 부분	—	
31	{ }	object	Y	문서	—	
32	name	string	Y	객체 차종은 총 6종임	"car","bus","small_truck","large_truck","bike","unknown"	

	구분	속성명	타입	필수여부	설명	유효값	비고
	33	attributes	object	N	라벨링 객체 속성	-	
	34	segments	object	N	의미없는 항목 (가공툴상에서 라벨링 종류(polygon), 가공툴에 Default로 나오는 값)	-	
	35	segment	object	N	의미없는 항목 (가공툴상에서 Default로 나오는 값)	-	
	36	stop	string	N	의미없는 항목 (가공툴상에서 Default로 나오는 값)	-	
	37	start	string	N	의미없는 항목 (가공툴상에서 Default로 나오는 값)	-	
	38	id	string	N	라벨링 객체 ID	-	
	39	url	string	N	의미없는 항목 (가공툴상에서 Default로 나오는 값)	-	
	40	mode	string	N	라벨링 종류	-	
	41	overlap	string	N	의미없는 항목(이미지 중복 판별)	-	
	42	size	string	Y	어노테이션 정보를 포함하고 있는 가공 이미지의 개수	-	
	43	name	string	Y	가공 도구 상에서의 작업단위(Task)명	-	
	44	bugtracker	string	N	의미없는 항목 (버그추적)	-	
	45	id	string	N	task ID	-	
	46	stop_frame	string	N	의미없는 항목(종료 프레임)	-	
	47	updated	string	N	의미없는 항목(최종 수정일)	-	
	48	dumped	string	N	의미없는 항목 (작업 완료 제출일)	-	
	49	version	string	N	의미없는 항목(xml 버전)	-	
데이터셋 구축 수행기관 담당자	주관기관	기관명	책임자명	전화번호 (유선전화번호기입)		메일주소	담당업무
		(주)유아이네트웍스	전용현	02-586-0461		kipon7@naver.com	과제 총괄
	참여기관	기관명	담당업무	기관명	담당업무		
		(주)라온로드	데이터 수집				
		(주)테스트웍스	데이터 정제,가공				
		한국전자기술연구원	데이터 검사				