

메타데이터 정보 (다중기입가능)	분야	데이터 유형 ¹⁾	구축 데이터량	원천데이터 형식 ²⁾	라벨링 형식 ³⁾	라벨링 유형 ⁴⁾
	<i>NIA 기입</i>	이미지	431,120건	jpg	json	폴리곤 세그멘테이션
	데이터 출처 ⁵⁾	데이터 구축년도	구축기관(총괄)	가공기관	검수기관	
	자체 수집	<i>2021년</i>	(주)씨유박스	(주)딥핑소스	(주)딥핑소스	
	데이터 문의처	기관명	문의담당자명	전화번호 (유선전화번호기입)	메일주소	
		(주)씨유박스	이기훈	02-6277-7841	ghlee@cubox.aero	
	데이터 소개	장애물(예, 철조망, 유리창 등)에 가려지거나 후방에 존재하는 객체의 인식을 향상 및 텍스처 바이어스(Texture-bias)를 줄이기 위한 장애물에 가려진 데이터셋 구축				
	주요키워드	장애물에 가려진 물체 인식, 장애물에 가려진 물체 형상 추정, occluded object, 객체 인식, object detection, semantic segmentation				
카테고리 정의서		카테고리 정의서 엑셀파일에 데이터카테고리 작성하여 제출				

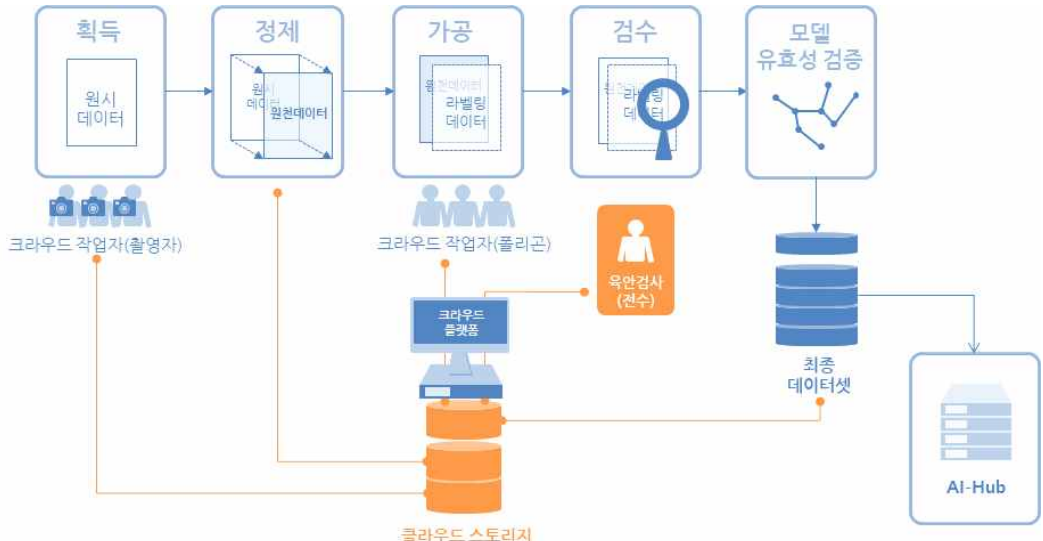
1) 텍스트, 오디오, 이미지, 비디오,

2) txt, jpg,.....

3) json, csv,.....

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

5) 4대 언론기사, 자체 수집,.....

데이터셋명	국문영문	장애물에 가려진 물체 형상 추정 데이터																																	
		Occluded Object Dataset																																	
구축목적	자율주행, 로봇틱스 등에서 사용이 확대되고 있는 객체 인식 기술은 객체가 장애물에 가려질 경우 텍스처 바이어스로 인한 오인식 등 인식 정확도가 떨어짐. 객체 인식 성능을 높이기 위해서 장애물 (예, 철조망, 유리창 등)에 가려진 객체로 구성된 데이터셋을 확보할 필요가 있음																																		
활용서비스	자율 주행 중 표지판 인식 기술 성능 개선, 로봇틱스에 적용하여 로봇의 객체 인식 성능 정확도 개선, 스마트 CCTV를 이용한 보안 시스템 등 장애물로 인한 텍스처 바이어스가 발생할 가능성이 높은 인식 상황을 처리하는 AI 응용서비스 일체																																		
소개	□ 데이터 구축 ○ (수집) ImageNet 기반으로 국내에서의 수집 적합성 확인 후 200종 객체 카테고리 선정 후 각 카테고리별로 100개의 물체 인스턴스 촬영. 각 인스턴스당 5종류의 장애물(장애물 없음 None, 반투명 Semi-transparent, 촘촘한 망 Wire-dense, 중간 간격 망 Wire-medium, 느슨한 망 Wire-Loose)를 변경하면서 4 방향에서 수집 ○ (정제) 원시데이터 검수 기준에 따라 촬영 이미지 품질 검수(반려된 이미지는 재촬영 후 재검수 진행) 및 얼굴, 전화번호, 자동차 번호판 등 개인정보 또는 민감정보에 대한 비식별화 작업 진행 ○ (가공) 객체의 가시 영역 외곽선을 따라 점을 찍어 다각형 모양으로 그리는 폴리곤 라벨링 진행 ○ (검수) 폴리곤 라벨링에 대해 육안 검수 후 통과 또는 반려함. 반려된 작업물은 재작업 후 다시 검수 과정을 거침 																																		
	1. 데이터 구축 규모 <table><tr><th>구분</th><th>상세</th><th>형식</th><th>총 수량</th></tr><tr><td>원천 이미지</td><td rowspan="3">200종 객체 카테고리 및 객체당 2,000장 이상의 장애물에 가려진/가려지지 않은 객체를 촬영한 이미지 및 라벨링 데이터</td><td>JPG</td><td>431,120</td></tr><tr><td>폴리곤 좌표값</td><td>JSON</td><td>431,120</td></tr><tr><td>촬영 메타 정보</td><td>JSON</td><td>431,120</td></tr></table> <table><tr><td>No.</td><td>클래스</td><td>No.</td><td>클래스</td><td>No.</td><td>클래스</td><td>No.</td><td>클래스</td><td>No.</td><td>클래스</td></tr><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>		구분	상세	형식	총 수량	원천 이미지	200종 객체 카테고리 및 객체당 2,000장 이상의 장애물에 가려진/가려지지 않은 객체를 촬영한 이미지 및 라벨링 데이터	JPG	431,120	폴리곤 좌표값	JSON	431,120	촬영 메타 정보	JSON	431,120	No.	클래스	No.	클래스	No.	클래스	No.	클래스	No.	클래스									
구분	상세	형식	총 수량																																
원천 이미지	200종 객체 카테고리 및 객체당 2,000장 이상의 장애물에 가려진/가려지지 않은 객체를 촬영한 이미지 및 라벨링 데이터	JPG	431,120																																
폴리곤 좌표값		JSON	431,120																																
촬영 메타 정보		JSON	431,120																																
No.	클래스	No.	클래스	No.	클래스	No.	클래스	No.	클래스																										

No.	클래스	No.	클래스	No.	클래스	No.	클래스	No.	클래스
1	AirFryer	41	DigitalDoorlock	81	Iron	121	Perfume	161	SportsBall
2	Airtight Container	42	DiningTable	82	Jug	122	PillBottle	162	Sprayer
3	Apple	43	DishDrainer/ DryingRack	83	Ketchup	123	Pillow	163	Stapler
4	Backpack	44	Dishsoap	84	Kleenex	124	Pizza	164	Stove
5	Banana	45	DisposableMask	85	Knife	125	Plate	165	Suitcase
6	BandAid	46	Doll	86	Ladle	126	Plunger	166	Sunglasses
7	BathMat	47	DoorHandle	87	Laptop Computer	127	Pole	167	Sunscreen
8	BeanCurd	48	Dumbbell	88	Laundry Detergent	128	PopBottle	168	Supplementary Battery
9	BeerBottle	49	ElectricFan	89	Lemon	129	Popsicle	169	SweatShirt
10	BellPepper	50	ElectricKettle	90	Loaf	130	Pot	170	TabletPC
11	Bicycle	51	Electric RiceCooker	91	Loafer	131	Potato	171	Tape
12	Blender	52	ElectricSwitch	92	Lotion	132	Printer	172	Teacup
13	Book	53	Envelope	93	Macaron	133	Projector	173	Teapot
14	Broccoli	54	EspressoMaker	94	Magnifier	134	Puncher	174	Television
15	Broom	55	Eyepatch	95	Mailbag	135	Ramen	175	Thermometer
16	Brush	56	FacePowder	96	Mailbox	136	Razor	176	Tie
17	Bucket	57	Fan	97	Manhole Cover	137	Remote Control	177	Toaster
18	Cabbage	58	Flashlight	98	MicroPhone	138	RiceScoop	178	Toilet
19	Calculator	59	FlowerPot	99	Microwave	139	RiceSoup Bowl	179	ToiletTissue
20	Camera	60	FluorescentLamp	100	Milk	140	RubberEraser	180	Tomato
21	Can	61	Fork	101	Mitten	141	Ruler	181	Toolkit
22	Cardigan	62	Frame	102	MixingBowl	142	RunningShoe	182	Toothbrush
23	Carrot	63	Frypan	103	Mobilephone	143	SaltShaker	183	Toothpaste
24	CarSideMirror	64	FuseBox	104	Monitor	144	Sandal	184	Towel
25	Carwheel	65	HairBand	105	Mop	145	Scale	185	Tray
26	Cereal	66	HairBrushComb	106	Mosquitocide	146	Scissors	186	T-shirt
27	Chair	67	HairDryer	107	MotorScooter	147	ScouringPad	187	Umbrella
28	Cheese	68	HairSpray	108	Muffler	148	Screwdriver	188	USB
29	ChoppingBoard	69	Hamburger	109	Mug	149	SesameOil	189	Vase
30	Cigarettes	70	Hammer	110	Mushroom	150	Sewer Drainage	190	Wallet
31	Clock	71	Handbag	111	One-piece	151	Shampoo	191	Washbasin
32	Coat	72	Hanger	112	Orange	152	Shopping Cart	192	Wastebin
33	Computer Keyboard	73	Hankerchief	113	Pants	153	ShowerHead	193	Watch
34	ComputerMouse	74	Hat	114	PaperBag	154	Skirt	194	WaterBottle

No.	클래스	No.	클래스	No.	클래스	No.	클래스	No.	클래스
35	CookingOil	75	Headphone	115	Paper Notebook	155	Slippers	195	WaterPurifier
36	Corn	76	Hotdog	116	ParkBench	156	Soap	196	WetTissue
37	Cucumber	77	Hourglass	117	Pen	157	Sock	197	Whisk
38	Cucurbita	78	HwatuCard	118	PencilCase	158	Spatula	198	WineBottle
39	DeskLamp	79	Icecream	119	Pencil Sharpener	159	Speaker	199	Wineglass
40	Desktop Computer	80	InstantRice	120	PennyBank	160	Spoon	200	Wok

<객체 200종 카테고리>

2. 데이터 분포

항목	조건구분	비율(%)
클래스별 이미지 분포	객체 200종	0.46~0.83% (2,000~3,570)
	총계	100.0% (총 431,120건)
장애물 환경별 이미지 분포	장애물 없음	20.0%
	장애물 - 반투명	20.0%
	장애물 - 촘촘한 망	20.0%
	장애물 - 중간 간격 망	20.0%
	장애물 - 느슨한 망	20.0%
	총계	100.0% (총 431,120건)

1. 대표도면

분류	원본이미지	폴리곤 좌표값 JSON	메타정보 JSON
예시		<pre>{ "additionalInformation": { "id": "AirFryer218987_AirFryer_321_81_None_Seg", "imageName": "218987_AirFryer_321_81_None" }, "figures": [{ "id": 1, "type": "shape", "faces": "polygon", "shape": { "vertices": [{ "x": 140, "y": 140, "order": 1 }, { "x": 140, "y": 160, "order": 2 }, { "x": 160, "y": 160, "order": 3 }, { "x": 160, "y": 140, "order": 4 }, { "x": 140, "y": 140, "order": 5 }] }, "occluder": "None", "coordinates": { "x": 140, "y": 140 } }] }</pre>	<pre>{ "class": "AirFryer", "occluder": "None" }</pre>
구축량	431,120 SET (원본이미지/폴리곤 좌표값 JSON/메타정보 JSON 각 1:1:1 구축)		
포맷	JPG	JSON	JSON
폴더 구조	- 상기 데이터 3종은 SET로 구축되어 동일한 폴더구조로 구성 - 폴더구조 : 총 200개 클래스의 폴더로 구분되어있으며, 파일명 내 클래스명 및 장애물정보 구분자를 포함		

2. 라벨링데이터 구성

폴리곤 좌표값 JSON 항목

구분	항목명		타입	필수 여부	설명	범위	비고
1	additional Information		object				
	1.1	s3Url	string		원본이미지 저장경로		
	1.2	imageName	string	Y	원본이미지 파일명		
2	figures		array				
	2.1	id	number	Y	각각의 폴리곤에 대한 id값	1~100	
	2.2	type	string		형태 관련 좌표값임을 명시	shape	
	2.3	label	string	Y	레이블 관련 정보 명시	polygon	
	2.4	shape		object			
		2.4.1	keyframe	boolean			
		2.4.2	type	string	레이블 관련 정보 명시	polygon	
		2.4.3	occluded	boolean			
		2.4.4	coordinates	array	폴리곤 좌표값을 포함하는 항목		
			L\$value\$	number	폴리곤 점(point) (x.y) 좌표값		
	2.5	attributes		array			
	2.6	displayIndex		number			
3	attributes		array				

메타정보 JSON 항목

구분	항목명	타입	필수여부	설명	범위	비고
1	class	String	Y	클래스명	200종	
2	occluder	String	Y	장애물 종류	5종	

3 라벨링데이터 실제예시

폴리곤 좌표값 JSON 항목

```

{
  "additionalInformation": {
    "s3Url": "AirFryer/210907_AirFryer_921_01_None.jpg",
    "imageName": "210907_AirFryer_921_01_None"
  },
  "figures": [
    {
      "id": 1,
      "type": "shape",
      "label": "polygon",
      "shape": {
        "keyframe": true,
        "type": "polygon",
        "occluded": false,
        "coordinates": [
          [
            1449.5135498046875,
            1334.6683349609375
          ],
          [
            1456.473876953125,
            1393.5633544921875
          ],
          [
            1465.5758056640625,
            1418.7275390625
          ],
          [
            1475.2132568359375,
            1429.435791015625
          ],
          [
            1486.456787109375,
            1431.0419921875
          ]
        ]
      }
    }
  ],
  "metaInfo": {
    "class": "AirFryer",
    "occluder": "None"
  }
}

```

■ 메타정보 JSON 항목

```

{
  "class": "AirFryer",
  "occluder": "None"
}

```

데이터셋 구축 수행기관 담당자	주관기관	기관명	책임자명	전화번호 (유선전화번호기입)	메일주소	담당업무
		(주)씨유박스	박준석	02-6277-7835	jspark@cubox.aero	데이터 설계/수집/ 품질관리/모델개발
	참여기관	기관명	담당업무	기관명	담당업무	
		(주)인터마인즈	데이터수집	(주)답핑소스	데이터가공/검수	