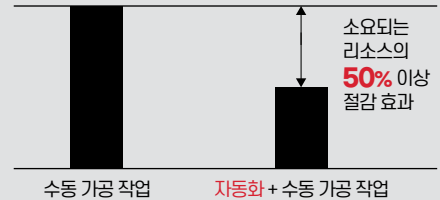


blackolive

인공지능의 **자양분** 고품질 학습 데이터를 위한 **블랙올리브**

인공지능 개발 및 성능 향상에 필요한 AI 학습용 데이터셋
블랙올리브로 더 빠르게, 더 편리하게, 더 정확하게
가공하고 관리할 수 있습니다.

By up to **50% +**



대규모

데이터 작업 생성, 데이터 업로드, 데이터 작업 할당까지

- 대규모 데이터 작업을 더 효율적으로 생성 진행 관리

자동화

최신 객체 검출 기술로 고도화된 라벨링 자동화 모델

- 다량의 자동 라벨링 후 더 세밀하게 작업 및 검수

고품질

자동화를 통한 단순 가공 리소스 절감으로 검수 과정 개선

- 데이터 가공 검수 시 소통 공간을 통한 협업 및 품질 향상

고도화

사용자를 배려한 작업 화면과 편리한 메뉴 구성

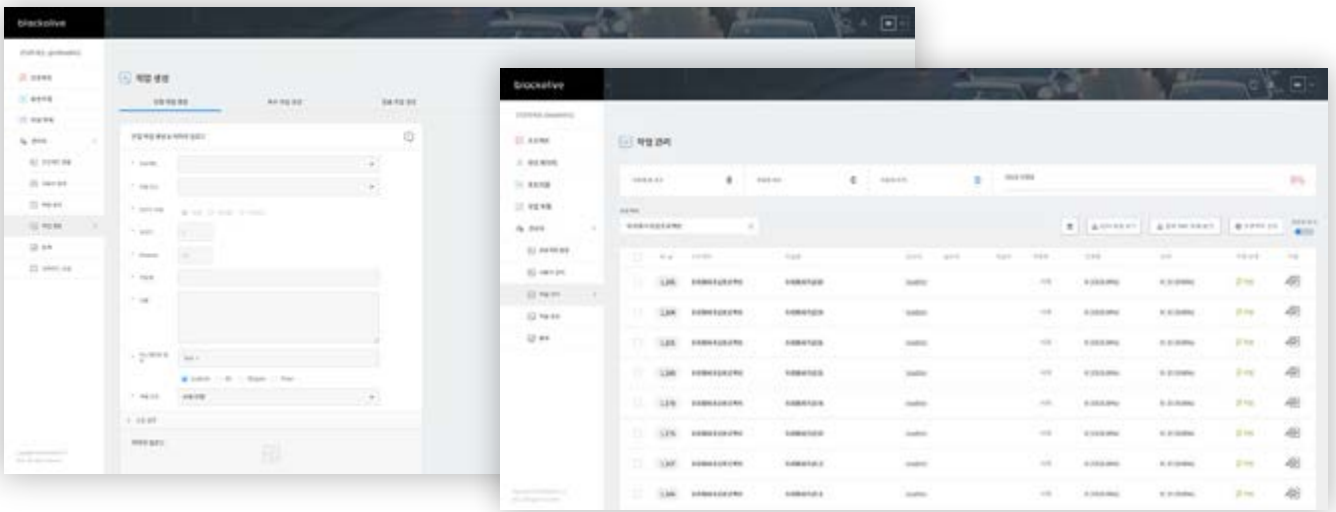
- 자동 라벨링 기능과 단축키를 통해 더 빠르게 작업



대규모

다량의 이미지 및 영상 등이 포함된 폴더를 클릭 또는 드래그 앤 드롭으로 업로드하여 손쉽게 작업 생성 (Azure Storage 사용으로 보안 유지)

- ▶ 간편한 프로젝트 생성
- ▶ 단일, 복수, 일괄 작업 중 프로젝트 규모에 따라 선택 가능
- ▶ 프로젝트 생성 시 세부 항목 설정 및 데이터셋 정보를 효율적 관리
xml, 작업 목록, 가공 타입별 통계, 라벨별 통계 등



자동화

자동화 모델을 통해 다량의 이미지/영상 데이터의 1차 가공 작업을 빠르게 수행

- ▶ 인도보행, 자율주행, 의료, 보안 등 7개 분야 20개의 카테고리 약 260여 개의 객체를 인식할 수 있는 자동화 모델 보유
- ▶ 사전에 학습되지 않은 객체도 Semi-auto 기능을 통해 자동 라벨링 후 수동으로 보정 작업 가능
- ▶ Semi-auto로 라벨링한 데이터셋을 객체 자동 인식 모델에 학습시키면 대규모 자동 라벨링 가능



[대상 선택]



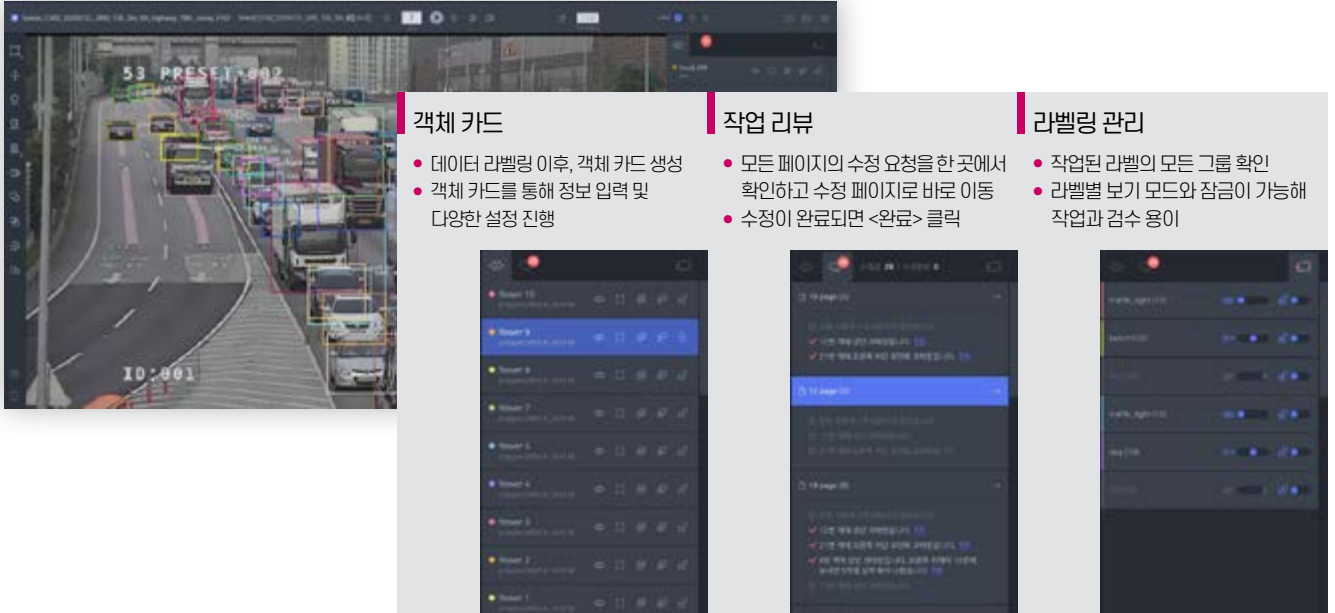
[자동으로 가공된 결과]

Semi-auto Polygon 기능 예시
대상 객체를 시각형 형태로 선택 시
자동으로 배경 분리 및 Polygon으로 가공

고품질

1차 자동 라벨링을 거쳐 수동 보정 및 검수에 더 집중하여 품질 확보

- ▶ 최신 객체 검출 기술과 자체 데이터셋으로 더 고도화된 자동 라벨링으로 1차 가공의 정확도 개선
- ▶ 작업 화면 내 라벨링 및 검수 리뷰를 위한 협업 공간 제공으로 빠르게 작업 진행



고도화

사용자를 배려한 편리한 라벨링 기능, 사용자 메뉴, 작업 화면을 통해 생산성 향상

- ▶ 트렌디한 UI와 사용자 경험을 최적화한 UX

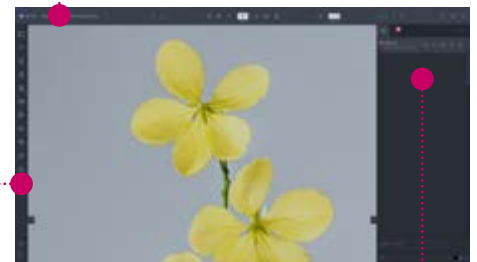
- 1 다크 모드(Dark Mode)로 몰입감 증대
- 2 단축키의 효율성 개선 및 사용성 증대
- 3 전문적으로 세분화된 관리 기능
- 4 쉽고 직관적인 아이콘과 기능

축색 조정 영역

데이터 라벨링 작업 시 객체의 생성과 작업을 위한 이미지 조정 아이콘 도구 구성

상단 컨트롤 영역

작업의 기본 정보, 페이지 이동, 단축키, 통계, 설정 팝업 제공



우측 가공 관리 영역

라벨링 객체 생성, 객체의 추가 정보 입력, 검수를 통한 수정 요청 → 객체별 상세한 라벨링 관리 기능

다양한 고급 설정 기능

- 우선 순위 라벨 배치
- 오픈포즈(OpenPose) 학습을 위한 데이터셋 구성
- 오버랩 사이즈와 세그먼트 사이즈 설정
- 이미지 화질 설정
- 동영상 업로드 초당 프레임 설정



관리자 통계 화면

- 해당 프로젝트에 대한 전체 작업 수, 전체 라벨 수, 전체 오류 수, 전체 프레임 수, 라벨 리뷰 등에 대한 요약 정보를 시각적인 그래프로 확인

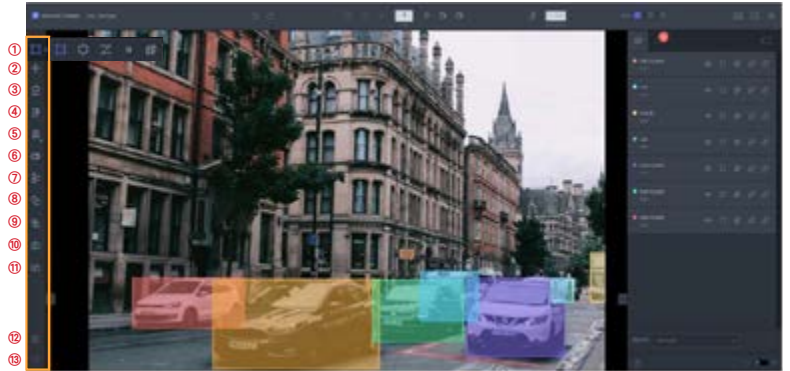


블랙올리브 데이터 가공 작업 화면

» 좌측 조정 영역

주요 메뉴

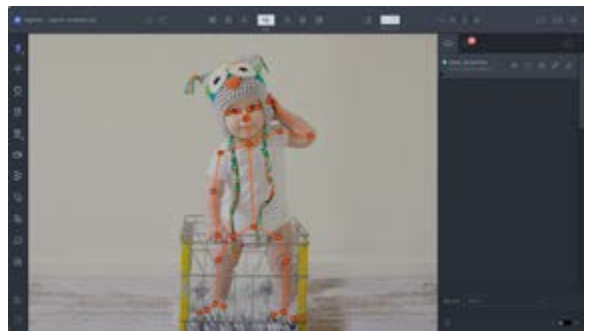
- ① 라벨링 선택 ⑧ 병합
- ② 객체/이미지 이동 ⑨ 그룹
- ③ 밝기 조절 ⑩ 오른쪽 회전
- ④ 불투명도 ⑪ 왼쪽 회전
- ⑤ 격자 보기 ⑫ 저장
- ⑥ 배치 순서 표시 ⑬ 나가기
- ⑦ 우선 순위 정렬



» 최신 라벨링 기능

종류	라벨링 기능	활용 사례
Box	 특정 객체를 직사각형 모양 안에 포함되도록 그리는 라벨링 방식	특정 객체의 사각형 영역 캡처가 필요한 경우 사용(가장 일반적으로 이용되는 기능)
Polygon	 특정 객체 테두리의 외곽선을 따라 점을 찍어 그리는 라벨링 방식	인공지능 모델링 방법에 따라 식별하고자 하는 객체 이외에 포함된 빈공간으로 생기는 오류를 줄이기 위해 사용
Polyline	 연결된 하나의 선을 그려 객체를 구분하는 라벨링 방식	인도, 차선 등을 구분하기 위해 주로 사용
Points	 특정 지점에 점을 생성하는 라벨링 방식 단, 점을 연속한 위치에 군집으로 생성해도 다각형은 형성할 수 없음	안면 인식을 통한 감정 분석과 같이 정밀하고 섬세한 작업이 요구되는 기술에 사용
Cuboid	 2D 작업의 한계가 있는 입체적인 객체를 정육면체로 생성하는 3D 라벨링 방식	자동차, 건물 등 입체적인 객체들의 3D 라벨링 기술에 사용
Body	 사람의 신체에 스켈레톤 형태로 객체를 생성하는 라벨링 방식	전체적인 모션 캡처나 이상행동 등 사람의 움직임을 검출하는 데 사용
Face	 사람 얼굴의 부위 별 객체를 생성하는 라벨링 방식	기계 학습 기반의 안면 응용기술 등에 활용하기 위해 얼굴의 특징점을 검출하는 데 사용
Hand	 사람의 손에 객체를 생성하는 라벨링 방식	수어와 같이 손의 움직임을 파악하기 위해 손의 마디를 검출하는 데 사용
Animal	 스켈레톤 형태로 동물의 객체를 생성하는 라벨링 방식	동물의 얼굴, 몸, 관절 등의 움직임을 검출할 때 사용

라벨링 기능 활용 예



기업 고객을 위한 블랙올리브 가격 정책

블랙올리브는 국내 최초 자율주행 시 데이터 가공 경험 노하우를 기반으로 대규모 데이터 구축의 효율성과 사용자 편리성을 고려해 자체 데이터셋으로 고도화시킨 시 데이터 가공 관리를 위한 최적의 솔루션입니다.

» blackolive는 기업 고객의 입장에서 인공지능 데이터 가공과 관리의 생산성을 높여주는 최적의 솔루션을 지원합니다.

온-프레미스 (On-premise)

- blackolive 가공 관리 솔루션 사용 비용
 - 연간 라이선스 비용 산정
- + 기술 지원비
- + 교육비

클라우드 (SaaS)

- blackolive 가공 관리 솔루션 사용 비용
 - 가공 데이터 객체 수, 워크로드, 스토리지 용량 과금 기준에 따른 비용 산정
- + 유지 관리비 (기술 지원)
- + 교육비

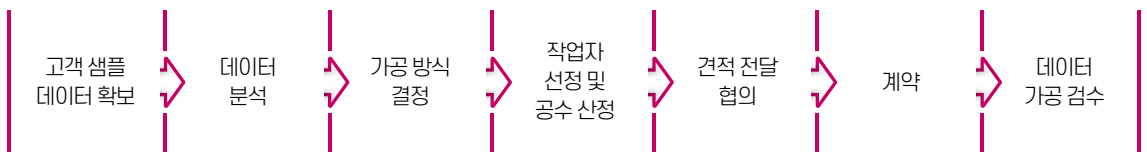
» 학습 데이터 가공 서비스

- 가공 데이터의 종류에 따라 작업 방식과 단위별로 과금이 적용됩니다.

데이터	작업 방식	단위
이미지	Box, Polygon, Polyline, Points, Cuboid, Body, Hands, Face	객체, 포인트
동영상	Box, Polygon, Polyline, Points, Cuboid, Body, Hands, Face	객체, 포인트
텍스트	의미(Context) 파악, 대화 판별, 질의 및 응답	문장
사운드	전사(받아쓰기), 문장 자르기, 녹음 하기, 질문 가부 확인	구간, 문장 길이, 질문

» 학습 데이터 가공 PoC (Proof of Concept) 프로세스

- 정확한 가공 비용 산정을 위해 고객의 샘플 데이터 검토 후, 특성을 분석하여 가공 방식과 작업자 선정 및 공수를 산정하여 견적을 제공합니다.



* 상세 비용은 별도 문의

블랙올리브 활용 사례

대규모 데이터셋 구축

» SideGuide

이동취약계층을 위한 대규모 데이터셋

- 2019 NIA '인도 보행 영상 시 데이터' 주관



구분	산출물		성과 기준	결과	
가공	데이터셋	객체 인식	Bounding Box	350,000장	352,810장
			Polygon	100,000장	100,712장
			Surface Masking	50,000장	51,318장
		깊이 인식	Depth Prediction	150,000세트	174,080세트

» Sign Language dataset

수어 인식 모델 개발을 위한 수어 영상 데이터셋

- 2020 NIA '수어 영상 시 데이터' 주관



과정	산출물	상세	수량	수행 결과
수집	수어 영상	직접 촬영 문장·단어	500,000클립	536,000 클립 데이터 구축 완료
		클라우드-소싱 지숫자·문자	21,000클립	
		가상 데이터 수어·문장	15,000클립	
가공	학습용 데이터셋	어노테이션 파일셋	536,000개	

» 교통 안전 데이터셋

교통 문제 해결을 위한 고속도로 CCTV 교통 데이터셋

- 2020 NIA '교통 안전 시 데이터' 참여



[고속도로 Bounding Box 데이터]

이미지 수	객체 수(Class)			
	car	bus	truck	합계
301,329	1,413,751	55,028	637,579	2,106,358

[고속도로 Polygon Segmentation 데이터]

이미지 수	객체 수(Class)			
	car	bus	truck	합계
265,618	1,053,640	47,370	509,846	1,610,856