

Original Article

<https://doi.org/10.12985/ksaa.2026.34.3.160>
ISSN 1225-9705(print) ISSN 2466-1791(online)

항공기 지상조업 관제를 위한 YOLO 기반 데이터셋 구축 및 객체 인식 통합 프레임워크

이호준*, 정재훈**, 안순홍***

Aircraft Ground Handling Surveillance: An Integrated YOLO-Based Framework for Dataset Construction and Object Detection

Hojun Lee*, Jaehun Jung**, Soonhong An***

ABSTRACT

Ground handling complexity and reliance on manual visual control lead to persistent collision risks and reduced airport efficiency. To overcome these limitations, we propose an integrated YOLO-based framework, which merges auto-labeling(pseudo-labeling) and object detection into a single pipeline. This framework dramatically cuts manual data costs, enabling a gradual dataset expansion from v_0 (3,003 pairs) to the final v_4 (87,410 pairs). We confirmed this expansion strategy effectively improved object detection performance, achieving $mAP@0.5:0.95$ of 0.94. The final v_4 dataset secured both quantitative efficiency and qualitative balance(Gini coefficient converged to the initial level). This integrated framework is anticipated to establish a self-learning system for smart airport surveillance, maximizing long-term safety and operational efficiency.

Key Words : Airport Ground Handling(공항 지상조업), Object Detection(객체 탐지), Pseudo-labeling(의사 라벨링), Dataset Expansion(데이터셋 확장)

1. 서 론

지상조업은 항공기가 공항에 도착한 직후부터 다음 비행을 위해 이륙하기 전까지 주기장 내에서 수행되는 일련의 현장 지원 활동으로, 항공 운항의 정시성과 안전성을 좌우하는 핵심 요소이다(IATA, 2025). 그러나 조업 환경은 제한된 공간적 특성상 다수의 작업자와

다양한 특수장비가 동시다발적으로 교차하며 작업 구역을 공유하기 때문에 고도의 복잡성을 띤다. 최근 항공 수요의 회복과 함께 이러한 복잡도가 더욱 커지면서, 근접 작업 중 발생하는 장비 간 충돌 및 작업자 안전사고 위험이 지속적으로 보고되고 있다. 더불어 지속되는 인력난과 육안 감시에 의존하는 수동적 통제 방식은 주기장 내 사각지대 발생과 상황 인지 지연을 초래하여 공항 운영 전반의 효율을 저하시키는 주요 원인으로 지목된다. 이러한 물리적 제약과 인적 한계를 극복하고 예측 불가능한 현장 상황에 즉각적으로 대응하기 위해, 비전 기술 기반의 인공지능을 활용한 지능형 관제 시스템 도입이 시급한 실정이다(IATA, 2025; He et al., 2025).

Received: 8. Jun. 2026, Revised: 21. Aug. 2026,

Accepted: 28. Aug. 2026

* 아시아나IDT AI빅데이터연구소 연구원

** 아시아나HDT AI빅데이터연구소 주임연구원

*** 아시아나IDT AI빅데이터연구소 연구소장

연락처 E-mail : ansh@asianaidt.com

연락처 주소 : 서울시 종로구 우정국로 26

국토교통부는 계류장 전반에 대한 지능형 관제 고도화를 핵심 과제로 정의했으며, 디지털트윈 기반의 사전 예방형 안전 체계 구축을 세부 전략목표로 제시하였다(국토교통부, 2025). 이러한 흐름에 맞춰 인천국제공항은 CCTV를 기반으로 항공기와 지상조업 차량 간의 경로 이탈을 예측 및 경보하는 통합관제 플랫폼을 단계적으로 구축하고 있다(차중환, 2022).

이러한 AI 기반의 관제 시스템을 실제 조업 현장에 성공적으로 적용하기 위해서는 주기장 내 혼재된 다양한 객체(항공기, 특수 차량, 작업자 등)를 실시간으로 정확하게 인식하고 추적할 수 있어야 한다(Redmon et al., 2016). 특히 지상조업 구역은 야간 조명, 기상 악화 등 외부 환경 변화가 극심하고, 장비 간의 빈번한 가림이 발생하므로, 해당 환경에 특화된 고품질의 학습 데이터셋이 대량으로 요구된다(Bingol et al., 2026). 하지만 현장 영상을 수동으로 라벨링하는 기존의 방식은 막대한 시간과 비용을 소모할 뿐만 아니라, 인적 오류로 인한 데이터 품질 저하를 유발할 수 있다(Roh et al., 2019).

이에 본 연구는 지상조업 관제를 위한 YOLO(you only look once) 기반의 데이터셋 구축 및 객체 인식 통합 프레임워크를 제안한다. 이는 라벨링 과정을 자동화함으로써 데이터 구축 비용을 획기적으로 절감할 수 있으며, 지상조업 객체의 안정적인 인식 성능을 보장한다.

본 연구의 주요 기여는 다음과 같다. 첫째, 현업 담당자의 요구를 반영한 지상조업 객체 14개를 정의하고, 수동 라벨링을 시작으로 단계적인 의사 라벨링을 통해 데이터셋 확장을 수행했다. 둘째, 지상조업 환경에 딥러닝 기반의 객체 탐지 모델을 적용하고, 이를 라벨 생성 모델로 활용하여 통합 프레임워크를 구성했다.

II. 본 론

2.1 선행연구 고찰

단일 단계 구조로 실시간성을 확보할 수 있는 YOLO 알고리즘을 기반으로 다양한 지상조업 객체를 인식하는 연구가 수행되고 있다. He et al.(2025)은 YOLOv5 기반의 DT-YOLO를 제안했으며, 작업자와 소형 객체의 탐지 성능을 향상시켰다. Zhou et al.(2023)은 공항 지상 감시를 위해 ASS 데이터셋을 구축했으며, ASSD-YOLO를 제안했다. 이는 소형 장

비의 정밀도를 높이기 위해 소형 전용 탐지 레이어와 어텐션 기반 모듈을 추가했다. 그 결과, 93.5%의 mAP와 다양한 기상에서 탐지 정확도를 확보했다. 또한, Zhao et al.(2026)은 EIDM(edge information detection module)과 SFFM(selective feature fusion module)을 YOLO11에 더하여 다양한 시간대에서 소형 객체의 탐지 성능을 향상시켰다.

한편, 라벨링 시간과 비용을 줄이기 위해 제로샷(zero-shot), 퓨샷(few-shot), 오픈 보캐블러리(open-vocabulary) 객체 탐지 기법 등이 제시되고 있다. 그러나, 이는 도메인 특화 객체를 대상으로 성능이 낮은 한계를 보인다(Bianchi et al., 2024; Zhang et al., 2025). 이에 따라, YOLO 기반의 라벨링을 통해 데이터셋 구축을 최적화하는 시도가 지속되고 있다. Pham et al.(2024)은 YOLOv5 기반의 모델 보조 라벨링 프로세스를 수립하여 대규모 데이터셋을 구축하는 과정에서 라벨링 비용을 70% 이상 감소시켰다. Lester et al.(2025)은 사전 학습된 YOLOv8 기반의 반자동 라벨링 프로세스를 제안하여 철도 결함 탐지 데이터셋 구축의 효율성을 높였다. 위 연구는 모델의 자동 감지와 인간의 검수를 결합하여 오류를 최소화하는 방식으로 데이터셋을 점진적으로 확장했다. 이를 통해 이미지당 라벨링 시간을 기존 2~4분에서 30초~2분 수준으로 대폭 단축함으로써, 대규모 데이터셋 구축을 위한 비용 효율적인 대안을 제시했다.

2.2 방법론

본 연구는 주기장 내 지상조업 과정에서 다양한 객체의 오토라벨링과 실시간 탐지를 위해 YOLO11을 적용했다. YOLO는 단일 단계 구조로, 이미지 전체를 한번에 분석하여 객체의 위치와 종류를 예측하는 알고리즘이다. 특히, YOLO11은 YOLOv8 이후에 아키텍처와 학습 측면의 개선을 통해 정확도와 연산 효율의 균형을 강화한 모델로(Fig. 1), 다중 객체가 혼재되어 가

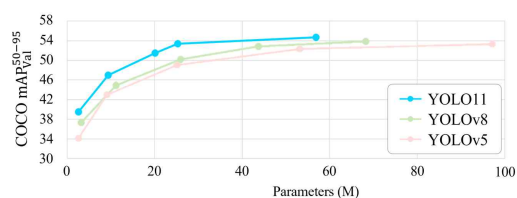


Fig. 1. Comparison of YOLO model performance

림이 빈번한 지상조업 환경에서 실시간 탐지 성능을 확보하는 데 유리하다.

한편, Ultralytics는 2026년 1월 YOLO26을 공개했으나, 배포 단순화 및 엣지 환경 친화성에 초점을 두었다. 또한, 공개 직후 모델로 운용 안정성은 추가 검증이 필요하다고 판단했다. 이에 본 연구는 YOLO11을 방법론으로 선정했다.

2.3 실험

2.3.1 프레임워크

본 연구에서 제안하는 프레임워크는 Fig. 2와 같다. 지상조업 객체를 수동으로 라벨링하여 기초 데이터셋(v_0_D)을 생성하고, YOLO 모델을 학습 및 최적화하여 기준 YOLO 모델(v_0_M)을 구축한다. 이들은 데이터셋 규모 및 불균형 정도와 객체 탐지 성능의 개선 여부를 정량적으로 비교하는데 활용된다.

이후, 여러 이미지(v_n_I)를 현행 모델(v_n_M)에 입력하여 객체 인식을 수행하고, 이를 기반으로 오토라벨링 데이터($v_n_AL_D$)를 생성한다. 이때, 초기 모델의 확증 편향을 완화하고, 성능 붕괴 최소화를 위해 데이터를 단계적으로 추가하는 점진적 확장 전략을 채택했다. 오토라벨링 데이터($v_n_AL_D$)는 기존 데이터셋(v_n_D)과 병합되어 차기 데이터셋(v_{n+1}_D)으로 구축되며, 재학습과 최적화를 거쳐 신규 YOLO 모델(v_{n+1}_M)을 생성한다. 이러한 순환 과정을 반복함에 따라 신규 모델이 현행 모델(v_n_M)을 지속적으로 대체하며, 모델은 확장된 데이터 분포를 반영하여 객체 탐지 성능이 점진적으로 고도화된다. 즉, YOLO 모델은 차기 학습을 위한 라벨

생성기인 동시에 실시간 객체 탐지 모델의 역할을 수행한다.

2.3.2 데이터 수집 및 전처리

데이터 수집을 위해 김포공항 주기장에서 지상조업 전체 과정을 촬영했다. 항공기를 기준으로 정면, 우측 후방, 좌측후방 시점을 기준으로 총 120GB 규모의 24개의 영상을 확보했다. 또한, 시간대, 기상, 형태 등의 조건에서 데이터 다양성을 확보하기 위해 지상조업 객체가 포함된 외부 데이터셋을 추가로 조사했다. 그 결과, AGVS 데이터셋(Zhang et al. 2022)과 AI-Hub의 활주로 내 이상물체 감지 객체 이미지 데이터를 확보했으며, AI-Hub는 약 974GB의 규모로 이미지-라벨 쌍과 이미지가 혼재되어 있다.

다음으로 항공사 램프마스터와 협의를 통해 지상조업과 관련성이 높은 14개 객체를 선정하였다(Table 1). 이후 방향 별 영상을 이미지로 저장한 뒤, 바운딩박스 라벨링을 수행했다. 그 결과, 초기 학습을 위한 이미지-라벨 데이터 3,003쌍을 구축했다. 이를 바탕으로 초기 YOLO 모델을 구축했으며, 추가 데이터셋은 김포공항 주기장 영상, AGVS 데이터셋, AI-Hub 데이터를 활용했다. AI-Hub 데이터는 14개 객체가 존재하는 이미지를 선별했으며, 라벨 데이터가 없는 경우에는 직접 라벨링을 수행하여 데이터를 확보했다. 이 과정에서 항공

Table 1. Ground handling objects

Index (label)	Object
0	Chock
1	Refueling car
2	Bridge
3	Safety cone
4	Car
5	Worker
6	ULD loader
7	Belt loader
8	ULD
9	Catering truck
10	Tug car
11	Towing car
12	Palette dolly
13	Aircraft

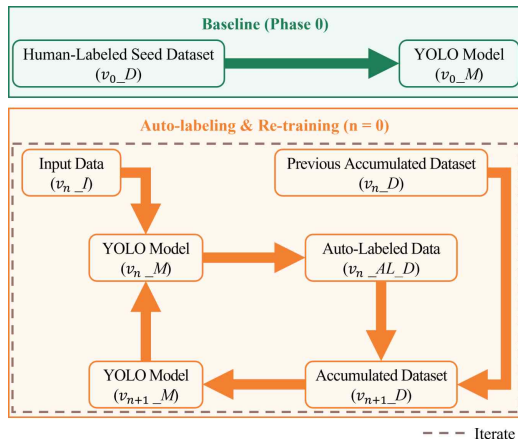


Fig. 2. Framework

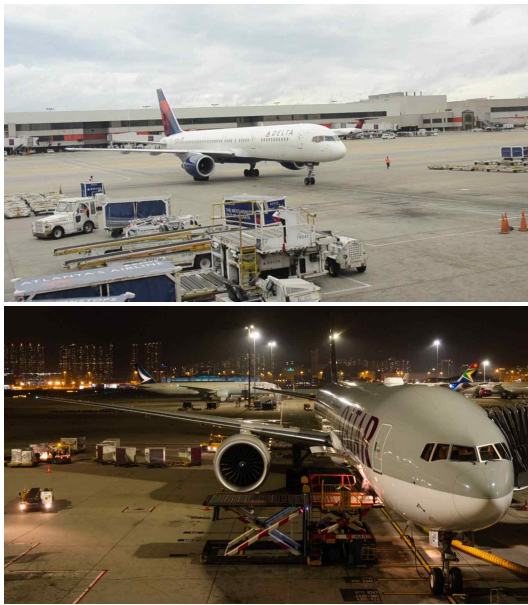


Fig. 3. Ground handling image data

기, 카고로더, 카고컨베이어, 토잉카 등 지상조업 객체의 외형이 다르게 나타났으며, 이를 통해 객체 형태의 다양성을 확보했다(Fig. 3).

2.3.3 모델 학습 및 최적화

본 연구는 YOLO11 모델의 객체 탐지 성능을 향상시키고, 데이터셋의 신뢰도를 확보하기 위해 다양한 파라미터를 조합했다(Table 2).

우선, 5개 크기의 모델(n, s, m, l, x)을 활용했으며, 입력 이미지의 해상도는 320, 640, 1,280 px로 설정했다. 가중치 업데이트를 위한 옵티마이저는 전통적인 SGD부터 수렴 속도가 빠른 Adam, 일반화 성능이 개선된 AdamW를 적용했다. 또한, 학습률(learning rate)을 조절하는 코사인 스케줄(cos_lr) 여부를 선택

Table 2. YOLO training parameters

Parameter	Detail
Model type	n, s, m, l, x
Epoch	100, 200, 300
Image size	320, 640, 1,280
Bach size	-1, 2, 4, 8, 16, 32, 64
Early stopping patience	10, 20, 30
Optimizer	SGD, Adam, AdamW
Cos_lr	True, False

Table 3. Pseudo labeling process

Training dataset	Process	Output
v_0	v_0_M training	v_0_AL
$v_1(v_0+v_0_AL)$	v_1_M training	v_1_AL
$v_2(v_1+v_1_AL)$	v_2_M training	v_2_AL
$v_3(v_2+v_2_AL)$	v_3_M training	v_3_AL
-	-	$v_4(v_3+v_3_AL)$

하여 최적의 수렴 경로를 탐색했다. 학습 횟수(epoch)는 100에서 300까지 설정하고, 가중치를 업데이트하는 배치 크기(batch size)를 자동 최적화(-1)하거나 2에서 64까지 다양화했다. 특히, 모델의 성능 개선이 정체될 때 학습을 중단하는 조기 종료(early stopping)를 도입하고, 대기시간(patience)을 10에서 30까지 설정하여 과적합을 방지하고 학습 효율을 높였다.

본 연구는 4개의 데이터셋을 바탕으로 4개의 YOLO 모델을 학습 및 최적화했다(Table 3). v_4 데이터셋은 마지막 YOLO 모델(v_3_M)의 의사 라벨링을 통해 도출된 것으로, 이전 데이터가 모두 포함되어 있다. 인간이 라벨링한 v_0 데이터셋으로 v_0 모델을 최적화하고, 신규 이미지의 의사 라벨링(v_0_AL)을 v_0 과 통합하여 v_1 데이터셋을 구축했다. v_1 데이터셋은 v_1 모델(v_1_M) 최적화에 활용되었으며, v_1 모델로 생성한 의사 라벨링(v_1_AL)을 기존 데이터와 병합하여 v_2 데이터셋을 구축했다. 이러한 과정을 반복하여 v_3 모델(v_3_M)과 v_4 데이터셋을 구축했다.

2.4 결과 분석

본 장에서는 제안 방법론의 타당성을 검증하기 위해 두 가지 관점으로 결과를 분석했다. 우선, 모델 관점에서 precision, mAP 등의 평가 지표를 통해 객체 탐지 성능을 분석했다. 이후, 객체별 데이터 개수와 불균형 지수 등을 분석하여 데이터셋의 적합성과 유효성을 평가했다.

2.4.1 객체 탐지 성능 결과

YOLO 모델의 학습 결과는 Table 4와 같다. 학습 결과, 데이터셋이 확장됨에 따라 모델 성능이 향상하는 경향을 보였다. 4번, 7번 11번, 15번 모델은 각각 0.88, 0.93, 0.85, 0.94를 기록했으며, v_3 데이터셋에서 최고 성능이 도출되었다. 이는 데이터셋의 점진적 확장이 객체 검출 성능 향상에 영향을 미치는 것을 시

Table 4. YOLO training result

No.	Dataset	Training parameter						Evaluation metric		
		Model	Epoch	Image size	Batch size	Patience	Optimizer	Precision	Recall	mAP @0.5:0.95
1	v ₀	m	300	640	32	10	AdamW	0.95	0.89	0.79
2	v ₀	m	200	640	32	10	AdamW	0.95	0.89	0.79
3	v ₀	m	200	1,280	32	10	AdamW	0.96	0.95	0.86
4	v ₀	l	200	1,280	16	10	AdamW	0.97	0.96	0.88
5	v ₁	s	200	1,280	8	20	AdamW	0.94	0.89	0.83
6	v ₁	l	200	640	8	20	AdamW	0.95	0.91	0.88
7	v ₁	l	200	1,280	16	10	AdamW	0.95	0.96	0.93
8	v ₂	s	200	640	32	10	AdamW	0.92	0.84	0.82
9	v ₂	l	200	1,280	8	20	AdamW	0.93	0.86	0.84
10	v ₂	l	200	640	16	20	AdamW	0.90	0.93	0.79
11	v ₂	l	200	1,280	16	30	AdamW	0.91	0.90	0.85
12	v ₃	m	200	1,280	-1	10	AdamW	0.94	0.92	0.92
13	v ₃	s	300	640	-1	30	AdamW	0.93	0.91	0.88
14	v ₃	m	300	640	-1	10	AdamW	0.84	0.85	0.76
15	v ₃	s	200	1,280	-1	20	AdamW	0.95	0.95	0.94

사한다. 이후, 세부 파라미터에 따른 모델의 성능을 분석한 결과, 다음과 같은 특징이 도출되었다. 첫 번째로, 입력 이미지의 해상도가 객체 검출 결과에 영향을 미쳤다. 데이터셋이 동일하더라도 입력 이미지의 크기를 640에서 1,280으로 확장했을 때, 모델의 종류와 상관 없이 성능 향상이 나타났다. 특히, v₀에서 해상도를 높였을 때(2번, 3번), mAP가 약 10% p 상승했다. 이는 고해상도 이미지가 작은 객체나 복잡한 배경에서 특징 추출에 유리한 것을 의미한다. 두 번째로, 데이터셋의 점진적 확장이 모델 규모에 따른 성능 차이를 상쇄하는 효과가 있는 것으로 나타났다. v₀~v₂에서는 상대적으로 큰 YOLO11l 모델의 성능이 높게 도출되었고, 모델 크기에 따른 성능 편차도 존재했다. 그러나, 데이터셋의 확장에 따라 해당 격차는 줄어드는 양상을 보였다. 특히, v₃ 데이터셋은 경량 모델인 YOLO11s가 mAP 0.94를 기록하며, 최고 성능을 도출했다(15번). 이는 양질의 데이터를 확보하는 것이 성능 향상에 필수적인 것을 시사한다. 특히, 데이터셋 버전을 기준으로 최고 성능은 4번, 7번, 11번, 15번 모델인 것을 확인할 수 있다. 4번 모델과 7번 모델은 데이터셋 버전만

v₀과 v₁으로 다르고, 모델의 종류, 입력 해상도, batch size 등 동일한 조건으로, mAP가 0.88에서 0.93으로 상승한 것을 확인할 수 있다. 이를 통해 학습 데이터 규모의 증가와 상대적으로 적었던 지상조업 장비 데이터가 보강된 것이 성능 향상에 도움이 되는 것으로 해석할 수 있다. 또한, 다양한 학습 파라미터의 조합을 통해 v₃에서 최고 성능인 0.94가 도출된 것은 데이터셋의 양적 확장이 탐지 성능 개선과 관련될 수 있음을 시사한다.

추가로, Table 4의 15개 모델을 대상으로 정답(true positive)인 바운딩박스의 폭, 높이, 면적을 원본 이미지 좌표계에서 산출했다. 이를 위해 각 데이터셋에서 활용한 테스트 이미지를 활용했으며, 이는 Full HD(1,920×1,080)급에서 4K(3,840×2,160)급 해상도이다. Table 5는 최소 바운딩박스의 폭(w), 높이(h), 면적(area)을 픽셀 단위로 정리한 결과이다. 분석 결과, 입력 크기가 640인 모델은 11~13 px의 폭, 9~12 px의 높이, 210~270 px² 면적으로, 2번 모델의 바운딩박스가 가장 작은 것을 확인했다. 입력 크기가 1,280인 경우, 5~10 px의 폭, 8~10 px의 높이, 면적은 66~192 px² 범위로 나타났고, 4번 모델의 바운딩박스가

Table 5. Minimum bounding-box sizes

No.	Image size	Min_w (px)	Min_h (px)	Min_area (px ²)
1	640	12.59	10.37	266.74
2	640	12.44	9.26	210.63
3	1,280	6.25	8.84	96.55
4	1,280	4.69	8.88	66.10
5	1,280	6.82	8.34	102.53
6	640	13.33	8.63	210.94
7	1,280	6.67	8.31	101.08
8	640	12.31	10.32	219.90
9	1,280	9.31	10.23	174.86
10	640	11.20	12.06	251.90
11	1,280	10.34	9.08	178.97
12	1,280	7.92	9.78	192.15
13	640	12.01	10.98	269.77
14	640	12.58	10.18	228.28
15	1,280	8.20	7.75	178.46

가장 작은 것을 확인했다(66.10 px²). 특히, 대부분 바운딩박스의 폭, 너비, 면적은 1,280이 작은 것으로 보아 입력 해상도가 증가할수록 더 작은 픽셀 규모의 객체까지 탐지가 가능한 것을 시사한다. 그러나, 2번 모델의 최소 바운딩박스는 210.63 px² 크기의 고임목이다. 해당 이미지의 해상도는 2,336×1,080 px로, 전체 이미지 크기의 약 1/12,000 수준이다. 따라서, 입력이 640 이라고 하더라도 실질적인 탐지 성능을 확보할 수 있다는 것을 확인했다.

2.4.2 학습 데이터셋 분석 결과

본 연구는 지상조업에 특화된 객체 인식 모델과 데이터셋을 구축하기 위해 5단계의 점진적 확장 전략을 채택했다. 이 과정에서 최적화가 완료된 4개의 YOLO 모델을 활용하여 의사 라벨링을 수행했다(Table 3). v₀ 부터 v₄까지 연구를 통해 구축한 데이터셋의 세부 구성은 Table 6과 같다. 초기 단계인 v₀ 데이터셋은 인간이 직접 라벨링을 수행했으며, 3,003쌍의 이미지-라벨과 32,544개의 객체가 포함되어 있다. 이후, 점진적 확장을 통해 87,410쌍의 이미지-라벨과 849,183개의

Table 6. Dataset statistics and object distributions across versions v₀ to v₄

Dataset version		v ₀	v ₁	v ₂	v ₃	v ₄
Image-Label pairs		3,003	6,525	18,134	57,383	87,410
Annotations	Total	32,544	89,334	150,458	547,819	849,183
	Chock	3,321	9,183	10,136	38,032	64,195
	Refueling car	742	1,053	3,291	12,695	29,529
	Bridge	1,360	6,222	13,483	44,643	74,149
	Safety cone	8,755	38,379	60,647	175,360	278,982
	Car	2,050	3,089	3,810	24,092	24,240
	Worker	6,620	12,737	22,405	99,442	129,804
	ULD loader	1,469	1,991	2,788	25,562	37,853
	Belt conveyor	1,459	1,512	1,754	15,738	33,897
	ULD	937	2,420	5,998	16,745	18,235
	Catering truck	710	736	934	12,774	22,814
	Tug car	659	2,172	5,603	14,007	18,263
	Towing car	1,310	2,448	3,994	17,248	30,474
	Palette dolly	785	4,163	7,276	16,042	23,447
	Aircraft	2,367	3,229	8,339	35,439	63,301

객체로 구성된 v_4 데이터셋을 구축했으며, 이는 초기 대비 약 26배에서 30배 증가한 수치이다. 또한, v_0 데이터셋에서 급유차(742개), 컨테이너(937개), 케이터링 트럭(710개), 터그카(659개), 팔레트 돌리(785개)는 안전콘(8,755개)과 작업자(6,620개)에 비해 표본 수가 매우 적었다. 이에, v_4 데이터셋 구축 과정에서 해당 클래스의 표본을 확보하는 것에 중점을 두었으며, 최종적으로 약 18,000개~30,000개로 증가했다. 특히, 시간이 지날수록 모든 클래스가 단계별로 증가하는 양상을 확인할 수 있다. 최종적으로 구축한 v_4 데이터셋은 안전콘(278,982개)과 작업자(129,804개)가 가장 많았으며, 컨테이너(18,235개), 터그카(18,263개)가 상대적으로 낮은 분포를 보였다. 이처럼 클래스 간 빈도 차이는 지상조업 프로세스의 운영적 특성에서 기인한 것으로 분석된다. 안전콘과 작업자는 조업의 전 과정에서 빈번하게 노출되는 반면, 컨테이너는 수하물 하기 및 탑재, 터그카는 항공기 푸시백 과정 등 특정 작업 구간에서만 일시적으로 등장한다. 이러한 특성에도 불구하고, v_4 데이터셋은 지상조업 특수장비인 터그카, 케이터링 트럭, 토잉카, 팔레트 돌리 등이 초기 대비 15배에서 30배 이상 증가한 것을 확인할 수 있다. 특히, v_3 데이터셋에서 v_4 데이터셋으로 확장되는 과정에서 항공기와 지상조업 특화 장비의 개수가 지속적으로 증가했다. 이는 본 연구가 일반 객체에 매몰되지 않고, 특수장비까지 고려하여 지상조업 데이터셋으로서의 가치를 확보했다는 것을 입증한다.

모델의 학습 과정에서 클래스 간 데이터 불균형은 수렴 속도가 느려지고, 일반화 성능을 저하시킨다(Oksuz et al., 2020). 이에, Table 6를 기반으로 5개 데이터셋의 클래스 간 불균형 정도를 측정했으며, 로렌츠 곡선(Lorenz curve)과 지니계수(Gini coefficient)를 활용했다. 로렌츠 곡선은 Fig. 4와 같으며, 가로축

은 클래스의 누적 비율을, 세로축에 해당 클래스가 점유하는 개수의 누적 비율을 의미한다. 또한, 로렌츠 곡선은 모든 클래스가 동일한 빈도를 가질 시, 45도의 직선이 그려진다. 이 직선에서 멀어질수록 데이터의 분포가 특정 클래스에 편중된 것을 의미한다.

지니계수(G)는 로렌츠 곡선과 완전 평등선 사이의 면적을 이용하여 산출되는 정량적 지표이다. 이는 전체 클래스 수(n), 각 클래스의 객체 수(x_i), 클래스 별 객체 수의 산술 평균($\bar{x}$)을 이용하여 산출되며, 0에서 1 사이의 값을 가진다(식 (1)). 지니계수가 0에 가까울수록 모든 객체가 균등하게 분포함을 의미하며, 1에 가까울수록 특정 객체가 전체 데이터의 대부분을 차지하는 롱테일 분포를 형성하고 있는 것을 나타낸다.

$$G = \frac{2 \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{2n^2 \bar{x}} \quad (1)$$

분석 결과, 인간이 직접 라벨링을 수행한 초기 데이터셋(v_0)의 지니계수는 0.462가 도출됐다. 이는 작업자, 안전콘 등 특정 객체가 빈번하게 노출되는 주기장의 특성에 기인하고, 원천 데이터 자체에 일정 수준의 데이터 불균형이 존재하는 것을 알 수 있다. 이후, YOLO를 활용한 의사 라벨링을 통해 구축한 v_1 데이터셋은 지니계수가 0.583까지 상승하며 불균형이 다소 심화되었으나, v_2 , v_3 를 거치며 지니계수가 점진적으로 하락하는 추세를 보였다. 최종 데이터셋인 v_4 의 지니계수는 0.470이 도출되어 의사 라벨링을 수행한 데이터셋 중 가장 낮은 값을 기록했으며, 초기 데이터셋(v_0) 수준의 분포에 근접하게 수렴했다.

이러한 데이터 불균형의 변화는 Table 4의 데이터셋 버전별 최고 성능과 함께 살펴볼 수 있다. 데이터셋 버전이 올라갈수록 학습에 활용되는 데이터 양은 증가하며, 불균형은 v_1 에서 심화된 뒤 v_2 이후 점차 완화되는 양상을 보였다. 안전콘은 v_0 에서 v_1 으로 확장될 때, 약 27%에서 43%로 비중이 확대된 반면, 케이터링 트럭과 카고컨베이어는 같은 구간에서 개체 수 증가가 제한적이었다(Table 6). 이후, 데이터셋이 확장되면서 초기 대비 적었던 지상조업 특화 장비가 보강되었고, 지니계수는 다시 낮아지며 v_4 는 v_0 과 유사한 수준에 근접하였다. 이러한 흐름은 버전별 최고 성능과도 연계된다. v_0 에서 v_1 으로 확장되며 mAP가 0.88에서 0.93으로 상승한 점은 데이터 규모의 증가와 객체 구성의 확장이 탐지 성능 개선과 함께 나타난 결과로 볼 수 있

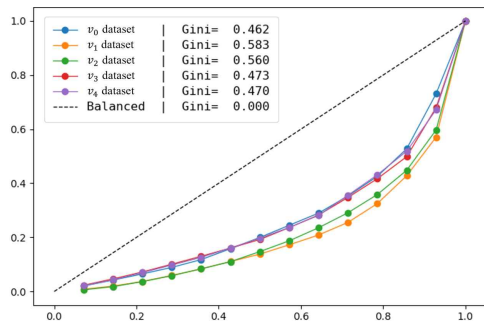


Fig. 4. Lorenz curve

다. 또한, v_3 에서 0.94의 최고 mAP가 도출된 점 역시 양적 확장과 클래스 분포의 안정화가 병행된 구간에 해당한다. 따라서 본 연구의 데이터셋 확장 과정은 양적 향상과 질적 균형을 동시에 확보했으며, 탐지 성능이 개선되는 경향을 확인했다는 점에서 의의가 있다.

III. 결 론

본 연구는 항공기 지상조업 관제의 복잡성 및 인적 한계를 극복하기 위해 YOLO 기반 객체 인식 통합 프레임워크를 제안했다. YOLO 모델의 최적화를 통한 객체 인식과 의사 라벨링을 통해 데이터셋을 점진적으로 확장했으며, 이를 통해 고품질 데이터셋을 효과적으로 구축했다. 실험 결과, 데이터셋의 점진적 확장은 객체 탐지 성능 향상에 기여한 것으로 해석되며, 최종적으로 mAP@0.5:0.95 0.94를 도출했다. 특히, 본 연구에서 구축한 데이터셋은 초기 대비 약 30배 확장됐을 뿐만 아니라, 지니계수가 초기 수준에 근접하게 도출되며 양적 효율성과 질적 균형을 동시에 확보한 것의 의의가 있다.

본 연구에서 제시한 통합 프레임워크는 실제 공항 주기장의 지상조업 환경을 촬영하여 지상조업 객체의 종류와 위치를 영상으로 인식함으로써, 관제자의 상황 인지를 보완하는 보조 정보로 활용될 수 있을 것으로 기대된다. 특히, 다수 장비와 작업자가 동시에 교차하는 혼잡 구간에서 객체의 출현 여부를 정량적으로 파악하면, 위험 및 공정 진행 상황을 보다 신속히 확인하는 데 기여할 수 있다. 또한, 탐지 결과를 관제 시스템과 연계할 경우, 스케줄 운영 효율과 작업 안전성 제고에 도움이 될 수 있으며, 의사 라벨링을 통한 데이터셋의 지속적 확장은 현장 데이터 구축 부담을 완화하고 모델의 갱신을 지원하는 기반이 될 수 있다.

본 연구의 결과는 실제 공항 지상조업 관제에 바로 적용하기에는 다음과 같은 한계가 있다. 우선, 실제 주기장의 운영 CCTV와 동일한 설치 조건(카메라, 높이, 화각 등)에서 현장 검증을 수행하지 못했다. 또한, 지상조업 영상을 최대한 많이 확보했으나, 야간, 기상 변화 등 탐지환경의 모든 제약을 검토하지 못했다. 마지막으로, 제안 프레임워크는 실제 관제시스템과 연계과정에서 필수적인 통신체계, 운영 절차 등은 고려하지 않았다.

이러한 한계를 바탕으로 향후 연구로는, 본 연구에서 활용한 영상보다 낮은 해상도의 조건에서 객체 탐

지 성능을 검증할 필요가 있다. 이와 함께 객체의 실제 크기, 카메라 설치 거리, 촬영 각도를 함께 고려하여 탐지 가능 범위를 정량화함으로써, 실제 공항 운영 환경에서 최소 기준을 제시할 필요가 있다. 두 번째로, 실제 관제 시스템과의 연계를 위해서는 현장에서의 유효성 검증과 운영 요건 확인이 선행되어야 하며, 통신 체계, 관제 시스템의 구성 요소 및 연계 방안은 후속 연구에서 단계적으로 구체화할 필요가 있다. 세 번째로, YOLO26 등 지속적으로 발전하는 후속 모델을 본 프레임워크에 적용한다면, 해당 모델로 교체한 뒤 탐지 성능과 의사 라벨링 성능을 검증하는 연구가 필요하다. 마지막으로, 향후 다양한 형태의 외부 자료를 활용할 수 있다면, 본 연구에서 제시한 프레임워크에 적용하여 그 범위를 확대할 필요가 있다.

References

1. IATA, "2025 Vision for the Future of Air Cargo Facilities", 2025, p.9.
2. He, Z., He, Y., and Lv, Y., "DT-YOLO: An improved object detection algorithm for key components of aircraft and staff in airport scenes based on YOLOv5", *Sensors*, 25(6), 2025, Article 1705.
3. (주)에어로솔루션즈, "AI·XR 기반 비행장 통합관제 운용기술 개발 기획연구 최종보고서[스마트공항 4.0 기술개발 기획]", 2022, pp.28-29.
4. 국토교통부, "제4차 항공정책기본계획(2025-2029)", 2025, p.85.
5. 차중환, "인천국제공항 관제시스템 더 똑똑해진다", 뉴스1, 2022, Available from: <https://www.koit.co.kr/news/articleView.html?idxno=98328>
6. Redmon, J., Divvala, S., Girshick, R., and Farhadi, A., "You only look once: Unified, real-time object detection", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp.779-788.
7. Bingol, E. C., Al-Rawashidy, H., and Banitsas, K., "Vision-based dual-mode collision risk-warning for aircraft apron monitoring", *Drones*, 10(3), 2026, Article 173.
8. Roh, Y., Heo, G., and Whang, S. E., "A survey

- on data collection for machine learning: A big data-AI integration perspective", *IEEE Transactions on Knowledge and Data Engineering*, 33(4), 2019, pp.1328-1347.
9. Oksuz, K., Cam, B. C., Kalkan, S., and Akbas, E., "Imbalance problems in object detection: A review", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10), 2020, pp.3388-3415.
 10. Zhao, Y., Kong, Z., Hu, Y., Wen, Z., and Wang, S., "Airport-YOLO: A fixed-camera solution for airport object detection", *IET Image Processing*, 20(1), 2026, Article e70387.
 11. Zhou, W., Cai, C., Zheng, L., Li, C., and Zeng, D., "ASSD-YOLO: A small object detection method based on improved YOLOv7 for airport surface surveillance", *Multimedia Tools and Applications*, 83(18), 2024, pp. 55527-55548.
 12. Lester, D., Gao, J., Sutphin, S., Zhu, P., Narman, H., and Alzarrad, A., "A YOLO-based semi-automated labeling approach to improve fault detection efficiency in railroad videos", *arXiv Preprint arXiv:2504.01010*, 2025.
 13. Zhang, X., Shu, C., Li, S., Wu, C., and Liu, Z., "AGVS: A new change detection dataset for airport ground video surveillance", *IEEE Transactions on Intelligent Transportation Systems*, 23(11), 2022, pp.20588-20600.
 14. Pham, T. N., Nguyen, V. H., Kwon, K. R., Kim, J. H., and Huh, J. H., "Improved YOLOv5 based deep learning system for jellyfish detection", *IEEE Access*, 12, 2024, pp.87838-87849.
 15. Bianchi, L., Carrara, F., Messina, N., Gennaro, C., and Falchi, F., "The devil is in the fine-grained details: Evaluating open-vocabulary object detectors for finegrained understanding", *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp.22520-22529.
 16. Zhang, Y., Han, R., Zhou, F., Wang, S., Feng, W., and Wan, L., "ODOV: Towards open-domain open-vocabulary object detection", *arXiv Preprint arXiv:2508.01253*, 2025.