

# 동관 브레이징 용접 위치 검출을 위한 멀티클래스 YOLO 모델과 클래스별 YOLO 결정 수준 모델의 성능 비교

## Performance Comparison of a Multi-class YOLO Model and Class-wise YOLO Model with Decision-level Output Selection for Welding Position Detection in Copper Pipe Brazing

양수빈<sup>1,†</sup>, 강환희<sup>1,†</sup>, 김영곤<sup>1,2</sup>, 김명진<sup>1</sup>, 김유진<sup>1</sup>

(Subin Yang<sup>1,\*</sup>, Hwanhee Kang<sup>1,2,\*</sup>, Younggon Kim<sup>1</sup>, Myeongjin Kim<sup>1</sup>, and Eugene Kim<sup>1,\*</sup>)

<sup>1</sup>Purpose Built Mobility group, Korea Institute of Industrial Technology

<sup>2</sup>Robot Engineering, Korea National University of Science and Technology

**Abstract:** This study investigates welding position detection for robotic copper pipe brazing using YOLOv11n-OBB object detection models. Two structures were compared, including a single multi-class model (Model A) and a class-wise trained model with decision-level selection (Model B). Both models were evaluated under identical training conditions in controlled and uncontrolled industrial environments. Results confirmed that while both models achieve high accuracy in controlled settings, Model A provided faster inference and more stable performance under environmental variations, whereas Model B displayed higher localization precision. These findings indicate Model A to be a more practical choice for real-time robotic brazing applications.

**Keywords:** copper pipe brazing, robotic welding, YOLO, decision-level detection, welding position detection

### I. 서론

동관(copper pipe) 브레이징은 가전제품, 난방·환기·공조 시스템(heating, ventilation, and air conditioning, HVAC), 자동차 열관리 장치 등 다양한 산업에서 핵심 제조 공정으로 중요한 역할을 한다[1]. 브레이징 품질은 제품의 구조적 신뢰성, 내구성, 열적 안정성에 직결되므로 공정 조건의 정밀한 제어와 관련된 품질 확보가 필수적이다[2]. 따라서 작업자 숙련도, 동관 형상 편차, 조명·반사 조건 변화에 적응하는 자동화 시스템이 요구되고 있다. 이러한 이유로 스마트 팩토리 및 지능형 제조 시스템에서는 비전 기반 용접부 인식과 품질 검사 기술의 수요가 증가하고 있으며, 고품질의 용접 자동화 기술이 빠르게 발전하고 있다[3,4].

최근 YOLO (You Only Look Once) 기반 객체 탐지 기술은 실시간 용접부 인식과 로봇 제어에 폭넓게 사용되고 있다[5,6]. 또한 선행 연구에서는 브레이징 타겟 검출[7], 카메라-로봇 좌표 변환 기반 위치 보정[8], 합성곱 신경망(convolutional neural

network, CNN) 모델을 사용해 용접 결함을 자동 분류하는 연구[9] 등 다양한 통합 프레임워크가 제안되었다. 이들 연구 중 다수는 단일 멀티클래스 객체 탐지 모델을 기반으로 용접 위치 인식 및 로봇 제어를 수행하거나 센서 기반 영상 분석을 통해 용접 품질을 평가하는 구조를 취하고 있다. 일반적으로 단일 멀티클래스 모델은 클래스 간 간섭, 불균형, 미세 시각 차이 등의 영향으로 성능이 저하가 발생할 수 있는 것으로 알려져 있다. 이에 따라 여러 단일 모델의 출력을 결합하는 앙상블 또는 다중 모델 융합(multi-model fusion) 접근이 단일 모델 대비 더 안정적인 성능을 제공함이 보고되었다[10-13]. 또한 최근에는 YOLO 기반 객체 탐지 모델을 추적 알고리즘과 결합하여 다중 객체 환경에서 탐지-추적 통합 성능을 분석하는 연구가 보고되었다[14]. 아울러 센서 기반 객체 추적 연구도 보고되었다[15]. 이들 연구는 탐지 결과를 후속 알고리즘과 연계하여 시스템 성능을 평가하였다는 점에서 의미가 있다.

그러나 기존 연구들은 주로 탐지 성능 향상, 추적 통합 또는 센서 및 기능 융합 구조에 초점을 두고 있어서 단일 멀티

†Equal Contribution

\*Corresponding Author

Manuscript received January 12, 2026; revised February 2, 2026; accepted March 10, 2026

양수빈: 한국생산기술연구원 서남기술실용화본부 목적기반모빌리티그룹 자체인턴(ysb25@kitech.re.kr, ORCID<sup>®</sup> 0009-0005-7823-6608)

강환희: 과학기술연합대학 로봇전공 UST학생(hwan@kitech.re.kr, ORCID<sup>®</sup> 0009-0003-5035-7059)

김영곤: 한국생산기술연구원 서남기술실용화본부 목적기반모빌리티그룹 수석연구원(ygkim1@kitech.re.kr, ORCID<sup>®</sup> 0000-0002-1344-5725)

김명진: 한국생산기술연구원 서남기술실용화본부 목적기반모빌리티그룹 선임연구원(mjkim002@kitech.re.kr, ORCID<sup>®</sup> 0000-0003-3451-4926)

김유진: 한국생산기술연구원 서남기술실용화본부 목적기반모빌리티그룹 선임연구원(egkim@kitech.re.kr, ORCID<sup>®</sup> 0000-0002-1441-4250)

※ 본 논문은 가전 공장의 비정형 생산제품의 고품질 관리를 위한 AI 기반 장비 제어 및 자율제조 운영 기술 개발과제 (KM250256)에서 지원하여 연구하였음.

클래스 모델과 추론 단계에서의 결정 수준(decision-level) 결과 결합 전략을 동일 조건에서 체계적으로 비교·분석한 연구는 제한적이다. 특히 자동화 용접 비전 기술에 대한 선행 연구에서도 다양한 딥러닝 기반 인식 및 검출 기법이 폭넓게 소개되었으나, YOLO 계열 모델을 중심으로 한 다중 모델 기반의 결정 수준 결과 결합 접근은 주요 연구 주제로 다루어지지 않았다[16,17].

이에 본 연구는 객체 탐지 알고리즘의 구조적 개선이 아닌 모델 구성 방식에 따른 성능 특성 차이를 분석하는 데 목적을 둔다. 이를 위해 모든 클래스를 하나의 모델로 통합 학습하는 단일 멀티클래스 YOLO 모델(이하 모델 A)과 각 클래스를 독립적으로 학습한 후 추론 단계에서 결과를 결정 수준에서 결합하는 클래스별 학습 기반 결정 수준 YOLO 모델(이하 모델 B) 구조를 비교 대상으로 설정하였다. 두 모델 구조에 대해 F1 점수(F1-score), IoU (Intersection over Union), mAP (mean Average Precision), FPS (Frames Per Second), 회전 각도 오차를 정량적으로 비교함으로써, 실제 동관 브레이징 공정에 적용 가능한 객체 탐지 구조의 특성과 구현 관점에서의 활용 가능성을 분석하고자 한다.

II. 데이터 수집 및 전처리

본 연구에서 사용된 데이터는 실제 산업 환경의 산업용 용접 로봇의 엔드 이펙터에 장착된 RGB 비전 센서를 통해 취득되었으며, 용접 대상 파이프 및 용접 위치 정보를 포함한다. 그림 1은 본 연구에서 사용한 학습 데이터 취득 환경을 나타낸다.

수집된 데이터는 RGB 이미지 데이터와 이에 대응하는 객체 위치 라벨로 구성된다. 이미지 데이터는 JPG 형식으로 저장되었으며 라벨 데이터는 YOLO 객체 탐지 학습에 사용 가능한 텍스트(txt) 형식으로 구성되었다. 각 라벨은 용접이 수행되어야 할 파이프 이음부(joint)와 필러(filler) 영역을 각각 OBB (Oriented Bounding Box) 형태로 정의하였다.

본 연구에서는 객체 유형 구분을 위해 파이프 형상과 필러 링의 조합을 기준으로 클래스를 정의하였다. 파이프 형상은 T형과 P형의 두 가지로 구분되며, 필러 링은 내경 크기에 따라 총 5종(R1-R5)으로 구분된다. 이 중 R3는 다른 필러 링과 달리 은납 재질을 사용한다. 그림 2는 본 연구에서 정의한 파이프 형상(T형, P형)과 필러 링의 종류를 나타낸다.

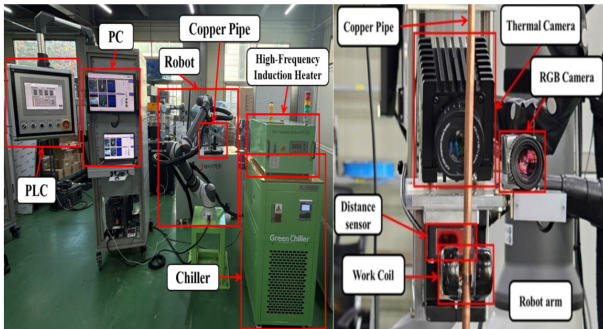


그림 1. 데이터 수집 환경 및 장치 구성.  
Fig. 1. Data collection environment and equipment configuration.



그림 2. 파이프 형상(T형, P형)과 필러 링들(R1-R5).  
Fig. 2. Pipe shapes (T-type, P-type) and filler rings (R1-R5).



그림 3. 파이프 형상(T형, P형)과 필러 링에 따른 클래스 구성 예시.  
Fig. 3. Examples of class configurations according to pipe shapes (T-type, P-type) and filler rings.

그림 2의 파이프와 필러 링을 크기에 따라 조합하여 그림 3과 같이 총 5개의 클래스를 구성하였다.

T1 파이프는 은납 재질을 사용하는 조합으로 T2 파이프 조합과 동일하여 C4 클래스로 통합하였다. 각 클래스에 사용된 파이프와 필러 링의 종류를 표 1에 정리하였다.

데이터 과적합을 방지하고 실제 산업 환경에서 발생할 수 있는 다양한 외부 조건에 대한 모델의 일반화 성능을 향상시키기 위해 데이터 증강(image augmentation)을 수행하였다. 본 연구에서는 표 2에 제시된 다섯 가지 증강 기법을 서로 다른 적용 확률로 설정하여 사용하였으며, 각 기법에 대해 랜덤 값을 선택하여 복합적으로 적용하였다. 그 결과, 각 원본 이미지에 대해 평균 3~4장의 증강 이미지가 생성되었다.

표 1. 파이프-필러 링 조합 클래스 구성.

Table 1. Pipe - filler ring combination classes.

| 클래스 | 파이프    | 필러 링 |
|-----|--------|------|
| C1  | P1, P3 | R4   |
| C2  | P2, P3 | R4   |
| C3  | P4, T2 | R2   |
| C4  | P6, T2 | R3   |
| C5  | P5, T1 | R1   |

표 2. 데이터 증강 기법 및 적용 조건.

Table 2. Data augmentation methods and parameters.

| 종류            | 적용 확률 (%) | 범위                                |
|---------------|-----------|-----------------------------------|
| 노출(Exposure)  | 100       | -40% ~ +100%                      |
| 회전(Rotation)  | 100       | -180° ~ +180°                     |
| 대비(Contrast)  | 50        | -40% ~ +40%                       |
| 노이즈(Noise)    | 30        | $\sigma$ 0.01 ~ 0.05<br>(표준편차 비율) |
| 스케일(Scale/축소) | 50        | 0% ~ 50%                          |

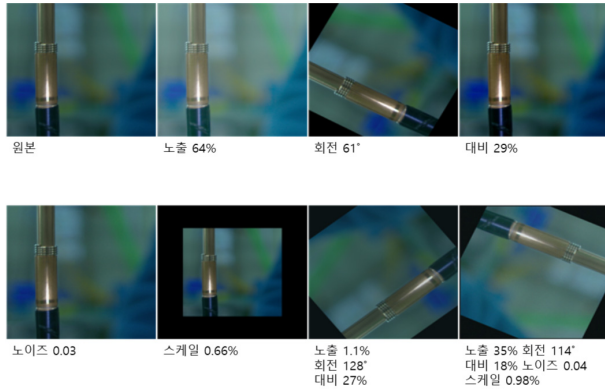


그림 4. 데이터 증강 기법 적용 예시.

Fig. 4. Examples of image augmentation techniques.

회전 및 스케일 변환과 같이 객체 위치에 영향을 미치는 증강 기법의 경우, 이미지와 라벨 데이터에 동일한 변환을 적용하여 객체 위치 정보의 정확성을 유지하였다. 여백 처리 시에는 검정 픽셀 채움과 거울 반사 방식을 랜덤 적용하였으며, 반사 처리는 약 25% 확률로 적용되도록 하였다. 이 과정에서 객체 영역을 침범하지 않도록 처리하였다. 그림 4는 데이터 증강 기법 적용 예시를 보여준다.

데이터 전처리 및 증강 후, 학습(training), 검증(validation), 테스트(test) 데이터셋은 총 2,138장으로 구성되었다. 학습 데이터는 원본 448장과 증강 1,393장을 포함한 1,841장, 검증 데이터는 원본 61장과 증강 189장을 포함한 250장, 테스트 데이터는 총 47장으로 구성하였다. 또한 모델 B의 클래스별 학습 구조를 고려하여, 클래스별 데이터 분포를 표 3에 정리하였다.

표 3. 클래스별 데이터 분포.

Table 3. Class-wise data distribution.

|         | Class 1 | Class 2 | Class 3 | Class 4 | Class 5 | 총합    |
|---------|---------|---------|---------|---------|---------|-------|
| 학습 (원본) | 70      | 93      | 108     | 92      | 85      | 448   |
| 학습 (증강) | 271     | 266     | 339     | 259     | 258     | 1,393 |
| 검증 (원본) | 9       | 15      | 17      | 8       | 12      | 61    |
| 검증 (증강) | 41      | 49      | 31      | 30      | 38      | 189   |
| 테스트     | 9       | 13      | 5       | 11      | 9       | 47    |

### III. 모델 학습 및 평가

전체 데이터셋은 학습, 검증, 테스트 데이터로 분할하였으며 약 86%, 12%, 2% 내외로 설정하였다. 모델 A와 B는 동일한 데이터 분할 및 전처리 조건을 사용하여 공정한 성능 비교가 가능하도록 하였다. 특히 테스트 데이터는 총 47장으로 구성되어 대규모 객체 탐지 연구 대비 상대적으로 작은 규모에 해당한다. 이에 따라 본 연구에서는 절대적 일반화 성능 산출보다는 동일 조건 하에서 모델 구조 간 상대적 성능 비교에 초점을 두었다.

본 연구는 산업 현장 적용을 고려하여 실시간 추론이 가능한 모델을 대상으로 하였다. Transformer 기반 탐지 모델들은 높은 성능이 보고되었으나[18], 높은 연산 복잡도와 추론 비용으로 인해 본 연구에서 요구하는 실시간 산업 환경에 적합하지 않다고 판단하였다. 따라서 OBB를 직접 지원하면서도 경량 구조를 갖는 YOLO 계열 모델을 대상으로 하였다.

그림 5는 모델 A의 구조를 나타낸다. 모델 A는 모든 클래스를 단일 YOLOv11n-OBB 모델로 통합 학습하는 멀티클래스 구조이다. 하나의 모델이 모든 클래스를 동시에 예측하므로, YOLO 기본 추론 과정에 포함된 NMS (Non-Maximum Suppression) 외에 추가적인 후처리는 적용하지 않는다. 따라서 모델이 출력한 원본 탐지 결과를 그대로 최종 평가에 사용하였다.

반면, 그림 6은 모델 B의 구조를 나타낸다. 모델 B는 클래스 별로 각각 학습된 총 5개의 단일 클래스 YOLOv11n-OBB 모델의 출력을 결합하는 구조이다. 각 모델의 예측 결과에 대해 우선 NMS를 적용하여 중복 검출을 제거한 후, 모든 클래스

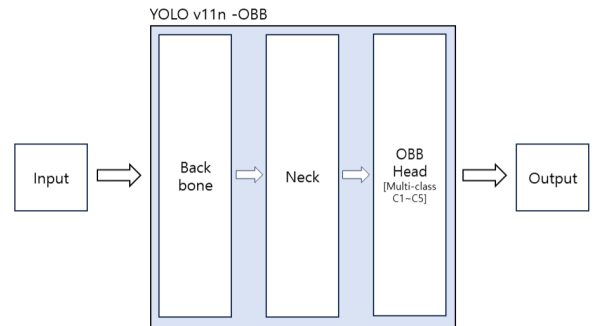


그림 5. 모델 A 구조도.

Fig. 5. Structure of the model A.

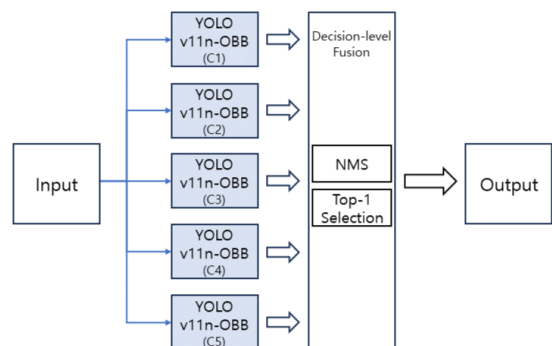


그림 6. 모델 B 구조도.

Fig. 6. Structure of the model B.

표 4. 학습 하이퍼 파라미터 설정.

Table 4. Training hyperparameter settings.

| 항목                    | 값                                    |
|-----------------------|--------------------------------------|
| 에폭                    | 200                                  |
| 입력 해상도                | 800×800                              |
| 배치 크기                 | 16                                   |
| 최적화 알고리즘              | AdamW                                |
| Momentum              | 0.9                                  |
| Weight decay          | 0.0005                               |
| Initial learning rate | 0.001667                             |
| LR Scheduler          | 3-epoch linear warmup + linear decay |
| Loss function         | Box + Cls + DFL + Angle              |
| Loss weights          | box=7.5, cls=0.5, dfl=1.5            |

출력 중 신뢰도(confidence) 값이 가장 높은 탐지 결과만을 최종 출력으로 선택하는 계열별(category-wise) Top-1 필터링을 수행하였다. 그 결과, 필터와 이음부 계열에서 각각 1개의 탐지 결과가 최종 출력된다. 이러한 후처리 전략은 실제 산업 환경에서 관심 영역(region of interest)을 통해 단일 용접 타겟이 존재하는 상황을 가정한 설계이다.

본 연구에서는 Ultralytics 기반 YOLOv11n-OBb 모델을 활용하였으며, 사전학습 가중치(yolo11n-obb.pt)를 초기 파라미터로 사용하여 학습을 수행하였다. 학습은 입력 해상도 800×800, 배치 크기(batch size) 16, 총 200 에폭(epoch) 동안 진행되었다. 최적화 알고리즘(optimizer)은 Ultralytics의 auto 설정에 따라 AdamW가 자동 선택되었으며, 초기 3 epoch 동안 linear warmup을 수행한 후 선형 감소(linear decay) 기반의 학습률 스케줄을 적용하였다. 그 외 세부 학습 하이퍼 파라미터는 표 4에 제시하였다.

IV. 모델 성능 비교 결과

본 장에서는 제안한 모델 구조의 타당성을 검증하기 위해 통제된 테스트셋에서의 정량적·정성적 성능 비교와 추론 속도 분석을 수행하였으며, 추가적으로 모델 크기별 및 세대별 구조 비교를 통해 모델 선정의 근거를 제시하였다. 또한 비통제 산업 환경에서 수집된 데이터에 대한 일반화 성능을 평가하였다.

학습·검증 단계에서의 F1·신뢰도 곡선 기반 최대 F1 점수를 비교한 결과, 모델 B의 단일 클래스 모델들이 개별적으로 더 높은 F1 점수를 보였다. 이는 단일 클래스 학습 구조의 특성상 예측 난이도가 낮기 때문으로 해석된다. 그러나 해당 결과는 학습 데이터 분포 내에서의 수렴 특성을 나타내는 지표이며, 실제 산업 적용 단계의 최종 결정 성능을 직접적으로 반영하지는 않는다. 이에 따라 모델 간 실질적 성능 비교는 통제된 테스트셋 및 비통제 산업 환경 데이터를 기반으로 수행하였다.

1. 모델 구조에 따른 성능 비교

본 절에서는 모델 A와 모델 B의 탐지 성능 및 추론 속도를 비교하여 구조적 차이에 따른 성능 특성을 평가하였다.

1.1 정량적 성능 비교 및 추론 속도 분석

탐지 성능은 IoU 기준 정밀도(precision), 재현율(recall), F1 점수 및 mAP50-95를 활용하여 평가하였으며, 추론 속도는

표 5. 모델 구조에 따른 종합 정량적 성능 비교.

Table 5. Overall quantitative performance comparison according to model structure.

| Metric          | A (단일 멀티클래스) | B (결정 수준) |
|-----------------|--------------|-----------|
| 정밀도 (IoU=0.5)   | 0.9895       | 0.9894    |
| 재현율 (IoU=0.5)   | 1.0000       | 0.9894    |
| F1 점수 (IoU=0.5) | 0.9947       | 0.9894    |
| mIoU (IoU=0.5)  | 0.9342       | 0.8935    |
| mAP50-95        | 0.9091       | 0.7841    |
| 회전 각도 오차(°)     | 1.2466       | 3.9642    |

동일한 NVIDIA GeForce RTX 4090 환경에서 파일 입출력 및 후처리를 포함한 end-to-end 실행 시간을 기준으로 FPS를 산출하였다. 그 결과, 모델 A는 55.00 FPS, 모델 B는 10.75 FPS로 약 5배의 처리 속도 차이를 보였다.

표 5에 따르면, 두 모델은 IoU 0.5 기준 정밀도에서 거의 동일한 수준(0.9895 vs. 0.9894)을 보였으나, 재현율, F1 점수, mIoU, mAP50-95 및 회전 각도 오차에서 모델 A가 전반적으로 우수하였다. 이는 모델 A가 탐지 정확도 및 위치·각도 정밀도 측면에서 상대적 우위를 가짐을 의미한다.

1.2 정성적 탐지 결과 비교

그림 7은 두 모델의 탐지 결과를 비교한 예시이다. 모델 A는 이음부 객체의 위치를 정확히 탐지한 반면, 모델 B는 동일한 신뢰도 값의 복수 후보 중 상단 객체를 선택하였다. 이는 결정 수준 구조의 특성상 동일 신뢰도 후보 존재 시 실제 용접 대상과 다른 객체가 선택될 가능성이 있음을 보여준다.

2. 모델 선정 근거 및 구조 비교 분석

본 절에서는 모델 크기별 및 세대별 비교를 통해 최종 모델 선정 근거를 제시하였다.

2.1 YOLOv11 모델 크기별 성능 비교

YOLOv11 모델 크기에 따른 성능 특성을 비교하기 위해 nano(n), small(s), large(l) 모델을 동일한 학습 조건(에폭=50)에서 학습하였고 결과는 표 6에 제시하였다.

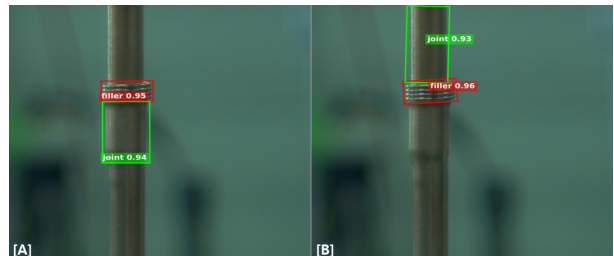


그림 7. 통제환경 이음부 탐지 사례 비교 (모델A: 정상, 모델B: 오검출).

Fig. 7. Controlled joint location comparison (ModelA: Normal, ModelB: Detection Error).

표 6. 동일 아키텍처, 크기별 성능 비교.

Table 6. Performance comparison by model size under the same architecture.

| Model       | mAP50-95 | FPS   |
|-------------|----------|-------|
| YOLOv11n-OB | 0.9800   | 268.1 |
| YOLOv11s-OB | 0.9834   | 245.6 |
| YOLOv11l-OB | 0.9839   | 129.2 |

mAP50-95 차이는 최대 0.0039로 제한적이었으나 추론 속도는 모델 크기에 따라 큰 차이를 보였다. YOLOv11n-OB는 268 FPS로 가장 빠른 처리 속도를 기록하였으며, 정확도 차이가 미미한 점을 고려하여 최종 모델로 선정하였다.

2.2 YOLOv8n-OB와 YOLOv11n-OB의 세대별 성능 비교

선정된 YOLOv11n-OB 모델의 채택 타당성을 검증하기 위해, 이전 세대 경량 모델인 YOLOv8n-OB와 동일한 학습 조건에서 비교 실험을 수행하였다. 추론 속도 측정 조건은 2.1 절과 동일하게 설정하였으며, 결과는 표 7에 제시하였다.

표 7에 따르면, 두 모델은 0.98 이상의 유사한 mAP50-95를 보였으며 정확도 차이는 0.0019로 미미하였다. YOLOv8n-OB가 약 13% 높은 FPS를 기록하였으나, 아키텍처 확장성과 운용 효율성을 고려하여 YOLOv11n-OB를 본 연구의 기본 모델로 채택하였다.

3. 비통제 산업 환경 조건에서의 일반화 성능 분석

통제된 테스트셋과 별도로, 실제 산업 생산 공정에서 비통제 조건으로 촬영된 412장의 이미지를 활용하여 일반화 성능을 평가하였다. 이는 통제 테스트셋(47장) 대비 약 8.8배 규모이며, 전체 5개 클래스 중 C2, C3, C4만 포함되었다(C2: 195개, C3: 114개, C4: 103개). 본 실험은 추가 학습(additional training)이나 파인튜닝(fine-tuning) 없이 기존 학습 모델을 그대로 적용하였으며, 정량적 결과는 표 8에 제시하였다.

표 8에 따르면, 모델 A는 정밀도(0.7815), F1 점수(0.7564), mAP50-95(0.5158) 지표에서 모델 B 대비 상대적으로 높은 값을 나타냈다. 이는 비통제 환경에서 전반적인 탐지 성능 유지 측면에서 우위를 보였음을 의미한다. 반면, 모델 B는 mIoU(0.9321)와 회전 각도 오차(4.74°)에서 더 우수한 결과를 나타냈다. 이는 검출된 객체에 한 해 위치 및 방향 추정 정밀도가 상대적으로 높았음을 시사한다. 즉, 모델 A는 전반적인 검출 성능에서, 모델 B는 개별 객체의 위치 및 각도 정밀도에서 상대적 강점을 가지는 것으로 분석된다.

정성적 분석 결과, 배경 내 파이프와 유사한 형상, 바코드 및 배선 구조 등으로 인한 오검출 사례가 일부 확인되었다. 모델 B는 결정 수준 기반 단일 선택 구조의 특성으로 인해 객체 이음부 위치를 오탐지하는 사례가 관찰되었으며, 그림 8은 해당 사례를 보여준다.

표 7. 모델 세대별 성능 비교.

Table 7. Performance comparison across model generations.

| Model       | mAP50-95 | FPS   |
|-------------|----------|-------|
| YOLOv11n-OB | 0.9800   | 268.1 |
| YOLOv8n-OB  | 0.9819   | 303.5 |

표 8. 비통제 산업 환경 조건에서의 종합 정량적 성능 비교.

Table 8. Overall quantitative performance comparison under uncontrolled industrial conditions.

| Metric          | A (단일 멀티클래스) | B (결정 수준) |
|-----------------|--------------|-----------|
| 정밀도 (IoU=0.5)   | 0.7815       | 0.7572    |
| 재현율 (IoU=0.5)   | 0.7328       | 0.7341    |
| F1 점수 (IoU=0.5) | 0.7564       | 0.7455    |
| mIoU (IoU=0.5)  | 0.8496       | 0.9321    |
| mAP50-95        | 0.5158       | 0.4477    |
| 회전 각도 오차(°)     | 6.1650       | 4.7439    |

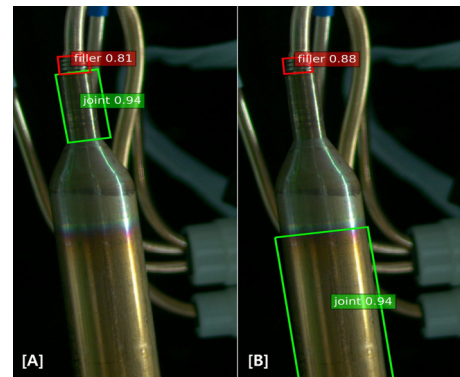


그림 8. 비통제 데이터의 이음부 탐지 사례 비교 (모델A: 정상, 모델B: 오검출).

Fig. 8. Uncontrolled joint location comparison (ModelA: Normal, ModelB: Detection Error).

통제된 실험 환경과 비교할 때 전반적인 성능은 상대적으로 감소하는 경향을 보였으며, 이는 학습 데이터 대비 실제 환경의 배경 복잡도 증가 및 객체 밀도 차이에 기인한 것으로 해석된다. 이러한 결과는 모델 구조 차이가 산업 환경에서의 일반화 성능에 직접적인 영향을 미칠 수 있음을 시사한다.

V. 결론

본 연구에서는 산업용 로봇 기반 동관 브레이징 공정에서 YOLOv11n-OB 기반 객체 탐지 모델의 구조적 차이에 따른 성능 특성을 비교·분석하였다. 단일 멀티클래스 구조의 모델 A와 클래스별 개별 학습 후 결정 수준에서 결합하는 모델 B를 동일 조건에서 실험적으로 평가하였다.

통제된 환경에서는 두 모델 모두 높은 성능을 보였으나, 모델 A가 전반적인 검출 안정성과 추론 효율 측면에서 우수한 특성을 나타냈다. 특히 실시간 공정 적용을 고려할 때 추론 속도에서의 차이는 구조 선택에 중요한 요소로 작용할 수 있다.

비통제 산업 환경 데이터 실험에서는 두 모델 모두 성능 저하가 관찰되었으나, 모델 A는 탐지 성능 유지 측면에서 상대적으로 안정적인 경향을 보였다. 반면, 모델 B는 검출 성공 객

체에 대한 위치 및 방향 추정 정밀도 측면에서 강점을 보였다.

종합적으로 볼 때, 실시간 자동화 공정과 같이 연속적인 처리 및 안정적 검출이 요구되는 산업 환경에서는 단일 멀티클래스 구조가 보다 적합한 구조임을 시사하였다. 향후 연구에서는 대규모 데이터 기반 학습 및 다양한 산업 환경 조건에서의 추가 검증을 통해 구조적 선택 기준을 보다 정교화할 필요가 있다.

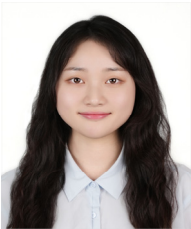
## REFERENCES

- [1] A. B. Pereira, J. M. S. Dias, J. P. Rios, N. M. Silva, S. Duraisamy, and A. Horovistiz, "Brazing of copper pipes for heat pump and refrigeration applications," *Metals*, vol. 14, no. 2, p. 171, Jan. 2024.  
doi: <https://doi.org/10.3390/met14020171>
- [2] E. A. Lemma, J. M. S. Dias, A. Bastos, B. Mascate, and A. Horovistiz, "Advances in induction brazing of copper and dissimilar metals: Challenges and emerging trends," *Journal of Advanced Joining Processes*, vol. 11, p. 100302, Jun. 2025.  
doi: <https://doi.org/10.1016/j.jajp.2025.100302>
- [3] J. Wang, L. Li, and P. Xu, "Visual sensing and depth perception for welding robots and their industrial applications," *Sensors*, vol. 23, no. 24, p. 9700, Dec. 2023.  
doi: <https://doi.org/10.3390/s23249700>
- [4] Y. Xu and Z. Wang, "Visual sensing technologies in robotic welding: Recent research developments and future interests," *Sensors and Actuators A: Physical*, vol. 320, p. 112551, Jan. 2021.  
doi: <https://doi.org/10.1016/j.sna.2021.112551>
- [5] M. Zhao, X. Liu, K. Wang, Z. Liu, Q. Dong, P. Wang, and Y. Su, "Welding seam tracking and inspection robot based on improved YOLOv8s-Seg model," *Sensors*, vol. 24, no. 14, p. 4690, Jul. 2024.  
doi: <https://doi.org/10.3390/s24144690>
- [6] M. Li, J. Huang, L. Xue, and R. Zhang, "A guidance system for robotic welding based on an improved YOLOv5 algorithm with a RealSense depth camera," *Scientific Reports*, vol. 13, p. 21299, Dec. 2023.  
doi: <https://doi.org/10.1038/s41598-023-48318-8>
- [7] A. Gao, Z. Fan, A. Li, Q. Le, D. Wu, and F. Du, "YOLO-weld: A modified YOLOv5-based weld feature detection network for extreme weld noise," *Sensors*, vol. 23, no. 12, p. 5640, Jun. 2023.  
doi: <https://doi.org/10.3390/s23125640>
- [8] C. Kim, S. Hwang, J. Chung, and H. Sohn, "Development of automatic crack detection technology in welded surface using laser active thermography and CNN deep learning," *Journal of the Korean Society for Nondestructive Testing (in Korean)*, vol. 40, no. 3, pp. 163–173, Jun. 2020.  
doi: <https://doi.org/10.7779/JKSNT.2020.40.3.163>
- [9] G. A. Elhendawy and Y. El-Taybany, "Machine vision-assisted welding defect detection system with convolutional neural networks," *International Journal of Precision Engineering and Manufacturing*, vol. 26, pp. 3185–3194, Jun. 2025.  
doi: <https://doi.org/10.1007/s12541-025-01281-y>
- [10] T. Deng, B. Fu, M. Liu, H. He, D. Fan, L. Li, L. Huang, and E. Gao, "Comparison of multi-class and fusion of multiple single-class SegNet model for mapping karst wetland vegetation using UAV images," *Scientific Reports*, vol. 12, p. 13270, Aug. 2022.  
doi: <https://doi.org/10.1038/s41598-022-17620-2>
- [11] W.-C. Lin, C.-C. Wang, M.-C. Tsai, C.-Y. Huang, C.-C. Lin, and M.-H. Tseng, "A YOLO ensemble framework for detection of Barrett's esophagus lesions in endoscopic images," *Diagnostics*, vol. 15, no. 18, p. 2290, Sep. 2025.  
doi: <https://doi.org/10.3390/diagnostics15182290>
- [12] W. Zhang and Y. Wang, "Multi-model fusion fine-grained image classification method based on migration learning," *IEEE Access*, vol. 12, pp. 31977–31987, Feb. 2024.  
doi: <https://doi.org/10.1109/ACCESS.2024.3369032>
- [13] D. Koguciuk, Ł. Chechliński, and T. El-Gaaly, "3D object recognition with ensemble learning—a study of point cloud-based deep learning models," in *Advances in Visual Computing: 14th International Symposium on Visual Computing, Proceedings, Lecture Notes in Computer Science*, Cham, Switzerland: Springer, pp. 100–114, Oct. 2019.  
doi: [https://doi.org/10.1007/978-3-030-33723-0\\_9](https://doi.org/10.1007/978-3-030-33723-0_9)
- [14] S. Y. Kim and C. H. Kang, "Performance analysis of Re-ID in YOLOv9 and BoT-SORT algorithms for multiple-object detection and tracking," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 31, no. 2, pp. 160–165, Feb. 2025.  
doi: <https://doi.org/10.5302/J.ICROS.2025.24.0234>
- [15] G. Lee, K. Kim, J. Lee, and H. Kim, "Real-time and integrated framework for LiDAR-based object tracking," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 31, no. 3, pp. 196–205, Mar. 2025.  
doi: <https://doi.org/10.5302/J.ICROS.2025.24.0284>
- [16] Q. Guo, Z. Yang, J. Xu, Y. Jiang, W. Wang, Z. Liu, W. Zhao, and Y. Sun, "Progress, challenges and trends on vision sensing technologies in automatic/intelligent robotic welding: State-of-the-art review," *Robotics and Computer-Integrated Manufacturing*, vol. 89, p. 102767, Oct. 2024.  
doi: <https://doi.org/10.1016/j.rcim.2024.102767>
- [17] T. Liu, P. Zheng, and J. Bao, "Deep learning-based welding image recognition: A comprehensive review," *Journal of Manufacturing Systems*, vol. 68, pp. 601–625, Jun. 2023.  
doi: <https://doi.org/10.1016/j.jmsy.2023.05.026>
- [18] T. Shehzadi, K. A. Hashmi, M. Liwicki, D. Stricker, and M. Z. Afzal, "Object detection with transformers: A review," *Sensors*, vol. 25, no. 19, p. 6025, Oct. 2025.  
doi: <https://doi.org/10.3390/s25196025>



**양수빈**

2022년 전남대 전자컴퓨터공학부 졸업.  
2025년~현재 한국생산기술연구원 자체  
인턴. 관심분야는 자동화 테스트 및 스마  
트 팩토리.



**강환희**

2023년 건국대 스마트운행체공학과 졸  
업. 2026년 UST 대학원 석사. 2024년~현  
재 한국생산기술연구원 서남기술실용화  
본부장실 UST학생. 관심분야는 자율주  
행 및 자율제조.



**김영곤**

1997년 전북대 금속공학과 졸업. 2001년,  
2006년 오사카대 대학원 석사/박사. 2012  
년~현재 한국생산기술연구원 목적기반  
모빌리티그룹 수석연구원. 관심분야는  
저항용접공정 및 열분포해석.



**김명진**

2017년 동아대학교 전자공학과 졸업. 2019  
년, 2024년 대구경북과학기술원(DGIST)  
로봇공학 및 메카트로닉스 공학 석사/  
박사. 현재 한국생산기술연구원 선임연  
구원. 관심분야는 생체모방 로봇 및 산업  
용 로봇.



**김유진**

2016년 나고야공업대학 기계과 졸업. 2018  
년, 2021년 나고야대학 대학원 석사/박  
사. 현재 한국생산기술연구원 목적기반  
모빌리티그룹 선임연구원. 관심분야는  
로봇 안전 및 응용, 자율주행, 인간 공학.