

논문 2024-19-01

카메라 기반 객체의 위치인식을 위한 왜곡제거 및 오검출 필터링 기법 (Distortion Removal and False Positive Filtering for Camera-based Object Position Estimation)

† 진 실, † 송 지 민, 최 지 호, 진 용 식, 정 재 진*, 이 상 준*
(Sil Jin, Jimin Song, Jiho Choi, Yongsik Jin, Jae Jin Jeong, Sang Jun Lee)

Abstract : Robotic arms have been widely utilized in various labor-intensive industries such as manufacturing, agriculture, and food services, contributing to increasing productivity. In the development of industrial robotic arms, camera sensors have many advantages due to their cost-effectiveness and small sizes. However, estimating object positions is a challenging problem, and it critically affects to the robustness of object manipulation functions. This paper proposes a method for estimating the 3D positions of objects, and it is applied to a pick-and-place task. A deep learning model is utilized to detect 2D bounding boxes in the image plane, and the pinhole camera model is employed to compute the object positions. To improve the robustness of measuring the 3D positions of objects, we analyze the effect of lens distortion and introduce a false positive filtering process. Experiments were conducted on a real-world scenario for moving medicine bottles by using a camera-based manipulator. Experimental results demonstrated that the distortion removal and false positive filtering are effective to improve the position estimation precision and the manipulation success rate.

Keywords : Deep learning, Computer vision, Manipulator robot, Distance estimation, Object detection

1. 서 론

로봇팔은 다양한 산업에서 여러 형태와 구성요소로 설치되어 반복적인 작업수행 자동화를 위한 사용이 늘어나는 추세이다. 제조업에서는 제품의 생산, 조립, 포장 그리고 운반 등의 영역에 활용하여 생산성을 높이고 인력 비용을 절감하는데 도움을 준다. 또한 의료 분야에서는 정밀하고 정확한 제어기능으로 수술 보조를 수행하여 의사의 부담을 완화하고 환자의 안전을 보장하는데 이바지하고 있다. 농업에서도 작물 수확, 식물 관리 등 작물 생산과 관리작업을 부분 자동화하여 농작업의 효율성을 향상시키고 있다. 이 외에도 로봇팔은 요식업에서도 조리과정에 사용되거나 주문처리 및 서빙에 사용되며 일관된 품질과 정확성으로 인력을 보조하거나 대체함으로써 현대의 다양한 산업 분야에서 그 효용성을

을 나타내고 있다. 이처럼 로봇팔이 쓰이는 산업 분야, 작업의 복잡성, 요구되는 정확도 등에 따라 로봇팔의 형태와 구성을 달리하여 우리 삶에서 사람의 반복적인 노동 작업을 줄이는 데에 유용하게 사용되고 있다. 이러한 로봇팔 시스템은 센서, 모터, 제어시스템, 인공지능 등 다양한 기술들로 이루어져 있으며 컴퓨터 프로그래밍을 통하여 움직임과 작업을 제어할 수 있다.

카메라 센서는 저렴한 단가와 작은 크기에서 오는 다양한 이점으로 산업용 로봇팔 구축에 빈번하게 사용된다. 또한 카메라 센서는 다른 센서들에 비해서 높은 공간분해능을 가지기 때문에 관심 객체의 정밀한 2차원 위치정보를 알 수 있다는 장점이 있다. 이와 같은 장점을 실제 산업에 적용하기 위해서 카메라 기반의 물체 인식에 대한 다양한 연구가 진행되고 있다. Friedrich et al. [1]은 로봇이 실시간으로 감지되는 경로를 탐색할 때 정밀한 제어를 위해 보정된 시스템을 갖는 것의 중요성을 강조한다. 해당 연구에서는 관심 객체에 카메라 센서의 초점을 맞추며 로봇팔이 객체를 따라 이동해야 할 때 필요한 기술을 제안한다. 이동 경로를 인식하는 센서로 카메라가 사용될 때 정확한 경로 추정을 위해서 카메라 캘리브레이션을 통해 얻는 내부 파라미터를 로봇팔 제어에 실시간으로 적용하는 시스템을 구축했다. Park et al. [2]은 자율주행을 위한 동적 객체 인식 방법을 연구했다. 이와 같은 도로위의 객체검출 기술은 차량 운행 중 장애물을 인지하고 회피하는 기술 구현에 활용할 수 있다. Hwang et al. [3]은 차량의 긴급 제동 장치에 활용될 수 있는 딥러닝 기반의 객체 검출 및 거리 추정을 위한 기술을 연구했다.

*Corresponding Author (jeong@kumoh.ac.kr, sj.lee@jbnu.ac.kr)

† Equal contribution

Received: Oct. 20, 2023, Revised: Nov. 9, 2023, Accepted: Nov. 15, 2023.

S. Jin: Jeonbuk National University (Undergraduate Student)

J. M. Song: Jeonbuk National University (M.S. Student)

J. Choi: Jeonbuk National University (Ph.D. Student)

Y. Jin: ETRI (Researcher)

J. J. Jeong: Kumoh National Institute of Technology (Assist. Prof.)

S. J. Lee: Jeonbuk National University (Assist. Prof.)

* This work was supported by Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government [23ZD1130, Regional Industry ICT Convergence Technology Advancement and Support Project in Daegu-GyeongBuk (Robot)]. 본 과제 (결과물)는 교육부와 한국연구재단의 지원으로 지원을 받아 수행된 사회맞춤형 산학협력 선도대학 (LINC+) 육성사업의 연구결과입니다. 이 연구는 2023년도 산업통상자원부 및 산업기술평가관리원 (KEIT) 연구비 지원에 의한 연구임 (20023305).

위의 연구들은 공통적으로 왜곡 보정의 중요성을 시사하고 있다. 본 논문에서는 카메라 캘리브레이션을 통해서 추정된 왜곡 계수를 반영한 왜곡 보정이 객체 검출 및 거리추정 성능에 미치는 영향을 실험하고 분석했다.

카메라 영상 기반의 딥러닝 기술은 다양한 분야에서 활발한 응용 연구가 진행되는 추세이다. 특히 로봇팔의 시스템에 딥러닝 기술을 접목하여 제품의 결함 검출과 검사로 불량제품의 발생률을 낮추거나 [4], 딥러닝 학습을 통해 로봇의 이동 경로나 작업방식을 최적화 할 수 있다 [5]. 본 논문에서는 객체를 검출하기 위하여 딥뉴럴 네트워크를 사용하였다. 딥러닝 객체 검출 연구에 활용되는 대표적인 알고리즘은 R-CNN [6], SSD [7] 그리고 YOLO [8] 등이 있다. 영역 별 Convolutional Neural Network (CNN) 기반의 분류기를 적용하는 R-CNN [6]은 객체 검출을 위한 주요 딥러닝 알고리즘 중 하나이다. R-CNN은 입력 이미지에서 객체의 경계 상자와 해당 객체의 클래스를 예측하는 데 사용된다. R-CNN은 객체 분류와 경계 박스 추정 두 가지 단계를 나누어 제안하는 2-stage detector이고, 각 후보 영역을 독립적으로 처리하므로 연산량이 많고 느린 단점을 가진다. SSD [7]는 Single Shot Multi Box Detector의 약자로, 다수의 객체를 실시간으로 탐지하는 데 유용한 1-stage detector 모델 중 하나이다. R-CNN과는 달리 객체 분류와 경계박스 검출을 동시에 수행함으로써 객체를 탐지하기 때문에 상대적으로 빠른 추론속도를 보이는 특징이 있다. YOLO [8]는 입력 이미지를 여러 크기의 그리드 셀로 분할하고, 경계박스의 후보군인 셀들에서 Non-Maximum Suppression (NMS)을 통해서 객체당 하나의 경계박스를 선택하는 1-stage detector 모델이다. 본 논문에서는 연산자원의 효율적인 사용이 요구될 수 있음을 고려하여 매니플레이션 시스템에 1-stage detector 모델인 YOLO [8] 기반의 모델을 활용하여 실험을 진행하였다.

딥러닝 모델 자체의 추론 정확도를 향상시키기 위한 연구와는 별개로 추론 결과에 대한 후처리 알고리즘을 적용하여 성능을 향상시키기 위한 연구도 진행되고 있다. Kim et al. [9]은 적외선 카메라 영상을 RGB 영상으로 변환하는 색상화의 성능 향상을 위한 후처리 방법을 연구했다. Jung et al. [10]은 카테고리 정보 추정 정확도를 향상시키기 위해서 검출한 객체 사이의 관계를 고려하는 후처리 알고리즘을 제안했다. Conditional Random Field (CRF) 기반의 후처리 알고리즘을 통해 객체의 확률을 계산하고 가장 높은 값을 갖는 카테고리 정보로 변경한다. 하지만 제안한 알고리즘은 제대로 검출된 경우(True Positive)도 카테고리 정보를 변경하여 실질적으로 큰 성능향상을 보이지 않았다. 본 연구에서는 2차원 영상으로부터 객체의 3차원 정보를 계산하고 알고리즘은 객체의 형태와 비교하여 오검출 (False Positive)을 제거하는 후처리 기법을 제안한다. 제안하는 알고리즘은 정검출된 객체는 유지하는 동시에 카메라 기하학에 기반하여 잘못 검출된 객체들을 필터링 함으로써 적은 연산량으로 검출 정확도를 개선할 수 있다.

영상에서 2차원 객체 검출 후 다양한 종류의 센서 정보를

기반으로 3차원 위치를 알아낼 수 있다. 대표적인 센서의 종류로 라이다 센서, GPS 센서, 카메라 센서 등이 있다. 라이다는 빠른 반응 속도와 정밀한 거리 측정이 가능하지만, 카메라나 GPS 센서에 비해 고가로 시스템 구축에 많은 비용이 든다는 단점이 있다. GPS 센서는 넓은 영역에서 위치 추적이 가능하며 비용이 저렴하지만, 실내와 같은 신호 차단이 있는 환경에서는 정확도가 떨어지고 센서의 반응 속도가 느려 위치 업데이트 속도가 더딜 수 있다. 본 논문에서는 스마트 팩토리 내에서 객체 검출을 시행하는 실험 환경을 가정하므로 라이다 센서와 GPS 센서가 갖는 비용적 측면의 한계와 실내에서 사용하지 못한다는 한계를 보완할 수 있는 카메라 센서를 사용한다. 카메라 센서는 비용이 비교적 저렴하며 실내에서도 사용 가능하고, 반응 속도도 빨라 데이터 처리에 용이한 장점을 가지고 있다. 또한, 센서의 크기가 소형으로 로봇팔 적용에 적합하다. 최근에는 자율 주행 및 로봇의 주변 환경 인지 연구에서 주로 사용되며 중요성이 증가하고 있어 다양한 분야에서 연구되고 있다.

카메라를 이용하여 얻은 영상에서 물체의 거리를 추정하는 연구는 다양한 방법으로 제시되고 있다. He et al. [11]은 단일 카메라로 얻은 영상 속 물체 위치 추정 과정에서 고정 초점 길이 데이터셋으로부터 가변 초점 길이 데이터셋을 생성하는 새로운 딥 뉴럴 네트워크를 제안하였다. Moises et al. [12]은 색상정보를 활용한 세그멘테이션과 퍼지 기반의 군집화 기술을 결합하여 교통신호등을 감지한 뒤, 베이지안 필터를 교통 신호등에 적용하여 자율 주행 차량과의 거리를 추정한다.

본 논문의 연구 목적은 특정 물체에 최적화된 딥러닝 기반 객체 검출 및 거리 추정 알고리즘을 개발하여 제조 공정에 사용될 로봇팔의 매니플레이션 성능을 향상하는 것이다. 본 논문의 전체적인 실험 개요는 그림 1로 설명할 수 있다. 먼저, 객체 검출 속도를 우선 시 하여 채택한 딥러닝 모델 YOLOv5를 학습시킨다. 둘째로, 본 논문에서 제안하는 후처리 방법, False Positive Filtering (FPF)를 통해 객체 검출 성능을 높여 실시간 객체 탐지를 진행한다. 마지막으로, 이미지 왜곡을 보정하는 전처리를 거친 후 핀홀 카메라 모델을 가정한 기하학 정보를 이용하여 단일 카메라로 획득한 영상에서 물체 거리 추정 방법을 제안한다. 제안하는 방법의 유효성을 보이기 위해서 실제 로봇팔을 활용하여 추정된 물체와의 거리를 기반으로 매니플레이션 실험을 수행했다.

II. 딥러닝 기반 객체의 위치인식 기법

1. 데이터셋 수집 방법

본 논문에서는 약병의 거리 추정 알고리즘을 개발하고 로봇팔을 활용한 매니플레이션 실험을 위하여 실제 데이터셋을 수집 및 레이블링 하였다. 카메라 기반 매니플레이터 개발에 있어서 카메라의 위치는 크게 로봇팔 외부와 그리퍼 위치에 구성할 수 있다. 카메라를 로봇팔 외부에 설치할 경우 로봇 좌표계와 카메라 좌표계사이의 캘리브레이션을 필

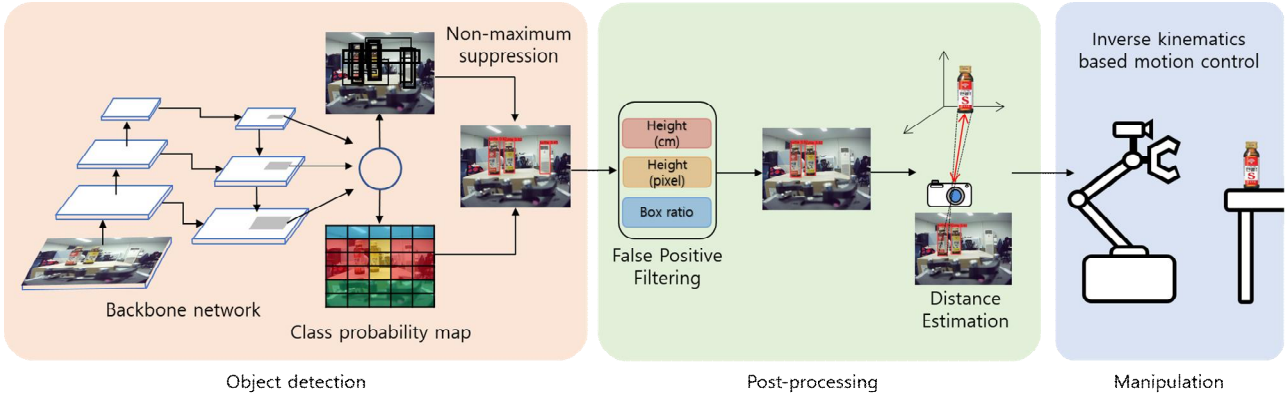


그림 1. 제안하는 알고리즘의 개요
Fig. 1. Overview of the proposed method

요로 하기 때문에, 본 논문에서는 그리퍼 좌표계의 원점으로부터 약 3cm 위의 위치에 카메라를 고정하여 실험을 수행하였다. 로봇팔에 고정된 카메라로부터 총 700장의 RGB 영상을 획득하였으며 각 영상에 대하여 객체 위치를 레이블링 하여 객체검출모델 훈련에 필요한 데이터셋을 수집하였다. 약병 검출을 위한 딥러닝 모델의 훈련을 위하여 20가지의 서로 다른 배경을 가진 상황에서 영상데이터를 획득하였다. 상황마다 35장의 영상 데이터를 촬영하였고, 객체 검출뿐만 아니라 거리 추정의 ground truth를 얻기 위해 각 35장의 상황마다 미리 지정해 놓은 순서대로 약병이 놓이는 위치와 개수를 정하여 촬영하였다. 영상의 크기는 640 x 480 픽셀 크기로 하나의 영상에는 다양한 개수의 약병이 나타나도록 실험 환경을 구성하였다. 각 RGB 영상은 최대 3개의 약병을 포함하고 있으며 각 약병 객체에 대하여 2차원 경계 상자 (Bounding Box)의 형태로 객체 위치에 대한 정보를 기록하였다. 경계 상자의 정보는 x, y, w, h 총 네 가지 정보를 포함한다. 여기서 x와 y는 경계 상자의 중앙 위치를 가리키는 픽셀 값이고 w와 h는 경계 상자의 픽셀단위 너비와 높이를 의미한다.

2. 2차원 객체검출을 위한 딥러닝 모델

본 연구에서는 영상에서의 2차원 객체 검출을 위하여 YOLOv5 기반의 딥러닝 모델을 사용하였다. 객체 검출 학습에 사용한 모델 YOLOv5는 크게 backbone, neck 그리고 head로 구성된다. Backbone 네트워크에서는 입력으로 들어온 영상으로부터 특징 맵을 추출하는데, 여러 가지 backbone 네트워크들 중 CSPDarkNet을 사용하여 파라미터 수를 감소시켰다.

Common Objects in Context (COCO) 데이터셋에서 미리 학습된 모델 파라미터 (pre-trained parameter)를 기반으로 실제 데이터셋에 대하여 훈련함으로써 객체의 매니플레이션이 가능한 정도의 정확도를 달성하는 동시에 빠른 속도로 실시간 추론이 가능한 경량 딥러닝 모델을 개발하였다. YOLOv5 기반의 딥러닝 모델의 종류에는 모델의 크기에 따라 YOLOv5s, YOLOv5m, YOLOv5l 그리고 YOLOv5x의 총

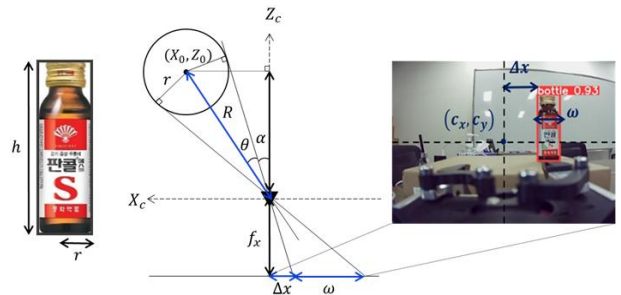


그림 2. 단일영상 기반 거리추정 기법
Fig. 2. Distance Estimation with Single-View Geometry

4가지로 구성된다. 해당 종류는 depth multiple과 width multiple의 두 가지 변수에 의하여 결정되고, 여러 가지 모델들 중 YOLOv5s 모델은 두 가지 변수가 모두 가장 작은 모델로 모델의 크기가 작아짐에 따라서 객체 검출 속도가 빠르다는 장점을 가진다. 본 실험에서는 약병 객체의 매니플레이션이 가능한 정도의 정확도를 달성하는 동시에 실시간 처리가 가능한 기술을 개발하기 위하여 YOLOv5s 모델을 채택하고 학습에 사용하였다.

3. 거리 계산을 통한 객체의 3차원 위치 추정

본 논문에서는 2차원 객체검출 결과를 기반으로 약병의 3차원 위치를 계산하기 위하여 카메라 기하학 기반의 연산방법을 사용하였다. 그림 2는 카메라와 약병의 위치를 위에서 내려다본 개념도를 보여준다. 카메라에서 약병 중심까지의 거리인 R을 계산하는 과정에서는 실제 약병의 물리적인 반지름 길이인 r과 2차원 영상에서 알 수 있는 경계 상자 픽셀너비 (w), xc 축 방향의 주점인 cx로부터 경계 상자까지의 거리 (Δx), 카메라 캘리브레이션을 통해 알아낸 카메라 내부 파라미터 중 xc 축 방향의 초점 거리 (fx)를 이용한다. 우선 fx, w, Δx의 값에 대하여 삼각함수 수학적식을 기반으로 α와 θ의 값을 계산할 수 있다. 수식 (1)을 활용하여 α의 값을 먼저 계산할 수 있으며, fx, w, Δx의 값들과 수식

(1)에서 계산한 α 값을 기반으로 수식 (2)를 촬영하여 θ 의 값을 계산할 수 있다.

$$f_x \times \tan\alpha = \Delta x. \quad (1)$$

$$f_x \times \tan(2\theta + \alpha) = \omega + \Delta x. \quad (2)$$

수식 (2)에서 계산된 θ 를 활용하여 카메라 원점으로부터 약병 중심까지의 거리 R 을 계산할 수 있으며 그 과정은 아래의 수식으로 표현할 수 있다.

$$R = r / \sin\theta. \quad (3)$$

카메라 초점거리 f_x 및 측정값인 w , Δx 의 픽셀거리들을 기반으로 계산한 α , θ , R 의 값을 이용하여 카메라 좌표계에서 약병의 3차원 위치를 추정할 수 있으며, 그 과정은 아래 수식과 같다.

$$X_C = R \times \sin(\theta + \alpha). \quad (4)$$

$$Z_C = R \times \cos(\theta + \alpha). \quad (5)$$

실험환경에서 약병들은 매니플레이터와 동일한 높이에 위치하기 때문에, 카메라 좌표계에서 아래쪽 방향을 의미하는 Y_C 의 값은 알고 있다고 가정한다. 카메라 좌표계를 중심으로 3차원 위치를 계산하였기 때문에 해당 위치에 로봇팔이 도달할 수 있도록 카메라 좌표계 기준의 약병의 3차원 좌표를 로봇 좌표계로 변환 한다. 카메라 좌표계와 로봇 좌표계 사이의 관계는 로봇의 하드웨어 스펙정보에 기반한 유클리디안 변환으로 계산할 수 있다.

본 연구에서는 단일 영상 기반 관심객체 위치추정의 정밀도를 개선하기 위하여 렌즈의 왜곡계수를 고려한 알고리즘을 개발하였다. 카메라 캘리브레이션 과정에서 왜곡계수를 미리 계산하였으며, 객체까지의 거리를 추정하기 전에 실제 거리에 근사한 추정을 할 수 있도록 영상의 왜곡을 보정하는 전처리를 진행하였다. 카메라 캘리브레이션을 통해 얻은 왜곡계수를 기반으로 전처리 함으로써 실험적으로 객체까지의 거리추정 정확도 및 매니플레이션 성공률을 개선할 수 있었다.

4. 매니플레이션 정확도 개선을 위한 후처리 기술

영상에서 2차원 객체검출의 정확도를 개선하기 위하여 오검출 필터링 기법을 도입하였다. 본 논문에서 제안하는 오검출 필터링 기법은 관심객체를 정면에서 바라봤을 때 핀홀 카메라 모델의 기하학을 기반으로 한다. 그림 3은 오검출을 필터링 하기 위한 후처리 과정을 보여준다. 후처리 방법에서는 옆에서 바라본 핀홀 카메라 모델 수식을 활용하였다. 후처리 방법에 사용된 수식은 영상에서의 경계 상자 크기와 실제 약병의 크기가 비례한다는 점을 기반으로 한다. 객체 검출 시 약병이 아닌 객체를 검출하는 오검출을 거르기 위해 영상에서의 경계 상자 정보를 활용하여 두 가지 단계로 구성된 후처리를 진행하였다. 우선 약병이 놓일 수 없는 위치에 약병이 있다고 검출하는 오검출을 거르기 위하여 약병이 놓이는 실제 높이 H 를 이용하였다. 이때 H 는 카메라 좌

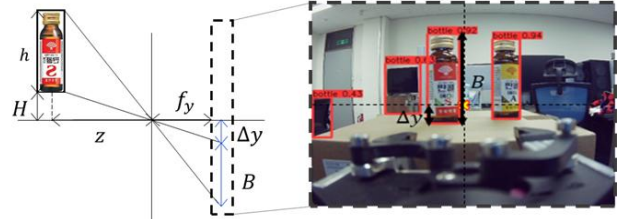


그림 3. 오검출 필터링을 위한 후처리 기법
Fig. 3. Post-processing for false positive filtering

표계 Z_C 축으로부터의 높이이다.

$$H = (Z \times \Delta y) / f_y. \quad (6)$$

실제 높이 H 를 측정하는 수식 (6)의 픽셀높이 Δy 는 물리적으로 같은 높이에 약병이 위치해 있어도 카메라와의 거리에 따라 다른 크기로 나타난다. 예를 들어, 동일한 H 높이에 있는 물체일지라도 카메라에서 가까운 거리에 위치할수록 픽셀높이 Δy 가 커지는 경향을 보인다. 본 논문에서 제안하는 후처리 기법에서는 약병이 놓일 수 있는 H 의 최대, 최소 범위를 지정하여 해당 범위에 속하지 않으면 오검출로 간주하고 필터링 하였다. 그림 3에서의 h 와 B 는 각각 약병의 실제 높이와 영상에서 검출된 약병의 경계 상자 픽셀높이를 의미한다. 후처리의 두 번째 단계에서는 경계 상자의 높이와 너비 비율을 활용하여 오검출을 필터링 한다. 카메라로부터 약병의 거리가 최대일 때와 최소일 때를 기준으로 경계 상자의 높이와 너비와 높이의 비율 값인 상자 비율의 최대, 최소를 각각 구하여 해당 범위에 속하지 않는 경계 상자는 오검출로 필터링 하였다. 두 단계의 후처리 알고리즘을 결합하여 실제 높이 H , 경계 상자 높이, 경계 상자 비율의 지정 범위에 속하지 않는 경계 상자는 오검출로 간주하여 관심 객체의 후보군에서 제외하는 알고리즘을 본 논문에서는 False Positive Filtering의 약자로 FPF 알고리즘이라 부른다. 그림 4에서는 딥러닝 모델로 검출한 객체에 FPF 알고리즘을 적용함으로써 처리되는 과정을 보여준다. 예를 들어, 첫 번째 예시에서는 관심객체와 비슷한 시각특징을 가진 비타민 통이 오검출 되었는데 본 논문에서 제안하는 FPF를 통하여 오검출이 올바르게 필터링 된 것을 확인할 수 있다. 또한, 다른 예시들에서도 딥러닝 모델에 의하여 잘못 검출된 배경 영역이 FPF를 통하여 제거되는 것을 확인하였다. 올바르게 검출한 결과를 유지하는 동시에 오검출을 제거함으로써 객체검출의 정확도를 개선할 수 있었다.

III. 실험

1. 실험 환경

객체 검출 실험은 Nvidia GeForce RTX 3070 GPU를 포함한 하드웨어 환경에서 수행하였다. 딥러닝 모델의 훈련과정에서는 객체의 클래스 분류를 위한 크로스 엔트로피 손실함수와 경계 상자의 회귀분석을 위한 중첩 영역의 비율



그림 4. FPF 알고리즘 사용 전후 비교

Fig. 4. Comparison detection results before and after applying the FPF algorithm

(intersection over union) 기반의 손실함수를 사용하였다. 훈련 과정에서는 COCO 데이터셋에 대하여 훈련된 모델 파라미터를 전이학습 하였으며, 딥러닝 모델은 stochastic gradient descent의 방법으로 최적화 하였다. 딥러닝 모델의 훈련을 위하여 하이퍼 파라미터는 배치크기 8, 학습 epoch는 100으로 설정하여 실험하였다.

2. 평가지표

본 논문의 실험에 사용된 평가지표로는 객체 검출 성능을 확인하기 위한 정밀도-재현율 곡선 (Precision-Recall curve) 과 mAP50-95를 사용하였다. 거리 추정 성능에는 MAE (mean absolute error, 평균절대오차)와 MAPE (mean absolute percentage error, 평균 절대 백분율 오차)를 사

용하였고, 매니플레이션의 성공률 측정에는 manipulation accuracy로 직접 만든 평가지표를 사용하며 약자 MA로 표기 하였다.

우선 객체 검출 성능에 사용된 Precision-Recall curve는 분류 모델의 임계값을 변경함에 따라 정밀도와 재현율 간의 상충 관계 (trade-off)를 보여주므로 모델 성능을 전반적으로 평가하는데 도움을 준다. Precision-Recall curve의 곡선 아래 면적을 기반으로 평균정밀도 (mean average precision)를 계산할 수 있다. 평균정밀도는 예측된 객체와 실제 객체 사이의 겹치는 영역인 IoU (Intersection over Union) 임계값을 0부터 1까지 변화시키면서 정밀도 값을 평균 내어 계산한 것이다. mAP50-95는 IoU 임계값을 0.5부터 0.95까지 변화시키면서 평균정밀도를 계산한 값이다. 학습한 객체 검출 모델이 일관된 성능을 보이는지 확인하기 위해 IoU 0.5-0.95까지의 범위에서 모델 성능을 평가한 mAP50-95도 함께 측정했다. 거리 추정 성능을 평가하기 위해 사용한 Mean Absolute Error (MAE)는 예측값과 실제 값 간의 절대적인 오차를 측정하여 평균적인 차이를 나타내고, Mean Absolute Percentage Error (MAPE)는 예측값과 실제 값 간의 상대적인 오차를 측정하여 평균적인 백분율 차이를 나타낸다. 본 실험에서는 거리 추정이 얼마나 정확하고 상대적으로 안정적인 수행을 하는지 포괄적인 정보를 얻기 위해 두 측정지표를 함께 고려하였다. 마지막으로 매니플레이션 성능을 측정하기 위해 사용된 Manipulation accuracy (MA)는 직접 측정기준을 세운 평가지표이다. 측정기준은 로봇 팔이 거리 추정을 잘못하여 약병에 닿지 못하거나 약병이 아닌 곳에 가서 매니플레이션을 하려는 동작을 취하면 실패로 간주하여 매니플레이션의 성공과 실패를 측정한다. 측정이 끝난 뒤 성공 백분율을 나타내는 평가지표이다.

3. 실험 방법

본 논문에서는 객체 검출, 거리 추정, 매니플레이션 성능을 측정하기 위해 두 가지 종류의 실험을 진행했다. 첫 번째 실험은 객체 검출과 거리 추정 성능을 측정하기 위한 실험이다. 영상에 나타난 약병의 개수와 위치는 모두 다르지만 거리 추정 성능을 평가하기 위해 사전에 지정해 둔 자리에 약병을 놓고 영상을 획득하는 방법을 사용하여 거리 실측 데이터 (ground truth data)를 획득하였다. 두 번째 실험은 매니플레이션 성능을 측정하기 위한 실험으로 FPF 알고리즘 적용여부와 이미지 왜곡 보정 여부에 따른 성능 비교 실험을 수행하였다. FPF 알고리즘과 왜곡 보정의 효과를 확인하기 위한 실험에서는 각각 총 150번의 매니플레이션 시행에 따른 결과를 나타내고 있다. 네 가지 실험 모두 150번 매니플레이션 시행 중 장애물의 위치, 약병의 거리는 미리 지정해 둔 위치에 놓아 실험 환경을 같도록 설정하였다.

4. 실험 결과

본 실험에서는 약병을 검출하고 거리 추정 과정을 통해 매니플레이션 하는 과정을 다루고 있다. 최종 매니플레이션

표 1. FPF 여부에 따른 객체 검출 성능 결과
Table 1. Effectiveness of FPF for 2D object detection

FPF	Precision (%)	Recall (%)	mAP (%)	mAP 50-95 (%)
-	88.2	99.7	98.4	89.4
✓	99.7	98.2	99.0	91.3

표 2. 이미지 왜곡 보정 여부에 따른 거리추정 성능 결과
Table 2. Effectiveness of image distortion correction

Distortion correction	MAE (mm)	MAPE (%)
-	30.04	9.07
✓	19.50	6.13

성능을 높이기 위해 두 가지 방법을 제안한다. 첫 번째는 FPF 알고리즘을 통해 약병 검출 성능을 높여 최종 매니플레이션 성능이 높아지는 것을 기대하는 방법이고, 두 번째는 이미지 왜곡을 보정하여 더 정확한 거리 추정을 통해 매니플레이션 성능을 높이는 것이다.

먼저 최종 매니플레이션 성능을 높이는 첫 번째 방법인 FPF 알고리즘 사용 여부에 따른 약병 검출 성능을 표 1에서 보여준다. 오검출을 필터링하는 FPF 알고리즘을 적용하여 약병 검출 성능을 측정할 때 사용한 평가지표는 Precision-Recall curve, mAP50-95이다. Precision-Recall curve의 성능은 곡선 아래 면적을 수치화한 mAP로 나타내었다. FPF 알고리즘을 적용하지 않았을 때는 Precision-Recall curve 측면의 정확도가 98.4%로 측정되었고, mAP50-95는 89.4%였다. FPF 알고리즘을 적용했을 때 약병 검출 성능으로 Precision-Recall curve의 정확도는 99.0%, mAP50-95는 91.3%로 FPF 알고리즘을 적용하지 않았을 때보다 성능이 개선된 것을 관찰할 수 있었다. 이는 FPF 알고리즘을 적용했을 때 객체 검출 성능이 올라감으로 로봇팔의 매니플레이션 관찰 시 약병이 없는 곳에 매니플레이션을 수행 할 확률을 낮출 것으로 기대할 수 있다.

최종 매니플레이션 성능을 높이는 두 번째 방법으로 이미지 왜곡 보정을 통한 거리 추정 성능 결과를 표 2에서 나타낸다. 평가지표로는 MAE와 MAPE를 사용하였다. 왜곡된 이미지에서 거리 추정 결과 MAE는 30.04mm, MAPE는 9.07%로 측정되었고, 카메라 캘리브레이션을 통한 이미지 영상의 왜곡 보정 후 거리 추정 결과 MAE는 19.50mm, MAPE는 6.13%로 성능이 향상된 것을 확인할 수 있다. 따라서 이미지 왜곡 보정 시 더 정확한 거리 추정으로 로봇이 약병을 집을 수 있는 확률이 높아질 것을 기대할 수 있다.

FPF 알고리즘과 이미지 왜곡 보정을 통한 객체 검출과 거리 추정 성능 향상이 최종적인 약병 매니플레이션에 영향을 끼치는지 보기 위해 아래 표 3과 같이 FPF 알고리즘 사용 여부와 이미지 왜곡 보정 여부에 따른 4가지 비교 실험을 통해 매니플레이션 성능을 관찰하였다. 해당 실험에 사용된 평가지표는 로봇팔의 엔드 이펙터가 약병에 정확히 위

표 3. FPF 과정와 이미지 왜곡 보정 여부에 따른 매니플레이션 성능 비교

Table 3. Ablation study for manipulation accuracy with and without FPF and image distortion correction

FPF	Distortion correction	MAE (mm)	MAPE (%)	MA (%)
-	-	17.60	5.08	73.33
✓	-	19.56	5.76	76.66
-	✓	15.70	4.54	75.33
✓	✓	11.58	3.23	83.33

치 하는지 관찰하기 위해 거리 추정 결과 관측 시 사용했던 MAE, MAPE와 매니플레이션 성공 백분율을 나타내는 MA이다. 실험의 기본 모델은 FPF 알고리즘과 이미지 왜곡 보정을 하지 않은 경우로 4가지 실험 중 기본 모델이 MA, MAE, MAPE 모두 가장 낮게 나온 것을 확인할 수 있다. 차례로 기본 모델에서 FPF 알고리즘만 사용하였을 때와 기본 모델에서 이미지 왜곡 보정만 사용했을 때를 비교해 보면 FPF 알고리즘만 사용했을 때 MA 성능이 이미지 왜곡 보정만 사용한 실험보다 높지만, MAE와 MAPE 성능은 더 낮은 것을 볼 수 있다. 해당 실험 결과를 통하여 FPF 알고리즘이 이미지 왜곡 보정을 하는 것보다 MA에 직접적인 영향을 주는 것을 알 수 있다. 또한, 이미지 왜곡 보정을 통해 기본 모델에서 거리 오차로 매니플레이션 하지 못했던 횟수를 줄여 기본 모델보다 MA 성능이 향상되는 것을 확인할 수 있다. 따라서 마지막 실험처럼 FPF 알고리즘과 이미지 왜곡 보정을 모두 사용한 모델이 MA, MAE 그리고 MAPE 모두 가장 좋은 성능을 나타내는 것을 확인할 수 있다. 결과적으로 매니플레이션 성능을 향상시키기 위해서는 객체 검출 성능을 향상시키는 FPF 알고리즘과 객체의 거리 추정 성능을 향상시키는 이미지 왜곡 보정이 적용되는 것이 유의미하다.

IV. 결론

본 논문에서는 2차원 객체 검출 알고리즘 및 3차원 위치 추정을 이용하여 로봇팔로 물체를 잡는 알고리즘을 제안하였다. 영상에서 객체의 픽셀위치를 검출 과정에서 영상의 왜곡을 보정하는 전처리 과정을 제안하여 더욱 정확한 거리 정보를 추정할 수 있었다. 객체 검출 성능을 개선하기 위해 카메라 기하학을 기반으로 하는 알고리즘을 사용하여 오검출을 선별하는 후처리 과정을 제안하였다. 실제 데이터를 활용한 실험을 통하여 전처리와 후처리를 모두 사용한 알고리즘이 로봇의 매니플레이션 성공률을 개선하는 데에 도움이 됨을 확인하였다. 본 논문에서 제안하는 기술을 기반으로 제조공정 과정에서 특정 물체를 매니플레이션 하는 로봇팔 제어 기술을 활용하여 생산성과 효율성을 높일 수 있을 것이라 기대한다.

References

- [1] L. Friedrich, H. Gerd, "Calibration and Synchronization of a Robot-Mounted Camera for Fast Sensor-Based Robot Motion," IEEE International Conference On Robotics and Automation, Vol. 4, pp. 3900-3905, 2005.
- [2] S. J. Park, S. B. Park, J. H. Kim, "A Study on the Motion Object Detection Method for Autonomous Driving," Journal of The Korean Society of Industry Convergence, Vol. 24, No. 5, pp. 547-553, 2021.
- [3] S. H. Hwang, J. H. Choi, H. S. Park, J. H. Kim, "A Study on the Implementation of Camera Sensor for Object Detection Based on Deep Learning and Distance Estimation," Transactions of the Korean Society of Automotive Engineers, Vol. 30, No. 8, pp. 659-665, 2022.
- [4] S. H. Park, K. H. Lee, J. S. Park, Y. S. Shin, "Deep Learning-Based Defect Detection for Sustainable Smart Manufacturing," Sustainability, Vol. 14, No. 5, pp. 2697, 2022.
- [5] N. Liu, L. Li, B. Hao, L. Yang, T. Hu, T. Xue, S. Wang, "Modeling and Simulation of Robot Inverse Dynamics Using LSTM-Based Deep Learning Algorithm for Smart Cities and Factories," IEEE ACCESS, Vol. 7, pp. 173989-173998, 2019.
- [6] R. Girshick, J. Donahue, T. Darrell, J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580-587, 2013.
- [7] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, A. C. Berg, "SSD: Single Shot MultiBox Detector," ECCV, pp. 21-37, 2016.
- [8] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, "You Only Look Once: Unified, Real-time Object Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779-788, 2016.
- [9] Y. B. Kim, H. Choi, J. H. Cho. "High-performance of Deep Learning Colorization With Wavelet Fusion," IEMEK J. Embed. Sys. Appl., Vol. 13, No. 6, pp. 313-319, 2018.
- [10] J. S. Jung, H. J. Park, J. H. Hwang, G. T. Park, N. J. Kwak. "객체 사이의 관계를 고려한 객체검출 알고리즘," 30th Workshop on Image Processing and Image Understanding, Vol. 30, 2018.
- [11] L. He, G. Wang, Z. Hu, "Learning Depth from Single Images with Deep Neural Network Embedding Focal Length," IEEE Transactions on Image Processing, Vol. 27, No. 9, pp. 4676-4689, 2018.
- [12] M. Diaz-Cabrera, P. Cerri, P. Medici, "Robust Real-time Traffic Light Detection and Distance Estimation Using a Single Camera," Expert Systems with Applications, Vol. 42, No. 8, pp. 3911-3923, 2015.

Sil Jin (진 실)



2024 Division of Electronic Engineering from Jeonbuk National University (B.S.)

Field of Interests: Artificial intelligence, Computer vision, Deep learning, Robotics
Email: jku951@jbnu.ac.kr

Jimin Song (송 지 민)



2022 Division of Electronic Engineering from Jeonbuk National University (B.S.)
2022~Division of Electronics and Information Engineering from Jeonbuk National University (M.S.)

Field of Interests: Artificial intelligence, Computer vision, Deep learning, Robotics
Email: jimn_song@jbnu.ac.kr

Jiho Choi (최 지 호)



2021 Division of Electronic Engineering from Jeonbuk National University (B.S.)
2023 Division of Electronics and Information Engineering from Jeonbuk National University (M.S.)
2023~Division of Electronics and Information Engineering from Jeonbuk National University (Ph.D.)

Field of Interests: Generative learning, Image-to-image translation, Weakly supervised learning, Time-series signal analysis
Email: jihochoi@jbnu.ac.kr

Yongsik Jin (진 용 식)



2014 School of Electronic and Electrical Engineering from Daegu University (B.S.)
2017 School of Electronic engineering from Kyungpook National University (M.S.)
2022 School of Electronic engineering from Kyungpook National University (Ph.D.)

2019~Electronics and Telecommunications Research Institute (Researcher)
Field of Interests: Artificial intelligence, Robotics
Email: yongsik@etri.re.kr

Jae Jin Jeong (정재진)

2008 Electrical and Electronic Engineering
from Kyungpook National University
(B.S.)

2011 Electrical and Electronic Engineering
from POSTECH (M.S.)

2016 Electrical and Electronic Engineering
from POSTECH (Ph.D.)

2022~School of Electronic Engineering at Kumoh National
Institute of Technology (Assistant Professor)

Career:

2016~2019 Senior Researcher, Defense Agency for Technology
and Quality

2019~2022 School of Electronic and Electrical Engineering at
Daegu Catholic University (Assistant Professor)

Field of Interests: Signal Processing, Deep Learning

Email: jeong@kumoh.ac.kr

Sang Jun Lee (이상준)

2011 Electrical Engineering from POSTECH
(B.S.)

2018 Electrical Engineering from POSTECH
(Ph.D.)

Career:

2018~2020 Samsung Advanced Institute of Technology
(Senior Researcher)

2020~ Jeonbuk National University (Assistant Professor)

Field of Interests: Artificial intelligence, Computer vision, Deep
learning, Robotics

Email: sj.lee@jbnu.ac.kr