



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0032241
(43) 공개일자 2022년03월15일

(51) 국제특허분류(Int. Cl.)

B25J 9/16 (2006.01)

(52) CPC특허분류

B25J 9/1612 (2013.01)

B25J 9/161 (2013.01)

(21) 출원번호 10-2020-0113796

(22) 출원일자 2020년09월07일

심사청구일자 2020년09월07일

(71) 출원인

울산과학기술원

울산광역시 울주군 언양읍 유니스트길 50

(72) 발명자

전세영

울산광역시 울주군 언양읍 유니스트길 50

박동원

울산광역시 울주군 언양읍 유니스트길 50

서용혁

울산광역시 울주군 언양읍 유니스트길 50

(74) 대리인

특허법인 무한

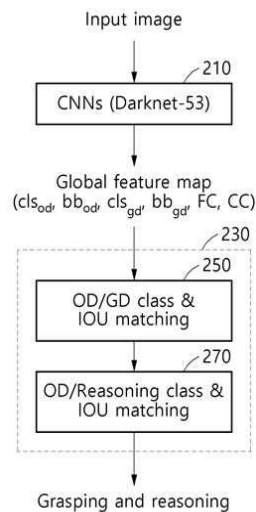
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 단일 뉴럴 네트워크를 이용한 로봇 파지 방법 및 이를 수행하는 장치

(57) 요약

단일 뉴럴 네트워크를 이용한 로봇 파지 방법 및 이를 수행하는 장치가 개시된다. 일 실시예에 따른 로봇 파지 방법은, 하나 이상의 객체를 포함하는 영역에 대한 이미지를 획득하는 단계와, 상기 이미지를 단일 뉴럴 네트워크에 입력하여 상기 하나 이상의 객체를 파지하기 위한 정보를 획득하는 단계와 상기 정보에 기초하여 로봇 파지를 위한 객체 관계 그래프를 생성하는 단계를 포함한다.

대표도 - 도2



(52) CPC특허분류

B25J 9/1653 (2013.01)

B25J 9/1697 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1415167691
과제번호	10077533
부처명	산업통상자원부
과제관리(전문)기관명	한국산업기술평가관리원
연구사업명	로봇산업핵심기술개발(R&D)
연구과제명	시각 및 촉각 정보를 이용하여 다양한 부품을 파지/조립할 수 있는 인공지능 기반의
부품 핸들링 기술 개발	
기 여 율	1/2
과제수행기관명	한국로봇융합연구원
연구기간	2020.01.01 ~ 2020.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1465027921
과제번호	HI18C0316030019
부처명	보건복지부
과제관리(전문)기관명	한국보건산업진흥원
연구사업명	첨단의료기술개발(R&D)
연구과제명	다중 병리요소 및 진단군을 대상으로 한 인공지능 기반 멀티 사이트 병리영상 고속
분석 기술 개발	
기 여 율	1/2
과제수행기관명	울산과학기술원
연구기간	2019.01.01 ~ 2019.12.31

명세서

청구범위

청구항 1

하나 이상의 객체를 포함하는 영역에 대한 이미지를 획득하는 단계;
상기 이미지를 단일 뉴럴 네트워크에 입력하여 상기 하나 이상의 객체를 파악하기 위한 정보를 획득하는 단계;
및
상기 정보에 기초하여 로봇 파지를 위한 객체 관계 그래프를 생성하는 단계
를 포함하는, 로봇 파지 방법.

청구항 2

제1항에 있어서,
상기 정보는,
대한 객체 검출(object detection) 정보, 파지점 검출(grasping detection) 정보 및 관계 추론(relationship reasoning) 정보를 포함하는 로봇 파지 방법.

청구항 3

제1항에 있어서,
상기 정보는,
상기 이미지를 분할하는 복수의 그리드에 대해 각 그리드 내 객체가 존재할 확률, 파지 가능할 확률 및 객체 위나 아래에 존재하는 객체의 정보를 포함하는, 로봇 파지 방법.

청구항 4

제1항에 있어서,
상기 획득하는 단계는,
객체 검출을 위한 제1 앵커 박스, 파지점 검출을 위한 제2 앵커 박스 및 파지각 검출을 위한 제3 앵커 박스를 이용하여 상기 정보를 획득하는 단계
를 포함하는, 로봇 파지 방법.

청구항 5

제2항에 있어서,
상기 객체 검출 정보는,
객체의 위치, 앵커 박스의 폭과 높이, 앵커 박스 내 객체가 존재할 확률 및 객체의 클래스에 연관된 파라미터들을 포함하는, 로봇 파지 방법.

청구항 6

제2항에 있어서,

상기 파지점 검출 정보는,

객체의 위치, 앵커 박스의 폭과 높이, 앵커 박스 내 파지점이 존재할 확률 및 객체의 클래스 및 앵커 박스의 각도에 연관된 파라미터들을 포함하는, 로봇 파지 방법.

청구항 7

제2항에 있어서,

상기 관계 추론 정보는,

객체의 위에 존재하는 객체의 클래스 및 객체의 아래에 존재하는 객체의 클래스 중 적어도 하나를 포함하는, 로봇 파지 방법.

청구항 8

제2항에 있어서,

상기 생성하는 단계는,

상기 객체 검출 정보 및 상기 파지점 검출 정보에 포함된 불필요한 검출 결과를 제거하는 단계; 및

객체가 검출된 객체 바운딩 박스 및 파지점이 검출된 파지점 바운딩 박스를 그룹화하는 단계

를 포함하는, 로봇 파지 방법.

청구항 9

제8항에 있어서,

상기 제거하는 단계는,

비-최대 억제 알고리즘(non-maximum suppression algorithm)에 기초하여 상기 불필요한 검출 결과를 제거하는 단계

를 포함하는, 로봇 파지 방법.

청구항 10

제8항에 있어서,

상기 그룹화하는 단계는,

객체의 클래스 정보에 기초하여 상기 객체 바운딩 박스 및 상기 파지점 바운딩 박스를 매칭하는 단계;

상기 객체 바운딩 박스 및 상기 파지점 바운딩 박스의 IoU(intersection over Union)를 계산하는 단계; 및

상기 IoU가 임계 값에 기초하여 상기 객체 바운딩 박스 및 상기 파지점 바운딩 박스를 매칭하는 단계

를 포함하는, 로봇 파지 방법.

청구항 11

인스트럭션들을 포함하는 메모리; 및

상기 인스트럭션들을 실행하기 위한 프로세서

를 포함하고,

상기 프로세서에 의해 상기 인스트럭션들이 실행될 때, 상기 프로세서는,

하나 이상의 객체를 포함하는 영역에 대한 이미지를 획득하고, 상기 이미지를 단일 뉴럴 네트워크에 입력하여 상기 하나 이상의 객체를 파지하기 위한 정보를 획득하고, 상기 정보에 기초하여 로봇 파지를 위한 객체 관계 그래프를 생성하는, 로봇 파지 정보 생성 장치.

청구항 12

제11항에 있어서,

상기 정보는,

대한 객체 검출(object detection) 정보, 파지점 검출(grasping detection) 정보 및 관계 추론(relationship reasoning) 정보를 포함하는, 로봇 파지 정보 생성 장치.

청구항 13

제11항에 있어서,

상기 정보는,

상기 이미지를 분할하는 복수의 그리드에 대해 각 그리드 내 특정 위치에 객체가 존재할 확률, 파지 가능할 확률 및 객체 위나 아래에 존재하는 객체의 정보를 포함하는, 로봇 파지 정보 생성 장치.

청구항 14

제11항에 있어서,

상기 프로세서는,

객체 검출을 위한 제1 앵커 박스, 파지점 검출을 위한 제2 앵커 박스 및 파지각 검출을 위한 제3 앵커 박스를 이용하여 상기 정보를 획득하는, 로봇 파지 정보 생성 장치.

청구항 15

제12항에 있어서,

상기 객체 검출 정보는,

객체의 위치, 앵커 박스의 폭과 높이, 앵커 박스 내 객체가 존재할 확률 및 객체의 클래스에 연관된 파라미터들을 포함하는, 로봇 파지 정보 생성 장치.

청구항 16

제12항에 있어서,

상기 파지점 검출 정보는,

객체의 위치, 앵커 박스의 폭과 높이, 앵커 박스 내 파지점이 존재할 확률 및 객체의 클래스 및 앵커 박스의 각도에 연관된 파라미터들을 포함하는, 로봇 파지 정보 생성 장치.

청구항 17

제12항에 있어서,

상기 관계 추론 정보는,

객체의 위에 존재하는 객체의 클래스 및 객체의 아래에 존재하는 객체의 클래스 중 적어도 하나를 포함하는, 로봇 파지 정보 생성 장치.

청구항 18

제12항에 있어서,

상기 프로세서는,

상기 객체 검출 정보 및 상기 파지점 검출 정보에 포함된 불필요한 검출 결과를 제거하고, 객체가 검출된 객체 바운딩 박스 및 파지점이 검출된 파지점 바운딩 박스를 그룹화하는, 로봇 파지 정보 생성 장치.

청구항 19

제18항에 있어서,

상기 프로세서는,

비-최대 억제 알고리즘(non-maximum suppression algorithm)에 기초하여 상기 불필요한 검출 결과를 제거하는, 로봇 파지 정보 생성 장치.

청구항 20

제18항에 있어서,

상기 프로세서는,

객체의 클래스 정보에 기초하여 상기 객체 바운딩 박스 및 상기 파지점 바운딩 박스를 매칭하고, 상기 객체 바운딩 박스 및 상기 파지점 바운딩 박스의 IoU(intersection over Union)를 계산하고, 상기 IoU가 임계 값에 기초하여 상기 객체 바운딩 박스 및 상기 파지점 바운딩 박스를 매칭하는, 로봇 파지 정보 생성 장치.

발명의 설명

기술 분야

[0001] 본 개시는 단일 뉴럴 네트워크를 이용한 로봇 파지 방법 및 이를 수행하는 장치에 관한 것이다.

배경 기술

[0002] 머신 러닝(machine learning)을 이용하여, 로봇의 동작을 제어하기 위한 다양한 연구들이 진행되고 있다. 그리고 기계 학습 방법 중에서도, 심층 지도 학습을 통해 로봇의 파지 동작을 제어하는 연구들이 있다. 심층 지도 학습이란 심층 학습(Deep learning)과 지도 학습(Supervised learning)이 결합한 형태의 기계 학습 방법이다.

[0003] 로봇은 인공 신경망 즉 뉴럴 네트워크(neural network)를 이용하여 심층 지도 학습을 수행할 수 있다. 로봇은 카메라로부터 획득한 이미지를 통해 파지 대상인 목표 물체의 특징을 추출하며, 레이블(label), 즉 정답으로 제시된 목표 물체에 대한 파지 동작을 이용하여 지도 학습을 수행할 수 있다. 그리고 이미지로부터 목표 물체의 특징을 추출하기 위해, CNN(Convolutional Neural Network)이 이용될 수 있다.

- [0004] 최근 로봇 파지 정보뿐만 아니라 물체 인식 및 관계성 파악까지 동시에 하는 방식이 제안되었다. 다만, 각각의 태스크에 대해서 독립적인 뉴럴 네트워크를 사용하는 방식은 공유될 수 있는 물체에 대한 정보들이 따로 처리되어 효율성을 감소시킨다.
- [0005] 관련 선행기술로, 한국 등록특허공보 제10-2023588호(발명의 명칭: 로봇 파지용 심층 기계 학습 방법 및 장치)가 있다.

발명의 내용

해결하려는 과제

- [0006] 아래 실시예들은 로봇 파지를 위한 파지점 검출, 물체 검출, 물체 관계 추론을 단일 네트워크를 이용하여 수행하는 기술을 제공할 수 있다.
- [0007] 다만, 기술적 과제는 상술한 기술적 과제들로 한정되는 것은 아니며, 또 다른 기술적 과제들이 존재할 수 있다.

과제의 해결 수단

- [0008] 일 실시예에 따른 로봇 파지 방법은, 하나 이상의 객체를 포함하는 영역에 대한 이미지를 획득하는 단계와, 상기 이미지를 단일 뉴럴 네트워크에 입력하여 상기 하나 이상의 객체를 파지하기 위한 정보를 획득하는 단계와 상기 정보에 기초하여 로봇 파지를 위한 객체 관계 그래프를 생성하는 단계를 포함한다.
- [0009] 상기 정보는, 대한 객체 검출(object detection) 정보, 파지점 검출(grasping detection) 정보 및 관계 추론(relationship reasoning) 정보를 포함할 수 있다.
- [0010] 상기 정보는, 상기 이미지를 분할하는 복수의 그리드에 대해 각 그리드 내 특정 위치에 객체가 존재할 확률, 파지 가능할 확률 및 객체 위나 아래에 존재하는 객체의 정보를 포함할 수 있다.
- [0011] 상기 획득하는 단계는, 객체 검출을 위한 제1 앵커 박스, 파지점 검출을 위한 제2 앵커 박스 및 파지각 검출을 위한 제3 앵커 박스를 이용하여 상기 정보를 획득하는 단계를 포함할 수 있다.
- [0012] 상기 객체 검출 정보는, 객체의 위치, 앵커 박스의 폭과 높이, 앵커 박스 내 객체가 존재할 확률 및 객체의 클래스에 연관된 파라미터들을 포함할 수 있다.
- [0013] 상기 파지점 검출 정보는, 객체의 위치, 앵커 박스의 폭과 높이, 앵커 박스 내 파지점이 존재할 확률 및 객체의 클래스 및 앵커 박스의 각도에 연관된 파라미터들을 포함할 수 있다.
- [0014] 상기 관계 추론 정보는, 객체의 위에 존재하는 객체의 클래스 및 객체의 아래에 존재하는 객체의 클래스 중 적어도 하나를 포함할 수 있다.
- [0015] 상기 생성하는 단계는, 상기 객체 검출 정보 및 상기 파지점 검출 정보에 포함된 불필요한 검출 결과를 제거하는 단계와, 객체가 검출된 객체 바운딩 박스 및 파지점이 검출된 파지점 바운딩 박스를 그룹화하는 단계를 포함할 수 있다.
- [0016] 상기 제거하는 단계는, 비-최대 억제 알고리즘(non-maximum suppression algorithm)에 기초하여 상기 불필요한 검출 결과를 제거하는 단계를 포함할 수 있다.
- [0017] 상기 그룹화하는 단계는, 객체의 클래스 정보에 기초하여 상기 객체 바운딩 박스 및 상기 파지점 바운딩 박스를 매칭하는 단계와, 상기 객체 바운딩 박스 및 상기 파지점 바운딩 박스의 IoU(intersection over Union)를 계산하는 단계와, 상기 IoU가 임계 값에 기초하여 상기 객체 바운딩 박스 및 상기 파지점 바운딩 박스를 매칭하는 단계를 포함할 수 있다.
- [0018] 일 실시예에 따른 로봇 파지 정보 생성 장치는, 인스트럭션들을 포함하는 메모리와, 상기 인스트럭션들을 실행하기 위한 프로세서를 포함하고, 상기 프로세서에 의해 상기 인스트럭션들이 실행될 때, 상기 프로세서는, 하나 이상의 객체를 포함하는 영역에 대한 이미지를 획득하고, 상기 이미지를 단일 뉴럴 네트워크에 입력하여 상기 하나 이상의 객체를 파지하기 위한 정보를 획득하고, 상기 정보에 기초하여 로봇 파지를 위한 객체 관계 그래프를 생성한다.
- [0019] 상기 정보는, 대한 객체 검출(object detection) 정보, 파지점 검출(grasping detection) 정보 및 관계 추론(relationship reasoning) 정보를 포함할 수 있다.

- [0020] 상기 정보는, 상기 이미지를 분할하는 복수의 그리드에 대해 각 그리드 내 특정 위치에 객체가 존재할 확률, 파괴 가능할 확률 및 객체 위나 아래에 존재하는 객체의 정보를 포함할 수 있다.
- [0021] 상기 프로세서는, 객체 검출을 위한 제1 앵커 박스, 파괴점 검출을 위한 제2 앵커 박스 및 파괴각 검출을 위한 제3 앵커 박스를 이용하여 상기 정보를 획득할 수 있다.
- [0022] 상기 객체 검출 정보는, 객체의 위치, 앵커 박스의 폭과 높이, 앵커 박스 내 객체가 존재할 확률 및 객체의 클래스에 연관된 파라미터들을 포함할 수 있다.
- [0023] 상기 파괴점 검출 정보는, 객체의 위치, 앵커 박스의 폭과 높이, 앵커 박스 내 파괴점이 존재할 확률 및 객체의 클래스 및 앵커 박스의 각도에 연관된 파라미터들을 포함할 수 있다.
- [0024] 상기 관계 추론 정보는, 객체의 위에 존재하는 객체의 클래스 및 객체의 아래에 존재하는 객체의 클래스 중 적어도 하나를 포함할 수 있다.
- [0025] 상기 프로세서는, 상기 객체 검출 정보 및 상기 파괴점 검출 정보에 포함된 불필요한 검출 결과를 제거하고, 객체가 검출된 객체 바운딩 박스 및 파괴점이 검출된 파괴점 바운딩 박스를 그룹화할 수 있다.
- [0026] 상기 프로세서는, 비-최대 억제 알고리즘(non-maximum suppression algorithm)에 기초하여 상기 불필요한 검출 결과를 제거할 수 있다.
- [0027] 상기 프로세서는, 객체의 클래스 정보에 기초하여 상기 객체 바운딩 박스 및 상기 파괴점 바운딩 박스를 매칭하고, 상기 객체 바운딩 박스 및 상기 파괴점 바운딩 박스의 IoU(intersection over Union)를 계산하고, 상기 IoU가 임계 값에 기초하여 상기 객체 바운딩 박스 및 상기 파괴점 바운딩 박스를 매칭할 수 있다.

도면의 간단한 설명

- [0028] 도 1은 로봇 파괴를 위한 객체 관계 그래프를 생성하는 장치를 개략적으로 도시한 도면이다.
- 도 2는 도 1에 도시된 장치의 동작을 설명하기 위한 도면이다.
- 도 3은 도 1에 도시된 장치가 이용하는 뉴럴 네트워크의 구조를 나타내는 도면이다.
- 도 4a 내지 도 4d는 도 1에 도시된 장치의 성능을 보여주기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0029] 이하에서, 첨부된 도면을 참조하여 실시예들을 상세하게 설명한다. 그러나, 실시예들에는 다양한 변경이 가해질 수 있어서 특허출원의 권리 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 실시예들에 대한 모든 변경, 균등물 내지 대체물이 권리 범위에 포함되는 것으로 이해되어야 한다.
- [0030] 실시예에서 사용한 용어는 단지 설명을 목적으로 사용된 것으로, 한정하려는 의도로 해석되어서는 안된다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 명세서 상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0031] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0032] 또한, 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일한 참조부호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다. 실시예를 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 실시예의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.
- [0033] 또한, 실시 예의 구성 요소를 설명하는 데 있어서, 제 1, 제 2, A, B, (a), (b) 등의 용어를 사용할 수 있다. 이러한 용어는 그 구성 요소를 다른 구성 요소와 구별하기 위한 것일 뿐, 그 용어에 의해 해당 구성 요소의 본질이나 차례 또는 순서 등이 한정되지 않는다. 어떤 구성 요소가 다른 구성요소에 "연결", "결합" 또는 "접

속"된다고 기재된 경우, 그 구성 요소는 그 다른 구성요소에 직접적으로 연결되거나 접속될 수 있지만, 각 구성 요소 사이에 또 다른 구성 요소가 "연결", "결합" 또는 "접속"될 수도 있다고 이해되어야 할 것이다.

- [0034] 어느 하나의 실시 예에 포함된 구성요소와, 공통적인 기능을 포함하는 구성요소는, 다른 실시 예에서 동일한 명칭을 사용하여 설명하기로 한다. 반대되는 기재가 없는 이상, 어느 하나의 실시 예에 기재한 설명은 다른 실시 예에도 적용될 수 있으며, 중복되는 범위에서 구체적인 설명은 생략하기로 한다.
- [0036] 도 1은 로봇 파지를 위한 객체 관계 그래프를 생성하는 장치를 개략적으로 도시한 도면이다.
- [0037] 로봇은 특정 물체를 파지(grasping)하기 위해 객체를 검출하고(Object Detection(OD)), 객체 간의 관계를 파악하고(Relationship reasoning), 객체에 대한 파지점을 검출(Grasping Detection(GD))해야 할 수 있다.
- [0038] 로봇 파지를 위한 객체 관계 그래프는 객체 검출 정보, 관계 추론 정보 및 파지점 검출 정보를 포함할 수 있다. 즉, 로봇은 객체 관계 그래프에 기초하여 물체를 파지할 수 있다.
- [0039] 객체 관계 그래프를 생성하기 위한 장치(100)는 메모리(130) 및 프로세서(150)를 포함한다.
- [0040] 프로세서(150)는 메모리(130)에 저장된 컴퓨터로 읽을 수 있는 코드(예를 들어, 소프트웨어) 및 프로세서에 의해 유발된 인스트럭션(instruction)들을 실행할 수 있다.
- [0041] 프로세서(150)는 목적하는 동작들(desired operations)을 실행시키기 위한 물리적인 구조를 갖는 회로를 가지는 하드웨어로 구현된 데이터 처리 장치일 수 있다. 예를 들어, 목적하는 동작들은 프로그램에 포함된 코드(code) 또는 인스트럭션들(instructions)을 포함할 수 있다.
- [0042] 예를 들어, 하드웨어로 구현된 데이터 처리 장치는 마이크로프로세서(microprocessor), 중앙 처리 장치(central processing unit), 프로세서 코어(processor core), 멀티-코어 프로세서(multi-core processor), 멀티프로세서(multiprocessor), ASIC(Application-Specific Integrated Circuit), FPGA(Field Programmable Gate Array)를 포함할 수 있다.
- [0043] 메모리(130)는 프로세서에 의해 실행가능한 인스트럭션들(또는 프로그램)을 저장할 수 있다. 예를 들어, 인스트럭션들은 프로세서의 동작 및/또는 프로세서의 각 구성의 동작을 실행하기 위한 인스트럭션들을 포함할 수 있다.
- [0044] 메모리(130)는 휘발성 메모리 장치 또는 불휘발성 메모리 장치로 구현될 수 있다.
- [0045] 휘발성 메모리 장치는 DRAM(dynamic random access memory), SRAM(static random access memory), T-RAM(thyristor RAM), Z-RAM(zero capacitor RAM), 또는 TTRAM(Twin Transistor RAM)으로 구현될 수 있다.
- [0046] 불휘발성 메모리 장치는 EEPROM(Electrically Erasable Programmable Read-Only Memory), 플래시(flash) 메모리, MRAM(Magnetic RAM), 스핀전달토크 MRAM(Spin-Transfer Torque(STT)-MRAM), Conductive Bridging RAM(CBRAM), FeRAM(Ferroelectric RAM), PRAM(Phase change RAM), 저항 메모리(Resistive RAM(RRAM)), 나노 튜브 RRAM(Nanotube RRAM), 폴리머 RAM(Polymer RAM(PoRAM)), 나노 부유 게이트 메모리(Nano Floating Gate Memory(NFGM)), 홀로그래픽 메모리(holographic memory), 분자 전자 메모리 소자(Molecular Electronic Memory Device), 또는 절연 저항 변화 메모리(Insulator Resistance Change Memory)로 구현될 수 있다.
- [0047] 프로세서(150)는 하나 이상의 객체를 포함하는 영역에 대한 이미지에 기초하여 로봇이 객체를 파지하기 위한 정보를 획득할 수 있다.
- [0048] 프로세서(150)는 단일 심층 뉴럴 네트워크(deep neural network(DNN))를 이용하여 파지점 검출(GD), 객체 검출(OD), 객체 사이 관계 추론을 포함하는 다중 태스크를 처리할 수 있다. 프로세서(150)는 YOLOv3를 기반으로 하는 단일 DNN을 이용하여 RGB-D 이미지로부터 파지점 검출(GD), 객체 검출(OD), 객체 사이 관계 추론을 수행할 수 있다.
- [0049] 프로세서(150)는 뉴럴 네트워크를 이용하여 획득한 피쳐맵(feature map)에 기초하여 객체 관계 그래프를 획득할 수 있다. 프로세서(150)는 피쳐맵에 포함된 정보를 후처리하여 로봇 파지를 위한 정보를 추출할 수 있다. 피쳐맵은 객체가 검출된 바운딩 박스(bounding box) 및 파지점을 검출하는 바운딩 박스, 각 바운딩 박스 내 객체의 클래스 정보 및 객체 간 관계 추론 정보를 포함할 수 있다.
- [0050] 프로세서(150)는 단일 DNN을 이용함으로써 메모리 사용량을 줄이고 더 높은 정확도 및 더 빠른 계산 속도로 파

지점 검출(GD), 객체 검출(OD), 객체 사이 관계 추론을 수행할 수 있다.

- [0052] 도 2는 도 1에 도시된 장치의 동작을 설명하기 위한 도면이고, 도 3은 도 1에 도시된 장치가 이용하는 뉴럴 네트워크의 구조를 나타내는 도면이다.
- [0053] 로봇이 단일 물체를 파지하기 위한 파지점 검출(GD)는 평행 그리퍼(parallel gripper)와 함께 5D 로봇 파지 표현법(5D robotic grasp representation)이 이용될 수 있다. 5D 로봇 파지 표현법은 벡터 $\{x_{gd}, y_{gd}, \theta_{gd}, w_{gd}, h_{gd}\}$ 일 수 있다. (x_{gd}, y_{gd}) 는 그리드 내 객체의 위치(position)를 나타내고, θ_{gd} 는 방향(orientation)을 나타내고, w_{gd} 는 그리퍼 개방 너비(gripper opening width)를 나타내고, h_{gd} 는 평행 그리퍼 플레이트 사이즈(parallel gripper plate size)를 나타낼 수 있다.
- [0054] 로봇이 물체들이 임의의 배치로 쌓여 있는 환경에서 특정 대상 물체를 파지하기 위한 경우 하나 이상의 물체의 파지 정보가 필요하며, 순차적인 파지 계획을 세우기 위해 물체의 클래스(class) 및 관계 추론 정보가 필요할 수 있다. 이 경우 5D 로봇 파지 표현법에 물체 클래스(cls_{gd}) 및 쌓인 순서(ord_{gd})를 더 포함시킬 수 있다.
- [0055] 프로세서(150)는 입력 이미지(input image)를 컨볼루션 뉴럴 네트워크(Convolution Neural Network(CNN); 210)에 입력할 수 있다. 입력 이미지는 RGB-D 카메라를 통해 획득한 2차원 이미지일 수 있다. CNN(210)은 객체 검출, 파지점 검출 및 관계 추론을 수행할 수 있도록 학습된 단일 다중-작업 DNN(multi-task DNN)일 수 있다.
- [0056] CNN(210)은 15D 표현법에 기초하여 연산이 수행될 수 있다. 15D 표현법은 객체 검출 파라미터, 파지점 검출 파라미터 및 관계 추론 파라미터의 재매개변수화(reparametrization) 파라미터들을 포함할 수 있다.
- [0057] CNN(210)은 앵커 박스(anchor box)를 이용하여 그리드(grid)로 분할된 입력 이미지에 대해 각 그리드에서의 파라미터들을 계산할 수 있다.
- [0058] 객체 검출 파라미터는 $\{x_{od}, y_{od}, w_{od}, h_{od}, cls_{od}, pr_{od}\}$ 이고, 파지점 검출 파라미터는 $\{x_{gd}, y_{gd}, w_{gd}, h_{gd}, pr_{gd}, ord_{gd}, \theta_{gd}\}$ 일 수 있다. (x_{od}, y_{od}) 및 (x_{gd}, y_{gd}) 각각 객체 및 파지점의 위치를 나타내고, (w_{od}, h_{od}) 및 (w_{gd}, h_{gd}) 는 각각 객체 검출을 위한 앵커 박스의 크기 및 파지점 검출을 위한 앵커 박스의 크기를 나타내고, θ_{gd} 는 파지점 검출을 위한 앵커 박스의 방향을 나타내고, pr_{od} 는 객체가 존재할 확률을 나타내고, pr_{gd} 는 파지 가능할 확률을 나타낼 수 있다. 관계 추론 파라미터는 $\{cls_{fc}, cls_{cc}\}$ 일 수 있다. cls_{fc} 및 cls_{cc} 는 각각 객체의 아래에 있는 객체의 클래스(Father Class(FC)) 및 위에 있는 객체의 클래스(Children Class(CC))를 나타낼 수 있다.
- [0059] 객체 검출 파라미터(OD), 파지점 검출 파라미터(GD) 및 관계 추론 파라미터(R)의 재매개변수화 파라미터들은 수학적 식 1 내지 수학적 식 3과 같이 표현될 수 있다.

수학적 식 1

$$OD = \{t_{od}^x, t_{od}^y, t_{od}^w, t_{od}^h, t_{od}^{pr}, t_{od}^{cls}\}$$

수학적 식 2

$$GD = \{t_{gd}^x, t_{gd}^y, t_{gd}^w, t_{gd}^h, t_{gd}^{pr}, t_{gd}^{cls}, t_{gd}^{\theta}\}$$

수학식 3

$$R = \{t_{fa}^{cls}, t_{cc}^{cls}\}$$

이때, $x_j = \sigma(t_j^x) + c_j^x, y_j = \sigma(t_j^y) + c_j^y$, $\sigma(\cdot)$ 는 시그모이드 함수(sigmoid function)이고, $w_j = p_j^w \exp(t_j^w), h_j = p_j^h \exp(t_j^h)$, $\theta_{gd} = p_{gd}^\theta + t_{gd}^\theta$, $cls_{rs} = \sigma(t_{rs}^{cls})$, $pr_j = \sigma(t_j^{pr})$, $j \in \{od, gd\}$ and $rs \in \{fc, cc\}$ 이고, p_j^h , p_j^w 및 p_{gd}^θ 는 각각 앵커 박스의 미리 정의된 높이, 너비 및 방향이고, (c_j^x, c_j^y) 는 각 그리드 셀의 왼쪽 상단 모서리 위치일 수 있다.

CNN(210)은 $\{x_j, y_j, \theta_{gd}, w_j, h_j, pr_j\}$ 대신 $\{t_j^x, t_j^y, t_j^\theta, t_j^w, t_j^h, t_j^{pr}\}$ 를 추정할 수 있다. x_j, y_j, w_j, h_j 는 각 그리드의 크기가 1X1이 되도록 정규화될 수 있고, θ_{gd} 는 불연속 값으로 모델링될 수 있다.

CNN(210)은 모든 그리드 내 앵커 박스 후보군 중에서 가장 파지 성공 확률이 높은 지점을 선택할 수 있다. 이때, 재매개변수화는 회귀 문제(regression problem)을 w_j, h_j 에 대한 회귀 + 분류 문제(regression + classification problem)으로 변경할 수 있다.

CNN(210)은 유한한 수의 각도 후보 중 가장 파지 성공 확률이 높은 각도 θ_{gd} 를 추정할 수 있다. CNN(210)은 방향 앵커 박스(orientation anchor box)를 이용하여 θ_{gd} 를 추정할 수 있다. 예를 들어, CNN(210)은 분류(classification) 및 불연속 앵커 상자 회전(discrete anchor box rotations)를 이용하여 θ_{gd} 를 결정할 수 있다.

CNN(210)은 각각의 객체 검출 박스(object detection box) 및/또는 파지 검출 박스(grasp detection box)에 대해 객체의 클래스를 추정할 수 있다. 예를 들어, CNN(210)은 클래스 활성화 함수(class activation function)으로 소프트맥스(softmax)를 이용할 수 있다.

CNN(210)은 각각의 대상 객체의 위 및/또는 아래에 있는 객체의 클래스를 검출할 수 있다. CNN(210)은 객체가 존재하지 않는 경우를 포함하여 $(t_{fc}^{cls}, t_{cc}^{cls}) \in \{class_1, class_2, ..., class_{none}\}$ 를 검출할 수 있다. CNN(210)은 객체의 위에 있는 객체의 클래스(CC)만을 추정할 수도 있다.

CNN(210)은 도 3과 같이, 사전 학습된 관계 추론(reasoning) 및 파지점 검출을 포함하는 다중-태스크 객체 검출을 위해 확장된 객체 검출에 대한 YOLOv3의 Darknet-53을 상속하는 FCNN일 수 있다.

CNN(210)은 다중 태스크 예측(multi-task predictions)을 위해, 피쳐 피라미드 네트워크(feature pyramid networks)를 이용하여 객체 검출을 스케일에 걸쳐 예측(prediction across scales)할 뿐만 아니라 각각의 스케일에서 관계 추론 및 파지점 검출을 수행할 수 있도록 확장될 수 있다.

CNN(210)은 서로 다른 스케일(scale)에서 서로 다른 앵커 박스를 이용하여 객체 검출, 파지점 검출 및 관계 추론 결과를 출력할 수 있다. 예를 들어, CNN(210)은 저해상도 스케일(low-resolution scale), 중간 해상도 스케일(mid-resolution scale) 및 고해상도 스케일(high-resolution scale)에서 서로 다른 앵커 박스를 이용하여 각각의 경우에 객체가 검출된 바운딩 박스, 파지점이 검출된 바운딩 박스 및 관계 정보를 출력할 수 있다.

CNN(210)은 객체 검출을 위한 세 개의 앵커 박스, 파지점 검출을 위한 한 개의 앵커 박스 및 파지 각도(grasping angle)를 검출하기 위한 네 개의 앵커 박스를 이용할 수 있다.

[0077] CNN(210)는 저해상도 스케일에서 수학적식 4와 같이 앵커 박스를 예측할 수 있다.

수학적식 4

$$(p_{od}^w, p_{od}^h) \in \{(540, 540), (480, 480), (420, 420)\},$$

$$(p_{gd}^w, p_{gd}^h) \in \{(300, 300)\}, p_{gd}^\theta \in \{0, \pi/4, 2\pi/4, 3\pi/4\}$$

[0079]

[0081] CNN(210)는 X2 바이리니어 스케일 업(X2 bilinear scale up)을 수행한 후 중간 해상도 스케일에서 수학적식 5와 같이 앵커 박스를 예측할 수 있다.

수학적식 5

$$(p_{od}^w, p_{od}^h) \in \{(360, 360), (300, 300), (240, 240)\},$$

$$(p_{gd}^w, p_{gd}^h) \in \{(100, 100)\}, p_{gd}^\theta \in \{0, \pi/4, 2\pi/4, 3\pi/4\}$$

[0083]

[0085] CNN(210)는 X4 바이리니어 스케일 업(X4 bilinear scale up)을 수행한 후 고 해상도 스케일에서 수학적식 6와 같이 앵커 박스를 예측할 수 있다.

수학적식 6

$$(p_{od}^w, p_{od}^h) \in \{(180, 180), (120, 120), (60, 60)\}$$

[0087]

[0089] CNN(210)은 객체 검출 및 관계 추론을 X1, X2 및 X4 스케일에서 수행할 수 있고, 파지점 검출은 X1 및 X2 스케일에서 수행할 수 있다.

[0090] CNN(210)은 관계 추론을 포함하는 객체 검출을 위한 네 개의 바운딩 박스 오프셋, 객체 검출 확률, 객체의 클래스(클래스 번호), (클래스 번호+1) X2 관계 추론 클래스(FC, CC)로 아홉 개의 앵커 박스가 예측할 수 있다.

[0091] CNN(210)은 파지점 검출을 위한 네 개의 바운딩 박스 오프셋, 방향(orientation), 파지점 검출 확률 및 객체의 클래스(클래스 번호)로 여덟 개의 앵커 박스를 예측할 수 있다.

[0092] CNN(210)은 수학적식 7과 같은 재매개화된 파라미터(OD, GD 및 R)과 기반 진실(Ground Truth(GT) OD_{gt}, GD_{gt} 및 R_{gt})에 대한 손실 함수에 기초하여 학습될 수 있다.

수학식 7

$$\begin{aligned} & \sum_{i \in \Omega} \sum_{j \in \{od, gd\}} \{ \sum_{k \in \{x, y, w, h\}} \text{MSE}(k_j^i, k_{j,gt}^i) + \\ & \sum_{k \in pr} (-\log k_j^i) + \sum_{k \in cls_{ob}} \text{FocLoss}(k_j^i, k_{j,gt}^i) \} + \\ & \lambda_n \sum_{i \in \Omega^c} \sum_{j \in \{od, gd\}} \sum_{k \in pr} (-\log(1 - k_j^i)) + \\ & \sum_{i \in \Omega} \sum_{j \in R} \sum_{k \in cls_{fc}, cls_{cc}} \text{BiFocLoss}(k_j^i, k_{j,gt}^i) + \\ & \sum_{i \in \Omega} \sum_{j \in gd} \sum_{k \in \theta} \text{MSE}(k_j^i, k_{j,gt}^i) \end{aligned}$$

[0094]

[0096]

이때, x, y, w, h 및 z는 각각 t^x , t^y , t^w , t^h 및 t^z 의 함수이고, Ω 은 파지 대상 객체가 위치하는 그리드일 수 있다. FocLoss는 수학식 8과 같이 표현될 수 있다. 즉, cls_{fc} 및 cls_{cc} 는 다중 클래스이므로, binary focal loss(BiFocLoss)가 이용될 수 있고, 이때, 포컬 로스 감마(focal loss gamma)은 2로 설정되고, $\lambda_n = 100$ 로 설정될 수 있다.

수학식 8

$$\text{FocLoss}(p_t) = -(1 - p_t)^\gamma \log p_t$$

[0098]

[0100]

프로세서(150)는 CNN(210)에서 출력된 정보를 후처리(230)하여 객체 관계 그래프를 획득할 수 있다. 예를 들어, 프로세서(150)는 객체 검출 결과와 파지점 검출 결과의 클래스 정보 및 관계 추론 결과에 기초하여 객체 검출 및 파지점 검출 바운딩 박스 쌍을 획득할 수 있다.

[0101]

프로세서(150)는 비-최대 억제 알고리즘(non-maximum suppression algorithm)에 기초하여 불필요한 검출 결과를 제거할 수 있다.

[0102]

프로세서(150)는 객체 검출(OD) 및 파지점 검출(GD) 바운딩 박스(bbod, bbgd)를 각각의 바운딩 박스의 클래스 정보(cls_{od} , cls_{gd})에 기초하여 그룹화할 수 있다(250). 프로세서(150)는 수학식 9와 같이 표현되는 IOU(Intersection Over Union)에 기초하여 객체 검출(OD) 및 파지점 검출(GD) 바운딩 박스(bbod, bbgd) 쌍을 추가적으로 그룹화할 수 있다.

수학식 9

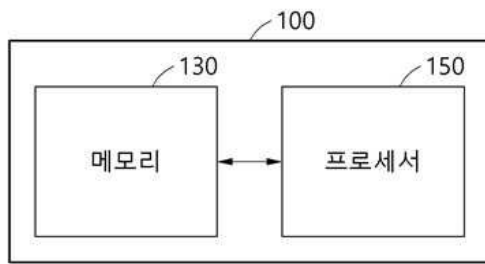
$$IOU = \frac{bb_{od} \cap bb_{gd}}{bb_{od} \cup bb_{gd}}$$

[0104]

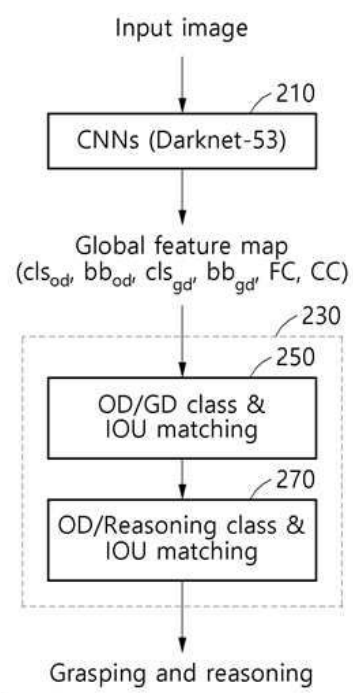
- [0106] 프로세서(150)는 IOU가 임계 값을 초과하는 파지점 검출(GD) 바운딩 박스 후보 중 파지점 검출(GD)에 대한 최상의 확률을 선택하여 최종적으로 객체 검출(OD) 및 파지점 검출(GD) 바운딩 박스(bbod, bbgd) 쌍을 선택할 수 있다.
- [0107] 프로세서(150)는 이와 유사하게 IOU 임계 값을 이용하여 바운딩 박스를 매칭함으로써 객체의 클래스와 객체 위의 객체의 클래스(CC)를 비교하여 관계를 획득할 수 있다.
- [0109] 도 4a 내지 도 4d는 도 1에 도시된 장치의 성능을 보여주기 위한 도면이다.
- [0110] 프로세서(150)는 도 4a에 표시된 파지점의 후보들(실선 박스)로부터 도 4b에 표시된 추정된 객체의 클래스와 최적의 파지점(파선 박스)를 검출할 수 있다. 또한, 프로세서(150)는 도 4c와 같이 각 객체 사이의 관계를 추론할 수 있다.
- [0111] 이때, 도 4d를 참조하면, 프로세서(150)는 상술한 방법을 이용하여 다른 방식(a, b, 및 c)과 비교하여 계산 속도(FPS) 대 예측 정확도(mAP)에서 높은 성능을 보이는 것을 확인할 수 있다.
- [0113] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0115] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.
- [0117] 이상과 같이 실시예들이 비록 한정된 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기를 기초로 다양한 기술적 수정 및 변형을 적용할 수 있다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.
- [0118] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 청구범위의 범위에 속한다.

도면

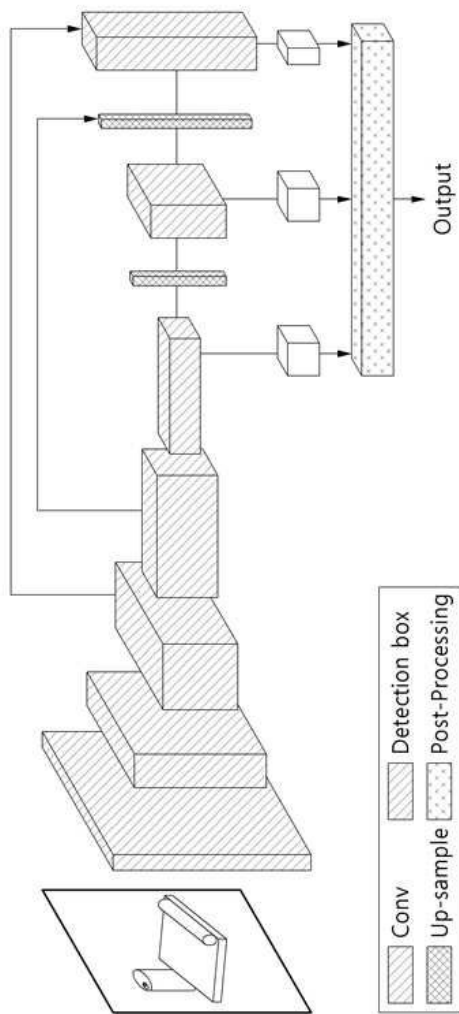
도면1



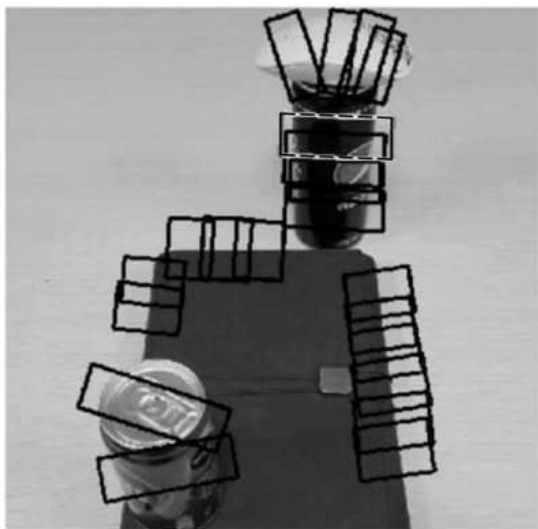
도면2



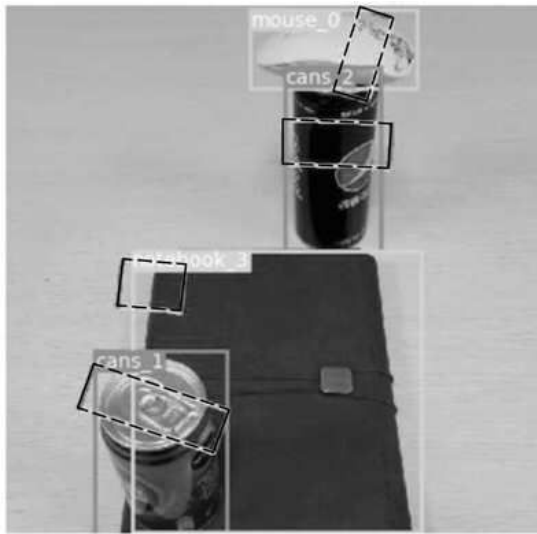
도면3



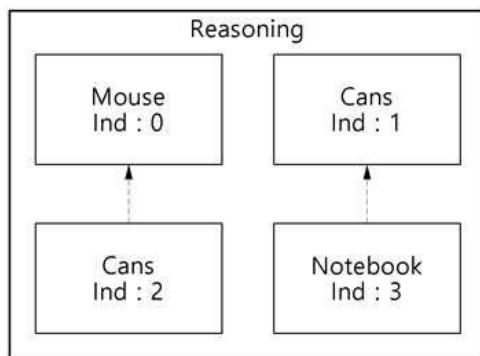
도면4a



도면4b



도면4c



도면4d

