



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0002970
(43) 공개일자 2021년01월11일

(51) 국제특허분류(Int. Cl.)
G06N 3/08 (2006.01) G06N 3/04 (2006.01)
(52) CPC특허분류
G06N 3/08 (2013.01)
G06K 9/00456 (2013.01)
(21) 출원번호 10-2019-0078910
(22) 출원일자 2019년07월01일
심사청구일자 없음

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암
동5가)
(72) 발명자
송재복
서울특별시 강남구 언주로30길 26
김다윗
경기도 고양시 일산서구 킨텍스로 340, 705동
1306호
조현준
서울특별시 성북구 인촌로26길 34, 2층 7호
(74) 대리인
심경식, 홍성욱

전체 청구항 수 : 총 5 항

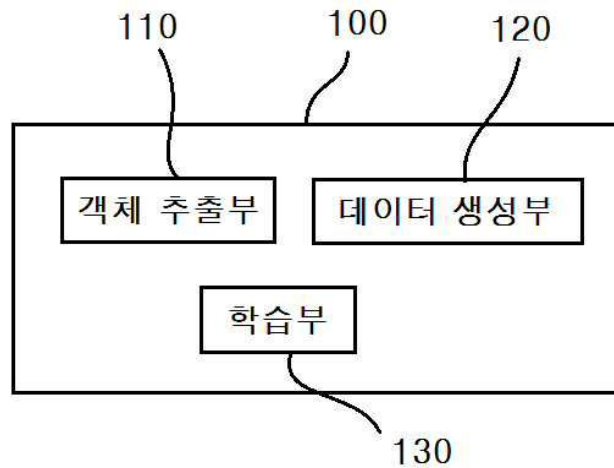
(54) 발명의 명칭 영상 학습 장치

(57) 요약

실시예에 따른 영상 학습 장치는 레퍼런스 이미지로부터 객체 이미지를 추출하는 객체 추출부와, 상기 객체 이미지에 서로 다른 배경 이미지들을 각각 합성하여 입력 데이터를 생성하는 데이터 생성부와, 인공 신경망을 이용하여 상기 입력 데이터를 학습시키는 학습부를 포함할 수 있다.

실시예는 다양한 학습 데이터를 효과적으로 수집함으로써, 객체 인식 성능을 향상시킬 수 있는 효과가 있다.

대표도 - 도1



(52) CPC특허분류

G06N 3/04 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	2018-0-00622
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기술진흥센터
연구사업명	ICT융합산업원천기술개발
연구과제명	촉각이 가능한 로봇 손으로 다양한 물체를 다루는 방법과 절차를 학습하는 로봇 손
조작 지능 개발	
기 여 율	1/1
과제수행기관명	고려대학교 산학협력단
연구기간	2018.04.01 ~ 2021.12.31

명세서

청구범위

청구항 1

레퍼런스 이미지로부터 객체 이미지를 추출하는 객체 추출부;
상기 객체 이미지에 서로 다른 배경 이미지들을 각각 합성하여 입력 데이터를 생성하는 데이터 생성부; 및
인공 신경망을 이용하여 상기 입력 데이터를 학습시키는 학습부;를 포함하는 영상 학습 장치.

청구항 2

제1항에 있어서,
상기 객체 이미지는 상기 레퍼런스 이미지로부터 배경 제거 알고리즘을 사용하여 상기 배경 이미지를 제거하고,
상기 배경 제거 알고리즘은 상기 객체 이미지의 윤곽선을 검출하고, 상기 윤곽선의 바깥쪽 영역의 투명도를 최대로 설정하여 상기 배경 이미지를 제거하는 영상 학습 장치.

청구항 3

제2항에 있어서,
상기 입력 데이터는 레이블링 파일을 더 포함하고, 상기 데이터 생성부는 상기 레이블링 파일을 상기 입력 데이터와 동시에 생성하는 영상 학습 장치.

청구항 4

제3항에 있어서,
상기 레이블링 파일은 상기 객체 이미지의 위치를 나타낸 마스크 파일과 상기 객체 이미지의 위치를 기록한 텍스트 파일을 포함하는 영상 학습 장치.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서,
상기 입력 데이터는 하나의 상기 객체 이미지에 대해 서로 다른 각도의 객체 이미지들, 서로 다른 크기의 객체 이미지들, 서로 다른 밝기의 배경 이미지들을 포함하는 영상 학습 장치.

발명의 설명

기술 분야

[0001] 실시예는 객체 인식을 수행하기 위한 영상 학습 장치에 관한 것이다.

배경 기술

[0002] 근래에 들어 딥러닝은 얼굴 인식, 전신 인식, 자세 인식, 음성 인식, 객체 인식, 데이터 마이닝 등 다양한 인식에 적용되고 있다. 특히, 영상으로부터 특정 객체를 인식하는 객체 인식과 딥러닝 학습 네트워크를 접목시키는 연구는 다양한 방식으로 이루어지고 있다.

[0003] 객체를 효과적으로 인식하기 위해서는 대량의 데이터를 사용해야 좋은 학습 결과를 얻을 수 있다는 점이 알려져 있다. 하지만, 충분한 양의 데이터를 수집하기 위해서는 많은 비용과 시간을 필요로 하게 된다.

[0004] 또한, 이미지 수집 후에는 이미지 상의 객체들의 위치와 종류를 표시하는 레이블 과정이 필요하나, 종래에는 이를 수작업으로 진행하여 비용 및 시간이 더 소요되는 문제점이 있다.

발명의 내용

해결하려는 과제

- [0005] 상술한 문제점을 해결하기 위해 실시예는 시간 및 비용을 절감하면서 양질의 데이터를 수집하기 위한 영상 학습 장치를 제공하는 것을 그 목적으로 한다.
- [0006] 또한, 실시예는 객체 인식 효율을 향상시키기 위한 영상 학습 장치를 제공하는 것을 그 목적으로 한다.

과제의 해결 수단

- [0007] 실시예에 따른 영상 학습 장치는 레퍼런스 이미지로부터 객체 이미지를 추출하는 객체 추출부와, 상기 객체 이미지에 서로 다른 배경 이미지들을 각각 합성하여 입력 데이터를 생성하는 데이터 생성부와, 인공 신경망을 이용하여 상기 입력 데이터를 학습시키는 학습부를 포함할 수 있다.
- [0008] 상기 객체 이미지는 상기 레퍼런스 이미지로부터 배경 제거 알고리즘을 사용하여 상기 배경 이미지를 제거하고, 상기 배경 제거 알고리즘은 상기 객체 이미지의 윤곽선을 검출하고, 상기 윤곽선의 바깥쪽 영역의 투명도를 최대로 설정하여 상기 배경 이미지를 제거할 수 있다.
- [0009] 상기 입력 데이터는 레이블링 파일을 더 포함하고, 상기 데이터 생성부는 상기 레이블링 파일을 상기 입력 데이터와 동시에 생성할 수 있다.
- [0010] 상기 레이블링 파일은 상기 객체 이미지의 위치를 나타낸 마스크 파일과 상기 객체 이미지의 위치를 기록한 텍스트 파일을 포함할 수 있다.
- [0011] 상기 입력 데이터는 하나의 상기 객체 이미지에 대해 서로 다른 각도의 객체 이미지들, 서로 다른 크기의 객체 이미지들, 서로 다른 밝기의 배경 이미지들을 포함할 수 있다.

발명의 효과

- [0012] 실시예는 객체의 이미지를 조작하여 다양한 이미지를 생성함으로써, 시간 및 비용을 효과적으로 줄일 수 있는 효과가 있다.
- [0013] 또한, 실시예는 학습될 학습 데이터를 형성함과 동시에 레이블링 파일을 형성함으로써, 데이터 생성에 소요되는 시간을 현저하게 줄일 수 있는 효과가 있다.
- [0014] 또한, 실시예는 다양한 학습 데이터를 효과적으로 수집함으로써, 객체 인식 성능을 향상시킬 수 있는 효과가 있다.

도면의 간단한 설명

- [0015] 도 1은 실시예에 따른 영상 학습 장치를 개략적으로 나타낸 블록도이다.
- 도 2는 실시예에 따른 영상 학습 장치의 배경 제거 알고리즘을 도시한 도면이다.
- 도 3은 실시예에 따른 영상 학습 장치의 배경 제거 알고리즘을 설명하기 위한 도면이다.
- 도 4는 도 2의 배경 제거 알고리즘이 적용된 객체 이미지들을 나타낸 도면이다.
- 도 5는 하나의 객체 이미지에 대해 서로 다른 각도의 객체 이미지들을 나타낸 도면이다.
- 도 6은 실시예에 따른 합성 기법에 의한 입력 데이터를 생성하는 과정을 나타낸 도면이다.
- 도 7은 실시예에 따른 합성 기법에 의한 생성된 최종 입력 데이터들을 나타낸 도면이다.
- 도 8은 실시예에 따른 합성 기법에 의한 생성된 레이블링 파일을 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0016] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어

야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.

- [0017] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [0018] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.
- [0019] 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 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.
- [0020] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 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.
- [0022] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시예를 상세하게 설명한다.
- [0023] 도 1은 실시예에 따른 영상 학습 장치를 개략적으로 나타낸 블록도이고, 도 2는 실시예에 따른 영상 학습 장치의 배경 제거 알고리즘을 도시한 도면이고, 도 3은 실시예에 따른 영상 학습 장치의 배경 제거 알고리즘을 설명하기 위한 도면이고, 도 4는 도 2의 배경 제거 알고리즘이 적용된 객체 이미지들을 나타낸 도면이고, 도 5는 하나의 객체 이미지에 대해 서로 다른 각도의 객체 이미지들을 나타낸 도면이고, 도 6은 실시예에 따른 합성 기법에 의한 입력 데이터를 생성하는 과정을 나타낸 도면이고, 도 7은 실시예에 따른 합성 기법에 의한 생성된 최종 입력 데이터들을 나타낸 도면이고, 도 8은 실시예에 따른 합성 기법에 의한 생성된 레이블링 파일을 나타낸 도면이다.
- [0024] 도 1 내지 도 3을 참조하면, 실시예에 따른 영상 학습 장치(100)는 레퍼런스 이미지로부터 객체 이미지를 추출하는 객체 추출부(110)와, 상기 객체 이미지에 서로 다른 배경 이미지들을 각각 합성하여 입력 데이터를 생성하는 데이터 생성부(120)와, 인공 신경망을 이용하여 상기 입력 데이터를 학습시키는 학습부(130)를 포함할 수 있다.
- [0025] 객체 추출부(110)는 배경 이미지(20) 및 객체 이미지(10)가 포함된 레퍼런스 이미지로부터 배경 이미지(20)를 제거하여 객체 이미지(10)를 추출할 수 있다. 레퍼런스 이미지는 카메라 또는 센서에 의해 촬영될 수 있다. 레퍼런스 이미지는 우수한 인식 성능을 위해 다양한 면을 가지도록 다양한 각도에서 촬영된 이미지일 수 있다.
- [0026] 레퍼런스 이미지는 효과적으로 배경 이미지(20)를 제거하기 위해 PNG(Portable Network Graphics) 형태의 이미지를 포함할 수 있다.
- [0027] 객체 추출부(110)는 배경 제거 알고리즘을 이용하여 레퍼런스 이미지의 배경 이미지(20)를 제거할 수 있다. 객체 추출부(110)는 투명도를 조정하여 객체 이미지(10)만 남기고 배경 이미지(20)를 제거할 수 있다. 배경 제거 알고리즘은 도 2에 도시된 알고리즘을 사용할 수 있다.
- [0028] 먼저, 레퍼런스 이미지를 흑백 이미지로 변경할 수 있다. 흑백 이미지로 변환된 이미지는 도 3과 같다.
- [0029] 이어서, 배경 제거 알고리즘을 이용하여 1번 방향을 따라서 미리 정해 놓은 임계값(Threshold) 보다 큰 첫 번째 픽셀의 위치를 저장할 수 있다. 여기서 임계값은 10 내지 20의 값을 사용할 수 있다.
- [0030] 마찬가지로 배경 제거 알고리즘을 이용하여 2번 방향, 3번 방향 및 4번 방향을 따라 배경 제거 알고리즘을 이용

하여 픽셀의 위치를 저장할 수 있다.

- [0031] 상기와 같이 저장된 픽셀의 위치를 이용하여 객체의 윤곽선을 검출할 수 있다. 즉, 저장된 픽셀의 위치를 서로 연결하게 되면 윤곽선이 형성될 수 있다. 윤곽선은 객체 이미지(10)와 배경 이미지(20)를 구분할 수 있게 된다.
- [0032] 윤곽선의 안쪽 영역은 투명도를 최소로하고, 윤곽선의 바깥 영역은 투명도를 최대로 할 수 있다. 즉, 객체 이미지(10) 영역의 투명도는 최소로 하고, 배경 이미지 영역(20)의 투명도는 최대로 할 수 있다.
- [0033] 윤곽선 바깥 영역은 투명도가 최대로 형성됨으로써, 배경 이미지(20)를 효과적으로 제거할 수 있다.
- [0034] 이에 따라 객체 이미지(10)는 도 4에 도시된 바와 같이, 실제 객체의 크기와 동일하게 생성될 수 있게 된다.
- [0035] 이와 마찬가지로 다양한 각도의 이미지를 포함하는 레퍼런스 이미지들에 배경 제거 알고리즘을 이용하여 배경 이미지(20)를 제거하게 되면 도 5에 도시된 바와 같이, 다양한 각도에서의 객체 이미지(10)들을 얻을 수 있게 된다.
- [0036] 데이터 생성부(120)는 학습에 필요한 입력 데이터를 생성시키는 역할을 한다. 데이터 생성부(120)는 다양한 각도의 이미지를 포함하는 객체 이미지(10)를 다양한 환경의 배경 이미지와 합성하여 입력 데이터를 생성할 수 있다. 여기서, 배경 이미지는 별도의 구성 예컨대 배경 생성부에 의해 배경을 수집할 수 있다.
- [0037] 배경 이미지는 다양한 환경의 이미지를 포함할 수 있다. 이로 인해 입력 데이터는 하나의 객체에 대해 다양한 배경 이미지를 포함할 수 있다.
- [0038] 데이터 생성부(120)는 객체 이미지(10)와 다양한 환경의 배경 이미지를 합성하는 동시에 객체 이미지를 변화시키는 이미지 조작을 통해 다양한 환경에서의 객체 이미지를 수집할 수 있다. 데이터 생성부는 객체 이미지와 배경 이미지를 회전시키거나 크기를 변화시키거나 밝기를 변화시킬 수 있다.
- [0039] 도 6에 도시된 바와 같이, 먼저, 합성할 객체 이미지(10)와 다양한 환경의 배경 이미지(30)를 랜덤으로 선택하게 된다. 여기서, 다양한 환경의 배경 이미지는 레퍼런스 이미지의 배경 이미지와 다른 배경 이미지일 수 있다. 이에, 다양한 환경의 배경 이미지는 레퍼런스 이미지의 배경 이미지와 다른 도면 부호를 부여한다.
- [0040] 객체 이미지(10)는 수집된 객체 이미지(10) 중에 선택할 수 있다. 예를 들어 하나의 합성 이미지에 객체 3개가 포함되도록 하고 싶다면 n개의 개체 중 3개의 객체 이미지를 랜덤으로 고르도록 할 수 있다. 이때, 객체 이미지(10)의 중복 선택을 허용하게 하여 동일한 객체 이미지(10)가 여러 번 선정될 수 있다. 배경 이미지(30)는 m개의 배경 이미지(20) 중 랜덤으로 선택할 수 있다.
- [0041] 객체 이미지(10)와 배경 이미지(30)가 선정되면 이미지 조작을 통해 다양한 이미지를 생성시킬 수 있다. 이미지 조작은 객체 이미지(10)의 회전, 크기를 변화시켜 다양한 형태의 객체 10이미지()를 생성할 수 있다 또한, 배경 30이미지()도 밝기를 변화시켜 다양한 형태의 배경 이미지(30)를 생성할 수 있다.
- [0042] 예컨대, 객체 이미지(10)를 회전시킬 경우 0도~360도 범위 내에서 객체 이미지(10)를 회전시켜 다양한 각도의 이미지가 생성시킬 수 있다. 객체 이미지(10)의 크기를 변화시킬 경우, 원하는 픽셀 범위를 정하고 객체 이미지(10)의 장축이 정해진 크기를 만족하도록 조작할 수 있다.
- [0043] 객체 이미지(10)에 대해 회전과 크기가 정해지면 객체 이미지(10)의 크기를 고려하여 합성 위치를 랜덤으로 정하게 된다. 이 과정을 통해 배경 이미지(30)에 객체 이미지(10)가 합성되면 밝기 조작을 통해 예컨대 기존 대비 50% 내지 150%의 밝기를 가지도록 하여 도 7에 도시된 바와 같이, 최종 합성 이미지를 획득할 수 있다.
- [0044] 데이터 생성부(120)는 합성 이미지 외에 객체의 레이블링 파일을 동시에 생성할 수 있다. 도 8에 도시된 바와 같이, 레이블링 파일로는 객체의 위치를 나타낸 복수의 마스크 파일(40)을 포함할 수 있다. 레이블링 파일로는 객체의 위치를 기록한 텍스트 파일을 포함할 수 있다.
- [0045] 데이터 생성부(120)는 마스크 파일(40)과 텍스트 파일을 생성하고, 이를 객체 인식을 위한 학습 데이터로 사용할 수 있다.
- [0046] 상기와 같이, 20종의 객체 인식을 위한 학습 데이터를 생성한다고 했을 때, 객체 이미지 480장(객체 당 24장)을 이용하여 30,000장의 입력 데이터를 생성하는 데 걸리는 시간은 약 20분정도 걸릴 수 있다. 이때, 한 이미지 당 객체가 4~5개가 합성되므로 총 135,000개의 객체가 포함된 양의 데이터이며, 이를 수잡업으로 레이블링 할 경우 약 135시간이 소요될 수 있다.
- [0047] 따라서, 실시예에서 제시된 방법으로 학습할 입력 데이터를 생성할 경우 400배의 효율을 얻을 수 있는 효과가

있다.

[0048] 학습부는 수집된 입력 데이터를 이용하여 학습을 수행할 수 있다. 학습부는 DNN(Deep Neural Network) 및 CNN(Convolutional Neural Network)과 같은 다양한 종류의 인공 신경망을 포함할 수 있다.

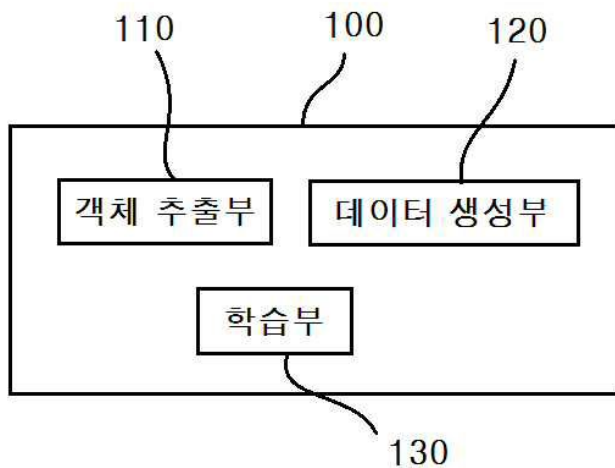
[0049] 따라서, 실시예는 객체 인식 성능을 향상시킬 수 있는 효과가 있다.

부호의 설명

[0050] 10: 객체 이미지
20: 배경 이미지
110: 객체 추출부
120: 데이터 생성부
130: 학습부

도면

도면1



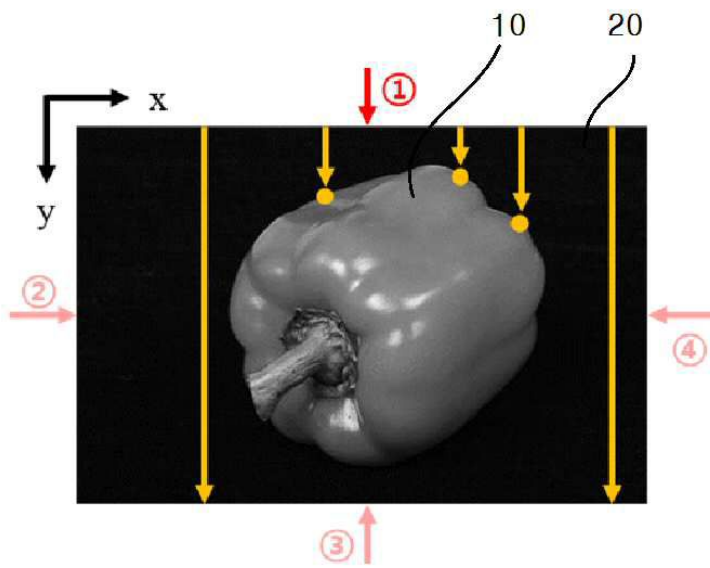
도면2

Algorithm 1 Background removal algorithm

```

1: Input: the image of object
2: rows  $\leftarrow$  the width of image
3: cols  $\leftarrow$  the height of image
4: points  $\leftarrow$  an empty list
5:
6: for y  $\leftarrow$  1 to rows do
7:   for x  $\leftarrow$  1 to cols do
8:     if  $P_{x,y} > \text{threshold}$  then
9:       points.append((x,y))
10: return points
    
```

도면3



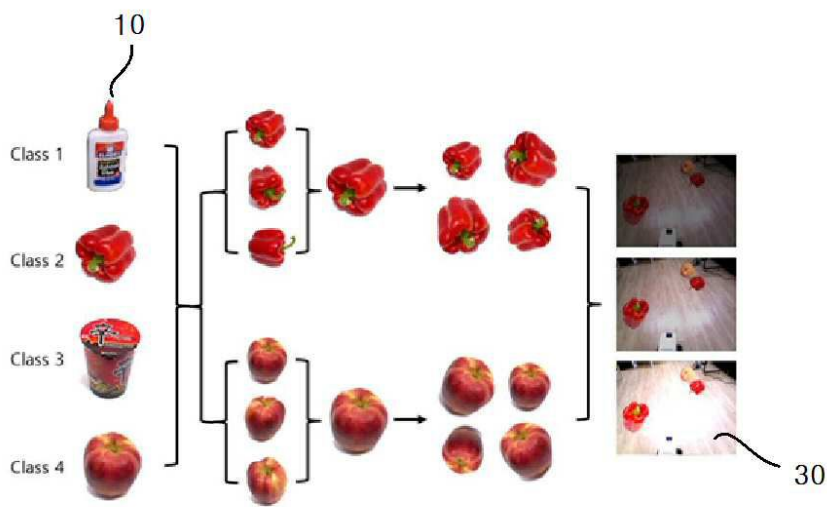
도면4



도면5



도면6



도면7



도면8

