(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0062959
(43) 공개일자 2023년05월09일

(51) 국제특허분류(Int. Cl.)
G06Q 50/26 (2012.01) *G06F 18/00* (2023.01)
G06N 3/08 (2023.01) *G06T 7/246* (2017.01)
G06T 7/60 (2017.01)
(52) CPC특허분류
G06Q 50/26 (2013.01)
G06N 3/08 (2023.01)
(21) 출원번호 10-2021-0147676
(22) 출원일자 2021년11월01일
심사청구일자 2021년11월01일

(71) 출원인
인천대학교 산학협력단
인천광역시 연수구 아카데미로 119 (송도동)
(72) 발명자
전광길
인천광역시 연수구 해송로 70, 209동 805호(송도동, 송도웰카운티2단지)
(74) 대리인
특허법인충정

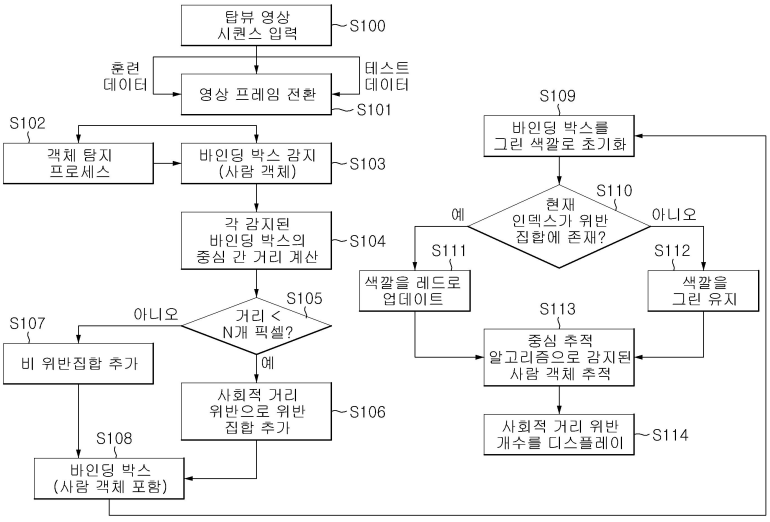
전체 청구항 수 : 총 9 항

(54) 발명의 명칭 **딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템 및 방법**

(57) 요 약

딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템 및 방법은 영상 시퀀스에서 사회적 거리두기 위반을 자동으로 감지하고, 감지된 위반 데이터를 활용하여 바이러스 확산에 대한 예방 조치를 실행할 수 있다.

대 표 도



(52) CPC특허분류
G06T 7/246 (2017.01)
G06T 7/60 (2013.01)
G06V 20/46 (2022.01)
G06V 20/53 (2022.01)
G06T 2207/20084 (2013.01)

명 세 서

청구범위

청구항 1

탑뷰에 설치된 카메라부를 이용하여 탑뷰 영상 프레임을 생성하는 탑뷰 영상 생성부;

사전 훈련된 데이터셋에서 영상 프레임의 제1 특징맵을 추출하고, 추출한 제1 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출하고, 상기 탑뷰 영상 프레임의 제2 특징맵을 추출하고, 상기 제1 특징맵과 상기 제2 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출하여 바운딩 박스를 예측하는 객체 탐지부;

상기 바운딩 박스를 감지하여 바운딩 박스 중심을 계산하고, 상기 감지된 각 바운딩 박스 중심 쌍 사이의 거리를 픽셀의 개수로 계산하는 거리 계산부; 및

상기 계산한 거리가 기설정된 임계값인 N개 픽셀의 개수보다 작은지 판단하고, 상기 계산한 거리가 상기 N개 픽셀의 개수보다 작은 경우, 사회적 거리 위반으로 판단하여 위반 집합에 추가하는 제어부를 포함하는 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템.

청구항 2

청구항 1에 있어서,

상기 제어부는 상기 사회적 거리 위반의 개수를 카운트하여 누적하고, 상기 누적된 사회적 거리 위반의 개수가 기설정된 기준 사회적 거리 위반수치와 비교하여 넘어서는 경우, 위험 신호를 생성하여 전송하는 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템.

청구항 3

청구항 1에 있어서,

상기 제어부는 상기 감지한 바운딩 박스 중심을 유클리드 거리(Euclidean Distance)를 이용하여 계산하고, 상기 바운딩 박스 중심이 계산되면, 해당 바운딩 박스에 인덱스를 각각 할당하고, 상기 할당한 인덱스를 상기 위반 집합에 추가하여 관리하는 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템.

청구항 4

청구항 3에 있어서,

상기 제어부는 바운딩 박스를 그린(Green) 색깔로 초기화하고, 상기 인덱스가 상기 위반 집합에 존재하는지 판단하고, 상기 인덱스가 상기 위반 집합에 존재하는 경우, 상기 바운딩 박스의 색깔을 레드(Red)로 업데이트하고, 상기 인덱스가 상기 위반 집합에 존재하지 않는 경우, 상기 바운딩 박스의 색깔을 그린으로 유지하는 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템.

청구항 5

청구항 1에 있어서,

상기 제어부는 상기 사회적 거리 위반으로 판단된 경우, 상기 바운딩 박스의 색깔을 레드(Red)로 업데이트하고, 상기 사회적 거리 위반으로 판단되지 않은 경우, 상기 바운딩 박스의 색깔을 그린으로 유지하고

상기 사회적 거리 위반의 개수를 카운트하여 누적하고, 상기 누적된 사회적 거리 위반의 개수를 디스플레이부에 표시하며, 상기 누적된 사회적 거리 위반의 개수가 기설정된 기준 사회적 거리 위반수치와 비교하여 넘어서는 경우, 위험 신호를 생성하여 외부의 중앙 서버로 전송하는 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템.

청구항 6

청구항 5에 있어서,

상기 탑뷰 영상 생성부, 상기 객체 탐지부, 상기 거리 계산부, 상기 제어부로 이루어진 사회적 거리 모니터링 장치는 여러 지역에 각각 설치되고,

상기 중앙 서버는 상기 각각의 사회적 거리 모니터링 장치로부터 상기 위험 신호, 상기 누적된 사회적 거리 위반의 개수를 포함한 사회적 거리 모니터링 정보를 수신하여 지역별로 데이터베이스하는 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템.

청구항 7

탑뷰 영상 생성부는 탑뷰에 설치된 카메라부를 이용하여 탑뷰 영상 프레임을 생성하는 단계;

객체 탐지부는 사전 훈련된 데이터셋에서 영상 프레임의 제1 특징맵을 추출하고, 추출한 제1 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출하고, 상기 탑뷰 영상 프레임의 제2 특징맵을 추출하고, 상기 제1 특징맵과 상기 제2 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출하여 바운딩 박스를 예측하는 단계; 및

거리 계산부는 상기 바운딩 박스를 감지하여 바운딩 박스 중심을 계산하고, 상기 감지된 각 바운딩 박스 중심 쌍 사이의 거리를 픽셀의 개수로 계산하는 단계;

제어부는 상기 계산한 거리가 기설정된 임계값인 N개 픽셀의 개수보다 작은지 판단하고, 상기 계산한 거리가 상기 N개 픽셀의 개수보다 작은 경우, 사회적 거리 위반으로 판단하여 위반 집합에 추가하는 단계를 포함하는 딥러닝 기반의 사회적 거리두기를 모니터링하는 방법.

청구항 8

청구항 7에 있어서,

상기 제어부는 상기 사회적 거리 위반의 개수를 카운트하여 누적하는 단계; 및

상기 누적된 사회적 거리 위반의 개수가 기설정된 기준 사회적 거리 위반수치와 비교하여 넘어서는 경우, 위험 신호를 생성하여 외부로 전송하거나 색깔을 다르게 하여 표시하는 단계를 포함하는 딥러닝 기반의 사회적 거리두기를 모니터링하는 방법.

청구항 9

청구항 7에 있어서,

상기 제어부는 상기 사회적 거리 위반으로 판단된 경우, 상기 바운딩 박스의 색깔을 레드(Red)로 업데이트하고, 상기 사회적 거리 위반으로 판단되지 않은 경우, 상기 바운딩 박스의 색깔을 그린으로 유지하는 단계; 및

상기 제어부는 상기 사회적 거리 위반의 개수를 카운트하여 누적하고, 상기 누적된 사회적 거리 위반의 개수를 디스플레이부에 표시하는 단계; 및

상기 제어부는 상기 누적된 사회적 거리 위반의 개수가 기설정된 기준 사회적 거리 위반수치와 비교하여 넘어서는 경우, 위험 신호를 생성하여 외부의 중앙 서버로 전송하는 단계를 포함하는 딥러닝 기반의 사회적 거리두기를 모니터링하는 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템 및 방법에 관한 것으로서, 더욱 상세하게는 영상 시퀀스에서 사회적 거리두기 위반을 자동으로 감지하고, 감지된 위반 데이터를 활용하여 바이러스 확산에 대한 예방 조치를 실행할 수 있는 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템 및 방법에 관한 것이다.

배 경 기 술

[0002] COVID-19 코로나 바이러스 발병은 치명적인 확산과 함께 전 세계적인 재앙을 일으켰다. COVID-19는 효과적인 치

료제의 부재와 바이러스에 대한 예방 접종의 부족으로 인해 인구 취약성이 증가한다. 현재 상황에서는 사용할 수 있는 백신이 없다.

[0003] 따라서, 사회적 거리두기는 대유행 바이러스의 확산에 대한 적절한 예방 조치(규범)로 생각된다. 사람들 사이의 물리적 접촉을 피함으로써 바이러스 확산의 위험을 최소화할 수 있다.

[0004] 2019년 12월 말 이후 COVID-19 대유행의 부상 이후, 사회적 거리두기는 전염성 바이러스 전파를 예방하기 위한 가장 신뢰할 수 있는 관행으로 간주되어 2020년 1월 23일에 표준 관행으로 채택되었다.

[0005] 그러나 COVID-19 발생 이후에 2년이 되는 시점에도 바이러스 유행이 끝나지 않고 있다.

[0006] 코로나 발생 환자는 2천명 정도되는 상황이 6개월 이상 지속되고 있으며, 지역 경제에 회복될 수 없을 정도로 피해가 막심하다.

[0007] 이러한 문제를 해결하기 위하여 최근 정부는 코로나 바이러스 감염증 방역을 위해 실시해 온 사회적 거리두기를 단계적으로 완화하여 일상을 회복하는 방향으로 위드 코로나로 전환을 검토하고 있다.

[0008] 그러나 위드 코로나와 사회적 거리두기는 양립하기 어려운 문제이므로 위드 코로나 전환 후, 바이러스 환자가 급증하게 된다면, 사회적 거리두기를 어느 정도로 실행해야 하는지 다시 고민해야 문제점이 발생할 수도 있다.

선행기술문헌

특허문헌

[0009] (특허문헌 0001) 한국 등록특허번호 제10-2230200호

발명의 내용

해결하려는 과제

[0010] 이와 같은 문제점을 해결하기 위하여, 본 발명은 영상 시퀀스에서 사회적 거리두기 위반을 자동으로 감지하고, 감지된 위반 데이터를 활용하여 바이러스 확산에 대한 예방 조치를 실행할 수 있는 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템 및 방법을 제공하는데 그 목적이 있다.

과제의 해결 수단

[0011] 상기 목적을 달성하기 위한 본 발명의 특징에 따른 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템은,

[0012] 탑뷰에 설치된 카메라부를 이용하여 탑뷰 영상 프레임을 생성하는 탑뷰 영상 생성부;

[0013] 사전 훈련된 데이터셋에서 영상 프레임의 제1 특징맵을 추출하고, 추출한 제1 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출하고, 상기 탑뷰 영상 프레임의 제2 특징맵을 추출하고, 상기 제1 특징맵과 상기 제2 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출하여 바운딩 박스를 예측하는 객체 탐지부;

[0014] 상기 바운딩 박스를 감지하여 바운딩 박스 중심을 계산하고, 상기 감지된 각 바운딩 박스 중심 쌍 사이의 거리를 픽셀의 개수로 계산하는 거리 계산부; 및

[0015] 상기 계산한 거리가 기설정된 임계값인 N개 픽셀의 개수보다 작은지 판단하고, 상기 계산한 거리가 상기 N개 픽셀의 개수보다 작은 경우, 사회적 거리 위반으로 판단하여 위반 집합에 추가하는 제어부를 포함한다.

[0017] 본 발명의 특징에 따른 딥러닝 기반의 사회적 거리두기를 모니터링하는 방법은,

[0018] 탑뷰 영상 생성부는 탑뷰에 설치된 카메라부를 이용하여 탑뷰 영상 프레임을 생성하는 단계;

[0019] 객체 탐지부는 사전 훈련된 데이터셋에서 영상 프레임의 제1 특징맵을 추출하고, 추출한 제1 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출하고, 상기 탑뷰 영상 프레임의 제2 특징맵을 추출하고, 상기 제1 특징맵과 상기 제2 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의

영역을 추출하여 바운딩 박스를 예측하는 단계; 및

[0020] 거리 계산부는 상기 바운딩 박스를 감지하여 바운딩 박스 중심을 계산하고, 상기 감지된 각 바운딩 박스 중심 쌍 사이의 거리를 픽셀의 개수로 계산하는 단계;

[0021] 제어부는 상기 계산한 거리가 기설정된 임계값인 N개 픽셀의 개수보다 작은지 판단하고, 상기 계산한 거리가 상기 N개 픽셀의 개수보다 작은 경우, 사회적 거리 위반으로 판단하여 위반 집합에 추가하는 단계를 포함한다.

발명의 효과

[0022] 전술한 구성에 의하여, 본 발명은

[0023] 영상 시퀀스에서 사회적 거리두기 위반을 자동으로 감지하여 경고함으로써 바이러스 확산에 대한 예방 조치를 미리 구현할 수 있는 효과가 있다.

도면의 간단한 설명

[0024] 도 1은 본 발명의 실시예에 따른 COVID-19를 위한 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템의 구성을 나타낸 도면이다.

도 2는 본 발명의 실시예에 따른 사회적 거리 모니터링 장치의 내부 구성을 간략하게 나타낸 블록도이다.

도 3은 본 발명의 실시예에 따른 탑뷰 영상에서 사람 객체를 탐지하는 YOLOv3의 구조를 간략하게 나타낸 도면이다.

도 4는 본 발명의 실시예에 따른 사회적 거리 모니터링 방법을 설명하기 위한 도면이다.

도 5는 본 발명의 실시예에 따른 감지된 바운딩 박스의 사람 객체 좌표를 나타낸 도면이다.

도 6은 본 발명의 실시예에 따른 바운딩 박스의 각 중심 쌍 사이의 거리를 빨간색 선으로 나타낸 도면이다.

도 7은 본 발명의 실시예에 따른 사회적 거리 모니터링 결과를 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

[0025] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.

[0026] 도 1은 본 발명의 실시예에 따른 COVID-19를 위한 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템의 구성을 나타낸 도면이다.

[0027] 도 1은 본 발명의 실시예에 따른 COVID-19를 위한 딥러닝 기반의 사회적 거리두기를 모니터링하는 시스템은 지역별로 설치된 사회적 거리 모니터링 장치(100), 통신망(101) 및 중앙 서버(200)를 포함한다.

[0028] 사회적 거리 모니터링 장치(100)는 탑뷰에 설치된 카메라부를 이용하여 탑뷰 영상 프레임을 생성한다.

[0029] 사회적 거리 모니터링 장치(100)는 사전 훈련된 데이터셋에서 영상 프레임의 제1 특징맵을 추출하고, 추출한 제1 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출하고, 탑뷰 영상 프레임의 제2 특징맵을 추출하고, 제1 특징맵과 제2 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출하여 바운딩 박스를 예측한다.

[0030] 사회적 거리 모니터링 장치(100)는 바운딩 박스를 감지하여 바운딩 박스 중심을 계산하고, 감지된 각 바운딩 박스 중심 쌍 사이의 거리를 픽셀의 개수로 계산한다.

[0031] 사회적 거리 모니터링 장치(100)는 계산한 거리가 기설정된 임계값인 N개 픽셀의 개수보다 작은지 판단하고, 계산한 거리가 상기 N개 픽셀의 개수보다 작은 경우, 사회적 거리 위반으로 판단하여 위반 집합에 추가한다.

[0032] 사회적 거리 모니터링 장치(100)는 사회적 거리 위반의 개수를 카운트하여 누적하고, 누적된 사회적 거리 위반의 개수가 기설정된 기준 사회적 거리 위반수치와 비교하여 넘어서는 경우, 위험 신호를 생성하여 통신망(101)을 통해 중앙 서버(200)로 전송한다.

[0033] 중앙 서버(200)는 각각의 사회적 거리 모니터링 장치(100)로부터 위험 신호, 누적된 사회적 거리 위반의 개수를 포함한 사회적 거리 모니터링 정보를 수신하여 지역별로 데이터베이스한다.

[0035] 도 2는 본 발명의 실시예에 따른 사회적 거리 모니터링 장치의 내부 구성을 간략하게 나타낸 블록도이고, 도 3은 본 발명의 실시예에 따른 탑뷰 영상에서 사람 객체를 탐지하는 YOLOv3의 구조를 간략하게 나타낸 도면이고, 도 4는 본 발명의 실시예에 따른 사회적 거리 모니터링 방법을 설명하기 위한 도면이다.

[0036] 본 발명의 실시예에 따른 사회적 거리 모니터링 장치(100)는 탑뷰 영상 생성부(110), 영상 프레임 입력부(120), COCO 데이터셋 데이터베이스부(130), 객체 탐지부(140), 전이 학습 모델부(150), 제어부(160), 거리 계산부(170), 디스플레이부(180), 통신부(190), 데이터베이스부(192) 및 객체 추적부(194)를 포함한다.

[0037] 탑뷰 영상 생성부(110)는 탑뷰에 설치된 카메라부를 이용하여 탑뷰 영상 시퀀스를 생성한다(S100). 여기서, 탑뷰(Top View)는 상부에서 내려다 보는 시야이다.

[0038] 영상 프레임 입력부(120)는 탑뷰 영상 생성부(110)로부터 수신한 탑뷰 영상 시퀀스가 훈련 데이터 또는 테스트 데이터인지 판단하여 영상 프레임으로 변환한다(S101).

[0039] COCO 데이터셋 데이터베이스부(130)는 사전 훈련된 데이터셋으로 객체 탐지, 세그먼테이션, 키포인트 탐지 등의 컴퓨터 비전 분야의 Task를 목적으로 만들어진 데이터셋이며, 학습 데이터셋, 검증 데이터셋, 테스트 데이터셋을 저장하고 있다.

[0040] 객체 탐지부(140)는 COCO 데이터셋 데이터베이스부(130)로부터 학습 데이터셋, 검증 데이터셋, 테스트 데이터셋을 수집하고, 사전 훈련된 YOLOv3를 이용하여 영상 프레임의 특징맵을 추출하고, 추출한 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출하여 바운딩 박스로 표시한다.

[0041] 전이 학습 모델부(150)는 객체 탐지부(140)로부터 사전 학습 완료된 학습 모델로부터 학습된 결과에 대한 가중치를 생성하고, 생성한 가중치를 객체 탐지부(140)로 전송한다. 가중치는 객체 탐지부(140)에서 YOLOv3를 이용하여 사람 객체를 탐지할 때 사용한다.

[0042] 전이 학습 모델부(150)는 객체 탐지부(140)에서 출력된 특징맵이 COCO 데이터셋 데이터베이스부(210)에서 사전 학습 완료된 특징맵을 닮아가도록 전이학습을 수행한다.

[0043] 객체 탐지부(140)는 COCO 데이터셋 데이터베이스부(130), 전이 학습 모델부(150)를 이용하여 영상 프레임에서 사람 객체를 추출하고, 바운딩 박스로 표시한다(S102, S103).

[0044] 객체 탐지부(140)는 도 3과 같이, 객체 분류를 위해 Softmax가 아닌 다양한 독립 로지스틱을 사용하고, 컨볼루션 레이어를 이용하여 특징 학습이 수행한다.

[0045] 객체 탐지부(140)는 복수의 컨볼루션 레이어와 스킵 연결로 구성되어 3개의 개별 규모에서 특징을 감지하고, 컨볼루션 레이어에 의해 특징맵을 다운 샘플링하고, 불변 크기 특징을 전송하기 위해 실행된다.

[0046] 객체 탐지부(140)는 영상 프레임에서 객체가 존재할 것으로 추정되는 영역을 추출하고, 추출된 영역으로부터 특징을 나타내는 특징맵을 추출한다.

[0047] 객체 탐지부(140)는 추출한 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출한다. 영역을 추출하는 방법은 예를 들어 faster RCNN, SSD(Single Shot MultiBox Detector), YOLO(You Only Look Once) 등이 있을 수 있으며, 본 발명은 YOLO 객체 인식모듈(YOLOv3)을 일례로 하고 있다.

[0048] 객체 탐지부(140)는 특징맵 중에서 영상의 영역별 클래스의 좌표를 포함하는 특징맵을 선정하고, 선정된 특징맵으로부터 영역을 구별하는 좌표를 식별한 뒤, 식별된 좌표를 개체의 존재가 추정되는 영역으로 추출할 수 있다.

[0049] 객체 탐지부(140)는 COCO 데이터셋 데이터베이스부(130)의 사전 훈련된 데이터셋에서 영상 프레임의 제1 특징맵을 추출하고, 추출한 제1 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출한다.

[0050] 객체 탐지부(140)는 탑뷰 영상 프레임의 제2 특징맵을 추출하고, 제1 특징맵과 제2 특징맵을 기초로 영상에서 사람 객체의 존재가 추정되는 적어도 하나의 영역을 추출하여 바운딩 박스를 예측한다.

[0051] 객체 탐지부(140)는 추출된 적어도 하나의 영역 각각에 대해서, 해당 객체의 최외곽을 둘러싸는 바운딩 박스(Bounding Box)로서 표시할 수 있다.

[0052] 각각의 바운딩 박스는 영상에서 해당 바운딩 박스의 위치에 개체의 존재 가능성이 있음을 나타낸다.

[0053] YOLOv3 모델은 단일 네트워크 아키텍처를 사용하여 전체 입력 영상에 대한 클래스 확률과 해당 바운딩 박스를 예측한다.

[0054] YOLOv3 모델은 24개의 컨볼루션 레이어와 2개의 완전히 연결된 레이어가 포함되어 있다. 컨볼루션 레이어는 영상 추출에 사용되는 반면 완전 연결 레이어는 클래스 예측과 확률을 계산한다.

[0055] 객체 탐지부(140)는 사람 객체를 포함한 바운딩 박스를 감지한다.

[0056] 객체 탐지부(140)는 감지된 바운딩 박스의 정도를 사람으로 나타내는 신뢰도 값(Conf(p))을 수학식 1로 정의한다.

수학식 1

[0057]
$$Conf(p) = \Pr(p) \times IOU(pred, actual)$$

[0058] 여기서, Pr(p)는 예측된 바운딩 박스에 사람이 있는지 여부(예: 1, 아니오: 0)를 나타낸다.

[0059] IOU(pred, actual)는 실제 바운딩 박스와 예측된 바운딩 박스의 IOU를 결정한다.

[0060] IOU(pred, actual)는 예측된 바운딩 박스와 실제 바운딩 박스의 겹치는 넓이를 두 영역을 합친 넓이로 나눈 IoU 중첩 비율이다.

[0061] 도 5와 같이, 객체 탐지부(140)는 사람 객체가 식별하는 동안, 입력 영상 프레임을 그리드 셀이라고도 하는 S × S 영역으로 나눈다. 이러한 그리드 셀은 바운딩 박자 예측 및 클래스 확률과 관련된다. 객체 탐지부(140)는 각 셀이 사람의 중심이 격자 셀에 있는지 여부의 확률을 예측한다.

[0062] 객체 탐지부(140)는 탑뷰 영상 프레임에서 사람 감지를 위해 적합한 영역이 선택되어 예측되고, 예측 후 신뢰도 값을 사용하여 원하는 바운딩 박스를 획득한다.

[0063] 객체 탐지부(140)는 예측이 양성이면 바운딩 박스와 각 양성 검출에 대한 신뢰도 값이 예측된다.

수학식 2

[0064]
$$IoU(pred, actual) = area\frac{BoxT \cap BoxP}{BoxT \cup BoxP}$$

[0065] 여기서, BoxT로 표시된 훈련 데이터 세트에서 수동으로 레이블이 지정된 그라운드 트루(Ground Truth) 박스(실제)이고, 예측된 바운딩 박스는 BoxP로 표시된다. area(영역)은 교차 영역을 나타낸다.

[0066] 그라운드 트루(Ground Truth)는 신경망 모델이 예측한 값이 아닌 실제 정답 레벨(Label)을 의미한다. 즉, 어떤 객체의 실제 위치를 나타낸다.

[0067] IoU은 객체 탐지에서 바운딩 박스로 객체의 위치를 나타내고, 이때 예측한 바운딩 박스와 실제 그라운드 트루 바운딩 박스의 영역을 비교하고, 두 영역을 비교했을 때, 겹치는 넓이를 두 영역을 합친 넓이로 나눈 IoU 중첩 비율이 기설정된 임계값(0.5 이상)일 때 매치되었다고 한다.

[0068] IoU 값이 0.5 이상이면, 해당 Region을 객체로 바라보고, 그라운드 트루와 같은 클래스로 레이블링한다.

[0069] 객체 탐지부(140)는 입력 영상 프레임에서 감지된 각 사람 객체에 대해 허용 영역이 예측되고 결정된다.

[0070] 신뢰도 값은 최적의 바운딩 박스를 달성하기 위해 예측 후에 적용된다.

[0071] 객체 탐지부(140)는 각 예측된 바운딩 박스에 대해 h,w,x,y가 추정된다.

[0072] 바운딩 박스 좌표는 x,y로 정의되고 너비와 높이는 w,h로 결정된다.

[0073] 객체 탐지부(140)는 탑뷰 영상 프레임에서 사람 감지를 위해 적합한 영역이 선택되고 예측되고, 예측 후 신뢰도

값을 사용하여 원하는 바운딩 박스를 획득한다.

[0074] 객체 탐지부(140)는 h, w, x, y 및 신뢰도 값을 포함하여 모든 바운딩 박스에 대해 5개의 값이 예측된다.

[0075] 객체 탐지부(140)는 임계값을 정의하여 낮은 점수의 신뢰도 값을 버리고 나머지 여러 개의 높은 신뢰도 바운딩 박스를 처리한다.

[0076] 객체 탐지부(140)는 도 5 및 하기의 수학식 3과 같은 예측된 바운딩 박스 값을 생성한다. 도 5는 감지된 바운딩 박스의 사람 객체 좌표를 표시한 것이다.

수학식 3

[0077]
$$b_x=\sigma(t_x)+c_x$$
$$b_y=\sigma(t_y)+c_y$$
$$b_w=p_w e^{t_w}$$
$$b_h=p_h e^{t_h}$$

[0078] 여기서, b_x, b_y, b_w, b_h 는 예측된 바운딩 박스 좌표이고, 좌표의 중심은 x,y로 표시되고 너비와 높이는 w,h로 표시된다. t_w, t_h, t_x, t_y 는 네트워크 출력을 정의한다. $\sigma(t_x)$ 와 $\sigma(t_y)$ 는 바운딩 박스 중심에서 하나의 그리드 셀의 X 길이와 Y 길이를 나타낸다.

[0079] c_x, c_y 는 도 5와 같이, 그리드 셀의 왼쪽 상단 좌표에 해당하는데 사용되며, p_w and p_h 는 앵커의 너비와 높이이다.

[0080] 높은 신뢰도 값을 처리하고 낮은 신뢰도 값을 버리는 임계값이 정의된다.

[0081] 비최대 억제를 사용하여 감지된 경계 상자에 대한 최종 위치 매개변수가 파생된다. 객체 탐지부(140)는 검출된 바운딩 박스에 대한 손실 함수가 계산된다.

[0082] 객체 탐지부(140)는 계산한 손실 함수의 총합이 최소가 되도록 영상 프레임의 특징을 학습한다.

[0083] 손실 함수는 분류 손실(수학식 4), 회귀 손실(수학식 5) 및 신뢰도 손실(수학식 6, 수학식 7)이라는 세 가지 함수의 합이다.

[0084] 객체 탐지부(140)는 손실 함수를 사용하여 각 참 양성에 대한 손실을 계산한다.

[0085] 객체 탐지부(140)는 계산한 손실 함수의 총합(분류 손실 + 회귀 손실 + 신뢰도 손실)이 최소가 되도록 영상 프레임의 특징을 학습한다.

[0086] 객체 탐지부(140)는 각 그리드 셀에서 객체가 감지되면 분류 손실을 조건부 클래스 확률의 제곱 오차로 계산하고 다음의 수학식 4와 같이 계산한다.

수학식 4

[0087]
$$\mathscr{L}_{cls} = \sum_{i=0}^{S^2} 1_{ij}^{obj} \sum_{c \varepsilon class} 1_i^{obj} (p_i(c) - p*_i(c))^2$$

[0088] 수학식 4에서, 그리드 셀 i에서 사람이 감지되면 $1_{ij}^{obj} = 1$ = 1이고, 사람이 감지되지 않으면 0이다. 그리드 셀 i에서 클래스 c에 대한 조건부 클래스 확률은 $p*_i(c)$로 표시된다. 회귀(Localization) 손실은 예측된 바운딩 박스 크기와 위치의 실패를 추정한다. 감지된 객체(즉, 사람)를 포함한 바운딩 박스가 추가된다.

[0089] 객체 탐지부(140)는 회귀 손실을 하기의 수학식 5에 의해 계산된다.

수학식 5

[0090]
$$\mathscr{L}_{loc} = \lambda_{coord} 1_{ij}^{obj} \sum_{i=0}^{S^2} \sum_{j=0}^{B} [(x_i - x_i^*)^2 + (y_i - y_i^*)^2 + (\sqrt{w_i} - \sqrt{w_i^*})^2 + (\sqrt{h_i} - \sqrt{h_i^*})^2]$$

[0091] 객체 탐지부(140)는 예측된 바운딩 박스 위치(x,y)와 관련된 손실을 계산한다. 객체 탐지부(140)는 각 그리드 셀(i=0,...,S^2)의 각 바운딩 박스 예측자(j=0,..,B)에 대한 합계를 계산한다. 수학식 5의 두 번째 부분은 예측된 바운딩 박스의 너비와 높이와 관련된 손실이다.

[0092] 위의 수학식 5에서 $1_{ij}^{obj} = 1$는 그리드 셀 i의 j번째 바운딩 박스가 객체 감지에 사용되는 경우 1이고, 그렇지 않으면 0이다.

[0093] 객체 탐지부(140)는 단순한 높이와 너비를 예측하는 대신에 바운딩 박스 너비와 높이의 제곱근을 예측한다.

[0094] 수학식 5는 스케일 매개변수 λ_{coord}는 바운딩 박스의 예측에 사용되며, 스케일 매개 변수를 5로 정의한다.

[0095] λ_{coord}는 바운딩 박스의 좌표 예측에 사용하는 스케일 매개변수이다($\lambda_{coord} = 5$). λ의 파라미터는 손실 함수에서 모델 안정성을 높이기 위해 필요하다.

[0096] 가장 높은 패널티는 좌표 예측($\lambda_{coord} = 5$)이고, 객체가 없을 때(λ_{noobj}=0.5), 신뢰 예측에 대한 가장 낮은 패널티이다.

[0097] x_i, y_i, h_i, w_i는 i번째 셀에서 감지된 바운딩 박스의 예측 위치이고, $x_i^*, y_i^*, h_i^*, w_i^*$는 i번째 셀에서 바운딩 박스의 실제 위치이다.

[0098] 객체 탐지부(140)는 수학식 5에서 좌표값 x,y를 갖는 예측된 바운딩 박스의 손실 함수를 계산한다.

[0099] 1_{ij}^{obj}는 j번째 바운딩 박스에서 감지된 사람의 가능성을 나타내기 위해 사용되고, λ는 일정한 상수이다.

[0100] 수학식 5는 각 그리드 셀(i = 0 to S^2)에 대한 예측자 (j = 0 to B)를 사용하여 각 바운딩 박스에 대한 합계를 계산한다.

[0101] 객체 탐지부(140)는 하기의 수학식 6을 이용하여 신뢰도 손실을 계산한다.

수학식 6

[0102]
$$\mathscr{L}_{conf} = \sum_{i=0}^{S^2}\sum_{j=0}^{B} 1_{ij}^{obj}(C_i - C_i^*)^2$$

[0103] 여기서, 신뢰도 점수는 그리드 셀 i의 j번째 바운딩 박스에 대해 C^*로 정의된다. 1_{ij}^{obj}는 셀 i에서 j번째 바운딩 박스가 객체를 감지하는 경우, 1과 같고, 감지되지 않는 경우, 0이 된다.

[0104] C는 신뢰도 점수이고, C^*는 예측 바운딩 박스와 그라운드 트루의 합집합에 대한 교차점이다.

[0105] 객체가 감지되지 않으면, 신뢰도 손실은 다음의 수학식 7과 같이 제공된다.

수학식 7

[0106]
$$\mathscr{L}_{conf} = \lambda_{noobj}\sum_{i=0}^{S^2}\sum_{j=0}^{B} 1_{ij}^{noobj}(C_i - C_i^*)^2$$

[0107] 수학식 7에서 1_{ij}^{noobj}는 1_{ij}^{obj}의 보수(Complement)로 정의된다.

[0108] 셀 i에서 바운딩 박스의 신뢰도 점수 C^*와 λ_{noobj}는 배경을 감지하는 동안 손실을 줄이는데 사용된다.

[0109] 감지된 대부분의 경우와 같이, 바운딩 박스에는 클래스 불균형 문제를 일으키는 객체가 포함되어 있지 않다.

[0110] 따라서, 이러한 모델은 객체를 감지하는 것보다 배경을 감지하도록 더 자주 훈련된다.

[0111] 이를 해결하기 위해서 손실은 인수 λ_{noobj}(디폴트: 0.5)에 의해 가중치를 줄인다.

[0112] 영상 프레임에서 사람을 감지한 후, 다음 단계에서 녹색 상자로 표시된 각 감지된 사람 바운딩 박스의 중심이 도 6의 (b)와 같이, 거리 계산에 사용된다.

[0113] 감지된 바운딩 박스 좌표는 바운딩 박스의 중심을 계산하는데 사용된다.

[0114] 도 6의 (c)는 바운딩 박스 좌표들을 받아들이고, 중심을 계산한 후, 감지된 각 바운딩 박스에 고유한 인덱스가 할당된다.

[0115] 다음 단계에서는 거리 계산부(170)는 유클리드 거리를 사용하여 감지된 각 중심 사이의 거리를 계산한다.

[0116] 거리 계산부(170)는 바운딩 박스를 감지하여 바운딩 박스 중심을 계산하고, 감지된 각 바운딩 박스 중심 쌍 사이의 거리를 픽셀의 개수로 계산한다.

[0117] 거리 계산부(170)는 영상 스트림의 모든 후속 프레임에 대해서 도 6의 (c)에 표시된 바운딩 박스 중심을 계산한다. 그런 다음, 거리 계산부(170)는 감지된 각 바운딩 박스 중심 쌍 사이의 거리(빨간색 선으로 강조 표시됨)를 계산한다(도 6의 (d))(S104).

[0118] 제어부(160)는 영상 프레임에 바운딩 박스(색깔 표시)와, 바운딩 박스 중심 간 유클리드 거리를 색깔을 다르게 하여 디스플레이부(180)에 표시한다.

[0119] 도 6의 (b)는 딥러닝 알고리즘을 사용하여 감지된 사람 객체의 바인딩 박스, 도 6의 (c)는 감지된 각 바운딩 박

스의 중심 계산을 나타내고, 도 6의 (d)는 바운딩 박스의 각 중심 쌍 사이의 거리를 빨간색 선으로 나타낸다.

[0120] 거리 계산부(170)는 각 중심의 정보는 목록 형태로 데이터베이스부(192)에 저장된다.

[0121] 제어부(160)는 계산한 각 바운딩 박스 중심 쌍 사이의 거리가 기설정된 임계값인 N개 픽셀의 개수보다 작은지 판단하고(S105), 계산한 거리가 N개 픽셀의 개수보다 작은 경우, 사회적 거리 위반으로 판단하여 위반 집합에 추가하여 데이터베이스부(192)에 저장한다(S106).

[0122] 제어부(160)는 계산한 거리가 N개 픽셀의 개수보다 작지 않은 경우, 사회적 거리 위반이 아니므로 비위반 집합에 추가하여 데이터베이스부(192)에 저장한다(S106).

[0123] 제어부(160)는 사회적 거리 위반의 개수를 카운트하여 누적하고, 누적된 사회적 거리 위반의 개수가 기설정된 기준 사회적 거리 위반수치와 비교하여 넘어서는 경우, 위험 신호를 생성한다.

[0124] 다른 실시예로서, 제어부(160)는 사회적 거리 위반의 개수를 카운트하여 누적하고, 누적된 사회적 거리 위반의 개수가 기설정된 기준 사회적 거리 위반수치와 비교하여 넘어서는 경우, 다른 지역으로 이동을 알리는 알람 신호를 생성하여 스피커(미도시)를 통해 출력한다. 예를 들어, 현재 지역에 사람들이 혼잡하므로 특정한 방향의 한적한 지역으로 이동할 것으로 알린다.

[0125] 제어부(160)는 감지한 바운딩 박스 중심을 유클리드 거리(Euclidean Distance)를 이용하여 계산하고, 바운딩 박스 중심이 계산되면, 해당 바운딩 박스에 인덱스를 각각 할당하고, 할당한 인덱스를 위반 집합에 추가하여 관리한다.

[0126] 제어부(180)는 계산한 거리가 기설정된 최소 사회적 거리를 위반하거나, 두 사람이 너무 가까우면 정보가 위반 집합에 추가되고, 바운딩 박스 색상은 녹색으로 초기화된다.

[0127] 도 7은 본 발명의 실시예에 따른 사회적 거리 모니터링 결과를 나타낸 도면이다.

[0128] 제어부(160)는 사람 객체를 포함한 바운딩 박스를 그린(Green) 색깔로 초기화하고(S108, S109), 인덱스가 위반 집합에 존재하는지 판단하고(S110), 인덱스가 위반 집합에 존재하는 경우, 바운딩 박스의 색깔을 레드(Red)로 업데이트하고(S111), 인덱스가 위반 집합에 존재하지 않는 경우, 바운딩 박스의 색깔을 그린으로 유지한다(S112).

[0129] 도 7에서 녹색 사각형은 사회적 거리를 유지하고 빨간색 사각형은 사회적 거리를 위반/위반하는 사람들이다.

[0130] 제어부(160)는 사회적 거리 위반으로 판단된 경우, 바운딩 박스의 색깔을 레드(Red)로 업데이트하고, 사회적 거리 위반으로 판단되지 않은 경우, 바운딩 박스의 색깔을 그린으로 유지한다.

[0131] 제어부(160)는 사회적 거리 위반의 개수를 카운트하여 누적하고, 누적된 사회적 거리 위반의 개수, 녹색 사각형, 빨간색 사각형을 디스플레이부(180)에 표시하며, 누적된 사회적 거리 위반의 개수가 기설정된 기준 사회적 거리 위반수치와 비교하여 넘어서는 경우, 위험 신호를 생성하여 외부의 중앙 서버(200)로 전송한다.

[0132] 제어부(160)는 위험 신호, 누적된 사회적 거리 위반의 개수를 포함한 사회적 거리 모니터링 정보를 생성하여 통신부(190)를 통해 중앙 서버(200)로 전송한다.

[0133] 객체 추적부(194)는 중심 추적 알고리즘을 이용하여 영상 시퀀스에서 감지된 사람을 추적하는데 사용된다. 중심 추적 알고리즘은 사회적 거리 임계값을 위반하는 사람들을 추적하는 데 도움이 된다. 출력에서 모델은 사회적 거리두기 위반의 총 개수에 대한 정보를 표시한다.

[0134] 객체 추적부(194)는 바운딩 박스 정보를 추적하기 위해 칼만 필터를 사용할 수도 있다.

[0135] 객체 추적부(194)는 사람 객체를 바운딩 박스로 표시한 영상 프레임을 수신하고, 바운딩 박스 정보를 추적하기 위해 칼만 필터를 사용하여 시간적으로 수신되는 복수의 영상 프레임에 포함된 바운딩 박스 정보를 추적한다.

[0136] 객체 추적부(194)는 컨볼루션 신경망(Convolutional Proposal Network, CNN)과 복수개의 필터들로 이루어져 바운딩 박스의 공간 정보 및 추적 정보를 수신하여 바운딩 박스의 모양(외형, Appearance) 정보를 포함한 특징 벡터를 추출한다.

[0137] CNN을 구성하는 컨볼루션 레이어는 커널을 이용하여 입력과 연관된 컨볼루션 연산을 수행할 수 있다.

[0138] 객체 추적부(194)는 CNN을 사용하여 바운딩 박스의 모양(외형, Appearance) 정보를 포함한 특징 벡터를 추출한다.

[0139] 객체 추적부(194)는 중심 추적 알고리즘을 이용하여 감지된 사람 객체를 추적하고(S113), 누적된 사회적 거리 위반의 개수를 디스플레이부(180)에 표시한다(S114).

[0140] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

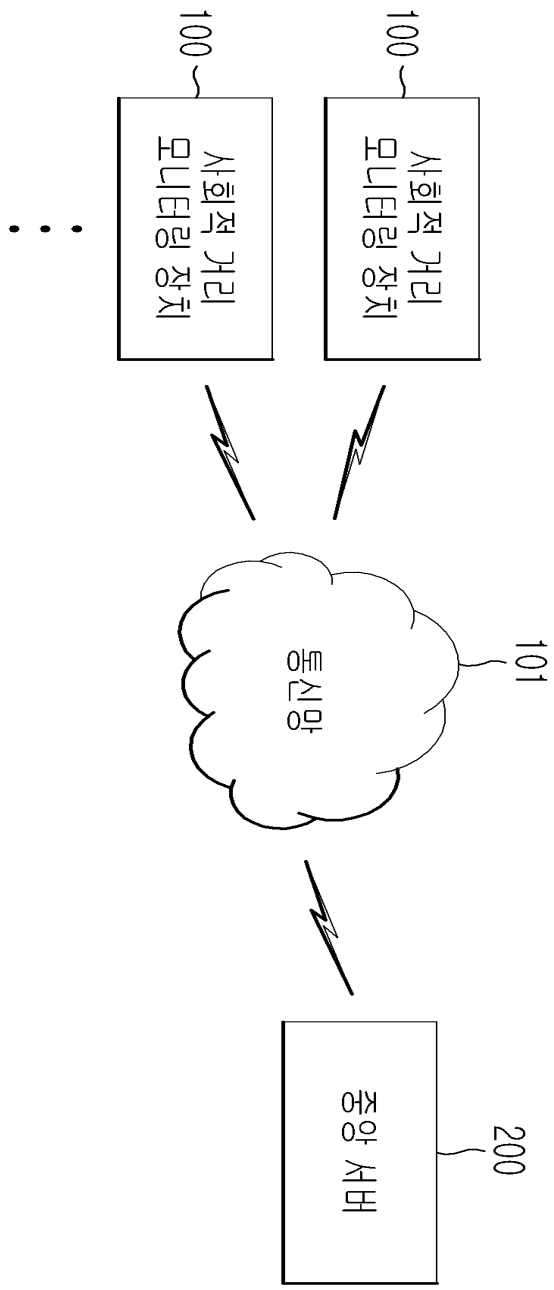
부호의 설명

[0141] 100: 사회적 거리 모니터링 장치

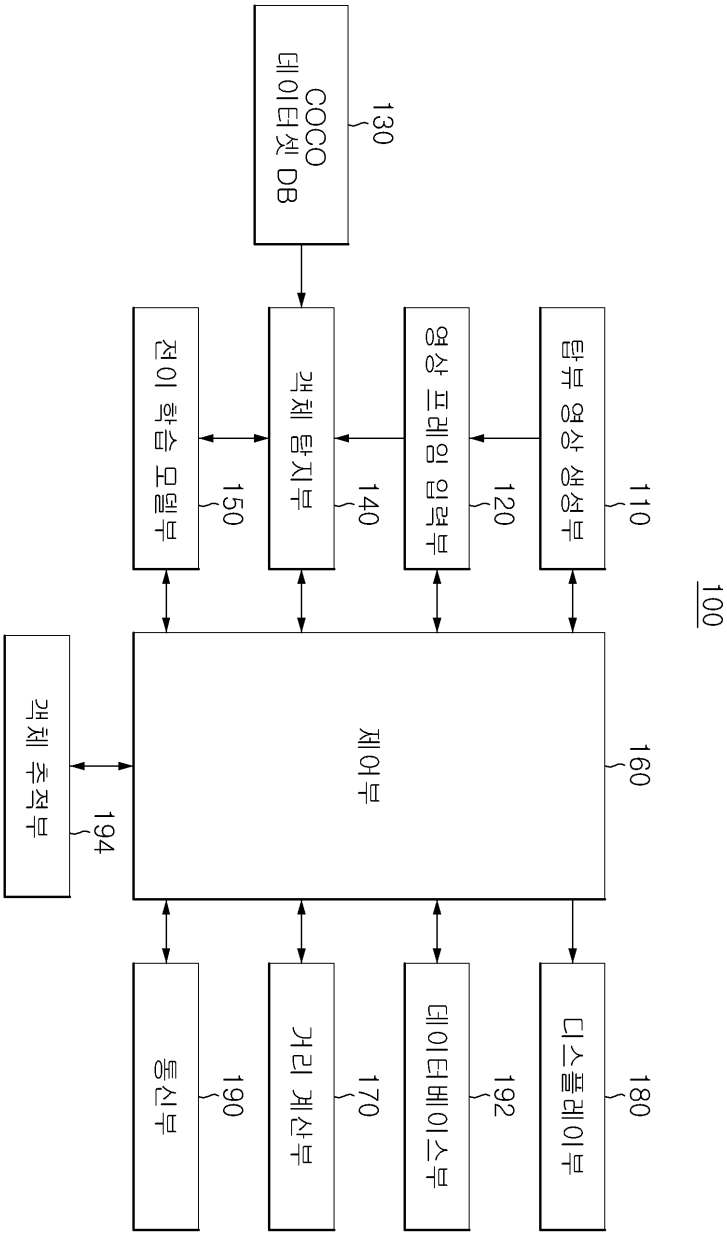
200: 중앙 서버

도면

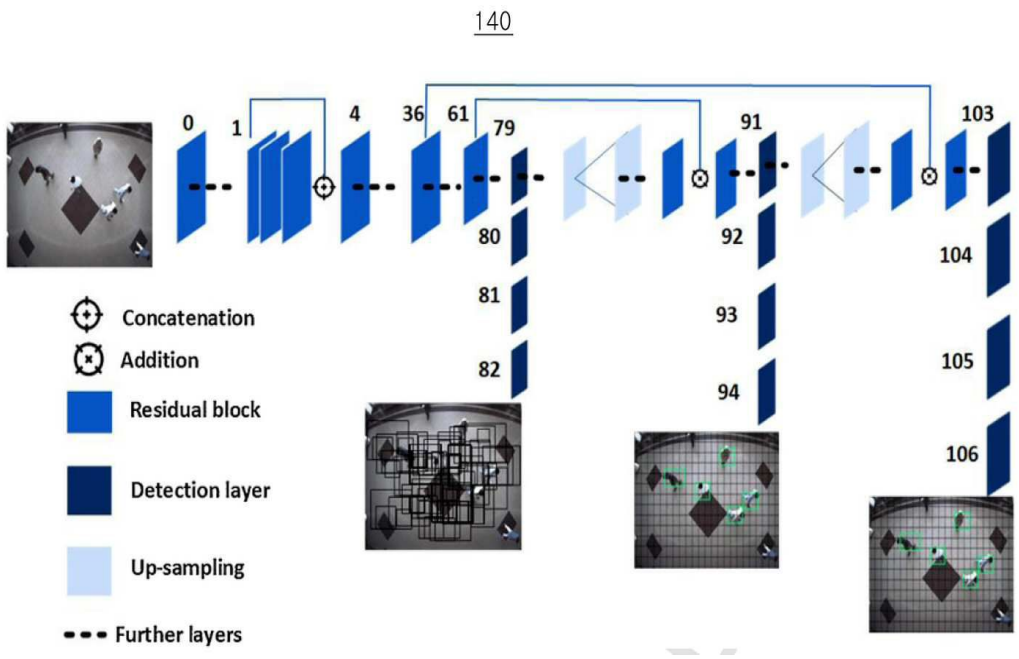
도면1



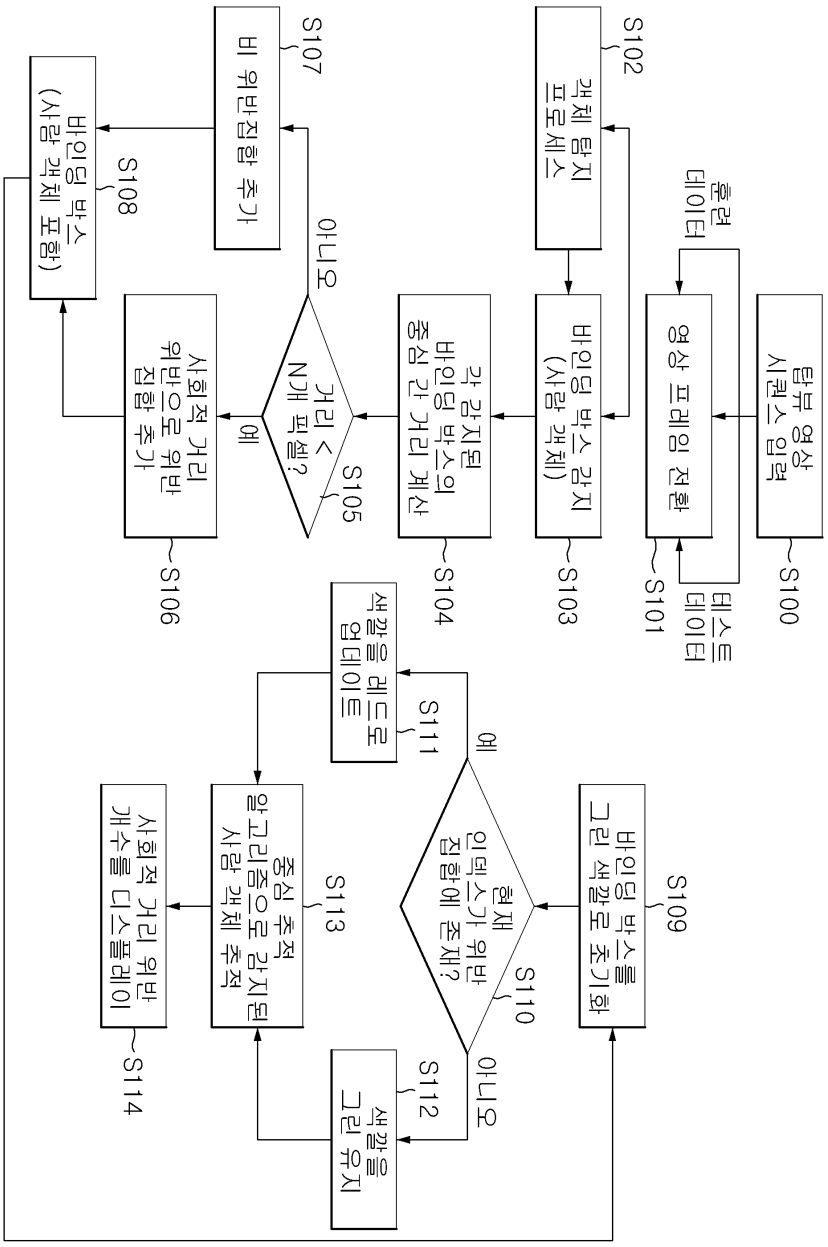
도면2


100
탐뷰 영상 생성부
110
영상 프레임 입력부
120
객체 탐지부
140
COCO 데이터셋 DB
130
전이 학습 모델부
150
제어부
160
디스플레이부
180
데이터베이스부
192
거리 계산부
170
통신부
190
객체 추적부
194

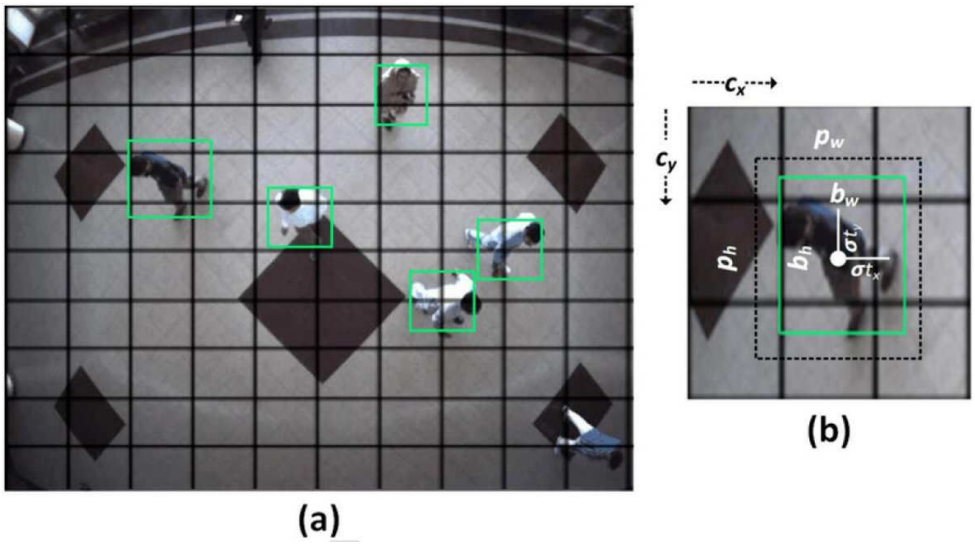
도면3


140
0
1
4
36
61
79
91
103
80
81
82
92
93
94
104
105
106
Concatenation
Addition
Residual block
Detection layer
Up-sampling
Further layers

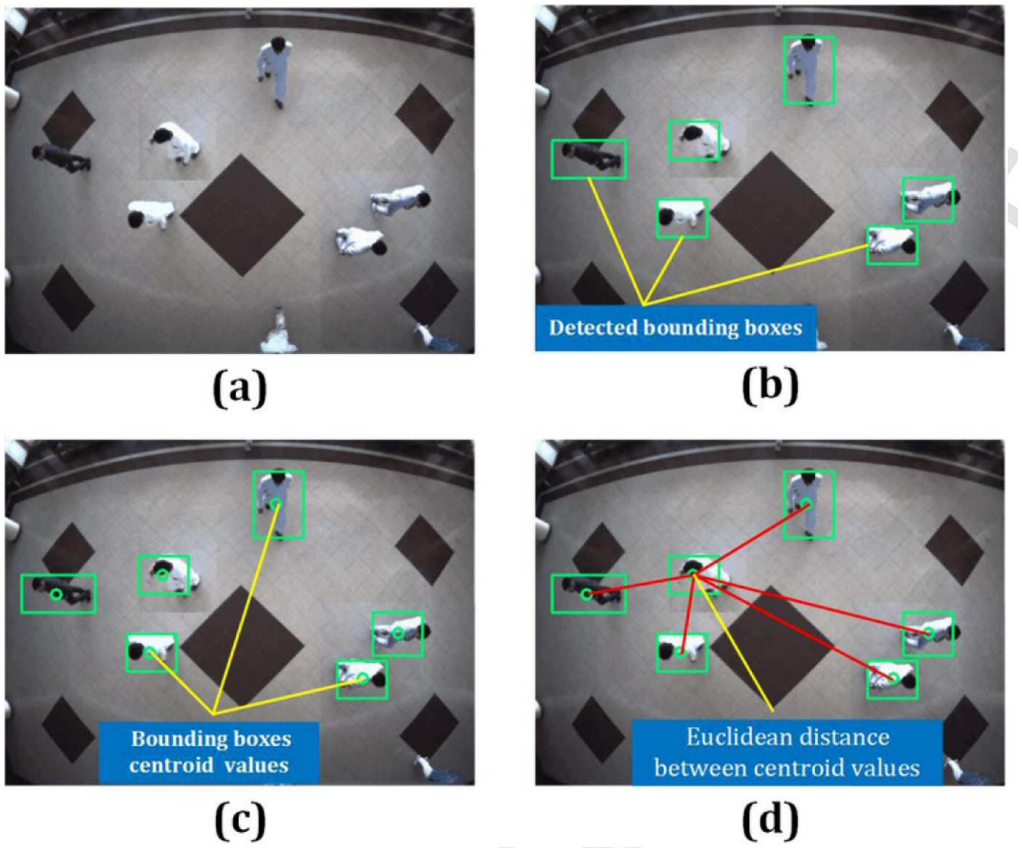
도면4


탑뷰 영상 시퀀스 입력
S100
훈련 데이터
영상 프레임 전환
S101
테스트 데이터
바인딩 박스 감지 (사람 객체)
S103
객체 탐지 프로세스
S102
각 감지된 바인딩 박스의 중심 간 거리 계산
S104
거리 < N개 픽셀?
S105
아니오
예
사회적 거리 위반으로 위반 집합 추가
S106
비 위반집합 추가
S107
바인딩 박스 (사람 객체 포함)
S108
바인딩 박스를 그린 색깔로 초기화
S109
현재 인덱스가 위반 집합에 존재?
S110
아니오
예
색깔을 그린 유지
S112
색깔을 레드로 업데이트
S111
중심 추적 알고리즘으로 감지된 사람 객체 추적
S113
사회적 거리 위반 개수를 디스플레이
S114

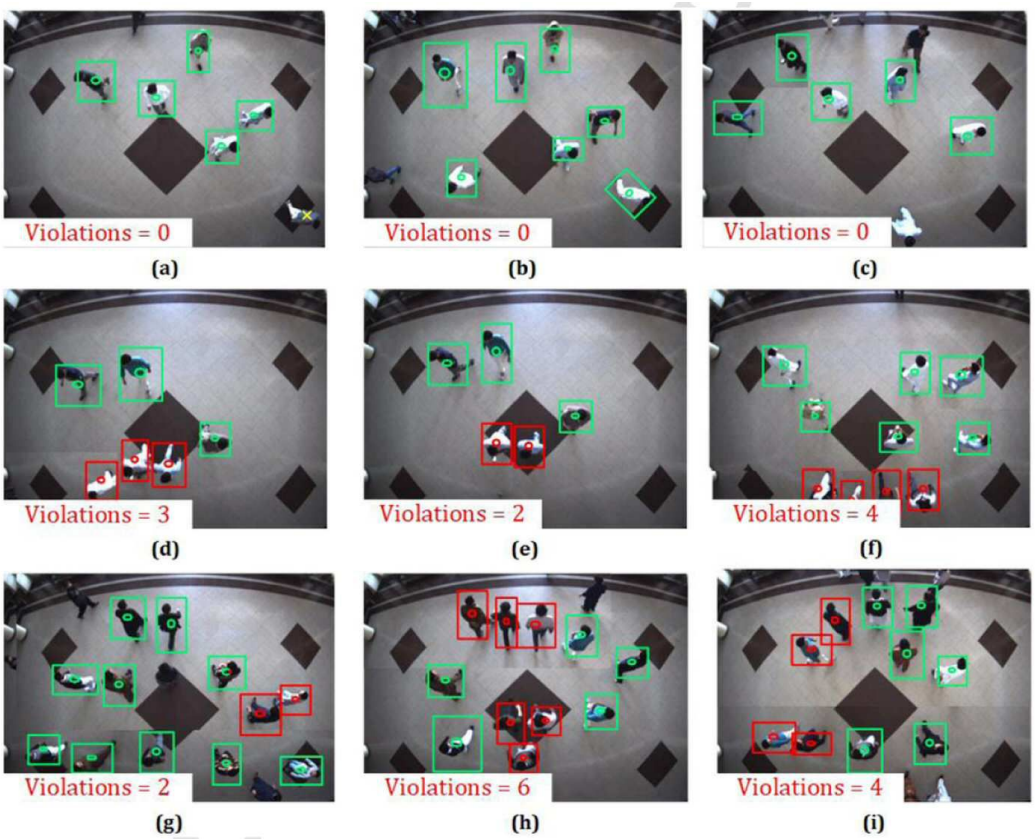
도면5


c_x
c_y
p_w
p_h
b_w
b_h
σt_y
σt_x
(a)
(b)

도면6


Detected bounding boxes
Bounding boxes centroid values
Euclidean distance between centroid values
(a)
(b)
(c)
(d)

도면7


Violations = 0
(a)
Violations = 0
(b)
Violations = 0
(c)
Violations = 3
(d)
Violations = 2
(e)
Violations = 4
(f)
Violations = 2
(g)
Violations = 6
(h)
Violations = 4
(i)