

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2024-0087443
(43) 공개일자 2024년06월19일

- (51) 국제특허분류(Int. Cl.)
G06V 10/762 (2022.01) *G06N 3/0464* (2023.01)
G06N 3/08 (2023.01) *G06V 10/40* (2022.01)
G06V 10/764 (2022.01) *G06V 10/774* (2022.01)
- (52) CPC특허분류
G06V 10/762 (2023.08)
G06N 3/0464 (2023.01)
- (21) 출원번호 10-2022-0173165
(22) 출원일자 2022년12월12일
심사청구일자 2022년12월12일
- (71) 출원인
울산과학기술원
울산광역시 울주군 언양읍 유니스트길 50
- (72) 발명자
심재영
울산광역시 울주군 언양읍 유니스트길 50
양재원
울산광역시 울주군 언양읍 유니스트길 50
김두현
울산광역시 울주군 언양읍 유니스트길 50
- (74) 대리인
특허법인 무한

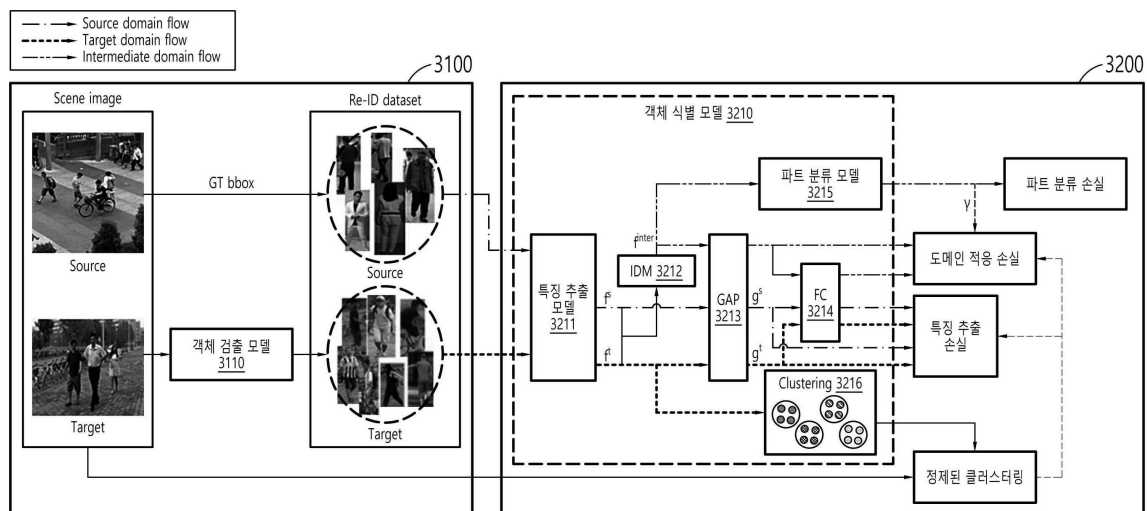
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 방법 및 장치

(57) 요약

일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 방법은 트레이닝이 완료된 객체 검출 모델에 기초하여 소스 도메인(source domain)에 속하는 소스 데이터의 소스 객체를 검출하는 단계, 상기 객체 검출 모델에 기초하여 타겟 도메인(target domain)에 속하는 타겟 데이터의 타겟 객체를 검출하는 단계, 특징 추출(뒷면에 계속)

대표도



모델에 기초하여 상기 검출된 소스 객체로부터 소스 특징 데이터, 상기 검출된 타겟 객체로부터 타겟 특징 데이터를 추출하는 단계, 상기 소스 특징 데이터 및 상기 타겟 특징 데이터로부터 생성된 중간 특징 데이터에 기초하여 상기 검출된 타겟 객체에 대한 신뢰성 스코어를 결정하는 단계, 상기 타겟 특징 데이터에 기초하여 상기 타겟 데이터로부터 검출된 상기 타겟 객체를 클러스터링(clustering)하는 단계, 유일성(uniqueness) 논리를 적용하여 상기 타겟 객체에 대한 클러스터를 정제(refine)하는 단계 및 상기 정제된 클러스터링 결과 및 상기 신뢰성 스코어에 기초한 목적 함수(objective function)를 이용하여 상기 특징 추출 모델을 포함하는 객체 식별 모델을 트레이닝시키는 단계를 포함할 수 있다.

(52) CPC특허분류

G06N 3/088 (2023.01)

G06V 10/40 (2023.08)

G06V 10/764 (2023.08)

G06V 10/774 (2023.08)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711152597
과제번호	2017-0-00667-006
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	디지털콘텐츠원천기술개발
연구과제명	이동 객체에 대한 증강현실 서비스 실현을 위한 정보-영상 정합기술 연구
기 여 율	1/3
과제수행기관명	울산과학기술원
연구기간	2022.01.01 ~ 2022.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711166373
과제번호	2020R1A2B5B01002725
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)
연구과제명	딥러닝 기반 유리 왜곡 모델링 및 영상 복원 기술 개발
기 여 율	1/3
과제수행기관명	울산과학기술원
연구기간	2022.03.01 ~ 2023.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1711152473
과제번호	2020-0-01336-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	인공지능대학원지원(울산과학기술원)
기 여 율	1/3
과제수행기관명	울산과학기술원
연구기간	2022.01.01 ~ 2022.12.31

명세서

청구범위

청구항 1

프로세서에 의해 수행되는 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 방법에 있어서,
 트레이닝이 완료된 객체 검출 모델에 기초하여 소스 도메인(source domain)에 속하는 소스 데이터의 소스 객체를 검출하는 단계;
 상기 객체 검출 모델에 기초하여 타겟 도메인(target domain)에 속하는 타겟 데이터의 타겟 객체를 검출하는 단계;
 특징 추출 모델에 기초하여 상기 검출된 소스 객체로부터 소스 특징 데이터, 상기 검출된 타겟 객체로부터 타겟 특징 데이터를 추출하는 단계;
 상기 소스 특징 데이터 및 상기 타겟 특징 데이터로부터 생성된 중간 특징 데이터에 기초하여 상기 검출된 타겟 객체에 대한 신뢰성 스코어를 결정하는 단계;
 상기 타겟 특징 데이터에 기초하여 상기 타겟 데이터로부터 검출된 상기 타겟 객체를 클러스터링(clustering)하는 단계;
 유일성(uniqueness) 논리를 적용하여 상기 타겟 객체에 대한 클러스터를 정제(refine)하는 단계; 및
 상기 정제된 클러스터링 결과 및 상기 신뢰성 스코어에 기초한 목적 함수(objective function)를 이용하여 상기 특징 추출 모델을 포함하는 객체 식별 모델을 트레이닝시키는 단계를 포함하는 객체 검색 모델의 트레이닝 방법.

청구항 2

제1항에 있어서,
 상기 소스 도메인의 트레이닝 데이터를 이용하여 상기 객체 검출 모델에 대한 지도 트레이닝을 수행하는 단계를 포함하는 객체 검색 모델의 트레이닝 방법.

청구항 3

제2항에 있어서,
 상기 객체 검출 모델의 트레이닝 후에, 상기 소스 특징 데이터 및 상기 타겟 특징 데이터를 이용하여 상기 객체 식별 모델의 비지도 트레이닝을 수행하는 단계를 더 포함하는 객체 검색 모델의 트레이닝 방법.

청구항 4

제1항에 있어서,
 상기 신뢰성 스코어를 결정하는 단계는,
 파트(part) 분류 모델에 기초하여 상기 중간 특징 데이터를 복수개의 파트(part)들에 대한 분류 결과를 생성하는 단계; 및
 상기 분류 결과에서 상기 복수개의 파트들에 대한 정답 예측 확률의 평균에 기초하여 상기 신뢰성 스코어를 산출하는 단계

를 포함하는 객체 검색 모델의 트레이닝 방법.

청구항 5

제4항에 있어서,

상기 파트(part)로 분류하는 단계는,

상기 복수개의 파트가 분류되는 상기 정답 예측 확률에 기초하여 산출되는 파트(part) 손실함수가 감소하는 방향으로 상기 파트(part) 분류 모델을 트레이닝시키는 단계

를 포함하는 객체 검색 모델의 트레이닝 방법.

청구항 6

제1항에 있어서,

상기 클러스터를 정제하는 단계는,

상기 클러스터링된 타겟 객체를 이미지 별로 그룹화하고 각 그룹에 대하여 유일성 논리를 적용함에 따라 상기 타겟 객체에 대한 클러스터를 정제하는 단계; 및

상기 정제된 클러스터를 기준으로 부정(negative) 객체 및 긍정(positive) 객체를 재설정하는 단계

를 포함하는 객체 검색 모델의 트레이닝 방법.

청구항 7

제6항에 있어서,

상기 객체 식별 모델을 트레이닝시키는 단계는,

상기 부정 객체와 기준 객체의 특징 벡터들 간의 거리가 멀어지고, 상기 긍정 객체와 상기 기준 객체의 특징 벡터들 간의 거리가 가까워지도록, 상기 특징 추출 모델의 파라미터를 업데이트하는 단계

를 더 포함하는 객체 검색 모델의 트레이닝 방법.

청구항 8

제1항에 있어서,

상기 클러스터를 정제하는 단계는,

상기 정제된 클러스터링 결과에 기초하여 산출된 손실함수가 감소하는 방향으로 상기 특징 추출 모델을 트레이닝시키는 단계

를 더 포함하는 객체 검색 모델의 트레이닝 방법.

청구항 9

제1항에 있어서,

상기 객체 식별 모델을 트레이닝시키는 단계는,

상기 신뢰성 스코어 및 상기 정제된 클러스터링 결과에 기초하여 산출되는 분류 도메인 적응 손실함수가 감소하는 방향으로 상기 객체 식별 모델을 트레이닝시키는 단계

를 더 포함하는 객체 검색 모델의 트레이닝 방법.

청구항 10

제1항에 있어서,

상기 객체 식별 모델을 트레이닝시키는 단계는,

상기 신뢰성 스코어 및 도메인 간의 거리 값에 기초하여 산출되는 특징 도메인 적응 손실함수가 감소하는 방향으로 상기 객체 식별 모델을 트레이닝시키는 단계

를 더 포함하는 객체 검색 모델의 트레이닝 방법.

청구항 11

제1항 내지 제10항 중 어느 한 항의 방법을 수행하기 위한 명령어를 포함하는 하나 이상의 컴퓨터 프로그램을 저장한 컴퓨터 판독 가능 기록 매체.

청구항 12

비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 장치에 있어서,

객체 검출 모델 및 객체 식별 모델을 트레이닝시키는 인스트럭션들(instruction)이 저장된 메모리; 및

상기 메모리에 저장된 인스트럭션들을 실행하는 프로세서

를 포함하고,

상기 프로세서는,

상기 메모리에 저장된 인스트럭션들을 실행하여 트레이닝된 상기 객체 검출 모델로부터 소스 도메인에 속하는 소스 데이터의 소스 객체 및 타겟 도메인에 속하는 타겟 데이터의 타겟 객체를 검출하고, 특징 추출 모델에 기초하여 상기 소스 객체로부터 소스 특징 데이터 및 상기 타겟 객체로부터 타겟 특징 데이터를 추출하고, 상기 소스 특징 데이터 및 상기 타겟 특징 데이터로부터 생성된 중간 데이터에 기초하여 상기 타겟 객체에 대한 신뢰성 스코어를 결정하고, 상기 타겟 특징 데이터에 기초하여 상기 타겟 객체를 클러스터링 하고, 유일성 논리를 적용하여 상기 타겟 객체에 대한 클러스터를 정제하며, 상기 정제된 클러스터링 결과 및 상기 신뢰성 스코어에 기초한 목적 함수를 이용하여 상기 특징 추출 모델을 포함하는 객체 식별 모델을 트레이닝시키는 객체 검색 모델의 트레이닝 장치.

청구항 13

제12항에 있어서,

상기 프로세서는,

상기 소스 도메인의 트레이닝 데이터를 이용하여 상기 객체 검출 모델에 대한 지도 트레이닝을 수행하고, 상기 소스 특징 데이터 및 상기 타겟 특징 데이터를 이용하여 상기 객체 식별 모델에 대한 비지도 트레이닝을 수행하는 객체 검색 모델의 트레이닝 장치.

청구항 14

제12항에 있어서,

상기 프로세서는,

파트 분류 모델에 기초하여 상기 중간 특징 데이터를 복수개의 파트(part)들에 대한 분류 결과를 생성하고, 상기 분류 결과에서 상기 복수개의 파트들에 대한 정답 예측 확률의 평균에 기초하여 상기 신뢰성 스코어를 산출하는 객체 검색 모델의 트레이닝 장치.

청구항 15

제14항에 있어서,

상기 프로세서는,

상기 복수개의 파트가 분류되는 상기 정답 예측 확률에 기초하여 산출되는 부분(part) 분류 손실 함수가 감소하는 방향으로 상기 파트(part) 분류 모델을 트레이닝시키는 객체 검색 모델의 트레이닝 장치.

청구항 16

제12항에 있어서,

상기 프로세서는,

상기 클러스터링된 타겟 객체를 이미지 별로 그룹화하고 각 그룹에 대하여 유일성 논리를 적용함에 따라 상기 타겟 객체에 대한 클러스터를 정제하며, 상기 정제된 클러스터를 기준으로 부정 객체 및 긍정 객체를 재설정하는 객체 검색 모델의 트레이닝 장치.

청구항 17

제16항에 있어서,

상기 프로세서는,

상기 부정 객체와 기준 객체의 특징 벡터들 간의 거리가 멀어지고, 상기 긍정 객체와 상기 기준 객체의 특징 벡터들 간의 거리가 가까워지도록 상기 특징 추출 모델의 파라미터를 업데이트하는 객체 검색 모델의 트레이닝 장치.

청구항 18

제12항에 있어서,

상기 프로세서는,

상기 정제된 클러스터링 결과에 기초하여 산출된 손실함수가 감소하는 방향으로 상기 특징 추출 모델을 트레이닝시키는 객체 검색 모델의 트레이닝 장치.

청구항 19

제12항에 있어서,

상기 프로세서는,

상기 신뢰성 스코어 및 상기 정제된 클러스터링 결과에 기초하여 산출되는 분류 도메인 적응 손실함수가 감소하는 방향으로 상기 객체 식별 모델을 트레이닝시키는 객체 검색 모델의 트레이닝 장치.

청구항 20

제12항에 있어서,

상기 프로세서는,

상기 신뢰성 스코어 및 도메인 간의 거리 값에 기초하여 산출되는 특징 도메인 적응 손실함수가 감소하는 방향으로 상기 객체 식별 모델을 트레이닝시키는 객체 검색 모델의 트레이닝 장치.

발명의 설명

기술 분야

[0001] 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 방법 및 장치가 제공된다.

배경 기술

[0003] 영상 데이터 기반 AI 서비스는 AI 기술들을 활용하여 영상과 이미지에 존재하는 객체의 종류와 특징들을 추출하고, 추출한 특징들로부터 유의미한 정보를 산출하여 사용자에게 제공하는 서비스를 의미할 수 있다. 이를 통해 AI 서비스 사용자는 높은 인식률과 정확도를 가진 의료, 보안, 불량검출, 범죄인지, 상황인지 등의 서비스를 제공할 수 있다. 또한, 사람이 등장하는 영상들로부터 특정 사람을 찾아내는 기술은 영상 데이터 기반 AI 서비스 중 AI 영상처리 기술의 일부분으로 시각 지능이 이미 인간 수준을 넘은 시각 인식률을 달성하는 것과 같이 현재 비약적인 발전을 이루고 있다.

[0004] 영상들로부터 특정 사람을 찾아내는 기술은 크게 사람 식별(person re-identification) 기술과 사람 검색(person search) 기술로 나눌 수 있다. 사람 식별 기술은 주어진 사람을 찾기 위하여 한 명의 사람만 촬영된 영상이나 전체 영상에서 미리 검출된 사람의 영상을 사용하는 등 제한된 촬영 환경과 조건을 필요로 하는 단점이 있다. 그러나 사람 검색 기술은 다수의 사람이 등장하는 CCTV 영상과 같은 영상으로부터 사람을 검출한 뒤 그 중에서 주어진 사람을 찾으므로 기술 적용에 별도의 제약이 없는 점에서 차이가 있다. 따라서 사람 검색 기술은 사람 식별 기술보다 실제 생활에 적용이 용이하고 적용되는 분야가 폭넓을 수 있다. 특히, 사람 검색 기술은 CCTV, 블랙박스 등으로부터 수집된 영상을 분석하여 사람 또는 발생 상황 등을 인식할 수 있는 광범위 보안 및 감시 시스템에 적용될 수 있다. 한편, 사람 검색 기술은 사람 검출(person detection) 기술과 사람 식별 기술이 함께 이용되는 기술로 전술한 사람 식별 기술보다 훨씬 복잡하고 다양한 도전적인 과제를 가지고 있다. 이에 사람 검색 기술은 도전적인 과제를 해결하기 위한 다양한 기술 개발을 필요로 한다.

[0005] 한편, 머신러닝(machine learning)은 AI의 한 분야로 소스 도메인(source domain)을 이용하여 학습하고, 타겟 도메인(target domain)을 이용하여 학습 결과를 확인할 수 있다. 또한, 머신러닝(machine learning)은 크게 지도 학습(supervised supervision) 방식 및 비지도 학습(unsupervised supervision) 방식으로 학습 방식을 분류할 수 있다. 지도 학습은 기계학습 모델에 데이터와 정답에 해당하는 라벨(label)을 함께 제공하여 학습하는 방식에 해당할 수 있다. 비지도 학습은 라벨없이 데이터를 분석함에 따라 특징을 구분하고, 구분한 특징을 이용하여 학습하는 방식에 해당할 수 있다. 사람 검색 기술은 지도 학습과 비지도 학습을 각각 응용하여 기계학습 모델에 적용하거나 상호 혼합하여 응용한 학습 방식을 기계학습 모델에 적용할 수 있다.

[0006] 위에서 설명한 배경기술은 발명자가 본원의 개시 내용을 도출하는 과정에서 보유하거나 습득한 것으로서, 반드시 본 출원 전에 일반 공중에 공개된 공지기술이라고 할 수는 없다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 대한민국 공개 특허공보 10- 2022-0068373(2022년 05월 26일 공고)에는 Cctv 환경에서의 보행자 추적 장치 및 방법이 제시된다.

(특허문헌 0002) 대한민국 등록 특허공보 10-2112859(2020년 05월 13일 공고)에는 레이블링 작업을 위해 딥러닝 모델을 트레이닝하는 방법 및 그를 이용한 장치가 제시된다.

발명의 내용

해결하려는 과제

- [0009] 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 방법 및 장치는 객체 식별 학습 시 유일성 정보를 적용하여 라벨링을 위한 군집화 과정이 개선될 수 있으며, 부분 분류화 기법으로 산출된 신뢰성이 적용된 목적함수를 사용하여 도메인 갭에 의한 성능 저하를 개선할 수 있다.
- [0010] 다만, 기술적 과제는 상술한 기술적 과제들로 한정되는 것은 아니며, 또 다른 기술적 과제들이 존재할 수 있다.

과제의 해결 수단

- [0012] 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 방법은 트레이닝이 완료된 객체 검출 모델에 기초하여 소스 도메인(source domain)에 속하는 소스 데이터의 소스 객체를 검출하는 단계, 상기 객체 검출 모델에 기초하여 타겟 도메인(target domain)에 속하는 타겟 데이터의 타겟 객체를 검출하는 단계, 특징 추출 모델에 기초하여 상기 검출된 소스 객체로부터 소스 특징 데이터, 상기 검출된 타겟 객체로부터 타겟 특징 데이터를 추출하는 단계, 상기 소스 특징 데이터 및 상기 타겟 특징 데이터로부터 생성된 중간 특징 데이터에 기초하여 상기 검출된 타겟 객체에 대한 신뢰성 스코어를 결정하는 단계, 상기 타겟 특징 데이터에 기초하여 상기 타겟 데이터로부터 검출된 상기 타겟 객체를 클러스터링(clustering)하는 단계, 유일성(uniqueness) 논리를 적용하여 상기 타겟 객체에 대한 클러스터를 정제(refine)하는 단계 및 상기 정제된 클러스터링 결과 및 상기 신뢰성 스코어에 기초한 목적 함수(objective function)를 이용하여 상기 특징 추출 모델을 포함하는 객체 식별 모델을 트레이닝시키는 단계를 포함할 수 있다.
- [0013] 상기 객체 검색 모델의 트레이닝 방법은 상기 소스 도메인의 트레이닝 데이터를 이용하여 상기 객체 검출 모델에 대한 지도 트레이닝을 수행하는 단계를 포함할 수 있다.
- [0014] 상기 객체 검색 모델의 트레이닝 방법은 상기 객체 검출 모델의 트레이닝 후에, 상기 소스 특징 데이터 및 상기 타겟 특징 데이터를 이용하여 상기 객체 식별 모델의 비지도 트레이닝을 수행하는 단계를 더 포함할 수 있다.
- [0015] 상기 신뢰성 스코어를 결정하는 단계는, 파트(part) 분류 모델에 기초하여 상기 중간 특징 데이터를 복수개의 파트(part)들에 대한 분류 결과를 생성하는 단계 및 상기 분류 결과에서 상기 복수개의 파트들에 대한 정답 예측 확률의 평균에 기초하여 상기 신뢰성 스코어를 산출하는 단계를 포함할 수 있다.
- [0016] 상기 파트(part)로 분류하는 단계는, 상기 복수개의 파트가 분류되는 상기 정답 예측 확률에 기초하여 산출되는 파트(part) 손실함수가 감소하는 방향으로 상기 파트(part) 분류 모델을 트레이닝시키는 단계를 포함할 수 있다.
- [0017] 상기 클러스터를 정제하는 단계는, 상기 클러스터링된 타겟 객체를 이미지 별로 그룹화하고 각 그룹에 대하여 유일성 논리를 적용함에 따라 상기 타겟 객체에 대한 클러스터를 정제하는 단계 및 상기 정제된 클러스터를 기준으로 부정(negative) 객체 및 긍정(positive) 객체를 재설정하는 단계를 포함할 수 있다.
- [0018] 상기 객체 식별 모델을 트레이닝시키는 단계는, 상기 부정 객체와 기준 객체의 특징 벡터들 간의 거리가 멀어지고, 상기 긍정 객체와 상기 기준 객체의 특징 벡터들 간의 거리가 가까워지도록, 상기 특징 추출 모델의 파라미터를 업데이트하는 단계를 더 포함할 수 있다.
- [0019] 상기 클러스터를 정제하는 단계는, 상기 정제된 클러스터링 결과에 기초하여 산출된 손실함수가 감소하는 방향으로 상기 특징 추출 모델을 트레이닝시키는 단계를 더 포함할 수 있다.
- [0020] 상기 객체 식별 모델을 트레이닝시키는 단계는, 상기 신뢰성 스코어 및 상기 정제된 클러스터링 결과에 기초하여 산출되는 분류 도메인 적응 손실함수가 감소하는 방향으로 상기 객체 식별 모델을 트레이닝시키는 단계를 더 포함할 수 있다.
- [0021] 상기 객체 식별 모델을 트레이닝시키는 단계는, 상기 신뢰성 스코어 및 도메인 간의 거리 값에 기초하여 산출되는 특징 도메인 적응 손실함수가 감소하는 방향으로 상기 객체 식별 모델을 트레이닝시키는 단계를 더 포함할 수 있다.

- [0022] 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 장치는 객체 검출 모델 및 객체 식별 모델을 트레이닝시키는 인스트럭션들(instruction)이 저장된 메모리 및 상기 메모리에 저장된 인스트럭션들을 실행하는 프로세서를 포함하고, 상기 프로세서는, 상기 메모리에 저장된 인스트럭션들을 실행하여 트레이닝된 상기 객체 검출 모델로부터 소스 도메인에 속하는 소스 데이터의 소스 객체 및 타겟 도메인에 속하는 타겟 데이터의 타겟 객체를 검출하고, 특징 추출 모델에 기초하여 상기 소스 객체로부터 소스 특징 데이터 및 상기 타겟 객체로부터 타겟 특징 데이터를 추출하고, 상기 소스 특징 데이터 및 상기 타겟 특징 데이터로부터 생성된 중간 데이터에 기초하여 상기 타겟 객체에 대한 신뢰성 스코어를 결정하고, 상기 타겟 특징 데이터에 기초하여 상기 타겟 객체를 클러스터링 하고, 유일성 논리를 적용하여 상기 타겟 객체에 대한 클러스터를 정제하며, 상기 정제된 클러스터링 결과 및 상기 신뢰성 스코어에 기초한 목적 함수를 이용하여 상기 특징 추출 모델을 포함하는 객체 식별 모델을 트레이닝시킬 수 있다.
- [0023] 상기 프로세서는 상기 소스 도메인의 트레이닝 데이터를 이용하여 상기 객체 검출 모델에 대한 지도 트레이닝을 수행하고, 상기 소스 특징 데이터 및 상기 타겟 특징 데이터를 이용하여 상기 객체 식별 모델에 대한 비지도 트레이닝을 수행할 수 있다.
- [0024] 상기 프로세서는 파트 분류 모델에 기초하여 상기 중간 특징 데이터를 복수개의 파트(part)들에 대한 분류 결과를 생성하고, 상기 분류 결과에서 상기 복수개의 파트들에 대한 정답 예측 확률의 평균에 기초하여 상기 신뢰성 스코어를 산출할 수 있다.
- [0025] 상기 프로세서는 상기 복수개의 파트가 분류되는 상기 정답 예측 확률에 기초하여 산출되는 부분(part) 분류 손실 함수가 감소하는 방향으로 상기 파트(part) 분류 모델을 트레이닝시킬 수 있다.
- [0026] 상기 프로세서는 상기 클러스터링된 타겟 객체를 이미지 별로 그룹화하고 각 그룹에 대하여 유일성 논리를 적용함에 따라 상기 타겟 객체에 대한 클러스터를 정제하며, 상기 정제된 클러스터를 기준으로 부정 객체 및 긍정 객체를 재설정할 수 있다.
- [0027] 상기 프로세서는 상기 부정 객체와 기준 객체의 특징 벡터들 간의 거리가 멀어지고, 상기 긍정 객체와 상기 기준 객체의 특징 벡터들 간의 거리가 가까워지도록 상기 특징 추출 모델의 파라미터를 업데이트할 수 있다.
- [0028] 상기 프로세서는 상기 정제된 클러스터링 결과에 기초하여 산출된 손실함수가 감소하는 방향으로 상기 특징 추출 모델을 트레이닝시킬 수 있다.
- [0029] 상기 프로세서는 상기 신뢰성 스코어 및 상기 정제된 클러스터링 결과에 기초하여 산출되는 분류 도메인 적응 손실함수가 감소하는 방향으로 상기 객체 식별 모델을 트레이닝시킬 수 있다.
- [0030] 상기 프로세서는 상기 신뢰성 스코어 및 도메인 간의 거리 값에 기초하여 산출되는 특징 도메인 적응 손실함수가 감소하는 방향으로 상기 객체 식별 모델을 트레이닝시킬 수 있다.
- [0031] 일 실시예에 따른 하나 이상의 컴퓨터 프로그램을 저장한 컴퓨터 판독 가능 기록 매체는, 방법을 수행하기 위한 명령어를 포함할 수 있다.

도면의 간단한 설명

- [0033] 도 1은 비지도 도메인 적응 기술에 대한 예시도를 도시한다.
- 도 2는 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 장치의 블록도를 도시한다.
- 도 3은 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 장치 내 프로세서에 의해 동작되는 도메인 기준에 따른 흐름도를 도시한다.
- 도 4는 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 방법의 흐름도를 도시한다.
- 도 5는 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 방법 중 클러스터링 방식 및 클러스터링된 클러스터를 정제하는 방식의 예시도를 도시한다.
- 도 6은 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 방법의 추론 단계의 흐름도를 도시한다.
- 도 7은 일 실시예에 따른 트레이닝 방법이 적용된 객체 검색 모델의 성능을 다른 객체 검색 모델과 비교한 결과

이다.

도 8은 일 실시예에 따른 트레이닝 방법을 순차적으로 적용한 객체 검색 모델의 성능을 비교한 결과이다.

발명을 실시하기 위한 구체적인 내용

- [0034] 실시예들에 대한 특정한 구조적 또는 기능적 설명들은 단지 예시를 위한 목적으로 개시된 것으로서, 다양한 형태로 변경되어 구현될 수 있다. 따라서, 실제 구현되는 형태는 개시된 특정 실시예로만 한정되는 것이 아니며, 본 명세서의 범위는 실시예들로 설명한 기술적 사상에 포함되는 변경, 균등물, 또는 대체물을 포함한다.
- [0035] 제1 또는 제2 등의 용어를 다양한 구성요소들을 설명하는데 사용될 수 있지만, 이런 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 해석되어야 한다. 예를 들어, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.
- [0036] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다.
- [0037] 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 설명된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함으로써 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0038] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 해당 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0039] 이하, 실시예들을 첨부된 도면들을 참조하여 상세하게 설명한다. 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일한 참조 부호를 부여하고, 이에 대한 중복되는 설명은 생략하기로 한다.
- [0040] 사람 검색 기술은 기계학습 모델의 학습 방식에 따라 크게 두가지로 분류될 수 있다. 사람 검색 기술은 예시적으로 지도 사람 검색 기술(Supervised Person Search) 및 약한 지도 사람 검색 기술(Weakly Supervised Person Search)로 분류할 수 있다.
- [0041] 지도 사람 검색 기술에서는 사람 검출의 기계학습 모델과 사람 식별 기계학습 모델 모두가 지도 학습을 이용하여 학습될 수 있다. 지도 사람 검색 기술은 소스 도메인의 확률 분포 타겟 도메인의 확률 분포가 일치 또는 유사한 경우에 높은 인식률을 나타낼 수 있으며, 사람의 위치 정보를 나타내는 바운딩 박스(bounding box) 라벨과 사람의 신원 정보를 나타내는 ID 라벨을 모두 기계학습 모델에 제공하는 특성을 가질 수 있다. 지도 사람 검색 기술은 사람 식별 기술과 다르게 대규모 데이터셋을 사용하며, 종단간 네트워크(end-to-end network)를 이용하여 사람 검출과 사람 식별에 대한 두가지 작업을 함께 수행하는 사람 검색 기술의 초창기 기술이다. 또한, 지도 사람 검색 기술은 응용기술로 어텐션(attention) 구조를 사용한 네트워크를 통해 잘못된 사람 검출 결과의 영향을 완화시키는 기술과 가중치 기법을 통해 검출된 결과의 품질을 사람 식별 단계에 반영하는 기술에 대한 개발이 이루어졌다. 또한, 지도 사람 검색 기술은 사람 검출 단계에서 앵커 프리(anchor-free)기반의 기계학습 모델을 제안하여 기존 Faster R-CNN(Region-Based Convolutional Neural Network)의 높은 계산 비용을 완화시키는 기술에 대한 개발 또한 이루어졌다.
- [0042] 한편, 소스 도메인과 타겟 도메인은 항상 같은 확률 분포를 가질 수 없고, 데이터 수가 충분하지 않을 경우에는 다른 도메인으로 학습해야 하는 경우가 있을 수 있다. 이에 소스 도메인과 타겟 도메인의 확률분포가 다른 것을 도메인 갭(domain gap)이라고 하며, 도메인 갭은 성능 저하의 원인이 될 수 있다. 그러나 지도 사람 검색 기술과 관련하여, 소스 도메인과 타겟 도메인의 확률 분포가 일치하지 않은 경우에 대한 기술 개발이 이루어지지 않고 있다. 또한, 동시에 지도 사람 검색 기술은 학습을 위해 대규모 데이터 셋(data set)의 라벨을 생성함에 따라 엄청난 경제적, 시간적 비용을 수반하는 문제가 있다.
- [0043] 이를 해결하고자 약한 지도 사람 검색 기술은 타겟 도메인의 ID 라벨없이 바운딩 박스 라벨만 제공하는 특성을 가진다. 이에 약한 지도 사람 검색 기술은 사람 검출 모델에는 지도 학습을 이용하고, 사람 식별 모델에는 비

지도 학습을 이용할 수 있다. 약한 지도 사람 검색 기술은 삼 네트워크(Siamese network)를 통해 군집화 단계에서 사람이 등장한 배경의 문맥을 개체마다 일관성을 가지도록 학습할 수 있는 기술 개발을 이루었으며, 하나의 영상에서 같이 등장한 사람은 다른 영상에서도 등장할 확률이 높다는 논리를 이용한 군집화 기법에 대한 기술 개발을 이루었다. 또한, 약한 지도 사람 검색 기술은 여러가지 문맥적 상황을 이용하여 배경 정보에서 벗어난 사람의 특징만을 학습하는 기술 개발 또한 이루었다. 그러나 약한 지도 사람 검색 기술은 학습을 위하여 여전히 타겟 도메인에 바운딩 박스 라벨을 생성해야 하는 문제가 존재하고 있으며, 소스 도메인과 타겟 도메인의 확률 분포 불일치에 대한 기술 개발 또한 여전히 이루어지지 않은 문제가 있다.

[0044] 소스 도메인과 타겟 도메인의 확률 분포 불일치 문제를 해결하고자 하는 목적은 타겟 도메인에서의 테스트 적용 시 좋은 결과를 도출하는 것일 수 있다. 이에 도메인 적응(domain adaptation) 시 소스 도메인을 통해 획득한 지식을 타겟 도메인으로 용이하게 전이(transfer)하는 것은 전술한 문제를 해결하는 방안이 될 수 있다. 이러한 도메인 적응 기법은 사람 검색 기술 분야가 아닌 사람 식별 기술 분야에 적용하는 연구가 활발히 이루어지고 있다.

[0045] 도 1은 비지도 도메인 적응 기술에 대한 예시도를 도시한다.

[0046] 일 실시예에 따르면, 비지도 도메인 적응 기술은 소스 도메인과 타겟 도메인이 다른 상황을 가정할 수 있다. 소스 도메인의 이미지(예: 소스 이미지)(110) 및 타겟 도메인의 이미지(예: 타겟 이미지)(130)의 둘 다에서 바운딩 박스 라벨이 검출될 수 있으나, 바운딩 박스 별 참값 라벨(ground truth label)은 소스 도메인(110)에만 있는 경우를 가정할 수 있다. 달리 말해, 소스 도메인에 속하는 트레이닝 데이터는 소스 이미지 및 소스 이미지에 대해 미리 주어진 참값 바운딩 박스라벨과 참값 바운딩 박스(ground truth bounding box) 별 참값 라벨(ground truth label)을 포함할 수 있다.

[0047] 예시적으로, 소스 도메인은 임의의 환경(예: 소스 환경)에서 획득된 데이터들(예: 소스 데이터, 소스 이미지)이 분포되는 도메인을 나타낼 수 있고, 타겟 도메인은 소스 환경과는 다른 환경(예: 타겟 환경)에서 획득된 데이터들(예: 타겟 데이터, 타겟 이미지)이 분포되는 도메인을 나타낼 수 있다. 수학적으로는 소스 도메인에 속하는 데이터들의 확률 분포와 타겟 도메인에 속하는 데이터들의 확률 분포가 서로 다른 것으로 해석될 수 있다. 예를 들어, 감시 카메라에 의해 획득된 이미지들과 스마트폰의 카메라에 의해 획득된 이미지들은 서로 다른 특성을 나타내므로, 서로 다른 도메인에 속할 수 있다.

[0048] 예를 들어, 기계 학습 모델(예: 하기 도 2의 객체 검출 모델)은 전술된 소스 도메인에 속하는 트레이닝 데이터를 이용하여 트레이닝될 수 있다. 예를 들어, 기계 학습 모델은 소스 도메인의 각 이미지(111, 112)로부터 바운딩 박스 라벨의 검출 결과 및/또는 검출된 바운딩 박스 별 ID 라벨(ID1, ID2, ID3, ID4)의 추출 결과를 생성할 수 있다. 기계 학습 모델은 바운딩 박스 라벨의 검출 결과 및/또는 검출된 바운딩 박스 별 ID 라벨(ID1, ID2, ID3, ID4)의 추출 결과에 기초하여 소스 객체의 특징을 추출할 수 있다. 소스 도메인 스페이스(120)에서는 특징 스페이스(feature space)에서 각 이미지(111, 112)로부터 추출된 소스 객체의 특징 벡터를 이용한 분포가 나타날 수 있다. 기계 학습 모델은 소스 객체의 특징 벡터를 이미지 별로 그룹(121, 122)으로 분류할 수 있다. 예를 들면, 이미지(111)는 그룹(121)으로 분류될 수 있고, 이미지(112)는 그룹(122)으로 분류될 수 있다. 전술된 바와 같이 트레이닝된 기계 학습 모델은 타겟 도메인에 속하는 이미지에 대해서도 사용될 수 있다. 소스 도메인(110)에 대해 트레이닝된 기계 학습 모델을 타겟 도메인(130)에 속하는 데이터(예: 이미지)(131, 132, 133)에 대해 사용하는 것은, 소스 도메인(110)을 이용하여 학습된 지식(knowledge)이 타겟 도메인(130)으로 전이(transfer)되는 것으로 해석될 수 있다. 따라서, 참값 라벨을 확보하기 어렵거나, 참값 라벨이 확보될 수 없는 타겟 도메인에 속하는 이미지(131, 132, 133)는 전술된 기계 학습 모델을 이용하여 타겟 도메인 스페이스(140)와 같이 그룹(141, 142, 143)으로 분류될 수 있다. 타겟 도메인 스페이스(140)에서는 특징 스페이스에서 각 이미지(131, 132, 133)로부터 추출된 타겟 객체의 특징 벡터를 이용한 분포가 나타날 수 있다. 타겟 도메인 스페이스(140)에서는 소스 도메인(110)으로부터 전이된 지식을 이용하여 타겟 객체의 특징 벡터가 이미지별 그룹(141, 142, 143)으로 분류될 수 있다. 예를 들면, 이미지(131)는 그룹(141)으로 분류될 수 있고, 이미지(132)는 그룹(142)으로 분류될 수 있으며, 이미지(133)는 그룹(143)으로 분류될 수 있다. 이에 따라 기계 학습 모델의 트레이닝 과정에서 타겟 도메인은 라벨이 필요하지 않을 수 있다.

[0049] 비지도 도메인 적응형 사람 식별 기술(Unsupervised Domain Adaptive Person Re-Identification)은 타겟 도메인의 라벨이 없는 상태에서 소스 도메인과 타겟 도메인 간의 거리(예: 차이)를 줄이는 방식을 포함할 수 있다. 소스 도메인과 타겟 도메인 간의 거리(예: 차이)를 줄이는 방식 중 하나는 생성 모델을 사용하여 소스 도메인의 스타일을 타겟 도메인처럼 변형하거나 소스 도메인과 타겟 도메인을 함께 학습에 이용하는 방식이 있

다. 그러나 비지도 도메인 적응형 사람 식별 기술은 사람 검색 기술이 아닌 사람 식별 기술로 사람 검출 단계가 없으므로 영상에 한 사람만 존재하는 제한적인 상황에서만 적용가능한 문제가 있다. 이에 본 발명은 비지도 도메인 적응형 사람 식별 기술을 응용하여 전술한 문제들에 대한 해결 방안을 도출하였다.

[0050] 도 2는 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 장치의 블록도를 도시한다.

[0051] 일 실시예에 따른 전자 장치(예: 객체 검색 모델의 트레이닝 장치(200))는 통신부(210), 메모리(220) 및 프로세서(230)를 포함할 수 있다. 객체 검색 모델의 트레이닝 장치(200)는 객체(예: 사람) 검색 전 객체 검색 모델을 트레이닝시킬 수 있다. 객체 검색 모델은 객체 검출 모델 및 객체 식별 모델을 포함할 수 있다. 예를 들어, 전자 장치는 트레이닝된 객체 검출 모델을 이용하여 이미지로부터 객체를 검출하고, 트레이닝된 객체 식별 모델을 이용하여 객체 검출 모델에 의해 검출된 객체를 식별할 수 있다.

[0052] 통신부(210)는 외부의 장치와 연결되어 외부 장치로부터 객체 검색 모델의 트레이닝 및 추론에 필요한 이미지들을 수신할 수 있다. 또한, 객체 검색 모델의 트레이닝 장치(200)는 하기 도 6에서 설명하는 바와 같이 객체 검색 장치로도 이용될 수 있으므로, 통신부(210)를 통해 객체 검색이 필요한 이미지들을 수신할 수 있다.

[0053] 메모리(220)는 객체 검출 모델 및 객체 식별 모델을 트레이닝시키는 인스트럭션들(instruction) 및 트레이닝에 필요한 하이퍼 파라미터 값을 저장할 수 있다. 하이퍼파라미터는 모델링할 때 사용자가 직접 세팅해주는 값으로 일 실시예에 따르면 η, τ, β 및 λ 가 포함되나 이에 한정하지 않는다. 추가로, 메모리(220)는 통신부(210)를 통해 외부 장치로부터 수신한 데이터를 임시 또는 장기로 저장할 수 있다. 메모리(220)는 전원이 공급되지 않아도 저장된 정보를 계속 유지하는 비휘발성 저장장치 및 휘발성 저장장치일 수 있다. 예를 들면, 메모리(220)는 콤팩트 플래시(compact flash; CF) 카드, SD(secure digital) 카드, 메모리 스틱(memorystick), 솔리드 스테이트 드라이브(solid-state drive; SSD) 및 마이크로(micro) SD 카드 등과 같은 낸드 플래시 메모리(NAND flash memory), 하드 디스크 드라이브(hard disk drive; HDD) 등과 같은 마그네틱 컴퓨터 기억장치 및 CD-ROM, DVD-ROM 등과 같은 광학 디스크 드라이브(optical disc drive) 등을 포함할 수 있다.

[0054] 인스트럭션들은 프로세서(230)에서 수행하는 모든 동작들을 포함할 수 있다. 프로세서(230)는 메모리(220)에 저장된 인스트럭션들을 실행하여 객체 검출 모델 및 객체 식별 모델을 트레이닝시킬 수 있다.

[0055] 도 3은 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 장치 내 프로세서에 의해 동작되는 도메인 기준에 따른 흐름도를 도시한다.

[0056] 프로세서(230)는 객체 검출 모델(3110)을 이용하여 동작(3100)을 수행할 수 있고, 객체 식별 모델(3210)을 이용하여 동작(3200)을 수행할 수 있다.

[0057] 동작(3100)에서, 트레이닝 장치는 소스 도메인(source domain)에 속하는 트레이닝 데이터를 이용하여 객체 검출 모델(3110)의 지도 트레이닝을 수행할 수 있다. 지도 트레이닝 과정은 정답이 있는 소스 도메인을 이용하여 수행될 수 있다. 지도 트레이닝이 완료된 객체 검출 모델(3110)은 후술하는 객체 식별 모델(3210)의 트레이닝시 정답이 없는 타겟 도메인의 데이터에 대해서도 이용될 수 있다.

[0058] 객체 검출 모델(3110)은 입력된 이미지에 대해 바운딩 박스와 객체 정보(예: 객체의 종류를 지시하는 ID 라벨)를 출력하도록 설계 및 트레이닝될 수 있다. 예를 들어, 입력이 소스 도메인(예: 전체 영상(scene image))인 경우, 객체 검출 모델(3110)의 출력은 객체가 있을 것으로 예측된 소스 바운딩 박스(bounding box) 라벨과 소스 객체 검출 ID 라벨(예: 검출된 객체의 종류)을 포함할 수 있다. 예를 들어, 프로세서(230)는 소스 도메인의 참값(ground truth, GT) 라벨(예: 바운딩 박스 라벨 및 객체 검출 ID 라벨)을 이용하여 객체 검출 모델(3110)을 트레이닝시킬 수 있다. 참고로, 사람 검색에서는 검출된 객체가 사람이어야 하고, 사람이 아니면 배경으로 취급될 수 있다. 트레이닝 장치는 객체 검출 ID 라벨이 사람을 지시하는 것으로 검출된 바운딩 박스 외의 나머지 객체 검출 ID 라벨에 대응하는 바운딩 박스에 대해서는 배경으로 처리하여 객체 검출 모델(3110)을 트레이닝시킬 수 있다.

[0059] 한편, 트레이닝이 완료된 객체 검출 모델(3110)은 후술하는 객체 식별 모델(3210)의 트레이닝 동작 및/또는 객체 검색 모델을 이용한 추론 동작에 사용될 수 있다. 이 때, 객체 검출 모델(3110)에는 소스 도메인이 아닌 타겟 도메인에 속하는 데이터가 입력될 수 있다. 입력이 타겟 도메인(target domain)인 경우, 객체 검출 모델(3110)의 출력은 객체의 검출 결과(예: 타겟 바운딩 박스 라벨 및 타겟 객체 검출 ID 라벨)일 수 있다. 프로세서(230)는 타겟 도메인에서 검출된 객체를 이용하여 동작(3200)을 수행할 수 있다.

[0060] 또한, 객체 검출 모델은 Faster R-CNN 모델을 포함할 수 있다. Faster R-CNN은 딥러닝 기반의 영역 제안 네트

워크(Region Proposal Network, RPN)와 Fast R-CNN 모델로 구성될 수 있다. 딥러닝 기반의 영역 제안 네트워크(RPN)는 객체를 포함할 가능성이 높은 영역을 선택적으로 탐색하여 객체를 탐지할 수 있다. 그런 후, Fast R-CNN 모델은 탐지된 객체를 전달받아 객체의 클래스(class)와 위치를 예측할 수 있다. Faster R-CNN은 매우 빠른 단일 통합 네트워크이므로 하나의 네트워크만 트레이닝하고 실행하면 된다.

[0061] 동작(3200)에서, 트레이닝 장치는 소스 도메인, 타겟 도메인 및 중간 도메인을 이용하여 객체 식별 모델(3210)의 비지도 트레이닝을 수행할 수 있다. 비지도 트레이닝 과정은 객체 식별 모델(3210)을 이용하여 수행할 수 있으며, 추론 단계는 도6에서 후술한다.

[0062] 프로세서(230)는, 객체 검출 모델(3110)에 의해 검출된 객체를 식별하기 위하여, 객체 식별 모델(3210)을 트레이닝시킬 수 있다. 객체 식별 모델(3210)은 객체 검출 결과(예: 사람으로 분류된 바운딩 박스)에 대응하는 이미지(예: 바운딩 박스에 대응하는 패치 이미지(patch image))를 입력받아 객체 식별 ID라벨을 출력하도록 설계 및 트레이닝될 수 있다. 후술하겠으나, 객체 식별 ID라벨은, 객체의 실제 신원을 지시하는 정보라기 보다, 여러 이미지들에서 검출된 바운딩 박스들의 객체들을 구분하기 위한 일종의 임시 식별자일 수 있다. 예를 들어, 복수의 이미지들에서 검출된 바운딩 박스들 중 같은 객체를 포함하는 것으로 판단된 바운딩 박스에 대해서는 같은 객체 식별 ID라벨이 부여될 수 있다. 다른 예를 들어, 복수의 이미지들에서 검출된 바운딩 박스들 중 서로 다른 객체를 포함하는 것으로 판단된 바운딩 박스들에 대해서는 서로 다른 객체 식별 ID라벨이 부여될 수 있다.

[0063] 일 실시예에 따른 객체 식별 모델(3210)은 특징 추출 모델(3211), IDM(Intermediate Domain Module)(3212), GAP layer(3213), FC layer(3214), 파트 분류 모델(3215) 및 클러스터링 알고리즘(3216)을 포함할 수 있다. 프로세서(230)는 객체 검출 모델(3110)에 의해 검출된 소스 객체 및 타겟 객체의 특징을 특징 추출 모델(3211)에 의해 추출할 수 있다. 소스 객체는 소스 이미지로부터 객체 검출 모델(3110)에 기초하여 검출된 객체를 나타내고 타겟 객체는 타겟 이미지로부터 객체 검출 모델(3110)에 기초하여 검출된 객체를 나타낼 수 있다. 프로세서(230)는 소스 이미지로부터 소스 객체에 대응하는 부분을 크롭(crop)함으로써 생성된 소스 패치 이미지를 특징 추출 모델(3211)에 입력함으로써, 소스 객체에 대한 특징 데이터(예: 특징 벡터)를 추출할 수 있다. 유사하게, 프로세서(230)는 타겟 이미지로부터 타겟 객체에 대응하는 부분을 크롭함으로써 생성된 타겟 패치 이미지를 특징 추출 모델(3211)에 입력함으로써, 타겟 객체에 대한 특징 데이터를 추출할 수 있다.

[0064] 프로세서(230)는 특징 추출 모델(3211)에 의해 추출된 소스 특징 데이터 및 타겟 특징 데이터를 혼합함으로써 중간 특징 데이터를 생성할 수 있다. 중간 특징 데이터는 중간 도메인에 속하는 특징 데이터를 나타낼 수 있다. 예시적으로 프로세서(230)는 IDM(3212)에 기초하여 소스 특징 데이터 및 타겟 특징 데이터를 혼합함으로써 중간 특징 데이터를 생성할 수 있다. IDM(3212)에 기초한 중간 특징 데이터의 생성은, 'Dai, Y.; Liu, J.; Sun, Y.; Tong, Z.; Zhang, C.; and Duan, L.-Y. 2021. Idm: An intermediate domain module for domain adaptive person re-id. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 11864- 11874'에서 참조될 수 있다. 중간 도메인은 소스 도메인과 타겟 도메인을 매개하는 도메인을 나타낼 수 있다. 소스 도메인을 이용하여 트레이닝된 지식은 중간 도메인을 경유하여 도메인 적응(Domain Adaptation) 시 타겟 도메인으로 용이하게 전이(transfer)될 수 있다.

[0065] 프로세서(230)는 소스 특징 데이터, 타겟 특징 데이터 및 중간 특징 데이터를 개별적으로 전역 평균 풀링(Global average pooling, GAP) 레이어(3213)에 통과시켜 특징의 수를 감소시킬 수 있다. 프로세서(230)는 특징의 수가 감소된 중간 특징 데이터를 바로 도메인 적응 손실 함수에 적용할 수 있다. 또한, 프로세서(230)는 특징의 수가 감소된 소스 특징 데이터 및 타겟 특징 데이터를 바로 특징 추출 손실 함수에 적용할 수 있다. FC(fully-connected) 레이어(3214)는 마지막 특징과 매트릭스(matrix) 곱을 한 결과를 출력할 수 있다. 이에 프로세서(230)는 GAP 레이어(3213)를 통과하여 특징의 수가 감소한 소스 특징 데이터, 타겟 특징 데이터 및 중간 특징 데이터를 FC 레이어(3214)에 각각 개별적으로 통과시켜 매트릭스 형태의 결과 값을 출력할 수 있다. 이렇게 출력된 매트릭스 형태의 결과는 도메인 적응 손실 함수 및 특징 추출 손실 함수의 값을 산출하는데 각각 사용될 수 있다.

[0066] 프로세서(230)는 전술한 중간 특징 데이터를 파트 분류 모델(3215)에 적용하여 복수개의 파트(part)들로 분류할 수 있다. 이러한 분류 결과에 따라 프로세서(230)는 복수개의 파트들에 대한 정답 예측 확률을 파트(part) 분류 손실 함수에 적용할 수 있다. 프로세서(230)는 파트(part) 분류 손실 함수를 이용하여 파트 분류 모델(3215)을 트레이닝시킬 수 있다. 추가로 프로세서(230)는 복수의 파트들에 대한 정답 예측 확률의 평균에 기초하여 신뢰성 스코어를 산출할 수 있다. 또한, 프로세서(230)는 타겟 특징 데이터를 클러스터링 알고리즘(3216)에 통과시켜 타겟 특징 데이터에 기초하여 타겟 객체에 대한 클러스터링을 수행할 수 있다. 프로세서(230)는

클러스터링된 클러스터를 정제하여 정제된 클러스터를 형성할 수 있다. 프로세서(230)는 정제된 클러스터링 결과를 특징 추출 손실 함수에 적용하여 특징 추출 모델(3211)을 트레이닝시킬 수 있다. 또한, 프로세서(230)는 정제된 클러스터링 결과와 전술한 신뢰성 스코어를 도메인 적응 손실 함수에 적용하여 객체 식별 모델(3210)을 트레이닝시킬 수 있다. 결과적으로 프로세서(230)는 객체 식별 모델(3210)을 트레이닝시키면서 객체 식별 모델(3210) 내에 있는 모델 또한 트레이닝시킬 수 있다. 또한, 전술한 특징 추출 모델(3211), IDM(3212), GAP layer(3213), FC layer(3214), 파트 분류 모델(3215) 및 클러스터링 알고리즘(3216)에 의한 동작들은 동시에 이루어질 수 있으나, 이에 한정하지 않고, 순차적으로도 이루어질 수 있다.

[0067] 도 4는 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 방법의 흐름도를 도시한다.

[0068] 단계(410)에서, 프로세서는 트레이닝이 완료된 객체 검출 모델에 기초하여 소스 도메인에 속하는 소스 데이터의 소스 객체를 검출할 수 있다. 프로세서는 소스 도메인의 GT 라벨(예: 바운딩 박스 라벨, ID 라벨)을 이용하여 객체 검출 모델에 지도 트레이닝을 수행할 수 있다. 소스 객체는 객체 검출 모델에 의한 검출없이 트레이닝을 위한 GT 라벨을 다음단계에 사용할 수 있다.

[0069] 단계(420)에서, 프로세서는 객체 검출 모델에 기초하여 타겟 도메인에 속하는 타겟 데이터의 타겟 객체를 검출할 수 있다. 도메인에는 복수의 이미지들이 있고 객체는, 예를 들면, 한 명 이상의 사람일 수 있다.

[0070] 단계(430)에서, 프로세서는 특징 추출 모델에 기초하여 검출된 소스 객체로부터 소스 특징 데이터, 검출된 타겟 객체로부터 타겟 특징 데이터를 추출할 수 있다. 예를 들면, 프로세서는 소스 특징 데이터 및 타겟 특징 데이터를 Re-ID 모델을 이용하여 추출할 수 있다.

[0071] 단계(440)에서, 프로세서는 소스 특징 데이터 및 타겟 특징 데이터로부터 생성된 중간 특징 데이터에 기초하여 타겟 객체에 대한 신뢰성 스코어를 결정할 수 있다. 프로세서는 IDM에 소스 특징 데이터 및 타겟 특징 데이터를 입력하여 중간 특징 데이터를 생성할 수 있다. 신뢰성 스코어는 수학적식 1 및 2에서 후술한다.

[0072] 단계(450)에서, 프로세서는 타겟 특징 데이터에 기초하여 타겟 데이터로부터 검출된 타겟 객체를 클러스터링할 수 있다. 프로세서는 클러스터링을 통해 타겟 객체에 수도 라벨(pseudo label)을 부여할 수 있다.

[0073] 단계(460)에서, 프로세서는 유일성(uniqueness) 논리를 적용하여 타겟 객체에 대한 클러스터를 정제(refine)할 수 있다. 유일성 논리는 “한 영상에서 같은 사람은 등장할 수 없다” 는 의미를 내포하고, 유일성 논리를 적용하여 클러스터를 정제하는 방식은 도5에서 후술한다.

[0074] 단계(470)에서, 프로세서는 정제된 클러스터링 결과 및 신뢰성 스코어에 기초한 목적 함수(objective function)를 이용하여 특징 추출 모델을 포함하는 객체 식별 모델을 트레이닝시킬 수 있다. 목적함수는 트레이닝한 모델의 최적화 알고리즘에서 최대값 또는 최소값을 찾아야하는 함수로, 본 발명에서는 손실 함수(loss function)를 사용할 수 있다. 손실(또는 에러(Error))은 0이 완벽한 모델로 0과 1 사이의 숫자로 측정될 수 있다. 이에 모델은 손실이 가능한 0에 가깝게 되도록 최적화할 수 있다. 신뢰성 스코어에 기초한 목적 함수는 수학적식 4에서 후술한다.

[0075] 객체 검출 모델을 통한 타겟 도메인의 객체 검출은 타겟 도메인에 바운딩 박스 라벨 정보가 없기 때문에 객체 검출 결과가 부정확할 확률이 높다. 또한, 객체 식별 모델을 이용한 객체 식별은 객체 검출 모델을 이용하여 검출된 객체에 기초하여 수행되므로 연쇄적으로 낮은 식별 결과를 갖게 될 수 있다. 이를 해결하기 위하여 프로세서는 신뢰성 스코어 값을 산출하여 객체 검출 결과를 확인할 수 있다. 프로세서는 파트(part) 분류 모델에 의해 분류된 파트(part)의 행렬 값에 기초하여 신뢰성 스코어를 결정할 수 있다.

[0076] 파트(part) 분류 모델은 중간 특징 데이터에 파셜 최대 풀링(partial max pooling)을 적용하여 여러 개의 나누어진 부분을 하나의 파트(part)으로 지정할 수 있다. 예를 들면, 파트(part) 분류 모델은 검출된 객체(예: 사람)를 가로로 일정하게 나누어 각 부분을 라벨링할 수 있다. 이에 예를 들면, 첫번째 라벨은 0이 될 수 있고, 두번째 라벨은 1이 될 수 있다. 또한, 프로세서는 다층 퍼셉트론(Multi-Layer Perceptron, MLP) 층 하나로 이루어진 부분 분류화(part classification)를 이용하여 파트(part)를 분류할 수 있다. 미니배치(Mini-batch)가 주어졌을 때 파트(part) 분류 모델의 출력은 P 이고 P 는 $N_p * N_p$ 의 예측 스코어 행렬을 나타낼 수 있다. N_p 는 파트의 개수일 수 있고, 행렬 P 의 원소 P_{ij} 는 i 번째 파트(part)가 j 번째 파트(part)로 분류될 확률을 나타낼 수 있다. 또한, 프로세서는 파트(part) 분류 모델의 손실함수가 감소하는 방향으로 파트(part) 분류 모델을 트레이닝시킬 수 있다. 파트(part) 분류 모델의 손실함수는 수학적식1과 같이 표현될 수 있다.

수학식 1

$$\mathcal{L}_{\text{part}} = - \sum_{i=1}^{N_p} \log \left(\frac{\exp(P_{i p_i} / \tau_p)}{\sum_{j=1}^{N_p} \exp(P_{i j} / \tau_p)} \right)$$

[0077]

[0078] 전술한 수학식 1에서 P_i 는 i 번째 부분의 정답 라벨(GT label)을 나타낼 수 있고, $P_{i p_i}$ 는 정답 예측 확률을 나타낼 수 있으며, τ_p 는 하이퍼파라미터로 0.04일 수 있다.

[0079] 한편, 신뢰성 스코어는 예측 확률의 평균으로 결정될 수 있으며, 수학식2와 같이 표현될 수 있다.

수학식 2

$$\gamma = \frac{1}{N_p} \sum_{i=1}^{N_p} P_{ii}$$

[0080]

[0081] 전술한 수학식 2에서 P_{ii} 는 전술한 $P_{i p_i}$ 와 동일하게 정답 예측 확률을 나타낼 수 있고, 신뢰성 스코어 γ 는 파트(part)의 분류 정도를 또한 나타낼 수 있다. 이에 γ 값이 1인 경우, 모든 파트가 잘 분류되었고 검출된 타겟 객체는 정확하게 검출된 것을 의미할 수 있다. 만약 γ 값이 1보다 작은 경우, 검출된 타겟 객체는 부정확하게 검출되었음을 의미할 수 있다. 예를 들어, 검출된 타겟 객체가 사람이 아닌 배경인 경우, 파트(part) 분류 모델은 파트를 분류하기 어려워 정확하게 분류될 확률 또한 낮아질 수 있으므로 γ 는 1보다 작아질 수 있다.

[0082] 프로세서는 객체 식별 모델을 트레이닝시키기 위해서 ID 라벨이 필요하나 본 발명의 타겟 도메인에는 ID 라벨이 없으므로 ID 라벨의 할당이 필요하다. 프로세서는 밀도 기반 클러스터링(DBSCAN, Density-based spatial clustering of applications with noise)이라는 군집화 알고리즘을 이용하여 수도 라벨(예: 객체 식별 ID 라벨)을 할당할 수 있다. 이에 프로세서는 밀도 차이를 이용하여 클러스터링을 수행할 수 있다. 그러나 이러한 클러스터링 방식은 클러스터링의 정확도가 떨어지는 단점이 있다. 따라서 이를 개선하기 위하여 본 발명에서는 유일성(uniqueness) 논리를 적용하여 클러스터링된 클러스터를 정제할 수 있다.

[0083] 도 5는 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 방법 중 클러스터링 방식 및 클러스터링된 클러스터를 정제하는 방식의 예시도를 도시한다.

[0084] (510), (520)은 특징 스페이스(feature space)에서 이미지(I_1, I_2)로부터 검출된 타겟 객체의 특징 벡터를 이용한 분포를 나타낼 수 있다. 프로세서는 클러스터링 결과에 기초하여 ID라벨을 할당할 수 있다. 프로세서는, 예를 들면, (510), (520)에서 기준(Anchor) 객체를 (511) 객체로 지정할 수 있다. 클러스터링의 밀도는 이동수단 특징을 중심으로 집중될 수 있다. 이에 따라 클러스터링 알고리즘은 동일한 이동수단을 이용하는 (511) 객체, (512) 객체, (513) 객체, (514) 객체를 동일 객체로 분류하여 하나의 클러스터로 클러스터링할 수 있다. 하나의 클러스터에는 동일한 ID 라벨이 할당되므로 (511) 객체, (512) 객체, (513) 객체, (514) 객체는 동일한 ID 라벨을 가질 수 있다. 프로세서는 특징 추출 모델의 최적화를 위하여 다른 ID 라벨을 가진 객체들 중에 가장 거리가 가까운 객체를 부정(negative) 객체로 지정할 수 있다. 이에 따라 (510)에서 프로세서는 (515) 객체를 부정 객체로 지정할 수 있다. 또한, 프로세서는 같은 ID 라벨을 가진 객체들 중에 가장 거리가 먼 객체를 긍정(positive) 객체로 지정할 수 있다. 이에 따라 (510)에서 프로세서는 (512) 객체를 긍정 객체로 지정할 수 있다.

[0085] 그러나 (511) 객체와 (513) 객체는 상호 다른 객체이지만 같은 클러스터 내에 있으므로 이에 대한 구분이 필요하다. 이를 위하여 프로세서는 클러스터 내에 각 객체를 이미지 별로 그룹화할 수 있다. 이에 "같은 영상에 등장한 사람들은 서로 다른 사람이다"라는 유일성 논리에 따라 하나의 이미지(I_1)에 대한 그룹(521)에서 기준 객체인 (511) 객체 외에는 클러스터에서 제외시킬 수 있다. 따라서 정제된 클러스터링 결과는 (511) 객체, (512)

객체, (514) 객체는 동일한 ID 라벨을 가질 수 있다. 이러한 결과에 따라 프로세서는 부정(negative) 객체를 (515) 객체에서 (513) 객체로 변경할 수 있다.

[0086] 한편, 특징 추출 모델은 소스 도메인, 타겟 도메인 및 중간 도메인의 특징을 추출할 수 있다. 프로세서는 정제된 클러스터링 결과에 기초하여 산출된 손실함수가 감소하는 방향으로 특징 추출 모델을 트레이닝시킬 수 있다. 특징 추출 모델의 손실함수는 배치 하드 삼중항 손실(batch-hard triplet loss) 함수와 크로스 엔트로피 기반의 분류화 손실(classification loss) 함수로 구성될 수 있다.

[0087] 배치 하드 삼중항 손실(batch-hard triplet loss) 함수는 수학식3과 같이 표현될 수 있다.

수학식 3

$$L_{triplet} = \sum_{i=1}^P \sum_{a=1}^K \left[m + \overbrace{\max_{p=1 \dots K} D(x_a^i, x_p^j)}^{\text{hardest positive}} - \overbrace{\min_{\substack{j=1 \dots P \\ n=1 \dots N}} D(x_a^i, x_n^j)}^{\text{hardest negative}} \right]$$

[0088]

[0089] 전술한 수학식 3에서 프로세서는 배치 크기가 PK인 미니 배치에 대해 P개의 ID에서 K개의 특징 벡터를 무작위로 선택할 수 있다. x_a^i 는 기준 객체, x_a^j 는 긍정 객체, x_n^j 는 부정 객체를 나타낼 수 있다. $D(\cdot)$ 는 유클리드 거리를 나타내고, m은 하이퍼파라미터로 사용자에게 의해 지정될 수 있다.

[0090] 프로세서는 삼중항 손실(batch-hard triplet loss) 함수를 통해 전술한 도5에서의 (513) 객체(예: 부정 객체)와 (511) 객체(예: 기준 객체) 간의 거리가 멀어지도록 특징 추출 모델을 트레이닝시킬 수 있다. 또한, 프로세서는 삼중항 손실(batch-hard triplet loss) 함수를 통해 (512) 객체(예: 긍정 객체)와 (511) 객체(예: 기준 객체)간의 거리가 가까워지도록 특징 추출 모델을 트레이닝시킬 수 있다. 또한, 손실 함수는 실제 값과 예측 값 간의 차이를 수치화하는 함수로, 손실을 감소시키면서 모델을 최적화할 수 있다. 또한, 모델이 트레이닝되면 트레이닝에 의한 파라미터가 정해지는데, 모델을 최적화하는 것은 정해진 파라미터를 최적화하는 것과 동일한 의미일 수 있다. 이에 모델을 최적화하는 것은 손실함수를 줄여들게 하는 방향으로 파라미터를 업데이트하는 것을 의미할 수 있다. 결국, 특징 추출 모델을 트레이닝시키는 것은 특징 추출 모델의 파라미터를 업데이트하는 것을 의미할 수 있다.

[0091] 크로스 엔트로피 기반의 분류화 손실(classification loss) 함수는 수학식4와 같이 표현될 수 있다.

수학식 4

$$\mathcal{L}_{\text{class}} = - \sum_{k \in \{s, t\}} \left\{ \beta^k \log \left(\frac{\exp(C_{y^k} / \tau_c)}{\sum_{i=1}^N \exp(C_i / \tau_c)} \right) \right\}$$

[0092]

[0093] 전술한 수학식 4에서 C_i 는 i번째 ID일 확률을 나타내고, k는 소스도메인과 타겟 도메인을 나타낼 수 있다. y^k 는 소스도메인일 경우에 정답인 ID이며, 타겟 도메인일 경우에 클러스터링의 결과로 부여된 허위 ID인 것을 나타낼 수 있다. N은 ID의 개수를 나타낼 수 있고, τ_c 는 하이퍼파라미터로 1일 수 있다. β^k 는 입력된 도메인이 소스 도메인일 경우에 항상 정답 클러스터를 알고 있으므로 1일 수 있다. 그러나 입력된 도메인이 타겟 도메인일 경우에는 유일성 논리가 적용되어 같은 클러스터로 묶인 객체들(예: 같은 ID를 가진 객체들)이 같은 이미지에서 나온 경우에는 0일 수 있다. 또한, 입력된 도메인이 타겟 도메인일 경우, 같은 클러스터로 묶인 객체들(예: 같은 ID를 가진 객체들)이 다른 이미지에서 나온 경우에는 1일 수 있다.

[0094] 한편, 객체 식별 모델은 목적함수로 도메인 적응 손실(domain adaptation loss) 함수를 적용할 수 있다. 도메인 적응 손실 함수는 분류 부분과 특징 부분으로 나뉘어 손실을 산출할 수 있다. 분류 부분에 있어서 도메인 적응 손실 함수는 전술한 신뢰성 스코어 및 정제된 클러스터링 결과에 기초하여 수학식5와 같이 표현될 수

있다.

수학식 5

$$\mathcal{L}_{\text{adapt}}^c = -\gamma \sum_{k \in \{s, t\}} \left\{ \beta^k r^k \log \left(\frac{\exp(A_{y^k}^{\text{inter}} / \tau_a)}{\sum_{i=1}^N \exp(A_i^{\text{inter}} / \tau_a)} \right) \right\}$$

[0095]

전술한 수학식 5에서 A_i^{inter} 는 중간도메인의 특징이 i번째 ID의 라벨을 가지고 있을 예측 값을 나타낼 수 있다. r^k 은 중간 도메인을 만들 때 소스 도메인과 타겟 도메인을 얼마나 섞을지에 대한 가중치를 나타낼 수 있고, τ_a 는 하이퍼파라미터로 1일 수 있다.

[0096]

또한, 특징 부분에 있어서 도메인 적응 손실 함수는 신뢰성 스코어 및 도메인 간의 거리 값에 기초하여 수학식6과 같이 표현될 수 있다.

[0097]

수학식 6

$$\mathcal{L}_{\text{adapt}}^f = \gamma \sum_{k \in \{s, t\}} \{r^k \|g^k - g^{\text{inter}}\|_2\}$$

[0098]

전술한 수학식 6에서 g^k 는 소스 도메인과 타겟 도메인의 특징을 나타내고, g^{inter} 는 중간 도메인의 특징을 나타낼 수 있다. 특징 도메인 적응 손실 함수는 L2 norm(Euclidean Distance)를 이용하여 도메인의 특징 간의 거리를 산출할 수 있다. 이에 특징 도메인 적응 손실 함수는 중간 도메인의 특징과 소스 도메인의 특징 간의 거리 또는 중간 도메인의 특징과 타겟 도메인의 특징 간의 거리를 산출할 수 있다. 신뢰성 감마는 미니배치 내 샘플들에 각각 산출되며 도메인 적응 손실은 미니배치 내 모든 샘플들에 대해서 산출될 수 있다. 결과적으로 도메인 적응 손실함수는 중간도메인 특징들의 신뢰성 스코어가 부족할 때 모델 트레이닝에 영향을 주는 정도를 제한할 수 있다.

[0099]

도 6은 일 실시예에 따른 비지도 도메인 적응을 위한 객체 검색 모델의 트레이닝 방법의 추론 단계의 흐름도를 도시한다.

[0100]

프로세서는 객체 식별 모델의 트레이닝이 끝나면 테스트 데이터를 이용하여 추론 단계를 수행할 수 있다. 테스트 데이터에는 하나의 객체(예: 사람) 이미지가 있는 쿼리와 복수의 객체 영상이 포함된 갤러리 영상들이 포함될 수 있다. 추론 단계는 갤러리 영상들에서 쿼리와 같은 객체를 찾아내는 테스트 방식일 수 있다. 프로세서는 객체 검출 모델에 테스트 데이터를 입력하면, 테스트 데이터의 테스트 객체를 검출할 수 있다. 프로세서는 검출된 테스트 객체를 트레이닝된 특징 추출 모델(610)에 입력하여 테스트 특징 데이터(611)를 추출할 수 있다. 프로세서는 추출된 테스트 특징 데이터(611)에 기초하여 특징 데이터 간의 유사도를 계산할 수 있다. 프로세서는 계산된 유사도가 높은 순으로 쿼리와 같은 객체를 찾을 수 있다. 이를 통해 사용자는 복수의 객체 사이에서 쿼리와 같은 객체를 정확하게 찾을 수 있다. 추론 단계에서는 전술한 트레이닝 단계와 다르게 객체식별 모델(600)이 특징 추출 모델(610)으로만 구성될 수 있다.

[0101]

도 7은 일 실시예에 따른 트레이닝 방법이 적용된 객체 검색 모델의 성능을 다른 객체 검색 모델과 비교한 결과이다.

[0102]

본 발명의 일 실시예에 따른 트레이닝 방법이 적용된 객체 검색 모델의 성능은 벤치마크 데이터셋인 CUHK-SYSU와 PRW를 이용하여 확인할 수 있다. CUHK-SYSU 데이터 셋은 18,184개의 이미지를 제공하며 96,143개의 바운딩 박스 라벨과 8,432개의 ID 라벨을 포함할 수 있다. 그 중 학습데이터(예: 소스 도메인)는 11,204장이고 5,532개의 ID 라벨을 포함하며, 테스트데이터(예: 타겟 도메인)는 6,978장의 갤러리 이미지와 2,900장의 쿼리 이미지를 포함할 수 있다. PRW 데이터셋은 6개의 다른 카메라에서 촬영된 영상에서 캡처한 것이며, 11,816개의 영상 개수와 43,110개의 바운딩 박스 라벨과 932개의 ID 라벨이 포함할 수 있다. 학습데이터는 5,704장이고 482개의 ID 라벨을 포함하며, 테스트데이터는 6,112장의 갤러리 이미지와 2057개의 쿼리 이미지를 포함할 수 있

[0103]

다. 또한, 성능지표는 mAP와 Top-k를 성능지표로 사용했다. CUHK-SYSU 데이터 셋에서는 타겟 도메인이 라벨 없이도 지도 기계학습 모델 및 약한지도 기계학습 모델의 성능과 유사하게 달성할 수 있다. 또한, PRW 데이터 셋에서는 다른 지도 기계학습 모델 및 약한지도 기계학습 모델보다 가장 뛰어난 성능을 보여주고 있다.

[0104] 도 8은 일 실시예에 따른 트레이닝 방법을 순차적으로 적용한 객체 검색 모델의 성능을 비교한 결과이다.

[0105] 도 8에서 baseline기법은 객체 식별 모델에 기본적인 IDM만 적용한 것을 의미할 수 있다. with DA(domain adaptation) 기법은 부분 분류화를 이용하여 신뢰성 스코어를 산출하는 파트 분류 모델을 객체 식별 모델에 적용한 것을 의미할 수 있다. with CR(clustering refinement) 기법은 유일성 논리를 적용하여 클러스터를 정제하고 그 결과를 객체 식별 모델에 적용한 것을 의미할 수 있다. CUHK → PRW은 CHUCK 데이터 셋을 소스 도메인으로 설정하고, PRW 데이터 셋을 타겟 도메인으로 설정한 것을 의미할 수 있다. 또한, 이와 반대로 PRW → CUHK은 PRW 데이터 셋을 소스 도메인으로 설정하고, CHUCK 데이터 셋을 타겟 도메인으로 설정한 것을 의미할 수 있다. 성능결과는 각각의 기법이 개별적으로 적용된 것 또한 Baseline 보다 우수한 것을 보여주고 있다. 그러나 최종적인 성능 결과는 with DA 기법과 with CR 기법이 다 적용된 제안 기법(본 발명의 일 실시예에 따른 트레이닝 방법)이 제일 뛰어난 것을 보여주고 있다.

[0106] 이상에서 설명된 실시예들은 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치, 방법 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0107] 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, 가상 장치(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 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0108] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있으며 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.

[0109] 위에서 설명한 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 또는 복수의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

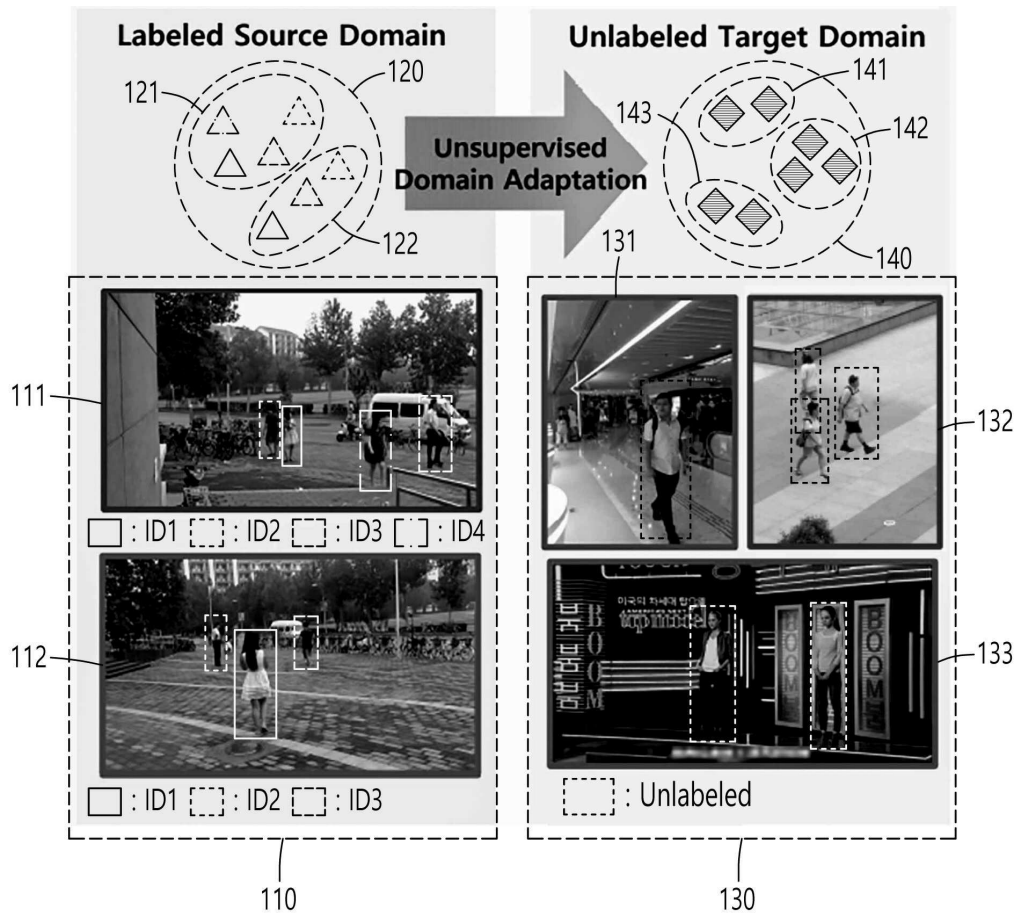
[0110] 이상과 같이 실시예들이 비록 한정된 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 이를 기초로 다양한 기술적 수정 및 변형을 적용할 수 있다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수

있다.

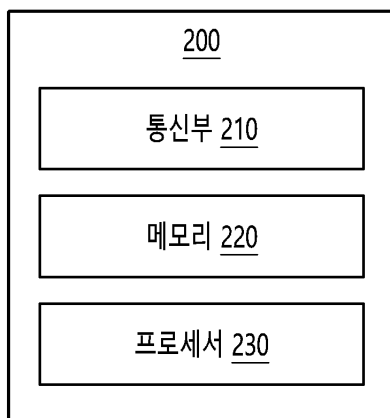
[0111] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

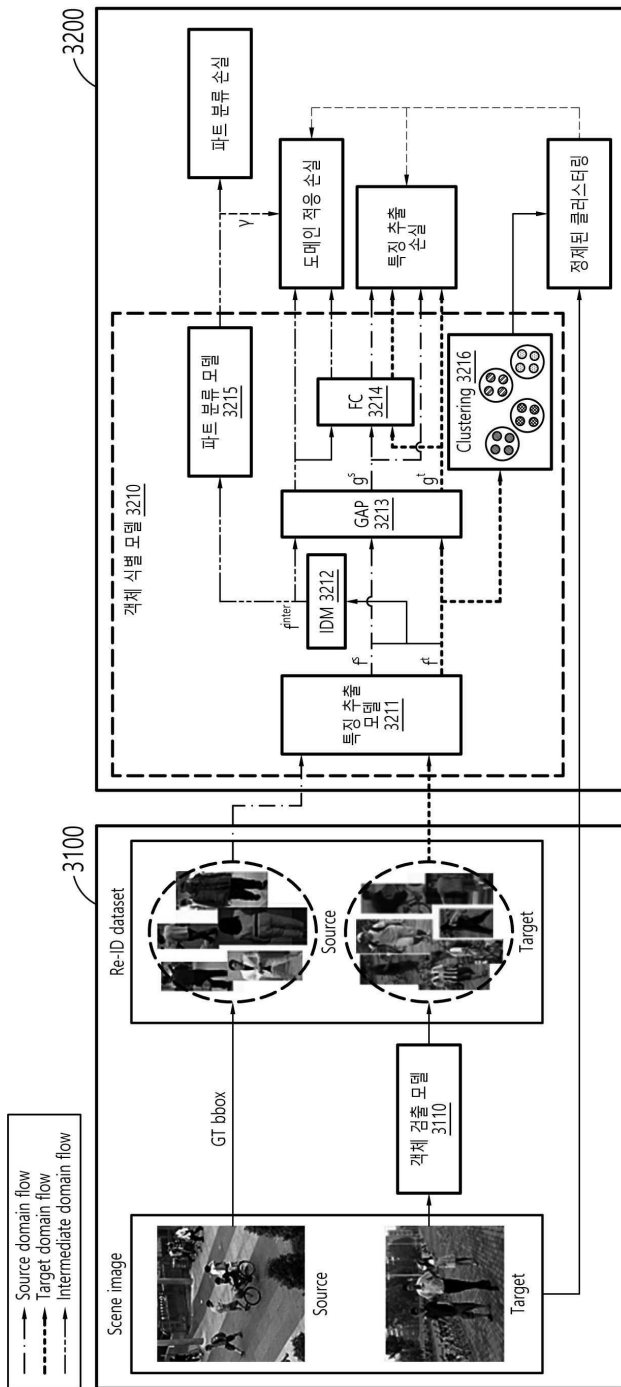
도면1



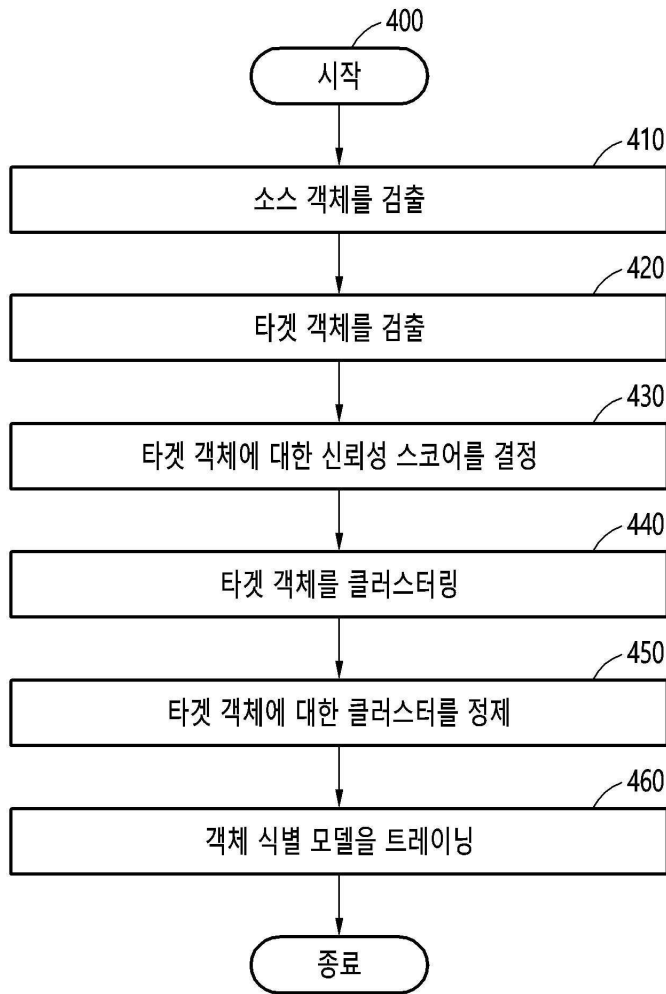
도면2



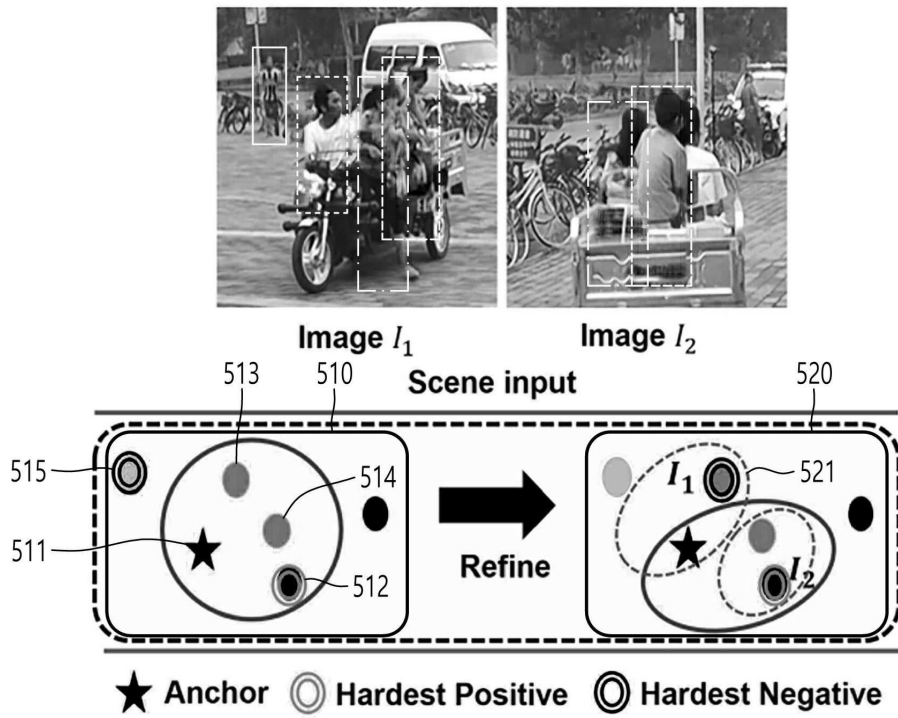
도면3



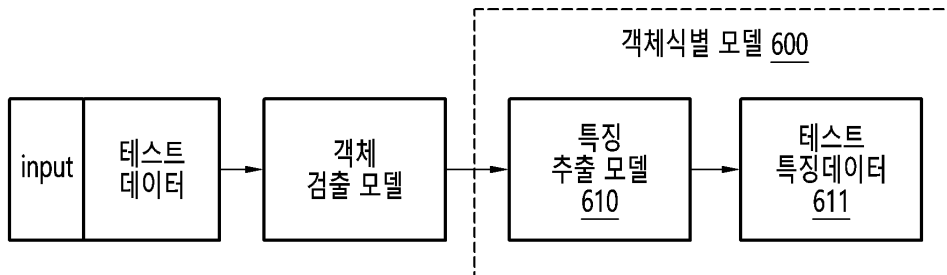
도면4



도면5



도면6



도면7

기법		CUHK-SYSU		PRW	
		mAP	Top-1	mAP	Top-1
지도학습	OIM	75.5	78.7	21.3	49.9
	IAN	76.3	80.1	23.0	61.9
	NAE	92.1	92.9	44.0	81.1
	HOIM	89.7	90.8	39.8	80.4
	SeqNet	94.8	95.7	47.6	87.6
	AlignPS	94.0	94.5	46.1	82.1
	AGWF	93.3	94.2	53.3	87.7
약한지도 학습	RSN	86.0	87.1	21.4	75.2
	CUCPS	81.1	83.2	41.7	86.0
	CGPS	80.0	82.3	16.2	68.0
제안기법		75.7	78.2	57.2	90.3

도면8

기법	CUHK → PRW		PRW → CUHK	
	mAP	Top-1	mAP	Top-1
Baseline	53.5	89.2	70.8	73.5
with DA (domain adaptation)	54.8	89.3	73.5	78.2
with CR (clustering refinement)	55.5	89.1	74.6	76.8
제안기법	57.2	90.3	75.7	76.8