



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0100200
(43) 공개일자 2023년07월05일

- | | |
|--|---|
| <p>(51) 국제특허분류(Int. Cl.)
 $G06V\ 10/774$ (2022.01) $G06N\ 20/00$ (2019.01)
 $G06V\ 10/25$ (2022.01) $G06V\ 10/40$ (2022.01)
 $G06V\ 10/46$ (2022.01) $G06V\ 10/70$ (2022.01)
 $G06V\ 10/74$ (2022.01) $G06V\ 40/10$ (2022.01)</p> <p>(52) CPC특허분류
 $G06V\ 10/774$ (2023.01)
 $G06N\ 20/00$ (2021.08)</p> <p>(21) 출원번호 10-2021-0189891
 (22) 출원일자 2021년12월28일
 심사청구일자 2021년12월28일</p> | <p>(71) 출원인
 울산과학기술원
 울산광역시 울주군 언양읍 유니스트길 50</p> <p>(72) 발명자
 심재영
 울산광역시 울주군 언양읍 유니스트길 50
 한병주
 울산광역시 울주군 언양읍 유니스트길 50
 고규현
 울산광역시 울주군 언양읍 유니스트길 50</p> <p>(74) 대리인
 특허법인 무한</p> |
|--|---|

전체 청구항 수 : 총 23 항

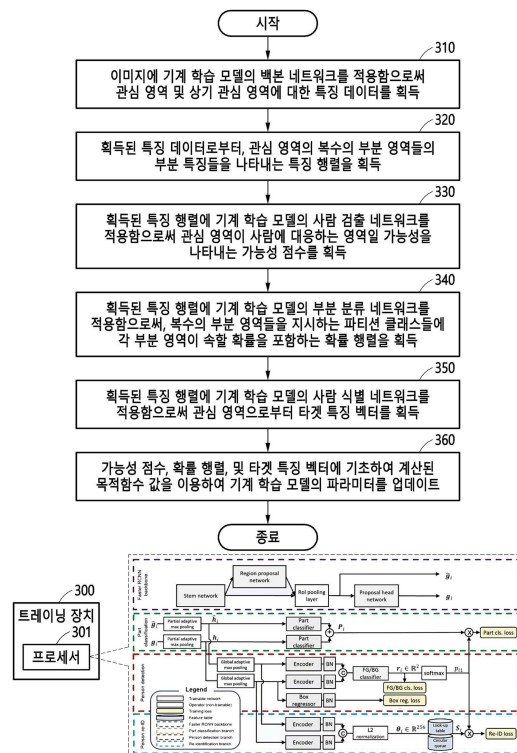
(54) 발명의 명칭 적응형 그래디언트 전파를 이용한 사람 검색을 위한 기계 학습 모델의 트레이닝 방법 및 장치

(57) 요약

일 실시예에 따른 기계 학습 모델을 트레이닝시키는 장치는, 이미지에 기계 학습 모델의 백본 네트워크(backbone network)를 적용함으로써 관심 영역(region of interest; ROI) 및 상기 관심 영역에 대한 특징 데이터(feature data)를 획득하고, 상기 획득된 특징 데이터로부터, 상기 관심 영역의 복수의 부분 영역들의 부분 특징들을 나타내는 특징 행렬을 획득

(뒷면에 계속)

대표도 - 도3



내는 특징 행렬을 획득하며, 상기 획득된 특징 행렬에 상기 기계 학습 모델의 사람 검출 네트워크(person detection network)를 적용함으로써 상기 관심 영역이 사람에 대응하는 영역일 가능성을 나타내는 가능성 점수(possibility score)를 획득하고, 상기 획득된 특징 행렬에 상기 기계 학습 모델의 부분 분류 네트워크(part classification network)를 적용함으로써, 상기 복수의 부분 영역들을 지시하는 파티션 클래스들(partition classes)에 각 부분 영역이 속할 확률을 포함하는 확률 행렬(probability matrix)을 획득하며, 상기 획득된 특징 행렬에 상기 기계 학습 모델의 사람 식별 네트워크(person re-identification network)를 적용함으로써 상기 관심 영역으로부터 타겟 특징 벡터(target feature vector)를 획득하고, 상기 가능성 점수, 상기 확률 행렬, 및 상기 타겟 특징 벡터에 기초하여 계산된 목적함수 값을 이용하여 상기 기계 학습 모델의 파라미터를 업데이트하는 프로세서를 포함할 수 있다.

(52) CPC특허분류

G06V 10/25 (2023.01)
G06V 10/40 (2023.01)
G06V 10/469 (2023.01)
G06V 10/70 (2023.01)
G06V 10/761 (2023.01)
G06V 40/103 (2022.01)
G06V 2201/07 (2022.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711126214
과제번호	2017-0-00667-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	디지털콘텐츠원천기술개발(R&D, 정보화)
연구과제명	이동 객체에 대한 증강현실 서비스 실현을 위한 정보-영상 정합기술 연구
기 여 율	1/3
과제수행기관명	울산과학기술원
연구기간	2021.01.01 ~ 2021.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711142910
과제번호	2020R1A2B5B01002725
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)(R&D)
연구과제명	딥러닝 기반 유리 왜곡 모델링 및 영상 복원 기술 개발
기 여 율	1/3
과제수행기관명	울산과학기술원
연구기간	2021.03.01 ~ 2022.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1711126319
과제번호	2020-0-01336-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성(R&D)
연구과제명	인공지능대학원지원(울산과학기술원)
기 여 율	1/3
과제수행기관명	울산과학기술원
연구기간	2021.01.01 ~ 2021.12.31

명세서

청구범위

청구항 1

이미지에 기계 학습 모델의 백본 네트워크(backbone network)를 적용함으로써 관심 영역(region of interest; ROI) 및 상기 관심 영역에 대한 특징 데이터(feature data)를 획득하는 단계;

상기 획득된 특징 데이터로부터, 상기 관심 영역의 복수의 부분 영역들의 부분 특징들을 나타내는 특징 행렬을 획득하는 단계;

상기 획득된 특징 행렬에 상기 기계 학습 모델의 사람 검출 네트워크(person detection network)를 적용함으로써 상기 관심 영역이 사람에 대응하는 영역일 가능성을 나타내는 가능성 점수(possibility score)를 획득하는 단계;

상기 획득된 특징 행렬에 상기 기계 학습 모델의 부분 분류 네트워크(part classification network)를 적용함으로써, 상기 복수의 부분 영역들을 지시하는 파티션 클래스들(partition classes)에 각 부분 영역이 속할 확률을 포함하는 확률 행렬(probability matrix)을 획득하는 단계;

상기 획득된 특징 행렬에 상기 기계 학습 모델의 사람 식별 네트워크(person re-identification network)를 적용함으로써 상기 관심 영역으로부터 타겟 특징 벡터(target feature vector)를 획득하는 단계; 및

상기 가능성 점수, 상기 확률 행렬, 및 상기 타겟 특징 벡터에 기초하여 계산된 목적함수 값을 이용하여 상기 기계 학습 모델의 파라미터를 업데이트하는 단계

를 포함하는 기계 학습 모델을 트레이닝시키는 방법.

청구항 2

제1항에 있어서,

상기 기계 학습 모델의 파라미터를 업데이트하는 단계는,

상기 가능성 점수에 기초하여 상기 관심 영역의 상기 사람 검출 네트워크에 대한 검출 손실 값(detection loss value)을 계산하는 단계;

상기 확률 행렬에 기초하여 상기 부분 분류 네트워크에 관한 부분 분류 손실 값(part classification loss value)을 계산하는 단계; 및

상기 타겟 특징 벡터에 기초하여 상기 사람 식별 네트워크에 관한 식별 손실 값(re-identification loss value)을 계산하는 단계를 포함하는,

기계 학습 모델을 트레이닝시키는 방법.

청구항 3

제2항에 있어서,

상기 부분 분류 손실 값을 계산하는 단계는,

상기 확률 행렬과 함께 상기 가능성 점수에 더 기초하여 상기 부분 분류 손실 값을 계산하는 단계를 포함하는,

기계 학습 모델을 트레이닝시키는 방법.

청구항 4

제3항에 있어서,

상기 가능성 점수에 더 기초하여 상기 부분 분류 손실 값을 계산하는 단계는,

상기 확률 행렬의 확률에 상기 가능성 점수를 곱한 값을 부분 영역이 파티션 클래스에 속할 가능성을 나타내는 보정된 확률로서 이용하여 상기 부분 분류 손실 값을 계산하는 단계를 포함하는,

기계 학습 모델을 트레이닝시키는 방법.

청구항 5

제2항에 있어서,

상기 식별 손실 값을 계산하는 단계는,

상기 타겟 특징 벡터와 함께 상기 가능성 점수에 더 기초하여 상기 식별 손실 값을 계산하는 단계를 포함하는,

기계 학습 모델을 트레이닝시키는 방법.

청구항 6

제5항에 있어서,

상기 가능성 점수에 더 기초하여 상기 식별 손실 값을 계산하는 단계는,

상기 타겟 특징 벡터와 다른 특징 벡터의 유사도에 상기 가능성 점수를 곱한 값을 상기 타겟 특징 벡터 및 상기 다른 특징 벡터 간의 보정된 유사도로서 이용하여 상기 식별 손실 값을 계산하는 단계를 포함하는,

기계 학습 모델을 트레이닝시키는 방법.

청구항 7

제2항에 있어서,

상기 식별 손실 값을 계산하는 단계는,

상기 관심 영역에 대한 사람의 신원(identity)을 지시하는 참값 신원 클래스(identity class)가 존재하는 경우에 응답하여, 상기 타겟 특징 벡터 및 다른 특징 벡터 간의 유사도 및 상기 타겟 특징 벡터 및 상기 참값 신원 클래스의 대표 특징 벡터(representative feature vector) 간의 유사도의 차이에 기초하여 상기 식별 손실 값을 계산하는 단계를 포함하는,

기계 학습 모델을 트레이닝시키는 방법.

청구항 8

제7항에 있어서,

상기 식별 손실 값을 계산하는 단계는,

상기 관심 영역이 상기 참값 신원 클래스로 분류될지 여부에 대한 바이너리 분류 손실 값(binary classification loss value)을 계산하는 단계를 더 포함하는,

기계 학습 모델을 트레이닝시키는 방법.

청구항 9

제2항에 있어서,

상기 식별 손실 값을 계산하는 단계는,

상기 관심 영역에 대한 참값 신원 클래스가 존재하지 않는 경우에 응답하여, 상기 타겟 특징 벡터 및 각 신원 클래스의 대표 특징 벡터 간의 유사도에 기초하여 상기 식별 손실 값을 계산하는 단계를 포함하는,

기계 학습 모델을 트레이닝시키는 방법.

청구항 10

제9항에 있어서,

상기 타겟 특징 벡터 및 각 신원 클래스의 대표 특징 벡터 간의 유사도에 기초하여 상기 식별 손실 값을 계산하는 단계는,

상기 이미지의 다른 관심 영역에 대한 신원 클래스가 존재하는 경우에 응답하여, 상기 다른 관심 영역의 타겟 특징 벡터 및 상기 다른 관심 영역의 신원 클래스의 대표 특징 벡터 간의 유사도를 계산하는 단계; 및

상기 관심 영역의 상기 타겟 특징 벡터 및 각 신원 클래스의 대표 특징 벡터 간의 유사도와 상기 다른 관심 영역의 타겟 특징 벡터 및 상기 다른 관심 영역의 신원 클래스의 대표 특징 벡터 간의 유사도의 차이에 기초하여, 상기 식별 손실 값을 계산하는 단계를 포함하는,

기계 학습 모델을 트레이닝시키는 방법.

청구항 11

제1항에 있어서,

계산된 목적함수 값을 이용하여 상기 기계 학습 모델의 파라미터를 업데이트하는 단계는,

상기 이미지로부터 복수의 관심 영역들이 획득된 경우에 응답하여, 복수의 관심 영역들에 대하여 계산된 목적함수 값들에 기초하여 상기 이미지에 대한 목적함수 값을 계산하는 단계; 및

상기 이미지에 대하여 계산된 목적함수 값을 이용하여 상기 기계 학습 모델의 파라미터를 업데이트하는 단계를 포함하는,

기계 학습 모델을 트레이닝시키는 방법.

청구항 12

하드웨어와 결합되어 제1항 내지 제11항 중 어느 하나의 항의 방법을 실행시키기 위하여 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램.

청구항 13

이미지에 기계 학습 모델의 백본 네트워크(backbone network)를 적용함으로써 관심 영역(region of interest; ROI) 및 상기 관심 영역에 대한 특징 데이터(feature data)를 획득하고, 상기 획득된 특징 데이터로부터, 상기 관심 영역의 복수의 부분 영역들의 부분 특징들을 나타내는 특징 행렬을 획득하며, 상기 획득된 특징 행렬에 상기 기계 학습 모델의 사람 검출 네트워크(person detection network)를 적용함으로써 상기 관심 영역이 사람에 대응하는 영역일 가능성을 나타내는 가능성 점수(possibility score)를 획득하고, 상기 획득된 특징 행렬에 상기 기계 학습 모델의 부분 분류 네트워크(part classification network)를 적용함으로써, 상기 복수의 부분 영역들을 지시하는 파티션 클래스들(partition classes)에 각 부분 영역이 속할 확률을 포함하는 확률 행렬(probability matrix)을 획득하며, 상기 획득된 특징 행렬에 상기 기계 학습 모델의 사람 식별 네트워크(person re-identification network)를 적용함으로써 상기 관심 영역으로부터 타겟 특징 벡터(target feature vector)

를 획득하고, 상기 가능성 점수, 상기 확률 행렬, 및 상기 타겟 특징 벡터에 기초하여 계산된 목적함수 값을 이용하여 상기 기계 학습 모델의 파라미터를 업데이트하는 프로세서를 포함하는 기계 학습 모델을 트레이닝시키는 장치.

청구항 14

제13항에 있어서,

상기 프로세서는,

상기 가능성 점수 및 참값(ground truth) 가능성 점수의 비교에 기초하여 상기 관심 영역의 상기 사람 검출 네트워크에 대한 검출 손실 값(detection loss value)을 계산하고, 상기 확률 행렬에 기초하여 상기 부분 분류 네트워크에 관한 부분 분류 손실 값(part classification loss value)을 계산하며, 상기 타겟 특징 벡터에 기초하여 상기 사람 식별 네트워크에 관한 식별 손실 값(re-identification loss value)을 계산하는,

기계 학습 모델을 트레이닝시키는 장치.

청구항 15

제14항에 있어서,

상기 프로세서는,

상기 확률 행렬과 함께 상기 가능성 점수에 더 기초하여 상기 부분 분류 손실 값을 계산하는,

기계 학습 모델을 트레이닝시키는 장치.

청구항 16

제15항에 있어서,

상기 프로세서는,

상기 확률 행렬의 확률에 상기 가능성 점수를 곱한 값을 부분 영역이 파티션 클래스에 속할 가능성을 나타내는 보정된 확률로서 이용하여 상기 부분 분류 손실 값을 계산하는,

기계 학습 모델을 트레이닝시키는 장치.

청구항 17

제14항에 있어서,

상기 프로세서는,

상기 타겟 특징 벡터와 함께 상기 가능성 점수에 더 기초하여 상기 식별 손실 값을 계산하는,

기계 학습 모델을 트레이닝시키는 장치.

청구항 18

제17항에 있어서,

상기 프로세서는,

상기 타겟 특징 벡터와 다른 특징 벡터의 유사도에 상기 가능성 점수를 곱한 값을 상기 타겟 특징 벡터 및 상기

다른 특징 벡터 간의 보정된 유사도로서 이용하여 상기 식별 손실 값을 계산하는,
기계 학습 모델을 트레이닝시키는 장치.

청구항 19

제14항에 있어서,

상기 프로세서는,

상기 관심 영역에 대한 사람의 신원(identity)을 지시하는 참값 신원 클래스(identity class)가 존재하는 경우에 응답하여, 상기 타겟 특징 벡터 및 다른 특징 벡터 간의 유사도 및 상기 타겟 특징 벡터 및 상기 참값 신원 클래스의 대표 특징 벡터(representative feature vector) 간의 유사도의 차이에 기초하여 상기 식별 손실 값을 계산하는,

기계 학습 모델을 트레이닝시키는 장치.

청구항 20

제19항에 있어서,

상기 프로세서는,

상기 관심 영역이 상기 참값 신원 클래스로 분류될지 여부에 대한 바이너리 분류 손실 값(binary classification loss value)을 계산하는,

기계 학습 모델을 트레이닝시키는 장치.

청구항 21

제14항에 있어서,

상기 프로세서는,

상기 관심 영역에 대한 참값 신원 클래스가 존재하지 않는 경우에 응답하여, 상기 타겟 특징 벡터 및 각 신원 클래스의 대표 특징 벡터 간의 유사도에 기초하여 상기 식별 손실 값을 계산하는,

기계 학습 모델을 트레이닝시키는 장치.

청구항 22

제18항에 있어서,

상기 프로세서는,

상기 이미지의 다른 관심 영역에 대한 신원 클래스가 존재하는 경우에 응답하여 상기 다른 관심 영역의 타겟 특징 벡터 및 상기 다른 관심 영역의 신원 클래스의 대표 특징 벡터 간의 유사도를 계산하고, 상기 관심 영역의 상기 타겟 특징 벡터 및 각 신원 클래스의 대표 특징 벡터 간의 유사도와 상기 다른 관심 영역의 타겟 특징 벡터 및 상기 다른 관심 영역의 신원 클래스의 대표 특징 벡터 간의 유사도의 차이에 기초하여 상기 식별 손실 값을 계산하는,

기계 학습 모델을 트레이닝시키는 장치.

청구항 23

제13항에 있어서,

상기 프로세서는,

상기 이미지로부터 복수의 관심 영역들이 획득된 경우에 응답하여, 복수의 관심 영역들에 대하여 계산된 목적함수 값들에 기초하여 상기 이미지에 대한 목적함수 값을 계산하고, 상기 이미지에 대하여 계산된 목적함수 값을 이용하여 상기 기계 학습 모델의 파라미터를 업데이트하는,

기계 학습 모델을 트레이닝시키는 장치.

발명의 설명

기술 분야

[0001] 이하, 사람 검색을 위한 기계 학습 모델의 트레이닝 방법 및 장치에 대한 기술이 개시된다.

배경 기술

[0003] 본 발명과 관련된 기술들은 2가지로 정리될 수 있다. 첫 번째는 사람 검출(person detection) 및 사람 식별(person re-identification)을 구분하여 학습하는 투-스텝(two-step) 기법들이고, 두 번째는 사람 검출 및 사람 식별을 통합하여 학습하는 end-to-end 기법들이다.

[0004] 투-스텝 기법에서, 먼저 영상에서 사람의 위치를 찾기 위하여 사람 검출 기술이 이용될 수 있다. 그 후에, 찾은 사람들은 사람이 누구인지 식별하기 위하여, 사람 식별 기술이 이용될 수 있다. 사람 검색 기술 분야에서 사용되는 벤치마크 PRW 데이터 셋이 제공될 수 있고, 사람 검출 기술 및 사람 식별 기술을 개별적으로 학습시키며 사람 검색(person search)의 결과를 분석하기 위하여 다양한 실험이 진행될 수 있다. 사람(person) 및 배경(background)을 분리하여 학습을 진행하기 위하여, 사람에 대응하는 영역을 따로 분할(segmentation) 작업을 함으로써 인공지능 모델은 사람에게 더 집중하여 트레이닝될 수 있는 기법이 이용될 수 있다. 다양한 스케일들로 추출된 특징들이 사람 검색에 이용될 수 있고, 사람 검출기 및 사람 식별기 간의 태스크(task) 불일치를 해결하기 위해 혼합된 학습 데이터 세트가 구축될 수 있다.

[0005] 엔드-투-엔드 기법들은 사람 검출 기술 및 사람 식별 기술을 통합하여 사람 검색을 수행하고, 엔드-투-엔드 기법에 따른 사람 검색 모델은 트레이닝에서도 사람 검출 및 사람 식별이 통합되어 트레이닝될 수 있다. 사람 검출 기술 및 사람 식별 기술이 통합된 딥러닝 기반 네트워크를 제안될 수 있고, 사람 검색 기술 분야에서 사용하는 벤치마크 CUHK-SYSU 데이터 세트가 제공될 수 있다. 배경은 사람 검색 모델을 학습시키는 과정에서 부정적인 영향을 끼칠 수 있고, 인공지능 모델이 트레이닝될 때 배경보다 사람의 특징을 추출할 수 있는 기법이 제안될 수 있다. 객체 분할(instance segmentation) 기법 및 사람의 관절을 예측하는 기법(keypoint detection) 기법이 추가적으로 도입됨으로써, 사람 검색을 위한 유용한 특징이 추출될 수 있다. 사람들의 추출된 특징들을 크기(magnitude)와 각도(angle)로 나누는 기법을 통해 사람 검출 기술 및 사람 식별 기술의 task 불일치가 완화될 수 있다. 사람 검색에서 사람의 일부분만 보이는, 가려진 사람(occluded person)에 대해 대응하기 위한 기법이 제안될 수 있고, 사람 검색 분야에서 데이터 세트 안에 가려진 사람들이 많이 등장하는 또 하나의 벤치마크 데이터 세트가 제공될 수 있다.

[0006] 엔드 투 엔드(End-to-end) 방식의 기계 학습 모델은, 사람 검출 및 사람 식별이 통합되어 함께 트레이닝될 수 있는 장점을 가질 수 있다. 다만, 사람의 공통된 특징을 찾는 사람 검출 기술과 사람들 간의 차별화된 특징을 찾는 사람 식별 기술 간의 충돌이 발생할 수 있다. 반대로, 투 스텝(two-step) 방식의 기계 학습 모델은 사람 검출 및 사람 식별에 대하여 별도로 설계되어 개별적으로 트레이닝될 수 있으므로 사람 검출 및 사람 식별 간의 충돌이 일어나지 않는 장점을 가질 수 있다. 다만, 두 개의 기술을 따로 학습을 해야 한다는 단점이 있다.

[0007] 본 개시에서, 사람 검색 기술에 관한 새로운 엔드 투 엔드 기반의 학습 네트워크가 제안될 수 있다. 본 발명에 따른 기계 학습 모델은, 사람 검출 및 사람 식별과 함께 사람의 부분 분류기(part classifier)를 통해 다수의 태스크들을 통합적으로 구성함으로써 사람마다 중요한 대표 특징을 추출하며, 사람 검출 및 사람 식별 태스크 간의 충돌을 방지할 수 있다.

[0008] 실험적으로, 제안된 기법은 기존의 엔드 투 엔드 기법들보다 높은 성능을 기록할 수 있다.

[0009] 컴퓨터의 GPU 성능 향상과 더불어 최근 딥러닝 기반 인공지능 기술이 다양한 영역에 활용될 수 있다. 구체적으로, 기계의 시각 지능을 연구하는 컴퓨터 비전 분야에서, 최근 들어 사람 식별(person re-identification) 기술은 개인 보안, 엔터테인먼트 서비스 등 다양한 분야에서 활용될 수 있어 중요한 문제로 간주될 수 있다. 하지만, 종래의 사람 식별 기술들은 정제된 환경에서 단일 사람만 촬영이 된 데이터 또는 제한된 촬영 환경을 요구할 수 있다. 그리하여 최근에는 이러한 문제를 해결하기 위하여, 데이터 및 촬영 환경에 대한 제약이 없고 한 영상에 다수의 사람이 등장하는 일반 영상에서 원하는 사람을 찾아내는 사람 검색(person search) 기술이 주목을 받을 수 있다. 사람 검색 기술은 사람 식별에서 다루는 영상과 다르게 CCTV 영상과 같이 다수의 사람들이 다양한 자세들과 다양한 각도들로 등장하기 때문에 종래의 사람 식별 문제보다 도전적인 문제로 간주될 수 있다. 사람 검색 기술은 다양한 감시 카메라들의 출현으로 매일 대규모의 수많은 보행자 이미지들이 제공됨에 따라 감시 및 추적에 활용도가 높고, 스마트폰에 있는 증강 현실 기술과도 결합할 수 있는 시각 엔터테인먼트 서비스 기술을 위한 핵심 기술일 수 있다.

발명의 내용

해결하려는 과제

과제의 해결 수단

- [0011] 일 실시예에 따른 기계 학습 모델을 트레이닝시키는 방법은, 이미지에 기계 학습 모델의 백본 네트워크(backbone network)를 적용함으로써 관심 영역(region of interest; ROI) 및 상기 관심 영역에 대한 특징 데이터(feature data)를 획득하는 단계, 상기 획득된 특징 데이터로부터, 상기 관심 영역의 복수의 부분 영역들의 부분 특징들을 나타내는 특징 행렬을 획득하는 단계, 상기 획득된 특징 행렬에 상기 기계 학습 모델의 사람 검출 네트워크(person detection network)를 적용함으로써 상기 관심 영역이 사람에 대응하는 영역일 가능성을 나타내는 가능성 점수(possibility score)를 획득하는 단계, 상기 획득된 특징 행렬에 상기 기계 학습 모델의 부분 분류 네트워크(part classification network)를 적용함으로써, 상기 복수의 부분 영역들을 지시하는 파티션 클래스들(partition classes)에 각 부분 영역이 속할 확률을 포함하는 확률 행렬(probability matrix)을 획득하는 단계, 상기 획득된 특징 행렬에 상기 기계 학습 모델의 사람 식별 네트워크(person re-identification network)를 적용함으로써 상기 관심 영역으로부터 타겟 특징 벡터(target feature vector)를 획득하는 단계, 및 상기 가능성 점수, 상기 확률 행렬, 및 상기 타겟 특징 벡터에 기초하여 계산된 목적함수 값을 이용하여 상기 기계 학습 모델의 파라미터를 업데이트하는 단계를 포함할 수 있다.
- [0012] 상기 기계 학습 모델의 파라미터를 업데이트하는 단계는, 상기 가능성 점수에 기초하여 상기 관심 영역의 상기 사람 검출 네트워크에 대한 검출 손실 값(detection loss value)을 계산하는 단계, 상기 확률 행렬에 기초하여 상기 부분 분류 네트워크에 관한 부분 분류 손실 값(part classification loss value)을 계산하는 단계, 및 상기 타겟 특징 벡터에 기초하여 상기 사람 식별 네트워크에 관한 식별 손실 값(re-identification loss value)을 계산하는 단계를 포함할 수 있다.
- [0013] 상기 부분 분류 손실 값을 계산하는 단계는 상기 확률 행렬과 함께 상기 가능성 점수에 더 기초하여 상기 부분 분류 손실 값을 계산하는 단계를 포함할 수 있다.
- [0014] 상기 가능성 점수에 더 기초하여 상기 부분 분류 손실 값을 계산하는 단계는, 상기 확률 행렬의 확률에 상기 가능성 점수를 곱한 값을 부분 영역이 파티션 클래스에 속할 가능성을 나타내는 보정된 확률로서 이용하여 상기 부분 분류 손실 값을 계산하는 단계를 포함할 수 있다.
- [0015] 상기 식별 손실 값을 계산하는 단계는, 상기 타겟 특징 벡터와 함께 상기 가능성 점수에 더 기초하여 상기 식별 손실 값을 계산하는 단계를 포함할 수 있다.
- [0016] 상기 가능성 점수에 더 기초하여 상기 식별 손실 값을 계산하는 단계는, 상기 타겟 특징 벡터와 다른 특징 벡터의 유사도에 상기 가능성 점수를 곱한 값을 상기 타겟 특징 벡터 및 상기 다른 특징 벡터 간의 보정된 유사도로서 이용하여 상기 식별 손실 값을 계산하는 단계를 포함할 수 있다.
- [0017] 상기 식별 손실 값을 계산하는 단계는, 상기 관심 영역에 대한 사람의 신원(identity)을 지시하는 참값 신원 클래스(identity class)가 존재하는 경우에 응답하여, 상기 타겟 특징 벡터 및 다른 특징 벡터 간의 유사도 및 상

기 타겟 특징 벡터 및 상기 참값 신원 클래스의 대표 특징 벡터(representative feature vector) 간의 유사도의 차이에 기초하여 상기 식별 손실 값을 계산하는 단계를 포함할 수 있다.

[0018] 상기 식별 손실 값을 계산하는 단계는, 상기 관심 영역이 상기 참값 신원 클래스로 분류될지 여부에 대한 바이너리 분류 손실 값(binary classification loss value)을 계산하는 단계를 더 포함할 수 있다.

[0019] 상기 식별 손실 값을 계산하는 단계는, 상기 관심 영역에 대한 참값 신원 클래스가 존재하지 않는 경우에 응답하여, 상기 타겟 특징 벡터 및 각 신원 클래스의 대표 특징 벡터 간의 유사도에 기초하여 상기 식별 손실 값을 계산하는 단계를 포함할 수 있다.

[0020] 상기 타겟 특징 벡터 및 각 신원 클래스의 대표 특징 벡터 간의 유사도에 기초하여 상기 식별 손실 값을 계산하는 단계는, 상기 이미지의 다른 관심 영역에 대한 신원 클래스가 존재하는 경우에 응답하여, 상기 다른 관심 영역의 타겟 특징 벡터 및 상기 다른 관심 영역의 신원 클래스의 대표 특징 벡터 간의 유사도를 계산하는 단계 및 상기 관심 영역의 상기 타겟 특징 벡터 및 각 신원 클래스의 대표 특징 벡터 간의 유사도와 상기 다른 관심 영역의 타겟 특징 벡터 및 상기 다른 관심 영역의 신원 클래스의 대표 특징 벡터 간의 유사도의 차이에 기초하여, 상기 식별 손실 값을 계산하는 단계를 포함할 수 있다.

[0021] 계산된 목적함수 값을 이용하여 상기 기계 학습 모델의 파라미터를 업데이트하는 단계는, 상기 이미지로부터 복수의 관심 영역들이 획득된 경우에 응답하여, 복수의 관심 영역들에 대하여 계산된 목적함수 값들에 기초하여 상기 이미지에 대한 목적함수 값을 계산하는 단계 및 상기 이미지에 대하여 계산된 목적함수 값을 이용하여 상기 기계 학습 모델의 파라미터를 업데이트하는 단계를 포함할 수 있다.

[0022] 일 실시예에 따른 기계 학습 모델을 트레이닝시키는 장치는, 이미지에 기계 학습 모델의 백본 네트워크(backbone network)를 적용함으로써 관심 영역(region of interest; ROI) 및 상기 관심 영역에 대한 특징 데이터(feature data)를 획득하고, 상기 획득된 특징 데이터로부터, 상기 관심 영역의 복수의 부분 영역들의 부분 특징들을 나타내는 특징 행렬을 획득하며, 상기 획득된 특징 행렬에 상기 기계 학습 모델의 사람 검출 네트워크(person detection network)를 적용함으로써 상기 관심 영역이 사람에게 대응하는 영역일 가능성을 나타내는 가능성 점수(possibility score)를 획득하고, 상기 획득된 특징 행렬에 상기 기계 학습 모델의 부분 분류 네트워크(part classification network)를 적용함으로써, 상기 복수의 부분 영역들을 지시하는 파티션 클래스들(partition classes)에 각 부분 영역이 속할 확률을 포함하는 확률 행렬(probability matrix)을 획득하며, 상기 획득된 특징 행렬에 상기 기계 학습 모델의 사람 식별 네트워크(person re-identification network)를 적용함으로써 상기 관심 영역으로부터 타겟 특징 벡터(target feature vector)를 획득하고, 상기 가능성 점수, 상기 확률 행렬, 및 상기 타겟 특징 벡터에 기초하여 계산된 목적함수 값을 이용하여 상기 기계 학습 모델의 파라미터를 업데이트하는 프로세서를 포함할 수 있다.

[0023] 상기 프로세서는, 상기 가능성 점수 및 참값(ground truth) 가능성 점수의 비교에 기초하여 상기 관심 영역의 상기 사람 검출 네트워크에 대한 검출 손실 값(detection loss value)을 계산하고, 상기 확률 행렬에 기초하여 상기 부분 분류 네트워크에 관한 부분 분류 손실 값(part classification loss value)을 계산하며, 상기 타겟 특징 벡터에 기초하여 상기 사람 식별 네트워크에 관한 식별 손실 값(re-identification loss value)을 계산할 수 있다.

[0024] 상기 프로세서는, 상기 확률 행렬과 함께 상기 가능성 점수에 더 기초하여 상기 부분 분류 손실 값을 계산할 수 있다.

[0025] 상기 프로세서는, 상기 확률 행렬의 확률에 상기 가능성 점수를 곱한 값을 부분 영역이 파티션 클래스에 속할 가능성을 나타내는 보정된 확률로서 이용하여 상기 부분 분류 손실 값을 계산할 수 있다.

[0026] 상기 프로세서는, 상기 타겟 특징 벡터와 함께 상기 가능성 점수에 더 기초하여 상기 식별 손실 값을 계산할 수 있다.

[0027] 상기 프로세서는, 상기 타겟 특징 벡터와 다른 특징 벡터의 유사도에 상기 가능성 점수를 곱한 값을 상기 타겟 특징 벡터 및 상기 다른 특징 벡터 간의 보정된 유사도로서 이용하여 상기 식별 손실 값을 계산할 수 있다.

[0028] 상기 프로세서는, 상기 관심 영역에 대한 사람의 신원(identity)을 지시하는 참값 신원 클래스(identity class)가 존재하는 경우에 응답하여, 상기 타겟 특징 벡터 및 다른 특징 벡터 간의 유사도 및 상기 타겟 특징 벡터 및 상기 참값 신원 클래스의 대표 특징 벡터(representative feature vector) 간의 유사도의 차이에 기초하여 상기 식별 손실 값을 계산할 수 있다.

- [0029] 상기 프로세서는, 상기 관심 영역이 상기 참값 신원 클래스로 분류될지 여부에 대한 바이너리 분류 손실 값(binary classification loss value)을 계산할 수 있다.
- [0030] 상기 프로세서는, 상기 관심 영역에 대한 참값 신원 클래스가 존재하지 않는 경우에 응답하여, 상기 타겟 특징 벡터 및 각 신원 클래스의 대표 특징 벡터 간의 유사도에 기초하여 상기 식별 손실 값을 계산할 수 있다.
- [0031] 상기 프로세서는, 상기 이미지의 다른 관심 영역에 대한 신원 클래스가 존재하는 경우에 응답하여 상기 다른 관심 영역의 타겟 특징 벡터 및 상기 다른 관심 영역의 신원 클래스의 대표 특징 벡터 간의 유사도를 계산하고, 상기 관심 영역의 상기 타겟 특징 벡터 및 각 신원 클래스의 대표 특징 벡터 간의 유사도와 상기 다른 관심 영역의 타겟 특징 벡터 및 상기 다른 관심 영역의 신원 클래스의 대표 특징 벡터 간의 유사도의 차이에 기초하여 상기 식별 손실 값을 계산할 수 있다.
- [0032] 상기 프로세서는, 상기 이미지로부터 복수의 관심 영역들이 획득된 경우에 응답하여, 복수의 관심 영역들에 대하여 계산된 목적함수 값들에 기초하여 상기 이미지에 대한 목적함수 값을 계산하고, 상기 이미지에 대하여 계산된 목적함수 값을 이용하여 상기 기계 학습 모델의 파라미터를 업데이트할 수 있다.

발명의 효과

- [0035] 본 발명은 투-스텝으로 사람 검색 모델을 트레이닝시킴으로써 발생할 수 있는 불필요한 시간 소모 및 비효율성을 고려하여 엔드 투 엔드(end-to-end) 기법으로 설계될 수 있다. 일반적으로 투 스텝 기법은 엔드 투 엔드 기법들보다 높은 성능을 도출하는 경향이 있으나, 본 발명에 따른 사람 검색 모델은 엔드 투 엔드 기법임에도 불구하고 투 스텝 기법들보다 높은 성능을 기록할 수 있다.

도면의 간단한 설명

- [0037] 도 1은 일 실시예에 따른 기계 학습 모델을 이용한 사람 검색 방법을 나타낸다.
- 도 2는 비교 실시예 및 일 실시예에 따른 사람 검색을 위한 기계 학습 모델의 트레이닝을 나타낸다.
- 도 3은 일 실시예에 따른 사람 검색을 위한 기계 학습 모델의 트레이닝을 나타낸다.
- 도 4는 일 실시예에 따른 관심 영역이 분할된 부분 영역 및 부분 분류를 나타낸다.
- 도 5는 사람 검출의 신뢰도에 따른 기계 학습 모델의 목적함수 계산을 나타낸다.
- 도 6는 일 실시예에 따른 사람 검색을 위한 기계 학습 모델의 벤치마크(benchmark) 데이터 세트들을 나타낸다.
- 도 7은 비교 실시예 및 일 실시예에 따른 기계 학습 모델의 성능을 정성적으로 비교하는 것을 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0038] 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.
- [0039] 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.
- [0040] 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.
- [0041] 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 존재함으로 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을

을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

- [0042] 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.
- [0043] 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.
- [0044] 사람 검색(person search)은 이미지로부터 사람에 대응하는 관심 영역(Region of Interest; ROI)을 추출하는 사람 검출(person detection) 및 추출된 관심 영역에 대한 사람 식별(person re-identification)을 함께 수행하는 것을 나타낼 수 있다. 이미지는, 복수의 사람들이 함께 촬영된 이미지로서, 사람들에게 대한 복수의 관심 영역들을 포함할 수 있다. 다만, 이에 한정하는 것은 아니고 이미지는 한 사람이 촬영된 이미지로서 하나의 관심 영역을 포함할 수도 있다.
- [0045] 사람 검출(person detection)은 이미지로부터 사람에 대응하는 영역을 포함하는 관심 영역을 추출하는 동작을 나타낼 수 있다. 복수의 관심 영역들이 하나의 이미지로부터 추출될 수 있다. 관심 영역은 미리 정의된 형태로 추출될 수 있다. 예시적으로, 관심 영역은 직사각형의 형태로 추출될 수 있다. 다만, 이에 한정하는 것은 아니고, 관심 영역이 다른 형태(예를 들어, 타원)로 추출되거나 관심 영역마다 결정되는 형태로 추출될 수도 있다.
- [0046] 사람 식별(person re-identification)은 검출된 관심 영역이 쿼리(query) 사람에 대한 것인지를 결정하는 동작을 나타낼 수 있다. 관심 영역이 쿼리 사람에 대한 것인지 아닌지 여부는 특징 벡터(feature vector) 간의 유사도에 기초하여 결정될 수 있다. 쿼리 사람에 대한 특징 벡터는 외부 장치로부터 수신될 수 있다. 후술하겠으나, 이미지로부터 추출된 관심 영역의 타겟 특징 벡터(target feature vector)는 기계 학습 모델을 통해 획득될 수 있다. 쿼리 사람의 특징 벡터 및 이미지로부터 추출된 관심 영역의 타겟 특징 벡터 간의 유사도가 임계 이상인 경우, 해당 관심 영역은 쿼리 사람에 대한 관심 영역으로 결정될 수 있다.
- [0047] 도 1은 일 실시예에 따른 기계 학습 모델을 이용한 사람 검색 방법을 나타낸다.
- [0048] 사람 검색 장치는 이미지에 대하여 사람 검색을 수행하는 장치를 나타낼 수 있다. 사람 검색 장치는 기계 학습 모델을 이용할 수 있다.
- [0049] 기계 학습 모델은 이미지(110)에 적용됨으로써 복수의 관심 영역들을 출력할 수 있다. 복수의 관심 영역들은 미리 정해진 개수(예를 들어, 도 1의 N)로 추출될 수 있다. 또한, 기계 학습 모델은 추출된 복수의 관심 영역들 각각에 대한 타겟 특징 벡터 및 가능성 점수(possibility score)를 더 출력할 수 있다. 예시적으로, 기계 학습 모델은 이미지로부터 도3의 Faster R-CNN 백본 네트워크(backbone network), 사람 검출 네트워크(person detection network), 부분 분류 네트워크(part classification network), 및 사람 식별 네트워크(person re-identification network)에 기초하여 구현될 수 있다. 일 실시예에 따른 기계 학습 모델의 구조는 도3에서 자세히 설명한다.
- [0050] 복수의 관심 영역들은 이미지의 일부 영역으로서, 사람에 대응하는 관심 영역들, 배경에 대응하는 관심 영역들, 및 사람에 대응하는 영역과 일부 중첩되는 영역들을 포함할 수 있다. 배경은 이미지 중 사람에 대응하는 영역을 제외한 나머지를 나타낼 수 있다. 또한, 복수의 이미지들은 하나의 사람에 대한 서로 중첩되는 복수의 관심 영역들을 포함할 수 있다.
- [0051] 가능성 점수(possibility score)는 사람 검출에 관한 평가 지표로서, 추출된 해당 관심 영역에 대한 신뢰도(confidence)를 나타낼 수 있다. 가능성 점수는 사람에 대응하는 영역이 해당 관심 영역에 포함될 가능성을 나타낼 수 있다. 예를 들어, 사람에 대응하는 영역이 포함된 제1 관심 영역에 대한 가능성 점수는 1에 가까운 값을 가질 수 있다. 다른 예를 들어, 배경에 대응하는 영역이 포함된 제2 관심 영역에 대한 가능성 점수는 0에 가까운 값을 가질 수 있다.
- [0052] 사람 검색 장치는 가능성 점수에 기초하여 획득된 복수의 관심 영역들 중 적어도 하나를 후보 관심 영역(121)으로 결정할 수 있다. 사람 검색 장치는 가능성 점수가 임계 이상인 경우, 해당 영역을 사람에 대응하는 영역으로

로 추정되는 후보 관심 영역(121)으로 결정할 수 있다. 다만, 도 1의 이미지(120)에서 나타난 바와 같이, 한 사람에 대한 서로 중첩되는 복수의 관심 영역들이 포함될 수 있으므로, 중첩되는 복수의 후보 관심 영역들 중 하나를 선택하는 것이 요구될 수 있다.

- [0053] 사람 검색 장치는 후보 관심 영역들이 서로 중첩되는 경우, 중첩되는 후보 관심 영역들 중 하나의 후보 관심 영역을 추정 관심 영역(131)으로 결정할 수 있다. 예를 들어, 후보 관심 영역들 간의 IoU(Intersection over Union)이 임계 이상인 경우, 최대 가능성 점수를 갖는 후보 관심 영역이 추정 관심 영역으로 선택될 수 있다.
- [0054] 사람 검색 장치는 추정 관심 영역의 타겟 특징 벡터 및 쿼리 사람의 특징 벡터 간의 유사도에 기초하여, 추정 관심 영역이 쿼리 사람에 대한 영역인지 여부를 결정할 수 있다.
- [0055] 이하에서는, 일 실시예에 따라 사람 검색을 위한 엔드-투-엔드(end-to-end) 기반의 기계 학습 모델을 트레이닝시키는 방법을 도 2 내지 도 4을 참조하여 상세히 설명한다.
- [0056] 도 2는 비교 실시예 및 일 실시예에 따른 사람 검색을 위한 기계 학습 모델의 트레이닝을 나타낸다.
- [0057] 비교 실시예에 따른 기계 학습 모델(210)은 사람 검출(person detection) 및 사람 식별(person re-identification)을 함께 학습할 수 있다. 사람 검출은 사람의 공통된 특징을 찾는 것을 요구할 수 있다. 사람 식별은 사람들 간의 차별화된 특징을 찾는 것을 요구할 수 있다. 따라서, 비교 실시예에 따른 엔드-투-엔드 기법의 기계 학습 모델(210)에서, 사람 검출 태스크(task) 및 사람 식별 태스크 간의 충돌이 발생할 수 있다.
- [0058] 일 실시예에 따른 기계 학습 모델(220)에서, 사람 검출 및 사람 식별 간의 태스크 불일치를 감소시키기 위한 기법이 적용될 수 있다. 도 2에서 나타난 바와 같이, 기계 학습 모델(220)은 사람 식별에서 사람 검출의 결과(예를 들어, 사람 검출의 신뢰도)를 반영할 수 있다. 또한, 기계 학습 모델(220)은 사람의 부분에 관한 특징을 추출하기 위하여, 사람에 대응하는 영역이 분할된 부분 영역으로부터 특징을 추출하는 부분 분류(part classification)를 이용하여 학습할 수 있다. 기계 학습 모델(220)은 부분 분류에서도 사람 검출의 결과(예를 들어, 사람 검출의 신뢰도)를 반영할 수 있다. 실험적으로, 일 실시예에 따른 기계 학습 모델(220)은 비교 실시예에 따른 기계 학습 모델(210)보다 향상된 성능을 가질 수 있다.
- [0059] 도 3는 일 실시예에 따른 사람 검색을 위한 기계 학습 모델의 트레이닝을 나타낸다.
- [0060] 트레이닝 장치(300)는, 사람 검색을 위한 기계 학습 모델의 트레이닝을 수행하기 위한 장치로서, 프로세서(301)를 포함할 수 있다.
- [0061] 일 실시예에 따른 사람 검색을 위한 기계 학습 모델은 백본 네트워크(backbone network), 사람 검출 네트워크(person detection network), 부분 분류 네트워크(part classification network), 및 사람 식별 네트워크(person re-identification network)를 포함할 수 있다.
- [0062] 단계(310)에서, 프로세서(301)는 이미지에 백본 네트워크를 적용함으로써 관심 영역, 관심 영역의 특징 데이터(feature data), 및 관심 영역의 보다 더 깊은 특징 데이터를 추출할 수 있다. 백본 네트워크는 Faster R-CNN 및 프로포절 헤드 네트워크(proposal head network)를 포함할 수 있다. Faster R-CNN은 이미지로부터 관심 영역(예를 들어, 이미지의 복수의 관심 영역들 중 i 번째 관심 영역) 및 관심 영역의 특징 데이터(예를 들어, i 번째 관심 영역의 특징 데이터($\hat{g}_i$))를 추출할 수 있다. 프로포절 헤드 네트워크는 Faster R-CNN으로부터 추출된 관심 영역의 특징 데이터보다 더 깊은 특징 데이터(예를 들어, 더 많은 개수의 채널(channel)을 갖는 i 번째 관심 영역의 특징 데이터(g_i))를 추출할 수 있다. 결과적으로, 기계 학습 모델은 멀티-스케일 피처링(multi-scale featurig)에 기초하여 관심 영역의 제1 특징 데이터($\hat{g}_i$) 및 제2 특징 데이터(g_i)를 모두 이용할 수 있다.
- [0063] 단계(320)에서, 프로세서(301)는 백본 네트워크로부터 획득된 특징 데이터로부터 특징 행렬(feature matrix)을 획득할 수 있다. i 번째 관심 영역의 특징 행렬($\hat{h}_i$)은 특징 데이터($\hat{g}_i$)에 부분적 적응형 최대 풀링(partial adaptive max pooling)을 적용함으로써 획득될 수 있다. 또한, i 번째 관심 영역의 다른 특징 행렬(h_i)은 더 깊은 특징 데이터(g_i)에 부분적 적응형 최대 풀링을 적용함으로써 획득될 수도 있다. 특징 행렬은 관심 영역의 부분 특징들을 포함할 수 있다. 부분 특징은, 관심 영역이 분할된 부분 영역의 특징 데이터를 나타낼 수 있

다. 예를 들어, i 번째 관심 영역의 특징 행렬($\mathbf{h}_i$)의 j 번째 로우($\mathbf{h}_{ij}$)는, i 번째 관심 영역의 j 번째 파티션(partition)의 특징 데이터를 나타낼 수 있다.

[0064] 단계들(330, 340, 350)에서, 프로세서(301)는 획득된 특징 행렬에 사람 검출 네트워크, 부분 분류 네트워크, 및 사람 식별 네트워크를 적용할 수 있다.

[0065] 단계(330)에서, 프로세서(301)는 특징 행렬에 사람 검출 네트워크를 적용함으로써 가능성 점수를 획득할 수 있다. 사람 검출 네트워크는 글로벌 적응형 최대 풀링(global adaptive max pooling), 인코더(encoder), 배치 نرم(batch normalization; BN), 연결(concatenation)(도 3에서 C로 나타남), 및 사람 및 배경 분류기(foreground and background classifier)(도 3에서 FG/BG classifier로 나타남)를 포함할 수 있다. 가능성 점수는 관심 영역이 사람에 대응하는 영역일 가능성을 나타낼 수 있다. 가능성 점수는, 백본 네트워크를 적용함으로써 획득된 관심 영역이 사람에 대응하는 영역일 사람 검출에 대한 신뢰도(confidence)를 나타낼 수 있다.

[0066] 참고로, 사람 검출 네트워크에 관하여, 분류 손실(classification loss)(도 3에서, FG/BG cls. loss로 표현됨) 및 회귀 손실(regression loss)(도 3에서, Box reg. loss로 표현됨)이 계산될 수 있다. 기계 학습 모델의 목적함수는 분류 손실 및 회귀 손실을 포함할 수 있다.

[0067] 단계(340)에서, 프로세서(301)는 특징 행렬에 부분 분류 네트워크를 적용함으로써 확률 행렬(probability matrix)을 획득할 수 있다. 확률 행렬은, 관심 영역의 부분 영역이 파티션 클래스에 속할 확률을 포함할 수 있다. 부분 영역 및 파티션 클래스는 도 4에서 설명한다. 프로세서(301)는 확률 행렬에 기초하여 부분 분류 손실 값(part classification loss value)(도 3에서, Part cls. loss로 표시됨)을 계산할 수 있다. 부분 분류 손실 값의 계산은 도 5에서 설명한다.

[0068] 단계(350)에서, 프로세서(301)는 특징 행렬에 사람 식별 네트워크를 적용함으로써 타겟 특징 벡터(target feature vector)를 획득할 수 있다. 타겟 특징 벡터는, 관심 영역의 특징을 나타내는 벡터로서, 쿼리 사람의 특징 벡터와 비교 대상일 수 있다. 부분 영역의 특징 데이터를 가지는 특징 행렬로부터 획득된 타겟 특징 벡터는, 관심 영역 중 일부 영역(예를 들어, 하나의 부분 영역)에 치우치지 않고 부분 영역들 각각의 특징을 나타낼 수 있는 특징 벡터로 획득될 수 있다.

[0069] 단계(360)에서, 프로세서(301)는 목적함수 값을 이용하여 기계 학습 모델의 파라미터를 업데이트할 수 있다. 목적함수는, 사람 검출 네트워크에 관하여 계산된 분류 손실 값(도 3에서 FG/BG cls. loss로 표현됨) 및 회귀 손실 값(도 3에서 Box reg. loss로 표현됨), 부분 분류 네트워크에 관하여 계산된 부분 분류 손실 값, 사람 식별 네트워크에 관하여 계산된 식별 손실 값을 포함할 수 있다. 기계 학습 모델의 목적함수 계산은 도 5에서 후술한다.

[0070] 도 4는 일 실시예에 따른 관심 영역이 분할된 부분 영역 및 부분 분류를 나타낸다.

[0071] 부분 영역들(410)은 관심 영역이 분할됨으로써 획득될 수 있다. 예를 들어, 부분 영역은 관심 영역을 복수의 파티션들(partitions)로 수평 분할한 관심 영역의 일부를 나타낼 수 있다.

[0072] 파티션 클래스는 사람에 대응하는 관심 영역 중 부분 영역이 차지하는 위치(예를 들어, 관심 영역 중 첫번째 부분 영역)를 지시할 수 있다. 참고로, 파티션 클래스는 머리 또는 다리와 같은 특정한 신체 부위를 지시할 수도 있으나, 이에 한정되지 않고, 신체 부위 중 하나로 정의하기 어려운 신체의 일부를 나타낼 수 있다.

[0073] 파티션 클래스는 배경에 대응하는 관심 영역의 부분 영역을 지시하는 더미 클래스(dummy class)를 포함할 수 있다. 예를 들어, 관심 영역이 미리 정의된 개수(N_p)로 분할된 부분 영역들이 있는 경우, 사람에 대응하는 관심 영역의 부분 영역들의 미리 정의된 개수(N_p)개의 파티션 클래스들 및 배경에 대응하는 관심 영역의 부분 영역들의 1개의 파티션 클래스를 포함하는 N_p+1 개의 파티션 클래스들이 존재할 수 있다.

[0074] 부분 분류 네트워크가 적용됨으로써 획득된 확률 행렬은 부분 영역이 파티션 클래스에 속할 확률을 포함할 수 있다. 이미지의 i 번째 관심 영역에 대한 확률 행렬($\mathbf{P}_i$)의 j 행 k 열 원소(P_{ijk})는, i 번째 관심 영역의 j 번째 부분 영역이 k 번째 파티션 클래스에 속할 확률을 나타낼 수 있다.

[0075] 도 5는 사람 검출의 신뢰도에 따른 기계 학습 모델의 목적함수 계산을 나타낸다.

[0076] 크로스 엔트로피 손실(cross-entropy loss)은 분류를 위한 기계 학습 모델의 트레이닝에서 사용될 수 있다.

$\mathbf{S}_i$ 는 i 번째 관심 영역의 스코어 벡터를 나타내고, $\mathbf{S}_i$ 의 j 번째 원소 S_{ij} 는 i 번째 관심 영역이 j 번째 클래스에 속할 스코어를 나타낼 때, 소프트맥스 확률들(softmax probabilities)을 갖는 크로스 엔트로피 손실은 다음과 같이 정의될 수 있다:

$$\mathcal{L}(\mathbf{S}_i, y_i) = -\log \left(\frac{\exp(S_{iy_i}/\tau)}{\sum_j \exp(S_{ij}/\tau)} \right)$$

여기서, y_i 는 i 번째 관심 영역의 참값 클래스 인덱스를 나타내고, τ 는 손실 함수의 민감도(sensitivity)를 제어하는 온도 계수(temperature coefficient)를 나타낼 수 있다.

크로스 엔트로피 손실은 다음과 같이 표현될 수도 있다:

$$\mathcal{L}(\mathbf{S}_i, y_i) = \log \left(1 + \sum_{j \neq y_i} \exp((S_{ij} - S_{iy_i})/\tau) \right)$$

여기서, 크로스 엔트로피 손실은 참값 클래스의 스코어에 대한 다른 클래스들의 상대적인 스코어들에 의존하는 것을 알 수 있다. 손실 함수는, 관심 영역이 다른 클래스들에 속할 것으로 예측되는 스코어가 참값 클래스의 스코어보다 상대적으로 높을 때(예를 들어, $S_{ij} > S_{iy_i}, \forall j \neq y_i$ 일 때), 높은 값을 가질 수 있다.

트레이닝 동안, 기계 학습 모델의 파라미터들은 다음과 같은 체인 룰(chain rule)에 기초하여 후방 전파 그래디언트들(back-propagated gradients)에 의하여 업데이트될 수 있다:

$$\frac{\partial \mathcal{L}(\mathbf{S}_i, y_i)}{\partial \boldsymbol{\pi}} = \sum_k \frac{\partial \mathcal{L}(\mathbf{S}_i, y_i)}{\partial S_{ik}} \cdot \frac{\partial S_{ik}}{\partial \boldsymbol{\pi}}$$

여기서, $\boldsymbol{\pi}$ 는 기계 학습 모델의 파라미터들을 나타낼 수 있다.

$F_k(\mathbf{S}_i, y_i)$ 는 크로스 엔트로피 손실에 기초한 손실 함수의 S_{ik} 에 대한 도함수(derivative)로 정의되면, 기계 학습 모델의 파라미터와 연관된 그래디언트($\frac{\partial S_{ik}}{\partial \boldsymbol{\pi}}$)의 가중치 함수(weighting function)로 간주될 수 있다. 가중치 함수의 동작을 해석하기 위하여, $F_k(\mathbf{S}_i, y_i)$ 는 다음과 같이 상대적인 스코어들의 소프트맥스와 유사한(softmax-like) 함수로 도출될 수 있다:

$$F_k(\mathbf{S}_i, y_i) = \frac{\partial \mathcal{L}(\mathbf{S}_i, y_i)}{\partial S_{ik}} = \begin{cases} \frac{1}{\tau} \frac{\exp((S_{ik} - S_{iy_i})/\tau)}{1 + \sum_{j \neq y_i} \exp((S_{ij} - S_{iy_i})/\tau)}, & \text{if } k \neq y_i, \\ -\sum_{j \neq y_i} F_j(\mathbf{S}_i, y_i), & \text{otherwise.} \end{cases}$$

도 5의 그래프들(510, 520)는 참값 클래스는 1이고 다른 클래스는 0인 간단한 이항 분류(binary classification) 태스크의 예시에 따른, 일반 가중치 함수($F_0(\mathbf{S}_i, 1)$) 및 적응형 그래디언트 가중치 함수($G_0(\mathbf{S}_i, \alpha_i, 1)$)의 모양을 나타낼 수 있다. 그래프들(510, 520)는 참값 클래스일 확률(S_{i1})가 0.5이고 온도 계수(τ)가 0.3으로 고정될 때 사람에 대응하는 다른 클래스일 확률(S_{i0})의 변화에 따른 일반 가중치 함수($F_0(\mathbf{S}_i, 1)$) 및 다양한 신뢰도(α_i) 값들에 따른 적응형 그래디언트 가중치 함수($G_0(\mathbf{S}_i, \alpha_i, 1)$)의 플롯(plot)을 나타낼 수 있다.

기계 학습 모델을 트레이닝할 때 그래디언트들의 흐름을 제어하는 표준 가중치 함수로서 $F_k(\mathbf{S}_i, y_i)$ 를 고려하면, 사람 식별 네트워크 및 부분 분류 네트워크를 적응적으로 트레이닝시키기 위하여, 다음과 같은 적응형 그래디언트 가중치 함수(adaptive gradient weighting function; AGWF)가 이용될 수 있다:

- [0089] $G_k(\mathbf{S}_i, \alpha_i, y_i) = \alpha_i \cdot F_k(\alpha_i \cdot \mathbf{S}_i, y_i)$
- [0090] 여기서, α_i 는 i 번째 관심 영역에 대한 사람 검출의 신뢰도로서, i 번째 관심 영역이 사람에 대응하는 영역이 포함될 확률(예를 들어, 가능성 점수)을 나타낼 수 있다.
- [0091] 그래프(510)에서, 다른 클래스일 확률(S_{i0})이 참값 클래스일 확률(S_{i1})보다 상대적으로 클 때, 기계 학습 모델의 파라미터들은, 온도 계수의 역수($\frac{1}{T}$)에 가까운 큰 가중치를 갖는 최대로 그래디언트($\frac{\partial S_{ik}}{\partial \pi}$)를 전파시킴으로써 상당히(significantly) 벗어날(deviate) 수 있다. 대조적으로, 다른 클래스일 확률(S_{i0})가 참값 클래스일 확률(S_{i1})보다 작을 때, 0에 가까운 작은 가중치는 그래디언트 후방 전파를 억제할 수 있다.
- [0092] 그래프(520)는, 다양한 신뢰도(α_i) 값들에 따른 적응형 그래디언트 가중치 함수($G_0(\mathbf{S}_i, \alpha_i, 1)$)의 모양들을 나타낼 수 있다. 일반 가중치 함수($F_0(\mathbf{S}_i, 1)$)의 그래프(510)에 비하여, 입력에 대한 민감도(sensitivity) 및 출력 범위는 모두 나타날 수 있다. 신뢰도(α_i)가 감소함에 따라, 적응형 그래디언트 가중치 함수의 출력 범위가 감소하고 적응형 그래디언트 가중치 함수의 모양이 완만해질 수 있다. 적응형 그래디언트 가중치 함수는, 사람 검출의 품질뿐만 아니라 공유된 특징 맵들의 트레이닝 상태(state)도 내재적으로(implicitly) 반영할 수 있다. 예를 들어, 네트워크가 트레이닝의 초기 단계(early phase)이고 검출된 관심 영역이 상당한 부분이 배경에 대응하는 영역을 포함하는 경우, 신뢰도(α_i)는 낮아질 수 있고, 적응형 그래디언트 가중치 함수는 식별 손실 값으로부터의 후방 전파를 억제하면서 신원 클래스들에 대한 관심 영역의 특징 벡터 유사도들에 둔감(Insensitive)해질 수 있다.
- [0093] 사람 검색에서, 사람 검출에 의하여 획득된 관심 영역에 기초하여 사람 식별이 수행되므로, 사람 검출의 품질(quality)은 사람 식별의 품질에 큰 영향을 미칠 수 있다. 특히, 엔드-투-엔드 기반의 기계 학습 모델에서, 사람 검출 및 사람 식별에 대하여 공유되는 특징 맵(feature map)이 사람 검출에 대하여 과적합(overfit)되는 경우, 사람 식별에 대한 특징 맵의 대표성이 열화(degrade)될 수 있다. 일 실시예에 따른 기계 학습 모델의 트레이닝은 사람 검출에 대한 신뢰도(예를 들어, 사람 검출 네트워크에 의하여 획득된 가능성 점수)에 따라 사람 식별 네트워크 및 부분 분류 네트워크를 적응적으로(adaptively) 트레이닝시키는 것을 포함할 수 있다. 이하, 사람 검출에 대한 신뢰도에 따라 사람 식별 네트워크 및 부분 분류 네트워크를 적응적으로 트레이닝하기 위한 목적함수 계산에 대하여 설명한다.
- [0094] 기계 학습 모델의 목적함수는 검출 손실 값(detection loss value), 식별 손실 값(re-identification loss value), 및 부분 분류 손실 값(part classification loss value)을 포함할 수 있다.
- [0095] 검출 손실 값은, 사람 검출 네트워크에 관한 손실로서, 도 3에서 전술한 바와 같이 분류 손실 및 회귀 손실을 포함할 수 있다. 분류 손실은 관심 영역이 사람에 대응하는 영역인지 아닌지를 분류하는 것에 대한 손실 값을 나타낼 수 있다. 회귀 손실은 관심 영역의 경계(예를 들어, 바운딩 박스)의 회귀에 대한 손실 값을 나타낼 수 있다. 검출 손실 값은 가능성 점수에 기초하여 계산될 수 있다.
- [0096] 식별 손실 값은, 사람 식별 네트워크에 관한 손실로서, 관심 영역의 타겟 특징 벡터에 기초하여 계산될 수 있다. 식별 손실 값은, 적응형 그래디언트 가중치 함수로서, i 번째 관심 영역의 사람 검출에 대한 신뢰도(α_i)에 기초하여 계산될 수 있다. 식별 손실 값은 타겟 특징 벡터와 함께 가능성 점수에 더 기초하여 계산될 수 있다.
- [0097] 식별 손실 값은 관심 영역의 타겟 특징 벡터 및 참값 신원 클래스(identity class)의 대표 특징 벡터(representative feature vector) 간의 유사도가 타겟 특징 벡터 및 다른 신원 클래스의 특징 벡터 간의 유사도보다 높도록 기계 학습 모델을 트레이닝시키기 위하여 설계될 수 있다. 신원 클래스는 관심 영역에 대한 사람의 신원을 지시하는 클래스를 나타낼 수 있다. 기계 학습 모델의 트레이닝을 위하여, 신원 클래스의 대표 특징 벡터는 외부 장치 또는 트레이닝 장치의 메모리로부터 획득될 수 있다. 관심 영역의 참값 신원 클래스가 존재하는 경우에 대하여 주로 설명하였으나, 후술될 바와 같이, 관심 영역의 참값 신원 클래스가 존재하지 않는 경우에도 식별 손실 값이 계산될 수 있다.

[0098] 식별 손실 값은 타겟 특징 벡터 및 다른 특징 벡터 간의 유사도를 이용할 수 있다. 타겟 특징 벡터 및 다른 특징 벡터 간의 유사도에 가능성 점수를 곱한 값은 보정된 유사도로서 이용될 수 있다.

[0099] 다른 특징 벡터는, 신원 클래스가 알려진 경우에 응답하여, 신원 클래스에 매핑된 대표 특징 벡터(representative feature vector)를 포함할 수 있다. 신원 클래스에 매핑된 대표 특징 벡터는 룩 업 테이블(look up table)에 저장될 수 있다. 또한, 다른 특징 벡터는 신원 클래스가 알려지지 않은(unlabeled) 특징 벡터를 포함할 수 있다. 라벨링되지 않은 다른 특징 벡터는 원형 큐(circular queue)에 저장될 수 있다.

[0100] 식별 손실 값은 관심 영역에 대하여 계산될 수 있고, 관심 영역에 대한 참값 신원 클래스는 존재할 수도 있고, 존재하지 않을 수도 있다.

[0101] 관심 영역의 참값 신원 클래스가 존재하는 경우, 식별 손실 값은 기계 학습 모델의 사람 식별 네트워크로부터 획득된 관심 영역의 타겟 특징 벡터 및 참값 신원 클래스의 대표 특징 벡터 간의 유사도를 기준 유사도로서 계산될 수 있다. 식별 손실 값은 타겟 특징 벡터 및 다른 특징 벡터 간의 유사도와 기준 유사도의 차이에 기초하여 계산될 수 있다.

[0102] 가능성 점수에 더 기초한 식별 손실 값은, 기준 유사도(예를 들어, 타겟 특징 벡터 및 참값 신원 클래스의 대표 특징 벡터 간의 유사도)에 가능성 점수를 곱한 값을 보정된 기준 유사도로서 이용할 수 있다. 또한, 가능성 점수에 더 기초한 식별 손실 값은 타겟 특징 벡터 및 다른 특징 벡터 간의 유사도에 가능성 점수를 곱한 값을 보정된 유사도로서 이용할 수 있다. 예를 들어, 참값 신원 클래스가 존재하는 i 번째 관심 영역에 대한 식별 손실 값은 다음과 같이 계산될 수 있다:

$$\mathcal{L}_{id}(\mathbf{S}_i, \alpha_i, y_i) = \log \left(1 + \sum_{\substack{k \in \{\Omega_L \cup \Omega_Q\} \\ k \neq y_i}} \exp(\alpha_i(S_{ik} - S_{iy_i})/\tau_1) \right)$$

[0103] 여기서, Ω_L 는 신원 클래스가 알려진 다른 특징 벡터(예를 들어, 룩 업 테이블에 저장된 다른 특징 벡터)의 인덱스 세트를 나타내고, Ω_Q 는 신원 클래스가 알려지지 않은 다른 특징 벡터(예를 들어, 원형 큐에 저장된 다른 특징 벡터)의 인덱스 세트를 나타내고, τ_1 는 온도 계수를 나타낼 수 있다.

[0105] 참값 신원 클래스가 존재하는 관심 영역의 식별 손실 값($\mathcal{L}_{id}$)은 상대적인 유사도에 의존하고, 기준 유사도(예를 들어, 타겟 특징 벡터 및 참값 신원 클래스의 대표 특징 벡터 간의 유사도(S_{iy_i}))가 그 자체로 높은 스코어에 도달하기 전에 포화(saturate)될 수 있다. 기준 유사도가 높은 스코어에 도달하기 전에 포화되는 것을 방지하기 위하여, 다음과 같은 바이너리 분류 손실 값(binary classification loss value)($\mathcal{L}_{bin}$)이 이용될 수 있다:

$$\mathcal{L}_{bin}(\mathbf{S}_i, \alpha_i, y_i) = \log(1 + \exp((s_\kappa - \alpha_i S_{iy_i})/\tau_2))$$

[0107] 여기서, 상수 s_κ 는 1의 최대 유사도로 설정될 수 있고, τ_2 는 온도 계수를 나타낼 수 있다.

[0108] 관심 영역의 참값 신원 클래스가 존재하지 않는 경우, 식별 손실 값은 다른 관심 영역에 대한 타겟 특징 벡터 및 다른 관심 영역의 신원 클래스의 대표 특징 벡터 간의 유사도를 기준 유사도로 계산될 수 있다. 다른 관심 영역은 사람에 대응하는 관심 영역으로서, 관심 영역이 획득된 이미지와 같은 미니 배치(mini-batch)에 속하는 이미지로부터 획득될 수 있다. 미니 배치에 속하는 이미지에서 획득된 신원 클래스가 존재하는 다른 관심 영역이 복수개인 경우, 다른 관심 영역의 타겟 특징 벡터 및 다른 관심 영역의 신원 클래스의 대표 특징 벡터 간의 유사도들의 평균이 기준 유사도로서 이용될 수 있다.

[0109] 참고로, 미니 배치는 기계 학습 모델의 파라미터를 업데이트할 목적함수를 계산하는 단위를 나타낼 수 있다. 미니 배치는 하나 이상의 이미지들을 포함할 수 있다. 하나의 미니 배치로부터 획득된 관심 영역에 대한 목적함수를 계산할 수 있다. 하나의 미니 배치에 대하여 계산된 목적함수에 기초하여 기계 학습 모델의 파라미터가 업데이트될 수 있다.

[0110] 식별 손실 값은 관심 영역의 타겟 특징 벡터 및 다른 특징 벡터 간의 유사도와 기준 유사도의 차이에 기초하여

계산될 수 있다. 다른 특징 벡터는, 신원 클래스가 알려진 경우에 응답하여, 신원 클래스에 매핑된 대표 특징 벡터(representative feature vector)를 포함할 수 있다. 신원 클래스에 매핑된 대표 특징 벡터는 룩 업 테이블(look up table)에 저장될 수 있다.

[0111] 가능성 점수에 더 기초한 식별 손실 값은, 기준 유사도(예를 들어, 다른 관심 영역에 대한 타겟 특징 벡터 및 다른 관심 영역의 신원 클래스의 대표 특징 벡터 간의 유사도)에 가능성 점수를 곱한 값을 보정된 기준 유사도로서 이용할 수 있다. 또한, 가능성 점수에 더 기초한 식별 손실 값은 타겟 특징 벡터 및 다른 특징 벡터 간의 유사도에 가능성 점수를 곱한 값을 보정된 유사도로서 이용할 수 있다. 예를 들어, 참값 신원 클래스가 존재하지 않는 i 번째 관심 영역에 대한 식별 손실 값은 다음과 같은 식에 따라 계산될 수 있다:

$$\mathcal{L}_{un}(\mathbf{S}_i, \alpha_i) = \log \left(1 + \sum_{k \in \Omega_L} \exp(\alpha_i(S_{ik} - \bar{S})/\tau_1) \right)$$

$$\bar{S} = \frac{1}{|\Psi_L|} \sum_{i \in \Psi_L} S_{iy_i}$$

[0113] 여기서, Ψ_L 는 미니 배치의 관심 영역 중 신원 클래스가 존재하는 관심 영역의 인덱스 세트를 나타낼 수 있다.

[0114] 식별 손실 값은 이미지에 대하여 계산될 수 있다. 이미지로부터 복수의 관심 영역들이 획득된 경우, 각 관심 영역에 대하여 계산된 식별 손실 값에 기초하여 이미지에 대한 식별 손실 값이 계산될 수 있다. 예를 들어, 관심 영역에 대한 식별 손실 값들이 평균됨으로써 이미지에 대한 식별 손실 값이 계산될 수 있다. 신원 클래스가 존재하는 관심 영역들의 식별 손실 값들의 평균 및 신원 클래스가 존재하지 않는 관심 영역들의 식별 손실 값들의 평균을 덧셈함으로써 이미지에 대한 식별 손실 값이 계산될 수 있다. 이미지에 대한 식별 손실 값은 다음의 식으로 정의될 수 있다:

$$\mathcal{L}_{re-id}(\mathbf{S}_i, \alpha_i, y_i) = \frac{1}{|\Psi_L|} \sum_{i \in \Psi_L} \{ \mathcal{L}_{id}(\mathbf{S}_i, \alpha_i, y_i) + \mathcal{L}_{bin}(\mathbf{S}_i, \alpha_i, y_i) \} + \frac{\lambda}{|\Psi_U|} \sum_{i \in \Psi_U} \mathcal{L}_{un}(\mathbf{S}_i, \alpha_i)$$

[0116] 여기서, Ψ_L 는 미니 배치의 관심 영역 중 신원 클래스가 존재하는 관심 영역의 인덱스 세트를 나타낼 수 있고, Ψ_U 는 미니 배치의 관심 영역 중 신원 클래스가 존재하지 않는 관심 영역의 인덱스 세트를 나타낼 수 있고, λ 는 0.1로 실험적으로 설정될 수 있다.

[0117] 결과적으로, 적응형 그래디언트 가중치 함수는 사람 검출의 결과가 신뢰할 수 있을 때(예를 들어, 가중치 점수가 클 때), 기계 학습 모델의 파라미터를 업데이트하기 위하여 식별 손실 값의 기여도(contribution)를 증가시킬 수 있다. 도 3에서 나타난 바와 같이, 사람 검출 네트워크 및 사람 식별 네트워크 간의 스킵 연결(skip connection)은 식별 손실 값에 대한 적응형 그래디언트 가중치 함수의 적용(application)을 나타낼 수 있다.

[0118] 부분 분류 손실 값은, 부분 분류 네트워크에 관한 손실로서, 확률 행렬에 기초하여 계산될 수 있다. 부분 분류 손실 값은 부분 영역이 참값 파티션 클래스에 속할 확률이 다른 파티션 클래스에 속할 확률보다 크도록 기계 학습 모델을 트레이닝시키기 위하여 설계될 수 있다. 부분 분류 손실 값은 기준 확률 및 부분 영역이 다른 파티션 클래스에 속할 확률 간의 차이에 기초하여 계산될 수 있다. 기준 확률은 부분 분류 손실 값은 참값 파티션 클래스에 속할 확률을 나타낼 수 있다.

[0119] 부분 분류 손실 값은 확률 행렬과 함께 가능성 점수에 더 기초하여 계산될 수 있다. 가능성 점수에 더 기초한 부분 분류 손실 값은, 기준 확률(예를 들어, 부분 영역이 참값 파티션 클래스에 속할 확률)에 가능성 점수를 곱한 값을 보정된 기준 확률로서 이용할 수 있다. 또한, 가능성 점수에 더 기초한 부분 분류 손실 값은, 부분 영역이 다른 파티션 클래스에 포함될 확률에 가능성 점수를 곱한 값을 보정된 확률로서 이용할 수 있다. 부분 분류 손실 값은 크로스 엔트로피 기반의 분류 손실을 이용할 수 있다. 예를 들어, 확률 행렬 및 가능성 점수에 기초한 부분 분류 손실 값은 다음과 같이 계산될 수 있다:

$$\mathcal{L}_{\text{part}} = \frac{1}{N} \sum_{\substack{i \in \Psi \\ 1 \leq j \leq N_p}} \log \left(1 + \sum_{k \neq z_{ij}} \exp(\alpha_i (P_{ijk} - P_{ijz_{ij}}) / \tau_3) \right)$$

[0120]

[0121]

여기서, P_{ijk} 는 i 번째 관심 영역의 확률 행렬의 j 행 k 열 원소(예를 들어, i 번째 관심 영역의 j 번째 부분 영역이 k 번째 파티션 클래스에 속할 확률)를 나타내고, Ψ 는 미니 배치로부터 획득된 관심 영역들의 세트를 나타내며, τ_3 는 온도 계수를 나타내고, N 는 미니 배치의 총 파티션들의 개수를 나타내며, N_p 는 관심 영역마다 부분 영역의 개수를 나타낼 수 있다.

[0122]

도 6는 일 실시예에 따른 사람 검색을 위한 기계 학습 모델의 벤치마크(benchmark) 데이터 세트들을 나타낸다.

[0123]

데이터 세트(610)는, CUHK-SYSU 데이터 세트로서, 영화 및 도시의 길거리 이미지들을 포함할 수 있다. 총 18,184개의 이미지들이 포함될 수 있고, 96,143개의 사람들에 대응하는 관심 영역들(예를 들어, 바운딩 박스(bounding box)) 및 8,432개의 신원들(identities; IDs)이 제공될 수 있다. 데이터 세트 중 11,206개의 이미지들 및 5,532개의 신원들이 기계 학습 모델의 트레이닝에 사용될 수 있고, 6,978장의 이미지들 및 2,900명의 쿼리 사람들이 기계 학습 모델의 평가(evaluation)에 사용될 수 있다.

[0124]

데이터 세트(620)는, PRW 데이터 세트로서, 6개의 각도들로 촬영된 이미지들을 포함할 수 있다. 총 11,816장의 이미지들이 포함될 수 있고, 43,110개의 사람들에 대응하는 관심 영역들(예를 들어, 바운딩 박스(bounding box)) 및 932개의 신원들이 제공될 수 있다. 데이터 세트 중 5,704장의 이미지들 및 482개의 신원들이 기계 학습 모델의 트레이닝에 사용될 수 있고, 6,112장의 이미지들 및 2,057명의 쿼리 사람들이 기계 학습 모델의 평가에 사용될 수 있다.

[0125]

데이터 세트(620)는, 데이터 세트(610)와 달리, 레귤러(regular) 갤러리(gallery) 및 멀티뷰(Multiview) 갤러리를 포함하는 평가를 위한 데이터 세트를 설계할 수 있다. 레귤러 갤러리는 쿼리 사람과 같은 카메라 각도에서 촬영된 이미지들을 포함할 수 있고, 이미지 내의 라벨링되지 않은 사람에 대응하는 관심 영역들(예를 들어, 참값 신원 클래스가 존재하지 않는 관심 영역들)은 평가에서 제외될 수 있다. 멀티뷰 갤러리는 쿼리 사람과 같은 카메라 각도에서 촬영된 이미지들이 제외될 수 있고, 라벨링되지 않은 사람에 대응하는 관심 영역들은 포함되어 평가에 사용될 수 있다. 기계 학습 모델의 평가에서, 멀티뷰 갤러리는 레귤러 갤러리보다 더 도전적인(challenging) 세팅(setting)일 수 있다.

표 1

[0126]

	학습 방법	CUHK-SYSU		PRW	
		mAP	Rank-1	mAP	Rank-1
비교 실시예 1	Two-step	-	-	20.5	48.3
비교 실시예 2	Two-step	83.0	83.7	32.6	72.1
비교 실시예 3	Two-step	87.2	88.5	38.7	65.0
비교 실시예 4	Two-step	93.0	94.2	42.9	70.2
비교 실시예 5	Two-step	93.9	95.1	46.8	87.5
비교 실시예 6	Two-step	90.3	91.4	47.2	87.0
비교 실시예 7	End-to-end	75.5	78.7	21.3	49.9
비교 실시예 8	End-to-end	76.3	80.1	23.0	61.9
비교 실시예 9	End-to-end	77.9	81.2	24.2	53.1
비교 실시예 10	End-to-end	79.3	81.3	-	-
비교 실시예 11	End-to-end	84.1	86.5	33.4	73.6
비교 실시예 12	End-to-end	90.0	90.7	45.3	81.7
비교 실시예 13	End-to-end	90.5	88.4	48.5	87.9
비교 실시예 14	End-to-end	88.7	89.6	36.0	76.1
비교 실시예 15	End-to-end	88.9	89.3	41.2	81.4
비교 실시예 16	End-to-end	92.1	92.9	44.0	81.1
비교 실시예 17	End-to-end	89.7	90.8	39.8	80.4
일 실시예	End-to-end	93.3	94.2	53.3	87.7
비교 실시예 16*	End-to-end	-	-	40.0	67.5
비교 실시예 17*	End-to-end	-	-	36.5	65.0

일 실시예*	End-to-end	-	-	48.0	73.2
--------	------------	---	---	------	------

- [0127] 표 1은 비교 실시예들 및 일 실시예에 따른 사람 검색을 위한 기계 학습 모델의 데이터 세트들의 성능 비교를 나타낸다.
- [0128] 비교 실시예 1 내지 6에 따른 모델은 사람 검출 및 사람 식별을 개별적으로 학습하는 투 스텝(two-step) 기반의 모델로서, 각각 다음과 같은 논문에 개시된 모델을 나타낼 수 있다. 비교 실시예 1에 따른 모델은 논문 'Tong Xiao, Shuang Li, Bochao Wang, Liang Lin, and Xiaogang Wang. Joint detection and identification feature learning for person search. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017'에 개시된 모델을 나타낼 수 있다. 비교 실시예 2에 따른 모델은 논문 'Di Chen, Shanshan Zhang, Jian Yang, and Ying Tai. Person search via a mask-guided two-stream cnn model. In Proceedings of the European Conference on Computer Vision, 2018'에 개시된 모델을 나타낼 수 있다. 비교 실시예 3에 따른 모델은 논문 'Xu Lan, Xiatian Zhu, and Shaogang Gong. Person search by multi-scale matching. In Proceedings of the European Conference on Computer Vision, 2018'에 개시된 모델을 나타낼 수 있다. 비교 실시예 4에 따른 모델은 논문 'Chuchu Han, Jiacheng Ye, Yunshan Zhong, Xin Tan, Chi Zhang, Changxin Gao, and Nong Sang. Re-id driven localization refinement for person search. In Proceedings of the IEEE Conference on International Conference on Computer Vision, 2019'에 개시된 모델을 나타낼 수 있다. 비교 실시예 5에 따른 모델은 논문 'Cheng Wang, Bingpeng Ma, Hong Chang, Shiguang Shan, and Xilin Chen. Tcts: A task-consistent two-stage framework for person search. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020'에 개시된 모델을 나타낼 수 있다. 비교 실시예 6에 따른 모델은 논문 'Wenkai Dong, Zhaoxiang Zhang, Chunfeng Song, and Tieniu Tan. Instance guided proposal network for person search. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020'에 개시된 모델을 나타낼 수 있다.
- [0129] 비교 실시예 7 내지 17에 따른 모델은, 사람 검출 및 사람 식별을 한 번에 학습하는 엔드 투 엔드(end-to-end) 기반의 모델로서, 각각 다음과 같은 논문에 개시된 모델을 나타낼 수 있다. 비교 실시예 7에 따른 모델은 논문 'Tong Xiao, Shuang Li, Bochao Wang, Liang Lin, and Xiaogang Wang. Joint detection and identification feature learning for person search. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017'에 개시된 모델을 나타낼 수 있다. 비교 실시예 8에 따른 모델은 논문 'Jimin Xiao, Yanchun Xie, Tammam Tillo, Kaizhu Huang, Yunchao Wei, and Jiashi Feng. Ian: the individual aggregation network for person search. Pattern Recognition, 87:332- 340, 2019'에 개시된 모델을 나타낼 수 있다. 비교 실시예 9에 따른 모델은 논문 'Hao Liu, Jiashi Feng, Zequn Jie, Karlekar Jayashree, Bo Zhao, Meibin Qi, Jianguo Jiang, and Shuicheng Yan. Neural person search machines. In Proceedings of the IEEE Conference on International Conference on Computer Vision, 2017'에 개시된 모델을 나타낼 수 있다. 비교 실시예 10에 따른 모델은 논문 'Xiaojun Chang, Po-Yao Huang, Yi-Dong Shen, Xiaodan Liang, Yi Yang, and Alexander G. Hauptmann. Rcaa: Relational context-aware agents for person search. In Proceedings of the European Conference on Computer Vision, 2018'에 개시된 모델을 나타낼 수 있다. 비교 실시예 11에 따른 모델은 논문 'Yichao Yan, Qiang Zhang, Bingbing Ni, Wendong Zhang, Minghao Xu, and Xiaokang Yang. Learning context graph for person search. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019'에 개시된 모델을 나타낼 수 있다. 비교 실시예 12에 따른 모델은 논문 'Wenkai Dong, Zhaoxiang Zhang, Chunfeng Song, and Tieniu Tan. Bi-directional interaction network for person search. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020'에 개시된 모델을 나타낼 수 있다. 비교 실시예 13에 따른 모델은 논문 'Kun Tian, Houjing Huang, Yun Ye, Shiyu Li, Jinbin Lin, and Guan Huang. End-to-end thorough body perception for person search. In Proceedings of the AAAI Conference on Artificial Intelligence, 2020'에 개시된 모델을 나타낼 수 있다. 비교 실시예 14에 따른 모델은 논문 'Ju Dai, Pingping Zhang, Huchuan Lu, and Hongyu Wang. Dynamic imposter based online instance matching for person search. Pattern Recognition, 100:107120, 2020'에 개시된 모델을 나타낼 수 있다. 비교 실시예 15에 따른 모델은 논문 'Yingji Zhong, Xiaoyu Wang, and Shiliang Zhang. Robust partial matching for person search in the wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020'에 개시된 모델을 나타낼 수 있다. 비교 실시예 16에 따른 모델은 논문 'Di Chen, Shanshan Zhang, Jian Yang, and Bernt Schiele. Norm-

aware embedding for efficient person search. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2020'에 개시된 모델을 나타낼 수 있다. 비교 실시예 17에 따른 모델은 논문 'Di Chen, Shanshan Zhang, Wanli Ouyang, Jian Yang, and Bernt Schiele. Hierarchical online instance matching for person search. In Proceedings of the AAAI Conference on Artificial Intelligence, 2020'에 개시된 모델을 나타낼 수 있다.

[0130] 표 1에서 나타난 바와 같이, 비교 실시예들 및 일 실시예에 따른 기계 학습 모델의 평가 성능이 데이터 세트들 (610, 620)로 측정되어 정량적으로 비교될 수 있다. 투 스텝 기법 및 엔드 투 엔드 기법으로 분류된 비교 실시예들 및 일 실시예의 성능이 비교될 수 있다.

[0131] 일 실시예에 따른 기계 학습 모델은 CUHK-SYSU 데이터 세트(610)에 대하여 비교 실시예들에 따른 엔드 투 엔드 기법의 기계 학습 모델들 중 최고 성능을 가질 수 있다. 일 실시예에 따른 기계 학습 모델은 또한, 투 스텝 기법의 기계 학습 모델들 중 최고 성능과도 비슷한 성능을 가질 수 있다. 투 스텝 기반의 기계 학습 모델은, 사람 검출 네트워크 및 사람 식별 네트워크를 개별적으로 학습하는 불편함을 가질 수 있다. 따라서, 일 실시예에 따른 기계 학습 모델은, 엔드 투 엔드 기법에 기반한 모델로서, 효율성 및 성능에 있어서 사람 검색을 위한 기계 학습 모델의 트레이닝에 대하여 우수한 성과를 가질 수 있다.

[0132] 또한, 일 실시예에 따른 기계 학습 모델은 PRW 데이터 세트(620)에 대하여 일 실시예 및 비교 실시예들에 따른 기계 학습 모델들 중 최고 성능을 가질 수 있다. 표 1에서, PRW 데이터 세트(620)의 레귤러 갤러리 및 멀티뷰 갤러리(표 1에서 *로 표시됨) 각각에 대한 성능이 비교될 수 있다. 레귤러 갤러리보다 더 도전적인 멀티뷰 갤러리에 대한 평가 성능은, 레귤러 갤러리에 대한 평가 성능보다 낮은 성능을 가질 수 있다. 표 1에서 나타난 바와 같이, 일 실시예에 따른 기계 학습 모델은 멀티뷰 갤러리에 대한 평가에서도 비교 실시예에 따른 기계 학습 모델들보다 우수한 성능을 가질 수 있다. CUHK와 마찬가지로 본 연구팀이 제안한 end-to-end 기법은 PRW 데이터 셋에서 최고 성능을 기록하며 사람 검색 기술에서 우수한 성과를 도출한 것으로 판단될 수 있다.

[0133] [표 3]

Detection				
Method	AP	Recall	Δ AP	Δ Recall
Proposed	94.54	97.49		
w/o Part cls.	93.22	97.02	-1.32	-0.47
w/o AGWF	92.30	95.20	-2.24	-2.29
w/o AGWF & Part cls.	92.67	96.36	-1.87	-1.13
Re-identification				
Method	mAP	Top-1	Δ mAP	Δ Top-1
Proposed	48.04	73.21		
w/o Part cls.	46.52	72.19	-1.52	-1.02
w/o AGWF	46.87	72.92	-1.17	-0.29
w/o AGWF & Part cls.	45.82	70.24	-2.22	-2.97

[0134]

[0135] 표 3은 부분 분류 태스크 및 적응형 그래디언트 가중치 함수가 사람 검출 및 사람 식별에 미치는 영향을 나타낸다.

[0136] 사람 검색의 성능은 mAP(mean Average Precision) 및 Top-k 스코어로 평가될 수 있다. 여기서 mAP는 매칭 결과들에 대한 정밀도-재현율 곡선 아래 영역(area under the precision-recall curve)을 측정할 수 있고, Top-k 스코어는 쿼리 사람과 상위 유사도를 갖는 적어도 k개의 관심 영역들이 신원 매칭에 성공한 비율을 나타낼 수 있다. 참값 바운딩 박스에 대한 IoU(Intersection over Union)가 0.5 미만인 관심 영역은, Top-k 스코어에 계산되지 않을 수 있다.

[0137] 도 7은 비교 실시예 및 일 실시예에 따른 기계 학습 모델의 성능을 정성적으로 비교하는 것을 나타낸다.

- [0138] 비교 실시예에 따른 기계 학습 모델은 백본 네트워크, 사람 검출 네트워크, 및 사람 식별 네트워크를 포함하지 만 부분 분류 네트워크가 구현되지 않은 기계 학습 모델을 나타낼 수 있다. 일 실시예에 따른 기계 학습 모델 은, 백본 네트워크, 사람 검출 네트워크, 및 사람 식별 네트워크와 함께 부분 분류 네트워크를 포함하는 기계 학습 모델을 나타낼 수 있다.
- [0139] 표 3에서, 사람 검출에 대한 성능은 mAP 및 재현율(recall) 스코어로 측정될 수 있고, 사람 식별에 대한 성능은 mAP와 Top-1 스코어로 측정될 수 있다. 일 실시예에 따른 기계 학습 모델의 사람 검출에 대한 성능은 mAP 및 재현율 스코어는 각각 1.32% 및 0.47% 만큼 비교 실시예에 따른 기계 학습 모델보다 향상될 수 있다. 일 실시 예에 따른 기계 학습 모델의 사람 식별에 대한 성능은, mAP 및 Top-1 스코어는 각각 1.52% 및 1.02% 만큼 비교 실시예에 따른 기계 학습 모델보다 향상될 수 있다.
- [0140] 도 7의 행(column)은 쿼리 사람, 비교 실시예에 따른 기계 학습 모델 및 일 실시예에 따른 기계 학습 모델에 의 하여 추정된 최고 유사도를 갖는 관심 영역을 나타낸다. 도 7에서, 초록색 외곽선을 갖는 관심 영역은 쿼리 사 람에 대응하는 관심 영역을 검색, 다시 말해, 검출 및 식별한 경우를 나타낼 수 있다. 빨간색 외곽선을 갖는 관심 영역은, 쿼리 사람에 대응하는 관심 영역을 검색하지 못한 경우를 나타낼 수 있다.
- [0141] 일 실시예에 따른 기계 학습 모델은 부분 분류 네트워크에 의하여, 공유된 특징 맵을 보다 더 설명적으로 (descriptive) 트레이닝될 수 있다. 도 7에서, 일 실시예에 따른 기계 학습 모델은, 옷의 패턴 및 액세서리(예 를 들어, 가방 및/또는 우산)와 같은 뚜렷한(distinct) 부분적 특징들을 갖는 쿼리 사람에 대한 관심 영역을 정 확하게 찾을 수 있다. 이와 달리, 부분 분류 네트워크가 포함되지 않은 비교 실시예에 따른 기계 학습 모델은, 쿼리 사람과 옷 색상이 유사한 잘못된 사람에 대한 관심 영역을 반환될 수 있다.
- [0142] 표 3에서, 적응형 그래디언트 가중치 함수(AGWF)의 효과가 나타날 수 있다. 적응형 그래디언트 가중치 함수는 사람 검출 및 사람 식별 모두에 대하여 효과적일 수 있다.
- [0143] 적응형 그래디언트 가중치 함수가 비활성화(deactivate)되는 경우(예를 들어, 일반 가중치 함수가 적용되는 경 우), 적응형 그래디언트 가중치 함수가 활성화된 경우보다, 사람 검출의 성능은 mAP 및 재현율 스코어에서 각각 2.24% 및 2.29% 감소될 수 있고, 사람 식별의 성능이 mAP 및 Top-1 스코어에서 각각 1.17% 및 0.29% 감소될 수 있다.
- [0144] 또한, 부분 분류 네트워크가 포함되지 않은 기계 학습 모델은, 적응형 그래디언트 가중치 함수가 비활성화되는 경우, 부분 분류 네트워크가 포함된 기계 학습 모델보다 더 저하되므로, 사람 식별에 대한 낮은 성능을 가질 수 있다. 다만, 엔드 투 엔드 기법의 기계 학습 모델에서 태스크의 개수가 증가하면, 태스크 간의 충돌 위험이 발 생할 수 있다. 따라서, 적응형 그래디언트 가중치 함수가 비활성화된 부분 분류 네트워크를 포함하는 기계 학 습 모델은, 적응형 그래디언트 가중치 함수가 비활성화된 부분 분류 네트워크를 포함하지 않는 기계 학습 모델 보다 사람 검출에 대한 낮은 성능을 가질 수 있다. 적응형 그래디언트 가중치 함수는 일 실시예에 따른 기계 학습 모델뿐만 아니라, 종래 사람 검출 및 사람 식별의 2개의 태스크들을 위한 네트워크에 적용될 수 있다.

[0145] [표 4]

Method	mAP	Top-1	Δ mAP	Δ Top-1
Proposed	48.04	73.21		
w/o $\mathcal{L}_{bin}$ & $\mathcal{L}_{un}$	47.44	71.85	-0.60	-1.36

- [0146]
- [0147] 표 4는 참값 신원 클래스를 갖는 관심 영역에 대한 바이너리 분류 손실 값($\mathcal{L}_{bin}$) 및 참값 신원 클래스가 존재하 지 않는 관심 영역에 대한 식별 손실 값($\mathcal{L}_{un}$)이 사람 검색을 위한 기계 학습 모델의 성능에 미치는 영향을 나 타낸다.
- [0148] 참값 신원 클래스를 갖는 관심 영역에 대한 바이너리 분류 손실 값($\mathcal{L}_{bin}$)은 타겟 특징 벡터 및 참값 신원 클래 스와 다른 신원 클래스들(예를 들어, 네거티브 클래스들(negative classes))의 대표 특징 벡터와의 유사도에 관 계없이 참값 신원 클래스(예를 들어, 포지티브 클래스(positive class))의 대표 특징 벡터와의 유사도를 높이도 록 설계될 수 있다. 이와 달리, 참값 신원 클래스가 존재하지 않는 관심 영역에 대한 식별 손실 값($\mathcal{L}_{un}$)은 존 재하지 않는 참값 신원 클래스로 인해, 다른 신원 클래스들(예를 들어, 네거티브 클래스)의 대표 특징 벡터와의 유사도를 줄이도록 설계될 수 있다. 그러므로, 기계 학습 모델의 성능 평가에 있어서, 참값 신원 클래스를 갖

는 관심 영역에 대한 바이너리 분류 손실 값($\mathcal{L}_{bin}$) 및 참값 신원 클래스가 존재하지 않는 관심 영역에 대한 식별 손실 값($\mathcal{L}_{un}$)은 모두, 원하지 않는(undesired) 바이어스(bias)를 피하기 위하여 보완적인 값으로 함께 사용되는 것이 보다 더 타당(reasonable)할 수 있다. 표 4에서 나타난 바와 같이, 바이너리 분류 손실 값($\mathcal{L}_{bin}$) 및 식별 손실 값($\mathcal{L}_{un}$)을 사용하지 않고 트레이닝된 기계 학습 모델은, 일 실시예에 따른 기계 학습 모델보다 mAP 및 Top-1 스코어에서 각각 0.6% 및 1.36% 낮은 성능을 가질 수 있다.

[0149] 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.

[0150] 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(embodiment)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.

[0151] 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.

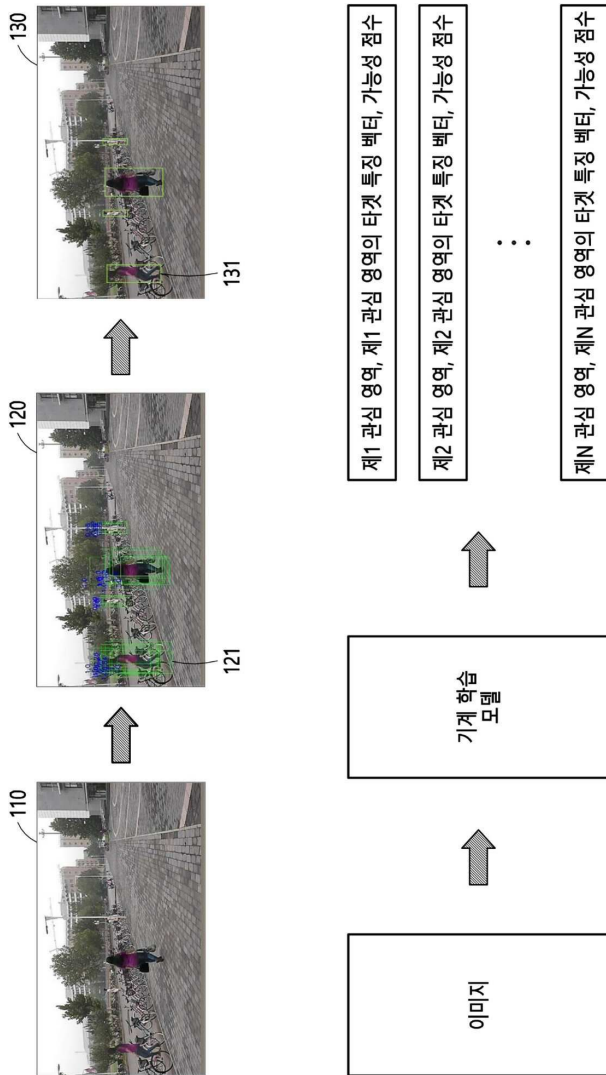
[0152] 위에서 설명한 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 또는 복수의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0153] 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.

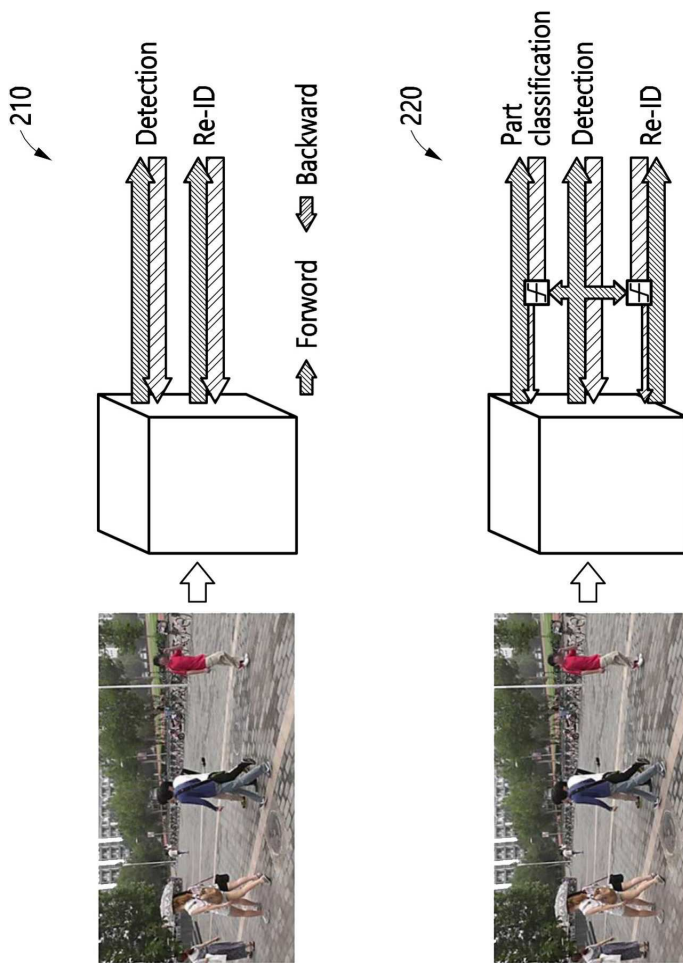
[0154] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

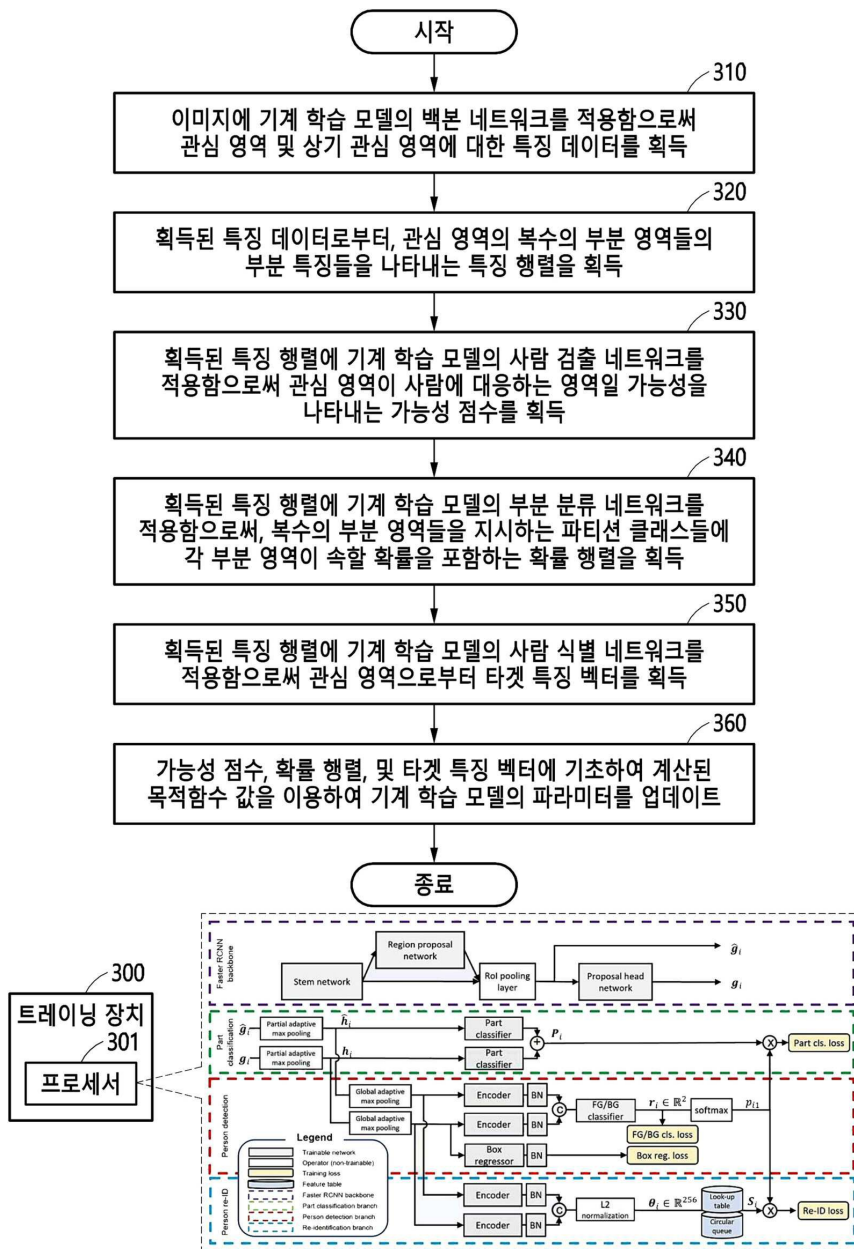
도면1



도면2



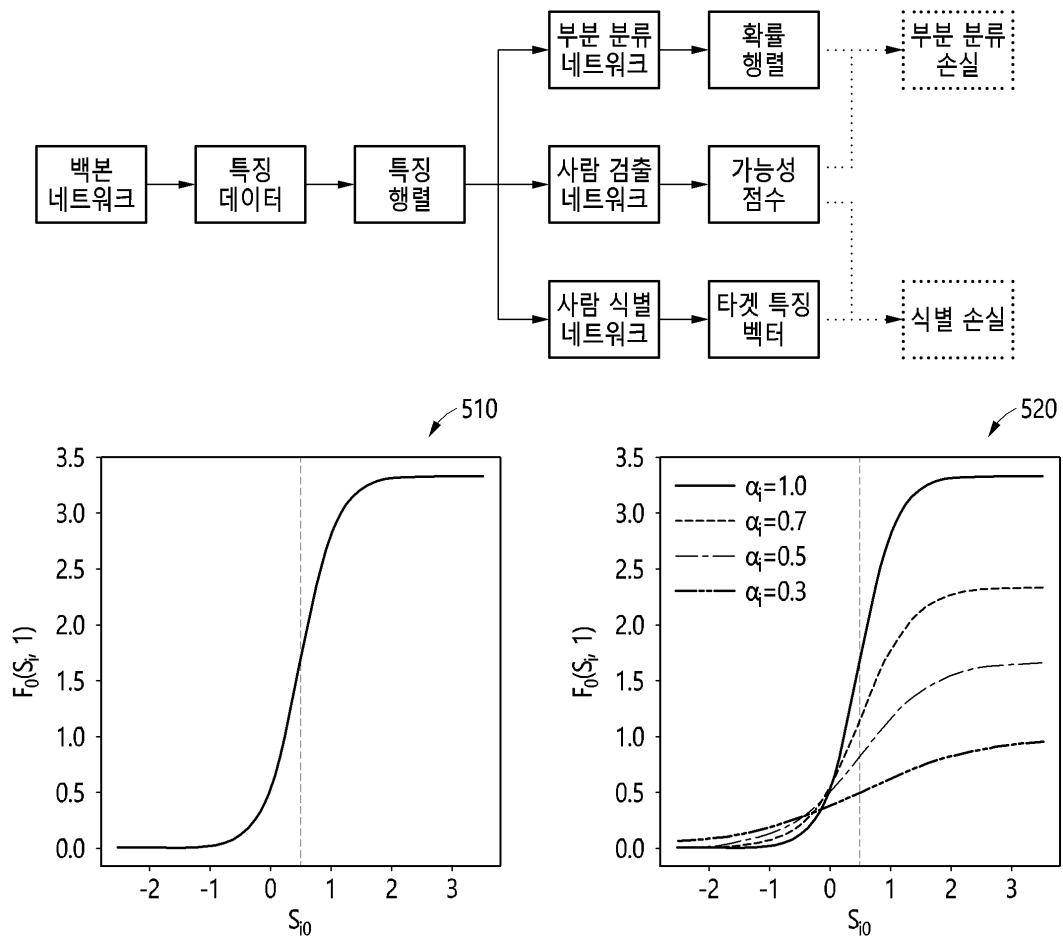
도면3



도면4



도면5



도면6

610



620



도면7

710 ↗

