



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0077439
(43) 공개일자 2022년06월09일

- (51) 국제특허분류(Int. Cl.)
G06N 3/08 (2006.01) G06F 16/23 (2019.01)
G06F 16/28 (2019.01) G06N 3/04 (2006.01)
- (52) CPC특허분류
G06N 3/08 (2013.01)
G06F 16/23 (2019.01)
- (21) 출원번호 10-2020-0166373
(22) 출원일자 2020년12월02일
심사청구일자 2020년12월02일
- (71) 출원인
울산과학기술원
울산광역시 울주군 언양읍 유니스트길 50
- (72) 발명자
심재영
울산광역시 울주군 언양읍 유니스트길 50
고규현
울산광역시 울주군 언양읍 유니스트길 50
- (74) 대리인
특허법인 무한

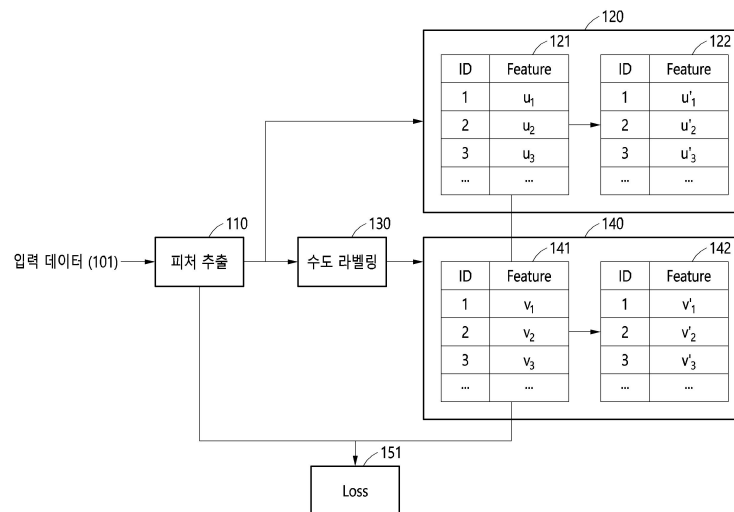
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 객체 검색 모델 및 그 학습 방법

(57) 요약

실시예는 객체 검색 모델 및 그 학습 방법에 관한 것이다. 일 실시예에 따른 객체 검색 모델의 학습 방법은 객체 검색 모델을 이용하여, 입력 데이터에 포함된 객체 영역의 피처를 추출하는 단계, 객체 영역에 대응하는 라벨이 존재하는 경우, 라벨 및 추출된 피처에 기초하여 제1 룩업 테이블을 갱신하는 단계, 객체 영역에 대응하는 라벨이 존재하지 않는 경우, 제2 룩업 테이블에 기초하여, 객체 영역에 수도 라벨을 부여하여, 수도 라벨 및 추출된 피처에 기초하여 제2 룩업 테이블을 갱신하는 단계, 및 추출된 피처, 제1 룩업 테이블 및 제2 룩업 테이블에 기초하여, 객체 검색 모델을 학습시키는 단계를 포함한다.

대표도



(52) CPC특허분류

G06F 16/289 (2019.01)

G06N 3/04 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711103020
과제번호	2017-0-00667-004
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	70/100
과제수행기관명	울산과학기술원
연구기간	2020.01.01 ~ 2020.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711116816
과제번호	2020-0-01336-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성(R&D)
연구과제명	인공지능대학원지원(울산과학기술원)
기 여 율	30/100
과제수행기관명	울산과학기술원
연구기간	2020.04.01 ~ 2024.12.31

명세서

청구범위

청구항 1

객체 검색(object search) 모델을 이용하여, 입력 데이터에 포함된 객체 영역의 피처(feature)를 추출하는 단계;

상기 객체 영역에 대응하는 라벨이 존재하는 경우,

상기 라벨 및 상기 추출된 피처에 기초하여 제1 룩업 테이블을 갱신하는 단계;

상기 객체 영역에 대응하는 라벨이 존재하지 않는 경우,

제2 룩업 테이블에 기초하여, 상기 객체 영역에 수도 라벨을 부여하는 단계; 및

상기 수도 라벨 및 상기 추출된 피처에 기초하여 상기 제2 룩업 테이블을 갱신하는 단계; 및

상기 추출된 피처, 상기 제1 룩업 테이블 및 상기 제2 룩업 테이블에 기초하여, 상기 객체 검색 모델을 학습시키는 단계

를 포함하는,

학습 방법.

청구항 2

제1항에 있어서,

상기 객체 영역에 수도 라벨을 부여하는 단계는

상기 추출된 피처와 상기 제2 룩업 테이블에 저장된 피처 사이의 유사도 및 미리 정해진 임계 값에 기초하여, 상기 객체 영역에 부여될 수도 라벨을 결정하는 단계

를 더 포함하는,

학습 방법.

청구항 3

제2항에 있어서,

상기 객체 영역에 부여될 수도 라벨을 결정하는 단계는

상기 제2 룩업 테이블에 포함된 수도 라벨들 각각에 대응하여, 상기 추출된 피처와 해당하는 수도 라벨에 대하여 저장된 피처의 유사도를 계산하는 단계; 및

계산된 유사도들 중 상기 임계 값보다 큰 유사도가 없는 경우, 상기 제2 룩업 테이블에 포함되지 않은 새로운 수도 라벨을 상기 객체 영역에 부여될 수도 라벨로 결정하는 단계

를 포함하는,

학습 방법.

청구항 4

제3항에 있어서,

상기 계산된 유사도들 중 상기 임계 값보다 큰 유사도가 있는 경우, 상기 계산된 유사도들 중 가장 높은 유사도에 대응하는 수도 라벨을 상기 객체 영역에 부여될 수도 라벨로 결정하는 단계

를 더 포함하는,

학습 방법.

청구항 5

제2항에 있어서,

상기 미리 정해진 임계 값은

학습 데이터 내 라벨이 존재하는 객체 영역의 피처와 상기 제1 특업 테이블에 해당 라벨에 대응하여 저장된 피처 사이의 유사도에 기초하여 결정된,

학습 방법.

청구항 6

제1항에 있어서,

상기 객체 검색 모델을 학습시키는 단계는

상기 추출된 피처와 상기 제1 특업 테이블에 저장된 피처 사이의 유사도 및 상기 추출된 피처와 상기 제2 특업 테이블에 저장된 피처 사이의 유사도에 기초하여, 손실 함수를 결정하는 단계; 및

상기 손실 함수를 최소화하는 방향으로 상기 객체 검색 모델의 네트워크 파라미터를 학습시키는 단계

를 포함하는,

학습 방법.

청구항 7

제6항에 있어서,

상기 손실 함수는

상기 객체 영역에 대응하는 라벨이 존재하는 경우, 상기 추출된 피처와 상기 제1 특업 테이블에 상기 라벨에 대응하여 저장된 피처 사이의 유사도가 클수록 작게 결정되고,

상기 객체 영역에 대응하는 수도 라벨이 존재하는 경우, 상기 추출된 피처와 상기 제2 특업 테이블에 상기 수도 라벨에 대응하여 저장된 피처와의 유사도가 클수록 작게 결정되는,

학습 방법.

청구항 8

제1항에 있어서,

상기 제1 특업 테이블을 갱신하는 단계는

상기 제1 특업 테이블에 상기 라벨에 대응하는 피처가 저장되어 있는 경우, 상기 추출된 피처를 반영하여 상기 제1 특업 테이블에 상기 라벨에 대응하여 저장된 피처를 갱신하는 단계; 및

상기 제1 특업 테이블에 상기 라벨에 대응하는 피처가 저장되어 있지 않은 경우, 상기 제1 특업 테이블에 상기 추출된 피처를 상기 라벨에 대응하여 저장하는 단계

를 포함하는,
학습 방법.

청구항 9

제1항에 있어서,

상기 제2 룩업 테이블을 갱신하는 단계는

상기 제2 룩업 테이블에 상기 수도 라벨에 대응하는 피처가 저장되어 있는 경우, 상기 추출된 피처를 반영하여 상기 제2 룩업 테이블에 상기 수도 라벨에 대응하여 저장된 피처를 갱신하는 단계; 및

상기 제2 룩업 테이블에 상기 수도 라벨에 대응하는 피처가 저장되어 있지 않은 경우, 상기 제2 룩업 테이블에 상기 추출된 피처를 상기 수도 라벨에 대응하여 저장하는 단계

를 포함하는,

학습 방법.

청구항 10

검색 대상을 포함하는 쿼리 이미지 및 검색용 이미지를 수신하는 단계;

상기 쿼리 이미지에서, 인코더를 이용하여 상기 검색 대상의 제1 피처를 획득하는 단계;

상기 검색용 이미지에서, 상기 인코더를 이용하여 상기 검색용 이미지에 포함된 적어도 하나의 객체 영역의 제2 피처를 획득하는 단계; 및

상기 제1 피처 및 상기 적어도 하나의 제2 피처에 기초하여, 상기 검색용 이미지에서 상기 검색 대상에 해당하는 영역을 획득하는 단계

를 포함하고,

상기 인코더는 라벨링되지 않은 학습 데이터를 수도 라벨링하여, 다른 라벨에 대응하는 데이터와 구분되는 피처를 추출하도록 학습된 뉴럴 네트워크로 구성된,

객체 검색 방법.

청구항 11

제10항에 있어서,

상기 뉴럴 네트워크는 라벨링된 학습 데이터의 피처를 해당 라벨에 대응하여 저장하는 제1 룩업 테이블 및 라벨링되지 않은 학습 데이터를 수도 라벨링하여, 상기 라벨링되지 않은 학습 데이터의 피처를 해당 수도 라벨에 대응하여 저장하는 제2 룩업 테이블에 기초하여 학습된,

객체 검색 방법.

청구항 12

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

청구항 13

라벨링되지 않은 학습 데이터를 수도 라벨링하여, 다른 라벨에 대응하는 데이터와 구분되는 피처를 추출하도록 학습된 뉴럴 네트워크로 구성된 인코더; 및

검색 대상을 포함하는 쿼리 이미지 및 검색용 이미지를 수신하고, 상기 쿼리 이미지에서, 상기 인코더를 이용하여 상기 검색 대상의 제1 피처를 획득하고, 상기 검색용 이미지에서, 상기 인코더를 이용하여 상기 검색용 이미지에 포함된 적어도 하나의 객체 영역의 제2 피처를 획득하며, 상기 제1 피처 및 상기 적어도 하나의 제2 피처에 기초하여, 상기 검색용 이미지에서 상기 검색 대상에 해당하는 영역을 획득하는, 적어도 하나의 프로세서

를 포함하는,

객체 검색 장치.

청구항 14

제13항에 있어서,

상기 뉴럴 네트워크는,

라벨링된 학습 데이터의 피처를 해당 라벨에 대응하여 저장하는 제1 룩업 테이블 및 라벨링되지 않은 학습 데이터를 수도 라벨링하여, 상기 라벨링되지 않은 학습 데이터의 피처를 해당 수도 라벨에 대응하여 저장하는 제2 룩업 테이블에 기초하여 학습된,

객체 검색 장치.

청구항 15

입력 데이터에서 객체 영역을 검출하고, 상기 객체 영역에서 피처를 추출하는 뉴럴 네트워크;

라벨을 키(key)로 하여, 피처를 값(value)으로 저장하는 제1 룩업 테이블 및 수도 라벨을 키로 하여, 피처를 값으로 저장하는 제2 룩업 테이블을 기록하는 메모리; 및

상기 제1 룩업 테이블 및 상기 제2 룩업 테이블에 기초하여, 상기 뉴럴 네트워크를 학습시키는 적어도 하나의 프로세서

를 포함하는,

객체 검색을 위한 학습 모델.

청구항 16

제15항에 있어서,

상기 프로세서는,

상기 뉴럴 네트워크를 학습시킴에 있어서,

상기 입력 데이터를 상기 뉴럴 네트워크에 인가하여, 상기 입력 데이터에 포함된 객체 영역의 피처를 획득하고,

상기 객체 영역에 대응하는 라벨이 존재하는 경우, 상기 라벨 및 상기 추출된 피처에 기초하여 상기 제1 룩업 테이블을 갱신하고,

상기 객체 영역에 대응하는 라벨이 존재하지 않는 경우, 상기 제2 룩업 테이블에 기초하여, 상기 객체 영역에 수도 라벨을 부여하고, 상기 수도 라벨 및 상기 추출된 피처에 기초하여 상기 제2 룩업 테이블을 갱신하며,

상기 추출된 피처, 상기 제1 룩업 테이블 및 상기 제2 룩업 테이블에 기초하여, 상기 뉴럴 네트워크를 학습시키는,

객체 검색을 위한 학습 모델.

청구항 17

제16항에 있어서,

상기 프로세서는,

상기 객체 영역에 수도 라벨을 부여함에 있어서,

상기 제2 룩업 테이블에 포함된 수도 라벨들 각각에 대응하여, 상기 추출된 피처와 해당하는 수도 라벨에 대응하여 저장된 피처의 유사도를 계산하고,

계산된 유사도들 중 상기 임계 값보다 큰 유사도가 없는 경우, 상기 제2 룩업 테이블에 포함되지 않은 새로운 수도 라벨을 상기 객체 영역에 부여될 수도 라벨로 결정하고,

상기 계산된 유사도들 중 상기 임계 값보다 큰 유사도가 있는 경우, 상기 계산된 유사도들 중 가장 높은 유사도에 대응하는 수도 라벨을 상기 객체 영역에 부여될 수도 라벨로 결정하는,

객체 검색을 위한 학습 모델.

청구항 18

제16항에 있어서,

상기 프로세서는,

상기 객체 검색 모델을 학습시킴에 있어서,

상기 추출된 피처와 상기 제1 룩업 테이블에 저장된 피처 사이의 유사도 및 상기 추출된 피처와 상기 제2 룩업 테이블에 저장된 피처 사이의 유사도에 기초하여, 손실 함수를 결정하고,

상기 손실 함수를 최소화하는 방향으로 상기 객체 검색 모델의 네트워크 파라미터를 학습시키는,

객체 검색을 위한 학습 모델.

청구항 19

제16항에 있어서,

상기 프로세서는,

상기 제1 룩업 테이블을 갱신함에 있어서,

상기 제1 룩업 테이블에 상기 라벨에 대응하는 피처가 저장되어 있는 경우, 상기 추출된 피처를 반영하여 상기 제1 룩업 테이블에 상기 라벨에 대응하여 저장된 피처를 갱신하고,

상기 제1 룩업 테이블에 상기 라벨에 대응하는 피처가 저장되어 있지 않은 경우, 상기 제1 룩업 테이블에 상기 추출된 피처를 상기 라벨에 대응하여 저장하는,

객체 검색을 위한 학습 모델.

청구항 20

제16항에 있어서,

상기 프로세서는,

상기 제2 룩업 테이블을 갱신함에 있어서,

상기 제2 룩업 테이블에 상기 수도 라벨에 대응하는 피처가 저장되어 있는 경우, 상기 추출된 피처를 반영하여 상기 제2 룩업 테이블에 상기 수도 라벨에 대응하여 저장된 피처를 갱신하고,

상기 제2 룩업 테이블에 상기 수도 라벨에 대응하는 피처가 저장되어 있지 않은 경우, 상기 제2 룩업 테이블에 상기 추출된 피처를 상기 수도 라벨에 대응하여 저장하는,

객체 검색을 위한 학습 모델.

발명의 설명

기술 분야

[0001] 객체 검색 모델 및 그 학습 방법에 관한 것이다.

배경 기술

[0002] 사람 검색(person search) 기술은 영상에서 원하는 사람을 찾아내는 기술로, 복수의 사람이 임의의 자세와 임의의 각도에서 등장하는 CCTV, 블랙박스 등으로부터 수집된 다양한 영상에서 자동으로 타겟을 검출할 수 있어 보안 및 감시 시스템에 광범위하게 이용될 수 있다.

[0003] 딥러닝 기반의 사람 검색 기술을 개발하기 위해서, 영상에 나타난 사람의 위치를 가리키는 바운딩 박스(bounding box) 라벨 및 사람의 식별 정보를 의미하는 ID 라벨을 포함한 영상 데이터가 필요하다. 그러나 라벨을 포함한 학습 데이터는 사람에 의해 생성되는 것으로 그 양이 많지 않으며, 라벨링 작업이 완벽하게 되지 않은 학습 데이터는 사람 검색 기술의 개발을 방해하는 요인으로 지적된다.

발명의 내용

해결하려는 과제

[0004] 실시예는 유사도를 기반으로 ID 라벨이 부여되지 않은 객체에 수도(pseudo) ID 라벨을 할당함으로써, 객체 검색 모델의 학습에 ID 라벨이 없는 데이터를 효과적으로 활용하는 학습 방법을 제안한다.

[0005] 다만, 기술적 과제는 상술한 기술적 과제들로 한정되는 것은 아니며, 또 다른 기술적 과제들이 존재할 수 있다.

과제의 해결 수단

[0006] 일 측에 따른 학습 방법은 객체 검색(object search) 모델을 이용하여, 입력 데이터에 포함된 객체 영역의 피처(feature)를 추출하는 단계; 상기 객체 영역에 대응하는 라벨이 존재하는 경우, 상기 라벨 및 상기 추출된 피처에 기초하여 제1 룩업 테이블을 갱신하는 단계; 상기 객체 영역에 대응하는 라벨이 존재하지 않는 경우, 제2 룩업 테이블에 기초하여, 상기 객체 영역에 수도 라벨을 부여하는 단계; 및 상기 수도 라벨 및 상기 추출된 피처에 기초하여 상기 제2 룩업 테이블을 갱신하는 단계; 및 상기 추출된 피처, 상기 제1 룩업 테이블 및 상기 제2 룩업 테이블에 기초하여, 상기 객체 검색 모델을 학습시키는 단계를 포함한다.

[0007] 상기 객체 영역에 수도 라벨을 부여하는 단계는 상기 추출된 피처와 상기 제2 룩업 테이블에 저장된 피처 사이의 유사도 및 미리 정해진 임계 값에 기초하여, 상기 객체 영역에 부여될 수도 라벨을 결정하는 단계를 더 포함할 수 있다.

[0008] 상기 객체 영역에 부여될 수도 라벨을 결정하는 단계는 상기 제2 룩업 테이블에 포함된 수도 라벨들 각각에 대응하여, 상기 추출된 피처와 해당하는 수도 라벨에 대응하여 저장된 피처의 유사도를 계산하는 단계; 및 계산된 유사도들 중 상기 임계 값보다 큰 유사도가 없는 경우, 상기 제2 룩업 테이블에 포함되지 않은 새로운 수도 라벨을 상기 객체 영역에 부여될 수도 라벨로 결정하는 단계를 포함할 수 있다.

[0009] 상기 상기 객체 영역에 부여될 수도 라벨을 결정하는 단계는 상기 계산된 유사도들 중 상기 임계 값보다 큰 유사도가 있는 경우, 상기 계산된 유사도들 중 가장 높은 유사도에 대응하는 수도 라벨을 상기 객체 영역에 부여될 수도 라벨로 결정하는 단계를 더 포함할 수 있다.

- [0010] 상기 미리 정해진 임계 값은 학습 데이터 내 라벨이 존재하는 객체 영역의 피처와 상기 제1 룩업 테이블에 해당 라벨에 대응하여 저장된 피처 사이의 유사도에 기초하여 결정될 수 있다.
- [0011] 상기 객체 검색 모델을 학습시키는 단계는 상기 추출된 피처와 상기 제1 룩업 테이블에 저장된 피처 사이의 유사도 및 상기 추출된 피처와 상기 제2 룩업 테이블에 저장된 피처 사이의 유사도에 기초하여, 손실 함수를 결정하는 단계; 및 상기 손실 함수를 최소화하는 방향으로 상기 객체 검색 모델의 네트워크 파라미터를 학습시키는 단계를 포함할 수 있다.
- [0012] 상기 손실 함수는 상기 객체 영역에 대응하는 라벨이 존재하는 경우, 상기 추출된 피처와 상기 제1 룩업 테이블에 상기 라벨에 대응하여 저장된 피처 사이의 유사도가 클수록 작게 결정되고, 상기 객체 영역에 대응하는 수도 라벨이 존재하는 경우, 상기 추출된 피처와 상기 제2 룩업 테이블에 상기 수도 라벨에 대응하여 저장된 피처와의 유사도가 클수록 작게 결정될 수 있다.
- [0013] 상기 제1 룩업 테이블을 갱신하는 단계는 상기 제1 룩업 테이블에 상기 라벨에 대응하는 피처가 저장되어 있는 경우, 상기 추출된 피처를 반영하여 상기 제1 룩업 테이블에 상기 라벨에 대응하여 저장된 피처를 갱신하는 단계; 및 상기 제1 룩업 테이블에 상기 라벨에 대응하는 피처가 저장되어 있지 않은 경우, 상기 제1 룩업 테이블에 상기 추출된 피처를 상기 라벨에 대응하여 저장하는 단계를 포함할 수 있다.
- [0014] 상기 제2 룩업 테이블을 갱신하는 단계는 상기 제2 룩업 테이블에 상기 수도 라벨에 대응하는 피처가 저장되어 있는 경우, 상기 추출된 피처를 반영하여 상기 제2 룩업 테이블에 상기 수도 라벨에 대응하여 저장된 피처를 갱신하는 단계; 및 상기 제2 룩업 테이블에 상기 수도 라벨에 대응하는 피처가 저장되어 있지 않은 경우, 상기 제2 룩업 테이블에 상기 추출된 피처를 상기 수도 라벨에 대응하여 저장하는 단계를 포함할 수 있다.
- [0016] 일 측에 따른 객체 검색 방법은 검색 대상을 포함하는 쿼리 이미지 및 검색용 이미지를 수신하는 단계; 상기 쿼리 이미지에서, 인코더를 이용하여 상기 검색 대상의 제1 피처를 획득하는 단계; 상기 검색용 이미지에서, 상기 인코더를 이용하여 상기 검색용 이미지에 포함된 적어도 하나의 객체 영역의 제2 피처를 획득하는 단계; 및 상기 제1 피처 및 상기 적어도 하나의 제2 피처에 기초하여, 상기 검색용 이미지에서 상기 검색 대상에 해당하는 영역을 획득하는 단계를 포함한다.
- [0017] 상기 인코더는 라벨링되지 않은 학습 데이터를 수도 라벨링하여, 다른 라벨에 대응하는 데이터와 구분되는 피처를 추출하도록 학습된 뉴럴 네트워크로 구성된다.
- [0018] 상기 뉴럴 네트워크는 라벨링된 학습 데이터의 피처를 해당 라벨에 대응하여 저장하는 제1 룩업 테이블 및 라벨링되지 않은 학습 데이터를 수도 라벨링하여, 상기 라벨링되지 않은 학습 데이터의 피처를 해당 수도 라벨에 대응하여 저장하는 제2 룩업 테이블에 기초하여 학습될 수 있다.
- [0020] 일 측에 따른 객체 검색 장치는 라벨링되지 않은 학습 데이터를 수도 라벨링하여, 다른 라벨에 대응하는 데이터와 구분되는 피처를 추출하도록 학습된 뉴럴 네트워크로 구성된 인코더; 및 검색 대상을 포함하는 쿼리 이미지 및 검색용 이미지를 수신하고, 상기 쿼리 이미지에서, 상기 인코더를 이용하여 상기 검색 대상의 제1 피처를 획득하고, 상기 검색용 이미지에서, 상기 인코더를 이용하여 상기 검색용 이미지에 포함된 적어도 하나의 객체 영역의 제2 피처를 획득하며, 상기 제1 피처 및 상기 적어도 하나의 제2 피처에 기초하여, 상기 검색용 이미지에서 상기 검색 대상에 해당하는 영역을 획득하는, 적어도 하나의 프로세서를 포함한다.
- [0021] 상기 뉴럴 네트워크는, 라벨링된 학습 데이터의 피처를 해당 라벨에 대응하여 저장하는 제1 룩업 테이블 및 라벨링되지 않은 학습 데이터를 수도 라벨링하여, 상기 라벨링되지 않은 학습 데이터의 피처를 해당 수도 라벨에 대응하여 저장하는 제2 룩업 테이블에 기초하여 학습될 수 있다.
- [0023] 일 측에 따른 객체 검색을 위한 학습 모델은 입력 데이터에서 객체 영역을 검출하고, 상기 객체 영역에서 피처를 추출하는 뉴럴 네트워크; 라벨을 키(key)로 하여, 피처를 값(value)으로 저장하는 제1 룩업 테이블 및 수도 라벨을 키로 하여, 피처를 값으로 저장하는 제2 룩업 테이블을 기록하는 메모리; 및 상기 제1 룩업 테이블 및 상기 제2 룩업 테이블에 기초하여, 상기 뉴럴 네트워크를 학습시키는 적어도 하나의 프로세서를 포함한다.
- [0024] 상기 프로세서는, 상기 뉴럴 네트워크를 학습시킴에 있어서, 상기 입력 데이터를 상기 뉴럴 네트워크에 인가하

여, 상기 입력 데이터에 포함된 객체 영역의 피처를 획득하고, 상기 객체 영역에 대응하는 라벨이 존재하는 경우, 상기 라벨 및 상기 추출된 피처에 기초하여 상기 제1 룩업 테이블을 갱신하고, 상기 객체 영역에 대응하는 라벨이 존재하지 않는 경우, 상기 제2 룩업 테이블에 기초하여, 상기 객체 영역에 수도 라벨을 부여하고, 상기 수도 라벨 및 상기 추출된 피처에 기초하여 상기 제2 룩업 테이블을 갱신하며, 상기 추출된 피처, 상기 제1 룩업 테이블 및 상기 제2 룩업 테이블에 기초하여, 상기 뉴럴 네트워크를 학습시킬 수 있다.

[0025] 상기 프로세서는, 상기 객체 영역에 수도 라벨을 부여함에 있어서, 상기 제2 룩업 테이블에 포함된 수도 라벨들 각각에 대응하여, 상기 추출된 피처와 해당하는 수도 라벨에 대응하여 저장된 피처의 유사도를 계산하고, 계산된 유사도들 중 상기 임계 값보다 큰 유사도가 없는 경우, 상기 제2 룩업 테이블에 포함되지 않은 새로운 수도 라벨을 상기 객체 영역에 부여될 수도 라벨로 결정하고, 상기 계산된 유사도들 중 상기 임계 값보다 큰 유사도가 있는 경우, 상기 계산된 유사도들 중 가장 높은 유사도에 대응하는 수도 라벨을 상기 객체 영역에 부여될 수도 라벨로 결정할 수 있다.

[0026] 상기 프로세서는, 상기 객체 검색 모델을 학습시킴에 있어서, 상기 추출된 피처와 상기 제1 룩업 테이블에 저장된 피처 사이의 유사도 및 상기 추출된 피처와 상기 제2 룩업 테이블에 저장된 피처 사이의 유사도에 기초하여, 손실 함수를 결정하고, 상기 손실 함수를 최소화하는 방향으로 상기 객체 검색 모델의 네트워크 파라미터를 학습시킬 수 있다.

[0027] 상기 프로세서는, 상기 제1 룩업 테이블을 갱신함에 있어서, 상기 제1 룩업 테이블에 상기 라벨에 대응하는 피처가 저장되어 있는 경우, 상기 추출된 피처를 반영하여 상기 제1 룩업 테이블에 상기 라벨에 대응하여 저장된 피처를 갱신하고, 상기 제1 룩업 테이블에 상기 라벨에 대응하는 피처가 저장되어 있지 않은 경우, 상기 제1 룩업 테이블에 상기 추출된 피처를 상기 라벨에 대응하여 저장할 수 있다.

[0028] 상기 프로세서는, 상기 제2 룩업 테이블을 갱신함에 있어서, 상기 제2 룩업 테이블에 상기 수도 라벨에 대응하는 피처가 저장되어 있는 경우, 상기 추출된 피처를 반영하여 상기 제2 룩업 테이블에 상기 수도 라벨에 대응하여 저장된 피처를 갱신하고, 상기 제2 룩업 테이블에 상기 수도 라벨에 대응하는 피처가 저장되어 있지 않은 경우, 상기 제2 룩업 테이블에 상기 추출된 피처를 상기 수도 라벨에 대응하여 저장할 수 있다.

발명의 효과

[0029] 실시예에 따른 학습 방법으로 학습된 객체 검색 모델은 기존의 사람 검색 기술보다 성능이 향상된 사람 검색 기술을 제공할 수 있으며, 라벨링된 학습 데이터를 활용하기 어려운 다양한 상황에서 활용 가능하다.

[0030] 또한, 다른 이미지에 등장한 같은 사람에게 같은 ID 라벨을 부여할 수 있기 때문에 ID 라벨이 없는 데이터를 보다 효과적으로 활용할 수 있다.

도면의 간단한 설명

[0031] 도 1은 일실시예에 따른 객체 검색 모델의 학습 방법을 설명하기 위한 도면.

도 2는 일실시예에 따른 객체 검색 모델을 이용한 객체 검색 방법을 설명하기 위한 도면.

도 3은 일실시예에 따른 객체 검색 모델의 구조를 예시한 도면.

도 4a는 일실시예에 따른 객체 검색 모델을 이용하여 찾고자 하는 사람의 영상을 도시한 도면.

도 4b는 일실시예에 따른 객체 검색 모델을 이용하여 입력 영상에서 검색 대상을 검출한 결과를 도시한 도면.

도 5는 일실시예에 따른 장치의 구성의 예시도.

발명을 실시하기 위한 구체적인 내용

[0032] 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.

[0033] 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.

- [0034] 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.
- [0035] 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.
- [0036] 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.
- [0037] 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.
- [0039] 이하에서, $\hat{x}$ 는 $\hat{x}$ 로 표시될 수 있다.
- [0041] 도 1은 일실시예에 따른 객체 검색 모델의 학습 방법을 설명하기 위한 도면이다.
- [0042] 도 1을 참조하면, 일실시예에 따른 객체 검색(object search) 모델의 학습 방법은 객체 검색 모델을 이용하여, 입력 데이터에 포함된 객체 영역의 피처(feature)를 추출하는 단계(110), 객체 영역에 대응하는 라벨(label)이 존재하는 경우, 라벨 및 추출된 피처에 기초하여 제1 룩업 테이블(look-up table)을 갱신하는 단계(120), 객체 영역에 대응하는 라벨이 존재하지 않는 경우, 객체 영역에 수도 라벨(pseudo label)을 부여(130)하여, 수도 라벨 및 추출된 피처에 기초하여 제2 룩업 테이블을 갱신하는 단계(140) 및 추출된 피처, 제1 룩업 테이블 및 제2 룩업 테이블에 기초하여, 객체 검색 모델을 학습시키는 단계를 포함한다.
- [0043] 일실시예에 따른 학습 방법은 라벨(예를 들어, 객체의 식별 정보를 나타내는 ID 라벨)이 있는 객체 영역의 피처를 라벨 별로 저장하는 제1 룩업 테이블과 라벨이 없는 객체들의 피처를 수도 라벨 별로 저장하는 제2 룩업 테이블을 이용할 수 있다. 수도 라벨은 라벨이 존재하지 않는 객체 영역에서 추출된 특징에 따라 객체 검색 모델에서 생성된 가짜 라벨로, 라벨링(labeling)된 데이터 내 객체 영역에 부여된 진짜 라벨과 구분될 수 있다. 제1 룩업 테이블은 라벨을 키(key)로 하여, 해당 라벨로 라벨링된 객체 영역에서 추출된 피처를 값(value)으로 저장하는 룩업 테이블에 해당할 수 있다. 제2 룩업 테이블은 라벨이 없는 객체 영역에 임의로 부여된 수도 라벨을 키로 하여, 해당 수도 라벨로 라벨링된 객체 영역에서 추출된 피처를 값으로 저장하는 룩업 테이블에 해당할 수 있다.
- [0045] 일실시예에 따른 수도 라벨링(130)은 라벨이 없는 데이터가 주어졌을 때, 객체 검색 모델에서 데이터에 맞는 라벨을 자동으로 부여하는 과정이다. 일실시예에 따른 객체 검색 모델은 라벨이 없는 입력 데이터가 주어지면, 입력 데이터에 포함된 객체 영역의 피처를 추출하고, 추출된 피처를 제2 룩업 테이블에 기 저장된 수도 라벨들에 대응하는 피처들과 비교하여, 제2 룩업 테이블에 이미 저장된 수도 라벨들 중 어느 하나를 부여하거나, 제2 룩업 테이블에 저장되지 않은 새로운 수도 라벨을 부여할 수 있다.
- [0046] 보다 구체적으로, 수도 라벨을 부여하는 단계(130)는 추출된 피처와 제2 룩업 테이블에 저장된 피처 사이의 유사도 및 미리 정해진 임계 값에 기초하여, 객체 영역에 부여될 수도 라벨을 결정하는 단계를 더 포함할 수 있다. 객체 영역에 부여될 수도 라벨을 결정하는 단계는 제2 룩업 테이블에 포함된 수도 라벨들 각각에 대응하여, 추출된 피처와 해당하는 수도 라벨에 대응하여 저장된 피처의 유사도를 계산하는 단계, 계산된 유사도들 중 임계 값보다 큰 유사도가 없는 경우, 제2 룩업 테이블에 포함되지 않은 새로운 수도 라벨을 객체 영역에 부여될 수도 라벨로 결정하는 단계 및 계산된 유사도들 중 임계 값보다 큰 유사도가 있는 경우, 계산된 유사도들 중 가장 높은 유사도에 대응하는 수도 라벨을 상기 객체 영역에 부여될 수도 라벨로 결정하는 단계를 포함할 수

있다.

[0047] 다시 말해, 일실시예에 따른 객체 검색 모델은 라벨 없는 입력 데이터가 주어졌을 때, 입력 데이터의 객체 영역에서 추출된 피처와 제2 룩업 테이블에 저장된 피처 사이의 유사도를 계산하여 유사한 피처가 제2 룩업 테이블에 있는 경우 해당 피처의 수도 라벨을 입력 데이터의 객체 영역에 부여하며, 추출된 피처가 제2 룩업 테이블에 저장된 모든 피처와 충분히 다른 경우 새로운 수도 라벨을 입력 데이터의 객체 영역에 부여한다.

[0048] 일실시예에 따른 때, 피처들의 유사 여부는 피처들 사이의 유사도를 미리 정해진 임계 값과 비교하여 결정될 수 있다. 예를 들어, 피처들 사이의 유사도가 임계 값보다 큰 경우 해당 피처들은 유사한 것으로 판단될 수 있고, 그렇지 않은 경우 유사하지 않은 것으로 판단될 수 있다. 미리 정해진 임계 값은 학습 데이터 내 라벨이 존재하는 객체 영역의 피처와 제1 룩업 테이블에 해당 라벨에 대응하여 저장된 피처 사이의 유사도에 기초하여 결정될 수 있다. 예를 들어, 피처들 사이의 유사도 판단의 기준이 되는 임계 값 α 는 아래의 수학적 식 1과 같이 정의될 수 있다.

수학적 식 1

$$\alpha = \frac{1}{|\Omega|} \sum_{(x_i, j) \in \Omega} S_L(x_i, L_j)$$

[0049] 여기서, Ω 는 학습 데이터에 포함된 라벨링된 데이터 i 에서 추출된 피처 x_i 와 대응하는 라벨 인덱스 j 의 쌍의 집합, $S_L(x_i, L_j)$ 는 x_i 와 제1 룩업 테이블에 기 저장된 라벨 인덱스 j 에 대응하는 피처 L_j 와의 유사도를 나타낸다. 다시 말해, 피처들 사이의 유사도 판단의 기준이 되는 임계 값 α 는 학습 데이터에 포함된 라벨링된 데이터에서 추출된 피처와 제1 룩업 테이블에 기 저장된 해당 라벨의 피처 사이의 유사도의 평균에 해당할 수 있다. 일실시예에 따른 때, 임계 값은 이전 학습 에포크(epoch)에서 수학적 식 1에 따라 계산된 α 값을 이용할 수 있다.

[0051] 라벨링되지 않은 데이터에서 추출된 피처의 수도 라벨 인덱스 $\hat{j}$ 는 수학적 식 2과 같이 정의될 수 있다.

수학적 식 2

$$\hat{j} = \begin{cases} j^* & \text{if } S_P(\hat{x}_i, P_{j^*}) > \alpha \\ |P_j| + 1 & \text{otherwise} \end{cases}$$

[0053] 여기서, $|P_j|$ 는 제2 룩업 테이블에 저장된 수도 라벨의 수 또는 피처의 수를 나타내며, $j^* = \operatorname{argmax}_j S_P(\hat{x}_i, P_j)$, 즉 제2 룩업 테이블에 기 저장된 수도 라벨 인덱스 j 에 대응하는 피처 P_j 중 임계 값 α 보다 큰 유사도를 가지면서, 추출된 피처 $\hat{x}_i$ 와 가장 높은 유사도를 갖는 피처 P_{j^*} 에 대응하는 수도 라벨을 의미한다.

[0054] 새로운 수도 라벨이 부여되는 경우, 제2 룩업 테이블의 크기는 동적으로 커질 수 있다. 다시 말해, 제1 룩업 테이블은 학습 데이터에 포함된 라벨의 수에 따라 고정된 크기를 갖는 반면, 제2 룩업 테이블은 할당되는 수도 라벨의 수에 따라 동적으로 커지거나 작아질 수 있다.

[0055] 일실시예에 따른 때, 학습 초기에 잘못된 수도 라벨링에 의한 에러를 줄이기 위해 라벨이 없는 데이터에 대한 수도 라벨링은 매 에포크(epoch)마다 새롭게 수행될 수 있다.

[0057] 일실시예에 따른 객체 검색 모델을 학습시키는 단계는 추출된 피처와 제1 룩업 테이블(121)에 저장된 피처 사이의 유사도 및 추출된 피처와 제2 룩업 테이블(141)에 저장된 피처 사이의 유사도에 기초하여, 손실 함수(151)를 결정하는 단계 및 손실 함수(151)를 최소화하는 방향으로 객체 검색 모델의 네트워크 파라미터를 학습시키는 단계를 포함할 수 있다. 이하에서 상술하겠으나, 일실시예에 따른 제1 룩업 테이블 및 제2 룩업 테이블은 추출된

피처를 반영하여 갱신될 수 있으며, 도 1에 도시된 바와 같이 손실 함수(151)는 갱신되기 이전의 제1 룩업 테이블(121) 및 갱신되기 이전의 제2 룩업 테이블(141)에 기초하여 결정될 수 있다. 한편, 일실시예에 따를 때, 도 1에 도시된 바와 달리 손실 함수(151)는 갱신된 이후의 제1 룩업 테이블(122) 및 갱신된 이후의 제2 룩업 테이블(142)에 기초하여 결정될 수도 있다.

[0058] 일실시예에 따른 손실 함수(151)는 객체 영역에 대응하는 라벨이 존재하는 경우, 추출된 피처와 제1 룩업 테이블에 해당 라벨에 대응하여 저장된 피처 사이의 유사도가 클수록 작게 결정될 수 있다. 다시 말해, 입력 데이터 i에서 추출된 피처 x_i 가 라벨 인덱스 j로 라벨링된 경우, 제1 룩업 테이블에 기 저장된 라벨 인덱스 j에 대응하는 피처와 유사하고, 제1 룩업 테이블 및 제2 룩업 테이블에 기 저장된 다른 피처와 다르게 x_i 가 추출되도록 객체 검색 모델의 네트워크 파라미터의 학습이 진행될 수 있다. 라벨링된 학습 데이터 입력에 대응하는 객체 검색 모델의 학습을 위한 손실 함수는 다음의 수학적 식 3과 같이 정의될 수 있다.

수학적 식 3

$$Loss(x_i, j, L, P) = -\log \frac{\exp\left(\frac{S(x_i, L_j)}{\tau}\right)}{\sum_{k_1} \exp\left(\frac{S(x_i, L_{k_1})}{\tau}\right) + \sum_{k_2} \exp\left(\frac{S(x_i, P_{k_2})}{\tau}\right)}$$

[0059] 여기서, $S_{k1}(x_i, L_{k1})$ 는 x_i 와 제1 룩업 테이블에 기 저장된 라벨 인덱스 k1에 대응하는 피처 L_{k1} 와의 유사도, $S_{k2}(x_i, L_{k2})$ 는 x_i 와 제2 룩업 테이블에 기 저장된 라벨 인덱스 k2에 대응하는 피처 L_{k2} 와의 유사도를 나타낸다.

[0061] 수학적 식 3과 같이 정의된 손실 함수를 최소화함으로써, j 라벨 인덱스로 라벨링된 입력 데이터의 객체 영역으로부터 추출한 피처 x_i 가 제1 룩업 테이블에 기 저장된 라벨 인덱스 j에 대응하여 저장된 피처에 가까워지도록 네트워크 파라미터가 학습될 수 있다.

[0062] 한편, 객체 영역에 대응하는 수도 라벨이 존재하는 경우, 추출된 피처와 제2 룩업 테이블에 해당 수도 라벨에 대응하여 저장된 피처와의 유사도가 클수록 작게 결정될 수 있다. 다시 말해, 입력 데이터 i에서 추출된 피처 $\hat{x}_i$ 가 수도 라벨 인덱스 $\hat{j}$ 로 수도 라벨링된 경우, 제2 룩업 테이블에 기 저장된 수도 라벨 인덱스 $\hat{j}$ 에 대응하는 피처와 유사하고, 제1 룩업 테이블 및 제2 룩업 테이블에 기 저장된 다른 피처와 다르게 $\hat{x}_i$ 가 추출되도록 객체 검색 모델의 네트워크 파라미터의 학습이 진행될 수 있다. 수도 라벨링된 학습 데이터 입력에 대응하는 객체 검색 모델의 학습을 위한 손실 함수는 다음의 수학적 식 4와 같이 정의될 수 있다.

수학적 식 4

$$Loss(\hat{x}_i, \hat{j}, L, P) = -\log \frac{\exp\left(\frac{S(\hat{x}_i, P_{\hat{j}})}{\tau}\right)}{\sum_{k_1} \exp\left(\frac{S(\hat{x}_i, L_{k_1})}{\tau}\right) + \sum_{k_2} \exp\left(\frac{S(\hat{x}_i, P_{k_2})}{\tau}\right)}$$

[0063] 여기서, $S_{k1}(\hat{x}_i, L_{k1})$ 는 $\hat{x}_i$ 와 제1 룩업 테이블에 기 저장된 라벨 인덱스 k1에 대응하는 피처 L_{k1} 와의 유사도, $S_{k2}(\hat{x}_i, L_{k2})$ 는 $\hat{x}_i$ 와 제2 룩업 테이블에 기 저장된 라벨 인덱스 k2에 대응하는 피처 L_{k2} 와의 유사도를 나타낸다.

[0065] 수학적 식 4와 같이 정의된 손실 함수를 최소화함으로써, $\hat{j}$ 수도 라벨 인덱스로 수도 라벨링된 입력 데이터의 객체 영역으로부터 추출한 피처 $\hat{x}_i$ 가 제2 룩업 테이블에 기 저장된 수도 라벨 인덱스 $\hat{j}$ 에 대응하여 저장된 피처에 가까워지도록 네트워크 파라미터가 학습될 수 있다.

[0066] 일실시예에 따를 때, 객체 검색 모델의 네트워크 파라미터는 SGD optimizer를 사용하여 수학적 식 3 및 4의 손실 함수를 최소화하는 방향으로 학습될 수 있다.

- [0067] 일실시예에 따른 때, 객체 검색 모델의 네트워크 파라미터가 학습 이터레이션(iteration)마다 학습되는 것과 마찬가지로 제1 룩업 테이블 및 제2 룩업 테이블도 학습 이터레이션마다 갱신될 수 있다.
- [0068] 일실시예에 따른 제1 룩업 테이블을 갱신하는 단계(120)는 제1 룩업 테이블에 추출된 피처의 라벨에 대응하는 피처가 저장되어 있는 경우, 추출된 피처를 반영하여 제1 룩업 테이블에 해당 라벨에 대응하여 저장된 피처를 갱신하는 단계 및 제1 룩업 테이블에 추출된 피처의 라벨에 대응하는 피처가 저장되어 있지 않은 경우, 제1 룩업 테이블에 추출된 피처를 해당 라벨에 대응하여 저장하는 단계를 포함할 수 있다.
- [0069] 일실시예에 따른 제2 룩업 테이블을 갱신하는 단계(140)는 제2 룩업 테이블에 추출된 피처의 수도 라벨에 대응하는 피처가 저장되어 있는 경우, 추출된 피처를 반영하여 제2 룩업 테이블에 해당 수도 라벨에 대응하여 저장된 피처를 갱신하는 단계 및 제2 룩업 테이블에 추출된 피처의 수도 라벨에 대응하는 피처가 저장되어 있지 않은 경우, 제2 룩업 테이블에 추출된 피처를 해당 수도 라벨에 대응하여 저장하는 단계를 포함할 수 있다.
- [0070] 다시 말해, 추출된 피처 x_i 가 제1 룩업 테이블(또는 제2 룩업 테이블)의 라벨 인덱스(또는 수도 라벨 인덱스) j 를 지닐 때, 제1 룩업 테이블(또는 제2 룩업 테이블)에서 라벨 인덱스(또는 수도 라벨 인덱스)에 대응하여 저장된 피처는 아래의 수학적 식 5와 같이 갱신될 수 있다.

수학적 식 5

[0071]
$$T_j \leftarrow \frac{T_j + x_i}{\|T_j + x_i\|_2}$$

- [0072] 도 2는 일실시예에 따른 객체 검색 모델을 이용한 객체 검색 방법을 설명하기 위한 도면이다.
- [0073] 도 2를 참조하면, 일실시예에 따른 객체 검색 방법은 검색 대상을 포함하는 쿼리 이미지(201) 및 검색용 이미지(202)를 수신하는 단계, 쿼리 이미지(201)에서, 인코더(210)를 이용하여 검색 대상의 제1 피처를 추출하는 단계, 검색용 이미지(202)에서, 인코더(210)를 이용하여 검색용 이미지에 포함된 적어도 하나의 객체 영역의 제2 피처를 추출하는 단계, 및 제1 피처 및 적어도 하나의 제2 피처에 기초하여, 검색용 이미지에서 검색 대상에 해당하는 영역을 획득하는 단계를 포함한다.
- [0074] 일실시예에 따른 인코더(210)는 상술한 학습 방법에 따라 학습된 뉴럴 네트워크로 구성될 수 있다. 다시 말해, 인코더(210)는 라벨링되지 않은 학습 데이터를 수도 라벨링하여, 다른 라벨에 대응하는 데이터와 구분되는 피처를 추출하도록 학습된 뉴럴 네트워크로 구성될 수 있다. 여기서, 뉴럴 네트워크는 라벨링된 학습 데이터의 피처를 해당 라벨에 대응하여 저장하는 제1 룩업 테이블 및 라벨링되지 않은 학습 데이터를 수도 라벨링하여, 라벨링되지 않은 학습 데이터의 피처를 해당 수도 라벨에 대응하여 저장하는 제2 룩업 테이블에 기초하여 학습될 수 있다.
- [0075] 일실시예에 따른 검색 대상 검출 모듈(220)은 인코더(210)를 통해 추출된 검색 대상의 제1 피처와 적어도 하나의 제2 피처의 유사도를 비교하여, 검색용 이미지에 포함된 적어도 하나의 객체 영역 중 검색 대상에 해당하는 것으로 판단되는 객체 영역을 검출하는 모듈에 해당할 수 있다. 검색 대상 검출 모듈(220)에 의해 검색용 이미지(202)에서 검색 대상에 해당하는 영역이 표시된 출력 이미지(203)가 획득될 수 있다.
- [0077] 도 3은 일실시예에 따른 객체 검색 모델의 구조를 예시한 도면이다.
- [0078] 이하에서, 객체는 사람인 경우를 예로 들어 설명한다. 그러나, 일실시예에 따른 객체 검색 모델의 검색 대상을 사람으로 한정하는 것은 아니며, 학습 데이터에 포함된 다양한 유형의 객체가 객체 검색 모델의 검색 대상이 될 수 있다.
- [0079] 일실시예에 따른 객체 검색 모델을 이용한 사람 검색 방법은 입력 영상에서 사람을 검출하고, 검출된 사람이 찾고자 하는 사람인지 구분해야 하는 바, 입력 영상에서 사람 영역을 검출하는 사람 검출(detection) 과정과 검출된 사람 영역에서 찾고자 하는 사람을 구분하는 사람 재인식(re-identification) 과정으로 이루어질 수 있다.
- [0080] 도 3을 참조하면, 일실시예에 따른 사람 검색 방법을 수행하는 객체 검색 모델은 입력된 영상(301) 내 사람 영

역의 피처를 추출하기 위한 Faster R-CNN(310), Faster R-CNN(310)을 기반으로 추출된 피처에 기초하여 사람을 재인식(re-identification)하기 위한 네트워크(320) 및 사람 영역을 검출(detection)하기 위한 네트워크(330)를 포함할 수 있다.

[0081] 일실시예에 따른 재인식 네트워크(320)는 Faster R-CNN(310)에 의해 추출된 피처에 기초하여, 입력된 영상 내 사람 영역을 수도 라벨링하는 네트워크에 해당할 수 있다. Faster R-CNN(310) 및 재인식 네트워크(320)는 상술한 학습 방법에 따라 제1 룩업 테이블(302) 및 제2 룩업 테이블(303)을 이용하여 학습될 수 있다.

[0082] 일실시예에 따른 때, 재인식 네트워크(320)는 추출된 피처를 기 저장된 식별 정보에 대응하는 피처와 비교하여, 유사도가 높은 피처의 식별 정보로 입력된 영상 내 사람 영역을 수도 라벨링할 수 있다. 수도 라벨링을 통해 부여된 수도 라벨은 해당 영역의 식별 정보에 대응할 수 있으며, 객체 검색 모델은 입력 영상에서 검출된 사람 영역의 수도 라벨을 출력함으로써, 사람 영역의 식별 정보를 제공할 수 있다.

[0083] 일실시예에 따른 때, 재인식 네트워크(320)는 검색 대상의 이미지에서 추출된 검색 대상의 피처를 수도 라벨링하여 저장하고, 입력된 검색용 이미지의 사람 영역에서 추출된 피처를 수도 라벨링하여, 검색 대상과 검색용 이미지 내 사람이 일치하는지에 관한 정보를 출력할 수 있다. 보다 구체적으로, 검색용 이미지의 사람 영역에 검색 대상과 동일한 수도 라벨이 부여된 경우, 검색용 이미지 내 해당 영역을 검색 대상에 해당하는 영역으로 출력할 수 있다.

[0085] 도 4a는 일실시예에 따른 객체 검색 모델을 이용하여 찾고자 하는 사람의 영상을 도식한 도면이고, 도 4b는 일실시예에 따른 객체 검색 모델을 이용하여 입력 영상에서 검색 대상을 검출한 결과를 도식한 도면이다.

[0086] 도 4a 및 도 4b를 참조하면, 일실시예에 따른 객체 검색 모델은 도 4a와 같이 찾고자 하는 사람이 주어졌을 때, 도 4b와 같이 검색용 이미지에서 다른 자세로 있는 검색 대상(401)을 검출하여 검색용 이미지에 검색 대상(401)을 표시한 영상을 출력할 수 있다. 도 4b에 도시된 바와 같이, 일실시예에 따른 객체 검색 모델은 복수의 사람들을 포함하는 입력 영상에서 복수의 사람 영역들(401 내지 404)을 검출하고, 복수의 사람 영역들(401 내지 404)이 검색 대상에 해당할 확률을 계산함으로써, 입력 영상에서 검색 대상(401)을 검출할 수 있다. 도 4b는 일실시예에 따른 객체 검색 모델의 출력 결과를 예시한 도면으로, 출력 결과는 이에 한정되는 것은 아니며, 입력 영상에서 검색 대상을 검출한 결과의 출력 방법은 다양한 형태로 구현될 수 있다.

[0088] 도 5는 일실시예에 따른 장치의 구성의 예시도이다.

[0089] 도 5를 참조하면, 장치(500)는 프로세서(501), 메모리(503) 및 입출력 장치(505)를 포함한다.

[0090] 일실시예에 따른 장치(500)는 상술한 객체 검색 모델이 구현된 장치로, 서버 및 사용자의 디바이스(예를 들어, 휴대폰, 컴퓨터 등)를 포함할 수 있다. 프로세서(501)는 도 1 내지 도 4를 통하여 기술한 적어도 하나의 객체 검색 방법 및 학습 방법을 수행할 수 있다. 메모리(503)는 객체 검색과 관련된 정보 또는 객체 검색 모델의 학습을 위한 정보를 저장하거나 상술한 객체 검색 방법 또는 객체 검색 모델의 학습 방법이 구현된 프로그램을 저장할 수 있다. 메모리(503)는 휘발성 메모리 또는 비휘발성 메모리일 수 있다.

[0091] 프로세서(501)는 프로그램을 실행하고, 장치(500)를 제어할 수 있다. 프로세서(501)에 의하여 실행되는 프로그램의 코드는 메모리(503)에 저장될 수 있다. 장치(500)는 입출력 장치(505)를 통하여 외부 장치(예를 들어, 퍼스널 컴퓨터 또는 네트워크)에 연결되고, 데이터를 교환할 수 있다.

[0093] 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, 컨트롤러, 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.

[0094] 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.

[0095] 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.

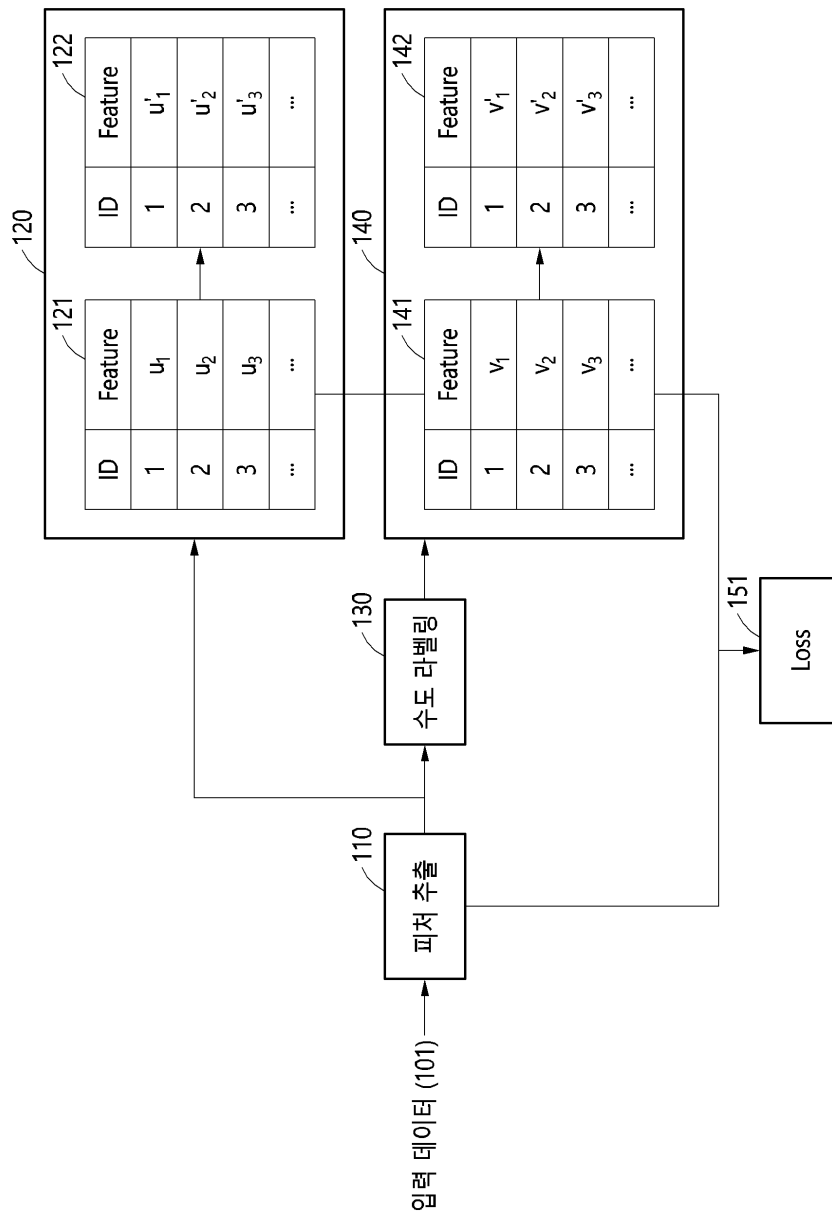
[0096] 위에서 설명한 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 또는 복수의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0097] 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.

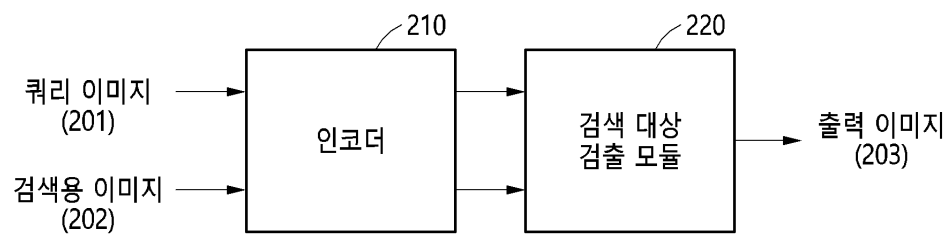
[0098] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

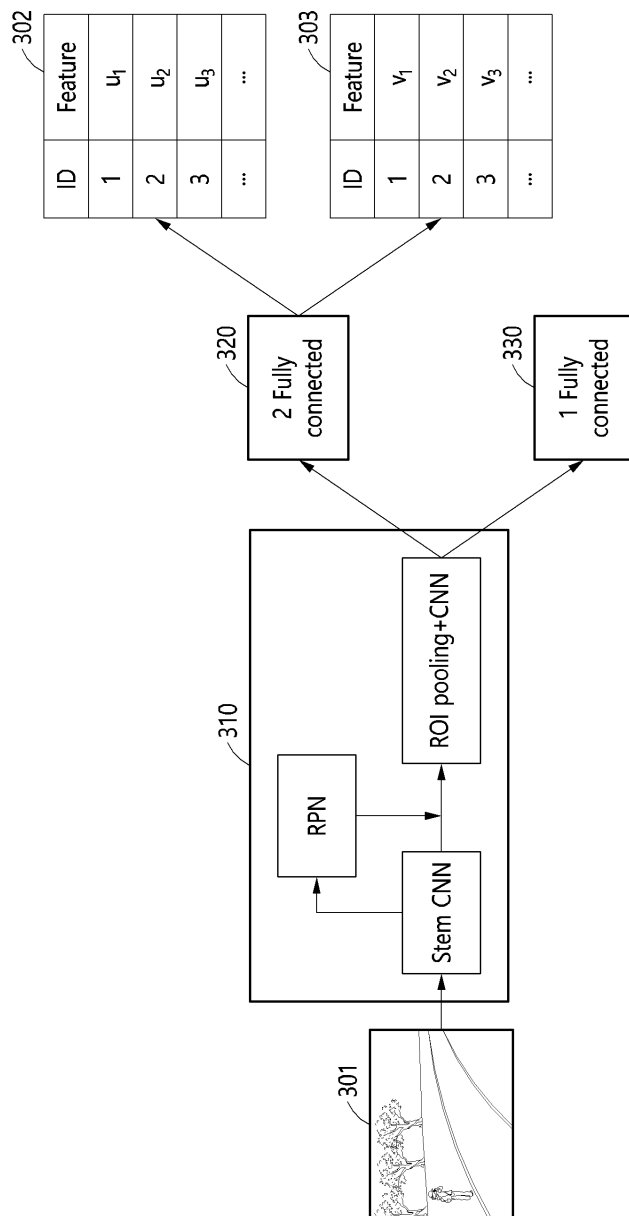
도면1



도면2



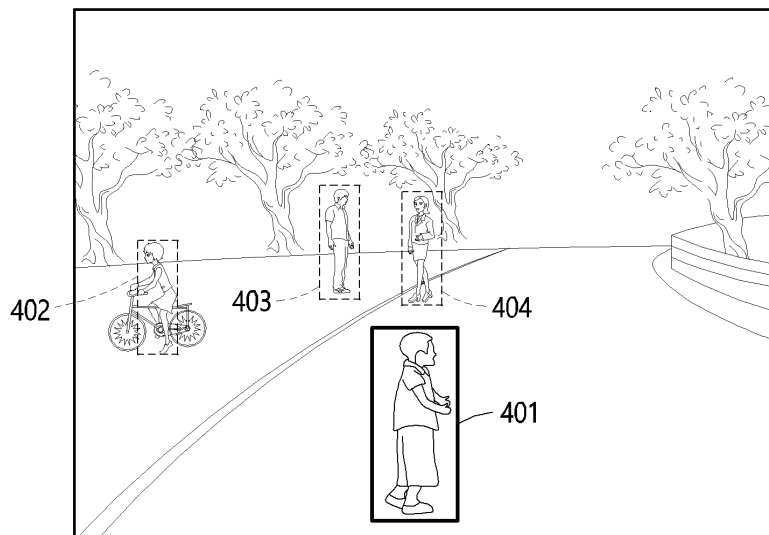
도면3



도면4a



도면4b



도면5

