



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0007284
(43) 공개일자 2025년01월14일

(51) 국제특허분류(Int. Cl.)
G06V 10/764 (2022.01) G06N 3/08 (2023.01)
G06T 5/00 (2024.01) G06V 10/26 (2022.01)
G06V 10/40 (2022.01) G06V 10/774 (2022.01)
G06V 20/56 (2022.01)
(52) CPC특허분류
G06V 10/764 (2023.08)
G06N 3/08 (2023.01)
(21) 출원번호 10-2023-0087216
(22) 출원일자 2023년07월05일
심사청구일자 없음

(71) 출원인
현대자동차주식회사
서울특별시 서초구 현릉로 12 (양재동)
기아 주식회사
서울특별시 서초구 현릉로 12 (양재동)
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
박진호
서울특별시 강동구 구천면로24길 9, 엘림스퀘어 II 1308호
김진솔
경기도 화성시 남양읍 시청로102번길 11, 204-901 (뒷면에 계속)
(74) 대리인
특허법인태평양

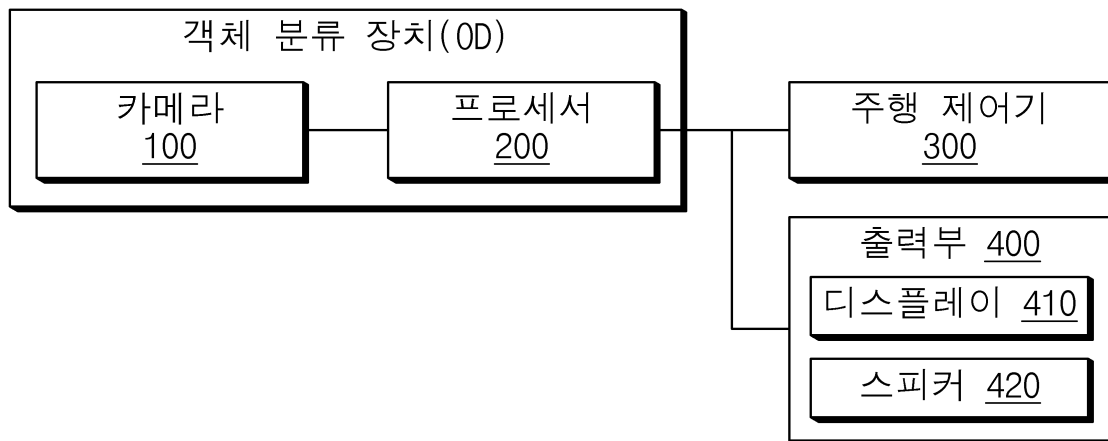
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 객체 분류 장치 및 그 방법

(57) 요약

본 발명은 객체 분류 장치, 및 그 방법에 관한 것으로, 본 발명의 실시 예에 의한 객체 분류 장치는 이미지를 획득하는 카메라, 및 이미지를 딥러닝 학습하는 프로세서를 포함할 수 있다. 프로세서는 객체 분류 모델을 이용하여 이미지를 학습한 것을 바탕으로 제1 깊이값을 획득하고, 이미지의 일부 영역을 마스킹한 부분 이미지를 획득하며, 객체 분류 모델을 이용하여 부분 이미지를 학습한 것을 바탕으로 제2 깊이값을 획득하고, 제1 깊이값 및 제2 깊이값의 편차를 줄이도록 객체 분류 모델을 학습할 수 있다.

대표도 - 도1



(52) CPC특허분류

G06T 5/75 (2024.01)
G06V 10/26 (2023.08)
G06V 10/40 (2023.08)
G06V 10/774 (2023.08)
G06V 20/56 (2023.08)
G06T 2207/10028 (2013.01)
G06T 2207/30252 (2013.01)

(72) 발명자

김장윤

서울특별시 강동구 상암로 53-10, 203호

김승룡

서울특별시 성북구 고려대로22길 6-7, 과학도서관
404C

백중범

서울특별시 동대문구 한빛로 50-2, 204호

박성훈

서울특별시 노원구 상계로23길 57-19, 이도빌라
503호

김경년

서울특별시 성북구 고려대로 47, 207호

명세서

청구범위

청구항 1

이미지를 획득하는 카메라; 및

상기 이미지를 딥러닝 학습하는 프로세서;를 포함하고,

상기 프로세서는

객체 분류 모델을 이용하여, 이미지를 학습한 것을 바탕으로 제1 깊이값을 획득하고,

상기 이미지의 일부 영역을 마스킹한 부분 이미지를 획득하며,

상기 객체 분류 모델을 이용하여, 상기 부분 이미지를 학습한 것을 바탕으로 제2 깊이값을 획득하고,

상기 제1 깊이값 및 상기 제2 깊이값의 편차를 줄이도록 상기 객체 분류 모델을 학습하는 것을 특징으로 하는 객체 분류 장치.

청구항 2

제 1 항에 있어서,

상기 프로세서는

상기 이미지의 차원 변환 또는 상기 이미지의 계조 변환을 바탕으로 확장된 학습 데이터를 학습하여 상기 제1 깊이값을 획득하는 것을 특징으로 하는 객체 분류 장치.

청구항 3

제 1 항에 있어서,

상기 프로세서는

상기 이미지를 패치 단위로 분할하고, 상기 복수의 패치들을 상기 객체 분류 모델의 인코더에 제공하며, 상기 인코더를 이용하여 제1 특징값을 출력하고, 상기 제1 특징값을 상기 객체 분류 모델의 디코더에 제공하며, 상기 디코더를 이용하여 상기 제1 깊이값을 획득하는 것을 특징으로 하는 객체 분류 장치.

청구항 4

제 3 항에 있어서,

상기 프로세서는

마스크를 이용하여 상기 패치 단위로 분할된 상기 이미지의 일부 영역을 샘플링하여, 상기 부분 이미지들을 획득하는 것을 특징으로 하는 객체 분류 장치.

청구항 5

제 4 항에 있어서,

상기 프로세서는

상기 부분 이미지들 각각에서 샘플링 된 상기 패치들이 서로 중복되지 않도록 설정된 상기 마스크들을 이용하는

것을 특징으로 하는 객체 분류 장치.

청구항 6

제 4 항에 있어서,

상기 프로세서는

상기 부분 이미지들을 상기 인코더에 제공하고, 상기 인코더를 이용하여 상기 부분 이미지들에 대한 부분 특징값을 추출하며, 상기 부분 특징값들을 결합하여 제2 특징값을 추출하고, 상기 제2 특징값을 상기 디코더에 제공하며, 상기 디코더를 이용하여 상기 제2 깊이값을 획득하는 것을 특징으로 하는 객체 분류 장치.

청구항 7

제 6 항에 있어서,

상기 프로세서는

상기 제1 특징값과 상기 제2 특징값의 편차에 비례하는 제1 손실함수의 크기를 계산하고, 상기 제1 손실함수의 크기를 줄이도록 상기 인코더의 파라미터를 조절하는 것을 특징으로 하는 객체 분류 장치.

청구항 8

제 1 항에 있어서,

상기 프로세서는

상기 제1 깊이값과 상기 제2 깊이값의 편차에 비례하는 제2 손실함수의 크기를 계산고, 상기 제2 손실함수의 크기를 줄이도록 상기 객체 분류 모델의 파라미터를 조절하는 것을 특징으로 하는 객체 분류 장치.

청구항 9

제 8 항에 있어서,

상기 프로세서는

상기 제2 손실함수를 바탕으로 상기 파라미터가 조절된 상기 객체 분류 모델의 불확실 정도를 판단하고, 상기 불확실 정도를 바탕으로 상기 제2 손실함수의 가중치를 조절하는 것을 특징으로 하는 객체 분류 장치.

청구항 10

제 1 항에 있어서,

상기 프로세서는

상기 이미지를 학습한 것을 바탕으로 제1 클래스를 획득하고, 상기 부분 이미지를 학습한 것을 바탕으로 제2 클래스를 획득하며, 상기 제2 클래스와 상기 제2 클래스의 편차를 줄이도록 상기 객체 분류 모델의 파라미터를 조절하는 것을 특징으로 하는 객체 분류 장치.

청구항 11

객체 분류 모델을 이용하여, 이미지를 학습한 것을 바탕으로 제1 깊이값을 획득하는 단계;

상기 이미지의 일부 영역을 마스킹한 부분 이미지들을 획득하는 단계;

상기 객체 분류 모델을 이용하여, 상기 부분 이미지를 학습한 것을 바탕으로 제2 깊이값을 획득하는 단계; 및
상기 제1 깊이값 및 상기 제2 깊이값의 편차를 줄이도록 상기 객체 분류 모델을 학습하는 단계;
를 포함하는 객체 분류 방법.

청구항 12

제 11 항에 있어서,

상기 제1 깊이값을 획득하는 단계는

상기 이미지의 차원 변환 또는 상기 이미지의 계조 변환을 바탕으로 확장된 학습 데이터를 학습하는 단계를 더 포함하는 것을 특징으로 하는 객체 분류 방법.

청구항 13

제 11 항에 있어서,

상기 제1 깊이값을 획득하는 단계는

상기 이미지를 패치 단위로 분할하는 단계;

상기 복수의 패치들을 상기 객체 분류 모델의 인코더에 제공하는 단계;

상기 인코더를 이용하여 제1 특징값을 출력하는 단계;

상기 제1 특징값을 상기 객체 분류 모델의 디코더에 제공하는 단계; 및

상기 디코더를 이용하여 상기 제1 깊이값을 획득하는 단계;를 포함하는 것을 특징으로 하는 객체 분류 방법.

청구항 14

제 13 항에 있어서,

상기 부분 이미지들을 획득하는 단계는

마스크를 이용하여, 상기 패치 단위로 분할된 상기 이미지의 일부 영역을 샘플링하는 단계를 포함하는 것을 특징으로 하는 객체 분류 방법.

청구항 15

제 14 항에 있어서,

상기 부분 이미지들을 획득하는 단계는

상기 부분 이미지들 각각에서 샘플링 된 상기 패치들이 서로 중복되지 않도록 설정된 상기 마스크들을 이용하는 것을 특징으로 하는 객체 분류 방법.

청구항 16

제 14 항에 있어서,

상기 제2 깊이값을 획득하는 단계는

상기 부분 이미지들을 상기 인코더에 제공하는 단계;

상기 인코더를 이용하여 상기 부분 이미지들에 대한 부분 특징값을 추출하는 단계;
 상기 부분 특징값들을 결합하여, 제2 특징값을 출력하는 단계;
 상기 제2 특징값을 상기 디코더에 제공하는 단계; 및
 상기 디코더를 이용하여 상기 제2 깊이값을 획득하는 단계;를 포함하는 것을 특징으로 하는 객체 분류 방법.

청구항 17

제 16 항에 있어서,
 상기 객체 분류 모델을 학습하는 단계는
 상기 제1 특징값과 상기 제2 특징값의 편차에 비례하는 제1 손실함수의 크기를 계산하는 단계; 및
 상기 제1 손실함수의 크기를 줄이도록 상기 인코더의 파라미터를 조절하는 단계;를 포함하는 것을 특징으로 하는 객체 분류 방법.

청구항 18

제 11 항에 있어서,
 상기 객체 분류 모델을 학습하는 단계는
 상기 제1 깊이값과 상기 제2 깊이값의 편차에 비례하는 제2 손실함수의 크기를 계산하는 단계; 및
 상기 제2 손실함수의 크기를 줄이도록 상기 객체 분류 모델의 파라미터를 조절하는 단계;를 포함하는 것을 특징으로 하는 객체 분류 방법.

청구항 19

제 18 항에 있어서,
 상기 제2 손실함수를 바탕으로 상기 파라미터가 조절된 상기 객체 분류 모델의 불확실 정도를 판단하는 단계; 및
 상기 불확실 정도를 바탕으로 상기 제2 손실함수의 가중치를 조절하는 단계;를 더 포함하는 것을 특징으로 하는 객체 분류 방법.

청구항 20

제 11 항에 있어서,
 상기 이미지를 학습한 것을 바탕으로 제1 클래스를 획득하는 단계;
 상기 부분 이미지를 학습한 것을 바탕으로 제2 클래스를 획득하는 단계; 및
 상기 제2 클래스와 상기 제2 클래스의 편차를 줄이도록 상기 객체 분류 모델의 파라미터를 조절하는 단계;를 더 포함하는 것을 특징으로 하는 객체 분류 방법.

발명의 설명

기술 분야

[0001] 본 발명은 객체 분류 장치 및 그 방법에 관한 것으로, 보다 상세하게는 객체 분류 모델의 성능을 높이기 위한 기술에 관한 것이다.

배경 기술

- [0002] 자율주행 자동차(Autonomous Vehicle)란 운전자 또는 승객의 조작 없이 자동차 스스로 운행이 가능한 자동차를 말하며, 자율주행시스템(Automated Vehicle & Highway Systems)은 이러한 자율주행자동차가 스스로 운행될 수 있도록 모니터링하고 제어하는 시스템을 말한다. 또한, 운전자의 운전 보조를 위해서 차량의 외부를 모니터링하고, 모니터링 된 차량 외부 환경을 바탕으로 다양한 운전 보조 수단을 동작하는 기술들이 제안되고 있다.
- [0003] 자율주행 시스템 또는 운전 보조 수단의 동작은 차량 외부를 모니터링한 결과에 따라 제어될 수 있다.
- [0004] 차량 외부를 모니터링하기 위해서, 카메라, 라이다(Light Detection and Ranging; LiDAR), 또는 레이더(Radio Detection And Ranging; RADAR) 등을 이용하기도 한다. 이 중에서 카메라를 이용하는 모니터링 기술이 널리 연구되고 있다. 카메라는 적은 비용을 통해서 외부 객체를 분류할 수 있는 장점이 있지만, 일반적인 2D 카메라는 2차원 정보만을 포함하기 때문에 깊이값(depth)을 정확하게 판단하기 어렵다.
- [0005] 또한, 차량의 카메라 이외에도 증강현실, 가상 현실, 의료용 영상 등에 이미지를 활용하는 과정에서 깊이값을 정확하게 추정하는 기술은 매우 중요하다.
- [0006] 깊이값을 보다 정확하게 추정하기 위해서는 스테레오 영상을 이용하거나, 정답 데이터를 많이 확보하여야 하는 조건이 필요하다.
- [0007] 따라서, 스테레오 영상이 아닌 단안 영상에서 깊이값을 추정하거나, 정답 데이터를 많이 활용하지 않으면서도 깊이값을 보다 정확하게 추정할 수 있는 기술이 요구되고 있다.

발명의 내용

해결하려는 과제

- [0008] 본 발명은 단안 영상에서 깊이값을 보다 정확하게 추정할 수 있는 객체 분류 장치 및 그 방법을 제공하기 위한 것이다.
- [0009] 또한, 본 발명은 정답 데이터가 부족한 상태에서도 깊이값을 보다 정확하게 추정할 수 있는 객체 분류 장치 및 그 방법을 제공하기 위한 것이다.
- [0010] 본 발명의 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재들로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0011] 본 발명의 실시 예에 의한 객체 분류 장치는 이미지를 획득하는 카메라, 및 이미지를 딥러닝 학습하는 프로세서를 포함할 수 있다. 프로세서는 객체 분류 모델을 이용하여 이미지를 학습한 것을 바탕으로 제1 깊이값을 획득하고, 이미지의 일부 영역을 마스킹한 부분 이미지를 획득하며, 객체 분류 모델을 이용하여 부분 이미지를 학습한 것을 바탕으로 제2 깊이값을 획득하고, 제1 깊이값 및 제2 깊이값의 편차를 줄이도록 객체 분류 모델을 학습할 수 있다.
- [0012] 실시 예에 의하면, 상기 프로세서는 상기 이미지의 차원 변환 또는 상기 이미지의 계조 변환을 바탕으로 확장된 학습 데이터를 학습하여 상기 제1 깊이값을 획득할 수 있다.
- [0013] 실시 예에 의하면, 상기 프로세서는 상기 이미지를 패치 단위로 분할하고, 상기 복수의 패치들을 상기 객체 분류 모델의 인코더에 제공하며, 상기 인코더를 이용하여 제1 특징값을 출력하고, 상기 제1 특징값을 상기 객체 분류 모델의 디코더에 제공하며, 상기 디코더를 이용하여 상기 제1 깊이값을 획득할 수 있다.
- [0014] 실시 예에 의하면, 상기 프로세서는 마스크를 이용하여 상기 패치 단위로 분할된 상기 이미지의 일부 영역을 샘플링하여, 상기 부분 이미지들을 획득할 수 있다.
- [0015] 실시 예에 의하면, 상기 프로세서는 상기 부분 이미지들 각각에서 샘플링 된 상기 패치들이 서로 중복되지 않도록 설정된 상기 마스크들을 이용할 수 있다.
- [0016] 실시 예에 의하면, 상기 프로세서는 상기 부분 이미지들을 상기 인코더에 제공하고, 상기 인코더를 이용하여 상기 부분 이미지들에 대한 부분 특징값을 추출하며, 상기 부분 특징값들을 결합하여 제2 특징값을 추출하고, 상

기 제2 특징값을 상기 디코더에 제공하며, 상기 디코더를 이용하여 상기 제2 깊이값을 획득할 수 있다.

- [0017] 실시 예에 의하면, 상기 프로세서는 상기 제1 특징값과 상기 제2 특징값의 편차에 비례하는 제1 손실함수의 크기를 계산하고, 상기 제1 손실함수의 크기를 줄이도록 상기 인코더의 파라미터를 조절할 수 있다.
- [0018] 실시 예에 의하면, 상기 프로세서는 상기 제1 깊이값과 상기 제2 깊이값의 편차에 비례하는 제2 손실함수의 크기를 계산고, 상기 제2 손실함수의 크기를 줄이도록 상기 객체 분류 모델의 파라미터를 조절할 수 있다.
- [0019] 실시 예에 의하면, 상기 프로세서는 상기 제2 손실함수를 바탕으로 상기 파라미터가 조절된 상기 객체 분류 모델의 불확실 정도를 판단하고, 상기 불확실 정도를 바탕으로 상기 제2 손실함수의 가중치를 조절할 수 있다.
- [0020] 실시 예에 의하면, 상기 프로세서는 상기 이미지를 학습한 것을 바탕으로 제1 클래스를 획득하고, 상기 부분 이미지를 학습한 것을 바탕으로 제2 클래스를 획득하며, 상기 제2 클래스와 상기 제2 클래스의 편차를 줄이도록 상기 객체 분류 모델의 파라미터를 조절할 수 있다.
- [0021] 본 발명의 실시 예에 의한 객체 분류 방법은 객체 분류 모델을 이용하여, 이미지를 학습한 것을 바탕으로 제1 깊이값을 획득하는 단계, 상기 이미지의 일부 영역을 마스킹한 부분 이미지들을 획득하는 단계, 상기 객체 분류 모델을 이용하여, 상기 부분 이미지를 학습한 것을 바탕으로 제2 깊이값을 획득하는 단계 및 상기 제1 깊이값 및 상기 제2 깊이값의 편차를 줄이도록 상기 객체 분류 모델을 학습하는 단계를 포함할 수 있다.
- [0022] 실시 예에 의하면, 상기 제1 깊이값을 획득하는 단계는 상기 이미지의 차원 변환 또는 상기 이미지의 계조 변환을 바탕으로 확장된 학습 데이터를 학습하는 단계를 더 포함할 수 있다.
- [0023] 실시 예에 의하면, 상기 제1 깊이값을 획득하는 단계는 상기 이미지를 패치 단위로 분할하는 단계, 상기 복수의 패치들을 상기 객체 분류 모델의 인코더에 제공하는 단계, 상기 인코더를 이용하여 제1 특징값을 출력하는 단계, 상기 제1 특징값을 상기 객체 분류 모델의 디코더에 제공하는 단계, 및 상기 디코더를 이용하여 상기 제1 깊이값을 획득하는 단계를 포함할 수 있다.
- [0024] 실시 예에 의하면, 상기 부분 이미지들을 획득하는 단계는 마스크를 이용하여, 상기 패치 단위로 분할된 상기 이미지의 일부 영역을 샘플링하는 단계를 포함할 수 있다.
- [0025] 실시 예에 의하면, 상기 부분 이미지들을 획득하는 단계는 상기 부분 이미지들 각각에서 샘플링 된 상기 패치들이 서로 중복되지 않도록 설정된 상기 마스크들을 이용할 수 있다.
- [0026] 실시 예에 의하면, 상기 제2 깊이값을 획득하는 단계는 상기 부분 이미지들을 상기 인코더에 제공하는 단계, 상기 인코더를 이용하여 상기 부분 이미지들에 대한 부분 특징값을 추출하는 단계, 상기 부분 특징값들을 결합하여, 제2 특징값을 출력하는 단계, 상기 제2 특징값을 상기 디코더에 제공하는 단계, 및 상기 디코더를 이용하여 상기 제2 깊이값을 획득하는 단계를 포함할 수 있다.
- [0027] 실시 예에 의하면, 상기 객체 분류 모델을 학습하는 단계는 상기 제1 특징값과 상기 제2 특징값의 편차에 비례하는 제1 손실함수의 크기를 계산하는 단계, 및 상기 제1 손실함수의 크기를 줄이도록 상기 인코더의 파라미터를 조절하는 단계를 포함할 수 있다.
- [0028] 실시 예에 의하면, 상기 객체 분류 모델을 학습하는 단계는 상기 제1 깊이값과 상기 제2 깊이값의 편차에 비례하는 제2 손실함수의 크기를 계산하는 단계, 및 상기 제2 손실함수의 크기를 줄이도록 상기 객체 분류 모델의 파라미터를 조절하는 단계를 포함할 수 있다.
- [0029] 실시 예에 의하면, 상기 제2 손실함수를 바탕으로 상기 파라미터가 조절된 상기 객체 분류 모델의 불확실 정도를 판단하는 단계, 및 상기 불확실 정도를 바탕으로 상기 제2 손실함수의 가중치를 조절하는 단계를 더 포함할 수 있다.
- [0030] 실시 예에 의하면, 상기 이미지를 학습한 것을 바탕으로 제1 클래스를 획득하는 단계, 상기 부분 이미지를 학습한 것을 바탕으로 제2 클래스를 획득하는 단계, 및 상기 제2 클래스와 상기 제2 클래스의 편차를 줄이도록 상기 객체 분류 모델의 파라미터를 조절하는 단계를 더 포함할 수 있다.

발명의 효과

- [0031] 본 발명의 실시 예에 의하면, 단일 이미지를 이용하여 깊이값을 추정할 수 있기 때문에, 스테레오 카메라가 아닌 2D 카메라가 획득한 이미지로부터 깊이값을 추정할 수 있다.

[0032] 또한, 본 발명의 실시 예에 의하면 영상의 특징이 누락된 부분 이미지를 학습한 결과와 원본 이미지를 학습한 결과를 일치시키도록 딥러닝이 진행되기 때문에, 객체 분류 모델의 정확도를 높여서 정답 데이터를 줄이면서도 깊이값을 보다 정확하게 추정할 수 있다.

[0033] 이 외에, 본 문서를 통해 직접적 또는 간접적으로 파악되는 다양한 효과들이 제공될 수 있다.

도면의 간단한 설명

[0034] 도 1은 본 발명의 실시 예에 의한 객체 검출 장치의 구성을 나타내는 블록도이다.

도 2는 본 발명의 실시 예에 의한 객체 검출 장치가 장착된 차량을 나타내는 도면이다.

도 3은 본 발명의 실시 예에 의한 프로세서의 구성을 나타내는 블록도이다.

도 4는 본 발명의 실시 예에 의한 프로세서의 딥러닝 동작을 설명하기 위한 도면이다.

도 5는 본 발명의 실시 예에 의한 객체 검출 방법을 설명하기 위한 순서도이다.

도 6은 이미지 패치들의 실시 예를 나타내는 도면이다.

도 7은 이미지 패치들을 시퀀스 데이터로 인코더에 제공하는 실시 예를 설명하기 위한 도면이다.

도 8은 제1 특징값을 추출하는 과정에서 활용되는 데이터 증강 기법을 설명하기 위한 도면이다.

도 9는 이미지 패치들을 시퀀스 데이터로 인코더에 제공하는 다른 실시 예를 설명하기 위한 도면이다.

도 10은 인코더를 이용하여 제2 특징값을 추출하는 과정을 설명하기 위한 도면이다.

도 11은 제1 손실함수를 바탕으로 인코더를 학습하는 방식을 설명하기 위한 도면이다.

도 12는 제2 손실함수를 바탕으로 객체 분류 모델을 학습하는 설명하기 위한 도면이다.

도 13은 본 발명의 다른 실시 예에 의한 프로세서의 딥러닝 동작을 설명하기 위한 도면이다.

도 14는 본 발명의 다른 실시 예에 의한 딥러닝 동작을 설명하기 위한 도면이다.

도 15는 본 발명의 일 실시 예에 따른 컴퓨팅 시스템을 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0035] 이하, 본 발명의 일부 실시 예들을 예시적인 도면을 통해 상세하게 설명한다. 각 도면의 구성요소들에 참조부호를 부가함에 있어서, 동일한 구성요소들에 대해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 발명의 실시 예를 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 실시 예에 대한 이해를 방해한다고 판단되는 경우에는 그 상세한 설명은 생략한다.

[0036] 본 발명의 실시 예의 구성 요소를 설명하는 데 있어서, 제 1, 제 2, A, B, (a), (b) 등의 용어를 사용할 수 있다. 이러한 용어는 그 구성 요소를 다른 구성 요소와 구별하기 위한 것일 뿐, 그 용어에 의해 해당 구성 요소의 본질이나 차례 또는 순서 등이 한정되지 않는다. 또한, 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 가진 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0037] 이하, 도 1 내지 도 15를 참조하여, 본 발명의 실시 예들을 구체적으로 설명하기로 한다.

[0038] 도 1은 본 발명의 실시 예에 의한 객체 검출 장치의 구성을 나타내는 블록도이다. 도 2는 본 발명의 실시 예에 의한 객체 검출 장치가 장착된 차량을 나타내는 도면이다. 이하, 객체 검출 장치는 도 2에서와 같이 차량에 장착된 실시 예를 중심으로 설명하지만, 객체 검출 장치가 활용될 수 있는 영역은 이에 한정되지 않을 수 있다.

[0039] 도 1 및 도 2를 참조하면, 본 발명의 실시 예에 의한 객체 검출 장치(OD)는 카메라(100), 및 프로세서(200)를 포함할 수 있다.

[0040] 카메라(100)는 차량의 외부 이미지를 획득하기 위한 것으로, 프런트 윈드 쉴드에 근접하게 배치되거나 프런트

범퍼 또는 라디에이터 그릴 주변에 배치될 수 있다. 외부 이미지는 2차원의 이미지 평면에 표현된 것일 수 있으며, 이미지 평면의 각 픽셀들은 영상 좌표로 표현될 수 있다.

[0041] 프로세서(200)는 카메라(100)가 획득한 외부 이미지에서 객체를 분류하기 위한 객체 분류 모델을 이용할 수 있다. 객체 분류 모델은 프로세서(200)에 포함되거나, 외부 메모리에 저장될 수 있다. 객체 분류 모델은 비전 트랜스포머(Vision Transformer)를 이용할 수 있다. 비전 트랜스포머는 CNN(Convolutional Neural Networks)에 의존하지 않으면서 이미지 패치의 시퀀스를 입력으로 이용할 수 있다.

[0042] 프로세서(200)는 객체 분류 모델과 같은 하나 이상의 인공지능(artificial intelligence; 이하, AI) 프로세서를 포함할 수 있다. AI 프로세서는 미리 저장된 프로그램을 이용하여 신경망을 학습할 수 있다. 타겟 차량 및 위험 차량을 검출하기 위한 신경망은 인간의 뇌 구조를 컴퓨터 상에서 모의하도록 설계될 수 있으며, 인간의 신경망의 뉴런(neuron)을 모의하는, 가중치를 갖는 복수의 네트워크 노드들을 포함할 수 있다. 복수의 네트워크 노드들은 뉴런이 시냅스(synapse)를 통해 신호를 주고 받는 뉴런의 시냅틱 활동을 모의하도록 각각 연결 관계에 따라 데이터를 주고 받을 수 있다. 신경망은 신경망 모델에서 발전한 딥러닝 모델을 포함할 수 있다. 딥러닝 모델에서 복수의 네트워크 노드들은 서로 다른 레이어에 위치하면서 컨볼루션(convolution) 연결 관계에 따라 데이터를 주고 받을 수 있다. 신경망 모델의 예는 심층 신경망(DNN, deep neural networks), 합성곱 신경망(CNN, convolutional deep neural networks), 순환 신경망(RNN, Recurrent Boltzmann Machine), 제한 볼츠만 머신(RBM, Restricted Boltzmann Machine), 심층 신뢰 신경망(DBN, deep belief networks), 심층 Q-네트워크(Deep Q-Network)와 같은 다양한 딥러닝 기법들을 포함할 수 있다.

[0043] 이를 위해서, 프로세서(200)는 알고리즘 및 AI 프로세서를 저장하기 위한 메모리를 포함할 수 있다. 메모리는 하드 디스크 드라이브, 플래시 메모리, EEPROM(Electrically erasable programmable read-only memory), SRAM(Static RAM), FRAM (Ferro-electric RAM), PRAM (Phase-change RAM), MRAM(Magnetic RAM), DRAM(Dynamic Random Access Memory), SDRAM(Synchronous Dynamic Random Access Memory), DDR-SDRAM(Double Data Rate-SDRAM) 등을 이용할 수 있다.

[0044] 특히, 본 발명의 실시 예에 의한 프로세서(200)는 카메라(100)가 획득한 원본 이미지와 원본 이미지의 일부를 마스킹한 부분 이미지들을 학습한 결과가 유사하도록 객체 분류 모델을 학습함으로써, 깊이값을 보다 정확하게 획득할 수 있다. 이를 위해서, 프로세서(200)는 객체 분류 모델(MD)을 이용하여 이미지를 학습한 것을 바탕으로 제1 깊이값을 획득할 수 있다. 또한, 프로세서(200)는 이미지의 일부 영역을 마스킹한 부분 이미지를 획득하고, 부분 이미지를 학습한 것을 바탕으로 제2 깊이값을 획득할 수 있다. 또한, 프로세서(200)는 제1 깊이값 및 제2 깊이값의 편차를 줄이도록 객체 분류 모델을 학습할 수 있다.

[0045] 또한, 프로세서(200)는 검출된 객체에 따라 결정되는 시나리오를 바탕으로, 주행 제어기(300)를 제어할 수 있다. 예를 들어, 프로세서(200)는 검출된 객체가 차량의 주행에 방해가 되는 장애물일 경우, 장애물을 회피하도록 주행을 제어할 수 있다.

[0046] 주행 제어기(300)는 프로세서(200)로부터의 제어 신호에 응답하여 차량의 주행을 제어하기 위한 것으로, 조향 제어기, 엔진 제어기, 제동 제어기 및 변속 제어기(Transmission Control Module)를 포함할 수 있다. 주행 제어기(300)는 자동차 공학회에서 규정한 자율주행 레벨에 따라 주행하는 차량에 장착된 장치를 한정하는 것은 아니며, 프로세서(200)의 제어에 따라 사용자 편의성을 높이는 운전 보조 장치를 통칭하는 의미일 수 있다.

[0047] 조향 제어기는 유압펌프에 의해 형성된 유압을 이용하여 조향을 제어하는 유압식 조향(HPS; Hydraulic Power Steering) 시스템 및 전동모터의 출력 토크를 이용하여 조향을 제어하는 전동식 조향(MDPS; Motor Driven Power Steering System, 이하 'MDPS'라 함)으로 구분될 수 있다.

[0048] 엔진 제어기는 차량의 엔진을 제어하는 액추에이터로, 차량의 가속을 제어한다. 엔진 제어기는 EMS(Engine Management System)으로 구현될 수 있다. 엔진 제어기는 가속 페달 위치 센서로부터 출력되는 가속 페달 위치 정보에 따라 엔진의 구동토크를 제어한다. 엔진 제어기는 자율주행시 프로세서(100)로부터 요청받은 차량의 주행속도를 추종하기 위해 엔진 출력을 제어한다.

[0049] 제동 제어기는 차량의 감속을 제어하는 액추에이터로, 전자식 주행 안정화 컨트롤(Electronic Stability Control, ESC)로 구현될 수 있다. 제동 제어기는 프로세서(200)로부터 요청받은 목표 속도를 추종하기 위해 제동 압력을 제어한다. 즉, 제동 제어기는 차량의 감속을 제어한다.

[0050] 변속 제어기는 차량의 변속기를 제어하기 위한 액추에이터로, 전기식 시프터(Shift By Wire, SBW)로 구현될 수

있다. 변속 제어기는 기어 위치 및 기어 상태 범위에 따라 차량의 변속을 제어한다.

- [0051] 출력부(400)는 프로세서(200)의 제어에 따라 이미지를 바탕으로 검출된 객체 정보를 출력하기 위한 것으로, 디스플레이(410) 및 스피커(420)를 포함할 수 있다. 프로세서(200)는 디스플레이(410)를 통해서 검출된 객체를 시각적으로 표현할 수 있다. 또한, 프로세서(200)는 스피커(420)를 통해서 안전에 위협을 줄 수 있는 장애물을 검출한 경우 경고음을 출력할 수 있다.
- [0052] 프로세서(200)의 구조 및 동작에 대해서 보다 자세하게 설명하면 다음과 같다.
- [0053] 도 3은 본 발명의 실시 예에 의한 프로세서의 구성을 나타내는 블록도이고, 도 4는 본 발명의 실시 예에 의한 프로세서의 딥러닝(deep learning) 동작을 설명하기 위한 도면이다.
- [0054] 도 3 및 도 4를 참조하면, 본 발명의 실시 예에 의한 프로세서(200)는 패치 분할부(210), 마스킹부(220), 객체 분류 모델(MD)을 포함할 수 있다. 객체 분류 모델(MD)은 인코더(230), 결합부(240), 및 디코더(250)를 포함할 수 있다.
- [0055] 패치 분할부(210)는 이미지(IMG)를 패치 단위로 분할하여, 복수 개의 패치를 생성할 수 있다. 이미지(IMG)는 카메라(100)가 획득한 영상 데이터를 의미할 수 있다.
- [0056] 마스킹부(220)는 복수 개의 마스크들(M₁~M_k)을 이용하여 마스크 개수에 대응하는 부분 이미지들을 생성할 수 있다. 마스킹부(220)는 패치들의 일부를 제거하고, 샘플링 된 패치들을 남기는 방식으로 부분 이미지를 생성할 수 있다. 각각의 부분 이미지들에서 샘플링 된 패치들은 서로 중복되지 않을 수 있다.
- [0057] 인코더(230)는 패치 단위로 이미지를 제공받고, 패치들을 학습하여, 특징값을 추출할 수 있다. 인코더(230)는 이미지 전체의 패치들을 학습하여 제1 특징값(Ft1)을 추출할 수 있다. 또한, 인코더(230)는 마스킹부(220)가 생성한 부분 이미지들을 학습하여 부분 이미지들 각각에 대한 부분 특징값들을 생성할 수 있다.
- [0058] 결합부(240)는 부분 특징값들을 결합하여 제2 특징값(Ft2)을 생성할 수 있다.
- [0059] 프로세서(200)는 제1 손실함수(Lf)를 이용하여 객체 분류 모델(MD)을 학습할 수 있다. 제1 손실함수(Lf)는 제1 특징값(Ft1)과 제2 특징값(Ft2)의 편차에 따라 크기가 결정되는 함수일 수 있다. 프로세서(200)는 제1 손실함수(Lf)의 크기가 작아지도록 인코더(230)의 파라미터를 조절할 수 있다.
- [0060] 디코더(250)는 제1 특징값을 바탕으로 제1 깊이값(Dep1)을 추출할 수 있고, 제2 특징값(Ft2)을 바탕으로 제2 깊이값(Dep2)을 추출할 수 있다.
- [0061] 프로세서(200)는 제2 손실함수(Ld)를 이용하여 객체 분류 모델(MD)을 학습할 수 있다. 제2 손실함수(Ld)는 제1 깊이값(Dep1)과 제2 깊이값(Dep2)의 편차에 따라 크기가 결정되는 함수일 수 있다. 프로세서(200)는 제2 손실함수(Ld)의 크기가 작아지도록 인코더(230) 및 디코더(250)의 파라미터를 조절할 수 있다.
- [0062] 도 5는 본 발명의 실시 예에 의한 객체 검출 방법을 설명하기 위한 순서도이다. 도 5를 참조하여, 본 발명의 실시 예에 의한 객체 검출 방법을 살펴보면 다음과 같다.
- [0063] S310에서, 프로세서(200)는 객체 분류 모델(MD)을 이용하여 이미지를 학습한 것을 바탕으로 제1 깊이값을 획득할 수 있다.
- [0064] 이를 위해서, 객체 분류 모델(MD)의 인코더(230)는 패치 분할부(210)가 생성한 이미지 패치들을 제공받을 수 있다. 인코더(230)는 이미지 패치를 학습하여 제1 특징값(Ft1)을 추출할 수 있다. 또한, 디코더(250)는 제1 특징값(Ft1)을 바탕으로 제1 깊이값(Dep1)을 추출할 수 있다.
- [0065] 제1 깊이값(Dep1)은 이미지를 바탕으로 확장된 학습 데이터를 학습한 결과일 수 있다. 예를 들어, 프로세서(200)는 이미지의 차원 변환 또는 이미지의 계조 변환을 바탕으로 확장된 학습 데이터를 생성하고, 확장된 학습 데이터를 학습하여 제1 깊이값(Dep1)을 추출할 수 있다.
- [0066] S320에서, 프로세서(200)는 이미지의 일부 영역을 마스킹한 부분 이미지를 획득할 수 있다.
- [0067] 마스킹부(220)는 패치 분할부(210)가 생성한 이미지 패치에서 일부 패치를 샘플링하여 남기고, 샘플링 된 패치들 이외의 패치들을 제거하여 부분 이미지를 생성할 수 있다. 마스킹부(220)는 미리 설정된 복수 개의 마스크를 이용하여 부분 이미지를 생성할 수 있다.
- [0068] S330에서, 프로세서(200)는 부분 이미지를 학습한 것을 바탕으로 제2 깊이값(Dep2)을 획득할 수 있다.

- [0069] 인코더(230)는 마스크부(220)가 생성한 부분 이미지들을 학습하여 부분 특징값들을 생성할 수 있다. 결합부(240)는 부분 특징값들을 결합하여 제2 특징값(Ft_2)을 생성할 수 있다.
- [0070] 프로세서(200)는 제1 특징값(Ft_1)과 제2 특징값(Ft_2)의 편차를 줄이도록 인코더(230)의 파라미터를 수정할 수 있다. 이를 위해서 프로세서(200)는 제1 특징값(Ft_1)과 제2 특징값(Ft_2)의 편차에 비례하는 제2 손실함수를 이용할 수 있다. 프로세서(200)는 제1 손실함수(L_f)의 크기가 작아지도록 인코더(230)의 파라미터를 조절할 수 있다.
- [0071] S340에서, 프로세서(200)는 제1 깊이값(Dep_1) 및 제2 깊이값(Dep_2)의 편차를 줄이도록 객체 분류 모델(MD)을 학습할 수 있다.
- [0072] 이를 위해서, 제1 깊이값(Dep_1)과 제2 깊이값(Dep_2)의 편차에 비례하는 제2 손실함수(L_d)를 이용할 수 있다. 프로세서(200)는 제2 손실함수(L_d)의 크기가 작아지도록 인코더(230)의 파라미터 및 디코더(250)의 파라미터를 조절할 수 있다.
- [0073] 또한, 디코더(250)가 출력하는 제1 깊이값(Dep_1)은 실제 깊이값과의 편차를 줄이기 위한 제3 손실함수(L_{gt})를 이용하여 학습될 수 있다. 즉, 프로세서(200)는 제1 깊이값(Dep_1)과 실제 깊이값과의 편차에 비례하는 제3 손실함수(L_{gt})의 크기를 줄이도록 인코더(230)의 파라미터 및 디코더(250)의 파라미터를 조절할 수 있다.
- [0074] 이하, 각 절차들에 대한 세부적인 실시 예를 살펴보면 다음과 같다.
- [0075] 도 6은 이미지 패치들의 실시 예를 나타내는 도면이다.
- [0076] 도 6을 참조하면, 패치 분할부(210)는 이미지를 복수 개의 패치로 분할할 수 있다. 도 6은 이미지를 로우(row) 방향과 컬럼(column) 방향으로 각각 5개씩으로 분할하여, 25개의 패치를 생성한 실시 예를 도시하고 있다.
- [0077] 패치들의 개수는 이미지의 크기를 패치의 크기(p)로 나눈 값이 될 수 있다. 예를 들어, 이미지는 (C, H, W)와 같이 채널(C), 높이(H), 폭(W)으로 표현될 수 있다. 그리고, 이미지가 (3, 256, 256)으로 표현될 경우, 패치 크기를 16으로 설정하여 이미지를 분할하면, 이미지 패치들 각각의 크기는 (3, 16, 16)이 될 수 있고, 이미지 패치의 개수는 16×16 개일 수 있다.
- [0078] 도 7은 이미지 패치들을 시퀀스 데이터로 인코더에 제공하는 실시 예를 설명하기 위한 도면이다. 도 7은 제1 특징값을 획득하기 위한 입력값을 인코더에 제공하는 실시 예를 설명하는 도면일 수 있다.
- [0079] 도 7을 참조하면, 프로세서(200)는 이미지 패치들을 시퀀스 데이터로 인코더(230)에 제공할 수 있다.
- [0080] 예를 들어, 프로세서(200)는 이미지의 좌측 상단에서부터 우측 하단으로의 패치들을 순서대로 정렬하여 시퀀스 데이터를 생성할 수 있다. 예를 들어, 프로세서(200)는 첫 번째 로우의 패치들을 (R_1, C_1), (R_1, C_2), (R_1, C_3), (R_1, C_4), (R_1, C_5)로 정렬할 수 있다. 그리고 두 번째 로우의 패치들을 컬럼 순서대로 정렬하며, 이와 같은 방식으로 마지막 로우의 패치들을 (R_5, C_1), (R_5, C_2), (R_5, C_3), (R_5, C_4), (R_5, C_5)로 정렬할 수 있다.
- [0081] 이어서, 프로세서(200)는 각 패치들을 평탄화(flatten)하여 벡터로 변환할 수 있다.
- [0082] 채널이 3개이고, 패치의 개수가 16×16 개일 경우, 벡터의 크기는 $3 \times 16 \times 16$ 이며, 벡터의 개수는 16×16 개가 될 수 있다. 그리고, 시퀀스 데이터는 (256, 768)로 표현될 수 있다.
- [0083] 프로세서(200)는 각각의 벡터들에 선형(linear) 연산을 수행하여, 임베딩(embedding) 절차를 진행할 수 있다. 그리고, 프로세서(200)는 임베딩 절차의 결과에 클래스(class)를 예측하기 위한 클래스 토큰을 추가할 수 있다.
- [0084] 또한, 프로세서(200)는 각 패치들의 위치 정보를 보존하기 위해서, 인코더(230)의 입력값에 포지셔널 임베딩(positional embedding)을 추가할 수 있다.
- [0085] 인코더(230)는 이미지의 모든 패치들을 바탕으로 생성된 시퀀스 데이터를 바탕으로, 제1 특징값을 추출할 수 있다.
- [0086] 또한, 제1 특징값을 추출하기 위한 이미지(IMG)는 데이터 증강 기법(Data Augmentatio)을 통해서 확장될 수 있다. 즉, 객체 분류 모델(MD)은 데이터 증강 기법을 통해서 다양하게 변화된 학습 데이터를 학습하여 객체 검출 성능을 높일 수 있다. 데이터 증강 기법을 도 8을 참조하여 설명하면 다음과 같다.
- [0087] 도 8은 제1 특징값을 추출하는 과정에서 활용되는 데이터 증강 기법을 설명하기 위한 도면이다.
- [0088] 도 8에서와 같이, 프로세서(200)는 이미지(IMG)를 반전하는 플리핑(Flipping) 기법을 이용하여 학습 데이터를

확장할 수 있다.

- [0089] 또한, 프로세서(200)는 이미지(IMG)를 z축 방향으로 회전(rotation)하여 학습 데이터를 확장할 수 있다. z축 방향은 이미지의 2차원 평면과 수직인 축을 의미할 수 있다.
- [0090] 또한, 프로세서(200)는 이미지(IMG)의 각 픽셀들을 동일한 위치만큼 이동시키는 쉬프트(shifting) 기법으로 학습 데이터를 확장할 수 있다.
- [0091] 도 8에 도시된 증강 기법을 이용하여 생성된 학습 데이터는 이미지(IMG)의 특징을 유지할 수 있다.
- [0092] 도 8에 도시된 방식 이외에도, 프로세서(200)는 이미지의 전체 영역에 화이트 노이즈를 랜덤으로 추가하거나, 블랙 노이즈를 랜덤으로 추가하여 강화된 학습 데이터를 생성할 수 있다. 이와 같이, 노이즈의 추가를 통해서 생성된 학습 데이터 또한 이미지(IMG)의 특징을 유지할 수 있다.
- [0093] 객체 분류 모델(MD)은 이미지(IMG)의 특징을 유지하는 확장된 학습 데이터를 이용하여 제1 특징값(Ft_1)의 정확도를 높일 수 있다.
- [0094] 도 9는 이미지 패치들을 시퀀스 데이터로 인코더에 제공하는 다른 실시 예를 설명하기 위한 도면이다. 도 9는 제2 특징값을 획득하기 위한 입력값을 인코더에 제공하는 실시 예를 설명하는 도면일 수 있다.
- [0095] 도 9를 참조하면, 프로세서(200)는 부분 이미지의 패치들을 시퀀스 데이터로 인코더(230)에 제공할 수 있다.
- [0096] 도 9에 도시된 부분 이미지는 (R1, C4), (R3, C1), (R3, C3), (R4, C4), (R5, C2)의 패치들이 샘플링 되고, 샘플링 된 패치들 이외의 패치들은 제거된 것을 도시하고 있다.
- [0097] 프로세서(200)는 부분 이미지의 좌측 상단에서부터 우측 하단으로의 패치들을 순서대로 정렬하여 시퀀스 데이터를 생성할 수 있다. 예를 들어, 프로세서(200)는 (R1, C4), (R3, C1), (R3, C3), (R4, C4), (R5, C2) 순서대로 샘플링 된 패치들을 정렬할 수 있다.
- [0098] 이어서, 프로세서(200)는 샘플링 된 패치들을 평탄화(flatten)하여 벡터로 변환할 수 있다.
- [0099] 프로세서(200)는 각각의 벡터들에 선형(linear) 연산을 수행하여, 임베딩(embedding) 절차를 진행할 수 있다.
- [0100] 또한, 프로세서(200)는 각 패치들의 위치 정보를 보존하기 위해서, 인코더(230)의 입력값에 포지셔널 임베딩(positional embedding)을 추가할 수 있다.
- [0101] 도 10은 인코더를 이용하여 제2 특징값을 추출하는 과정을 설명하기 위한 도면이다.
- [0102] 도 10을 참조하면, 트랜스포머의 인코더(230)는 부분 이미지들($Ip_1 \sim Ip_k$)을 학습하여 부분 특징값들을 생성할 수 있다. 부분 이미지들($Ip_1 \sim Ip_k$)은 마스크부(220)에서 마스크들($M_1 \sim M_k$)을 이용하여 생성될 수 있다.
- [0103] 마스크부(220)는 제1 마스크(M_1) 내지 제k(k는 자연수) 마스크들($M_1 \sim M_k$)을 이용하여 제1 부분 이미지(Ip_1) 내지 제k 부분 이미지(Ip_k)들을 생성할 수 있다.
- [0104] 이미지가 n개의 패치들을 포함하고, 제1 내지 제k 마스크들($M_1 \sim M_k$) 각각이 i(i는 자연수)개의 패치들을 샘플링 할 경우, 마스크의 개수(k)는 " n/i "가 될 수 있다.
- [0105] 그리고, 각각의 부분 이미지들($Ip_1 \sim Ip_k$)에서 샘플링 된 패치는 서로 중복되지 않을 수 있다. 예를 들어, 제1 부분 이미지(Ip_1)에서 샘플링 된 (R1, C4) 위치의 패치는 다른 부분 이미지들에 포함되지 않을 수 있다.
- [0106] 각각의 부분 이미지들($Ip_1 \sim Ip_k$)이 포함하는 패치들은 전술한 바와 같이 임베딩 된 위치 정보들을 바탕으로 인코더(230)에 제공될 수 있다.
- [0107] 인코더(230)는 제1 부분 이미지(Ip_1)를 학습하여 제1 부분 특징값(Ftp_1)을 추출할 수 있고, 동일한 방식을 반복하여 제k 부분 이미지(Ip_k)를 학습하여 제k 부분 특징값(Ftp_k)을 추출할 수 있다.
- [0108] 각각의 부분 특징값들($Ftp_1 \sim Ftp_k$)은 이미지 전체의 정보를 추정하는 것이 어려울 수 있다. 따라서, 각각의 부분 특징값($Ftp_1 \sim Ftp_k$)들을 바탕으로 이미지 전체의 특징값을 추정하기 위한 절차가 추가될 수 있다. 이를 위해서, 제1 내지 제k 부분 특징값들($Ftp_1 \sim Ftp_k$)은 결합부(240)에 제공될 수 있다.
- [0109] 결합부(240)는 제1 내지 제k 부분 특징값들($Ftp_1 \sim Ftp_k$)의 위치 정보를 이용하여 부분 특징값들을 결합하여 제2 특징값을 생성할 수 있다.
- [0110] 결과적으로, 객체 분류 모델(MD)은 부분 이미지들($Ip_1 \sim Ip_k$)을 학습하여, 이미지 전체에 대한 제2 특징값을 획득

득할 수 있다.

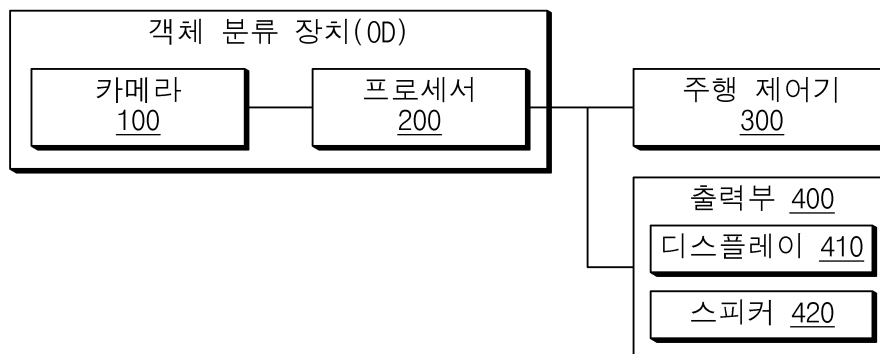
- [0111] 도 11은 제1 손실함수를 바탕으로 인코더를 학습하는 방식을 설명하기 위한 도면이고, 도 12는 제2 손실함수를 바탕으로 객체 분류 모델을 학습하는 방식을 설명하기 위한 도면이다.
- [0112] 도 11을 참조하면, 본 발명의 실시 예에 의한 인코더(230)는 이미지(IMG)의 패치들을 학습하여 제1 특징값(Ft1)을 추출할 수 있다.
- [0113] 또한, 인코더(230)는 이미지(IMG)의 일부가 삭제된 부분 이미지들(Ip_1~Ip_k)을 학습하여 부분 특징값들(Ftp_1~Ftp_k)을 추출할 수 있고, 결합부(240)는 부분 특징값들(Ftp_1~ Ftp_k)을 결합하여 제2 특징값(Ft2)을 획득할 수 있다.
- [0114] 프로세서(200)는 제1 손실함수(Lf)를 이용하여 인코더(230)의 파라미터를 조절할 수 있다. 제1 손실함수(Lf)는 제1 특징값(Ft1)과 제2 특징값(Ft2)의 편차에 비례하여 큰 값을 도출할 수 있다. 프로세서(200)는 제1 특징값(Ft1) 및 제2 특징값(Ft2)을 바탕으로 제1 손실함수(Lf)를 계산할 수 있다. 그리고, 프로세서(200)는 제1 손실함수(Lf)의 크기를 줄이도록 인코더(230)의 파라미터들을 조절할 수 있다. 인코더(230)의 파라미터들은 이미지 또는 부분 이미지를 바탕으로 도출되는 특징값을 결정하기 위한 인자들일 수 있다.
- [0115] 제1 특징값(Ft1)은 이미지 전체를 학습한 결과일 수 있고, 이미지 증강 기법을 통해서 이미지의 특징을 유지한 학습 데이터를 추가적으로 학습한 결과일 수 있다.
- [0116] 제2 특징값(Ft2)은 이미지의 일부가 생략된 부분 이미지를 학습한 결과들을 결합한 것일 수 있다. 따라서, 제2 특징값(Ft2)은 이미지의 특징이 누락된 부분 이미지를 학습한 결과일 수 있다.
- [0117] 즉, 본 발명의 실시 예에 의한 객체 분류 모델(MD)은 이미지의 특징을 유지한 학습 데이터로부터 도출된 제1 특징값(Ft1)과 이미지의 특징이 누락될 수 있는 부분 이미지로부터 도출된 제2 특징값(Ft2)을 일치시키는 방향으로 학습을 진행할 수 있다. 따라서, 본 발명의 실시 예에 의한 객체 분류 모델(MD)은 이미지의 일부가 누락되어 객체의 특징이 누락된 데이터를 이용하여도 이미지 학습 성능을 높일 수 있고, 이미지의 깊이값을 보다 정확하게 도출시킬 수 있는 특징값을 생성할 수 있다.
- [0118] 도 12를 참조하면, 본 발명의 실시 예에 의한 객체 분류 모델(MD)은 이미지(IMG)의 패치들을 학습하여 제1 깊이값(Dep1)을 추출할 수 있다.
- [0119] 또한, 객체 분류 모델(MD)은 이미지(IMG)의 일부가 삭제된 부분 이미지들(Ip_1~Ip_k)을 학습하여 제2 깊이값(Dep2)을 획득할 수 있다.
- [0120] 프로세서(200)는 제2 손실함수(Ld)를 이용하여 객체 분류 모델(MD)의 파라미터들을 조절할 수 있다. 예를 들어, 프로세서(200)는 제2 손실함수(Ld)의 크기를 줄이도록 인코더(230) 및 디코더(250)의 파라미터들을 조절할 수 있다. 인코더(230) 및 디코더(250)의 파라미터들은 이미지 또는 부분 이미지를 바탕으로 도출되는 깊이값을 결정하기 위한 인자들일 수 있다.
- [0121] 제1 깊이값(Dep1)은 이미지 전체를 학습한 결과일 수 있고, 이미지 증강 기법을 통해서 이미지의 특징을 유지한 학습 데이터를 추가적으로 학습한 결과일 수 있다. 제2 깊이값(Dep2)은 이미지의 일부가 생략된 부분 이미지를 학습한 결과들을 결합한 것일 수 있다. 따라서, 제2 깊이값(Dep2)은 이미지의 특징이 누락된 부분 이미지를 학습한 결과일 수 있다.
- [0122] 즉, 본 발명의 실시 예에 의한 객체 분류 모델(MD)은 이미지의 특징을 유지한 학습 데이터로부터 도출된 제1 깊이값(Dep1)과 이미지의 특징이 누락될 수 있는 부분 이미지로부터 도출된 제2 깊이값(Dep2)을 일치시키는 방향으로 학습을 진행할 수 있다. 따라서, 본 발명의 실시 예에 의한 객체 분류 모델(MD)은 이미지의 일부가 누락되어 객체의 특징이 누락된 데이터를 이용하여도 이미지 학습 성능을 높일 수 있고, 이미지의 깊이값을 보다 정확하게 추출할 수 있다.
- [0123] 도 13은 본 발명의 다른 실시 예에 의한 프로세서의 딥러닝 동작을 설명하기 위한 도면이다.
- [0124] 도 13을 참조하면, 본 발명의 실시 예에 의한 프로세서(200)는 패치 분할부(210), 마스크부(220), 객체 분류 모델(MD)을 포함할 수 있다. 객체 분류 모델(MD)은 비전 트랜스포머(Vision Transformer)를 이용할 수 있고, 인코더(230), 결합부(240), 및 디코더(250)를 포함할 수 있다. 도 13의 실시 예에서 전술한 실시 예와 실질적으로 동일한 구성에 대해서는 자세한 설명을 생략하기로 한다.

- [0125] 프로세서(200)는 제1 손실함수(Lf)를 이용하여 객체 분류 모델(MD)을 학습할 수 있다. 제1 손실함수(Lf)는 제1 특징값과 제2 특징값의 편차에 따라 크기가 결정되는 함수일 수 있다. 프로세서(200)는 제1 손실함수(Lf)의 크기가 작아지도록 인코더(230)의 파라미터를 조절할 수 있다.
- [0126] 디코더(250)는 제1 특징값을 바탕으로 제1 깊이값을 추출할 수 있고, 제2 특징값을 바탕으로 제2 깊이값을 추출할 수 있다.
- [0127] 프로세서(200)는 제2 손실함수(Ld)를 이용하여 객체 분류 모델(MD)을 학습할 수 있다. 제2 손실함수(Ld)는 제1 깊이값과 제2 깊이값의 편차에 따라 크기가 결정되는 함수일 수 있다. 프로세서(200)는 제2 손실함수(Ld)의 크기가 작아지도록 인코더(230) 및 디코더(250)의 파라미터를 조절할 수 있다.
- [0128] 신뢰도 판단부(260)는 제2 손실함수(Ld)를 줄이도록 학습된 객체 분류 모델(MD)의 불확실 정도(U)를 판단할 수 있고, 불확실 정도(U)에 따라 제2 손실함수(Ld)를 조절할 수 있다. 예를 들어, 신뢰도 판단부(260)는 픽셀 단위 또는 패치 단위로 불확실 정도(U)를 판단할 수 있고, 불확실 정도(U)를 반영하여 제2 손실함수(Ld)에서 임의의 픽셀 또는 임의의 패치에 대한 가중치를 조절할 수 있다.
- [0129] 신뢰도 판단부(260)는 불확실 정도에 따라 제2 손실함수(Ld)를 조절함으로써, 제2 손실함수(Ld)가 감소하였음에도 불구하고 추출된 깊이값이 줄어드는 현상을 줄일 수 있다.
- [0130] 도 14는 본 발명의 다른 실시 예에 의한 딥러닝 동작을 설명하기 위한 도면이다.
- [0131] 도 14를 참조하면, 본 발명의 실시 예에 의한 프로세서(200)는 패치 분할부(210), 마스킹부(220), 객체 분류 모델(MD)을 포함할 수 있다. 객체 분류 모델(MD)은 비전 트랜스포머(Vision Transformer)를 이용할 수 있고, 인코더(230), 결합부(240), 및 제1 디코더(250), 및 제2 디코더(152)를 포함할 수 있다. 도 14의 실시 예에서 전술한 실시 예와 실질적으로 동일한 구성에 대해서는 자세한 설명을 생략하기로 한다.
- [0132] 프로세서(200)는 제1 손실함수(Lf)를 이용하여 객체 분류 모델(MD)을 학습할 수 있다. 제1 손실함수(Lf)는 제1 특징값과 제2 특징값의 편차에 따라 크기가 결정되는 함수일 수 있다. 프로세서(200)는 제1 손실함수(Lf)의 크기가 작아지도록 인코더(230)의 파라미터를 조절할 수 있다.
- [0133] 제1 디코더(250)는 제1 특징값을 바탕으로 제1 깊이값(Dep1)을 추출할 수 있고, 제2 특징값을 바탕으로 제2 깊이값(Dep2)을 추출할 수 있다. 프로세서(200)는 제2 손실함수(Ld)를 이용하여 객체 분류 모델(MD)을 학습할 수 있다. 제2 손실함수(Ld)는 제1 깊이값과 제2 깊이값(Dep2)의 편차에 따라 크기가 결정되는 함수일 수 있다. 프로세서(200)는 제2 손실함수(Ld)의 크기가 작아지도록 인코더(230) 및 디코더(250)의 파라미터를 조절할 수 있다.
- [0134] 제2 디코더(152)는 제1 특징값을 바탕으로 제1 클래스(Cls1)를 추출할 수 있고, 제2 특징값을 바탕으로 제2 클래스(Cls2)를 추출할 수 있다. 프로세서(200)는 제4 손실함수(Lc)를 이용하여 객체 분류 모델(MD)을 학습할 수 있다. 제4 손실함수(Lc)는 제1 클래스(Cls1)와 제2 클래스(Cls2)의 편차에 따라 크기가 결정되는 함수일 수 있다. 프로세서(200)는 제4 손실함수(Lc)의 크기가 작아지도록 인코더(230) 및 디코더(250)의 파라미터를 조절할 수 있다.
- [0135] 제1 클래스(Cls1)는 이미지 전체를 학습한 결과일 수 있고, 이미지 증강 기법을 통해서 이미지의 특징을 유지한 학습 데이터를 추가적으로 학습한 결과일 수 있다. 제2 클래스(Cls2)는 이미지의 일부가 생략된 부분 이미지를 학습한 결과들을 결합한 것일 수 있다. 따라서, 제2 클래스(Cls2)는 이미지의 특징이 누락된 부분 이미지를 학습한 결과일 수 있다.
- [0136] 즉, 본 발명의 실시 예에 의한 객체 분류 모델(MD)은 이미지의 특징을 유지한 학습 데이터로부터 도출된 제1 클래스(Cls1)와 이미지의 특징이 누락될 수 있는 부분 이미지로부터 도출된 제2 클래스(Cls2)를 일치시키는 방향으로 학습을 진행할 수 있다. 따라서, 본 발명의 실시 예에 의한 객체 분류 모델(MD)은 이미지의 일부가 누락되어 객체의 특징이 누락된 데이터를 이용하여도 이미지 학습 성능을 높일 수 있고, 이미지의 클래스를 보다 정확하게 추출할 수 있다.
- [0137] 도 15는 본 발명의 일 실시 예에 따른 컴퓨팅 시스템을 도시한다.
- [0138] 도 15를 참조하면, 컴퓨팅 시스템(1000)은 버스(1200)를 통해 연결되는 적어도 하나의 프로세서(1100), 메모리(1300), 사용자 인터페이스 입력 장치(1400), 사용자 인터페이스 출력 장치(1500), 스토리지(1600), 및 네트워크 인터페이스(1700)를 포함할 수 있다.

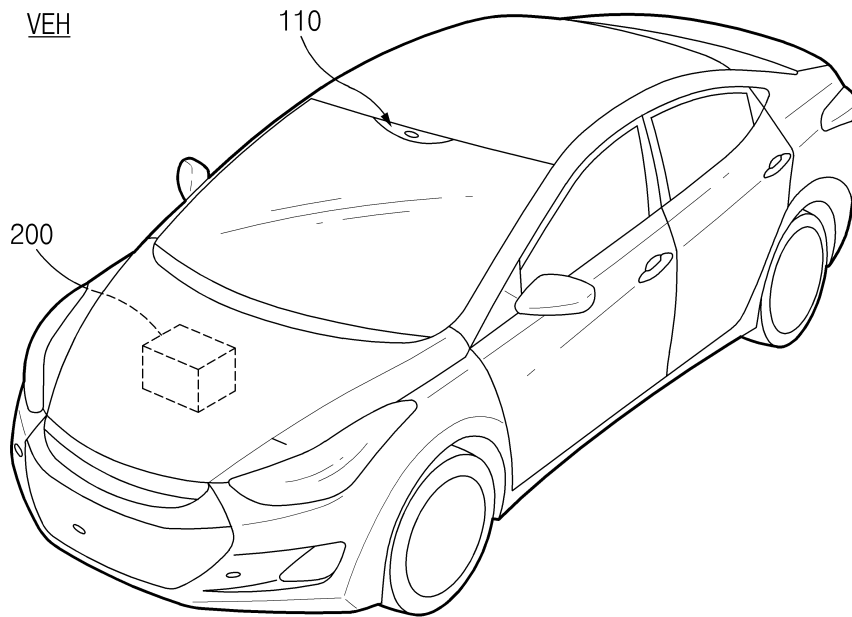
- [0139] 프로세서(1100)는 중앙 처리 장치(CPU) 또는 메모리(1300) 및/또는 스토리지(1600)에 저장된 명령어들에 대한 처리를 실행하는 반도체 장치일 수 있다. 메모리(1300) 및 스토리지(1600)는 다양한 종류의 휘발성 또는 불휘발성 저장 매체를 포함할 수 있다. 예를 들어, 메모리(1300)는 ROM(Read Only Memory) 및 RAM(Random Access Memory)을 포함할 수 있다.
- [0140] 따라서, 본 명세서에 개시된 실시 예들과 관련하여 설명된 방법 또는 알고리즘의 단계는 프로세서(1100)에 의해 실행되는 하드웨어, 소프트웨어 모듈, 또는 그 2 개의 결합으로 직접 구현될 수 있다. 소프트웨어 모듈은 RAM 메모리, 플래시 메모리, ROM 메모리, EPROM 메모리, EEPROM 메모리, 레지스터, 하드 디스크, 착탈형 디스크, CD-ROM과 같은 저장 매체(즉, 메모리(1300) 및/또는 스토리지(1600))에 상주할 수도 있다.
- [0141] 예시적인 저장 매체는 프로세서(1100)에 커플링되며, 그 프로세서(1100)는 저장 매체로부터 정보를 판독할 수 있고 저장 매체에 정보를 기입할 수 있다. 다른 방법으로, 저장 매체는 프로세서(1100)와 일체형일 수도 있다. 프로세서 및 저장 매체는 주문형 집적회로(ASIC) 내에 상주할 수도 있다. ASIC는 사용자 단말기 내에 상주할 수도 있다. 다른 방법으로, 프로세서 및 저장 매체는 사용자 단말기 내에 개별 컴포넌트로서 상주할 수도 있다.
- [0142] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다.
- [0143] 따라서, 본 발명에 개시된 실시 예들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시 예에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

도면

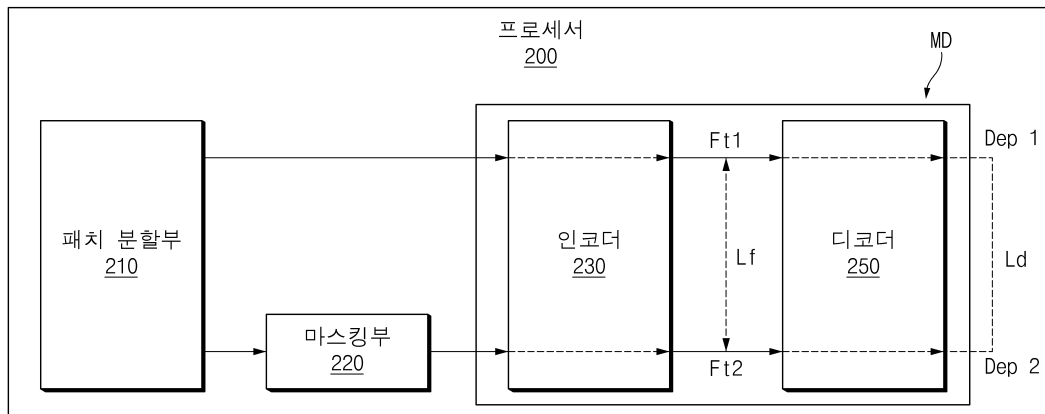
도면1



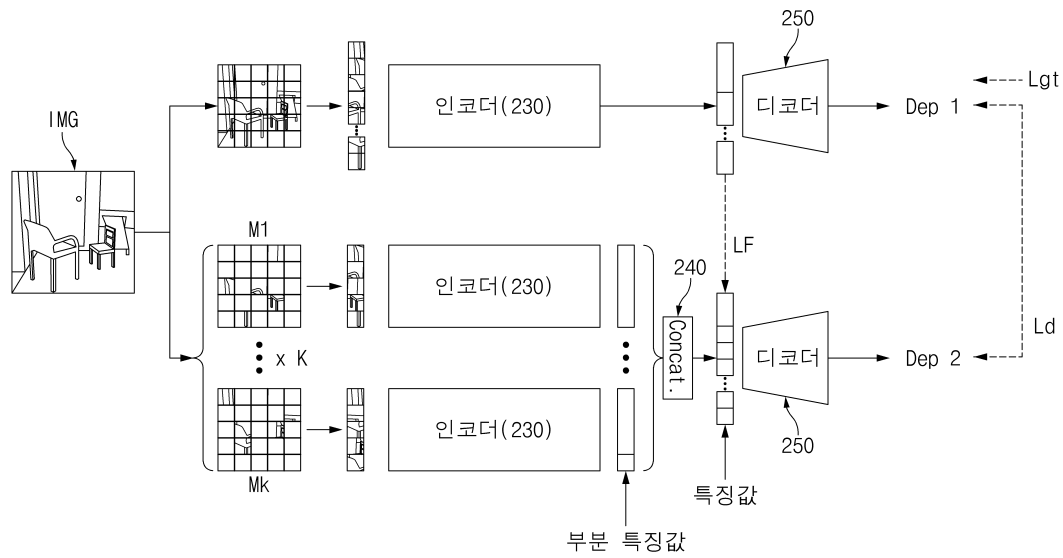
도면2



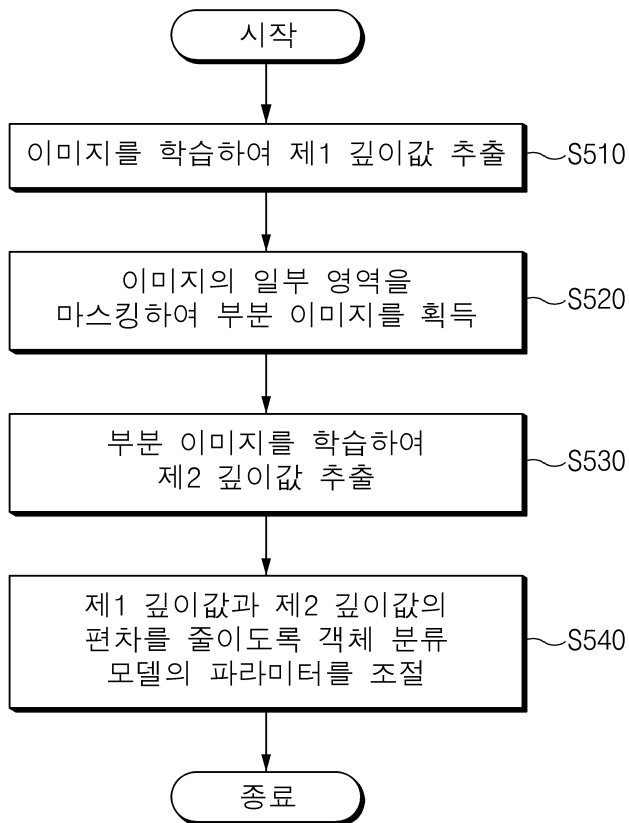
도면3



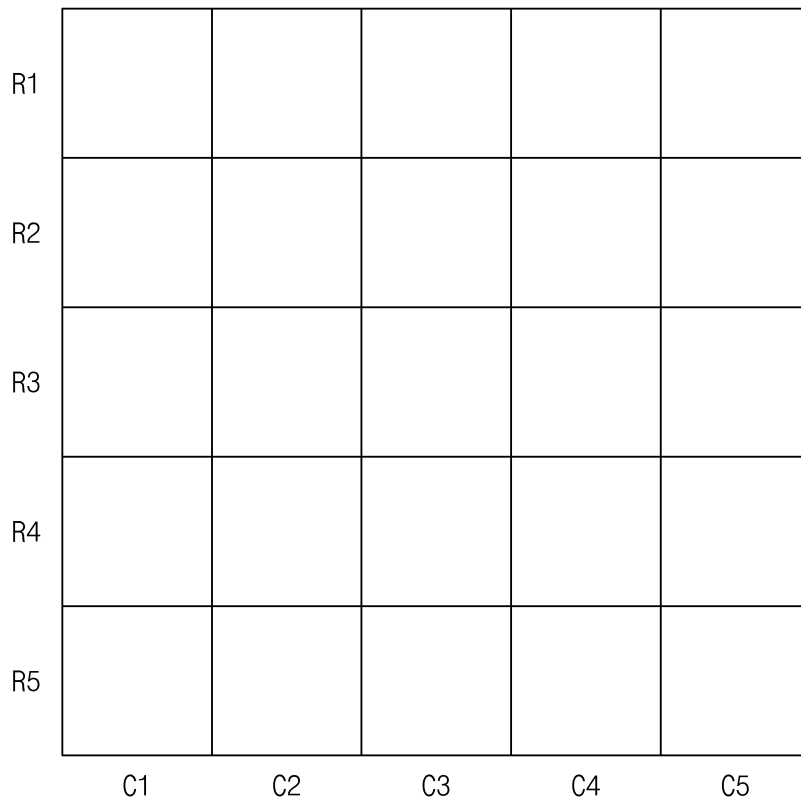
도면4



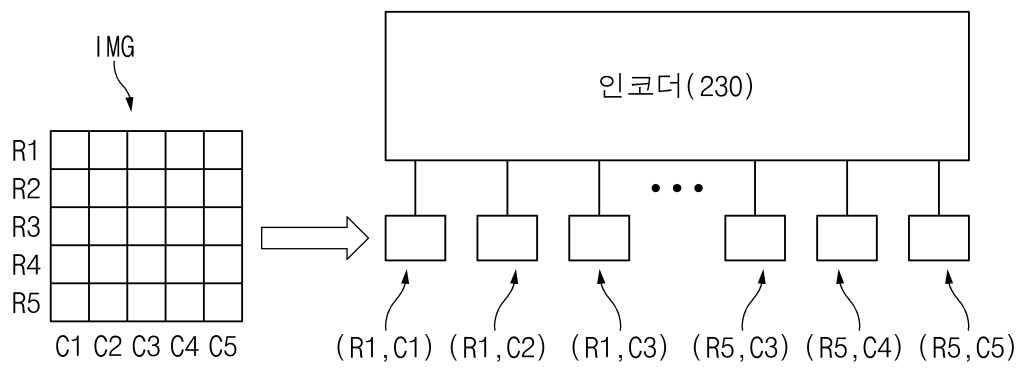
도면5



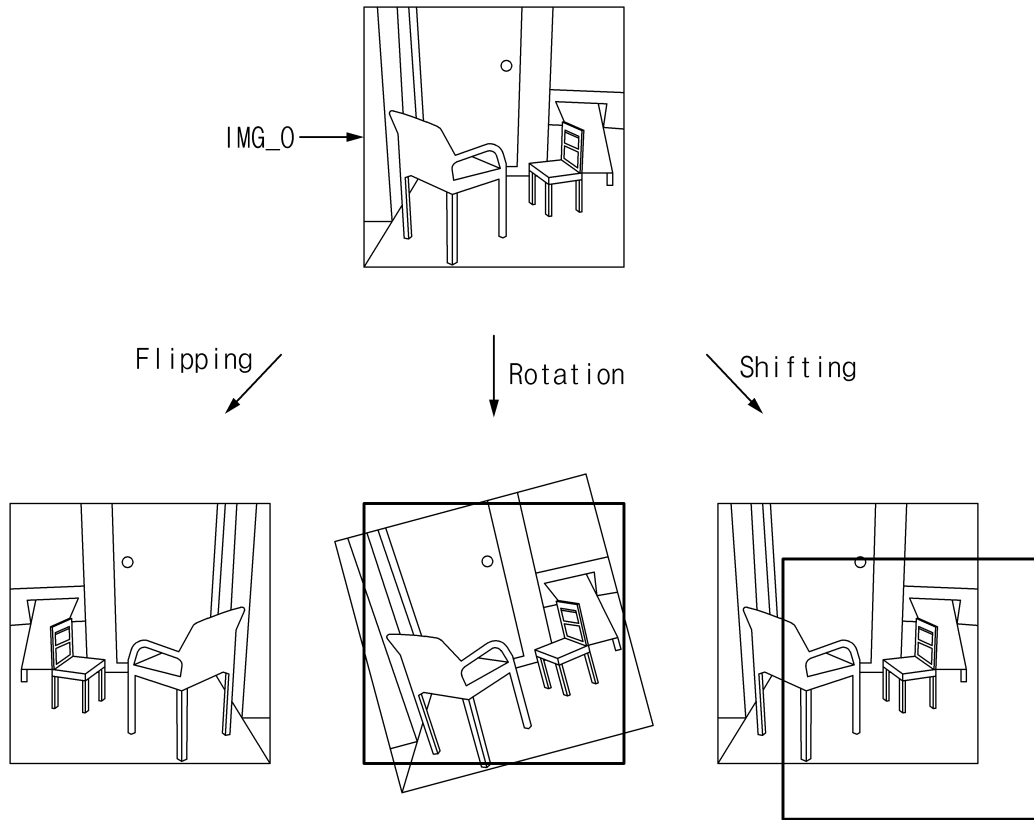
도면6



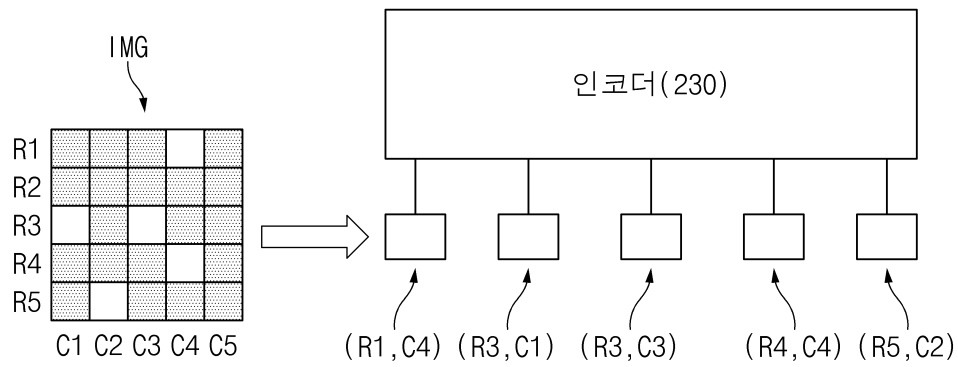
도면7



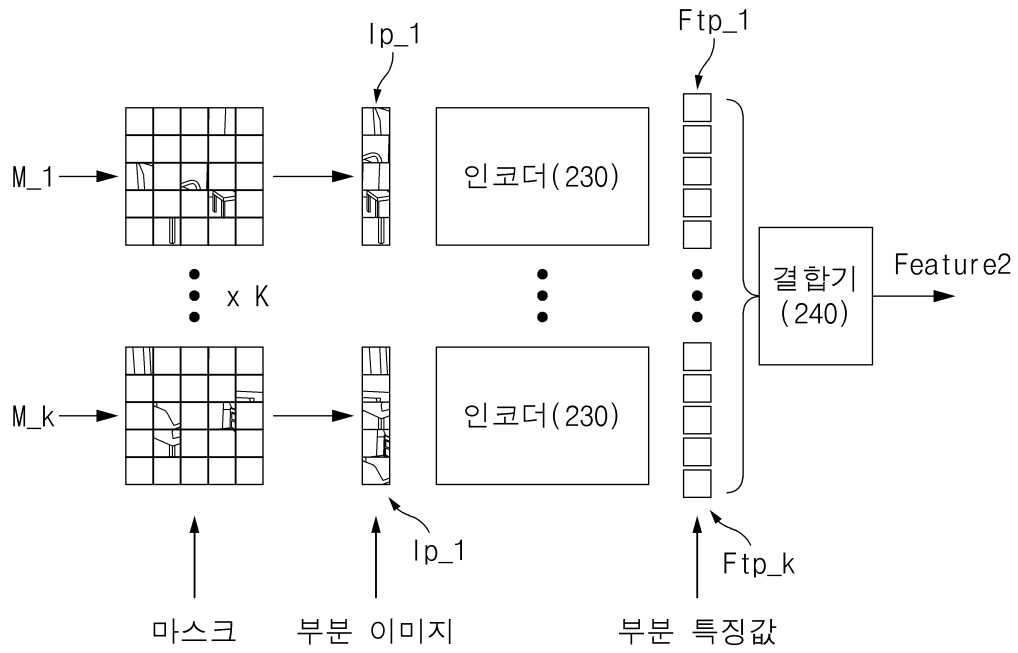
도면8



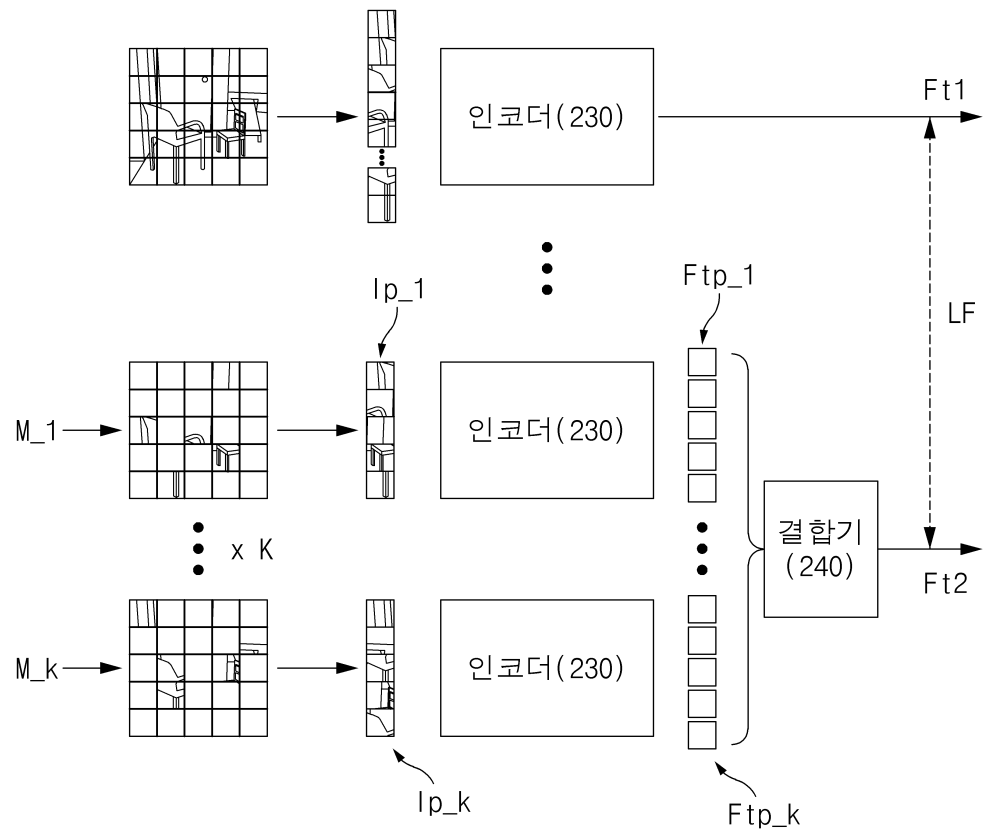
도면9



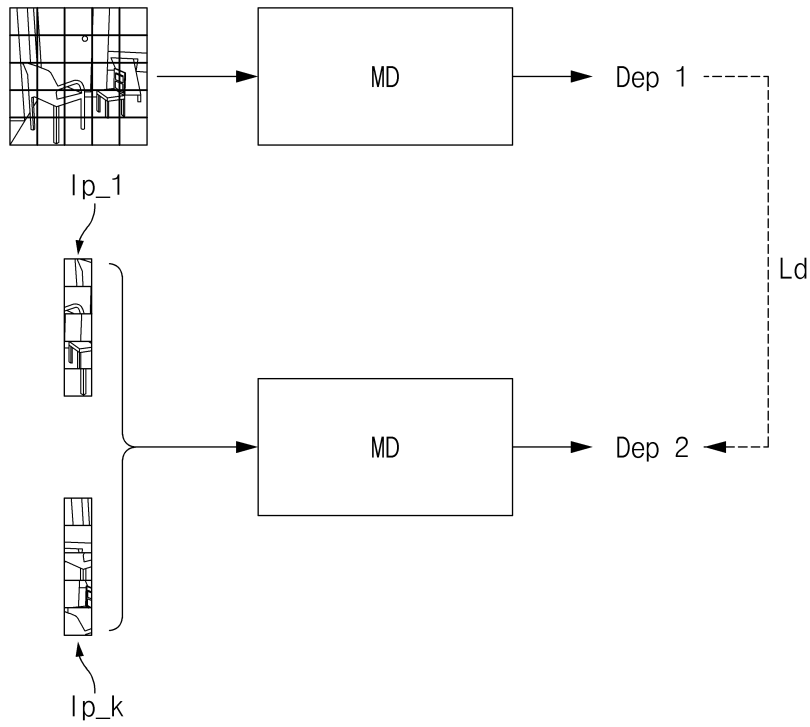
도면10



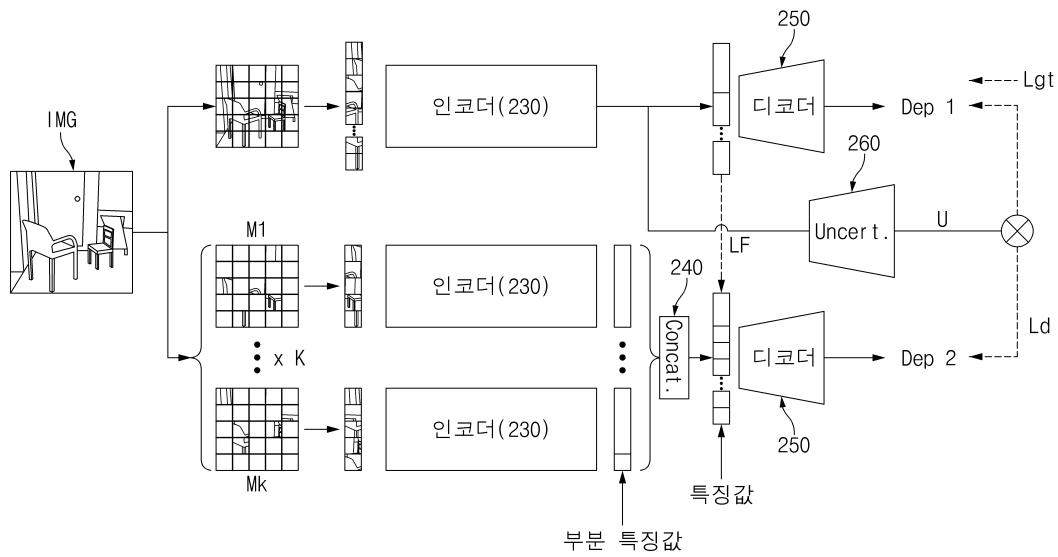
도면11



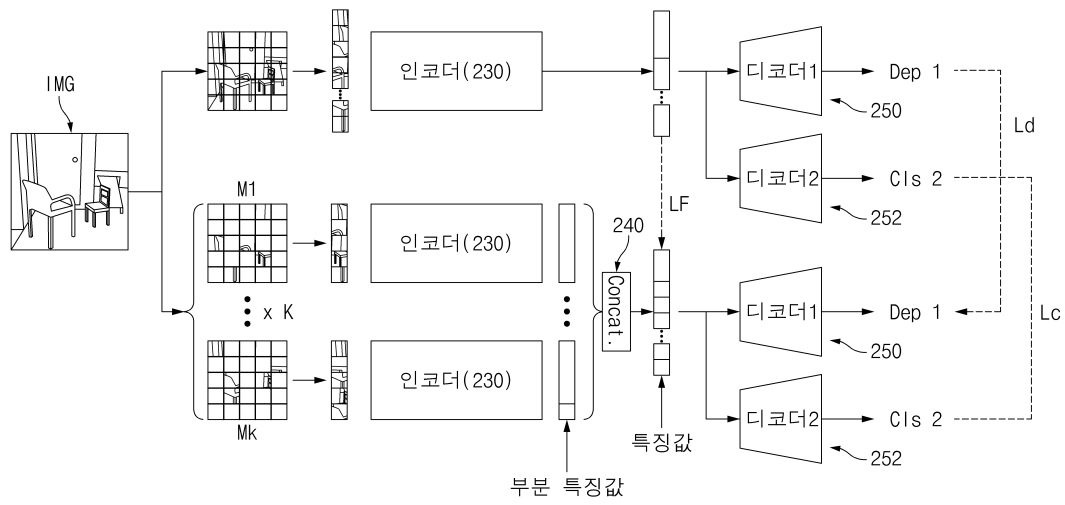
도면12



도면13



도면14



도면15

