



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0050885
(43) 공개일자 2021년05월10일

(51) 국제특허분류(Int. Cl.)
G06T 7/174 (2017.01) G06K 9/00 (2006.01)
G06T 7/11 (2017.01) G06T 7/20 (2017.01)
G06T 9/00 (2019.01)
(52) CPC특허분류
G06T 7/174 (2017.01)
G06K 9/00335 (2013.01)
(21) 출원번호 10-2019-0135439
(22) 출원일자 2019년10월29일
심사청구일자 2019년10월29일

(71) 출원인
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
이상윤
서울특별시 서초구 서초중앙로24길 43, 101동 1103호(서초동, 유원서초아파트)
김우진
서울특별시 마포구 백범로1길 83, 505호(노고산동, 신촌다울노블리움)
(뒷면에 계속)
(74) 대리인
민영준

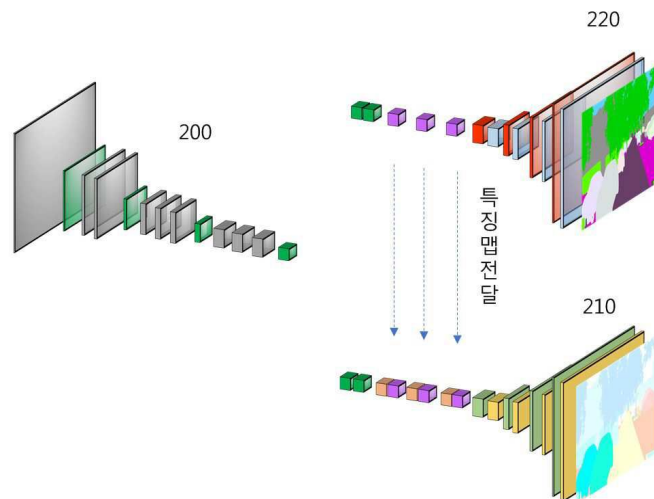
전체 청구항 수 : 총 17 항

(54) 발명의 명칭 동적 객체 검출 장치 및 방법

(57) 요약

동적 객체 검출 장치 및 방법이 개시된다. 개시된 장치는, 두 개의 프레임 이미지를 입력받아 상기 두 개의 프레임 이미지에 포함된 객체별 광학 플로우를 연산하도록 학습되어 있는 객체별 광학 플로우 네트워크 모듈; 상기 두 개의 프레임 이미지를 입력받아 상기 두 개의 프레임 이미지에 포함된 객체들의 객체 영역을 검출하도록 학습되어 있는 객체 검출 네트워크 모듈; 및 상기 객체별 광학 플로우 네트워크 모듈 및 상기 객체 검출 네트워크 모듈의 출력을 이용하여 상기 두 개의 프레임 이미지의 동적 객체를 검출하는 동적 객체 검출 모듈을 포함하되, 상기 두 개의 프레임 이미지는 동일한 카메라에 의해 촬영되고 시간적 갭이 있는 이미지이다. 개시된 장치 및 방법에 의하면, 실시간으로 동적 객체를 검출할 수 있으며, 조도 변화 및 환경 변화에 강인하게 동적 객체를 검출할 수 있는 장점이 있다.

대표도 - 도5



(52) CPC특허분류

G06T 7/11 (2017.01)

G06T 7/20 (2013.01)

G06T 9/00 (2019.01)

G06T 2207/20081 (2013.01)

(72) 발명자

이준협

서울특별시 서대문구 연희로10길 29-5, 205호(연희동)

황상원

서울특별시 강서구 공항대로45길 68-6, 106동 605호(등촌동, 서광등촌마을아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711080916

부처명 과학기술정보통신부

과제관리(전문)기관명 정보통신기획평가원(한국연구재단부설)

연구사업명 정보통신방송연구개발사업

연구과제명 스마트카 다중 센서와 딥러닝을 이용한 초정밀 내추럴 3D 뷰 생성 기술 개발 (창조
씨앗형 2단계)(3/5)

기 여 율 1/1

과제수행기관명 연세대학교 산학협력단

연구기간 2019.01.01 ~ 2019.12.31

명세서

청구범위

청구항 1

두 개의 프레임 이미지를 입력받아 상기 두 개의 프레임 이미지에 포함된 객체별 광학 플로우를 연산하도록 학습되어 있는 객체별 광학 플로우 네트워크 모듈;

상기 두 개의 프레임 이미지를 입력받아 상기 두 개의 프레임 이미지에 포함된 객체들의 객체 영역을 검출하도록 학습되어 있는 객체 검출 네트워크 모듈; 및

상기 객체별 광학 플로우 네트워크 모듈 및 상기 객체 검출 네트워크 모듈의 출력을 이용하여 상기 두 개의 프레임 이미지의 동적 객체를 검출하는 동적 객체 검출 모듈을 포함하되,

상기 두 개의 프레임 이미지는 동일한 카메라에 의해 촬영되고 시간적 갭이 있는 이미지인 것을 특징으로 하는 동적 객체 검출 장치.

청구항 2

제1항에 있어서,

상기 객체별 광학 플로우 네트워크 모듈 및 상기 객체 검출 네트워크 모듈은 각각 인코더 네트워크 및 디코더 네트워크를 포함하며, 상기 객체별 광학 플로우 네트워크 모듈 및 상기 객체 검출 네트워크 모듈은 동일한 인코더 네트워크를 공유하는 것을 특징으로 하는 동적 객체 검출 장치.

청구항 3

제2항에 있어서,

상기 객체 검출 네트워크 모듈의 객체 검출 디코더 네트워크는 디코딩 과정에서 생성되는 적어도 하나의 특징맵을 상기 객체별 광학 플로우 네트워크 모듈의 객체별 광학 플로우 디코더 네트워크에 전달하는 것을 특징으로 하는 동적 객체 검출 장치.

청구항 4

제3항에 있어서,

상기 객체 검출 디코더 네트워크의 레이어 수와 상기 광학 플로우 디코더 네트워크의 레이어 수는 동일하며, 상기 객체 검출 디코더 네트워크의 각 레이어의 특징맵은 동일 차수의 상기 객체별 광학 플로우 디코더 네트워크의 레이어로 전달되는 것을 특징으로 하는 동적 객체 검출 장치.

청구항 5

제4항에 있어서,

상기 객체 검출 디코더 네트워크의 각 레이어의 특징맵 디멘션은 상기 객체별 광학 플로우 디코더 네트워크의 각 레이어의 특징맵 디멘션과 동일한 것을 특징으로 하는 동적 객체 검출 장치.

청구항 6

제5항에 있어서,

상기 객체별 광학 플로우 디코더 네트워크는 상기 객체 검출 디코더 네트워크로부터 전달되는 특정 레이어의 특징맵을 상기 특정 레이어에 상응하는 레이어의 특징맵과 결합(Concatenation)하여 디코딩을 수행하는 것을 특징으로 하는 동적 객체 검출 장치.

청구항 7

제1항에 있어서,

상기 객체별 광학 플로우 네트워크 모듈 및 상기 객체 검출 모듈은 지도 학습에 의해 객체별 광학 플로우 및 객체 검출을 학습하는 것을 특징으로 하는 동적 객체 검출 장치.

청구항 8

제1 항에 있어서,

상기 객체별 광학 플로우 네트워크 모듈을 학습하기 위한 객체별 광 플로우 참값은 상기 두 개의 프레임 이미지 간 광학 플로우를 연산한 후 객체별로 획득되는 다수의 광학 플로우 값들 중 어느 하나를 대표값으로 설정하고, 상기 설정된 대표값을 객체별 광학 플로우 값으로 설정하여 획득되는 것을 특징으로 하는 동적 객체 검출 장치.

청구항 9

두 개의 프레임 이미지를 입력받아 상기 두 개의 프레임 이미지에 포함된 객체별 광학 플로우를 연산하도록 학습되어 있는 객체별 광학 플로우 네트워크 모듈을 통해 동적 영역을 검출하는 단계;

상기 두 개의 프레임 이미지를 입력받아 상기 두 개의 프레임 이미지에 포함된 객체들의 객체 영역을 검출하도록 학습되어 있는 객체 검출 네트워크 모듈을 통해 객체 영역을 검출하는 단계; 및

상기 검출되는 동적 영역 및 객체 영역을 이용하여 동적 객체를 검출하는 단계를 포함하되,

상기 두 개의 프레임 이미지는 동일한 카메라에 의해 촬영되고 시간적 갭이 있는 이미지인 것을 특징으로 하는 동적 객체 검출 방법.

청구항 10

제9항에 있어서,

상기 객체별 광학 플로우 네트워크 모듈 및 상기 객체 검출 네트워크 모듈은 각각 인코더 네트워크 및 디코더 네트워크를 포함하며, 상기 객체별 광학 플로우 네트워크 모듈 및 상기 객체 검출 네트워크 모듈은 동일한 인코더 네트워크를 공유하는 것을 특징으로 하는 동적 객체 검출 방법.

청구항 11

제10항에 있어서,

상기 객체 검출 네트워크 모듈의 객체 검출 디코더 네트워크는 디코딩 과정에서 생성되는 적어도 하나의 특징맵을 상기 객체별 광학 플로우 네트워크 모듈의 객체별 광학 플로우 디코더 네트워크에 전달하는 것을 특징으로 하는 동적 객체 검출 방법.

청구항 12

제11항에 있어서,

상기 객체 검출 디코더 네트워크의 레이어 수와 상기 광학 플로우 디코더 네트워크의 레이어 수는 동일하며, 상기 객체 검출 디코더 네트워크의 각 레이어의 특징맵은 동일 차수의 상기 객체별 광학 플로우 디코더 네트워크의 레이어로 전달되는 것을 특징으로 하는 동적 객체 검출 방법.

청구항 13

제12항에 있어서,

상기 객체 검출 디코더 네트워크의 각 레이어의 특징맵 디멘션은 상기 객체별 광학 플로우 디코더 네트워크의 각 레이어의 특징맵 디멘션과 동일한 것을 특징으로 하는 동적 객체 검출 방법.

청구항 14

제13항에 있어서,

상기 객체별 광학 플로우 디코더 네트워크는 상기 객체 검출 디코더 네트워크로부터 전달되는 특정 레이어의 특징맵을 상기 특정 레이어에 상응하는 레이어의 특징맵과 결합(Concatenation)하여 디코딩을 수행하는 것을 특징으로 하는 동적 객체 검출 방법.

청구항 15

제9항에 있어서,

상기 객체별 광학 플로우 네트워크 모듈 및 상기 객체 검출 모듈은 지도 학습에 의해 객체별 광학 플로우 및 객체 검출을 학습하는 것을 특징으로 하는 동적 객체 검출 방법.

청구항 16

제9항에 있어서,

상기 객체별 광학 플로우 네트워크 모듈을 학습하기 위한 객체별 광 플로우 참값은 상기 두 개의 프레임 이미지 간 광학 플로우를 연산한 후 객체별로 획득되는 다수의 광학 플로우 값들 중 어느 하나를 대표값으로 설정하고, 상기 설정된 대표값을 객체별 광학 플로우 값으로 설정하여 획득되는 것을 특징으로 하는 동적 객체 검출 방법.

청구항 17

제9항 내지 제16항 중 어느 한 항의 방법을 실행하기 위한 프로그램이 기록되어 있으며 디지털 데이터 처리 장치에 의해 판독 가능한 프로그램이 기록된 기록매체.

발명의 설명

기술 분야

본 발명은 동적 객체 검출 장치 및 방법에 관한 것으로서, 더욱 상세하게는 동일한 카메라에 의해 획득되는 프레임들을 분석하여 동적 객체를 검출하는 장치 및 방법에 관한 것이다.

배경 기술

[0001]

- [0003] 동적 객체를 정확하게 검출하는 것은 어려운 기술이며, 특히 카메라가 비고정인 상태에서 동적 객체를 검출하는 것은 더욱 어려운 문제이다. 카메라가 비고정인 상태에서 동적 객체를 검출하기 위해 광학 플로우(Optical flow)를 이용한 방법들이 제안되어 왔다.
- [0004] 그러나, 기존의 광학 플로우 기반의 동적 객체 검출 기술은 현재 프레임과 다음 프레임의 픽셀 값의 차이를 연산하여 움직임을 검출하는 방식이기에 상당한 연산량을 요구하게 된다. 근래에 들어, 자율 주행에 대한 연구가 활성화되면서 주행 방향을 결정하기 위해 동적 객체의 검출이 필수적으로 요구되고 있으나 모든 픽셀에 대해 광학 플로우가 연산될 경우 실시간 검출이 어려워지는 문제가 있다. 특히, 인공 신경망을 이용하여 동적 객체를 검출하게 될 경우 이러한 연산량 증가에 대한 문제는 더욱 심화되게 된다.
- [0005] 또한, 조도 변화에 취약한 픽셀값의 특성으로 인해 인접한 프레임 이미지에서 동일 화소인지 여부를 판단하기 어려운 문제가 있고, 특정 화소가 동일한 객체에 속하는 것인지에 대한 판별 역시 어려운 문제가 있어 정확한 동적 객체 검출이 이루어지지 않는 문제점이 있었다.

발명의 내용

해결하려는 과제

- [0008] 본 발명은 실시간으로 동적 객체를 검출할 수 있는 동적 객체 검출 장치 및 방법을 제안한다.
- [0009] 또한, 본 발명은 조도 변화 및 환경 변화에 강인하게 동적 객체를 검출할 수 있는 방법 및 장치를 제안한다.

과제의 해결 수단

- [0011] 상기와 같은 목적을 달성하기 위해, 본 발명의 일 측면에 따르면, 두 개의 프레임 이미지를 입력받아 상기 두 개의 프레임 이미지에 포함된 객체별 광학 플로우를 연산하도록 학습되어 있는 객체별 광학 플로우 네트워크 모듈; 상기 두 개의 프레임 이미지를 입력받아 상기 두 개의 프레임 이미지에 포함된 객체들의 객체 영역을 검출하도록 학습되어 있는 객체 검출 네트워크 모듈; 및 상기 객체별 광학 플로우 네트워크 모듈 및 상기 객체 검출 네트워크 모듈의 출력을 이용하여 상기 두 개의 프레임 이미지의 동적 객체를 검출하는 동적 객체 검출 모듈을 포함하되, 상기 두 개의 프레임 이미지는 동일한 카메라에 의해 촬영되고 시간적 겹이 있는 이미지인 동적 객체 검출 장치가 제공된다.
- [0012] 상기 객체별 광학 플로우 네트워크 모듈 및 상기 객체 검출 네트워크 모듈은 각각 인코더 네트워크 및 디코더 네트워크를 포함하며, 상기 객체별 광학 플로우 네트워크 모듈 및 상기 객체 검출 네트워크 모듈은 동일한 인코더 네트워크를 공유한다.
- [0013] 상기 객체 검출 네트워크 모듈의 객체 검출 디코더 네트워크는 디코딩 과정에서 생성되는 적어도 하나의 특징맵을 상기 객체별 광학 플로우 네트워크 모듈의 객체별 광학 플로우 디코더 네트워크에 전달한다.
- [0014] 상기 객체 검출 디코더 네트워크의 레이어 수와 상기 광학 플로우 디코더 네트워크의 레이어 수는 동일하며, 상기 객체 검출 디코더 네트워크의 각 레이어의 특징맵은 동일 차수의 상기 객체별 광학 플로우 디코더 네트워크의 레이어로 전달된다.
- [0015] 상기 객체 검출 디코더 네트워크의 각 레이어의 특징맵 디멘션은 상기 객체별 광학 플로우 디코더 네트워크의 각 레이어의 특징맵 디멘션과 동일하다.
- [0016] 상기 객체별 광학 플로우 디코더 네트워크는 상기 객체 검출 디코더 네트워크로부터 전달되는 특정 레이어의 특징맵을 상기 특정 레이어에 상응하는 레이어의 특징맵과 결합(Concatenation)하여 디코딩을 수행한다.
- [0017] 상기 객체별 광학 플로우 네트워크 모듈 및 상기 객체 검출 모듈은 지도 학습에 의해 객체별 광학 플로우 및 객체 검출을 학습한다.
- [0018] 상기 객체별 광학 플로우 네트워크 모듈을 학습하기 위한 객체별 광 플로우 참값은 상기 두 개의 프레임 이미지 간 광학 플로우를 연산한 후 객체별로 획득되는 다수의 광학 플로우 값들 중 어느 하나를 대표값으로 설정하고,

상기 설정된 대표값을 객체별 광학 플로우 값으로 설정하여 획득된다.

[0019] 본 발명의 다른 측면에 따르면, 두 개의 프레임 이미지를 입력받아 상기 두 개의 프레임 이미지에 포함된 객체별 광학 플로우를 연산하도록 학습되어 있는 객체별 광학 플로우 네트워크 모듈을 통해 동적 영역을 검출하는 단계; 상기 두 개의 프레임 이미지를 입력받아 상기 두 개의 프레임 이미지에 포함된 객체들의 객체 영역을 검출하도록 학습되어 있는 객체 검출 네트워크 모듈을 통해 객체 영역을 검출하는 단계; 및 상기 검출되는 동적 영역 및 객체 영역을 이용하여 동적 객체를 검출하는 단계를 포함하되, 상기 두 개의 프레임 이미지는 동일한 카메라에 의해 촬영되고 시간적 갭이 있는 이미지인 동적 객체 검출 방법이 제공된다.

발명의 효과

[0021] 본 발명의 실시예들에 따르면, 실시간으로 동적 객체를 검출할 수 있으며, 조도 변화 및 환경 변화에 강인하게 동적 객체를 검출할 수 있는 장점이 있다.

도면의 간단한 설명

[0023] 도 1은 본 발명의 일 실시예에 따른 동적 객체 검출 장치의 전체적인 구조를 도시한 블록도.
 도 2는 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 객체별 광학 플로우 네트워크 모듈의 학습 구조를 나타낸 도면.
 도 3은 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 객체 검출 네트워크 모듈의 학습 구조를 나타낸 도면.
 도 4는 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 객체별 광학 플로우 네트워크 모듈 및 객체 검출 네트워크 모듈의 구조를 나타낸 도면.
 도 5는 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 객체별 광학 플로우 네트워크 모듈 및 객체 검출 네트워크 모듈의 상세 구조를 나타낸 도면.
 도 6은 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 객체별 광학 플로우의 참값을 획득하는 방법을 나타낸 도면.
 도 7은 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 객체별 광학 플로우의 대표값을 설정하는 일례를 나타낸 도면.
 도 8은 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 동적 객체를 검출한 결과를 나타낸 도면.
 도 9는 본 발명의 일 실시예에 따른 동적 객체 검출 방법의 전체적인 흐름을 나타낸 도면.

발명을 실시하기 위한 구체적인 내용

[0024] 이하에서는 첨부한 도면을 참조하여 본 발명을 설명하기로 한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 따라서 여기에서 설명하는 실시예로 한정되는 것은 아니다.
 [0025] 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
 [0026] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 부재를 사이에 두고 "간접적으로 연결"되어 있는 경우도 포함한다.
 [0027] 또한 어떤 부분이 어떤 구성 요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성 요소를 제외하는 것이 아니라 다른 구성 요소를 더 구비할 수 있다는 것을 의미한다.
 [0028] 이하 첨부된 도면을 참고하여 본 발명의 실시예를 상세히 설명하기로 한다.
 [0029] 도 1은 본 발명의 일 실시예에 따른 동적 객체 검출 장치의 전체적인 구조를 도시한 블록도이다.
 [0030] 도 1을 참조하면, 본 발명의 일 실시예에 따른 동적 객체 검출 장치는 객체별 광학 플로우 네트워크 모듈(100),

객체 검출 네트워크 모듈(110) 및 동적 객체 검출 모듈(120)을 포함한다.

- [0031] 동적 객체의 검출은 다양한 분야에서 요구되고 있으며, 특히 자율 주행 분야에서는 실시간으로 동적 객체와 비동적 객체를 구분하는 작업이 필수적으로 요구된다.
- [0032] 본 발명의 동적 객체 검출 장치로는 두 개의 프레임 이미지가 입력된다. 두 개의 프레임 이미지는 시간적으로 차이가 나는 프레임 이미지이며, 예를 들어, t 프레임 이미지와 $(t+1)$ 이미지일 수 있다. 물론, 입력되는 두 개의 프레임 이미지가 연속하는 프레임 이미지인 것에 한정되는 것은 아니며 시간적 갭을 가지면서 동일한 카메라에 의해 획득되는 프레임 이미지라면 본 발명의 입력 이미지로 사용될 수 있을 것이다.
- [0033] 만일 프레임 이미지에 포착된 특정 객체가 동적 객체라면 제1 프레임과 제2 프레임에서 위치의 변화가 있을 것이다. 반면에 정적 객체라면 제1 프레임과 제2 프레임에서 위치의 변화가 발생하지 않을 것이다.
- [0034] 이러한 위치의 변화를 감지하기 위해 종래에는 광학 플로우가 이용되었다. 그러나, 발명의 배경의 기술에서 설명한 바와 같이, 광학 플로우만으로 동적 객체를 정확히 검출하기에는 한계가 있었다.
- [0035] 광학 플로우는 픽셀 단위로 위치의 변화를 감지하는 방법이기에 광학 플로우만으로는 객체 단위의 위치 변화를 감지할 수 없다. 또한, 광학 플로우는 제1 프레임과 제2 프레임의 모든 픽셀에 대해 그 차이값을 연산하는 방식이기에 매우 높은 연산량을 요구하므로 동적 객체의 실시간 검출 연산에도 적절하지 않으며 특히 신경망 연산이 도입될 경우에는 그 복잡도가 더욱 심화되는 문제점이 있다.
- [0036] 아울러, 광학 플로우는 특정 프레임에서의 물체의 픽셀값이 다음 프레임에서 변하지 않는다는 전제 하에서 그 정확성을 담보할 수 있는 것이나 조도 변화와 같은 환경 변화에 따라 동일한 물체라도 프레임간 픽셀값의 차이가 발생하는 경우는 빈번하게 발생한다.
- [0037] 본 발명은 이와 같은 문제를 해결하기 위해 두 개의 신경망을 이용하며, 이는 도 1에 도시된 객체별 광학 플로우 네트워크 모듈(100) 및 객체 검출 네트워크 모듈(110)이다.
- [0038] 객체별 광학 플로우 네트워크 모듈(100)은 프레임 이미지에 포함된 객체별로 광학 플로우를 출력한다. 객체별 광학 플로우 네트워크 모듈(100)은 신경망 연산을 통해 객체별 광학 플로우를 출력한다. 객체별 광학 플로우 네트워크 모듈(100)은 예를 들어, CNN(Convolutional Neural Network)를 이용하여 객체별 광학 플로우를 출력할 수 있으나 다른 종류의 신경망이 이용될 수도 있다는 점은 당업자에게 있어 자명할 것이다.
- [0039] CNN이 이용될 경우, 객체별 광학 플로우는 컨볼루션 가중치를 프레임의 각 픽셀에 적용하는 방식으로 객체별 광학 플로우를 출력할 수 있을 것이다.
- [0040] 객체별 광학 플로우 네트워크 모듈(100)은 학습에 의해 입력 프레임으로부터 객체별 광학 플로우를 출력하며, 일례로 지도 학습에 의해 객체별 광학 플로우를 출력할 수 있을 것이다.
- [0041] 도 2는 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 객체별 광학 플로우 네트워크 모듈의 학습 구조를 나타낸 도면이다.
- [0042] 도 2를 참조하면, 본 발명의 일 실시예에 따른 객체별 광학 플로우 네트워크 모듈(100)은 신경망 연산을 통해 객체별 광학 플로우를 출력한다.
- [0043] 출력되는 객체별 광학 플로우는 미리 준비된 참값(Ground Truth)과 비교되고, 참값과의 차이인 에러가 연산된다. 연산된 에러는 객체별 광학 플로우 네트워크 모듈로 역전파되고 역전파되는 에러에 기초하여 객체별 광학 플로우 네트워크 모듈(100)의 가중치가 갱신된다. 알려진 바와 같이, 에러 함수의 그래디언트가 최소화되는 방향으로 가중치 갱신이 이루어질 수 있을 것이다.
- [0044] 갱신된 가중치에 기초하여 객체별 광학 플로우는 다시 연산되고, 연산된 결과를 다시 참값(Ground Truth)과 비교하여 가중치를 갱신하면서 학습이 이루어진다.
- [0045] 객체 검출 네트워크 모듈(110)은 프레임 이미지에 포함된 객체를 검출하는 기능을 한다. 프레임 이미지에는 배경 및 객체들이 존재하며, 객체 검출 네트워크 모듈(110)은 배경을 제외한 객체들만을 검출하는 것이다. 구체적으로 객체 검출 네트워크 모듈(110)은 이미지에서 객체 영역만을 검출하는 것으로 그 기능을 정의할 수 있을 것이다.
- [0046] 객체 검출 네트워크 모듈(110)도 신경망 연산을 통해 객체 영역을 검출한다. 예를 들어, CNN(Convolutional Neural Network)을 이용하여 객체 영역을 검출할 수 있을 것이나, 다른 종류의 신경망이 이용될 수도 있다는 점

은 당업자에게 있어 자명할 것이다.

- [0047] CNN이 이용될 경우, 객체 검출 네트워크 모듈(110) 역시 콘볼루션 가중치를 프레임의 각 픽셀에 적용하는 방식으로 객체 영역을 검출할 수 있을 것이다.
- [0048] 객체 검출 네트워크 모듈(110)은 학습에 의해 입력 프레임으로부터 객체 영역을 검출하고, 일례로 지도 학습에 의해 객체 영역을 검출할 수 있을 것이다.
- [0049] 도 3은 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 객체 검출 네트워크 모듈의 학습 구조를 나타낸 도면이다.
- [0050] 도 3을 참조하면, 본 발명의 일 실시예에 따른 객체 검출 네트워크 모듈(110)은 신경망 연산을 통해 객체별 광학 플로우를 출력한다.
- [0051] 출력되는 객체 검출 영역은 미리 준비된 참값(Ground Truth)과 비교되고, 참값과의 차이인 에러가 연산된다. 연산된 에러는 객체 검출 네트워크 모듈로 역전파되고 역전파되는 에러에 기초하여 객체 검출 네트워크 모듈(110)의 가중치가 갱신된다. 알려진 바와 같이, 에러 함수의 그래디언트가 최소화되는 방향으로 가중치 갱신이 이루어질 수 있을 것이다.
- [0052] 갱신된 가중치에 기초하여 객체 검출 영역은 다시 연산되고, 연산된 결과를 다시 참값(Ground Truth)과 비교하여 가중치를 갱신하면서 학습이 이루어진다.
- [0053] 도 4는 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 객체별 광학 플로우 네트워크 모듈 및 객체 검출 네트워크 모듈의 구조를 나타낸 도면이다
- [0054] 도 4를 참조하면, 본 발명의 일 실시예에 따른 객체별 광학 플로우 네트워크 모듈(100) 및 객체 검출 네트워크 모듈(110)은 인코더 네트워크(200), 객체별 광학 플로우 디코더 네트워크(210) 및 객체 검출 디코더 네트워크(220)를 포함한다.
- [0055] 객체별 광학 플로우 네트워크 모듈(100)은 인코더 네트워크(200) 및 객체별 광학 플로우 디코더 네트워크(210)를 포함하며, 객체 검출 네트워크 모듈(110)은 인코더 네트워크(200) 및 객체 검출 디코더 네트워크(220)를 포함한다.
- [0056] 도 4를 통해 확인되는 바와 같이, 본 발명의 객체별 광학 플로우 네트워크 모듈(100)과 객체 검출 네트워크 모듈(110)은 인코더 네트워크(200)를 공유한다. 다만, 디코더 네트워크는 별개로 형성되는 것이다.
- [0057] 입력 프레임 이미지로부터 인코더 네트워크(200)는 특징맵을 인코딩하는 일반적인 네트워크를 사용할 수 있을 것이며, 인코더 네트워크(200)에 대한 별도의 학습이 이루어지지 않을 수 있다. 물론, 인코더 네트워크 역시 그 가중치가 학습될 수도 있을 것이다.
- [0058] 객체별 광학 플로우 디코더 네트워크(210) 및 객체 검출 디코더 네트워크(220)는 각각 다른 참값(Ground Truth)를 이용하여 학습된다는 점은 이미 위에서 설명하였다. 객체별 광학 플로우 디코더 네트워크(210)와 객체 검출 디코더 네트워크(220)는 동일한 수의 레이어를 가지고 각 레이어의 특징맵의 디멘션(Dimension)은 동일하게 설정된다. 네트워크의 레이어의 수와 각 레이어에서의 특징맵의 디멘션은 디코더 네트워크 설정 시에 미리 정해지며, 본 발명은 광학 플로우 디코더 네트워크(210)와 객체 검출 디코더 네트워크(220)의 레이어의 수와 각 레이어의 특징맵 디멘션이 동일하도록 미리 설정하는 것이다.
- [0059] 본 발명은 객체 검출 디코더 네트워크(220)의 각 레이어의 특징맵을 동일한 차수의 객체별 광학 플로우 디코더 네트워크(210)의 레이어에 전달한다. 광학 플로우 디코더 네트워크(210)는 객체 검출 디코더 네트워크(220)로부터 전달되는 특징맵을 참조하여 디코딩을 수행한다.
- [0060] 이와 같이 객체 검출 디코더 네트워크(220)의 특징맵을 객체별 광학 플로우 디코더 네트워크(210)에 전달하는 것은 객체별 광학 플로우를 추론할 때 객체 검출 디코더 네트워크(220)로부터 제공되는 특징맵을 이용하여 객체 영역을 확인하기 위해서이다.
- [0061] 이와 같은 객체 검출 디코더 네트워크(220)로부터 객체별 광학 플로우 디코더 네트워크(210)로의 특징맵 전달을 통해 객체별 광학 플로우에 대한 학습 효율을 향상시킬 수 있을 뿐만 아니라 보다 정확한 객체별 광학 플로우 추론이 가능해진다.
- [0062] 도 5는 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 객체별 광학 플로우 네트워크 모듈 및 객체 검출

네트워크 모듈의 상세 구조를 나타낸 도면이다.

- [0063] 도 5를 참조하면, 본 발명의 일 실시예에 따른 객체별 광학 플로우 디코더 네트워크(210)와 객체 검출 디코더 네트워크(220)는 다수의 레이어로 이루어져 있다. 통상적인 디코더 네트워크와 같이 객체별 광학 플로우 디코더 네트워크(210)와 객체 검출 디코더 네트워크(220)는 레이어가 증가함에 따라 특징맵의 디멘션이 증가하게 된다.
- [0064] 각 디코더 네트워크(210, 220)의 최종 레이어의 출력은 각 네트워크 모듈(100, 110)의 출력으로 작용한다.
- [0065] 앞서 설명한 바와 같이, 객체별 광학 플로우 디코더 네트워크(210)의 레이어 수와 객체 검출 디코더 네트워크(220)의 레이어 수는 동일하다. 또한, 각 레이어의 특징맵의 디멘션은 동일하게 설정되는 것을 도 5로부터 확인할 수 있다.
- [0066] 객체별 광학 플로우 디코더 네트워크(210)와 객체 검출 디코더 네트워크(220)에서 각 레이어의 디멘션을 동일하게 설정하는 것은 객체 검출 디코더 네트워크(220)에서 전달하는 특징맵과 객체별 광학 플로우 디코더 네트워크(210)의 특징맵과의 결합(Concatenation)을 위해서이다.
- [0067] 도 5를 참조하면, 객체별 광학 플로우 디코더 네트워크(210) 및 객체 검출 디코더 네트워크(220)의 레이어별로 객체 검출 디코더 네트워크(220)의 특징맵이 객체별 광학 플로우 디코더 네트워크(210)로 전달되고, 객체별 광학 플로우 디코더 네트워크(210)는 양 디코더 네트워크(220)는 양 레이어의 특징맵을 결합(Concatenation)한 후 다음 레이어로의 디코딩을 수행한다.
- [0068] 동적 객체 검출 모듈(120)은 객체별 광학 플로우 네트워크 모듈(100)의 출력 및 객체 검출 네트워크 모듈(110)의 출력을 이용하여 입력된 프레임으로부터 동적 객체를 검출한다.
- [0069] 객체 검출 네트워크 모듈(110)은 입력된 프레임 이미지로부터 객체 영역을 검출하고, 객체별 광학 플로우 네트워크 모듈(100)은 객체별 광학 플로우를 통해 동적임 움직임이 있는 영역을 검출하기 때문에, 동적 객체 검출 모듈(120)은 객체별 광학 플로우 네트워크 모듈(100)로부터 출력하는 동적 영역에 존재하는 객체 검출 네트워크 모듈(110)의 객체들을 동적 객체로 판단한다.
- [0070] 도 8은 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 동적 객체를 검출한 결과를 나타낸 도면이다.
- [0071] 도 8에 도시된 바와 같이, 객체별 광학 플로우 네트워크 모듈(100)의 동적 영역 및 객체 검출 네트워크 모듈(110)의 객체 영역 정보를 이용하여 동적 객체가 검출되며, 도 8에는 동적 객체가 하얀색으로 표시되어 있다.
- [0072] 이와 같이 검출되는 동적 객체는 자율 주행에서 움직이는 사람 또는 차량으로 인지되고 검출된 정보를 이용하여 주행 방향을 결정할 수 있게 된다.
- [0073] 도 6은 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 객체별 광학 플로우의 참값을 획득하는 방법을 나타낸 도면이다.
- [0074] 광학 플로우는 픽셀 단위로 획득되는 것이기에 객체별 광학 플로우는 참값(Ground Truth)을 획득하기 어려운 측면이 있다. 가장 정확한 객체별 광학 플로우는 각 객체의 순간 속도를 직접 측정하는 것에 의해 획득될 수 있으나 이 역시 현실적으로 어려운 작업이다.
- [0075] 본 발명에서는 객체별 광학 플로우 네트워크 모듈의 학습을 위해 객체별 광학 플로우의 참값을 획득하는 방법을 제시한다.
- [0076] 도 6과 같은 방법을 통해 획득되는 객체별 광학 플로우 참값은 학습 단계에서만 이용되는 것이다. 또한, 도 6에 도시된 객체별 광학 플로우 참값 획득 방법은 예시적인 것이며, 다른 다양한 방법에 의해 각 객체의 속도에 상응하는 파라미터를 획득할 수 있다는 점은 당업자에게 있어 자명할 것이다.
- [0077] 도 6을 참조하면, 먼저 학습 대상 프레임 영상으로부터 객체 영역을 선택한다(단계 600).
- [0078] 객체 영역 선택이 이루어지면, 기존의 광학 플로우 연산 방식에 의해 광학 플로우를 연산한다(단계 602). 앞서 설명한 바와 같이, 광학 플로우 연산은 픽셀 단위로 이루어지기에 앞서 수행한 객체 영역 선택과는 무관하게 광학 플로우가 연산된다.
- [0079] 프레임 이미지의 모든 픽셀에 대해 광학 플로우가 연산되면, 선택한 객체 영역별로 연산된 광학 플로우 값들을 추출한다(단계 604). 예를 들어, 프레임 이미지의 세 개의 객체가 포함되어 있는 경우, 세 개의 객체 영역별로 광학 플로우 값들을 추출하는 것이다.

- [0080] 각 객체 영역별로 광학 플로우 값이 추출되면, 각 객체 영역의 광학 플로우 값들의 분포를 확인한다(단계 606). 예를 들어, HOF(Histogram of Oriented Optical Flow)를 이용하여 광학 플로우 값들의 분포를 확인할 수 있을 것이다. 물론 HOF 이외에도 다수의 값들의 분포를 확인할 수 있는 다양한 방법이 사용될 수 있을 것이다.
- [0081] 객체 영역의 광학 플로우 값들의 분포에 기초하여 각 객체 영역의 광학 플로우 대표값을 설정한다(단계 608). 대표값 설정은 다양한 정규화 방식을 이용하여 이루어질 수 있을 것이다. 일례로, 가장 많이 분포되는 광학 플로우 값을 해당 객체 영역의 대표값으로 설정할 수 있을 것이다.
- [0082] 선택된 각 객체의 대표값을 각 객체의 광학 플로우 참값으로 결정한다(단계 610).
- [0083] 도 7은 본 발명의 일 실시예에 따른 동적 객체 검출 장치에서 객체별 광학 플로우의 대표값을 설정하는 일례를 나타낸 도면이다.
- [0084] 도 7의 (a)는 특정 객체 영역의 다양한 광학 플로우 값들을 나타낸 것이다. 도 7의 (a)와 같이 일반적인 광학 플로우 연산 방식에 의할 경우 같은 객체이더라도 다양한 광학 플로우 분포를 가지게 된다.
- [0085] 도 7의 (b)와 같이 본 발명은 특정 객체 영역 내의 다양한 광학 플로우 값들에 대한 대표값을 설정하여 해당 대표값을 해당 객체의 광학 플로우 값으로 지정하는 것이며, 지정된 대표값을 해당 객체의 광학 플로우 참값으로 이용하여 학습을 수행한다.
- [0086] 도 9는 본 발명의 일 실시예에 따른 동적 객체 검출 방법의 전체적인 흐름을 나타낸 도면이다.
- [0087] 도 9는 객체별 광학 플로우 네트워크 모듈(100) 및 객체 검출 네트워크 모듈(110)의 학습이 완료된 후 동적 객체를 실질적으로 검출하는 과정을 나타낸 순서도이고 학습 과정은 도 9에 도시되어 있지 않다.
- [0088] 도 9를 참조하면, 시간적 갭이 있는 두 개의 프레임 이미지를 입력받는다(단계 900).
- [0089] 두 개의 프레임 이미지가 입력되면, 입력된 이미지에 대해 인코더를 이용한 인코딩을 수행함으로써 특징맵을 생성한다(단계 902). 인코딩은 다수의 레이어를 거치면서 이루어질 수 있으며, 레이어의 차수가 증가할수록 인코딩되는 특징맵의 디멘션은 감소할 수 있다.
- [0090] 인코딩을 통해 특징맵이 생성되면, 객체 검출 디코더 네트워크에서 인코딩된 특징맵에 대한 디코딩을 수행한다(단계 904). 앞서 설명한 바와 같이, 디코딩은 다수의 레이어를 거치면서 이루어진다. 객체 검출 디코더 네트워크의 각 레이어에서 생성되는 특징맵은 객체별 광학 플로우 디코더 네트워크로 전달된다.
- [0091] 객체별 광학 플로우 디코더 네트워크는 인코딩된 특징맵에 대한 디코딩을 수행한다(단계 906). 객체별 광학 플로우 디코더 네트워크에서의 디코딩도 다수의 레이어를 거치면서 이루어진다. 앞서 설명한 바와 같이 객체 검출 디코더 네트워크의 각 레이어의 특징맵 디멘션은 객체별 광학 플로우 디코더 네트워크의 각 레이어의 디멘션과 동일하게 설정된다.
- [0092] 객체 검출 디코더 네트워크로부터 전달되는 각 레이어의 특징맵은 객체별 광학 플로우 디코더 네트워크의 각 레이어의 특징맵과 결합되어 디코딩이 이루어진다.
- [0093] 한편, 위에서는 객체 검출 디코더 네트워크의 모든 레이어의 특징맵이 객체별 광학 플로우에 전달되는 경우에 대해 설명하였으나, 모든 레이어의 특징맵이 전달되는 것에 한정되는 것은 아니며 부분적 전달될 수도 있을 것이다.
- [0094] 객체별 광학 플로우 네트워크 모듈의 출력 및 객체 검출 네트워크 모듈의 출력을 이용하여 동적 객체를 검출한다(단계 908). 앞서 설명한 바와 같이, 객체 검출 네트워크 모듈로부터 확인되는 다수의 객체 중 객체별 광학 플로우 네트워크 모듈로부터 확인되는 동적 영역에 위치하는 객체를 동적 객체로 판단한다.
- [0095] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다.
- [0096] 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다.
- [0097] 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.
- [0098] 본 발명의 범위는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등

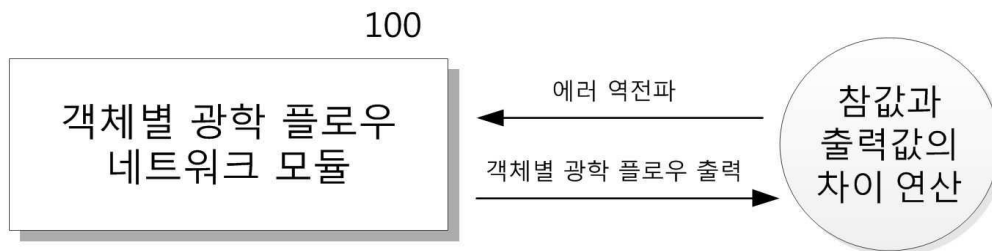
개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

도면

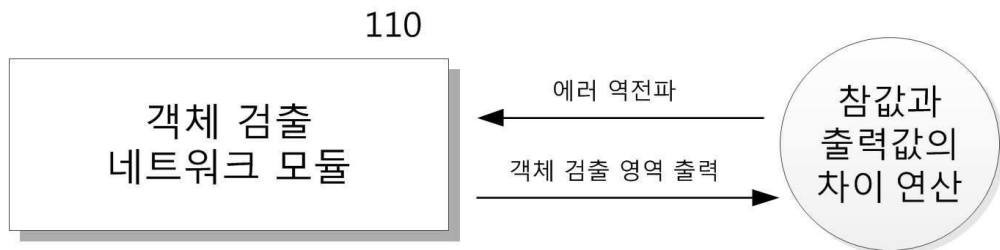
도면1



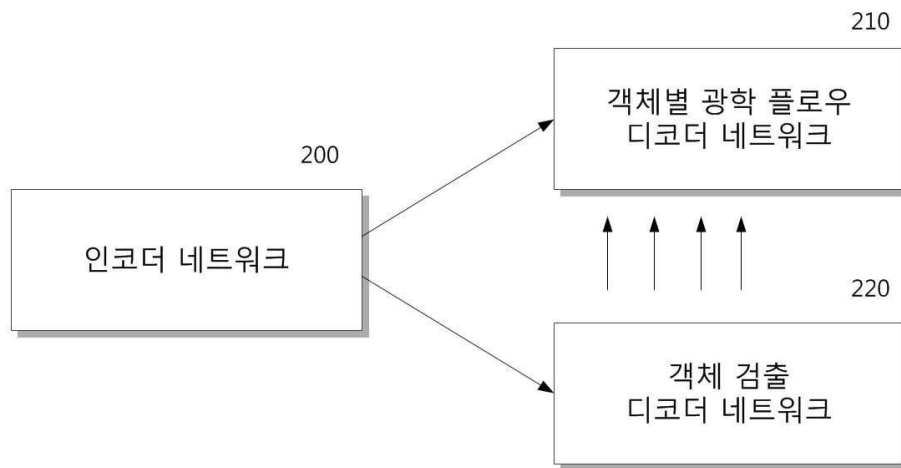
도면2



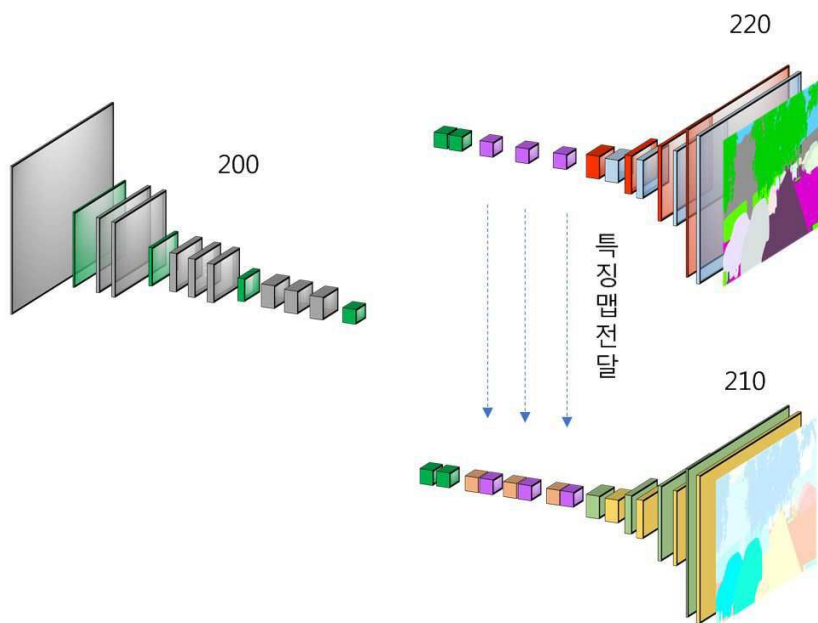
도면3



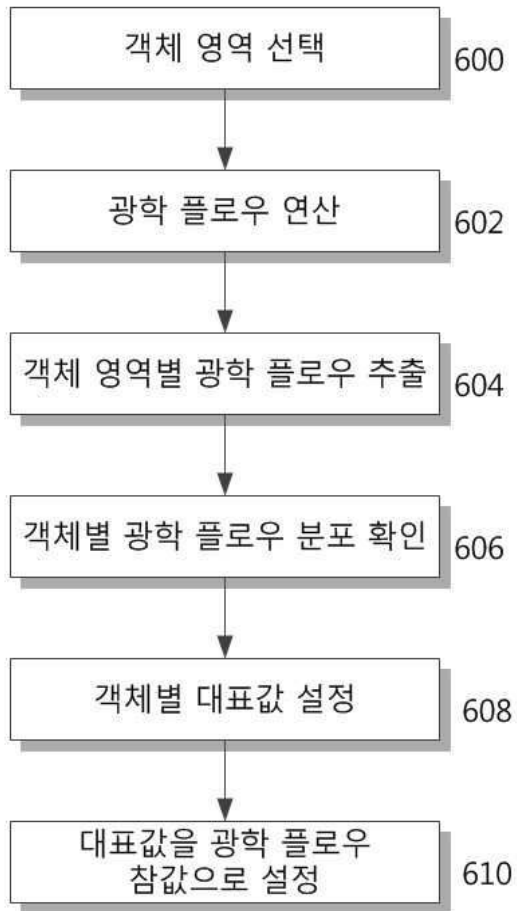
도면4



도면5



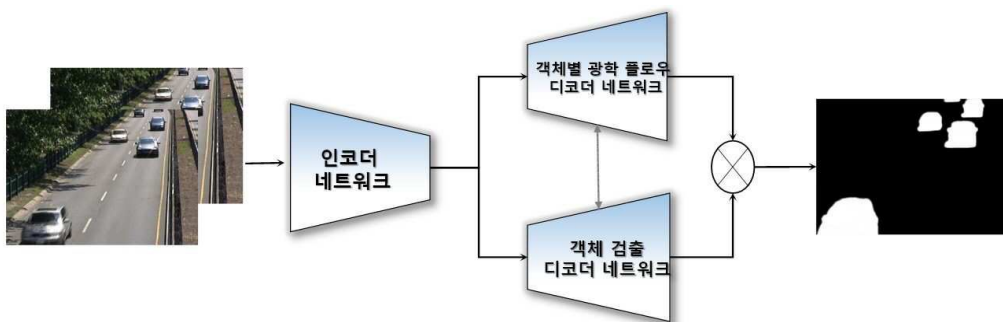
도면6



도면7



도면8



도면9

