



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0046818
(43) 공개일자 2023년04월06일

- (51) 국제특허분류(Int. Cl.)
G06N 3/08 (2023.01) G06N 3/04 (2023.01)
G06N 5/02 (2023.01) G06T 7/10 (2021.01)
- (52) CPC특허분류
G06N 3/08 (2023.01)
G06N 3/04 (2023.01)
- (21) 출원번호 10-2021-0130314
- (22) 출원일자 2021년09월30일
심사청구일자 없음

- (71) 출원인
엘지디스플레이 주식회사
서울특별시 영등포구 여의대로 128(여의도동)
포항공과대학교 산학협력단
경상북도 포항시 남구 청암로 77 (지곡동)
- (72) 발명자
오태현
경상북도 포항시 남구 청암로 77 포항공과대학교
지광환
경기도 파주시 월롱면 엘지로 245
(뒷면에 계속)
- (74) 대리인
특허법인(유한) 대아

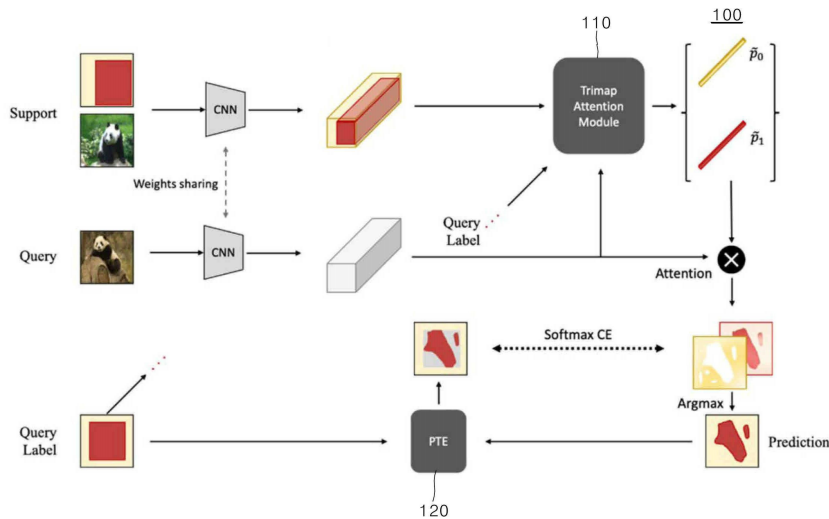
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 의미론적 영상 분할을 위한 데이터 학습 장치 및 방법

(57) 요약

적은 수의 샘플에서부터 세그먼테이션을 수행하는 능력을 학습하여 기존에 보지 못한 클래스에 대해서도 적은 수의 샘플로부터 세그먼테이션 마스크를 추론할 수 있는 의미론적 영상 분할을 위한 데이터 학습 장치 및 방법이 개시된다. 의미론적 영상 분할을 위한 데이터 학습 장치는, 에피소드의 서포트 세트와 쿼리 세트의 라벨을 교정하여 추정 세그먼테이션을 획득하고, 추정 세그먼테이션을 이용하여 트라이맵 리프리젠테이션을 획득하는 트라이맵 어텐션 모듈; 및 추정 세그먼테이션과 바운딩 박스를 이용하여 트라이맵 리프리젠테이션에서 배경 영역, 전경 영역 및 그레이 영역을 추정하는 의사 트라이맵 추정 모듈을 포함한다.

대표도



(52) CPC특허분류

G06N 5/02 (2023.01)

G06T 7/10 (2021.01)

(72) 발명자

김학진

경기도 파주시 월롱면 엘지로 245

한별이

경상북도 포항시 남구 청암로 77 포항공과대학교

명세서

청구범위

청구항 1

에피소드의 서포트 세트와 쿼리 세트의 라벨을 교정하여 추정 세그먼테이션(segmentation)을 획득하고, 상기 추정 세그먼테이션을 이용하여 트라이맵 리프리젠테이션(trimap representation)을 획득하는 트라이맵 어텐션 모듈(trimap attention module); 및

상기 추정 세그먼테이션과 바운딩 박스(bounding box)를 이용하여 상기 트라이맵 리프리젠테이션에서 배경 영역, 전경 영역 및 그레이 영역을 추정하는 의사 트라이맵 추정 모듈(Pseudo trimap estimation module)을 포함하는,

의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 2

제 1 항에 있어서,

상기 트라이맵 어텐션 모듈은,

상기 바운딩 박스가 주어진 메타 학습 과정 동안 상기 서포트 세트의 교정된 대표 특징 벡터를 구하고, 상기 대표 특징 벡터와 픽셀 별 어텐션(attention) 계산을 통하여 상기 쿼리 세트에 대한 상기 추정 세그먼테이션을 구하는,

의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 3

제 2 항에 있어서,

상기 트라이맵 어텐션 모듈은,

상기 추정 세그먼테이션을 이용하여 상기 쿼리 세트의 상기 바운딩 박스에 대한 상기 트라이맵 리프리젠테이션을 구하는,

의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 4

제 1 항에 있어서,

상기 트라이맵 어텐션 모듈은,

베이스라인의 초기 예측을 이용하여 상기 에피소드의 상기 서포트 세트와 상기 쿼리 세트의 라벨을 반복적 기법으로 교정하는 의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 5

제 1 항에 있어서,

상기 트라이맵 어텐션 모듈은,

합성곱 신경망의 상기 추정 세그먼테이션을 구하기 위한 적어도 하나의 트라이맵-어텐션 블록(Trimap-attention block)을 포함하는 의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 6

제 5 항에 있어서,

상기 트라이맵-어텐션 블록은,

상기 서포트 세트, 상기 쿼리 세트 및 베이스라인에서 구해진 상기 쿼리 세트의 초기 세그먼테이션을 기초로 목표 도메인과 참고 도메인을 교차하여 짝수 번 씹음으로써 상기 서포트 세트와 상기 쿼리 세트에 대한 상기 트라이맵 리프리젠테이션을 구하는 의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 7

제 6 항에 있어서,

상기 트라이맵-어텐션 블록은,

상기 목표 도메인과 상기 참고 도메인을 서로를 교정하는 참고로 설정하여 데이터 세트의 바운딩 박스에서 발생되는 배경 노이즈를 교정하는 의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 8

제 5 항에 있어서,

상기 트라이맵-어텐션 블록은,

참고 도메인의 라벨을 상기 트라이맵 리프리젠테이션으로 나타내고, 상기 트라이맵 리프리젠테이션 상에서 전경 영역과 배경 영역에 포함된 픽셀의 특징 벡터를 평균하는 방식으로 대표 특징 벡터를 계산하며,

상기 대표 특징 벡터와 목표 도메인의 픽셀 별 특징 벡터와의 어텐션을 구하여 상기 목표 도메인의 픽셀 별 영역을 추정하여 합성곱 신경망의 상기 추정 세그먼테이션을 획득하는 의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 9

제 1 항에 있어서,

상기 의사 트라이맵 추정 모듈은,

상기 추정 세그먼테이션과 상기 바운딩 박스를 그라운드 트루스 마스크(ground truth mask)의 노이즈 추정자로 이용하는 의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 10

제 1 항에 있어서,

상기 의사 트라이맵 추정 모듈은,

상기 바운딩 박스의 외부를 상기 배경 영역으로 추정하고 상기 바운딩 박스의 내부를 상기 전경 영역과 상기 그레이 영역으로 구분하는 의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 11

제 10 항에 있어서,

상기 의사 트라이맵 추정 모듈은,

상기 바운딩 박스 내부 픽셀 중 합성곱 신경망이 배경이라고 예측한 부분을 상기 그레이 영역으로 추정하고, 상기 합성곱 신경망이 전경으로 예측한 부분을 상기 전경 영역으로 추정하는 의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 12

제 1 항에 있어서,

상기 의사 트라이맵 추정 모듈은,

상기 쿼리 세트에 대한 상기 트라이맵 리프리젠테이션에서 전경과 배경에 해당하는 픽셀들의 특징 벡터가 각각 전경과 배경 대표 특징 벡터와 가까워지도록 소프트맥스 크로스 엔트로피(softmax cross entropy)를 계산하여 합성곱 신경망을 업데이트하는 의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 13

제 12 항에 있어서,

상기 의사 트라이맵 추정 모듈은,

상기 바운딩 박스 내의 픽셀이 상기 배경으로 추정된 경우 상기 픽셀을 상기 그레이 영역으로 정의하고 상기 합성곱 신경망 업데이트 및 대표 특징 벡터 계산에서 제외하는 의미론적 영상 분할을 위한 데이터 학습 장치.

청구항 14

에피소드의 서포트 세트와 쿼리 세트의 라벨을 교정하여 추정 세그먼테이션을 획득하는 단계;

상기 추정 세그먼테이션을 이용하여 트라이맵 리프리젠테이션을 획득하는 단계; 및

상기 추정 세그먼테이션과 바운딩 박스를 이용하여 상기 트라이맵 리프리젠테이션에서 배경 영역, 전경 영역 및 그레이 영역을 추정하는 단계를 포함하는,

의미론적 영상 분할을 위한 데이터 학습 방법.

청구항 15

제 14 항에 있어서,

상기 추정 세그먼테이션을 획득하는 단계는,

상기 바운딩 박스가 주어진 메타 학습 과정 동안 상기 서포트 세트의 교정된 대표 특징 벡터를 구하는 단계; 및

상기 대표 특징 벡터와 픽셀 별 어텐션 계산을 통하여 상기 쿼리 세트에 대한 상기 추정 세그먼테이션을 구하는 단계를 포함하는,

의미론적 영상 분할을 위한 데이터 학습 방법.

청구항 16

제 14 항에 있어서,

상기 트라이맵 리프리젠테이션을 획득하는 단계는,

상기 서포트 세트, 상기 쿼리 세트 및 베이스라인에서 구해진 상기 쿼리 세트의 초기 세그먼테이션을 기초로 목표 도메인과 참고 도메인을 교차하여 짝수 번 쌓는 단계를 포함하는 의미론적 영상 분할을 위한 데이터 학습 방법.

청구항 17

제 14 항에 있어서,

상기 트라이맵 리프리젠테이션에서 배경 영역, 전경 영역 및 그레이 영역을 추정하는 단계는,

상기 추정 세그먼테이션과 상기 바운딩 박스를 그라운드 트루스 마스크(ground truth mask)의 노이즈 추정자로 이용하는 단계를 포함하는 의미론적 영상 분할을 위한 데이터 학습 방법.

청구항 18

제 14 항에 있어서,

상기 트라이맵 리프리젠테이션에서 배경 영역, 전경 영역 및 그레이 영역을 추정하는 단계는,

상기 바운딩 박스의 외부를 상기 배경 영역으로 추정하는 단계;

상기 바운딩 박스 내부 픽셀 중 합성곱 신경망이 배경이라고 예측한 부분을 상기 그레이 영역으로 추정하는 단계; 및

상기 합성곱 신경망이 전경으로 예측한 부분을 상기 전경 영역으로 추정하는 단계를 포함하는 의미론적 영상 분할을 위한 데이터 학습 방법.

청구항 19

제 14 항에 있어서,

상기 쿼리 세트에 대한 상기 트라이맵 리프리젠테이션에서 전경과 배경에 해당하는 픽셀들의 특징 벡터가 각각 전경과 배경 대표 특징 벡터와 가까워지도록 소프트맥스 크로스 엔트로피(softmax cross entropy)를 계산하여 합성곱 신경망을 업데이트하는 단계를 더 포함하는 의미론적 영상 분할을 위한 데이터 학습 방법.

청구항 20

제 19 항에 있어서,

상기 바운딩 박스 내의 픽셀이 상기 배경으로 추정된 경우 상기 픽셀을 상기 그레이 영역으로 정의하고 상기 합성곱 신경망 업데이트 및 대표 특징 벡터 계산에서 제외하는 단계를 더 포함하는 의미론적 영상 분할을 위한 데이터 학습 방법.

발명의 설명

기술 분야

[0001] 본 명세서에 개시되는 실시예는 데이터 학습 장치 및 방법에 관한 것으로, 더 상세하게는 의미론적 영상 분할을 위한 데이터 학습 장치 및 방법에 관한 것이다.

배경 기술

[0002] 최근 들어, 인공지능이 발전됨에 따라 다양한 분야에서 인공지능을 적용하고 있다. 이때, 인공지능의 최적의 성능을 내기 위해서는 인공지능이 처리할 작업에 대해 최적화 되어야 한다.

[0003] 이를 위해, 인공지능이 처리할 작업과 관련된 정보를 이용하여 인공지능을 학습하는 과정이 필요하며, 효율적인 인공지능의 학습을 위해 메타학습 방법이 이용되고 있다.

[0004] 메타 학습이란 메타데이터를 사용하여 학습하는 과정을 학습하는 기계 학습(machine learning)으로 복수의 작업을 학습하여 새로운 작업을 일반화(generalize)하여 학습을 수행함으로써 학습 효율을 높인다.

발명의 내용

해결하려는 과제

[0005] 메타 학습의 목표는 소량의 라벨링된 학습 정보인 메타데이터를 이용하여 새로운 작업을 해결할 수 있도록 모델을 학습시키는 것이다.

[0006] 퓨-샷 세그먼테이션(Few-shot segmentation)은 적은 수의 샘플만 보고도 일반화 시킬 수 있는 능력을 갖추도록 학습하는 기법인 퓨-샷 학습(few-shot learning)을 세그먼테이션으로 확장한 것이다.

[0007] 이러한 퓨-샷 세그먼테이션 방법은 적은 수의 마스크와 샘플만 보고 세그먼테이션을 수행한다. 픽셀 단위의 라벨링이 들어가야 하는 세그먼테이션 마스크(segmentation mask)의 특성상 세그먼테이션을 위한 큰 규모의 데이터세트(dataset)를 제작하기 힘든 데, 퓨-샷 세그먼테이션을 이용하여 관심 있는 물체의 사진 몇 장만 라벨링하면 퓨-샷 세그먼테이션을 수행할 수 있다.

[0008] 퓨-샷 세그먼테이션은 메타 학습에 활용되기 위한 클래스(객체의 종류)들에 대하여 많은 수의 세그먼테이션 마스크가 라벨링 되어있는 이미지를 필요로 한다. 따라서 퓨-샷 세그먼테이션은 비록 관심 있는 클래스에 대하여 몇 장의 라벨링된 이미지가 필요함에도, 여전히 비슷한 분포를 따르는 클래스에 대하여 많은 수의 라벨이 필요하다는 단점이 있다.

[0009] 약-지도 세그먼테이션(Weakly-supervised segmentation)은 다른 태스크(task)에 비해 특히 세그먼테이션을 위한 라벨은 주체마다 라벨링의 결과가 다를 수 있으며, 물체의 위치를 픽셀 단위로 정교하게 마스크해야 하기 때문에 라벨링에 많은 비용과 시간이 소모된다는 단점이 있다.

[0010] 이에 세그먼테이션을 학습하는 데에 필요한 라벨인 마스크보다 라벨링 비용이 적으며 담고 있는 정보량이 적은 위크(weak) 라벨을 이용하여 세그먼테이션을 학습하는 약-지도 세그먼테이션에 대한 연구가 진행되고 있다.

- [0011] 기존의 약-지도 세그먼테이션은 관심 있는 클래스가 많은 수의 이미지에 대하여 미리 학습되어 있어야만 세그먼테이션을 수행 가능하다. 그러나, 오엘이디(OLED) 결함 검출 등의 문제에서는 결함의 종류에 따라 발생 빈도가 달라지는 클래스 불균형(class imbalance)이 있을 수 있어서, 특정 클래스에 대하여는 충분한 수의 학습 샘플을 확보하기 어려울 수 있다.
- [0012] 따라서, 본 실시예가 해결하고자 하는 과제는 메타 학습 동안 바운딩 박스(bounding box)만 접근 가능한 한정된 상황에서 적은 수의 샘플에서부터 세그먼테이션을 수행하는 능력을 학습하여 기존에 보지 못한 클래스에 대해서도 적은 수의 샘플로부터 세그먼테이션 마스크를 추론할 수 있는 의미론적 영상 분할을 위한 데이터 학습 장치 및 방법을 제공하는데 있다.
- [0013] 또한, 본 실시예가 해결하고자 하는 과제는 새로운 클래스의 이미지와 그 중 일부의 위치 정보가 주어졌을 때, 바운딩 박스에 의해 주어진 대략적인 위치 정보만을 이용하여 강인한 대표 특징 벡터를 뽑아서 세그먼테이션을 하도록 학습할 수 있는 의미론적 영상 분할을 위한 데이터 학습 장치 및 방법을 제공하는데 있다.
- [0014] 본 명세서의 일 실시예에 따른 해결 과제들은 이상에서 언급한 과제들로 제한되지 않으며, 언급되지 않은 또 다른 과제들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0015] 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치는, 에피소드의 서포트 세트와 쿼리 세트의 라벨을 교정하여 추정 세그먼테이션을 획득하고, 상기 추정 세그먼테이션을 이용하여 트라이맵 리프리젠테이션(trimap representation)을 획득하는 트라이맵 어텐션 모듈(trimap attention module); 및 상기 추정 세그먼테이션과 바운딩 박스를 이용하여 상기 트라이맵 리프리젠테이션에서 배경 영역(background zone), 전경 영역(foreground) 및 그레이 영역(gray zone)을 추정하는 의사 트라이맵 추정 모듈(Pseudo trimap estimation module)을 포함한다.
- [0016] 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 방법은, 에피소드의 서포트 세트와 쿼리 세트의 라벨을 교정하여 추정 세그먼테이션을 획득하는 단계; 상기 추정 세그먼테이션을 이용하여 트라이맵 리프리젠테이션을 획득하는 단계; 및 상기 추정 세그먼테이션과 바운딩 박스를 이용하여 상기 트라이맵 리프리젠테이션에서 배경 영역, 전경 영역 및 그레이 영역을 추정하는 단계를 포함한다.

발명의 효과

- [0017] 실시예에 따르면, 메타 학습 동안 바운딩 박스만 접근 가능한 한정된 상황에서 적은 수의 샘플에서부터 세그먼테이션을 수행하는 능력을 학습하여 기존에 보지 못한 클래스에 대해서도 적은 수의 샘플로부터 세그먼테이션 마스크를 추론할 수 있다.
- [0018] 또한, 새로운 클래스의 이미지와 그 중 일부의 위치 정보가 주어졌을 때 바운딩 박스에 의해 주어진 대략적인 위치 정보만을 이용하여 강인한 대표 특징 벡터를 뽑아서 세그먼테이션을 하도록 학습할 수 있다.
- [0019] 또한, 메타 학습 과정에서 바운딩 박스를 이용하는 퓨-샷 세그먼테이션 프레임워크를 제시할 수 있다. 이는 라벨링 코스트 측면에서 효율적이므로 오엘이디 디스플레이 결함 검출, 개인화된 컴퓨터이셔널 포토그래피 등의 실생활 어플리케이션으로 확장을 용이하게 할 있다.
- [0020] 또한, 바운딩 박스에 포함된 배경 부분이 전경에 포함되어 발생하는 배경 노이즈를 고려한 메타 학습 기법을 제시할 수 있다. 이는 메타 학습 과정에서 효과적인 트라이맵 표현을 추정할 수 있다. 따라서 백그라운드 노이즈에 강인하게 학습할 수 있다.
- [0021] 또한, 배경 노이즈가 존재하는 것을 고려하며 메타 학습을 진행하여 기 학습되지 않은 테스트 클래스에 대해서도 퓨-샷 세그먼테이션 성능을 향상시킬 수 있다.
- [0022] 본 명세서의 효과는 이상에서 언급한 효과에 제한되지 않으며, 언급되지 않은 또 다른 효과는 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [0023] 도 1은 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치의 베이스라인에서 메타 학습하는 과정을 나타낸다.

도 2는 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치의 베이스라인에서 추론하는 과정을 나타낸다.

도 3은 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치의 전체 프레임워크에서 메타 학습하는 과정을 나타낸다.

도 4는 도 3에 도시된 트라이맵-어텐션 모듈에서 베이스 라인 구조의 초기 예측을 이용하여 에피소드 서포트 세트와 쿼리 세트의 라벨을 교정하는 과정을 나타낸다.

도 5는 도 4에 도시된 트라이맵-어텐션 블록에서 바운딩 박스의 트라이맵 리프레젠테이션을 얻는 과정을 나타낸다.

도 6은 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치에서 의사 트라이맵을 추정하는 과정을 설명하기 위한 도면이다.

도 7은 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치에서 메타 학습 과정에서의 의사 트라이맵 추정 결과를 나타낸다.

도 8은 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치에서 테스트 클래스에서의 퓨-샷 세그먼테이션 결과를 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0024] 본 명세서의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 명세서는 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이며, 단지 본 실시예들은 본 명세서의 개시가 완전하도록 하며, 본 명세서가 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 명세서는 청구항의 범주에 의해 정의될 뿐이다.
- [0025] 본 명세서의 실시예를 설명하기 위한 도면에 개시된 형상, 크기, 비율, 각도, 개수 등은 예시적인 것이므로 본 명세서가 도시된 사항에 한정되는 것은 아니다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다. 또한, 본 명세서를 설명함에 있어서, 관련된 공지 기술에 대한 구체적인 설명이 본 명세서의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명은 생략한다. 본 명세서 상에서 언급된 '포함한다', '갖는다', '이루어진다' 등이 사용되는 경우 '~만'이 사용되지 않는 이상 다른 부분이 추가될 수 있다. 구성 요소를 단수로 표현한 경우에 특별히 명시적인 기재 사항이 없는 한 복수를 포함하는 경우를 포함한다.
- [0026] 구성 요소를 해석함에 있어서, 별도의 명시적 기재가 없더라도 오차 범위를 포함하는 것으로 해석한다.
- [0027] 위치 관계에 대한 설명일 경우, 예를 들어, '~상에', '~상부에', '~하부에', '~옆에' 등으로 두 부분의 위치 관계가 설명되는 경우, '바로' 또는 '직접'이 사용되지 않는 이상 두 부분 사이에 하나 이상의 다른 부분이 위치할 수도 있다.
- [0028] 시간 관계에 대한 설명일 경우, 예를 들어, '~후에', '~에 이어서', '~다음에', '~전에' 등으로 시간 적 선후 관계가 설명되는 경우, '바로' 또는 '직접'이 사용되지 않는 이상 연속적이지 않은 경우도 포함할 수 있다.
- [0029] 제1, 제2 등이 다양한 구성요소들을 서술하기 위해서 사용되나, 이들 구성요소들은 이들 용어에 의해 제한되지 않는다. 이들 용어들은 단지 하나의 구성요소를 다른 구성요소와 구별하기 위하여 사용하는 것이다. 따라서, 이하에서 언급되는 제1 구성요소는 본 명세서의 기술적 사상 내에서 제2 구성요소일 수도 있다.
- [0030] 본 명세서의 여러 실시예들의 각각 특징들이 부분적으로 또는 전체적으로 서로 결합 또는 조합 가능하고, 기술적으로 다양한 연동 및 구동이 가능하며, 각 실시예들이 서로에 대하여 독립적으로 실시 가능할 수도 있고 연관 관계로 함께 실시할 수도 있다.
- [0031] 이하에서는, 적은 수의 샘플에서부터 세그먼테이션을 수행하는 능력을 학습하여 기준에 보지 못한 클래스에 대해서도 적은 수의 샘플로부터 세그먼테이션 마스크를 추론할 수 있는 의미론적 영상 분할을 위한 데이터 학습 장치 및 방법이 개시된다.
- [0032] 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치 및 방법을 설명하기에 앞서, 본 명세서에서 사용되는 용어들의 의미를 정의한다.

- [0033] 세그먼테이션(Segmentation)은 이미지로부터 물체의 영역을 구분하는 작업으로 정의될 수 있고, 퓨-샷 학습(Few-shot learning)은 적은 수의 샘플을 이용하여 일반화 시키는 학습 기법으로 정의될 수 있다.
- [0034] 약-지도 세그먼테이션(Weakly-supervised Segmentation)은 세그먼테이션을 학습하는 데에 필요한 라벨(label)인 마스크(mask)보다 라벨링 비용이 적으며 담고 있는 정보량이 적은 약한(weak) 라벨을 이용하여 세그먼테이션을 학습하는 기법으로 정의될 수 있다.
- [0035] 서포트 세트(support set)는 추론 시간(inference time)에 참조할 적어도 하나의 학습 예제를 포함하고 있는 참조 세트로 정의될 수 있고, 쿼리 세트(query set)는 현재 세그먼테이션을 수행하려고 하는 대상 영상으로 정의될 수 있다.
- [0036] 트라이맵(trimap) 추정은 쿼리 세트에서 전경 영역(foreground zone), 미지의 영역(unknown zone) 및 배경 영역(background zone)을 추정하는 것으로 정의될 수 있다. 여기에서, 미지의 영역은 그레이 영역(gray zone)으로 명명될 수 있다.
- [0037] 에피소드는 작업에 대한 분류를 대표하는 데이터로 데이터세트에서 샘플링된 예시인 쿼리 세트와 쿼리 세트에 대응되는 예시를 포함하는 서포트 세트의 모음으로 정의될 수 있다.
- [0038] 명세서에서, 모듈 및 블록은 적어도 하나의 프로세서 또는 적어도 하나의 프로그램으로 정의될 수 있다.
- [0039] 도 1은 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치의 베이스라인에서 메타 학습하는 과정을 나타낸다. 도 2는 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치의 베이스라인에서 추론하는 과정을 나타낸다.
- [0040] 베이스라인은 메타 학습 과정에서 바운딩 박스(bounding box)를 세그먼테이션 마스크로 간주한다.
- [0041] 먼저, 서포트 세트의 특징 벡터와 바운딩 박스 간의 마스크된 평균 풀링(masked average pooling)을 통해 전경 및 배경(foreground and background) 대표 특징 벡터를 구한다.
- [0042] 다음으로, 쿼리 세트의 픽셀 별 특징 벡터와 전경 및 배경 대표 특징 벡터간의 거리에 반비례하도록 클래스 확률(class probability)을 구한다.
- [0043] 추정 세그먼테이션을 구할 때에는 픽셀 별로 더 가까운 대표 특징 벡터를 갖는 쪽으로 분류하여 마스크(mask)를 계산한다. 에피소드 학습법의 특성(메타 학습 과정에서 테스트 상황을 모방하여 학습함)에 따라, 추론하는 과정은 메타 학습 과정과 동일하다.
- [0044] 의미론적 영상 분할을 위한 데이터 학습 장치에 적용되는 알고리즘은, 먼저 바운딩 박스 라벨에서 확실한 부분과 불확실한 부분을 구분하는 의사 트라이맵 추정을 실시한다.
- [0045] 이 과정에서 합성곱 신경망(Convolutional Neural Network, CNN)의 추론 값과 바운딩 박스를 실제 세그먼테이션 마스크(segmentation mask)에 대한 두 독립된 추정자(estimator)로 가정한다.
- [0046] 그 후 학습 과정에서 불확실한 부분을 제하고 확실한 부분만을 이용하는 방식으로 디노이징을 실시하며 개선된 퓨-샷 세그먼테이션(few-shot segmentation) 성능을 얻을 수 있다.
- [0047] 합성곱 신경망은 시각적 영상을 분석하는데 사용되는 다층의 피드-포워드적인 인공신경망 중 하나이다. 딥 러닝에서 심층 신경망으로 분류되며, 시각적 영상 분석에 주로 적용된다. 또한 공유 가중치 구조와 변환 불변성 특성에 기초하여 변이 불변 또는 공간 불변 인공 신경망으로 이용될 수 있다. 영상 및 동영상 인식, 추천 시스템, 영상 분류, 의료 영상 분석 및 자연어 처리 등에 응용된다.
- [0048] 합성곱 신경망은 정규화 된 버전의 다층 퍼셉트론이다. 다층 퍼셉트론은 일반적으로 완전히 연결된 네트워크, 즉 한 계층의 각 뉴런이 다음 계층의 모든 뉴런에 연결되는 신경망 구조이다. 이와 같이 네트워크가 완전 연결된 경우 주어진 데이터에 과적합 되는 경향이 있다.
- [0049] 일반적인 정규화를 위해 최적화 함수에 특정 척도를 추가하는 방법이 흔히 쓰이지만, 합성곱 신경망은 정규화를 위한 다른 접근 방식을 취한다. 데이터에서 계층적 패턴을 활용하고 더 작고 간단한 패턴을 사용하여 더 복잡한 패턴을 표현함으로써 정규화와 같은 효과를 내는 것이다. 따라서 합성곱 신경망의 연결 구조의 복잡성은 유사한 기능의 다층 퍼셉트론에 비해 극단적으로 낮다.
- [0050] 합성곱 신경망을 이용한 영상 분류는 다른 영상 분류 알고리즘에 비해 상대적으로 전처리를 거의 사용하지 않는

다. 이는 신경망이 기존 알고리즘에서 수작업으로 제작된 필터를 학습한다는 것을 의미한다. 영상 분류 알고리즘에서 설계자가 영상의 특징들을 미리 이해해 알고리즘을 만드는 과정이 없는 것이 합성곱 신경망의 주요한 장점이다.

- [0051] 합성곱 신경망은 크게 합성곱층(Convolution layer)과 풀링층(Pooling layer)으로 구성될 수 있다.
- [0052] 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치는 가중치를 공유하는 제1 합성곱 신경망 및 제2 합성곱 신경망을 포함할 수 있다. 일례로, 제1 합성곱 신경망은 서포트 세트에 대한 데이터를 처리할 수 있고, 제2 합성곱 신경망은 쿼리 세트에 대한 데이터를 처리할 수 있다.
- [0053] 도 3은 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치(100)의 전체 프레임워크에서 메타 학습하는 과정을 나타낸다.
- [0054] 바운딩 박스가 주어진 메타 학습 과정 동안, 트라이맵 어텐션 모듈(trimap attention module)(110)을 이용하여 서포트 세트의 교정된 대표 특징 벡터를 구한다.
- [0055] 구해진 대표 특징 벡터와 픽셀 별 어텐션(attention) 계산을 통하여 쿼리 이미지에 대한 초기 세그먼테이션을 구하고 이를 이용하여 쿼리 세트의 바운딩 박스에 대한 트라이맵 리프리젠테이션을 구한다.
- [0056] 그리고 쿼리의 트라이맵에서 전경과 배경에 해당하는 픽셀들의 특징 벡터가 각각 전경과 배경 대표 특징 벡터와 가까워지도록 소프트맥스 크로스 엔트로피(softmax cross entropy)를 계산하여 합성곱 신경망을 업데이트 한다. 테스트(추론) 과정은 베이스라인과 동일하다.
- [0057] 도 4는 도 3에 도시된 트라이맵-어텐션 모듈에서 베이스 라인 구조의 초기 예측을 이용하여 에피소드 서포트 세트와 쿼리 세트의 라벨을 교정하는 과정을 나타낸다.
- [0058] 트라이맵-어텐션 모듈(trimap attention module, 110)은 베이스라인 구조의 초기 예측을 이용하여 에피소드의 서포트 세트와 쿼리 세트의 라벨을 반복적 기법으로 교정한다. 트라이맵-어텐션 모듈의 입력으로 서포트 세트와 쿼리 세트, 및 베이스라인 구조에서 얻어진 쿼리 세트의 초기 세그먼테이션이 입력될 수 있다.
- [0059] 이때 트라이맵-어텐션 블록(Trimap-attention block, 112)의 목표 도메인과 참고 도메인을 교차하여 짝수 번 씩 음으로써 서포트 세트와 쿼리 세트의 트라이맵 리프리젠테이션을 얻을 수 있다. 이로써 두 도메인이 서로를 교정하는 참고(reference)로 작용함으로써 데이터 세트의 바운딩 박스에서 오는 배경 노이즈(background noise)를 교정한다.
- [0060] 도 5는 도 4에 도시된 트라이맵-어텐션 블록에서 바운딩 박스의 트라이맵 리프리젠테이션을 얻는 과정을 나타낸다.
- [0061] 트라이맵-어텐션 블록(Trimap-attention block, 112)은 바운딩 박스의 트라이맵 리프리젠테이션(trimap representation)을 얻기 위해 합성곱 신경망의 추정 세그먼테이션을 얻는 블록이다.
- [0062] 이때 에피소드 학습법 하에서 추정 세그먼테이션을 얻기 위해서는 서포트 세트와 쿼리 세트가 필요하기 때문에, 쿼리 세트로 쓰일 목표 도메인에 추가적으로 서포트 세트로 기능하고 대표 특징 벡터를 제공할 참고 도메인이 필요하다. 이때 참고 도메인은 초기 컨볼루션 신경망 세그먼테이션을 가지고 있는 것으로 가정한다.
- [0063] 트라이맵-어텐션 블록(112) 다음과 같이 동작한다. 먼저 의사 트라이맵 추정자(pseudo trimap estimator)를 이용하여 참고 도메인의 라벨을 트라이맵 리프리젠테이션(trimap representation)으로 나타낸다.
- [0064] 그 후, 그에 따른 대표 특징 벡터를 계산한다. 전경과 배경의 대표 특징 벡터는 트라이맵 상에서 각각 전경 영역과 배경 영역에 포함된 픽셀의 특징 벡터를 평균하는 방식으로 계산된다(masked average pooling).
- [0065] 이렇게 목표 도메인을 이용하여 구해진 대표 특징 벡터와 목표 도메인의 픽셀 별 특징 벡터와의 어텐션(attention)을 구하여 더 어텐션이 높은 쪽으로 분류하면 목표 도메인의 픽셀 별 영역 추정, 즉 합성곱 신경망의 추정 세그먼테이션이 얻어진다.
- [0066] 도 6은 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치에서 의사 트라이맵을 추정하는 과정을 설명하는 도면이다.
- [0067] 의사 트라이맵 추정 모듈(Pseudo trimap estimation module, 120)은 추정 세그먼테이션과 바운딩 박스를 이용하여 트라이맵 리프리젠테이션에서 배경 영역, 전경 영역 및 그레이 영역을 추정한다.

- [0068] 기본적인 가정은 바운딩 박스와 합성곱 신경망의 추정 세그먼테이션을 그라운드 트루스 마스크(ground truth mask)의 독립적인 노이지 추정자(noisy estimator)로 보는 것이다.
- [0069] 이때, 바운딩 박스의 바깥 부분은 확실한 배경 영역으로 가정하고, 바운딩 박스의 내부를 전경과 그레이 영역(gray zone)으로 나누어 트라이맵 리프리젠테이션(trimap representation)의 영역을 추정한다.
- [0070] 바운딩 박스 내부 픽셀 중 합성곱 신경망이 배경이라고 예측한 픽셀은 그레이 영역(gray zone), 그리고 합성곱 신경망이 전경(foreground)로 예측한 부분은 전경 영역(foreground zone)으로 나타낸다.
- [0071] 도 6에서 이어지는 그림은 팬더의 바운딩 박스와 그라운드 트루스(ground truth segmentation)이 들어왔을 때의 의사 트라이맵 추정(pseudo trimap estimation)을 나타낸 것으로 검정색은 배경 영역(background zone), 흰색은 전경 영역(foreground zone), 그리고 회색은 그레이 영역(gray zone)을 나타낸다.
- [0072] 만약 바운딩 박스 내의 픽셀이 배경으로 추정된 경우(false negative error) (A) 해당 픽셀이 바운딩 박스에 포함된 배경이거나 (B) 합성곱 신경망의 예측이 틀리고 픽셀이 실제 전경에 속한 경우로 나뉜다. (B)의 경우는 특별한 조치가 필요하지 않지만, (A)인 경우, 이 픽셀에 대하여 합성곱 신경망의 파라미터를 업데이트하는 것이 합성곱 신경망의 학습을 방해한다.
- [0073] 따라서 이 경우를 방지하기 위하여 바운딩 박스 내의 픽셀이 배경으로 추정된 경우 발생한 픽셀들을 그레이 영역으로 정의하고 합성곱 신경망 업데이트 및 대표 특징 벡터 계산에서 제외한다.
- [0074] 도 7은 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치에서 메타 학습 과정에서의 의사 트라이맵 추정 결과를 나타낸다. 도 8은 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치에서 테스트 클래스에서의 퓨-샷 세그먼테이션 결과를 나타낸다.
- [0075] 아래는 파스칼(Pascal-5ⁱ) 데이터셋에서 테스트 클래스에서의 퓨-샷 세그먼테이션 결과(정량적)를 나타낸다. 표 1은 원-샷 시나리오(one-shot scenario)를 나타내고, 표 2는 파이브-샷 시나리오(five-shot scenario)를 나타낸다.

[표 1]

mIoU (%)	Split-0	Split-1	Split-2	Split-3	mean
Baseline	36.74	51.89	46.63	37.03	43.07
Ours	36.96	52.24	49.06	35.23	43.37
Δ	+0.22	+0.35	+2.43	-1.80	+0.30

[0077]

[표 2]

mIoU (%)	Split-0	Split-1	Split-2	Split-3	mean
Baseline	45.83	57.62	56.06	41.70	50.30
Ours	46.48	58.99	58.19	42.97	51.66
Δ	+0.65	+1.37	+2.13	+1.27	+1.36

[0079]

[0080] 실시예에 따르면, 메타 학습 동안 바운딩 박스만 접근 가능한 한정된 상황에서 적은 수의 샘플에서부터 세그먼테이션을 수행하는 능력을 학습하여 기존에 보지 못한 클래스에 대해서도 적은 수의 샘플로부터 세그먼테이션 마스크를 추론할 수 있다.

[0081] 또한, 새로운 클래스의 이미지와 그 중 일부의 위치 정보가 주어졌을 때 바운딩 박스에 의해 주어진 대략적인 위치 정보만을 이용하여 강한 대표 특징 벡터를 뽑아서 세그먼테이션을 하도록 학습할 수 있다.

[0082] 또한, 메타 학습 과정에서 바운딩 박스를 이용하는 퓨-샷 세그먼테이션 프레임워크를 제시할 수 있다. 이는 라벨링 코스트 측면에서 효율적이므로 오엘이디 디스플레이 결함 검출, 개인화된 컴퓨터이셔널 포토그래피 등의 실생활 어플리케이션으로 확장을 용이하게 할 있다.

[0083] 또한, 바운딩 박스에 포함된 배경 부분이 전경에 포함되어 발생하는 배경 노이즈를 고려한 메타 학습 기법을 제

시할 수 있다. 이는 메타 학습 과정에서 효과적인 트라이맵 표현을 추정할 수 있다. 따라서 백그라운드 노이즈에 강인하게 학습할 수 있다.

- [0084] 또한, 배경 노이즈가 존재하는 것을 고려하며 메타 학습을 진행하여 기 학습되지 않은 테스트 클래스에 대해서도 퓨-샷 세그멘테이션 성능을 향상시킬 수 있다.
- [0085] 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 장치(100)는, 에피소드의 서포트 세트와 쿼리 세트의 라벨을 교정하여 추정 세그멘테이션을 획득하고 추정 세그멘테이션을 이용하여 트라이맵 리프리젠테이션을 획득하는 트라이맵 어텐션 모듈(110), 및 추정 세그멘테이션과 바운딩 박스를 이용하여 트라이맵 리프리젠테이션에서 배경 영역, 전경 영역 및 그레이 영역을 추정하는 의사 트라이맵 추정 모듈(120)을 포함한다.
- [0086] 트라이맵 어텐션 모듈(110)은 바운딩 박스가 주어진 메타 학습 과정 동안 서포트 세트의 교정된 대표 특징 벡터를 구하고 대표 특징 벡터와 픽셀 별 어텐션 계산을 통하여 쿼리 세트에 대한 추정 세그멘테이션을 구한다.
- [0087] 트라이맵 어텐션 모듈(110)은 추정 세그멘테이션을 이용하여 쿼리 세트의 바운딩 박스에 대한 트라이맵 리프리젠테이션을 구한다.
- [0088] 트라이맵 어텐션 모듈(110)은 베이스라인의 초기 예측을 이용하여 에피소드의 서포트 세트와 쿼리 세트의 라벨을 반복적 기법으로 교정한다.
- [0089] 트라이맵 어텐션 모듈(110)은 합성곱 신경망의 추정 세그멘테이션을 구하기 위한 적어도 하나의 트라이맵-어텐션 블록(112)을 포함한다.
- [0090] 트라이맵-어텐션 블록(112)은 서포트 세트, 쿼리 세트 및 베이스라인에서 구해진 쿼리 세트의 초기 세그멘테이션을 기초로 목표 도메인과 참고 도메인을 교차하여 짝수 번 쌓음으로써 서포트 세트와 쿼리 세트에 대한 트라이맵 리프리젠테이션을 구한다.
- [0091] 트라이맵-어텐션 블록(112)은 목표 도메인과 참고 도메인을 서로를 교정하는 참고로 설정하여 데이터 세트의 바운딩 박스에서 발생하는 배경 노이즈를 교정한다.
- [0092] 트라이맵-어텐션 블록(112)은 참고 도메인의 라벨을 트라이맵 리프리젠테이션으로 나타내고, 트라이맵 리프리젠테이션 상에서 전경 영역과 배경 영역에 포함된 픽셀의 특징 벡터를 평균하는 방식으로 대표 특징 벡터를 계산한다. 그리고, 트라이맵-어텐션 블록(112)은 대표 특징 벡터와 목표 도메인의 픽셀 별 특징 벡터와의 어텐션을 구하여 목표 도메인의 픽셀 별 영역을 추정하여 합성곱 신경망의 추정 세그멘테이션을 획득한다.
- [0093] 의사 트라이맵 추정 모듈(120)은 추정 세그멘테이션과 바운딩 박스를 그라운드 트루스 마스크의 노이즈 추정자료로 이용한다.
- [0094] 의사 트라이맵 추정 모듈(120)은 바운딩 박스의 외부를 배경 영역으로 추정하고 바운딩 박스의 내부를 전경 영역과 그레이 영역으로 구분한다.
- [0095] 의사 트라이맵 추정 모듈(120)은 바운딩 박스 내부 픽셀 중 합성곱 신경망이 배경이라고 예측한 부분을 그레이 영역으로 추정하고 합성곱 신경망이 전경으로 예측한 부분을 전경 영역으로 추정한다.
- [0096] 의사 트라이맵 추정 모듈(120)은 쿼리 세트에 대한 트라이맵 리프리젠테이션에서 전경과 배경에 해당하는 픽셀들의 특징 벡터가 각각 전경과 배경 대표 특징 벡터와 가까워지도록 소프트맥스 크로스 엔트로피를 계산하여 합성곱 신경망을 업데이트한다.
- [0097] 의사 트라이맵 추정 모듈(120)은 바운딩 박스 내의 픽셀이 배경으로 추정된 경우 픽셀을 그레이 영역으로 정의하고 합성곱 신경망 업데이트 및 대표 특징 벡터 계산에서 제외한다.
- [0098] 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 방법은, 에피소드의 서포트 세트와 쿼리 세트의 라벨을 교정하여 추정 세그멘테이션을 획득하는 단계, 추정 세그멘테이션을 이용하여 트라이맵 리프리젠테이션을 획득하는 단계, 및 추정 세그멘테이션과 바운딩 박스를 이용하여 트라이맵 리프리젠테이션에서 배경 영역, 전경 영역 및 그레이 영역을 추정하는 단계를 포함한다.
- [0099] 추정 세그멘테이션을 획득하는 단계는, 바운딩 박스가 주어진 메타 학습 과정 동안 서포트 세트의 교정된 대표 특징 벡터를 구하는 단계, 대표 특징 벡터와 픽셀 별 어텐션 계산을 통하여 쿼리 세트에 대한 추정 세그멘테이션을 구하는 단계를 포함한다.
- [0100] 트라이맵 리프리젠테이션을 획득하는 단계는, 서포트 세트, 상기 쿼리 세트 및 베이스라인에서 구해진 상기 쿼리

리 세트의 초기 세그먼테이션을 기초로 목표 도메인과 참고 도메인을 교차하여 짝수 번 쌓는 단계를 포함한다.

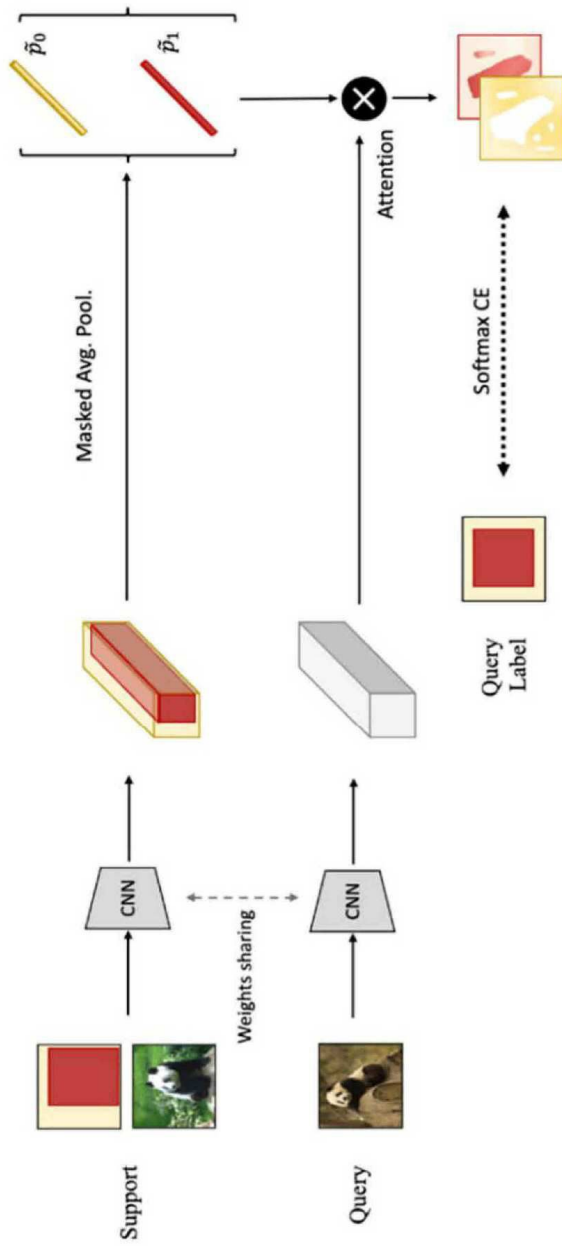
- [0101] 트라이맵 리프리젠테이션에서 배경 영역, 전경 영역 및 그레이 영역을 추정하는 단계는, 추정 세그먼테이션과 바운딩 박스를 그래운드 트루스 마스크의 노이즈 추정자로 이용하는 단계를 포함한다.
- [0102] 트라이맵 리프리젠테이션에서 배경 영역, 전경 영역 및 그레이 영역을 추정하는 단계는, 바운딩 박스의 외부를 상기 배경 영역으로 추정하는 단계, 바운딩 박스 내부 픽셀 중 합성곱 신경망이 배경이라고 예측한 부분을 그레이 영역으로 추정하는 단계, 및 합성곱 신경망이 전경으로 예측한 부분을 전경 영역으로 추정하는 단계를 포함한다.
- [0103] 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 방법은, 쿼리 세트에 대한 트라이맵 리프리젠테이션에서 전경과 배경에 해당하는 픽셀들의 특징 벡터가 각각 전경과 배경 대표 특징 벡터와 가까워지도록 소프트맥스 크로스 엔트로피를 계산하여 합성곱 신경망을 업데이트하는 단계를 더 포함한다.
- [0104] 일 실시예에 따른 의미론적 영상 분할을 위한 데이터 학습 방법은, 바운딩 박스 내의 픽셀이 배경으로 추정된 경우 픽셀을 그레이 영역으로 정의하고 합성곱 신경망 업데이트 및 대표 특징 벡터 계산에서 제외하는 단계를 더 포함한다.
- [0105] 이상 첨부된 도면을 참조하여 본 명세서의 실시예들을 더욱 상세하게 설명하였으나, 본 명세서는 반드시 이러한 실시예로 국한되는 것은 아니고, 본 명세서의 기술사상을 벗어나지 않는 범위 내에서 다양하게 변형 실시될 수 있다. 따라서, 본 명세서에 개시된 실시예들은 본 명세서의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 명세서의 기술 사상의 범위가 한정되는 것은 아니다. 그러므로, 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 본 명세서의 보호 범위는 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 명세서의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

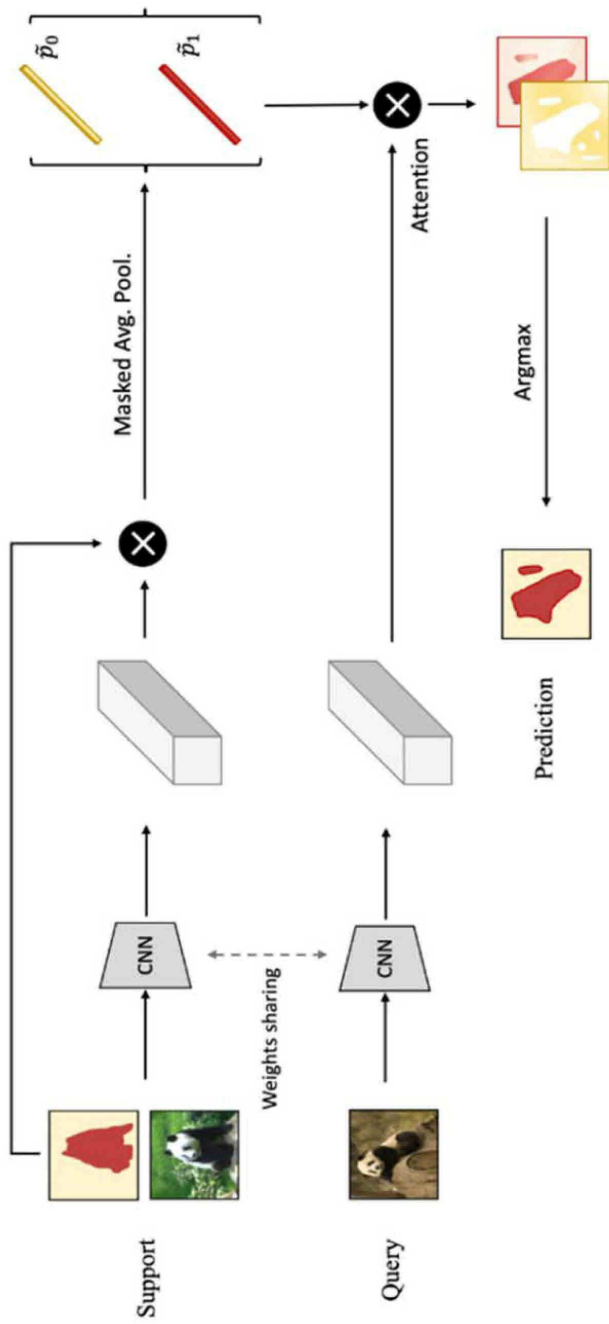
- [0106] 100: 메타 학습 장치
- 110: 트라이맵 어텐션 모듈
- 112: 트라이맵 어텐션 블록
- 120: 의사 트라이맵 추정 모듈
- CNN: 합성곱 신경망

도면

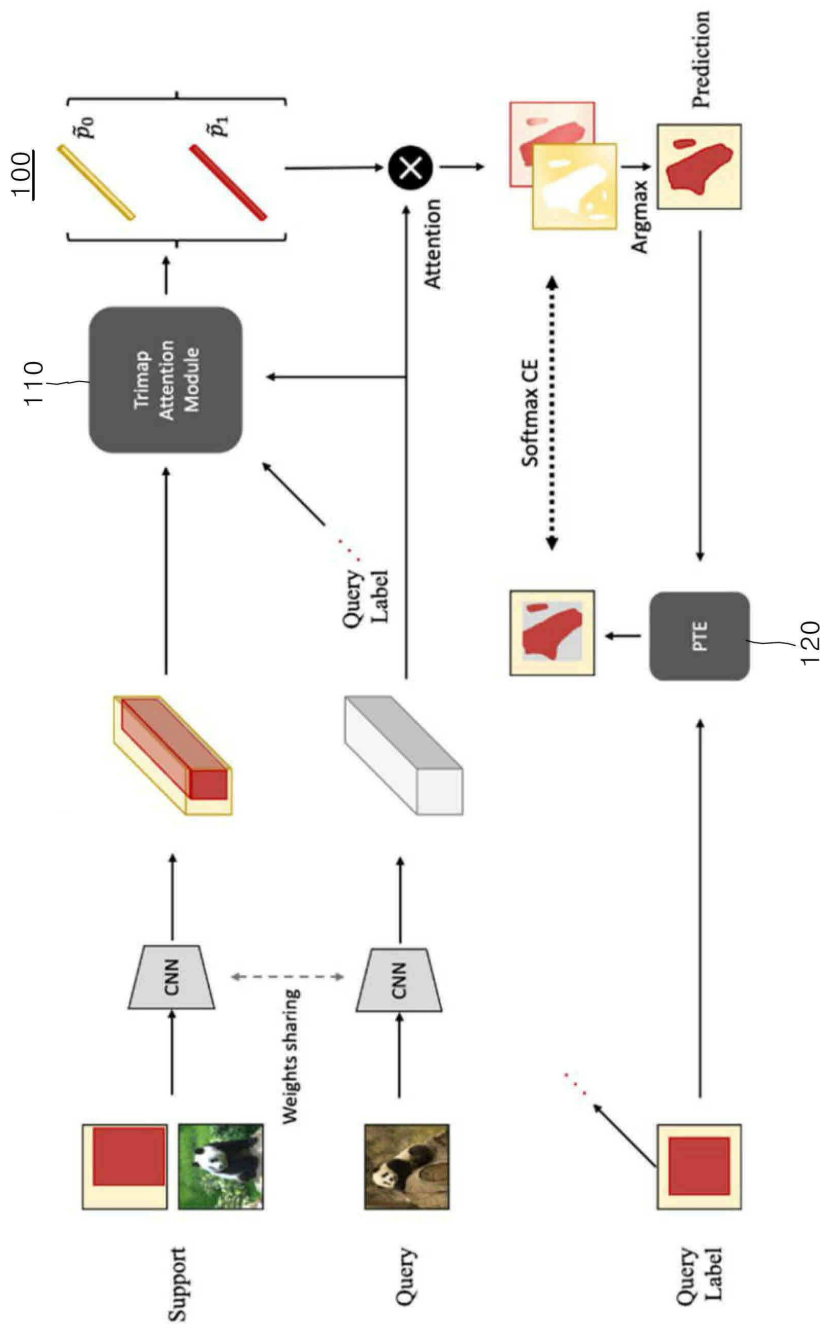
도면1



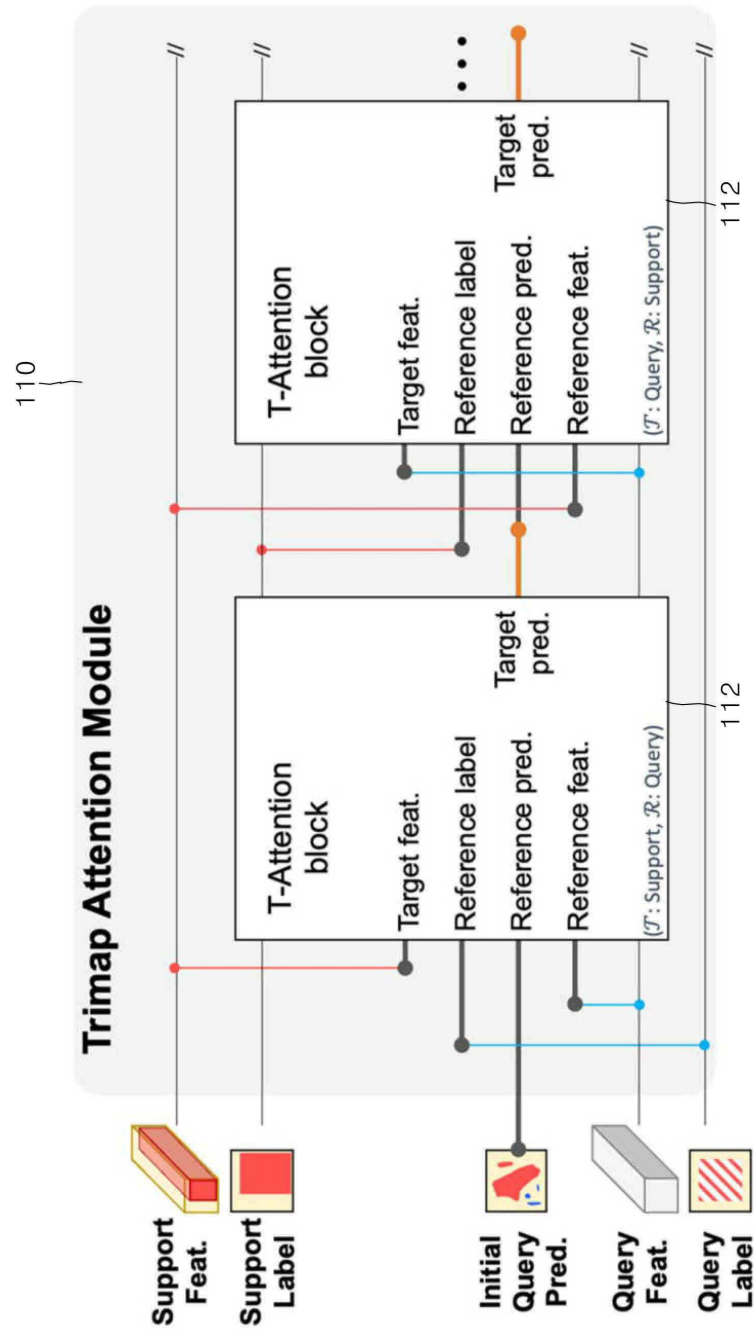
도면2



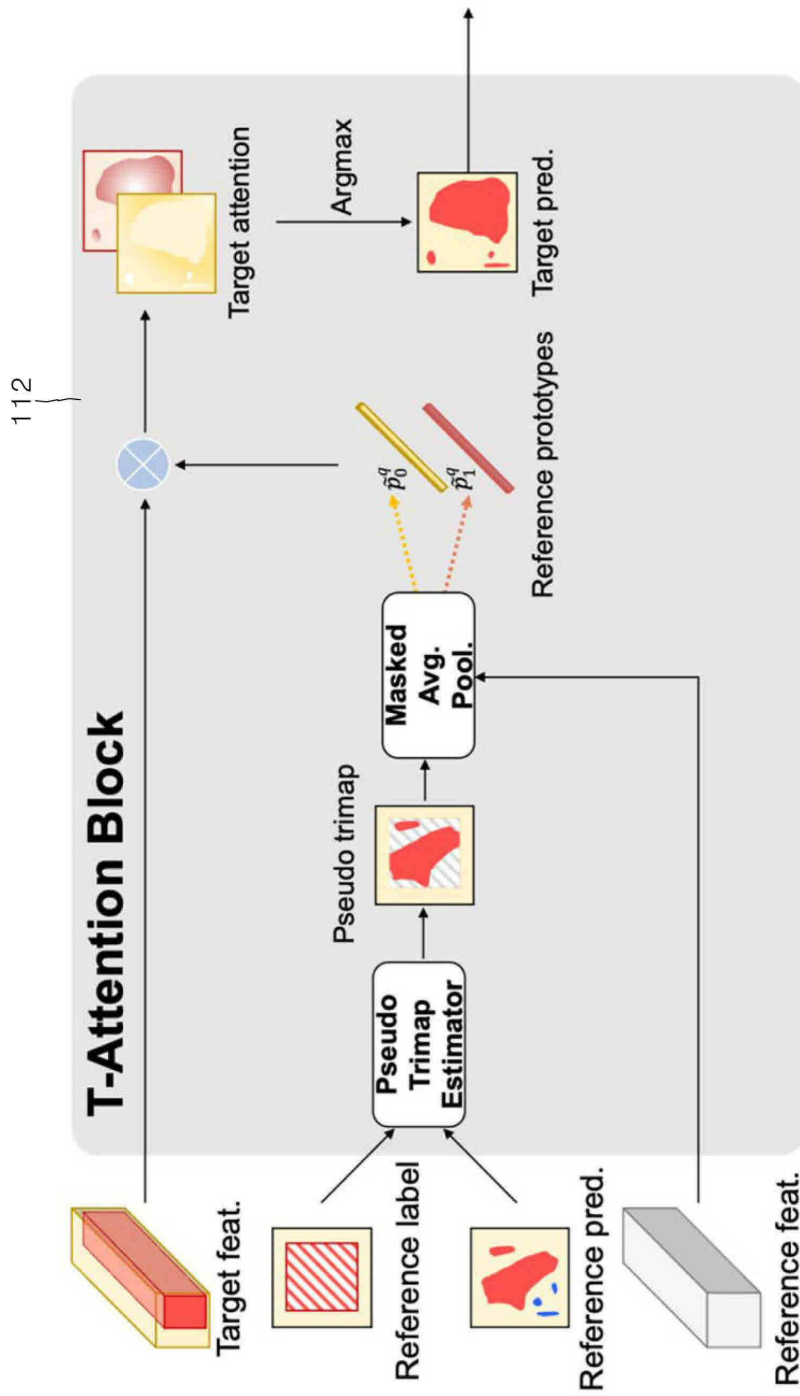
도면3



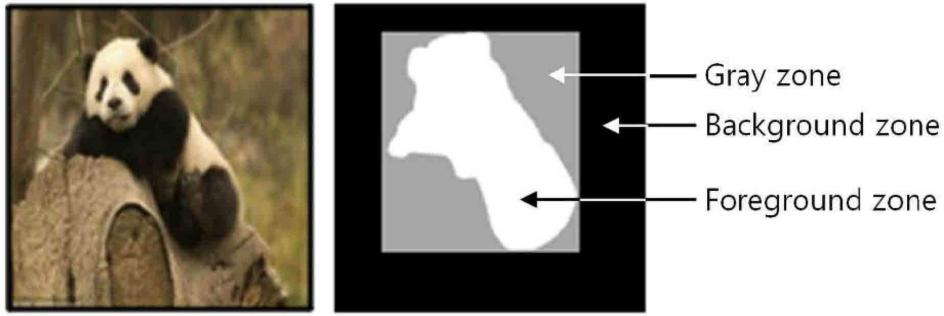
도면4



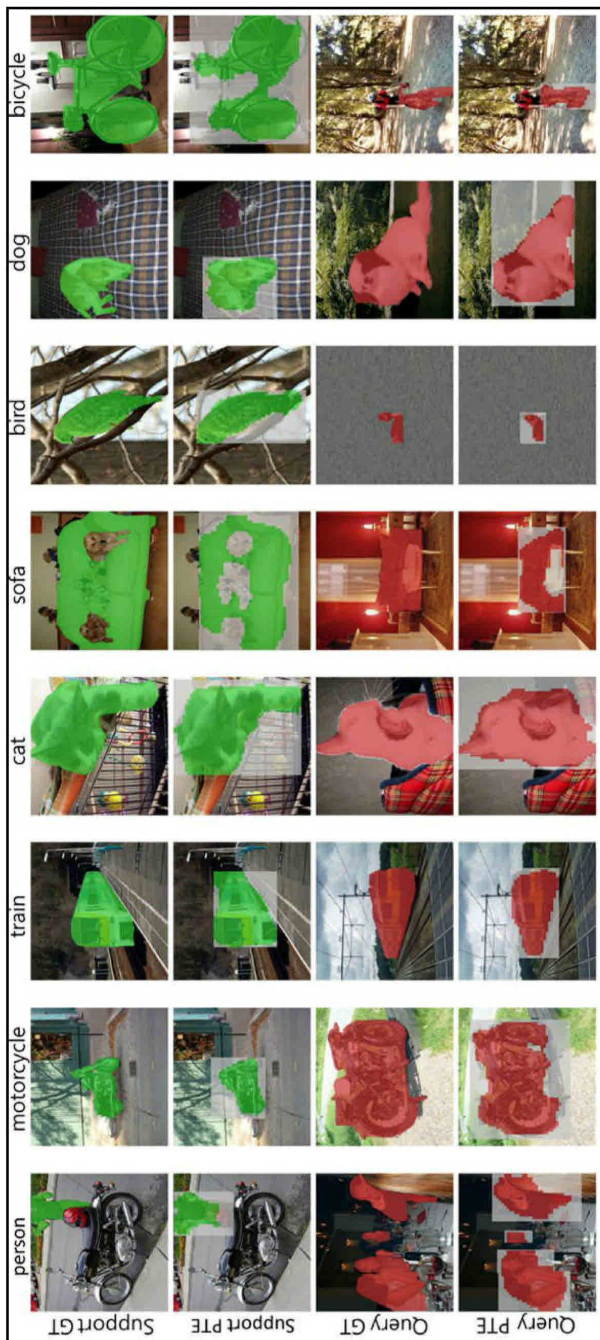
도면5



도면6



도면7



도면8

