(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0114320
(43) 공개일자 2022년08월17일

(51) 국제특허분류(Int. Cl.)
G06N 5/02 (2006.01) *G06N 20/00* (2019.01)
G06T 7/194 (2017.01) *G06T 7/20* (2017.01)
(52) CPC특허분류
G06N 5/02 (2019.01)
G06N 20/00 (2021.08)
(21) 출원번호 10-2021-0017666
(22) 출원일자 2021년02월08일
심사청구일자 2021년02월08일

(71) 출원인
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
함범섭
서울특별시 강남구 압구정로61길 37, 72동 506호 (압구정동, 한양아파트)
오영민
서울특별시 마포구 성미산로32길 10, 201호(연남동, 그린하우스)
김범준
서울특별시 서초구 잠원로14길 32, 102동 402호(잠원동, 래미안 신반포 리오센트)
(74) 대리인
민영준

전체 청구항 수 : 총 20 항

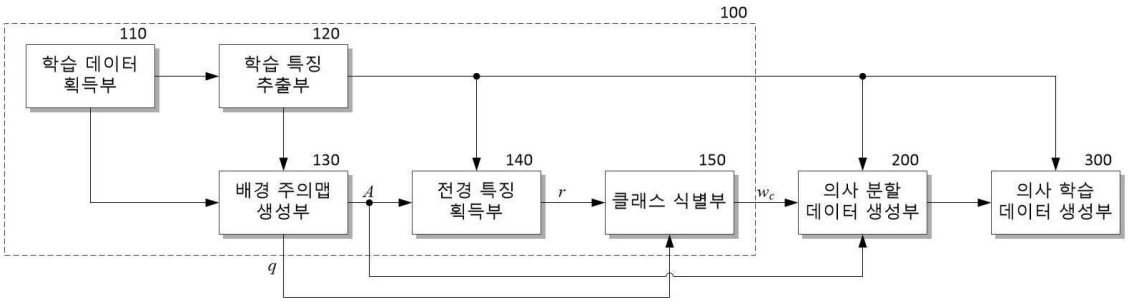
(54) 발명의 명칭 **약지도 학습 기반 시멘틱 영상 분할 학습 데이터 생성 장치 및 방법**

(57) 요 약

본 발명은 학습 영상에 포함된 객체 영역이 경계 박스 형태로 레이블된 경계 박스 수준의 학습 데이터를 인가받아 미리 학습된 방식에 따라 학습 영상의 특징을 추출하여 학습 특징맵을 획득하고, 특징맵과 경계 박스를 이용하여, 학습 특징맵에서 경계 박스 내부 영역 중 경계 박스 외부의 배경 영역과 유사한 영역을 나타내는 배경 주

(뒷면에 계속)

대 표 도



의맵을 획득하고, 학습 특징맵과 배경 주의맵으로부터 경계 박스 내부의 전경 영역에 대한 특징인 전경 특징을 획득하며, 전경 특징을 학습에 의해 미리 지정된 가중치에 따라 분류하여 객체 및 배경 클래스를 식별하는 클래스 분할부 및 가중치 중 객체 클래스 각각의 식별에 이용된 개별 가중치를 학습 특징맵에 가중하고 기지정된 방식으로 경계 박스에 대응하도록 정규화하여 객체 클래스 주의맵을 획득하고, 객체 클래스 주의맵과 배경 주의맵을 미리 학습된 방식에 따라 조합하여 시멘틱 영상 분할 장치를 지도 학습시키기 위한 가상의 픽셀 수준 학습 데이터인 의사 분할 데이터를 생성하는 의사 분할 데이터 생성부를 포함하여, 저비용의 약지도 학습 데이터인 경계 박스 수준 학습 데이터로부터 픽셀 수준 학습 데이터와 유사한 의사 픽셀 수준 학습 데이터를 획득할 수 있는 시멘틱 영상 분할 학습 데이터 생성 장치 및 방법을 제공할 수 있다.

(52) CPC특허분류

G06T 7/194 (2017.01)

G06T 7/20 (2013.01)

G06T 2207/20081 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711108723
과제번호	2019R1A2C2084816
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	중견연구자지원사업
연구과제명	딥러닝 한계 극복을 통한 탄력적 시멘틱 영상 분할 연구(2/3)(2019.9.1~2022.2.28)
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2020.03.01 ~ 2021.02.28

명 세 서

청구범위

청구항 1

학습 영상에 포함된 객체 영역이 경계 박스 형태로 레이블된 경계 박스 수준의 학습 데이터를 인가받아 미리 학습된 방식에 따라 상기 학습 영상의 특징을 추출하여 학습 특징맵을 획득하고, 상기 특징맵과 상기 경계 박스를 이용하여, 상기 학습 특징맵에서 경계 박스 내부 영역 중 경계 박스 외부의 배경 영역과 유사한 영역을 나타내는 배경 주의맵을 획득하고, 상기 학습 특징맵과 상기 배경 주의맵으로부터 상기 경계 박스 내부의 전경 영역에 대한 특징인 전경 특징을 획득하며, 상기 전경 특징을 학습에 의해 미리 지정된 가중치에 따라 분류하여 객체 및 배경 클래스를 식별하는 클래스 분할부; 및

상기 가중치 중 상기 객체 클래스 각각의 식별에 이용된 개별 가중치를 상기 학습 특징맵에 가중하고 기지정된 방식으로 상기 경계 박스에 대응하도록 정규화하여 객체 클래스 주의맵을 획득하고, 상기 객체 클래스 주의맵과 상기 배경 주의맵을 미리 학습된 방식에 따라 조합하여 시멘틱 영상 분할 장치를 지도 학습시키기 위한 가상의 픽셀 수준 학습 데이터인 의사 분할 데이터를 생성하는 의사 분할 데이터 생성부를 포함하는 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 2

제1항에 있어서, 상기 클래스 분할부는

상기 경계 박스 수준의 학습 데이터를 획득하는 학습 데이터 획득부;

인공 신경망으로 구현되어, 미리 학습된 방식에 따라 상기 학습 데이터의 특징을 추출하여 상기 학습 특징맵을 획득하는 특징 추출부;

상기 경계 박스를 이용하여 마스크를 생성하고, 상기 마스크를 이용하여 상기 특징맵에서 경계 박스 외부 배경 영역에 대한 배경 특징을 획득하고, 획득된 배경 특징을 기반으로 상기 특징맵의 상기 경계 박스 내부 영역 중 배경 특징과 유사한 영역을 나타내는 상기 배경 주의맵을 생성하는 배경 주의맵 생성부;

상기 특징맵의 상기 경계 박스 내부 영역 중 상기 배경 주의맵에 의해 배경으로 지정되는 영역을 제거하여 전경 영역을 획득하고, 전경 영역의 특징을 기지정된 방식으로 계산하여 전경 특징을 획득하는 전경 특징 획득부; 및

인공 신경망으로 구현되어, 미리 학습된 방식에 따라 전경 특징과 배경 특징 각각에 대응하는 클래스를 식별하는 클래스 식별부를 포함하는 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 3

제2항에 있어서, 상기 배경 주의맵 생성부는

상기 학습 데이터에 레이블된 경계 박스 영역을 기반으로 상기 마스크를 획득하는 마스크 획득부;

상기 특징맵에 상기 마스크를 적용하고, 상기 마스크가 적용된 특징맵을 기지정된 크기의 다수의 그리드로 구분하고, 구분된 다수의 그리드 각각에 대한 다수의 배경 특징을 추출하는 배경 특징 추출부;

상기 다수의 배경 특징 각각과 상기 학습 특징맵의 경계 박스 영역 사이의 유사도를 기반으로 상기 학습 특징맵의 경계 박스 영역에서 다수의 배경 특징 각각에 대응하는 영역을 추출하여 다수의 배경 상관맵을 획득하는 배경 상관맵 획득부; 및

상기 다수의 배경 상관맵을 누적 평균하여 상기 배경 주의맵을 획득하는 배경 주의맵 획득부를 포함하는 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 4

제3항에 있어서, 상기 배경 특징 추출부는

다수의 그리드(G(j)) 각각에 대한 다수의 배경 특징(q_j)을 수학식

$$q_j = \frac{\sum_{\mathbf{p}\in G(j)} M(\mathbf{p}) f(\mathbf{p})}{\sum_{\mathbf{p}\in G(j)} M(\mathbf{p})}$$

(여기서 M(p)와 f(p)는 마스크(M)와 학습 특징맵(f)의 픽셀을 나타낸다.)

에 따라 획득하고,

상기 배경 상관맵 획득부는

상기 다수의 배경 특징(q_j) 각각에 대응하는 다수의 배경 상관맵(A_j)을 수학식

$$A_j(\mathbf{p}) = \begin{cases} \mathrm{ReLU}\left(\frac{f(\mathbf{p})}{\|f(\mathbf{p})\|} \cdot \frac{q_j}{\|q_j\|}\right) & , \mathbf{p} \in \mathcal{B} \\ 1 & , \mathbf{p} \notin \mathcal{B} \end{cases}$$

(여기서 B는 경계 박스를 나타내고, $\|\cdot\|$는 L_2 놈 함수를 나타내며, ReLU(Rectified Linear Unit)는 정류 함수를 나타낸다.)

에 따라 획득하며,

상기 배경 주의맵 획득부는

상기 배경 주의맵(A)을 수학식

$$A(\mathbf{p}) = \frac{1}{J}\sum_j A_j(\mathbf{p})$$

에 따라 획득하는 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 5

제3항에 있어서, 상기 전경 특징 획득부는

상기 배경 주의맵을 픽셀 반전하고, 픽셀 반전된 배경 주의맵 각각을 상기 학습 특징맵을 가중하여 다수의 전경 주의맵을 획득하는 전경 영역 강조부; 및

상기 전경 주의맵의 특징을 집계하여 각 경계 박스에 대한 특징인 전경 특징을 추출하는 전경 특징 추출부를 포함하는 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 6

제5항에 있어서, 상기 전경 특징 획득부는

각 경계 박스(B_i)에 대한 상기 배경 주의맵(A)과 상기 학습 특징맵(f)으로부터 상기 전경 특징(r_i)을 수학식

$$r_i = \frac{\sum_{\mathbf{p}\in B_i} (1 - A(\mathbf{p})) f(\mathbf{p})}{\sum_{\mathbf{p}\in B_i} (1 - A(\mathbf{p}))}$$

에 따라 계산하여 획득하는 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 7

제5항에 있어서, 상기 의사 분할 데이터 생성부는

상기 가중치 중 상기 객체 클래스 각각의 식별에 이용된 개별 가중치를 상기 학습 특징맵에 가중하여 식별된 객체 클래스 각각에 대한 클래스 활성맵을 획득하는 클래스 활성맵 획득부;

상기 클래스 활성맵을 대응하는 경계 박스에 대응하여 정규화하여 객체 클래스 주의맵을 획득하는 객체 클래스 주의맵 획득부; 및

상기 객체 클래스 주의맵과 상기 배경 주의맵을 조합하여 상기 의사 분할 데이터를 획득하는 분할 데이터 획득부를 포함하는 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 8

제7항에 있어서, 상기 클래스 활성맵 획득부는

식별된 각 클래스(c)에 대한 상기 클래스 활성맵(CAM_c)을 수학식

$$\mathbf{CAM}_c(\mathbf{p}) = \mathbf{ReLU}(f(\mathbf{p}) \cdot w_c)$$

(여기서 f(p)는 학습 특징맵의 픽셀, w_c는 클래스(c) 식별에 이용된 가중치를 나타낸다.)

에 따라 계산하여 획득하고,

상기 객체 클래스 주의맵 획득부는

식별된 각 클래스(c)에 대한 상기 객체 클래스 주의맵(u_c)을 수학식

$$u_c(\mathbf{p}) = \begin{cases} \frac{\mathbf{CAM}_c(\mathbf{p})}{\max_{\mathbf{p}}(\mathbf{CAM}_c(\mathbf{p}))} & , \mathbf{p} \in \mathcal{B}_c \\ 0 & , \mathbf{p} \notin \mathcal{B}_c \end{cases}$$

(여기서 B_c 는 내부에 포함된 객체 클래스(c)가 식별된 경계 박스)

에 따라 획득하는 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 9

제7항에 있어서, 상기 분할 데이터 획득부는

상기 객체 클래스 주의맵과 상기 배경 주의맵을 밀집 CFR(dense CRF: dense Conditional Random Fields) 기법에 따라 조합하여 상기 의사 분할 데이터를 획득하는 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 10

제7항에 있어서, 상기 시멘틱 영상 분할 학습 데이터 생성 장치는

상기 의사 분할 데이터의 각 클래스 영역에 대한 클래스 특징을 추출하고, 추출된 클래스 특징과 상기 학습 특징맵 사이의 유사도를 기반으로 상기 학습 특징맵의 각 클래스별 영역을 다시 재구분하여, 상기 의사 분할 데이터와 함께 상기 시멘틱 영상 분할 장치를 지도 학습시키기 위한 가상의 픽셀 수준 학습 데이터인 의사 학습 데이터를 생성하는 의사 학습 데이터 생성부를 더 포함하는 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 11

제10항에 있어서, 상기 의사 학습 데이터 생성부는

상기 의사 분할 데이터를 인가받아 각 클래스별 영역을 구분하여 추출하는 클래스 영역 추출부;

상기 학습 특징맵에서 구분되어 추출된 각 클래스 영역에 대응하는 영역의 클래스 특징을 계산하여 획득하는 클래스 특징 획득부; 및

상기 학습 특징맵과 계산된 클래스 특징 사이의 유사도를 계산하여 클래스 상관맵을 획득하고, 클래스 상관맵에 따라 상기 학습 특징맵에서 클래스 특징에 유사한 영역을 구분하여 상기 의사 학습 데이터를 획득하는 의사 학

습 데이터 획득부를 포함하는 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 12

제11항에 있어서, 상기 클래스 특징 획득부는

상기 클래스 영역(Q_c)에 대응하는 영역의 클래스 특징(q_c)을 수학식

$$q_c = \frac{1}{|\mathcal{Q}_c|} \sum_{\mathbf{p} \in \mathcal{Q}_c} f(\mathbf{p})$$

(여기서 $|Q_c|$는 클래스 영역(Q_c)에 대응하는 학습 특징맵(f) 영역의 픽셀 개수)

에 따라 계산하고,

상기 의사 학습 데이터 획득부는

상기 클래스 상관맵(C_c)을 수학식

$$C_c(\mathbf{p}) = \frac{f(\mathbf{p})}{\|f(\mathbf{p})\|} \cdot \frac{q_c}{\|q_c\|}$$

(여기서 $\|\cdot\|$는 L_2 놈 함수)

에 따라 계산하여, 상기 클래스 상관맵(C_c)에 argmax 함수를 적용하여 상기 의사 학습 데이터를 획득하는 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 13

제1항에 있어서, 상기 클래스 분할부는

상기 경계 박스 수준의 학습 데이터를 인가받아 약지도 학습 방식으로 미리 학습된 시멘틱 영상 분할 학습 데이터 생성 장치.

청구항 14

약지도 학습 방식으로 미리 학습된 인공 신경망을 포함하여, 학습 영상에 포함된 객체 영역이 경계 박스 형태로 레이블된 경계 박스 수준의 학습 데이터를 인가받아 미리 학습된 방식에 따라 상기 학습 영상의 특징을 추출하여 학습 특징맵을 획득하고, 상기 특징맵과 상기 경계 박스를 이용하여, 상기 학습 특징맵에서 경계 박스 내부 영역 중 경계 박스 외부의 배경 영역과 유사한 영역을 나타내는 배경 주의맵을 획득하고, 상기 학습 특징맵과 상기 배경 주의맵으로부터 상기 경계 박스 내부의 전경 영역에 대한 특징인 전경 특징을 획득하며, 상기 전경 특징을 학습에 의해 미리 지정된 가중치에 따라 분류하여 객체 및 배경 클래스를 식별하는 단계; 및

상기 가중치 중 상기 객체 클래스 각각의 식별에 이용된 개별 가중치를 상기 학습 특징맵에 가중하고 기지정된 방식으로 상기 경계 박스에 대응하도록 정규화하여 객체 클래스 주의맵을 획득하고, 상기 객체 클래스 주의맵과 상기 배경 주의맵을 미리 학습된 방식에 따라 조합하여 시멘틱 영상 분할 방법을 지도 학습시키기 위한 가상의 픽셀 수준 학습 데이터인 의사 분할 데이터를 생성하는 단계를 포함하는 시멘틱 영상 분할 학습 데이터 생성 방법.

청구항 15

제14항에 있어서, 상기 클래스를 식별하는 단계는

상기 경계 박스 수준의 학습 데이터를 획득하는 단계;

미리 학습된 인공 신경망을 이용하여, 학습된 방식에 따라 상기 학습 데이터의 특징을 추출하여 상기 학습 특징맵을 획득하는 단계;

상기 경계 박스를 이용하여 마스크를 생성하고, 상기 마스크를 이용하여 상기 특징맵에서 경계 박스 외부 배경

영역에 대한 배경 특징을 획득하고, 획득된 배경 특징을 기반으로 상기 특징맵의 상기 경계 박스 내부 영역 중 배경 특징과 유사한 영역을 나타내는 상기 배경 주의맵을 생성하는 단계;

상기 특징맵의 상기 경계 박스 내부 영역 중 상기 배경 주의맵에 의해 배경으로 지정되는 영역을 제거하여 전경 영역을 획득하고, 전경 영역의 특징을 기지정된 방식으로 계산하여 전경 특징을 획득하는 단계; 및

인공 신경망으로 구현되어, 미리 학습된 방식에 따라 전경 특징과 배경 특징 각각의 클래스를 판단하는 단계를 포함하는 시멘틱 영상 분할 학습 데이터 생성 방법.

청구항 16

제15항에 있어서, 상기 배경 주의맵을 생성하는 단계는

상기 학습 데이터에 레이블된 경계 박스 영역을 기반으로 상기 마스크를 획득하는 단계;

상기 특징맵에 상기 마스크를 적용하고, 상기 마스크가 적용된 특징맵을 기지정된 크기의 다수의 그리드로 구분하고, 구분된 다수의 그리드 각각에 대한 다수의 배경 특징을 추출하는 단계;

상기 다수의 배경 특징 각각과 상기 학습 특징맵의 경계 박스 영역 사이의 유사도를 기반으로 상기 학습 특징맵의 경계 박스 영역에서 다수의 배경 특징 각각에 대응하는 영역을 추출하여 다수의 배경 상관맵을 획득하는 단계; 및

상기 다수의 배경 상관맵을 누적 평균하여 상기 배경 주의맵을 획득하는 단계를 포함하는 시멘틱 영상 분할 학습 데이터 생성 방법.

청구항 17

제15항에 있어서, 상기 전경 특징을 획득하는 단계는

상기 배경 주의맵을 픽셀 반전하고, 픽셀 반전된 배경 주의맵 각각을 상기 학습 특징맵을 가중하여 다수의 전경 주의맵을 획득하는 단계; 및

상기 전경 주의맵의 특징을 집계하여 각 경계 박스에 대한 특징인 전경 특징을 추출하는 단계를 포함하는 시멘틱 영상 분할 학습 데이터 생성 방법.

청구항 18

제15항에 있어서, 상기 의사 분할 데이터를 생성하는 단계는

상기 가중치 중 상기 객체 클래스 각각의 식별에 이용된 개별 가중치를 상기 학습 특징맵에 가중하여 식별된 객체 클래스 각각에 대한 클래스 활성맵을 획득하는 단계;

상기 클래스 활성맵을 대응하는 경계 박스에 대응하여 정규화하여 객체 클래스 주의맵을 획득하는 단계; 및

상기 객체 클래스 주의맵과 상기 배경 주의맵을 조합하여 상기 의사 분할 데이터를 획득하는 단계를 포함하는 시멘틱 영상 분할 학습 데이터 생성 방법.

청구항 19

제18항에 있어서, 상기 시멘틱 영상 분할 학습 데이터 생성 방법은

상기 의사 분할 데이터를 생성하는 단계 이후, 상기 의사 분할 데이터의 각 클래스 영역에 대한 클래스 특징을 추출하고, 추출된 클래스 특징과 상기 학습 특징맵 사이의 유사도를 기반으로 상기 학습 특징맵의 각 클래스별 영역을 다시 재구분하여, 상기 의사 분할 데이터와 함께 상기 시멘틱 영상 분할 방법을 지도 학습시키기 위한 가상의 픽셀 수준 학습 데이터인 의사 학습 데이터를 생성하는 단계를 더 포함하는 시멘틱 영상 분할 학습 데이터 생성 방법.

청구항 20

제19항에 있어서, 상기 의사 학습 데이터를 생성하는 단계는

상기 의사 분할 데이터를 인가받아 각 클래스별 영역을 구분하여 추출하는 단계;

상기 학습 특징맵에서 구분되어 추출된 각 클래스 영역에 대응하는 영역의 클래스 특징을 계산하여 획득하는 단계; 및

상기 학습 특징맵과 계산된 클래스 특징 사이의 유사도를 계산하여 클래스 상관맵을 획득하고, 클래스 상관맵에 따라 상기 학습 특징맵에서 클래스 특징에 유사한 영역을 구분하여 의사 학습 데이터를 획득하는 단계를 포함하는 시멘틱 영상 분할 학습 데이터 생성 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 시멘틱 영상 분할 학습 데이터 생성 장치 및 방법에 관한 것으로, 약지도 학습 기반 시멘틱 영상 분할 학습 데이터 생성 장치 및 방법에 관한 것이다.

배 경 기 술

[0002] 시멘틱 영상 분할(semantic segmentation)은 입력 영상을 식별 가능한 클래스 각각에 대응하는 영역별로 분할하는 것을 의미하며, 자율 주행, 의료 영상, 영상 편집 등 다양한 응용 분야에 적용될 수 있다. 이러한 시멘틱 영상 분할은 입력 영상의 다수의 픽셀 각각을 사람, 자동차, 자전거 등과 같은 객체를 지정된 클래스로 분류하여 레이블링하는 것을 목표로 한다.

[0003] CNN(convolutional neural networks) 과 같은 인공 신경망을 이용하는 딥 러닝 기반 시멘틱 영상 분할 기술들은 우수한 성능을 나타내지만, 학습을 위해서는 각 객체의 클래스가 픽셀 단위로 레이블되어 클래스별 객체 영역이 정확하게 표현된 픽셀 수준(pixel-level)의 학습 데이터가 대량으로 필요하다. 그러나 픽셀 수준의 학습 데이터는 사람의 수작업으로 획득되어야 하므로 고비용이 소요되는 문제가 있다.

[0004] 이에 영상에 포함된 각 객체 영역을 픽셀 수준에서 정확하게 레이블하지 않고서도 학습을 수행할 수 있도록 하는 약지도 학습(Weakly Supervised learning) 방식이 제안되었다. 시멘틱 영상 분할에서 기존의 약지도 학습 방식으로는 객체의 영역을 지정하지 않고 영상 내에 포함된 객체의 클래스만이 레이블링된 영상 수준(image-level) 학습 데이터를 이용하는 방법과 영상에서 각 클래스의 객체가 포함된 영역이 박스형으로 레이블링되는 경계 박스 수준(bounding box-level) 학습 데이터를 이용하는 방식이 주로 이용되고 있다.

[0005] 영상 수준 학습 데이터를 이용하는 기법에서는 영상에 포함된 객체의 클래스만을 레이블링하면되므로 학습 데이터를 저비용으로 매우 용이하게 획득할 수 있는 반면, 객체 영역을 정확하게 분할하기 어렵다는 한계가 있다.

[0006] 그리고 경계 박스 수준 학습 데이터를 이용하는 기법은 영상 수준 학습 데이터에 비해서는 높은 비용이 요구되지만, 픽셀 수준 학습 데이터에 비해서는 대략 1/15 수준의 저비용을 학습 데이터를 획득할 수 있으며, 영상에서 개략적일지라도 객체 영역이 지정됨으로써, 영상 수준 학습 데이터를 이용하는 경우에 비해 우수한 학습 성능을 획득할 수 있다. 그러나 경계 박스 수준 학습 데이터는 객체 영역과 배경 영역을 정확하게 구분하지 않으므로, 경계 박스 내부에는 객체에 해당하는 전경과 배경이 혼재하여 포함된다. 따라서 픽셀 수준 학습 데이터를 이용하는 경우에 비해 낮은 학습 성능을 나타내게 된다는 한계가 있다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 한국 공개 특허 제10-2020-0112646호 (2020.10.05 공개)

발명의 내용

해결하려는 과제

[0008] 본 발명의 목적은 저비용으로 높은 학습 성능을 갖는 학습 데이터를 획득할 수 있는 시멘틱 영상 분할 학습 데이터 생성 장치 및 방법을 제공하는데 있다.

[0009] 본 발명의 다른 목적은 경계 박스 수준의 약지도 학습 데이터를 기반으로 픽셀 수준 학습 데이터와 유사한 의사 픽셀 수준 학습 데이터를 용이하게 획득할 수 있는 시멘틱 영상 분할 학습 데이터 생성 장치 및 방법을 제공하

는데 있다.

과제의 해결 수단

[0010] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 장치는 학습 영상에 포함된 객체 영역이 경계 박스 형태로 레이블된 경계 박스 수준의 학습 데이터를 인가받아 미리 학습된 방식에 따라 상기 학습 영상의 특징을 추출하여 학습 특징맵을 획득하고, 상기 특징맵과 상기 경계 박스를 이용하여, 상기 학습 특징맵에서 경계 박스 내부 영역 중 경계 박스 외부의 배경 영역과 유사한 영역을 나타내는 배경 주의맵을 획득하고, 상기 학습 특징맵과 상기 배경 주의맵으로부터 상기 경계 박스 내부의 전경 영역에 대한 특징인 전경 특징을 획득하며, 상기 전경 특징을 학습에 의해 미리 지정된 가중치에 따라 분류하여 객체 및 배경 클래스를 식별하는 클래스 분할부; 및 상기 가중치 중 상기 객체 클래스 각각의 식별에 이용된 개별 가중치를 상기 학습 특징맵에 가중하고 기지정된 방식으로 상기 경계 박스에 대응하도록 정규화하여 객체 클래스 주의맵을 획득하고, 상기 객체 클래스 주의맵과 상기 배경 주의맵을 미리 학습된 방식에 따라 조합하여 시멘틱 영상 분할 장치를 지도 학습시키기 위한 가상의 픽셀 수준 학습 데이터인 의사 분할 데이터를 생성하는 의사 분할 데이터 생성부를 포함한다.

[0011] 상기 클래스 분할부는 상기 경계 박스 수준의 학습 데이터를 획득하는 학습 데이터 획득부; 인공 신경망으로 구현되어, 미리 학습된 방식에 따라 상기 학습 데이터의 특징을 추출하여 상기 학습 특징맵을 획득하는 특징 추출부; 상기 경계 박스를 이용하여 마스크를 생성하고, 상기 마스크를 이용하여 상기 특징맵에서 경계 박스 외부 배경 영역에 대한 배경 특징을 획득하고, 획득된 배경 특징을 기반으로 상기 특징맵의 상기 경계 박스 내부 영역 중 배경 특징과 유사한 영역을 나타내는 상기 배경 주의맵을 생성하는 배경 주의맵 생성부; 상기 특징맵의 상기 경계 박스 내부 영역 중 상기 배경 주의맵에 의해 배경으로 지정되는 영역을 제거하여 전경 영역을 획득하고, 전경 영역의 특징을 기지정된 방식으로 계산하여 전경 특징을 획득하는 전경 특징 획득부; 및 인공 신경망으로 구현되어, 미리 학습된 방식에 따라 전경 특징과 배경 특징 각각에 대응하는 클래스를 식별하는 클래스 식별부를 포함할 수 있다.

[0012] 상기 배경 주의맵 생성부는 상기 학습 데이터에 레이블된 경계 박스 영역을 기반으로 상기 마스크를 획득하는 마스크 획득부; 상기 특징맵에 상기 마스크를 적용하고, 상기 마스크가 적용된 특징맵을 기지정된 크기의 다수의 그리드로 구분하고, 구분된 다수의 그리드 각각에 대한 다수의 배경 특징을 추출하는 배경 특징 추출부; 상기 다수의 배경 특징 각각과 상기 학습 특징맵의 경계 박스 영역 사이의 유사도를 기반으로 상기 학습 특징맵의 경계 박스 영역에서 다수의 배경 특징 각각에 대응하는 영역을 추출하여 다수의 배경 상관맵을 획득하는 배경 상관맵 획득부; 및 상기 다수의 배경 상관맵을 누적 평균하여 상기 배경 주의맵을 획득하는 배경 주의맵 획득부를 포함할 수 있다.

[0013] 상기 전경 특징 획득부는 상기 배경 주의맵을 픽셀 반전하고, 픽셀 반전된 배경 주의맵 각각을 상기 학습 특징맵을 가중하여 다수의 전경 주의맵을 획득하는 전경 영역 강조부; 및 상기 전경 주의맵의 특징을 집계하여 각 경계 박스에 대한 특징인 전경 특징을 추출하는 전경 특징 추출부를 포함할 수 있다.

[0014] 상기 의사 분할 데이터 생성부는 상기 가중치 중 상기 객체 클래스 각각의 식별에 이용된 개별 가중치를 상기 학습 특징맵에 가중하여 식별된 객체 클래스 각각에 대한 클래스 활성맵을 획득하는 클래스 활성맵 획득부; 상기 클래스 활성맵을 대응하는 경계 박스에 대응하여 정규화하여 객체 클래스 주의맵을 획득하는 객체 클래스 주의맵 획득부; 및 상기 객체 클래스 주의맵과 상기 배경 주의맵을 조합하여 상기 의사 분할 데이터를 획득하는 분할 데이터 획득부를 포함할 수 있다.

[0015] 상기 시멘틱 영상 분할 학습 데이터 생성 장치는 상기 의사 분할 데이터의 각 클래스 영역에 대한 클래스 특징을 추출하고, 추출된 클래스 특징과 상기 학습 특징맵 사이의 유사도를 기반으로 상기 학습 특징맵의 각 클래스별 영역을 다시 재구분하여, 상기 의사 분할 데이터와 함께 상기 시멘틱 영상 분할 장치를 지도 학습시키기 위한 가상의 픽셀 수준 학습 데이터인 의사 학습 데이터를 생성하는 의사 학습 데이터 생성부를 더 포함할 수 있다.

[0016] 상기 의사 학습 데이터 생성부는 상기 의사 분할 데이터를 인가받아 각 클래스별 영역을 구분하여 추출하는 클래스 영역 추출부; 상기 학습 특징맵에서 구분되어 추출된 각 클래스 영역에 대응하는 영역의 클래스 특징을 계산하여 획득하는 클래스 특징 획득부; 및 상기 학습 특징맵과 계산된 클래스 특징 사이의 유사도를 계산하여 클래스 상관맵을 획득하고, 클래스 상관맵에 따라 상기 학습 특징맵에서 클래스 특징에 유사한 영역을 구분하여 상기 의사 학습 데이터를 획득하는 의사 학습 데이터 획득부를 포함할 수 있다.

[0017] 상기 클래스 분할부는 상기 경계 박스 수준의 학습 데이터를 인가받아 약지도 학습 방식으로 미리 학습될 수 있다.

[0018] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 방법은 약지도 학습 방식으로 미리 학습된 인공 신경망을 포함하여, 학습 영상에 포함된 객체 영역이 경계 박스 형태로 레이블된 경계 박스 수준의 학습 데이터를 인가받아 미리 학습된 방식에 따라 상기 학습 영상의 특징을 추출하여 학습 특징맵을 획득하고, 상기 특징맵과 상기 경계 박스를 이용하여, 상기 학습 특징맵에서 경계 박스 내부 영역 중 경계 박스 외부의 배경 영역과 유사한 영역을 나타내는 배경 주의맵을 획득하고, 상기 학습 특징맵과 상기 배경 주의맵으로부터 상기 경계 박스 내부의 전경 영역에 대한 특징인 전경 특징을 획득하며, 상기 전경 특징을 학습에 의해 미리 지정된 가중치에 따라 분류하여 객체 및 배경 클래스를 식별하는 단계; 및 상기 가중치 중 상기 객체 클래스 각각의 식별에 이용된 개별 가중치를 상기 학습 특징맵에 가중하고 기지정된 방식으로 상기 경계 박스에 대응하도록 정규화하여 객체 클래스 주의맵을 획득하고, 상기 객체 클래스 주의맵과 상기 배경 주의맵을 미리 학습된 방식에 따라 조합하여 시멘틱 영상 분할 방법을 지도 학습시키기 위한 가상의 픽셀 수준 학습 데이터인 의사 분할 데이터를 생성하는 단계를 포함한다.

발명의 효과

[0019] 따라서, 본 발명의 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 장치 및 방법은 경계 박스 수준의 약지도 학습 데이터에서 각 경계 박스의 전경과 배경을 구분하여 전경 영역만을 추출함으로써, 픽셀 수준 학습 데이터와 유사한 의사 픽셀 수준 학습 데이터를 획득할 수 있다. 그러므로 저비용의 약지도 학습 데이터인 경계 박스 수준 학습 데이터로도 시멘틱 영상 분할 장치를 높은 학습 성능으로 학습시킬 수 있다.

도면의 간단한 설명

[0020] 도 1은 본 발명의 일 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 장치의 개략적 구성을 나타낸다.

도 2는 도 1의 클래스 분할부의 상세 구조를 나타낸다.

도 3은 도 2의 영역 분할부의 동작을 설명하기 위한 도면이다.

도 4는 도 1의 의사 분할 데이터 생성부와 의사 학습 데이터 생성부의 상세 구성의 일 예를 나타낸다.

도 5는 도 4의 의사 분할 데이터 생성부와 의사 학습 데이터 생성부의 동작을 설명하기 위한 도면이다.

도 6은 본 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 장치에 의해 생성된 의사 학습 데이터의 일 예를 나타낸다.

도 7은 본 발명의 일 실시예에 따른 의사 학습 데이터에 의해 학습되는 시멘틱 영상 분할 장치의 개략적 구성을 나타낸다.

도 8은 본 발명의 일 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 방법을 나타낸다.

도 9는 본 실시예에 따라 획득된 시멘틱 영상 분할 학습 데이터를 이용한 시멘틱 영상 분할 장치 학습 방법의 일 예를 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0021] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.

[0022] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.

[0023] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

[0024] 도 1은 본 발명의 일 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 장치의 개략적 구성을 나타낸다.

[0025] 본 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 장치는 영상에 포함된 각 객체에 대한 객체 영역이 경계 박스(B) 형태로 지정되고, 각 경계 박스에 객체 클래스가 레이블된 약지도 학습용 경계 박스 수준의 학습 데이터를 인가받아 각 객체와 배경 영역이 픽셀 수준에서 구분된 고품질의 학습 데이터를 생성할 수 있다. 여기서 학습 데이터에는 학습 영상에 포함된 객체의 개수에 따라 다수의 경계 박스($B = \{B_1, B_2, \cdots, B_K\}$)가 지정될 수 있다.

[0026] 도 1을 참조하면, 본 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 장치는 클래스 분할부(100), 의사 분할 데이터 생성부(200) 및 의사 학습 데이터 생성부(300)를 포함할 수 있다.

[0027] 우선 클래스 분할부(100)는 영상의 객체 영역이 박스 형태로 지정되고 객체 클래스가 레이블된 경계 박스 수준의 학습 데이터를 인가받아 영상에서 객체 영역과 배경 영역을 구분하고, 구분된 객체 영역과 배경 영역의 클래스를 식별한다. 여기서 클래스 분할부(100)는 경계 박스 수준의 학습 데이터를 인가받아 약지도 학습 방식으로 미리 학습될 수 있다.

[0028] 그리고 약지도 학습 방식으로 학습된 클래스 분할부(100)는 학습된 방식에 따라 학습 데이터의 학습 영상으로부터 학습 특징맵(f)을 획득하고, 획득된 학습 특징맵(f)과 경계 박스(B)를 기반으로 영상의 배경 영역을 지정하는 배경 주의맵(A)과 전경 영역의 객체의 클래스를 식별하기 위해 이용되는 클래스 분류 가중치(w)를 획득한다. 특히 본 실시예에서 클래스 분할부(100)는 학습 특징맵(f)에서 경계 박스(B) 외부 영역에 해당하는 배경의 특징을 추출하여 경계 박스(B) 내부와의 유사성을 비교함으로써 경계 박스(B) 내부의 배경까지 인식하도록 함으로써 배경 영역과 객체 영역을 구분할 수 있다.

[0029] 클래스 분할부(100)는 학습 데이터 획득부(110), 학습 특징 추출부(120), 배경 주의맵 생성부(130), 전경 특징 획득부(140) 및 클래스 식별부(150)를 포함할 수 있다.

[0030] 학습 데이터 획득부(110)는 학습 영상의 객체 영역이 경계 박스(B)로 지정되고, 각 객체 영역에 대한 객체 클래스가 레이블된 경계 박스 수준의 학습 데이터를 획득한다. 상기한 바와 같이 경계 박스 수준의 학습 데이터는 기존의 약지도 학습 방식에 이용되는 학습 데이터로서 픽셀 수준 학습 데이터에 비해 상대적으로 매우 저비용으로 용이하게 획득될 수 있다.

[0031] 학습 특징 추출부(120)는 학습 영상으로부터 미리 학습된 방식에 따라 특징을 추출하여 학습 특징맵(f)을 추출한다. 여기서 학습 특징 추출부(120)는 CNN(Convolutional Neural Networks)과 같은 인공 신경망으로 구현될 수 있다.

[0032] 배경 주의맵 생성부(130)는 학습 특징 추출부(120)에서 획득된 학습 특징맵(f)에서 학습 데이터의 각 객체 영역에 지정된 경계 박스(B)로 설정된 영역을 제외한 나머지 영역을 기반으로 학습 영상의 배경 특징(q)을 추출하고, 추출된 배경 특징(q)과 학습 특징맵(f)의 경계 박스(B) 영역 사이의 유사도에 따라 경계 박스(B) 내에 포함된 배경 영역을 지정하는 배경 주의맵(A)을 획득한다.

[0033] 전경 특징 획득부(140)는 학습 특징맵(f)과 배경 주의맵(A)을 이용하여 경계 박스 내의 객체 영역을 검출하고, 검출된 객체 영역에 대한 객체 특징(r)을 추출한다.

[0034] 그리고 클래스 식별부(150)는 배경 주의맵 생성부(130)에서 추출된 배경 특징(q)과 전경 특징 획득부(140)에서 추출된 객체 특징(r)으로부터 배경과 객체 각각의 클래스를 식별한다. 클래스 식별부(150)는 미리 학습된 인공 신경망의 완전 연결 레이어(Fully Connected Layer)로 구현될 수 있으며, 이때 배경 특징(q)과 객체 특징(r)으로부터 개별 클래스를 식별하도록 학습에 의해 클래스 분류 가중치(w)가 설정될 수 있다.

[0035] 한편 의사 분할 데이터 생성부(200)는 클래스 분할부(100)에서 획득된 학습 특징맵(f)과 클래스 분류 가중치(w)를 기반으로 학습 영상의 각 객체 영역을 추출하기 위한 관심이 집중된 클래스 활성맵(CAM)을 획득하고, 획득된 클래스 활성맵(CAM)과 배경 주의맵(A)을 기반으로 픽셀 수준에서 배경 영역과 객체 영역이 구분되어 레이블된 의사 분할 데이터(Pseudo Segmentation Data)(Y_{crf})를 생성한다.

[0036] 의사 학습 데이터 생성부(300)는 의사 분할 데이터 생성부(200)에서 생성된 의사 분할 데이터(Y_{crf})의 각 클래스 영역에 대한 클래스 특징을 추출하고, 추출된 클래스 특징과 클래스 분할부(100)에서 획득된 학습 특징맵(f) 사이의 유사도를 기반으로 클래스 영역을 다시 재구분하여, 재구분된 픽셀 수준의 분할 영상인 의사 학습 데이터

(Y_{ret})를 생성한다.

[0037] 여기서 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})는 모두 픽셀 수준으로 배경 영역과 각 객체에 대한 객체 영역이 클래스 영역으로 구분되고 대응하는 클래스가 레이블된 데이터이므로, 지도 학습용 학습 데이터로 이용될 수 있다. 즉 의사 분할 데이터(Y_{crf})만으로도 지도 학습 데이터로 이용될 수도 있다. 다만 클래스 활성맵(CAM)과 배경 주의맵(A)으로부터 획득된 의사 분할 데이터(Y_{crf})에서는 객체의 색상 및 질감과 같은 로우 레벨 특징에 따라서는 클래스 영역이 오분할될 가능성이 있어 잘못된 학습이 수행될 수 있다.

[0038] 이러한 문제를 보완하기 위하여 본 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 장치는 의사 분할 데이터(Y_{crf})와 학습 영상으로부터 획득된 하이 레벨의 학습 특징맵(f)을 이용하여 의사 학습 데이터(Y_{ret})를 다시 획득하는 의사 학습 데이터 생성부(300)를 추가로 구비한다.

[0039] 의사 학습 데이터(Y_{ret})는 의사 분할 데이터(Y_{crf})를 보완하여 생성된 데이터로서 더욱 효과적인 픽셀 수준 학습 데이터로 이용될 수 있다. 그러나 의사 학습 데이터(Y_{ret})만으로 학습된 시멘틱 영상 분할 장치는 입력되는 영상 품질에 크게 성능이 좌우될 수 있다는 한계가 있다.

[0040] 즉 시멘틱 영상 분할 장치의 경우, 다양한 노이즈가 포함된 영상으로부터도 강건하게 클래스별로 영상을 분할할 수 있도록 영상 품질에 둔감하도록 학습되어야 하지만, 의사 학습 데이터(Y_{ret})만으로 학습된 시멘틱 영상 분할 장치의 경우 영상 품질에 따라 영상 분할 결과에 큰 오차를 유발할 수 있다.

[0041] 이에 본 실시예에서는 다양한 품질의 영상에서 정확하게 영상 분할을 수행할 수 있도록 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})를 함께 이용하여 시멘틱 영상 분할 장치를 학습시키는 것으로 가정한다.

[0042] 도 2는 도 1의 클래스 분할부의 상세 구조를 나타내고, 도 3은 도 2의 영역 분할부의 동작을 설명하기 위한 도면이다.

[0043] 도 1에서 설명한 바와 같이, 클래스 분할부(100)는 학습 데이터 획득부(110), 학습 특징 추출부(120), 배경 주의맵 생성부(130), 전경 특징 획득부(140) 및 클래스 식별부(150)를 포함할 수 있다.

[0044] 도 2 및 도 3을 참조하면, 배경 주의맵 생성부(130)는 마스크 획득부(131), 배경 특징 추출부(132), 배경 상관맵 획득부(133) 및 배경 주의맵 획득부(134)를 포함할 수 있다.

[0045] 마스크 획득부(131)는 학습 데이터에 레이블된 경계 박스 영역을 기반으로 마스크(M)를 획득한다. 경계 박스 수준의 학습 데이터에서 경계 박스는 학습 영상에서 객체가 포함된 영역을 사각 박스 형태로 지시하고 있으며, 이에 마스크 획득부(131)는 객체가 포함된 경계 박스 영역을 마스크(M)로 설정한다. 이때 마스크(M)는 경계 박스 내부 영역의 픽셀값이 0인 이진 마스크일 수 있다.

[0046] 배경 특징 추출부(132)는 학습 특징 추출부(120)에서 획득된 학습 특징맵(f)에 마스크 획득부(131)에서 획득된 마스크(M)를 적용하여 학습 특징맵(f)에서 경계 박스 영역을 제거함으로써 명확한 배경 영역을 추출하고, 추출된 배경 영역에 대한 배경 특징(q)을 추출한다. 배경 특징 추출부(132)는 학습 특징맵(f)과 마스크(M)를 원소 곱(Element-wise Multiplication)하여 배경 영역을 추출할 수 있다. 학습 데이터에서 객체 영역을 명시적으로 나타내는 경계 박스를 기반으로 획득된 마스크(M)가 학습 특징맵(f)에 원소 곱으로 적용되면, 해당 영역의 픽셀값은 모두 0이 된다. 따라서 학습 특징맵(f)에서 객체가 포함된 영역은 모두 제거되어 배경 영역만이 남게 된다. 그리고 배경 특징 추출부(132)는 마스크(M)가 적용된 학습 특징맵(f)의 특징을 추출하여 배경 특징(q)을 획득할 수 있다.

[0047] 이때 배경 특징 추출부(132)는 학습 특징맵(f)을 기지정된 크기의 그리드(grid)로 구분하고, 구분된 그리드 위치에 따른 개별 배경 특징(q_j)을 구분하여 획득할 수 있다. 배경 특징 추출부(132)는 학습 특징맵(f)을 $N \times N$ 크기로 분할하여 다수의 그리드 셀($G(j)$)(여기서 $1 \leq j \leq N^2$)로 구분할 수 있다. 그리고 구분된 개별 그리드 각각에 대한 배경 특징(q_j)을 수학식 1에 따라 획득할 수 있다.

수학식 1

$$q_j = \frac{\sum_{\mathbf{p}\in G(j)} M(\mathbf{p}) f(\mathbf{p})}{\sum_{\mathbf{p}\in G(j)} M(\mathbf{p})}$$

[0048]

[0049] 여기서 p는 마스크(M)와 학습 특징맵(f)의 픽셀을 나타낸다.

[0050] 즉 수학식 1에 따르면, 배경 특징(q_j)은 학습 특징맵(f)에서 대응하는 그리드 셀(G(j))의 가중 평균으로 획득될 수 있다.

[0051] 배경 상관맵 획득부(133)는 배경 특징 추출부(132)에서 획득된 다수의 배경 특징(q_j) 각각과 학습 특징맵(f)의 경계 박스(B) 영역 사이의 유사도를 기반으로 학습 특징맵(f)의 경계 박스(B) 영역에서 다수의 배경 특징(q_j) 각각에 대응하는 영역을 추출하는 배경 상관맵(A_j)을 수학식 2에 따라 계산하여 획득할 수 있다.

수학식 2

$$A_j(\mathbf{p}) = \begin{cases} \mathrm{ReLU}\left(\frac{f(\mathbf{p})}{\|f(\mathbf{p})\|} \cdot \frac{q_j}{\|q_j\|}\right) & , \mathbf{p} \in \mathcal{B} \\ 1 & , \mathbf{p} \notin \mathcal{B} \end{cases}$$

[0052]

[0053] 여기서 $\|\cdot\|$는 L_2 놈 함수를 나타내며, ReLU(Rectified Linear Unit)는 정류 함수를 나타낸다.

[0054] 배경 주의맵 획득부(134)는 다수의 배경 상관맵(A_j)을 수학식 3과 같이 누적 평균하여 배경 주의맵(A)을 획득한다.

수학식 3

$$A(\mathbf{p}) = \frac{1}{J}\sum_j A_j(\mathbf{p})$$

[0055]

[0056] 여기서 각각은 다수의 배경 특징(q_j)은 학습 특징맵(f)에서 명시적으로 객체가 포함되지 않은 배경 영역을 다수의 그리드로 구분하여 추출한 특징이므로 다양한 배경 영역별 특징이고, 다수의 배경 상관맵(A_j)은 마스크(M)에 의해 지정되는 학습 특징맵(f)의 경계 박스 영역에서 각 배경 특징(q_j)과 유사한 영역이 강조된 맵이므로, 배경 주의맵(A)은 객체가 포함된 영역을 지정하는 경계 박스 영역 내에 함께 포함된 배경 영역을 나타내는 맵으로 볼 수 있다.

[0057] 여기서 배경 주의맵(A)의 픽셀값이 1에 가까울수록 배경 영역에 해당할 가능성이 크다고 볼 수 있다.

[0058] 그리고 전경 특징 획득부(140)는 전경 영역 강조부(141) 및 전경 특징 추출부(142)를 포함할 수 있다.

[0059] 전경 영역 강조부(141)는 학습 특징맵(f)과 배경 주의맵(A)을 인가받고, 배경 주의맵(A)을 반전하여 학습 특징맵(f)과의 유사도를 분석하여, 경계 박스(B) 내의 객체 영역을 검출한다. 전경 영역 강조부(141)는 우선 경계 박스(B) 내의 배경 영역이 강조된 배경 주의맵(A)을 픽셀 반전(1-A)하여, 전경 영역, 즉 객체 영역이 표출되도록 하고, 픽셀 반전된 배경 주의맵(A)과 학습 특징맵(f)을 원소 곱하여 각 경계 박스(B_i)(여기서 i = {1, 2, …, K}) 내의 객체 영역이 강조되도록 한다. 경계 박스(B_i)에서 객체 영역이 강조된 맵을 전경 주의맵((1-

A(p))f(p))이라고도 할 수 있다.

[0060] 그리고 전경 특징 추출부(142)는 전경 주의맵의 특징을 집계하여 각 경계 박스(B_i)에 대한 특징인 전경 특징(r_i)을 계산한다. 전경 특징 획득부(140)는 수학식 4에 따라 각 경계 박스(B_i)에 대응하는 전경 특징(r_i)을 획득할 수 있다.

수학식 4

[0061]
$$r_i = \frac{\sum_{\mathbf{p}\in B_i}(1-A(\mathbf{p}))f(\mathbf{p})}{\sum_{\mathbf{p}\in B_i}(1-A(\mathbf{p}))}$$

[0062] 클래스 식별부(150)는 상기한 바와 같이, 각 경계 박스(B_i)에 대응하는 전경 특징(r_i)을 학습에 의해 미리 지정된 가중치(w)에 따라 분류하여, 전경 특징(r_i)에 대응하는 객체의 클래스를 식별한다. 즉 각 경계 박스(B_i)에 포함된 객체를 식별한다. 또한 클래스 식별부(150)는 배경 특징(q_j)을 인가받아 배경 또한 함께 식별한다.

[0063] 결과적으로, 도 2에 도시된 클래스 분할부(100)는 경계 박스를 이용하여 배경 특징을 추출하고, 추출된 배경 특징을 기반으로 경계 박스 내에 포함된 배경을 인지하여 풀링함으로써, 경계 박스에서 전경 영역과 배경 영역을 구분하여 클래스를 식별하도록 한다. 이와 같은 방식을 본 실시예에서는 배경 인지 풀링(Background-Aware Pooling: BAP) 기법이라 한다.

[0064] 도 4는 도 1의 의사 분할 데이터 생성부와 의사 학습 데이터 생성부의 상세 구성의 일 예를 나타내고, 도 5는 도 4의 의사 분할 데이터 생성부와 의사 학습 데이터 생성부의 동작을 설명하기 위한 도면이다.

[0065] 도 4 및 도 5를 참조하면, 의사 분할 데이터 생성부(200)는 클래스 활성맵 획득부(210), 객체 클래스 주의맵 획득부(220), 분할 데이터 획득부(230)를 포함할 수 있다.

[0066] 클래스 활성맵 획득부(210)는 학습 특징맵(f)과 클래스 식별부(150)가 식별한 객체 클래스(c)에 대응하는 가중치(w_c)를 인가받아, 학습 특징맵(f)에서 클래스(c)를 식별하기 위해 주목된 영역을 나타내는 클래스 활성맵(CAM_c)을 수학식 5에 따라 획득한다.

수학식 5

[0067]
$$\mathbf{CAM}_c(\mathbf{p}) = \mathbf{ReLU}(f(\mathbf{p})\cdot w_c)$$

[0068] 본 실시예에서 객체 클래스(c)는 각 경계 박스(B_i)에 포함된 객체 특징(r_i)을 추출하여 식별된 결과이므로, 특정 객체 클래스(c)에 대한 클래스 활성맵(CAM_c)은 해당 클래스로 식별된 특정 경계 박스(B_c) 내의 객체 특징(r_c)으로부터 획득된다. 따라서 객체 클래스 주의맵 획득부(220)는 수학식 6에 따라 클래스 활성맵(CAM_c)을 경계 박스(B_c) 영역 내의 정규화된 값을 갖는 객체 클래스 주의맵(u_c)으로 변환할 수 있다.

수학식 6

[0069]
$$u_c(\mathbf{p}) = \begin{cases} \frac{\mathbf{CAM}_c(\mathbf{p})}{\max_{\mathbf{p}}(\mathbf{CAM}_c(\mathbf{p}))} & ,\mathbf{p}\in\mathcal{B}_c \\ 0 & ,\mathbf{p}\notin\mathcal{B}_c \end{cases}$$

[0070] 분할 데이터 획득부(230)는 각 객체 클래스별(c)로 획득된 객체 클래스 주의맵(u_c)과 함께 배경 주의맵(A)을 배경 클래스 주의맵(u_0 = A)으로 인가받고, 기지정된 방식으로 조합하여 의사 분할 데이터(Y_{crf})를 획득한다. 도 5의 (a)에 도시된 바와 같이, 분할 데이터 획득부(230)는 객체를 식별하기 위해 주목된 영역을 나타내는 객체 클래스 주의맵(u_c)과 배경을 식별하기 위해 주목된 배경 클래스 주의맵(u_0)을 결합하여 각 객체에 대한 영역과 배경 영역이 구분되고 해당 클래스가 레이블된 의사 분할 데이터(Y_{crf})를 생성할 수 있다.

[0071] 분할 데이터 획득부(230)는 일 예로 밀집 CFR(dense CRF: dense Conditional Random Fields) 기법을 적용하여 의사 분할 데이터(Y_{crf})를 생성할 수 있다. 밀집 CFR은 공지된 기술이므로 여기서는 상세하게 설명하지 않는다.

[0072] 한편, 의사 학습 데이터 생성부(300)는 클래스 영역 추출부(310), 클래스 특징 획득부(320) 및 의사 학습 데이터 획득부(330)를 포함할 수 있다.

[0073] 클래스 영역 추출부(310)는 의사 분할 데이터(Y_{crf})를 인가받고, 의사 분할 데이터(Y_{crf})의 각 클래스별 영역(Q_c)을 별도로 추출한다.

[0074] 의사 분할 데이터(Y_{crf})는 이미 객체 영역과 배경 영역이 구분되고 대응하는 클래스가 레이블되어 있으므로, 클래스 영역 추출부(310)는 도 5의 (b)에 도시된 바와 같이, 레이블된 클래스에 따른 클래스 영역(Q_c)을 개별적으로 선택하여 추출할 수 있다.

[0075] 클래스 특징 획득부(320)는 학습 특징맵(f)을 인가받아, 학습 특징맵(f)에서 별도로 추출된 각 클래스 영역(Q_c)에 대응하는 영역의 클래스 특징(q_c)을 계산하여 획득한다. 클래스 특징 획득부(320)는 수학식 7과 같이 각 클래스 영역(Q_c)에 대응하는 학습 특징맵(f) 영역의 픽셀 평균값을 계산하여 클래스 특징(q_c)을 획득할 수 있다.

수학식 7

[0076]
$$q_c = \frac{1}{|\mathcal{Q}_c|} \sum_{\mathbf{p} \in \mathcal{Q}_c} f(\mathbf{p})$$

[0077] 여기서 $|Q_c|$는 클래스 영역(Q_c)에 대응하는 학습 특징맵(f) 영역의 픽셀 개수를 나타낸다.

[0078] 의사 학습 데이터 획득부(330)는 학습 특징맵(f)과 각 클래스별 클래스 특징(q_c) 사이의 유사도를 계산하여 클래스 상관맵(C_c)을 수학식 8에 따라 획득한다.

수학식 8

[0079]
$$C_c(\mathbf{p}) = \frac{f(\mathbf{p})}{\|f(\mathbf{p})\|} \cdot \frac{q_c}{\|q_c\|}$$

[0080] 그리고 획득된 클래스 상관맵(C_c)에 argmax 함수를 적용함으로써, 의사 학습 데이터(Y_{ret})를 획득한다.

[0081] 도 6은 본 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 장치에 의해 생성된 의사 학습 데이터의 일 예를 나타낸다.

[0082] 도 6에서 (a)는 입력 이미지로서 경계 박스 수준의 학습 데이터이고, (b)와 (c)는 각각 객체 클래스 주의맵(u_c)과 배경 클래스 주의맵(u_0)을 나타낸다. (d)와 (e)는 각각 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})를 나타내며, (f)는 진리값(Ground Truth)을 나타낸다. 도 6에 도시된 바와 같이, 객체 클래스 주의맵(u_c)과 배경 클래스 주의맵(u_0)으로부터 획득되는 의사 분할 데이터(Y_{crf})도 대부분의 경우, 안정적으로 클래스를 식별할 수 있으나, 위로부터 5번째와 같이 특정 질감에 대해서는 영상 분할을 수행하지 못하는 경우도 발생할 수 있다.

그에 반해 의사 학습 데이터(Y_{ret})는 상대적으로 정확하게 영상 분할을 수행할 수 있다. 다만, 3번째의 경우, 의사 학습 데이터(Y_{ret})보다 의사 분할 데이터(Y_{crf})가 더욱 정확하게 영상 분할을 수행하였으며, 이로부터 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})는 상호 보완적 학습 데이터로 이용될 수 있음을 알 수 있다.

[0083] 도 7은 본 발명의 일 실시예에 따른 의사 학습 데이터에 의해 학습되는 시멘틱 영상 분할 장치의 개략적 구성을 나타낸다.

[0084] 도 7을 참조하면, 시멘틱 영상 분할 장치(400)는 영상 획득부(410), 특징 추출부(420), 클래스 확률맵 획득부(430) 및 시멘틱 분할 영상 획득부(440)를 포함할 수 있다.

[0085] 영상 획득부(410)는 시멘틱 분할 대상이 되는 영상을 획득한다. 특징 추출부(420)는 영상 획득부(410)에서 획득된 영상에 대해 미리 학습된 방식에 따라 특징을 추출하여 특징맵(ϕ)을 획득한다. 클래스 확률맵 획득부(430)는 학습된 방식에 따라 특징맵(ϕ)의 각 픽셀이 각 클래스에 대응할 확률을 나타내는 클래스 확률맵(H_c)을 획득한다. 클래스 확률맵 획득부(430)는 일 예로 소프트 맥스(softmax) 분류 기법에 따라 클래스 확률맵(H_c)을 획득할 수 있다. 시멘틱 분할 영상 획득부(440)는 학습된 방식에 다라 클래스 확률맵(H_c)을 기반으로 각 픽셀이 해당하는 클래스를 식별하고, 식별된 클래스에 따라 영역을 구분하여 시멘틱 분할 영상을 획득한다.

[0086] 도 7은 시멘틱 영상 분할 장치(400)의 일 예에 따른 구성으로 시멘틱 분할 영상 장치는 기존에 공지된 기법에 따라 다른 구성을 가질 수도 있다.

[0087] 한편, 시멘틱 영상 분할 장치(400)는 학습 시에 학습부(500)가 더 포함될 수 있다. 본 실시예에 따른 영상 분할 장치는 도 1에 도시된 시멘틱 영상 분할 학습 데이터 생성 장치에서 획득된 가상의 픽셀 수준 학습 데이터인 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})를 기반으로 지도 학습이 수행될 수 있다.

[0088] 상기한 바와 같이, 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})는 기존에 약지도 학습을 수행하기 위해 획득된 경계 박스 수준의 학습 데이터로부터 가상으로 획득된 유사 픽셀 수준 학습 데이터이다. 또한 상기한 바와 같이, 시멘틱 영상 분할 장치(400)의 학습 시에는 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret}) 중 하나만을 이용할 수도 있으나, 본 실시예에서는 상호 보완이 가능하도록 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})를 모두 이용함으로써, 시멘틱 영상 분할 장치(400)를 강건하게 학습시킬 수 있다.

[0089] 이에 시멘틱 영상 분할 장치(400)의 학습 시에 영상 획득부(410)는 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})를 인가받는다. 그리고 학습부(500)는 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret}) 각각에서 획득된 클래스 확률맵(H_c)을 인가받아 손실을 계산하고, 계산된 손실을 역전파하여 시멘틱 영상 분할 장치(400)를 학습시킬 수 있다.

[0090] 학습부(500)는 공통 영역 판별부(510), 공통 영역 손실 계산부(520), 비공통 영역 손실 계산부(530) 및 손실 역전파부(540)를 포함할 수 있다.

[0091] 공통 영역 판별부(510)는 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})를 인가받고, 인가된 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})에서 동일한 클래스 레이블을 제공하는 공통 영역(S)을 판별하여 획득한다. 이때 공통 영역 판별부(510)는 공통 영역(S)을 클래스 레이블에 따라 클래스 공통 영역(S_c)으로 구분하여 획득할 수 있다. 또한 공통 영역 판별부(510)는 공통 영역(S)으로 판별되지 않은 비공통 영역(~S) 또한 구분하여 획득한다.

[0092] 여기서 공통 영역 판별부(510)가 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})에서 동일 클래스로 식별된 공통 영역(S)과 서로 다른 클래스로 식별된 비공통 영역(~S)을 구분하는 것은 잘못된 픽셀 수준의 분할 정보가 레이블될 수 있는 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})에서 상대적으로 더욱 신뢰할 수 있는 영역을 추출하기 위해서이다.

[0093] 즉 공통 영역(S)은 비공통 영역(~S)에 비해 상대적으로 높은 신뢰도를 갖는 학습 데이터로 판별할 수 있기 때문에 공통 영역(S)을 기반으로 우선 손실을 계산할 수 있다.

[0094] 이에 공통 영역 손실 계산부(520)는 공통 영역 판별부(510)에서 판별된 클래스 공통 영역(S_c)과 대응하는 클래스 확률맵(H_c)을 기반으로 공통 영역 손실(L_{ce})을 수학식 9에 따라 계산하여 획득한다.

수학식 9

[0095]
$$\mathcal{L}_{\text{ce}} = -\frac{1}{\sum_c |\mathcal{S}_c|} \sum_c \sum_{\mathbf{p} \in \mathcal{S}_c} \log H_c(\mathbf{p})$$

[0096] 여기서 H_c는 각 클래스(c)에 대한 확률이고, S_c는 공통 영역(S)에서 특정 클래스로 레이블된 픽셀 집합이다.

[0097] 한편, 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})에서 서로 다른 클래스로 식별된 비공통 영역(~S)은 공통 영역(S)에 비해 신뢰도가 낮음에도 정확한 레이블이 지정되었을 가능성이 있다. 이에 본 실시예에서 학습부(500)는 비공통 영역(~S)을 기반으로 추가적인 손실을 계산한다.

[0098] 비공통 영역 손실 계산부(530)는 클래스 상관맵 생성부(531), 신뢰도맵 생성부(532) 및 크로스 엔트로피 손실 계산부(533)를 포함할 수 있다.

[0099] 클래스 상관맵 생성부(531)는 시멘틱 분할 영상 획득부(440)에서 각 객체 클래스를 식별하기 위해 학습에 의해 지정되는 클래스 분류 가중치(W_c)가 해당 클래스의 대표 특징인 것으로 간주하여 특징 추출부(420)에서 획득된 특징맵(ϕ)과 클래스 분류 가중치(W_c) 사이의 코사인 유사도에 기반하여 상관 관계맵(D_c)을 수학식 10에 따라 생성한다.

수학식 10

[0100]
$$D_c(\mathbf{p}) = 1 + \left(\frac{\phi(\mathbf{p})}{\|\phi(\mathbf{p})\|} \cdot \frac{W_c}{\|W_c\|} \right)$$

[0101] 그리고 신뢰도맵 생성부(532)는 상관 관계맵(D_c)으로부터 의사 분할 데이터(Y_{crf})에 레이블된 클래스가 정확할 가능성을 나타내는 신뢰도를 수학식 11에 따라 계산하여 신뢰도맵(σ)을 획득한다.

수학식 11

[0102]
$$\sigma(\mathbf{p}) = \left(\frac{D_{c^*}(\mathbf{p})}{\max_c(D_c(\mathbf{p}))} \right)^{\gamma}$$

[0103] 여기서 c* 는 의사 분할 데이터(Y_{crf})에 레이블된 클래스를 나타내고, c 는 의사 학습 데이터(Y_{ret})에 레이블된 클래스를 나타내며, γ ($\gamma \geq 1$)는 감쇠 파라미터이다.

[0104] 수학식 11은 의사 분할 데이터(Y_{crf})에 레이블된 클래스(c*)가 신뢰할 수 있는 경우, 픽셀별 유사도($D_c(p)$)와 최대 픽셀별 유사도($\max(D_c(p))$)의 상관값이 유사하고 그 반대도 성립할 수 있다는 점에 근거한다. 그리고 감쇠 파라미터(γ)는 신뢰도를 조절하기 위한 파라미터로서, 감쇠 파라미터(γ)가 증가할수록 신뢰할 수 있는 클래스 레이블만이 고려되도록 한다. 즉 감쇠 파라미터(γ)가 매우 큰 값을 가지면, 의사 분할 데이터(Y_{crf})에 대한 클래스 레이블에 따른 픽셀별 유사도 $D_{c^*}(p) = \max_c(D_c(p))$ 일 때만 해당 픽셀 신뢰도($\sigma(p)$)가 1이고 이외에는 $\sigma(p) \fallingdotseq 0$으로 계산된다.

[0105] 그리고 크로스 엔트로피 손실 계산부(533)는 신뢰도 맵(σ)을 기반으로 비공통 영역(~S)의 픽셀별 신뢰도(σ

(p))와 클래스 확률맵(H_c)을 이용하여, 비공통 영역(~S)의 크로스 엔트로피를 수학식 12에 따라 계산하여 비공통 영역 손실(L_{wce})을 획득한다.

수학식 12

[0106]
$$\mathcal{L}_{\text{wce}} = -\frac{1}{\sum_c \sum_{\mathbf{p}\in\sim\mathcal{S}_c} \sigma(\mathbf{p})} \sum_c \sum_{\mathbf{p}\in\sim\mathcal{S}_c} \sigma(\mathbf{p}) \log H_c(\mathbf{p})$$

[0107] 손실 역전파부(540)는 공통 영역 손실(L_{ce})과 비공통 영역 손실(L_{wce})을 인가받아 총 손실(L)을 수학식 13에 따라 계산하고, 계산된 총 손실(L)을 역전파하여 시멘틱 영상 분할 장치(400)를 학습시킨다.

수학식 13

[0108]
$$\mathcal{L} = \mathcal{L}_{\text{ce}} + \lambda \mathcal{L}_{\text{wce}}$$

[0109] 여기서 λ는 손실 가중치이다.

[0110] 도 8은 본 발명의 일 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 방법을 나타낸다.

[0111] 도 1 내지 도 6을 참조하면, 본 실시예에 따른 시멘틱 영상 분할 학습 데이터 생성 방법은 픽셀 수준의 시멘틱 영상 분할 학습 데이터를 생성하기에 앞서 먼저 경계 박스 수준의 학습 데이터로부터 각 객체 영역과 각 객체 영역의 클래스 식별할 수 있도록 약지도 학습 단계(S10)를 수행한다.

[0112] 약지도 학습 단계에서는 우선 경계 박스 수준의 학습 데이터를 획득한다(S11). 그리고 인공 신경망을 이용하여 획득된 학습 데이터의 학습 영상으로부터 학습되는 방식에 따라 특징을 추출하여 학습 특징맵(f)을 획득한다(S12). 학습 특징맵(f)이 획득되면 학습 데이터에 레이블된 경계 박스(B)와 학습 특징맵(f)을 이용하여, 객체가 포함되지 않은 배경 영역의 위치별 배경 특징(q_j)을 추출하고, 추출된 배경 특징(q_j)과 학습 특징맵(f) 사이의 유사도에 기반하여, 경계 박스(B)내의 배경 영역을 지정하는 배경 주의맵(A)을 생성한다(S13). 배경 주의맵(A)이 생성되면, 배경 주의맵(A)을 픽셀 반전하여 학습 특징맵(f)을 이용하여 경계 박스 내에서 배경 영역을 배제하고 객체 영역이 강조된 전경 주의맵을 획득하고, 기지정된 방식으로 전경 주의맵의 특징을 계산하여 전경 특징(r_i)을 획득한다(S14).

[0113] 각 경계 박스에 대한 전경 특징(r_i)이 획득되면, 학습되는 방식에 따라 가중치(w)가 가변되어 설정되는 인공 신경망을 이용하여 전경 특징(r_i)과 배경 특징(q_j)을 분류함으로써, 객체 및 배경에 대한 클래스를 식별한다(S15). 그리고 식별된 클래스와 학습 데이터에 레이블된 클래스 사이의 차이를 기반으로 기지정된 방식에 따라 손실을 계산하고 계산된 손실을 역전파하여 인공 신경망을 약지도 학습시킨다(S16).

[0114] 이후 약지도 학습의 종료 여부를 판별한다(S17). 여기서 약지도 학습의 종료는 기지정된 기준 횟수만큼 반복 학습이 수행되었는지 여부로 판별될 수 있다. 만일 학습이 종료되지 않은 것으로 판단되면, 다시 학습 데이터를 획득하여 약지도 학습을 수행한다. 그러나 약지도 학습이 종료된 것으로 판별되면, 가상의 픽셀 수준 학습 데이터 획득 단계(S20)를 수행한다.

[0115] 가상의 픽셀 수준 학습 데이터 획득 단계(S20)에서는 우선 경계 박스 수준으로 객체 영역이 레이블된 입력 데이터를 인가받는다(S21). 그리고 인가된 입력 데이터로부터 학습 특징맵(f)과 배경 주의맵(A) 및 전경 특징(r_i)을 획득하여 입력 데이터에 대한 객체 영역과 배경 영역을 구분하고 각 영역의 클래스를 식별한다(S22).

[0116] 입력 데이터의 객체 영역과 배경 영역이 구분되고 각 영역의 클래스가 식별되면, 학습 특징맵(f)에 각 클래스 식별에 이용된 가중치(w_c)를 가중하여, 학습 특징맵(f)에서 클래스 식별에 주목된 영역을 나타내는 클래스 활성맵(CAM_c)을 획득하고, 획득된 클래스 활성맵(CAM_c)을 대응하는 경계 박스(B_c) 영역 내의 정규화된 맵으로 변환하

여 객체 클래스 주의맵(u_c)을 생성하고, 배경 주의맵(A)을 배경 클래스 주의맵(u_0)으로 획득한다(S23).

[0117] 객체 클래스 주의맵(u_c)과 배경 클래스 주의맵(u_0)이 획득되면, 객체 클래스 주의맵(u_c)과 배경 클래스 주의맵(u_0)을 밀집 CFR 기법과 같이 기지정된 방식으로 조합하여 각 객체에 대한 영역과 배경 영역이 구분되고 해당 클래스가 레이블된 가상의 픽셀 수준 학습 데이터인 의사 분할 데이터(Y_{crf})를 생성한다(S24).

[0118] 그리고 의사 분할 데이터(Y_{crf})에서 구분된 객체 영역과 배경 영역 각각을 개별적으로 추출하여 개별 클래스 영역(Q_c)을 획득한다(S25). 이후, 학습 특징맵(f)에서 획득된 클래스 영역(Q_c)에 대응하는 픽셀 평균값을 계산하여 클래스 특징(q_c)을 획득한다(S26). 학습 특징맵(f)과 각 클래스별 클래스 특징(q_c) 사이의 유사도를 계산하여 클래스 상관맵(C_c)을 획득하고, 획득된 클래스 상관맵(C_c)으로부터 또 다른 가상의 픽셀 수준 학습 데이터인 의사 학습 데이터(Y_{ret})를 생성한다(S27).

[0119] 도 9는 본 실시예에 따라 획득된 시멘틱 영상 분할 학습 데이터를 이용한 시멘틱 영상 분할 장치(400) 학습 방법의 일 예를 나타낸다.

[0120] 도 8에 도시된 시멘틱 영상 분할 학습 데이터 생성 방법에 따라 가상의 픽셀 수준 학습 데이터인 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})가 획득되면, 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})를 지도 학습 데이터로 이용하여 시멘틱 영상 분할 장치(400)를 학습시키는 시멘틱 영상 분할 학습 단계를 수행한다(S30).

[0121] 시멘틱 영상 분할 학습 단계에서는 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})를 시멘틱 영상 분할 장치(400)를 학습시키기 위한 지도 학습 데이터로서 인가한다(S31).

[0122] 그리고 시멘틱 영상 분할 장치(400)는 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})를 인가받아 의사 특징맵(ϕ)을 획득하고, 특징맵(ϕ)의 각 픽셀이 각 클래스에 대응할 확률을 나타내는 클래스 확률맵(H_c)을 획득하며, 획득된 클래스 확률맵(H_c)의 클래스를 학습에 의해 업데이트되는 클래스 가중치(W_c)에 따라 식별하여, 각 객체 영역과 배경 영역을 구분하고 클래스를 판별하여 시멘틱 분할 영상을 획득한다(S32).

[0123] 한편, 시멘틱 영상 분할 장치(400)를 학습시키기 위해서는 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})에서 동일한 클래스(c)가 레이블된 공통 영역(S)과 서로 다른 클래스(c)가 레이블된 비공통 영역(~S)을 판별한다(S33).

[0124] 그리고 공통 영역(S_c)과 대응하는 클래스 확률맵(H_c)을 기반으로 공통 영역 손실(L_{ce})을 수학식 9에 따라 계산하여 획득한다(S34).

[0125] 또한 특징맵(ϕ)과 클래스 분류 가중치(W_c) 사이의 코사인 유사도에 기반하여 상관 관계맵(D_c)을 수학식 10에 따라 생성한다(S35). 상관 관계맵(D_c)이 계산되면, 의사 분할 데이터(Y_{crf})에 대한 상관 관계맵(D_{c*})과 의사 학습 데이터(Y_{ret})에 대한 상관 관계맵(D_c)을 이용하여 특징맵(ϕ)의 각 픽셀별 신뢰도를 나타내는 신뢰도맵(σ)을 수학식 11에 따라 획득한다(S36).

[0126] 이후 비공통 영역(~S)의 픽셀별 신뢰도($\sigma(p)$)와 클래스 확률맵(H_c)을 이용하여, 비공통 영역(~S)의 크로스 엔트로피를 수학식 12에 따라 계산하여 비공통 영역 손실(L_{wce})을 계산한다(S37).

[0127] 공통 영역 손실(L_{ce})과 비공통 영역 손실(L_{wce})이 계산되면, 수학식 13에 따라 총 손실(L)을 계산하고, 계산된 총 손실(L)을 역전파하여 시멘틱 영상 분할 장치(400)를 지도 학습시킨다(S38). 그리고 지도 학습이 완료되었는지 판별한다(S39).

[0128] 지도 학습 또한 기지정된 횟수로 반복 학습되었는지 여부에 따라 학습 완료 여부가 판별될 수 있으며, 학습이 완료되지 않은 것으로 판별되면 다시 의사 분할 데이터(Y_{crf})와 의사 학습 데이터(Y_{ret})를 이용하여 지도 학습을 수행한다.

[0129] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스 될 수 있는 임의의 가용 매체일 수 있고, 또한 컴퓨터 저장

매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.

[0130] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다.

[0131] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

[0132] 100: 클래스 분할부 110: 학습 데이터 획득부

120: 학습 특징 추출부 130: 배경 주의맵 생성부

131: 마스크 획득부 132: 배경 특징 추출부

133: 배경 상관맵 획득부 133: 배경 주의맵 획득부

140: 전경 특징 획득부 141: 전경 영역 강조부

142: 전경 특징 추출부 150: 클래스 식별부

200: 의사 분할 데이터 생성부 210: 클래스 활성맵 획득부

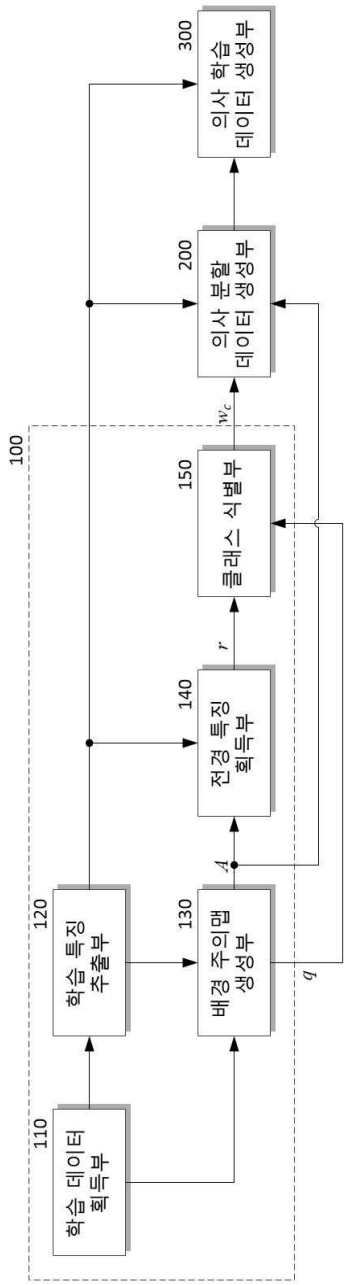
220: 객체 클래스 주의맵 획득부 230: 분할 데이터 획득부

300: 의사 학습 데이터 생성부 310: 클래스 영역 추출부

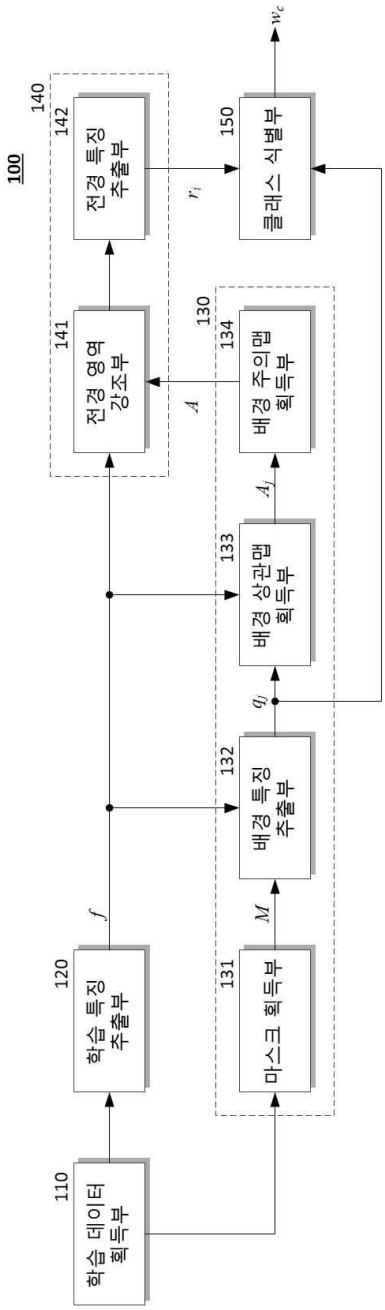
320: 클래스 특징 획득부 330: 의사 학습 데이터 획득부

도면

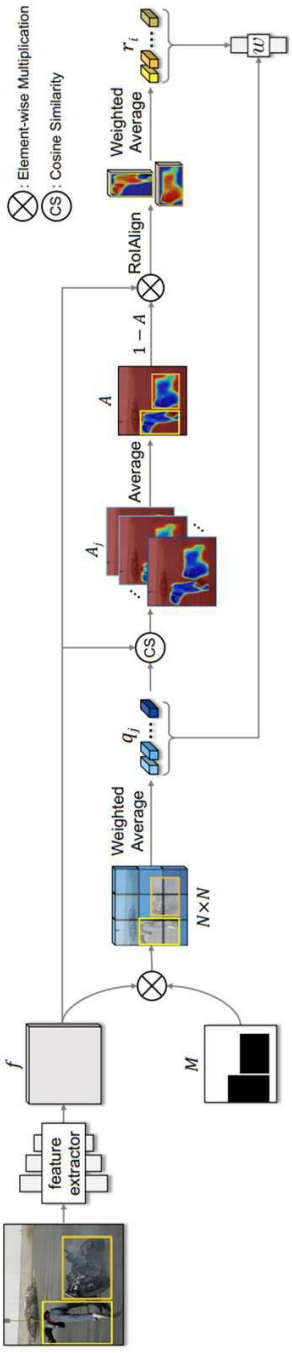
도면1


100
110
학습 데이터 획득부
120
학습 특징 추출부
130
배경 주의맵 생성부
140
전경 특징 획득부
150
클래스 식별부
200
의사 분할 데이터 생성부
300
의사 학습 데이터 생성부
w_c
r
A
q

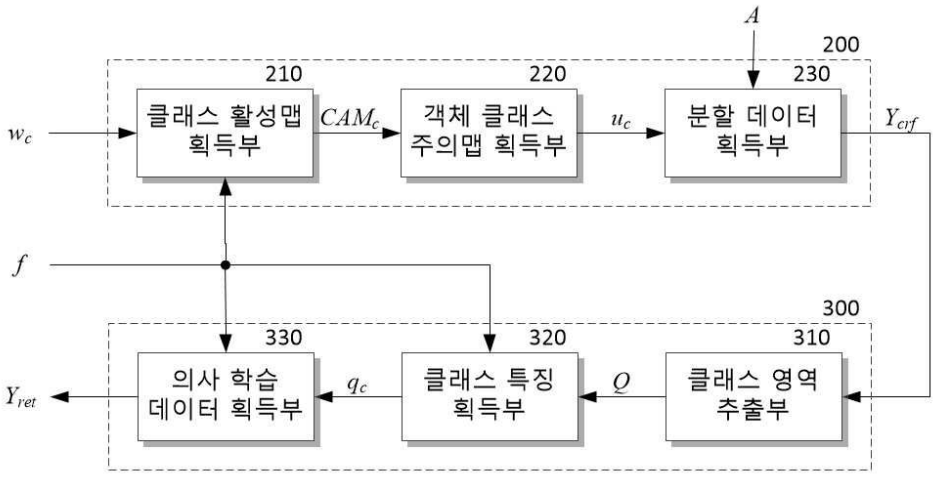
도면2


100
110
학습 데이터 획득부
120
학습 특징 추출부
130
131
마스크 획득부
132
배경 특징 추출부
133
배경 상관맵 획득부
134
배경 주의맵 획득부
140
141
전경 영역 강조부
142
전경 특징 추출부
150
클래스 식별부
f
M
q_j
A_j
A
r_i
w_c

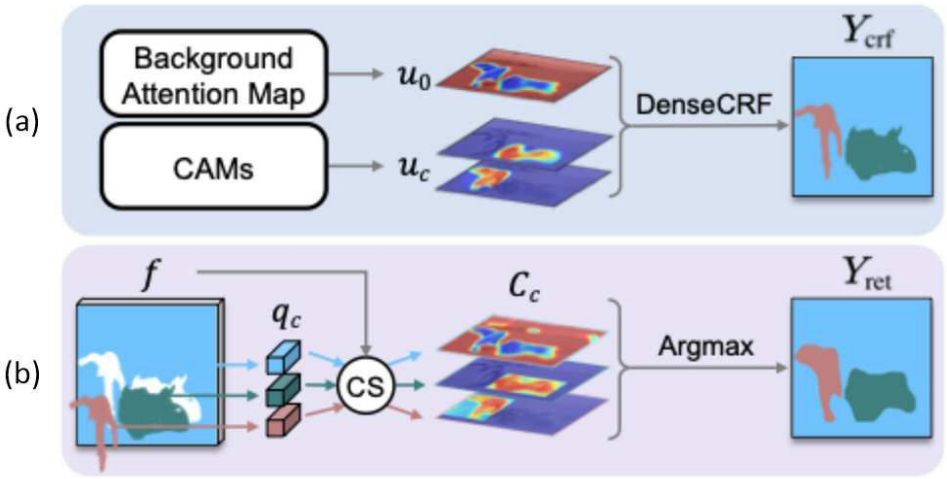
도면3


: Element-wise Multiplication
: Cosine Similarity
feature extractor
f
M
Weighted Average
N×N
q_j
CS
A_j
Average
A
1 − A
RoIAlign
Weighted Average
r_i
u

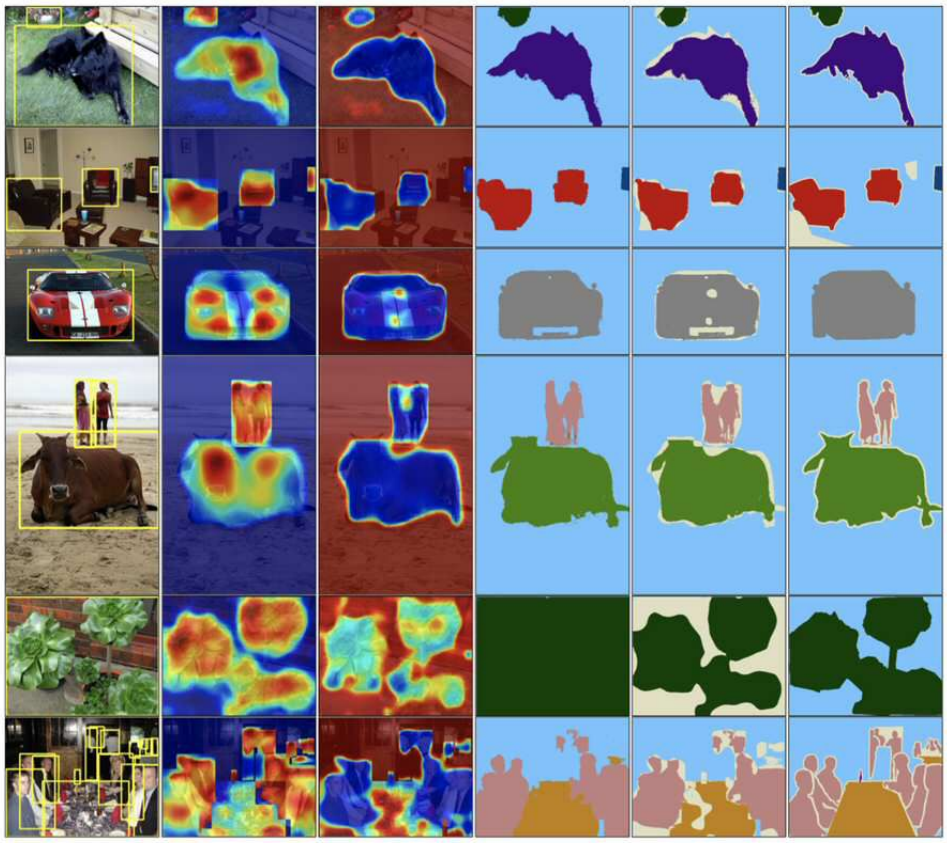
도면4


A
200
210
220
230
w_c
클래스 활성맵 획득부
CAM_c
객체 클래스 주의맵 획득부
u_c
분할 데이터 획득부
Y_{crf}
f
300
330
320
310
Y_{ret}
의사 학습 데이터 획득부
q_c
클래스 특징 획득부
Q
클래스 영역 추출부

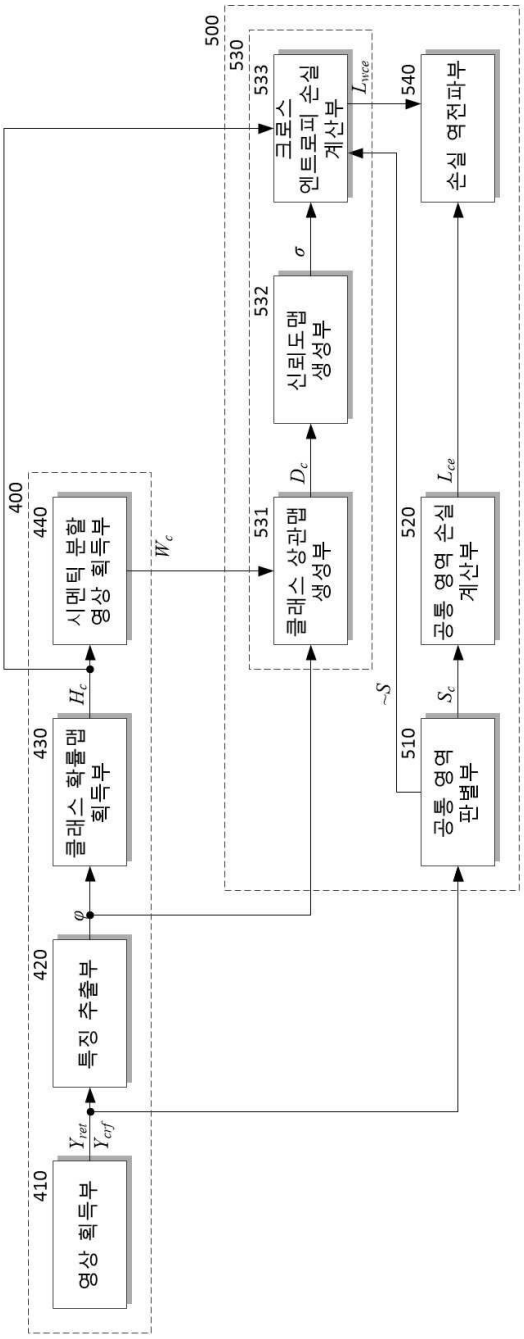
도면5


(a)
Background Attention Map
CAMs
u_0
u_c
DenseCRF
Y_{crf}
(b)
f
q_c
CS
C_c
Argmax
Y_{ret}

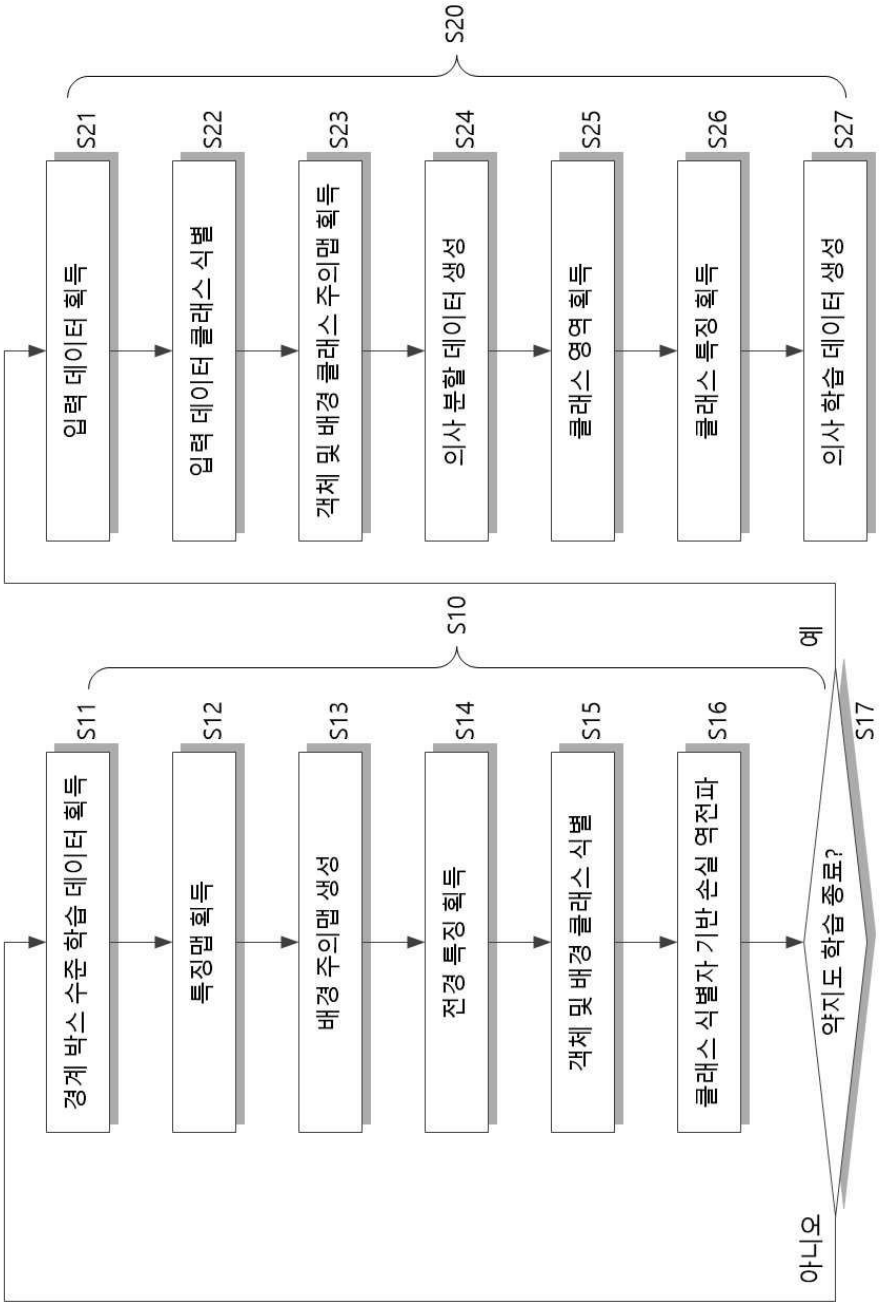
도면6



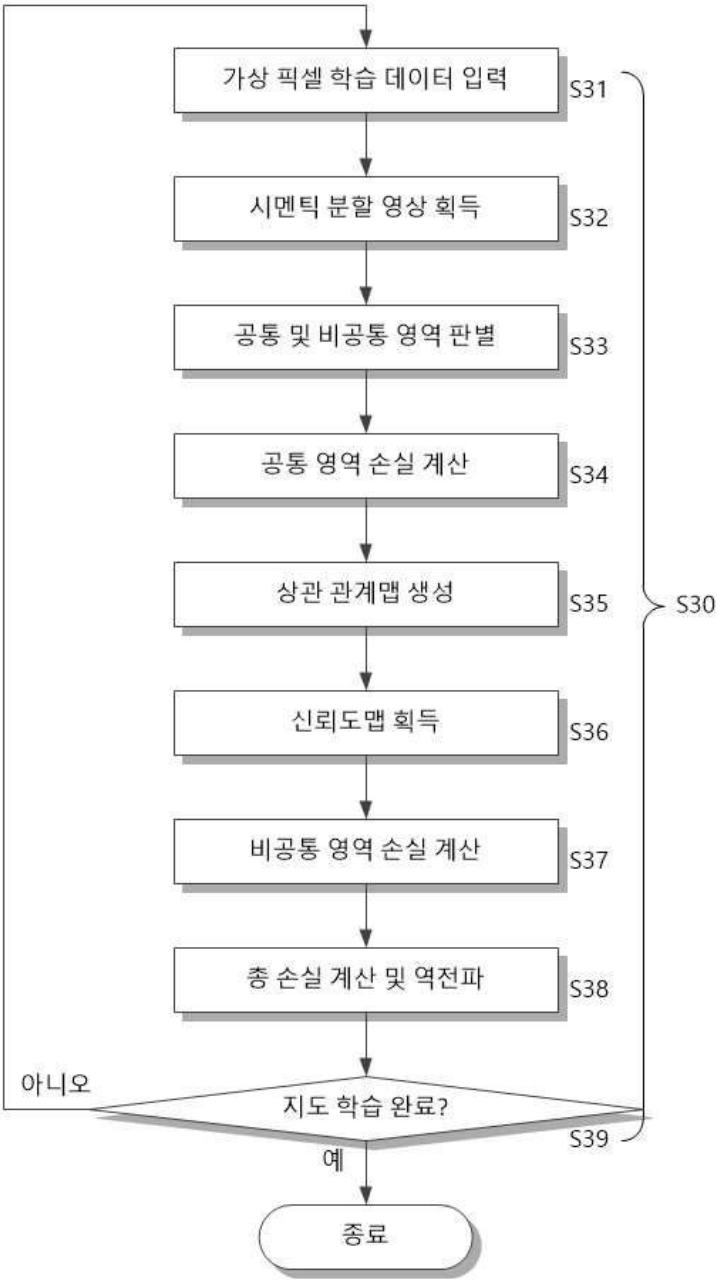
도면7


400
410
영상 획득부
420
특징 추출부
430
클래스 확률맵 획득부
440
시멘틱 분할 영상 획득부
500
510
공통 영역 판별부
520
공통 영역 손실 계산부
530
531
클래스 상관맵 생성부
532
신뢰도맵 생성부
533
크로스 엔트로피 손실 계산부
540
손실 역전파부
Y_{est}
Y_{ref}
φ
H_c
W_c
D_c
σ
S_c
$\sim S$
L_{ce}
L_{wce}

도면8


S20
S21
입력 데이터 획득
S22
입력 데이터 클래스 식별
S23
객체 및 배경 클래스 주의맵 획득
S24
의사 분할 데이터 생성
S25
클래스 영역 획득
S26
클래스 특징 획득
S27
의사 학습 데이터 생성
S10
S11
경계 박스 수준 학습 데이터 획득
S12
특징맵 획득
S13
배경 주의맵 생성
S14
전경 특징 획득
S15
객체 및 배경 클래스 식별
S16
클래스 식별자 기반 손실 역전파
S17
약지도 학습 종료?
예
아니오

도면9


가상 픽셀 학습 데이터 입력
S31
시멘틱 분할 영상 획득
S32
공통 및 비공통 영역 판별
S33
공통 영역 손실 계산
S34
상관 관계맵 생성
S35
S30
신뢰도맵 획득
S36
비공통 영역 손실 계산
S37
총 손실 계산 및 역전파
S38
아니오
지도 학습 완료?
S39
예
종료