(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0030722
(43) 공개일자 2025년03월05일

(51) 국제특허분류(Int. Cl.)
G06T 7/70 (2017.01) *G06N 3/0895* (2023.01)
G06T 7/11 (2017.01) *G06V 10/40* (2022.01)
(52) CPC특허분류
G06T 7/70 (2017.01)
G06N 3/0895 (2023.01)
(21) 출원번호 10-2023-0112036
(22) 출원일자 2023년08월25일
심사청구일자 2023년08월25일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
석흥일
서울특별시 성북구 길음로 118, 411동 1202호(길음동, 길음뉴타운)
손정효
서울특별시 동대문구 왕산로22길 16, 706호(용두동, 가우디캐슬)
(74) 대리인
송인호, 윤형근, 최관락

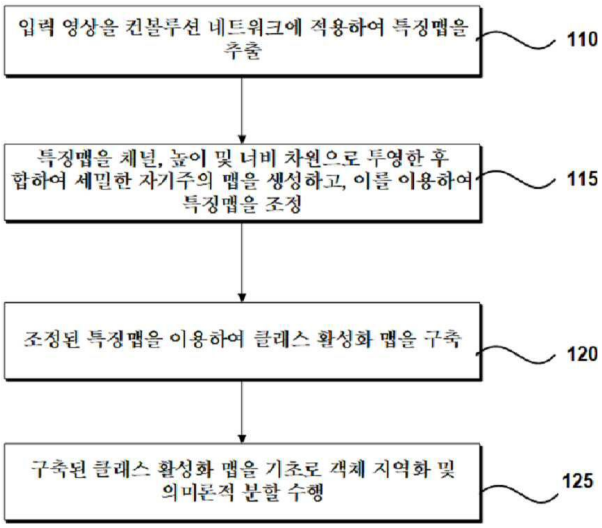
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 **약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법 및 그 장치**

(57) 요 약

약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법 및 그 장치가 개시된다. 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법은, (a) 입력 영상에 대한 입력 특징맵을 추출하는 단계; (b) 상기 입력 특징맵에서 채널, 높이 및 너비 차원으로 상호보완적인 특징맵을 추출하여 상기 입력 특징맵을 조정하는 단계; (c) 상기 조정된 특징맵을 이용하여 클래스 활성화 맵을 구성하는 단계; 및 (d) 상기 클래스 활성화 맵을 기초로 객체 지역화 및 의미론적 분할을 수행하는 단계를 포함한다.

대 표 도 - 도1



(52) CPC특허분류
G06T 7/11 (2017.01)
G06V 10/40 (2023.08)
G06T 2207/20081 (2013.01)
G06T 2210/12 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193973
과제번호	2019-0-00079-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	인공지능대학원지원(고려대학교)
기 여 율	1/1
과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명 세 서

청구범위

청구항 1

(a) 입력 영상에 대한 입력 특징맵을 추출하는 단계;

(b) 상기 입력 특징맵에서 채널, 높이 및 너비 차원으로 상호보완적인 특징맵을 추출하여 상기 입력 특징맵을 조정하는 단계;

(c) 상기 조정된 특징맵을 이용하여 클래스 활성화 맵을 구성하는 단계; 및

(d) 상기 클래스 활성화 맵을 기초로 객체 지역화 및 의미론적 분할을 수행하는 단계를 포함하는 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법.

청구항 2

제1 항에 있어서,

상기 (d) 단계는,

(d1) 상기 클래스 활성화 맵을 기초로 상기 객체 지역화를 위한 경계 상자를 생성하는 단계;

(d2) 상기 클래스 활성화 맵을 기초로 픽셀 카테고리에 해당하는 의사 마스크를 생성하는 단계; 및

(d3) 상기 생성된 의사 마스크를 라벨로 사용하여 상기 입력 영상에 대한 의미론적 분할(semantic segmentation)을 수행하는 단계를 포함하되,

상기 (d2) 및 (d3)는 상기 (d1)과 병렬로 수행되는 것을 특징으로 하는 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법.

청구항 3

제1 항에 있어서,

상기 (b) 단계는,

상기 입력 특징맵을 상기 채널, 높이 및 너비 각 차원으로 투영하여 각 차원 특징맵을 생성하는 단계;

상기 채널, 높이 및 너비 차원 특징맵을 제로패딩을 사용하여 1D 컨볼루션을 적용하고 배치 정규화 및 비선형 활성화 함수를 적용하여 채널, 높이 및 너비 차원에 대한 주의 맵을 생성하는 단계; 및

상기 채널, 높이 및 너비 차원의 주의 맵을 확장한 후 합산하여 세밀한 자기주의 맵을 생성하고, 상기 세밀한 자기주의 맵과 상기 입력 특징맵의 하다마드 곱 연산을 수행한 결과에 상기 입력 특징맵과의 요소별 합산을 통해 상기 입력 특징맵을 조정하는 단계를 포함하는 것을 특징으로 하는 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법.

청구항 4

제3 항에 있어서,

상기 채널, 높이 및 너비 차원의 주의 맵은 상기 입력 특징맵의 크기로 확장되는 것을 특징으로 하는 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법.

청구항 5

적어도 하나의 명령어를 저장하는 메모리; 및

상기 메모리에 저장된 명령어를 실행하는 프로세서를 포함하되,

상기 프로세서에 의해 실행된 명령어는 각각,

(a) 입력 영상에 대한 입력 특징맵을 추출하는 단계;

(b) 상기 입력 특징맵에서 채널, 높이 및 너비 차원으로 상호보완적인 특징맵을 추출하여 상기 입력 특징맵을 조정하는 단계;

(c) 상기 조정된 특징맵을 이용하여 클래스 활성화 맵을 구성하는 단계; 및

(d) 상기 클래스 활성화 맵을 기초로 객체 지역화 및 의미론적 분할을 수행하는 단계를 수행하는 것을 특징으로 하는 컴퓨팅 장치.

청구항 6

제5 항에 있어서,

상기 (d) 단계는,

(d1) 상기 클래스 활성화 맵을 기초로 상기 객체 지역화를 위한 경계 상자를 생성하는 단계;

(d2) 상기 클래스 활성화 맵을 기초로 픽셀 카테고리에 해당하는 의사 마스크를 생성하는 단계; 및

(d3) 상기 생성된 의사 마스크를 라벨로 사용하여 상기 입력 영상에 대한 의미론적 분할(semantic segmentation)을 수행하는 단계를 포함하되,

상기 (d2) 및 (d3)는 상기 (d1)과 병렬로 수행되는 것을 특징으로 하는 컴퓨팅 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법 및 그 장치에 관한 것이다.

배 경 기 술

[0003] 객체 지역화(object localization)는 컴퓨터 비전 분야에서 중요한 기술로 영상에서 특정 객체의 위치를 결정해 해당 객체의 경계 상자를 찾는 것을 목표로 한다. 또한, 의미론적 분할(semantic segmentation)은 영상 내 각 픽셀을 특정 클래스 또는 카테고리에 할당하는 것을 목표로 하는 컴퓨터 비전 분야의 기술이다.

[0004] 종래에는 객체 지역화와 의미론적 분할을 위해 객체의 위치 및 각 픽셀의 클래스에 해당하는 정보를 포함한 데이터 셋을 만들어 학습에 사용하는 완전 지도학습(fully supervised) 기반의 방법을 사용하였다. 그러나, 종래 방법은 데이터 셋의 생성에 많은 시간과 노동력을 필요로 하여 다양한 문제에 적용하기 어려운 문제점이 있다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법 및 그 장치를 제공하기 위한 것이다.

[0007] 또한, 본 발명은 특징 맵의 서로 다른 차원에 분포된 정보를 활용하여 조정함으로써 객체 지역화 성능 향상 및

의미론적 분할 성능 향상이 가능한 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법 및 그 장치를 제공하기 위한 것이다.

과제의 해결 수단

[0009] 본 발명의 일 측면에 따르면, 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법이 제공된다.

[0010] 본 발명의 일 실시예에 따르면, (a) 입력 영상에 대한 입력 특징맵을 추출하는 단계; (b) 상기 입력 특징맵에서 채널, 높이 및 너비 차원으로 상호보완적인 특징맵을 추출하여 상기 입력 특징맵을 조정하는 단계; (c) 상기 조정된 특징맵을 이용하여 클래스 활성화 맵을 구성하는 단계; 및 (d) 상기 클래스 활성화 맵을 기초로 객체 지역화 및 의미론적 분할을 수행하는 단계를 포함하는 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법이 제공될 수 있다.

[0011] 상기 (d) 단계는, (d1) 상기 클래스 활성화 맵을 기초로 상기 객체 지역화를 위한 경계 상자를 생성하는 단계; (d2) 상기 클래스 활성화 맵을 기초로 픽셀 카테고리에 해당하는 의사 마스크를 생성하는 단계; 및 (d3) 상기 생성된 의사 마스크를 라벨로 사용하여 상기 입력 영상에 대한 의미론적 분할(semantic segmentation)을 수행하는 단계를 포함하되, 상기 (d2) 및 (d3)는 상기 (d1)과 병렬로 수행될 수 있다.

[0012] 상기 (b) 단계는, 상기 입력 특징맵을 상기 채널, 높이 및 너비 각 차원으로 투영하여 각 차원 특징맵을 생성하는 단계; 상기 채널, 높이 및 너비 차원 특징맵을 제로패딩을 사용하여 1D 컨볼루션을 적용하고 배치 정규화 및 비선형 활성화 함수를 적용하여 채널, 높이 및 너비 차원에 대한 주의 맵을 생성하는 단계; 및 상기 채널, 높이 및 너비 차원의 주의 맵을 확장한 후 합산하여 세밀한 자기주의 맵을 생성하고, 상기 세밀한 자기주의 맵과 상기 입력 특징맵의 하다마드 곱 연산을 수행한 결과에 상기 입력 특징맵과의 요소별 합산을 통해 상기 입력 특징맵을 조정하는 단계를 포함할 수 있다.

[0013] 상기 채널, 높이 및 너비 차원의 주의 맵은 상기 입력 특징맵의 크기로 확장될 수 있다.

[0015] 본 발명의 다른 측면에 따르면, 약한 지도 학습 기반 객체 지역화와 의미론적 분할을 위한 장치가 제공된다.

[0016] 본 발명의 일 실시예에 따르면, 적어도 하나의 명령어를 저장하는 메모리; 및 상기 메모리에 저장된 명령어를 실행하는 프로세서를 포함하되, 상기 프로세서에 의해 실행된 명령어는 각각, (a) 입력 영상에 대한 입력 특징맵을 추출하는 단계; (b) 상기 입력 특징맵에서 채널, 높이 및 너비 차원으로 상호보완적인 특징맵을 추출하여 상기 입력 특징맵을 조정하는 단계; (c) 상기 조정된 특징맵을 이용하여 클래스 활성화 맵을 구성하는 단계; 및 (d) 상기 클래스 활성화 맵을 기초로 객체 지역화 및 의미론적 분할을 수행하는 단계를 수행하는 것을 특징으로 하는 컴퓨팅 장치가 제공될 수 있다.

발명의 효과

[0018] 본 발명의 일 실시예에 따른 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법 및 그 장치를 제공함으로써, 특징 맵의 서로 다른 차원에 분포된 정보를 활용하여 특징맵을 조정한 후 객체 지역화 및 의미론적 분할에 적용하여 성능을 향상시킬 수 있는 이점이 있다.

[0019] 또한, 본 발명의 특징맵의 상호보완적인 정보로 생성된 자기주의 맵을 활용하여 객체 전체 영역에 대한 세밀한 정보를 획득할 수 있는 이점도 있다.

도면의 간단한 설명

[0021] 도 1은 본 발명의 일 실시예에 따른 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법을 나타낸 순서도.

도 2는 종래와 본 발명의 일 실시예에 따른 주의맵의 차이를 설명하기 위해 도시한 도면.

도 3은 본 발명의 일 실시예에 따른 특징맵 조정 프레임워크를 도시한 도면.

도 4 및 도 5는 종래와 본 발명의 일 실시예에 따른 객체 지역화 결과를 도시한 도면.

도 6은 종래와 본 발명의 일 실시예에 따른 의사 마스크를 비교한 도면.

도 7은 종래와 본 발명의 일 실시예에 따른 의미론적 분할 수행 결과를 비교한 도면.

도 8은 본 발명의 일 실시예에 따른 약한 지도 학습 기반 객체 지역화와 의미론적 분할을 수행하는 컴퓨팅 장치의 내부 구성을 개략적으로 도시한 블록도.

발명을 실시하기 위한 구체적인 내용

[0022] 본 명세서에서 사용되는 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "구성된다" 또는 "포함한다" 등의 용어는 명세서상에 기재된 여러 구성 요소들, 또는 여러 단계들을 반드시 모두 포함하는 것으로 해석되지 않아야 하며, 그 중 일부 구성 요소들 또는 일부 단계들은 포함되지 않을 수도 있고, 또는 추가적인 구성 요소 또는 단계들을 더 포함할 수 있는 것으로 해석되어야 한다. 또한, 명세서에 기재된 "...부", "모듈" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어 또는 소프트웨어로 구현되거나 하드웨어와 소프트웨어의 결합으로 구현될 수 있다.

[0023] 이하, 첨부된 도면들을 참조하여 본 발명의 실시예를 상세히 설명한다.

[0025] 도 1은 본 발명의 일 실시예에 따른 약한 지도 학습 기반 객체 지역화와 의미론적 분할 방법을 나타낸 순서도이고 도 2는 종래와 본 발명의 일 실시예에 따른 주의맵의 차이를 설명하기 위해 도시한 도면이고, 도 3은 본 발명의 일 실시예에 따른 특징맵 조정 프레임워크를 도시한 도면이며, 도 4 및 도 5는 종래와 본 발명의 일 실시예에 따른 객체 지역화 결과를 도시한 도면이고, 도 6 및 도 7는 종래와 본 발명의 일 실시예에 따른 의사 마스크 및 의미론적 분할 수행 결과를 비교한 도면이다.

[0026] 단계 110에서 컴퓨팅 장치(100)는 입력 영상을 컨볼루션 네트워크에 적용하여 특징맵을 추출한다. 컨볼루션 네트워크를 통해 입력 영상에 대한 특징맵을 추출하는 방법 자체는 당업자에게는 자명한 사항이므로 이에 대한 별도의 설명은 생략하기로 한다.

[0027] 단계 115에서 컴퓨팅 장치(100)는 특징맵을 채널, 높이 및 너비 차원으로 투영한 후 합하여 세밀한 자기주의 맵을 생성하고, 이를 이용하여 특징맵을 조정한다.

[0028] 이에 대해 보다 상세히 설명하기로 한다.

[0029] 예를 들어, 입력 특징맵을 $\mathbf{X} \in \mathbb{R}^{C \times H \times W}$ 로 나타내기로 한다. 여기서, C, H, W는 각각 채널, 높이 및 너비의 차원을 나타낸다.

[0030] 세가지 관점에서 입력 특징맵 X의 전역 분포를 압축하기 위해, 채널, 높이 및 너비 차원 각각에 대해 평균 풀링을 적용할 수 있다. 이를 수학식으로 나타내면 수학식 1 내지 수학식 3과 같다.

수학식 1

[0031] $$\mathbf{c} = \mathrm{AvgPool}_{\mathrm{w,h}}(\mathbf{X})$$

수학식 2

[0032] $$\mathbf{h} = \mathrm{AvgPool}_{\mathrm{w}}(\mathbf{X})$$

수학식 3

[0033] $$\mathbf{w} = \mathrm{AvgPool}_{\mathrm{h}}(\mathbf{X})$$

[0034] 여기서, $\mathrm{AvgPool}_{\{\cdot\}}$는 $\{\cdot\}$ 차원에 대한 평균 풀링 연산자를 나타낸다. $\mathbf{c} \in \mathbb{R}^{C\times 1\times 1}$, $\mathbf{h} \in \mathbb{R}^{C\times H\times 1}$, $\mathbf{w} \in \mathbb{R}^{C\times 1\times W}$의 세가지 평균 풀링된 특징은 서로 다른 관점에서 X에 대한 추출된 특징을 요약한 것으로 간주될 수 있다.

[0035] 즉, 3가지 관점은 입력 특징맵 X에 분산된 서로 다른 정보를 전달할 수 있다. 즉, C는 어떤 특징 표현이 고도로 활성화되었는지를 요약하고, h와 w는 채널 전체에 독립적으로 수직 및 수평으로 분포된 구별된 특징을 반영할 수 있다.

[0036] 풀링된 각 차원 특징의 단위 간 로컬 상호 작용을 활용하기 위해 커널 크기 k와 편향 없이 제로 패딩을 사용하여 1D 컨볼루션을 적용하여 차원성을 유지한 후 배치 정규화 및 비선형 활성화 함수를 적용하면 수학식 4 내지 수학식 6과 같이 나타낼 수 있다.

수학식 4

[0037] $$\mathbf{z}_{\mathrm{c}} = \sigma(\mathrm{BN}(\mathbf{W}_{\mathrm{c}}(\mathbf{c})))$$

수학식 5

[0038] $$\mathbf{z}_{\mathrm{h}} = \sigma(\mathrm{BN}(\mathbf{W}_{\mathrm{h}}(\mathbf{h})))$$

수학식 6

[0039] $$\mathbf{z}_{\mathrm{w}} = \sigma(\mathrm{BN}(\mathbf{W}_{\mathrm{w}}(\mathbf{w})))$$

[0040] 수학식 4 내지 수학식 6에서 $\sigma(\cdot)$는 시그모이드 함수를 나타내고, $\mathbf{W}_{\{\cdot\}}$는 풀링된 특징의 1D 컨볼루션 레이어를 나타낸다. $\mathbf{z}_{\mathrm{c}} \in \mathbb{R}^{C\times 1\times 1}$, $\mathbf{z}_{\mathrm{h}} \in \mathbb{R}^{C\times H\times 1}$, $\mathbf{z}_{\mathrm{w}} \in \mathbb{R}^{C\times 1\times W}$는 3가지 관점의 주의 맵을 나타낸다.

[0041] 3가지 관점의 주의 맵을 확장한 후 합 함수(f)를 통해 합하여 입력 특징맵(X)와 동일한 크기를 가지는 자기주의 맵을 생성할 수 있다.

[0042] 이를 수학식으로 나타내면 수학식 7과 같다.

수학식 7

$$\mathbf{M} = f(\mathbf{z}_c, \mathbf{z}_h, \mathbf{z}_w) = \left[z_c^{(i,1,1)} + z_h^{(i,j,1)} + z_w^{(i,1,k)} \right] = \left[m^{(i,j,k)} \right]$$

[0043]

[0044] 여기서, $z_c^{(i,1,1)}, z_h^{(i,j,1)}, z_w^{(i,1,k)}$, $m^{(i,j,k)}$는 각각 $\mathbf{z}_c, \mathbf{z}_h, \mathbf{z}_w$, $\mathbf{M}$의요소를 나타내고, i, j, k는 채널, 높이, 너비의 인덱스를 나타낸다. 이로 인해, 본 발명의 일 실시예에 따른 자기주의 맵 M의 값은 서로 다를 가능성이 높으며, 그 결과 세분화된 주의 맵이 생성될 수 있다.

[0045] 컴퓨팅 장치(100)는 생성된 세밀한 자기주의 맵을 이용하여 입력 특징맵(X)를 조정할 수 있다.

[0046] 이를 수학식으로 나타내면 수학식 8과 같이 나타낼 수 있다.

수학식 8

$$\hat{\mathbf{X}} = \mathbf{X} \oplus (\mathbf{X} \odot \mathbf{M})$$

[0047]

[0048] 여기서, $\odot$, $\oplus$는 각각 하다마드 곱(Hadamard product)와 요소별 합산(element-wise summation)을 나타낸다.

[0049] 본 발명의 일 실시예에 따르면, 세밀한 자기주의 맵을 이용하여 요소별 세부적인 특징맵을 보정할 수 있다. 다양한 관점에서 획득된 변별 특징맵의 합인 축 기반 특징맵 $(\mathbf{X} \odot \mathbf{M})$은 풍부한 특징 표현을 가진다. 또한, 요소별 합산은 변별 특징에 대한 정보가 이미 포함된 입력 특징맵을 추가함으로써 스케일링항의 식별의 덜한 영역을 활성화하는데 도움을 줄 수 있다.

[0050] 즉, 세밀한 자기주의 맵은 각 요소에 대해 다른 주의 값을 할당하는 미세한 특성을 보여주며, 원래 입력 특징맵에 더해져 자기주의 특징맵의 가장 구별되는 영역뿐만 아니라 객체의 덜 구별적인 영역에도 초점을 맞출 수 있다.

[0051] 도 2를 참조하여 설명하기로 한다.

[0052] 도 2의 (a)는 채널 주의 맵을 나타내고, (b)는 공간 주의 맵을 나타내며, (c)는 본 발명의 일 실시예에 따른 자기주의 맵을 나타낸다. 도 2의 (c)에서 보여지는 바와 같이, 축 기반 상호 보완적인 정보를 가지고 생성된 세밀한 자기주의 맵을 통해 입력 특징맵을 보정할 수 있다.

[0053] 이에 대한 프레임워크는 도 3에 도시된 바와 같다.

[0054] 단계 120에서 컴퓨팅 장치(100)는 조정된 특징맵을 이용하여 클래스 활성화 맵을 구축한다. 예를 들어, 컴퓨팅 장치(100)는 조정된 특징맵을 기반으로 GAP(Global Average Pooling) 또는 GMP(Global Max Pooling)를 적용하여 클래스 활성화 맵을 생성할 수 있다.

[0055] 단계 125에서 컴퓨팅 장치(100)는 구축된 클래스 활성화 맵을 기초로 객체 지역화 및 의미론적 분할을 수행한다.

[0056] 지면 설명의 특성상 순차적으로 설명되나, 객체 지역화와 의미론적 분할은 병렬로 수행될 수 있다.

[0057] 컴퓨팅 장치(100)는 구축된 클래스 활성화 맵을 기초로 객체 지역화를 위한 경계 박스를 생성할 수 있다.

[0058] 또한, 컴퓨팅 장치(100)는 구축된 클래스 활성화 맵을 기초로 픽셀 카테고리에 해당하는 의사 마스크를 생성하고, 이를 라벨 (label)로 사용하여 영상에 대한 의미론적 분할을 수행한다.

[0059] 도 4 및 도 5는 종래와 본 발명의 일 실시예에 따른 객체 지역화 결과를 도시한 도면이다. 도 4 및 도 5의 (a)는 입력 영상을 나타내고, (b)는 종래 기술의 클래스 활성화 맵을 도시한 도면이고, (c)는 본 발명의 일 실시예에 따른 조정된 특징맵을 이용한 클래스 활성화 맵을 도시한 도면이며, (d)는 종래 기술의 객체 지역화 결과이며, (e)는 본 발명의 일 실시예에 따른 객체 지역화 결과를 도시한 도면이다. 도 4 및 도 5의 (d) 및 (e)에서 붉은색 경계 박스는 실제 객체 지역화를 나타내고, 초록색 경계 박스는 종래 기술과 본 발명의 객체 지역화 결과를 나타낸다. 도 4 및 도 5에서 IoU는 실제 경계 박스와 예측된 경계 박스의 일치 정도를 정량적으로 나타낸 지표로 0과 1 사이의 값으로 표시되며 1에 가까울수록 완전히 일치하게 된다.

[0060] 도 4 및 도 5에서 보여지는 바와 같이, 본 발명의 일 실시예에 따른 세밀한 자기주의 맵으로 조정된 특징맵을 기초로 생성된 클래스 활성화 맵이 종래의 클래스 활성화 맵보다 객체 전체에 걸쳐 활성화 맵이 나타나는 것을 알 수 있다. 또한, 실제 객체 지역화 결과에서도 본 발명의 일 실시예에 따른 객체 지역화 결과, 즉 경계 박스가 실제와 더 일치하는 것을 알 수 있다.

[0061] 도 6 및 도 7는 종래와 본 발명의 일 실시예에 따른 의사 마스크 및 의미론적 분할 수행 결과를 비교한 도면이다.

[0062] 도 6에서 (a)는 입력 영상을 나타내고, (b)는 종래 기술의 의사 마스크를 나타내고, (c)는 본 발명의 일 실시예에 따른 의사 마스크를 나타내고, (d)는 실제 마스크를 나타낸다.

[0063] 도 7에서 (a)는 입력 영상을 나타내고, (b)는 종래 기술의 의미론적 분할 결과를 나타내고, (c)는 본 발명의 일 실시예에 따른 의미론적 분할 결과를 나타내고, (d)는 실제 값을 나타낸다.

[0064] 도 6 및 도 7에서 보여지는 바와 같이, 본 발명의 일 실시예에 따른 세밀한 자기주의 맵으로 조정된 특징맵을 기초로 생성된 클래스 활성화 맵을 기반으로 한 의사 마스크 및 의미론적 분할 결과가 실제 값에 더 가깝고 정확한 것을 알 수 있다.

[0066] 도 8은 본 발명의 일 실시예에 따른 약한 지도 학습 기반 객체 지역화와 의미론적 분할을 수행하는 컴퓨팅 장치의 내부 구성을 개략적으로 도시한 블록도이다.

[0067] 도 8을 참조하면, 본 발명의 일 실시예에 따른 약한 지도 학습 기반 객체 지역화와 의미론적 분할을 수행하는 컴퓨팅 장치(100)는 메모리(810) 및 프로세서(820)를 포함하여 구성된다.

[0068] 메모리(810)는 적어도 하나의 명령어를 저장한다.

[0069] 프로세서(820)는 메모리(810)에 저장된 명령어를 실행할 수 있다. 프로세서(820)에 의해 실행된 명령어는 입력 영상에 대한 입력 특징맵을 추출하고, 입력 특징맵에서 채널, 높이 및 너비 차원으로 상호 보완적인 특징맵을 추출하여 상기 입력 특징맵을 조정하며, 조정된 특징맵을 이용하여 클래스 활성화 맵을 구성한 후 클래스 활성화 맵을 기초로 객체 지역화 및 의미론적 분할을 수행하는 일련의 과정을 수행할 수 있다. 이는 도 1을 참조하여 설명한 바와 동일하므로 중복되는 설명은 생략하기로 한다.

[0071] 본 발명의 실시 예에 따른 장치 및 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 컴퓨터 판독 가능 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 분야 통상의 기술자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media) 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴

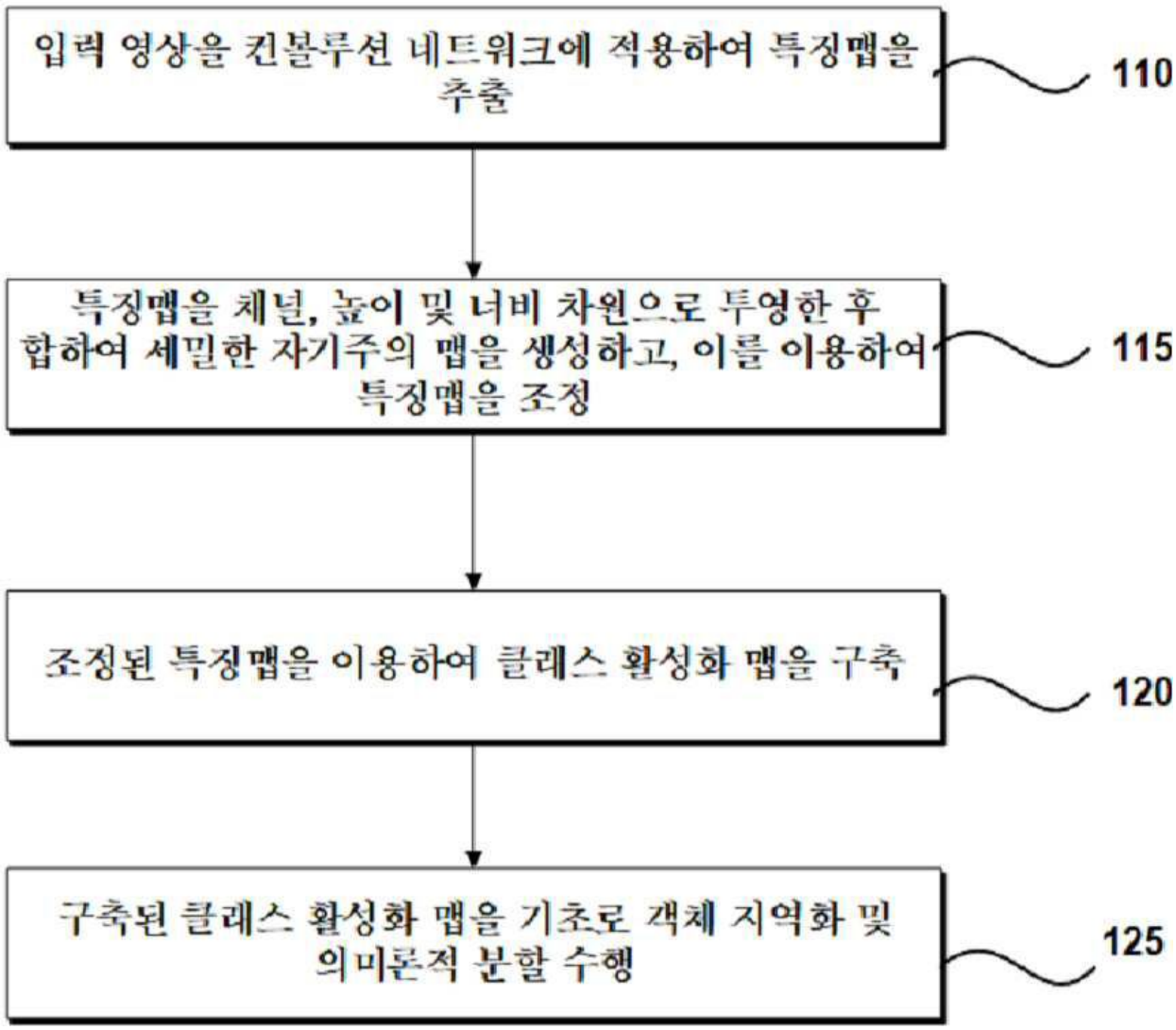
파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0072] 상술한 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

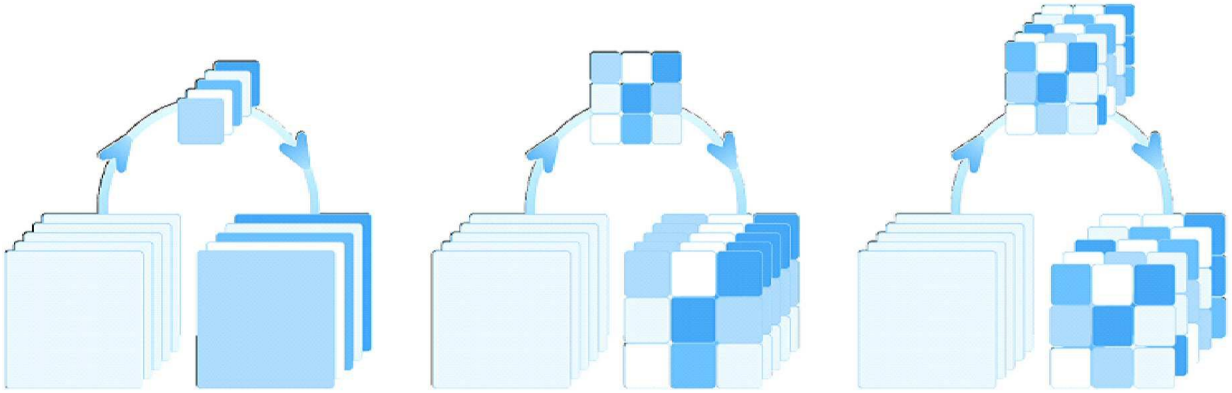
[0073] 이제까지 본 발명에 대하여 그 실시 예들을 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시 예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

도면

도면1

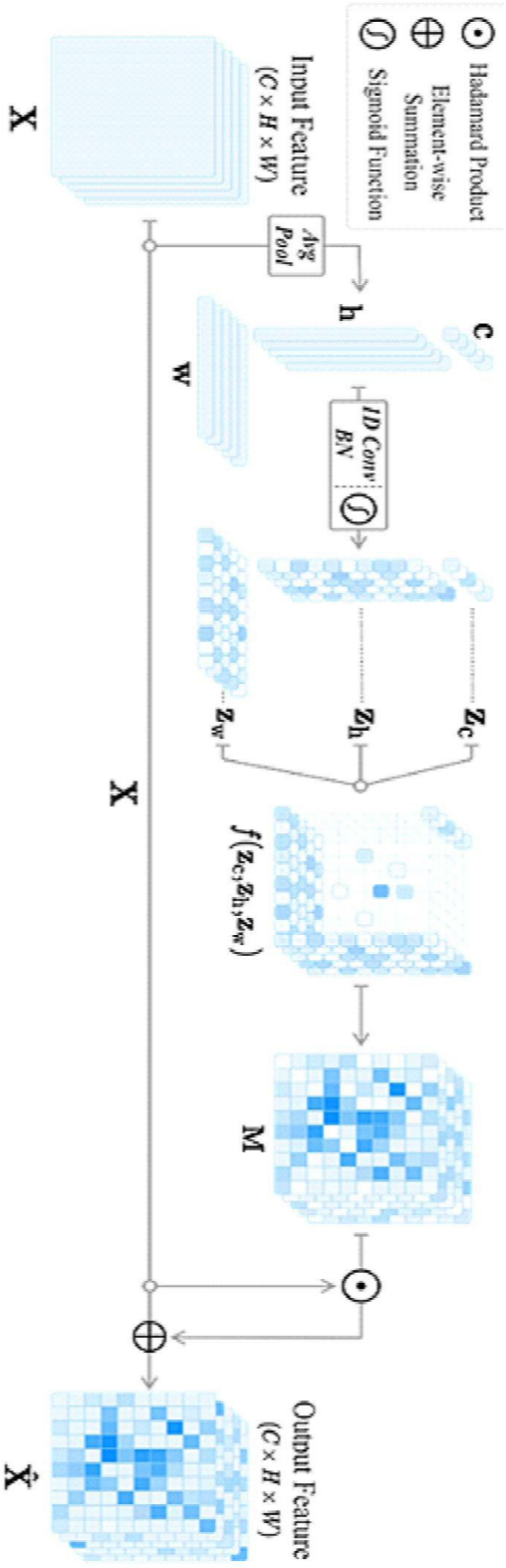


도면2

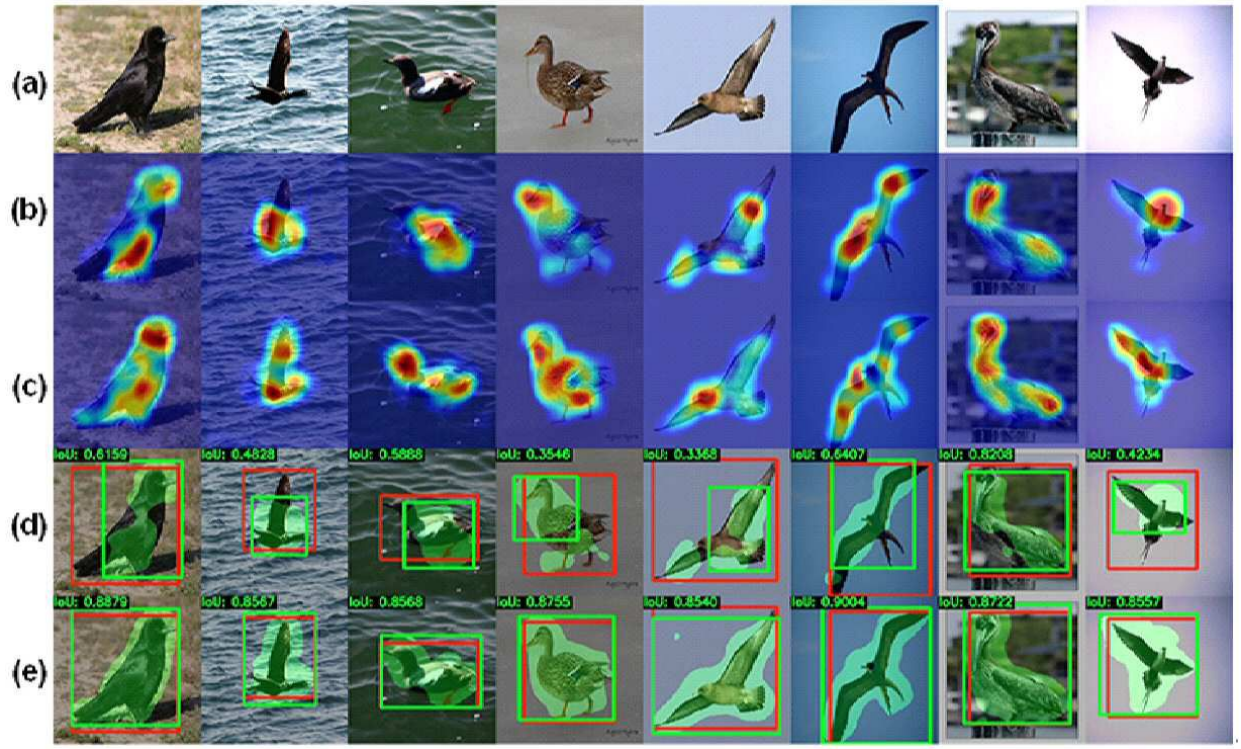


(a) Channel Attention (b) Spatial Attention (c) MoANA (Ours)

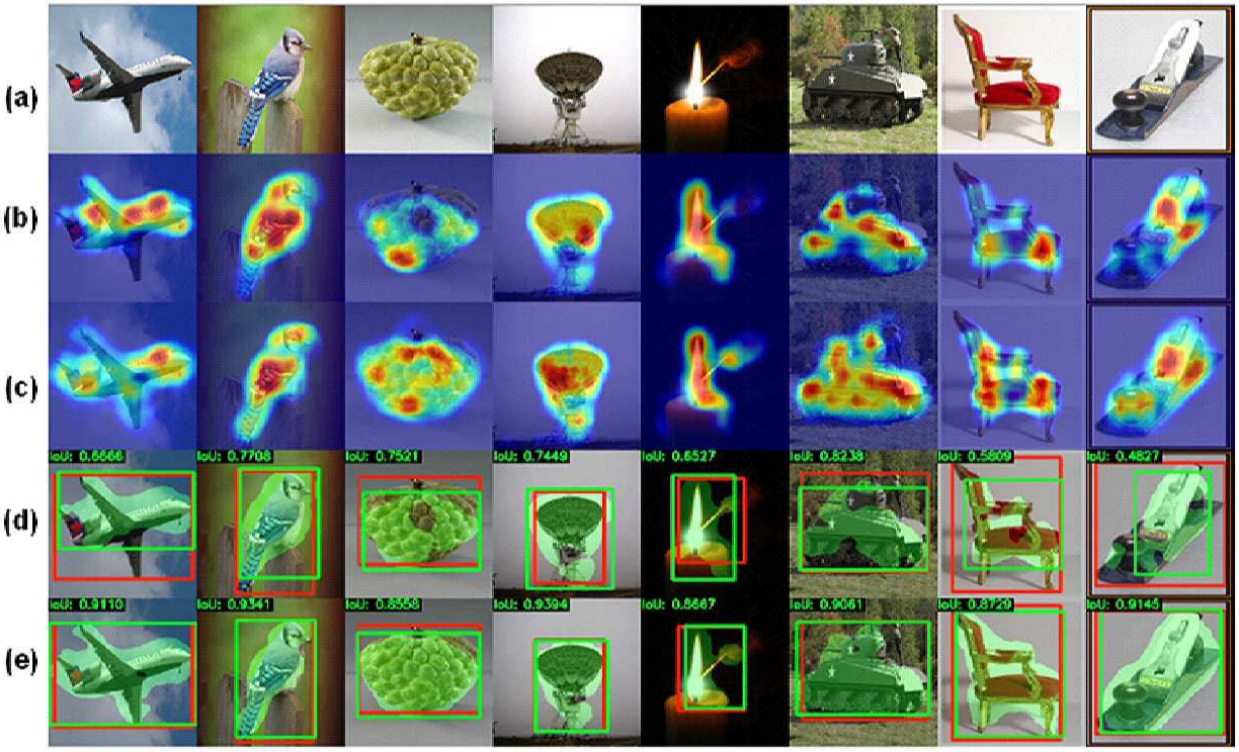
도면3


Hadamard Product
Element-wise Summation
Sigmoid Function
X
Input Feature ($C \times H \times W$)
Avg Pool
h
c
w
1D Conv BN
z_w
z_h
z_c
X
$f(z_c, z_h, z_w)$
M
$\hat{X}$
Output Feature ($C \times H \times W$)

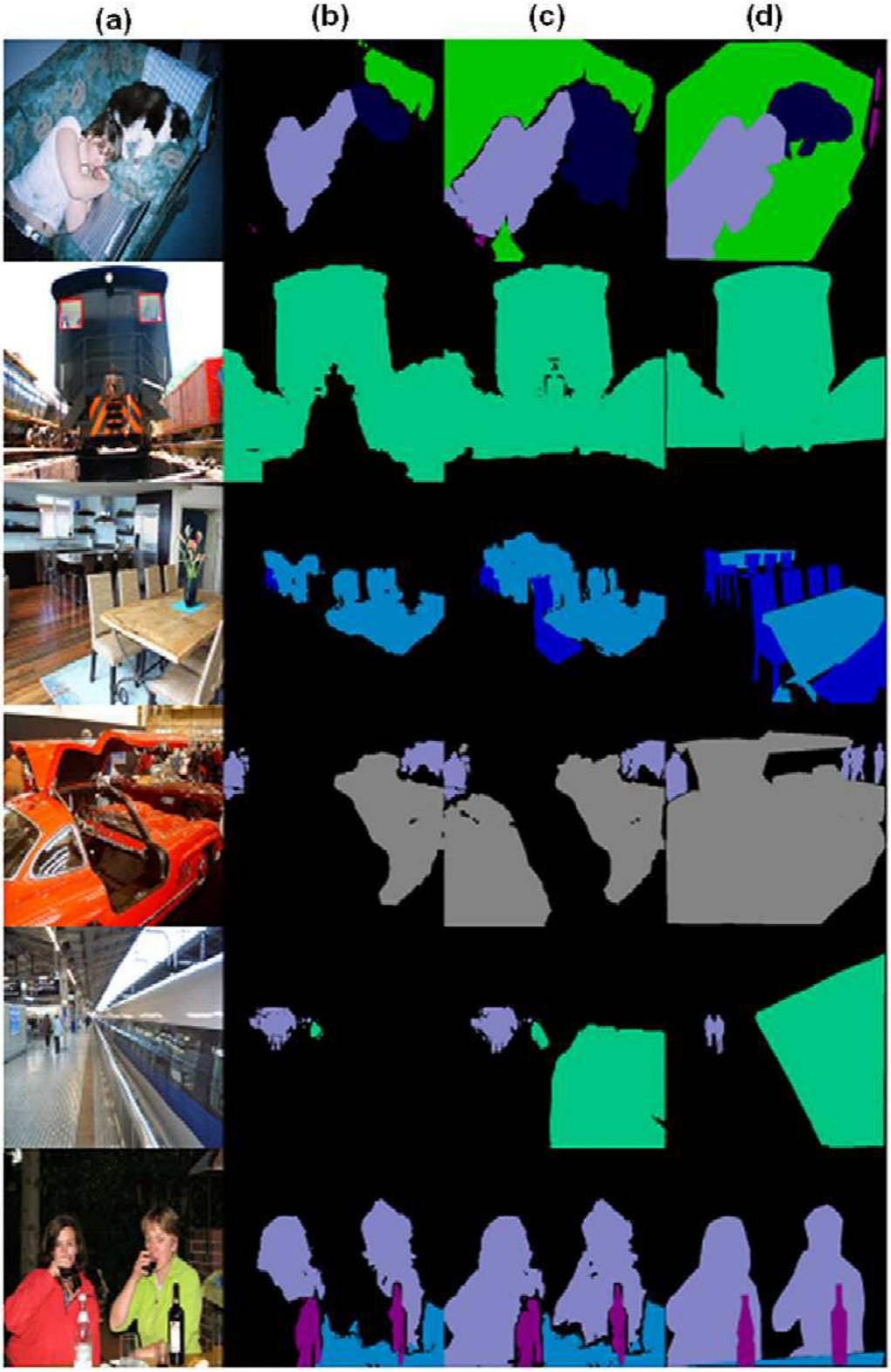
도면4


(a)
(b)
(c)
(d)
IoU: 0.6150
IoU: 0.4828
IoU: 0.5888
IoU: 0.3546
IoU: 0.3368
IoU: 0.6407
IoU: 0.8208
IoU: 0.4234
(e)
IoU: 0.8879
IoU: 0.8567
IoU: 0.8568
IoU: 0.8755
IoU: 0.8540
IoU: 0.9004
IoU: 0.8722
IoU: 0.8557

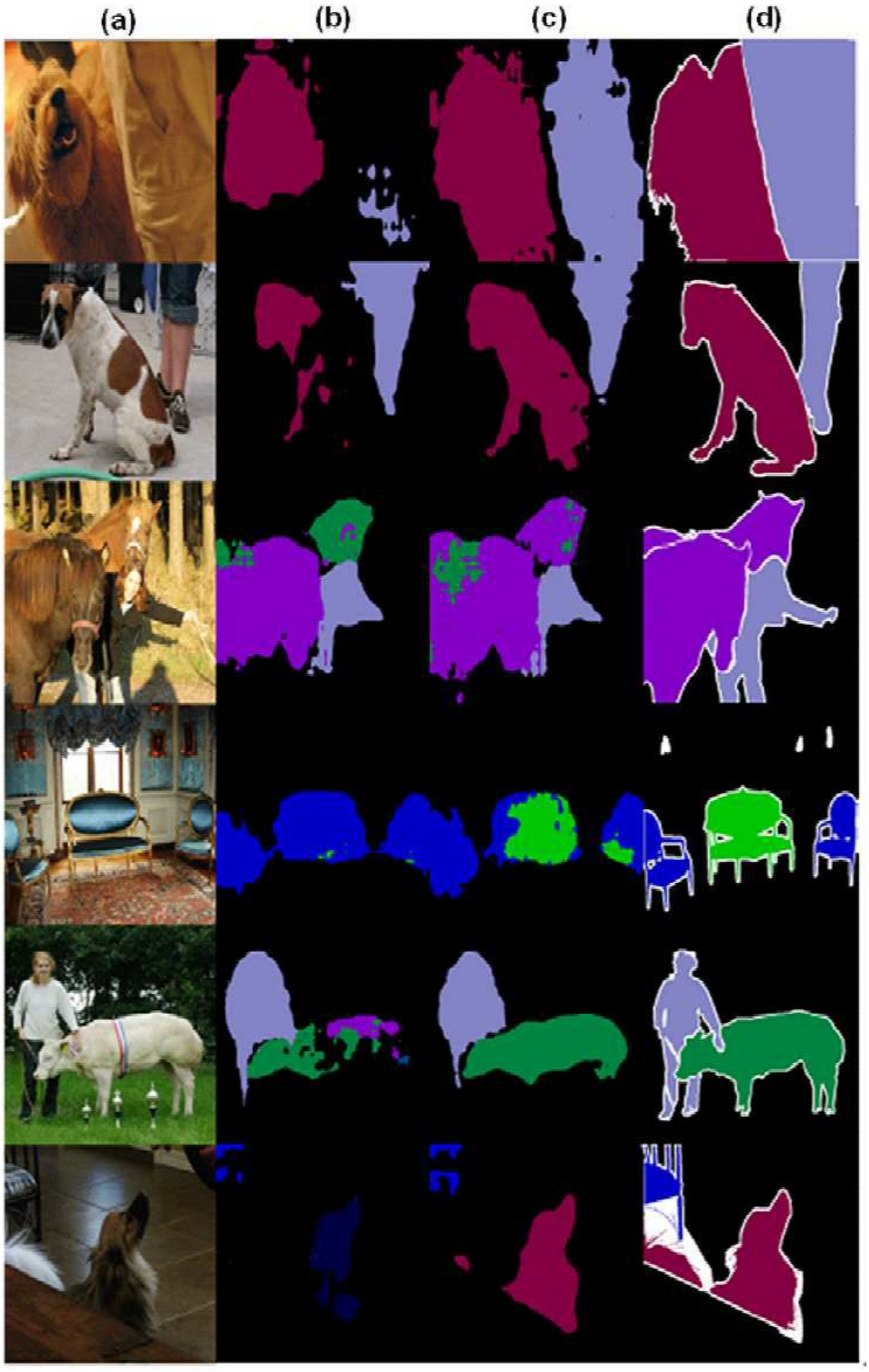
도면5


(a)
(b)
(c)
(d)
IoU: 0.8666
IoU: 0.7708
IoU: 0.7521
IoU: 0.7449
IoU: 0.6527
IoU: 0.8238
IoU: 0.5809
IoU: 0.4827
(e)
IoU: 0.9110
IoU: 0.9341
IoU: 0.8558
IoU: 0.9394
IoU: 0.8667
IoU: 0.9061
IoU: 0.8729
IoU: 0.9145

도면6



도면7



도면8

