

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2024-0043536
(43) 공개일자 2024년04월03일

(51) 국제특허분류(Int. Cl.)
G06N 3/08 (2023.01) G06N 3/04 (2023.01)
G06T 7/20 (2017.01)
(52) CPC특허분류
G06N 3/08 (2023.01)
G06N 3/045 (2023.01)
(21) 출원번호 10-2022-0122752
(22) 출원일자 2022년09월27일
심사청구일자 2022년09월27일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
고성제
서울특별시 서초구 방배로 270, 나-304호 (방배동, 삼호아파트)
조성진
경기도 성남시 수정구 위례중안로 216, 5313동 1102호 (창곡동, 위례 호반베르디움)
(74) 대리인
윤귀상

전체 청구항 수 : 총 10 항

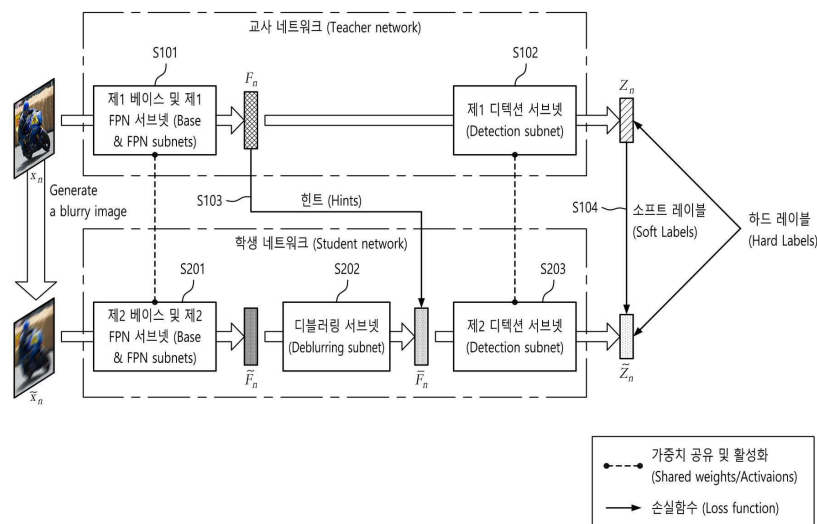
(54) 발명의 명칭 교사 모델 및 학생 모델의 학습을 위한 지식중류 기반 시스템

(57) 요약

잔상이 있는 영상에서도 관심 객체 검출 성능이 향상될 수 있는 시스템에 대하여 개시한다.

개시하는 잔상이 있는 영상에서도 관심 객체 검출 성능이 향상될 수 있는 시스템의 일 실시예에 따르면, 제1 베이스 서브넷(Base Subnet), 제1 피쳐 피라미드 서브넷(Feature Pyramid Subnet) 및 제1 디텍션 서브넷(Detection Subnet)을 포함하는 교사 네트워크(teacher network) 및 제2 베이스 서브넷, 제2 피쳐 피라미드 서브넷, 디블러링 서브넷(Deblurring Subnet) 및 제2 디텍션 서브넷을 포함하는 학생 네트워크(student network)를 포함하고, 상기 교사 네트워크에서 상기 학생 네트워크로 힌트(hints)를 제공할 수 있다.

대표도 - 도2



(52) CPC특허분류

G06T 7/20 (2013.01)

G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711126469
과제번호	2014-3-00077-008
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	혁신성장동력프로젝트(R&D)(과기정통부)
연구과제명	(딥뷰-2세부) 대규모 실시간 비디오 분석에 의한 전역적 다중 관심객체 추적 및 상
항 예측 기술 개발	
기 여 율	1/1
과제수행기관명	광주과학기술원
연구기간	2022.01.01 ~ 2022.12.31

명세서

청구범위

청구항 1

제1 베이스 서브넷(Base Subnet), 제1 피쳐 피라미드 서브넷(Feature Pyramid Subnet) 및 제1 디텍션 서브넷(Detection Subnet)을 포함하는 교사 네트워크(teacher network); 및

제2 베이스 서브넷, 제2 피쳐 피라미드 서브넷, 디블러링 서브넷(Deblurring Subnet) 및 제2 디텍션 서브넷을 포함하는 학생 네트워크(student network);를 포함하고,

상기 교사 네트워크에서 상기 학생 네트워크로 힌트(hints)를 제공하는 것을 특징으로 하는, 시스템.

청구항 2

청구항 1에 있어서,

상기 디블러링 서브넷은,

1x1 컨볼루션 커널(1x1 convolution kernel); 및

3x3 디포머블 컨볼루션 커널(3x3 Deformable kernel)를 포함하고,

상기 3x3 디포머블 컨볼루션 커널은 3x3 컨볼루션 커널(3x3 convolution kernel)과 오프셋 필드(Offset field)를 포함하고,

상기 오프셋 필드는 상기 3x3 컨볼루션 커널(3x3 convolution kernel)을 통해 추정되는 것을 특징으로 하는 시스템.

청구항 3

청구항 1에 있어서,

상기 제1 베이스 서브넷 및 제1 피쳐 피라미드 서브넷은 아래 수학식 1로 표현되는 과정을 수행하고,

상기 제2 베이스 서브넷 및 제2 피쳐 피라미드 서브넷은 아래 수학식 2로 표현되는 과정을 수행하는 것을 특징으로 하는, 시스템.

(수학식 1)

$$F_n \triangleq \{f_n^l\}_{l=1}^L = \mathcal{E}(\mathbf{x}_n)$$

(수학식 2)

$$\tilde{F}_n \triangleq \{\tilde{f}_n^l\}_{l=1}^L = \mathcal{E}(\tilde{\mathbf{x}}_n)$$

(여기서, F_n 는 상기 수학식 1을 통해 도출된 제1 파라미터, $\tilde{F}_n$ 는 상기 수학식 2를 통해 도출된 제2 파라미터, $\{f_n^l\}_{l=1}^L$ 는 제1 피쳐 피라미드 서브넷으로부터 l 번째 추출된 피쳐, $\{\tilde{f}_n^l\}_{l=1}^L$ 는 제2 피쳐 피라미드 서브넷으로부터 l 번째 추출된 피쳐, $\mathcal{E}(\cdot)$ 는 베이스 서브넷과 피쳐피라미드 서브넷의 혼합 함수, $\mathbf{x}_n$ 는 n 번째 선명한 이미지, $\tilde{\mathbf{x}}_n$ 는 n 번째 블러링된 이미지를 의미함)

청구항 4

청구항 3에 있어서,

상기 제1 디텍션 서브넷은 아래 수학적 식 3으로 표현되는 과정을 수행하고,

상기 제2 디텍션 서브넷은 아래 수학적 식 4로 표현되는 과정을 수행하는 것을 특징으로 하는 시스템.

(수학적 식 3)

$$Z_n \triangleq \{(\mathbf{z}_n^k, \mathbf{b}_n^k)\}_{k=1}^K = \mathcal{D}(F_n)$$

(수학적 식 4)

$$\tilde{Z}_n \triangleq \{(\tilde{\mathbf{z}}_n^k, \tilde{\mathbf{b}}_n^k)\}_{k=1}^K = \mathcal{D}(\mathcal{H}(\tilde{F}_n))$$

(여기서, Z_n 는 상기 수학적식3을 통해 도출된 제3 파라미터, $\tilde{Z}_n$ 는 상기 수학적 식 4를 통해 도출된 제4 파라미터, $\mathbf{z}_n^k$ 는 교사 네트워크에서 소프트맥스(softmax)함수를 적용하기 전의 k 번째 분류 로짓(logit), $\tilde{\mathbf{z}}_n^k$ 는 학생 네트워크에서 소프트맥스(softmax)함수를 적용하기 전의 k 번째 분류 로짓(logit), $\mathbf{b}_n^k$ 는 n 번째 선명한 이미지에서 측정된 객체 경계 상자(bounding box), $\tilde{\mathbf{b}}_n^k$ 는 n 번째 블러링 이미지에서 측정된 객체 경계 상자, $\mathcal{D}(\cdot)$ 는 제1 및 제2 디텍션 서브넷을 도출하기 위한 함수, $\mathcal{H}(\cdot)$ 는 디블러링 서브넷을 도출하기 위한 함수를 의미함)

청구항 5

청구항 1에 있어서,

상기 시스템의 전체 손실함수가 수학적 식 5로 표현되는 것을 특징으로 하는, 시스템.

(수학적 식 5)

$$L_{\text{total}} = L_{\text{det}} + \lambda_1 L_{\text{SG}} + \lambda_2 L_{\text{deblur}}$$

(여기서, L_{total} 는 전체 손실함수, L_{det} 는 제1 디텍션 서브넷 및 제2 디텍션 서브넷의 손실 함수, L_{SG} 는 확률 분포의 평활도를 제어하는 매개 변수, L_{deblur} 는 디블러링 서브넷의 손실함수, λ_1 및 λ_2 는 0.5를 의미함)

청구항 6

청구항 5에 있어서,

상기 L_{deblur} 는 아래 수학적 식 6으로 표현되는 과정을 수행하는 것을 특징으로 하는 시스템.

(수학적 식 6)

$$L_{\text{deblur}} = \sum_{n=1}^N \sum_{l=1}^L \left\| \tilde{\mathbf{f}}_n^l - \mathbf{f}_n^l \right\|_1$$

(여기서, M 은 $\sum_{n=1}^N \sum_{k=1}^K w_n^k$, $\mathbf{f}_n^l$ 는 n 번째 선명한 이미지에서 측정된 객체 경계 상자(bounding box), $\tilde{\mathbf{f}}_n^l$ 는 n 번째 블러링 이미지에서 측정된 객체 경계 상자, $\|\cdot\|_1$ 는 "1-norm"을 의미함)

청구항 7

청구항 5에 있어서,

상기 L_{SG} 는 아래 수학적 식 7로 표현되는 것을 특징으로 하는 시스템.

(수학적 식 7)

$$L_{SG} = L_{\text{conf}} + L_{\text{reg}}$$

(여기서, L_{SG} 는 자기주도학습(self-guided learning)의 전체손실함수, L_{conf} 는 선명한 피쳐(feature)에서 블러링된 피쳐로 분류 지식을 전달하기 위한 목적 함수, L_{reg} 는 자기주도학습의 회귀 손실함수를 의미함)

청구항 8

청구항 7에 있어서,

상기 학생 네트워크는 상기 교사 네트워크로부터 소프트 레이블(soft-labels)에 의해 지도를 받으며,

상기 소프트 레이블은 아래 수학적 식 8의 과정으로 연산되는 것을 특징으로 하는 시스템.

(수학적 식 8)

$$\mathbf{p}_{n,c}^k = \frac{\exp(\mathbf{z}_{n,c}^k / \tau)}{\sum_{c=1}^C \exp(\mathbf{z}_{n,c}^k / \tau)}$$

(여기서, $\mathbf{p}_n^k$ 는 n번째 선명한 이미지로부터 얻은 소프트 추정치, $\mathbf{z}_n^k$ 는 교사 네트워크에서 소프트맥스(softmax) 함수를 적용하기 전의 k번째 분류 로짓(logit), τ 는 확률 분포의 평활도를 제어하는 매개 변수를 의미함)

청구항 9

청구항 8에 있어서,

상기 L_{conf} 는 아래 수학적 식 9로 표현되는 것을 특징으로 하는 시스템.

(수학적 식 9)

$$L_{\text{conf}} = -\frac{1}{M} \sum_{n=1}^N \sum_{k=1}^K (w_n^k + v_n^k) \text{CE}(\tilde{\mathbf{p}}_n^k, \mathbf{p}_n^k)$$

(여기서, M 은 $\sum_{n=1}^N \sum_{k=1}^K w_n^k$, $\tilde{\mathbf{p}}_n^k$ 는 n번째 블러링된 이미지로부터 얻은 소프트 추정치, $\text{CE}(\cdot, \cdot)$ 는 교차 엔트로피 손실함수를 의미함)

청구항 10

청구항 5에 있어서,

상기 L_{reg} 는 아래 수학적 식 10으로 표현되는 것을 특징으로 하는 시스템.

(수학식 10)

$$L_{\text{reg}} = \frac{1}{M} \sum_{n=1}^N \sum_{k=1}^K w_n^k \left\| \tilde{\mathbf{b}}_n^k - \mathbf{b}_n^k \right\|_1$$

(여기서, M 은 $\sum_{n=1}^N \sum_{k=1}^K w_n^k$, $\mathbf{b}_n^k$ 는 n 번째 선명한 이미지에서 측정된 객체 경계 상자(bounding box), $\tilde{\mathbf{b}}_n^k$ 는 n 번째 블러링 이미지에서 측정된 객체 경계 상자, $\|\cdot\|_1$ 는 "1-norm"을 의미함)

발명의 설명

기술 분야

[0001] 본 발명은 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템으로서, 보다 상세하게는 지식증류(Knowledge Distillation) 기반으로 교사 네트워크에서 습득한 지식을 활용하여 학생 네트워크를 트레이닝함으로써 잔상이 있는 영상에서도 관심 객체 검출 성능이 향상될 수 있는 시스템에 관한 것이다.

배경 기술

[0002] 관심 객체 검출 기술은 단일 영상을 입력으로 하여 영상 내에 존재하는 객체의 위치 및 종류를 특정하는 기술이다. 최근, 자율 주행 자동차 시장 규모가 확대되고 보안 및 감시 분야의 산업이 확장됨에 따라 보다 나은 성능을 위해 객체 검출 기술의 성능 고도화의 필요성이 대두되고 있다.

[0003] 관심 객체 검출 기술의 핵심은 빠르고 정확하게 영상 내에서 관심이 있는 종류의 객체들의 위치와 카테고리를 특정하는 것이다. 관심 객체 검출 기술은 대부분 상용화된 카메라로 취득된 영상에서 수행되는데, 취득된 영상은 카메라의 흔들림, 혹은 객체의 빠른 움직임 등으로 인해 영상에 잔상이 생길 수 있다. 취득한 영상의 잔상은 보기에만 안좋은 뿐 아니라 검출 시스템 전체의 성능을 저하시킬 수 있다.

[0004] PACAL VOC(비특허문헌 1)과 COCO(비특허문헌 2)의 관심 객체 감지를 위한 일반적인 데이터 세트(Data set)는 이미지의 노이즈 및 모션 블러(motion blur)를 적절하게 고려하지 않고 있다. 이에 따라, 위 문헌에서는 움직임으로 인한 영상에서 잔상이 생겨 흐릿한 이미지가 입력된 경우 객체 감지가 실패할 수 있다.

[0005] 한편, 지식 증류법(knowledge distillation)은 학습된 인공 신경망을 실제 서비스에 적용하기 위해 효과적으로 모델을 압축하여 속도를 향상시키는 네트워크 압축 방법 중 하나이다. 지식 증류 방법은 교사 네트워크(Teacher Network)에서 습득한 지식을 활용하여 학생 네트워크(Student Network)를 트레이닝 한다. 학생 네트워크는 사전에 트레이닝 된 교사 네트워크를 모방함으로써, 상대적으로 적은 파라미터를 가지고 있어도 학습 모델의 성능을 향상시킬 수 있다. 또한, 학습된 학생 네트워크는 단독으로 학습되었을 때 보다 우수한 성능을 보인다고 알려져 있다.

[0006] 지식 증류법은 교사 네트워크라고 불리는 비교적 성능이 뛰어난 네트워크를 이용하여 학생 네트워크라고 불리는 연산량 및 메모리 사용량이 적은 네트워크를 학습시킬 수 있다. 이때 교사 네트워크의 특징 맵과 학생 네트워크에서 특징 맵을 비교하는 손실 함수를 정의하고, 이를 이용하여 학생 네트워크를 학습시키게 된다.

선행기술문헌

비특허문헌

[0007] (비특허문헌 0001) M. Everingham, L.Van Gool, C. K. I.Williams, J.Winn, andW. Zisserman, "The PASCAL visual object classes (VOC) challenge," Int. J. Comput. Vis., vol. 88, no. 2, Sep. 2010.

(비특허문헌 0002) T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Doll ar, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in Proc. Eur. Conf. Comput. Vis., 2014.

발명의 내용

해결하려는 과제

[0008] 상기와 같은 문제점을 해결하기 위한 본 발명의 목적은 지식증류기반으로 교사 네트워크에서 습득한 지식을 활용하여 학생 네트워크를 트레이닝함으로써 잔상이 있는 영상에서도 관심 객체 검출 성능이 향상될 수 있는 시스템을 제공하는 것이다.

과제의 해결 수단

[0009] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 시스템은 제1 베이스 서브넷(Base Subnet), 제1 피쳐 피라미드 서브넷(Feature Pyramid Subnet) 및 제1 디텍션 서브넷(Detection Subnet)을 포함하는 교사 네트워크(teacher network) 및 제2 베이스 서브넷, 제2 피쳐 피라미드 서브넷, 디블러링 서브넷(Deblurring Subnet) 및 제2 디텍션 서브넷을 포함하는 학생 네트워크(student network)를 포함하고, 상기 교사 네트워크에서 상기 학생 네트워크로 힌트(hints)를 제공할 수 있다.

[0010] 여기서, 상기 디블러링 서브넷은, 1x1 컨볼루션 커널(1x1 convolution kernel) 및 3x3 디포머블 컨볼루션 커널(3x3 Deformable kernel)를 포함하고, 상기 3x3 디포머블 컨볼루션 커널은 3x3 컨볼루션 커널(3x3 convolution kernel)과 오프셋 필드(Offset field)를 포함하고, 상기 오프셋 필드는 상기 3x3 컨볼루션 커널(3x3 convolution kernel)을 통해 추정될 수 있다.

[0011] 여기서, 상기 제1 베이스 서브넷 및 제1 피쳐 피라미드 서브넷은 아래 수학식 1로 표현되는 과정을 수행하고, 상기 제2 베이스 서브넷 및 제2 피쳐 피라미드 서브넷은 아래 수학식 2로 표현되는 과정을 수행할 수 있다.

[0012] (수학식 1)

$$F_n \triangleq \{\mathbf{f}_n^l\}_{l=1}^L = \mathcal{E}(\mathbf{x}_n)$$

[0013]

[0014] (수학식 2)

$$\tilde{F}_n \triangleq \{\tilde{\mathbf{f}}_n^l\}_{l=1}^L = \mathcal{E}(\tilde{\mathbf{x}}_n)$$

[0015]

[0016] 여기서, F_n 는 상기 수학식 1을 통해 도출된 제1 파라미터, $\tilde{F}_n$ 는 상기 수학식 2를 통해 도출된 제2 파라미터, $\{\mathbf{f}_n^l\}_{l=1}^L$ 는 제1 피쳐 피라미드 서브넷으로부터 l 번째 추출된 피쳐, $\{\tilde{\mathbf{f}}_n^l\}_{l=1}^L$ 는 제2 피쳐 피라미드 서브넷으로부터 l 번째 추출된 피쳐, $\mathcal{E}(\cdot)$ 는 베이스 서브넷과 피쳐피라미드 서브넷의 혼합 함수, $\mathbf{x}_n$ 는 n 번째 선명한 이미지, $\tilde{\mathbf{x}}_n$ 는 n 번째 블러링된 이미지를 의미한다.

[0017] 여기서, 상기 제1 디텍션 서브넷은 아래 수학식 3으로 표현되는 과정을 수행하고, 상기 제2 디텍션 서브넷은 아래 수학식 4로 표현되는 과정을 수행한다.

[0018] (수학식 3)

$$Z_n \triangleq \{(\mathbf{z}_n^k, \mathbf{b}_n^k)\}_{k=1}^K = \mathcal{D}(F_n)$$

[0019]

[0020] (수학식 4)

$$\tilde{Z}_n \triangleq \{(\tilde{\mathbf{z}}_n^k, \tilde{\mathbf{b}}_n^k)\}_{k=1}^K = \mathcal{D}(\mathcal{H}(\tilde{F}_n))$$

[0021]

[0022] 여기서, Z_n 는 상기 수학식3을 통해 도출된 제3 파라미터, $\tilde{Z}_n$ 는 상기 수학식 4를 통해 도출된 제4 파라미터,

$\mathbf{z}_n^k$ 는 교사 네트워크에서 소프트맥스(softmax)함수를 적용하기 전의 k 번째 분류 로짓(logit), $\tilde{\mathbf{z}}_n^k$ 는 학생 네트워크에서 소프트맥스(softmax)함수를 적용하기 전의 k 번째 분류 로짓(logit), $\mathbf{b}_n^k$ 는 n 번째 선명한 이미지에서 측정된 객체 경계 상자(bounding box), $\tilde{\mathbf{b}}_n^k$ 는 n 번째 블러링 이미지에서 측정된 객체 경계 상자, $\mathcal{D}(\cdot)$ 는 제1 및 제2 디텍션 서브넷을 도출하기 위한 함수, $\mathcal{H}(\cdot)$ 는 디블러링 서브넷을 도출하기 위한 함수를 의미한다.

여기서, 상기 시스템의 전체 손실함수가 수학식 5로 표현된다.

(수학식 5)

$$L_{\text{total}} = L_{\text{det}} + \lambda_1 L_{\text{SG}} + \lambda_2 L_{\text{deblur}}$$

여기서, L_{total} 는 전체 손실함수, L_{det} 는 제1 디텍션 서브넷 및 제2 디텍션 서브넷의 손실 함수, L_{SG} 는 확률 분포의 평활도를 제어하는 매개 변수, L_{deblur} 는 디블러링 서브넷의 손실함수, λ_1 및 λ_2 는 0.5를 의미한다.

여기서, 상기 L_{deblur} 는 아래 수학식 6으로 표현되는 과정을 수행한다.

(수학식 6)

$$L_{\text{deblur}} = \sum_{n=1}^N \sum_{l=1}^L \left\| \tilde{\mathbf{f}}_n^l - \mathbf{f}_n^l \right\|_1$$

여기서, M 은 $\sum_{n=1}^N \sum_{k=1}^K w_n^k$, $\mathbf{f}_n^l$ 는 n 번째 선명한 이미지에서 측정된 객체 경계 상자(bounding box), $\tilde{\mathbf{f}}_n^l$ 는 n 번째 블러링 이미지에서 측정된 객체 경계 상자, $\|\cdot\|_1$ 는 "1-norm"을 의미한다.

여기서, 상기 L_{SG} 는 아래 수학식 7로 표현될 수 있다.

(수학식 7)

$$L_{\text{SG}} = L_{\text{conf}} + L_{\text{reg}}$$

여기서, L_{SG} 는 자기주도학습(self-guided learning)의 전체손실함수, L_{conf} 는 선명한 피쳐(feature)에서 블러링된 피쳐로 분류 지식을 전달하기 위한 목적 함수, L_{reg} 는 자기주도학습의 회귀 손실함수를 의미함)

상기 학생 네트워크는 상기 교사 네트워크로부터 소프트 레이블(soft-labels)에 의해 지도를 받으며,

여기서, 상기 소프트 레이블은 아래 수학식 8의 과정으로 연산되는 것을 특징으로 하는 시스템.

(수학식 8)

$$\mathbf{p}_{n,c}^k = \frac{\exp(\mathbf{z}_{n,c}^k / \tau)}{\sum_{c=1}^C \exp(\mathbf{z}_{n,c}^k / \tau)}$$

(여기서, $\mathbf{p}_n^k$ 는 n 번째 선명한 이미지로부터 얻은 소프트 추정치, $\mathbf{z}_n^k$ 는 교사 네트워크에서 소프트맥스

(softmax) 함수를 적용하기 전의 k 번째 분류 로짓(logit), τ 는 확률 분포의 평활도를 제어하는 매개 변수를 의미함)

여기서, 상기 L_{conf} 는 아래 수학적 식 9로 표현되는 것을 특징으로 하는 시스템.

(수학적 식 9)

$$L_{\text{conf}} = -\frac{1}{M} \sum_{n=1}^N \sum_{k=1}^K (w_n^k + v_n^k) \text{CE}(\tilde{\mathbf{p}}_n^k, \mathbf{p}_n^k)$$

(여기서, M 은 $\sum_{n=1}^N \sum_{k=1}^K w_n^k$, $\tilde{\mathbf{p}}_n^k$ 는 n 번째 블러링된 이미지로부터 얻은 소프트 추정치, $\text{CE}(\cdot, \cdot)$ 는 교차 엔트로피 손실함수를 의미함)

여기서, 상기 L_{reg} 는 아래 수학적 식 10으로 표현되는 것을 특징으로 하는 시스템.

(수학적 식 10)

$$L_{\text{reg}} = \frac{1}{M} \sum_{n=1}^N \sum_{k=1}^K w_n^k \left\| \tilde{\mathbf{b}}_n^k - \mathbf{b}_n^k \right\|_1$$

(여기서, M 은 $\sum_{n=1}^N \sum_{k=1}^K w_n^k$, $\mathbf{b}_n^k$ 는 n 번째 선명한 이미지에서 측정된 객체 경계 상자(bounding box), $\tilde{\mathbf{b}}_n^k$ 는 n 번째 블러링 이미지에서 측정된 객체 경계 상자, $\|\cdot\|_1$ 는 "1-norm"을 의미함)

발명의 효과

본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템은 지식증류 기반으로 교사 네트워크에서 습득한 지식을 활용하여 학생 네트워크를 트레이닝함으로써 잔상이 있는 영상에서도 관심 객체 검출 성능이 향상될 수 있다.

다만, 본 발명의 실시 예들에 따른 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템이 달성할 수 있는 효과는 이상에서 언급한 것들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

본 발명에 관한 이해를 돕기 위해 상세한 설명의 일부로 포함되는, 첨부도면은 본 발명에 대한 실시예를 제공하고, 상세한 설명과 함께 본 발명의 기술적 사상을 설명한다.

도 1은 본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템의 블록도를 도시한 것이다.

도 2는 본 발명에 따른 더블러링 서브넷의 구조를 도시한 것이다.

도 3은 본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템의 개념도를 도시한 것이다.

도 4는 본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템의 제1 실시예를 도시한 것이다.

도 5는 본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템의 제2 실시예를 도시한 것이다.

도 6은 본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식중류 기반 시스템의 제3 실시예를 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

- [0051] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0052] 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [0053] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.
- [0054] 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0055] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가진 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0056] 이하, 첨부한 도면들을 참조하여, 본 발명의 바람직한 실시예를 보다 상세하게 설명하고자 한다. 본 발명을 설명함에 있어 전체적인 이해를 용이하게 하기 위하여 도면상의 동일한 구성요소에 대해서는 동일한 참조부호를 사용하고 동일한 구성요소에 대해서 중복된 설명은 생략한다.
- [0058] 도 1은 본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식중류 기반 시스템의 블록도를 도시한 것이다.
- [0059] 도 1을 참조하면, 본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식중류 기반 시스템(1000)은 교사 네트워크(100)와 학생 네트워크(200)를 포함할 수 있다. 상기 교사 네트워크(100)는 제1 베이스 서브넷과 제1 피쳐피라미드 서브넷(110) 및 제1 디텍션 서브넷(120)을 포함할 수 있다. 상기 학생 네트워크(200)는 제2 베이스 서브넷과 제2 피쳐피라미드 서브넷(210), 디블러링 서브넷(220) 및 제2 디텍션 서브넷(230)을 포함할 수 있다.
- [0061] 도 2는 본 발명에 따른 디블러링 서브넷(220)의 구조를 도시한 것이다.
- [0062] 도 2를 참조하면, 상기 디블러링 서브넷(220)은 $r=4$ 의 감소 비율로 변형 가능한 컨볼루션의 계산 복잡성을 줄이기 위해 1×1 컨볼루션 커널(1×1 convolution kernel) 및 3×3 디포머블 컨볼루션 커널(3×3 Deformable kernel)을 포함할 수 있다.
- [0063] 상기 3×3 디포머블 컨볼루션 커널은 3×3 컨볼루션 커널(3×3 convolution kernel)과 오프셋 필드(Offset field)를 포함할 수 있다. 상기 3×3 디포머블 컨볼루션 커널에서는 상기 3×3 컨볼루션 커널과 오프셋 필드를 기반으로 커널모양을 변형하고, 변형된 커널을 사용하여 피쳐를 추출할 수 있다.

[0064] 상기 오프셋 필드는 채널이 축소된 피쳐 맵에서 적응형 컨볼루션의 샘플링 동작을 위해 3x3 컨볼루션 커널을 통해 추정될 수 있다. 상기 오프셋 필드는 채널이 축소된 피쳐 맵(feature map)과 동일한 공간적 크기를 가지며 각 픽셀은 3x3 컨볼루션 커널의 수평 및 수직 오프셋으로 구성될 수 있다.

[0066] 도 3은 본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템의 개념도를 도시한 것이다.

[0067] 도3을 참조하면, 교사 네트워크(100)의 제1 베이스 서브넷 및 제1 피쳐피라미드 서브넷(110)에 선명한 이미지가 입력될 수 있으며, 상기 제1베이스 서브넷 및 제1 피쳐피라미드 서브넷(110)은 아래 수학식 2와 같은 과정(S101)을 수행할 수 있다.

[0068] (수학식 1)

$$F_n \triangleq \{\mathbf{f}_n^l\}_{l=1}^L = \mathcal{E}(\mathbf{x}_n)$$

[0070] 여기서, F_n 는 상기 수학식 1을 통해 도출된 제1 파라미터, $\{\mathbf{f}_n^l\}_{l=1}^L$ 는 제1 피쳐 피라미드 서브넷으로부터 l 번째 추출된 피쳐, $\mathcal{E}(\cdot)$ 는 베이스 서브넷과 피쳐피라미드 서브넷의 혼합 함수, $\mathbf{x}_n$ 는 n 번째 선명한 이미지를 의미한다.

[0071] 상기 S101단계를 통해 연산된 제1 파라미터는 제1 디텍션 서브넷(120)에 입력될 수 있으며, 상기 제1 디텍션 서브넷(120)은 아래 수학식 3과 같은 과정(S102)을 수행할 수 있다.

[0072] (수학식 3)

$$Z_n \triangleq \{(\mathbf{z}_n^k, \mathbf{b}_n^k)\}_{k=1}^K = \mathcal{D}(F_n)$$

[0074] 여기서, Z_n 는 상기 수학식3을 통해 도출된 제3 파라미터, $\mathbf{z}_n^k$ 는 교사 네트워크에서 소프트맥스(softmax)함수를 적용하기 전의 k 번째 분류 로짓(logit), $\mathbf{b}_n^k$ 는 n 번째 선명한 이미지에서 측정된 객체 경계 상자(bounding box, $\mathcal{D}(\cdot)$ 는 제1 및 제2 디텍션 서브넷을 도출하기 위한 함수를 의미한다.

[0076] 상기 학생 네트워크(200)의 제2 베이스 서브넷 및 제2 피쳐피라미드 서브넷(210)에 블러링된 이미지가 입력될 수 있으며, 상기 제2 베이스 서브넷 및 제2 피쳐피라미드 서브넷(210)은 아래 수학식 2와 같은 과정(S201)을 수행할 수 있다.

[0077] (수학식 2)

$$\tilde{F}_n \triangleq \{\tilde{\mathbf{f}}_n^l\}_{l=1}^L = \mathcal{E}(\tilde{\mathbf{x}}_n)$$

[0079] 여기서, $\tilde{F}_n$ 는 상기 수학식 2를 통해 도출된 제2 파라미터, $\{\tilde{\mathbf{f}}_n^l\}_{l=1}^L$ 는 제2 피쳐 피라미드 서브넷으로부터 l 번째 추출된 피쳐, $\mathcal{E}(\cdot)$ 는 베이스 서브넷과 피쳐피라미드 서브넷의 혼합 함수, $\tilde{\mathbf{x}}_n$ 는 n 번째 블러링된 이미지를 의미한다.

[0080] 상기 S201 단계를 통해 연산된 제2 파라미터는 디블러링 서브넷(220)에 입력될 수 있으며, 상기 디블러링 서브넷(220)이 동작하는 과정(S202)에서의 손실함수는 아래 수학식 6으로 표현될 수 있다.

[0081] (수학식 6)

$$L_{\text{deblur}} = \sum_{n=1}^N \sum_{l=1}^L \left\| \tilde{\mathbf{f}}_n^l - \mathbf{f}_n^l \right\|_1$$

[0082]

[0083] 여기서, L_{deblur} 는 디블러링 서브넷의 손실함수, M 은 $\sum_{n=1}^N \sum_{k=1}^K w_n^k$, $\mathbf{f}_n^l$ 는 n 번째 선명한 이미지에
서 측정된 객체 경계 상자(bounding box), $\tilde{\mathbf{f}}_n^l$ 는 n 번째 블러링 이미지에서 측정된 객체 경계 상자, $\|\cdot\|_1$ 는
"1-norm"을 의미한다.

[0084] 상기 교사 네트워크(100)에서 상기 학생 네트워크(200)로 힌트(hints)를 제공할 수 있다. 상기 디블러링 서브넷
(220)이 동작하는 과정(S202)과 상기 교사 네트워크(100)에서 제공하는 힌트(hints)를 통해 $\bar{F}_n$ 이 도출될 수
있다. 상기 $\bar{F}_n$ 은 제2 디텍션 서브넷(230)에 입력될 수 있으며, 제2 디텍션 서브넷(230)은 아래 수학식 4와
같은 과정(S203)을 수행할 수 있다.

[0085] (수학식 4)

$$\tilde{Z}_n \triangleq \{(\tilde{\mathbf{z}}_n^k, \tilde{\mathbf{b}}_n^k)\}_{k=1}^K = \mathcal{D}(\mathcal{H}(\tilde{F}_n))$$

[0086]

[0087] 여기서, $\tilde{Z}_n$ 는 상기 수학식 4를 통해 도출된 제4 파라미터, $\tilde{\mathbf{z}}_n^k$ 는 학생 네트워크에서 소프트맥스(softmax)함수
를 적용하기 전의 k 번째 분류 로짓(logit), $\tilde{\mathbf{b}}_n^k$ 는 n 번째 블러링 이미지에서 측정된 객체 경계 상자, $\mathcal{D}(\cdot)$ 는
제1 및 제2 디텍션 서브넷을 도출하기 위한 함수, $\mathcal{H}(\cdot)$ 는 디블러링 서브넷을 도출하기 위한 함수, $\tilde{F}_n$ 는 상기
수학식 2를 통해 도출된 제2 파라미터를 의미한다.

[0089] 상기 시스템(1000)의 전체 손실함수는 아래 수학식 5로 표현될 수 있다.

[0090] (수학식 5)

$$L_{\text{total}} = L_{\text{det}} + \lambda_1 L_{\text{SG}} + \lambda_2 L_{\text{deblur}}$$

[0091]

[0092] 여기서, L_{total} 는 전체 손실함수, L_{det} 는 제1 디텍션 서브넷 및 제2 디텍션 서브넷의 손실 함수, L_{SG} 는 자
기지도학습(self-guided learning)의 전체손실함수, L_{deblur} 는 상기 디블러링 서브넷의 손실함수, λ_1 및 λ_2
는 확률 분포의 평활도를 제어하는 변수로 0.5를 의미한다.

[0093] 상기 자기지도학습의 전체손실함수는 아래 수학식 7로 표현될 수 있다.

[0094] (수학식 7)

$$L_{\text{SG}} = L_{\text{conf}} + L_{\text{reg}}$$

[0095]

[0096] 여기서, L_{SG} 는 자기지도학습의 전체손실함수, L_{conf} 는 선명한 피쳐(feature)에서 블러링된 피쳐로 분류 지
식을 전달하기 위한 목적 함수, L_{reg} 는 자기지도학습의 회귀 손실함수를 의미한다.

[0097] 상기 학생 네트워크(200)는 상기 교사 네트워크(100)로부터 소프트 레이블(soft-labels) 및/또는 하드 레이블

(hard-labels)에 의해 지도를 받을 수 있다. 상기 소프트 레이블은 아래 수학적 8의 과정으로 연산될 수 있다.

(수학적 8)

$$\mathbf{p}_{n,c}^k = \frac{\exp(\mathbf{z}_{n,c}^k/\tau)}{\sum_{c=1}^C \exp(\mathbf{z}_{n,c}^k/\tau)}$$

여기서, $\mathbf{p}_n^k$ 는 n번째 선명한 이미지로부터 얻은 소프트 추정치, $\mathbf{z}_n^k$ 는 교사 네트워크에서 소프트맥스(softmax) 함수를 적용하기 전의 k 번째 분류 로짓(logit), τ 는 확률 분포의 평활도를 제어하는 매개 변수를 의미한다.

상기 L_{conf} 는 아래 수학적 9로 표현될 수 있다.

(수학적 9)

$$L_{\text{conf}} = -\frac{1}{M} \sum_{n=1}^N \sum_{k=1}^K (w_n^k + v_n^k) \text{CE}(\tilde{\mathbf{p}}_n^k, \mathbf{p}_n^k)$$

여기서, M 은 $\sum_{n=1}^N \sum_{k=1}^K w_n^k$, $\tilde{\mathbf{p}}_n^k$ 는 n번째 블러링된 이미지로부터 얻은 소프트 추정치, $\text{CE}(\cdot, \cdot)$ 는 교차 엔트로피 손실함수를 의미한다.

상기 L_{reg} 는 아래 수학적 10으로 표현될 수 있다.

(수학적 10)

$$L_{\text{reg}} = \frac{1}{M} \sum_{n=1}^N \sum_{k=1}^K w_n^k \|\tilde{\mathbf{b}}_n^k - \mathbf{b}_n^k\|_1$$

여기서, M 은 $\sum_{n=1}^N \sum_{k=1}^K w_n^k$, $\mathbf{b}_n^k$ 는 n번째 선명한 이미지에서 측정된 객체 경계 상자(bounding box), $\tilde{\mathbf{b}}_n^k$ 는 n번째 블러링 이미지에서 측정된 객체 경계 상자, $\|\cdot\|_1$ 는 "1-norm"을 의미한다. 이하, 도면을 참조하여, 본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템을 상세히 설명한다.

도 4는 본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템의 제1 실시예를 도시한 것이다. 도 4의 (a)는 선명한 이미지를 입력하였을 때 객체 감지 기술의 감지 성능을 나타낸 사진이다. 도 4의 (b)는 블러링 처리가된 이미지를 입력하였을 때 객체 감지 기술의 감지 성능을 나타낸 사진이다. 도 4의 (c)는 블러링된 이미지를 디블러링 처리한 후 입력하였을 때, 객체 감지 기술의 감지 성능을 나타낸 사진이다. 도 4의 (d)는 본 발명의 일 실시예에 따른 지식 증류 시스템에 블러링된 사진을 입력하였을 때, 객체 감지 기술의 감지 성능을 나타낸 사진이다.

도 4의 (a)를 참조하면, 선명한 이미지가 입력되는 경우에는 사람과 차량을 올바르게 감지하는 것을 알 수 있다. 도 4의 (b)를 참조하면, 흐릿한 이미지가 입력되는 경우 사람과 차량을 올바르게 감지하지 못하는 것을 알 수 있다. 도 4의 (c)를 참조하면, 블러링된 이미지를 디블러링 처리한 후 입력하였을 때, 사람과 차량을 인식하지 못하거나 정확하게 인식하지 못하는 것을 알 수 있다. 즉, 블러링된 이미지를 디블러링하여 입력하는 방법의 경우에는 다수의 이미지를 디블러링하는 작업이 어려운 문제가 있을 뿐만 아니라, 객체 감지 기술의 감지 성능이 저하되는 것을 알 수 있다. 도 4의 (d)를 참조하면, 본 발명의 일 실시예에 따른 시스템으로 학습된 객체 감지 기술의 경우에는 사람과 차량을 정확하게 인식하는 것을 알 수 있다.

- [0113] 도 5는 본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템의 제2 실시예를 도시한 것이다.
- [0114] 도 5에서, Ground truth(GT)는 객체를 정확하게 표시한 정답 사진이다. Original(Sharp)는 선명한 이미지로만 학습된 객체 검출기를 사용하여 객체 감지여부를 나타낸 사진이다. Original(Blurry)는 블러링 처리되어 흐릿한 이미지로만 학습된 객체 검출기를 사용하여 객체 감지여부를 나타낸 사진이다. Proposed method는 본 발명의 일 실시예에 따른 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템이다. 객체 검출기로는 잘 알려진 SSD(Single Shot MultiBox Detector), RetinaNet(Focal Loss for Dense Object Detection) 및 ATSS(Adaptive Training Sample Selection)를 사용하였다. 학습데이터 셋으로는 PASCAL VOC test 2007 데이터 셋을 사용하였다.
- [0115] 도 5를 참조하면, 선명한 이미지로만 학습된 SSD의 경우에는 흐릿한 사진의 객체의 범위를 올바르게 인식하지 못하는 문제가 있다. 흐릿한 이미지로만 학습된 SSD의 경우에는 흐릿한 사진의 일부 객체를 인식하지 못하는 문제가 있다. 그러나 본 발명에 일 실시예에 따른 시스템으로 학습된 SSD의 경우에는 흐릿한 이미지의 객체를 정확하게 인식할 수 있다.
- [0116] 선명한 이미지로만 학습된 RetinaNet의 경우에는 흐릿한 사진을 대부분 인식하지 못하는 문제점이 있다. 구체적으로, 선명한 이미지로만 학습된 RetinaNet의 경우에는 흐릿한 사진에서 차량과 사람을 포함한 총 6개의 객체 중 1개 밖에 인식하지 못하였다. 흐릿한 이미지로만 학습된 RetinaNet의 경우에는 객체가 아닌 것을 인식하는 문제가 있다. 그러나 본 발명에 일 실시예에 따른 시스템으로 학습된 RetinaNet의 경우에는 흐릿한 이미지의 객체를 비교적 정확하게 인식할 수 있다.
- [0117] 선명한 이미지로만 학습된 ATTS의 경우에는 흐릿한 사진에서 객체를 아예 인식하지 못하였다. 흐릿한 이미지로만 학습된 ATTS의 경우에는 흐릿한 사진에서 소형화된 객체를 인식하지 못하는 문제가 있다. 그러나, 본 발명에 일 실시예에 따른 시스템으로 학습된 ATTS의 경우에는 흐릿한 이미지의 객체를 정확하게 인식할 수 있다.
- [0118] 하기 표 1은 학습방법, 학습 데이터셋 및 객체 검출기에 따라 선명한 사진 및 흐릿한 사진에서의 객체 검출 성능을 비교한 표이다. 표 1에서 행은 데이터 셋을 기재하였으며, VOC_{sharp}는 선명한 PASCAL VOC test 2007 데이터 셋을 의미하며, VOC_{blurry}는 블러링 처리된 PASCAL VOC test 2007 데이터 셋을 의미한다. 표 1에서 값은 mAP(mean Average Precision)을 의미하며, 단위는 %이다.

표 1

Training methods	Training data	SSD		RetinaNet		ATSS	
		VOC _{sharp}	VOC _{blurry}	VOC _{sharp}	VOC _{blurry}	VOC _{sharp}	VOC _{blurry}
Original	VOC _{sharp}	77.33	34.2	79.9	34.0	79.1	31.2
	VOC _{blurry}	73.4	66.5	71.2	62.1	67.0	58.0
Proposed	VOC _{pair}	78.0	69.0	79.9	67.1	77.0	64.2

- [0119]
- [0120] 상기 표 1을 참조하면, 제안한 시스템으로 객체 검출기를 학습할 경우 흐릿한 이미지에서의 객체 검출 성능이 우수함을 알 수 있다. ATSS가 객체 검출기일 때, 제안한 시스템으로 ATSS를 학습할 경우보다 선명한 사진으로만 학습한 경우 미세하게 성능이 우수함을 알 수 있다. 그러나, 흐릿한 사진으로만 학습된 ATSS의 경우에는 31.2인 반면, 제안한 시스템으로 ATSS를 학습할 경우에는 64.2로 도출되됨에 따라 흐릿한 사진의 객체 탐지 성능이 월등히 우수함을 알 수 있다.
- [0122] 도 6은 본 발명에 따른 교사 모델 및 학생 모델의 학습을 위한 지식증류 기반 시스템의 제3 실시예를 도시한 것이다. 도 6에서, 학습데이터 셋으로는 COCO val2017 데이터 셋을 사용하였다. 그 외는 상기 제2 실시예에서와 동일한 조건으로 실험을 진행하였다.
- [0123] 도 6을 참조하면, 선명한 이미지로만 학습된 SSD의 경우에는 흐릿한 사진의 객체의 범위를 올바르게 인식하지 못하는 문제가 있다. 흐릿한 이미지로만 학습된 SSD의 경우에는 흐릿한 사진의 일부 객체를 인식하지 못하는 문제가 있다. 그러나 본 발명에 일 실시예에 따른 시스템으로 학습된 SSD의 경우에는 흐릿한 이미지의 객체를 정확하게 인식할 수 있다.

[0124] 선명한 이미지로만 또는 흐릿한 이미지로만 학습된 RetinaNet의 경우에는 객체가 아닌 것을 인식하는 문제가 있다. 그러나 본 발명에 일 실시예에 따른 시스템으로 학습된 RetinaNet의 경우에는 흐릿한 이미지의 객체를 정확하게 인식할 수 있다.

[0125] 선명한 이미지로만 또는 흐릿한 이미지로만 학습된 ATSS의 경우에는 흐릿한 사진에서 객체를 아예 인식하지 못하였다. 그러나, 본 발명에 일 실시예에 따른 시스템으로 학습된 ATSS의 경우에는 흐릿한 이미지의 객체를 정확하게 인식할 수 있다.

[0126] 하기 표 2는 학습방법, 학습 데이터셋 및 객체 검출기에 따라 선명한 사진 및 흐릿한 사진에서의 객체 검출 성능을 비교한 표이다. 표 1에서 COCO_{sharp}는 선명한 COCOval2017 데이터 셋을 의미하며, COCO_{blurry}은 블러링 처리된 COCOval2017데이터 셋을 의미한다.

표 2

Training methods	Training data	SSD		RetinaNet		ATSS	
		COCO _{sharp}	COCO _{blurry}	COCO _{sharp}	COCO _{blurry}	COCO _{sharp}	COCO _{blurry}
Original	COCO _{sharp}	25.5	11.7	35.4	14.6	39.4	15.1
	COCO _{blurry}	21.4	18.0	29.2	23.3	30.7	24.2
Proposed	COCO _{pair}	26.9	21.5	36.4	27.1	36.5	27.3

[0127]

[0128] 상기 표 2를 참조하면, 제안한 시스템으로 객체 검출기를 학습할 경우 흐릿한 이미지에서의 객체 검출 성능이 우수함을 알 수 있다. ATSS가 객체 검출기일 때, 제안한 시스템으로 ATSS를 학습할 경우보다 선명한 사진으로만 학습한 경우 미세하게 성능이 우수함을 알 수 있다. 그러나, 흐릿한 사진으로만 학습된 ATSS의 경우에는 15.1인 반면, 제안한 시스템으로 ATSS를 학습할 경우에는 27.3으로 도출되됨에 따라 흐릿한 사진의 객체 탐지 성능이 월등히 우수함을 알 수 있다.

[0129] 이상 본 발명의 바람직한 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 특허 청구의 범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

부호의 설명

[0131]

1000: 시스템

100: 교사네트워크

110: 제1 베이스 서버넷 및 제1 피쳐피라미드 서버넷

120: 제1 디텍션 서버넷

200: 학생 네트워크

210: 제2 베이스 서버넷 및 제2 피쳐피라미드 서버넷

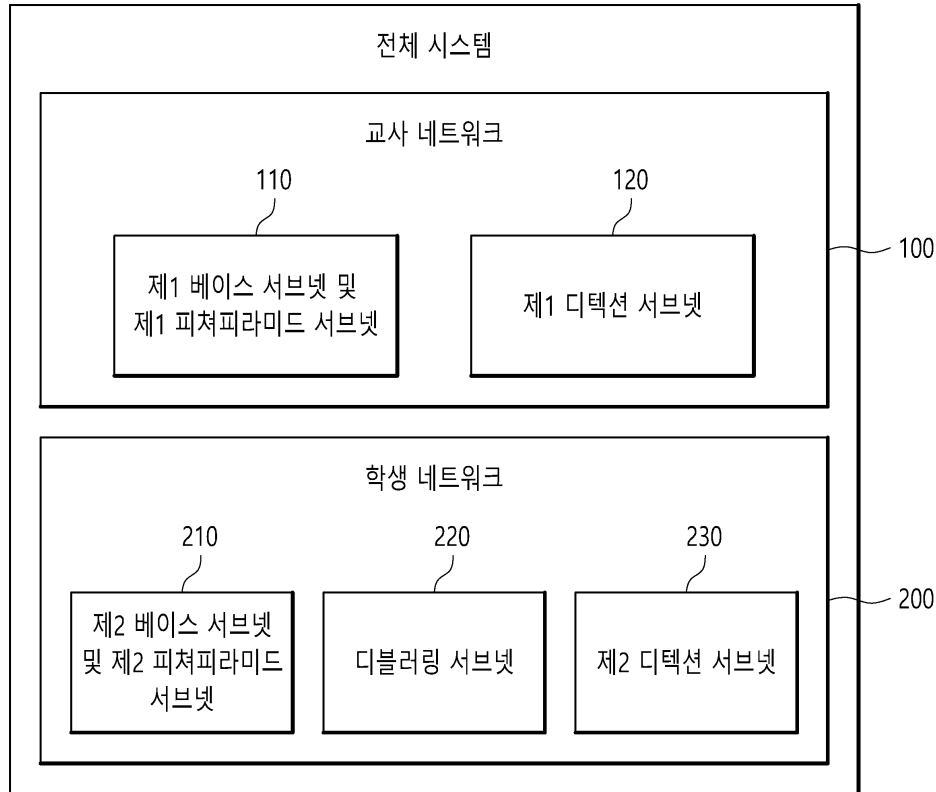
220: 디블러링 서버넷

230: 제2 디텍션 서버넷

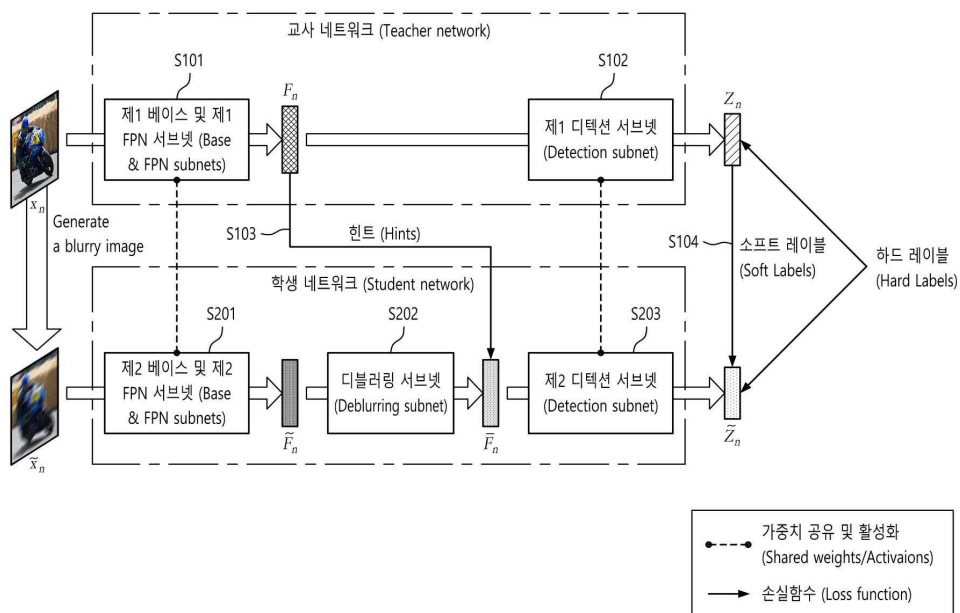
도면

도면1

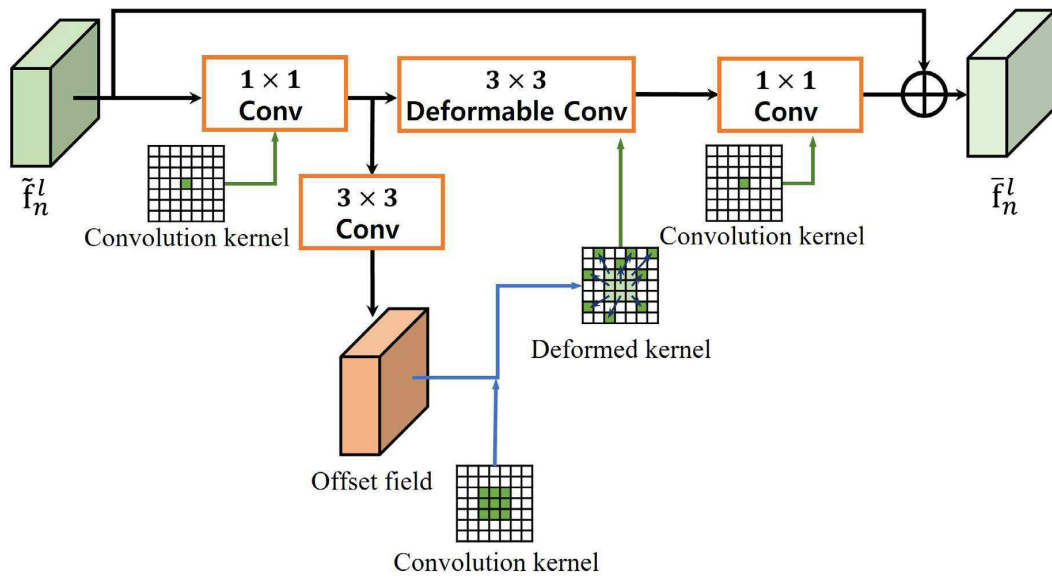
1000



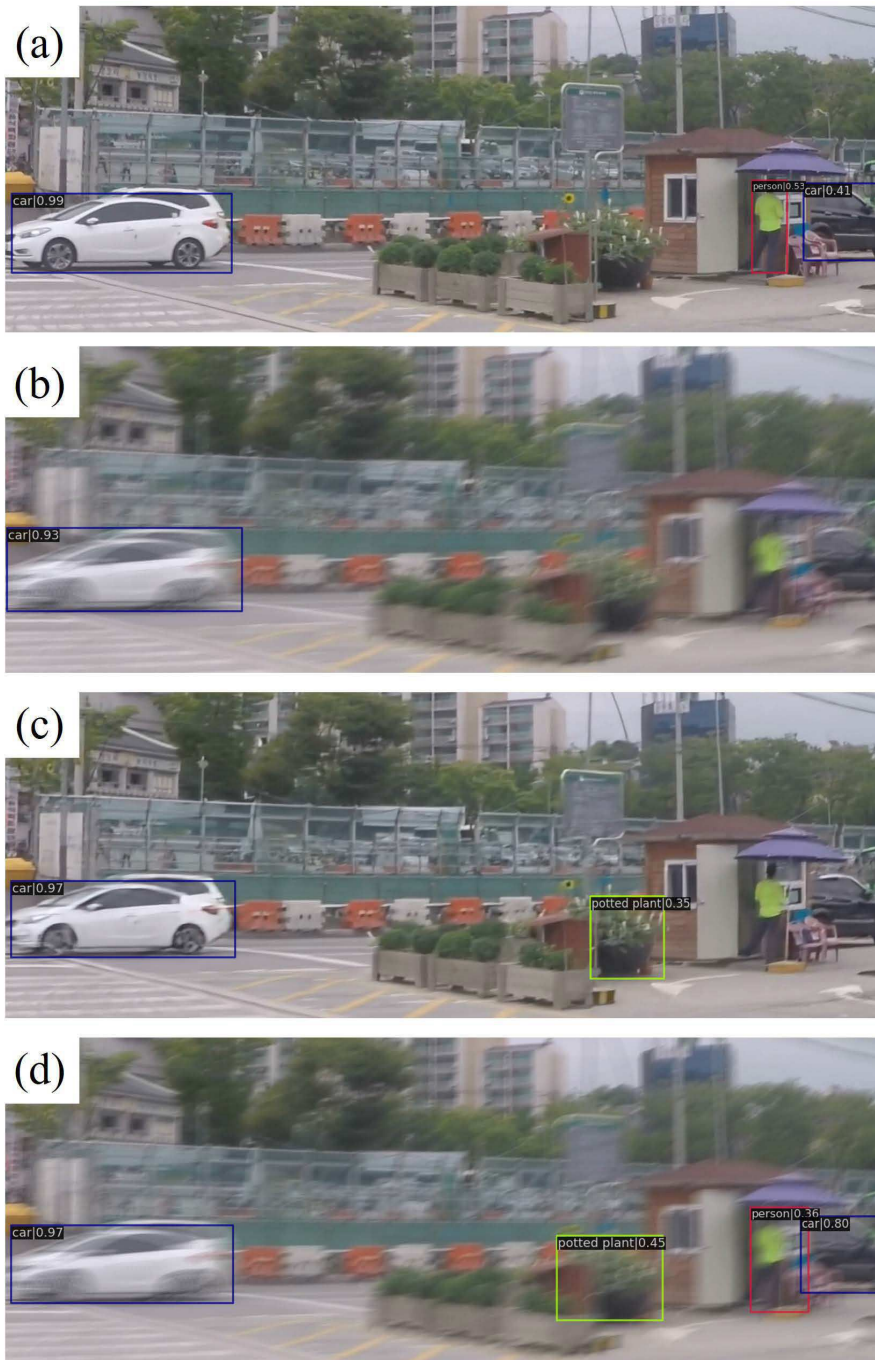
도면2



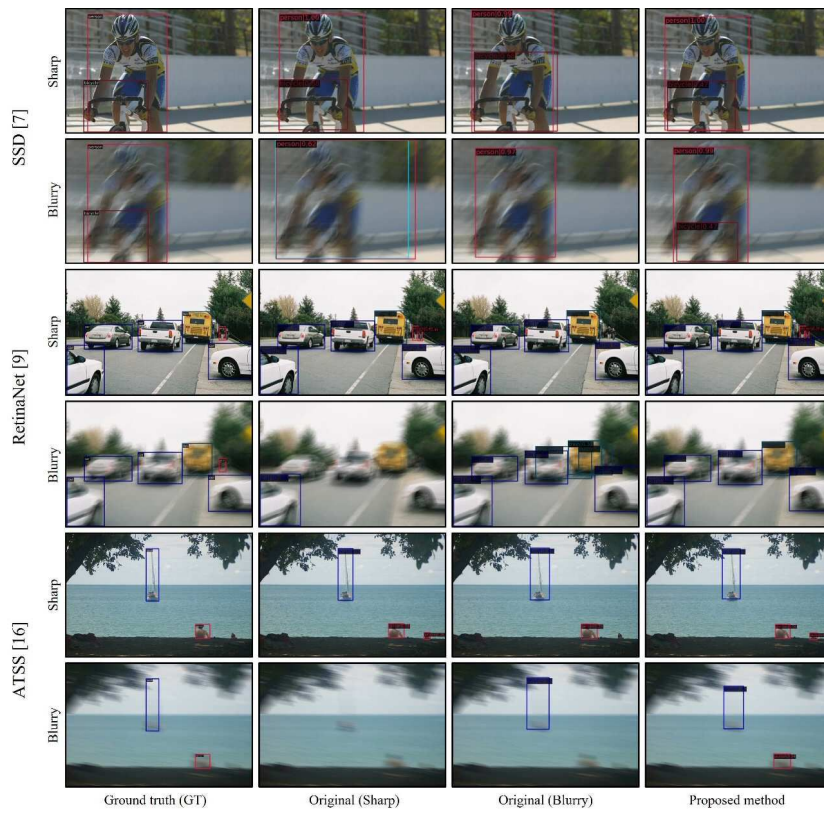
도면3



도면4



도면5



도면6

