



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년11월04일
(11) 등록번호 10-2464359
(24) 등록일자 2022년11월02일

- (51) 국제특허분류(Int. Cl.)
G06T 3/00 (2019.01) G06K 9/00 (2022.01)
G06N 20/00 (2019.01) G06T 5/50 (2006.01)
- (52) CPC특허분류
G06T 3/00 (2019.01)
G06N 20/00 (2021.08)
- (21) 출원번호 10-2020-0108051
(22) 출원일자 2020년08월26일
심사청구일자 2020년08월26일
- (65) 공개번호 10-2022-0026917
(43) 공개일자 2022년03월07일
- (56) 선행기술조사문헌
Yunsheng Li et al., 'Bidirectional Learning for Domain Adaptation of Semantic Segmentation', CVPR, pp. 6936-6945, June 2019.*
Hongyi Zhang et al., 'mixup: BEYOND EMPIRICAL RISK MINIMIZATION', arXiv:1710.09412v2 [cs.LG] 27 Apr 2018.*
Robert Geirhos et al., 'IMAGENET-TRAINED CNNs ARE BIASED TOWARDS TEXTURE: INCREASING SHAPE BIAS IMPROVES ACCURACY AND ROBUSTNESS', arXiv:1811.12231v2 [cs.CV] 14 Jan. 2019.*
*는 심사관에 의하여 인용된 문헌

- (73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
- (72) 발명자
변혜란
서울특별시 서대문구 연세로 50, 제4공학관 821호(신촌동, 연세대학교)
- 김명진
서울특별시 서대문구 연세로 50, 제4공학관 821호(신촌동, 연세대학교)
- 전석규
서울특별시 서대문구 연세로 50, 제4공학관 821호(신촌동, 연세대학교)
- (74) 대리인
특허법인우인

전체 청구항 수 : 총 8 항

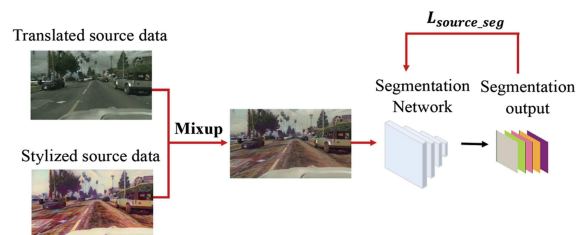
심사관 : 정성윤

(54) 발명의 명칭 **합성 도메인에서 실제 도메인으로 적응하기 위한 이미지 변환 장치 및 방법**

(57) 요약

본 실시예들은 합성 도메인에서 학습한 모델이 실제 도메인에 정확하게 작동하도록 만드는 도메인 적응에 관한 것으로 상이한 형식의 이미지를 혼합하여 시맨틱 분할 모델을 학습시켜, 실제 도메인에서도 정확도를 향상시키는 이미지 변환 방법 및 장치를 제공한다.

대표도



(52) CPC특허분류

G06T 5/50 (2013.01)

G06V 20/41 (2022.01)

G06T 2207/20081 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711112608

과제번호 2019R1A2C2003760

부처명 과학기술정보통신부

과제관리(전문)기관명 한국연구재단

연구사업명 중견연구자지원사업

연구과제명 특성 정보 자동 생성을 통한 처음 보는 복합카테고리의 이미지와 비디오 생성 및 인
식을 위한 제로샷 학습 기술 연구(2/3)

기 여 율 1/1

과제수행기관명 연세대학교

연구기간 2020.03.01 ~ 2021.02.28

공지예외적용 : 있음

명세서

청구범위

청구항 1

컴퓨팅 디바이스에 의한 이미지 변환 방법에 있어서,

제1 형식의 이미지와 제2 형식의 이미지 간의 혼합비를 설정하는 단계;

상기 제1 형식의 이미지와 상기 제2 형식의 이미지를 혼합하여 혼합 이미지를 생성하는 단계;

상기 혼합 이미지를 이용하여 시맨틱 분할 모델을 학습하는 제1 학습 단계;

실제 도메인(real domain)에 해당하는 레이블이 없는 대상 이미지에 대해서 상기 시맨틱 분할 모델로부터 기준치를 만족하는 슈도 레이블을 추출하는 단계;

상기 시맨틱 분할 모델에 상기 대상 이미지와 상기 대상 이미지에 대한 픽셀 레벨의 검증 자료를 이용하여 정의된 제2 손실 함수를 적용하며, 상기 슈도 레이블에 따라 자가 학습하는 제2 학습 단계; 및

상기 슈도 레이블에 따라 자가 학습된 시맨틱 분할 모델을 통해 상기 실제 도메인에 해당하는 상기 대상 이미지를 변환하여 변환된 이미지를 출력하는 단계를 포함하며,

상기 시맨틱 분할 모델은 상기 혼합 이미지와 상기 혼합 이미지에 대한 픽셀 레벨의 검증 자료를 이용하여 정의된 제1 손실 함수를 갖고, 상기 시맨틱 분할 모델은 상기 혼합 이미지의 픽셀에 클래스 레이블을 매칭하는 것을 특징으로 하는 이미지 변환 방법.

청구항 2

제1항에 있어서,

상기 제1 형식의 이미지는 합성 도메인(synthetic domain)에 해당하며 레이블이 매칭된 이미지이고,

상기 제2 형식의 이미지는 상기 합성 도메인에 해당하며 레이블이 매칭되고 질감 다양화를 적용한 이미지이고,

상기 제1 형식의 이미지와 상기 제2 형식의 이미지는 픽셀의 시맨틱 정보가 동일한 시맨틱 레이아웃을 갖는 것을 특징으로 하는 이미지 변환 방법.

청구항 3

제2항에 있어서,

상기 제1 형식의 이미지와 상기 제2 형식의 이미지 간의 혼합비에 따라 픽셀 레벨에서 세부 정보와 질감 표현이 결정되고,

상기 제1 형식의 이미지의 비중을 0.5 내지 0.75로 설정하고,

상기 제2 형식의 이미지의 비중을 0.25 내지 0.5로 설정하는 것을 특징으로 하는 이미지 변환 방법.

청구항 4

삭제

청구항 5

제1항에 있어서,

상기 시맨틱 분할 모델은 상기 제1 손실 함수를 산출할 때 클래스에 대해서 클래스 가중치를 적용하는 것을 특징으로 하는 이미지 변환 방법.

청구항 6

삭제

청구항 7

삭제

청구항 8

하나 이상의 프로세서 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 이미지 변환 장치에 있어서,

상기 프로세서는,

제1 형식의 이미지와 제2 형식의 이미지 간의 혼합비를 설정하고,

상기 제1 형식의 이미지와 상기 제2 형식의 이미지를 혼합하여 혼합 이미지를 생성하고,

상기 혼합 이미지를 이용하여 시맨틱 분할 모델을 학습하며,

실제 도메인(real domain)에 해당하는 레이블이 없는 대상 이미지에 대해서 상기 시맨틱 분할 모델로부터 기준치를 만족하는 슈도 레이블을 추출하고,

상기 시맨틱 분할 모델에 상기 대상 이미지와 상기 대상 이미지에 대한 픽셀 레벨의 검증 자료를 이용하여 정의된 제2 손실 함수를 적용하며, 상기 슈도 레이블에 따라 자가 학습하며,

상기 슈도 레이블에 따라 자가 학습된 시맨틱 분할 모델을 통해 상기 실제 도메인에 해당하는 상기 대상 이미지를 변환하여 변환된 이미지를 출력하며,

상기 시맨틱 분할 모델은 상기 혼합 이미지와 상기 혼합 이미지에 대한 픽셀 레벨의 검증 자료를 이용하여 정의된 제1 손실 함수를 갖고, 상기 시맨틱 분할 모델은 상기 혼합 이미지의 픽셀에 클래스 레이블을 매칭하는 것을 특징으로 하는 이미지 변환 장치.

청구항 9

제8항에 있어서,

상기 제1 형식의 이미지는 합성 도메인(synthetic domain)에 해당하며 레이블이 매칭된 이미지이고,

상기 제2 형식의 이미지는 상기 합성 도메인에 해당하며 레이블이 매칭되고 질감 다양화를 적용한 이미지이고,

상기 제1 형식의 이미지와 상기 제2 형식의 이미지는 픽셀의 시맨틱 정보가 동일한 시맨틱 레이아웃을 갖는 것을 특징으로 하는 이미지 변환 장치.

청구항 10

제9항에 있어서,

상기 제1 형식의 이미지와 상기 제2 형식의 이미지 간의 혼합비에 따라 픽셀 레벨에서 세부 정보와 질감 표현이 결정되고,

상기 제1 형식의 이미지의 비중을 0.5 내지 0.75로 설정하고,

상기 제2 형식의 이미지의 비중을 0.25 내지 0.5로 설정하는 것을 특징으로 하는 이미지 변환 장치.

청구항 11

삭제

청구항 12

제8항에 있어서,

상기 시맨틱 분할 모델은 상기 제1 손실 함수를 산출할 때 클래스에 대해서 클래스 가중치를 적용하는 것을 특징으로 하는 이미지 변환 장치.

청구항 13

삭제

청구항 14

삭제

발명의 설명

기술 분야

[0001] 본 발명이 속하는 기술 분야는 시맨틱 분할 및 도메인 적응을 수행하는 이미지 변환 장치 및 방법에 관한 것이다.

배경 기술

[0002] 이 부분에 기술된 내용은 단순히 본 실시예에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.

[0003] 시맨틱 분할(Semantic segmentation)은 입력 이미지의 각 픽셀에 클래스 레이블을 매칭하는 컴퓨터 비전 작업의 하나이다. 마스크라고 불리는 예측 결과는 자율 주행, 의료 영상 분석 등의 다양한 분야에 적용될 수 있다.

[0004] 도메인 적응(Domain adaptation)은 어떤 도메인에서 학습한 모델을 다른 도메인에 적용하는 것을 의미한다. 예컨대 레이블이 매칭된 합성 데이터에서 학습한 모델을 실제 데이터에 적용할 수 있다. 도메인 적응의 목적은 훈련 데이터와 테스트 데이터 분포 간의 갭 차이에서 분별력있는 예측 변수를 학습하는 것이다.

선행기술문헌

특허문헌

[0005] (특허문헌 0001) 한국공개특허공보 제10-2020-0065433호 (2020.06.09.)

(특허문헌 0002) 한국등록특허공보 제10-0924689호 (2009.10.27.)

발명의 내용

해결하려는 과제

[0006] 본 발명의 실시예들은 합성 도메인에서 학습한 모델이 실제 도메인에 정확하게 작동하도록 만드는 도메인 적응에 관한 것으로 상이한 형식의 이미지를 혼합하여 시맨틱 분할 모델을 학습시켜, 실제 도메인에서도 정확도를 향상시키는 데 주된 목적이 있다.

[0007] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다.

과제의 해결 수단

[0008] 본 실시예의 일 측면에 의하면, 컴퓨팅 디바이스에 의한 이미지 변환 방법에 있어서, 제1 형식의 이미지와 제2 형식의 이미지 간의 혼합비를 설정하는 단계, 상기 제1 형식의 이미지와 상기 제2 형식의 이미지를 혼합하여 혼합 이미지를 생성하는 단계, 및 상기 혼합 이미지를 이용하여 시맨틱 분할 모델을 학습하는 제1 학습 단계를 포함하는 이미지 변환 방법을 제공한다.

[0009] 상기 제1 형식의 이미지는 합성 도메인(synthetic domain)에 해당하며 레이블이 매칭된 이미지이고, 상기 제2 형식의 이미지는 상기 합성 도메인에 해당하며 레이블이 매칭되고 질감 다양화를 적용한 이미지이고, 상기 제1 형식의 이미지와 상기 제2 형식의 이미지는 픽셀의 시맨틱 정보가 동일한 시맨틱 레이어아웃을 가질 수 있다.

[0010] 상기 제1 형식의 이미지와 상기 제2 형식의 이미지 간의 혼합비에 따라 픽셀 레벨에서 세부 정보와 질감 표현이 결정되고, 상기 제1 형식의 이미지의 비중을 0.5 내지 0.75로 설정하고, 상기 제2 형식의 이미지의 비중을 0.25

내지 0.5로 설정할 수 있다.

- [0011] 상기 시맨틱 분할 모델은 상기 혼합 이미지와 상기 혼합 이미지에 대한 픽셀 레벨의 검증 자료를 이용하여 정의된 제1 손실 함수를 갖는 시맨틱 분할 모델이며, 상기 시맨틱 분할 모델은 상기 혼합 이미지의 픽셀에 클래스 레이블을 매칭할 수 있다.
- [0012] 상기 시맨틱 분할 모델은 상기 제1 손실 함수를 산출할 때 클래스에 대해서 클래스 가중치를 적용할 수 있다.
- [0013] 상기 이미지 변환 방법은 실제 도메인(real domain)에 해당하는 레이블이 없는 대상 이미지에 대해서 상기 시맨틱 분할 모델로부터 기준치를 만족하는 슈도 레이블을 추출하는 단계 및 상기 시맨틱 분할 모델에 상기 대상 이미지와 상기 대상 이미지에 대한 픽셀 레벨의 검증 자료를 이용하여 정의된 제2 손실 함수를 적용하며, 상기 슈도 레이블에 따라 자가 학습하는 제2 학습 단계를 포함할 수 있다.
- [0014] 상기 이미지 변환 방법은 상기 슈도 레이블에 따라 자가 학습된 시맨틱 분할 모델을 통해 실제 도메인에 해당하는 대상 이미지를 변환하여 변환된 이미지를 출력하는 단계를 포함할 수 있다.
- [0015] 본 실시예의 다른 측면에 의하면, 하나 이상의 프로세서 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 이미지 변환 장치에 있어서, 상기 프로세서는, 제1 형식의 이미지와 제2 형식의 이미지 간의 혼합비를 설정하고, 상기 제1 형식의 이미지와 상기 제2 형식의 이미지를 혼합하여 혼합 이미지를 생성하고, 상기 혼합 이미지를 이용하여 시맨틱 분할 모델을 학습하는 것을 특징으로 하는 이미지 변환 장치를 제공한다.

발명의 효과

- [0016] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 합성 도메인에서 학습한 모델이 실제 도메인에 정확하게 작동하도록 만드는 도메인 적응에 관한 것으로 상이한 형식의 이미지를 혼합하여 시맨틱 분할 모델을 학습시켜, 실제 도메인에서도 정확도를 향상시킬 수 있는 효과가 있다.
- [0017] 여기에서 명시적으로 언급되지 않은 효과라 하더라도, 본 발명의 기술적 특징에 의해 기대되는 이하의 명세서에서 기재된 효과 및 그 잠정적인 효과는 본 발명의 명세서에 기재된 것과 같이 취급된다.

도면의 간단한 설명

- [0018] 도 1은 본 발명의 일 실시예에 따른 이미지 변환 장치를 예시한 블록도이다.
- 도 2 및 도 3은 본 발명의 일 실시예에 따른 이미지 변환 장치가 혼합한 이미지를 예시한 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 이미지 변환 장치의 시맨틱 분할 모델을 예시한 도면이다.
- 도 5는 본 발명의 다른 실시예에 따른 이미지 변환 방법을 예시한 흐름도이다.
- 도 6 및 도 7은 본 발명의 실시예들에 따른 시뮬레이션 결과를 예시한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0019] 이하, 본 발명을 설명함에 있어서 관련된 공지기능에 대하여 이 분야의 기술자에게 자명한 사항으로서 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하고, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다.
- [0020] 도 1은 본 발명의 일 실시예에 따른 이미지 변환 장치를 예시한 블록도이다.
- [0021] 이미지 변환 장치(110)는 적어도 하나의 프로세서(120), 컴퓨터 판독 가능한 저장매체(130) 및 통신 버스(170)를 포함한다.
- [0022] 프로세서(120)는 이미지 변환 장치(110)로 동작하도록 제어할 수 있다. 예컨대, 프로세서(120)는 컴퓨터 판독 가능한 저장 매체(130)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 컴퓨터 실행 가능 명령어는 프로세서(120)에 의해 실행되는 경우 이미지 변환 장치(110)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.
- [0023] 컴퓨터 판독 가능한 저장 매체(130)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능한 저장 매체(130)에 저장된 프로그램(140)

은 프로세서(120)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독한 가능 저장 매체(130)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 이미지 변환 장치(110)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적절한 조합일 수 있다.

[0024] 통신 버스(170)는 프로세서(120), 컴퓨터 판독 가능한 저장 매체(140)를 포함하여 이미지 변환 장치(110)의 다른 다양한 컴포넌트들을 상호 연결한다.

[0025] 이미지 변환 장치(110)는 또한 하나 이상의 입출력 장치(24)를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(150) 및 하나 이상의 통신 인터페이스(160)를 포함할 수 있다. 입출력 인터페이스(150) 및 통신 인터페이스(160)는 통신 버스(170)에 연결된다. 입출력 장치(미도시)는 입출력 인터페이스(150)를 통해 이미지 변환 장치(110)의 다른 컴포넌트들에 연결될 수 있다.

[0026] 이미지 변환 장치(110)는 합성 도메인에서 학습한 모델이 실제 도메인에서 잘 작동할 수 있도록 입력 데이터를 변형한다. 이미지 변환 장치(110)는 두 가지 방식에서 사용하는 입력 이미지를 섞는 방식을 사용한다. 하나의 방식은 이미지 변환(Image Translation) 알고리즘을 사용해 두 도메인의 입력 이미지 간의 스타일 차이를 줄이는 방식이고, 다른 하나의 방식은 합성 도메인에 스타일 변환(Style Transfer) 알고리즘을 이용해 질감 다양화(Texture Randomization)를 통해 합성 도메인의 질감에 과적합(Overfit)되지 않게 만드는 질감 불변 표현(texture-invariant representation)을 학습하는 방식이다. 두 방식의 장점을 모두 얻을 수 있어서 개선된 성능을 확보할 수 있다.

[0027] 도 2 및 도 3은 본 발명의 일 실시예에 따른 이미지 변환 장치가 혼합한 이미지를 예시한 도면이다.

[0028] 이미지 변환 장치는 제1 형식의 이미지와 제2 형식의 이미지 간의 혼합비를 설정한다. 이미지 변환 장치는 제1 형식의 이미지와 제2 형식의 이미지를 혼합하여 혼합 이미지를 생성한다.

[0029] 제1 형식의 이미지는 합성 도메인(synthetic domain)에 해당하며 레이블이 매칭된 이미지이고, 제2 형식의 이미지는 합성 도메인에 해당하며 레이블이 매칭되고 질감 다양화를 적용한 이미지이다. 제1 형식의 이미지와 제2 형식의 이미지는 픽셀의 시맨틱 정보가 동일한 시맨틱 레이아웃을 갖는다. 두 이미지 간에 동일한 시맨틱 레이아웃을 갖지 않으면 시맨틱 분할 과정에서 픽셀 레벨의 시맨틱 정보가 손상되므로, 혼합된 이미지를 사용할 수 없다.

[0030] 레이블에 따른 이미지 혼합은 수학적 식 1과 같이 표현된다.

수학적 식 1

$$\tilde{x} = \lambda x_i + (1 - \lambda)x_j$$

$$\tilde{y} = \lambda y_i + (1 - \lambda)y_j$$

[0031]

[0032] x_i 와 x_j 는 두 유형의 입력 이미지의 픽셀이고, y_i 와 y_j 는 픽셀에 대응하는 레이블이다. λ 는 혼합비이다.

[0033] 제1 형식의 이미지와 제2 형식의 이미지 간의 혼합비에 따라 픽셀 레벨에서 세부 정보와 질감 표현의 비중이 결정된다. 제1 형식의 이미지의 비중을 0.5 내지 0.75로 설정하고, 제2 형식의 이미지의 비중을 0.25 내지 0.5로 설정할 수 있다.

[0034] 표 1은 SYNTHIA 데이터에서 Cityscapes 데이터로 도메인 적응을 수행한 결과이다.

표 1

SYNTHIA → Cityscapes				
Source Type	road	side.	t-light	t-sign
Original(0.75) + Stylized(0.25)	69.2	30.6	11.9	19.7
Original(0.5) + Stylized(0.5)	71.8	27.9	10.3	17.4
Original(0.25) + Stylized(0.75)	75.4	30.7	4.8	14.
Translated(0.75) + Stylized(0.25)	80.7	33.6	10.0	15.6
Translated(0.5) + Stylized(0.5)	81.5	37.3	8.8	16.0
Translated(0.25) + Stylized(0.75)	80.7	35.9	8.1	14.0

[0035]

[0036]

제1 학습 단계에서 질감에 영향을 크게 받는 큰 클래스인 큰 질감 민감성 클래스(예컨대, road, sidewalk 등)와 작은 질감 민감성 클래스(예컨대, traffic-light, traffic-sign 등)가 혼합비에 따라 상이한 변환 결과를 나타낼 수 있다.

[0037]

표 2은 GTA5 데이터에서 Cityscapes 데이터로 SYNTHIA 데이터에서 Cityscapes 데이터로 도메인 적응을 수행한 결과이다.

표 2

Method	GTA5 → Cityscapes	SYNTHIA → Cityscapes
Original(O)	36.6	38.6
Translated(T)	41.0	40.2
Stylized(S)	42.5	37.7
O(0.75) + S(0.25)	43.7	42.4
O(0.5) + S(0.5)	44.1	44.0
O(0.25) + S(0.75)	43.9	43.5
T(0.75) + S(0.25)	45.2	46.9
T(0.5) + S(0.5)	44.9	46.9
T(0.25) + S(0.75)	44.7	45.3
O(0.5) + S(0.5) + Weighted Loss	44.9	46.9
T(0.75) + S(0.25) + Weighted Loss	46.5	48.0
T(0.5) + S(0.5) + Weighted Loss	46.5	47.1

[0038]

[0039]

제1 형식의 이미지와 제2 형식의 이미지 간의 혼합비를 조절하여 픽셀 레벨에서 세부 정보를 살리면서 질감 표현을 적용할 수 있다. 제1 형식의 이미지의 비중을 0.5 내지 0.75로 설정하고, 제2 형식의 이미지의 비중을 0.25 내지 0.5로 설정하는 것이 바람직하다.

[0040]

도 4는 본 발명의 일 실시예에 따른 이미지 변환 장치의 시맨틱 분할 모델을 예시한 도면이다.

[0041]

시맨틱 분할 모델은 특징 추출 모델을 포함할 수 있으며, 특징 추출 모델은 뉴럴 네트워크로 구현될 수 있다. 특징 추출 모델은 다수의 레이어가 네트워크로 연결되며 하든 레이어를 포함한다. 레이어는 파라미터를 포함할 수 있고, 레이어의 파라미터는 학습가능한 필터 집합을 포함한다. 파라미터는 노드 간의 가중치 및/또는 바이어스를 포함한다.

[0042]

시맨틱 분할 모델은 혼합 이미지를 이용하여 학습하는 제1 학습 단계에서 제1 손실 함수를 최소화하도록 학습하여 최적화된 파라미터를 찾는다.

[0043]

시맨틱 분할 모델은 혼합 이미지와 혼합 이미지에 대한 픽셀 레벨의 검증 자료를 이용하여 정의된 제1 손실 함수

수를 갖는 시맨틱 분할 모델이며, 시맨틱 분할 모델은 혼합 이미지의 픽셀에 클래스 레이블을 매칭한다.

제1 손실 함수는 수학식 2와 같이 표현된다.

수학식 2

$$L_{seg}(I_s) = - \sum_{h,w} \sum_{c=1}^C y_s^{h,w,c} \log P_s^{(h,w,c)}$$

P_s 는 혼합 이미지이고, $y_s^{h,w,c}$ 는 혼합 이미지에 대한 픽셀 레벨의 검증 자료이다. H는 이미지의 높이이고, W는 이미지의 폭이고, C는 클래스의 개수이다.

시맨틱 분할 모델은 제1 손실 함수를 산출할 때 클래스에 대해서 클래스 가중치를 적용할 수 있다. 제1 손실 함수를 계산할 때 적은 클래스에 대해 큰 손실을 부여해 클래스 불균형을 해결한다. 클래스 가중치는 수학식 3과 같이 표현된다.

수학식 3

$$w_i = \left(\frac{\# \text{ of pixels of all classes}}{\# \text{ of pixels of class } i} \right)^{\frac{1}{4}}$$

클래스 가중치는 전체 클래스의 픽셀 개수를 해당 클래스의 픽셀 개수를 나눈 값을 이용한다.

이미지 변환 장치는 합성 이미지에 대한 제1 학습 단계 이후에 실제 이미지에 대한 제2 학습 단계를 수행한다.

이미지 변환 장치는 실제 도메인(real domain)에 해당하는 레이블이 없는 대상 이미지에 대해서 시맨틱 분할 모델로부터 기준치를 만족하는 슈도 레이블을 추출한다. 예컨대, 기준치 0.9를 초과하는 예측값을 슈도 레이블로 설정할 수 있다.

이미지 변환 장치는 시맨틱 분할 모델에 대상 이미지와 대상 이미지에 대한 픽셀 레벨의 검증 자료를 이용하여 정의된 제2 손실 함수를 적용하며, 슈도 레이블에 따라 자가 학습하는 제2 학습 단계를 수행할 수 있다.

시맨틱 분할 모델은 제2 학습 단계에서 제2 손실 함수를 최소화하도록 학습하여 최적화된 파라미터를 찾는다.

제2 손실 함수는 수학식 4과 같이 표현된다.

수학식 4

$$L_{ST}(I_t) = - \sum_{h,w} \mathbb{1}_{pseudo} \sum_{c=1}^C \hat{y}_t^{h,w,c} \log P_t^{(h,w,c)}$$

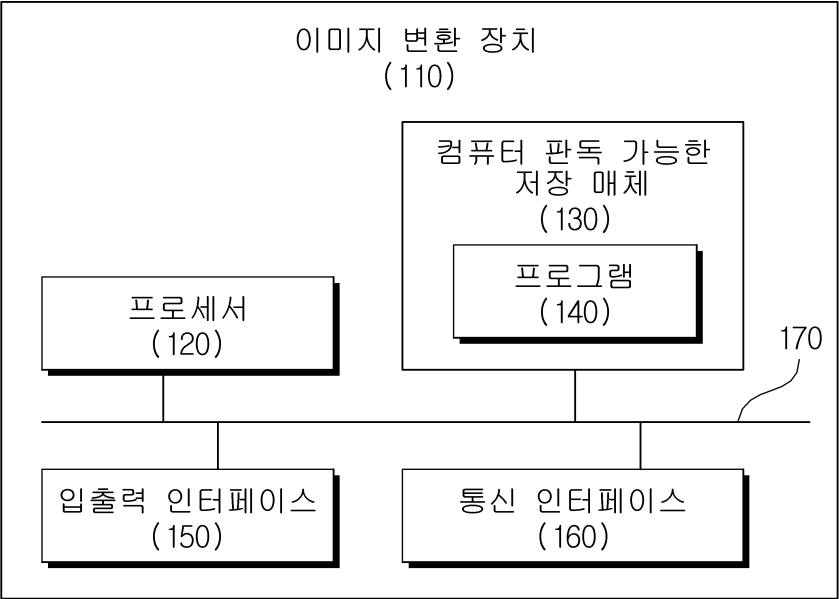
P_t 는 대상 이미지이고, $\hat{y}_t^{h,w,c}$ 는 대상 이미지에 대한 픽셀 레벨의 검증 자료이다. H는 이미지의 높이이고, W는 이미지의 폭이고, C는 클래스의 개수이다. $\mathbb{1}_{pseudo}$ 는 대상 훈련 이미지의 각 픽셀에서 슈도 레이블의

가능여부에 관한 지표이다.

- [0057] 이미지 변환 장치는 슈도 레이블에 따라 자가 학습된 시맨틱 분할 모델을 통해 실제 도메인에 해당하는 대상 이미지를 변환하여 변환된 이미지를 출력한다.
- [0058] 도 5는 본 발명의 다른 실시예에 따른 이미지 변환 방법을 예시한 흐름도이다. 이미지 변환 방법은 컴퓨팅 디바이스에 의하여 수행될 수 있으며, 이미지 변환 장치와 동일한 방식으로 동작한다.
- [0059] 단계 S210에서 프로세서는 제1 형식의 이미지와 제2 형식의 이미지 간의 혼합비를 설정한다.
- [0060] 단계 S220에서 프로세서는 제1 형식의 이미지와 제2 형식의 이미지를 혼합하여 혼합 이미지를 생성한다.
- [0061] 단계 S230에서 프로세서는 혼합 이미지를 이용하여 시맨틱 변환 모델을 학습한다.
- [0062] 단계 S240에서 프로세서는 실제 도메인에 해당하는 레이블이 없는 대상 이미지에 대해서 시맨틱 분할 모델로부터 기준치를 만족하는 슈도 레이블을 추출한다.
- [0063] 단계 S250에서 프로세서는 시맨틱 분할 모델에 대상 이미지와 대상 이미지에 대한 픽셀 레벨의 검증 자료를 이용하여 정의된 제2 손실 함수를 적용하며, 슈도 레이블에 따라 자가 학습한다.
- [0064] 단계 S260에서 프로세서는 슈도 레이블에 따라 자가 학습된 시맨틱 분할 모델을 통해 실제 도메인에 해당하는 대상 이미지를 변환하여 변환된 이미지를 출력한다.
- [0065] 도 6 및 도 7은 본 발명의 실시예들에 따른 시뮬레이션 결과를 예시한 도면이다.
- [0066] 도 6에 도시된 바와 같이 본 실시예에 따른 이미지 변환 장치는 노이즈를 갖는 입력 이미지에 대해서도 강인하게 시맨틱 분할을 수행하는 것을 쉽게 파악할 수 있다.
- [0067] 도 7에 도시된 바와 같이 본 실시예에 따른 이미지 변환 장치는 검증 자료(Ground truth)에 가장 유사하게 시맨틱 분할을 수행하는 것을 쉽게 파악할 수 있다.
- [0068] 이미지 변환 장치는 하드웨어, 펌웨어, 소프트웨어 또는 이들의 조합에 의해 로직회로 내에서 구현될 수 있고, 범용 또는 특정 목적 컴퓨터를 이용하여 구현될 수도 있다. 장치는 고정배선형(Hardwired) 기기, 필드 프로그램 가능한 게이트 어레이(Field Programmable Gate Array, FPGA), 주문형 반도체(Application Specific Integrated Circuit, ASIC) 등을 이용하여 구현될 수 있다. 또한, 장치는 하나 이상의 프로세서 및 컨트롤러를 포함한 시스템온칩(System on Chip, SoC)으로 구현될 수 있다.
- [0069] 이미지 변환 장치는 하드웨어적 요소가 마련된 컴퓨팅 디바이스 또는 서버에 소프트웨어, 하드웨어, 또는 이들의 조합하는 형태로 탑재될 수 있다. 컴퓨팅 디바이스 또는 서버는 각종 기기 또는 유무선 통신망과 통신을 수행하기 위한 통신 모듈 등의 통신장치, 프로그램을 실행하기 위한 데이터를 저장하는 메모리, 프로그램을 실행하여 연산 및 명령하기 위한 마이크로프로세서 등을 전부 또는 일부 포함한 다양한 장치를 의미할 수 있다.
- [0070] 도 5에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 5에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이다.
- [0071] 본 실시예들에 따른 동작은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능한 매체에 기록될 수 있다. 컴퓨터 판독 가능한 매체는 실행을 위해 프로세서에 명령어를 제공하는 데 참여한 임의의 매체를 나타낸다. 컴퓨터 판독 가능한 매체는 프로그램 명령, 데이터 파일, 데이터 구조 또는 이들의 조합을 포함할 수 있다. 예를 들면, 자기 매체, 광기록 매체, 메모리 등이 있을 수 있다. 컴퓨터 프로그램은 네트워크로 연결된 컴퓨터 시스템 상에 분산되어 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다. 본 실시예를 구현하기 위한 기능적인(Functional) 프로그램, 코드, 및 코드 세그먼트들은 본 실시예가 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있을 것이다.
- [0072] 본 실시예들은 본 실시예의 기술 사상을 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

도면

도면1



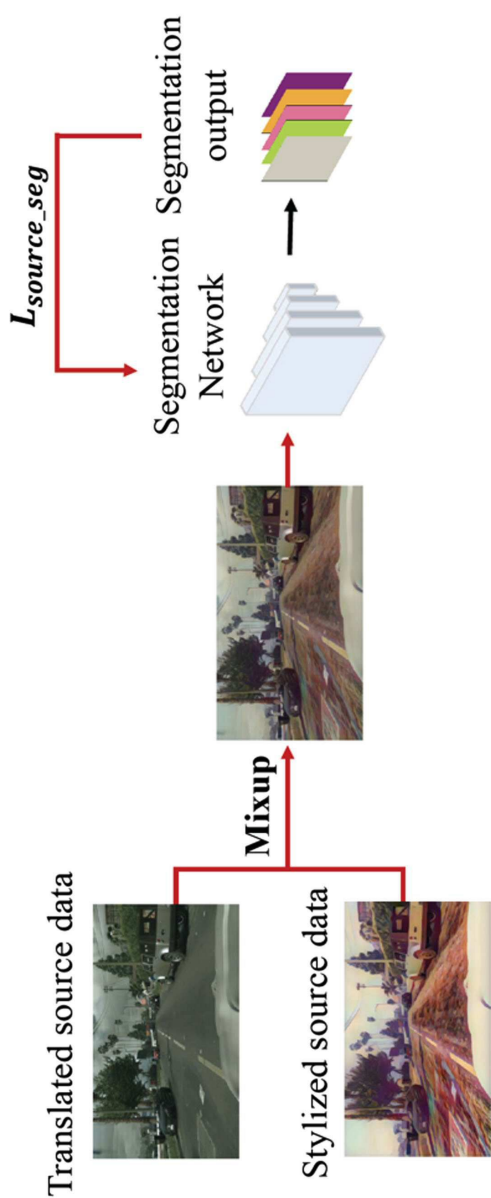
도면2



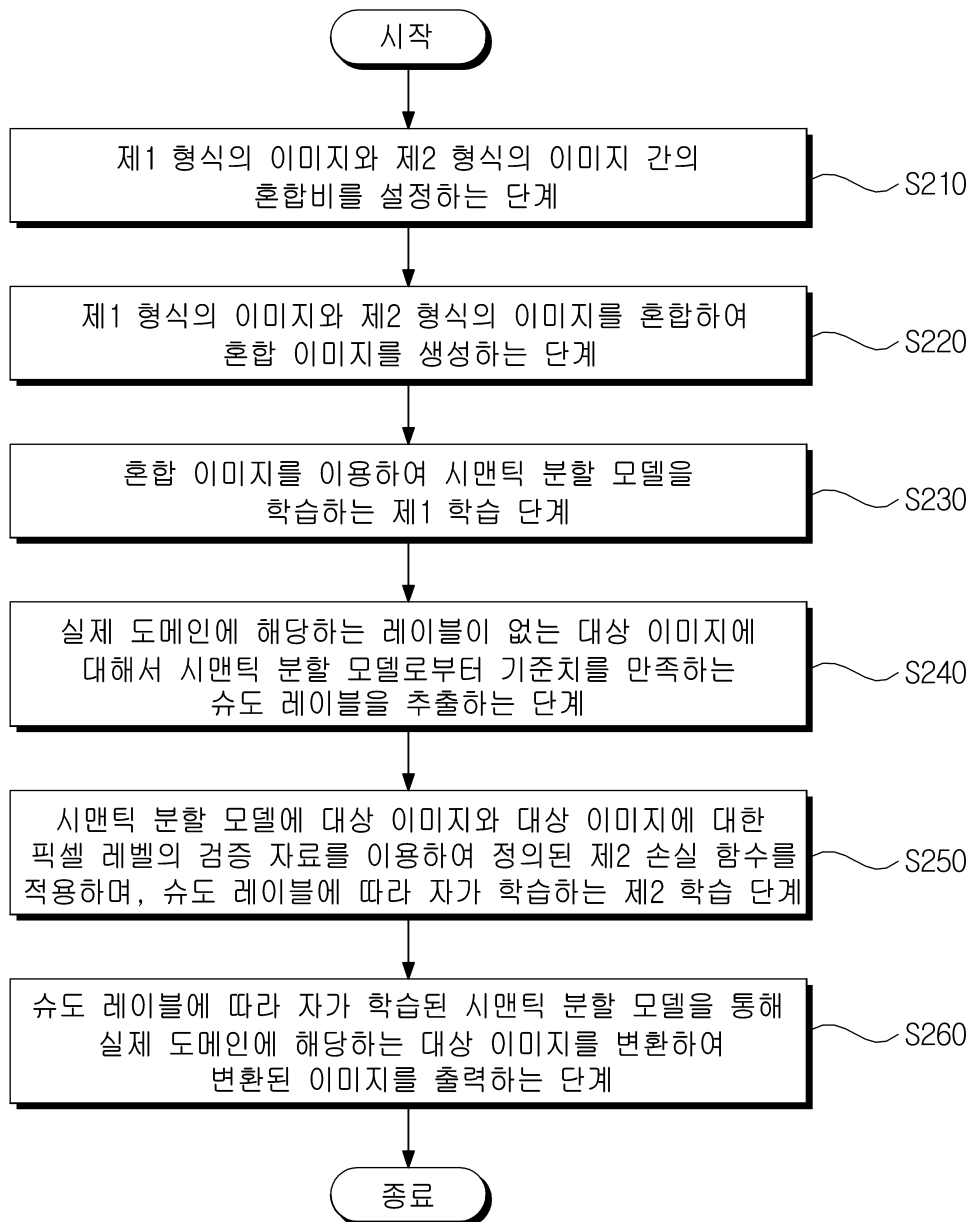
도면3

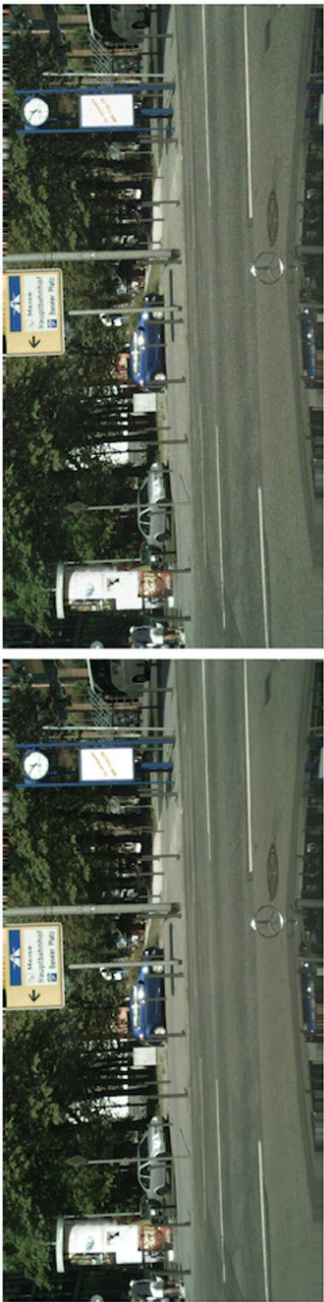


도면4

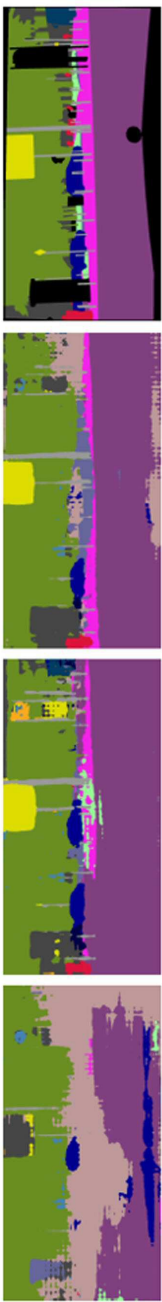


도면5





(a) Original image (b) Image w/Gaussian noise



(c) AdaptSegNet (d) LTIR (e) $T(0.5) + S(0.5)$ (f) Ground Truth

도면7

