



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0101256
(43) 공개일자 2022년07월19일

(51) 국제특허분류(Int. Cl.)
G06K 9/00 (2022.01) G06N 20/00 (2019.01)
(52) CPC특허분류
G06V 20/41 (2022.01)
G06N 20/00 (2021.08)
(21) 출원번호 10-2021-0003078
(22) 출원일자 2021년01월11일
심사청구일자 2021년01월11일

(71) 출원인
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
김은태
서울특별시 서대문구 연세로 50, 제3공학관 C607호(신촌동, 연세대학교)
이수현
서울특별시 서대문구 연세로 50, 제3공학관 C607호(신촌동, 연세대학교)
(74) 대리인
특허법인우인

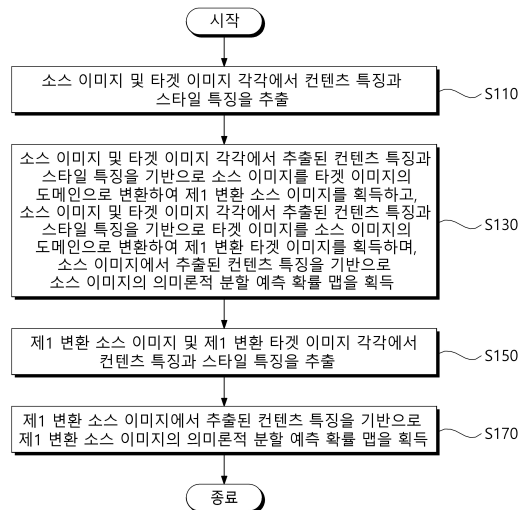
전체 청구항 수 : 총 5 항

(54) 발명의 명칭 제로 스타일 로스를 이용한 도메인 적응 장치 및 방법

(57) 요약

본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 장치 및 방법은, 제로 스타일 로스(zero style loss)를 통해 이미지를 콘텐츠와 스타일로 완전히 분리하고, 소스 이미지와 타겟 이미지가 콘텐츠 공간에서 정렬되도록 하여, 비지도 도메인 적응(Unsupervised Domain Adaptation)에 사용함으로써, 의미론적 분할(Semantic Segmentation) 신경망이 타겟 이미지에 대해 지도 학습되는 효과를 획득할 수 있다.

대표도 - 도4



(52) CPC특허분류

G06V 20/46 (2022.01)

(72) 발명자

현준혁

서울특별시 서대문구 연세로 50, 제3공학관 C607
호(신촌동, 연세대학교)

성홍제

서울특별시 서대문구 연세로 50, 제3공학관 C607
호(신촌동, 연세대학교)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711110608
과제번호	2019R1A2C1007153
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	중견연구자지원사업
연구과제명	자율주행차량에서 사고 방지를 위한 강인한 객체 검출 및 추적 시스템 개발(2/3)
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2020.03.01 ~ 2021.02.28

명세서

청구범위

청구항 1

소스 이미지 및 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출하는 제1 특징 추출부;

상기 소스 이미지 및 상기 타겟 이미지 각각에 대해 상기 제1 특징 추출부를 통해 추출된 콘텐츠 특징과 스타일 특징을 기반으로 상기 소스 이미지를 상기 타겟 이미지의 도메인으로 변환하여 제1 변환 소스 이미지를 획득하고, 상기 소스 이미지 및 상기 타겟 이미지 각각에 대해 상기 제1 특징 추출부를 통해 추출된 콘텐츠 특징과 스타일 특징을 기반으로 상기 타겟 이미지를 상기 소스 이미지의 도메인으로 변환하여 제1 변환 타겟 이미지를 획득하며, 상기 소스 이미지에 대해 상기 제1 특징 추출부를 통해 추출된 콘텐츠 특징을 기반으로 상기 소스 이미지의 의미론적 분할 예측 확률 맵을 획득하는 제1 학습부;

상기 제1 학습부를 통해 획득된 상기 제1 변환 소스 이미지 및 상기 제1 변환 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출하는 제2 특징 추출부; 및

상기 제1 변환 소스 이미지에 대해 상기 제2 특징 추출부를 통해 추출된 콘텐츠 특징을 기반으로 상기 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵을 획득하는 제2 학습부;

를 포함하는 제로 스타일 로스를 이용한 도메인 적응 장치.

청구항 2

제1항에서,

상기 제1 특징 추출부는,

상기 소스 이미지로부터 상기 타겟 이미지의 스타일을 추출하면 0이 되도록 하고, 상기 타겟 이미지로부터 상기 소스 이미지의 스타일을 추출하면 0이 되도록 하는 제1 로스(loss)를 기반으로, 상기 소스 이미지 및 상기 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출하는,

제로 스타일 로스를 이용한 도메인 적응 장치.

청구항 3

제2항에서,

상기 제2 특징 추출부는,

상기 제1 변환 소스 이미지로부터 상기 타겟 이미지의 스타일을 추출하면 0이 되도록 하고, 상기 제1 변환 타겟 이미지로부터 상기 소스 이미지의 스타일을 추출하면 0이 되도록 하는 제2 로스(loss)를 기반으로, 상기 제1 변환 소스 이미지 및 상기 제1 변환 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출하는,

제로 스타일 로스를 이용한 도메인 적응 장치.

청구항 4

제3항에서,

상기 제1 학습부는,

상기 소스 이미지의 의미론적 분할 예측 확률 맵과 상기 소스 이미지의 의미론적 분할 정답값의 차이가 최소화 되도록 하는 제3 로스(loss)를 기반으로, 상기 소스 이미지의 의미론적 분할 예측 확률 맵을 획득하고,

상기 제2 학습부는,

상기 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵과 상기 제1 변환 소스 이미지의 의미론적 분할 정답값의 차이가 최소화 되도록 하는 제4 로스(loss)를 기반으로, 상기 제1 변환 소스 이미지의 의미론적 분할 예측

확률 맵을 획득하는,

제로 스타일 로스를 이용한 도메인 적응 장치.

청구항 5

소스 이미지 및 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출하는 단계;

상기 소스 이미지 및 상기 타겟 이미지 각각에서 추출된 콘텐츠 특징과 스타일 특징을 기반으로 상기 소스 이미지를 상기 타겟 이미지의 도메인으로 변환하여 제1 변환 소스 이미지를 획득하고, 상기 소스 이미지 및 상기 타겟 이미지 각각에서 추출된 콘텐츠 특징과 스타일 특징을 기반으로 상기 타겟 이미지를 상기 소스 이미지의 도메인으로 변환하여 제1 변환 타겟 이미지를 획득하며, 상기 소스 이미지에서 추출된 콘텐츠 특징을 기반으로 상기 소스 이미지의 의미론적 분할 예측 확률 맵을 획득하는 단계;

상기 제1 변환 소스 이미지 및 상기 제1 변환 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출하는 단계; 및

상기 제1 변환 소스 이미지에서 추출된 콘텐츠 특징을 기반으로 상기 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵을 획득하는 단계;

를 포함하는 제로 스타일 로스를 이용한 도메인 적응 방법.

발명의 설명

기술 분야

[0001] 본 발명은 제로 스타일 로스를 이용한 도메인 적응 장치 및 방법에 관한 것으로서, 더욱 상세하게는 서로 다른 이미지 세트 사이의 도메인 차이를 줄이는, 장치 및 방법에 관한 것이다.

배경 기술

[0002] 의미론적 분할(Semantic Segmentation)은 이미지의 시맨틱 카테고리를 픽셀 단위로 분류하는 문제이다. 의미론적 분할을 위한 깊은 신경망(Deep Neural Network)을 학습하기 위해서는 아주 많은 양의 픽셀 단위 카테고리가 표시된 데이터 세트가 필요하다. 픽셀 단위의 정답값 취득은 굉장히 노동 집약적이고 어려운 일이다. 따라서, 정답값 취득이 간편한 가상 데이터 세트로 깊은 신경망을 학습하여 타겟 데이터 세트에서의 의미론적 분할 성능을 높이기 위한 연구가 진행되어 오고 있다.

[0003] 서로 다른 데이터 세트 사이에는 도메인 차이가 존재한다. 가상의 소스(Source) 데이터 세트와 실제 타겟(Target) 데이터 세트 사이의 도메인 차이는 소스 데이터 세트로 학습된 의미론적 분할 신경망이 타겟 데이터 세트에서의 높은 성능을 내는 것을 방해한다. 비지도 도메인 적응 방법(Unsupervised Domain Adaptation)은 두 데이터 세트 사이의 도메인 차이를 줄여 타겟 데이터 세트에서의 높은 의미론적 분할 성능을 얻기 위해 제안되었다. 도메인 차이를 줄이기 위한 방법으로는 두 데이터 세트를 특징 단계에서 정렬시키는 방법(feature-level alignment)이 있다.

발명의 내용

해결하려는 과제

[0004] 본 발명이 이루고자 하는 목적은, 제로 스타일 로스(zero style loss)를 통해 이미지를 콘텐츠와 스타일로 완전히 분리하고, 소스 이미지와 타겟 이미지가 콘텐츠 공간에서 정렬되도록 하여, 비지도 도메인 적응(Unsupervised Domain Adaptation)에 사용하는, 제로 스타일 로스를 이용한 도메인 적응 장치 및 방법을 제공하는 데 있다.

[0005] 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

- [0006] 상기의 목적을 달성하기 위한 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 장치는, 소스 이미지 및 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출하는 제1 특징 추출부; 상기 소스 이미지 및 상기 타겟 이미지 각각에 대해 상기 제1 특징 추출부를 통해 추출된 콘텐츠 특징과 스타일 특징을 기반으로 상기 소스 이미지를 상기 타겟 이미지의 도메인으로 변환하여 제1 변환 소스 이미지를 획득하고, 상기 소스 이미지 및 상기 타겟 이미지 각각에 대해 상기 제1 특징 추출부를 통해 추출된 콘텐츠 특징과 스타일 특징을 기반으로 상기 타겟 이미지를 상기 소스 이미지의 도메인으로 변환하여 제1 변환 타겟 이미지를 획득하며, 상기 소스 이미지에 대해 상기 제1 특징 추출부를 통해 추출된 콘텐츠 특징을 기반으로 상기 소스 이미지의 의미론적 분할 예측 확률 맵을 획득하는 제1 학습부; 상기 제1 학습부를 통해 획득된 상기 제1 변환 소스 이미지 및 상기 제1 변환 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출하는 제2 특징 추출부; 및 상기 제1 변환 소스 이미지에 대해 상기 제2 특징 추출부를 통해 추출된 콘텐츠 특징을 기반으로 상기 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵을 획득하는 제2 학습부;를 포함한다.
- [0007] 여기서, 상기 제1 특징 추출부는, 상기 소스 이미지로부터 상기 타겟 이미지의 스타일을 추출하면 0이 되도록 하고, 상기 타겟 이미지로부터 상기 소스 이미지의 스타일을 추출하면 0이 되도록 하는 제1 로스(loss)를 기반으로, 상기 소스 이미지 및 상기 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출할 수 있다.
- [0008] 여기서, 상기 제2 특징 추출부는, 상기 제1 변환 소스 이미지로부터 상기 타겟 이미지의 스타일을 추출하면 0이 되도록 하고, 상기 제1 변환 타겟 이미지로부터 상기 소스 이미지의 스타일을 추출하면 0이 되도록 하는 제2 로스(loss)를 기반으로, 상기 제1 변환 소스 이미지 및 상기 제1 변환 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출할 수 있다.
- [0009] 여기서, 상기 제1 학습부는, 상기 소스 이미지의 의미론적 분할 예측 확률 맵과 상기 소스 이미지의 의미론적 분할 정답값의 차이가 최소화 되도록 하는 제3 로스(loss)를 기반으로, 상기 소스 이미지의 의미론적 분할 예측 확률 맵을 획득하고, 상기 제2 학습부는, 상기 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵과 상기 제1 변환 소스 이미지의 의미론적 분할 정답값의 차이가 최소화 되도록 하는 제4 로스(loss)를 기반으로, 상기 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵을 획득할 수 있다.
- [0011] 상기의 목적을 달성하기 위한 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 방법은, 소스 이미지 및 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출하는 단계; 상기 소스 이미지 및 상기 타겟 이미지 각각에서 추출된 콘텐츠 특징과 스타일 특징을 기반으로 상기 소스 이미지를 상기 타겟 이미지의 도메인으로 변환하여 제1 변환 소스 이미지를 획득하고, 상기 소스 이미지 및 상기 타겟 이미지 각각에서 추출된 콘텐츠 특징과 스타일 특징을 기반으로 상기 타겟 이미지를 상기 소스 이미지의 도메인으로 변환하여 제1 변환 타겟 이미지를 획득하며, 상기 소스 이미지에서 추출된 콘텐츠 특징을 기반으로 상기 소스 이미지의 의미론적 분할 예측 확률 맵을 획득하는 단계; 상기 제1 변환 소스 이미지 및 상기 제1 변환 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출하는 단계; 및 상기 제1 변환 소스 이미지에서 추출된 콘텐츠 특징을 기반으로 상기 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵을 획득하는 단계;를 포함한다.

발명의 효과

- [0012] 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 장치 및 방법에 의하면, 제로 스타일 로스(zero style loss)를 통해 이미지를 콘텐츠와 스타일로 완전히 분리하고, 소스 이미지와 타겟 이미지가 콘텐츠 공간에서 정렬되도록 하여, 비지도 도메인 적응(Unsupervised Domain Adaptation)에 사용함으로써, 의미론적 분할(Semantic Segmentation) 신경망이 타겟 이미지에 대해 지도 학습되는 효과를 획득할 수 있다.
- [0013] 본 발명의 효과들은 이상에서 언급한 효과들로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [0014] 도 1은 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 장치를 설명하기 위한 블록도이다.
- 도 2는 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 과정을 설명하기 위한 도면이다.

도 3은 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 동작의 성능을 설명하기 위한 도면이다.

도 4는 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0015] 이하, 첨부된 도면을 참조하여 본 발명의 실시 예를 상세히 설명한다. 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시 예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.
- [0016] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다. 또 일반적으로 사용되는 사전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한 이상적으로 또는 과도하게 해석되지 않는다.
- [0017] 본 명세서에서 "제1", "제2" 등의 용어는 하나의 구성요소를 다른 구성요소로부터 구별하기 위한 것으로, 이들 용어들에 의해 권리범위가 한정되어서는 아니 된다. 예를 들어, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다.
- [0018] 본 명세서에서 각 단계들에 있어 식별부호(예를 들어, a, b, c 등)는 설명의 편의를 위하여 사용되는 것으로 식별부호는 각 단계들의 순서를 설명하는 것이 아니며, 각 단계들은 문맥상 명백하게 특정 순서를 기재하지 않는 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 단계들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [0019] 본 명세서에서, "가진다", "가질 수 있다", "포함한다" 또는 "포함할 수 있다"등의 표현은 해당 특징(예: 수치, 기능, 동작, 또는 부품 등의 구성요소)의 존재를 가리키며, 추가적인 특징의 존재를 배제하지 않는다.
- [0020] 또한, 본 명세서에 기재된 '~부'라는 용어는 소프트웨어 또는 FPGA(field-programmable gate array) 또는 ASIC과 같은 하드웨어 구성요소를 의미하며, '~부'는 어떤 역할들을 수행한다. 그렇지만 '~부'는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. '~부'는 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 '~부'는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로코드, 회로, 데이터 구조들 및 변수들을 포함한다. 구성요소들과 '~부'들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 '~부'들로 결합되거나 추가적인 구성요소들과 '~부'들로 더 분리될 수 있다.
- [0022] 이하에서 첨부한 도면을 참조하여 본 발명에 따른 제로 스타일 로스를 이용한 도메인 적응 장치 및 방법의 바람직한 실시예에 대해 상세하게 설명한다.
- [0024] 먼저, 도 1을 참조하여 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 장치에 대하여 설명한다.
- [0025] 도 1은 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 장치를 설명하기 위한 블록도이다.
- [0026] 도 1을 참조하면, 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 장치(이하 '도메인 적응 장치'라 한다)(100)는 제로 스타일 로스(zero style loss)를 통해 이미지를 콘텐츠(content)와 스타일(style)로 완전히 분리하고, 소스 이미지와 타겟 이미지가 콘텐츠 공간에서 정렬되도록 하여, 비지도 도메인 적응(Unsupervised Domain Adaptation)에 사용한다. 이에 따라, 본 발명은 의미론적 분할(Semantic

Segmentation) 신경망이 타겟 이미지에 대해 지도 학습되는 효과를 획득할 수 있다.

- [0028] 이를 위해, 도메인 적응 장치(100)는 제1 특징 추출부(110), 제1 학습부(130), 제2 특징 추출부(150) 및 제2 학습부(170)를 포함할 수 있다.
- [0029] 제1 특징 추출부(110)는 소스 이미지 및 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출한다.
- [0030] 이때, 제1 특징 추출부(110)는 소스 이미지로부터 타겟 이미지의 스타일을 추출하면 0이 되도록 하고, 타겟 이미지로부터 소스 이미지의 스타일을 추출하면 0이 되도록 하는 제1 로스(loss)를 기반으로, 소스 이미지 및 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출할 수 있다.
- [0032] 제1 학습부(130)는 소스 이미지 및 타겟 이미지 각각에 대해 제1 특징 추출부(110)를 통해 추출된 콘텐츠 특징과 스타일 특징을 기반으로 소스 이미지를 타겟 이미지의 도메인으로 변환하여 제1 변환 소스 이미지를 획득한다.
- [0033] 그리고, 제1 학습부(130)는 소스 이미지 및 타겟 이미지 각각에 대해 제1 특징 추출부(110)를 통해 추출된 콘텐츠 특징과 스타일 특징을 기반으로 타겟 이미지를 소스 이미지의 도메인으로 변환하여 제1 변환 타겟 이미지를 획득한다.
- [0034] 그리고, 제1 학습부(130)는 소스 이미지에 대해 제1 특징 추출부(110)를 통해 추출된 콘텐츠 특징을 기반으로 소스 이미지의 의미론적 분할 예측 확률 맵을 획득한다.
- [0035] 이때, 제1 학습부(130)는 소스 이미지의 의미론적 분할 예측 확률 맵과 소스 이미지의 의미론적 분할 정답값의 차이가 최소화 되도록 하는 제3 로스(loss)를 기반으로, 소스 이미지의 의미론적 분할 예측 확률 맵을 획득할 수 있다.
- [0036] 또한, 제1 학습부(130)는 타겟 이미지에 대해 제1 특징 추출부(110)를 통해 추출된 콘텐츠 특징을 기반으로 타겟 이미지의 의미론적 분할 예측 확률 맵을 획득할 수 있다.
- [0038] 제2 특징 추출부(150)는 제1 학습부(130)를 통해 획득된 제1 변환 소스 이미지 및 제1 변환 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출한다.
- [0039] 이때, 제2 특징 추출부(150)는 제1 변환 소스 이미지로부터 타겟 이미지의 스타일을 추출하면 0이 되도록 하고, 제1 변환 타겟 이미지로부터 소스 이미지의 스타일을 추출하면 0이 되도록 하는 제2 로스(loss)를 기반으로, 제1 변환 소스 이미지 및 제1 변환 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출할 수 있다.
- [0041] 제2 학습부(170)는 제1 변환 소스 이미지에 대해 제2 특징 추출부(150)를 통해 추출된 콘텐츠 특징을 기반으로 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵을 획득한다.
- [0042] 이때, 제2 학습부(170)는 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵과 제1 변환 소스 이미지의 의미론적 분할 정답값의 차이가 최소화 되도록 하는 제4 로스(loss)를 기반으로, 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵을 획득할 수 있다.
- [0043] 또한, 제2 학습부(170)는 제1 변환 타겟 이미지에 대해 제2 특징 추출부(150)를 통해 추출된 콘텐츠 특징을 기반으로 제2 변환 타겟 이미지의 의미론적 분할 예측 확률 맵을 획득할 수 있다.
- [0046] 그러면, 도 2를 참조하여 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 과정에 대하여 보다 자세하게 설명한다.
- [0047] 도 2는 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 과정을 설명하기 위한 도면이다.

[0048] 이미지는 콘텐츠 특징과 스타일 특징으로 분리될 수 있다. 콘텐츠는 스타일과 상관없이 의미론적 분할만을 위한 정보를 포함한다. 스타일은 의미론적 분할과 관계없이 도메인의 특성을 내포하고, 이는 도메인 차이를 만들게 된다. 따라서, 콘텐츠와 스타일의 완전한 분리는 의미론적 분할을 위한 콘텐츠 특징이 도메인에 상관없이 콘텐츠 공간에서 정렬되도록 한다. 소스 이미지의 도메인과 타겟 이미지의 도메인이 콘텐츠 공간에서 정렬되면, 소스 이미지의 정답값으로 학습된 신경망은 마치 타겟 이미지의 정답값으로 지도 학습된 것과 같은 효과를 얻을 수 있다. 본 발명은 제로 스타일 로스를 이용한 비지도 도메인 적응 과정을 통해 이미지의 스타일과 콘텐츠를 완전히 분리함으로써, 타겟 이미지에서의 의미론적 분할 성능을 향상시킬 수 있다.

[0050] 본 발명은 제로 스타일 로스를 통해 이미지를 콘텐츠 특징과 스타일 특징으로 완전히 분리한다. 그리고, 본 발명은 분리된 콘텐츠 특징과 스타일 특징을 조합하여, 도메인이 변환된 이미지들(소스 이미지의 스타일로 변환된 타겟 이미지, 타겟 이미지의 스타일로 변환된 소스 이미지)를 획득한다. 그리고, 본 발명은 도메인이 변환된 이미지들에 제로 스타일 로스를 적용하여 스타일과 콘텐츠와 완전한 분리를 강제한다. 또한, 본 발명은 콘텐츠 특징이 의미론적 분할을 위한 정보를 모두 내포하도록 하는 로스를 함께 사용한다.

[0051] 즉, 본 발명의 제로 스타일 로스를 이용한 도메인 적응 과정의 전체적인 구조는 도 2에 도시된 바와 같다. 본 발명은 이미지에서 소스 이미지의 스타일을 추출하는 인코더 E_s , 타겟 이미지의 스타일을 추출하는 인코더 E_t , 이미지에서 콘텐츠를 추출하는 인코더 E_c , 소스 이미지의 도메인으로 변환된 이미지를 생성하는 디코더 G_s , 타겟 이미지의 도메인으로 변환된 이미지를 생성하는 디코더 G_t , 의미론적 분할을 예측하기 위한 디코더 G_c 로 구성될 수 있다.

[0052] 본 발명이 다른 비지도 도메인 적응 방법과 차이가 있는 부분은 제로 스타일 로스의 유무이다. 본 발명에 따른 제로 스타일 로스는 세 부분으로 이루어져 있으며, 1)원본 이미지에서 스타일과 콘텐츠를 완전히 분리하기 위한 부분, 2)도메인이 변환된 이미지에서 스타일과 콘텐츠를 완전히 분리하기 위한 부분, 및 3)콘텐츠 특징이 의미론적 분할을 위한 정보를 전부 내포하도록 유도하기 위한 부분이다.

[0053] 먼저, 원본 이미지에서 스타일과 콘텐츠 분리를 위한 로스는 아래의 [수학식 1]과 같다.

수학식 1

$$L_{zero} = \|E_t(I_s) - 0\|_1 + \|E_s(I_t) - 0\|_1$$

[0055] 여기서, I_s 는 소스 이미지를 나타낸다. I_t 는 타겟 이미지를 나타낸다.

[0056] 즉, 소스 이미지로부터 타겟 이미지의 스타일을 추출하면 0이 되도록, 타겟 이미지로부터 소스 이미지의 스타일을 추출하면 0이 되도록 강제한다. 이를 통해, 두 도메인의 스타일 특징을 완벽히 분리하여 추출할 수 있다.

[0057] 소스 이미지 I_s 와 타겟 이미지 I_t 로부터 스타일 특징과 콘텐츠 특징을 분리하여 추출하면, 아래의 [수학식 2]와 같이 타겟 이미지의 도메인으로 변환된 소스 이미지 I_{s2t} 와 소스 이미지의 도메인으로 변환된 타겟 이미지 I_{t2s} 를 생성할 수 있다.

수학식 2

$$I_{s2t} = G_t(E_t(I_t), E_c(I_s))$$

$$I_{t2s} = G_s(E_s(I_s), E_c(I_t))$$

[0060] 도메인이 변환된 이미지에 대해서도 아래의 [수학식 3]과 같이 스타일과 콘텐츠 분리를 위한 로스를 설정한다.

수학식 3

$$L_{zero_trans} = \|E_s(I_{s2t}) - 0\|_1 + \|E_t(I_{t2s}) - 0\|_1$$

위의 [수학식 1]과 [수학식 3]은 원본 이미지와 도메인이 변환된 이미지로부터 스타일과 콘텐츠를 완전히 분리하도록 강제하지만, 콘텐츠 특징이 의미론적 분할을 위한 모든 정보를 내포하는 것을 보장하지는 않는다.

따라서, 콘텐츠 특징이 의미론적 분할을 위한 모든 정보를 내포하도록 하기 위해, 아래의 [수학식 4]와 같이 의미론적 분할 로스를 추가한다.

수학식 4

$$L_{seg} = L(P_s, Y_s) + L(P_{s2t}, Y_{s2t})$$

여기서, $P_s = G_c(E_c(I_s))$ 는 I_s 의 의미론적 분할 예측 확률 맵을 나타낸다. $P_{s2t} = G_c(E_c(I_{s2t}))$ 는 I_{s2t} 의 의미론적 분할 예측 확률 맵을 나타낸다. Y_s 는 I_s 의 의미론적 분할 정답값을 나타낸다. Y_{s2t} 는 I_{s2t} 의 의미론적 분할 정답값을 나타낸다.

I_{s2t} 는 I_s 의 도메인만을 타겟 이미지의 도메인으로 변환한 이미지이기 때문에, Y_s 와 Y_{s2t} 는 서로 같다. 이때, 의미론적 분할 예측 확률 맵과 의미론적 분할 정답값 사이의 로스는 크로스 엔트로피와 같은 함수로 계산될 수 있다.

본 발명에 다른 최종적인 제로 스타일 로스는 아래의 [수학식 5]와 같다.

수학식 5

$$L_{zero_style} = L_{zero} + L_{zero_trans} + L_{seg}$$

그러면, 도 3을 참조하여 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 동작의 성능에 대하여 설명한다.

도 3은 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 동작의 성능을 설명하기 위한 도면이다.

본 발명에 따른 제로 스타일 로스를 이용한 도메인 적응 동작이 적용된 비지도 적응 학습의 결과는 도 3에 도시된 바와 같다.

도 3을 참조하면, 소스 이미지로 GTA5(Richter, Stephan R., et al. "Playing for data: Ground truth from computer games." European conference on computer vision. Springer, Cham, 2016. 참조) 데이터 세트를 이용하고 타겟 이미지로 Cityscapes(Cordts, Marius, et al. "The cityscapes dataset for semantic urban scene understanding." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. 참조) 데이터 세트를 이용하여 학습한 신경망의 타겟 이미지에 대한 의미론적 분할 예측 결과를 확인할 수 있다.

즉, 도 3에 도시된 바와 같이, 본 발명에 따른 제로 스타일 로스를 이용한 도메인 적응 동작을 적용하는 경우, 타겟 이미지 I_t 에 대한 의미론적 분할 정답값 Y_t 과 의미론적 분할 예측 확률 맵 P_t 의 차이가 크지 않는 것을 확인할 수 있다.

[0076] 아래의 [표 1]은 제로 스타일 로스의 사용 유무에 따른 비지도 도메인 적응 학습의 성능을 나타낸다.

표 1

[0077]

	제로 스타일 로스 X	제로 스타일 로스 0
GTA5 -> Cityscapes	45.4	46.6
SYNTHIA -> Cityscapes	41.5	43.3

[0078] [표 1]을 참조하면, 타겟 이미지로 Cityscapes 데이터 세트를 이용하고 소스 이미지로 GTA5 데이터 세트와 SYNTHIA(Ros, German, et al. "The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. 참조) 데이터 세트를 각각 이용하였을 때, Cityscapes 데이터 세트의 평가 데이터 세트 (evaluation data set)에서의 의미론적 분할 성능을 mIoU(mean of lass-wise Intersection-over-Union)로 계산한 결과를 확인할 수 있다.

[0079] 즉, [표 1]을 통해 확인할 수 있듯이, 본 발명에 따른 제로 스타일 로스를 이용한 도메인 적응 동작을 적용하는 경우, 의미론적 분할 성능이 향상되는 것을 확인할 수 있다.

[0082] 그러면, 도 4를 참조하여 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 방법에 대하여 설명한다.

[0083] 도 4는 본 발명의 바람직한 실시예에 따른 제로 스타일 로스를 이용한 도메인 적응 방법을 설명하기 위한 흐름도이다.

[0084] 도 4를 참조하면, 도메인 적응 장치(100)는 소스 이미지 및 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출한다(S110).

[0085] 이때, 도메인 적응 장치(100)는 소스 이미지로부터 타겟 이미지의 스타일을 추출하면 0이 되도록 하고, 타겟 이미지로부터 소스 이미지의 스타일을 추출하면 0이 되도록 하는 제1 로스(loss)를 기반으로, 소스 이미지 및 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출할 수 있다.

[0087] 그런 다음, 도메인 적응 장치(100)는 소스 이미지 및 타겟 이미지 각각에서 추출된 콘텐츠 특징과 스타일 특징을 기반으로 소스 이미지를 타겟 이미지의 도메인으로 변환하여 제1 변환 소스 이미지를 획득하고, 소스 이미지 및 타겟 이미지 각각에서 추출된 콘텐츠 특징과 스타일 특징을 기반으로 타겟 이미지를 소스 이미지의 도메인으로 변환하여 제1 변환 타겟 이미지를 획득하며, 소스 이미지에서 추출된 콘텐츠 특징을 기반으로 소스 이미지의 의미론적 분할 예측 확률 맵을 획득한다(S130).

[0088] 이때, 도메인 적응 장치(100)는 소스 이미지의 의미론적 분할 예측 확률 맵과 소스 이미지의 의미론적 분할 정답값의 차이가 최소화 되도록 하는 제3 로스(loss)를 기반으로, 소스 이미지의 의미론적 분할 예측 확률 맵을 획득할 수 있다.

[0089] 또한, 도메인 적응 장치(100)는 타겟 이미지에서 추출된 콘텐츠 특징을 기반으로 타겟 이미지의 의미론적 분할 예측 확률 맵을 획득할 수 있다.

[0091] 이후, 도메인 적응 장치(100)는 제1 변환 소스 이미지 및 제1 변환 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출한다(S150).

[0092] 이때, 도메인 적응 장치(100)는 제1 변환 소스 이미지로부터 타겟 이미지의 스타일을 추출하면 0이 되도록 하고, 제1 변환 타겟 이미지로부터 소스 이미지의 스타일을 추출하면 0이 되도록 하는 제2 로스(loss)를 기반으로, 제1 변환 소스 이미지 및 제1 변환 타겟 이미지 각각에서 콘텐츠 특징과 스타일 특징을 추출할 수 있다.

- [0094] 그런 다음, 도메인 적응 장치(100)는 제1 변환 소스 이미지에서 추출된 콘텐츠 특징을 기반으로 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵을 획득한다(S170).
- [0095] 이때, 도메인 적응 장치(100)는 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵과 제1 변환 소스 이미지의 의미론적 분할 정답값의 차이가 최소화 되도록 하는 제4 로스(loss)를 기반으로, 제1 변환 소스 이미지의 의미론적 분할 예측 확률 맵을 획득할 수 있다.
- [0096] 또한, 도메인 적응 장치(100)는 제1 변환 타겟 이미지에서 추출된 콘텐츠 특징을 기반으로 제2 변환 타겟 이미지의 의미론적 분할 예측 확률 맵을 획득할 수 있다.
- [0099] 이상에서 설명한 본 발명의 실시예를 구성하는 모든 구성요소들이 하나로 결합하거나 결합하여 동작하는 것으로 기재되어 있다고 해서, 본 발명이 반드시 이러한 실시예에 한정되는 것은 아니다. 즉, 본 발명의 목적 범위 안에서라면, 그 모든 구성요소들이 하나 이상으로 선택적으로 결합하여 동작할 수도 있다. 또한, 그 모든 구성요소들이 각각 하나의 독립적인 하드웨어로 구현될 수 있지만, 각 구성요소들의 그 일부 또는 전부가 선택적으로 조합되어 하나 또는 복수개의 하드웨어에서 조합된 일부 또는 전부의 기능을 수행하는 프로그램 모듈을 갖는 컴퓨터 프로그램으로서 구현될 수도 있다. 또한, 이와 같은 컴퓨터 프로그램은 USB 메모리, CD 디스크, 플래쉬 메모리 등과 같은 컴퓨터가 읽을 수 있는 기록 매체(Computer Readable Media)에 저장되어 컴퓨터에 의하여 읽혀지고 실행됨으로써, 본 발명의 실시예를 구현할 수 있다. 컴퓨터 프로그램의 기록 매체로서는 자기기록매체, 광기록매체 등이 포함될 수 있다.
- [0100] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위 내에서 다양한 수정, 변경 및 치환이 가능할 것이다. 따라서, 본 발명에 개시된 실시예 및 첨부된 도면들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예 및 첨부된 도면에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

- [0101] 100 : 도메인 적응 장치,
110 : 제1 특징 추출부,
130 : 제1 학습부,
150 : 제2 특징 추출부,
170 : 제2 학습부

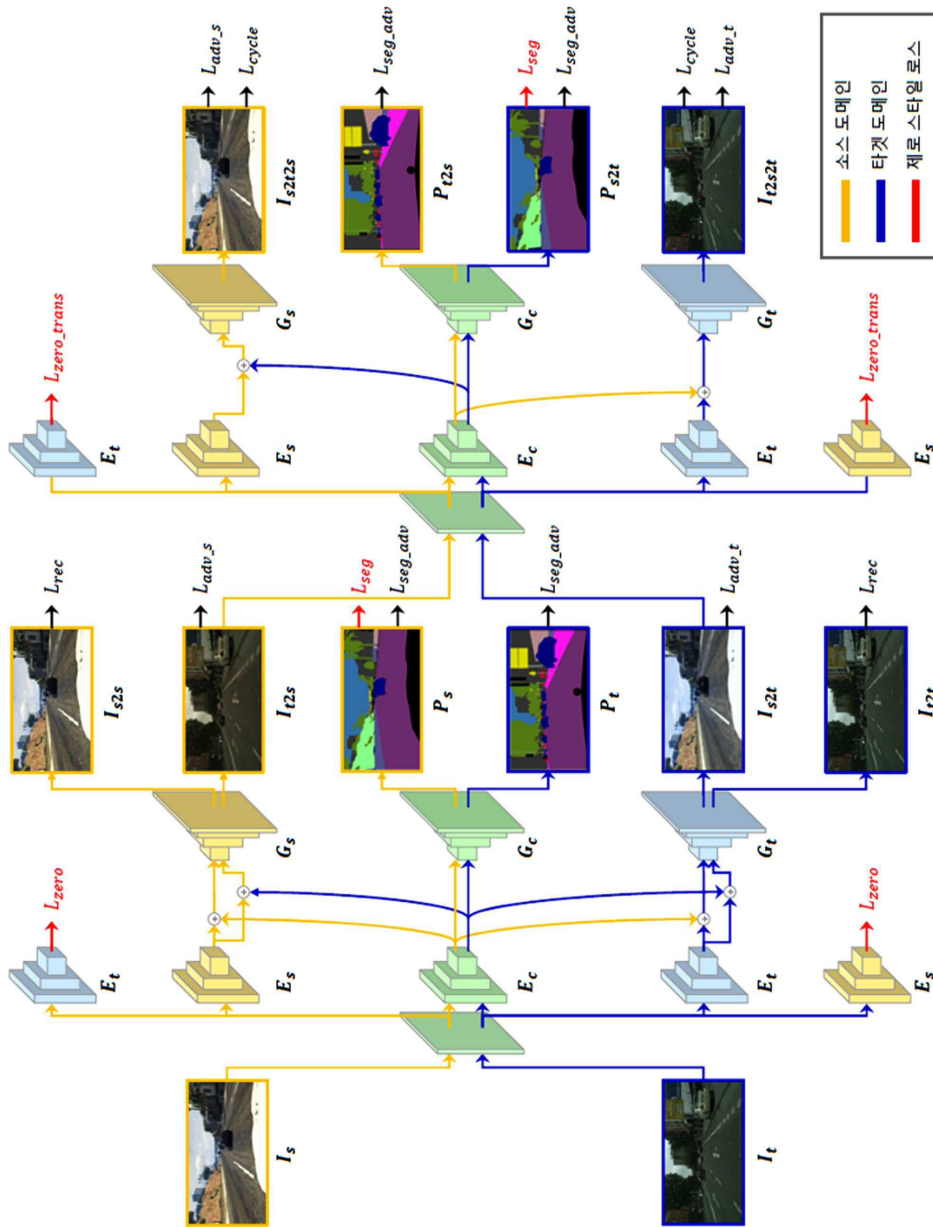
도면

도면1

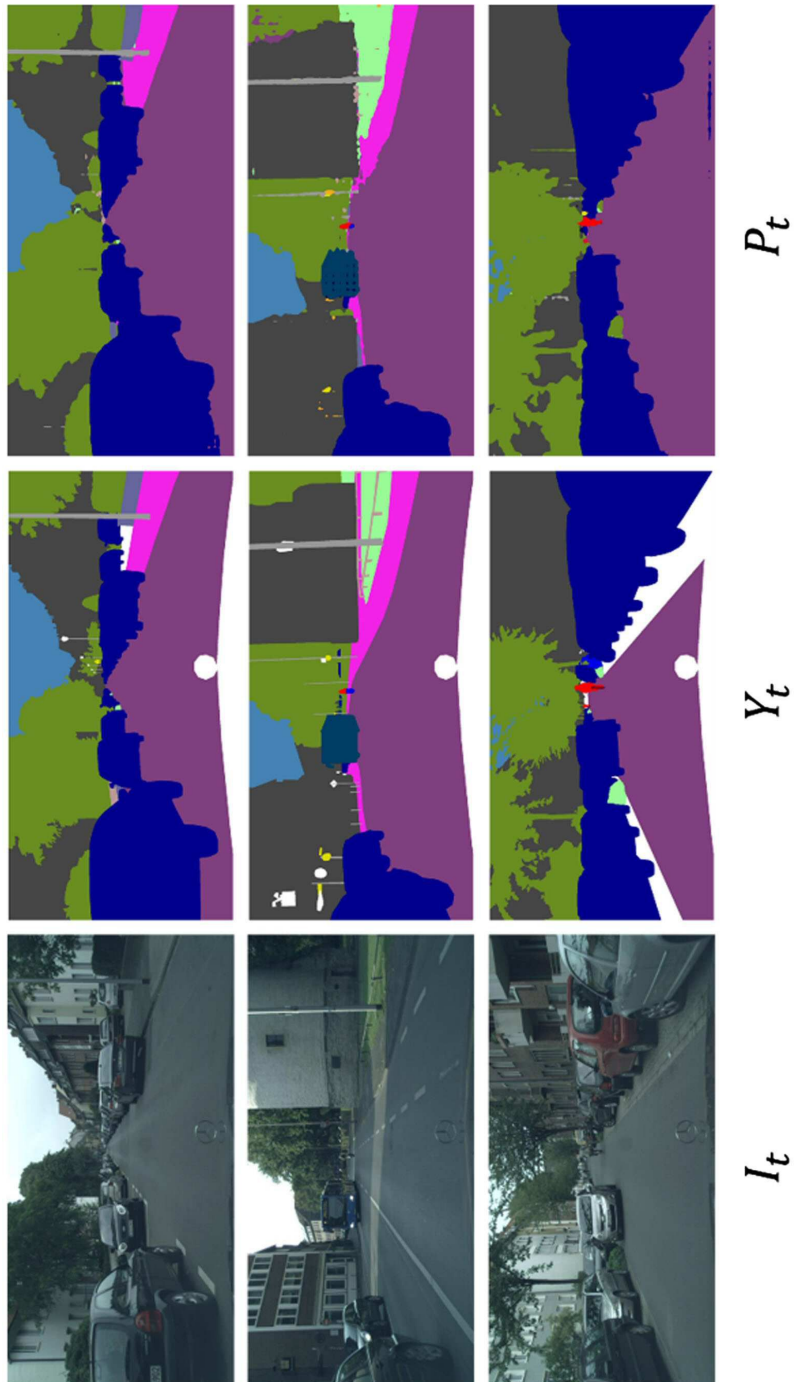
100



도면2



도면3



도면4

