



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년12월07일
(11) 등록번호 10-2611893
(24) 등록일자 2023년12월05일

- (51) 국제특허분류(Int. Cl.)
G06T 3/40 (2006.01) G06N 3/08 (2023.01)
G06V 10/40 (2022.01) G06V 10/74 (2022.01)
G06V 10/774 (2022.01)
- (52) CPC특허분류
G06T 3/4046 (2013.01)
G06N 3/088 (2023.01)
- (21) 출원번호 10-2023-0090553
(22) 출원일자 2023년07월12일
심사청구일자 2023년07월12일
- (56) 선행기술조사문헌
Xueqi Hu, et al., "QS-Attn: Query-Selected Attention for Contrastive Learning in I2I Translation", Computer Vision and Pattern Recognition, (2022.3.6.)
Sanghyun Woo, et al., "CBAM: Convolutional Block Attention Module", Computer Vision and Pattern Recognition, (2018.07.18.)
Shihao Yang, et al., "An Unpaired Thermal Infrared Image Translation Method Using GMA-CycleGAN", <https://doi.org/10.3390/rs15030663>, (2023.01.22.)
hen Zhao, et al., "Spectral Normalized Dual Contrastive Regularization for Image-to-Image Translation", Computer Vision and Pattern Recognition, (2023.05.23.)

- (73) 특허권자
경상국립대학교산학협력단
경상남도 진주시 진주대로 501 (가좌동)
- (72) 발명자
서현
경상남도 진주시 새평거로 25
오윤석
경상남도 진주시 영천강로 172
오선혜
경상남도 진주시 에나로 100
- (74) 대리인
김종석

전체 청구항 수 : 총 8 항

심사관 : 이정은

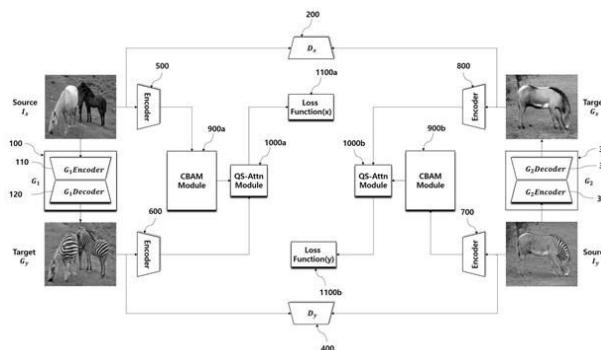
(54) 발명의 명칭 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치 및 방법

(57) 요약

본 발명은 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치 및 방법에 관한 것으로, 보다 구체적으로 제1 소스 이미지의 제1 도메인으로부터 제1 소스 특징을 추출하는 제1 소스 인코더, 제1 타겟 이미지의 제2 도메인으로부터 제1 타겟 특징을 추출하는 제1 타겟 인코더, 제2 소스 이미지의 제2 도메인으로부터 제2 소

(뒷면에 계속)

대표도 - 도1



스 특징을 추출하는 제2 소스 인코더, 제2 타겟 이미지의 제1 도메인으로부터 제2 타겟 특징을 추출하는 제2 타겟 인코더, 각 소스 인코더로부터 출력된 각 소스 특징으로부터 각 소스 이미지에서 어디에 초점을 맞춰야 하는지를 나타내는 리파인 특징을 각각 추출하는 컨볼루션 블록 어텐션 모듈 및 주어진 쿼리(Query)를 이용하여 각 리파인 특징에 대한 임시 어텐션 매트릭스를 각각 생성하고, 엔트로피를 기반으로 각 임시 어텐션 매트릭스를 정렬한 후 선택된 행을 갖는 최종 어텐션 매트릭스를 출력하는 쿼리 어텐션 모듈을 포함하는 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치 및 방법에 관한 것이다.

(52) CPC특허분류

G06V 10/40 (2023.08)

G06V 10/761 (2023.08)

G06V 10/774 (2023.08)

G06T 2207/20081 (2013.01)

G06T 2207/20084 (2013.01)

명세서

청구범위

청구항 1

제1 소스 이미지의 제1 도메인으로부터 제1 소스 특징을 추출하는 제1 소스 인코더;

제1 타겟 이미지의 제2 도메인으로부터 제1 타겟 특징을 추출하는 제1 타겟 인코더;

제2 소스 이미지의 제2 도메인으로부터 제2 소스 특징을 추출하는 제2 소스 인코더;

제2 타겟 이미지의 제1 도메인으로부터 제2 타겟 특징을 추출하는 제2 타겟 인코더;

각 소스 인코더로부터 출력된 각 소스 특징의 채널 및 픽셀 위치에 대한 확률값에 따라 각 소스 특징을 인코딩하고 리파인 특징을 각각 추출하는 컨볼루션 블록 어텐션 모듈; 및

주어진 쿼리(Query)를 이용하여 각 리파인 특징에 대한 어텐션 연산이 수행됨으로써 특징 클래스에 속할 확률값이 각 픽셀에 표시된 임시 어텐션 매트릭스를 각각 생성하고, 엔트로피를 기반으로 각 임시 어텐션 매트릭스를 정렬한 후 선택된 행을 갖는 최종 어텐션 매트릭스를 출력하는 쿼리 어텐션 모듈;을 포함하는 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치.

청구항 2

제 1항에 있어서,

상기 제1 타겟 인코더에 상기 제1 타겟 이미지가 입력되도록 상기 제1 소스 이미지로부터 상기 제2 도메인을 포함하는 상기 제1 타겟 이미지를 생성하는 제1 생성자;

상기 제2 타겟 인코더에 상기 제2 타겟 이미지가 입력되도록 상기 제2 소스 이미지로부터 상기 제1 도메인을 포함하는 상기 제2 타겟 이미지를 생성하는 제2 생성자;

상기 제1 소스 이미지와 상기 제2 생성자로부터 생성된 상기 제2 타겟 이미지 간의 진위여부를 판별하는 제1 판별자; 및

상기 제2 소스 이미지와 상기 제1 생성자로부터 생성된 상기 제1 타겟 이미지 간의 진위여부를 판별하는 제2 판별자;를 포함하는 것을 특징으로 하는 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치.

청구항 3

제 1항에 있어서,

상기 쿼리 어텐션 모듈은,

각 리파인 특징의 값(Value)과 각 최종 어텐션 매트릭스를 연산하여 소스 어텐션 매트릭스를 각각 출력하는 소스 어텐션 매트릭스 출력부; 및

각 타겟 특징의 값(Value)과 각 최종 어텐션 매트릭스를 연산하여 타겟 어텐션 매트릭스를 각각 출력하는 타겟 어텐션 매트릭스 출력부;를 포함하는 것을 특징으로 하는 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치.

청구항 4

제 3항에 있어서,

각 소스 어텐션 매트릭스와 각 타겟 어텐션 매트릭스 내 쌍(pair)의 유사도를 연산하고, 손실함수를 이용하여 각 최종 어텐션 매트릭스의 손실값을 줄이는 방향으로 학습하는 학습부;를 더 포함하는 것을 특징으로 하는 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치.

청구항 5

제1 소스 인코더에 의하여, 제1 소스 이미지의 제1 도메인으로부터 제1 소스 특징이 추출되는 제1 소스 특징 추출단계;

제1 타겟 인코더에 의하여, 제1 타겟 이미지의 제2 도메인으로부터 제1 타겟 특징이 추출되는 제1 타겟 특징 추출단계;

제2 소스 인코더에 의하여, 제2 소스 이미지의 제2 도메인으로부터 제2 소스 특징이 추출되는 제2 소스 특징 추출단계;

제2 타겟 인코더에 의하여, 제2 타겟 이미지의 제1 도메인으로부터 제2 타겟 특징이 추출되는 제2 타겟 특징 추출단계;

컨볼루션 블록 어텐션 모듈에 의하여, 각 소스 인코더로부터 출력된 각 소스 특징의 채널 및 픽셀 위치에 대한 확률값에 따라 각 소스 특징을 인코딩하고 리파인 특징이 각각 추출되는 리파인 특징 추출단계;

쿼리 어텐션 모듈에 의하여, 주어진 쿼리(Query)가 이용되어 각 리파인 특징에 대한 어텐션 연산이 수행됨으로써 특징 클래스에 속할 확률값이 각 픽셀에 표시된 임시 어텐션 매트릭스가 각각 생성되는 임시 어텐션 매트릭스 생성단계; 및

상기 쿼리 어텐션 모듈에 의하여, 엔트로피를 기반으로 각 임시 어텐션 매트릭스가 정렬된 후 선택된 행을 갖는 최종 어텐션 매트릭스가 출력되는 최종 어텐션 매트릭스 출력단계;를 포함하는 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 방법.

청구항 6

제 5항에 있어서,

상기 쿼리 어텐션 모듈에 의하여, 각 리파인 특징의 값(Value)과 각 최종 어텐션 매트릭스가 연산되어 소스 어텐션 매트릭스가 각각 출력되는 소스 어텐션 매트릭스 출력단계; 및

상기 쿼리 어텐션 모듈에 의하여, 각 타겟 특징의 값(Value)과 각 최종 어텐션 매트릭스가 연산되어 타겟 어텐션 매트릭스가 각각 출력되는 타겟 어텐션 매트릭스 출력단계;를 더 포함하는 것을 특징으로 하는 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 방법.

청구항 7

제 6항에 있어서,

학습부에 의하여, 각 소스 어텐션 매트릭스와 각 타겟 어텐션 매트릭스 내 쌍(pair)의 유사도가 연산되는 유사도 연산단계; 및

상기 학습부에 의하여, 손실함수가 이용되어 긍정적인 쌍(Positive pair)의 유사도가 작아지도록 학습하고, 부정적인 쌍(Negative pair)의 유사도가 높아지도록 학습하는 학습단계;를 더 포함하는 것을 특징으로 하는 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 방법.

청구항 8

제 5항에 있어서,

제1 생성자에 의하여, 상기 제1 소스 이미지로부터 제2 도메인을 포함하는 제1 타겟 이미지가 생성되는 제1 타겟 이미지 생성단계;

제2 생성자에 의하여, 상기 제2 소스 이미지로부터 제1 도메인을 포함하는 제2 타겟 이미지가 생성되는 제2 타겟 이미지 생성단계;

제1 판별자에 의하여, 상기 제1 소스 이미지와 상기 제2 타겟 이미지 간의 진위여부가 판별되는 제1 판별단계; 및

제2 판별자에 의하여, 상기 제2 소스 이미지와 상기 제1 타겟 이미지 간의 진위여부가 판별되는 제2 판별단계;를 더 포함하는 것을 특징으로 하는 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 방법.

발명의 설명

기술 분야

[0001] 본 발명은 생성적 적대 신경망(GAN)을 대칭적으로 배치함으로써, 두 도메인에 대해 동시에 학습이 가능하도록 하는 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치 및 방법에 관한 것이다.

배경 기술

[0002] Image-to-Image(I2I) 변환은 소스 도메인에서 대상 도메인으로 이미지를 생성하는 것을 목표로 하는 컴퓨터 비전과 기계학습의 분야 중 하나이다. 이는, 생성적 적대 신경망(Generative Adversarial Networks;GAN)의 출현으로 크게 발전하였다.

[0003] 생성적 적대 신경망(GAN)은 비지도 학습을 기반으로 하고, 임의의 랜덤 노이즈로부터 가상의 데이터를 생성하는 생성자(Generator)가 더 실제와 같은 데이터를 생성할 수 있도록 진위여부를 판별할 수 있는 구별자(Discriminator)를 붙여 경쟁적으로 학습시키는 모델이다. 상대적으로 규모가 작은 데이터셋에서는 좋은 성능을 발휘함으로써 효율적인 면에서 유리한 모델이다. 반면에, 생성적 적대 신경망(GAN)은 생성자와 판별자를 동시에 경쟁적으로 학습시킨다는 점에서 학습이 굉장히 불안정하고, 대규모 데이터셋으로 학습시키기 힘들다는 단점이 존재한다.

[0004] Image-to-Image(I2I) 변환에 있어서, 생성적 적대 신경망(GAN)은 입력값과 출력값을 한 쌍으로 하는 데이터를 이용한 지도 학습에서 품질 좋은 이미지를 생성할 수 있다. 그러나 지도 학습을 위해 픽셀 대 픽셀 매핑으로 한 쌍을 이룬 데이터를 수집하는 것이 상당히 어렵고, 이러한 데이터를 구축하는데 상당한 비용이 드는 기술적 한계가 존재한다.

[0005] Image-to-Image(I2I) 변환에 있어서, 생성적 적대 신경망(GAN)은 사전 정보 즉, 레이블(Label)이 없는 데이터를 활용한 비지도 학습 역시 가능하나, 적대적 손실이 제한되지 않고 도메인 간에 여러 가능한 매핑이 있어 품질이 좋지 않은 이미지를 생성할 수 있다. 이러한 문제를 해결하기 위해서 개발된 CycleGAN, DiscoGAN 및 DualGAN은 주기 일관성 가정을 하게 되는데, 이 또한 형상 변경을 수행하는 기능을 제한할 수 있으며 두 도메인 간의 관계를 항상 이상적인 전단사(bijection)로 만들지는 않는다.

[0006] CycleGAN, DiscoGAN 및 DualGAN의 기술적 한계를 보완하기 위해서 DCLGAN과 QS-Attn이 개발되었다. DCLGAN은 대조 학습에서 대상 앵커(Anchor)를 선택하지 않으며, 각 앵커 특징은 다른 포지션과의 관계를 고려하지 않고 수용영역만 제한된다는 문제점이 있다. QS-Attn은 두 개의 개별 도메인에 대해 하나의 임베딩을 사용하여 도메인 격차를 포착하지 못하여 훈련 안정성이 감소할 수 있다. 그리고 QS-Attn은 앵커 특징을 얻었을 때 해당 특징이 도메인 특성을 반영하는 데 얼마나 중요한지를 측정하기 위해 엔트로피 분포를 매트릭스로 계산한 후 엔트로피를 정렬하여 이미지에서 가장 작은 N 포인트를 선택한다. 엔트로피 매트릭스를 분석했을 때 샘플 수를 초과하는 제로 엔트로피 포인트를 발견하였는데 이는 중요한 특징 선택이 누락된 것이다. 즉, QS-Attn은 대상이 아닌 배경에 해당하는 부분에 누락된 특징들이 집중되어 있을 수 있다.

[0007] 이에, Image-to-Image(I2I) 변환에 있어서 비지도 학습이 가능한 GAN을 이용하여 보다 품질이 좋은 이미지를 생성할 수 있는 기술이 본 기술분야에서 절실히 필요한 실정이다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 대한민국 공개특허공보 제10-2023-0070128호
(특허문헌 0002) 대한민국 등록특허공보 제10-1975186호

발명의 내용

해결하려는 과제

[0009] 본 발명은 상기와 같은 문제점을 해결하기 위한 것으로, 비지도 학습 기반으로 보다 품질이 좋은 이미지를 생성

할 수 있도록 생성적 적대 신경망(GAN)을 대칭적으로 배치하여 두 도메인에 대해 동시에 학습이 가능하도록 하는 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치 및 방법을 얻고자 하는 것을 목적으로 한다.

[0010] 본 발명이 이루고자 하는 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 본 발명의 기재로부터 당해 분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있다.

과제의 해결 수단

[0011] 상기 목적을 달성하기 위하여, 본 발명의 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치는 제1 소스 이미지의 제1 도메인으로부터 제1 소스 특징을 추출하는 제1 소스 인코더; 제1 타겟 이미지의 제2 도메인으로부터 제1 타겟 특징을 추출하는 제1 타겟 인코더; 제2 소스 이미지의 제2 도메인으로부터 제2 소스 특징을 추출하는 제2 소스 인코더; 제2 타겟 이미지의 제1 도메인으로부터 제2 타겟 특징을 추출하는 제2 타겟 인코더; 각 소스 인코더로부터 출력된 각 소스 특징으로부터 각 소스 이미지에서 어디에 초점을 맞춰야 하는지를 나타내는 리파인 특징을 각각 추출하는 컨볼루션 블록 어텐션 모듈; 및 주어진 쿼리(Query)를 이용하여 각 리파인 특징에 대한 임시 어텐션 매트릭스를 각각 생성하고, 엔트로피를 기반으로 각 임시 어텐션 매트릭스를 정렬한 후 선택된 행을 갖는 최종 어텐션 매트릭스를 출력하는 쿼리 어텐션 모듈;을 제공한다.

[0012] 상기 목적을 달성하기 위하여, 본 발명의 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 방법은 제1 소스 인코더에 의하여, 제1 소스 이미지의 제1 도메인으로부터 제1 소스 특징이 추출되는 제1 소스 특징 추출단계; 제1 타겟 인코더에 의하여, 제1 타겟 이미지의 제2 도메인으로부터 제1 타겟 특징이 추출되는 제1 타겟 특징 추출단계; 제2 소스 인코더에 의하여, 제2 소스 이미지의 제2 도메인으로부터 제2 소스 특징이 추출되는 제2 소스 특징 추출단계; 제2 타겟 인코더에 의하여, 제2 타겟 이미지의 제1 도메인으로부터 제2 타겟 특징이 추출되는 제2 타겟 특징 추출단계; 컨볼루션 블록 어텐션 모듈에 의하여, 각 소스 인코더로부터 출력된 각 소스 특징으로부터 각 소스 이미지에서 어디에 초점을 맞춰야 하는지를 나타내는 리파인 특징이 각각 추출되는 리파인 특징 추출단계; 쿼리 어텐션 모듈에 의하여, 주어진 쿼리(Query)가 이용하여 각 리파인 특징에 대한 임시 어텐션 매트릭스가 각각 생성되는 임시 어텐션 매트릭스 생성단계; 및 상기 쿼리 어텐션 모듈에 의하여, 엔트로피를 기반으로 각 임시 어텐션 매트릭스가 정렬된 후 선택된 행을 갖는 최종 어텐션 매트릭스가 출력되는 최종 어텐션 매트릭스 출력단계;를 제공한다.

발명의 효과

[0013] 이상과 같이 본 발명에 의하면 생성적 적대 신경망(GAN)을 대칭적으로 배치하여 두 도메인에 대해 동시에 학습이 가능하도록 함으로써, 비지도 학습 기반으로 보다 품질이 좋은 이미지를 생성할 수 있는 효과가 있다.

[0014] 본 발명의 효과들은 이상에서 언급한 효과들로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 상세한 설명 및 청구범위의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0015] 도 1은 본 발명의 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치 구성도이다.
 도 2는 본 발명의 일실시예에 따른 제1 소스 이미지를 처리하는 제1 소스 인코더와, 제1 소스 특징을 처리하는 컨볼루션 블록 어텐션 모듈을 표시한 도면이다.
 도 3은 본 발명의 일실시예에 따른 제2 소스 이미지를 처리하는 제2 소스 인코더와, 제2 소스 특징을 처리하는 컨볼루션 블록 어텐션 모듈을 표시한 도면이다.
 도 4는 본 발명의 일실시예에 따른 제1 리파인 특징을 처리하는 쿼리 어텐션 모듈과 제1 타겟 이미지를 처리하는 제1 타겟 인코더와 학습부를 표시한 도면이다.
 도 5는 본 발명의 일실시예에 따른 제2 리파인 특징을 처리하는 쿼리 어텐션 모듈과 제2 타겟 이미지를 처리하는 제2 타겟 인코더와 학습부를 표시한 도면이다.
 도 6은 본 발명의 일실시예에 따른 제1 소스 이미지와 제1 타겟 이미지에 대한 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 방법 흐름도이다.

도 7은 본 발명의 일실시예에 따른 제2 소스 이미지와 제2 타겟 이미지에 대한 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 방법 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0016] 본 명세서에서 사용되는 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나, 이는 당 분야에 종사하는 기술자의 의도 또는 판례, 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 발명의 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서 본 발명에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 발명의 전반에 걸친 내용을 토대로 정의되어야 한다.
- [0017] 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 문맥상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0018] 이하, 본 발명에 따른 실시예를 첨부한 도면을 참조하여 상세히 설명하기로 한다. 도 1은 본 발명의 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치 구성도이다.
- [0019] 도 2는 본 발명의 일실시예에 따른 제1 소스 이미지(I_x)를 처리하는 제1 소스 인코더(500)와, 제1 소스 특징($F_{x,1}$)을 처리하는 컨볼루션층 블록 어텐션 모듈(900a)을 표시한 도면이다. 도 3은 본 발명의 일실시예에 따른 제2 소스 이미지(I_y)를 처리하는 제2 소스 인코더(700)와, 제2 소스 특징($F_{y,2}$)을 처리하는 컨볼루션층 블록 어텐션 모듈(900b)을 표시한 도면이다.
- [0020] 도 4는 본 발명의 일실시예에 따른 제1 리파인 특징(Refined $F_{x,1}$)을 처리하는 쿼리 어텐션 모듈(1000a)과 제1 타겟 이미지(G_y)를 처리하는 제1 타겟 인코더(600)와 학습부(1100a)를 표시한 도면이다. 도 5는 본 발명의 일실시예에 따른 제2 리파인 특징(Refined $F_{y,2}$)을 처리하는 쿼리 어텐션 모듈(1000b)과 제2 타겟 이미지(G_x)를 처리하는 제2 타겟 인코더(800)와 학습부(1100b)를 표시한 도면이다.
- [0021] 도 6은 본 발명의 일실시예에 따른 제1 소스 이미지(I_x)와 제1 타겟 이미지(G_y)에 대한 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 방법 흐름도이다. 도 7은 본 발명의 또 다른 일실시예에 따른 제2 소스 이미지(I_y)와 제2 타겟 이미지(G_x)에 대한 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 방법 흐름도이다.
- [0023] **이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치**
- [0024] 도 1을 보면, 본 발명의 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 장치는 제1 소스 인코더(500), 제1 타겟 인코더(600), 제2 소스 인코더(700) 및 제2 타겟 인코더(800), 컨볼루션층 블록 어텐션 모듈(900a, 900b) 및 쿼리 어텐션 모듈(1000a, 1000b)을 포함한다.
- [0025] 우선, 상기 제1 소스 인코더(500)는 제1 소스 이미지(I_x)의 제1 도메인(x)으로부터 제1 소스 특징($F_{x,1}$)을 추출한다. 상기 제1 타겟 인코더(600)는 제1 타겟 이미지(G_y)의 제2 도메인(y)으로부터 제1 타겟 특징($F_{y,1}$)을 추출한다.
- [0026] 다음으로, 상기 제2 소스 인코더(700)는 제2 소스 이미지(I_y)의 제2 도메인(y)으로부터 제2 소스 특징($F_{y,2}$)을 추출한다. 상기 제2 타겟 인코더(800)는 제2 타겟 이미지(G_x)의 제1 도메인(x)으로부터 제2 타겟 특징($F_{x,2}$)을 추출한다.
- [0027] 다음으로, 상기 컨볼루션층 블록 어텐션 모듈(900a, 900b)는 각 소스 인코더(500, 700)로부터 출력된 각 소스 특징($F_{x,1}$, $F_{y,2}$)으로부터 각 소스 이미지(I_x , I_y)에서 어디에 초점을 맞춰야 하는지를 나타내는 리파인 특징(Refined $F_{x,1}$, Refined $F_{y,2}$)을 각각 추출한다.

[0028] 도 2의 일실시예를 보면, 하나의 컨볼루셔널 블록 어텐션 모듈(900a)은 제1 소스 인코더(500)로부터 출력된 제1 소스 특징($F_{x,1}$)이 입력된다. 여기서, 상기 제1 소스 특징($F_{x,1}$)은 높이(H)*폭(W)*채널(C) 크기를 가질 수 있다. 하나의 컨볼루셔널 블록 어텐션 모듈(900a)은 하나의 채널 어텐션 모듈(910a)과 하나의 공간 어텐션 모듈(920a)를 포함할 수 있다.

[0029] 우선, 하나의 채널 어텐션 모듈(910a)은 상기 제1 소스 특징($F_{x,1}$)의 채널 간의 관계를 살펴 어떤 채널에 더 초점을 맞춰야 하는지 인코딩하기 위함이다. 이는 하기 [수학식 1]과 같다.

[0030] [수학식 1]

$$M_c(F_{x,1}) = \sigma(MLP(Maxpool(F_{x,1})))$$

[0031]

[0032] 우선 하나의 채널 어텐션 모듈(910a)은 최대 풀링(Max pooling)을 통해서 제1 소스 특징($F_{x,1}$)에서 채널별 최대값만을 벡터로 출력할 수 있다. 그리고 하나의 채널 어텐션 모듈(910a)은 최대 풀링(Max pooling)된 벡터들을 2개의 층(layer)을 갖는 다중 퍼셉트론(Multi-Layer Perceptron; MLP)을 통해서 차원을 축소할 수 있다. 하나의 채널 어텐션 모듈(910a)은 차원이 축소된 벡터들을 시그모이드 함수(Sigmoid function)를 이용하여 확률값으로 변환할 수 있다. 그리고 하나의 채널 어텐션 모듈(910a)은 채널별 확률값 중에서 가장 큰 확률값을 갖는 채널에 초점을 맞춰 상기 제1 소스 특징($F_{x,1}$)을 인코딩할 수 있다.

[0033] 이에 따라, 하나의 컨볼루셔널 블록 어텐션 모듈(900a)은 상기 제1 소스 특징($F_{x,1}$)과 하나의 채널 어텐션 모듈(910a)로부터 채널이 인코딩된 제1 소스 특징을 곱산하여 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)을 출력할 수 있다.

[0034] 또한, 하나의 공간 어텐션 모듈(920a)은 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)의 H*W*1 크기에서 다수 개의 픽셀 중 위치관계를 살펴 어떤 위치에 더 초점을 맞춰야 하는지 인코딩하기 위함이다. 이는 하기 [수학식 2]와 같다.

[0035] [수학식 2]

$$M_s(Channel_refined F_{x,1}) = \sigma(f^{7*7}(Maxpool(Channel_refined F_{x,1})))$$

[0036]

[0037] 우선 하나의 공간 어텐션 모듈(920a)은 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)이 입력될 수 있다. 하나의 공간 어텐션 모듈(920a)은 H*W*1 크기를 대상으로 하는 최대 풀링(Max pooling)을 통해서 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)에서 최대값만을 벡터로 출력할 수 있다. 그리고 하나의 공간 어텐션 모듈(920a)은 H*W*1 크기를 갖는 두 채널을 연결하여 H*W*2 크기의 특징맵을 생성할 수 있다. 하나의 공간 어텐션 모듈(920a)은 특징맵이 2채널이므로 7*7 컨볼루션 연산을 수행할 수 있다. 하나의 공간 어텐션 모듈(920a)은 시그모이드 함수(Sigmoid function)를 이용하여 7*7 컨볼루션 연산된 특징맵을 확률값으로 변환할 수 있다. 그리고 하나의 공간 어텐션 모듈(920a)은 확률값 중에서 가장 큰 확률값을 갖는 위치에 초점을 맞춰 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)을 인코딩할 수 있다.

[0038] 이에 따라, 하나의 컨볼루셔널 블록 어텐션 모듈(900a)은 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)과 하나의 공간 어텐션 모듈(920a)로부터 위치가 인코딩된 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)을 곱산하여 하나의 리파인 특징(Refined $F_{x,1}$)을 출력할 수 있다.

[0039] 도 3의 일실시예를 보면, 또 다른 하나의 컨볼루셔널 블록 어텐션 모듈(900b)은 제2 소스 인코더(700)로부터 출력된 제2 소스 특징($F_{y,2}$)이 입력된다. 여기서, 상기 제2 소스 특징($F_{y,2}$)은 높이(H)*폭(W)*채널(C) 크기를 가질 수 있다. 또 다른 하나의 컨볼루셔널 블록 어텐션 모듈(900b)은 또 다른 하나의 채널 어텐션 모듈(910b)과 또 다른 하나의 공간 어텐션 모듈(920b)를 포함할 수 있다.

[0040] 우선, 또 다른 하나의 채널 어텐션 모듈(910b)은 상기 제2 소스 특징($F_{y,2}$)의 채널 간의 관계를 살펴 어떤 채널에 더 초점을 맞춰야 하는지 인코딩하기 위함이다. 이는 하기 [수학식 3]과 같다.

[0041] [수학식 3]

$$M_C(F_{y,2}) = \sigma(MLP(Maxpool(F_{y,2})))$$

[0042]

[0043] 우선 또 다른 하나의 채널 어텐션 모듈(910b)은 최대 풀링(Max pooling)을 통해서 제2 소스 특징($F_{y,2}$)에서 채널별 최대값만을 벡터로 출력할 수 있다. 다음으로 또 다른 하나의 채널 어텐션 모듈(910b)은 최대 풀링(Max pooling)된 벡터들을 각각 2개의 층(layer)을 갖는 다중 퍼셉트론(Multi-Layer Perceptron; MLP)을 통해서 차원을 축소할 수 있다. 또 다른 하나의 채널 어텐션 모듈(910b)은 차원이 축소된 벡터들을 시그모이드 함수(Sigmoid function)를 이용하여 확률값으로 변환할 수 있다. 그리고 또 다른 하나의 채널 어텐션 모듈(910b)은 채널별 확률값 중에서 가장 큰 확률값을 갖는 채널에 초점을 맞춰 상기 제2 소스 특징($F_{y,2}$)을 인코딩할 수 있다.

[0044] 이에 따라, 또 다른 하나의 컨볼루션 블록 어텐션 모듈(900b)은 상기 제2 소스 특징($F_{y,2}$)과 또 다른 하나의 채널 어텐션 모듈(910b)로부터 채널이 인코딩된 제2 소스 특징을 곱산하여 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)을 출력할 수 있다.

[0045] 또한, 또 다른 하나의 공간 어텐션 모듈(920b)은 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)의 H*W*1 크기에서 다수 개의 픽셀 중 위치관계를 살펴 어떤 위치에 더 초점을 맞춰야 하는지 인코딩하기 위함이다. 이는 하기 [수학식 4]와 같다.

[0046] [수학식 4]

$$M_S(Channel_refined\ F_{y,2}) = \sigma(f^{7*7}(Maxpool(Channel_refined\ F_{y,2})))$$

[0047]

[0048] 우선 또 다른 하나의 공간 어텐션 모듈(920b)은 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)이 입력될 수 있다. 또 다른 하나의 공간 어텐션 모듈(920b)은 H*W*1 크기를 대상으로 하는 최대 풀링(Max pooling)을 통해서 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)에서 최대값만을 벡터로 출력할 수 있다. 그리고 또 다른 하나의 공간 어텐션 모듈(920b)은 H*W*1 크기를 갖는 두 채널을 연결하여 H*W*2 크기의 특징맵을 생성할 수 있다. 또 다른 하나의 공간 어텐션 모듈(920b)은 특징맵이 2채널이므로 7*7 컨볼루션 연산을 수행할 수 있다. 또 다른 하나의 공간 어텐션 모듈(920b)은 시그모이드 함수(Sigmoid function)를 이용하여 7*7 컨볼루션 연산된 특징맵을 확률값으로 변환할 수 있다. 그리고 또 다른 하나의 공간 어텐션 모듈(920b)은 확률값 중에서 가장 큰 확률값을 갖는 위치에 초점을 맞춰 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)을 인코딩할 수 있다.

[0049] 이에 따라, 또 다른 하나의 컨볼루션 블록 어텐션 모듈(900b)은 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)과 또 다른 하나의 공간 어텐션 모듈(920b)로부터 위치가 인코딩된 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)을 곱산하여 또 다른 하나의 리파인 특징(Refined $F_{y,2}$)을 출력할 수 있다.

[0050] 다음으로, 상기 쿼리 어텐션 모듈(1000a, 1000b)은 주어진 쿼리(Query)를 이용하여 각 리파인 특징(Refined $F_{x,1}$, Refined $F_{y,2}$)에 대한 임시 어텐션 매트릭스를 각각 생성하고, 엔트로피를 기반으로 각 임시 어텐션 매트릭스를 정렬한 후 선택된 행을 갖는 최종 어텐션 매트릭스를 출력한다.

[0051] 바람직하게, 상기 쿼리 어텐션 모듈(1000a, 1000b)은, 각 리파인 특징(Refined $F_{x,1}$, Refined $F_{y,2}$)의 값(Value)과 각 최종 어텐션 매트릭스를 연산하여 소스 어텐션 매트릭스를 각각 출력하는 소스 어텐션 매트릭스 출력부(1010a, 1010b) 및 각 타겟 특징($F_{y,1}$, $F_{x,2}$)의 값(Value)과 각 최종 어텐션 매트릭스를 연산하여 타겟 어텐션 매트릭스를 각각 출력하는 타겟 어텐션 매트릭스 출력부(1020a, 1020b)를 포함하는 것을 특징으로 한다.

[0052] 다음으로, 상기 학습부(1100a, 1100b)는 각 소스 어텐션 매트릭스와 각 타겟 어텐션 매트릭스 내 쌍(pair)의 유사도를 연산하고, 손실함수를 이용하여 각 최종 어텐션 매트릭스의 손실값을 줄이는 방향으로 학습하는 것을 특징으로 한다.

[0053] 일반적으로, 어텐션 연산에는 쿼리(Query), 키(key) 및 값(Value) 3가지 벡터가 사용될 수 있다. 쿼리(Query)는 현재 시점의 토큰이고, 키(key) 및 값(Value)은 어텐션을 구하고자 하는 대상 토큰이다. 다시 말하면, 쿼리(Query)는 고정되어 있는 하나의 토큰이고, 해당 쿼리(Query)와 가장 부합하는, 즉 어텐션이 가장 높은 토큰을

찾기 위해서 키(key) 및 값(Value)을 처음부터 끝까지 탐색시키는 것이다. 쿼리(Query), 키(key) 및 값(Value)은 동일한 크기를 갖는다.

- [0054] 도 4의 일실시예를 보면, 하나의 쿼리 어텐션 모듈(1000a)은 $(n \times d_k)$ 크기를 갖는 쿼리(Query)를 이용하여 행과 열이 변환되어 $(d_k \times n)$ 크기를 갖는 키(key)와 곱산할 수 있다. 그리고 하나의 쿼리 어텐션 모듈(1000a)은 소프트맥스(softmax) 함수를 이용하여 앞서 곱산한 값을 0 내지 1 사이의 정규화된 값으로 변환할 수 있다. 이때, 소프트맥스(Softmax) 함수로 정규화된 값을 모두 더하면 1이 된다.
- [0055] 즉, 하나의 쿼리 어텐션 모듈(1000a)로부터 출력된 $(n \times n)$ 크기를 갖는 하나의 임시 어텐션 매트릭스의 각 픽셀은 전체 확률에서 특정 클래스에 속할 확률을 갖고, 0 내지 1 사이의 정규화된 값을 갖는다. 하나의 임시 어텐션 매트릭스는 각 픽셀이 속한 클래스마다 서로 다른 색으로 표시될 수 있고, 동일한 클래스에서 높은 값을 가질수록 진한 색으로 표시될 수 있고, 동일한 클래스에서 낮은 값을 가질수록 옅은 색으로 표시될 수 있다. 여기서, 클래스는 긍정(Positive) 및 부정(Negative)을 포함할 수 있다.
- [0056] 또한, 하나의 쿼리 어텐션 모듈(1000a)은 $(n \times n)$ 크기를 갖는 하나의 임시 어텐션 매트릭스를 엔트로피를 기반으로 정렬할 수 있고, 선택된 행을 갖는 하나의 최종 어텐션 매트릭스를 출력할 수 있다. 예컨대, 선택된 행이 d_k 개일 수 있다. 즉, 하나의 최종 어텐션 매트릭스는 $(d_k \times n)$ 크기를 가질 수 있다.
- [0057] 또한, 하나의 소스 어텐션 매트릭스 출력부(1010a)는 하나의 리파인 특징(Refined_F_{x1})에 대한 $(n \times d_k)$ 크기를 갖는 값(Value)과 $(d_k \times n)$ 크기를 갖는 최종 어텐션 매트릭스를 곱산하여 $(n \times n)$ 크기를 갖는 하나의 소스 어텐션 매트릭스(Source attention matrix)를 출력할 수 있다.
- [0058] 또한, 하나의 타겟 어텐션 매트릭스 출력부(1020a)는 제1 타겟 이미지(G_y)로부터 출력된 제1 타겟 특징(F_{y1})에 대한 $(n \times d_k)$ 크기를 갖는 값(Value)과 $(d_k \times n)$ 크기를 갖는 하나의 최종 어텐션 매트릭스를 곱산하여 $(n \times n)$ 크기를 갖는 하나의 타겟 어텐션 매트릭스(Target attention matrix)를 출력할 수 있다.
- [0059] 또한, 하나의 학습부(1110a)는 $(n \times n)$ 크기를 갖는 하나의 소스 어텐션 매트릭스(Source attention matrix)와 $(n \times n)$ 크기를 갖는 하나의 타겟 어텐션 매트릭스(Target attention matrix) 내 픽셀의 쌍(pair)의 유사도를 연산할 수 있다. 그리고 상기 학습부(1110a)는 대조 학습(Contrastive learning) 기반으로 하고, 손실함수(loss function(x))를 이용하여 하나의 최종 어텐션 매트릭스의 손실값을 줄이는 방향으로 학습할 수 있다.
- [0060] 도 5의 일실시예를 보면, 또 다른 하나의 쿼리 어텐션 모듈(1000b)은 $(n \times d_k)$ 크기를 갖는 쿼리(Query)를 이용하여 $(d_k \times n)$ 크기를 갖는 키(key)와 곱산할 수 있다. 그리고 또 다른 하나의 쿼리 어텐션 모듈(1000b)은 소프트맥스(softmax) 함수를 이용하여 앞서 곱산한 값을 0 내지 1 사이의 정규화된 값으로 변환할 수 있다. 이때, 소프트맥스(Softmax) 함수로 정규화된 값을 모두 더하면 1이 된다.
- [0061] 즉, 또 다른 하나의 쿼리 어텐션 모듈(1000b)로부터 출력된 $(n \times n)$ 크기를 갖는 또 다른 하나의 임시 어텐션 매트릭스의 각 픽셀은 전체 확률에서 특정 클래스에 속할 확률을 갖고, 0 내지 1 사이의 정규화된 값을 갖는다. 또 다른 하나의 임시 어텐션 매트릭스는 각 픽셀이 속한 클래스마다 서로 다른 색으로 표시될 수 있고, 동일한 클래스에서 높은 값을 가질수록 진한 색으로 표시될 수 있고, 동일한 클래스에서 낮은 값을 가질수록 옅은 색으로 표시될 수 있다. 여기서, 클래스는 긍정(Positive) 및 부정(Negative)을 포함할 수 있다.
- [0062] 또한, 또 다른 하나의 쿼리 어텐션 모듈(1000b)은 $(n \times n)$ 크기를 갖는 또 다른 하나의 임시 어텐션 매트릭스를 엔트로피를 기반으로 정렬할 수 있고, 선택된 행을 갖는 또 다른 하나의 최종 어텐션 매트릭스를 출력할 수 있다. 예컨대, 선택된 행이 d_k 개일 수 있다. 즉, 또 다른 하나의 최종 어텐션 매트릭스는 $(d_k \times n)$ 크기를 가질 수 있다.
- [0063] 또한, 또 다른 하나의 소스 어텐션 매트릭스 출력부(1010b)는 또 다른 하나의 리파인 특징(Refined_F_{y2})에 대한 $(n \times d_k)$ 크기를 갖는 값(Value)과 $(d_k \times n)$ 크기를 갖는 또 다른 하나의 최종 어텐션 매트릭스를 곱산하여 $(n \times n)$ 크기를 갖는 또 다른 하나의 소스 어텐션 매트릭스(Source attention matrix)를 출력할 수 있다.
- [0064] 또한, 또 다른 하나의 타겟 어텐션 매트릭스 출력부(1020b)는 제2 타겟 이미지(G_x)로부터 출력된 제2 타겟 특징(F_{x2})에 대한 $(n \times d_k)$ 크기를 갖는 값(Value)과 $(d_k \times n)$ 크기를 갖는 또 다른 하나의 최종 어텐션 매트릭스를 곱

산하여 ($n \times n$) 크기를 갖는 또 다른 하나의 타겟 어텐션 매트릭스(Target attention matrix)를 출력할 수 있다.

[0065] 또한, 또 다른 하나의 학습부(1110b)는 또 다른 하나의 소스 어텐션 매트릭스(Source attention matrix)와 또 다른 하나의 타겟 어텐션 매트릭스(Target attention matrix) 내 픽셀의 쌍(pair)의 유사도를 연산할 수 있다. 그리고 상기 학습부(1110b)는 대조 학습(Contrastive learning) 기반으로 하고, 손실함수(loss function(y))를 이용하여 또 다른 하나의 최종 어텐션 매트릭스의 손실값을 줄이는 방향으로 학습할 수 있다.

[0066] 다시 도 1을 보면, 본 발명은 제1 생성자(100) 및 제1 판별자(200)를 포함하는 생성적 적대 신경망(GAN) 1개와 제2 생성자(300) 및 제2 판별자(400)를 포함하는 생성적 적대 신경망(GAN) 1개가 대칭적으로 배치될 수 있다.

[0067] 여기서, 상기 제1 생성자(100)는 상기 제1 소스 이미지(I_x)로부터 상기 제2 도메인(y)을 포함하는 제1 타겟 이미지(G_y)를 생성할 수 있다. 상기 제1 생성자(100)는 선단에 인코딩(Encoding) 하는 제1 생성자 인코더(110)와 말단에 디코딩(Decoding) 하는 제1 생성자 디코더(120)를 포함할 수 있다.

[0068] 다음으로, 상기 제2 생성자(300)는 상기 제2 소스 이미지(I_y)로부터 상기 제1 도메인(x)을 포함하는 제2 타겟 이미지(G_x)를 생성할 수 있다. 상기 제2 생성자(300)의 선단에 인코딩(Encoding) 하는 제2 생성자 인코더(310)와 말단에 디코딩(Decoding) 하는 제2 생성자 디코더(320)를 포함할 수 있다.

[0069] 다음으로, 상기 제1 판별자(200)는 상기 제1 소스 이미지(I_x)와 상기 제2 타겟 이미지(G_x) 간의 진위여부를 판별할 수 있다. 즉, 상기 제1 판별자(200)는 상기 제2 생성자(300)로부터 생성된 제2 타겟 이미지(G_x)가 진짜(Real)인지 가짜(Fake)인지 판단하는 것이다. 그리고 상기 제2 생성자(300)는 상기 제1 판별자(200)가 진짜(Real)인지 가짜(Fake)인지 구분하지 못하는 제2 타겟 이미지(G_x)를 생성하는 것을 학습 목표로 한다.

[0070] 다음으로, 상기 제2 판별자(400)는 상기 제2 소스 이미지(I_y)와 상기 제1 타겟 이미지(G_y) 간의 진위여부를 판별할 수 있다. 즉, 상기 제2 판별자(400)는 상기 제1 생성자(100)로부터 생성된 제1 타겟 이미지(G_y)가 진짜(Real)인지 가짜(Fake)인지 판단하는 것이다. 그리고 상기 제1 생성자(100)는 상기 제2 판별자(400)가 진짜(Real)인지 가짜(Fake)인지 구분하지 못하는 제1 타겟 이미지(G_y)를 생성하는 것을 학습 목표로 한다.

[0072] 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 방법

[0073] 우선, 본 발명의 일실시예에 따르면, 본 발명의 이중 학습이 가능한 생성적 적대 신경망(GAN) 기반 이미지 생성 방법은 제1 생성자(100)에 의하여, 상기 제1 소스 이미지(I_x)로부터 제2 도메인(y)을 포함하는 제1 타겟 이미지(G_y)가 생성되는 제1 타겟 이미지 생성단계(S100)와, 제2 생성자(300)에 의하여, 상기 제2 소스 이미지(I_y)로부터 제1 도메인(x)을 포함하는 제2 타겟 이미지(G_x)가 생성되는 제2 타겟 이미지 생성단계(S200)를 포함할 수 있다.

[0074] 또한, 본 발명의 일실시예에 따르면 본 발명은 제1 판별자(200)에 의하여, 상기 제1 소스 이미지(I_x)와 상기 제2 타겟 이미지(G_x) 간의 진위여부가 판별되는 제1 판별단계(S1400)와, 제2 판별자(400)에 의하여, 상기 제2 소스 이미지(I_y)와 상기 제1 타겟 이미지(G_y) 간의 진위여부가 판별되는 제2 판별단계(S1500)를 더 포함할 수 있다.

[0075] 즉, 상기 제1 판별단계(S1400)는 상기 제2 생성자(300)로부터 생성된 제2 타겟 이미지(G_x)가 진짜(Real)인지 가짜(Fake)인지 판단되는 것이다. 그리고 본 발명은 상기 제1 판별단계(S1400)로부터 상기 제1 판별자(200)가 진짜(Real)인지 가짜(Fake)인지 구분하지 못하는 제2 타겟 이미지(G_x)가 생성되도록 하는 것을 학습 목표로 한다.

[0076] 그리고 상기 제2 판별단계(S1500)는 상기 제1 생성자(100)로부터 생성된 제1 타겟 이미지(G_y)가 진짜(Real)인지 가짜(Fake)인지 판단되는 것이다. 그리고 본 발명은 상기 제2 판별단계(S1500)로부터 상기 제2 판별자(400)가 진짜(Real)인지 가짜(Fake)인지 구분하지 못하는 제1 타겟 이미지(G_y)가 생성되도록 하는 것을 학습 목표로 한다.

[0078] (1) 제1 소스 이미지(I_x)와 제1 타겟 이미지(G_y)에 대한 학습

- [0079] 도 6의 일실시예를 보면, 상기 제1 소스 이미지(I_x)와 상기 제1 타겟 이미지(G_y)에 대해서, 본 발명은 제1 소스 인코더(500)에 의하여, 제1 소스 이미지(I_x)의 제1 도메인(x)으로부터 제1 소스 특징($F_{x,1}$)이 추출되는 제1 소스 특징 추출단계(S300)와, 제1 타겟 인코더(600)에 의하여, 제1 타겟 이미지(G_y)의 제2 도메인(y)으로부터 제1 타겟 특징($F_{y,1}$)이 추출되는 제1 타겟 특징 추출단계(S400)를 포함한다.
- [0080] 다음으로, 본 발명은 컨볼루션 블록 어텐션 모듈(900a, 900b)에 의하여, 각 소스 인코더(500, 700)로부터 출력된 각 소스 특징($F_{x,1}$, $F_{y,2}$)으로부터 각 소스 이미지(I_x , I_y)에서 어디에 초점을 맞춰야 하는지를 나타내는 리파인 특징(Refined $F_{x,1}$, Refined $F_{y,2}$)이 각각 추출되는 리파인 특징 추출단계(S700a, S700b)를 더 포함한다.
- [0081] 일실시예에 따른 리파인 특징 추출단계(S700a)는 제1 소스 특징($F_{x,1}$)으로부터 제1 소스 이미지(I_x)에서 어디에 초점을 맞춰야 하는지를 나타내는 하나의 리파인 특징(Refined $F_{x,1}$)이 추출될 수 있다. 여기서, 상기 제1 소스 특징($F_{x,1}$)은 높이(H)*폭(W)*채널(C) 크기를 가질 수 있다.
- [0082] 일실시예에 따른 리파인 특징 추출단계(S700a)는 하나의 컨볼루션 블록 어텐션 모듈(900a) 내 채널 어텐션 모듈(910a)에 의하여, 상기 제1 소스 특징($F_{x,1}$)의 채널 간의 관계를 살펴 어떤 채널에 더 초점을 맞춰야 하는지 인코딩되는 채널 어텐션 단계(S710a)를 포함하는 것을 특징으로 한다. 이는, 상기 [수학식 1]과 같다.
- [0083] 우선 일실시예에 따른 채널 어텐션 단계(S710a)는 최대 풀링(Max pooling)을 통해서 제1 소스 특징($F_{x,1}$)에서 채널별 최대값만이 벡터로 출력될 수 있다. 다음으로 일실시예에 따른 채널 어텐션 단계(S710a)는 최대 풀링(Max pooling)된 벡터들이 2개의 층(layer)을 갖는 다중 퍼셉트론(Multi-Layer Perceptron; MLP)을 통해서 차원이 축소될 수 있다. 일실시예에 따른 채널 어텐션 단계(S710a)는 차원이 축소된 벡터들이 시그모이드 함수(Sigmoid function)가 이용되어 확률값으로 변환될 수 있다. 그리고 일실시예에 따른 채널 어텐션 단계(S710a)는 채널별 확률값 중에서 가장 큰 확률값을 갖는 채널에 초점을 맞춰 상기 제1 소스 특징($F_{x,1}$)이 인코딩될 수 있다.
- [0084] 이에 따라, 일실시예에 따른 채널 어텐션 단계(S710a)는 상기 제1 소스 특징($F_{x,1}$)과 하나의 채널 어텐션 모듈(910a)로부터 채널이 인코딩된 제1 소스 특징이 곱산되어 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)이 출력될 수 있다.
- [0085] 또한, 일실시예에 따른 리파인 특징 추출단계(S700a)는 하나의 컨볼루션 블록 어텐션 모듈(900a) 내 공간 어텐션 모듈(920a)에 의하여, 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)의 H*W*1 크기에서 다수 개의 픽셀 중 위치관계를 살펴 어떤 위치에 더 초점을 맞춰야 하는지 인코딩되는 공간 어텐션 단계(S720a)를 더 포함하는 것을 특징으로 한다. 이는 상기 [수학식 2]와 같다.
- [0086] 우선 일실시예에 따른 공간 어텐션 단계(S720a)는 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)이 입력될 수 있다. 일실시예에 따른 공간 어텐션 단계(S720a)는 H*W*1 크기를 대상으로 하는 최대 풀링(Max pooling)을 통해서 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)에서 최대값만이 벡터로 출력될 수 있다. 그리고 일실시예에 따른 공간 어텐션 단계(S720a)는 H*W*1 크기를 갖는 두 채널이 연결되어 H*W*2 크기의 특징맵이 생성될 수 있다. 일실시예에 따른 공간 어텐션 단계(S720a)는 특징맵이 2채널이므로 7*7 컨볼루션 연산이 수행될 수 있다. 일실시예에 따른 공간 어텐션 단계(S720a)는 시그모이드 함수(Sigmoid function)가 이용되어 7*7 컨볼루션 연산된 특징맵이 확률값으로 변환될 수 있다. 그리고 일실시예에 따른 공간 어텐션 단계(S720a)는 확률값 중에서 가장 큰 확률값을 갖는 위치에 초점을 맞춰 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)이 인코딩될 수 있다.
- [0087] 이에 따라, 일실시예에 따른 공간 어텐션 단계(S720a)는 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)과 하나의 공간 어텐션 모듈(920a)로부터 위치가 인코딩된 채널 정제된 제1 소스 특징(Channel_refined $F_{x,1}$)이 곱산되어 하나의 리파인 특징(Refined $F_{x,1}$)이 출력될 수 있다.
- [0088] 다음으로, 본 발명은 쿼리 어텐션 모듈(1000a, 1000b)에 의하여, 주어진 쿼리(Query)가 이용되어 각 리파인 특징(Refined $F_{x,1}$, $F_{y,2}$)에 대한 임시 어텐션 매트릭스가 각각 생성되는 임시 어텐션 매트릭스 생성단계(S800a, S800b) 및 상기 쿼리 어텐션 모듈(1000a, 1000b)에 의하여, 엔트로피를 기반으로 각 임시 어텐션 매트릭스가 정

렬된 후 선택된 행을 갖는 최종 어텐션 매트릭스가 각각 출력되는 최종 어텐션 매트릭스 출력단계(S900a, S900b)를 더 포함한다.

[0089] 일실시예에 따른 임시 어텐션 매트릭스 생성단계(S800a)는 $(n \times d_k)$ 크기를 갖는 쿼리(Query)가 이용되어 $(d_k \times n)$ 크기를 갖는 키(key)와 곱산될 수 있다. 그리고 일실시예에 따른 임시 어텐션 매트릭스 생성단계(S800a)는 소프트맥스(softmax) 함수가 이용되어 앞서 곱산된 값이 0 내지 1 사이의 정규화된 값으로 변환될 수 있다. 이때, 소프트맥스(Softmax) 함수로 정규화된 값을 모두 더하면 1이 된다.

[0090] 즉, 하나의 쿼리 어텐션 모듈(1000a)로부터 출력된 $(n \times n)$ 크기를 갖는 하나의 임시 어텐션 매트릭스의 각 픽셀은 전체 확률에서 특정 클래스에 속할 확률을 갖고, 0 내지 1 사이의 정규화된 값을 갖는다. 하나의 임시 어텐션 매트릭스는 각 픽셀이 속한 클래스마다 서로 다른 색으로 표시될 수 있고, 동일한 클래스에서 높은 값을 가질수록 진한 색으로 표시될 수 있고, 동일한 클래스에서 낮은 값을 가질수록 옅은 색으로 표시될 수 있다. 여기서, 클래스는 긍정(Positive) 및 부정(Negative)을 포함할 수 있다.

[0091] 또한, 일실시예에 따른 최종 어텐션 매트릭스 생성단계(S900a)는 $(n \times n)$ 크기를 갖는 하나의 임시 어텐션 매트릭스가 엔트로피를 기반으로 정렬될 수 있고, 선택된 행을 갖는 하나의 최종 어텐션 매트릭스가 출력될 수 있다. 예컨대, 선택된 행이 d_k 개일 수 있다. 즉, 하나의 최종 어텐션 매트릭스는 $(d_k \times n)$ 크기를 가질 수 있다.

[0092] 다음으로, 본 발명은 상기 쿼리 어텐션 모듈(1000a, 1000b)에 의하여, 각 리파인 특징(Refined F_{x1} , Refined F_{y2})의 값(Value)과 각 최종 어텐션 매트릭스가 연산되어 소스 어텐션 매트릭스가 각각 출력되는 소스 어텐션 매트릭스 출력단계(S1000a, 1000b) 및 상기 쿼리 어텐션 모듈(1000a, 1000b)에 의하여, 각 타겟 특징(F_{y1} , F_{x2})의 값(Value)과 각 최종 어텐션 매트릭스가 연산되어 타겟 어텐션 매트릭스가 각각 출력되는 타겟 어텐션 매트릭스 출력단계(S1100a, S1100b)를 더 포함할 수 있다.

[0093] 일실시예에 따른 소스 어텐션 매트릭스 출력단계(S1000a)는 하나의 소스 어텐션 매트릭스 출력부(1010a)에 의하여, 하나의 리파인 특징(Refined F_{x1})에 대한 $(n \times d_k)$ 크기를 갖는 값(Value)과 $(d_k \times n)$ 크기를 갖는 최종 어텐션 매트릭스가 곱산되어 $(n \times n)$ 크기를 갖는 하나의 소스 어텐션 매트릭스(Source attention matrix)가 출력될 수 있다.

[0094] 또한, 일실시예에 따른 타겟 어텐션 매트릭스 출력단계(S1100a)는 하나의 타겟 어텐션 매트릭스 출력부(1020a)에 의하여, 제1 타겟 이미지(G_y)로부터 출력된 제1 타겟 특징(F_{y1})에 대한 $(n \times d_k)$ 크기를 갖는 값(Value)과 $(d_k \times n)$ 크기를 갖는 하나의 최종 어텐션 매트릭스가 곱산되어 $(n \times n)$ 크기를 갖는 하나의 타겟 어텐션 매트릭스(Target attention matrix)가 출력될 수 있다.

[0095] 다음으로, 본 발명은 학습부(1100a, 1100b)에 의하여, 각 소스 어텐션 매트릭스와 각 타겟 어텐션 매트릭스 내 쌍(pair)의 유사도가 연산되는 유사도 연산단계(S1200a, 1200b) 및 상기 학습부(1100a, 1100b)에 의하여, 손실 함수가 이용되어 긍정적인 쌍(Positive pair)의 유사도가 작아지도록 학습하고, 부정적인 쌍(Negative pair)의 유사도가 높아지도록 학습하는 학습단계(S1300a, S1300b)를 더 포함할 수 있다.

[0096] 일실시예에 따른 유사도 연산단계(S1200a)는 하나의 소스 어텐션 매트릭스와 하나의 타겟 어텐션 매트릭스 내 쌍(pair)의 유사도가 연산될 수 있다. 그리고 일실시예에 따른 학습단계(S1300a)는 손실 함수가 이용되어 긍정적인 쌍(Positive pair)의 유사도가 작아지도록 학습하고, 부정적인 쌍(Negative pair)의 유사도가 높아지도록 학습할 수 있다.

[0098] (2) 제2 소스 이미지(I_y)와 제2 타겟 이미지(G_x)에 대한 학습

[0099] 도 7의 일실시예를 보면, 상기 제2 소스 이미지(I_y)와 상기 제2 타겟 이미지(G_x)에 대해서, 본 발명은 제2 소스 인코더(700)에 의하여, 제2 소스 이미지(I_y)의 제2 도메인(y)으로부터 제2 소스 특징(F_{y2})이 추출되는 제2 소스 특징 추출단계(S500), 제2 타겟 인코더(800)에 의하여, 제2 타겟 이미지(G_x)의 제1 도메인(x)으로부터 제2 타겟 특징(F_{x2})이 추출되는 제2 타겟 특징 추출단계(S600)를 포함한다.

[0100] 또 다른 일실시예에 따른 리파인 특징 추출단계(S700b)는 제2 소스 특징(F_{y2})으로부터 제2 소스 이미지(I_x)에서 어디에 초점을 맞춰야 하는지를 나타내는 또 다른 하나의 리파인 특징(Refined F_{y2})이 추출될 수 있다. 여기서,

상기 제2 소스 특징($F_{y,2}$)은 높이(H)*폭(W)*채널(C) 크기를 가질 수 있다.

- [0101] 또 다른 일실시예에 따른 리파인 특징 추출단계(S700b)는 또 다른 하나의 컨볼루션 블록 어텐션 모듈(900b) 내 채널 어텐션 모듈(910b)에 의하여, 상기 제2 소스 특징($F_{y,2}$)의 채널 간의 관계를 살펴 어떤 채널에 더 초점을 맞춰야 하는지 인코딩되는 채널 어텐션 단계(S710b)를 포함하는 것을 특징으로 한다. 이는, 상기 [수학식 1]과 같다.
- [0102] 우선 또 다른 일실시예에 따른 채널 어텐션 단계(S710b)는 최대 풀링(Max pooling)을 통해서 제2 소스 특징($F_{y,2}$)에서 채널별 최대값만이 벡터로 출력될 수 있다. 다음으로 또 다른 일실시예에 따른 채널 어텐션 단계(S710b)는 최대 풀링(Max pooling)된 벡터들이 2개의 층(layer)을 갖는 다중 퍼셉트론(Multi-Layer Perceptron; MLP)을 통해서 차원이 축소될 수 있다. 또 다른 일실시예에 따른 채널 어텐션 단계(S710b)는 차원이 축소된 벡터들이 시그모이드 함수(Sigmoid function)가 이용되어 확률값으로 변환될 수 있다. 그리고 또 다른 일실시예에 따른 채널 어텐션 단계(S710b)는 채널별 확률값 중에서 가장 큰 확률값을 갖는 채널에 초점을 맞춰 상기 제2 소스 특징($F_{y,2}$)이 인코딩될 수 있다.
- [0103] 이에 따라, 또 다른 일실시예에 따른 채널 어텐션 단계(S710b)는 상기 제2 소스 특징($F_{y,2}$)과 또 다른 하나의 채널 어텐션 모듈(910b)로부터 채널이 인코딩된 제2 소스 특징이 곱산되어 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)이 출력될 수 있다.
- [0104] 또한, 또 다른 일실시예에 따른 리파인 특징 추출단계(S700b)는 하나의 컨볼루션 블록 어텐션 모듈(900b) 내 공간 어텐션 모듈(920b)에 의하여, 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)의 $H*W*1$ 크기에서 다수 개의 픽셀 중 위치관계를 살펴 어떤 위치에 더 초점을 맞춰야 하는지 인코딩되는 공간 어텐션 단계(S720b)를 더 포함하는 것을 특징으로 한다. 이는 상기 [수학식 2]와 같다.
- [0105] 우선 또 다른 일실시예에 따른 공간 어텐션 단계(S720b)는 채널 정제된 제1 소스 특징(Channel_refined $F_{y,2}$)이 입력될 수 있다. 또 다른 일실시예에 따른 공간 어텐션 단계(S720b)는 $H*W*1$ 크기를 대상으로 하는 최대 풀링(Max pooling)을 통해서 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)에서 최대값만이 벡터로 출력될 수 있다. 그리고 또 다른 일실시예에 따른 공간 어텐션 단계(S720b)는 $H*W*1$ 크기를 갖는 두 채널이 연결되어 $H*W*2$ 크기의 특징맵이 생성될 수 있다. 또 다른 일실시예에 따른 공간 어텐션 단계(S720b)는 특징맵이 2채널이므로 $7*7$ 컨볼루션 연산이 수행될 수 있다. 또 다른 일실시예에 따른 공간 어텐션 단계(S720b)는 시그모이드 함수(Sigmoid function)가 이용되어 $7*7$ 컨볼루션 연산된 특징맵이 확률값으로 변환될 수 있다. 그리고 또 다른 일실시예에 따른 공간 어텐션 단계(S720b)는 확률값 중에서 가장 큰 확률값을 갖는 위치에 초점을 맞춰 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)이 인코딩될 수 있다.
- [0106] 이에 따라, 또 다른 일실시예에 따른 공간 어텐션 단계(S720b)는 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)과 또 다른 하나의 공간 어텐션 모듈(920b)로부터 위치가 인코딩된 채널 정제된 제2 소스 특징(Channel_refined $F_{y,2}$)이 곱산되어 또 다른 하나의 리파인 특징(Refined $F_{y,2}$)이 출력될 수 있다.
- [0107] 다음으로, 또 다른 일실시예에 따른 임시 어텐션 매트릭스 생성단계(S800b)는 $(n \times d_k)$ 크기를 갖는 쿼리(Query)가 이용되어 $(d_k \times n)$ 크기를 갖는 키(key)와 곱산될 수 있다. 그리고 또 다른 일실시예에 따른 임시 어텐션 매트릭스 생성단계(S800b)는 소프트맥스(softmax) 함수가 이용되어 앞서 곱산된 값이 0 내지 1 사이의 정규화된 값으로 변환될 수 있다. 이때, 소프트맥스(Softmax) 함수로 정규화된 값을 모두 더하면 1이 된다.
- [0108] 즉, 하나의 쿼리 어텐션 모듈(1000b)로부터 출력된 $(n \times n)$ 크기를 갖는 또 다른 하나의 임시 어텐션 매트릭스의 각 픽셀은 전체 확률에서 특정 클래스에 속할 확률을 갖고, 0 내지 1 사이의 정규화된 값을 갖는다. 또 다른 하나의 임시 어텐션 매트릭스는 각 픽셀이 속한 클래스마다 서로 다른 색으로 표시될 수 있고, 동일한 클래스에서 높은 값을 가질수록 진한 색으로 표시될 수 있고, 동일한 클래스에서 낮은 값을 가질수록 옅은 색으로 표시될 수 있다. 여기서, 클래스는 긍정(Positive) 및 부정(Negative)을 포함할 수 있다.
- [0109] 다음으로, 또 다른 일실시예에 따른 최종 어텐션 매트릭스 생성단계(S900b)는 $(n \times n)$ 크기를 갖는 또 다른 하나의 임시 어텐션 매트릭스가 엔트로피를 기반으로 정렬될 수 있고, 선택된 행을 갖는 또 다른 하나의 최종 어텐션 매트릭스가 출력될 수 있다. 예컨대, 선택된 행이 d_k 개일 수 있다. 즉, 또 다른 하나의 최종 어텐션 매트릭

스는 ($d_k \times n$) 크기를 가질 수 있다.

- [0110] 다음으로, 또 다른 일실시예에 따른 소스 어텐션 매트릭스 출력단계(S1000b)는 또 다른 하나의 소스 어텐션 매트릭스 출력부(1010b)에 의하여, 또 다른 하나의 리파인 특징(Refined $F_{y,2}$)에 대한 ($n \times d_k$) 크기를 갖는 값(Value)과 ($d_k \times n$) 크기를 갖는 최종 어텐션 매트릭스가 곱산되어 ($n \times n$) 크기를 갖는 또 다른 하나의 소스 어텐션 매트릭스(Source attention matrix)가 출력될 수 있다.
- [0111] 다음으로, 또 다른 일실시예에 따른 타겟 어텐션 매트릭스 출력단계(S1100b)는 또 다른 하나의 타겟 어텐션 매트릭스 출력부(1020b)에 의하여, 제2 타겟 이미지(G_x)로부터 출력된 제2 타겟 특징($F_{x,2}$)에 대한 ($n \times d_k$) 크기를 갖는 값(Value)과 ($d_k \times n$) 크기를 갖는 또 다른 하나의 최종 어텐션 매트릭스가 곱산되어 ($n \times n$) 크기를 갖는 또 다른 하나의 타겟 어텐션 매트릭스(Target attention matrix)가 출력될 수 있다.
- [0112] 다음으로, 또 다른 일실시예에 따른 유사도 연산단계(S1200b)는 또 다른 하나의 소스 어텐션 매트릭스와 또 다른 하나의 타겟 어텐션 매트릭스 내 쌍(pair)의 유사도가 연산될 수 있다. 그리고 또 다른 일실시예에 따른 학습단계(S1300b)는 손실함수가 이용되어 긍정적인 쌍(Positive pair)의 유사도가 작아지도록 학습하고, 부정적인 쌍(Negative pair)의 유사도가 높아지도록 학습할 수 있다.
- [0114] 실시예들은 하드웨어, 소프트웨어, 펌웨어, 미들웨어, 마이크로코드, 하드웨어 기술 언어, 또는 이들의 임의의 조합에 의해 구현될 수 있다. 소프트웨어, 펌웨어, 미들웨어 또는 마이크로코드로 구현되는 경우, 필요한 작업을 수행하는 프로그램 코드 또는 코드 세그먼트들은 컴퓨터 판독 가능 저장 매체에 저장되고 하나 이상의 프로세서에 의해 실행될 수 있다.
- [0115] 그리고 본 명세서에 설명된 주제의 양태들은 컴퓨터에 의해 실행되는 프로그램 모듈 또는 컴포넌트와 같은 컴퓨터 실행 가능 명령어들의 일반적인 맥락에서 설명될 수 있다. 일반적으로, 프로그램 모듈 또는 컴포넌트들은 특정 작업을 수행하거나 특정 데이터 형식을 구현하는 루틴, 프로그램, 객체, 데이터 구조를 포함한다. 본 명세서에 설명된 주제의 양태들은 통신 네트워크를 통해 링크되는 원격 처리 디바이스들에 의해 작업들이 수행되는 분산 컴퓨팅 환경들에서 실시될 수도 있다. 분산 컴퓨팅 환경에서, 프로그램 모듈들은 메모리 저장 디바이스들을 포함하는 로컬 및 원격 컴퓨터 저장 매체에 둘 다에 위치할 수 있다.
- [0116] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 으로 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.
- [0117] 그러므로 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

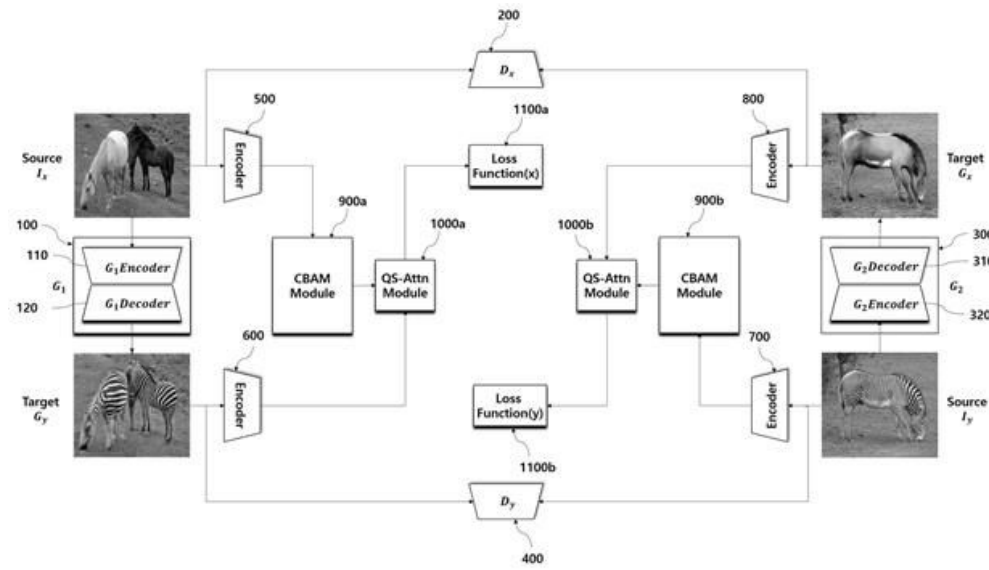
부호의 설명

- [0118] 100.. 제1 생성자
110.. 제1 생성자 인코더
120.. 제1 생성자 디코더
200.. 제1 판별자
300.. 제2 생성자
310.. 제2 생성자 인코더
320.. 제2 생성자 디코더
400.. 제2 판별자
500.. 제1 소스 인코더
600.. 제1 타겟 인코더

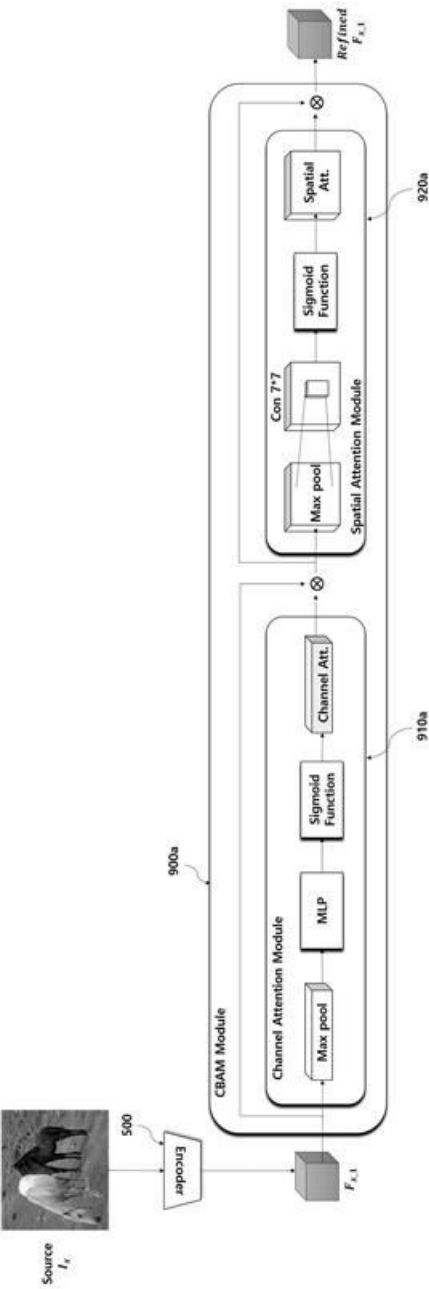
700.. 제2 소스 인코더
 800.. 제2 타겟 인코더
 900a, 900b.. 컨볼루셔널 블록 어텐션 모듈
 910a, 910b.. 채널 어텐션 모듈
 920a, 920b.. 공간 어텐션 모듈
 1000a, 1000b.. 쿼리 어텐션 모듈
 1010a, 1010b.. 소스 어텐션 매트릭스 출력부
 1020a, 1020b.. 타겟 어텐션 매트릭스 출력부
 1100a, 1100b.. 학습부
 S100.. 제1 타겟 이미지 생성단계
 S200.. 제2 타겟 이미지 생성단계
 S300.. 제1 소스 특징 추출단계
 S400.. 제1 타겟 특징 추출단계
 S500.. 제2 소스 특징 추출단계
 S600.. 제2 타겟 특징 추출단계
 S700a, S700b.. 리파인 특징 추출단계
 S710a, S710b.. 채널 어텐션 단계
 S720a, S720b.. 공간 어텐션 단계
 S800a, S800b.. 임시 어텐션 매트릭스 생성단계
 S900a, S900b.. 최종 어텐션 매트릭스 출력단계
 S1000a, S1000b.. 소스 어텐션 매트릭스 출력단계
 S1100a, S1100b.. 타겟 어텐션 매트릭스 출력단계
 S1200a, S1200b.. 유사도 연산단계
 S1300a, S1300b.. 학습단계
 S1400.. 제1 판별단계
 S1500.. 제2 판별단계

도면

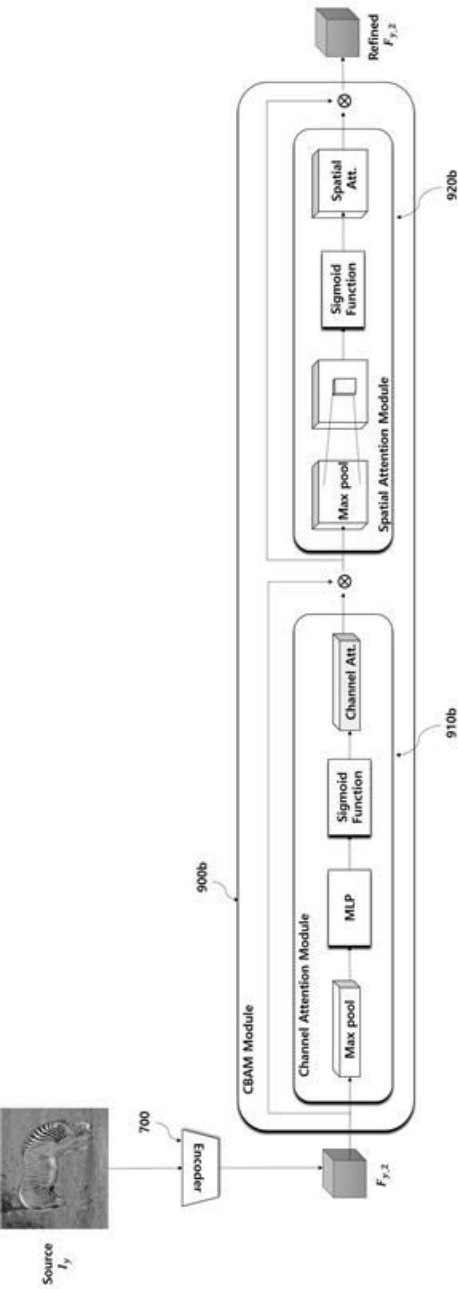
도면1



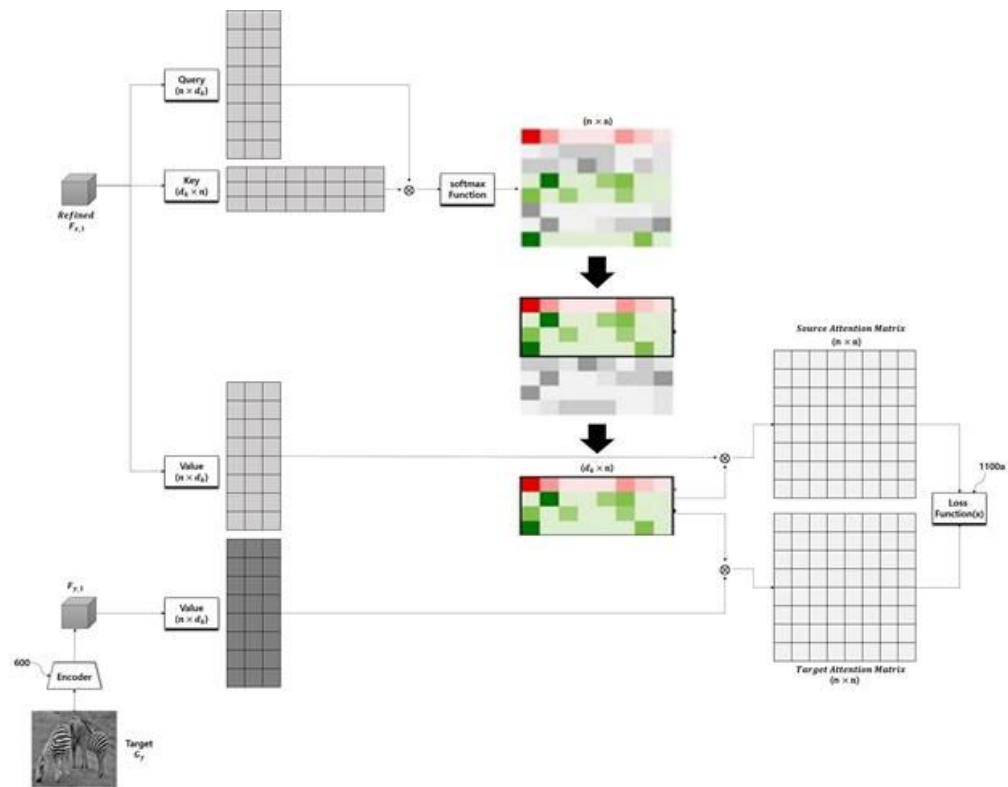
도면2



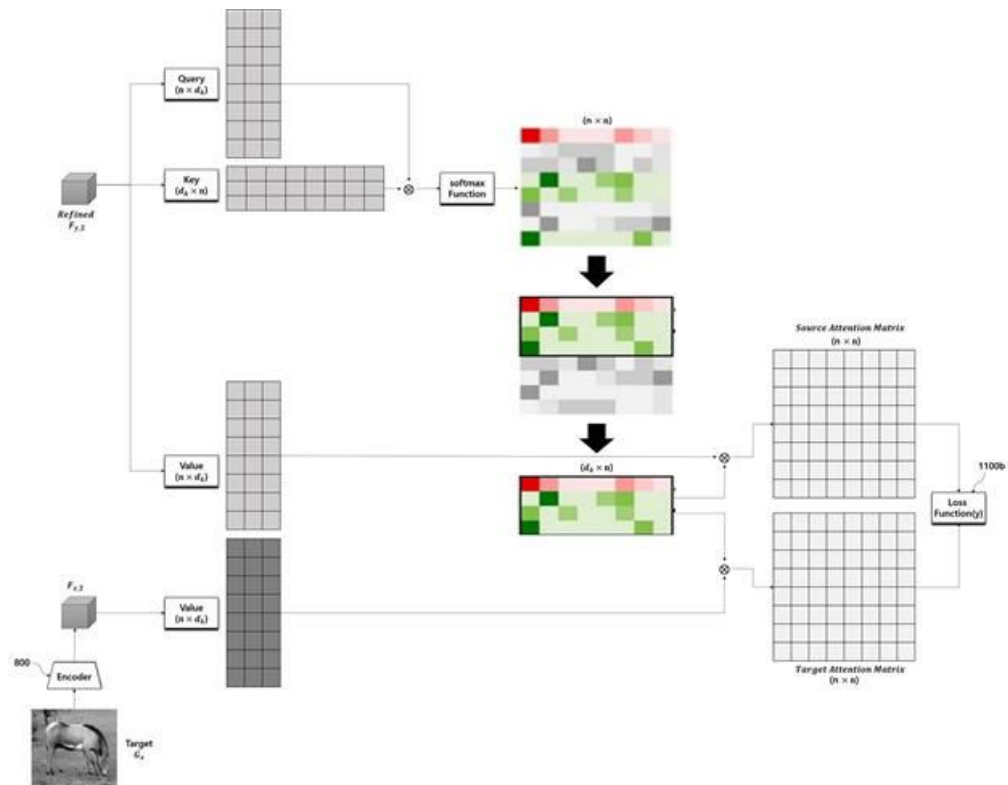
도면3



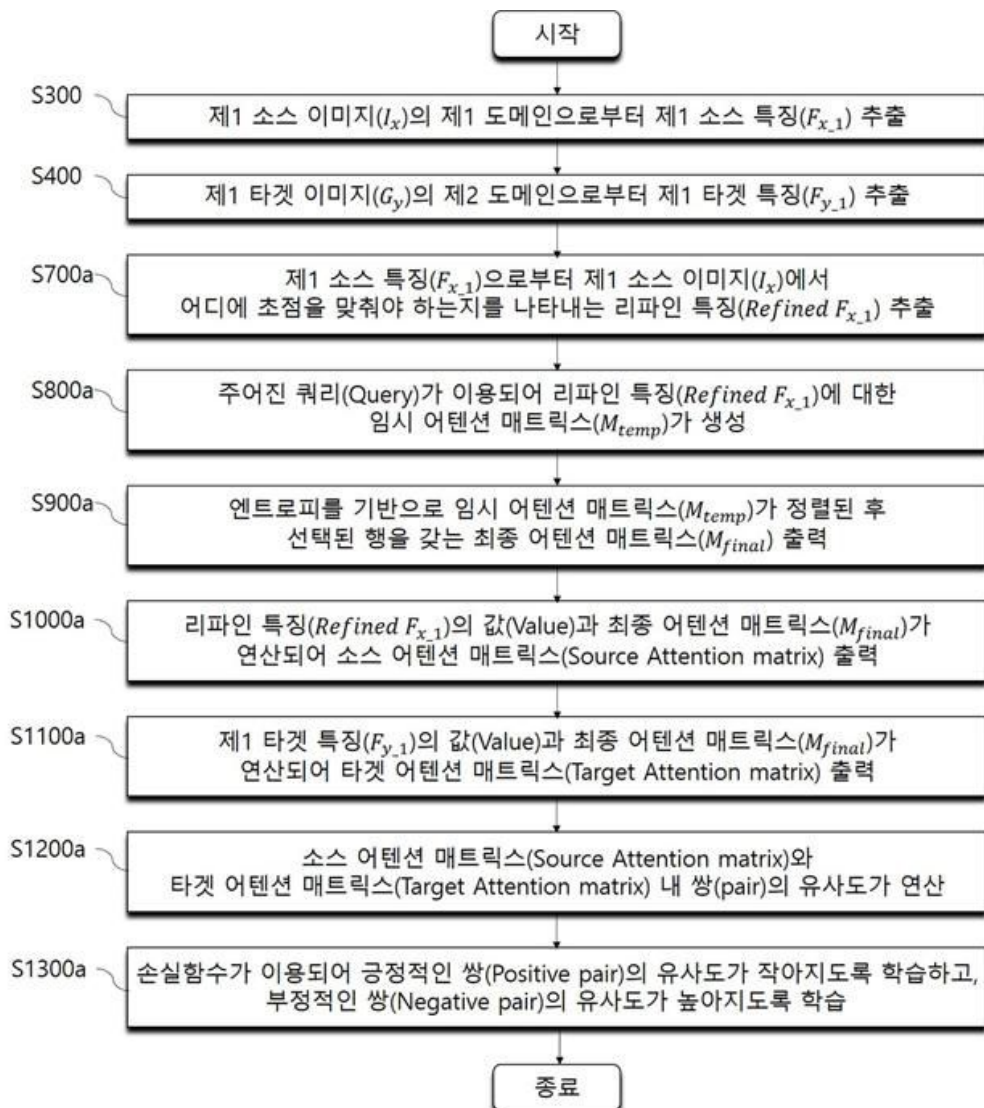
도면4



도면5



도면6



도면7

