



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0146428
(43) 공개일자 2024년10월08일

- (51) 국제특허분류(Int. Cl.)
G06N 3/096 (2023.01) G06N 3/0895 (2023.01)
G06N 3/091 (2023.01) G06N 3/098 (2023.01)
- (52) CPC특허분류
G06N 3/096 (2023.01)
G06N 3/0895 (2023.01)
- (21) 출원번호 10-2023-0041383
(22) 출원일자 2023년03월29일
심사청구일자 2023년03월29일
- (71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
- (72) 발명자
홍성은
서울특별시 성동구 독서당로62길 43, 9동 307호(응봉동, 대림아파트)
- 한경탁
전라남도 여수시 웅천로 261, 201동 1502호 (웅천동, 여수웅천 부영2차)
- (74) 대리인
양성보

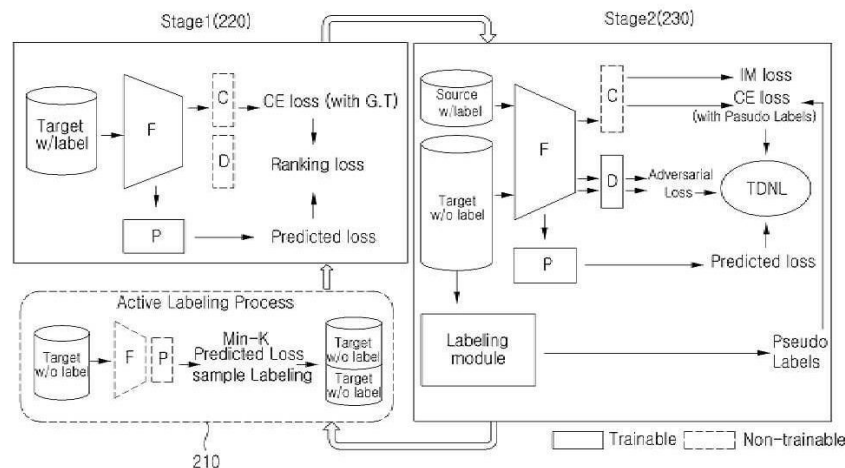
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 전이 가능한 손실 기반 능동적 도메인 적응 방법 및 시스템

(57) 요약

전이 가능한 손실 기반 능동적 도메인 적응 방법 및 시스템이 제시된다. 본 발명에서 제안하는 전이 가능한 손실 기반 능동적 도메인 적응 방법은 쿼리 선택부를 통해 레이블이 지정되지 않은 샘플의 손실을 추론한 후 전이 가능 이중 중첩 손실(Transferable doubly nested loss; TDNL)을 통한 손실 기반 쿼리를 선택하는 단계, 제1 모델 훈련부를 통해 상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 얻은 후, 레이블이 지정된 대상 도메인을 사용하여 모델을 훈련하는 단계 및 제2 모델 훈련부를 통해 레이블이 지정된 대상 샘플로 모델을 순차적으로 훈련한 후 상기 TDNL을 통해 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플로 모델을 훈련하는 단계를 포함한다.

대표도



(52) CPC특허분류

G06N 3/091 (2023.01)

G06N 3/098 (2023.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711179347
과제번호	00155915
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신인재양성(교육운영수익)
연구과제명	[Ezbaro][정부+민간] 인공지능융합혁신인재양성(1차년도)
기 여 율	1/2
과제수행기관명	인하대학교 산학협력단
연구기간	2022.07.01 ~ 2023.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1711186835
과제번호	2022R1A4A1033549
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	기초연구실육성사업(BRL)
연구과제명	[Ezbaro] 인공지능 융합형 메타휴먼 플랫폼 연구실
기 여 율	1/2
과제수행기관명	인하대학교
연구기간	2022.06.01 ~ 2023.02.28

명세서

청구범위

청구항 1

쿼리 선택부를 통해 레이블이 지정되지 않은 샘플의 손실을 추론한 후 전이 가능 이중 중첩 손실(Transferable doubly nested loss; TDNL)을 통한 손실 기반 쿼리를 선택하는 단계;

제1 모델 훈련부를 통해 상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 얻은 후, 레이블이 지정된 대상 도메인을 사용하여 모델을 훈련하는 단계; 및

제2 모델 훈련부를 통해 레이블이 지정된 대상 샘플로 모델을 순차적으로 훈련한 후 상기 TDNL을 통해 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플로 모델을 훈련하는 단계

를 포함하는 능동적 도메인 적응 방법.

청구항 2

제1항에 있어서,

상기 제1 모델 훈련부를 통해 상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 얻은 후, 레이블이 지정된 대상 도메인을 사용하여 모델을 훈련하는 단계는,

상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 사용하여 제1 모델 훈련부의 보조 손실 예측기를 훈련하고, 오라클(oracle)을 사용하여 레이블이 지정된 대상 도메인을 획득한 후, 교차점 손실 값을 사용하여 제1 모델 훈련부의 특징 추출기를 훈련하며,

상기 제1 모델 훈련부의 특징 추출기를 통해 획득한 실제 손실 값을 상기 제1 모델 훈련부의 보조 손실 예측기를 통해 획득한 보조 손실 예측 변수에 대한 레이블로 사용하고, 상기 제1 모델 훈련부의 특징 추출기가 도메인 이동에서 손실을 학습하도록 하여 레이블이 지정되지 않은 대상 샘플의 유사 레이블을 생성하는

능동적 도메인 적응 방법.

청구항 3

제1항에 있어서,

상기 제2 모델 훈련부를 통해 레이블이 지정된 대상 샘플로 모델을 순차적으로 훈련한 후 상기 TDNL을 통해 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플로 모델을 훈련하는 단계는,

레이블이 지정되지 않은 대상 샘플의 대상 샘플의 유사 레이블을 사용하여 제2 모델 훈련부의 특징 추출기를 자체 훈련 방식으로 훈련시키고,

레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플 간의 불일치를 완화시키고, 상기 유사 레이블의 신뢰성을 높이기 위해 셀프-자기 엔트로피 및 클래스에 따른 정보 최대화(Information Maximization; IM) 손실 값을 사용하는

능동적 도메인 적응 방법.

청구항 4

제1항에 있어서,

상기 쿼리 선택부를 통해 레이블이 지정되지 않은 샘플의 손실을 추론한 후 전이 가능 이중 중첩 손실을 통한 손실 기반 쿼리를 선택하는 단계는,

레이블이 지정된 샘플에 대하여 훈련된 보조 손실 예측기를 통해 레이블이 지정되지 않은 샘플의 손실을 추론한 후, 손실이 가장 큰 샘플을 샘플로 선택하는 전이 가능 이중 중첩 손실을 통한 손실 기반 쿼리를 선택하는

능동적 도메인 적응 방법.

청구항 5

레이블이 지정되지 않은 샘플의 손실을 추론한 후 전이 가능 이중 중첩 손실(Transferable doubly nested loss; TDNL)을 통한 손실 기반 쿼리를 선택하는 쿼리 선택부;

상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 얻은 후, 레이블이 지정된 대상 도메인을 사용하여 모델을 훈련하는 제1 모델 훈련부; 및

레이블이 지정된 대상 샘플로 모델을 순차적으로 훈련한 후 상기 TDNL을 통해 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플로 모델을 훈련하는 제2 모델 훈련부

를 포함하는 능동적 도메인 적응 시스템.

청구항 6

제5항에 있어서,

상기 제1 모델 훈련부는,

상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 사용하여 제1 모델 훈련부의 보조 손실 예측기를 훈련하고, 오라클(oracle)을 사용하여 레이블이 지정된 대상 도메인을 획득한 후, 교차점 손실 값을 사용하여 제1 모델 훈련부의 특징 추출기를 훈련하며,

상기 제1 모델 훈련부의 특징 추출기를 통해 획득한 실제 손실 값을 상기 제1 모델 훈련부의 보조 손실 예측기를 통해 획득한 보조 손실 예측 변수에 대한 레이블로 사용하고, 상기 제1 모델 훈련부의 특징 추출기가 도메인 이동에서 손실을 학습하도록 하여 레이블이 지정되지 않은 대상 샘플의 유사 레이블을 생성하는

능동적 도메인 적응 시스템.

청구항 7

제5항에 있어서,

상기 제2 모델 훈련부는,

레이블이 지정되지 않은 대상 샘플의 대상 샘플의 유사 레이블을 사용하여 제2 모델 훈련부의 특징 추출기를 자체 훈련 방식으로 훈련시키고,

레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플 간의 불일치를 완화시키고, 상기 유사 레이블의 신뢰성을 높이기 위해 셀프-자기 엔트로피 및 클래스에 따른 정보 최대화(Information Maximization; IM) 손실 값을 사용하는

능동적 도메인 적응 시스템.

청구항 8

제5항에 있어서,

상기 쿼리 선택부는,

레이블이 지정된 샘플에 대하여 훈련된 보조 손실 예측기를 통해 레이블이 지정되지 않은 샘플의 손실을 추론한 후, 손실이 가장 큰 샘플을 샘플로 선택하는 전이 가능 이중 중첩 손실을 통한 손실 기반 쿼리를 선택하는

능동적 도메인 적응 시스템.

발명의 설명**기술 분야**

[0001]

본 발명은 전이 가능한 손실 기반 능동적 도메인 적응 방법 및 시스템에 관한 것이다.

배경 기술

[0002] 비지도 도메인 적응(Unsupervised Domain Adaptation; UDA)은 대상 도메인 레이블이 필요하지 않다는 장점으로 많은 관심을 받았다. UDA는 다양한 분야에서 유망한 결과를 보여주었지만, UDA와 지도 설정(즉, 대상 레이블 사용) 사이의 성능 격차는 상당하다. 이 문제를 완화하기 위해 몇 가지 대상 샘플의 레이블에 액세스할 수 있는 준지도 도메인 적응이 제안되었다. 준지도 설정에서 레이블이 지정할 대상 샘플은 무작위로 선택되거나 모든 클래스에 걸쳐 균등한 비율로 샘플링된다. 그러나 랜덤 레이블 지정은 제한된 주식 예산을 가진 특정 기준에 기초한 샘플링보다 효율성이 떨어진다. 또한, 대상 샘플의 레이블을 알 수 없기 때문에 클래스에 걸쳐 균일한 샘플링을 적용하는 것은 비현실적이다.

[0003] 능동적 학습에 의해 동기 부여된 능동적 도메인 적응(Active Domain Adaptation; ADA)은 도메인 이동 하에서 가장 유용한 대상 샘플에 주석을 다는 것을 목표로 하고 있으며, 점점 더 주목을 받고 있다. 오라클을 사용하여 레이블이 지정된 대상 샘플을 얻은 후, ADA 방법은 레이블이 지정된 소스 도메인과 부분적으로 레이블이 지정된 대상 도메인 사이의 도메인 격차를 줄인다. 능동 학습과 ADA 모두 특정 쿼리 선택 기준에 따라 레이블이 지정되지 않은 샘플을 자동으로 선택하여 모델에 도움이 될 수 있다. 그러나 ADA에서 쿼리 선택은 단일 도메인만 고려하는 능동 학습과 달리 도메인 전이 시나리오를 고려해야 한다. 또한 레이블이 지정된 소스, 레이블이 지정되지 않은 대상 및 레이블이 지정된 대상 도메인을 처리하는 것은 ADA 태스크에서 중요한 문제로 남아 있다.

[0004] 주식 예산이 제한된 ADA 시나리오에서 가장 유용한 대상 샘플을 선택하기 위해 상당한 노력이 기울여졌다. 종래 기술에서는 불확실성 또는 다양성을 쿼리 선택 기준으로 사용했다. 그러나 대상 도메인에 대한 불확실성 추정은 일반적으로 잘못 보정되어 샘플링 이상치 또는 중복 사례가 발생한다. 다양성으로 사용되는 도메인 유사성은 도메인 이동에 따른 차별적 특징 공간을 보장할 수 없다. 최근의 ADA 방법은 불확실성과 다양성을 모두 쿼리 선택 기준으로 채택했다. 종래기술에서는 여러 예측 간의 합의를 활용하여 불확실성과 도메인 유사성을 가진 커미티 방식별 쿼리를 채택한다.

선행기술문헌

비특허문헌

[0005] (비특허문헌 0001) [1] Shao, Jie-Jing, Guo, Lan-Zhe, Yang, Xiao-Wen, & Li, Yu-Feng LOG: Active model adaptation for label-efficient OOD generalization. In Thirty-sixth conference on neural information processing systems.

(비특허문헌 0002) [2] Yoo, Donggeun, & Kweon, In So (2019). Learning loss for active learning. In Proc. of computer vision and pattern recognition (pp. 93-102).

발명의 내용

해결하려는 과제

[0006] 본 발명이 이루고자 하는 기술적 과제는 전이 가능한 손실 기반 능동적 도메인 적응(Transferable Loss-based Active Domain Adaptation; TL-ADA) 방법 및 시스템을 제공하는데 있다. 제안하는 접근 방식은 능동 학습에서 우수한 결과를 보여준 손실 기반 쿼리 선택을 이용하지만, 손실 기반 쿼리 선택을 ADA 시나리오에 직접 적용하면 전이 가능성 문제와 낮은 다양성으로 인해 모델에 기여하지 않는 고손실 샘플이 축적되는 과제를 해결하기 위해, 대상 유사(target pseudo) 레이블과 도메인 적대적 손실을 통합하는 전이 가능한 이중 중첩 손실을 제안한다.

과제의 해결 수단

[0007] 일 측면에 있어서, 본 발명에서 제안하는 전이 가능한 손실 기반 능동적 도메인 적응 방법은 쿼리 선택부를 통해 레이블이 지정되지 않은 샘플의 손실을 추론한 후 전이 가능 이중 중첩 손실(Transferable doubly nested loss; TDNL)을 통한 손실 기반 쿼리를 선택하는 단계, 제1 모델 훈련부를 통해 상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 얻은 후, 레이블이 지정된 대상 도메인을 사용하여 모델을 훈련하는 단계 및 제2 모델 훈련부를 통해 레이블이 지정된 대상 샘플로 모델을 순차적으로 훈련한 후 상기 TDNL을 통해 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플로 모델을 훈련하는 단계를 포함한다.

[0008] 상기 제1 모델 훈련부를 통해 상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 얻은 후, 레이블이 지정된 대상 도메인을 사용하여 모델을 훈련하는 단계는 상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 사용하여 제1 모델 훈련부의 보조 손실 예측기를 훈련하고, 오라클(oracle)을 사용하여 레이블이 지정된 대상 도메인을 획득한 후, 교차점 손실 값을 사용하여 제1 모델 훈련부의 특징 추출기를 훈련하며, 상기 제1 모델 훈련부의 특징 추출기를 통해 획득한 실제 손실 값을 상기 제1 모델 훈련부의 보조 손실 예측기를 통해 획득한 보조 손실 예측 변수에 대한 레이블로 사용하고, 상기 제1 모델 훈련부의 특징 추출기가 도메인 이동에서 손실을 학습하도록 하여 레이블이 지정되지 않은 대상 샘플의 유사 레이블을 생성한다.

[0009] 상기 제2 모델 훈련부를 통해 레이블이 지정된 대상 샘플로 모델을 순차적으로 훈련한 후 상기 TDNL을 통해 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플로 모델을 훈련하는 단계는 레이블이 지정되지 않은 대상 샘플의 대상 샘플의 유사 레이블을 사용하여 제2 모델 훈련부의 특징 추출기를 자체 훈련 방식으로 훈련시키고, 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플 간의 불일치를 완화시키고, 상기 유사 레이블의 신뢰성을 높이기 위해 셀프-자기 엔트로피 및 클래스에 따른 정보 최대화(Information Maximization; IM) 손실 값을 사용한다.

[0010] 상기 쿼리 선택부를 통해 레이블이 지정되지 않은 샘플의 손실을 추론한 후 전이 가능 이중 중첩 손실을 통한 손실 기반 쿼리를 선택하는 단계는 레이블이 지정된 샘플에 대하여 훈련된 보조 손실 예측기를 통해 레이블이 지정되지 않은 샘플의 손실을 추론한 후, 손실이 가장 큰 샘플을 샘플로 선택하는 전이 가능 이중 중첩 손실을 통한 손실 기반 쿼리를 선택한다.

[0011] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 전이 가능한 손실 기반 능동적 도메인 적응 시스템은 레이블이 지정되지 않은 샘플의 손실을 추론한 후 전이 가능 이중 중첩 손실(Transferable doubly nested loss; TDNL)을 통한 손실 기반 쿼리를 선택하는 쿼리 선택부, 상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 얻은 후, 레이블이 지정된 대상 도메인을 사용하여 모델을 훈련하는 제1 모델 훈련부 및 레이블이 지정된 대상 샘플로 모델을 순차적으로 훈련한 후 상기 TDNL을 통해 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플로 모델을 훈련하는 제2 모델 훈련부를 포함한다.

발명의 효과

[0012] 본 발명의 실시예들에 따르면 전이 가능한 손실 기반 능동적 도메인 적응(Transferable Loss-based Active Domain Adaptation; TL-ADA) 방법 및 시스템을 통해 대상 유사(target pseudo) 레이블과 도메인 적대적 손실을 통합하는 전이 가능한 이중 중첩 손실을 제안함으로써, 손실 기반 쿼리 선택을 ADA 시나리오에 직접 적용하면 전이 가능성 문제와 낮은 다양성으로 인해 모델에 기여하지 않는 고손실 샘플이 축적되는 문제를 해결할 수 있다.

도면의 간단한 설명

[0013] 도 1은 종래기술에 따른 ADA에서 손실 기반 쿼리 선택의 한계를 설명하기 위한 도면이다.

도 2는 본 발명의 일 실시예에 따른 전이 가능한 손실 기반 능동적 도메인 적응 시스템의 구성을 나타내는 도면이다.

도 3은 본 발명의 일 실시예에 따른 전이 가능한 손실 기반 능동적 도메인 적응 방법을 설명하기 위한 흐름도이다.

도 4는 본 발명의 일 실시예에 따른 레이블이 지정된 대상 도메인을 사용한 모델 훈련 알고리즘을 나타내는 도면이다.

도 5는 본 발명의 일 실시예에 따른 레이블이 지정되지 않은 대상 도메인을 사용한 모델 훈련 알고리즘을 나타내는 도면이다.

도 6은 본 발명의 일 실시예에 따른 쿼리 선택 알고리즘을 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0014] 능동 도메인 적응(Active Domain Adaptation; ADA) 분야는 지도 학습 설정과 비지도 학습 설정 사이의 성능 격차를 줄이는 방법을 연구해 왔다. 이전의 ADA 연구는 주로 쿼리 선택에 중점을 두었지만 레이블이 지정된 소스 샘플과 레이블이 지정되지 않은 대상 샘플을 모두 사용하여 새로 레이블이 지정된 대상 샘플을 효과적으로 훈련

하는 방법에 대한 조사는 거의 없었다.

- [0015] 본 발명에서는 새로운 전이 가능 손실 기반 ADA(Transferable Loss-based ADA; TL-ADA) 프레임워크를 제시한다. 제안하는 접근 방식은 능동 학습에서 유망한 결과를 보여준 손실 기반 쿼리 선택에서 영감을 받았다. 그러나 손실 기반 쿼리 선택을 ADA 시나리오에 직접 적용하면 전이 가능성 문제와 낮은 다양성으로 인해 모델에 기여하지 않는 고손실 샘플이 축적된다. 이러한 과제를 해결하기 위해, 본 발명에서는 대상 유사(pseudo) 레이블과 도메인 적대적 손실을 통합하는 전이 가능한 이중 중첩 손실을 제안한다.
- [0016] 제안하는 TL-ADA 프레임워크는 도메인 유형(소스(source)/대상(target))과 레이블의 가용성(레이블 된(labeled)/레이블 되지 않은(unlabeled))을 모두 고려하여 모델을 순차적으로 훈련시킨다. 또한, 유사 레이블이 신뢰성을 향상시키기 위해 낮은 자기 엔트로피와 다양한 클래스 분포를 가질 것을 권장한다. 여러 벤치마크 데이터셋에 대한 실험은 제안하는 TL-ADA 모델이 이전 ADA 방법을 능가한다는 것을 보여주며, 심층 분석은 제안하는 접근 방식의 효과를 뒷받침한다. 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.
- [0017] 능동 학습(Active Learning; AL)은 레이블이 지정되지 않은 데이터 풀에서 가장 유용한 샘플에 주석을 달아 모델 훈련에 적극 활용하는 것을 목표로 한다. 기존 AL 방법은 불확실성(uncertainty), 커미티(committee), 다양성(diversity)의 세 가지 범주로 나눌 수 있다. 신뢰도에 기초한 불확실성 기반 접근법, 엔트로피 또는 best-vs-second-best는 단순성과 높은 계산 효율성 때문에 널리 사용되어 왔다. 그러나 대부분의 불확실성 기반 접근법은 작업별 불확실성 측정을 사용한다. 여러 분류기의 앙상블 간의 합의를 활용하는 커미티별 쿼리는 이 문제를 처리할 수 있으므로 광범위한 응용이 가능하다. 다양성 기반 선택 방법은 레이블이 지정되지 않은 데이터 풀의 전체 분포를 나타내는 샘플을 선택한다. 심층망은 작업 수, 작업 유형 및 모델 복잡성에 관계없이 손실을 최소화하여 훈련되기 때문에 최근에는 손실 기반 접근법을 학습하는 것이 주목 받고 있다. 모델의 손실 값을 불확실성으로 고려하여, 그들은 손실 예측 모델을 메인 네트워크에 부착하고 이 모델을 훈련시켜 레이블이 지정되지 않은 샘플의 손실을 추정했다.
- [0018] 특히 AL에서 사용되는 기존 기준은 단일 도메인을 고려하여 설계된다. 따라서 그것들은 전이할 수 없으며 도메인 적응 시나리오에서 충분히 효율적이지 않다. 그것들은 도메인 이동 하에서 주석을 위한 가장 유용한 샘플을 선택할 수 없었다. 이 문제를 해결하기 위해, 본 발명에서는 전이 가능한 손실 기반 활성 도메인 적응을 제안한다. 기존의 손실 기반 접근 방식과 제안하는 접근 방식의 주요 차이점은 제안하는 모델이 손실 예측 변수를 훈련하기 위해 레이블이 지정되지 않은 대상 샘플의 유사 레이블을 적극적으로 활용한다는 것이며, 이는 도메인 불일치를 완화할 수 있는 적대적 손실을 수반한다는 것이다. 본 발명에서는 기존의 손실 기반 접근 방식과 달리 제안된 전이 가능 이중 중첩 손실이 전이 가능 문제를 효과적으로 해결한다는 것을 정량적, 정성적으로 보여준다.
- [0019] 비지도 도메인 적응(Unsupervised Domain Adaptation; UDA)은 레이블이 지정된 소스 도메인의 지식을 레이블이 지정되지 않은 대상 도메인으로 전이할 수 있기 때문에 적극적으로 조사되었다. 초기 방법은 최대 평균 불일치(Maximum Mean Discrepancy; MMD)에 의존하는 반면, 최근의 접근 방식은 도메인 적대적 훈련을 활용한다. 도메인 판별기를 사용하여 두 도메인 간의 특징 분포를 조정함으로써, 적대적 훈련은 비지도 도메인 적응에서 인상적인 성능을 보였다. 유망한 결과에도 불구하고, UDA와 지도 전이 학습 사이의 성능 격차는 상당하다. 무시할 수 없는 성능 격차는 실제 문제에 UDA를 적용하는 데 주요 장애물이다. 이를 고려하여, 다양한 영역에서 소수의 레이블이 지정된 대상 샘플을 사용할 수 있는 준지도 도메인 적응(Semi-Supervised Domain Adaptation; SSDA)과 퓨샷 도메인 적응(Few-Shot Domain Adaptation; FSDA)이 제안되었다. ADA는 SSDA와 유사하다. 그러나 SSDA 및 FSDA 접근 방식은 레이블이 지정되지 않은 대상 샘플에 레이블을 랜덤 또는 미리 결정된 규칙에 따라 할당하는 반면, ADA는 도메인 이동을 최소화하기 위해 가장 유용한 대상 샘플을 자동으로 선택한다.
- [0020] 제한된 주석 예산으로 도메인 이동에서 가장 유용한 대상 샘플을 선택하는 방법은 ADA 작업의 주요 과제 중 하나이다. 초기 ADA 모델은 사전 훈련된 소스 분류기 또는 중요도 가중치를 사용하여 도메인 간 정보를 활용한다. 최근의 방법은 불확실성과 다양성을 쿼리 선택 기준으로 활용했다.
- [0021] 커미티 방식에 의한 쿼리는 ADA에도 사용되었으며, 레이블이 지정된 소스와 레이블이 지정된 대상만 훈련하면 도메인 적응에서 성능이 향상된다. 그러나 커미티의 쿼리에 대한 여러 분류기를 구성하려면 높은 계산 비용이 필요하다. 최근에는 쿼리 선택을 위해 소스 클래스 센터의 큰 마진인 샘플을 사용하였다.
- [0022] 종래기술에서는 의사 결정 경계에 가까운 하드 인스턴스를 저장하는 SVM과 유사하게 작동하는 방법을 보여준다. 또한 레이블에 대한 관련 쿼리를 선택하고 다른 데이터를 훈련하고 적응하는 방법에 대한 지침을 제공하는 2단

계 프레임워크를 개발했다. 기존의 ADA 접근 방식과 달리, 본 발명에서는 ADA에서 탐구되지 않은 전이 가능한 손실 기반 쿼리 선택을 도입한다. 또한 이전 ADA 방법에서와 같이 쿼리 선택 기준에 중점을 둘 뿐만 아니라 도메인 유형(소스/대상)과 레이블(레이블 지정/레이블 지정되지 않음)의 접근성을 모두 고려한 2단계 훈련 프로세스를 제안한다.

[0024] 도 1은 종래기술에 따른 ADA에서 손실 기반 쿼리 선택의 한계를 설명하기 위한 도면이다.

[0025] 손실 기반 쿼리 선택(즉, 학습 손실)은 심층 네트워크에서 작업에 구매 받지 않고 효율적이라는 장점이 있지만, 이전의 ADA 연구에서는 적용되지 않았다. 이 접근법의 핵심 아이디어는 레이블이 지정된 샘플에 대해 훈련된 보조 손실 예측기에 의해 레이블이 지정되지 않은 샘플의 손실을 추론한 다음 손실이 가장 큰 샘플을 가장 유용한 샘플로 선택하는 것이다. 결정적으로, 전통적인 학습 손실은 도 1(a)와 같이 도메인 이동 하에서 전이 가능한 문제와 낮은 다양성을 초래한다. 도 1(b)와 같이, 잘못 예측된 저손실 샘플의 수는 증가하는 반면, 주식 예산이 증가함에 따라 예측된 고손실 샘플의 수는 감소한다. 이러한 결과는 손실 기반 쿼리 선택에 의해 선택된 손실이 높은 샘플이 모델 훈련에 도움이 되지 않으며 전이 가능한 문제를 처리할 수 없음을 보여준다.

[0026] 이전의 ADA 연구는 주로 쿼리 선택 기준에 중점을 두었지만 쿼리 선택 후 레이블이 지정된 소스, 레이블이 지정된 대상 및 레이블이 지정되지 않은 대상 도메인을 처리하는 방법을 조사하지 않았다. 예를 들어, 새로 레이블이 지정된 대상 도메인과 레이블이 지정된 소스 도메인을 통합하여 도메인 적응을 수행했다. 레이블이 지정된 소스 도메인과 레이블이 지정된 대상 도메인을 직접 통합하지는 않았지만, 이를 단일 데이터 풀로 처리할 수도 있다. 그러나, 본 발명에서는 도메인 이동을 설명하면서 정보를 제공하는 레이블이 없는 대상 샘플을 선택하는 방법뿐만 아니라 레이블이 지정된 대상 샘플을 처리하는 방안도 제안한다.

[0027] 본 발명에서는 전이 가능한 손실 기반 능동적 도메인 적응(Transferable Loss-based Active Domain Adaptation; TL-ADA)을 제안한다. 먼저 원래 학습 손실의 전이 가능한 문제를 처리하기 위해 전이 가능 이중 중첩 손실(Transferable doubly nested loss; TDNL)을 도입한다. 본 발명에서는 도메인 격차를 완화할 수 있도록 특별히 설계된 기준의 특성으로 '전이 가능'이라는 용어를 정의한다. 또한 도메인 격차를 줄일 수 없는 특성을 '전이 불가'로 정의한다. 소수의 레이블이 지정된 샘플만으로 손실 예측 변수를 학습하는 기존 접근 방식과 비교하여 TDNL은 레이블이 지정되지 않고 레이블이 지정된 샘플의 유사 레이블을 통해 훈련되며 도메인 불일치를 완화할 수 있는 적대적 손실을 수반한다. 이를 통해 우리 모델은 선택된 샘플의 다양성과 전이 가능성을 높일 수 있다. 또한 도메인 유형(소스/대상)과 레이블(레이블 된/레이블 되지 않은)의 접근성을 모두 고려하는 순차 학습 체계인 2 스테이지 적응(Two-Stage-Adaptation; TSA)을 소개한다. 제안된 쿼리 선택에 의해 선택된 레이블이 지정된 대상 샘플은 소스 및 대상 샘플에 대한 차별적 대상 표현 또는 의미 정보를 풍부하게 하기 위한 정보를 포함할 수 있다. 따라서 이전 ADA 접근 방식에서와 같이 레이블이 지정된 소스 및 레이블이 지정된 대상 도메인을 하나의 데이터 풀로 처리하면 전이 가능한 문제를 처리할 수 없다. 제안하는 TSA 전략은 먼저 레이블이 지정된 대상 샘플의 풍부한 정보를 활용한 다음 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플의 모델 파라미터를 조정하여 도메인 불일치를 줄인다. 제안된 방법의 주요 특징 중 하나는 대상 도메인의 유사 레이블을 적극적으로 활용한다는 것이다. 이를 고려하여, 정보 극대화를 활용하여 셀프-엔트로피와 다양한 클래스 분포를 향상시켜 대상 유사 레이블의 신뢰성을 높인다.

[0028] 제안하는 방법이 다양한 구성 중에서 효과적이라는 것을 경험적으로 보여주며, 제안된 방식은 대상 샘플의 의사 레이블의 신뢰성을 점진적으로 증가시킨다. 또한 레이블의 도메인 유형과 접근성을 고려한 순차 학습의 효과를 보여준다. 전반적으로, 제안하는 방법은 다양한 ADA 데이터셋에서 다른 접근 방식을 크게 능가한다.

[0029] 본 발명의 실시예에 따르면, 도메인 이동에 따른 기존 손실 기반 쿼리 선택의 한계를 종합적으로 분석하고 레이블이 지정되지 않고 레이블이 지정된 대상 샘플의 수많은 유사 레이블을 적극적으로 활용하는 TDNL을 제안한다.

[0030] 본 발명의 실시예에 따르면, 레이블의 도메인 유형(소스/대상)과 접근성(레이블 지정/레이블 지정되지 않은)을 고려하는 새로운 순차 학습을 제안한다.

[0031] 본 발명의 실시예에 따르면, ADA 방법은 다양한 데이터셋, 특히 5% 주식 예산(주 실험 프로토콜)에서 큰 격차로 종래기술보다 우수하다. 제안하는 TDNL과 TSA가 다양한 ADA 시나리오에서 중요한 문제를 효과적으로 해결하는 것을 정량적, 정성적으로 확인하였다.

- [0033] 본 발명의 실시예에 따른 ADA 시나리오에서는 레이블이 지정된 소스 도메인 $\mathcal{X}_s = \{(x_s^i, y_s^i)\}_{i=1}^{N_s}$ 및 레이블이 지정되지 않은 대상 도메인 $\mathcal{X}_u = \{(x_u^i)\}_{i=1}^{N_u}$ 이 주어진다. 여기서 N_s 와 N_u 는 각각 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 도메인의 총 수이다. 모델 훈련이 진행됨에 따라 대상 도메인의 일부는 쿼리 선택에 따라 오라클에 의해 레이블이 지정되고 $\mathcal{X}_t = \{(x_t^i, y_t^i)\}_{i=1}^{N_t}$ 로 표시된다. 레이블이 지정된 대상 도메인 $\mathcal{X}_t$ 를 얻은 후 모델은 기존 $\mathcal{X}_s$ 와 나머지 $\mathcal{X}_u \leftarrow \mathcal{X}_u \setminus \mathcal{X}_t$ 로 준지도 방식으로 훈련된다.
- [0035] 도 2는 본 발명의 일 실시예에 따른 전이 가능한 손실 기반 능동적 도메인 적응 시스템의 구성을 나타내는 도면이다.
- [0036] 제안하는 전이 가능한 손실 기반 능동적 도메인 적응 시스템은 쿼리 선택부(210), 제1 모델 훈련부(220) 및 제2 모델 훈련부(230)를 포함한다.
- [0037] 본 발명의 실시예에 따르면, 레이블이 지정된 소스 도메인과 부분적으로 레이블이 지정된 대상 도메인 사이의 도메인 불일치를 완화하는 것을 목표로 한다.
- [0038] 쿼리 선택부(210)에서는 레이블이 지정된 샘플에 대하여 훈련된 보조 손실 예측기(P)를 통해 레이블이 지정되지 않은 샘플의 손실을 추론한 다음 손실이 가장 큰 샘플을 샘플로 선택하는 전이 가능 이중 중첩 손실(Transferable doubly nested loss; TDNL)을 통한 손실 기반 쿼리를 선택한다.
- [0039] 먼저, 보조 손실 예측 변수 P와 적대적 도메인 판별기 D를 활용하여 TDNL을 도입했다. 도메인 적대적 훈련과 함께 레이블이 지정되지 않은 샘플의 유사 레이블을 적극적으로 활용하는 쿼리 선택은 기존의 손실 기반 쿼리 선택의 전이 가능한 문제를 효과적으로 해결한다.
- [0040] 또한 레이블의 도메인 유형과 접근성을 모두 고려하는 TSA(Two-stage adaptation)라는 순차 학습 체계를 도입했다. 제1 모델 훈련부(220)에서, 쿼리 선택에 의해 선택된 레이블이 지정된 대상 샘플로 특징 추출기 F와 손실 예측기 P를 훈련시켰다. 모델이 레이블이 지정된 몇 개의 대상 샘플에 대해서만 훈련된 경우, 모델은 다양성이 낮은 소수의 레이블이 지정된 샘플에 의해 과대 장착될 수 있다.
- [0041] 제2 모델 훈련부(230)에서는 분류기 C에서 유사 레이블 지정 및 정보 최대화 손실이 있는 레이블 지정되지 않은 대상 샘플의 대다수를 활용했다. 마지막으로, 손실 예측 변수 P에 의한 최소-K 손실을 가진 샘플에 주석을 달았고, 최대 주석 예산에 도달할 때까지 위의 과정을 반복한다.
- [0042] 제안된 프레임워크 실행 순서는 초기 단계에서 이용 가능한 레이블 지정 데이터가 없기 때문에 레이블 지정 프로세스인 제2 모델 훈련부(230)에서 초기 단계를 시작하고, 이후 제1 모델 훈련부(220)를 통한 훈련을 시작한다. 레이블이 지정된 대상 데이터셋이 존재하면 도 2의 프레임워크에 표시된 것처럼 실행 순서가 따른다.
- [0043] 본 발명의 실시예에 따른 전이 가능 이중 중첩 손실(Transferable doubly nested loss; TDNL)에 있어서, 손실 기반 쿼리 선택의 직관은 손실이 높은 샘플이 모델에 의해 잘못 인식될 가능성이 높기 때문에 손실이 높은 샘플에 적극적으로 레이블을 지정하는 것이다. 그러나 ADA 방법은 소스 도메인과 대상 도메인을 모두 고려해야 하므로 기존의 손실 기반 쿼리 선택을 ADA에 적용하는 것은 현실적으로 불가능하다. 즉, ADA 작업에 대한 유용한 쿼리는 두 도메인 모두에서 전이 가능한 샘플이 될 것이다. 이 문제를 완화하기 위해 TDNL을 제안하여 전이 가능한 손실을 대상 작업에 차별화하고 도메인 이동에 대해 무차별적으로 만든다.
- [0044] 먼저, 본 발명에서는 이전에 쿼리 선택을 통해 얻은 레이블이 지정된 대상 샘플 $\mathcal{X}_t$ 를 사용하여 보조 손실 예측 변수 P를 훈련시켰다. 오라클을 사용하여 레이블이 지정된 대상 도메인 $\mathcal{X}_t = \{(x_t^i, y_t^i)\}_{i=1}^{N_t}$ 을 얻은 후 기존의 교차점 손실 $\mathcal{L}_{ce}$ 를 사용하여 특징 추출기 F를 훈련시켰다. 그런 다음, 얻은 실제 손실 값 $\mathcal{L}_{ce}$ 를 보조 손실 예측 변수 P에 대한 ground-truth 레이블로 사용했다. 실제 손실 $\mathcal{L}_{ce}$ 의 규모는 모델 교육이 진행됨에 따라 꾸준히 감소했다. 따라서, 다음과 같이 표현되는 마진 순위 손실을 사용했다.

$$L_{loss} = \mathbb{E}_{x_t \in \mathcal{X}_t} \max(0, -\mathbb{1}(l_n, l_m) \cdot (\hat{l}_n - \hat{l}_m) + \Delta)$$

$$s.t. \mathbb{1}(l_n, l_m) = \begin{cases} +1, & \text{if } l_n > l_m \\ -1, & \text{otherwise,} \end{cases} \quad (1)$$

여기서 n 과 m 은 훈련 미니 배치의 순위 쌍 인덱스이고 Δ 는 사전 정의된 마진이다. 기존의 손실 기반 쿼리 선택과 비교하여 손실 예측 변수 P 의 그래디언트가 특징 추출기 F 로 흐르기 때문에 공동으로 학습된다. 이 접근 방식은 특징 추출기 F 가 도메인 이동에서 손실을 학습하도록 하여 레이블이 지정되지 않은 대상 샘플의 유사 레이블을 생성하는 데 효과적이다.

레이블이 지정된 대상 샘플의 수가 적기 때문에 레이블이 지정된 대상 샘플로만 모델을 교육하면 전이 가능한 문제를 처리할 수 없다. 따라서 유사 레이블을 활용하여 레이블이 지정되지 않은 대상 샘플의 대부분을 활용했다. 손실 예측 변수 P 를 훈련하기 위해 샘플의 실제 레이블이 필요하지만 레이블이 지정되지 않은 대상 도메인 $\mathcal{X}_{tu}$ 에는 레이블이 없다.

이 문제를 해결하기 위해, 본 발명에서는 다음과 같이 얻은 레이블이 지정되지 않은 대상 샘플 x_{tu} 의 대상 유사 레이블 $\tilde{y}_t$ 를 사용했다.

$$\tilde{y}_t = \underset{c}{\operatorname{argmax}}(C(F(x_{tu}))_c), \quad (2)$$

여기서 F , C , c 는 각각 특징 추출기, 분류기, 클래스 인덱스를 나타낸다. 유사 레이블은 모든 단계에서 미리 결정된 간격 $\mathcal{V}$ 로 업데이트되며, 특징 추출기 F 를 자체 훈련(self-training) 방식으로 훈련시키는 데 사용된다. 이후, 유사 레이블과 모델의 추론 사이에서 유사 교차 엔트로피(pseudo-cross-entropy) 손실 $\mathcal{L}_{ce}$ 를 사용했다. $\mathcal{L}_{ce}$ 는 손실 예측 변수 P 의 ground-truth 레이블로 사용된다.

또한, 레이블이 지정된 소스 도메인 $\mathcal{X}_s$ 와 레이블이 지정되지 않은 대상 도메인 $\mathcal{X}_{tu}$ 사이의 도메인 불일치를 완화하기 위해 특징 추출기 F 와 도메인 판별기 D 사이의 기존 미니-맥스 게임(mini-max game)을 활용했다. 도메인 판별기 D 는 주어진 샘플의 도메인 레이블을 분류하려고 시도하는 반면, 특징 추출기 F 는 다음과 같이 도메인 식별자 D 를 속이려고 시도한다:

$$L_{dis} = -\mathbb{E}_{x_s \in \mathcal{X}_s} [\log D(F(x_s))] - \mathbb{E}_{x_{tu} \in \mathcal{X}_{tu}} [\log(1 - D(F(x_{tu})))] \quad (3)$$

$$L_{adv} = -L_{dis} \quad (4)$$

최종 TDNL은 다음 식으로 계산된다:

$$L_{TDNL} = L_{loss} + \lambda L_{adv}, \quad (5)$$

여기서 λ 는 예측 손실과 적대적 손실 사이의 균형 파라미터이다. 이 접근 방식은 도메인 이동에 따른 감독을 제공하여 전이 가능한 손실을 생성하고 다양한 문제를 처리하는 데 효과적이다.

모델 훈련을 위한 유사 레이블의 신뢰성과 관련하여 자연스러운 질문(question)이 발생할 수 있다. 대상 유사 레이블의 신뢰성을 높이기 위해 다음과 같이 표현되는 자기 엔트로피와 클래스 다양성을 고려한 정보 최대화(Information Maximization; IM) 손실을 사용했다:

$$L_{im} = -\mathbb{E}_{x_{tu} \in \mathcal{X}_{tu}} \sum_{c=1}^C [\mathbb{1}_{[c=\tilde{y}_t]} \log(\sigma(C(F(x_{tu}))))] + \xi D_{KL}(\hat{y}_t, \frac{1}{C} \mathbf{1}_C) - \log(C). \quad (6)$$

[0059] 첫 번째 항은 자체 엔트로피를 나타내며, 이는 모델이 x_u 의 특징 표현에 이질적인 단일 핫 인코딩을 할당하도록 장려한다. 두 번째 항은 대상 유사 레이블 $\tilde{y}_t$ 가 소수의 클래스에만 할당되는 상황(즉, 낮은 다양성)을 피하기 위해 사용된다. 위의 식에서 σ 는 Softmax 함수이고 $\mathbb{1}[\cdot]$ 은 지시 함수이다. C 는 클래스의 수이고 $\mathbf{1}_C$ 는 모든 요소가 1인 벡터이며 클래스의 수와 크기가 같다. $\hat{y}_t = \mathbb{E}_{x_t \in x_u} [\sigma(C(F(x_u)))]$ 는 레이블이 지정되지 않은 전체 대상 도메인의 평균 출력 확률이다. D_{KL} 은 쿨백-라이블러 발산이며 ξ 는 두 손실 사이의 제어 변수이다. 결과적으로, 제안하는 모델은 다양한 클래스에 대한 정보 최대화 손실로 더 신뢰할 수 있는 대상 유사 레이블 $\tilde{y}_t$ 를 생성한다. 보조 손실 예측 변수 P 를 포함한 ADA 프레임워크를 훈련한 후 모델 성능을 향상시킬 수 있는 레이블이 지정되지 않은 대상 샘플을 선택했다. 본 발명의 실시예에 따른 경험적 결과에서 손실 예측기에 의한 높은 손실 샘플이 도메인 이동에서 가장 유용한 샘플이 아니라는 것을 밝혔다.

[0061] 도 3은 본 발명의 일 실시예에 따른 전이 가능한 손실 기반 능동적 도메인 적응 방법을 설명하기 위한 흐름도이다.

[0062] 제안하는 전이 가능한 손실 기반 능동적 도메인 적응 방법은 쿼리 선택부를 통해 레이블이 지정되지 않은 샘플의 손실을 추론한 후 전이 가능 이중 중첩 손실(Transferable doubly nested loss; TDNL)을 통한 손실 기반 쿼리를 선택하는 단계(310), 제1 모델 훈련부를 통해 상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 얻은 후, 레이블이 지정된 대상 도메인을 사용하여 모델을 훈련하는 단계(320) 및 제2 모델 훈련부를 통해 레이블이 지정된 대상 샘플로 모델을 순차적으로 훈련한 후 상기 TDNL을 통해 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플로 모델을 훈련하는 단계(330)를 포함한다.

[0063] 단계(310)에서, 쿼리 선택부를 통해 레이블이 지정되지 않은 샘플의 손실을 추론한 후 전이 가능 이중 중첩 손실(Transferable doubly nested loss; TDNL)을 통한 손실 기반 쿼리를 선택한다.

[0064] 본 발명의 실시예에 따른 ADA의 주요 목적은 도메인 불일치를 완화하는 데 도움이 되는 가장 유용한 샘플을 선택하는 것이다. 선택된 대상 샘플은 대상 도메인을 잘 나타낼 수 있는 몇 가지 대상 샘플 또는 도메인 적대적 훈련으로 인해 소스 도메인과 유사한 분포를 가진 대상 샘플의 두 가지 특성 중 하나를 가질 수 있다. 대표적인 정보로 대상 샘플을 학습하면 대상 도메인의 클래스를 구별할 수 있는 능력이 향상된다. 한편, 소스 도메인과 유사한 분포를 가진 대상 샘플은 두 도메인에 걸쳐 일반화 측면에서 모델에 도움이 될 수 있다. 이러한 시나리오에서 레이블이 지정된 대상 샘플과 레이블이 지정된 소스 샘플을 단일 데이터 풀로 처리하면 전이 가능한 문제를 효과적으로 해결할 수 없다. 대신, 본 발명에서는 선택된 대상 샘플을 조화롭게 활용하기 위해 TSA이라는 순차적 훈련 방법을 제안하고, 이는 DA 성능을 크게 향상시킬 수 있다.

[0065] 단계(320)에서, 제1 모델 훈련부를 통해 상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 얻은 후, 레이블이 지정된 대상 도메인을 사용하여 모델을 훈련한다.

[0066] 상기 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 대상 샘플을 사용하여 제1 모델 훈련부의 보조 손실 예측기를 훈련한다. 오라클(oracle)을 사용하여 레이블이 지정된 대상 도메인을 획득한 후, 교차점 손실 값을 사용하여 제1 모델 훈련부의 특징 추출기를 훈련한다.

[0067] 상기 제1 모델 훈련부의 특징 추출기를 통해 획득한 실제 손실 값을 상기 제1 모델 훈련부의 보조 손실 예측기를 통해 획득한 보조 손실 예측 변수에 대한 레이블로 사용하고, 상기 제1 모델 훈련부의 특징 추출기가 도메인 이동에서 손실을 학습하도록 하여 레이블이 지정되지 않은 대상 샘플의 유사 레이블을 생성한다.

[0069] 도 4는 본 발명의 일 실시예에 따른 레이블이 지정된 대상 도메인을 사용한 모델 훈련 알고리즘을 나타내는 도면이다.

[0070] 손실 기반 쿼리 선택에 의해 새로운 레이블이 지정된 샘플을 얻은 후, 제안하는 모델은 도 4의 알고리즘1을 통해 새롭게 레이블 지정된 샘플 x_t 로 훈련된다. 훈련 가능한 모든 파라미터가 다음과 같이 업데이트 된다:

$$\theta_F^1, \theta_P^1 \leftarrow (F, P; \mathcal{X}_t). \quad (7)$$

여기서 θ_F , θ_P 는 각각 각 네트워크의 가중치(F, P)이다. 특징 추출기 F는 레이블이 지정된 샘플과 ground truth 교차 엔트로피 $\mathcal{L}_{ce}$ 에 의해 계산된 순위 손실과 예측 손실에 의해 훈련된 손실 예측기 P를 사용하여 훈련된다. 레이블이 지정되지 않은 대상 샘플은 이 단계에서 사용되지 않는다. 제1 모델 훈련부에서 제안하는 모델은 레이블이 지정된 대상 샘플에 더 잘 맞도록 훈련되었다. 이를 통해 모델은 레이블이 지정된 대상 도메인에서 의미 있는 정보를 캡처할 수 있다. 그러나 레이블이 지정된 대상 도메인을 학습하는 것이 도메인 적응 성능만 향상시키는 것은 아니라는 것을 알게 되었다. 따라서 레이블이 지정된 샘플과 레이블이 지정되지 않은 샘플 간의 합의가 되는 파라미터를 정규화 함으로써 DA 성능을 향상시키기 위해 레이블 지정 및 레이블 지정되지 않은 대상 도메인에 대한 방법을 추가로 조사했다.

단계(330)에서, 제2 모델 훈련부를 통해 레이블이 지정된 대상 샘플로 모델을 순차적으로 훈련한 후 상기 TDNL을 통해 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플로 모델을 훈련한다.

레이블이 지정되지 않은 대상 샘플의 대상 샘플의 유사 레이블을 사용하여 제2 모델 훈련부의 특징 추출기를 자체 훈련 방식으로 훈련시킨다.

레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플 간의 불일치를 완화시키고, 상기 유사 레이블의 신뢰성을 높이기 위해 셀프-자기 엔트로피 및 클래스에 따른 정보 최대화(Information Maximization; IM) 손실 값을 사용한다.

도 5는 본 발명의 일 실시예에 따른 레이블이 지정되지 않은 대상 도메인을 사용한 모델 훈련 알고리즘을 나타내는 도면이다.

이 단계는 레이블이 지정되거나 레이블이 지정되지 않은 샘플에 의존하지 않고 합의가 되도록 파라미터를 조작하는 것이다. 모델 파라미터는 제1 모델 훈련부를 통해 레이블이 지정된 대상에 맞도록 훈련되므로 네트워크는 DA 작업의 일반화를 위한 파라미터를 해제하고 성능을 저하시키지 않아야 한다. 제2 모델 훈련부는 도 5의 알고리즘 2 프로세스를 통해 레이블이 지정되지 않은 대상 도메인 $\mathcal{X}_u$ 및 $\mathcal{X}_s$ 로만 훈련된다. 모든 훈련 가능한 파라미터는 다음과 같이 업데이트 된다.

$$\theta_F^2, \theta_P^2, \theta_D^2 \leftarrow (F, P, D; \mathcal{X}_u, \mathcal{X}_s) \quad (8)$$

여기서 θ^1 은 제1 모델 훈련부의 파라미터이고, θ^2 는 레이블이 지정되지 않은 대상 및 소스 도메인을 학습하는데 새로 계산된 파라미터이다.

요약하면, 본 발명의 실시예에서는 대상 도메인의 유사 레이블과 도메인 판별기의 적대적 손실을 사용하여 샘플 손실의 순위를 예측하여 전이 가능한 문제를 처리하기 위해 TDNL을 도입했다. 레이블이 지정된 대상 샘플로 모델을 순차적으로 훈련한 다음 TDNL 및 IM 손실을 통해 레이블이 지정된 소스 및 레이블이 지정되지 않은 대상 샘플로 모델을 훈련하여 손실 및 도메인 이동의 잘못된 예측을 완화했다.

도 6은 본 발명의 일 실시예에 따른 쿼리 선택 알고리즘을 나타내는 도면이다.

본 발명의 실시예에 따른 쿼리 선택 방법은 보조 모듈(즉, 보조 손실 예측기)을 활용하여 샘플의 예상 손실을 예측하는 학습 손실을 기반으로 한다. 본 발명에서는 제안하는 전이 가능 이중 중첩 손실(TDNL)로 각 모듈을 훈련시킨다. 프레임워크로 훈련한 후 모델은 레이블이 지정되지 않은 대상 도메인에서 예상 손실을 측정하고 레이블 지정을 위한 K 샘플을 선택하여 유용한 쿼리를 선택한다. 본 발명의 실시예에 따른 쿼리 선택은 도 6의 알고리즘3을 따른다.

이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어

구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0088] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

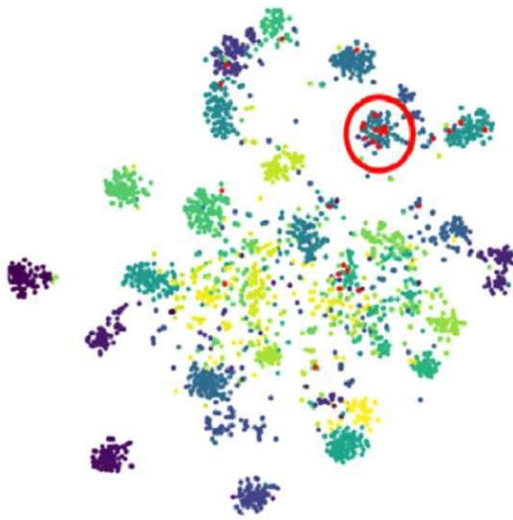
[0089] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0090] 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 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.

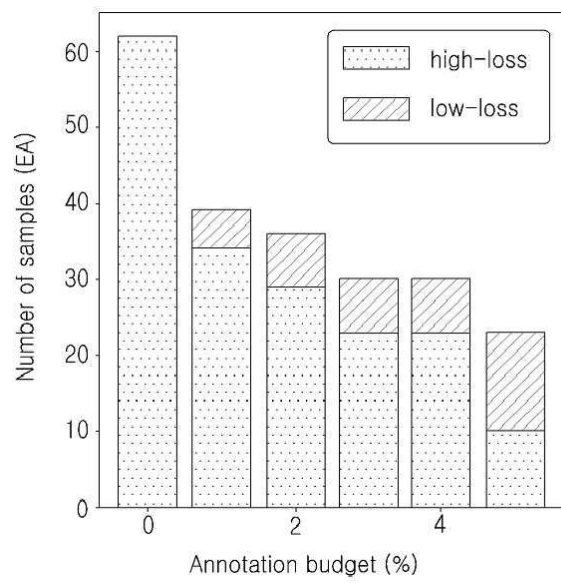
[0091] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

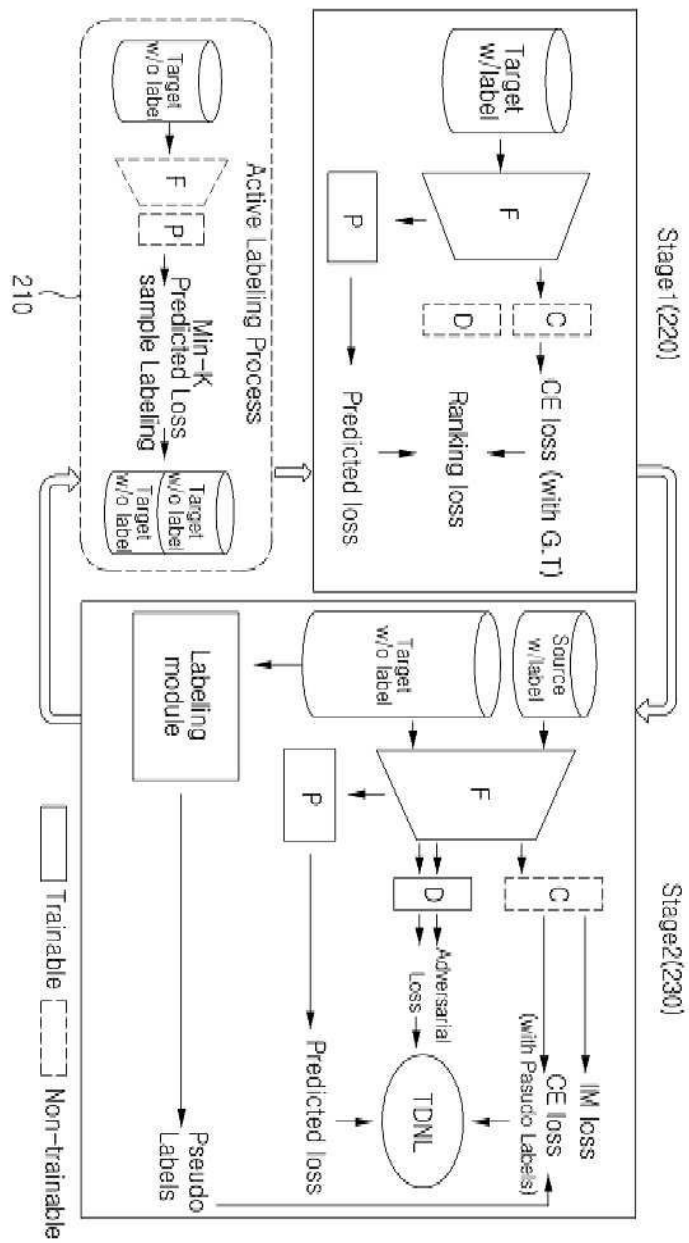
도면1



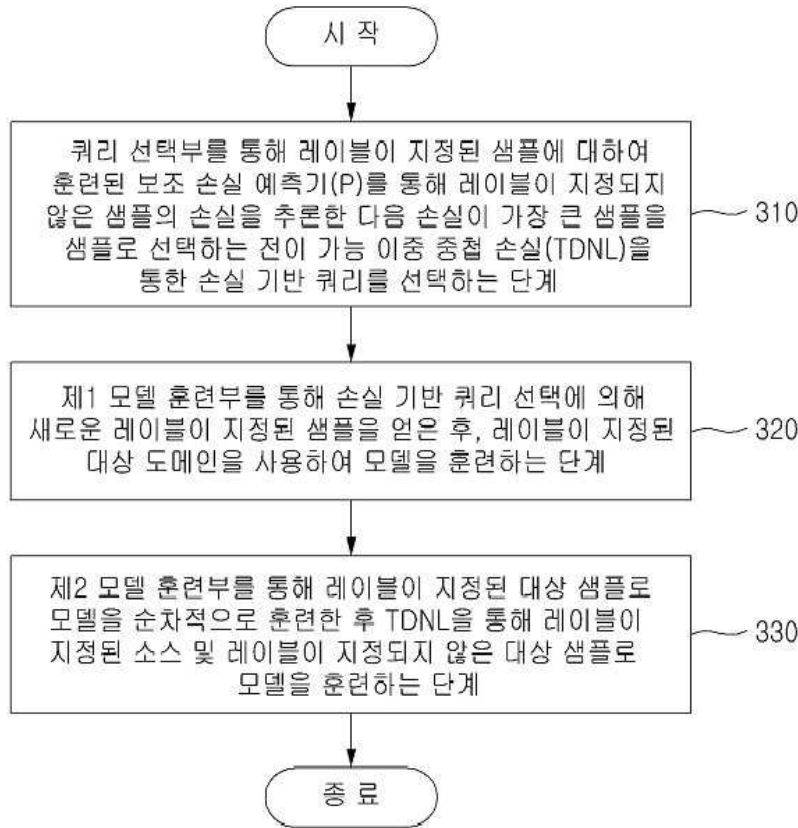
도면1b



도면2



도면3



도면4

Algorithm 1 Model training with labeled target domain

Input: labeled target domain: X_t , trainable modules: (F, P) , non-trainable module: (C, D) total iterations N , $\theta_{(F,P)}$: parameters of each module.

- 1: for $n = 1$ to N do
 - 2: $\hat{I}_i \leftarrow P(F(x_t))$
 - 3: $L_{ce} \leftarrow$ with target labels
 - 4: $L_{loss} \leftarrow L_{loss}(L_{ce}, \hat{I}_i)$ via ranking loss
 - 5: optimize three modules F , and P in turn:

$$\theta_F^1 \leftarrow \theta_F^1 - \eta \nabla_{\theta_F^1} (L_{ce})$$

$$\theta_P^1 \leftarrow \theta_P^1 - \eta \nabla_{\theta_P^1} (L_{loss})$$
 - 6: end for
-

도면5

Algorithm 2 Model training with unlabeled target domain

Input: labeled source domain: X_s , unlabeled target domain: X_t , trainable modules: (F, P, D) , non-trainable module: C , total iterations N .

$\theta_{(F, P, D)}$: parameters of each module, γ : intervals for pseudo labeling.

```

1: Obtain models trained by Stage 1
2:
3: for  $n = 1$  to  $N$  do
4:   if  $n \% \gamma == 0$  then
5:     get pseudo labels using Eq. (2)
6:   end if
7:    $\hat{d}_i \leftarrow D(F(x_t, x_s))$ 
8:    $\hat{I}_i \leftarrow P(F(x_t))$ 
9:    $I_{G.T} \leftarrow L_{ce}$  with target pseudo labels
10:
11:    $L_{im} \leftarrow L_{ent} + \xi L_{class\_div}$  via Eq. (6)
12:    $L_{dis} \leftarrow -\log(\hat{d}_i) - \log(1 - \hat{d}_i)$  via Eq. (3)
13:    $L_{adv} \leftarrow -L_{dis}$  via Eq. (4)
14:    $L_{TDNL} \leftarrow L_{loss}(I_{G.T}, \hat{I}_i) + \lambda L_{adv}$  via Eq. (5)
15:
16:   optimize three modules  $F$ ,  $P$ , and  $D$  in turn:
        $\theta_F \leftarrow \theta_F - \eta \nabla_{\theta_F} (L_{im} + L_{adv})$ 
        $\theta_P \leftarrow \theta_P - \eta \nabla_{\theta_P} (L_{TDNL})$ 
        $\theta_D \leftarrow \theta_D - \eta \nabla_{\theta_D} (L_{dis})$ 
17: end for

```

도면6

Algorithm 3 Query selection

Input: unlabeled target domain: X_{t_u} , non-trainable modules: (F, P) ,
 annotation budget: B
 Output: set of queries $Q(X)$

- 1: $L = []$ // initialize a predicted loss set
- 2: $Q = []$ // initialize an active sample set
- 3: for $i = 1$ to N_{t_u} do
- 4: $\hat{l}^i = P(F(x_{t_u}^i))$
- 5: $L \leftarrow L \cup \hat{l}^i$
- 6: end for
- 7: for $i = 1$ to B do
- 8: $Q \leftarrow Q \cup \text{argmin}_{x_{t_u}^i}(L)$
- 9: $L \leftarrow L \setminus \hat{l}^i$
- 10: end for
