



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0125008
(43) 공개일자 2022년09월14일

(51) 국제특허분류(Int. Cl.)
G06N 3/08 (2006.01) G06N 3/04 (2006.01)
G06T 7/20 (2017.01)
(52) CPC특허분류
G06N 3/084 (2013.01)
G06N 3/04 (2013.01)
(21) 출원번호 10-2021-0028872
(22) 출원일자 2021년03월04일
심사청구일자 2021년03월04일

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
이필규
서울특별시 동작구 상도로53길 8, 314동 601호(상도동, 래미안상도3차아파트)
(74) 대리인
특허법인아이더스

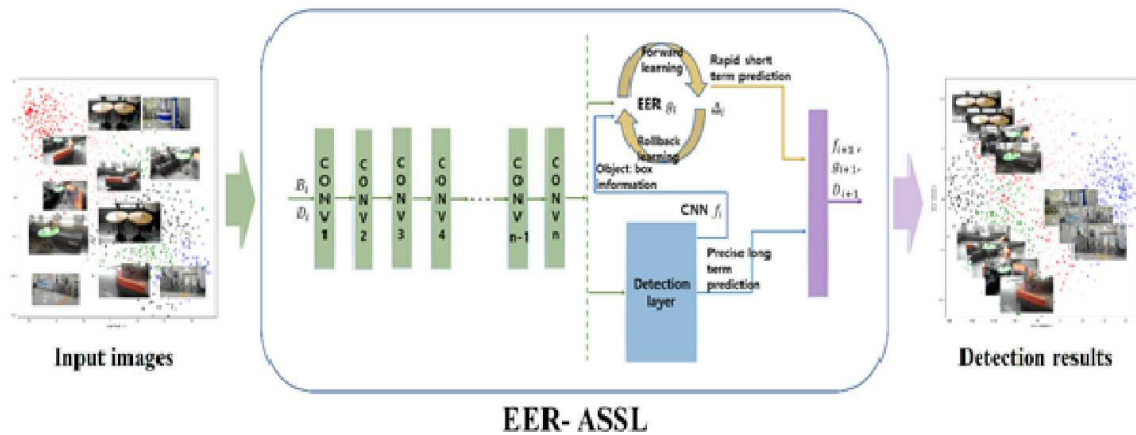
전체 청구항 수 : 총 4 항

(54) 발명의 명칭 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법

(57) 요약

본 발명은 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법에 관한 것으로, 능동학습(AL) 및 롤백 기반의 준 지도 학습(SSL)을 활용함으로써 AL을 통해 불확실성과 다양성을 측정하는 보다 유익하고 대표적인 샘플을 선택할 수 있고 SSL은 선택한 스트리밍 이미지 샘플을 빈(Bin)으로 나누고 각 빈(Bin)은 EER 롤백 학습 알고리즘과의 수렴 및 통합이 달성될 때까지 EER 및 CNN 모델의 식별 정보를 다음 빈으로 반복적으로 전송하며 EER 모델은 빠른 단기 근시 적응을 제공하고 CNN 모델은 점진적인 장기 성능 향상을 시킴으로써 다양한 데이터 분포에서 노이즈가 많고 편향된 레이블을 극복 할 수 있는 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법을 제공한다.

대표도 - 도2



(52) CPC특허분류

G06T 7/20 (2013.01)

G06T 2207/20084 (2013.01)

명세서

청구범위

청구항 1

능동학습(AL)을 통해 입력 데이터 스트림에서 불확실성과 다양성을 측정하는 스트리밍 이미지 샘플을 선택하는 제1단계; 및

준 지도 학습(SSL)을 통해 상기 선택한 스트리밍 이미지 샘플을 빈(Bin)으로 나누고 각 빈(Bin)은 EER 롤백 학습 알고리즘과의 수렴될 때까지 EER 및 CNN 모델의 식별 정보를 다음 빈으로 반복적으로 전송하는 제2단계;를 포함하는 것을 특징으로 하는 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법.

청구항 2

제 1항에 있어서,

상기 제1단계는 데이터 샘플 배치가 수집되고 공동 샘플링 알고리즘에 의해 공동 샘플링을 통해 스트리밍 이미지 샘플을 선택하는 단계인 것을 특징으로 하는 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법.

청구항 3

제 1항에 있어서,

상기 제2단계는 롤백 기능이 있는 준 지도 학습(SSL) 알고리즘을 통해 EER 기반으로 순방향 학습(forward learning) 및 롤백 학습(rollback learning)하여 샘플의 재선택 및 레이블 재지정하는 것을 특징으로 하는 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법.

청구항 4

제 1항에 있어서,

상기 롤백 학습(rollback learning)은 상기 빈(Bin)에서 샘플 제거, 레이블 변경 및 재선택하는 것을 특징으로 하는 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법.

발명의 설명

기술 분야

[0001]

본 발명은 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법에 관한 것으로서, 더욱 상세하게는 스트리밍 객체 감지시 학습 성능을 향상시키기 위해 롤백 학습 기반의 EER과 ASSL을 결합하여 신뢰할 수 있는 레이블 추정 및 성능을 향상시킬 수 있는 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법에 관한 것이다.

배경 기술

[0003]

최근 인공지능과 머신 러닝에서 주목할 만한 발전이 있었다. 패턴 분류 기술, 패턴 추출 및 패턴에서 습득한 지식을 기반으로 예측을 하는 작업은 특히 복잡하고 변화하는 실제 세계에서 여전히 어려운 문제이다. 객체 감지는 컴퓨터 비전의 중요한 하위 영역 중 하나이다. 지난 몇 년 동안 딥 러닝 기술은 Krishnaswamy 등이 2012년에 장벽을 깬 이후 많은 컴퓨터 비전 분야에 성공적으로 적용되었다. 대부분의 성능 향상은 간단한 정적 데이터 분포를 가정하여 올바르게 레이블링된 대규모 데이터 세트의 가용성에 의존한다. 그러나 실제 세계의 기본 분포는

매우 다양하고 불균형하며 시스템을 재구축하려면 어렵고 시간이 많이 걸리는 노동과 노력이 필요하다. 또한 일부 수집된 샘플은 편향되거나 잘못 표시되는 경향이 있으며 성능 저하로 이어질 수 있다. 완전히 레이블링된 재교육(retraining)은 비용과 시간 제약으로 인해 실제로는 불가능하지만 레이블이 없는 데이터를 수집하는 것은 상대적으로 저렴하다. 레이블이 지정된 데이터와 레이블이 없는 데이터를 모두 활용할 수 있다면 매우 실용적인 가치가 될 수 있다. 레이블이 없는 데이터를 레이블링된 데이터와 함께 사용할 수 있도록 하는 반지도 학습(SSL)은 학습 성능을 크게 향상시킬 수 있다.

[0004] 능동학습(AL)은 준지도 학습의 특별한 경우로 생각할 수 있다. 능동학습(AL)은 모델 성능을 최적화하기 위해 알 수 없는 데이터의 레이블을 얻기 위해 Oracle에 대한 최소 쿼리를 사용한다. 능동학습(AL) 프로세스에서 성공적인 모델 수렴을 위해서는 대부분의 경우 숙련된 전문가의 직관이 중요하다. 인적 레이블링은 노동 집약적이고 시간 소모적이며 오류가 발생하기 쉬우므로 항상 많은 비용이 필요하다. 따라서 일반적인 인간의 노력을 효과적으로 활용하는 방법은 능동학습(AL) 연구의 주요 관심사 중 하나이다. 능동학습(AL)의 장점은 레이블이 없는 각 샘플의 잠재력을 고려하여 레이블이 없는 샘플을 탐색하는 것이다. 능동학습(AL)은 수동적인 방식으로 레이블링된 데이터에 의존하는 대신 최소 레이블링 비용으로 가장 유익한 샘플을 탐색하기 위해 선택적 샘플링을 사용한다. 능동학습(AL)에서 트레이닝 데이터의 레이블링 노력은 기존의 지도학습 방법과 비교할 때 많이 감소한다. 레이블이 지정되지 않은 데이터 샘플에 레이블이 지정되면 순방향 학습 조작에 사용되지만 잡음이 있는 샘플 또는 레이블이 있는 경우 모델의 성능 저하로 이어질 수 있다. 레이블이 잘못 지정된 샘플 또는 이상치의 노이즈는 일반화 능력 구축에 부정적인 영향을 미친다. 또한 그것은 부정적인 영향을 미칠 수 있으며, 이에 최근 모델 일반화 능력 향상을 위해 양방향 능동학습(BDAL; Bi-Directional Active Learning)이 제안되었다. 양방향 능동학습(BDAL)은 EER 기반 불확실성 공식을 기반으로 순방향 및 역방향 학습 프로세스를 결합하고 단방향 노력에 비해 우수한 성능을 달성하였다. 그러나 BDAL은 여전히 데이터 선택에서 다시 실행하고 데이터 레이블링에서 실행 취소하는 데 시간이 많이 걸리며 변화하는 스트리밍 데이터 환경에서 사용할 수 없으므로 빠른 적응 학습 기능이 필요하다. 레이블이 약한 비디오의 모션 정보는 고정밀 개체 제안을 학습하는 데 사용할 수 있다. Krishna 등은 후보 객체 인식과 약한지도 학습을 통합하면 탐지 성능이 향상됨을 보여준다. CNN과 같은 심층 신경망(Deep Neural Networks)은 일반화 기능을 제공하는 가장 강력한 계층형 학습 네트워크 중 하나이며 30 년 동안 그렇게 해왔다. 그러나 심층 신경망의 느린 수렴은 여전히 매우 어려운 문제이며 특히 실시간 데이터 스트림에서 중요한 장애물이다.

[0005] ASSL이라고 하는 AL과 SSL의 결합은 다양한 데이터 분포에서 객체 감지 성능을 효율적으로 개선하여 대규모 레이블 데이터 세트와 긴 학습 시간이 필요한 CNN 접근 방식의 한계를 극복 할 수 있다. 최근 증분 ASSL 접근 방식이 제안되었으며 CNN의 지식 전달 능력과 증분 학습으로 인해 시변 문제를 효율적으로 처리 할 수 있음을 보여준다. 그러나 AL과 SSL이 결합된 방법은 여전히 긴 CNN 딥 러닝 단계를 필요로 하며 빠른 적응 학습 기능을 충족하기 위해 사소한 양의 트레이닝 시간을 소비한다.

[0006] 이상과 관련하여 참고문헌 1 내지 3이 제안된 바 있다.

선행기술문헌

특허문헌

- [0008] (특허문헌 0001) 참고문헌 1: 한국등록특허 제10-2174048호
(특허문헌 0002) 참고문헌 2: 한국등록특허 제10-1990123호
(특허문헌 0003) 참고문헌 3: 한국등록특허 제10-1882743호

발명의 내용

해결하려는 과제

[0009] 따라서 이러한 종래기술들의 문제 및 한계들을 해결하기 위한 것으로서 본 발명은 실제 다양한 불균형한 데이터 분포에서 노이즈가 많고 편향된 레이블을 극복 할 수 있는 신속 적응형 객체 감지를 위한 물백 딥 러닝 방법을 제공하는데 그 목적이 있다.

과제의 해결 수단

- [0011] 이와 같은 기술적 과제를 해결하기 위해 본 발명은;
- [0012] 능동학습(AL)을 통해 입력 데이터 스트림에서 불확실성과 다양성을 측정하는 스트리밍 이미지 샘플을 선택하는 제1단계; 및 준 지도 학습(SSL)을 통해 상기 선택한 스트리밍 이미지 샘플을 빈(Bin)으로 나누고 각 빈(Bin)은 EER 롤백 학습 알고리즘과의 수렴될 때까지 EER 및 CNN 모델의 식별 정보를 다음 빈으로 반복적으로 전송하는 제2단계;를 포함하는 것을 특징으로 하는 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법을 제공한다.
- [0013] 이때 상기 제1단계는 데이터 샘플 배치가 수집되고 공동 샘플링 알고리즘에 의해 공동 샘플링을 통해 스트리밍 이미지 샘플을 선택하는 단계인 것을 특징으로 한다.
- [0014] 그리고 상기 제2단계는 롤백 기능이 있는 준 지도 학습(SSL) 알고리즘을 통해 EER 기반으로 순방향 학습(forward learning) 및 롤백 학습(rollback learning)하여 샘플의 재선택 및 레이블 재지정하는 것을 특징으로 한다.
- [0015] 또한 상기 롤백 학습(rollback learning)은 상기 빈(Bin)에서 샘플 제거, 레이블 변경 및 재선택하는 것을 특징으로 한다.

발명의 효과

- [0017] 본 발명에 따르면, 노이즈가 많은 데이터 분포가 있는 경우 ERR기반 롤백 학습과 CNN 객체 감지를 위한 빈 기반 SSL을 결합함으로써 실제 다양한 불균형한 데이터 분포에서 노이즈가 많고 편향된 레이블을 극복하면서도 기존의 최첨단 탐지기에 비해 더 높은 정확도를 달성하고 사람의 노력도 줄일 수 있다.
- [0018] 특히 본 발명은 빠른 단기 근시 적응을 지원하고 CNN은 점진적인 장기 성능 향상을 기대할 수 있어 고속 피드포워드 네트워크 및 익스트림 학습 기계와 협력하여 적응형 및 개선된 딥 러닝 아키텍처를 구축하여 잡음이 많은 데이터 분포에서보다 유연하고 빠른 적응형 학습 시퀀스를 달성할 수 있다.

도면의 간단한 설명

- [0020] 도 1은 본 발명에 따른 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법을 설명하기 위해 (a) 순방향 딥 러닝만 사용하는 빈 기반 SSL과 (b) 순방향 딥 러닝과 결합된 순방향 재선택, 롤백 제거 및 롤백 재지정으로 구성된 EER 롤백 학습의 정보 흐름 스냅 샷이다.
- 도 2는 본 발명에 따른 신속 적응형 객체 감지를 위한 롤백 딥 러닝 아키텍처를 도시한 도면이다.
- 도 3은 Adam과 SGD 옵티마이저를 사용하여 학습율을 다르게 한 실험 결과 그래프이다.
- 도 4는 롤백 기반 ASSL의 정보 흐름도이다.
- 도 5는 레이블링 결과가 있는 노이즈가 많은 이미지 샘플을 도시한 도면이다.
- 도 6은 노이즈가 많은 로컬 이미지를 사용하여 본 발명(EER-ASSL)의 평균 정밀도(AP)로 측정한 성능 비교 그래프이다.
- 도 7은 의자, 소파 및 테이블의 로컬 데이터 세트에 대한 IASSL의 탐지 결과를 도시한 도면이다.
- 도 8은 벤치 마크 데이터 세트에 대한 mAP 측정시 최첨단 객체 감지기와 본 발명(EER-ASSL)의 성능을 비교한 그래프이다.

발명을 실시하기 위한 구체적인 내용

- [0021] 이하, 본 발명에 따른 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법을 첨부한 도면을 참고로 하여 상세히 기술되는 실시 예에 의하여 그 특징들을 이해할 수 있을 것이다.

- [0022] 이에 앞서 본 명세서 및 청구범위에 사용된 용어나 단어는 통상적이거나 사전적인 의미로 한정해서 해석되어서는 아니 되며, 발명자는 그 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념을 적절하게 정의할 수 있다는 원칙에 입각하여 본 발명의 기술적 사상에 부합하는 의미와 개념으로 해석되어야만 한다.
- [0023] 따라서, 본 명세서에 기재된 실시 예와 도면에 도시된 구성은 본 발명의 가장 바람직한 일 실시 예에 불과할 뿐이고, 본 발명의 기술적 사상을 모두 대변하는 것은 아니므로, 본 출원시점에 있어서 이들은 대체할 수 있는 다양한 균등물과 변형 예들이 있을 수 있음을 이해하여야 한다.
- [0025] 본 발명에 따른 신속 적응형 객체 감지를 위한 롤백 딥 러닝 방법은 객체 감지를 위한 신속 적응형 학습 프레임워크인 EER 학습과 딥 러닝 SSL 알고리즘을 결합한 방식(EER-ASSL)을 제안한다. 본 발명은 BDAL(Bi-Directional Active Learning)과 마찬가지로 도 1의 (b)에서와 같이 레이블링된 데이터 세트에서 의심스러운 데이터 샘플을 검사하면 롤백 학습이 수행되어 재선택 또는 레이블 재지정 메커니즘을 통해 모델을 재구성한다. 롤백 빈(Bin) 기반 SSL에서 선택한 배치 샘플은 여러 빈으로 분할되고 각 빈은 빠른 적응을 위해 CNN 딥 러닝 및 EER 기반 롤백 학습에 대한 차별적 지식을 반복적으로 전달한다. 이때 롤백 학습 방법은 빈 기반 SSL에 내장되어 불균형 분포에서 잡음이 있는 불확실한 샘플의 영향을 신속하게 제거한다. 수렴이 이루어질 때까지 위의 과정이 반복된다. 본 발명에서 제안하는 방식(EER-ASSL)의 특징은 다음과 같다.
- [0026] 우선 본 발명의 ERR-ASSL은 변화하는 환경에서 효율적인 객체 감지를 위한 신속 적응형 학습 프레임워크를 제공한다. 이 방법은 대부분의 최첨단 탐지 기술에 채택한 정적 데이터 분포의 가정에 의존하지 않는다. EER 롤백 학습과 CNN 딥 러닝을 결합하여 수렴될 때까지 현재 빈에서 다음 빈으로 모델의 동적 차별적 지식을 전송한다. 따라서 본 발명의 ERR-ASSL은 알려지지 않은 데이터 분포에서 노이즈가 많고 편향된 레이블을 극복할 수 있다. 여기서 EER 모델은 빠른 단기 적응을 가능하게 하고 CNN 모델은 점진적인 장기 성능 향상을 제공하는데 도움이 된다.
- [0027] 그리고 본 발명의 ERR-ASSL에서 EER를 사용하는 롤백 학습 방법은 이상값을 제거하거나 레이블이 잘못된 샘플을 다시 레이블링하여 차별적 기능을 효과적으로 활용한다. 이는 신속 적응형 학습 기능을 제공하는 반면, 도 1의 (a)에서와 같은 단방향 레이블링 및 양방향 접근 방식은 데이터 선택 및 학습 프로세스에서 실행 취소에 시간 소모적인 재실행이 필요하다. 본 발명의 EER 기반 ASSL은 EER의 근시 적응형 학습기능을 효과적으로 활용하고 딥 러닝 객체 감지기의 긴 학습시간의 단점을 극복한다.
- [0028] 아울러 본 발명의 ERR-ASSL 방법은 Pascal VOC 2007, 2012, MS COCO 벤치 마크 데이터 셋 및 로컬 데이터 셋을 사용하여 Faster RCNN, SSD300 및 YOLOv2 와 같은 최첨단 탐지 방법과 비교되는 것으로 오류 수를 대폭 줄이고 잘못된 객체 감지율을 개선한다.
- [0030] 이하에서는 본 발명과 관련된 최신 관련 작업에 대해 간략하게 설명하고 본 발명의 시스템 개요를 상세히 설명하며 이후 본 발명의 EER-ASSL 알고리즘에 대한 세부 정보를 언급하며 그 결과를 보여주고 이후 결론을 설명한다.
- [0032] 우선 본 발명과 관련된 최신 관련 작업을 살펴본다.
- [0033] 먼저 ASSL(Active Semi-Supervised Learning)에 관하여 살펴본다. 많은 경우 레이블링된 데이터 세트가 제공되고 고정 될 것으로 예상하지만, 항상 그런 것은 아니다. 전문 오라클 또는 알고리즘 레이블을 사용하고 능동학습(AL)을 사용하면 레이블링된 데이터 세트가 변경 될 수 있다. 능동학습(AL)은 노이즈가 많은 데이터를 제거하고 더 많은 데이터 유형에 올바르게 레이블링을 할 수 있다. 능동학습(AL)은 더 적은 양의 데이터에 레이블링해야 할 때 중요한 역할을 하고 알고리즘은 레이블링할 시기와 레이블을 지정하지 않을 시기를 결정할 수 있다. Zhu 등은 SSL(Semi-Supervised Learning)과 능동학습(AL:Active Learning)의 효과적인 결합 아이디어를 2003년에 처음 소개했다. 그는 레이블이 없는 데이터가 고조파 에너지 최소화 기능의 위험을 최소화할 수 있도록 AL과 함께 가우스 랜덤 필드 모델을 적용했다. 풀 기반(pool-based) 능동 학습 또는 선택적 샘플링을 사용하여 쿼리 선택 수를 최소화할 수 있다. AL과 SSL을 함께 결합하면 레이블링된 데이터와 레이블이 없는 데이터를 모두 활용하여 분류 성능을 향상시킬 수 있다. Tong과 Koller 등은 SVM의 버전 공간 크기를 줄이는 유사한 연구를 진행한 바 있다. Cohn 등은 분산 성분을 줄여 추정 된 일반화 오류를 최소화한다. 쿼리가 선택되면 이러한 능동학습

방법의 대부분은 많은 양의 레이블링 되지 않은 데이터를 악용할 가능성이 없다. 일부 연구자들은 트레이닝 단계에서 SSL(Semi-Supervised Learning)을 적용했다. Chaloner와 Verdinelli는 베이지안 접근법을 1995년에 적용했다. Gaussian 랜덤 필드는 SSL과 AL의 조합으로 인해 조화로운 기능에 잘 어울린다. ASSL은 쿼리 지점을 효율적으로 추정할 다음 모호한 샘플을 무작위로 선택하고 예상되는 일반화 오류를 추정할 수 있기 때문에 ASSL에는 많은 장점이 있다. 더 나은 레이블 선택 기준은 샘플 분포의 모호성과 불균형을 제거하고 불완전한 레이블링 및 선택 편향을 극복하면서 객체 탐지기를 트레이닝 할 수 있다.

[0034] 다음으로 EER(Expected Error Reduction)은 유지 기반(retaining-based) 능동학습(AL)을 기반으로 새 샘플에 레이블링할 때 선택 기준으로 일반화 오류를 줄인다. 이 방법의 효과는 텍스트 분류 응용 프로그램에서 입증되었다. 이론적 근거는 미래의 일반화 오류를 최소화하는 샘플을 선택하는 것이다. 레이블링 되지 않은 풀은 테스트 분포를 나타내며 검증 세트에 사용된다. 레이블링되지 않은 샘플의 레이블에 대한 지식이 없기 때문에 EER은 평균 사례, 최악 사례 또는 최고 사례의 관점에서 평균 사례 잠재적 손실을 추정한다. 이러한 EER은 현재 모델의 최대 변경으로 이어지는 샘플을 선택하여 예상 오류 모델의 변경을 탐색한다. 유사한 접근 방식은 출력 분산을 최소화하려는 분산 감소 방법에서 찾을 수 있다. 분산 및 로지스틱 회귀를 채택한 후 로지스틱 회귀에 대한 예상 분산 감소 방법으로 확장된다. 최소 최대보기 활성 학습에서 최악의 경우 기준은 새 샘플에 레이블링한 후 목적 함수를 최소화하는 샘플을 선택하도록 조정된다. 이 방법은 목적 함수를 계산할 때 레이블이 없는 모든 데이터를 고려하여 확장된다. 그러나 EER 기반 재교육 접근 방식은 모든 레이블링되지 않은 데이터와 가능한 모든 레이블을 탐색하기 때문에 많은 계산이 필요하다.

[0036] 이하 도 2를 참고로 본 발명의 시스템 개요를 상세히 설명한다.

[0037] 본 발명을 위한 시스템은 신속 적응형 객체 감지를 위한 룰백 딥 러닝 방법의 구현을 위해 공동 샘플링과 EER-ASSL로 구성된다. 본 발명에서 제안된 방법(EER-ASSL)은 EER 방법의 신속한 모델링 기능과 CNN의 정밀하고 점진적인 학습 기능을 통합한다. 본 발명에서 제안된 방법(EER-ASSL)은 AL 알고리즘의 탐색 기능과 룰백 기능이 있는 bin 기반 SSL 알고리즘의 단계별 활용을 이용한다. 이 방법은 교육 시간뿐만 아니라 비용이 많이 드는 인적 노력을 최소화하는 동시에 적응형 학습 프로세스에서 고품질 레이블 데이터 세트와 고정밀 객체 탐지기를 유지한다. 불확실성, 다양성 및 신뢰도 기준을 고려한 공동 샘플링 방법은 중복성이 낮은 보다 정보적이고 신뢰할 수 있는 샘플을 선택할 수 있다. 불확실성 기준은 잡음이 있거나 중복된 샘플의 선택을 유발할 수 있으므로 클러스터링 기반 중복 제거 알고리즘에 다양성 기준이 적용된다. 본 발명에서 공동 샘플링 기반 AL 학습 알고리즘은 룰백 bin 기반 SSL 알고리즘과 통합된다. 동적으로 변화하는 환경에서 신속한 적응형 객체 감지를 위한 EER-ASSL의 간략한 스케치가 도 2에 도시된다. 한 번에 하나의 이미지가 아닌 이미지 스트림의 불확실성과 다양성 샘플링을 기반으로 샘플 배치가 선택된다. 공동 샘플링 후 샘플링된 배치 스트림은 룰백 SSL 알고리즘을 위해 bin(Bin)으로 나뉜다. 이미지 bin의 샘플은 레이블링되지 않은 샘플이며 CNN 검출기와 EER 기반 룰백 학습 방법을 모두 사용하는 bin 기반 SSL에 의해 의사(pseudo) 레이블링된다. 그러나 의사(pseudo) 데이터 세트에는 올바르게 레이블링된 샘플들 뿐만 아니라 불균형 데이터 분포에서 편향된 레이블도 포함한다. 이러한 샘플은 유해한 영향을 미치고 더 나은 객체 감지기를 구축하는 데 어떠한 기여도하지 않기 때문에 잡음이 있는 샘플은 신뢰 샘플에서 제외되어야 한다. 본 발명에서 제안된 EER 기반 학습 방법은 재선택 및 레이블 재지정을 통해보다 유의한 샘플링 및 학습을 하도록 빠른 순방향 학습(forward learning) 및 룰백 학습(rollback learning)을 위해 채택되었다.

[0039] 다음으로 본 발명에서 제안된 방법(EER-ASSL)의 학습을 설명한다.

[0040] 본 발명에서는 EER 모델과 미세 조정된 CNN 모델의 트레이닝을 위해 매우 제한된 수의 레이블링된 데이터를 사용한다. 미세 조정된 CNN 모델 및 EER 모델은 제한된 레이블링된 데이터 샘플을 사용하여 구축된다. 앙상블 네트워크는 CNN 감지기와 EER 모델로 구성되어 객체 감지를 수행한다. 증분 ASSL이 채택되어 입력 데이터 스트림에서 데이터 샘플 배치가 수집되고 공동 샘플링 알고리즘에 의해 선택적으로 샘플링된다. 선택된 샘플은 이미지 캡처 품질에 따라 배치 및 bin의 크기가 결정되는 bin으로 분할된다. CNN 및 EER 앙상블에 의해 처음 레이블링된 데이터 세트와 의사 트레이닝 데이터 세트는 새로운 CNN 및 EER 모델을 트레이닝하는 데 사용된다. 즉, 다음 bin 주기를 위해 모델을 업데이트하는 데 사용된다. 새로운 EER 모델은 필요한 경우 bin에서 샘플 제거, 레이블 변경 및 재선택하는 것으로 구성된 룰백 학습 프로세스에 포함된다. bin 기반 증분 학습은 샘플 재선택을 위한 순방향 학습과 제거 및 레이블 재지정을 위한 룰백 학습과 통합되어 처리된다(도 3 참조). 새로운 CNN 모델은 새로운

공동 샘플링 과정에서도 사용된다. 이 과정은 수렴될 때까지 반복된다.

[0042] 우선 EER 순방향 학습(forward learning) 프로세스를 설명한다. 본 발명에서는 패턴 분류 문제에서 제안되고 사용된 EER 방법을 채택한다. EER 방법의 목적은 향후 단계에서 일반화 오류를 최소화하는 샘플을 선택하는 것이다. 테스트 데이터를 미리 사용할 수 없기 때문에 스트리밍 데이터 세트의 일부가 향후 오류를 추정하기 위한 유효성 검사 데이터 세트로 사용된다. 레이블링되지 않은 데이터 세트의 실제 레이블은 알 수 없으며 향후 오류는 레이블이 없는 데이터에 대한 예상 로그 손실을 사용하여 대략적으로 추정된다. $UD = \{x_i\}_{i=m+1}^n$ 은 레이블링된 트레이닝 데이터 세트를 나타내고, $LD = \{(x_i, y_i)\}_{i=1}^m$ 은 레이블이 없는 데이터 세트이다 (where $m \ll n$). 만약 선택한 샘플 x 가 y 로 레이블링되고 LD에 추가되면, $LD^+ = LD \cup (x, y)$ 로 표시된다. g_{LD} 는 LD의 EER모델과 LD^+ 의 g_{LD^+} 를 나타낸다. 가장 유익한 데이터 샘플은 레이블링되지 않은 데이터 세트를 사용하여 예상 엔트로피를 최소화하여 예상 오류 감소를 극대화하는 것으로 가정한다. 레이블링되지 않은 데이터 세트에서 가장 유익한 데이터 샘플이 다음의 방정식 (1)을 충족하는 샘플로 선택된다.

$$x^* = \underset{x \in UL}{\operatorname{argmin}} \sum_{y \in C} P(y|x; g_{LD}) \times (-\sum_{x' \in UL, y' \in C} P(y'|x'; g_{LD^+}) \log(y', x', g_{LD^+})), \quad (1)$$

[0044]

[0046] 여기서 C는 객체 클래스를 나타내고, 첫 번째 항 $P(y|x; g_{LD})$ 은 현재 모델의 레이블 정보를 나타내고 두 번째 항은 모델 g_{LD^+} 의 레이블이 없는 데이터 UD에 대한 예상 엔트로피의 합이다. 위 방정식 (1)은 직렬 모드 학습 프로세스를 공식화하는 반면 새 모델은 UD에서 각 새 데이터 샘플의 레이블링 직후에 업데이트된다. 그러나 학습의 철저한 반복은 실제로 무거운 계산 오버 헤드로 인해 어려움을 겪는다. 본 발명에서는 레이블이 없는 스트림 데이터 세트를 배치로 나누고, 각 배치는 도 3과 같이 UD에서 레이블이 없는 전체 데이터 샘플을 처리하는 대신 빈으로 분할된다.

[0047] 룰백 SSL에서 레이블링되지 않은 데이터 샘플의 빈이 각 학습 단계에 사용된다. 시간 단계 i 의 경우 방정식 (1)은 빈(bin) B_i 를 고려하여 다음의 방정식 (2)와 같이 다시 작성된다.

$$x_{B_i}^* = \underset{x \in B_i}{\operatorname{argmin}} \sum_{y \in C} P(y|x; g_{LD}) \times (-\sum_{x' \in B_i, y' \in C} P(y'|x'; g_{LD^+}) \log(y', x', g_{LD^+})) \quad (2)$$

[0049]

[0051] 여기서 첫 번째 항은 현재 모델의 레이블 정보를 나타내고, 두 번째 항은 레이블이 없는 데이터 빈(bin) B_i 에서 모델 g_{LD^+} 에 대한 예상 엔트로피의 합이다. 공동 샘플링을 적용한 후 방정식 (2)를 반복적으로 적용하는 빈(bin) 데이터 세트에 대한 유사 레이블 집합 $\Delta_i = \{x_1, \dots, x_k\}$ 을 결정할 수 있다. 그러나 B_i 의 각 데이터 샘플에 대한 모델을 빌드하려면 여전히 많은 계산 오버 헤드가 필요하다. 따라서 방정식 (2)는 다음의 방정식 (3)과 같이 유사 레이블링된 집합 Δ_i 의 선택된 샘플에 대한 모델을 구축하여 근사치를 구한다.

$$\Delta_i = \underset{\{x_1, \dots, x_k\} \in B_i}{\operatorname{argmin}} \sum_{x_i \in B_i, y' \in C} P(y|x; g_{LD}) \times \left(- \sum_{x' \in B_i, y' \in C} P(y'|x'; g_{LD+\Delta_i}) \log(y', x', g_{LD+\Delta_i}) \right), \quad (3)$$

여기서 첫 번째 항은 레이블링된 의사 집합 Δ_i 의 선택된 샘플에 대한 현재 모델의 레이블 정보를 나타내고, 두 번째 항은 가중치 모델 $g_{LD+\Delta_i}$ 과 레이블링 되지 않은 데이터 B_i 에 대한 예상 엔트로피의 합이다.

선택한 샘플 $\{x_1, \dots, x_k\}$ 레이블이 $\{y_1, \dots, y_k\}$ 이고 LD에 추가 된 경우 $LD+\Delta_i = LD \cup \{(x_1, y_1) \dots, (x_k, y_k)\}$ 로 표시된다. 방정식 (1)을 사용하여 레이블이 없는 전체 데이터 샘플에 대한 모델을 빌드하는 대신 모델을 한번 빌드한다는 것을 알 수 있다. EER 순방향 학습 프로세스는 CNN 검출기에 의해 실패한 bin의 재교육을 위한 Δ_i 의 재선택 알고리즘에 사용된다. 현재 레이블링된 데이터 세트에 추가된 재 선택된 샘플과 결합된 데이터 세트는 bin 기반 SSL 단계에 대한 CNN 모델을 유지하는 데 사용된다.

다음으로 EER 롤백 학습(Rollback learning) 프로세서를 설명한다. 이 프로세서의 목표는 현재 모델을 방해하는 가장 불확실한 레이블링 된 샘플을 조사하고 가장 불확실한 샘플을 대체하거나 최근에 의사 레이블링된 샘플을 다시 레이블링하는 새 샘플을 선택하는 것이다. 롤백 학습을 위해 EER 추정을 사용하여 의사 레이블링된 데이터 샘플에 대한 예상 엔트로피를 최소화한다. 계산 시간을 고려할 때 레이블링되지 않은 전체 데이터 세트를 고려하지 않고 가장 최근의 유사 레이블 데이터 세트만 검사한다. 롤백 프로세스에서는 최근 의사 레이블링 된 샘플에서 제거 또는 레이블을 재지정할 후보 샘플을 찾는다. 롤백 학습 프로세스는 롤백 샘플의 레이블을 다시 지정하거나 이웃에서 다시 선택하여 롤백 샘플의 레이블 인증을 수행한다. 롤백 학습은 1) 제거 프로세스와 2) 레이블 재지정 프로세스의 두 가지 유형으로 나뉜다.

우선 제거 프로세스를 살펴보면, 제거 롤백 프로세스는 롤백 샘플이 EER 모델을 방해하는 것으로 의심되기 때문에 레이블링된 데이터 세트에 추가된 현재 의사 데이터 세트에서 데이터 샘플을 제거하여 데이터 샘플의 나쁜 영향을 취소하는 것이다. 이전에 레이블링된 데이터 세트를 사용하는 모델은 마지막 레이블링된 데이터 세트보다 더 안정적인 것으로 예상되므로 모델을 업데이트하는 데 사용된다. 롤백 학습을 제거하는 동안 롤백 샘플의 영향은 모델의 의사 데이터 세트에서 제거된다. 제거 롤백 프로세스는 의사 레이블링된 데이터 세트에서 신뢰할 수 없는 샘플을 삭제하고 신뢰할 수 없는 샘플이 감지된다. 롤백 제거 프로세스는 다음의 방정식 (4), (5)와 같이 구성된다.

$$x^\dagger = \underset{x \in LD}{\operatorname{argmin}} \sum_{y \in C} P(y|x; g_{LD \setminus (x, y^\dagger)}) \times \quad (4)$$

$$\left(- \sum_{x' \in UL, y' \in C} P(y'|x'; g_{LD \setminus (x, y^\dagger)}) \log(y', x', g_{LD \setminus (x, y^\dagger)}) \right), \quad (5)$$

여기서 $LD \setminus (x, y^\dagger)$ 은 $(x, y^\dagger)$ 을 제거한 후 레이블링된 데이터 세트를 나타낸다. 방정식 (4)는 실제로 계산할 수 없는 많은 계산 시간이 필요하기 때문에 롤백 샘플, R_Δ^{remov} 은 레이블링되지 않은 전체 데이터 세트의 샘플 대신 현재 단계의 의사 레이블링된 데이터 샘플에서 선택된다. 트레이닝 데이터 세트는 LD에서 $LD \setminus R_\Delta^{remov}$ 로 롤백하고 현재 기능 공간의 이웃에서 새 샘플을 대체한다. 제거로 인해 레이블링되지 않은 데이터 세트에 대한 엔트로피를 최소화하는 제거할 레이블링된 샘플 풀을 찾으며, 다른 유사 레이블링된 샘플은 재선택 프로세스에

서 다시 선택된다. 제거를 위한 룰백 샘플은 다음과 같이 분류 모델인 방정식 (6)을 사용하여 마지막 의사 레이블링된 샘플의 빈에서만 선택된다.

$$R_{\Delta}^{remov} = \underset{\{x_1, \dots, x_r\} \in \Delta_t}{\operatorname{argmin}} \sum_{x_i \in R_{\Delta}^{remov}, y \in C} P(y|x; g_{LD \setminus R_{\Delta}^{remov}}) \times (-\sum_{x' \in \Delta_t, y' \in C} P(y'|x'; g_{LD \setminus R_{\Delta}^{remov}}) \log(y', x', g_{LD \setminus R_{\Delta}^{remov}})), \quad (6)$$

여기서 R_{Δ}^{remov} 은 재선택 프로세스를 위해 제거할 룰백 샘플을 나타낸다. 선택한 룰백 샘플이 각각 $\{y_1, \dots, y_r\}$ 에 의해 의사 레이블링된 $\{x_1, \dots, x_r\}$ 이고 LD에서 제거된 경우 집합 차이 $LD \setminus R_{\Delta} = LD \setminus \{(x_1, y_1) \dots, (x_r, y_r)\}$ 으로 표시된다.

다음으로 레이블 재지정 프로세스를 살펴보면 R_{Δ} 에서 선택한 레이블 재지정 샘플이 LD에 업데이트되거나 추가된다. 만약 룰백 학습 레이블 재지정 프로세스 후 룰백 샘플의 레이블이 변경되면 순방향 학습 레이블이 새 레이블로 대체된다. 이러한 방식으로 포워드 레이블링 오류가 수정된다. 만약 새 레이블이 정방향 학습 레이블과 동일하면 룰백 샘플이 새 샘플로 처리되고, LD에 추가될 것이다. 부스팅에 대한 유사한 아이디어가 논의되며, 잘못된 레이블이 붙은 샘플 후보가 더 집중 될 것이다. 한 샘플에 대한 레이블 재지정 룰백 학습은 다음의 방정식 (7)과 유사한 공식에 따라 수행된다.

$$x^{\dagger} = \underset{x \in LD}{\operatorname{argmin}} \frac{1}{Z} \sum_{y_i \in C, i \neq i^{\dagger}} P(y_i|x; g_{LD \setminus (x, y_i^{\dagger})}) \times (-\sum_{x' \in UL, y' \in C} P(y'|x'; g_{LD \setminus (x, y_i)}) \log(y', x', g_{LD \setminus (x, y_i)})) \quad (7)$$

여기서 $LD \setminus (x, y_i)$ 는 y_i 에 의해 할당된 의사 레이블링된 x 를 나타낸다. Z는 다음에 의해 계산된 정규화 계수이다.

$$Z = \sum_{i \neq i^{\dagger}, y_i \in C} P(y_i|x; g_{LD \setminus (x, y_i^{\dagger})}) = 1 - (y_i^{\dagger}|x; g_{LD \setminus (x, y_i^{\dagger})})$$

레이블이 재지정된 각 후보에 대한 모델 계산의 계산 오버 헤드를 고려하여 다음과 같이 Δ_i 의 레이블이 재지정된 후보 풀 측면에서 아래의 방정식 (8)에서와 같이 레이블 재지정 룰백 학습 프로세스를 공식화한다.

$$R_{\Delta}^{relab} = \underset{\{x_1, \dots, x_r\} \in \Delta_i}{\operatorname{argmin}} \frac{1}{Z} \sum_{y_i \in C, i \in R_{\Delta}^{relab}} P(y_i|x; g_{LD \setminus R_{\Delta}^{relab}}) \times (-\sum_{x' \in \Delta_i, y' \in C} P(y'|x'; g_{LD \setminus (x, y_i)}) \log(y', x', g_{LD \setminus (x, y_i)})), \quad (8)$$

여기서 $LD \setminus (x, y_i)$ 는 y_i 에 의해 할당된 의사 레이블링된 x 를 나타낸다. Z는 다음에 의해 계산된 정규화 계수이

다.

$$Z = \sum_{i \in R_{\Delta}^{relab}, y_i \in C} P(y_i | x; g_{LD \setminus R_{\Delta}^{relab}}) \\ = 1 - P(y_i^{\dagger} | x; g_{LD \setminus R_{\Delta}^{relab}}), \text{ where } y_i^{\dagger} \in R_{\Delta}^{relab}. \quad (9)$$

이하에서는 본 발명에 따른 방법(EER-ASSL)의 알고리즘을 설명한다.

본 발명에서 제안하는 EER-ASSL의 주요 작업은 공동 샘플링 기반 AL과 롤백 빈 기반 SSL로 구성된다. AL 프로세스에서 데이터 샘플의 배치는 입력 데이터 스트림에서 수집되며, 최소한의 중복성을 가진 정보 샘플에 대한 공동 샘플링 알고리즘에 의해 처리되며 빈으로 분할된다. 롤백 SSL에서 EER 기반 롤백 학습과 빈 기반 SSL은 빠른 적응 학습을 위해 결합된다. 제한된 레이블링된 샘플은 CNN 모델 및 EER 모델을 초기화하는 데 사용된다. 모델은 빈 기반 증분 SSL 체계에서 학습된다. CNN 모델 학습에서 성능 기준을 위반하면 빠른 롤백 학습을 위해 EER 방법이 활성화된다. 신뢰할 수 있는 레이블링된 데이터 세트의 볼륨은 유사 레이블링된 데이터 샘플을 추가하여 증가한다. 확대 된 LD는 다음 CNN 및 EER 모델을 구축하는 데 사용된다. 이 과정은 수렴 될 때까지 반복된다. EER 롤백 모델은 빠른 단기 적응을 제공하고 CNN 검출기 모델은 각각 점진적인 장기 성능 향상을 제공한다는 것을 알 수 있다. D_{div} 는 공동 샘플링 후 현재 배치에서 채굴 된 샘플을 나타낸다. 여기서는 롤백 빈 기반 SSL 알고리즘에 중점을 둔다. D_{Δ} 는 빈 기반 SSL의 기밀 배치 데이터 세트를 나타내며 기밀 배치 데이터 세트의 컨테이너로 사용된다. D_{Δ} 의 카디널리티가 신뢰 매개변수 γ 가 되면 신뢰 표본 선택 프로세스가 중지된다. D_{Δ} 는 $x_{top} = \operatorname{argmax}_{x \in D_{div}} f(x)$, $x_{top} \in D_{div}$ 을 만족하는 샘플로 초기화된다. 기밀 샘플링 전략은 D_{div} 에서 샘플을 선택하고 $x_{top} = \operatorname{argmax}_{x \in D_{div}} \{ \max_{x_i, x_j \in D_{\Delta}} d(x_i, x_j) \}$ 을 사용하여 현재 덩 피쳐 공간의 거리 메트릭에 따라 D_{Δ} 에 추가한다. 여기서 $d(x_i, x_j)$ 는 덩 피쳐 공간에서 두 샘플 x_i 와 x_j 간의 유클리드 거리이다. CNN은 D_{Δ} 에 있는 기밀 샘플의 빈 시퀀스를 사용하여 다시 트레이닝 된다. 기밀 샘플은 빈으로 분할되고 $B(=[B_j]_{j \in B})$ or $= \{B_0, \dots, B_j, \dots, B_J\}$ 으로 표시된 빈 풀에 저장된다. 각 롤백 빈 기반 SSL 단계에서 신뢰도 점수는 현재 CNN 감지기에 의해 의사 샘플에 할당된다. 레이블링된 데이터 D_0 은 각각 처음에 CNN 검출기 모델 f_0 및 EER 모델 g_0 을 초기화하는데 사용된다. Acc_0 은 검증 데이터를 사용하여 f_0 에 의해 계산된다. 각 빈에 대해 각각 $D_0 \cup B_j$ 를 사용하여 CNN 모델 $\{f_0^{B_j}\}_{j=1}^J$ 을 세운다. Acc_1 은 $\{f_0^{B_j}\}_{j=1}^J$ 에 의해 계산된 빈 점수 중 최대 정확도 즉, $Acc_1 = \max_{B_j} \{Acc_0^{B_j}\}$ 를 나타낸다. 만약 성능이 향상 즉 $Acc_1 \geq Acc_0$ 이면, $D_1 = D_0 \cup B^*$ 및 $f_1 = f_0^{B^*}$ 를 업데이트하여 다음 단계로 이동한다. 시간 단계 i 에서 각 빈에 대해 $D_i \cup B_j$ 및 $Acc_{i+1} = \max_{B_j} \{Acc_i^{B_j}\}$ 를 사용하여 CNN 모델 $\{f_i^{B_j}\}_{j=1}^J$ 을 세우고, 세 가지 사례인 사례 1) $Acc_{i+1} \geq Acc_i$, 사례 2) $Acc_i - \tau < Acc_{i+1} < Acc_i$ 및 사례 3) $Acc_{i+1} \leq Acc_i - \tau$ 로 나뉜다. 여기서 τ 은 탐사 가능성에 대한 허용 한계이다.

사례 1: 다음 단계를 위한 최상의 빈을 얻고 $D_{i+1} = D_i \cup B^*$ 및 $B_i = B^*$; $B = B \setminus B_i$ 를 업데이트한다. 선택한 빈을 제거하면 빈 풀 B 이 줄어드는 것을 알 수 있다.

[0091] 사례 2: 다음 하위 단계를 수행한다. 1) 방정식 (7)을 사용한 롤백 학습 프로세스를 사용하여 Δ_i 에서 제거 샘플을 찾고, 2) 레이블 재지정 샘플을 찾고 방정식 (5)를 기반으로 한 롤백 학습 프로세스를 사용하여 Δ_i 에서 새 레이블을 할당하고, 방정식 (3)에 기반한 EER 순방향 학습 프로세스를 사용하여 재선택하여 Δ_i 을 업데이트한다.

[0092] 위의 정방향 롤백 학습 프로세스는 $Acc_{i+1} \geq Acc_i$ 또는 $Acc_{i+1} \leq Acc_i - \tau$ 조건을 만족하기 전 또는 시간 제한까지 반복한다. 만약 조건 $Acc_{i+1} \geq Acc_i$ 이 충족되면, $D_{i+1} = D_i \cup \Delta_j, f_{i+1} = f_i^{\Delta_j}, g_{i+1} = g_i^{\Delta_j}$ 와 $B = B \setminus B_i$ 를 업데이트한다.

[0094] 사례 3: Oracle은 B^* 의 레이블링된 데이터를 잘못 레이블하고 $f_{i+1} = f_i, g_{i+1} = g_i, D_{i+1}$ 을 업데이트한다.

[0095] 사례 2의 롤백 프로세스는 Oracle 레이블링 단계를 현저하게 줄일 수 있다. $(D_i \cup \Delta_i)$ 는 시간 t에서 f_{i+1} 및 g_{i+1} 의 트레이닝에 사용되는 트레이닝 데이터 세트 D_{i+1} 을 세우는데 사용된다. 이 과정은 수렴 될 때까지 반복된다. 마지막으로 롤백 빈 기반 SSL은 두 가지 모델 f 및 g와 확장된 레이블 데이터 세트 LD를 생성한다. EER 기반 롤백 학습과 빈 기반 SSL을 결합하면 동적으로 변화하는 환경에서 노이즈가 있는 스트리밍 샘플에서도 신속 적응형 객체 감지기를 얻을 수 있다. 롤백 빈 기반 SSL 알고리즘은 아래의 알고리즘 1에 요약된다.

Algorithm 1. Rollback bin-based SSL
Input: bin pool B
Output: CNN model f , EER model g , and labeled dataset LD .
Repeat until $B \neq \phi$
 1. For each bin, $B_i \in B$, build f_{i+1} using $D_i \cup B_i$, and calculate $Acc_i^{B_i}$.
 2. $Acc_{i+1} = \max_{B_j} \{Acc_i^{B_j}\}$.
 3. If $Acc_{i+1} \geq Acc_i$,
 $B^* = \operatorname{argmax}_{B_j} \{Acc_i^{B_j}\}, D_{i+1} = D_i \cup B^*;$
 $f_{i+1} = f_i^{B^*}; B_i = B^*$, and $B = B \setminus B_i$.
 4. Else if $Acc_i - \tau < Acc_{i+1} < Acc_i$,
 While $Acc_i - \tau < Acc_{i+1} < Acc_i$,
 4.1 Remove the samples from Δ_i , i.e., the removing rollback process using Eq. (5).
 4.2 Relabel the samples in Δ_i i.e., the relabeling rollback process using Eq. (7).
 4.3 Reselect the samples from B_i using the forward learning process using Eq. (3).
 4.4. If $Acc_{i+1} \geq Acc_i$, $D_{i+1} = D_i \cup \Delta_j, f_{i+1} = f_i^{\Delta_j}, g_{i+1} = g_i^{\Delta_j}$, and $B = B \setminus B_i, i++$
 Else if $Acc_{i+1} < Acc_i - \tau$ or time limit, oracle labels incorrectly labeled data in B^* .
 $f_{i+1} = f_i, g_{i+1} = g_i, D_{i+1}, i++$.
Return $\{f = f_{i+1}, g = g_{i+1}, LD = D_{i+1}\}$

[0097]

[0099] 이하, 본 발명에 따른 방법(EER-ASSL) 등의 실험 예를 설명한다.

[0100] 본 발명에 따른 방법(EER-ASSL)은 PASCAL VOC 및 로컬 데이터 세트와 같은 벤치 마크 데이터 세트를 사용하여 광범위한 실험을 수행하고 성능을 Faster RCNN, SSD300 및 YOLOv2와 같은 최첨단 감지기 기술과 비교한다. 실험적 구현은 cuDNN 및 Tensorflow를 갖는 단일 NVIDIA TITAN X가 있는 단일 서버를 사용하여 수행되었다. 실험 설정을 Darknet-19 CNN 모델로 사용했으며 기본 감지기는 최첨단 객체 감지기인 YOLOv2이다.

[0102] <벤치 마크 데이터 세트>

[0103] - PASCAL VOC 데이터 세트: 유명한 PASCAL VOC 벤치 마크에는 Pascal VOC 2007 및 2012의 두 가지 버전이 있다. Pascal VOC 2007은 총 9963 개의 이미지(train/validation/test)로 구성된 20개의 클래스와 24,640 개의 주석이 달린 객체로 구성된다. Pascal VOC 2012에는 27,450 개의 주석이 달린 객체가 포함 된 11,530 개의 이미지 (train/validation/test)가 있는 20 개의 클래스가 있다. YOLOv2 모델은 PASCAL VOC 2007 데이터 세트와 PASCAL VOC 2012 데이터 세트를 사용하여 트레이닝 되었다. Pascal VOC 2007 데이터 세트에는 사람, 동물, 차량, 실내의 네 가지 수퍼 클래스가 있다. 본 발명의 실험은 실내 환경(예: 병, 의자, 식탁, 화분, 소파 및 TV 모니터)의 개체 등급에 초점을 맞추었다.

[0104] - 로컬 데이터 세트: 450개의 의자 이미지, 450개의 화분 이미지, 450개 개찰구 이미지 및 450 개의 테이블 이미지 데이터 세트가 로컬 영역에서 선택된다. 본 발명은 416 × 416 픽셀의 입력 이미지 해상도를 사용한다. 12 개의 의자 이미지를 초기 레이블 데이터 샘플로 사용했다. 98개의 이미지가 무작위로 선택되어 유효성 검사 데이터 세트에 사용된다. 각 클래스의 나머지 340개 이미지는 실험을 위해 레이블링되지 않은 데이터 세트에 사용되었다. PASCAL VOC 데이터 세트를 사용하여 트레이닝 할 때, 로컬 데이터 세트는 YOLOv2에서 매우 열악한 감지 결과를 생성한다.

[0106] <실험 매개 변수 설정>

[0107] 본 발명의 객체 감지 방법은 PASCAL VOC 챌린지 평가를 사용했다. 이것은 균일한 간격의 리콜 레벨 세트에 걸쳐 정밀도를 평균하여 아래의 방정식 (10)의 평균 정밀도가 계산되는 경우에 적용된다. 아래의 방정식 (10)은 평균 정밀도에 관한 식이다.

[0109]
$$AveP = \frac{1}{11} \sum_{r \in \{0.0, 0.1, \dots, 1.0\}} P_{interp}^{(r)} \quad (10)$$

[0111] 여기서 $P_{interp}^{(r)}$ 은 r보다 큰 모든 재현율에 대해 최대 정밀도를 취하는 보간 정밀도이다. 본 발명 등에서는 실측 자료와 예측이라는 두 경계 사이의 중첩을 계산하기 위해 IoU (Intersection over Union)를 사용한다. 도 7에서와 같이 적색 경계 박스는 Ground Truth, 검은색 경계 박스는 EER-ASSL, 노란색 경계 박스는 YOLOv2 VOC 모델을 나타낸다. 실험에서 IoU 임계 값은 0.5(IoU => 0.5)로 미리 정의된 것으로 간주한다. 예측 성능이 임계 값을 초과하면 올바른 것으로 간주하고 그렇지 않으면 잘못된 것으로 간주된다. 최상의 성능을 얻기 위해 기울기 기반 최적화 방법 Adam과 SGD(stochastic gradient descent)를 적용한다. 이 두 옵티마이저에 대해 학습률 0.001을 선택한다. 그러나 SGD 옵티마이저는 도 3에서 볼 수 있듯이 500 epoch의 Adam에 비해 실험에서 훨씬 느리다.

[0112] 도 3은 Adam 및 SGD 최적화 프로그램을 사용하여 로컬 데이터 세트를 트레이닝하는 동안 다양한 학습율(Learning Rate)의 실험 결과를 나타낸다. 본 발명은 학습률 0.001이 더 높고 안정적인 성능을 가지므로 학습률이 0.01 및 0.001 인 Adam 옵티마이저를 선택한다. 트레이닝 과정은 1 단계와 2 단계와 같은 여러 단계로 나뉜다. SGD와 Adam 최적화 프로그램은 서로 다른 수의 빈을 트레이닝하는 데 사용되었다. 두 단계의 경우 총 빈 수는 20개이다. 이 실험은 Adam 옵티마이저가 AP가 더 높은 SGD 옵티마이저보다 더 빠른 (시간)수렴을 가지고 있음을 보여준다. 이러한 이유로 학습률이 0.001 인 Adam 최적화 프로그램을 선택했다. 레이블링된 전체 데이터 세트를 두 단계로 나누었다. 1단계에서 매개 변수 조합은 불확실성, 다양성 및 신뢰도 [0.8, 0.8, 0.8]이다. 성능에 따라 매개 변수 조합을 [0.8, 0.6, 0.8] 또는 [0.8, 0.8, 0.6]으로 변경한다.

[0114] <본 발명에 따른 방법(EER-ASSL)의 실험 효과>

[0115] 본 발명에서 제안한 EER-ASSL, IASSL 및 단순 SSL의 성능은 도 4에서 비교된다. 모든 실험에 대한 불확실성, 다양성 및 신뢰도를 위해 협업 샘플링 매개 변수는 [0.8, 0.6, 0.8]로 설정된다. Adam 옵티마이저는 학습율 0.001로 사용되었다. EER-ASSL이 증분 ASSL에 비해 훨씬 향상된 성능을 가지고 있음을 알 수 있다. 단순 SSL의 성능은 거의 향상되지 않았다.

[0117] <노이즈가 있는 로컬 이미지 테스트>

[0118] 도 6의 로컬 노이즈 이미지를 사용하여 노이즈 이미지에서 훨씬 개선된 EER-ASSL, IASSL 및 ICL(Iterative Cross learning)을 비교했다. ICL은 이미지 분류 성능으로 테스트되는 반면 EER-ASSL은 경계 박스 위치와 클래스 레이블이 모두 노이즈가 있는 객체 감지 성능으로 평가된다. Adam Optimizer의 학습율은 0.001로 설정된다. 실험 결과는 도 3과 같다. 처음에는 EER-ASSL이 노이즈 데이터의 영향을 받는 낮은 성능(AP)을 보여 주지만 첫 번째 단계의 4 번째 빈 이후에는 다른 방법보다 성능이 우수하다. 실험 결과는 도 7에 반영되어 있다. EER-ASSL은 다양한 조명 변화에서 뛰어난 성능을 보여준다. 도 7은 의자, 소파 및 테이블의 로컬 데이터 세트에 대한 IASSL의 탐지 결과로서, 모든 경우에 적색 경계 박스는 실측 자료, 검은색 경계 박스는 EER-ASSL, 노란색 경계 박스는 YOLOv2 VOC 모델을 나타낸다.

[0120] <본 발명에 따른 방법(EER-ASSL)과 최첨단 기술과의 비교>

[0121] 본 발명에 따른 방법(EER-ASSL)은 여러 첨단 객체 탐지기와 비교된다. 4 개의 개체 각각에는 100 개의 레이블링 된 데이터 샘플과 300 개의 레이블링 되지 않은 데이터 샘플이 있으며, 공정한 평가를 위해 로컬 의자, 소파 및 테이블 이미지가 PASCAL VOC 테스트 데이터와 혼합되었다. 탐지기는 공정한 평가를 위해 동일한 벤치마크 데이터 세트와 로컬 데이터 세트로 트레이닝 되었다. 하기의 표 1은 각 열이 벤치마크 데이터와 로컬 데이터의 구성 비율을 나타내는 비교 결과를 보여준다.

[0123] 표 1. mAP 측정 측면에서 최첨단 객체 감지기와 EER-ASSL의 성능 비교

Method	(0,100)*	(10,90)	(25,75)	(50,50)	(75,25)	(90,10)
Faster RCNN	67.7	67.9	68.1	68.3	68.9	69.3
SSD300	68.4	68.8	69.1	69.3	69.8	70.3
YOLOv2	70.1	70.4	70.6	71.1	71.5	71.5
EER-ASSL	71.9	71.7	70.0	69.9	68.6	67.6

*In the first row (a, b) indicates the composition ratio for all data, in which "a" represents VOC 2007 test data, and "b" represents the ratio of local data

[0124]

[0126] YOLOv2는 ImageNet 1000 클래스 분류 데이터 세트를 사용하여 네트워크를 트레이닝 시킨 다음 탐지를 수행하기 위해 네트워크를 수정한다. 이 공동 트레이닝 분류 및 감지 데이터는 소파 또는 개찰구와 같이 각 클래스에 대해 100개의 이미지로 제한되는 로컬 데이터 세트보다 훨씬 크다. 결과적으로 로컬 학습 데이터를 줄이면 실험에서 YOLOv2 모델에 큰 영향을 주지 않는 반면에 본 발명의 EER-ASSL 모델은 이미 로컬 데이터를 채택했으며 구성 데이터 비율이 50% 이상이면 EER-ASSL은 추가열 (10,90) 및 (25,75)에서 볼 수 있는 다른 최첨단 방법을 능가한다.

[0127] mAP 측면에서 EER-ASSL과 Faster RCNN, SSD 300 및 YOLOv2와 같은 최첨단 객체 감지기를 비교한 결과는 하기의 표 2와 같다. 실험은 Fast RCNN 및 Faster RCNN을 사용한 점진적 학습을 고려하여 이전 작업과 유사하게 수행되었다. 마찬가지로 본 발명은 로컬 및 벤치 마크 데이터 세트를 모두 사용하여 기준 감지기를 YOLOv2로 사용한다. PASCAL VOC 2007 테스트 데이터와 로컬 데이터가 모두 사용된다. 공동 샘플링 매개 변수는 불확실성, 다양성 및 신뢰도에 대해 각각 0.8, 0.6 및 0.8로 설정된다. EER-ASSL은 다른 객체 탐지기보다 크게

향상되었다.

[0129] 표 2. 로컬 데이터 세트의 mAP 측정시 EER-ASSL 성능

Method (local data)	mAP	Batch size	Trained-on	Speed (fps)	#Boxes	Input resolution
Faster RCNN	67.7	1	07+12	5	~6000	~1000×600
SSD300	68.4	8	07+12	59	8732	300×300
YOLOv2	71.0	1	07+12	67	845	416×416
Ours (EER-ASSL)	71.9	1	07+12+local	42	845	416×416

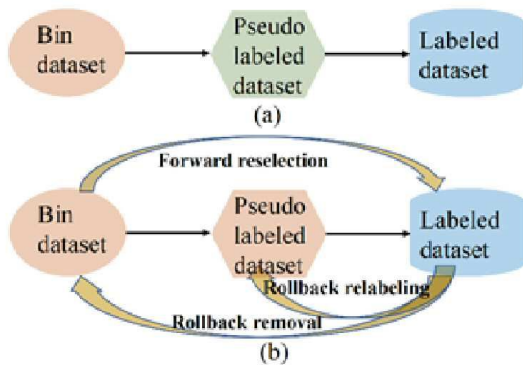
[0130]

[0132] 상기 표 2는 로컬 테스트 데이터 세트에 대한 최신 방법의 평균 정밀도를 요약한다. 본 발명에서 제안한 EER-ASSL 방법은 유사한 환경에서 소파 또는 개찰구와 같은 새로운 객체(물체)를 사용하여 더 높은 mAP 결과에서 성능 향상을 보여준다. 표 2에 표시된 것처럼 EER-ASSL 방법의 적응형 속성은 로컬 데이터 세트 및 해당 환경에서 더 빠른 계산 속도로 탐지 성능을 크게 향상시킨다.

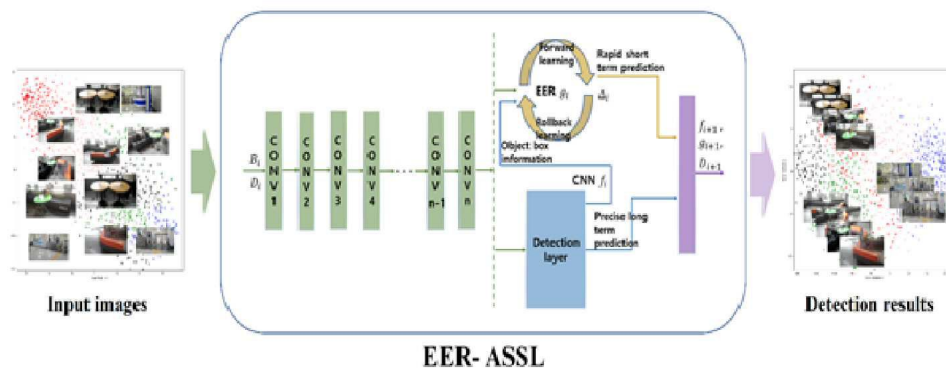
[0134] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형 가능한 것으로, 본 발명의 보호범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

도면

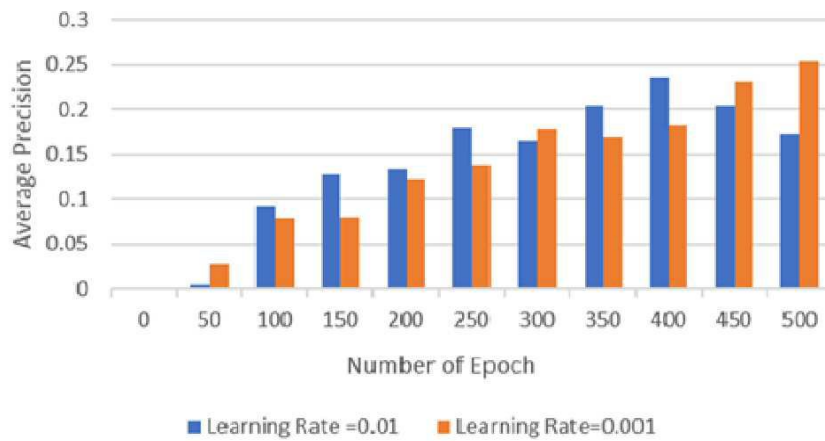
도면1



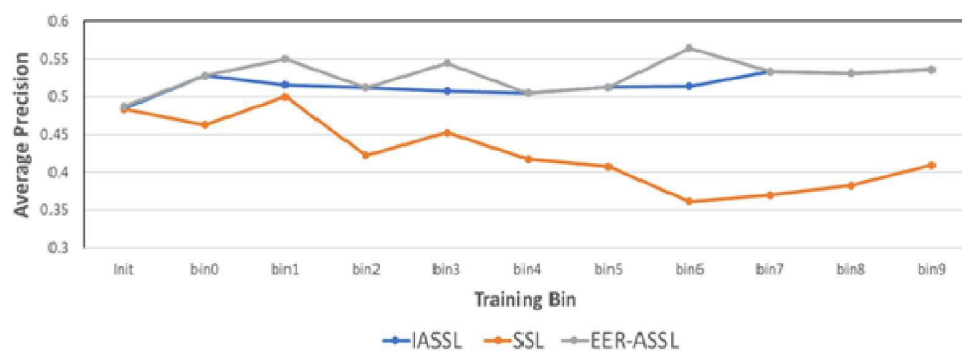
도면2



도면3



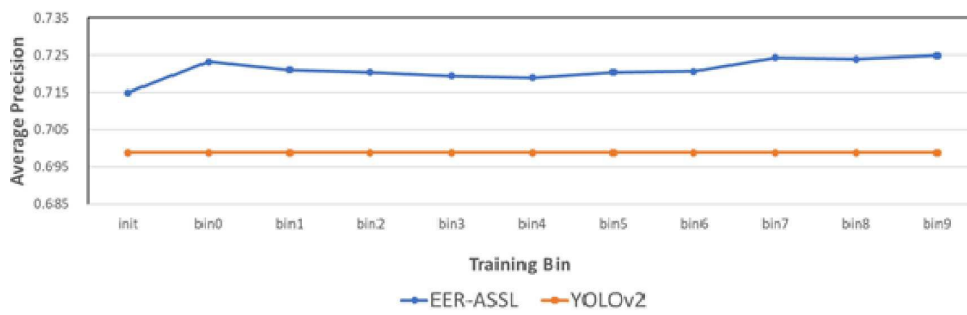
도면4



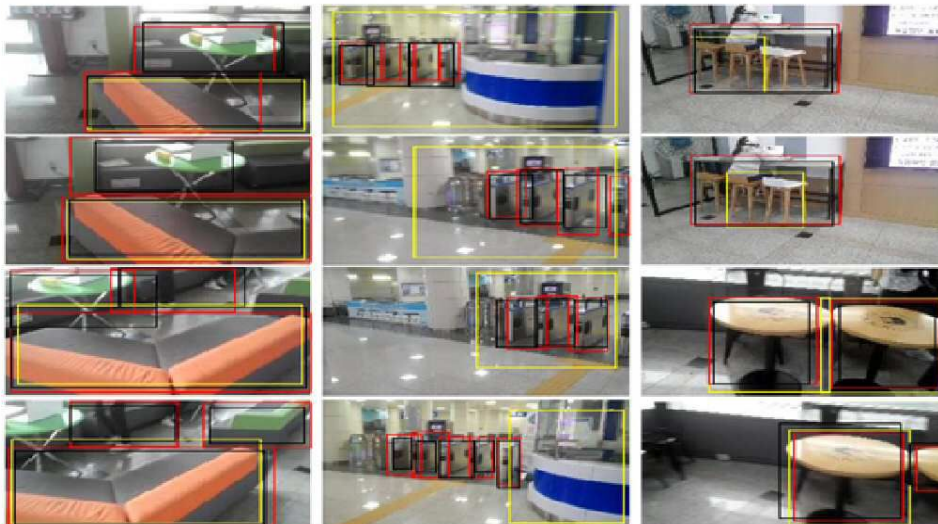
도면5



도면6



도면7



도면8

