



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0060938
(43) 공개일자 2021년05월27일

(51) 국제특허분류(Int. Cl.)
G06T 19/00 (2011.01) G06N 20/00 (2019.01)
G06T 19/20 (2011.01) G06T 7/11 (2017.01)
G06T 7/70 (2017.01)
(52) CPC특허분류
G06T 19/006 (2013.01)
G06N 20/00 (2019.01)
(21) 출원번호 10-2019-0148532
(22) 출원일자 2019년11월19일
심사청구일자 2019년11월19일

(71) 출원인
서강대학교산학협력단
서울특별시 마포구 백범로 35 (신수동, 서강대학교)
(72) 발명자
강석주
서울특별시 마포구 백범로 35 서강대학교 최양업관 419호
박재서
서울특별시 마포구 백범로 35 서강대학교 최양업관 419호
허준호
서울특별시 마포구 백범로 35 서강대학교 최양업관 419호
(74) 대리인
이철희

전체 청구항 수 : 총 11 항

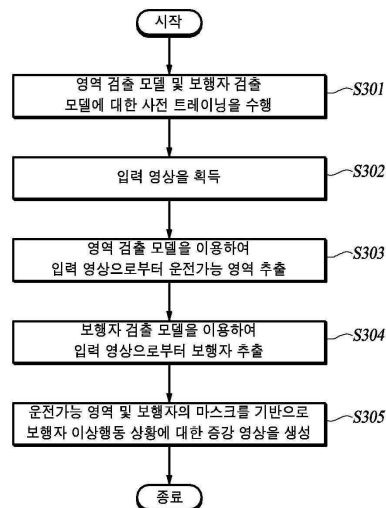
(54) 발명의 명칭 딥러닝 기반 보행자 영상 데이터 증강방법

(57) 요약

딥러닝 기반 보행자 영상 데이터 증강방법을 개시한다.

본 실시예는, 딥러닝(deep learning) 기반 객체 검출 모델(object detection models)을 이용하여 입력 영상으로부터 운전가능 영역(drivable area) 및 보행자(pedestrian)를 검출한다. 검출된 운전가능 영역 및 보행자에 대한 마스크(masks)를 기반으로 보행자 이상행동(abnormal behavior) 상황에 대한 증강 영상을 자동으로 생성하는 것이 가능한 보행자 영상 데이터 증강방법을 제공한다.

대표도 - 도3



(52) CPC특허분류

G06T 19/20 (2013.01)

G06T 7/11 (2017.01)

G06T 7/70 (2017.01)

G06T 2207/20081 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	19PQW0-B153359-01-000000
부처명	행정안전부
과제관리(전문)기관명	국토교통과학기술진흥원
연구사업명	스마트 도로조명 플랫폼 개발 및 실증연구
연구과제명	(4세부) 스마트 도로조명 활용 도시재난안전관리 연계 기술 개발
기 여 율	1/2
과제수행기관명	한국건설기술연구원
연구기간	2019.05.10 ~ 2019.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711093078
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	대학ICT연구센터지원사업
연구과제명	인공지능 서비스 실현을 위한 지능형 반도체 설계 핵심기술 개발
기 여 율	1/2
과제수행기관명	서강대학교 산학협력단
연구기간	2018.06.01 ~ 2019.12.31

명세서

청구범위

청구항 1

보행자 영상 데이터 증강장치를 이용하는 영상 증강방법에 있어서,

입력 영상(input image)을 획득하는 과정;

딥러닝(deep learning) 기반 영역 검출 모델(area detection model)을 이용하여 상기 입력 영상으로부터 운전 가능 영역(drivable area)을 추출하는 과정;

딥러닝 기반 보행자 검출 모델(pedestrian detection model)을 이용하여 상기 입력 영상으로부터 보행자(pedestrian)를 추출하는 과정; 및

상기 운전가능 영역 및 상기 보행자에 대한 마스크(masks)를 기반으로 상기 보행자의 이상행동(abnormal behavior) 상황에 대한 증강 영상(augmented images)을 생성하는 과정

을 포함하는 것을 특징으로 하는, 컴퓨터 상에 구현되는 영상 증강방법.

청구항 2

제1항에 있어서,

학습용 데이터세트(datasets)를 기반으로 상기 영역 검출 모델 및 상기 보행자 검출 모델에 대한 사전 트레이닝을 수행하는 과정을 더 포함하는 것을 특징으로 하는, 컴퓨터 상에 구현되는 영상 증강방법.

청구항 3

제1항에 있어서,

상기 입력 영상은,

상기 운전가능 영역 및/또는 적어도 한 명의 보행자를 포함하는 것을 특징으로 하는, 컴퓨터 상에 구현되는 영상 증강방법.

청구항 4

제1항에 있어서,

상기 입력 영상이 복수의 보행자를 포함하는 경우, 상기 입력 영상 중에 상기 운전가능 영역을 제외한 영역에서 가장 먼저 추출된 한 명의 보행자를 선택하는 것을 특징으로 하는, 컴퓨터 상에 구현되는 영상 증강방법.

청구항 5

제1항에 있어서,

상기 증강 영상을 생성하는 과정은,

상기 운전가능 영역에 대한 마스크 내에 상기 보행자에 대한 마스크를 위치시키는 것을 특징으로 하는, 컴퓨터 상에 구현되는 영상 증강방법.

청구항 6

제1항에 있어서,

상기 증강 영상을 생성하는 과정은,

상기 보행자의 마스크 및 가로 플립(horizontal flip)이 적용된 보행자의 마스크와 상기 운전가능 영역의 마스크를 상호 결합함으로써 상기 보행자의 이상행동 상황에 대한 증강 영상을 생성하는 것을 특징으로 하는, 컴퓨터 상에 구현되는 영상 증강방법.

청구항 7

제1항에 있어서,

상기 증강 영상은,

비정상 상황으로 분류되고, 상기 이상행동을 실행하는 보행자의 위치 정보에 대한 레이블을 포함하는 것을 특징으로 하는, 컴퓨터 상에 구현되는 영상 증강방법.

청구항 8

입력 영상(image)을 획득하는 입력부;

영역 검출 모델(area detection model)을 이용하여 상기 입력 영상으로부터 운전가능 영역(drivable area)을 추출하는 영역추출부;

보행자 검출 모델(pedestrian detection model)을 이용하여 상기 입력 영상으로부터 보행자(pedestrian)를 추출하는 보행자추출부; 및

상기 운전가능 영역 및 상기 보행자 마스크(masks)를 기반으로 상기 보행자의 이상행동(abnormal behavior) 상황에 대한 증강 영상(augmented images)을 생성하는 데이터증강부

를 포함하는 것을 특징으로 하는 보행자 영상 데이터 증강장치.

청구항 9

제8항에 있어서,

상기 영역 검출 모델 및 상기 보행자 검출 모델은 딥러닝(deep learning) 기반의 신경망으로 구현되되, 학습용 데이터세트(datasets)를 기반으로 사전에 트레이닝되는 것을 특징으로 하는 보행자 영상 데이터 증강장치.

청구항 10

제8항에 있어서,

상기 증강 영상은,

비정상 상황으로 분류되고, 상기 이상행동을 실행하는 보행자의 위치 정보에 대한 레이블을 포함하는 것을 특징으로 하는 보행자 영상 데이터 증강장치.

청구항 11

제1항 내지 제7항 중 어느 한 항에 따른 영상 증강방법이 포함하는 각 단계를 실행시키기 위하여 컴퓨터로 읽을 수 있는 기록매체에 저장된 컴퓨터프로그램.

발명의 설명

기술 분야

[0001] 본 발명은 딥러닝 기반 보행자 영상 데이터 증강방법에 관한 것이다.

배경 기술

[0002] 이하에 기술되는 내용은 단순히 본 발명과 관련되는 배경 정보만을 제공할 뿐 종래기술을 구성하는 것이 아니다.

[0003] 도로 교통안전(road traffic safety)에 관련된 정보를 수집하고 분석하는 지능형 시스템(intellectual system)이 널리 연구되고 있다. 지능형 교통안전 시스템은 CCTV(Closed-Circuit Television)로 촬영된 영상을 기반으로, 도로에서 발생하는 비정상적인 상황을 자동으로 탐지하는 도로 이상 탐지(road anomaly detection)를 수행한다. 비정상적인 상황은 그 발생 주기가 매우 긴 경우가 대부분이므로, 사람이 비정상적인 상황을 지속적으로 감시하고, 탐지하는 것은 매우 어렵다. 따라서 지능형 교통안전 시스템은 보행자의 이상행동과 같은 비정상적인 상황을 자동으로 탐지함으로써 사람이 지속적으로 감시할 필요를 줄이는 것을 목표로 한다.

- [0004] 한편, CNN(Convolutional Neural Network)으로 대표되는 딥러닝(deep learning) 모델 기반의 영상처리(image processing) 기술이 지능형 시스템에 널리 이용되고 있다. 딥러닝 모델의 구조 또는 학습 방법 등이 도로 이상 탐지에 대한 성능을 결정할 수 있으므로, 딥러닝 모델을 채용하는 지능형 시스템은 주로 딥러닝 모델의 이용 및 개선에 집중하였다. 그러나, 대다수의 지능형 시스템은 이용되는 딥러닝 모델의 학습에 사용되는 데이터세트(dataset)의 양적 또는 질적 제약이라는 문제점을 안고 있다. 따라서, 일부 지능형 시스템은 데이터세트에 대한 연구를 병행하기도 한다(비특허문헌 3 참조).
- [0005] 보행자 관련 데이터세트로 가장 널리 이용되는 것은 UCSD(University of California, San Diego)의 보행자 이상(pedestrian anomaly) 데이터세트이다(비특허문헌 1 참조). UCSD 데이터세트는 두 개의 서로 다른 정적 카메라로 촬영된 비디오(각각 Ped1 및 Ped2로 호칭함)를 포함하며 각각의 비디오는 보행자 통로를 보여준다. Ped1 비디오는 238 x 158 해상도의 34 개의 학습용 비디오 및 36 개의 평가용 비디오를 포함하고, Ped2 비디오는 360 x 240 해상도의 16 개의 학습용 비디오 및 12 개의 평가용 비디오를 포함한다. 각 비디오는 120 ~ 200 개의 프레임으로 구성된다.
- [0006] UCSD 데이터세트의 학습용 비디오는 보행자 통로를 따라 걷는 사람들이 포함된 이미지이다. 평가용 비디오는 비정상(abnormal) 프레임을 포함하며, 프레임 별로 각각 비정상 객체의 위치에 대한 픽셀 단위 공간 레이블(pixelwise spatial labels)이 있다.
- [0007] UCSD 데이터세트는 각 비디오를 구성하는 전체 프레임의 수가 적고 비정상 상황의 종류가 한정적이라는 단점을 지닌다. UCSD 데이터세트는 가방을 위로 던지는 행동과 같은 5 가지 유형의 비정상 상황을 포함하는데, 전체 프레임 대비 비정상 상황에 대한 프레임 수가 많으므로, 비디오에 포함된 상황이 부자연스럽다. 또한 위치 정보에 대한 레이블이 일부 테스트 비디오에 대해서만 제공된다는 단점이 있다.
- [0008] 다른 데이터세트로, 주로 개찰구를 통해 출입하는 사람들을 주목하는, 지하철 입구 및 출구에 대한 두 개의 긴 비디오를 포함하는 지하철 데이터세트(subway dataset)가 있다(비특허문헌 2 참조). 지하철 데이터세트에서의 비정상적 상황은 개찰구 주변을 뛰어 다니는 사람, 잘못된 방향으로 걷는 사람, 또는 벽을 청소하는 사람을 포함한다.
- [0009] 지하철 데이터세트에는 두 개의 긴 비디오만 제공하므로, 프레임을 추출하는 속도에 대한 정의가 추가로 필요하며, 정확히 어떤 프레임이 비정상으로 표시되는지도 알 수 없다. 또한 비정상 상황의 대상이 되는 객체의 위치 정보에 대한 레이블이 없다는 단점이 있다.
- [0010] 또다른 데이터세트로, 보행자 통로를 포함하는 건물에 대하여, 건물의 측면을 바라보는 단일 야외 카메라에서 찍은 짧은 비디오 클립을 포함하는 CUHK(Chinese University of Hong Kong)의 어베뉴(avenue) 데이터세트가 있다(비특허문헌 3 참조). CUHK의 어베뉴 데이터세트는 건물을 출입하는 사람들이 포함된, 640 x 360 해상도를 가지는 16 개의 학습용 비디오와 21 개의 평가용 비디오를 포함한다. 평가용 비디오에는 총 47 개의 이상 이벤트가 포함되나, 특이한 행동들이 많이 발생하기 때문에 비디오에 포함된 상황이 부자연스럽다라는 단점이 있다.
- [0011] 또다른 데이터세트로, 야외 필드, 야외 마당 또는 실내 로비 주변에서 촬영된 세 가지 짧은 동영상 클립 11 개를 포함하는 UMN(University of Minnesota) 데이터세트가 있다(비특허문헌 4 참조). 각 클립 영상에서의 비정상 상황 영상은 대피 시나리오로서 갑자기 도망가는 사람을 포함한다. 각 동영상 클립은 한 차례의 비정상 상황 이벤트가 발생하는 시점을 포함한다.
- [0012] UMN 데이터세트는 학습용 프레임과 평가용 프레임 간의 분할이 명확하지 않다. 또한 비정상 상황에 대한 레이블은 각 동영상의 전체 프레임에 대해 일시적으로만 제공되므로 학습에 사용하기에 적절하지 못하다는 단점이 있다.
- [0013] 전술한 문제점 외에도 대부분의 공개된 데이터세트는 보행자의 위치에 따른 이상행동, 예컨대 보행자가 횡단보도가 아닌 차도를 건너는 무단횡단(jaywalking)과 같은 교통사고 위험 상황에 해당하는 데이터를 포함하지 않는다. 따라서 무단횡단과 같은 상황을 검출하기 위한 다양한 학습용 데이터를 생성하는 방법을 필요로 한다.

선행기술문헌

비특허문헌

- [0014] (비특허문헌 0001) 비특허문헌 1: W. Li, V. Mahadevan, and N. Vasconcelos. Anomaly detection and localization in crowded scenes. PAMI, 2014.

(비특허문헌 0002) 비특허문헌 2: A. Adam, E. Rivlin, I. Shimshoni, and D. Reinitz. Robust real-time unusual event detection using multiple fixed-location monitors. PAMI, 2008.

(비특허문헌 0003) 비특허문헌 3: C. Lu, J. Shi, and J. Jia. Abnormal event detection at 150 fps in matlab. In ICCV, 2013.

(비특허문헌 0004) 비특허문헌 4: Mehran, Ramin, Alexis Oyama, and Mubarak Shah. Abnormal crowd behavior detection using social force model. 2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2009.

(비특허문헌 0005) 비특허문헌 5: He, Kaiming, et al. "Mask r-cnn." Proceedings of the IEEE international conference on computer vision. 2017.

(비특허문헌 0006) 비특허문헌 6: Szegedy, Christian, et al "Going deeper with convolutions" Proceedings of the IEEE conference on computer vision and pattern recognition 2015.

발명의 내용

해결하려는 과제

[0015] 본 개시는, 딥러닝(deep learning) 기반 객체 검출 모델(object detection models)을 이용하여 입력 영상으로부터 운전가능 영역(drivable area) 및 보행자(pedestrian)를 검출한다. 검출된 운전가능 영역 및 보행자에 대한 마스크(masks)를 기반으로 보행자 이상행동(abnormal behavior) 상황에 대한 증강 영상을 자동으로 생성함으로써, 지능형 교통안전 시스템(intelligent traffic safety system)의 도로 이상(road anomaly)에 대한 탐지 성능(detection performance) 향상이 가능한 보행자 영상 데이터 증강방법을 제공하는 데 주된 목적이 있다.

과제의 해결 수단

[0016] 본 발명의 실시예에 따르면, 보행자 영상 데이터 증강장치를 이용하는 영상 증강방법에 있어서, 입력 영상(input image)을 획득하는 과정; 딥러닝(deep learning) 기반 영역 검출 모델(area detection model)을 이용하여 상기 입력 영상으로부터 운전가능 영역(drivable area)을 추출하는 과정; 딥러닝 기반 보행자 검출 모델(pedestrian detection model)을 이용하여 상기 입력 영상으로부터 보행자(pedestrian)를 추출하는 과정; 및 상기 운전가능 영역 및 상기 보행자에 대한 마스크(masks)를 기반으로 상기 보행자의 이상행동(abnormal behavior) 상황에 대한 증강 영상(augmented images)을 생성하는 과정을 포함하는 것을 특징으로 하는, 컴퓨터 상에 구현되는 영상 증강방법을 제공한다.

[0017] 본 발명의 다른 실시예에 따르면, 입력 영상(image)을 획득하는 입력부; 영역 검출 모델(area detection model)을 이용하여 상기 입력 영상으로부터 운전가능 영역(drivable area)을 추출하는 영역추출부; 보행자 검출 모델(pedestrian detection model)을 이용하여 상기 입력 영상으로부터 보행자(pedestrian)를 추출하는 보행자추출부; 및 상기 운전가능 영역 및 상기 보행자 마스크(masks)를 기반으로 상기 보행자의 이상행동(abnormal behavior) 상황에 대한 증강 영상(augmented images)을 생성하는 데이터증강부를 포함하는 것을 특징으로 하는 보행자 영상 데이터 증강장치를 제공한다.

[0018] 본 발명의 다른 실시예에 따르면, 영상 증강방법이 포함하는 각 단계를 실행시키기 위하여 컴퓨터로 읽을 수 있는 기록매체에 저장된 컴퓨터프로그램을 제공한다.

발명의 효과

[0019] 이상에서 설명한 바와 같이 본 실시예에 따르면, 딥러닝(deep learning)을 기반으로 보행자(pedestrian) 이상행동(abnormal behavior) 상황에 대한 증강 영상을 생성하는 것이 가능한 보행자 영상 데이터 증강방법을 제공함으로써, 증강된 데이터셋을 이용하는 학습 과정에 기반하여, 교통상황 중 보행자의 이상행동에 대한 지능형 교통안전 시스템의 탐지 성능(detection performance) 향상이 가능해지는 효과가 있다.

[0020] 또한 본 실시예에 따르면, 딥러닝(deep learning)을 기반으로 다양한 장소에서 발생 가능한 보행자 이상행동 상황에 대한 증강 영상을 생성하는 것이 가능한 보행자 영상 데이터 증강방법을 제공함으로써, 발생 가능성이 매우 낮은 보행자 이상행동의 감지에 범용적으로 적용 가능한 감지 시스템의 구축이 가능해지는 효과가 있다.

도면의 간단한 설명

- [0021] 도 1은 본 발명의 일 실시예에 따른 보행자 영상 데이터 증강장치에 대한 구성도이다.
- 도 2는 본 발명의 일 실시예에 따른 영역 검출 모델 및 보행자 검출 모델의 기반이 되는 Mask R-CNN에 대한 구성도이다.
- 도 3은 본 발명의 일 실시예에 따른 보행자 영상 데이터 증강방법에 대한 순서도이다.
- 도 4는 본 발명의 일 실시예에 따른 영상 증강방법에 따른 증강 영상 생성 과정을 보여 주는 예시도이다.
- 도 5는 본 발명의 일 실시예에 따른 분류기 모델의 기반이 되는 GoogLeNet에 대한 구성도이다.

발명을 실시하기 위한 구체적인 내용

- [0022] 이하, 본 발명의 실시예들을 예시적인 도면을 참조하여 상세하게 설명한다. 각 도면의 구성요소들에 참조부호를 부가함에 있어서, 동일한 구성요소들에 대해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 실시예들을 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 실시예들의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.
- [0023] 또한, 본 실시예들의 구성요소를 설명하는 데 있어서, 제 1, 제 2, A, B, (a), (b) 등의 용어를 사용할 수 있다. 이러한 용어는 그 구성요소를 다른 구성요소와 구별하기 위한 것일 뿐, 그 용어에 의해 해당 구성요소의 본질이나 차례 또는 순서 등이 한정되지 않는다. 명세서 전체에서, 어떤 부분이 어떤 구성요소를 '포함', '구비'한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 '...부', '모듈' 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [0024] 첨부된 도면과 함께 이하에 개시될 상세한 설명은 본 발명의 예시적인 실시형태를 설명하고자 하는 것이며, 본 발명이 실시될 수 있는 유일한 실시형태를 나타내고자 하는 것이 아니다.
- [0025] 본 실시예는 딥러닝(deep learning) 기반 보행자 영상 데이터 증강방법에 내용을 개시한다. 보다 자세하게는, 딥러닝 기반 객체 검출 모델(object detection models)을 이용하여 입력 영상으로부터 운전가능 영역(drivable area) 및 보행자(pedestrian)를 검출한다. 검출된 운전가능 영역 및 사람에 대한 마스크(masks)를 기반으로 보행자 이상행동(abnormal behavior) 상황에 대한 증강 영상을 자동으로 생성하는 것이 가능한 보행자 영상 데이터 증강방법을 제공한다.
- [0026] 도 1은 본 발명의 일 실시예에 따른 보행자 영상 데이터 증강장치에 대한 구성도이다.
- [0027] 본 발명에 대한 실시예에 있어서, 보행자 영상 데이터 증강장치(100, 이하 '영상 증강장치')는 입력 영상으로부터 운전가능 영역 및 사람을 검출하고, 검출된 운전가능 영역 및 사람에 대한 마스크를 기반으로 보행자 이상행동 상황에 대한 증강 영상을 생성한다. 영상 증강장치(100)는 입력부(101), 영역추출부(102), 보행자추출부(103) 및 데이터증강부(104)의 전부 또는 일부를 포함한다. 여기서, 본 실시예에 따른 영상 증강장치(100)에 포함되는 구성요소가 반드시 이에 한정되는 것은 아니다. 예컨대, 영상 증강장치(100) 상에 검출 모델의 트레이닝을 위한 트레이닝부(미도시)를 추가로 구비하거나, 외부의 트레이닝부와 연동되는 형태로 구현될 수 있다.
- [0028] 본 실시예에 따른 입력부(101)는 입력 영상을 획득한다. 입력 영상은 정상 상황을 포함하며, 운전가능 영역 및/또는 복수의 보행자를 포함하는 것으로 가정한다.
- [0029] 본 실시예에 따른 영역추출부(102)는 영역 검출 모델(area detection model)을 기반으로 입력 영상으로부터 운전가능 영역을 추출한다. 영역 검출 모델은 딥러닝 기반의 객체 검출 모델로서, 운전가능 영역이 포함된 학습용 데이터를 이용하여 사전에 트레이닝될 수 있다. 영역추출부(102)는 영역 검출 모델로서, Mask R-CNN(Mask Region-based Convolutional Neural Network)을 이용한다(비특허문헌 5 참조).
- [0030] 본 실시예에 따른 보행자추출부(103)는 보행자 검출 모델(pedestrian detection model)을 기반으로 입력 영상으로부터 운전가능 영역을 추출한다. 보행자 검출 모델은 딥러닝 기반의 객체 검출 모델로서, 사람이 포함된 학습용 데이터를 이용하여 사전에 트레이닝될 수 있다. 보행자검출부(103)는 보행자 검출 모델로서 역시 Mask R-CNN을 이용한다.

- [0031] 도 2는 본 발명의 일 실시예에 따른 영역 검출 모델 및 보행자 검출 모델의 기반이 되는 Mask R-CNN에 대한 구성도이다. Mask R-CNN은 RoI(Region of Interest)를 기반으로 객체의 ID(identification) 및 바운딩 박스를 생성하는 객체 검출 경로와 병렬로 객체에 대한 마스크(mask)를 검출하는 경로를 포함한다.
- [0032] 본 실시예에 따른 데이터증강부(104)는 추출된 운전가능 영역 및 보행자에 대한 마스크를 기반으로 무단횡단(jaywalking)과 같은 보행자 이상행동 상황에 대한 증강 영상을 생성한다.
- [0033] 본 실시예에 따른 영상 증강장치(100)가 탑재되는 디바이스(미도시)는 프로그램가능 컴퓨터일 수 있으며, 서버(미도시)와 연결이 가능한 적어도 한 개의 통신 인터페이스를 포함한다.
- [0034] 전술한 바와 같은 영역 검출 모델 및 보행자 검출 모델에 대한 트레이닝은, 영상 증강장치(100)가 탑재되는 디바이스의 컴퓨팅 파워를 이용하여 디바이스에서 진행될 수 있다.
- [0035] 전술한 바와 같은 영역 검출 모델 및 보행자 검출 모델에 대한 트레이닝은 서버에서 진행될 수도 있다. 디바이스 상에 탑재된 영상 증강장치(100)의 구성요소인 영역 검출 모델 및 보행자 검출 모델과 동일한 구조의 딥러닝 모델에 대하여 서버의 트레이닝부는 트레이닝을 수행할 수 있다. 디바이스와 연결되는 통신 인터페이스를 이용하여 서버는 트레이닝된 딥러닝 모델의 파라미터를 디바이스로 전달하고, 전달받은 파라미터를 이용하여 영상 증강장치(100)는 영역 검출 모델 및 보행자 검출 모델의 파라미터를 업데이트할 수 있다. 또한 디바이스 출하시점 또는 영상 증강장치(100)가 디바이스에 탑재되는 시점에, 영역 검출 모델 및 보행자 검출 모델의 파라미터가 설정될 수 있다.
- [0036] 이하, 도 3 및 도 4를 참조하여, 영상 증강장치(100)을 이용하는 보행자 영상 데이터 증강방법(이하 '영상 증강방법')에 대하여 설명한다.
- [0037] 도 3은 본 발명의 일 실시예에 따른 보행자 영상 데이터 증강방법에 대한 순서도이다. 도 4는 본 발명의 일 실시예에 따른 영상 증강방법에 따른 증강 영상 생성 과정을 보여 주는 예시도이다.
- [0038] 영상 증강장치(100)의 트레이닝부는 영역 검출 모델 및 보행자 검출 모델에 대한 사전 트레이닝을 수행한다(S301).
- [0039] 영역 검출 모델 및 보행자 검출 모델 각각은 Mask R-CNN에 기반하는 딥러닝 기반 객체 검출 모델이다. 영역 검출 모델은 BDD100K(Black box Driving Dataset 100K)라는 차량용 블랙박스 데이터셋을 학습용 데이터로 이용하여 사전에 트레이닝된다. BDD100K는 대략 십만 개의 블랙박스 동영상 클립을 포함하며, 대부분의 영상이 도로를 포함하므로, 트레이닝부는 운전가능 영역의 검출에 특화되도록 영역 검출 모델을 사전에 트레이닝시킬 수 있다.
- [0040] 보행자 검출 모델은 MS-COCO 2014(Microsoft Common Object in Context 2014)라는 데이터셋을 학습용 데이터로 이용하여 사전에 트레이닝된다. MS-COCO 데이터셋은 객체 검출 및 분할(segmentation) 등에 이용하기 위한 영상을 포함하며, 데이터셋에 포함된 영상 중 사람이 포함된 영상을 이용하여, 트레이닝부는 보행자의 검출에 특화되도록 보행자 검출 모델을 사전에 트레이닝시킬 수 있다.
- [0041] 영상 증강장치(100)는 입력 영상을 획득한다(S302). 입력 영상으로는 CUHK에서 수집된 어베뉴 CCTV 영상(비특허 문헌 3 참조) 및 Y시의 거리에서 수집된 CCTV 영상이 이용되었다. 입력 영상은 정상 상황을 포함하며, 도 4에 도시된 바와 같이 운전가능 영역 및 적어도 한 명의 보행자를 포함한다.
- [0042] 영상 증강장치(100)는 영역 검출 모델을 이용하여 입력 영상으로부터 운전가능 영역을 검출한다(S303). 영역 검출 모델은 도 4에 도시된 바와 같이 입력 영상으로부터 운전가능 영역에 대한 마스크를 검출할 수 있다.
- [0043] 영상 증강장치(100)는 보행자 검출 모델을 이용하여 입력 영상으로부터 보행자를 검출한다(S304). 보행자 검출 모델은 도 4에 도시된 바와 같이 입력 영상으로부터 보행자에 대한 마스크를 검출할 수 있다.
- [0044] 입력 영상이 복수의 보행자를 포함하는 경우, 다음과 같은 규칙을 따른다. 보행자 검출 모델은 포함된 보행자들 모두 검출하되, 검출 순서에 따라 검출된 보행자를 O_k 로 표현한 후, 수학식 1에 표현된 규칙을 기반으로 한 명의 보행자를 선택한다.

수학식 1

If $O_k (|k| > 1)$ exist, choose $\min k (O_k \notin D)$

[0045]

[0046]

여기서 D는 입력 영상에 포함된 운전가능 영역이다. 수학식 1에 표현된 규칙에 따르면, 운전가능 영역에 포함되지 않는, 검출된 보행자에 대하여 검출 순서가 가장 빠른 보행자가 선택된다.

[0047]

영상 증강장치(100)는 운전가능 영역 및 보행자의 마스크를 기반으로 보행자 이상행동 상황에 대한 증강 영상을 생성한다(S305). 영상 증강장치(100)는 도 4에 도시된 바와 같이 보행자의 마스크를 운전가능 영역 마스크 내에 위치하도록 함으로써 무단횡단(jaywalking) 상황에 대한 합성 영상을 생성한다. 또한 복수의 보행자 마스크 및 가로 플립(horizontal flip)이 적용된 보행자 마스크와 복수의 운전가능 영역 마스크를 상호 결합함으로써 보행자 이상행동 상황에 대한 증강 영상을 생성할 수 있다.

[0048]

무단횡단 상황에 대한 증강 영상은 비정상 상황으로 분류되고, 이상행동을 실행하는 보행자의 위치 정보에 대한 레이블도 포함한다. 증강 영상 및 증강 영상에 포함된 레이블은 차후 지능형 교통안전 시스템에 대한 트레이닝 과정에서 사용될 수 있다.

[0049]

이하, 본 실시예에 따른 영상 증강방법에 기반하여 생성된 증강 영상을 포함하는 데이터셋을 지능형 교통안전 시스템의 트레이닝에 적용한 실험예 및 실험 결과를 설명한다.

[0050]

실험예에서는 교통안전 시스템의 예로서 입력 영상으로부터 무단횡단 상황과 정상 상황을 구분하는 분류기 모델(classifier model)을 이용하였다. 분류기 모델로는 딥러닝 기반의 GoogLeNet이 이용되었다(비특허문헌 6 참조). 도 5에 도시된 바와 같이 GoogLeNet은 깊은 층(deep layered) 구조 기반의 CNN으로서, 분류 성능을 향상시키기 위하여 신경망의 깊이, 즉 레이어의 수를 증가시키되, 각 레이어를 단순화하여 연산량의 증가를 억제시킨 CNN이다.

[0051]

분류기 모델은 20,000 개 이상의 이미지 카테고리(categories)를 포함하는 ImageNet 데이터셋을 기반으로 사전 트레이닝된 후, 무단횡단 여부를 구분하기 위한 레이어가 추가된 채로 리트레이닝(retraining)된다.

[0052]

분류기 모델을 리트레이닝하기 위한 증강 영상 데이터셋, 즉 무단횡단 상황을 포함한 영상은 전술한 바와 같은 영상 증강방법에 기반하여 생성될 수 있다. 실험예에서는 전체 500 장의 증강 무단횡단 영상을 생성하였다. 증강 무단횡단 영상 및 동수의 정상 상황 영상을 이용하여 분류기 모델을 리트레이닝한 후, 검증(validation)용 데이터셋을 이용하여 분류 정확도(classification accuracy, 이하 '정확도')를 측정하였다.

[0053]

검증용 데이터셋은 337 장의 자연(natural) 무단횡단 영상 및 400 장의 정상 상황 영상을 포함한다. 검증용 데이터셋은 구글 또는 Y시의 CCTV 영상으로부터 수집되었다.

[0054]

비교 대상으로는 자연 무단횡단 영상 및 정상 상황 영상을 포함하는 자연 데이터셋을 이용하여 리트레이닝된 분류기 모델을 이용하였다.

[0055]

표 1은 본 실시예가 적용된 분류기 모델 및 자연 데이터셋을 이용하는 분류기 모델에 대한 측정된 정확도를 나타내고 있다.

표 1

	TP	FN	TN	FP	정확도
비교 대상	328	9	319	81	0.878
본 실시예	327	10	386	14	0.967
비교 대상: 자연 데이터셋을 이용하여 분류기 모델을 리트레이닝					
본 실시예: 증강 및 자연 데이터셋을 이용하여 분류기 모델을 리트레이닝					

[0056]

[0057]

표 1에서 정확도는 수학식 2에 의하여 표현된다.

수학식 2

$$\text{정확도} = \frac{\text{TP} + \text{TN}}{\text{positive} + \text{negative}}$$

[0058]

[0059]

여기서, 긍정(positive)은 무단횡단 상황을 부정(negative)은 정상 상황을 의미한다. 또한, TP 및 TN 각각은 무단횡단 및 정상 상황을 제대로 분류한 것이고, FN는 무단횡단을 정상 상황으로, FP은 정상 상황을 무단 횡단으로 분류한 것이다.

[0060]

표 1에 나타낸 바와 같이 본 실시예에 따른 영상 증강방법을 이용하는 분류기 모델의 정확도가 비교 대상에 대하여 8 % 이상 향상되었다. 특히 정상 상황을 무단횡단 상황으로 분류하는 FP 경우가 비교 대상에 비하여 약 1/5 이하로 감소됨으로써 무단횡단 분류에 대한 분류기 모델의 전체적인 성능이 향상되었다.

[0061]

이상에서 설명한 바와 같이 본 실시예에 따르면, 딥러닝(deep learning)을 기반으로 보행자(pedestrian) 이상행동(abnormal behavior) 상황에 대한 증강 영상을 생성하는 것이 가능한 보행자 영상 데이터 증강방법을 제공함으로써, 증강된 데이터셋을 이용하는 학습 과정에 기반하여, 교통상황 중 보행자의 이상행동에 대한 지능형 교통안전 시스템의 탐지 성능(detection performance) 향상이 가능해지는 효과가 있다.

[0062]

또한 본 실시예에 따르면, 딥러닝(deep learning)을 기반으로 다양한 장소에서 발생 가능한 보행자 이상행동 상황에 대한 증강 영상을 생성하는 것이 가능한 보행자 영상 데이터 증강방법을 제공함으로써, 발생 가능성이 매우 낮은 보행자 이상행동의 감지에 범용적으로 적용 가능한 감지 시스템의 구축이 가능해지는 효과가 있다.

[0063]

본 실시예에 따른 각 순서도에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나, 반드시 이에 한정되는 것은 아니다. 다시 말해, 순서도에 기재된 과정을 변경하여 실행하거나 하나 이상의 과정을 병렬적으로 실행하는 것이 적용 가능할 것이므로, 순서도는 시계열적인 순서로 한정되는 것은 아니다.

[0064]

본 명세서에 설명되는 시스템들 및 기법들의 다양한 구현예들은, 디지털 전자 회로, 집적 회로, FPGA(field programmable gate array), ASIC(application specific integrated circuit), 컴퓨터 하드웨어, 펌웨어, 소프트웨어, 및/또는 이들의 조합으로 실현될 수 있다. 이러한 다양한 구현예들은 프로그래밍가능 시스템 상에서 실행가능한 하나 이상의 컴퓨터 프로그램들로 구현되는 것을 포함할 수 있다. 프로그래밍가능 시스템은, 저장 시스템, 적어도 하나의 입력 디바이스, 그리고 적어도 하나의 출력 디바이스로부터 데이터 및 명령들을 수신하고 이들에게 데이터 및 명령들을 전송하도록 결합되는 적어도 하나의 프로그래밍가능 프로세서(이것은 특수 목적 프로세서일 수 있거나 혹은 범용 프로세서일 수 있음)를 포함한다. 컴퓨터 프로그램들(이것은 또한 프로그램들, 소프트웨어, 소프트웨어 애플리케이션들 혹은 코드로서 알려져 있음)은 프로그래밍가능 프로세서에 대한 명령어들을 포함하며 "컴퓨터가 읽을 수 있는 기록매체"에 저장된다.

[0065]

컴퓨터가 읽을 수 있는 기록매체는, 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 이러한 컴퓨터가 읽을 수 있는 기록매체는 ROM, CD-ROM, 자기 테이프, 플로피디스크, 메모리 카드, 하드 디스크, 광자기 디스크, 스토리지 디바이스 등의 비휘발성(non-volatile) 또는 비일시적인(non-transitory) 매체일 수 있으며, 또한 캐리어 웨이브(예를 들어, 인터넷을 통한 전송) 및 데이터 전송 매체(data transmission medium)와 같은 일시적인(transitory) 매체를 더 포함할 수도 있다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다.

[0066]

본 명세서에 설명되는 시스템들 및 기법들의 다양한 구현예들은, 프로그램가능 컴퓨터에 의하여 구현될 수 있다. 여기서, 컴퓨터는 프로그램가능 프로세서, 데이터 저장 시스템(휘발성 메모리, 비휘발성 메모리, 또는 다른 종류의 저장 시스템이거나 이들의 조합을 포함함) 및 적어도 한 개의 커뮤니케이션 인터페이스를 포함한다. 예컨대, 프로그램가능 컴퓨터는 서버, 네트워크 기기, 셋탑 박스, 내장형 장치, 컴퓨터 확장 모듈, 개인용 컴퓨터, 랩탑, PDA(Personal Data Assistant), 클라우드 컴퓨팅 시스템 또는 모바일 장치 중 하나일 수 있다.

[0067]

이상의 설명은 본 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 실시예들은 본 실시예의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호

범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

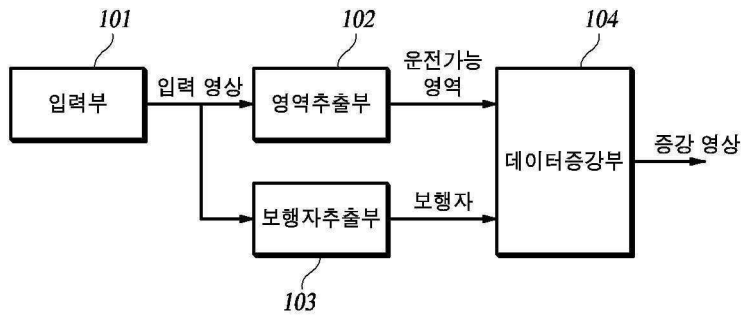
부호의 설명

100: 영상 증강장치 101: 입력부
102: 영역추출부 103: 보행자추출부
104: 데이터증강부

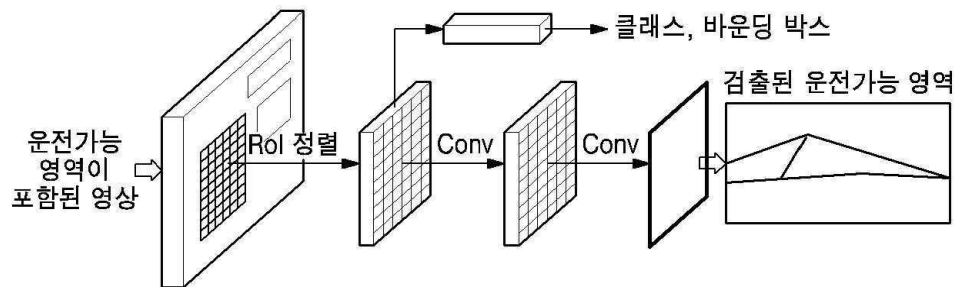
도면

도면1

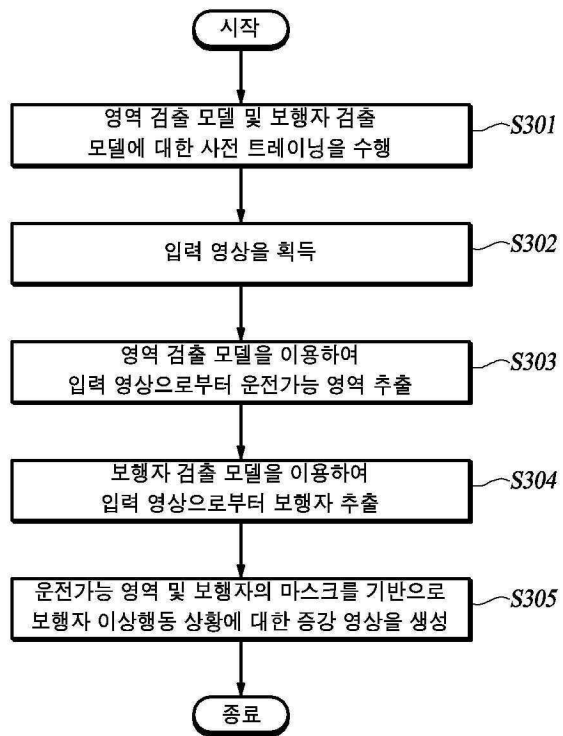
100



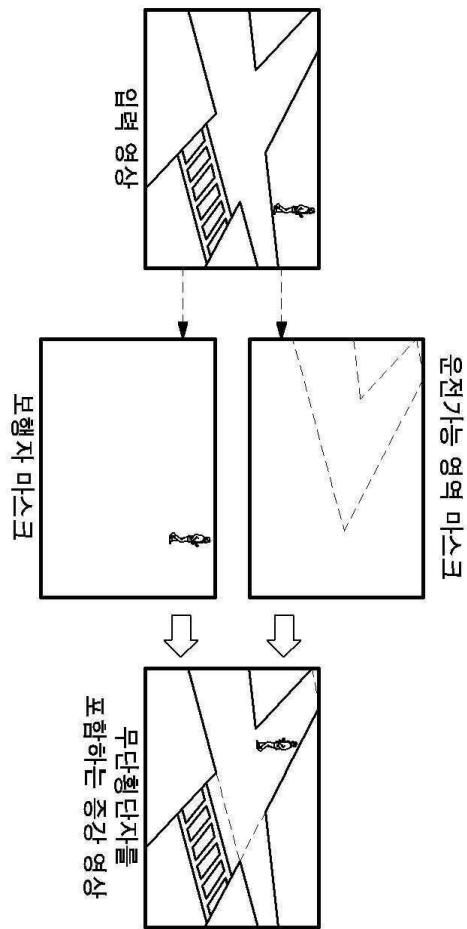
도면2



도면3



도면4



도면5

