



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0076862
(43) 공개일자 2024년05월31일

- (51) 국제특허분류(Int. Cl.)
G06V 20/52 (2022.01) G06N 3/08 (2023.01)
G06Q 50/10 (2012.01) G06T 7/70 (2017.01)
G06V 10/774 (2022.01) G06V 40/10 (2022.01)
- (52) CPC특허분류
G06V 20/52 (2023.08)
G06N 3/08 (2023.01)
- (21) 출원번호 10-2022-0157946
- (22) 출원일자 2022년11월23일
심사청구일자 2022년11월23일

- (71) 출원인
포항공과대학교 산학협력단
경상북도 포항시 남구 청암로 77 (지곡동)
주식회사 브라이언
경상북도 경산시 삼풍로 27, 경북테크노파크 글로벌벤처동 2513호 (삼풍동)
- (72) 발명자
라이언
경상북도 포항시 남구 청암로 77 포항공과대학교 서영주
경상북도 포항시 남구 청암로 77 포항공과대학교 (뒷면에 계속)
- (74) 대리인
관현기, 임동우

전체 청구항 수 : 총 10 항

(54) 발명의 명칭 딥러닝을 이용한 위급상황 모니터링 시스템 및 그 방법

(57) 요약

본 발명은 딥러닝을 이용한 위급상황 모니터링 시스템 및 그 방법에 관한 것으로, 설치된 위치에서 촬영된 감시 영상을 제공하는 CCTV 영상카메라; 및 상기 CCTV 영상카메라와 엣지컴퓨팅환경으로 구축되며, 기준데이터셋을 이용하여 딥러닝모델을 학습시킨 후에, 상기 감시영상이 제공될 경우 사람객체를 검출하고, 상기 검출된 사람객체에 개별아이디를 부여하면서 골격좌표를 추출하며, 상기 추출된 골격좌표를 상기 딥러닝모델에 입력하여 사람자세를 구분한 후에, 구분결과를 제공하는 관제센터;를 포함함으로써, 이미지분류의 단점을 보완하여 유사행동에 대한 분류정확도를 향상시킬 수 있고, 다수의 사람들에 대한 행동을 인식 및 구분하여 정교한 이상행동 탐지를 수행할 수 있다.

대표도 - 도1



(52) CPC특허분류

G06Q 50/10 (2015.01)
G06T 7/70 (2017.01)
G06V 10/774 (2023.08)
G06V 40/103 (2022.01)
G06T 2207/20044 (2013.01)
G06T 2207/30196 (2013.01)
G06T 2210/12 (2013.01)

(72) 발명자

황도경

경상북도 포항시 남구 청암로 77 포항공과대학교

김동주

경상북도 포항시 남구 청암로 77 포항공과대학교

임익기

경상북도 경산시 삼성현로 347 101동 2505호

이 발명을 지원한 국가연구개발사업

과제고유번호	1345354019
과제번호	2022R1A6A1A03052954
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	이공학학술연구기반구축
연구과제명	인공지능연구원
기여율	1/1
과제수행기관명	포항공과대학교
연구기간	2022.06.01 ~ 2023.02.28

명세서

청구범위

청구항 1

설치된 위치에서 촬영된 감시영상을 제공하는 CCTV 영상카메라; 및

상기 CCTV 영상카메라와 엣지컴퓨팅환경으로 구축되며, 기준데이터셋을 이용하여 딥러닝모델을 학습시킨 후에, 상기 감시영상이 제공될 경우 사람객체를 검출하고, 상기 검출된 사람객체에 개별아이드를 부여하면서 골격좌표를 추출하며, 상기 추출된 골격좌표를 상기 딥러닝모델에 입력하여 사람자세를 구분한 후에, 구분결과를 제공하는 관제센터;

를 포함하는 딥러닝을 이용한 위급상황 모니터링 시스템.

청구항 2

청구항 1에 있어서,

상기 딥러닝모델은,

Bi-LSTM모델(bidirectional long short-term memory)을 포함하는

딥러닝을 이용한 위급상황 모니터링 시스템.

청구항 3

청구항 2에 있어서,

상기 관제센터는,

Openpose모델을 이용하여 상기 사람객체를 검출하여 경계상자(bounding box)로 표시하는

딥러닝을 이용한 위급상황 모니터링 시스템.

청구항 4

청구항 3에 있어서,

상기 관제센터는,

Strong sort모델을 이용한 상기 사람객체의 위치 비교를 통해 상기 개별아이드를 부여한 후에, Openpose모델을 이용하여 상기 골격좌표를 추출하여 단위벡터로 변환하고, 시계열데이터로 변환하는

딥러닝을 이용한 위급상황 모니터링 시스템.

청구항 5

청구항 4에 있어서,

상기 관제센터는,

상기 개별아이드에 따라 상기 시계열데이터를 취합하고, 학습된 상기 딥러닝모델에 입력하여 상기 사람자세를 구분하는

딥러닝을 이용한 위급상황 모니터링 시스템.

청구항 6

CCTV 영상카메라와 관제센터를 엣지컴퓨팅환경으로 구축하는 단계;

상기 관제센터에서 기준데이터셋을 이용하여 딥러닝모델을 학습시키는 단계;

상기 관제센터에서 상기 CCTV 영상카메라를 통해 촬영된 감시영상을 수신하는 단계;

상기 관제센터에서 상기 감시영상이 수신될 경우 사람객체를 검출하는 단계;

상기 관제센터에서 상기 검출된 사람객체에 개별아이디를 부여하면서 골격좌표를 추출하는 단계; 및

상기 관제센터에서 상기 추출된 골격좌표를 상기 딥러닝모델에 입력하여 사람자세를 구분한 후에, 구분결과를 제공하는 단계;

를 포함하는 딥러닝을 이용한 위급상황 모니터링 방법.

청구항 7

청구항 6에 있어서,

상기 딥러닝모델을 학습시키는 단계는,

Bi-LSTM모델(bidirectional long short-term memory)을 포함하는 상기 딥러닝모델을 학습시키는

딥러닝을 이용한 위급상황 모니터링 방법.

청구항 8

청구항 7에 있어서,

상기 사람객체를 검출하는 단계는,

Openpose모델을 이용하여 상기 사람객체를 검출하여 경계상자(bounding box)로 표시하는

딥러닝을 이용한 위급상황 모니터링 방법.

청구항 9

청구항 8에 있어서,

상기 골격좌표를 추출하는 단계는,

Strong sort모델을 이용한 상기 사람객체의 위치 비교를 통해 상기 개별아이디를 부여한 후에, Openpose모델을 이용하여 상기 골격좌표를 추출하여 단위벡터로 변환하고, 시계열데이터로 변환하는

딥러닝을 이용한 위급상황 모니터링 방법.

청구항 10

청구항 9에 있어서,

상기 구분결과를 제공하는 단계는,

상기 개별아이디에 따라 상기 시계열데이터를 취합하고, 학습된 상기 딥러닝모델에 입력하여 상기 사람자세를

구분하는

딥러닝을 이용한 위급상황 모니터링 방법.

발명의 설명

기술분야

[0001] 본 발명은 CCTV 영상카메라와 관제센터가 엣지컴퓨팅환경으로 구축되며, 관제센터에서 기준데이터셋을 이용하여 딥러닝모델을 학습시킨 후에, CCTV 영상카메라로부터 감시영상이 제공될 경우 사람객체를 검출하고, 검출된 사람객체에 개별아이디를 부여하면서 골격좌표를 추출하며, 추출된 골격좌표를 딥러닝모델에 입력하여 사람자세를 구분한 후에, 구분결과를 제공함으로써, 이미지분류의 단점을 보완하여 유사행동에 대한 분류정확도를 향상시킬 수 있고, 다수의 사람들에 대한 행동을 인식 및 구분하여 정교한 이상행동 탐지를 수행할 수 있는 딥러닝을 이용한 위급상황 모니터링 시스템 및 그 방법에 관한 것이다.

배경기술

[0002] 잘 알려진 바와 같이, 컴퓨터비전 기술이 급성장함에 따라 이를 활용하는 어린이, 노인, 장애인 등이 생활하는 생활시설의 모니터링 분야에서 영상을 통한 관제에 대한 수요가 급속하게 증가하고 있으며, 이와 관련하여 컴퓨터비전 기술을 활용하여 영상 속 사람의 행동에 대해 인식 및 구분하는 연구가 활발히 진행되고 있다.

[0003] 예를 들면, 생활시설 현장에서 위험 상황을 탐지하기 위해 카메라를 활용한 안전장비 착용 여부, CCTV를 통한 위험 지역에서의 돌발상황 알람 등이 존재하는데, CCTV를 활용하여 관제하는 경우, CCTV가 설치되어 촬영하는 영역 외의 지역에서의 위험 상황들을 관제실 등에서 인식하기에는 어려움이 존재한다.

[0004] 그리고, 생활시설 내 사각지대와 같이 CCTV가 관찰하기 힘든 지역에서는 더욱이 관제하기 어려운 실정이다.

[0005] 이러한 문제점을 해결하기 위해 종래에는 하나의 객체(한 사람)에 대해 행동 인식 및 구분하는 연구들이 주로 진행되고 있으며, 다수의 객체에 대해 인식 및 구분하기에는 적합하지 않은 문제점이 있다.

[0006] 또한, 종래의 연구들은 이미지분류(image classification)를 통한 행동 구분을 하는 방식을 취하고 있기 때문에, 유사한 행동들(누워있는 행위와 쓰러지는 행위, 수영과 물에 빠지는 행위)을 분류하는 데 어려움이 있으며, 이는 특정 이미지데이터에 대하여 과적합된 것이 대부분이다.

[0007] 이에 따라, 이미지분류의 단점을 보완하고, 다수의 사람들에 대한 행동을 인식 및 구분하여 정교한 이상행동 탐지를 수행할 수 있는 방안이 시급히 요구되고 있다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 1. 한국등록특허 제10-2374201호(2022.03.08.등록)

발명의 내용

해결하려는 과제

[0009] 본 발명은 CCTV 영상카메라와 관제센터가 엣지컴퓨팅환경으로 구축되며, 관제센터에서 기준데이터셋을 이용하여 딥러닝모델을 학습시킨 후에, CCTV 영상카메라로부터 감시영상이 제공될 경우 사람객체를 검출하고, 검출된 사람객체에 개별아이디를 부여하면서 골격좌표를 추출하며, 추출된 골격좌표를 딥러닝모델에 입력하여 사람자세를 구분한 후에, 구분결과를 제공함으로써, 이미지분류의 단점을 보완하여 유사행동에 대한 분류정확도를 향상시킬 수 있고, 다수의 사람들에 대한 행동을 인식 및 구분하여 정교한 이상행동 탐지를 수행할 수 있는 딥러닝을 이용한 위급상황 모니터링 시스템 및 그 방법을 제공하고자 한다.

[0010] 본 발명의 실시예들의 목적은 이상에서 언급한 목적으로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0011] 본 발명의 일 측면에 따르면, 설치된 위치에서 촬영된 감시영상을 제공하는 CCTV 영상카메라; 및 상기 CCTV 영상카메라와 엣지컴퓨팅환경으로 구축되며, 기준데이터셋을 이용하여 딥러닝모델을 학습시킨 후에, 상기 감시영상이 제공될 경우 사람객체를 검출하고, 상기 검출된 사람객체에 개별아이디를 부여하면서 골격좌표를 추출하며, 상기 추출된 골격좌표를 상기 딥러닝모델에 입력하여 사람자세를 구분한 후에, 구분결과를 제공하는 관제센터;를 포함하는 딥러닝을 이용한 위급상황 모니터링 시스템이 제공될 수 있다.
- [0012] 또한, 본 발명의 일 측면에 따르면, 상기 딥러닝모델은, Bi-LSTM모델(bidirectional long short-term memory)을 포함하는 딥러닝을 이용한 위급상황 모니터링 시스템이 제공될 수 있다.
- [0013] 또한, 본 발명의 일 측면에 따르면, 상기 관제센터는, Openpose모델을 이용하여 상기 사람객체를 검출하여 경계상자(bounding box)로 표시하는 딥러닝을 이용한 위급상황 모니터링 시스템이 제공될 수 있다.
- [0014] 또한, 본 발명의 일 측면에 따르면, 상기 관제센터는, Strong sort모델을 이용한 상기 사람객체의 위치 비교를 통해 상기 개별아이디를 부여한 후에, Openpose모델을 이용하여 상기 골격좌표를 추출하여 단위벡터로 변환하고, 시계열데이터로 변환하는 딥러닝을 이용한 위급상황 모니터링 시스템이 제공될 수 있다.
- [0015] 또한, 본 발명의 일 측면에 따르면, 상기 관제센터는, 상기 개별아이디에 따라 상기 시계열데이터를 취합하고, 학습된 상기 딥러닝모델에 입력하여 상기 사람자세를 구분하는 딥러닝을 이용한 위급상황 모니터링 시스템이 제공될 수 있다.
- [0016] 본 발명의 다른 측면에 따르면, CCTV 영상카메라와 관제센터를 엣지컴퓨팅환경으로 구축하는 단계; 상기 관제센터에서 기준데이터셋을 이용하여 딥러닝모델을 학습시키는 단계; 상기 관제센터에서 상기 CCTV 영상카메라를 통해 촬영된 감시영상을 수신하는 단계; 상기 관제센터에서 상기 감시영상이 수신될 경우 사람객체를 검출하는 단계; 상기 관제센터에서 상기 검출된 사람객체에 개별아이디를 부여하면서 골격좌표를 추출하는 단계; 및 상기 관제센터에서 상기 추출된 골격좌표를 상기 딥러닝모델에 입력하여 사람자세를 구분한 후에, 구분결과를 제공하는 단계;를 포함하는 딥러닝을 이용한 위급상황 모니터링 방법이 제공될 수 있다.
- [0017] 또한, 본 발명의 다른 측면에 따르면, 상기 딥러닝모델을 학습시키는 단계는, Bi-LSTM모델(bidirectional long short-term memory)을 포함하는 상기 딥러닝모델을 학습시키는 딥러닝을 이용한 위급상황 모니터링 방법이 제공될 수 있다.
- [0018] 또한, 본 발명의 다른 측면에 따르면, 상기 사람객체를 검출하는 단계는, Openpose모델을 이용하여 상기 사람객체를 검출하여 경계상자(bounding box)로 표시하는 딥러닝을 이용한 위급상황 모니터링 방법이 제공될 수 있다.
- [0019] 또한, 본 발명의 다른 측면에 따르면, 상기 골격좌표를 추출하는 단계는, Strong sort모델을 이용한 상기 사람객체의 위치 비교를 통해 상기 개별아이디를 부여한 후에, Openpose모델을 이용하여 상기 골격좌표를 추출하여 단위벡터로 변환하고, 시계열데이터로 변환하는 딥러닝을 이용한 위급상황 모니터링 방법이 제공될 수 있다.
- [0020] 또한, 본 발명의 다른 측면에 따르면, 상기 구분결과를 제공하는 단계는, 상기 개별아이디에 따라 상기 시계열데이터를 취합하고, 학습된 상기 딥러닝모델에 입력하여 상기 사람자세를 구분하는 딥러닝을 이용한 위급상황 모니터링 방법이 제공될 수 있다.

발명의 효과

- [0021] 본 발명은 CCTV 영상카메라와 관제센터가 엣지컴퓨팅환경으로 구축되며, 관제센터에서 기준데이터셋을 이용하여 딥러닝모델을 학습시킨 후에, CCTV 영상카메라로부터 감시영상이 제공될 경우 사람객체를 검출하고, 검출된 사람객체에 개별아이디를 부여하면서 골격좌표를 추출하며, 추출된 골격좌표를 딥러닝모델에 입력하여 사람자세를 구분한 후에, 구분결과를 제공함으로써, 이미지분류의 단점을 보완하여 유사행동에 대한 분류정확도를 향상시킬 수 있고, 다수의 사람들에 대한 행동을 인식 및 구분하여 정교한 이상행동 탐지를 수행할 수 있다.
- [0022] 또한, 본 발명은 엣지컴퓨팅 환경에서 CCTV 영상을 실시간으로 분석함으로써, 사회적 약자의 이상행동을 탐지하여 안전사고를 예방할 수 있고, 치료 및 진단의 자료를 축적할 수 있다.
- [0023] 또한, 기존의 행동구분 기술은 이미지 분류 및 하나의 객체(한 명의 사람)에 대한연구로 다양한 환경에서 다수의 객체에 적용하기에는 어려움이 존재하는 반면에, 본 발명에서는 CCTV 영상을 통해 상황의 맥락을 파악할 수 있고, 영상에서 다수의 사람들이 하는 행동에 대해 인식 및 구분할 수 있다. 또한 본 발명의 목적인 어린이, 노

인, 장애인 등이 이용하는 편의시설에서 사회적 약자들의 이상행동을 실시간으로 효과적으로 탐지할 수 있을 뿐만 아니라, 2차 사고 예방의 목적을 넘어, 사회적 약자의 생활 시설에만 국한되는 것이 아니라 유동 인구가 많은 시설에서 CCTV를 활용하여 광범위의 지역에 대해 안전 관리 및 모니터링 플랫폼을 구축함에 있어 유용하게 적용될 수 있다.

도면의 간단한 설명

- [0024] 도 1은 본 발명의 일 실시예에 따른 딥러닝을 이용한 위급상황 모니터링 시스템의 블록구성도이고,
- 도 2 내지 도 4는 본 발명의 일 실시예에 따른 딥러닝을 이용한 위급상황 모니터링 시스템의 딥러닝모델을 설명하기 위한 도면이며,
- 도 5는 본 발명의 다른 실시예에 따른 딥러닝을 이용한 위급상황 모니터링 방법을 나타낸 플로우차트이고,
- 도 6은 본 발명의 다른 실시예에 따른 딥러닝을 이용한 위급상황 모니터링 과정에서 사람의 행동을 구분하는 것을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0025] 본 발명의 실시예들에 대한 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성요소를 지칭한다.
- [0026] 본 발명의 실시예들을 설명함에 있어서 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략할 것이다. 그리고 후술되는 용어들은 본 발명의 실시예에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.
- [0027] 이하, 첨부된 도면을 참조하여 본 발명의 실시예를 상세히 설명하기로 한다.
- [0028] 도 1은 본 발명의 일 실시예에 따른 딥러닝을 이용한 위급상황 모니터링 시스템의 블록구성도이고, 도 2 내지 도 4는 본 발명의 일 실시예에 따른 딥러닝을 이용한 위급상황 모니터링 시스템의 딥러닝모델을 설명하기 위한 도면이다.
- [0029] 도 1 내지 도 4를 참조하면, 본 발명의 일 실시예에 따른 딥러닝을 이용한 위급상황 모니터링 시스템은 CCTV 영상카메라(100), 통신망(200), 관제센터(300) 등을 포함할 수 있다.
- [0030] CCTV 영상카메라(100)는 설치된 위치에서 촬영된 감시영상을 제공하는 감시카메라로서, CCTV(closed-circuit television)는 영상카메라를 통해 촬영된 감시영상을 DVR(digital video recorder)에 저장(녹화)하는 장치를 의미하며, 방법, 감시, 화재예방 등 다양한 안전사고를 예방하기 위해 필요한 위치에 설치될 수 있다.
- [0031] 통신망(200)은 유선통신망과 무선통신방식을 지원하는 네트워크망으로, 예를 들어 TCP/IP 등의 인터넷프로토콜을 이용한 유선통신망을 포함할 수 있고, 예를 들어 기지국, 이동교환국, 홈위치등록기 등을 포함하는 이동통신망 뿐만 아니라 RF, 지그비(zigbee), 블루투스(bluetooth), 와이파이(wi-fi), 와이브로(wibro), NFC(near field communication), BLE(bluetooth low energy) 등의 무선통신망을 포함할 수 있으며, 이를 통해 CCTV 영상카메라(100)와 관제센터(200) 간의 원활한 무선통신환경을 제공할 수 있다.
- [0032] 관제센터(300)는 CCTV 영상카메라(100)와 통신망(200)을 통해 엡지컴퓨팅환경으로 구축되며, 기준데이터셋을 이용하여 딥러닝모델을 학습시킨 후에, 감시영상이 제공될 경우 사람객체를 검출하고, 검출된 사람객체에 개별 아이디를 부여하면서 골격좌표를 추출하며, 추출된 골격좌표를 딥러닝모델에 입력하여 사람자세를 구분한 후에, 구분결과를 제공할 수 있다.
- [0033] 여기에서, 엡지컴퓨팅은 사용자 또는 데이터 소스의 물리적인 위치나 그 근처에서 컴퓨팅을 수행하는 통신환경을 의미하는데, CCTV 영상카메라(100)와 가까운 위치에 위치하는 관제센터(300)를 통해 컴퓨팅 서비스를 빠르고 안정적으로 처리할 수 있는 장점이 있으며, 여러 위치에서 공통의 리소스를 사용하여 데이터연산 및 처리를

분산시킬 수 있는 장점이 있다.

- [0034] 이러한 엣지컴퓨팅은 중앙집중식 서버나 클라우드가 아닌 데이터가 수집되고 분석되는 물리적위치 근처에서 컴퓨팅이 이루어지는 분산 컴퓨팅모델을 의미하며, 이를 사용하는 이유는 관찰하고자 하는 대상 혹은 장소가 서버 혹은 컴퓨팅장비의 설치가 원활하지 않은 사회적 약자들의 공간에서 쉽게 운용할 수 있게 하기 위함이다.
- [0035] 그리고, 딥러닝모델은 예를 들면, Bi-LSTM모델(bidirectional long short-term memory)을 포함할 수 있는데, 사람이 쓰러지는 영상데이터셋과 수어 구분을 위한 영상데이터셋을 포함하는 기존 영상데이터셋을 이용하여 학습시킬 수 있다.
- [0036] 예를 들면, 사람이 쓰러지는 영상데이터셋은 예를 들면, UR Fall Detection 등에서 제공하는 영상이미지셋을 이용할 수 있고, 수어 구분을 위한 영상데이터셋은 예를 들면, 데이콘의 수어 구분 경진대회 등에서 제공하는 영상이미지셋을 이용할 수 있는데, UR Fall Detection은 폴란드의 제슈프 대학에서 사람의 쓰러짐을 인식하기 위해 제작한 데이터셋으로, 6-10초 길이의 사람이 넘어지는 영상들로 구성되어 있으며, 데이콘의 수어 구분은 인공지능 플랫폼인 DACON에서 수신호를 구분하기 위해 제공한 오픈 데이터셋으로, 영상의 프레임마다 캡처본과 사람의 골격좌표가 저장되어 있다.
- [0037] 상술한 바와 같은 Bi-LSTM모델의 학습에 대해 설명하면, Bi-LSTM(양방향 LSTM)모델은 정방향으로만 학습을 진행하는 대신, 마지막 노드에서 뒤에서 앞으로(즉, 역방향) 실행되는 다른 LSTM을 추가하는 것으로, LSTM모델 대비 역방향으로 정보를 전달하는 히든레이어(hidden layer)를 추가함으로써, 정보를 보다 유연하게 처리할 수 있다. 즉, 각 시점에서 히든스테이트(hidden state)가 이전 시점과 미래 시점의 정보를 모두 갖는 효과가 있기 때문에, 모델을 전체 시계열 데이터로부터 학습할 수 있는 장점이 있다.
- [0038] 여기에서, 히든스테이트는 RNN(recurrent neural network, 순환신경망-시계열의 기본예측모델)의 은닉층(hidden layer)을 의미하며, 이전 상태에서부터 전달된 값을 나타낸다.
- [0039] 이러한 Bi-LSTM모델에 대해 부연하여 설명하면, LSTM은 시계열성 데이터를 분석하는 모델로, 순차적인 데이터를 학습하여 맥락을 파악할 수 있고, Bi-LSTM은 순차적인 데이터를 역순으로도 학습하는 추가과정을 통해 맥락을 더욱 잘 파악할 수 있도록 수정된 LSTM계열의 모델이다.
- [0040] 예를 들면, 도 2는 Bi-LSTM모델의 모식도로서, 시계열성 골격 데이터를 입력받을 경우 행동을 구분하는 과정을 나타내고 있는데, Bi-LSTM은 시계열 데이터를 연산하는 부분을 나타내고, 완전연결계층(FC layer)은 LSTM의 최종 출력물을 로짓(즉, 확률)으로 변환하는 부분을 나타낸다.
- [0041] 이러한 Bi-LSTM모델에 시계열 데이터가 입력될 경우 LSTM에 순차적으로 한번 입력하고, 역순으로 한번 더 입력하게 되며, 해당 과정 중 각 셀에서는 입력받은 데이터들 가운데 어떤 데이터를 버릴지, 저장할지를 결정한 후에 다음 셀로 데이터를 전달할 수 있고, 모든 셀에서의 연산이 끝나고 Bi-LSTM의 최종적인 출력인 마지막 셀의 출력을 완전연결계층(FC layer)에 통과시켜 로짓으로 변환하여 행동을 분류할 수 있다.
- [0042] 한편, UR Fall Detection 등에서 제공하는 영상이미지셋과 같이 골격좌표를 포함하지 않은 데이터의 경우 Openpose모델을 이용하여 사람객체를 검출하고, 검출된 사람객체에 대한 골격좌표를 추출한 후에, 딥러닝모델(즉, Bi-LSTM모델)을 학습시킬 수 있다.
- [0043] 여기에서, 딥러닝 기반의 Openpose모델은 도 3에 도시한 바와 같이 카메라로 입력되는 2D 사진을 이용하여 사람의 자세데이터를 추론하여 32개의 관절포인트를 획득할 수 있는데, 여러인종, 옷 및 자세들이 포함된 대량의 데이터로 학습된 인공신경망 모델이기 때문에, CCTV 안에 다양하게 나타나는 인물의 성별, 체격, 옷 배경 안에서도 효과적으로 사람의 골격을 추론할 수 있다.
- [0044] 이러한 Openpose모델에서 각 관절포인트는 x좌표와 y좌표로 구성되어 있으며, 이미지사이즈클 기준으로 0-1 사이로 노멀라이즈(normalize)되어 있어 추후 Bi-LSTM 모델의 입력데이터로 사용될 수 있다. 도 4는 Openpose모델을 통해 추론되는 신체의 골격좌표값 터미널 출력창을 보여주고 있다.
- [0045] 한편, 관계센터(300)는 CCTV 영상카메라(100)로부터 통신망(200)을 통해 수신되는 감시영상에서 Openpose모델을 이용하여 사람객체를 검출하여 경계상자(bounding box)로 표시할 수 있는데, 감시영상에서 Openpose모델을 통해 이미지프레임 내 사람객체만을 탐지하고, 탐지된 사람객체의 경계선에 따라 경계상자를 마킹하여 사람이 내부에 포함된 박스로 하여 사람객체를 표시할 수 있다.
- [0046] 또한, 관계센터(300)는 Strong sort모델을 이용한 사람객체의 위치 비교를 통해 개별아이디를 부여한 후에,

Openpose모델을 이용하여 골격좌표를 추출하여 단위벡터로 변환할 수 있는데, Strong sort모델을 이용하여 이전 프레임과 현재프레임의 객체 위치를 비교하여 일관성 있는 아이디를 부여할 수 있는 장점이 있다.

- [0047] 이러한 Strong sort모델은 연속된 프레임(영상)에서 객체들의 관계를 파악하여 객체에 ID를 부여하는 실시간 다중객체 추적 기술로, 둘 이상의 객체가 교차하여 겹치는 상황에서도 객체들의 ID를 유지하여 다중객체 추적 분야에서 정확성이 높은 기술로 알려져 있는데, 여러사람이 동시에 교차하는 영상에서도 좋은 성능을 나타낼 수 있고, 일관성 있는 아이디를 객체별로 부여함으로써, 각 아이디별로 데이터를 취합할 수 있는 장점이 있고, 여러사람에 대한 자세 구분을 동시에 수행할 수 있는 장점이 있다.
- [0048] 또한, 관계센터(300)는 사람객체에서 골격좌표를 추출하여 단위벡터로 변환한 후에 사람의 골격을 자세별로 레이블링할 수 있는데, Openpose모델을 이용하여 사람의 자세데이터를 추론하여 32개의 관절포인트를 획득할 수 있으며, 이러한 관절포인트를 통해 사람의 골격좌표를 획득할 수 있고, 추출된 골격좌표를 단위벡터로 변환할 수 있다.
- [0049] 이러한 Openpose모델에서 추론된 사람의 골격 정보들은 감시영상 내에서의 x, y 위치를 나타내며, 이 좌표들을 전처리하지 않고 Bi-LSTM모델에 입력하여 시계열 분석을 하게 되면 학습데이터에 있는 사람위치와 실제 환경에서의 사람위치가 다른 경우 모델의 추론에 악영향을 끼치게 되는데, 이러한 현상이 발생하는 이유는 Bi-LSTM모델이 감시영상 내의 사람 위치까지 같이 학습하여 사람의 자세 추론과 관계없는 사람의 위치정보가 학습에 영향을 미치기 때문이다.
- [0050] 이를 방지하기 위해 사람의 골격좌표를 벡터형식으로 변환하여 감시영상 내 사람의 위치정보가 Bi-LSTM모델에서 학습되지 않도록 전처리할 수 있고, 이 후 사람의 위치정보를 제거한 벡터형식의 자세데이터를 실제 어떤 자세를 취하고 있는지에 대해 자세별로 레이블링이 수행될 수 있다.
- [0051] 여기에서, 관계센터(300)는 다수의 객체에 대한 행동을 구분하기 위해서 사람객체에 부여된 아이디별로 상술한 바와 같이 레이블링이 완료된 벡터형식의 자세데이터를 취합할 수 있고, 취합된 단위벡터는 Bi-LSTM모델에서 처리할 수 있는 텐서(tensor)형태의 시계열데이터로 변환될 수 있다. 여기에서, 단위벡터는 영상 2-3초 길이(예를 들면, 30 프레임 기준으로 60-90장 이미지)로 취합될 수 있다.
- [0052] 다음에, 관계센터(300)는 개별아이디에 따라 시계열데이터를 취합하고, 학습된 딥러닝모델에 입력하여 사람자세를 구분할 수 있는데, 아이디별로 취합된 시계열데이터를 학습된 Bi-LSTM모델에 입력하여 사람의 행위 맥락을 파악하고, 이를 통해 사람행동을 구분한 후에, CCTV 영상카메라(100)로부터 수신된 감시영상에 실시간으로 구분 결과를 표시하여 별도의 디스플레이장치를 통해 제공(디스플레이)할 수 있다.
- [0053] 여기에서, Bi-LSTM모델은 학습 시, 이전 시점과 미래 시점의 정보 또한 학습할 수 있기 때문에, 사고 및 범죄 발생 전에 사람행동을 효과적으로 예측할 수 있으며, 이를 통해 감시영상과 딥러닝모델을 이용하여 어린이, 노인, 장애인 등의 안전사고 및 범죄를 실시간으로 탐지할 수 있다.
- [0054] 상술한 바와 같은 사람행동의 구분은 공격, 자해, 반사회적, 심정지, 수면 이상, 섭식 이상행동 등을 딥러닝모델을 통해 실시간으로 탐지하여 2차 사고를 미연에 방지할 수 있다.
- [0055] 따라서, 본 발명의 일 실시예에 따르면, CCTV 영상카메라와 관계센터가 엣지컴퓨팅환경으로 구축되며, 관계센터에서 기준데이터셋을 이용하여 딥러닝모델을 학습시킨 후에, CCTV 영상카메라로부터 감시영상이 제공될 경우 사람객체를 검출하고, 검출된 사람객체에 개별아이디를 부여하면서 골격좌표를 추출하며, 추출된 골격좌표를 딥러닝모델에 입력하여 사람자세를 구분한 후에, 구분결과를 제공함으로써, 이미지분류의 단점을 보완하여 유사행동에 대한 분류정확도를 향상시킬 수 있고, 다수의 사람들에 대한 행동을 인식 및 구분하여 정교한 이상행동 탐지를 수행할 수 있다.
- [0056] 도 5는 본 발명의 다른 실시예에 따른 딥러닝을 이용한 위급상황 모니터링 방법을 나타낸 플로우차트이고, 도 6은 본 발명의 다른 실시예에 따른 딥러닝을 이용한 위급상황 모니터링 과정에서 사람의 행동을 구분하는 것을 설명하기 위한 도면이다.
- [0057] 도 5 및 도 6을 참조하면, CCTV 영상카메라(100)와 관계센터(300)를 엣지컴퓨팅환경으로 구축할 수 있다(단계 310).
- [0058] 이러한 엣지컴퓨팅은 중앙집중식 서버나 클라우드가 아닌 데이터가 수집되고 분석되는 물리적위치 근처에서 컴퓨팅이 이루어지는 분산 컴퓨팅모델을 의미하며, 이를 사용하는 이유는 관찰하고자 하는 대상 혹은 장소가 서버

혹은 컴퓨팅장비의 설치가 원활하지 않은 사회적 약자들의 공간에서 쉽게 운용할 수 있게 하기 위함이다.

- [0059] 그리고, 관제센터(300)에서 기준데이터셋을 이용하여 딥러닝모델을 학습시킬 수 있다(단계320).
- [0060] 예를 들면, 딥러닝모델은 예를 들면, Bi-LSTM모델을 포함할 수 있는데, 사람이 쓰러지는 영상데이터셋과 수어 구분을 위한 영상데이터셋을 포함하는 기준 영상데이터셋을 이용하여 학습시킬 수 있다.
- [0061] 여기에서, 사람이 쓰러지는 영상데이터셋은 예를 들면, UR Fall Detection 등에서 제공하는 영상이미지셋을 이용할 수 있고, 수어 구분을 위한 영상데이터셋은 예를 들면, 데이콘의 수어 구분 경진대회 등에서 제공하는 영상이미지셋을 이용할 수 있다.
- [0062] 상술한 바와 같은 Bi-LSTM모델의 학습에 대해 설명하면, Bi-LSTM(양방향 LSTM)모델은 정방향으로만 학습을 진행하는 대신, 마지막 노드에서 뒤에서 앞으로(즉, 역방향) 실행되는 다른 LSTM을 추가하는 것으로, LSTM모델 대비 역방향으로 정보를 전달하는 히든레이어를 추가함으로써, 정보를 보다 유연하게 처리할 수 있다. 즉, 각 시점에서 히든스테이트가 이전 시점과 미래 시점의 정보를 모두 갖는 효과가 있기 때문에, 모델을 전체 시계열 데이터로부터 학습할 수 있는 장점이 있다.
- [0063] 한편, UR Fall Detection 등에서 제공하는 영상이미지셋과 같이 골격좌표를 포함하지 않은 데이터의 경우 Openpose모델을 이용하여 사람객체를 검출하고, 검출된 사람객체에 대한 골격좌표를 추출한 후에, 딥러닝모델(즉, Bi-LSTM모델)을 학습시킬 수 있다.
- [0064] 다음에, 관제센터(300)에서 CCTV 영상카메라(100)를 통해 촬영된 감시영상을 수신할 수 있다(단계330).
- [0065] 그리고, 관제센터(300)에서 감시영상이 수신될 경우 사람객체를 검출할 수 있다(단계340).
- [0066] 여기에서, 관제센터(300)에서는 CCTV 영상카메라(100)로부터 통신망(200)을 통해 수신되는 감시영상에서 Openpose모델을 이용하여 사람객체를 검출하여 경계상자로 표시할 수 있는데, 감시영상에서 Openpose모델을 통해 이미지프레임 내 사람객체만을 탐지하고, 탐지된 사람객체의 경계선에 따라 경계상자를 마킹하여 사람이 내부에 포함된 박스로 하여 사람객체를 표시할 수 있다.
- [0067] 또한, 관제센터(300)에서 검출된 사람객체에 개별아이디를 부여하면서 골격좌표를 추출할 수 있다(단계350).
- [0068] 여기에서, 관제센터(300)에서는 Strong sort모델을 이용한 사람객체의 위치 비교를 통해 개별아이디를 부여한 후에, Openpose모델을 이용하여 골격좌표를 추출하여 단위벡터로 변환하고, 시계열데이터로 변환할 수 있는데, Strong sort모델을 이용하여 이전프레임과 현재프레임의 객체 위치를 비교하여 일관성 있는 아이디를 부여할 수 있는 장점이 있다.
- [0069] 또한, 관제센터(300)에서는 사람객체에서 골격좌표를 추출하여 단위벡터로 변환한 후에 사람의 골격을 자세별로 레이블링할 수 있는데, Openpose모델을 이용하여 사람의 자세데이터를 추론하여 32개의 관절포인트를 획득할 수 있으며, 이러한 관절포인트를 통해 사람의 골격좌표를 획득할 수 있고, 추출된 골격좌표를 단위벡터로 변환할 수 있다.
- [0070] 이 후 사람의 위치정보를 제거한 벡터형식의 자세데이터를 실제 어떤 자세를 취하고 있는지에 대해 자세별로 레이블링이 수행될 수 있다.
- [0071] 여기에서, 관제센터(300)에서는 다수의 객체에 대한 행동을 구분하기 위해서 사람객체에 부여된 아이디별로 상술한 바와 같이 레이블링이 완료된 벡터형식의 자세데이터를 취합할 수 있고, 취합된 단위벡터는 Bi-LSTM모델에서 처리할 수 있는 텐서(tensor)형태의 시계열데이터로 변환될 수 있다. 여기에서, 단위벡터는 영상 2-3초 길이(예를 들면, 30 프레임 기준으로 60-90장 이미지)로 취합될 수 있다.
- [0072] 이어서, 관제센터(300)에서 추출된 골격좌표를 딥러닝모델에 입력하여 사람자세를 구분한 후에, 구분결과를 제공할 수 있다(단계360).
- [0073] 여기에서, 관제센터(300)에서는 개별아이디에 따라 시계열데이터를 취합하고, 학습된 딥러닝모델에 입력하여 사람자세를 구분할 수 있는데, 아이디별로 취합된 시계열데이터를 학습된 Bi-LSTM모델에 입력하여 사람의 행위 맥락을 파악하고, 이를 통해 사람행동을 구분한 후에, CCTV 영상카메라(100)로부터 수신된 감시영상에 실시간으로 구분결과를 표시하여 별도의 디스플레이장치를 통해 제공(디스플레이)할 수 있다.
- [0074] 여기에서, Bi-LSTM모델은 학습 시, 이전 시점과 미래 시점의 정보 또한 학습할 수 있기 때문에, 사고 및 범죄 발생 전에 사람행동을 효과적으로 예측할 수 있으며, 이를 통해 감시영상과 딥러닝모델을 이용하여 어린이, 노

인, 장애인 등의 안전사고 및 범죄를 실시간으로 탐지할 수 있다.

[0075] 상술한 바와 같은 사람행동의 구분은 공격, 자해, 반사회적, 심정지, 수면 이상, 섭식 이상행동 등을 딥러닝모델을 통해 실시간으로 탐지하여 2차 사고를 미연에 방지할 수 있다.

[0076] 한편, Bi-LSTM모델을 이용하여 사람의 행동을 구분하는 것에 대해 도 6을 참조하여 설명하면, 기존 CNN모델을 기반으로 하는 객체 검출(object detection)을 이용하는 방식에서는 단일 프레임만 분석이 가능하기 때문에, 사람이 넘어졌을 때 의도적으로 누운 것인지, 부상을 당하여 넘어진 것인지 구분할 수 없지만, 이러한 문제점을 보완하는 본 발명에서는 Bi-LSTM모델이 CCTV 감시영상의 단일프레임이 아닌 시계열성 특징을 갖는 다중 프레임으로 입력되며, 복수의 프레임으로 이루어진 동영상과 같은 경향성 분석이 이루어질 수 있으며, 이를 통해 어린이, 노인, 장애인 등의 쓰러짐, 누워있기, 밀고 당기기, 머리를 때리는 자해, 발차기, 주먹질 등의 이상행동을 오로지 CCTV 감시영상만으로 효과적으로 구분할 수 있다.

[0077] 따라서, 본 발명의 다른 실시예에 따르면, CCTV 영상카메라와 관제센터가 엣지컴퓨팅환경으로 구축되며, 관제센터에서 기준데이터셋을 이용하여 딥러닝모델을 학습시킨 후에, CCTV 영상카메라로부터 감시영상이 제공될 경우 사람객체를 검출하고, 검출된 사람객체에 개별아이디를 부여하면서 골격좌표를 추출하며, 추출된 골격좌표를 딥러닝모델에 입력하여 사람자세를 구분한 후에, 구분결과를 제공함으로써, 이미지분류의 단점을 보완하여 유사행동에 대한 분류정확도를 향상시킬 수 있고, 다수의 사람들에 대한 행동을 인식 및 구분하여 정교한 이상행동 탐지를 수행할 수 있다.

[0078] 이상의 설명에서는 본 발명의 다양한 실시예들을 제시하여 설명하였으나 본 발명이 반드시 이에 한정되는 것은 아니며, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자라면 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 여러 가지 치환, 변형 및 변경이 가능함을 쉽게 알 수 있을 것이다.

부호의 설명

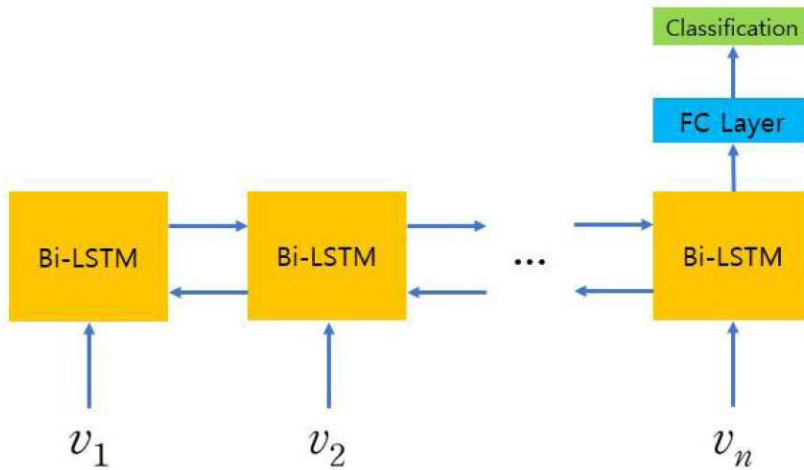
- [0079] 100 : CCTV 영상카메라
- 200 : 통신망
- 300 : 관제센터

도면

도면1



도면2



도면3

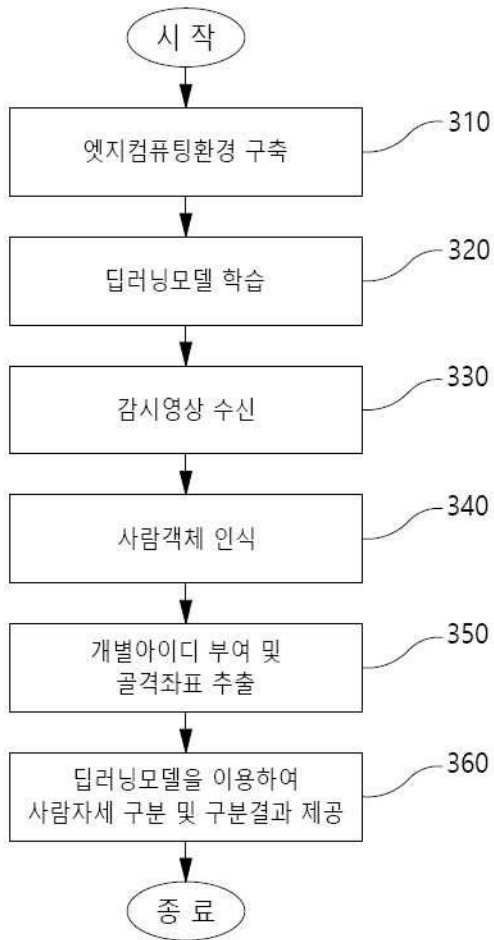


도면4

```

brighten@brightenai: ~/brighten-internship/action-classifier
left_knee - x : 1252.548583984375 y : 640.9558715820312
right_knee - x : 1192.2548828125 y : 648.1455688476562
left_ankle - x : 1246.8062744140625 y : 731.54638671875
right_ankle - x : 1197.9971923828125 y : 751.6776123046875
----- keypoints -----
nose - x : 906.8392333984375 y : 340.4401550292969
left_eye - x : 916.84326171875 y : 331.79156494140625
right_eye - x : 901.1226196289062 y : 331.79156494140625
left_ear - x : 945.4262084960938 y : 336.1158447265625
right_ear - x : 896.835205078125 y : 336.1158447265625
left_shoulder - x : 968.2926025390625 y : 390.8902893066406
right_shoulder - x : 888.2603149414062 y : 386.5660095214844
left_elbow - x : 966.8634033203125 y : 460.0790710449219
right_elbow - x : 869.681396484375 y : 458.63763427734375
left_wrist - x : 916.84326171875 y : 411.0703430175781
right_wrist - x : 885.4020385742188 y : 408.1874694824219
left_hip - x : 932.5638427734375 y : 537.9164428710938
right_hip - x : 885.4020385742188 y : 535.0336303710938
left_knee - x : 941.1387939453125 y : 647.46533203125
right_knee - x : 899.6935424804688 y : 646.02392578125
left_ankle - x : 959.7177124023438 y : 755.57275390625
right_ankle - x : 909.697509765625 y : 744.0413818359375
    
```

도면5



도면6

