



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년11월09일
(11) 등록번호 10-2601233
(24) 등록일자 2023년11월07일

(51) 국제특허분류(Int. Cl.)
G06V 10/46 (2022.01) G06N 3/08 (2023.01)
G06T 7/00 (2017.01) G06V 10/32 (2022.01)
H04N 7/18 (2023.01)
(52) CPC특허분류
G06V 10/469 (2023.08)
G06N 3/08 (2023.01)
(21) 출원번호 10-2021-0118882
(22) 출원일자 2021년09월07일
심사청구일자 2021년09월07일
(65) 공개번호 10-2023-0036249
(43) 공개일자 2023년03월14일
(56) 선행기술조사문헌
KR1020190127510 A*
KR1020190139539 A*
KR1020200103194 A*
KR102189373 B1*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
백성욱
서울특별시 광진구 아차산로 262, B동 1304호 (자양동, 더샵스타시티)
이미영
서울특별시 강남구 학동로82길 6 (삼성동, 삼성리치빌)
(뒷면에 계속)
(74) 대리인
특허법인위더피플

전체 청구항 수 : 총 5 항

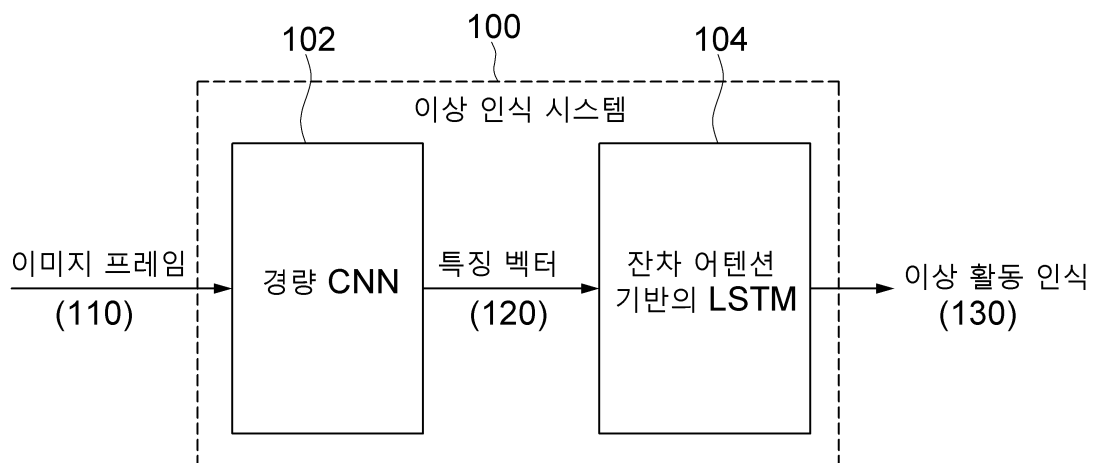
심사관 : 강현일

(54) 발명의 명칭 LSTM 기반의 이상 인식 방법 및 시스템

(57) 요약

본 발명은 LSTM 기반의 이상 인식 방법에 관한 것이다. LSTM 기반의 이상 인식 방법은, 이상 인식의 판단 대상이 되는 연속된 복수의 이미지 프레임을 수신하는 단계, 수신된 복수의 이미지 프레임을 학습된 경량 CNN 모델에 제공하여, 복수의 이미지 프레임의 특징을 나타내는 특징 벡터를 생성하는 단계 및 생성된 특징 벡터를 잔차 어텐션 기반의 LSTM에 제공하여 복수의 이미지 프레임에 대응하는 이상 활동을 인식하는 단계를 포함한다.

대표도 - 도1



- (52) CPC특허분류
G06T 7/0004 (2013.01)
G06V 10/32 (2023.08)
H04N 7/18 (2023.01)

줄피카 아마드 칸

서울특별시 광진구 능동로 209 세종대학교 대양AI
센터 411호

- (72) 발명자

와심 올라

서울특별시 광진구 능동로 209 세종대학교 대양AI
센터 411호

탄비어 후세인

서울특별시 광진구 능동로 209 세종대학교 대양AI
센터 411호

이 발명을 지원한 국가연구개발사업

| | |
|-------------|--------------------------------|
| 과제고유번호 | 1711103369 |
| 과제번호 | 2019-0-00136-003 |
| 부처명 | 과학기술정보통신부 |
| 과제관리(전문)기관명 | 정보통신기획평가원 |
| 연구사업명 | 인공지능융합선도프로젝트(R&D) |
| 연구과제명 | 스마트시티 산업 생산성 혁신을 위한 AI융합 기술 개발 |
| 기 여 율 | 1/1 |
| 과제수행기관명 | 세종대학교 산학협력단 |
| 연구기간 | 2021.01.01 ~ 2021.12.31 |

공지예외적용 : 있음

명세서

청구범위

청구항 1

적어도 하나의 프로세서에 의해 수행되는 LSTM 기반의 이상 인식(anomaly recognition) 방법으로서,

이상 인식의 판단 대상이 되는 연속된 복수의 이미지 프레임을 수신하는 단계;

상기 수신된 복수의 이미지 프레임을 학습된 경량 CNN(convolutional neural network) 모델에 제공하여, 상기 복수의 이미지 프레임의 특징을 나타내는 특징 벡터(feature vector)를 생성하는 단계; 및

상기 생성된 특징 벡터를 잔차 함수를 기초로 학습된 잔차 어텐션(residual attention) 기반의 LSTM(long short-term memory)에 제공하여 상기 복수의 이미지 프레임에 대응하는 이상 활동을 인식하는 단계;

를 포함하고,

상기 잔차 어텐션 기반의 LSTM은 훈련 시간을 감소시키기 위해 상기 특징 벡터를 정규화하기 위한 정규화 레이어(normalization layer)를 포함하고,

상기 잔차 어텐션 기반의 LSTM은 셀프 어텐션 레이어(self-attention layer)와 연관되고,

상기 셀프 어텐션 레이어는 상기 특징 벡터를 기초로 상기 복수의 이미지 프레임의 연속적인 특징에 대한 상황 인식(context-aware) 벡터를 생성하고,

상기 잔차 어텐션 기반의 LSTM은 이상 활동을 분류하기 위한 소프트 맥스 레이어(Softmax layer)와 연관되고,

상기 이상 활동을 인식하는 단계는,

상기 생성된 특징 벡터를 상기 잔차 어텐션 기반의 LSTM에 제공하여 출력된 출력 데이터를 상기 소프트 맥스 레이어에 입력하여 상기 이상 활동을 인식하는 단계;

를 포함하고,

상기 잔차 어텐션 기반의 LSTM은 Adam 최적화(optimization) 함수 및 크로스 엔트로피(cross-entropy) 손실 함수와 연관된, LSTM 기반의 이상 인식 방법.

청구항 2

제1항에 있어서,

상기 경량 CNN은 깊이 별(depth-wise) 구별 가능한 컨볼루션 블록(convolutional blocks)을 포함하고,

상기 컨볼루션 블록은 3 x 3 필터(filter)와 연관된 깊이 별(depth-wise) 컨볼루션 레이어 및 상기 깊이 별 컨볼루션 레이어에 의해 필터링된 값을 병합하기 위한 1 x 1 필터와 연관된 포인트 별(point-wise) 컨볼루션 레이어를 포함하는, LSTM 기반의 이상 인식 방법.

청구항 3

제1항에 있어서,

상기 경량 CNN은 입력 데이터의 채널 수를 확장하기 위한 확장 레이어(expansion layer)를 포함하는, LSTM 기반의 이상 인식 방법.

청구항 4

삭제

청구항 5

삭제

청구항 6

삭제

청구항 7

삭제

청구항 8

제1항 내지 제3항 중 어느 한 항에 따른 LSTM 기반의 이상 인식 방법을 컴퓨터에서 실행하기 위해 컴퓨터 판독 가능한 기록 매체에 저장된 컴퓨터 프로그램.

청구항 9

이상 인식 시스템으로서,

통신 모듈;

메모리; 및

상기 메모리와 연결되고, 상기 메모리에 포함된 컴퓨터 판독 가능한 적어도 하나의 프로그램을 실행하도록 구성된 적어도 하나의 프로세서

를 포함하고,

상기 적어도 하나의 프로그램은,

이상 인식의 판단 대상이 되는 연속된 복수의 이미지 프레임을 수신하고,

상기 수신된 복수의 이미지 프레임을 학습된 경량 CNN 모델에 제공하여, 상기 복수의 이미지 프레임의 특징을 나타내는 특징 벡터를 생성하고,

상기 생성된 특징 벡터를 잔차 함수를 기초로 학습된 잔차 어텐션 기반의 LSTM에 제공하여 상기 복수의 이미지 프레임에 대응하는 이상 활동을 인식하기 위한 명령어들을 포함하고,

상기 잔차 어텐션 기반의 LSTM은 훈련 시간을 감소시키기 위해 상기 특징 벡터를 정규화하기 위한 정규화 레이어(normalization layer)를 포함하고,

상기 잔차 어텐션 기반의 LSTM은 셀프 어텐션 레이어(self-attention layer)와 연관되고,

상기 셀프 어텐션 레이어는 상기 특징 벡터를 기초로 상기 복수의 이미지 프레임의 연속적인 특징에 대한 상황 인식(context-aware) 벡터를 생성하고,

상기 잔차 어텐션 기반의 LSTM은 이상 활동을 분류하기 위한 소프트 맥스 레이어(Softmax layer)와 연관되고,

상기 이상 활동을 인식하는 것은,

상기 생성된 특징 벡터를 상기 잔차 어텐션 기반의 LSTM에 제공하여 출력된 출력 데이터를 상기 소프트 맥스 레이어에 입력하여 상기 이상 활동을 인식하는 것을 포함하고,

상기 잔차 어텐션 기반의 LSTM은 Adam 최적화(optimization) 함수 및 크로스 엔트로피(cross-entropy) 손실 함수와 연관된, 이상 인식 시스템.

발명의 설명

기술분야

[0001] 본 발명은 LSTM 기반의 이상 인식 방법 및 시스템에 관한 것으로, 구체적으로, 경량 CNN 및 새로운 유형의 LSTM을 이용하여 이미지 내에 포함된 이상 활동을 인식하는 방법 및 시스템에 관한 것이다.

배경기술

[0003] 안전한 스마트 시티 환경을 위해 감시 카메라를 통해 촬영된 영상에서 폭행, 범죄, 도로 사고 등의 이상 활동을 인식하는 것이 요구된다. 그러나, 실생활에서 이러한 이상 활동은 다양하고, 복잡하며, 빈번하게 발생하지 않으므로, 임의의 시스템이 이상 활동을 정확히 인식하는 것은 어렵다.

[0004] 한편, 공공의 안전을 위해 감시 카메라에 의해 촬영된 영상을 실시간으로 분석하여 이상 활동을 인식하는 것은 중요하다. 그러나, 공공에 설치된 대부분의 감시 카메라는 녹화 기능만을 지원하고, 실시간 모니터링 기능을 지원하지 못한다. 이에 따라, 이상 활동을 인식하기 위해, 전문가가 녹화된 영상을 직접 보고 이상 활동이 발생했는지 여부를 판단해야 한다. 이와 같이, 전문가가 직접 판단하는 경우, 이상 활동의 발생을 인식하는데 소요되는 시간이 급격히 증가하게 된다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 상기와 같은 문제점을 해결하기 위한 LSTM 기반의 이상 인식 방법, 기록매체에 저장된 컴퓨터 프로그램 및 시스템(장치)을 제공한다.

과제의 해결 수단

[0008] 본 발명은 방법, 시스템(장치) 또는 판독 가능한 저장(기록) 매체에 저장된 컴퓨터 프로그램을 포함한 다양한 방식으로 구현될 수 있다.

[0009] 본 발명의 일 실시예에 따르면, 적어도 하나의 프로세서에 의해 수행되는 LSTM 기반의 이상 인식 방법은, 이상 인식의 판단 대상이 되는 연속된 복수의 이미지 프레임을 수신하는 단계, 수신된 복수의 이미지 프레임을 학습된 경량 CNN 모델에 제공하여, 복수의 이미지 프레임의 특징을 나타내는 특징 벡터를 생성하는 단계 및 생성된 특징 벡터를 잔차 어텐션 기반의 LSTM에 제공하여 복수의 이미지 프레임에 대응하는 이상 활동을 인식하는 단계를 포함한다.

[0010] 본 발명의 일 실시예에 따르면, 경량 CNN은 깊이 별 구별 가능한 컨볼루션 블록을 포함한다. 컨볼루션 블록은 3 x 3 필터와 연관된 깊이 별 컨볼루션 레이어 및 깊이 별 컨볼루션 레이어에 의해 필터링된 값을 병합하기 위한 1 x 1 필터와 연관된 포인트 별 컨볼루션 레이어를 포함한다.

[0011] 본 발명의 일 실시예에 따르면, 경량 CNN은 입력 데이터의 채널 수를 확장하기 위한 확장 레이어를 포함한다.

[0012] 본 발명의 일 실시예에 따르면, 잔차 어텐션 기반의 LSTM은 훈련 시간을 감소시키기 위해 특징 벡터를 정규화하기 위한 정규화 레이어를 포함한다.

[0013] 본 발명의 일 실시예에 따르면, 잔차 어텐션 기반의 LSTM은 셀프 어텐션 레이어와 연관된다. 셀프 어텐션 레이어는 특징 벡터를 기초로 복수의 이미지 프레임의 연속적인 특징에 대한 상황 인식 벡터를 생성한다.

[0014] 본 발명의 일 실시예에 따르면, 잔차 어텐션 기반의 LSTM은 이상 활동을 분류하기 위한 소프트 맥스 레이어와 연관된다. 이상 활동을 인식하는 단계는, 생성된 특징 벡터를 잔차 어텐션 기반의 LSTM에 제공하여 출력된 출력 데이터를 소프트 맥스 레이어에 입력하여 이상 활동을 인식하는 단계를 포함한다.

[0015] 본 발명의 일 실시예에 따르면, 잔차 어텐션 기반의 LSTM은 Adam 최적화 함수 및 크로스 엔트로피 손실 함수와 연관된다.

[0016] 본 발명의 일 실시예에 따른 상술된 LSTM 기반의 이상 인식 방법을 컴퓨터에서 실행하기 위해 컴퓨터 판독 가능한 기록 매체에 저장된 컴퓨터 프로그램이 제공된다.

[0017] 본 발명의 일 실시예에 따른 이상 인식 시스템은, 통신 모듈, 메모리 및 메모리와 연결되고, 메모리에 포함된 컴퓨터 판독 가능한 적어도 하나의 프로그램을 실행하도록 구성된 적어도 하나의 프로세서를 포함한다. 적어도

하나의 프로그램은, 이상 인식의 판단 대상이 되는 연속된 복수의 이미지 프레임을 수신하고, 수신된 복수의 이미지 프레임을 학습된 경량 CNN 모델에 제공하여, 복수의 이미지 프레임의 특징을 나타내는 특징 벡터를 생성하고, 생성된 특징 벡터를 잔차 어텐션 기반의 LSTM에 제공하여 복수의 이미지 프레임에 대응하는 이상 활동을 인식하기 위한 명령어들을 포함한다.

발명의 효과

- [0019] 본 발명의 다양한 실시예에서 이상 인식 시스템은 경량화된 CNN 및 새로운 유형의 LSTM을 이용하여 실시간으로 발생하는 이상 활동을 높은 정확도로 인식할 수 있다.
- [0020] 본 발명의 다양한 실시예에서 셀프 어텐션 레이어를 이용함으로써, 잔차 어텐션 기반의 LSTM은 복수의 입력이 요구되는 다른 모델과 달리, 특징 벡터와 같은 하나의 입력만을 수신하여 효율적으로 이상 활동을 인식할 수 있다.
- [0021] 본 발명의 다양한 실시예에서 Adam 최적화 함수 및 크로스 엔트로피 손실 함수가 적용됨으로써, 이상 활동 인식 성능이 극대화될 수 있다.
- [0022] 본 발명의 다양한 실시예에서 잔차를 이용함으로써 학습이 진행되면서 각 파라미터에 대한 가중치가 변형되는 기울기 소실 문제가 효과적으로 해결될 수 있다.
- [0023] 본 발명의 다양한 실시예에서 잔차 어텐션 기반의 LSTM은 더 적은 컴퓨팅 파워로도 실시간으로 이상 활동을 인식할 수 있다.
- [0024] 본 발명의 효과는 이상에서 언급한 효과로 제한되지 않으며, 언급되지 않은 다른 효과들은 청구범위의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자("통상의 기술자"라 함)에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [0026] 본 발명의 실시예들은, 이하 설명하는 첨부 도면들을 참조하여 설명될 것이며, 여기서 유사한 참조 번호는 유사한 요소들을 나타내지만, 이에 한정되지는 않는다.

도 1은 본 발명의 일 실시예에 따른 이상 인식 시스템의 예시를 나타내는 도면이다.
 도 2는 본 발명의 일 실시예에 따른 경량 CNN의 내부 구성을 나타내는 예시적인 도면이다.
 도 3은 본 발명의 일 실시예에 따른 잔차 어텐션 기반의 LSTM의 내부 구성을 나타내는 예시적인 도면이다.
 도 4는 본 발명의 일 실시예에 따른 잔차 어텐션 기반의 LSTM의 특징 추출 과정을 나타내는 예시적인 도면이다.
 도 5는 본 개시의 일 실시예에 따른 잔차 어텐션 기반의 LSTM의 이상 활동에 대한 예측 결과를 나타내는 예시적인 혼동 행렬이다.
 도 6은 본 발명의 일 실시예에 따른 LSTM 기반의 이상 인식 방법의 예시를 나타내는 흐름도이다.
 도 7은 본 발명의 일 실시예에 따른 이상 인식 시스템의 내부 구성을 나타내는 블록도이다.

발명을 실시하기 위한 구체적인 내용

- [0027] 이하, 본 발명의 실시를 위한 구체적인 내용을 첨부된 도면을 참조하여 상세히 설명한다. 다만, 이하의 설명에서는 본 발명의 요지를 불필요하게 흐릴 우려가 있는 경우, 널리 알려진 기능이나 구성에 관한 구체적 설명은 생략하기로 한다.
- [0028] 첨부된 도면에서, 동일하거나 대응하는 구성요소에는 동일한 참조부호가 부여되어 있다. 또한, 이하의 실시예들의 설명에 있어서, 동일하거나 대응되는 구성요소를 중복하여 기술하는 것이 생략될 수 있다. 그러나, 구성요소에 관한 기술이 생략되어도, 그러한 구성요소가 어떤 실시예에 포함되지 않는 것으로 의도되지는 않는다.
- [0029] 개시된 실시예의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 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 발명이 완전하도록 하고, 본 발명이 통상의 기술자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것일 뿐이다.

- [0030] 본 명세서에서 사용되는 용어에 대해 간략히 설명하고, 개시된 실시예에 대해 구체적으로 설명하기로 한다. 본 명세서에서 사용되는 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나, 이는 관련 분야에 종사하는 기술자의 의도 또는 관례, 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 발명의 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서, 본 발명에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 발명의 전반에 걸친 내용을 토대로 정의되어야 한다.
- [0031] 본 명세서에서의 단수의 표현은 문맥상 명백하게 단수인 것으로 특정하지 않는 한, 복수의 표현을 포함한다. 또한, 복수의 표현은 문맥상 명백하게 복수인 것으로 특정하지 않는 한, 단수의 표현을 포함한다. 명세서 전체에서 어떤 부분이 어떤 구성요소를 포함한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있음을 의미한다.
- [0032] 본 발명에서, "포함하다", "포함하는" 등의 용어는 특징들, 단계들, 동작들, 요소들 및/또는 구성 요소들이 존재하는 것을 나타낼 수 있으나, 이러한 용어가 하나 이상의 다른 기능들, 단계들, 동작들, 요소들, 구성 요소들 및/또는 이들의 조합이 추가되는 것을 배제하지는 않는다.
- [0033] 본 발명에서, 특정 구성 요소가 임의의 다른 구성 요소에 "결합", "조합", "연결" 되거나, "반응" 하는 것으로 언급된 경우, 특정 구성 요소는 다른 구성 요소에 직접 결합, 조합 및/또는 연결되거나, 반응할 수 있으나, 이에 한정되지 않는다. 예를 들어, 특정 구성 요소와 다른 구성 요소 사이에 하나 이상의 중간 구성 요소가 존재할 수 있다. 또한, 본 발명에서 "및/또는"은 열거된 하나 이상의 항목의 각각 또는 하나 이상의 항목의 적어도 일부의 조합을 포함할 수 있다.
- [0034] 도 1은 본 발명의 일 실시예에 따른 이상 인식 시스템(100)의 예시를 나타내는 도면이다. 이상 인식 시스템(100)은 이미지 센서(예: 감시 카메라 등)에 의해 촬영된 영상을 수신하고, 수신된 영상을 기초로 이상 활동의 발생 여부를 인식하기 위한 시스템을 지칭할 수 있다. 예를 들어, 이상 인식 시스템(100)은 이미지 센서 내부에 포함되거나, 하나 이상의 이미지 센서와 연관될 수 있다. 도시된 바와 같이, 이상 인식 시스템(100)은 경량 CNN(light-weight convolutional neural network)(102)과 잔차 어텐션(residual attention) 기반의 LSTM(long short-term memory)(104)을 포함할 수 있다.
- [0035] 일 실시예에 따르면, 이상 인식 시스템(100)에 포함된 경량 CNN(102)은 이상 인식의 판단 대상이 되는 연속된 복수의 이미지 프레임(110)을 수신할 수 있다. 예를 들어, 연속된 복수의 이미지 프레임(110)은 영상으로부터 추출된 연속된 30개의 이미지 프레임일 수 있으나, 이에 한정되지 않는다. 또한, 경량 CNN(102)은 연산량과 연산에 소요되는 컴퓨팅 파워, 연산 시간 등을 감소시키기 위해 경량화된 인공신경망을 지칭할 수 있다. 예를 들어, 경량 CNN(102)은 거대한 컨볼루션 레이어(convolutional layer)가 깊이 별 구별가능한 컨볼루션 블록(depth-wise distinguishable convolution block)으로 대체된 MobileNet을 포함할 수 있다.
- [0036] 이상 인식 시스템(100)은 수신된 복수의 이미지 프레임(110)을 학습된 경량 CNN 모델에 제공하여, 복수의 이미지 프레임(110)의 특징을 나타내는 특징 벡터(feature vector)(120)를 생성할 수 있다. 다시 말해, 경량 CNN(102)은 영상으로부터 추출된 복수의 이미지 프레임(110)을 수신하고, 해당 복수의 이미지 프레임(110)에 대응하는 특징 벡터(120)를 생성할 수 있다. 예를 들어, 하나의 영상에서 복수의 세트의 이미지 프레임들이 추출될 수 있으며, 경량 CNN(102)은 이러한 복수의 세트의 이미지 프레임 각각에 대하여 대응하는 특징 벡터를 생성할 수 있다. 여기서, 특징 벡터(120)는 복수의 이미지 각각에 포함된 특징점(feature point)을 기초로 생성된 벡터일 수 있으며, 특징점은 이미지 내의 코너(corner), 엣지(edge) 등을 기초로 추출될 수 있으나, 이에 한정되지 않는다.
- [0037] 이상 인식 시스템(100)은 생성된 특징 벡터(120)를 잔차 어텐션 기반의 LSTM(104)에 제공하여 복수의 이미지 프레임(110)에 대응하는 이상 활동을 인식할 수 있다(130). 다시 말해, 잔차 어텐션 기반의 LSTM(104)은 경량 CNN(102)으로부터 특징 벡터(120)를 수신하고, 특징 벡터(120)에 포함된 정보를 기초로 복수의 이미지 프레임(110)에 대응하는 이상 활동을 인식할 수 있다. 여기서, 잔차 어텐션 기반의 LSTM(104)은 셀프 어텐션 레이어(self-attention layer)와 연관된 잔차 LSTM으로서, 하나의 정보가 아닌 연속된 정보(예: 연속된 복수의 이미지 프레임 등)에 대한 연산을 수행하기 위한 인공신경망일 수 있다. 즉, 잔차 어텐션 기반의 LSTM(104)은 특징 벡터(120)를 기초로 단일의 이미지 프레임이 아닌 연속된 복수의 이미지 프레임(110)에 포함된 정보를 고려하여 이상 활동의 발생 여부를 인식할 수 있다.
- [0038] 도 1에서는 이상 인식 시스템(100)이 이상 활동의 발생 여부를 인식하는 것으로 상술되었으나, 이에 한정되지

않으며, 이상 인식 시스템(100)은 어떤 종류의 이상 활동이 발생했는지 여부를 인식할 수도 있다. 이와 같은 구성에 의해, 이상 인식 시스템(100)은 경량화된 CNN(102) 및 새로운 유형의 LSTM(104)을 이용하여 실시간으로 발생하는 이상 활동을 높은 정확도로 인식할 수 있다.

[0039] 도 2는 본 발명의 일 실시예에 따른 경량 CNN(220)의 내부 구성을 나타내는 예시적인 도면이다. 상술된 바와 같이, 경량 CNN(220)은 복수의 세트의 이미지 프레임들(210)(예: 복수의 이미지 프레임(i), 복수의 이미지 프레임(i+1), 복수의 이미지 프레임(N) 등)을 수신하고, 각 세트의 이미지 프레임들에 대한 특징 벡터(230)를 생성할 수 있다. 예를 들어, 경량 CNN(220)은 MobileNetV2와 연관되거나, MobileNetV2를 기초로 생성된 인공신경망일 수 있다.

[0040] 일 실시예에 따르면, 경량 CNN(220)은 컨볼루션 레이어가 대체된 깊이 별 구별 가능한 컨볼루션 블록을 포함할 수 있다. 이러한 컨볼루션 블록은 3 x 3 필터(filter)와 연관된 깊이 별 컨볼루션 레이어 및 깊이 별 컨볼루션 레이어에 의해 필터링된 값을 병합하기 위한 1 x 1 필터와 연관된 포인트 별 컨볼루션 레이어를 포함할 수 있다. 즉, 경량 CNN(220)은 거대한 컨볼루션 레이어를 이러한 컨볼루션 블록으로 대체하여 경량화될 수 있으며, 이에 따라 적은 컴퓨팅 파워로 빠르게 동작할 수 있다.

[0041] 일 실시예에 따르면, 경량 CNN(220)은 입력 데이터의 채널 수를 확장하기 위한 확장 레이어(expansion layer)를 포함할 수 있다. 예를 들어, 확장 레이어는 깊이 별 컨볼루션 레이어로 데이터가 전달되기 이전에, 채널 수를 확장하여 데이터를 전달할 수 있다. 다시 말해, 확장 레이어는 주어진 입력 채널의 수보다 더 많은 출력 채널을 생성할 수 있다. 또한, 포인트 별 컨볼루션 레이어는 프로젝션(projection) 레이어의 부분일 수 있다. 즉 포인트 별 컨볼루션 레이어는 다수의 채널을 더 작은 수의 채널로 프로젝팅하여 데이터를 전달할 수 있다.

[0042] 경량 CNN(220)은 병목 잔차 블록(bottleneck residual block)을 포함할 수 있다. 이러한 병목 잔차 블록은 특징 벡터(230)를 생성하기 위해 사용되는 파라미터의 수를 감소시키거나 매트릭스 곱(matrix multiplication)의 수를 감소시킬 수 있다. 또한, 경량 CNN(220)은 활성화 함수로 사용되는 ReLU 및 정규화 레이어(예: 배치 정규화 레이어) 등을 포함할 수 있다. 또한, 경량 CNN(220)은 오버피팅(overfitting) 문제를 감소시키는 전역 평균 풀링 레이어(global average pooling layer)를 포함할 수 있다.

[0043] 도 3은 본 발명의 일 실시예에 따른 잔차 어텐션 기반의 LSTM(310)의 내부 구성을 나타내는 예시적인 도면이다. 상술된 바와 같이, 잔차 어텐션 기반의 LSTM(310)은 특징 벡터(230)를 수신하고, 수신된 특징 벡터(230)를 기초로 이상 활동(330)을 인식할 수 있다. 예를 들어, LSTM은 RNN(recurrent neural network)의 일종으로서, 입력 게이트(input gate) 및 망각 게이트(forget gate)에 의해 제어되는 메모리 셀(memory cell)을 포함할 수 있다.

[0044] 일 실시예에 따르면, 셀 상태(cell state)는 망각 게이트에 의해 변경되거나, 입력 데이터에 의해 수정될 수 있다. 여기서, 망각 게이트는 이전 메모리에서 다음 시간 단계로 전달해야 하는 정보 및/또는 데이터의 양을 결정할 수 있다. 즉, 이전 메모리에 저장된 정보 및/또는 데이터가 다음 시간 단계에서 연산 등을 위해 필요한 정도를 판정하고, 판정된 양의 정보 및/또는 데이터가 전달될 수 있다. 또한, 입력 게이트는 메모리 셀에 입력해야 하는 새로운 정보 및/또는 데이터의 양을 결정할 수 있다. 이와 같은 입력 게이트 및 망각 게이트를 이용하여, LSTM은 순차적인 정보의 장단기 종속성의 문제를 효과적으로 처리할 수 있다. 이와 같은 LSTM은 다음의 수학적 1과 같이 결정될 수 있다.

수학적 1

[0046]
$$C_t^s = \tanh(\hat{w}_g h_t + \hat{w}_{xg} * x_t^s)$$

[0047]
$$f_t^s = \sigma(\hat{w}_f h_t + \hat{w}_{xf} * x_t^s)$$

[0048]
$$i_t^s = \sigma(\hat{w}_i h_t + \hat{w}_{xi} * x_t^s)$$

[0049]
$$\bar{O}_t^s = \sigma(\hat{w}_o h_t + \hat{w}_{xo} * x_t^s)$$

[0050]
$$C_t^s = f_t^s \odot C_{t-1}^s + i_t^s * C_t^s$$

[0051]
$$h_t^s = \bar{O}_t^s \odot \tanh(C_t^s)$$

[0053] 여기서, $\hat{w}h_g$, $\hat{w}h_f$, $\hat{w}h_i$, $\hat{w}h_o$ 는 히든 상태(hidden state)에 관한 이전 시간 단계의 정보를 제어하고 모니터링하는데 사용되는 파라미터일 수 있으며, \hat{w}_{xg}^s , \hat{w}_{xf}^s , \hat{w}_{xi}^s , \hat{w}_{xo}^s 는 현재 입력 시간 단계의 가중치 매트릭(matrices)에 적용되는 파라미터일 수 있다. 또한, 위 첨자 s는 입력 시퀀스 정보를 나타내고, 아래 첨자 t는 시간 단계 정보를 나타낼 수 있다. σ 는 시그모이드 함수에 사용될 수 있으며, \odot 는 요소별 곱셈을 뜻하는 Hadamard product 연산자일 수 있다. 이와 같은 구성에 의해, LSTM은 각 시간 단계의 반복 모듈 간의 상호작용을 통해 이전 단계에서 얻은 정보가 다음 단계에서도 지속되도록 할 수 있으며, 그에 따라, 긴 기간의 의존성(long-term dependencies) 문제를 해결할 수 있다.

[0054] 일 실시예에 따르면, 잔차 어텐션 기반의 LSTM(310)은 정규화 레이어(normalization layer)를 포함할 수 있다. 정규화 레이어에 의해 정보가 정규화 되는 경우, 잔차 어텐션 기반의 LSTM(310)의 훈련 시간이 감소될 수 있다. 예를 들어, 다음의 수학적 식 2에 의해 정규화가 수행될 수 있다.

수학적 식 2

[0056]
$$\tilde{n}_t = \frac{1}{h} \sum_{i=0}^h (ht)i$$

[0057]
$$\delta_t = \sqrt{\frac{1}{h} \sum_{i=0}^h ((ht)i - \tilde{n}_t)^2}$$

[0058]
$$\hat{y}_t = f\left(\frac{\hat{g}}{\delta_t} \odot (ht - \tilde{n}_t) + b\right)$$

[0060] 여기서, $(ht)i$ 은 i번째 뉴런의 LSTM(310)의 각 레이어의 히든 상태일 수 있으며, \hat{g} 및 b 는 활성화 함수 f에 대한 입력 시퀀스를 리스케일(rescale)하기 위해 사용되는 훈련 가능한 가중치일 수 있다. 또한, t는 시간 단계(time step)를 나타낼 수 있다. 예를 들어, 오버 피팅을 감소시키기 위해, 잔차 어텐션 기반의 LSTM(310)의 각 레이어에 0.5의 탈락 임계값(dropout threshold)이 적용될 수 있다.

[0061] 일 실시예에 따르면, 잔차 어텐션 기반의 LSTM(310)은 셀프 어텐션 레이어(self-attention layer)와 연관될 수 있다. 이 경우, 셀프 어텐션 레이어는 특징 벡터(230)를 기초로 복수의 이미지 프레임의 연속적인 특징에 대한 상황 인식(context-aware) 벡터를 생성할 수 있다. 여기서, 상황 인식 벡터는 연속된 복수의 이미지 프레임의 순차적인 특징(예: 특징 간의 상관관계 등)을 포함하는 벡터일 수 있다. 이와 같은 셀프 어텐션 레이어를 이용함으로써, 잔차 어텐션 기반의 LSTM(310)은 복수의 입력이 요구되는 다른 모델과 달리, 특징 벡터(230)와 같은 하나의 입력만을 수신하여 효율적으로 이상 활동을 인식할 수 있다.

[0062] 일 실시예에 따르면, 잔차 어텐션 기반의 LSTM(310)은 이상 활동을 분류하기 위한 소프트 맥스 레이어(Softmax layer)(320)와 연관되고, 소프트 맥스 레이어(320)는 잔차 어텐션 기반의 LSTM(310)에 의해 출력된 정보를 기초로 이상 활동을 인식할 수 있다. 예를 들어, 잔차 어텐션 기반의 LSTM(310) 및/또는 소프트 맥스 레이어(320)는 다음의 수학적 식 3에 의해 이상 활동을 인식할 수 있다.

수학적 식 3

[0064]
$$k_t = \frac{1}{h} \sum_{i=1}^h \hat{w}_t^i f_i$$

[0065]
$$\hat{S}_t = \hat{w}^T \tanh(\hat{w}_b ht + M_b R_f + b_b)$$

[0066]
$$A_t = \text{softmax}(S_t)$$

[0068] 여기서, \hat{w}^T , \hat{w}_b , M_b , b_b 는 어텐션 가중치 \hat{w}_t^i 에 따른 프레임 특징 f_i 에 대하여 학습된 파라미터이고, \hat{s}_t 는 이상 활동과 연관된 스코어를 나타낼 수 있다. 또한, A_t 는 소프트 맥스 레이어(320)로부터 획득된 이상 활동의 확률을 나타낼 수 있다. 다시 말해, 잔차 어텐션 기반의 LSTM(310)에 의해 추출된 복수의 이미지 프레임에 대한 정보는 해당 이미지 프레임이 이상 활동 및/또는 정상 활동을 포함하는지 여부를 식별하는데 사용될 수 있으며, 소프트 맥스 레이어(320)는 잔차 어텐션 기반의 LSTM(310)에 의해 출력된 정보를 기초로 최종적인 분류 및/또는 예측을 통해 이상 활동을 인식할 수 있다. 또한, 잔차 어텐션 기반의 LSTM(310)은 Adam 최적화(optimization) 함수 및 크로스 엔트로피(cross-entropy) 손실 함수와 연관될 수 있다. 즉, Adam 최적화 함수 및 크로스 엔트로피 손실 함수가 적용됨으로써, 이상 활동 인식 성능이 극대화될 수 있다.

[0069] 도 4는 본 발명의 일 실시예에 따른 잔차 어텐션 기반의 LSTM의 특징 추출 과정(400)을 나타내는 예시적인 도면이다. 도시된 예에서, X_{t-1} , X_t , X_{t+1} , X_n 은 각 시간 단계의 새로운 입력으로서, 복수의 이미지 프레임으로부터 추출된 특징 벡터의 적어도 일부를 나타낼 수 있다. 즉, 잔차 어텐션 기반의 LSTM은 반복적으로 실행되며, 각 시간 단계에서 새로운 입력을 수신하여 처리할 수 있다. 이 경우, 각 시간 단계에서의 LSTM은 서로 연관될 수 있으며, 이전 시간 단계의 정보 중 적어도 일부가 다음 시간 단계에서 이용될 수 있다.

[0070] 일 실시예에 따르면, 잔차 어텐션 기반의 LSTM은 셀프 어텐션 레이어와 연관될 수 있다. 셀프 어텐션 레이어와 연관되는 경우, 잔차 어텐션 기반의 LSTM은 시간 단계의 길이가 길어진 경우에도 이전의 정보를 효과적으로 활용할 수 있으며, 잔차 어텐션 기반의 LSTM 및/또는 셀프 어텐션 레이어는 복수의 이미지 프레임에서 추출된 연속된 정보들 중 서로 연관성이 있는 요소들을 결정할 수 있다. 이와 같이 잔차 어텐션 기반의 LSTM에 의해 추출된 복수의 이미지 프레임의 이상 활동과 연관된 특징은 결합(concatenate) 및/또는 처리되어 소프트 맥스 레이어에 전달될 수 있다.

[0071] 또한, LSTM은 잔차 기반의 LSTM일 수 있다. 여기서, 잔차(residual)는 최상위 레이어의 순차 정보를 특성화하고 입력 레이어로 주어진 잔차 함수를 발견하여 레이어를 재구성하는데 사용될 수 있다. 예를 들어, 잔차 함수 학습은 아래의 수학적 식 4에 의해 결정될 수 있다.

수학적 식 4

[0073]
$$y = f(\vec{x}, \vec{w}) + \vec{x}$$

[0075] 여기서, \vec{x} 는 입력 데이터일 수 있으며, y 는 레이어의 결과적인 순차 정보 벡터일 수 있다. 또한, $f(\vec{x}, \vec{w})$ 는 관련된 레이어에서 학습된 잔차를 나타낼 수 있다. 이러한 잔차 학습에 의해 레이어들 사이의 효과적인 훈련을 위한 숏컷(shortcut) 함수가 생성될 수 있다. 이와 같은 구성에 의해, 잔차를 이용함으로써 학습이 진행되면서 각 파라미터에 대한 가중치가 변형되는 기울기 소실(vanishing gradient) 문제가 효과적으로 해결될 수 있다.

[0076] 도 5는 본 개시의 일 실시예에 따른 잔차 어텐션 기반의 LSTM의 이상 활동에 대한 예측 결과를 나타내는 예시적인 혼동 행렬(confusion matrix)이다. 도시된 예에서, 잔차 어텐션 기반의 LSTM은 UCF-Crime 데이터셋을 이용하여, 복수의 이미지 프레임을 기초로 이상 활동(예: 이상 활동의 발생 여부 및/또는 발생된 이상 활동의 종류 등)을 예측할 수 있다. 도시된 바와 같이, 잔차 어텐션 기반의 LSTM은 복수의 이미지 프레임을 이용하여 공격(assaults), 폭발(explosion), 다툼(fighting), 정상(normal), 도로 사고(road accident) 등을 높은 정확도로 예측할 수 있다.

[0077] 추가적으로, 잔차 어텐션 기반의 LSTM의 성능은 다른 종류의 LSTM보다 높게 측정될 수 있다. 이와 관련하여, LSTM, 양방향 LSTM, 잔차 LSTM 및 잔차 어텐션 기반의 LSTM 각각에 대한 성능이 평가되었다. 아래의 표 1과 같이 성능 평가는 UCF-Crime 데이터셋, UMN 및 Avenue 데이터셋을 이용하여 수행되었다.

표 1

[0079]

| 모델 | 데이터셋 | 리콜(%) | 정확성(%) | F1 점수(%) | AUC(%) |
|----|------|-------|--------|----------|--------|
|----|------|-------|--------|----------|--------|

| | | | | | |
|-----------------|-----------------|----|----|----|----|
| LSTM | UCF-Crime 데이터 셋 | 86 | 74 | 77 | 88 |
| BD-LSTM | | 79 | 84 | 76 | 87 |
| 잔차 LSTM | | 91 | 78 | 82 | 95 |
| 잔차 어텐션 기반의 LSTM | | 78 | 87 | 81 | 96 |
| LSTM | UMN | 87 | 77 | 81 | 86 |
| BD-LSTM | | 88 | 81 | 84 | 88 |
| 잔차 LSTM | | 94 | 95 | 94 | 96 |
| 잔차 어텐션 기반의 LSTM | | 98 | 98 | 98 | 98 |
| LSTM | Avenue | 91 | 93 | 92 | 91 |
| BD-LSTM | | 94 | 95 | 94 | 94 |
| 잔차 LSTM | | 93 | 94 | 94 | 94 |
| 잔차 어텐션 기반의 LSTM | | 98 | 99 | 99 | 98 |

[0080] 표 1와 같이, 잔차 어텐션 기반의 LSTM이 종래의 다른 LSTM보다 대부분의 지표에서 높은 성능을 나타냈다. 또한, 아래의 표 2와 같이 종래의 다른 모델과 비교하여, 잔차 어텐션 기반의 LSTM의 효율성이 증가하였다.

표 2

| 모델 | 시간 복잡성(seconds) | 모델 사이즈(MB) | 파라미터 (Millions) | FLOPs(Mega) |
|-----------------|-----------------|------------|-----------------|-------------|
| VGG-16 | - | 528 | 138 | |
| VGG-19 | - | 549 | 143 | |
| FlowNet | - | 638.5 | 162.49 | |
| DEARESt | - | 1187.5 | 305.49 | |
| 잔차 어텐션 기반의 LSTM | 0.263 | 12.8 | 3.3 | 618.3 |

[0083] 즉, 표 2와 같이, 시간 복잡성, 모델 사이즈, 파라미터, FLOPs의 모든 측면에서 잔차 어텐션 기반의 LSTM이 높은 성능을 나타냈으며, 이에 따라, 잔차 어텐션 기반의 LSTM은 더 적은 컴퓨팅 파워로도 실시간으로 이상 활동을 인식할 수 있다. 또한, 아래의 표 3과 같이, 종래의 다른 모델과 비교하여, 잔차 어텐션 기반의 LSTM의 예측 정확도가 증가하였다.

표 3

| 모델 | 정확성(%) | | |
|----------------------|-----------|-------|--------|
| | UCF-Crime | UMN | Avenue |
| VGG-16 | 72.66 | - | - |
| VGG-19 | 71.66 | - | - |
| FlowNet | 71.33 | - | - |
| DEARESt | 76.66 | - | - |
| Nandedkar and Bansod | - | 96.99 | - |
| Kyung Joo Cheoi | - | 96.50 | 90.18 |
| Al-Dhamari et al. | - | 97.44 | - |
| 잔차 어텐션 기반의 LSTM | 78.43 | 98.20 | 98.80 |

[0086] 즉, 잔차 어텐션 기반의 LSTM은 더 적은 컴퓨팅 파워를 이용하여 종래의 다른 모델보다 더 높은 정확도로 이상 활동을 인식할 수 있다.

[0087] 도 6은 본 발명의 일 실시예에 따른 LSTM 기반의 이상 인식 방법(600)의 예시를 나타내는 흐름도이다. LSTM 기반의 이상 인식 방법(600)은 프로세서(예를 들어, 이상 인식 시스템의 적어도 하나의 프로세서)에 의해 수행될 수 있다. 도시된 바와 같이, LSTM 기반의 이상 인식 방법(600)은 프로세서가 이상 인식의 판단 대상이 되는 연속된 복수의 이미지 프레임을 수신할 수 있다(S610).

- [0088] 프로세서는 수신된 복수의 이미지 프레임을 학습된 경량 CNN 모델에 제공하여, 복수의 이미지 프레임의 특징을 나타내는 특징 벡터를 생성할 수 있다(S620). 여기서, 경량 CNN은 깊이 별 구별 가능한 컨볼루션 블록을 포함할 수 있으며, 컨볼루션 블록은 3 x 3 필터와 연관된 깊이 별 컨볼루션 레이어 및 깊이 별 컨볼루션 레이어에 의해 필터링된 값을 병합하기 위한 1 x 1 필터와 연관된 포인트 별 컨볼루션 레이어를 포함할 수 있다. 또한, 경량 CNN은 입력 데이터의 채널 수를 확장하기 위한 확장 레이어를 포함할 수 있다.
- [0089] 프로세서는 생성된 특징 벡터를 잔차 어텐션 기반의 LSTM에 제공하여 복수의 이미지 프레임에 대응하는 이상 활동을 인식할 수 있다(S630). 예를 들어, 잔차 어텐션 기반의 LSTM은 훈련 시간을 감소시키기 위해 특징 벡터를 정규화하기 위한 정규화 레이어를 포함할 수 있다. 또한, 잔차 어텐션 기반의 LSTM은 셀프 어텐션 레이어(self-attention layer)와 연관될 수 있다. 여기서, 셀프 어텐션 레이어는 특징 벡터를 기초로 복수의 이미지 프레임의 연속적인 특징에 대한 상황 인식 벡터를 생성할 수 있다.
- [0090] 일 실시예에 따르면, 잔차 어텐션 기반의 LSTM은 이상 활동을 분류하기 위한 소프트 맥스 레이어와 연관될 수 있다. 이 경우, 프로세서는 생성된 특징 벡터를 잔차 어텐션 기반의 LSTM에 제공하여 출력된 출력 데이터를 소프트 맥스 레이어에 입력하여 이상 활동을 인식할 수 있다. 여기서, 잔차 어텐션 기반의 LSTM은 Adam 최적화 함수 및 크로스 엔트로피 손실 함수와 연관될 수 있다.
- [0091] 도 7은 본 발명의 일 실시예에 따른 이상 인식 시스템(700)의 내부 구성을 나타내는 블록도이다. 이상 인식 시스템(700)은 메모리(710), 프로세서(720), 통신 모듈(730) 및 입출력 인터페이스(740)를 포함할 수 있다. 도 7에 도시된 바와 같이, 이상 인식 시스템(700)은 통신 모듈(730)을 이용하여 네트워크를 통해 정보 및/또는 데이터를 통신할 수 있도록 구성될 수 있다.
- [0092] 메모리(710)는 비-일시적인 임의의 컴퓨터 판독 가능한 기록매체를 포함할 수 있다. 일 실시예에 따르면, 메모리(710)는 RAM(random access memory), ROM(read only memory), 디스크 드라이브, SSD(solid state drive), 플래시 메모리(flash memory) 등과 같은 비소멸성 대용량 저장 장치(permanent mass storage device)를 포함할 수 있다. 다른 예로서, ROM, SSD, 플래시 메모리, 디스크 드라이브 등과 같은 비소멸성 대용량 저장 장치는 메모리와의 구분되는 별도의 영구 저장 장치로서 이상 인식 시스템(700)에 포함될 수 있다. 또한, 메모리(710)에는 운영체제와 적어도 하나의 프로그램 코드가 저장될 수 있다.
- [0093] 이러한 소프트웨어 구성요소들은 메모리(710)와는 별도의 컴퓨터에서 판독 가능한 기록매체로부터 로딩될 수 있다. 이러한 별도의 컴퓨터에서 판독 가능한 기록매체는 이러한 이상 인식 시스템(700)에 직접 연결가능한 기록매체를 포함할 수 있는데, 예를 들어, 플로피 드라이브, 디스크, 테이프, DVD/CD-ROM 드라이브, 메모리 카드 등의 컴퓨터에서 판독 가능한 기록매체를 포함할 수 있다. 다른 예로서, 소프트웨어 구성요소들은 컴퓨터에서 판독 가능한 기록매체가 아닌 통신 모듈(730)을 통해 메모리(710)에 로딩될 수도 있다. 예를 들어, 적어도 하나의 프로그램은 개발자들 또는 어플리케이션의 설치 파일을 배포하는 파일 배포 시스템이 통신 모듈(730)을 통해 제공하는 파일들에 의해 설치되는 컴퓨터 프로그램에 기반하여 메모리(710)에 로딩될 수 있다.
- [0094] 프로세서(720)는 기본적인 산술, 로직 및 입출력 연산을 수행함으로써, 컴퓨터 프로그램의 명령을 처리하도록 구성될 수 있다. 명령은 메모리(710) 또는 통신 모듈(730)에 의해 사용자 단말(미도시) 또는 다른 외부 시스템으로 제공될 수 있다.
- [0095] 통신 모듈(730)은 네트워크를 통해 외부 장치 및/또는 사용자 단말(미도시)과 이상 인식 시스템(700)이 서로 통신하기 위한 구성 또는 기능을 제공할 수 있으며, 이상 인식 시스템(700)이 외부 시스템(일례로 별도의 클라우드 시스템 등)과 통신하기 위한 구성 또는 기능을 제공할 수 있다. 일례로, 이상 인식 시스템(700)의 프로세서(720)의 제어에 따라 제공되는 제어 신호, 명령, 데이터 등이 통신 모듈(730)과 네트워크를 거쳐 사용자 단말 및/또는 외부 시스템의 통신 모듈을 통해 사용자 단말 및/또는 외부 시스템으로 전송될 수 있다.
- [0096] 또한, 이상 인식 시스템(700)의 입출력 인터페이스(740)는 이상 인식 시스템(700)과 연결되거나 이상 인식 시스템(700)이 포함할 수 있는 입력 또는 출력을 위한 장치(미도시)와의 인터페이스를 위한 수단일 수 있다. 도 7에서는 입출력 인터페이스(740)가 프로세서(720)와 별도로 구성된 요소로서 도시되었으나, 이에 한정되지 않으며, 입출력 인터페이스(740)가 프로세서(720)에 포함되도록 구성될 수 있다. 이상 인식 시스템(700)은 도 7의 구성요소들보다 더 많은 구성요소들을 포함할 수 있다. 그러나, 대부분의 종래기술적 구성요소들을 명확하게 도시할 필요성은 없다.
- [0097] 이상 인식 시스템(700)의 프로세서(720)는 복수의 사용자 단말 및/또는 복수의 외부 시스템으로부터 수신된 정보 및/또는 데이터를 관리, 처리 및/또는 저장하도록 구성될 수 있다.

- [0098] 상술된 방법 및/또는 다양한 실시예들은, 디지털 전자 회로, 컴퓨터 하드웨어, 펌웨어, 소프트웨어 및/또는 이들의 조합으로 실현될 수 있다. 본 발명의 다양한 실시예들은 데이터 처리 장치, 예를 들어, 프로그래밍 가능한 하나 이상의 프로세서 및/또는 하나 이상의 컴퓨팅 장치에 의해 실행되거나, 컴퓨터 판독 가능한 기록 매체 및/또는 컴퓨터 판독 가능한 기록 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 상술된 컴퓨터 프로그램은 컴파일된 언어 또는 해석된 언어를 포함하여 임의의 형태의 프로그래밍 언어로 작성될 수 있으며, 독립 실행형 프로그램, 모듈, 서브 루틴 등의 임의의 형태로 배포될 수 있다. 컴퓨터 프로그램은 하나의 컴퓨팅 장치, 동일한 네트워크를 통해 연결된 복수의 컴퓨팅 장치 및/또는 복수의 상이한 네트워크를 통해 연결되도록 분산된 복수의 컴퓨팅 장치를 통해 배포될 수 있다.
- [0099] 상술된 방법 및/또는 다양한 실시예들은, 입력 데이터를 기초로 동작하거나 출력 데이터를 생성함으로써, 임의의 기능, 함수 등을 처리, 저장 및/또는 관리하는 하나 이상의 컴퓨터 프로그램을 실행하도록 구성된 하나 이상의 프로세서에 의해 수행될 수 있다. 예를 들어, 본 발명의 방법 및/또는 다양한 실시예는 FPGA(Field Programmable Gate Array) 또는 ASIC(Application Specific Integrated Circuit)과 같은 특수 목적 논리 회로에 의해 수행될 수 있으며, 본 발명의 방법 및/또는 실시예들을 수행하기 위한 장치 및/또는 시스템은 FPGA 또는 ASIC와 같은 특수 목적 논리 회로로서 구현될 수 있다.
- [0100] 컴퓨터 프로그램을 실행하는 하나 이상의 프로세서는, 범용 목적 또는 특수 목적의 마이크로 프로세서 및/또는 임의의 종류의 디지털 컴퓨팅 장치의 하나 이상의 프로세서를 포함할 수 있다. 프로세서는 읽기 전용 메모리, 랜덤 액세스 메모리의 각각으로부터 명령 및/또는 데이터를 수신하거나, 읽기 전용 메모리와 랜덤 액세스 메모리로부터 명령 및/또는 데이터를 수신할 수 있다. 본 발명에서, 방법 및/또는 실시예들을 수행하는 컴퓨팅 장치의 구성 요소들은 명령어들을 실행하기 위한 하나 이상의 프로세서, 명령어들 및/또는 데이터를 저장하기 위한 하나 이상의 메모리 디바이스를 포함할 수 있다.
- [0101] 일 실시예에 따르면, 컴퓨팅 장치는 데이터를 저장하기 위한 하나 이상의 대용량 저장 장치와 데이터를 주고받을 수 있다. 예를 들어, 컴퓨팅 장치는 자기 디스크(magnetic disc) 또는 광 디스크(optical disc)로부터 데이터를 수신하거나/수신하고, 자기 디스크 또는 광 디스크로 데이터를 전송할 수 있다. 컴퓨터 프로그램과 연관된 명령어들 및/또는 데이터를 저장하기에 적합한 컴퓨터 판독 가능한 저장 매체는, EPROM(Erasable Programmable Read-Only Memory), EEPROM(Electrically Erasable PROM), 플래시 메모리 장치 등의 반도체 메모리 장치를 포함하는 임의의 형태의 비 휘발성 메모리를 포함할 수 있으나, 이에 한정되지 않는다. 예를 들어, 컴퓨터 판독 가능한 저장 매체는 내부 하드 디스크 또는 이동식 디스크와 같은 자기 디스크, 광 자기 디스크, CD-ROM 및 DVD-ROM 디스크를 포함할 수 있다.
- [0102] 사용자와의 상호 작용을 제공하기 위해, 컴퓨팅 장치는 정보를 사용자에게 제공하거나 디스플레이하기 위한 디스플레이 장치(예를 들어, CRT (Cathode Ray Tube), LCD(Liquid Crystal Display) 등) 및 사용자가 컴퓨팅 장치 상에 입력 및/또는 명령 등을 제공할 수 있는 포인팅 장치(예를 들어, 키보드, 마우스, 트랙볼 등)를 포함할 수 있으나, 이에 한정되지 않는다. 즉, 컴퓨팅 장치는 사용자와의 상호 작용을 제공하기 위한 임의의 다른 종류의 장치들을 더 포함할 수 있다. 예를 들어, 컴퓨팅 장치는 사용자와의 상호 작용을 위해, 시각적 피드백, 청각 피드백 및/또는 촉각 피드백 등을 포함하는 임의의 형태의 감각 피드백을 사용자에게 제공할 수 있다. 이에 대해, 사용자는 시각, 음성, 동작 등의 다양한 제스처를 통해 컴퓨팅 장치로 입력을 제공할 수 있다.
- [0103] 본 발명에서, 다양한 실시예들은 백엔드 구성 요소(예: 데이터 서버), 미들웨어 구성 요소(예: 애플리케이션 서버) 및/또는 프론트 엔드 구성 요소를 포함하는 컴퓨팅 시스템에서 구현될 수 있다. 이 경우, 구성 요소들은 통신 네트워크와 같은 디지털 데이터 통신의 임의의 형태 또는 매체에 의해 상호 연결될 수 있다. 예를 들어, 통신 네트워크는 LAN(Local Area Network), WAN(Wide Area Network) 등을 포함할 수 있다.
- [0104] 본 명세서에서 기술된 예시적인 실시예들에 기반한 컴퓨팅 장치는, 사용자 디바이스, 사용자 인터페이스(UI) 디바이스, 사용자 단말 또는 클라이언트 디바이스를 포함하여 사용자와 상호 작용하도록 구성된 하드웨어 및/또는 소프트웨어를 사용하여 구현될 수 있다. 예를 들어, 컴퓨팅 장치는 랩톱(laptop) 컴퓨터와 같은 휴대용 컴퓨팅 장치를 포함할 수 있다. 추가적으로 또는 대안적으로, 컴퓨팅 장치는, PDA(Personal Digital Assistants), 태블릿 PC, 게임 콘솔(game console), 웨어러블 디바이스(wearable device), IoT(internet of things) 디바이스, VR(virtual reality) 디바이스, AR(augmented reality) 디바이스 등을 포함할 수 있으나, 이에 한정되지 않는다. 컴퓨팅 장치는 사용자와 상호 작용하도록 구성된 다른 유형의 장치를 더 포함할 수 있다. 또한, 컴퓨팅 장치는 이동 통신 네트워크 등의 네트워크를 통한 무선 통신에 적합한 휴대용 통신 디바이스(예를 들어, 이동 전화, 스마트 전화, 무선 셀룰러 전화 등) 등을 포함할 수 있다. 컴퓨팅 장치는, 무선 주파수(RF; Radio

Frequency), 마이크로파 주파수(MWF; Microwave Frequency) 및/또는 적외선 주파수(IRF; Infrared Ray Frequency)와 같은 무선 통신 기술들 및/또는 프로토콜들을 사용하여 네트워크 서버와 무선으로 통신하도록 구성될 수 있다.

[0105] 본 발명에서 특정 구조적 및 기능적 세부 사항을 포함하는 다양한 실시예들은 예시적인 것이다. 따라서, 본 발명의 실시예들은 상술된 것으로 한정되지 않으며, 여러 가지 다른 형태로 구현될 수 있다. 또한, 본 발명에서 사용된 용어는 일부 실시예를 설명하기 위한 것이며 실시예를 제한하는 것으로 해석되지 않는다. 예를 들어, 단수형 단어 및 상기는 문맥상 달리 명확하게 나타내지 않는 한 복수형도 포함하는 것으로 해석될 수 있다.

[0106] 본 발명에서, 달리 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함하여 본 명세서에서 사용되는 모든 용어는 이러한 개념이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 갖는다. 또한, 사전에 정의된 용어와 같이 일반적으로 사용되는 용어들은 관련 기술의 맥락에서의 의미와 일치하는 의미를 갖는 것으로 해석되어야 한다.

[0107] 본 명세서에서는 본 발명이 일부 실시예들과 관련하여 설명되었지만, 본 발명의 발명이 속하는 기술분야의 통상의 기술자가 이해할 수 있는 본 발명의 범위를 벗어나지 않는 범위에서 다양한 변형 및 변경이 이루어질 수 있다. 또한, 그러한 변형 및 변경은 본 명세서에 첨부된 특허청구의 범위 내에 속하는 것으로 생각되어야 한다.

부호의 설명

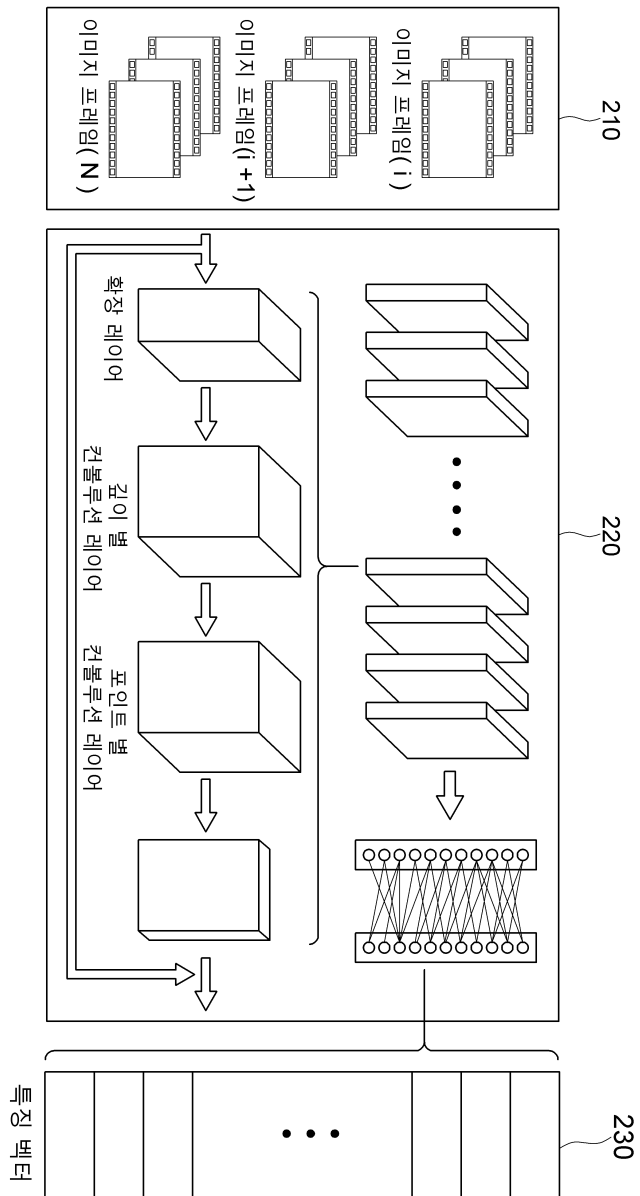
- [0109] 100: 이상 인식 시스템
- 102: 경량 CNN
- 104: 잔차 어텐션 기반의 LSTM
- 110: 이미지 프레임
- 120: 특징 벡터
- 130: 이상 활동 인식

도면

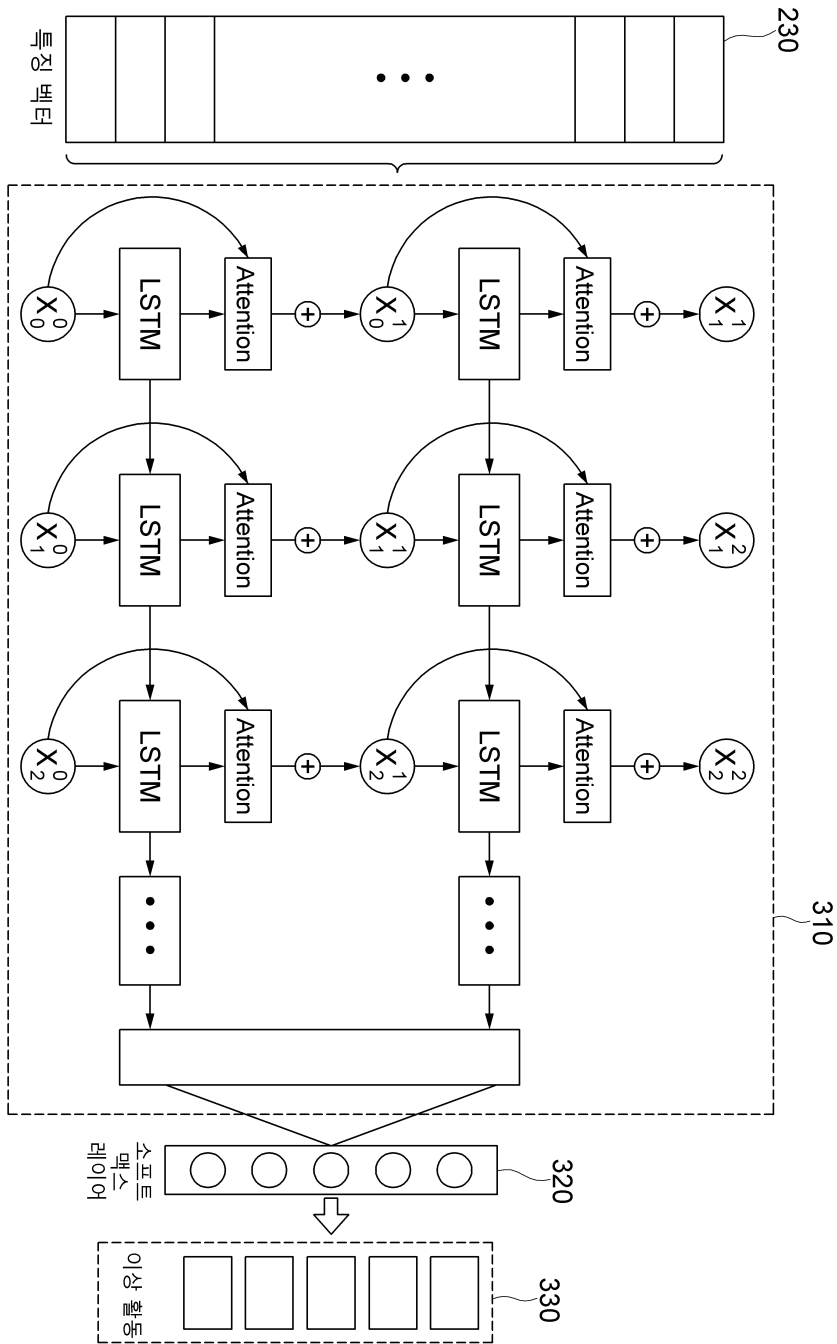
도면1



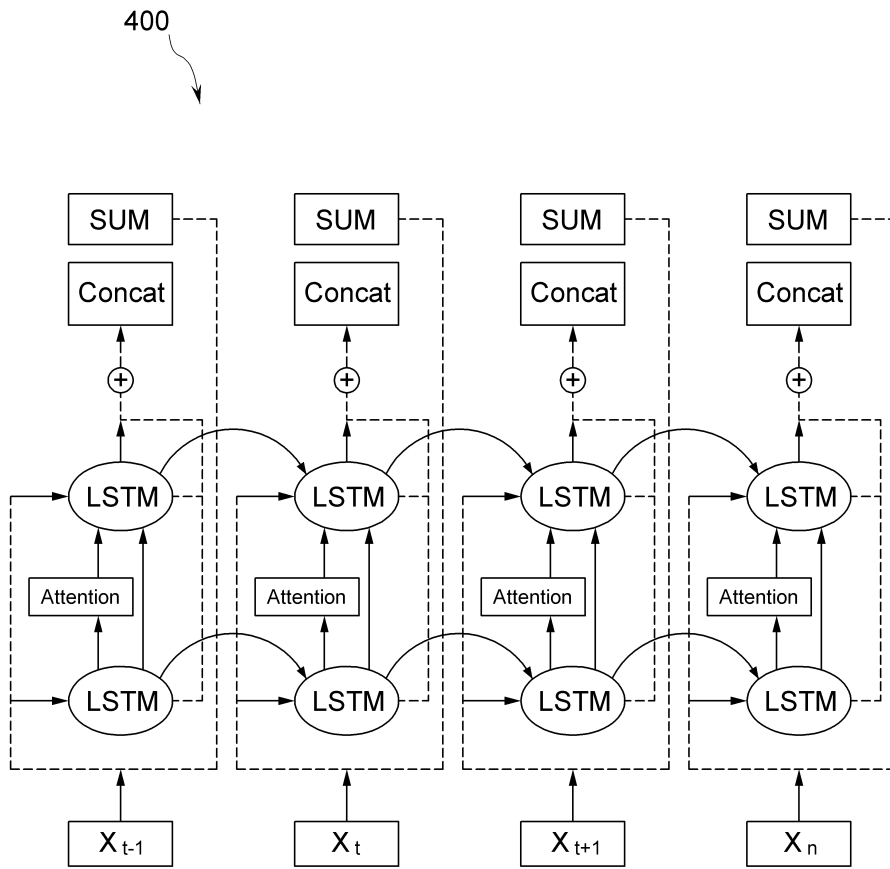
도면2



도면3



도면4

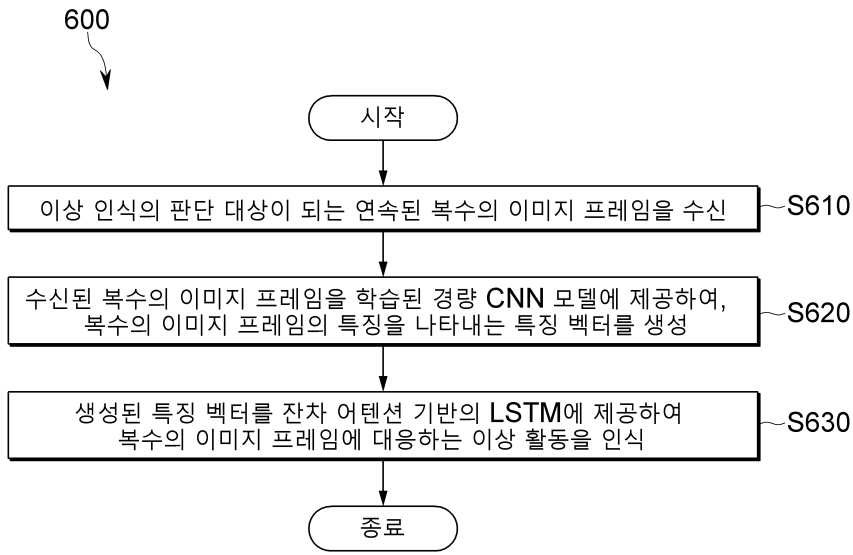


도면5

500

| | | | | | |
|-----------------------|--------------|---------------|--------------|------------|-----------------------|
| 공격(assaults) | 94.69 | 0.00 | 4.78 | 0.00 | 0.53 |
| 폭발(explosion) | 2.67 | 93.76 | 3.57 | 0.00 | 0.00 |
| 다툼(fighting) | 1.93 | 1.29 | 95.49 | 1.029 | 0.00 |
| 정상(normal) | 3.08 | 10.78 | 1.88 | 73.04 | 11.21 |
| 도로 사고(road accidents) | 9.15 | 5.09 | 12.51 | 11.19 | 60.06 |
| | 공격(assaults) | 폭발(explosion) | 다툼(fighting) | 정상(normal) | 도로 사고(road accidents) |

도면6



도면7

