(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0124699
(43) 공개일자 2024년08월19일

(51) 국제특허분류(Int. Cl.)
G08B 21/04 (2006.01) *G06V 10/764* (2022.01)
G10L 25/18 (2013.01) *H04L 67/50* (2022.01)
(52) CPC특허분류
G08B 21/0484 (2013.01)
G06V 10/764 (2023.08)
(21) 출원번호 10-2023-0017497
(22) 출원일자 2023년02월09일
심사청구일자 2023년02월09일

(71) 출원인
울산대학교 산학협력단
울산광역시 남구 대학로 93(무거동)
(72) 발명자
지영준
울산광역시 울주군 범서읍 상동길 94-8
현승호
울산광역시 북구 호계매곡1로 31 드림인시티에일린의뜰1차, 108동 3201호
(74) 대리인
특허법인태백

전체 청구항 수 : 총 8 항

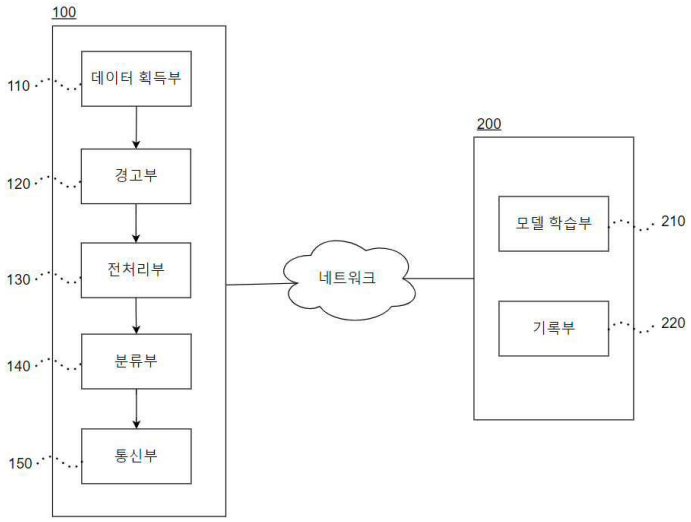
(54) 발명의 명칭 **물 사용 소리 기반 일상 활동 검출 방법 및 시스템**

(57) 요 약

본 발명은 물 사용 소리 기반 일상 활동 검출 방법 및 시스템에 관한 것으로, 본 발명에 따르면, 사용자의 거주지에 설치된 음향 센서를 통해 물 사용과 관련된 일상 소리를 녹음하여 음향 데이터를 획득하는 데이터 획득부, 상기 데이터 획득부를 통해 획득한 음향 데이터에 대한 전처리를 통해 이미지 데이터를 생성하는 전처리부 및 상기 데이터 획득부를 통해 실시간으로 획득한 음향 데이터를 기초로 상기 전처리부에 의해 생성된 이미지 데이터를 이용하여 기 학습된 SED(Sound based Event Detector) 모델을 통하여 사용자 활동을 분류하는 분류부를 포함한다.

본 발명에 따르면, 프라이버시를 침해하지 않고 사람들의 일상 생활 활동을 인식하여 비상상황일 경우 알람을 울려 적절한 조치를 취할 수 있게 할 수 있다.

대 표 도 - 도1



(52) CPC특허분류
G10L 25/18 (2013.01)
H04L 67/50 (2022.05)

이 발명을 지원한 국가연구개발사업
과제고유번호 1415183499
과제번호 P0014158
부처명 산업통상자원부
과제관리(전문)기관명 한국산업기술진흥원
연구사업명 산업기술국제협력
연구과제명 시니어 케어 서비스를 위한 포그-클라우드 기반 센서 네트워크 및 인공지능 데이터 분석을 이용한 SAFE-HOME 시스템 개발
기 여 율 1/1
과제수행기관명 (주)아이알엠
연구기간 2022.01.01 ~ 2022.12.31

명 세 서

청구범위

청구항 1

사용자의 거주지에 설치된 음향 센서를 통해 물 사용과 관련된 일상 소리를 녹음하여 음향 데이터를 획득하는 데이터 획득부;

상기 데이터 획득부를 통해 획득한 음향 데이터에 대한 전처리를 통해 이미지 데이터를 생성하는 전처리부; 및

상기 데이터 획득부를 통해 실시간으로 획득한 음향 데이터를 기초로 상기 전처리부에 의해 생성된 이미지 데이터를 이용하여 기 학습된 SED(Sound based Event Detector) 모델을 통하여 사용자 활동을 분류하는 분류부를 포함하는 물 사용 소리 기반 일상 활동 검출 시스템.

청구항 2

제 1 항에 있어서,

상기 전처리부는 상기 이미지 데이터에 대한 데이터 증강을 더 수행하는 물 사용 소리 기반 일상 활동 검출 시스템.

청구항 3

제 2 항에 있어서,

상기 전처리부는 상기 이미지 데이터에 대해 프레임 시프트(frame shift), 시간 마스크(time mask), 주파수 마스크(frequency mast) 및 믹스업(Mix-up) 중 적어도 하나의 증강 기법을 적용하여 데이터 증강을 수행하는 물 사용 소리 기반 일상 활동 검출 시스템.

청구항 4

제 1 항에 있어서,

상기 전처리부는,

멜 스펙트로그램(mel Spectrogram)를 사용하여 기 설정된 단위시간 동안의 음향 데이터를 이미지 데이터로 변환하는 물 사용 소리 기반 일상 활동 검출 시스템.

청구항 5

제 4 항에 있어서,

상기 분류부는, 상기 SED 모델을 통해 상기 이미지 데이터의 프레임 단위로 사용자 활동을 분류하고, 상기 단위시간 동안의 프레임 분류 결과의 최빈 값에 해당하는 활동으로 사용자 활동을 분류하는 물 사용 소리 기반 일상 활동 검출 시스템.

청구항 6

제 1 항에 있어서,

상기 데이터 획득부를 통해 획득한 음향 데이터가 기 설정된 조건을 만족할 경우 경고 알람을 제공하는 경고부를 더 포함하는 물 사용 소리 기반 일상 활동 검출 시스템.

청구항 7

제 1 항에 있어서,

상기 분류부에 의한 사용자 활동 분류 결과와 상기 경고부에 의한 경고 알람 제공에 대한 기록을 네트워크를 통하여 클라우드 서버로 전송하는 통신부를 더 포함하는 물 사용 소리 기반 일상 활동 검출 시스템.

청구항 8

사용자의 거주지에 설치된 음향 센서를 통해 물 사용과 관련된 일상 소리를 녹음하여 음향 데이터를 획득하는 단계;

상기 획득한 음향 데이터가 기 설정된 조건을 만족하는지 판단하는 단계;

상기 획득한 음향 데이터에 대한 전처리를 통해 이미지 데이터를 생성하는 단계; 및

상기 실시간으로 획득한 음향 데이터를 기초로 상기 전처리를 통해 생성된 이미지 데이터를 이용하여 학습 완료된 상기 SED 모델을 통하여 사용자 활동을 분류하는 단계를 포함하는 물 사용 소리 기반 일상 활동 검출 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 물 사용 소리 기반 일상 활동 검출 방법 및 시스템에 관한 것으로서, 더욱 상세하게 설명하면 대상자의 거주지에서 일상 활동에서의 물 소리를 녹음하여 학습 데이터를 확보하고, 확보한 학습 데이터를 이용하여 음향 기반 상황 인지(Sound-based Event Detector, 이하 SED) 알고리즘을 학습시켜서 물 사용 소리 기반으로 일상 활동을 검출하고 비상 상황을 감지하여 알람을 제공하는 물 사용 소리 기반 일상 활동 검출 방법 및 시스템에 관한 것이다.

배 경 기 술

[0002] 최근 급격한 고령화와 저출산 현상에 따라 가구 구조가 변화하면서 1인 가구가 보편화되고 있다. 앞으로 1인 가구는 지속적으로 늘어날 것으로 예상되는데, 1인 가구의 특성 상 다양한 주거안전에 취약하다. 이에 국가적인 대책뿐만 아니라 개인적 차원에서의 대응책 마련도 요구되고 있는 실정이다.

[0003] 특히, 고령자나 취약계층의 독신 거주자의 경우 안전과 건강에 관련된 일상 활동 모니터링에 대한 필요성이 계속적으로 증가하고 있으며, 이와 관련된 다양한 연구가 진행되고 있다.

[0004] 일 예로, 거주자의 일상 활동을 모니터링하기 위하여 거주지에 음향 센서, 움직임 감지 센서 등과 같은 다양한 센서를 설치하여 거주자의 일상 활동을 감지하는 기술이 제안된 바 있다.

[0005] 그러나 소리를 기반으로 하는 일상 활동 감지의 경우 거주지 환경의 특성이나 거주자의 행동 특성에 따라 감지 결과가 달라질 수 있어서 정확성이 떨어진다는 한계가 있다.

발명의 내용

해결하려는 과제

[0006] 본 발명에 따르면, 대상자의 거주지에서 일상 활동에서의 물 소리를 녹음하여 학습 데이터를 확보하고, 확보한 학습 데이터를 이용하여 SED 알고리즘을 학습시켜서 물 사용 소리 기반으로 일상 활동을 검출하고 비상 상황을 감지하여 알람을 제공하는 물 사용 소리 기반 일상 활동 검출 방법 및 시스템을 제공하기 위한 것이다.

과제의 해결 수단

[0007] 이러한 기술적 과제를 이루기 위한 본 발명의 실시예에 따르면 물 사용 소리 기반 일상 활동 검출 시스템에 있어서, 사용자의 거주지에 설치된 음향 센서를 통해 물 사용과 관련된 일상 소리를 녹음하여 음향 데이터를 획득하는 데이터 획득부, 상기 데이터 획득부를 통해 획득한 음향 데이터에 대한 전처리를 통해 이미지 데이터를 생성하는 전처리부 및 상기 데이터 획득부를 통해 실시간으로 획득한 음향 데이터를 기초로 상기 전처리부에 의해 생성된 이미지 데이터를 이용하여 기 학습된 SED(Sound based Event Detector) 모델을 통하여 사용자 활동을 분류하는 분류부를 포함한다.

[0008] 또한, 본 발명의 실시예에 따르면 물 사용 소리 기반 일상 활동 검출 방법에 있어서, 사용자의 거주지에 설치된 음향 센서를 통해 물 사용과 관련된 일상 소리를 녹음하여 음향 데이터를 획득하는 단계, 상기 획득한 음향 데이터가 기 설정된 조건을 만족하는지 판단하는 단계, 상기 획득한 음향 데이터에 대한 전처리를 통해 이미지 데이터를 생성하는 단계 및 상기 실시간으로 획득한 음향 데이터를 기초로 상기 전처리를 통해 생성된 이미지 데

이터를 이용하여 학습 완료된 상기 SED 모델을 통하여 사용자 활동을 분류하는 단계를 포함한다.

발명의 효과

[0009] 이와 같이 본 발명에 따르면, 프라이버시를 침해하지 않고 사람들의 일상 활동을 감지할 수 있으며, 비상 상황 발생시에 알람을 제공함으로써 적절한 조치를 취하도록 할 수 있다.

도면의 간단한 설명

[0010] 도 1은 본 발명의 일 실시예에 따른 물 사용 소리 기반 일상 활동 검출 시스템의 구성도이다.

도 2는 본 발명의 일 실시예에 따라 멜 스펙트로그램을 사용하여 데이터 전처리를 진행하는 예시도이다.

도 3은 본 발명의 일 실시예에 따른 CRNN 모델의 구성도이다.

도 4는 본 발명의 일 실시예에 따른 BiGRU(Bidirectional Gated Recurrent Unit)의 구성도이다.

도 5는 본 발명의 일 실시예에 따라 분류 결과에 대한 후처리를 진행하는 예시도이다.

도 6은 본 발명의 일 실시예에 따른 물 사용 소리 기반 일상 활동 검출 방법의 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0011] 이하 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시예를 상세히 설명하기로 한다. 이 과정에서 도면에 도시된 선들의 두께나 구성요소의 크기 등은 설명의 명료성과 편의상 과장되게 도시되어 있을 수 있다.

[0012] 또한 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서, 이는 사용자, 운용자의 의도 또는 관례에 따라 달라질 수 있다. 그러므로 이러한 용어들에 대한 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.

[0013] 도 1은 본 발명의 일 실시예에 따른 물 사용 소리 기반 일상 활동 검출 시스템의 구성도이다.

[0014] 도 1을 참조하면, 물 사용 소리 기반 일상 활동 검출 시스템(100)은 데이터 획득부(110), 전처리부(130) 및 분류부(140)를 포함하며, 경고부(120) 및 통신부(150)를 추가로 포함할 수 있다.

[0015] 또한, 클라우드 서버(200)가 네트워크를 통해 물 사용 소리 기반 일상 활동 검출 시스템(100)과 연결될 수 있으며, 클라우드 서버(200)는 모델 학습부(210)와 기록부(220)를 포함할 수 있다.

[0016] 데이터 획득부(110)는 사용자의 거주지에 설치된 음향 센서를 통해 물 사용과 관련된 일상 소리를 녹음하여 음향 데이터를 획득한다.

[0017] 여기서, 음향 센서는 예를 들어 사용자의 거주지의 화장실에 설치되어 사용자의 일상 활동에 따라 발생되는 물 사용 소리를 녹음할 수 있다.

[0018] 경고부(120)는 데이터 획득부(110)에서 획득한 음향 데이터가 기 설정된 조건(예를 들어, 일정 음량 이상)을 만족할 경우 경고 알람을 제공한다.

[0019] 전처리부(130)는 데이터 획득부(110)를 통해 획득한 음향 데이터에 대한 전처리를 통해 이미지 데이터를 생성한다.

[0020] 도 2는 본 발명의 일 실시예에 따라 멜 스펙트로그램을 사용하여 데이터 전처리를 진행하는 예시도이다.

[0021] 도 2를 참조하면, 전처리부(130)는 멜 스펙트로그램(mel Spectrogram)를 사용하여 기 설정된 단위시간(예를 들어, 1 초) 동안의 음향 데이터를 학습이 가능한 특징 벡터, 즉, 이미지 데이터로 변환할 수 있다.

[0022] 이를 위해, 먼저 음향 데이터에 대해서 STFT(Short Time Fourier Transform)를 적용하여 파워 스펙트로그램(power Spectrogram)을 획득하고, 획득한 파워 스펙트로그램에 멜 필터 뱅크(Mel Filter Bank)를 적용하여 기 설정된 해상도의 멜 스펙트로그램을 획득한다.

[0023] 이 때, STFT는 데이터에서 시간에 대해 구간을 짧게 나누어 나누어진 여러 구간의 데이터를 각각 푸리에 변환하는 방법을 의미하고, 파워 스펙트로그램은 신호나 잡음의 전력을 주파수 함수로 나타낸 스펙트로그램을 의미한다.

[0024] 또한, 멜 필터 뱅크는 각각의 프레임에 대해 얻어낸 주파수들에 대해서 Mel 값을 얻어내기 위한 필터를 의미한

다.

[0025] STFT 적용, 파워 스펙트로그램 획득 및 멜 필터 뱅크의 적용을 통해 멜 스펙트로그램을 획득하는 방법은 통상의 기술자에게 알려진 기술을 적용할 수 있는 바 이에 대한 구체적인 설명은 생략한다.

[0026] 또한, 전처리부(130)에서는 이미지 데이터에 대해 프레임 시프트(frame shift), 시간 마스크(time mask), 주파수 마스크(frequency mask) 및 믹스업(Mix-up) 중 적어도 하나의 증강 기법을 적용하여 데이터 증강을 수행한다.

[0027] 한편, 클라우드 서버(200)에 포함된 모델 학습부(210)는 통신부(150)를 통하여 전처리부(130)에서 생성된 이미지 데이터를 전송받고, 전송받은 이미지 데이터를 기반으로 SED 모델을 학습시킬 수 있다.

[0028] 도 3은 본 발명의 일 실시예에 따른 CRNN 모델의 구성도이고, 도 4는 본 발명의 일 실시예에 따른 BiGRU(Bidirectional Gated Recurrent Unit)의 구성도이다.

[0029] 도 3을 참조하면, SED 모델은 전처리부(130)에서 생성된 이미지 데이터에서 CNN을 통해 특징을 추출한 후, 이를 RNN을 통해 분류하는 방식의 모델인 CRNN모델을 이용하여 학습하도록 구성될 수 있다.

[0030] 도 4를 참조하면, 본 발명에서 사용되는 CRNN모델의 RNN은 메모리 셀(memory cell)로 BiGRU셀을 사용한다.

[0031] BiGRU셀의 상태와 각 레이어의 출력은 하기의 수학식 1에 의해 계산될 수 있다.

수학식 1

$$\mathbf{r}_t = \sigma\left(\mathbf{W}_{xr}^T \cdot \mathbf{x}_t + \mathbf{W}_{hr}^T \cdot \mathbf{h}_{t-1} + \mathbf{b}_r\right)$$
$$\mathbf{z}_t = \sigma\left(\mathbf{W}_{xz}^T \cdot \mathbf{x}_t + \mathbf{W}_{hz}^T \cdot \mathbf{h}_{t-1} + \mathbf{b}_z\right)$$
$$\mathbf{g}_t = \tanh\left(\mathbf{W}_{xg}^T \cdot \mathbf{x}_t + \mathbf{W}_{hg}^T \cdot (\mathbf{r}_t \otimes \mathbf{h}_{t-1}) + \mathbf{b}_g\right)$$
$$\mathbf{h}_t = \mathbf{z}_t \otimes \mathbf{h}_{t-1} + (1 - \mathbf{z}_t) \otimes \mathbf{g}_t$$

[0032]

[0033] 이 때, h_t는 상태 벡터를 나타내고, 하나의 게이트 제어기인 z_t가 망각(forget) 게이트와 입력(input) 게이트를 모두 제어한다. z_t가 1을 출력하면 망각 게이트가 열리고 입력 게이트가 닫히며, z_t가 0일 경우 반대로 망각 게이트가 닫히고 입력 게이트가 열린다. 즉, 이전(t-1)의 기억이 저장 될 때 마다 타임 스텝 t의 입력은 삭제된다.

[0034] BiGRU셀은 출력(output) 게이트가 없어 전체 상태 벡터 h_t가 타임 스텝마다 출력되며, 이전 상태 h_{t-1}의 어느 부분이 출력될지 제어하는 새로운 게이트 컨트롤러인 r_t를 포함한다.

[0035] 분류부(140)는 데이터 획득부(110)를 통해 실시간으로 획득한 음향 데이터를 기초로 전처리부(130)에 의해 생성된 이미지 데이터를 이용하여 기 학습된 SED 모델을 통하여 사용자 활동을 분류한다.

[0036] 구체적으로, 분류부(140)는 클라우드 서버(200) 상에서 모델 학습부(210)를 통해 학습시킨 SED 모델을 탑재하고 이를 통해 이미지 데이터의 프레임 단위로 사용자 활동을 분류하고, 기 설정된 단위시간 동안의 프레임 분류 결과의 최빈 값에 해당하는 활동으로 사용자 활동을 분류할 수 있다.

[0037] 도 5는 본 발명의 일 실시예에 따라 분류 결과에 대한 후처리를 진행하는 예시도이다.

[0038] 도 5를 참조하면, 이미지 데이터에 포함된 프레임 분류 결과의 최빈 값에 해당하는 활동인 Flushing으로 사용

자 활동을 분류할 수 있다.

[0039] 클라우드 서버(200)에 포함된 기록부(220)는 통신부(150)를 통해 분류부(140)에 의한 사용자 활동 분류 결과와 경고부(120)에 의한 경고 알람 제공에 대한 기록을 네트워크를 통하여 전송 받아 로그를 기록한다.

[0040] 도 6은 본 발명의 일 실시예에 따른 물 사용 소리 기반 일상 활동 검출 방법의 흐름도이다.

[0041] 도 6을 참조하면, 물 사용 소리 기반 일상 활동 검출 방법은 사용자의 거주지에 설치된 음향 센서를 통해 물 사용과 관련된 일상 소리를 녹음하여 음향 데이터를 획득하는 단계(S210), 상기 음향 데이터를 획득하는 단계(S210)를 통해 획득한 음향 데이터가 기 설정된 조건을 만족하는지 판단하는 단계(S220), 상기 획득한 음향 데이터에 대한 전처리를 통해 이미지 데이터를 생성하는 단계(S230) 및 상기 음향 데이터를 획득하는 단계(S210)를 통해 실시간으로 획득한 음향 데이터를 기초로 상기 전처리를 통해 이미지 데이터를 생성하는 단계(S230)에 의해 생성된 이미지 데이터를 이용하여 기 학습된 SED 모델을 통하여 사용자 활동을 분류하는 단계(S240)를 포함하여 구성될 수 있다.

[0042] 또한, 각 단계에서 수행되는 구체적인 동작은 도 1 내지 도 5를 참조하여 상술한 바와 동일하므로 이에 대한 중복적인 설명은 생략한다.

[0043] 본 발명은 도면에 도시된 실시예를 참고로 하여 설명되었으나 이는 예시적인 것에 불과하며, 당해 기술이 속하는 분야에서 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 따라서 본 발명의 진정한 기술적 보호범위는 아래의 특허청구범위의 기술적 사상에 의하여 정해져야 할 것이다.

부호의 설명

[0044] 100: 물 사용 소리 기반 일상 활동 검출 시스템

110: 데이터 획득부

120: 경고부

130: 전처리부

140: 분류부

150: 통신부

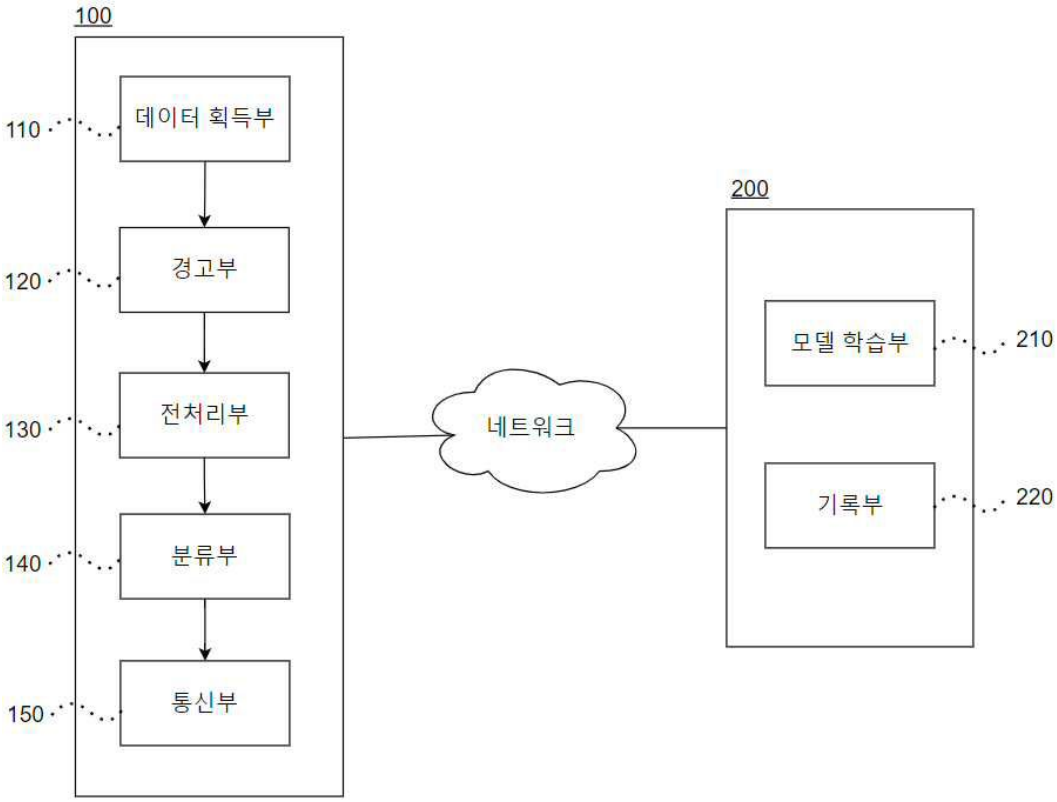
200: 클라우드 서버

210: 모델 학습부

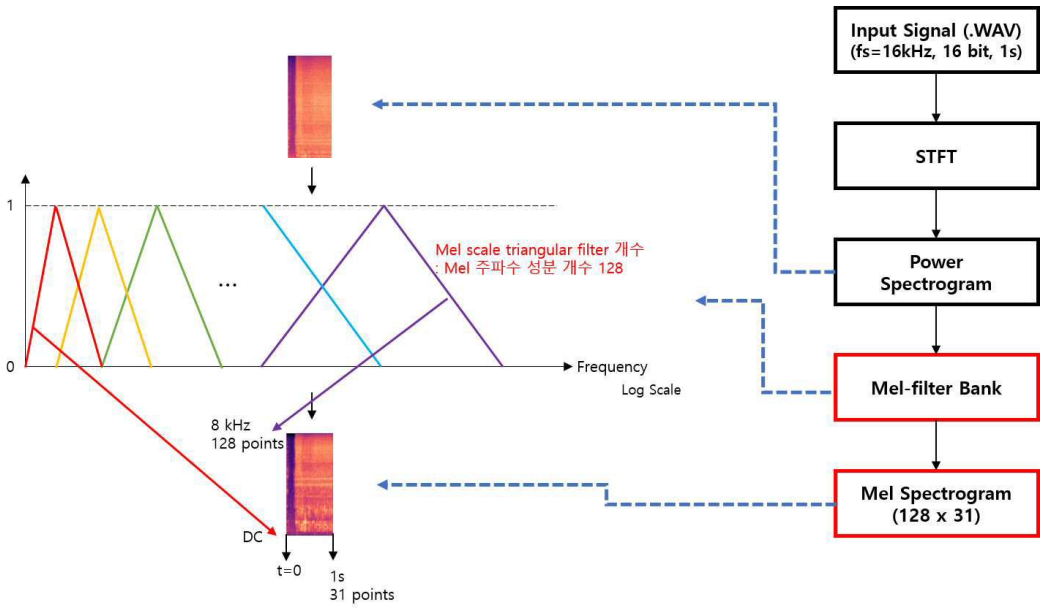
220: 기록부

도면

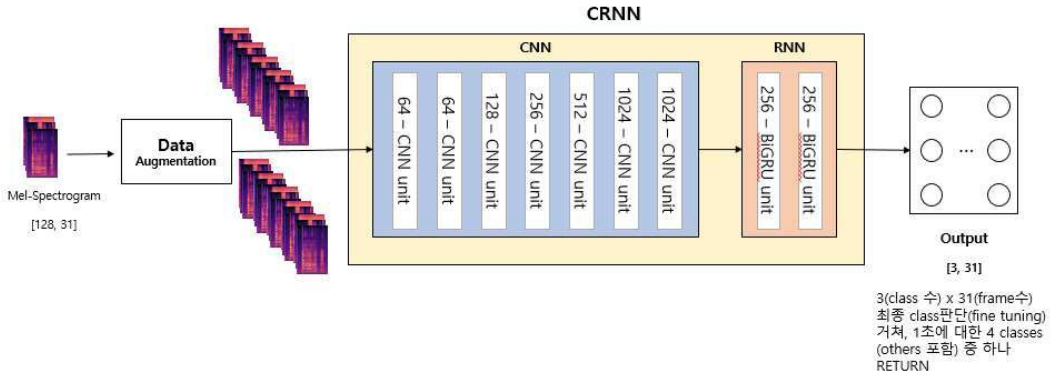
도면1


100
110
데이터 획득부
120
경고부
130
전처리부
140
분류부
150
통신부
네트워크
200
모델 학습부
210
기록부
220

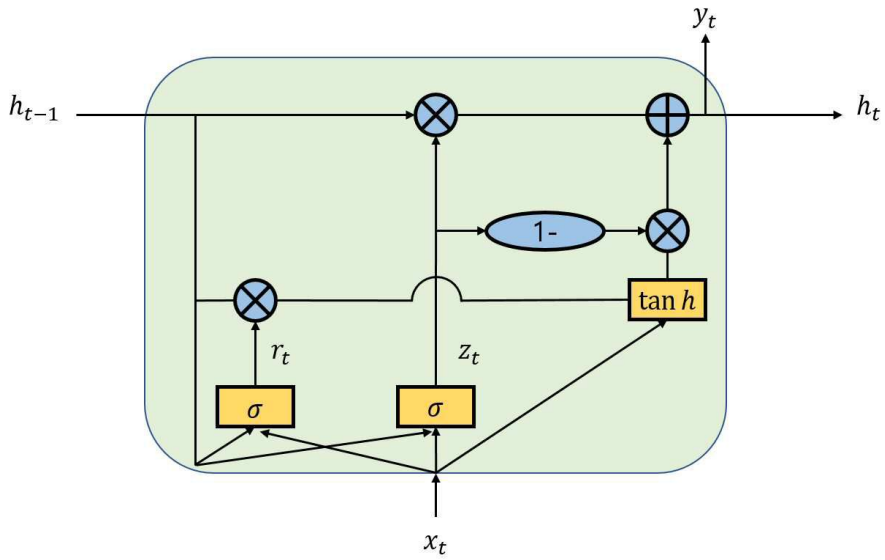
도면2


Input Signal (.WAV)
(fs=16kHz, 16 bit, 1s)
STFT
Power Spectrogram
Mel-filter Bank
Mel Spectrogram (128 x 31)
Mel scale triangular filter 개수
: Mel 주파수 성분 개수 128
1
0
Frequency
Log Scale
8 kHz
128 points
DC
t=0
1s
31 points

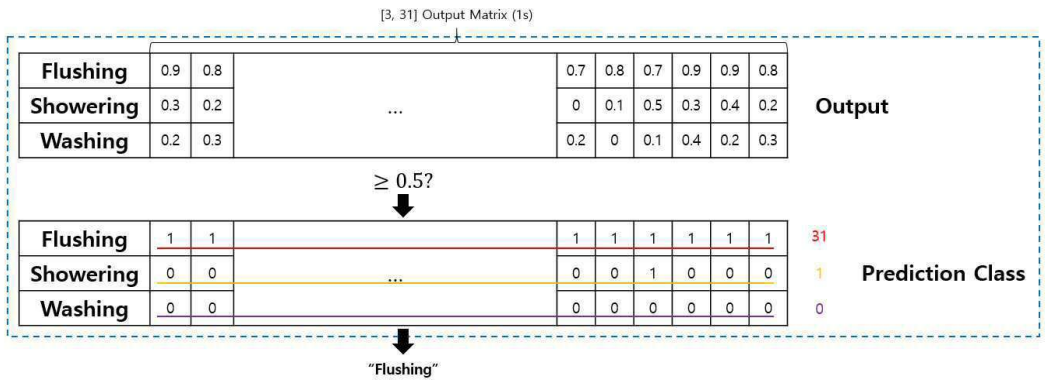
도면3


CRNN
CNN
RNN
Mel-Spectrogram
[128, 31]
Data Augmentation
64 – CNN unit
64 – CNN unit
128 – CNN unit
256 – CNN unit
512 – CNN unit
1024 – CNN unit
1024 – CNN unit
256 – BiGRU unit
256 – BiGRU unit
Output
[3, 31]
3(class 수) x 31(frame수)
최종 class판단(fine tuning)
거쳐, 1초에 대한 4 classes
(others 포함) 중 하나
RETURN

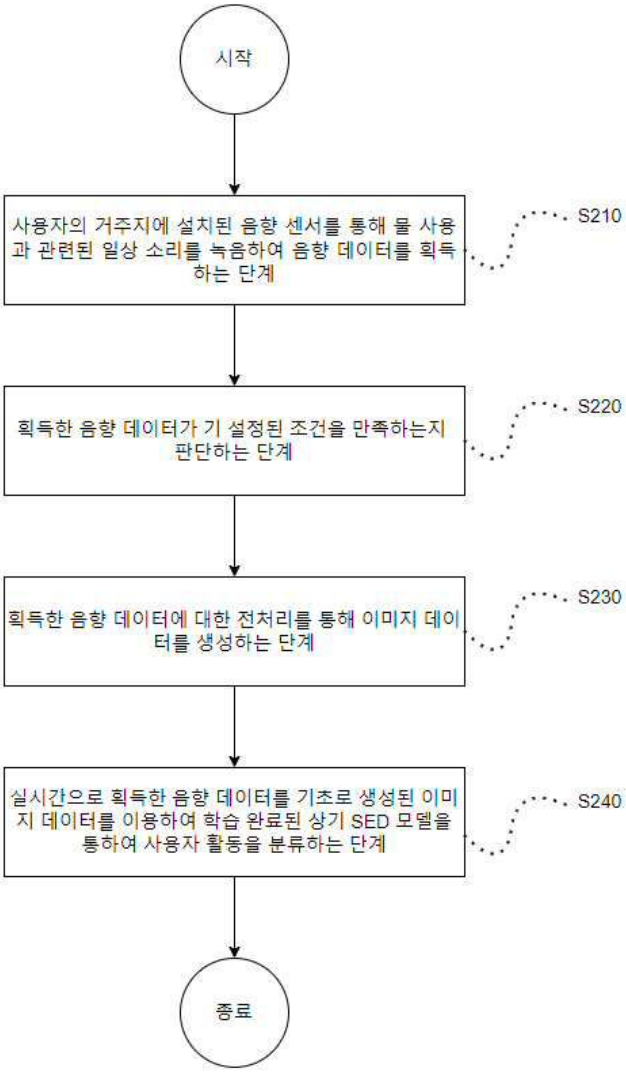
도면4


y_t
h_{t-1}
h_t
1-
tan h
r_t
z_t
σ
σ
x_t

도면5


[3, 31] Output Matrix (1s)
Flushing 0.9 0.8 ... 0.7 0.8 0.7 0.9 0.9 0.8
Showering 0.3 0.2 ... 0 0.1 0.5 0.3 0.4 0.2
Washing 0.2 0.3 ... 0.2 0 0.1 0.4 0.2 0.3
Output
≥ 0.5?
Flushing 1 1 ... 1 1 1 1 1 1 31
Showering 0 0 ... 0 0 1 0 0 0 1
Washing 0 0 ... 0 0 0 0 0 0 0
Prediction Class
"Flushing"

도면6


시작
사용자의 거주지에 설치된 음향 센서를 통해 물 사용과 관련된 일상 소리를 녹음하여 음향 데이터를 획득하는 단계
S210
획득한 음향 데이터가 기 설정된 조건을 만족하는지 판단하는 단계
S220
획득한 음향 데이터에 대한 전처리를 통해 이미지 데이터를 생성하는 단계
S230
실시간으로 획득한 음향 데이터를 기초로 생성된 이미지 데이터를 이용하여 학습 완료된 상기 SED 모델을 통하여 사용자 활동을 분류하는 단계
S240
종료