



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년11월04일
(11) 등록번호 10-2463143
(24) 등록일자 2022년11월01일

(51) 국제특허분류(Int. Cl.)
G06V 20/59 (2022.01) G06V 10/44 (2022.01)
G06V 10/82 (2022.01) G06V 20/70 (2022.01)
G06V 40/16 (2022.01)
(52) CPC특허분류
G06V 20/597 (2022.01)
G06V 10/454 (2022.01)
(21) 출원번호 10-2020-0132467
(22) 출원일자 2020년10월14일
심사청구일자 2020년10월14일
(65) 공개번호 10-2021-0155319
(43) 공개일자 2021년12월22일
(30) 우선권주장
1020200072252 2020년06월15일 대한민국(KR)
(56) 선행기술조사문헌
KR1020180119258 A
(뒷면에 계속)
전체 청구항 수 : 총 8 항

(73) 특허권자
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
김용국
경기도 성남시 분당구 동판교로 156, 911동 503호(삼평동, 봇들마을9단지금호어울림아파트)
(74) 대리인
송인호, 최관탁

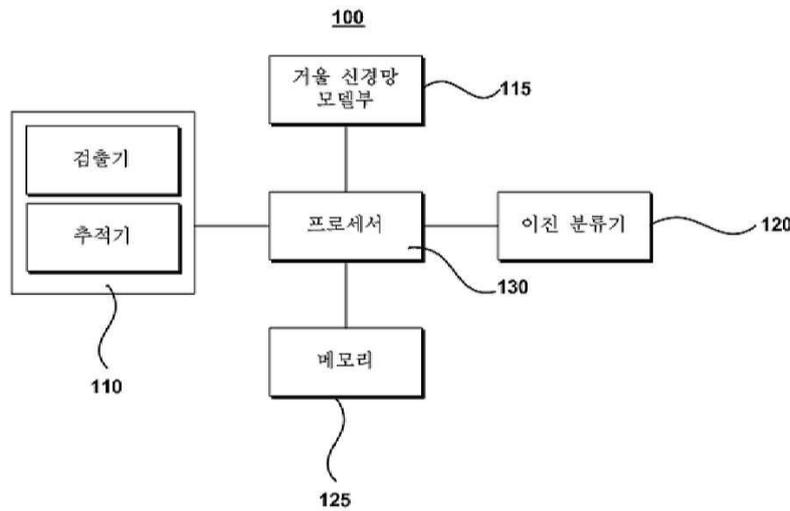
심사관 : 하은주

(54) 발명의 명칭 졸음 운전 감지 방법 및 그 장치

(57) 요약

졸음 운전 감지 방법 및 그 장치가 개시된다. 졸음 운전 감지 장치는, 입력 영상에서 얼굴 영역을 검출하는 검출부; 상기 얼굴 영역을 분석한 후 얼굴 특징값을 PB 벡터로서 출력하는 거울 신경망 모델부; 및 상기 PB 벡터를 스파이크(spike)로 변환한 후 분류 모델을 통해 졸음 상태 및 정상 상태 중 어느 하나로 분류하는 이진 분류기를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06V 10/82 (2022.01)
 G06V 20/70 (2022.01)
 G06V 40/169 (2022.01)

(56) 선행기술조사문헌

KR1020200095383 A
 KR101775874 B1
 KR1020190100982 A
 KR1020200003572 A
 JP6446126 B2
 KR102096617 B1

이 발명을 지원한 국가연구개발사업

과제고유번호	1711090675
과제번호	2017-0-00731-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	방송통신산업기술개발(R&D)
연구과제명	딥러닝 기반 시청자 주의 및 감정상태 판별을 이용한 맞춤형 광고 플랫폼 연구 개발
기여율	1/1
과제수행기관명	세종대학교 산학협력단
연구기간	2019.01.01 ~ 2019.12.31

공지예외적용 : 있음

명세서

청구범위

청구항 1

입력 영상에서 얼굴 영역을 검출하는 검출부;

상기 얼굴 영역을 분석한 후 얼굴 특징값을 PB 벡터로서 출력하는 거울 신경망 모델부; 및

상기 PB 벡터를 스파이크(spike)로 변환한 후 STDP(spike-timing-dependent plasticity) 기반 분류 모델을 통해 졸음 상태 및 정상 상태 중 어느 하나로 분류하는 이진 분류기를 포함하되,

상기 PB 벡터는 0보다 큰 경우 상기 STDP 기반 분류 모델에 입력 가능한 형태인 스파이크 뉴런으로 변환되되,

상기 STDP 기반 분류 모델은 복수의 뉴런을 가지되, 상기 변환된 스파이크 뉴런에 의해 복수의 뉴런의 활성화 상태를 결정하여 졸음 상태 및 정상 상태 중 어느 하나로 분류하는 것을 특징으로 하는 졸음 운전 감지 장치.

청구항 2

제1 항에 있어서,

상기 거울 신경망 모델부는,

상기 얼굴 영역을 복수의 콘볼루션 층을 통해 분석하여 특징값을 추출하는 인코더;

상기 인코더의 출력인 상기 특징값을 PB 노드에 누적하는 2D 그리드 장단기 메모리;

상기 특징값을 이용하여 얼굴 영역을 재구성하는 디코더; 및

상기 얼굴 영역과 상기 재구성된 얼굴 영역간의 에러를 저장하는 에러 PB 노드를 포함하는 것을 특징으로 하는 졸음 운전 감지 장치.

청구항 3

제1 항에 있어서,

상기 거울 신경망 모델부는

모든 시간 스텝(time step)의 PB 노드에 대한 역전과 오류를 누적하며,

상기 PB 벡터는 하기 수학적식에 의해 갱신되는 것을 특징으로 하는 졸음 운전 감지 장치.

$$p_{x_i} = p_{x_i}^{old} + \delta p_{x_i}$$

여기서, $p_{x_i}^{old}$ 는 갱신 이전의 PB 벡터의 값을 나타내며,

$$\delta p_{x_i} = \frac{1}{l_i} \sum_{t=0}^{l_i-1} error_{p_{x_i}}(t).$$

는 모든 시간 스텝을 통해 누적된 역전과 오류값의 평균값을 나타내며, $error_{p_{x_i}}(t)$ 는 t 시간 스텝에서의 역전과 오류를 나타냄.

청구항 4

제2 항에 있어서,

상기 2D 그리드 장단기 메모리는 상기 얼굴 영역에 상응하는 졸음 또는 정상 상태를 컨텍스트 정보로서 입력받는 것을 특징으로 하는 졸음 운전 감지 장치.

청구항 5

삭제

청구항 6

제1 항에 있어서,

상기 PB 벡터를 스파이크 누적으로 변환하되, 상기 STDP 기반 분류 모델에 입력 가능한 포아송 분산 스파이크 트레이너로 표현하는 것을 특징으로 하는 졸음 운전 감지 장치.

청구항 7

삭제

청구항 8

졸음 운전 감지 장치에서 수행되는 졸음 운전 감지 방법에 있어서,

- (a) 입력 영상에서 얼굴 영역을 검출하는 단계;
- (b) 상기 얼굴 영역을 거울 신경망 모델에 적용하여 얼굴 특징값을 추출한 후 PB 벡터를 출력하는 단계; 및
- (c) 상기 PB 벡터를 스파이크(spike)로 변환한 후 STDP(spike-timing-dependent plasticity) 기반 분류 모델을 통해 졸음 상태 및 정상 상태 중 어느 하나로 분류하는 단계를 포함하되,

상기 (c) 단계는,

상기 PB 벡터는 0보다 큰 경우 상기 STDP 기반 분류 모델에 입력 가능한 형태인 스파이크 뉴런으로 변환되되, 상기 STDP 기반 분류 모델은 복수의 뉴런을 가지되, 상기 변환된 스파이크 뉴런에 의해 복수의 뉴런의 활성화 상태를 결정하여 졸음 상태 및 정상 상태 중 어느 하나로 분류하는 것을 특징으로 하는 졸음 운전 감지 방법.

청구항 9

제8 항에 있어서,

상기 거울 신경망 모델부는,

RNNPB 기반 구조에 2D 그리드 장단기 메모리가 결합되며, 원시 영상이 상기 얼굴 영역을 입력받아 특징값을 추출하기 위해 인코더가 상기 RNNPB의 입력단에 결합되고,

상기 2D 그리드 장단기메모리의 출력은 상기 RNNPB의 히든계층에 완전하게 연결되어 상기 특징값을 PB 노드에 누적하는 것을 특징으로 하는 졸음 운전 감지 방법.

청구항 10

제8 항에 따른 방법을 수행하기 위해 필요한 프로그램 코드를 기록한 컴퓨터로 판독 가능한 매체에 저장된 컴퓨터 프로그램.

발명의 설명

기술 분야

[0001] 본 발명은 거울 신경망 및 STDP에 기반한 졸음 운전 감지 방법 및 그 장치에 관한 것이다.

배경 기술

[0003] 현재 교통 사고로 인한 사망자의 30% 이상이 졸음 운전에 기인한다. 2008년 미국 고속도로 교통 안전국에서 발행된 차량 충돌에 대한 경찰 보고서에 따르면, 10만건이 운전자 졸음의 직접적인 결과로 1550명이 사망하고, 71000명의 부상 및 125억 달러의 금전적 손실을 초래한 것으로 추정하고 있다.

[0004] 이러한 이유로 주요 자동차 제조업체는 졸음 운전을 방지할 수 있는 다양한 시스템을 개발하고 있다. 그러나, 종래의 졸음 운전 방지는 얼굴을 특징점을 분리하는 방식으로 얼굴 전체 영상을 입력으로 사용하는 제한으로 인해, 운전 중 운전자 얼굴 전체 영상을 정확하게 획득하기 어려운 문제로 인해 실제 적용에는 한계가 있다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 거울 신경망 및 STDP에 기반한 졸음 운전 감지 방법 및 그 장치를 제공하기 위한 것이다.

[0007] 또한, 본 발명은 운전자의 다양한 포즈에서도 졸음 운전을 정확하게 인식할 수 있는 졸음 운전 감지 방법 및 그 장치를 제공하기 위한 것이다.

과제의 해결 수단

[0009] 본 발명의 일 측면에 따르면, 거울 신경망 및 STDP에 기반한 졸음 운전 감지 장치가 제공된다.

[0010] 본 발명의 일 실시예에 따르면, 입력 영상에서 얼굴 영역을 검출하는 검출부; 상기 얼굴 영역을 분석한 후 얼굴 특징값을 PB 벡터로서 출력하는 거울 신경망 모델부; 및 상기 PB 벡터를 스파이크(spike)로 변환한 후 분류 모델을 통해 졸음 상태 및 정상 상태 중 어느 하나로 분류하는 이진 분류하는 이진 분류기를 포함하는 졸음 운전 감지 장치가 제공될 수 있다.

[0011] 상기 거울 신경망 모델부는, 상기 얼굴 영역을 복수의 콘볼루션 층을 통해 분석하여 특징값을 추출하는 인코더; 상기 인코더의 출력인 상기 특징값을 PB 노드에 누적하는 2D 그리드 장단기 메모리; 상기 특징값을 이용하여 얼굴 영역을 재구성하는 디코더; 및 상기 얼굴 영역과 상기 재구성된 얼굴 영역간의 에러를 저장하는 에러 PB 노드를 포함할 수 있다.

[0012] 상기 거울 신경망 모델부는 모든 시간 스텝(time step)의 PB 노드에 대한 역전파 오류를 누적하며, 상기 PB 벡터는 하기 수학식에 의해 갱신되며,

$$p_{x_i} = p_{x_i}^{old} + \delta p_{x_i}$$

[0013]

$$\delta p_{x_i} = \frac{1}{l_i} \sum_{t=0}^{l_i-1} error_{p_{x_i}}(t).$$

[0014] 여기서, $p_{x_i}^{old}$ 는 갱신 이전의 PB 벡터의 값을 나타내며, δp_{x_i} 는 모든 시간 스텝을 통해 누적된 역전파 오류값의 평균값을 나타내며, $error_{p_{x_i}}(t)$ 는 t 시간 스텝에서의 역전파 오류를 나타낸다.

[0015] 상기 2D 그리드 장단기 메모리는 상기 얼굴 영역에 상응하는 졸음 또는 정상 상태를 컨텍스트 정보로서 입력 받을 수 있다.

[0016] 상기 분류 모델은 STDP 기반 구조로 형성될 수 있다.

[0017] PB 벡터 리스트를 스파이크 누적으로 변환하고, 포아송 분산 스파이크 트레인으로 표현하여 상기 분류 모델에

적용할 수 있다.

- [0018] 상기 PB 벡터 리스트를 스파이크 누적으로 변환함에 있어, PB 벡터의 값이 0보다 큰 경우 스파이크된 뉴런으로 변환될 수 있다.
- [0020] 본 발명의 다른 측면에 따르면, 거울 신경망 및 STDP에 기반한 졸음 운전 감지 방법이 제공된다.
- [0021] 본 발명의 일 실시예에 따르면, 입력 영상에서 얼굴 영역을 검출하는 단계; 상기 얼굴 영역을 거울 신경망 모델에 적용하여 얼굴 특징값을 추출한 후 PB 벡터를 출력하는 단계; 상기 PB 벡터를 스파이크(spike)로 변환한 후 분류 모델을 통해 졸음 상태 및 정상 상태 중 어느 하나로 분류하는 이진 분류하는 단계를 포함하는 졸음 운전 감지 방법이 제공될 수 있다.
- [0022] 상기 거울 신경망 모델부는, RNNPB 기반 구조에 2D 그리드 장단기 메모리가 결합되며, 원시 영상이 상기 얼굴 영역을 입력받아 특징값을 추출하기 위해 인코더가 상기 RNNPB의 입력단에 결합되고, 상기 2D 그리드 장단기 메모리의 출력은 상기 RNNPB의 히든계층에 완전하게 연결되어 상기 특징값을 PB 노드에 누적할 수 있다.

발명의 효과

- [0024] 본 발명의 일 실시예에 따른 졸음 운전 감지 방법 및 그 장치를 제공함으로써, 거울 신경망 및 STDP를 기반으로 정확하게 졸음 운전을 인식할 수 있는 이점이 있다.
- [0025] 또한, 본 발명은 운전자의 다양한 포즈에서도 졸음 운전을 정확하게 인식할 수 있는 이점도 있다.

도면의 간단한 설명

- [0027] 도 1은 본 발명의 일 실시예에 따른 졸음 운전 인식 장치의 내부 구성을 개략적으로 도시한 블록도.
- 도 2는 본 발명의 일 실시예에 따른 거울 신경망 모델부의 세부 구조를 도시한 도면.
- 도 3은 본 발명의 일 실시예에 따른 학습 과정에 대한 의사 코드를 도시한 도면.
- 도 4는 본 발명의 일 실시예에 따른 인식 과정에 대한 의사 코드를 도시한 도면.
- 도 5는 오토인코더에 대한 그래픽 확률 모델을 설명하기 위해 도시한 도면.
- 도 6은 본 발명의 일 실시예에 따른 인코더와 디코더의 상세 구조를 도시한 도면.
- 도 7은 본 발명의 일 실시예에 따른 분류 모델의 세부 구조를 도시한 도면.
- 도 8 내지 도 10은 종래와 본 발명의 일 실시예에 따른 졸음 운전 감지 방법에 따른 졸음 분별 정확도를 비교한 결과를 나타낸 도면.
- 도 11은 본 발명의 일 실시예에 따른 졸음 운전 감지 방법을 나타낸 순서도.

발명을 실시하기 위한 구체적인 내용

- [0028] 본 명세서에서 사용되는 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "구성된다" 또는 "포함한다" 등의 용어는 명세서상에 기재된 여러 구성 요소들, 또는 여러 단계들을 반드시 모두 포함하는 것으로 해석되지 않아야 하며, 그 중 일부 구성 요소들 또는 일부 단계들은 포함되지 않을 수도 있고, 또는 추가적인 구성 요소 또는 단계들을 더 포함할 수 있는 것으로 해석되어야 한다. 또한, 명세서에 기재된 "...부", "모듈" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어 또는 소프트웨어로 구현되거나 하드웨어와 소프트웨어의 결합으로 구현될 수 있다.
- [0029] 이하, 첨부된 도면들을 참조하여 본 발명의 실시예를 상세히 설명한다.
- [0031] 도 1은 본 발명의 일 실시예에 따른 졸음 운전 감지 장치의 내부 구성을 개략적으로 도시한 블록도이며, 도 2는 본 발명의 일 실시예에 따른 거울 신경망 모델부의 세부 구조를 도시한 도면이고, 도 3은 본 발명의 일 실시예에 따른 학습 과정에 대한 의사 코드를 도시한 도면이고, 도 4는 본 발명의 일 실시예에 따른 인식 과정에 대한 의사 코드를 도시한 도면이며, 도 5는 오토인코더에 대한 그래픽 확률 모델을 설명하기 위해 도시한 도면이고, 도 6은 본 발명의 일 실시예에 따른 인코더와 디코더의 상세 구조를 도시한 도면이며, 도 7은 본 발명의 일 실시예에 따른 분류 모델의 세부 구조를 도시한 도면이고, 도 8 내지 도 10은 종래와 본 발명의 일 실시예에 따른 졸음 운전 감지 방법에 따른 졸음 분별 정확도를 비교한 결과를 나타낸 도면이다.

- [0032] 도 1을 참조하면, 본 발명의 일 실시예에 따른 줄음 운전 감지 장치(100)는 검출부(110), 거울 신경망 모델부(115), 이진 분류기(120), 메모리(125) 및 프로세서(130)를 포함하여 구성된다.
- [0033] 검출부(110)는 각각의 프레임에서 얼굴 영역을 검출하기 위한 수단이다.
- [0034] 검출부(110)는 입력 영상을 분석하여 얼굴 영역을 검출한 후 이를 추적할 수 있다. 따라서, 검출부(110)는 검출기와 추적기를 포함하여 구성될 수 있다. 따라서, 검출기에 의해 얼굴 영역이 검출되면, 추적기를 이용하여 얼굴이 추적될 수 있다.
- [0035] 예를 들어, 본 발명의 일 실시예에서는 KK 추적기(KK tracker)가 포함되는 것을 가정하기로 한다. 이러한 KK 추적기는 칼만 필터와 결합된 KCF(Kernelized Correlation Filter)가 사용된다.
- [0036] 이와 같이, 검출부(110)가 검출기와 추적기를 이용하여 구성됨으로써, 검출기에 의해 얼굴 영역이 검출된 이후, 이를 기초로 얼굴을 추적하도록 하여 선글라스 등을 이용하여 얼굴 일부 영역이 가려지거나 정면이 아닌 측면 등 다른 형상으로 얼굴 영역이 입력되는 경우에도 정확하게 얼굴 영역을 검출할 수 있는 이점이 있다.
- [0037] 거울 신경망 모델부(115)는 검출된 얼굴 영상을 분석하여 특징값을 추출하고, 추출된 특징값을 PB 벡터로서 출력하기 위한 수단이다.
- [0038] 거울 신경망 모델부(115)의 상세 구조는 도 2에 도시된 바와 같다.
- [0039] 도 2를 참조하여 거울 신경망 모델부(115)의 구조에 대해 설명하기로 한다.
- [0040] 거울 신경망 모델부(115)는 RNNPB 기반 구조에 2D 그리드 장단기메모리가 결합된 구조이며, 원시 영상인 얼굴 영역을 처리하기 위해 인코더가 RNNPB의 입력단으로 결합된 구조이다.
- [0041] RNNPB 아키텍처는 Jordan 유형과 유사하다. 즉, RNNPB는 재귀 피드 포워드 신경망 기반 구조와 유사하지만 입력 계층에 PB 노드가 포함되는 점에서 차이가 있다. PB 노드는 각 시간 시퀀스에서 특정 상수 벡터를 가져와 고정된 길이의 벡터와 시간 시퀀스를 매핑할 수 있다. 따라서, PB 노드는 자체 구성 프로세스를 통해 시간 시퀀스를 인코딩할 수 있다.
- [0042] 거울 신경망 모델부(115)는 인코더를 포함한다. 인코더는 얼굴 랜드 마크가 아닌 원시 얼굴 영상을 분석하여 특징값을 추출한다.
- [0043] 예를 들어, 인코더는 Rest34 기반 구조로 형성될 수 있다. 인코더에 의해 추출된 특징값은 장단기 메모리를 통해 PB 노드에 누적될 수 있다.
- [0044] PB 벡터의 학습은 BPTT(Back-Propagation Through Time) 알고리즘을 변형하여 수행될 수 있다.
- [0045] BPTT 알고리즘은 PB 계층의 고유값을 조정하고 나가는 연결의 가중치를 고정하여 PB 벡터를 갱신할 수 있다. PB 노드와 관련된 역전과 오류는 모든 시간 동안 누적될 수 있다.
- [0046] i번째 학습 시간 시퀀스(x_i)가 인코딩된 PB 벡터(p_{x_i})는 수학식 1과 수학식 2에 따라 갱신될 수 있다.

수학식 1

$$\delta p_{x_i} = \frac{1}{l_i} \sum_{t=0}^{l_i-1} error_{p_{x_i}}(t)$$

- [0047]
- [0048] 여기서, δp_{x_i} 는 모든 PB 노드에 대한 평균 역전과 오류값을 나타내고, t는 현재 시간 스텝을 나타내며, $error_{p_{x_i}}(t)$ 는 현재 시간 스텝에서의 에러값을 나타내고, l_i 는 이포크(epoch)로 학습시 100으로 사용될 수 있다.

수학식 2

$$p_{x_i} = p_{x_i}^{old} + \delta p_{x_i}$$

[0049]

여기서, $p_{x_i}^{old}$ 는 갱신되기 이전의 PB 벡터값을 나타낸다.

[0050]

본 발명의 일 실시예에 따른 거울 신경망 모델부(115)는 PB 벡터에 상응하는 시간 시퀀스를 생성한다. 생성 프로세서는 적절한 PB 벡터, 고정된 초기 컨텍스트 벡터 및 입력 벡터를 포함하는 2D 그리드 장단기메모리가 이용될 수 있다. 얼굴 이미지는 입력 벡터를 위해 외부 정보로서 활용될 수 있다.

[0052]

본 발명의 일 실시예에 따른 거울 신경망 모델부(115)는 시퀀스 생성 및 인식 과정에 모두 사용될 수 있다. 수학식 1에 의해 연결 가중치가 고정되면 대응하는 PB 벡터는 주어진 시퀀스에 대해 획득될 수 있다. 학습 과정은 PB 공간의 훈련 시간 시퀀스들 사이의 관계 구조와 필수적인 특징들이 추출될 수 있다. 거울 신경망 모델부(115)는 거울 신경(mirror neurons)의 속성으로 인해, 동적 시스템 접근 방식을 대표할 수 있다.

[0053]

본 발명의 일 실시예에 따른 거울 신경망 모델부(115)는 학습 모드, 생성 모드, 인식 모드와 같은 세가지 운영 모드를 가지고 있다. 거울 신경망 모델부(115)는 학습 데이터 셋을 이용하여 줄음 또는 정상 상태에 대한 라벨(y)에 매핑되어 학습될 수 있다. 또한, 학습이 완료된 거울 신경망 모델부(115)에 얼굴 영상이 적용되어 얼굴 특징값을 추출한 후 PB 벡터에 누적하여 출력할 수 있다.

[0054]

본 발명의 일 실시예에 따른 거울 신경망 모델부(115)의 학습 과정에 대한 의사 코드는 도 3에 도시된 바와 같으며, 인식 과정에 대한 의사 코드는 도4에 도시된 바와 같다.

[0055]

인코더h의 출력은 잠재 변수(latent variable)이다. 잠재 변수는 조인트 분산(joint distribution), $P(x, c) = P(x|c)P(c)$,을 통해 입력(x)에 관련될 수 있다.

[0056]

도 5의 (A)에서 보여지는 바와 같이, 잠재 변수에 대한 분산($P(c)$)과 잠재 변수를 입력에 관련시키는 분산($P(x|c)$)은 분산($P(x)$)의 모양을 결정할 수 있다.

[0057]

잠재 변수와 레이블(y) 사이의 연관은 도 5의 (B)에 도시된 바와 같이, 거울 신경망 모델부(115)의 파라메트릭 바이어스를 통해 조사될 수 있다.

[0058]

잠재 분산($P(c)$)는 수학식 3과 같이 나타낼 수 있다.

수학식 3

$$P(c) = \sum_y P(c|PB = y)P(PB = y)$$

[0059]

따라서, 조인트 분산은 수학식 4와 같이 나타낼 수 있다.

수학식 4

$$P(x, c) = P(x|c) \sum_y P(c|PB = y)P(PB = y)$$

[0061]

- [0062] 파라메트릭 바이어스 분산($P(PB)$)은 데이터베이스내의 라벨(y)(졸음 또는 정상)의 분산에 의해 결정될 수 있다. 수학적 식 4는 거울 신경망 모델부의 훈련 및 학습 과정 동안 사용될 수 있다.
- [0063] 거울 신경망 모델부(115)에 포함되는 인코더와 디코더의 상세 구조는 도 6에 도시된 바와 같다. 인코더는 얼굴 영상을 분석하여 특징맵(m)을 추출한다. 즉, 인코더는 얼굴 영상을 다수의 컨볼루션 레이어에 적용하여 특징맵(m)을 추출할 수 있다. Resnet34 인코더 구조는 도 6에 도시된 바와 같다. Resnet의 기능은 당업자에게는 자명한 사항이므로 이에 대한 별도의 설명은 생략하기로 한다.
- [0064] 디코더는 특징맵을 이용하여 얼굴 이미지를 재구성하기 위한 수단이다. 디코더의 세부 구조 또한 도 6에 도시된 바와 같다. 디코더 층에서의 분해 계층 구조를 구축함으로써, 거울 신경망 학습을 위한 풍부한 특징이 학습될 수 있다. 거울 신경망 모델부(115)는 인코더 함수($h = f(x)$)와 디코더 함수($r = g(h)$)를 포함한다.
- [0065] 도 2에 도시된 바와 같이, 인코더로 입력되는 원시 영상인 얼굴 영역과 디코더에 의해 재구성되는 얼굴 영역 사이의 오차가 역전파되어 PB 노드가 학습될 수 있다. 모든 시간 스텝에 대한 정보들을 저장하기 위해 2D 그리드 장단기메모리가 활용될 수 있다. 또한, 2D 그리드 장단기메모리의 출력은 RNNPB의 히든 계층에 완전히 연결될 수 있다.
- [0066] 거울 신경망 모델부(115)는 0.001의 초기 학습과 크기 16인 미니 배치를 사용하여 인코더, 2D 그리드 장단기 메모리 및 PB 노드의 가중치를 갱신하였다. 본 발명의 일 실시예에 따르면 거울 신경망 모델부(115)의 2D 그리드 장단기 메모리와 PB 노드는 1000회 반복 한 후에도 검증 손실이 감소하지 않는 경우 학습률을 0.9 팩터(factor) 감소시켰다. 100번 반복을 포함하는 100에포크 이후에 학습이 중단되었다.
- [0067] 이진 분류기(120)는 분류 모델을 포함하며 PB 벡터를 스파이크(spike)로 변환하고, 변환된 스파이크를 분류 모델에 적용하여 졸음 상태 또는 정상 상태로 이진 분류할 수 있다.
- [0068] 본 발명의 일 실시예에 따르면, 분류 모델은 STDP(spike-timing-dependent plasticity)에 기반한다. 도 7에는 분류 모델의 상세 구조가 도시되어 있다. 이에 대해 간략하게 설명하기로 한다. 통합 및 재모델링은 뉴런 활동 연구의 필수적 측면이 되었으며, 하위 임계값 막위 전위 역할을 모델링하는 더 간단한 작업만 남기는 모델 뉴런 역할에 매우 유용하다.
- [0069] 시냅스 입력을 포함하여 모든 활성막 전도도를 무시하면 모델을 수동 또는 LIG 모델이라고 한다. 본 발명의 일 실시예에서는 수학적 식 3에 따라 멤브레인 볼트 V를 사용했다.

수학적 식 5

[0070]
$$\tau \frac{dV}{dt} = (E_{rest} - V) C g_e (E_{exc} - V) C g_i (E_{inh} - V)$$

[0071] C는 커패시턴스(capacitance)를 나타낸다.

[0072] 여기서, E_{rest} 는 휴지막 전위(resting membrane potential)을 나타내고, E_{exc} 와 E_{inh} 는 각각 흥분성 및 억제성 시냅스의 평형 전위를 나타낸다. 또한, g_e 와 g_i 는 각각 흥분성 및 억제성 시냅스의 전도도를 나타낸다. 이론적으로 시간 상수는 억제성 뉴런보다 흥분성 뉴런에 대해 더 길다. 시냅스 전 활동 전위 후 시간 t에서 시냅스 전도도는 기하급수적으로 감소한다. 전시냅스 뉴런(presynaptic neuron)은 활동 전위를 발산하므로, 전도도(g_e)의 역할은 수학적 식 6과 같이 나타낼 수 있다.

수학식 6

$$\tau_{g_e} \frac{dg_e}{dt} = -g_e$$

[0073]

[0074]

[0075]

[0076]

[0077]

[0078]

[0079]

[0080]

[0081]

[0082]

[0083]

[0084]

[0085]

[0087]

여기서, τ_{g_e} 는 흥분성 시냅스 후 전위(excitatory post-synaptic potential)를 나타낸다. 수학식 6에서 시냅스 전 뉴런이 억제성으로 주어지는 경우, 갱신을 위해 τ_{g_e} 가 사용된다.

분류 모델은 입력 계층과 처리 계층과 같이 두개의 계층으로 구성된다. 입력 계층은 PB 특징 개수에 적정하도록 64개의 뉴런으로 구성될 수 있다. 처리 계층의 주요 특징은 가변적인 뉴런인 흥분성 뉴런과 억제성 뉴런이다.

우선, 각 입력은 포아송 스파이크 형태로 흥분성 뉴런에 공급된다. 그 후 흥분성 뉴런은 일대일 방식으로 억제성 뉴런에 연결되며, 흥분성 뉴런의 각 스파이크는 해당 억제성 뉴런에 상응하는 스파이크를 활성화한다. 마지막으로, 각 억제성 뉴런은 일대 다 방식으로 흥분성 뉴런에 연결될 수 있다. 또한, 측면 여기가 제공될 수 있다. STDP 학습 원리는 입력 뉴런에서 흥분성 뉴런까지 모든 시냅스에 사용되었다.

메모리(125)는 본 발명의 일 실시예에 따른 거울 신경망 및 STDP에 기반한 줄음 운전 감지 방법을 수행하기 위한 적어도 하나의 명령어(프로그램 코드)를 저장하기 위한 수단이다.

프로세서(130)는 본 발명의 일 실시예에 따른 줄음 운전 감지 장치(100)의 내부 구성 요소들(예를 들어, 검출부(110), 거울 신경망 모델부(115), 이진 분류기(120), 메모리(125) 등)을 제어하기 위한 수단이다.

도 8은 종래와 본 발명의 일 실시예에 따른 줄음 운전 감지 방법에 따른 줄음 분별 정확도를 비교한 결과를 도시한 도면이고, 도 9는 종래와 본 발명의 일 실시예에 따른 줄음 운전 감지 방법에 따른 줄음 분별 정확도를 비교한 표이며, 도 10은 종래와 본 발명의 일 실시예에 따른 줄음 운전 감지 방법에 따른 줄음 분별 정확도를 비교한 그래프이다. 도 8에서 노란색은 정상 상태를 나타내고, 보라색은 줄음을 나타낸다. 본 발명의 일 실시예에 따른 방법이 줄음 운전 상태를 명확하게 분리할 수 있음을 보여준다. 도 9 및 도 10에서 보여지는 바와 같이, 본 발명의 일 실시예에 따른 방법이 종래에 비해 85.73% 정확도를 보여주는 것을 알 수 있다. 도 11은 본 발명의 일 실시예에 따른 줄음 운전 감지 방법을 나타낸 순서도이다. 단계 1110에서 줄음 운전 감지 장치(100)는 입력 영상에서 얼굴 영역을 검출한다.

단계 1115에서 줄음 운전 감지 장치(100)는 원시 영상 형태인 얼굴 영역을 학습된 거울 신경망 모델에 적용하여 얼굴 특징값을 추출하고, 이를 PB 벡터에 누적하여 출력한다.

거울 신경망 모델은, RNNPB 기반 구조에 2D 그리드 장단기 메모리가 결합되며, 원시 영상인 얼굴 영역을 처리하기 위해 입력단에 인코더가 결합된 구조이다. 또한, 2D 그리드 장단기메모리의 출력은 RNNPB의 히든계층에 완전하게 연결되어 특징값을 PB 노드에 누적하도록 할 수 있다.

또한, 거울 신경망 모델은 BPTT를 기반으로 모든 시간 스텝(time step)의 PB 노드에 대한 역전파 오류를 누적하여 PB 벡터가 갱신될 수 있다.

단계 1120에서 줄음 운전 감지 장치(100)는 PB 벡터를 스파이크로 변환한다.

단계 1125에서 줄음 운전 감지 장치(100)는 변환된 스파이크를 분류 모델에 적용하여 줄음 상태 또는 정상 상태로 이진 분류한다. 여기서, 분류 모델은 STDP 기반 구조로 형성될 수 있다.

PB 벡터 리스트를 스파이크로 누적 변환하고, 포아송 분산 스파이크 트레인으로 표현하여 분류 모델에 적용할 수 있다. 또한, PB 벡터 리스트를 스파이크 누적으로 변환함에 있어, PB 벡터의 값이 0보다 큰 경우 스파이크된 뉴런으로 변환되어 분류 모델을 통해 줄음 상태 또는 정상 상태로 이진 분류될 수 있다.

본 발명의 실시 예에 따른 장치 및 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 컴퓨터 판독 가능 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 분야 통상의 기술자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프

와 같은 자기매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플로포티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media) 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0088] 상술한 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

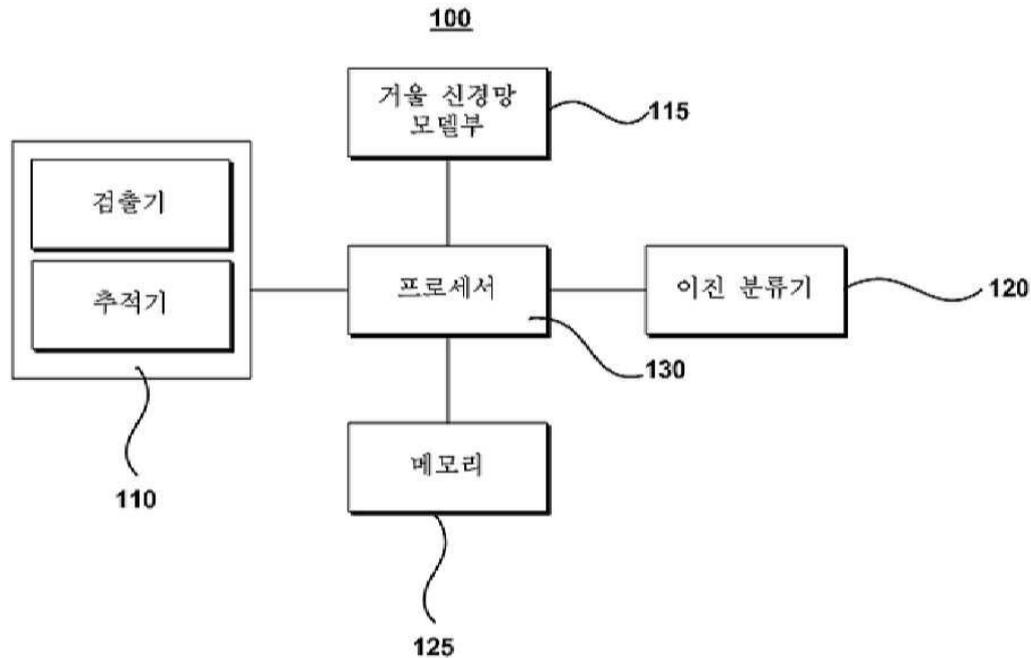
[0089] 이제까지 본 발명에 대하여 그 실시 예들을 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시 예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

부호의 설명

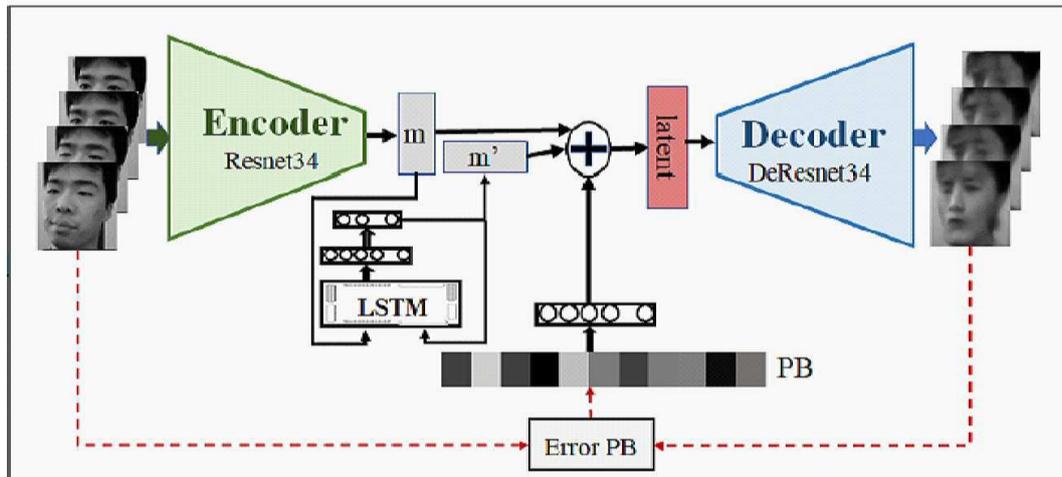
- [0091] 100: 줄음 운전 감지 장치
- 110: 검출부
- 115: 거울 신경망 모델부
- 120: 이진 분류기
- 125: 메모리
- 130: 프로세서

도면

도면1



도면2

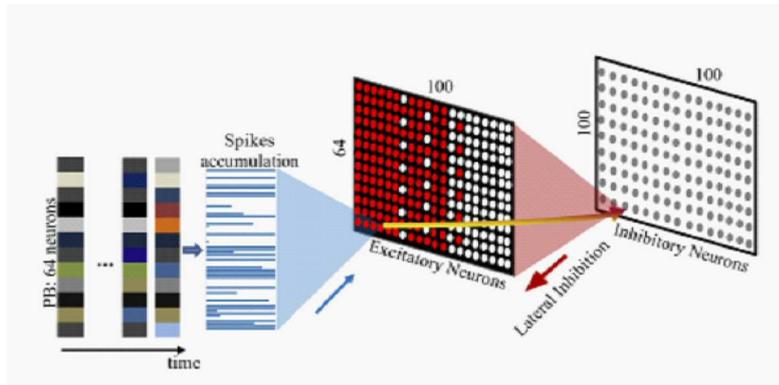


도면3

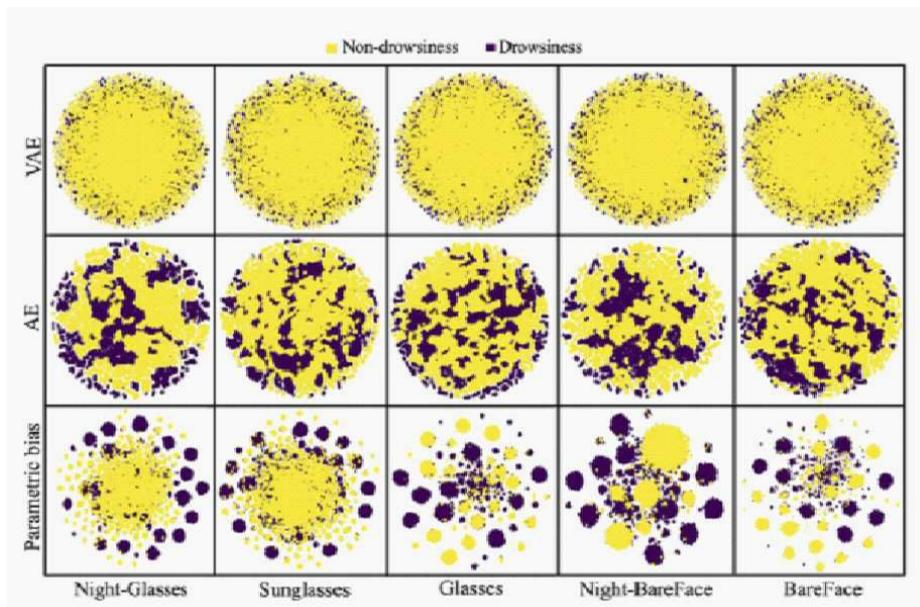
Algorithm 1: Training phase of AE-PB

Input: Image sequences of training data
Output: PB
Network: AE-PB
labels= [Non-drowsy, Drowsy]
foreach *c* **in** labels **do**
| $PB[c] \leftarrow [0 \dots 0]$
end
Divide Training data to Train_Set and Val_Set
for $epoch=0$; $epoch < 100$; $epoch=epoch+1$ **do**
| **for** $iteration=0$; $iteration < 100$; $iteration=iteration+1$ **do**
| | Select image sequence from Train_Set with label *c* randomly
| | output = AE-PB(image sequence, PB[*c*])
| | Error PB = MSE(output, image sequence)
| | Gradients propagated back from AE-PB with Error PB
| | Update the weights of AE-PB
| | Get the gradient of $PB[c]$ as Equation 1
| | Update $PB[c]$ as Equation 2
| **end**
| val_loss=0
| **foreach** *Image sequence* **in** Val_Set **do**
| | output = AE-PB(Image sequence, PB[*c*])
| | Error PB = MSE(output, Image sequence)
| | val_loss = val_loss + Error PB
| **end**
| **if** *val_loss was not decreased after each 10 epochs* **then**
| | Decrease learning rate
| **end**
end

도면7



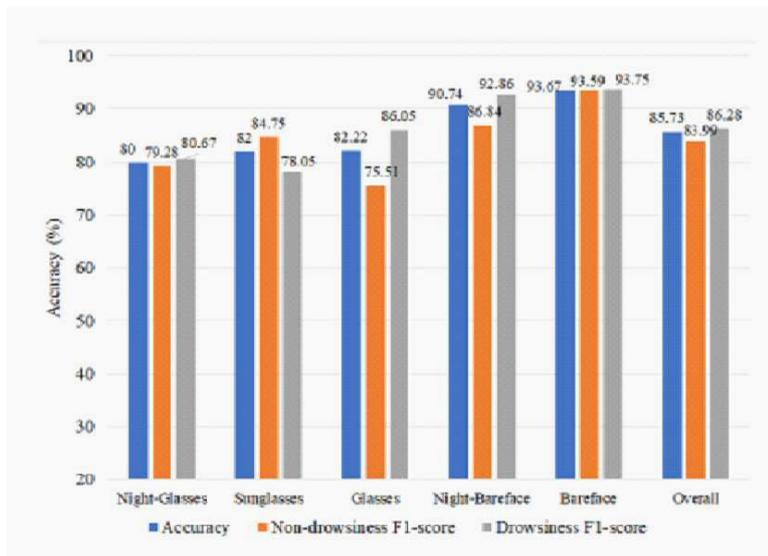
도면8



도면9

Authors	Drowsiness Model	Accuracy (%)
Park et al. (Park et al., 2017)	VGG-FaceNet	73.06
Yao et al. (Yao et al., 2018)	3D CNN	78.48
Reta et al. (Reta Cárcamo, 2018)	PCA, Fisher, SVM	79.84
Celona et al. (Celona et al., 2018)	Multi-task CNN	83.44
Weng et al. (Weng et al., 2017)	CNN, Markov models	84.82
Shih et al. (Shih & Hsu, 2017)	CNN and LSTM	85.52
Ours	AE-PB and STDP	85.73

도면10



도면11

