(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0080714
(43) 공개일자 2024년06월07일

(51) 국제특허분류(Int. Cl.)
A61B 5/16 (2006.01) *A61B 5/00* (2021.01)
G06N 3/0442 (2023.01) *G06N 3/0464* (2023.01)
G16H 50/20 (2018.01) *G16H 50/50* (2018.01)
(52) CPC특허분류
A61B 5/165 (2013.01)
A61B 5/4803 (2013.01)
(21) 출원번호 10-2022-0164300
(22) 출원일자 2022년11월30일
심사청구일자 없음

(71) 출원인
조선대학교산학협력단
광주광역시 동구 조선대길 126(서석동, 산학협력단)
(72) 발명자
곽근창
광주광역시 서구 풍암2로 66, 202동 703호(풍암동, 금호타운)
조아현
전라남도 순천시 백연길 105, 201동 405호(조례동, 연동2차대주파크빌)
(74) 대리인
특허법인지원

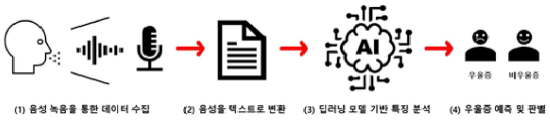
전체 청구항 수 : 총 3 항

(54) 발명의 명칭 **딥러닝기술 기반 음성 및 텍스트를 이용한 우울증 판별 시스템**

(57) 요 약

본 발명은 일상생활에서 쉽게 취득할 수 있는 음성 및 텍스트를 이용하여 딥러닝 기술을 기반으로 우울증을 판별하는 시스템에 관한 것으로, 음성과 텍스트와 같은 언어적인 데이터에서는 우울증 환자들의 특징이 잘 드러나기 때문에 이러한 특징을 분석하여 우울증을 예측 및 판별하는 방법에 대한 것이다.

대 표 도 - 도1



(52) CPC특허분류
A61B 5/7275 (2013.01)
G06N 3/0442 (2023.01)
G06N 3/0464 (2023.01)
G16H 50/20 (2018.01)
G16H 50/50 (2018.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1345347532
과제번호 2017R1A6A1A03015496
부처명 교육부
과제관리(전문)기관명 한국연구재단
연구사업명 이공학학술연구기반구축
연구과제명 IT연구소
기 여 율 1/2
과제수행기관명 조선대학교 산학협력단
연구기간 2022.03.01 ~ 2023.02.28
이 발명을 지원한 국가연구개발사업
과제고유번호 1711171054
과제번호 S0254-22-1001
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신산업진흥원
연구사업명 인공지능중심산업융합집적단지조성(R&D)
연구과제명 AI 기반 다중센싱을 이용한 Brain-Body 인터페이스 기술 개발
기 여 율 1/2
과제수행기관명 조선대학교 산학협력단
연구기간 2022.01.01 ~ 2022.12.31

명 세 서

청구범위

청구항 1

우울증 환자들의 특징이 잘 드러나는 음성 및 텍스트와 같은 언어적인 데이터의 특징을 분석하고 이를 바탕으로 우울증의 여부를 판별하는 방법.

청구항 2

분류 모델로서 음성의 1차원 2차원 특징, 텍스트의 특징을 각각 양방향장단기기억 네트워크와 합성곱신경망을 통해 분류한 뒤 각 소프트맥스 스코어를 마지막 단에서 결합하는 방법.

청구항 3

단어를 예측하고, 오타를 발견할 수 있고, 단어의 순서를 무시하는 방법의 단점을 보완할 수 있는 텍스트를 이용한 n-gram 기반 합성곱신경망 설계 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 일상생활에서 쉽게 취득할 수 있는 음성 및 텍스트를 이용하여 딥러닝 기술을 기반으로 우울증을 판별하는 시스템에 관한 것으로, 음성과 텍스트와 같은 언어적인 데이터에서는 우울증 환자들의 특징이 잘 드러나기 때문에 이러한 특징을 분석하여 우울증을 예측 및 판별하는 방법에 대한 것이다.

배 경 기 술

[0002] 우울증은 모든 사람에게 발생할 수 있는 흔한 정신질환으로 이를 다양한 데이터 및 특징을 고려해 객관적으로 판별할 방법이 필요하다. 따라서 본 발명을 통한 시스템은 사용자의 음성을 녹음해 수집하고 이를 텍스트로 변환한 뒤, 이러한 언어적인 데이터에서 우울증 환자들의 특징들을 분석하고 딥러닝 기술을 기반으로 우울증을 예측 및 판별할 수 있어 우울증을 포함한 여러 정신 장애에 대한 증상을 정확하게 감지하고 객관화하는 데 유용하게 사용될 것으로 기대된다.

선행기술문헌

특허문헌

[0003] (특허문헌 0001) KR등록 10-2041848

발명의 내용

해결하려는 과제

[0004] 현재 우울증을 진단하고 치료하기 위해서는 환자가 병원에 방문하여 직접 증상을 말하거나, 상담을 통한 의사의 주관적인 판단, 우울증 진단 관련 설문지를 통한 진단하는 등 다양한 방법을 사용한다. 하지만, 이러한 방법은 객관적인 의견이 배제되어 제한적이고 정확한 진단이 어렵다는 단점이 있다. 따라서 우울증을 초기 단계에서 진단하고 적절한 시기에 조치하기 위해서 적은 비용으로 쉽게 취득할 수 있고, 우울증 환자들의 특성이 잘 드러나는 음성 및 텍스트와 같은 언어적인 데이터의 다양한 특징 및 정보를 딥러닝 모델을 기반으로 분석함으로써 우울증의 여부를 보다 객관적이고 정확하게 판별 가능할 것으로 기대된다.

과제의 해결 수단

[0005] 본 발명의 일 실시예에 따른 우울증의 여부를 판별하는 방법은 우울증 환자들의 특징이 잘 드러나는 음성 및 텍스트와 같은 언어적인 데이터의 특징을 분석하고 이를 바탕으로 하는 것일 수 있다.

[0006] 본 발명은 분류 모델로서 음성의 1차원 2차원 특징, 텍스트의 특징을 각각 양방향장단기기억 네트워크와 합성곱신경망을 통해 분류한 뒤 각 소프트맥스 스코어를 마지막 단에서 결합하는 방법을 포함할 수 있다.

[0007] 본 발명은 단어를 예측하고, 오타를 발견할 수 있고, 단어의 순서를 무시하는 방법의 단점을 보완할 수 있는 텍스트를 이용한 n-gram 기반 합성곱신경망 설계 방법을 포함할 수 있다.

발명의 효과

[0008] 본 발명의 효과는 우울증을 보다 정확하고 객관적으로 예측 및 판별하기 위해 한 개의 데이터만으로 분석하지 않고, 음성 및 텍스트를 이용하여 우울증의 특징을 분석한다. 우울증은 복합적인 정신질환이기 때문에 다양한 데이터 및 특징 정보를 고려하여 진단하는 것이 중요하다. 특히 음성 또는 텍스트와 같은 언어적인 데이터에서는 음역대 및 속도 감소, 자주 사용하는 단어 등 우울증 환자들의 특징들이 잘 드러난다. 따라서 녹음을 통해 음성 데이터를 수집한 뒤, 텍스트 데이터로 변환하여 이 두 가지의 데이터를 이용하여 우울증의 여부를 판별한다. 구체적으로 음성의 1차원 스펙트럼 특징을 이영한 양방향장단기기억 네트워크와 시간-주파수 변환 기반 2차원 스펙트로그램 특징을 이용한 합성곱신경망 기반 전이학습 네트워크 분류 모델을 설계한다. 또한 텍스트 데이터를 토큰화 및 전처리과정을 거친 후 각각 양방향장단기기억 네트워크와 n-gram 기반 합성곱신경망 분류 모델을 설계한다. 최종적으로 앞선 4개의 네트워크를 결합함으로써 보다 다양한 정보와 특징을 사용하여 객관적이고 정확하게 우울증을 판별할 수 있을 것으로 기대된다.

도면의 간단한 설명

[0009] 도 1은 음성 및 텍스트 특징을 분석 후 우울증을 판별하는 시스템을 도시한 것이다.

도 2는 음성의 1차원 스펙트럼 특징 및 2차원 스펙트로그램 특징을 이용한 양방향 장단기 기억과 합성곱 신경망 기반 전이학습 분류 모델을 도시한 것이다.

도 3는 텍스트 특징을 이용한 양방향장단기기억과 n-gram기반 합성곱신경망 분류 모델을 도시한 것이다.

도 4는 각 분류 모델의 소프트맥스 스코어를 융합한 최종적으로 우울증을 판별하는 흐름도를 도시한 것이다.

도 5는 n-gram 언어 모델의 개념을 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

[0010] 본 발명은 도 1과 같이 음성과 텍스트에서 확인할 수 있는 우울증 환자들의 특징을 딥러닝 모델을 통해 분석하여 우울증의 여부를 판별하는 시스템이다. 우울증은 복합적인 정신질환이기 때문에 하나의 데이터를 이용해 특정 측면에서 분석하는 것은 효과적인 판별을 위해 충분하지 않을 수 있다. 따라서 더욱더 객관적이고 명확하게 우울증을 판별하기 위해 음성 및 텍스트를 사용하여 다양한 데이터 및 특징을 고려하여 우울증 여부의 판별을 진행한다. 본 시스템은 먼저 사림들에게 녹음을 통해 음성 데이터를 수집한 뒤, 음성의 1차원 스펙트럼 특징은 양방향장단기기억 네트워크로 2차원 스펙트로그램 특징은 합성곱신경망기반 전이학습 모델을 통해 특징을 분석한다. 다음 단계로 수집된 음성을 텍스트로 변환한 뒤 텍스트에서 우울증 환자들의 특징을 양방향장단기기억 네트워크와 n-gram 기반 합성곱신경망을 기반으로 분석한다. 마지막으로 우울증 판별은 각 모델의 소프트맥스 스코어를 융합하여 최종적으로 분류하는 방법으로 구성되어있다.

[0011] 도 1은 음성 및 텍스트 데이터의 특징을 분석한 뒤 우울증을 판별하는 시스템을 보여준다. 녹음을 통해 음성 데이터를 취득하고 이를 텍스트 데이터로 변환한 뒤, 두 데이터에 대한 여러 가지 특징들을 추출하여 양방향장단기기억 네트워크와 합성곱신경망을 기반으로 분석함으로써 우울증의 여부를 판단하는 시스템을 구현한다.

[0012] 도 2는 음성신호의 1차원 스펙트럼 특징을 얻어 양방향장단기기억 네트워크를 통해 분류하고, 음성신호를 시간-주파수 변환을 기반으로 2차원 이미지인 스펙트로그램 특징으로 변환하여 합성곱신경망 전이학습 모델을 통해 분류하는 모델을 보여준다. 음성의 1차원 스펙트럼 특징에서는 음성의 배음 구조와 음소의 형태를 분석할 수 있다. 또한 음성을 시간-주파수 영역의 2차원 이미지인 스펙트로그램 특징으로 변환하여 분석함으로써 비선형성 특성이 있는 음성의 시간에 따라 변하는 주파수 특징을 고려할 수 있다. 이 두 가지 특징을 동시에 고려함에 따라 많은 정보를 활용하여 더 정확하게 우울증을 판별할 수 있는 모델을 설계할 수 있다.

[0013] 도 3은 수집한 음성을 텍스트로 변환한 뒤, 토큰화 및 전처리과정을 거친 텍스트를 순차적으로 처리할 수 있는 양방향장단기기억 네트워크와 시간 차원에 대해 합성곱 연산을 취하는 1차원 합성곱신경망을 기반으로 분류하는 모델을 보여준다. 이는 텍스트의 순차적인 특징을 분석할 수 있고 n-gram 개념을 이용하여 텍스트 데이터에서 단어의 순서를 고려하여 다음 단어를 예측하고 오타를 발견할 수 있다. 따라서 입력값을 순차적으로 처리할 수 있을 뿐만 아니라 단어 및 표현의 등장순서도 학습에 반영할 수 있는 모델을 설계할 수 있다.

[0014] 도 4는 앞 서 4개 모델 각각의 소프트맥스 분류 스코어를 한 개의 스코어로 융합하여 최종적으로 우울증을 판별할 수 있는 앙상블 모델의 구조를 보여준다. 음성의 1차원 및 2차원 특징들과 텍스트 데이터의 순차적 특징과 단어의 순서 예측 및 표현 등장순서의 특징 등을 분석하여 이 모든 결과를 융합함으로써 우울증 환자의 다양한 특징 및 정보를 통해 우울증의 여부를 더 효과적으로 판단할 것으로 기대할 수 있다.

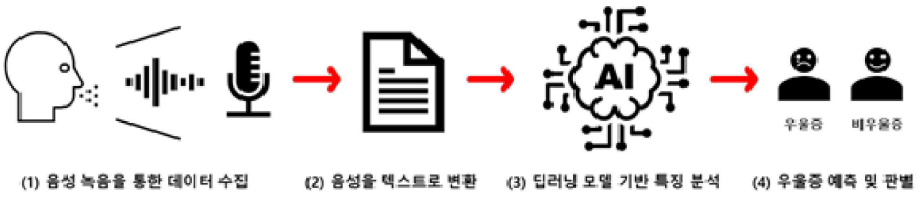
[0015] 도 5는 텍스트 분류의 합성곱신경망 설계에 사용된 n-gram 언어모델의 개념을 간단한 예시를 통해 보여준다. n-gram이란 연속적인 n개의 단어로 구성된 문장을 횟수를 사용하여 벡터로 표현하는 방법이다. 이는 n의 수에 단어 시퀀스를 몇 개의 단어로 구성할지 결정한다. n = 1, 2, 3인 경우를 각각 Uni-gram, Bi-gram, Tri-gram이라 부르며 도 5를 보면 단어 시퀀스가 어떻게 구성되는지 확인할 수 있다. n-gram은 다음 단어를 예측하고, 오타를 발견할 수 있고, 단어의 순서를 무시하는 방법의 단점을 보완할 수 있다. 실제 네트워크 학습 시 n-gram 개념을 적용한 합성곱신경망의 성능이 조금 더 향상됨을 확인하였다.

[0016] 최종적으로 도 4의 각 모델의 소프트맥스 값을 융합하는 앙상블 모델을 음성 및 텍스트 분석 및 우울증을 판별에 사용하여 우울증 환자의 다양한 특징 및 정보를 통해 우울증을 정확하고 객관적으로 예측 및 판별할 수 있는 강력한 시스템을 구현한다.

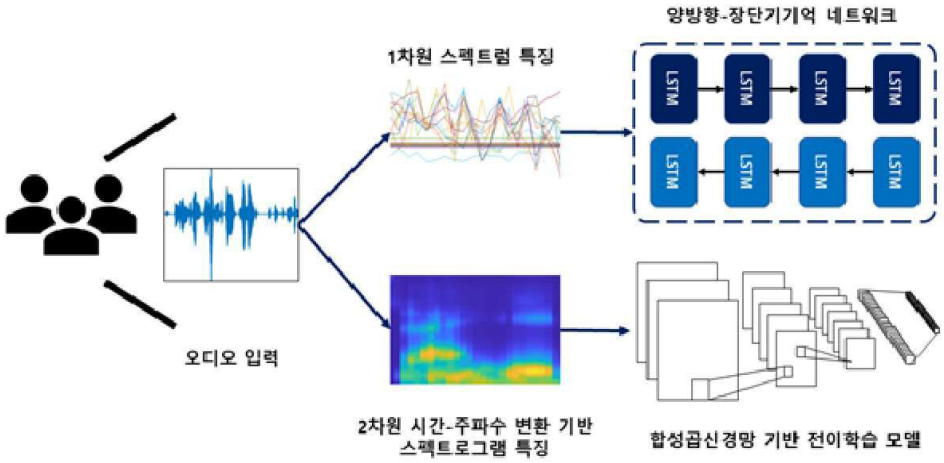
[0017] 본 발명의 효과는 우울증을 보다 정확하고 객관적으로 예측 및 판별하기 위해 한 개의 데이터만으로 분석하지 않고, 음성 및 텍스트를 이용하여 우울증의 특징을 분석한다. 우울증은 복합적인 정신질환이기 때문에 다양한 데이터 및 특징 정보를 고려하여 진단하는 것이 중요하다. 특히 음성 또는 텍스트와 같은 언어적인 데이터에서는 음역대 및 속도 감소, 자주 사용하는 단어 등 우울증 환자들의 특징들이 잘 드러난다. 따라서 녹음을 통해 음성 데이터를 수집한 뒤, 텍스트 데이터로 변환하여 이 두 가지의 데이터를 이용하여 우울증의 여부를 판별한다. 구체적으로 음성의 1차원 스펙트럼 특징을 이영한 양방향장단기기억 네트워크와 시간-주파수 변환 기반 2차원 스펙트로그램 특징을 이용한 합성곱신경망 기반 전이학습 네트워크 분류 모델을 설계한다. 또한 텍스트 데이터를 토큰화 및 전처리과정을 거친 후 각각 양방향장단기기억 네트워크와 n-gram 기반 합성곱신경망 분류 모델을 설계한다. 최종적으로 앞선 4개의 네트워크를 결합함으로써 보다 다양한 정보와 특징을 사용하여 객관적이고 정확하게 우울증을 판별할 수 있을 것으로 기대된다.

도면

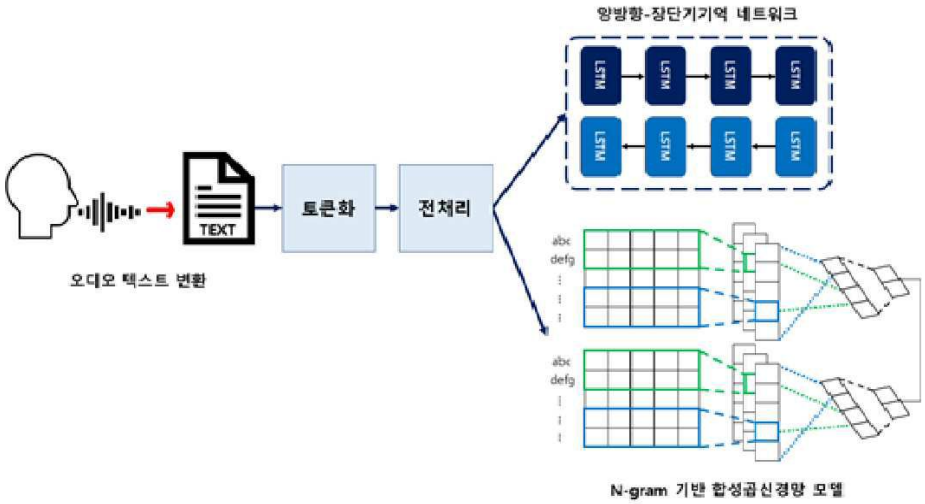
도면1



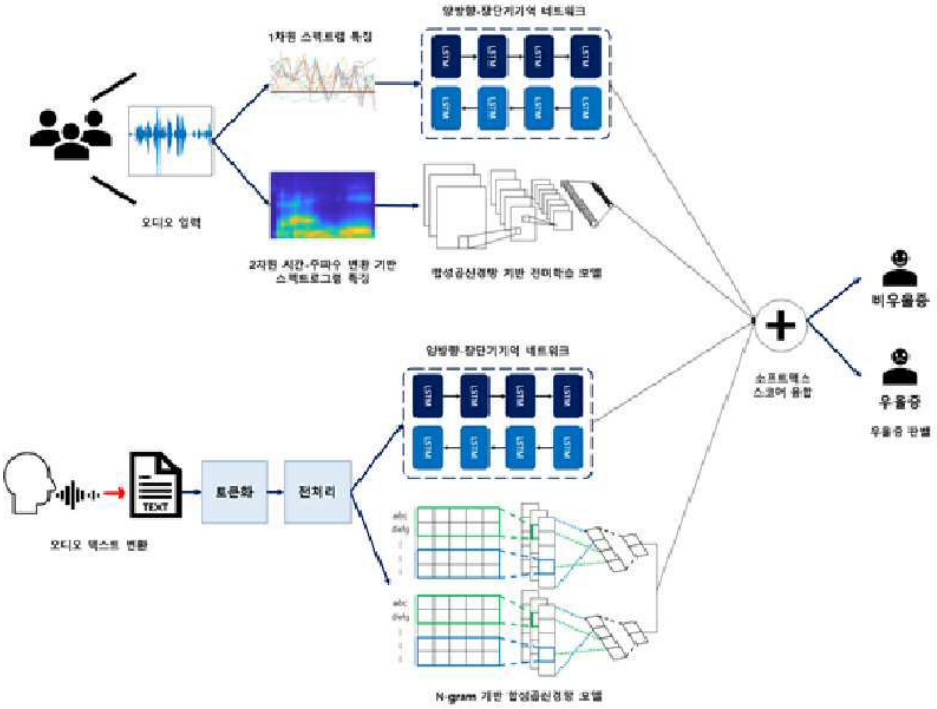
도면2


양방향-장단기기억 네트워크
1차원 스펙트럼 특징
LSTM
오디오 입력
2차원 시간-주파수 변환 기반
스펙트로그램 특징
합성곱신경망 기반 전이학습 모델

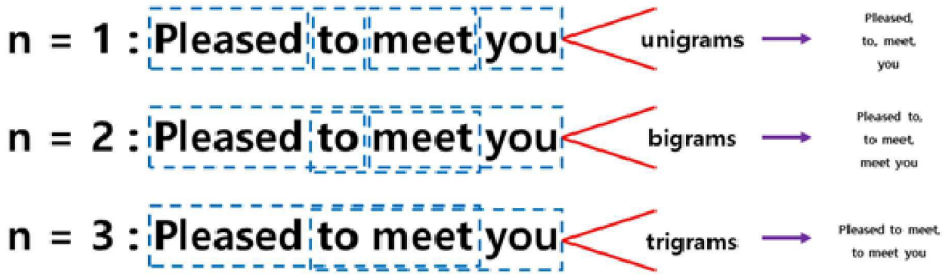
도면3


양방향-장단기기억 네트워크
LSTM
TEXT
토큰화
전처리
오디오 텍스트 변환
abc
defg
N-gram 기반 합성곱신경망 모델

도면4


양방향-장단기기억 네트워크
1차원 스펙트럼 특징
오디오 입력
2차원 시간-주파수 변환 기반 스펙트로그램 특징
멀티 공신경망 기반 정미학습 모델
양방향-장단기기억 네트워크
오디오 텍스트 변환
TEXT
토큰화
전처리
N-gram 기반 합성곱신경망 모델
소프트맥스 스코어 융합
비우울증
우울증
우울증 판별

도면5


n = 1 : Pleased to meet you
unigrams
Pleased, to, meet, you
n = 2 : Pleased to meet you
bigrams
Pleased to, to meet, meet you
n = 3 : Pleased to meet you
trigrams
Pleased to meet, to meet you