(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0120370
(43) 공개일자 2024년08월07일

(51) 국제특허분류(Int. Cl.)
G16H 50/80 (2018.01) *A61B 5/00* (2021.01)
A61B 5/024 (2006.01) *A61B 5/11* (2006.01)
G06N 20/00 (2019.01) *G16H 10/60* (2018.01)
(52) CPC특허분류
G16H 50/80 (2018.01)
A61B 5/024 (2013.01)
(21) 출원번호 10-2023-0012960
(22) 출원일자 2023년01월31일
심사청구일자 2023년01월31일

(71) 출원인
경상국립대학교산학협력단
경상남도 진주시 진주대로 501 (가좌동)
주식회사 딥노이드
서울특별시 구로구 디지털로33길 55, 13층 제1305호(구로동, 이앤씨벤처드림타워2차)
재단법인 정석연구재단
경상남도 김해시 주촌면 골든루트로 80-59 ,705호(김해의생명센터테크노타운운동)
(72) 발명자
김진현
경남 통영시 광도면 향교옆길 33-31
한용섭
경남 창원시 성산구 삼정자로 11
(뒷면에 계속)
(74) 대리인
김종석

전체 청구항 수 : 총 8 항

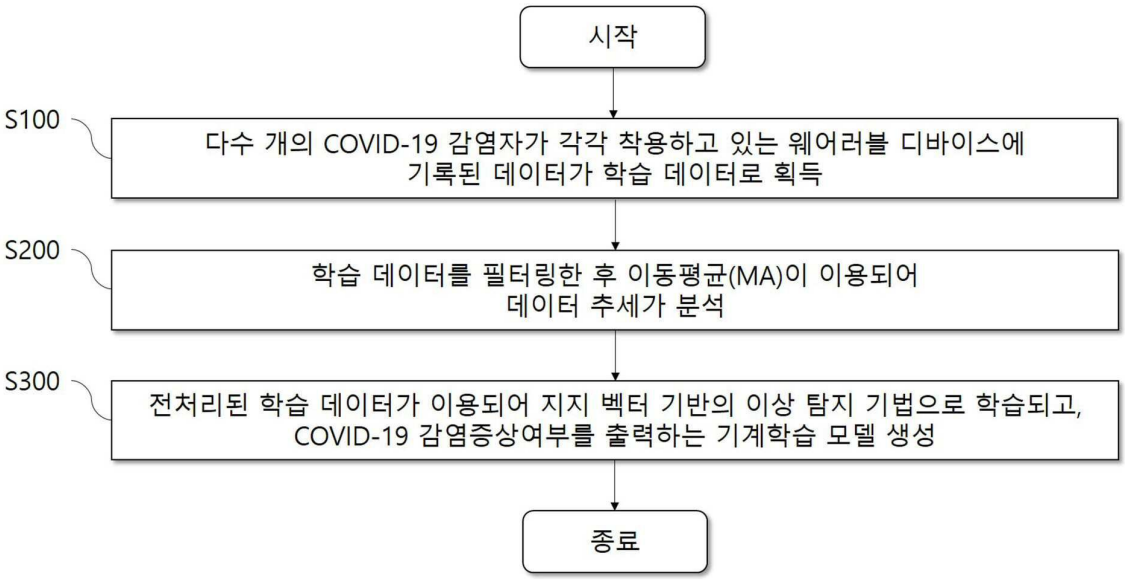
(54) 발명의 명칭 **기계학습 모델 학습 방법, 기계학습 기반 COVID-19 전조증상 감지 방법 및 이를 수행하는 프로그램이 기록된 컴퓨터 판독이 가능한 기록매체**

(57) 요 약

본 발명은 기계학습 모델 학습 방법, 기계학습 기반 COVID-19 전조증상 감지 방법 및 이를 수행하는 프로그램이 기록된 컴퓨터 판독이 가능한 기록매체에 관한 것으로, 보다 구체적으로, 적어도 하나의 프로세서에 의하여, 다수의 COVID-19 감염자가 각각 착용하고 있는 웨어러블 디바이스에 기록된 데이터가 학습 데이터로 획득되는 학습

(뒷면에 계속)

대 표 도 - 도1



데이터 획득단계, 상기 적어도 하나의 프로세서에 의하여, 상기 학습 데이터를 필터링한 후 이동평균(Moving Average; MA)이 이용되어 데이터 추세가 분석되는 학습 데이터 전처리 단계 및 상기 적어도 하나의 프로세서에 의하여, 전처리된 학습 데이터가 이용되어 지지 벡터 기반의 이상 탐지 기법으로 학습되고, COVID-19 감염증상여부를 출력하는 기계학습 모델이 생성되는 모델 학습 단계를 포함하는 기계학습 모델 학습 방법, 기계학습 기반 COVID-19 전조증상 감지 방법 및 이를 수행하는 프로그램이 기록된 컴퓨터 판독이 가능한 기록매체에 관한 것이다.

(52) CPC특허분류
A61B 5/112 (2013.01)
A61B 5/6822 (2013.01)
A61B 5/6824 (2013.01)
A61B 5/6829 (2013.01)
G06N 20/00 (2021.08)
G16H 10/60 (2021.08)

(72) 발명자
강태신
경남 창원시 성산구 삼정자로52번길 11-30
조형래
경남 창원시 마산합포구 중앙북1길 19
윤혜린
경남 통영시 통영해안로 2
김태규
경기 용인시 기흥구 동백죽전대로 455-17
최현주
경남 김해시 무계로 80
최우식
인천 계양구 새풀로 30
이승환
부산 북구 금곡대로 268

이 발명을 지원한 국가연구개발사업
과제고유번호 1711162341
과제번호 2020R1A2C1014855
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 개인기초연구(과기정통부)
연구과제명 자율주행자동차 엣지 간 핸드오버를 위한 동적 시간민감네트워크(TSN) 및 네트워크 슬라이싱 기술 및 개발
기 여 율 1/2
과제수행기관명 경상국립대학교
연구기간 2022.03.01 ~ 2023.02.28
이 발명을 지원한 국가연구개발사업
과제고유번호 1345341781
과제번호 2021RIS-003
부처명 교육부
과제관리(전문)기관명 한국연구재단
연구사업명 지자체-대학협력기반지역혁신사업
연구과제명 (울산경남지역혁신플랫폼)경상국립대학교
기 여 율 1/2
과제수행기관명 경상대학
연구기간 2021.06.01 ~ 2022.04.30

명 세 서

청구범위

청구항 1

적어도 하나의 프로세서에 의하여, 다수의 COVID-19 감염자가 각각 착용하고 있는 웨어러블 디바이스에 기록된 데이터가 학습 데이터로 획득되는 학습 데이터 획득단계;

상기 적어도 하나의 프로세서에 의하여, 상기 학습 데이터를 필터링한 후 이동평균(Moving Average; MA)이 이용되어 데이터 추세가 분석되는 학습 데이터 전처리 단계; 및

상기 적어도 하나의 프로세서에 의하여, 전처리된 학습 데이터가 이용되어 지지 벡터 기반의 이상 탐지 기법으로 학습되고, COVID-19 감염증상여부를 출력하는 기계학습 모델이 생성되는 모델 학습 단계;를 포함하는 기계학습 모델 학습 방법.

청구항 2

제 1항에 있어서,

상기 학습 데이터는,

시간별 걸음수, 시간별 심박수, COVID-19 증상발현일자 및 COVID-19 확진일자 중 적어도 하나를 포함하는 것을 특징으로 하는 기계학습 모델 학습 방법.

청구항 3

제 1항에 있어서,

상기 학습 데이터 전처리 단계는,

가양성(False Positives; FP)을 증가시키지 않음과 동시에 최소 이상 탐지 기간(Minimum Anomaly Detection Period; MADP)을 줄일 수 있도록 상기 이동평균(MA)의 크기(k_p)가 최적화되는 최적화 단계;를 포함하고,

이동평균(MA)의 최적크기(k_o)는,

300 내지 400 범위이내인 것을 특징으로 하는 기계학습 모델 학습 방법.

청구항 4

적어도 하나의 프로세서에 의하여, 다수의 COVID-19 감염자가 각각 착용하고 있는 웨어러블 디바이스에 기록된 데이터가 학습 데이터로 획득되는 학습 데이터 획득단계;

상기 적어도 하나의 프로세서에 의하여, 상기 학습 데이터를 필터링한 후 이동평균(Moving Average; MA)이 이용되어 데이터 추세가 분석되는 학습 데이터 전처리 단계;

상기 적어도 하나의 프로세서에 의하여, 전처리된 학습 데이터가 이용되어 지지 벡터 기반의 이상 탐지 기법으로 학습되고, COVID-19 감염증상여부를 출력하는 기계학습 모델이 생성되는 모델 학습 단계;

상기 적어도 하나의 프로세서에 의하여, 피검사자가 착용하고 있는 웨어러블 디바이스로부터 측정된 측정 데이터가 획득되는 측정 데이터 획득단계;

상기 적어도 하나의 프로세서에 의하여, 피검사자의 상기 측정 데이터를 필터링한 후 이동평균(MA)이 이용되어 데이터 추세가 분석되는 측정 데이터 전처리 단계; 및

상기 적어도 하나의 프로세서에 의하여, 상기 모델 학습 단계로부터 학습된 기계학습 모델에 피검사자의 전처리된 측정 데이터가 입력됨으로써, COVID-19 전조증상여부가 출력되는 결과 출력단계;를 포함하는 기계학습 기반 COVID-19 전조증상 감지 방법.

청구항 5

제 4항에 있어서,

상기 학습 데이터는,

시간별 걸음수, 시간별 심박수, COVID-19 증상발현일자 및 COVID-19 확진일자 중 적어도 하나를 포함하는 것을 특징으로 하는 기계학습 기반 COVID-19 전조증상 감지 방법.

청구항 6

제 4항에 있어서,

피검사자의 상기 측정 데이터는,

시간별 걸음수, 시간별 심박수를 포함하는 것을 특징으로 하는 기계학습 기반 COVID-19 전조증상 감지 방법.

청구항 7

제 4항에 있어서,

상기 결과 출력단계는,

COVID-19 전조증상이 있는 피검사자의 경우 피검사자 단말 및 검사자 단말 중 적어도 하나에 알람을 제공하는 것을 특징으로 하는 기계학습 기반 COVID-19 전조증상 감지 방법.

청구항 8

제1항 내지 제7항 중 어느 한 항의 기계학습 모델 학습 방법 또는 기계학습 기반 COVID-19 전조증상 감지 방법을 수행하는 프로그램을 기록한 컴퓨터 장치로 읽을 수 있는 기록매체.

발명의 설명

기 술 분 야

[0001] 본 발명은 기계학습 모델 학습 방법, 기계학습 기반 COVID-19 전조증상 감지 방법 및 이를 수행하는 프로그램이 기록된 컴퓨터 판독이 가능한 기록매체에 관한 것이다.

배 경 기 술

[0002] 2019년 중국 우한에서 신규 코로나바이러스에 의한 폐렴이 출현하여, 세계보건기구(WHO)는 2020년 2월 이 폐렴을 코로나바이러스 질환(Coronavirus disease 2019; COVID-19)라 명명하였다. COVID-19는 감염된 여행자들에 의해 전 세계로 빠르게 확산되었다. 감염 발생 국가에 따라 사망률이 1% 이하에서 10% 이상인 국가들이 있으며, 평균 6%의 사망률을 보인다.

[0003] COVID-19의 주요 사망 원인은 급성 호흡기 증후군에 의한 중증 폐렴이지만, 환자의 약 80%는 발열, 감기, 인후통, 근육통 등의 경미한 증상 또는 무증상이 나타나고, 나머지 20%의 환자에서만이 보충적 산소 요법을 필요로 하는 중증 폐렴으로 진행된다.

[0004] 즉, 일반적으로 신체 건강한 사람들이 감염된 경우 COVID-19의 증상이 감기와 유사한 경미한 증상 또는 무증상으로 시작된다. 경미한 증상자 및 무증상자는 자신이 감염된 지도 모른 채 가족, 회사동료, 지인들과의 밀접한 접촉을 하고, 국내외를 돌아다니면서 불특정 다수와 접촉하면서 감염시키게 된다. 이는, COVID-19 감염의 확산세가 줄어들지 않는 원인 중 하나이다.

[0005] COVID-19의 전염병이 전 세계적으로 크게 유행하는 팬데믹(Pandemic)인 만큼 COVID-19의 전조증상을 미리 파악하여 COVID-19 감염의 확산세를 줄일 수 있는 방법이 절실히 필요한 실정이다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 대한민국 공개특허공보 제10-2021-0144562호

(특허문헌 0002) 대한민국 공개특허공보 제10-2022-0054401호

발명의 내용

해결하려는 과제

[0007] 본 발명은 상기와 같은 문제점을 해결하기 위한 것으로 가양성(False Positives; FP)을 증가시키지 않음과 동시에 최소 이상 탐지 기간(Minimum Anomaly Detection Period; MADP)을 줄여 많은 량의 시계열 데이터가 획득되지 않아도 보다 신속 정확하게 COVID-19 감염증상여부를 출력할 수 있는 기계학습 모델을 생성하는 기계학습 모델 학습 방법을 얻고자 하는 것을 목적으로 한다.

[0008] 또한, 본 발명의 목적은 COVID-19의 전조증상을 미리 파악하여 COVID-19 감염의 확산세가 줄어들 수 있도록 피검사자가 착용하고 있는 상기 웨어러블 디바이스로부터 측정된 측정 데이터가 획득되고, 학습된 기계학습 모델에 피검사자의 전처리된 측정 데이터가 입력됨으로써 COVID-19 전조증상여부가 출력되는 기계학습 기반 COVID-19 전조증상 감지 방법을 제공하는 것이다.

[0009] 본 발명이 이루고자 하는 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 본 발명의 기재로부터 당해 분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있다.

과제의 해결 수단

[0010] 상기 목적을 달성하기 위하여, 본 발명의 기계학습 모델 학습 방법은 적어도 하나의 프로세서에 의하여, 다수의 COVID-19 감염자가 각각 착용하고 있는 웨어러블 디바이스에 기록된 데이터가 학습 데이터로 획득되는 학습 데이터 획득단계; 상기 적어도 하나의 프로세서에 의하여, 상기 학습 데이터를 필터링한 후 이동평균(Moving Average; MA)이 이용되어 데이터 추세가 분석되는 학습 데이터 전처리 단계; 및 상기 적어도 하나의 프로세서에 의하여, 전처리된 학습 데이터가 이용되어 지지 벡터 기반의 이상 탐지 기법으로 학습되고, COVID-19 감염증상여부를 출력하는 기계학습 모델이 생성되는 모델 학습 단계;를 제공한다.

[0011] 상기 목적을 달성하기 위하여, 본 발명의 기계학습 기반 COVID-19 전조증상 감지 방법은 적어도 하나의 프로세서에 의하여, 다수의 COVID-19 감염자가 각각 착용하고 있는 웨어러블 디바이스에 기록된 데이터가 학습 데이터로 획득되는 학습 데이터 획득단계; 상기 적어도 하나의 프로세서에 의하여, 상기 학습 데이터를 필터링한 후 이동평균(Moving Average; MA)이 이용되어 데이터 추세가 분석되는 학습 데이터 전처리 단계; 상기 적어도 하나의 프로세서에 의하여, 전처리된 학습 데이터가 이용되어 지지 벡터 기반의 이상 탐지 기법으로 학습되고, COVID-19 감염증상여부를 출력하는 기계학습 모델이 생성되는 모델 학습 단계; 상기 적어도 하나의 프로세서에 의하여, 피검사자가 착용하고 있는 상기 웨어러블 디바이스로부터 측정된 측정 데이터가 획득되는 측정 데이터 획득단계; 상기 적어도 하나의 프로세서에 의하여, 피검사자의 상기 측정 데이터를 필터링한 후 이동평균(MA)이 이용되어 데이터 추세가 분석되는 측정 데이터 전처리 단계; 및 상기 적어도 하나의 프로세서에 의하여, 상기 모델 학습 단계로부터 학습된 기계학습 모델에 피검사자의 전처리된 측정 데이터가 입력됨으로써, COVID-19 전조증상여부가 출력되는 결과 출력단계;를 제공한다.

[0012] 상기 목적을 달성하기 위하여, 본 발명은 기계학습 모델 학습 방법 또는 기계학습 기반 COVID-19 전조증상 감지 방법을 수행하는 프로그램을 기록한 컴퓨터 장치로 읽을 수 있는 기록매체를 제공한다.

발명의 효과

[0013] 이상과 같이 본 발명에 의하면 학습 데이터를 필터링한 후 이동평균(Moving Average; MA)이 이용되어 데이터 추세가 분석되고, 전처리된 학습 데이터가 이용되어 지지 벡터 기반의 이상 탐지 기법으로 학습되도록 함으로써, 가양성(False Positives; FP)을 증가시키지 않음과 동시에 최소 이상 탐지 기간(Minimum Anomaly Detection Period; MADP)을 줄여 많은 량의 시계열 데이터가 획득되지 않아도 보다 신속 정확하게 COVID-19 감염증상여부를 출력할 수 있는 효과가 있다.

[0014] 또한, 본 발명은 피검사자가 착용하고 있는 상기 웨어러블 디바이스로부터 측정된 측정 데이터가 획득되고, 학습된 기계학습 모델에 피검사자의 전처리된 측정 데이터가 입력되면 COVID-19 전조증상여부가 출력됨으로써,

COVID-19의 전조증상을 미리 파악하여 COVID-19 감염의 확산세가 줄어들 수 있는 현저한 효과가 있다.

[0015] 본 발명의 효과들은 이상에서 언급한 효과들로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 상세한 설명 및 청구범위의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0016] 도 1은 본 발명의 기계학습 모델 학습 방법 흐름도이다.

도 2는 본 발명의 기계학습 모델을 이용한 COVID-19 전조증상 감지 방법 흐름도이다.

도 3은 본 발명의 일실시예에 따른 전처리된 학습 데이터 중에서 시간별 휴식 심박수(RHR)(a) 및 시간별 걸음 심박수(HROS)(b)가 표시된 도면이다.

발명을 실시하기 위한 구체적인 내용

[0017] 본 명세서에서 사용되는 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나, 이는 당 분야에 종사하는 기술자의 의도 또는 판례, 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.

[0018] 다르게 정의되지 않는 한 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0019] 이하, 본 발명에 따른 실시예를 첨부한 도면을 참조하여 상세히 설명하기로 한다. 도 1은 본 발명의 기계학습 모델 학습 방법 흐름도이다. 도 2는 본 발명의 기계학습 모델을 이용한 COVID-19 전조증상 감지 방법 흐름도이다. 도 3은 본 발명의 일실시예에 따른 전처리된 학습 데이터 중에서 시간별 휴식 심박수(RHR)(a) 및 시간별 걸음 심박수(HROS)(b)가 표시된 도면이다.

[0020] 우선, 본 발명은 기계학습 기반 COVID-19 전조증상 감지 방법을 수행하는 프로그램을 기록한 컴퓨터 장치(100)로 읽을 수 있는 기록매체(120)를 포함한다. 상기 기록매체(120)는 예컨대, CD, DVD, 하드디스크, 블루레이 디스크, USB, 메모리 카드, ROM 등일 수 있다. 그리고 본 발명의 기계학습 기반 COVID-19 전조증상 감지 방법은 컴퓨터 장치(100) 내 적어도 하나의 프로세서(110)가 상기 기록매체(120)를 읽음으로써 구현될 수 있다.

[0022] **기계학습 모델 학습 방법**

[0023] 도 1을 보면, 본 발명의 기계학습 모델 학습 방법은 학습 데이터 획득단계(S100), 학습 데이터 전처리 단계(S200) 및 모델 학습 단계(S300)를 포함한다.

[0024] 보다 구체적으로, 상기 학습 데이터 획득단계(S100)는 적어도 하나의 프로세서(110)에 의하여, 다수의 COVID-19 감염자가 각각 착용하고 있는 웨어러블 디바이스(200)에 기록된 데이터가 학습 데이터로 획득된다.

[0025] COVID-19 감염자는 발열, 감기, 인후통, 근육통 등 경미한 증상 또는 무증상이 나타나고, 감염 초기에 자체적으로 느끼기에는 그 증상이 상당히 미미하다. 따라서 이를 미리 확인할 수 있는 유의미한 지표를 선정하는 것이 상당히 중요하다. 일반적으로 사람의 체온이 상승하게 되면 심박수와 호흡수가 증가함으로, 심박수와 호흡수는 COVID-19 전조증상을 확인하는데 유의미한 지표가 될 수 있다. 그리고 연속적인 시간에서의 심박수의 변동성(HRV) 역시 유의미한 지표가 될 수 있다. 이때, 사람의 심박수와 호흡수가 가장 안정되는 시간인 수면시간에서의 심박수와 호흡수 비교를 통해서 보다 정확하게 COVID-19 전조증상을 감지할 수 있다.

[0026] 따라서 상기 학습 데이터는, 시간별 걸음수, 시간별 심박수, COVID-19 증상발현일자 및 COVID-19 확진일자 중 적어도 하나를 포함하는 것을 특징으로 한다. 본 발명에서 언급하는 상기 COVID-19 증상발현일자는 자체적으로 발열, 감기, 인후통, 근육통 등 COVID-19로 의심되는 증상들이 발현되었다고 판단하고 상기 웨어러블 디바이스(200)를 통해서 기록한 일자이고, 상기 COVID-19 확진일자는 의료기관 및 COVID-19 검사기관으로부터 COVID-19

양성으로 판정받은 일자를 상기 웨어러블 디바이스(200)를 통해서 기록한 일자이다. 즉, 상기 학습 데이터는 시간을 X축으로 하고, 심박수를 Y축으로 하는 시계열 데이터이다.

[0027] 또한, 상기 웨어러블 디바이스(200)는 COVID-19 감염자가 시계, 목걸이, 팔찌, 발찌 등의 형태로 일상생활에서 불편함을 느끼지 않고 24시간 내내 착용할 수 있는 장치이다. 다시 말하면, 상기 웨어러블 디바이스(200)는 사람이 착용하고 있는 동안 심박수, 체온, 호흡수, 걸음수에 대한 측정 및 기록이 가능하고, 측정 및 기록된 데이터를 가공하여 화면에 표시할 수 있고, 유무선 통신방식을 통해서 외부기기로 측정 및 기록된 데이터를 전송할 수 있다.

[0028] 다음으로, 상기 학습 데이터 전처리 단계(S200)는 상기 적어도 하나의 프로세서(110)에 의하여, 상기 학습 데이터가 필터링된 후 이동평균(Moving Average; MA)이 이용되어 데이터 추세가 분석된다.

[0029] 우선, 상기 학습 데이터 전처리 단계(S200)는, 상기 학습 데이터 중 시간별 걸음수가 이용되어 시간별 심박수가 시간별 휴식 심박수(Reating Heart Rate; RHR)와 시간별 걸음 심박수(Heart Rate Of Steps; HROS)로 필터링될 수 있다. 본 발명에서 언급하는 상기 휴식 심박수(RHR)는 걸음수가 0일 때의 측정된 심박수로, 눕기, 앉기, 수면 등의 상태에서 시간에 따라 연속적으로 측정된 심박수일 수 있다. 본 발명에 언급하는 상기 걸음 심박수(HROS)는 걸음수가 1 이상일 때의 심박수로, 걷기, 뛰기, 운동 등 몸을 움직이는 상태에서 시간에 따라 연속적으로 측정된 심박수일 수 있다. 즉, 상기 학습 데이터 전처리 단계(S200)로부터 시간별 심박수가 필터링됨으로써, 동작변화에 따라 이상치로 분류되는 것을 확인할 수 있어 학습 정확도를 향상시키는 현저한 효과가 있다.

[0030] 또한, 일반적으로, 이동평균(MA)은 추세의 변동을 알 수 있도록 구간을 옮겨 가면서 구하는 평균이다. 예컨대, 이동평균(MA)의 크기(k_p)가 5이고 지난 5시간 동안의 시간별 심박수 B_1, B_2, B_3, B_4, B_5의 평균이 AVG_1이라면, 가장 오래된 심박수 B_1 대신에 오늘의 심박수 B_6을 넣어서 새로운 5시간 B_2, B_3, B_4, B_5, B_6의 평균 AVG_2를 구할 수 있다. 이렇게, 구한 AVG_1, AVG_2...를 이동평균(MA)이라고 한다. 즉, 상기 학습 데이터 전처리 단계(S200)로부터 상기 이동평균(MA)이 이용되는 것은 변동이 심하고 불규칙한 비정상이 많은 학습 데이터의 경향을 보여줌으로써 정상 시계열 데이터와 비정상 시계열 데이터를 쉽게 구분할 수 있기 때문이다.

[0031] 이때, 상기 학습 데이터 전처리 단계(S200)는, 가양성(False Positives; FP)을 증가시키지 않음과 동시에 최소 이상 탐지 기간(Minimum Anomaly Detection Period; MADP)을 줄일 수 있도록 상기 이동평균(MA)의 크기(k_p)가 최적화되는 최적화 단계(S210)를 포함하고, 이동평균(MA)의 최적크기(k_o)는, 300 내지 400 범위이내인 것을 특징으로 한다.

[0032] 본 발명에서 언급하는 상기 최소 이상 탐지 기간(Minimum Anomaly Detection Period; MADP)은 초기 유효 이상 탐지에 필요한 최소한의 학습 데이터를 수집하는 기간을 의미한다. 즉, 상기 최소 이상 탐지 기간(MADP)은 COVID-19 감염자별 증상발현일자로부터 그 이전시점까지 초기 유효 이상 탐지에 필요한 최소한의 학습 데이터를 수집하는 기간이거나, COVID-19 감염자별 확진일자로부터 그 이전시점까지 초기 유효 이상 탐지에 필요한 최소한의 학습 데이터를 수집하는 기간일 수 있다. 상기 최소 이상 탐지 기간(MADP)이 줄어들수록 COVID-19 전조증상을 더 빨리 감지할 수 있다.

[0033] 상기 최소 이상 탐지 기간(MADP)은 윈도우 크기(Window size)를 최소화함으로써 줄어들 수 있는데, 본 발명의 상기 윈도우 크기(Window size)는 이동평균(Moving Average; MA)의 크기(k_p)이다. 상기 학습 데이터 전처리 단계(S200)는 가양성(FP)을 증가시키지 않음과 동시에 최소 이상 탐지 기간(MADP)을 줄일 수 있도록 상기 이동평균(MA)의 최적크기(k_o)를 이용하여 데이터 추세가 분석될 수 있다. 상기 이동평균(MA)의 최적크기(k_o)는 300 내지 400 범위 이내일 수 있다.

[0034] 그리고 가양성(False Positives; FP)은 일반적으로 하나의 기준변인상에서 어떤 점수를 기준으로 실패와 성공으로 구분되는 경우에 이를 하나의 검사에 의해서 예언하고자 할 때 기준변인상에서는 실패인데 검사에서는 성공자로 잘못 예언하는 사례를 일컫는 말이다. 본 발명에서 언급하는 가양성(False Positives; FP)은 COVID-19 음성일 수 있으나 양성으로 결과를 출력하는 경우를 말한다. 즉, 가양성이 높아지면 격리하지 않아도 되는 자가 격리를 하게 됨으로 한 사람에게 큰 영향을 미칠 수 있으므로 이를 지양할 수 있도록 한다. 예컨대, 이동평균(MA)의 크기에 따라 심박수의 급격한 변화로 인식되면 그렇지 않은 상태임에도 불구하고 COVID 감염증상이 있는 것으로 분류될 수 있으므로 이동평균(MA)의 크기가 적절한 크기로 최적화되는 것이 상당히 중요하다.

[0035] 한편, 상기 학습 데이터 전처리 단계(S200)는 필터링된 학습 데이터를 대상으로 상기 최적크기(k_o)로 다수 개의

이동평균(MA)이 연산되고, 상기 다수 개의 이동평균(MA)을 이은 이동평균선이 획득되는 이동평균선 획득단계(S220), 상기 이동평균선이 다운샘플링(downsampling)되는 다운샘플링 단계(S230) 및 다운샘플링(downsampling)된 이동평균선이 이용되어 Z점수가 연산됨으로써, 상기 학습 데이터가 정규화되는 정규화 단계(S240)를 더 포함할 수 있다.

[0036] 예컨대, 상기 최적크기(k_o)가 300이라면, 상기 이동평균선 획득단계(S220)는 연속된 300개의 시간별 휴식 심박수(RHR)의 이동평균(MA)이 연산되거나 연속된 300개의 시간별 걸음 심박수(HROS)의 이동평균(MA)이 연산될 수 있다. 그리고 시간별 휴식 심박수(RHR)의 다수 개의 이동평균(MA)을 이은 이동평균선이 획득되거나 시간별 걸음 심박수(HROS)의 다수 개의 이동평균(MA)을 이은 이동평균선이 획득될 수 있다. 이때, 검사자가 확인하고자 하는 심박수의 종류에 따라 획득되는 이동평균선이 선정될 수 있다.

[0037] 또한, 상기 심박수별 다수 개의 이동평균(MA)이 많을수록 이동평균선이 상당히 복잡할 수 있고, 이는 모델 학습의 정확도와 속도를 낮추는 문제점이 있다. 따라서 상기 이동평균선의 경향성을 유지함과 동시에 학습 속도를 높일 수 있도록 상기 다운샘플링 단계(S230)는 이동평균선이 다운샘플링 되어 대략적인 곡선으로 표시될 수 있다. 예컨대, 10000개의 이동평균(MA)을 포함하는 이동평균선의 경우 1000개의 주요한 이동평균(MA)만이 표시됨으로써 보다 대략적인 곡선으로 표시될 수 있는 것이다.

[0038] 마지막으로, 상기 정규화 단계(S240)는 다운샘플링(downsampling)된 이동평균선이 이용되어 Z점수가 연산될 수 있다. 일반적으로, Z점수는 표준편차 측면에서 데이터 점이 평균으로부터 떨어진 거리를 측정한 수치로, 평균이 0이고 분산이 1이 되어 데이터 표준화가 가능하도록 한다. 즉, 상기 정규화 단계(S240)는 Z점수가 연산됨으로써, 심박수별 다수 개의 이동평균(MA)에 대한 데이터 정규화가 가능하도록 한다.

[0039] 따라서 도 3과 같이 상기 학습 데이터 전처리 단계(S200)는 상기 학습 데이터 획득단계(S100)로부터 획득된 학습 데이터가 상기 모델 학습 단계(S300)에 가장 적합한 형태로 전처리 되어 모델 학습의 정확도 및 속도를 향상시킬 수 있는 현저한 효과가 있다.

[0040] 다음으로, 상기 모델 학습 단계(S300)는 상기 적어도 하나의 프로세서(110)에 의하여, 전처리된 학습 데이터가 이용되어 지지 벡터 기반의 이상 탐지 기법으로 학습되고, COVID-19 감염증상여부를 출력하는 기계학습 모델이 생성된다.

[0041] 상기 지지 벡터 기반의 이상 탐지 기법을 포함하는 기계학습 모델(One Class-Support Vector Machine; OC-SVM)은 신규성 감지를 위한 결정 기능을 학습하는 비지도 알고리즘을 포함할 수 있다.

[0042] 일반적으로, SVM(Support Vector Machine)은 모든 데이터 포인트를 특징 공간 F 내 원점에서 분리하고 이 초평면에서 원점까지의 거리를 최대화 한다. 초평면을 구성하려면 RBF(Radial Basic Function)와 가우스 커널에 대한 계수인 g를 결정해야 한다. 여기서, RBF는 단순한 네트워크 구조로, 숨겨진 레이어(Hidden layer)가 하나이고, 선형 출력 레이어를 가지고 있으므로 가중치 계산에 용이하다. 그리고 계수 g가 크면 원점에 가까운 데이터를 지원 벡터(Support vector)로 간주하고, g가 작으면 원점에서 멀리 있는 데이터를 지원 벡터(Support vector)로 간주한다. n은 초평면 외부에 있는 데이터의 최대 비율을 결정한다.

[0043] 즉, 본 발명의 파라미터 값은 최적의 값으로 기 설정될 수 있고, n=0.1, g=0.15, kernal=RBF일 수 있다. 그리고 상기 모델 학습 단계(S300)는 상기 학습 데이터 전처리 단계(S200)로부터 전처리된 학습 데이터가 초평면 상에 구현될 수 있고, 데이터의 공변성(Covariance)에 의해 경계원이 생성될 수 있다. 생성된 경계원은 정상(Normal)과 비정상(Abnormal)을 구분하는 기준이 될 수 있다. 데이터가 경계원 안에 있으면 정상(Normal)으로 레이블이 지정될 수 있고 반대로, 데이터가 경계원 밖에 있으면 비정상(Abnormal)으로 레이블이 지정될 수 있다.

[0044] 따라서 본 발명의 상기 모델 학습 단계(S300)는 새로운 데이터가 입력되면 학습 데이터를 기준으로 정상(Normal)과 비정상(Abnormal)을 구분하여 COVID-19 감염증상여부를 출력할 수 있는 기계학습 모델이 생성될 수 있다.

[0045] 따라서 본 발명에 따르면, 다수의 COVID-19 감염자가 각각 착용하고 있는 웨어러블 디바이스(200)에 기록된 데이터가 학습 데이터로 획득됨으로써, COVID-19 감염자별 생리학적 데이터를 용이하게 획득할 수 있고, 시계열 데이터 형식의 측정 데이터 기반으로 한 기계학습 모델이 생성될 수 있다.

[0046] 또한, 본 발명은 학습 데이터가 필터링된 후 이동평균(MA)이 이용되어 데이터 추세가 분석됨으로써, 동작변화에 따라 이상치로 분류되는 것을 확인할 수 있고, 변동이 심하고 불규칙한 비정상이 많은 학습 데이터의 경향성을 보다 정확하게 확인할 수 있다. 즉, 가양성(FP)을 증가시키지 않음으로써 보다 정확성이 향상된 학습이 가능하

고, 최적화된 기계학습 모델이 생성될 수 있는 현저한 효과가 있다.

[0048] **기계학습 기반 COVID-19 전조증상 감지 방법**

[0049] 도 2를 보면, 본 발명의 기계학습 모델을 이용한 COVID-19 전조증상 감지 방법은 학습 데이터 획득단계(S100), 학습 데이터 전처리 단계(S200) 및 모델 학습 단계(S300), 측정 데이터 획득단계(S400), 측정 데이터 전처리 단계(S500) 및 분석결과 출력단계(S600)를 포함한다.

[0050] 상기 학습 데이터 획득단계(S100)는 적어도 하나의 프로세서(110)에 의하여, 다수의 COVID-19 감염자가 각각 착용하고 있는 웨어러블 디바이스(200)에 기록된 데이터가 학습 데이터로 획득된다. 상기 학습 데이터 전처리 단계(S200)는 상기 적어도 하나의 프로세서(110)에 의하여, 상기 학습 데이터를 필터링한 후 이동평균(Moving Average; MA)이 이용되어 데이터 추세가 분석된다. 상기 모델 학습 단계(S300)는 상기 적어도 하나의 프로세서(110)에 의하여, 전처리된 학습 데이터가 이용되어 지지 벡터 기반의 이상 탐지 기법으로 학습되고, COVID-19 감염증상여부를 출력하는 기계학습 모델이 생성된다. 이때, 기계학습 모델을 이용한 COVID-19 전조증상 감지 방법의 학습 데이터 획득단계(S100), 학습 데이터 전처리 단계(S200) 및 모델 학습 단계(S300)는 상술한 기계학습 모델 학습 방법과 동일하다.

[0051] 다음으로, 상기 측정 데이터 획득단계(S400)는 상기 적어도 하나의 프로세서(110)에 의하여, 피검사자가 착용하고 있는 상기 웨어러블 디바이스(200)로부터 측정된 측정 데이터가 획득된다.

[0052] 피검사자의 상기 측정 데이터는, 시간별 걸음수, 시간별 심박수를 포함하는 것을 특징으로 한다. 즉, 상기 측정 데이터는 시간을 X축으로 하고, 심박수를 Y축으로 하는 시계열 데이터이다.

[0053] 또한, 상기 웨어러블 디바이스(200)는 피검사자가 시계, 목걸이, 팔찌, 발찌 등의 형태로 일상생활에서 불편함을 느끼지 않고 24시간 내내 착용할 수 있는 장치이다. 다시 말하면, 상기 웨어러블 디바이스(200)는 피검사자가 착용하고 있는 동안 심박수, 체온, 호흡수, 걸음수에 대한 측정 및 기록이 가능하고, 측정 및 기록된 데이터를 가공하여 화면에 표시할 수 있고, 유무선 통신방식을 통해서 외부기기로 측정 및 기록된 데이터를 전송할 수 있다.

[0054] 다음으로, 상기 측정 데이터 전처리 단계(S500)는 상기 적어도 하나의 프로세서(110)에 의하여, 상기 측정 데이터가 필터링된 후 이동평균(Moving Average; MA)이 이용되어 데이터 추세가 분석된다.

[0055] 우선, 상기 측정 데이터 전처리 단계(S500)는, 상기 측정 데이터 중 시간별 걸음수가 이용되어 시간별 심박수가 시간별 휴식 심박수(Reating Heart Rate; RHR)와 시간별 걸음 심박수(Heart Rate Of Steps; HROS)로 필터링될 수 있다. 본 발명에서 언급하는 상기 휴식 심박수(RHR)는 걸음수가 0일 때의 측정된 심박수로, 눕기, 앉기, 수면 등의 상태에서 시간에 따라 연속적으로 측정된 심박수일 수 있다. 본 발명에 언급하는 상기 걸음 심박수(HROS)는 걸음수가 1 이상일 때의 심박수로, 걷기, 뛰기, 운동 등 몸을 움직이는 상태에서 시간에 따라 연속적으로 측정된 심박수일 수 있다. 즉, 상기 측정 데이터 전처리 단계(S500)로부터 시간별 심박수가 필터링됨으로써, 동작변화에 따른 이상치를 분류할 수 있고, 보다 정확한 결과가 출력될 수 있도록 하는 현저한 효과가 있다. 이때, 검사자가 확인하고자 하는 심박수의 종류에 따라 휴식 심박수(Reating Heart Rate; RHR)와 시간별 걸음 심박수(Heart Rate Of Steps; HROS) 중 하나가 선정될 수 있다.

[0056] 또한, 일반적으로, 이동평균(MA)은 추세의 변동을 알 수 있도록 구간을 옮겨 가면서 구하는 평균이다. 예컨대, 이동평균(MA)의 크기(k_p)가 5이고 지난 5시간 동안의 시간별 심박수 B_1, B_2, B_3, B_4, B_5의 평균이 AVG_1이라면, 가장 오래된 심박수 B_1 대신에 오늘의 심박수 B_6을 넣어서 새로운 5시간 B_2, B_3, B_4, B_5, B_6의 평균 AVG_2를 구할 수 있다. 이렇게, 구한 AVG_1, AVG_2...를 이동평균(MA)이라고 한다. 즉, 상기 측정 데이터 전처리 단계(S500)로부터 상기 이동평균(MA)이 이용되는 것은 변동이 심하고 불규칙한 비정상이 많은 심박수 측정 데이터의 경향을 보여줌으로써 정상 시계열 데이터와 비정상 시계열 데이터를 쉽게 구분하기 위함이다.

[0057] 한편, 상기 측정 데이터 전처리 단계(S500)는 필터링된 측정 데이터를 대상으로 최적크기(k_o)로 다수 개의 이동평균(MA)이 연산되고, 상기 다수 개의 이동평균(MA)을 이은 이동평균선이 획득되는 이동평균선 획득단계(S510), 상기 필터링된 측정 데이터에 대한 이동평균선이 다운샘플링(downsampling)되는 다운샘플링 단계(S520) 및 다운샘플링(downsampling)된 이동평균선이 이용되어 Z점수가 연산됨으로써, 피검사자의 상기 측정 데이터가 정규화되는 정규화 단계(S530)를 더 포함할 수 있다.

[0058] 예컨대, 상기 최적크기(k_o)가 300이라면, 상기 이동평균선 획득단계(S510)는 연속된 300개의 시간별 휴식 심박수(RHR)의 이동평균(MA)이 연산되거나 연속된 300개의 시간별 걸음 심박수(HROS)의 이동평균(MA)이 연산될 수 있다. 그리고 시간별 휴식 심박수(RHR)의 다수 개의 이동평균(MA)을 이은 이동평균선이 획득되거나 시간별 걸음 심박수(HROS)의 다수 개의 이동평균(MA)을 이은 이동평균선이 획득될 수 있다. 이때, 검사자가 확인하고자 하는 심박수의 종류에 따라 획득되는 이동평균선이 선정될 수 있다.

[0059] 또한, 상기 심박수별 다수 개의 이동평균(MA)이 많을수록 이동평균선이 상당히 복잡할 수 있고, 이는 상기 기계학습 모델의 결과 정확도와 결과 출력 속도를 낮추는 문제점이 있다. 따라서 상기 이동평균선의 경향성을 유지함과 동시에 결과 출력 속도를 높일 수 있도록 상기 다운샘플링 단계(S520)는 이동평균선이 다운샘플링 되어 대략적인 곡선으로 표시될 수 있다. 예컨대, 10000개의 이동평균(MA)을 포함하는 이동평균선의 경우 1000개의 주요한 이동평균(MA)만이 표시됨으로써 보다 대략적인 곡선으로 표시될 수 있는 것이다.

[0060] 마지막으로, 상기 정규화 단계(S530)는 다운샘플링(downsampling)된 이동평균선이 이용되어 Z점수가 연산될 수 있다. 일반적으로, Z점수는 표준편차 측면에서 데이터 점이 평균으로부터 떨어진 거리를 측정한 수치로, 평균이 0이고 분산이 1이 되어 데이터 표준화가 가능하도록 한다. 즉, 상기 정규화 단계(S530)는 Z점수가 연산됨으로써, 심박수별 다수 개의 이동평균(MA)에 대한 데이터 정규화가 가능하도록 한다.

[0061] 따라서 상기 측정 데이터 전처리 단계(S500)는 상기 측정 데이터 획득단계(S400)로부터 획득된 피검사자의 상기 측정 데이터가 상기 기계학습 모델에 입력될 수 있는 가장 적합한 형태로 전처리 되어 결과 출력의 정확도 및 속도를 향상시킬 수 있는 현저한 효과가 있다.

[0062] 다음으로, 상기 결과 출력단계(S600)는 상기 적어도 하나의 프로세서(110)에 의하여, 상기 모델 학습 단계(S300)로부터 학습된 기계학습 모델에 피검사자의 전처리된 측정 데이터가 입력됨으로써, COVID-19 전조증상여부가 출력된다.

[0063] 상기 결과 출력단계(S600)는, COVID-19 전조증상이 있는 피검사자의 경우 피검사자 단말(300) 및 검사자 단말(400) 중 적어도 하나에 알람을 제공하는 것을 특징으로 한다. 종래 신체 건강한 사람들이 감염된 경우 COVID-19의 증상이 감기와 유사한 경미한 증상 또는 무증상으로 시작된다. 경미한 증상자 및 무증상자는 자신이 감염된 지도 모른 채 가족, 회사동료, 지인들과의 밀접한 접촉을 하고, 국내외를 돌아다니면서 불특정 다수와 접촉하면서 감염시키게 된다. 이는, 감염의 확산세가 줄어들지 않는 원인 중 하나이다. 그러나 본 발명은 학습된 기계학습 모델을 통해서 피검사자가 자체적으로 경미한 증상이 느껴지기 전부터 COVID-19 전조증상여부를 파악할 수 있으므로, 보다 신속한 격리가 가능하고, 감염의 확산세를 줄일 수 있는 현저한 효과가 있다.

[0065] 실시예들은 하드웨어, 소프트웨어, 펌웨어, 미들웨어, 마이크로코드, 하드웨어 기술 언어, 또는 이들의 임의의 조합에 의해 구현될 수 있다. 소프트웨어, 펌웨어, 미들웨어 또는 마이크로코드로 구현되는 경우, 필요한 작업을 수행하는 프로그램 코드 또는 코드 세그먼트들은 컴퓨터 판독 가능 저장 매체에 저장되고 하나 이상의 프로세서에 의해 실행될 수 있다.

[0066] 그리고 본 명세서에 설명된 주제의 양태들은 컴퓨터에 의해 실행되는 프로그램 모듈 또는 컴포넌트와 같은 컴퓨터 실행 가능 명령어들의 일반적인 맥락에서 설명될 수 있다. 일반적으로, 프로그램 모듈 또는 컴포넌트들은 특정 작업을 수행하거나 특정 데이터 형식을 구현하는 루틴, 프로그램, 객체, 데이터 구조를 포함한다. 본 명세서에 설명된 주제의 양태들은 통신 네트워크를 통해 링크되는 원격 처리 디바이스들에 의해 작업들이 수행되는 분산 컴퓨팅 환경들에서 실시될 수도 있다. 분산 컴퓨팅 환경에서, 프로그램 모듈들은 메모리 저장 디바이스들을 포함하는 로컬 및 원격 컴퓨터 저장 매체에 둘 다에 위치할 수 있다.

[0067] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 으로 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0068] 그러므로 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

부호의 설명

[0069] 100.. 컴퓨터 장치

110.. 적어도 하나의 프로세서

120.. 기록매체

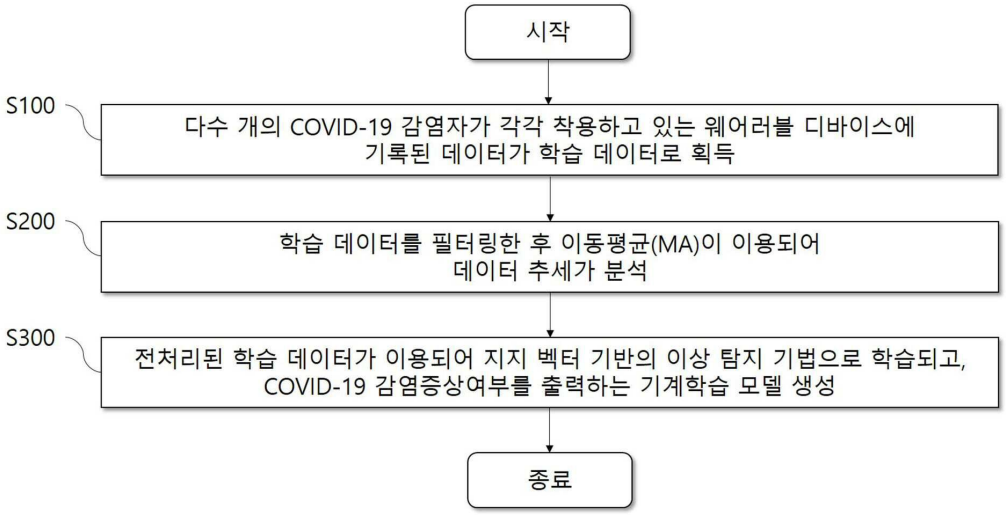
200.. 웨어러블 디바이스

300.. 피검사자 단말

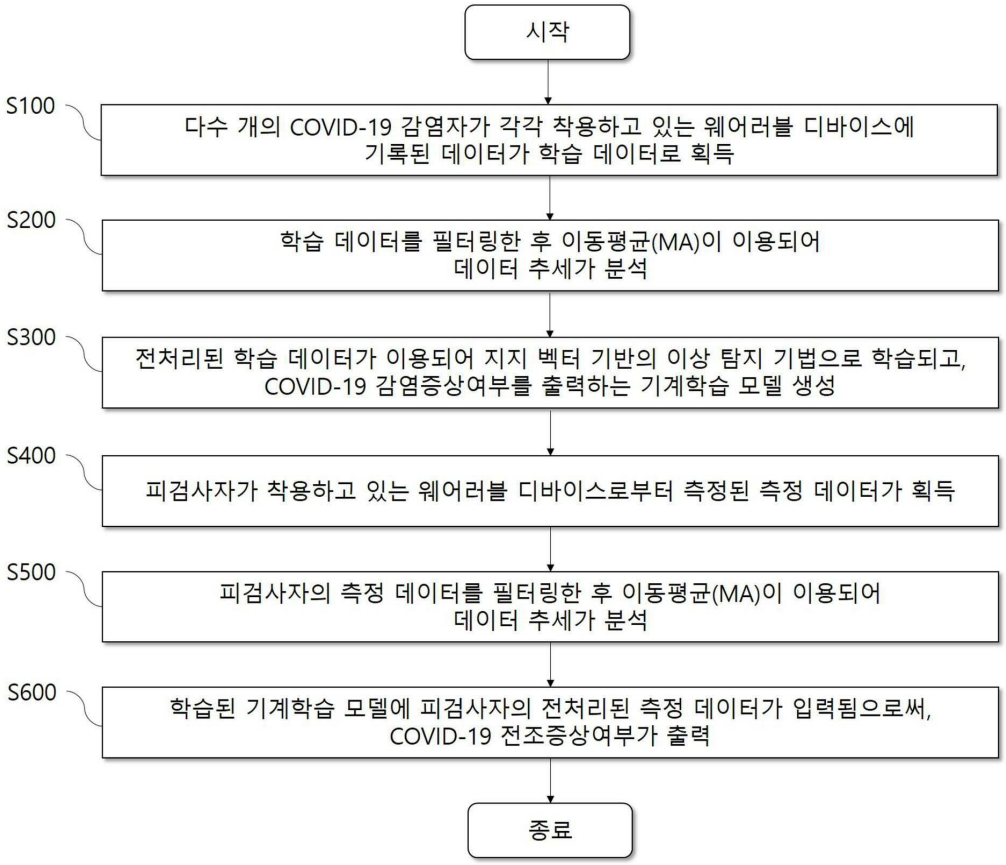
400.. 검사자 단말

도면

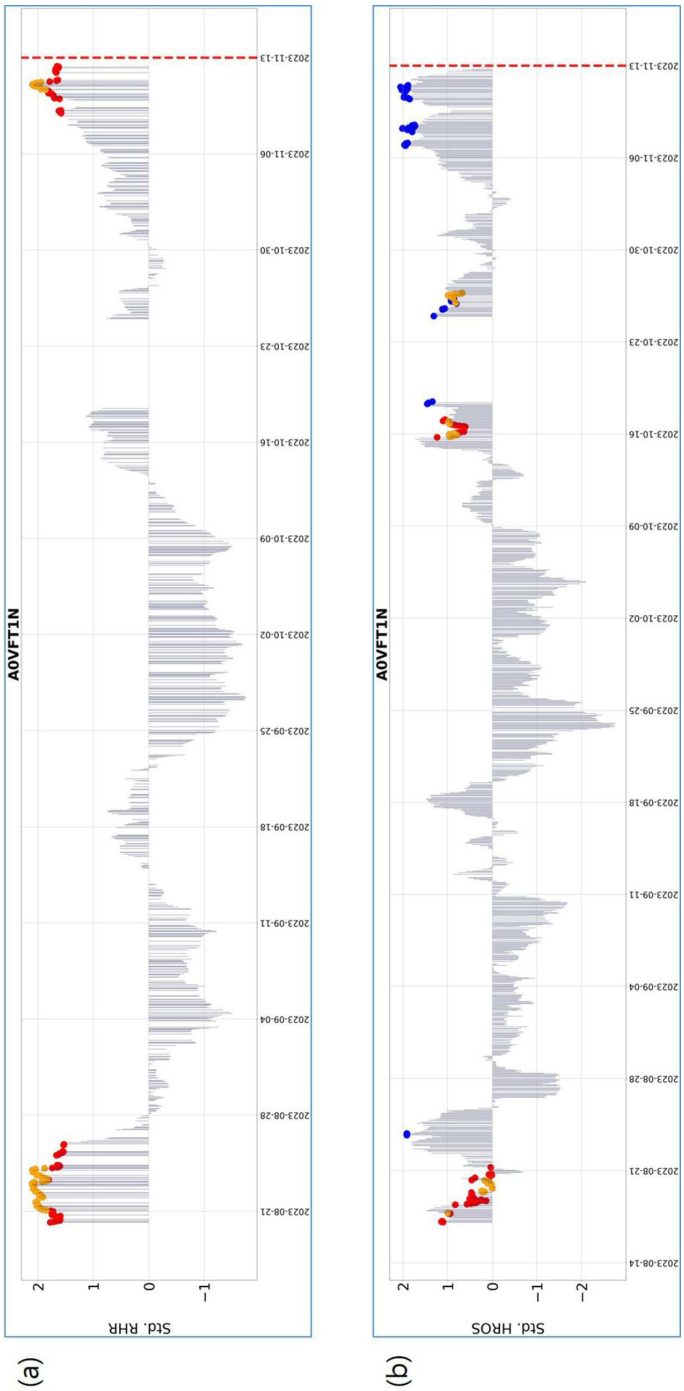
도면1



도면2


시작
S100
다수 개의 COVID-19 감염자가 각각 착용하고 있는 웨어러블 디바이스에 기록된 데이터가 학습 데이터로 획득
S200
학습 데이터를 필터링한 후 이동평균(MA)이 이용되어 데이터 추세가 분석
S300
전처리된 학습 데이터가 이용되어 지지 벡터 기반의 이상 탐지 기법으로 학습되고, COVID-19 감염증상여부를 출력하는 기계학습 모델 생성
S400
피검사자가 착용하고 있는 웨어러블 디바이스로부터 측정된 측정 데이터가 획득
S500
피검사자의 측정 데이터를 필터링한 후 이동평균(MA)이 이용되어 데이터 추세가 분석
S600
학습된 기계학습 모델에 피검사자의 전처리된 측정 데이터가 입력됨으로써, COVID-19 전조증상여부가 출력
종료

도면3


A0VFT1N
Std. RHR
2
1
0
−1
2023-11-13
2023-11-06
2023-10-30
2023-10-23
2023-10-16
2023-10-09
2023-10-02
2023-09-25
2023-09-18
2023-09-11
2023-09-04
2023-08-28
2023-08-21
(a)
A0VFT1N
Std. HROS
2
1
0
−1
−2
2023-11-13
2023-11-06
2023-10-30
2023-10-23
2023-10-16
2023-10-09
2023-10-02
2023-09-25
2023-09-18
2023-09-11
2023-09-04
2023-08-28
2023-08-21
2023-08-14
(b)