

(51) 국제특허분류:

G16H 50/80 (2018.01) G16H 30/40 (2018.01)
G16H 50/20 (2018.01) G06N 20/20 (2019.01)
G16H 50/50 (2018.01) G06T 7/00 (2006.01)
G16H 50/70 (2018.01) A61B 5/00 (2006.01)

(21) 국제출원번호: PCT/KR2024/001776

(22) 국제출원일: 2024 년 2 월 6 일 (06.02.2024)

(25) 출원언어: 한국어

(26) 공개언어: 한국어

(30) 우선권정보:

10-2023-0093794 2023 년 7 월 19 일 (19.07.2023) KR

(71) 출원인: 가톨릭대학교 산학협력단 (THE CATHOLIC UNIVERSITY OF KOREA INDUSTRY-ACADEMIC COOPERATION FOUNDATION) [KR/KR]; 06591 서울특별시 서초구 반포대로 222, 가톨릭대학교 성의교정, Seoul (KR). 이화여자대학교 산학협력단 (EWHHA UNIVERSITY-INDUSTRY COLLABORATION FOUNDATION) [KR/KR]; 03760 서울특별시 서대문구 이화여대길 52, Seoul (KR). 동국대학교 산학협력단 (DONGGUK UNIVERSITY

INDUSTRY-ACADEMIC COOPERATION FOUNDATION) [KR/KR]; 04620 서울특별시 중구 필동로1길 30, Seoul (KR).

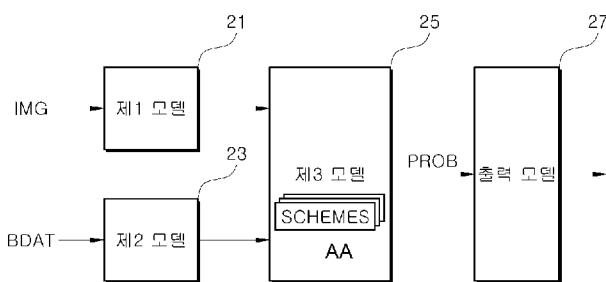
(72) 발명자: 박동진 (PARK, Dongjin); 06621 서울특별시 서초구 서초대로74길 29, 803호, Seoul (KR). 허희진 (HUH, Hee Jin); 10322 경기도 고양시 일산동구 위시티 4로 46, 213동 203호, Gyeonggi-do (KR). 박설희 (PARK, Sholhui); 07804 서울특별시 강서구 공항대로 260, 진단검사의학과, Seoul (KR). 백승민 (BAIK, Seung Min); 07981 서울특별시 양천구 목동서로70, 219동 506호, Seoul (KR).

(74) 대리인: 특허법인 비엘티 (BLT PATENT & LAW FIRM); 06595 서울특별시 서초구 법원로3길 25, 1층 (서초동, 태흥빌딩), Seoul (KR).

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD,

(54) Title: ELECTRONIC DEVICE TO WHICH MACHINE LEARNING ENSEMBLE MODEL FOR MALARIA DIAGNOSIS AND DETECTION IS APPLICABLE, AND OPERATION METHOD THEREOF

(54) 발명의 명칭: 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치 및 그 동작 방법



21 ... First model
23 ... Second model
25 ... Third model
27 ... Output model
AA ... SCHEMES
BB ... "Positive" or "negative"

(57) Abstract: An electronic device to which a machine learning ensemble model is applicable, according to an exemplary embodiment of the present disclosure, may comprise: a first model neural-network-trained on the basis of training images; a second model trained on the basis of training test data; a third model using voting based on a plurality of statistical methodologies so as to output the infection probability of malaria infection; and a processor, which inputs, into the first model, a blood image to be tested, uses the first model to make a primary determination of malaria infection about the blood image, inputs, into the second model, blood data to be tested, uses the second model to make a secondary determination of malaria infection about the blood data, and uses the third model to output the malaria infection probability.

(57) 요약서: 본 개시의 예시적 실시예에 따른 머신러닝 앙상블 모델이 적용가능한 전자 장치는, 학습용이미지에 기반하여 뉴럴네트워크 학습된 제1 모델; 학습용검사데이터에 기반하여 학습된 제2 모델; 복수의 통계적 방법론에 대한 투표(voting) 방식으로 말라리아 감염여부에 관한 감염확률을 출력하도록 구성된 제3 모델; 및 상기 제1 모델에 검사대상인 혈액이미지를 입력하고, 상기 제1 모델을 이용해 상기 혈액이미지에 대해 말라리아 감염여부를 1차 판정하고, 상기 제2 모델에 검사대상인 혈액데이터를 입력하고, 상기 제2 모델을 이용해 상기 혈액데이터에 대해 말라리아 감염여부를 2차 판정하며, 상기 제3 모델을 이용해 상기 말라리아 감염확률을 출력하는 프로세서를 포함하도록 구성될 수 있다.



MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO,
NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW,
SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의
역내 권리의 보호를 위하여): ARIPO (BW, CV, GH, GM,
KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ,
UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ,
TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC,
ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

발명의 설명

발명의 명칭: 말라리아 진단 및 탐지를 위한 머신러닝 양상블 모델이 적용가능한 전자 장치 및 그 동작 방법

기술분야

- [1] 본 개시는 말라리아 진단 및 탐지를 위한 머신러닝 양상블 모델이 적용가능한 전자 장치 및 그 동작 방법에 관한 것이다.

배경기술

- [2] 우리나라에서는 삼일열 말라리아 원충(*Plasmodium vivax*)에 의한 삼일열 말라리아가 자주 유행하고 있다. 말라리아 감염에 의한 후유증으로 비장 비대로 인한 파열, 중추신경계 이상증상인 기억상실, 경련, 정신 분열 같은 이상 행동 등이 발생하기도 하므로, 신속 정확하게 말라리아를 탐지하는 기술이 요구되는 실정이다.
- [3] 삼일열 말라리아 진단 및 탐지를 위해 특허출원 제10-2018-0071448호 등의 기술이 제안되고 있으나, 진단용 시약, 키트 등 전통적인 방식의 의학적 검출 방법에 의존하는 경우가 대부분이다.

발명의 내용

기술적 과제

- [4] 삼일열 말라리아 등 말라리아 질환의 정확한 진단 및 신속한 탐지를 위해 말라리아 환자의 말초혈액 슬라이드 사진을 기반으로 한 객체 탐지(object detection) 기반 CNN(convolutional neural network) 모델과 말라리아 진단을 위해 혈액검사 결과 및 임상정보를 기반으로 한 머신러닝 모델을 결합한 양상블 모델에 대한 장치 및 방법을 제공하는 데에 목적이 있다.

과제 해결 수단

- [5] 상술한 과제를 해결하기 위한 본 개시에 따른 머신러닝 양상블 모델이 적용가능한 전자 장치는, 학습용이미지에 기반하여 뉴럴네트워크 학습된 제1 모델; 학습용검사데이터에 기반하여 학습된 제2 모델; 복수의 통계적 방법론에 대한 투표(voting) 방식으로 말라리아 감염여부에 관한 감염확률을 출력하도록 구성된 제3 모델; 및 상기 제1 모델에 검사대상인 혈액이미지를 입력하고, 상기 제1 모델을 이용해 상기 혈액이미지에 대해 말라리아 감염여부를 1차 판정하고, 상기 제2 모델에 검사대상인 혈액데이터를 입력하고, 상기 제2 모델을 이용해 상기 혈액데이터에 대해 말라리아 감염여부를 2차 판정하며, 상기 제3 모델을 이용해 상기 말라리아 감염확률을 출력하는 프로세서를 포함할 수 있다.
- [6] 또한, 상술한 과제를 해결하기 위한 본 개시에 따른 전자 장치에 의해 수행되는 머신러닝 양상블 모델 기반 말라리아 진단 및 탐지 방법은, 학습용이미지에 기반하여 뉴럴네트워크 학습된 제1 모델에 검사대상인 혈액이미지를 입력하는 단계; 학습된 상기 제1 모델을 이용해 상기 혈액이미지에 대해 말라리아 감염여부

를 1차 판정하는 단계; 학습용검사데이터에 기반하여 학습된 제2 모델에 검사대상인 혈액데이터를 입력하는 단계; 학습된 상기 제2 모델을 이용해 상기 혈액데이터에 대해 말라리아 감염여부를 2차 판정하는 단계; 및 복수의 통계적 방법론에 대해 투표(voting) 방식을 적용하는 제3 모델을 이용해 말라리아 감염확률을 출력하는 단계를 포함할 수 있다.

[7] 이 외에도, 본 발명을 구현하기 위한 다른 방법, 다른 시스템 및 상기 방법을 실행하기 위한 컴퓨터 프로그램을 기록하는 컴퓨터 판독 가능한 기록 매체가 더 제공될 수 있다.

[8] 이 외에도, 본 발명을 구현하기 위한 방법이 컴퓨터 상에서 수행되도록 매체에 저장되는 컴퓨터 프로그램이 더 제공될 수 있다.

발명의 효과

[9] 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치 및 그 동작 방법에 따르면 말라리아 감염의 의심되는 경우 신속한 진단 및 정확한 탐지를 달성할 수 있다.

[10] 또한 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치 및 그 동작 방법에 따르면 딥러닝 기반 실시간 객체 인식 기술을 통해 적혈구의 개수에 따른 상대적인 말라리아 원충을 계수할 수 있으므로 말라리아 감염 정도의 정확한 보고가 가능하고, 딥러닝에서 정확히 탐지하지 못한 말라리아 원충 여부 탐지를 보완하기 위하여 혈액검사 결과를 바탕으로 한 머신러닝을 활용하여 보다 정확한 말라리아 원충을 탐지할 수 있다.

[11] 본 개시의 효과들은 이상에서 언급된 효과로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[12] 도 1은 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치를 도시하는 블록도이다.

[13] 도 2는 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델에 적용되는 인공신경망을 상세히 도시하는 블록도이다.

[14] 도 3은 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치의 성능을 나타내는 도면이다.

[15] 도 4는 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치의 성능을 나타내는 도면이다.

[16] 도 5는 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치의 성능을 나타내는 도면이다.

[17] 도 6a 내지 도 6c는 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델을 통한 타입별 객체 인식을 설명하는 그림이다.

[18] 도 7은 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델에 적용되는 인공신경망을 상세히 도시하는 블록도이다.

[19] 도 8은 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치를 도시하는 블록도이다.

발명의 실시를 위한 형태

[20] 본 개시 전체에 걸쳐 동일 참조 부호는 동일 구성요소를 지칭한다. 본 개시가 실시예들의 모든 요소들을 설명하는 것은 아니며, 본 개시가 속하는 기술분야에서 일반적인 내용 또는 실시예들 간에 중복되는 내용은 생략한다. 명세서에서 사용되는 '부, 모듈, 부재, 블록'이라는 용어는 소프트웨어 또는 하드웨어로 구현될 수 있으며, 실시예들에 따라 복수의 '부, 모듈, 부재, 블록'이 하나의 구성요소로 구현되거나, 하나의 '부, 모듈, 부재, 블록'이 복수의 구성요소들을 포함하는 것도 가능하다. 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 직접적으로 연결되어 있는 경우뿐 아니라, 간접적으로 연결되어 있는 경우를 포함하고, 간접적인 연결은 무선 통신망을 통해 연결되는 것을 포함한다.

[21] 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.

[22] 명세서 전체에서, 어떤 부재가 다른 부재 "상에" 위치하고 있다고 할 때, 이는 어떤 부재가 다른 부재에 접해 있는 경우뿐 아니라 두 부재 사이에 또 다른 부재가 존재하는 경우도 포함한다.

[23] 제 1, 제 2 등의 용어는 하나의 구성요소를 다른 구성요소로부터 구별하기 위해 사용되는 것으로, 구성요소가 전술된 용어들에 의해 제한되는 것은 아니다.

[24] 단수의 표현은 문맥상 명백하게 예외가 있지 않는 한, 복수의 표현을 포함한다.

[25] 각 단계들에 있어 식별부호는 설명의 편의를 위하여 사용되는 것으로 식별부호는 각 단계들의 순서를 설명하는 것이 아니며, 각 단계들은 문맥상 명백하게 특정 순서를 기재하지 않는 이상 명기된 순서와 다르게 실시될 수 있다.

[26] 이하 첨부된 도면들을 참고하여 본 개시의 작용 원리 및 실시예들에 대해 설명한다.

[27] 본 명세서에서 '본 개시에 따른 장치'는 연산처리를 수행하여 사용자에게 결과를 제공할 수 있는 다양한 장치들이 모두 포함된다. 예를 들어, 본 개시에 따른 장치는, 컴퓨터, 서버 장치 및 휴대용 단말기를 모두 포함하거나, 또는 어느 하나의 형태가 될 수 있다.

[28] 여기에서, 상기 컴퓨터는 예를 들어, 웹 브라우저(WEB Browser)가 탑재된 노트북, 데스크톱(desktop), 랩톱(laptop), 태블릿 PC, 슬레이트 PC 등을 포함할 수 있다.

- [29] 상기 서버 장치는 외부 장치와 통신을 수행하여 정보를 처리하는 서버로써, 애플리케이션 서버, 컴퓨팅 서버, 데이터베이스 서버, 파일 서버, 게임 서버, 메일 서버, 프록시 서버 및 웹 서버 등을 포함할 수 있다.
- [30] 상기 휴대용 단말기는 예를 들어, 휴대성과 이동성이 보장되는 무선 통신 장치로서, PCS(Personal Communication System), GSM(Global System for Mobile communications), PDC(Personal Digital Cellular), PHS(Personal Handyphone System), PD A(Personal Digital Assistant), IMT(International Mobile Telecommunication)-2000, C DMA(Code Division Multiple Access)-2000, W-CDMA(W-Code Division Multiple Access), WiBro(Wireless Broadband Internet) 단말, 스마트 폰(Smart Phone) 등과 같은 모든 종류의 핸드헬드(Handheld) 기반의 무선 통신 장치와 시계, 반지, 팔찌, 발찌, 목걸이, 안경, 콘택트 렌즈, 또는 머리 착용형 장치(head-mounted-device(HMD)) 등과 같은 웨어러블 장치를 포함할 수 있다.
- [31] 본 개시에 따른 인공지능과 관련된 기능은 프로세서와 메모리를 통해 동작된다. 프로세서는 하나 또는 복수의 프로세서로 구성될 수 있다. 이때, 하나 또는 복수의 프로세서는 CPU, AP, DSP(Digital Signal Processor) 등과 같은 범용 프로세서, GPU, VPU(Vision Processing Unit)와 같은 그래픽 전용 프로세서 또는 NPU와 같은 인공지능 전용 프로세서일 수 있다. 하나 또는 복수의 프로세서는, 메모리에 저장된 기 정의된 동작 규칙 또는 인공지능 모델에 따라, 입력 데이터를 처리하도록 제어한다. 또는, 하나 또는 복수의 프로세서가 인공지능 전용 프로세서인 경우, 인공지능 전용 프로세서는, 특정 인공지능 모델의 처리에 특화된 하드웨어 구조로 설계될 수 있다.
- [32] 본 개시에 따른 인공지능과 관련된 기능은 프로세서와 메모리를 통해 동작된다. 프로세서는 하나 또는 복수의 프로세서로 구성될 수 있다. 이때, 하나 또는 복수의 프로세서는 CPU, AP, DSP(Digital Signal Processor) 등과 같은 범용 프로세서, GPU, VPU(Vision Processing Unit)와 같은 그래픽 전용 프로세서 또는 NPU와 같은 인공지능 전용 프로세서일 수 있다. 하나 또는 복수의 프로세서는, 메모리에 저장된 기 정의된 동작 규칙 또는 인공지능 모델에 따라, 입력 데이터를 처리하도록 제어한다. 또는, 하나 또는 복수의 프로세서가 인공지능 전용 프로세서인 경우, 인공지능 전용 프로세서는, 특정 인공지능 모델의 처리에 특화된 하드웨어 구조로 설계될 수 있다.
- [33] 인공지능 모델은 학습을 통해 만들어진 것을 특징으로 한다. 여기서, 학습을 통해 만들어진다는 것은, 기본 인공지능 모델이 학습 알고리즘에 의하여 다수의 학습 데이터들을 이용하여 학습됨으로써, 원하는 특성(또는, 목적)을 수행하도록 설정된 기 정의된 동작 규칙 또는 인공지능 모델이 만들어짐을 의미한다. 이러한 학습은 본 개시에 따른 인공지능이 수행되는 기기 자체에서 이루어질 수도 있고, 별도의 서버 및/또는 시스템을 통해 이루어 질 수도 있다. 학습 알고리즘의 예로는, 지도형 학습(supervised learning), 비지도형 학습(unsupervised learning),

준지도형 학습(semi-supervised learning) 또는 강화 학습(reinforcement learning)이 있으나, 전술한 예에 한정되지 않는다.

- [34] 인공지능 모델은, 복수의 신경망 레이어들로 구성될 수 있다. 복수의 신경망 레이어들 각각은 복수의 가중치들 (weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행한다. 복수의 신경망 레이어들이 갖고 있는 복수의 가중치들은 인공지능 모델의 학습 결과에 의해 최적화될 수 있다. 예를 들어, 학습 과정 동안 인공지능 모델에서 획득한 로스(loss) 값 또는 코스트(cost) 값이 감소 또는 최소화되도록 복수의 가중치들이 갱신될 수 있다. 인공 신경망은 심층 신경망(DNN:Deep Neural Network)를 포함할 수 있으며, 예를 들어, CNN (Convolutional Neural Network), DNN (Deep Neural Network), RNN (Recurrent Neural Network), RBM (Restricted Boltzmann Machine), DBN (Deep Belief Network), BRDNN(Bidirectional Recurrent Deep Neural Network) 또는 심층 Q-네트워크 (Deep Q-Networks) 등이 있으나, 전술한 예에 한정되지 않는다.
- [35] 본 개시의 예시적인 실시예에 따르면, 프로세서는 인공지능을 구현할 수 있다. 인공지능이란 사람의 신경세포(biological neuron)를 모사하여 기계가 학습하도록 하는 인공신경망(Artificial Neural Network) 기반의 기계 학습법을 의미한다. 인공지능의 방법론에는 학습 방식에 따라 훈련데이터로서 입력데이터와 출력데이터가 같이 제공됨으로써 문제(입력데이터)의 해답(출력데이터)이 정해져 있는 지도학습(supervised learning), 및 출력데이터 없이 입력데이터만 제공되어 문제(입력데이터)의 해답(출력데이터)이 정해지지 않는 비지도학습(unsupervised learning), 및 현재의 상태(State)에서 어떤 행동(Action)을 취할 때마다 외부 환경에서 보상(Reward)이 주어지는데, 이러한 보상을 최대화하는 방향으로 학습을 진행하는 강화학습(reinforcement learning)으로 구분될 수 있다. 또한, 인공지능의 방법론은 학습 모델의 구조인 아키텍처에 따라 구분될 수도 있는데, 널리 이용되는 딥러닝 기술의 아키텍처는, 합성곱신경망(CNN; Convolutional Neural Network), 순환신경망(RNN; Recurrent Neural Network), 트랜스포머(Transformer), 생성적 대립 신경망(GAN; generative adversarial networks) 등으로 구분될 수 있다.
- [36] 본 장치와 시스템은 인공지능 모델을 포함할 수 있다. 인공지능 모델은 하나의 인공지능 모델일 수 있고, 복수의 인공지능 모델로 구현될 수도 있다. 인공지능 모델은 뉴럴 네트워크(또는 인공 신경망)로 구성될 수 있으며, 기계학습과 인지과학에서 생물학의 신경을 모방한 통계학적 학습 알고리즘을 포함할 수 있다. 뉴럴 네트워크는 시냅스의 결합으로 네트워크를 형성한 인공 뉴런(노드)이 학습을 통해 시냅스의 결합 세기를 변화시켜, 문제 해결 능력을 가지는 모델 전반을 의미할 수 있다. 뉴럴 네트워크의 뉴런은 가중치 또는 바이어스의 조합을 포함할 수 있다. 뉴럴 네트워크는 하나 이상의 뉴런 또는 노드로 구성된 하나 이상의 레이어(layer)를 포함할 수 있다. 예시적으로, 장치는 input layer, hidden layer, output layer를 포함할 수 있다. 장치를 구성하는 뉴럴 네트워크는 뉴런의 가중치를 학습

을 통해 변화시킴으로써 임의의 입력(input)으로부터 예측하고자 하는 결과(output)를 추론할 수 있다.

- [37] 프로세서는 뉴럴 네트워크를 생성하거나, 뉴럴 네트워크를 훈련(train, 또는 학습(learn)하거나, 수신되는 입력 데이터를 기초로 연산을 수행하고, 수행 결과를 기초로 정보 신호(information signal)를 생성하거나, 뉴럴 네트워크를 재훈련(retrain)할 수 있다. 뉴럴 네트워크의 모델들은 GoogleNet, AlexNet, VGG Network 등과 같은 CNN(Convolution Neural Network), R-CNN(Region with Convolution Neural Network), RPN(Region Proposal Network), RNN(Recurrent Neural Network), S-DNN(Stacking-based deep Neural Network), S-SDNN(State-Space Dynamic Neural Network), Deconvolution Network, DBN(Deep Belief Network), RBM(Restricted Boltzman Machine), Fully Convolutional Network, LSTM(Long Short-Term Memory) Network, Classification Network 등 다양한 종류의 모델들을 포함할 수 있으나 이에 제한되는 것은 아니다. 프로세서는 뉴럴 네트워크의 모델들에 따른 연산을 수행하기 위한 하나 이상의 프로세서를 포함할 수 있다. 예를 들어 뉴럴 네트워크는 심층 뉴럴 네트워크 (Deep Neural Network)를 포함할 수 있다.
- [38] 뉴럴 네트워크는 CNN(Convolutional Neural Network), RNN(Recurrent Neural Network), 퍼셉트론(perceptron), 다층 퍼셉트론(multilayer perceptron), FF(Feed Forward), RBF(Radial Basis Network), DFF(Deep Feed Forward), LSTM(Long Short Term Memory), GRU(Gated Recurrent Unit), AE(Auto Encoder), VAE(Variational Auto Encoder), DAE(Denoising Auto Encoder), SAE(Sparse Auto Encoder), MC(Markov Chain), HN(Hopfield Network), BM(Boltzmann Machine), RBM(Restricted Boltzmann Machine), DBN(Deep Belief Network), DCN(Deep Convolutional Network), DN(Deconvolutional Network), DCIGN(Deep Convolutional Inverse Graphics Network), GAN(Generative Adversarial Network), LSM(Liquid State Machine), ELM(Extreme Learning Machine), ESN(Echo State Network), DRN(Deep Residual Network), DNC(Differentiable Neural Computer), NTM(Neural Turing Machine), CN(Capsule Network), KN(Kohonen Network) 및 AN(Attention Network)를 포함할 수 있으나 이에 한정되는 것이 아닌 임의의 뉴럴 네트워크를 포함할 수 있음은 통상의 기술자가 이해할 것이다.
- [39] 본 개시의 예시적인 실시예에 따르면, 프로세서는 GoogleNet, AlexNet, VGG Network 등과 같은 CNN(Convolution Neural Network), R-CNN(Region with Convolution Neural Network), RPN(Region Proposal Network), RNN(Recurrent Neural Network), S-DNN(Stacking-based deep Neural Network), S-SDNN(State-Space Dynamic Neural Network), Deconvolution Network, DBN(Deep Belief Network), RBM(Restricted Boltzman Machine), Fully Convolutional Network, LSTM(Long Short-Term Memory) Network, Classification Network, Generative Modeling, eXplainable AI, Continual AI, Representation Learning, AI for Material Design, 자연어 처리를 위한 BERT, SP-BERT, MRC/QA, Text Analysis, Dialog System, GPT-3, GPT-4, 비전 처리를 위한 V

isual Analytics, Visual Understanding, Video Synthesis, ResNet 데이터 지능을 위한 Anomaly Detection, Prediction, Time-Series Forecasting, Optimization, Recommendation, Data Creation 등 다양한 인공지능 구조 및 알고리즘을 이용할 수 있으며, 이에 제한되지 않는다.

- [40] 이하, 첨부된 도면을 참조하여 본 개시의 실시예를 상세하게 설명한다.
- [41] 도 1은 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치(10)를 도시하는 블록도이다.
- [42] 도 1을 참고하면 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치(10)는 입/출력부(110), 통신부(120), 식별부(130), 데이터베이스(140) 및 프로세서(150)를 포함할 수 있다. 이하에서는, 전자 장치(10)는 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델을 구현하는 전자적 장치이며, 머신러닝 앙상블 모델이 적용된 말라리아 진단 및 탐지 방법은 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치(10)를 통해 구현되는 것으로 가정된다.
- [43] 입/출력부(110)는 사용자 입력을 받거나 또는 사용자에게 정보를 출력하는 각종 인터페이스나 연결 포트 등일 수 있다. 입/출력부(110)는 입력 모듈과 출력 모듈로 구분될 수 있다.
- [44] 입력 모듈은 사용자로부터 사용자 입력을 수신한다. 입력 모듈은 영상 정보(또는 신호), 오디오 정보(또는 신호), 데이터, 또는 사용자로부터 입력되는 정보의 입력을 위한 것으로서, 적어도 하나의 카메라, 적어도 하나의 마이크로폰 및 사용자 입력부 중 적어도 하나를 포함할 수 있다. 입력부에서 수집한 음성 데이터나 이미지 데이터는 분석되어 사용자의 제어명령으로 처리될 수 있다.
- [45] 사용자 입력은 키 입력, 터치 입력, 음성 입력을 비롯한 다양한 형태로 이루어질 수 있다. 이러한 사용자 입력을 받을 수 있는 입력 모듈의 예로는 전통적인 형태의 키패드나 키보드, 마우스는 물론, 사용자의 터치를 감지하는 터치 센서, 음성 신호를 입력 받는 마이크, 영상 인식을 통해 제스처 등을 인식하는 카메라, 사용자 접근을 감지하는 조도 센서나 적외선 센서 등으로 구성되는 근접 센서, 가속도 센서나 자이로 센서 등을 통해 사용자 동작을 인식하는 모션 센서 및 그 외의 다양한 형태의 사용자 입력을 감지하거나 입력 받는 다양한 형태의 입력 수단을 모두 포함하는 포괄적인 개념이다.
- [46] 여기서, 터치 센서는 디스플레이 패널에 부착되는 터치 패널이나 터치 필름을 통해 터치를 감지하는 압전식 또는 정전식 터치 센서, 광학적인 방식에 의해 터치를 감지하는 광학식 터치 센서 등으로 구현될 수 있다. 이외에도 입력 모듈은 자체적으로 사용자 입력을 감지하는 장치 대신 사용자 입력을 입력 받는 외부의 입력 장치를 연결시키는 입력 인터페이스(USB 포트, PS/2 포트 등)의 형태로 구현될 수도 있다.
- [47] 출력 모듈은 각종 정보를 출력해 사용자에게 이를 제공할 수 있다. 출력 모듈은 영상을 출력하는 디스플레이, 소리를 출력하는 스피커(및/또는 이와 연결된 증폭

기(amplifier)), 진동을 발생시키는 햅틱 장치 및 그 외의 다양한 형태의 출력 수단을 모두 포함하는 포괄적인 개념이다. 이외에도 출력 모듈은 상술한 개별 출력 수단을 연결시키는 포트 타입의 출력 인터페이스의 형태로 구현될 수도 있다.

[48] 일 예로, 디스플레이 형태의 출력 모듈은 텍스트, 정지 영상, 동영상을 디스플레이 할 수 있다. 디스플레이는 액정 디스플레이(LCD: Liquid Crystal Display), 발광 다이오드(LED: light emitting diode) 디스플레이, 유기 발광 다이오드(OLED: Organic Light Emitting Diode) 디스플레이, 평판 디스플레이(FPD: Flat Panel Display), 투명 디스플레이(transparent display), 곡면 디스플레이(Curved Display), 플렉시블 디스플레이(flexible display), 3차원 디스플레이(3D display), 홀로그래픽 디스플레이(holographic display), 프로젝터 및 그 외의 영상 출력 기능을 수행할 수 있는 다양한 형태의 장치를 모두 포함하는 광의의 영상 표시 장치를 의미하는 개념이다. 이러한 디스플레이는 입력 모듈의 터치 센서와 일체로 구성된 터치 디스플레이의 형태일 수도 있다.

[49] 통신부(120)는 외부 기기와 통신할 수 있다. 따라서, 장치(디바이스)는 통신부를 통해 외부 기기와 정보를 송수신할 수 있다. 예를 들어, 장치는 통신부를 이용해 전기차 충전관리 시스템 내에 저장 및 생성된 정보들이 공유되도록 외부 기기와 통신을 수행할 수 있다. 통신부(120)는 예를 들어, 유선통신 모듈, 무선통신 모듈, 근거리 통신 모듈, 위치정보 모듈 중 적어도 하나를 포함할 수 있다.

[50] 여기서, 통신, 즉 데이터의 송수신은 유선 또는 무선으로 이루어질 수 있다. 이를 위해 통신부는 LAN(Local Area Network)를 통해 인터넷 등에 접속하는 유선 통신 모듈, 이동 통신 기지국을 거쳐 이동 통신 네트워크에 접속하여 데이터를 송수신하는 이동 통신 모듈, 와이파이(Wi-Fi) 같은 WLAN(Wireless Local Area Network) 계열의 통신 방식이나 블루투스(Bluetooth), 직비(Zigbee)와 같은 WPAN(Wireless Personal Area Network) 계열의 통신 방식을 이용하는 근거리 통신 모듈, GPS(Global Positioning System)와 같은 GNSS(Global Navigation Satellite System)를 이용하는 위성 통신 모듈 또는 이들의 조합으로 구성될 수 있다. 통신에 사용되는 무선 통신 기술은 저전력 통신을 위한 NB-IoT(Narrowband Internet of Things)를 포함할 수 있다. 이때, 예를 들어 NB-IoT 기술은 LPWAN(Low Power Wide Area Network) 기술의 일례일 수 있고, LTE Cat(category) NB1 및/또는 LTE Cat NB2 등의 규격으로 구현될 수 있으며, 상술한 명칭에 한정되는 것은 아니다. 추가적으로 또는 대체적으로, 다양한 실시예들에 따른 무선 기기에서 구현되는 무선 통신 기술은 LTE-M 기술을 기반으로 통신을 수행할 수 있다. 이때, 일 예로, LTE-M 기술은 LPWAN 기술의 일례일 수 있고, eMTC(enhanced Machine Type Communication) 등의 다양한 명칭으로 불릴 수 있다. 예를 들어, LTE-M 기술은 1) LTE CAT 0, 2) LTE Cat M1, 3) LTE Cat M2, 4) LTE non-BL(non-Bandwidth Limited), 5) LTE-MTC, 6) LTE Machine Type Communication, 및/또는 7) LTE M 등의 다양한 규격 중 적어도 어느 하나로 구현될 수 있으며 상술한 명칭에 한정되는 것은 아니다. 추가적으로 또는 대체적으로, 다양한 실시예들에 따른 무선 기기에서 구현되

는 무선 통신 기술은 저전력 통신을 고려한 지그비(ZigBee), 블루투스(Bluetooth) 및 저전력 광역 통신망(Low Power Wide Area Network, LPWAN) 중 적어도 어느 하나를 포함할 수 있으며, 상술한 명칭에 한정되는 것은 아니다. 일 예로 ZigBee 기술은 IEEE 802.15.4 등의 다양한 규격을 기반으로 소형/저-파워 디지털 통신에 관련된 PAN(personal area networks)을 생성할 수 있으며, 다양한 명칭으로 불릴 수 있다.

- [51] 유선 통신 모듈은, 지역 통신(Local Area Network; LAN) 모듈, 광역 통신(Wide Area Network; WAN) 모듈 또는 부가가치 통신(Value Added Network; VAN) 모듈 등 다양한 유선 통신 모듈뿐만 아니라, USB(Universal Serial Bus), HDMI(High Definition Multimedia Interface), DVI(Digital Visual Interface), RS-232(recommended standard 232), 전력선 통신, 또는 POTS(plain old telephone service) 등 다양한 케이블 통신 모듈을 포함할 수 있다.
- [52] 무선 통신 모듈은 와이파이(Wifi) 모듈, 와이프로(Wireless broadband) 모듈 외에도, GSM(global System for Mobile Communication), CDMA(Code Division Multiple Access), WCDMA(Wideband Code Division Multiple Access), UMTS(universal mobile telecommunications system), TDMA(Time Division Multiple Access), LTE(Long Term Evolution), 4G, 5G, 6G 등 다양한 무선 통신 방식을 지원하는 무선 통신 모듈을 포함할 수 있다.
- [53] 무선 통신 모듈은 신호를 송신하는 안테나 및 송신기(Transmitter)를 포함하는 무선 통신 인터페이스를 포함할 수 있다. 또한, 무선 통신 모듈은 제어부의 제어에 따라 무선 통신 인터페이스를 통해 제어부로부터 출력된 디지털 제어 신호를 아날로그 형태의 무선 신호로 변조하는 신호 변환 모듈을 더 포함할 수 있다.
- [54] 무선 통신 모듈은 신호를 수신하는 안테나 및 수신기(Receiver)를 포함하는 무선 통신 인터페이스를 포함할 수 있다. 또한, 무선 통신 모듈은 무선 통신 인터페이스를 통하여 수신한 아날로그 형태의 무선 신호를 디지털 제어 신호로 복조하기 위한 신호 변환 모듈을 더 포함할 수 있다.
- [55] 근거리 통신 모듈은 근거리 통신(Short range communication)을 위한 것으로서, 블루투스(Bluetooth™), RFID(Radio Frequency Identification), 적외선 통신(Infrared Data Association; IrDA), UWB(Ultra Wideband), ZigBee, NFC(Near Field Communication), Wi-Fi(Wireless-Fidelity), Wi-Fi Direct, Wireless USB(Wireless Universal Serial Bus) 기술 중 적어도 하나를 이용하여, 근거리 통신을 지원할 수 있다.
- [56] 위치정보 모듈은 본 개시에 따른 전자 장치(10)의 위치(또는 현재 위치)를 획득하기 위한 모듈로서, 그의 대표적인 예로는 GPS(Global Positioning System) 모듈 또는 WiFi(Wireless Fidelity) 모듈이 있다. 예를 들어, GPS모듈을 활용하면, GPS 위성에서 보내는 신호를 이용하여 본 전자 장치의 위치를 획득할 수 있다. 다른 예로서, Wi-Fi모듈을 활용하면, Wi-Fi모듈과 무선신호를 송신 또는 수신하는 무선 AP(Wireless Access Point)의 정보에 기반하여, 본 전자 장치(10)의 위치를 획득할 수 있다. 필요에 따라서, 위치정보모듈은 치환 또는 부가적으로 본 장치의 위

치에 관한 데이터를 얻기 위해 통신부의 다른 모듈 중 어느 기능을 수행할 수 있다. 위치정보모듈은 본 장치의 위치(또는 현재 위치)를 획득하기 위해 이용되는 모듈로, 본 장치의 위치를 직접적으로 계산하거나 획득하는 모듈로 한정되지는 않는다.

- [57] 식별부(130)는 영상 인식을 통해 오브젝트 등을 인식하는 카메라, 오브젝트 접근을 감지하는 감지 센서, 사용자의 입력에 따른 터치 센서 및 그 외의 다양한 형태의 외부 입력을 감지하거나 입력 받는 다양한 형태의 식별/센싱 수단을 모두 포함하는 포괄적인 개념일 수 있다.
- [58] 카메라는 촬영 모드에서 이미지 센서에 의해 얻어지는 정지영상 또는 동영상 등의 화상 프레임을 처리한다. 처리된 화상 프레임은 디스플레이부(또는 본 개시의 전자 장치의 화면)에 표시되거나 메모리에 저장될 수 있다.
- [59] 한편, 상기 카메라가 복수개일 경우, 매트릭스 구조를 이루도록 배치될 수 있으며, 이와 같이 매트릭스 구조를 이루는 카메라들을 통해 다양한 각도 또는 초점을 갖는 복수의 영상정보가 입력될 수 있고, 또한 상기 카메라들은 3차원의 입체 영상을 구현하기 위한 좌 영상 및 우 영상을 획득하도록, 스테레오 구조로 배치될 수도 있다.
- [60] 식별부는 입/출력부(110) 내의 입력 모듈과 동일한 것으로 이해될 수 있거나 및/또는 입력 모듈과는 별도의 것으로 이해될 수도 있다. 식별부(130)는 지자기 센서(Magnetic sensor), 가속도 센서(Acceleration sensor), 온/습도 센서, 적외선 센서, 자이로스코프 센서, 위치 센서(예컨대, GPS), 기압 센서, 근접 센서, RGB 센서(illuminance sensor), 라이다(radar) 센서, 조도 센서, 및 전류 센서 중 하나 이상을 더 포함할 수 있으나, 이에 한정되는 것은 아니다. 각 센서들의 기능은 그 명칭으로부터 당업자가 직관적으로 추론할 수 있으므로, 구체적인 설명은 생략하기로 한다.
- [61] 데이터베이스(140)는 각종 정보를 저장할 수 있다. 데이터베이스는 데이터를 임시적으로 또는 반영구적으로 저장할 수 있다. 예를 들어, 데이터베이스에는 본 디바이스를 구동하기 위한 운영 프로그램(OS: Operating System), 웹 사이트를 호스팅하기 위한 데이터나 점자 생성을 위한 프로그램 내지는 어플리케이션(예를 들어, 웹 어플리케이션)에 관한 데이터 등이 저장될 수 있다. 또, 데이터베이스는 상술한 바와 같이 모듈들을 컴퓨터 코드 형태로 저장할 수 있다.
- [62] 데이터베이스(140)의 예로는 하드 디스크(HDD: Hard Disk Drive), SSD(Solid State Drive), 플래쉬 메모리(flash memory), 롬(ROM: Read-Only Memory), 램(RAM: Random Access Memory) 등이 있을 수 있다. 이러한 데이터베이스는 내장 타입 또는 탈부착 가능한 타입으로 제공될 수 있다.
- [63] 프로세서(150)는 전자 장치(10)의 전반적인 동작을 제어한다. 이를 위해 프로세서(150)는 각종 정보의 연산 및 처리를 수행하고 본 디바이스의 구성요소들의 동작을 제어할 수 있다. 예를 들어, 프로세서(150)는 이하에서 설명되는 제1 모델, 제2 모델 및 제3 모델을 통해 말라리아 감염이 의심되는 경우 신속한 진단 및 정확한 탐지를 달성할 수 있다.

- [64] 프로세서(150)는 하드웨어 소프트웨어 또는 이들의 조합에 따라 컴퓨터나 이와 유사한 장치로 구현될 수 있다. 하드웨어적으로 프로세서(150)는 전기적인 신호를 처리하여 제어 기능을 수행하는 전자 회로 형태로 제공될 수 있으며, 소프트웨어적으로는 하드웨어적인 프로세서를 구동시키는 프로그램 형태로 제공될 수 있다. 한편, 이하의 설명에서 특별한 언급이 없는 경우에는 본 전자 장치(10)의 동작은 프로세서(150)의 제어에 의해 수행되는 것으로 해석될 수 있다.
- [65] 프로세서(150)는 본 장치 내의 구성요소들의 동작을 제어하기 위한 알고리즘 또는 알고리즘을 재현한 프로그램에 대한 데이터를 저장하는 메모리, 및 메모리에 저장된 데이터를 이용하여 전술한 동작을 수행하는 적어도 하나의 세부프로세서(미도시)로 구현될 수 있다. 이때, 메모리와 프로세서는 각각 별개의 칩으로 구현될 수 있다. 또는, 메모리와 프로세서는 단일 칩으로 구현될 수도 있다.
- [66] 또한, 프로세서(150)는 이하의 도 2 내지 도 8에서 설명될 본 개시에 따른 다양한 실시 예들을 본 장치 상에서 구현하기 위하여, 위에서 살펴본 구성요소들을 중 어느 하나 또는 복수를 조합하여 제어할 수 있다.
- [67] 우리나라에서는 삼일열 말라리아 원충(*Plasmodium vivax*)에 의한 삼일열 말라리아가 자주 유행하고 있다. 말라리아 감염에 의한 후유증으로 비장 비대로 인한 파열, 중추신경계 이상증상인 기억상실, 경련, 정신 분열 같은 이상 행동 등이 발생하기도 하므로, 말라리아 감염의 신속한 진단 및 정확한 탐지는 치료 및 환자의 예후에 절대적 영향을 미치는 매우 중요한 요소이다.
- [68] 환자의 말라리아 감염 여부를 진단하기 위한 다양한 진단방법이 있으나 일반적으로는 말라리아 원충이 환자의 말초 혈액에서 관찰되는지 여부에 따라 결정될 수 있다. 따라서 본 개시는 환자의 말초 혈액 이미지로부터 말라리아 원충을 탐지하기 위해 객체 인식(object detection) 기술, 및 딥러닝(deep learning) 중 이미지 처리에 특화되어 있는 CNN(Convolutional Neural Network)을 접목하는 방법을 제안한다. 또한, 본 개시는 이미지 데이터 뿐만 아니라, 말라리아 진단을 위한 혈액 검사 결과데이터를 기계학습(machine learning; 머신러닝)을 이용해 보조/보완적으로 활용함으로써 보다 성능이 좋은 말라리아 감염 탐지 방법을 제안한다.
- [69] 결과적으로, 말라리아 진단을 위해 이미지 기반의 뉴럴네트워크 모델, 및 데이터 머신러닝을 접목하는 앙상블 모델(ensemble model)이 구축될 수 있다.
- [70] 본 개시의 예시적인 실시예에 따르면, 머신러닝 앙상블 모델이 적용가능한 전자 장치(10)는 적어도 하나의 뉴럴네트워크 모델을 포함하는 복수의 학습 및/또는 예측모델을 이용하여 구성될 수 있다.
- [71] 예를 들어, 전자 장치(10)는 학습용이미지에 기반하여 학습된 제1 모델, 학습용 검사데이터에 기반하여 학습된 제2 모델을 포함할 수 있다.
- [72] 예를 들어, 전자 장치(10)는 이미지 모델 및 데이터 모델의 결과 중 어느 하나를 선택하거나 조합가능한 제3 모델을 이용할 수 있다. 복수의 데이터 결과 중 어느 하나를 선택하거나 조합가능한 모델은 앙상블(ensemble) 모델로도 지칭될 수 있다. 예를 들어, 제3 모델은 복수의 통계적 방법론에 대한 투표(voting) 방식으로

말라리아 감염여부에 관한 감염확률을 출력하도록 구성될 수 있다. 다만 이러한 예시의 복수의 모델들이 제한 해석되지는 않는다. 복수의 학습 및/또는 예측모델에 관하여는 도 2 및 도 7을 통해 보다 상세히 설명될 것이다.

- [73] 상기와 같은 전자 장치(10)의 프로세서(150)는 상기 제1 모델을 이용해 상기 혈액이미지에 대해 말라리아 감염여부를 1차 판정하고, 상기 제2 모델에 검사대상인 혈액데이터를 입력하고, 상기 제2 모델을 이용해 상기 혈액데이터에 대해 말라리아 감염여부를 2차 판정하며, 상기 제3 모델을 이용해 상기 말라리아 감염확률을 출력할 수 있다.
- [74] 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치 및 그 동작 방법에 따르면 말라리아 감염의 의심되는 경우 신속한 진단 및 정확한 탐지를 달성할 수 있다.
- [75] 또한 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치 및 그 동작 방법에 따르면 딥러닝 기반 실시간 객체 인식 기술을 통해 적혈구의 개수에 따른 상대적인 말라리아 원충을 계수할 수 있으므로 말라리아 감염 정도의 정확한 보고가 가능하고, 딥러닝에서 정확히 탐지하지 못한 말라리아 원충 여부 탐지를 보완하기 위하여 혈액검사 결과를 바탕으로 한 머신러닝을 활용하여 보다 정확한 말라리아 원충을 탐지할 수 있다.
- [76] 도 2는 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델에 적용되는 인공신경망을 상세히 도시하는 블록도이다.
- [77] 도 2를 참조하면, 본 개시의 예시적 실시예에 따른 머신러닝 앙상블 모델이 적용가능한 전자 장치(20)는 제1 모델(21), 제2 모델(23), 제3 모델(25), 출력 모델(27)을 포함할 수 있다. 그러나 이는 예시적 실시예에 불과하며, 적어도 하나의 모델이 생략되거나, 다른 모델들이 더 포함될 수 있음을 배제하지 않는다.
- [78] 본 개시의 예시적 실시예에 따른 전자 장치(20)에 의해 수행되는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법은, 학습용이미지에 기반하여 뉴럴네트워크 학습된 제1 모델에 검사대상인 혈액이미지를 입력하는 단계, 학습된 제1 모델을 이용해 혈액이미지에 대해 말라리아 감염여부를 1차 판정하는 단계, 학습용검사데이터에 기반하여 뉴럴네트워크 학습된 제2 모델에 검사대상인 혈액데이터를 입력하는 단계, 학습된 제2 모델을 이용해 혈액데이터에 대해 말라리아 감염여부를 2차 판정하는 단계, 및 복수의 통계적 방법론에 대해 투표(voting) 방식을 적용하는 제3 모델을 이용해 말라리아 감염확률을 출력하는 단계를 포함할 수 있다.
- [79] 본 개시의 예시적 실시예에 따르면, 제1 모델(21)은 학습용이미지에 기반해 학습될 수 있다. 학습용이미지는 말라리아 감염환자의 말초혈액을 분절(slide)하여 구성된 슬라이드 이미지일 수 있다.
- [80] 본 개시의 예시적 실시예에 따르면, 학습용이미지는 말라리아 양성 증례 168건, 음성 증례 120건 총 288건을 학습에 이용하였으며, 각 개별 증례에 대해서 40

장의 현미경 슬라이드 영상을 사용하여 전체 총 11,520 장의 현미경 슬라이드 이미지로 학습데이터로서 이용되었다

- [81] 본 개시의 예시적 실시예에 따르면, 제1 모델(21)의 학습에는 말라리아 개별 원충에 대한 이미지데이터가 이용될 수 있다. 각 이미지 데이터에는 참값(ground truth)에 대한 라벨링이 부여될 수 있다. 예를 들어, 제1 모델(21)은 원형(ring), 영양형(trophozoite), 생식모세포(gametocyte), 열충(schizont) 및 백혈구, 적혈구의 타입별 말라리아 개별 원충에 대한 라벨링데이터를 전처리된 학습데이터로 하여 학습될 수 있다. 본 개시의 예시적 실시예에 따른 전자 장치(20)에 의해 수행되는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법에 있어서, 제1 모델(21)은, 백혈구, 적혈구 및 원형(ring), 영양형(trophozoite), 생식모세포(gametocyte), 및 열충(schizont)를 포함하는 말라리아 개별 원충의 라벨링이 수행된 결과를 이용할 수 있다.
- [82] 본 개시의 예시적 실시예에 따르면, 예를 들어, 제1 모델(21)은 모델의 성능 향상을 위해 데이터를 증강(augmentation)할 수 있다. 예컨대, 제1 모델(21)은 수평접기(horizontal flip) 방법론, 밝기 범위(brightness range) 방법론 중 적어도 하나를 이용해 데이터 증강(data augmentation)이 수행된 결과를 이용함으로써 라벨링된 데이터의 양을 증강시켜 학습 및 추론값의 향상을 도모할 수 있다.
- [83] 본 개시의 예시적 실시예에 따르면, 제1 모델(21)은 전처리된 데이터들, 라벨링된 데이터들, 증강된 데이터들을 기반으로 학습대상 이미지사이즈 및 배치(batch) 사이즈를 확정 후 인공지능망 기반으로 학습될 수 있다. 본 개시의 예시적 실시예에 따르면, 제1 모델에는 CNN(Convolutional Neural Network) 중 Yolo-v5가 이용되었다. 구체적으로, 전체 학습용이미지는 전체 증례를 8:2의 비율로 학습용 데이터셋과 테스트용 데이터셋으로 분리되었으며, 학습용 데이터셋의 20%를 Yolo-v5의 검증용 데이터로 활용하여 모델 성능이 검증되었다. 모델의 성능에 관하여는 도 4 내지 도 5를 통해 보다 상세히 설명될 것이다.
- [84] 본 개시의 예시적 실시예에 따르면, 제1 모델(21)은 진단대상인 혈액이미지(IMG)에 포함된 적혈구에 대한 제1 계수치에 기반해 말라리아 원충에 대한 제2 계수치를 산출함으로써 말라리아 감염여부 예측결과를 출력할 수 있다. 여기에서, 제1 계수치는 제2 계수치와 통계적 상관관계(correlation)를 가질 수 있다.
- [85] 본 개시의 예시적 실시예에 따르면, 제1 모델(21)은 학습된 제1 뉴럴 네트워크 모델을 이용해 검사대상인 혈액이미지에 포함된 적혈구에 대한 제1 계수치를 출력하는 단계, 제1 계수치에 기반해 말라리아 원충에 대한 제2 계수치를 출력하는 단계, 및 제2 계수치에 기반해 말라리아 감염여부 예측결과를 출력함으로써, 혈액이미지에 대해 말라리아 감염여부를 1차 판정할 수 있다.
- [86] 본 개시의 예시적 실시예에 따르면, 제2 모델(23)은 학습용검사데이터에 기반해 학습될 수 있다. 본 개시의 예시적 실시예에 따르면, 제2 모델(23)의 기계학습을 위해 288 증례의 혈액검사데이터에서 백혈구, 호중구(%), 림프구(%), 단핵구(%), 호산구(%), 호염기구(%), 혈색소, 적혈구, 혈소판 수치(Plt), C-반응성 단백질

(CRP), 알라닌아미노전이효소, 아스파테이트아미노전이효소, 총 빌리루빈, LDH, 나이 및 성별을 포함한 다양한 정보가 수집될 수 있다. 여기에서, 총 데이터셋에서 결측치는 422 (8.62%, 422/4896)개로 최종적으로 4474개의 데이터가 획득되었다.

- [87] 본 개시의 예시적 실시예에 따르면, 제2 모델(23)은 각 피쳐별 상이한 중요도를 가지고 학습하도록 구성되고, 결측치를 대체(replace)하지 않아도 모델링 가능한 적어도 하나의 알고리즘을 채택할 수 있다. 제2 모델(23)은 학습된 제2 머신러닝 모델을 이용해 혈액데이터에 대해 말라리아 감염여부를 2차 판정하는 단계의 일환으로서 결측치를 대체(replace)하지 않아도 모델링 가능한 적어도 하나의 알고리즘을 채택하는 단계를 수행할 수 있다.
- [88] 예컨대, 제2 모델(23)은 XGBoost(eXtreme Gradient Boosting) 모델, LGBM(Light Gradient Boosting Model) 모델 중 적어도 하나를 사용할 수 있고, 이에 국한되지 않는다. eXtreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LGBM), CNN 등 결측치를 별도로 전 처리하지 않고 빈칸처리(blank)하여도 모델링이 용이한 모델들이 제1 및/또는 제2 모델의 학습에 이용될 수 있다.
- [89] 본 개시의 예시적 실시예에 따르면, 제2 모델(23)의 학습에는 말라리아 양성 168건, 말라리아 음성 120건 총 288 증례를 8:2의 비율로 학습용 데이터셋 및 테스트용 데이터셋으로 분리하고, 학습용 데이터 셋은 9:1의 비율로 학습 및 검증을 10회 진행하는 방법이 수행되었다(stratified K-fold validation). 학습완료된 제2 모델(23)은 별도로 분류된 20%의 테스트용 데이터셋을 이용해 성능이 검증되었다. 성능예측에 영향을 미친 파라미터는 특징중요도(feature importance) 기술을 이용해 분석되었다.
- [90] 본 개시의 예시적 실시예에 따르면, 제2 모델(23)은 혈액데이터(BDAT)에서 CRP(C Reactive Protein), Plt(Platelet)를 포함하는 적어도 하나의 세부혈액데이터가 결손되었다고 판단되는 경우 보충을 요구할 수 있다. 여기에서 CRP(C Reactive Protein)는 C-반응 단백질로 염증이 체내에 발생할 경우 간에서 합성되어 혈액 상에 용해되는 단백질의 일종이며, Plt는 혈소판수치이다. 실험데이터의 각종 약어 및 수치에 관하여는 도 4를 통해 보다 상세히 설명될 것이다.
- [91] 본 개시의 예시적 실시예에 따르면, 제3 모델(25)은 복수의 통계적 방법론에 대한 투표(voting) 방식으로 말라리아 감염여부에 관한 감염확률(PROB)을 출력하도록 구성될 수 있다.
- [92] 본 개시의 예시적 실시예에 따르면, 제3 모델(25)은 딥러닝 기반의 제1 모델(21) 및 머신러닝 기반의 제2 모델(23)에서 나온 말라리아 진단 결과를 상호 보완하는 기능 모듈로 동작할 수 있다. 예를 들어, 제3 모델(35)은 말라리아 검출 평가기준인 AUROC, accuracy, precision, recall, F1 score를 향상시키는데 기여를 할 수 있다.
- [93] 본 개시의 예시적 실시예에 따르면, 제3 모델(25)은 상기 제1 모델의 예측결과 중 진단기준치에 미달하되 임계범위 내에 있는 의심결과값에 대해 재판단함으

로써 말라리아 감염여부 검증결과를 출력할 수 있다. 예를 들어, 제3 모델(25)은 제1 모델(21)에서 스코어링 점수가 “0.8”에는 달하지 못하여 “양성” 판정이 되지 않는 않았으나, 임계범위로서 오차범위인 -10%~+10% 범위 내의 수치인 “0.75”를 가지는 객체를 의심결과값을 설정하고, 이에 대해 제2 모델(23)의 혈액 실험데이터를 기반으로 재판단해, 말라리아 감염여부를 검증할 수 있다. 제3 모델(23)은 제1 모델(21)에서 “음성” 판정된 의심 수치를 제2 모델(23)을 이용해 검증해 “양성”판정하여 말라리아 감염 판정의 정확도를 향상시킬 수 있다.

- [94] 본 개시의 예시적 실시예에 따르면,출력 모델(27)은 감염확률이 임계값을 초과하는지 여부에 따라 말라리아 감염을 의미하는 “양성”클래스 또는 말라리아 비감염을 의미하는 “음성”클래스를 출력하도록 구성될 수 있다. 여기에서, “음성”클래스는, 제2 모델(23)의 출력값에 비해 제1 모델(21)의 출력값에 더 큰 의존도(dependency)를 가지며 결정될 수 있다. 또한, “양성”클래스는, 상기 복수의 통계적 방법론 각각을 조합하는 앙상블 모델에 대한 투표 결과로부터 추론될 수 있다.
- [95] 본 개시의 예시적 실시예에 따른 전자 장치(20)에 의해 수행되는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법에 있어서, 복수의 통계적 방법론에 대해 투표(voting) 방식을 적용하는 제3 모델을 이용해 말라리아 감염확률을 출력하는 단계는, 감염확률이 임계값 초과인 경우 말라리아 감염을 의미하는 “양성”클래스를 출력하는 단계, 및 감염확률이 임계값 미만인 경우 말라리아 비감염을 의미하는 “음성”클래스를 출력하는 단계를 포함할 수 있다.
- [96] 본 개시의 예시적 실시예에 따르면, 출력 모델(27)은 말라리아 원충이 하나라도 검출되었다고 판단하는 경우, 말라리아 감염 여부를 “양성”이라고 판단할 수 있다. 출력 모델(27)은 말라리아 원충이 하나도 검출되지 않았더라도, 이미지데이터(IMG) 및 혈액데이터(BDAT) 중 어느하나 또는 이의 조합으로부터 말라리아 감염 여부에 대한 판정을 실시할 수 있다.
- [97] 도 3는 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치(도 1, 10)의 성능을 나타내는 도면이다. 도 3을 참조하면, 혈액검사 결과 및 임상정보를 기반으로 한 머신 러닝 모델인 제 2 모델(23)의 성능이 도시된다.
- [98] 도 3에는 말라리아 검출 평가치인 AUROC, accuracy, precision, recall, F1 score 값, 정확도, 전체평균, 가중평균 등의 기준에 따라 도시되어 있다. 도 3의 데이터를 간략하게 표로 요약하면 다음 표 1과 같다.

[99] [표1]

Model	AUROC	Accuracy	Precision	Recall	F1 score
XGBoost	0.9994	0.9828	0.9857	0.9792	0.9821
LGBM	1.0	0.9828	0.9800	0.9853	0.9823

[100] 표 1은 혈액검사 결과 및 임상정보를 기반으로 한 머신 러닝 모델 성능을 나타낸다. 제2 모델(23)에서 XGBoost는 0.9828, LGBM은 0.9828의 정확도를 보임을 확인할 수 있다. 한편, 제1 모델(21)의 성능은 다음 표 2와 같다.

[101] [표2]

Model	Accuracy	Precision	Recall	F1 score
Yolo v5	0.879	0.829	1.0	0.907

[102] 표 2를 참조하면, 제1 모델(21)은 말초혈액 슬라이드 사진을 기반으로 한 YOLO v5 모델의 말라리아 분류 성능을 나타낸다. 제1 모델(21)의 검증에는 58건의 중례가 이용되었다. 학습된 제1 모델(21)은 총 58명 중 41명에 있어서는 양성 또는 음성을 정확히 진단할 수 있었고, 음성환자 24명에서 7명을 양성으로 진단하였다.

[103] 이 경우 정확도(accuracy)는 $87.9\% (= (34+17)/58)$ 이다. 정밀도(Precision)는 $82.9\% (= 34/(34+7))$ 이다. 예측도는 positive predicted value로서, 양성으로 예측한 것이 실제 양성인 것을 의미한다.

[104] 재현율(recall)은 민감도(sensitivity)로서, 분모는 (true positive + false negative; 실제 양성), 분자는 (true positive)로 구성되는데 이 경우 $100\% (= 34/34)$ 이다.

[105] F1 score는 정밀도와 재현율의 조화평균으로 $0.907 (= 2 \times (0.829 \times 1) / (0.829+1))$ 가 스코어링되었다.

[106] 도 4는 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치(도 1, 10)의 성능을 나타내는 도면이다. 학습완료된 제2 모델(23)은 별도로 분류된 20%의 테스트용 데이터셋을 이용해 성능이 검증되었는데, 성능예측에 영향을 미친 파라미터는 특징중요도(feature importance) 기술을 이용해 분석될 수 있음은 전술한 바와 같다.

[107] "특징 중요도(feature importance)"는 앙상블 모델에서 변수를 해석하기 위한 기술적 방법론(technique)으로, 모델 해석시 변수의 중요도를 추출하기 위한 방법의 일종이다.

[108] 도 4는 구체적으로 XGboost의 특징 중요도값을 나타낸다. 혈액검사데이터에서 백혈구, 호중구(%), 림프구(%), 단핵구(%), 호산구(%), 호염기구(%), 혈색소, 적혈구, 혈소판 수치(Plt), C-반응성 단백질(CRP), 알라닌아미노전이효소, 아스파테이트아미노전이효소, 총 빌리루빈, LDH, 나이 및 성별이 도시되어 있는데, 이 중 수치가 높은 파라미터가 제2 모델(23)의 구현 및 해석에 중대한 영향을 미치는 것이다.

[109] 도 4에 따르면, CRP(C Reactive Protein)의 특징 중요도는 0.289505, Plt(Platelet)의 특징 중요도는 0.203164로, 다른 변수들에 비해 상대적으로 중요성이 큰 것을 확인할 수 있다. 따라서, 제2 모델(23)은 CRP(C Reaction Protein), Plt(Platelet)를 포함하는 적어도 하나의 세부혈액데이터가 결손되었다고 판단되는 경우 보충을 요구할 수 있다.

[110] 도 5은 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치의 성능을 나타내는 도면이다.

[111] 도 5는 YOLO v5의 세포 분류 성능 결과치를 설명한다. 도 5는 세포의 타입/클래스 별 예측도, 재현율 등이 도시되어 있는데 이를 정리하면 이하의 표 3과 같다.

[112] [표3]

Cell	Precision	Recall	mAP (0.5)	mAP (0.5-0.95)	Epoch
All	0.829	0.700	0.747	0.572	100
Gametocyte	0.805	0.927	0.916	0.671	100
Ring	0.749	0.783	0.79	0.571	100
Schizont	1.000	0	0.242	0.218	100
Trophozoite	0.641	0.788	0.816	0.613	100
WBC	0.953	1.000	0.969	0.787	100

[113] 도 6a 내지 도 6c는 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델을 통한 타입별 객체 인식을 설명하는 그림이다. 도 6a는 YOLO v5의 말라리아 탐지 실제 예시를 나타낸다. 특히 도 6a는 원형(Ring) 말라리아 원충에 대한 검출 결과를 나타내는데, 바운딩 박스 상에 “ring”스코어링 점수는 “0.84”를 기록함을 나타낸다. 도 6b는 YOLO v5의 말라리아 탐지 실제 예시를 나타낸다. 특히 도 6b는 영양형(Trophozoite) 말라리아 원충에 대한 검출 결과를 나타내는데, 바운딩 박스 상에 “Trophozoite”스코어링 점수는 “0.87”를 기록함을 나타낸다.

[114] 도 6c는 YOLO v5의 말라리아 탐지 실제 예시를 나타낸다. 특히 도 6c는 생식모세포(Gametocyte) 말라리아 원충에 대한 검출 결과를 나타내는데, 바운딩 박스 상에 “Trophozoite”스코어링 점수는 “0.87”를 기록함을 나타낸다.

[115] 도 7은 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델에 적용되는 방법론을 상세히 도시하는 블록도이다.

[116] 본 개시의 예시적 실시예에 따르면, 제3 모델(도 2, 25)은 결정 트리(Decision Tree), K-최근린(K-Nearest Neighbor), 로지스틱 회귀분석(Logistic Regression) 각각의 예측치에 대해 투표(voting) 방식을 적용해 최종 예측값을 결정하는 구조를 가진다. 결정 트리(Decision Tree), K-최근린(K-Nearest Neighbor), 로지스틱 회귀분석(Logistic Regression) 각각의 예측치는 데이터셋으로 축적될 수 있다.

[117] 본 개시의 예시적 실시예에 따르면, 제3 모델(25)은 hard voting 방법을 적용할 수 있다. 본 개시에서 총 3가지 모델(CNN, XGB, LGB)에 대해 hard voting 방법을 적용함을 예시한 결과는 이하의 표 4와 같다.

[118] [표4]

Model	Accuracy	Precision	Recall	F1 score
Ensemble	1.0	1.0	1.0	1.0

- [119] 객체 인식(CNN계열 Yolo v5)에서는 정밀도(accuracy)가 0.881이었고, 머신러닝에서 XGBoost는 0.9828, LGBM은 0.9828을 보였다. 최종적으로 hard voting을 수행하여 3개 모델중에 2개 모델 이상에서 voting하면 그것을 정답 값으로 분류한 결과는 표 4가 같으며, 정밀도는 100%를 보여, 모두 정확한 결과를 보였다. 결국 앙상블 모델을 적용한 본 개시의 전자 장치가 최적의 성능을 보임이 검증되었다. 도 8은 본 개시의 예시적 실시예에 따른 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치를 도시하는 블록도이다. 본 개시에 따른 머신러닝 앙상블 모델이 적용가능한 전자 장치의 동작 방법, 또는 머신러닝 앙상블 모델이 적용된 말라리아 감염 진단 및 탐지 방법은 컴퓨터에 의해 실행 가능한 명령어를 저장하는 기록매체의 형태로 구현될 수 있다. 명령어는 프로그램 코드의 형태로 저장될 수 있으며, 프로세서에 의해 실행되었을 때, 프로그램 모듈을 생성하여 개시된 실시예들의 동작을 수행할 수 있다. 기록매체는 컴퓨터로 읽을 수 있는 기록매체로 구현될 수 있다.
- [120] 한편, 개시된 실시예들은 컴퓨터에 의해 실행 가능한 명령어를 저장하는 기록매체의 형태로 구현될 수 있다. 명령어는 프로그램 코드의 형태로 저장될 수 있으며, 프로세서에 의해 실행되었을 때, 프로그램 모듈을 생성하여 개시된 실시예들의 동작을 수행할 수 있다. 기록매체는 컴퓨터로 읽을 수 있는 기록매체로 구현될 수 있다.
- [121] 컴퓨터가 읽을 수 있는 기록매체로는 컴퓨터에 의하여 해독될 수 있는 명령어가 저장된 모든 종류의 기록 매체를 포함한다. 예를 들어, ROM(Read Only Memory), RAM(Random Access Memory), 자기 테이프, 자기 디스크, 플래쉬 메모리, 광 데이터 저장장치 등이 있을 수 있다.
- [122] 본 발명에 따른 전자 장치(도 1, 10)는 컴퓨팅 장치(30)에 대응될 수 있으며, 컴퓨팅 장치(30)는 적어도 하나의 프로세서(31), 프로그램(33)을 포함하는 컴퓨터 판독 가능 저장 매체(32) 및 통신 버스(39)를 포함할 수 있다. 또한, 컴퓨팅 장치(30)는 입출력 장치(35)를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(34) 및 하나 이상의 네트워크 통신 인터페이스(36)를 포함할 수 있다.
- [123] 본 발명에 따른 장치의 각 부품은 상술한 하나의 컴퓨팅 장치(30)에 모두 포함되거나 각각 별개의 장치에 구현될 수도 있다.
- [124] 이상에서와 같이 첨부된 도면을 참조하여 개시된 실시예들을 설명하였다. 본 개시가 속하는 기술분야에서 통상의 지식을 가진 자는 본 개시의 기술적 사상이나 필수적인 특징을 변경하지 않고도, 개시된 실시예들과 다른 형태로 본 개시가 실시될 수 있음을 이해할 것이다. 개시된 실시예들은 예시적인 것이며, 한정적으로 해석되어서는 안 된다.

청구범위

- [청구항 1] 머신러닝 앙상블 모델이 적용가능한 전자 장치에 있어서,
학습용이미지에 기반하여 뉴럴네트워크 학습된 제1 모델;
학습용검사데이터에 기반하여 학습된 제2 모델;
복수의 통계적 방법론에 대한 투표(voting) 방식으로 말라리아 감염여부에 관한 감염확률을 출력하도록 구성된 제3 모델; 및
상기 제1 모델에 검사대상인 혈액이미지를 입력하고, 상기 제1 모델을 이용해 상기 혈액이미지에 대해 말라리아 감염여부를 1차 판정하고, 상기 제2 모델에 검사대상인 혈액데이터를 입력하고, 상기 제2 모델을 이용해 상기 혈액데이터에 대해 말라리아 감염여부를 2차 판정하며, 상기 제3 모델을 이용해 상기 말라리아 감염확률을 출력하는 프로세서를 포함하는, 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치.
- [청구항 2] 제1항에 있어서,
상기 제1 모델은,
진단대상인 혈액이미지에 포함된 적혈구에 대한 제1 계수치에 기반해 말라리아 원충에 대한 제2 계수치를 산출함으로써 말라리아 감염여부 예측 결과를 출력하는 것을 특징으로 하는, 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치.
- [청구항 3] 제1항에 있어서,
상기 학습용이미지는,
말라리아 감염환자의 말초혈액을 분절(slide)하여 구성된 것을 특징으로 하는, 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치.
- [청구항 4] 제1항에 있어서,
상기 제2 모델은,
각 피쳐별 상이한 중요도를 가지고 학습하도록 구성되고, 결측치를 대체(replace)하지 않아도 모델링 가능한 적어도 하나의 알고리즘을 채택하는 것을 특징으로 하는, 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치.
- [청구항 5] 제4항에 있어서,
상기 제2 모델은,
XGBoost(eXtreme Gradient Boosting) 모델, LGBM(Light Gradient Boosting Model) 모델 중 적어도 하나를 사용하는 것을 특징으로 하는, 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치.
- [청구항 6] 제5항에 있어서,
상기 제2 모델은,

검사대상 혈액데이터에서 CRP(C-Reactive Protein), Plt(Platelet)를 포함하는 적어도 하나의 세부혈액데이터가 결손되었다고 판단되는 경우 보충을 요구하는 것을 특징으로 하는, 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치.

[청구항 7] 제1항에 있어서,
상기 제3 모델은,
상기 제1 모델의 예측결과 중 진단기준치에 미달하되 임계범위 내에 있는 의심결과값에 대해 재판단함으로써 말라리아 감염여부 검증결과를 출력하는 것을 특징으로 하는, 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치.

[청구항 8] 제1항에 있어서,
상기 제1 모델은,
말라리아 개별 원충(ring, trophozoite, gametocyte, schizont) 및 백혈구, 적혈구의 라벨링이 수행된 결과를 이용하는 것을 특징으로 하는, 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치.

[청구항 9] 제8항에 있어서,
상기 제1 모델은,
수평 접기(horizontal flip) 방법론, 밝기 범위(brightness range) 방법론 중 적어도 하나를 이용해 데이터 증강(data augmentation)이 수행된 결과를 이용하는 것을 특징으로 하는, 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치.

[청구항 10] 제1항에 있어서,
상기 감염확률이 임계값을 초과하는지 여부에 따라 말라리아 감염을 의미하는 양성 클래스 또는 말라리아 비감염을 의미하는 음성 클래스를 출력하도록 구성된 출력 모듈을 더 포함하는 것을 특징으로 하는, 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치.

[청구항 11] 제10항에 있어서,
상기 음성 클래스는,
상기 제2 모델의 출력값에 비해 상기 제1 모델의 출력값에 더 큰 의존도(dependency)를 가지는 것을 특징으로 하는, 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치.

[청구항 12] 제10항에 있어서,
상기 양성 클래스는,
상기 복수의 통계적 방법론 각각을 조합하는 앙상블 모델에 대한 투표 결과로부터 추론되는 것을 특징으로 하는, 말라리아 진단 및 탐지를 위한 머신러닝 앙상블 모델이 적용가능한 전자 장치.

[청구항 13] 전자 장치에 의해 수행되는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법에 있어서,

학습용이미지에 기반하여 뉴럴네트워크 학습된 제1 모델에 검사대상인 혈액이미지를 입력하는 단계;
 학습된 상기 제1 모델을 이용해 상기 혈액이미지에 대해 말라리아 감염여부를 1차 판정하는 단계;
 학습용검사데이터에 기반하여 학습된 제2 모델에 검사대상인 혈액데이터를 입력하는 단계;
 학습된 상기 제2 모델을 이용해 상기 혈액데이터에 대해 말라리아 감염여부를 2차 판정하는 단계; 및
 복수의 통계적 방법론에 대해 투표(voting) 방식을 적용하는 제3 모델을 이용해 말라리아 감염확률을 출력하는 단계를 포함하는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법.

- [청구항 14] 제13항에 있어서,
 상기 학습된 제1 모델을 이용해 상기 혈액이미지에 대해 말라리아 감염여부를 1차 판정하는 단계는,
 검사대상인 혈액이미지에 포함된 적혈구에 대한 제1 계수치를 출력하는 단계;
 상기 제1 계수치에 기반해 말라리아 원충에 대한 제2 계수치를 출력하는 단계; 및
 상기 제2 계수치에 기반해 말라리아 감염여부 예측결과를 출력하는 것을 특징으로 하는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법.
- [청구항 15] 제13항에 있어서,
 상기 학습용이미지는,
 말라리아 감염환자의 말초혈액을 분절(slide)하여 구성된 것을 특징으로 하는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법.
- [청구항 16] 제13항에 있어서,
 상기 학습된 제2 모델을 이용해 상기 혈액데이터에 대해 말라리아 감염여부를 2차 판정하는 단계는,
 결측치를 대체(replace)하지 않아도 모델링 가능한 적어도 하나의 알고리즘을 채택하는 단계를 포함하는 것을 특징으로 하는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법.
- [청구항 17] 제16항에 있어서,
 상기 제2 모델은,
 각 피처별 상이한 중요도를 가지고 학습하도록 구성되고, XGBoost(eXtreme Gradient Boosting) 모델, LGBM(Light Gradient Boosting Model) 모델 중 적어도 하나를 사용하는 것을 특징으로 하는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법.
- [청구항 18] 제17항에 있어서,
 상기 제2 모델은,

상기 혈액데이터에서 CRP(C-Reactive Protein), Plt(Platelet)를 포함하는 적어도 하나의 세부혈액데이터가 결손되었다고 판단되는 경우 보충을 요구하는 것을 특징으로 하는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법.

[청구항 19] 제13항에 있어서,
복수의 통계적 방법론에 대해 투표(voting) 방식을 적용하는 제3 모델을 이용해 말라리아 감염확률을 출력하는 단계는,
상기 제1 모델의 예측결과 중 진단기준치에 미달하되 임계범위 내에 있는 의심결과값에 대해 재판단함으로써 말라리아 감염여부 검증결과를 출력하는 단계를 포함하는 것을 특징으로 하는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법.

[청구항 20] 제13항에 있어서,
상기 제1 모델은,
백혈구, 적혈구 및 원형(ring), 영양형(trophozoite), 생식모세포(gametocyte), 및 열충(schizont)를 포함하는 말라리아 개별 원충의 라벨링이 수행된 결과를 이용하는 것을 특징으로 하는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법.

[청구항 21] 제20항에 있어서,
상기 제1 모델은,
수평 접기(horizontal flip) 방법론, 밝기 범위(brightness range) 방법론 중 적어도 하나를 이용해 데이터 증강(data augmentation)이 수행된 결과를 이용하는 것을 특징으로 하는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법.

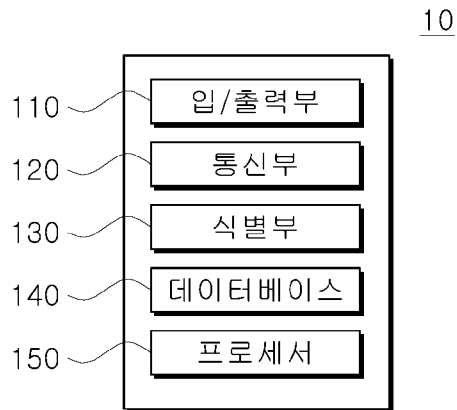
[청구항 22] 제13항에 있어서,
상기 복수의 통계적 방법론에 대해 투표(voting) 방식을 적용하는 제3 모델을 이용해 말라리아 감염확률을 출력하는 단계는,
상기 감염확률이 임계값 초과인 경우 말라리아 감염을 의미하는 “양성” 클래스를 출력하는 단계; 및
상기 감염확률이 임계값 미만인 경우 말라리아 비감염을 의미하는 “음성” 클래스를 출력하는 단계를 포함하는 것을 특징으로 하는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법.

[청구항 23] 제22항에 있어서,
상기 음성 클래스는,
상기 제2 모델의 출력값에 비해 상기 제1 모델의 출력값에 더 큰 의존도(dependency)를 가지는 것을 특징으로 하는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법.

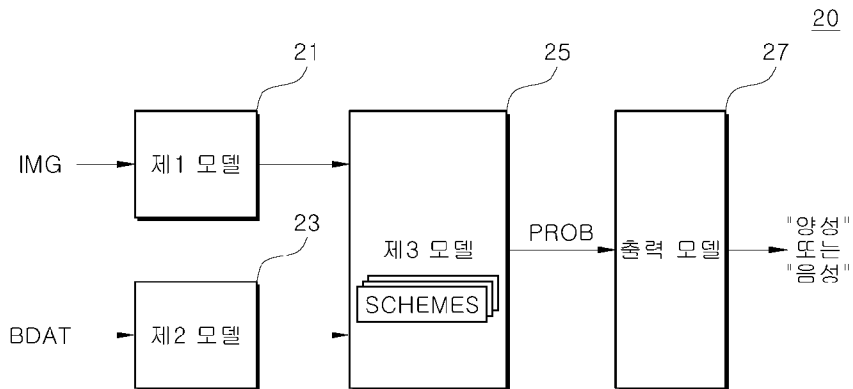
[청구항 24] 제22항에 있어서,
상기 양성 클래스는,

상기 복수의 통계적 방법론 각각을 조합하는 앙상블 모델에 대한 투표 결과로부터 추론되는 것을 특징으로 하는, 머신러닝 앙상블 모델 기반 말라리아 진단 및 탐지 방법.

[도1]



[도2]



[도3]

0.9993872549019608					
	precision	recall	f1-score	support	
0	1.0000	0.9583	0.9787	24	
1	0.9714	1.0000	0.9855	34	
accuracy			0.9828	58	
macro avg	0.9857	0.9792	0.9821	58	
weighted avg	0.9833	0.9828	0.9827	58	
1.0					
	precision	recall	f1-score	support	
0	0.9600	1.0000	0.9796	24	
1	1.0000	0.9706	0.9851	34	
accuracy			0.9828	58	
macro avg	0.9800	0.9853	0.9823	58	
weighted avg	0.9834	0.9828	0.9828	58	

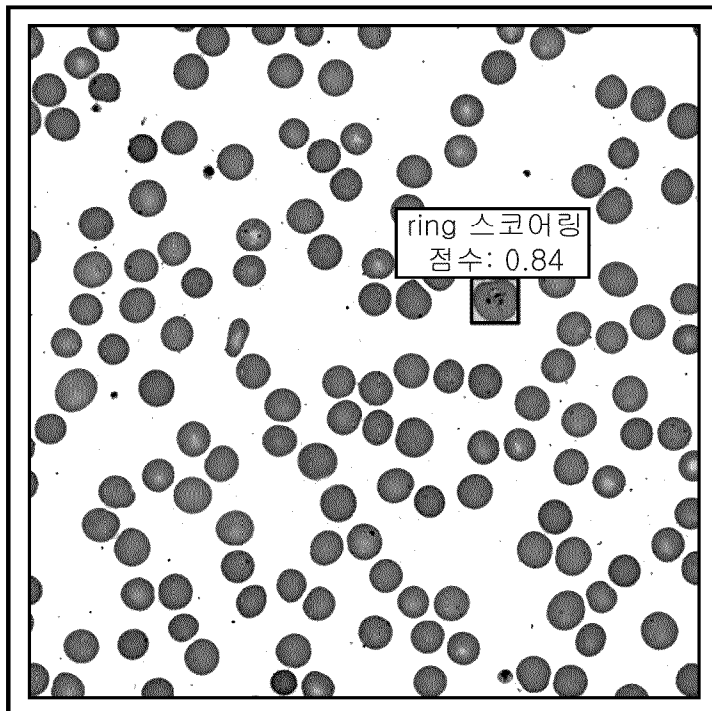
[도4]

CRP	0.289505
Plt	0.203164
LDH	0.162625
RBC	0.060911
AST	0.038452
ALT	0.035824
나이	0.029715
lymphocyte(%)	0.029087
Neutrophil(%)	0.028927
total bilirubin	0.028641
Hb	0.025013
성별	0.019116
monocyte(%)	0.017810
eo(%)	0.010987
WBC	0.010611
baso(%)	0.009613

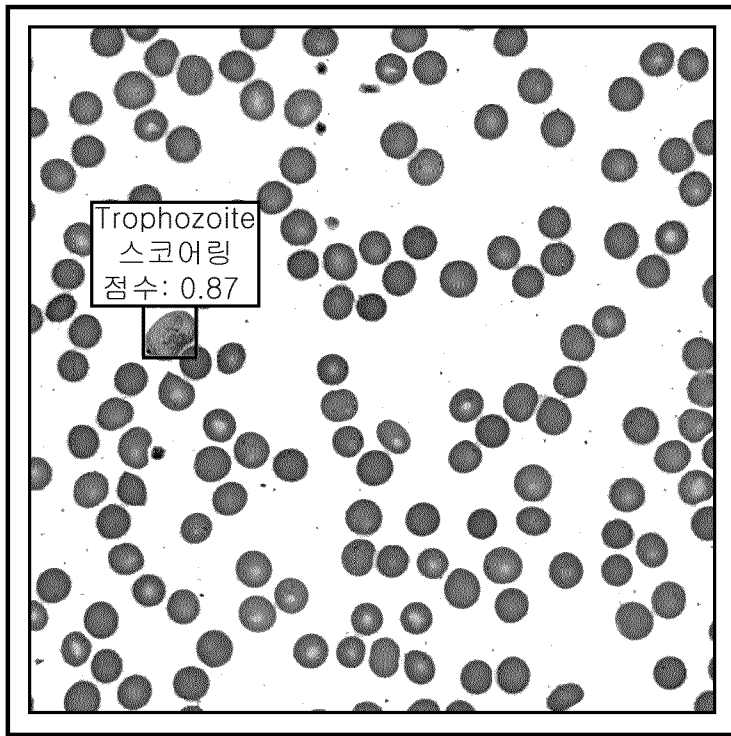
[도5]

Class	P	R	mAP@.5	mAP@.5:.95:
all	0.829	0.7	0.747	0.572
gametocyte	0.805	0.927	0.916	0.671
ring	0.749	0.783	0.79	0.571
schizont	1	0	0.242	0.218
trophozoite	0.641	0.788	0.816	0.613
wbc	0.953	1	0.969	0.787

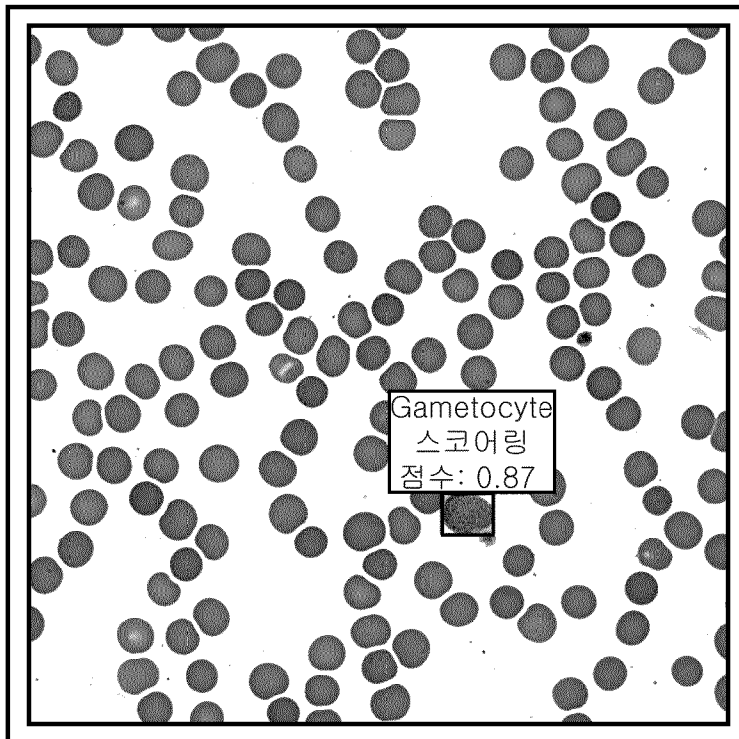
[도6a]



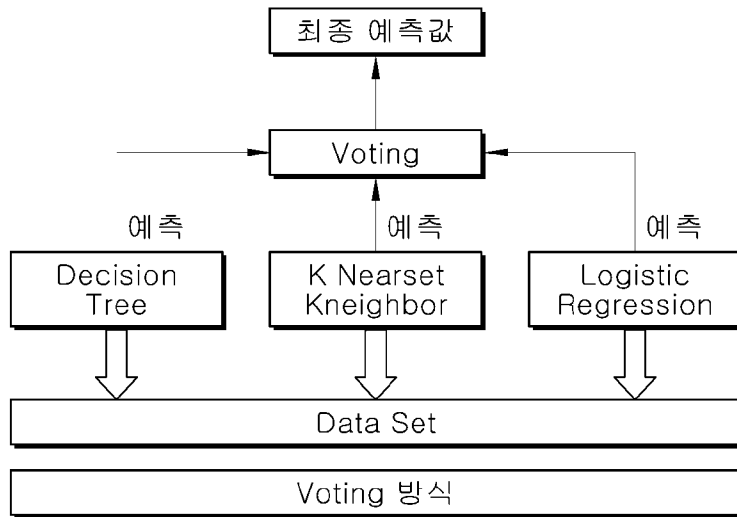
[도6b]

규칙 91,
13.03.2024

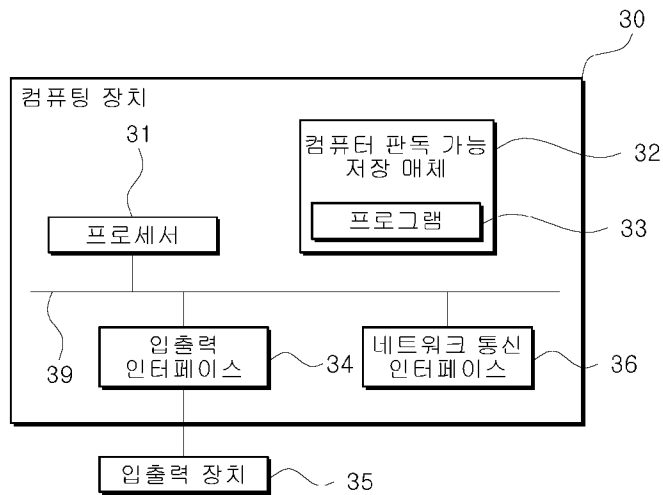
[도6c]

규칙 91,
13.03.2024

[도7]



[도8]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2024/001776

A. CLASSIFICATION OF SUBJECT MATTER

G16H 50/80(2018.01)i; **G16H 50/20**(2018.01)i; **G16H 50/50**(2018.01)i; **G16H 50/70**(2018.01)i; **G16H 30/40**(2018.01)i;
G06N 20/20(2019.01)i; **G06T 7/00**(2006.01)i; **A61B 5/00**(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G16H 50/80(2018.01); G01N 30/72(2006.01); G01N 33/487(2006.01); G16H 30/20(2018.01); G16H 50/20(2018.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models: IPC as above
Japanese utility models and applications for utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & keywords: 말라리아(malaria), 진단(diagnosis), 기계학습(machine learning), 앙상블(ensemble)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2019-0117229 A (POSTECH RESEARCH AND BUSINESS DEVELOPMENT FOUNDATION) 16 October 2019 (2019-10-16) See paragraphs [0090], [0102]-[0106] and [0112].	1-3,8-10,12-15,20-22,24
A		4-7,11,16-19,23
Y	RAGB, Hussin K. et al. Deep Convolutional Neural Network Ensemble for Improved Malaria Parasite Detection. [online]. IEEE Xplore. 10 May 2021. [Retrieved on 22 April 2024]. Retrieved from the Internet < https://doi.org/10.1109/AIPR50011.2020.9425273 >. See pages 3, 5 and 7-8 and figure 4.	1-3,8-10,12-15,20-22,24
Y	송혜정. 패턴인식 기반 말라리아 자동 진단용 소프트웨어의 개발. [online]. scienceon.kisti.re.kr. 30 September 2014, non-official translation (SONG, Hye-Jung. Development of Pattern Recognition-Based Automatic Malaria Diagnosis Software). [Retrieved on 22 April 2024]. Retrieved from the Internet <URL: ">https://scienceon.kisti.re.kr/srch/selectPORSrchReport.do?cn=TRKO201400029839#> . See page 5.	2,8-9,14,20-21

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
“D” document cited by the applicant in the international application
“E” earlier application or patent but published on or after the international filing date
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
“O” document referring to an oral disclosure, use, exhibition or other means
“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“&” document member of the same patent family

Date of the actual completion of the international search

02 May 2024

Date of mailing of the international search report

07 May 2024

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon Building 4, 189 Cheongsaro, Seo-gu, Daejeon 35208

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	김민규 등. 딥러닝 기반 의료영상 분석을 위한 데이터 증강 기법. [online]. www.ncbi.nlm.nih.gov. 30 November 2020, vol. 81, no. 6 (KIM, Min Gyu et al. Data Augmentation Techniques for Deep Learning-Based Medical Image Analyses). [Retrieved on 22 April 2024]. Retrieved from the Internet <URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9431833>. See page 1 and figure 1.	9,21
A	WO 2021-202620 A1 (THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY) 07 October 2021 (2021-10-07) See entire document.	1-24

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2024/001776

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
KR	10-2019-0117229	A	16 October 2019	KR	10-2052902	B1	11 December 2019
WO	2021-202620	A1	07 October 2021	None			

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G16H 50/80(2018.01)i; G16H 50/20(2018.01)i; G16H 50/50(2018.01)i; G16H 50/70(2018.01)i; G16H 30/40(2018.01)i;
G06N 20/20(2019.01)i; G06T 7/00(2006.01)i; A61B 5/00(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G16H 50/80(2018.01); G01N 30/72(2006.01); G01N 33/487(2006.01); G16H 30/20(2018.01); G16H 50/20(2018.01)

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC
일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 말라리아(malaria), 진단(diagnosis), 기계학습(machine learning), 앙상블(ensemble)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-2019-0117229 A (포항공과대학교 산학협력단) 2019.10.16 단락 90, 102-106, 112 참조.	1-3,8-10,12-15,20-22,24
A		4-7,11,16-19,23
Y	Hussin K. Ragb 등, Deep Convolutional Neural Network Ensemble for Improved Malaria Parasite Detection, [온라인], IEEE Xplore, 2021.05.10., [검색일: 2024.04.22], 인터넷 검색, <10.1109/AIPR50011.2020.9425273> 페이지 3, 5, 7-8 및 도면 4 참조.	1-3,8-10,12-15,20-22,24
Y	송혜정, 패턴인식 기반 말라리아 자동 진단용 소프트웨어의 개발, [온라인], scienceon.kisti.re.kr, 2014.09.30, [검색일: 2024.04.22], 인터넷 검색, <URL: https://scienceon.kisti.re.kr/srch/selectPORSrchReport.do?cn=TRKO201400029839#> 페이지 5 참조.	2,8-9,14,20-21

☒ 추가 문헌이 C(계속)에 기재되어 있습니다.

☒ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

“A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌

“D” 본 국제출원에서 출원인이 인용한 문헌

“E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌

“L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌

“O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌

“P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

“T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌

“X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.

“Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.

“&” 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일

2024년05월02일(02.05.2024)

국제조사보고서 발송일

2024년05월07일(07.05.2024)

ISA/KR의 명칭 및 우편주소

대한민국 특허청
(35208) 대전광역시 서구 청사로 189, 4동 (둔산동,
정부대전청사)

팩스 번호 +82-42-481-8578

심사관

이강하

전화번호 +82-42-481-5003

C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	김민규 등, 딥러닝 기반 의료영상 분석을 위한 데이터 증강 기법, [온라인], www.ncbi.nlm.nih.gov, 2020.11.30, 제81권 6호, [검색일 : 2024.04.22], 인터넷 검색, <URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9431833> 페이지 1 및 도면 1 참조.	9,21
A	WO 2021-202620 A1 (THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY) 2021.10.07 전체문헌 참조.	1-24

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2019-0117229 A	2019/10/16	KR 10-2052902 B1	2019/12/11
WO 2021-202620 A1	2021/10/07	없음	