

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2020년 5월 22일 (22.05.2020)



(10) 국제공개번호
WO 2020/101457 A2

- (51) 국제특허분류:
G16H 50/20 (2018.01)
- (21) 국제출원번호: PCT/KR2019/015737
- (22) 국제출원일: 2019년 11월 18일 (18.11.2019)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2018-0142095 2018년 11월 16일 (16.11.2018)KR
- (71) 출원인: 주식회사 딥바이오 (DEEP BIO INC.) [KR/KR];
08394 서울시 구로구 디지털로 242, 1013호(구로동),
Seoul (KR).
- (72) 발명자: 김선우 (KIM, Sun Woo); 13599 경기도 성남시
분당구 내정로 152, 129동 502호(수내동, 파크타운롯데
아파트), Gyeonggi-do (KR).
- (74) 대리인: 심충섭 (SHIM, Choong Sup); 06729 서울시 서
초구 효령로 428, 501호 (서초동, 광림빌딩), Seoul (KR).
- (81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국
내 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, CZ, DE, DJ, DK, DM, DO, DZ, EC,
EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU,
ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KW, KZ,
LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK,
MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR,
TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역
내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE,

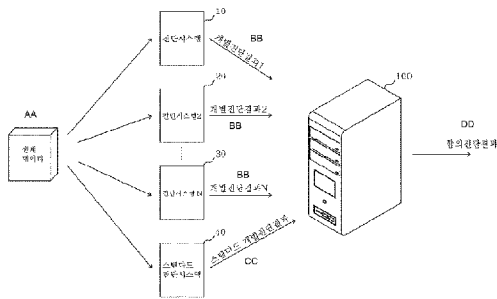
LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM,
ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유
럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK,
MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML,
MR, NE, SN, TD, TG).

공개:

— 국제조사보고서 없이 공개하며 보고서 접수 후 이를 별
도 공개함 (규칙 48.2(g))

(54) Title: SUPERVISED LEARNING-BASED CONSENSUS DIAGNOSIS METHOD AND SYSTEM THEREOF

(54) 발명의 명칭: 지도학습기반의 합의 진단방법 및 그 시스템



10 ... Diagnosis system
 20 ... Diagnosis system 2
 30 ... Diagnosis system N
 40 ... Standard diagnosis system
 AA ... Biological data
 BB ... Individual diagnosis result
 CC ... Standard individual diagnosis result
 DD ... Consensus diagnosis result

(57) Abstract: Disclosed are a supervised learning-based consensus diagnosis method and a system thereof. The supervised learning-based consensus diagnosis method comprises: a step of confirming, by a consensus diagnostic system, N individual diagnosis results in which each of N (N is an integer of 2 or more) diagnostic systems receives and outputs predetermined biological data, wherein the N diagnostic systems, respectively, are systems that are each trained with learning data annotated by different annotation subjects; and a step of outputting a consensus diagnosis result of the biological data on the basis of the individual diagnosis results confirmed by the consensus diagnosis system.

(57) 요약서: 지도학습기반 합의 진단방법 및 그 시스템이 개시된다. 상기 지도학습기반 합의 진단방법은 합의 진단 시스템이 N(N은 2 이상의 정수) 개의 진단 시스템들-N 개의 진단시스템들 각각은 서로 다른 어노테이션(annotation) 주체에 의해 어노테이션된 학습 데이터로 각각 학습된 시스템들임- 각각이 소정의 생체 데이터를 각각 입력받아 출력하는 N 개의 개별진단결과들을 확인하는 단계, 상기 합의 진단 시스템이 확인한 상기 개별진단결과들에 기초하여 상기 생체 데이터의 합의진단결과를 출력하는 단계를 포함한다.

WO 2020/101457 A2

명세서

발명의 명칭: 지도학습기반의 합의 진단방법 및 그 시스템

기술분야

- [1] 본 발명은 학습을 통해 질병을 진단하는 방법 및 그 시스템에 관한 것이다. 보다 상세하게는 뉴럴 네트워크를 통해 지도학습기반으로 학습된 진단 시스템을 복수개 구비하고, 복수 개의 진단 시스템들의 합의를 통해 최종적인 진단결과를 도출함으로써 지도학습기반의 학습의 취약점을 보완할 수 있는 방법 및 시스템에 관한 것이다.

[2]

배경기술

- [3] 병리학 또는 병리과에서 수행하는 주요한 업무 중 하나는 환자의 생체이미지를 판독하여 특정 질병에 대한 상태 또는 징후를 판단하는 진단을 수행하는 일이다. 이러한 진단은 오랜기간 숙련된 의료인의 경험과 지식에 의해 의존되는 방식이다.
- [4] 최근에는 기계학습의 발달로 인해 이미지를 인식하거나 분류하는 등의 업무를 컴퓨터 시스템에 의해 자동화하고자 하는 시도가 활발히 이루어지고 있다. 특히 기계학습의 일종인 뉴럴 네트워크(예컨대, 컨벌루션 뉴럴 네트워크(Convolution neural network, CNN)를 이용한 딥러닝 방식)를 이용하여 숙련된 의료인이 수행하던 진단을 자동화하기 위한 시도가 이루어지고 있다.
- [5] 특히 뉴럴 네트워크(예컨대, CNN)를 이용한 딥러닝을 통한 진단은 종래에 숙련된 의료인의 경험과 지식을 단순히 자동화하는 것이 아니라, 스스로 학습을 통해 특징적인 요소들을 찾아내어 원하는 해답을 도출한다는 점에 있어서 오히려 숙련된 의료인이 알지 못하던 질병인자의 특징을 이미지에서 찾아내는 경우도 있다.
- [6] 일반적으로 생체 데이터(예컨대, 생체 이미지)를 이용하는 뉴럴 네트워크를 통한 질병의 진단은 생체 데이터에 숙련된 의료인이 특정 질병의 상태(예컨대, 암이 발현되었는지 여부)를 어노테이션(annotation)하고, 이러한 어노테이션된 다수의 데이터들을 학습 데이터로 이용하여 뉴럴 네트워크를 학습하게 된다. 즉, 학습을 위해 학습 데이터에 어노테이션을 수행하여 어노테이션된 학습 데이터를 통한 학습이 주로 이용되는데 이러한 학습방식을 지도학습(supervised learning)이라고 한다.
- [7] 하지만 지도학습은 학습된 시스템의 성능이 어노테이션을 수행한 어노테이터의 판단 성향 또는 경향에 지배적인 영향을 받게 된다. 즉, 학습된 진단 시스템의 진단 결과는 어노테이터의 어노테이션 경향이나 성향에 의존적이게 된다.
- [8] 하지만 실제 어노테이션을 수행할 때 어노테이터에 따라 서로 다른

어노테이션을 수행할 정도로, 학습을 할 생체 데이터에는 명확하게 분류되기 힘든 생체 데이터가 상당히 존재하는 것이 현실이다. 즉, 숙련된 의료인이 어노테이션을 수행한다고 하더라도 생체 데이터에 따라 질병의 발현 여부 또는 질병의 진행 정도를 분류함에 있어서 의견이 다를 수 있다.

[9] 그럼에도 지도학습 기반으로 학습된 질병의 진단시스템의 경우에는 어노테이터의 경향에 의존적인 판단을 진단결과로 출력하게 되는 취약점이 있다.

[10] 따라서 이러한 취약점을 개선할 수 있는 기술적 사상이 요구된다.

[11]

[12] *선행기술문헌

[13] -특허문헌

[14] 한국등록특허 10-20170057399 "모듈화된 강화학습을 통한 질병 진단 시스템"

발명의 상세한 설명

기술적 과제

[15] 본 발명이 이루고자 하는 기술적인 과제는 서로 다른 어노테이션 주체의 의해 학습된 지도학습기반의 진단 시스템을 복수 개 구비하고, 소정의 생체 데이터에 대해 상기 복수 개의 진단 시스템들이 각각 수행한 진단 결과에 기초하여 최종 진단결과 즉, 진단 시스템들간의 합의된 진단결과를 도출하도록 함으로써 지도학습기반의 취약점을 보완할 수 있는 방법 및 시스템을 제공하는 것이다.

[16] 또한 이러한 진단시스템들의 합의를 통한 진단결과를 이용하여 개별 진단 시스템들을 재학습할 수 있도록 함으로써, 개별 진단 시스템들의 성능 역시 향상시킬 수 있는 방법 및 시스템을 제공하는 것이다.

[17]

과제 해결 수단

[18] 상기 기술적 과제를 달성하기 위한 지도학습을 통해 학습된 진단 시스템을 이용해 질병을 진단하는 방법은 합의 진단 시스템이 $N(N$ 은 2이상의 정수) 개의 진단 시스템들- N 개의 진단시스템들 각각은 서로 다른 어노테이션(annotation) 주체에 의해 어노테이션된 학습 데이터로 각각 학습된 시스템들임- 각각이 소정의 생체 데이터를 각각 입력받아 출력하는 N 개의 개별진단결과들을 확인하는 단계, 상기 합의 진단 시스템이 확인한 상기 개별진단결과들에 기초하여 상기 생체 데이터의 합의진단결과를 출력하는 단계를 포함한다.

[19] 상기 지도학습기반 합의 진단방법은 골드 스탠다드(gold standard) 학습 데이터로 학습된 스탠다드 진단 시스템이 상기 생체 데이터를 입력받아 출력하는 스탠다드 개별진단결과를 확인하는 단계를 더 포함하며, 상기 합의 진단 시스템이 확인한 상기 개별진단결과들에 기초하여 상기 생체 데이터의 합의진단결과를 출력하는 단계는, 상기 개별진단결과들 및 상기 스탠다드 개별진단결과에 기초하여 상기 합의진단결과를 출력하는 단계를 포함할 수

있다.

- [20] 상기 합의 진단 시스템이 확인한 상기 개별진단결과들에 기초하여 상기 생체 데이터의 합의진단결과를 출력하는 단계는 상기 진단 시스템들 각각에 부여된 가중치에 기초하여 상기 합의진단결과를 출력하는 것을 특징으로 하며, 상기 가중치는 특정 진단 시스템이 합의진단결과와 다른 개별진단결과를 출력한 횟수에 기초하여 결정될 수 있다.
- [21] 상기 가중치는 상기 특정 진단 시스템이 합의진단결과와 다른 개별진단결과를 출력한 경우, 상기 전체 개별진단결과들 중 상기 다른 개별진단결과를 출력한 타 진단 시스템의 개수 또는 비중에 더 기초하여 결정될 수 있다.
- [22] 상기 합의 진단 시스템이 확인한 상기 개별진단결과들에 기초하여 상기 생체 데이터의 합의진단결과를 출력하는 단계는 상기 진단 시스템들 및 상기 스탠다드 진단 시스템 각각에 부여된 가중치에 기초하여 상기 합의진단결과를 출력하는 것을 특징으로 하며, 상기 스탠다드 진단 시스템에는 상기 진단 시스템들에 비해 높은 가중치가 부여된 것을 특징으로 할 수 있다.
- [23] 상기 진단 시스템들은 소정의 학습 데이터 세트가 N 개로 분할된 분할 학습데이터 세트들-임의의 분할 학습데이터 세트는 타 분할 학습데이터 세트에는 포함되지 않은 학습데이터를 적어도 한 개 포함하며 각각이 서로 다른 주체에 의해 어노테이션된 학습 데이터 세트임-을 각각 학습하여 질병의 진단결과를 출력하는 시스템인 것을 특징으로 할 수 있다.
- [24] 상기 지도학습기반 합의 진단방법은 특정 생체 데이터의 합의진단결과와 상기 특정 생체 데이터의 개별진단결과가 상이한 진단 시스템이, 상기 합의진단결과로 어노테이션된 상기 특정 생체 데이터를 포함하는 학습 데이터로 재훈련되는 단계를 더 포함할 수 있다.
- [25] 상기의 방법은 데이터 처리장치에 설치되며 컴퓨터 판독가능한 기록매체에 저장된 컴퓨터 프로그램에 의해 구현될 수 있다.
- [26] 상기의 기술적 과제를 해결하기 위한 시스템은 프로세서 및 프로그램이 저장된 저장장치를 포함하며, 상기 프로그램은 $N(N$ 은 2이상의 정수) 개의 진단 시스템들- N 개의 진단시스템들 각각은 서로 다른 어노테이션(annotation) 주체에 의해 어노테이션된 학습 데이터로 각각 학습된 시스템들임- 각각이 소정의 생체 데이터를 각각 입력받아 출력하는 N 개의 개별진단결과들을 확인하고, 확인한 상기 개별진단결과들에 기초하여 상기 생체 데이터의 합의진단결과를 출력한다.
- [27] 상기 프로그램은 골드 스탠다드(gold standard) 학습 데이터로 학습된 스탠다드 진단 시스템이 상기 생체 데이터를 입력받아 출력하는 스탠다드 개별진단결과를 더 확인하고, 상기 개별진단결과들 및 상기 스탠다드 개별진단결과에 기초하여 상기 합의진단결과를 출력할 수 있다.
- [28] 상기 프로그램은 상기 진단 시스템들 각각에 부여된 가중치에 기초하여 상기 합의진단결과를 출력하는 것을 특징으로 하며, 상기 가중치는 상기

합의진단결과와 상기 진단 시스템들 각각의 개별진단결과가 다른 경우의 횟수에 기초하여 결정될 수 있다.

[29] 상기 가중치는 상기 특정 진단 시스템이 합의진단결과와 다른 개별진단결과를 출력한 경우, 상기 전체 개별진단결과들 중 상기 다른 개별진단결과를 출력한 타 진단 시스템의 개수 또는 비중에 더 기초하여 결정될 수 있다.

[30] 상기 프로그램은 상기 진단 시스템들 및 상기 스탠다드 진단 시스템 각각에 부여된 가중치에 기초하여 상기 합의진단결과를 출력하는 것을 특징으로 하며, 상기 스탠다드 진단 시스템에는 상기 진단 시스템들에 비해 높은 가중치가 부여된 것을 특징으로 할 수 있다.

[31] 또한, 상기 합의 진단시스템의 진단에 이용되는 생체 데이터 중 특정 생체 데이터는, 상기 합의진단결과와 상기 진단 시스템들 중 특정 진단시스템의 개별진단결과가 상이하며, 상기 합의진단결과로 어노테이션된 상기 특정 생체 데이터가 포함된 학습 데이터로 상기 특정 진단시스템이 재훈련되는 것을 특징으로 할 수 있다.

[32]

발명의 효과

[33] 본 발명의 기술적 사상에 따르면 어느 한 어노테이션 주체의 의해 어노테이션된 학습 데이터를 이용하여 학습된 지도학습기반의 진단 시스템을 통해 질병을 진단하는 경우 발생할 수 있는 진단결과의 어노테이터 의존성을 보완할 수 있는 효과가 있다.

[34] 또한 본 발명의 기술적 사상에 따라 합의를 통한 진단결과를 이용하여 개별 진단 시스템들을 재학습할 수 있으며, 이러한 경우 개별 진단 시스템들의 성능 역시 향상시킬 수 있는 효과가 있다.

[35]

도면의 간단한 설명

[36] 본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 간단한 설명이 제공된다.

[37] 도 1은 본 발명의 기술적 사상에 따른 지도학습기반 합의 진단방법을 구현하기 위한 개략적인 시스템 구성을 나타내는 도면이다.

[38] 도 2는 본 발명의 실시 예에 따른 진단 시스템들 각각의 개략적인 구성을 설명하기 위한 도면이다.

[39] 도 3은 본 발명의 실시 예에 따라 개별 진단 결과에 기초한 합의를 도출하는 개념을 설명하기 위한 도면이다.

[40] 도 4는 본 발명의 실시 예에 따른 학습 데이터의 분할 어노테이션 개념을 설명하기 위한 도면이다.

[41] 도 5는 본 발명의 실시 예에 따라 합의 진단결과를 이용하여 개별 진단 시스템을 재학습하는 개념을 설명하기 위한 도면이다.

[42]

발명의 실시를 위한 최선의 형태

[43]

본 발명은 다양한 변환을 가할 수 있고 여러 가지 실시 예를 가질 수 있는 바, 특정 실시 예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변환, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

[44]

제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 아니 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.

[45]

본 출원에서 사용한 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다.

[46]

본 명세서에 있어서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[47]

또한, 본 명세서에 있어서는 어느 하나의 구성요소가 다른 구성요소로 데이터를 '전송'하는 경우에는 상기 구성요소는 상기 다른 구성요소로 직접 상기 데이터를 전송할 수도 있고, 적어도 하나의 또 다른 구성요소를 통하여 상기 데이터를 상기 다른 구성요소로 전송할 수도 있는 것을 의미한다. 반대로 어느 하나의 구성요소가 다른 구성요소로 데이터를 '직접 전송'하는 경우에는 상기 구성요소에서 다른 구성요소를 통하지 않고 상기 다른 구성요소로 상기 데이터가 전송되는 것을 의미한다.

[48]

이하, 첨부된 도면들을 참조하여 본 발명의 실시 예들을 중심으로 본 발명을 상세히 설명한다. 각 도면에 제시된 동일한 참조부호는 동일한 부재를 나타낸다.

[49]

도 1은 본 발명의 기술적 사상에 따른 지도학습기반 합의 진단방법을 구현하기 위한 개략적인 시스템 구성을 나타내는 도면이다.

[50]

도 1을 참조하면, 본 발명의 기술적 사상에 따른 지도학습기반 합의 진단방법을 구현하기 위해서는 합의 진단시스템(100)이 구현될 수 있다.

[51]

상기 합의 진단 시스템(100)은 본 발명의 기술적 사상에 따라 복수(예컨대, N 개, N은 2이상의 정수)의 개별 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30)으로부터 개별진단결과들(예컨대, 개별진단결과 1 내지 개별진단결과 N)를 수신할 수 있다. 그리고 수신한 개별진단결과들(예컨대,

- 개별진단결과 1 내지 개별진단결과 N)에 기초하여 합의 진단 결과를 출력할 수 있다.
- [52] 상기 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 각각은 지도학습 기반으로 학습되어 소정의 생체 데이터(예컨대, 생체 이미지)가 입력되면, 상기 생체 데이터를 학습된 뉴럴 네트워크를 통해 미리 정해진 분류(예컨대, 질병의 발현여부 또는 질병의 진행도에 따른 분류)로 분류하여 그 결과를 진단결과로 출력하는 시스템일 수 있다.
- [53] 상기 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30)은 각각 동일한 생체 데이터를 입력받아, 각각의 학습된 뉴럴 네트워크를 통해 진단결과들 즉, 상기 개별진단결과들(예컨대, 개별진단결과 1 내지 개별진단결과 N)을 출력할 수 있다.
- [54] 상술한 바와 같이 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 각각의 진단결과는 어노테이터에 따른 어노테이션에 의존적일 수 있다. 즉, 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 각각은 어노테이션된 학습 데이터에 의존성을 갖도록 개별진단결과를 출력할 수 있다.
- [55] 이러한 취약점을 보완하기 위해 본 발명의 기술적 사상에 따른 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 각각은 서로 다른 어노테이션 주체에 의해 어노테이션된 학습 데이터 세트로 학습된 시스템일 수 있다.
- [56] 예컨대, 진단 시스템 1(10)은 제1어노테이터에 의해 어노테이션된 제1학습 데이터 세트로 학습된 시스템일 수 있고, 진단 시스템 2(20)는 제2어노테이터에 의해 어노테이션된 제2학습 데이터 세트로 학습된 시스템일 수 있으며, 진단 시스템 N(30)은 제N어노테이터에 의해 어노테이션된 제N학습 데이터 세트로 학습된 시스템일 수 있다.
- [57] 이러한 경우 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 각각은 학습에 이용된 학습 데이터의 어노테이터의 판단성향 또는 경향에 의존적인 특징을 가질 수 있다.
- [58] 예컨대, 동일한 생체 데이터(예컨대, 생체 이미지)를 입력받은 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30)은 이상적으로는 동일한 진단결과를 출력하는 것이 바람직하지만, 상술한 바와 같이 어노테이터의 판단성향이 진단결과에 영향을 미쳐서 어떤 경우에는 서로 다른 진단 결과를 출력할 수도 있다. 즉, 어노테이터의 성향에 따라 동일한 생체 데이터에 대해 서로 다른 진단결과를 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30)이 출력할 수도 있다.
- [59] 하지만 이러한 취약점은 본 발명의 기술적 사상에 따라 복수의 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30)의 개별진단결과를 이용한 합의진단을 통해 보완될 수 있다.
- [60] 상기 합의 진단 시스템(100)은 각각의 진단 시스템들(예컨대, 진단 시스템 1

내지 진단 시스템 N, 10 내지 30)이 출력한 개별진단결과들에 기초하여 합의진단결과를 출력할 수 있다. 합의진단결과는 개별진단결과들 전부 또는 일부에 기초하여 생성된 최종적인 진단결과를 의미할 수 있다.

- [61] 합의진단결과를 출력하기 위해 상기 합의 진단 시스템(100)은 개별진단결과들이 서로 다른 진단결과를 갖는 경우에는 단순히 다수결에 따른 합의진단결과를 도출할 수도 있다. 또는 상기 합의 진단 시스템(100)은 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 각각에 대해 가중치를 부여하고, 부여된 가중치를 상응하는 개별진단결과에 적용한 합의진단결과를 출력할 수도 있다.
- [62] 가중치는 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 각각의 진단이력에 따라 부여될 수 있다. 예컨대, 합의진단결과와 다른 진단결과를 출력한 진단 시스템에는 가중치가 소정의 기준만큼 감소되도록 가중치 부여 프로세스가 상기 합의 진단 시스템(100)에 의해 구동될 수 있다. 또한 가중치 부여 프로세스는 단순히 합의진단결과와 다른지 여부만을 고려할 수도 있지만, 전체 개별진단결과들 중 진단 결과의 종류별 개수를 더 고려할 수도 있다.
- [63] 예컨대, 10개의 진단 시스템이 구현되어 합의진단을 구현할 경우, 10개의 진단 시스템들 중 특정 진단 시스템 어느 하나만 특정 생체 데이터에 대해 질병이 발생되지 않았다고 판단하고, 나머지 진단 시스템들은 모두 질병이 발생되었다고 판단할 수 있다. 이러한 경우에는 상기 특정 진단 시스템의 가중치를 미리 정해진 제1기준 값(예컨대, 0.1)만큼 감소시킬 수 있다.
- [64] 하지만 10개의 진단 시스템들 중 4개의 진단 시스템이 상기 특정 생체 데이터에 대해 질병이 발생되지 않았다고 판단하고, 6개의 진단 시스템들은 모두 질병이 발생되었다고 판단할 수 있다. 이러한 경우에는 합의진단결과가 질병의 발현이라 할지라도 4개의 진단 시스템의 가중치는 제1기준 값(예컨대, 0.1)보다는 작은 소정의 제2기준 값(예컨대, 0.05)만 감소될 수도 있다.
- [65] 즉, 합의진단결과를 정답으로 가정할 경우, 단순히 오답을 출력한 횟수에 따라 진단 시스템의 가중치를 조절할 수도 있지만, 특정 진단 시스템이 오답을 출력하더라도 오답을 출력한 진단 시스템의 수를 더 고려하여 가중치의 조절정도를 달리하거나 경우에 따라서는 오답이 아니라고 판단하여 가중치를 조절하지 않을 수도 있다.
- [66] 이외에도 합의진단결과를 도출하기 위한 방식 및 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 각각에 가중치를 부여하는 방식들에 대해서는 다양한 실시 예가 가능할 수 있음을 본 발명의 기술분야의 평균적 전문가가 용이하게 추론할 수 있을 것이다.
- [67] 실시 예에 따라, 상기 합의 진단 시스템(100)은 후술할 바와 같은 스탠다드 진단 시스템(40)으로부터 스탠다드 개별진단결과를 더 수신할 수도 있다. 그리고 수신한 스탠다드 개별진단결과에 더 기초하여 합의진단결과를 출력할 수도

있다.

- [68] 스탠다드 진단 시스템(40)은 골드 스탠다드(gold standard) 학습 데이터로 학습된 진단 시스템을 의미할 수 있다. 골드 스탠다는 학습 데이터는 표준적인 또는 이상적인 어노테이션이 수행된 학습 데이터를 의미할 수 있다. 이러한 골드 스탠다드 학습 데이터는 어느 하나의 주체에 의해 단순히 어노테이션된 것이 아니라, 복수의 주체에 의해 합의된 어노테이션이 수행되거나 또는 다수의 주체에 의해 합의된 기준에 의해 어노테이션이 수행된 데이터를 의미할 수 있다. 일반적으로는 이러한 골드 스탠다드 학습 데이터를 이용하여 학습을 수행하는 것이 바람직할 수 있지만, 실제로는 골드 스탠다드 학습 데이터를 생성하는 것이 많은 비용이 들 수도 있으며, 오히려 골드 스탠다드 학습 데이터를 이용하여 학습된 하나의 진단 시스템을 구축하는 것보다 개별 주체에 의해 학습된 진단 시스템들을 복수 개 구비하여 합의 진단 결과를 구축하는 것이 오히려 더 높은 진단성능을 발휘할 수도 있다. 이는 골드 스탠다드 학습 데이터의 경우는 특정 생체 데이터에 대한 어노테이션에 대해 서로 다른 어노테이션 의견이 있는 경우에는 소수의 의견은 무시된 채 합의된 어노테이션만이 학습에 반영되게 되지만, 본원발명의 기술적 사상과 같이 개별 주체에 의한 어노테이션을 통해 개별적으로 학습된 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30)이 구축된 경우에는 소수의 의견이 적어도 개별 진단 시스템에는 학습에 반영되기 때문이다.
- [69] 더욱이 합의진단을 위한 복수의 진단 시스템들 중에 스탠다드 진단 시스템(40)이 포함될 경우에는 보다 높은 진단 성능을 보일 수 있다.
- [70] 또한, 합의 진단 시스템(100)은 스탠다드 진단 시스템(40)에 대해서는 개별 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 각각에 비해 보다 높은 가중치를 부여하여, 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30)별로 부여된 가중치를 반영하여 합의결과를 도출할 수도 있다.
- [71] 한편, 상기 합의 진단 시스템(100)은 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 또는 스탠다드 진단 시스템(40)와는 별개의 물리적 장치로 구현되는 경우를 도 1에서 도시하고 있지만, 상기 합의 진단 시스템(100)은 상기 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 또는 상기 스탠다드 진단 시스템(40) 중 적어도 하나와 물리적으로 동일한 장치에 구현될 수도 있음은 본 발명의 기술분야의 평균적 전문가가 용이하게 추론할 수 있을 것이다.
- [72] 상기 합의 진단 시스템(100)은 본 명세서에서 정의된 기능을 수행할 수만 있으면 다양한 데이터 처리시스템(예컨대, 컴퓨터, 서버, 스마트폰, 또는 전용장치 등)로 구현될 수 있음은 물론이다.
- [73] 도 2는 본 발명의 실시 예에 따른 진단 시스템들 각각의 개략적인 구성을 설명하기 위한 도면이다.

- [74] 우선 도 2a를 참조하면, 상기 개별 진단 시스템들(예컨대, 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30))은 도 2a에 도시된 바와 같은 구성을 포함할 수 있다. 스탠다드 진단 시스템(40) 역시 상기 개별 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30)의 구성과 동일한 구성으로 구현될 수 있으므로, 본 명세서에서는 개별 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30)에 대해서만 설명하도록 한다. 또한, 개별 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 역시 동일하거나 유사한 구성을 포함할 수 있으므로 본 명세서에서는 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 중 어느 하나의 구성만을 설명하도록 한다.
- [75] 제1진단 시스템(10)은 프로세서(11) 및 저장장치(12)를 포함할 수 있다. 상기 제1진단 시스템(10)은 본 발명의 기술적 사상을 구현하기 위한 연산능력을 가진 데이터 처리장치를 의미하며, 일반적으로 네트워크를 통해 클라이언트가 접속가능한 데이터 처리장치뿐만 아니라 개인용 컴퓨터, 휴대 단말 등과 같이 특정 서비스를 수행할 수 있는 어떠한 장치로도 구현될 수 있음을 본 발명의 기술분야의 평균적 전문가가 용이하게 추론할 수 있을 것이다.
- [76] 상기 프로세서(11)는 본 발명의 기술적 사상을 구현하기 위한 프로그램(12-1)을 구동시킬 수 있는 연산장치를 의미할 수 있으며, 상기 프로세서(11)는 상기 프로그램(12-1)과 본 발명의 기술적 사상에 의해 정의되는 뉴럴 네트워크(Neural Network, 12-2)를 이용해 진단을 수행할 수 있다. 상기 뉴럴 네트워크는 컨벌루션 뉴럴 네트워크일 수 있으며, 생체 데이터(예컨대, 이미지)를 입력받으면 학습된 뉴럴 네트워크를 통해 진단결과를 출력할 수 있다.
- [77] 상기 프로그램(12-1)은 지도학습을 통해 뉴럴 네트워크(12-2)를 학습시키거나 또는 학습된 뉴럴 네트워크(12-2)를 이용하여 진단을 수행할 수 있도록 정의되는 소프트웨어를 의미할 수 있다.
- [78] 상기 저장장치(12)는 상기 프로그램(12-1) 및 뉴럴 네트워크(12-2)를 저장할 수 있는 데이터 저장수단을 의미할 수 있으며, 구현 예에 따라 복수의 저장수단으로 구현될 수도 있다. 또한 상기 저장장치(12)는 상기 제1진단 시스템(10)에 포함된 주기억장치 뿐만 아니라, 상기 프로세서(11)에 포함될 수 있는 임시 저장장치 또는 메모리 등을 포함하는 의미일 수도 있다.
- [79] 상기 진단 시스템(100)은 도 2에서는 어느 하나의 물리적 장치로 구현된 것으로 도시하였지만, 필요에 따라 복수의 물리적 장치가 유기적으로 결합되어 본 발명의 기술적 사상에 따른 제1진단 시스템(10)을 구현할 수 있음을 본 발명의 기술분야의 평균적 전문가가 용이하게 추론할 수 있을 것이다.
- [80] 이하 본 명세서에서 상기 진단 시스템(예컨대, 10)이 소정의 기능을 수행한다고 함은, 진단 시스템(예컨대, 10)에 구비된 프로세서(예컨대, 11)가 상기 프로그램(예컨대, 12-1)을 이용하여 소정의 기능을 수행함을 의미할 수 있음은 물론이다.

- [81] 본 명세서에서 상기 제1진단 시스템(10)이 진단을 수행한다고 함은 생체 데이터를 입력받아 본 명세서에서 정의된 출력 데이터 예컨대, 진단결과를 출력하는 일련의 프로세스를 의미할 수 있다.
- [82] 상기 제1진단 시스템(10)은 생체 데이터를 소정의 단위 유닛별로 입력받을 수 있다. 단위 유닛은 예컨대, 픽셀 단위일 수도 있고 패치, 또는 슬라이드 단위일 수도 있다.
- [83] 제1진단 시스템(10)의 진단결과는 질병의 종류에 따라 단순히 질병의 발현여부 또는 이에 상응하는 값(예컨대, 확률 등)일 수도 있고, 또는 질병이 발현된 경우에는 질병의 상태 정도를 나타내는 상태정보일 수도 있다.
- [84] 예컨대, 후술할 바와 같이 본 발명의 기술적 사상이 전립선 암의 진단에 이용되는 경우, 전립선 암의 진행 정도를 나타내는 지표인 글리슨 패턴(Gleason Pattern) 또는 글리슨 스코어(Gleason Score)가 상기 상태정보에 포함될 수 있다. 예컨대, 글리슨 패턴은 2 내지 5의 값을 가지며, 숫자가 클수록 전립선 암이 발현된 정도가 심한 것을 나타낸다. 따라서 상기 상태정보는 진단의 대상이 되는 단위 유닛에 해당하는 생체조직이 글리슨 패턴의 특정 값(예컨대, 3, 4, 또는 5)에 해당할 확률에 상응하는 정보 또는 노멀(즉, 질병이 발현되지 않은 경우)에 해당할 확률에 상응하는 정보를 포함할 수 있다.
- [85] 어떠한 경우든 상기 제1진단 시스템(10)은 학습된 뉴럴 네트워크(12-2)를 통해 진단결과를 출력할 수 있다.
- [86] 한편, 본 발명의 기술적 사상에 따른 합의 진단 시스템(100)은 도 2b에 도시된 구성을 포함할 수 있다.
- [87] 합의 진단 시스템(100) 역시 프로세서(110) 및 프로그램(121)이 저장된 저장장치(120)를 포함할 수 있다.
- [88] 상기 합의 진단 시스템(100)은 유무선 네트워크를 통해 개별 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 및/또는 스탠다드 진단 시스템(40)과 연결되어 본 발명의 기술적 사상을 구현하기 위해 필요한 정보들을 송수신할 수 있다. 구현 예에 따라 상기 합의 진단 시스템(100)은 상기 개별 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 또는 스탠다드 진단 시스템(40)에 설치되어 구현될 수도 있다. 이러한 경우 상기 합의 진단 시스템(100)은 개별 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 또는 스탠다드 진단 시스템(40)의 하드웨어를 공유할 수 있다.
- [89] 상기 합의 진단 시스템(100)의 저장장치(120)에 저장된 프로그램(121)은 개별 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30) 및/또는 스탠다드 진단 시스템(40)으로부터 각각 진단결과(예컨대, 개별진단결과 1 내지 N 및/또는 스탠다드 개별진단결과)를 수신하고 확인할 수 있다.
- [90] 상기 합의 진단 시스템(100)은 서로 다른 어노테이션 주체들에 의해 어노테이션된 학습 데이터 세트로 각각 학습된 진단 시스템들(예컨대, 진단

시스템 1 내지 진단 시스템 N, 10 내지 30)만으로부터 개별진단결과를 수신할 수도 있지만, 전술한 바와 같이 스탠다드 진단 시스템(40)으로부터 스탠다드 개별진단결과를 더 수신할 수도 있다.

- [91] 어떠한 경우든 상기 합의 진단 시스템(100)은 수신한 진단결과들에 기초하여 합의진단결과를 생성하여 출력할 수 있다.
- [92] 상기 합의 진단 시스템(100)이 합의진단결과를 생성하는 일 예는 도 3을 참조하여 설명하도록 한다.
- [93] 도 3은 본 발명의 실시 예에 따라 개별 진단 결과에 기초한 합의를 도출하는 개념을 설명하기 위한 도면이다.
- [94] 도 3을 참조하면, 본 발명의 기술적 사상을 구현하기 위해 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30)이 구현될 수 있으며, 합의 진단 시스템(100)은 각각의 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30)로부터 동일한 생체 데이터에 대한 진단결과를 수신할 수 있다. 또한, 도 3에는 합의 진단 시스템(100)이 스탠다드 진단 시스템(STD, 40)으로부터 스탠다드 개별진단결과를 더 수신하는 경우를 예시적으로 설명하고 있다.
- [95] 도 3에서는 진단결과가 특정 질병의 발현여부를 나타내는 경우를 예시하고 있지만, 뉴럴 네트워크(12-2)의 구현 예에 따라 다양한 형태의 정보가 진단결과로 출력될 수 있음은 전술한 바와 같다.
- [96] 일 예에 의하면, 상기 합의 진단 시스템(100)은 단순히 진단결과와 종류들(예컨대, 질병이 발현인 O, 질병의 미발현인 X) 중 더 많은 수를 갖는 종류를 합의진단결과로 특정할 수도 있다.
- [97] 구현 예에 따라 상기 합의 진단 시스템(100)은 각 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30, 스탠다드 진단 시스템(40))에 대해 가중치를 부여할 수도 있다.
- [98] 이러한 가중치는 다수의 생체 데이터에 대해 합의진단결과를 도출하면서, 합의진단결과와 다른 개별진단결과를 출력한 횟수에 기초하여 조절될 수 있다. 예컨대, 진단 시스템 2는 가중치가 0.9일 수 있고, 진단 시스템 1은 가중치가 1일 수 있는데 이는 진단 시스템 2는 합의진단결과와 다른 개별진단결과를 출력한 경우가 적어도 한 번 존재했음을 의미할 수 있다.
- [99] 상기 합의 진단 시스템(100)은 가중치가 웨이트 팩터로 반영된 진단결과에 기초하여 합의진단결과를 생성할 수 있다. 예컨대, 특정 진단 시스템이 X로 진단결과를 출력했는데 상기 특정 진단 시스템의 가중치가 a(예컨대, 0.9)라면, 합의진단결과를 도출하기 위한 과정에서 상기 특정 진단 시스템의 진단결과는 a개(0.9개)의 X로 판단될 수 있다. 이처럼 가중치를 반영하여 보다 많은 진단결과와 종류(예컨대, O 또는 X)를 합의진단결과로 판단할 수 있다.
- [100] 상기 합의 진단 시스템(100)은 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30 및/또는 스탠다드 진단 시스템(40))별로, 단순히

합의진단결과와 다른 개별진단결과를 출력한 경우마다 미리 정해진 값(예컨대, 0.1)을 낮춤으로써 가중치를 조절할 수도 있다.

- [101] 또는 전술한 바와 같이 특정 진단 시스템(예컨대, 진단 시스템 2)이 합의진단결과(예컨대, O)와 다른 개별진단결과(예컨대, X)를 출력한 경우라도, 상기 다른 개별진단결과(예컨대, X)를 출력한 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30 및/또는 스탠다드 진단 시스템(40))의 개수 또는 비중(전체 진단 시스템들 개수 대비 상기 특정 진단 시스템(예컨대, 진단 시스템 2)과 동일한 종류의 진단결과를 출력한 진단 시스템의 개수)를 더 고려하여 가중치를 조절할 수도 있다. 예컨대, 10개의 진단 시스템 중 1개만 다른 진단결과를 출력한 경우에는 해당 시스템의 가중치를 제1값(예컨대, 0.1)만큼 조절하고, 10개의 진단 시스템 중 일정 개수(예컨대, 3개) 이상 또는 일정 비중(예컨대, 30%) 이상 다른 진단결과를 출력한 경우에는 해당 시스템들 각각의 가중치를 제2값(예컨대, 0.03)만큼만 조절할 수도 있다. 또는 일정 개수 또는 일정 비중 이상의 경우에는 합의진단결과와 다른 진단 결과를 출력하였어도 해당 시스템의 가중치를 아예 조절하지 않을 수도 있다.
- [102] 즉, 후자의 경우는 실제로 판단에 따라 다른 진단결과로 해석될 수 있는 여지가 있으므로, 해당 진단 시스템의 신뢰도에 해당하는 가중치를 적응적으로 조절함으로써 전체적으로 보다 신뢰성 있는 진단결과를 도출할 수 있다.
- [103] 또한, 스탠다드 진단 시스템(40)의 진단결과를 더 이용할 경우에는, 상기 합의 진단 시스템(100)은 스탠다드 진단 시스템(40)의 가중치는 다른 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30)에 비해 높은 가중치를 부여할 수도 있다.
- [104] 한편, 본 발명의 기술적 사상에 따른 합의 진단 방법의 경우에는 서로 다른 학습 데이터 세트로 학습된 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30 및/또는 스탠다드 진단 시스템(40))을 이용하는 경우에도 유용하게 진단결과를 도출할 수 있다. 이는 후술할 바와 같이 합의 진단 시스템(100)이 합의진단결과와 다른 진단결과를 출력하는 진단 시스템에 대해 재학습을 수행하는 경우에 더욱 유용할 수 있다.
- [105] 도 4는 본 발명의 실시 예에 따른 학습 데이터의 분할 어노테이션 개념을 설명하기 위한 도면이다.
- [106] 도 4를 참조하면, 본 발명의 기술적 사상에 따른 합의 진단 방법을 적용할 경우, 전술한 바와 같이 어노테이터별로 서로 다르게 어노테이션된 학습 데이터를 이용하여 각각 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30 및/또는 스탠다드 진단 시스템(40))들을 학습시키는 경우 뿐만 아니라, 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30 및/또는 스탠다드 진단 시스템(40))별로 서로 다른 학습 데이터를 이용하여 학습하는 경우에도 유용할 수 있다.
- [107] 예컨대, 도 4에 도시된 바와 같이 학습된 엔진 즉, 진단 시스템의 성능은 학습에

이용되는 학습 데이터를 어노테이션하는 어노테이터의 성향에 따라 달라질 수도 있지만, 동일한 어노테이터가 어노테이션을 수행하더라도 학습 데이터를 어떤 것을 사용하는지에 따라서도 영향을 받게 된다.

- [108] 따라서 도 4에 도시된 바와 같이 진단 시스템 1은 학습 데이터 세트 1로 학습된 시스템이고, 진단 시스템 2는 학습 데이터 세트 2로 학습된 시스템이고, 진단 시스템 3은 학습 데이터 세트 3으로 학습된 시스템이며, 진단 시스템 4는 학습 데이터 세트 4로 학습된 시스템일 수 있다.
- [109] 각각의 학습 데이터 세트는 중복된 학습 데이터가 존재할 수도 있고 그렇지 않을 수도 있으며, 학습 데이터 세트별로 어노테이터가 동일할 수도 있고 다를 수도 있다.
- [110] 어떠한 경우든 학습 데이터 세트별로 각각 학습된 진단 시스템의 진단결과는 차이가 있을 수 있는데, 이를 위해서 종래에는 학습 데이터 세트를 선정할때부터 다양한 케이스들을 모두 포함하는 광범위한 학습 데이터 세트를 결정하는데 상당한 노력과 비용을 소요하여야 했다.
- [111] 하지만 본 발명의 기술적 사상에 의하면, 진단 시스템들 각각을 학습시키기 위해 어느 정도 충분한 정도의 학습 데이터 세트로만 학습을 한다면, 서로 다른 학습 데이터 세트로 학습된 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30 및/또는 스탠다드 진단 시스템(40))이라 하더라도 합의 진단 결과를 통해 보다 정확한 진단결과를 도출할 수 있게 된다.
- [112] 즉, 학습 데이터 세트 자체의 편향성이 합의 진단 결과를 도출하는 과정에서 상당히 무마될 수 있는 효과가 있다.
- [113] 더욱이 특정 생체 데이터에 대해 합의진단결과와 다른 진단결과를 출력한 진단 시스템에 대해서는 상기 특정 생체 데이터를 합의진단결과로 어노테이션하여 재학습을 수행할 수 있고, 이러한 경우 어노테이터에 의한 편향성 및/또는 학습 데이터에 의한 편향성은 시간이 갈수록 무마될 수 있어서 궁극적으로 개별 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30 및/또는 스탠다드 진단 시스템(40))의 성능 역시 향상시킬 수 있는 효과가 있다.
- [114] 이러한 일 예는 도 5를 참조하여 설명하도록 한다.
- [115] 도 5는 본 발명의 실시 예에 따라 합의 진단결과를 이용하여 개별 진단 시스템을 재학습하는 개념을 설명하기 위한 도면이다.
- [116] 도 5를 참조하면, 예컨대, 특정 생체 데이터가 각각의 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30 및/또는 스탠다드 진단 시스템(40))로 입력되고, 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30 및/또는 스탠다드 진단 시스템(40)) 각각은 합의 진단 시스템(100)으로 개별진단결과를 출력할 수 있다. 그리고 이때 합의 진단 시스템(100)은 합의진단결과로 O를 출력했으며, 진단 시스템 2(20)는 개별진단결과로 X를 출력했을 수 있다.
- [117] 이러한 경우 상기 합의 진단 시스템(100)은 합의진단결과와 다른

개별진단결과를 출력한 진단 시스템(예컨대, 진단 시스템 2)에 대한 정보, 이때의 상기 특정 생체 데이터, 및 합의진단결과(예컨대, O)를 저장할 수 있다.

[118] 그리고 저장된 정보를 이용해 상기 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30 및/또는 스탠다드 진단 시스템(40))을 재학습시킬 수 있다.

[119] 예컨대, 소정의 시간 경과 후에 상기 진단 시스템 2는 상기 합의진단결과(예컨대, O)로 라벨링 된 상기 특정 생체 데이터를 포함하는 재학습 데이터 세트를 이용하여 재학습될 수 있다.

[120] 즉, 본 발명의 기술적 사상에 의하면, 합의진단결과와 다른 진단결과를 출력했던 생체 데이터들을 모아서 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30 및/또는 스탠다드 진단 시스템(40)) 각각을 다시 재학습할 수 있다. 즉, 합의 진단 시스템(100)에 의해 저장된, 합의진단결과와 다른 진단결과를 출력하게 한 생체 데이터들이 학습 데이터로써 개별 진단 시스템에 피드백될 수 있다. 이러한 경우 진단 시스템들(예컨대, 진단 시스템 1 내지 진단 시스템 N, 10 내지 30 및/또는 스탠다드 진단 시스템(40)) 각각이 갖는 편향성(예컨대, 어노테이터의 성향에 따라 또는 학습 데이터 자체의 편향성에 따라)이 개선되게 되는 효과가 있다.

[121] 본 발명의 실시 예에 따른 합의 진단 방법은 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 하드 디스크, 플로피 디스크, 광 데이터 저장장치 등이 있다. 또한, 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다. 그리고 본 발명을 구현하기 위한 기능적인(functional) 프로그램, 코드 및 코드 세그먼트들은 본 발명이 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있다.

[122] 본 발명은 도면에 도시된 일 실시 예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시 예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

[123]

산업상 이용가능성

[124] 본 발명은 "지도학습기반의 합의 진단방법 및 그 시스템"에 이용될 수 있다.

청구범위

- [청구항 1] 지도학습을 통해 학습된 진단 시스템을 이용해 질병을 진단하는 방법에 있어서,
 합의 진단 시스템이 N (N 은 2이상의 정수) 개의 진단 시스템들- N 개의 진단시스템들 각각은 서로 다른 어노테이션(annotation) 주체에 의해 어노테이션된 학습 데이터로 각각 학습된 시스템들임- 각각이 소정의 생체 데이터를 각각 입력받아 출력하는 N 개의 개별진단결과들을 확인하는 단계;
 상기 합의 진단 시스템이 확인한 상기 개별진단결과들에 기초하여 상기 생체 데이터의 합의진단결과를 출력하는 단계를 포함하는 지도학습기반 합의 진단방법.
- [청구항 2] 제1항에 있어서, 상기 지도학습기반 합의 진단방법은,
 골드 스탠다드(gold standard) 학습 데이터로 학습된 스탠다드 진단 시스템이 상기 생체 데이터를 입력받아 출력하는 스탠다드 개별진단결과를 확인하는 단계를 더 포함하며,
 상기 합의 진단 시스템이 확인한 상기 개별진단결과들에 기초하여 상기 생체 데이터의 합의진단결과를 출력하는 단계는,
 상기 개별진단결과들 및 상기 스탠다드 개별진단결과에 기초하여 상기 합의진단결과를 출력하는 단계를 포함하는 지도학습기반 합의 진단방법.
- [청구항 3] 제1항에 있어서, 상기 합의 진단 시스템이 확인한 상기 개별진단결과들에 기초하여 상기 생체 데이터의 합의진단결과를 출력하는 단계는,
 상기 진단 시스템들 각각에 부여된 가중치에 기초하여 상기 합의진단결과를 출력하는 것을 특징으로 하며,
 상기 가중치는,
 특정 진단 시스템이 합의진단결과와 다른 개별진단결과를 출력한 횟수에 기초하여 결정되는 지도학습기반 합의 진단방법.
- [청구항 4] 제3항에 있어서, 상기 가중치는,
 상기 특정 진단 시스템이 합의진단결과와 다른 개별진단결과를 출력한 경우, 상기 전체 개별진단결과들 중 상기 다른 개별진단결과를 출력한 타 진단 시스템의 개수 또는 비중에 더 기초하여 결정되는 지도학습기반 합의 진단방법.
- [청구항 5] 제2항에 있어서, 상기 합의 진단 시스템이 확인한 상기 개별진단결과들에 기초하여 상기 생체 데이터의 합의진단결과를 출력하는 단계는,
 상기 진단 시스템들 및 상기 스탠다드 진단 시스템 각각에 부여된 가중치에 기초하여 상기 합의진단결과를 출력하는 것을 특징으로 하며,
 상기 스탠다드 진단 시스템에는 상기 진단 시스템들에 비해 높은 가중치가 부여된 것을 특징으로 하는 지도학습기반 합의 진단방법.

- [청구항 6] 제1항에 있어서, 상기 진단 시스템들은, 소정의 학습 데이터 세트가 N개로 분할된 분할 학습데이터 세트들-임의의 분할 학습데이터 세트는 타 분할 학습데이터 세트에는 포함되지 않은 학습데이터를 적어도 한 개 포함하며 각각이 서로 다른 주체에 의해 어노테이션된 학습 데이터 세트임-을 각각 학습하여 질병의 진단결과를 출력하는 시스템인 것을 특징으로 하는 지도학습기반 합의 진단방법.
- [청구항 7] 제1항에 있어서, 상기 지도학습기반 합의 진단방법은, 특정 생체 데이터의 합의진단결과와 상기 특정 생체 데이터의 개별진단결과가 상이한 진단 시스템이, 상기 합의진단결과로 어노테이션된 상기 특정 생체 데이터를 포함하는 학습 데이터로 재훈련되는 단계를 더 포함하는 지도학습기반 합의 진단방법.
- [청구항 8] 데이터 처리장치에 설치되며 제1항 내지 제7항 중 어느 한 항에 기재된 방법을 수행하기 위한 컴퓨터 판독가능한 기록매체에 저장된 컴퓨터 프로그램.
- [청구항 9] 프로세서; 및 프로그램이 저장된 저장장치를 포함하며, 상기 프로세서에 의해 실행되는 상기 프로그램에 의해 상기 제1항 내지 제6항 중 어느 한 항에 기재된 방법이 수행되는 데이터 처리 시스템.
- [청구항 10] 프로세서; 상기 프로세서에 의해 실행되는 프로그램이 저장된 저장장치를 포함하며, 상기 프로그램은, N(N은 2이상의 정수) 개의 진단 시스템들-N개의 진단시스템들 각각은 서로 다른 어노테이션(annotation) 주체에 의해 어노테이션된 학습 데이터로 각각 학습된 시스템들임- 각각이 소정의 생체 데이터를 각각 입력받아 출력하는 N개의 개별진단결과들을 확인하고, 확인한 상기 개별진단결과들에 기초하여 상기 생체 데이터의 합의진단결과를 출력하는 합의 진단시스템.
- [청구항 11] 제10항에 있어서, 상기 프로그램은, 골드 스탠다드(gold standard) 학습 데이터로 학습된 스탠다드 진단 시스템이 상기 생체 데이터를 입력받아 출력하는 스탠다드 개별진단결과를 더 확인하고, 상기 개별진단결과들 및 상기 스탠다드 개별진단결과에 기초하여 상기 합의진단결과를 출력하는 합의 진단시스템.
- [청구항 12] 제10항에 있어서, 상기 프로그램은, 상기 진단 시스템들 각각에 부여된 가중치에 기초하여 상기 합의진단결과를 출력하는 것을 특징으로 하며,

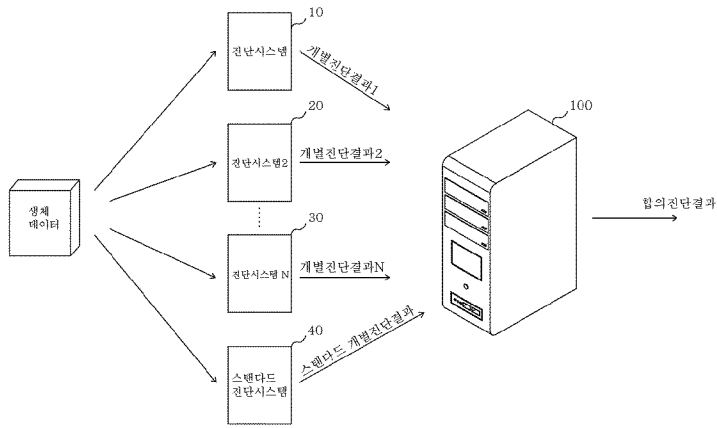
상기 가중치는,
상기 합의진단결과와 상기 진단 시스템들 각각의 개별진단결과가 다른
경우의 횟수에 기초하여 결정되는 합의 진단시스템.

[청구항 13] 제12항에 있어서, 상기 가중치는,
상기 특정 진단 시스템이 합의진단결과와 다른 개별진단결과를 출력한
경우, 상기 전체 개별진단결과들 중 상기 다른 개별진단결과를 출력한 타
진단 시스템의 개수 또는 비중에 더 기초하여 결정되는 지도학습기반
합의 진단시스템.

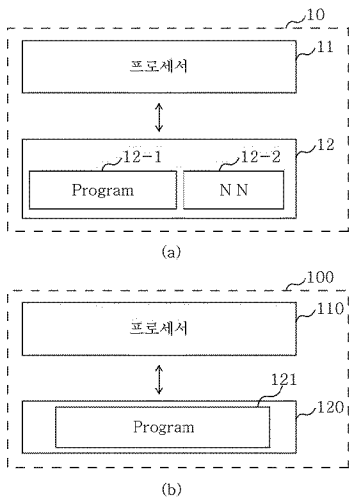
[청구항 14] 제11항에 있어서, 상기 프로그램은,
상기 진단 시스템들 및 상기 스탠다드 진단 시스템 각각에 부여된
가중치에 기초하여 상기 합의진단결과를 출력하는 것을 특징으로 하며,
상기 스탠다드 진단 시스템에는 상기 진단 시스템들에 비해 높은
가중치가 부여된 것을 특징으로 하는 합의 진단시스템.

[청구항 15] 제10항에 있어서, 상기 합의 진단시스템의 진단에 이용되는 생체 데이터
중 특정 생체 데이터는,
상기 합의진단결과와 상기 진단 시스템들 중 특정 진단시스템의
개별진단결과가 상이하며,
상기 합의진단결과로 어노테이션된 상기 특정 생체 데이터가 포함된
학습 데이터로 상기 특정 진단시스템이 재훈련되는 것을 특징으로 하는
지도학습기반 합의 진단시스템.

[도1]



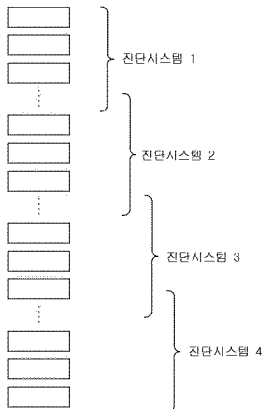
[도2]



[도3]

진단시스템	진단결과	가중치
1	O	1
2	X	0.9
3	O	1
⋮	⋮	⋮
N	X	0.5
STD	O	2

[도4]



[도5]

