(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2019 년 12 월 26 일 (26.12.2019)

WIPO | PCT

(10) 국제공개번호
WO 2019/245186 A1

(51) 국제특허분류:
G06N 20/00 (2019.01)

(21) 국제출원번호: PCT/KR2019/006500

(22) 국제출원일: 2019 년 5 월 30 일 (30.05.2019)

(25) 출원언어: 한국어

(26) 공개언어: 한국어

(30) 우선권정보:
10-2018-0070374 2018 년 6 월 19 일 (19.06.2018) KR

(71) 출원인: 삼성전자주식회사 (**SAMSUNG ELECTRONICS CO., LTD.**) [KR/KR]; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR).

(72) 발명자: **박치연 (PARK, Chiyoun)**; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). **김재덕 (KIM, Jaedeok)**; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). **정현주 (JUNG, Hyunjoo)**; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). **최인권 (CHOI, Inkwon)**; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR).

(74) 대리인: **김태헌 등 (KIM, Tae-hun** et al.); 06626 서울시 서초구 강남대로 343 신덕빌딩 9층, Seoul (KR).

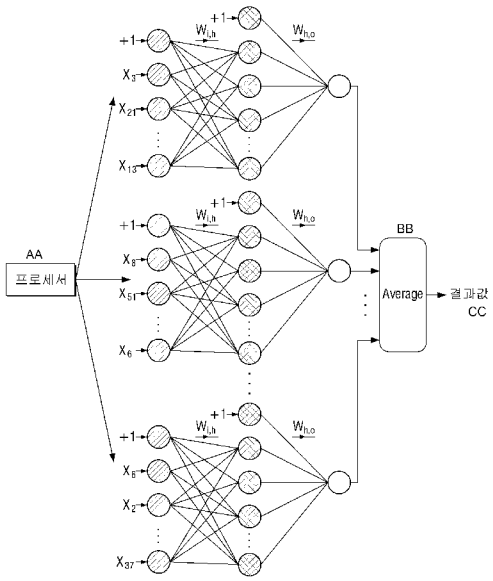
(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR,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

(54) **Title:** ELECTRONIC DEVICE AND CONTROL METHOD THEREOF

(54) **발명의 명칭:** 전자 장치 및 그의 제어 방법



AA ... Processor
BB ... Average
CC ... Result value

(57) **Abstract:** Disclosed are an electronic device for generating an artificial intelligence model and a control method thereof.The control method of an electronic device, according to the present disclosure, comprises the steps of: when a user command for using a plurality of artificial intelligence models is input, generating a plurality of artificial intelligence models; and inputting input data to the generated artificial intelligence models to obtain output data for the input data, wherein the generating step includes generating layers at corresponding positions among layers included in each of the plurality of artificial intelligence models, on the basis of a reference value for obtaining layer parameters for the layers at corresponding positions among the plurality of layers included in each of the plurality of artificial intelligence models.

(57) **요약서:** 인공 지능 모델을 생성하기 위한 전자 장치 및 그의 제어 방법이 개시된다. 본 개시에 따른 전자 장치의 제어 방법은, 복수의 인공 지능 모델을 이용하기 위한 사용자 명령이 입력되면, 복수의 인공 지능 모델을 생성하는 단계, 생성된 인공 지능 모델에 입력 데이터를 입력하여 입력 데이터에 대한 출력 데이터를 획득하는 단계를 포함하며, 생성하는 단계는, 복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 레이어 파라미터를 획득하기 위한 기준값을 바탕으로, 복수의 인공 지능 모델 각각에 포함된 레이어 중 대응되는 위치의 레이어들을 생성한다.

WO 2019/245186 A1

명세서

발명의 명칭: 전자 장치 및 그의 제어 방법

기술분야

[1] 본 개시는 전자 장치 및 그의 제어 방법에 관한 것으로, 보다 상세하게는 사용자 단말과 같은 전자 장치에서 인공 지능 모델을 효율적으로 사용하기 위한 방법에 관한 것이다.

[2] 또한, 본 개시는 기계학습 알고리즘을 활용하여 인간 두뇌의 인지, 판단 등의 기능을 모사하는 인공 지능(Artificial Intelligence, AI) 시스템 및 그 응용에 관한 것이다.

배경기술

[3] 인공 지능(Artificial Intelligence, AI) 시스템은 인간 수준의 지능을 구현하는 컴퓨터 시스템이며, 기존 규칙 기반 스마트 시스템과 달리 기계가 스스로 학습하고 판단하며 똑똑해지는 시스템이다. 인공 지능 시스템은 사용할수록 인식률이 향상되고 사용자 취향을 보다 정확하게 이해할 수 있게 되어, 기존 규칙 기반 스마트 시스템은 점차 딥러닝 기반 인공 지능 시스템으로 대체되고 있다.

[4] 인공 지능 기술은 기계학습(딥러닝) 및 기계 학습을 활용한 요소 기술들로 구성된다.

[5] 기계 학습은 입력 데이터들의 특징을 스스로 분류/학습하는 알고리즘 기술이며, 요소 기술은 딥러닝 등의 기계학습 알고리즘을 활용하는 기술로서, 언어적 이해, 시각적 이해, 추론/예측, 지식 표현, 동작 제어 등의 기술 분야로 구성된다.

[6] 인공 지능 기술이 응용되는 다양한 분야는 다음과 같다. 언어적 이해는 인간의 언어/문자를 인식하고 응용/처리하는 기술로서, 자연어 처리, 기계 번역, 대화 시스템, 질의 응답, 음성 인식/합성 등을 포함한다. 시각적 이해는 사물을 인간의 시각처럼 인식하여 처리하는 기술로서, 객체 인식, 객체 추적, 영상 검색, 사람 인식, 장면 이해, 공간 이해, 영상 개선 등을 포함한다. 추론 예측은 정보를 판단하여 논리적으로 추론하고 예측하는 기술로서, 지식/확률 기반 추론, 최적화 예측, 선호 기반 계획, 추천 등을 포함한다. 지식 표현은 인간의 경험 정보를 지식데이터로 자동화 처리하는 기술로서, 지식 구축(데이터 생성/분류), 지식 관리(데이터 활용) 등을 포함한다. 동작 제어는 차량의 자율 주행, 로봇의 움직임을 제어하는 기술로서, 움직임 제어(항법, 충돌, 주행), 조작 제어(행동 제어) 등을 포함한다.

[7] 한편, 종래의 인공 지능 기술은 복수의 인공 지능 모델을 결합하여 사용되어 많은 학습 시간 및 큰 저장 공간을 필요로 하였다. 따라서, 종래의 인공 지능 기술은 큰 저장 공간 및 높은 연산 수행이 가능한 외부 서버에서 이루어지는

경우가 많았으며, 스마트폰과 같은 사용자 단말 장치에서 효율적으로 인공 지능 모델을 이용하기 위한 방법에 대한 논의가 필요하게 되었다.

발명의 상세한 설명

기술적 과제

[8] 본 개시는 상술한 문제점을 해결하기 위한 것으로, 한정된 저장 공간에서 인공 지능 모델을 사용하기 위한 방법에 관한 것이다

과제 해결 수단

[9] 상기 목적을 달성하기 위한 인공 지능 모델을 생성하기 위한 전자 장치의 제어 방법은, 복수의 인공 지능 모델을 이용하기 위한 사용자 명령이 입력되면, 상기 복수의 인공 지능 모델을 생성하는 단계, 상기 생성된 인공 지능 모델에 입력 데이터를 입력하여 상기 입력 데이터에 대한 출력 데이터를 획득하는 단계를 포함하며, 상기 생성하는 단계는, 상기 복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 레이어 파라미터를 획득하기 위한 기준값을 바탕으로, 상기 복수의 인공 지능 모델 각각에 포함된 레이어 중 상기 대응되는 위치의 레이어들을 생성한다.

[10] 이때, 상기 기준값은, 상기 대응되는 위치의 레이어들에 대한 레이어 파라미터의 평균값 및 분산값을 바탕으로 결정될 수 있다.

[11] 이때, 상기 복수의 레이어 파라미터의 분산값이 기 설정된 값 이하인 경우, 상기 기준값은 상기 복수의 레이어 파라미터의 평균값을 바탕으로 결정될 수 있다.

[12] 이때, 상기 복수의 레이어 파라미터의 분산값이 기 설정된 값 이상인 경우, 상기 복수의 레이어 파라미터간의 분포를 판단하는 단계, 상기 분포가 기 설정된 조건을 만족하면, 상기 복수의 레이어 파라미터의 분산값을 바탕으로 기준값을 획득하는 단계를 더 포함할 수 있다.

[13] 이때, 상기 분포가 상기 기 설정된 조건을 만족하지 않는 경우, 상기 복수의 레이어 파라미터의 평균값 및 상기 복수의 레이어 파라미터의 분산값을 바탕으로 기준값을 획득하는 단계를 더 포함할 수 있다.

[14] 이때, 상기 생성하는 단계는, 상기 전자 장치의 메모리 용량을 바탕으로 생성되는 인공 지능 모델의 개수를 판단하는 단계를 더 포함할 수 있다.

[15] 이때, 상기 제어 방법은, 상기 생성된 인공 지능 모델에 입력값이 입력되면, 상기 복수의 레이어를 앙상블(Ensemble)하여 결과값을 출력하는 단계를 더 포함할 수 있다.

[16] 한편, 상술한 목적을 달성하기 위한 본 개시의 또 다른 실시예에 따른 인공 지능 모델을 생성하기 위한 전자 장치는, 통신부, 메모리 및 외부 서버로부터, 상기 통신부를 통해 복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 레이어 파라미터를 획득하기 위한 기준값을 수신하는 프로세서를 포함하고, 상기 프로세서는, 복수의 인공 지능 모델을 이용하기 위한 사용자 명령이 입력되면, 상기 복수의 인공 지능 모델을 생성하고, 상기 생성된

인공 지능 모델에 입력 데이터를 입력하여 상기 입력 데이터에 대한 출력 데이터를 획득하고,상기 복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 레이어 파라미터를 획득하기 위한 기준값을 바탕으로, 상기 복수의 인공 지능 모델 각각에 포함된 레이어 중 상기 대응되는 위치의 레이어들을 생성한다.

[17] 이때, 상기 기준값은, 상기 대응되는 위치의 레이어들에 대한 레이어 파라미터의 평균값 및 분산값을 바탕으로 결정될 수 있다.

[18] 이때, 상기 프로세서는, 상기 복수의 레이어 파라미터의 분산값이 기 설정된 값 이하인 경우, 상기 기준값은 상기 복수의 레이어 파라미터의 평균값을 바탕으로 결정할 수 있다.

[19] 이때, 상기 프로세서는, 상기 복수의 레이어 파라미터의 분산값이 기 설정된 값 이상인 경우, 상기 복수의 레이어 파라미터간의 분포를 판단하고, 상기 분포가 기 설정된 조건을 만족하면, 상기 복수의 레이어 파라미터의 분산값을 바탕으로 기준값을 획득할 수 있다.

[20] 이때, 상기 프로세서는, 상기 분포가 상기 기 설정된 조건을 만족하지 않는 경우, 상기 복수의 레이어 파라미터의 평균값 및 상기 복수의 레이어 파라미터의 분산값을 바탕으로 기준값을 획득할 수 있다.

[21] 이때, 상기 프로세서는, 상기 메모리 용량을 바탕으로 생성되는 인공 지능 모델의 개수를 판단할 수 있다.

[22] 이때, 상기 프로세서는, 상기 생성된 인공 지능 모델에 입력값이 입력되면, 상기 복수의 레이어를 앙상블(Ensemble)하여 결과값을 출력할 수 있다.

[23] 한편, 상술한 목적을 달성하기 위한 본 개시의 또 다른 실시예에 따른 서버의 제어 방법은, 복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 복수의 레이어 파라미터를 획득하는 단계, 상기 획득된 복수의 레이어 파라미터를 바탕으로, 상기 대응되는 위치의 새로운 레이어에 대한 레이어 파라미터를 생성하기 위한 기준값을 획득하는 단계 및 상기 획득된 기준값을 사용자 단말로 전송하는 단계를 포함한다.

[24] 이때, 상기 기준값은, 상기 대응되는 위치의 레이어들에 대한 레이어 파라미터의 평균값 및 분산값을 바탕으로 결정될 수 있다.

[25] 이때, 상기 복수의 레이어 파라미터의 분산값이 기 설정된 값 이하인 경우, 상기 기준값은 상기 복수의 레이어 파라미터의 평균값을 바탕으로 결정될 수 있다.

[26] 이때, 상기 복수의 레이어 파라미터의 분산값이 기 설정된 값 이상인 경우, 상기 복수의 레이어 파라미터간의 분포를 판단하는 단계, 상기 분포가 기 설정된 조건을 만족하면, 상기 복수의 레이어 파라미터의 분산값을 바탕으로 기준값을 획득하는 단계를 포함할 수 있다.

[27] 이때, 상기 분포가 상기 기 설정된 조건을 만족하지 않는 경우, 상기 복수의 레이어 파라미터의 평균값 및 상기 복수의 레이어 파라미터의 분산값을 바탕으로 기준값을 획득하는 단계를 포함할 수 있다.

발명의 효과

[28] 상기와 같은 본 개시의 다양한 실시예에 따라, 전자 장치는 한정된 저장 공간에서도 인공 지능 모델을 효율적으로 사용할 수 있다.

도면의 간단한 설명

[29] 도 1a 내지 도 1c는 본 개시의 일 실시예에 따른 종래 서버에서 동작하는 인공 지능 모델의 결과 출력 방법을 설명하기 위한 예시도이다.

[30] 도 2는 본 개시의 일 실시예에 따른 전자 장치의 구성을 간략히 도시한 블록도이다.

[31] 도 3은 본 개시의 일 실시예에 따른 전자 장치(200)의 구성을 상세히 도시한 블록도이다.

[32] 도 4a 및 도 4b는 본 개시의 일 실시예에 따른, 메모리 저장공간에 따른 인공 지능 모델 생성 방법을 설명하기 위한 예시도이다.

[33] 도 5a 내지 도 5c는 본 개시의 일 실시예에 따른, 사용자 명령에 따라 생성되는 인공 지능 모델을 조절할 수 있는 UI를 설명하기 위한 예시도이다.

[34] 도 6a 내지 도 6e는 본 개시의 일 실시예에 따른 레이어의 앙상블 방법을 설명하기 위한 예시도이다.

[35] 도 7은 본 개시의 일 실시예에 따른 서버의 구성을 간략히 도시한 블록도이다.

[36] 도 8은 본 개시의 일 실시예에 따른 전자 장치의 동작으로 설명하기 위한 흐름도이다.

[37] 도 9는 본 개시의 일 실시예에 따른 서버의 동작을 설명하기 위한 흐름도이다.

[38] 도 10은 본 개시의 일 실시예에 따른 서버 및 전자 장치의 동작을 설명하기 위한 시스템도이다.

발명의 실시를 위한 최선의 형태

[39] 본 명세서에서 사용되는 용어에 대해 간략히 설명하고, 본 개시에 대해 구체적으로 설명하기로 한다.

[40] 본 개시의 실시 예에서 사용되는 용어는 본 개시에서의 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 개시의 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서 본 개시에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 개시의 전반에 걸친 내용을 토대로 정의되어야 한다.

[41] 본 개시의 실시 예들은 다양한 변환을 가할 수 있고 여러 가지 실시 예를 가질 수 있는바, 특정 실시 예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나 이는 특정한 실시 형태에 대해 범위를 한정하려는 것이 아니며, 발명된 사상 및 기술 범위에 포함되는 모든 변환, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 실시 예들을 설명함에 있어서 관련된 공지

기술에 대한 구체적인 설명이 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

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

[43] 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "구성되다" 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.

[44] 본 개시에서 "모듈" 혹은 "부"는 적어도 하나의 기능이나 동작을 수행하며, 하드웨어 또는 소프트웨어로 구현되거나 하드웨어와 소프트웨어의 결합으로 구현될 수 있다. 또한, 복수의 "모듈" 혹은 복수의 "부"는 특정한 하드웨어로 구현될 필요가 있는 "모듈" 혹은 "부"를 제외하고는 적어도 하나의 모듈로 일체화되어 적어도 하나의 프로세서(미도시)로 구현될 수 있다.

[45] 본 개시에서 앙상블 모델(Ensemble Model)이라 함은 적어도 두개 이상의 인공 지능 모델의 출력을 연산하여 이용하는 모델을 의미할 수 있다. 이때, 복수개의 모델을 결합하여 이용하는 것을 앙상블이라 할 수 있다. 앙상블 모델을 구성하는 각각의 인공 지능 모델을 개별 모델 또는 인공 지능 모델이라고 할 수 있다 앙상블 모델은 각각의 인공 지능 모델의 출력값을 Averaging, Weight Sum, Boosting 방식 등을 적용하여 최종 출력값을 출력할 수 있다.

[46] 본 개시에서 레이어(Layer)라 함은 다층 신경망 구조에서 하나의 층을 의미할 수 있다. 즉, 복수개의 앙상블 모델에서 하나의 층을 레이어라 정의할 수 있다. 일반적으로 레이어는 입력 벡터 x와 출력 벡터 y에 대하여 y=f(Wx+b)의 형태로 정의 될 수 있다. 이때, W 및 b는 레이어 파라미터라고 할 수 있다.

[47] 본 개시에서 앙상블 레이어(Ensemble Layer)라 함은 복수의 인공 지능 모델에서 둘 이상의 레이어 출력을 합하여 이용하는 경우, 출력이 합하여 지는 레이어를 의미할 수 있다.

[48] 본 개시에서, 레이어 구조(Layer Scheme 또는 Neural Network Scheme)라 함은 앙상블 모델 또는 앙상블 모델에서 개별 모델의 개수, 레이어간의 연결성 등, 인공 지능 모델의 구조를 의미할 수 있다.

[49] 본 개시에서 레이어 함수(Layer Function)라 함은, 상기 레이어의 정의(y=f(Wx+b))를 일반화 하여 임의의 함수 g로 표현한 것을 의미할 수 있다. 예를 들어, 레이어 함수는 y = g(x) 혹은 y_j= g(x1, x2, ... x_N)와 같이 표현될 수 있다. 이때, 레이어 함수 g는 입력 x1, ..., x_N의 선형 함수로 표현되나, 다른 형태로 표현도 가능함은 물론이다.

[50] 본 개시에서 확률적 레이어 함수(Stochastic Layer Function)이라 함은, 레이어

함수의 출력이 확률적으로 결정되는 함수를 의미할 수 있다. 예를 들어, 확률적 레이어 함수는 y_j~N(f_mean(x_1...x_N), f_var(x_1...X_N)와 같이 x_1, ... x_N에 대해서 출력 y_j의 확률 분포로 이루어 질 수 있다.

[51] 아래에서는 첨부한 도면을 참고하여 본 개시의 실시 예에 대하여 본 개시가 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 개시는 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시 예에 한정되지 않는다. 그리고 도면에서 본 개시를 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[52]

[53] 도 1a 내지 도 1c는 본 개시의 일 실시예에 따른 종래 서버에서 동작하는 인공 지능 모델의 결과 출력 방법을 설명하기 위한 예시도이다.

[54] 도 1a에 도시된 바와 같이, 종래 서버는 하나의 인공 지능 모델에 데이터를 입력하여, 결과값을 출력하는 방법이 존재할 수 있다. 구체적으로, 종래 서버는 하나의 인공 지능 모델을 학습시키고, 학습된 모델에 데이터를 입력하여 결과값을 출력할 수 있다. 그러나, 종래 서버가 하나의 단일 모델을 사용하는 경우, 작은 노이즈에도 쉽게 인식 결과가 변할 수 있어, 결과값에 대한 정확도 및 안정성이 낮은 문제점이 존재하였다.

[55] 인공 지능 모델의 정확도 및 안정성을 향상시키기 위해서는, 종래 서버는 도 1b에 도시된 바와 같이, 복수개의 모델을 앙상블하여 결과값을 출력할 수 있다. 구체적으로, 앙상블 모델은 별도로 학습된 복수의 모델에 데이터를 입력하여 복수개의 결과 값을 출력하고, 출력된 복수개의 결과값을 바탕으로 최종 결과를 출력할 수 있다. 이때, 최종 결과물은 복수개의 결과값에 대하여 Averaging, Weight Sum, Boosting 방식 등을 적용하여 출력될 수 있다.

[56] 예를 들어, 앙상블 모델은 각각의 인공 지능 모델을 통해 출력된 결과값들의 평균값을 최종 출력값으로 결정할 수 있다.

[57] 또는, 앙상블 모델은 각각의 인공지능 모델을 통해 출력된 결과값 중 특정 출력값에 특정 가중치 값을 곱하여 합산한 값을 최종 출력값으로 출력할 수 있다.

[58] 또는, 각각의 인공 지능 모델을 통해 출력된 각각의 결과값이 복수의 파라미터를 포함하고 있는 경우, 앙상블 모델은 복수의 파라미터 각각에 대하여 최종 출력 파라미터를 결정할 수 있다. 예를 들어, 앙상블 모델이 3개의 인공지능 모델을 포함하고, 3개의 인공 지능 모델을 통해 출력된 결과값이 5개의 파라미터를 포함하는 경우를 가정할 수 있다. 이때, 제1 모델은 제1 파라미터 내지 제5 파라미터를 포함하는 제1 결과값을 출력하고, 제2 모델은 제1 파라미터 내지 제5 파라미터를 포함하는 제2 결과값을 출력하고, 제3 모델은 제1 파라미터 내지 제5 파라미터를 포함하는 제3 결과값을 출력할 수 있다. 이때, 앙상블 모델은 제1 모델의 제1 파라미터, 제2 모델의 제1 파라미터 및 제3 모델의 제1

파라미터를 바탕으로 최종 출력값의 제1 파라미터를 획득하고, 제1 모델의 제2 파라미터, 제2 모델의 제2 파라미터 및 제3 모델의 제2 파라미터를 바탕으로 최종 출력값의 제2 파라미터를 획득할 수 있다. 같은 방법으로, 최종 출력값의 제3 파라미터 내지 제5 파라미터가 결정되면, 앙상블 모델은 결정된 제1 파라미터 내지 제5 파라미터를 최종 출력값으로 획득할 수 있다.

[59] 앙상블 모델을 사용하는 경우, 하나의 단일 모델을 사용하는 경우보다 나은 안정성 및 높은 정확도를 가질 수 있으며, 훈련 과정의 Overfitting 문제가 적으며, 외부의 Adversarial Attack등에 대하여도 강한 안정성을 가질 수 있다.

[60] 그러나, 앙상블 모델은, 앙상블 되는 모델의 개수가 증가할수록, 모델에 사용되는 레이어 파라미터의 개수 및 연산량이 증가할 수 있다. 또한, 복수의 모델이 별도로 학습되었다고 하더라도, 앙상블 모델은 하나의 목표(예를 들어, 음성 인식, 얼굴 인식 등)를 달성하기 위해 학습된 모델이기 때문에, 별도로 학습된 복수의 모델의 학습 결과가 유사한 결과로 수렴할 가능성이 높아 획기적인 성능 향상을 보장하지 않는 경우가 많다.

[61] 따라서, 앙상블 모델을 사용하는 경우, 기대되는 성능 향상에 비하여 큰 저장 공간과 많은 연산 과정을 필요로 하므로, 스마트폰, 디스플레이 장치, IoT 시스템과 같이 상대적으로 적은 저장 공간과 연산처리 속도를 가지는 장치에서 앙상블 모델을 이용하는 것은 부적절한 측면이 있다.

[62] 따라서, 상술한 사용자 단말과 같은 장치(이하 전자 장치)에서 인공 지능 모델을 안정적으로 이용하기 위해서는 앙상블 모델을 사용하면서도 적은 저장 공간 및 적은 연산량을 확보하여야 한다.

[63] 따라서, 이하에서 설명하는 바와 같이, 전자 장치(200)는 서버(100)로부터 복수의 모델을 수신하여 저장하는 대신, 복수의 모델을 생성할 수 있는 레이어 생성부를 수신하여 저장하고, 앙상블 모델이 필요한 경우, 프로세서를 통해 앙상블 모델을 생성하여 적은 저장 공간에서 유사한 성능을 유지할 수 있는 앙상블 모델을 생성할 수 있다. 즉, 도 1c에 도시된 바와 같이, 전자 장치(200)는 복수개의 레이어를 저장하고 있지 않고, 레이어 파라미터를 생성할 수 있는 기준값 및 특정 조건에 대한 데이터 및 레이어 파라미터 생성 함수 등을 서버(100)로부터 수신하여, 필요한 경우, 레이어 파라미터를 생성할 수 있다. 생성된 레이어 파라미터로부터 레이어를 생성하기 위하여, 전자 장치(200)는 레이어 구조에 대한 데이터를 서버(100)로부터 수신할 수 있음은 물론이다.

[64] 한편, 본 개시에 따른 앙상블 모델은 다양한 분야에서 사용될 수 있음은 물론이다. 일 실시예로, 전자 장치(200)에서 생성되는 앙상블 모델은 음성 인식을 위한 앙상블 모델일 수 있다. 일반적인 음성 인식 인공 지능 모델은 용량이 커 서버에서 이용될 수 있다. 그러나, 음성 인식이 서버에서 이루어 지는 경우, 지속적인 데이터 사용 또는 개인 정보 유출 문제가 발생할 수 있다. 본 개시에 따라, 전자 장치(200)가 음성 인식을 위한 앙상블 모델을 생성하는 경우, 서버(100)로부터 앙상블 모델을 생성하기 위한 데이터를 수신하는 것만으로도

앙상블 모델을 통해 음성 인식을 수행할 수 있다.

[65] 또 다른 실시예로, 전자 장치(200)에서 생성되는 앙상블 모델은 언어 번역을 위한 앙상블 모델일 수 있다. 언어 번역이 서버(100)와 별개로 전자 장치(200)에서 이루어 지는 경우, 오프라인 상황에서도 번역 기능을 사용할 수 있다는 이점이 있다.

[66] 또 다른 실시예로, 전자 장치(200)에서 생성되는 앙상블 모델은 영상 인식 및 처리를 위한 앙상블 모델일 수 있다. 일반적인 영상 인식/처리를 위한 인공 지능 모델은 용량이 커 서버에서 이용되고 있으며, 입력되는 영상이 고화질 영상인 경우, 서버로 전송하는데 많은 시간과 비용이 소요될 수 있다. 또한, 개인 사진과 같은 영상은 보안 문제를 가지고 있어 서버로 전송하기 적합하지 않을 여지가 있다. 본 개시에 따라, 전자 장치(200)가 영상 인식/처리를 위한 앙상블 모델을 생성하는 경우, 서버(100)로부터 앙상블 모델을 생성하기 위한 데이터를 수신하는 것만으로도 앙상블 모델을 통해 영상 인식/처리를 수행할 수 있다.

[67]

[68] 도 2는 본 개시의 일 실시예에 따른 전자 장치의 구성을 간략히 도시한 블록도이다.

[69] 도 2에 도시된 바와 같이, 전자 장치(200)는 통신부(210), 메모리(220) 및 프로세서(230)를 포함할 수 있다. 본 개시의 다양한 실시예들에 따른 전자 장치(200)는, 예를 들면, 스마트폰, 태블릿 PC, 이동 전화기, 영상 전화기, 전자책 리더기, 데스크탑 PC, 랩탑 PC, 넷북 컴퓨터, 워크스테이션, 서버, PDA, PMP(portable multimedia player), MP3 플레이어, 의료기기, 카메라, 또는 웨어러블 장치 중 적어도 하나를 포함할 수 있다. 웨어러블 장치는 액세서리형(예: 시계, 반지, 팔찌, 발찌, 목걸이, 안경, 콘택트 렌즈, 또는 머리 착용형 장치(head-mounted-device(HMD)), 직물 또는 의류 일체형(예: 전자 의복), 신체 부착형(예: 스킨 패드 또는 문신), 또는 생체 이식형 회로 중 적어도 하나를 포함할 수 있다.

[70] 통신부(210)는 외부 장치와의 통신을 수행하기 위한 구성이다. 이 경우, 통신부(210)가 외부 장치와 통신 연결되는 것은 제3 기기(예로, 중계기, 허브, 엑세스 포인트, 서버 또는 게이트웨이 등)를 거쳐서 통신하는 것을 포함할 수 있다.

[71] 무선 통신은, 예를 들면, LTE, LTE-A(LTE Advance), CDMA(code division multiple access), WCDMA(wideband CDMA), UMTS(universal mobile telecommunications system), WiBro(Wireless Broadband), 또는 GSM(Global System for Mobile Communications) 등 중 적어도 하나를 사용하는 셀룰러 통신을 포함할 수 있다. 일 실시예에 따르면, 무선 통신은, 예를 들면, WiFi(wireless fidelity), 블루투스, 블루투스 저전력(BLE), 지그비(Zigbee), NFC(near field communication), 자력 시큐어 트랜스미션(Magnetic Secure Transmission), 라디오 프리퀀시(RF), 또는 보디 에어리어 네트워크(BAN) 중 적어도 하나를 포함할 수 있다. 유선

통신은, 예를 들면, USB(universal serial bus), HDMI(high definition multimedia interface), RS-232(recommended standard232), 전력선 통신, 또는 POTS(plain old telephone service) 등 중 적어도 하나를 포함할 수 있다.

[72] 무선 통신 또는 유선 통신이 수행되는 네트워크는 텔레커뮤니케이션 네트워크, 예를 들면, 컴퓨터 네트워크(예: LAN 또는 WAN), 인터넷, 또는 텔레폰 네트워크 중 적어도 하나를 포함할 수 있다.

[73] 메모리(220)는, 예를 들면, 전자 장치(200)의 적어도 하나의 다른 구성요소에 관계된 명령 또는 데이터를 저장할 수 있다. 일 실시예에 따르면, 메모리(220)는 소프트웨어 및/또는 프로그램을 저장할 수 있다. 프로그램은, 예를 들면, 커널, 미들웨어, 어플리케이션 프로그래밍 인터페이스(API) 및/또는 어플리케이션 프로그램(또는 "어플리케이션") 등을 포함할 수 있다. 커널, 미들웨어 또는 API의 적어도 일부는, 운영 시스템으로 지칭될 수 있다. 커널은, 예를 들면, 다른 프로그램들에 구현된 동작 또는 기능을 실행하는 데 사용되는 시스템 리소스들을 제어 또는 관리할 수 있다. 또한, 커널은 미들웨어, API, 또는 어플리케이션 프로그램에서 전자 장치(200)의 개별 구성요소에 접근함으로써, 시스템 리소스들을 제어 또는 관리할 수 있는 인터페이스를 제공할 수 있다.

[74] 미들웨어는, 예를 들면, API또는 어플리케이션 프로그램이 커널과 통신하여 데이터를 주고받을 수 있도록 중개 역할을 수행할 수 있다. 또한, 미들웨어는 어플리케이션 프로그램으로부터 수신된 하나 이상의 작업 요청들을 우선 순위에 따라 처리할 수 있다. 예를 들면, 미들웨어는 어플리케이션 프로그램 중 전자 장치(200)의 시스템 리소스를 사용할 수 있는 우선 순위를 부여하고, 상기 하나 이상의 작업 요청들을 처리할 수 있다. API는 어플리케이션이 커널 또는 미들웨어에서 제공되는 기능을 제어하기 위한 인터페이스로, 예를 들면, 파일 제어, 창 제어, 영상 처리, 또는 문자 제어 등을 위한 적어도 하나의 인터페이스 또는 함수(예: 명령어)를 포함할 수 있다.

[75] 한편, 메모리(220)는 본 개시에 따른 레이어를 생성하기 위한 다양한 모듈을 포함할 수 있다. 일 실시예로, 메모리(220)는 레이어 파라미터를 생성하기 위한 레이어 파라미터 생성 모듈, 레이어 함수를 생성하기 위한 레이어 함수 생성 모듈, 레이어의 구조를 변경하기 위한 레이어 구조 변경 모듈, 레이어 생성 모듈의 출력이 특정 확률 분포 값에 따라 변경되도록 제어하기 위한 확률적 앙상블 생성 모듈 등 다양한 모듈을 포함할 수 있다. 상술한 다양한 모듈은 프로세서에 의해 제어되어 레이어 파라미터, 레이어 함수, 레이어 구조 등을 생성할 수 있다.

[76] 프로세서(230)는, 중앙처리장치, 어플리케이션 프로세서, 또는 커뮤니케이션 프로세서(communication processor(CP)) 중 하나 또는 그 이상을 포함할 수 있다.

[77] 또한, 프로세서(230)는 주문형 집적 회로(application specific integrated circuit, ASIC), 임베디드 프로세서, 마이크로프로세서, 하드웨어 컨트롤 로직, 하드웨어 유한 상태 기계(hardware finite state machine, FSM), 디지털 신호 프로세서(digital

signal processor, DSP), 중 적어도 하나로 구현될 수 있다. 도시하진 않았으나, 프로세서(230)는 각 구성들과 통신을 위한 버스(bus)와 같은 인터페이스를 더 포함할 수 있다.

[78] 프로세서(230)는, 예를 들면, 운영 체제 또는 응용 프로그램을 구동하여 프로세서(230)에 연결된 다수의 하드웨어 또는 소프트웨어 구성요소들을 제어할 수 있고, 각종 데이터 처리 및 연산을 수행할 수 있다. 프로세서(230)는, 예를 들면, SoC(system on chip) 로 구현될 수 있다. 일 실시예에 따르면, 프로세서(230)는 GPU(graphic processing unit) 및/또는 이미지 신호 프로세서를 더 포함할 수 있다. 프로세서(230)는 다른 구성요소들(예: 비휘발성 메모리) 중 적어도 하나로부터 수신된 명령 또는 데이터를 휘발성 메모리에 로드하여 처리하고, 결과 데이터를 비휘발성 메모리에 저장할 수 있다.

[79] 한편, 프로세서(230)는 인공 지능(AI; artificial intelligence)을 위한 전용 프로세서를 포함하거나, 기존의 범용 프로세서(예: CPU 또는 application processor) 또는 그래픽 전용 프로세서(예: GPU)의 일부로 제작될 수 있다. 이 때, 인공 지능을 위한 전용 프로세서는 확률 연산에 특화된 전용 프로세서로서, 기존의 범용 프로세서보다 병렬처리 성능이 높아 기계 학습과 같은 인공 지능 분야의 연산 작업을 빠르게 처리할 수 있다.

[80] 프로세서(230)는 외부 서버로부터 통신부(210)를 통해 복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어에 대한 레이어 파라미터를 획득하기 위한 기준값을 수신할 수 있다. 이때, 기준값은 외부 서버에서 생성된 것으로, 레이어 파라미터를 생성하기 위한 값이며, 예를 들어, 기준값은 앙상블 모델의 대응되는 위치의 레이어들에 대한 레이어 파라미터의 평균값 또는 분산값 중 적어도 하나일 수 있다.

[81] 복수의 인공 지능 모델을 이용하기 위한 사용자 명령이 입력되면, 프로세서(230)는 외부 서버로부터 수신한 기준값을 바탕으로 앙상블 모델(또는 복수의 인공 지능 모델)을 생성할 수 있다. 생성된 앙상블 모델에 입력 데이터가 입력되면, 프로세서(230)는 입력데이터에 대한 출력 데이터를 획득할 수 있다.

[82] 구체적으로, 프로세서(230)는 앙상블 모델에 포함된 복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중, 대응되는 위치의 레이어들에 대한 레이어 파라미터를 획득하기 위한 기준값을 바탕으로, 복수의 인공 지능 모델 각각에 포함된 레이어 중 대응되는 위치의 레이어를 생성할 수 있다.

[83] 이때, 기준값은, 대응되는 위치의 레이어들에 대한 레이어 파라미터의 평균값 또는 분산값일 수 있다. 기준값은 다양한 조건에 따라 다양하게 결정될 수 있다. 일 실시예로, 복수의 레이어 파라미터의 분산값이 기 설정된 값 이하인 경우, 기준값은 복수의 레이어 파라미터의 평균값을 바탕으로 결정될 수 있다. 또 다른 실시예로, 기준값은, 복수의 레이어 파라미터의 분산값이 기 설정된 값 이상이고, 복수의 레이어 파라미터의 분포가 기 설정된 조건을 만족하면, 복수의 레이어 파라미터의 분산값을 바탕으로 결정될 수 있다. 또 다른 실시예로,

기준값은, 복수의 레이어 파라미터의 분산값이 기 설정된 값 이상이고, 복수의 레이어 파라미터의 분포가 기 설정된 조건을 만족하지 않으면, 복수의 레이어 파라미터의 평균값 및 분산값을 바탕으로 결정될 수 있다.

[84] 기준값이 복수의 레이어 파라미터의 평균값을 바탕으로 결정된 경우, 프로세서(230)는 복수의 레이어 파라미터의 평균값을 바탕으로 결정된 기준값을 기초로 레이어 파라미터를 생성할 수 있다. 예를 들어, 생성된 레이어 파라미터는 복수의 레이어 파라미터의 평균값을 바탕으로 결정된 기준값을 입력 값으로 가우시안 분포를 가지는 함수에 대입한 임의의 출력값일 수 있다. 이때, 생성된 레이어 파라미터 W'는 예를 들어, $W'=\{w_{_ij}\sim N(m,v)\}_{_ij}$와 같이 표현될 수 있다.

[85] 또 다른 예로, 생성된 레이어 파라미터는 복수의 레이어 파라미터의 평균값을 바탕으로 결정된 기준값 및 레이어의 Transformation에 대한 분포에 따라 결정될 수 있다. 이 경우, 생성된 레이어 파라미터 W'은 임의의 함수 A(t)에 대한 연산값일 수 있다. 즉, W'=A(t)*W의 형태일 수 있다.

[86] 한편, 기준값이 복수의 레이어 파라미터의 분산값을 바탕으로 결정된 경우, 프로세서(230)는 복수의 레이어 파라미터의 분산값을 바탕으로 결정된 기준값을 기초로 레이어 파라미터를 생성할 수 있다.

[87] 한편, 기준값이 복수의 레이어 파라미터의 평균값 및 분산값을 바탕으로 결정된 경우, 프로세서(230)는 복수의 레이어 파라미터의 평균값 및 분산값을 바탕으로 결정된 기준값을 기초로 레이어 파라미터를 생성할 수 있다.

[88] 예를 들어, 기준값이 복수의 레이어 파라미터의 평균값 및 분산값을 바탕으로 결정된 경우, 프로세서(230)는 확률적 레이어 함수(예를 들어 y_j~N(f_mean(x_1...x_N), f_var(x_1...X_N))의 출력값을 레이어 파라미터로 생성할 수 있다.

[89] 상술한 방법에 따라, 프로세서(230)는 앙상블 모델에 포함된 모든 레이어의 레이어 파라미터를 생성할 수 있다. 즉, 프로세서(230)는 서버(100)로부터 수신한 기준값 및 레이어 파라미터 생성 방법에 관한 정보만을 가지고 레이어 파라미터를 생성하므로, 앙상블 모델이 포함하는 모든 레이어 파라미터를 저장할 필요가 없으므로 저장 공간 및 연산량을 줄일 수 있다.

[90] 한편, 프로세서(230)는 전자 장치의 메모리 용량을 바탕으로 생성되는 인공 지능 모델의 개수를 판단할 수 있다. 즉, 프로세서(230)는 메모리(220)의 공간을 판단하고, 판단된 메모리(220) 공간에 따라 앙상블 모델에 필요한 인공 지능 모델의 개수를 결정할 수 있다. 메모리(220) 사용량이 충분한 경우, 프로세서(230)는 많은 인공 지능 모델을 생성하여 정확도 및 안정성을 높일 수 있으며, 메모리(220) 공간이 충분치 않은 경우, 메모리(220) 공간에 맞게 인공 지능 모델을 생성하여 빠른 결과를 출력할 수 있다.

[91] 프로세서(230)는 생성된 앙상블 모델에 입력값이 입력되면, 결과값을 출력할 수 있다.

[92]

[93] 도 3은 본 개시의 일 실시예에 따른 전자 장치(200)의 구성을 상세히 도시한 블록도이다.

[94] 전자 장치(200)는 통신부(210), 메모리(220) 및 프로세서(230) 이외에도 입력부(240) 및 출력부(250)을 더 포함할 수 있다. 그러나 상술한 구성에 한정되는 것은 아니며, 필요에 따라 일부 구성이 추가 또는 생략될 수 있음은 물론이다.

[95] 상술한 바와 같이, 통신부(210)는 외부 서버와의 통신을 수행한다. 특히, 통신부(210)는 와이파이 칩, 블루투스 칩, 무선 통신 칩, NFC칩 등과 같은 무선 통신을 수행하기 위한 다양한 통신 칩을 포함할 수 있다. 이때, 와이파이 칩, 블루투스 칩, NFC 칩은 각각 LAN 방식, WiFi 방식, 블루투스 방식, NFC 방식으로 통신을 수행한다. 와이파이 칩이나 블루투스칩을 이용하는 경우에는 SSID 및 세션 키 등과 같은 각종 연결 정보를 먼저 송수신 하여, 이를 이용하여 통신 연결한 후 각종 정보들을 송수신할 수 있다. 무선 통신칩은 IEEE, 지그비, 3G(3rd Generation), 3GPP(3rd Generation Partnership Project), LTE(Long Term Evolution) 등과 같은 다양한 통신 규격에 따라 통신을 수행하는 칩을 의미한다.

[96] 메모리(220)는, 내장 메모리 및 외장 메모리 중 적어도 하나를 포함할 수 있다. 내장 메모리는, 예를 들면, 휘발성 메모리(예: DRAM, SRAM, 또는 SDRAM 등), 비휘발성 메모리(예: OTPROM(one time programmable ROM), PROM, EPROM, EEPROM, mask ROM, flash ROM, 플래시 메모리, 하드 드라이브, 또는 솔리드 스테이트 드라이브(SSD) 중 적어도 하나를 포함할 수 있다. 외장 메모리는 플래시 드라이브(flash drive), 예를 들면, CF(compact flash), SD(secure digital), Micro-SD, Mini-SD, xD(extreme digital), MMC(multi-media card) 또는 메모리 스틱 등을 포함할 수 있다. 외장 메모리는 다양한 인터페이스를 통하여 전자 장치(200)와 기능적으로 또는 물리적으로 연결될 수 있다.

[97] 입력부(240)는 사용자 명령을 입력받기 위한 구성이다. 이때, 입력부(240)는 카메라(241), 마이크(242) 및 터치패널(243)을 포함할 수 있다.

[98] 카메라(241)는 전자 장치(200) 주변의 영상 데이터를 획득하기 위한 구성이다. 카메라(241)는 정지 영상 및 동영상을 촬영할 수 있다. 예로, 카메라(241)는 하나 이상의 이미지 센서(예: 전면 센서 또는 후면 센서), 렌즈, 이미지 시그널 프로세서(ISP), 또는 플래시(예: LED 또는 xenon lamp 등)를 포함할 수 있다. 마이크(242)는 전자 장치(200) 주변의 소리를 획득하기 위한 구성이다. 마이크(242)는 외부의 음향 신호를 입력 받아 전기적인 음성 정보를 생성할 수 있다. 마이크(242)는 외부의 음향 신호를 입력 받는 과정에서 발생되는 잡음(noise)를 제거하기 위한 다양한 잡음 제거 알고리즘을 이용할 수 있다. 카메라(241) 또는 마이크(242)를 통해 입력된 이미지 정보 또는 음성 정보는 인공지능 모델의 입력값으로 입력될 수 있다.

[99] 터치 패널(243)은 다양한 사용자 입력을 입력 받을 수 있는 구성이다. 터치

패널(243)는 사용자 조작에 의한 데이터를 입력 받을 수 있다. 터치 패널(243)은 후술하는 디스플레이와 결합하여 구성될 수도 있다.

[100] 입력부(240)는 상술한 카메라(241), 마이크(242), 터치 패널(243) 외에도 다양한 데이터를 입력 받기 위한 다양한 구성일 수 있음은 물론이다.

[101] 출력부(250)는 오디오 출력부(251) 또는 디스플레이(252)를 포함할 수 있으며, 다양한 데이터를 출력할 수 있는 구성이다.

[102] 오디오 출력부(251)는 입력 데이터에 대한 다양한 소리를 출력하기 위한 구성이다. 오디오 출력부(251)는 오디오 처리부(미도시) 에 의해 디코딩이나 증폭, 노이즈 필터링과 같은 다양한 처리 작업이 수행된 각종 오디오 데이터뿐만 아니라 각종 알림 음이나 음성 메시지를 출력할 수 있다. 특히, 오디오 출력부(251)는 스피커로 구현될 수 있으나, 이는 일 실시 예에 불과할 뿐, 오디오 데이터를 출력할 수 있는 출력 단자로 구현될 수 있다.

[103] 디스플레이(252)는 입력 데이터에 대한 다양한 영상을 출력하기 위한 구성이다. 다양한 영상을 제공하기 위한 디스플레이(252)는 다양한 형태의 디스플레이 패널로 구현될 수 있다. 예를 들어, 디스플레이 패널은 LCD(Liquid Crystal Display), OLED(Organic Light Emitting Diodes), AM-OLED(Active-Matrix Organic Light-Emitting Diode), LcoS(Liquid Crystal on Silicon) 또는 DLP(Digital Light Processing) 등과 같은 다양한 디스플레이 기술로 구현될 수 있다. 또한, 디스플레이(252)는 플렉서블 디스플레이(flexible display)의 형태로 전자 장치(100)의 전면 영역 및, 측면 영역 및 후면 영역 중 적어도 하나에 결합될 수도 있다.

[104]

[105] 이하에서는 도 4a 내지 도 6e를 이용하여 전자 장치의 동작을 상세히 설명한다.

[106] 도 4a 및 도 4b는 본 개시의 일 실시예에 따른, 메모리 저장공간에 따른 인공 지능 모델 생성 방법을 설명하기 위한 예시도이다.

[107] 도 4a 및 도 4b에 도시된 바와 같이, 전자 장치(200)는 사용할 인공 지능 모델의 개수를 메모리(220) 저장 공간에 따라 변경할 수 있다. 구체적으로, 전자 장치(220)는 인공 지능 모델을 이용하기 위한 사용자 명령이 입력되면, 메모리(220)의 저장 공간을 판단하고, 판단된 저장 공간에 따라 앙상블 할 인공 지능 모델의 개수를 결정할 수 있다. 구체적으로, 도 4a에 도시된 바와 같이 메모리(220)가 사용 가능한 저장 공간이 부족한 경우, 전자 장치(220)는 사용 가능한 저장 공간에 따라 적은 수의 인공 지능 모델을 생성할 수 있다. 또는, 도 4b에 도시된 바와 같이 메모리(220)가 사용 가능한 저장 공간이 충분한 경우, 전자 장치(220)는 사용 가능한 저장 공간에 따라 많은 수의 인공 지능 모델을 생성할 수 있다.

[108] 예를 들어, 사용 가능한 메모리 공간이 앙상블 모델의 크기보다 작은 경우, 전자 장치(200)는 앙상블 모델에 포함된 복수의 인공 지능 모델 중 일부를 제거하여 메모리 공간에 맞는 앙상블 모델을 생산할 수 있다.

[109] 또는, 사용 가능한 메모리 공간이 앙상블 모델의 크기보다 큰 경우, 전자 장치(200)는 앙상블 모델을 그대로 사용할 수 있다.

[110] 또는, 사용 가능한 메모리 공간이 앙상블 모델의 크기보다 큰 경우, 전자 장치(200)는 성능 향상을 위해 기존 앙상블 모델에 포함된 복수의 인공 지능 모델 뿐만 아니라, 사용 가능만 메모리 공간을 바탕으로 추가 인공 지능 모델을 생성하고, 기존 앙상블 모델에 포함된 복수의 인공 지능 모델과 추가 인공 지능 모델을 앙상블 할 수 있다.

[111]

[112] 도 5a 내지 도 5c는 본 개시의 일 실시예에 따른, 사용자 명령에 따라 생성되는 인공 지능 모델을 조절할 수 있는 UI를 설명하기 위한 예시도이다.

[113] 구체적으로, 도 5a에 도시된 바와 같이, 전자 장치(200)는 사용자 선택에 따라 인공 지능 모델의 종류 및 앙상블 되는 레이어의 개수를 조절할 수 있다. 이때, 인공 지능 모델의 종류란, 음성 인식을 위한 인공 지능 모델, 영상 분석을 위한 인공 지능 모델 등에 대한 종류일 수 있다.

[114] 한편, 도 5b에 도시된 바와 같이, 전자 장치(200)는 전자 장치(200)의 메모리(220), 성능 등으로 고려하여 사용자에게 최적화된 인공 지능 모델을 앙상블 할 수 있는 UI를 제공할 수 있다. 즉, 도 5b에 도시된 바와 같이, 전자 장치(200)는 인공지능 모델에 사용하고자 하는 메모리의 크기 및 원하는 성능을 조작하기 위한 바 UI(bar UI)를 표시할 수 있다. 다만, 바 UI의 형태는 도 5b에 도시된 예에 한정되지 않으며, 도 5c에 도시된 바와 같이 단계별로 설정 가능한 UI의 형태로 표시될 수도 있다. 즉, 도 5c와 같이 사용자는 메모리, 성능 등의 지표에 대하여 다섯 단계 중 하나의 단계(불연속적으로)를 선택할 수도 있음은 물론이다.

[115] 한편, 도 5a 내지 도 5c에는 개시되어 있지 않으나, 전자 장치(200)는 앙상블 모델에 필요한 레이어의 개수, 사용하고자 하는 인공 지능 모델의 종류 등을 직접 지정할 수 있는 UI를 표시할 수 있음은 물론이다.

[116] 상술한 앙상블 모델 생성과 관련된 설정은 다양한 방법으로 변경될 수 있다. 일 실시예로, 인공 지능 서비스가 특정 어플리케이션에서 동작하는 경우, 전자 장치(200)는 특정 어플리케이션 내부의 설정을 변경하여 앙상블 모델 설정 환경을 변경할 수 있다. 또 다른 실시예로, 전자 장치(200)는 전자 장치(200)자체의 설정 환경(예를 들어, OS 설정 환경)에서 앙상블 모델 설정 환경을 변경할 수 있음은 물론이다.

[117]

[118] 도 6a 내지 도 6e는 본 개시의 일 실시예에 따른 레이어의 앙상블 방법을 설명하기 위한 예시도이다.

[119] 상술한 바와 같이, 전자 장치(200)는 기준값을 통해 레이어 파라미터를 생성할 수 있을 뿐만 아니라, 나아가 레이어의 구조(Layer Scheme)을 생성할 수 있다. 즉, 전자 장치(200)는 레이어 구조를 서버(100)로부터 수신하여 메모리(220)에

저장하고 있으나, 필요에 따라, 도 6a 내지 도 6e와 같이 레이어의 구조를 변경하여 사용할 수 있다. 도 6a 내지 도 6e는 3개의 인공 지능 모델을 앙상블하는 경우에 대하여 설명하였으나, 3개 이상의 모델이 앙상블되거나 두개의 모델이 앙상블 될 수 있음은 물론이다.

[120] 도 6a는 일반적인 앙상블(Simple Ensemble) 모델에 따른 레이어 구조를 나타낸다. 도 6a에 따른 앙상블 모델은 각각의 인공 지능 모델로부터 출력값을 획득하여, 획득된 출력 값을 앙상블할 수 있다. 도 6b는 스텝 와이즈 앙상블(Stepwise Ensemble)모델에 따른 앙상블 구조를 나타낸다. 도 6b에 따른 앙상블 모델은 예를 들어, ReLU 레이어의 출력값을 각각 다음 레이어로 전달하는 것이 아니라, 각각의 ReLU 레이어의 출력값을 앙상블한 값을 다음 레이어로 전달할 수 있다. 이 경우, 런타임 메모리를 절약할 수 있는 효과가 있다. 도 6c는 부분적 스텝 와이즈 앙상블(Stepwise Partial Ensemble) 모델에 따른 앙상블 구조를 나타낸다. 스텝 와이즈 앙상블(Stepwise Ensemble)모델과 유사하나, 필요에 따라 모든 레이어를 앙상블 하는 것이 적절하지 않은 경우, 도 6c와 같은 앙상블 모델이 이용될 수 있다. 도 6d는 부분적 캐스캐이드 앙상블(Cascade Partial Ensemble) 모델에 따른 앙상블 구조를 나타낸다. 캐스캐이드 앙상블 모델은 단계가 거듭될수록 인공 지능 모델의 개수가 줄어들어 메모리를 절약할 수 있다. 결과값이 유사하다고 예측되는 인공 지능 모델이 복수개 존재하는 경우, 도 6d와 같은 앙상블 모델이 이용될 수 있을 것이다. 도 6e는 확률적 앙상블(Stochastic Partial Ensemble) 모델에 따른 앙상블 구조를 나타낸다. 도 6e에 따른 앙상블 모델을 이용하는 경우, 안정적인 결과값을 도출하는데 용이할 수 있다.

[121] 전자 장치(200)는 생성된 레이어 파라미터, 현재 전자 장치(200)의 성능, 인공 지능 모델의 종류 및 목적에 따라 다양한 레이어 구조를 선택하여 적용할 수 있음은 물론이다.

[122]

[123] 도 7은 본 개시의 일 실시예에 따른 서버의 구성을 간략히 도시한 블록도이다.

[124] 도 7에 도시된 바와 같이, 서버(100)는 메모리(110), 통신부(120) 및 프로세서(130)를 포함할 수 있다. 다만, 상술한 구성 외에도 다양한 구성이 추가될 수 있음은 물론이다.

[125] 메모리(110)는 복수의 인공 지능 모델을 저장하기 위한 구성이다. 메모리(110)의 동작 및 구성은 전술한 전자 장치(200)의 메모리(220)의 동작 및 구성과 유사하다. 메모리(110)는 다양한 인공 지능 모델을 저장할 수 있다.

[126] 통신부(120)는 전자 장치(200)와 통신하기 위한 구성이다. 통신부(120)의 동작 및 구성은 전술한 전자 장치(200)의 통신부(210)의 동작 및 구성과 유사하다.

[127] 프로세서(130)는 서버(100)의 동작을 제어하기 위한 구성이다. 프로세서(130)의 동작 및 구성은 전술한 전자 장치(200)의 프로세서(230)의 동작 및 구성을 포함할 수 있다.

[128] 한편, 프로세서(130)는 데이터 학습부 및 데이터 인식부를 포함할 수 있다. 데이터 학습부는 데이터 인식 모델이 특정 목적에 따른 기준을 갖도록 학습시킬 수 있다. 이때, 특정 목적이란 음성 인식, 번역, 영상 인식, 상황 인식 등과 관련된 목적을 포함할 수 있다. 데이터 학습부는 상술한 목적에 따른 동작을 결정하기 위하여 학습 데이터를 데이터 인식 모델에 적용하여, 판단 기준을 갖는 데이터 인식 모델을 생성할 수 있다. 데이터 인식부는 인식 데이터에 기초하여 특정 목적에 대한 상황을 판단할 수 있다. 데이터 인식부는 학습된 데이터 인식 모델을 이용하여, 소정의 인식 데이터로부터 상황을 판단할 수 있다. 데이터 인식부는 기 설정된 기준에 따라 소정의 인식 데이터를 획득하고, 획득된 인식 데이터를 입력 값으로 하여 데이터 인식 모델에 적용함으로써, 소정의 인식 데이터에 기초한 소정의 상황을 판단할 수 있다(또는, 추정(estimate)할 수 있다). 또한, 획득된 인식 데이터를 입력 값으로 데이터 인식 모델에 적용하여 출력된 결과 값은, 데이터 인식 모델을 갱신하는데 이용될 수 있다.

[129] 한편, 데이터 학습부의 적어도 일부 및 데이터 인식부의 적어도 일부는, 소프트웨어 모듈로 구현되거나 적어도 하나의 하드웨어 칩 형태로 제작되어 전자 장치에 탑재될 수 있다. 예를 들어, 데이터 학습부 및 데이터 인식부 중 적어도 하나는 인공 지능(AI; artificial intelligence)을 위한 전용 하드웨어 칩 형태로 제작될 수도 있고, 또는 기존의 범용 프로세서(예: CPU 또는 application processor) 또는 그래픽 전용 프로세서(예: GPU)의 일부로 제작되어 전술한 각종 전자 장치에 탑재될 수도 있다. 이 때, 인공 지능을 위한 전용 하드웨어 칩은 확률 연산에 특화된 전용 프로세서로서, 기존의 범용 프로세서보다 병렬처리 성능이 높아 기계 학습과 같은 인공 지능 분야의 연산 작업을 빠르게 처리할 수 있다. 데이터 학습부 및 데이터 인식부가 소프트웨어 모듈(또는, 인스트럭션(instruction) 포함하는 프로그램 모듈)로 구현되는 경우, 소프트웨어 모듈은 컴퓨터로 읽을 수 있는 판독 가능한 비일시적 판독 가능 기록매체(non-transitory computer readable media)에 저장될 수 있다. 이 경우, 소프트웨어 모듈은 OS(Operating System)에 의해 제공되거나, 소정의 애플리케이션에 의해 제공될 수 있다. 또는, 소프트웨어 모듈 중 일부는 OS(Operating System)에 의해 제공되고, 나머지 일부는 소정의 애플리케이션에 의해 제공될 수 있다.

[130] 이 경우, 데이터 학습부 및 데이터 인식부는 하나의 서버(100)에 탑재될 수도 있으며, 또는 별개의 서버들에 각각 탑재될 수도 있다. 예를 들어, 데이터 학습부 및 데이터 인식부 중 하나는 전자 장치(200)에 포함되고, 나머지 하나는 외부의 서버에 포함될 수 있다. 또한, 데이터 학습부 및 데이터 인식부는 유선 또는 무선으로 통하여, 데이터 학습부가 구축한 모델 정보를 데이터 인식부로 제공할 수도 있고, 데이터 인식부로 입력된 데이터가 추가 학습 데이터로서 데이터 학습부로 제공될 수도 있다.

[131] 한편, 데이터 학습부는 데이터 획득부, 전처리부, 학습 데이터 선택부, 모델

학습부 및 모델 평가부를 더 포함할 수 있다. 데이터 획득부는 특정 목적에 따른 학습 데이터를 획득하기 위한 구성이다. 전처리부는 획득부에서 획득된 데이터를 기 정의된 포맷으로 전처리하기 위한 구성이다. 학습 데이터 선택부는 전처리된 데이터 중 학습에 필요한 데이터를 선별하기 위한 구성이다. 모델 학습부는 학습 데이터를 이용하여 데이터 인식 모델을 학습시키기 위한 구성이다. 모델 평가부는 데이터 인식 모델의 결과를 향상 시키기 위한 구성이다. 전술한, 데이터 획득부, 전처리부, 학습 데이터 선택부, 모델 학습부 및 모델 평가부 중 적어도 하나는, 소프트웨어 모듈로 구현되거나 적어도 하나의 하드웨어 칩 형태로 제작되어 전자 장치에 탑재될 수 있다. 예를 들어, 데이터 획득부, 전처리부, 학습 데이터 선택부, 모델 학습부 및 모델 평가부 중 적어도 하나는 인공 지능(AI; artificial intelligence)을 위한 전용 하드웨어 칩 형태로 제작될 수도 있고, 또는 기존의 범용 프로세서(예: CPU 또는 application processor) 또는 그래픽 전용 프로세서(예: GPU)의 일부로 제작되어 전술한 각종 전자 장치에 탑재될 수도 있다.

[132] 또한, 데이터 인식부는 데이터 획득부, 전처리부, 인식 데이터 선택부, 인식 결과 제공부 및 모델 갱신부를 더 포함할 수 있다. 데이터 획득부는 인식 데이터를 획득하기 위한 구성이다. 전처리부는 획득부에서 획득된 데이터를 기 정의된 포맷으로 전처리하기 위한 구성이다. 인식 데이터 선택부는 전처리된 데이터 중 인식에 필요한 데이터를 선별하기 위한 구성이다. 인식 결과 제공부는 인식 데이터로부터 선택된 데이터를 수신할 수 있는 구성이다. 모델 갱신부는 인식 결과 제공부로부터 제공된 인식 결과에 대한 평가에 기초하여 데이터 인식 모델을 갱신하기 위한 구성이다. 전술한, 데이터 획득부, 전처리부, 인식 데이터 선택부, 인식 결과 제공부 및 모델 갱신부 중 적어도 하나는, 소프트웨어 모듈로 구현되거나 적어도 하나의 하드웨어 칩 형태로 제작되어 전자 장치에 탑재될 수 있다. 예를 들어, 데이터 획득부, 전처리부, 학습 데이터 선택부, 모델 학습부 및 모델 평가부 중 적어도 하나는 인공 지능(AI; artificial intelligence)을 위한 전용 하드웨어 칩 형태로 제작될 수도 있고, 또는 기존의 범용 프로세서(예: CPU 또는 application processor) 또는 그래픽 전용 프로세서(예: GPU)의 일부로 제작되어 전술한 각종 전자 장치에 탑재될 수도 있다.

[133] 이때, 데이터 인식 모델이란 단일 모델 또는 앙상블 모델 중 어느 하나일 수 있다. 또한, 데이터 인식 모델은 인식 모델의 적용 분야, 학습의 목적 또는 장치의 컴퓨터 성능 등을 고려하여 구축될 수 있다. 데이터 인식 모델은, 예를 들어, 신경망(Neural Network)을 기반으로 하는 모델일 수 있다. 데이터 인식 모델은 인간의 뇌 구조를 컴퓨터 상에서 모의하도록 설계될 수 있다. 데이터 인식 모델은 인간의 신경망의 뉴런(neuron)을 모의하는, 가중치를 가지는 복수의 네트워크 노드들을 포함할 수 있다. 복수의 네트워크 노드들은 뉴런이 시냅스(synapse)를 통하여 신호를 주고 받는 시냅틱(synaptic) 활동을 모의하도록 각각 연결 관계를 형성할 수 있다. 데이터 인식 모델은, 일 예로, 신경망 모델,

또는 신경망 모델에서 발전한 딥 러닝 모델을 포함할 수 있다. 딥 러닝 모델에서 복수의 네트워크 노드들은 서로 다른 깊이(또는, 레이어)에 위치하면서 컨볼루션(convolution) 연결 관계에 따라 데이터를 주고 받을 수 있다.

[134] 예컨대, DNN(Deep Neural Network), RNN(Recurrent Neural Network), BRDNN(Bidirectional Recurrent Deep Neural Network)과 같은 모델이 데이터 인식 모델로서 사용될 수 있으나, 이에 한정되지 않는다.

[135] 한편, 프로세서(130)는 전술한 데이터 학습부 및 데이터 인식부에 의해 학습된 인공 지능 모델을 바탕으로, 전자 장치(200)가 레이어를 생성하기 위한 기준값을 획득할 수 있다.

[136] 이때, 기준값은, 대응되는 위치의 레이어들에 대한 레이어 파라미터의 평균값 및 분산값일 수 있다. 기준값은 다양한 조건에 따라 다양하게 결정될 수 있다. 일 실시예로, 복수의 레이어 파라미터의 분산값이 기 설정된 값 이하인 경우, 기준값은 복수의 레이어 파라미터의 평균값을 바탕으로 결정될 수 있다. 또 다른 실시예로, 기준값은, 복수의 레이어 파라미터의 분산값이 기 설정된 값 이상이고, 복수의 레이어 파라미터의 분포가 기 설정된 조건을 만족하면, 복수의 레이어 파라미터의 분산값을 바탕으로 결정될 수 있다. 또 다른 실시예로, 기준값은, 복수의 레이어 파라미터의 분산값이 기 설정된 값 이상이고, 복수의 레이어 파라미터의 분포가 기 설정된 조건을 만족하지 않으면, 복수의 레이어 파라미터의 평균값 및 분산값을 바탕으로 결정될 수 있다.

[137] 구체적으로, 프로세서(130)는 학습된 모델에 포함된 레이어 파라미터를 바탕으로 기준값을 획득할 수 있다. 구체적으로, 도 1b를 참조하면, 프로세서(130)는 복수의 인공 지능 모델에 포함된 레이어 중 동일 경로에 대한 레이어 파라미터를 바탕으로 기준값을 획득할 수 있다. 예를 들어, 앙상블 모델이 3개의 인공 지능 모델(제1 인공 지능 모델 내지 제3 인공 지능 모델)을 포함하고, 각각의 인공 지능 모델은 10개의 레이어 파라미터(제1 레이어 파라미터 내지 제10 레이어 파라미터)를 포함하는 경우를 가정할 수 있다. 이때, 각각의 인공 지능 모델의 제n 레이어 파라미터는 동일 경로를 가지는 레이어 파라미터일 수 있다.

[138] 이때, 프로세서(130)는 제1 인공 지능 모델의 제1 레이어 파라미터, 제2 인공 지능 모델의 제1 레이어 파라미터 및 제3 인공 지능 모델의 제1 레이어 파라미터를 바탕으로 제1 기준값을 획득할 수 있다. 또한, 프로세서(130)는 제1 인공 지능 모델의 제2 레이어 파라미터, 제2 인공 지능 모델의 제2 레이어 파라미터 및 제3 인공 지능 모델의 제2 레이어 파라미터를 바탕으로 제2 기준값을 획득할 수 있다. 같은 방법으로, 프로세서(130)는 제3 기준값 내지 제10 기준값을 획득할 수 있다. 이하 본 개시에서는 하나의 기준값이 처리되는 과정을 설명하나, 기준값이 복수개 존재하는 경우에도 본 개시의 기술적 사상이 적용될 수 있음은 물론이다.

[139] 이때, 기준값은, 레이어 파라미터의 평균값 또는 분산값을 바탕으로 획득될 수

있다. 다만, 기준값은 다양한 방법을 통해 결정될 수 있다. 예를 들어, 기준값은 레이어 파라미터 w1, w2에 대하여 Polygonal Chain(수학식 1 참조) 또는 Bezier Curve(수학식 2 참조)를 이용하여 획득될 수 있다.

[140] [수식1]

$$\Phi_{\theta}(t)=\begin{cases}2(t\theta+(0.5-t)\,w_1), 0\le t\le 0.5.\\ 2((t-0.5)\,w_2+(1-t)\theta, 0.5\le t\le 1.\end{cases}$$

[141] [수식2]

$$\Phi_{\theta}(t)=(1-t)^2\,w_1+2t(1-t)\theta+t^2\,w_2, 0\le t\le 1.$$

[142] 또는 기준 값은 하기 수식에 의해 결정될 수도 있다.

[143] [수식3]

$$p(w_{ij})\propto\frac{1}{|w_{ij}|},$$

$$q(w_{ij})=N(w_{ij}\mid \mu_{ij},\alpha\mu^2{}_{ij}).$$

[144] 설명의 편의를 위해 본 개시에서는 기준값이 레이어 파라미터의 평균값 또는 분산값을 바탕으로 획득되는 경우로 한정하여 설명하나, 상술한 다양한 방법을 통해 기준값이 결정될 수 있음은 물론이다.

[145] 프로세서(130)는 동일 경로를 가지는 레이어 파라미터별 분산값을 판단할 수 있다. 이때, 분산값이 기 설정된 값보다 낮은 경우, 기준값은 레이어 파라미터의 평균값을 바탕으로 획득될 수 있다. 즉, 분산값이 낮다는 의미는, 각각의 레이어 파라미터가 평균값과 차이가 크지 않다는 의미이므로, 프로세서(130)는 기준값을 평균값으로 획득할 수 있다. 전자 장치(200)는 서버(100)로부터 기준값을 수신하고, 기준값을 바탕으로 레이어 파라미터를 생성할 수 있다. 즉, 분산값이 기 설정된 값보다 낮은 경우, 전자 장치(200)는 레이어 파라미터 전부를 서버(100)로부터 수신하여 저장하는 대신, 기준값만을 수신하여 저장하고, 수신한 기준값으로부터 복수의 레이어 파라미터를 생성할 수 있다.

[146] 한편, 프로세서(130)는 분산값이 기 설정된 값보다 큰 경우, 해당 레이어 파라미터가 기 설정된 조건을 만족하는지 여부를 판단할 수 있다. 이때, 기 설정된 조건은 각각의 레이어 파라미터의 임팩트(Impact)값과 관련된 조건일 수 있다. 예를 들어, 프로세서(130)는 임팩트 값에 따라, Coefficient Vector를 이용하여 레이어 파라미터를 생성하는 확률적 함수를 생성할 수 있다. 즉, 프로세서(130)는 레이어 파라미터 값에 대한 레이어의 출력값을 바탕으로 확률적 출력 값을 획득할 수 있다. 임팩트값이라 함은, 결과값에 영향을 주는 레이어 파라미터에 대한 특정 값을 의미한다. 예로, 임팩트값은 하기와 같이 정의될 수 있다.

[147] [수식4]

$$\kappa_{ij} = \sigma^2_{ij} \frac{d^2}{dW^2_{ij}} [\log p(y \mid x, W, W_{net})] \mid_{W=M}$$

[148] 그러나 임팩트값은 상술한 수학식 4에 한정되는 것은 아니며, 결과값에 영향을 주는 레이어 파라미터를 구분할 수 있는 다양한 값에도 적용될 수 있음은 물론이다.

[149] 이때, 임팩트값이 기 설정된 값 이상인 경우, 프로세서(130)는 기 설정된 값 이상의 임팩트 값을 가지는 레이어 파라미터를 전자 장치(200)로 전송할 수 있다. 즉, 기 설정된 값 이상의 임팩트 값을 가지는 레이어 파라미터는 결과 값에 영향을 줄 수 있는 파라미터이므로, 전자 장치(200)는 레이어 파라미터를 기준값으로부터 생성하는 것이 아니라, 레이어 파라미터 원본을 서버(100)로부터 수신할 수 있다.

[150] 한편, 임팩트값이 기 설정된 값 이하인 경우, 프로세서(130)는 기 설정된 값 이하의 임팩트 값을 가지는 레이어 파라미터 삭제하고, 확률적 레이어 함수로 대체한다. 즉, 기 설정된 값 이하의 임팩트 값을 가지는 레이어 파라미터는 결과값에 큰 영향을 주지 않으므로, 전자 장치(200)는 서버(100)로부터 기준값 및 확률적 레이어 함수를 수신하여 저장하고, 수신된 기준값 및 확률적 레이어 함수로부터 레이어 파라미터를 생성할 수 있다.

[151] 상술한 방법을 통해 프로세서(130)는 복수개의 레이어 파라미터로부터 기준값 및 기 설정된 조건 만족 여부(예를 들어 임팩트 값)을 획득하고, 전자 장치(100)는 기준값 및 기 설정된 조건만으로 레이어를 생성할 수 있어 전자 장치의 메모리를 절약할 수 있다.

[152]

[153] 도 8은 본 개시의 일 실시예에 따른 전자 장치의 동작으로 설명하기 위한 흐름도이다.

[154] 먼저, 전자 장치(200)는 복수의 인공 지능 모델을 이용하기 위한 사용자 명령을 입력 받을 수 있다(S810). 인공 지능 모델을 이용하기 위한 사용자 명령이란, 인공 지능 모델을 이용하는 어플리케이션 실행 명령일 수 있다. 또는, 인공 지능 모델을 이용하기 위한 사용자 명령이란, 인공 지능 모델에 입력되는 입력 데이터를 선택하는 사용자 명령일 수 있다.

[155] 사용자 명령이 입력되면, 전자 장치(200)는 서버(100)로부터 수신한 기준값을 바탕으로 복수의 인공 지능 모델 각각에 포함된 레이어 중 기준값에 대응되는 레이어에 대한 레이어 파라미터를 생성할 수 있다(S820). 예를 들어, 복수의 인공 지능 모델이 제1 인공 지능 모델 내지 제m 인공 지능 모델이고, 각각의 모델은 제1 레이어 내지 제n 레이어로 구성된 경우를 가정할 수 있다. 전자 장치(200)는 제1 인공 지능 모델 내지 제m 인공 지능 모델 각각의 제1 레이어들에 대한 제1 기준값을 바탕으로 m개의 제1 레이어 파라미터를 생성하고, 같은 방법으로

m개의 제2 레이어 파라미터 내지 제n 레이어 파라미터를 생성할 수 있다.

[156] 전자 장치(200)는 생성된 인공 지능 모델에 입력 데이터를 입력하여 입력 데이터에 대한 출력 데이터를 획득할 수 있다(S830). 구체적으로, 전자 장치(100)는 레이어 파라미터가 생성되면, 레이어 구조 및 앙상블 방법에 따라 인공 지능 모델을 생성할 수 있다. 이때, 레이어의 구조는 서버(100)로부터 수신하여 저장하고 있을 수 있다. 앙상블 방법 또한 서버(100)로부터 수신하여 저장하고 있거나, 전자 장치(200)의 상황에 따라 적절한 방법을 선택하여 사용할 수 있다.

[157]

[158] 도 9는 본 개시의 일 실시예에 따른 서버의 동작을 설명하기 위한 흐름도이다.

[159] 서버(100)는 복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 복수의 레이어 파라미터를 획득할 수 있다(S910). 예를 들어, 복수의 인공 지능 모델이 제1 인공 지능 모델 내지 제m 인공 지능 모델이고, 각각의 모델은 제1 레이어 내지 제n 레이어로 구성된 경우를 가정할 수 있다. 이때, 서버(100)는 제1 인공 지능 모델의 제1 레이어에 대한 제1 레이어 파라미터, 제2 인공 지능 모델의 제1 레이어에 대한 제1 레이어 파라미터 내지 제m 인공 지능 모델의 제1 레이어에 대한 제1 레이어 파라미터를 획득할 수 있다.

[160] 서버(100)는 획득된 레이어 파라미터를 바탕으로, 대응되는 위치의 새로운 레이어에 대한 레이어 파라미터를 생성하기 위한 기준값을 획득할 수 있다(S920). 예를 들어, 서버(100)는 획득한 m개의 레이어 파라미터를 바탕으로 기준값을 획득할 수 있다. 이때, 기준값은 m개의 파라미터에 대한 평균값 및 분산 값일 수 있다.

[161] 서버(100)는 획득된 기준값을 전자 장치로 전송할 수 있다(S930). 다만, 상술한 바와 같이, 서버(100)는 기준값 이외에도 임팩트값을 획득하여 전자 장치(200)로 전송할 수 있음은 물론이다.

[162]

[163] 도 10은 본 개시의 일 실시예에 따른 서버 및 전자 장치의 동작을 설명하기 위한 시스템도이다.

[164] 서버(100)는 복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 복수의 레이어 파라미터를 획득할 수 있다(S1010).

[165] 서버(100)는 획득된 레이어 파라미터를 바탕으로, 대응되는 위치의 새로운 레이어에 대한 레이어 파라미터를 생성하기 위한 기준값을 획득할 수 있다(S1020). 서버(100)는 획득된 기준값을 전자 장치(200)로 전송할 수 있다(S1030).

[166] 전자 장치(200)는 복수의 인공 지능 모델을 이용하기 위한 사용자 명령을 입력받을 수 있다(S1040).

[167] 사용자 명령이 입력되면, 전자 장치(200)는 서버(100)로부터 수신한 기준값을

바탕으로 복수의 인공 지능 모델 각각에 포함된 레이어 중 기준값에 대응되는 레이어에 대한 레이어 파라미터를 생성할 수 있다(S1050).

[168] 전자 장치(200)는 레이어 파라미터를 바탕으로 생성된 인공 지능 모델에 입력 데이터를 입력하여 입력 데이터에 대한 출력 데이터를 획득할 수 있다(S1060).

[169]

[170] 한편, 본 개시의 실시 예를 구성하는 모든 구성 요소들이 하나로 결합하거나 결합하여 동작하는 것으로 설명되었다고 해서, 본 발명이 반드시 이러한 실시 예에 한정되는 것은 아니다. 즉, 본 개시의 목적 범위 안에서라면, 그 모든 구성 요소들이 하나 이상으로 선택적으로 결합하여 동작할 수도 있다. 또한, 그 모든 구성요소들이 각각 하나의 독립적인 하드웨어로 구현될 수 있지만, 각 구성 요소들의 그 일부 또는 전부가 선택적으로 조합되어 하나 또는 복수 개의 하드웨어에서 조합된 일부 또는 전부의 기능을 수행하는 프로그램 모듈을 갖는 컴퓨터 프로그램으로서 구현될 수도 있다.

[171] 다양한 실시예에 따른 장치(예: 모듈들 또는 그 기능들) 또는 방법(예: 동작들)의 적어도 일부는 프로그램 모듈의 형태로 컴퓨터로 판독 가능한 비일시적 판독 가능 기록매체(non-transitory computer readable media)에 저장된 명령어로 구현될 수 있다. 명령어가 프로세서(예: 프로세서(130))에 의해 실행될 경우, 프로세서가 상기 명령어에 해당하는 기능을 수행할 수 있다.

[172] 여기서, 프로그램은 컴퓨터가 읽을 수 있는 비일시적 기록매체에 저장되어 컴퓨터에 의하여 읽혀지고 실행됨으로써, 본 개시의 실시 예를 구현할 수 있다.

[173] 여기서 비일시적 판독 가능 기록매체란, 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미할 뿐만 아니라 레지스터, 캐쉬, 버퍼 등을 포함하며, 신호(signal), 전류(current) 등과 같은 전송 매개체는 포함하지 않는다.

[174] 구체적으로, 상술한 프로그램들은 CD, DVD, 하드 디스크, 블루레이 디스크, USB, 내장 메모리(예: 메모리(150)), 메모리 카드, ROM 또는 RAM 등과 같은 비일시적 판독가능 기록매체에 저장되어 제공될 수 있다.

[175] 또한, 개시된 실시예들에 따른 방법은 컴퓨터 프로그램 제품(computer program product)으로 제공될 수 있다.

[176] 컴퓨터 프로그램 제품은 S/W 프로그램, S/W 프로그램이 저장된 컴퓨터로 읽을 수 있는 저장 매체 또는 판매자 및 구매자 간에 거래되는 상품을 포함할 수 있다.

[177] 예를 들어, 컴퓨터 프로그램 제품은 전자 장치 또는 전자 장치의 제조사 또는 전자 마켓(예, 구글 플레이 스토어, 앱 스토어)을 통해 전자적으로 배포되는 S/W 프로그램 형태의 상품(예, 다운로더블 앱)을 포함할 수 있다. 전자적 배포를 위하여, S/W 프로그램의 적어도 일부는 저장 매체에 저장되거나, 임시적으로 생성될 수 있다. 이 경우, 저장 매체는 제조사 또는 전자 마켓의 서버, 또는 중계 서버의 저장매체가 될 수 있다.

[178] 이상에서는 본 개시의 실시 예에 대하여 도시하고 설명하였지만, 본 발명은

상술한 특정의 실시 예에 한정되지 아니하며, 청구범위에 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어서는 안 될 것이다.

청구범위

[청구항 1] 인공 지능 모델을 생성하기 위한 전자 장치의 제어 방법에 있어서,
복수의 인공 지능 모델을 이용하기 위한 사용자 명령이 입력되면, 상기 복수의 인공 지능 모델을 생성하는 단계;
상기 생성된 인공 지능 모델에 입력 데이터를 입력하여 상기 입력 데이터에 대한 출력 데이터를 획득하는 단계;를 포함하며,
상기 생성하는 단계는,
상기 복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 레이어 파라미터를 획득하기 위한 기준값을 바탕으로, 상기 복수의 인공 지능 모델 각각에 포함된 레이어 중 상기 대응되는 위치의 레이어들을 생성하는 제어 방법.

[청구항 2] 제1항에 있어서,
상기 기준값은, 상기 대응되는 위치의 레이어들에 대한 레이어 파라미터의 평균값 및 분산값을 바탕으로 결정되는 것을 특징으로 하는 제어 방법.

[청구항 3] 제2항에 있어서,
상기 복수의 레이어 파라미터의 분산값이 기 설정된 값 이하인 경우,
상기 기준값은 상기 복수의 레이어 파라미터의 평균값을 바탕으로 결정되는 것을 특징으로 하는 제어 방법.

[청구항 4] 제2항에 있어서,
상기 복수의 레이어 파라미터의 분산값이 기 설정된 값 이상인 경우,
상기 복수의 레이어 파라미터간의 분포를 판단하는 단계;
상기 분포가 기 설정된 조건을 만족하면, 상기 복수의 레이어 파라미터의 분산값을 바탕으로 기준값을 획득하는 단계; 를 포함하는 제어 방법.

[청구항 5] 제4항에 있어서,
상기 분포가 상기 기 설정된 조건을 만족하지 않는 경우, 상기 복수의 레이어 파라미터의 평균값 및 상기 복수의 레이어 파라미터의 분산값을 바탕으로 기준값을 획득하는 단계; 를 포함하는 제어 방법.

[청구항 6] 제1항에 있어서,
상기 생성하는 단계는,
상기 전자 장치의 메모리 용량을 바탕으로 생성되는 인공 지능 모델의 개수를 판단하는 단계; 를 더 포함하는 제어 방법.

[청구항 7] 제1항에 있어서,
상기 제어 방법은,
상기 생성된 인공 지능 모델에 입력값이 입력되면, 상기 복수의 레이어를 앙상블(ensemble)하여 결과값을 출력하는 단계; 를 더 포함하는 제어 방법.

[청구항 8] 인공 지능 모델을 생성하기 위한 전자 장치에 있어서,
통신부;
메모리; 및
외부 서버로부터, 상기 통신부를 통해 복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 레이어 파라미터를 획득하기 위한 기준값을 수신하는 프로세서; 를 포함하고,
상기 프로세서는,
복수의 인공 지능 모델을 이용하기 위한 사용자 명령이 입력되면, 상기 복수의 인공 지능 모델을 생성하고, 상기 생성된 인공 지능 모델에 입력 데이터를 입력하여 상기 입력 데이터에 대한 출력 데이터를 획득하고,
상기 복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 레이어 파라미터를 획득하기 위한 기준값을 바탕으로, 상기 복수의 인공 지능 모델 각각에 포함된 레이어 중 상기 대응되는 위치의 레이어들을 생성하는 전자 장치.

[청구항 9] 제8항에 있어서,
상기 기준값은, 상기 대응되는 위치의 레이어들에 대한 레이어 파라미터의 평균값 및 분산값을 바탕으로 결정되는 것을 특징으로 하는 전자 장치.

[청구항 10] 제9항에 있어서,
상기 프로세서는,
상기 복수의 레이어 파라미터의 분산값이 기 설정된 값 이하인 경우, 상기 기준값은 상기 복수의 레이어 파라미터의 평균값을 바탕으로 결정하는 것을 특징으로 하는 전자 장치.

[청구항 11] 제9항에 있어서,
상기 프로세서는,
상기 복수의 레이어 파라미터의 분산값이 기 설정된 값 이상인 경우, 상기 복수의 레이어 파라미터간의 분포를 판단하고, 상기 분포가 기 설정된 조건을 만족하면, 상기 복수의 레이어 파라미터의 분산값을 바탕으로 기준값을 획득하는 전자 장치.

[청구항 12] 제11항에 있어서,
상기 프로세서는, 상기 분포가 상기 기 설정된 조건을 만족하지 않는 경우, 상기 복수의 레이어 파라미터의 평균값 및 상기 복수의 레이어 파라미터의 분산값을 바탕으로 기준값을 획득하는 전자 장치.

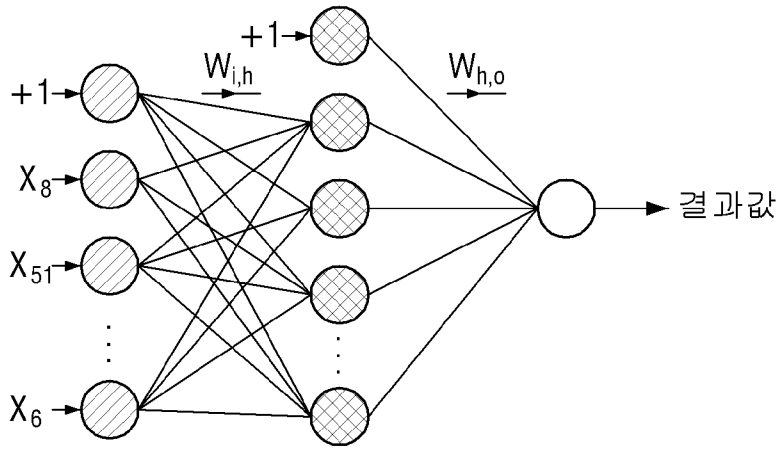
[청구항 13] 제8항에 있어서,
상기 프로세서는,
상기 메모리 용량을 바탕으로 생성되는 인공 지능 모델의 개수를 판단하는전자 장치.

[청구항 14] 제8항에 있어서,

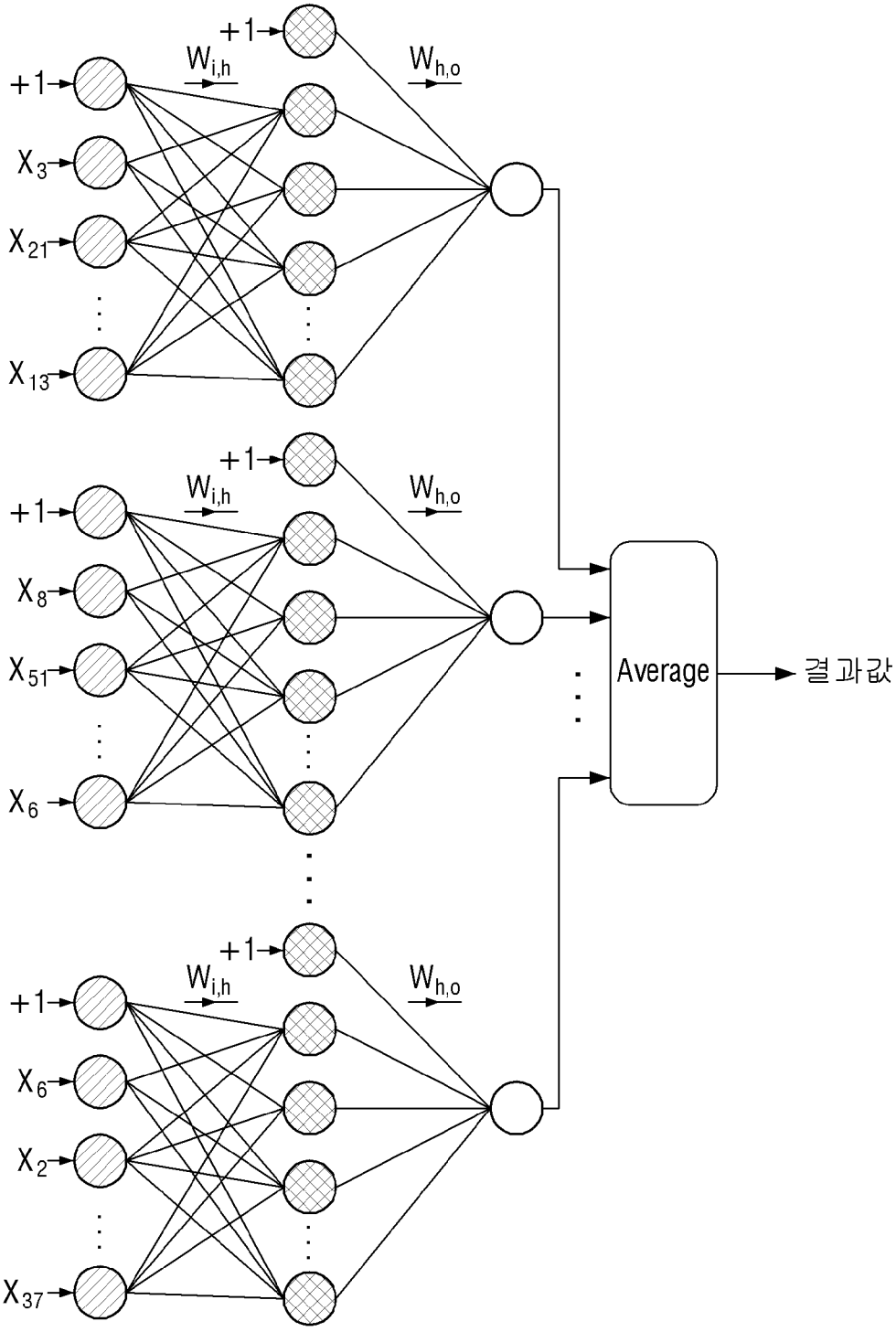
상기 프로세서는,
상기 생성된 인공 지능 모델에 입력값이 입력되면, 상기 복수의 레이어를 앙상블(ensemble)하여 결과값을 출력하는 전자 장치.

[청구항 15] 서버의 제어 방법에 있어서,
복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 복수의 레이어 파라미터를 획득하는 단계;
상기 획득된 복수의 레이어 파라미터를 바탕으로, 상기 대응되는 위치의 새로운 레이어에 대한 레이어 파라미터를 생성하기 위한 기준값을 획득하는 단계; 및
상기 획득된 기준값을 사용자 단말로 전송하는 단계;를 포함하는 제어 방법.

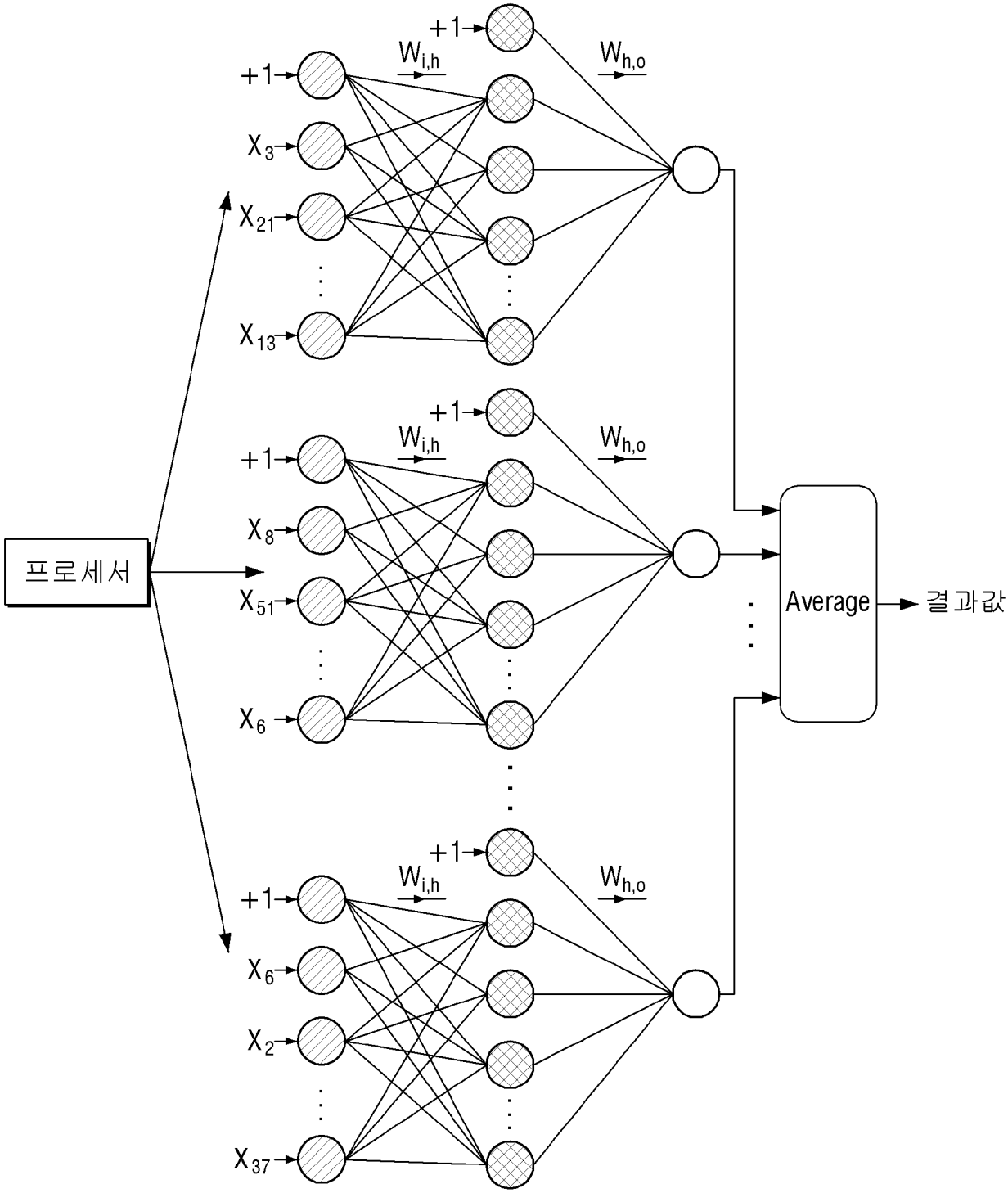
[도1a]



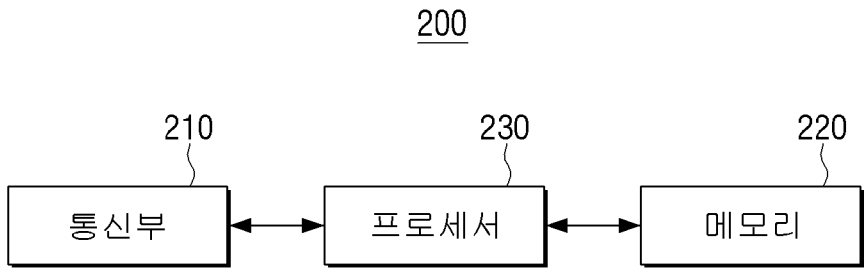
[도1b]


+1
W i,h
W h,o
+1
X 3
X 21
X 13
+1
W i,h
W h,o
+1
X 8
X 51
X 6
+1
W i,h
W h,o
+1
X 6
X 2
X 37
Average
결과값

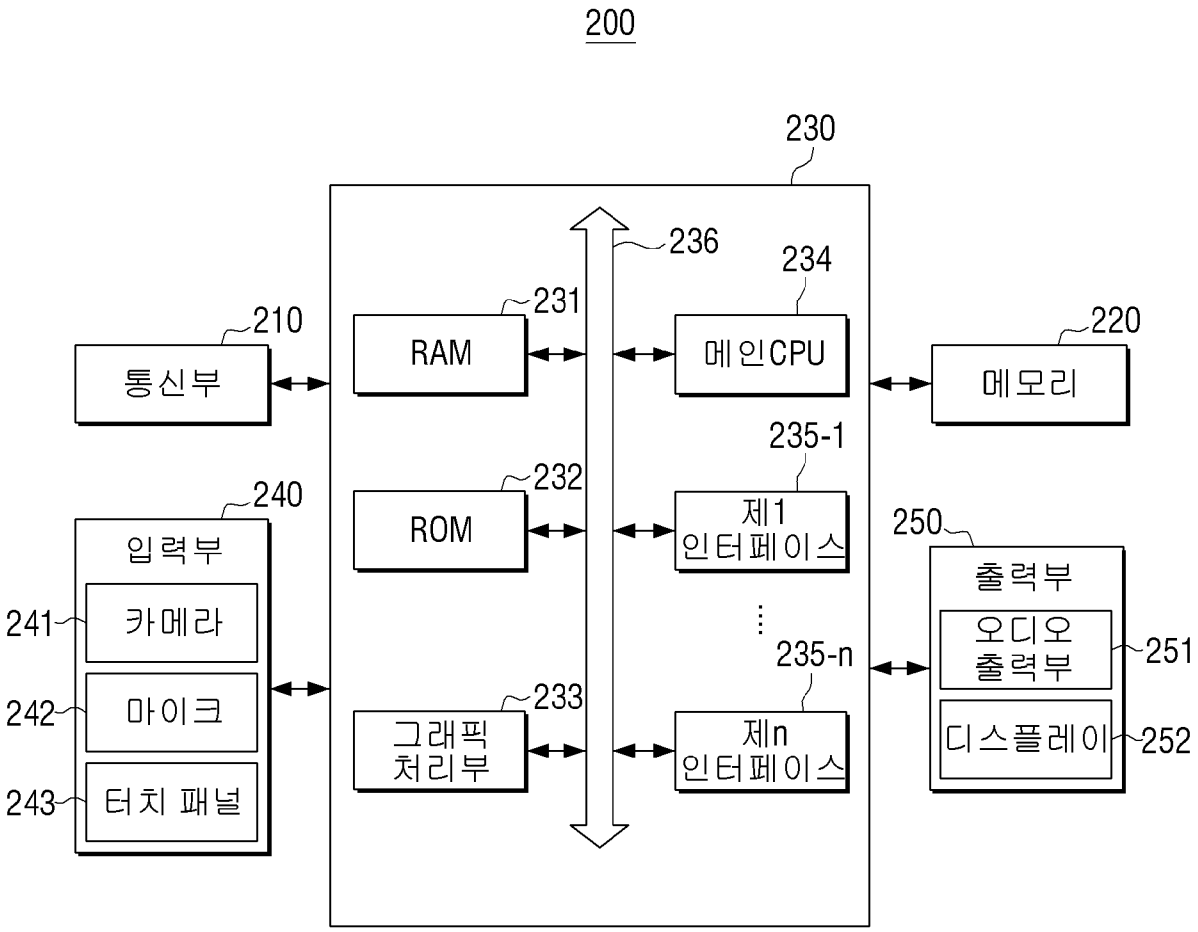
[도1c]


+1
$W_{i,h}$
$W_{h,o}$
+1
X_3
X_{21}
X_{13}
+1
$W_{i,h}$
$W_{h,o}$
+1
X_8
X_{51}
X_6
프로세서
+1
$W_{i,h}$
$W_{h,o}$
+1
X_6
X_2
X_{37}
Average
결과값

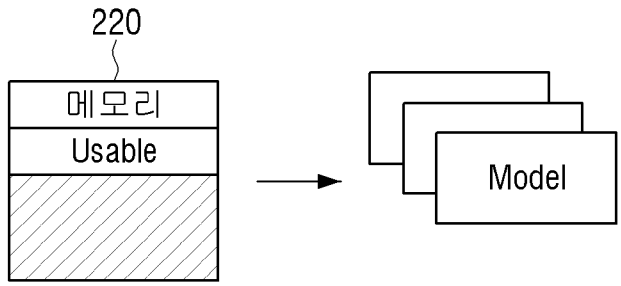
[도2]


200
210
230
220
통신부
프로세서
메모리

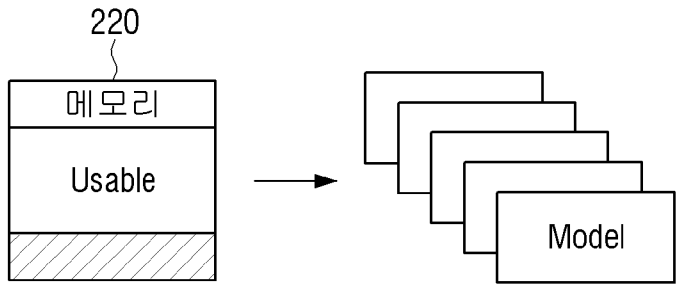
[도3]


200
230
236
234
231
210
RAM
메인CPU
220
통신부
메모리
235-1
232
240
제1
인터페이스
ROM
250
입력부
출력부
241
카메라
오디오
출력부
251
235-n
242
마이크
233
그래픽
처리부
제n
인터페이스
디스플레이
252
243
터치 패널

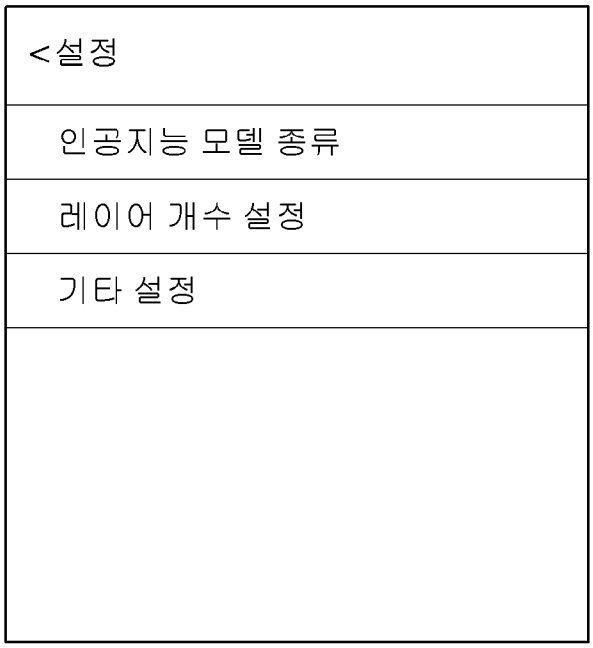
[도4a]


220
메모리
Usable
Model

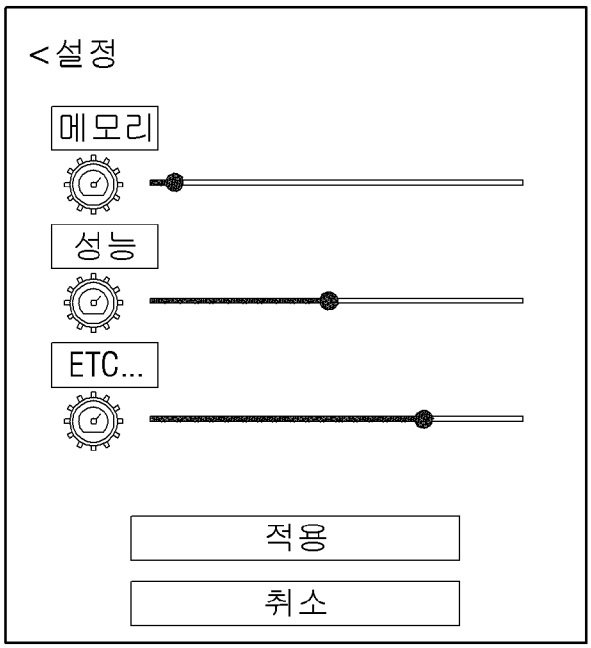
[도4b]


220
메모리
Usable
Model

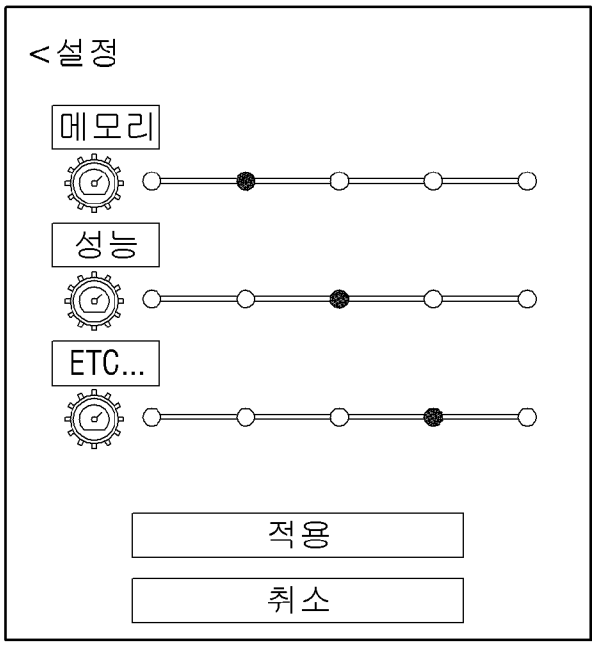
[도5a]



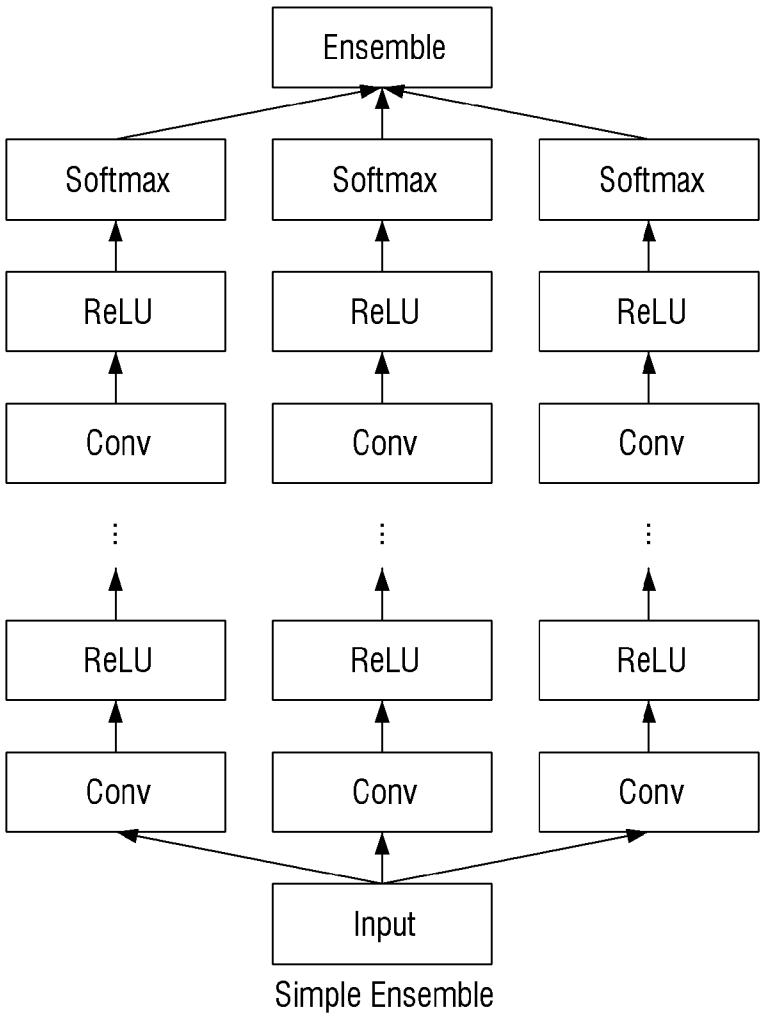
[도5b]



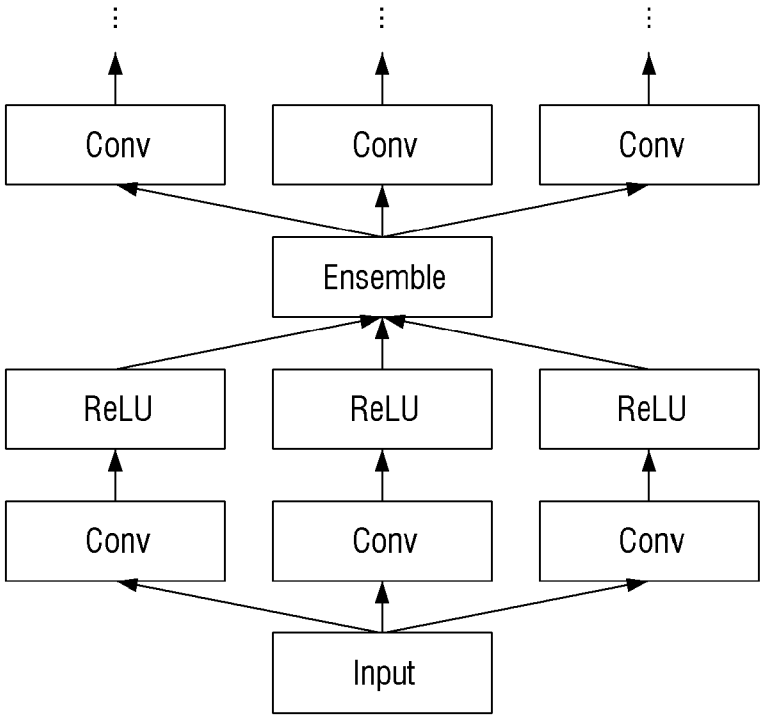
[도5c]


< 설정
메모리
성능
ETC...
적용
취소

[도6a]

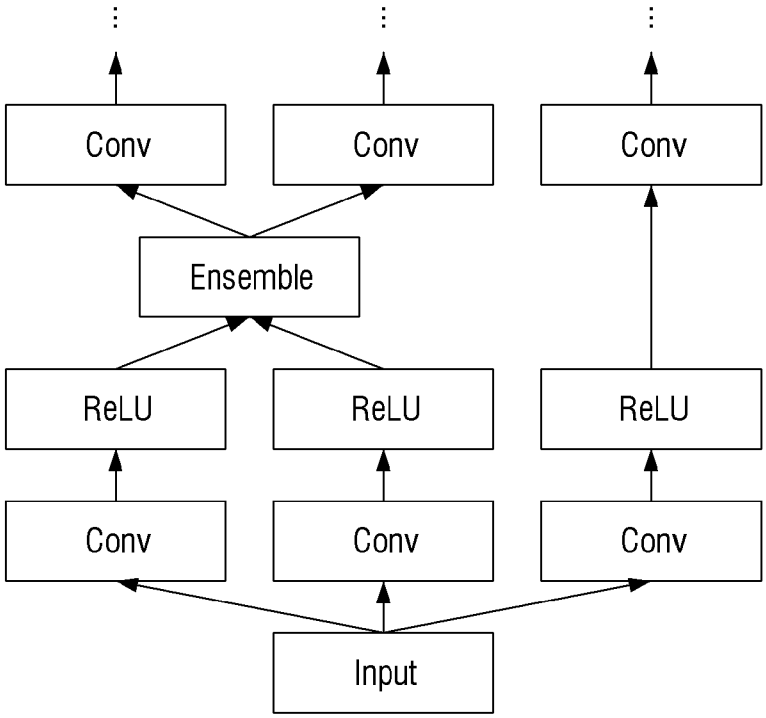

Ensemble
Softmax
Softmax
Softmax
ReLU
ReLU
ReLU
Conv
Conv
Conv
ReLU
ReLU
ReLU
Conv
Conv
Conv
Input
Simple Ensemble

[도6b]



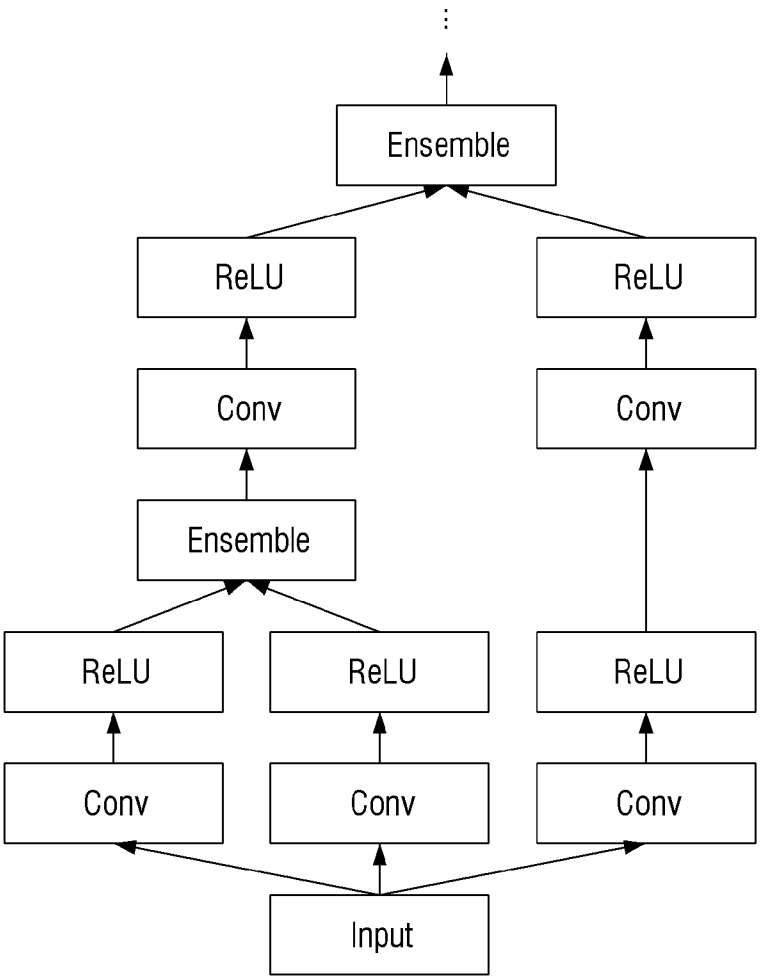
Stepwise Ensemble

[도6c]



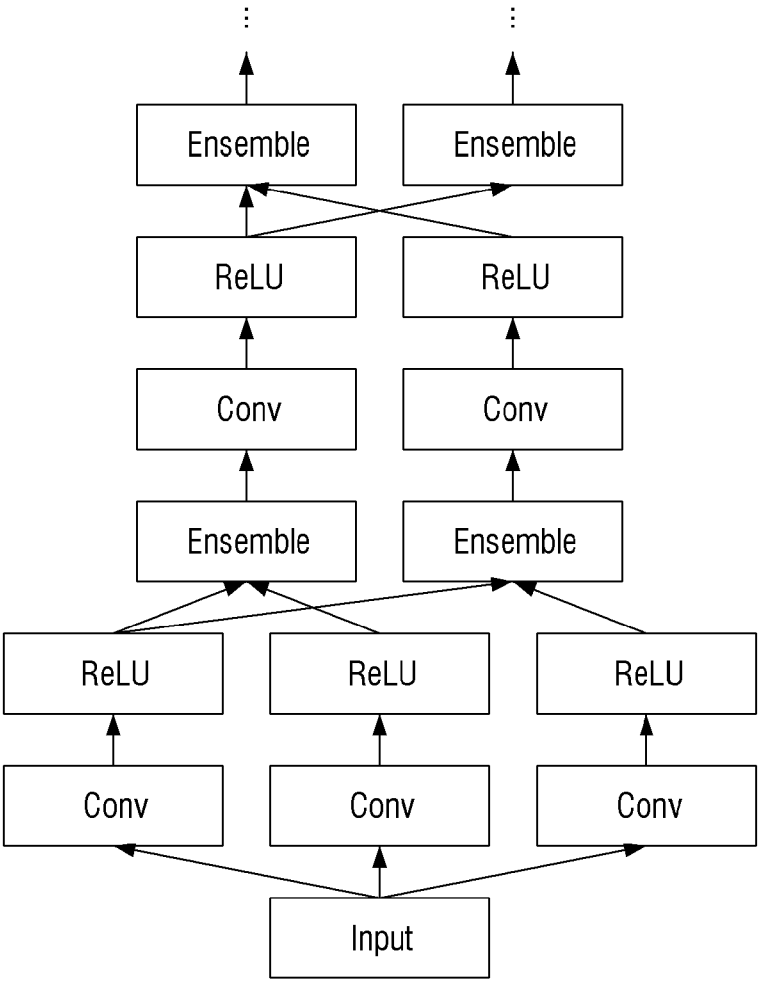
Stepwise Partial Ensemble

[도6d]



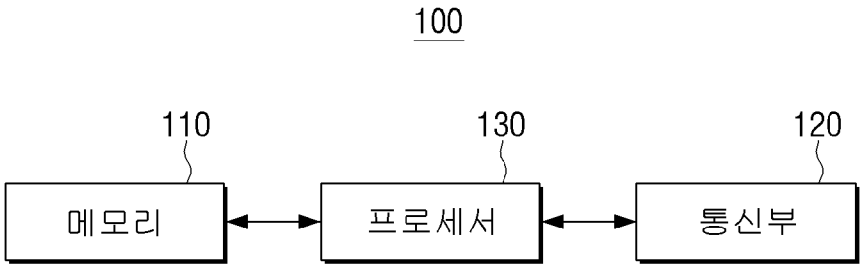
Cascaded Partial Ensemble

[도6e]

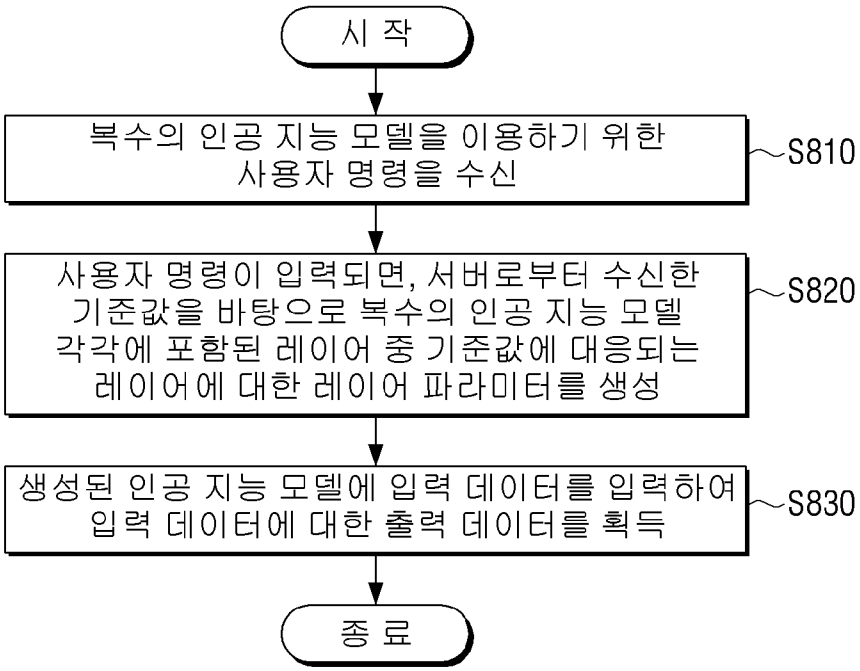


Stochastic Partial Ensemble

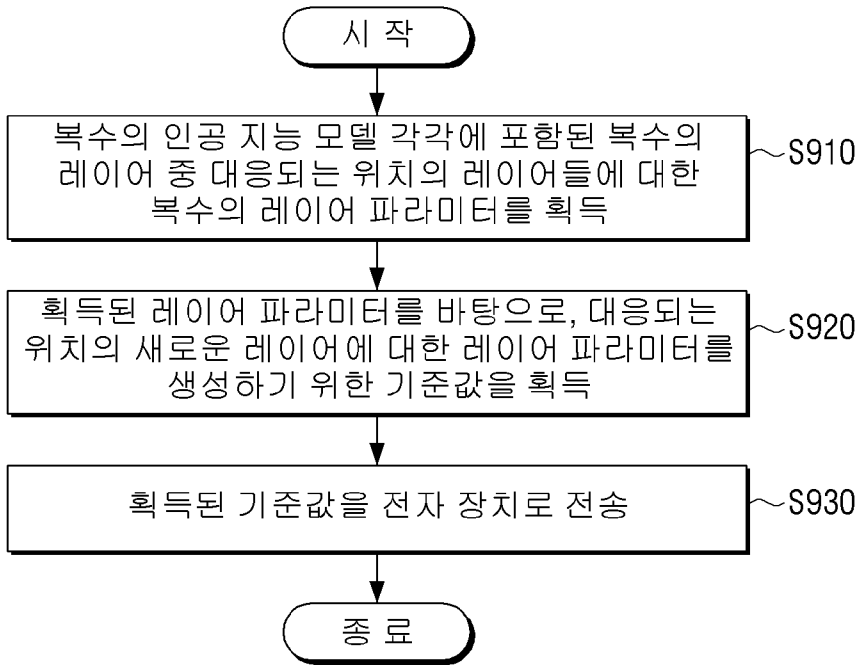
[도7]



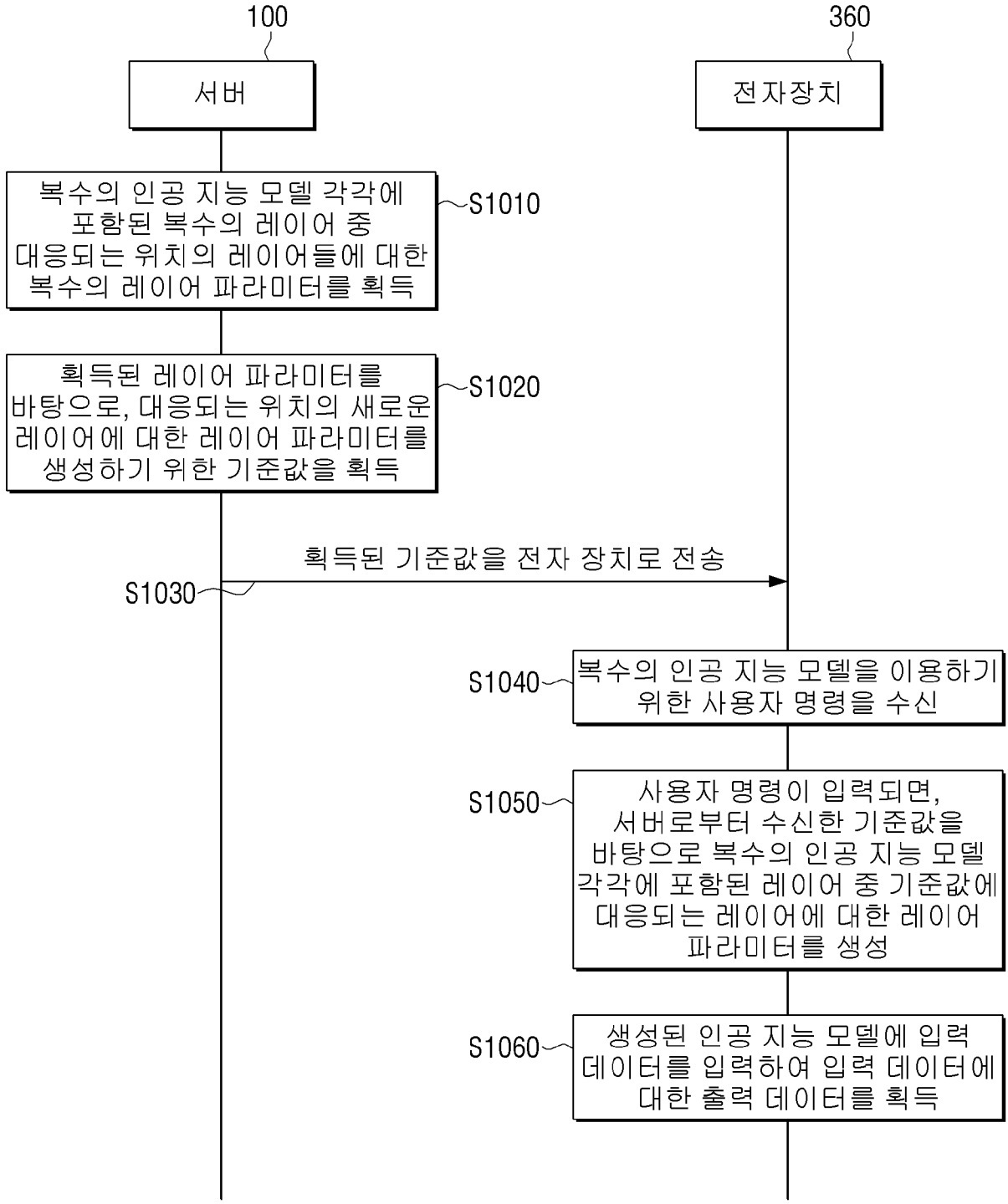
[도8]


시 작
복수의 인공 지능 모델을 이용하기 위한 사용자 명령을 수신
S810
사용자 명령이 입력되면, 서버로부터 수신한 기준값을 바탕으로 복수의 인공 지능 모델 각각에 포함된 레이어 중 기준값에 대응되는 레이어에 대한 레이어 파라미터를 생성
S820
생성된 인공 지능 모델에 입력 데이터를 입력하여 입력 데이터에 대한 출력 데이터를 획득
S830
종 료

[도9]


시 작
복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 복수의 레이어 파라미터를 획득
S910
획득된 레이어 파라미터를 바탕으로, 대응되는 위치의 새로운 레이어에 대한 레이어 파라미터를 생성하기 위한 기준값을 획득
S920
획득된 기준값을 전자 장치로 전송
S930
종 료

[도10]


100
서버
360
전자장치
복수의 인공 지능 모델 각각에 포함된 복수의 레이어 중 대응되는 위치의 레이어들에 대한 복수의 레이어 파라미터를 획득
S1010
획득된 레이어 파라미터를 바탕으로, 대응되는 위치의 새로운 레이어에 대한 레이어 파라미터를 생성하기 위한 기준값을 획득
S1020
획득된 기준값을 전자 장치로 전송
S1030
S1040
복수의 인공 지능 모델을 이용하기 위한 사용자 명령을 수신
S1050
사용자 명령이 입력되면, 서버로부터 수신한 기준값을 바탕으로 복수의 인공 지능 모델 각각에 포함된 레이어 중 기준값에 대응되는 레이어에 대한 레이어 파라미터를 생성
S1060
생성된 인공 지능 모델에 입력 데이터를 입력하여 입력 데이터에 대한 출력 데이터를 획득

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2019/006500

A. CLASSIFICATION OF SUBJECT MATTER

G06N 20/00(2019.01)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)

G06N 20/00; G06N 3/04; G06N 3/063; G06N 3/08; G06N 3/12; G06Q 50/22

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: artificial intelligence, model, layer, parameter, 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 10-1763869 B1 (THE DNA SYSTEM) 07 August 2017 See paragraphs [0039]-[0060]; and claim 1.	1,6-8,13-15
A		2-5,9-12
Y	KR 10-2018-0045635 A (SAMSUNG ELECTRONICS CO., LTD.) 04 May 2018 See paragraphs [0091], [0094]; and claim 1.	1,6-8,13-15
A	WO 2018-005433 A1 (YOUNG, Robin) 04 January 2018 See paragraph [0027]; claim 1; and figure 2.	1-15
A	KR 10-2018-0049786 A (SAMSUNG ELECTRONICS CO., LTD.) 11 May 2018 See paragraphs [0101]-[0106]; and claim 1.	1-15
A	US 2018-0053090 A1 (APPLIED BRAIN RESEARCH INC.) 22 February 2018 See claim 1.	1-15

☐ 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

"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	Date of mailing of the international search report
23 SEPTEMBER 2019 (23.09.2019)	**24 SEPTEMBER 2019 (24.09.2019)**

Name and mailing address of the ISA/KR	Authorized officer
Korean Intellectual Property Office Government Complex Daejeon Building 4, 189, Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea Facsimile No. +82-42-481-8578	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2019/006500

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-1763869 B1	07/08/2017	WO 2018-151366 A1	23/08/2018
KR 10-2018-0045635 A	04/05/2018	US 2018-0114110 A1	26/04/2018
WO 2018-005433 A1	04/01/2018	CN 109716365 A	03/05/2019
		EP 3475883 A1	01/05/2019
		US 2019-0171928 A1	06/06/2019
KR 10-2018-0049786 A	11/05/2018	US 2018-0121732 A1	03/05/2018
		WO 2018-084577 A1	11/05/2018
US 2018-0053090 A1	22/02/2018	EP 3287957 A1	28/02/2018

국 제 조 사 보 고 서

국제출원번호
PCT/KR2019/006500

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G06N 20/00(2019.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G06N 20/00; G06N 3/04; G06N 3/063; G06N 3/08; G06N 3/12; G06Q 50/22

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: 인공지능(artificial intelligence), 모델(model), 레이어(layer), 파라미터(parameter), 앙상블(ensemble)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-1763869 B1 (주식회사 더디엔에이시스템) 2017.08.07 단락 [0039]-[0060]; 및 청구항 1 참조.	1,6-8,13-15
A		2-5,9-12
Y	KR 10-2018-0045635 A (삼성전자주식회사) 2018.05.04 단락 [0091], [0094]; 및 청구항 1 참조.	1,6-8,13-15
A	WO 2018-005433 A1 (YOUNG, ROBIN) 2018.01.04 단락 [0027]; 청구항 1; 및 도면 2 참조.	1-15
A	KR 10-2018-0049786 A (삼성전자주식회사) 2018.05.11 단락 [0101]-[0106]; 및 청구항 1 참조.	1-15
A	US 2018-0053090 A1 (APPLIED BRAIN RESEARCH INC.) 2018.02.22 청구항 1 참조.	1-15

☐ 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

"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	국제조사보고서 발송일
2019년 09월 23일 (23.09.2019)	2019년 09월 24일 (24.09.2019)

ISA/KR의 명칭 및 우편주소
대한민국 특허청
(35208) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)
팩스 번호 +82-42-481-8578

심사관
권성호
전화번호 +82-42-481-3547

국제조사보고서
대응특허에 관한 정보

국제출원번호
PCT/KR2019/006500

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-1763869 B1	2017/08/07	WO 2018-151366 A1	2018/08/23
KR 10-2018-0045635 A	2018/05/04	US 2018-0114110 A1	2018/04/26
WO 2018-005433 A1	2018/01/04	CN 109716365 A EP 3475883 A1 US 2019-0171928 A1	2019/05/03 2019/05/01 2019/06/06
KR 10-2018-0049786 A	2018/05/11	US 2018-0121732 A1 WO 2018-084577 A1	2018/05/03 2018/05/11
US 2018-0053090 A1	2018/02/22	EP 3287957 A1	2018/02/28