

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국



(10) 국제공개번호

WO 2019/194356 A1

2019 년 10 월 10 일 (10.10.2019) WIPO | PCT

(51) 국제특허분류:

G06K 9/00 (2006.01)

G06T 5/00 (2006.01)

G06K 9/38 (2006.01)

G06N 3/08 (2006.01)

(21) 국제출원번호:

PCT/KR2018/006080

(22) 국제출원일:

2018 년 5 월 29 일 (29.05.2018)

(25) 출원언어:

한국어

(26) 공개언어:

한국어

(30) 우선권정보:

10-2018-0038256 2018 년 4 월 2 일 (02.04.2018) KR

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

(72) 발명자: 코멘코비아체슬라브 (KHOMENKO, Viacheslav); 01032 키예프 르바 톨스토코 스트리트 57, Kyiv (UA). 보칸코스타안틴 (BOKHAN, Kostiantyn); 61183 하르키우 에이피.27 드루즈비 나로디브 스트리트 232, Kharkiv (UA). 라디보네킴올가 (RADYVONENKO, Olga); 04209 키예프 에이피.321 보가티르스카 스트리트 6 에이, Kyiv (UA). 오사드취바덤 (OSADCHIY, Vadym); 07400 키예프 리전 브로바리 에이피.27 라구노보이 스트리트 5, Kyiv Region (UA). 비엣차니노브올렉산드

르 (VIATCHANINOV, Oleksandr); 03189 키예프 에이피.801 로모노소바 스트리트 46/1, Kyiv (UA).

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

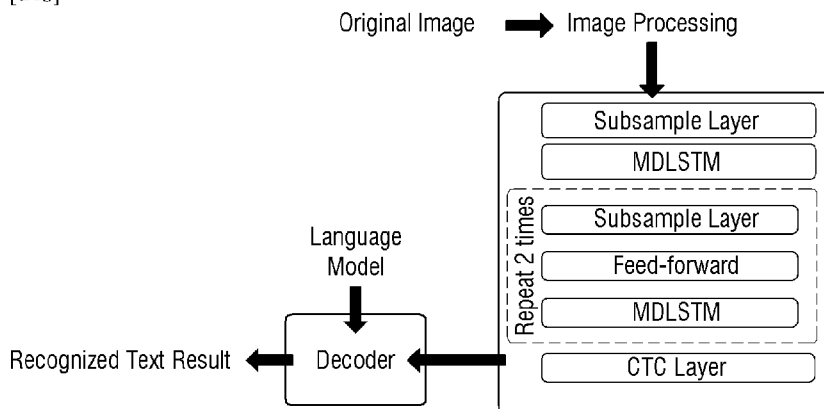
(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR,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

(54) Title: ELECTRONIC DEVICE AND CONTROL METHOD THEREOF

(54) 발명의 명칭: 전자 장치 및 그 제어 방법

[도6]



(57) Abstract: An electronic device is disclosed. The electronic device comprises: a storage unit for storing a training model of a multi-dimensional long short-term memory (MDLSTM); and a processor for acquiring an image including at least one of handwritten text and printed text, identifying each text line region in the image through image processing, and recognizing text included in the each identified text line region, on the basis of the training model.

(57) 요약서: 전자 장치가 개시된다. 전자 장치는, MDLSTM (Multi-dimensional Long Short-Term Memory) 의 트레이닝 모델이 저장된 저장부 및 필기 텍스트 및 인쇄 텍스트 중 적어도 하나를 포함하는 이미지를 획득하고, 이미지 프로세싱을 통해 이미지에서 각 텍스트 라인 영역을 식별하고, 트레이닝 모델에 기초하여 식별된 각 텍스트 라인 영역에 포함된 텍스트를 인식하는 프로세서를 포함한다.

[다음 쪽 계속]

WO 2019/194356 A1

(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML,
MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

명세서

발명의 명칭: 전자 장치 및 그 제어 방법

기술분야

- [1] 본 발명은 텍스트를 인식하는 전자 장치 및 그 제어 방법에 관한 것이다.
구체적으로, 인공신경망에 기초하여 텍스트를 인식하는 기술에 관한 것이다.

배경기술

- [2] 전자 장치에 직접 타이핑 되지 않은 텍스트를 인식하는 기술은 종래에도 존재하였다.
- [3] 다만, 종래에는 손으로 쓰여진 텍스트나 기울어진 텍스트를 제대로 인식하지 못하는 문제점이 있었다.
- [4] 인공지능 기술의 일종인 CNN (Convolutional Neural Network)에 기반한 텍스트 인식 기술이 있으나, 이 또한, 겹치게 기재된 문자 또는 손으로 쓴 필기 텍스트를 제대로 인식하지 못하는 문제점이 있었다.
- [5] 따라서, 필기 텍스트를 포함하는 다양한 형태의 텍스트를 정확하고 빠르게 인식할 수 있는 기술의 필요성이 대두되었다.

발명의 상세한 설명

기술적 과제

- [6] 본 발명은 상술한 필요성에 따른 것으로, 본 발명의 목적은, RNN(Recurrent Neural Network)의 일종인 LSTM(Long-Short Term Memory)에 기반하여 이미지에 포함된 텍스트를 정확하고 빠르게 인식할 수 있는 전자 장치 및 그 제공 방법을 제공함에 있다.

과제 해결 수단

- [7] 상술한 목적을 달성하기 위한 본 개시의 일 실시 예에 따른 전자 장치는, MDLSTM (Multi-dimensional Long Short-Term Memory)의 트레이닝 모델이 저장된 저장부 및 필기 텍스트 및 인쇄 텍스트 중 적어도 하나를 포함하는 이미지를 획득하고, 이미지 프로세싱을 통해 상기 이미지에서 각 텍스트 라인 영역을 식별하고, 상기 트레이닝 모델에 기초하여 식별된 각 텍스트 라인 영역에 포함된 텍스트를 인식하는 프로세서를 포함할 수 있다.
- [8] 상기 프로세서는, 상기 이미지에서 텍스트 영역을 이진화하고, 상기 이진화된 텍스트 영역에 대한 보정 처리에 의해 상기 이미지 프로세싱을 수행할 수 있다.
- [9] 상기 보정 처리는, 상기 텍스트 영역에 대한 기울기 제거, 높이 표준화 및 골격화(skeletonization) 중 적어도 하나를 포함할 수 있다.
- [10] 상기 프로세서는, 상기 식별된 텍스트 라인을 포함하는 영역을 복수 개의 수직 블록으로 식별하고, 각 수직 블록의 픽셀 값에 기초하여 상기 각 수직 블록에서 추정되는 문자의 확률을 산출할 수 있다.
- [11] 상기 프로세서는, 제1 수직 블록에서 제1 문자의 존재 확률, 상기 제1 수직

블록에서 상기 제1 문자의 부존재 확률, 제2 수직 블록에서 제2 문자의 존재 확률 및 상기 제2 수직 블록에서 상기 제2 문자의 부존재 확률을 조합하여 상기 텍스트 라인 영역에 제1 및 제2 문자 중 적어도 하나가 포함될 복수의 확률을 산출할 수 있다.

- [12] 상기 프로세서는, 상기 산출된 복수의 확률에 토큰 패스 디코딩(token passing decoding) 및 기설정된 언어 모델을 적용하여 상기 텍스트 라인 영역에서 인식되는 텍스트를 획득할 수 있다.
- [13] 전자 장치는, 카메라 및 디스플레이를 더 포함하고, 상기 필기 텍스트 및 인쇄 텍스트 중 적어도 하나를 포함하는 이미지는, 상기 카메라에 의해 획득된 이미지이거나, 상기 디스플레이에 입력된 필기 이미지일 수 있다.
- [14] 상기 프로세서는, 상기 인식된 텍스트와 관련된 기능을 식별하고, 상기 식별된 기능을 실행할 수 있다.
- [15] 본 개시의 일 실시 예에 따른 MDLSTM (Multi-dimensional Long Short-Term Memory)의 트레이닝 모델이 저장된 전자 장치의 제어 방법은, 필기 텍스트 및 인쇄 텍스트 중 적어도 하나를 포함하는 이미지를 획득하는 단계, 이미지 프로세싱을 통해 상기 이미지에서 각 텍스트 라인 영역을 식별하는 단계 및 상기 트레이닝 모델에 기초하여 식별된 각 텍스트 라인 영역에 포함된 텍스트를 인식하는 단계를 포함할 수 있다.
- [16] 상기 각 텍스트 라인 영역을 식별하는 단계는, 상기 이미지에서 텍스트 영역을 이진화하고, 상기 이진화된 텍스트 영역에 대한 보정 처리에 의해 상기 이미지 프로세싱을 수행할 수 있다.
- [17] 상기 보정 처리는, 상기 텍스트 영역에 대한 기울기 제거, 높이 표준화 및 골격화(skeletonization) 중 적어도 하나를 포함할 수 있다.
- [18] 상기 식별된 텍스트 라인을 포함하는 영역을 복수 개의 수직 블록으로 식별하고, 각 수직 블록의 픽셀 값에 기초하여 상기 각 수직 블록에서 추정되는 문자의 확률을 산출하는 단계를 더 포함할 수 있다.
- [19] 상기 확률을 산출하는 단계는, 제1 수직 블록에서 제1 문자의 존재 확률, 상기 제1 수직 블록에서 상기 제1 문자의 부존재 확률, 제2 수직 블록에서 제2 문자의 존재 확률 및 상기 제2 수직 블록에서 상기 제2 문자의 부존재 확률을 조합하여 상기 텍스트 라인 영역에 제1 및 제2 문자 중 적어도 하나가 포함될 복수의 확률을 산출할 수 있다.
- [20] 상기 텍스트를 인식하는 단계는, 상기 산출된 복수의 확률에 토큰 패스 디코딩(token passing decoding) 및 기설정된 언어 모델을 적용하여 상기 텍스트 라인 영역에서 인식되는 텍스트를 획득할 수 있다.
- [21] 상기 제어 방법은 상기 인식된 텍스트와 관련된 기능을 식별하고, 상기 식별된 기능을 실행하는 단계를 더 포함할 수 있다.

발명의 효과

- [22] 상술한 바와 같이 본 발명의 다양한 실시 예에 따르면 LSTM에 기반한 텍스트 인식 기술을 통해 신속하고 정확하게 인쇄된 텍스트뿐만 아니라 필기체도 인식하고 카메라에 의해 촬영된 이미지에 포함된 텍스트를 인식할 수 있으므로 사용자에게 편의를 제공할 수 있다.

도면의 간단한 설명

- [23] 도 1은 본 발명의 일 실시 예에 따른 텍스트를 포함하는 이미지가 입력된 전자 장치를 나타내는 도면이다.
- [24] 도 2a 및 도 2b는 본 발명의 일 실시 예에 따른 전자 장치의 구성을 나타내는 블록도이다.
- [25] 도 3은 본 개시의 일 실시 예에 따른 이미지 프로세싱을 설명하기 위한 도면이다.
- [26] 도 4는 본 개시의 일 실시 예에 따른 각 텍스트 라인 영역을 식별하는 동작을 설명하기 위한 도면이다.
- [27] 도 5는 본 개시의 일 실시 예에 따른 각 수직 블록에서 추정되는 문자의 확률을 산출하는 동작을 설명하기 위한 도면이다.
- [28] 도 6은 본 개시의 일 실시 예에 따른 이미지에 포함된 텍스트가 인식되는 과정을 설명하기 위한 도면이다.
- [29] 도 7은 본 개시의 일 실시 예에 따른 필기 텍스트가 인식되는 동작을 설명하기 위한 도면이다.
- [30] 도 8은 전자 장치에서 인식된 텍스트와 관련된 기능을 식별하고 식별된 기능의 실행을 설명하기 위한 도면이다.
- [31] 도 9는 본 발명의 일 실시 예에 따른 전자 장치를 제어하는 방법을 설명하기 위한 흐름도이다.

발명의 실시를 위한 최선의 형태

- [32] -

발명의 실시를 위한 형태

- [33] 이하에서는 도면을 참조하여 본 개시에 대해 상세히 설명하도록 한다. 본 개시를 설명함에 있어서, 관련된 공지 기능 혹은 구성에 대한 구체적인 설명이 본 개시의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그에 대한 상세한 설명은 생략한다. 덧붙여, 하기 실시 예는 여러 가지 다른 형태로 변형될 수 있으며, 본 개시의 기술적 사상의 범위가 하기 실시 예에 한정되는 것은 아니다. 오히려, 이들 실시 예는 본 개시를 더욱 충실하고 완전하게 하고, 당업자에게 본 개시의 기술적 사상을 완전하게 전달하기 위하여 제공되는 것이다.
- [34] 또한, 어떤 구성요소를 '포함'한다는 것은, 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있다는 것을 의미한다. 나아가, 도면에서의 다양한 요소와 영역은 개략적으로 그려진 것이다. 따라서, 본 개시의 기술적 사상은 첨부한 도면에 그려진 상대적인

크기나 간격에 의해 제한되지 않는다.

- [35] 또한, 'a, b 및 c 중 적어도 하나'라는 표현은 'a', 'b', 'c', 'a 및 b', 'a 및 c', 'b 및 c' 또는 'a, b 및 c'로 해석될 수 있다.
- [36] 이하, 첨부된 도면을 이용하여 본 개시에 대하여 구체적으로 설명한다.
- [37] 도 1은 본 발명의 일 실시 예에 따른 텍스트를 포함하는 이미지가 입력된 전자 장치를 나타내는 도면이다.
- [38] 전자 장치(100)는 텍스트가 포함된 이미지로부터 텍스트를 인식할 수 있는 스마트폰으로 구현될 수 있으나, 이에 한정되는 것은 아니며, 랩탑 컴퓨터, 태블릿, PC, 카메라 등으로 구현될 수도 있다.
- [39] 이미지에 포함된 텍스트는 필기 텍스트 및 인쇄 텍스트 중 적어도 하나일 수 있다. 여기서, 필기 텍스트는 손으로 쓰여진(written) 글씨체 형태의 텍스트이고, 인쇄 텍스트는 프린트된 글씨체 형태의 텍스트일 수 있다.
- [40] 본 개시의 일 실시 예에 따르면, 전자 장치(100)는 RNN(Recurrent Neural Network)의 일종인 LSTM(Long-Short Term Memory)에 기반하여 이미지에 포함된 텍스트를 신속하고 정확하게 인식할 수 있는데, 이하에서는 도면을 참조하여 본 개시의 다양한 실시 예에 대해 자세히 설명하도록 한다.
- [41] 여기서, RNN은 순환 신경망을 의미하며, 시계열 데이터와 같이 시간의 흐름에 따라 변화하는 데이터를 학습하기 위한 딥 러닝 모델의 일종이다. LSTM은 RNN의 히든 레이어 단계에 cell-state를 추가하여 RNN의 vanishing gradient problem을 해결하기 위한 RNN의 일종이다. 여기서, vanishing gradient problem은 RNN에서 관련 정보와 그 정보를 사용하는 지점 사이의 거리가 먼 경우 역전파시 그라디언트가 점차 줄어 학습 능력이 저하되는 문제이다.
- [42] 도 2a 및 도 2b는 본 개시의 일 실시 예에 따른 전자 장치의 구성을 나타내는 블록도이다.
- [43] 도 2a에 따르면, 전자 장치(100)는 저장부(110) 및 프로세서(120)를 포함한다.
- [44] 저장부(110)는 전자 장치(100)를 구동/제어하기 위한 다양한 데이터, 프로그램 또는 어플리케이션을 저장할 수 있다. 저장부(110)는 전자 장치(100) 및 프로세서(120)의 제어를 위한 제어 프로그램, 제조사에서 최초 제공되거나 외부에서부터 다운로드 받은 어플리케이션, 데이터베이스들 또는 관련 데이터들을 저장할 수 있다.
- [45] 특히, 저장부(110)는 MDLSTM (Multi-dimensional Long Short-Term Memory)의 트레이닝 모델을 저장할 수 있다. 또한, 저장부(110)는 카메라(미도시)에 의해 촬영되거나 전자 장치(100)의 화면에서 캡처된 이미지를 저장할 수 있다. 여기서, MDLSTM 방식은 종래 기술이므로 구체적인 설명은 생략한다.
- [46] 프로세서(120)는 전자 장치(100)의 전반적인 동작을 제어한다.
- [47] 일 실시 예에 따라 프로세서(120)는 디지털 신호를 처리하는 디지털 시그널 프로세서(digital signal processor(DSP), 마이크로 프로세서(microprocessor), TCON(Time controller)으로 구현될 수 있다. 다만, 이에 한정되는 것은 아니며,

중앙처리장치(central processing unit(CPU)), MCU(Micro Controller Unit), MPU(micro processing unit), 컨트롤러(controller), 어플리케이션 프로세서(application processor(AP)), 또는 커뮤니케이션 프로세서(communication processor(CP)), ARM 프로세서 중 하나 또는 그 이상을 포함하거나, 해당 용어로 정의될 수 있다. 또한, 프로세서(120)는 프로세싱 알고리즘이 내장된 SoC(System on Chip), LSI(large scale integration)로 구현될 수도 있고, FPGA(Field Programmable gate array) 형태로 구현될 수도 있다.

- [48] 프로세서(120)는 필기 텍스트 및 인쇄 텍스트 중 적어도 하나를 포함하는 이미지를 획득할 수 있다.
- [49] 여기서, 필기 텍스트 및 인쇄 텍스트 중 적어도 하나를 포함하는 이미지는, 카메라(미도시)에 의해 획득된 이미지이거나, 디스플레이(미도시) 상에 입력된 텍스트를 포함하는 이미지일 수 있다. 또는, 디스플레이 상의 화면이 캡처된 이미지일 수 있다.
- [50] 예를 들어, 프로세서(120)는 카메라에 의해 촬영된 이미지에 포함된 텍스트를 인식할 수 있고, 디스플레이 상에 손 터치 또는 펜 터치에 의해 입력된 텍스트를 인식할 수도 있다. 또한, 프로세서(120)는 디스플레이 상에 디스플레이되는 화면을 캡처한 이미지에 포함된 텍스트를 인식할 수도 있다.
- [51] 프로세서(120)는 이미지 프로세싱을 통해 이미지에서 각 텍스트 라인 영역을 식별할 수 있다.
- [52] 일 실시 예에 따르면, 프로세서(120)는 이미지에서 텍스트 영역을 이진화하고, 이진화된 텍스트 영역에 대한 보정 처리에 의해 이미지 프로세싱을 수행할 수 있다.
- [53] 여기서, 이진화는 텍스트 영역과 나머지 영역을 구분하는 동작일 수 있다. 즉, 프로세서(120)는 이미지에서 텍스트 영역을 식별하여 텍스트 및 나머지 영역을 구별하는 이진화 작업을 수행할 수 있다.
- [54] 보정 처리는, 텍스트 영역에 대한 기울기 제거, 높이 표준화 및 골격화(skeletonization) 중 적어도 하나를 포함할 수 있다.
- [55] 예를 들어, 프로세서(120)는 수평면에서 30도 기울어진 텍스트 라인의 기울기가 0도가 되도록 기울기를 제거하고, 각각의 텍스트 높이를 일정하게 표준화하는 보정을 수행할 수 있다.
- [56] 골격화는 텍스트를 형성하는 픽셀이 동일한 컬러의 픽셀들로 촘촘하게 구성되도록 하는 동작이다. 예를 들어, 자음"ㄱ"을 형성하는 픽셀이 검정색인 경우, "ㄱ"을 확대하면 "ㄱ"을 형성하는 모든 픽셀이 검정색으로 구성되지 않을 수 있다. 일 예로, 연필을 통해 낮은 압력으로 쓰여진 "ㄱ"은 흐릿하게 표시되므로 "ㄱ"을 형성하는 모든 픽셀들이 검정색으로 구성되는 것은 아닐 수 있다. 프로세서(120)는 이를 명확하게 하기 위해 "ㄱ"을 형성하는 픽셀들의 컬러를 모두 검정색으로 보정 처리할 수 있다. 이러한 골격화 작업을 통해 프로세서(120)는 명확하게 해당 문자를 표현하고 인식할 수 있다. 일 예로,

프로세서(120)는 골격화를 통해 텍스트를 형성하는 픽셀을 검정색 픽셀, 나머지 영역을 나타내는 픽셀을 흰색 픽셀로 제공할 수 있다.

[57] 프로세서(120)는 상술한 이미지 프로세싱을 통해 보정 처리된 텍스트 영역에서 각각의 텍스트 라인 영역을 식별할 수 있다.

[58] 이미지 프로세싱이 수행된 이미지에 포함된 텍스트 라인이 복수인 경우 프로세서(120)는 각 텍스트 라인을 개별적으로 식별하여 텍스트 인식률을 높일 수 있다.

[59] 예를 들어, 프로세서(120)는 이미지 프로세싱을 통해 표준화된 텍스트 높이의 평균 값을 산출하고, 산출된 평균 값 높이의 수평 라인, 평균 값 높이의 수평 라인을 기준으로 상, 하 방향으로 균등한 간격으로 이격된 수평 라인들을 텍스트 라인 영역에 적용하여 각 텍스트 라인을 식별할 수 있다. 즉, 프로세서(120)는 3개의 라인을 텍스트 라인 영역에 적용하여 3개의 라인에 걸치는 텍스트를 하나의 텍스트 라인으로 식별할 수 있다. 이에 관하여 후술하는 도 4에서 자세히 설명하도록 한다.

[60] 프로세서(120)는 식별된 텍스트 라인을 포함하는 영역을 복수 개의 수직 블록으로 식별하고, 각 수직 블록의 픽셀 값에 기초하여 각 수직 블록에서 추정되는 문자의 확률을 산출할 수 있다. 프로세서(120)는 MDMDLSTM(Multi-dimensional Multi-directional Long Short-Term Memory)에 기반하여 각 수직 블록에서 추정되는 문자의 확률을 산출할 수 있다. 여기서, "Multi-directional"은 전자 장치(100)에 입력된 이미지를 여러 방향에서 스캔함을 의미하고, 이에 따라 프로세서(120)는 이미지에 포함된 텍스트의 인식률을 높일 수 있다.

[61] 예를 들어, "The cat"이라는 필기 텍스트를 포함하는 영역에 26개의 수직 블록을 적용하고, 프로세서(120)는 각 수직 블록 상의 픽셀 값에 기초하여 각 수직 블록에서 추정되는 문자의 확률을 산출할 수 있다.

[62] 프로세서(120)는 제1 수직 블록에서 제1 문자의 존재 확률, 제1 수직 블록에서 제1 문자의 부존재 확률, 제2 수직 블록에서 제2 문자의 존재 확률 및 제2 수직 블록에서 제2 문자의 부존재 확률을 조합하여 텍스트 라인 영역에 제1 및 제2 문자 중 적어도 하나가 포함될 복수의 확률을 산출할 수 있다.

[63] 프로세서(120)는 산출된 복수의 확률에 토큰 패스 디코딩(token passing decoding) 및 기설정된 언어 모델을 적용하여 텍스트 라인 영역에서 인식되는 텍스트를 획득할 수 있다.

[64] 프로세서(120)는 토큰 패스 디코딩에 기초하여 각 수직 블록에서 추정되는 문자 중 가장 높은 확률의 문자를 식별할 수 있다. 특정 수직 블록에서 가장 높은 확률이 해당 블록에서 문자의 부존재인 경우, 프로세서(120)는 해당 블록에서 문자가 부존재하는 것으로 식별할 수 있다. 이에 따라, 프로세서(120)는 각 블록마다 가장 높은 확률로 추정되는 문자를 식별할 수 있다.

[65] 예를 들어, 첫번째 수직 블록에서 어떠한 문자도 포함되지 않을 확률이 가장

높은 경우 프로세서(120)는 첫번째 수직 블록에서 문자가 부존재하는 것으로 식별하고, 세번째 수직 블록에서 알파벳 "T"의 존재 확률이 가장 높은 경우 프로세서(120)는 세번째 수직 블록에서 "T"가 존재하는 것으로 식별할 수 있다. 이에 관하여 후술하는 도 5에서 자세히 설명하도록 한다.

- [66] 프로세서(120)는 각 수직 블록에서 가장 높은 확률로 추정되는 문자를 조합하여 후보 텍스트를 제공할 수 있다. 예를 들어, 프로세서(120)는 "The eat"을 후보 텍스트로 제공할 수 있다.
- [67] 프로세서(120)는 기설정된 언어 모델에 기초하여 제공된 후보 텍스트를 수정할 수 있다. 예를 들어, 프로세서(120)는 언어 모델에 기초하여 "The eat"을 "The cat"으로 수정할 수 있다.
- [68] 여기서, 토큰 패스 디코딩은 기설정된 언어 모델과 Viterbi 알고리즘을 사용하여 각 수직 블록 상의 픽셀 값에 기초하여 텍스트 라인 영역의 텍스트를 획득할 수 있다.
- [69] 여기서, 기설정된 언어 모델은 실제 사용자가 사용하는 언어, 예를 들어 문장, 구 등을 모델링한 데이터일 수 있다. 언어 모델을 이용하면 순차적으로 입력된 텍스트에 기초하여 입력된 텍스트의 다음 워드로 가장 적절한 텍스트를 식별할 수 있다. 본 개시의 일 실시 예에 따르면, 기설정된 언어 모델은 N-Grams 언어 모델일 수 있으나 이에 한정되는 것은 아니다. 상술한 Viterbi 알고리즘 및 N-Grams 언어 모델은 종래 기술이므로 자세한 설명은 생략하도록 한다.
- [70] 프로세서(120)는 이미지에 필기 텍스트 및 인쇄 텍스트가 모두 존재하는 경우, 필기 텍스트 및 인쇄 텍스트를 모두 인식할 수 있다. 또한, 프로세서(120)는 이미지에 서로 다른 언어의 텍스트가 포함된 경우, 각 텍스트의 언어를 식별하여 이미지에 포함된 텍스트를 인식할 수도 있다.
- [71] 프로세서(120)는 상술한 과정을 통해 식별된 텍스트를 제공할 수 있다. 예를 들어, 프로세서(120)는 식별된 텍스트를 디스플레이(미도시) 상에 제공하거나, 식별된 텍스트의 의미를 인식하여 이에 대응되는 동작을 수행할 수 있다.
- [72] 즉, 프로세서(120)는 인식된 텍스트와 관련된 기능을 식별하고, 식별된 기능을 실행할 수 있다.
- [73] 예를 들어, 식별된 텍스트가 "아침 6시에 기상"인 경우, 프로세서(120)는 오전 6시에 알람을 설정할 수 있다. 또는, 식별된 텍스트가 "AAA와 9월 25일 2시에 미팅"인 경우, 프로세서(120)는 스케줄러 상의 9월 25일에 AAA와 2시 미팅이라는 스케줄을 등록할 수도 있다.
- [74] 또한, 식별된 텍스트가 "쇼핑 리스트: 계란, 우유, 사과, 빵"인 경우, 프로세서(120)는 이를 쇼핑 리스트로 인식하고 메모장에 "쇼핑 리스트: 계란, 우유, 사과, 빵"을 저장할 수 있다.
- [75] 도 2b는 도 2a의 전자 장치의 세부 구성의 일 예를 나타내는 블록도이다.
- [76] 도 2b에 따르면, 전자 장치(100)는 저장부(110), 프로세서(120), 통신부(130), 카메라(140) 및 디스플레이(150)를 포함한다. 도 2b에 도시된 구성 중, 도 2a에

도시된 구성과 중복되는 구성에 대해서는 자세한 설명을 생략하도록 한다.

- [77] 저장부(110)는 프로세서(120)에 포함된 롬(ROM), 램(RAM) 등의 내부 메모리로 구현되거나, 프로세서(120)와 별도의 메모리로 구현될 수도 있다. 이 경우, 저장부(110)는 데이터 저장 용도에 따라 전자 장치(100)에 임베디드된 메모리 형태로 구현되거나, 전자 장치(100)에 탈부착이 가능한 메모리 형태로 구현될 수도 있다. 예를 들어, 전자 장치(100)의 구동을 위한 데이터의 경우 전자 장치(100)에 임베디드된 메모리에 저장되고, 전자 장치(100)의 확장 기능을 위한 데이터의 경우 전자 장치(100)에 탈부착이 가능한 메모리에 저장될 수 있다. 한편, 전자 장치(100)에 임베디드된 메모리의 경우 비휘발성 메모리, 휘발성 메모리, 하드 디스크 드라이브(HDD) 또는 솔리드 스테이트 드라이브(SSD) 등과 같은 형태로 구현될 수 있다.
- [78] 프로세서(120)는 CPU(121), 전자 장치(100)의 제어를 위한 제어 프로그램이 저장된 롬(ROM, 또는 비 휘발성 메모리) 및 전자 장치(100)의 외부에서부터 입력되는 데이터를 저장하거나 전자 장치(100)에서 수행되는 다양한 작업에 대응되는 저장 영역으로 사용되는 램(RAM, 또는 휘발성 메모리)을 포함할 수 있다.
- [79] 프로세서(120)는 기설정된 이벤트가 발생되면 저장부(110)에 저장된 OS(Operating System), 프로그램 및 다양한 어플리케이션을 실행할 수 있다. 프로세서(120)는 싱글 코어, 듀얼 코어, 트리플 코어, 쿼드 코어 및 그 배수의 코어를 포함할 수 있다.
- [80] CPU(121)는 저장부(110)에 액세스하여, 저장부(110)에 저장된 O/S를 이용하여 부팅을 수행한다. 그리고, 저장부(110)에 저장된 각종 프로그램, 콘텐츠, 데이터 등을 이용하여 다양한 동작을 수행한다.
- [81] 통신부(130)는 외부 전자 장치와 통신을 수행한다. 통신부(130)를 통해 외부 전자 장치로부터 텍스트가 포함된 이미지를 수신하여 프로세서(120)에 의해 이미지에 포함된 텍스트가 인식될 수 있다. 또한, 전자 장치(100)상에 입력된 이미지에 포함된 텍스트를 인식하고 통신부(130)를 통해 외부 전자 장치로 인식된 텍스트 또는 텍스트와 관련된 정보를 제공할 수 있다.
- [82] 여기서, 통신부(130)는 BT(BlueTooth), WI-FI(Wireless Fidelity), Zigbee, IR(Infrared) 와 같은 통신 방식으로 외부 전자 장치(미도시)와 무선 통신을 수행할 수 있으나, Serial Interface, USB(Universal Serial Bus), NFC(Near Field Communication) 등과 같은 다양한 통신 방식을 통해 외부 전자 장치와 통신을 수행할 수도 있다.
- [83] 예를 들어, 통신부(130)는 기설정된 이벤트가 발생하면, 외부 전자 장치와의 기정의된 통신 방식에 따라 통신을 수행하여 연동 상태가 될 수 있다. 여기서, 연동은 전자 장치(100)와 외부 전자 장치 간에 통신이 초기화되는 동작, 네트워크가 형성되는 동작, 기기 페어링이 수행되는 동작 등 통신이 가능한 상태가 되는 모든 상태를 의미할 수 있다. 예를 들어, 외부 전자 장치의 기기 식별

- 정보가 전자 장치(100)로 제공되고, 그에 따라 양 기기 간의 페어링 절차가 수행될 수 있다. 예를 들어, 전자 장치(100) 또는 외부 전자 장치에서 기설정된 이벤트가 발생하면, DLNA(Digital Living Network Alliance) 기술을 통해 주변 기기를 탐색하고 탐색된 기기와 페어링을 수행하여 연동 상태가 될 수 있다.
- [84] 카메라(140)는 필기 텍스트 및 인쇄 텍스트 중 적어도 하나를 포함하는 이미지를 촬영하여 획득할 수 있다.
- [85] 디스플레이(150)는 LCD(liquid crystal display), OLED(organic light-emitting diode), LCoS(Liquid Crystal on Silicon), DLP(Digital Light Processing), QD(quantum dot) 디스플레이 패널 등과 같은 다양한 형태로 구현될 수 있다.
- [86] 디스플레이(150)는 터치 패드와 상호 레이어 구조를 이루는 터치 스크린 형태로 구현될 수 있다. 이 경우, 디스플레이(150)는 출력 장치 이외에 사용자 인터페이스(미도시)로 사용될 수 있다. 여기서, 터치 스크린은 터치 입력 위치 및 면적뿐만 아니라 터치 입력의 압력까지도 검출할 수 있다.
- [87] 이에 따라, 손 터치 또는 펜 터치에 의해 디스플레이(150) 상에 입력되는 텍스트는 프로세서(120)에 의해 인식될 수 있다.
- [88] 도 3은 본 개시의 일 실시 예에 따른 이미지 프로세싱을 설명하기 위한 도면이다.
- [89] 도 3에 따르면, 전자 장치(100)는 이미지 프로세싱을 수행하여 이미지에 포함된 텍스트를 좀 더 명확하게 표현할 수 있다.
- [90] 원본 이미지(310)는 다양한 방법을 통해 전자 장치(100)에 입력될 수 있다. 예를 들어, 카메라(140)에 의해 촬영되거나 전자 장치(100)의 디스플레이(150) 상에 손 터치 또는 펜 터치에 의해 텍스트가 입력되거나, 타 전자 장치로부터 텍스트가 포함된 원본 이미지를 수신할 수도 있다.
- [91] 전자 장치(100)는 이미지에서 텍스트 영역을 이진화하고, 이진화된 텍스트 영역에 대한 보정 처리에 의해 이미지 프로세싱을 수행할 수 있다.
- [92] 여기서, 이진화는 텍스트 영역과 나머지 영역을 구분하는 동작일 수 있다. 즉, 전자 장치(100)는 이미지에서 텍스트 영역을 식별하여 텍스트 및 나머지 영역을 구별하는 이진화 작업을 수행할 수 있다.
- [93] 여기서, 보정 처리는, 텍스트 영역에 대한 기울기 제거(330), 높이 표준화(340) 및 골격화(350) 중 적어도 하나를 포함할 수 있다.
- [94] 예를 들어, 수평면에서 30도 기울어진 텍스트 라인의 기울기가 0도가 되도록 기울기를 제거하고, 각각의 텍스트 높이를 일정하게 표준화하는 보정을 수행할 수 있다.
- [95] 골격화는 텍스트를 형성하는 픽셀이 동일한 컬러의 픽셀들로 촘촘하게 구성되도록 하는 동작이다. 예를 들어, 자음 "ㄱ"을 형성하는 픽셀이 검정색인 경우, "ㄱ"을 확대하면 "ㄱ"을 형성하는 모든 픽셀이 검정색으로 구성되지 않을 수 있다. 일 예로, 연필을 통해 낮은 압력으로 쓰여진 "ㄱ"은 흐릿하게 표시되므로 "ㄱ"을 형성하는 모든 픽셀들이 검정색으로 구성되는 것은 아닐 수

있다. 전자 장치(100)는 이를 명확하게 하기 위해 "ㄱ"을 형성하는 픽셀들의 컬러를 모두 검정색으로 보정 처리할 수 있다. 이러한 골격화 작업을 통해 전자 장치(100)는 명확하게 해당 문자를 표현하고 인식할 수 있다. 일 예로, 전자 장치(100)는 골격화를 통해 텍스트를 형성하는 픽셀을 검정색 픽셀, 나머지 영역을 나타내는 픽셀을 흰색 픽셀로 제공할 수 있다.

[96] 도 4는 본 개시의 일 실시 예에 따른 각 텍스트 라인 영역을 식별하는 동작을 설명하기 위한 도면이다.

[97] 전자 장치(100)는 이미지 프로세싱이 수행된 이미지(410)에 포함된 텍스트 라인이 복수인 경우, 각 텍스트 라인을 개별적으로 식별하여 텍스트 인식률을 높일 수 있다.

[98] 예를 들어, 전자 장치(100)는 이미지 프로세싱을 통해 표준화된 텍스트 높이의 평균 값을 산출하고, 산출된 평균 값 높이의 수평 라인, 평균 값 높이의 수평 라인을 기준으로 상, 하 방향으로 균등한 간격으로 이격된 수평 라인들을 텍스트 라인 영역에 적용하여 각 텍스트 라인을 식별할 수 있다. 즉, 전자 장치(100)는 3개의 라인을 텍스트 라인 영역에 적용하여 3개의 라인에 걸치는 텍스트를 하나의 텍스트 라인으로 식별할 수 있다(420).

[99] 예를 들어, 도 4에 따르면, 전자 장치(100)는 "pleasure and pain will", "give you a complete" 및 "account to the system"을 개별적인 텍스트 라인으로 식별할 수 있다.

[100] 전자 장치(100)는 첫번째 텍스트 라인 만을 독립적으로 식별할 수 있다(430). 각 텍스트 라인 영역이 개별적으로 식별되지 않는 경우, 전자 장치(100)의 텍스트 인식률이 감소될 수 있다.

[101] 예를 들어, 각 텍스트 라인 영역이 개별적으로 식별되지 않는 경우, 두번째 줄의 텍스트가 첫번째 줄의 텍스트의 인식을 방해하여 전자 장치(100)는 첫번째 줄의 텍스트를 잘못 인식할 수 있다.

[102] 도 5는 본 개시의 일 실시 예에 따른 각 수직 블록에서 추정되는 문자의 확률을 산출하는 동작을 설명하기 위한 도면이다.

[103] 도 5에 따르면, 텍스트 영역(510)은 이미지 프로세싱이 수행된 텍스트일 수 있다.

[104] 전자 장치(100)는 식별된 텍스트 라인을 포함하는 영역을 복수 개의 수직 블록으로 식별하고, 각 수직 블록의 픽셀 값에 기초하여 각 수직 블록에서 추정되는 문자의 확률을 산출할 수 있다.

[105] 도 5에 따르면, "The cat"이라는 필기 텍스트를 포함하는 영역에 26개의 수직 블록을 적용하고, 전자 장치(100)는 각 수직 블록 상의 픽셀 값에 기초하여 각 수직 블록에서 추정되는 문자의 확률을 산출할 수 있다.

[106] 전자 장치(100)는 제1 수직 블록에서 제1 문자의 존재 확률, 제1 수직 블록에서 제1 문자의 부존재 확률, 제2 수직 블록에서 제2 문자의 존재 확률 및 제2 수직 블록에서 제2 문자의 부존재 확률을 조합하여 텍스트 라인 영역에 제1 및 제2 문자 중 적어도 하나가 포함될 복수의 확률을 산출할 수 있다.

- [107] 예를 들어, 첫번째 수직 블록에서 어떠한 문자도 포함되지 않을 확률이 가장 높은 경우 전자 장치(100)는 첫번째 수직 블록에서 문자가 부존재하는 것으로 식별하고, 세번째 수직 블록에서 알파벳 "T"의 존재 확률이 가장 높은 경우 전자 장치(100)는 세번째 수직 블록에서 "T"가 존재하는 것으로 식별할 수 있다. 전자 장치(100)는 각 수직 블록에서 가장 높은 확률로 존재 또는 부존재하는 문자를 식별하여 조합할 수 있다.
- [108] 전자 장치(100)는 520에 따른 복수의 확률을 각 수직 블록마다 식별하여 각 수직 블록에서 가장 높은 확률의 문자를 식별할 수 있다.
- [109] 전자 장치(100)는 산출된 복수의 확률에 토큰 패스 디코딩(token passing decoding) 및 기설정된 언어 모델을 적용하여 텍스트 라인 영역에서 인식되는 텍스트를 획득할 수 있다.
- [110] 전자 장치(100)는 토큰 패스 디코딩에 기초하여 각 수직 블록에서 추정되는 문자 중 가장 높은 확률의 문자를 식별할 수 있다. 특정 수직 블록에서 가장 높은 확률이 해당 블록에서 문자의 부존재인 경우, 전자 장치(100)는 해당 블록에서 문자가 부존재하는 것으로 식별할 수 있다. 이에 따라, 전자 장치(100)는 각 블록마다 가장 높은 확률로 추정되는 문자를 식별할 수 있다. 전자 장치(100)는 가장 높은 확률로 추정되는 문자를 조합하여 후보 텍스트를 제공할 수 있다.
- [111] 전자 장치(100)는 기설정된 언어 모델에 기초하여 제공된 후보 텍스트를 수정할 수 있다. 예를 들어, 각 블록마다 가장 높은 확률로 추정되는 문자에 기초하여 식별된 후보 텍스트가 "The eat"인 경우, 전자 장치(100)는 언어 모델에 기초하여 "The eat"을 "The cat"으로 수정할 수 있다.
- [112] 도 6은 본 개시의 일 실시 예에 따른 이미지에 포함된 텍스트가 인식되는 과정을 설명하기 위한 도면이다.
- [113] 원본 이미지(Original Image)가 전자 장치(100)에 입력되면, 전자 장치(100)는 이미지 프로세싱(Image Processing) 동작을 수행할 수 있다.
- [114] 이후, 전자 장치(100) MDLSTM에 기반하여 텍스트 영역에 적용된 각 수직 블록상의 픽셀 값에 기초하여 각 수직 블록에서 추정되는 문자의 확률을 산출할 수 있다.
- [115] 도 6에 따른 Subsample Layer, Feed-forward, MDLSTM 및 CTC Layer는 신경망 구조를 형성하는데 이용되는 레이어의 유형이다. 여기서, CTC(Connectionist Temporal Classification) Layer는 훈련 신경망에만 사용되는 비용 함수(cost function)의 일종일 수 있다.
- [116] 전자 장치(100)는 Subsample Layer 1, Feed-forward 1, MDLSTM 1, Subsample Layer 2, Feed-forward 2, MDLSTM 2의 순서로 Subsample Layer, Feed-forward, MDLSTM 구간을 2회 반복할 수 있다. 이러한 반복을 통하여 전자 장치(100) 상의 과부하를 방지할 수 있다.
- [117] 전자 장치(100)는 이러한 구간의 반복을 통해 원본 이미지에 포함된 각 텍스트의 확률을 산출할 수 있다. 여기서, 확률은 도 5에서 각 수직 블록에서

산출된 확률을 의미한다.

- [118] 전자 장치(100)는 토큰 패스 디코딩에 기초하여 각 수직 블록마다 가장 높은 확률로 추정되는 문자를 식별할 수 있다. 전자 장치(100)는 가장 높은 확률로 추정되는 문자를 조합하여 후보 텍스트를 제공할 수 있다.
- [119] 전자 장치(100)는 기설정된 언어 모델에 기초하여 제공된 후보 텍스트를 수정할 수 있다. 예를 들어, 각 블록마다 가장 높은 확률로 추정되는 문자에 기초하여 식별된 후보 텍스트가 "The eat"인 경우, 전자 장치(100)는 언어 모델에 기초하여 "The eat"을 "The cat"으로 수정할 수 있다.
- [120] 여기서, 토큰 패스 디코딩은 기설정된 언어 모델과 Viterbi 알고리즘을 사용하여 텍스트 라인 영역에서 인식되는 텍스트를 획득할 수 있다.
- [121] 여기서, 기설정된 언어 모델은 실제 사용자가 사용하는 언어, 예를 들어 문장, 구 등을 모델링한 데이터일 수 있다. 언어 모델을 이용하면 순차적으로 입력된 텍스트에 기초하여 입력된 텍스트의 다음 워드로 가장 적절한 텍스트를 식별할 수 있다. 본 개시의 일 실시 예에 따르면, 기설정된 언어 모델은 N-Grams 언어 모델일 수 있으나 이에 한정되는 것은 아니다. 상술한 Viterbi 알고리즘 및 N-Grams 언어 모델은 종래 기술이므로 자세한 설명은 생략하도록 한다.
- [122] 전자 장치(100)는 언어 모델에 의해 수정된 텍스트를 인식할 수 있다.
- [123] 도 7은 본 개시의 일 실시 예에 따른 필기 텍스트가 인식되는 동작을 설명하기 위한 도면이다.
- [124] 도 7a은 카메라(140)에 의해 촬영된 이미지, 디스플레이(150) 상에 손 터치 또는 펜 터치에 의해 입력된 이미지 또는 디스플레이(150) 상에 출력되는 이미지가 캡처된 이미지에 포함된 필기 텍스트 및 인쇄 텍스트 중 적어도 하나가 전자 장치(100)에 입력된 도면을 나타낸다.
- [125] 전자 장치(100)는 보정 처리를 포함하는 이미지 프로세싱을 통해 각 텍스트 라인 영역을 식별하고, 트레이닝 모델에 기초하여 각 텍스트 라인 영역에 포함된 텍스트를 인식할 수 있다.
- [126] 도 7b는 전자 장치(100)에 입력된 이미지에 포함된 텍스트가 인식된 도면을 나타낸다. 전자 장치(100)는 입력된 텍스트를 인식하여 도 7b와 같이 타이핑된 형태의 텍스트로 제공할 수 있다.
- [127] 텍스트를 인식하는 각 단계는 상술한 바 있으므로 자세한 설명은 생략하도록 한다.
- [128] 도 8은 전자 장치에서 인식된 텍스트와 관련된 기능을 식별하고 식별된 기능의 실행을 설명하기 위한 도면이다.
- [129] 도 8a에 따르면, 필기 텍스트 및 인쇄 텍스트 중 적어도 하나가 전자 장치(100)에 입력되면, 전자 장치(100)는 상술한 동작에 기초하여 텍스트를 인식할 수 있다. 전자 장치(100)는 인식된 텍스트와 관련된 기능을 식별할 수 있다.
- [130] 도 8b에 따르면, 식별된 텍스트가 "내일 오후 10시에 Mr. Brown과 만남"인

경우, 전자 장치(100)는 내일 오후 10시에 "Mr. Brown과의 만남"이라는 제목의 알람을 설정할 수 있다. 또한, 전자 장치(100)는 스케줄러 상에 내일에 해당되는 날짜 및 해당 날짜의 오후 10시에 "Mr. Brown과의 만남"이라는 스케줄을 설정할 수도 있다.

[131] 도 9는 본 발명의 일 실시 예에 따른 전자 장치를 제어하는 방법을 설명하기 위한 흐름도이다.

[132] 도 9에 따르면, 전자 장치는 필기 텍스트 및 인쇄 텍스트 중 적어도 하나를 포함하는 이미지를 획득할 수 있다(S910).

[133] 전자 장치는 이미지 프로세싱을 통해 이미지에서 각 텍스트 라인 영역을 식별할 수 있다(S920).

[134] 전자 장치는 이미지에서 텍스트 영역을 이진화하고, 이진화된 텍스트 영역에 대한 보정 처리에 의해 이미지 프로세싱을 수행할 수 있다.

[135] 여기서, 보정 처리는, 상기 텍스트 영역에 대한 기울기 제거, 높이 표준화 및 골격화(skeletonization) 중 적어도 하나를 포함할 수 있다.

[136] 전자 장치는 식별된 텍스트 라인을 포함하는 영역을 복수 개의 수직 블록으로 식별하고, 각 수직 블록의 픽셀 값에 기초하여 각 수직 블록에서 추정되는 문자의 확률을 산출할 수 있다.

[137] 구체적으로, 전자 장치는 제1 수직 블록에서 제1 문자의 존재 확률, 제1 수직 블록에서 제1 문자의 부존재 확률, 제2 수직 블록에서 제2 문자의 존재 확률 및 제2 수직 블록에서 제2 문자의 부존재 확률을 조합하여 텍스트 라인 영역에 제1 및 제2 문자 중 적어도 하나가 포함될 복수의 확률을 산출할 수 있다.

[138] 전자 장치는 트레이닝 모델에 기초하여 식별된 각 텍스트 라인 영역에 포함된 텍스트를 인식할 수 있다(S930).

[139] 전자 장치는 산출된 복수의 확률에 토큰 패스 디코딩(token passing decoding) 및 기설정된 언어 모델을 적용하여 텍스트 라인 영역에서 인식되는 텍스트를 획득할 수 있다.

[140] 전자 장치는 인식된 텍스트와 관련된 기능을 식별하고, 식별된 기능을 실행할 수 있다.

[141] 예를 들어, 식별된 텍스트가 "아침 6시에 기상"인 경우, 전자 장치는 오전 6시에 알람을 설정할 수 있다. 또는, 식별된 텍스트가 "AAA와 3월 2일 2시에 미팅"인 경우, 전자 장치는 스케줄러 상의 3월 2일에 AAA와 2시 미팅이라는 스케줄을 등록할 수도 있다.

[142] 또한, 식별된 텍스트가 "쇼핑 리스트, 계란, 우유, 사과, 빵"인 경우, 전자 장치는 이를 쇼핑 리스트로 인식하고 메모장에 이를 필기 텍스트 형식 그대로 저장하거나 인쇄 텍스트 형식으로 변경하여 저장할 수 있다.

[143] 각 단계의 상세 동작에 대해서는 상술한 바 있으므로 자세한 설명은 생략하도록 한다.

[144] 한편, 상술한 본 개시의 다양한 실시 예들에 따른 방법들 중 적어도 일부 구성은

기존 전자 장치에 설치 가능한 어플리케이션 형태로 구현될 수 있다.

[145] 또한, 상술한 본 개시의 다양한 실시 예들에 따른 방법들 중 적어도 일부 구성은, 기존 전자 장치에 대한 소프트웨어 업그레이드, 또는 하드웨어 업그레이드 만으로도 구현될 수 있다.

[146] 또한, 상술한 본 개시의 다양한 실시 예들 중 적어도 일부 구성은 전자 장치에 구비된 임베디드 서버, 또는 전자 장치의 외부 서버를 통해 수행되는 것도 가능하다.

[147] 한편, 이상에서 설명된 다양한 실시 예들은 소프트웨어(software), 하드웨어(hardware) 또는 이들의 조합을 이용하여 컴퓨터(computer) 또는 이와 유사한 장치로 읽을 수 있는 기록 매체 내에서 구현될 수 있다. 일부 경우에 있어 본 명세서에서 설명되는 실시 예들이 프로세서 자체로 구현될 수 있다. 소프트웨어적인 구현에 의하면, 본 명세서에서 설명되는 절차 및 기능과 같은 실시 예들은 별도의 소프트웨어 모듈들로 구현될 수 있다. 소프트웨어 모듈들 각각은 본 명세서에서 설명되는 하나 이상의 기능 및 작동을 수행할 수 있다.

[148] 한편, 상술한 본 개시의 다양한 실시 예들에 따른 처리 동작을 수행하기 위한 컴퓨터 명령어(computer instructions)는 비일시적 컴퓨터 판독 가능 매체(non-transitory computer-readable medium)에 저장될 수 있다. 이러한 비일시적 컴퓨터 판독 가능 매체에 저장된 컴퓨터 명령어는 프로세서에 의해 실행되었을 때 상술한 다양한 실시 예에 따른 처리 동작을 특정 기기가 수행하도록 할 수 있다.

[149] 한편, 상술한 본 개시의 다양한 실시 예들에 따른 처리 동작을 수행하기 위한 컴퓨터 명령어(computer instructions)는 비일시적 컴퓨터 판독 가능 매체(non-transitory computer-readable medium)에 저장될 수 있다. 이러한 비일시적 컴퓨터 판독 가능 매체에 저장된 컴퓨터 명령어는 프로세서에 의해 실행되었을 때 상술한 다양한 실시 예에 따른 처리 동작을 특정 기기가 수행하도록 할 수 있다.

[150] 제1 전자 장치의 프로세서에 의해 실행되는 경우, 제1 전자 장치가 동작을 수행하도록 하는 컴퓨터 명령을 저장하는 비일시적 컴퓨터 판독 가능 매체의 동작은, 제2 전자 장치의 오류 관련 정보로부터 타겟 시간 구간에 대한 제1 오류 관련 정보 및 타겟 시간 구간 및 타겟 시간 구간 외의 시간 구간을 포함하는 기준 시간 구간에 대한 제2 오류 관련 정보를 획득하고, 제1 오류 관련 정보에 기초하여 타겟 시간 구간에 대한 오류 발생 횟수 별 빈도 수 정보를 획득하고, 제2 오류 관련 정보에 기초하여 기준 시간 구간에 대한 오류 발생 횟수 별 빈도 수 정보를 획득하고, 타겟 시간 구간에 대한 오류 발생 횟수 별 빈도 수 정보 및 기준 시간 구간에 대한 오류 발생 횟수 별 빈도 수 정보를 비교하여 타겟 시간 구간에 대한 오류 발생 수준을 식별하는 동작을 포함할 수 있다.

[151] 비일시적 컴퓨터 판독 가능 매체란 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며,

기기에 의해 판독(reading)이 가능한 매체를 의미한다. 비일시적 컴퓨터 판독 가능 매체의 구체적인 예로는, CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM 등이 있을 수 있다.

- [152] 이상에서는 본 개시의 바람직한 실시 예에 대하여 도시하고 설명하였지만, 본 개시는 상술한 특정의 실시 예에 한정되지 아니하며, 청구범위에서 청구하는 본 개시의 요지를 벗어남이 없이 당해 개시에 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 개시의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

청구범위

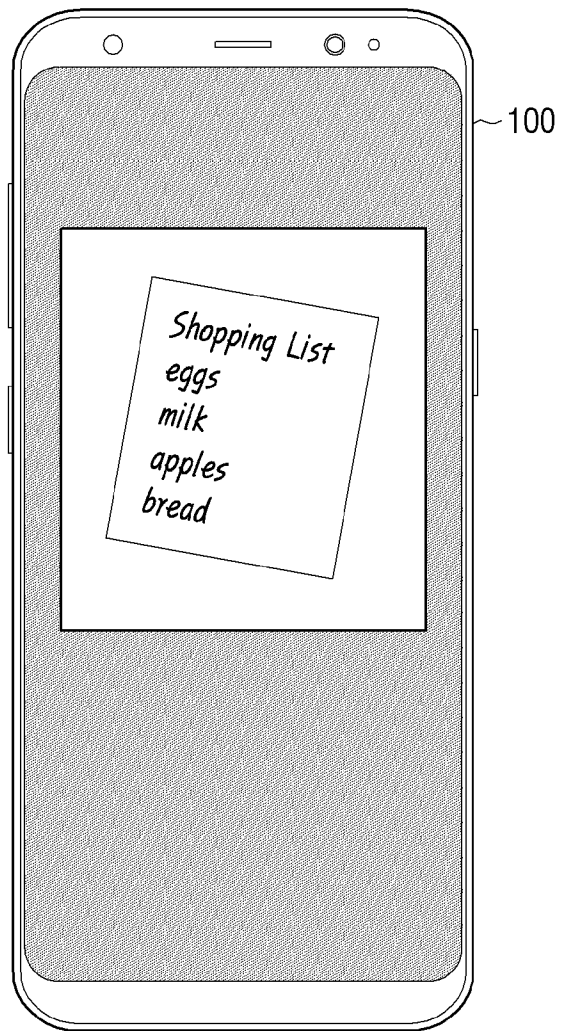
- [청구항 1] 전자 장치에 있어서,
MDLSTM (Multi-dimensional Long Short-Term Memory)의 트레이닝
모델이 저장된 저장부; 및
필기 텍스트 및 인쇄 텍스트 중 적어도 하나를 포함하는 이미지를
획득하고,
이미지 프로세싱을 통해 상기 이미지에서 각 텍스트 라인 영역을
식별하고,
상기 트레이닝 모델에 기초하여 식별된 각 텍스트 라인 영역에 포함된
텍스트를 인식하는 프로세서;를 포함하는 전자 장치.
- [청구항 2] 제1항에 있어서,
상기 프로세서는,
상기 이미지에서 텍스트 영역을 이진화하고, 상기 이진화된 텍스트
영역에 대한 보정 처리에 의해 상기 이미지 프로세싱을 수행하는, 전자
장치.
- [청구항 3] 제2항에 있어서,
상기 보정 처리는,
상기 텍스트 영역에 대한 기울기 제거, 높이 표준화 및
골격화(skeletonization) 중 적어도 하나를 포함하는, 전자 장치.
- [청구항 4] 제1항에 있어서,
상기 프로세서는,
상기 식별된 텍스트 라인을 포함하는 영역을 복수 개의 수직 블록으로
식별하고, 각 수직 블록의 픽셀 값에 기초하여 상기 각 수직 블록에서
추정되는 문자의 확률을 산출하는, 전자 장치.
- [청구항 5] 제4항에 있어서,
상기 프로세서는,
제1 수직 블록에서 제1 문자의 존재 확률, 상기 제1 수직 블록에서 상기
제1 문자의 부존재 확률, 제2 수직 블록에서 제2 문자의 존재 확률 및 상기
제2 수직 블록에서 상기 제2 문자의 부존재 확률을 조합하여 상기 텍스트
라인 영역에 제1 및 제2 문자 중 적어도 하나가 포함될 복수의 확률을
산출하는, 전자 장치.
- [청구항 6] 제5항에 있어서,
상기 프로세서는,
상기 산출된 복수의 확률에 토큰 패스 디코딩(token passing decoding) 및
기설정된 언어 모델을 적용하여 상기 텍스트 라인 영역에서 인식되는
텍스트를 획득하는, 전자 장치.
- [청구항 7] 제1항에 있어서,

- 카메라; 및
디스플레이;를 더 포함하고,
상기 필기 텍스트 및 인쇄 텍스트 중 적어도 하나를 포함하는 이미지는,
상기 카메라에 의해 획득된 이미지이거나, 상기 디스플레이에 입력된
필기 이미지인, 전자 장치.
- [청구항 8] 제1항에 있어서,
상기 프로세서는,
상기 인식된 텍스트와 관련된 기능을 식별하고, 상기 식별된 기능을
실행하는, 전자 장치.
- [청구항 9] MDLSTM (Multi-dimensional Long Short-Term Memory)의 트레이닝
모델이 저장된 전자 장치의 제어 방법에 있어서,
필기 텍스트 및 인쇄 텍스트 중 적어도 하나를 포함하는 이미지를
획득하는 단계;
이미지 프로세싱을 통해 상기 이미지에서 각 텍스트 라인 영역을
식별하는 단계; 및
상기 트레이닝 모델에 기초하여 식별된 각 텍스트 라인 영역에 포함된
텍스트를 인식하는 단계;를 포함하는 전자 장치의 제어 방법.
- [청구항 10] 제9항에 있어서,
상기 각 텍스트 라인 영역을 식별하는 단계는,
상기 이미지에서 텍스트 영역을 이진화하고, 상기 이진화된 텍스트
영역에 대한 보정 처리에 의해 상기 이미지 프로세싱을 수행하는, 제어
방법.
- [청구항 11] 제10항에 있어서,
상기 보정 처리는,
상기 텍스트 영역에 대한 기울기 제거, 높이 표준화 및
골격화(skeletonization) 중 적어도 하나를 포함하는, 제어 방법.
- [청구항 12] 제9항에 있어서,
상기 식별된 텍스트 라인을 포함하는 영역을 복수 개의 수직 블록으로
식별하고, 각 수직 블록의 픽셀 값에 기초하여 상기 각 수직 블록에서
추정되는 문자의 확률을 산출하는 단계;를 더 포함하는, 제어 방법.
- [청구항 13] 제12항에 있어서,
상기 확률을 산출하는 단계는,
제1 수직 블록에서 제1 문자의 존재 확률, 상기 제1 수직 블록에서 상기
제1 문자의 부존재 확률, 제2 수직 블록에서 제2 문자의 존재 확률 및 상기
제2 수직 블록에서 상기 제2 문자의 부존재 확률을 조합하여 상기 텍스트
라인 영역에 제1 및 제2 문자 중 적어도 하나가 포함될 복수의 확률을
산출하는, 제어 방법.
- [청구항 14] 제13항에 있어서,

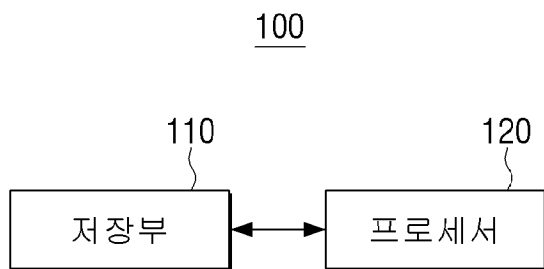
상기 텍스트를 인식하는 단계는,
상기 산출된 복수의 확률에 토큰 패스 디코딩(token passing decoding) 및
기설정된 언어 모델을 적용하여 상기 텍스트 라인 영역에서 인식되는
텍스트를 획득하는, 제어 방법.

[청구항 15] 제9항에 있어서,
상기 인식된 텍스트와 관련된 기능을 식별하고, 상기 식별된 기능을
실행하는 단계;를 더 포함하는, 제어 방법.

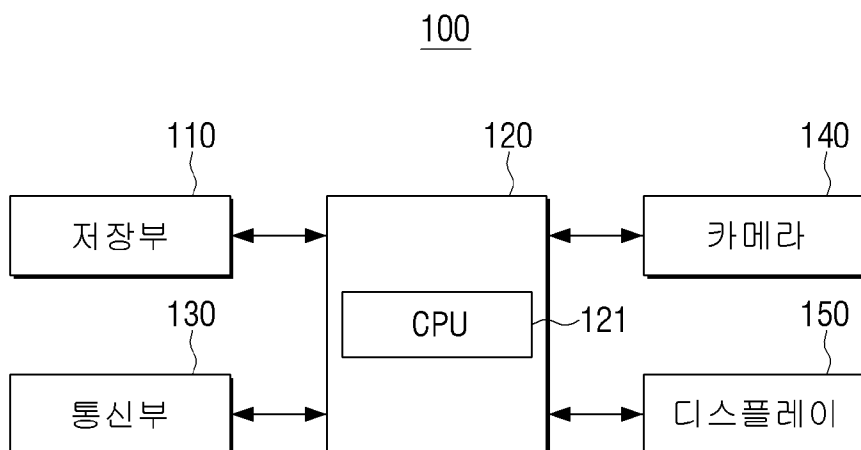
[도1]



[도2a]



[도2b]



[도3]

Make sure it doesn't go off ! ~ 310

Make sure it doesn't go off ! ~ 320

Make sure it doesn't go off ! ~ 330

Make sure it doesn't go off ! ~ 340

Make sure it doesn't go off ! ~ 350

[도4]

pleasure and pain will
give you a complete
account to the system ~ 410

~~pleasure and pain will~~
~~give you a complete~~
~~account to the system~~ ~ 420

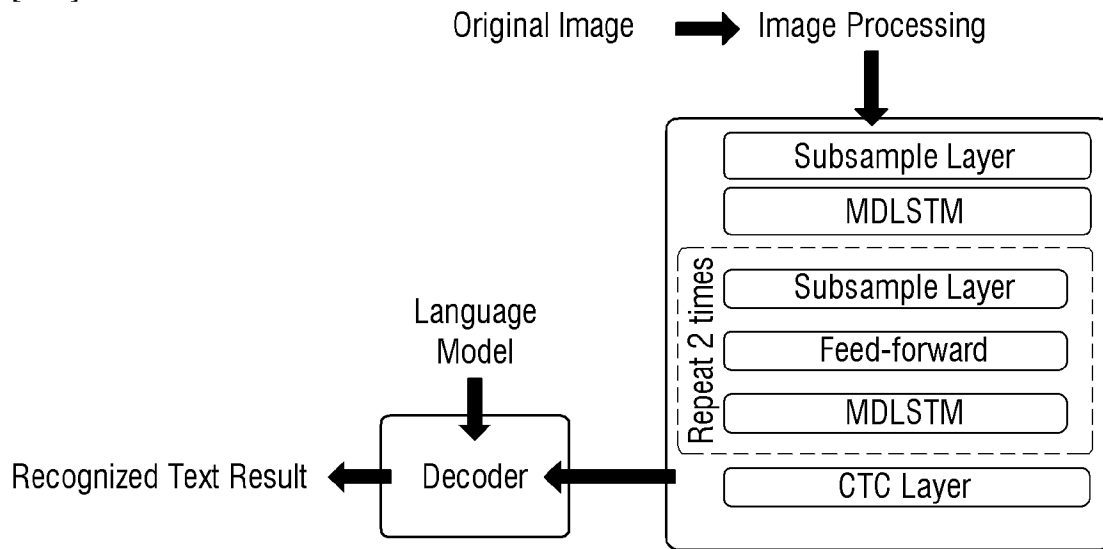
pleasure and pain will ~ 430

[도5]

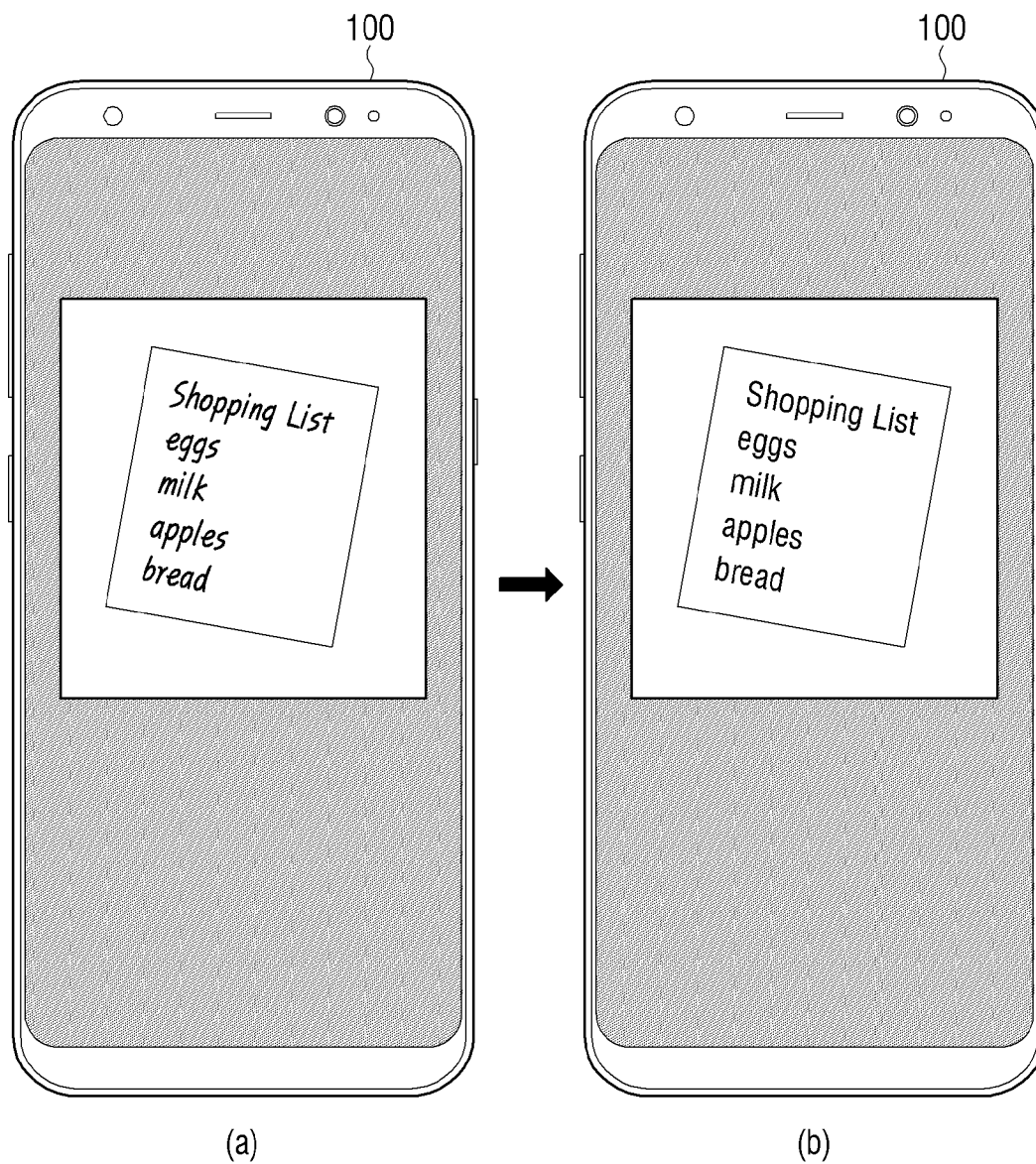
The cat ~ 510

$$\begin{array}{c}
 P(\_TH\_ \_E\_C\_AAA\_TT\_ \_) \\
 \vdots \\
 P(\_T\_H\_EE\_ \_C\_AA\_T\_ \_)
 \end{array}
 \left. \begin{array}{l} \sim 520 \\ P(TH - CAT -) \end{array} \right\}$$

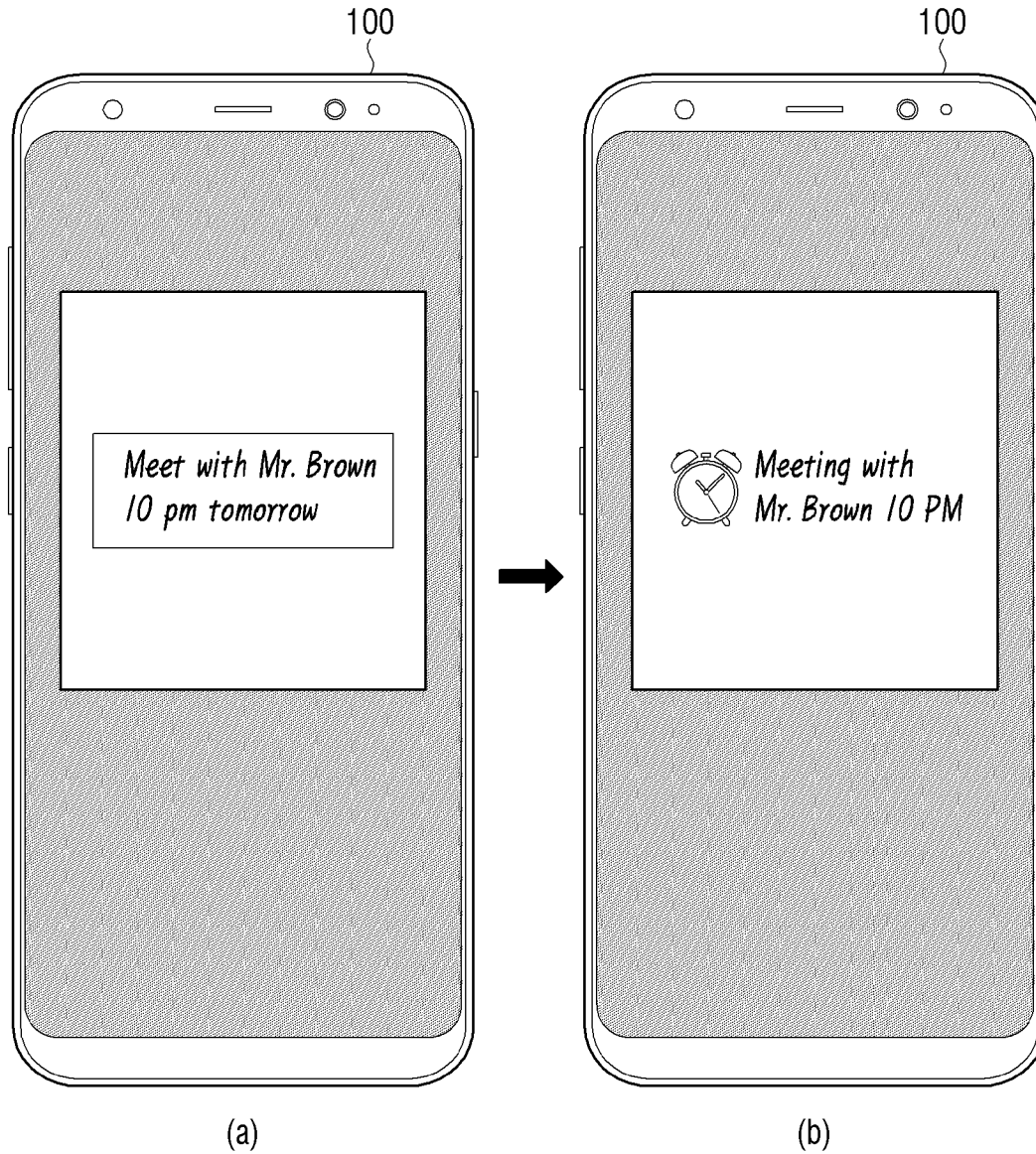
[도6]



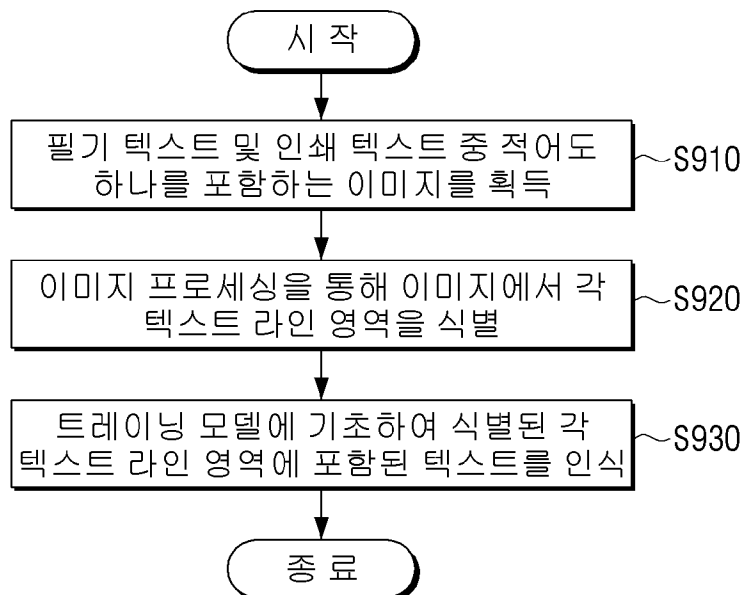
[도7]



[도8]



[도9]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2018/006080

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i, G06K 9/38(2006.01)i, G06T 5/00(2006.01)i, G06N 3/08(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)

G06K 9/00; G06F 17/20; G06F 17/27; G06F 3/01; G06K 9/18; G06K 9/46; G06K 9/62; H04N 1/387; G06K 9/38; G06T 5/00; G06N 3/08

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: text, recognition, MDLSTM(Multi-dimensional Long Short Term Memory), training, RNN(Recurrent Neural network)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2018-0005082 A1 (A2IA S.A.S.) 04 January 2018 See paragraphs [0015], [0017], [0040], [0060], [0072] and claim 1.	1,4-9,12-15
Y		2,3,10,11
Y	KR 10-2013-0082223 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 19 July 2013 See claims 1, 2.	2,3,10,11
A	WO 2016-197381 A1 (SENSETIME GROUP LIMITED) 15 December 2016 See paragraphs [0004], [0005] and claims 1, 10.	1-15
A	US 2017-0140240 A1 (SALESFORCE.COM, INC.) 18 May 2017 See paragraphs [0020]-[0024], [0048]-[0053] and claims 1, 11, 21.	1-15
A	US 2016-0378195 A1 (ORANGE) 29 December 2016 See claims 1, 5, 6.	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

10 JANUARY 2019 (10.01.2019)

Date of mailing of the international search report

10 JANUARY 2019 (10.01.2019)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex Daejeon Building 4, 189, Cheongsu-ro, Seo-gu,
Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2018/006080

Patent document cited in search report	Publication date	Patent family member	Publication date
US 2018-0005082 A1	04/01/2018	CA 2963808 A1	11/10/2017
KR 10-2013-0082223 A	19/07/2013	KR 10-1905426 B1	10/10/2018
		US 2013-0177247 A1	11/07/2013
		US 8934716 B2	13/01/2015
WO 2016-197381 A1	15/12/2016	CN 107636691 A	26/01/2018
US 2017-0140240 A1	18/05/2017	US 2016-0350653 A1	01/12/2016
		US 2017-0024645 A1	26/01/2017
		US 2017-0032280 A1	02/02/2017
		US 2018-0096219 A1	05/04/2018
US 2016-0378195 A1	29/12/2016	EP 3109797 A1	28/12/2016
		FR 3038100 A1	30/12/2016
		US 10126825 B2	13/11/2018

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G06K 9/00(2006.01)i, G06K 9/38(2006.01)i, G06T 5/00(2006.01)i, G06N 3/08(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G06K 9/00; G06F 17/20; G06F 17/27; G06F 3/01; G06K 9/18; G06K 9/46; G06K 9/62; H04N 1/387; G06K 9/38; G06T 5/00; G06N 3/08

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: 텍스트, 인식, MDLSTM(Multi-dimensional Long Short Term Memory), 트레이닝, RNN(Recurrent Neural network)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
X	US 2018-0005082 A1 (A2IA S.A.S.) 2018.01.04 단락 [0015], [0017], [0040], [0060], [0072] 및 청구항 1 참조.	1,4-9,12-15
Y		2,3,10,11
Y	KR 10-2013-0082223 A (한국전자통신연구원) 2013.07.19 청구항 1, 2 참조.	2,3,10,11
A	WO 2016-197381 A1 (SENSETIME GROUP LIMITED) 2016.12.15 단락 [0004], [0005] 및 청구항 1, 10 참조.	1-15
A	US 2017-0140240 A1 (SALESFORCE.COM, INC.) 2017.05.18 단락 [0020]-[0024], [0048]-[0053] 및 청구항 1, 11, 21 참조.	1-15
A	US 2016-0378195 A1 (ORANGE) 2016.12.29 청구항 1, 5, 6 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

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

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

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

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

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

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

“&” 동일한 대응특허문헌에 속하는 문헌

“P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

국제조사의 실제 완료일

2019년 01월 10일 (10.01.2019)

국제조사보고서 발송일

2019년 01월 10일 (10.01.2019)

ISA/KR의 명칭 및 우편주소

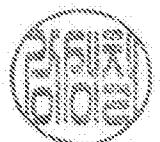
대한민국 특허청
(35208) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)

팩스 번호 +82-42-481-8578

심사관

강성철

전화번호 +82-42-481-8405



국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
US 2018-0005082 A1	2018/01/04	CA 2963808 A1	2017/10/11
KR 10-2013-0082223 A	2013/07/19	KR 10-1905426 B1	2018/10/10
		US 2013-0177247 A1	2013/07/11
		US 8934716 B2	2015/01/13
WO 2016-197381 A1	2016/12/15	CN 107636691 A	2018/01/26
US 2017-0140240 A1	2017/05/18	US 2016-0350653 A1	2016/12/01
		US 2017-0024645 A1	2017/01/26
		US 2017-0032280 A1	2017/02/02
		US 2018-0096219 A1	2018/04/05
US 2016-0378195 A1	2016/12/29	EP 3109797 A1	2016/12/28
		FR 3038100 A1	2016/12/30
		US 10126825 B2	2018/11/13