

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2020년 9월 3일 (03.09.2020)



(10) 국제공개번호
WO 2020/175806 A1

(51) 국제특허분류:

G06K 9/32 (2006.01)

G06K 9/00 (2006.01)

G06T 7/11 (2017.01)

G06T 11/60 (2006.01)

G06N 20/00 (2019.01)

(21) 국제출원번호:

PCT/KR2020/001333

(22) 국제출원일:

2020년 1월 29일 (29.01.2020)

(25) 출원언어:

한국어

(26) 공개언어:

한국어

(30) 우선권정보:

10-2019-0022102 2019년 2월 25일 (25.02.2019) KR

(71) 출원인: 네이버 주식회사 (NAVER CORPORATION)
[KR/KR]: 13561 경기도 성남시 분당구 불정로 6, Gyeonggi-do (KR).

(72) 발명자: 백영민 (BAEK, Young Min); 13561 경기도 성남시 분당구 불정로 6, Gyeonggi-do (KR). 이활석 (LEE, Hwal Suk); 13561 경기도 성남시 분당구 불정로 6, Gyeonggi-do (KR). 신승 (SHIN, Seung); 13561 경기도 성남시 분당구 불정로 6, Gyeonggi-do (KR). 이영무 (LEE, Young Moo); 01709 서울시 노원구 덕릉로71길 30, 102동 1401호, Seoul (KR).

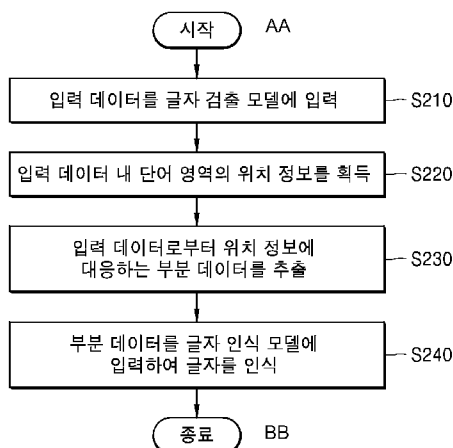
(74) 대리인: 리앤목 특허법인 (Y.P.LEE, MOCK & PARTNERS); 06292 서울시 강남구 언주로 30길 13 대림아크로텔 12층, Seoul (KR).

(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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 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 (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: DEVICE FOR RECOGNIZING CHARACTERS AND METHOD FOR RECOGNIZING CHARACTERS BY MEANS OF SAME

(54) 발명의 명칭: 글자 인식 장치 및 이에 의한 글자 인식 방법



(57) Abstract: Disclosed is a method for recognizing characters in input data by means of a device for recognizing characters. A method for recognizing characters, according to one embodiment, comprises the steps of: inputting input data into a character detection model; obtaining location information of a word area in the input data on the basis of output data outputted by the character detection model; extracting, from the input data, partial data corresponding to the obtained location information; and inputting the partial data into a character recognition model and recognizing characters in the partial data.

(57) 요약서: 글자 인식 장치에 의한 입력 데이터 내 글자를 인식하는 방법에 있어서, 입력 데이터를 글자 검출 모델에 입력하는 단계; 글자 검출 모델에서 출력되는 출력 데이터에 기초하여 입력 데이터 내 단어 영역의 위치 정보를 획득하는 단계; 획득한 위치 정보에 대응하는 부분 데이터를 입력 데이터로부터 추출하는 단계; 및 부분 데이터를 글자 인식 모델에 입력하여 부분 데이터 내 글자를 인식하는 단계를 포함하는 것을 특징으로 하는 일 실시예에 따른 글자 인식 방법이 개시된다.

S210 ... Input input data into character detection model
S220 ... Obtain location information of word area in input data
S230 ... Extract, from input data, partial data corresponding to location information
S240 ... Input partial data into character recognition model, and recognize characters
AA ... Start
BB ... End

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

명세서

발명의 명칭: 글자 인식 장치 및 이에 의한 글자 인식 방법

기술분야

- [1] 본 개시는 데이터 처리 분야에 관한 것이다. 보다 구체적으로, 본 개시는 이미지 등의 데이터에서 글자를 인식하는 글자 인식 장치 및 방법에 관한 것이다.

배경기술

- [2] 핀테크(fintech) 기술의 발전에 따라, 휴대폰 등에 카드 정보를 저장하여 놓고 간편하게 결제할 수 있게 하는 서비스가 제공되고 있다. 신용카드, 체크카드 등의 실물 카드 이미지에서 카드 번호 및 유효 기간 등의 정보를 인식 및 저장하는 기술은 간편 결제 서비스를 위한 핵심이 되는 기술 중 하나이다.
- [3] 그러나, 카드 이미지에서 글자를 인식함에 있어, 카드 내 양각으로 인쇄된 글자가 다수 존재하고, 카드 배경이 다양하므로, 카드 번호 및 유효 기간을 정확하게 인식하는 것에 기술 장벽이 존재한다.

발명의 상세한 설명

기술적 과제

- [4] 일 실시예에 따른 글자 인식 장치 및 이에 의한 글자 인식 방법은 이미지 등의 데이터에서 글자를 정확하고 신속하게 인식하는 것을 기술적 과제로 한다.
- [5] 또한, 일 실시예에 따른 글자 인식 장치 및 이에 의한 글자 인식 방법은 실물 카드의 이미지 내에서 글자를 정확하게 인식하여 핀테크 산업의 발전에 기여하는 것을 기술적 과제로 한다.

과제 해결 수단

- [6] 일 실시예에 따른 글자 인식 방법은, 입력 데이터를 글자 검출 모델에 입력하는 단계; 상기 글자 검출 모델에서 출력되는 출력 데이터에 기초하여 상기 입력 데이터 내 단어 영역의 위치 정보를 획득하는 단계; 상기 획득한 위치 정보에 대응하는 부분 데이터를 상기 입력 데이터로부터 추출하는 단계; 및 상기 부분 데이터를 글자 인식 모델에 입력하여 상기 부분 데이터 내 글자를 인식하는 단계를 포함할 수 있다.

발명의 효과

- [7] 일 실시예에 따른 글자 인식 장치 및 이에 의한 글자 인식 방법은 이미지 등의 데이터에서 글자를 정확하고 신속하게 인식할 수 있다.
- [8] 또한, 일 실시예에 따른 글자 인식 장치 및 이에 의한 글자 인식 방법은 실물 카드의 이미지 내에서 글자를 정확하게 인식하여 핀테크 산업의 발전에 기여할 수 있다.
- [9] 다만, 일 실시예에 따른 글자 인식 장치 및 이에 의한 글자 인식 방법이 달성할 수 있는 효과는 이상에서 언급한 것들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 개시가 속하는 기술분야에서 통상의

지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [10] 본 명세서에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 간단한 설명이 제공된다.
- [11] 도 1은 일 실시예에 따른 글자 인식 장치를 도시하는 도면이다.
- [12] 도 2는 일 실시예에 따른 글자 인식 방법을 설명하기 위한 순서도이다.
- [13] 도 3은 일 실시예에 따른 글자 인식 장치를 통해 글자가 인식되는 과정을 설명하기 위한 도면이다.
- [14] 도 4는 글자 검출 모델에 의해 출력되는 출력 데이터를 도시하는 예시적인 도면이다.
- [15] 도 5는 글자 검출 모델에서 출력된 출력 데이터에 기초하여 입력 데이터 내 단어 영역의 위치 정보를 획득하는 방법을 설명하기 위한 도면이다.
- [16] 도 6은 도 5에 도시된 이진화 과정 및 병합 과정을 설명하기 위한 도면이다.
- [17] 도 7은 도 5에 도시된 단어 박스 결정 과정을 설명하기 위한 도면이다.
- [18] 도 8은 일 실시예에 따른 특징 추출 모델의 구조를 설명하기 위한 도면이다.
- [19] 도 9는 일 실시예에 따른 글자 인식 모델의 구조를 설명하기 위한 도면이다.
- [20] 도 10은 일 실시예에 따른 글자 검출 모델의 훈련 방법을 설명하기 위한 순서도이다.
- [21] 도 11은 제 1 GT 스코어 맵을 생성하는 방법을 설명하기 위한 도면이다.
- [22] 도 12는 제 2 GT 스코어 맵을 생성하는 방법을 설명하기 위한 도면이다.
- [23] 도 13은 서로 인접한 글자 박스들 사이에서 연결 박스를 결정하는 방법을 설명하기 위한 도면이다.
- [24] 도 14는 일 실시예에 따른 글자 인식 장치의 구성을 도시하는 블록도이다.
- [25] 도 15는 일 실시예에 따른 글자 인식 장치가 적용될 수 있는 서버 장치 및 클라이언트 장치를 도시하는 도면이다.

발명의 실시를 위한 최선의 형태

- [26] 일 실시예에 따른 글자 인식 방법은, 입력 데이터를 글자 검출 모델에 입력하는 단계; 상기 글자 검출 모델에서 출력되는 출력 데이터에 기초하여 상기 입력 데이터 내 단어 영역의 위치 정보를 획득하는 단계; 상기 획득한 위치 정보에 대응하는 부분 데이터를 상기 입력 데이터로부터 추출하는 단계; 및 상기 부분 데이터를 글자 인식 모델에 입력하여 상기 부분 데이터 내 글자를 인식하는 단계를 포함할 수 있다.
- [27] 다른 실시예에 따른 글자 인식 장치는, 프로세서; 및 적어도 하나의 인스트럭션을 저장하는 메모리를 포함하되, 상기 프로세서는 상기 적어도 하나의 인스트럭션에 따라, 입력 데이터를 글자 검출 모델에 입력하고, 상기 글자 검출 모델에서 출력되는 출력 데이터에 기초하여 상기 입력 데이터 내 단어 영역의 위치 정보를 획득하고, 상기 획득한 위치 정보에 대응하는 부분 데이터를

상기 입력 데이터로부터 추출하고, 상기 부분 데이터를 글자 인식 모델에 입력하여 상기 부분 데이터 내 글자를 인식할 수 있다.

발명의 실시를 위한 형태

- [28] 본 개시는 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고, 이를 상세한 설명을 통해 설명하고자 한다. 그러나, 이는 본 개시를 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 개시의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [29] 실시예를 설명함에 있어서, 관련된 공지 기술에 대한 구체적인 설명이 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다. 또한, 실시예의 설명 과정에서 이용되는 숫자(예를 들어, 제 1, 제 2 등)는 하나의 구성요소를 다른 구성요소와 구분하기 위한 식별기호에 불과하다.
- [30] 또한, 본 명세서에서 일 구성요소가 다른 구성요소와 "연결된다" 거나 "접속된다" 등으로 언급된 때에는, 상기 일 구성요소가 상기 다른 구성요소와 직접 연결되거나 또는 직접 접속될 수도 있지만, 특별히 반대되는 기재가 존재하지 않는 이상, 중간에 또 다른 구성요소를 매개하여 연결되거나 또는 접속될 수도 있다고 이해되어야 할 것이다.
- [31] 또한, 본 명세서에서 '~부(유닛)', '모듈' 등으로 표현되는 구성요소는 2개 이상의 구성요소가 하나의 구성요소로 합쳐지거나 또는 하나의 구성요소가 보다 세분화된 기능별로 2개 이상으로 분화될 수도 있다. 또한, 이하에서 설명할 구성요소 각각은 자신이 담당하는 주기능 이외에도 다른 구성요소가 담당하는 기능 중 일부 또는 전부의 기능을 추가적으로 수행할 수도 있으며, 구성요소 각각이 담당하는 주기능 중 일부 기능이 다른 구성요소에 의해 전담되어 수행될 수도 있음은 물론이다.
- [32] 또한, 본 명세서에서 '글자'는 단어나 문장을 구성하는 기본 문자 단위를 의미할 수 있다. 예를 들어, 영어의 경우에는 각각의 알파벳이 글자에 해당할 수 있고, 숫자의 경우에는 '0' 내지 '9'의 숫자 각각이 글자에 해당할 수 있고, 한국어의 경우에는 자음과 모음이 결합된 문자(예를 들어, '가'), 자음, 모음 및 자음이 결합된 문자(예를 들어, '강'), 단독으로 기재된 자음(예를 들어, 'ㄱ'), 단독으로 기재된 모음(예를 들어, 'ㅏ')이 글자에 해당할 수 있다. 또한, 글자는 기호(예를 들어, '/', '.' 등)를 포함할 수도 있다.
- [33] 또한, 본 명세서에서 '단어'는 적어도 하나의 글자를 포함하는 문자 단위를 의미할 수 있다. '단어'를 구성하는 글자들은 서로 간에 소정 간격 이상 이격되어 있지 않을 수 있다. '단어'는 하나의 글자로 이루어질 수도 있다. 예를 들어, 영어의 부정사 'a'는 하나의 글자로 이루어졌지만 주변 글자와 소정 거리 이상 이격되어 있는 경우 '단어'에 해당할 수 있다.
- [34] 또한, 본 명세서에서 '글자 그룹'은 후술하는 어느 하나의 부분 데이터에서

인식된 적어도 하나의 글자들을 의미할 수 있다.

- [35] 이하, 본 개시의 기술적 사상에 의한 실시예들을 차례로 상세히 설명한다.
- [36] 도 1은 일 실시예에 따른 글자 인식 장치(100)를 도시하는 도면이다.
- [37] 일 실시예에 따른 글자 인식 장치(100)는 입력 데이터(10)를 획득하고, 입력 데이터(10) 내에서 글자(50)를 인식한다. 입력 데이터(10)는 체크카드, 신용카드 등의 실물 카드를 촬영한 이미지를 포함할 수 있으며, 또는 후술하는 바와 같이, 실물 카드 등을 촬영한 이미지에 기초하여 특징 추출 모델(800)에서 출력된 특징 맵(feature map)을 포함할 수도 있다.
- [38] 글자 인식 장치(100)는 입력 데이터(10)에서 카드 번호, 유효 기간 등의 카드 정보를 인식 및 저장할 수 있다. 글자 인식 장치(100)에 의해 인식 및 저장된 카드 정보는 물건 등의 구매를 위해 대금을 지불하는데 이용될 수 있다.
- [39] 이하에서는, 도 2 및 도 3을 참조하여 글자 인식 장치(100)의 동작에 대해 설명한다.
- [40] 도 2는 일 실시예에 따른 글자 인식 방법을 설명하기 위한 순서도이고, 도 3은 일 실시예에 따른 글자 인식 장치(100)를 통해 글자가 인식되는 과정을 설명하기 위한 도면이다.
- [41] S210 단계에서, 글자 인식 장치(100)는 입력 데이터(10)를 글자 검출 모델(410)에 입력한다. 글자 인식 장치(100)는 글자 검출 모델(410)을 미리 저장할 수 있다. 글자 검출 모델(410)은 학습용 데이터에 기초하여 훈련될 수 있다.
- [42] S220 단계에서, 글자 인식 장치(100)는 글자 검출 모델(410)에서 출력되는 출력 데이터(30)에 기초하여 입력 데이터(10) 내 단어 영역의 위치 정보를 획득한다.
- [43] 글자 검출 모델(410)에서 출력되는 출력 데이터(30)는, 입력 데이터(10) 내 글자가 존재할 것으로 예상되는 지점의 위치를 나타낸다. 글자 인식 장치(100)는 출력 데이터(30)에 기초하여 입력 데이터(10) 내 적어도 하나의 글자를 포함하는 단어 영역의 위치 정보를 획득한다.
- [44] S230 단계에서, 글자 인식 장치(100)는 단어 영역의 위치 정보에 대응하는 부분 데이터(40)를 입력 데이터(10)로부터 추출한다. 일 실시예에서, 단어 영역의 위치 정보가 복수 개로 획득된 경우, 각 위치 정보에 대응하는 복수의 부분 데이터들(40)이 입력 데이터(10)로부터 추출될 수 있다.
- [45] S240 단계에서, 글자 인식 장치(100)는 부분 데이터(40)를 글자 인식 모델(420)에 입력하여 부분 데이터(40)에 포함된 글자(50)를 인식한다. 부분 데이터(40)가 복수 개인 경우, 글자 인식 장치(100)는 복수의 부분 데이터(40) 각각을 글자 인식 모델(420)에 입력하여 복수의 부분 데이터(40) 각각에 포함된 글자(50)들을 인식할 수 있다.
- [46] 일 실시예에서, 글자 인식 장치(100)는 글자 검출 모델(410)의 출력 데이터(30)를 부분 데이터(40)와 함께 글자 인식 모델(420)로 입력할 수도 있다. 글자 검출 모델(410)의 출력 데이터(30)는 입력 데이터(10) 내 개별 글자의 위치

정보를 포함할 수 있으므로, 글자 인식 모델(420)의 글자 인식의 정확도가 보다 향상될 수 있다.

- [47] 글자 인식 장치(100)는 인식된 글자를 저장하거나, 네트워크를 통해 외부 장치로 전송할 수 있다.
- [48] 도 4는 글자 검출 모델(410)에 의해 출력되는 출력 데이터(30)의 일 예를 도시하는 예시적인 도면이다.
- [49] 출력 데이터(30)는 입력 데이터(10) 내 글자가 존재할 확률을 입력 데이터(10)에 대응되는 데이터 공간(예를 들어, 이미지 공간)상에 나타내는 제 1 스코어 맵(31), 및 입력 데이터(10) 내 글자들 사이의 연결성(connectivity)을 입력 데이터(10)에 대응되는 데이터 공간상에 나타내는 제 2 스코어 맵(33)을 포함할 수 있다.
- [50] 제 1 스코어 맵(31) 내 각 위치에 저장된 값(예를 들어, 픽셀 값)은 해당 위치에 대응하는 입력 데이터(10)에 글자가 존재할 확률을 나타낼 수 있다. 또한, 제 2 스코어 맵(33) 내 각 위치에 저장된 값(예를 들어, 픽셀 값)은 해당 위치에 대응하는 입력 데이터(10) 내에서 복수의 글자들이 서로 인접할 확률을 나타낼 수 있다.
- [51] 위치 대응 관계에 대한 계산을 용이하게 하기 위하여 제 1 스코어 맵(31) 및 제 2 스코어 맵(33)의 크기는 입력 데이터(10)와 동일하게 할 수 있다.
- [52] 후술하는 바와 같이, 글자 검출 모델(410)은 학습용 데이터에 대응하여 생성된 제 1 GT(ground truth) 스코어 맵 및 제 2 GT 스코어 맵과 유사한 제 1 스코어 맵(31) 및 제 2 스코어 맵(33)이 출력되도록 훈련될 수 있다.
- [53] 글자 인식 장치(100)는 제 1 스코어 맵(31) 및 제 2 스코어 맵(33)에 기초하여 입력 데이터(10) 내 단어 영역의 위치 정보를 결정할 수 있는데, 이에 대해서는 도 5 내지 도 7을 참조하여 설명한다.
- [54] 도 5는 글자 검출 모델(410)에서 출력된 출력 데이터(30)에 기초하여 입력 데이터(10) 내 단어 영역의 위치 정보를 획득하는 방법을 설명하기 위한 도면이고, 도 6은 도 5에 도시된 이진화 과정 및 병합 과정을 설명하기 위한 도면이고, 도 7은 도 5에 도시된 단어 박스 결정 과정을 설명하기 위한 도면이다.
- [55] S510 단계 및 S520 단계에서, 글자 인식 장치(100)는 제 1 스코어 맵(31) 내 데이터 값들을 임계 값과 비교하여 제 1 스코어 맵(31)을 이진화(binrarization)하고, 제 2 스코어 맵(33) 내 데이터 값들을 임계 값과 비교하여 제 2 스코어 맵(33)을 이진화한다. 일 예에서, 글자 인식 장치(100)는 제 1 스코어 맵(31) 및 제 2 스코어 맵(33) 내 데이터 값들 중 임계 값 이상의 데이터 값들을 제 1 값으로 변경하고, 임계 값 미만의 데이터 값들을 제 2 값으로 변경할 수 있다.
- [56] 도 6에 도시된 바와 같이, 제 1 스코어 맵(31) 및 제 2 스코어 맵(33)에서 임계 값 이상의 값을 갖는 데이터들은 이진화된 제 1 스코어 맵(601) 및 이진화된 제 2 스코어 맵(603)에서 제 1 값을 갖도록 변경되고, 제 1 스코어 맵(31) 및 제 2 스코어 맵(33)에서 임계 값 미만의 값을 갖는 데이터들은 이진화된 제 1 스코어

- 맵(601) 및 이진화된 제 2 스코어 맵(603)에서 제 2 값을 갖도록 변경될 수 있다.
- [57] 제 1 스코어 맵(31)의 이진화를 위한 임계 값과 제 2 스코어 맵(33)의 이진화를 위한 임계 값은 서로 동일하거나 서로 상이할 수 있다.
- [58] S530 단계에서, 글자 인식 장치(100)는 이진화된 제 1 스코어 맵(601)와 이진화된 제 2 스코어 맵(603)을 병합(merge)한다. 예를 들어, 글자 인식 장치(100)는 이진화된 제 1 스코어 맵(601)과 이진화된 제 2 스코어 맵(603) 내 데이터 값들을 더하거나, or 연산을 하여 병합 맵(605)을 생성할 수 있다. 예를 들어 도 6에 도시된 바와 같이, 이진화된 제 1 스코어 맵(601) 및 이진화된 제 2 스코어 맵(603) 내 제 1 값을 갖는 데이터들이 병합 맵(605)에 함께 포함될 수 있다. 이러한 방법으로 병합 맵(605)은 입력 데이터(10) 내 글자가 존재할 가능성이 높은 영역(606)들과 그렇지 않은 영역들로 구분될 수 있다.
- [59] S540 단계에서, 글자 인식 장치(100)는 병합 맵(605)을 이용하여 글자가 포함된 영역을 나타내는 단어 박스(610)를 결정하게 된다.
- [60] 예를 들면, 병합 맵(605) 내에서 동일한(내지는 동일한 범위의) 값을 가지고 서로 인접하게 연결된 영역(606)들의 적어도 일부를 단어 영역으로 결정하고, 결정된 단어 영역을 포함하는 단어 박스(610)를 결정할 수 있다. 일 실시예에서, 글자 인식 장치(100)는 병합 맵(605) 내 단어 영역들의 구분을 위해 단어 영역 각각에 대해 레이블링(labeling)을 할 수도 있다.
- [61] 일 실시예에서, 글자 인식 장치(100)는 병합 맵(605)을 이용하여 인식된 영역(606) 각각이 실제 단어를 포함하는지를 검증하기 위해, 추가 확인을 할 수 있다. 구체적으로 예를 들어, 병합 맵(605) 내에서 동일한(내지는 동일한 범위의) 값을 가지고 서로 인접하게 연결된 영역(606)을 단어 후보 영역으로 두고, 단어 후보 영역 내의 각 데이터에 대응하는 제 1 스코어 맵(601)의 값들 중에 정해진 임계치보다 높은 것이 하나 이상 존재하면 해당 단어 후보 영역을 단어 영역으로 결정할 수 있다. 즉, 각 단어 후보 영역에 대응하는 제 1 스코어 맵(601)의 값들 중 가장 큰 값과 임계값을 비교하여 각 단어 후보 영역이 단어 영역에 해당하는지 여부를 검증할 수 있다.
- [62] 이렇게 하면 글자와 유사한 배경이 있어 단어 후보 영역으로 결정된 경우들을 필터링할 수 있게 된다.
- [63] 일 실시예에서, 글자 인식 장치(100)는 단어 영역에 해당하는 것으로 검증된 데이터들의 영역을 포함하는 최소 크기의 단어 박스(610)를 결정할 수 있다.
- [64] 글자 인식 장치(100)는 결정된 단어 박스(610)의 위치 정보(예를 들어, 입력 데이터(10) 또는 병합 맵(605) 상에서의 단어 박스(610)의 모서리들의 위치 값)를 단어 영역의 위치 정보로 결정할 수 있다.
- [65] 단어 영역의 위치 정보가 결정되면, 글자 인식 장치(100)는 해당 위치 정보에 대응하는 부분 데이터(40)를 입력 데이터(10)로부터 추출하고, 추출된 부분 데이터(40)를 글자 인식 모델(420)에 입력하여 부분 데이터(40) 내 글자를 인식할 수 있다.

- [66] 앞서 설명한 바와 같이, 글자 검출 모델(410)로 입력되는 입력 데이터(10)는 원본 이미지에 기초하여 특징 검출 모델(800)에서 출력되는 특징 맵(feature map)을 포함할 수 있다. 도 8은 특징 검출 모델(800)의 구조를 설명하기 위한 도면이다.
- [67] 원본 이미지(20)는 특징 검출 모델(800)로 입력될 수 있다. 여기서, 원본 이미지(20)는 특징 검출 모델(800)로 입력되는 이미지를 의미하는 것이며, 최초 카드 등을 촬영한 이미지를 복사한 이미지 또는 변형한 이미지가 아님을 의미하는 것은 아니다.
- [68] 원본 이미지(20)는 제 1 컨볼루션층(805), 제 2 컨볼루션층(810), 제 3 컨볼루션층(815), 제 4 컨볼루션층(820), 제 5 컨볼루션층(825) 및 제 6 컨볼루션층(830)에서 컨볼루션 처리가 된다. 제 6 컨볼루션층(830)의 출력과 제 5 컨볼루션층(825)의 출력이 연결(concatenation) 연산되어 제 1 업 컨볼루션층(835)으로 입력되고, 제 1 업 컨볼루션층(835)으로 입력된 값들은 컨볼루션 처리(836), 배치 정규화(normalization)(837), 컨볼루션 처리(838) 및 배치 정규화(839)를 통해 제 1 업 샘플링층(840)으로 입력된다. 제 1 업 샘플링층(840)의 출력은 제 4 컨볼루션층(820)의 출력과 연결 연산되어 제 2 업 컨볼루션층(845) 및 제 2 업 샘플링층(850)에서 처리된다. 제 2 업 샘플링층(850)의 출력은 제 3 컨볼루션층(815)의 출력과 연결 연산되어 제 3 업 컨볼루션층(855)과 제 3 업 샘플링층(860)에서 처리되고, 처리 결과는 제 2 컨볼루션층(810)의 출력과 연결 연산되어 제 4 업 컨볼루션층(865)에 입력된다. 그리고, 제 4 업 컨볼루션층(865)으로부터 출력된 결과를 입력 데이터(10)로 사용할 수 있다.
- [69] 일 실시예에서, 입력 데이터(10)의 가로 크기 및 세로 크기는 원본 이미지(20)의 가로 크기 및 세로 크기의 1/2일 수 있으나, 이에 한정되는 것은 아니다.
- [70] 도 8에 도시된 특징 검출 모델(800)의 구조는 하나의 예시일 뿐이며, 컨볼루션층, 업 컨볼루션층, 업 샘플링층의 개수 및 처리 순서는 다양하게 변형될 수 있다.
- [71] 도 9는 일 실시예에 따른 글자 인식 모델(420)의 구조를 설명하기 위한 도면이다.
- [72] 글자 인식 모델(420)은 입력 데이터(10)로부터 추출된 부분 데이터(40)를 입력받고, 부분 데이터(40) 내 글자(50)들을 인식한다. 글자 인식 모델(420)은 컨볼루션 네트워크(421), RNN(recurrent neural network)(423) 및 디코더(decoder)(425)를 포함할 수 있다.
- [73] 컨볼루션 네트워크(421)는 적어도 하나의 컨볼루션 층을 포함하고, 부분 데이터(40)를 컨볼루션 처리하여 특징 맵을 추출한다. 일 예시에서, 컨볼루션 네트워크(421)는 공지된 VGG, ResNet 등을 포함할 수 있지만, 일 실시예에서 글자 인식 모델(420)은 원본 이미지(20)의 특징 맵(즉, 입력 데이터)으로부터 추출된 부분 데이터(40)를 입력 받을 수 있으므로, 필요로 하는 컨볼루션 층의

개수는 적을 수 있다.

- [74] RNN(423)은 부분 데이터(40)에 대응하는 특징 맵으로부터 특징 벡터의 시퀀스를 추출한다. RNN(423)은 bi-LSTM (bidirectional Long-short-term memory)을 통해 연속되는 특징 벡터들의 컨텍스트(context) 관계를 파악할 수 있다.
- [75] 디코더(425)는 특징 벡터들의 시퀀스 정보에서 글자를 추출한다. 디코더(425)는 어텐션(attention) 단계 및 생성(generation) 단계를 수행할 수 있는데, 어텐션 단계에서, 디코더(425)는 어떤 시퀀스에서 정보를 뽑을지 여부를 나타내는 웨이트를 계산하고, 생성 단계에서 가중치를 시퀀스에 적용하고, LSTM(Long-short-term memory)을 통해 개별 글자를 추출할 수 있다.
- [76] 한편, 일 실시예에서, 글자 인식 장치(100)는 여러 부분 데이터(40)들 각각에서 인식된 글자 그룹들을 소정 기준에 따라 분류할 수 있다. 일 예에서, 글자 인식 장치(100)는 어느 부분 데이터(40)에서 인식된 글자 그룹에 소정의 기호(예를 들어, /)가 포함되어 있으면, 해당 글자 그룹을 제 1 종류의 정보로 결정할 수 있다. 카드 내 유효 기간에는 년도와 월을 구분하기 위한 소정의 기호가 포함되어 있는 것이 일반적이므로, 글자 인식 장치(100)는 어느 부분 데이터(40)에서 인식된 글자 그룹에 소정 기호가 포함되어 있으면 해당 글자 그룹을 유효 기간 정보로 결정할 수 있는 것이다.
- [77] 만약, 소정의 기호가 포함되어 있는 글자 그룹의 개수가 복수 개인 경우, 글자 인식 장치(100)는 년도에 해당하는 숫자(예를 들어, 기호를 기준으로 우측에 위치하는 숫자)가 큰 글자 그룹을 유효 기간 정보로 결정할 수 있다. 카드에 유효 기간과 발급 일자가 포함되어 있는 경우, 유효 기간에 포함된 년도가 발급 일자에 포함된 년도보다 클 것이므로, 글자 인식 장치(100)는 년도에 해당하는 숫자가 큰 글자 그룹을 유효 기간 정보로 결정할 수 있는 것이다.
- [78] 또한, 일 실시예에서, 글자 인식 장치(100)는 복수의 부분 데이터(40) 각각에서 인식된 글자 그룹들 중 소정의 기호를 포함하고 있지 않은 글자 그룹들을 제 2 종류의 정보로 결정할 수 있다. 제 2 종류의 정보는 예를 들어, 카드 번호 정보를 포함할 수 있다.
- [79] 또한, 일 실시예에서, 글자 인식 장치(100)는 복수의 부분 데이터(40)들 각각에서 인식된 글자 그룹들을, 입력 데이터(10) 내에서의 복수의 부분 데이터(40)들의 위치에 따라 정렬할 수 있다. 일 예로, 글자 인식 장치(100)는 입력 데이터(10) 내 좌상단을 기준으로 Z 스캔 방식으로 글자 그룹들을 정렬할 수 있다.
- [80] 글자 인식 장치(100)는 정렬된 글자 그룹들 중 연속으로 정렬된 소정 개수의 글자 그룹에 포함된 글자의 개수에 기초하여 글자의 재인식이 필요한지 여부를 결정할 수 있다. 일 예로, 글자 인식 장치(100)는 정렬된 글자 그룹들 중 소정 개수의 숫자를 각각 포함하면서 연속으로 정렬된 소정 개수의 글자 그룹이 존재하는 경우, 글자 인식이 정확히 수행되어 글자의 재인식이 필요하지 않은

- 것으로 결정할 수 있다. 일반적으로 카드 번호는 16개의 숫자들을 포함하되, 4개의 숫자끼리 하나의 글자 그룹을 이룬다는 면에서, 글자 인식 장치(100)는 정렬된 글자 그룹들 중 4개의 숫자를 포함하는 4개의 글자 그룹이 연속으로 정렬되어 있는 경우, 글자의 재인식이 필요하지 않은 것으로 결정할 수 있다.
- [81] 또한, 일 실시예에서, 글자 인식 장치(100)는 여러 부분 데이터(40)들에서 인식된 글자 그룹들에 소정의 기호가 존재하지 않으면, 글자의 재인식이 필요한 것으로 결정할 수 있다.
- [82] 글자의 재인식이 필요한 경우, 글자 인식 장치(100)는 이미지의 재촬영이 필요하다는 정보를 스피커, 모니터 등을 통해 출력하거나 네트워크를 통해 외부 장치로 알릴 수 있다. 일 실시예에서, 글자 인식 장치(100)가 카메라의 프리뷰 이미지로부터 글자를 인식하는 중에, 글자의 재인식이 필요한 것으로 결정된 경우, 카메라를 통해 연속적으로 촬영되고 있는 프리뷰 이미지로부터 글자를 재인식할 수도 있다.
- [83] 이하에서는, 도 10 내지 도 13을 참조하여 글자 검출 모델(410)을 훈련시키는 방법에 대해 설명한다.
- [84] 도 10은 일 실시예에 따른 글자 검출 모델(410)의 훈련 방법을 설명하기 위한 순서도이다.
- [85] S1010 단계에서, 글자 인식 장치(100)는 학습용 데이터(60) 내 글자가 존재할 확률을 데이터 공간상에 나타내는 제 1 GT 스코어 맵(71), 및 학습용 데이터(60) 내 글자들 사이의 연결성을 데이터 공간상에 나타내는 제 2 GT 스코어 맵(73)을 획득한다. 학습용 데이터(60)의 가로 크기 및 세로 크기는 입력 데이터(10)의 가로 크기 및 세로 크기와 동일할 수 있다. 또한, 학습용 데이터(60)의 가로 크기 및 세로 크기는 제 1 GT 스코어 맵(71)의 가로 크기 및 세로 크기와 동일하고, 제 2 GT 스코어 맵(73)의 가로 크기 및 세로 크기와도 동일할 수 있다.
- [86] 일 실시예에서, 학습용 데이터(60)는 전술한 원본 이미지(20)와 마찬가지로 카드 등의 대상체를 촬영한 이미지 또는 해당 이미지에 기초하여 추출된 특징 맵을 포함할 수 있다.
- [87] 글자 인식 장치(100)는 학습용 데이터(60)로부터 제 1 GT 스코어 맵(71) 및 제 2 GT 스코어 맵(73) 중 적어도 하나를 직접 생성할 수도 있고, 또는 네트워크나 외부 관리자를 통해 제 1 GT 스코어 맵(71) 및 제 2 GT 스코어 맵(73) 중 적어도 하나를 수신할 수도 있다.
- [88] 제 1 GT 스코어 맵(71) 내 값들은 해당 지점에서 학습용 데이터(60)에 글자가 위치할 확률을 나타낼 수 있다. 또한, 제 2 GT 스코어 맵(73) 내 값들은 해당 지점에서 복수의 글자들이 서로 인접할 확률을 나타낼 수 있다.
- [89] S1020 단계에서, 글자 인식 장치(100)는 학습용 데이터(60)를 글자 검출 모델(410)에 입력한다.
- [90] S1030 단계에서, 학습용 데이터(60)에 대응하여 글자 검출 모델(410)에서 출력되는 제 1 스코어 맵 및 제 2 스코어 맵 각각과 제 1 GT 스코어 맵(71) 및 제 2

GT 스코어 맵(73)의 비교 결과에 따라 글자 검출 모델(410)의 내부 가중치가 갱신될 수 있다.

- [91] 제 1 스코어 맵 및 제 2 스코어 맵 각각과 제 1 GT 스코어 맵(71) 및 제 2 GT 스코어 맵(73)의 비교 결과에 따라 로스(loss) 값이 산출될 수 있다. 로스 값은 예를 들어, L2 Loss 값에 해당할 수 있다. 로스 값은 그 외에도, L1 loss, smooth L1 loss 등 다양한 방법을 이용할 수 있다. 산출된 로스 값은 글자 검출 모델(410)에 입력되고, 글자 검출 모델(410)은 로스 값에 따라 내부 가중치를 갱신할 수 있다.
- [92] 도 11은 제 1 GT 스코어 맵(71)을 생성하는 방법을 설명하기 위한 도면이고, 도 12는 제 2 GT 스코어 맵(73)을 생성하는 방법을 설명하기 위한 도면이다. 또한, 도 13은 서로 인접한 글자 박스들(62a, 62b) 사이에서 연결 박스(63a)를 결정하는 방법을 설명하기 위한 도면이다.
- [93] 도 11을 참조하면, 학습용 데이터(60) 내 적어도 하나의 글자들을 포함하는 단어 영역들에 대해 단어 박스들(61a, 61b, 61c, 61d, 61e)이 결정된다. 그리고, 단어 박스(61a, 61b, 61c, 61d, 61e) 내 포함된 글자들의 개수에 따라 단어 박스(61a, 61b, 61c, 61d, 61e)가 적어도 하나의 글자 박스(62a, 62b, 62c, 62d)로 분할된다. 예를 들어, 어느 하나의 단어 박스 내에 4개의 글자들이 포함되어 있는 경우, 해당 단어 박스는 총 4개의 글자 박스로 분할될 수 있다. 글자 박스(62a, 62b, 62c, 62d) 각각에 소정의 이미지(1100), 예를 들어, 2D 가우시안 이미지가 합성되어 제 1 GT 스코어 맵(71)이 생성될 수 있다.
- [94] 도 12 및 도 13을 참조하면, 복수의 글자 박스들(62a, 62b, 62c, 62d) 중 서로 인접한 글자 박스들 사이의 경계선(L) 상에 위치하는 연결 박스(63a, 63b, 63c)가 결정되고, 연결 박스(63a, 63b, 63c)에 소정 이미지(1100), 예를 들어, 2D 가우시안 이미지가 합성되어 제 2 GT 스코어 맵(73)이 생성될 수 있다.
- [95] 연결 박스(63a, 63b, 63c)는, 서로 인접한 글자 박스들의 내부 공간에 설정된 복수의 지점들을 연결함으로써 결정될 수 있다. 구체적으로, 도 13에 도시된 바와 같이, 서로 인접한 글자 박스들(62a, 62b) 중 좌측 글자 박스(62a) 내 2개의 지점 및 우측 글자 박스(62b) 내 2개의 지점을 연결한 연결 박스(63a)가 결정될 수 있다.
- [96] 일 예에서, 서로 인접한 좌측 글자 박스(62a) 및 우측 글자 박스(62b)의 모서리들 중 좌측 하단 모서리와 우측 상단 모서리를 연결하고, 좌측 상단 모서리와 우측 하단 모서리를 연결하여 상부 및 하부의 삼각형들을 결정하고, 해당 삼각형들의 중점들을 연결함으로써 연결 박스(63a)가 결정될 수도 있다.
- [97] 도 14는 일 실시예에 따른 글자 인식 장치(100)의 구성을 도시하는 블록도이다.
- [98] 도 14를 참조하면, 글자 인식 장치(100)는 메모리(1410), 통신 모듈(1430) 및 프로세서(1450)를 포함할 수 있다. 메모리(1410)에는 적어도 하나의 인스트럭션이 저장될 수 있고, 프로세서(1450)는 적어도 하나의 인스트럭션에 따라 글자 검출 및 글자 검출 모델(410)의 훈련을 제어할 수 있다.
- [99] 도 14는 하나의 메모리(1410)와 하나의 프로세서(1450)만을 도시하고 있으나,

글자 인식 장치(100)는 복수의 메모리 및/또는 복수의 프로세서를 포함할 수도 있다.

[100] 메모리(1410)는 글자 검출 모델(410) 및 글자 인식 모델(420)을 저장할 수 있다. 또한, 메모리(1410)는 특징 추출 모델(800)을 더 저장할 수 있다.

[101] 프로세서(1450)는 글자 검출 모델(410)로 입력 데이터(10)를 입력하고, 글자 검출 모델(410)에서 출력되는 출력 데이터에 기초하여 입력 데이터(10) 내 단어 영역의 위치 정보를 획득할 수 있다. 그리고, 프로세서(1450)는 획득한 위치 정보에 대응하는 부분 데이터를 글자 인식 모델(420)에 입력하고, 글자 인식 모델(420)에서 출력된 글자 정보를 메모리(1410) 또는 기타 저장 장치에 저장할 수 있다.

[102] 일 실시예에서, 프로세서(1450)는 학습용 데이터(60)에 기초하여 글자 검출 모델(410), 글자 인식 모델(420) 및 특징 추출 모델(800) 중 적어도 하나를 훈련시킬 수 있다.

[103] 통신 모듈(1430)은 네트워크를 통해 외부 장치와 데이터를 송수신한다. 예를 들어, 통신 모듈(1430)은 외부 장치와 이미지를 송수신하거나, 입력 데이터(10) 내에서 인식된 글자 정보를 외부 장치와 송수신할 수 있다.

[104] 도 15는 일 실시예에 따른 글자 인식 장치(100)가 적용될 수 있는 서버 장치(1510) 및 클라이언트 장치(1520)를 도시하는 도면이다.

[105] 글자 인식 장치(100)는 서버 장치(1510)로 구현되거나 또는 클라이언트 장치(1520)로 구현될 수 있다.

[106] 글자 인식 장치(100)가 서버 장치(1510)로 구현되는 경우, 서버 장치(1510)는 클라이언트 장치(1520)로부터 이미지를 수신하고, 수신된 이미지 내에서 글자를 인식하여 저장할 수 있다. 일 예에서, 서버 장치(1510)는 클라이언트 장치(1520)로부터 수신된 이미지 내에서 인식된 글자 정보를 클라이언트 장치(1520)로 전송할 수도 있다. 또한, 서버 장치(1510)는 클라이언트 장치(1520)를 포함한 외부 장치로부터 학습용 데이터를 수신하거나, 또는 내부에 저장된 학습용 데이터를 이용하여 글자 검출 모델(410), 글자 인식 모델(420) 및 특징 추출 모델(800) 중 적어도 하나의 훈련을 제어할 수도 있다.

[107] 글자 인식 장치(100)가 클라이언트 장치(1520)로 구현되는 경우, 클라이언트 장치(1520)는 클라이언트 장치(1520)의 카메라에 의해 촬영된 이미지 또는 클라이언트 장치(1520)에 저장된 이미지 내에서 글자를 인식하여 저장할 수 있다.

[108] 일 실시예에서, 클라이언트 장치(1520)는 글자 검출 모델(410), 글자 인식 모델(420) 및 특징 추출 모델(800) 중 적어도 하나의 실행을 위한 데이터를 서버 장치(1510)로부터 수신할 수 있다. 클라이언트 장치(1520)는 카메라 모듈을 통해 촬영된 이미지, 내부 메모리에 저장된 이미지 또는 외부 장치로부터 수신된 이미지를 글자 검출 모델(410), 글자 인식 모델(420) 및 특징 추출 모델(800) 중 적어도 하나에 입력시켜 글자를 인식할 수 있다.

- [109] 클라이언트 장치(1520)는 외부 장치로부터 학습용 데이터를 수신하거나, 또는 내부에 저장된 학습용 데이터를 이용하여 글자 검출 모델(410), 글자 인식 모델(420) 및 특징 추출 모델(800) 중 적어도 하나의 훈련을 제어할 수도 있다. 구현예에 따라, 글자 검출 모델(410), 글자 인식 모델(420) 및 특징 추출 모델(800) 중 적어도 하나의 실행을 위한 데이터를 클라이언트 장치(1520)로 제공한 서버 장치(1510)가 학습용 데이터에 기초하여 글자 검출 모델(410), 글자 인식 모델(420) 및 특징 추출 모델(800) 중 적어도 하나의 훈련을 제어할 수도 있다. 이 경우, 서버 장치(1510)는 훈련 결과 갱신된 가중치 정보만을 클라이언트 장치(1520)로 전송하고, 클라이언트 장치(1520)는 수신된 정보에 따라 글자 검출 모델(410), 글자 인식 모델(420) 및 특징 추출 모델(800) 중 적어도 하나를 갱신할 수 있다.
- [110] 도 15는 클라이언트 장치(1520)로서, 데스크탑 PC를 도시하고 있으나, 이에 한정되는 것은 아니고 클라이언트 장치(1520)는 노트북, 스마트폰, 태블릿 PC, AI(artificial intelligence) 로봇, AI 스피커, 웨어러블 기기 등을 포함할 수 있다.
- [111] 한편, 상술한 본 개시의 실시예들은 컴퓨터에서 실행될 수 있는 프로그램으로 작성가능하고, 작성된 프로그램은 매체에 저장될 수 있다.
- [112] 매체는 컴퓨터로 실행 가능한 프로그램을 계속 저장하거나, 실행 또는 다운로드를 위해 임시 저장하는 것일 수도 있다. 또한, 매체는 단일 또는 수개 하드웨어가 결합된 형태의 다양한 기록수단 또는 저장수단일 수 있는데, 어떤 컴퓨터 시스템에 직접 접속되는 매체에 한정되지 않고, 네트워크 상에 분산 존재하는 것일 수도 있다. 매체의 예시로는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium), 및 ROM, RAM, 플래시 메모리 등을 포함하여 프로그램 명령어가 저장되도록 구성된 것이 있을 수 있다. 또한, 다른 매체의 예시로, 애플리케이션을 유통하는 앱 스토어나 기타 다양한 소프트웨어를 공급 내지 유통하는 사이트, 서버 등에서 관리하는 기록매체 내지 저장매체도 들 수 있다.
- [113] 이상, 본 개시의 기술적 사상을 바람직한 실시예를 들어 상세하게 설명하였으나, 본 개시의 기술적 사상은 상기 실시예들에 한정되지 않고, 본 개시의 기술적 사상의 범위 내에서 당 분야에서 통상의 지식을 가진 자에 의하여 여러 가지 변형 및 변경이 가능하다.

청구범위

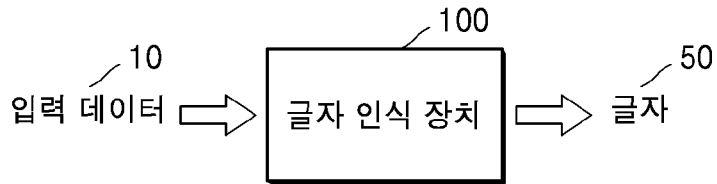
- [청구항 1] 글자 인식 장치에 의한 입력 데이터 내 글자를 인식하는 방법에 있어서,
 입력 데이터를 글자 검출 모델에 입력하는 단계;
 상기 글자 검출 모델에서 출력되는 출력 데이터에 기초하여 상기 입력 데이터 내 단어 영역의 위치 정보를 획득하는 단계;
 상기 획득한 위치 정보에 대응하는 부분 데이터를 상기 입력 데이터로부터 추출하는 단계; 및
 상기 부분 데이터를 글자 인식 모델에 입력하여 상기 부분 데이터 내 글자를 인식하는 단계를 포함하는 것을 특징으로 하는 글자 인식 방법.
- [청구항 2] 제1항에 있어서,
 상기 출력 데이터는,
 상기 입력 데이터 내 글자가 존재할 확률을 상기 입력 데이터에 대응되는 데이터 공간상에 나타내는 제 1 스코어 맵, 및 상기 입력 데이터 내 글자들 사이의 연결성을 상기 입력 데이터에 대응되는 데이터 공간상에 나타내는 제 2 스코어 맵을 포함하는 것을 특징으로 하는 글자 인식 방법.
- [청구항 3] 제2항에 있어서,
 상기 단어 영역의 위치 정보를 획득하는 단계는,
 상기 제 1 스코어 맵 및 상기 제 2 스코어 맵 내 값들과 임계 값의 비교 결과에 따라 상기 제 1 스코어 맵 및 상기 제 2 스코어 맵을 이진화하는 단계;
 이진화된 상기 제 1 스코어 맵과 이진화된 상기 제 2 스코어 맵을 병합하는 단계;
 병합 맵 내에서 소정 값을 갖는 영역을 결정하는 단계; 및
 상기 결정된 영역을 포함하는 단어 영역의 위치 정보를 결정하는 단계를 포함하는 것을 특징으로 하는 글자 인식 방법.
- [청구항 4] 제3항에 있어서,
 상기 단어 영역의 위치 정보를 결정하는 단계는,
 상기 결정된 영역을 포함하는 최소 크기의 단어 박스를 결정하는 단계; 및
 상기 결정된 단어 박스의 위치 정보를 상기 단어 영역의 위치 정보로 결정하는 단계를 포함하는 것을 특징으로 하는 글자 인식 방법.
- [청구항 5] 제2항에 있어서,
 상기 글자 인식 방법은,
 학습용 데이터 내 글자가 존재할 확률을 데이터 공간상에 나타내는 제 1 GT 스코어 맵, 및 상기 학습용 데이터 내 글자들 사이의 연결성을 데이터 공간상에 나타내는 제 2 GT 스코어 맵을 획득하는 단계; 및
 상기 학습용 데이터를 상기 글자 검출 모델에 입력하는 단계를 더 포함하되,

상기 학습용 데이터에 대응하여 상기 글자 검출 모델에서 출력되는 제 1 스코어 맵 및 제 2 스코어 맵 각각과 상기 제 1 GT 스코어 맵 및 상기 제 2 GT 스코어 맵의 비교 결과에 따라 상기 글자 검출 모델의 내부 가중치가 갱신되는 것을 특징으로 하는 글자 인식 방법.

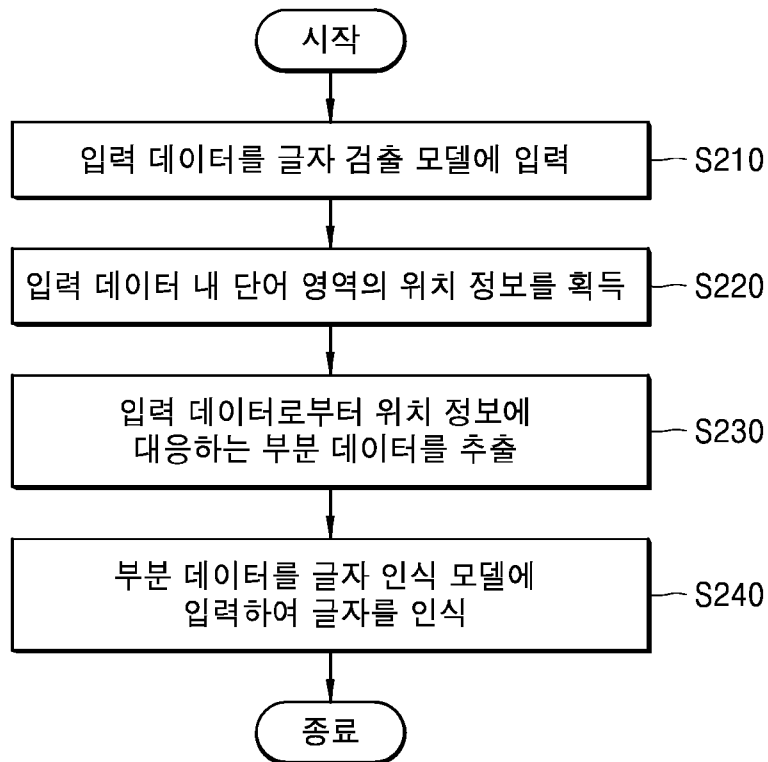
- [청구항 6] 제5항에 있어서,
상기 제 1 GT 스코어 맵을 획득하는 단계는,
상기 학습용 데이터 내 단어를 포함하는 단어 박스를 결정하는 단계;
상기 결정된 단어 박스에 포함된 글자의 개수에 따라 상기 단어 박스를 복수의 글자 박스로 분할하는 단계; 및
상기 복수의 글자 박스 각각에 소정의 이미지를 합성하여 상기 제 1 GT 스코어 맵을 생성하는 단계를 포함하는 것을 특징으로 하는 글자 인식 방법.
- [청구항 7] 제6항에 있어서,
상기 제 2 GT 스코어 맵을 생성하는 단계는,
상기 복수의 글자 박스들 중 서로 인접한 글자 박스들 사이의 경계선 상에 위치하는 연결 박스를 결정하는 단계; 및
상기 연결 박스에 소정의 이미지를 합성하여 상기 제 2 GT 스코어 맵을 생성하는 단계를 포함하는 것을 특징으로 하는 글자 인식 방법.
- [청구항 8] 제1항에 있어서,
상기 글자 인식 방법은,
상기 부분 데이터 내에서 인식된 글자 그룹에 소정의 기호가 포함되어 있는 경우, 해당 글자 그룹을 제 1 종류의 정보로 결정하는 단계를 더 포함하는 것을 특징으로 하는 글자 인식 방법.
- [청구항 9] 제1항에 있어서,
상기 입력 데이터로부터 추출된 부분 데이터의 개수는 복수 개이되,
상기 글자 인식 방법은,
복수의 부분 데이터들 각각에서 인식된 글자 그룹들을, 상기 입력 데이터 내에서의 상기 복수의 부분 데이터들의 위치에 따라 정렬하는 단계를 더 포함하는 것을 특징으로 하는 글자 인식 방법.
- [청구항 10] 제9항에 있어서,
상기 글자 인식 방법은,
상기 정렬된 글자 그룹들 중 연속으로 정렬된 소정 개수의 글자 그룹에 포함된 글자의 개수에 기초하여 글자의 재인식이 필요한지 여부를 결정하는 단계를 더 포함하는 것을 특징으로 하는 글자 인식 방법.
- [청구항 11] 제1항에 있어서,
상기 글자를 인식하는 단계는,
상기 글자 검출 모델에서 출력되는 출력 데이터를 상기 글자 인식 모델로 더 입력시켜 상기 부분 데이터 내 글자를 인식하는 단계를 포함하는 것을

- 특징으로 하는 글자 인식 방법.
- [청구항 12] 제1항에 있어서,
상기 입력 데이터는,
원본 이미지에 대응하여 특징 추출 모델로부터 출력된 특징 맵을
포함하는 것을 특징으로 하는 글자 인식 방법.
- [청구항 13] 하드웨어와 결합하여 제1항의 글자 인식 방법을 실행하기 위하여 매체에
저장된 프로그램.
- [청구항 14] 프로세서; 및
적어도 하나의 인스트럭션을 저장하는 메모리를 포함하되,
상기 프로세서는 상기 적어도 하나의 인스트럭션에 따라,
입력 데이터를 글자 검출 모델에 입력하고,
상기 글자 검출 모델에서 출력되는 출력 데이터에 기초하여 상기 입력
데이터 내 단어 영역의 위치 정보를 획득하고,
상기 획득한 위치 정보에 대응하는 부분 데이터를 상기 입력
데이터로부터 추출하고,
상기 부분 데이터를 글자 인식 모델에 입력하여 상기 부분 데이터 내
글자를 인식하는 것을 특징으로 하는 글자 인식 장치.

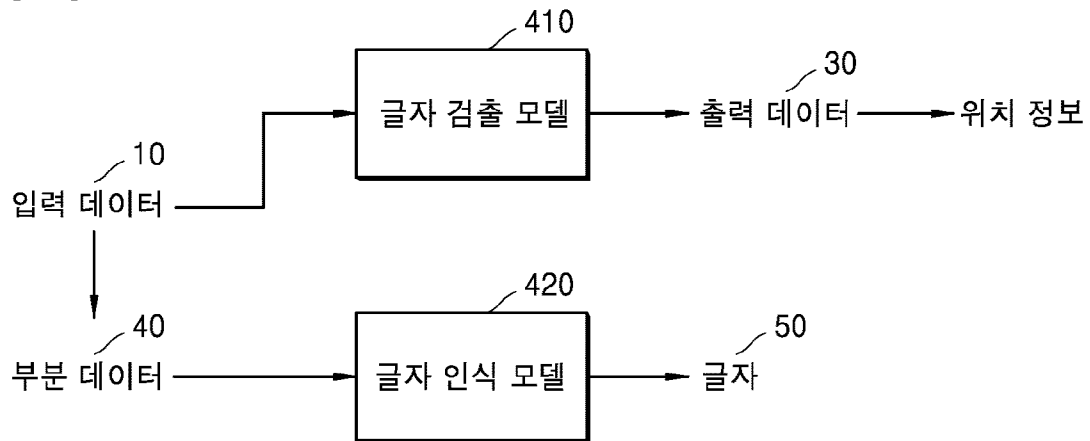
[도1]



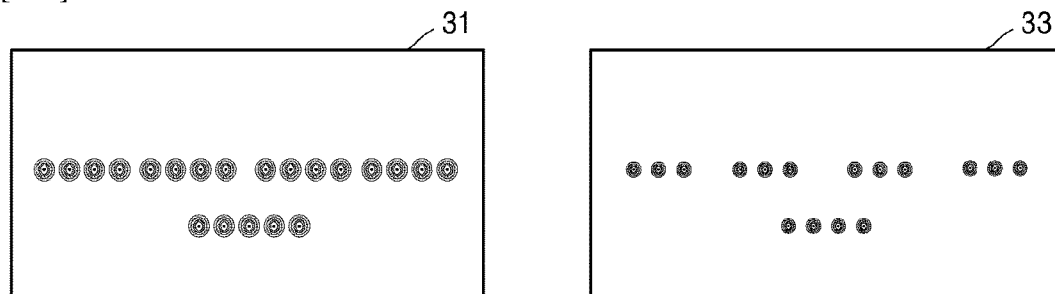
[도2]



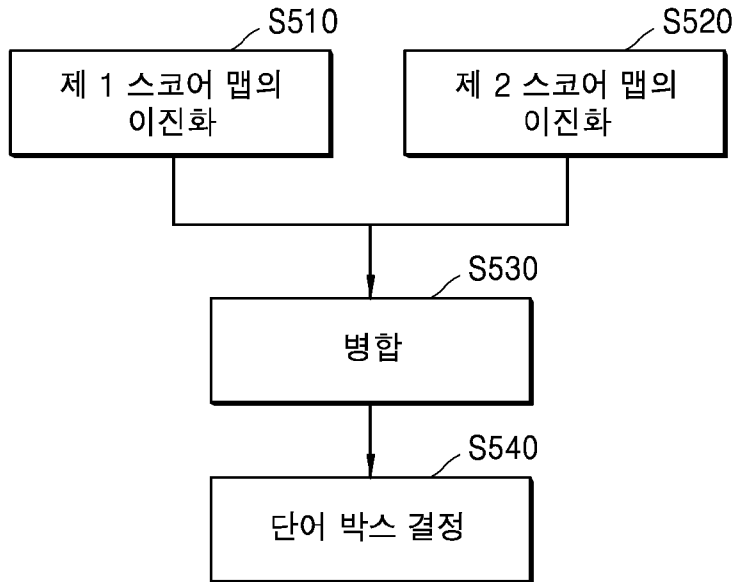
[도3]



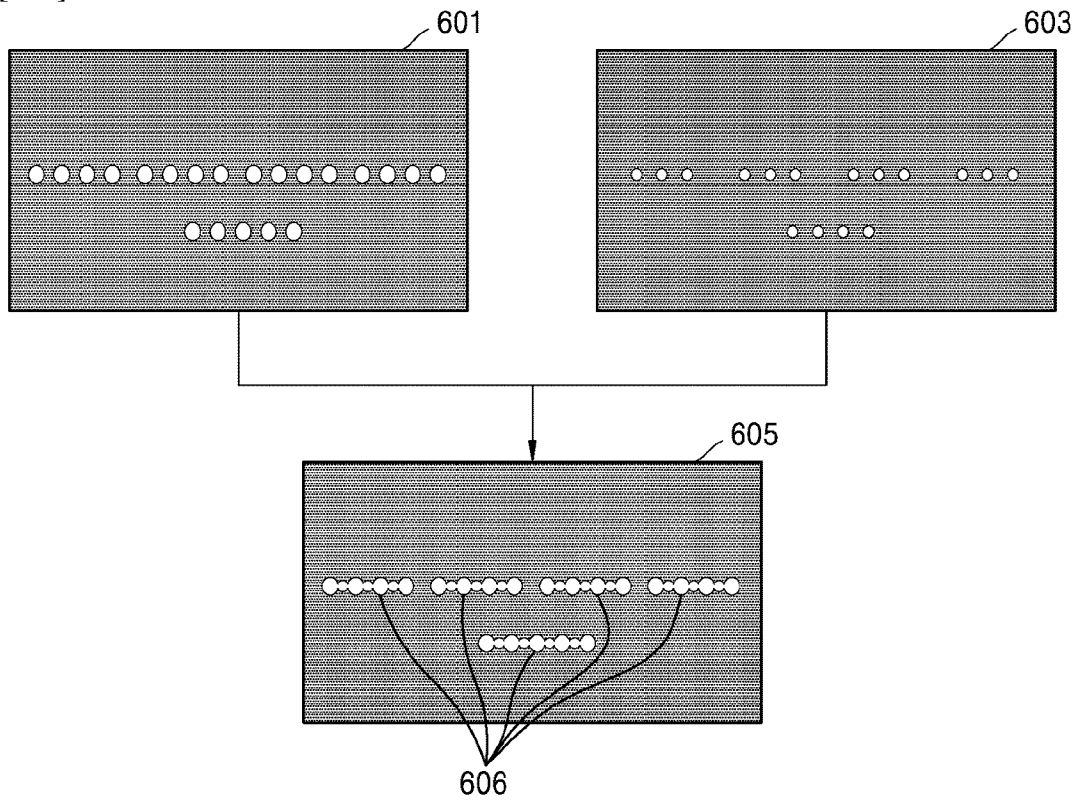
[도4]



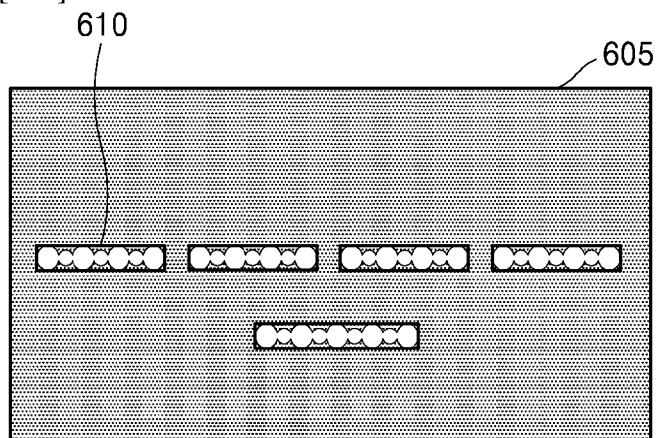
[도5]



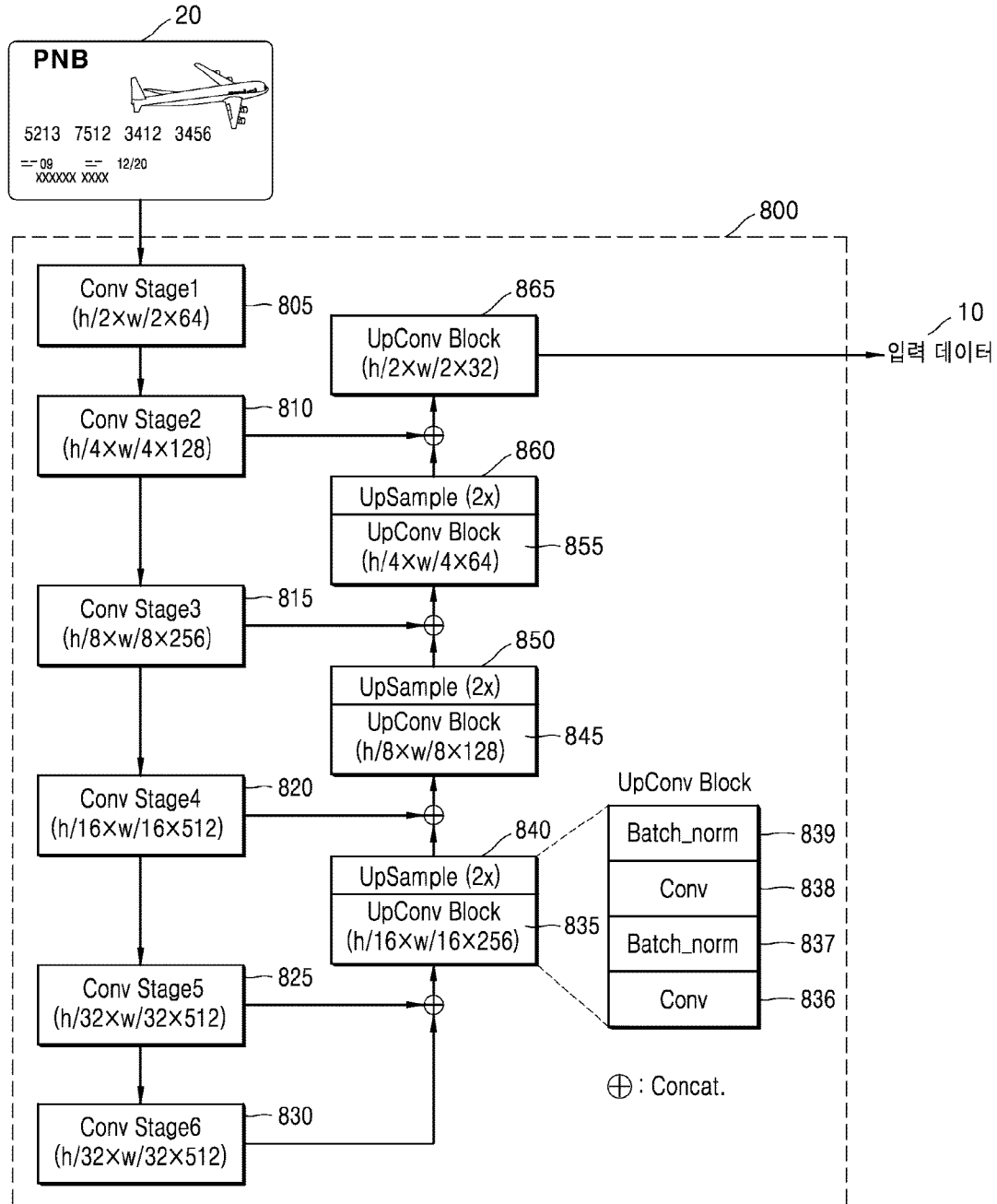
[도6]



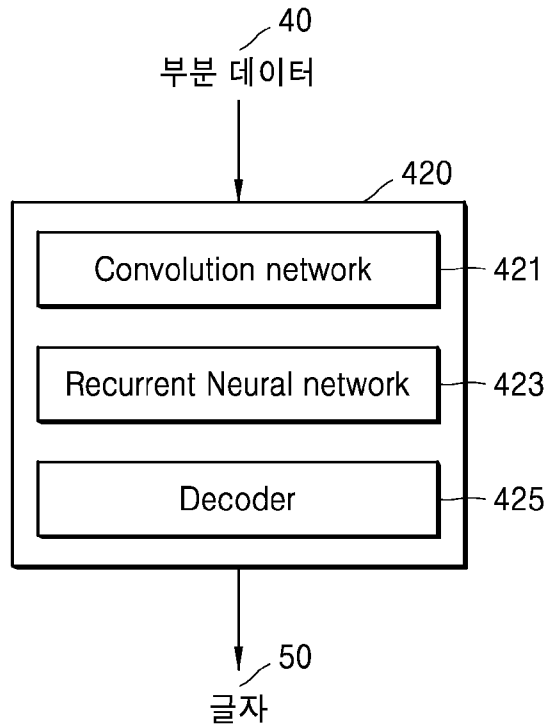
[도7]



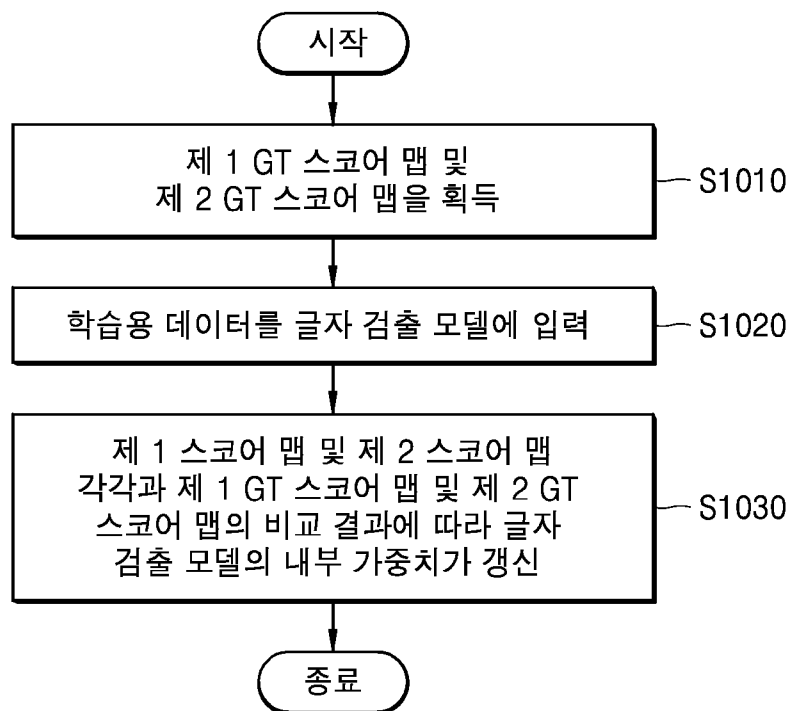
[도8]



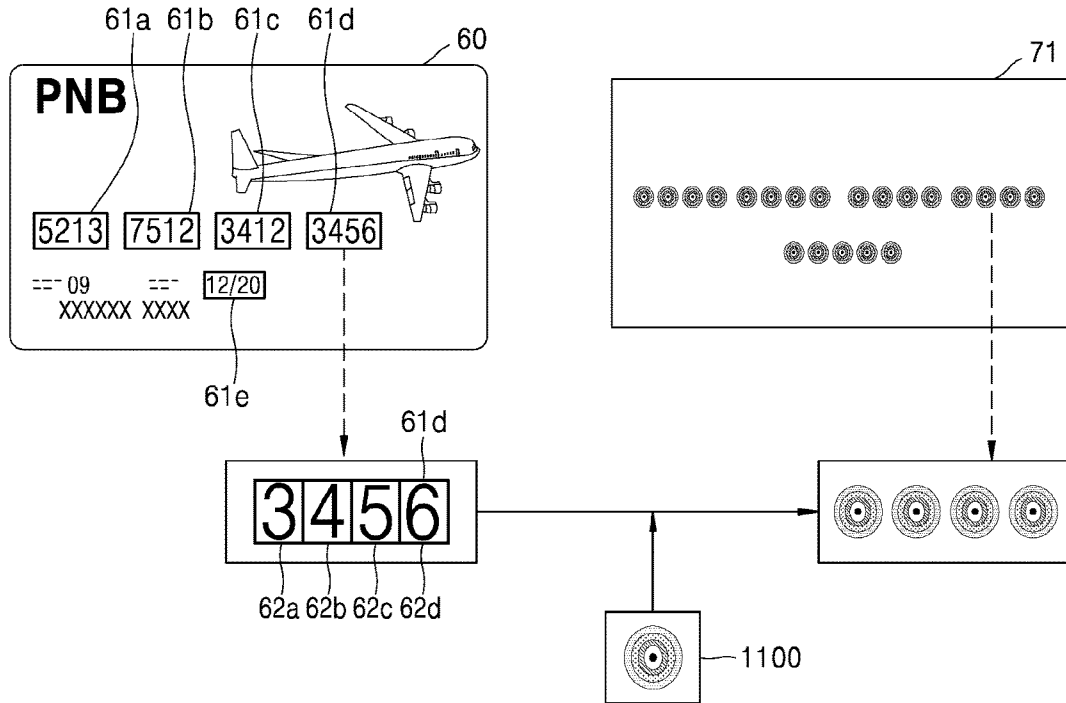
[도9]



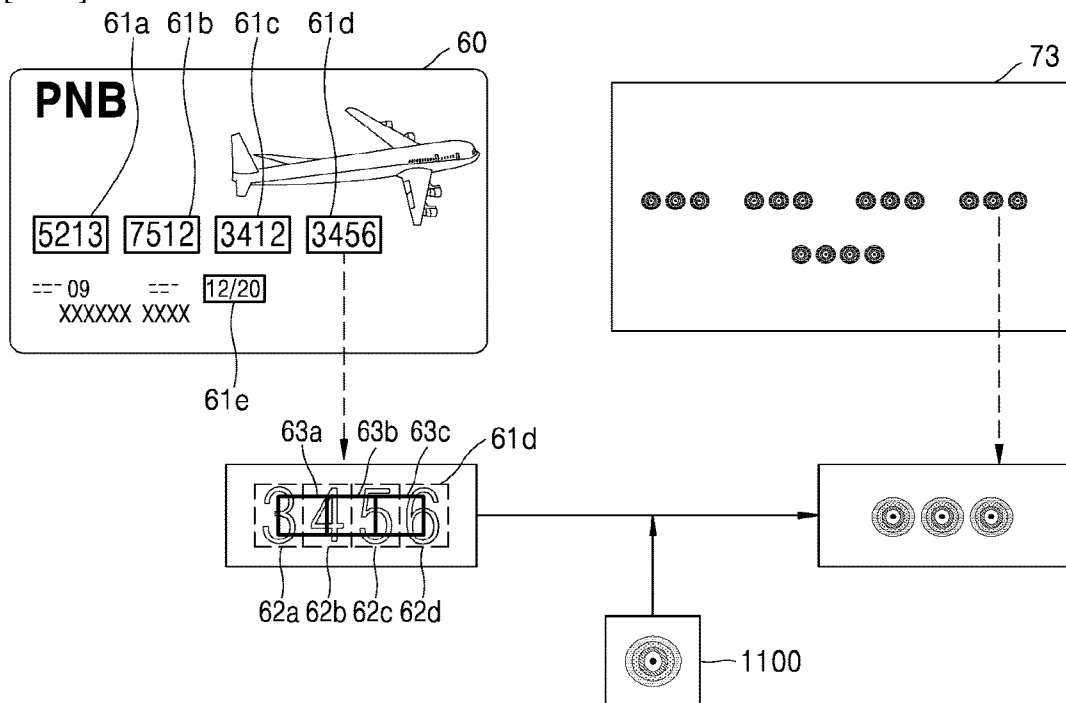
[도10]



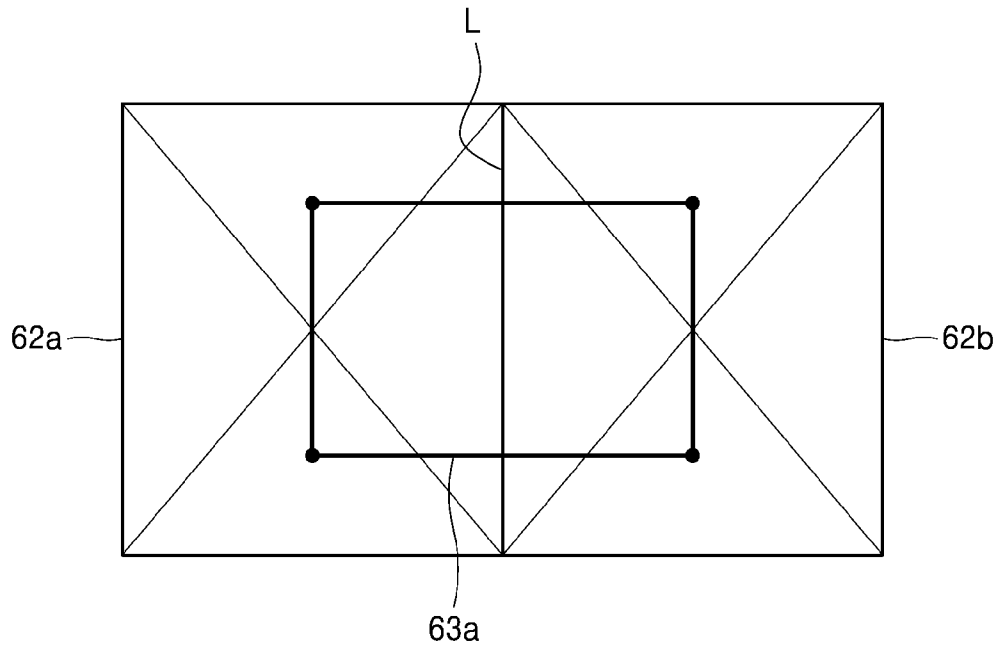
[도11]



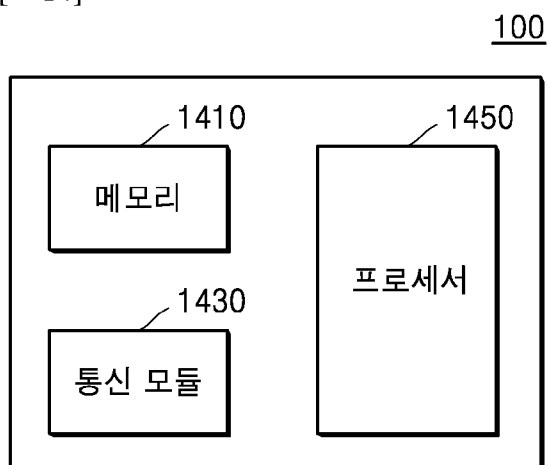
[도12]



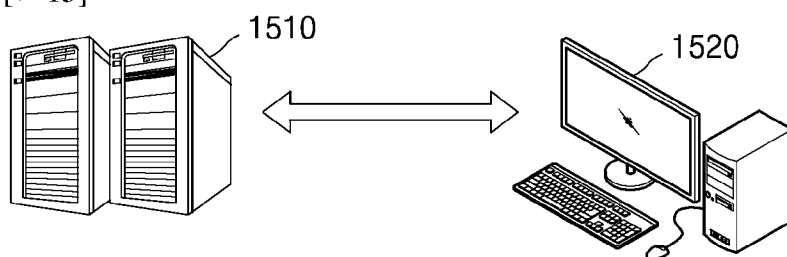
[도13]



[도14]



[도15]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2020/001333

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/32(2006.01)i, G06K 9/00(2006.01)i, G06T 7/11(2017.01)i, G06T 11/60(2006.01)i, 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)

G06K 9/32; G06F 17/27; G06F 17/30; G06F 3/023; G06K 9/18; G06K 9/34; G06K 9/46; G06K 9/58; G06K 9/66; G06K 9/00; G06T 7/11; G06T 11/60; G06N 20/00

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: character detection model, position information, partial data, probability

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KR 10-2017-0029947 A (SK PLANET CO., LTD.) 16 March 2017 See paragraphs [0122]-[0125], claims 1-10 and figures 1-9.	1,8-9,11-13
Y		14
A		2-7,10
Y	KR 10-1805318 B1 (POSTECH ACADEMY-INDUSTRY FOUNDATION) 06 December 2017 See paragraphs [0070]-[0071], claim 11 and figure 1.	14
A	KR 10-1295000 B1 (KGMOBILIANS. CO., LTD.) 09 August 2013 See claim 1 and figures 1-7.	1-14
A	KR 10-2018-0112590 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 12 October 2018 See claim 1 and figures 1-8.	1-14
A	KR 10-2016-0065174 A (MICROSOFT TECHNOLOGY LICENSING, LLC.) 08 June 2016 See claim 1 and figures 1-12.	1-14



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

25 MAY 2020 (25.05.2020)

Date of mailing of the international search report

25 MAY 2020 (25.05.2020)

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/KR2020/001333

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2017-0029947 A	16/03/2017	CN 106503703 A KR 10-1824600 B1 US 10121087 B2 US 2017-0068867 A1	15/03/2017 14/03/2018 06/11/2018 09/03/2017
KR 10-1805318 B1	06/12/2017	None	
KR 10-1295000 B1	09/08/2013	None	
KR 10-2018-0112590 A	12/10/2018	US 2018-0285744 A1	04/10/2018
KR 10-2016-0065174 A	08/06/2016	CN 105683874 A EP 3053009 A1 US 2015-0100537 A1 WO 2015-050910 A1	15/06/2016 10/08/2016 09/04/2015 09/04/2015

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G06K 9/32(2006.01)i, G06K 9/00(2006.01)i, G06T 7/11(2017.01)i, G06T 11/60(2006.01)i, G06N 20/00(2019.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G06K 9/32; G06F 17/27; G06F 17/30; G06F 3/023; G06K 9/18; G06K 9/34; G06K 9/46; G06K 9/58; G06K 9/66; G06K 9/00; G06T 7/11; G06T 11/60; G06N 20/00

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: 글자 검출 모델(character detection model), 위치 정보(position information), 부분 데이터(partial data), 확률(probability)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
X	KR 10-2017-0029947 A (에스케이플래닛 주식회사) 2017.03.16 단락 [0122]-[0125], 청구항 1-10 및 도면 1-9	1, 8-9, 11-13
Y		14
A		2-7, 10
Y	KR 10-1805318 B1 (포항공과대학교 산학협력단) 2017.12.06 단락 [0070]-[0071], 청구항 11 및 도면 1	14
A	KR 10-1295000 B1 (주식회사 케이지모빌리언스) 2013.08.09 청구항 1 및 도면 1-7	1-14
A	KR 10-2018-0112590 A (한국전자통신연구원) 2018.10.12 청구항 1 및 도면 1-8	1-14
A	KR 10-2016-0065174 A (마이크로소프트 테크놀로지 라이선싱, 엘엘씨) 2016.06.08 청구항 1 및 도면 1-12	1-14

☐ 추가 문헌이 C(계속)에 기재되어 있습니다.

☒ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

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

“D” 본 국제출원에서 출원인이 인용한 문헌

“E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후 “X”에 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

2020년 05월 25일 (25.05.2020)

국제조사보고서 발송일

2020년 05월 25일 (25.05.2020)

ISA/KR의 명칭 및 우편주소



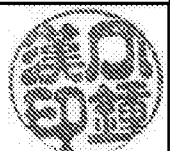
대한민국 특허청
(35208) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)

팩스 번호 +82-42-481-8578

심사관

정종환

전화번호 +82-42-481-5642



국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2017-0029947 A	2017/03/16	CN 106503703 A KR 10-1824600 B1 US 10121087 B2 US 2017-0068867 A1	2017/03/15 2018/03/14 2018/11/06 2017/03/09
KR 10-1805318 B1	2017/12/06	없음	
KR 10-1295000 B1	2013/08/09	없음	
KR 10-2018-0112590 A	2018/10/12	US 2018-0285744 A1	2018/10/04
KR 10-2016-0065174 A	2016/06/08	CN 105683874 A EP 3053009 A1 US 2015-0100537 A1 WO 2015-050910 A1	2016/06/15 2016/08/10 2015/04/09 2015/04/09