(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0093754
(43) 공개일자 2023년06월27일

(51) 국제특허분류(Int. Cl.)
G06F 40/58 (2020.01) *G06F 40/237* (2020.01)
G06F 40/263 (2020.01) *G06F 40/268* (2020.01)
G06F 40/284 (2020.01) *G06N 3/08* (2023.01)
(52) CPC특허분류
G06F 40/58 (2020.01)
G06F 40/237 (2022.01)
(21) 출원번호 10-2021-0182628
(22) 출원일자 2021년12월20일
심사청구일자 2021년12월20일

(71) 출원인
한림대학교 산학협력단
강원도 춘천시 한림대학길 1, 한림대학교(옥천동)
(72) 발명자
김철호
강원도 춘천시 서부대성로 34, 104동 204호(요선동, 한신휴플러스아파트)
김유섭
강원도 춘천시 후석로 325, 104동 108호(후평동, 춘천더샾아파트)
(뒷면에 계속)
(74) 대리인
특허법인본

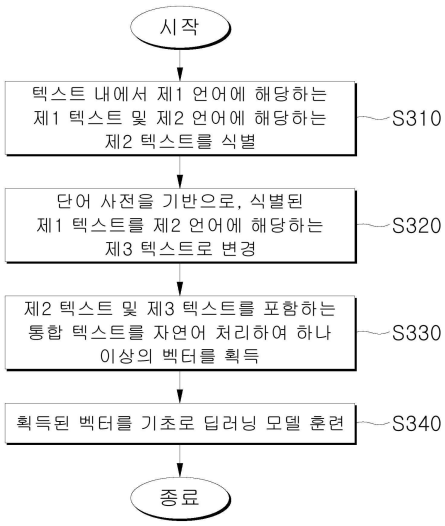
전체 청구항 수 : 총 11 항

(54) 발명의 명칭 **전자 의무 기록을 구성하는 텍스트의 분석과 관련된 전처리를 수행하는 전자 장치**

(57) 요 약

전자 장치가 개시된다. 본 전자 장치는, 제1 언어의 단어와 제2 언어의 단어 간의 매칭 관계를 나타내는 단어 사전, 및 상기 제2 언어의 텍스트를 처리하기 위한 적어도 하나의 딥러닝 모델이 저장된 메모리, 메모리와 연결된 프로세서를 포함한다. 프로세서는, 전체 텍스트 내에서 제1 언어에 해당하는 제1 텍스트 및 제2 언어에 해당하는 제2 텍스트를 식별하고, 단어 사전을 기반으로, 식별된 제1 텍스트를 제2 언어에 해당하는 제3 텍스트로 변환하고, 제2 텍스트 및 제3 텍스트를 포함하는 통합 텍스트를 자연어 처리하여 하나 이상의 벡터를 획득하고, 획득된 벡터를 기초로 딥러닝 모델을 훈련시킨다.

대 표 도 - 도3



(52) CPC특허분류
G06F 40/263 (2020.01)
G06F 40/268 (2020.01)
G06F 40/284 (2020.01)
G06N 3/08 (2023.01)

(72) 발명자
최정명
강원도 춘천시 향교옆길 21번길 7, 203호(교동)
서수영
강원도 춘천시 동면 소양강로 102, 207동 902(장학엘에이치해온채)
이재준
강원도 춘천시 삭주로 77 본관 10층 병원장실

이 발명을 지원한 국가연구개발사업

과제고유번호	1711116275
과제번호	2019R1G1A1097707
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	생애 첫 연구
연구과제명	딥러닝 기반 자연어처리를 통한 뇌경색 자동 예후예측 시스템 개발
기 여 율	1/2
과제수행기관명	한림대학교 산학협력단
연구기간	2019.09.01 ~ 2022.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1465034383
과제번호	HR21C0198040021
부처명	보건복지부
과제관리(전문)기관명	한국보건산업진흥원
연구사업명	연구중심병원육성R&D
연구과제명	유전체 · 인공지능 기반 환자맞춤형 신치료 개발 및 관리 기술 개발
기 여 율	1/2
과제수행기관명	한림대학교 산학협력단
연구기간	2021.07.01 ~ 2029.12.31

명 세 서

청구범위

청구항 1

전자 장치에 있어서,

제1 언어의 단어와 제2 언어의 단어 간의 매칭 관계를 나타내는 단어 사전, 및 상기 제2 언어의 텍스트를 처리하기 위한 적어도 하나의 딥러닝 모델이 저장된 메모리; 및

상기 메모리와 연결된 프로세서를 포함하고,

상기 프로세서는,

전체 텍스트 내에서 상기 제1 언어에 해당하는 제1 텍스트 및 상기 제2 언어에 해당하는 제2 텍스트를 식별하고,

상기 단어 사전을 기반으로, 상기 식별된 제1 텍스트를 상기 제2 언어에 해당하는 제3 텍스트로 변환하고,

상기 제2 텍스트 및 상기 제3 텍스트를 포함하는 통합 텍스트를 자연어 처리하여 하나 이상의 벡터를 획득하고,

상기 획득된 벡터를 기초로 상기 딥러닝 모델을 훈련시키는, 전자 장치.

청구항 2

제1항에 있어서,

상기 프로세서는,

상기 제1 텍스트 내에서, 상기 단어 사전에 포함된 적어도 하나의 제1 키워드를 식별하고,

상기 단어 사전을 기반으로, 상기 제1 텍스트에 포함된 상기 제1 키워드를 상기 제2 언어에 해당하는 제2 키워드로 변환하고,

상기 제2 키워드가 하나의 개체로 설정된 상태에서, 상기 제1 텍스트를 상기 제2 언어에 해당하는 상기 제3 텍스트로 변환하는, 전자 장치.

청구항 3

제2항에 있어서,

상기 프로세서는,

상기 획득된 벡터를 기초로 훈련된 상기 딥러닝 모델의 훈련 전후 정확도를 비교하고,

훈련 전후의 정확도에 대한 비교 결과 상기 딥러닝 모델의 정확도가 낮아진 경우, 상기 제1 키워드에 매칭되는 상기 제2 언어의 키워드가 변경되도록 상기 단어 사전을 업데이트하는, 전자 장치.

청구항 4

제2항에 있어서,

상기 프로세서는,

상기 단어 사전에 따라 상기 제1 키워드에 매칭되는 복수의 제2 키워드가 식별되면, 상기 복수의 제2 키워드 각각을 독립적으로 반영하여 복수의 제3 텍스트를 획득하고,

상기 복수의 제3 텍스트 각각에 매칭되는 벡터가 독립적으로 입력됨에 따른 상기 딥러닝 모델의 출력 간의 차이가 임계치 미만인 경우, 상기 단어 사전을 유지하고,

상기 복수의 제3 텍스트 각각에 매칭되는 벡터가 독립적으로 입력됨에 따른 상기 딥러닝 모델의 출력 간의 차이가 상기 임계치 이상인 경우, 상기 단어 사전을 업데이트하는, 전자 장치.

청구항 5

제1항에 있어서,

상기 프로세서는,

상기 통합 텍스트에 대하여 tokenization, 품사 태깅, 및 word embedding을 수행하여 하나 이상의 벡터를 획득하는, 전자 장치.

청구항 6

제1항에 있어서,

상기 메모리는,

제1 전문 분야에 대한 제1 단어 사전, 제2 전문 분야에 대한 제2 단어 사전, 및 제3 전문 분야에 대한 제3 단어 사전을 포함하고,

상기 프로세서는,

상기 전체 텍스트가 상기 제1 전문 분야로 설정된 경우, 상기 제1 단어 사전을 활용하여 상기 제1 텍스트를 상기 제2 언어에 해당하는 제3 텍스트로 변환하는, 전자 장치.

청구항 7

제6항에 있어서,

상기 프로세서는,

상기 제1 텍스트 내에 상기 제1 단어 사전에 포함된 키워드가 존재하지 않는 경우, 상기 제2 단어 사전 및 상기 제3 단어 사전 중 적어도 하나를 활용하여 상기 제1 텍스트를 상기 제2 언어에 해당하는 제3 텍스트로 변환하는, 전자 장치.

청구항 8

제7항에 있어서,

상기 프로세서는,

상기 제2 단어 사전 및 상기 제3 단어 사전 중, 상기 제1 텍스트에 포함된 키워드와 관련된 적어도 하나의 단어 사전을 선택하고,

상기 선택된 단어 사전을 활용하여 상기 제1 텍스트를 상기 제2 언어에 해당하는 제3 텍스트로 변환하고,

상기 선택된 단어 사전 내에서 상기 키워드와 매칭되는 상기 제2 언어의 키워드를 바탕으로, 상기 제1 단어 사전을 업데이트하는, 전자 장치.

청구항 9

전자 장치의 제어 방법에 있어서,

상기 전자 장치는,

제1 언어의 단어와 제2 언어의 단어 간의 매칭 관계를 나타내는 단어 사전, 및 상기 제2 언어의 텍스트를 처리하기 위한 적어도 하나의 딥러닝 모델을 포함하고,

상기 제어 방법은,

전체 텍스트 내에서 상기 제1 언어에 해당하는 제1 텍스트 및 상기 제2 언어에 해당하는 제2 텍스트를 식별하는 단계;

상기 단어 사전을 기반으로, 상기 식별된 제1 텍스트를 상기 제2 언어에 해당하는 제3 텍스트로 변환하는 단계;

상기 제2 텍스트 및 상기 제3 텍스트를 포함하는 통합 텍스트를 자연어 처리하여 하나 이상의 벡터를 획득하는 단계; 및

상기 획득된 벡터를 기초로 상기 딥러닝 모델을 훈련시키는 단계;를 포함하는, 전자 장치의 제어 방법.

청구항 10

컴퓨터 판독 가능 매체에 저장된 컴퓨터 프로그램에 있어서,

전자 장치의 프로세서에 의해 실행되어 상기 전자 장치로 하여금 제9항의 제어 방법을 수행하도록 하는, 컴퓨터 판독 가능 매체에 저장된 컴퓨터 프로그램.

청구항 11

전자 장치에 있어서,

제1 언어의 단어와 제2 언어의 단어 간의 매칭 관계를 나타내는 단어 사전이 저장된 메모리; 및

상기 메모리와 연결된 프로세서를 포함하고,

상기 프로세서는,

전체 텍스트 내에서 상기 제1 언어에 해당하는 제1 텍스트 및 상기 제2 언어에 해당하는 제2 텍스트를 식별하고,

상기 단어 사전을 기반으로, 상기 식별된 제1 텍스트를 상기 제2 언어에 해당하는 제3 텍스트로 변환하고,

상기 제2 텍스트 및 상기 제3 텍스트를 포함하는 통합 텍스트를 자연어 처리하여 하나 이상의 벡터를 획득하는, 전자 장치.

발명의 설명

기 술 분 야

[0001] 본 개시는 텍스트의 분석을 위한 전처리를 수행하는 전자 장치에 관한 것으로, 보다 상세하게는, 전문 분야에 특화된 단어 사전을 바탕으로 특정 언어에 대한 번역을 수행함으로써 복수의 언어가 혼재된 텍스트 전체를 효과적으로 처리할 수 있는 전자 장치에 관한 것이다.

배 경 기 술

[0002] 뇌경색 등 다양한 질환의 환자들의 전자 의무 기록을 활용함에 있어, 일반적인 기계번역으로는 정확한 의학용어를 영문으로 치환하는 것이 효율적이지 않았다.

[0003] 특히, 일반적인 대화 또는 기사 등에서 자주 사용하게 되는 단어들의 경우 의학 용어와는 상당히 동떨어져 있으며, 이러한 일반적인 텍스트 전처리 방법을 사용할 경우에는 불명확한 한글의 영문 전환으로 인한 텍스트의 의미가 기계적으로 정확하게 인식될 수 없다.

[0004] 인공어에 비하여 자연어의 경우, 특히 한글의 경우는 의학, 예술 등과 같은 각각의 분야에서도 동일한 단어가 다른 의미로 사용되어 기계적인 번역을 적용할 경우에는 잘못된 벡터로의 전환을 일으켜 효율적인 텍스트 데이터를 구현하는 것이 불가능하다.

[0005] 도 1a는 종래의 영문 자연어처리의 과정을 설명하기 위한 모식도, 도 1b는 종래의 한글 자연어처리의 과정을 설명하기 위한 모식도이다.

[0006] 도 1a를 참조하면, Tokenization, Stemming, Stop Word Removal, Part-Of-Speech Tagging 등을 통하여 전처리된 영문은 영문 텍스트 벡터로 사용될 수 있다.

[0007] 그리고, 도 1b를 참조하면, 한글 텍스트의 경우 Tokenization, POS Tagging 외에 Entity Detection, Relation Detection 등의 과정을 거쳐 벡터로 사용될 수 있다.

[0008] 다만, 한영이 혼재되어 있는 경우에는 token 제작에 한글과 영문 token을 각각 따로 제작해야 하며, 이러한 경우에는 한글, 영문의 순서가 뒤섞여 정확한 문맥의 해석을 기대할 수 없다.

선행기술문헌

특허문헌

[0009] (특허문헌 0001) 등록 특허 공보 제10-233144호(신경망을 이용한 텍스트 인식 시스템 및 그 방법)

발명의 내용

해결하려는 과제

[0010] 본 개시는 복수의 언어(ex. 한글, 영어 등)가 포함된 텍스트를 전처리함에 있어, 전문 분야에 특화된 단어 사전을 활용하여 특정 언어로 모두 변환한 뒤, 통합 텍스트를 활용하여 딥러닝 모델을 구축하는 전자 장치 및 제어 방법을 제공한다.

[0011] 본 개시의 목적들은 이상에서 언급한 목적으로 제한되지 않으며, 언급되지 않은 본 개시의 다른 목적 및 장점들은 하기의 설명에 의해서 이해될 수 있고, 본 개시의 실시 예에 의해 보다 분명하게 이해될 것이다. 또한, 본 개시의 목적 및 장점들은 특허 청구 범위에 나타낸 수단 및 그 조합에 의해 실현될 수 있음을 쉽게 알 수 있을 것이다.

과제의 해결 수단

[0012] 본 개시의 일 실시 예에 따른 전자 장치는, 제1 언어의 단어와 제2 언어의 단어 간의 매칭 관계를 나타내는 단어 사전, 및 상기 제2 언어의 텍스트를 처리하기 위한 적어도 하나의 딥러닝 모델이 저장된 메모리, 상기 메모리와 연결된 프로세서를 포함한다. 상기 프로세서는, 전체 텍스트 내에서 상기 제1 언어에 해당하는 제1 텍스트 및 상기 제2 언어에 해당하는 제2 텍스트를 식별하고, 상기 단어 사전을 기반으로, 상기 식별된 제1 텍스트를 상기 제2 언어에 해당하는 제3 텍스트로 변환하고, 상기 제2 텍스트 및 상기 제3 텍스트를 포함하는 통합 텍스트를 자연어 처리하여 하나 이상의 벡터를 획득하고, 상기 획득된 벡터를 기초로 상기 딥러닝 모델을 훈련시킨다.

[0013] 상기 프로세서는, 상기 제1 텍스트 내에서, 상기 단어 사전에 포함된 적어도 하나의 제1 키워드를 식별하고, 상기 단어 사전을 기반으로, 상기 제1 텍스트에 포함된 상기 제1 키워드를 상기 제2 언어에 해당하는 제2 키워드로 변환하고, 상기 제2 키워드가 하나의 개체로 설정된 상태에서, 상기 제1 텍스트를 상기 제2 언어에 해당하는 상기 제3 텍스트로 변환할 수 있다.

[0014] 이 경우, 상기 프로세서는, 상기 획득된 벡터를 기초로 훈련된 상기 딥러닝 모델의 훈련 전후 정확도를 비교하고, 훈련 전후의 정확도에 대한 비교 결과 상기 딥러닝 모델의 정확도가 낮아진 경우, 상기 제1 키워드에 매칭되는 상기 제2 언어의 키워드가 변경되도록 상기 단어 사전을 업데이트할 수 있다.

[0015] 또한, 상기 프로세서는, 상기 단어 사전에 따라 상기 제1 키워드에 매칭되는 복수의 제2 키워드가 식별되면, 상기 복수의 제2 키워드 각각을 독립적으로 반영하여 복수의 제3 텍스트를 획득하고, 상기 복수의 제3 텍스트 각각에 매칭되는 벡터가 독립적으로 입력됨에 따른 상기 딥러닝 모델의 출력 간의 차이가 임계치 미만인 경우, 상기 단어 사전을 유지하고, 상기 복수의 제3 텍스트 각각에 매칭되는 벡터가 독립적으로 입력됨에 따른 상기 딥러닝 모델의 출력 간의 차이가 상기 임계치 이상인 경우, 상기 단어 사전을 업데이트할 수도 있다.

[0016] 상기 프로세서는, 상기 통합 텍스트에 대하여 tokenization, 품사 태깅, 및 word embedding을 수행하여 하나 이상의 벡터를 획득할 수 있다.

[0017] 한편, 상기 메모리는, 제1 전문 분야에 대한 제1 단어 사전, 제2 전문 분야에 대한 제2 단어 사전, 및 제3 전문 분야에 대한 제3 단어 사전을 포함할 수도 있다. 이 경우, 상기 프로세서는, 상기 전체 텍스트가 상기 제1 전문 분야로 설정된 경우, 상기 제1 단어 사전을 활용하여 상기 제1 텍스트를 상기 제2 언어에 해당하는 제3 텍스트로 변환할 수 있다.

[0018] 여기서, 상기 프로세서는, 상기 제1 텍스트 내에 상기 제1 단어 사전에 포함된 키워드가 존재하지 않는 경우, 상기 제2 단어 사전 및 상기 제3 단어 사전 중 적어도 하나를 활용하여 상기 제1 텍스트를 상기 제2 언어에 해당하는 제3 텍스트로 변환할 수도 있다.

[0019] 이 경우, 상기 프로세서는, 상기 제2 단어 사전 및 상기 제3 단어 사전 중, 상기 제1 텍스트에 포함된 키워드와 관련된 적어도 하나의 단어 사전을 선택하고, 상기 선택된 단어 사전을 활용하여 상기 제1 텍스트를 상기 제2

언어에 해당하는 제3 텍스트로 변환하고, 상기 선택된 단어 사전 내에서 상기 키워드와 매칭되는 상기 제2 언어의 키워드를 바탕으로, 상기 제1 단어 사전을 업데이트할 수도 있다.

[0020] 본 개시의 일 실시 예에 따라 제1 언어의 단어와 제2 언어의 단어 간의 매칭 관계를 나타내는 단어 사전, 및 상기 제2 언어의 텍스트를 처리하기 위한 적어도 하나의 딥러닝 모델을 포함하는 전자 장치의 제어 방법은, 전체 텍스트 내에서 상기 제1 언어에 해당하는 제1 텍스트 및 상기 제2 언어에 해당하는 제2 텍스트를 식별하는 단계, 상기 단어 사전을 기반으로, 상기 식별된 제1 텍스트를 상기 제2 언어에 해당하는 제3 텍스트로 변환하는 단계, 상기 제2 텍스트 및 상기 제3 텍스트를 포함하는 통합 텍스트를 자연어 처리하여 하나 이상의 벡터를 획득하는 단계, 상기 획득된 벡터를 기초로 상기 딥러닝 모델을 훈련시키는 단계를 포함한다.

[0021] 본 개시의 일 실시 예에 따른 전자 장치는, 제1 언어의 단어와 제2 언어의 단어 간의 매칭 관계를 나타내는 단어 사전이 저장된 메모리, 상기 메모리와 연결된 프로세서를 포함한다. 상기 프로세서는, 전체 텍스트 내에서 상기 제1 언어에 해당하는 제1 텍스트 및 상기 제2 언어에 해당하는 제2 텍스트를 식별하고, 상기 단어 사전을 기반으로, 상기 식별된 제1 텍스트를 상기 제2 언어에 해당하는 제3 텍스트로 변환하고, 상기 제2 텍스트 및 상기 제3 텍스트를 포함하는 통합 텍스트를 자연어 처리하여 하나 이상의 벡터를 획득할 수 있다.

발명의 효과

[0022] 본 개시에 따른 전자 장치 및 제어 방법은, 복수의 언어가 혼재된 전체 텍스트를 취급함에 있어서, 전문 분야에 최적화된 단어 사전이 활용함으로써 언어의 종류를 일원화할 수 있고, 그 결과 맥락이 통일된 정밀한 자연처 처리가 수행됨으로써 딥러닝 모델의 훈련 환경을 개선하는 효과가 있다.

[0023] 의학 전문 분야에 활용되는 경우, 본 개시에 따른 전자 장치는, 의학용어의 의미를 정확하게 전달할 수 있는 지식기반 딕셔너리를 제작한 이후 이를 전자의무기록에 적용한 이후 추가적인 기계번역을 통한 통합 전처리방법을 통하여 한영혼재 전자의무기록 자연어처리의 효율성을 확보할 수 있다. 이 경우, 단어들의 개별적인 의미 외에도 문장의 의미나 문맥 등이 최대한 반영될 수 있다. 본 개시에 따른 전자 장치는 뇌경색 등 뇌질환을 포함하여 다양한 종류의 질환에 대하여 응용/적용될 수 있는 전처리 과정을 수행한다.

도면의 간단한 설명

[0024] 도 1a는 종래의 영문 자연어처리의 과정을 설명하기 위한 모식도,

도 1b는 종래의 한글 자연어처리의 과정을 설명하기 위한 모식도,

도 2는 본 개시의 일 실시 예에 따라 단어 사전을 포함하는 전자 장치의 구성을 설명하기 위한 블록도,

도 3은 본 개시의 일 실시 예에 따른 전자 장치의 동작을 설명하기 위한 흐름도,

도 4는 본 개시의 일 실시 예에 따른 전자 장치가 단어 사전을 활용하여 한영 변환을 수행하는 동작을 설명하기 위한 도면, 그리고

도 5는 본 개시의 일 실시 예에 따른 전자 장치가 전문 분야 별로 구분된 단어 사전을 선택적으로 활용하는 동작을 설명하기 위한 알고리즘이다.

발명을 실시하기 위한 구체적인 내용

[0025] 본 개시에 대하여 구체적으로 설명하기에 앞서, 본 명세서 및 도면의 기재 방법에 대하여 설명한다.

[0026] 먼저, 본 명세서 및 청구범위에서 사용되는 용어는 본 개시의 다양한 실시 예들에서의 기능을 고려하여 일반적인 용어들을 선택하였다. 하지만, 이러한 용어들은 당해 기술 분야에 종사하는 기술자의 의도나 법률적 또는 기술적 해석 및 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 일부 용어는 출원인이 임의로 선정한 용어도 있다. 이러한 용어에 대해서는 본 명세서에서 정의된 의미로 해석될 수 있으며, 구체적인 용어 정의가 없으면 본 명세서의 전반적인 내용 및 당해 기술 분야의 통상적인 기술 상식을 토대로 해석될 수도 있다.

[0027] 또한, 본 명세서에 첨부된 각 도면에 기재된 동일한 참조번호 또는 부호는 실질적으로 동일한 기능을 수행하는 부품 또는 구성요소를 나타낸다. 설명 및 이해의 편의를 위해서 서로 다른 실시 예들에서도 동일한 참조번호 또는 부호를 사용하여 설명한다. 즉, 복수의 도면에서 동일한 참조 번호를 가지는 구성요소를 모두 도시되어 있다고 하더라도, 복수의 도면들이 하나의 실시 예를 의미하는 것은 아니다.

[0028] 또한, 본 명세서 및 청구범위에서는 구성요소들 간의 구별을 위하여 "제1", "제2" 등과 같이 서수를 포함하는

용어가 사용될 수 있다. 이러한 서수는 동일 또는 유사한 구성요소들을 서로 구별하기 위하여 사용하는 것이며 이러한 서수 사용으로 인하여 용어의 의미가 한정 해석되어서는 안 된다. 일 예로, 이러한 서수와 결합된 구성요소는 그 숫자에 의해 사용 순서나 배치 순서 등이 제한되어서는 안 된다. 필요에 따라서는, 각 서수들은 서로 교체되어 사용될 수도 있다.

[0029] 본 명세서에서 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.

[0030] 본 개시의 실시 예에서 "모듈", "유닛", "부(part)" 등과 같은 용어는 적어도 하나의 기능이나 동작을 수행하는 구성요소를 지칭하기 위한 용어이며, 이러한 구성요소는 하드웨어 또는 소프트웨어로 구현되거나 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다. 또한, 복수의 "모듈", "유닛", "부(part)" 등은 각각이 개별적인 특정한 하드웨어로 구현될 필요가 있는 경우를 제외하고는, 적어도 하나의 모듈이나 칩으로 일체화되어 적어도 하나의 프로세서로 구현될 수 있다.

[0031] 또한, 본 개시의 실시 예에서, 어떤 부분이 다른 부분과 연결되어 있다고 할 때, 이는 직접적인 연결뿐 아니라, 다른 매체를 통한 간접적인 연결의 경우도 포함한다. 또한, 어떤 부분이 어떤 구성요소를 포함한다는 의미는, 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.

[0032] 도 2는 본 개시의 일 실시 예에 따라 단어 사전을 포함하는 전자 장치의 구성을 설명하기 위한 블록도이다.

[0033] 도 2를 참조하면, 전자 장치(100)는 메모리(110) 및 프로세서(120)를 포함할 수 있다.

[0034] 전자 장치(100)는 서버에 해당할 수 있으며, 스마트폰, 스마트 스피커, 데스크탑 PC, 노트북 PC, 태블릿 PC 등 다양한 단말 기기에 해당할 수 있다.

[0035] 또한, 전자 장치(100)는 하나 이상의 컴퓨터를 포함하는 시스템으로 구현될 수 있다. 예를 들어, 전자 장치(100)는 적어도 하나의 병원, 의료기관, 또는 공공기관에서 운영하는 전자 의무 기록을 관리하는 시스템에 해당할 수 있으나, 이에 한정되지 않는다.

[0036] 메모리(110)는 전자 장치(100)의 구성요소들의 전반적인 동작을 제어하기 위한 운영체제(OS: Operating System) 및 전자 장치(100)의 구성요소와 관련된 적어도 하나의 인스트럭션 또는 데이터를 저장하기 위한 구성이다.

[0037] 메모리(110)는 ROM, 플래시 메모리 등의 비휘발성 메모리를 포함할 수 있으며, DRAM 등으로 구성된 휘발성 메모리를 포함할 수 있다. 또한, 메모리(110)는 하드 디스크, SSD(Solid state drive) 등을 포함할 수도 있다.

[0038] 도 2를 참조하면, 메모리(110)는, 적어도 하나의 단어 사전(111), 적어도 하나의 딥러닝 모델(112)을 포함할 수 있다.

[0039] 단어 사전(111)은 서로 다른 언어로 된 단어들 간의 매칭 관계에 대한 정보를 포함할 수 있다. 예를 들어, 단어 사전(111)은 제1 언어(ex. 한글)로 된 단어와, 해당 단어와 매칭되는 제2 언어(ex. 영어)의 단어를 각각 포함할 수 있다.

[0040] 단어 사전(111)은, 전문 분야 별로 각각 구비될 수 있다. 예를 들어, 메모리(110)는 뇌질환과 관련된 단어들이 포함된 단어 사전, 심장질환과 관련된 단어들이 포함된 단어 사전 등을 각각 별도로 포함할 수 있다.

[0041] 딥러닝 모델(112)은 특정 언어의 텍스트를 처리하기 위한 네트워크 모델에 해당할 수 있다. 예를 들어, 단어 사전(111)이 한글 단어 및 영어 단어 간의 매칭 관계를 포함하는 경우, 딥러닝 모델(112)은 영어로 구성된 텍스트를 처리하기 위한 네트워크 모델에 해당할 수 있다.

[0042] 딥러닝 모델(112)은, 텍스트의 분류, 텍스트의 요약, 텍스트의 인식/변환, 텍스트와 관련된 대화 생성 등 다양한 목적에 따라 텍스트를 처리하도록 설계된 것일 수 있다.

[0043] 딥러닝 모델(112)은 다양한 학습 알고리즘을 통해 전자 장치(100) 또는 별도의 서버/시스템을 통해 학습된 것일 수 있다.

[0044] 학습 알고리즘의 예로는, 지도형 학습(supervised learning), 비지도형 학습(unsupervised learning), 준지도형 학습(semi-supervised learning) 또는 강화 학습(reinforcement learning)이 있으며, 본 개시에서의 학습

알고리즘은 명시한 경우를 제외하고 전술한 예에 한정되지 않는다.

[0045] 딥러닝 모델(112)은 신경망(Neural Network)을 기반으로 하는 네트워크 모델(신경망 모델)이며, 가중치를 가지는 복수의 네트워크 노드들을 포함할 수 있다. 복수의 네트워크 노드들은 서로 다른 레이어의 노드 간 가중치를 기반으로 연결 관계를 형성할 수 있다.

[0046] 프로세서(120)는 전자 장치(100)를 전반적으로 제어하기 위한 구성이다. 구체적으로, 프로세서(120)는 메모리(110)와 연결되는 한편 메모리(110)에 저장된 적어도 하나의 인스트럭션을 실행함으로써 본 개시의 다양한 실시예들에 따른 동작을 수행할 수 있다.

[0047] 프로세서(120)는 CPU, AP, DSP(Digital Signal Processor) 등과 같은 범용 프로세서, GPU, VPU(Vision Processing Unit) 등과 같은 그래픽 전용 프로세서 또는 NPU와 같은 인공지능 전용 프로세서 등을 포함할 수 있다. 인공지능 전용 프로세서는, 특정 인공지능 모델의 훈련 내지는 이용에 특화된 하드웨어 구조로 설계될 수 있다.

[0048] 도 3은 본 개시의 일 실시 예에 따른 전자 장치의 동작을 설명하기 위한 흐름도이다.

[0049] 도 3을 참조하면, 전자 장치(100)는 전체 텍스트 내에서 제1 언어에 해당하는 제1 텍스트 및 상기 제2 언어에 해당하는 제2 텍스트를 식별할 수 있다(S310).

[0050] 일 예로, 전체 텍스트는 한글과 영어가 혼재된 텍스트일 수 있고, 제1 언어는 한글, 제2 언어는 영어일 수 있으나, 이에 한정되지 않고 다양한 종류의 언어가 식별될 수 있다.

[0051] 그리고, 전자 장치(100)는 상술한 단어 사전(111)을 기반으로 제1 텍스트를 제2 언어에 해당하는 제3 텍스트로 변환할 수 있다(S320).

[0052] 구체적으로, 전자 장치(100)는 제1 텍스트 내에서 단어 사전(111)에 포함된 적어도 하나의 제1 키워드를 식별할 수 있다. 그리고, 단어 사전(111)을 기반으로, 전자 장치(100)는 제1 텍스트에 포함된 제1 키워드를 제2 언어에 해당하는 제2 키워드로 변환할 수 있다.

[0053] 이때, 제2 키워드가 하나의 개체로 설정된 상태에서, 전자 장치(100)는 제1 텍스트를 제2 언어에 해당하는 제3 텍스트로 변환할 수 있다.

[0054] 관련하여, 도 4는 본 개시의 일 실시 예에 따른 전자 장치가 단어 사전을 활용하여 한영 변환을 수행하는 동작을 설명하기 위한 도면이다.

[0055] 도 4를 참조하면, 단어 사전(405)은 뇌질환과 관련된 단어들을 포함하며, 구체적으로 한글 단어와 영어 단어 간의 매칭 관계를 포함한다.

[0056] 리스트(410)는 한글에 해당하는 텍스트들을 포함하고, 리스트(410')는 리스트(410)의 텍스트들이 영어로 변환된 결과를 포함한다.

[0057] 도 4를 참조하면, 리스트(410)에 포함된 각 텍스트들은 단어 사전(405)에 정의된 매칭 관계에 따라 리스트(410')에 포함된 각 텍스트로 변환될 수 있다.

[0058] 예를 들어, 상술한 제1 텍스트에 “활력증상”이라는 키워드가 포함된 경우, 전자 장치(100)는 단어 사전(405)을 활용하여 “활력증상”을 “vital sign”으로 변환할 수 있다.

[0059] 여기서, 전자 장치(100)는 “활력증상”이 “vital sign”으로 대체된 제1 텍스트에서 “vital sign”을 하나의 개체로 설정한 상태로 제1 텍스트를 영어로 변환할 수 있다. 이렇듯, 단어 사전(405)에 따라 확보된 “vital sign” 자체가 하나의 개체로 고정된 상태로 번역이 수행되는 경우, 번역의 정확도가 상승할 수 있다.

[0060] 이때, 전자 장치(100)는 상술한 단어 사전(111, 405) 외에 한영 번역을 수행하기 위한 적어도 하나의 인공지능 모델을 활용할 수도 있다(ex. 종래의 기계 번역 내지는 통계기반 번역).

[0061] 상술한 실시 예에 따라 제1 언어의 제1 텍스트가 제3 텍스트(: 제2 언어)로 변환되면, 전자 장치(100)는 제2 텍스트 및 제3 텍스트를 포함하는 통합 텍스트(: 제2 언어)를 자연어 처리하여 하나 이상의 벡터를 획득할 수 있다(S330).

[0062] 구체적으로, 전자 장치(100)는 통합 텍스트에 대하여 tokenization, 품사 태깅(POS Tagging)을 수행한 이후, word embedding을 수행하여 하나 이상의 벡터를 획득할 수 있다.

[0063] Word embedding을 통한 벡터 변환 이전에, 전자 장치(100)는 통합 텍스트에 대하여 Stop Word Removal, N-gram 기반의 단어 선택 내지는 단어 간 결합 판단 등을 수행할 수도 있다.

[0064] 그리고, 상술한 바와 같이 통합 텍스트가 변환된 벡터가 획득되면, 전자 장치(100)는 획득된 벡터를 기초로 딥러닝 모델(112)을 훈련시킬 수 있다(S340).

[0065] 제1 언어 및 제2 언어가 혼재하는 전체 텍스트에 대해서 일괄적으로 벡터를 획득하는 경우와 비교했을 때, 본 개시에 따른 전자 장치의 제어 방법(도 3)은, 단어 변환의 전문성이 담보된 단어 사전(111)을 활용하여 하나의 언어로 통일한 뒤 벡터를 획득한다는 점에서, 벡터 변환 전 잘못된 자연어 처리의 가능성이 줄어든다는 장점이 있다. 이러한 장점은, 텍스트가 변환된 벡터를 통해 훈련되는 딥러닝 모델의 훈련 성과의 증대로 이어질 수 있다.

[0066] 한편, 일 실시 예로, 전자 장치(100)는 상술한 통합 텍스트의 벡터를 통해 훈련된 딥러닝 모델(112)의 훈련 전후 정확도를 비교하여 훈련의 효과를 검증할 수 있다

[0067] 예를 들어, 전자 장치(100)는 딥러닝 모델(112)의 검증을 위해 구비된 다양한 텍스트의 벡터를 딥러닝 모델(112)에 입력할 수 있으며, 딥러닝 모델(112)의 출력에 따라 정확도를 식별할 수 있다.

[0068] 만약 상술한 통합 텍스트의 벡터를 통해 훈련이 수행된 결과 딥러닝 모델(112)의 정확도가 낮아진 경우, 전자 장치(100)는 단어 사전(111)을 업데이트할 수 있다.

[0069] 구체적으로, 제1 언어에 해당하는 제1 텍스트에 포함된 제1 키워드(ex. 활력 증상)가 단어 사전(111)에 따라 제2 언어의 제2 키워드(ex. vital sign)로 변환된 결과, 통합 텍스트(: 제2 언어)가 획득된 경우를 가정한다.

[0070] 여기서, 통합 텍스트의 벡터에 따라 훈련된 딥러닝 모델(112)의 정확도가 오히려 낮아지는 경우, 전자 장치(100)는 단어 사전(111) 내에서 제1 키워드에 매칭되는 제2 언어의 키워드를 상술한 제2 키워드가 아닌 다른 키워드로 변경할 수 있다.

[0071] 이때, 제1 키워드에 매칭되는 키워드는, 한 명 이상의 전문가의 사용자 입력에 따라 지정될 수 있다.

[0072] 예를 들어, 전자 장치(100)는 제1 키워드에 대한 정보를 복수의 전문가 단말로 전송할 수 있다. 그리고, 전자 장치(100)는 제1 키워드가 제2 언어로 변환된 키워드를 각 전문가 단말로부터 수신할 수 있다. 이때, 전자 장치(100)는 가장 많은 수의 전문가 단말로부터 수신된 키워드에 따라 단어 사전(111)을 업데이트할 수 있다.

[0073] 한편, 단어 사전(111)에 따라, 제1 키워드(: 제1 언어)에 매칭되는 제2 키워드(: 제2 언어)가 복수 개 매칭되어 있을 수도 있다.

[0074] 일 예로, 상술한 S320 단계에서, 제1 언어에 해당하는 제1 키워드가 식별되고, 단어 사전(111)에 따라 제1 키워드에 매칭되는 복수의 제2 키워드가 식별된 경우를 가정한다.

[0075] 이 경우, 전자 장치(100)는 복수의 제2 키워드 각각을 독립적으로 반영하여 복수의 제3 텍스트를 획득할 수 있다. 여기서, 복수의 제3 텍스트 각각이 제2 언어에 해당하는 제2 텍스트와 결합되어 복수의 통합 텍스트가 획득될 수 있다.

[0076] 이 경우, 서로 다른 제3 텍스트를 포함하는 복수의 통합 텍스트가 각각 별도로 벡터로 변환될 수 있으며, 전자 장치(100)는 각 통합 텍스트가 변환된 벡터를 딥러닝 모델(112)에 독립적으로 입력할 수 있다.

[0077] 여기서, 각 통합 텍스트가 변환된 벡터에 대한 딥러닝 모델(112)의 출력 간의 차이가 임계치 미만인 경우, 단어 사전(111)은 기존과 동일하게 유지될 수 있다. 즉, 단어 사전(111) 내에서, 제1 키워드에 매칭되는 것으로 설정된 복수의 제2 키워드가 그대로 유지될 수 있다.

[0078] 반면, 각 통합 텍스트가 변환된 벡터에 대한 딥러닝 모델(112)의 출력 간의 차이가 임계치 이상인 경우라면, 단어 사전(111)은 업데이트될 수 있다.

[0079] 구체적인 예로, 전자 장치(100)는, 앞서 서로 다른 제2 키워드에 따라 도출된 복수의 통합 텍스트 중, 정답에 가장 가까운 딥러닝 모델(112)의 출력을 유발한 통합 텍스트를 선택할 수 있다.

[0080] 이때, 전자 장치(100)는, 선택된 통합 텍스트를 제외한 나머지 통합 텍스트 중, 정답에 가장 가까운 상술한 출력과 임계치 이상 차이가 나는 출력이 기인하는 (저품질의) 통합 텍스트를 식별할 수 있다. 그리고, 전자 장치(100)는 식별된 (저품질의) 통합 텍스트와 관련된 제2 키워드를, 단어 사전(111) 내 제1 키워드와 매칭되는 키

워드 리스트에서 삭제할 수 있다.

[0081] 한편, 일 실시 예로, 메모리(110)는 전문 분야 별로 구분된 복수의 단어 사전을 포함할 수 있다.

[0082] 예를 들어, 메모리(110)는 제1 전문 분야에 대한 제1 단어 사전, 제2 전문 분야에 대한 제2 단어 사전, 및 제3 전문 분야에 대한 제3 단어 사전을 포함할 수 있다.

[0083] 구체적인 예로, 제1 전문 분야는 뇌질환, 제2 전문 분야는 심장질환, 제3 전문 분야는 정신질환에 해당할 수 있다. 다만, 전문 분야는 이 밖에도 다양한 의료 전문 분야를 포함할 수 있으며, 의료 전문 분야가 아닌 전혀 다른 분야의 전문 분야 역시 포함될 수 있음은 물론이다.

[0084] 이렇듯, 다양한 전문 분야에 해당하는 단어 사전이 기저장된 경우, 전자 장치(100)는 전체 텍스트(: 제1 언어와 제2 언어가 혼재)가 속하는 전문 분야에 맞는 단어 사전을 우선적으로 활용하여 통합 텍스트를 획득할 수 있다. 다만, 상황에 따라서는 다른 전문 분야의 단어 사전도 활용될 수 있는 바, 관련 실시 예는 이하 도 5를 통해 설명한다.

[0085] 도 5는 본 개시의 일 실시 예에 따른 전자 장치가 전문 분야 별로 구분된 단어 사전을 선택적으로 활용하는 동작을 설명하기 위한 알고리즘이다.

[0086] 도 5를 참조하면, 전자 장치(100)는 전체 텍스트(: 제1 언어와 제2 언어 혼재)의 전문 분야를 설정할 수 있다(S510). 이때, 전문 분야는 사용자 입력에 따라 설정될 수도 있고, 전체 텍스트와 함께 저장된 메타데이터에 따라 기설정된 것일 수도 있다.

[0087] 또는, 전자 장치(100)는 전문 분야를 선택/분류하도록 훈련된 적어도 하나의 분류기 모델에 상술한 전체 텍스트를 입력하여, 전문 분야를 식별할 수도 있다.

[0088] 그리고, 전자 장치(100)는 전체 텍스트에 포함된 제1 언어의 제1 텍스트를 식별할 수 있다(S520). 본 과정은 상술한 S310 과정에 포함될 수 있다.

[0089] 여기서, 전자 장치(100)는 제1 텍스트를 제2 언어의 제3 텍스트로 변환할 수 있으며, 본 과정에서 앞서 설정된 전문 분야에 맞는 단어 사전이 선택되어 활용될 수 있다(S530).

[0090] 예를 들어, 전체 텍스트가 제1 전문 분야로 설정된 경우, 전자 장치(100)는 제1 전문 분야에 맞는 제1 단어 사전을 선택할 수 있으며, 선택된 제1 단어 사전을 활용하여 제1 텍스트를 제2 언어의 제3 텍스트로 변환할 수 있다.

[0091] 구체적으로, 제1 텍스트에 제1 단어 사전의 키워드가 존재하는 경우(S540 - Y), 전자 장치(100)는 제1 단어 사전을 활용하여 해당 키워드를 변환함으로써 제1 텍스트를 제2 언어로 변환할 수 있다(S550).

[0092] 반면, 선택된 제1 단어 사전의 키워드가 제1 텍스트 내에 존재하지 않을 수도 있다(S540 - N).

[0093] 이 경우, 전자 장치(100)는 다른 전문 분야에 해당하는 적어도 하나의 단어 사전을 선택할 수 있다(S560).

[0094] 그리고, 전자 장치(100)는 선택된 단어 사전을 바탕으로, 제1 텍스트를 제2 언어로 변환할 수 있다(S550).

[0095] 일 예로, 전자 장치(100)는 제1 텍스트에 포함된 키워드들 중 적어도 하나를 포함하는 단어 사전을 선택할 수 있다. 구체적인 예로, 제1 텍스트에 포함된 키워드가 제2 전문 분야에 해당하는 제2 단어 사전에 포함된 경우, 전자 장치(100)는 제2 단어 사전을 활용하여 제1 텍스트를 제2 언어로 변환함으로써 통합 텍스트를 획득할 수 있다.

[0096] 또한, 일 예로, 전자 장치(100)는 전체 텍스트에 대하여 설정된 제1 전문 분야에 대한 연관도가 높은 적어도 하나의 전문 분야를 선택하고, 선택된 전문 분야의 단어 사전을 선택할 수도 있다.

[0097] 전문 분야 간의 연관도는, 각 전문 분양에 매칭되는 단어 사전의 내용 간의 유사도에 따라 산출될 수 있다.

[0098] 구체적으로, 전자 장치(100)는 제1 전문 분야의 제1 단어 사전에 포함된 키워드들 각각을 제2 전문 분야의 제2 단어 사전에 포함된 키워드들 각각과 비교함으로써 키워드 간 유사도의 평균(또는 기타 통계치)을 산출할 수 있다.

[0099] 그리고, 전자 장치(100)는 키워드 간 유사도의 평균에 따라 제1 전문 분야 및 제2 전문 분야 간의 연관도를 설정할 수 있다. 예를 들어, 키워드 간 유사도의 평균이 클수록 제1 전문 분야 및 제2 전문 분야 간의 연관도도 더 크게 산출될 수 있다.

[0100] 일 예로, 제1 전문 분야와 제2 전문 분야 간의 연관도가, 제1 전문 분야와 제3 전문 분야 간의 연관도보다 큰 경우를 가정한다.

[0101] 여기서, 상술한 바와 같이 전체 텍스트에 포함된 제1 텍스트의 키워드가 앞서 선택된 제1 전문 분야의 제1 단어 사전에 존재하지 않는 경우, 전자 장치(100)는 제1 전문 분야와 연관도가 가장 높은 제2 전문 분야의 제2 단어 사전을 선택할 수 있다.

[0102] 이때, 전자 장치(100)는 선택된 제2 단어 사전을 활용하여 제1 텍스트를 제2 언어로 변환할 수 있다.

[0103] 다만, 제2 단어 사전 역시 제1 텍스트 내 키워드를 포함하지 않는 경우도 발생할 수 있다. 이 경우, 전자 장치(100)는 제1 전문 분야와의 연관도가 두 번째로 높은 제3 전문 분야의 제3 단어 사전을 선택할 수 있다. 그리고, 전자 장치(100)는 제3 단어 사전을 활용하여 제1 텍스트를 제2 언어로 변환할 수 있다.

[0104] 한편, 상술하였듯 전체 텍스트의 전문 분야에 따라 (최초로) 선택된 단어 사전(ex. 제1 단어 사전)을 통해 키워드 변환이 수행되지 않고 다른 전문 분야의 단어 사전을 통해 키워드 변환이 수행된 경우(ex. S540 - N -> S560 -> S550), 전자 장치(100)는 최종적으로 선택된 단어 사전(ex. 제2 단어 사전 및/또는 제3 단어 사전)을 기초로 최초로 선택된 단어 사전(ex. 제1 단어 사전)을 업데이트할 수 있다.

[0105] 구체적으로, 최종적으로 선택된 다른 전문 분야의 단어 사전(ex. 제2 단어 사전 및/또는 제3 단어 사전) 내에서, 전자 장치(100)는 상술한 제1 텍스트에 포함된 키워드와 매칭되는 제2 언어의 키워드를 식별할 수 있다.

[0106] 그리고, 전자 장치(100)는 상술한 제1 텍스트에 포함된 키워드 및 (다른 단어 사전을 통해 매칭된) 제2 언어의 키워드 간의 매칭 관계를, 최초로 선택된 단어 사전 내에 추가할 수 있다.

[0107] 그 결과, 전체 텍스트의 전문 분야(ex. 제1 전문 분야)에 매칭되는 단어 사전(ex. 제1 단어 사전)이 지속적으로 보완될 수 있다.

[0108] 한편, 이상에서 설명된 다양한 실시 예들은 서로 저촉되거나 모순되지 않는 한 두 개 이상의 실시 예가 서로 결합되어 구현될 수 있다.

[0109] 한편, 이상에서 설명된 다양한 실시 예들은 소프트웨어(software), 하드웨어(hardware) 또는 이들의 조합된 것을 이용하여 컴퓨터(computer) 또는 이와 유사한 장치로 읽을 수 있는 기록 매체 내에서 구현될 수 있다.

[0110] 하드웨어적인 구현에 의하면, 본 개시에서 설명되는 실시 예들은 ASICs(Application Specific Integrated Circuits), DSPs(digital signal processors), DSPDs(digital signal processing devices), PLDs(programmable logic devices), FPGAs(field programmable gate arrays), 프로세서(processors), 제어기(controllers), 마이크로 컨트롤러(micro-controllers), 마이크로 프로세서(microprocessors), 기타 기능 수행을 위한 전기적인 유닛(unit) 중 적어도 하나를 이용하여 구현될 수 있다.

[0111] 일부의 경우에 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. 상술한 소프트웨어 모듈들 각각은 본 명세서에서 설명되는 하나 이상의 기능 및 작동을 수행할 수 있다.

[0112] 한편, 상술한 본 개시의 다양한 실시 예들에 따른 전자 장치(100)에서의 처리동작을 수행하기 위한 컴퓨터 명령어(computer instructions) 또는 컴퓨터 프로그램은 비일시적 컴퓨터 판독 가능 매체(non-transitory computer-readable medium)에 저장될 수 있다. 이러한 비일시적 컴퓨터 판독 가능 매체에 저장된 컴퓨터 명령어 또는 컴퓨터 프로그램은 특정 기기의 프로세서에 의해 실행되었을 때 상술한 다양한 실시 예에 따른 전자 장치(100)에서의 처리 동작을 상술한 특정 기기가 수행하도록 한다.

[0113] 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.

[0114] 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 가진 자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 개시의 기술

적 사상이나 전망으로부터 개별적으로 이해되어서는 안될 것이다.

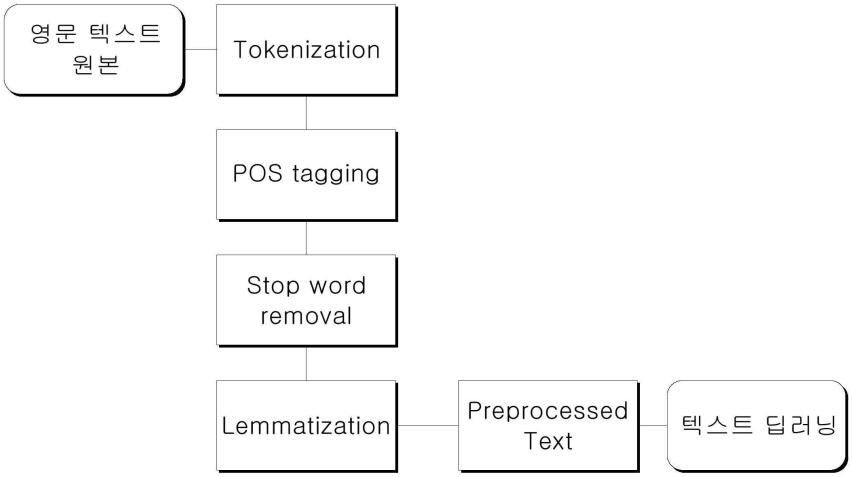
부호의 설명

[0115] 100: 전자 장치 110: 메모리

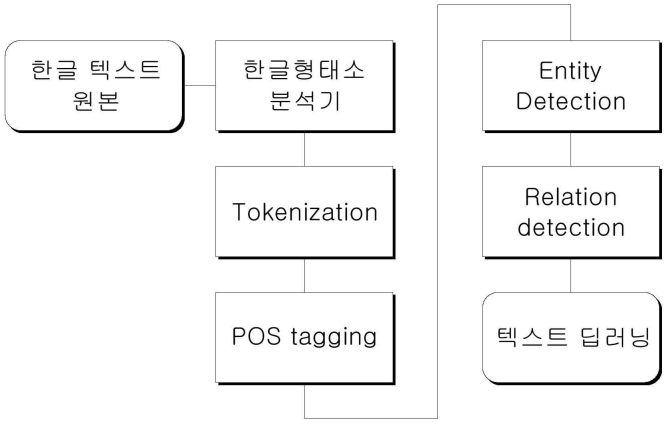
111: 단어 사전 120: 프로세서

도면

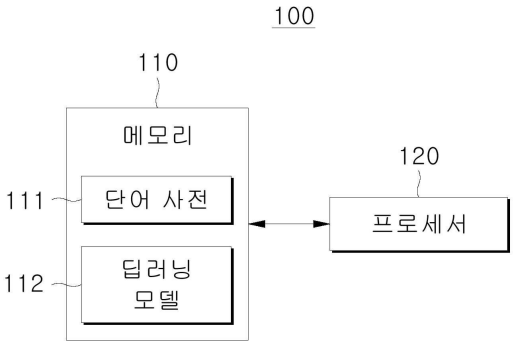
도면1a



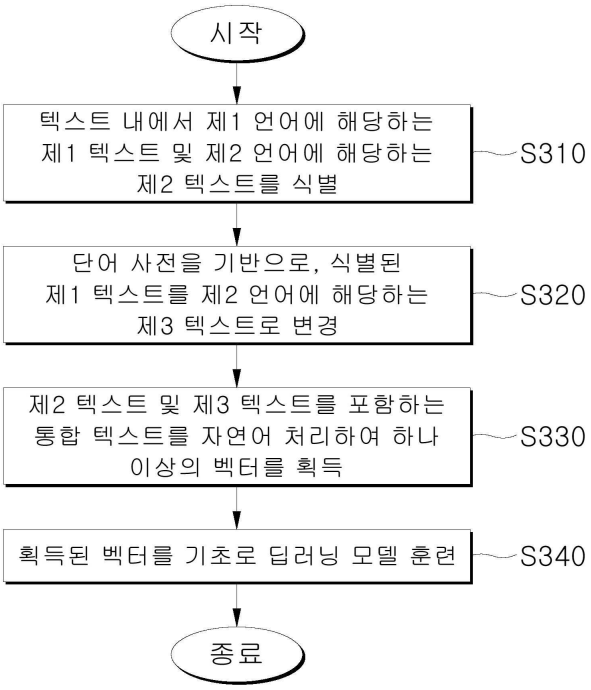
도면1b



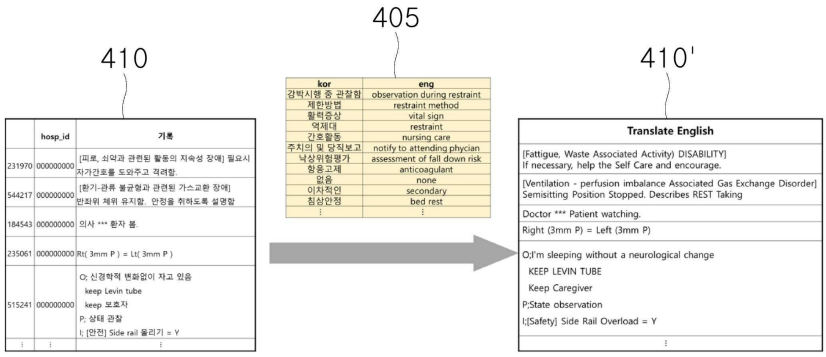
도면2



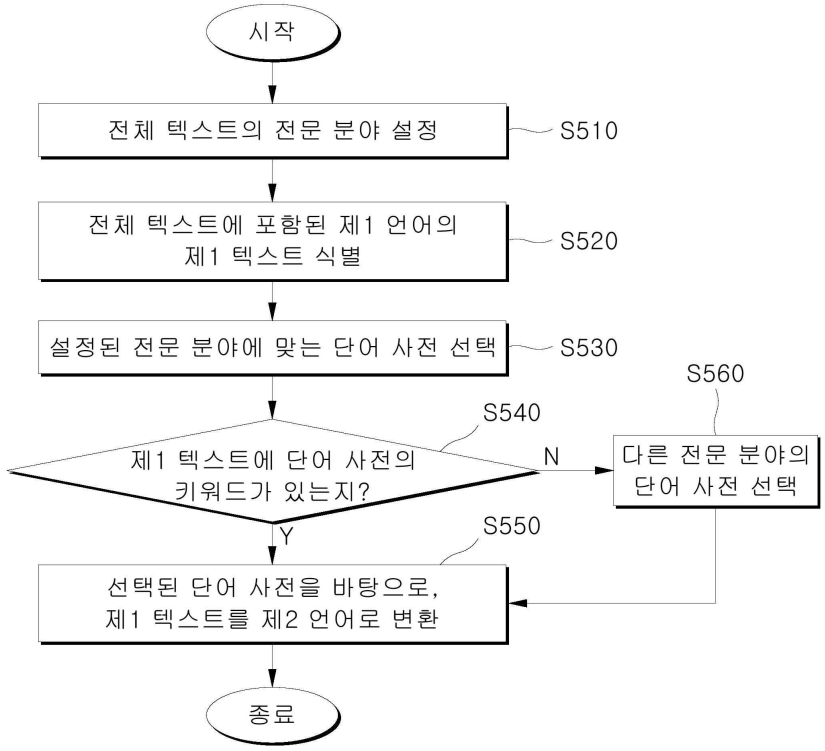
도면3


시작
텍스트 내에서 제1 언어에 해당하는 제1 텍스트 및 제2 언어에 해당하는 제2 텍스트를 식별
S310
단어 사전을 기반으로, 식별된 제1 텍스트를 제2 언어에 해당하는 제3 텍스트로 변경
S320
제2 텍스트 및 제3 텍스트를 포함하는 통합 텍스트를 자연어 처리하여 하나 이상의 벡터를 획득
S330
획득된 벡터를 기초로 딥러닝 모델 훈련
S340
종료

도면4


405
410
410'
kor
eng
강박시행 중 관찰함
observation during restraint
제한방법
restraint method
활력증상
vital sign
억제대
restraint
간호활동
nursing care
주치의 및 당직보고
notify to attending phycian
낙상위험평가
assessment of fall down risk
항응고제
anticoagulant
없음
none
이차적인
secondary
침상안정
bed rest
hosp_id
기록
231970 000000000 [피로, 쇠약과 관련된 활동의 지속성 장애] 필요시 자가간호를 도와주고 격려함.
544217 000000000 [환기-관류 불균형과 관련된 가스교환 장애] 반좌위 체위 유지함. 안정을 취하도록 설명함
184543 000000000 의사 *** 환자 봄.
235061 000000000 Rt(3mm P) = Lt(3mm P)
515241 000000000 O; 신경학적 변화없이 자고 있음
keep Levin tube
keep 보호자
P; 상태 관찰
I; [안전] Side rail 올리기 = Y
Translate English
[Fattigue, Waste Associated Activity) DISABILITY]
If necessary, help the Self Care and encourage.
[Ventilation - perfusion imbalance Associated Gas Exchange Disorder]
Semisitting Position Stopped. Describes REST Taking
Doctor *** Patient watching.
Right (3mm P) = Left (3mm P)
O;I'm sleeping without a neurological change
KEEP LEVIN TUBE
Keep Caregiver
P;State observation
I;[Safety] Side Rail Overload = Y

도면5


시작
전체 텍스트의 전문 분야 설정
S510
전체 텍스트에 포함된 제1 언어의 제1 텍스트 식별
S520
설정된 전문 분야에 맞는 단어 사전 선택
S530
S540
제1 텍스트에 단어 사전의 키워드가 있는지?
N
S560
다른 전문 분야의 단어 사전 선택
Y
S550
선택된 단어 사전을 바탕으로, 제1 텍스트를 제2 언어로 변환
종료