



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0138746
(43) 공개일자 2024년09월20일

(51) 국제특허분류(Int. Cl.)

G06Q 10/06 (2012.01) G06F 40/20 (2020.01)
G06F 40/30 (2020.01) G06N 3/00 (2022.01)
G06T 19/00 (2011.01) G06V 30/41 (2022.01)
G06V 30/42 (2022.01)

(52) CPC특허분류

G06Q 10/06375 (2013.01)
G06F 40/20 (2022.01)

(21) 출원번호 10-2023-0032363

(22) 출원일자 2023년03월13일

심사청구일자 없음

(71) 출원인

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

청주대학교 산학협력단

충청북도 청주시 청원구 대성로 298 (내덕동)

(72) 발명자

김영호

서울특별시 성북구 동소문로 211, A동 602호

강지호

서울특별시 용산구 효창원로37가길 7, 203호

박상성

충청북도 청주시 청원구 공향로 159-12, 526호

(74) 대리인

김준석, 박민욱

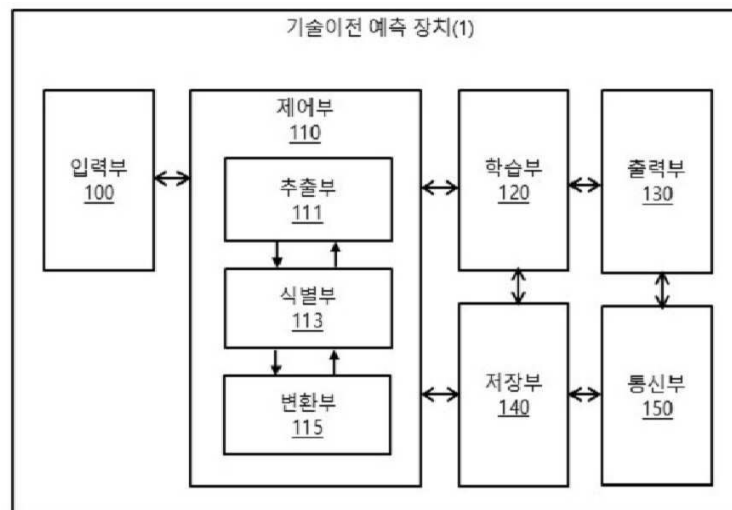
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 기술이전 예측 장치 및 그 방법

(57) 요약

본 발명은 기술이전 예측 장치 및 그 방법에 관한 것으로, 미리 수집된 특허 데이터를 입력받는 단계; 특허 데이터로부터 정량정보 및 텍스트 정보를 추출하는 단계; 텍스트 정보로부터 세부기술을 식별하는 단계; 정량정보 및 세부기술을 결합하여 이미지로 변환하는 단계; 변환된 이미지를 증강하는 단계; 및 증강된 이미지를 예측 모델에 학습하여 기술이전을 예측하는 단계;를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06F 40/30 (2020.01)
G06N 3/006 (2023.01)
G06T 19/006 (2013.01)
G06V 30/41 (2023.08)
G06V 30/42 (2022.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1345353640
과제번호	2022R1I1A1A01062652
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	이공학학술연구기반구축
연구과제명	R&D 전략 수립 지원을 위한 기술문서 이벤트 분석 시스템 개발
기 여 율	100/100
과제수행기관명	고려대학교
연구기간	2023.03.01 ~ 2024.02.29

명세서

청구범위

청구항 1

미리 수집된 특허 데이터를 입력받는 단계;
 상기 특허 데이터로부터 정량정보 및 텍스트 정보를 추출하는 단계;
 상기 텍스트 정보로부터 세부기술을 식별하는 단계;
 상기 정량정보 및 상기 세부기술을 결합하여 이미지로 변환하는 단계;
 상기 변환된 이미지를 증강하는 단계; 및
 상기 증강된 이미지를 예측 모델에 학습하여 기술이전을 예측하는 단계;
 를 포함하는 기술이전 예측 방법.

청구항 2

제1항에 있어서,
 상기 세부기술을 식별하는 단계는,
 사전 훈련된 BERT 모델을 통해 상기 텍스트 정보를 임베딩하는 단계;
 K-Means 군집화를 통해 복수의 기술 군집을 형성하는 단계; 및
 상기 기술 군집 별로 가장 많이 출현된 단어를 통해 세부기술을 식별하는 단계;
 를 포함하는 기술이전 예측 방법.

청구항 3

제1항에 있어서,
 상기 이미지로 변환하는 단계는,
 상기 식별된 세부기술에 대한 레이블을 더미변수로 변환하는 단계;
 상기 더미변수와 상기 정량정보를 결합하여 매트릭스로 도출하는 단계;
 상기 매트릭스에 min-max 정규화를 통해 변수 값을 정규화하는 단계; 및
 tSNE를 적용하여 저차원에 매핑 및 DeepInsight를 적용하여 이미지로 변환하는 단계;
 를 포함하는 기술이전 예측 방법.

청구항 4

제1항에 있어서,
 상기 변환된 이미지를 증강하는 단계;
 상기 변환된 이미지 중에서 소수의 레이블에 해당하는 이미지를 추출하여 반전, 회전을 통해 상기 소수의 레이블에 해당하는 이미지를 증강하는 기술이전 예측 방법.

청구항 5

미리 수집된 특허 데이터를 입력받는 입력부,
 상기 특허 데이터로부터 정량정보 및 텍스트 정보를 추출하는 추출부,

상기 텍스트 정보로부터 세부기술을 식별하는 식별부,
상기 정량정보 및 상기 세부기술을 결합하여 이미지로 변환하는 변환부,
상기 변환된 이미지를 증강하고, 상기 증강된 이미지를 예측 모델에 학습하여 기술이전을 예측하는 학습부를 포함하는 기술이전 예측 장치.

청구항 6

제5항에 있어서,
상기 식별부는,
사전 훈련된 BERT 모델을 통해 상기 텍스트 정보를 임베딩하고, K-Means 군집화를 통해 복수의 기술 군집을 형성하며, 상기 기술 군집 별로 가장 많이 출현된 단어를 통해 세부기술을 식별하는 기술이전 예측 장치.

청구항 7

제5항에 있어서,
상기 변환부는,
상기 식별된 세부기술에 대한 레이블을 더미변수로 변환하고, 상기 더미변수와 상기 정량정보를 결합하여 매트릭스로 도출하며, 상기 매트릭스에 min-max 정규화를 통해 변수 값을 정규화하고, tSNE를 적용하여 저차원에 매핑 및 DeepInsight를 적용하여 이미지로 변환하는 기술이전 예측 장치.

청구항 8

제5항에 있어서,
상기 학습부는,
상기 변환된 이미지 중에서 소수의 레이블에 해당하는 이미지를 추출하여 반전, 회전을 통해 상기 소수의 레이블에 해당하는 이미지를 증강하는 기술이전 예측 장치.

발명의 설명

기술 분야

[0001] 본 발명은 기술이전 예측 장치 및 그 방법에 관한 것이다.

배경 기술

[0002] 기술사업화는 연구개발 투자를 통해 개발된 기술을 직간접적으로 이전하여 사업활동에 활용함으로써 경제적 이익을 창출하는 행위로, R&D 투자와 부가가치 창출을 연결하며, 연구성과의 확산과 국가 및 기업 경쟁력 강화를 위해 수행되는 중요한 경영전략이다.

[0003] 따라서, 많은 국가들이 기술을 상용화하기 위해 노력하고 있다.

[0004] 그러나, 국가적 지원에도 불구하고, 연구개발기관은 기술사업화에 대한 연구자들의 인식과 성과관리에 대한 문제점때문에 연구성과의 사업화에 어려움이 있다.

[0005] 또한, 기술이전은 R&D 결과를 소비자와 산업체에 전파하여 기술사업화 성공률을 높이기 위한 것으로, 기술 이전의 형태에는 라이선스, M&A, 지적 재산권의 이전이 포함된다.

[0006] 이 중 지적재산권의 이전, 즉 특허권의 타인 이전은 상용화를 촉진하는 중요한 기술이전 형태로, 이미 개발된 연구개발 성과의 가치를 인정받아 타인에게 이전하기 때문에 연구 성과를 확산시키고 비용 효율성을 극대화할 수 있다.

[0007] 이로 인해, 많은 국가와 기관들은 R&D를 직접 상용화로 연결하기보다는 기술을 이전하여 상용화하는 것을 선호

하고 있으며, 기술이전은 중복연구의 방지, 연구개발비 절감, 전문적인 사업화의 가능성의 장점이 있다.

[0008] 사업화를 위해서는 상당한 양의 기술이전이 이루어져야 하므로, 정부는 기술이전 성공률을 높이기 위한 지원체계를 구축하여 운영하고 있다.

[0009] 그러나, 기존의 대체 기술과 시장의 불확실성 때문에 성공률은 여전히 높지 않은 문제가 있으며, 취득할 특허의 가치를 분석하는데 많은 시간과 비용이 소비되며, 분석가들의 판단은 주관적이라는 문제점이 있다.

[0010] 또한, 최근에는 대용량의 빅데이터를 기계학습 또는 인공지능에 적용하여 다양한 가치를 창출하고 있다.

[0011] 그러나, 특허의 정보 획득 및 세부기술 식별을 통해 내재되어 있는 정보를 식별하는 것은 어렵고, 예측하고자 하는 특허 관련 이벤트에 대한 모델 학습 데이터의 건수가 부족한 경우 효과적인 모델 구축이 어렵다는 문제가 있다.

선행기술문헌

특허문헌

[0012] (특허문헌 0001) 대한민국등록특허공보 제10-2224684호 ("기계학습 기반의 기술이전 예측모델 생성 장치 및 생성 방법", 청주대학교 산학협력단, 공개일자: 2021.01.11)

발명의 내용

해결하려는 과제

[0013] 본 발명은 특허에 부여된 정량정보를 기반으로 대용량의 특허 데이터를 이미지로 변환하여 기술이전을 예측하는 방법을 제공하는 것을 해결하고자 하는 과제로 한다.

과제의 해결 수단

[0014] 상술한 과제를 해결하기 위하여 기술이전 예측 장치 및 기술이전 예측 방법이 제공된다.

[0015] 기술이전 예측 방법은 미리 수집된 특허 데이터를 입력받는 단계; 상기 특허 데이터로부터 정량정보 및 텍스트 정보를 추출하는 단계; 상기 텍스트 정보로부터 세부기술을 식별하는 단계; 상기 정량정보 및 상기 세부기술을 결합하여 이미지로 변환하는 단계; 상기 변환된 이미지를 증강하는 단계; 및 상기 증강된 이미지를 예측 모델에 학습하여 기술이전을 예측하는 단계;를 포함한다.

[0016] 일 실시예에서, 상기 세부기술을 식별하는 단계는, 사전 훈련된 BERT 모델을 통해 상기 텍스트 정보를 임베딩하는 단계; K-Means 군집화를 통해 복수의 기술 군집을 형성하는 단계; 및 상기 기술 군집 별로 가장 많이 출현된 단어를 통해 세부기술을 식별하는 단계;를 포함할 수 있다.

[0017] 일 실시예에서, 상기 이미지로 변환하는 단계는, 상기 식별된 세부기술에 대한 레이블을 더미변수로 변환하는 단계; 상기 더미변수와 상기 정량정보를 결합하여 매트릭스로 도출하는 단계; 상기 매트릭스에 min-max 정규화를 통해 변수 값을 정규화하는 단계; 및 tSNE를 적용하여 저차원에 매핑 및 DeepInsight를 적용하여 이미지로 변환하는 단계;를 포함할 수 있다.

[0018] 일 실시예에서, 상기 변환된 이미지를 증강하는 단계; 상기 변환된 이미지 중에서 소수의 레이블에 해당하는 이미지를 추출하여 반전, 회전을 통해 상기 소수의 레이블에 해당하는 이미지를 증강할 수 있다.

[0019] 기술이전 예측 장치는 미리 수집된 특허 데이터를 입력받는 입력부, 상기 특허 데이터로부터 정량정보 및 텍스트 정보를 추출하는 추출부, 상기 텍스트 정보로부터 세부기술을 식별하는 식별부, 상기 정량정보 및 상기 세부기술을 결합하여 이미지로 변환하는 변환부, 상기 변환된 이미지를 증강하고, 상기 증강된 이미지를 예측 모델에 학습하여 기술이전을 예측하는 학습부를 포함한다.

[0020] 일 실시예에서, 상기 식별부는, 사전 훈련된 BERT 모델을 통해 상기 텍스트 정보를 임베딩하고, K-Means 군집화를 통해 복수의 기술 군집을 형성하며, 상기 기술 군집 별로 가장 많이 출현된 단어를 통해 세부기술을 식별할 수 있다.

[0021] 일 실시예에서, 상기 변환부는, 상기 식별된 세부기술에 대한 레이블을 더미변수로 변환하고, 상기 더미변수와

상기 정량정보를 결합하여 매트릭스로 도출하며, 상기 매트릭스에 min-max 정규화를 통해 변수 값을 정규화하고, tSNE를 적용하여 저차원에 매핑 및 DeepInsight를 적용하여 이미지로 변환할 수 있다.

- [0022] 일 실시예에서, 상기 학습부는, 상기 변환된 이미지 중에서 소수의 레이블에 해당하는 이미지를 추출하여 반전, 회전을 통해 상기 소수의 레이블에 해당하는 이미지를 증강할 수 있다.

발명의 효과

- [0023] 상술한 기술이전 예측 장치 및 기술이전 예측 방법에 의하면, 특히 데이터를 이미지로 변환하여 예측 모델에 학습함으로써, 효과적인 기술이전 예측을 수행할 수 있도록 하고, 레이블 불균형 문제가 있을 시, 이미지 데이터 증강 기법 적용을 통해 해결함으로써, 자동화된 기술이전 예측에 활용할 수 있다.

도면의 간단한 설명

- [0024] 도 1은 본 발명의 일 실시예에 따른 기술이전 예측 장치를 설명하기 위한 블록도이다.
 도 2는 정량 정보를 설명하기 위한 도면이다.
 도 3는 세부기술을 식별하는 과정을 설명하기 위한 순서도이다.
 도 4는 이미지를 변환하는 과정을 설명하기 위한 순서도이다.
 도 5는 본 발명의 일 실시예에 따른 기술이전 예측 방법을 설명하기 위한 순서도이다.

발명을 실시하기 위한 구체적인 내용

- [0025] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다.
- [0026] 본 명세서에서 사용되는 용어에 대해 간략히 설명하고, 본 발명에 대해 구체적으로 설명하기로 한다.
- [0027] 본 발명에서 사용되는 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나, 이는 당 분야에 종사하는 기술자의 의도 또는 관례, 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 발명의 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서 본 발명에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 발명의 전반에 걸친 내용을 토대로 정의되어야 한다.
- [0028] 명세서 전체에서 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있음을 의미한다. 또한, 명세서에서 사용되는 "부", "모듈", "유닛" 등의 용어는 적어도 하나의 기능 또는 동작을 처리하는 단위를 의미하며, 소프트웨어, FPGA 또는 ASIC과 같은 하드웨어 구성요소, 또는 소프트웨어와 하드웨어의 결합으로 구현될 수 있다. 그렇지만 "부", "모듈", "유닛" 등의 용어가 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. "부", "모듈", "유닛" 등은 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 "부", "모듈", "유닛" 등의 용어는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로 코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들 및 변수들을 포함한다.
- [0029] 아래에서는 첨부한 도면을 참고하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략한다.
- [0030] "제1", "제2" 등과 같이 서수를 포함하는 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 구성 요소들은 용어들에 의해 한정되지는 않는다. 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. "및/또는"이라는 용어는 복수의 관련된 항목

들의 조합 또는 복수의 관련된 항목들 중의 어느 하나의 항목을 포함한다.

- [0032] 이하, 도면을 참조하여 기술이전 예측 장치(1)을 설명하도록 한다.
- [0033] 도 1은 본 발명의 일 실시예에 따른 기술이전 예측 장치를 설명하기 위한 블록도이다.
- [0034] 도 1에서 도시한 바와 같이, 본 발명의 일 실시예에 따른 기술이전 예측 장치(1)는 입력부(100), 제어부(110), 학습부(120), 출력부(130), 저장부(140) 및 통신부(150)를 포함한다.
- [0035] 입력부(100), 제어부(110), 학습부(120), 출력부(130) 및 저장부(140) 중 적어도 둘은 회로, 케이블 또는 무선 통신 네트워크 등을 통해 일방으로 또는 쌍방으로 전기적 신호를 전달하여 명령이나 데이터 등을 전송할 수 있다. 필요에 따라, 입력부(100), 출력부(130) 및 저장부(140) 중 적어도 하나는 생략 가능하다.
- [0036] 입력부(100)는 미리 수집된 특허 데이터를 입력받는다.
- [0037] 예를 들어, 입력부(100)는 미리 수집된 총 1만 5,193건의 미국 특허청에 등록된 AI 기술 관련 특허를 입력받을 수 있고, 특허 정보 사이트로부터 특허 데이터를 입력받을 수 있다.
- [0038] 제어부(110)는 추출부(111), 식별부(113) 및 변환부(115)를 포함할 수 있다.
- [0039] 추출부(111)는 특허 데이터로부터 정량정보 및 텍스트 정보를 추출한다.
- [0040] 도 2는 정량 정보를 설명하기 위한 도면이다.
- [0041] 도 2에서 도시한 바와 같이, 추출부(111)는 app_to_regi(days), all_claim_num, applicant_um, inventor_num, current_owner_num, IPC_num, b_citation_num, f_citation_num, fam_nation_num, fam_doc_num, alone_app_yn, stand_patent_yn, lit_yn을 포함하는 정량정보를 추출할 수 있다.
- [0042] 이때, app_to_regi(day)는 특허출원 후 등록까지 경과한 일수이고, all_claim_num은 권리 범위를 나타내는 클레임 수이며, applicant_num은 출원자의 수이고, inventor_num는 특허에 참여한 발명자의 수이며, current_owner_num은 현재 특허 보유자의 수이고, IPC_num은 해당 특허와 관련된 기술 필드 코드 수이다.
- [0043] 또한, b_citation_num은 인용 수이고, f_citation_num은 피인용 수이며, fam_nation_num 및 fam_doc_num은 패밀리 특허에 포함된 국가 및 특허 수이고, alone_app_yn은 특허 출원이 단독 출원인지 여부를 나타내고, stand_patent_yn은 표준 특허인지 여부를 나타내며, lit_yn은 특허 소송이 제기되었는지 여부를 나타낸다.
- [0044] 또한, 추출부(111)는 기술에 대한 요약, 제목, 청구항을 포함하는 텍스트 정보를 추출할 수 있다.
- [0045] 즉, 추출부(111)는 특허에 부여된 출원 후 등록까지 소요일수, 전체 청구항 수, 출원인 수, 발명자 수, 현재 권리자 수, IPC 수, 인용 수, 피인용 수, 패밀리 국가 수, 패밀리 문헌 수, 단독출원 여부, 표준 특허 여부, 소송 발생 여부를 포함하는 정량 정보와 특허의 제목, 요약, 청구항을 포함하는 텍스트 정보를 추출할 수 있다.
- [0046] 식별부(113)는 텍스트 정보로부터 세부기술을 식별한다.
- [0047] 도 3는 세부기술을 식별하는 과정을 설명하기 위한 순서도이다.
- [0048] 식별부(113)는 사전 학습된 BERT 모델을 통해 텍스트 정보를 임베딩할 수 있다(S110).
- [0049] 이때, 사전 훈련된 BERT는 다양한 거대한 텍스트 데이터 세트를 이용하여 학습될 수 있고, 384차원 공간에 실제 숫자로 분석할 문서 세트인 15,193개의 특허가 저장될 수 있다.
- [0050] 식별부(113)는 K-Means 군집화를 통해 복수의 기술 군집을 형성할 수 있다(S120).
- [0051] 식별부(113)는 임베딩된 텍스트 정보를 K-Means 군집화를 적용하여 복수의 군집을 형성할 수 있다.
- [0052] 식별부(113)는 군집별로 가장 많이 출현하는 단어를 통해 세부기술을 식별할 수 있다(S130).
- [0053] 식별부(113)는 군집별로 가장 많이 출현하는 단어를 도출하여 세부기술을 식별할 수 있다.
- [0054] 변환부(115)는 정량정보 및 세부기술을 결합하여 이미지로 변환한다.
- [0055] 도 4는 이미지를 변환하는 과정을 설명하기 위한 순서도이다.
- [0056] 변환부(115)는 식별된 세부기술에 대한 레이블을 더미변수로 변환할 수 있다(S210).

- [0057] 변환부(115)는 더미변수와 정량정보와 결합하여 매트릭스로 도출할 수 있다(S220).
- [0058] 변환부(115)는 매트릭스에 min-max 정규화를 통해 변수 값을 정규화할 수 있다(S230).
- [0059] 이때, 변환부(115)는 다음의 수학적 식 1을 통해 정규화할 수 있다.

수학적 식 1

$$\frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

- [0060]
- [0061] 여기서, $x_{\min}$ 은 데이터의 최소값이고, $x_{\max}$ 는 데이터의 최대값이다.
- [0062] 변환부(115)는 t-SNE를 적용하여 저차원에 매핑 및 DeepInsight를 적용하여 이미지로 변환할 수 있다(S240).
- [0063] 더욱 자세하게는, 변환부(115)는 정규화된 값에 t-SNE를 적용하여 2차원의 공간에 매핑할 수 있다.
- [0064] 변환부(115)는 매핑된 공간 정보에 Convex hull 알고리즘을 이용하여 모든 데이터를 구성하는 가장 작은 사각형을 추출하고, 작은 사각형을 회전시킨 후에 픽셀(pixel)에 매핑하는 과정을 포함하는 DeepInsight를 적용하여 이미지로 변환할 수 있다.
- [0065] 이때, 이미지는 120X120의 크기와 RGB를 포함하는 3가지 채널을 지닐 수 있고, 변환된 이미지의 수는 특허의 수와 동일할 수 있다.
- [0066] 학습부(120)는 변환된 이미지를 증강할 수 있다.
- [0067] 학습부(120)는 이미지화된 변수 중에서 소수의 레이블인 데이터만을 추출하여 이미지 반전, 회전을 통해 데이터를 증강할 수 있다.
- [0068] 학습부(120)는 증강된 이미지를 예측 모델에 학습하여 기술이전을 예측할 수 있다.
- [0069] 학습부(120)는 레이블별 데이터 개수를 동일하게 하여 예측 모델을 통해 학습하여 기술이전을 예측할 수 있다.
- [0070] 이때, 예측 모델은 이미지 분류 수행이 가능한 기계학습 또는 CNN 등의 인공지능 모델일 수 있다.
- [0071] 출력부(130)는 예측 모델을 통해 기술이전 예측 결과를 출력할 수 있다.
- [0072] 출력부(130)는, 예를 들어, 디스플레이, 프린터 장치, 스피커 장치, 영상 출력 단자, 데이터 입출력 단자 또는 통신 모듈 등을 포함할 수 있으나, 이에 한정되는 것은 아니다.
- [0074] 저장부(140)는 기술이전 예측 장치의 운영 프로그램, 기술이전 예측 프로그램, 영상처리 프로그램 등과 같이 장치의 운영에 필요한 정보를 저장할 수 있으며, 제어부(120)로부터 기술이전 예측 결과를 저장할 수 있다.
- [0075] 저장부(140)는 입력부(100)로부터 입력받은 미리 수집된 특허 데이터를 저장할 수 있다.
- [0076] 저장부(140)는 제어부(110)를 통해 변환된 이미지를 저장할 수 있다.
- [0077] 저장부(140)는 학습부(120)를 통해 학습된 예측 모델을 저장할 수 있다.
- [0078] 저장부(140)는 주기억장치 및 보조기억장치 중 적어도 하나를 포함할 수 있다. 주기억장치는 예를 들어, 롬(ROM) 및/또는 램(RAM)과 같은 반도체 저장 매체를 이용하여 구현된 것일 수 있고, 보조기억장치는, 플래시 메모리 장치(솔리드 스테이트 드라이브(SSD, Solid State Drive) 등), SD(Secure Digital) 카드, 하드 디스크 드라이브(HDD, Hard Disc Drive), 콤팩트 디스크, 디브이디(DVD) 또는 레이저 디스크 등과 같이 데이터를 영구적 또는 반영구적으로 저장 가능한 장치를 기반으로 구현될 수 있다.
- [0080] 통신부(150)는 입력받은 특허 데이터를 제어부(110)로 전송하고, 변환된 이미지를 학습부(12)로 전송할 수 있다.

- [0081] 통신부(150)는 학습부(120)를 통해 학습된 예측 모델을 통해 추출된 기술이전 예측에 대한 결과값을 출력부(130)로 전송할 수 있다.
- [0082] 통신부(150)는 제어부(110)의 추출부(111), 식별부(113) 및 변환부(115)가 서로 연결되어 데이터를 송수신할 수 있다.
- [0083] 통신부(150)는 유무선 통신망을 포함할 수 있고, 유선 통신망은 케이블망이나 공중 전화망(PSTN)과 같은 인터넷망을 포함할 수 있으며, 무선 통신망은 CDMA, WCDMA, GSM, EPC(Evolved Packet Core), LTE(Long Term Evolution), 와이브로(Wibro) 망, 5G 통신망 등을 포함할 수 있다.
- [0084] 통신부(150)는 이에 한정되는 것이 아니며, 향후 구현될 차세대 이동통신 장치의 접속망으로서 가령 클라우드 컴퓨팅 환경하의 클라우드 컴퓨팅망, 5G망 등에 사용될 수 있고, 통신부(150)가 유선 통신망인 경우 통신부(150) 내의 액세스포인트는 전화국의 교환국 등에 접속할 수 있지만, 무선 통신망인 경우에는 통신사에서 운용하는 SGSN 또는 GGSN(Gateway GPRS Support Node)에 접속하여 데이터를 처리하거나, BTS(Base Station Transmission), NodeB, e-NodeB 등의 다양한 중계기에 접속하여 데이터를 처리할 수 있다.
- [0086] 이하에서는, 도면을 참고하여 본 발명의 일 실시예에 따른 기술이전 예측 방법을 설명하도록 한다.
- [0087] 도 5는 본 발명의 일 실시예에 따른 기술이전 예측 방법을 설명하기 위한 순서도이다.
- [0088] 입력부(100)는 특허 데이터를 입력받는다(S310).
- [0089] S310 단계는 도 1의 입력부(100)에서 이미 설명하였는 바, 자세한 설명은 생략한다.
- [0090] 추출부(111)는 특허 데이터로부터 정량정보 및 텍스트 정보를 추출한다(S320).
- [0091] S320 단계는 도 1의 추출부(111)에서 이미 설명하였는 바, 자세한 설명은 생략한다.
- [0092] 식별부(113)는 텍스트 정보로부터 세부기술을 식별한다(S330)
- [0093] S330 단계는 도 1의 식별부(113)에서 이미 설명하였는 바, 자세한 설명은 생략한다.
- [0094] 변환부(115)는 정량정보 및 세부기술을 결합하여 이미지로 변환한다(S340).
- [0095] S340 단계는 도 1의 변환부(115)에서 이미 설명하였는 바, 자세한 설명은 생략한다.
- [0096] 학습부(120)는 변환된 이미지를 증강한다(S350).
- [0097] 학습부(120)는 증강된 이미지를 예측 모델에 학습하여 기술이전을 예측한다(S360).
- [0098] S350 및 S360 단계는 도 1의 학습부(120)에서 이미 설명하였는 바, 자세한 설명은 생략한다.
- [0100] 이상에서 설명한 바와 같이, 본 발명에 따르면, 특허 데이터를 이미지로 변환하여 예측 모델에 학습함으로써, 효과적인 기술이전 예측을 수행할 수 있도록 하고, 레이블 불균형 문제가 있을 시, 이미지 데이터 증강 기법 적용을 통해 해결함으로써, 자동화된 기술이전 예측에 활용할 수 있다.
- [0102] 본원 발명의 실시예들과 관련된 기술 분야에서 통상의 지식을 가진 자는 상기 기재의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로, 개시된 방법들은 한정적인 관점이 아닌 설명적 관점에서 고려되어야 한다. 본 발명의 범위는 발명의 상세한 설명이 아닌 특허청구 범위에 나타나며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

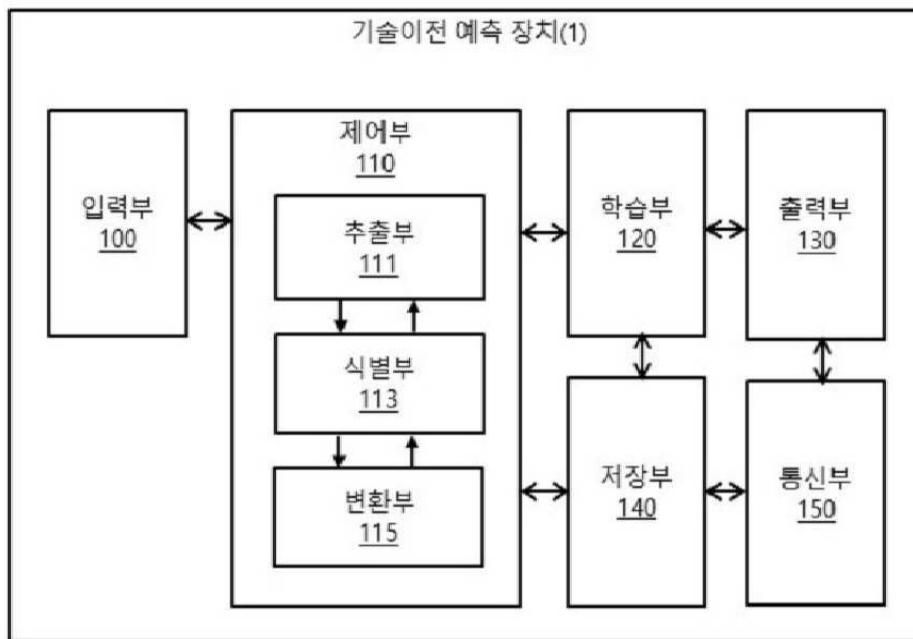
부호의 설명

- [0104] 1: 기술이전 예측 장치
100: 입력부

200: 제어부
 210: 추출부
 220: 식별부
 230: 변환부
 300: 학습부
 400: 출력부
 500: 저장부
 600: 통신부

도면

도면1



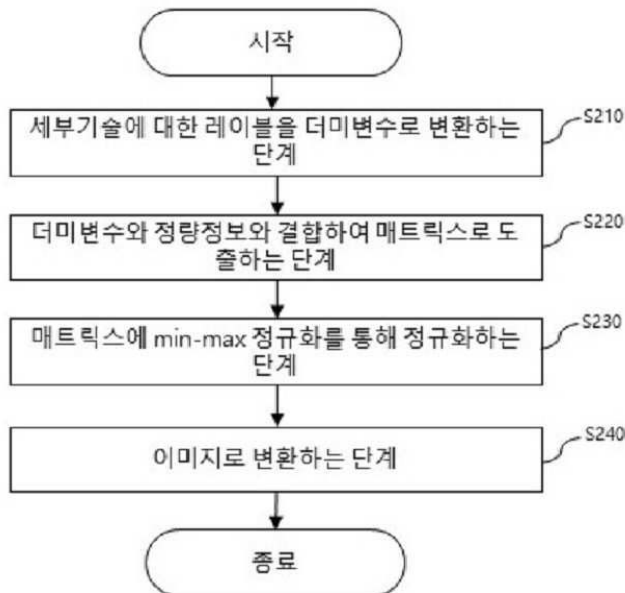
도면2

Indicator	Description	Measurable Value Information
app_to_regi (days)	Time required from application to registration	Utility value
all_claim_num	Number of claims	Rights
applicant_num	Number of applicants	Utility
inventor_num	Number of inventors	Sustainable development
current_owner_num	Number of current rights owners	Market impact
IPC_num	Number of IPC codes	Technology scalability
b_citation_num	Number of backward citations	Technology impact
f_citation_num	Number of forward citations	Technology impact
fam_nation_num	Number of family nations	Market impact
fam_doc_num	Number of family patents	Market impact
alone_app_yn	Sole application status (dummy)	Utility value
stand_patent_yn	Standard patent status (dummy)	Technology impact
lit_yn	Litigation status (dummy)	Utility value

도면3



도면4



도면5

