

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2024-0064989
(43) 공개일자 2024년05월14일

(51) 국제특허분류(Int. Cl.)

G16H 30/40 (2018.01) G06N 3/08 (2023.01)
G06V 10/82 (2022.01) G06V 30/18 (2022.01)
G16H 10/20 (2018.01) G16H 15/00 (2018.01)
G16H 30/20 (2018.01) G16H 50/20 (2018.01)
G16H 50/50 (2018.01) G16H 50/70 (2018.01)
G16H 70/60 (2018.01)

(52) CPC특허분류

G16H 30/40 (2018.01)
G06N 3/08 (2023.01)

(21) 출원번호 10-2022-0146478

(22) 출원일자 2022년11월04일

심사청구일자 2022년11월04일

(71) 출원인

경희대학교 산학협력단

경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)

(72) 발명자

이승룡

경기도 성남시 분당구 구미로144번길 25, 102동 102호(구미 동, 삼환빌라)

후아 캄 하오

경기도 수원시 영통구 매영로398번길 14, B04호(영통동)

레호만 우베드

경기도 수원시 영통구 매영로398번길 17, b04호(영통동)

(74) 대리인

두호특허법인

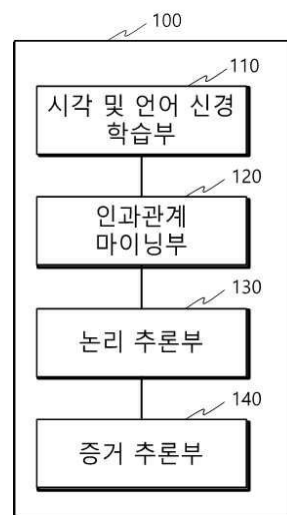
전체 청구항 수 : 총 17 항

(54) 발명의 명칭 **의학적 결정 지원 장치 및 방법**

(57) 요약

의료 영상 이미지에 대한 사용자 질문에 대응한 증거가 포함된 답변을 생성하는 의학적 결정 지원 장치 및 방법이 개시된다. 일 실시예에 따른 의학적 결정 지원 장치는 의료 영상 이미지를 수신하여 시각 임베딩(embedding)을 추출하는 시각 및 언어 신경 학습부; 시각 임베딩 및 사전 학습된 임베딩에 기초하여 소정 인덱스를 추출하며, 추출된 소정 인덱스 및 하나 이상의 인덱스-개념 쌍을 포함하는 해시맵(hash map)을 기초로 의료 영상 이미지에 대한 소정 인과관계(cause)를 출력하는 인과관계 마이닝부; 시각 임베딩, 배경 지식 그래프 및 소정 인과관계에 기초하여 의료 보고서를 생성하는 논리 추론부; 및 소정 인과관계 및 의료 보고서에 기초하여 의료 영상 이미지에 대한 사용자의 질문에 대한 증거(proof)가 포함된 답변을 생성하는 증거 추론부를 포함할 수 있다.

대표도 - 도1



(52) CPC특허분류

G06V 10/82 (2022.01)
G06V 30/18 (2023.08)
G16H 10/20 (2021.08)
G16H 15/00 (2021.08)
G16H 30/20 (2018.01)
G16H 50/20 (2018.01)
G16H 50/50 (2018.01)
G16H 50/70 (2018.01)
G16H 70/60 (2021.08)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711160619
과제번호	2022-0-00078-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	사람중심인공지능핵심원천기술개발
연구과제명	의료지식 생성을 위한 설명가능한 논리추론 기술개발
기 여 율	1/1
과제수행기관명	경희대학교 산학협력단
연구기간	2022.04.01 ~ 2022.12.31

명세서

청구범위

청구항 1

의료 영상 이미지를 수신하여 시각 임베딩(embedding)을 추출하는 시각 및 언어 신경 학습부;

상기 시각 임베딩 및 사전 학습된 임베딩에 기초하여 소정 인덱스를 추출하며, 추출된 소정 인덱스 및 하나 이상의 인덱스-개념 쌍을 포함하는 해시맵(hash map)을 기초로 상기 의료 영상 이미지에 대한 소정 인과관계(cause)를 출력하는 인과관계 마이닝부;

상기 시각 임베딩, 배경 지식 그래프 및 상기 소정 인과관계에 기초하여 의료 보고서를 생성하는 논리 추론부; 및

상기 소정 인과관계 및 상기 의료 보고서에 기초하여 상기 의료 영상 이미지에 대한 사용자의 질문에 대한 증거(proof)가 포함된 답변을 생성하는 증거 추론부를 포함하는, 의학적 결정 지원 장치.

청구항 2

제 1 항에 있어서,

상기 시각 및 언어 신경 학습부는

의료 영상 이미지로부터 시각 임베딩을 추출하는 시각 인코더;

의료 영상 이미지에 대한 텍스트 데이터로부터 텍스트 임베딩을 추출하는 텍스트 인코더; 및

상기 시각 임베딩 및 상기 텍스트 임베딩에 기초하여 멀티 모달 임베딩을 추출하는 교차 모달 디코더(Cross-Modal Decoder)를 포함하는, 의학적 결정 지원 장치.

청구항 3

제 2 항에 있어서,

상기 시각 인코더는

컨볼루션 커널, 배치 정규화 레이어 및 비선형 활성화 레이어를 포함하는 하나 이상의 컨볼루션 레이어 및 전역 풀링 레이어(Global Pooling Layer)를 포함하는, 의학적 결정 지원 장치.

청구항 4

제 2 항에 있어서,

상기 텍스트 인코더는

단어 임베딩(Word Embedding) 레이어, 위치 임베딩(Positional Embedding) 레이어 및 추가(Add) 레이어를 포함하는, 의학적 결정 지원 장치.

청구항 5

제 2 항에 있어서,

상기 교차 모달 디코더는

하나 이상의 디코더를 포함하며,

상기 하나 이상의 디코더 각각은 상기 시각 인코더의 출력인 시각 임베딩, 상기 텍스트 인코더의 출력인 텍스트 임베딩 및 이전 단계 디코더의 출력인 멀티 모달 임베딩 중 적어도 하나에 기초하여 멀티 모달 임베딩을 출력하는, 의학적 결정 지원 장치.

청구항 6

제 1 항에 있어서,

상기 인과관계 마이닝부는

시각 임베딩을 입력 받아 인과관계 개념을 출력하는 모달리티-불가지론(Modality-Agnostic) 인과관계 디코더;

시각 임베딩, 인과관계 개념을 저장하는 인과관계 트리플 임베딩 및 개념 저장기;

코사인 유사도 및 유클리디언 유사도 중 적어도 하나를 이용하여 인과관계 트리플 임베딩 및 개념 저장기에 저장된 사전 학습된 임베딩 중 입력된 시각 임베딩과 가장 유사한 사전 학습된 임베딩의 인덱스를 추출하는 인과관계 후보 검색기; 및

상기 추출된 인덱스 및 상기 인과관계 트리플 임베딩 및 개념 저장기에 저장된 해시맵에 기초하여 상기 추출된 인덱스에서 개념을 매칭하고, 매칭된 개념을 문장으로의 변환하여 연관된 인과관계 표현을 출력하는 인과관계 개념 추론기를 포함하는, 의학적 결정 지원 장치.

청구항 7

제 1 항에 있어서,

상기 논리 추론부는

시각 임베딩, 텍스트 임베딩 및 그래프 기반 임베딩을 기초로 멀티모달 임베딩을 생성하고, 멀티모달 임베딩에 기초하여 진단 레이블을 생성하는 멀티모달 인식기; 및

상기 진단 레이블 및 병리학적 논리에 기초하여 논리 개념을 생성하며, 상기 논리 개념 및 상기 소정 인과관계에 기초하여 생성된 병리학적 메타 규칙을 자연어로 변환하여 의료 보고서를 생성하는 병리학적 추론기를 포함하는, 의학적 결정 지원 장치.

청구항 8

제 1 항에 있어서,

상기 증거 추론부는

의료 영상 이미지에 대한 사용자의 질문을 입력 받아 시각적 질문과 답변의 유형을 추출하여 시각적 질문 답변(Visual Question Answer, VQA) 유형에 따른 VQA 트리플을 결정하는 시각적 질문 답변 통역기(Visual Question Answer Interpreter);

상기 VQA 트리플 및 상기 소정 인과관계 및 의료 보고서를 입력 받아 상기 소정 인과관계 및 의료 보고서로부터 상기 시각적 질문 답변과 관련된 문장을 추출하며, 추출된 문장에 대한 엔터티(Entity), 문장 및 컨텍스트 형태의 계층을 구성하는 의미 분석기(Semantic Parser); 및

상기 VQA 트리플 및 중복 필터링된 엔터티(Entity), 문장 및 컨텍스트 형태의 계층을 입력 받아 사용자의 질문에 대한 증거를 생성하는 사실 추론 및 검증기(Factual Reasoning & Verifier)를 포함하는, 의학적 결정 지원 장치.

청구항 9

의료 영상 이미지를 수신하여 시각 임베딩(embedding)을 추출하는 시각 및 언어 신경 학습 단계;

상기 시각 임베딩 및 사전 학습된 임베딩에 기초하여 소정 인덱스를 추출하며, 추출된 소정 인덱스 및 하나 이상의 인덱스-개념 쌍을 포함하는 해시맵(hash map)을 기초로 상기 의료 영상 이미지에 대한 소정 인과관계(cause)를 출력하는 인과관계 마이닝 단계;

상기 시각 임베딩, 배경 지식 그래프 및 상기 소정 인과관계에 기초하여 의료 보고서를 생성하는 논리 추론 단계; 및

상기 소정 인과관계 및 상기 의료 보고서에 기초하여 상기 의료 영상 이미지에 대한 사용자의 질문에 대한 증거(proof)가 포함된 답변을 생성하는 증거 추론 단계를 포함하는, 의학적 결정 지원 방법.

청구항 10

제 9 항에 있어서,

상기 시각 및 언어 신경 학습 단계는

의료 영상 이미지로부터 시각 임베딩을 추출하며,

의료 영상 이미지에 대한 텍스트 데이터로부터 텍스트 임베딩을 추출하며,

상기 시각 임베딩 및 상기 텍스트 임베딩에 기초하여 멀티 모달 임베딩을 추출하는, 의학적 결정 지원 방법.

청구항 11

제 10 항에 있어서,

상기 시각 및 언어 신경 학습 단계는

컨볼루션 커널, 배치 정규화 레이어 및 비선형 활성화 레이어를 포함하는 하나 이상의 컨볼루션 레이어 및 전역 풀링 레이어(Global Pooling Layer)를 포함하는 시각 인코더를 이용하는, 의학적 결정 지원 방법.

청구항 12

제 10 항에 있어서,

상기 시각 및 언어 신경 학습 단계는

단어 임베딩(Word Embedding) 레이어, 위치 임베딩(Positional Embedding) 레이어 및 추가(Add) 레이어를 포함하는 텍스트 인코더를 이용하는, 의학적 결정 지원 방법.

청구항 13

제 10 항에 있어서,

상기 시각 및 언어 신경 학습 단계는

하나 이상의 디코더를 포함하며, 상기 하나 이상의 디코더 각각은 시각 인코더의 출력인 시각 임베딩, 텍스트 인코더의 출력인 텍스트 임베딩 및 이전 단계 디코더의 출력인 멀티 모달 임베딩 중 적어도 하나에 기초하여 멀티 모달 임베딩을 출력하는 교차 모달 디코더를 이용하는, 의학적 결정 지원 방법.

청구항 14

제 9 항에 있어서,

상기 인과관계 마이닝 단계는

시각 임베딩을 입력 받아 인과관계 개념을 출력하며,

시각 임베딩, 인과관계 개념을 저장하며,

코사인 유사도 및 유클리디언 유사도 중 적어도 하나를 이용하여 인과관계 트리플 임베딩 및 개념 저장기에 저장된 사전 학습된 임베딩 중 입력된 시각 임베딩과 가장 유사한 사전 학습된 임베딩의 인덱스를 추출하며,

상기 추출된 인덱스 및 인과관계 트리플 임베딩 및 개념 저장기에 저장된 해시맵에 기초하여 상기 추출된 인덱스에서 개념을 매칭하고, 매칭된 개념을 문장으로의 변환하여 연관된 인과관계 표현을 출력하는, 의학적 결정 지원 방법.

청구항 15

제 9 항에 있어서,

상기 논리 추론 단계는

시각 임베딩, 텍스트 임베딩 및 그래프 기반 임베딩을 기초로 멀티모달 임베딩을 생성하고, 멀티모달 임베딩에 기초하여 진단 레이블을 생성하며,

상기 진단 레이블 및 병리학적 논리에 기초하여 논리 개념을 생성하며, 상기 논리 개념 및 상기 소정 인과관계에 기초하여 생성된 병리학적 메타 규칙을 자연어로 변환하여 의료 보고서를 생성하는, 의학적 결정 지원 방법.

청구항 16

제 9 항에 있어서,

상기 증거 추론 단계는

의료 영상 이미지에 대한 사용자의 질문을 입력 받아 시각적 질문과 답변의 유형을 추출하여 시각적 질문 답변(Visual Question Answer, VQA) 유형에 따른 VQA 트리플을 결정하며,

상기 VQA 트리플 및 상기 소정 인과관계 및 의료 보고서를 입력 받아 상기 소정 인과관계 및 의료 보고서로부터 상기 시각적 질문 답변과 관련된 문장을 추출하며, 추출된 문장에 대한 엔터티(Entity), 문장 및 컨텍스트 형태의 계층을 구성하며,

상기 VQA 트리플 및 중복 필터링된 엔터티(Entity), 문장 및 컨텍스트 형태의 계층을 입력 받아 사용자의 질문에 대한 증거를 생성하는, 의학적 결정 지원 방법.

청구항 17

비일시적 컴퓨터 판독 가능한 저장 매체(non-transitory computer readable storage medium)에 저장된 컴퓨터 프로그램으로서,

상기 컴퓨터 프로그램은 하나 이상의 명령어들을 포함하고, 상기 명령어들은 하나 이상의 프로세서들을 갖는 컴퓨팅 장치에 의해 실행될 때, 상기 컴퓨팅 장치로 하여금,

의료 영상 이미지를 수신하여 시각 임베딩(embedding)을 추출하는 시각 및 언어 신경 학습 단계;

상기 시각 임베딩 및 사전 학습된 임베딩에 기초하여 소정 인덱스를 추출하며, 추출된 소정 인덱스 및 하나 이상의 인덱스-개념 쌍을 포함하는 해시맵(hash map)을 기초로 상기 의료 영상 이미지에 대한 소정 인과관계(cause)를 출력하는 인과관계 마이닝 단계;

상기 시각 임베딩, 배경 지식 그래프 및 상기 소정 인과관계에 기초하여 의료 보고서를 생성하는 논리 추론 단

계; 및

상기 소정 인과관계 및 상기 의료 보고서에 기초하여 상기 의료 영상 이미지에 대한 사용자의 질문에 대한 증거(proof)가 포함된 답변을 생성하는 증거 추론 단계를 수행하도록 하는, 컴퓨터 프로그램.

발명의 설명

기술 분야

[0001] 의료 영상 이미지에 대한 사용자 질문에 대응한 증거가 포함된 답변을 생성하는 의학적 결정 지원 장치 및 방법에 관한 것이다.

배경 기술

[0003] 최근 인공 신경망을 이용하여 의료 영상을 분석하는 기술이 활발히 연구되고 있다. 공개특허 제10-2022-0105723호의 경우, 미리 학습된 의료 영상 분석 모델을 이용하여 동일 진단 대상자에 대해 다른 시점에 촬영된 의료 영상 사이의 변화를 판별하는 방법을 제시하고 있다.

[0004] 그러나, 이러한 종래의 기술은 영상 분석에 대한 근거를 제공하지 않은 채 결론만을 제공하여 의사가 임상 단계에서 분석 결과에 대한 신뢰성을 가지기 어려운 문제가 있었다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 한국공개특허공보 제10-2022-0105723호(2022.07.28)

발명의 내용

해결하려는 과제

[0007] 의료 영상 이미지에 대한 사용자 질문에 대응한 증거가 포함된 답변을 생성하는 의학적 결정 지원 장치 및 방법을 제공하는데 목적이 있다.

과제의 해결 수단

[0009] 일 양상에 따르면, 의학적 결정 지원 장치는 의료 영상 이미지를 수신하여 시각 임베딩(embedding)을 추출하는 시각 및 언어 신경 학습부; 시각 임베딩 및 사전 학습된 임베딩에 기초하여 소정 인덱스를 추출하며, 추출된 소정 인덱스 및 하나 이상의 인덱스-개념 쌍을 포함하는 해시맵(hash map)을 기초로 의료 영상 이미지에 대한 소정 인과관계(cause)를 출력하는 인과관계 마이닝부; 시각 임베딩, 배경 지식 그래프 및 소정 인과관계에 기초하여 의료 보고서를 생성하는 논리 추론부; 및 소정 인과관계 및 의료 보고서에 기초하여 의료 영상 이미지에 대한 사용자의 질문에 대한 증거(proof)가 포함된 답변을 생성하는 증거 추론부를 포함할 수 있다.

[0010] 시각 및 언어 신경 학습부는 의료 영상 이미지로부터 시각 임베딩을 추출하는 시각 인코더; 의료 영상 이미지에 대한 텍스트 데이터로부터 텍스트 임베딩을 추출하는 텍스트 인코더 및 시각 임베딩 및 텍스트 임베딩에 기초하여 멀티 모달 임베딩을 추출하는 교차 모달 디코더(Cross-Modal Decoder)를 포함할 수 있다.

[0011] 시각 인코더는 컨볼루션 커널, 배치 정규화 레이어 및 비선형 활성화 레이어를 포함하는 하나 이상의 컨볼루션 레이어 및 전역 풀링 레이어(Global Pooling Layer)를 포함할 수 있다.

[0012] 텍스트 인코더는 단어 임베딩(Word Embedding) 레이어, 위치 임베딩(Positional Embedding) 레이어 및 추가

(Add) 레이어를 포함할 수 있다.

- [0013] 교차 모달 디코더는 하나 이상의 디코더를 포함하며, 하나 이상의 디코더 각각은 시각 인코더의 출력인 시각 임베딩, 텍스트 인코더의 출력인 텍스트 임베딩 및 이전 단계 디코더의 출력인 멀티 모달 임베딩 중 적어도 하나에 기초하여 멀티 모달 임베딩을 출력할 수 있다.
- [0014] 인과관계 마이닝부는 시각 임베딩을 입력 받아 인과관계 개념을 출력하는 모달리티-불가지론(Modality-Agnostic) 인과관계 디코더; 시각 임베딩, 인과관계 개념을 저장하는 인과관계 트리플 임베딩 및 개념 저장기; 코사인 유사도 및 유클리디언 유사도 중 적어도 하나를 이용하여 인과관계 트리플 임베딩 및 개념 저장기에 저장된 사전 학습된 임베딩 중 입력된 시각 임베딩과 가장 유사한 사전 학습된 임베딩의 인덱스를 추출하는 인과관계 후보 검색기; 및 추출된 인덱스 및 인과관계 트리플 임베딩 및 개념 저장기에 저장된 해시맵에 기초하여 추출된 인덱스에서 개념을 매칭하고, 매칭된 개념을 문장으로의 변환하여 연관된 인과관계 표현을 출력하는 인과관계 개념 추론기를 포함할 수 있다.
- [0015] 논리 추론부는 시각 임베딩, 텍스트 임베딩 및 그래프 기반 임베딩을 기초로 멀티모달 임베딩을 생성하고, 멀티모달 임베딩에 기초하여 진단 레이블을 생성하는 멀티모달 인식기; 및 진단 레이블 및 병리학적 논리에 기초하여 논리 개념을 생성하며, 논리 개념 및 소정 인과관계에 기초하여 생성된 병리학적 메타 규칙을 자연어로 변환하여 의료 보고서를 생성하는 병리학적 추론기를 포함할 수 있다.
- [0016] 증거 추론부는 의료 영상 이미지에 대한 사용자의 질문을 입력 받아 시각적 질문과 답변의 유형을 추출하여 시각적 질문 답변(Visual Question Answer, VQA) 유형에 따른 VQA 트리플을 결정하는 시각적 질문 답변 통역기(Visual Question Answer Interpreter); VQA 트리플 및 소정 인과관계 및 의료 보고서를 입력 받아 소정 인과관계 및 의료 보고서로부터 시각적 질문 답변과 관련된 문장을 추출하며, 추출된 문장에 대한 엔터티(Entity), 문장 및 컨텍스트 형태의 계층을 구성하는 의미 분석기(Semantic Parser); 및 VQA 트리플 및 중복 필터링된 엔터티(Entity), 문장 및 컨텍스트 형태의 계층을 입력 받아 사용자의 질문에 대한 증거를 생성하는 사실 추론 및 검증기(Factual Reasoning & Verifier)를 포함할 수 있다.
- [0017] 일 양상에 따르면, 의학적 결정 지원 방법은 의료 영상 이미지를 수신하여 시각 임베딩(embedding)을 추출하는 시각 및 언어 신경 학습 단계; 시각 임베딩 및 사전 학습된 임베딩에 기초하여 소정 인덱스를 추출하며, 추출된 소정 인덱스 및 하나 이상의 인덱스-개념 쌍을 포함하는 해시맵(hash map)을 기초로 의료 영상 이미지에 대한 소정 인과관계(cause)를 출력하는 인과관계 마이닝 단계; 시각 임베딩, 배경 지식 그래프 및 소정 인과관계에 기초하여 의료 보고서를 생성하는 논리 추론 단계; 및 소정 인과관계 및 의료 보고서에 기초하여 의료 영상 이미지에 대한 사용자의 질문에 대한 증거(proof)가 포함된 답변을 생성하는 증거 추론 단계를 포함할 수 있다.
- [0018] 시각 및 언어 신경 학습 단계는 의료 영상 이미지로부터 시각 임베딩을 추출하며, 의료 영상 이미지에 대한 텍스트 데이터로부터 텍스트 임베딩을 추출하며, 시각 임베딩 및 텍스트 임베딩에 기초하여 멀티 모달 임베딩을 추출할 수 있다.
- [0019] 시각 및 언어 신경 학습 단계는 컨볼루션 커널, 배치 정규화 레이어 및 비선형 활성화 레이어를 포함하는 하나 이상의 컨볼루션 레이어 및 전역 풀링 레이어(Global Pooling Layer)를 포함하는 시각 인코더를 이용할 수 있다.
- [0020] 시각 및 언어 신경 학습 단계는 단어 임베딩(Word Embedding) 레이어, 위치 임베딩(Positional Embedding) 레이어 및 추가(Add) 레이어를 포함하는 텍스트 인코더를 이용할 수 있다.
- [0021] 시각 및 언어 신경 학습 단계는 하나 이상의 디코더를 포함하며, 하나 이상의 디코더 각각은 시각 인코더의 출력인 시각 임베딩, 텍스트 인코더의 출력인 텍스트 임베딩 및 이전 단계 디코더의 출력인 멀티 모달 임베딩 중 적어도 하나에 기초하여 멀티 모달 임베딩을 출력하는 교차 모달 디코더를 이용할 수 있다.
- [0022] 인과관계 마이닝 단계는 시각 임베딩을 입력 받아 인과관계 개념을 출력하며, 시각 임베딩, 인과관계 개념을 저장하며, 코사인 유사도 및 유클리디언 유사도 중 적어도 하나를 이용하여 인과관계 트리플 임베딩 및 개념 저장기에 저장된 사전 학습된 임베딩 중 입력된 시각 임베딩과 가장 유사한 사전 학습된 임베딩의 인덱스를 추출하며, 추출된 인덱스 및 인과관계 트리플 임베딩 및 개념 저장기에 저장된 해시맵에 기초하여 추출된 인덱스에서 개념을 매칭하고, 매칭된 개념을 문장으로의 변환하여 연관된 인과관계 표현을 출력할 수 있다.
- [0023] 논리 추론 단계는 시각 임베딩, 텍스트 임베딩 및 그래프 기반 임베딩을 기초로 멀티모달 임베딩을 생성하고, 멀티모달 임베딩에 기초하여 진단 레이블을 생성하며, 진단 레이블 및 병리학적 논리에 기초하여 논리 개념을

생성하며, 논리 개념 및 소정 인과관계에 기초하여 생성된 병리학적 메타 규칙을 자연어로 변환하여 의료 보고서를 생성할 수 있다.

[0024] 증거 추론 단계는 의료 영상 이미지에 대한 사용자의 질문을 입력 받아 시각적 질문과 답변의 유형을 추출하여 시각적 질문 답변(Visual Question Answer, VQA) 유형에 따른 VQA 트리플을 결정하며, VQA 트리플 및 소정 인과관계 및 의료 보고서를 입력 받아 소정 인과관계 및 의료 보고서로부터 시각적 질문 답변과 관련된 문장을 추출하며, 추출된 문장에 대한 엔터티(Entity), 문장 및 컨텍스트 형태의 계층을 구성하며, VQA 트리플 및 중복 필터링된 엔터티(Entity), 문장 및 컨텍스트 형태의 계층을 입력 받아 사용자의 질문에 대한 증거를 생성할 수 있다.

발명의 효과

[0026] 의료 영상 이미지에 대한 증거가 포함된 답변을 제공함으로써 사용자가 신뢰성을 가지고 분석 결과를 이용할 수 있는 효과가 있다.

도면의 간단한 설명

[0028] 도 1은 일 실시예에 따른 의학적 결정 지원 장치의 구성도이다.
 도 2는 일 실시예에 따른 의학적 결정 지원 장치의 세부 구성을 설명하기 위한 구성도이다.
 도 3은 일 실시예에 따른 시각 및 언어 신경 학습부의 구성도이다.
 도 4는 일 실시예에 따른 인과관계 마이닝부 구성도이다.
 도 5는 일 실시예에 따른 논리 추론부의 구성도이다.
 도 6은 일 실시예에 따른 증거 추론부의 구성도이다.
 도 7은 일 실시예에 따른 의학적 결정 지원 방법을 도시한 흐름도이다.
 도 8은 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 장치를 포함하는 컴퓨팅 환경을 예시하여 설명하기 위한 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0029] 이하, 첨부된 도면을 참조하여 본 발명의 일 실시예를 상세하게 설명한다. 본 발명을 설명함에 있어 관련된 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략할 것이다. 또한, 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로, 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.

[0030] 이하, 의학적 결정 지원 및 방법의 실시예들을 도면들을 참고하여 자세히 설명한다.

[0031] 도 1은 일 실시예에 따른 의학적 결정 지원 장치의 구성도이다.

[0032] 도 1을 참조하면, 의학적 결정 지원 장치(100)는 의료 영상 이미지를 수신하여 시각 임베딩(embedding)을 추출하는 시각 및 언어 신경 학습부(110), 시각 임베딩 및 사전 학습된 임베딩에 기초하여 소정 인텍스를 추출하며, 추출된 소정 인텍스 및 하나 이상의 인텍스-개념 쌍을 포함하는 해시맵(hash map)을 기초로 의료 영상 이미지에 대한 소정 인과관계(cause)를 출력하는 인과관계 마이닝부(120), 시각 임베딩, 배경 지식 그래프 및 소정 인과관계에 기초하여 의료 보고서를 생성하는 논리 추론부(130) 및 소정 인과관계 및 의료 보고서에 기초하여 의료 영상 이미지에 대한 사용자의 질문에 대한 증거(proof)가 포함된 답변을 생성하는 증거 추론부(140)를 포함할 수 있다.

[0033] 일 예에 따르면, 시각 언어 신경 학습부(110)는 신경 학습 프로세스를 수행하기 위하여 의료 이미지와 의료 이미지에 대한 임상 텍스트를 입력 받을 수 있으며, 의료 데이터의 기본 속성을 나타내는 특징 벡터인 임베딩을 추출할 수 있다. 인과관계 마이닝부(120)는 시각 및 언어 신경 학습부 (110)의 수치적 결과를 사용하여 의료 영

상 이미지에서 시각적 소건의 소정 인과관계를 추론할 수 있다.

- [0034] 일 예에 따르면, 논리 추론부(130)는 시각 언어 신경 학습 과정의 암묵적 발견에 대해 자연어 또는 인간이 해석할 수 있는 개념의 소정 인과관계를 배경 지식 그래프를 이용하여 파악한 후 논리적 추론 모델을 통해 의료 보고서를 생성할 수 있다. 증거 추론부(140)는 사용자인 의사가 제기한 의료 영상 이미지 또는 진단 결과에 대한 하나 이상의 질문에 대하여 추론된 답변을 뒷받침하는 증거(proof)와 함께 제공할 수 있다.
- [0035] 일 예로, 의학적 결정 지원 장치(100)가 제공하는 신경-심볼적 일관성 추론(Neural-Symbolic Coherent Reasoning) 방법의 최종 출력은 소정 인과관계, 의료 보고서 및 증거가 있는 답변(입력된 의료 영상 이미지의 시각적 내용과 관련된 하나 이상의 질문이 제공되는 경우)으로 구성될 수 있다. 이러한 결과는 의사가 의학적 결정을 내리는 데 도움을 줄 수 있다.
- [0036] 도 2는 일 실시예에 따른 의학적 결정 지원 장치의 세부 구성을 설명하기 위한 구성도이다.
- [0037] 일 실시예에 따르면, 시각 및 언어 신경 학습부(110)는 시각 인코더(203)를 이용하여 입력된 의료 영상 이미지(201)의 고유한 속성을 수치 벡터 형태로 전달할 수 있는 시각 임베딩을 추출할 수 있다. 텍스트 인코더(204)는 입력된 의료 영상에 대한 텍스트 데이터(202)가 주어지면 텍스트 임베딩을 생성할 수 있다.
- [0038] 일 예로, 시각 임베딩 및 텍스트 임베딩은 교차-모달 디코더(Cross-Modal Decoder)(205)에 입력되며, 교차-모달 디코더(205)의 출력은 시각 언어 작업을 훈련하는데 이용될 수 있다. 서로 다른 시각 언어 작업으로 학습한 사전 훈련된 교차-모달 디코더(205)는 모델 전이 학습 기법을 통해 인과관계 마이닝부(120)의 모달리티-불가지론 인과관계 디코더(206)로 채택될 수 있다.
- [0039] 일 예로, 학습(오프라인) 단계에서 시각 임베딩 및 텍스트 임베딩 중 적어도 하나가 주어지면 모달리티-불가지론 인과관계 디코더(Modality-Agnostic Causality Decoder)(206)는 인과관계 임베딩 및 개념을 추출한 다음 인과관계 트리플 임베딩 및 개념 저장기(Causal Triple Embedding & Concept Repository)(208)에 저장할 수 있다. 인과관계 후보 검색기(Nearest Causal Candidate Retrieval)(207)는 추론(온라인) 단계에서 인과관계 트리플 임베딩 및 개념 저장기(208)에 사전 학습된 임베딩을 참조하여 시각적 임베딩 입력 받아 최상의 소정 인텍스를 생성할 수 있다.
- [0040] 일 예에 따르면, 인과관계 개념 추론기(Causal Concept Inference)(209)는 인과관계 트리플 임베딩 및 개념 저장기(208)에서 인텍스와 개념의 쌍으로 구성된 해시맵을 트리거하여 입력 의료 이미지의 심각도 상태에 대한 소정 인과관계를 생성할 수 있다.
- [0041] 일 예에 따르면, 논리적 추론부(130)의 멀티모달 인식기(Multimodal Recognizer)(212)는 배경 지식 그래프(background knowledge graph)(211)와 함께 시각 임베딩 및 텍스트 임베딩을 입력 받아 진단 레이블을 출력할 수 있다. 병리학적 추론기(Pathological Reasoner)(213)는 출력된 진단 레이블, 인과 관계 마이닝부(120)서 추론한 소정 인과관계 및 배경 지식 그래프를 입력 받아 자연 언어로 의료 보고서를 생성할 수 있다.
- [0042] 일 예로, 제안된 신경 기호 일관성 추론(Neural-Symbolic Coherent Reasoning) 방법론의 추론 결과에는 소정 인과관계와 의료 보고서가 포함될 수 있다. 또한, 의사가 입력 의료 이미지에 대해 하나 이상의 임상 질문을 입력하는 경우 추론 절차는 증거 추론부(140)에 의해 추가로 활용될 수 있다. 예를 들어, 시각적 질문 답변 통역기(Visual Question Answer Interpreter)(216)는 후속 구성요소인 의미 분석기(Semantic Parser)(218) 및 사실적 추론 및 검증기(Factual Reasoning & Verifier)(219)의 입력으로서 시각적 질의응답 트리플을 생성하도록 활성화될 수 있다. 의미 분석기(218)는 필터링된 계층적 트리플을 추론하기 위해 이전 추론 결과, 즉 소정 인과관계 및 의료 보고서를 이용할 수 있다.
- [0043] 일 예에 따르면, 사실적 추론 및 검증기(219)는 이전 작업에서 검색된 주어진 기호 정보를 기반으로 입력 질문에 대한 답변의 이론적 근거를 지원하기 위해 최종 예측된 증거를 생성할 수 있다.
- [0044] 이에 따라, 일 실시예에 따른 경우, 의학적 결정 지원 장치(100)는 인과관계, 논리, 의료 보고서 및 증거(VQA 작업에 대한 추론된 답변 지원)의 텍스트 표현으로 구성된 일관된 다차원 결과의 추론을 통해 임상 진단 문제에 대해 추론할 수 있다.
- [0045] 도 3은 일 실시예에 따른 시각 및 언어 신경 학습부의 구성도이다.
- [0046] 일 실시예에 따르면, 시각 및 언어 신경 학습부(110)는 의료 영상 이미지로부터 시각 임베딩을 추출하는 시각 인코더(203), 의료 영상 이미지에 대한 텍스트 데이터로부터 텍스트 임베딩을 추출하는 텍스트 인코더(204) 및

시각 임베딩 및 텍스트 임베딩에 기초하여 멀티 모달 임베딩을 추출하는 교차 모달 디코더(Cross-Modal Decoder)(205)를 포함할 수 있다.

- [0047] 도 3을 참조하면, 시각 및 언어 신경 학습부(110)는 시각적 인코더(203), 텍스트 인코더(204) 및 교차 모드 디코더(205)로 구성될 수 있다. 각각의 구성 요소는 다양한 태스크 특정 출력 헤드와 유연하게 연결되어 시각 태스크(예: 이미지 인식 또는 분류 등), 언어 태스크(예: 감정 분석, 텍스트 분류 등) 및 시각-언어 태스크(예: 시각적 질문에 대한 답변, 이미지 캡션 등)를 동시에 또는 독립적으로 처리할 수 있다.
- [0048] 일 예에 따르면, 시각 및 언어 신경 학습부(110)의 출력은 태스크 특정 레이블 또는 텍스트이며, 이 중 실제 대응물은 해당 공개 데이터 세트에서 널리 사용할 수 있다. 따라서 시각 및 언어 신경 학습부(110)는 종단 간 방식으로 독립적으로 각 단일 입력 양식에 해당하는 강력한 기능을 학습할 수 있다. 결과적으로, 시각 및 언어 신경 학습부(110)는 멀티모달 데이터의 풍부한 주석에 크게 의존하지 않고 교차 모드 표현을 효과적으로 획득할 수 있다.
- [0049] 일 실시예에 따르면, 시각 인코더(203)는 컨볼루션 커널, 배치 정규화 레이어 및 비선형 활성화 레이어를 포함하는 하나 이상의 컨볼루션 레이어 및 전역 풀링 레이어(Global Pooling Layer)를 포함할 수 있다.
- [0050] 일 예로, 시각 인코더(203)에 입력되는 의료 영상 이미지(201)는 미리 정의된 하나 이상의 컨볼루션 계층의 순차적 스택에 의해 조작되며, 각 계층은 각각 컨볼루션 필터/커널(301), 배치 정규화 계층, 비선형 활성화 계층(예: Sigmoid, Rectified Linear Unit 등)으로 구성될 수 있다. 전역 풀링 계층(Global Pooling Layer)(302)는 의미론적으로 시각적인 특징을 전체 이미지 콘텐츠의 전체적인 맥락 정보를 나타내는 시각 임베딩으로 변환할 수 있다. 전역 풀링 계층(302)의 출력은 앞에서 언급한 것처럼 컨볼루션 계층(Convolutional Layers)의 학습 가능한 매개변수를 훈련하기 위해 시각 태스크를 학습하는 데 사용될 수 있다.
- [0051] 일 실시예에 따르면, 텍스트 인코더(204)는 단어 임베딩(Word Embedding) 레이어, 위치 임베딩(Positional Embedding) 레이어 및 추가(Add) 레이어를 포함할 수 있다.
- [0052] 일 예로, 텍스트 인코더(204)는 단어 임베딩(Word Embedding) 계층(304), 위치 임베딩(Positional Embedding) 계층(305), Add 계층(306) 및 미리 정의된 하나 이상의 트랜스포머(Transformer) 기반 인코더 계층의 순차적 스택을 기반으로 입력 텍스트를 처리하여 텍스트 임베딩을 생성할 수 있다. 여기서, 중간 결과는 위에서 언급한 언어 태스크를 학습하여 트랜스포머 기반 인코더 계층의 학습 가능한 매개변수를 학습하는 데 활용될 수 있다.
- [0053] 일 실시예에 따르면, 교차 모달 디코더(205)는 하나 이상의 디코더를 포함하며, 하나 이상의 디코더 각각은 시각 인코더의 출력인 시각 임베딩, 텍스트 인코더의 출력인 텍스트 임베딩 및 이전 단계 디코더의 출력인 멀티모달 임베딩 중 적어도 하나에 기초하여 멀티 모달 임베딩을 출력할 수 있다.
- [0054] 일 예로, 교차-모달 디코더(205)는 미리 정의된 수의 하나 이상의 디코더 계층의 순차적 스택을 포함할 수 있으며, 시각 인코더 및 텍스트 인코더에 의해 각각 추출된 시각 임베딩 및 텍스트 임베딩을 이용하여 멀티모달 임베딩을 추출할 수 있다. 교차-모달 디코더(205)의 최종 출력은 디코더 계층의 학습 가능한 매개변수를 훈련하기 위해 시각 언어 태스크를 학습하는 데 사용될 수 있다.
- [0055] 일 예로, 모달리티-불가지론 인과관계 디코더(206)는 모델 전이 학습 기법을 통해 인과관계 마이닝부(120)에서 미리 훈련된 교차-모달 디코더(205)를 채택하여 구성될 수 있다.
- [0056] 도 4는 일 실시예에 따른 인과관계 마이닝부 구성도이다.
- [0057] 일 실시예에 따르면, 인과관계 마이닝부(120)는 시각 임베딩을 입력 받아 인과관계 개념을 출력하는 모달리티-불가지론(Modality-Agnostic) 인과관계 디코더(206), 시각 임베딩, 인과관계 개념을 저장하는 인과관계 트리플 임베딩 및 개념 저장기(208), 코사인 유사도 및 유클리디언 유사도 중 적어도 하나를 이용하여 인과관계 트리플 임베딩 및 개념 저장기에 저장된 사전 학습된 임베딩 중 입력된 시각 임베딩과 가장 유사한 사전 학습된 임베딩의 인덱스를 추출하는 인과관계 후보 검색기(207) 및 추출된 인덱스 및 인과관계 트리플 임베딩 및 개념 저장기(208)에 저장된 해시맵에 기초하여 추출된 인덱스에서 개념을 매칭하고, 매칭된 개념을 문장으로의 변환하여 연관된 인과관계 표현을 출력하는 인과관계 개념 추론기(209)를 포함할 수 있다.
- [0058] 일 예로, 인과관계 마이닝부(120)는 귀추적 추론(abductive reasoning)을 이용하여 인과관계 마이닝을 수행할 수 있다.
- [0059] 일 예에 따르면, 학습(오프라인) 단계에서 모달리티-불가지론 인과관계 디코더(Modality-Agnostic Causality

Decoder)(206)는 시각적 임베딩(x)이 사전 정의된 인과관계 개념(c)으로 투영될 수 있도록 훈련된다.

[0060] 일 예에 따르면, 모달리티-불가지론 인과관계 디코더(206)는 m개의 디코더 계층의 순차 스택, 마지막 디코더 계층에 이어지는 Add 계층(306) 및 소프트맥스 분류기(Softmax Classifier)(402)를 포함할 수 있다.

[0061] 일 예로, 하나 이상의 디코더(401)는 각각 자기 주의(Self-Attention) 계층(403), 요소별 추가(Element-wise Addition)를 수행하는 추가(Add) 계층(306), 정규화(Normalization) 계층(404) 및 하나 이상의 완전 연결 계층(Fully Connected Layers)(405)의 순차적 스택을 포함할 수 있다. 학습 가능한 쿼리 행렬 W_q , 키 행렬 W_k 및 시 각 임베딩(x)과 차원이 일치하는 값 행렬 W_v 로 구성될 수 있는 자기 주의(Self-Attention) 계층(403)의 연산은 아래 수식과 같이 표현될 수 있다.

[0062] [수식식 1]

[0063]
$$Q = X \times W_q$$

[0064]
$$K = X \times W_k$$

[0065]
$$V = X \times W_v$$

[0066]
$$M = \frac{Q \times K^T}{\sqrt{d_k}}$$

[0067]
$$S = \text{Softmax}(M) = \frac{e^M}{\sum_{j=1}^{d_M} e^M}$$

[0068]
$$Z = S \times V$$

[0070] 여기서, Q, K, V, M, S, Z는 추출된 중간 임베딩 행렬을 나타내며, d_k 및 d_v 는 각각 K 및 V 행렬의 차원을 나타낸다. 일 예로, 자기 주의(Self-Attention)를 통하여 입력 X 자체가 다른 학습 가능한 행렬에 의해 조작되어 임베딩의 중요한 항목에 더 높은 값의 계수로 가중치를 부여하여 더 많은 주의를 기울이고 덜 중요한 항목은 더 낮은 값의 계수를 곱하여 약화시키는 것을 알 수 있습니다.

[0071] 일 예로, 추가 계층(306)은 인코딩된 정보가 예기치 않게 누락되는 것을 방지하기 위해 초기 입력 X를 재사용하는 데 관여한다. 이후, 정규화 계층(405)을 적용하여 입력 임베딩의 항목을 0에서 1 사이의 값으로 확장하여 특성 학습 프로세스의 안정성을 향상시킬 수 있다. 이어서, 디코더 계층에서 학습된 기능을 완성하기 위해 하나 이상의 완전 연결 계층(402)이 수행될 수 있다.

[0072] 일 예에 따르면, 디코더 계층 외부의 작업 워크플로와 관련하여 각 디코더 계층의 출력 $Y_i(i=1, \dots, m)$ 은 추가 계층(306)에서 함께 추가되며, 이는 다른 디코딩 단계에서 획득한 기능을 재사용 가능하게 유지하여 후속 소프트맥스 분류기(402)가 X에 대한 다른 인과관계 개념 간의 구별에 대한 더 많은 정보를 얻을 수 있도록 할 수 있다.

[0073] 일 예에 따르면, 학습 단계에서 모달리티-불가지론 인과관계 디코더(206)는 입력된 인과관계 임베딩과 인과관계 개념 간의 매핑을 키-값 쌍의 형태로 인과관계 트리플 임베딩 및 개념 저장기(208)에 저장할 수 있다.

[0074] 일 예에 따르면, 추론(온라인) 단계에서 인과관계 마이닝부(120)는 시각 임베딩에서 인과관계 개념을 추출할 수 있다. 예를 들어, 인과관계 후보 검색기(Nearest Causal Candidate Retrieval)(207)는 코사인/유클리디안 유사도 측정(Cosine/Euclidean similarity Measures)(406)를 사용한 비교를 통해 입력된 시각 임베딩과 가장 유사한 인과관계 트리플 임베딩 및 개념 저장기(208)에서 임베딩의 최상의 소정 인덱스를 수집할 수 있다. 이후, 인과적 개념 추론이 활성화되어 인덱스-개념 매퍼(Index-to-Concept Mapper)(407)의 순차 작업과 개념-문장 변환(Concept-to-Sentence Transformation)(408)을 통해 연관된 소정 인과관계를 출력할 수 있다.. 특히, 인덱스-개념 매퍼(407)는 인덱스(일명 키)와 인과관계 개념(일명 값)의 쌍을 포함하는 사전 정의된 해시맵을 스캔하여 이전 작업에서 검색된 최상의 소정 인덱스에 해당하는 최상의 인과관계 개념을 산출할 수 있다. 이후 개념에서 문장으로의 변환이 트리거 되어 주어(원인) + 동사 + 목적어(효과)로 구성되는 트리플(triple) 형태의 인과관계

개념을 소정 인과관계를 나타내는 인간 친화적인 문장으로 변환할 수 있다.

- [0075] 도 5는 일 실시예에 따른 논리 추론부의 구성도이다.
- [0076] 일 실시예에 따르면, 논리 추론부(130)는 시각 임베딩, 텍스트 임베딩 및 그래프 기반 임베딩을 기초로 멀티모달 임베딩을 생성하고, 멀티모달 임베딩에 기초하여 진단 레이블을 생성하는 멀티모달 인식기(212) 및 진단 레이블 및 병리학적 논리에 기초하여 논리 개념을 생성하며, 논리 개념 및 소정 인과관계에 기초하여 생성된 병리학적 메타 규칙을 자연어로 변환하여 의료 보고서를 생성하는 병리학적 추론기(213)를 포함할 수 있다.
- [0077] 일 예에 따르면, 논리 추론부(130)는 귀납적 추론을 수행할 수 있다. 도 5를 참조하면, 논리 추론부(130)은 두 개의 주요 구성요소, 즉 또 다른 신경 학습 절차로 작용하는 멀티모달 인식기(Multimodal Recognizer)(212)와 상징적 추론 메커니즘에 따라 수행하는 병리학적 추론기(Pathological Reasoner)(213)로 구성될 수 있다.
- [0078] 일 예에 따르면, 멀티모달 인식기(212)는 시각적 임베딩, 텍스트 임베딩 및 배경 지식 그래프를 포함하는 복수의 입력을 획득할 수 있다. 멀티모달 인식기(212)는 그래프 신경망(501)을 포함하며, 배경 지식 그래프(211)를 숫자 표현, 즉 그래프의 모든 노드(정점)와 모서리 사이의 복잡한 관계를 표현하기 위한 그래프 기반 임베딩으로 암묵적으로 변환할 수 있다. 멀티모달 주의(Multimodal Attention) 계층(502)은 시각적 임베딩, 텍스트 임베딩 및 사전에 추론된 그래프 기반 임베딩을 획득하여 멀티모달 임베딩을 생성할 수 있다. 멀티모달 주의(Multimodal Attention) 계층(502)의 작동 단계는 대부분 수학적 1의 자기 주의 계층과 유사하지만, Q, K, V가 시각적 임베딩인 다른 입력(텍스트 임베딩 및 이전에 추론된 그래프 기반 임베딩)에서 계산된다는 점에서 차이가 있다. 일 예로, 분류기(503)는 훈련 데이터 세트의 실측 주석과 관련하여 진단 레이블을 제공할 수 있다.
- [0079] 일 예에 따르면, 배경 지식 그래프(211)는 지식 그래프의 다양한 노드 및 모서리에서 병리 관련 컨텍스트를 포함하는 병리학 특정 논리 풀(Pathology-specific Logic Pool)(504)의 구성을 위해 이용될 수 있다. 예를 들어, 하나 이상의 질병 범주 또는 심각도가 관련 시각적 또는 텍스트 세부 정보와 함께 입력 데이터에 첨부될 수 있다. 예를 들어, 특정 질병 A가 의료 이미지의 특정 영역에 있는 이상 징후의 존재와 연관될 수 있다.
- [0080] 일 예에 따르면, 1차 논리 프로그래밍(First Order Logic Programming) 계층(505)은 술어, 등식, 부정, 이항 접속사, 수량자 기호 등으로 정의된 잘 형성된 공식을 유도하기 위한 변수 및 함수로서 분류기(503) 및 병리 특정 논리 풀(504)에서 수신한 병리학적 논리를 기초로 진단 레이블을 획득할 수 있다. 예를 들어, 1차 논리 프로그래밍 계층(505)의 출력은 진단된 질병 레이블과 비정상 징후, 증상, 의학적 상태, 설명 또는 입력 데이터의 집단적 물리적 또는 심리적 발생 사이의 가능한 모든 연관성을 보여주는 상징적 표현의 형태인 논리 개념일 수 있다.
- [0081] 일 예에 따르면, 가설 귀납(Relationship Hypothesis Induction) 계층(506)은 논리 개념과 함께 추론된 소정 인과관계(210)를 획득하여 시각적 내용을 넘어 심각성의 원인을 설명하는 모든 측면을 포함하는 시간적, 공간적 관점에서 병리학적 메타 규칙을 생성할 수 있다.
- [0082] 일 예에 따르면, 텍스트 추론(Textual Inference) 계층(507)은 병리학적 메타 규칙을 자연어로 변환하여 의료 보고서를 생성할 수 있다. 이 보고서는 입력된 의료 이미지의 시각적 내용에 대한 모든 병리학적 발견과 사람이 해석 가능한 추론을 종합적으로 설명할 수 있다.
- [0083] 도 6은 일 실시예에 따른 증거 추론부의 구성도이다.
- [0084] 일 실시예에 따르면, 증거 추론부(140)는 의료 영상 이미지에 대한 사용자의 질문을 입력 받아 시각적 질문과 답변의 유형을 추출하여 시각적 질문 답변(Visual Question Answer, VQA) 유형에 따른 VQA 트리플을 결정하는 시각적 질문 답변 통역기(Visual Question Answer Interpreter)(216), VQA 트리플 및 소정 인과관계 및 의료 보고서를 입력 받아 소정 인과관계 및 의료 보고서로부터 시각적 질문 답변과 관련된 문장을 추출하며, 추출된 문장에 대한 엔터티(Entity), 문장 및 컨텍스트 형태의 계층을 구성하는 의미 분석기(Semantic Parser)(208) 및 VQA 트리플 및 중복 필터링된 엔터티(Entity), 문장 및 컨텍스트 형태의 계층을 입력 받아 사용자의 질문에 대한 증거를 생성하는 사실 추론 및 검증기(Factual Reasoning & Verifier)(219)를 포함할 수 있다.
- [0085] 일 예로, 증거 추론부(140)는 의사인 사용자로부터 의료 영상 이미지인 시각적 콘텐츠와 관련된 하나 이상의 병리학적 질문을 수신할 수 있으며, 임상 에서 높은 견고성과 신뢰성을 가질 수 있는 VQA 시나리오를 처리할 수 있다. 즉, 증거 추론부(140)는 정확성과 투명성이 보장되는 동시에 기존 진단 절차가 가속화될 수 있도록 증거 지원을 통해 신속하게 답변을 생성할 수 있다.
- [0086] 일 예에 따르면, 시각적 질문 답변 통역기(216)는 VQA 작업으로 사전 훈련된 시각 및 언어 신경 학습부(110)를

채택하여 의료 영상의 특정 시각적 콘텐츠에 대한 입력 질문에 해당하는 답변을 생성할 수 있다. 시각적 질문-답변 유형 추출기 (Visual Question-Answering Type Extractor)(701)는 시각적 질문과 답변의 유형을 추출할 수 있다. 이후, 트리플 식별기(Triple Identifier)(702)가 VQA 유형에 따라 트리플을 결정할 수 있다. 트리플 식별기(702)는 Generative Pretrained Transformer(이하 "GPT") 또는 Bidirectional Encoder Representations from Transformers(이하 "BERT")일 수 있다.

- [0087] 일 예에 따르면, 의미 분석기(218)는 식별된 트리플 정보와 추론된 소정 인과관계 및 의료 보고서를 다음 작업에 대한 입력으로 사용할 수 있다.
- [0088] 일 예로, 증거 문장 선택기(Evidence Sentence Selection)(703)는 GPT 또는 BERT와 같이 사전 훈련된 다른 언어 모델로 구성될 수 있다. 증거 문장 선택기(703)는 소정 인과관계 및 의료 보고서 내에서 관련 문장을 식별할 수 있다.
- [0089] 일 예로, 계층 구조 작성기(Hierarchy Builder)(704)는 지정된 언어 모델을 사용하여 엔터티, 문장 및 컨텍스트 형태의 계층을 구성할 수 있다. 중복 필터(705)는 사전 정의된 게이팅 메커니즘을 사용하여 계층의 각 수준에서 필터링되며 필터링된 계층은 사실적 추론 및 검증기(219)로 전달될 수 있다.
- [0090] 일 예에 따르면, 사실 추론 및 검증기(219)에서 증거 저장기(Evidence Repository)(706)는 트리플 식별기(702)에서 추출한 VQA 트리플 및 대칭적 상호작용 주의(Symmetrical Interaction Attention)(707) 계층에 대한 입력으로 얻은 필터링된 계층 외에도 사전 정의된 도메인별 트리플/정제된 지식(triple/refined knowledge)을 제공할 수 있다.
- [0091] 일 예에 따르면, 다층 퍼셉트론을 기반으로 구축된 퓨전 네트워크(Fusion Network)(708)는 증거 간의 정보 전파를 처리하기 위해 특징 벡터의 형태로 식별된 주장-증거 관계를 수신할 수 있다. 하며, 퓨전 네트워크(708)의 각 노드는 하나의 증거를 나타낼 수 있으며, 이를 통하여 각 증거 문장에 대한 노드 특성을 생성할 수 있다.
- [0092] 일 예에 따르면, 사실 수집기(Facts Aggregator)(709)는 모든 노드 기능을 결합한 다음 증거 문장 사이에서 평가 절차를 수행하여 추론된 답변을 뒷받침하는 추론된 결과(즉, 소정 인과관계 및 의료 보고서)에 존재하는 하나 이상의 특정 증거를 결정할 수 있다. 이를 통하여, 사실 수집기(709)는 의사에게 신뢰할 수 있고 설득력 있는 방식으로 증거가 있는 답변을 제공할 수 있다.
- [0093] 도 7은 일 실시예에 따른 의학적 결정 지원 방법을 도시한 흐름도이다.
- [0094] 일 실시예에 따르면, 의학적 결정 지원 장치는 의료 영상 이미지를 수신하여 시각 임베딩(embedding)을 추출할 수 있다(710). 이후, 의학적 결정 지원 장치는 시각 임베딩 및 사전 학습된 임베딩에 기초하여 소정 인덱스를 추출하며(720), 추출된 소정 인덱스 및 하나 이상의 인덱스-개념 쌍을 포함하는 해시맵(hash map)을 기초로 의료 영상 이미지에 대한 소정 인과관계(cause)를 출력할 수 있다(730).
- [0095] 의학적 결정 지원 장치는 시각 임베딩, 배경 지식 그래프 및 소정 인과관계에 기초하여 의료 보고서를 생성할 수 있으며(740), 소정 인과관계 및 의료 보고서에 기초하여 의료 영상 이미지에 대한 사용자의 질문에 대한 증거(proof)가 포함된 답변을 생성할 수 있다(750).
- [0096] 일 예에 따르면, 의학적 결정 지원 장치는 의료 영상 이미지로부터 시각 임베딩을 추출하며, 의료 영상 이미지에 대한 텍스트 데이터로부터 텍스트 임베딩을 추출하며, 시각 임베딩 및 텍스트 임베딩에 기초하여 멀티 모달 임베딩을 추출할 수 있다.
- [0097] 일 예에 따르면, 의학적 결정 지원 장치는 컨볼루션 커널, 배치 정규화 레이어 및 비선형 활성화 레이어를 포함하는 하나 이상의 컨볼루션 레이어 및 전역 풀링 레이어(Global Pooling Layer)를 포함하는 시각 인코더, 의학적 결정 지원 장치는 단어 임베딩(Word Embedding) 레이어, 위치 임베딩(Positional Embedding) 레이어 및 추가(Add) 레이어를 포함하는 텍스트 인코더를 포함할 수 있다. 또한, 의학적 결정 지원 장치는 하나 이상의 디코더를 포함하며, 하나 이상의 디코더 각각은 시각 인코더의 출력인 시각 임베딩, 텍스트 인코더의 출력인 텍스트 임베딩 및 이전 단계 디코더의 출력인 멀티 모달 임베딩 중 적어도 하나에 기초하여 멀티 모달 임베딩을 출력하는 교차 모달 디코더를 포함할 수 있다.
- [0098] 일 예로, 의학적 결정 지원 장치는 시각 임베딩을 입력 받아 인과관계 개념을 출력하며, 시각 임베딩, 인과관계 개념을 저장하며, 코사인 유사도 및 유클리디언 유사도 중 적어도 하나를 이용하여 인과관계 트리플 임베딩 및 개념 저장기에 저장된 사전 학습된 임베딩 중 입력된 시각 임베딩과 가장 유사한 사전 학습된 임베딩의 인덱스를 추출하며, 추출된 인덱스 및 인과관계 트리플 임베딩 및 개념 저장기에 저장된 해시맵에 기초하여 추출된 인

텍스에서 개념을 매칭하고, 매칭된 개념을 문장으로의 변환하여 연관된 인과관계 표현을 출력할 수 있다.

- [0099] 일 예에 따르면, 의학적 결정 지원 장치는 시각 임베딩, 텍스트 임베딩 및 그래프 기반 임베딩을 기초로 멀티모달 임베딩을 생성하고, 멀티모달 임베딩에 기초하여 진단 레이블을 생성하며, 진단 레이블 및 병리학적 논리에 기초하여 논리 개념을 생성하며, 논리 개념 및 소정 인과관계에 기초하여 생성된 병리학적 메타 규칙을 자연어로 변환하여 의료 보고서를 생성할 수 있다.
- [0100] 일 예에 따르면, 의학적 결정 지원 장치는 의료 영상 이미지에 대한 사용자의 질문을 입력 받아 시각적 질문과 답변의 유형을 추출하여 시각적 질문 답변(Visual Question Answer, VQA) 유형에 따른 VQA 트리플을 결정하며, VQA 트리플 및 소정 인과관계 및 의료 보고서를 입력 받아 소정 인과관계 및 의료 보고서로부터 시각적 질문 답변과 관련된 문장을 추출하며, 추출된 문장에 대한 엔터티(Entity), 문장 및 컨텍스트 형태의 계층을 구성하며, VQA 트리플 및 중복 필터링된 엔터티(Entity), 문장 및 컨텍스트 형태의 계층을 입력 받아 사용자의 질문에 대한 증거를 생성할 수 있다.
- [0101] 도 7을 실시예 중 도 1 내지 도 6을 참조하여 설명한 내용과 중복되는 설명은 생략하였다.
- [0102] 도 8은 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 장치를 포함하는 컴퓨팅 환경(10)을 예시하여 설명하기 위한 블록도이다. 도시된 실시예에서, 각 컴포넌트들은 이하에 기술된 것 이외에 상이한 기능 및 능력을 가질 수 있고, 이하에 기술된 것 이외에도 추가적인 컴포넌트를 포함할 수 있다.
- [0103] 도시된 컴퓨팅 환경(10)은 컴퓨팅 장치(12)를 포함한다. 일 실시예에서, 컴퓨팅 장치(12)는 의학적 결정 지원 장치(100)일 수 있다.
- [0104] 컴퓨팅 장치(12)는 적어도 하나의 프로세서(14), 컴퓨터 판독 가능 저장 매체(16) 및 통신 버스(18)를 포함한다. 프로세서(14)는 컴퓨팅 장치(12)로 하여금 앞서 언급된 예시적인 실시예에 따라 동작하도록 할 수 있다. 예컨대, 프로세서(14)는 컴퓨터 판독 가능 저장 매체(16)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 상기 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 상기 컴퓨터 실행 가능 명령어는 프로세서(14)에 의해 실행되는 경우 컴퓨팅 장치(12)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.
- [0105] 컴퓨터 판독 가능 저장 매체(16)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능 저장 매체(16)에 저장된 프로그램(20)은 프로세서(14)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능 저장 매체(16)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 컴퓨팅 장치(12)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.
- [0106] 통신 버스(18)는 프로세서(14), 컴퓨터 판독 가능 저장 매체(16)를 포함하여 컴퓨팅 장치(12)의 다른 다양한 컴포넌트들을 상호 연결한다.
- [0107] 컴퓨팅 장치(12)는 또한 하나 이상의 입출력 장치(24)를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(22) 및 하나 이상의 네트워크 통신 인터페이스(26)를 포함할 수 있다. 입출력 인터페이스(22) 및 네트워크 통신 인터페이스(26)는 통신 버스(18)에 연결된다. 입출력 장치(24)는 입출력 인터페이스(22)를 통해 컴퓨팅 장치(12)의 다른 컴포넌트들에 연결될 수 있다. 예시적인 입출력 장치(24)는 포인팅 장치(마우스 또는 트랙패드 등), 키보드, 터치 입력 장치(터치패드 또는 터치스크린 등), 음성 또는 소리 입력 장치, 다양한 종류의 센서 장치 및/또는 촬영 장치와 같은 입력 장치, 및/또는 디스플레이 장치, 프린터, 스피커 및/또는 네트워크 카드와 같은 출력 장치를 포함할 수 있다. 예시적인 입출력 장치(24)는 컴퓨팅 장치(12)를 구성하는 일 컴포넌트로서 컴퓨팅 장치(12)의 내부에 포함될 수도 있고, 컴퓨팅 장치(12)와는 구별되는 별개의 장치로 컴퓨팅 장치(12)와 연결될 수도 있다.
- [0108] 이제까지 본 발명에 대하여 그 바람직한 실시 예를 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 따라서, 본 발명의 범위는 전술한 실시 예에 한정되지 않고 특허 청구범위에 기재된 내용과 동등한 범위 내에 있는 다양한 실시 형태가 포함되도록 해석되어야 할 것이다.

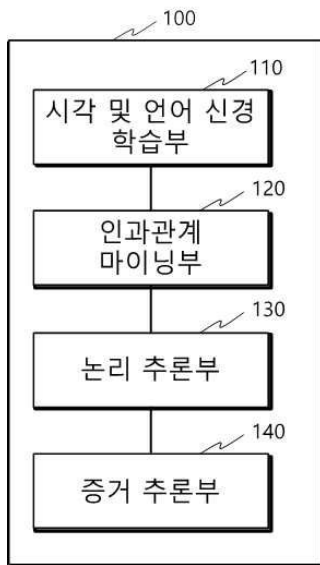
부호의 설명

[0110]

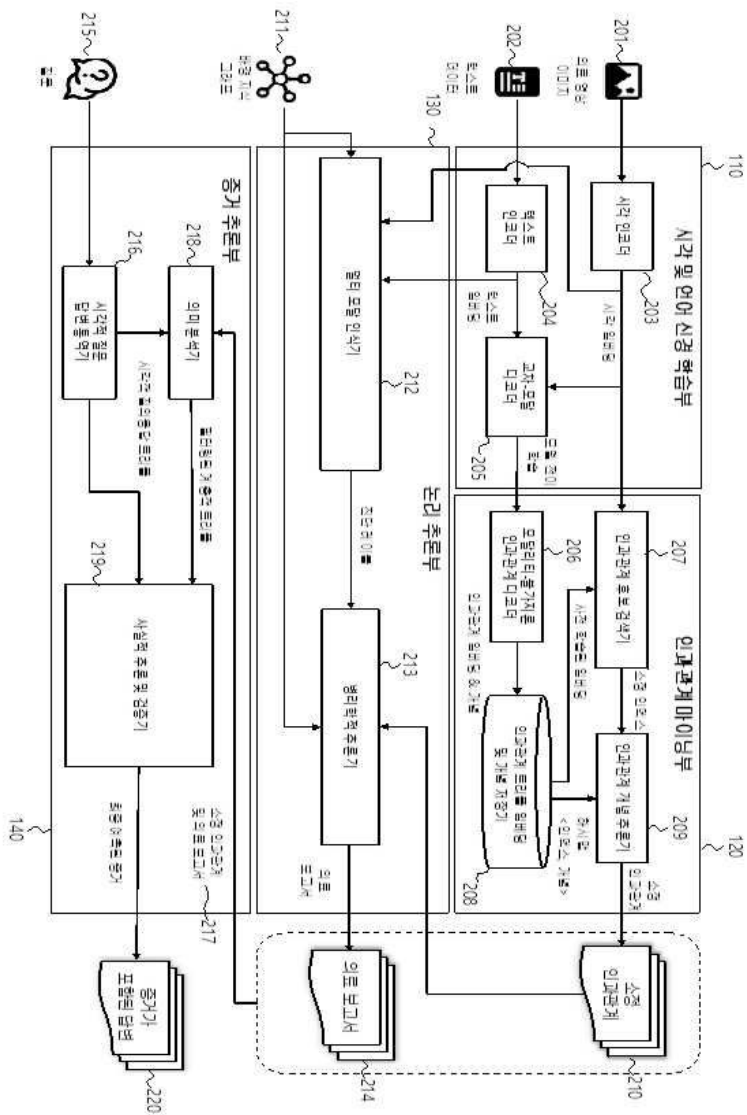
- 100: 의학적 결정 지원 장치
- 110: 시각 및 언어 신경 학습부
- 120: 인과관계 마이닝부
- 130: 논리 추론부
- 140: 증거 추론부

도면

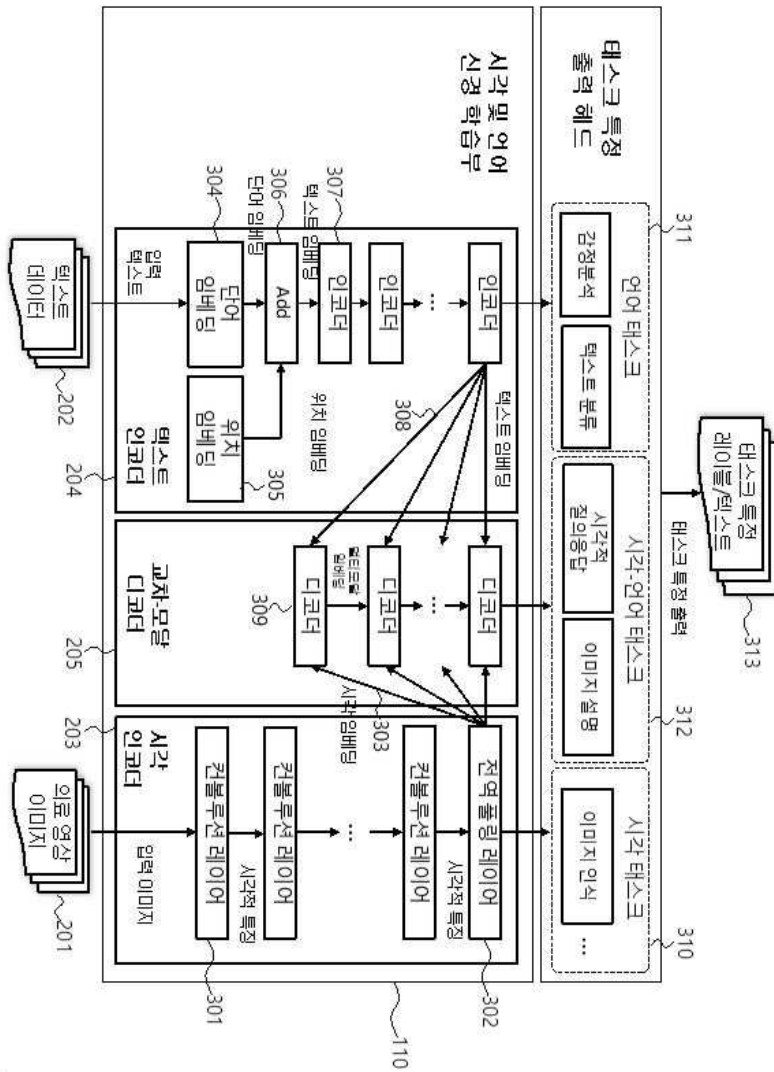
도면1



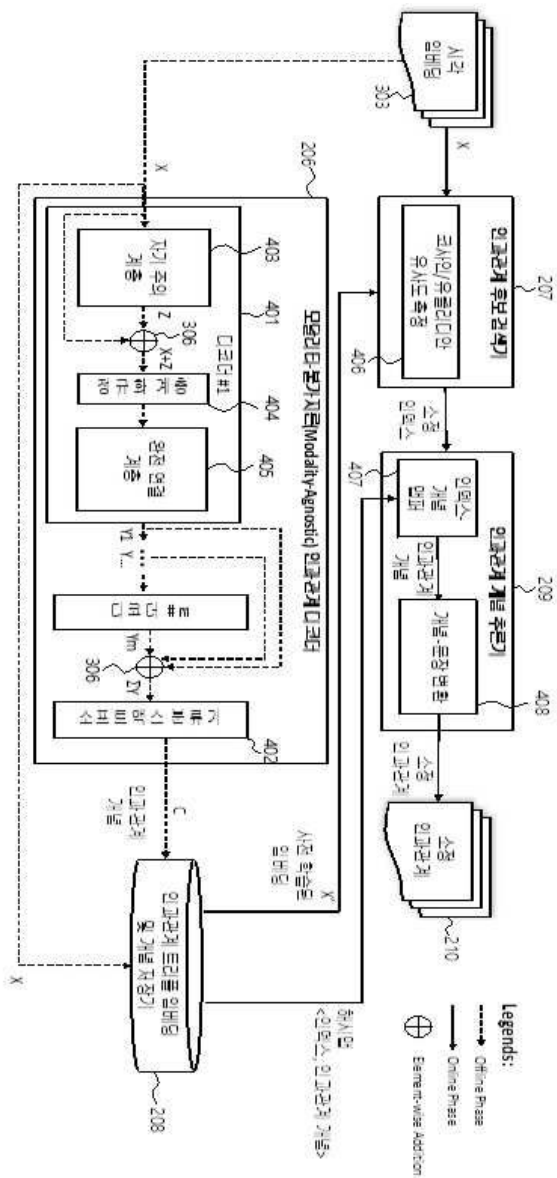
도면2



도면3

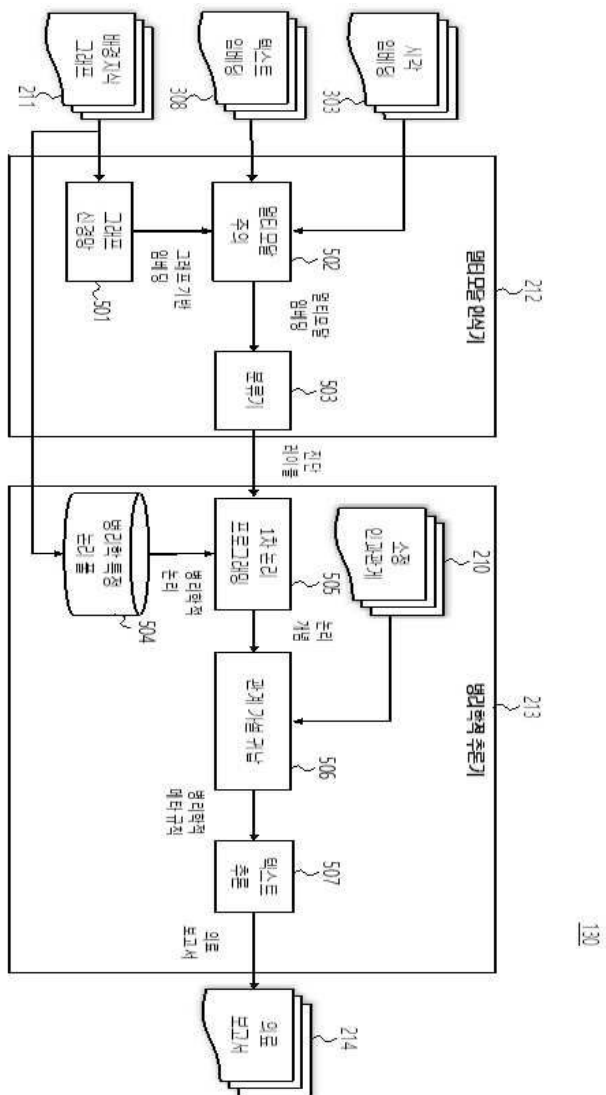


도면4

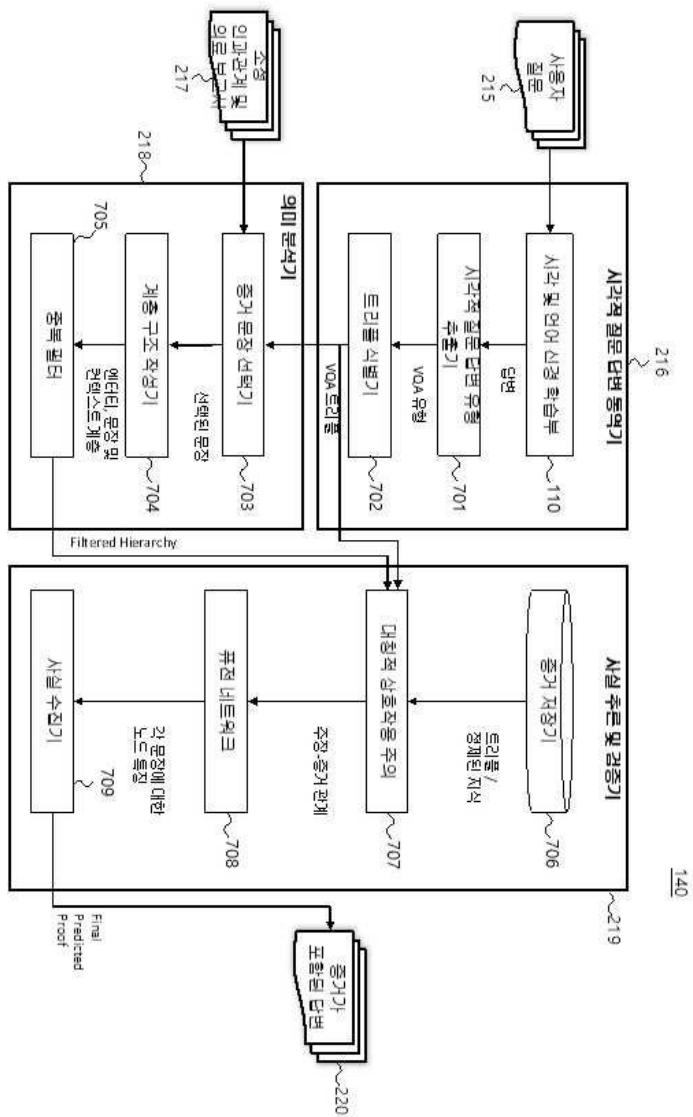


120

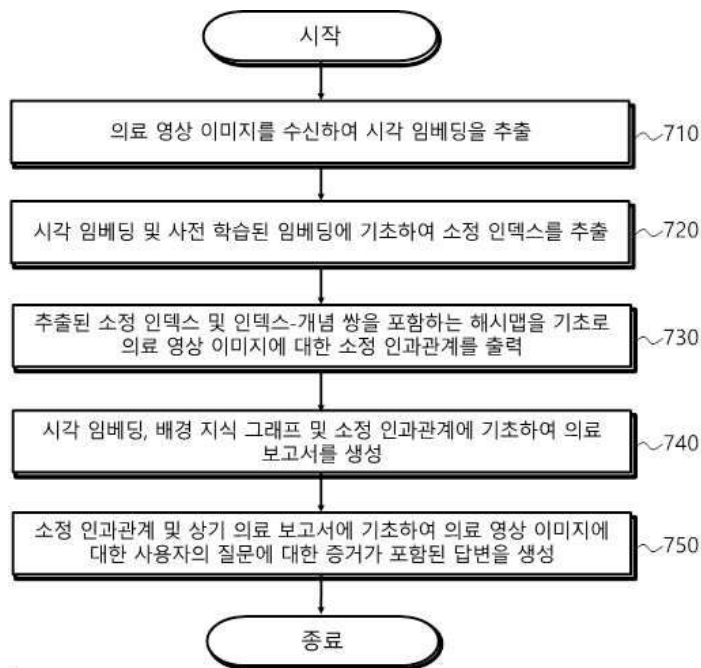
도면5



도면6



도면7



도면8

10

