



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0016324
(43) 공개일자 2023년02월02일

(51) 국제특허분류(Int. Cl.)

G06N 3/08 (2023.01) G06N 3/04 (2023.01)
G06N 5/02 (2023.01) G16H 30/20 (2018.01)
G16H 30/40 (2018.01) G16H 50/50 (2018.01)

(52) CPC특허분류

G06N 3/082 (2023.01)
G06N 3/042 (2023.01)

(21) 출원번호 10-2021-0097644

(22) 출원일자 2021년07월26일

심사청구일자 2021년07월26일

(71) 출원인

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자

석홍일

서울특별시 성북구 길음로 118, 411동 1202호(길음동, 길음뉴타운)

전은지

경기도 고양시 덕양구 화신로 170, 2105동 903호(행신동, 햇빛마을21단지아파트)

(74) 대리인

송인호, 윤형근, 최관락

전체 청구항 수 : 총 7 항

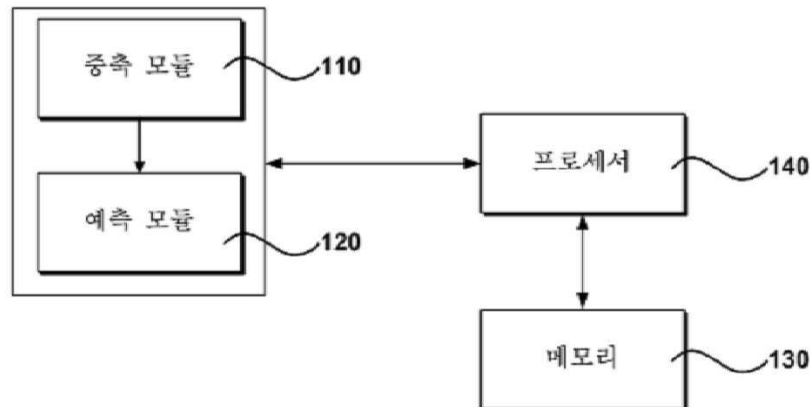
(54) 발명의 명칭 3차원 의료 영상 분석을 위한 자가학습 기반 전이 학습 방법 및 그 장치

(57) 요약

3차원 의료 영상 분석을 위한 자가학습 기반 전이 학습 방법 및 그 장치가 개시된다. 3차원 의료 영상 분석을 위한 자가학습 기반 전이 학습 프레임워크 장치는, 3차원 의료 영상을 3개의 평면으로 나눈 후 각 평면마다 2차원의 순차적 데이터로 표현한 후 3차원 공간 특징 정보로 통합하도록 사전 표현 학습(Representation learning)되는 중측 모듈; 및 완전 연결 계층으로 구성되며, 상기 3차원 공간 특징 정보를 기반으로 목표 태스크에 대한 예측을 수행하는 예측 모듈을 포함하되, 상기 사전 학습된 중측 모듈과 상기 예측 모듈은 상기 목표 태스크에 따라 미세 조정될 수 있다.

대표도 - 도1

100



(52) CPC특허분류

G06N 5/025 (2023.01)

G16H 30/20 (2018.01)

G16H 30/40 (2018.01)

G16H 50/50 (2018.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711090729
과제번호	2019R1A2C1006543
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)(R&D)
연구과제명	멀티-모달, 멀티-스케일, 멀티-도메인 뇌 연결망 분석 및 설명 가능한 알츠하이머병
조기 진단 기술 개발	
기 여 율	2/3
과제수행기관명	고려대학교 산학협력단
연구기간	2021.01.01 ~ 2021.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711126110
과제번호	2019-0-00079-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	인공지능대학원지원(고려대학교)
기 여 율	1/3
과제수행기관명	고려대학교 산학협력단
연구기간	2021.01.01 ~ 2021.12.31

명세서

청구범위

청구항 1

3차원 의료 영상을 3개의 평면으로 나눈 후 각 평면마다 2차원의 순차적 데이터로 표현한 후 3차원 공간 특징 정보로 통합하도록 사전 표현 학습(Representation learning)되는 중축 모델;

완전 연결 계층으로 구성되며, 상기 3차원 공간 특징 정보를 기반으로 목표 태스크에 대한 예측을 수행하는 예측 모듈을 포함하되,

상기 사전 학습된 중축 모듈과 상기 예측 모듈은 상기 목표 태스크에 따라 미세 조정되는 것을 특징으로 하는 자가학습 기반 전이 학습 프레임워크 장치.

청구항 2

제1 항에 있어서,

상기 중축 모듈은,

상기 각 평면에 대한 2차원 슬라이스 영상에 대한 인코딩 벡터를 추출하도록 자가 학습 방식으로 학습되는 2차원 합성곱 신경망 모듈; 및

상기 인코딩 벡터를 콤비네이션하여 상기 3차원 공간 특징 정보로 통합하도록 표현 학습하는 트랜스포머를 포함하되,

상기 2차원 합성곱 신경망 모듈은 트리플렛 손실 함수를 기반으로 사전 학습되는 것을 특징으로 하는 자가학습 기반 전이 학습 프레임워크 장치.

청구항 3

제2 항에 있어서,

상기 트랜스포머는 2차원 슬라이스 영상 중 일부를 무작위로 마스킹처리한 후 마스킹되지 않은 슬라이스의 인코딩 벡터를 이용하여 상기 마스킹된 슬라이스 영상의 인코딩 벡터를 예측하도록 학습되는 것을 특징으로 하는 자가학습 기반 전이 학습 프레임워크 장치.

청구항 4

제3 항에 있어서,

상기 트랜스포머 학습시, 상기 2차원 합성곱 신경망 모듈은 학습되지 않도록 동결되는 것을 특징으로 하는 자가학습 기반 전이 학습 프레임워크 장치.

청구항 5

제2 항에 있어서,

상기 사전 학습된 중축 모듈과 상기 예측 모듈은 엔드투엔드방식으로 각 목표 태스크에 맞도록 미세 조정되며,

상기 목표 태스크는 뇌 질환 진단, 뇌 나이 예측 및 뇌 종양 분할 중 적어도 하나인 것을 특징으로 하는 자가학습 기반 전이 학습 프레임워크 장치.

청구항 6

제5 항에 있어서,

상기 예측 모듈은,

상기 목표 태스크에 따라 상기 2차원 합성곱 신경망 모듈의 높은 수준(high-level) 인코딩 벡터를 이용하는 싱글 스케일(single-scale) 기법과 상기 2차원 합성곱 신경망 모듈의 낮은 수준(low-level) 인코딩 벡터들을 통합하는 다중 스케일(multi-scale) 기법 중 어느 하나가 적용되는 것을 특징으로 하는 자가학습 기반 전이 학습 프레임워크 장치.

청구항 7

3차원 의료 영상을 3개의 평면으로 나눈 후 각 평면마다 2차원의 순차적 데이터로 표현한 후 3차원 공간 특징 정보로 통합하도록 모델을 사전 표현 학습(Representation learning)하는 단계; 및

상기 3차원 공간 특징 정보를 기반으로 목표 태스크에 대한 예측을 수행하도록 상기 목표 태스크에 따라 상기 모델을 미세 조정하는 단계를 포함하는 자가학습 기반 전이 학습 방법.

발명의 설명

기술 분야

[0001] 본 발명은 3차원 의료 영상 분석을 위한 자가학습 기반 전이 학습 방법 및 그 장치에 관한 것이다.

배경 기술

[0003] 전이 학습이란 학습 데이터가 부족한 모델 구축을 위해 데이터가 풍부한 분야에서 훈련된 모델을 재사용하는 학습 기법을 말한다. 이는 특정한 태스크에 대하여 학습한 모델을 다른 태스크로 전이하여 해당 모델을 사후적으로 학습하는 개념을 포괄한다.

[0004] 전이 학습을 통해 학습 시간이 비교적 오래 걸리는 딥러닝 모델의 학습 시간을 줄여주고, 적은 수의 학습 샘플만으로 높은 예측 성능을 보임으로써 딥러닝의 다양한 분야에서 연구가 활발히 이루어지고 있다.

[0005] 최근 라벨링된 샘플이 적고 고차원인 의료 영상 분석 분야에서도 전이 학습이 주목받고 있는데, 이를 통해, 영상의 분류(Classification), 분할(Segmentation) 등의 다양한 의료 영상 분석 관련 태스크들에 활발하게 활용되고 있다.

[0006] 이러한 의료 영상 분석을 위한 전이 학습 프레임워크는 대부분 3차원 합성곱 신경망 (Convolutional Neural Network, CNN) 네트워크를 지도 학습(Supervised learning) 또는 자가지도 학습 방식을 통해 사전 학습하여 활용하고 있다.

[0007] 하지만 지도 학습 기반 사전 학습 방식은 학습을 위해 많은 수의 라벨링된 학습 데이터셋을 요구하는 문제점이 있다. 또한 3차원 합성곱 신경망 네트워크 기반 방법들은 모델 학습에 필요한 매개변수 수가 많아 실제로 고차원의 의료 영상 데이터셋을 학습시키기 어려운 한계가 존재한다.

발명의 내용

해결하려는 과제

[0009] 본 발명은 라벨링된 학습 데이터셋이 부족한 3차원 의료 영상 분석 분야에서 적은 학습 샘플만으로 높은 성능을

도출할 수 있는 자가학습 기반 전이 학습 방법 및 그 장치를 제공하기 위한 것이다.

[0010] 또한, 본 발명은 3차원 의료 영상을 3개의 평면(Plane)으로 나누고, 각 평면마다 2차원의 순차적 데이터로 표현한 후, 3차원 공간 특징으로 통합하여 표현 학습(Representation learning)함으로써, 3차원 공간 특징을 효과적으로 학습할 뿐만 아니라 학습 매개변수를 효율적으로 줄일 수 있는 자가학습 기반 전이 학습 방법 및 그 장치를 제공하기 위한 것이다.

과제의 해결 수단

- [0012] 본 발명의 일 측면에 따르면, 3차원 의료 영상 분석을 위한 자가학습 기반 전이 학습을 위한 장치가 제공된다.
- [0013] 본 발명의 일 실시예에 따르면, 3차원 의료 영상을 3개의 평면으로 나눈 후 각 평면마다 2차원의 순차적 데이터로 표현한 후 3차원 공간 특징 정보로 통합하여 사전 표현 학습(Representation learning)하는 중측 모델; 완전 연결 계층으로 구성되며, 상기 3차원 공간 특징 정보를 기반으로 목표 태스크에 대한 예측을 수행하는 예측 모듈을 포함하되, 상기 사전 학습된 중측 모듈과 상기 예측 모듈은 상기 목표 태스크에 따라 미세 조정되는 것을 특징으로 하는 자가학습 기반 전이 학습 프레임워크 장치가 제공될 수 있다.
- [0014] 상기 중측 모듈은, 상기 각 평면에 대한 2차원 슬라이스 영상에 대한 인코딩 벡터를 추출하도록 자가 학습 방식으로 학습되는 2차원 합성곱 신경망 모듈; 및 상기 인코딩 벡터를 콤비네이션하여 상기 3차원 공간 특징 정보로 통합하도록 표현 학습하는 트랜스포머를 포함하되, 상기 2차원 합성곱 신경망 모듈은 트리플렛 손실 함수를 기반으로 사전 학습될 수 있다.
- [0015] 상기 트랜스포머는 2차원 슬라이스 영상 중 일부를 무작위로 마스크처리한 후 마스킹되지 않은 슬라이스의 인코딩 벡터를 이용하여 상기 마스킹된 슬라이스 영상의 인코딩 벡터를 예측하도록 학습될 수 있다.
- [0016] 상기 트랜스포머 학습시, 상기 2차원 합성곱 신경망 모듈은 학습되지 않도록 동결될 수 있다.
- [0017] 상기 사전 학습된 중측 모듈과 상기 예측 모듈은 엔드투엔드방식으로 각 목표 태스크에 맞도록 미세 조정되되, 상기 목표 태스크는 뇌 질환 진단, 뇌 나이 예측 및 뇌 종양 분할 중 적어도 하나일 수 있다.
- [0018] 상기 예측 모듈은, 상기 목표 태스크에 따라 상기 2차원 합성곱 신경망 모듈의 높은 수준(high-level) 인코딩 벡터를 이용하는 싱글 스케일(single-scale) 기법과 상기 2차원 합성곱 신경망 모듈의 낮은 수준(low-level) 인코딩 벡터들을 통합하는 다중 스케일(multi-scale) 기법 중 어느 하나가 적용될 수 있다.
- [0020] 본 발명의 다른 측면에 따르면, 3차원 의료 영상 분석을 위한 자가학습 기반 전이 학습 방법이 제공된다.
- [0021] 본 발명의 일 실시예에 따르면, 3차원 의료 영상을 3개의 평면으로 나눈 후 각 평면마다 2차원의 순차적 데이터로 표현한 후 3차원 공간 특징 정보로 통합하도록 모델을 사전 표현 학습(Representation learning)하는 단계; 및 상기 3차원 공간 특징 정보를 기반으로 목표 태스크에 대한 예측을 수행하도록 상기 목표 태스크에 따라 상기 모델을 미세 조정하는 단계를 포함하는 자가학습 기반 전이 학습 방법이 제공될 수 있다.

발명의 효과

- [0023] 본 발명의 일 실시예에 따른 3차원 의료 영상 분석을 위한 자가학습 기반 전이 학습 방법 및 그 장치를 제공함으로써, 라벨링된 학습 데이터셋이 부족한 3차원 의료 영상 분석 분야에서 적은 학습 샘플만으로 높은 성능을 도출할 수 있다.
- [0024] 또한, 본 발명은 3차원 의료 영상을 3개의 평면(Plane)으로 나누고, 각 평면마다 2차원의 순차적 데이터로 표현한 후, 3차원 공간 특징으로 통합하여 표현 학습(Representation learning)함으로써, 3차원 공간 특징을 효과적으로 학습할 뿐만 아니라 학습 매개변수를 효율적으로 줄일 수 있는 이점도 있다.

도면의 간단한 설명

- [0026] 도 1은 본 발명의 일 실시예에 따른 자가학습 기반 전이 학습 프레임워크 장치의 내부 구성을 개략적으로 도시한 블록도.
- 도 2는 본 발명의 일 실시예에 따른 중측 모듈의 내부 구성을 도시한 도면.
- 도 3은 본 발명의 일 실시예에 따른 2차원 합성곱 신경망 모듈의 트리플렛 손실 기반 학습을 설명하기 위해 도시한 도면.
- 도 4는 본 발명의 일 실시예에 따른 미세 조정을 설명하기 위해 도시한 도면.
- 도 5는 본 발명의 일 실시예에 따른 자가학습 기반 전이 학습 방법을 나타낸 순서도.
- 도 6은 본 발명의 일 실시예에 따른 자가학습 기반 전이 학습 프레임워크의 전체 플로우를 도시한 도면.
- 도 7은 종래와 본 발명의 뇌 종양 분할 결과를 비교한 도면.

발명을 실시하기 위한 구체적인 내용

- [0027] 본 명세서에서 사용되는 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "구성된다" 또는 "포함한다" 등의 용어는 명세서상에 기재된 여러 구성 요소들, 또는 여러 단계들을 반드시 모두 포함하는 것으로 해석되지 않아야 하며, 그 중 일부 구성 요소들 또는 일부 단계들은 포함되지 않을 수도 있고, 또는 추가적인 구성 요소 또는 단계들을 더 포함할 수 있는 것으로 해석되어야 한다. 또한, 명세서에 기재된 "...부", "모듈" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어 또는 소프트웨어로 구현되거나 하드웨어와 소프트웨어의 결합으로 구현될 수 있다.
- [0028] 이하, 첨부된 도면들을 참조하여 본 발명의 실시예를 상세히 설명한다.
- [0030] 도 1은 본 발명의 일 실시예에 따른 자가학습 기반 전이 학습 프레임워크 장치의 내부 구성을 개략적으로 도시한 블록도이고, 도 2는 본 발명의 일 실시예에 따른 중측 모듈의 내부 구성을 도시한 도면이고, 도 3은 본 발명의 일 실시예에 따른 2차원 합성곱 신경망 모듈의 트리플렛 손실 기반 학습을 설명하기 위해 도시한 도면이며, 도 4는 본 발명의 일 실시예에 따른 미세 조정을 설명하기 위해 도시한 도면이다.
- [0031] 도 1을 참조하면, 본 발명의 일 실시예에 따른 자가학습 기반 전이 학습 프레임워크 장치(100)는 중측 모듈(110), 예측 모듈(120), 메모리(130) 및 프로세서(140)를 포함하여 구성된다.
- [0032] 본 발명의 일 실시예에 따르면, 자가학습 기반 전이 학습 프레임워크 장치(100)는 SSL 프록시 작업에 의해 사전 훈련되는 중측 모듈(110)과 최종 예측 위한 완전 연결 계층 기반 예측 모듈(120)을 포함할 수 있다.
- [0033] 중측 모듈(110)은 3차원 의료 영상을 각 평면으로 나눈 후 각 평면에 대한 2차원 슬라이스의 고수준 공간 특징을 추출하여 슬라이스간 종속성을 모델링하여 관계형 특징을 추출할 수 있다.
- [0034] 3차원 의료 영상은 각 평면에 대해 서로 다른 뷰를 갖는 축, 관상, 시상 3개의 평면으로 구성되기 때문에 중측 모듈(110)은 매개 변수를 공유하지 않는 각 평면에 대해 독립적인 합성곱 인코더가 사용될 수 있다.
- [0035] 합성곱 인코더(212)에 의해 추출된 공간 특징은 트랜스포머(220)로 공급되기 전에 슬라이스 순서 지정을 위한 위치 인코딩과 식별 평면에 대한 세그먼트 인코딩이 수행될 수 있다.
- [0036] 트랜스포머(220)를 통과한 후 모든 평면에 대한 슬라이스별 특징의 조합으로 3차원 공간 특징 맵(정보)이 형성되며, 예측 모듈(120)은 해당 3차원 공간 특징 맵을 이용하여 목표 태스크에 따른 예측을 수행할 수 있다.
- [0037] 이하에서 보다 상세히 설명하기로 한다.
- [0038] 본 발명의 일 실시예에 따르면, 중측 모듈(110)은 3차원 의료 영상을 3개의 평면으로 각각 나눈 후 각 평면에 대한 2차원의 순차적인 데이터로 표현한 후 이를 3차원 공간 특징 정보로 통합하도록 사전 표현 학습된다.
- [0039] 중측 모듈(110)의 사전 학습이 완료되면, 사전 학습 완료된 중측 모듈(110)과 예측 모듈(120)이 목표 태스크에 적합하도록 미세 조정될 수 있다.
- [0040] 예측 모듈(120)은 완전 연결 계층으로 구성되며, 사전 학습된 중측 모듈(110)과 예측 모듈은 엔드투엔드 방식으로 목표 태스크에 적합하도록 미세 조정될 수 있다.

[0041] 도 2에 중축 모듈(110)의 내부 구조가 도시되어 있다. 도 2에 도시된 바와 같이, 중축 모듈(110)은 2차원 합성곱 신경망 모듈(210)과 트랜스포머(220)를 포함하여 구성된다.

[0042] 2차원 합성곱 신경망 모듈(210)은 3차원 의료 영상을 3개의 평면으로 나눈 후 각 평면에 대해 생성된 2차원 슬라이스 영상의 공간적 특징으로 슬라이스 인코딩 벡터(slice encoding vector)들을 추출한다.

[0043] 트랜스포머(220)는 각 평면에 상응하는 슬라이스 인코딩 벡터(slice encoding vector)들을 콤비네이션하여 3차원 공간 특징 정보로 통합하기 위한 수단이다.

[0044] 예를 들어, 3차원 의료 영상을 $\mathcal{X}^{pl} (pl \in \{\text{cor, sag, axial}\})$ 라 가정하면, 각 평면의 2D 슬라이스 영상의 집합을 $\mathcal{X}^{pl} = \{X_1^{pl}, \dots, X_{N_{pl}}^{pl}\}$ 로 나타낼 수 있다. $(\mathcal{X}^{pl} \in \mathbb{R}^{w \times d \times h})$ 이다. N_{pl} 는 상응하는 평면의 슬라이스 개수를 나타내고, w, d, h는 3차원 의료 영상의 너비, 깊이, 및 높이를 각각 나타낸다.

[0045] 2차원 슬라이스 영상이 2차원 합성곱 신경망 모듈(210)로 전달되며, 2차원 합성곱 신경망 모듈(210)은 각 평면에 대한 슬라이스 영상에 대한 임베딩 벡터 $Z^{pl} = \{z_1^{pl}, \dots, z_{N_{pl}}^{pl}\} (z_n^{pl} \in \mathbb{R}^{d_{\text{emb}}})$ 를 추출할 수 있다. 여기서, d_{emb} 는 임베딩 벡터의 차원을 나타낸다.

[0046] 이어, 2차원 합성곱 신경망 모듈(210)은 각 평면의 슬라이스 영상에 대한 임베딩 벡터에 대해 각각 위치 및 세그먼트 인코딩을 수행하여 인코딩 벡터 $E^{pl} = \{e_1^{pl}, \dots, e_{N_{pl}}^{pl}\} (e_n^{pl} \in \mathbb{R}^{d_{\text{emb}}})$ 를 생성할 수 있으며, 해당 인코딩 벡터가 트랜스포머(220)로 전달될 수 있다.

[0047] 본 발명의 일 실시예에 따른 중축 모듈(110)은 사전 학습될 수 있다. 이하에서는 중축 모듈(110)의 사전 학습 과정에 대해 우선 설명하기로 한다.

[0048] 중축 모듈(110)은 두번의 자가 학습에 기반하여 사전 학습될 수 있다.

[0049] 도 3을 참조하여 이에 대해 보다 상세히 설명하기로 한다.

[0050] 2차원 합성곱 신경망 모듈(210)이 공간적으로 의미 있는 임베딩 벡터를 추출하도록 자가 학습 방식으로 2차원 합성곱 신경망 모듈(210)이 학습될 수 있다.

[0051] 2차원 합성곱 신경망 모듈(210)은 합성곱 인코더(212)를 포함하되, 합성곱 인코더(212)는 트리플렛 손실 함수를 기반으로 학습될 수 있다. 각 평면에 대해 모든 슬라이스별 특징이 평균 풀링되고, 포지티브 페어(변환된 특징)와 네거티브 페어(다른 배치로부터의 샘플 특징)와 비교하여 수학적 1과 같이 트리플렛 손실($\mathcal{L}^{pl}$)이 계산될 수 있다.

수학적식 1

$$\mathcal{L}_{pl} = \frac{1}{N_T} \sum_{i=1}^{N_T} \max\{0, D(\tilde{z}_{(i)}^{pl}, \tilde{z}_{(i)}^{pl,+}) - D(\tilde{z}_{(i)}^{pl}, \tilde{z}_{(i)}^{pl,-}) + \alpha\}$$

[0052] 여기서, $\tilde{z}_{(i)}^{pl}$ 는 슬라이스에 대해 평균을 낸 랜덤 훈련 샘플의 임베딩 벡터를 나타내고, $\tilde{z}_{(i)}^{pl,+}$ 는 트랜스포머 버전(포지티브 예제)을 나타내고, $\tilde{z}_{(i)}^{pl,-}$ 는 데이터세트로부터의 다른 샘플(네거티브 예제)을 나타내며, N_T 는 훈련된 샘플의 개수를 나타내고, α 는 포지티브 샘플과 네거티브 샘플 사이의 마진을 나타내며, D는 거리

측정을 위한 함수로 예를 들어, L2거리일 수 있다. 최종 손실은 세 평면의 모든 손실을 평균하여 정의될 수 있다.

[0054] 합성곱 인코더(212)에 의해 각 평면에 대한 2차원 슬라이스 공간 특징(임베딩 벡터)가 추출된 후, 슬라이스 순서 지정을 위한 위치 인코딩과 식별 평면에 대한 세그먼트 인코딩이 수행되어 출력값인 인코딩 벡터가 트랜스포머(220)로 전달될 수 있다. 슬라이스 단위 인코딩 벡터 e_n^{pl} 를 토큰으로 간주하고, 이미지 수준 특징 집합인 인코딩 벡터의 집합 $E^{pl} = \{e_1^{pl}, \dots, e_{N_{pl}}^{pl}\}$ 을 문장으로 간주할 수 있다.

[0055] 트리플렛 손실 함수를 기반으로 2차원 합성곱 신경망 모듈(210)의 학습의 학습 수행된 후 학습된 합성곱 신경망 모듈(210)이 학습되지 않도록 동결시킨 후 임의로 마스킹한 인코더 벡터를 예측하는 마스킹된 인코딩 벡터 예측 태스크를 통해 트랜스포머(220)를 학습을 수행할 수 있다.

[0056] 이와 같이, 2차원 합성곱 신경망 모듈(210)이 사전 학습된 후, 트랜스포머(220)가 학습될 수 있다.

[0057] 본 발명의 일 실시예에 따른 트랜스포머(220)의 사전 학습시, 각 평면에 대한 일부 슬라이스 영상을 무작위로 마스킹하고, 나머지 마스킹되지 않은 슬라이스 영상에 대한 인코딩 벡터에서 마스킹된 슬라이스 영상의 인코딩 벡터를 예측하도록 트랜스포머(220)가 학습될 수 있다.

[0058] 이때, 이전 단계에서 사전 훈련된 2차원 합성곱 신경망 모듈(210)은 고정되며, 트랜스포머(220)만 마스킹된 인코딩 벡터를 예측하도록 훈련되므로 트랜스포머(220)는 슬라이스간 종속성을 파악하여 마스킹된 인코딩 벡터를 예측할 수 있다.

[0059] 예를 들어, 슬라이스 단위 인코딩 벡터를

$$E^p = \{e_n^p \in \mathbb{R}^{d_{emb}}\}_{n=1}^{N_p}, \quad e_n^p = z_n^p + PSE(n, p)$$

라고 가정하기로 한다.

[0060] 트랜스포머(220)는 집중 매커니즘을 통해 슬라이스간 관계(종속성)를 모델링할 수 있다.

[0061] 이를 수학식으로 나타내면, 수학식 2와 같이 나타낼 수 있다.

수학식 2

$$\bar{E}^p = E^p + \text{Softmax} \left(\frac{(W_Q E^p)(W_K E^p)^T}{\sqrt{d_{emb}}} \right) (W_V E^p)$$

$$\hat{E}^p = \bar{E}^p + \text{ReLU}(\bar{E}^p F_1) F_2$$

[0062]

[0063] 여기서, $\{W_Q, W_K, W_V\}$ 는 학습 가능한 가중치 행렬의 집합을 나타내고, $F_1 \in \mathbb{R}^{d_{emb} \times d_{ff}}$ 와 $F_2 \in \mathbb{R}^{d_{ff} \times d_{emb}}$ 는 두개의 위치 단위 피드포워드 합성곱(point-wise feed-forward convolutions)을 나타낸다.

[0064] 마스킹된 인코딩 벡터 예측에 대한 손실 함수는 수학식 3과 같이 나타낼 수 있다.

수학식 3

$$\mathcal{L}_{mask} = \frac{1}{|M|} \sum_{m \in M} (\hat{e}_m^p - z_m)^2$$

[0065]

- [0066] 여기서, M 는 마스킹된 슬라이스들의 인덱스 세트를 나타낸다.
- [0067] 이와 같이, 중축 모듈(110)에 대한 사전 학습이 완료된 후 사전 학습된 중축 모듈(110)과 예측 모듈(120)이 미세 조정될 수 있다.
- [0068] 이미 전술한 바와 같이, 사전 학습된 중축 모듈(110)과 예측 모듈(120)은 엔드투엔드방식으로 목표 태스크에 따라 미세 조정될 수 있다.
- [0069] 이에 대해 도 4를 참조하여 보다 상세히 설명하기로 한다.
- [0070] 목표 태스크는 뇌 질환 진단, 뇌 나이 예측 및 뇌 종양 분할 중 어느 하나일 수 있다.
- [0071] 즉, 사전 학습된 중축 모듈(110)과 예측 모듈(120)은 뇌 질환 진단, 뇌 나이 예측 및 뇌 종양 분할 중 어느 하나에 따라 상이하게 미세 조정될 수 있다.
- [0072] 목표 태스크에 따라 낮은 수준(low-level) 및 높은 수준(high-level) 특징을 모두 캡처하는 멀티 스케일 접근 방식과 분류 및 회귀 작업에 대한 높은 수준의 특징을 캡처하는 싱글 스케일 접근 방식이 취해질 수 있다.
- [0073] 사전 학습 및 미세 조정을 위한 전체 플로우는 도 4에 도시된 바와 같다.
- [0074] 메모리(130)는 본 발명의 일 실시예에 따른 자가학습 기반 전이 학습 방법을 위한 다양한 명령어들을 저장한다.
- [0075] 프로세서(140)는 본 발명의 일 실시예에 따른 자가학습 기반 전이 학습 프레임워크 장치(100)의 내부 구성 요소들(예를 들어, 중축 모듈(110), 예측 모듈(120), 메모리(130) 등)을 제어하기 위한 수단이다.
- [0076] 또한, 프로세서(140)는 중축 모듈(110)이 사전 학습되도록 제어하며, 사전 학습된 중축 모듈(110)과 예측 모듈(120)이 목표 태스크에 따라 미세 조정되도록 제어할 수 있다.
- [0078] 도 5는 본 발명의 일 실시예에 따른 자가학습 기반 전이 학습 방법을 나타낸 순서도이고, 도 6은 본 발명의 일 실시예에 따른 자가학습 기반 전이 학습 프레임워크의 전체 플로우를 도시한 도면이고, 도 7은 종래와 본 발명의 뇌 종양 분할 결과를 비교한 도면이다.
- [0079] 단계 510에서 자가학습 기반 전이 학습 프레임워크 장치(100)는 백본 네트워크를 사전 학습시킨다.
- [0080] 도 6에 도시된 바와 같이, 백본 네트워크는 중축 모듈(110)로, 2차원 합성곱 신경망 모듈(210)과 트랜스포머(220)를 포함하여 구성된다. 2차원 합성곱 신경망 모듈(210)은 3차원 의료 영상의 각 평면에 대한 2차원 슬라이스의 고수준 공간 특징을 추출하고, 트랜스포머(220)를 통해 슬라이스간 종속성을 모델링하여 관계형 특징을 추출하도록 사전학습될 수 있다.
- [0081] 3차원 의료 영상은 각 평면에 대해 서로 다른 뷰를 갖는 축, 관상 및 시상 3개의 평면으로 구성되므로 매개변수를 공유하지 않는 각 평면에 대해 독립적인 2차원 합성곱 신경망 모듈(210)을 사용한다.
- [0082] 공간 특징은 트랜스포머(220)로 전달되기 전에 슬라이스의 순서 지정을 위한 위치 인코딩과 식별 평면에 대한 세그먼트 인코딩을 거친 후 트랜스포머(220)로 전달될 수 있다.
- [0083] 트랜스포머(220)는 자가주의 매커니즘을 통해 인접 슬라이스와 원거리 슬라이스간 종속성을 캡처하도록 학습될 수 있다. 즉, 트랜스포머(220)는 마스킹된 일부 슬라이스의 인코딩 벡터를 마스킹되지 않은 슬라이스의 인코딩 벡터를 이용하여 예측하도록 학습될 수 있다.
- [0084] 이미 전술한 바와 같이, 트랜스포머(220)의 사전 학습시, 2차원 합성곱 신경망 모듈(210)은 동결되어 학습되지 않도록 할 수 있다.
- [0085] 단계 515에서 자가학습 기반 전이 학습 프레임워크 장치(100)는 사전 학습된 중축 모듈(110)과 예측 모듈(120)을 목표 태스크에 따라 미세 조정한다.
- [0086] 예를 들어, 목표 태스크는 뇌 질환 진단, 뇌 나이 예측 및 뇌 종양 분할일 수 있다.
- [0087] 사전 학습된 중축 모듈(110)과 예측 모듈(120)은 목표 태스크에 따라 낮은 수준(low-level) 특징과 높은 수준(high-level)의 특징을 모두 이용하는 멀티 스케일 접근 방식과 높은 수준의 특징을 이용하는 싱글 스케일 접근

방식이 이용될 수 있다.

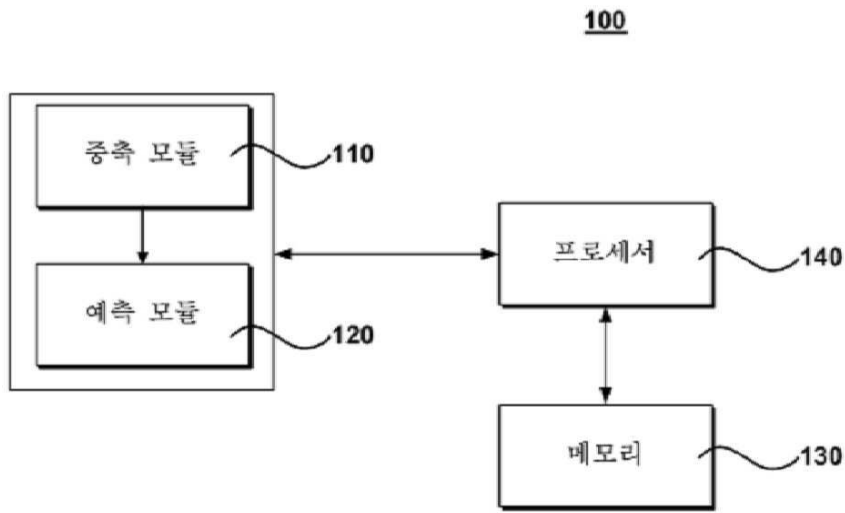
- [0088] 따라서, 목표 태스크에 따라 사전 학습된 중축 모듈(110)과 예측 모듈(120)이 엔드투엔드 방식으로 미세 조정될 수 있다.
- [0089] 이와 같이, 미세 조정이 완료되면, 중축 모듈(110)을 통해 3차원 의료 영상을 각 평면으로 나눈 후 각 평면에 대한 2차원 슬라이스의 고수준 공간 특징을 추출한 후 이를 조합하여 3차원 공간 특징 정보를 형성할 수 있다.
- [0090] 이어, 예측 모듈(120)은 3차원 공간 특징 정보를 이용하여 목표 태스크를 예측할 수 있다.
- [0091] 도 7은 종래와 본 발명의 일 실시예에 따른 3차원 의료 영상에 대한 뇌 종양 분할 결과를 비교한 것이다. 본 발명의 뇌 종양 분할 결과가 3D-SSL의 분할 결과와 유사하며 종래와 비교하여 종양 유형이 가장 잘 구별되는 것을 알 수 있다.
- [0093] 본 발명의 실시 예에 따른 장치 및 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 컴퓨터 판독 가능 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 분야 통상의 기술자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media) 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.
- [0094] 상술한 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0095] 이제까지 본 발명에 대하여 그 실시 예들을 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. 그러므로 개시된 실시 예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

부호의 설명

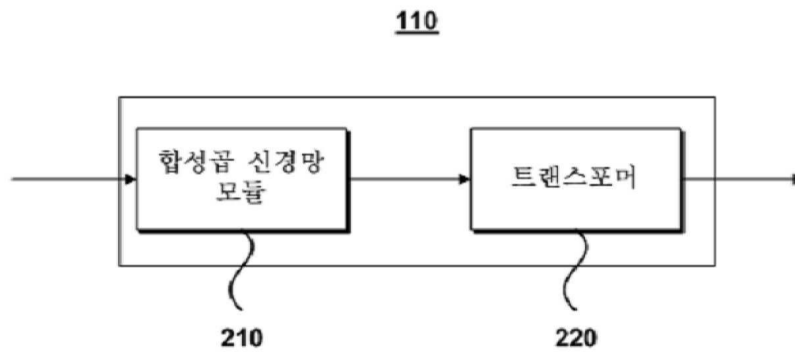
- [0097] 100: 자가학습 기반 전이 학습 프레임워크 장치.
- 110: 중축 모듈
- 120: 예측 모듈
- 130: 메모리
- 140: 프로세서

도면

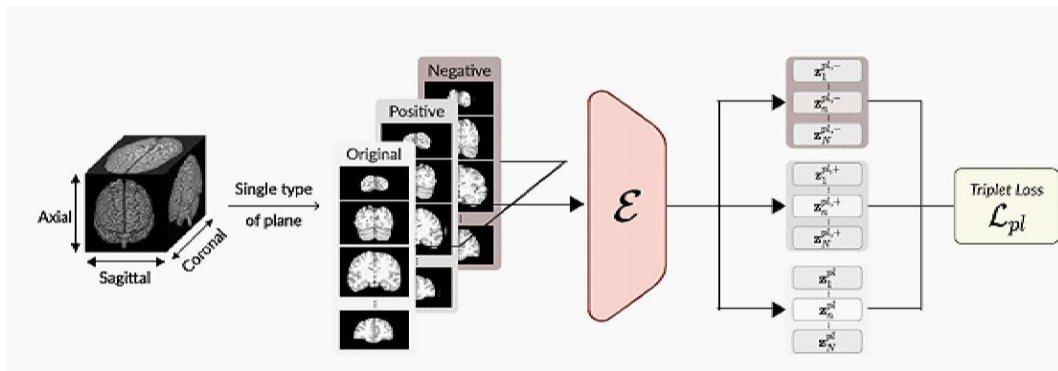
도면1



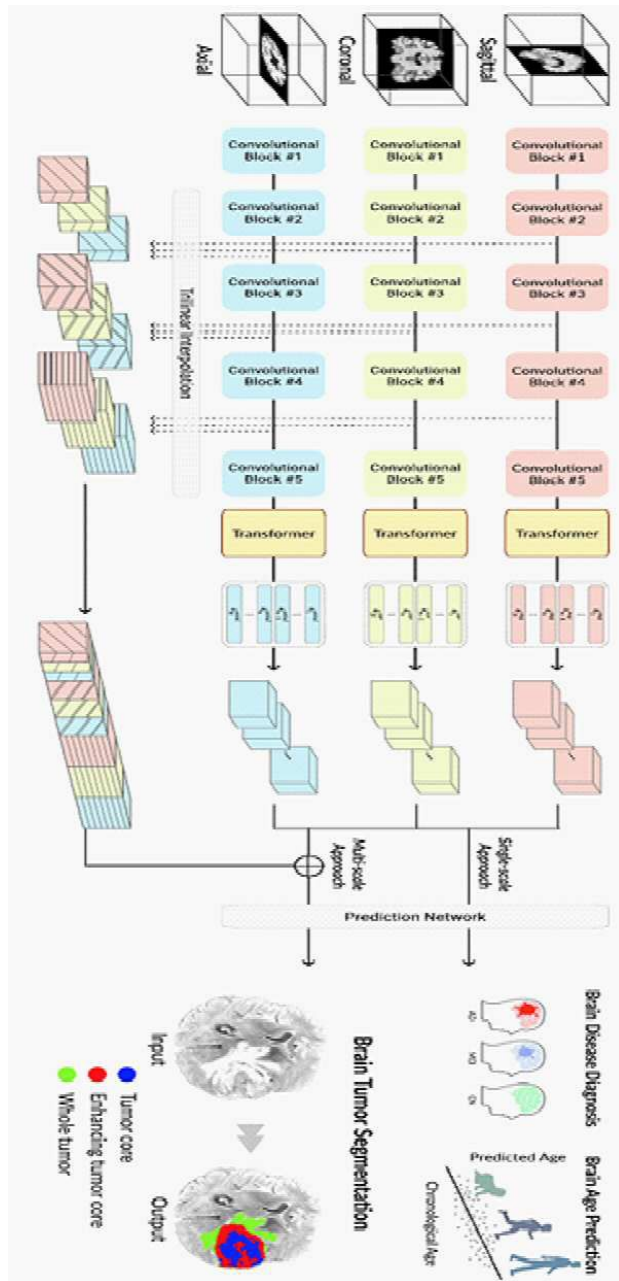
도면2



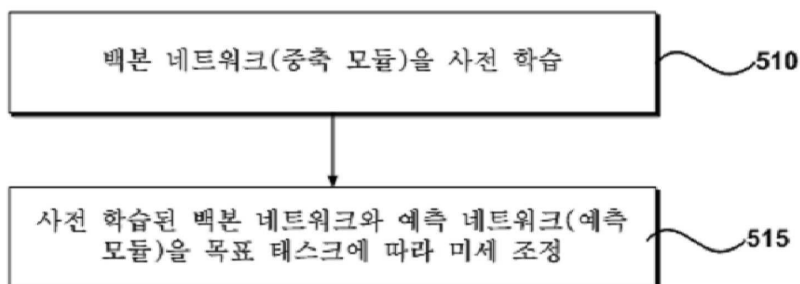
도면3



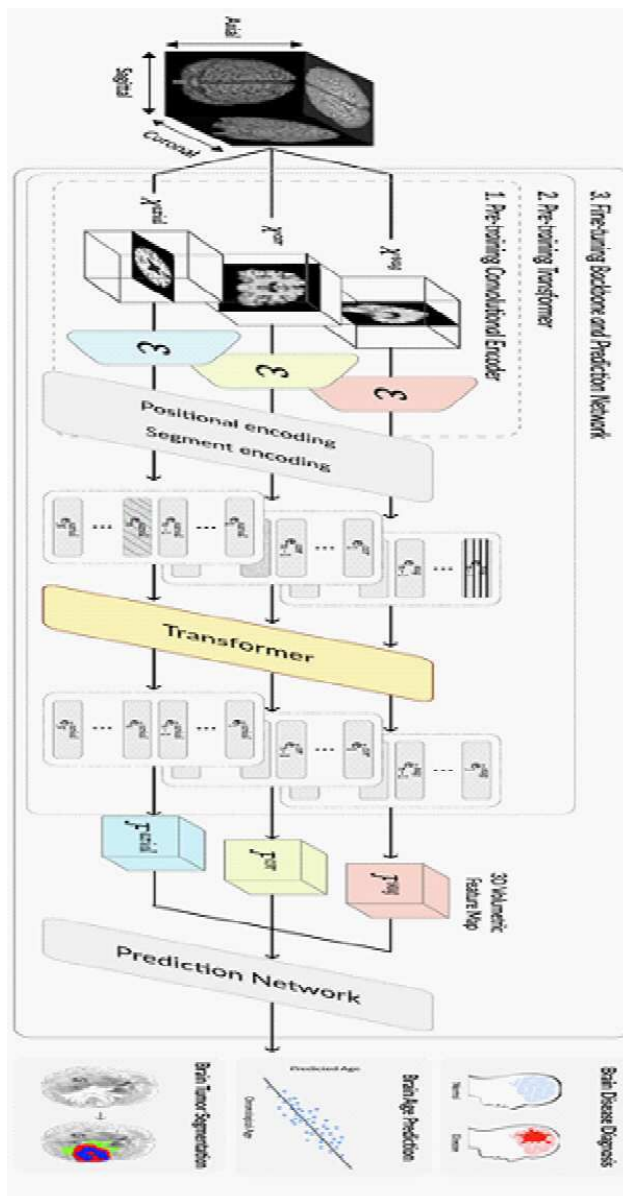
도면4



도면5



도면6



도면7

