(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0026045
(43) 공개일자 2025년02월25일

(51) 국제특허분류(Int. Cl.)
G06N 3/094 (2023.01) *G06N 3/0455* (2023.01)
G06N 3/0464 (2023.01) *G06N 3/0475* (2023.01)
G06N 3/08 (2023.01)
(52) CPC특허분류
G06N 3/094 (2023.01)
G06N 3/0455 (2023.01)
(21) 출원번호 10-2023-0107233
(22) 출원일자 2023년08월16일
심사청구일자 2023년08월16일

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
신병석
서울특별시 서초구
김지온
인천광역시 미추홀구 재넘이길 86, 207호(용현동)
(74) 대리인
양성보

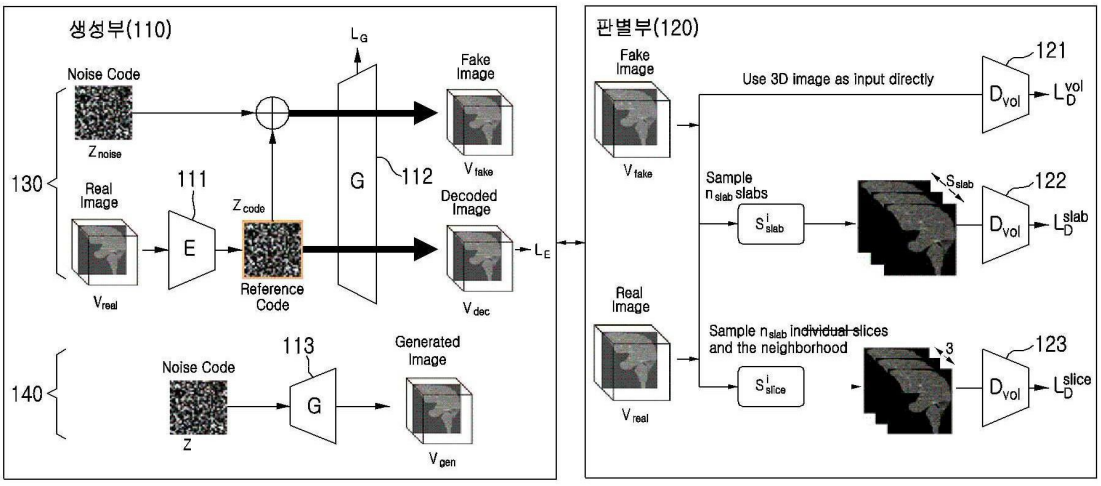
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 **의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 방법 및 시스템**

(57) 요 약

의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 방법 및 시스템이 제시된다. 본 발명에서 제안하는 의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 시스템은 생성부 및 판별부를 포함하고, 상기 생성부는 실제 이미지를 입력으로 받아 특징을 추출하여 참조 코드를 생성하고, 상기 참조 코드를 생성기에 제공하는 인코더 및 상기 참조 코드 및 노이즈 코드를 입력으로 받아 가상 이미지를 생성하기 위한 학습을 수행하는 생성기를 포함하고, 상기 판별부는 상기 생성된 가상 이미지와 실제 이미지를 구별하기 위한 학습을 수행하는 볼륨 판별기(volume discriminator), 슬랩 판별기(slab discriminator) 및 슬라이스 판별기(slice discriminator)를 포함하며, 상기 학습이 완료되면 상기 생성부의 생성기는 상기 노이즈 코드만을 입력으로 받아 이미지를 합성한다.

대 표 도



(52) CPC특허분류
G06N 3/0464 (2023.01)
G06N 3/0475 (2023.01)
G06N 3/088 (2023.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711182418
과제번호 2022R1A2B5B01001553
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 중견후속연구
연구과제명 [Ezbaro] 의료용 메타휴먼 응용을 위한 초고실감 인체모델링기술개발
기 여 율 45/100
과제수행기관명 인하대학교
연구기간 2023.03.01 ~ 2024.02.29
이 발명을 지원한 국가연구개발사업
과제고유번호 1711186835
과제번호 2022R1A4A1033549
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 기초연구실육성사업(BRL)
연구과제명 [Ezbaro] 인공지능 융합형 메타휴먼 플랫폼 연구실
기 여 율 45/100
과제수행기관명 인하대학교
연구기간 2023.03.01 ~ 2024.02.29
이 발명을 지원한 국가연구개발사업
과제고유번호 1711179347
과제번호 00155915
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 정보통신인재양성(교육운영수익)
연구과제명 [Ezbaro][정부+민간] 인공지능융합혁신인재양성(2차년도)
기 여 율 10/100
과제수행기관명 인하대학교 산학협력단
연구기간 2023.01.01 ~ 2023.12.31

명 세 서

청구범위

청구항 1

생성적 적대 네트워크의 학습 시스템에 있어서,

생성부 및 판별부를 포함하고,

상기 생성부는,

실제 이미지를 입력으로 받아 특징을 추출하여 참조 코드를 생성하고, 상기 참조 코드를 생성기에 제공하는 인코더; 및

상기 참조 코드 및 노이즈 코드를 입력으로 받아 가상 이미지를 생성하기 위한 학습을 수행하는 생성기

를 포함하고,

상기 판별부는,

상기 생성된 가상 이미지와 실제 이미지를 구별하기 위한 학습을 수행하는 볼륨 판별기(volume discriminator), 슬랩 판별기(slab discriminator) 및 슬라이스 판별기(slice discriminator)

를 포함하며,

상기 학습이 완료되면 상기 생성부의 생성기는 상기 노이즈 코드만을 입력으로 받아 이미지를 합성하는

생성적 적대 네트워크의 학습 시스템.

청구항 2

제1항에 있어서,

상기 생성부의 생성기는,

상기 참조 코드를 입력으로 받아 3D 전치 컨볼루션 레이어(transposed convolution layer)를 사용하여 3D 가상 이미지로 업샘플링하고,

학습 과정에서 상기 참조 코드와 가우시안 기반의 노이즈 코드의 평균을 입력으로 사용하여 3D 가상 이미지를 생성하기 위한 학습을 수행하는

생성적 적대 네트워크의 학습 시스템.

청구항 3

제1항에 있어서,

상기 생성부의 인코더는,

3D 컨볼루션 레이어를 사용하여 상기 실제 이미지에서 특징을 추출하여 참조 코드를 생성하고,

상기 참조 코드는 상기 노이즈 코드와의 합산 과정 없이 생성기에 입력되도록 하여 디코딩된 이미지를 생성하며,

상기 디코딩된 이미지와 상기 실제 이미지의 비교 과정을 통해 상기 생성된 참조 코드를 정확도를 확인하는

생성적 적대 네트워크의 학습 시스템.

청구항 4

제1항에 있어서,

상기 판별부는,

상기 볼륨 판별기, 슬랩 판별기 및 슬라이스 판별기를 포함하고,

상기 볼륨 판별기는 상기 3D 가상 이미지와 상기 실제 이미지의 전체 슬라이스를 입력으로 사용하며 3D 컨볼루션 레이어를 통해 실제 이미지 여부를 나타내는 확률을 추출하고,

상기 슬랩 판별기는 상기 3D 가상 이미지 및 상기 실제 이미지에서 추출한 특정 z축 범위의 연속된 슬라이스 집합인 슬랩을 사용하고,

상기 슬라이스 판별기는 상기 3D 가상 이미지와 상기 실제 이미지 내 랜덤 z-위치를 가지는 특정 슬라이스들과 상기 특정 슬라이스들과 인접한 슬라이스들을 사용하며,

상기 슬랩 판별기와 상기 슬라이스 판별기는 2D 컨볼루션 레이어를 사용하여 각각 상기 3D 가상 이미지와 상기 실제 이미지에서 추출된 슬랩과 슬라이스들을 판별하는

생성적 적대 네트워크의 학습 시스템.

청구항 5

인코더를 통해 실제 이미지를 입력으로 받아 특징을 추출하여 참조 코드를 생성하고, 상기 참조 코드를 생성기에 제공하는 단계;

생성기를 통해 상기 참조 코드 및 노이즈 코드를 입력으로 받아 가상 이미지를 생성하기 위한 학습을 수행하는 단계;

볼륨 판별기(volume discriminator), 슬랩 판별기(slab discriminator) 및 슬라이스 판별기(slice discriminator)를 포함하는 판별부를 통해 상기 생성된 가상 이미지와 실제 이미지를 구별하기 위한 학습을 수행하는 단계; 및

상기 학습이 완료되면 상기 생성기를 통해 상기 노이즈 코드만을 입력으로 받아 이미지를 합성하는 단계

를 포함하는 생성적 적대 네트워크의 학습 방법.

청구항 6

제5항에 있어서,

상기 인코더를 통해 실제 이미지를 입력으로 받아 특징을 추출하여 참조 코드를 생성하고, 상기 참조 코드를 생성기에 제공하는 단계는,

3D 컨볼루션 레이어를 사용하여 상기 실제 이미지에서 특징을 추출하여 참조 코드를 생성하고,

상기 참조 코드는 상기 노이즈 코드와의 합산 과정 없이 생성기에 입력되도록 하여 디코딩된 이미지를 생성하며,

상기 디코딩된 이미지와 상기 실제 이미지의 비교 과정을 통해 상기 생성된 참조 코드를 정확도를 확인하는

생성적 적대 네트워크의 학습 방법.

청구항 7

제5항에 있어서,

상기 생성기를 통해 상기 참조 코드 및 노이즈 코드를 입력으로 받아 가상 이미지를 생성하기 위한 학습을 수행하는 단계는,

상기 참조 코드를 입력으로 받아 3D 전치 컨볼루션 레이어(transposed convolution layer)를 사용하여 3D 가상 이미지로 업샘플링하고,

학습 과정에서 상기 참조 코드와 가우시안 기반의 노이즈 코드의 평균을 입력으로 사용하여 3D 가상 이미지를 생성하기 위한 학습을 수행하는

생성적 적대 네트워크의 학습 방법.

청구항 8

제5항에 있어서,

상기 볼륨 판별기, 슬랩 판별기 및 슬라이스 판별기를 포함하는 판별부를 통해 상기 생성된 가상 이미지와 실제 이미지를 구별하기 위한 학습을 수행하는 단계는,

상기 볼륨 판별기는 상기 3D 가상 이미지와 상기 실제 이미지의 전체 슬라이스를 입력으로 사용하며 3D 컨볼루션 레이어를 통해 실제 이미지 여부를 나타내는 확률을 추출하고,

상기 슬랩 판별기는 상기 3D 가상 이미지 및 상기 실제 이미지에서 추출한 특정 z축 범위의 연속된 슬라이스 집합인 슬랩을 사용하고,

상기 슬라이스 판별기는 상기 3D 가상 이미지와 상기 실제 이미지 내 랜덤 z-위치를 가지는 특정 슬라이스들과 상기 특정 슬라이스들과 인접한 슬라이스들을 사용하며,

상기 슬랩 판별기와 상기 슬라이스 판별기는 2D 컨볼루션 레이어를 사용하여 각각 상기 3D 가상 이미지와 상기 실제 이미지에서 추출된 슬랩과 슬라이스들을 판별하는

생성적 적대 네트워크의 학습 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 방법 및 시스템에 관한 것이다.

배 경 기 술

[0002] 현재 의료 영상은 개인 환자 정보의 프라이버시 문제로 인해 제한된 가용성(availability)을 가지고 있다. 의료 목적 이외의 용도로 개인 환자 정보를 사용할 수 없기 때문에 연구에 활용할 수 있는 데이터가 제한되며, 이로 인해 의료 영상 연구의 발전이 저해되고 있다. 제한된 가용성 문제를 해결하고 충분한 의료 영상을 확보하기 위해 GAN(Generative Adversarial Network) 기반 기법들이 활용되었다. 그러나 기존의 기법들은 주로 2D 이미지를 생성하여 3D 해부학적 구조와 인접 슬라이스 간의 연관성을 완전히 반영하는 데 어려움이 있다. 따라서 슬라이스 간에 불연속성이 있는 이미지가 생성되는 문제가 있다.

[0003] 이러한 문제를 해결하기 위해서 기존에 훈련 과정에서 3D 이미지를 활용하는 생성 네트워크가 제시되었다. 그러나 적은 3D 이미지가 훈련에 사용될 경우 결과가 대상 물체에 수렴하지 않고 발산할 수 있다. 그러므로 3D 이미지 생성 과정에서 수렴도(convergence rate)를 높이는 네트워크 구조가 필요하다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 한국등록특허 제10-2095097호(2020.03.24)

발명의 내용

해결하려는 과제

[0005] 본 발명이 이루고자 하는 기술적 과제는 3D 이미지 생성 프로세스의 수렴도(convergence rate)를 높이는 GAN(Generative Adversarial Network) 구조를 통해 의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 방법 및 시스템을 제공하는데 있다. 본 발명의 실시예들에 따르면 인코딩(encoding) 과정을 통해 실제 영상에서 특징(feature)을 추출하고 디코딩(decoding) 과정을 통해 특징 추출 과정의 정확도를 검증한다. 인코딩 과정에서 추출된 특징을 참조 코드(reference code)라고 명명하며, 이는 생성기(generator)에서 3D 이미지를 생성하는 데 사용된다. 또한 제안하는 기법은 판별기(discriminator) 구조를 3개의 판별기로 나누어 세밀한 판별을 수행하며 생성되는 이미지의 수렴도을 높이고자 한다

과제의 해결 수단

[0006] 일 측면에 있어서, 본 발명에서 제안하는 의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 시스템은 생성부 및 판별부를 포함하고, 상기 생성부는 실제 이미지를 입력으로 받아 특징을 추출하여 참조 코드를 생성하고, 상기 참조 코드를 생성기에 제공하는 인코더 및 상기 참조 코드 및 노이즈 코드를 입력으로 받아 가상 이미지를 생성하기 위한 학습을 수행하는 생성기를 포함하고, 상기 판별부는 상기 생성된 가상 이미지와 실제 이미지를 구별하기 위한 학습을 수행하는 볼륨 판별기(volume discriminator), 슬랩 판별기(slab discriminator) 및 슬라이스 판별기(slice discriminator)를 포함하며, 상기 학습이 완료되면 상기 생성부의 생성기는 상기 노이즈 코드만을 입력으로 받아 이미지를 합성한다.

[0007] 상기 생성부의 생성기는 상기 참조 코드를 입력으로 받아 3D 전치 컨볼루션 레이어(transposed convolution layer)를 사용하여 3D 가상 이미지로 업샘플링하고, 학습 과정에서 상기 참조 코드와 가우시안 기반의 노이즈 코드의 평균을 입력으로 사용하여 3D 가상 이미지를 생성하기 위한 학습을 수행한다.

[0008] 상기 생성부의 인코더는 3D 컨볼루션 레이어를 사용하여 상기 실제 이미지에서 특징을 추출하여 참조 코드를 생성하고, 상기 참조 코드는 상기 노이즈 코드와의 합산 과정 없이 생성기에 입력되도록 하여 디코딩된 이미지를 생성하며, 상기 디코딩된 이미지와 상기 실제 이미지의 비교 과정을 통해 상기 생성된 참조 코드를 정확도를 확인한다.

[0009] 상기 판별부는 상기 볼륨 판별기, 슬랩 판별기 및 슬라이스 판별기를 포함하고, 상기 볼륨 판별기는 상기 3D 가상 이미지와 상기 실제 이미지의 전체 슬라이스를 입력으로 사용하며 3D 컨볼루션 레이어를 통해 실제 이미지 여부를 나타내는 확률을 추출하고, 상기 슬랩 판별기는 상기 3D 가상 이미지 및 상기 실제 이미지에서 추출한 특정 z축 범위의 연속된 슬라이스 집합인 슬랩을 사용하고, 상기 슬라이스 판별기는 상기 3D 가상 이미지와 상기 실제 이미지 내 랜덤 z-위치를 가지는 특정 슬라이스들과 상기 특정 슬라이스들과 인접한 슬라이스들을 사용하며, 상기 슬랩 판별기와 상기 슬라이스 판별기는 2D 컨볼루션 레이어를 사용하여 각각 상기 3D 가상 이미지와 상기 실제 이미지에서 추출된 슬랩과 슬라이스들을 판별한다.

[0010] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 방법은 인코더를 통해 실제 이미지를 입력으로 받아 특징을 추출하여 참조 코드를 생성하고, 상기 참조 코드를 생성기에 제공하는 단계, 생성기를 통해 상기 참조 코드 및 노이즈 코드를 입력으로 받아 가상 이미지를 생성하기 위한 학습을 수행하는 단계, 볼륨 판별기(volume discriminator), 슬랩 판별기(slab discriminator) 및 슬라이스 판별기(slice discriminator)를 포함하는 판별부를 통해 상기 생성된 가상 이미지와 실제 이미지를 구별하기 위한 학습을 수행하는 단계 및 상기 학습이 완료되면 상기 생성기를 통해 상기 노이즈 코드만을 입력으로 받아 이미지를 합성하는 단계를 포함한다.

발명의 효과

[0011] 본 발명의 실시예들에 따르면 생성적 적대 네트워크의 학습 방법 및 시스템을 통해 실제 이미지에서 추출한 특징인 참조 코드를 인코딩, 디코딩 과정이 포함된 훈련 과정에 적용하여 생성 프로세스의 수렴도을 높일 수 있다. 또한, 판별기 구조를 볼륨, 슬랩, 슬라이스 판별기로 나누어 세밀한 판별을 수행하며 생성되는 이미지의 수렴도를 높일 수 있다.

도면의 간단한 설명

[0012] 도 1은 본 발명의 일 실시예에 따른 의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 시스템의 구성을 설명하기 위한 도면이다.

도 2는 본 발명의 일 실시예에 따른 생성기, 인코더 및 복수의 서로 다른 판별기의 구조를 설명하기 위한 도면이다.

도 3은 본 발명의 일 실시예에 따른 의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 방법을 설명하기 위한 흐름도이다.

도 4는 본 발명의 일 실시예에 따른 3D 이미지 생성 기법으로 생성된 이미지를 종래기술과 비교하기 위한 도면이다.

도 5는 본 발명의 일 실시예에 따른 3D 이미지 생성 기법으로 생성된 이미지의 픽셀 값 분포를 종래기술과 비교

하기 위한 도면이다.

도 6은 본 발명의 일 실시예에 따른 픽셀 값 분포를 이용한 다차원 척도 분석 결과를 종래기술과 비교하기 위한 도면이다.

도 7은 본 발명의 일 실시예에 따른 슬랩 및 슬라이스 매개변수를 변경하여 얻은 실험 결과의 시각적 비교를 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0013] 본 발명의 실시예에 따르면, 3D 이미지 생성 프로세스의 수렴도(convergence rate)를 높이는 GAN(Generative Adversarial Network) 구조를 제안한다. 제안하는 생성적 적대 네트워크의 학습 방법 및 시스템은 인코딩(encoding) 과정을 통해 실제 영상에서 특징(feature)을 추출하고 디코딩(decoding) 과정을 통해 특징 추출 과정의 정확도를 검증한다. 인코딩 과정에서 추출된 특징을 참조 코드(reference code)라고 명명하며, 이는 생성기(generator)에서 3D 이미지를 생성하는 데 사용된다. 또한 제안하는 기법은 판별기(discriminator) 구조를 3개의 판별기로 나누어 세밀한 판별을 수행하며 생성되는 이미지의 수렴도을 높인다. 판별기는 볼륨 판별기(volume discriminator), 슬랩 판별기(slab discriminator), 슬라이스 판별기(slice discriminator)로 나누어진다. 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.

[0015] 도 1은 본 발명의 일 실시예에 따른 의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 시스템의 구성을 설명하기 위한 도면이다.

[0016] 본 발명의 실시예에 따른 의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 시스템은 생성부(110) 및 판별부(120)를 포함한다.

[0017] 본 발명의 실시예에 따른 생성부(110)는 학습 과정(130)에서 실제 이미지를 입력으로 받아 특징을 추출하여 참조 코드를 생성하고, 상기 참조 코드를 생성기에 제공하는 인코더(111) 및 상기 참조 코드 및 노이즈 코드를 입력으로 받아 가상 이미지를 생성하기 위한 학습을 수행하는 생성기(112)를 포함한다.

[0018] 본 발명의 실시예에 따른 생성부(110)의 인코더(111)는 3D 컨볼루션 레이어를 사용하여 상기 실제 이미지에서 특징을 추출하여 참조 코드를 생성하고, 상기 참조 코드는 상기 노이즈 코드와의 합산 과정 없이 생성기에 입력되도록 하여 디코딩된 이미지를 생성하며, 상기 디코딩된 이미지와 상기 실제 이미지의 비교 과정을 통해 상기 생성된 참조 코드를 정확도를 확인한다.

[0019] 본 발명의 실시예에 따른 생성부(110)의 생성기(112)는 상기 참조 코드를 입력으로 받아 3D 전치 컨볼루션 레이어(transposed convolution layer)를 사용하여 3D 가상 이미지로 업샘플링하고, 학습 과정(130)에서 상기 참조 코드와 가우시안 기반의 노이즈 코드의 평균을 입력으로 사용하여 3D 가상 이미지를 생성하기 위한 학습을 수행한다.

[0020] 본 발명의 실시예에 따른 상기 판별부(120)는 상기 생성된 가상 이미지와 실제 이미지를 구별하기 위한 학습을 수행하는 볼륨 판별기(volume discriminator)(121), 슬랩 판별기(slab discriminator)(122) 및 슬라이스 판별기(slice discriminator)(123)를 포함한다.

[0021] 본 발명의 실시예에 따른 볼륨 판별기(121)는 상기 3D 가상 이미지와 상기 실제 이미지의 전체 슬라이스를 입력으로 사용하며 3D 컨볼루션 레이어를 통해 실제 이미지 여부를 나타내는 확률을 추출한다.

[0022] 본 발명의 실시예에 따른 슬랩 판별기(122)는 상기 3D 가상 이미지 및 상기 실제 이미지에서 추출한 특정 z축 범위의 연속된 슬라이스 집합인 슬랩을 사용한다.

[0023] 본 발명의 실시예에 따른 슬라이스 판별기(123)는 상기 3D 가상 이미지와 상기 실제 이미지 내 랜덤 z-위치를 가지는 특정 슬라이스들과 상기 특정 슬라이스들과 인접한 슬라이스들을 사용한다.

[0024] 본 발명의 실시예에 따른 슬랩 판별기(122) 및 슬라이스 판별기(123)는 2D 컨볼루션 레이어를 사용하여 각각 상기 3D 가상 이미지와 상기 실제 이미지에서 추출된 슬랩과 슬라이스들을 판별한다.

[0025] 본 발명의 실시예에 따른 학습이 완료되면, 이미지 합성 과정(140)에서 상기 생성부(110)의 생성기(113)는 상기 노이즈 코드만을 입력으로 받아 이미지를 합성한다.

[0027] 도 2는 본 발명의 일 실시예에 따른 생성기, 인코더 및 복수의 서로 다른 판별기의 구조를 설명하기 위한 도면이다.

[0028] 본 발명의 실시예에 따른 생성기(G)(210)와 판별기(D)(D_{vol}, D_{slab}, D_{slice})(220) 및 인코더(E)(230)의 내부 구조는 도 2(a)와 같다.

[0029] 도 2(b)는 도 2(a)에 도시된 생성기(G)(210)와 판별기(D)(D_{vol}, D_{slab}, D_{slice})(220) 및 인코더(E)(230) 각각의 내부 구조에 따른 레이어 특성을 나타내는 도면이다.

[0030] 본 발명의 실시예에 따르면, 3D 컨볼루션(convolution)은 인코더(E)와 볼륨 판별기(D_{vol})에서 사용되며 2D 컨볼루션은 슬랩 판별기(D_{slab}) 및 슬라이스 판별기(D_{slice})에서 사용된다. 도 3을 참조하여 본 발명의 실시예에 따른 의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 과정을 더욱 상세히 설명한다.

[0032] 도 3은 본 발명의 일 실시예에 따른 의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 방법을 설명하기 위한 흐름도이다.

[0033] 본 발명의 실시예에 따른 의료 이미지 생성을 위한 높은 수렴도를 지니는 생성적 적대 네트워크의 학습 방법은 인코더를 통해 실제 이미지를 입력으로 받아 특징을 추출하여 참조 코드를 생성하고, 상기 참조 코드를 생성기에 제공하는 단계(310), 생성기를 통해 상기 참조 코드 및 노이즈 코드를 입력으로 받아 가상 이미지를 생성하기 위한 학습을 수행하는 단계(320), 볼륨 판별기(volume discriminator), 슬랩 판별기(slab discriminator) 및 슬라이스 판별기(slice discriminator)를 포함하는 판별부를 통해 상기 생성된 가상 이미지와 실제 이미지를 구별하기 위한 학습을 수행하는 단계(330) 및 상기 학습이 완료되면 상기 생성기를 통해 상기 노이즈 코드만을 입력으로 받아 이미지를 합성하는 단계(340)를 포함한다.

[0034] 단계(310)에서, 인코더를 통해 실제 이미지를 입력으로 받아 특징을 추출하여 참조 코드를 생성하고, 상기 참조 코드를 생성기에 제공한다.

[0035] 본 발명의 실시예에 따르면, 3D 컨볼루션 레이어를 사용하여 상기 실제 이미지에서 특징을 추출하여 참조 코드를 생성하고, 상기 참조 코드는 상기 노이즈 코드와의 합산 과정 없이 생성기에 입력되도록 하여 디코딩된 이미지를 생성하며, 상기 디코딩된 이미지와 상기 실제 이미지의 비교 과정을 통해 상기 생성된 참조 코드를 정확도를 확인한다.

[0036] 단계(320)에서, 상기 생성기를 통해 상기 참조 코드 및 노이즈 코드를 입력으로 받아 가상 이미지를 생성하기 위한 학습을 수행한다.

[0037] 본 발명의 실시예에 따르면, 상기 참조 코드를 입력으로 받아 3D 전치 컨볼루션 레이어(transposed convolution layer)를 사용하여 3D 가상 이미지로 업샘플링하고, 학습 과정에서 상기 참조 코드와 가우시안 기반의 노이즈 코드의 평균을 입력으로 사용하여 3D 가상 이미지를 생성하기 위한 학습을 수행한다.

[0038] 단계(330)에서, 상기 볼륨 판별기, 슬랩 판별기 및 슬라이스 판별기를 포함하는 판별부를 통해 상기 생성된 가상 이미지와 실제 이미지를 구별하기 위한 학습을 수행한다.

[0039] 본 발명의 실시예에 따르면, 상기 볼륨 판별기는 상기 3D 가상 이미지와 상기 실제 이미지의 전체 슬라이스를 입력으로 사용하며 3D 컨볼루션 레이어를 통해 실제 이미지 여부를 나타내는 확률을 추출한다.

[0040] 본 발명의 실시예에 따르면, 상기 슬랩 판별기는 상기 3D 가상 이미지 및 상기 실제 이미지에서 추출한 특정 z 축 범위의 연속된 슬라이스 집합인 슬랩을 사용한다.

[0041] 본 발명의 실시예에 따르면, 상기 슬라이스 판별기는 상기 3D 가상 이미지와 상기 실제 이미지 내 랜덤 z-위치를 가지는 특정 슬라이스들과 상기 특정 슬라이스들과 인접한 슬라이스들을 사용한다.

[0042] 본 발명의 실시예에 따른 상기 슬랩 판별기와 상기 슬라이스 판별기는 2D 컨볼루션 레이어를 사용하여 각각 상기 3D 가상 이미지와 상기 실제 이미지에서 추출된 슬랩과 슬라이스들을 판별한다.

[0043] 본 발명의 실시예에 따른 생성기 G는 참조 코드(reference code) Z_{code}를 입력으로 받아 3D 전치 컨볼루션 레이

어(transposed convolution layer)를 사용하여 3D 이미지 V_{fake}로 업샘플링(up-sampling)한다. 본 발명의 실시예에 따른 생성기 G는 훈련 과정에서 참조 코드(reference code) Z_{code}와 가우시안(Gaussian) 기반의 노이즈 코드(noise code) Z_{noise}의 평균을 입력으로 사용한다. Z_{code}는 실제 이미지 V_{real}을 인코더 E를 사용하여 인코딩함으로써 생성되며 훈련 과정에 실제 이미지의 특징을 제공한다. 본 발명의 실시예에 따른 생성기 G의 학습 과정은 도 1(130)에 도시되었고, 식(1)과 같이 나타낼 수 있다:

[0044]
$$V_{\text{fake}} = G(Z), Z = \frac{Z_{\text{noise}} + Z_{\text{code}}}{2}$$
$$Z_{\text{code}} = E(V_{\text{real}}), Z_{\text{noise}} \sim N(0,1) \in \mathbb{R}^3 \quad (1)$$

[0045] 본 발명의 실시예에 따른 이미지 합성 과정에서는 가우시안 기반 노이즈 코드 Z_{noise}만이 생성기 G의 입력으로 사용된다. 이미지 합성 과정은 도 1(140)에 도시되었고, 식(2)와 같이 나타낼 수 있다:

[0046]
$$V_{\text{gen}} = G(Z_{\text{noise}}), Z_{\text{noise}} \sim N(0,1) \in \mathbb{R}^3 \quad (2)$$

[0047] 본 발명의 실시예에 따르면, 3D 이미지가 실제인지 아닌지 여부를 구분하기 위해 볼륨 판별기(D_{vol}), 슬랩 판별기(D_{slab}), 그리고 슬라이스 판별기(D_{slice})의 세 가지 판별기를 사용한다.

[0048] 본 발명의 실시예에 따른 D_{vol}은 V_{fake}와 V_{real}의 전체 슬라이스를 입력으로 사용하며 3D 컨볼루션 레이어를 통해 진위 여부를 나타내는 확률을 추출한다.

[0049] 본 발명의 실시예에 따른 D_{slab}은 V_{fake} 및 V_{real}에서 추출한 특정 z축 범위의 연속된 슬라이스 집합인 슬랩(slab)을 사용한다.

[0050] 본 발명의 실시예에 따른 D_{slice}는 V_{fake}와 V_{real} 내 랜덤 z-위치를 가진 특정 슬라이스들과 이들과 인접한 슬라이스들을 사용한다.

[0051] 본 발명의 실시예에 따른 D_{slab}과 D_{slice}는 2D 컨볼루션 레이어를 사용하여 각각 V_{fake}와 V_{real}에서 추출된 슬랩과 슬라이스들을 판별한다. 제안하는 판별기 구조는 3차원 영상에서 특정 슬랩와 슬라이스 뿐만 아니라 3D 이미지 내의 전체 슬라이스까지 고려하여 진위 여부를 정확하게 판단할 수 있다. 또한 복수의 서로 다른 판별기를 사용함으로써 3D 이미지 생성 과정에서 수렴도를 향상시킬 수 있다.

[0052] 본 발명의 실시예에 따른 인코더 E는 3D 컨볼루션 레이어를 사용하여 실제 이미지에서 특징을 추출하여 참조 코드 Z_{code}를 생성한다. Z_{code}는 노이즈 코드 Z_{noise}와 함께 생성기 훈련 과정의 입력으로 사용된다. 또한 Z_{code}는 Z_{noise}와의 합산 과정 없이 생성기의 입력으로 사용하여 디코딩된 이미지 V_{dec}를 생성한다. V_{dec}과 실제 이미지 V_{real}의 비교 과정을 통해 Z_{code}가 제대로 생성되었는지 여부를 확인한다. 이 프로세스를 통해 인코더는 실제 이미지의 특징을 정확하게 추출할 수 있다. 생성기와 인코더의 손실 함수(loss function) L_G와 L_E는 식(3)과 같이 나타낼 수 있다:

[0053]
$$L_G = -\mathbb{E}_{\tilde{v} \sim V_{\text{fake}}} [D_{\text{vol}}(\tilde{v})] + \alpha L_{\text{recon}}$$
$$L_E = \mathbb{E}_{\bar{z} \sim Z_{\text{code}}, v \sim V_{\text{real}}} \|G(\bar{z}) - v\|_1^1 \quad (3)$$

[0054] 본 발명의 실시예에 따른 복수의 서로 다른 판별기를 포함하는 판별부 L_D의 손실 함수는 식(4)와 같이 표현된다. L_D^{vol}, L_D^{slab}, L_D^{slice}는 각각 D_{vol}, D_{slab}, D_{slice}의 손실을 나타낸다. 최종 판별부 손실은 각 판별기 손실의 평균으로 계산된다. λ는 그래디언트 패널티(gradient panelty)의 가중치를 나타낸다.

[0055]
$$L_D = \frac{L_D^{\text{vol}} + L_D^{\text{slab}} + L_D^{\text{slice}}}{3} +$$
$$\lambda \mathbb{E}_{\tilde{v} \sim V_{\text{fake}}} (\|\nabla_{\tilde{v}} D_{\text{vol}}(\tilde{v})\|_2 - 1)^2 + \alpha L_{\text{recon}} \quad (4)$$

[0056] 본 발명의 실시예에 따른 각 판별기의 손실함수 L_D^{vol}, L_D^{slab}, L_D^{slice}는 식(5), 식(6), 식(7)로 정의된다. 셀렉터(selector) S_{slab}^{i} 및 S_{slice}^{i}는 각각 i번째 슬랩과 슬라이스들(랜덤 z-위치 슬라이스들과 이와 인접한 슬라이스들)을 추출한다. N_{slab}과 N_{slices}는 슬랩과 슬라이스들의 총 개수를 나타낸다.

[0057] $$L_D^{\mathrm{vol}} = -\mathbb{E}_{v\sim V_{\mathrm{real}}}\left[D_{\mathrm{vol}}(v)\right] + \mathbb{E}_{\tilde{v}\sim V_{\mathrm{fake}}}\left[D_{\mathrm{vol}}(\tilde{v})\right] \quad (5)$$

[0058] $$L_D^{\mathrm{slab}} = \frac{1}{N_{\mathrm{slab}}}\sum_{i=1}^{N_{\mathrm{slab}}}\{-\mathbb{E}_{v\sim V_{\mathrm{real}}}\left[D_{\mathrm{slab}}(S_{\mathrm{slab}}^{i}(v))\right] + \mathbb{E}_{\tilde{v}\sim V_{\mathrm{fake}}}\left[D_{\mathrm{slab}}(S_{\mathrm{slab}}^{i}(\tilde{v}))\right]\} \quad (6)$$

[0059] $$L_D^{\mathrm{slice}} = \frac{1}{N_{\mathrm{slice}}}\sum_{i=1}^{N_{\mathrm{slice}}}\{-\mathbb{E}_{v\sim V_{\mathrm{real}}}\left[D_{\mathrm{slice}}(S_{\mathrm{slice}}^{i}(v))\right] + \mathbb{E}_{\tilde{v}\sim V_{\mathrm{fake}}}\left[D_{\mathrm{slice}}(S_{\mathrm{slice}}^{i}(\tilde{v}))\right]\} \quad (7)$$

[0060] 재구성 손실(reconstruction loss) L_{recon}은 생성된 이미지와 실제 이미지 간의 차이를 측정하며 식(8)과 같이 나타낼 수 있다:

[0061] $$L_{\mathrm{recon}} = L_{\mathrm{dist}} + L_{\mathrm{sim}} + L_{\mathrm{cont}} \quad (8)$$

[0062] 거리 손실(distance loss) L_{dist}는 식(9)와 같이 정의된다. L_{dist}는 실제 이미지 V_{real}과 생성된 이미지 V_{fake} 간의 L1 error를 계산한다.

[0063] $$L_{\mathrm{dist}} = \mathbb{E}_{v\sim V_{\mathrm{real}},\tilde{v}\sim V_{\mathrm{fake}}}\left\|v-\tilde{v}\right\|_1^1 \quad (9)$$

[0064] 유사성 손실(similarity loss)은 식(10)으로 정의되며 SSIM(Structural Similarity Index Map)을 사용하여 V_{real}과 V_{fake}의 내부 슬라이스 간 품질 차이를 측정한다. 모든 슬라이스에서 측정된 SSIM 값의 평균을 유사성 손실에 사용한다. $v[k]$는 k번째 이미지 슬라이스를 나타내고, r_z는 3D 이미지의 z축 방향 해상도이다. 유사성 손실을 사용하여 실제 이미지의 내부 특성을 훈련 과정에 반영할 수 있다.

[0065] $$L_{\mathrm{sim}} = 1 - \mathbb{E}_{v\sim V_{\mathrm{real}},\tilde{v}\sim V_{\mathrm{fake}}}\left[\frac{1}{r_z}\sum_{k=0}^{r_z} SSIM(v[k],\tilde{v}[k])\right] \quad (10)$$

[0066] 연속 손실(continuous loss) L_{cont}는 식(11)과 같이 정의되며, 인접한 슬라이스 사이에 변화가 적은 3D 이미지의 연속적인 특성을 반영한다. L_{cont}는 V_{real}과 V_{fake}에서 각각 (k-1)번째 슬라이스와 k번째 슬라이스를 추출한다. 추출된 두 인접한 슬라이스 사이의 L1 오류를 계산한다. V_{real}과 V_{fake}에서 각각 계산된 L1 오류 값의 차이를 L_{cont}로 사용한다.

[0067] $$L_{\mathrm{cont}} = \mathbb{E}_{v\sim V_{\mathrm{real}},\tilde{v}\sim V_{\mathrm{fake}}}\frac{1}{r_z}\sum_{k=1}^{r_z}\left|\left\|v[k-1]-v[k]\right\|_1^1 - \left\|\tilde{v}[k-1]-\tilde{v}[k]\right\|_1^1\right| \quad (11)$$

[0069] 본 발명의 실시예에 따른 실험을 위해 140명의 간 CT 이미지 19,312개(1인당 평균 137개의 이미지)와 128명의 척추 CT 이미지 4,116개(1인당 평균 32개의 이미지)로 구성된 척추 데이터로 구성된 CT-ORG 데이터 셋을 사용하여 훈련되었다. 간이나 척추가 포함되지 않은 부분을 잘라낸 다음 크기를 조정하고 쌓아서 1283 해상도를 가진 3D 이미지를 만들었다. 생성된 3D 이미지의 복셀 값은 [0, 1]로 정규화하였다. 훈련 과정에서는 RMSprop 최적화

기법이 적용되었으며 48GB 메모리를 가진 NVIDIA Quadro RTX 8000 GPU가 사용되었다. 생성기, 판별기, 인코더의 학습률(learning rate)는 각각 2×10^{-5}, 2×10^{-5}, 2×10^{-6}으로 설정했다. 제안하는 기법은 19,213MB GPU 메모리를 사용하여 400 에포크 동안 약 20시간 동안 훈련되었다. 판별 과정에서 간의 경우 28개의 슬랩과 36개의 슬라이스들을 추출하였고, 척추의 경우 28개의 슬랩과 56개의 슬라이스들을 추출하였다.

[0071] 도 4는 본 발명의 일 실시예에 따른 3D 이미지 생성 기법으로 생성된 이미지를 종래기술과 비교하기 위한 도면이다.

[0072] 도 4(a)는 제안하는 기법과 종래기술에 따른 3D 이미지 생성 기법으로 생성된 간 이미지를 나타내고, 도 4(b)는 제안하는 기법과 종래기술에 따른 3D 이미지 생성 기법으로 생성된 척추 이미지의 시각적 비교를 위한 도면이다.

[0073] 첫 번째 행은 볼륨 렌더링으로 생성된 3D 뷰를 나타낸다. 두 번째 내지 네 번째 행은 각각 3D 이미지의 중심에 위치한 coronal, sagittal, axial 슬라이스를 나타낸다. 시각적 비교에 활용된 이미지 생성 기법은 3D-GAN(첫 번째 열), HA-GAN(두 번째 열), 3D-α-WGAN(세 번째 열), 제안하는 기법(네 번째 열)이다. 이미지 생성 과정에서 에포크는 모든 기법에 대해 동일하게 400으로 설정하였다.

[0074] 3D-GAN은 경계에 많은 아티팩트(artifact)를 보여주며 척추 내부 조직 사이에 존재하는 빈 공간을 정확하게 표현하지 못한다. HA-GAN과 3D-α-WGAN은 3D-GAN보다 대상 객체(간, 척추)에 더 가까운 결과를 생성했다. 그러나 생성된 간 이미지가 실제 이미지의 형태와 많이 다르고, 척추 좌우의 트래버스 프로세스(traverse process)가 명확하게 표현되지 않는다(즉, 축(axial) 슬라이스에서 이러한 문제점이 명확히 나타난다). 또한 척추 조직 사이에 존재하는 일부 빈 공간을 정확하게 나타내지 못했다(이는 3D 뷰에서 확인할 수 있다). 제안하는 방법은 실제 이미지의 간의 형태와 매우 유사한 이미지를 생성하며, 척추의 트래버스 프로세스(traverse process)와 조직 사이에 존재하는 빈 공간을 정확하게 표현한다. 이를 통해 제안하는 방법이 기존 방법보다 대상 객체에 더 가까운 3D 이미지를 생성할 수 있음을 확인하였다.

[0076] 표 1 및 표 2는 3D-GAN, HA-GAN, 3D-α-WGAN 및 제안하는 기법에 의해 생성된 이미지에 평가 인자를 사용하여 비교한 결과를 보여준다.

[0077] <표 1> 제안하는 기법과 기존 기법으로 생성된 간 이미지의 정량적 비교

[0078]

	MMD ↓	FID ↓	LPIPS ↓	PSNR ↑
3D-GAN	0.0623	1.3725	0.4406	12.8007
HA-GAN	0.0875	2.4813	0.4273	13.1136
3D-α-WGAN	0.0393	0.6562	0.4315	10.0632
Proposed	**0.0303**	**0.4721**	**0.3635**	**16.1946**

[0079] <표 2> 제안하는 기법과 기존 기법으로 생성된 척추 이미지의 정량적 비교

[0080]

	MMD ↓	FID ↓	LPIPS ↓	PSNR ↑
3D-GAN	0.0685	1.4361	0.3832	13.0321
HA-GAN	0.0394	0.6445	0.3573	6.1534
3D-α-WGAN	0.0477	0.8792	**0.3455**	8.0951
Proposed	**0.0383**	**0.5978**	0.3502	**13.1341**

[0081] FID(Frechet inception distance), MMD(maximum mean discrepancy), LPIPS(Learned Perceptual Image Patch Similarity)는 생성된 이미지가 실제 이미지와 얼마나 유사한지를 평가하는 지표이다. FID는 생성된 이미지와 실제 이미지의 분포 간 차이(disparity)를 측정한다. MMD는 생성된 이미지와 실제 이미지가 나타내는 분포의 포인트별 거리(pointwise distance)를 계산한다. LPIPS는 생성된 이미지와 실제 이미지 간의 지각 유사도(perceptual similarity)를 산출한다. 낮은 FID, MMD 및 LPIPS 값은 생성된 이미지가 대상 객체와 더욱 유사함을 나타낸다. PSNR(Peak Signal-to-Noise Ratio)는 이미지 생성 과정에서 품질 손실을 측정하며 평균 제곱 오차

(MSE)를 기반으로 계산된다. PSNR 값이 높을수록 이미지 생성 중 실제 이미지의 특성 손실이 감소하여 품질이 높아진다.

[0082] 제안하는 기법은 3D-GAN 및 HA-GAN과 비교하여 모든 평가 지표에서 더 높은 성능을 보여주었다. LPIPS의 경우 제안하는 기법과 비교하여 3D-α-WGAN이 더 낮았지만 다른 지표에서 성능을 능가하였다. 따라서 제안하는 기법은 3D-α-WGAN보다 더 높은 유사성과 품질을 가진 이미지를 생성하는 것을 보여준다. 결론적으로 제안하는 기법은 기존 기법과 비교하여 대상 객체와 더 유사하며 높은 품질의 이미지를 생성하는 것을 확인할 수 있다.

[0084] 도 5는 본 발명의 일 실시예에 따른 3D 이미지 생성 기법으로 생성된 이미지의 픽셀 값 분포를 종래기술과 비교하기 위한 도면이다.

[0085] 도 5(a)는 제안된 방법과 종래기술에 따른 3D 이미지 생성 방법에 따라 생성된 간 이미지의 픽셀 값 분포를 박스 플롯(box plot)으로 나타낸 도면이고, 도 5(b)는 제안된 방법과 종래기술에 따른 3D 이미지 생성 방법에 따라 생성된 척추 이미지의 픽셀 값 분포를 박스 플롯으로 나타낸 도면이다. 각 박스 플롯에서 중앙값은 선으로 표시되고 상자의 가장자리는 픽셀 값 분포의 첫 번째 사분위수(25%)와 세 번째 사분위수(75%)를 나타낸다.

[0086] 도 5에서 제안하는 기법이 간과 척추 이미지에서 실제 데이터에 가장 가까운 중앙값을 가지고 있었고 IQR(Inter-Quartile Range)도 가장 많이 겹친다는 것을 확인할 수 있다. 도 5의 결과를 바탕으로 제안하는 기법은 기존 기법에 비해 실제 데이터의 픽셀 값 분포와 더욱 유사한 3D 이미지를 생성하는 것을 확인할 수 있다.

[0088] 도 6은 본 발명의 일 실시예에 따른 픽셀 값 분포를 이용한 다차원 척도 분석 결과를 종래기술과 비교하기 위한 도면이다.

[0089] 표 3 및 표 4는 각각 간 및 척추에 대해 인코더 및 판별기 구조의 절제(ablation) 효과를 나타낸다.

[0090] <표 3> 생성된 간 이미지에 대한 절제 연구

[0091]

	MMD ↓	FID ↓	LPIPS ↓	PSNR ↑
w/o encoding	0.0624	1.4560	0.5080	11.4651
w/o slice	0.0326	0.5427	0.4031	15.2971
w/o slab	0.0372	0.6307	0.4030	15.2749
w/o slab, slice	0.0393	0.6831	0.3992	15.3060
Proposed	**0.0303**	**0.4721**	**0.3635**	**16.1946**

[0092] <표 4> 생성된 척추 이미지에 대한 절제 연구

[0093]

	MMD ↓	FID ↓	LPIPS ↓	PSNR ↑
w/o encoding	0.0394	0.7473	0.3688	12.0311
w/o slice	0.0473	0.8355	0.3722	13.3977
w/o slab	0.0475	0.8411	0.3749	**13.5009**
w/o slab, slice	0.0419	0.6783	0.3513	13.1443
Proposed	**0.0383**	**0.5978**	**0.3502**	13.1341

[0094] 첫 번째 행은 인코딩 과정 없이 노이즈만으로 생성된 이미지의 결과를 보여준다. 두 번째 행은 슬라이스 판별기가 이미지 생성에 사용되지 않을 때 얻은 결과를 나타낸다. 세 번째 행은 슬랩 판별기가가 이미지 생성에 사용되지 않을 때 얻은 결과이다. 네 번째 행은 슬라이스 및 슬랩 판별기 둘 다 사용하지 않고 얻은 결과를 나타낸다. 마지막으로 다섯 번째 행은 본 발명에서 제안한 인코딩 과정과 판별기 구조를 모두 사용하여 얻은 결과를 보여준다. 척추의 경우 슬랩 판별기 또는 슬라이스 판별기를 따로 사용하는 것이 둘 다 사용하지 않는 것보다 성능이 낮다. 그러나 간의 경우 슬랩 또는 슬라이스 판별기를 개별적으로 사용하면 사용하지 않는 것에 비해 성능이 향상된다. 이러한 결과는 척추와 같은 복잡한 모양의 대상 객체를 처리할 때 생성된 이미지의 품질을 향상시키기 위해 슬랩 판별기와 슬라이스 판별기를 결합하는 것이 필요함을 시사한다.

[0095] 표 3 및 표 4의 결과는 제안된 인코딩 과정과 판별기 구조를 모두 활용하면 대부분의 평가 지표에서 향상된 결과를 얻을 수 있음을 보여준다. 슬라이스 판별기의 성능 향상은 간에서 상대적으로 적었으며, 척추에 슬라이스 판별기를 사용하면 간보다 성능이 더 크게 향상되었다. 이러한 결과는 대상 객체의 복잡성이 증가함에 따라 슬라이스 판별기가 성능이 더 증가하는 것을 나타낸다. 간의 경우 제안된 판별기를 채택하면 PSNR이 더 증가한다. 그러나 척추의 경우 제안된 판별기가 PSNR의 증가에 기여한다고 보기 어렵다.

[0096] 도 6(a)는 간 이미지에 대하여 제안된 방법과 종래기술로 생성된 3D 이미지의 픽셀 값 분포를 이용한 다차원 척도 분석(Multi-Dimensional Scaling; MDS) 결과를 나타내고, 도 6(b)는 척추 이미지에 대하여 제안된 방법과 종래기술로 생성된 3D 이미지의 픽셀 값 분포를 이용한 다차원 척도 분석 결과를 나타낸다.

[0097] MDS 값의 범위는 convex hull로 표시된다. 3D 이미지 간의 MDS 거리는 생성된 데이터의 다양성을 나타낸다. 제안한 방법이 간 이미지에서 두 번째로 큰 MDS 분포 범위를 보였으며, 척추 이미지에서 가장 큰 MDS 분포 범위를 보임을 알 수 있다. 다른 기법들과 달리 3D-GAN은 실제 이미지에서 추출된 특징을 학습 과정에서 활용하지 않는다. 실제 이미지의 특징은 생성된 이미지가 대상 개체의 모양으로 수렴시키는 역할을 한다. 따라서 실제 이미지의 특징을 사용하지 않는 3D-GAN으로 생성된 이미지는 대상 객체의 모양에 덜 수렴되어 역으로 다양성이 증대된다. 결론적으로 도 6의 결과는 제안하는 방법이 종래기술에 따른 방법보다 대부분의 경우 더 다양한 모양의 이미지를 생성함을 보여준다.

[0099] 도 7은 본 발명의 일 실시예에 따른 슬랩 및 슬라이스 매개변수를 변경하여 얻은 실험 결과의 시각적 비교를 위한 도면이다.

[0100] 표 5 내지 표 8은 판별기의 슬랩 및 슬라이스 매개변수를 조정하여 얻은 실험 결과를 나타낸다.

[0101] 도 7(a) 내지 도7(d)는 각각 표 5 내지 표 8에 제시된 결과를 시각적으로 보여준다. 각 평가 요소의 측정 값인 표 5 내지 표 8의 열 값을 기준으로 한 상대적인 비율을 나타낸다.

[0102] <표 5> 간 이미지에서 슬랩 매개변수를 변경한 결과

Slabs	MMD ↓	FID ↓	LPIPS ↓	PSNR ↑
8	0.0365	0.6322	0.4014	15.4904
18	0.0336	0.5571	0.3971	15.1051
28	**0.0303**	**0.4721**	**0.3635**	**16.1946**
38	0.0312	0.5171	0.3794	15.8167
48	0.0361	0.6393	0.4007	14.6764

[0103]

[0104] <표 6> 척추 이미지에서 슬랩 매개변수를 변경한 결과

Slices	MMD ↓	FID ↓	LPIPS ↓	PSNR ↑
16	0.0328	0.5281	0.3811	**16.2126**
26	0.0320	0.5079	0.3884	15.6555
36	**0.0303**	**0.4721**	**0.3635**	16.1946
46	0.0309	0.4885	0.3757	14.9179
56	0.0329	0.5375	0.3798	15.8764

[0105]

[0106] <표 7> 간 이미지에서 슬라이스 매개변수를 변경한 결과

Slabs	MMD ↓	FID ↓	LPIPS ↓	PSNR ↑
8	0.0717	1.5906	0.4300	11.0781
18	0.0523	0.9781	0.3869	12.2150
28	**0.0383**	**0.5978**	**0.3502**	**13.1341**
38	0.0475	0.8515	0.3700	12.2538
48	0.0473	0.8097	0.3870	13.0054

[0107]

[0108] <표 8> 척추 이미지에서 슬라이스 매개변수를 변경한 결과

Slices	MMD ↓	FID ↓	LPIPS ↓	PSNR ↑
36	0.0604	1.2253	0.4139	11.1866
46	0.0537	1.0050	0.3940	12.1558
56	**0.0383**	**0.5978**	**0.3502**	**13.1341**
66	0.0530	0.9781	0.3782	12.4567
76	0.0504	0.9243	0.3844	12.1858

[0109]

[0110] 슬랩 매개변수는 슬랩 판별기가 추출하는 연속된 위치의 슬라이스 수를 결정하는 반면, 슬라이스 매개변수는 슬라이스 판별기가 추출하는 무작위로 선택된 슬라이스의 수를 제어한다.

[0111] 본 발명의 실시예에 따른 실험에서 슬랩 매개변수의 크기는 간과 척추에 대해 모두 8~48개(3D 이미지의 128개 슬라이스 중 약 6%~38%)의 범위로 설정했다. 슬라이스 매개변수는 간의 경우 16~56개(총 슬라이스의 약 13%~44%), 척추의 경우 36~76개(총 슬라이스의 약 28%~59%)로 설정했다. 본 발명의 실시예에 따른 실험에서 얻은 최적 슬랩 및 슬라이스 매개변수 값은 간의 경우 28개, 38개이고 척추의 경우 28개, 56개이다. 간과 비교하여 척추는 상대적으로 복잡한 모양을 가지고 있으므로 상대적으로 더 큰 최적 슬라이스 매개변수 값을 가진다.

[0112] 간의 경우 슬랩 매개변수가 최적값(28개)에서 벗어남에 따라 MMD, FID, LPIPS가 증가했다. 이는 슬랩 매개변수의 변화가 생성된 간 이미지의 유사도에 영향을 미친다는 것을 의미한다. PSNR은 슬랩 매개변수가 최적값에서 멀어질수록 감소하고 최적값에서 최대값에 도달한다. 슬라이스 매개변수가 최적값(36개)에서 벗어남에 따라 MMD와 FID가 모두 증가했다. 슬랩 매개변수에 대한 MMD, FID의 증가는 슬라이스 매개변수보다 더 크다. 이는 슬랩 매개변수의 변화가 슬라이스 매개변수보다 생성된 간 이미지의 유사성에 더 큰 영향을 미쳤음을 나타낸다. LPIPS는 슬라이스 매개변수가 최적값보다 낮을 때 상대적으로 높은 값을 보였고, 최적값에서 최소값에 도달했으며, 매개변수가 최적값보다 커질수록 점차 증가했다. PSNR는 슬라이스 매개변수의 변화에 따른 명확한 증감를 나타내지 않았으며, 이는 슬라이스 매개변수가 생성된 간 이미지의 PSNR에 큰 영향을 미치지 않음을 보여준다.

[0113] 척추의 경우 슬랩 매개변수가 최적 값보다 낮을 때 MMD, FID, LPIPS가 증가하고 PSNR이 감소했다. 이는 슬랩 판별기에서 더 작은 슬랩 매개변수를 사용할 때 이미지 유사성과 품질이 감소함을 나타낸다. 이를 통해 특정 z 범위 내의 연속된 슬라이스를 기반으로 3D 이미지를 평가하는 슬랩 판별기의 성능은 사용된 슬라이스 수에 영향을 받으며, 슬랩 판별기가 효과적으로 수행하기 위해 충분한 수의 슬라이스를 갖는 것이 중요하다는 것을 보여준다. 슬랩 매개변수가 최적값을 초과할 때 MMD, FID, LPIPS는 증가했지만 PSNR은 크게 변하지 않았다. 이러한 결과는 슬랩 크기를 과도하게 늘리면 역으로 성능이 저하될 수도 있음을 보여준다. 슬라이스 매개변수에 대한 결과는 슬랩 매개변수와 유사한 경향을 나타냈다. 그러나 슬랩 매개변수와 비교하여 슬라이스 매개변수가 최적값(56개) 이하일 때 MMD와 FID의 증가폭이 적었고, 최적값 이상일 때 증가폭이 컸다. 이는 슬랩 매개변수에 비해 슬라이스 매개변수의 경우 최적값보다 커질 경우 성능에 더 큰 영향을 미친다는 것을 나타낸다. 슬라이스 매개변수가 최적값에서 벗어남에 따라 PSNR이 감소했다.

[0114] 종래기술에 따른 GAN 기반 방법은 주로 2D 이미지를 생성하는 데 사용되었다. 하지만 제안하는 방법은 3D 이미지를 생성하는 능력을 갖추고 있다. 이는 의료 영상에서 중요한 3D 해부학적 구조와 인접 슬라이스 간의 연관성을 보다 정확하게 반영할 수 있음을 의미한다. 제안하는 방법은 실제 이미지에서 추출한 특징인 참조 코드를 인코딩 및 디코딩 과정을 포함하는 전체 학습 과정에 통합한다. 이를 통해 실제 이미지가 가진 특징을 학습 과정

에 정확히 반영하여 수렴도를 향상시킨다. 제안하는 방법은 판별기 구조를 볼륨, 슬랩, 슬라이스 판별기로 구분하여 세밀한 판별이 가능하다. 이로 인해 3D 이미지 내 미세한 세부 사항을 학습과정에 더 잘 반영할 수 있다. 슬랩 판별기는 연속적인 위치의 슬라이스를 평가하고, 슬라이스 판별기는 비연속적인 위치의 슬라이스에서 작동하므로 인접한 슬라이스 간의 연속성과 특정 슬라이스의 세부 사항을 모두 고려할 수 있다. 결론적으로 제안하는 방법은 학습 과정에 참조 코드 적용 및 세밀한 판별 기능을 통해 수렴도를 향상시키는 장점을 가지고 있다.

[0115] 또한, 제안하는 기법은 기존 3D 생성 기법들과 비교하여 더욱 사실적인 3D 이미지를 생성할 수 있음을 실험 결과에서 확인할 수 있다. 제안하는 방법은 높은 유사도와 품질을 가진 3D 이미지 생성함으로써 의료 영상 분야 연구에 도움을 줄 수 있다. 특히 의료 진단, 수술 계획 및 훈련, 의료 영상 데이터 분석 등 다양한 응용 분야에서 제안하는 기법이 생성한 3D 이미지가 유용하게 활용될 수 있다.

[0117] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0118] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

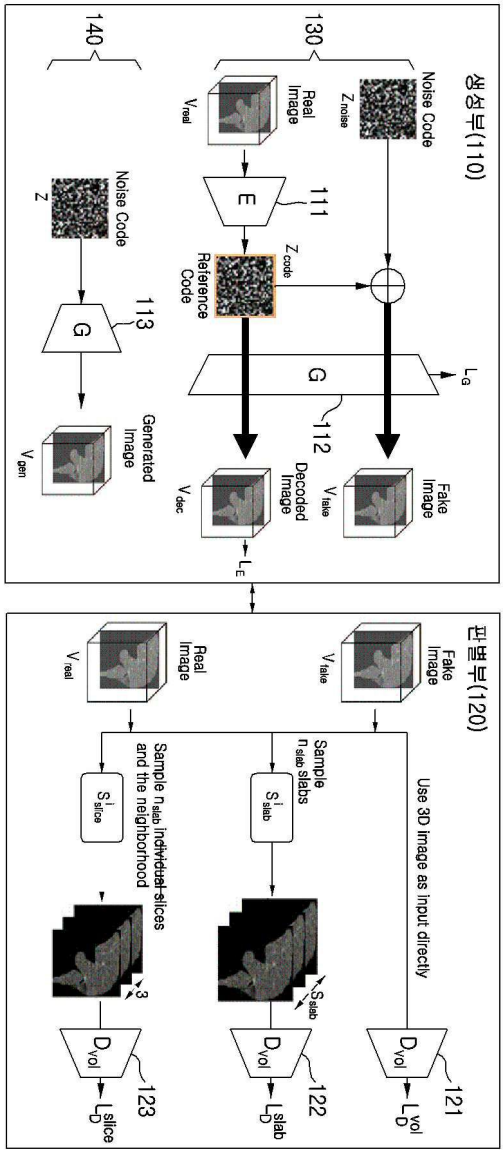
[0119] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic 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.

[0120] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

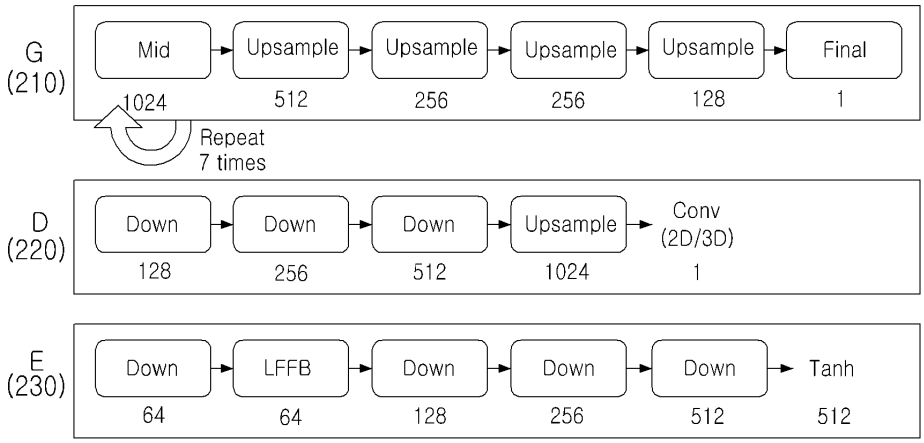
[0121] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

도면1


생성부(110)
130
140
Noise Code
Real Image
Reference Code
Decoded Image
Fake Image
Generated Image
111
112
113
E
G
Noise Code
Generated Image
판별부(120)
Fake Image
Real Image
Use 3D image as input directly
Sample n slab slabs
Sample n slab individual slices and the neighborhood
121
122
123

도면2a

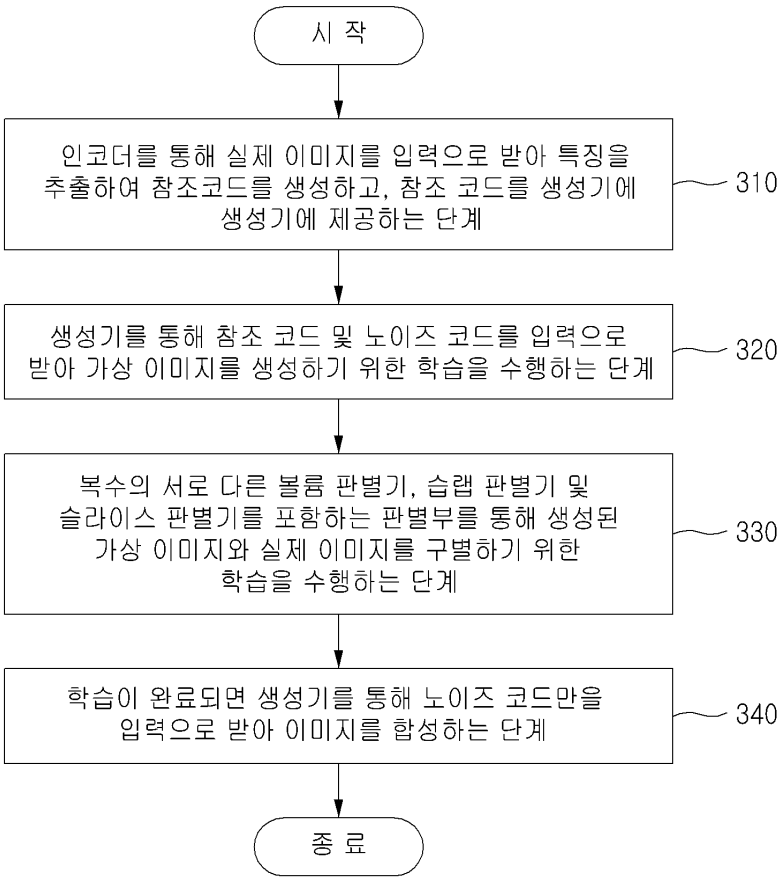


도면2b

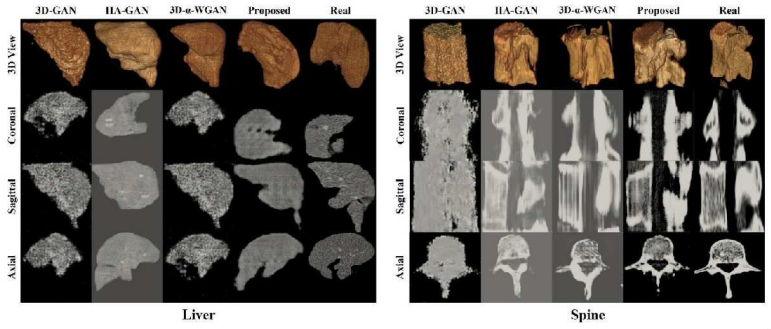
(Layer Specification)

Down	Mid	Upsample	Final
(2D/3D) Conv InslanceNorm3D LeakyReLU Dropout	3D Conv InslanceNorm3D LeakyReLU Dropout (Skip Connection)	3D TransposeConv InslanceNorm3D ReLU Dropout	3D TransposeConv Tanh

도면3

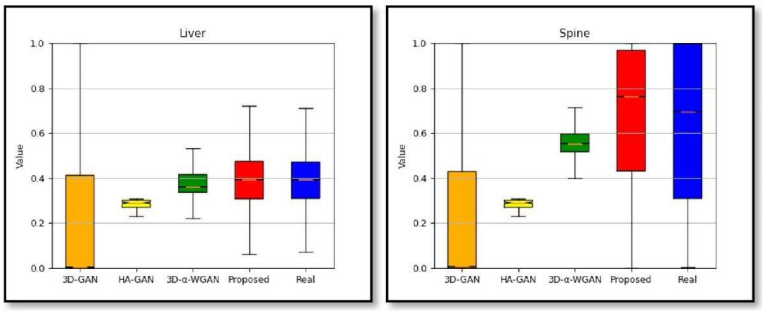

시 작
인코더를 통해 실제 이미지를 입력으로 받아 특징을 추출하여 참조코드를 생성하고, 참조 코드를 생성기에 생성기에 제공하는 단계
310
생성기를 통해 참조 코드 및 노이즈 코드를 입력으로 받아 가상 이미지를 생성하기 위한 학습을 수행하는 단계
320
복수의 서로 다른 볼륨 판별기, 습랩 판별기 및 슬라이스 판별기를 포함하는 판별부를 통해 생성된 가상 이미지와 실제 이미지를 구별하기 위한 학습을 수행하는 단계
330
학습이 완료되면 생성기를 통해 노이즈 코드만을 입력으로 받아 이미지를 합성하는 단계
340
종 료

도면4


3D-GAN
IIA-GAN
3D-α-WGAN
Proposed
Real
3D View
Coronal
Sagittal
Axial
Liver
Spine

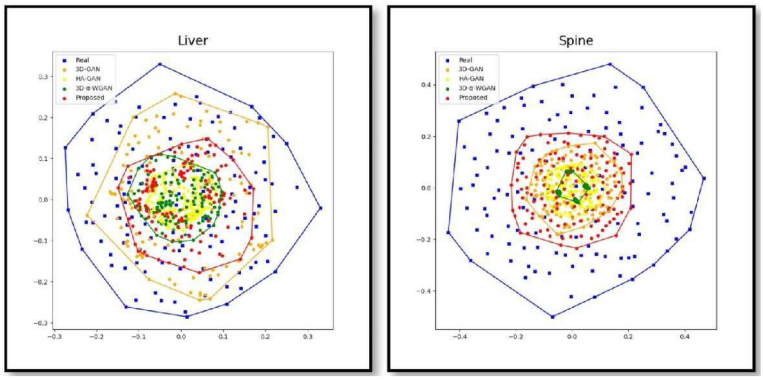
(a) (b)

도면5


Liver
Value
1.0
0.8
0.6
0.4
0.2
0.0
3D-GAN
HA-GAN
3D-α-WGAN
Proposed
Real
Spine
Value
1.0
0.8
0.6
0.4
0.2
0.0
3D-GAN
HA-GAN
3D-α-WGAN
Proposed
Real

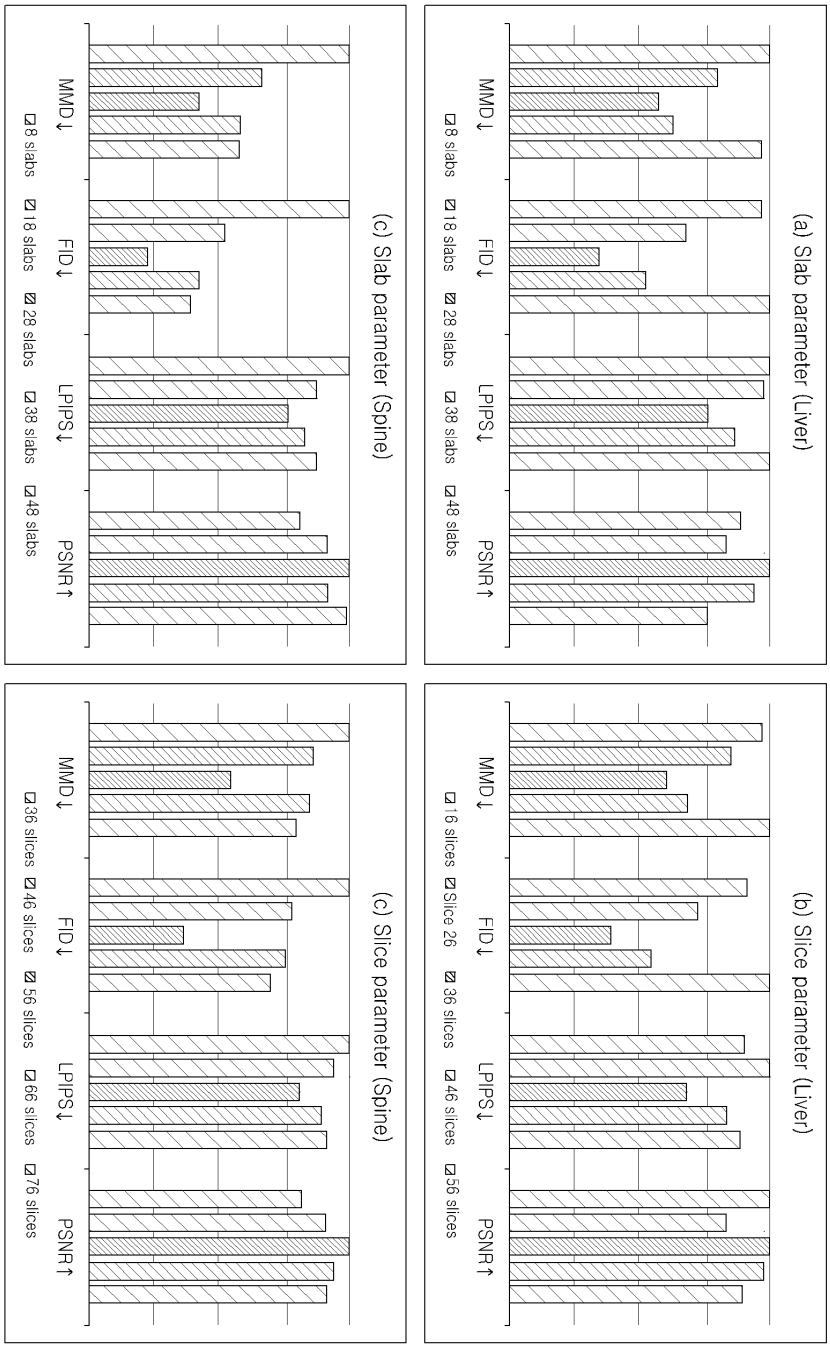
(a) (b)

도면6


Liver
Real
3D-GAN
HA-GAN
3D-α-WGAN
Proposed
Spine
Real
3D-GAN
HA-GAN
3D-α-WGAN
Proposed

(a) (b)

도면7


(a) Slab parameter (Liver)
MMD↓
FID↓
LPIPS↓
PSNR↑
8 slabs
18 slabs
28 slabs
38 slabs
48 slabs
(b) Slice parameter (Liver)
16 slices
Slice 26
36 slices
46 slices
56 slices
(c) Slab parameter (Spine)
(c) Slice parameter (Spine)
36 slices
46 slices
56 slices
66 slices
76 slices