



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년09월20일
(11) 등록번호 10-2708543
(24) 등록일자 2024년09월13일

- | | |
|--|--|
| <p>(51) 국제특허분류(Int. Cl.)
G06T 3/40 (2024.01) G06N 3/0464 (2023.01)
G06T 11/00 (2006.01)</p> <p>(52) CPC특허분류
G06T 3/4053 (2024.01)
G06N 3/0464 (2023.01)</p> <p>(21) 출원번호 10-2023-0095421</p> <p>(22) 출원일자 2023년07월21일
심사청구일자 2023년07월21일</p> <p>(56) 선행기술조사문헌
KR1020220155737 A
SRDiff: Single image super-resolution with diffusion probabilistic models](2022.03.28)
High-Resolution Image Synthesis With Latent Diffusion Models](2022.06.19)</p> | <p>(73) 특허권자
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)</p> <p>(72) 발명자
조대현
서울특별시 동대문구 천호대로25길 81, 814호
류희성
서울특별시 강북구 월계로21가길 41, 106동 1001호
월러번 크리스티안
02465, 대한민국 서울특별시 동대문구 회기로2길 5, 샤인빌304호</p> <p>(74) 대리인
김등용</p> |
|--|--|

전체 청구항 수 : 총 5 항

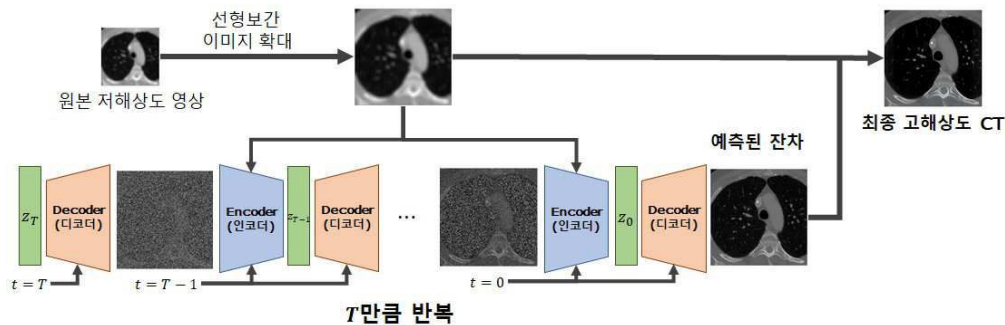
심사관 : 이정은

(54) 발명의 명칭 확산 모델 기반 컴퓨터 단층촬영 고해상도 변환 장치 및 방법

(57) 요약

확산 모델 기반 컴퓨터 단층촬영 영상의 고해상도 변환 장치 및 방법이 개시된다. 상기 고해상도 변환 방법은 적어도 프로세서(processor)를 포함하는 컴퓨팅 장치에 의해 수행되고, 확산 모델(Diffusion Model)을 이용하여 원본 이미지로부터 학습 데이터를 생성하는 단계, 상기 학습 데이터를 이용하여 인코더(endoder)와 디코더(decoder)를 포함하는 합성곱 신경망 모델(Convolutional Neural Network Model, CNN 모델)을 학습하는 단계, 입력 이미지를 선형 보간(linear interpolation)하여 확대된 이미지를 생성하는 단계, 가우시안 분포로부터 추출된 임의의 벡터와 상기 확대된 이미지를 이용하여 고해상도 잔차를 생성하는 단계, 및 상기 확대된 이미지와 상기 고해상도 잔차를 결합하여 고해상도 이미지를 생성하는 단계를 포함한다.

대표도 - 도4



(52) CPC특허분류
G06T 11/003 (2013.01)
G06T 3/4007 (2024.01)
G06T 3/4046 (2024.01)
G06T 2207/10081 (2013.01)
G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193973
과제번호	2019-0-00079-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	인공지능대학원지원(고려대학교)
기 여 율	1/1
과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

적어도 프로세서(processor)를 포함하는 컴퓨팅 장치에 의해 수행되는 고해상도 변환 방법에 있어서,

확산 모델(Diffusion Model)을 이용하여 원본 이미지로부터 학습 데이터를 생성하는 단계;

상기 학습 데이터를 이용하여 인코더(endoder)와 디코더(decoder)를 포함하는 합성곱 신경망 모델(Convolutional Neural Network Model, CNN 모델)을 학습하는 단계;

입력 이미지를 선형 보간(linear interpolation)하여 확대된 이미지를 생성하는 단계;

가우시안 분포로부터 추출된 임의의 벡터와 상기 확대된 이미지를 이용하여 고해상도 잔차를 생성하는 단계; 및

상기 확대된 이미지와 상기 고해상도 잔차를 결합하여 고해상도 이미지를 생성하는 단계를 포함하고,

상기 학습 데이터는 제1 학습 데이터를 포함하고,

상기 학습 데이터를 생성하는 단계는 상기 원본 이미지에 대응하는 잔차 이미지로부터 부가된 노이즈가 상이한 복수의 이미지들을 생성하는 단계를 포함하고,

상기 제1 학습 데이터는 상기 잔차 이미지와 상기 복수의 이미지들을 포함하고,

상기 학습하는 단계는 상기 제1 학습 데이터를 이용하여 상기 합성곱 신경망 모델을 학습하되, 확산 과정에서 노이즈가 덜 부가된 이전 단계의 이미지를 추론하도록 학습하고,

상기 고해상도 잔차를 생성하는 단계는, 학습이 완료된 모델을 이용하여 확산 과정에서 노이즈가 부가되는 단계의 개수에 해당하는 횟수만큼 추론을 수행하되, 상기 확대된 이미지의 임베딩이 인코더에 입력되는,

고해상도 변환 방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

적어도 프로세서(processor)를 포함하는 컴퓨팅 장치에 의해 수행되는 고해상도 변환 방법에 있어서,

확산 모델(Diffusion Model)을 이용하여 원본 이미지로부터 학습 데이터를 생성하는 단계;

상기 학습 데이터를 이용하여 인코더(endoder)와 디코더(decoder)를 포함하는 합성곱 신경망 모델(Convolutional Neural Network Model, CNN 모델)을 학습하는 단계;

입력 이미지를 선형 보간(linear interpolation)하여 확대된 이미지를 생성하는 단계;

가우시안 분포로부터 추출된 임의의 벡터와 상기 확대된 이미지를 이용하여 고해상도 잔차를 생성하는 단계; 및

상기 확대된 이미지와 상기 고해상도 잔차를 결합하여 고해상도 이미지를 생성하는 단계를 포함하고,

상기 학습 데이터는 제2 학습 데이터를 포함하고,

상기 학습 데이터를 생성하는 단계는 상기 원본 이미지에 대응하는 잔차 이미지의 임베딩에 확산 모델을 적용하여 복수의 임베딩들을 생성하는 단계를 포함하고,

상기 제2 학습 데이터는 상기 원본 이미지에 대응하는 잔차 이미지의 임베딩, 및 상기 복수의 임베딩들을 포함

하고,

상기 학습하는 단계는 상기 제2 학습 데이터를 이용하여 상기 합성곱 신경망 모델을 학습하되, 확산 과정에서 노이즈가 덜 부가된 이전 단계의 임베딩을 추론하도록 학습하고,

상기 고해상도 잔차를 생성하는 단계는, 학습이 완료된 모델을 이용하여 확산 과정에서 노이즈가 부가되는 단계의 개수에 해당하는 횟수만큼 추론을 수행하되, 상기 확대된 이미지의 임베딩이 잠재 인코더에 입력되는,

고해상도 변환 방법.

청구항 5

제1항에 있어서,

상기 고해상도 잔차를 생성하는 단계에서 수행되는 추론 동작들 중 제1 추론 동작은 상기 추출된 임의의 벡터를 상기 학습이 완료된 모델의 디코더에 입력하여 수행되는,

고해상도 변환 방법.

청구항 6

삭제

청구항 7

삭제

청구항 8

삭제

청구항 9

제4항에 있어서,

상기 고해상도 잔차를 생성하는 단계에서 수행되는 추론 동작들 중 제1 추론 동작은 상기 추출된 임의의 벡터를 상기 학습이 완료된 모델의 잠재 디코더에 입력하여 수행되는,

고해상도 변환 방법.

청구항 10

제1항에 있어서,

상기 원본 이미지에 대응하는 잔차 이미지는,

선형보간을 통해 상기 원본 이미지를 저해상도 이미지로 변환하는 단계;

선형보간을 통해 변환된 상기 저해상도 이미지를 상기 원본 이미지의 해상도로 확대하는 단계; 및

상기 원본 이미지와 상기 원본 이미지의 해상도로 확대된 이미지의 차를 통해 상기 잔차 이미지를 생성하는 단계를 통해 생성되는,

고해상도 변환 방법.

발명의 설명

기술 분야

본 발명은 컴퓨터 단층촬영(Computed Tomography, CT) 영상의 잡음(noise)을 저감하는 방법에 관한 것으로, CT 영상에서 생기는 잡음을 제거하는 방식을 통해 고해상도 CT 영상을 생성하는 기법에 관한 것이다. 저해상도 CT 영상의 질을 개선하여 의료 현장에서 보다 정확한 진단을 기대할 수 있다.

배경 기술

[0001]

[0002] CT(Computed Tomography)와 같은 비침습(non-invasive) 3D 이미징 기술은 신체 내부를 재구성하기 위해 투과된 신호를 사용한다. 이 기술은 실제로 신체를 직접 들여다보는 것이 아니라, 투과된 신호를 분석하여 내부를 재구성하는 방식을 사용한다. 이 방법은 성공적으로 신체를 투영해 내었지만, 여전히 세밀한 부위를 확인하기에는 잡음(noise)이 많은 상태이다. 특히, CT의 경우 빠르게 촬영되는 장점이 있으나, 다른 영상 기법(e.g. 자기공명 영상(magnetic resonance imaging))과 비교하였을 때 더 많은 잡음을 포함한다. 일부 질병들은 병세가 많이 진행되어 증상이 뚜렷해 쉽게 찾아낼 수 있으므로, 이러한 질병들의 경우 저해상도 영상 혹은 잡음이 많이 섞여 있는 영상을 통해서도 진단이 가능하다. 하지만, 질환들의 조기 진단을 이러한 영상들로 수행하기에는 아직 많은 어려움이 따른다.

[0003] 따라서, 이미 낮은 화질로 재구성된 영상들의 경우 사후 검증을 통한 추가적인 연구를 이루어 내기 위해 저해상도 이미지를 고해상도로 변환하는 기술이 필요하다. 이를 위해 본 발명은 잡음이 많이 섞인 저해상도의 CT 영상을 고해상도 CT 영상으로 변환하는 문제를 해결하고자 한다. 이를 위해 확산 모델 기반의 딥러닝 방법론을 사용하며, 저해상도 영상을 고해상도로 변환하고 세밀한 부위까지 재구성할 수 있도록 한다. 이러한 기술은 질환의 조기 진단과 치료에 매우 중요한 역할을 할 수 있다. 또한, 해상도가 향상된 영상을 통해 역으로 영상 취득 장치의 기술 향상을 기대할 수 있다.

선행기술문헌

특허문헌

- [0004] (특허문헌 0001) 대한민국 등록특허 제2414299호 (2022.06.29. 공고)
 (특허문헌 0002) 대한민국 등록특허 제2059529호 (2019.12.26. 공고)
 (특허문헌 0003) 대한민국 등록특허 제2399792호 (2022.05.23. 공고)
 (특허문헌 0004) 대한민국 공개특허 제2022-0131809호 (2022.09.29. 공개)

발명의 내용

해결하려는 과제

[0005] 본 발명이 이루고자 하는 기술적인 과제는 확산 모델 기반의 컴퓨터 단층촬영 고해상도 변환 장치 및 방법을 제공하는 것이다.

과제의 해결 수단

[0006] 본 발명의 일 실시예에 따른 고해상도 변환 방법은 적어도 프로세서(processor)를 포함하는 컴퓨팅 장치에 의해 수행되고, 확산 모델(Diffusion Model)을 이용하여 원본 이미지로부터 학습 데이터를 생성하는 단계, 상기 학습 데이터를 이용하여 인코더(endoder)와 디코더(decoder)를 포함하는 합성곱 신경망 모델(Convolutional Neural Network Model, CNN 모델)을 학습하는 단계, 입력 이미지를 선형 보간(linear interpolation)하여 확대된 이미지를 생성하는 단계, 가우시안 분포로부터 추출된 임의의 벡터와 상기 확대된 이미지를 이용하여 고해상도 잔차를 생성하는 단계, 및 상기 확대된 이미지와 상기 고해상도 잔차를 결합하여 고해상도 이미지를 생성하는 단계를 포함한다.

발명의 효과

[0007] 본 발명의 실시예에 따른 확산 모델 기반의 컴퓨터 단층촬영 고해상도 변환 장치 및 방법에 의할 경우, 저해상도의 컴퓨터 단층촬영 이미지로부터 고해상도의 컴퓨터 단층촬영 이미지를 생성할 수 있다.

도면의 간단한 설명

- [0008] 본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 상세한 설명이 제공된다.
 도 1은 고해상도 이미지에 대한 확산 과정을 설명하기 위한 도면이다.
 도 2a는 학습 모델의 학습 과정을 설명하기 위한 개념도이다.

도 2b는 학습 모델의 학습 과정을 보다 상세하게 설명하기 위한 개념도이다.

도 2c는 학습 모델에 포함되는 인코더의 구조를 도시한다.

도 2d는 학습 모델에 포함되는 디코더의 구조를 도시한다.

도 3은 다양한 HU 윈도우에 따른 CT 영상을 도시한다.

도 4는 학습된 모델을 이용하여 추론하는 과정을 설명하기 위한 개념도이다.

도 5에는 잠재 확산을 이용한 학습 방법에 대한 개요도이다.

도 6은 본 발명의 일 실시예에 따른 고해상도 변환 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0009] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시예들에 대해서 특정한 구조적 또는 기능적 설명들은 단지 본 발명의 개념에 따른 실시예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시예들은 다양한 형태들로 실시될 수 있으며 본 명세서에 설명된 실시예들에 한정되지 않는다.
- [0010] 본 발명의 개념에 따른 실시예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시예들을 도면에 예시하고 본 명세서에서 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시예들을 특정한 개시 형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물, 또는 대체물을 포함한다.
- [0011] 제1 또는 제2 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 상기 구성 요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소로부터 구별하는 목적으로만, 예컨대 본 발명의 개념에 따른 권리 범위로부터 벗어나지 않은 채, 제1 구성 요소는 제2 구성 요소로 명명될 수 있고 유사하게 제2 구성 요소는 제1 구성 요소로도 명명될 수 있다.
- [0012] 어떤 구성 요소가 다른 구성 요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성 요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성 요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성 요소가 다른 구성 요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는 중간에 다른 구성 요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성 요소들 간의 관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.
- [0013] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 본 명세서에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0014] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0015] 이하, 본 명세서에 첨부된 도면들을 참조하여 본 발명의 실시예들을 상세히 설명한다. 그러나, 특허출원의 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0016] 딥러닝(deep learning) 모델들은 주어진 입력에 대해 대응되는 정답을 맞추는 방향으로 학습된다. 고해상도 모델(super-resolution model)의 경우도 마찬가지로, 저해상도의 이미지가 제공되었을 때 기대되는 고해상도 이미지를 예측하도록 학습된다. 이때, 학습에 이용되는 한 쌍의 저해상도/고해상도 이미지는 원본 고해상도 이미지를 고정된 과정을 통해 저해상도 이미지로 변환함으로써 생성된다. 저해상도 이미지로 변환하는 과정은 흐림(blurring) 효과 부가, 잡음(예컨대, 가우시안 잡음(Gaussian noise)) 추가, 및/또는 다운 샘플링(down-

sampling) 방법을 통해 진행된다.

- [0017] 상술한 방법을 통해 저해상도 이미지를 생성하는 과정에서 발생하는 가장 큰 문제는, 저해상도 이미지와 고해상도 이미지는 일대일 대응이 아닌 일대다 대응이기 때문에 다양한 고해상도 이미지가 생성될 수 있다는 것이다. 기존의 합성곱 기반 고해상도 모델은 이를 일대일로 해결하려는 문제를 가지고 있다. 이는 의도적인 이미지 손상이 확률 과정을 거치지 않고 고정된 일련의 과정을 반복하기 때문이다. 이렇게 일대일 대응의 형태로 학습을 진행하면 매우 많은 데이터가 요구된다. 또한, 의료 영상의 특성상 일반적인 이미지 데이터보다 그 수가 매우 부족하기 때문에 해당 방법론을 적용하기는 어려움이 있다. 또 다른 문제는 저해상도 이미지를 단 한 번의 단계로 고해상도로 끌어올리는 과정에서 흐림(blurring)이 발생하는 것이다.
- [0018] 본 발명에서 제안하는 고해상도 모델은 고해상도 이미지 및/또는 잔차 이미지를 여러 단계에 걸쳐 순수한 잡음을 만드는 확산 과정을 수행한다. 잔차 이미지를 생성하기 위해서, 원본 고해상도 이미지(미리 정해진 해상도(예컨대, 512*512 해상도)를 갖는 고해상도 이미지로써 제1 영상이라 명명할 수 있음)를 선형보간을 통해 저해상도(미리 정해진 해상도(예컨대, 256*256 해상도) 이미지(제2 영상이라 명명할 수 있음)를 생성하고, 제2 영상을 다시 선형보간을 통해 확대 이미지(원본 고해상도 이미지와 동일한 해상도를 갖도록 확대된 이미지로써, 제3 영상이라 명명할 수 있음)를 생성할 수 있다. 결국, 제1 영상에서 제3 영상을 뺀(각 대응되는 픽셀 값을 차감하는 것으로 이해될 수 있음) 잔차 영상(잔차 이미지)을 생성할 수 있다.
- [0019] 이때, 확산 과정에 포함되는 단계의 수는 T (T 는 임의의 자연수로서, 예시적인 값은 500 내지 1000의 값을 갖음)개이다. 이와 관련하여, 고해상도 이미지(예컨대, 고해상도 CT 이미지)에 대한 확산 과정을 설명하기 위한 도면이 도 1에 도시되어 있다. 본 발명에서 제안하는 모델은 상술한 과정(확산 과정)의 역순서, 즉 잡음이 있는 영상으로부터 잡음을 조금씩 덜어내어진 이미지를 예측하도록 학습된다. 물론, 본 발명에서는 잔차 이미지에 확산 과정을 적용한다. 잡음을 제거하는 과정이 고해상도 이미지를 생성하는 이미지로 연결되는 이유는, 일반적으로 선형 보간을 통해 저해상도 영상을 고해상도로 강제로 업샘플링(upsampling)했을 때 뿌옇게 처리되는 부분들이 발생하는 데, 이러한 잡음 또한 확산 과정에서 생긴 잡음을 제거하는 학습을 통해 처리가 가능하기 때문이다.
- [0020] 본 발명에서 제안하는 방법을 통해 상술한 두 가지 문제를 해결할 수 있다. 첫째로, 확산 과정은 하나의 고정된 경로가 아닌 다양한 경로로 발생할 수 있다. 확산 과정은 확률적인 과정으로 잡음을 입히는 과정이 수행될 때마다 다른 방향으로 작동할 수 있어 훈련 과정 중에 다양한 복원 방식을 학습하는 것이 가능하다. 이는 고정된 방법으로 이미지를 손상시킨 기존의 방법들과 차이를 가지며, 이를 통해 일대다 문제를 해결하는 것이 가능해진다. 둘째로, 잡음이 제거된 영상을 여러 단계에 걸쳐 해결함으로써 모델이 흐릿한 영상을 만들지 않도록 조절할 수 있다.
- [0021] 이하에서는, 본 발명에 대해 보다 구체적으로 설명한다.
- [0022] 확산 모델을 통한 고해상도 변환은 훈련 단계와 추론 단계로 구성된다.
- [0023] **1. 훈련 단계**
- [0024] 확산 모델은 어떤 영상을 입력으로 받았을 때, 여러 단계에 걸쳐 영상에 대응하는 잔차 영상을 순수한 잡음이 되도록 한다. 순수한 잡음이 되는 확산 과정은 도 1에 도시되어 있다. 본 발명에서는, 확산 과정을 통해 생성된 복수의 이미지를 중 임의의 이미지를 선택하여 잡음이 제거된 이전 단계의 영상을 예측하도록 모델을 학습시킨다. 이렇게 잡음을 제거하는 방식을 학습하여 최종적으로 저해상도 CT 영상에서 발생하는 잡음을 제거할 수 있다. 이때, 실시예에 따라, 확산 과정을 통해 생성된 전체 복수의 이미지들 각각에 대하여 학습이 수행될 수 있다. 이 경우, 단일의 원본 고해상도 이미지에 대하여, 확산 과정에 포함된 단계들의 개수만큼의 학습이 진행될 수 있다. 구체적인 예로, 500($T=500$) 단계에 걸쳐서 원본 고해상도 이미지에 대한 잔차 영상이 순수한 잡음으로 변환되었다면, 생성된 500 개의 이미지들 각각에 대하여 전 단계의 이미지(즉, 잡음이 제거된 잔차 영상)를 예측하도록 학습될 수 있다.
- [0025] 학습 모델은 인코더(encoder)와 디코더(decoder)를 포함하는 적어도 두 부분으로 구성될 수 있다. 본 발명에서 사용되는 인코더-디코더 구조는 흔히 U-Net으로 알려진 영상 분할에 사용되는 대표적인 모델이다. 다시 말하면, 학습 모델은 인코더-디코더 구조의 합성곱 신경망(Convolutional Neural Network, CNN) 모델일 수 있다. 인코더는 주어진 입력에 대응하는 저차원의 벡터(잠재 벡터(latent vector)로 명명될 수 있음)를 출력한다. 디코더는 인코더가 출력한 벡터를 다시 입력 영상의 사이즈 영상으로 되돌리는 역할을 수행한다. 따라서, 모델은 잡음이 포함된 이미지를 받아, 이를 저차원 벡터로 출력한 후, 디코더가 이 벡터로부터 잡음이 제거된 이미지를 출력할 수 있도록 학습한다. 학습 모델의 학습 과정은 도 2a에 도시되어 있다. 도 2를 참조하면, 모델은 임의의

단계(예컨대, $t+1$ 시점)의 잡음 영상을 입력으로 받아 잡음이 한 단계 제거된(또는 건너진) 이전 단계(예컨대, t 시점)의 영상을 예측하도록 학습될 수 있다. 이때, 모델은 임의의 단계(예컨대, $t+1$ 시점)의 잡음 영상을 입력으로 받고, 추가적으로 시점 정보(예컨대, $t+1$)를 더 입력으로 받을 수도 있다.

[0026] 학습 모델의 학습 과정을 보다 상세하게 도시하는 도 2b를 참조하면, 모델은 임의의 단계의 잡음 영상으로부터 잡음이 한 단계 건너진 영상을 예측하도록 학습된다. 모델의 입력에는 $t+1$ 시점의 잡음 영상, 시점 정보(즉, $t+1$), 및 선형보간으로 확대된 이미지 총 3 가지를 포함할 수 있다. 선형보간으로 확대된 이미지는 원본 고해상도 이미지를 선형보간을 통해 축소 및 확대하는 과정을 통해 강제로 품질을 낮춘 고해상도 이미지를 의미할 수 있다. 또한, 도 2c에는 학습 모델에 포함되는 인코더의 구조가, 도 2d에는 학습 모델에 포함되는 디코더의 구조가 도시되어 있다. 특히, nn.Conv2d의 출력은 매개변수에 의해 입력대비 사이즈를 조절할 수 있다.

[0027] 이때, 확산에 사용되는 잡음(노이즈(noise))을 첨가하는 과정은, 일반적인 확산 모델에서 사용되는 과정과 상이할 수 있다. 일반적인 영상은 하나의 픽셀값이 0부터 255 사이의 정수값을 갖고, 이 값을 -1과 1 사이로 보낸 후(정규화한 후) 가우시안 잡음($N(0,1)$)을 첨가한다. 그러나, CT 영상의 최소 단위 픽셀은 0부터 255가 아닌 -1000부터 3000 사이의 HU(Hounsfield Unit) 값으로 이는 방사선 감쇄계수를 뜻한다. 이 값은 조직에 따라 다른 값을 부여받기 때문에, 보고자 하는 부위에 따라 가시화 가능한 HU 영역(이하, HU 윈도우(window))을 다르게 두어 영상을 보는 것이 가능하다(도 3 참조). 도 3의 좌측은 폐 내의 조직을 더 자세히 볼 수 있는 폐 HU 윈도우 설정이고, 도 3의 우측은 연조직을 더 자세히 보여주는 종격부 HU 윈도우이다.

[0028] 이렇게 같은 부분을 보더라도 HU 윈도우를 조절하여 목적으로 하는 부위를 다르게 보는 것이 가능하다. 예컨대, 도 3의 좌측과 같이 HU 윈도우를 Lung window로 조절하면 폐 내의 조직을 가시화하는 것이 가능하다. 반대로, 도 3의 우측처럼 Mediastinal Window를 사용하면 소위 폐의 종격부라 불리는 폐 외의 주변 연조직을 더 자세히 보는 것이 가능하다. 이러한 특성에 맞추어 본 발명은 일반적인 영상에 입히는 확산 과정과 다르게, HU 윈도우를 다르게 조절하여 다양한 이미지에 대해 잡음을 제거하는 방법을 제안한다. 즉, 하나의 영상으로 다양한 HU 윈도우 설정에 따른 잡음 제거 학습을 진행할 수 있다. 또한, 본 발명에서는 설정된 HU 윈도우를 미리 정해진 범위를 갖도록 정규화한 후 노이즈를 첨가할 수 있다.

[0029] 2. 추론 단계

[0030] 본 발명에서는 선형 보간을 통해 이미지 사이즈를 키운 저해상도 이미지와 확산 과정을 통해 학습된 딥러닝 모델의 고해상도 잔차를 더해 고해상도 CT 이미지를 생성한다. 딥러닝 모델의 출력만으로 고해상도 영상을 만들기도 하는, 확대된 선형보간된 영상과 모델이 예측한 잔차를 합하여 최종적인 고해상도 영상을 만들어 주는 것이 더욱 선명한 이미지를 얻을 수 있다. 다만, 본 발명이 이에 제한되는 것은 아니며, 선형 보간된 이미지를 이용하지 않고, 학습 모델의 결과만을 이용해 고해상도 CT 영상을 생성하는 것도 가능하다.

[0031] 보다 상세한 과정은 다음과 같다. 추론은 가우시안 분포로부터 하나의 벡터(z_0)를 추출하는 것으로 시작된다. 이를 학습된 모델에 반복적으로 통과시켜 잡음이 점차 제거되는 이미지(잔차 이미지를 의미할 수 있음)를 생성한다. 이때, 선형보간으로 확대된 원본 이미지를 인코더에 같이 통과시켜 추론 시 원본과 이미지가 상이하지 않도록 조절할 수 있다. 추론 과정에 대한 개략도는 도 4에 도시되어 있다. 추론 과정에서도, 학습 과정과 유사하게, 시점 정보가 인코더와 디코더의 입력으로 이용될 수 있다.

[0032] 이때, 확산 과정에 포함된 단계의 개수에 해당하는 횟수만큼의 추론 동작이 수행될 수 있다. 즉, 순수한 잡음(추출된 벡터(z_0)를 의미할 수 있음)으로부터 고해상도 잔차를 추론하기 까지 총 T 회의 추론 동작이 수행된다. 복수의 추론 동작 중 제1 추론 동작은 디코더만을 이용하여 수행되고, 제1 추론 동작을 제외한 나머지 추론 동작들은 이전 추론의 결과를 입력으로 받아 노이즈가 제거된 영상(잔차 영상)을 출력한다. 즉, 제1 추론 동작을 제외한 나머지 추론 동작들은 인코더와 디코더를 모두 사용하는 추론 동작이다.

[0033] [변형예]

[0034] 본 발명에서는 잠재 확산(Latent Diffusion)과 같은 방법을 사용할 수 있다. 이 방법은 직접 이미지(잔차 이미지를 의미할 수 있음)에 확산 과정을 입히는 것이 아니라, 인코더를 한 번 통과한 이미지 임베딩에 확산 과정을 거쳐준 후 잡음이 입혀진 임베딩을 만드는 것이다. 그 후 상술한 학습 과정과 유사하게 잡음이 제거된 임베딩을 예측한다. 이는 일반적으로 256×256 에서 512×512 차원을 가지는 고차원의 영상을 예측하는 것보다 훨씬 적은 저차원의 임베딩을 예측하기 때문에 속도가 많이 개선되는 이점을 갖는다. 이와 관련하여, 도 5에는 잠재 확산을 이용한 학습 방법에 대한 개요도가 도시되어 있다. 변형예에서 사용되는 모델은, 원본 이미지(잔차 이미지를

의미할 수도 있음)로부터 임베딩을 생성하는 인코더, 임베딩으로부터 고해상도 이미지(잔차 이미지를 의미할 수도 있음)를 추론하는 디코더, 및 임의의 단계의 임베딩으로부터 이전 단계의 임베딩을 추론하는 잠재 인코더(latent encoder, 제2 인코더라 명명될 수도 있음) 및 잠재 디코더(latent decoder, 제2 디코더라 명명될 수도 있음)를 포함한다.

[0035] 즉, 본 실시예에서, 순수한 잡음이 될 때까지의 T회에 걸친 확산 동작은 입력 이미지(고해상도 이미지로써 잔차 이미지를 의미할 수 있음)가 아닌 입력 이미지에 대한 임베딩(인코더의 출력인 잠재 벡터를 의미할 수 있음)을 대상으로 수행된다. 또한, 학습 모델은 임의의 단계(예컨대, t+1 시점)의 잠재 벡터의 잡음을 제거하여 이전 단계(예컨대, t 시점)의 잠재 벡터를 추론하도록 학습될 수 있다.

[0036] 또한, 학습이 완료된 모델을 이용한 추론은 가우시안 분포로부터 하나의 벡터를 추출하는 것으로 시작된다. 이를 학습된 모델에 반복적으로 통과시켜 잡음이 점차 제거된 임베딩 벡터를 생성한다. 이때, 선형보간으로 확대된 원본 이미지를 인코더에 통과시켜 얻은 원본 영상의 임베딩을 함께 디코더의 입력으로 활용하여 추론시 원본과 이미지가 상이하지 않도록 조절할 수 있다. 최종적으로 획득된 임베딩(z_0)을 디코더에 통과시킴으로써, 고해상도 이미지 또는 고해상도 잔차를 획득할 수 있다. 이때, 디코더의 출력을 최종 고해상도 이미지로 이용하거나, 예측된 고해상도 잔차를 선형보간된 이미지와 합하여 최종 고해상도 이미지를 생성할 수 있다.

[0037] 도 6은 본 발명의 일 실시예에 따른 고해상도 변환 방법을 설명하기 위한 흐름도이다. 고해상도 변환 방법은 고해상도 이미지 변환 방법, CT 고해상도 변환 방법 등과 같이 다양한 명칭으로 명명될 수 있다. 고해상도 변환 방법은 적어도 프로세서(processor) 및/또는 메모리(memory)를 포함하는 컴퓨팅 장치로 구현되는 고해상도 변환 장치에 의해 수행될 수 있다. 따라서, 고해상도 변환 방법을 이루는 단계들 중 적어도 일부는 컴퓨팅 장치에 포함되는 프로세서의 동작으로 이해될 수 있다. 컴퓨팅 장치는 PC(Personal Computer), 서버(server), 랩탑 컴퓨터, 태블릿 PC, 노트북 등을 포함한다. 이하, 고해상도 변환 방법을 설명함에 있어, 앞선 기재와 중복되는 내용에 관하여는 그 구체적인 설명을 생략하기로 한다.

[0038] 우선, 학습 데이터가 생성된다(S110). 학습 데이터는 제1 학습 데이터 및/또는 제2 학습 데이터를 포함할 수 있다.

[0039] 제1 학습 데이터는 원본 이미지 및/또는 확산 모델(Diffusion Model)을 이용하여 단계별로 노이즈가 부가된 이미지들(잔차 이미지들을 의미할 수 있음)을 포함한다. 즉, 임의의 단계에서 생성된 이미지(잔차 이미지를 의미할 수 있음)에 추가적인 노이즈가 부가됨으로써 다음 단계의 이미지(잔차 이미지를 의미할 수 있음)가 생성된다. 확산 모델을 적용한 결과, 원본 이미지(원본 이미지에 대응하는 잔차 이미지를 의미할 수 있음)로부터 순수한 잡음을 나타내는 이미지까지 복수의 이미지들이 생성될 수 있다. 따라서, 원본 이미지별로 복수의 이미지들이 생성될 수 있다. 원본 이미지는 미리 정해진 해상도를 갖는 고해상도 이미지를 나타낸다. 예시적으로 원본 이미지는 고해상도 CT 이미지를 의미할 수 있다.

[0040] 제2 학습 데이터는 원본 이미지(원본 이미지에 대응하는 잔차 이미지를 의미할 수 있음), 원본 이미지의 임베딩(잔차 이미지의 임베딩을 의미할 수 있음), 및/또는 확산 모델을 이용하여 단계별로 노이즈가 부가된 임베딩들을 포함한다. 즉, 임의의 단계에서 생성된 임베딩에 추가적인 노이즈가 부가됨으로써 다음 단계의 임베딩이 생성된다. 확산 모델을 적용한 결과, 원본 이미지(원본 이미지에 대응하는 잔차 이미지를 의미할 수 있음)의 임베딩으로부터 순수한 잡음을 나타내는 임베딩까지 복수의 임베딩들이 생성될 수 있다. 따라서, 원본 이미지별로 복수의 임베딩들이 생성될 수 있다. 원본 이미지는 미리 정해진 해상도를 갖는 고해상도 이미지를 나타낸다. 예시적으로 원본 이미지는 고해상도 CT 이미지를 의미할 수 있다.

[0041] 실시예에 따라, 원본 이미지에 대한 임의의 미리 정해진 HU 윈도우가 설정된 후 학습 데이터가 생성될 수도 있다. 이를 통해, 단일의 원본 이미지로부터 서로 다른 HU 윈도우를 갖는 복수의 학습 데이터들이 생성될 수 있다. 이를 통해 생성된 학습 데이터는 HU 윈도우가 설정된 이미지를 추론하도록 학습될 수 있다.

[0042] 모델에 대한 학습이 수행된다(S120). 모델은 인코더-디코더 구조의 신경망 모델을 의미할 수 있다. 예컨대, 모델은 인코더-디코더 구조를 갖는 합성곱 신경망일 수 있다. 다만, 본 발명이 이용되는 모델에 제한되는 것은 아니며, 모델은 인코더-디코더 구조를 갖는 것으로 충분하다.

[0043] 제1 학습 데이터를 이용하여 학습을 수행하는 경우, 원본 이미지, 원본 이미지에 대응하는 잔차 이미지, 및 확산 모델을 통해 생성된 이미지들 중 적어도 일부를 이용하여 학습을 진행한다. 구체적으로, 임의의 단계의 이미지로부터 노이즈가 덜 부가된 이전 단계의 이미지를 추론하도록 학습이 진행될 수 있다. 학습은, 각 원본 이미지별로, 순수한 노이즈를 나타내는 T번째 이미지부터 원본 이미지(원본 이미지에 대응하는 잔차 이미지를 의미

할 수 있음)까지의 모든 이미지를 이용하여 진행될 수 있다.

- [0044] 제2 학습 데이터를 이용하여 학습을 수행하는 경우, 원본 이미지, 원본 이미지에 대응하는 잔차 이미지, 및 확산 모델을 통해 생성된 임베딩들 중 적어도 일부를 이용하여 학습을 진행한다. 구체적으로, 잠재 인코더와 잠재 디코더가, 임의의 단계의 임베딩으로부터 노이즈가 덜 부가된 이전 단계의 임베딩을 추론하도록 학습이 진행될 수 있다. 학습은, 각 원본 이미지별로, 순수한 노이즈를 나타내는 T번째 임베딩부터 원본 이미지의 잔차 이미지에 대한 임베딩까지의 모든 임베딩을 이용하여 진행될 수 있다.
- [0045] 모델에 대한 학습이 완료된 후, 입력 데이터가 수신된다(S130). 입력 데이터는 미리 정해진 해상도를 갖는 (저해상도) 이미지로서, 예시적으로 저해상도 CT 이미지일 수 있다. 입력 데이터는 소정의 유무선 통신망을 통하여 소정의 단말로부터 수신되거나, 본 방법을 수행하는 컴퓨팅 장치에 미리 저장되어 있을 수 있다.
- [0046] 입력 데이터에 대한 선형 보간을 수행하여 선형 보간된 이미지가 생성된다(S140). 선형 보간된 이미지는 선형 보간을 통해 확대된 이미지로써, 예시적으로 입력 데이터의 사이즈가 256×256 이라면 선형 보간된 이미지의 사이즈는 512×512 일 수 있다. 또한, 선형 보간된 이미지를 학습 모델에 포함된 인코더에 투영하여 입력 이미지 (또는 선형 보간된 이미지)에 대응하는 임베딩을 생성할 수 있다.
- [0047] 추론 동작이 수행된다(S150). 추론 동작은 두 가지 동작 중 어느 하나를 의미할 수 있다. 제1 추론 동작은 제1 학습 데이터를 이용하여 학습된 모델을 이용한 추론 동작이고, 제2 추론 동작은 제2 학습 데이터를 이용하여 학습된 모델을 이용한 추론 동작을 의미할 수 있다.
- [0048] 제1 추론 동작을 위해, 가우시안 분포로부터 임의의 벡터가 추출된다. 다음으로, 학습된 모델을 이용하여 제1 학습 데이터를 생성하기 위한 확산 과정에서 생성된 이미지들의 개수에 해당하는 만큼의 추론 동작을 수행한다. 이때, 각 추론 동작에서, 모델의 인코더는 입력 이미지에 대응하는 임베딩(선형보간된 확대 이미지의 임베딩을 의미할 수 있음)을 입력으로 더 이용한다. 물론, 첫번째 추론 동작의 경우, 추출된 벡터를 모델의 디코더에 입력하여 잡음이 완화된 이미지를 생성한다. 즉, 첫번째 추론에서 인코더의 역할은 임의의 벡터를 추출하는 것으로 대체된다. T회에 걸친 추론 동작에 따라, 고해상도 이미지 및/또는 고해상도 잔차가 생성된다.
- [0049] 제2 추론 동작을 위해, 가우시안 분포로부터 임의의 벡터가 추출된다. 다음으로, 학습된 모델을 이용하여 제2 학습 데이터를 생성하기 위한 확산 과정에서 생성된 임베딩들의 개수에 해당하는 만큼의 추론 동작이 수행된다. 이때, 각 추론 동작에서, 모델의 잠재 인코더는 입력 이미지에 대응하는 임베딩(선형보간된 확대 이미지의 임베딩을 의미할 수 있음)을 입력으로 더 이용할 수 있다. 물론, 첫번째 추론 동작의 경우, 추출될 벡터를 모델의 잠재 디코더에 입력하여 잡음이 완화된 임베딩을 생성한다. 즉, 첫번째 추론에서 잠재 인코더의 역할은 임의의 벡터를 추출하는 것으로 대체될 수 있다. T회에 걸친 추론 동작에 따라, 최종 임베딩이 생성되고, 임베딩을 모델의 디코더에 입력함으로써 고해상도 이미지 및/또는 고해상도 잔차가 생성될 수 있다. 이 경우에도, 디코더는 최종 임베딩과 입력 데이터에 대응하는 임베딩을 입력으로 이용할 수 있다.
- [0050] 마지막으로, 최종 고해상도 이미지, 예컨대 고해상도 CT 이미지가 생성된다(S160). 최종 고해상도 이미지는 선형 보간된 이미지와 고해상도 잔차의 결합으로 생성될 수 있다.
- [0051] 이상에서 설명된 장치는 하드웨어 구성 요소, 소프트웨어 구성 요소, 및/또는 하드웨어 구성 요소 및 소프트웨어 구성 요소의 집합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성 요소는, 예를 들어, 프로세서, 컨트롤러, ALU(Arithmetic Logic Unit), 디지털 신호 프로세서(Digital Signal Processor), 마이크로 컴퓨터, FPA(Field Programmable 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(Operation System, 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 기술 분야에서 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 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(Parallel Processor)와 같은, 다른 처리 구성(Processing Configuration)도 가능하다.
- [0052] 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 처

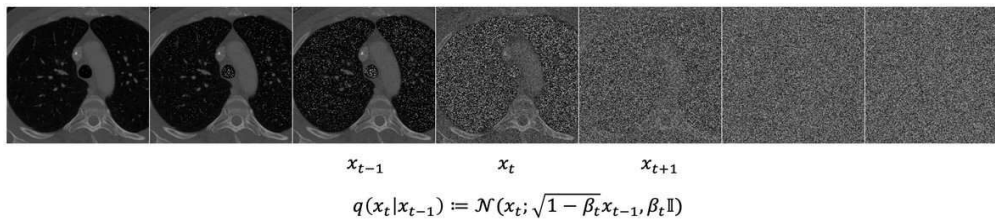
리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성 요소(Component), 물리적 장치, 가상 장치(Virtual Equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(Signal Wave)에 영구적으로, 또는 일시적으로 구체화(Embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0053] 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), 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. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

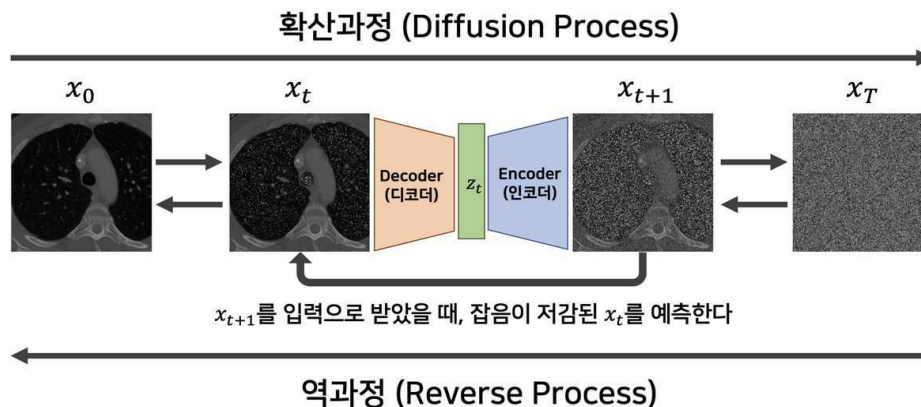
[0054] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성 요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성 요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

도면

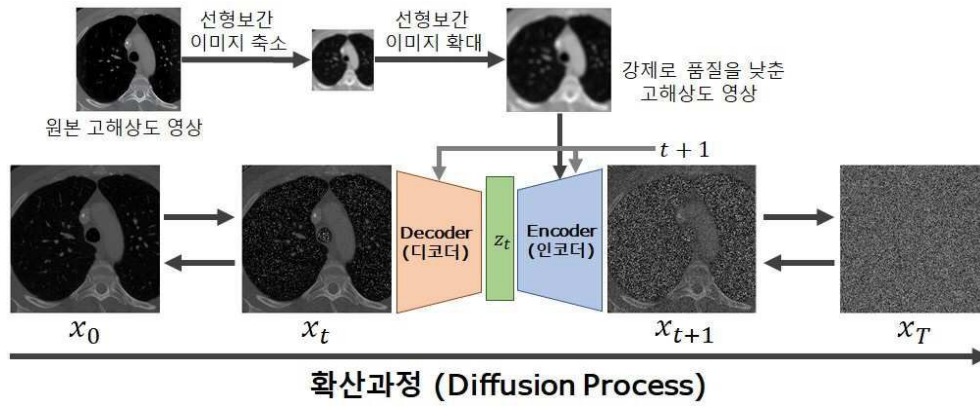
도면1



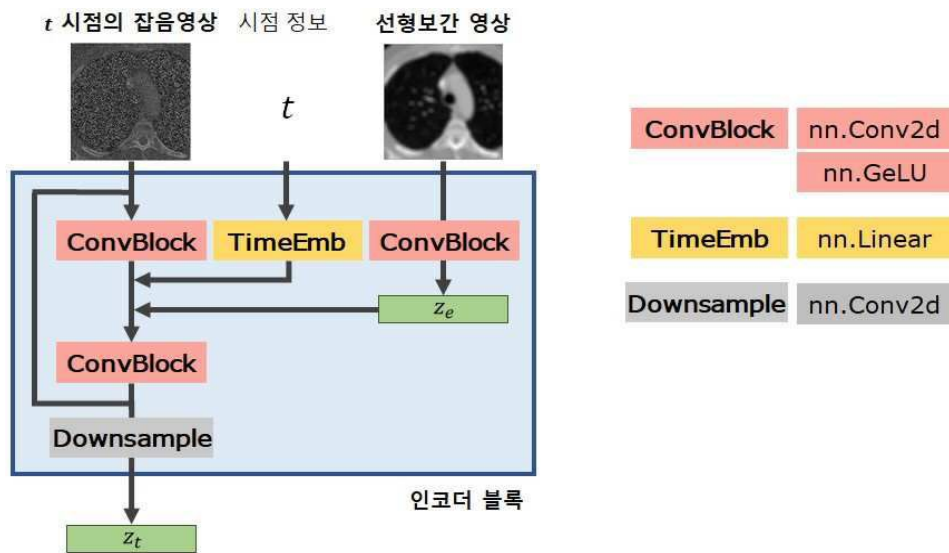
도면2a



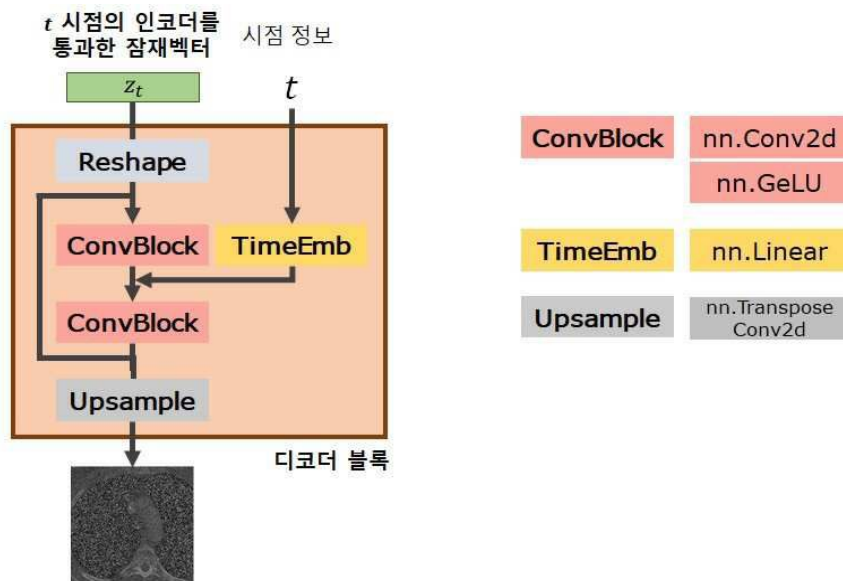
도면2b



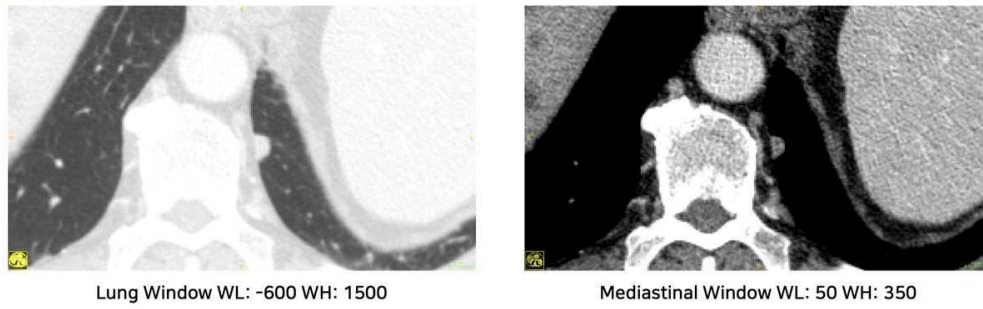
도면2c



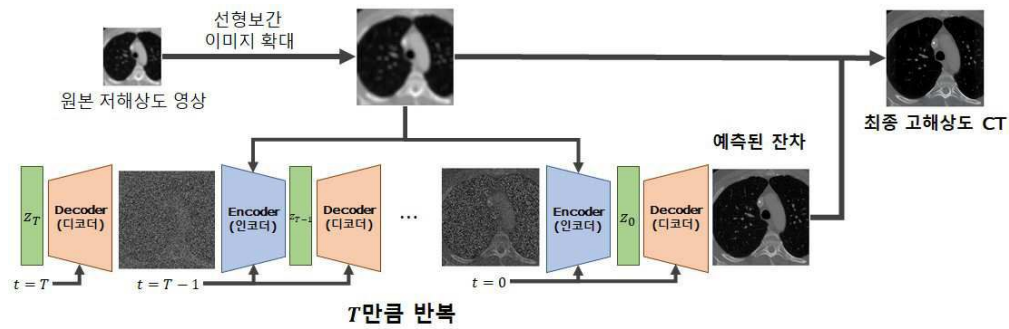
도면2d



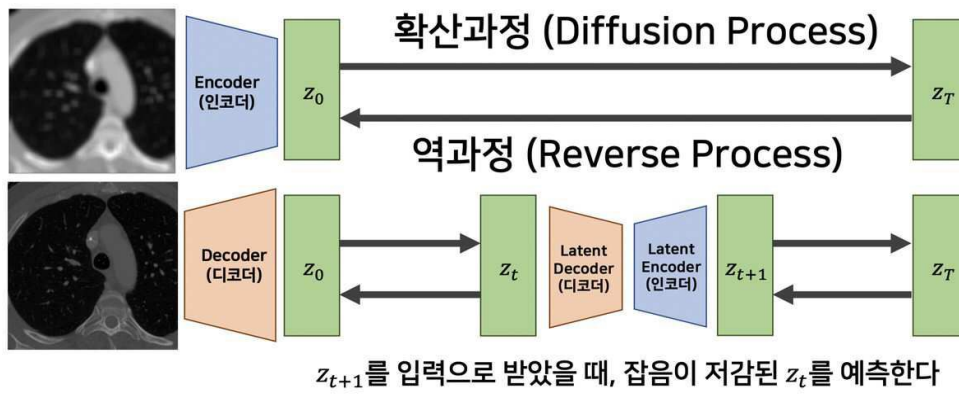
도면3



도면4



도면5



도면6

