



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0048823
(43) 공개일자 2022년04월20일

(51) 국제특허분류(Int. Cl.)
G06T 3/00 (2019.01) G06K 9/00 (2022.01)
G06T 11/60 (2006.01)
(52) CPC특허분류
G06T 3/0093 (2013.01)
G06T 11/60 (2013.01)
(21) 출원번호 10-2020-0132128
(22) 출원일자 2020년10월13일
심사청구일자 2020년10월13일

(71) 출원인
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
변혜란
서울특별시 서대문구 연세로 50, 제4공학관 810호(신촌동, 연세대학교)
전석규
서울특별시 서대문구 연세로 50, 제4공학관 810호(신촌동, 연세대학교)
(74) 대리인
특허법인우인
(뒷면에 계속)

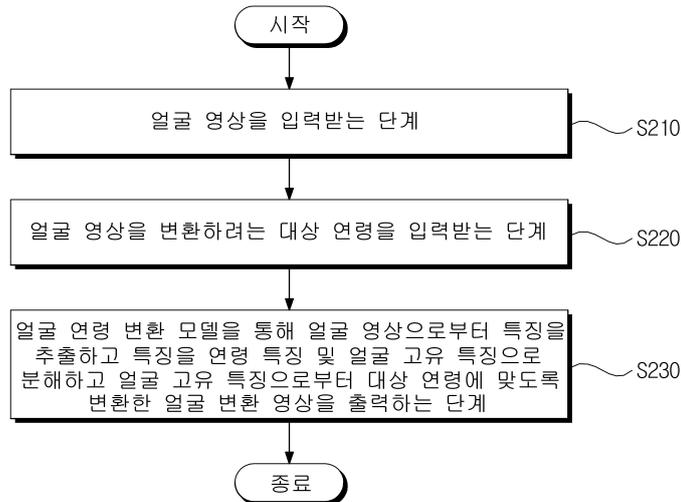
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 **특징 표현 단순화와 다중 작업 학습을 통한 얼굴 영상 연령 변환 방법 및 장치**

(57) 요약

본 실시예들은 연속적 변수인 대상 연령의 특징을 분석하고 영상의 연령 특징과 인물 특징을 특징 공간에서 분리하여 인물의 고유 특징을 보존하면서 대상 연령으로 변환하는 얼굴 영상 연령 변환 방법 및 장치를 제공한다.

대표도 - 도7



(52) CPC특허분류

G06V 40/168 (2022.01)
 G06T 2207/20081 (2013.01)
 G06T 2207/30201 (2013.01)

도미래

서울특별시 서대문구 연세로 50, 제4공학관 810호
 (신촌동, 연세대학교)

(72) 발명자

홍기범

서울특별시 서대문구 연세로 50, 제4공학관 810호
 (신촌동, 연세대학교)

김명진

서울특별시 서대문구 연세로 50, 제4공학관 810호
 (신촌동, 연세대학교)

이 발명을 지원한 국가연구개발사업

| | |
|-------------------------|---|
| 과제고유번호 | 1711112608 |
| 과제번호 | 2019R1A2C2003760 |
| 부처명 | 과학기술정보통신부 |
| 과제관리(전문)기관명 | 한국연구재단 |
| 연구사업명 | 중견연구자지원사업 |
| 연구과제명 | 특성 정보 자동 생성을 통한 처음 보는 복합카테고리의 이미지와 비디오 생성 및 인 |
| 식을 위한 제로샷 학습 기술 연구(2/3) | |
| 기 여 율 | 1/1 |
| 과제수행기관명 | 연세대학교 |
| 연구기간 | 2020.03.01 ~ 2021.02.28 |

명세서

청구범위

청구항 1

컴퓨팅 디바이스에 의한 얼굴 영상 연령 변환 방법에 있어서,

얼굴 영상을 입력받는 단계;

상기 얼굴 영상을 변환하려는 대상 연령을 입력받는 단계; 및

얼굴 영상 변환 모델을 통해 상기 얼굴 영상으로부터 특징을 추출하고 상기 특징을 연령 특징 및 얼굴 고유 특징으로 분해하고 상기 얼굴 고유 특징으로부터 상기 대상 연령에 맞도록 변환한 얼굴 변환 영상을 출력하는 단계를 포함하는 얼굴 영상 연령 변환 방법.

청구항 2

제1항에 있어서,

상기 얼굴 영상 변환 모델은 상기 연령 특징을 상기 특징의 놈(Norm)으로 산출하고, 상기 얼굴 고유 특징을 상기 특징에서 상기 특징의 놈으로 나눠서 산출하는 것을 특징으로 하는 얼굴 영상 연령 변환 방법.

청구항 3

제1항에 있어서,

상기 얼굴 영상을 입력받는 단계는 상기 얼굴 영상에 대한 인물의 원래 연령을 입력받는 단계 및 상기 얼굴 영상에 대한 인물의 고유 정보를 입력받는 단계를 수행하고,

상기 얼굴 영상 변환 모델은 상기 원래 연령을 기초로 상기 연령 특징으로부터 예상 연령을 추정하도록 학습되며,

상기 얼굴 영상 변환 모델은 상기 고유 정보를 기초로 상기 얼굴 고유 특징으로부터 예상 고유 정보를 추정하도록 학습되는 것을 특징으로 하는 얼굴 영상 연령 변환 방법.

청구항 4

제3항에 있어서,

상기 얼굴 영상 변환 모델은 제1 인코더(Encoder) 및 상기 제1 인코더에 연결된 디코더(Decoder)를 포함하는 학습 네트워크 구조이고, 상기 제1 인코더에 연결된 제1 연령 회귀자(Age Regressor) 및 상기 제1 인코더에 연결된 제1 신원 분류자(Identity Classifier)를 포함하는 것을 특징으로 하는 얼굴 영상 연령 변환 방법.

청구항 5

제4항에 있어서,

상기 얼굴 영상 변환 모델은,

상기 제1 인코더에 의해 추출한 특징을 이용하여 상기 제1 연령 회귀자에 의해 추정한 예상 연령에 대한 제1 연령 회귀 손실 함수를 최소화하도록 학습되고,

상기 제1 인코더에 의해 추출한 특징을 이용하여 상기 제1 신원 분류자에 의해 추정한 예상 고유 정보에 대한 제1 고유 정보 손실 함수를 최소화하도록 학습되는 것을 특징으로 하는 얼굴 영상 연령 변환 방법.

청구항 6

제4항에 있어서,

상기 얼굴 영상 변환 모델은 얼굴 영상 판별 모델과 상호 작용하여 학습되며,

상기 얼굴 연령 판별 모델은 제2 인코더 및 상기 제2 인코더에 연결된 판별자 레이어를 포함하는 학습 네트워크 구조이고, 상기 제2 인코더에 연결된 제2 연령 회귀자 및 상기 제2 인코더에 연결된 제2 신원 분류자를 포함하는 것을 특징으로 하는 얼굴 영상 연령 변환 방법.

청구항 7

제6항에 있어서,

상기 얼굴 연령 변환 모델은,

상기 얼굴 연령 변환 모델이 변환한 상기 얼굴 변환 영상이 상기 얼굴 연령 판별 모델에 의해 실제 영상으로 판단되도록 하는 제1 생성 손실 함수를 최소화하도록 학습되는 것을 특징으로 하는 얼굴 영상 연령 변환 방법.

청구항 8

제6항에 있어서,

상기 얼굴 연령 변환 모델은,

상기 얼굴 연령 변환 모델이 변환한 상기 얼굴 변환 영상이 상기 얼굴 연령 판별 모델에 의해 상기 대상 연령으로 판별되도록 하는 연령 손실 함수를 최소화하도록 학습되는 것을 특징으로 하는 얼굴 영상 연령 변환 방법.

청구항 9

제6항에 있어서,

상기 얼굴 연령 변환 모델은,

상기 얼굴 연령 변환 모델이 변환한 상기 얼굴 변환 영상을 상기 얼굴 연령 변환 모델을 통해 상기 원래 연령으로 변환할 때 상기 입력받은 얼굴 영상과 동일하도록 하는 일치성 손실 함수를 최소화하도록 학습되는 것을 특징으로 하는 얼굴 영상 연령 변환 방법.

청구항 10

제6항에 있어서,

상기 얼굴 연령 변환 모델은,

상기 얼굴 연령 변환 모델이 상기 원본 연령을 상기 대상 연령으로 설정하여 변환한 상기 얼굴 변환 영상이 상기 입력받은 얼굴 영상과 동일하도록 하는 고유성 매핑 손실 함수를 최소화하도록 학습되는 것을 특징으로 하는 얼굴 영상 연령 변환 방법.

청구항 11

제6항에 있어서,

상기 얼굴 연령 변환 모델은,

상기 얼굴 연령 변환 모델이 변환한 상기 얼굴 변환 영상의 얼굴 고유 특징과 상기 입력받은 얼굴 영상의 얼굴 고유 특징이 유사하도록 하는 고유성 재구성 손실 함수를 최소화하도록 학습되는 것을 특징으로 하는 얼굴 영상 연령 변환 방법.

청구항 12

제6항에 있어서,

상기 얼굴 연령 판별 모델은,

상기 제2 인코더에 의해 추출한 특징을 이용하여 상기 제2 연령 회귀자에 의해 추정된 예상 연령에 대한 제2 연령 회귀 손실 함수를 최소화하도록 학습되고,

상기 제2 인코더에 의해 추출한 특징을 이용하여 상기 제2 신원 분류자에 의해 추정된 예상 고유 정보에 대한 제2 고유 정보 손실 함수를 최소화하도록 학습되는 것을 특징으로 하는 얼굴 영상 연령 변환 방법.

청구항 13

제6항에 있어서,

상기 얼굴 연령 판별 모델은,

상기 입력받은 얼굴 영상이 실제 영상으로 판단되고 상기 얼굴 연령 변환 모델이 변환한 상기 얼굴 변환 영상이 모조 영상으로 판단되도록 하는 제2 생성 손실 함수를 최소화하도록 학습되는 것을 특징으로 하는 얼굴 영상 연령 변환 방법.

청구항 14

하나 이상의 프로세서 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 얼굴 영상 연령 변환 장치에 있어서,

상기 프로세서는,

얼굴 영상을 입력받고,

상기 얼굴 영상을 변환하려는 대상 연령을 입력받고,

얼굴 연령 변환 모델을 통해 상기 얼굴 영상으로부터 특징을 추출하고 상기 특징을 연령 특징 및 얼굴 고유 특징으로 분해하고 상기 얼굴 고유 특징으로부터 상기 대상 연령에 맞도록 변환한 얼굴 변환 영상을 출력하는 것을 특징으로 하는 얼굴 영상 연령 변환 장치.

발명의 설명

기술 분야

[0001] 본 발명이 속하는 기술 분야는 특징 표현 단순화와 다중 작업 학습을 통한 얼굴 영상 연령 변환 방법 및 장치에 관한 것이다.

배경 기술

[0002] 이 부분에 기술된 내용은 단순히 본 실시예에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.

[0003] 기존의 연령 변환 모델은 그룹 단위로 카테고리가 정해진 영상들을 이용하여 연령 변환을 적용한다. 기존의 연령 변환 모델은 특정 연령으로 변환한 영상에 대해 정확한 연령 추정이 불가능하며, 특정한 연령으로의 변환이 제한적인 문제가 있다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 한국등록특허공보 제10-1981634호 (2019.05.17)

발명의 내용

해결하려는 과제

[0005] 본 발명의 실시예들은 연속적 변수인 대상 연령의 특징을 분석하고 영상의 연령 특징과 인물 특징을 특징 공간에서 분리하여 인물의 고유 특징을 보존하면서 대상 연령으로 변환하는 데 주된 목적이 있다.

[0006] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다.

과제의 해결 수단

- [0007] 본 실시예의 일 측면에 의하면, 컴퓨팅 디바이스에 의한 얼굴 영상 연령 변환 방법에 있어서, 얼굴 영상을 입력 받는 단계, 상기 얼굴 영상을 변환하려는 대상 연령을 입력받는 단계, 및 얼굴 연령 변환 모델을 통해 상기 얼굴 영상으로부터 특징을 추출하고 상기 특징을 연령 특징 및 얼굴 고유 특징으로 분해하고 상기 얼굴 고유 특징으로부터 상기 대상 연령에 맞도록 변환한 얼굴 변환 영상을 출력하는 단계를 포함하는 얼굴 영상 연령 변환 방법을 제공한다.
- [0008] 상기 얼굴 연령 변환 모델은 상기 연령 특징을 상기 특징의 놈(Norm)으로 산출하고, 상기 얼굴 고유 특징을 상기 특징에서 상기 특징의 놈으로 나눠서 산출할 수 있다.
- [0009] 상기 얼굴 영상을 입력받는 단계는 상기 얼굴 영상에 대한 인물의 원래 연령을 입력받는 단계 및 상기 얼굴 영상에 대한 인물의 고유 정보를 입력받는 단계를 수행할 수 있다.
- [0010] 상기 얼굴 연령 변환 모델은 상기 원래 연령을 기초로 상기 연령 특징으로부터 예상 연령을 추정하도록 학습되며, 상기 얼굴 연령 변환 모델은 상기 고유 정보를 기초로 상기 얼굴 고유 특징으로부터 예상 고유 정보를 추정하도록 학습될 수 있다.
- [0011] 상기 얼굴 연령 변환 모델은 제1 인코더(Encoder) 및 상기 제1 인코더에 연결된 디코더(Decoder)를 포함하는 학습 네트워크 구조이고, 상기 제1 인코더에 연결된 제1 연령 회귀자(Age Regressor) 및 상기 제1 인코더에 연결된 제1 신원 분류자(Identity Classifier)를 포함할 수 있다.
- [0012] 상기 얼굴 연령 변환 모델은, 상기 제1 인코더에 의해 추출한 특징을 이용하여 상기 제1 연령 회귀자에 의해 추정된 예상 연령에 대한 제1 연령 회귀 손실 함수를 최소화하도록 학습되고, 상기 제1 인코더에 의해 추출한 특징을 이용하여 상기 제1 신원 분류자에 의해 추정된 예상 고유 정보에 대한 제1 고유 정보 손실 함수를 최소화하도록 학습될 수 있다.
- [0013] 상기 얼굴 연령 변환 모델은 얼굴 연령 판별 모델과 상호 작용하여 학습되며, 상기 얼굴 연령 판별 모델은 제2 인코더 및 상기 제2 인코더에 연결된 판별자 레이어를 포함하는 학습 네트워크 구조이고, 상기 제2 인코더에 연결된 제2 연령 회귀자 및 상기 제2 인코더에 연결된 제2 신원 분류자를 포함할 수 있다.
- [0014] 상기 얼굴 연령 변환 모델은, 상기 얼굴 연령 변환 모델이 변환한 상기 얼굴 변환 영상이 상기 얼굴 연령 판별 모델에 의해 실제 영상으로 판단되도록 하는 제1 생성 손실 함수를 최소화하도록 학습될 수 있다.
- [0015] 상기 얼굴 연령 변환 모델은, 상기 얼굴 연령 변환 모델이 변환한 상기 얼굴 변환 영상이 상기 얼굴 연령 판별 모델에 의해 상기 대상 연령으로 판별되도록 하는 연령 손실 함수를 최소화하도록 학습될 수 있다.
- [0016] 상기 얼굴 연령 변환 모델은, 상기 얼굴 연령 변환 모델이 변환한 상기 얼굴 변환 영상을 상기 얼굴 연령 변환 모델을 통해 상기 원래 연령으로 변환할 때 상기 입력받은 얼굴 영상과 동일하도록 하는 일치성 손실 함수를 최소화하도록 학습될 수 있다.
- [0017] 상기 얼굴 연령 변환 모델은, 상기 얼굴 연령 변환 모델이 상기 원본 연령을 상기 대상 연령으로 설정하여 변환한 상기 얼굴 변환 영상이 상기 입력받은 얼굴 영상과 동일하도록 하는 고유성 매핑 손실 함수를 최소화하도록 학습될 수 있다.
- [0018] 상기 얼굴 연령 변환 모델은, 상기 얼굴 연령 변환 모델이 변환한 상기 얼굴 변환 영상의 얼굴 고유 특징과 상기 입력받은 얼굴 영상의 얼굴 고유 특징이 유사하도록 하는 고유성 재구성 손실 함수를 최소화하도록 학습될 수 있다.
- [0019] 상기 얼굴 연령 판별 모델은, 상기 제2 인코더에 의해 추출한 특징을 이용하여 상기 제2 연령 회귀자에 의해 추정된 예상 연령에 대한 제2 연령 회귀 손실 함수를 최소화하도록 학습되고, 상기 제2 인코더에 의해 추출한 특징을 이용하여 상기 제2 신원 분류자에 의해 추정된 예상 고유 정보에 대한 제2 고유 정보 손실 함수를 최소화하도록 학습될 수 있다.
- [0020] 상기 얼굴 연령 판별 모델은, 상기 입력받은 얼굴 영상이 실제 영상으로 판단되고 상기 얼굴 연령 변환 모델이 변환한 상기 얼굴 변환 영상이 모조 영상으로 판단되도록 하는 제2 생성 손실 함수를 최소화하도록 학습될 수 있다.
- [0021] 본 실시예의 다른 측면에 의하면, 하나 이상의 프로세서 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 얼굴 영상 연령 변환 장치에 있어서, 상기 프로세서는, 얼굴 영상을 입력받고, 상기 얼굴 영상을 변환하려는 대상 연령을 입력받고, 얼굴 연령 변환 모델을 통해 상기 얼굴 영

상으로부터 특징을 추출하고 상기 특징을 연령 특징 및 얼굴 고유 특징으로 분해하고 상기 얼굴 고유 특징으로부터 상기 대상 연령에 맞도록 변환한 얼굴 변환 영상을 출력하는 것을 특징으로 하는 얼굴 영상 연령 변환 장치를 제공한다.

발명의 효과

[0022] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 연속적 변수인 대상 연령의 특징을 분석하고 영상의 연령 특징과 인물 특징을 특징 공간에서 분리하여 인물의 고유 특징을 보존하면서 대상 연령으로 변환할 수 있는 효과가 있다.

[0023] 여기에서 명시적으로 언급되지 않은 효과라 하더라도, 본 발명의 기술적 특징에 의해 기대되는 이하의 명세서에서 기재된 효과 및 그 잠정적인 효과는 본 발명의 명세서에 기재된 것과 같이 취급된다.

도면의 간단한 설명

- [0024] 도 1은 본 발명의 일 실시예에 따른 얼굴 영상 연령 변환 장치를 예시한 블록도이다.
- 도 2 및 도 3은 본 발명의 일 실시예에 따른 얼굴 영상 연령 변환 장치의 얼굴 연령 변환 모델을 예시한 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 얼굴 영상 연령 변환 장치의 얼굴 연령 판별 모델을 예시한 도면이다.
- 도 5는 본 발명의 일 실시예에 따른 얼굴 영상 연령 변환 장치의 얼굴 연령 변환 모델의 네트워크 구조를 예시한 도면이다.
- 도 6은 본 발명의 일 실시예에 따른 얼굴 영상 연령 변환 장치의 얼굴 연령 판별 모델의 네트워크 구조를 예시한 도면이다.
- 도 7은 본 발명의 다른 실시예에 따른 얼굴 영상 연령 변환 방법을 예시한 흐름도이다.
- 도 8은 본 발명의 실시예들에 따른 시뮬레이션 결과를 예시한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0025] 이하, 본 발명을 설명함에 있어서 관련된 공지기능에 대하여 이 분야의 기술자에게 자명한 사항으로서 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하고, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다.
- [0026] 본 발명은 그룹 단위 연령 변환으로부터 연속적 변수에 대한 연령 변환으로 확장한다. 기존의 연령 변환 모델은 그룹 단위로 카테고리가 정해진 영상들을 이용하여 연령 변환을 적용하였으나 특정 연령대로 변환한 영상에 대해 정확한 연령 추정이 불가능하며, 특정한 연령으로의 변환이 제한적이다.
- [0027] 본 발명은 기존의 그룹 단위가 아닌 연속적 변수인 대상 연령의 특징을 자동으로 분석하고, 영상을 인물 특성과 연령 특성으로 특징 공간에서 분리함으로써 인물 특성을 보존할 수 있다.
- [0028] 도 1은 본 발명의 일 실시예에 따른 얼굴 영상 연령 변환 장치를 예시한 블록도이다.
- [0029] 얼굴 영상 연령 변환 장치(110)는 적어도 하나의 프로세서(120), 컴퓨터 판독 가능한 저장매체(130) 및 통신 버스(170)를 포함한다.
- [0030] 프로세서(120)는 얼굴 영상 연령 변환 장치(110)로 동작하도록 제어할 수 있다. 예컨대, 프로세서(120)는 컴퓨터 판독 가능한 저장 매체(130)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 컴퓨터 실행 가능 명령어는 프로세서(120)에 의해 실행되는 경우 얼굴 영상 연령 변환 장치(110)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.
- [0031] 컴퓨터 판독 가능한 저장 매체(130)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능한 저장 매체(130)에 저장된 프로그램(140)은 프로세서(120)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능한 저장 매체(130)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 일

굴 영상 연령 변환 장치(110)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.

- [0032] 통신 버스(170)는 프로세서(120), 컴퓨터 판독 가능한 저장 매체(140)를 포함하여 얼굴 영상 연령 변환 장치(110)의 다른 다양한 컴포넌트들을 상호 연결한다.
- [0033] 얼굴 영상 연령 변환 장치(110)는 또한 하나 이상의 입출력 장치를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(150) 및 하나 이상의 통신 인터페이스(160)를 포함할 수 있다. 입출력 인터페이스(150) 및 통신 인터페이스(160)는 통신 버스(170)에 연결된다. 입출력 장치(미도시)는 입출력 인터페이스(150)를 통해 얼굴 영상 연령 변환 장치(110)의 다른 컴포넌트들에 연결될 수 있다.
- [0034] 얼굴 영상 연령 변환 장치(110)는 입력된 인물 영상에서 인물의 고유 특징을 보존하고, 해당 연령으로 변환한다. 인물의 고유 특징에 대한 표상(Representation)과 인물의 연령에 대한 특징에 대한 표상을 분리할 필요가 있다. 각각의 특징 정보에 대해 극좌표계와 같은 수직적 정보가 되도록 모델링한다. 인물의 고유 특징과 변환 대상 연령을 네트워크를 통해 최종적으로 해당 연령의 인물 영상을 생성한다. 얼굴 영상 연령 변환 장치(110)는 얼굴 연령 변환 모델과 얼굴 연령 판별 모델을 포함할 수 있다.
- [0035] 본 발명의 얼굴 연령 변환 모델은 연속적인 연령 정보에 맞춰 입력된 얼굴 영상을 변환한다.
- [0036] 도 2 및 도 3은 본 발명의 일 실시예에 따른 얼굴 영상 연령 변환 장치의 얼굴 연령 변환 모델을 예시한 도면이다.
- [0037] 얼굴 연령 변환 모델은 연령 특징을 특징의 놈(Norm)으로 산출하고, 얼굴 고유 특징을 특징에서 특징의 놈으로 나눠서 산출한다.
- [0038] 얼굴 영상 연령 변환 장치는 얼굴 영상을 입력받는 과정에서 얼굴 영상에 대한 인물의 원래 연령을 입력받고, 얼굴 영상에 대한 인물의 고유 정보를 입력받는다. 학습 과정에서 인물의 원래 연령 및 인물의 고유 정보를 적용한다.
- [0039] 얼굴 연령 변환 모델은 원래 연령을 기초로 연령 특징으로부터 예상 연령을 추정하도록 학습되며, 얼굴 연령 변환 모델은 고유 정보를 기초로 얼굴 고유 특징으로부터 예상 고유 정보를 추정하도록 학습된다.
- [0040] 얼굴 연령 변환 모델은 제1 인코더(Encoder) 및 제1 인코더에 연결된 디코더(Decoder)를 포함하는 학습 네트워크 구조이고, 제1 인코더에 연결된 제1 연령 회귀자(Age Regressor) 및 제1 인코더에 연결된 제1 신원 분류자(Identity Classifier)를 포함한다.
- [0041] 얼굴 연령 변환 모델의 손실 함수는 제1 연령 회귀 손실 함수 및 제1 고유 정보 손실 함수를 포함하며, 얼굴 연령 변환 모델은 제1 인코더, 제1 연령 회귀자, 및 제1 신원 분류자를 함께 학습한다.
- [0042] 얼굴 연령 변환 모델은 제1 인코더에 의해 추출한 특징을 이용하여 제1 연령 회귀자에 의해 추정된 예상 연령에 대한 제1 연령 회귀 손실 함수(L_{reg})를 최소화하도록 학습된다.

수학식 1

$$\mathcal{L}_{reg} = \left| f_{age}^G(\text{Enc}^G(X)) - \text{Age}_{org} \right|_2$$

[0043]

[0044] 얼굴 연령 변환 모델은 제1 인코더에 의해 추출한 특징을 이용하여 상기 제1 신원 분류자에 의해 추정된 예상 고유 정보에 대한 제1 고유 정보 손실 함수(L_{id})를 최소화하도록 학습된다.

수학식 2

$$\mathcal{L}_{id} = - \sum \text{id}_{org} f_{id}^G(\text{Enc}^G(X))$$

[0045]

[0046] 얼굴 연령 변환 모델의 손실 함수는 제1 생성 손실 함수, 연령 손실 함수, 일치성 손실 함수, 고유성 매핑 손실 함수, 및 재구성 손실 함수를 포함하며, 얼굴 연령 변환 모델은 제1 인코더 및 디코더를 함께 학습한다.

[0047] 얼굴 연령 변환 모델은 얼굴 연령 변환 모델이 변환한 얼굴 변환 영상이 얼굴 연령 판별 모델에 의해 실제 영상으로 판단되도록 하는 제1 생성 손실 함수(L_{GAN})를 최소화하도록 학습된다.

수학식 3

$$[0048] \quad \mathcal{L}_{GAN} = 1 - D(G(X, Age_{trg}))$$

[0049] 얼굴 연령 변환 모델은 얼굴 연령 변환 모델이 변환한 얼굴 변환 영상이 얼굴 연령 판별 모델에 의해 대상 연령으로 판별되도록 하는 연령 손실 함수(L_{age})를 최소화하도록 학습된다.

수학식 4

$$[0050] \quad \mathcal{L}_{age} = \left| f_{age}^D(G(X, Age_{trg})) - Age_{trg} \right|_2$$

[0051] 얼굴 연령 변환 모델은 얼굴 연령 변환 모델이 변환한 얼굴 변환 영상을 얼굴 연령 변환 모델을 통해 원래 연령으로 변환할 때 입력받은 얼굴 영상과 동일하도록 하는 일치성 손실 함수(L_{cyc})를 최소화하도록 학습된다.

수학식 5

$$[0052] \quad \mathcal{L}_{cyc} = \left| G(G(X, Age_{trg}), Age_{org}) - X \right|_2$$

[0053] 얼굴 연령 변환 모델은 얼굴 연령 변환 모델이 원본 연령을 대상 연령으로 설정하여 변환한 얼굴 변환 영상이 입력받은 얼굴 영상과 동일하도록 하는 고유성 매핑 손실 함수(L_{idt})를 최소화하도록 학습된다.

수학식 6

$$[0054] \quad \mathcal{L}_{idt} = \left| G(X, Age_{org}) - X \right|_2$$

[0055] 얼굴 연령 변환 모델은 얼굴 연령 변환 모델이 변환한 얼굴 변환 영상의 얼굴 고유 특징과 입력받은 얼굴 영상의 얼굴 고유 특징이 유사하도록 하는 고유성 재구성 손실 함수(L_{id_rec})를 최소화하도록 학습된다.

수학식 7

$$[0056] \quad \mathcal{L}_{id_rec} = 1 - CosSim(Enc^G(G(X, Age_{trg})), X)$$

수학식 8

$$\text{CosSim}(A, B) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

[0057]

[0058]

[0059]

[0060]

[0061]

[0062]

도 4는 본 발명의 일 실시예에 따른 얼굴 영상 연령 변환 장치의 얼굴 연령 판별 모델을 예시한 도면이다.

얼굴 연령 변환 모델은 얼굴 연령 판별 모델과 상호 작용하여 학습된다.

얼굴 연령 판별 모델은 제2 인코더 및 제2 인코더에 연결된 판별자 레이어를 포함하는 학습 네트워크 구조이고, 제2 인코더에 연결된 제2 연령 회귀자 및 제2 인코더에 연결된 제2 신원 분류자를 포함한다.

얼굴 연령 판별 모델의 손실 함수는 제2 연령 회귀 손실 함수, 및 제2 고유 정보 손실 함수를 포함하며, 얼굴 연령 판별 모델은 제2 인코더, 제2 연령 회귀자, 및 제2 신원 분류자를 함께 학습한다.

얼굴 연령 판별 모델은 제2 인코더에 의해 추출한 특징을 이용하여 제2 연령 회귀자에 의해 추정된 예상 연령에 대한 제2 연령 회귀 손실 함수(L_{reg})를 최소화하도록 학습된다.

수학식 9

$$\mathcal{L}_{reg} = \left| f_{age}^D(\text{Enc}^D(X)) - \text{Age}_{org} \right|_2$$

[0063]

[0064]

얼굴 연령 판별 모델은 제2 인코더에 의해 추출한 특징을 이용하여 제2 신원 분류자에 의해 추정된 예상 고유 정보에 대한 제2 고유 정보 손실 함수(L_{id})를 최소화하도록 학습된다.

수학식 10

$$\mathcal{L}_{id} = - \sum \text{id}_{org} f_{id}^D(\text{Enc}^D(X))$$

[0065]

[0066]

얼굴 연령 판별 모델의 손실 함수는 제2 생성 손실 함수를 포함하며, 얼굴 연령 판별 모델은 제2 인코더 및 판별자 레이어를 학습한다.

[0067]

얼굴 연령 판별 모델은 입력받은 얼굴 영상이 실제 영상으로 판단되고 얼굴 연령 변환 모델이 변환한 얼굴 변환 영상이 모조 영상으로 판단되도록 하는 제2 생성 손실 함수(L_{GAN})를 최소화하도록 학습된다.

수학식 11

$$\mathcal{L}_{GAN} = D(X) + (1 - D(G(X, \text{Age}_{trg})))$$

[0068]

[0069]

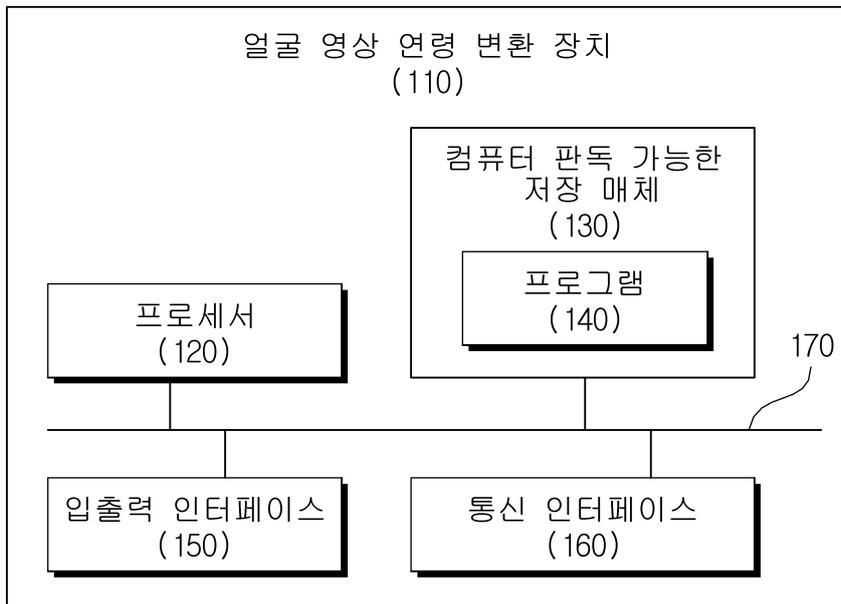
도 5는 본 발명의 일 실시예에 따른 얼굴 영상 연령 변환 장치의 얼굴 연령 변환 모델의 네트워크 구조(제1 인코더, 디코더, 제1 연령 회귀자, 제1 신원 분류자)를 예시한 도면이고, 도 6은 본 발명의 일 실시예에 따른 얼굴 영상 연령 변환 장치의 얼굴 연령 판별 모델의 네트워크 구조(제2 인코더, 제2 연령 회귀자, 제2 신원 분류

자, 판별자 레이어)를 예시한 도면이다.

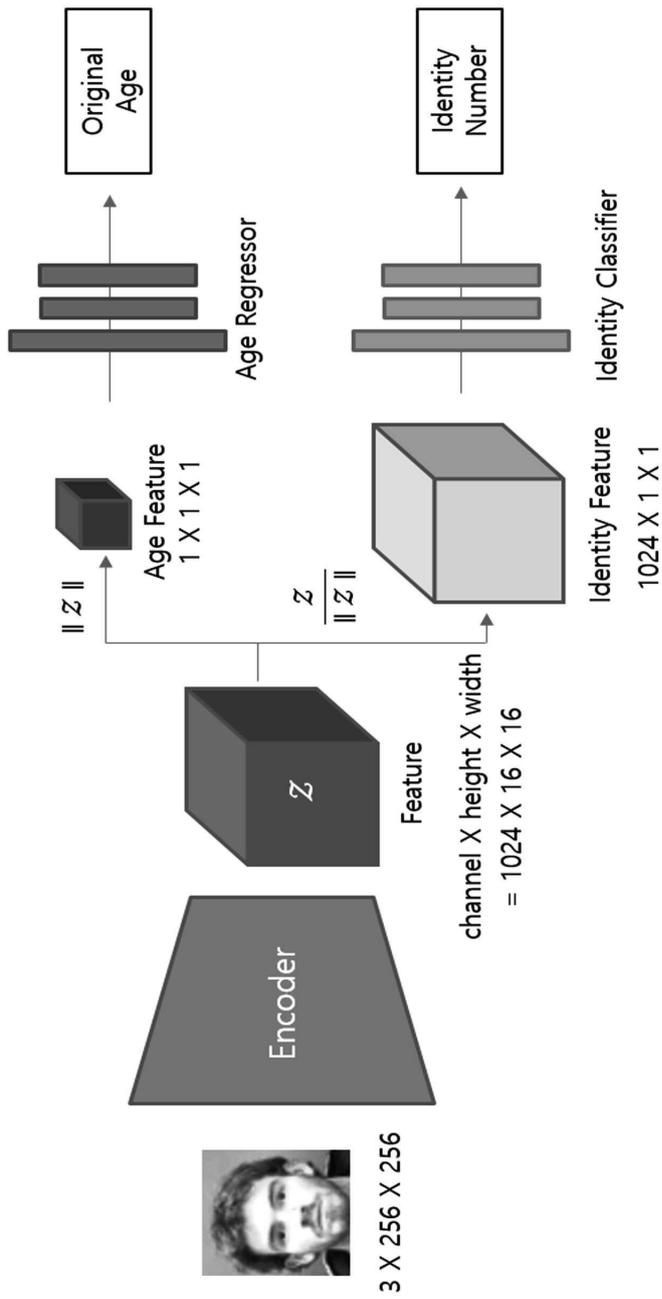
- [0070] 네트워크 모델은 다수의 레이어가 네트워크로 연결되며 컨볼루션 레이어를 포함할 수 있다. 레이어는 파라미터를 포함할 수 있고, 레이어의 파라미터는 학습가능한 필터 집합을 포함한다. 파라미터는 노드 간의 가중치 및/또는 바이어스를 포함한다. 네트워크 모델은 손실 함수를 최소화하는 방향으로 네트워크 가중치를 갱신한다.
- [0071] 네트워크 모델은 레지듀얼(Residual) 블록을 포함할 수 있다. 레지듀얼 블록은 네트워크 구조의 출력에 다시 입력을 더해서 다음 레이어로 넘기며, 레이어의 입력을 레이어의 출력에 바로 연결하는 스킵 구조를 가질 수 있다.
- [0072] 네트워크 모델은 풀리 커넥티드 레이어(Fully Connected Layer)를 포함할 수 있다.
- [0073] 도 7은 본 발명의 다른 실시예에 따른 얼굴 영상 연령 변환 방법을 예시한 흐름도이다. 얼굴 영상 연령 변환 방법은 컴퓨팅 디바이스에 의하여 수행될 수 있으며, 얼굴 영상 연령 변환 장치와 동일한 방식으로 동작한다.
- [0074] 단계 S210에서 프로세서는 얼굴 영상을 입력받는다.
- [0075] 단계 S220에서 프로세서는 얼굴 영상을 변환하려는 대상 연령을 입력받는다.
- [0076] 단계 S230에서 프로세서는 얼굴 연령 변환 모델을 통해 얼굴 영상으로부터 특징을 추출하고 특징을 연령 특징 및 얼굴 고유 특징으로 분해하고 얼굴 고유 특징으로부터 대상 연령에 맞도록 변환한 얼굴 변환 영상을 출력한다.
- [0077] 도 8은 본 발명의 실시예들에 따른 시뮬레이션 결과를 예시한 도면이다.
- [0078] 도 8에 도시된 바와 같이, 대상의 인물 고유 특징이 잘 보존되고, 인접한 나이 간의 차이가 과장되거나 부자연스럽지 않음을 확인할 수 있다.
- [0079] 얼굴 영상 연령 변환 장치는 하드웨어, 펌웨어, 소프트웨어 또는 이들의 조합에 의해 로직회로 내에서 구현될 수 있고, 범용 또는 특정 목적 컴퓨터를 이용하여 구현될 수도 있다. 장치는 고정배선형(Hardwired) 기기, 필드 프로그램 가능한 게이트 어레이(Field Programmable Gate Array, FPGA), 주문형 반도체(Application Specific Integrated Circuit, ASIC) 등을 이용하여 구현될 수 있다. 또한, 장치는 하나 이상의 프로세서 및 컨트롤러를 포함한 시스템온칩(System on Chip, SoC)으로 구현될 수 있다.
- [0080] 얼굴 영상 연령 변환 장치는 하드웨어적 요소가 마련된 컴퓨팅 디바이스 또는 서버에 소프트웨어, 하드웨어, 또는 이들의 조합하는 형태로 탑재될 수 있다. 컴퓨팅 디바이스 또는 서버는 각종 기기 또는 유무선 통신망과 통신을 수행하기 위한 통신 모듈 등의 통신장치, 프로그램을 실행하기 위한 데이터를 저장하는 메모리, 프로그램을 실행하여 연산 및 명령하기 위한 마이크로프로세서 등을 전부 또는 일부 포함한 다양한 장치를 의미할 수 있다.
- [0081] 도 7에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 7에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이다.
- [0082] 본 실시예들에 따른 동작은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능한 매체에 기록될 수 있다. 컴퓨터 판독 가능한 매체는 실행을 위해 프로세서에 명령어를 제공하는 데 참여한 임의의 매체를 나타낸다. 컴퓨터 판독 가능한 매체는 프로그램 명령, 데이터 파일, 데이터 구조 또는 이들의 조합을 포함할 수 있다. 예를 들면, 자기 매체, 광기록 매체, 메모리 등이 있을 수 있다. 컴퓨터 프로그램은 네트워크로 연결된 컴퓨터 시스템 상에 분산되어 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다. 본 실시예를 구현하기 위한 기능적인(Functional) 프로그램, 코드, 및 코드 세그먼트들은 본 실시예가 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있을 것이다.
- [0083] 본 실시예들은 본 실시예의 기술 사상을 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

도면

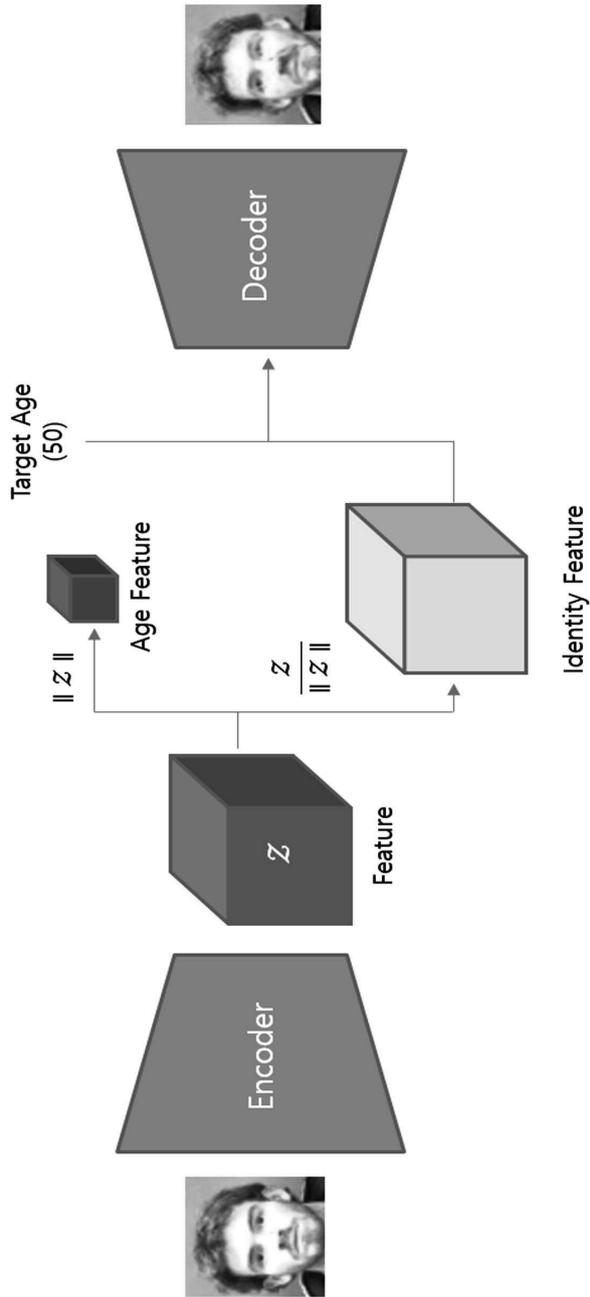
도면1



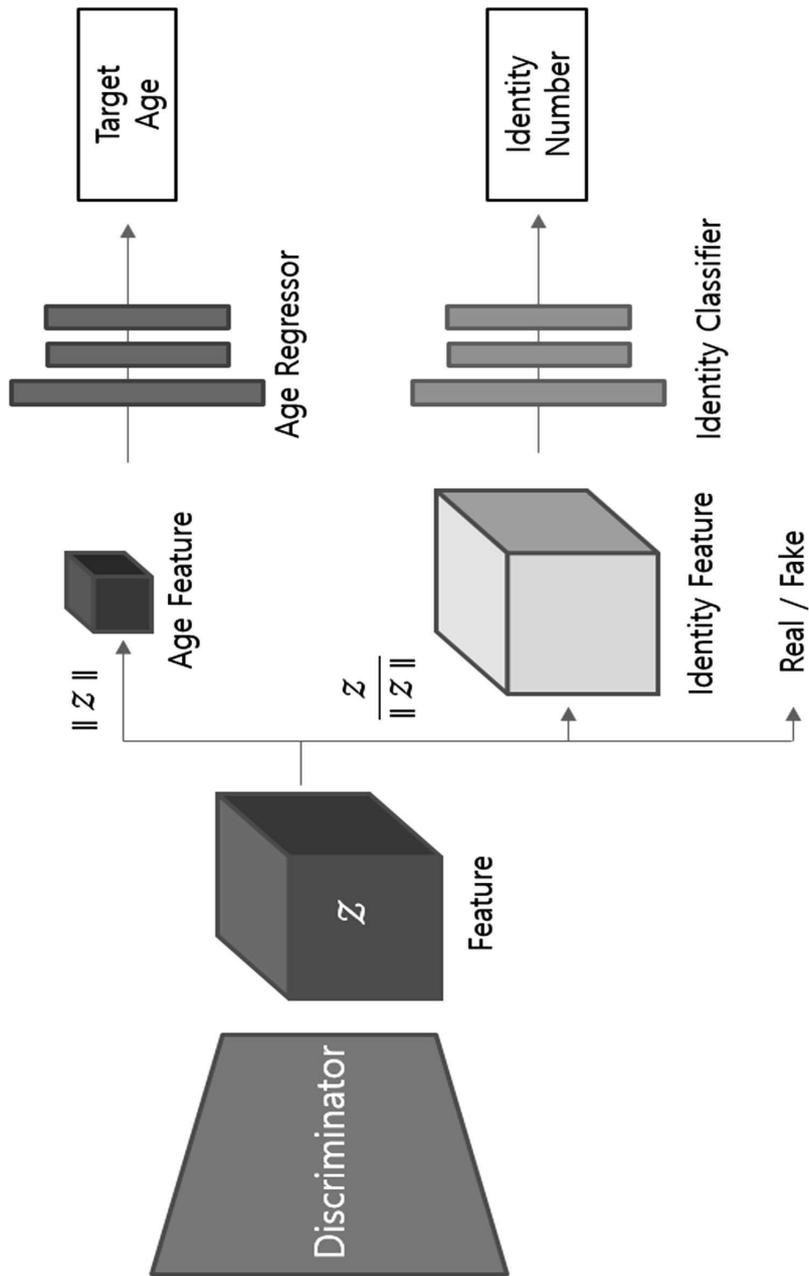
도면2



도면3



도면4



도면5

| Layer | (K, S, P) | Resample | Activation | Normalization | Output Shape |
|----------------|-----------|----------|------------|---------------|-----------------|
| Data x | - | - | - | - | 3 X 256 X 256 |
| Conv2D | 7, 1, 3 | - | ReLU | BatchNorm2D | 64 X 256 X 256 |
| Conv2D | 4, 2, 1 | - | ReLU | BatchNorm2D | 128 X 128 X 128 |
| Conv2D | 4, 2, 1 | - | ReLU | BatchNorm2D | 256 X 64 X 64 |
| Conv2D | 4, 2, 1 | - | ReLU | BatchNorm2D | 512 X 32 X 32 |
| Conv2D | 4, 2, 1 | - | ReLU | BatchNorm2D | 1024 X 16 X 16 |
| Residual Block | 3, 1, 1 | - | ReLU | BatchNorm2D | 1024 X 16 X 16 |

(a)

| Layer | (K, S, P) | Resample | Activation | Normalization | Output Shape |
|-----------------------|-----------|----------|------------|----------------|------------------|
| Data x + Target Age c | - | - | - | - | (1024+1) X 1 X 1 |
| ConvTranspose2D | 4, 2, 1 | - | ReLU | InstanceNorm2D | 1024 X 16 X 16 |
| ConvTranspose2D | 4, 2, 1 | - | ReLU | InstanceNorm2D | 512 X 32 X 32 |
| ConvTranspose2D | 4, 2, 1 | - | ReLU | InstanceNorm2D | 256 X 64 X 64 |
| ConvTranspose2D | 4, 2, 1 | - | ReLU | InstanceNorm2D | 128 X 128 X 128 |
| ConvTranspose2D | 4, 2, 1 | - | ReLU | InstanceNorm2D | 64 X 256 X 256 |
| Conv2D | 7, 1, 3 | - | Tanh | - | 3 X 256 X 256 |

(b)

| Layer | Activation | Normalization | Output Shape |
|-----------------------|------------|---------------|--------------|
| Data x | - | - | 1 |
| Fully Connected Layer | - | - | 1 |

(c)

| Layer | Activation | Normalization | Output Shape |
|-----------------------|------------|---------------|--------------|
| Data x | - | - | 1024 |
| Fully Connected Layer | - | - | 10000 |

(d)

도면6

| Layer | (K, S, P) | Resample | Activation | Normalization | Output Shape |
|----------------|-----------|----------|------------|---------------|-----------------|
| Data x | - | - | - | - | 3 X 256 X 256 |
| Conv2D | 7, 1, 3 | - | LeakyReLU | BatchNorm2D | 64 X 256 X 256 |
| Conv2D | 4, 2, 1 | - | LeakyReLU | BatchNorm2D | 128 X 128 X 128 |
| Conv2D | 4, 2, 1 | - | LeakyReLU | BatchNorm2D | 256 X 64 X 64 |
| Conv2D | 4, 2, 1 | - | LeakyReLU | BatchNorm2D | 512 X 32 X 32 |
| Conv2D | 4, 2, 1 | - | LeakyReLU | BatchNorm2D | 1024 X 16 X 16 |
| Residual Block | 3, 1, 1 | - | LeakyReLU | BatchNorm2D | 1024 X 16 X 16 |
| Residual Block | 3, 1, 1 | - | LeakyReLU | BatchNorm2D | 1024 X 16 X 16 |
| Residual Block | 3, 1, 1 | - | LeakyReLU | BatchNorm2D | 1024 X 16 X 16 |
| Residual Block | 3, 1, 1 | - | LeakyReLU | BatchNorm2D | 1024 X 16 X 16 |
| Residual Block | 3, 1, 1 | - | LeakyReLU | BatchNorm2D | 1024 X 16 X 16 |
| Residual Block | 3, 1, 1 | - | LeakyReLU | BatchNorm2D | 1024 X 16 X 16 |

(a)

| Layer | Activation | Normalization | Output Shape |
|-----------------------|------------|---------------|--------------|
| Data x | - | - | 1 |
| Fully Connected Layer | - | - | 1 |

(b)

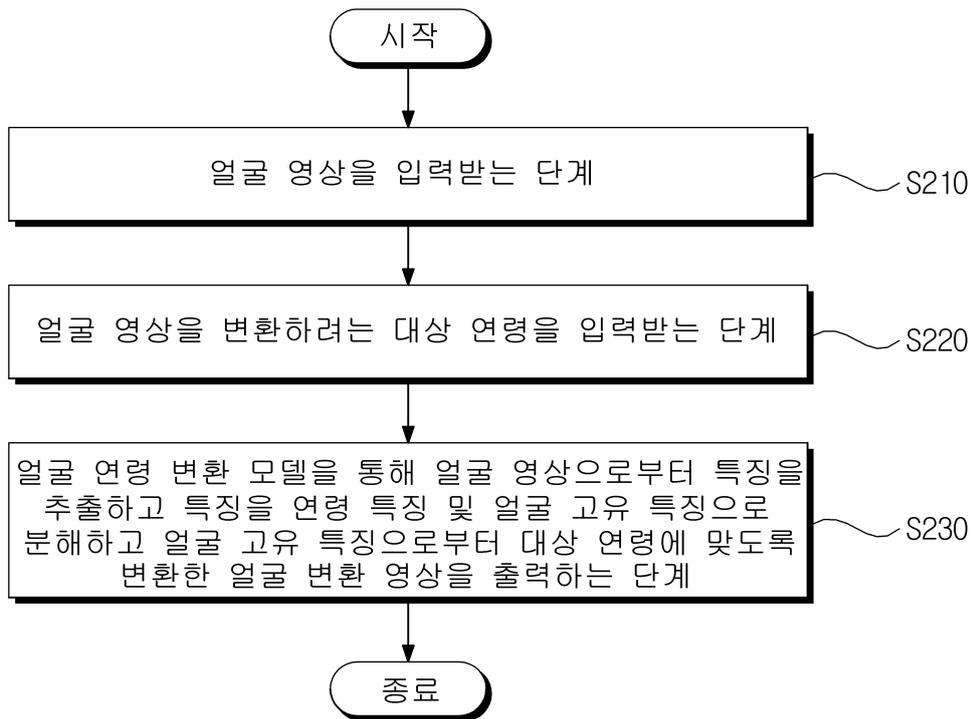
| Layer | Activation | Normalization | Output Shape |
|-----------------------|------------|---------------|--------------|
| Data x | - | - | 1024 |
| Fully Connected Layer | - | - | 10000 |

(c)

| Layer | (K, S, P) | Activation | Normalization | Output Shape |
|--------|-----------|------------|---------------|----------------|
| Data x | - | - | - | 1024 X 16 X 16 |
| Conv2D | 3, 1, 1 | - | - | 1 X 16 X 16 |

(d)

도면7



도면8

