

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국



(10) 국제공개번호

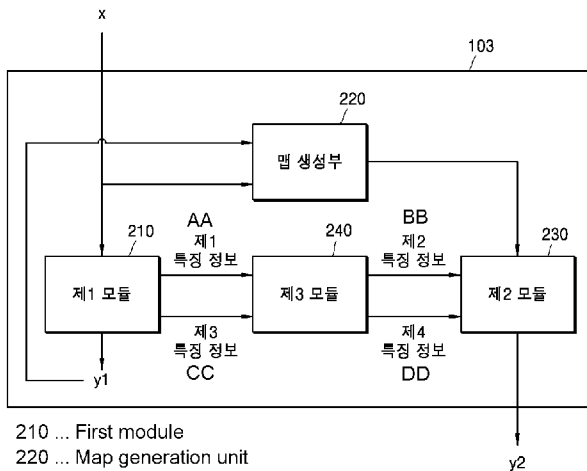
WO 2023/219276 A1

2023년 11월 16일 (16.11.2023) WIPO | PCT

- (51) 국제특허분류: G06T 3/40 (2006.01) G06T 5/00 (2006.01)
G06T 7/11 (2017.01) G06T 5/20 (2006.01)
G06N 3/08 (2006.01)
- (21) 국제출원번호: PCT/KR2023/004669
- (22) 국제출원일: 2023년 4월 6일 (06.04.2023)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보: 10-2022-0056889 2022년 5월 9일 (09.05.2022) KR
10-2022-0132739 2022년 10월 14일 (14.10.2022) KR
- (71) 출원인: 삼성전자 주식회사 (SAMSUNG ELECTRONICS CO., LTD.) [KR/KR]; 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR).
- (72) 발명자: 박재연 (PARK, Jaeyeon); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). 안일준 (AHN, Iljun); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). 이태미 (LEE, Tammy); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). 강수민 (KANG, Soomin); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). 송영찬 (SONG, Youngchan); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR). 신하늘 (SHIN, Hanul); 16677 경기도 수원시 영통구 삼성로 129, Gyeonggi-do (KR).
- (74) 대리인: 리앤목 특허법인 (Y.P.LEE, MOCK & PARTNERS); 06292 서울특별시 강남구 언주로30길 13 대림아크로텔 12층, Seoul (KR).
- (81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ,

(54) Title: IMAGE PROCESSING APPARATUS AND OPERATION METHOD THEREOF

(54) 발명의 명칭: 영상 처리 장치 및 그 동작 방법



- 210 ... First module
- 220 ... Map generation unit
- 230 ... Second module
- 240 ... Third module
- AA ... First feature information
- BB ... Second feature information
- CC ... Third feature information
- DD ... Fourth feature information

(57) Abstract: An image processing apparatus for processing an image by using one or more neural networks according to an embodiment may comprise a memory for storing one or more instructions, and at least one processor for executing one or more instructions stored in the memory, thereby: acquiring first feature information of a first image; generating an intermediate output image for the first image by performing first image processing on the first feature information; generating an attention map on the basis of the first image and the intermediate output image; acquiring second feature information by performing second image processing on the first feature information; acquiring fourth feature information by performing third image processing on third feature information extracted during the first image processing; and generating a second image having a higher image quality compared to the first image on the basis of the attention map, the second feature information, and the fourth feature information.

(57) 요약서: 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치는, 하나 이상의 인스트럭션들을 저장하는 메모리 및 메모리에 저장된 하나 이상의 인스트럭션들을 실행함으로써, 제1 영상의 제1 특징 정보를 획득하고, 제1 특징 정보에 제1 영상 처리를 수행함으로써, 제1 영상에 대한 중간 출력 영상을 생성하고, 제1 영상 및 중간 출력 영상에 기초하여, 어텐션 맵(attention map)을 생성하고, 제1 특징 정보에 제2 영상 처리를 수행함으로써, 제2 특징 정보를 획득하고, 제1 영상 처리 중에 추출되는 제3 특징 정보에 제3 영상 처리를 수행함으로써, 제4 특징 정보를 획득하고, 어텐션 맵, 제2 특징 정보 및 제4 특징 정보에 기초하여, 제1 영상보다 고화질인 제2 영상을 생성하는 적어도 하나의 프로세서를 포함할 수 있다.

EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

- 국제조사보고서와 함께 (조약 제21조(3))

명세서

발명의 명칭: 영상 처리 장치 및 그 동작 방법

기술분야

- [1] 본 개시는 영상을 처리하는 영상 처리 장치에 관한 것으로, 더 구체적으로는 뉴럴 네트워크를 이용하여, 영상을 처리하는 영상 처리 장치 및 영상 처리 장치의 동작 방법에 관한 것이다.

배경기술

- [2] 컴퓨터 기술의 발달과 함께 데이터 트래픽이 지수함수 형태로 증가하면서 인공지능 미래 혁신을 주도하는 중요한 트렌드로 자리잡았다. 인공지능은 사람의 사고방식을 모방하는 방식이며, 이와 같은 인공지능은 다양한 산업에 응용 가능하다. 인공지능 기술들의 예시들로는 패턴 인식, 기계 학습, 전문가 시스템, 뉴럴 네트워크, 자연어 처리 등이 있다.
- [3] 뉴럴 네트워크는 인간의 생물학적 신경 세포의 특성을 수학적 표현에 의해 모델링한 것으로, 인간이 가지고 있는 학습이라는 능력을 모방한 알고리즘을 이용한다. 이 알고리즘을 통하여, 뉴럴 네트워크는 입력 데이터와 출력 데이터 사이의 사상(mapping)을 생성할 수 있고, 이러한 사상을 생성하는 능력은 뉴럴 네트워크의 학습 능력이라고 표현될 수 있다. 또한, 뉴럴 네트워크는 학습된 결과에 기초하여, 학습에 이용되지 않았던 입력 데이터에 대하여, 올바른 출력 데이터를 생성할 수 있는 일반화 능력을 가진다.
- [4] 뉴럴 네트워크는 영상 처리에 이용될 수 있으며, 특히, 심층 신경망(DNN: Deep Neural Network)을 이용하여, 영상의 노이즈 또는 아티팩트를 제거하거나, 영상의 화질 또는 해상도를 증가시키는 영상 처리를 수행할 수 있다.

발명의 상세한 설명

과제 해결 수단

- [5] 일 실시예에 따른 영상 처리 장치는 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리할 수 있다.
- [6] 일 실시예에 따른 영상 처리 장치는 하나 이상의 인스트럭션들을 저장하는 메모리 및 상기 하나 이상의 인스트럭션들을 실행하는 적어도 하나의 프로세서를 포함할 수 있다.
- [7] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 제1 영상의 제1 특징 정보를 획득할 수 있다.
- [8] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 특징 정보에 제1 처리를 수행함으로써, 출력 영상을 생성할 수 있다.

- [9] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 영상 및 상기 출력 영상에 기초하여, 어텐션 맵(attention map)을 생성할 수 있다.
- [10] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 특징 정보에 제2 처리를 수행함으로써, 제2 특징 정보를 획득할 수 있다.
- [11] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 처리 중에 제3 특징 정보를 획득할 수 있다.
- [12] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제3 특징 정보에 제3 영상 처리를 수행하여, 제4 특징 정보를 획득할 수 있다.
- [13] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 어텐션 맵, 상기 제2 특징 정보 및 상기 제4 특징 정보에 기초하여, 상기 제1 영상의 제1 화질보다 고화질인 제2 화질을 가지는 제2 영상을 생성할 수 있다.
- [14] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 영상 및 상기 출력 영상에 대응하는 레지듀얼 맵(residual map)을 획득할 수 있다.
- [15] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 레지듀얼 맵의 로컬 맵(local map)과 글로벌(global) 분산 값들을 획득할 수 있다.
- [16] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 로컬 맵과 상기 글로벌 분산 값들에 기초하여, 상기 어텐션 맵을 생성할 수 있다.
- [17] 상기 로컬 맵은 상기 레지듀얼 맵의 로컬 분산 값들을 포함하고, 상기 로컬 분산 값들은 제1 크기를 가지는 윈도우를 이용하여 획득될 수 있다.
- [18] 상기 글로벌 분산 값들은 상기 제1 크기보다 큰 제2 크기를 가지는 패치를 이용하여 획득될 수 있다.
- [19] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 어텐션 맵에 제4 처리를 수행함으로써, 제5 특징 정보를 획득할 수 있다.
- [20] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제2 특징 정보와 상기 제5 특징 정보에 기초하여, 제6 특징 정보를 획득할 수 있다.
- [21] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제6 특징 정보 및 상기 제4 특징 정보에 기초하여, 상기 제2 영상을 생성할 수 있다.

- [22] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 어텐션 맵에 제5 처리를 수행함으로써, 제7 특징 정보를 획득할 수 있다.
- [23] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제2 특징 정보, 제4 특징 정보 및 제7 특징 정보에 기초하여, 상기 제2 영상을 생성할 수 있다.
- [24] 상기 하나 이상의 뉴럴 네트워크들은 하나 이상의 컨볼루션 뉴럴 네트워크들을 포함할 수 있다.
- [25] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 하나 이상의 컨볼루션 뉴럴 네트워크들을 이용하여, 상기 제1 영상의 제1 특징 정보를 획득할 수 있다.
- [26] 상기 제1 처리는 상기 제1 특징 정보의 업 샘플링을 포함하고, 상기 제1 영상 및 상기 출력 영상은 크기가 동일할 수 있다.
- [27] 상기 제2 처리는 하나 이상의 완전 연결 레이어들(fully connected layers)을 포함하는 MLP 모듈에 의해 수행될 수 있다.
- [28] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 하나의 완전 연결 레이어에 입력되는 입력 데이터와 상기 완전 연결 레이어에 포함되는 가중치 행렬과의 곱셈 연산 및 상기 완전 연결 레이어에 포함되는 바이어스와의 덧셈 연산을 수행할 수 있다.
- [29] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제3 특징 정보와 기 학습된 가중치에 기초하는 컨볼루션 연산을 수행함으로써, 상기 제4 특징 정보를 획득할 수 있다.
- [30] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제2 특징 정보에 제1 연산을 수행하여, 제8 특징 정보를 획득할 수 있다.
- [31] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제8 특징 정보 및 상기 제4 특징 정보에 제2 연산을 수행함으로써, 상기 제2 영상을 생성할 수 있다.
- [32] 상기 제1 연산은 AdaIN 연산을 포함하고, 상기 제2 연산은 SFT 연산을 포함할 수 있다.
- [33] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 영상의 분류 정보를 획득할 수 있다.
- [34] 상기 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 특징 정보 및 상기 분류 정보에 상기 제1 처리를 수행함으로써, 출력 영상을 생성할 수 있다.
- [35] 상기 제2 화질의 상기 제2 영상은, 상기 제1 화질의 상기 제1 영상과 비교할 때, 적은 아티팩트 또는 적은 노이즈를 가질 수 있다.

- [36] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 제1 영상의 제1 특징 정보를 획득하는 단계를 포함할 수 있다.
- [37] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 상기 제1 특징 정보에 제1 처리를 수행함으로써, 출력 영상을 생성하는 단계를 포함할 수 있다.
- [38] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 상기 제1 영상 및 상기 출력 영상에 기초하여, 어텐션 맵을 생성하는 단계를 포함할 수 있다.
- [39] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 상기 제1 특징 정보에 제2 영상 처리를 수행함으로써, 제2 특징 정보를 획득하는 단계를 포함할 수 있다.
- [40] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 상기 제1 처리 중에 제3 특징 정보를 획득하는 단계를 포함할 수 있다.
- [41] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 상기 제3 특징 정보에 제3 영상 처리를 수행함으로써, 제4 특징 정보를 획득하는 단계를 포함할 수 있다.
- [42] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 상기 어텐션 맵, 상기 제2 특징 정보 및 상기 제4 특징 정보에 기초하여, 상기 제1 영상의 제1 화질보다 고화질인 제2 화질을 가지는 제2 영상을 생성하는 단계를 포함할 수 있다.
- [43] 본 개시의 일 측면에 따르면, 제1 영상의 제1 특징 정보를 획득하는 단계; 상기 제1 특징 정보에 제1 처리를 수행함으로써, 출력 영상을 생성하는 단계; 상기 제1 영상 및 상기 출력 영상에 기초하여, 어텐션 맵을 생성하는 단계; 상기 제1 특징 정보에 제2 처리를 수행함으로써, 제2 특징 정보를 획득하는 단계; 상기 제1 처리 중에 제3 특징 정보를 획득하는 단계; 상기 제3 특징 정보에 제3 처리를 수행함으로써, 제4 특징 정보를 획득하는 단계; 및 상기 어텐션 맵, 상기 제2 특징 정보 및 상기 제4 특징 정보에 기초하여, 상기 제1 영상의 제1 화질보다 고화질인 제2 화질을 가지는 제2 영상을 생성하는 단계를 포함하는, 영상 처리 장치의 동작 방법을 컴퓨터에 의해 수행하기 위한 프로그램을 저장하는 컴퓨터로 읽을 수 있는 기록 매체를 제공할 수 있다.

도면의 간단한 설명

- [44] 도 1은 일 실시예에 따른 영상 처리 장치가 영상 처리 네트워크를 이용하여, 영상을 처리하는 동작을 나타내는 도면이다.
- [45] 도 2는 일 실시예에 따른 영상 처리 네트워크를 나타내는 도면이다.
- [46] 도 3은 일 실시예에 따른 제1 모듈을 나타내는 도면이다.

- [47] 도 4는 일 실시예에 따른 제1 특징 추출부를 나타내는 도면이다.
- [48] 도 5는 일 실시예에 따른 중간 영상 생성부를 나타내는 도면이다.
- [49] 도 6은 일 실시예에 따른 맵 생성부를 나타내는 도면이다.
- [50] 도 7은 일 실시예에 따른 제3 모듈이 제1 영상의 제1 특징 정보를 처리하는 방법을 나타내는 도면이다.
- [51] 도 8은 일 실시예에 따른 제3 모듈이 제3 특징 정보를 처리하는 방법을 나타내는 도면이다.
- [52] 도 9는 일 실시예에 따른 제2 모듈을 나타내는 도면이다.
- [53] 도 10은 일 실시예에 따른 Channel-Split SFT 연산부를 나타내는 도면이다.
- [54] 도 11은 일 실시예에 따른 영상 처리 네트워크를 나타내는 도면이다.
- [55] 도 12는 일 실시예에 따른 제3 모듈을 나타내는 도면이다.
- [56] 도 13은 일 실시예에 따른 제2 모듈을 나타내는 도면이다.
- [57] 도 14는 일 실시예에 따른 영상 처리 네트워크를 나타내는 도면이다.
- [58] 도 15는 일 실시예에 따른 영상 처리 장치의 동작 방법을 나타내는 흐름도이다.
- [59] 도 16은 일 실시예에 따른 영상 처리 장치의 구성을 나타내는 블록도이다.

발명의 실시를 위한 형태

- [60] 본 개시에서, "a, b 또는 c 중 적어도 하나" 표현은 "a", "b", "c", "a 및 b", "a 및 c", "b 및 c", "a, b 및 c 모두", 혹은 그 변형들을 지칭할 수 있다.
- [61] 본 명세서에서 사용되는 용어에 대해 간략히 설명하고, 본 발명에 대해 구체적으로 설명하기로 한다.
- [62] 본 발명에서 사용되는 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나, 이는 당 분야에 종사하는 기술자의 의도 또는 관례, 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 발명의 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서 본 발명에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 발명의 전반에 걸친 내용을 토대로 정의되어야 한다.
- [63] 명세서 전체에서 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있음을 의미한다. 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.
- [64] 아래에서는 첨부한 도면을 참고하여 실시예들에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해

서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

- [65] 도 1은 일 실시예에 따른 영상 처리 장치가 영상 처리 네트워크를 이용하여, 영상을 처리하는 동작을 나타내는 도면이다.
- [66] 도 1을 참조하면, 일 실시예에 따른 영상 처리 네트워크(103)는 제1 영상(101)을 입력 받아, 제1 영상(101)을 처리함으로써, 제2 영상(102)을 생성할 수 있다. 일 실시예에 따른, 제1 영상(101)은 노이즈 또는 아티팩트를 포함하는 영상일 수 있으며, 또는, 저해상도 영상, 또는 저화질 영상일 수 있다. 일 실시예에 따른 영상 처리 장치(100)는 영상 처리 네트워크(103)를 이용하여, 제1 영상(101)의 세밀한 가장자리(edge)와 텍스처를 유지하면서 노이즈를 제거하는 디노이징(denoising)을 수행함으로써, 제2 영상(102)을 생성할 수 있다. 제2 영상(102)은 제1 영상(101)보다 화질이 개선된 영상일 수 있다. 예를 들어, 제2 영상(102)은 제1 영상(101)보다 노이즈 및 아티팩트들을 덜 가질 수 있다. 또한, 제2 영상(102)은 제1 영상(101)보다 고해상도 영상일 수 있다. 다만, 이에 한정되지 않는다. 이와 같이, 다른 실시예에 따르면, 제2 영상(102)은 제1 영상(101)에 비해 개선된 하나 이상의 특성들 또는 특징들을 가질 수 있다.
- [67] 일 실시예에 따른 영상 처리 네트워크(103)는 하나 이상의 뉴럴 네트워크들을 포함할 수 있다. 예를 들어, 영상 처리 네트워크(103)는 맵 생성부, 제1 모듈, 제2 모듈, 및 제3 모듈을 포함할 수 있다. 다만, 이에 한정되지 않는다.
- [68] 이하, 도면을 참조하여 일 실시예에 따른 영상 처리 네트워크를 자세히 설명하기로 한다.
- [69] 도 2는 일 실시예에 따른 영상 처리 네트워크를 나타내는 도면이다.
- [70] 도 2를 참조하면, 일 실시예에 따른 영상 처리 네트워크(103)는 제1 모듈(210), 맵 생성부(220), 제2 모듈(230) 및 제3 모듈(240)을 포함할 수 있다.
- [71] 일 실시예에 따른 제1 모듈(210)로 제1 영상(x)이 입력될 수 있다. 제1 모듈(210)은 제1 영상(x)의 제1 특징 정보를 획득할 수 있다.
- [72] 제1 모듈(210)에서 획득된 제1 영상(x)의 제1 특징 정보는 제3 모듈(240)로 입력될 수 있다.
- [73] 또한, 제1 모듈(210)은 제1 영상(x)의 제1 특징 정보에 기초하여, 중간 출력 영상(y1)을 생성할 수 있다. 예를 들어, 제1 모듈(210)은 제1 영상(x)의 제1 특징 정보에 제1 영상 처리를 수행함으로써, 중간 출력 영상(y1)을 생성할 수 있다. 중간 출력 영상(y1)은 제1 영상(x)에서 아티팩트가 제거된 영상일 수 있다. 다만, 이에 한정되지 않는다. 이와 같이, 다른 실시예에 따르면, 제1 모듈(210)은 제1 영상(x)의 특성을 개선하기 위해 제1 영상(x)을 처리하고, 처리된 영상을 중간 출력 영상(y1)으로 출력할 수 있다.
- [74] 또한, 제1 모듈(210)은 제1 영상 처리의 중간에서 추출되는 제3 특징 정보를 획득하고, 제3 특징 정보는 제3 모듈(230)로 입력될 수 있다.

- [75] 이하, 도면을 참조하여 일 실시예에 따른 맵 생성부(220) 및 제1 내지 제3 모듈들(210, 230, 240) 각각의 세부적인 구성을 자세히 설명하기로 한다.
- [76] 도 3은 일 실시예에 따른 제1 모듈을 나타내는 도면이다.
- [77] 도 3을 참조하면, 제1 모듈(210)은 제1 특징 추출부(310) 및 중간 영상 생성부(320)를 포함할 수 있다.
- [78] 일 실시예에 따른 제1 특징 추출부(310)는 특징 추출 네트워크를 포함할 수 있으며, 특징 추출 네트워크는 하나 이상의 컨볼루션 레이어들을 포함할 수 있다. 제1 영상(x)은 제1 특징 추출부(310)로 입력될 수 있으며, 제1 특징 추출부(310)는 특징 추출 네트워크를 이용하여, 제1 영상의 제1 특징 정보(F_{latent})를 획득할 수 있다.
- [79] 또는, 일 실시예에 따른 제1 특징 추출부(310)는 특징 추출 네트워크를 이용하여, 제1 영상의 분류 정보를 획득할 수 있다. 이때, 제1 영상의 분류 정보는 제1 영상(x)이 기 설정된 복수의 카테고리들 각각에 포함될 확률을 나타내는 데이터를 포함할 수 있다.
- [80] 제1 특징 추출부(310)의 세부적인 구성에 대해서는 도 4를 참조하여 자세히 설명하기로 한다.
- [81] 도 4는 일 실시예에 따른 제1 특징 추출부를 나타내는 도면이다.
- [82] 도 4를 참조하면, 일 실시예에 따른 제1 특징 추출부(310)는 제1 영상(x)과 제1 컨볼루션 레이어(410)에 포함된 제1 가중치에 기초하는 컨볼루션 연산을 수행함으로써, 제1 특징 데이터를 획득할 수 있다. 예를 들어, 제1 특징 추출부(310)는 제1 영상(x)과 제1 가중치의 컨볼루션 연산을 수행하여, 제1 특징 데이터를 획득할 수 있다. 제1 특징 추출부(310)는 제1 특징 데이터와 제2 컨볼루션 레이어(420)에 포함된 제2 가중치에 기초하는 컨볼루션 연산을 수행함으로써, 제2 특징 데이터를 획득할 수 있다. 예를 들어, 제1 특징 추출부(310)는 제1 특징 데이터와 제2 가중치의 컨볼루션 연산을 수행하여, 제2 특징 데이터를 획득할 수 있다.
- [83] 또한, 제1 특징 추출부(310)는 제1 영상(x)과 제3 컨볼루션 레이어(430)에 포함된 제3 가중치에 기초하여 컨볼루션 연산을 수행함으로써, 제3 특징 데이터를 획득할 수 있다. 예를 들어, 제1 특징 추출부(310)는 제1 영상(x)과 제3 가중치의 컨볼루션 연산을 수행하여, 제3 특징 데이터를 획득할 수 있다.
- [84] 또한, 제1 특징 추출부(310)는 합산 레이어(440)에서 제2 특징 데이터와 제3 특징 데이터의 요소별 덧셈 연산을 수행할 수 있다.
- [85] 일 실시예에 따른 제1 특징 추출부(310)는 제1 컨볼루션 레이어(410), 제2 컨볼루션 레이어(420), 제3 컨볼루션 레이어(430) 및 합산 레이어(440)를 포함하는 모듈(401)을 포함할 수 있다. 또한, 일 실시예에 따르면, 제1 특징 추출부(310)는 모듈(401)이 직렬로 반복적으로 배열되는 구조를 포함할 수 있다.
- [86] 다만, 이에 한정되지 않으며, 제1 특징 추출부(310)는 제1 영상(x)에 대한 특징 정보를 추출할 수 있는 인코더(encoder)의 역할을 하는 다양한 구조를 포함할 수 있다. 예를 들어, 제1 특징 추출부(310)는 복수의 제1 컨볼루션 레이어들(410), 복

수의 제2 컨볼루션 레이어들(420), 복수의 제3 컨볼루션 레이어들(430), 및 복수의 합산 레이어들(440)을 포함하는 구조를 포함할 수 있다.

- [87] 이에 따라, 제1 특징 추출부(310)는 제1 영상의 제1 특징 정보(F_{latent})를 획득할 수 있다. 이때, 제1 특징 정보(F_{latent})는 벡터 형태의 데이터일 수 있다.
- [88] 또는, 일 실시예에 따른 제1 특징 추출부(310)는 제1 영상의 분류 정보를 획득할 수 있다. 이때, 제1 영상의 분류 정보는 제1 영상(x)이 복수의 카테고리들 각각에 포함될 확률을 나타내는 데이터를 포함할 수 있다. 일 실시예에 따르면, 카테고리들은 기 설정된 카테고리들일 수 있다. 이와 같이, 제1 영상(x)의 분류 정보는 제1 영상(x)이 기 설정된 복수의 카테고리들 각각에 포함될 확률을 나타내는 데이터를 포함할 수 있다.
- [89] 다시 도 3을 참조하면, 제1 영상의 제1 특징 정보(F_{latent})는 제3 모듈(240)로 입력될 수 있다.
- [90] 한편, 중간 영상 생성부(320)는 제1 특징 추출부(310)에서 획득된 제1 영상의 제1 특징 정보(F_{latent})에 기초하여, 중간 출력 영상(y_1)을 생성할 수 있다. 중간 영상 생성부(320)는 중간 영상 생성 네트워크를 이용하여, 중간 출력 영상(y_1)을 생성할 수 있다. 중간 영상 생성 네트워크는 하나 이상의 컨볼루션 레이어들을 포함할 수 있다.
- [91] 또는, 중간 영상 생성부(320)는 제1 영상의 제1 특징 정보(F_{latent}) 및 분류 정보에 기초하여, 중간 출력 영상(y_1)을 생성할 수 있다. 다만, 이에 한정되지 않는다.
- [92] 중간 영상 생성부(320)의 세부적인 구성에 대해서는 도 5를 참조하여 자세히 설명하기로 한다.
- [93] 도 5는 일 실시예에 따른 중간 영상 생성부를 나타내는 도면이다.
- [94] 도 5를 참조하면, 일 실시예에 따른 중간 영상 생성부(320)는 제1 영상의 제1 특징 정보(F_{latent})와 제4 컨볼루션 레이어(510)에 포함된 제4 가중치에 기초하는 컨볼루션 연산을 수행함으로써, 제4 특징 데이터를 획득할 수 있다. 예를 들어, 중간 영상 생성부(320)는 제1 영상의 제1 특징 정보(F_{latent})와 제4 컨볼루션 레이어(510)에 포함된 제4 가중치와의 컨볼루션 연산을 수행하여, 제4 특징 데이터를 획득할 수 있다.
- [95] 중간 영상 생성부(320)는 제1 업샘플링 레이어(520)에서, 제4 특징 데이터의 업샘플링을 수행하여, 제1 업샘플링된 데이터를 획득할 수 있다. 예를 들어, 중간 영상 생성부(320)는 바이큐빅(bicubic), 바이리니어(bilinear) 등의 보간법을 이용하여, 업샘플링을 수행할 수 있다. 또는, 중간 영상 생성부(320)는 컨볼루션 연산을 통한 업샘플링을 수행할 수 있다. 다만, 이에 한정되지 않는다. 이와 같이, 다른 실시예에 따르면, 중간 영상 생성부(320)는 다른 방법을 이용하여, 업샘플링을 수행할 수 있다.
- [96] 중간 영상 생성부(320)는 제1 업샘플링된 데이터와 제5 컨볼루션 레이어(530)에 포함된 제5 가중치에 기초하는 컨볼루션 연산을 수행함으로써, 제5 특징 데이

- 터를 획득할 수 있다. 예를 들어, 중간 영상 생성부(320)는 제1 업샘플링된 데이터와 제5 가중치와의 컨볼루션 연산을 수행하여, 제5 특징 데이터를 획득할 수 있다.
- [97] 중간 영상 생성부(320)는 제2 업샘플링 레이어(540)에서, 제1 영상의 제1 특징 정보(F_{latent})의 업샘플링을 수행하여, 제2 업샘플링된 데이터를 획득할 수 있다.
- [98] 예를 들어, 중간 영상 생성부(320)는 바이큐빅(bicubic), 바이리니어(bilinear) 등의 보간법을 이용하여, 업샘플링을 수행할 수 있다. 또는, 중간 영상 생성부(320)는 컨볼루션 연산을 통한 업샘플링을 수행할 수 있다. 다만, 이에 한정되지 않는다.
- [99] 중간 영상 생성부(320)는 업샘플링된 데이터와 제6 컨볼루션 레이어(550)에 포함된 제6 가중치에 기초하는 컨볼루션 연산을 수행함으로써, 제6 특징 데이터를 획득할 수 있다. 예를 들어, 중간 영상 생성부(320)는 제2 업샘플링된 데이터와 제6 가중치와의 컨볼루션 연산을 수행하여, 제6 특징 데이터를 획득할 수 있다.
- [100] 중간 영상 생성부(320)는 합산 레이어(560)에서, 제5 특징 데이터와 제6 특징 데이터의 요소별 덧셈 연산을 수행할 수 있다.
- [101] 일 실시예에 따른 중간 영상 생성부(320)는 제4 컨볼루션 레이어(510), 제1 업샘플링 레이어(520), 제5 컨볼루션 레이어(530), 제2 업샘플링 레이어(540), 제6 컨볼루션 레이어(550), 및 합산 레이어(560)를 포함하는 모듈(501)이 직렬로 반복적으로 배열되는 구조를 포함할 수 있다. 다만, 이에 한정되지 않는다.
- [102] 다만, 이에 한정되지 않으며, 중간 영상 생성부(320)는 제1 영상(x)에 대한 특징 정보로부터 제1 영상(x)에서 아티팩트를 제거한 중간 출력 영상을 생성할 수 있는 디코더(decoder)의 역할을 하는 다양한 구조를 포함할 수 있다.
- [103] 중간 영상 생성부(320)는 모듈(501)이 직렬로 반복적으로 배열되는 구조에 있어, 각 모듈(501)에서 출력되는 제3 특징 정보를 제3 모듈(240)로 입력할 수 있다.
- [104] 도 6은 일 실시예에 따른 맵 생성부를 나타내는 도면이다.
- [105] 도 6을 참조하면, 맵 생성부(220)는 제1 영상(x) 및 중간 출력 영상(y1)에 기초하여, 어텐션 맵(650)을 생성할 수 있다.
- [106] 일 실시예에 따른 맵 생성부(220)는 제1 영상(x)과 중간 출력 영상(y1)의 차이 정보를 나타내는 레지듀얼 맵(610)을 획득할 수 있다. 예를 들어, 맵 생성부(220)는 제1 영상(x)과 중간 출력 영상(y1)의 요소별 뺄셈 연산을 통해 레지듀얼 맵(610)을 획득할 수 있다.
- [107] 중간 출력 영상(y1)은 제1 영상(x)에서 아티팩트가 제거된 영상일 수 있으며, 이에 따라, 레지듀얼 맵(610)은 제1 영상의 아티팩트 정보를 나타낼 수 있다. 다만, 이에 한정되지 않는다.
- [108] 맵 생성부(220)는 레지듀얼 맵(610)을 이용하여, 로컬 맵(620)과 글로벌 분산 값들(630)을 획득할 수 있다.
- [109] 예를 들어, 맵 생성부(220)는 제1 크기를 가지는 윈도우를 레지듀얼 맵(610)에서 한 픽셀씩 이동시키면서, 레지듀얼 맵(610)의 픽셀들 중 윈도우에 포함되는 픽

셀들의 분산 값을 획득하고, 획득한 분산 값들을 포함하는 로컬 맵(620)을 생성할 수 있다. 여기서, 획득된 분산 값들은 로컬 분산 값들로 지칭될 수 있다. 이때, 획득한 로컬 분산 값들의 개수는 레지듀얼 맵(610)에 포함되는 픽셀들의 개수와 동일할 수 있으며, 로컬 맵(620)은 레지듀얼 맵(610)과 크기가 동일할 수 있다.

- [110] 또한, 맵 생성부(220)는 제1 크기보다 큰 제2 크기를 가지는 영역을 이용하여, 글로벌 분산 값들(630)을 획득할 수 있다. 예를 들어, 맵 생성부(220)는 레지듀얼 맵(610)을 제2 크기를 가지는 영역들로 분할하고, 분할된 영역들 각각에 대한 분산 값을 획득할 수 있다. 여기서, 획득된 분산 값들은 글로벌 분산 값들로 지칭될 수 있다. 이때, 획득한 글로벌 분산 값들의 개수는 분할된 영역들의 개수와 동일할 수 있다. 일 실시예에 따르면, 영역들은 패치들로 지칭될 수 있다.
- [111] 맵 생성부(220)는 로컬 맵(620)에 패치들 각각에 대응하는 글로벌 분산 값을 곱함으로써, 어텐션 맵(650)을 생성할 수 있다.
- [112] 맵 생성부(220)에서 생성된 어텐션 맵(650)은 제2 모듈(230)로 입력될 수 있다.
- [113] 도 7 및 도 8은 일 실시예에 따른 제3 모듈을 나타내는 도면들이다.
- [114] 도 7은 일 실시예에 따른 제3 모듈이 제1 영상의 제1 특징 정보(F_{latent})를 처리하는 방법을 나타내는 도면이다. 일 실시예에 따르면, 제3 모듈(240)은 정규화 레이어(710), 및 제2 완전 연결 레이어(721) 및 제2 완전 연결 레이어(722) 및 제n 완전 연결 레이어(729)를 포함하는 완전 연결 레이어들을 포함할 수 있다.
- [115] 도 7을 참조하면, 제3 모듈(240)은 정규화 레이어(710)에서, 제1 모듈(210)로부터 입력된 제1 영상의 제1 특징 정보(F_{latent})를 정규화할 수 있다. 예를 들어, 제3 모듈(240)은 제1 특징 정보(F_{latent})의 합이 1이 되도록 제1 특징 정보(F_{latent})를 정규화할 수 있다. 다만, 이에 한정되지 않으며, 다양한 정규화 방법을 이용하여, 제1 특징 정보(F_{latent})를 정규화할 수 있다. 정규화된 제1 특징 정보(F_{latent})(예를 들어, 제1 데이터)는 다중 퍼셉트론(MLP: Multi-Layer Perceptron) 모듈로 입력될 수 있다.
- [116] 일 실시예에 따른 제3 모듈(240)은 다중 퍼셉트론 모듈을 포함할 수 있으며, 다중 퍼셉트론 모듈은 n개의 완전 연결 레이어들(FCs: Fully Connected layers)을 포함할 수 있으며, 여기서 n은 정수이다. 제1 데이터는 제1 완전 연결 레이어(721)로 입력될 수 있다.
- [117] 제3 모듈(230)은 제1 데이터와 제1 완전 연결 레이어(721)에 포함되는 제1 가중치 행렬과의 곱셈 연산 및 제1 완전 연결 레이어(721)에 포함되는 바이어스와의 덧셈 연산을 수행할 수 있다.
- [118] 제3 모듈(240)은 제1 완전 연결 레이어(721)로부터 출력된 제2 데이터와 제2 완전 연결 레이어(722)에 포함되는 제2 가중치 행렬과의 곱셈 연산 및 제2 완전 연결 레이어(722)에 포함되는 바이어스와의 덧셈 연산을 수행할 수 있다.
- [119] 또한, 제3 모듈(240)은 제 n-1 완전 연결 레이어로부터 출력된 제3 데이터와 제 n 완전 연결 레이어(729)에 포함되는 제3 가중치 행렬과의 곱셈 연산 및 제n 완전

- 연결 레이어(729)에 포함되는 바이어스와의 덧셈 연산을 수행하여, 제2 특징 정보를 획득할 수 있다.
- [120] 제 n 완결 레이어(729)에서 출력되는 제2 특징 정보는 제2 모듈(230)로 입력될 수 있다.
- [121] 또는, 일 실시예에 따른 제3 모듈(240)은 제1 특징 정보와 분류 정보를 함께 처리하여, 제2 특징 정보를 획득할 수도 있다.
- [122] 도 8은 일 실시예에 따른 제3 모듈이 제3 특징 정보를 처리하는 방법을 나타내는 도면이다.
- [123] 도 8을 참조하면, 제1 모듈(210)의 중간 영상 생성부(320)로부터 출력된 제3 특징 정보는 제3 모듈(240)로 입력될 수 있다.
- [124] 일 실시예에 따른 제3 특징 정보는 제1 영상의 제1 특징 정보(F_{latent})로부터 중간 출력 영상(y_1)을 생성하는 과정의 중간에서 출력되는 복수의 중간 데이터들(제1 내지 제 n 중간 데이터들)을 포함할 수 있다. 예를 들어, 복수의 중간 데이터들은 n 개의 중간 데이터들을 포함할 수 있으며, 여기서 n 은 정수이다. 예를 들어, 복수의 중간 데이터들은 제1 중간 데이터, 제2 중간 데이터, ... 및 제 n 중간 데이터를 포함할 수 있다.
- [125] 예를 들어, 도 5에서 설명한 바와 같이, 중간 영상 생성부(320)에 포함된 복수의 모듈(501)들 각각에서 제1 내지 제 n 중간 데이터들이 출력될 수 있다.
- [126] 일 실시예에 따른 제3 모듈(240)은 제3 특징 정보에 포함되는 제1 내지 제 n 중간 데이터들 각각과 제1 내지 제 n 컨볼루션 레이어들(810, 820, ..., 890) 각각에 포함되는 가중치의 컨볼루션 연산을 수행함으로써, 제4 특징 정보를 획득할 수 있다.
- [127] 예를 들어, 제3 모듈(240)은 제1 중간 데이터와 제1 컨볼루션 레이어(810)에 포함되는 제1 가중치에 기초하는 컨볼루션 연산을 수행함으로써, 제1 중간 출력 데이터를 획득할 수 있다. 예를 들어, 제3 모듈(240)은 제1 중간 데이터 및 제1 컨볼루션 레이어(810)에 포함되는 제1 가중치와의 컨볼루션 연산을 수행하여, 제1 중간 출력 데이터를 획득할 수 있다.
- [128] 제3 모듈(240)은 제2 중간 데이터와 제2 컨볼루션 레이어(820)에 포함되는 제2 가중치에 기초하는 컨볼루션 연산을 수행함으로써, 제2 중간 출력 데이터를 획득할 수 있다. 예를 들어, 제3 모듈(240)은 제2 중간 데이터 및 제2 컨볼루션 레이어(820)에 포함되는 제2 가중치와의 컨볼루션 연산을 수행하여, 제2 중간 출력 데이터를 획득할 수 있다.
- [129] 또한, 제3 모듈(240)은 제 n 중간 데이터와 제 n 컨볼루션 레이어(890)에 포함되는 제 n 가중치에 기초하는 컨볼루션 연산을 수행함으로써, 제 n 중간 출력 데이터를 획득할 수 있다. 예를 들어, 제3 모듈(240)은 제 n 중간 데이터와 제 n 컨볼루션 레이어(890)에 포함되는 제 n 가중치와의 컨볼루션 연산을 수행하여, 제 n 중간 출력 데이터를 획득할 수 있다. 일 실시예에 따른 제4 특징 정보는 제1 내지 제 n 중간 출력 데이터들을 포함할 수 있다.

[130] 제3 모듈(240)에서 출력된 제1 내지 제 n 중간 출력 데이터들(제4 특징 정보)은 제2 모듈(230)로 입력될 수 있다.

[131] 제2 모듈(230)의 세부적인 구성에 대해서는 도 9를 참조하여 자세히 설명하기로 한다.

[132] 도 9는 일 실시예에 따른 제2 모듈을 나타내는 도면이다.

[133] 도 9를 참조하면, 제2 모듈(230)은 최종 영상 생성 네트워크(900)를 포함할 수 있다. 제2 모듈(230)은 최종 영상 생성 네트워크(900)를 이용하여, 제2 영상(y2)을 생성할 수 있다.

[134] 도 9에 도시된 바와 같이, 기 학습된 파라미터(910)는 최종 영상 생성 네트워크(900)로 입력될 수 있다. 예를 들어, 기 학습된 파라미터(910)의 크기는 512 x 4 x 4일 수 있다. 일 실시예에 따른 파라미터(910)는 영상 처리 네트워크(103)가 학습될 때 함께 학습되는 파라미터일 수 있다. 다만, 이에 한정되지 않는다.

[135] 일 실시예에 따른 제2 모듈(230)은 업샘플링 레이어(920)에서, 입력된 파라미터(910)의 업샘플링을 수행하여, 제1 데이터를 획득하여, 제3 업샘플링된 데이터를 획득할 수 있다.

[136] 제2 모듈(230)은 제1 데이터와 제3 모듈(240)로부터 입력된 제2 특징 정보에 기초하여, AdaIN(Adaptive Instance Normalization) 연산부(930)에서 AdaIN 연산을 수행할 수 있다. AdaIN 연산은 다음과 같은 수학적 식 1과 같이 나타낼 수 있다.

[137] [수식1]

$$\text{AdaIn}(x,y) = \sigma(y) \left(\frac{x - \mu(x)}{\sigma(x)} \right) + \mu(y)$$

[138] 수학적 식 1에서 $\mu(\cdot)$ 는 평균을 나타내고, $\sigma(\cdot)$ 는 표준 편차를 나타낸다. $\text{AdaIn}(x, y)$ 는 x값을 x의 평균과 표준편차를 이용하여, 정규화한 후, y의 평균과 표준편차를 이용하여, x값을 y의 분포로 변화하는 것을 의미할 수 있다. 즉, 특징 x를 특징 y의 특성으로 변조(modulation)하는 것을 의미할 수 있다.

[139] 따라서, 제2 모듈(230)은 AdaIN 연산부(930)에서 AdaIN 연산을 통해 제1 데이터를 제2 특징 정보의 특성으로 변조함으로써, 제2 데이터를 획득할 수 있다.

[140] 제2 모듈(230)은 제2 데이터와 컨볼루션 레이어(940)에 포함된 가중치에 기초하는 컨볼루션 연산을 수행함으로써, 제3 데이터를 획득할 수 있다.

[141] 제2 모듈(230)은 제3 데이터와 제3 모듈(240)로부터 입력된 제2 특징 정보에 기초하여, AdaIN 연산부(950)에서 AdaIN 연산을 수행할 수 있다. 이에 따라, 제2 모듈(230)은 AdaIN 연산을 통해 제3 데이터를 제2 특징 정보의 특성으로 변조함으로써, 제4 데이터를 획득할 수 있다.

[142] 또한, 제2 모듈(230)은 맵 생성부(220)로부터 입력된 어텐션 맵(650)을 스케일링(970)하여 제5 데이터를 획득할 수 있다. 제5 데이터는 Channel-Split SFT(Spatial Feature Transform) 연산부(960)로 입력될 수 있다.

- [143] 제2 모듈(230)은 제4 데이터와 제3 모듈(240)로부터 입력된 제4 특징 정보, 제5 데이터에 기초하여, Channel-Split SFT(Spatial Feature Transform) 연산부(960)에서 Channel-Split SFT 연산을 수행할 수 있다.
- [144] Channel-Split SFT(Spatial Feature Transform) 연산부(960)의 세부적인 구성에 대해서는 도 10을 참조하여 자세히 설명하기로 한다.
- [145] 도 10은 일 실시예에 따른 Channel-Split SFT 연산부를 나타내는 도면이다.
- [146] 도 10을 참조하면, 제2 모듈(230)은 스플릿 레이어(1010)에서 제4 데이터를 반으로 스플릿하여, 제6 데이터 및 제7 데이터를 획득할 수 있다.
- [147] 제2 모듈(230)은 스플릿된 제6 데이터와 제4 특징 정보의 SFT 연산을 수행할 수 있다. SFT 연산은 다음과 같은 수학적 식 2로 나타낼 수 있다.
- [148] [수식2]
- $$\mathbf{SFT}(\mathbf{F}|\gamma, \beta) = \gamma \odot \mathbf{F} + \beta$$
- [149] 수학적 식 2에서 \odot 는 요소별 곱셈 연산을 나타낼 수 있다. 또한, \mathbf{F} 는 제6 데이터를 의미하고, γ, β 는 제4 특징 정보에 기초한 데이터이다. 예를 들어, 제4 특징 정보를 반으로 나누어 각각 γ, β 에 할당할 수 있다. 따라서, $\mathbf{F}, \gamma, \beta$ 는 크기가 모두 동일할 수 있다.
- [150] $\mathbf{SFT}(\mathbf{F}|\gamma, \beta)$ 는 γ 와 β 를 이용하여, \mathbf{F} 를 어파인 변환(affine transform)시켜, \mathbf{F} 의 특성을 변조(modulation)시키는 것을 의미한다. 어파인 변환은 공간 리니어 매핑(spatial linear mapping) 방법 중 하나로 점들, 직선들, 면들이 보존될 수 있다.
- [151] 제2 모듈(230)은 제4 특징 정보를 이용하여, 제6 데이터의 특성을 변조함으로써, 제8 데이터를 획득할 수 있다.
- [152] 또한, 제2 모듈(230)은 스플릿된 제7 데이터와 제5 데이터의 요소별 곱셈 연산을 수행하여, 제9 데이터를 획득할 수 있다.
- [153] 제2 모듈(230)은 제8 데이터와 제9 데이터를 연결 레이어(1030)에서 연결시켜, 제10 데이터를 획득할 수 있다.
- [154] 다시 도 9를 참조하면, 최종 영상 생성 네트워크(900)는 업샘플링 레이어(920), 제1 AdaIn 연산부(930), 컨볼루션 레이어(940), 제2 AdaIN 연산부(950), 및 Channel-Split SFT 연산부(960)를 포함하는 모듈(901)이 직렬로 반복적으로 배열되는 구조를 포함할 수 있다. 다만, 이에 한정되지 않는다.
- [155] 한편, 제3 모듈(240)로부터 출력된 제4 특징 정보에 포함되는 제1 내지 제 n 중간 출력 데이터들은 각각 각 모듈(901)에 포함된 Channel-Split SFT 연산부(960)로 입력될 수 있다. 예를 들어, 제2 모듈(230)은 모듈(901) n개가 직렬로 배열되는 구조를 포함할 수 있으며, 제1 내지 제 n 중간 출력 데이터들이 각각 n개의 모듈들(901) 각각에 포함되는 Channel-Split SFT 연산부로 입력될 수 있다.
- [156] 제2 모듈(230)은 n개의 모듈들(901)을 통해, 제2 영상(y_2)을 생성할 수 있다.
- [157] 도 11은 일 실시예에 따른 영상 처리 네트워크를 나타내는 도면이다.

- [158] 도 11의 영상 처리 네트워크(1100)는 도 1의 영상 처리 네트워크(103)의 일 실시예일 수 있다.
- [159] 도 11을 참조하면, 영상 처리 네트워크(1100)는 제1 모듈(1110), 맵 생성부(1120), 제2 모듈(1130), 및 제3 모듈(1140)을 포함할 수 있다.
- [160] 일 실시예에 따른 제1 모듈(1110)로 제1 영상(x)이 입력될 수 있으며, 제1 모듈(1110)은 제1 영상(x)의 제1 특징 정보를 획득할 수 있다.
- [161] 제1 모듈(1110)에서 획득된 제1 영상(x)의 제1 특징 정보는 제3 모듈(1140)로 입력될 수 있다.
- [162] 또한, 제1 모듈(1110)은 제1 영상(x)의 제1 특징 정보에 기초하여, 중간 출력 영상(y1)을 생성할 수 있다. 예를 들어, 제1 모듈(1110)은 제1 영상(x)의 제1 특징 정보에 제1 영상 처리를 수행함으로써, 중간 출력 영상(y1)을 생성할 수 있다. 중간 출력 영상(y1)은 제1 영상(x)에서 아티팩트가 제거된 영상일 수 있다. 다만, 이에 한정되지 않는다.
- [163] 또한, 제1 모듈(1110)은 제1 영상 처리의 중간에서 추출되는 제3 특징 정보를 획득하고, 제3 특징 정보는 제3 모듈(1140)로 입력될 수 있다.
- [164] 도 11의 제1 모듈(1110)의 세부적인 구성 및 동작은 도 3 내지 도 5에서 도시하고 설명한 제1 모듈(210)의 세부적인 구성 및 동작과 동일할 수 있다. 예를 들어, 제1 모듈(1110)은 제1 특징 추출부 및 중간 영상 생성부를 포함할 수 있으며, 이에 대해서는 도 3 내지 도 5에서 자세히 설명하였으므로 동일한 설명은 생략하기로 한다.
- [165] 일 실시예에 따른 맵 생성부(1120)는 제1 영상(x)과 중간 출력 영상(y1)의 차이 정보에 기초한 레지듀얼 맵(residual map)을 획득하고, 레지듀얼 맵에 기초하여, 어텐션 맵을 생성할 수 있다. 맵 생성부(1120)에서 생성된 어텐션 맵은 제3 모듈(1140)로 입력될 수 있다.
- [166] 도 11의 맵 생성부(1120)의 세부적인 구성 및 동작은 도 6에서 도시하고 설명한 맵 생성부(220)의 세부적인 구성 및 동작과 동일할 수 있다. 예를 들어, 맵 생성부(1120)는 레지듀얼 맵을 이용하여, 로컬 맵과 글로벌 분산 값들을 획득하고, 로컬 맵과 글로벌 분산 값들을 이용하여, 어텐션 맵을 생성할 수 있다. 이에 대해서는 도 6에서 자세히 설명하였으므로 동일한 설명은 생략하기로 한다.
- [167] 일 실시예에 따른 제3 모듈(1140)은 제1 영상의 제1 특징 정보 및 맵 생성부(1120)로부터 획득한 어텐션 맵에 기초하여, 제6 특징 정보를 획득할 수 있다. 이에 대해서는 도 12를 참조하여 자세히 설명하기로 한다.
- [168] 도 12는 일 실시예에 따른 제3 모듈을 나타내는 도면이다.
- [169] 도 12를 참조하면, 일 실시예에 따른 제3 모듈(1140)은 어텐션 맵(1201)을 스케일링(1210)하여, 제1 정규화 레이어(1220)로 입력할 수 있다.
- [170] 제3 모듈(1140)은 스케일링된 어텐션 맵을 제1 정규화 레이어(1220)에서 정규화하여, 제1 데이터를 획득할 수 있다. 예를 들어, 제3 모듈(1140)은 스케일링된

어텐션 맵의 합이 1이 되도록 정규화할 수 있으나, 이에 한정되지 않으며, 다양한 정규화 방법을 이용하여, 스케일링된 어텐션 맵을 정규화할 수 있다.

- [171] 제1 데이터는 다중 퍼셉트론 모듈로 입력될 수 있다. 일 실시예에 따른 제3 모듈(1140)은 다중 퍼셉트론 모듈을 포함할 수 있으며, 다중 퍼셉트론 모듈은 n개의 완전 연결 레이어들을 포함할 수 있다. 제1 데이터는 제1 완전 연결 레이어(1231)로 입력될 수 있다.
- [172] 제3 모듈(1140)은 제1 데이터와 제1 완전 연결 레이어(1231)에 포함되는 제1 가중치 행렬과의 곱셈 연산 및 제1 완전 연결 레이어(1231)에 포함되는 바이어스와의 덧셈 연산을 수행할 수 있다.
- [173] 제3 모듈(1140)은 제1 완전 연결 레이어(1231)로부터 출력된 제2 데이터와 제2 완전 연결 레이어(1232)에 포함되는 제2 가중치 행렬과의 곱셈 연산 및 제2 완전 연결 레이어(1232)에 포함되는 바이어스와의 덧셈 연산을 수행할 수 있다.
- [174] 또한, 제3 모듈(1140)은 제n-1 완전 연결 레이어로부터 출력된 제3 데이터와 제n 완전 연결 레이어(1239)에 포함되는 제3 가중치 행렬과의 곱셈 연산 및 제n 완전 연결 레이어(1239)에 포함되는 바이어스와의 덧셈 연산을 수행하여, 제5 특징 정보를 획득할 수 있다.
- [175] 또한, 제3 모듈(1140)은 제2 정규화 레이어(1240)에서 제1 모듈(1110)로부터 입력된 제1 영상의 제1 특징 정보(F_{latent})를 정규화할 수 있다. 예를 들어, 제3 모듈(1140)은 제1 특징 정보(F_{latent})의 합이 1이 되도록 제1 특징 정보(F_{latent})를 정규화할 수 있다. 다만, 이에 한정되지 않으며, 다양한 정규화 방법을 이용하여, 제1 특징 정보(F_{latent})를 정규화할 수 있다. 정규화된 제1 특징 정보(F_{latent})(예를 들어, 제4 데이터)는 제2 다중 퍼셉트론(MLP: Multi-Layer Perceptron) 모듈로 입력될 수 있다.
- [176] 일 실시예에 따른 제3 모듈(1140)은 제2 다중 퍼셉트론 모듈을 포함할 수 있으며, 제2 다중 퍼셉트론 모듈은 n개의 완전 연결 레이어들을 포함할 수 있다. 제4 데이터는 제1 완전 연결 레이어(1251)로 입력될 수 있다.
- [177] 제3 모듈(1140)은 제4 데이터와 제1 완전 연결 레이어(1251)에 포함되는 제1 가중치 행렬과의 곱셈 연산 및 제1 완전 연결 레이어(1251)에 포함되는 바이어스와의 덧셈 연산을 수행할 수 있다.
- [178] 제3 모듈(1140)은 제1 완전 연결 레이어(1251)로부터 출력된 제5 데이터와 제2 완전 연결 레이어(1252)에 포함되는 제2 가중치 행렬과의 곱셈 연산 및 제2 완전 연결 레이어(1252)에 포함되는 바이어스와의 덧셈 연산을 수행할 수 있다.
- [179] 또한, 제3 모듈(1140)은 제n-1 완전 연결 레이어로부터 출력된 제6 데이터와 제n 완전 연결 레이어(1259)에 포함되는 제3 가중치 행렬과의 곱셈 연산 및 제n 완전 연결 레이어(1259)에 포함되는 바이어스와의 덧셈 연산을 수행하여, 제2 특징 정보를 획득할 수 있다.
- [180] 제3 모듈(1140)은 제5 특징 정보와 제2 특징 정보의 요소별 곱셈 연산을 수행함으로써, 제6 특징 정보를 획득할 수 있다.
- [181] 제6 특징 정보는 제2 모듈(1130)로 입력될 수 있다.

- [182] 다시 도 11을 참조하면, 제3 모듈(1140)은 제3 특징 정보에 기초하여, 제4 특징 정보를 획득할 수 있다. 제3 모듈(1140)은 제3 특징 정보에 제3 영상 처리를 수행함으로써, 제4 특징 정보를 획득할 수 있다. 도 11의 제3 모듈(1140)은 도 8에서 도시하고 설명한 제3 모듈(240)의 구조를 포함할 수 있다. 제3 모듈이 제3 특징 정보에 기초하여, 제4 특징 정보를 획득하는 동작에 대해서는 도 8에서 자세히 설명하였으므로 동일한 설명은 생략하기로 한다.
- [183] 제3 모듈(1140)에서 획득된 제4 특징 정보는 제2 모듈(1130)로 입력될 수 있다.
- [184] 제2 모듈(1130)은 제6 특징 정보 및 제4 특징 정보에 기초하여, 제2 영상(y_2)을 생성할 수 있다. 이에 대해서는 도 13를 참조하여 자세히 설명하기로 한다.
- [185] 도 13은 일 실시예에 따른 제2 모듈을 나타내는 도면이다.
- [186] 도 13을 참조하면, 기 학습된 파라미터(1310)가 제2 모듈(1130)로 입력될 수 있다. 예를 들어, 기 학습된 파라미터(1310)의 크기는 $512 \times 4 \times 4$ 일 수 있다. 일 실시예에 따른 파라미터(1310)는 영상 처리 네트워크(1100)가 학습될 때 함께 학습되는 파라미터일 수 있다. 다만, 이에 한정되지 않는다.
- [187] 일 실시예에 따른 제2 모듈(1130)은 업샘플링 레이어(1320)에서, 입력된 파라미터(1310)의 업샘플링을 수행하여, 제1 데이터를 획득할 수 있다.
- [188] 제2 모듈(1130)은 제1 데이터와 제3 모듈(1140)로부터 입력된 제6 특징 정보에 기초하여, AdaIN(Adaptive Instance Normalization) 연산부(1330)에서 AdaIN 연산을 수행할 수 있다. AdaIN 연산은 전술한 수학식 1과 같이 나타낼 수 있다.
- [189] 수학식 1의 AdaIN(x, y)는 x 값을 x 의 평균과 표준편차를 이용하여, 정규화한 후, y 의 평균과 표준편차를 이용하여, x 값을 y 의 분포로 변화하는 것을 의미할 수 있다. 즉, 특징 x 를 특징 y 의 특성으로 변조(modulation)하는 것을 의미할 수 있다.
- [190] 따라서, 제2 모듈(1130)은 AdaIN 연산부(1330)에서 AdaIN 연산을 통해 제1 데이터를 제6 특징 정보의 특성으로 변조함으로써, 제2 데이터를 획득할 수 있다.
- [191] 제2 모듈(1130)은 제2 데이터와 컨볼루션 레이어(1340)에 포함된 가중치에 기초하는 컨볼루션 연산을 수행함으로써, 제3 데이터를 획득할 수 있다.
- [192] 제2 모듈(1130)은 제3 데이터와 제3 모듈(1140)로부터 입력된 제6 특징 정보에 기초하여, AdaIN 연산부(1350)에서 AdaIN 연산을 수행할 수 있다. 이에 따라, 제2 모듈(1130)은 AdaIN 연산을 통해 제3 데이터를 제6 특징 정보의 특성으로 변조함으로써, 제4 데이터를 획득할 수 있다.
- [193] 제2 모듈(1130)은 제4 데이터와 제3 모듈(1140)로부터 입력된 제4 특징 정보에 기초하여, SFT(Spatial Feature Transform) 연산부(1360)에서 SFT 연산을 수행할 수 있다.
- [194] SFT 연산은 전술한 수학식 2와 같이 나타낼 수 있다.
- [195] 수학식 2에서 F 는 제4 데이터를 의미하고, γ, β 는 제4 특징 정보에 기초한 데이터이다. 예를 들어, 제4 특징 정보를 제4 데이터 크기(dimension)의 2배가 되도록 조절한 후, 2로 나누어 각각 γ, β 에 할당할 수 있다. 따라서, F, γ, β 는 크기가 모두 동일할 수 있다.

- [196] SFT($F\gamma, \beta$)는 γ 와 β 를 이용하여, F 를 어파인 변환(affine transform)시켜, F 의 특성을 변조(modulation)시키는 것을 의미한다. 어파인 변환은 공간 리니어 매핑(spatial linear mapping) 방법 중 하나로 점들, 직선들, 면들이 보존될 수 있다.
- [197] 제2 모듈(1130)은 제4 특징 정보를 이용하여, 제4 데이터의 특성을 변조함으로써, 제5 데이터를 획득할 수 있다.
- [198] 제2 모듈(1130)은 업샘플링 레이어(1320), 제1 AdaIn 연산부(1330), 컨볼루션 레이어(1340), 제2 AdaIn 연산부(1350), 및 SFT 연산부(1360)를 포함하는 모듈(1301)이 직렬로 반복적으로 배열되는 구조를 포함할 수 있다. 다만, 이에 한정되지 않는다.
- [199] 한편, 제3 모듈(1140)로부터 출력된 제4 특징 정보에 포함되는 제1 내지 제 n 중간 출력 데이터들은 각각 각 모듈(1301)에 포함된 SFT 연산부(1360)로 입력될 수 있다. 예를 들어, 제2 모듈(1130)은 모듈(1301) n 개가 직렬로 배열되는 구조를 포함할 수 있으며, 제1 내지 제 n 중간 출력 데이터들이 각각 n 개의 모듈들 각각에 포함되는 SFT 연산부로 입력될 수 있다.
- [200] 제2 모듈(1130)은 n 개의 모듈들을 통해, 제2 영상(y_2)을 생성할 수 있다.
- [201] 제2 영상(y_2)은 제1 영상(x)보다 고화질일 수 있다.
- [202] 도 14는 일 실시예에 따른 영상 처리 네트워크를 나타내는 도면이다.
- [203] 도 14의 영상 처리 네트워크(1400)는 도 1의 영상 처리 네트워크(103)의 일 실시예일 수 있다.
- [204] 도 14의 영상 처리 네트워크(1400)는 제1 모듈(1410), 맵 생성부(1420), 제2 모듈(1430), 및 제3 모듈(1440)을 포함할 수 있다.
- [205] 일 실시예에 따른 제1 모듈(1410)로 제1 영상(x)이 입력될 수 있으며, 제1 모듈(1410)은 제1 영상(x)의 제1 특징 정보를 획득할 수 있다.
- [206] 제1 모듈(1410)에서 획득된 제1 영상(x)의 제1 특징 정보는 제3 모듈(1440)로 입력될 수 있다.
- [207] 또한, 제1 모듈(1410)은 제1 영상(x)의 제1 특징 정보에 기초하여, 중간 출력 영상(y_1)을 생성할 수 있다. 예를 들어, 제1 모듈(1410)은 제1 영상(x)의 제1 특징 정보에 제1 영상 처리를 수행함으로써, 중간 출력 영상(y_1)을 생성할 수 있다. 중간 출력 영상(y_1)은 제1 영상(x)에서 아티팩트가 제거된 영상일 수 있다. 다만, 이에 한정되지 않는다.
- [208] 또한, 제1 모듈(1410)은 제1 영상 처리의 중간에서 추출되는 제3 특징 정보를 획득하고, 제3 특징 정보는 제3 모듈(1440)로 입력될 수 있다.
- [209] 도 14의 제1 모듈(1410)의 세부적인 구성 및 동작은 도 3 내지 도 5에서 도시하고 설명한 제1 모듈(210)의 세부적인 구성 및 동작과 동일할 수 있다. 예를 들어, 제1 모듈(1410)은 제1 특징 추출부 및 중간 영상 생성부를 포함할 수 있으며, 이에 대해서는 도 3 내지 도 5에서 자세히 설명하였으므로 동일한 설명은 생략하기로 한다.

- [210] 일 실시예에 따른 맵 생성부(1420)는 제1 영상(x)과 중간 출력 영상(y1)의 차이 정보에 기초한 레지듀얼 맵(residual map)을 획득하고, 레지듀얼 맵에 기초하여, 어텐션 맵을 생성할 수 있다. 맵 생성부(1420)에서 생성된 어텐션 맵은 제3 모듈(1440) 및 제2 모듈(1430)로 입력될 수 있다.
- [211] 도 14의 맵 생성부(1420)의 세부적인 구성 및 동작은 도 6에서 도시하고 설명한 맵 생성부(220)의 세부적인 구성 및 동작과 동일할 수 있다. 예를 들어, 맵 생성부(1420)는 레지듀얼 맵을 이용하여, 로컬 맵과 글로벌 분산 값들을 획득하고, 로컬 맵과 글로벌 분산 값들을 이용하여, 어텐션 맵을 생성할 수 있다. 이에 대해서는 도 6에서 자세히 설명하였으므로 동일한 설명은 생략하기로 한다.
- [212] 일 실시예에 따른 제3 모듈(1440)은 제1 영상의 제1 특징 정보 및 맵 생성부(1420)로부터 획득한 어텐션 맵에 기초하여, 제6 특징 정보를 획득할 수 있다. 제3 모듈(1440)의 세부적인 구성 및 동작은 도 12에서 도시하고 설명한 제3 모듈(1140)의 세부적인 구성 및 동작과 동일할 수 있다. 제1 특징 정보 및 어텐션 맵에 기초하여, 제6 특징 정보를 획득하는 동작에 대해서는 도 12에서 자세히 설명하였으므로, 동일한 설명은 생략하기로 한다. 제3 모듈(1440)에서 획득된 제6 특징 정보는 제2 모듈(1430)로 입력될 수 있다.
- [213] 또한, 제3 모듈(1440)은 제3 특징 정보에 기초하여, 제4 특징 정보를 획득할 수 있다. 제3 모듈(1440)은 제3 특징 정보에 제3 영상 처리를 수행함으로써, 제4 특징 정보를 획득할 수 있다. 도 14의 제3 모듈(1440)은 도 8에서 도시하고 설명한 제3 모듈(240)의 구조를 포함할 수 있다. 제3 모듈(240)이 제3 특징 정보에 기초하여, 제4 특징 정보를 획득하는 동작에 대해서는 도 8에서 자세히 설명하였으므로 동일한 설명은 생략하기로 한다.
- [214] 제3 모듈(1440)에서 획득된 제4 특징 정보는 제2 모듈(1430)로 입력될 수 있다.
- [215] 제2 모듈(1430)은 제6 특징 정보, 제4 특징 정보 및 어텐션 맵에 기초하여, 제2 영상(y2)을 생성할 수 있다. 제2 모듈(1430)의 세부적인 구성 및 동작은 도 9에서 도시하고 설명한 제2 모듈(230)의 세부적인 구성 및 동작과 동일할 수 있다. 다만, 제2 모듈(1430)에서는 도 9의 제2 모듈(230)과 달리 제2 특징 정보 대신에 제6 특징 정보가 AdaIN 연산부로 입력되며, 나머지 구성 및 동작은 동일할 수 있다. 제2 특징 정보, 제4 특징 정보 및 어텐션 맵에 기초하여, 제2 영상을 생성하는 동작에 대해서는 도 9에서 자세히 설명하였으므로, 동일한 설명은 생략하기로 한다.
- [216] 도 15는 일 실시예에 따른 영상 처리 장치의 동작 방법을 나타내는 흐름도이다.
- [217] 일 실시예에 따른 영상 처리 장치(100)는 하나 이상의 뉴럴 네트워크들을 이용하여, 제1 영상의 제1 특징 정보를 획득할 수 있다(S1510).
- [218] 영상 처리 장치(100)는 하나 이상의 컨볼루션 레이어들을 포함하는 특징 추출 네트워크를 이용하여, 제1 영상의 제1 특징 정보를 획득할 수 있다.
- [219] 예를 들어, 영상 처리 장치(100)는 제1 컨볼루션 레이어, 제2 컨볼루션 레이어, 제3 컨볼루션 레이어 및 합산 레이어를 이용하여, 제1 영상의 제1 특징 정보를 획득할 수 있다. 제1 컨볼루션 레이어, 제2 컨볼루션 레이어, 제3 컨볼루션 레이어

및 합산 레이어를 포함하는 모듈(401)에 대해서는 도 4에서 자세히 설명하였으므로 동일한 설명은 생략하기로 한다.

- [220] 또한, 일 실시예에 따른 영상 처리 장치(100)는 제1 영상의 분류 정보를 획득할 수 있다. 예를 들어, 영상 처리 장치(100)는 영상 추출 네트워크를 이용하여, 제1 영상의 제1 특징 정보와 분류 정보를 함께 추출할 수 있다. 다만, 이에 한정되지 않는다.
- [221] 일 실시예에 따른 영상 처리 장치(100)는 제1 영상의 제1 특징 정보에 제1 영상 처리를 수행함으로써, 중간 출력 영상을 생성할 수 있다(S1520).
- [222] 영상 처리 장치(100)는 하나 이상의 컨볼루션 레이어들을 포함하는 중간 영상 생성부를 이용하여, 중간 출력 영상을 생성할 수 있다. 예를 들어, 영상 처리 장치(100)는 중간 영상 생성부에 포함되는 제4 컨볼루션 레이어, 제1 업샘플링 레이어, 제5 컨볼루션 레이어, 제2 업샘플링 레이어, 제6 컨볼루션 레이어, 및 합산 레이어를 이용하여, 중간 출력 영상을 생성할 수 있다.
- [223] 중간 영상 생성부에 대해서는 도 5에서 자세히 설명하였으므로 동일한 설명은 생략하기로 한다.
- [224] 일 실시예에 따른 영상 처리 장치(100)가 분류 정보를 획득한 경우, 영상 처리 장치(100)는 제1 특징 정보와 분류 정보를 함께 중간 영상 생성부에 입력함으로써, 중간 출력 영상을 생성할 수 있다. 다만, 이에 한정되지 않는다.
- [225] 일 실시예에 따른 영상 처리 장치(100)는 제1 영상 및 중간 출력 영상에 기초하여, 어텐션 맵을 생성할 수 있다(S1530).
- [226] 일 실시예에 따른 영상 처리 장치(100)는 제1 영상과 중간 출력 영상의 차이 정보를 나타내는 레지듀얼 맵을 획득할 수 있다. 예를 들어, 영상 처리 장치(100)는 제1 영상과 중간 출력 영상의 요소별 뺄셈 연산을 통해 레지듀얼 맵을 획득할 수 있다.
- [227] 영상 처리 장치(100)는 레지듀얼 맵을 이용하여, 로컬 맵과 글로벌 분산 값들을 획득할 수 있다. 예를 들어, 영상 처리 장치(100)는 제1 크기를 가지는 윈도우를 이용하여, 로컬 분산 값들을 획득할 수 있으며, 로컬 분산 값들을 포함하는 로컬 맵을 획득할 수 있다. 또한, 영상 처리 장치(100)는 제1 크기보다 큰 제2 크기를 가지는 패치를 이용하여, 패치들에 각각 대응하는 글로벌 분산 값들을 획득할 수 있다. 또한, 영상 처리 장치(100)는 로컬 맵에 패치들 각각에 대응하는 글로벌 분산 값을 곱함으로써, 어텐션 맵을 생성할 수 있다. 어텐션 맵을 생성하는 구조 및 과정에 대해서는 도 6에서 자세히 설명하였으므로 동일한 설명은 생략하기로 한다.
- [228] 일 실시예에 따른 영상 처리 장치(100)는 제1 특징 정보에 제2 영상 처리를 수행함으로써, 제2 특징 정보를 획득할 수 있다(S1540).
- [229] 예를 들어, 영상 처리 장치(100)는 하나 이상의 뉴럴 네트워크들을 이용하여, 제2 특징 정보를 생성할 수 있다. 영상 처리 장치(100)는 정규화 레이어, 제1 내지 제 n 완전 연결 레이어들을 이용하여, 제2 특징 정보를 생성할 수 있다. 제2 특징 정

보를 획득하는 구조 및 과정에 대해서는 도 7에서 자세히 설명하였으므로, 동일한 설명은 생략하기로 한다.

- [230] 일 실시예에 따른 영상 처리 장치(100)는 제1 영상 처리의 중간에서 추출되는 제3 특징 정보에 제3 영상 처리를 수행하여, 제4 특징 정보를 획득할 수 있다(S1550).
- [231] 예를 들어, 영상 처리 장치(100)는 제1 영상의 제1 특징 정보로부터 중간 출력 영상을 생성하는 과정의 중간에서 제3 특징 정보를 획득할 수 있다. 또는 영상 처리 장치(100)는 제1 영상의 제1 특징 정보 및 분류 정보로부터 중간 출력 영상을 생성하는 과정의 중간에서 제3 특징 정보를 획득할 수도 있다.
- [232] 영상 처리 장치(100)는 제3 특징 정보에 컨볼루션 연산을 수행함으로써, 제4 특징 정보를 획득할 수 있다. 제4 특징 정보를 획득하는 구조 및 과정에 대해서는 도 8에서 자세히 설명하였으므로, 동일한 설명은 생략하기로 한다.
- [233] 일 실시예에 따른 영상 처리 장치(100)는 1530 단계(S1530)에서 획득된 어텐션 맵, 1540 단계(S1540)에서 획득된 제2 특징 정보 및 1550 단계(S1550)에서 획득된 제4 특징 정보에 기초하여, 제2 영상을 생성할 수 있다(S1560).
- [234] 예를 들어, 영상 처리 장치(100)는 어텐션 맵, 제2 특징 정보 및 제4 특징 정보를 이용하여, 스케일링, 업 샘플링, AdaIN 연산, 컨볼루션 연산, Channel-Split SFT 연산 등을 수행함으로써, 제2 영상을 생성할 수 있다. 제2 영상을 생성하는 구조 및 과정에 대해서는, 도 9에서 자세히 설명하였으므로, 동일한 설명은 생략하기로 한다.
- [235] 일 실시예에 따른 제2 영상은 제1 영상보다 고해상도 영상일 수 있으며, 제1 영상에서 아티팩트, 노이즈 등이 제거됨으로써, 제1 영상보다 화질이 개선된 영상일 수 있다. 예를 들어, 제2 영상의 아티팩트들의 수는 제1 영상의 아티팩트들의 수보다 적을 수 있다. 다른 예로, 제2 영상의 노이즈의 양은 제1 영상의 노이즈의 양보다 적을 수 있다.
- [236] 도 16은 일 실시예에 따른 영상 처리 장치의 구성을 나타내는 블록도이다.
- [237] 도 16의 영상 처리 장치(100)는 영상 처리 네트워크(103, 1100, 1400)를 이용하여, 영상 처리를 수행하는 장치일 수 있다. 일 실시예에 따른 영상 처리 네트워크(103)는 하나 이상의 뉴럴 네트워크들(103, 1100, 1400)을 포함할 수 있다. 예를 들어, 영상 처리 네트워크(103, 1100, 1400)는 제1 모듈(210, 1110, 1410), 맵 생성부(220, 1120, 1420), 제2 모듈(230, 1130, 1430), 및 제3 모듈(240, 1140, 1440)을 포함할 수 있다. 다만, 이에 한정되지 않는다.
- [238] 도 16을 참조하면, 일 실시예에 따른 영상 처리 장치(100)는 프로세서(110), 메모리(120) 및 디스플레이(130)를 포함할 수 있다.
- [239] 일 실시예에 따른 프로세서(110)는 영상 처리 장치(100)를 전반적으로 제어할 수 있다. 일 실시예에 따른 프로세서(110)는 메모리(120)에 저장되는 하나 이상의 프로그램들을 실행할 수 있다.

- [240] 일 실시예에 따른 메모리(120)는 영상 처리 장치(100)를 구동하고 제어하기 위한 다양한 데이터, 프로그램 또는 어플리케이션을 저장할 수 있다. 메모리(120)에 저장되는 프로그램은 하나 이상의 인스트럭션들을 포함할 수 있다. 메모리(120)에 저장된 프로그램(하나 이상의 인스트럭션들) 또는 어플리케이션은 프로세서(110)에 의해 실행될 수 있다.
- [241] 일 실시예에 따른 프로세서(110)는 CPU(Central Processing Unit), GPU(Graphic Processing Unit) 및 VPU(Video Processing Unit) 중 적어도 하나를 포함할 수 있다. 또는, 실시예에 따라, CPU, GPU 및 VPU 중 적어도 하나를 통합한 SoC(System On Chip) 형태로 구현될 수 있다. 또는, 프로세서(110)는 NPU(Neural Processing Unit)를 더 포함할 수 있다.
- [242] 일 실시예에 따른 프로세서(110)는 하나 이상의 뉴럴 네트워크들을 이용하여, 제1 영상을 처리함으로써, 제2 영상을 생성할 수 있다. 예를 들어, 프로세서(110)는 영상 처리 네트워크(103, 1100, 1400)를 이용하여, 제1 영상의 노이즈를 제거하면서, 세밀한 가장자리 처리와 텍스처를 유지하는 디노이징을 수행한 제2 영상을 생성할 수 있다. 일 실시예에 따른 제2 영상은 제1 영상보다 고화질일 수 있다. 또는, 프로세서(110)는 영상 처리 네트워크(103)를 이용하여, 제1 영상보다 해상도가 높은 제2 영상을 생성할 수 있다.
- [243] 일 실시예에 따른 프로세서(110)는 제1 모듈(210, 1110, 1410)을 이용하여, 제1 영상의 제1 특징 정보를 획득할 수 있다. 제1 모듈(210, 1110, 1410)은 하나 이상의 컨볼루션 레이어들을 포함하는 특징 추출 네트워크를 포함할 수 있으며, 프로세서(110)는 특징 추출 네트워크를 이용하여, 제1 영상의 제1 특징 정보를 획득할 수 있다. 특징 추출 네트워크의 구성에 대해서는 도 4에서 자세히 설명하였으므로 구체적인 설명은 생략하기로 한다.
- [244] 또한, 프로세서(110)는 제1 모듈(210, 1110, 1410)을 이용하여, 제1 영상의 분류 정보를 획득할 수 있다. 제1 영상의 분류 정보는 제1 영상(x)이 기 설정된 복수의 카테고리들 각각에 포함될 확률을 나타내는 데이터를 포함할 수 있다.
- [245] 또한, 프로세서(110)는 제1 모듈(210, 1110, 1410)을 이용하여, 제1 영상의 제1 특징 정보에 제1 영상 처리를 수행함으로써, 중간 출력 영상을 생성할 수 있다. 또는, 프로세서(110)는 제1 모듈(210, 1110, 1410)을 이용하여, 제1 영상의 제1 특징 정보 및 분류 정보에 제1 영상 처리를 수행함으로써, 중간 출력 영상을 생성할 수 있다.
- [246] 일 실시예에 따른 중간 출력 영상은 제1 영상에서 아티팩트가 제거된 영상일 수 있다. 다만, 이에 한정되지 않는다.
- [247] 제1 모듈(210, 1110, 1410)은 하나 이상의 컨볼루션 레이어들을 포함하는 중간 영상 생성부를 이용하여, 중간 출력 영상을 생성할 수 있다. 중간 영상 생성부는 도 5에서 자세히 설명하였으므로 구체적인 설명은 생략하기로 한다.
- [248] 또한, 프로세서(110)는 중간 영상 생성부를 이용하여, 제1 영상 처리를 수행하는 과정의 중간에서 제3 특징 정보를 추출할 수 있다.

- [249] 예를 들어, 프로세서(110)는 제1 영상 처리를 수행하는 과정의 중간에서 제1 내지 제 n 중간 데이터들을 포함하는 제3 특징 정보를 획득할 수 있다.
- [250] 일 실시예에 따른 프로세서(110)는 맵 생성부(220, 1120, 1420)를 이용하여, 제1 영상 및 중간 출력 영상에 기초하여, 어텐션 맵을 생성할 수 있다. 이에 대해서는 도 6에서 자세히 설명하였으므로 동일한 설명은 생략하기로 한다.
- [251] 일 실시예에 따른 프로세서(110)는 제3 모듈(240, 1140, 1440)을 이용하여, 제1 특징 정보에 제2 영상 처리를 수행함으로써, 제2 특징 정보를 획득할 수 있다. 이에 대해서는 도 8에서 자세히 설명하였으므로 동일한 설명은 생략하기로 한다.
- [252] 일 실시예에 따른 프로세서(110)는 제3 모듈(240, 1140, 1440)을 이용하여, 제1 내지 제 n 중간 데이터들을 포함하는 제3 특징 정보에 제3 영상 처리를 수행함으로써, 제4 특징 정보를 획득할 수 있다. 이에 대해서는 도 7에서 자세히 설명하였으므로 동일한 설명은 생략하기로 한다.
- [253] 일 실시예에 따른 프로세서(110)는 제2 모듈(230, 1130, 1430)을 이용하여, 제2 특징 정보 및 제4 특징 정보에 기초한 제2 영상을 생성할 수 있다.
- [254] 예를 들어, 프로세서(110)는 어텐션 맵, 제2 특징 정보 및 제4 특징 정보를 이용하여, 스케일링, 업 샘플링, AdaIN 연산, 컨볼루션 연산, Channel-Split SFT 연산 등을 수행함으로써, 제2 영상을 생성할 수 있다. 이에 대해서는, 도 9에서 자세히 설명하였으므로, 동일한 설명은 생략하기로 한다.
- [255] 한편, 일 실시예에 따른 영상 처리 네트워크(103, 1100, 1400)는, 서버 또는 외부 장치에 의해 훈련된 네트워크일 수 있다. 외부 장치는 훈련 데이터에 기초하여, 영상 처리 네트워크(103, 1100, 1400)를 학습시킬 수 있다. 이때, 훈련 데이터는 노이즈가 포함된 영상 데이터와 노이즈는 제거되면서, 엣지 특성이나 텍스처 특성은 보존되는 영상 데이터를 포함하는 복수의 데이터 세트들을 포함할 수 있다.
- [256] 서버 또는 외부 장치는 영상 처리 네트워크(103, 1100, 1400)에 포함된 복수의 컨볼루션 레이어들 각각에서 이용되는 커널들에 포함되는 파라미터 값들 및 리니어 레이어들 또는 완전 연결 레이어들 각각에서 이용되는 가중치 행렬들에 포함되는 파라미터 값들을 결정할 수 있다. 예를 들어, 서버 또는 외부 장치는 영상 처리 네트워크(103, 1100, 1400)에 의해 생성된 영상 데이터와 노이즈는 제거되면서, 엣지 특성은 보존되는 영상 데이터(훈련 데이터)의 차이(손실 정보)를 최소화하는 방향으로 파라미터 값들을 결정할 수 있다.
- [257] 일 실시예에 따른 영상 처리 장치(100)는 서버 또는 외부 장치로부터 훈련이 완료된 영상 처리 네트워크(103, 1100, 1400)를 수신하여, 메모리(120)에 저장할 수 있다. 예를 들어, 메모리(120)는 일 실시예에 따른 영상 처리 네트워크(103, 1100, 1400)의 구조 및 파라미터 값들을 저장할 수 있으며, 프로세서(110)는 메모리(120)에 저장된 파라미터 값들을 이용하여, 일 실시예에 따른 제1 영상으로부터 노이즈는 제거되면서, 엣지 특성은 보존되는 제2 영상을 생성할 수 있다.
- [258] 일 실시예에 따른 디스플레이(130)는, 프로세서(110)에서 처리된 영상 신호, 데이터 신호, OSD 신호, 제어 신호 등을 변환하여 구동 신호를 생성한다. 디스플레이

이(130)는 PDP, LCD, OLED, 플렉시블 디스플레이(flexible display)등으로 구현될 수 있으며, 또한, 3차원 디스플레이(3D display)로 구현될 수 있다. 또한, 디스플레이(130)는, 터치 스크린으로 구성되어 출력 장치 이외에 입력 장치로 사용되는 것도 가능하다.

- [259] 일 실시예에 따른 디스플레이(130)는 영상 처리 네트워크(103, 1100, 1400)를 이용하여, 영상 처리된 제2 영상을 표시할 수 있다.
- [260] 한편, 도 16에 도시된 영상 처리 장치(100)의 블록도는 일 실시예를 위한 블록도이다. 블록도의 각 구성요소는 실제 구현되는 영상 처리 장치(100)의 사양에 따라 통합, 추가, 또는 생략될 수 있다. 즉, 필요에 따라 2 이상의 구성요소가 하나의 구성요소로 합쳐지거나, 혹은 하나의 구성요소가 2 이상의 구성요소로 세분되어 구성될 수 있다. 또한, 각 블록에서 수행하는 기능은 실시예들을 설명하기 위한 것이며, 그 구체적인 동작이나 장치는 본 발명의 권리범위를 제한하지 아니한다.
- [261] 일 실시예에 따른 영상 처리 장치는 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리할 수 있다.
- [262] 일 실시예에 따른 영상 처리 장치는 하나 이상의 인스트럭션들을 저장하는 메모리 및 상기 하나 이상의 인스트럭션들을 실행하는 적어도 하나의 프로세서를 포함할 수 있다.
- [263] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 제1 영상의 제1 특징 정보를 획득할 수 있다.
- [264] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 특징 정보에 제1 영상 처리를 수행함으로써, 상기 제1 영상에 대한 중간 출력 영상을 생성할 수 있다.
- [265] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 영상 및 상기 중간 출력 영상에 기초하여, 어텐션 맵(attention map)을 생성할 수 있다.
- [266] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 특징 정보에 제2 영상 처리를 수행함으로써, 제2 특징 정보를 획득할 수 있다.
- [267] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 영상 처리의 중간에서 추출되는 제3 특징 정보에 제3 영상 처리를 수행하여, 제4 특징 정보를 획득할 수 있다.
- [268] 일 실시예에 따른 적어도 하나의 프로세서는 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 어텐션 맵, 상기 제2 특징 정보 및 상기 제4 특징 정보에 기초하여, 상기 제1 영상보다 고화질인 제2 영상을 생성할 수 있다.

- [269] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 영상 및 상기 중간 출력 영상의 레지듀얼 맵(residual map)을 획득할 수 있다.
- [270] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 레지듀얼 맵의 로컬 맵(local map)과 글로벌(global) 분산 값들을 획득할 수 있다.
- [271] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 로컬 맵과 상기 글로벌 분산 값들에 기초하여, 상기 어텐션 맵을 생성할 수 있다.
- [272] 상기 로컬 맵은 제1 크기를 가지는 윈도우를 이용하여 획득한 상기 레지듀얼 맵의 로컬 분산 값들을 포함할 수 있다.
- [273] 상기 글로벌 분산 값들은 상기 제1 크기보다 큰 제2 크기를 가지는 패치를 이용하여 획득될 수 있다.
- [274] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 어텐션 맵에 제4 영상 처리를 수행함으로써, 제5 특징 정보를 획득할 수 있다.
- [275] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제2 특징 정보와 상기 제5 특징 정보에 기초하여, 제6 특징 정보를 획득할 수 있다.
- [276] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제6 특징 정보 및 상기 제4 특징 정보에 기초하여, 상기 제2 영상을 생성할 수 있다.
- [277] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 어텐션 맵에 제5 영상 처리를 수행함으로써, 제7 특징 정보를 획득할 수 있다.
- [278] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제2 특징 정보, 제4 특징 정보 및 제7 특징 정보에 기초하여, 상기 제2 영상을 생성할 수 있다.
- [279] 상기 하나 이상의 뉴럴 네트워크들은 하나 이상의 컨볼루션 뉴럴 네트워크들을 포함할 수 있다.
- [280] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 하나 이상의 컨볼루션 뉴럴 네트워크들을 이용하여, 상기 제1 영상의 제1 특징 정보를 획득할 수 있다.
- [281] 상기 제1 영상 처리는 상기 제1 특징 정보의 업 샘플링을 포함하고, 상기 제1 영상 및 상기 중간 영상은 크기가 동일할 수 있다.
- [282] 상기 제2 영상 처리는 하나 이상의 완전 연결 레이어들(fully connected layers)을 포함하는 MLP 모듈에 의해 수행될 수 있다.

- [283] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 하나의 완전 연결 레이어에 입력되는 입력 데이터와 상기 완전 연결 레이어에 포함되는 가중치 행렬과의 곱셈 연산 및 상기 완전 연결 레이어에 포함되는 바이어스와의 덧셈 연산을 수행할 수 있다.
- [284] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제3 특징 정보와 기 학습된 가중치와의 컨볼루션 연산을 수행함으로써, 상기 제4 특징 정보를 획득할 수 있다.
- [285] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제2 특징 정보에 제1 연산을 수행하여, 제8 특징 정보를 획득할 수 있다.
- [286] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제8 특징 정보 및 상기 제4 특징 정보에 제2 연산을 수행함으로써, 상기 제2 영상을 생성할 수 있다.
- [287] 상기 제1 연산은 AdaIn 연산을 포함할 수 있다.
- [288] 상기 제2 연산은 SFT 연산을 포함할 수 있다.
- [289] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 영상의 분류 정보를 획득할 수 있다.
- [290] 일 실시예에 따른 적어도 하나의 프로세서는, 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제1 특징 정보 및 상기 분류 정보에 상기 제1 영상 처리를 수행함으로써, 상기 제1 영상에 대한 중간 출력 영상을 생성할 수 있다.
- [291] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 제1 영상의 제1 특징 정보를 획득하는 단계를 포함할 수 있다.
- [292] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 상기 제1 특징 정보에 제1 영상 처리를 수행함으로써, 상기 제1 영상에 대한 중간 출력 영상을 생성하는 단계를 포함할 수 있다.
- [293] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 상기 제1 영상 및 상기 중간 출력 영상에 기초하여, 어텐션 맵을 생성하는 단계를 포함할 수 있다.
- [294] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 상기 제1 특징 정보에 제2 영상 처리를 수행함으로써, 제2 특징 정보를 획득하는 단계를 포함할 수 있다.
- [295] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 상기 제1 영상 처리의 중간에서 추출되는 제3 특징 정보에 제3 영상 처리를 수행함으로써, 제4 특징 정보를 획득하는 단계를 포함할 수 있다.

- [296] 일 실시예에 따른 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법은, 상기 어텐션 맵, 상기 제2 특징 정보 및 상기 제4 특징 정보에 기초하여, 상기 제1 영상보다 고화질인 제2 영상을 생성하는 단계를 포함할 수 있다.
- [297] 상기 제1 영상 및 상기 중간 출력 영상에 기초하여, 어텐션 맵을 생성하는 단계는, 상기 제1 영상 및 상기 중간 출력 영상의 레지듀얼 맵(residual map)을 획득하는 단계를 포함할 수 있다.
- [298] 상기 제1 영상 및 상기 중간 출력 영상에 기초하여, 어텐션 맵을 생성하는 단계는, 상기 레지듀얼 맵의 로컬 맵(local map)과 글로벌(global) 분산 값들을 획득하는 단계를 포함할 수 있다.
- [299] 상기 제1 영상 및 상기 중간 출력 영상에 기초하여, 어텐션 맵을 생성하는 단계는, 상기 로컬 맵과 상기 글로벌 분산 값들에 기초하여, 상기 어텐션 맵을 생성하는 단계를 포함할 수 있다.
- [300] 상기 레지듀얼 맵의 로컬 맵(local map)과 글로벌(global) 분산 값들을 획득하는 단계는, 제1 크기를 가지는 윈도우를 이용하여 상기 레지듀얼 맵의 로컬 분산 값들을 획득하는 단계를 포함할 수 있다.
- [301] 상기 레지듀얼 맵의 로컬 맵(local map)과 글로벌(global) 분산 값들을 획득하는 단계는, 상기 제1 크기보다 큰 제2 크기를 가지는 패치를 이용하여 상기 글로벌 분산 값들을 획득하는 단계를 포함할 수 있다.
- [302] 상기 제2 영상을 생성하는 단계는, 상기 어텐션 맵에 제4 영상 처리를 수행함으로써, 제5 특징 정보를 획득하는 단계를 포함할 수 있다.
- [303] 상기 제2 영상을 생성하는 단계는, 상기 제2 특징 정보와 상기 제5 특징 정보에 기초하여, 제6 특징 정보를 획득하는 단계를 포함할 수 있다.
- [304] 상기 제2 영상을 생성하는 단계는, 상기 제6 특징 정보 및 상기 제4 특징 정보에 기초하여, 상기 제2 영상을 생성하는 단계를 포함할 수 있다.
- [305] 상기 제2 영상을 생성하는 단계는, 상기 어텐션 맵에 제5 영상 처리를 수행함으로써, 제7 특징 정보를 획득하는 단계를 포함할 수 있다.
- [306] 상기 제2 영상을 생성하는 단계는, 상기 제2 특징 정보, 제4 특징 정보 및 제7 특징 정보에 기초하여, 상기 제2 영상을 생성하는 단계를 포함할 수 있다.
- [307] 상기 제1 특징 정보를 획득하는 단계는, 상기 하나 이상의 컨볼루션 뉴럴 네트워크들을 이용하여, 상기 제1 영상의 제1 특징 정보를 획득하는 단계를 포함할 수 있다.
- [308] 상기 제1 영상 처리는 상기 제1 특징 정보의 업 샘플링을 포함하고, 상기 제1 영상 및 상기 중간 영상은 크기가 동일할 수 있다.
- [309] 상기 제2 영상 처리는 하나 이상의 완전 연결 레이어들(fully connected layers)을 포함하는 MLP 모듈에 의해 수행될 수 있다.
- [310] 상기 제2 특징 정보를 획득하는 단계는, 하나의 완전 연결 레이어에 입력되는 입력 데이터와 상기 완전 연결 레이어에 포함되는 가중치 행렬과의 곱셈 연산 및

상기 완전 연결 레이어에 포함되는 바이어스와의 덧셈 연산을 수행하는 단계를 포함할 수 있다.

- [311] 상기 제4 특징 정보를 획득하는 단계는, 상기 제3 특징 정보와 기 학습된 가중치와의 컨볼루션 연산을 수행함으로써, 상기 제4 특징 정보를 획득하는 단계를 포함할 수 있다.
- [312] 상기 제2 영상을 생성하는 단계는, 상기 제2 특징 정보에 제1 연산을 수행하여, 제8 특징 정보를 획득할 수 있다.
- [313] 상기 제2 영상을 생성하는 단계는, 상기 제8 특징 정보 및 상기 제4 특징 정보에 제2 연산을 수행함으로써, 상기 제2 영상을 생성할 수 있다.
- [314] 상기 제1 연산은 AdaIn 연산을 포함할 수 있다.
- [315] 상기 제2 연산은 SFT 연산을 포함할 수 있다.
- [316] 일 실시예에 따른 영상 처리 장치의 동작 방법은, 상기 제1 영상의 분류 정보를 획득하는 단계를 더 포함할 수 있다.
- [317] 상기 중간 출력 영상을 생성하는 단계는, 상기 제1 특징 정보 및 상기 분류 정보에 상기 제1 영상 처리를 수행함으로써, 상기 제1 영상에 대한 중간 출력 영상을 생성하는 단계를 포함할 수 있다.
- [318] 일 실시예에 따른 영상 처리 장치는 제1 영상의 아티팩트 정보를 나타내는 레지듀얼 맵(잔차 영상)의 로컬 분산 값들을 포함하는 로컬 정보와 글로벌 분산 값들을 포함하는 글로벌 정보를 함께 이용하여, 어텐션 맵을 생성하고, 생성된 어텐션 맵을 이용하여 제1 영상을 처리함으로써, 고화질의 제2 영상을 생성할 수 있다.
- [319] 이에 따라, 일 실시예에 따른 영상 처리의 성능은 기존의 영상 처리 기술들에 비해 향상될 수 있다. 예를 들어, 생성된 제2 영상의 화질의 개선 정도나 아티팩트(노이즈)의 제거 정도가 기존 영상 처리 기술에 의해 처리된 영상에 비해 증가할 수 있다.
- [320] 일 실시예에 따른 영상 처리 장치의 동작 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 본 발명을 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.

- [321] 또한, 개시된 실시예들에 따른 영상 처리 장치 및 영상 처리 장치의 동작 방법은 컴퓨터 프로그램 제품(computer program product)에 포함되어 제공될 수 있다. 컴퓨터 프로그램 제품은 상품으로서 판매자 및 구매자 간에 거래될 수 있다.
- [322] 컴퓨터 프로그램 제품은 S/W 프로그램, S/W 프로그램이 저장된 컴퓨터로 읽을 수 있는 저장 매체를 포함할 수 있다. 예를 들어, 컴퓨터 프로그램 제품은 전자 장치의 제조사 또는 전자 마켓(예, 구글 플레이 스토어, 앱 스토어)을 통해 전자적으로 배포되는 S/W 프로그램 형태의 상품(예, 다운로드블 앱)을 포함할 수 있다. 전자적 배포를 위하여, S/W 프로그램의 적어도 일부는 저장 매체에 저장되거나, 임시적으로 생성될 수 있다. 이 경우, 저장 매체는 제조사의 서버, 전자 마켓의 서버, 또는 SW 프로그램을 임시적으로 저장하는 중계 서버의 저장매체가 될 수 있다.
- [323] 컴퓨터 프로그램 제품은, 서버 및 클라이언트 장치로 구성되는 시스템에서, 서버의 저장매체 또는 클라이언트 장치의 저장매체를 포함할 수 있다. 또는, 서버 또는 클라이언트 장치와 통신 연결되는 제3 장치(예, 스마트폰)가 존재하는 경우, 컴퓨터 프로그램 제품은 제3 장치의 저장매체를 포함할 수 있다. 또는, 컴퓨터 프로그램 제품은 서버로부터 클라이언트 장치 또는 제3 장치로 전송되거나, 제3 장치로부터 클라이언트 장치로 전송되는 S/W 프로그램 자체를 포함할 수 있다.
- [324] 이 경우, 서버, 클라이언트 장치 및 제3 장치 중 하나가 컴퓨터 프로그램 제품을 실행하여 개시된 실시예들에 따른 방법을 수행할 수 있다. 또는, 서버, 클라이언트 장치 및 제3 장치 중 둘 이상이 컴퓨터 프로그램 제품을 실행하여 개시된 실시예들에 따른 방법을 분산하여 실시할 수 있다.
- [325] 예를 들면, 서버(예로, 클라우드 서버 또는 인공지능 서버 등)가 서버에 저장된 컴퓨터 프로그램 제품을 실행하여, 서버와 통신 연결된 클라이언트 장치가 개시된 실시예들에 따른 방법을 수행하도록 제어할 수 있다.
- [326] 이상에서 실시예들에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속한다.

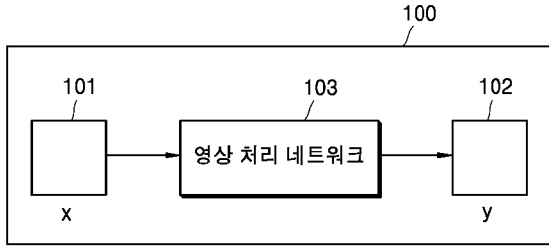
청구범위

- [청구항 1] 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치에 있어서,
 하나 이상의 인스트럭션들을 저장하는 메모리(120); 및
 상기 메모리에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 제1 영상의 제1 특징 정보를 획득하고,
 상기 제1 특징 정보에 제1 처리를 수행함으로써, 출력 영상을 생성하고,
 상기 제1 영상 및 상기 출력 영상에 기초하여, 어텐션 맵(attention map)을 생성하고,
 상기 제1 특징 정보에 제2 처리를 수행함으로써, 제2 특징 정보를 획득하고,
 상기 제1 처리 중에 제3 특징 정보를 획득하고,
 상기 제3 특징 정보에 제3 처리를 수행함으로써, 제4 특징 정보를 획득하고,
 상기 어텐션 맵, 상기 제2 특징 정보 및 상기 제4 특징 정보에 기초하여, 상기 제1 영상의 제1 화질보다 고화질인 제2 화질을 가지는 제2 영상을 생성하는 적어도 하나의 프로세서(110)를 포함하는, 영상 처리 장치.
- [청구항 2] 제1항에 있어서,
 상기 적어도 하나의 프로세서(110)는, 상기 메모리(120)에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써,
 상기 제1 영상 및 상기 출력 영상에 대응하는 레지듀얼 맵(residual map)을 획득하고,
 상기 레지듀얼 맵의 로컬 맵(local map)과 글로벌(global) 분산 값들을 획득하고,
 상기 로컬 맵과 상기 글로벌 분산 값들에 기초하여, 상기 어텐션 맵을 생성하는, 영상 처리 장치.
- [청구항 3] 제2항에 있어서,
 상기 로컬 맵은 상기 레지듀얼 맵의 로컬 분산 값들을 포함하고, 상기 로컬 분산 값들은 제1 크기를 가지는 윈도우를 이용하여 획득되며,
 상기 글로벌 분산 값들은 상기 제1 크기보다 큰 제2 크기를 가지는 패치를 이용하여 획득되는, 영상 처리 장치.
- [청구항 4] 제1항 내지 제3항 중 어느 한 항에 있어서,
 상기 적어도 하나의 프로세서(110)는, 상기 메모리(120)에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써,
 상기 어텐션 맵에 제4 처리를 수행함으로써, 제5 특징 정보를 획득하고,
 상기 제2 특징 정보와 상기 제5 특징 정보에 기초하여, 제6 특징 정보를 획득하고,

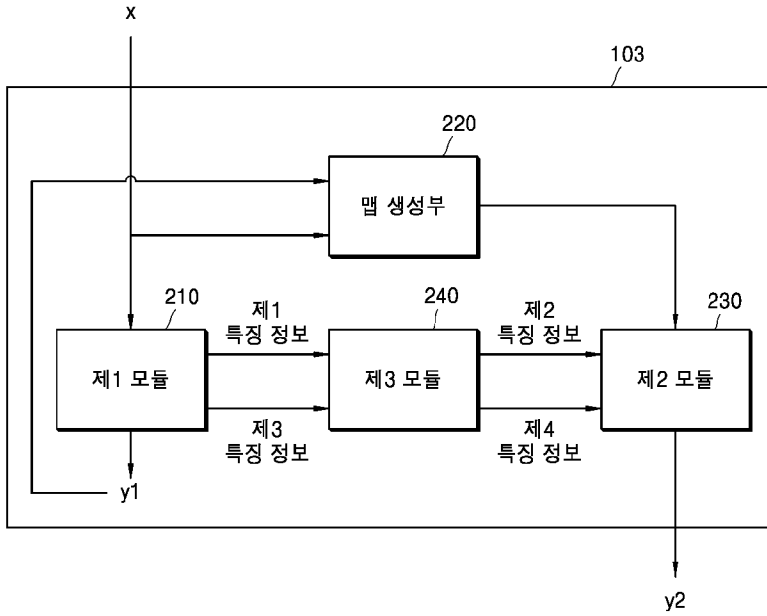
- 상기 제6 특징 정보 및 상기 제4 특징 정보에 기초하여, 상기 제2 영상을 생성하는, 영상 처리 장치.
- [청구항 5] 제1항 내지 제4항 중 어느 한 항에 있어서, 상기 적어도 하나의 프로세서(110)는, 상기 메모리(120)에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 어텐션 맵에 제5 처리를 수행함으로써, 제7 특징 정보를 획득하고, 상기 제2 특징 정보, 제4 특징 정보 및 제7 특징 정보에 기초하여, 상기 제2 영상을 생성하는, 영상 처리 장치.
- [청구항 6] 제1항 내지 제5항 중 어느 한 항에 있어서, 상기 하나 이상의 뉴럴 네트워크들은 하나 이상의 컨볼루션 뉴럴 네트워크들을 포함하고, 상기 적어도 하나의 프로세서(110)는, 상기 메모리(120)에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 하나 이상의 컨볼루션 뉴럴 네트워크들을 이용하여, 상기 제1 영상의 제1 특징 정보를 획득하는, 영상 처리 장치.
- [청구항 7] 제1항 내지 제6항 중 어느 한 항에 있어서, 상기 제1 처리는 상기 제1 특징 정보의 업 샘플링을 포함하고, 상기 제1 영상 및 상기 출력 영상은 크기가 동일한 영상 처리 장치.
- [청구항 8] 제1항 내지 제7항 중 어느 한 항에 있어서, 상기 제2 처리는 하나 이상의 완전 연결 레이어들(fully connected layers)을 포함하는 MLP 모듈에 의해 수행되며, 상기 적어도 하나의 프로세서(110)는, 상기 메모리(120)에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 하나의 완전 연결 레이어에 입력되는 입력 데이터와 상기 완전 연결 레이어에 포함되는 가중치 행렬과의 곱셈 연산 및 상기 완전 연결 레이어에 포함되는 바이어스와의 덧셈 연산을 수행하는, 영상 처리 장치.
- [청구항 9] 제1항 내지 제8항 중 어느 한 항에 있어서, 상기 적어도 하나의 프로세서(110)는, 상기 메모리(120)에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제3 특징 정보와 기 학습된 가중치에 기초하는 컨볼루션 연산을 수행함으로써, 상기 제4 특징 정보를 획득하는, 영상 처리 장치.
- [청구항 10] 제1항 내지 제9항 중 어느 한 항에 있어서, 상기 적어도 하나의 프로세서(110)는, 상기 메모리(120)에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써, 상기 제2 특징 정보에 제1 연산을 수행하여, 제8 특징 정보를 획득하고, 상기 제8 특징 정보 및 상기 제4 특징 정보에 제2 연산을 수행함으로써, 상기 제2 영상을 생성하는, 영상 처리 장치.
- [청구항 11] 제10항에 있어서,

- 상기 제1 연산은 AdaIN 연산을 포함하고,
 상기 제2 연산은 SFT 연산을 포함하는, 영상 처리 장치.
- [청구항 12] 제1항 내지 제11항 중 어느 한 항에 있어서,
 상기 적어도 하나의 프로세서(110)는, 상기 메모리(120)에 저장된 상기 하나 이상의 인스트럭션들을 실행함으로써,
 상기 제1 영상의 분류 정보를 획득하고,
 상기 제1 특징 정보 및 상기 분류 정보에 상기 제1 처리를 수행함으로써, 출력 영상을 생성하는, 영상 처리 장치.
- [청구항 13] 제1항 내지 제12항 중 어느 한 항에 있어서,
 상기 제2 화질의 상기 제2 영상은, 상기 제1 화질의 상기 제1 영상과 비교할 때, 적은 아티팩트 또는 적은 노이즈를 가지는, 영상 처리 장치.
- [청구항 14] 하나 이상의 뉴럴 네트워크들을 이용하여, 영상을 처리하는 영상 처리 장치의 동작 방법에 있어서,
 제1 영상의 제1 특징 정보를 획득하는 단계;
 상기 제1 특징 정보에 제1 처리를 수행함으로써, 출력 영상을 생성하는 단계;
 상기 제1 영상 및 상기 출력 영상에 기초하여, 어텐션 맵을 생성하는 단계;
 상기 제1 특징 정보에 제2 처리를 수행함으로써, 제2 특징 정보를 획득하는 단계;
 상기 제1 처리 중에 제3 특징 정보를 획득하는 단계;
 상기 제3 특징 정보에 제3 처리를 수행함으로써, 제4 특징 정보를 획득하는 단계; 및
 상기 어텐션 맵, 상기 제2 특징 정보 및 상기 제4 특징 정보에 기초하여, 상기 제1 영상의 제1 화질보다 고화질인 제2 화질을 가지는 제2 영상을 생성하는 단계를 포함하는, 영상 처리 장치의 동작 방법.
- [청구항 15] 제14항의 방법을 수행하도록 하는 프로그램이 저장된 컴퓨터로 읽을 수 있는 기록매체.

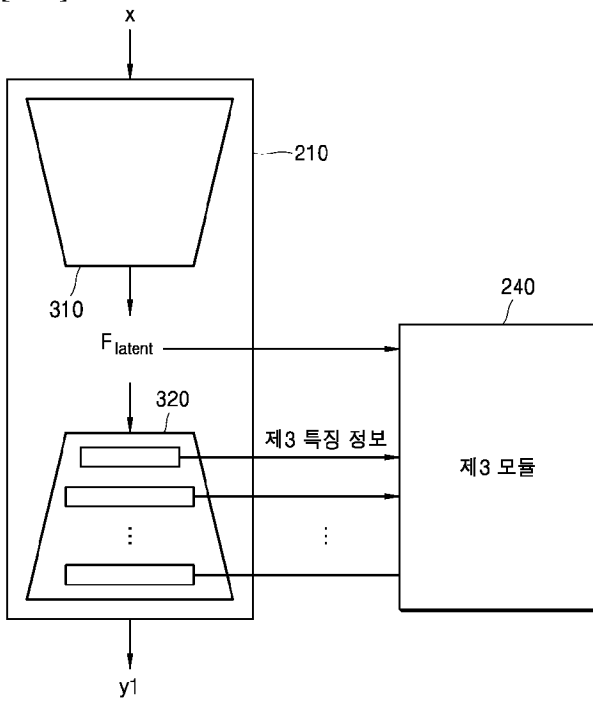
[도1]



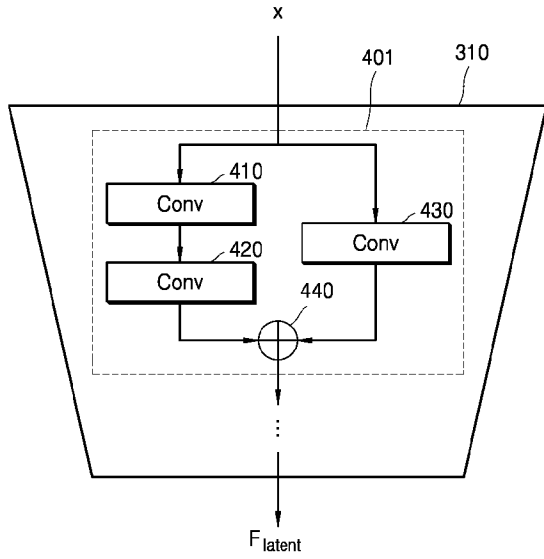
[도2]



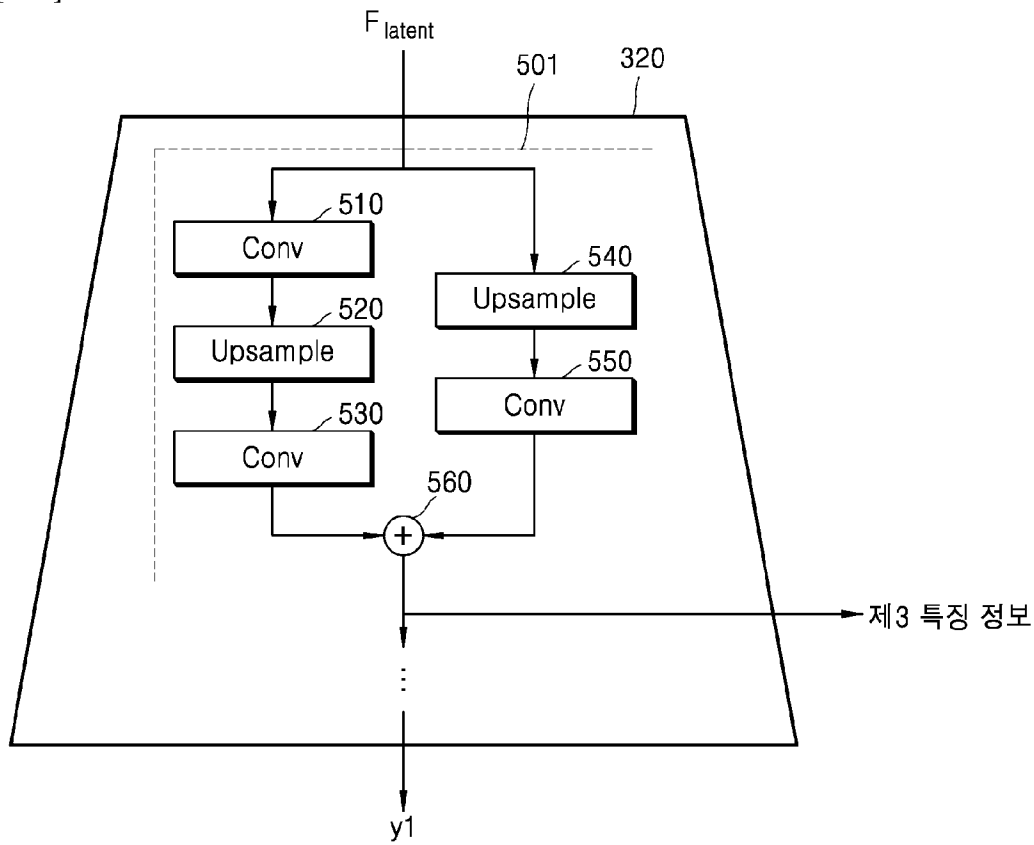
[도3]



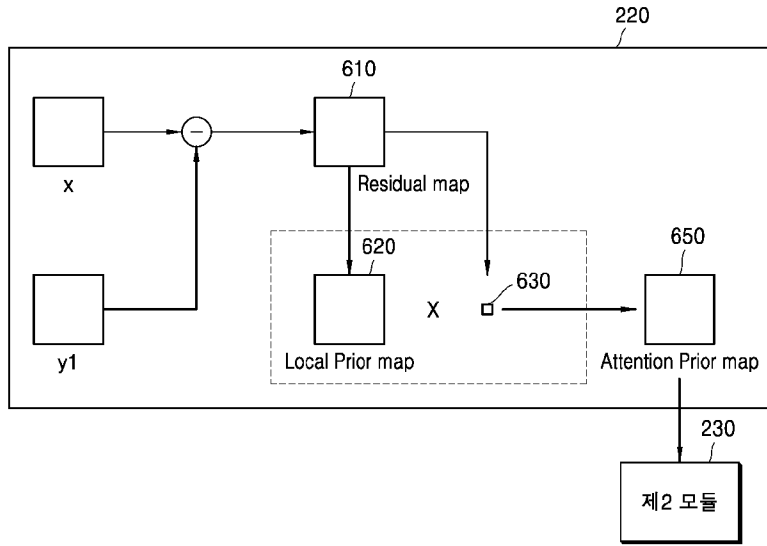
[도4]



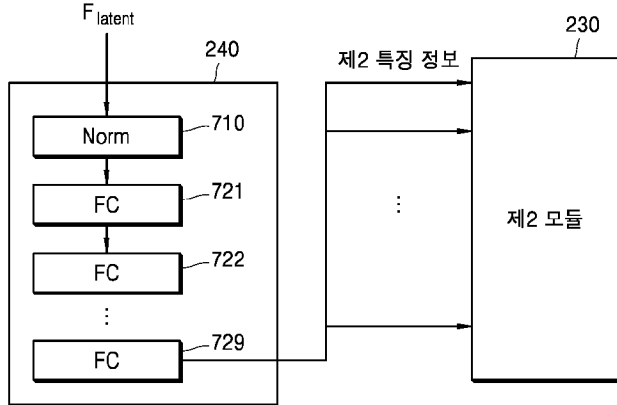
[도5]



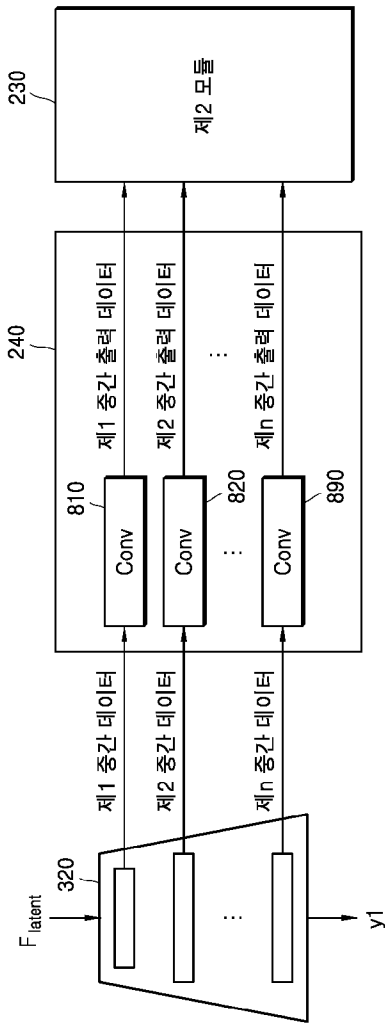
[도6]



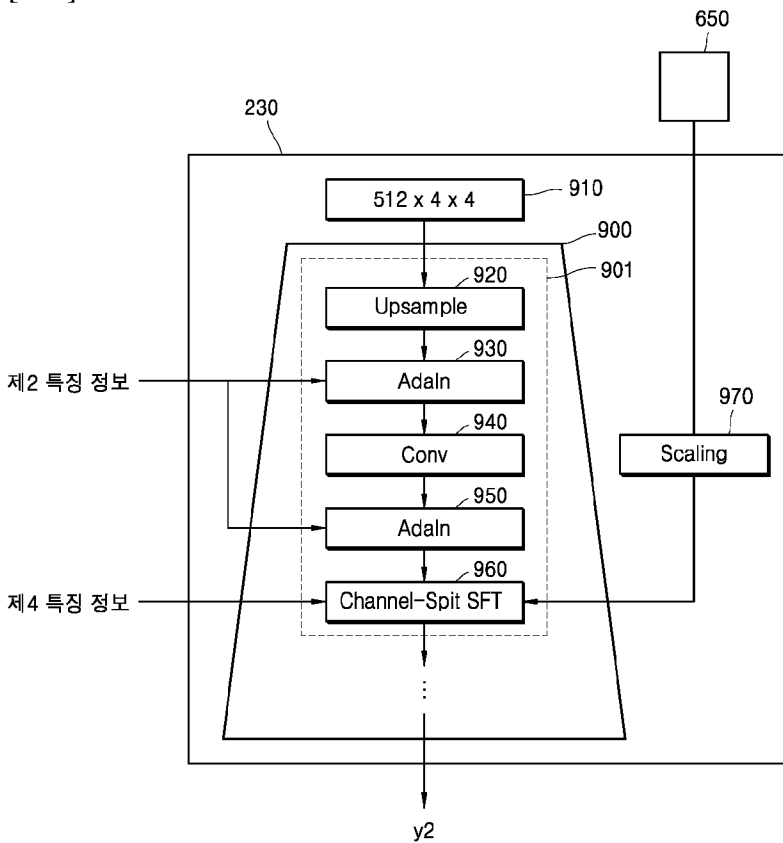
[도7]



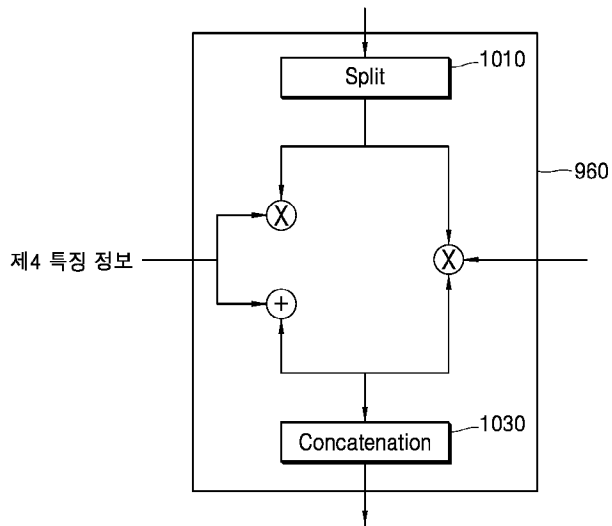
[도8]



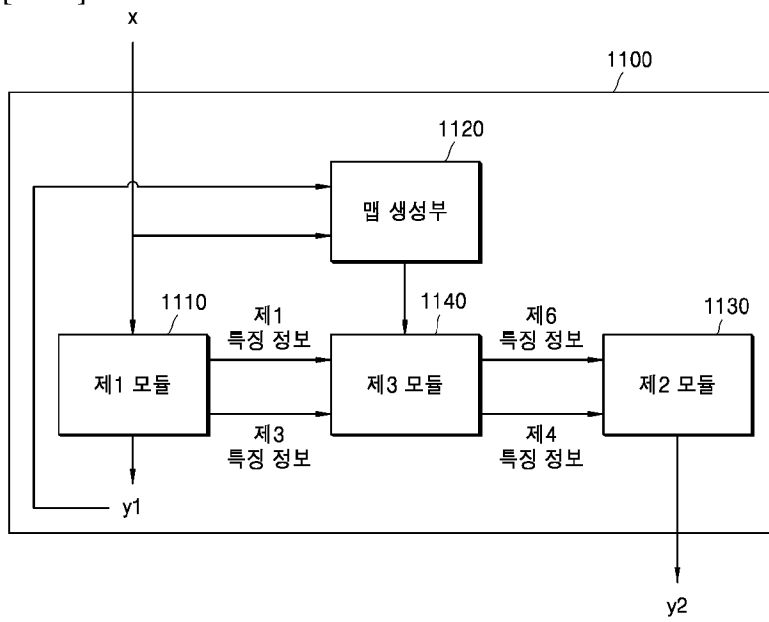
[도9]



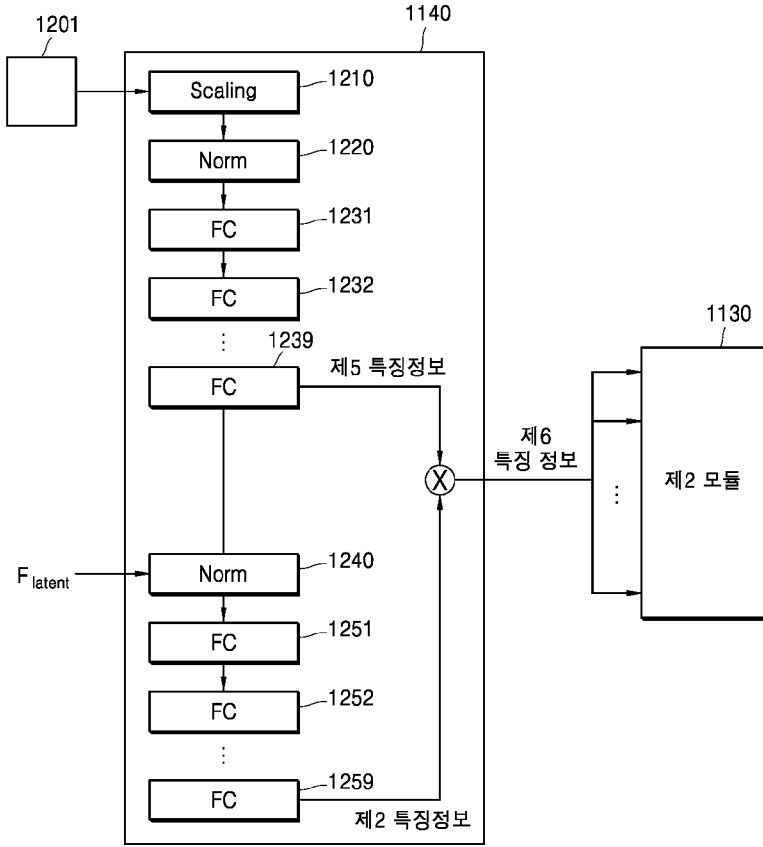
[도10]



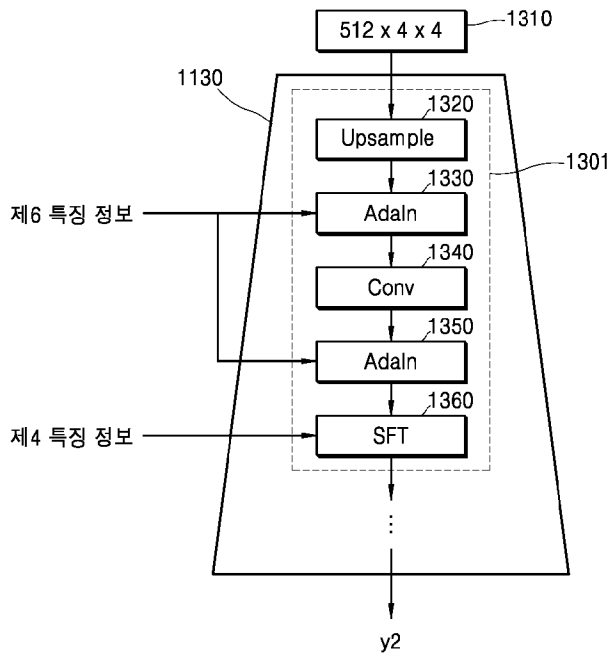
[도11]



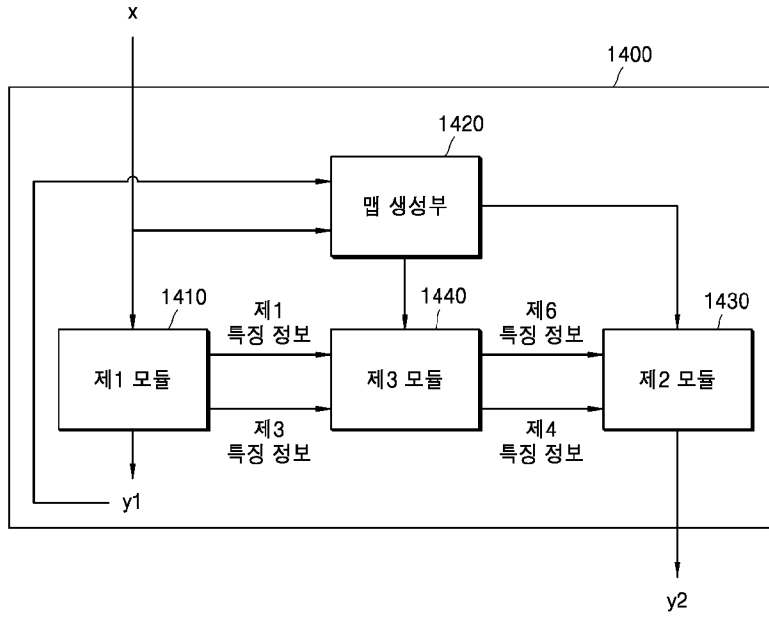
[도 12]



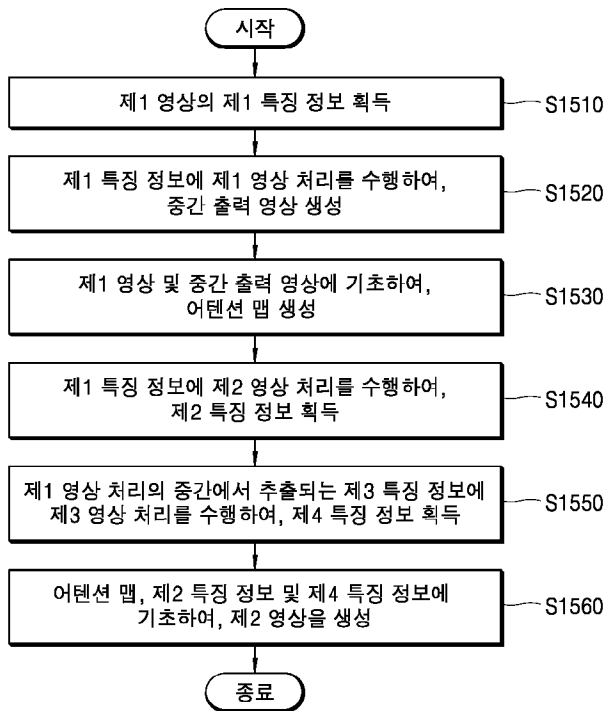
[도 13]



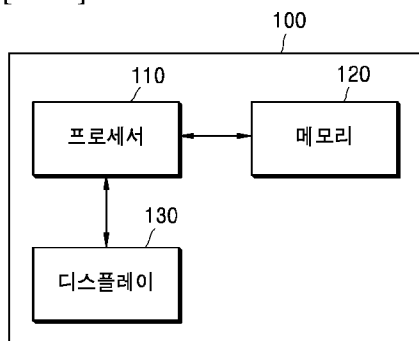
[도14]



[도15]



[도16]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2023/004669**A. CLASSIFICATION OF SUBJECT MATTER****G06T 3/40(2006.01)i; G06T 7/11(2017.01)i; G06N 3/08(2006.01)i; G06T 5/00(2006.01)i; G06T 5/20(2006.01)i**

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06T 3/40(2006.01); G06F 17/15(2006.01); G06N 3/04(2006.01); G06T 5/00(2006.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models: IPC as above

Japanese utility models and applications for utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & keywords: 뉴럴 네트워크(neural network), 특징 정보(feature information), 어텐션 맵(attention map), 레지듀얼 맵(residual map), 가중치(weight), 업샘플링(upsampling), 고화질(high quality)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2019-0113119 A (SAMSUNG ELECTRONICS CO., LTD. et al.) 08 October 2019 (2019-10-08) See paragraphs [0010], [0021], [0072]-[0073], [0087], [0165] and [0169]-[0171]; and figure 11.	1-15
Y	KR 10-2021-0114856 A (SAMSUNG ELECTRONICS CO., LTD.) 24 September 2021 (2021-09-24) See paragraphs [0006], [0053], [0060]-[0061], [0079] and [0081]; and figures 5 and 7.	1-15
Y	HUANG, Xun et al. Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization. 2017 IEEE International Conference on Computer Vision (ICCV). pp. 1510-1519, 22-29 October 2017. See page 1513; and figure 2.	10-11
A	KR 10-2020-0084434 A (POSTECH RESEARCH AND BUSINESS DEVELOPMENT FOUNDATION) 13 July 2020 (2020-07-13) See paragraphs [0014]-[0015].	1-15
A	KR 10-2296644 B1 (INDUSTRIAL COOPERATION FOUNDATION JEONBUK NATIONAL UNIVERSITY) 01 September 2021 (2021-09-01) See paragraph [0011].	1-15

 Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

06 July 2023

Date of mailing of the international search report

07 July 2023

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon Building 4, 189 Cheongsang-ro, Seo-gu, Daejeon 35208

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No. PCT/KR2023/004669

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
KR 10-2019-0113119 A	08 October 2019	None	
KR 10-2021-0114856 A	24 September 2021	CN 113379613 A	10 September 2021
		US 11508037 B2	22 November 2022
		US 2021-0287342 A1	16 September 2021
		US 2023-043310 A1	09 February 2023
KR 10-2020-0084434 A	13 July 2020	KR 10-2169242 B1	23 October 2020
KR 10-2296644 B1	01 September 2021	None	

A. 발명이 속하는 기술분류(국제특허분류(IPC)) G06T 3/40(2006.01)i; G06T 7/11(2017.01)i; G06N 3/08(2006.01)i; G06T 5/00(2006.01)i; G06T 5/20(2006.01)i		
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) G06T 3/40(2006.01); G06F 17/15(2006.01); G06N 3/04(2006.01); G06T 5/00(2006.01) 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) eKOMPASS(특허청 내부 검색시스템) & 키워드: 뉴럴 네트워크(neural network), 특징 정보(feature information), 어텐션 맵(attention map), 레지듀얼 맵(residual map), 가중치(weight), 업샘플링(upsampling), 고화질(high quality)		
C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-2019-0113119 A (삼성전자주식회사 등) 2019.10.08 단락 [0010], [0021], [0072]-[0073], [0087], [0165], [0169]-[0171]; 및 도면 11	1-15
Y	KR 10-2021-0114856 A (삼성전자주식회사) 2021.09.24 단락 [0006], [0053], [0060]-[0061], [0079], [0081]; 및 도면 5, 7	1-15
Y	XUN HUANG 등, `Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization`, 2017 IEEE International Conference on Computer Vision (ICCV), pp. 1510-1519, 2017.10.22-29 페이지 1513; 및 도면 2	10-11
A	KR 10-2020-0084434 A (포항공과대학교 산학협력단) 2020.07.13 단락 [0014]-[0015]	1-15
A	KR 10-2296644 B1 (전북대학교산학협력단) 2021.09.01 단락 [0011]	1-15
<input type="checkbox"/> 추가 문헌이 C(계속)에 기재되어 있습니다. <input checked="" type="checkbox"/> 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: "A" 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 "D" 본 국제출원에서 출원인이 인용한 문헌 "E" 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 "L" 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 "O" 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 "P" 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌 "T" 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌 "X" 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다. "Y" 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다. "&" 동일한 대응특허문헌에 속하는 문헌		
국제조사의 실제 완료일	국제조사보고서 발송일	
2023년07월06일(06.07.2023)	2023년07월07일(07.07.2023)	
ISA/KR의 명칭 및 우편주소	심사관	
대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사)	변성철	
팩스 번호 +82-42-481-8578	전화번호 +82-42-481-8262	

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2019-0113119 A	2019/10/08	없음	
KR 10-2021-0114856 A	2021/09/24	CN 113379613 A	2021/09/10
		US 11508037 B2	2022/11/22
		US 2021-0287342 A1	2021/09/16
		US 2023-043310 A1	2023/02/09
KR 10-2020-0084434 A	2020/07/13	KR 10-2169242 B1	2020/10/23
KR 10-2296644 B1	2021/09/01	없음	