(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2024 년 8 월 2 일 (02.08.2024)

WIPO | PCT

(10) 국제공개번호
WO 2024/158073 A1

(51) 국제특허분류:
G06V 20/70 (2022.01) *G06V 10/56* (2022.01)
G06V 20/40 (2022.01) *G06V 10/54* (2022.01)
G06V 10/774 (2022.01) *G06V 40/20* (2022.01)
G06V 10/74 (2022.01) *G06T 3/40* (2006.01)

(21) 국제출원번호: PCT/KR2023/001254

(22) 국제출원일: 2023 년 1 월 27 일 (27.01.2023)

(25) 출원언어: 한국어

(26) 공개언어: 한국어

(71) 출원인: 라쿠텐 심포니 주식회사 (**RAKUTEN SYMPHONY, INC.**) [JP/JP]; 158-0094 도쿄 세타가야쿠 다마가와 1-14-1, Seoul (JP).

(72) 발명자: 이병수 (**LEE, Byungsu**); 06611 서울특별시 서초구 사평대로 368, 10-11층, Seoul (KR).

(74) 대리인: 장덕순 등 (**CHANG, Duck Soon** et al.); 03170 서울특별시 종로구 사직로8길 39 세양빌딩 김.장법률사무소, Seoul (KR).

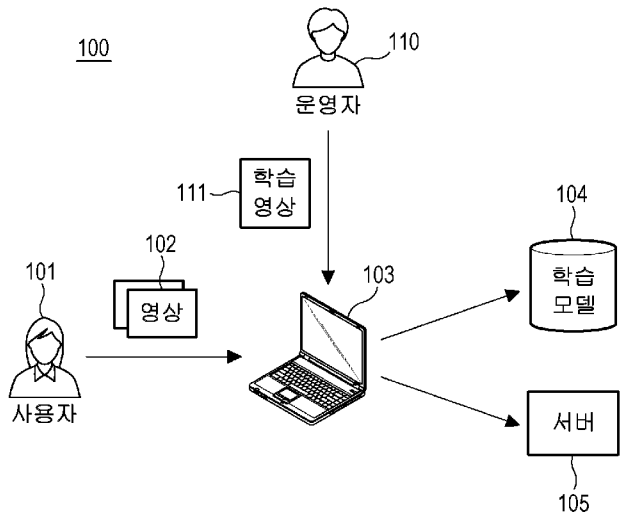
(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, 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))

(54) Title: DETERMINATION OF SIMILARITY BETWEEN IMAGES USING LEARNING MODEL

(54) 발명의 명칭: 학습 모델을 이용한 영상들 간의 유사도 판단



101 ... User
102 ... Image
104 ... Learning model
105 ... Server
110 ... Operator
111 ... Training image

(57) Abstract: An electronic apparatus according to various embodiments of the present disclosure comprises: one or more processors; and one or more memories that store instructions for, when executed by the processors, enabling the processors to perform operations, wherein when the instructions are executed by the processors, the processors may be configured to train a learning model to: sample, into a plurality of frames, a training image comprising a sequence of image frames; detect one or more objects appearing in each of the plurality of frames and generate object images corresponding to the one or more objects; and label the object images and classify the object images into one or more objects.

(57) 요약서: 본 개시의 다양한 실시예에 따른 전자 장치는 하나 이상의 프로세서; 및 프로세서에 의한 실행 시 프로세서가 연산을 수행하도록 하는 명령어들이 저장된 하나 이상의 메모리를 포함하고, 프로세서에 의해 상기 명령어들이 실행될 시, 프로세서는, 이미지 프레임의 시퀀스로 구성된 학습 영상을 복수의 프레임으로 샘플링하고, 복수의 프레임 각각에 나타난 하나 이상의 객체를 검출하여 상기 하나 이상의 객체에 대응하는 객체 이미지들을 생성하고, 객체 이미지들을 레이블링(labeling)하여 하나 이상의 객체 각각으로 분류되도록 학습 모델을 훈련하도록 구성될 수 있다.

WO 2024/158073 A1

명세서

발명의 명칭: 학습 모델을 이용한 영상들 간의 유사도 판단

기술분야

[1] 본 개시는 학습 모델을 이용하여 영상들 간의 유사도를 판단하기 위한 기술에 관한 것이다.

배경기술

[2] 웹상에서 다양한 컨텐츠가 자유롭게 배포 및 유통됨에 따라 컨텐츠에 대한 저작물 보호가 중요한 문제가 되었다. 컨텐츠는 복제, 유통 및 배포가 매우 용이하며 이와 같이 배포된 컨텐츠는 진본과 실질적으로 동일하므로 컨텐츠에 의한 저작권 침해는 저작권자의 권익을 심각하게 훼손하게 된다.

[3] 컨텐츠에 대한 저작권 보호를 위한 방안은 저작물의 복제, 유통, 배포가 어렵게 하는 사전적 조치와, 불법으로 복제, 유통, 배포된 저작물에 대해 검출하고 단속하는 사후적 조치로 나눌 수 있다. 사전적 조치는 예컨대 복제가 불가능하거나 복제 회수를 제한하기 위한 워터마킹 기술 등과 같이 기술적 측면에서 개발되는 방식으로서 많은 발전이 이루어져 왔다. 그러나 사전적 조치에 의한 방식은 제한을 해제하는 기술의 개발에 의해 대부분 무력화되고 있으며, 또한 저작물의 직접적 침해에 해당되지 않는 복제 등에 대해서도 구분을 두지 않고 금지하는 효과로 인해 현실적으로 적용이 부적절한 경우가 많다. 따라서 사후적 조치로서 저작권의 침해를 하고 있는 행위에 대해 검출 및 적발이 지속적으로 병행되어야 한다.

[4] 그런데 현재의 컨텐츠 저작물에 대한 침해 검출은 저작권자 스스로 또는 저작권을 위탁 받아 관리하는 위탁기관이 개별적으로 웹사이트들을 접속하여 검출하는 수작업에 의존하고 있다. 이러한 방식은 무수히 많은 웹사이트들에 대한 검출 및 적발을 매우 어렵게 하며, 기 검출된 웹사이트에서도 새로이 저작권 침해 사례가 추가되는 경우에 재접속 및 재검출을 하지 않는 한 지속적인 감시가 어렵게 된다. 나아가, 저작권자가 소자본 개인인 경우 위탁 기관에 자신의 저작물에 대한 권리 보호를 위탁시키는 것조차도 쉽지 않은 경우가 많다.

발명의 상세한 설명

기술적 과제

[5] 본 개시의 일 실시예를 통해 해결하고자 하는 기술적 과제는 학습 영상을 통해 영상들 간의 유사도를 판단할 수 있는 학습 모델을 훈련시키는 기술을 제공하는 것이다.

[6] 본 개시의 일 실시예를 통해 해결하고자 하는 기술적 과제는 학습 모델을 이용하여 영상에 나타난 객체들의 유사도를 판단하여 영상들 간의 유사도를 판단하는 기술을 제공하는 것이다.

과제 해결 수단

[7] 본 개시의 한 측면으로서, 전자 장치가 제안될 수 있다. 본 개시의 한 측면에 따른 전자 장치는, 하나 이상의 프로세서, 및 하나 이상의 메모리를 포함할 수 있다. 상기 프로세서는 이미지 프레임의 시퀀스로 구성된 학습 영상을 복수의 프레임으로 샘플링하고, 상기 복수의 프레임 각각에 나타난 하나 이상의 객체를 검출하여 상기 하나 이상의 객체에 대응하는 복수의 객체 이미지를 생성하고, 상기 복수의 객체 이미지를 레이블링하여 상기 하나 이상의 객체 각각으로 분류되도록 학습 모델을 훈련하도록 구성될 수 있다.

[8] 본 개시의 한 측면으로서, 전자 장치가 제안될 수 있다. 본 개시의 한 측면에 따른 전자 장치는, 하나 이상의 프로세서, 네트워크 인터페이스, 및 하나 이상의 메모리를 포함할 수 있다. 상기 메모리는 학습 모델을 저장할 수 있고, 학습 모델은 학습 영상에 포함된 객체와 레이블의 상관 관계가 학습된 모델일 수 있다. 상기 프로세서는 상기 메모리로부터 상기 학습 모델을 로드하고, 입력된 타겟 영상으로부터 획득한 복수의 이미지를 상기 학습 모델에 입력하여 하나 이상의 레이블 각각에 대한 유사도를 획득하고, 상기 하나 이상의 레이블 각각에 대한 유사도에 상기 하나 이상의 레이블 각각의 가중치를 곱하여 상기 타겟 영상의 상기 학습 영상에 대한 총 유사도를 산출하도록 구성될 수 있다.

[9] 본 개시의 한 측면으로서, 방법이 제안될 수 있다. 본 개시의 한 측면에 따른 방법은, 입력된 학습 영상을 복수의 프레임으로 샘플링하는 단계, 상기 복수의 프레임 각각에 나타난 하나 이상의 객체를 검출하여 상기 하나 이상의 객체에 대응하는 복수의 객체 이미지를 생성하는 단계, 상기 복수의 객체 이미지를 레이블링하여 상기 하나 이상의 객체 각각으로 분류되도록 학습 모델을 훈련시키는 단계, 입력된 타겟 영상으로부터 획득한 복수의 이미지를 상기 학습 모델에 입력하여 하나 이상의 레이블 각각에 대한 유사도를 획득하는 단계, 및 상기 하나 이상의 레이블 각각에 대한 유사도에 상기 하나 이상의 레이블 각각의 가중치를 곱하여, 상기 타겟 영상의 상기 학습 영상에 대한 총 유사도를 산출하는 단계를 포함할 수 있다.

발명의 효과

[10] 본 개시의 다양한 실시예에 따르면, 학습 영상을 통해 학습 모델이 이미지를 분류하도록 훈련시킬 수 있다.

[11] 본 개시의 다양한 실시예에 따르면, 학습 영상의 이미지를 다양하게 변형시켜서 학습 모델에 훈련시킴으로써 학습 모델의 정확도를 향상시킬 수 있다.

[12] 본 개시의 다양한 실시예에 따르면, 학습 모델에 미리 학습된 영상에 대한 입력된 영상의 유사도를 판단할 수 있다. 따라서, 입력된 영상이 특정 영상과 유사한지 여부를 판단할 수 있다.

[13] 본 개시의 다양한 실시예에 따르면, 학습 모델에 저작권이 있는 영상을 미리 입력하여 지도 훈련시킴으로써, 특정한 영상을 업로드하기 전에 해당 영상이

저작권을 침해하는 지 여부를 미리 판단할 수 있다.

[14] 본 개시의 다양한 실시예에 따르면, 저작권 침해가 문제될 수 있는 영상에 대해서는 업로드를 방지함으로써 저작권 침해를 방지할 수 있다.

[15] 본 개시의 다양한 실시예에 따르면, 영상의 유사 여부를 판단하여 이를 기초로 저작권 침해 여부를 판단함으로써 저작권 침해 사례 검출을 간편하게 할 수 있다.

도면의 간단한 설명

[16] 도 1은 본 개시의 일 실시예에 따른 영상의 유사 여부 판단을 위한 공유 플랫폼 시스템을 예시적으로 도시한 도면이다.

[17] 도 2는 본 개시의 일 실시예에 따른 전자 장치의 블록도이다.

[18] 도 3은 본 개시의 일 실시예에 따른 학습 영상을 샘플링하는 모습을 예시적으로 도시한 도면이다.

[19] 도 4a, 도 4b는 본 개시의 일 실시예에 따른 학습 모델 훈련 모습을 예시적으로 도시한 도면이다.

[20] 도 5는 본 개시의 일 실시예에 따른 이미지 그룹의 생성 방법을 도시한 도면이다.

[21] 도 6은 본 개시의 일 실시예에 따른 유사도 비교 방법을 도시한 도면이다.

[22] 도 7a, 도 7b는 본 개시의 일 실시예에 따른 유사도 판단 방법을 예시적으로 도시한 도면이다.

[23] 도 8은 본 개시의 일 실시예에 따른 입력 영상이 학습 영상과 유사한지를 판단하는 방법에 대한 흐름도이다.

발명의 실시를 위한 형태

[24] 본 개시에 기재된 다양한 실시예들은, 본 개시의 기술적 사상을 명확히 설명하기 위한 목적으로 예시된 것이며, 이를 특정한 실시 형태로 한정하려는 것이 아니다. 본 개시의 기술적 사상은, 본 개시에 기재된 각 실시예의 다양한 변경(modifications), 균등물(equivalents), 대체물(alternatives) 및 각 실시예의 전부 또는 일부로부터 선택적으로 조합된 실시예를 포함한다. 또한 본 개시의 기술적 사상의 권리 범위는 이하에 제시되는 다양한 실시예들이나 이에 대한 구체적 설명으로 한정되지 않는다.

[25] 개시에 사용되는 모든 기술적 용어들 및 과학적 용어들은, 달리 정의되지 않는 한, 본 개시가 속하는 기술 분야에서 통상의 지식을 가진 자에게 일반적으로 이해되는 의미를 갖는다. 본 개시에 사용되는 모든 용어들은 본 개시를 더욱 명확히 설명하기 위한 목적으로 선택된 것이며 본 개시에 따른 권리범위를 제한하기 위해 선택된 것이 아니다.

[26] 본 개시에서 사용되는 "포함하는", "구비하는", "갖는" 등과 같은 표현은, 해당 표현이 포함되는 어구 또는 문장에서 달리 언급되지 않는 한, 다른 실시예를 포함할 가능성을 내포하는 개방형 용어(open-ended terms)로 이해되어야 한다.

[27] 본 개시에서 기술된 단수형의 표현은 달리 언급하지 않는 한 복수형 의 의미를 포함할 수 있으며, 이는 청구의 범위에 기재된 단수형의 표현에도 마찬가지로 적용된다.

[28] 본 개시에서 사용되는 "제1", "제2" 등의 표현들은 복수의 구성요소들을 상호 구분하기 위해 사용되며, 해당 구성요소들의 순서 또는 중요도를 한정하는 것은 아니다.

[29] 본 개시에서 사용되는 "~에 기초하여"라는 표현은, 해당 표현이 포함되는 어구 또는 문장에서 기술되는, 결정, 판단의 행위 또는 동작에 영향을 주는 하나 이상의 인자를 기술하는데 사용되며, 이 표현은 결정, 판단의 행위 또는 동작에 영향을 주는 추가적인 인자를 배제하지 않는다.

[30] 본 개시에서, 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 경우, 상기 어떤 구성요소가 상기 다른 구성요소에 직접적으로 연결될 수 있거나 접속될 수 있는 것으로, 또는 새로운 다른 구성요소를 매개로 하여 연결될 수 있거나 접속될 수 있는 것으로 이해되어야 한다.

[31] 본 개시에서 사용된 표현 "~하도록 구성된(configured to)"은 문맥에 따라, "~하도록 설정된", "~하는 능력을 가지는", "~하도록 변경된", "~하도록 만들어진", "~를 할 수 있는" 등의 의미를 가질 수 있다. 이 표현은, "하드웨어적으로 특별히 설계된"의 의미로 제한되지 않으며, 예를 들어 특정 동작을 수행하도록 구성된 프로세서란, 소프트웨어를 실행함으로써 그 특정 동작을 수행할 수 있는 범용 프로세서(generic purpose processor)를 의미하거나, 그 특정 동작을 수행하도록 프로그래밍을 통해 구조화된 특수 목적 컴퓨터(special purpose computer)를 의미할 수 있다.

[32] 이하, 첨부한 도면들을 참조하여, 본 개시의 실시예들을 설명한다. 첨부된 도면에서, 동일하거나 대응하는 구성요소에는 동일한 참조번호가 부여되어 있다. 또한, 이하의 실시예들의 설명에 있어서, 동일하거나 대응하는 구성요소를 중복하여 기술하는 것이 생략될 수 있다. 그러나, 구성요소에 관한 기술이 생략되어도, 그러한 구성요소가 어떤 실시예에 포함되지 않는 것으로 의도되지는 않는다.

[33] 도 1은 본 개시의 일 실시예에 따른 영상의 유사 여부 판단을 위한 공유 플랫폼 시스템을 도시한 도면이다. 공유 플랫폼 시스템(100)은 전자 장치(103), 학습 모델(104), 서버(105)를 포함할 수 있다.

[34] 일 실시예에서 사용자(101)는 사용자 장치(미도시)를 사용하여 컨텐츠를 생성할 수 있다. 컨텐츠는 영상, 이미지, 음악 등을 포함할 수 있다. 영상 파일이란, 움직이는 영상물을 총칭하는 것으로서, 예를 들어, avi, mp4 또는 mkv의 확장자를 갖는 파일을 의미할 수 있다.

[35] 일 실시예에서, 사용자(101)는 영상(102)을 다른 사용자들과 공유하기 위해 공유 플랫폼 시스템(100)에 업로드할 수 있다. 예를 들어, 제1 사용자는

영상(102)을 공유 플랫폼 시스템(100)에 업로드할 수 있고, 제2 사용자는 제1 사용자가 입력한 영상(102)을 공유 플랫폼 시스템(100)에서 열람할 수도 있다.

[36] 일 실시예에서, 전자 장치(103)는 입력된 영상(102)을 서버(105)에 업로드할 수 있고, 서버(105)에 업로드하기 전 다양한 처리를 할 수도 있다. 복수의 구성 요소가 실제 물리적 환경에서는 서로 통합되는 형태로 구현될 수 있고, 필요에 따라 일부 구성 요소가 추가되거나 삭제될 수 있다.

[37] 본 개시의 전자 장치(103)는 다양한 형태의 장치가 될 수 있다. 예를 들어, 전자 장치(103)는 컴퓨팅 장치 또는 휴대용 통신 장치일 수 있다. 전자 장치(103)는 사용자(101)와 통신 연결되어 다양한 데이터를 송수신할 수 있다. 전자 장치(103)와 사용자(101)는 통신시 근거리 통신망(Local Area Network; LAN), 광역 통신망(Wide Area Network; WAN) 또는 부가가치 통신망(Value Added Network; VAN) 등과 같은 유선 네트워크나, 이동 통신망(mobile radio communication network), 위성 통신망, 블루투스(Bluetooth), Wibro(Wireless Broadband Internet), HSDPA(High Speed Downlink Packet Access) 등과 같은 모든 종류의 무선 네트워크를 통해 실행될 수 있다.

[38] 영상(102)이 다른 사람의 저작권과 관련되어 저작권을 침해할 가능성이 있는 경우, 전자 장치(103)는 영상(102)의 업로드를 금지할 수 있다. 예를 들어, 전자 장치(103)는 영상(102)을 수신하여 서버(105) 또는 다른 저장 장치(미도시)에 업로드하기 전에 다른 영상과의 유사 여부를 학습 모델(104)을 통해 확인할 수 있다. 입력되는 영상은 비교되는 다른 영상과 구분하여, 입력 영상 또는 타겟 영상 등으로 지칭될 수 있다.

[39] 학습 모델(104)은 전자 장치(103) 내에 통합되어 구현될 수도 있고, 별개로 분리되어 통신할 수도 있다. 전자 장치(103)는 영상(102)이 입력 되기 전에 학습 모델(104)을 미리 훈련(학습)시킬 수 있다. 학습 모델(104)은 이미지를 분류하기 위한 모델로, 다수의 이미지를 이용하여 학습될 수 있다. 이하에서는 설명의 편의를 위해 학습 모델(104)이 전자 장치(103)와 별도로 구현된 것으로 가정한다.

[40] 전자 장치(103)는 추후 명세서의 기재에 의해 구체화될 학습 영상(111)을 운영자(110)로부터 수신할 수 있다. 또는, 전자 장치(103)는 학습 영상(111)을 네트워크, 서버 등으로부터 다운로드 받을 수 있다. 학습 영상(111)은 시계열적으로 연속하는 다수의 프레임을 포함할 수 있다.

[41] 수신된 학습 영상(111)을 기초로 전자 장치(103)는 학습 모델(104)을 훈련시킬 수 있다. 전자 장치(103)는 학습 영상(111)을 복수의 프레임으로 샘플링할 수 있고, 샘플링된 복수의 프레임에 나타난 하나 이상의 객체를 검출하여 복수의 객체 이미지를 생성할 수 있다.

[42] 전자 장치(103)는 생성한 복수의 객체 이미지를 레이블링하여 학습 모델(104)을 훈련시킬 수 있다. 학습 모델(104)은 객체의 이미지와 객체의 레이블을 연관지어 상관 관계가 훈련될 수 있다. 학습 모델(104)은 훈련 후에 특정한 이미지에 포함된 객체를 보고, 정해진 레이블 중 하나로 분류할 수 있다.

[43] 학습이 된 학습 모델(104)은 영상 사이의 유사도를 판단할 수 있다. 전자 장치(103)는 학습 모델(104)이 판단한 영상 사이의 유사도를 기초로 저작권 침해 여부를 판단할 수 있다. 예를 들어, 학습 모델(104)은 사용자(101)가 업로드한 영상(102)의 객체에 대한 이미지들이 학습 영상(111) 속 객체의 이미지들과 얼마나 유사한지를 판단하여 영상(102)과 학습 영상(111)의 총 유사도를 산출할 수 있다. 전자 장치(103)는 산출된 총 유사도에 기초하여 영상(102)의 업로드 여부를 결정할 수 있다.

[44] 도 2는 본 개시의 일 실시예에 따른 전자 장치의 블록도이다. 도 2에서는 설명의 편의를 위해 전자 장치의 참조번호를 '200'으로 기재하였지만 전자 장치(103)과 전자 장치(200)은 서로 다른 장치 구성이 아니며 동일하거나 대응하는 장치일 수 있다. 도 2를 참조하면, 본 개시의 일 실시예에 따른 전자 장치(200)는 사용자에 의해 입력되는 영상(102)의 학습 영상(111)에 대한 유사 여부를 모니터링하여 저작권 침해 판단을 할 수 있는 서비스를 제공할 수 있다. 전자 장치(200)는 적어도 하나의 프로세서(210), 적어도 하나의 메모리(220), 및 적어도 하나의 네트워크 인터페이스(230)를 포함할 수 있다.

[45] 프로세서(210)는 메모리(220) 및/또는 네트워크 인터페이스(230)를 제어하며, 메모리(220)에 저장된 명령어를 실행하여 본 개시의 설명, 기능, 절차, 제안, 방법 및/또는 동작 순서도를 구현하도록 구성될 수 있다. 예를 들어, 프로세서(210)는 네트워크 인터페이스(230)를 통해 신호 및/또는 데이터를 수신하고, 신호 및/또는 데이터에 포함된 정보를 메모리(220)에 저장할 수 있다.

[46] 메모리(220)는 프로세서(210)와 연결될 수 있고, 프로세서(210)의 동작과 관련한 다양한 정보를 저장할 수 있다. 예를 들어, 메모리(220)는 프로세서(210)에 의해 제어되는 프로세스들 중 일부 또는 전부를 수행하거나, 본 개시의 설명, 기능, 절차, 제안, 방법 및/또는 동작 순서도들을 수행하기 위한 명령어들(instructions)을 포함하는 소프트웨어 코드를 저장할 수 있다. 메모리(220)는 비일시적인 컴퓨터 판독가능 매체, 예컨대 고속 랜덤 액세스 메모리 및/또는 비휘발성 컴퓨터 판독가능 저장 매체(예컨대, 하나 이상의 디스크 저장 장치, 플래쉬 메모리 장치, 또는 기타 비휘발성 솔리드 스테이트 메모리 장치)를 포함할 수 있다. 학습 모델(104)이 전자 장치(200) 내에 구현되는 경우 학습 모델(104)은 메모리(220) 내에 저장되어 구현될 수 있다.

[47] 네트워크 인터페이스(230)는 프로세서(210)와 연결될 수 있고, 유/무선 신호나 데이터를 전송 및/또는 수신할 수 있다. 예를 들어, 네트워크 인터페이스(230)는 유/무선 통신망을 통해 사용자 단말 및/또는 서버, 데이터베이스와 연결될 수 있다.

[48] 여기서, 무선 통신망은 이동 통신망, 무선 LAN, 근거리 무선 통신망 등을 포함할 수 있다. 예를 들어, 무선 통신망은 LTE, LTE-A(LTE Advance), CDMA(code division multiple access), WCDMA(wideband CDMA), UMTS(universal mobile telecommunications system), WiBro(Wireless Broadband), 및 GSM(Global

System for Mobile Communications) 등 중 적어도 하나를 사용하는 셀룰러 통신을 포함할 수 있다. 예를 들어, 무선 통신망은 WiFi(wireless fidelity), 블루투스, 블루투스 저전력(BLE), 지그비 (Zigbee), NFC(near field communication), 및 라디오 프리퀀시(RF) 중 적어도 하나를 포함할 수 있다.

[49] 여기서, 유선 통신망은 USB(Universal Serial Bus), USART(Universal Synchronous/Asynchronous Receiver Transmitter), 및 이더넷(ethernet) 중 적어도 하나를 포함할 수 있다.

[50] 적어도 하나의 네트워크 인터페이스(230) 각각은 상술한 유/무선 통신망에 대응될 수 있다.

[51] 이하에서 개시되는 전자 장치(200)에 포함되는 기능 단위들은 상술한 프로세서(210), 메모리(220) 및 네트워크 인터페이스(230)를 포함하는 하드웨어나 명령어들을 구현하기 위한 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

[52] 도 3은 일 실시예에 따른 프로세서가 학습 영상을 샘플링(sampling)하는 모습을 도시한 도면이다. 도 3을 참조하면 프로세서(210)는 수신한 영상(102)을 정해진 샘플링 주기로 샘플링할 수 있다. 샘플링은, 영상으로부터 프레임을 획득하는 것을 의미할 수 있다.

[53] 일 실시예에서, 샘플링의 주기는 미리 메모리(220)에 저장되어 있을 수 있고, 운영자(110)에 의해 입력을 받아 결정될 수 있다. 운영자(110)의 입력에 의해 샘플링 시점이 결정되는 경우, 샘플링의 주기는 주기적으로 이루어질 수 있고, 입력에 의해 비주기적으로 특정 시점의 프레임이 샘플링될 수 있다. 전체 프레임이 아닌 샘플링 된 프레임이 대해서만 처리를 함으로써 컴퓨팅 로드를 감소시킬 수 있다.

[54] 일 실시예에서, 프로세서(210)는 영상(102)에 있는 객체의 움직임에 대한 움직임 벡터를 취득하고 정해진 시간 동안 움직임 벡터의 크기를 측정할 수 있다. 정해진 시간 동안 움직임 벡터의 크기가 일정한 값 이상이면 객체의 움직임이 크다는 것을 의미할 수 있다. 예를 들어 움직임 벡터는 두드러진 움직임이 있는 물체를 배경에서 분리하기 위한 방법으로 가우시안 혼합 모델을 통해 생성된 배경 모델에 적응적 차영상을 통한 방법으로 구해질 수 있다. 정해진 시간 동안 움직임 벡터의 변화량이 제1 값 이상이면, 영상에서 객체의 중요성이 높다는 것을 의미할 수 있다. 프로세서(210)는 객체의 움직임에 대한 움직임 벡터의 변화량이 제1 값 이상이면 객체를 추적(tracking)해서 객체가 등장하는 프레임을 샘플링할 수 있다. 프로세서(210)는 객체를 추적하기 위해 점 추적, 커널(kernel) 추적, 실루엣 추적 등을 이용할 수 있다. 따라서, 프로세서(210)는 유의미한 프레임에 대해서만 샘플링을 함으로써, 컴퓨팅 로드를 감소시킬 수도 있다.

[55] 일 실시예에 따른 프로세서(210)는 영상(300)을 샘플링하여 서로 다른 시간에 대응하는 복수의 프레임(310)을 획득할 수 있다. 샘플링된 시간 T1, T2, T3, T4

사이의 간격은 일정할 수도 있고, 일정하지 않은 임의의 시간 간격일 수 있다. 복수의 프레임 중 각각의 프레임(311, 312, 313, 314)은 그 샘플링 된 시간에서의 영상의 프레임을 나타낸다.

[56] 도 4a, 도 4b는 본 개시의 일 실시예에 따른 학습 모델 훈련 모습을 예시적으로 도시한 도면이다. 도 4a에서는 설명의 편의를 위해 학습 모델의 참조번호를 '414'로 기재하였지만, 학습 모델(104)와 학습 모델(414)는 서로 다른 모델이 아니며 동일하거나 대응하는 학습 모델일 수 있다.

[57] 전술한 바와 같이, 프로세서(210)는 수신된 영상(102)으로부터 샘플링을 통해 복수의 프레임(311, 312, 313, 314)을 획득할 수 있다.

[58] 먼저 도 4a를 참조하면, 프레임 중 하나의 프레임(410)은 객체(411)를 포함할 수 있다. 예를 들어 객체(411)는 캐릭터, 인물, 특정한 형상, 의상, 건물, 또는 로고일 수 있다. 프로세서(210)는 프레임에서 객체(411)를 검출하기 위해 객체가 있는 영역(이하, 관심 영역(Region of Interest, ROI))을 식별(identify)할 수 있다.

[59] 일 실시예에서, 프로세서(210)는 프레임에서 색상, 질감, 패턴 중 적어도 어느 하나에 기초하여 유사한 영역을 찾고, 유사한 영역을 기초로 그룹화하여 ROI를 설정할 수 있다. 예를 들어, 객체(411)의 경계를 기초로 RGB 값에 급격한 차이가 날 수 있다. 즉, 프로세서(210)는 색이 비슷한 영역들을 유사한 영역으로 설정하고 영역화하여 ROI로 식별할 수 있다. 프로세서(210)는 유사한 영역을 판단하기 위해 프레임에서 색상값의 미분을 취하고, 미분값이 특정 임계값을 넘는 부분을 경계로 보고 유사한 영역을 그룹화할 수 있다. 이 경우, 유사 영역을 기반으로 ROI가 간편하게 설정될 수 있다.

[60] 예를 들어, 프로세서(210)는 객체의 윤곽을 검출하기 위해 소벨(Sobel) 필터를 사용하여 미분 연산자에 의한 픽셀의 밝기 값의 변화 또는 색상(RGB) 값의 변화를 이용하여 윤곽을 찾아낼 수 있다. 소벨 필터는 비선형 연산자로서 사용하는 마스크 창(Mask Window) 영역의 양 끝단에 속한 픽셀들 사이의 RGB 합의 차이를 구한 후 이를 수평과 수직 방향에 대하여 평균 크기를 구함으로써 윤곽 부위를 강조하는 기능을 수행한다. 소벨 필터의 기능에 대하여는 공지되어 있으므로 본 명세서에서는 상세한 설명을 생략한다.

[61] 일 실시예에서, 프로세서(210)는 프레임을 다수의 영역으로 분류하고 영역별로 질감 특징을 히스토그램화 하여 히스토그램의 값이 일정 크기 범위인 부분들을 하나의 영역으로 그룹화하여 ROI를 식별할 수도 있다.

[62] 일 실시예에서, 프로세서(210)는 운영자(110)로부터 입력을 받아서 ROI를 설정할 수도 있다. 이는 운영자(110)가 관심 있어하는 영역, 즉 객체가 있다고 판단되는 영역을 설정하면 그 부분으로 ROI가 설정될 수 있다.

[63] 일 실시예에서, 프로세서(210)는 설정된 ROI를 기초로 배경으로부터 객체를 분리할 수 있다. 프로세서(210)는 ROI를 식별하여 배경과 객체를 분리하는 경우 ROI를 기초로 박스(411)를 생성하고, 박스(411)를 기초로 이미지를 크롭(crop)할 수 있다. 크롭된 이미지의 사이즈는 상이할 수 있다.

[64] 일 실시예에서 프로세서(210)는 크롭된 이미지들의 사이즈를 일정한 사이즈로 조정하는 처리를 추가적으로 할 수 있다. 예를 들어, 학습 데이터를 위해 정해진 사이즈로 이미지가 조절될 수 있다.

[65] 일 실시예에서 프로세서(210)는 크롭된 이미지들이 특정한 객체에 대응하는 이미지인지 여부에 따라 훈련 전 사전 분류를 할 수 있다. 예를 들어, 프로세서(210)는 이미지들 사이의 특징점을 찾아 유사한 이미지들끼리 그룹을 생성할 수 있다. 프로세서(210)는 특정 객체를 지시하는 이미지 그룹을 생성하고 이미지 그룹을 기초로 학습 모델(104)을 훈련시킬 수 있다. 이미지의 그룹 생성에 대해서는 이하의 도 5에서 자세히 설명한다.

[66] 다른 실시예에서, 프로세서(210)는 프레임으로부터 ROI를 식별하여 이미지를 취득하는 것뿐만 아니라 이미지를 가공하거나 추가 이미지를 생성할 수도 있다.

[67] 일 실시예에서, 크롭된 이미지에 복수의 객체가 오버랩(overlap) 되어 있을 수 있다. 예를 들어, 복수의 객체로 제1 객체, 제2 객체가 오버랩 되어 있을 수 있다. 여기서 프로세서(210)는 해당 프레임과 다른 프레임에서 제1 객체에 대한 이미지를 획득할 수 있다. 프로세서(210)는 다른 시점의 프레임에서의 제1 객체에 대한 이미지에 기초하여 해당 프레임의 이미지를 분리할 수 있다. 예를 들어, 다른 시점은 해당 프레임 시점 이전의 시점일 수 있고, 해당 프레임 시점 이후 일부 시간 경과 후의 시점일 수도 있다. 예를 들어, 프로세서(210)는 이전 시점의 프레임에서의 제1 객체에 대한 이미지를 해당 프레임으로부터 분리하고 남은 부분을 제2 객체의 이미지로 생성하도록 할 수 있다. 이미지를 분리한다는 것은 해당 이미지에 대한 색을 제거하거나, 해당 이미지의 레이어를 구분하는 것을 의미할 수 있다. 이 경우, 객체가 겹쳐져 있는 경우에서도 다른 시점의 객체 이미지를 통해 각 객체를 분리해내어 훈련을 위한 이미지를 다양하게 생성할 수 있다.

[68] 일 실시예에서 프로세서(210)는 이미지를 병합하여 학습 데이터를 생성할 수도 있다. 제1 프레임에서 얻어진 객체의 일부 이미지와 제2 프레임에서 획득된 객체의 일부 이미지를 병합할 수 있다. 예를 들어, 프로세서(210)는 이미지가 연속성이 있다고 판단하는 경우, 또는 이미지에 공통된 형상이 있다고 판단하는 경우 이미지를 병합하여 학습 데이터로 생성할 수 있다.

[69] 프로세서(210)는 학습 영상(111)으로부터 샘플링되고, 크롭되어 얻어진 복수의 이미지 및 추가적인 처리를 거친 이미지들을 기초로 학습 모델(414)을 훈련시킬 수 있다.

[70] 따라서, 학습 모델(414)을 훈련시킬 때, 운영자(110)가 방대한 양의 학습 이미지를 따로 입력할 필요 없이 학습 영상(111) 하나를 입력하는 것만으로 충분히 훈련시킬 수 있다는 효과를 갖을 수 있다.

[71] 일 실시예에서 프로세서(210)는 다양한 객체에 대한 이미지들을 기초로 학습 모델(414)을 훈련시킬 수 있다. 프로세서(210)는 학습 모델(414)에 지도 학습(supervised learning)을 시킬 수 있다. 지도 학습이란, 정해진 레이블(label)에

따라 정해진 이미지들을 모델에 훈련시켜 모델이 특정 이미지를 특정 레이블로 분류할 수 있도록 훈련시키는 것을 의미한다.

[72] 일 실시예에서, 프로세서(210)는 이미지를 훈련시킬 때마다 운영자(110)로부터 이미지에 대한 레이블을 입력받을 수 있다. 예를 들어, 프로세서(210)는 디스플레이(미도시)를 통해 운영자가 크롭된 이미지 중 하나를 확인할 수 있도록 할 수 있다. 운영자(110)는 이미지를 보고 해당하는 레이블을 타이핑(typing), 클릭(click) 등의 방식으로 입력할 수 있다. 학습 모델(414)은 각각의 이미지에 따른 레이블을 입력 받아 지도 학습될 수 있다.

[73] 다른 실시예에서, 프로세서(210)는 유사하게 분류된 이미지들에 대해 임의의 레이블로 분류하여 학습 모델(414)을 훈련시킬 수도 있다. 즉, 이미지 그룹별로 레이블이 정해질 수 있고, 동일한 이미지 그룹에 속한 이미지들을 동일한 레이블로 학습 모델(414)을 훈련시킬 수 있다. 예를 들어, 프로세서(210)는 복수의 이미지(412, 413, 414)를 동일한 레이블에 해당한다고 분류하고 이미지들(412, 413, 414)을 "Label 1"으로 레이블링을 통해 학습 모델(414)을 훈련시킬 수 있다. 예를 들어, 학습 모델(414)은 복수의 이미지(412, 413, 414)와 유사한 이미지는 "Label 1"이라고 분류할 수 있도록 학습될 수 있다.

[74] 일 실시예에서, 프로세서(210)는 메모리(220)에 미리 저장되어 있는 레이블과 각 레이블에 해당하는 이미지를 로드(load)하여 훈련에 사용할 수 있다. 프로세서(210)는 로드한 레이블에 해당하는 이미지를 찾아내고, 레이블에 따라 1차적으로 분류를 진행한 뒤 분류된 이미지들을 학습 모델(414)의 학습 데이터로 사용할 수 있다.

[75] 도 4b를 참조하면, 프로세서(210)는 특정한 객체의 이미지가 추출되는 확률을 계산하고 이를 기초로 학습 대상을 결정할 수도 있다. 예를 들어, 프로세서(210)는 객체 A에 대한 이미지(422)가 전체 객체 이미지에서 차지하는 비율을 계산할 수 있다. 또는, 프로세서(210)는 획득한 이미지들 중 객체 A에 대한 이미지(422)에 해당하는 이미지의 개수를 계산할 수 있다. 예를 들어, 운영자로부터 레이블을 입력받은 경우에는, 프로세서(210)는 특정 레이블에 해당하는 이미지의 개수를 계산할 수 있다. 예를 들어, 프로세서(210)가 유사한 이미지끼리 분류하여 이미지 그룹을 만드는 경우라면, 프로세서(210)는 해당하는 이미지 그룹에 속하는 이미지의 개수를 계산할 수 있다.

[76] 일 실시예에서, 프로세서(210)는 특정한 객체의 이미지에 해당하는 이미지의 개수가 정해진 개수 이상이면 해당하는 특정한 객체를 학습 대상으로 선택할 수 있다. 학습 대상으로 선택한다는 것은, 학습 모델(414)을 훈련시키기 위한 입력(input)으로 사용한다는 것을 의미한다.

[77] 일 실시예에서, 프로세서(210)는 특정한 객체의 이미지에 해당하는 이미지가 전체 이미지에서 차지하는 비중이 일정한 비율 이상이면 해당하는 특정한 객체를 학습 대상으로 선택할 수 있다. 예를 들어, 샘플링된 전체 프레임에서 크롭된 모든 이미지들 중 객체 A에 대한 이미지가 차지하는 비율이 70% 이상일

수 있다. 프로세서(210)는 비율이 30%가 넘어가는 이미지들을 학습 대상으로 선택할 수 있고, A에 대한 이미지가 차지하는 비율이 해당하는 임계치를 초과하므로, A에 대한 이미지들을 학습 대상으로 선택할 수 있다. 그 후 A에 대한 이미지들을 훈련시킬 수 있다.

[78] 일 실시예에서 프로세서(210)는 특정한 객체의 이미지에 해당하는 확률을 계산하기 위해 전체 크롭된 이미지의 개수를 측정(measure)할 수 있고, 그 중에서 A로 라벨링된 이미지, 또는 A와 유사한 이미지로 그룹화된 이미지들의 개수를 측정할 수 있다. 프로세서(210)는 동일한 종류로 라벨링된 또는 그룹화된 이미지들을 전체 이미지의 개수로 나누어 확률을 계산할 수 있다.

[79] 일 실시예에서 프로세서(210)는 어떤 객체의 이미지가 자주 등장하지 않은 경우라면 위에서 정의되는 확률보다 낮다고 판단할 수 있다. 또는, 프로세서(210)는 어떤 객체의 이미지가 자주 등장하지 않은 경우 그 객체에 대한 이미지의 절대적인 개수가 정해진 값보다 적다고 판단하여 해당하는 객체의 이미지를 학습 대상에서 제외할 수 있다. 프로세서(210)는 메모리(220)에 크롭된 모든 이미지를 저장하고 필요시 로드해서 사용할 수 있는데 학습 대상에서 제외하는 경우 해당하는 이미지를 메모리(220)에서 제거할 수 있다.

[80] 프로세서(210)는 학습 모델(414)을 학습하기 전에 미리 필터링을 통해 학습 대상을 선택하고 일정 빈도 이상의 객체에 대해서만 학습 대상으로 선택함으로써, 불필요한 리소스(resource) 낭비를 방지할 수 있다. 또한, 프로세서(210)는 학습 대상을 일부로 제한함으로써 지나친 레이블의 필요성을 방지하여 추후 학습된 모델이 분류를 할 때 소요되는 시간을 절약할 수도 있다.

[81] 도 5는 본 개시의 일 실시예에 따른 이미지 그룹의 생성 방법을 도시한 도면이다. 도 5에서는 설명의 편의를 위해 학습 모델의 참조번호를 '530'으로 기재하였지만 학습 모델(104), 학습 모델(414), 학습 모델(530)은 서로 다른 모델이 아니며 동일하거나 대응하는 학습 모델일 수 있다. 일 실시예에서 학습 모델을 훈련 시키기 전에 미리 유사한 이미지들끼리 그룹화하는 경우 프로세서(210)는 유사한 객체를 한 그룹으로 분류할 수 있다.

[82] 훈련을 위해 준비된 이미지 사이의 유사도를 판단하기 위해, 이미지에서 각 객체의 특징점(feature point)이 추출될 수 있다. 이미지의 특징점을 추출하는 것은, 모서리, 코너, 객체의 크기, 객체의 색상, 객체의 강도(rigid), 조도 등을 추출하는 것일 수 있다. 예를 들어, 이미지 안에서 픽셀의 값이 갑자기 변하는 경계선을 추출하기 위해 이미지를 미분한 그레디언트(gradient)벡터의 크기를 이용할 수 있다. 이미지 사이에서 추출된 특징점이 유사하면 유사한 이미지로 분류될 수 있다.

[83] 객체마다 그룹으로 분류되는 경우, 그룹 1(510)에는 객체 A에 대한 서로 다른 이미지(511, 512)가 포함될 수 있다. 서로 다른 이미지(511, 512)는, 예를 들어, 그룹의 분류를 특징점으로 한 경우, 특징점이 유사하지만 전체적으로는 사이즈, 표정, 행동, 포즈 등이 다른 이미지일 수 있다. 그룹 2(520)에는 객체 B에 대한

서로 다른 이미지(521, 522)가 포함될 수 있다.

[84] 일 실시예에서, 프로세서(210)는 이미지들을 일차적으로 그룹별로 분류한 후, 다른 특징점을 이용해 유사도를 한 번 더 체크하는 과정을 거쳐 그룹에 상이한 객체의 이미지가 포함되지 않도록 추가로 처리할 수 있다. 예를 들어, 프로세서(210)가 일차적으로 코너를 중심으로 이미지들을 분류하였으면 이차적으로 에지를 중심으로 객체의 이미지 사이의 유사 여부를 판단할 수 있다.

[85] 다른 실시예에서, 프로세서(210)는 이미지들 간의 유사도를 판단하는 기준값을 조절함으로써, 이미지의 분류를 조정할 수 있다. 프로세서(210)가 이미지 사이의 유사도를 넓게 보면 전체 그룹이 감소함에 따라 레이블의 개수가 줄어들고, 학습 모델(530)의 출력 복잡도를 감소시킬 수 있다.

[86] 일 실시예에서, 프로세서(210)는 그룹을 설정하고(즉, 분류를 하고), 그룹화된 이미지를 학습 모델(530)의 학습 데이터로 입력할 수 있다. 프로세서(210)는 이미지를 분류하여 학습 데이터를 정리함으로써 학습 모델(530)을 효율적으로 훈련시킬 수 있는 효과가 있다. 따라서, 프로세서(210)가 이미지를 분류하는 경우 사용자가 일일이 분류할 필요가 없게 되어 사용자의 편의를 도모할 수 있다.

[87] 일 실시예에서 프로세서(210)는 그룹마다 레이블을 배정할 수 있다. 레이블은 의미 없는 순서를 나타낼 수도 있지만, 다른 경우에서 레이블은 객체의 성질을 나타낼 수도 있다. 예를 들어, 레이블을 운영자(110)로부터 입력받는 경우라면 운영자(110)는 그룹 1(510), 그룹 2(520)에 대해 대응하는 캐릭터 이름을 레이블로 설정할 수 있다. 즉, 레이블은 그룹의 객체에 대한 특성(예를 들어, 캐릭터의 이름, 인물의 이름, 로고 명칭, 특정한 글자 등)으로 설정될 수 있고, 단순히 서수나 알파벳으로 설정될 수도 있다.

[88] 일 실시예에서 특정한 그룹에 대해, 학습을 위한 이미지의 개수가 부족할 수 있다. 예를 들어, 그룹 1에 대한 이미지의 개수가 학습에 필요한 개수 미만일 수 있다. 예를 들어, 학습에 필요한 이미지의 개수는 프로세서(210)에 미리 정해져 있을 수 있다. 예를 들어, 학습에 필요한 최소 이미지의 개수는 50개 일 수 있고, 그룹 1에 대한 이미지의 개수가 50개 미만일 수 있다.

[89] 일 실시예에서, 학습에 필요한 이미지 개수가 부족한 경우, 프로세서(210)는 해당하는 레이블의 그룹에 대한 이미지를 추가적으로 생성할 수 있다. 예를 들어, 프로세서(210)는 객체에 대한 이미지의 개수가 소정의 개수 미만이면 이미지 처리 모델을 통해 추가적인 이미지를 생성하고 추가적인 이미지를 학습 모델(530)의 학습 데이터로 사용할 수 있다. 예를 들어, 이미지 처리 모델은 미리 메모리(220)에 저장되어 있고, 프로세서(210)에 의해 로드될 수 있다. 예를 들어, 이미지 처리 모델은 이미지의 크기를 변경하거나, 색상을 변경하거나, 색 변환을 실시하거나, 채도를 조절하거나, 이미지의 일부를 자르거나, 이미지를 회전시키거나, 이미지를 다른 이미지와 합치거나 이미지의 종횡비를 조절하는 처리를 할 수 있다. 프로세서(210)는 이미지 처리 모델을 통해 특정한 객체에 대한 추가적인 이미지를 생성할 수 있다. 추가적인 이미지는 객체에 대한 이미지

수를 증가시켜, 훈련을 위한 데이터가 부족한 경우에 있어서도 추가적인 데이터를 통해 학습 모델(530)을 훈련시키는데 사용될 수 있다.

[90] 일 실시예에서, 추가적인 학습을 위한 이미지를 생성하기 위해 프로세서(210)는 미리 정의된 이미지 생성 알고리즘을 사용할 수 있다. 이미지 생성 알고리즘은 위에서 설명된 이미지 변환뿐만 아니라 이미지에 다른 이미지, 레이어를 합성하는 동작을 수행할 수도 있다.

[91] 일 실시예에서, 학습 모델(530)은 필요한 레이블에 해당하는 이미지를 운영자(110)에게 추가적으로 요청할 수도 있다.

[92] 일 실시예에서, 학습 모델(530)은 선형 판별 분석, KNN(K Nearest Neighbor), 의사 결정 트리, 신경망, SVM(Support Vector Machine) 방식 중 적어도 하나에 의해 학습되도록 구성될 수 있다.

[93] 학습 모델(530)이 신경망 모델일 때 학습 모델(530)은 복수의 레이어들을 포함하는 딥 뉴럴 네트워크(deep neural network, DNN)에 해당할 수 있고, 간단히 '신경망'으로 지칭될 수 있다. 복수의 레이어들은 입력 레이어(input layer), 히든 레이어 (hidden layer), 및 출력 레이어(output layer)를 포함할 수 있다. 신경망은 완전 연결 네트워크(fully connected network, FCN), 컨볼루셔널 뉴럴 네트워크(convolutional neural network, CNN), 및 리커런트 뉴럴 네트워크(recurrent neural network, RNN)를 포함할 수 있다. 예를 들어, 신경망 내 복수의 레이어들 중 어느 일부는 컨볼루셔널 뉴럴 네트워크(CNN)에 해당할 수 있고, 다른 일부는 완전 연결 네트워크(FCN)에 해당할 수 있다. 이 경우, 컨볼루셔널 뉴럴 네트워크(CNN)는 컨볼루셔널 레이어로 지칭될 수 있고, 완전 연결 네트워크 (FCN)는 완전 연결 레이어로 지칭될 수 있다.

[94] 컨볼루셔널 뉴럴 네트워크(CNN)의 경우, 각 레이어에 입력되는 데이터는 입력 특징 맵(input feature map)으로 지칭될 수 있고, 각 레이어에서 출력되는 데이터는 출력 특징 맵(output feature map)으로 지칭될 수 있다. 입력 특징 맵 및 출력 특징 맵은 액티베이션 데이터(activation data)로 지칭될 수도 있다.

[95] 신경망 모델은 딥 러닝에 기반하여 트레이닝된 후, 비선형적 관계에 있는 입력 데이터 및 출력 데이터를 서로 매핑함으로써 트레이닝 목적에 맞는 추론(inference)을 수행해낼 수 있다. 딥 러닝은 빅 데이터 세트(Big Dataset)로부터 영상 또는 음성 인식과 같은 문제를 해결하기 위한 기계 학습 기법이다.

[96] 도 6은 본 개시의 일 실시예에 따른 유사도 비교 방법을 도시한 도면이다. 학습 모델(104)은 유사도를 비교하기 위한 영상(102) 입력에 앞서 미리 학습 영상(111)을 통해 분류를 위한 훈련이 된 것으로 가정한다. 영상(102)은 전자 장치(103)에 전송되어 학습 모델(104)에 입력될 수 있다. 전자 장치(103)가 영상(102)을 복수의 프레임으로 샘플링한 뒤 복수의 프레임을 학습 모델(104)에 입력해서 학습 모델(104)은 분류를 진행하고 분류 결과에 기초하여 유사도를 판단할 수 있다. 다만, 전자 장치(103)가 샘플링하여 프레임에 대한 이미지 또는

객체에 대한 이미지를 학습 모델(104)에 입력하는 것에 제한되는 것은 아니고, 학습 모델(104)이 영상(102) 자체를 입력받아 프레임을 샘플링하고 그 후에 유사도를 판단하는 동작까지 수행하게 구성될 수도 있다.

[97] 일 실시예에서, 샘플링된 복수의 프레임 중 일 프레임(611)에서 객체(612)의 이미지가 추출될 수 있다. 객체(612)는 도 4a에서 설명된 훈련을 위한 데이터를 만드는 것과 마찬가지의 방식으로 프레임에서 식별되어 크롭된 뒤 이미지(613)로 추출될 수 있다. 이미지(613)는 학습 모델(104)에 입력되어 하나 이상의 레이블(620) 마다의 복수의 이미지(621, 622, 623)와 비교될 수 있다. 학습 모델(104)은 입력된 이미지를, 학습을 위한 레이블링에서 입력된 모든 레이블의 모든 이미지와 각각 비교할 수 있다.

[98] 다른 실시예에서, 이미지(612)는 따로 크롭되지 않고 프레임(611) 전체에 대한 이미지가 학습 모델(104)에 입력되어 레이블(620)에 해당하는 복수의 이미지(621, 622, 623)와 비교될 수 있다. 즉, 학습 모델(104)은 입력된 프레임에 대한 이미지 내에 레이블(620)에 해당하는 복수의 이미지(621, 622, 623)가 포함되어 있는지를 판단할 수 있다. 예를 들어, 일 레이블에 해당하는 이미지가 입력된 프레임에 포함되었다고 판단되면, 해당 프레임은 판단된 레이블로 분류될 수 있다. 프레임이 하나 이상의 객체 이미지로 추출되지 않고, 프레임 전체로서 학습 모델(104)에 입력된 경우, 해당 프레임에 여러 레이블의 이미지가 있을 수 있다. 프레임에 하나 이상의 레이블에 대응되는 이미지가 포함되어 있는 경우, 해당 프레임은 후의 도 7b에서 설명될 가중치가 가장 높은 레이블에 해당하는 이미지로 분류될 수 있다.

[99] 일 실시예에서, 학습 모델(104)은 유사도 판단을 위해 입력된 이미지에 변경을 가할 수 있다. 예를 들어, 프레임(611)에서 객체에 대한 이미지가 추출되어 학습 모델(104)에 입력된 경우에 있어서, 이미지(613)는 특정한 객체의 뒤집어진 형상일 수 있다. 학습 모델(104)은 이미지(613)를 변경하여 레이블에 대한 이미지들과 비교할 수 있다. 예를 들어, 학습 모델(104)은 이미지(613)에 크기 변경, 색상 변경, 이미지 자르기, 이미지 회전을 통해 이미지를 변경하고 변경된 이미지를 각각의 레이블에 대한 이미지들과 비교할 수 있다. 예를 들어, 학습 모델(104)은 이미지(613)를 180도 회전하여 이미지(614)를 만들고, 회전된 이미지를 레이블(620)에 대한 복수의 이미지(621, 622, 623)와 비교할 수 있다. 이를 통해 변형된 이미지에 대해서도 유사도가 정확하게 계산될 수 있다. 다만, 학습 모델(104)이 이미지를 변형시키는 것에 한정되는 것은 아니고 프로세서(210)가 이미지를 입력하기 전에 미리 이미지를 변형(예: 크기 변경, 색상 변경, 이미지 자르기, 이미지 회전 등)하여 입력할 수도 있다.

[100] 일 실시예에서, 학습 모델(104)은 이미지의 특징점을 획득하고 레이블에 대응하는 이미지와 비교할 수 있다. 예를 들어, 학습 모델(104)은 이미지 중에서 코너를 특징점으로 추출하여 레이블에 대한 이미지와 비교할 수 있다.

[101] 도 7a, 도 7b는 본 개시의 일 실시예에 따른 유사도를 판단하는 방법을 나타내는

도면이다.

[102] 학습 모델(104)은 사용자가 입력한 영상(102)이 학습 모델(104)의 학습 데이터로 사용되었던 학습 영상(111)과 얼마나 유사한지에 관한 유사도를 판단할 수 있다. 이를 위해 학습 모델(104)은 영상(102)의 프레임들 중 객체에 대한 제 1 이미지를 입력받은 후 이미지의 특징점을 추출할 수 있다. 특징점은 코너, 모서리 등일 수 있다. 추출된 객체의 제1 이미지(701) 또는 제1 이미지(701)의 추출된 특징점은 학습 모델(104)에 저장되어 있는 레이블들의 복수의 이미지의 특징점과 비교될 수 있다. 예를 들어, 추출된 제1 이미지(701)의 특징점과 레이블에 해당하는 이미지들 사이의 매칭되는 특징점 개수에 기초하여 레이블에 대한 유사도를 계산할 수 있다.

[103] 일 실시예에서, 하나의 레이블에 복수의 이미지가 있을 때, 가장 높은 유사도를 갖는 이미지를 기초로 해당하는 레이블과의 유사도를 결정할 수 있다. 다른 경우에서, 하나의 레이블에 복수의 이미지가 있을 때, 복수의 이미지에 대한 유사도 평균을 구해서 해당하는 레이블과의 유사도를 결정할 수도 있다.

[104] 다른 실시예에서, 학습 모델(104)은 각 레이블에 해당하는 학습된 이미지들을 기초로 레이블에 대한 대표 이미지를 생성할 수 있다. 이러한 경우 유사도를 판단할 때 영상(102)의 프레임들 중 추출된 객체의 이미지는 레이블에 해당되는 모든 이미지들이 아닌 학습 모델(104)의 각 레이블에 대한 대표 이미지와 제1 이미지(701) 사이의 유사도가 판단될 수 있다. 예를 들어, 학습 모델(104)은 입력된 제1 이미지(701)와 레이블의 대표 이미지의 특징점을 추출하고 매칭되는 특징점 개수를 산출하여 유사도를 판단할 수 있다. 대표 이미지를 이용함으로써 학습 모델(104)의 비교 속도를 향상시킬 수 있다.

[105] 일 실시예에서, 유사도는 신경망 모델을 사용하여 모든 레이블에 대한 유사도의 합이 1이 되도록 하나 이상의레이블 각각에 대해 소프트맥스(softmax) 함수를 통해 계산될 수 있다. 예를 들어, 제1 이미지(701)는 레이블1, 2, 3에 대해 각각 0.75, 0.24, 0.01의 유사도를 가질 수 있다. 이는 제1 이미지(701)는 레이블 1의 이미지와 0.75만큼 즉, 75%만큼의 유사한 이미지임을 나타낼 수 있다. 예를 들어, 제2 이미지(711)는 레이블 1, 2, 3에 대해 각각 0.11, 0.37, 0.52의 유사도를 가질 수 있다. 이는 제2 이미지(711)는 레이블 1의 이미지와는 11%만큼 유사하고, 레이블 2의 이미지와는 37% 유사하며 레이블 3의 이미지와는 52% 유사함을 의미할 수 있다.

[106] 일 실시예에서, 각각의 레이블에 대해 계산된 유사도에 기초하여, 학습 모델(104)은 특정한 이미지가 어떤 레이블에 속하는지 분류할 수 있다. 일 실시예에서, 학습 모델(104)은 레이블 중 유사도가 가장 높은 레이블로 해당 이미지를 분류할 수 있다. 예를 들어, 제1 이미지(701)에 대해서는 레이블 1의 유사도가 가장 높으므로 제1 이미지(701)는 레이블 1로 분류될 수 있다. 예를 들어, 학습 모델(104)은 제2 이미지(711)에 대해서는 레이블 3에 대한 유사도가 가장 높으므로 제2 이미지(711)는 레이블 3으로 분류될 수 있다.

[107] 다른 실시예에서, 학습 모델(104)은 특정 레이블과의 유사도가 가장 높고 유사도가 일정한 임계값을 초과하면 해당 레이블로 분류할 수 있다. 반면, 학습 모델(104)은 특정 레이블과의 유사도가 가장 높아도 해당 유사도가 임계값 이하이면 어떠한 레이블로도 분류하지 않을 수 있다. 예를 들어, 학습 모델(104)은 레이블과의 유사도가 0.5(50%)를 넘으면 해당 레이블로 분류를 할 수 있다. 예를 들어, 제1 이미지(701)는 레이블 1과의 유사도가 0.75이므로 레이블 1(702)로 분류될 수 있다. 즉, 제1 이미지(701)는 레이블 1에 해당하는 것으로 판정될 수 있다. 예를 들어, 특정한 이미지(미도시)는 모든 레이블에 대해 유사도가 0.5를 넘지 않을 수 있다. 이러한 경우 해당하는 이미지는 어떠한 레이블로도 분류되지 않을 수 있다.

[108] 이를 통해 레이블 중 어느 것과도 유사하지 않음에도 학습 모델(104)의 특성 상 레이블 중 어느 하나로 분류되어 추후 총 유사도 판단에 사용되는 경우, 영상(102)이 학습 영상(111)과 유사하지 않음에도 유사하다고 판단될 수 있는 문제를 해결할 수 있다.

[109] 도 7b는 본 개시의 일 실시예에 따른, 영상(102) 속 하나 이상의 이미지 각각의 학습 모델에 훈련된 레이블들에 대한 유사도를 기초로 입력된 영상(102)의 학습 영상(111)에 대한 총 유사도를 판단하는 방법을 나타내는 도면이다.

[110] 일 실시예에서, 학습 모델(104)은 입력 영상(102)에서 추출된 각 객체의 이미지들을 도 7a에 도시된 것과 같이 하나 이상의 레이블 각각의 복수의 이미지와 비교한 뒤 또는 각각의 레이블의 대표 이미지와 비교한 뒤 객체의 이미지의 각각의 레이블에 대한 유사도를 판단할 수 있다. 일 실시예에서, 프로세서(210)는 학습 모델(104)로부터 출력된 영상(102)의 프레임들 속의 추출된 객체 모두가 각각의 레이블에 대해 판단된 유사도를 기초로 총 유사도를 판단할 수 있다. 총 유사도는 하나 이상의 객체가 포함된 입력 영상(102)의 학습 영상(111)에 대한 유사도를 의미한다.

[111] 총 유사도는, 레이블의 가중치의 함수일 수 있다. 예를 들어, 프로세서(210)는 레이블마다 가중치를 다르게 설정하고, 하나 이상의 레이블 각각에 대한 유사도에 하나 이상의 레이블의 가중치를 곱하여 입력 영상(102)의 학습 영상(111)에 대한 총 유사도를 계산할 수 있다. 이를 통해 입력 영상(102)의 학습 영상(111)과의 유사한 정도를 유의미하고 편리하게 계산할 수 있다.

[112] 예를 들어, 프로세서(210)는 학습 영상(111)에서 해당 레이블에 대한 객체 이미지가 얼마나 등장하였는지 빈도를 계산할 수 있다. 빈도를 기초로 가중치를 계산함으로써 합리적인 유사도 판단이 이루어질 수 있다. 프로세서(210)는 계산된 빈도를 기초로 자주 등장한 객체의 레이블에는 가중치를 높게 둘 수 있고, 빈도가 낮은 객체의 레이블에 대해서는 가중치를 낮게 둘 수 있다. 프로세서(210)는 레이블에 대한 가중치를 운영자(110)로부터 입력 받을 수도 있다. 프로세서(103)에는 미리 가중치에 대한 정보가 저장되어 있을 수 있다.

[113] 일 실시예에서, 프로세서(210)는 입력된 영상(102)에 자주 등장하는 객체에

대해 매칭되는 레이블의 가중치를 높게 설정할 수 있다.

[114] 일 실시예에서, 각 레이블을 평가하는 레이블 평가 모델이 있고, 해당 레이블 평가 모델이 레이블의 가중치를 설정할 수 있다. 예를 들어, 레이블 평가 모델은 영상(102)에서 레이블에 해당하는 객체의 빈도 등에 기초한 함수로 학습될 수 있다.

[115] 다른 실시예에서, 프로세서(210)는 레이블에 대해 가중치를 설정하는 것이 아니고 입력된 영상(102)으로부터의 객체(721, 722, 723)에 직접 가중치를 설정할 수도 있다. 프로세서(210)는 객체(721, 722, 723)의 정해진 시간 동안의 움직임 벡터를 측정하고, 움직임 벡터의 양이 정해진 값보다 큰 경우 중요성이 높다고 판단하여 가중치를 높게 설정할 수 있다.

[116] 일 실시예에서, 레이블 또는 객체(721, 722, 723)에 정해진 가중치를 기초로 입력된 영상(102)의 학습 영상(111)에 대한 총 유사도가 계산될 수 있다. 총 유사도를 계산하기 위해 학습 모델(104)은 하나 이상의 레이블 각각의 유사도에 레이블 각각의 가중치를 곱하고 이를 기초로 총 유사도를 계산할 수 있다. 예를 들어, 학습 모델(104)은 객체(721, 722, 723)의 각 유사도를 기초로 입력된 영상(102)이 학습 영상(111)과 73% 유사하다는 총 유사도(724)를 계산할 수 있다.

[117] 일 실시예에 따르면, 전자 장치(103)는 계산된 총 유사도(724)가 임계값 이상이면 입력된 영상(102)의 업로드를 제한할 수 있다. 예를 들어, 전자 장치(103)는 사용자(101)에게 영상 침해에 대한 경고 문구를 디스플레이할 수 있다. 예를 들어, 전자 장치(103)는 영상(102)을 서버(105)에 업로드하는 것을 금지할 수 있다. 예를 들어, 전자 장치(103)는 영상(102)을 전자 장치(103)의 메모리에 저장하는 것을 금지할 수 있다. 예를 들어, 전자 장치(103)는 학습 영상(111)의 일례가 될 수 있는 타인의 저작권이 문제되는 영상과 유사한 영상의 경우 서버(105)에 업로드를 금지하여 저작권 침해 문제를 방지할 수 있다.

[118] 도 8은 본 개시의 일 실시예에 따른, 입력 영상이 학습 영상과 유사한지를 판단하는 방법에 대한 흐름도이다. 이하 실시예에서, 각 단계들은 순차적으로 수행될 수도 있으나, 반드시 순차적으로 수행되는 것은 아니다. 예를 들어, 각 단계들의 순서가 변경될 수도 있으며, 적어도 두 단계들이 병렬적으로 수행될 수도 있다.

[119] 단계(S810)에서 전자 장치(103)는 업로드 영상(102)으로부터 복수의 프레임을 샘플링한다. 예를 들어, 전자 장치(103)는 샘플링 주기가 정해진 경우에는 해당 주기마다 프레임을 샘플링하여 취득할 수 있다. 샘플링 주기가 정해지지 않은 경우에는 전자 장치(103)는 운영자(110)로부터 원하는 프레임 시간을 입력 받아 샘플링할 수 있다.

[120] 단계(S820)에서는 전자 장치(103)가 샘플링 된 복수의 프레임에서 객체가 있는 영역을 크롭하여 객체에 대응하는 이미지를 생성할 수 있다. 객체가 있는 영역은 ROI로서 식별될 수 있고, 운영자(110)에 의해 마우스 드래그로 지정될 수 있다. 운영자(110)의 입력이 아닌, ROI 지정에 의해 객체가 있는 영역이 식별되는

경우에는, ROI를 찾는 알고리즘에 의해 객체가 있는 영역이 식별될 수 있다.

[121] 단계(S830)에서 전자 장치(103)는 객체에 대응하는 이미지들을 레이블링하여 학습 모델을 지도 훈련시킬 수 있다. 전자 장치(103)는 학습 데이터로서 객체에 대응하는 이미지와, 해당하는 이미지에 대한 레이블을 학습 모델에 입력할 수 있다. 학습 데이터를 이용하여 학습 모델(104)은 입력 변수와 출력 변수를 맵핑시키는 함수를 찾을 수 있다. 훈련은 유사도를 판단하고자 하는 영상(102)이 전자 장치(103)에 입력 되기 전에 이루어질 수 있다.

[122] 단계(S840)에서 학습 모델(104)을 통해 입력 영상(102) 내의 객체들의 각각의 레이블에 대한 유사도를 획득할 수 있다. 입력된 영상은 복수 개의 프레임으로 분리되어 학습 모델(104)에 입력될 수 있다. 입력된 영상은 복수 개의 프레임으로 분리된 뒤 객체에 대한 이미지들이 크롭되어 학습 모델(104)에 입력될 수도 있다. 학습 모델(104)은 출력으로서 해당 이미지에 대해, 레이블들 중 가장 유사한 레이블을 출력할 수 있다. 예를 들어, 학습 모델(104)이 레이블 A, B, C로 지도 학습된 경우라면 학습 모델(104)에 입력된 객체의 이미지는 A, B, C 중 하나로 분류될 수 있다. 일 실시예에서, 학습 모델(104)은 추가적인 출력으로 객체의 이미지가 각 레이블과 갖는 유사도를 출력할 수 있다. 예를 들어, A, B, C 레이블과의 각각의 유사도를 출력할 수 있다.

[123] 단계(S850)에서, 전자 장치(103)는 레이블 각각에 대한 가중치를 다르게 설정하여 타겟 영상의 훈련 영상에 대한 총 유사도를 산출할 수 있다. 예를 들어, 총 유사도는 하나 이상의 레이블 각각에 대한 유사도에 하나 이상의 레이블 각각의 가중치를 곱하여 구할 수 있다.

[124] 일 실시예에서, 전자 장치(103)는 레이블에 따른 가중치를 미리 전자 장치(103)의 메모리(220)에 저장된 값을 로드해서 사용할 수 있다. 일 실시예에서, 전자 장치(103)는 레이블에 따른 객체 이미지의 빈도에 기초하여 가중치를 설정할 수 있다. 일 실시예에서, 전자 장치(103)는 운영자(110)의 입력에 기초하여 가중치를 설정할 수도 있다. 예를 들어, 가중치에 따라 특정한 레이블에 해당하는 객체의 이미지가 입력 영상(102)에서 다른 이미지들보다 적게 등장하는 경우라도 입력 영상과 학습 영상(111)의 총 유사도를 높게 하는데 기여할 수 있다.

[125] 전자 장치(103)는 특정한 레이블에 대한 가중치가 소정값 이상인 경우에만, 해당하는 레이블과 일치하는 객체 이미지들의 개수에 기초하여 총 유사도를 계산할 수도 있다. 이러한 경우, 모든 레이블에 대한 유사도를 고려하여 총 유사도를 계산하는 것보다 유의미한 레이블에 대해서만 계산하게 되므로, 계산 속도를 빠르게 할 수 있고 컴퓨팅 로드를 줄일 수 있다는 효과를 갖는다.

[126] 이상에서 설명된 실시예들은 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치, 방법 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal

processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다. 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로 (collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 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 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[127] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단 독으로 또는 조합하여 저장할 수 있으며 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 위에서 설명한 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 또는 복수의 소프트웨어 모듈로서 작동하 도록 구성될 수 있으며, 그 역도 마찬가지이다. 이상과 같이 실시예들이 비록 한정된

도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 이를 기초로 다양한 기술적 수정 및 변형을 적용할 수 있다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다. 그러므로, 다른 구현들, 다른 실시예들 및 특허 청구범위와 균등한 것들도 후술하는 특허 청구의 범위의 범위에 속한다.

[128] 이하에서는, 본 개시의 다양한 실시예에 대해서 부기한다.

[129] [부기 1]

[130] 하나 이상의 프로세서; 및

[131] 상기 프로세서에 의한 실행 시 상기 프로세서가 연산을 수행하도록 하는 명령어들이 저장된 하나 이상의 메모리를 포함하고,

[132] 상기 프로세서에 의해 상기 명령어들이 실행될 시, 상기 프로세서는,

[133] 이미지 프레임의 시퀀스로 구성된 학습 영상을 복수의 프레임으로 샘플링하고,

[134] 상기 복수의 프레임 각각에 나타난 하나 이상의 객체를 검출하여 상기 하나 이상의 객체에 대응하는 복수의 객체 이미지를 생성하고,

[135] 상기 복수의 객체 이미지를 레이블링(labeling)하여 상기 하나 이상의 객체 각각으로 분류되도록 학습 모델을 훈련하도록 구성되는, 전자 장치.

[136] [부기 2]

[137] 제1항에 있어서,

[138] 상기 프로세서는,

[139] 상기 복수의 객체 이미지 사이의 특징점을 찾아 비교하여 특정 객체를 지시하는 유사한 이미지들인지 여부를 판단하고,

[140] 유사한 이미지들끼리 그룹화하여 상기 특정 객체를 지시하는 이미지 그룹을 생성하며,

[141] 상기 이미지 그룹에 기초하여 상기 학습 모델을 훈련시키도록 더 구성되는, 전자 장치.

[142] [부기 3]

[143] 제1항에 있어서,

[144] 상기 하나 이상의 객체를 검출하여 상기 하나 이상의 객체에 대응하는 복수의 객체 이미지를 생성하는 것은, 프레임 내에 색상, 질감, 또는 패턴 중 적어도 하나에 기초한 유사한 값을 갖는 영역을 관심 영역으로 그룹화하고, 상기 관심 영역을 배경과 분리함으로써 객체 이미지를 생성하는 것을 포함하는, 전자 장치.

[145] [부기 4]

[146] 제1항에 있어서,

[147] 상기 객체는 캐릭터, 인물, 특정한 형상, 의상, 건물, 또는 로고 중 적어도 하나에 해당하는, 전자 장치.

[148] [부기 5]

[149] 제1항에 있어서,

[150] 상기 프로세서는, 상기 복수의 프레임 중 제1 시점에서 샘플링된 제1 프레임에 제1 객체 및 제2 객체가 오버랩(overlap)되어 있는 경우, 상기 제1 시점과 다른 제2 시점에서의 상기 제1 객체가 나타난 제2 프레임 및 상기 제1 프레임을 이용하여, 상기 제1 프레임으로부터 상기 제2 객체의 이미지를 생성하도록 구성되는, 전자 장치.

[151] [부기 6]

[152] 제1항에 있어서,

[153] 상기 프로세서는,

[154] 상기 복수의 객체 이미지 중 상기 하나 이상의 객체 각각에 해당하는 이미지의 개수를 계산하고,

[155] 상기 이미지의 개수가 소정의 개수 이상인 객체에 대응하는 하나 이상의 객체의 이미지를 기초로 상기 학습 모델을 훈련하도록 구성되는, 전자 장치.

[156] [부기 7]

[157] 제1항에 있어서,

[158] 상기 프로세서는,

[159] 상기 학습 영상에 나타난 특정 객체의 움직임에 대한 움직임 벡터를 취득하고, 정해진 시간 동안의 상기 움직임 벡터의 크기가 기준값 이상이면 상기 특정 객체를 추적(tracking)하고,

[160] 상기 특정 객체가 나타난 프레임을 샘플링하도록 더 구성되는, 전자 장치.

[161] [부기 8]

[162] 제1항에 있어서,

[163] 상기 프로세서는, 미리 정해진 주기에 따라 상기 학습 영상을 샘플링하거나, 입력에 의해 특정 시점에 해당하는 프레임을 샘플링하도록 구성되는, 전자 장치.

[164] [부기 9]

[165] 제6항에 있어서,

[166] 상기 프로세서는, 상기 학습 모델을 훈련시키기 위한 특정 객체에 대한 이미지의 개수가 소정의 개수 이하인 경우, 이미지 처리 모델을 통해 추가적인 이미지를 생성하고 상기 특정 객체에 대한 추가적인 이미지로 상기 특정 객체에 대해 학습 모델을 훈련시키도록 더 구성되는, 전자 장치.

[167] [부기 10]

[168] 제9항에 있어서,

[169] 상기 이미지 처리 모델은 객체에 대한 이미지의 크기 변경, 색상 변경, 자르기, 또는 회전 중 적어도 하나를 수행하는, 전자 장치.

[170] [부기 11]

[171] 제1항에 있어서,

[172] 상기 학습 모델은 선형 판별 분석, KNN(K Nearest Neighbor), 의사 결정 트리, 신경망, SVM(Support Vector Machine) 방식 중 적어도 하나에 의해 훈련시키도록

더 구성되는, 전자 장치.

[173] [부기 12]

[174] 하나 이상의 프로세서;

[175] 네트워크 인터페이스;

[176] 학습 모델을 저장하는 하나 이상의 메모리를 포함하고,

[177] 상기 학습 모델은 학습 영상에 포함된 객체와 레이블의 상관 관계가 학습된 모델이고,

[178] 상기 프로세서는,

[179] 상기 메모리로부터 상기 학습 모델을 로드(load)하고,

[180] 입력된 타겟 영상으로부터 획득한 복수의 이미지를 상기 학습 모델에 입력하여 하나 이상의 레이블 각각에 대한 유사도를 획득하고,

[181] 상기 하나 이상의 레이블 각각에 대한 유사도에 상기 하나 이상의 레이블 각각의 가중치를 곱하여, 상기 타겟 영상의 상기 학습 영상에 대한 총 유사도를 산출하도록 구성되는, 전자 장치.

[182] [부기 13]

[183] 제12항에 있어서,

[184] 상기 프로세서는,

[185] 상기 타겟 영상으로부터 복수의 프레임을 획득하고,

[186] 상기 복수의 프레임으로부터 하나 이상의 객체에 대응하는 복수의 객체 이미지를 추출하여 상기 학습 모델에 입력하고,

[187] 상기 학습 모델은 상기 복수의 객체 이미지를 상기 하나 이상의 레이블 각각에 대응하는 이미지들과 비교하도록 구성되는, 전자 장치.

[188] [부기 14]

[189] 제13항에 있어서,

[190] 상기 학습 모델은 상기 복수의 객체 이미지에 크기 변경, 색상 변경, 자르기, 또는 회전 중 적어도 하나를 수행한 후, 상기 하나 이상의 레이블 각각에 대응하는 이미지들과 비교하도록 더 구성되는, 전자 장치.

[191] [부기 15]

[192] 제13항에 있어서,

[193] 상기 학습 모델은,

[194] 상기 복수의 객체 이미지에서 상기 하나 이상의 객체 각각의 특징점을 추출하고,

[195] 상기 특징점을 이용하여 상기 하나 이상의 레이블 각각에 대응하는 이미지들과 비교하여 유사도를 계산하도록 구성되는, 전자 장치.

[196] [부기 16]

[197] 제15항에 있어서,

[198] 상기 유사도는, 상기 특징점과 상기 하나 이상의 레이블 각각에 대응하는 이미지들의 특징점 사이의 매칭되는 특징점 개수에 기초하여 계산되는, 전자

장치.

[199] [부기 17]

[200] 제12항에 있어서,

[201] 상기 유사도는, 신경망 모델을 사용하여, 상기 하나 이상의 레이블에 대한 유사도의 합이 1이 되도록 상기 하나 이상의 레이블 각각에 대해 소프트맥스(softmax) 함수를 통해 계산되는, 전자 장치.

[202] [부기 18]

[203] 제12항에 있어서,

[204] 상기 하나 이상의 레이블 중 특정 레이블에 대한 가중치는, 상기 특정 레이블에 해당하는 객체의 빈도에 따라서 결정되거나, 입력을 받아 결정되는, 전자 장치.

[205] [부기 19]

[206] 제12항에 있어서,

[207] 상기 프로세서는,

[208] 상기 총 유사도가 임계값 이상이면, 상기 타겟 영상의 업로드를 제한하도록 더 구성되는, 전자 장치.

[209] [부기 20]

[210] 입력된 학습 영상을 복수의 프레임으로 샘플링하는 단계;

[211] 상기 복수의 프레임 각각에 나타난 하나 이상의 객체를 검출하여 상기 하나 이상의 객체에 대응하는 복수의 객체 이미지를 생성하는 단계;

[212] 상기 복수의 객체 이미지를 레이블링하여 상기 하나 이상의 객체 각각으로 분류되도록 학습 모델을 훈련시키는 단계;

[213] 입력된 타겟 영상으로부터 획득한 복수의 이미지를 상기 학습 모델에 입력하여 하나 이상의 레이블 각각에 대한 유사도를 획득하는 단계; 및

[214] 상기 하나 이상의 레이블 각각에 대한 유사도에 상기 하나 이상의 레이블 각각의 가중치를 곱하여, 상기 타겟 영상의 상기 학습 영상에 대한 총 유사도를 산출하는 단계

[215] 를 포함하는, 방법.

청구범위

[청구항 1] 하나 이상의 프로세서; 및
상기 프로세서에 의한 실행 시 상기 프로세서가 연산을 수행하도록 하는 명령어들이 저장된 하나 이상의 메모리를 포함하고,
상기 프로세서에 의해 상기 명령어들이 실행될 시, 상기 프로세서는,
이미지 프레임의 시퀀스로 구성된 학습 영상을 복수의 프레임으로 샘플링하고,
상기 복수의 프레임 각각에 나타난 하나 이상의 객체를 검출하여 상기 하나 이상의 객체에 대응하는 복수의 객체 이미지를 생성하고,
상기 복수의 객체 이미지를 레이블링(labeling)하여 상기 하나 이상의 객체 각각으로 분류되도록 학습 모델을 훈련하도록 구성되는, 전자 장치.

[청구항 2] 제1항에 있어서,
상기 프로세서는,
상기 복수의 객체 이미지 사이의 특징점을 찾아 비교하여 특정 객체를 지시하는 유사한 이미지들인지 여부를 판단하고,
유사한 이미지들끼리 그룹화하여 상기 특정 객체를 지시하는 이미지 그룹을 생성하며,
상기 이미지 그룹에 기초하여 상기 학습 모델을 훈련시키도록 더 구성되는, 전자 장치.

[청구항 3] 제1항에 있어서,
상기 하나 이상의 객체를 검출하여 상기 하나 이상의 객체에 대응하는 복수의 객체 이미지를 생성하는 것은, 프레임 내에 색상, 질감, 또는 패턴 중 적어도 하나에 기초한 유사한 값을 갖는 영역을 관심 영역으로 그룹화하고, 상기 관심 영역을 배경과 분리함으로써 객체 이미지를 생성하는 것을 포함하는, 전자 장치.

[청구항 4] 제1항에 있어서,
상기 객체는 캐릭터, 인물, 특정한 형상, 의상, 건물, 또는 로고 중 적어도 하나에 해당하는, 전자 장치.

[청구항 5] 제1항에 있어서,
상기 프로세서는, 상기 복수의 프레임 중 제1 시점에서 샘플링된 제1 프레임에 제1 객체 및 제2 객체가 오버랩(overlap)되어 있는 경우, 상기 제1 시점과 다른 제2 시점에서의 상기 제1 객체가 나타난 제2 프레임 및 상기 제1 프레임을 이용하여, 상기 제1 프레임으로부터 상기 제2 객체의 이미지를 생성하도록 구성되는, 전자 장치.

[청구항 6] 제1항에 있어서,
상기 프로세서는,
상기 복수의 객체 이미지 중 상기 하나 이상의 객체 각각에 해당하는

이미지의 개수를 계산하고,
상기 이미지의 개수가 소정의 개수 이상인 객체에 대응하는 하나 이상의 객체의 이미지를 기초로 상기 학습 모델을 훈련하도록 구성되는, 전자 장치.

[청구항 7] 제1항에 있어서,
상기 프로세서는,
상기 학습 영상에 나타난 특정 객체의 움직임에 대한 움직임 벡터를 취득하고, 정해진 시간 동안의 상기 움직임 벡터의 크기가 기준값 이상이면 상기 특정 객체를 추적(tracking)하고,
상기 특정 객체가 나타난 프레임을 샘플링하도록 더 구성되는, 전자 장치.

[청구항 8] 제1항에 있어서,
상기 프로세서는, 미리 정해진 주기에 따라 상기 학습 영상을 샘플링하거나, 입력에 의해 특정 시점에 해당하는 프레임을 샘플링하도록 구성되는, 전자 장치.

[청구항 9] 제6항에 있어서,
상기 프로세서는, 상기 학습 모델을 훈련시키기 위한 특정 객체에 대한 이미지의 개수가 소정의 개수 이하인 경우, 이미지 처리 모델을 통해 추가적인 이미지를 생성하고 상기 특정 객체에 대한 추가적인 이미지로 상기 특정 객체에 대해 학습 모델을 훈련시키도록 더 구성되는, 전자 장치.

[청구항 10] 제9항에 있어서,
상기 이미지 처리 모델은 객체에 대한 이미지의 크기 변경, 색상 변경, 자르기, 또는 회전 중 적어도 하나를 수행하는, 전자 장치.

[청구항 11] 제1항에 있어서,
상기 학습 모델은 선형 판별 분석, KNN(K Nearest Neighbor), 의사 결정 트리, 신경망, SVM(Support Vector Machine) 방식 중 적어도 하나에 의해 훈련시키도록 더 구성되는, 전자 장치.

[청구항 12] 하나 이상의 프로세서;
네트워크 인터페이스;
학습 모델을 저장하는 하나 이상의 메모리를 포함하고,
상기 학습 모델은 학습 영상에 포함된 객체와 레이블의 상관 관계가 학습된 모델이고,
상기 프로세서는,
상기 메모리로부터 상기 학습 모델을 로드(load)하고,
입력된 타겟 영상으로부터 획득한 복수의 이미지를 상기 학습 모델에 입력하여 하나 이상의 레이블 각각에 대한 유사도를 획득하고,
상기 하나 이상의 레이블 각각에 대한 유사도에 상기 하나 이상의 레이블 각각의 가중치를 곱하여, 상기 타겟 영상의 상기 학습 영상에 대한 총

유사도를 산출하도록 구성되는, 전자 장치.

[청구항 13] 제12항에 있어서,
상기 프로세서는,
상기 타겟 영상으로부터 복수의 프레임을 획득하고,
상기 복수의 프레임으로부터 하나 이상의 객체에 대응하는 복수의 객체 이미지를 추출하여 상기 학습 모델에 입력하고,
상기 학습 모델은 상기 복수의 객체 이미지를 상기 하나 이상의 레이블 각각에 대응하는 이미지들과 비교하도록 구성되는, 전자 장치.

[청구항 14] 제13항에 있어서,
상기 학습 모델은 상기 복수의 객체 이미지에 크기 변경, 색상 변경, 자르기, 또는 회전 중 적어도 하나를 수행한 후, 상기 하나 이상의 레이블 각각에 대응하는 이미지들과 비교하도록 더 구성되는, 전자 장치.

[청구항 15] 제13항에 있어서,
상기 학습 모델은,
상기 복수의 객체 이미지에서 상기 하나 이상의 객체 각각의 특징점을 추출하고,
상기 특징점을 이용하여 상기 하나 이상의 레이블 각각에 대응하는 이미지들과 비교하여 유사도를 계산하도록 구성되는, 전자 장치.

[청구항 16] 제15항에 있어서,
상기 유사도는, 상기 특징점과 상기 하나 이상의 레이블 각각에 대응하는 이미지들의 특징점 사이의 매칭되는 특징점 개수에 기초하여 계산되는, 전자 장치.

[청구항 17] 제12항에 있어서,
상기 유사도는, 신경망 모델을 사용하여, 상기 하나 이상의 레이블에 대한 유사도의 합이 1이 되도록 상기 하나 이상의 레이블 각각에 대해 소프트맥스(softmax) 함수를 통해 계산되는, 전자 장치.

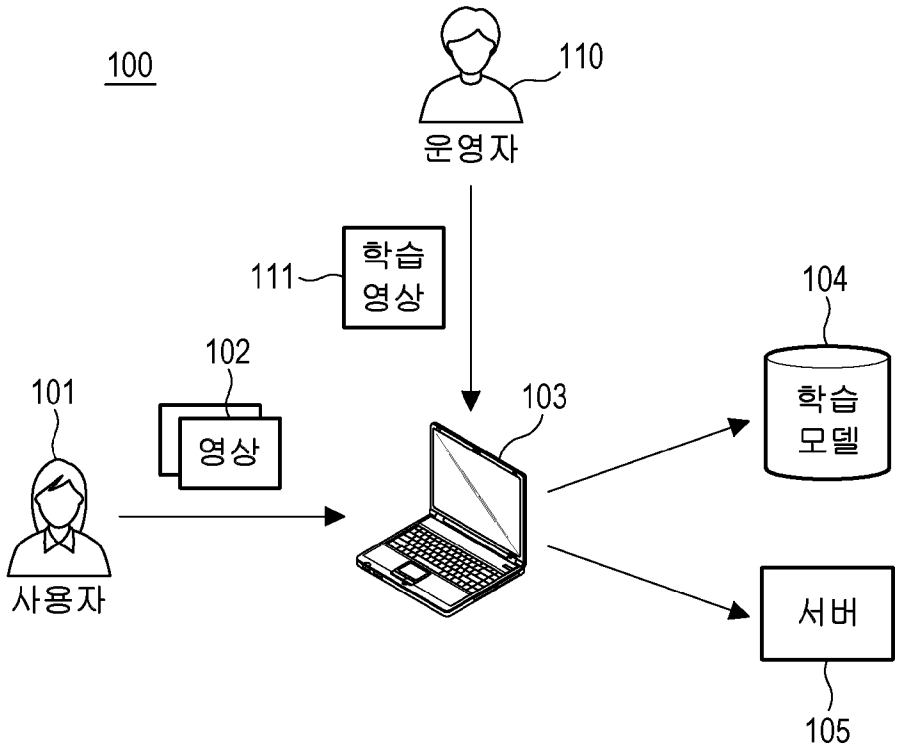
[청구항 18] 제12항에 있어서,
상기 하나 이상의 레이블 중 특정 레이블에 대한 가중치는, 상기 특정 레이블에 해당하는 객체의 빈도에 따라서 결정되거나, 입력을 받아 결정되는, 전자 장치.

[청구항 19] 제12항에 있어서,
상기 프로세서는,
상기 총 유사도가 임계값 이상이면, 상기 타겟 영상의 업로드를 제한하도록 더 구성되는, 전자 장치.

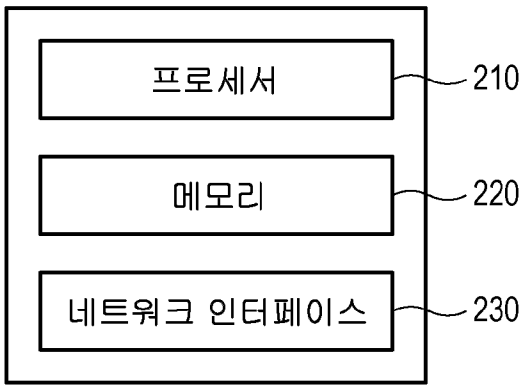
[청구항 20] 입력된 학습 영상을 복수의 프레임으로 샘플링하는 단계;
상기 복수의 프레임 각각에 나타난 하나 이상의 객체를 검출하여 상기 하나 이상의 객체에 대응하는 복수의 객체 이미지를 생성하는 단계;
상기 복수의 객체 이미지를 레이블링하여 상기 하나 이상의 객체

각각으로 분류되도록 학습 모델을 훈련시키는 단계;
입력된 타겟 영상으로부터 획득한 복수의 이미지를 상기 학습 모델에 입력하여 하나 이상의 레이블 각각에 대한 유사도를 획득하는 단계; 및
상기 하나 이상의 레이블 각각에 대한 유사도에 상기 하나 이상의 레이블 각각의 가중치를 곱하여, 상기 타겟 영상의 상기 학습 영상에 대한 총 유사도를 산출하는 단계
를 포함하는, 방법.

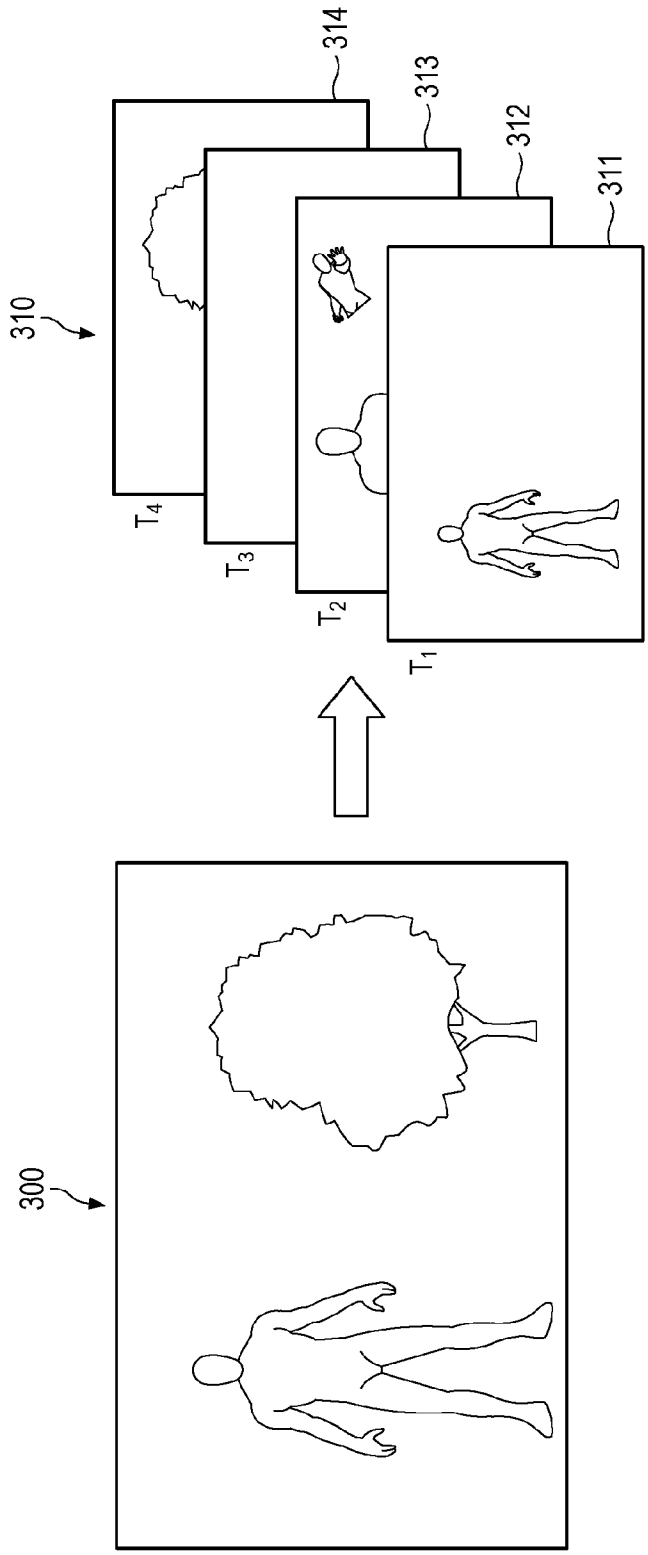
[도1]


100
110
운영자
111
학습
영상
104
102
101
영상
103
학습
모델
사용자
서버
105

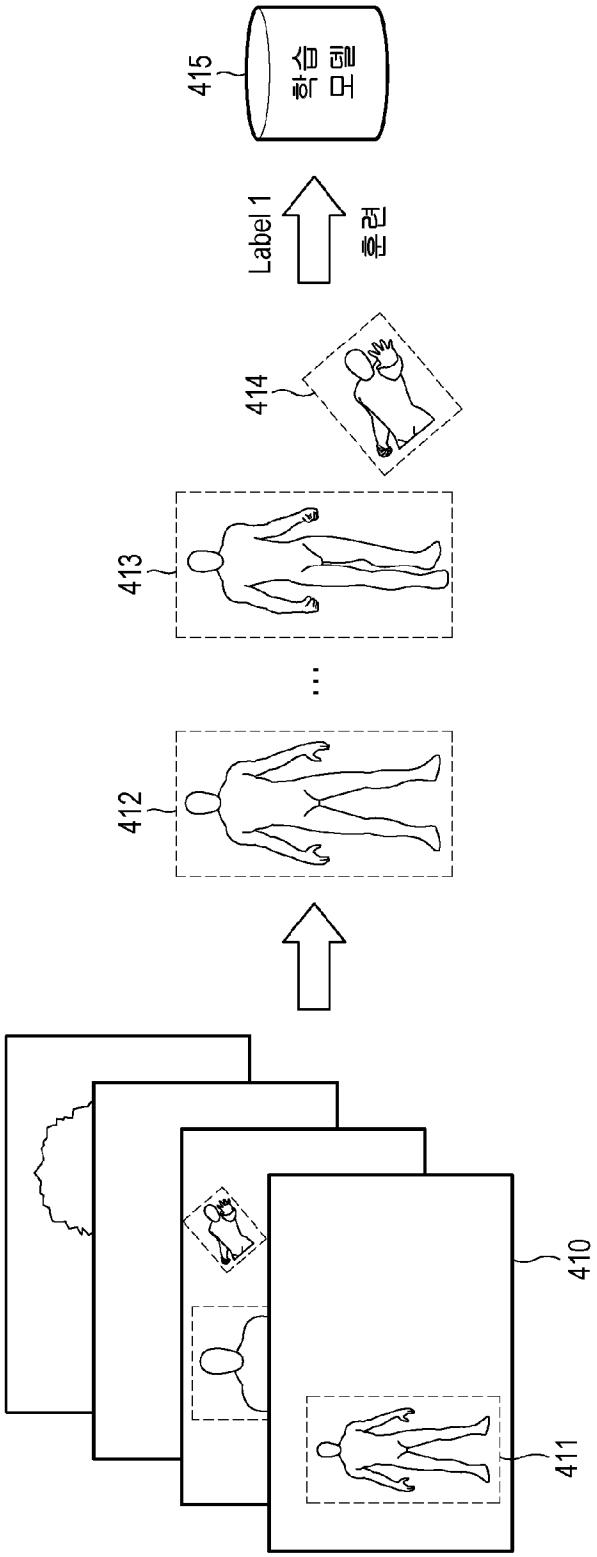
[도2]


200
프로세서
210
메모리
220
네트워크 인터페이스
230

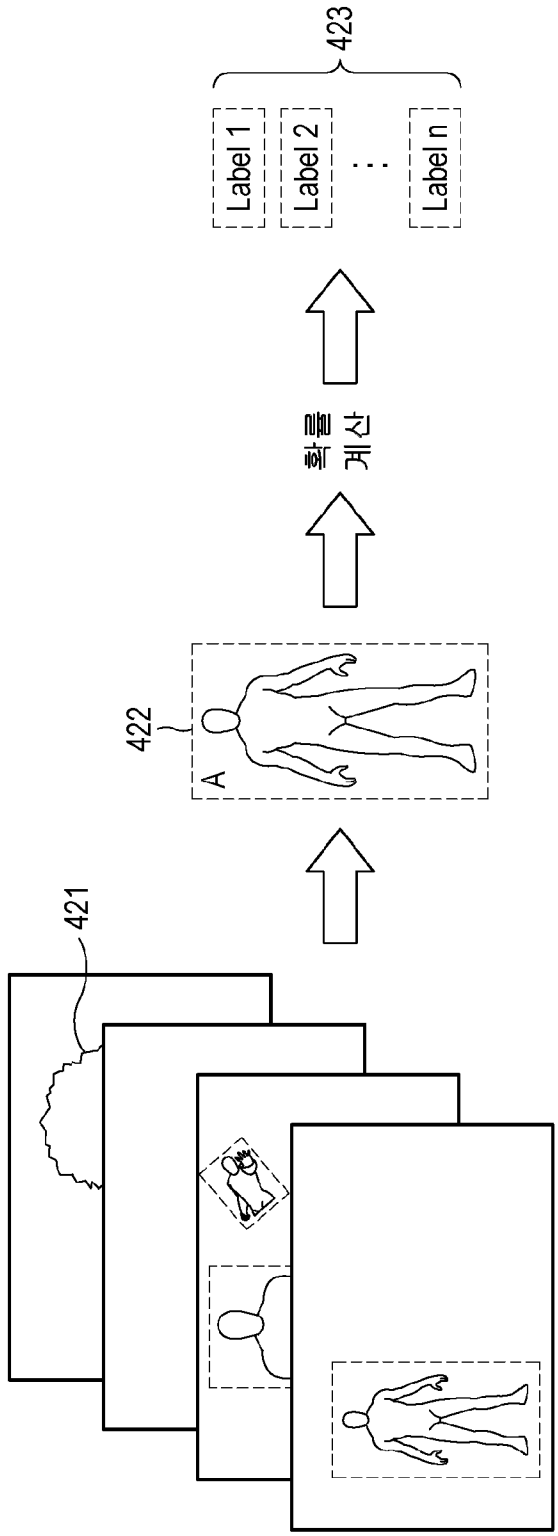
[도3]



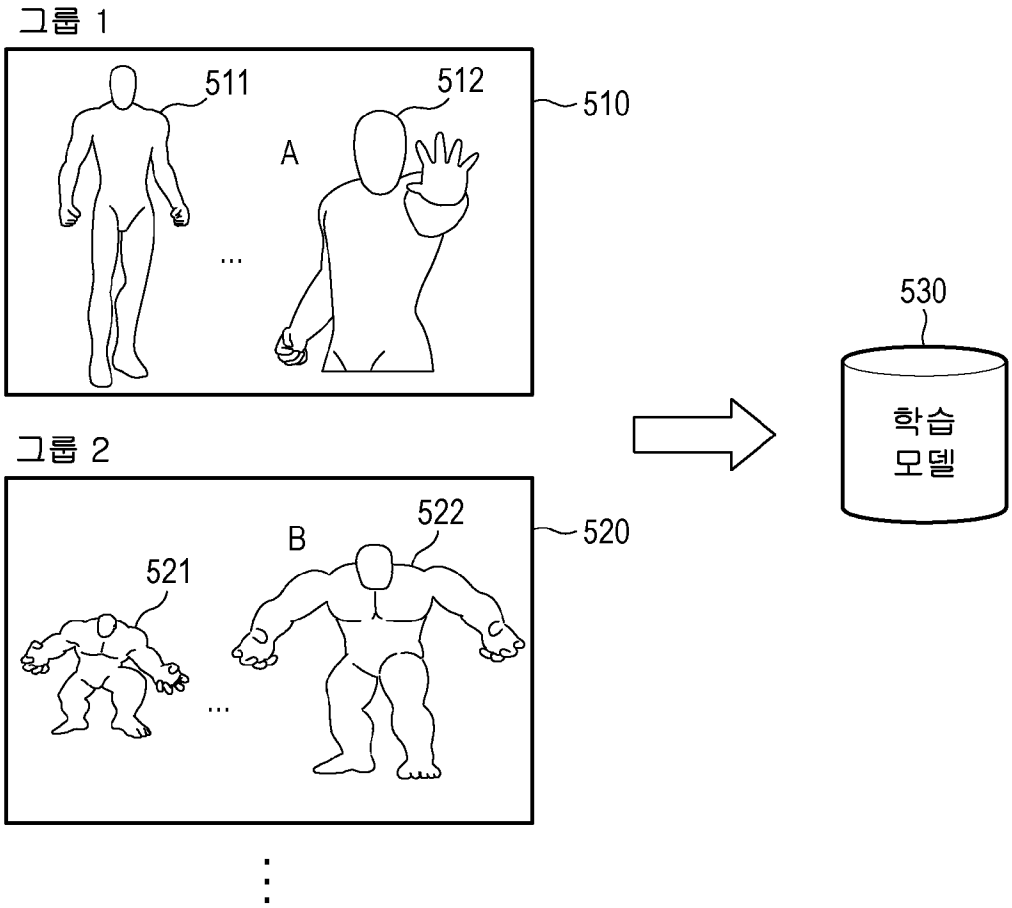
[도4a]



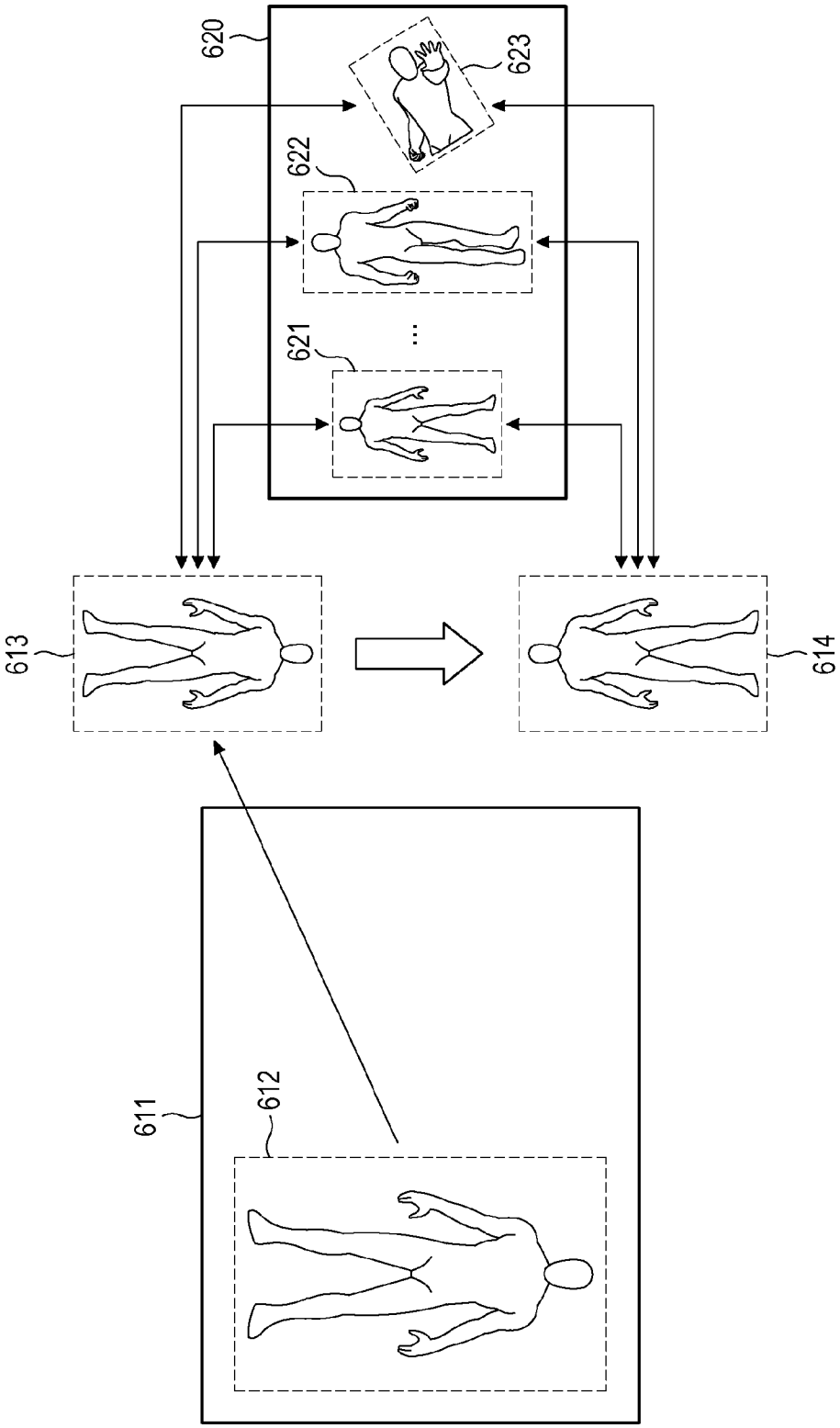
[도4b]



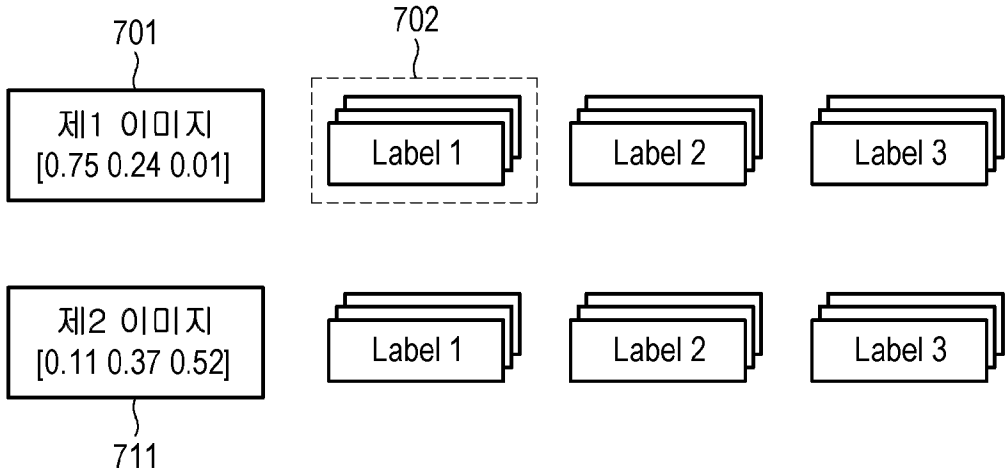
[도5]



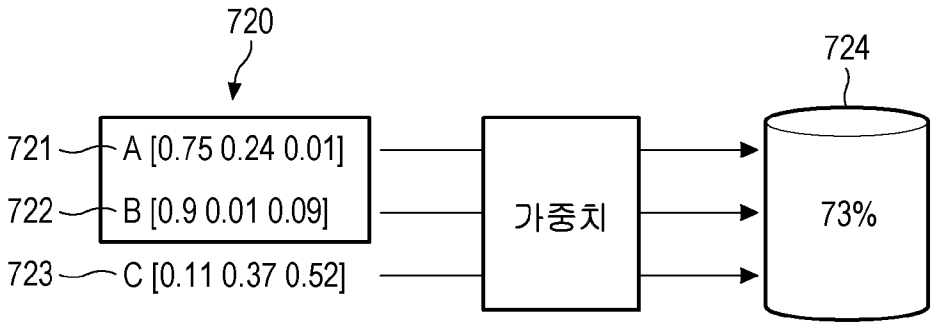
[도6]



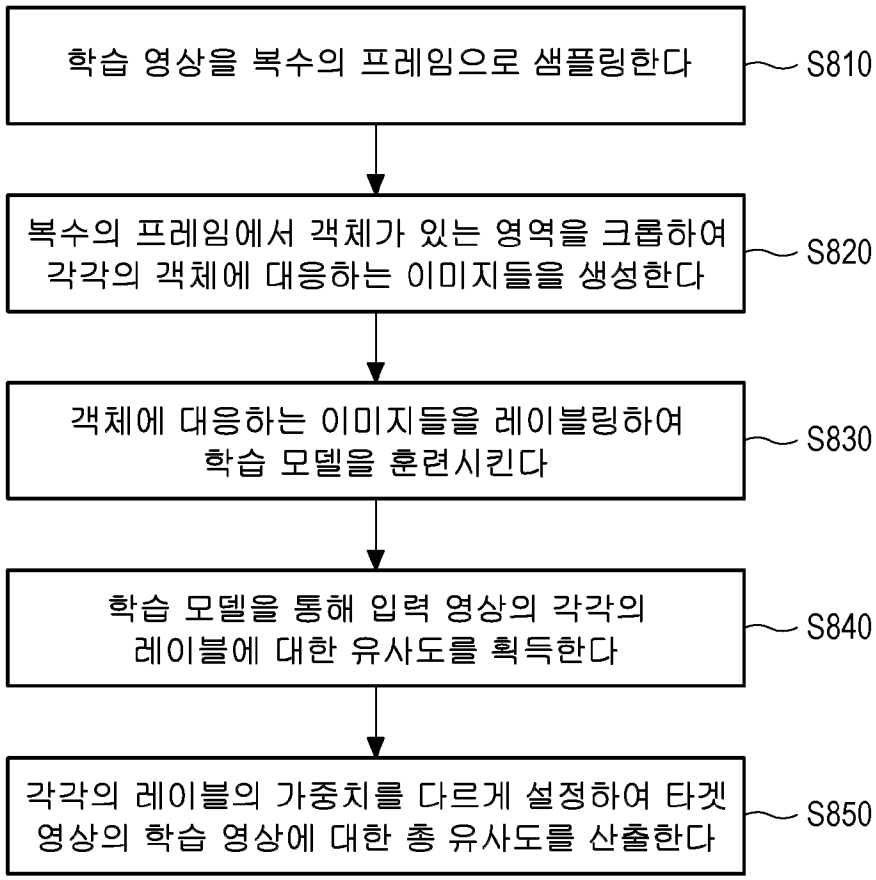
[도7a]


701
702
제1 이미지
[0.75 0.24 0.01]
Label 1
Label 2
Label 3
제2 이미지
[0.11 0.37 0.52]
Label 1
Label 2
Label 3
711

[도7b]


720
724
721
A [0.75 0.24 0.01]
722
B [0.9 0.01 0.09]
723
C [0.11 0.37 0.52]
가중치
73%

[도8]


학습 영상을 복수의 프레임으로 샘플링한다
S810
복수의 프레임에서 객체가 있는 영역을 크롭하여
각각의 객체에 대응하는 이미지들을 생성한다
S820
객체에 대응하는 이미지들을 레이블링하여
학습 모델을 훈련시킨다
S830
학습 모델을 통해 입력 영상의 각각의
레이블에 대한 유사도를 획득한다
S840
각각의 레이블의 가중치를 다르게 설정하여 타겟
영상의 학습 영상에 대한 총 유사도를 산출한다
S850

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2023/001254

A. CLASSIFICATION OF SUBJECT MATTER

G06V 20/70(2022.01)i; **G06V 20/40**(2022.01)i; **G06V 10/774**(2022.01)i; **G06V 10/74**(2022.01)i; **G06V 10/56**(2022.01)i; **G06V 10/54**(2022.01)i; **G06V 40/20**(2022.01)i; **G06T 3/40**(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)

G06V 20/70(2022.01); G06F 16/783(2019.01); G06N 3/00(2006.01); G06T 7/00(2006.01); G06T 7/20(2006.01); G06V 10/82(2022.01); G16H 10/60(2018.01); G16H 50/50(2018.01)

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: 영상(video), 레이블링(labeling), 학습 모델(learning model), 유사도(similarity), 가중치(weight)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KR 10-2021-0051473 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 10 May 2021 (2021-05-10) See paragraphs [0016]-[0104] and figures 2-4 and 9.	1-2,4,6,8,11
Y		3,5,7,9-10,12-20
Y	KR 10-2014-0047331 A (INDUSTRY FOUNDATION OF CHONNAM NATIONAL UNIVERSITY) 22 April 2014 (2014-04-22) See claims 1-2 and 8-9 and figures 1-6.	3
Y	KR 10-2016-0144149 A (INDUSTRY-ACADEMIC COOPERATION FOUNDATION, KUNSAN NATIONAL UNIVERSITY) 16 December 2016 (2016-12-16) See paragraphs [0027] and [0087] and figure 10.	5, 7
Y	KR 10-2022-0079542 A (UIPATH, INC.) 13 June 2022 (2022-06-13) See claims 1-4 and figures 7-8.	9-10,14

☑ 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	Date of mailing of the international search report
19 October 2023	**20 October 2023**

Name and mailing address of the ISA/KR	Authorized officer
Korean Intellectual Property Office **Government Complex-Daejeon Building 4, 189 Cheongsa-ro, Seo-gu, Daejeon 35208**	
Facsimile No. **+82-42-481-8578**	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2023/001254

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2179583 B1 (DEEPNOID CO., LTD.) 17 November 2020 (2020-11-17) See paragraph [0046] and figure 1.	12-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2023/001254

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
KR 10-2021-0051473 A	10 May 2021	None	
KR 10-2014-0047331 A	22 April 2014	KR 10-1394895 B1	14 May 2014
KR 10-2016-0144149 A	16 December 2016	KR 10-1764845 B1	03 August 2017
KR 10-2022-0079542 A	13 June 2022	CN 114556368 A	27 May 2022
		EP 4042335 A1	17 August 2022
		JP 2022-552133 A	15 December 2022
		US 10990876 B1	27 April 2021
		US 11599775 B2	07 March 2023
		US 2021-0103798 A1	08 April 2021
		US 2021-0209449 A1	08 July 2021
		WO 2021-071579 A1	15 April 2021
KR 10-2179583 B1	17 November 2020	None	

국 제 조 사 보 고 서

국제출원번호
PCT/KR2023/001254

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G06V 20/70(2022.01)i; **G06V 20/40**(2022.01)i; **G06V 10/774**(2022.01)i; **G06V 10/74**(2022.01)i; **G06V 10/56**(2022.01)i; **G06V 10/54**(2022.01)i; **G06V 40/20**(2022.01)i; **G06T 3/40**(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G06V 20/70(2022.01); G06F 16/783(2019.01); G06N 3/00(2006.01); G06T 7/00(2006.01); G06T 7/20(2006.01); G06V 10/82(2022.01); G16H 10/60(2018.01); G16H 50/50(2018.01)

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: 영상(video), 레이블링(labeling), 학습 모델(learning model), 유사도(similarity), 가중치(weight)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
X	KR 10-2021-0051473 A (한국전자통신연구원) 2021.05.10 단락 [0016]-[0104] 및 도면 2-4, 9	1-2,4,6,8,11
Y		3,5,7,9-10,12-20
Y	KR 10-2014-0047331 A (전남대학교산학협력단) 2014.04.22 청구항 1-2, 8-9 및 도면 1-6	3
Y	KR 10-2016-0144149 A (군산대학교산학협력단) 2016.12.16 단락 [0027], [0087] 및 도면 10	5,7
Y	KR 10-2022-0079542 A (유아이패스, 인크.) 2022.06.13 청구항 1-4 및 도면 7-8	9-10,14
Y	KR 10-2179583 B1 (주식회사 딥노이드) 2020.11.17 단락 [0046] 및 도면 1	12-20

☐ 추가 문헌이 C(계속)에 기재되어 있습니다.

☑ 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 문헌과조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.
"&" 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일	국제조사보고서 발송일
2023년10월19일(19.10.2023)	**2023년10월20일(20.10.2023)**

ISA/KR의 명칭 및 우편주소	심사관
대한민국 특허청 **(35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사)**	이강하
팩스 번호 **+82-42-481-8578**	전화번호 **+82-42-481-5003**

국 제 조 사 보 고 서
대응특허에 관한 정보

국제출원번호
PCT/KR2023/001254

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2021-0051473 A	2021/05/10	없음	
KR 10-2014-0047331 A	2014/04/22	KR 10-1394895 B1	2014/05/14
KR 10-2016-0144149 A	2016/12/16	KR 10-1764845 B1	2017/08/03
KR 10-2022-0079542 A	2022/06/13	CN 114556368 A	2022/05/27
		EP 4042335 A1	2022/08/17
		JP 2022-552133 A	2022/12/15
		US 10990876 B1	2021/04/27
		US 11599775 B2	2023/03/07
		US 2021-0103798 A1	2021/04/08
		US 2021-0209449 A1	2021/07/08
		WO 2021-071579 A1	2021/04/15
KR 10-2179583 B1	2020/11/17	없음	