

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국



(10) 국제공개번호

WO 2020/184828 A1

2020년 9월 17일 (17.09.2020)

- (51) 국제특허분류: G06T 7/00 (2006.01) G06T 1/60 (2006.01)
- (21) 국제출원번호: PCT/KR2020/000291
- (22) 국제출원일: 2020년 1월 7일 (07.01.2020)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보: 10-2019-0026852 2019년 3월 8일 (08.03.2019) KR
- (71) 출원인: 에스케이텔레콤 주식회사 (SK TELECOM CO., LTD.) [KR/KR]; 04539 서울시 중구 을지로 65, Seoul (KR). 서울대학교산학협력단 (SEOUL NATIONAL UNIVERSITY R&DB) [KR/KR]; 08826 서울시 관악구 관악로 1, Seoul (KR).
- (72) 발명자: 김태완 (KIM, Taewan); 04539 서울시 중구 을지로 65 에스케이텔레콤 주식회사, Seoul (KR). 김건희 (KIM, Gunhee); 08826 서울시 관악구 관악로 1 122A동 402호, Seoul (KR). 이수찬 (LEE, Soochan); 08787 서울시 관악구 남부순환로220길 38 309호, Seoul (KR).
- (74) 대리인: 제일특허법인(유) (FIRSTLAW P.C.); 06775 서울시 서초구 마방로 60, Seoul (KR).

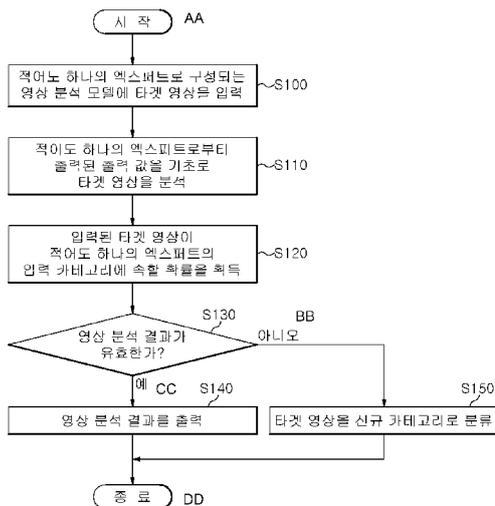
- (81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, 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 국내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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: IMAGE ANALYSIS DEVICE AND METHOD, AND METHOD FOR GENERATING IMAGE ANALYSIS MODEL USED FOR SAME

(54) 발명의 명칭: 영상 분석 장치, 방법 및 이에 이용되는 영상 분석 모델의 생성 방법



(57) Abstract: An image analysis method according to an embodiment includes: a step for inputting a target image to at least one expert composed of a generative model and a discriminative model generated by learning a image; a step for analyzing the target image on the basis of an output value output from the at least one expert; a step for determining the validity of the analysis result for the target image on the basis of the probability that the input target image belongs to an input category of the at least one expert; and a step for classifying the target image into a new category if the analysis result for the target image is determined not to be valid.

(57) 요약서: 일 실시예에 따른 영상 분석 방법은, 학습 영상을 학습하여 생성된 분류 모델(Discriminative Model) 및 생성 모델(Generative Model)로 구성되는 적어도 하나의 엑스퍼트(Expert)에 타겟 영상을 입력하는 단계; 상기 적어도 하나의 엑스퍼트로부터 출력된 출력 값을 기초로 상기 타겟 영상을 분석하는 단계; 상기 입력된 타겟 영상이 상기 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률에 기초하여, 상기 타겟 영상에 대한 분석 결과의 유효성을 결정하는 단계; 및 상기 타겟 영상에 대한 분석 결과가 유효하지 않은 것으로 결정되면, 상기 타겟 영상을 신규 카테고리로 분류하는 단계를 포함한다.

- S100 ... Input target image to image analysis model composed of at least one expert
- S110 ... Analyze target image on basis of output value output from at least one expert
- S120 ... Acquire probability that input target image belongs to input category of at least one expert
- S130 ... Image analysis result valid?
- S140 ... Output image analysis result
- S150 ... Classify target image into new category
- AA ... Start
- BB ... No
- CC ... Yes
- DD ... End

WO 2020/184828 A1

명세서

발명의 명칭: 영상 분석 장치, 방법 및 이에 이용되는 영상 분석 모델의 생성 방법

기술분야

- [1] 본 발명은 입력된 영상을 분석하는 장치, 방법 및 이에 이용되는 영상 분석 모델의 생성 방법에 관한 것이다.

배경기술

- [2] 컴퓨터 비전(Computer Vision)은 카메라, 스캐너 등의 시각 매체를 통해 입력 받은 영상에서 주변 물체와 환경 속성을 분석하여 유용한 정보를 생성하는 기술을 의미한다. 최근에는 컴퓨터 비전 기술이 CCTV나 휴대폰 카메라 등에 적용되어, 사람, 자동차, 동물 등 원하는 객체를 검출하고, 분류할 수 있게 되었다.
- [3] 딥러닝(Deep Learning)과 같은 기계 학습 기술이 소개되기 전에는 영상 내 다양한 특성(Feature)을 찾아 객체를 검출하였으며, 객체 검출에 이용된 대표적 특성으로는 HOG(Histogram of Oriented Gradients), ACF(Aggregated Channel Features), LBP(Local Binary Pattern) 등이 있다. 이러한 특성을 기반으로 입력된 영상으로부터 원하는 객체를 검출할 수 있으며, 딥러닝 기반의 객체 검출 방법의 높은 하드웨어 비용 및 레거시(Legacy) 등을 이유로 현재까지도 특징 기반의 객체 검출 방법이 널리 사용되고 있다.
- [4] 그럼에도 불구하고, 다양한 객체 검출 환경에서 보다 정확한 객체 검출 결과를 얻기 위해, 딥러닝 기반의 다양한 컴퓨터 비전 또는 영상 처리 기술에 대한 연구가 활발히 진행되고 있다.

발명의 상세한 설명

기술적 과제

- [5] 본 발명이 해결하고자 하는 과제는, 타겟 영상이 영상 분석 모델의 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률에 따라, 타겟 영상을 신규 카테고리로 분류할지 여부를 결정하는 영상 분석 장치, 방법 및 이에 이용되는 영상 분석 모델 생성 방법을 제공하는 것이다.
- [6] 다만, 본 발명이 해결하고자 하는 과제는 이상에서 언급한 것으로 제한되지 않으며, 언급되지 않은 또 다른 해결하고자 하는 과제는 아래의 기재로부터 본 발명이 속하는 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제 해결 수단

- [7] 일 실시예에 따른 영상 분석 방법은, 학습 영상을 학습하여 생성된 적어도 하나의 엑스퍼트(Expert)로 구성되는 영상 분석 모델에 타겟 영상을 입력하는 단계; 상기 적어도 하나의 엑스퍼트로부터 출력된 출력 값을 기초로 상기 타겟 영상을 분석하는 단계; 상기 입력된 타겟 영상이 상기 적어도 하나의 엑스퍼트의

- 입력 카테고리에 속할 확률에 기초하여, 상기 타겟 영상에 대한 분석 결과의 유효성을 결정하는 단계; 및 상기 타겟 영상에 대한 분석 결과가 유효하지 않은 것으로 결정되면, 상기 타겟 영상을 신규 카테고리로 분류하는 단계를 포함한다.
- [8] 또한, 상기 입력된 타겟 영상이 상기 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률에 기초하여, 상기 타겟 영상에 대한 분석 결과의 유효성을 랜덤(Random)하게 결정할 수 있다.
- [9] 또한, 상기 타겟 영상에 대한 분석 결과의 유효성을 결정하는 단계는, 상기 적어도 하나의 엑스퍼트 각각에 대한 입력 및 클래스(Class)의 우도(Likelihood)를 획득하는 단계; 상기 적어도 하나의 엑스퍼트 각각에 대한 사전 확률(Prior Probability)를 획득하는 단계; 및 상기 획득된 우도와 상기 사전 확률의 곱을 기초로 상기 입력된 타겟 영상이 상기 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률을 획득하는 단계를 포함할 수 있다.
- [10] 또한, 상기 적어도 하나의 엑스퍼트 각각에 대한 입력 및 클래스의 우도를 획득하는 단계는, 상기 적어도 하나의 엑스퍼트 각각의 분류 모델(Discriminative Model)을 이용하여, 상기 입력된 타겟 영상 및 상기 적어도 하나의 엑스퍼트 각각에 대한 클래스의 제 1 확률을 획득하는 단계; 상기 적어도 하나의 엑스퍼트 각각의 생성 모델(Generative Model)을 이용하여, 상기 적어도 하나의 엑스퍼트 각각에 대한 입력의 제 2 확률을 획득하는 단계; 및 상기 제 1 확률 및 상기 제 2 확률의 곱을 기초로 상기 적어도 하나의 엑스퍼트 각각에 대한 입력 및 클래스의 우도를 획득하는 단계를 포함할 수 있다.
- [11] 또한, 상기 제 2 확률을 획득하는 단계는, 상기 제 2 확률의 하한(Lower Bound)을 상기 제 2 확률로서 추정하는 단계를 포함할 수 있다.
- [12] 또한, 상기 신규 카테고리로 분류된 상기 타겟 영상을 메모리에 저장하는 단계; 및 상기 메모리에 저장된 타겟 영상의 누적 용량이 미리 정해진 기준 용량 이상이면, 상기 메모리에 저장된 타겟 영상을 학습하여 상기 신규 카테고리에 대한 엑스퍼트를 생성하는 단계를 더 포함할 수 있다.
- [13] 또한, 상기 신규 카테고리에 대한 엑스퍼트를 생성하는 단계는, 상기 메모리에 저장된 타겟 영상에 레이블링(Labeling)을 수행하는 단계; 및 상기 레이블링된 타겟 영상을 학습하여 상기 신규 카테고리에 대한 엑스퍼트를 생성하는 단계를 포함할 수 있다.
- [14] 본 발명의 일 실시예에 따른 영상 분석 장치는, 학습 영상을 학습하여 생성된 적어도 하나의 엑스퍼트(Expert)로 구성되는 영상 분석 모델을 기초로 타겟 영상을 분석하는 영상 분석부; 및 상기 타겟 영상이 상기 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률에 기초하여, 상기 타겟 영상에 대한 분석 결과의 유효성을 결정하는 결정부를 포함하고, 상기 영상 분석부는, 상기 타겟 영상에 대한 분석 결과가 유효하지 않은 것으로 결정되면, 상기 타겟 영상을 신규 카테고리로 분류한다.
- [15] 또한, 상기 신규 카테고리로 분류된 상기 타겟 영상을 저장하는 메모리를 더

포함할 수 있다.

[16] 또한, 상기 메모리에 저장된 타겟 영상의 누적 용량이 미리 정해진 기준 용량이상이면, 상기 메모리에 저장된 타겟 영상을 학습하여 상기 신규 카테고리에 대한 엑스퍼트를 생성하는 학습부를 더 포함할 수 있다.

[17] 또한, 제 1 입력 주기 동안 제 1 학습 영상을 순차적으로 입력 받는 단계; 상기 제 1 입력 주기 직후의 제 1 학습 주기 동안 상기 제 1 학습 영상을 학습하여, 제 1 카테고리에 대한 분류 모델(Discriminative Model) 및 생성 모델(Generative Model)로 구성되는 제 1 엑스퍼트(Expert)를 생성하는 단계; 상기 제 1 학습 주기 직후의 제 2 입력 주기 동안 제 2 학습 영상을 순차적으로 입력 받는 단계; 상기 제 2 학습 영상이 상기 제 1 카테고리에 속할 확률에 기초하여, 상기 제 1 엑스퍼트에 대하여 상기 제 2 학습 영상 각각을 학습할지 여부를 결정하는 단계; 및 상기 제 2 입력 주기 직후의 제 2 학습 주기 동안, 상기 제 2 학습 영상 중 상기 제 1 엑스퍼트에 대하여 학습하지 않을 것으로 결정된 학습 영상을 학습하여 제 2 카테고리에 대한 제 2 엑스퍼트를 생성하는 단계를 포함한다.

발명의 효과

[18] 본 발명의 일 실시예에 따르면, 영상 분석 모델을 구성하는 엑스퍼트의 입력 카테고리에 속하지 않는 타겟 영상을 신규 카테고리로 분류함으로써, 실사용 환경의 다양한 타겟 영상의 분석이 가능한 영상 분석 모델을 관리할 수 있다.

[19] 학습 영상을 학습하여 생성된 영상 분석 모델을 실사용 환경에서 이용할 경우, 입력되는 타겟 영상은 학습 영상과는 상이한 획득 조건에서 획득되거나, 학습 영상을 통해 학습하지 못한 객체를 포함할 수 있다. 그 결과, 영상 분석 모델은 분석의 정확도가 낮아질 뿐 아니라, 사용 과정에서 발생된 신규 객체 검출의 필요성에 적응적으로 동작할 수 없다.

[20] 그러나, 본 발명의 일 실시예에서는 타겟 영상을 신규 카테고리로 분류하고, 이를 신규 엑스퍼트 생성에 이용함으로써, 실사용 환경에서 입력될 수 있는 다양한 경우의 타겟 영상의 분석이 가능할 수 있다. 구체적으로, 본 발명의 일 실시예에 따른 영상 분석 장치는 딥러닝과 같은 기계 학습에 의해 생성된 영상 분석 모델 상에 새로운 정보를 추가하는 점진적 학습(Incremental Learning) 및 순차적 학습(Continual Learning) 방법을 채용함으로써, 입력되는 타겟 영상에 적응적으로 동작할 수 있고, 그 결과 실사용 환경에서의 활용성을 높일 수 있다.

[21] 특히, 카메라가 거의 움직이지 않는 고정형 CCTV 환경에서 본 발명의 일 실시예에 따른 IVA(Intelligent Video Analytics) 기술을 상용화할 경우, 설치된 카메라의 위치 및 각도와 무관하게 획득된 영상에 대한 정확한 분석이 수행될 수 있다.

도면의 간단한 설명

[22] 도 1은 본 발명의 일 실시예에 따른 영상 분석 장치의 기능 블록도이다.

[23] 도 2는 본 발명의 일 실시예에 따른 영상 분석 모델의 출력 값 출력 과정을

설명하기 위한 도면이다.

- [24] 도 3은 본 발명의 일 실시예에 따른 영상 분석 모델의 엑스퍼트를 설명하기 위한 도면이다.
- [25] 도 4 내지 도 6은 일 실시예에 따른 영상 분석 장치에 의해 수행되는 신규 카테고리에 대한 엑스퍼트 생성 방법을 설명하기 위한 도면이다.
- [26] 도 7은 본 발명의 일 실시예에 따른 영상 분석 방법의 흐름도이다.
- [27] 도 8은 본 발명의 일 실시예에 따른 영상 분석 모델 생성 방법의 흐름도이다.
- [28] 도 9는 본 발명의 일 실시예에 따른 영상 분석 방법의 효과를 검증하기 위한 실험에서 손글씨 영상으로 구성되는 입력을 설명하기 위한 도면이다.
- [29] 도 10은 도 9의 입력에 대한 본 발명의 일 실시예에 따른 영상 분석 방법의 효과를 검증하기 위한 실험에서 분석 정확도를 나타낸 그래프이다.
- [30] 도 11은 도 10의 대조군으로서, 종래 영상 분석 방법의 분석 정확도를 나타낸 그래프이다.
- [31] 도 12는 본 발명의 일 실시예에 따른 영상 분석 방법의 효과를 검증하기 위한 실험에서 CCTV 영상으로 구성되는 입력을 설명하기 위한 도면이다.
- [32] 도 13은 도 12의 입력에 대한 본 발명의 일 실시예에 따른 영상 분석 방법의 효과를 검증하기 위한 실험에서 분석 정확도를 나타낸 그래프이다.

발명의 실시를 위한 형태

- [33] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다.
- [34] 본 발명의 실시예들을 설명함에 있어서 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략할 것이다. 그리고 후술되는 용어들은 본 발명의 실시예에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.
- [35] 도 1은 본 발명의 일 실시예에 따른 영상 분석 장치의 기능 블록도이다.
- [36] 영상 분석 장치는 입력된 타겟 영상에 대한 분석을 수행하는 장치를 의미한다. 여기서, 타겟 영상이란 영상 분석의 대상이 되는 영상으로서, 정지 영상 및 정지 영상의 스트림(Stream)으로 구현되는 동영상을 포함할 수 있다. 또한, 영상 분석이란 타겟 영상 내 객체와 객체 이외의 배경에 대한 정보를 획득하는 모든 과정을 의미하며, 영상 내 객체의 분류(Classification) 및/또는 검출(Detection)을

포함할 수 있다.

- [37] 영상 분석의 정확도를 높이기 위해, 최근에는 영상 분석에 있어 딥러닝과 같은 기계 학습을 적용하려는 시도가 활발히 진행되고 있다. 딥러닝 기술이 적용된 영상 분석 모델의 예로 R-CNN, Faster R-CNN, Mask R-CNN, SSD, DSSD, Yolo 등이 있다.
- [38] 딥러닝 기술이 적용된 영상 분석 모델은 한정된 수의 객체를 분류 및/또는 검출하기 위해 오랜 시간 동안 수집된 학습 데이터베이스(DB)를 학습할 필요가 있다. 그러나, CCTV 서비스, 자율 주행 등과 같이 실제 서비스에 딥러닝 기반 영상 분석 모델을 적용하기 위해서는 그 환경에 맞는 학습 데이터베이스가 사전 구축되어야 하고, 구축되는 학습 데이터베이스는 한 객체 당 적어도 수만 장 이상의 학습 영상을 포함해야 한다.
- [39] 이와 같이, 분석 가능한 객체의 수에 비례하여 학습 영상의 수도 증가하므로, 실제 서비스에 적용되는 영상 분석 모델이 모든 객체를 분류 및/또는 검출하도록 학습되는 것은 현실적으로 불가능하다. 따라서, 실 사용에 제공되는 영상 분석 모델은 제한된 수의 객체에 대해서만 분류 및/또는 검출이 가능할 수 있다.
- [40] 만약, 소정 개수의 객체에 대하여 동작하는 영상 분석 모델을 이용하여 영상 분석 서비스를 제공하는 중 신규 객체에 대한 영상 분석이 필요한 경우, 다음의 세 가지 방법을 따를 수 있다. 이하에서는 사람을 객체로 하는 영상 분석 모델에 대하여 신규 객체인 자동차를 추가하고자 하는 경우를 가정하여 설명한다.
- [41] 첫 번째 방법은 사람에 대한 영상 분석 모델과는 별개로 자동차에 대한 영상 분석 모델을 학습하여 생성하는 것이다. 예를 들어, 다양한 자동차 영상을 수집, 학습하여 영상 분석 모델로서 '자동차 검출' 딥러닝 네트워크 모델을 생성할 수 있다. 이 경우, 사람과 자동차 모두에 대한 영상 분석이 가능하지만, 사람에 대한 영상 분석 모델과 자동차에 대한 영상 분석 모델 모두를 가동시켜야 하기 때문에, 하드웨어 리소스가 2배 이상 발생할 수 있다.
- [42] 두 번째 방법은 기존의 사람에 대한 영상 분석 모델의 학습에 이용된 학습 DB와 함께 새로운 자동차 학습 DB를 수집하여, 사람 및 자동차 모두를 객체로 하는 영상 분석 모델을 학습하는 것이다. 이 방법은 새롭게 자동차 학습 DB를 수집한다는 점이 첫 번째 방법과 유사하지만, 사람에 대한 영상 분석 모델의 학습에 이용된 학습 DB가 존재하는 경우에만 이용할 수 있다는 제약이 존재한다. 또한, 학습 완료 후 영상 분석 모델의 사람에 대한 분석 성능이 최초의 영상 분석 모델의 성능보다 감소할 가능성이 있다.
- [43] 세 번째 방법은 기존의 사람에 대한 영상 분석 모델 상에 자동차 DB를 학습하는 것이다. 이 방법은 기 사용된 사람에 대한 학습 DB가 필요 없다는 것이 장점이나, 사람에 대한 영상 분석 모델에 자동차 DB를 학습 시 결과물은 사람에 대한 분석 능력이 저하되는 Catastrophic Forgetting 현상이 발생한다는 점이다. 즉, 이미 생성된 사람에 대한 영상 분석 모델 상에 새로운 객체인 자동차의 학습 DB를 학습 시키면, 학습 결과에 따른 영상 분석 모델은 자동차에 대한 영상 분석

성능이 우수해지는 반면, 사람에 대한 영상 분석 성능은 저하될 수 있다.

[44] 이처럼 영상 분석 모델의 생성에 이용된 학습 DB와는 상이한 환경에서의 영상이 입력될 경우 상술한 세가지 방법 모두 분석 결과의 정확성을 담보하기 어렵다. 특히, 세 번째 방법의 경우 영상 분석 모델의 생성에 이용된 학습 DB와는 무관하게 생성된 영상 분석 모델만을 이용한다는 점에서 첫 번째 및 두 번째 방법에 비해 장점을 가짐에도 불구하고, 신규 객체에 대한 학습을 수행할수록 기존 객체에 대한 영상 분석 성능이 저하되므로 실사용 환경에 적용하기 어렵다.

[45] 이를 해결하기 위해, 본 발명의 일 실시예에 따른 영상 분석 장치는 클러스터링 알고리즘 중 하나인 DPMM(Dirichlet Process Mixture Model)의 기본 개념을 차용할 수 있다. 이하에서는 DPMM에 대하여 간략히 설명한다.

[46] DPMM은 클러스터의 수를 미리 정하지 않은 상태에서 클러스터링을 하는 클러스터링 알고리즘을 의미한다. 대표적인 클러스터링 알고리즘인 k-means 클러스터링 알고리즘이 클러스터의 수 k를 하이퍼파라미터(Hyper-Parameter)로 두는 것과 달리, DPMM의 하이퍼파라미터는 민감도 α 이다. DPMM은 민감도와 데이터에 따라 클러스터링 결과가 결정되며, α 의 값이 클수록 클러스터의 개수가 증가하는 경향이 있다.

[47] DPMM을 학습하는 대표적인 방법은 Gibbs Sampling 방법이다. 처음에 N개의 데이터가 주어지면 먼저 각각의 데이터에 임의로 클러스터를 할당하여 초기화를 진행한다. 이 때, 클러스터를 할당하는 방식은 학습 시간만 충분히 주어진다면 결과에 영향을 주지 않는다. 초기화가 완료되면 매 스텝마다 하나의 데이터를 임의로 선택하여 클러스터를 재할당한다. 이 때, 임의로 선택된 데이터는 새로운 클러스터에 할당이 될 수도 있고, 이미 존재하는 클러스터에 할당될 수도 있다. 다시 할당할 데이터를 x_i 라 하면, x_i 가 기존 클러스터 c에 할당될 확률 P_1 과 새로운 클러스터에 할당될 확률 P_2 는 수학식 1 및 수학식 2에 의해 정의된다.

[48] [수식1]

$$P_1 = b \frac{N_c}{N-1+\alpha} F(x_i, \phi_c)$$

[49] [수식2]

$$P_2 = b \frac{\alpha}{N-1+\alpha} \int F(x_i, \psi) dG_0(\psi)$$

[50] 여기서, α 는 민감도를 의미하고, b 는 정규화 상수를 의미하고, N 은 전체 데이터의 개수를 의미하고, N_c 는 클러스터 c에 할당되어 있는 데이터의 개수를 의미할 수 있다. 또한, Φ_c 는 클러스터 c의 파라미터로서, 각 클러스터가 정규 분포를 따를 때 평균과 분산을 의미할 수 있다.

[51] $F(x_i, \Phi_c)$ 는 Φ_c 에 대한 데이터 x_i 의 우도(Likelihood)를 의미하고,

$$\frac{N_c}{N-1+\alpha}$$

는 클러스터 c에 대한 사전 확률(Prior Probability)을 의미한다. 나아가, $G_0(\Phi)$ 는 또 다른 하이퍼파라미터인 Base Distribution으로서, Φ 의 분포에 대한 사전

확률을 의미할 수 있다.

- [52] 상술한 확률에 따라 새로운 클러스터가 할당되면, 각 클러스터의 파라미터 Φ_c 를 조정할 수 있다. 구체적으로, EM(Expectation maximization) 알고리즘의 최대화 단계(Maximization Step)와 같이, 해당 클러스터에 할당되어 있는 데이터들을 이용하여 최대 우도 추정(Maximum Likelihood Estimate, MLE) 또는 최대 사후 확률(Maximum A Posteriori, MAP) 추정에 따라 Φ_c 를 조정할 수 있다.
- [53] 이러한 DPMM 클러스터링 알고리즘이 클러스터의 수를 자동으로 결정하는 것처럼, 본 발명의 일 실시예에 따른 영상 분석 장치는 영상 분석 모델을 구성하는 엑스퍼트의 개수를 입력되는 학습 영상에 기초하여 자동으로 결정할 수 있다.
- [54] 도 1을 참조하면, 본 발명의 일 실시예에 따른 영상 분석 장치(100)는 영상 분석부(110), 결정부(120), 메모리(130) 및 학습부(140)를 포함할 수 있다.
- [55] 영상 분석부(110)는 적어도 하나의 엑스퍼트(Expert)로 구성되는 영상 분석 모델을 기초로 타겟 영상을 분석할 수 있다. 여기서, 엑스퍼트는 영상 분석을 위한 개별 모델을 의미할 수 있다.
- [56] 영상 분석부(110)가 이용하는 영상 분석 모델은 학습 영상을 학습하여 생성된 적어도 하나의 엑스퍼트로 구성되고, 영상 분석 장치(100)는 입력된 타겟 영상에 대응되는 적어도 하나의 엑스퍼트 각각에 타겟 영상을 입력하여, 출력된 출력 값을 기초로 타겟 영상을 분석할 수 있다.
- [57] 영상 분석 모델을 구성하는 적어도 하나의 엑스퍼트 중 입력된 타겟 영상에 대응되는 엑스퍼트를 이용하면, 영상 분석 결과의 정확도를 높일 수 있다. 특히, 타겟 영상과 유사한 학습 영상을 학습하여 생성된 엑스퍼트일수록, 효과적으로 타겟 영상을 분석할 수 있다.
- [58] 이에 기초하여, 적어도 하나의 엑스퍼트와 게이트 네트워크(Gate Network)를 결합한 MoE(Mixture of Experts)가 영상 분석 모델로 이용될 수 있다. 게이트 네트워크는 입력 데이터와 각각의 엑스퍼트의 성능을 기초로, 출력 값을 획득할 엑스퍼트를 결정할 수 있다. 구체적으로, MoE는 k번째 엑스퍼트에서 입력 값 x 와 k번째 엑스퍼트 E_k 에 대한 클래스(Class) y 의 확률 $p(y|x, E_k)$ 를 모델링하고, 게이트 네트워크에서 입력 값 x 에 대한 k번째 엑스퍼트 E_k 의 확률 $p(E_k|x)$ 를 모델링할 수 있다. 그 결과, MoE는 수학적 식 3에 따라 입력 값 x 에 대한 클래스 y 를 구할 수 있다.
- [59] [수식3]
- $$p(y|x) = \sum_k p(y|x, E_k) p(E_k|x)$$
- [60] 수학적 식 3을 참조하면, 게이트 네트워크는 엑스퍼트가 입력 값 x 를 입력 받기에 앞서, 입력 값 x 에 대응되는 엑스퍼트를 찾는 일종의 분류기(Classifier) 역할을 수행할 수 있다. 이 때, 특정 시점에 새로운 엑스퍼트를 생성하기 위해서는 해당 엑스퍼트에 대응되는 입력 값들을 새로 학습할 필요가 있으므로, 상술한

Catastrophic Forgetting 현상이 발생할 가능성이 있다. 다시 말해, 새로운 엑스퍼트에 대응되는 데이터를 학습하느라 영상 분석 모델이 기존에 학습한 내용을 잊어버릴 수 있다는 것이다.

- [61] 이를 해결 하기 위해, 본 발명의 일 실시예에 따른 영상 분석부(110)는 게이트 기반 MoE와 상이한 영상 분석 모델을 이용할 수 있다. 이하에서는 도 2 및 도 3을 참조하여, 본 발명의 일 실시예에 따른 영상 분석 모델을 설명한다.
- [62] 도 2는 본 발명의 일 실시예에 따른 영상 분석 모델의 출력 값 출력 과정을 설명하기 위한 도면이고, 도 3은 본 발명의 일 실시예에 따른 영상 분석 모델의 엑스퍼트를 설명하기 위한 도면이다.
- [63] 본 발명의 일 실시예에 따른 영상 분석부(110)는 입력 값 x 와 클래스 y 를 하나의 데이터 포인트(Data Point)로 설정하고, 이에 상술한 DPMM 클러스터링 알고리즘을 적용하는 영상 분석 모델을 이용할 수 있다. 이와 같은 영상 분석 모델은 각각의 엑스퍼트가 담당하는 데이터들을 클러스터링 하여, 적절한 클러스터를 찾도록 동작할 수 있다.
- [64] 이하에서는 이와 같은 영상 분석 모델을 DPMoE(Dirichlet Process Mixture of Experts) 영상 분석 모델이라 명명한다.
- [65] DPMoE 영상 분석 모델은 적어도 하나의 엑스퍼트 각각에 입력 값 x 를 입력할 수 있다. 도 2에서는 DPMoE 영상 분석 모델이 첫 번째 엑스퍼트인 E_1 , 두 번째 엑스퍼트인 E_2 및 세 번째 엑스퍼트인 E_3 로 구성되는 경우를 예시하며, 세 개의 엑스퍼트 E_1 , E_2 및 E_3 각각에 입력 값 x 가 입력될 수 있다.
- [66] 상술한 DPMM 클러스터링 알고리즘을 적용하기 위해, DPMoE 영상 분석 모델은 수학적 식 1의 $F(x_i, \Phi_j)$ 를 k 번째 엑스퍼트 E_k 에 대한 입력 값 x 및 클래스 y 의 확률 $P(x, y|E_k)$ 로 치환할 수 있다. 이를 위해, DPMoE 영상 분석 모델의 적어도 하나의 엑스퍼트 각각은 입력 값 x 와 k 번째 엑스퍼트 E_k 에 대한 클래스 y 의 확률 $p(y|x, E_k)$ 를 예측하는 분류 모델(Discriminative Model)과 k 번째 엑스퍼트 E_k 에 대한 입력 값 x 의 확률 $p(x|E_k)$ 를 예측하는 생성 모델(Generative Model)로 구성될 수 있다.
- [67] 도 2를 참조하면, k 번째 엑스퍼트 E_k 는 분류 모델 C_k 와 생성 모델 V_k 로 구성될 수 있다. 분류 모델 C_k 는 입력된 입력 값 x 에 대응되는 클래스 y 를 분류 및/또는 검출하는 분류기(Classifier) 및/또는 검출기(Detector)로 구현될 수 있으며, 출력 값으로서 확률 $p(y|x, E_k)$ 를 출력할 수 있다.
- [68] 일 실시예에 따른 분류 모델 C_k 는 딥러닝과 같은 기계 학습을 통해 학습됨으로써, CNN(Convolution Neural Network), DNN(Deep Neural Network), RNN(Recurrent Neural Network), BRDNN(Bidirectional Recurrent Deep Neural Network)과 같은 신경망 모델로 구현될 수 있다.
- [69] 생성 모델 V_k 는 k 번째 엑스퍼트 E_k 의 학습에 이용된 입력 값 x 와의 유사도를 획득하기 위해 VAE(Variational Auto Encoder) 등으로 구현될 수 있으며, 출력 값으로서 확률 $p(x|E_k)$ 를 출력할 수 있다. 이 때, 확률 $p(x|E_k)$ 를 정확히 예측하는

것이 어려울 수 있으므로, 본 발명의 일 실시예에 따른 DPMoE 영상 분석 모델의 생성 모델 V_k 는 확률 $p(x|E_k)$ 의 하한, 예를 들어 ELBO(Evidence Lower Bound)를 구한 후, 구해진 하한 값을 확률 $p(x|E_k)$ 로 추정할 수 있다.

[70] 다른 실시예에 따른 생성 모델 V_k 는 딥러닝과 같은 기계 학습에 의해 학습됨으로써, CNN, DNN, RNN, BRDNN과 같은 신경망 모델로 구현될 수 있다.

[71] 그 결과, DPMoE 영상 분석 모델의 적어도 하나의 엑스퍼트는 분류 모델 C_k 에 의해 출력된 확률 $p(y|x, E_k)$ 과 생성 모델 V_k 에 의해 출력된 확률 $p(x|E_k)$ 를 곱하여, 각각의 엑스퍼트 E_k 에 대한 입력 값 x 및 클래스 y 의 확률 $P(x, y|E_k)$ 를 출력할 수 있다. 도 2에서는, DPMoE 영상 분석 모델을 구성하는 엑스퍼트 E_1, E_2 및 E_3 가 분류 모델 C_1, C_2 및 C_3 각각에 의해 출력된 확률 $p(y|x, E_1), p(y|x, E_2)$ 및 $p(y|x, E_3)$ 각각과 생성 모델 V_k, V_2 및 V_3 각각에 의해 출력된 확률 $p(x|E_1), p(x|E_2)$ 및 $p(x|E_3)$ 각각을 곱하여, 엑스퍼트 E_1 에 대한 입력 값 x 및 클래스 y 의 확률 $P(x, y|E_1)$, 엑스퍼트 E_2 에 대한 입력 값 x 및 클래스 y 의 확률 $P(x, y|E_2)$ 및 엑스퍼트 E_3 에 대한 입력 값 x 및 클래스 y 의 확률 $P(x, y|E_3)$ 를 출력하는 경우를 예시한다.

[72] 또한, DPMoE 영상 분석 모델은 수학적 식 1의

$$\frac{N_k}{N + \alpha}$$

은 k 번째 엑스퍼트 E_k 일 사전 확률 $P(E_k)$ 로 치환할 수 있다. 이 때, 사전 확률 사전 확률 $P(E_k)$ 는 DPMM과 유사하게 k 번째 엑스퍼트 E_k 가 학습한 데이터에 수에 비례하도록 결정될 수 있으며, 구체적으로 수학적 식 4를 따를 수 있다.

[73] [수식4]

$$P(E_k) = \frac{N_k}{N + \alpha}$$

[74] 여기서, N 은 전체 데이터 수를 의미하고, N_k 는 k 번째 엑스퍼트 E_k 에서 학습한 데이터의 수를 의미하고, α 는 민감도를 의미할 수 있다.

[75] 상술한 바에 따라 DPMM 클러스터링 알고리즘을 적절히 변형하면, DPMoE 영상 분석 모델은 입력 값 x 에 대한 클래스 y 의 확률 $p(y|x)$ 을 출력 값으로서 출력할 수 있다. 구체적으로, DPMoE 영상 분석 모델은 이렇게 획득된 각각의 엑스퍼트에 대한 확률 $P(x, y|E_k)$ 과 $p(E_k)$ 을 곱한 후, 모두 더함으로써 입력 값 x 및 클래스 y 의 확률 $P(x, y)$ 를 구할 수 있다. 이렇게 구해진 $P(x, y)$ 을 이용하여, DPMoE 영상 분석 모델은 수학적 식 5에 따라 입력 값 x 에 대한 클래스 y 의 확률 $p(y|x)$ 을 출력 값으로서 출력할 수 있다.

[76] [수식5]

$$p(y|x) = \frac{p(x, y)}{p(x)} = \frac{\sum_k p(x, y|E_k)p(E_k)}{\sum_k p(x|E_k)p(E_k)}$$

[77] 최종적으로, 영상 분석부(110)는 입력 값 x , 즉 입력된 타겟 영상에 대한, 클래스 y , 즉 타겟 영상의 분석 결과의 확률을 획득할 수 있다. 만약, 영상 분석 모델이 분류기로 구현되는 경우, 영상 분석부(110)는 타겟 영상 내 객체의 종류에 따라 타겟 영상 분류 결과를 획득할 수 있다. 이와는 달리, 영상 분석 모델이 검출기로 구현되면, 영상 분석부(110)는 타겟 영상 내 특정 객체의 존재와 위치에 따라

타겟 영상 내 객체 검출 결과를 획득할 수 있다.

- [78] 결정부(120)는 영상 분석 모델에 의해 타겟 영상에 대한 분석 결과의 유효성을 결정할 수 있다. 구체적으로, 결정부(120)는 타겟 영상이 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률에 기초하여, 타겟 영상에 대한 분석 결과의 유효성을 결정할 수 있다. 여기서, 입력 카테고리란 해당 엑스퍼트의 학습에 이용된 학습 영상과 동종 영상으로 판단할 수 있는 카테고리를 의미하며, 객체의 종류, 객체의 위치, 영상 획득 환경 등에 의해 결정될 수 있다.
- [79] 또한, 결정부(120)는 타겟 영상에 대한 분석 결과가 유효하지 않은 것으로 결정되면, 타겟 영상을 신규 카테고리로 분류할 수 있다. 나아가, 결정부(120)는 신규 카테고리로 분류된 타겟 영상을 메모리(130)에 저장할 수 있다. 아울러, 학습부(140)는 메모리(130)에 저장된 타겟 영상을 학습하여 신규 카테고리에 대한 신규 엑스퍼트를 생성할 수 있다.
- [80] 이하에서는 도 4 내지 도 6을 참조하여, 신규 카테고리에 대한 엑스퍼트를 생성하는 방법을 설명한다.
- [81] 도 4 내지 도 6은 일 실시예에 따른 영상 분석 장치에 의해 수행되는 신규 카테고리에 대한 엑스퍼트 생성 방법을 설명하기 위한 도면이다.
- [82] 타겟 영상의 분석이 완료되면, 결정부(120)는 타겟 영상이 영상 분석 모델의 적어도 하나의 엑스퍼트의 입력 카테고리에 속하는지, 아니면 신규 카테고리에 속하는지 판단할 수 있다. 이를 위해, 결정부(120)는 타겟 영상이 입력 카테고리에 속할 확률 P_{11} 과 타겟 영상이 신규 카테고리에 속할 확률 P_{22} 를 구할 수 있다.
- [83] 이 때, 본 발명의 일 실시예에 따른 영상 분석 모델은 DPMM 클러스터링 알고리즘을 채용하는 바, 타겟 영상이 입력 카테고리에 속할 확률 P_{11} 은 수학식 1의 x_1 가 기존 클러스터 c 에 할당될 확률 P_1 에 대응되고, 타겟 영상이 신규 카테고리에 속할 확률 P_{22} 는 수학식 2의 x_1 가 새로운 클러스터에 할당될 확률 P_2 에 대응될 수 있다. 따라서, 결정부(120)는 타겟 영상이 입력 카테고리에 속할 확률 P_{11} 을

$$\frac{N_k}{N-\alpha} P(x, y | E_k)$$

(여기서, N 은 전체 데이터 수를 의미하고, N_k 는 k 번째 엑스퍼트 E_k 에서 학습한 데이터의 수를 의미하고, α 는 민감도를 의미함.)에 비례하여 결정하고, 타겟 영상이 신규 카테고리에 속할 확률 P_{22} 를

$$\frac{\alpha}{N-\alpha} P(x, y | E_0)$$

(여기서, N 은 전체 데이터 수를 의미하고, α 는 민감도를 의미하고, E_0 는 신규 카테고리를 가지는 신규 엑스퍼트를 의미함)에 비례하여 결정할 수 있다. 도 4의 경우, 결정부(120)는 하나의 엑스퍼트 E_1 의 입력 카테고리에 속할 확률 P_{11} 이

$$\frac{N_1}{N-\alpha} P(x, y | E_1)$$

에 비례함을 이용하고, 신규 카테고리에 속할 확률 P_{22} 가

$$\frac{\alpha}{N-\alpha} P(x, y | E_0)$$

에 비례함을 이용하는 경우를 예시한다.

- [84] 확률 P_{11} 과 확률 P_{22} 가 결정되면, 결정부(120)는 결정된 확률을 기초로 입력된 타겟 영상에 대한 분석 결과의 유효성을 랜덤(Random)하게 결정할 수 있다. 도 4에서 확률 P_{11} 이 확률 P_{22} 보다 높다면, 결정부(120)는 타겟 영상에 대한 분석 결과를 유효한 것으로 결정할 가능성이 높으나, 분석 결과가 유효하지 않은 것으로 결정한 후, 신규 카테고리 분류할 수도 있다. 이와는 달리, 확률 P_{22} 가 확률 P_{11} 보다 높다면, 결정부(120)는 타겟 영상에 대한 분석 결과를 유효하지 않은 것으로 결정할 가능성이 높으나, 분석 결과를 유효한 것으로 결정할 수도 있다.
- [85] 결정부(120)에 의해 타겟 영상에 대한 분석 결과가 유효한 것으로 결정되면, 분석 결과 자체를 사용자에게 제공하거나, 이를 활용하여 관련 서비스를 사용자에게 제공할 수 있다.
- [86] 반면, 타겟 영상에 대한 분석 결과가 유효하지 않은 것으로 결정되어, 해당 타겟 영상이 신규 카테고리 분류되면, 결정부(120)는 이를 메모리(130)에 저장할 수 있다. 만약, 입력되는 타겟 영상이 매 프레임마다 프레임 영상이 입력되는 데이터 스트림 형태인 경우, 결정부(120)는 매 프레임마다 타겟 영상에 대한 분석 결과의 유효성을 결정한 후, 분석 결과가 유효하지 않은 것으로 결정된 타겟 영상을 메모리(130)에 순차적으로 저장할 수 있다. 도 5를 참조하면, 메모리(130)는 순차적으로 입력되는 신규 카테고리에 대한 타겟 영상을 누적하여 저장할 수 있다.
- [87] 이를 위해, 본 발명의 일 실시예에 따른 메모리(130)는 램(Random Access Memory, RAM), SRAM(Static Random Access Memory), 롬(Read-Only Memory, ROM), PROM(Programmable Read-Only Memory), EEPROM(Electrically Erasable Programmable Read-Only Memory)와 같은 저장매체로 구현되거나, 클라우드 서버(Cloud Server)에 의해 가상 저장 공간으로서 구현될 수 있다. 이와는 달리 메모리(130)는 단기 기억 장치(Shortterm Memory)로 구현될 수 있다.
- [88] 학습부(140)는 메모리(130)에 저장된 타겟 영상을 학습하여 신규 카테고리에 대한 엑스퍼트를 생성할 수 있다. 이를 위해, 일 실시예에 따른 학습부(140)는 메모리(130)에 저장된 타겟 영상의 누적 용량이 미리 정해진 기준 용량 이상인 경우, 메모리(130)에 저장된 타겟 영상을 학습할 수 있다. 여기서, 기준 용량이란 신규 엑스퍼트를 생성하기에 충분한 타겟 영상 누적 용량의 최소 값을 의미할 수 있다.
- [89] 이와는 달리, 다른 실시예에 따른 학습부(140)는 미리 정해진 기준 시간이 초과하면, 메모리(130)에 저장된 타겟 영상을 학습할 수도 있다. 여기서, 기준 시간이란 신규 엑스퍼트를 생성하기에 충분한 타겟 영상 누적 시간의 최소 값을 의미할 수 있다. 또 다른 실시예에 따른 학습부(140)는 상술한 기준 시간이 초과하면, 메모리(130)에 저장된 타겟 영상의 누적 용량이 기준 용량 이상인 경우에만, 메모리(130)에 저장된 타겟 영상을 학습할 수도 있다.

- [90] 일 실시예에 따른 학습부(140)는 지도 학습(Supervised Learning)에 따라, 신규 엑스퍼트를 생성할 수 있다. 지도 학습 방법은 레이블링(Labeling)된 입력을 학습함으로써 수행되므로, 지도 학습에 앞서 메모리(130)에 누적된 타겟 영상에 객체에 대한 레이블링이 선행될 필요가 있다. 본 발명의 일 실시예에 따르면, 누적된 타겟 영상 각각에 대하여 외부의 입력을 기초로 레이블링이 수행될 수 있다. 이와는 달리, 누적된 타겟 영상을 레이블링을 위한 별도의 모델, 예를 들어 입력된 영상 내 객체에 대한 분류기 및/또는 검출기에 입력하여, 자동으로 레이블링이 수행될 수도 있다.
- [91] 지도 학습 방법을 따를 때, 학습부(140)는 메모리(130)에 누적된 타겟 영상을 컨볼루션 층에 대한 입력으로 하고, 레이블에 대한 미리 정해진 클래스와 입력을 대응시킴으로써 타겟 영상을 학습할 수 있다. 즉, 학습부(140)는 입력인 타겟 영상과 정답인 객체 사이의 관계를 학습함으로써, 입력되는 타겟 영상의 객체를 분석할 수 있는 신규 엑스퍼트를 생성할 수 있다.
- [92] 이와는 달리, 학습부(140)는 객체에 대한 레이블링이 된 일부 타겟 영상과 레이블링 되지 않은 나머지 타겟 영상을 함께 이용하여 학습하는 반지도 학습(Semisupervised Learning)에 의해 신규 엑스퍼트를 생성할 수도 있다. 또는, 학습부(140)는 학습 결과가 올바른지에 대한 피드백을 이용하는 강화 학습(Reinforcement Learning)에 의해 신규 엑스퍼트를 생성할 수도 있다.
- [93] 도 6에서는 학습부(140)가 신규 카테고리 분류되어 메모리(130)에 누적된 타겟 영상을 기초로 신규 엑스퍼트 E_2 를 생성한 경우를 예시한다.
- [94] 신규 엑스퍼트 E_2 가 생성된 후, 영상 분석부(110)는 새롭게 입력되는 타겟 영상을 신규 엑스퍼트 E_2 를 포함하는 영상 분석 모델을 이용하여 분석할 수 있다. 영상 분석 이후, 결정부(120)는 새로운 타겟 영상에 대한 분석 결과의 유효성을 결정할 수 있다. 구체적으로, 결정부(120)는 새로운 타겟 영상이 입력 카테고리에 속할 확률 P_{11} 과 타겟 영상이 신규 카테고리에 속할 확률 P_{22} 를 구할 수 있다. 이때, 결정부(120)는 입력된 타겟 영상이 각각의 엑스퍼트 E_1 및 E_2 의 입력 카테고리에 속할 확률 P_{11} 을 각각 구할 수 있다. 이렇게 구해진 확률에 기초하여, 결정부(120)는 분석 결과의 유효성을 랜덤하게 결정할 수 있다.
- [95] 지금까지는 신규 카테고리 분류된 타겟 영상을 메모리(130)에 누적하였다가, 학습부(140)가 이를 일시에 학습하는 경우를 설명하였다. 이와는 달리, 다른 실시예에 따른 영상 분석 장치(100)는 별도의 메모리(130) 구성을 생략하고, 타겟 영상이 신규 카테고리 분류되면 곧바로 이를 학습하여 신규 엑스퍼트를 생성하는 것도 가능할 수 있다.
- [96] 다시 도 1을 참조하면, 본 발명의 일 실시예에 따른 영상 분석 장치(100)의 각 구성은 마이크로프로세서(Microprocessor)를 포함하는 연산 장치로 구현될 수 있고, 예를 들어 중앙 처리 장치(Central Processing Unit, CPU) 및 그래픽 처리 장치(Graphic Processing Unit, GPU) 중 적어도 하나로 구현될 수 있다. 이와는 달리, 영상 분석 장치(100)의 각각의 구성 중 적어도 두 개가 SOC(System On

Chip)으로 구현되는 것도 가능할 수 있다.

- [97] 지금까지는 영상 분석 장치(100)의 구성에 대하여 설명하였다. 이하에서는 상술한 영상 분석 장치(100)에 의해 수행되는 영상 분석 방법과 이에 이용되는 영상 분석 모델 생성 방법을 설명한다.
- [98] 도 7은 본 발명의 일 실시예에 따른 영상 분석 방법의 흐름도이다.
- [99] 먼저, 영상 분석 장치(100)는 학습 영상을 학습하여 생성된 적어도 하나의 엑스퍼트로 구성되는 영상 분석 모델에 타겟 영상을 입력할 수 있다(S100). 여기서, 영상 분석 모델은 타겟 영상 내 객체와 객체 이외의 배경에 대한 정보를 획득하는 모든 과정을 의미하며, 영상 내 객체의 분류(Classification) 및/또는 검출(Detection)을 포함하는 영상 분석을 수행하는 모델로서, 학습 영상을 딥러닝과 같은 기계 학습에 의해 학습함으로써 생성될 수 있다. 이 때, 일 실시예에 따른 영상 분석 모델은 입력 값과 클래스를 하나의 데이터 포인트로 설정하고, 이에 DPMM 클러스터링 알고리즘을 적용하도록 구현될 수 있다.
- [100] 영상 분석 모델에 타겟 영상을 입력한 후, 영상 분석 장치(100)는 적어도 하나의 엑스퍼트로부터 출력된 출력 값을 기초로 타겟 영상을 분석할 수 있다(S110). 만약, 영상 분석 모델이 분류기로 구현되는 경우, 영상 분석 장치(100)는 타겟 영상 내 객체의 종류에 따라 타겟 영상 분류 결과를 획득할 수 있다. 이와는 달리, 영상 분석 모델이 검출기로 구현되면, 영상 분석 장치(100)는 타겟 영상 내 특정 객체의 존재와 위치에 따라 타겟 영상 내 객체 검출 결과를 획득할 수 있다.
- [101] 그 다음, 영상 분석 장치(100)는 입력된 타겟 영상이 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률을 획득할 수 있다(S120). 이를 위해, 영상 분석 장치(100)는 먼저 적어도 하나의 엑스퍼트 각각에 대한 입력 및 클래스의 우도를 획득할 수 있다. 구체적으로, 영상 분석 장치(100)는 엑스퍼트의 분류 모델을 이용하여 타겟 영상 및 엑스퍼트에 대한 클래스의 제 1 확률을 획득하고, 엑스퍼트의 생성 모델을 이용하여 엑스퍼트 각각에 대한 입력의 제 2 확률을 획득한 후, 제 1 확률 및 제 2 확률의 곱을 기초로 적어도 하나의 엑스퍼트 각각에 대한 입력 및 클래스의 우도를 획득할 수 있다. 이에 대해서는 도 2 및 3을 통해 설명한 바와 같다.
- [102] 타겟 영상이 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률을 획득한 후, 영상 분석 장치(100)는 입력된 타겟 영상에 대한 분석 결과의 유효성을 결정할 수 있다(S130). 이 때, 영상 분석 장치(100)는 입력된 타겟 영상이 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률에 기초하여, 입력된 타겟 영상의 분석 여부를 랜덤(Random)하게 결정할 수 있다. 즉, 확률에 따라 타겟 영상의 분석 여부의 가능성이 달라지겠지만, 타겟 영상이 입력 카테고리에 속할 확률이 높더라도 타겟 영상을 분석하지 않는 것으로 결정될 수 있고, 타겟 영상이 신규 카테고리에 속할 확률이 높더라도 타겟 영상을 분석하는 것으로 결정될 수도 있다.
- [103] 만약, 타겟 영상에 대한 분석 결과가 유효한 것으로 결정되면, 영상 분석

- 장치(100)는 유효한 영상 분석 결과를 출력할 수 있다(S140). 일 실시예에 따른 영상 분석 장치(100)는 타겟 영상에 대한 분석 결과를 출력하여 직접 사용자에게 제공하거나, 이를 활용하여 관련 서비스를 제공하는 사업자에게 제공할 수 있다.
- [104] 반면, 타겟 영상에 대한 분석 결과가 유효하지 않은 것으로 결정되면, 영상 분석 장치(100)는 타겟 영상을 신규 카테고리 분류할 수 있다(S150). 나아가, 영상 분석 장치(100)는 신규 카테고리 분류된 타겟 영상을 저장한 후, 이를 학습하여 신규 카테고리에 대한 신규 익스퍼트를 생성할 수 있다.
- [105] 이를 통해, 영상 분석 장치(100)는 영상 분석 모델의 생성 이후 신규 카테고리에 대한 타겟 영상이 입력되거나, 학습 영상과는 상이한 획득 조건의 타겟 영상이 입력되는 경우, 이에 대한 신규 카테고리를 용이하게 생성할 수 있다. 특히, Catastrophic Forgetting 현상이 발생하지 않으므로, 영상 분석 모델은 기존의 카테고리 내 타겟 영상에 대한 분석 성능을 유지하면서도, 신규 카테고리 내 타겟 영상에 대한 분석 성능을 새롭게 구비할 수 있다.
- [106] 도 8은 본 발명의 일 실시예에 따른 영상 분석 모델 생성 방법의 흐름도이다.
- [107] 먼저, 영상 분석 장치(100)는 제 1 입력 주기 동안 제 1 학습 영상을 순차적으로 입력 받을 수 있다(S200). 제 1 학습 영상이 입력되면, 영상 분석 장치(100)는 제 1 학습 주기 동안 제 1 학습 영상을 학습하여, 제 1 카테고리에 대한 분류 모델 및 생성 모델로 구성되는 제 1 익스퍼트를 생성할 수 있다(S210). 여기서, 제 1 학습 주기는 제 1 입력 주기 종료 시점으로부터 시작될 수 있다.
- [108] 그다음, 제 2 입력 주기 동안 제 2 학습 영상을 순차적으로 입력 받을 수 있다(S220). 여기서, 제 2 입력 주기는 제 1 학습 주기 종료 시점으로부터 시작될 수 있다. 제 2 학습 영상이 입력되면, 영상 분석 장치(100)는 제 2 학습 영상 각각이 제 1 카테고리에 속할 확률을 획득할 수 있다(S230). 이에 기초하여, 영상 분석 장치(100)는 제 1 익스퍼트에 대하여 제 2 학습 영상 각각을 학습할지 여부를 결정할 수 있다(S240).
- [109] 학습 여부가 결정되면, 영상 분석 장치(100)는 제 2 학습 영상 중 제 1 익스퍼트에 대하여 학습할 것으로 결정된 학습 영상을 제 2 학습 주기 동안 제 1 익스퍼트에 대하여 학습할 수 있다(S240). 여기서, 제 2 학습 주기는 제 2 입력 주기의 종료 시점으로부터 시작될 수 있다.
- [110] 반면, 영상 분석 장치(100)는 제 2 학습 영상 중 제 1 익스퍼트에 대하여 학습하지 않을 것으로 결정된 학습 영상을 제 2 학습 주기 동안 학습하여 제 2 카테고리에 대한 제 2 익스퍼트를 생성할 수 있다(S250).
- [111] 이와 같은 과정을 반복하여 수행함으로써, 영상 분석 장치(100)는 별도의 분류 과정 없이도, 학습 영상을 자동으로 클러스터링 하여 각각의 클러스터 별 익스퍼트를 생성할 수 있다. 즉, 입력 주기와 학습 주기가 반복적으로 발생하며, 영상 분석 장치(100)는 입력 주기에서 신규 카테고리의 학습 영상을 분류하고, 학습 주기에서 분류된 학습 영상을 학습하여 신규 카테고리를 생성할 수 있다. 그 결과, 영상 분석 모델의 영상 분석 정확도가 상승할 수 있다.

- [112] 지금까지는 학습 영상의 학습에 의해 생성된 영상 분석 모델을 이용하여 타겟 영상을 분석하는 장치 및 방법에 대하여 설명하였다. 그러나, 영상을 각종 데이터로 대체하면, 본 발명은 데이터 분석 장치와 방법으로 확장될 수 있다. 예를 들어, 본 발명의 일 실시예에 따른 데이터 분석 장치는 학습 음성의 학습에 의해 생성된 음성 분석 모델을 이용하여 타겟 음성을 분석할 수 있는 음성 분석 장치, 학습 텍스트에 의해 생성된 텍스트 분석 모델을 이용하여 타겟 텍스트를 분석할 수 있는 텍스트 분석 장치 등을 포함할 수 있다.
- [113] 지금까지는 상술한 영상 분석 장치(100)에 의해 수행되는 영상 분석 방법과 이에 이용되는 영상 분석 모델 생성 방법에 대하여 설명하였다. 이하에서는 상술한 영상 분석 방법의 효과에 대하여 설명한다.
- [114] 도 9는 본 발명의 일 실시예에 따른 영상 분석 방법의 효과를 검증하기 위한 실험에서 손글씨 영상으로 구성되는 입력을 설명하기 위한 도면이고, 도 10은 도 9의 입력에 대한 본 발명의 일 실시예에 따른 영상 분석 방법의 효과를 검증하기 위한 실험에서 분석 정확도를 나타낸 그래프이고, 도 11은 도 10의 대조군으로서, 종래 영상 분석 방법의 분석 정확도를 나타낸 그래프이다. 도 10 및 도 11에서 x 축은 학습 진행 시간이고, y 축은 성능, 즉 정확도를 의미한다.
- [115] 도 9를 참조하면, 입력되는 타겟 영상은 MNIST 데이터베이스(Modified National Institute of Standards and Technology database)로부터 획득된 MNIST 데이터 세트로 구성되고, 이는 총 5개의 서브 세트(subset)로 이루어진다. 첫 번째 서브 세트는 0, 및 1, 두 번째 서브 세트는 2 및 3, 세 번째 서브 세트는 4 및 5, 네 번째 서브 세트는 6 및 7, 다섯 번째 서브 세트는 8 및 9의 타겟 영상으로 구성된다.
- [116] 본 실험에서는 상술한 복수의 서브 세트를 순차적으로 3000 단계씩 학습하여, 본 발명의 일 실시예에 따른 영상 분석 방법을 적용하였다. 생성된 영상 분석 모델의 엑스퍼트 각각은 분류 모델로서 MLP 기반의 분류기(Classifier)를 가지고, 생성 모델로서 VAE를 갖도록 설계된다. 또한, 분류기는 단순한 3-layer MLP로 이루어진다.
- [117] 상술한 실험군과 비교하기 위한 대조군으로서, 종래 딥러닝 기반 실시간 객체 검출 네트워크 중 하나인 YOLO ver.2에 대해서도 동일한 실험을 진행한다.
- [118] 도 10을 참조하면, 본 발명의 일 실시예에 따른 영상 분석 모델을 적용할 경우, 각 객체를 학습한 이후에도 이전 객체의 분류 정확도가 거의 떨어지지 않는 것을 확인할 수 있으며, 전체 정확도가 0.87까지 꾸준히 상승하는 것을 볼 수 있다.
- [119] 반면, 도 11을 참조하면, 종래의 영상 분석 방법의 경우, 현재 학습 중인 서브 세트 만을 높은 정확도로 예측하고, 이전에 학습했던 서브 세트들에 대해서는 Catastrophic Forgetting 현상이 발생하는 것을 확인할 수 있다. 그 결과, 전체 정확도가 0.2 부근에서 형성되고 있다.
- [120] 도 12는 본 발명의 일 실시예에 따른 영상 분석 방법의 효과를 검증하기 위한 실험에서 CCTV 영상으로 구성되는 입력을 설명하기 위한 도면이고, 도 13은 도

12의 입력에 대한 본 발명의 일 실시예에 따른 영상 분석 방법의 효과를 검증하기 위한 실험에서 분석 정확도를 나타낸 그래프이다. 도 13에서 x 축은 학습 진행 시간이고, y 축은 성능, 즉 정확도를 의미한다.

- [121] 도 12를 참조하면, 입력되는 타겟 영상은 HDA Person 데이터 세트로부터 획득된 두 개의 서브 세트로 구성된다. 첫 번째 서브 세트(Camera 02)는 사람의 머리 위에서 아래 방향으로 향하는 Top-View 영상으로 구성되며, 이는 사람의 머리 위주로 촬영되어 사람 계수 등에 많이 이용되는 CCTV 각도에 대응되는 영상이다. 두 번째 서브 세트(Camera 17)는 Slanted-View 영상으로 구성되며, 이는 사람의 얼굴뿐만 아니라 몸의 일부가 함께 촬영된 영상으로서, 실제 많은 CCTV가 이와 같은 각도로 영상을 촬영하고 있다.
- [122] 본 실험에서는 먼저 Camera02에 대한 서브 세트를 이용하여 Top-View에서의 사람 객체 검출 네트워크를 학습 후, Camera17에 대한 서브 세트를 이용하여 Slanted-View에서도 사람이 검출될 수 있게 딥러닝 네트워크를 생성하고자 한다. 종래의 검출기는 Top-View에서 사람을 잘 검출하나, 이러한 검출기 상에 Slanted-View 영상을 학습할 경우, Top-View에서는 다시 사람 객체 검출 성능이 매우 저하되는 Catastrophic Forgetting 현상이 발생한 가능성이 높다.
- [123] 도 13에서 좌측 그래프는 Camera02에 대한 서브 세트에 대한 성능 그래프이고, 우측 그래프는 Camera17에 대한 서브 세트에 대한 성능 그래프로, 종래의 Baseline 검출기와 본 발명의 일 실시예에 따른 DPMoE 영상 분석 모델의 성능 변화를 각각 두 번씩 관찰하였다. 좌측 그래프를 보면, Camera17의 서브 세트가 입력되는 시점에서 위의 높은 성능을 보이던 종래 검출기의 정확도가 갑자기 낮아지는 것을 확인할 수 있다. 반면, 본 발명의 일 실시예에 따른 DPMoE 영상 분석 모델은 Camera17의 서브 세트가 입력되더라도, Camera02의 서브 세트에 대한 정확도가 저하되지 않음을 보인다. 우측 그래프를 보면, 추후 입력된 Camera17의 서브 세트에서는 DPMoE 영상 분석 모델뿐만 아니라 종래의 검출기도 정확도가 상승하나, Camera02의 서브 세트에서는 정확도의 차이가 확연하게 나는 것을 확인할 수 있다.
- [124] 상술한 여러 가지 실시예에 따른 영상 분석 장치, 방법 및 이에 이용되는 영상 분석 모델 생성 방법에 따르면, 영상 분석 모델을 구성하는 엑스퍼트의 입력 카테고리에 속하지 않는 타겟 영상을 신규 카테고리로 분류함으로써, 실사용 환경의 다양한 타겟 영상의 분석이 가능한 영상 분석 모델을 관리할 수 있다.
- [125] 학습 영상을 학습하여 생성된 영상 분석 모델을 실사용 환경에서 이용할 경우, 입력되는 타겟 영상은 학습 영상과는 상이한 획득 조건에서 획득되거나, 학습 영상을 통해 학습하지 못한 객체를 포함할 수 있다. 그 결과, 영상 분석 모델은 분석의 정확도가 낮아질 뿐 아니라, 사용 과정에서 발생된 신규 객체 검출의 필요성에 적응적으로 동작할 수 없다.
- [126] 그러나, 본 발명의 일 실시예에서는 타겟 영상을 신규 카테고리로 분류하고, 이를 신규 엑스퍼트 생성에 이용함으로써, 실사용 환경에서 입력될 수 있는

다양한 경우의 타겟 영상의 분석이 가능할 수 있다. 구체적으로, 본 발명의 일 실시예에 따른 영상 분석 장치는 딥러닝과 같은 기계 학습에 의해 생성된 영상 분석 모델 상에 새로운 정보를 추가하는 점진적 학습(Incremental Learning) 및 순차적 학습(Continual Learning) 방법을 채용함으로써, 입력되는 타겟 영상에 적응적으로 동작할 수 있고, 그 결과 실사용 환경에서의 활용성을 높일 수 있다.

[127] 특히, 카메라가 거의 움직이지 않는 고정형 CCTV 환경에서 본 발명의 일 실시예에 따른 IVA(Intelligent Video Analytics) 기술을 상용화할 경우, 설치된 카메라의 위치 및 각도와 무관하게 획득된 영상에 대한 정확한 분석이 수행될 수 있다.

[128] 한편, 상술한 실시예에 따른 영상 분석 방법 및 이에 이용되는 영상 분석 모델 생성 방법에 포함된 각각의 단계는, 이러한 단계를 수행하도록 프로그램된 컴퓨터 프로그램 또는 컴퓨터 프로그램을 기록하는 컴퓨터 판독가능한 기록매체에서 구현될 수 있다.

[129] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 품질에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 발명에 개시된 실시예들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 균등한 범위 내에 있는 모든 기술사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

산업상 이용가능성

[130] 일 실시예에 따르면, 상술한 영상 분석 장치, 방법 및 이에 이용되는 영상 분석 모델의 생성 방법은택내 또는 산업 현장 등 다양한 분야에서 이용될 수 있으므로 산업상 이용 가능성이 있다.

청구범위

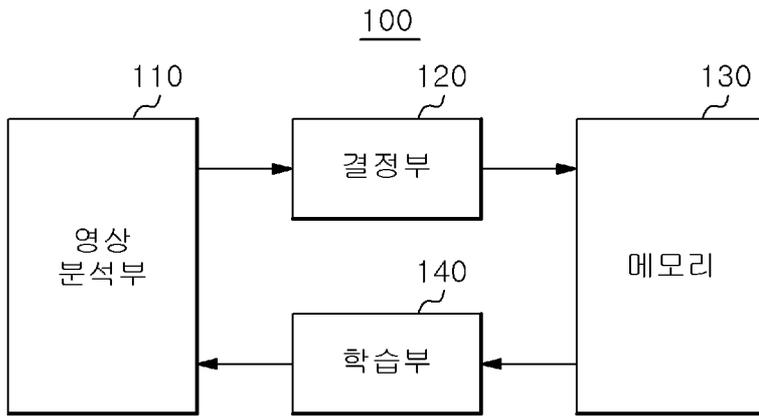
- [청구항 1] 학습 영상을 학습하여 생성된 적어도 하나의 엑스퍼트(Expert)로 구성되는 영상 분석 모델에 타겟 영상을 입력하는 단계;
 상기 적어도 하나의 엑스퍼트로부터 출력된 출력 값을 기초로 상기 타겟 영상을 분석하는 단계;
 상기 입력된 타겟 영상이 상기 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률에 기초하여, 상기 타겟 영상에 대한 분석 결과의 유효성을 결정하는 단계; 및
 상기 타겟 영상에 대한 분석 결과가 유효하지 않은 것으로 결정되면, 상기 타겟 영상을 신규 카테고리로 분류하는 단계를 포함하는 영상 분석 방법.
- [청구항 2] 제 1 항에 있어서,
 상기 타겟 영상에 대한 분석 결과의 유효성을 결정하는 단계는, 상기 입력된 타겟 영상이 상기 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률에 기초하여, 상기 타겟 영상에 대한 분석 결과의 유효성을 랜덤(Random)하게 결정하는 영상 분석 방법.
- [청구항 3] 제 1 항에 있어서,
 상기 타겟 영상에 대한 분석 결과의 유효성을 결정하는 단계는, 상기 적어도 하나의 엑스퍼트 각각에 대한 입력 및 클래스(Class)의 우도(Likelihood)를 획득하는 단계;
 상기 적어도 하나의 엑스퍼트 각각에 대한 사전 확률(Prior Probability)를 획득하는 단계; 및
 상기 획득된 우도와 상기 사전 확률을 기초로 상기 입력된 타겟 영상이 상기 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률을 획득하는 단계를 포함하는 영상 분석 방법.
- [청구항 4] 제 3 항에 있어서,
 상기 적어도 하나의 엑스퍼트 각각에 대한 입력 및 클래스의 우도를 획득하는 단계는,
 상기 적어도 하나의 엑스퍼트 각각의 분류 모델(Discriminative Model)을 이용하여, 상기 입력된 타겟 영상 및 상기 적어도 하나의 엑스퍼트 각각에 대한 클래스의 제 1 확률을 획득하는 단계;
 상기 적어도 하나의 엑스퍼트 각각의 생성 모델(Generative Model)을 이용하여, 상기 적어도 하나의 엑스퍼트 각각에 대한 입력의 제 2 확률을 획득하는 단계; 및
 상기 제 1 확률 및 상기 제 2 확률을 기초로 상기 적어도 하나의 엑스퍼트

- 각각에 대한 입력 및 클래스의 우도를 획득하는 단계를 포함하는 영상 분석 방법.
- [청구항 5] 제 4 항에 있어서,
상기 제 2 확률을 획득하는 단계는,
상기 제 2 확률의 하한(Lower Bound)을 상기 제 2 확률로서 추정하는 단계를 포함하는 영상 분석 방법.
- [청구항 6] 제 1 항에 있어서,
상기 신규 카테고리로 분류된 상기 타겟 영상을 메모리에 저장하는 단계; 및
상기 메모리에 저장된 타겟 영상의 누적 용량이 미리 정해진 기준 용량 이상이면, 상기 메모리에 저장된 타겟 영상을 학습하여 상기 신규 카테고리에 대한 엑스퍼트를 생성하는 단계를 더 포함하는 영상 분석 방법.
- [청구항 7] 제 6 항에 있어서,
상기 신규 카테고리에 대한 엑스퍼트를 생성하는 단계는,
상기 메모리에 저장된 타겟 영상에 레이블링(Labeling)을 수행하는 단계; 및
상기 레이블링된 타겟 영상을 학습하여 상기 신규 카테고리에 대한 엑스퍼트를 생성하는 단계를 포함하는 영상 분석 방법.
- [청구항 8] 학습 영상을 학습하여 생성된 적어도 하나의 엑스퍼트(Expert)로 구성되는 영상 분석 모델을 기초로 타겟 영상을 분석하는 영상 분석부; 및
상기 타겟 영상이 상기 적어도 하나의 엑스퍼트의 입력 카테고리에 속할 확률에 기초하여, 상기 타겟 영상에 대한 분석 결과의 유효성을 결정부를 포함하고,
상기 영상 분석부는
상기 타겟 영상에 대한 분석 결과가 유효하지 않은 것으로 결정되면, 상기 타겟 영상을 신규 카테고리로 분류하는 영상 분석 장치.
- [청구항 9] 제 8 항에 있어서,
상기 신규 카테고리로 분류된 상기 타겟 영상을 저장하는 메모리를 더 포함하는 영상 분석 장치.
- [청구항 10] 제 8 항에 있어서,
상기 메모리에 저장된 타겟 영상의 누적 용량이 미리 정해진 기준 용량 이상이면, 상기 메모리에 저장된 타겟 영상을 학습하여 상기 신규 카테고리에 대한 엑스퍼트를 생성하는 학습부를 더 포함하는

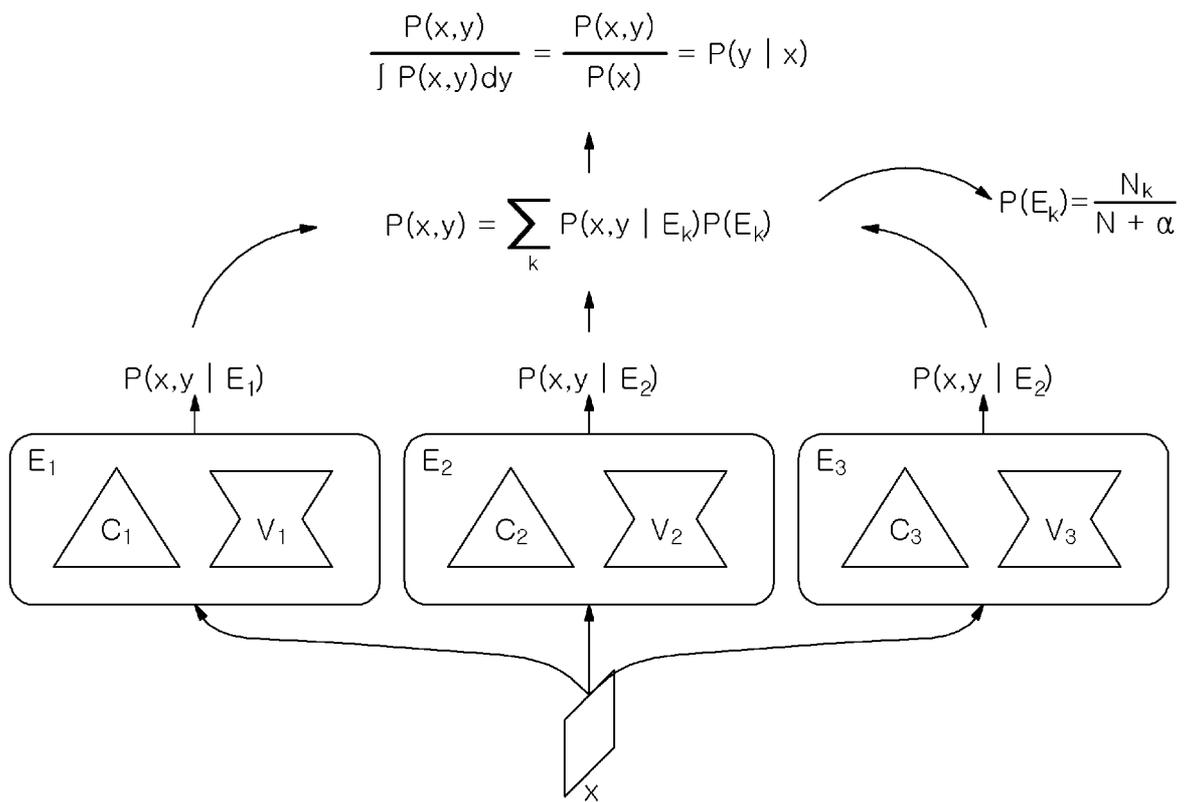
영상 분석 장치.

- [청구항 11] 제 1 입력 주기 동안 제 1 학습 영상을 순차적으로 입력 받는 단계;
 상기 제 1 입력 주기 직후의 제 1 학습 주기 동안 상기 제 1 학습 영상을 학습하여, 제 1 카테고리에 대한 분류 모델(Discriminative Model) 및 생성 모델(Generative Model)로 구성되는 제 1 엑스퍼트(Expert)를 생성하는 단계;
 상기 제 1 학습 주기 직후의 제 2 입력 주기 동안 제 2 학습 영상을 순차적으로 입력 받는 단계;
 상기 제 2 학습 영상이 상기 제 1 카테고리에 속할 확률에 기초하여, 상기 제 1 엑스퍼트에 대하여 상기 제 2 학습 영상 각각을 학습할지 여부를 결정하는 단계; 및
 상기 제 2 입력 주기 직후의 제 2 학습 주기 동안, 상기 제 2 학습 영상 중 상기 제 1 엑스퍼트에 대하여 학습하지 않을 것으로 결정된 학습 영상을 학습하여 제 2 카테고리에 대한 제 2 엑스퍼트를 생성하는 단계를 포함하는
 영상 분석 모델 생성 방법.
- [청구항 12] 제 1 항 내지 제 7 항 중 어느 한 항에 기재된 방법에 따른 각각의 단계를 수행하는 명령어를 포함하는 프로그램이 기록된 컴퓨터 판독 가능 기록매체.

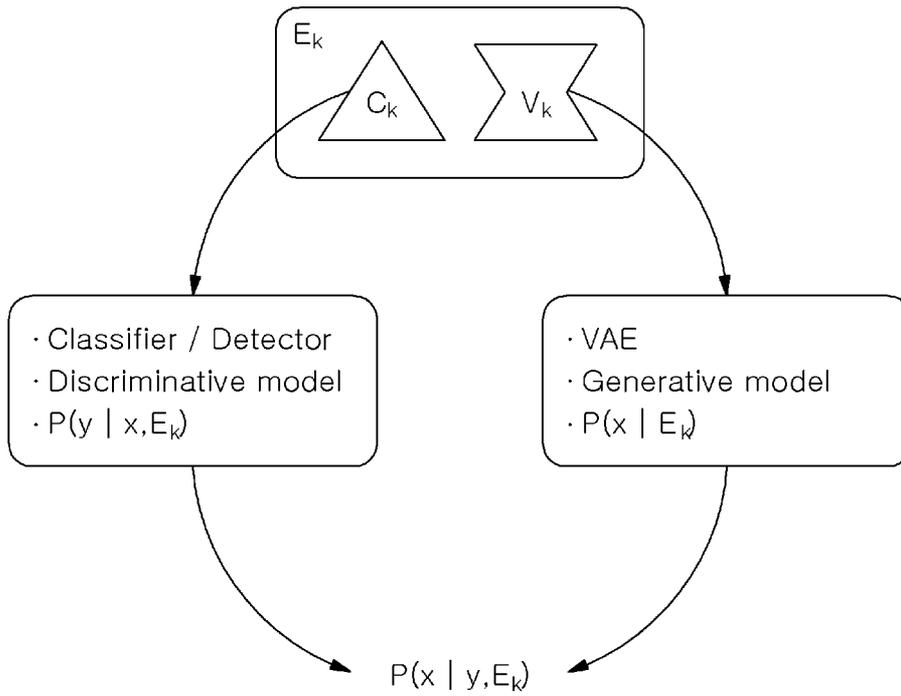
[도1]



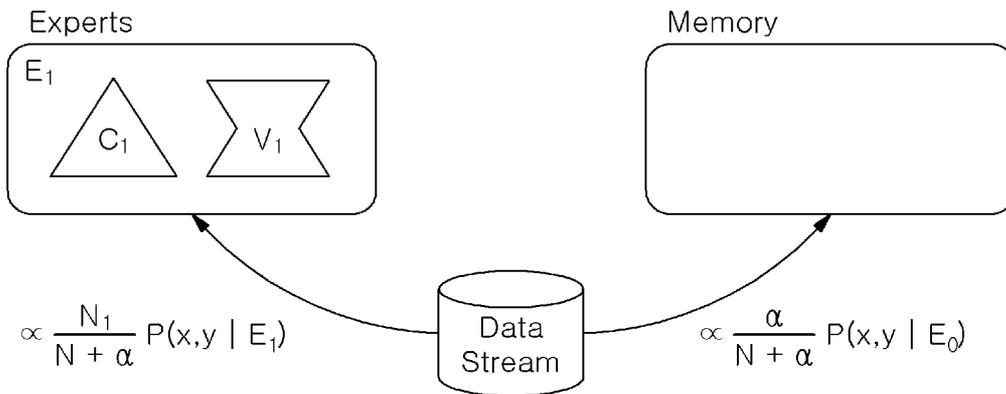
[도2]



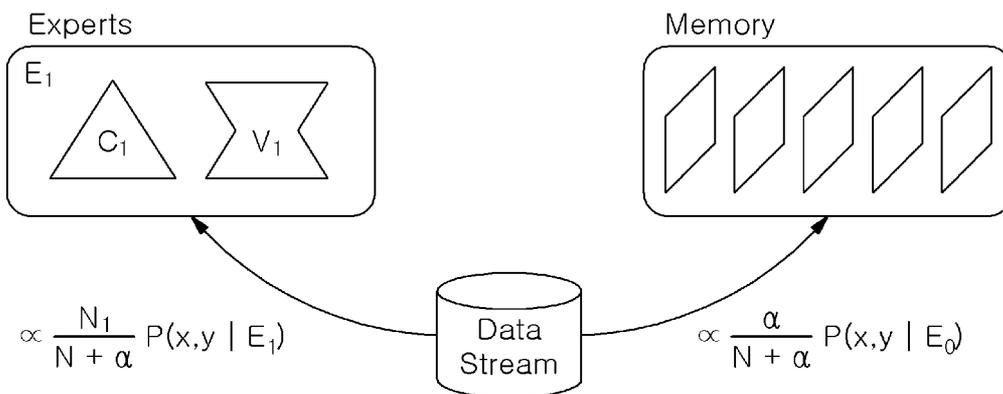
[도3]



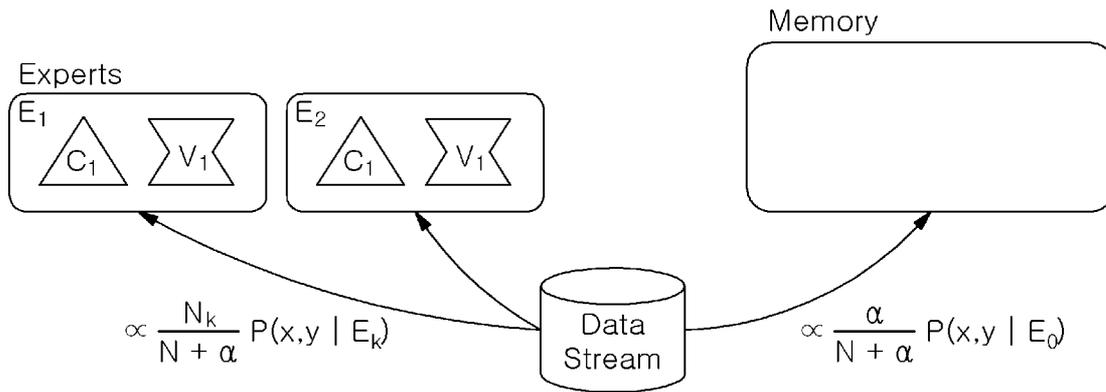
[도4]



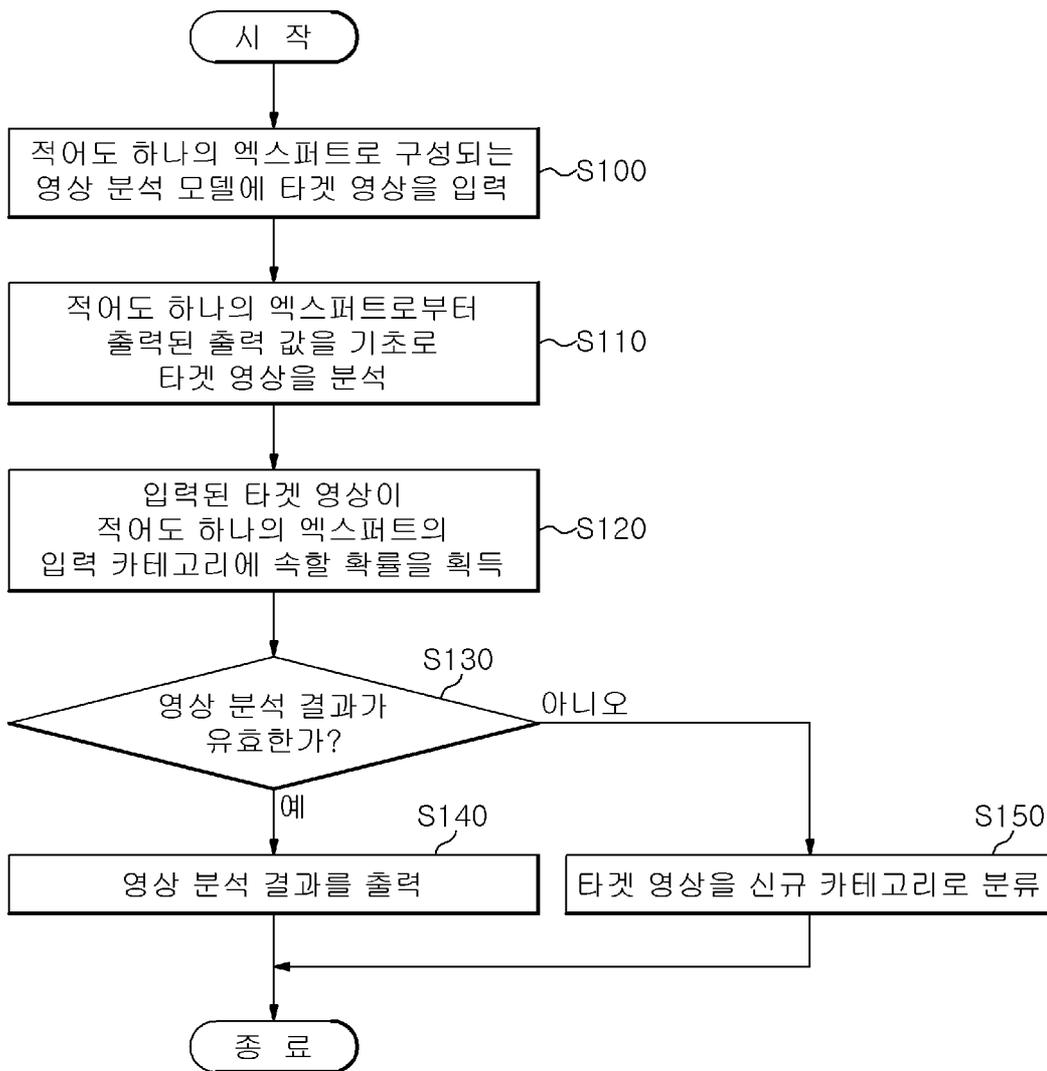
[도5]



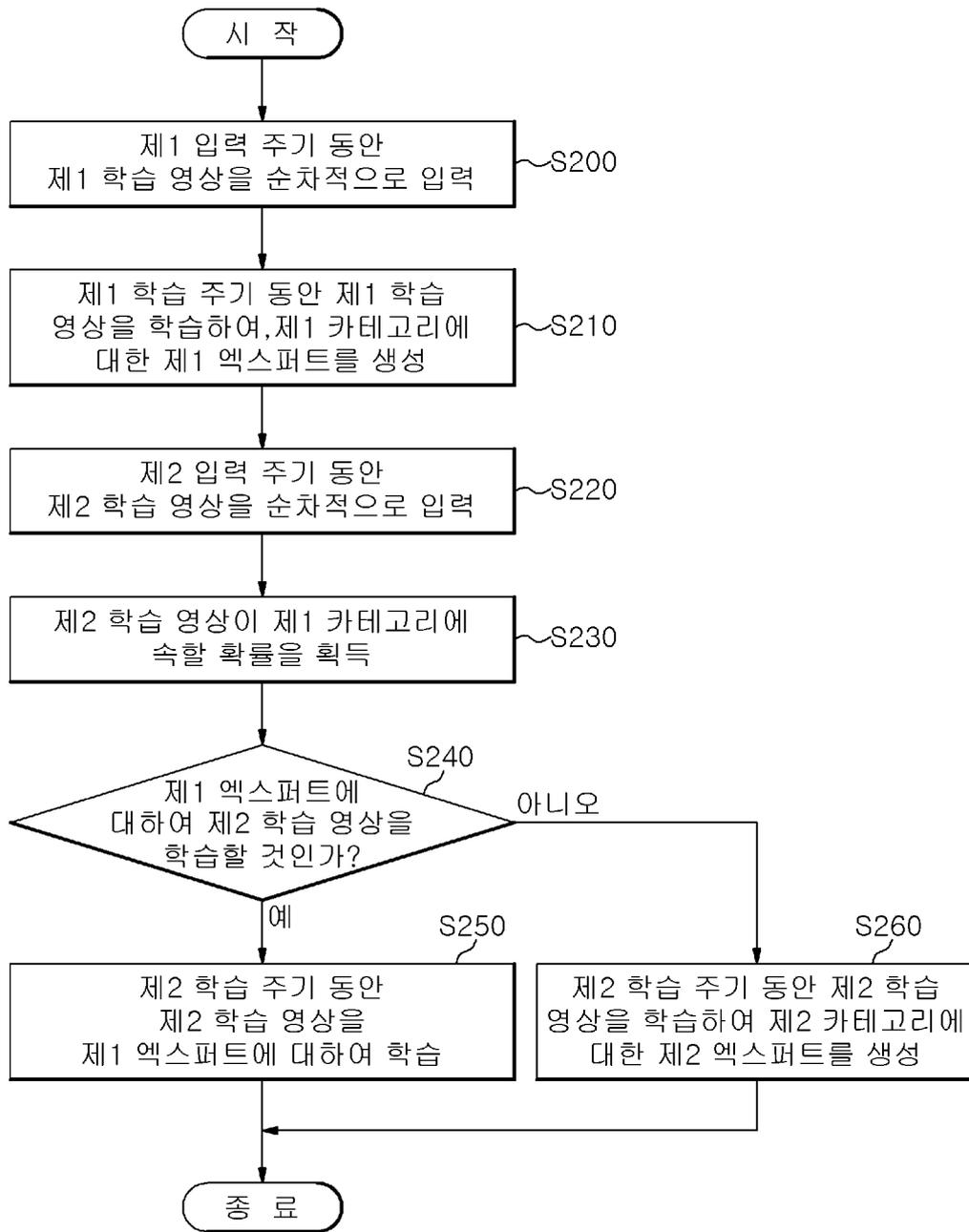
[도6]



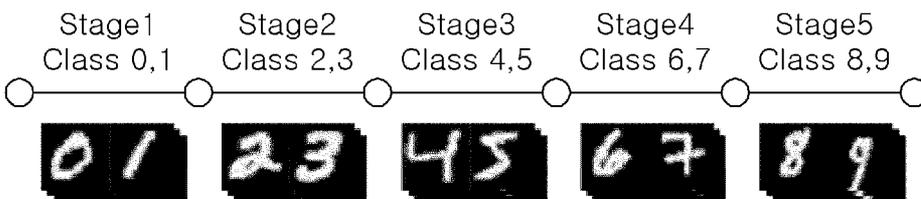
[도7]



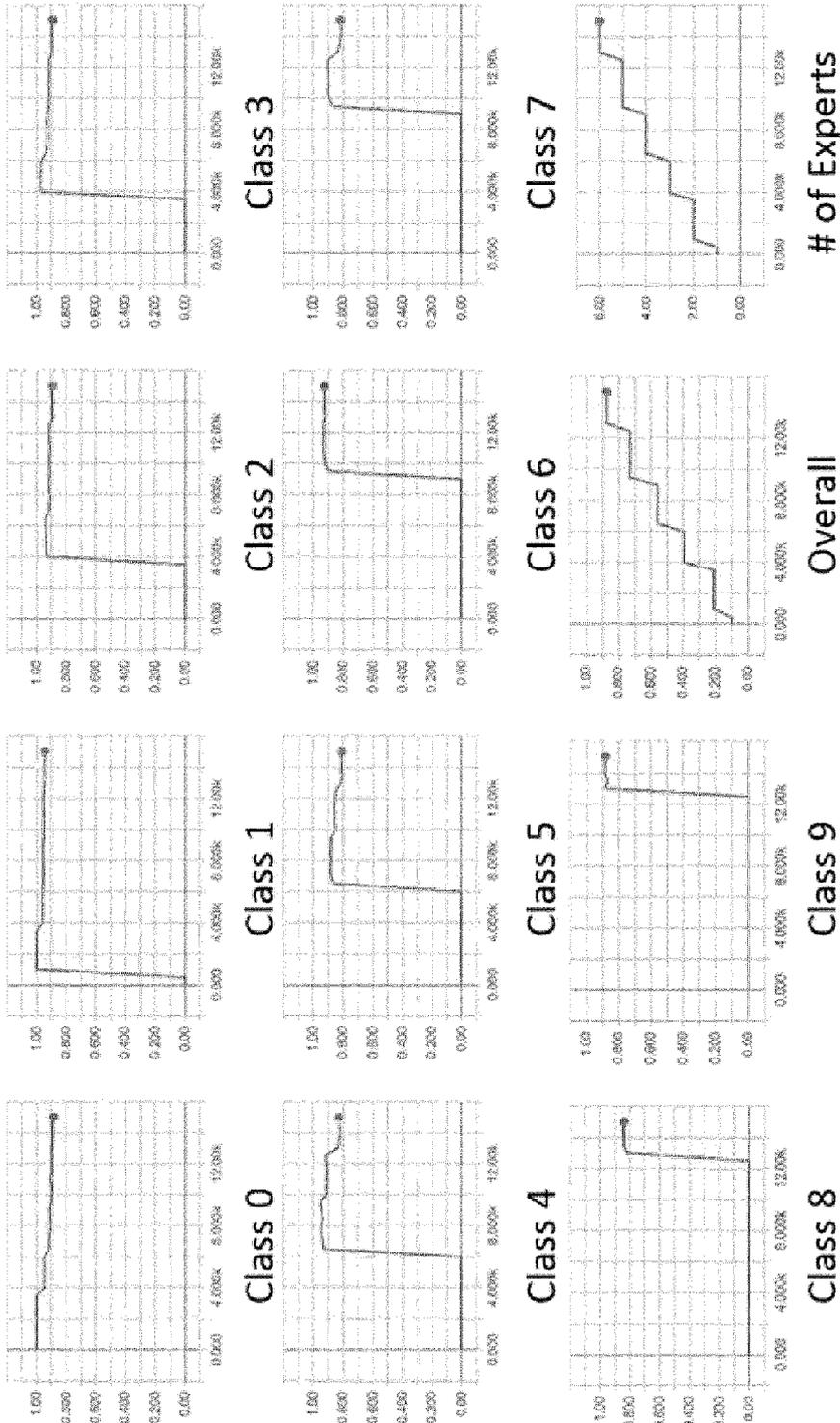
[도8]



[도9]



[도 10]



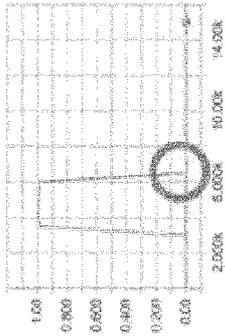
of Experts

Overall

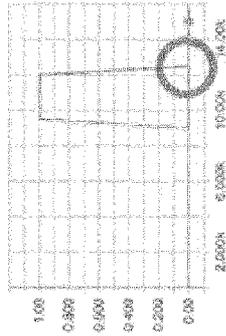
Class 9

Class 8

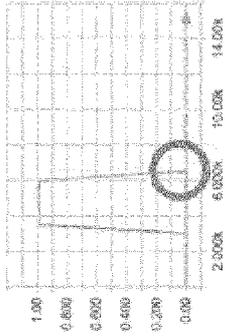
[도 11]



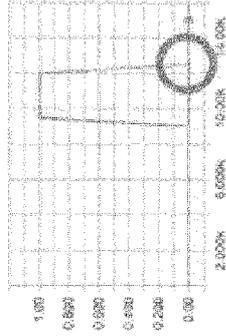
Class 3



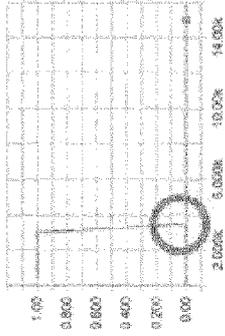
Class 7



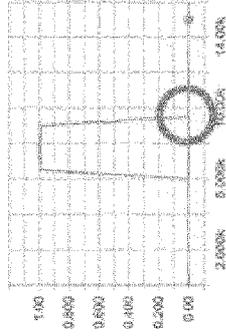
Class 2



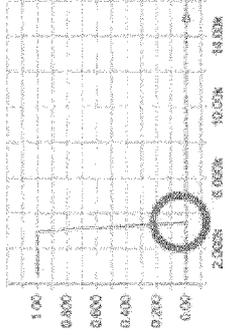
Class 6



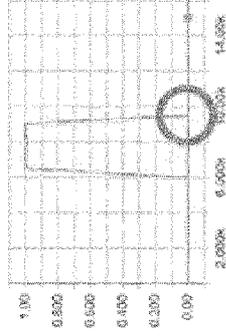
Class 1



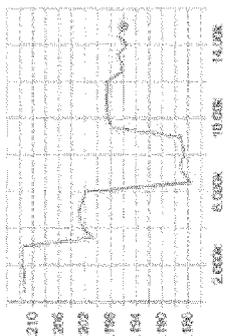
Class 5



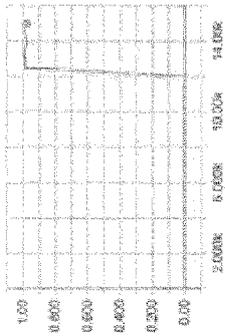
Class 0



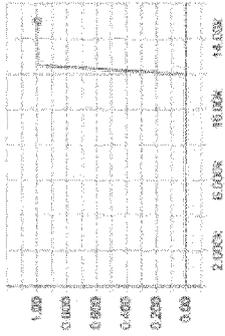
Class 4



Overall



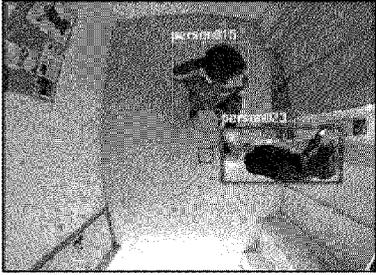
Class 9



Class 8

Catastrophic forgetting

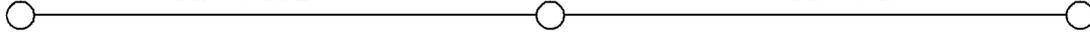
[도 12]



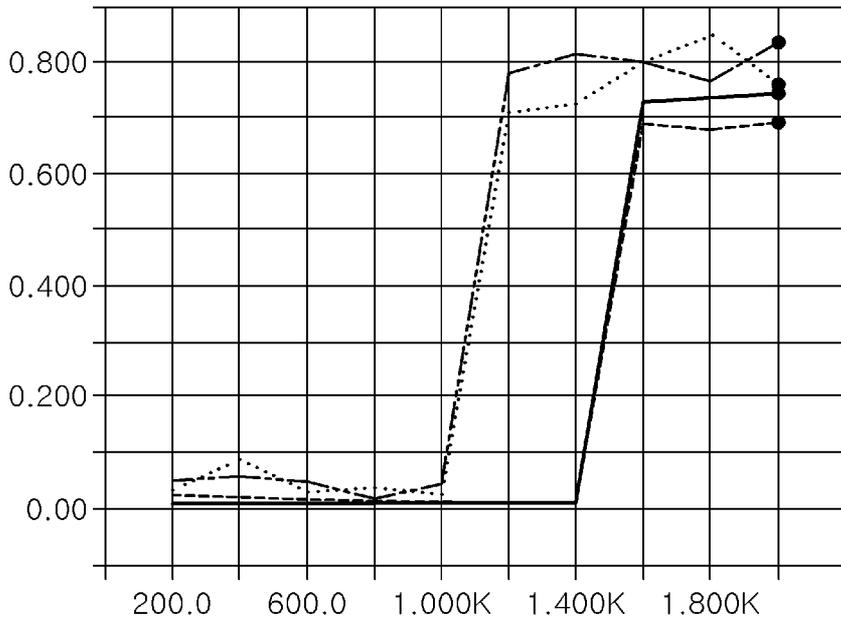
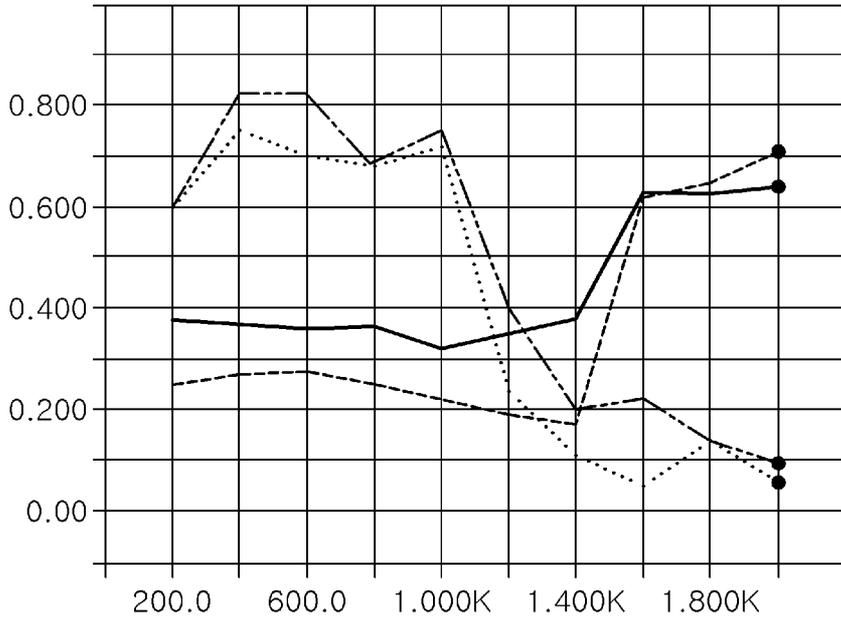
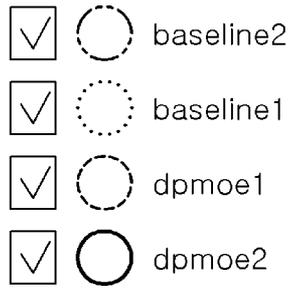
Stage1
camera02



Stage2
camera17



[도13]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2020/000291

A. CLASSIFICATION OF SUBJECT MATTER

G06T 7/00(2006.01)i, G06T 1/60(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 7/00; G06K 9/00; G06K 9/46; G06K 9/62; H04N 5/225; G06T 1/60

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: image, learn, model, cluster, category

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2009-0131626 A (SAMSUNG ELECTRONICS CO., LTD.) 29 December 2009 See paragraphs [0034], [0047]-[0048]; claims 1, 6, 22; and figures 2-3b.	1-12
Y	KR 10-2015-0108673 A (SAMSUNG ELECTRONICS CO., LTD.) 30 September 2015 See paragraphs [0065]-[0068]; claim 1; and figure 5.	1-12
Y	US 2011-0255759 A1 (NISHIMURA, Hirokazu et al.) 20 October 2011 See paragraphs [0063], [0091]-[0092], [0102]; and figures 1, 3, 11.	3-7,9-10
A	KR 10-2015-0107579 A (OMRON CORPORATION) 23 September 2015 See paragraphs [0073]-[0083]; claims 1-8; and figures 1-6.	1-12
A	US 2010-0158332 A1 (RICO, Dan et al.) 24 June 2010 See paragraphs [0037]-[0048]; claims 1-10; and figures 1-3, 5, 12-13.	1-12



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

"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

17 APRIL 2020 (17.04.2020)

Date of mailing of the international search report

17 APRIL 2020 (17.04.2020)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex Daejeon Building 4, 189, Cheongsa-ro, Seo-gu,
Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2020/000291

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2009-0131626 A	29/12/2009	US 08107726 B2	31/01/2012
		US 2009-0316988 A1	24/12/2009
KR 10-2015-0108673 A	30/09/2015	US 2016-0350610 A1	01/12/2016
		WO 2015-141892 A1	24/09/2015
US 2011-0255759 A1	20/10/2011	CN 101351149 A	21/01/2009
		CN 101351149 B	02/02/2011
		CN 101947113 A	19/01/2011
		CN 101947113 B	07/12/2011
		EP 1969992 A1	17/09/2008
		EP 1969992 A4	27/04/2011
		EP 1969992 B1	02/05/2012
		EP 2412300 A2	01/02/2012
		JP 04472631 B2	02/06/2010
		JP 04520405 B2	04/08/2010
		JP 2007-175432 A	12/07/2007
		JP 2007-175434 A	12/07/2007
		US 08300955 B2	30/10/2012
		US 08363962 B2	29/01/2013
		US 2007-0165932 A1	19/07/2007
		US 2011-0255757 A1	20/10/2011
		US 2011-0255758 A1	20/10/2011
		WO 2007-077672 A1	12/07/2007
		WO 2007-077672 A8	11/10/2007
		KR 10-2015-0107579 A	23/09/2015
CN 104915972 B	13/02/2018		
EP 2919162 A1	16/09/2015		
JP 2015-176169 A	05/10/2015		
JP 6330385 B2	30/05/2018		
KR 10-1640998 B1	29/07/2016		
US 09600746 B2	21/03/2017		
US 2015-0262039 A1	17/09/2015		
US 2010-0158332 A1	24/06/2010	CA 2783867 A1	01/07/2010
		CN 102438529 A	02/05/2012
		CN 102438529 B	22/10/2014
		CN 103824290 A	28/05/2014
		CN 103854028 A	11/06/2014
		EP 2378978 A1	26/10/2011
		EP 2378978 A4	22/05/2013
		EP 2378978 B1	12/11/2014
		HK 1199127 A1	19/06/2015
		JP 2012-512672 A	07/06/2012
		US 2013-0343626 A1	26/12/2013
		WO 2010-071999 A1	01/07/2010

A. 발명이 속하는 기술분류(국제특허분류(IPC))
G06T 7/00(2006.01)i, G06T 1/60(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)
G06T 7/00; G06K 9/00; G06K 9/46; G06K 9/62; H04N 5/225; G06T 1/60

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(image), 학습(learn), 모델(model), 클러스터(cluster), 카테고리(category)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-2009-0131626 A (삼성전자주식회사) 2009.12.29 단락 [0034], [0047]-[0048]; 청구항 1, 6, 22; 및 도면 2-3b	1-12
Y	KR 10-2015-0108673 A (삼성전자주식회사) 2015.09.30 단락 [0065]-[0068]; 청구항 1; 및 도면 5	1-12
Y	US 2011-0255759 A1 (HIROKAZU NISHIMURA 등) 2011.10.20 단락 [0063], [0091]-[0092], [0102]; 및 도면 1, 3, 11	3-7,9-10
A	KR 10-2015-0107579 A (오므론 가부시킴가이샤) 2015.09.23 단락 [0073]-[0083]; 청구항 1-8; 및 도면 1-6	1-12
A	US 2010-0158332 A1 (DAN RICO 등) 2010.06.24 단락 [0037]-[0048]; 청구항 1-10; 및 도면 1-3, 5, 12-13	1-12

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 이후 “X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.
 “L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌
 “Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.
 “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 2020년 04월 17일 (17.04.2020)	국제조사보고서 발송일 2020년 04월 17일 (17.04.2020)
--	---

ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578	심사관 김성희 전화번호 +82-42-481-3516
---	------------------------------------



국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2009-0131626 A	2009/12/29	US 08107726 B2 US 2009-0316988 A1	2012/01/31 2009/12/24
KR 10-2015-0108673 A	2015/09/30	US 2016-0350610 A1 WO 2015-141892 A1	2016/12/01 2015/09/24
US 2011-0255759 A1	2011/10/20	CN 101351149 A CN 101351149 B CN 101947113 A CN 101947113 B EP 1969992 A1 EP 1969992 A4 EP 1969992 B1 EP 2412300 A2 JP 04472631 B2 JP 04520405 B2 JP 2007-175432 A JP 2007-175434 A US 08300955 B2 US 08363962 B2 US 2007-0165932 A1 US 2011-0255757 A1 US 2011-0255758 A1 WO 2007-077672 A1 WO 2007-077672 A8	2009/01/21 2011/02/02 2011/01/19 2011/12/07 2008/09/17 2011/04/27 2012/05/02 2012/02/01 2010/06/02 2010/08/04 2007/07/12 2007/07/12 2012/10/30 2013/01/29 2007/07/19 2011/10/20 2011/10/20 2007/07/12 2007/10/11
KR 10-2015-0107579 A	2015/09/23	CN 104915972 A CN 104915972 B EP 2919162 A1 JP 2015-176169 A JP 6330385 B2 KR 10-1640998 B1 US 09600746 B2 US 2015-0262039 A1	2015/09/16 2018/02/13 2015/09/16 2015/10/05 2018/05/30 2016/07/29 2017/03/21 2015/09/17
US 2010-0158332 A1	2010/06/24	CA 2783867 A1 CN 102438529 A CN 102438529 B CN 103824290 A CN 103854028 A EP 2378978 A1 EP 2378978 A4 EP 2378978 B1 HK 1199127 A1 JP 2012-512672 A US 2013-0343626 A1 WO 2010-071999 A1	2010/07/01 2012/05/02 2014/10/22 2014/05/28 2014/06/11 2011/10/26 2013/05/22 2014/11/12 2015/06/19 2012/06/07 2013/12/26 2010/07/01