

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2020년 2월 13일 (13.02.2020)



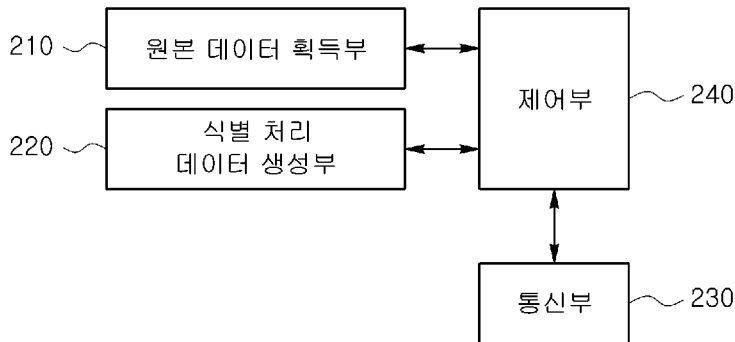
(10) 국제공개번호
WO 2020/032348 A1

- (51) 국제특허분류: *G06F 21/16* (2013.01) *G06N 3/08* (2006.01)
G06F 21/30 (2013.01)
- (21) 국제출원번호: PCT/KR2019/005367
- (22) 국제출원일: 2019년 5월 3일 (03.05.2019)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보: 10-2018-0093747 2018년 8월 10일 (10.08.2018) KR
- (71) 출원인: 주식회사 딥핑소스 (DEEPING SOURCE INC.) [KR/KR]; 06247 서울시 강남구 역삼로 165,2층, Seoul (KR).
- (72) 발명자: 김태훈 (KIM, Tae Hoon); 06275 서울시 강남구 도곡로 306, 104동 1002호, Seoul (KR).
- (74) 대리인: 모아특허법인 (MOA INTELLECTUAL PROPERTY LAW FIRM); 06109 서울시 강남구 봉은사로 213, 10층, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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,

(54) Title: METHOD, SYSTEM, AND NON-TRANSITORY COMPUTER-READABLE RECORDING MEDIUM FOR IDENTIFYING DATA

(54) 발명의 명칭: 데이터를 식별 처리하는 방법, 시스템 및 비일시성의 컴퓨터 판독 가능 기록 매체

200



- 210 ... Original data acquisition unit
- 220 ... Identified data generation unit
- 230 ... Communication unit
- 240 ... Control unit

(57) Abstract: An aspect of the present invention relates to a method for identifying data, comprising the steps of: acquiring original data; and, when synthesized data generated by synthesizing mark data with the original data is input to a learning model, generating, as identified data, synthesized data for which a result is output to be identical or similar to a result that is output as the original data is input to the learning model, wherein the mark data includes data for allowing the synthesized data and the original data to be differently recognized by a person.



WO 2020/032348 A1

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))

(57) 요약서: 본 발명의 일 태양에 따르면, 데이터를 식별(identification) 처리하는 방법으로서, 원본 데이터를 획득하는 단계, 및 상기 원본 데이터에 표지(mark) 데이터가 합성됨에 따라 생성되는 합성 데이터로서, 학습 모델(learning model)에 입력되면, 상기 학습 모델에 상기 원본 데이터가 입력됨에 따라 출력되는 결과와 서로 동일하거나 유사한 결과가 출력되는 합성 데이터를 식별 처리된 데이터로서 생성하는 단계를 포함하고, 상기 표지 데이터는, 상기 합성 데이터와 상기 원본 데이터가 사람에 의해 서로 다르게 인식되도록 하는 데이터를 포함하는 방법이 제공된다.

명세서

발명의 명칭: 데이터를 식별 처리하는 방법, 시스템 및 비밀시성의 컴퓨터 판독 가능 기록 매체

기술분야

- [1] 본 발명은 데이터를 식별 처리하는 방법, 시스템 및 비밀시성의 컴퓨터 판독 가능 기록 매체에 관한 것이다.

배경기술

- [2] 빅데이터(big data)란, 기존의 기업 환경이나 공공 기관에서 사용되는 정형화된 데이터는 물론, 전자 상거래 데이터, 메타 데이터, 웹 로그 데이터, 무선 식별(RFID; Radio Frequency Identification) 데이터, 센서 네트워크 데이터, 소셜 네트워크 데이터, 인터넷 텍스트와 문서에 관한 데이터, 인터넷 검색 인덱싱 데이터 등 기존에 미처 활용되지 못하던 비정형화 또는 반정형화된 데이터를 모두 포함하는 데이터로서, 이와 같은 데이터는 일반적으로 보통의 소프트웨어 툴 및 컴퓨터 시스템으로는 다루기 어려운 수준의 데이터 양을 갖게 된다는 의미에서 빅데이터라 칭하고 있다.
- [3] 근래에 들어, 인공 지능 기술에 대한 관심이 증대되면서, 그 기반이 되는 빅데이터의 거래 또는 공유가 활성화되고 있다.
- [4] 하지만, 데이터의 특성(예를 들어, 데이터의 양이 방대하고 누구로부터 제공받은 데이터인지 특정되기 어려운 특성 등)상 일단 제3자에게 넘어가게 되면, 복제 또는 복사가 용이할 뿐만 아니라 권리가 없는 자에게 무단으로 넘기더라도 이를 적발해내기가 쉬운 일이 아니었다. 이를 보완하기 위해 데이터에 소정의 식별 처리(예를 들어, 워터마크(watermark))를 하는 기술들이 등장하였다.
- [5] 이에 관한 종래 기술의 일 예로서, 한국공개특허공보 제2018-58116호에 개시된 기술을 예로 들 수 있는데, 이에 따르면, 제1 오디오 신호를 MCLT 변환하는 단계, 상기 MCLT 변환된 제1 오디오 신호에 워터마크의 비트열을 삽입하는 단계, 상기 비트열이 삽입된 제1 오디오 신호를 IMDCT 변환하는 단계, 및 상기 IMDCT 변환된 신호를 이웃 프레임 신호와 중첩 가산함으로써 제1 오디오 신호에 워터마크가 삽입된 제2 오디오 신호를 획득하는 단계를 포함하는 오디오 워터마크 삽입 방법이 제공된다.
- [6] 또한, 이에 관한 다른 종래 기술의 일 예로서, 한국등록특허공보 제1837939호에 개시된 기술을 예로 들 수 있는데, 이에 따르면, 소정 색상이 적용되는 문서를 복수 개 구역으로 분할하는 제1 단계, 상기 복수 개 구역으로 분할된 문서에 서로 중첩되지 않도록 워터마크를 선택 삽입하는 제2 단계, 및 상기 소정 색상이 적용된 문서의 색상과 일치하도록 워터마크 삽입 문서를 보정하되, 워터마크의 히든 정도에 따라 색상 보정값을 달리 적용하여 색상을 보정하는 제3 단계로

- 이루어지는 것을 특징으로 하는 문서에 워터마크를 삽입하는 방법이 제공된다.
- [7] 하지만, 위와 같은 종래 기술을 비롯하여 지금까지 소개된 기술에 의하면, 데이터에 식별 처리를 함으로써, 누구의 데이터인지 또는 누구로부터 받은 데이터인지 쉽게 식별될 수 있고, 무단 배포되는 것을 방지할 수 있었으나, 식별 처리 과정에서 원본 데이터의 변형 또는 손상을 야기시키게 되므로 인공지능을 위한 기계 학습, 심층 학습 또는 강화 학습 시 오인식(misrecognized)되거나 사용되기 어려운 문제가 있었다.
- [8] 이에 본 발명자(들)는, 원본 데이터와 표지 데이터를 합성하여 원본 데이터와 상이하게 인식되는 합성 데이터를 생성하되, 학습 모델에 원본 데이터가 입력됨에 따라 출력되는 결과와 해당 학습 모델에 위의 합성 데이터가 입력됨에 따라 출력되는 결과가 서로 동일하거나 유사하게 될 수 있도록 하는 기술을 제안하는 바이다.

발명의 상세한 설명

기술적 과제

- [9] 본 발명은, 전술한 종래 기술의 문제점을 모두 해결하는 것을 그 목적으로 한다.
- [10] 또한, 본 발명은, 식별 처리된 데이터를 통해 데이터의 출처를 쉽게 파악할 수 있도록 하는 것을 또 다른 목적으로 한다.
- [11] 또한, 본 발명은, 식별 처리된 데이터를 통해 기계 학습을 수행하더라도 원본 데이터를 통해 기계 학습을 수행한 결과와 동일하거나 유사한 결과를 얻을 수 있도록 하는 것을 또 다른 목적으로 한다.
- [12] 또한, 본 발명은, 사람에게서는 원본 데이터와 식별 처리된 데이터가 서로 다르게 인식되지만, 컴퓨터에게는 원본 데이터와 식별 처리된 데이터가 서로 동일하거나 유사하게 인식되는 것을 또 다른 목적으로 한다.
- [13] 또한, 본 발명은, 빅데이터 거래 시장에서 데이터 제공자가 안심하고 데이터를 거래 또는 공유할 수 있도록 지원하는 것을 또 다른 목적으로 한다.

과제 해결 수단

- [14] 상기 목적을 달성하기 위한 본 발명의 대표적인 구성은 다음과 같다.
- [15] 본 발명의 일 태양에 따르면, 데이터를 식별(identification) 처리하는 방법으로서, 원본 데이터를 획득하는 단계, 및 상기 원본 데이터에 표지(mark) 데이터가 합성됨에 따라 생성되는 합성 데이터로서, 학습 모델(learning model)에 입력되면, 상기 학습 모델에 상기 원본 데이터가 입력됨에 따라 출력되는 결과와 서로 동일하거나 유사한 결과가 출력되는 합성 데이터를 식별 처리된 데이터로서 생성하는 단계를 포함하고, 상기 표지 데이터는, 상기 합성 데이터와 상기 원본 데이터가 사람에 의해 서로 다르게 인식되도록 하는 데이터를 포함하는 방법이 제공된다.
- [16] 본 발명의 다른 태양에 따르면, 데이터를 식별(identification) 처리하는 방법으로서, 원본 데이터를 획득하는 원본 데이터 획득부, 및 상기 원본

데이터에 표지(mark) 데이터가 합성됨에 따라 생성되는 합성 데이터로서, 학습 모델(learning model)에 입력되면, 상기 학습 모델에 상기 원본 데이터가 입력됨에 따라 출력되는 결과와 서로 동일하거나 유사한 결과가 출력되는 합성 데이터를 식별 처리된 데이터로서 생성하는 단계를 포함하고, 상기 표지 데이터는, 상기 합성 데이터와 상기 원본 데이터가 사람에 의해 서로 다르게 인식되도록 하는 데이터를 포함하는 시스템이 제공된다.

- [17] 이 외에도, 본 발명을 구현하기 위한 다른 방법, 다른 시스템 및 상기 방법을 실행하기 위한 컴퓨터 프로그램을 기록하는 비일시성의 컴퓨터 판독 가능한 기록 매체가 더 제공된다.

발명의 효과

- [18] 본 발명에 의하면, 식별 처리된 데이터를 통해 데이터의 출처를 쉽게 파악할 수 있게 된다.
- [19] 또한, 본 식별 처리된 데이터를 통해 기계 학습을 수행하더라도 원본 데이터를 통해 기계 학습을 수행한 결과와 동일하거나 유사한 결과를 얻을 수 있게 된다.
- [20] 또한, 본 발명에 의하면, 사람에게에는 원본 데이터와 식별 처리된 데이터가 서로 다르게 인식되지만, 컴퓨터에게는 원본 데이터와 식별 처리된 데이터가 서로 동일하거나 유사하게 인식될 수 있게 된다.
- [21] 또한, 본 발명에 의하면, 빅데이터 거래 시장에서 데이터 제공자가 안심하고 데이터를 거래 또는 공유할 수 있도록 지원할 수 있게 된다.

도면의 간단한 설명

- [22] 도 1은 본 발명의 일 실시예에 따라 데이터를 식별(identification) 처리하기 위한 전체 시스템의 구성을 개략적으로 나타내는 도면이다.
- [23] 도 2는 본 발명의 일 실시예에 따른 식별 처리 시스템의 내부 구성을 예시적으로 나타내는 도면이다.
- [24] 도 3 및 도 4는 본 발명의 일 실시예에 따라 데이터가 식별 처리되는 과정의 일부를 예시적으로 나타내는 도면이다.
- [25] 도 5는 본 발명의 일 실시예에 따른 원본 이미지 데이터와 식별 처리된 이미지 데이터를 예시적으로 나타내는 도면이다.
- [26] <부호의 설명>
- [27] 100: 통신망
- [28] 200: 식별 처리 시스템
- [29] 210: 원본 데이터 획득부
- [30] 220: 식별 처리 데이터 생성부
- [31] 230: 통신부
- [32] 240: 제어부
- [33] 300: 디바이스

발명의 실시를 위한 형태

- [34] 후술하는 본 발명에 대한 상세한 설명은, 본 발명이 실시될 수 있는 특정 실시예를 예시로서 도시하는 첨부 도면을 참조한다. 이러한 실시예는 당업자가 본 발명을 실시할 수 있기에 충분하도록 상세히 설명된다. 본 발명의 다양한 실시예는 서로 다르지만 상호 배타적일 필요는 없음이 이해되어야 한다. 예를 들어, 본 명세서에 기재되어 있는 특정 형상, 구조 및 특성은 본 발명의 정신과 범위를 벗어나지 않으면서 일 실시예로부터 다른 실시예로 변경되어 구현될 수 있다. 또한, 각각의 실시예 내의 개별 구성요소의 위치 또는 배치도 본 발명의 정신과 범위를 벗어나지 않으면서 변경될 수 있음이 이해되어야 한다. 따라서, 후술하는 상세한 설명은 한정적인 의미로서 행하여지는 것이 아니며, 본 발명의 범위는 특허청구범위의 청구항들이 청구하는 범위 및 그와 균등한 모든 범위를 포괄하는 것으로 받아들여져야 한다. 도면에서 유사한 참조부호는 여러 측면에 걸쳐서 동일하거나 유사한 구성요소를 나타낸다.
- [35] 이하에서는, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명을 용이하게 실시할 수 있도록 하기 위하여, 본 발명의 여러 바람직한 실시예에 관하여 첨부된 도면을 참조하여 상세히 설명하기로 한다.
- [36] 전체 시스템의 구성
- [37] 도 1은 본 발명의 일 실시예에 따라 데이터를 식별(identification) 처리하기 위한 전체 시스템의 구성을 개략적으로 나타내는 도면이다.
- [38] 도 1에 도시된 바와 같이, 본 발명의 일 실시예에 따른 전체 시스템은 통신망(100), 식별(identification) 처리 시스템(200) 및 디바이스(300)를 포함할 수 있다.
- [39] 먼저, 본 발명의 일 실시예에 따르면, 통신망(100)은 유선 통신이나 무선 통신과 같은 통신 양태를 가리지 않고 구성될 수 있으며, 근거리 통신망(LAN; Local Area Network), 도시권 통신망(MAN; Metropolitan Area Network), 광역 통신망(WAN; Wide Area Network) 등 다양한 통신망으로 구성될 수 있다. 바람직하게는, 본 명세서에서 말하는 통신망(100)은 공지의 인터넷 또는 월드와이드웹(WWW; World Wide Web)일 수 있다. 그러나, 통신망(100)은, 굳이 이에 국한될 필요 없이, 공지의 유무선 데이터 통신망, 공지의 전화망 또는 공지의 유무선 텔레비전 통신망을 그 적어도 일부에 있어서 포함할 수도 있다.
- [40] 예를 들면, 통신망(100)은 무선 데이터 통신망으로서, 와이파이(WiFi) 통신, 와이파이 다이렉트(WiFi-Direct) 통신, 롱텀 에볼루션(LTE, Long Term Evolution) 통신, 블루투스 통신(더 구체적으로는, 저전력 블루투스(BLE; Bluetooth Low Energy)), 적외선 통신, 초음파 통신 등과 같은 종래의 통신 방법을 적어도 그 일부분에 있어서 구현하는 것일 수 있다.
- [41] 다음으로, 본 발명의 일 실시예에 따른 식별 처리 시스템(200)은 통신망(100)을 통하여 후술할 디바이스(300)와 통신을 수행할 수 있고, 원본 데이터를 획득하고, 그 원본 데이터에 표지(mark) 데이터가 합성됨에 따라 생성되는 합성 데이터로서, 학습 모델(learning model)에 입력되면, 그 학습 모델에 위의 원본

데이터가 입력됨에 따라 출력되는 결과와 서로 동일하거나 유사한 결과가 출력되는 합성 데이터를 식별 처리된 데이터로서 생성하는 기능을 수행할 수 있다. 예를 들어, 본 발명의 일 실시예에 따른 식별 처리 시스템(200)은 서버 시스템일 수 있다.

- [42] 또한, 본 발명의 일 실시예에 따른 원본 데이터는 텍스트, 이미지, 동영상, 소리 등 다양한 형태의 데이터를 포함할 수 있다. 또한, 본 발명의 일 실시예에 따른 원본 데이터는 기계 학습, 심층 학습, 강화 학습 등의 다양한 학습 모델을 통해 학습시키는데 사용되는 데이터를 포함할 수 있다.
- [43] 또한, 본 발명의 일 실시예에 따른 표지 데이터는, 합성 데이터와 원본 데이터가 사람에게 의해 서로 다르게 인식되도록 하는 데이터를 포함할 수 있다.
- [44] 구체적으로, 본 발명의 일 실시예에 따른 표지 데이터는 사람의 시각, 청각 및 촉각 중 적어도 하나의 감각에 의해 서로 다르게 인식되도록 하는 데이터를 포함할 수 있다.
- [45] 예를 들어, 본 발명의 일 실시예에 따르면, 이러한 표지 데이터는, 은닉되지 않고, 사람에게 의해 시각, 청각 및 촉각 중 적어도 하나의 감각에 의해 인식될 수 있는 워터마크(watermark) 또는 포렌식마크(forensicmark) 동일 수 있다.
- [46] 한편, 본 발명의 일 실시예에 따른 표지 데이터에는, 원본 데이터와 연관되는 권리자에 관한 정보가 포함될 수 있다.
- [47] 예를 들어, 본 발명의 일 실시예에 따르면, 위의 권리자에 관한 정보에는, 원본 데이터의 제작자, 배포자 또는 수령자에 관한 정보나 제작 일시, 배포 일시 또는 수령 일시에 관한 정보, 원본 데이터에 부여된 권리에 관한 정보 등이 포함될 수 있다.
- [48] 한편, 본 발명의 일 실시예에 따르면, 위의 학습 모델에는 기계 학습(machine learning) 모델, 강화 학습(reinforcement learning) 모델, 딥 러닝(deep learning) 모델, 신경망 학습(neural network learning) 모델 등이 포함될 수 있다.
- [49] 식별 처리 시스템(200)의 기능에 관하여는 아래에서 더 자세하게 알아보기로 한다. 한편, 식별 처리 시스템(200)에 관하여 위와 같이 설명되었으나, 이러한 설명은 예시적인 것이고, 식별 처리 시스템(200)에 요구되는 기능이나 구성요소의 적어도 일부가 필요에 따라 후술할 디바이스(300) 내에서 실현(예를 들어, 식별 처리 시스템(200)에 요구되는 기능이나 구성요소의 적어도 일부가 IC 칩(구체적으로는, 디바이스(300)에 포함된 IC 칩)에 의해 구현될 수 있다.)되거나 외부 시스템(미도시됨) 내에 포함될 수도 있음은 당업자에게 자명하다.
- [50] 다음으로, 본 발명의 일 실시예에 따른 디바이스(300)는 통신망(100)을 통해 식별 처리 시스템(200)에 접속한 후 통신할 수 있도록 하는 기능을 포함하는 디지털 기기로서, 스마트폰, 태블릿 PC 등과 같이 메모리 수단을 구비하고 마이크로 프로세서를 탑재하여 연산 능력을 갖춘 디지털 기기라면 얼마든지 본 발명에 따른 디바이스(300)로서 채택될 수 있다.
- [51] 한편, 본 발명의 일 실시예에 따르면, 디바이스(300)에는 본 발명에 따른

데이터의 식별 처리를 지원하기 위한 애플리케이션이 포함되어 있을 수 있다. 이와 같은 애플리케이션은 식별 처리 시스템(200) 또는 외부의 애플리케이션 배포 서버(미도시됨)로부터 다운로드된 것일 수 있다.

[52] 식별 처리 시스템의 구성

[53] 이하에서는, 본 발명의 구현을 위하여 중요한 기능을 수행하는 식별 처리 시스템(200)의 내부 구성 및 각 구성요소의 기능에 대하여 살펴보기로 한다.

[54] 도 2는 본 발명의 일 실시예에 따른 식별 처리 시스템(200)의 내부 구성을 예시적으로 나타내는 도면이다.

[55] 도 2를 참조하면, 본 발명의 일 실시예에 따른 식별 처리 시스템(200)은 원본 데이터 획득부(210), 식별 처리 데이터 생성부(220), 통신부(230) 및 제어부(240)를 포함할 수 있다. 본 발명의 일 실시예에 따르면, 원본 데이터 획득부(210), 식별 처리 데이터 생성부(220), 통신부(230) 및 제어부(240)는 그 중 적어도 일부가 외부 시스템(미도시됨)과 통신하는 프로그램 모듈들일 수 있다. 이러한 프로그램 모듈들은 운영 시스템, 응용 프로그램 모듈 및 기타 프로그램 모듈의 형태로 식별 처리 시스템(200)에 포함될 수 있으며, 물리적으로는 여러 가지 공지의 기억 장치 상에 저장될 수 있다. 또한, 이러한 프로그램 모듈들은 식별 처리 시스템(200)과 통신 가능한 원격 기억 장치에 저장될 수도 있다. 한편, 이러한 프로그램 모듈들은 본 발명에 따라 후술할 특정 업무를 수행하거나 특정 추상 데이터 유형을 실행하는 루틴, 서브루틴, 프로그램, 오브젝트, 컴포넌트, 데이터 구조 등을 포괄하지만, 이에 제한되지는 않는다.

[56] 먼저, 본 발명의 일 실시예에 따른 원본 데이터 획득부(210)는 원본 데이터를 획득하는 기능을 수행할 수 있다.

[57] 다음으로, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 원본 데이터에 표지(mark) 데이터가 합성됨에 따라 생성되는 합성 데이터로서, 학습 모델(learning model)에 입력되면, 그 학습 모델에 위의 원본 데이터가 입력됨에 따라 출력되는 결과와 서로 동일하거나 유사한 결과가 출력되는 합성 데이터를 식별 처리된 데이터로서 생성할 수 있다. 본 발명의 일 실시예에 따르면, 위의 결과는 위의 데이터(즉, 원본 데이터 또는 합성 데이터) 내 소정 특징과 연관된 확률(probability), 벡터(vector), 행렬(matrix) 및 좌표(coordinate) 중 적어도 하나에 관한 값을 포함할 수 있다. 예를 들어, 본 발명의 일 실시예에 따르면, 원본 데이터가 얼굴 이미지 데이터인 경우에, 위의 결과는, 얼굴 특징점(landmark point, 예를 들어, 눈의 양 끝점)에 관한 좌표 또는 얼굴 특징점에 관한 확률 분포(probability distribution)값 또는 로짓(logits)값일 수 있다.

[58] 구체적으로, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 원본 데이터에 표지 데이터를 합성하여 후보 합성 데이터를 생성할 수 있다. 또한, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 적어도 하나의 학습 모델에서 위의 후보 합성 데이터가 입력됨에 따라 출력되는 결과 및 위의 원본 데이터가 입력됨에 따라 출력되는 결과 사이의 차이를 산출하고, 그 산출되는

차이에 기초하여 위의 후보 합성 데이터를 보정할 수 있다.

- [59] 예를 들어, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 원본 데이터 x 에 표지 데이터 w 를 합성하여 후보 합성 데이터 x' 을 생성할 수 있다. 이 경우에, 본 발명의 일 실시예에 따르면, 위의 원본 데이터 x 에 위의 표지 데이터 w 를 합성하여 위의 후보 합성 데이터 x' 를 합성하는 목적 함수를 $W(x, w)$ 로 가정해 볼 수 있다. 그 다음에, 본 발명의 일 실시예에 따른 제1 학습 모델 F_1 에서 위의 후보 합성 데이터 x' 이 입력됨에 따라 출력되는 결과 및 위의 원본 데이터 x 가 입력됨에 따라 출력되는 결과 사이의 차이의 노름(norm)(즉, " $\|F_1(x) - F_1(x')\|$ ")을 에러값(err)으로서 산출할 수 있다.
- [60] 그 다음에, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 위의 에러값에 기초하여 위의 제1 학습 모델 F_1 에 원본 데이터가 입력됨에 따라 출력되는 결과와 위의 제1 학습 모델 F_1 에 후보 합성 데이터 x' 이 입력됨에 따라 출력되는 결과 사이의 차이를 감소(또는 최소화)시키도록 위의 목적 함수 $W(x, w)$ 를 보정할 수 있다. 즉, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 위의 에러값이 소정 수준 이하로 작아지도록 위의 목적 함수 $W(x, w)$ 를 보정할 수 있고, 이를 통해, 위의 제1 학습 모델 F_1 에 원본 데이터가 입력됨에 따라 출력되는 결과와 위의 제1 학습 모델 F_1 에 후보 합성 데이터 x' 이 입력됨에 따라 출력되는 결과 사이의 차이가 작아지도록(또는 최소화되도록) 할 수 있다.
- [61] 그 다음에, 위의 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 위의 보정 이후에 제2 내지 제N 학습 모델(즉 F_2 내지 F_N)에 대하여도 위와 동일한 과정을 반복하여 수행할 수 있고, 위의 과정을 반복하여 N회 보정된 목적 함수 $W(x, w)$ 에 원본 데이터 x 가 입력됨에 따라 출력되는 결과를 식별 처리된 데이터로서 생성할 수 있다. 한편, 본 발명의 일 실시예에 따르면, 위의 N은 제N 학습 모델 F_N 에 원본 데이터 x 가 입력됨에 따라 출력되는 결과와 N-1회 보정된 후보 합성 데이터 x' 이 입력됨에 따라 출력되는 결과 사이의 차이가 소정 수준 이하가 되는 값으로 결정될 수 있다. 한편, 본 발명의 일 실시예에 따르면, 위의 에러값(err)을 산출함에 있어서, 위의 제N 학습 모델 F_N 에서의 에러값은, 제1 내지 제N 학습 모델 각각에서의 노름의 합에 기초하여 산출될 수도 있다. 예를 들어, 본 발명의 일 실시예에 따르면, 위의 제N 학습 모델 F_N 에서의 에러값은 $\sum_N \|F_N(x) - F_N(x')\|$ 에 기초하여 산출될 수 있다.
- [62] 그 다음에, 본 발명의 일 실시예에 따르면, 식별 처리 시스템(200)은 위의 N회(또는 N-1회) 보정된 목적 함수 $W(x, w)$ 에 원본 데이터 x 가 입력됨에 따라 출력되는 결과를 식별 처리된 이미지 데이터로서 생성할 수 있다.
- [63] 한편, 앞서 살펴본 예에서는 결과 사이의 차이에 기초하여 에러값을 구하는 과정에서 노름(구체적으로는 L1 노름(norm))을 이용하였으나, 이에 한정되지 않고, 코사인 유사도(cosine similarity), L2 노름(norm) 등 거리 차이를 산출하는 다른 방식 또한 본 발명의 목적을 달성할 수 있는 범위 내에서 얼마든지 사용될

수 있다.

- [64] 한편, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 위의 학습 모델에 원본 데이터가 입력됨에 따라 출력되는 결과와 위의 학습 모델에 식별 처리된 데이터가 입력됨에 따라 출력되는 결과 사이의 차이를 감소(또는 최소화)시키고, 원본 데이터와 식별 처리된 데이터 사이의 차이를 증가(또는 최대화)시키도록 하여 위의 식별 처리된 데이터를 생성할 수 있다.
- [65] 예를 들어, 도 3을 참조하면, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 원본 데이터 x 에 표지 데이터 w 를 합성하여 후보 합성 데이터 x' 을 생성할 수 있다. 이 경우에, 본 발명의 일 실시예에 따르면, 위의 원본 데이터 x 에 위의 표지 데이터 w 를 합성하여 위의 후보 합성 데이터 x' 를 합성하는 목적 함수를 $W(x, w)$ 로 가정해 볼 수 있다(310)(예를 들어, 최초 목적 함수는 기설정될 수 있다.).
- [66] 그 다음에, 본 발명의 일 실시예에 따른 제1 학습 모델 F_1 에서 위의 후보 합성 데이터 x' 이 입력됨에 따라 출력되는 결과 및 위의 원본 데이터 x 가 입력됨에 따라 출력되는 결과 사이의 차이에 기초하여 에러값(err)을 산출할 수 있다. 예를 들어, 본 발명의 일 실시예에 따르면 이러한 에러값(err)은 $\sum_N \omega_N \text{loss function}_i(F_N(x), F_N(x'))$ (이 경우, 제1 학습 모델이므로 N 은 1일 수 있다.)(320)에 기초하여 산출될 수 있다. 보다 상세하게는, 본 발명의 일 실시예에 따른 에러값(err)은 $\sum_N \omega_N \|F_N(x) - F_N(x')\|$ (이 경우, 제1 학습 모델이므로 N 은 1일 수 있고, 가중치 ω_1 내지 ω_N 은 기설정될 수 있다. 즉 $\sum_i \omega_i \|F_i(x) - F_i(x')\|$)에 기초하여 산출될 수 있다.
- [67] 그 다음에, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 원본 데이터 x 와 후보 합성 데이터 x' 사이의 차이가 커지도록 하기 위하여 차이 극대화 모델 D (예를 들어, $D(x, x')$ 또는 $D(x, W(x, w))$)에 기초하여 위의 에러값(err)을 보정할 수 있다(330, 340). 예를 들어, 본 발명의 일 실시예에 따른 차이 극대화 모델 D 는 적대적 트레이닝(adversarial training)을 이용한 학습(예를 들어, 실제(real)/가짜(fake) 학습)에 기초하여 특정되는 모델일 수 있고, 그 모델 D 로부터 원본 데이터 x 및 후보 합성 데이터 x' 사이의 차이를 소정 수준 이상으로 증가(또는 최대화)시키기 위한 출력값이 산출될 수 있으며, 식별 처리 데이터 생성부(220)가 그 출력값을 위의 에러값에 반영시킴으로써, 위의 에러값을 보정할 수 있다. 즉, 위의 에러값에 위의 출력값을 반영시킴으로써, 원본 데이터 x 와 후보 합성 데이터 x' 사이의 차이를 증가시키고, 이를 통해, 원본 데이터 x 와 후보 합성 데이터 x' 가 사람의 시각, 청각 및 촉각 중 적어도 하나의 감각에 의해 다르게 인식되도록 할 수 있다.
- [68] 또한, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 원본 데이터 x 와 후보 합성 데이터 x' 사이의 차이가 커지도록 하기 위하여 유사도 모델(similarity model) S 를 사용할 수도 있다. 예를 들어, 본 발명의 일 실시예에 따른 유사도 모델 S 는 $\max(0, \text{Margin} - \|x - x'\|)$ 또는 $\exp(-\|x - x'\|/\delta)$ 에 기초하여 특정될 수 있고, 그 모델 S 로부터 원본 데이터 x 및 후보 합성 데이터 x' 사이의

차이를 소정 수준 이상으로 증가(또는 최대화)시키기 위한 출력값이 산출될 수 있으며 식별 처리 데이터 생성부(220)가 그 출력값을 위의 에러값에 반영시킴으로써, 위의 에러값을 보정할 수 있다(350). 즉, 위의 에러값에 위의 출력값을 반영시킴으로써, 원본 데이터 x 와 후보 합성 데이터 x' 사이의 차이를 증가시키고, 이를 통해, 원본 데이터 x 와 후보 합성 데이터 x' 가 사람의 시각, 청각 및 촉각 중 적어도 하나의 감각에 의해 다르게 인식될 수 있게 된다.

- [69] 그 다음에, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 위의 보정된 에러값에 기초하여 위의 제1 학습 모델 F_1 에 원본 데이터가 입력됨에 따라 출력되는 결과와 위의 제1 학습 모델 F_1 에 후보 합성 데이터 x' 가 입력됨에 따라 출력되는 결과 사이의 차이를 감소(또는 최소화)시키도록 위의 목적 함수 $W(x, w)$ 를 보정할 수 있다(360). 예를 들어, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 위의 보정된 에러값이 소정 수준 이하로 작아지도록 목적 함수 $W(x, w)$ 를 보정할 수 있고, 이를 통해, 위의 제1 학습 모델 F_1 에 원본 데이터 x 가 입력됨에 따라 출력되는 결과와 위의 제1 학습 모델 F_1 에 후보 합성 데이터 x' 가 입력됨에 따라 출력되는 결과 사이의 차이가 작아지도록(또는 최소화되도록) 할 수 있다.
- [70] 그 다음에, 위의 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 위의 보정 이후에 제2 내지 제N 학습 모델(즉 F_2 내지 F_N)에 대하여도 위와 동일한 과정을 반복하여 수행할 수 있고, 위의 과정을 반복하여 N회 보정된 목적 함수 $W(x, w)$ 에 원본 데이터 x 가 입력됨에 따라 출력되는 결과를 식별 처리된 데이터로서 생성할 수 있다(370).
- [71] 한편, 본 발명의 일 실시예에 따르면, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 위의 합성 데이터(또는, 식별 처리된 데이터)로부터 표지 데이터를 분리하여 원본 데이터가 추출되는 것을 방지하기 위하여 원본 데이터와 합성되기 전에 표지 데이터 상에 소정의 변형(perturbation)을 가할 수 있다.
- [72] 구체적으로, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 원본 데이터의 유형(예를 들어, 이미지, 소리 등)이나 특성(예를 들어, 픽셀, 주파수 등)을 참조하여 위의 표지 데이터 상에 가할 변형의 종류 또는 그 변형의 정도를 결정할 수 있다.
- [73] 예를 들어, 도 4를 참조하면, 본 발명의 일 실시예에 따른 원본 데이터가 이미지인 경우에, 식별 처리 데이터 생성부(220)는 표지 데이터(410) 상에 변형을 가할 픽셀 기반의 블렌딩 팩터(blending factor) 및 부분 픽셀 레벨(sub-pixel-level)의 공간 왜곡(spatial distortion) 정도를 결정하고, 그 결정된 정도에 따라 표지 데이터(410)를 변형함으로써 위의 합성에 사용될 표지 데이터(420)를 생성할 수 있다.
- [74] 다른 예를 들어, 본 발명의 일 실시예에 따른 식별 처리 데이터 생성부(220)는 원본 데이터로부터 추출되는 소정의 암호화 해시값(cryptographic hash value,

- 예를 들어, SHA-256)에 기초하여 위의 표지 데이터 상에 가할 임의 변형 정도(random perturbation)를 결정할 수 있고, 그 결정된 정도에 따라 위의 표지 데이터(410)를 변형함으로써 합성에 사용될 표지 데이터(420)를 생성할 수 있다.
- [75] 본 발명의 일 실시예에 따르면, 원본 데이터와 합성되는 표지 데이터 상에 위와 같은 변형(perturbation)을 가함으로써, 원본 데이터에 표지 데이터를 합성하여 생성되는 과정이 비가역적일 수 있게 된다.
- [76] 다음으로, 본 발명의 일 실시예에 따르면, 통신부(230)는 원본 데이터 획득부(210) 및 식별 처리 데이터 생성부(220)로부터의/로의 데이터 송수신이 가능하도록 하는 기능을 수행할 수 있다.
- [77] 마지막으로, 본 발명의 일 실시예에 따르면, 제어부(240)는 원본 데이터 획득부(210), 식별 처리 데이터 생성부(220) 및 통신부(230) 간의 데이터의 흐름을 제어하는 기능을 수행할 수 있다. 즉, 본 발명에 따른 제어부(240)는 식별 처리 시스템(200)의 외부로부터의/로의 데이터 흐름 또는 식별 처리 시스템(200)의 각 구성요소 간의 데이터 흐름을 제어함으로써, 원본 데이터 획득부(210), 식별 처리 데이터 생성부(220) 및 통신부(230)에서 각각 고유 기능을 수행하도록 제어할 수 있다.
- [78] 도 5는 본 발명의 일 실시예에 따른 원본 이미지 데이터와 식별 처리된 이미지 데이터를 예시적으로 나타내는 도면이다.
- [79] 도 5를 참조하면, 먼저, 본 발명의 일 실시예에 따른 식별 처리 시스템(200)은 원본 이미지 데이터 x (510)를 획득할 수 있다.
- [80] 그 다음에, 본 발명의 일 실시예에 따른 식별 처리 시스템(200)은 원본 이미지 데이터 x 의 유형이나 특성을 참조하여 표지 이미지 데이터 w 상에 변형을 가할 픽셀 기반의 블렌딩 팩터(blending factor) 및 부분 픽셀 레벨(sub-pixel-level)의 공간 왜곡(spatial distortion) 정도를 결정하고, 그 결정된 정도에 따라 표지 이미지 데이터 w 를 변형함으로써 합성에 사용될 표지 이미지 데이터 w' 을 생성할 수 있다.
- [81] 그 다음에, 본 발명의 일 실시예에 따른 식별 처리 시스템(200)은 원본 이미지 데이터 x 에 표지 이미지 데이터 w' 를 합성하여 후보 합성 이미지 데이터 x' 을 생성할 수 있다. 이 경우에, 본 발명의 일 실시예에 따르면, 위의 원본 이미지 데이터 x 에 위의 표지 이미지 데이터 w' 을 합성하여 위의 후보 합성 이미지 데이터 x' 를 합성하는 목적 함수를 $W(x, w')$ 로 가정해 볼 수 있다.
- [82] 그 다음에, 본 발명의 일 실시예에 따른 제1 학습 모델 F_1 에서 위의 후보 합성 이미지 데이터 x' 이 입력됨에 따라 출력되는 결과 및 위의 원본 이미지 데이터 x 가 입력됨에 따라 출력되는 결과 사이의 차이에 기초하여 에러값(err)을 산출할 수 있다. 예를 들어, 본 발명의 일 실시예에 따른 에러값(err)은 $\sum w_i \|F_1(x) - F_1(x')\|$ 에 기초하여 산출될 수 있다.
- [83] 그 다음에, 본 발명의 일 실시예에 따른 식별 처리 시스템(200)은 원본 이미지 데이터 x 와 후보 합성 이미지 데이터 x' 사이의 차이가 커지도록 하기 위하여

차이 극대화 모델 D(예를 들어, $D(x, x')$ 또는 $D(x, W(x, w))$)(330, 340)에 기초하여 위의 에러값(err)을 보정할 수 있다.

- [84] 한편, 본 발명의 일 실시예에 따르면, 위의 차이 극대화 모델 D는 적대적 트레이닝(adversarial training)을 이용한 학습(예를 들어, 실제(real)/가짜(fake) 학습)에 기초하여 특정되는 모델일 수 있다. 또한, 차이 극대화 모델 D와 위의 $W(x, w')$ 사이의 관계가 아래 수학적 식 1과 같이 설정될 수 있다.

[85]

[86] [수식1]

$$\min_w \max_D V(W, D) = E_{\mathbf{x} \sim p_{\text{desired}}}[\log D(\mathbf{x})] + E_{\mathbf{x} \sim p_{\mathbf{x}}(\mathbf{x})}[\log(1 - D(W(\mathbf{x}, \mathbf{w})))]$$

[87]

- [88] 이 경우에, 본 발명의 일 실시예에 따르면, $V(W, D)$ 는 목적 함수 $W(\cdot)$ 의 최소 및 차이 극대화 모델 $D(\cdot)$ 의 최대를 만족시키는 값을 찾는 가치 함수(value function)일 수 있고, p_{desired} 및 $p_{\mathbf{x}}$ 는 각각 원하는 분포(desired distribution) 및 데이터 분포(data distribution)를 의미할 수 있다.

- [89] 그 다음에, 본 발명의 일 실시예에 따른 식별 처리 시스템(200)은 위의 보정된 에러값에 기초하여 위의 제1 학습 모델 F_1 에 원본 이미지 데이터가 입력됨에 따라 출력되는 결과와 위의 제1 학습 모델 F_1 에 후보 합성 이미지 데이터 x' 이 입력됨에 따라 출력되는 결과 사이의 차이가 최소화되도록 위의 $W(x, w')$ 함수를 보정할 수 있다.

- [90] 그 다음에, 본 발명의 일 실시예에 따르면, 식별 처리 시스템(200)은 제2 내지 제N 학습 모델(즉 F_2 내지 F_N)에 대하여도 위와 동일한 과정을 반복하여 수행함으로써, $W(x, w')$ 함수를 재차 보정할 수 있다. 한편, 본 발명의 일 실시예에 따르면, 위의 N은 제N 학습 모델 F_N 에 원본 이미지 데이터가 입력됨에 따라 출력되는 결과와 N-1회 보정된 후보 합성 이미지 데이터 x' 이 입력됨에 따라 출력되는 결과 사이의 차이가 소정 수준 이하가 되는 값으로 결정될 수 있다.

- [91] 그 다음에, 본 발명의 일 실시예에 따르면, 식별 처리 시스템(200)은 위의 N회(또는 N-1회) 보정된 목적 함수 $W(x, w')$ 에 원본 이미지 데이터 x 가 입력됨에 따라 출력되는 결과를 식별 처리된 이미지 데이터로서 생성할 수 있다(520).

- [92] 도 5의 (b)는 본 발명의 일 실시예에 따라 원본 이미지 데이터가 식별 처리된 결과를 예시적으로 나타내는 도면이다.

- [93] 도 5의 (a) 및 도 5의 (b)를 참조하면, 본 발명의 일 실시예에 따른 도 5의 (b)의 식별 처리된 이미지 데이터와 도 5의 (a)의 원본 이미지 데이터는 사람에 의해 시각적으로 서로 다르게 인식될 수 있으나, 어느 학습 모델에 입력되더라도 동일하거나 유사한 결과를 산출할 수 있는 이미지 데이터일 수 있게 된다.

- [94] 한편, 본 발명에 따른 식별 처리 시스템(200)에 의해 식별 처리된 이미지 데이터가 기계 학습을 위하여 대량의 이미지 구매를 원하는 구매자에게

제공(또는 판매)될 수 있다.

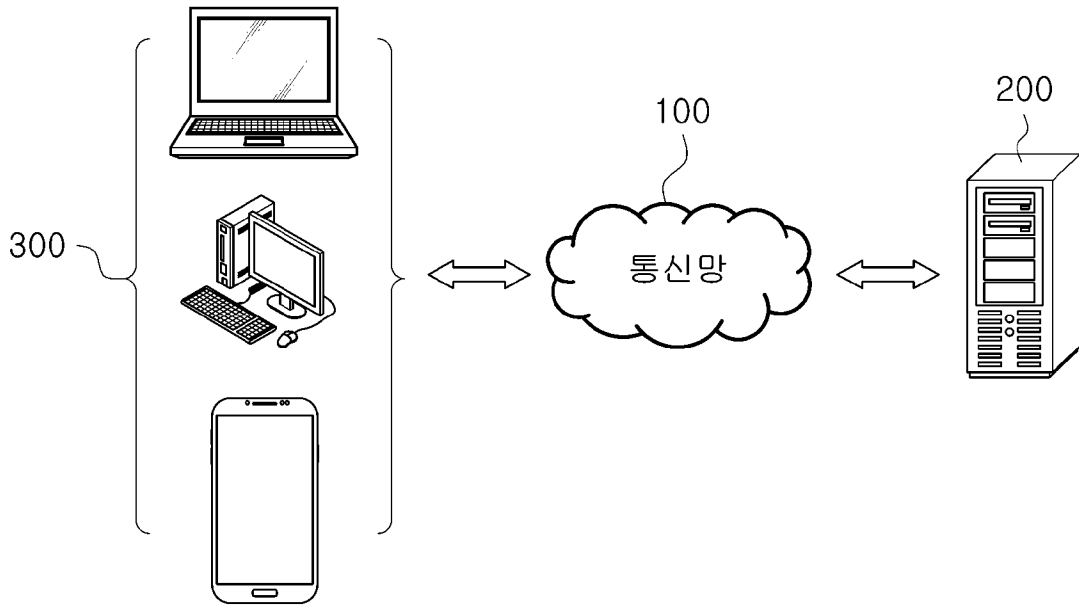
- [95] 이 경우에, 본 발명의 일 실시예에 따라 위의 판매된 이미지들이 식별 처리되어 있으므로, 구매자의 무단 복사 또는 무단 복제를 어렵게 할 뿐만 아니라 제3자에게 무단으로 제공되더라도, 향후에 쉽게 추적이 될 수 있게 된다. 또한, 본 발명의 일 실시예에 따르면, 사람에 의해서는 원본 이미지 데이터와 서로 상이하게 인식될 수 있으나, 기계에 의해서는(구체적으로는, 기계 학습 시) 원본 이미지 데이터와 서로 동일하거나 유사하게 인식될 수 있게 된다.
- [96] 이상 설명된 본 발명에 따른 실시예는 다양한 컴퓨터 구성요소를 통하여 실행될 수 있는 프로그램 명령어의 형태로 구현되어 컴퓨터 판독 가능한 기록 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능한 기록 매체는 프로그램 명령어, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 컴퓨터 판독 가능한 기록 매체에 기록되는 프로그램 명령어는 본 발명을 위하여 특별히 설계되고 구성된 것이거나 컴퓨터 소프트웨어 분야의 당업자에게 공지되어 사용 가능한 것일 수 있다. 컴퓨터 판독 가능한 기록 매체의 예에는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium), 및 ROM, RAM, 플래시 메모리 등과 같은, 프로그램 명령어를 저장하고 실행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령어의 예에는, 컴파일러에 의하여 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용하여 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드도 포함된다. 하드웨어 장치는 본 발명에 따른 처리를 수행하기 위하여 하나 이상의 소프트웨어 모듈로 변경될 수 있으며, 그 역도 마찬가지이다.
- [97] 이상에서 본 발명이 구체적인 구성요소 등과 같은 특정 사항과 한정된 실시예 및 도면에 의하여 설명되었으나, 이는 본 발명의 보다 전반적인 이해를 돕기 위하여 제공된 것일 뿐, 본 발명이 상기 실시예에 한정되는 것은 아니며, 본 발명이 속하는 기술분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정과 변경을 꾀할 수 있다.
- [98] 따라서, 본 발명의 사상은 상기 설명된 실시예에 국한되어 정해져서는 아니 되며, 후술하는 특허청구범위뿐만 아니라 이 특허청구범위와 균등한 또는 이로부터 등가적으로 변경된 모든 범위는 본 발명의 사상의 범주에 속한다고 할 것이다.

청구범위

- [청구항 1] 데이터를 식별(identification) 처리하는 방법으로서,
 원본 데이터를 획득하는 단계, 및
 상기 원본 데이터에 표지(mark) 데이터가 합성됨에 따라 생성되는 합성 데이터로서, 학습 모델(learning model)에 입력되면, 상기 학습 모델에 상기 원본 데이터가 입력됨에 따라 출력되는 결과와 서로 동일하거나 유사한 결과가 출력되는 합성 데이터를 식별 처리된 데이터로서 생성하는 단계를 포함하고,
 상기 표지 데이터는, 상기 합성 데이터와 상기 원본 데이터가 사람에 의해 서로 다르게 인식되도록 하는 데이터를 포함하는 방법.
- [청구항 2] 제1항에 있어서,
 상기 결과는, 상기 원본 데이터 또는 상기 합성 데이터 내 소정 특징과 연관된 확률(probability), 벡터(vector), 행렬(matrix) 및 좌표(coordinate) 중 적어도 하나에 관한 값을 포함하는 방법.
- [청구항 3] 제1항에 있어서,
 상기 표지 데이터에는 상기 원본 데이터와 연관되는 권리자에 관한 정보가 포함되는 방법.
- [청구항 4] 제1항에 있어서,
 상기 생성 단계에서, 상기 학습 모델에 상기 원본 데이터가 입력됨에 따라 출력되는 결과와 상기 학습 모델에 상기 식별 처리된 데이터가 입력됨에 따라 출력되는 결과 사이의 차이를 감소시키고, 상기 원본 데이터와 상기 식별 처리된 데이터 사이의 차이를 증가시켜 상기 식별 처리된 데이터를 생성하는 방법.
- [청구항 5] 제1항에 있어서,
 상기 생성 단계는, 상기 원본 데이터를 상기 표지 데이터와 합성하여 상기 합성 데이터를 생성하는 단계, 및
 적어도 하나의 학습 모델에서 상기 후보 합성 데이터가 입력됨에 따라 출력되는 결과 및 상기 원본 데이터가 입력됨에 따라 출력되는 결과 사이의 차이를 산출하고, 상기 산출되는 차이에 기초하여 상기 후보 합성 데이터를 보정하는 단계를 포함하는 방법.
- [청구항 6] 제1항에 있어서,
 상기 원본 데이터와 상기 표지 데이터 사이의 합성은 비가역적인

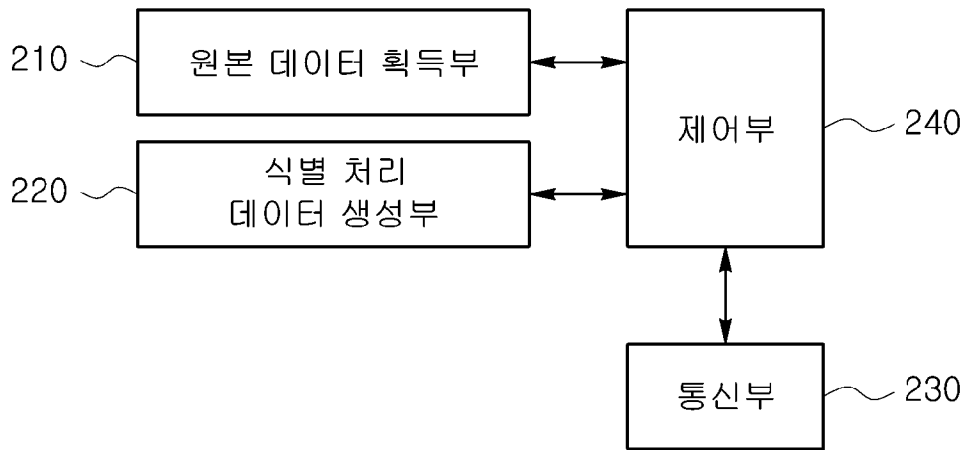
- 방법.
- [청구항 7] 제1항에 있어서,
상기 학습 모델에는, 기계 학습(machine learning) 모델 및 신경망 학습(neural network learning) 모델 중 적어도 하나가 포함되는 방법.
- [청구항 8] 제1항에 있어서,
상기 표지 데이터는 상기 사람의 시각, 청각 및 촉각 중 적어도 하나의 감각에 의해 서로 다르게 인식되도록 하는 데이터인 방법.
- [청구항 9] 제1항에 있어서,
상기 데이터는 이미지에 관한 데이터이고, 상기 표지 데이터는 워터마크(watermark)에 관한 데이터인 방법.
- [청구항 10] 제1항에 따른 방법을 실행하기 위한 컴퓨터 프로그램을 기록하는 비일시성의 컴퓨터 판독 가능 기록 매체.
- [청구항 11] 데이터를 식별(identification) 처리하는 방법으로서,
원본 데이터를 획득하는 원본 데이터 획득부, 및
상기 원본 데이터에 표지(mark) 데이터가 합성됨에 따라 생성되는 합성 데이터로서, 학습 모델(learning model)에 입력되면, 상기 학습 모델에 상기 원본 데이터가 입력됨에 따라 출력되는 결과와 서로 동일하거나 유사한 결과가 출력되는 합성 데이터를 식별 처리된 데이터로서 생성하는 단계를 포함하고,
상기 표지 데이터는, 상기 합성 데이터와 상기 원본 데이터가 사람에 의해 서로 다르게 인식되도록 하는 데이터를 포함하는 시스템.

[도1]

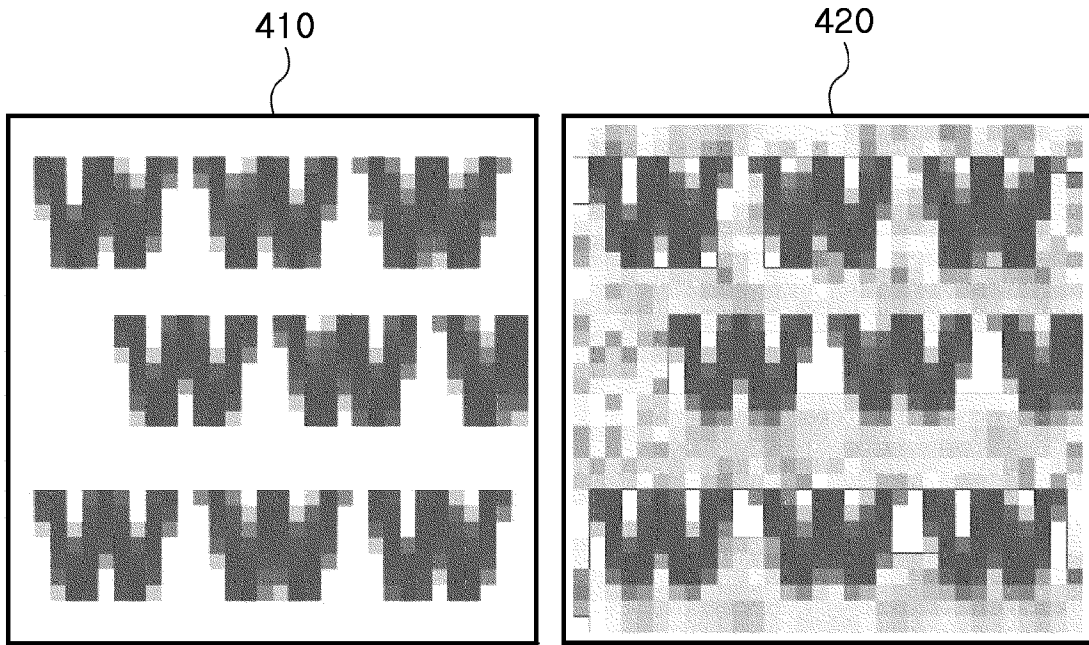


[도2]

200



[도4]



[도5]

(a) Original x (b) Watermarked x'

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2019/005367

A. CLASSIFICATION OF SUBJECT MATTER

G06F 21/16(2013.01)i, G06F 21/30(2013.01)i, G06N 3/08(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)

G06F 21/16; G06F 3/00; G06F 3/14; G06K 9/00; G06T 1/00; G06T 17/20; G06T 5/50; G06T 7/60; G06T 7/90; H04N 5/913; G06F 21/30; G06N 3/08

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: data identification, mark, learning model, synthesize, original

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2006-0269098 A1 (EBITANI, Kenji) 30 November 2006 See paragraphs [0002], [0031]-[0041]; and figure 3.	1-11
Y	KR 10-2017-0092631 A (EBAY INC.) 11 August 2017 See paragraphs [0107]-[0115]; claim 1; and figure 5.	1-11
A	KR 10-2017-0089752 A (KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY) 04 August 2017 See paragraphs [0100]-[0133]; and figures 4-5.	1-11
A	KR 10-2009-0109896 A (CHUNG-ANG UNIVERSITY INDUSTRY-ACADEMY COOPERATION FOUNDATION et al.) 21 October 2009 See claim 1.	1-11
A	KR 10-2014-0147461 A (SAMSUNG ELECTRONICS CO., LTD.) 30 December 2014 See paragraph [0132]; claims 1, 3; and figures 1, 8, 10.	1-11



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

05 AUGUST 2019 (05.08.2019)

Date of mailing of the international search report

05 AUGUST 2019 (05.08.2019)

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/KR2019/005367

Patent document cited in search report	Publication date	Patent family member	Publication date
US 2006-0269098 A1	30/11/2006	JP 2006-339711 A JP 4591211 B2	14/12/2006 01/12/2010
KR 10-2017-0092631 A	11/08/2017	CN 107408264 A EP 3241167 A1 US 2016-0189003 A1 US 2016-0314376 A1 US 9454713 B2 US 9767382 B2 WO 2016-109034 A1	28/11/2017 08/11/2017 30/06/2016 27/10/2016 27/09/2016 19/09/2017 07/07/2016
KR 10-2017-0089752 A	04/08/2017	CN 107067361 A KR 10-1871065 B1 US 2017-0213310 A1	18/08/2017 19/07/2018 27/07/2017
KR 10-2009-0109896 A	21/10/2009	KR 10-0950965 B1	02/04/2010
KR 10-2014-0147461 A	30/12/2014	None	

A. 발명이 속하는 기술분류(국제특허분류(IPC))
G06F 21/16(2013.01)i, G06F 21/30(2013.01)i, G06N 3/08(2006.01)i

B. 조사된 분야
조사된 최소문헌(국제특허분류를 기재)
G06F 21/16; G06F 3/00; G06F 3/14; G06K 9/00; G06T 1/00; G06T 17/20; G06T 5/50; G06T 7/60; G06T 7/90; H04N 5/913; G06F 21/30; G06N 3/08

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: 데이터 식별(data identification), 표지(mark), 학습 모델(learning model), 합성(synthesize), 원본(original)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	US 2006-0269098 A1 (KENJI EBITANI) 2006.11.30 단락 [0002], [0031]-[0041]; 및 도면 3 참조.	1-11
Y	KR 10-2017-0092631 A (이베이 인크.) 2017.08.11 단락 [0107]-[0115]; 청구항 1; 및 도면 5 참조.	1-11
A	KR 10-2017-0089752 A (한국과학기술원) 2017.08.04 단락 [0100]-[0133]; 및 도면 4-5 참조.	1-11
A	KR 10-2009-0109896 A (중앙대학교 산학협력단 등) 2009.10.21 청구항 1 참조.	1-11
A	KR 10-2014-0147461 A (삼성전자주식회사) 2014.12.30 단락 [0132]; 청구항 1, 3; 및 도면 1, 8, 10 참조.	1-11

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
 “E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌
 “L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌
 “O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌
 “P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌
 “T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌
 “X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.
 “Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.
 “&” 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일 2019년 08월 05일 (05.08.2019)	국제조사보고서 발송일 2019년 08월 05일 (05.08.2019)
--	---

ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578	심사관 진상범 전화번호 +82-42-481-8398
---	------------------------------------



국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
US 2006-0269098 A1	2006/11/30	JP 2006-339711 A JP 4591211 B2	2006/12/14 2010/12/01
KR 10-2017-0092631 A	2017/08/11	CN 107408264 A EP 3241167 A1 US 2016-0189003 A1 US 2016-0314376 A1 US 9454713 B2 US 9767382 B2 WO 2016-109034 A1	2017/11/28 2017/11/08 2016/06/30 2016/10/27 2016/09/27 2017/09/19 2016/07/07
KR 10-2017-0089752 A	2017/08/04	CN 107067361 A KR 10-1871065 B1 US 2017-0213310 A1	2017/08/18 2018/07/19 2017/07/27
KR 10-2009-0109896 A	2009/10/21	KR 10-0950965 B1	2010/04/02
KR 10-2014-0147461 A	2014/12/30	없음	