(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0059225
(43) 공개일자 2019년05월30일

(51) 국제특허분류(Int. Cl.)
G06N 3/08 (2006.01) *G06N 3/04* (2006.01)
(52) CPC특허분류
G06N 3/08 (2013.01)
G06N 3/04 (2013.01)
(21) 출원번호 10-2018-0140407
(22) 출원일자 2018년11월15일
심사청구일자 2018년11월15일
(30) 우선권주장
1020170156601 2017년11월22일 대한민국(KR)

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
이필규
인천광역시 미추홀구 매소홀로475번길 18, 17동 601호(학익동, 신동아아파트3차)
(74) 대리인
특허법인충정

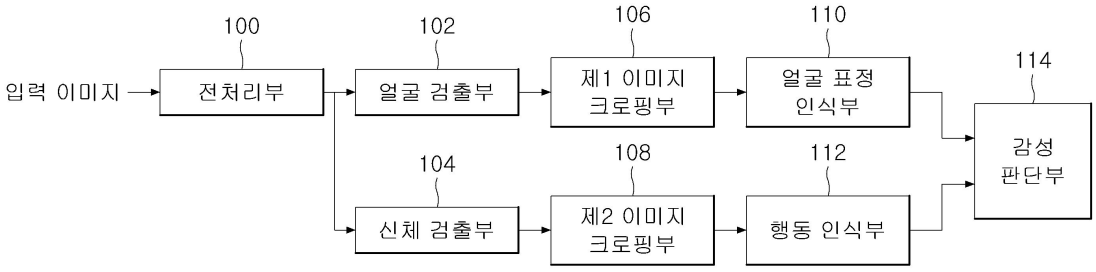
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 **점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법 및 장치**

(57) 요 약

본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치는, 감성 추정의 대상이 되는 사람이 포함된 이미지에서 사람의 얼굴 영역을 검출하기 위한 얼굴 검출부; 상기 감성 추정의 대상이 되는 사람이 포함된 이미지에서 사람의 신체 영역을 검출하기 위한 신체 검출부; 상기 검출된 얼굴 영역의 이미지를 입력으로 하여 얼굴 표정을 인식하기 위한 얼굴 표정 인식부; 상기 검출된 신체 영역의 이미지를 입력으로 하여 사람의 행동을 인식하기 위한 행동 인식부; 및 상기 인식된 얼굴 표정과 행동에 기반하여, 사람의 감성을 판단하기 위한 감성 판단부를 포함하여, 환경의 변화에 따라 감소한 인식률을 보다 적은 노력으로 빠르게 회복할 수 있어 환경 변화에 강인하게 영상을 인식할 수 있고, 점진적 딥러닝 방법을 통해 적은 데이터를 가지고 점진적으로 사람의 행동과 표정을 학습하여 두 가지 결과의 조합을 통해 사람의 감성을 정확히 추측할 수 있다.

대 표 도



이 발명을 지원한 국가연구개발사업

과제고유번호	1711055138
부처명	과학기술정보통신부
연구관리전문기관	정보통신기술진흥센터
연구사업명	ICT융합산업원천기술개발사업
연구과제명	보행자 위치공간 인지 증강 및 스포츠 경기력 분석을 위한 정밀측위 원천기술개발
기 여 율	1/1
주관기관	한국전자통신연구원
연구기간	2017.03.01 ~ 2017.12.31

명 세 서

청구범위

청구항 1

감성 추정의 대상이 되는 사람이 포함된 이미지에서 사람의 얼굴 영역을 검출하기 위한 얼굴 검출부;

상기 감성 추정의 대상이 되는 사람이 포함된 이미지에서 사람의 신체 영역을 검출하기 위한 신체 검출부;

상기 검출된 얼굴 영역의 이미지를 입력으로 하여 얼굴 표정을 인식하기 위한 얼굴 표정 인식부;

상기 검출된 신체 영역의 이미지를 입력으로 하여 사람의 행동을 인식하기 위한 행동 인식부; 및

상기 인식된 얼굴 표정과 행동에 기반하여, 사람의 감성을 판단하기 위한 감성 판단부를 포함하는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치.

청구항 2

청구항 1에 있어서,

조명으로 인한 상기 입력 이미지의 변화를 제거하기 위하여 상기 입력 이미지의 강도를 평준화하기 위한 전처리부를 더 포함하는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치.

청구항 3

청구항 2에 있어서,

이미지 크로핑을 통해 상기 검출된 얼굴 영역의 이미지에서 노이즈 영역을 제거하기 위한 제1 이미지 크로핑부; 및

이미지 크로핑을 통해 상기 검출된 신체 영역의 이미지에서 노이즈 영역을 제거하기 위한 제2 이미지 크로핑부를 더 포함하고,

상기 얼굴 표정 인식부는, 상기 노이즈 영역이 제거된 얼굴 영역의 이미지를 입력으로 하여 얼굴 표정을 인식하는 동작을 수행하며,

상기 행동 인식부는, 상기 노이즈 영역이 제거된 신체 영역의 이미지를 입력으로 하여 사람의 행동을 인식하는 동작을 수행하는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치.

청구항 4

청구항 1에 있어서,

상기 얼굴 표정 인식부는 제1 점진적 딥러닝 학습 모델에 기반하여 얼굴 표정을 인식하고,

상기 행동 인식부는 제2 점진적 딥러닝 학습 모델에 기반하여 행동을 인식하며,

상기 제1 및 제2 점진적 딥러닝 학습 모델 각각은,

(H1) 노이즈 제거 및 강도 평준화의 이미지 전처리 과정이 적용된 얼굴 및 신체 데이터베이스를 사용하여 얼굴 표정 또는 사람의 행동과 관련된 벡터 추출 능력을 지닌 사전 학습 모델을 생성하는 동작;

(H2) 상기 사전 학습 모델에 기반하여 라벨이 붙여진 초기 로컬 데이터를 사용하는 능동 학습을 통해 초기 학습 모델을 생성하는 동작;

(H3) 라벨이 붙여지지 않은 로컬 데이터를 사용하여 준지도 학습을 통해 상기 초기 학습 모델의 성능을 향상시키는 동작;

(H4) 상기 초기 학습 모델의 성능이 향상되었는지를 판단하는 동작;

(H5) 상기 초기 학습 모델의 성능이 향상되지 않은 경우, 해당 학습 데이터를 파기하고, 상기 동작 (H3)으로 진

행하는 동작;

(H6) 상기 초기 학습 모델의 성능이 향상된 경우, 라벨의 신뢰도가 소정 값 미만인지를 판단하는 동작;

(H7) 상기 라벨의 신뢰도가 소정 값 미만이 아닌 경우, 상기 동작 (H3)으로 진행하는 동작;

(H8) 상기 라벨의 신뢰도가 소정 값 미만인 경우, 능동 학습을 통해 신뢰도가 미흡한 라벨을 재검토하는 동작;

(H9) 실험 데이터를 사용하여 학습 모델의 최종 성능을 평가하는 동작;

(H10) 학습 모델의 최종 성능이 향상되지 않고 수렴하는지를 판단하는 동작;

(H11) 학습 모델의 최종 성능이 향상되는 경우, 상기 동작 (H3)으로 진행하는 동작;

(H12) 학습 모델의 최종 성능이 향상되지 않고 수렴하는 경우, 모든 환경의 학습 데이터가 사용되었는가를 판단하는 동작;

(H13) 모든 환경의 학습 데이터가 사용되지 않은 경우, 학습 데이터의 환경을 변경하고, 상기 동작 (H3)으로 진행하는 동작; 및

(H14) 모든 환경의 학습 데이터가 사용된 경우, 해당 학습 모델을 최종 학습 모델로서 저장하는 동작에 의해 생성되는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치.

청구항 5

청구항 4에 있어서,

상기 동작 (H3)은, 상기 사전 학습 모델에서 미리 학습된 파라미터를 전이학습 방식을 통해 전달받아 라벨이 붙여지지 않은 로컬 데이터를 사용하여 준지도 학습을 통해 상기 초기 학습 모델의 성능을 향상시키는 동작을 포함하는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치.

청구항 6

청구항 4에 있어서,

상기 동작 (H2) 이후에, 상기 초기 학습 모델을 사용하여 각 환경에서의 초기 성능을 측정하는 동작을 더 포함하고,

상기 측정된 초기 성능을 오름차순으로 정렬하여 가장 초기 성능이 낮은 환경부터 학습을 시작하는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치.

청구항 7

청구항 4에 있어서,

상기 동작 (H6)에서의 상기 소정 값은 0.9인, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치.

청구항 8

(A) 입력되는 복수의 샘플 이미지에서 사람의 얼굴 영역을 검출하고 검출된 얼굴 영역의 이미지들에 기반하여 얼굴 표정을 분류하기 위한 제1 점진적 딥러닝 학습 모델을 생성하는 단계;

(B) 입력되는 복수의 샘플 이미지에서 사람의 신체 영역을 검출하고 검출된 신체 영역의 이미지들에 기반하여 행동을 분류하기 위한 제2 점진적 딥러닝 학습 모델을 생성하는 단계;

(C) 감성 추정의 대상이 되는 사람이 포함된 이미지에서 사람의 얼굴 영역을 검출하는 단계;

(D) 상기 감성 추정의 대상이 되는 사람이 포함된 이미지에서 사람의 신체 영역을 검출하는 단계;

(E) 상기 검출된 얼굴 영역의 이미지를 입력으로 하여 얼굴 표정을 인식하는 단계;

(F) 상기 검출된 신체 영역의 이미지를 입력으로 하여 사람의 행동을 인식하는 단계; 및

(G) 상기 인식된 얼굴 표정과 행동에 기반하여, 사람의 감성을 판단하는 단계를 포함하는, 점진적 딥러닝 학습

을 이용한 적응적 영상 인식 기반 감성 추정 방법.

청구항 9

청구항 8에 있어서,

상기 단계 (C) 이전에, 조명으로 인한 상기 입력 이미지의 변화를 제거하기 위하여 상기 입력 이미지의 강도를 평준화하는 단계를 더 포함하는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법.

청구항 10

청구항 9에 있어서,

상기 단계 (C) 이후에, 이미지 크로핑을 통해 상기 검출된 얼굴 영역의 이미지에서 노이즈 영역을 제거하는 단계를 더 포함하고,

상기 단계 (D) 이후에, 이미지 크로핑을 통해 상기 검출된 신체 영역의 이미지에서 노이즈 영역을 제거하는 단계를 더 포함하며,

상기 단계 (E)는, 상기 노이즈 영역이 제거된 얼굴 영역의 이미지를 입력으로 하여 얼굴 표정을 인식하는 단계를 포함하고,

상기 단계 (F)는, 상기 노이즈 영역이 제거된 신체 영역의 이미지를 입력으로 하여 사람의 행동을 인식하는 단계를 포함하는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법.

청구항 11

청구항 8에 있어서,

상기 단계 (E)는 제1 점진적 딥러닝 학습 모델에 기반하여 얼굴 표정을 인식하는 단계를 포함하고,

상기 단계 (F)는 제2 점진적 딥러닝 학습 모델에 기반하여 행동을 인식하는 단계를 포함하며,

상기 제1 및 제2 점진적 딥러닝 학습 모델 각각은,

(H1) 노이즈 제거 및 강도 평준화의 이미지 전처리 과정이 적용된 얼굴 및 신체 데이터베이스를 사용하여 얼굴 표정 또는 사람의 행동과 관련된 벡터 추출 능력을 지닌 사전 학습 모델을 생성하는 단계;

(H2) 상기 사전 학습 모델에 기반하여 라벨이 붙여진 초기 로컬 데이터를 사용하는 능동 학습을 통해 초기 학습 모델을 생성하는 단계;

(H3) 라벨이 붙여지지 않은 로컬 데이터를 사용하여 준지도 학습을 통해 상기 초기 학습 모델의 성능을 향상시키는 단계;

(H4) 상기 초기 학습 모델의 성능이 향상되었는지를 판단하는 단계;

(H5) 상기 초기 학습 모델의 성능이 향상되지 않은 경우, 해당 학습 데이터를 파기하고, 상기 단계 (H3)으로 진행하는 단계;

(H6) 상기 초기 학습 모델의 성능이 향상된 경우, 라벨의 신뢰도가 소정 값 미만인지를 판단하는 단계;

(H7) 상기 라벨의 신뢰도가 소정 값 미만이 아닌 경우, 상기 단계 (H3)으로 진행하는 단계;

(H8) 상기 라벨의 신뢰도가 소정 값 미만인 경우, 능동 학습을 통해 신뢰도가 미흡한 라벨을 재검토하는 단계;

(H9) 실험 데이터를 사용하여 학습 모델의 최종 성능을 평가하는 단계;

(H10) 학습 모델의 최종 성능이 향상되지 않고 수렴하는지를 판단하는 단계;

(H11) 학습 모델의 최종 성능이 향상되는 경우, 상기 단계 (H3)으로 진행하는 단계;

(H12) 학습 모델의 최종 성능이 향상되지 않고 수렴하는 경우, 모든 환경의 학습 데이터가 사용되었는가를 판단하는 단계;

(H13) 모든 환경의 학습 데이터가 사용되지 않은 경우, 학습 데이터의 환경을 변경하고, 상기 단계 (H3)으로 진

행하는 단계; 및

(H14) 모든 환경의 학습 데이터가 사용된 경우, 해당 학습 모델을 최종 학습 모델로서 저장하는 단계에 의해 생성되는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법.

청구항 12

청구항 11에 있어서,

상기 단계 (H3)은, 상기 사전 학습 모델에서 미리 학습된 파라미터를 전이학습 방식을 통해 전달받아 라벨이 붙여지지 않은 로컬 데이터를 사용하여 준지도 학습을 통해 상기 초기 학습 모델의 성능을 향상시키는 단계를 포함하는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법.

청구항 13

청구항 11에 있어서,

상기 단계 (H2) 이후에, 상기 초기 학습 모델을 사용하여 각 환경에서의 초기 성능을 측정하는 단계를 더 포함하고,

상기 측정된 초기 성능을 오름차순으로 정렬하여 가장 초기 성능이 낮은 환경부터 학습을 시작하는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법.

청구항 14

청구항 11에 있어서,

상기 단계 (H6)에서의 상기 소정 값은 0.9인, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 기술이 발전하면서 현대 사회에서 기계가 차지하는 비중은 점점 더 커지고 있다. 특히 단순 노동 쪽 분야에 있어서는 이미 많은 기계가 사람의 노동력을 대신하고 있으며, 기계학습 기술이 발전하면서 점차 복잡한 작업들 또한 기계의 영역이 되어가고 있다.

[0003] 하지만 오검출(False Alarm) 문제 등을 비롯하여 아직까지 기계는 사람의 뇌가 갖고 있는 능력을 따라가진 못한다. 우리는 이러한 문제를 해결하기 위해 기계가 다양한 영상을 학습하고 인식하도록 연구해오고 있다. 그 중에서도 영상 인식 연구는 컴퓨터 비전은 물론 패턴 인식, 인공지능 분야에서도 매우 활발하게 연구가 이루어지고 있는 화제의 분야다. 이는 인간-컴퓨터 상호작용(Human-Computer Interaction), 인간 감성 분석(Human emotion analysis), 쌍방향 비디오(interactive video), 이미지 색인(image indexing)과 이미지 복구(image retrieval) 등 다양한 미래 기술과 연결될 수 있는 가능성을 지니고 있다.

[0004] 그래서 영상을 보다 자동적이고, 효율적이게, 그리고 정확하게 인식하기 위한 연구가 오래전부터 쭉 이어져 내려오고 있다. 특히 컨볼루션 신경망이 적용되면서부터는 인식률이 크게 상승했다. 하지만, 이러한 높은 인식률은 거의가 같은 도메인의 데이터베이스를 사용할 때만 유용하다.

[0005] 실제 세계에서는 실험 영상의 인식 대상의 다양성, 배경, 조명이나 각도 등 다양한 요인으로 인해 이보다 성능이 떨어지게 된다. 왜냐하면, 모델을 학습할 때 각 데이터가 가진 공통적인 특징에 근거하여 정보를 기록하는데, 위에 설명한 요인이 변하면 이러한 특징도 변하기 때문이다. 예를 들자면 조명의 경우, 어두운 곳에서 찍은 영상과 밝은 곳에서 찍은 영상은 그 특징이 서로 상이하다. 빛의 방향에 따라 곳곳에 그림자가 생기며, 빛의 강도에 따라 그 그림자의 진하기가 다르다. 또한, 그림자가 진 곳과 그렇지 않은 곳의 색깔 차이도 생

겨난다. 사람은 이러한 변화를 유연하게 받아들이기 때문에 약간의 환경적 요인이 변해도 인식이 가능하지만, 기계는 이러한 변화를 유연하게 받아들이지 못하기 때문에 인식률이 크게 떨어지게 된다. 이에 본 발명에서는 이러한 환경의 변화에 따라 감소한 인식률을 보다 적은 노력으로 빠르게 회복하기 위한 새로운 기법을 제안한다.

[0006] 또한, 자율주행 자동차, 인공지능 스피커, 사물 인터넷(Internet of Things) 등 이제는 익숙해진 기계학습 및 인공지능 기술이 우리의 삶에 스며들고 있다. 이러한 기술은 사람의 얼굴 표정, 음성, 나이, 성별, 행동 등 다양한 것을 인식하는 기술이 발전하면서 가능해졌다.

[0007] 이러한 인식 기술은 오래 전부터 쭉 연구되어 왔지만 최근 대두된 컨볼루션 신경망, 딥러닝 기술로 인해 크게 발전하였다. 딥러닝 기술과 빅 데이터의 결합이 인식 기술을 실용 가능한 영역까지 끌어올렸다.

[0008] 하지만, 모든 영역에 대해 빅 데이터가 존재하는 것은 아니다. 데이터의 부족으로 정확한 인식이 불가능한 경우는 많이 있다. 이를 해결하기 위해 준지도 학습(Semi-Supervised Learning)이나 능동 학습(Active Learning), 능동준지도 학습(Active Semi-Supervised Learning) 등 다양한 시도가 있어왔다.

[0009] 하지만, 이러한 방법을 통해서도 쉽게 파악하기 어려운 영역이 있으니, 그건 바로 사람의 감성이다. 사람은 말, 행동, 옷차림, 표정 등 다양한 요소에 의해서 감정을 드러낸다. 하지만, 사람은 자신의 감정을 숨길 줄 안다. 따라서 어느 한 가지만으로 상대방의 감성을 짐작하기란 쉽지 않은 일이다.

선행기술문헌

비특허문헌

[0010] (비특허문헌 0001) [1] Sinno Jialin Pan, Qiang Yang, "A Survey on Transfer Learning", IEEE Trans. Knowl. Data Eng. 22, 1345-1359 (2010).

(비특허문헌 0002) [2] Xiaoming Zhao, Shiqing Zhang, "A Review on Facial Expression Recognition - Feature Extraction and Classification", IETE Technical Review, 2016.

(비특허문헌 0003) [3] Changxing Ding, "Robust Face Recognition Via Multimodal Deep Face Representation", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 17, NO. 11, NOVEMBER 2015.

(비특허문헌 0004) [4] B. Fasel, Juergen Luettin, "Automatic facial expression analysis: a survey", International Journal of Computer Science & Engineering Survey (IJCSES), Vol.3, No.6, December 2012.

(비특허문헌 0005) [5] H. Rowley, S. Baluja, and T. Kanade, "Human face detection in visual scenes." Technical Report CMU-CS-95-158R, School of Computer Science, Carnegie Mellon University, Nov. 1995.

(비특허문헌 0006) [6] Andre Teixeira Lopes, Edilson de Aguiar, Thiago Oliveira-Santos, "A Facial Expression Recognition System Using Convolutional Networks", 28th SIBGRAPI Conference on Aug. 2015.

(비특허문헌 0007) [7] Ken Chatfield, Karen Simonyan, Andrea Vedaldi, Andrew Zisserman, "Return of the Devil in the Details: Delving Deep into Convolutional Nets", arXiv:1405.3531v4 [cs.CV] 5 Nov 2014.

(비특허문헌 0008) [8] Karen Simonyan, Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition", arXiv:1409.1556v6 [cs.CV] 10 Apr 2015.

(비특허문헌 0009) [9] Cohn, D., Ghahramani, Z., Jordan, M., "Active learning with statistical models." Journal of Artificial Intelligence Research, Vol.4, pp. 129-145, 1999.

(비특허문헌 0010) [10] Riccardi, G. and Hankkani-Tur, D., "Active learning: theory and applications to automatic speech recognition", IEEE Transactions on Speech and Audio Processing, Vol.13, No.4, pp. 504-511, 2005.

(비특허문헌 0011) [11] Andre Teixeira Lopes, Edilson de Aguiar, Alberto F. De Souza, Thiago Oliverira-Santos, "Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order", Pattern Recognition Volume 61, January 2017, Pages 610-628.

(비특허문헌 0012) [12] Gil-Jin Jang, Ahra Jo, Jeong-Sik Park, Yong-Ho Seo, "Video-based Facial

Emotion Recognition using Active Shape Models and StatisticalPattern Recognizers", JIIBC, VOL. 14 NO. 3,June 2014. pp.139-146.

발명의 내용

해결하려는 과제

[0011] 본 발명이 해결하고자 하는 과제는 환경의 변화에 따라 감소한 인식률을 보다 적은 노력으로 빠르게 회복할 수 있어 환경 변화에 강인하게 영상을 인식할 수 있고, 점진적 딥러닝 방법을 통해 적은 데이터를 가지고 점진적으로 사람의 행동과 표정을 학습하여 두 가지 결과의 조합을 통해 사람의 감성을 정확히 추측할 수 있는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법을 제공하는 것이다.

[0012] 본 발명이 해결하고자 하는 다른 과제는 환경의 변화에 따라 감소한 인식률을 보다 적은 노력으로 빠르게 회복할 수 있어 환경 변화에 강인하게 영상을 인식할 수 있고, 점진적 딥러닝 방법을 통해 적은 데이터를 가지고 점진적으로 사람의 행동과 표정을 학습하여 두 가지 결과의 조합을 통해 사람의 감성을 정확히 추측할 수 있는, 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치를 제공하는 것이다.

과제의 해결 수단

[0013] 상기 과제를 해결하기 위한 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치는,

[0014] 감성 추정의 대상이 되는 사람이 포함된 이미지에서 사람의 얼굴 영역을 검출하기 위한 얼굴 검출부;

[0015] 상기 감성 추정의 대상이 되는 사람이 포함된 이미지에서 사람의 신체 영역을 검출하기 위한 신체 검출부;

[0016] 상기 검출된 얼굴 영역의 이미지를 입력으로 하여 얼굴 표정을 인식하기 위한 얼굴 표정 인식부;

[0017] 상기 검출된 신체 영역의 이미지를 입력으로 하여 사람의 행동을 인식하기 위한 행동 인식부; 및

[0018] 상기 인식된 얼굴 표정과 행동에 기반하여, 사람의 감성을 판단하기 위한 감성 판단부를 포함한다.

[0019] 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치는, 조명으로 인한 상기 입력 이미지의 변화를 제거하기 위하여 상기 입력 이미지의 강도를 평준화하기 위한 전처리부를 더 포함할 수 있다.

[0020] 또한, 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치는,

[0021] 이미지 크로핑을 통해 상기 검출된 얼굴 영역의 이미지에서 노이즈 영역을 제거하기 위한 제1 이미지 크로핑부; 및

[0022] 이미지 크로핑을 통해 상기 검출된 신체 영역의 이미지에서 노이즈 영역을 제거하기 위한 제2 이미지 크로핑부를 더 포함하고,

[0023] 상기 얼굴 표정 인식부는, 상기 노이즈 영역이 제거된 얼굴 영역의 이미지를 입력으로 하여 얼굴 표정을 인식하는 동작을 수행하며,

[0024] 상기 행동 인식부는, 상기 노이즈 영역이 제거된 신체 영역의 이미지를 입력으로 하여 사람의 행동을 인식하는 동작을 수행할 수 있다.

[0025] 또한, 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치에 있어서, 상기 얼굴 표정 인식부는 제1 점진적 딥러닝 학습 모델에 기반하여 얼굴 표정을 인식하고, 상기 행동 인식부는 제2 점진적 딥러닝 학습 모델에 기반하여 행동을 인식하며,

[0026] 상기 제1 및 제2 점진적 딥러닝 학습 모델 각각은,

[0027] (H1) 노이즈 제거 및 강도 평준화의 이미지 전처리 과정이 적용된 얼굴 및 신체 데이터베이스를 사용하여 얼굴 표정 또는 사람의 행동과 관련된 벡터 추출 능력을 지닌 사전 학습 모델을 생성하는 동작;

[0028] (H2) 상기 사전 학습 모델에 기반하여 라벨이 붙여진 초기 로컬 데이터를 사용하는 능동 학습을 통해 초기 학습

모델을 생성하는 동작;

[0029] (H3) 라벨이 붙여지지 않은 로컬 데이터를 사용하여 준지도 학습을 통해 상기 초기 학습 모델의 성능을 향상시키는 동작;

[0030] (H4) 상기 초기 학습 모델의 성능이 향상되었는지를 판단하는 동작;

[0031] (H5) 상기 초기 학습 모델의 성능이 향상되지 않은 경우, 해당 학습 데이터를 파기하고, 상기 동작 (H3)으로 진행하는 동작;

[0032] (H6) 상기 초기 학습 모델의 성능이 향상된 경우, 라벨의 신뢰도가 소정 값 미만인지를 판단하는 동작;

[0033] (H7) 상기 라벨의 신뢰도가 소정 값 미만이 아닌 경우, 상기 동작 (H3)으로 진행하는 동작;

[0034] (H8) 상기 라벨의 신뢰도가 소정 값 미만인 경우, 능동 학습을 통해 신뢰도가 미흡한 라벨을 재검토하는 동작;

[0035] (H9) 실험 데이터를 사용하여 학습 모델의 최종 성능을 평가하는 동작;

[0036] (H10) 학습 모델의 최종 성능이 향상되지 않고 수렴하는지를 판단하는 동작;

[0037] (H11) 학습 모델의 최종 성능이 향상되는 경우, 상기 동작 (H3)으로 진행하는 동작;

[0038] (H12) 학습 모델의 최종 성능이 향상되지 않고 수렴하는 경우, 모든 환경의 학습 데이터가 사용되었는가를 판단하는 동작;

[0039] (H13) 모든 환경의 학습 데이터가 사용되지 않은 경우, 학습 데이터의 환경을 변경하고, 상기 동작 (H3)으로 진행하는 동작; 및

[0040] (H14) 모든 환경의 학습 데이터가 사용된 경우, 해당 학습 모델을 최종 학습 모델로서 저장하는 동작에 의해 생성될 수 있다.

[0041] 또한, 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치에 있어서, 상기 동작 (H3)은, 상기 사전 학습 모델에서 미리 학습된 파라미터를 전이학습 방식을 통해 전달받아 라벨이 붙여지지 않은 로컬 데이터를 사용하여 준지도 학습을 통해 상기 초기 학습 모델의 성능을 향상시키는 동작을 포함할 수 있다.

[0042] 또한, 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치에 있어서, 상기 동작 (H2) 이후에, 상기 초기 학습 모델을 사용하여 각 환경에서의 초기 성능을 측정하는 동작을 더 포함하고,

[0043] 상기 측정된 초기 성능을 오름차순으로 정렬하여 가장 초기 성능이 낮은 환경부터 학습을 시작할 수 있다.

[0044] 또한, 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치에 있어서, 상기 동작 (H6)에서의 상기 소정 값은 0.9일 수 있다.

[0045] 상기 다른 과제를 해결하기 위한 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법은,

[0046] (A) 입력되는 복수의 샘플 이미지에서 사람의 얼굴 영역을 검출하고 검출된 얼굴 영역의 이미지들에 기반하여 얼굴 표정을 분류하기 위한 제1 점진적 딥러닝 학습 모델을 생성하는 단계;

[0047] (B) 입력되는 복수의 샘플 이미지에서 사람의 신체 영역을 검출하고 검출된 신체 영역의 이미지들에 기반하여 행동을 분류하기 위한 제2 점진적 딥러닝 학습 모델을 생성하는 단계;

[0048] (C) 감성 추정의 대상이 되는 사람이 포함된 이미지에서 사람의 얼굴 영역을 검출하는 단계;

[0049] (D) 상기 감성 추정의 대상이 되는 사람이 포함된 이미지에서 사람의 신체 영역을 검출하는 단계;

[0050] (E) 상기 검출된 얼굴 영역의 이미지를 입력으로 하여 얼굴 표정을 인식하는 단계;

[0051] (F) 상기 검출된 신체 영역의 이미지를 입력으로 하여 사람의 행동을 인식하는 단계; 및

[0052] (G) 상기 인식된 얼굴 표정과 행동에 기반하여, 사람의 감성을 판단하는 단계를 포함할 수 있다.

[0053] 또한, 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법은, 상

기 단계 (C) 이전에, 조명으로 인한 상기 입력 이미지의 변화를 제거하기 위하여 상기 입력 이미지의 강도를 평준화하는 단계를 더 포함할 수 있다.

[0054] 또한, 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법은, 상기 단계 (C) 이후에, 이미지 크로핑을 통해 상기 검출된 얼굴 영역의 이미지에서 노이즈 영역을 제거하는 단계를 더 포함하고,

[0055] 상기 단계 (D) 이후에, 이미지 크로핑을 통해 상기 검출된 신체 영역의 이미지에서 노이즈 영역을 제거하는 단계를 더 포함하며,

[0056] 상기 단계 (E)는, 상기 노이즈 영역이 제거된 얼굴 영역의 이미지를 입력으로 하여 얼굴 표정을 인식하는 단계를 포함하고,

[0057] 상기 단계 (F)는, 상기 노이즈 영역이 제거된 신체 영역의 이미지를 입력으로 하여 사람의 행동을 인식하는 단계를 포함할 수 있다.

[0058] 또한, 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법에 있어서, 상기 단계 (E)는 제1 점진적 딥러닝 학습 모델에 기반하여 얼굴 표정을 인식하는 단계를 포함하고,

[0059] 상기 단계 (F)는 제2 점진적 딥러닝 학습 모델에 기반하여 행동을 인식하는 단계를 포함하며,

[0060] 상기 제1 및 제2 점진적 딥러닝 학습 모델 각각은,

[0061] (H1) 노이즈 제거 및 강도 평준화의 이미지 전처리 과정이 적용된 얼굴 및 신체 데이터베이스를 사용하여 얼굴 표정 또는 사람의 행동과 관련된 벡터 추출 능력을 지닌 사전 학습 모델을 생성하는 단계;

[0062] (H2) 상기 사전 학습 모델에 기반하여 라벨이 붙여진 초기 로컬 데이터를 사용하는 능동 학습을 통해 초기 학습 모델을 생성하는 단계;

[0063] (H3) 라벨이 붙여지지 않은 로컬 데이터를 사용하여 준지도 학습을 통해 상기 초기 학습 모델의 성능을 향상시키는 단계;

[0064] (H4) 상기 초기 학습 모델의 성능이 향상되었는지를 판단하는 단계;

[0065] (H5) 상기 초기 학습 모델의 성능이 향상되지 않은 경우, 해당 학습 데이터를 파기하고, 상기 단계 (H3)으로 진행하는 단계;

[0066] (H6) 상기 초기 학습 모델의 성능이 향상된 경우, 라벨의 신뢰도가 소정 값 미만인지를 판단하는 단계;

[0067] (H7) 상기 라벨의 신뢰도가 소정 값 미만이 아닌 경우, 상기 단계 (H3)으로 진행하는 단계;

[0068] (H8) 상기 라벨의 신뢰도가 소정 값 미만인 경우, 능동 학습을 통해 신뢰도가 미흡한 라벨을 재검토하는 단계;

[0069] (H9) 실험 데이터를 사용하여 학습 모델의 최종 성능을 평가하는 단계;

[0070] (H10) 학습 모델의 최종 성능이 향상되지 않고 수렴하는지를 판단하는 단계;

[0071] (H11) 학습 모델의 최종 성능이 향상되는 경우, 상기 단계 (H3)으로 진행하는 단계;

[0072] (H12) 학습 모델의 최종 성능이 향상되지 않고 수렴하는 경우, 모든 환경의 학습 데이터가 사용되었는가를 판단하는 단계;

[0073] (H13) 모든 환경의 학습 데이터가 사용되지 않은 경우, 학습 데이터의 환경을 변경하고, 상기 단계 (H3)으로 진행하는 단계; 및

[0074] (H14) 모든 환경의 학습 데이터가 사용된 경우, 해당 학습 모델을 최종 학습 모델로서 저장하는 단계에 의해 생성될 수 있다.

[0075] 또한, 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법에 있어서, 상기 단계 (H3)은, 상기 사전 학습 모델에서 미리 학습된 파라미터를 전이학습 방식을 통해 전달받아 라벨이 붙여지지 않은 로컬 데이터를 사용하여 준지도 학습을 통해 상기 초기 학습 모델의 성능을 향상시키는 단계를 포함할 수 있다.

[0076] 또한, 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법은, 상

기 단계 (H2) 이후에, 상기 초기 학습 모델을 사용하여 각 환경에서의 초기 성능을 측정하는 단계를 더 포함하고,

[0077] 상기 측정된 초기 성능을 오름차순으로 정렬하여 가장 초기 성능이 낮은 환경부터 학습을 시작할 수 있다.

[0078] 또한, 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법에 있어서, 상기 단계 (H6)에서의 상기 소정 값은 0.9일 수 있다.

발명의 효과

[0079] 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법 및 장치에 의하면, 환경의 변화에 따라 감소한 인식률을 보다 적은 노력으로 빠르게 회복할 수 있어 환경 변화에 강인하게 영상을 인식할 수 있고, 점진적 딥러닝 방법을 통해 적은 데이터를 가지고 점진적으로 사람의 행동과 표정을 학습하여 두 가지 결과의 조합을 통해 사람의 감성을 정확히 추측할 수 있다.

도면의 간단한 설명

[0080] 도 1은 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치의 블록도.

도 2는 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법의 흐름도.

도 3은 점진적 딥러닝 학습 모델 생성 과정을 도시한 도면.

도 4는 영상 인식 화면을 도시한 도면.

도 5는 VGG Verydeep 16 네트워크 모델을 도시한 도면.

도 6은 전이학습 방법을 설명하기 위한 도면.

도 7은 점진적 딥러닝 학습을 통한 적응적 영상 인식의 성능을 도시한 도면.

도 8은 환경 변화에 따른 초기 인식률 증가 그래프.

도 9는 필요 노동력 지표를 도시한 도면.

도 10은 행동 인식 그래프를 도시한 도면.

도 11은 얼굴 표정 인식 그래프를 도시한 도면.

발명을 실시하기 위한 구체적인 내용

[0081] 본 발명의 목적, 특정한 장점들 및 신규한 특징들은 첨부된 도면들과 연관되어지는 이하의 상세한 설명과 바람직한 실시예들로부터 더욱 명백해질 것이다.

[0082] 이에 앞서 본 명세서 및 청구범위에 사용된 용어나 단어는 통상적이고 사전적인 의미로 해석되어서는 아니되며, 발명자가 그 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념을 적절하게 정의할 수 있는 원칙에 입각하여 본 발명의 기술적 사상에 부합되는 의미와 개념으로 해석되어야 한다.

[0083] 본 명세서에서 각 도면의 구성요소들에 참조번호를 부가함에 있어서, 동일한 구성 요소들에 한해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 번호를 가지도록 하고 있음에 유의하여야 한다.

[0084] 또한, "제1", "제2", "일면", "타면" 등의 용어는, 하나의 구성요소를 다른 구성요소로부터 구별하기 위해 사용되는 것으로, 구성요소가 상기 용어들에 의해 제한되는 것은 아니다.

[0085] 이하, 본 발명을 설명함에 있어, 본 발명의 요지를 불필요하게 흐릴 수 있는 관련된 공지 기술에 대한 상세한 설명은 생략한다.

[0086] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시형태를 상세히 설명하기로 한다.

[0087] 본 발명에서는 보다 진솔한 사람의 감성을 이해하고 파악하기 위해 얼굴 표정과 행동에 주목했다. 행동과 표정은 부단한 노력과 훈련이 없으면 쉽게 감추거나 바꿀 수 없기 때문이다. 이에 본 발명에서는 보다 확실하게 사

람의 감정을 추측하기 위해 새로운 기법을 제안하였다.

[0088] 도 1은 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치의 블록도이고, 도 2는 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법의 흐름도이다.

[0089] 본 발명의 일 실시예에서 도 2에 도시된 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법은, 도 1에 도시된 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치에 의해 수행된다. 하지만, 본 발명은 이에 한정되지 않고, 도 2에 도시된 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법은 메모리를 포함하는 프로세서에 의해 수행될 수 있다.

[0090] 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치 및 방법에서는, 먼저 입력받은 이미지에서 사람의 신체와 얼굴을 잘 검출할 수 있도록 전처리 과정을 거친다. 그 후, 얼굴 검출부와 신체 검출부를 통해 각각 사각형 영역을 찾아낸다. 그렇게 찾아낸 영역을 잘라내어 적응적 딥러닝 학습 모델로 얼굴 표정과 행동을 인식한다. 인식이 끝나면 결과 값의 조합에 따라 종합적인 감성을 판단한다.

[0091] 도 1에 도시된 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치는, 감성 추정의 대상이 되는 사람이 포함된 이미지인 입력 이미지에서 사람의 얼굴 영역을 검출하기 위한 얼굴 검출부(102), 상기 감성 추정의 대상이 되는 사람이 포함된 이미지인 입력 이미지에서 사람의 신체 영역을 검출하기 위한 신체 검출부(104), 상기 검출된 얼굴 영역의 이미지를 입력으로 하여 얼굴 표정을 인식하기 위한 얼굴 표정 인식부(110), 상기 검출된 신체 영역의 이미지를 입력으로 하여 사람의 행동을 인식하기 위한 행동 인식부(112), 및 상기 인식된 얼굴 표정과 행동에 기반하여, 사람의 감성을 판단하기 위한 감성 판단부(114)를 포함한다.

[0092] 또한, 도 1에 도시된 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치는, 조명으로 인한 상기 입력 이미지의 변화를 제거하기 위하여 상기 입력 이미지의 강도를 평준화하기 위한 전처리부(100)를 더 포함한다.

[0093] 또한, 도 1에 도시된 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치는, 이미지 크로핑을 통해 상기 검출된 얼굴 영역의 이미지에서 노이즈 영역을 제거하기 위한 제1 이미지 크로핑부(106) 및 이미지 크로핑을 통해 상기 검출된 신체 영역의 이미지에서 노이즈 영역을 제거하기 위한 제2 이미지 크로핑부(108)를 더 포함하고, 상기 얼굴 표정 인식부(110)는, 상기 노이즈 영역이 제거된 얼굴 영역의 이미지를 입력으로 하여 얼굴 표정을 인식하는 동작을 수행하며, 상기 행동 인식부(112)는, 상기 노이즈 영역이 제거된 신체 영역의 이미지를 입력으로 하여 사람의 행동을 인식하는 동작을 수행한다.

[0094] 도 1 및 도 2를 참조하면, 단계 S200에서, 얼굴 표정 인식부(110)는, 입력되는 복수의 샘플 이미지에서 사람의 얼굴 영역을 검출하고 검출된 얼굴 영역의 이미지들에 기반하여 얼굴 표정을 분류하기 위한 제1 점진적 딥러닝 학습 모델을 생성한다.

[0095] 단계 S202에서, 행동 인식부(112)는, 입력되는 복수의 샘플 이미지에서 사람의 신체 영역을 검출하고 검출된 신체 영역의 이미지들에 기반하여 행동을 분류하기 위한 제2 점진적 딥러닝 학습 모델을 생성한다.

[0096] 단계 S204에서, 전처리부(100)는, 조명으로 인한 입력 이미지의 변화를 제거하기 위하여 상기 입력 이미지의 강도를 평준화한다.

[0097] 단계 S206에서, 얼굴 검출부(102)는, 강도가 평준화된 감성 추정의 대상이 되는 사람이 포함된 이미지에서 사람의 얼굴 영역을 검출한다.

[0098] 단계 S208에서, 신체 검출부(104)는, 강도가 평준화된 감성 추정의 대상이 되는 사람이 포함된 이미지에서 신체 영역을 검출한다.

[0099] 단계 S210에서, 제1 이미지 크로핑부(106)는, 이미지 크로핑을 통해 상기 검출된 얼굴 영역의 이미지에서 노이즈 영역을 제거한다.

[0100] 단계 S212에서, 제2 이미지 크로핑부(106)는, 이미지 크로핑을 통해 상기 검출된 신체 영역의 이미지에서 노이즈 영역을 제거한다.

[0101] 단계 S214에서, 얼굴 표정 인식부(110)는, 노이즈 영역이 제거된 얼굴 영역의 이미지를 입력으로 하여 제1 점진

적 딥러닝 학습 모델에 기반하여 얼굴 표정을 인식한다.

[0102] 단계 S216에서, 행동 인식부(112)는, 노이즈 영역이 제거된 신체 영역의 이미지를 입력으로 하여 제2 점진적 딥러닝 학습 모델에 기반하여 행동을 인식한다.

[0103] 단계 S218에서, 감성 판단부(114)는, 상기 인식된 얼굴 표정과 행동에 기반하여, 사람의 감성을 판단한다.

[0104] 하기에, 도 1에 도시된 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 장치 및 도 2에 도시된 본 발명의 일 실시예에 의한 점진적 딥러닝 학습을 이용한 적응적 영상 인식 기반 감성 추정 방법에 대해 상세히 설명하기로 한다.

[0106] 점진적 딥러닝 학습의 배경 이론

[0107] 컨볼루션 신경망을 통해 새로운 데이터를 학습할 때는 오버피팅 등의 문제가 발생하기 쉽다. 오버피팅이란 이미 학습하여 생성된 모델이 기존의 데이터 내에서는 데이터 라벨을 잘 분류해내지만, 새로운 데이터 내에서는 제 성능을 발휘하지 못하여 인식률이 크게 감소하는 것을 말한다. 학습을 너무 많이 했거나, 튜닝을 지나치게 하는 등 그 이유는 여러 가지 있을 수 있지만 가장 단순한 해결 방안은 바로 새로운 데이터도 잘 분류해낼 수 있도록 다양하고 많은 양의 학습 데이터를 사용하여 학습 모델을 생성하는 것이다.

[0108] 하지만 세상에는 매일같이 새로운 종류의 데이터가 쏟아져 나온다. 이 모든 것에 일일이 라벨을 붙여서 지도학습을 하는 것은 불가능하다. 너무 많은 노동력과 비용이 필요하기 때문이다. 사실 데이터에 라벨을 붙이는 이런 단순한 작업은 컴퓨터가 수행하기에 좋은 분야지만, 오버피팅 문제로 인해 정작 중요한 라벨을 제대로 분류해내질 못하니 믿고 맡길 수도 없다. 이러한 딜레마를 해결하고 노동력과 비용을 줄이면서 새로운 데이터에 라벨을 붙여 뛰어난 학습 모델을 생성하기 위한 방법이 바로 점진적 딥러닝 학습 방법이다.

[0110] 점진적 딥러닝 학습 모델

[0111] 지도 학습은 사람이 주어진 모든 데이터에 일일이 라벨을 붙여서 학습을 하는 가장 단순한 방법이다. 가장 확실하게 성능이 보장되지만 동시에 가장 많은 노동력을 필요로 한다.

[0112] 반면에 능동 학습은 컴퓨터가 어느 데이터에 라벨을 붙일지 능동적으로 판단하여 사람에게 그 라벨을 붙이도록 하는 방법이다. 능동 학습은 이러한 방법을 통해 학습에 사용하기 위해 라벨을 붙일 데이터의 수를 줄여 필요로 하는 노동력을 감소시킨다.

[0113] 또한 준지도 학습이란 최초에 감독관에 의해 라벨이 붙여진 데이터(Labeled Data)를 사용하여 학습한 모델을 통해 라벨이 붙여지지 않은 데이터(Unlabeled data)에 컴퓨터가 자동으로 새로운 라벨을 부여하여 반복적으로 학습함으로써 인식률을 높이는 방법이다. 이 방법은 컴퓨터에 의한 완전 자동화를 통해 새로운 데이터에 라벨을 붙이는데 필요한 노동력을 크게 감소시킨다. 바로 이때 문제가 되는 것이 바로 위에서 설명했던 오버피팅이다. 초기에 주어진 데이터와 유사하지 않은 새로운 데이터가 준지도 학습 과정에 들어가게 되면 매우 높은 확률로 잘못 분류된 라벨을 갖는 데이터가 생성되어 성능에 한계가 발생한다.

[0114] 마지막으로 점진적 딥러닝 학습은 능동 학습과 준지도 학습을 결합하여 장점만 취한 방법으로 준지도 학습을 통해 필요한 노동력을 감소시키고 능동 학습을 통해 올바른 라벨을 붙여줌으로써 이때 발생하는 오버피팅 문제를 해소한다.

[0116] 점진적 딥러닝 학습 모델 생성

[0117] 도 3은 얼굴 표정 인식부(110)에서 얼굴 표정을 인식하기 위하여 사용하는 제1 점진적 딥러닝 학습 모델 또는 행동 인식부(112)가 행동을 인식하기 위하여 사용하는 제2 점진적 딥러닝 학습 모델을 생성하는 과정을 도시한 도면이다. 제1 점진적 딥러닝 학습 모델의 생성 과정과 제2 점진적 딥러닝 학습 모델의 생성 과정은 분류 대상이 얼굴 표정을 분류하기 위한 것인지 또는 행동을 분류하기 위한 것인지만 상이할 뿐, 전반적인 과정은 동일하므로, 본 발명의 일 실시예에서는 도 3을 참조하여, 하나의 점진적 딥러닝 학습 모델에 대해 설명하기로 한다.

[0118] 먼저, 단계 S300에서 데이터베이스를 사용해 충분히 학습하여 뛰어난 벡터 추출 능력을 지닌 사전 학습 모델을 만들어 낸다. 다음, 단계 S302에서 사전 학습 모델을 기반으로 능동 학습을 통해 초기 학습 모델을 생성한다.

이후 단계 S304에서 준지도 학습과 능동 학습을 번갈아 사용하며 점진적으로 학습 모델의 성능을 향상시킨다. 단계 S306에서 초기 학습 모델의 성능이 향상되었는지를 판단하고, 만약 준지도 학습 결과 이전보다 성능이 감소하였을 경우에는 단계 S308에서 해당 학습 데이터를 파기하고 단계 S304로 진행하여 다음 학습 데이터를 사용하여 다시 준지도 학습을 실시한다.

[0119] 준지도 학습 결과 무사히 성능이 향상되었을 경우에는 단계 S312에서 능동 학습을 실시하여 신뢰도가 미흡한 라벨을 재검토한다. 물론 모든 학습 데이터의 라벨을 전부 재검토하지는 않는다. 작업의 효율성을 높이기 위해 단계 S310에서 준지도 학습 과정에서 예측한 각 라벨의 신뢰도가 0.9 미만인지를 판단하고, 신뢰도 판단 결과에 기반하여 재검토해야 할 후보를 결정짓는다. 신뢰도는 0에서 1사이의 숫자이며 본 발명의 일 실시예에서는 분류 기준을 0.9로 잡았다.

[0120] 능동 학습까지 마치고 나면 단계 S314에서 실험 데이터에 의해 최종 성능을 평가한다. 단계 S316에서, 학습 모델의 최종 성능이 향상되지 않고 수렴하는지를 판단하고, 학습 모델의 최종 성능이 향상되는 경우, 상기 단계 S304로 진행하며, 학습 모델의 최종 성능이 향상되지 않고 수렴하는 경우, 단계 S318에서, 모든 환경의 학습 데이터가 사용되었는가를 판단한다.

[0121] 모든 환경의 학습 데이터가 사용되지 않은 경우, 단계 S320에서, 학습 데이터의 환경을 변경하고, 상기 단계 S304로 진행하며, 모든 환경의 학습 데이터가 사용된 경우, 단계 S322에서, 해당 학습 모델을 최종 학습 모델로서 저장한다.

[0122] 본 발명의 일 실시예에서는, 본 발명에서 제안한 시스템이 한 바퀴를 돌 때마다 최종 성능을 기록하여 그 성능이 일차적으로 수렴하면 학습 데이터의 환경을 바꾸고, 이차적으로 수렴하여 환경의 영향에 강인해지면 최종 학습 모델로서 저장한다.

[0124] 영상 인식

[0125] 영상 인식 시스템

[0126] 영상 인식의 전반적인 절차는 다음과 같다. 우선 입력된 영상에 정규화(Normalization)나 분할(Segmentation)과 같은 전처리 과정을 거친다. 정규화는 영상의 배경이나 조명으로 인한 변화 등을 최대한 배제하여 입력된 영상이 되도록 균일한 특징을 가지도록 하는 기법을 말하며, 분할은 주어진 전체 이미지 중에서 필요로 하는 영상이 포함된 일부만을 추출하는 기법이다. 이러한 전처리 과정은 인식에 방해가 되는 노이즈 등을 제거하는 역할을 하기 때문에 정확성에 큰 영향을 준다.

[0127] 특징 벡터란 우리가 인식하고자 하는 영상이 갖는 특징을 기계가 사람처럼 파악할 수 있도록 기호화한 것이다. 따라서 영상을 인식하기 위해서는 전처리 과정을 마친 결과에 변형 추출(Deformation Extraction)이나 동작 추출(Motion Extraction) 등의 특징 추출 기법을 적용하여 명확하게 구분할 수 있는 특징 벡터를 뽑아내야 한다. 변형 추출은 영상의 기하학적인 변형을 포함하는 이미지나 모델을 특징 벡터로 표현하는 것이며, 동작 추출은 시간의 흐름에 따른 영상의 변화량을 특징 벡터로 표현한 것이다.

[0128] 이렇게 추출된 특징 벡터는 영상 특징 표현(Image feature representation)을 통해 분류(Classification)될 수 있게 정리된다. 이후에는 각종 분류기(Classifier)를 사용하여 영상 특징 표현을 입력받은 라벨 별로 구분하여 학습한다. 그렇게 학습이 끝나면 분류기는 각 라벨 별로 영상 특징 표현을 구분할 수 있는 어떤 값을 갖게 된다. 분류기는 이 값을 가지고 새로운 영상을 입력 받았을 때 추출된 특징 벡터를 분류하여 해당하는 라벨을 예측하게 된다.

[0130] 고효율 적응적 영상 인식 기법

[0131] 노이즈 제거

[0132] 일반적으로 아무런 영상 처리를 거치지 않은 이미지를 사용하여 영상 인식을 실행할 경우 그 인식률이 매우 저조하다. 그 이유는 원본 이미지가 배경 등의 노이즈를 가지고 있기 때문이다. 이러한 요소는 사실상 인식하고자 하는 대상과는 아무런 관련이 없으며 일정하지가 않다. 따라서 이러한 노이즈는 원활한 학습을 방해한다.

[0133] 왜냐하면, 학습 시 분류기는 이러한 노이즈도 모두 변수로 받아들여 학습하기 때문이다. 분류기는 입력 받은 이

미지의 어떠한 부분이 노이즈이고 또 어떠한 부분이 노이즈가 아닌지 구분할 수 없다. 따라서 분류기의 인식 성능을 높이기 위해서는 이러한 노이즈를 제거해줄 필요가 있다. 이러한 노이즈를 제거하는 방법 중 이미지 크로핑(Image cropping)이라는 기법이 있다. 사진 또는 삽화의 불필요한 부분을 다듬는 기술인데 이를 통해서 영상을 분류하기 위해 꼭 필요한 부분만 학습에 사용할 수 있다.

[0134]

[0135] 강도 평준화

[0136] 빛의 밝기는 이미지를 분류하는데 있어서 큰 장애가 되는 노이즈 요소다. 그림자가 생기면 해당 위치의 픽셀이 갖는 색에 해당하는 코드가 바뀌기 때문이다. 즉, 같은 장소를 찍은 영상이라고 하더라고 빛을 받아 밝은 톤의 이미지와 그림자가 진 어두운 톤의 이미지는 컴퓨터가 볼 때 전혀 다른 픽셀 값을 갖는 이미지가 되는 것이다. 이렇게 조명에 의한 간섭을 최소화하는 기법을 강도 평준화(Intensity normalization)라고 한다. 본 발명에서는 빛의 강도가 가장 약한 곳과 가장 강한 곳의 가우시안 가중치 평균값(Gaussian weighted average)을 사용하여 이러한 노이즈를 제거하였다.

[0138] 깊은 컨볼루션 신경망

[0139] 컨볼루션 신경망은 최근에 주목을 받고 있는 새로운 기계 학습 방식 중 하나이다. 컨볼루션 신경망은 사용자가 미리 정해놓은 네트워크의 구조를 따라 특징 벡터의 추출과 분류를 반복(iteration)을 하면서, 각 이미지에서 적절한 특징 벡터를 뽑아 실측 자료(Ground Truth)의 라벨로 분류가 되도록 네트워크를 학습해 나간다. 이렇게 학습된 네트워크는 학습 데이터에 따라 각각 물체나 사람 등 목적으로 하는 대상을 분류하기에 딱 맞는 전용 특징 벡터를 뽑을 수 있게끔 된다. 즉, 특징 벡터를 뽑아서 분류하는 기존의 학습 방식과 달리, 대상이 잘 분류가 되도록 조정된 특징 벡터를 뽑게끔 네트워크의 파라미터를 학습한다.

[0140] 영상 인식을 위한 학습에는 VGG verydeep 16 네트워크를 사용했다. 이 네트워크는 기존의 VGG 네트워크에 매우 작은(3x3) 컨볼루션 필터를 추가하여 대규모 이미지 인식에 있어서 보다 뛰어난 성능을 보인다. 도 5는 실험에 사용한 VGG verydeep 16 네트워크의 구조를 보여준다.

[0142] 전이 학습

[0143] 기계학습 시 학습 데이터와 테스트 데이터는 그 분포나 분산 등에 있어서 같은 조건을 가져야만 유의미한 결과를 얻을 수 있다. 만약 이러한 조건이 다르면 제대로 된 결과를 얻을 수 없으며, 이럴 때마다 새롭게 다시 데이터를 모으고 학습하는 것은 매우 큰 비용과 시간을 필요로 한다. 이러한 문제를 해결하여 비용과 시간을 줄이기 위한 방법 중 하나가 바로 이전에 학습한 모델이 가지고 있는 지식을 전달하는 전이 학습 기법이다. 전이 학습은 양이 풍부한 도메인의 데이터를 사용하여 사전학습한 뒤, 학습된 파라미터에 의해 양이 적은 도메인의 데이터를 미세 조정하여 보다 효과적이고 빠르게 네트워크 모델을 학습시킨다. 본 발명에서는 전이 학습 기법을 사용하여 이전에 학습한 파라미터들을 인계하였으며, 재학습 시간을 단축하였다. 도 6은 전이 학습을 도시한 도면이다.

[0145] 점진적 딥러닝 학습 모델 생성에 대한 실험 및 결과

[0146] 실험 절차

[0147] 영상의 인식률을 더 높이기 위해 데이터베이스에 노이즈 제거 및 강도 평준화 등의 이미지 전처리 과정을 적용하였다. 그 후, 이렇게 수정된 데이터를 VGG기반 딥러닝 네트워크 모델로 학습하여, 점진적 딥러닝 학습을 위한 초기 학습 모델을 생성하였다.

[0148] 초기 학습 모델을 학습하는 데는 새롭게 수집한 라벨이 없는 실험 데이터를 사용하였다. 이 실험 데이터는 약 200장의 영상으로 구성되었으며, 조명이 오른쪽에서 들어오는 환경이었기에 Right라고 이름 붙였다. 그리고 기존에 만들어진 사전 학습 모델에서 미리 학습된 파라미터를 전이학습 방식을 통해 전달받아 그 분류 능력을 더욱 끌어올렸다.

[0149] 이후, 본격적인 실험을 위해 6가지 환경에서 얼굴 이미지와 행동 이미지를 수집하였다. 빛을 등진 Front, 빛이

왼쪽에서 들어오는 Left, 안경을 쓴 Glasses, 모자를 쓴 Hat, 모자와 안경을 모드 쓴 Hat and Glasses, 마스크를 쓴 Mask가 6가지 환경이다.

[0150] 그리고 점진적 딥러닝 학습 방법을 통해 학습되는 도중에 지속적으로 그 성능을 측정하여 평균을 구했다. 그 과정에서 성능이 수렴할 때마다 실험 환경을 바꿔가며 새로운 환경에 대한 이전 학습 모델의 성능 변화를 계속하여 측정하였다.

[0152] 실험 결과

[0153] 실험 시 6가지 환경에 대해 어떤 순서로 학습을 시도할지 결정할 필요가 있다. 본 실험을 원활하게 진행하기 위해 먼저 초기 학습 모델을 사용하여 각 환경에서의 초기 성능을 측정하였다. 그리고 측정된 초기 성능을 오름차순으로 정렬하여 가장 초기 성능이 낮은 환경부터 학습을 시작하였다. 표 1은 6가지 환경에서 작성된 실험 데이터의 초기 성능과 그에 따른 학습 순서를 나타낸 것이다.

표 1

[0154]

구별	환경1	환경2	환경3	환경4	환경5	환경6
초기 성능	51.4	85.7	34.2	94.2	65.7	31.4
순서	3	5	2	6	4	1

[0155] 물론 이러한 방식으로 학습 순서를 정한 것에는 이유가 있다. 초기 성능이 높은 환경부터 점진적 딥러닝 학습을 시작하였을 경우, 준지도 학습 과정에서 라벨이 없는 학습 데이터에 높은 신뢰도로 잘못된 라벨을 부여하기 때문이다. 따라서 이러한 과정을 최소화하기 위해 초기 학습 모델을 구성하는데 쓰인 데이터와 가장 유사하지 않은 환경에서부터 점진적으로 학습을 시도하였다.

[0156] 도 7을 보면, 실험 도중 준지도 학습을 거치지 않고 바로 능동 학습으로 들어가는 부분이 있다. 이는 위와 같은 방법을 사용하였음에도 불구하고 높은 신뢰도를 가진 잘못된 라벨이 발생하였기 때문이다. 실험 결과 이러한 문제가 발생하였을 경우, 바로 이전에 학습한 모델보다 그 성능이 떨어지는 것을 발견하였고, 항상 두 학습 모델의 성능을 비교하여 성능이 감소하였을 경우의 학습 데이터는 파기하였다.

[0157] 실험 결과, 환경 변화 시 초기 성능의 감소폭이 점차 작아지면서 수렴하는 것을 알 수 있었다. 점진적 딥러닝 학습을 통해 자동적으로 다양한 환경의 데이터를 학습하였기 때문이다. 도 8은 환경 변화에 따른 초기 성능의 변화치를 나타낸 것이다.

[0158] 또한, 본 실험은 수집한 데이터에 모두 사람이 라벨을 붙이는 것보다 훨씬 적은 노동력을 필요로 한다. 도 9를 보면, 일일이 사람이 직접 라벨을 붙이는 것은 점진적 딥러닝 학습 방법에 비해 필요한 노동력이 3.5배나 됨을 확인할 수 있다.

[0159] 위와 같은 실험 결과로 볼 때, 점진적 딥러닝 학습 방법이 지도학습 방법보다 효율적이며 일반적인 딥러닝 학습 방법보다 환경 변화에 강인한 것으로 보인다.

[0160] 본 발명에서는 적은 노동력으로 다양한 환경 변화에 적응할 수 있는 학습 모델을 만들 수 있도록 점진적 딥러닝 학습 기법을 제안하였다. 제안한 실험을 실시하는 도중, 높은 신뢰도를 갖는 잘못된 라벨이 나타나는 문제가 발생하여 준지도 학습 결과가 좋지 않아 성능 향상에 문제를 일으키는 일이 있었지만, 초기 성능을 비교하여 실험 순서를 조절하고 이전 모델과의 성능 비교를 통해 성능이 낮아지는 경우의 데이터를 실험에서 배제하는 방법으로 이러한 문제를 해결할 수 있었다. 본 발명에서 얼굴 표정 및 행위를 인식하기 위하여 사용되는 점진적 딥러닝 학습 모델은, 다양한 환경에서 수많은 학습 데이터를 필요로 하지만 그 데이터의 확보에 어려움을 요하는 경우에 큰 도움을 줄 것이다.

[0162] 얼굴 표정 인식

[0163] 얼굴 표정 인식 인식 과정은 세 가지 단계로 나뉘는데, 전처리, 특징 추출, 결과 분류가 그것이다. 먼저, 전처리 과정에서 조명이나 배경 등의 노이즈를 제거하고, 특징 추출에서 얼굴 표정을 대표할 수 있는 특징 벡터를

뽑아내며, 결과 분류에서 추출된 특징들을 비교하여 가장 비슷한 결과를 찾아낸다.

[0164] 얼굴 표정 인식은 보통 미국의 심리학자 폴 에크만이 주장한 기본 6정서에 무표정을 더하여 7가지 라벨(분노, 역겨움, 공포, 행복, 슬픔, 놀람, 무표정)로 나타낸다. 본 발명에서는, 점진적 딥러닝 학습 모델에 기반하여 얼굴 표정을 인식한다.

[0166] 행동 인식

[0167] 행동 인식은 비디오 감시 시스템, 인간 행동 분석, 인간과 기계 간의 커뮤니케이션 등에 특히 중요한 분야로 지난 이십여 년 동안 꾸준히 연구되어 왔다. 주로 HOG(Histogram oriented gradient), STIP (Space time interest point), DPM (Deformable part model) 등의 기술을 사용해오다가 최근에는 딥러닝을 기반으로 하는 행동 인식 알고리즘이 개발되어 큰 성능 향상을 보였다. 본 발명에서는, 점진적 딥러닝 학습 모델에 기반하여 행동을 인식한다.

[0169] 얼굴 표정 및 행동 인식에 대한 실험 및 결과

[0170] 본 발명의 일 실시예에서는 VGG-16 네트워크를 사용하여 각각 행동 인식과 얼굴 표정 인식 모델을 학습하였다. 그 후, 종합적인 감성을 파악하기 위해 분노, 역겨움, 공포, 행복, 슬픔, 놀람, 무표정의 일곱 가지 감정 라벨과 전화 중, 책 읽는 중, 컴퓨터 사용 중, 사진 찍는 중, 점프 중의 다섯 가지 행동 라벨을 갖는 실험 데이터를 준비하여, 적응적 딥러닝 학습 기법을 통해 학습하였다.

[0171] 도 10과 같이 행동 인식의 경우, 뒤의 배경이 단순하느냐 복잡하느냐에 따라 그 인식률에 차이가 컸다. 영상에서 사람의 모습만 정확히 추출해낼 수 없었기 때문에 배경이 강력한 노이즈가 되어 인식을 방해했기 때문이다. 하지만, 적응적 딥러닝 학습 기법을 통해 80% 근처까지 성능을 향상시킬 수 있었다.

[0172] 도 11은 얼굴 표정 인식에 대한 그래프인데, 이 경우 얼굴선 안에서 이미지가 추출되므로 배경이 전혀 들어가지 않는다. 이를 통해 본 발명의 일 실시예에서는 얼굴 표정 인식률을 100%에 가깝게 끌어올릴 수 있었다.

[0173] 이상 본 발명을 구체적인 실시예를 통하여 상세하게 설명하였으나, 이는 본 발명을 구체적으로 설명하기 위한 것으로, 본 발명은 이에 한정되지 않으며, 본 발명의 기술적 사상 내에서 당 분야의 통상의 지식을 가진 자에 의해 그 변형이나 개량이 가능함은 명백하다고 할 것이다.

[0174] 본 발명의 단순한 변형 내지 변경은 모두 본 발명의 영역에 속하는 것으로, 본 발명의 구체적인 보호 범위는 첨부된 청구범위에 의하여 명확해질 것이다.

부호의 설명

[0175] 100 : 전처리부 102 : 얼굴 검출부

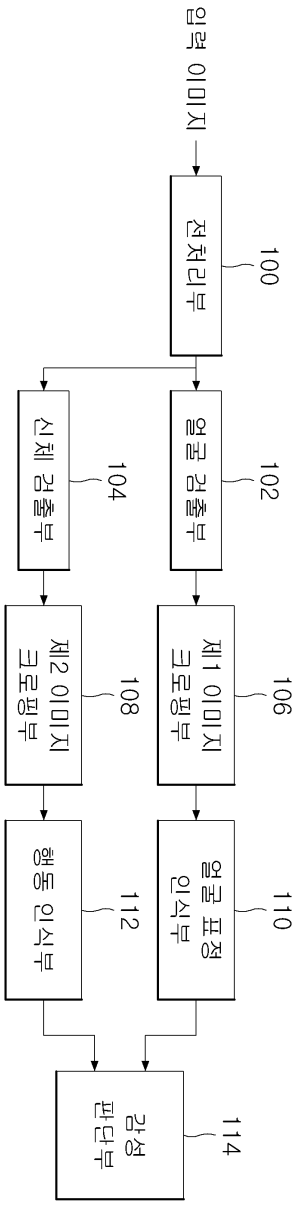
104 : 신체 검출부 106 : 제1 이미지 크로핑부

108 : 제2 이미지 크로핑부 110 : 얼굴 표정 인식부

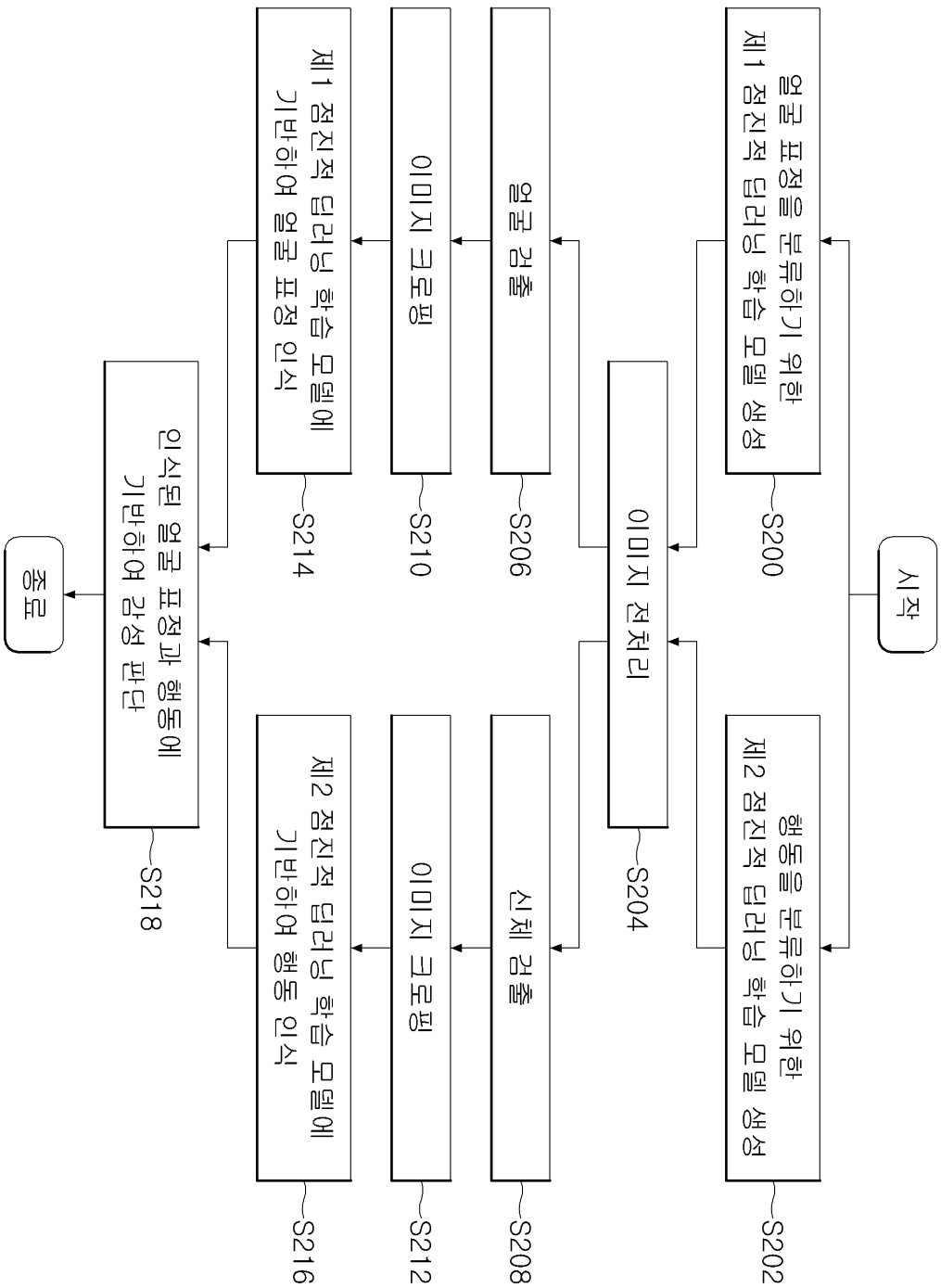
112 : 행동 인식부 114 : 감성 판단부

도면

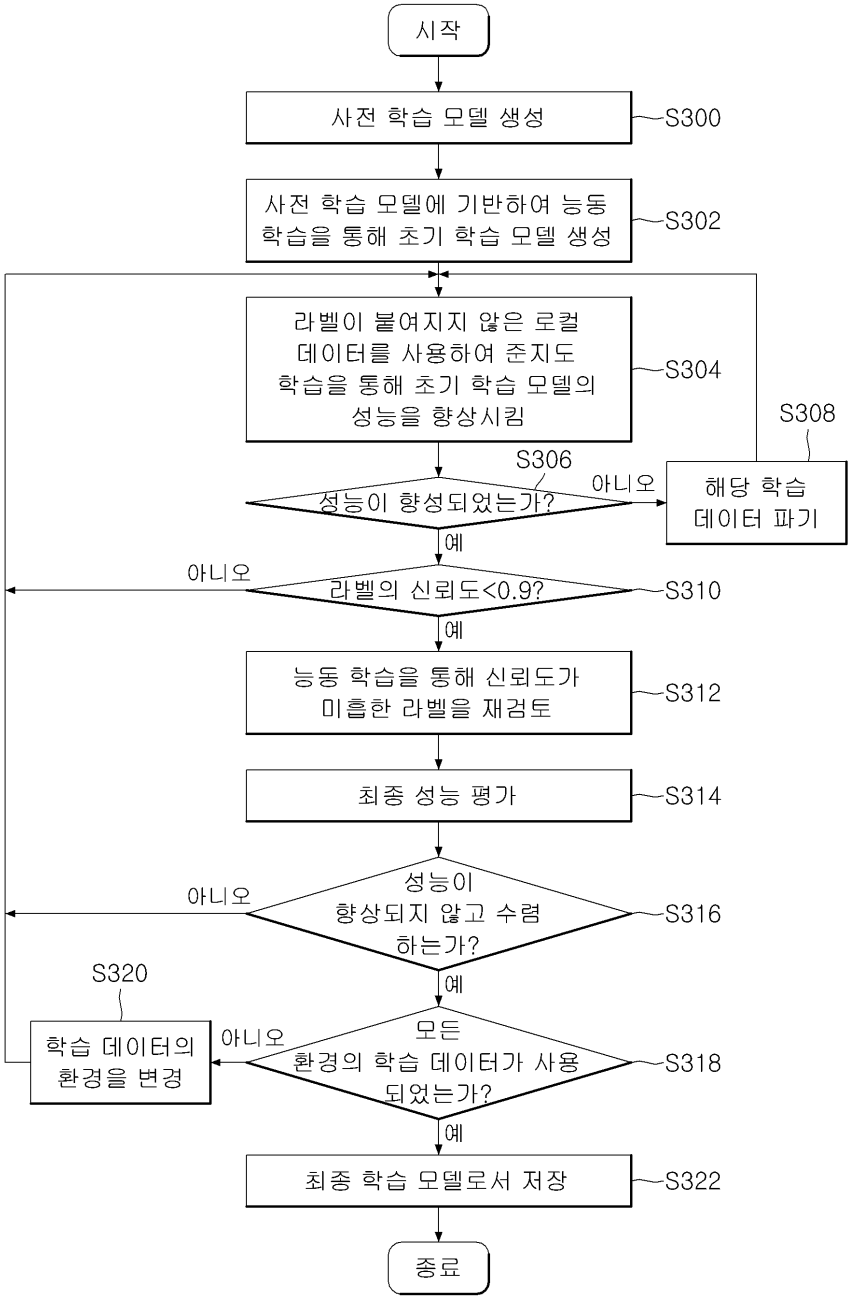
도면1


입력 이미지
전처리부
100
얼굴 검출부
102
신체 검출부
104
제1 이미지 크로핑부
106
제2 이미지 크로핑부
108
얼굴 표정 인식부
110
행동 인식부
112
감성 판단부
114

도면2


시작
얼굴 표정을 분류하기 위한 제1 점진적 딥러닝 학습 모델 생성
S200
행동을 분류하기 위한 제2 점진적 딥러닝 학습 모델 생성
S202
이미지 전처리
S204
얼굴 검출
S206
신체 검출
S208
이미지 크로핑
S210
이미지 크로핑
S212
제1 점진적 딥러닝 학습 모델에 기반하여 얼굴 표정 인식
S214
제2 점진적 딥러닝 학습 모델에 기반하여 행동 인식
S216
인식된 얼굴 표정과 행동에 기반하여 감성 판단
S218
종료

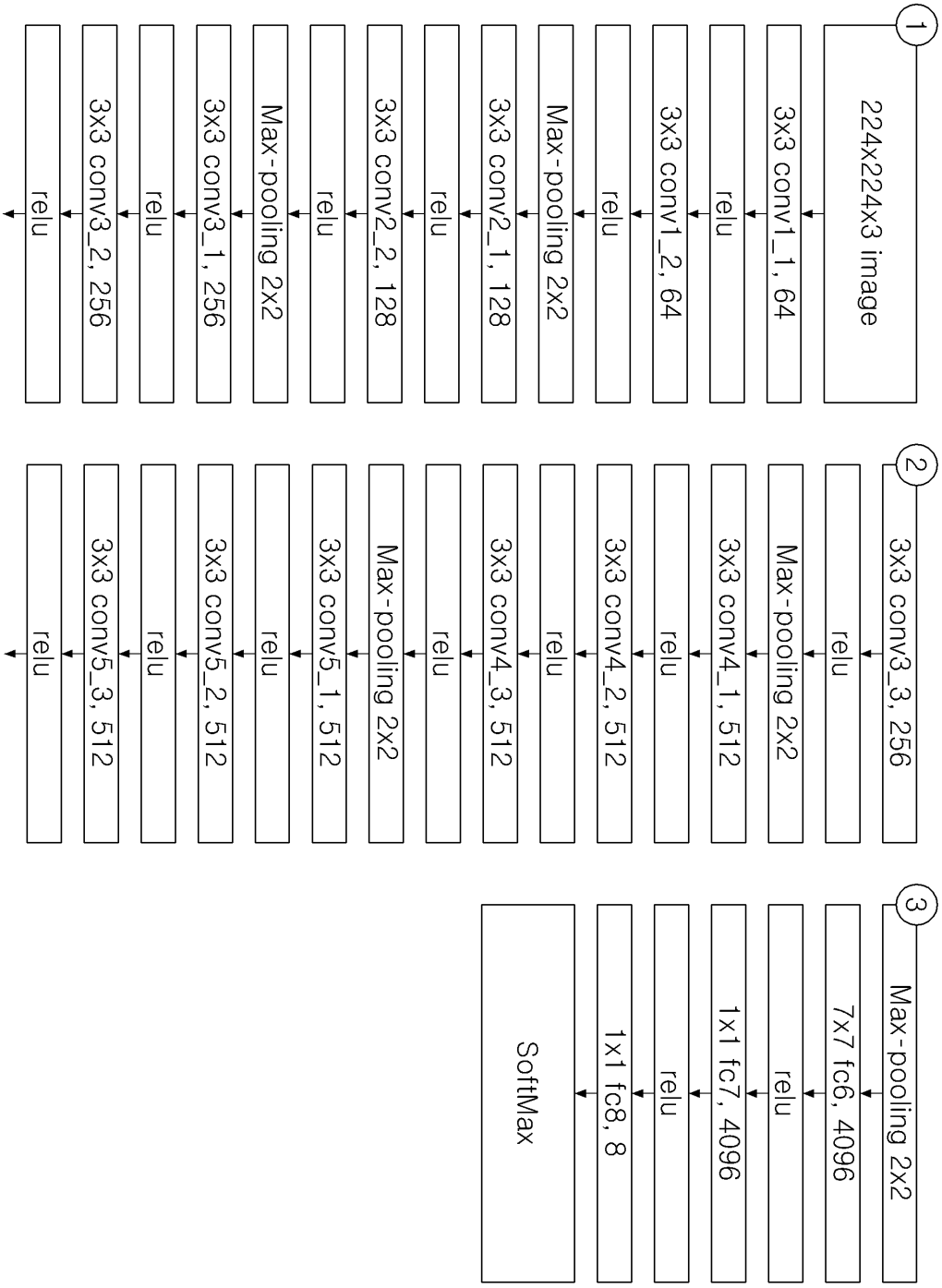
도면3


시작
사전 학습 모델 생성
S300
사전 학습 모델에 기반하여 능동 학습을 통해 초기 학습 모델 생성
S302
라벨이 붙여지지 않은 로컬 데이터를 사용하여 준지도 학습을 통해 초기 학습 모델의 성능을 향상시킴
S304
S306
성능이 향성되었는가?
아니오
S308
해당 학습 데이터 파기
예
아니오
라벨의 신뢰도<0.9?
S310
예
능동 학습을 통해 신뢰도가 미흡한 라벨을 재검토
S312
최종 성능 평가
S314
성능이 향상되지 않고 수렴하는가?
아니오
S316
예
S320
학습 데이터의 환경을 변경
아니오
모든 환경의 학습 데이터가 사용되었는가?
S318
예
최종 학습 모델로서 저장
S322
종료

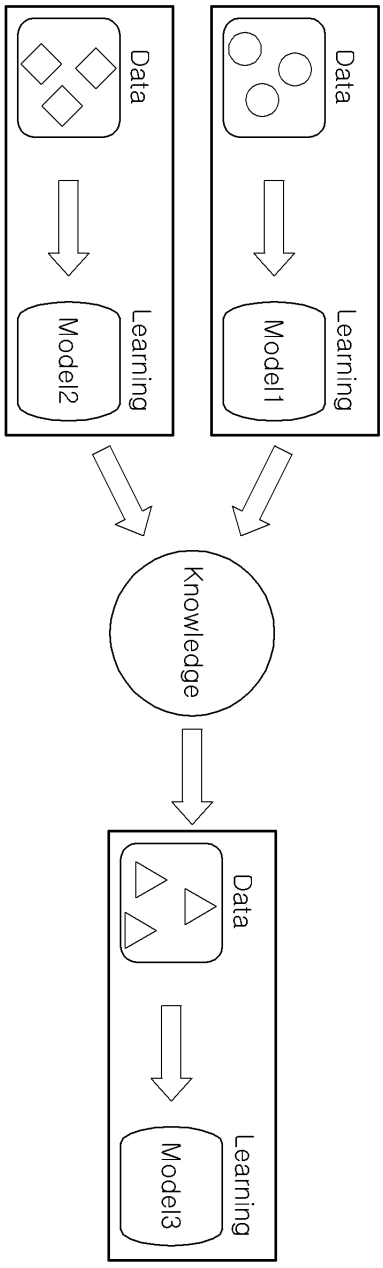
도면4



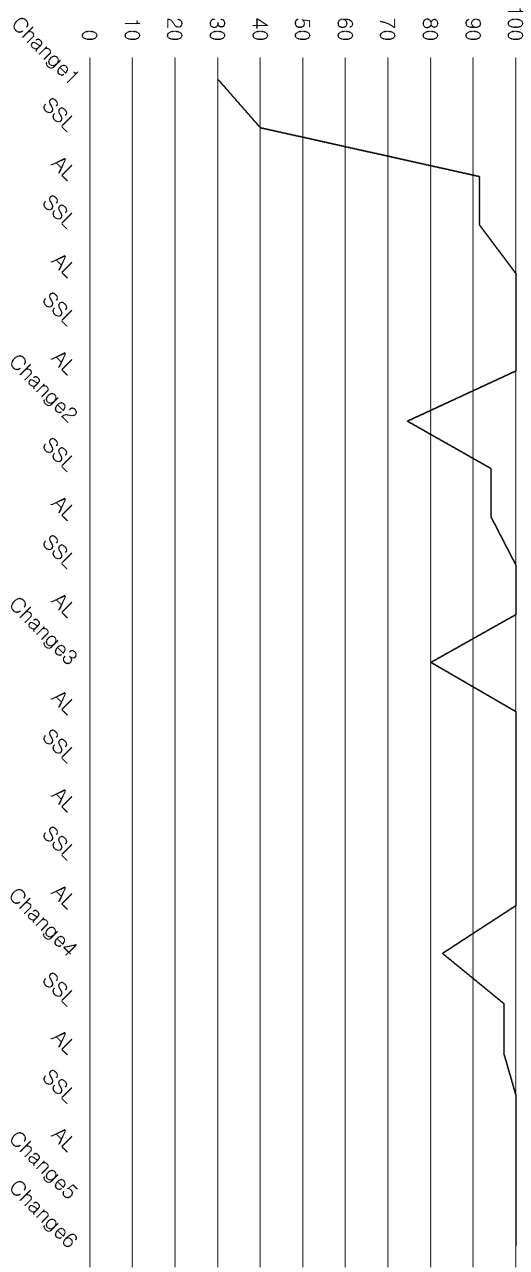
도면5


1
224x224x3 image
3x3 conv1_1, 64
relu
3x3 conv1_2, 64
relu
Max-pooling 2x2
3x3 conv2_1, 128
relu
3x3 conv2_2, 128
relu
Max-pooling 2x2
3x3 conv3_1, 256
relu
3x3 conv3_2, 256
relu
2
3x3 conv3_3, 256
relu
Max-pooling 2x2
3x3 conv4_1, 512
relu
3x3 conv4_2, 512
relu
3x3 conv4_3, 512
relu
Max-pooling 2x2
3x3 conv5_1, 512
relu
3x3 conv5_2, 512
relu
3x3 conv5_3, 512
relu
3
Max-pooling 2x2
7x7 fc6, 4096
relu
1x1 fc7, 4096
relu
1x1 fc8, 8
SoftMax

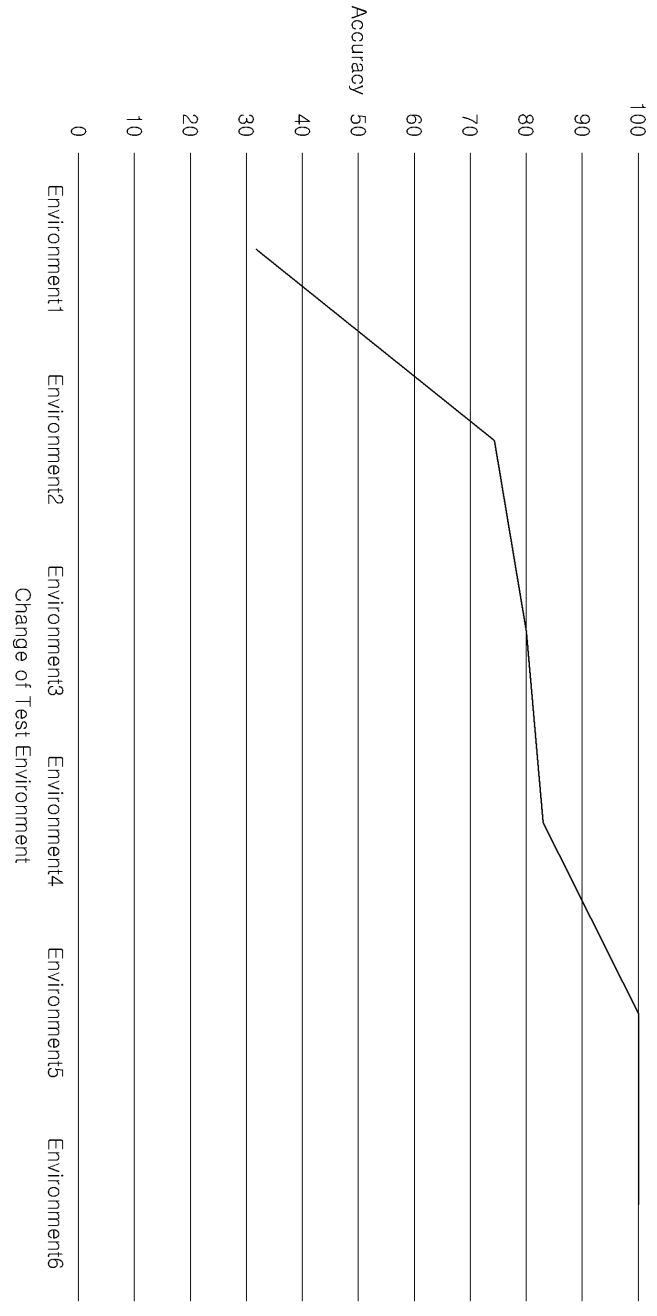
도면6


Data
Learning
Model2
Data
Learning
Model1
Knowledge
Data
Learning
Model3

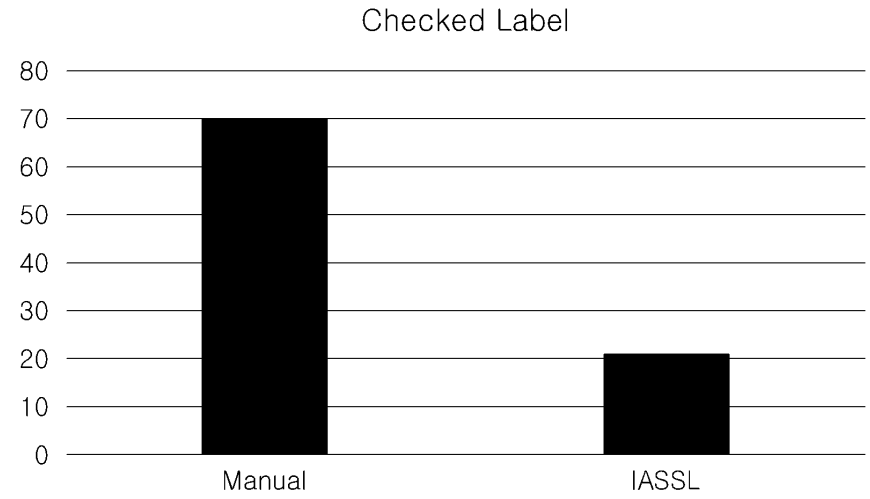
도면7


0
10
20
30
40
50
60
70
80
90
100
Change1
SSL
AL
SSL
AL
SSL
AL
Change2
SSL
AL
SSL
AL
Change3
AL
SSL
AL
SSL
AL
Change4
SSL
AL
SSL
AL
Change5
Change6

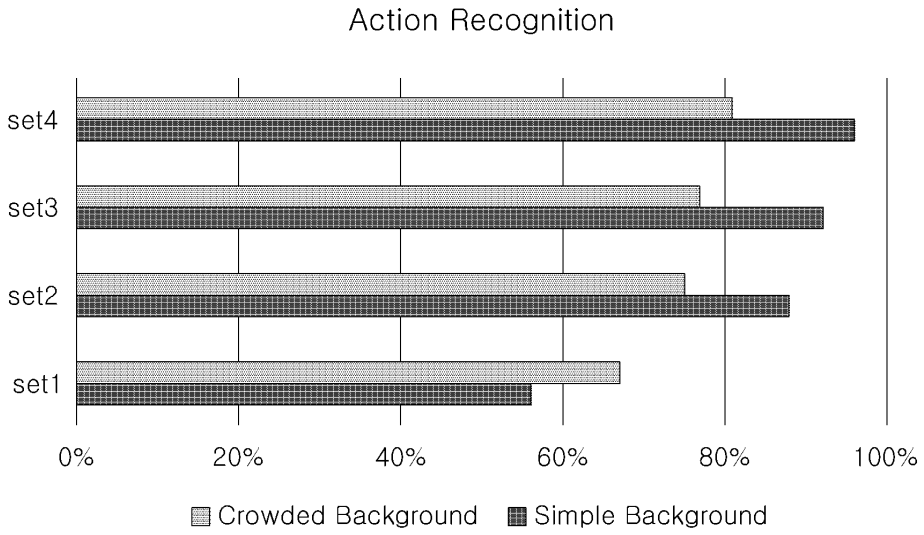
도면8


Accuracy
100
90
80
70
60
50
40
30
20
10
0
Environment1
Environment2
Environment3
Environment4
Environment5
Environment6
Change of Test Environment

도면9



도면10



도면11

