



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0123507
(43) 공개일자 2020년10월30일

(51) 국제특허분류(Int. Cl.)
G06K 9/00 (2006.01) G06N 3/08 (2006.01)
G06T 7/20 (2017.01)
(52) CPC특허분류
G06K 9/00362 (2013.01)
G06N 3/08 (2013.01)
(21) 출원번호 10-2019-0043970
(22) 출원일자 2019년04월15일
심사청구일자 2019년04월15일

(71) 출원인
계명대학교 산학협력단
대구광역시 달서구 달구벌대로 1095, 계명대학교 산학협력관 201호(신당동)
(72) 발명자
고병철
대구광역시 수성구 범어천로 107 (범어동, 범어 STX KAN) 103동 701호
남재열
대구광역시 수성구 수성로 412 수성보성타운 107동 503호
허두영
대구광역시 달서구 상화북로 196 제림아파트 101동 1302호
(74) 대리인
김건우

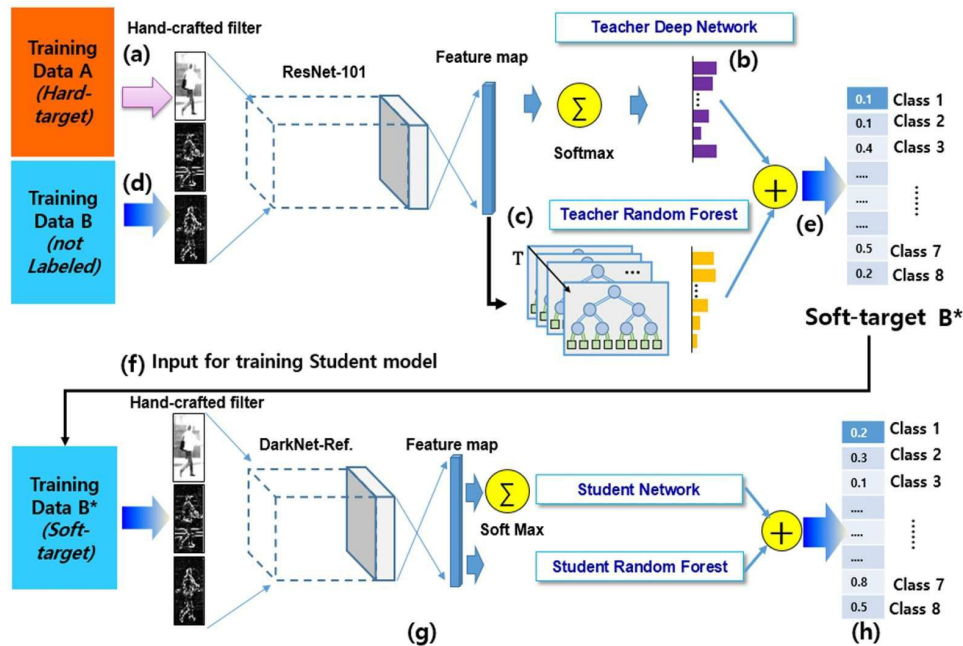
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템

(57) 요약

본 발명은 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법에 관한 것으로, 보다 구체적으로는 보행자 포즈 방향 추정 방법으로서, (1) 데이터 세트 A를 이용하여, 교사 심층 네트워크 및 교사 랜덤 포레스트로 구성되는 교사 모델을 학습시키는 단계; (2) 데이터 세트 B를 상기 단계 (1)에서 (뒷면에 계속)

대표도 - 도1



학습된 교사 모델에 입력하고, 출력된 소프트 출력을 이용하여 소프트 타겟 데이터 세트 B^* 를 생성하는 단계;
(3) 상기 단계 (2)에서 생성된 데이터 세트 B^* 를 이용하여, 학생 네트워크 및 학생 랜덤 포레스트로 구성되는 학생 모델을 학습시키는 단계; 및 (4) 상기 단계 (3)에서 학습된 학생 모델로 보행자 포즈 방향을 추정하는 단계를 포함하며, 상기 단계 (4)에서는, 최대 확률 클래스의 확률 및 상기 최대 확률 클래스와 인접한 클래스의 확률의 조합을 이용한 소프트 가중 방법을 이용해 상기 보행자 방향을 추정하는 것을 그 구성상의 특징으로 한다.

또한, 본 발명은 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템에 관한 것으로서, 보다 구체적으로는 보행자 포즈 방향 추정 시스템으로서, 데이터 세트 A를 이용하여, 교사 심층 네트워크 및 교사 랜덤 포레스트로 구성되는 교사 모델을 학습시키는 교사 학습 모듈; 데이터 세트 B를 상기 교사 학습 모듈에서 학습된 교사 모델에 입력하고, 출력된 소프트 출력을 이용하여 소프트 타겟 데이터 세트 B^* 를 생성하는 소프트 타겟 데이터 생성 모듈; 상기 소프트 타겟 데이터 생성 모듈에서 생성된 데이터 세트 B^* 를 이용하여, 학생 네트워크 및 학생 랜덤 포레스트로 구성되는 학생 모델을 학습시키는 학생 학습 모듈; 및 상기 학생 학습 모듈에서 학습된 학생 모델로 보행자 포즈 방향을 추정하는 방향 추정 모듈을 포함하며, 상기 방향 추정 모듈에서는, 최대 확률 클래스의 확률 및 상기 최대 확률 클래스와 인접한 클래스의 확률의 조합을 이용한 소프트 가중 방법을 이용해 상기 보행자 방향을 추정하는 것을 또 다른 구성상의 특징으로 한다.

본 발명에서 제안하고 있는 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템에 따르면, 교사 모델의 출력인 소프트 타겟 데이터 세트 B^* 를 입력으로 하여 학생 모델을 학습시킴으로써, 교사-학생 프레임워크를 통해 모델을 경량화하면서도 더 유연한 분류 결과를 출력하도록 할 수 있고, 클래스 레이블이 포함되는 데이터 세트 A와 클래스 레이블이 포함되지 않는 데이터 세트 B를 이용해 교사 모델의 오버 피팅(overfitting)을 방지할 수 있다.

또한, 본 발명에서 제안하고 있는 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템에 따르면, 최종 클래스 확률 및 인접한 클래스 확률의 조합을 이용한 소프트 가중 방법을 이용해 보행자 방향을 추정함으로써, 미리 구분된 클래스의 방향 중 어느 하나가 아닌, 실제 보행자의 방향에 가까운 값으로 보행자 포즈 방향을 추정할 수 있다.

(52) CPC특허분류

G06T 7/20 (2013.01)

G06T 2207/20081 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	B20160372
부처명	대구광역시
과제관리(전문)기관명	자동차부품연구원
연구사업명	미래형자동차 선도기술 개발사업
연구과제명	전후방 ADAS 센서를 연계한 보행자 보호 및 오토 트렁크 시스템 개발
기 여 율	1/1
과제수행기관명	평화정공(주)
연구기간	2018.06.01 ~ 2019.05.31

명세서

청구범위

청구항 1

보행자 포즈 방향 추정 방법으로서,

- (1) 데이터 세트 A를 이용하여, 교사 심층 네트워크 및 교사 랜덤 포레스트로 구성되는 교사 모델을 학습시키는 단계;
 - (2) 데이터 세트 B를 상기 단계 (1)에서 학습된 교사 모델에 입력하고, 출력된 소프트 출력을 이용하여 소프트 타겟 데이터 세트 B^* 를 생성하는 단계;
 - (3) 상기 단계 (2)에서 생성된 데이터 세트 B^* 를 이용하여, 학생 네트워크 및 학생 랜덤 포레스트로 구성되는 학생 모델을 학습시키는 단계; 및
 - (4) 상기 단계 (3)에서 학습된 학생 모델로 보행자 포즈 방향을 추정하는 단계를 포함하며,
- 상기 단계 (4)에서는, 최대 확률 클래스의 확률 및 상기 최대 확률 클래스와 인접한 클래스의 확률의 조합을 이용한 소프트 가중 방법을 이용해 상기 보행자 방향을 추정하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법.

청구항 2

제1항에 있어서, 상기 단계 (1)에서,

상기 데이터 세트 A는 클래스 레이블이 포함되는 하드 타겟 데이터 세트인 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법.

청구항 3

제1항에 있어서, 상기 단계 (2)에서,

상기 데이터 세트 B는 클래스 레이블이 포함되지 않는 데이터 세트인 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법.

청구항 4

제1항에 있어서, 상기 단계 (1)은,

- (1-1) 상기 데이터 세트 A를 이용하여 교사 심층 네트워크를 학습시키는 단계; 및
- (1-2) 상기 교사 심층 네트워크의 특징 맵(feature map)을 이용하여 교사 랜덤 포레스트를 학습시키는 단계를 포함하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법.

청구항 5

제1항에 있어서,

- (0) 웨이블릿 변환을 적용하여 입력 이미지에 대한 전처리를 수행하는 단계를 더 포함하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법.

청구항 6

제5항에 있어서,

상기 단계 (0)에서는, 하이 패스(high-pass) 필터된 2개의 서브이미지 및 로우 패스(low-pass) 필터된 하나의 서브이미지를 생성하며,

상기 단계 (1)에서는, 상기 단계 (0)에서 생성된 3개의 서브이미지를 이용하여 상기 교사 모델을 학습시키는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법.

청구항 7

제1항에 있어서, 상기 단계 (2)에서는,

상기 데이터 세트 B를 상기 단계 (1)에서 학습된 교사 심층 네트워크 및 교사 랜덤 포레스트에 입력하고, 상기 교사 심층 네트워크의 출력 및 상기 교사 랜덤 포레스트의 출력을 하나의 소프트 타겟 벡터로 결합하여, 각각의 클래스의 확률값인 소프트 타겟 데이터 세트 B^* 를 생성하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법.

청구항 8

제1항에 있어서, 상기 단계 (4)는,

(4-1) 상기 단계 (3)에서 학습된 학생 네트워크 및 학생 랜덤 포레스트의 출력값을 결합하여 최종 확률을 생성하는 단계;

(4-2) 보행자의 방향 추정을 위해 미리 정해진 각도로 구분된 클래스에서, 가장 높은 확률을 갖는 최대 확률 클래스(c_k) 및 상기 최대 확률 클래스에 인접한 두 클래스(c_{k-1} , c_{k+1})를 특징하는 단계;

(4-3) 상기 특징된 클래스(c_{k-1} , c_k , c_{k+1})의 가중치를 각각 산출하는 단계; 및

(4-4) 상기 단계 (4-3)에서 산출된 각 클래스의 가중치로부터 최종 보행자 방향을 추정하는 단계를 포함하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법.

청구항 9

제8항에 있어서, 상기 단계 (4-3)에서는,

다음 수학적식을 이용해 각 클래스의 가중치($w(c_k)$)를 산출하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법.

$$w(c_k) = 1/(\exp(F \cdot P(c_k)))$$

상기 수학적식에서, $P(c_k)$ 는 k번째 클래스 c_k 의 확률이고, F는 가중치 인자(weighting factor)임.

청구항 10

제8항에 있어서, 상기 단계 (4-4)에서는,

다음 수학적식을 이용해 최종 보행자 방향($\tilde{O}$)을 추정하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법.

$$\tilde{O} = \frac{1}{3} \sum_{k=1}^3 w(c_k) \cdot O_k$$

상기 수학적식에서, O_k 는 클래스 c_k 의 보행자 방향임.

청구항 11

보행자 포즈 방향 추정 시스템으로서,

데이터 세트 A를 이용하여, 교사 심층 네트워크 및 교사 랜덤 포레스트로 구성되는 교사 모델을 학습시키는 교사 학습 모듈(100);

데이터 세트 B를 상기 교사 학습 모듈(100)에서 학습된 교사 모델에 입력하고, 출력된 소프트 출력을 이용하여 소프트 타겟 데이터 세트 B^* 를 생성하는 소프트 타겟 데이터 생성 모듈(200);

상기 소프트 타겟 데이터 생성 모듈(200)에서 생성된 데이터 세트 B^* 를 이용하여, 학생 네트워크 및 학생 랜덤 포레스트로 구성되는 학생 모델을 학습시키는 학생 학습 모듈(300); 및

상기 학생 학습 모듈(300)에서 학습된 학생 모델로 보행자 포즈 방향을 추정하는 방향 추정 모듈(400)을 포함하며,

상기 방향 추정 모듈(400)에서는,

최대 확률 클래스의 확률 및 상기 최대 확률 클래스와 인접한 클래스의 확률의 조합을 이용한 소프트 가중 방법을 이용해 상기 보행자 방향을 추정하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템.

청구항 12

제11항에 있어서, 상기 교사 학습 모듈(100)에서,

상기 데이터 세트 A는 클래스 레이블이 포함되는 하드 타겟 데이터 세트인 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템.

청구항 13

제11항에 있어서, 상기 소프트 타겟 데이터 생성 모듈(200)에서,

상기 데이터 세트 B는 클래스 레이블이 포함되지 않는 데이터 세트인 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템.

청구항 14

제11항에 있어서, 상기 교사 학습 모듈(100)은,

상기 데이터 세트 A를 이용하여 교사 심층 네트워크를 학습시키는 제1 교사 학습부(110); 및

상기 교사 심층 네트워크의 특징 맵(feature map)을 이용하여 교사 랜덤 포레스트를 학습시키는 제2 교사 학습부(120)를 포함하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템.

청구항 15

제11항에 있어서,

웨이블렛 변환을 적용하여 입력 이미지에 대한 전처리를 수행하는 전처리 모듈(500)을 더 포함하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템.

청구항 16

제15항에 있어서,

상기 전처리 모듈(500)에서는, 하이 패스(high-pass) 필터된 2개의 서브이미지 및 로우 패스(low-pass) 필터된 하나의 서브이미지를 생성하며,

상기 교사 학습 모듈(100)에서는, 상기 전처리 모듈(500)에서 생성된 3개의 서브이미지를 이용하여 상기 교사 모델을 학습시키는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템.

청구항 17

제11항에 있어서, 상기 소프트 타겟 데이터 생성 모듈(200)에서는,

상기 데이터 세트 B를 상기 교사 학습 모듈(100)에서 학습된 교사 심층 네트워크 및 교사 랜덤 포레스트에 입력하고, 상기 교사 심층 네트워크의 출력 및 상기 교사 랜덤 포레스트의 출력을 하나의 소프트 타겟 벡터로 결합하여, 각각의 클래스의 확률값인 소프트 타겟 데이터 세트 B^* 를 생성하는 것을 특징으로 하는, 교사-학생 프레임

위크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템.

청구항 18

제11항에 있어서, 상기 방향 추정 모듈(400)은,

상기 학생 학습 모듈(300)에서 학습된 학생 네트워크 및 학생 랜덤 포레스트의 출력값을 결합하여 최종 확률을 생성하는 최종 확률 생성부(410);

보행자의 방향 추정을 위해 미리 정해진 각도로 구분된 클래스에서, 가장 높은 확률을 갖는 최대 확률 클래스(c_k) 및 상기 최대 확률 클래스에 인접한 두 클래스(c_{k-1} , c_{k+1})를 특징하는 클래스 검색부(420);

상기 특정된 클래스(c_{k-1} , c_k , c_{k+1})의 가중치를 각각 산출하는 가중치 산출부(430); 및

상기 가중치 산출부(430)에서 산출된 각 클래스의 가중치로부터 최종 보행자 방향을 추정하는 최종 방향 추정부(440)를 포함하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템.

청구항 19

제18항에 있어서, 상기 가중치 산출부(430)에서는,

다음 수학적식을 이용해 각 클래스의 가중치($w(c_k)$)를 산출하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템.

$$w(c_k) = 1/(\exp(F \cdot P(c_k)))$$

상기 수학적식에서, $P(c_k)$ 는 k 번째 클래스 c_k 의 확률이고, F 는 가중치 인자(weighting factor)임.

청구항 20

제18항에 있어서, 상기 최종 방향 추정부(440)에서는,

다음 수학적식을 이용해 최종 보행자 방향($\tilde{O}$)을 추정하는 것을 특징으로 하는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템.

$$\tilde{O} = \frac{1}{3} \sum_{k=1}^K w(c_k) \cdot O_k$$

상기 수학적식에서, O_k 는 클래스 c_k 의 보행자 방향임.

발명의 설명

기술 분야

[0001] 본 발명은 보행자 포즈 방향 추정 방법 및 시스템에 관한 것으로서, 보다 구체적으로는 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템에 관한 것이다.

배경 기술

[0002] 컴퓨터 비전을 기반으로 한 인간의 자세 추정(Pose Orientation Estimation; POE)은 인간의 의도와 행동에 대한 중요한 시각적 신호가 되기 때문에, 인간-로봇 상호작용, 비디오 감시, 자율 주행을 포함하여 많은 잠재적인 응용 분야를 가지고 있다. 예를 들어, 로봇은 사용자를 관찰하고 보다 자연스러운 상호작용을 달성하기 위해 스스로 움직이고 방향을 잡을 수 있으며, 감시 카메라는 사람들의 활동과 상호작용을 보다 정확하게 묘사할 수 있다. 자율 주행의 측면에서 볼 때는, 보행자의 보행 방향은 보행자가 다음에 무엇을 할 것인지에 대한 좋은 지표가 될 수 있다.

[0003] POE에는 2가지 유형의 접근 방식이 있다. 센서 기반 접근법은 환경, 조명 변화, 컴퓨터 변화 등과 같은 컴퓨터

비전의 근본적인 문제를 극복하는 데 도움이 되며, 스마트폰의 자이로스코프 센서 또는 키넥트 센서를 이용할 수 있다. 그러나 이러한 접근법은 스마트폰을 통해서만 수행 가능하며, 보행자와 센서 사이의 거리나 너무 멀면 인식률이 떨어질 수 있다.

[0004] 카메라 기반 POE 접근법을 사용하면, 더 먼 거리에 있는 보행자의 방향을 인식하고 추가 동작 정보(얼굴, 얼굴 표정, 자세 등)를 추정할 수 있다. 특정 접근법에서, 포즈 방향은 추적에 기반한 궤도 정보를 사용하여 예측될 수 있다. 이러한 접근법은 감시 카메라와 같이 카메라가 정적일 때 보행자가 의도한 움직임을 추론하는 데 효과적이다. 반면에, 로봇이나 지능형 차량과 같이 움직이는 카메라 환경에서는 일정 시간 동안 보행자를 올바르게 추적하고 보행자의 방향을 분석하기가 어려우므로 사용이 부적합하다. 그러므로 다른 POE 방법은 움직이는 카메라 환경에서 단일 프레임 접근법을 제안하였다. 단일 프레임 기반 POE 방법을 사용하면 카메라가 움직이는 지에 관계없이 보행자의 방향을 여러 프레임을 보지 않고도 복구할 수 있다.

[0005] 한편, 컨볼루션 뉴럴 네트워크(Convolutional Neural Network; CNN) 기반의 방향 추정은 학습 및 테스트를 위해 많은 수의 데이터 세트를 필요로 한다. 또한, 연산량이 많으므로 종래의 분류기와 비교하여 대규모의 고수준 컴퓨팅 장치를 필요로 하는 한계가 있다.

[0006] 따라서 차량 탑재 이동 카메라에서 포착된 단일 프레임 영상을 이용해, 보다 효율적으로 보행자 방향을 추정하기 위한 기술의 개발이 필요하다.

[0007] 한편, 본 발명과 관련된 선행기술로서, 등록특허 제10-1588648호(발명의 명칭: 지능형 영상 감시를 위한 보행자 검출 및 추적 방법) 등이 개시된 바 있다.

발명의 내용

해결하려는 과제

[0008] 본 발명은 기존에 제안된 방법들의 상기와 같은 문제점들을 해결하기 위해 제안된 것으로서, 교사 모델의 출력인 소프트 타겟 데이터 세트 B^* 를 입력으로 하여 학생 모델을 학습시킴으로써, 교사-학생 프레임워크를 통해 모델을 경량화하면서도 더 유연한 분류 결과를 출력하도록 할 수 있고, 클래스 레이블이 포함되는 데이터 세트 A와 클래스 레이블이 포함되지 않는 데이터 세트 B를 이용해 교사 모델의 오버 피팅(overfitting)을 방지할 수 있는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템을 제공하는 것을 그 목적으로 한다.

[0009] 또한, 본 발명은, 최종 클래스 확률 및 인접한 클래스 확률의 조합을 이용한 소프트 가중 방법을 이용해 보행자 방향을 추정함으로써, 미리 구분된 클래스의 방향 중 어느 하나가 아닌, 실제 보행자의 방향에 가까운 값으로 보행자 포즈 방향을 추정할 수 있는, 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템을 제공하는 것을 또 다른 목적으로 한다.

과제의 해결 수단

[0010] 상기한 목적을 달성하기 위한 본 발명의 특징에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법은,

[0011] 보행자 포즈 방향 추정 방법으로서,

[0012] (1) 데이터 세트 A를 이용하여, 교사 심층 네트워크 및 교사 랜덤 포레스트로 구성되는 교사 모델을 학습시키는 단계;

[0013] (2) 데이터 세트 B를 상기 단계 (1)에서 학습된 교사 모델에 입력하고, 출력된 소프트 출력을 이용하여 소프트

타겟 데이터 세트 B^* 를 생성하는 단계;

- [0014] (3) 상기 단계 (2)에서 생성된 데이터 세트 B^* 를 이용하여, 학생 네트워크 및 학생 랜덤 포레스트로 구성되는 학생 모델을 학습시키는 단계; 및
- [0015] (4) 상기 단계 (3)에서 학습된 학생 모델로 보행자 포즈 방향을 추정하는 단계를 포함하며,
- [0016] 상기 단계 (4)에서는, 최대 확률 클래스의 확률 및 상기 최대 확률 클래스와 인접한 클래스의 확률의 조합을 이용한 소프트 가중 방법을 이용해 상기 보행자 방향을 추정하는 것을 그 구성상의 특징으로 한다.
- [0017] 바람직하게는, 상기 단계 (1)에서,
- [0018] 상기 데이터 세트 A는 클래스 레이블이 포함되는 하드 타겟 데이터 세트일 수 있다.
- [0019] 바람직하게는 상기 단계 (2)에서,
- [0020] 상기 데이터 세트 B는 클래스 레이블이 포함되지 않는 데이터 세트일 수 있다.
- [0021] 바람직하게는, 상기 단계 (1)은,
- [0022] (1-1) 상기 데이터 세트 A를 이용하여 교사 심층 네트워크를 학습시키는 단계; 및
- [0023] (1-2) 상기 교사 심층 네트워크의 특징 맵(feature map)을 이용하여 교사 랜덤 포레스트를 학습시키는 단계를 포함할 수 있다.
- [0024] 바람직하게는,
- [0025] (0) 웨이블릿 변환을 적용하여 입력 이미지에 대한 전처리를 수행하는 단계를 더 포함할 수 있다.
- [0026] 더욱 바람직하게는,
- [0027] 상기 단계 (0)에서는, 하이 패스(high-pass) 필터된 2개의 서브이미지 및 로우 패스(low-pass) 필터된 하나의 서브이미지를 생성하며,
- [0028] 상기 단계 (1)에서는, 상기 단계 (0)에서 생성된 3개의 서브이미지를 이용하여 상기 교사 모델을 학습시킬 수 있다.
- [0029] 바람직하게는, 상기 단계 (2)에서는,
- [0030] 상기 데이터 세트 B를 상기 단계 (1)에서 학습된 교사 심층 네트워크 및 교사 랜덤 포레스트에 입력하고, 상기 교사 심층 네트워크의 출력 및 상기 교사 랜덤 포레스트의 출력을 하나의 소프트 타겟 벡터로 결합하여, 각각의 클래스의 확률값인 소프트 타겟 데이터 세트 B^* 를 생성할 수 있다.
- [0031] 바람직하게는, 상기 단계 (4)는,
- [0032] (4-1) 상기 단계 (3)에서 학습된 학생 네트워크 및 학생 랜덤 포레스트의 출력값을 결합하여 최종 확률을 생성하는 단계;
- [0033] (4-2) 보행자의 방향 추정을 위해 미리 정해진 각도로 구분된 클래스에서, 가장 높은 확률을 갖는 최대 확률 클래스(c_k) 및 상기 최대 확률 클래스에 인접한 두 클래스(c_{k-1} , c_{k+1})를 특징하는 단계;

- [0034] (4-3) 상기 특정된 클래스(c_{k-1} , c_k , c_{k+1})의 가중치를 각각 산출하는 단계; 및
- [0035] (4-4) 상기 단계 (4-3)에서 산출된 각 클래스의 가중치로부터 최종 보행자 방향을 추정하는 단계를 포함할 수 있다.
- [0036] 더욱 바람직하게는, 상기 단계 (4-3)에서는,
- [0037] 다음 수학적식을 이용해 각 클래스의 가중치($w(c_k)$)를 산출할 수 있다.
- [0038] $w(c_k) = 1/(\exp(F \cdot P(c_k)))$
- [0039] 상기 수학적식에서, $P(c_k)$ 는 k 번째 클래스 c_k 의 확률이고, F 는 가중치 인자(weighting factor)임.
- [0040] 더욱 바람직하게는, 상기 단계 (4-4)에서는,
- [0041] 다음 수학적식을 이용해 최종 보행자 방향($\tilde{O}$)을 추정할 수 있다.
- [0042] $\tilde{O} = \frac{1}{3} \sum_{k=1}^{k=1} w(c_k) \cdot O_k$
- [0043] 상기 수학적식에서, O_k 는 클래스 c_k 의 보행자 방향임.
- [0044] 상기한 목적을 달성하기 위한 본 발명의 특징에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템은,
- [0045] 보행자 포즈 방향 추정 시스템으로서,
- [0046] 데이터 세트 A를 이용하여, 교사 심층 네트워크 및 교사 랜덤 포레스트로 구성되는 교사 모델을 학습시키는 교사 학습 모듈;
- [0047] 데이터 세트 B를 상기 교사 학습 모듈에서 학습된 교사 모델에 입력하고, 출력된 소프트 출력을 이용하여 소프트 타겟 데이터 세트 B^* 를 생성하는 소프트 타겟 데이터 생성 모듈;
- [0048] 상기 소프트 타겟 데이터 생성 모듈에서 생성된 데이터 세트 B^* 를 이용하여, 학생 네트워크 및 학생 랜덤 포레스트로 구성되는 학생 모델을 학습시키는 학생 학습 모듈; 및
- [0049] 상기 학생 학습 모듈에서 학습된 학생 모델로 보행자 포즈 방향을 추정하는 방향 추정 모듈을 포함하며,
- [0050] 상기 방향 추정 모듈에서는,
- [0051] 최대 확률 클래스의 확률 및 상기 최대 확률 클래스와 인접한 클래스의 확률의 조합을 이용한 소프트 가중 방법을 이용해 상기 보행자 방향을 추정하는 것을 그 구성상의 특징으로 한다.
- [0052] 바람직하게는, 상기 교사 학습 모듈에서,
- [0053] 상기 데이터 세트 A는 클래스 레이블이 포함되는 하드 타겟 데이터 세트일 수 있다.
- [0054] 바람직하게는, 상기 소프트 타겟 데이터 생성 모듈에서,
- [0055] 상기 데이터 세트 B는 클래스 레이블이 포함되지 않는 데이터 세트일 수 있다.
- [0056] 바람직하게는, 상기 교사 학습 모듈은,

- [0057] 상기 데이터 세트 A를 이용하여 교사 심층 네트워크를 학습시키는 제1 교사 학습부; 및
- [0058] 상기 교사 심층 네트워크의 특징 맵(feature map)을 이용하여 교사 랜덤 포레스트를 학습시키는 제2 교사 학습부를 포함할 수 있다.
- [0059] 바람직하게는,
- [0060] 웨이블릿 변환을 적용하여 입력 이미지에 대한 전처리를 수행하는 전처리 모듈을 더 포함할 수 있다.
- [0061] 더욱 바람직하게는,
- [0062] 상기 전처리 모듈에서는, 하이 패스(high-pass) 필터된 2개의 서브이미지 및 로우 패스(low-pass) 필터된 하나의 서브이미지를 생성하며,
- [0063] 상기 교사 학습 모듈에서는, 상기 전처리 모듈에서 생성된 3개의 서브이미지를 이용하여 상기 교사 모델을 학습시킬 수 있다.
- [0064] 바람직하게는, 상기 소프트 타겟 데이터 생성 모듈에서는,
- [0065] 상기 데이터 세트 B를 상기 교사 학습 모듈에서 학습된 교사 심층 네트워크 및 교사 랜덤 포레스트에 입력하고, 상기 교사 심층 네트워크의 출력 및 상기 교사 랜덤 포레스트의 출력을 하나의 소프트 타겟 벡터로 결합하여, 각각의 클래스의 확률값인 소프트 타겟 데이터 세트 B^{*}를 생성할 수 있다.
- [0066] 더욱 바람직하게는, 상기 방향 추정 모듈은,
- [0067] 상기 학생 학습 모듈에서 학습된 학생 네트워크 및 학생 랜덤 포레스트의 출력값을 결합하여 최종 확률을 생성하는 최종 확률 생성부;
- [0068] 보행자의 방향 추정을 위해 미리 정해진 각도로 구분된 클래스에서, 가장 높은 확률을 갖는 최대 확률 클래스(c_k) 및 상기 최대 확률 클래스에 인접한 두 클래스(c_{k-1} , c_{k+1})를 특징하는 클래스 검색부;
- [0069] 상기 특정된 클래스(c_{k-1} , c_k , c_{k+1})의 가중치를 각각 산출하는 가중치 산출부; 및
- [0070] 상기 가중치 산출부에서 산출된 각 클래스의 가중치로부터 최종 보행자 방향을 추정하는 최종 방향 추정부를 포함할 수 있다.
- [0071] 더더욱 바람직하게는, 상기 가중치 산출부에서는,
- [0072] 다음 수학적식을 이용해 각 클래스의 가중치($w(c_k)$)를 산출할 수 있다.
- [0073]
$$w(c_k) = 1/(\exp(F \cdot P(c_k)))$$
- [0074] 상기 수학적식에서, $P(c_k)$ 는 k번째 클래스 c_k 의 확률이고, F는 가중치 인자(weighting factor)임.
- [0075] 더더욱 바람직하게는, 상기 최종 방향 추정부에서는,
- [0076] 다음 수학적식을 이용해 최종 보행자 방향($\vec{0}$)을 추정할 수 있다.
- [0077]
$$\vec{0} = \frac{1}{3} \sum_{k=1}^K w(c_k) \cdot O_k$$
- [0078] 상기 수학적식에서, O_k 는 클래스 c_k 의 보행자 방향임.

발명의 효과

[0079] 본 발명에서 제안하고 있는 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템에 따르면, 교사 모델의 출력인 소프트 타겟 데이터 세트 B^* 를 입력으로 하여 학생 모델을 학습시킴으로써, 교사-학생 프레임워크를 통해 모델을 경량화하면서도 더 유연한 분류 결과를 출력하도록 할 수 있고, 클래스 레이블이 포함되는 데이터 세트 A와 클래스 레이블이 포함되지 않는 데이터 세트 B를 이용해 교사 모델의 오버 피팅(overfitting)을 방지할 수 있다.

[0080] 또한, 본 발명에서 제안하고 있는 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템에 따르면, 최종 클래스 확률 및 인접한 클래스 확률의 조합을 이용한 소프트 가중 방법을 이용해 보행자 방향을 추정함으로써, 미리 구분된 클래스의 방향 중 어느 하나가 아닌, 실제 보행자의 방향에 가까운 값으로 보행자 포즈 방향을 추정할 수 있다.

도면의 간단한 설명

[0081] 도 1은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템을 도시한 도면.

도 2는 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법의 흐름을 도시한 도면.

도 3은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법에서, 단계 S100의 세부적인 흐름을 도시한 도면.

도 4는 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법에서, 학생 네트워크의 학습 절차를 설명한 알고리즘을 도시한 도면.

도 5는 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법에서, 학생 랜덤 포레스트의 학습 절차를 설명한 알고리즘을 도시한 도면.

도 6은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법에서, 보행자 방향 클래스 분류를 예를 들어 도시한 도면.

도 7은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법에서, 단계 S400의 세부적인 흐름을 도시한 도면.

도 8은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템의 구성을 도시한 도면.

도 9는 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법을 포함하는 8개의 실험의 보행자 방향 추정 결과를 비교한 도면.

도 10은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템에 따른, 방향 클래스별 POE 분류 정확도(Acc)를 confusion matrix로 도시한 도면.

도 11은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템에서, 학생 랜덤 포레스트의 트리 수 결정을 위한 실험 결과를 도시한 도면.

도 12는 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법을 포함하는 4개의 실험의 정확도, 파라미터의 수 및 연산 수를 비교한 도면.

도 13은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법을 포함하는 5개의 CNN 기반 방법에 대한 실험 결과를 요약한 도면.

도 14는 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템을 사용하여 (a) TUD 및 (b) KITTI 데이터 세트의 POE 결과를 도시한 도면.

발명을 실시하기 위한 구체적인 내용

- [0082] 이하, 첨부된 도면을 참조하여 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명을 용이하게 실시할 수 있도록 바람직한 실시예를 상세히 설명한다. 다만, 본 발명의 바람직한 실시예를 상세하게 설명함에 있어, 관련된 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략한다. 또한, 유사한 기능 및 작용을 하는 부분에 대해서는 도면 전체에 걸쳐 동일한 부호를 사용한다.
- [0083] 덧붙여, 명세서 전체에서, 어떤 부분이 다른 부분과 연결 되어 있다고 할 때, 이는 직접적으로 연결 되어 있는 경우뿐만 아니라, 그 중간에 다른 소자를 사이에 두고 간접적으로 연결 되어 있는 경우도 포함한다. 또한, 어떤 구성요소를 포함 한다는 것은, 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있다는 것을 의미한다.
- [0084] POE와 관련된 몇 가지 응용 프로그램 중 본 발명은, ADAS(Advanced Driver Assistant System)에서 충돌 회피를 위한 보행자의 자세 방향 추정에 중점을 두었다. ADAS에서 차량은 보행자를 탐지하고 POE를 기반으로 미리 보행자의 의도를 예측할 수 있다. 따라서 보행자가 차량을 알아차리지 않고 도로를 밟고 있을 때, 운전자에게 경고할 수 있으므로, 충돌 가능성이 크게 줄어들 수 있다.
- [0085] 본 발명은, 움직이는 차량에 의해 캡처된 단일 이미지에서 보행자의 포즈 방향 추정(Pose Orientation Estimation; POE)을 목적으로 하므로, 스테레오 카메라 또는 RGBD 센서를 사용하는 3D POE는 고려하지 않는다.
- [0086] 딥러닝 네트워크는 심층 모델 생성을 위해 많은 파라미터가 요구된다. 따라서 많은 양의 곱셈을 수행하기 위해 많은 양의 메모리와 시간이 필요하다. 본 발명에서는, 이러한 심층 네트워크 모델의 단점을 해결하기 위해, 교사-학생 프레임워크를 채택하여, 교사 심층 네트워크를 기반으로 동일한 수준의 성능을 갖춘 더 얇은 학생 모델을 구성하였다.
- [0087] 도 1은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템을 도시한 도면이다. 도 1에 도시된 바와 같이, 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템은, 교사 모델의 출력인 소프트 타겟 데이터 세트 B^* 를 입력으로 하여 학생 모델을 학습시킴으로써, 교사-학생 프레임워크를 통해 모델을 경량화하면서도 우수한 성능을 유지할 수 있고, 클래스 레이블이 포함되는 데이터 세트 A와 클래스 레이블이 포함되지 않는 데이터 세트 B를 이용해 교사 모델을 학습시켜 교사 모델의 오버 피팅(overfitting)을 방지할 수 있으며, 그 출력으로 소프트 타겟 데이터 세트 B^* 를 생성하여 학생 모델을 학습시켜, 학생 모델이 더 유연한 분류 결과를 출력하도록 할 수 있다.
- [0088] 도 1에 도시된 바와 같이, 교사 모델은 교사 심층 네트워크와 교사 랜덤 포레스트의 출력을 결합하여 각 클래스에 대한 소프트 타겟(확률 값)을 생성하고, 이러한 소프트 타겟 값을 입력하여 학생 모델을 훈련할 수 있다. 보다 구체적으로, 도 1을 상세하게 설명하면, (a) 하드 타겟으로 레이블 된 데이터 세트 A를 (b) 교사 심층 네트워크 및 (c) 교사 랜덤 포레스트에 입력하고, (d) 레이블이 지정되지 않은 데이터 세트 B를 학습된 2개의 교사 모델에 입력할 수 있다. (e) 두 교사(교사 심층 네트워크 및 교사 랜덤 포레스트)의 소프트 출력을 하나의 소프트 타겟 벡터로 결합하고, (f) 소프트 타겟 데이터 세트 B^* 를 학생 모델에 입력하여, (g) 학생 모델을 학습시켜, (h) 최종 클래스 확률을 얻을 수 있다. 이와 같이, 본 발명의 교사-학생 프레임워크는 네트워크의 크기를 줄일 수 있을 뿐 아니라, 교사 모델의 POE 기능을 모방할 수 있는 학생 모델을 쉽게 구성할 수 있다.
- [0089] 도 2는 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법의 흐름을 도시한 도면이다. 도 2에 도시된 바와 같이, 본 발명의 일실시예에 따른 교사-학생

프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법은, 데이터 세트 A를 이용하여 교사 모델을 학습시키는 단계(S100), 데이터 세트 B를 교사 모델에 입력하고, 출력된 소프트 출력을 이용하여 소프트 타겟 데이터 세트 B*를 생성하는 단계(S200), 데이터 세트 B*를 이용하여 학생 모델을 학습시키는 단계(S300) 및 학습된 학생 모델로 보행자 포즈 방향을 추정하는 단계(S400)를 포함하여 구현될 수 있으며, 웨이블릿 변환을 적용하여 입력 이미지에 대한 전처리를 수행하는 단계(S10)를 더 포함하여 구현될 수 있다.

[0090] 이하에서는, 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법의 각 흐름에 대하여 상세히 설명하도록 한다.

[0091] 단계 S100에서는, 데이터 세트 A를 이용하여, 교사 심층 네트워크 및 교사 랜덤 포레스트로 구성되는 교사 모델을 학습시킬 수 있다.

[0092] 교사 모델은 교사 심층 네트워크와, 많은 양의 학습 데이터를 기반으로 한 높은 수준의 성능을 가진 교사 랜덤 포레스트(Random Forest; RF)를 사용하여 구성될 수 있다. 단계 S100에서, 데이터 세트 A는 클래스 레이블이 포함되는 하드 타겟 데이터 세트일 수 있다. 즉, 단계 S100에서, 교사 모델은 0 또는 1로 레이블 된 데이터 세트 A를 사용하여 학습될 수 있다.

[0093] 교사 심층 네트워크 T는, 일반 CNN 모델의 소프트맥스(softmax) 함수와는 상이하게, 소프트 타겟(출력 확률(output probability))을 얻기 위해, 다음 수학적 식 1과 같은 연화된 소프트맥스 함수(softened softmax function)를 교사 사전-소프트맥스 활성화 벡터(the vector of the teacher pre-softmax activations) aT에 적용할 수 있다. 교사 심층 네트워크의 기본 아이디어는, 학생 네트워크가 실제 레이블에 의해 제공되는 정보뿐만 아니라, 교사 심층 네트워크에 의해 학습된 더 작은 구조를 포착할 수 있게 하는 것이다.

수학적 식 1

$$P_T^{\tau} = \frac{\exp\left(\frac{aT_i}{\tau}\right)}{\sum_j \exp\left(\frac{aT_j}{\tau}\right)}$$

[0094]

[0095] 여기서, aT는 표본의 진정한 레이블에 대한 하나의 하드 타겟 표현에 매우 가깝지만, 교사의 소프트 출력(softened output, P_T)은 템퍼러처(Temperature)($\tau > 1$)가 증가함에 따라 더 부드럽게 분포할 수 있다.

[0096] 이 방법은 교사 심층 네트워크의 출력에서 나오는 신호를 부드럽게 하고, 학생 모델을 학습하는 동안 학생 네트워크에 더 많은 정보를 제공할 수 있다. 그러나 학생 네트워크의 성능은 템퍼러처에 민감하기 때문에, 이 값은 모든 학습 데이터에 대해 경험적으로 결정되어야 하며, 최적의 템퍼러처를 예측하려면 상당한 노력이 요구된다.

[0097] 본 발명의 단계 S100에서는, 템퍼러처를 결정하고 교사 모델의 소프트 출력을 얻는데 필요한 노력을 줄이기 위해, 도 1에 도시된 바와 같이, 교사 심층 네트워크의 소프트 출력과 교사 랜덤 포레스트를 결합하여 새로운 소프트 출력을 선택하였다. 의사 결정 트리 양상블 분류기인 랜덤 포레스트는 기존의 분류기에 비해 높은 학습 속도로 매우 많은 양의 데이터를 처리하는 것으로 알려져 있다. 또한, 랜덤 포레스트는 본질적으로 특정 클래스에 대해 더 부드러운 분류 결과의 분포를 제공한다.

[0098] 도 3은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법에서, 단계 S100의 세부적인 흐름을 도시한 도면이다. 도 3에 도시된 바와 같이, 본 발명의 일

실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법의 단계 S100은, 데이터 세트 A를 이용하여 교사 심층 네트워크를 학습시키는 단계(S110) 및 교사 심층 네트워크의 특징 맵을 이용하여 교사 랜덤 포레스트를 학습시키는 단계(S120)를 포함하여 구현될 수 있다.

[0099] 단계 S110에서는, 데이터 세트 A를 이용하여 교사 심층 네트워크를 학습시킬 수 있다. 보다 구체적으로, 단계 S110에서는, 먼저 학습 데이터 세트 A를 사용하여 교사 심층 네트워크를 학습시키는데, 이때 데이터 세트 A는 클래스 레이블이 포함되는 하드 타겟 데이터 세트일 수 있다. 데이터 세트 $A = \{(x_i, y_i) | i=1, 2, \dots, N\}$ 는 M 차원 입력 벡터 $x_i = (x_{i1}, x_{i2}, \dots, x_{iM})$ 및 x_i 의 전문가가 표시한 스칼라 클래스 레이블 $y_i = \{g_1, g_2, \dots, g_c\}$ 로 구성될 수 있다.

[0100] 교사 심층 네트워크는 ResNet-101 모델(He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition, In Proceedings of IEEE Conference of Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV USA, 26 June 1 July 2016; pp. 770-778.)을 기반으로 클래스 레이블이 지정된(하드 타겟) 학습 데이터로 생성될 수 있다. 교사 심층 네트워크의 구조는 101개의 매개변수 계층(parameter layers), 하나의 평균 풀링 계층(average pooling layer), 및 하나의 완전 연결 계층(fully connected layer)로 구성될 수 있다. ResNet은 각 3×3 필터 쌍에 하나 이상의 레이어를 건너뛰는 바로가기 연결(shortcut connection)을 추가하지만, 기본 아키텍처는 일반 CNN과 동일할 수 있다. 또한, ResNet은 모든 바로가기와 제로 패딩(zero-padding)에 identity mapping을 사용하여 차원 수를 늘릴 수 있다. 짧은 연결의 출력은 쌓인 레이어의 출력에 추가될 수 있다.

[0101] 학습 데이터 세트 A가 주어지면, ImageNet에서 사전 교육된 ResNet-101 모델을 더 작은 데이터 세트 A로 미세 조정하여, 새로운 태스크를 위한 모든 네트워크 가중치를 업데이트할 수 있다. 교사 심층 네트워크를 학습한 후, 수학적 1은 출력 단위(클래스)에 소프트 출력을 제공할 수 있다. 즉, 미리 정해진 각도로 구분된 클래스, 예를 들어 45도 각도로 구분된 8개의 클래스에 출력 확률이 제공될 수 있다.

[0102] 단계 S120에서는, 교사 심층 네트워크의 특징 맵(feature map)을 이용하여 교사 랜덤 포레스트를 학습시킬 수 있다. 즉, 단계 S120에서, 두 번째 분류기로서 교사 랜덤 포레스트의 개별 의사 결정 트리는 하드 클래스 레이블 y_i 를 갖는 입력벡터 x_i 에 대한 최종 특징 벡터를 사용하여 학습될 수 있다. 의사 결정 트리의 학습은 정보 이득(information gain)을 사용하여 부분집합의 무작위 표본 추출 및 분리 함수 선택에 기반을 둔다. 샘플 x 의 최종 클래스 분포는 다음 수학적 2와 같이, 모든 트리 T 의 각 클래스 확률 분포 $p_t(c_i|x)$ 의 앙상블(산술평균)을 사용하여 생성될 수 있다.

수학적 2

[0103]
$$p(c_i|x) = \frac{1}{T} \sum_{t=1}^T p_t(c_i|x)$$

[0104] 단계 S200에서는, 데이터 세트 B를 단계 S100에서 학습된 교사 모델에 입력하고, 출력된 소프트 출력을 이용하여 소프트 타겟 데이터 세트 B^* 를 생성할 수 있다. 보다 구체적으로, 단계 S200에서는, 도 1의 (e)에서와 같이, 데이터 세트 B를 단계 S100에서 학습된 교사 심층 네트워크 및 교사 랜덤 포레스트에 입력하고, 교사 심층 네트워크의 출력 및 교사 랜덤 포레스트의 출력을 하나의 소프트 타겟 벡터로 결합하여, 각각의 클래스의 확률값인 소프트 타겟 데이터 세트 B^* 를 생성할 수 있다. 이때, 단계 S200에서, 데이터 세트 B는 클래스 레이블이 포함되지 않는 데이터 세트일 수 있다.

[0105] 보다 구체적으로, 교사 모델의 학습이 완료된 다음, 훨씬 더 크고 레이블이 없는 학습 데이터 세트 B가 교사 모

텔에 적용되고, 하드 타겟과 반대되는 소프트 타겟(클래스 확률)으로 구성된 새로운 데이터 세트 B^* 가 구성될 수 있다.

- [0106] 교사 모델에 하나의 데이터 세트 A만을 적용하여 소프트 타겟을 생성하는 알고리즘과는 달리, 본 발명의 접근법은 교사 모델의 오버 피팅(overfitting)을 방지하기 위해 추가 학습 데이터 세트 B를 사용할 수 있다. 단계 S200에서 생성된 새로운 소프트 타겟 데이터 세트 B^* 는, 다른 클래스들 사이의 관계를 유지함으로써, 원래 하드 타겟 데이터보다 더 많은 정보를 포착할 수 있다. 또한, 하드 타겟 데이터 세트를 사용하는 것보다 더 유연한(flexible) 분류 결과를 얻을 수 있다.
- [0107] 데이터 세트 B에 포함된 모든 M개의 샘플 x 가 학습된 후, 클래스 확률 p_i^* (소프트 타겟)로 표현된 새로운 데이터 세트 B^* 가 $B^* = \{(x_i, p_i^*) | i=1, 2, \dots, M\}$ 와 같이 생성될 수 있다.
- [0108] 인식 성능의 차이는 랜덤 포레스트에서 사용된 의사 결정 트리의 수에 따라 발생할 수 있으며, 트리 수는 300으로 설정할 수 있다.
- [0109] 한편, 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법은, 단계 S10을 더 포함하여 구현될 수 있다. 즉, 단계 S10에서는, 웨이블릿 변환을 적용하여 입력 이미지에 대한 전처리를 수행할 수 있다. 보다 구체적으로는, 단계 S10에서는 하이 패스(high-pass) 필터된 2개의 서브이미지 및 로우 패스(low-pass) 필터된 하나의 서브이미지를 생성하고, 단계 S100에서는 단계 S10에서 생성된 3개의 서브이미지를 이용하여 교사 모델을 학습시킬 수 있다.
- [0110] 구체적으로는, 연화된 소프트맥스 함수(soften softmax function) 외에도, 웨이블릿 변환의 세 가지 수작업 필터 응답을 모델에 제공할 수 있다. 즉, 2개의 하이 패스(high-pass) 필터된 서브이미지(sub-images)(LH 및 HL) 및 하나의 로우 패스(low-pass) 필터된 서브이미지(LL)를 사용하여, 도 1의 (a)에 도시된 바와 같은 회색 이미지와 함께 Daubechies D4 웨이블릿을 사용하여 적절한 수작업 특성을 제공하면, 특정 분류 문제에 대한 결과가 향상될 수 있다. 또한, 웨이블릿 변환은 양호한 공간 주파수 지역 특성을 가지며, 이미지의 공간 정보 및 기울기 정보를 보존할 수 있기 때문에, 다양한 밝기 조건에서 POE 성능을 향상시키는 데 도움이 될 수 있다.
- [0111] 단계 S300에서는, 단계 S200에서 생성된 데이터 세트 B^* 를 이용하여, 학생 네트워크 및 학생 랜덤 포레스트로 구성되는 학생 모델을 학습시킬 수 있다.
- [0112] 교사 모델을 학습시킨 후, 교사 모델에서 생성된 소프트 타겟 데이터 세트 B^* 를 사용하여, 보행자의 방향을 추정하는 학생 모델을 구성할 수 있다. 학생 모델은 교사 모델에서처럼 학생 네트워크 1개와 학생 랜덤 포레스트 1개로 구성될 수 있다.
- [0113] 학생 네트워크는 DarkNet 레퍼런스 모델(Darknet reference model. Available online: <https://pjreddie.com/darknet/imagenet/#reference> (accessed on 27 December 2018).)을 수정하여 생성할 수 있다. 이는 DarkNet 레퍼런스 모델의 계산 속도는, 파라미터의 개수가 1/5 및 1/10일 때 하나의 CPU에서 기존의 ResNet-101보다 16배 빠르고 AlexNet보다 2배 빠르기 때문이다. 따라서 교사 심층 네트워크를 압축하는 대신, 학생 네트워크로 알은 DarkNet 레퍼런스 모델을 사용하고, 교사 모델에서 생성한 소프트 타겟 데이터 세트를 사용하여 학생 네트워크를 다시 학습시킬 수 있다.

[0114] 학생 네트워크의 구조는 7개의 맥스 풀링 레이어(max pooling layer), 각 컨볼루션 레이어(convolution layer) 다음에 하나의 평균 풀링 레이어를 포함하는 총 8개의 컨볼루션 레이어로 구성될 수 있다. 전면의 7개의 컨볼루션 레이어는 33 크기의 컨볼루션 필터와 2×2 크기의 필터가 있는 맥스 풀링 레이어가 있으며, 마지막 컨볼루션 레이어는 1×1 크기의 컨볼루션 필터가 있고, 완전 연결 레이어(fully connected layer) 대신 평균 풀링 레이어가 있어서, 오버 피팅 문제를 방지하고 완전 연결 레이어의 학습 가능한 파라미터의 수를 줄일 수 있다. 또한, 각 컨볼루션 레이어에 배치 정규화(batch normalisation)가 적용되며, leaky ReLU(LReLU)는 dying ReLU 문제 해결을 위해 활성화 함수로 사용될 수 있다. LReLU 함수 $f(x)$ 는 다음 수학적 식 3과 같이 함수가 0인 대신 $x < 0$ 일 때 작은 음수값을 갖는다.

수학적 식 3

[0115] $f(x) = \max(0.01x, x)$

[0116] 학생 네트워크를 학습시키기 위해, ImageNet에서 사전 훈련된 컨볼루션 가중치를 사용하고, 소프트 타겟 데이터 세트 B^* 를 사용하여 미세 조정을 수행할 수 있다. 교차-엔트로피 기준(cross-entropy criterion)은 다음 수학적 식 4와 같이 하드 타겟 벡터를 소프트 타겟 벡터로 대체함으로써, 프레임 단위 최소화에 기반할 수 있다.

수학적 식 4

[0117] $\mathcal{L}(W_S) = -\sum_{i=1}^N \sum_{j=1}^C P_T(x_i|c_j) \log P_S(x_i|c_j)$

[0118] 여기서, N 은 데이터 세트 B^* 의 샘플 수이고 C 는 클래스 수이다. 또한, $P_T(x_i|c_j)$ 와 $P_S(x_i|c_j)$ 는 각각 입력 벡터 x_i 에 대한 교사와 학생의 후방(posterior) 클래스 확률이다.

[0119] 도 4는 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법에서, 학생 네트워크의 학습 절차를 설명한 알고리즘을 도시한 도면이다. 단계 S300에서는, 도 4에 도시된 바와 같은 알고리즘으로, 학생 네트워크를 먼저 학습시킬 수 있다.

[0120] 학생 네트워크와 마찬가지로, 학생 랜덤 포레스트의 초기 의사 결정 트리는 트리 수, 트리 깊이 및 개별 트리의 각 노드에 대한 스플릿 임계값을 갖는 스플릿 함수 등 교사 랜덤 포레스트와 동일한 구조를 사용할 수 있다.

[0121] 학생 랜덤 포레스트의 의사 결정 트리의 학습은, 입력 벡터 v 와 클래스 확률 p_i^* 를 입력으로 한다. 입력 벡터는 마지막 특징 맵(last feature maps)(4×4×8)으로부터 생성된 128 차원을 가지고, 클래스 확률은 교사 모델의 출력으로부터 추정될 수 있다. 소프트 타겟 데이터 세트 B^* 의 클래스 확률을 갖는 출력 벡터로 구성된 학습 데이터로부터, 학생 랜덤 포레스트의 의사 결정 트리는 샘플로부터 클래스 확률을 갖는 p 개의 변수를 무작위로 선택할 수 있다.

[0122] B'_0 가 노드 0에서 샘플을 나타낸다고 하자. 사전에 훈련되고 임의로 생성된 스플릿 함수(split function) $f(v_p)$ 는 랜덤 서브 세트(subset) B'_0 를 노드 0에서 좌측(B'_l) 및 우측(B'_r) 서브 세트로 반복적으로 분할할 수 있다. 최상의 스플릿 함수를 선택하기 위해, 노드 0의 엔트로피 $E(0)$ 는 확률 분포 P_j^* 를 갖는 p 변수만을 사용하여 추정될 수 있다. 본 발명에서, 노드 0의 엔트로피 $E(B'_0)$ 는 다음 수학적 식 5와 같이 정의될 수 있다.

수학식 5

$$E(B'_0) = -\sum_{i=1}^N \sum_{j=1}^C P_j^*(x_i) \log P_j^*(x_i)$$

동일한 방법을 사용하여, 노드 0의 좌측 및 우측 서브 세트는 B'_l 및 B'_r 로 분할되고, 엔트로피 $E(B'_l)$ 및 $E(B'_r)$ 이 계산될 수 있다. 3개의 엔트로피로부터, 노드 0의 정보 이득 E는 다음 수학식 6으로부터 계산될 수 있다.

수학식 6

$$\Delta E = E(B'_0) - \frac{|B'_l|}{|B'_0|} E(B'_l) - \frac{|B'_r|}{|B'_0|} E(B'_r)$$

이 과정은 후보 스플릿 함수의 수를 적용하는 동안 반복될 수 있으며, 최대 E를 갖는 함수를 노드 0에 대한 최상의 스플릿 함수 $f(v_p)$ 로 결정할 수 있다.

초기 의사 결정 트리 Tr_t 가 확장된 후에, C 클래스들의 확률 분포가 리프 노드(leaf node)에 저장될 수 있다. 그 다음에, 수학식 6에서 교사 랜덤 포레스트에 의해 기술된(transcribed) 데이터 세트 B^* 의 샘플 i의 j번째 클래스 분포를 나타내는 $P_{ij}^*(Te)$ 와, 구성된 의사 결정 트리 t에 기반을 둔 샘플 i의 j번째 클래스 분포를 나타내는 $P_{ij}^*(S_t)$ 로, 교차-엔트로피(cross-entropy)를 추정할 수 있다. 최종 교차-엔트로피의 일반적인 형태는 다음 수학식 7과 같다.

수학식 7

$$Tr(Te, S)_t = -\sum_{i=1}^N \sum_{j=1}^C P_{ij}^*(Te) \log (P_{ij}^*(S_t))$$

부스팅(boosted) 랜덤 포레스트의 높은 성능에 힘입어, $Tr(Te, S)_t$ 가 최소 기준 Θ 이하가 될 때까지 t번째 약한 의사 결정 트리를 업데이트하기 위해 부스팅을 반복할 수 있다. T개의 랜덤 의사 결정 트리가 완성되면, 학생 랜덤 포레스트는 최종적으로 클래스 당 확률 분포로 구성된 T 트리가 될 수 있다.

도 5는 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법에서, 학생 랜덤 포레스트의 학습 절차를 설명한 알고리즘을 도시한 도면이다. 단계 S300에서는, 도 5에 도시된 바와 같은 알고리즘으로 학생 랜덤 포레스트를 학습시킬 수 있다.

단계 S400에서는, 단계 S300에서 학습된 학생 모델로 보행자 포즈 방향을 추정할 수 있다.

도 6은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법에서, 보행자 방향 클래스 분류를 예를 들어 도시한 도면이다. 보행자 방향 추정에 있어서, 모든 방향이 예측될 때 클래스 수는 증가할 수 있다. 따라서 대부분의 기존 연구에서는 도 6에 도시된 바와 같이 방향을 N개의 클러스터로 나누어 인식하는 방법을 사용하였다. 예를 들어, TUD 데이터 세트는 테두리 상자의 뒤, 앞, 왼쪽, 오른쪽, 왼쪽 뒤, 오른쪽 뒤, 왼쪽 앞 및 오른쪽 앞과 같은 방향 주석이 있는 보행자 이미지 5,228개로 구성된다. TUD 데이터 세트의 경우, 방향 클래스가 45도로 나뉜다. 그러나 보행자의 특정 각도를

명확하게 판단할 수 없기 때문에 모호성이 발생할 수 있다. 따라서 모든 보행자의 각도를 0도 또는 45도로 표현하거나, 앞과 뒤와 같은 표현을 사용하는 것은 적절하지 않을 수 있다.

[0133] 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법의 단계 S400에서는, 이전의 POE 접근 방식과는 상이하게, 최대 확률 클래스의 확률 및 최대 확률 클래스와 인접한 클래스의 확률의 조합을 이용한 소프트 가중 방법을 이용해 보행자 방향을 추정할 수 있다. 즉, 최종 클래스 확률(도 1의 (h)) 및 인접한 조합을 이용한 소프트 가중 방법(soft weighting method)을 제안한다.

[0134] 도 7은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법에서, 단계 S400의 세부적인 흐름을 도시한 도면이다. 도 7에 도시된 바와 같이, 본 발명의 일 실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법의 단계 S400은, 학생 네트워크 및 학생 랜덤 포레스트의 출력값을 결합하여 최종 확률을 생성하는 단계(S410), 최대 확률 클래스 및 최대 확률 클래스에 인접한 두 클래스를 특징하는 단계(S420), 특징된 클래스의 가중치를 각각 산출하는 단계(S430) 및 각 클래스의 가중치로부터 최종 보행자 방향을 추정하는 단계(S440)를 포함하여 구현될 수 있다.

[0135] 단계 S410에서는, 단계 S300에서 학습된 학생 네트워크 및 학생 랜덤 포레스트의 출력값을 결합하여 최종 확률을 생성할 수 있다. 즉, 본 발명에서는, 학습된 학생 모델에 보행자 추정 대상이 되는 이미지를 입력하고, 학생 네트워크의 출력값과 학생 랜덤 포레스트의 출력값을 결합하여 최종 확률을 생성함으로써, 모호한 보행자 각도를 보다 정확하게 추정할 수 있다.

[0136] 단계 S420에서는, 보행자의 방향 추정을 위해 미리 정해진 각도로 구분된 클래스에서, 가장 높은 확률을 갖는 최대 확률 클래스(c_k) 및 최대 확률 클래스에 인접한 두 클래스(c_{k-1} , c_{k+1})를 특징할 수 있다. 즉, 단계 S420에서는, 각 클래스의 가장 높은 확률만을 검색하는 대신, 단계 S410에서 생성된 최종 확률에서 가장 높은 확률을 갖는 각도 클래스(c_k)를 먼저 찾고, 앞과 뒤의 인접한 두 클래스(c_{k-1} , c_{k+1})를 검색하여, 총 3개의 클래스의 확률에 따라 최종 보행자 각도를 예측할 수 있다.

[0137] 단계 S430에서는, 특징된 클래스(c_{k-1} , c_k , c_{k+1})의 가중치를 각각 산출할 수 있다. 보다 구체적으로, 단계 S430에서는, 다음 수학적 식 8을 이용해 각 클래스의 가중치 $w(c_k)$ 를 산출하며, 최대 확률 클래스의 가중치 $w(c_k)$ 는 다음 수학적 식 8에 따라 지수 가중 방식에 따르도록 추정될 수 있다.

수학적 식 8

$$w(c_k) = 1/(\exp(F \cdot P(c_k)))$$

[0138]

[0139] 여기서, $P(c_k)$ 는 k 번째 클래스 c_k 의 확률이고, F 는 가중치 인자(weighting factor)이다. 가중치 인자 F 는 $w(\cdot)$ 에서 $P(c_k)$ 의 영향을 최대화(최소화)하도록 선택될 수 있다. F 가 1일 때 $P(c_k)$ 의 변화는 $w(\cdot)$ 에 지수적으로 반영될 수 있다. 지수적 가중치는 지역 특징 관련성(local feature relevance)에서 변화에 더 민감하고, 더 큰 성능 향상을 가져올 수 있다. 본 발명에서는, 실험 결과에 따라 F 를 0.5로 설정하였다. 또한, 수학적 식 8은 최대 확률을 갖는 클래스에 기초하여 순방향(c_{k+1}) 및 역방향(c_{k-1}) 클래스의 확률값에도 적용될 수 있다. 따라서 수학적 식 8을 이용해, 특징된 3개의 클래스(c_{k-1} , c_k , c_{k+1})의 가중치 $w(c_{k-1})$, $w(c_k)$, $w(c_{k+1})$ 를 각각 산출할 수 있다.

[0140] 단계 S440에서는, 단계 S430에서 산출된 각 클래스의 가중치로부터 최종 보행자 방향을 추정할 수 있다. 즉, 3개의 클래스 및 각각의 가중치가 결정된 후, 단계 S440에서는, 다음 수학적 식 9를 이용해 최종 보행자 방향 ($\tilde{O}$)을 추정할 수 있다.

수학적 식 9

[0141]
$$\tilde{O} = \frac{1}{3} \sum_{k=1}^{K-1} w(c_k) \cdot O_k$$

[0142] 여기서, O_k 는 클래스 c_k 의 보행자 방향이다. 수학적 식 9에서, 최종 보행자 방향은 N 방향 중 하나로 결정되지 않고, 실제 보행자의 방향에 가까운 값을 갖도록 조정될 수 있다.

[0143] 도 8은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템의 구성을 도시한 도면이다. 도 8에 도시된 바와 같이, 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 시스템은, 데이터 세트 A를 이용하여, 교사 심층 네트워크 및 교사 랜덤 포레스트로 구성되는 교사 모델을 학습시키는 교사 학습 모듈(100), 데이터 세트 B를 교사 학습 모듈(100)에서 학습된 교사 모델에 입력하고, 출력된 소프트 출력을 이용하여 소프트 타겟 데이터 세트 B*를 생성하는 소프트 타겟 데이터 생성 모듈(200), 소프트 타겟 데이터 생성 모듈(200)에서 생성된 데이터 세트 B*를 이용하여, 학생 네트워크 및 학생 랜덤 포레스트로 구성되는 학생 모델을 학습시키는 학생 학습 모듈(300) 및 학생 학습 모듈(300)에서 학습된 학생 모델로 보행자 포즈 방향을 추정하는 방향 추정 모듈(400)을 포함하여 구성될 수 있으며, 웨이블릿 변환을 적용하여 입력 이미지에 대한 전처리를 수행하는 전처리 모듈(500)을 더 포함하여 구성될 수 있다.

[0144] 또한, 도 8에 도시된 바와 같이, 교사 학습 모듈(100)은, 데이터 세트 A를 이용하여 교사 심층 네트워크를 학습시키는 제1 교사 학습부(110), 및 교사 심층 네트워크의 특징 맵(feature map)을 이용하여 교사 랜덤 포레스트를 학습시키는 제2 교사 학습부(120)를 포함하여 구성될 수 있다. 또한, 방향 추정 모듈(400)은, 학생 학습 모듈(300)에서 학습된 학생 네트워크 및 학생 랜덤 포레스트의 출력값을 결합하여 최종 확률을 생성하는 최종 확률 생성부(410), 보행자의 방향 추정을 위해 미리 정해진 각도로 구분된 클래스에서, 가장 높은 확률을 갖는 최대 확률 클래스(c_k) 및 최대 확률 클래스에 인접한 두 클래스(c_{k-1} , c_{k+1})를 특징하는 클래스 검색부(420), 특징된 클래스(c_{k-1} , c_k , c_{k+1})의 가중치를 각각 산출하는 가중치 산출부(430), 및 가중치 산출부(430)에서 산출된 각 클래스의 가중치로부터 최종 보행자 방향을 추정하는 최종 방향 추정부(440)를 포함하여 구성될 수 있다.

[0145] 각각의 구성들과 관련된 상세한 내용들은, 앞서 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법의 각 단계와 관련하여 충분히 설명되었으므로, 상세한 설명은 생략하기로 한다.

[0146] 실험 결과

[0147] 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템의 효율성을 입증하기 위해, 벤치마크 데이터베이스를 사용하여 본 발명의 성능을 평가하고, 최근 연구에서 제시된 다른 접근법을 사용하여 비교 실험을 수행하였다.

[0148] 본 실험에서는, 먼저 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템이 다양한 보행자 자세의 방향 추정에 효과적임을 증명하기 위해, POE의

성능을 검증하였다. 이 실험은 Microsoft Windows 10을 실행하는 24GB RAM의 Intel Core i7 프로세서를 사용하여 수행되었다. 또한, 교사 랜덤 포레스트 및 학생 랜덤 포레스트를 포함한 모든 RF 접근 방식은 CPU를 기반으로 실행되었으며 교사 심층 네트워크는 하나의 Titan Xp GPU를 사용하여 실행되었다.

[0149] 교사 심층 네트워크의 학습을 위해, 배치 크기, 운동량, 학습률(learning rate) 및 가중치 감퇴(weight decay)는, 각각 32, 0.9, 0.001 및 0.0005로 설정하였다. 랜덤 포레스트의 경우, 성능 및 트리를 저장하는 데 필요한 메모리 측면에서 중요한 파라미터는 트리의 깊이 및 그 개수이다. 본 실험에서는, 최대 트리 깊이를 20, 교사 랜덤 포레스트의 트리 개수를 300으로 설정하였다. 학생 랜덤 포레스트의 트리 개수를 결정하기 위해, 나무의 개수를 250, 200, 150, 100, 70 및 50으로 순차적으로 줄였으며, 실험 결과에 기반하여 추후 설명할 바와 같이, 더 정확하고 빠른 계산을 위해 70으로 설정하였다.

[0150] 학생 네트워크와 학생 랜덤 포레스트는 도 4 및 도 5에 도시된 바와 같은 알고리즘 1 및 알고리즘 2를 기반으로 한 소프트 타겟 학습 데이터 세트 B^* 를 사용하여 재학습되었다. 보행자 감지와 관련된 많은 데이터 세트의 사용이 가능하지만, 상대적으로 보행자 방향 추정을 한 연구는 거의 없는 실정이다. 따라서 본 발명에서는, 테두리 상자과 이산 방향 주석이 있는 5,228개의 보행자 이미지로 구성된 가장 인기 있는 TUD 멀티 뷰 보행자 데이터 세트를 이용해 POE 실험을 수행하였다. 이 데이터 세트에는 학습을 위한 전신 보행자 이미지 4,732개, 유효성 검사 248개, 테스트 248개가 포함되어 있다. TUD 데이터 세트의 이미지는 실제 거리의 상황에서 촬영되었으며, 모든 이미지에는 다양한 포즈와 옷이 포함되어 있어서 데이터 세트를 훨씬 까다롭게 만들었다. 작은 데이터 세트로 훈련된 모델은 검증 및 테스트 세트의 데이터를 일반화하지 않아 오버 피팅되는 결과를 초래한다는 것은 일반적인 사실이다. 오버 피팅을 줄이기 위해, 본 발명에서는 이미지 이동, 확대 및 축소, -15° 에서 $+15^\circ$ 사이에서 임의의 각도로 회전, 왼쪽-오른쪽 뒤집기 및 자르기와 같은 데이터 확대를 적용하여 데이터 세트의 크기를 증가시켰다. 모든 학습 이미지에는 교사 모델에 제공된 원본 이미지와 복제 이미지가 포함된다. 앞서 언급한 데이터 증가에 의해, 데이터 세트 A에 4,732개의 이미지를, 데이터 세트 B에 4,732개의 이미지를 할당하였다.

[0151] 8개의 클래스의 방향 추정의 유효성을 검증하기 위해, TUD 데이터 세트의 정밀도(precision), 회수율(recall) 및 오탐율(False positive rate, FPR)을 측정하였다. 이 값은 일반적으로 물체 인식 성능을 평가하는 데 사용된다. 또한, 정확도(accuracy; Acc)는 포즈와 혼동 행렬(confusion matrices)을 평가하여 클래스 간의 성능을 비교하는 데 사용된다. 정확도는 조사된 전체 사례수에 대한 탐지 성공 비율이다.

[0152] TUD 데이터 세트에서 성능평가

[0153] 본 발명의 POE 방법의 효과를 검증하기 위해, 다섯 개의 최첨단 방법과 성능을 비교하였다. 각 실험은 다음과 같다. (1) 매우 무작위화된 트리 분류기의 배열을 사용하여 POE를 분류하는 MoAWG, (2) 랜덤 포레스트 분류기와 결합된 부분 최소 제곱 기반 모델을 사용하는 PLS-RF, (3) 신체 자세 방향을 인식하기 위해 희박한 표현 기법(sparse representation technique)을 사용하는 MACF, (4) 16개의 가중 CNN 레이어 및 저해상도 이미지를 갖는 CNN을 사용하는 VGG-16, (5) deep residual nets에 기반을 둔 ResNet-101, (6) 수작업 필터 없는 제안된 교사 모델, (7) 제안된 교사 모델(proposed T-Model), (8) 학생 네트워크 및 학생 랜덤 포레스트를 포함하는 제안된 학생 모델(proposed S-Model). 총 8가지 방법 중 (4) 내지 (8)의 방법은 CNN을 기반으로 한다.

[0154] 도 9는 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법을 포함하는 8개의 실험의 보행자 방향 추정 결과를 비교한 도면이다. 도 9는 평균 정밀도(average precision; AP), 평균 회수율(average recall; AR), 및 평균 FPR (AFPR)의 관점에서 8가지 접근법의 결과를 비교한 것이다. 도 9에 도시된 바와 같이, 모든 실험에서 CNN 기반 방법((4) 내지 (8)의 방법)이 종래의 수작업 및 분류기 기반 방법보다 우수한 분류 성능을 가진다는 것을 확인할 수 있다. MoAWG가 기존의 세 가

지 접근 방식(MoAWG, PLS-RF, MACF) 중 최고 성능을 달성하였으나, 심층 네트워크 기반 접근 방법 중 가장 낮은 성능을 보인 VGG-16보다 0.2%, 3.2% 및 0.6%의 낮은 성능을 보였다. VGG-16과 ResNet-101은 기존의 접근법보다는 나은 성능을 보였으나, 기본적인 CNN 모델을 사용하기 때문에, 그 성능이 본 발명보다 낮다는 것을 확인할 수 있다. 제안된 3가지 방법 중에서, T-Model 방법은 교사 심층 네트워크와 교사 랜덤 포레스트를 동시에 사용하기 때문에, 적용된 세 가지 평가 항목에서 다른 방법에 비해 최고의 성능을 보였다.

[0155] 수작업 필터를 사용하여 전처리하는 단계 S10이 없는 제안된 T-Model(Proposed T-Model without handcraft filters)의 경우 원래의 T-Model과 비교하여 세 가지 평가 항목 모두에서 성능이 떨어졌다. 결과에 기초하여, 웨이블릿 변환은 양호한 공간 주파수 위치 특성을 가지며, 이미지의 공간 정보 및 경도 정보를 보존할 수 있음을 알 수 있다.

[0156] 제안된 S-Model의 평가 결과는 T-Model과 비교할 때, AP와 AR 측면에서 7.3%와 5.3%의 약간 낮은 성능을 보였다. 그러나 모델의 크기 축소 비율에 비하여 성능 저하가 적기 때문에, 제안된 방법이 성능을 유지하면서 메모리 및 속도 요구 사항을 효과적으로 향상시킬 수 있다. 제안된 S-Model은, 다른 CNN 기반의 방법과 비교할 때, AP 및 AR은 상대적으로 높고 AFPR은 낮다. 이는 제안된 방법이 복잡한 배경 또는 흐릿한 보행자 외곽선에 대해 강건하다는 것을 나타낸다.

[0157] 도 10은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템에 따른, 방향 클래스별 POE 분류 정확도(Acc)를 confusion matrix로 도시한 도면이다. 도 10에 도시된 바와 같이, 'Back' 과 'Lback' 을 제외하고 대부분의 방향은 비슷한 분류 성능을 보였다. 다른 방향과 비교할 때 이 두 방향의 정확도가 낮은 주된 이유는 웨이블릿 변환이 CNN의 이전 단계에서 적용되더라도 두 방향이 비슷한 모양을 가졌기 때문이다. 반면에, 'Lfront' 와 'Rback' 은 외모의 차이로 인해 가장 우수한 분류 성능을 보였다.

[0158] 학생 RF에 대한 의사 결정 트리의 최적 수 결정

[0159] 학생 랜덤 포레스트의 경우 의사 결정 트리의 수는 처리 시간 및 메모리 절약을 위한 파라미터 수를 줄이는 데 중요한 요소이다. 학생 랜덤 포레스트의 최적 트리 수를 결정하기 위해 TUD 데이터 세트에서 정밀도(Precision), 회수율(recall) 및 정확도(accuracy) 성능을 비교하면서, 트리 수를 200, 150, 100, 70 및 50으로 순차적으로 줄여서 실험을 수행하였다.

[0160] 도 11은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템에서, 학생 랜덤 포레스트의 트리 수 결정을 위한 실험 결과를 도시한 도면이다. 도 11에 도시된 바와 같이, 트리의 수가 증가함에 따라 정밀도, 회수율 및 정확도가 증가하지만, 파라미터의 수가 상대적으로 증가하고 속도와 압축률이 감소하게 된다. 이러한 결과를 바탕으로, 70개의 트리가 다른 트리 수와 비슷하거나 약간 더 높은 성능을 나타내므로, 학생 랜덤 포레스트의 최적 트리 수라고 볼 수 있다. 따라서 본 발명에서는, 정확도를 높이고 파라미터의 수를 줄이기 위해 학생 랜덤 포레스트의 트리 수를 70으로 설정하였다.

[0161] 모델 압축 평가

[0162] 모델 압축의 목표는 교사 모델과 비슷한 성능으로 매개변수 및 연산이 적은 최적의 학생 모델을 생성하는 것이다. 따라서, 제안된 학생 모델을 인기 있는 모델 압축 방법인 MobileNet과 TUD 데이터 세트를 사용하여 파라미터의 개수 및 연산의 관점에서 비교하였다. 비교 모델은 사전 훈련된 파라미터를 기반으로 한 TUD 학습 데이터

를 사용하여 미세 조정되었다. MobileNets는 파라미터의 개수와 연산을 줄이기 위해 적용되는 separable depth-wise convolutions을 기반으로 한다. 본 실험에서는, 하나의 Titan-X GPU를 사용하여 세 가지 비교 방법을 실행하였다.

[0163] 도 12는 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법을 포함하는 4개의 실험의 정확도, 파라미터의 수 및 연산 수를 비교한 도면이다. 도 12에 도시된 바와 같이, 제안된 학생 모델은 교사 모델에 비해 파라미터의 수를 약 5배, 연산 수를 약 19.6배 줄일 수 있다. 즉, 학생 모델의 POE 정확도는 교사 모델의 POE 정확도보다 다소 낮지만, 요구되는 연산 및 파라미터의 수가 매우 적다는 것을 확인할 수 있다. 또한, 학생 모델은 MobileNet보다 POE 정확도가 17.9% 우수하며, 5배 적은 수의 연산을 사용한다. 그러나 학생 네트워크는 일반적인 컨볼루션 방법을 사용하는데, 이 방법은 연산 수를 19.6배 늘린다. 비교 결과에서 알 수 있듯이, 제안된 모델 압축 방법은 기존의 압축 방법에 비해 POE 인식을 및 연산 횟수 면에서 우수한 성능을 보임을 확인할 수 있다.

[0164] KITTI 데이터 세트에 대한 성능 평가

[0165] POE에서 사용된 알고리즘이 다른 데이터 세트에 효과적으로 적용될 수 있는지 여부를 검증하기 위해, 본 발명의 알고리즘을 KITTI 데이터 세트에도 적용하고 결과를 비교하였다.

[0166] 두 번째 데이터 세트로 사용한 KITTI 데이터 세트는, 스테레오 이미징, 옵티컬 플로(optical flow), 시각적 주행 측정(visual odometry), 3D 객체 탐지(3D object detection) 및 3D 추적(3D tracking)을 포함하는 실제 세계 컴퓨터 비전 벤치마크이다. 이용 가능한 9가지 카테고리 중에서, 보행자 카테고리에 대한 실험을 수행하였다. KITTI 데이터 세트의 보행자 카테고리를 5,415개의 이미지로 구성된 학습 데이터 세트와 2,065개의 이미지로 구성된 유효성 검사 세트로 나누었다. 또한, 학습 데이터 세트에만 데이터 증가를 적용하여 데이터 세트의 크기를 늘리고 4,732개의 이미지의 전체 학습 데이터 세트를 사용했다. 데이터 세트의 난이도는 크기, 폐색(occlusions) 및 절단 수준(truncation level)에 따라 “쉬움(easy)”, “보통(moderate)”, “어려움(hard)”으로 정의하였다. 중요하지 않은 영역의 탐지 또는 최소 크기보다 작은 탐지는 오탐지(false positive)로 간주하지 않는다. KITTI 데이터 세트에 대해 학생 모델을 학습하기 위해, 학습 데이터는 교사 모델에 적용되었고, 교사 모델의 출력인 소프트 타겟 데이터가 학생 네트워크 및 학생 랜덤 포레스트에 적용되었다. 모델이 학습되는 동안, 8개의 각도에서 보행자의 방향을 정규화하고, 수학적 8을 사용하여 연속된 방향 값을 추정하였다.

[0167] 8개의 클래스의 방향 추정의 유효성을 검증하기 위해, KITTI 데이터 세트의 경우, KITTI 데이터 세트의 보행자 데이터가 TUD와 다른 방향으로 계속 표시되어 있기 때문에, 평균 방향 유사성(Average Orientation Similarity; AOS)을 사용하였다.

[0168] 성능 평가를 위해 다음과 같은 최첨단 방법들과의 정확성을 비교하였다. (1) 모델 방법의 변형 가능한 부분을 확장하여 다른 관점을 다루는 DPM-VOC+VP, (2) CNN을 사용하여 단일 단안 영상(single monocular image)으로부터 3D 객체를 검출하는 Mono3D, (3) 하위 카테고리 인식 컨볼루션 신경망 기반의 SubCNN, (4) 고도로 최적화된 CNN 기반 탐지 프레임워크의 최상단에서 관점 추론을 사용한 FRCNN, (5) 제안된 학생 모델.

[0169] 도 13은 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법을 포함하는 5개의 CNN 기반 방법에 대한 실험 결과를 요약한 도면이다. 도 13에 도시된 바와 같이, KITTI 데이터 세트를 이용한 실험에서, DPM-VOC+VP와 FRCNN의 두가지 방법은, 보행자의 외곽이 흐릿하여 입력 이미지가 작을 때 다른 세가지 방법보다 낮은 AOS 비율을 보였다. 그러나 SubCNN 방법은 작은 크기의 보행자 검출하는 데에 이미지 피라미드를 사용하기 때문에, 다른 방법들보다 우수한 AOS 성능을 나타냈다. SubCNN은 KITTI 데이터 세트에 대해 상대적으로 우수한 AOS 비율을 나타냈으나, 본 발명에서 제안된 학생 모델

보다 네트워크 구조가 더 깊고 넓기 때문에, 지역 제안 및 객체 검출을 위한 추가적인 네트워크가 필요하다는 단점이 있다. 그러나 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법은, 교사-학생 구조를 적용하여 AOS 속도를 향상시키고, 두 개의 압축된 분류기(학생 네트워크 및 학생 랜덤 포레스트)가 다른 것의 단점을 보완하여, KITTI 데이터 세트의 쉬운, 보통 및 어려운 데이터에 대하여 우수한 성능을 나타냈다.

[0170] 도 14는 본 발명의 일실시예에 따른 교사-학생 프레임워크 기반의 소프트 타겟 학습방법을 이용한 보행자 포즈 방향 추정 방법 및 시스템을 사용하여 (a) TUD 및 (b) KITTI 데이터 세트의 POE 결과를 도시한 도면이다. 도 14에 도시된 바와 같이, 본 발명에서 제안된 학생 모델은, 보행자의 몸이 다른 보행자에 의해 왜곡되거나 부분적으로 가려졌을 때, 심지어 이미지가 흐릿하더라도 올바르게 보행자의 방향을 예측할 수 있다.

[0171] 이와 같이, 본 발명에서는, 교사 모델과 학생 모델로 구성된 프레임워크를 사용하여, 실제 주행 상황에서 적용할 수 있는 새로운 POE 방법을 제안하였다. 교사 모델은 교사 심층 네트워크의 출력과 교사 랜덤 포레스트의 출력을 결합하여 각 클래스의 확률값을 생성하고, 이와 같은 소프트 타겟 값을 입력하여 학생 모델을 학습시킬 수 있다. 두가지 다른 분류 모델을 결합함으로써, 모델 크기를 줄일 수 있을 뿐 아니라, 교사 모델의 POE 기능을 모방하는 학생 네트워크를 구성할 수 있다. 또한, 기존의 CNN 기반 POE 접근법과 달리, 본 발명은 교사 심층 네트워크와 교사 랜덤 포레스트의 출력을 결합하여 새로운 소프트 출력을 선택하고, 교사 모델을 기반으로 동등한 성능의 학생 모델을 구성할 수 있다.

[0172] 또한, 2개의 벤치마크 데이터 세트를 사용한 실험 결과, 제안된 알고리즘은 CNN뿐만 아니라, 기존의 분류기를 기반으로 한 다른 최첨단 방법에 비해 POE 성능을 향상시킨다는 것을 확인할 수 있었다. 또한, 제안된 학생 모델은 더 깊고 짧은 CNN 기반 네트워크 및 교사 모델에 비해 적은 메모리와 적은 연산이 필요하다는 것이 입증되었다. 따라서 본 발명은 지능형 차량의 임베디드 시스템뿐만 아니라, 감시 및 로봇 비전과 같은 다양한 분야에 적용할 수 있다.

[0173] 이상 설명한 본 발명은 본 발명이 속한 기술분야에서 통상의 지식을 가진 자에 의하여 다양한 변형이나 응용이 가능하며, 본 발명에 따른 기술적 사상의 범위는 아래의 특허청구범위에 의하여 정해져야 할 것이다.

부호의 설명

[0174] 100: 교사 학습 모듈
110: 제1 교사 학습부
120: 제2 교사 학습부
200: 소프트 타겟 데이터 생성 모듈
300: 학생 학습 모듈
400: 방향 추정 모듈
410: 최종 확률 생성부
420: 클래스 검색부
430: 가중치 산출부
440: 최종 방향 추정부
500: 전처리 모듈
S10: 웨이블렛 변환을 적용하여 입력 이미지에 대한 전처리를 수행하는 단계
S100: 데이터 세트 A를 이용하여 교사 모델을 학습시키는 단계

S110: 데이터 세트 A를 이용하여 교사 심층 네트워크를 학습시키는 단계

S120: 교사 심층 네트워크의 특징 맵을 이용하여 교사 랜덤 포레스트를 학습시키는 단계

S200: 데이터 세트 B를 교사 모델에 입력하고, 출력된 소프트 출력을 이용하여 소프트 타겟 데이터 세트 B^* 를 생성하는 단계

S300: 데이터 세트 B^* 를 이용하여 학생 모델을 학습시키는 단계

S400: 학습된 학생 모델로 보행자 포즈 방향을 추정하는 단계

S410: 학생 네트워크 및 학생 랜덤 포레스트의 출력값을 결합하여 최종 확률을 생성하는 단계

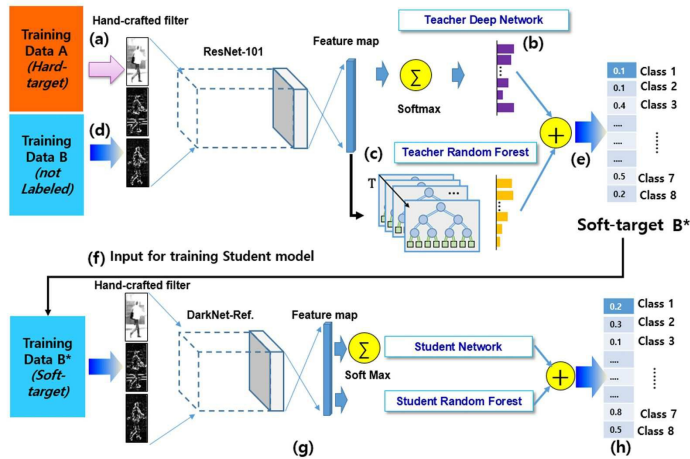
S420: 최대 확률 클래스 및 최대 확률 클래스에 인접한 두 클래스를 특정하는 단계

S430: 특정된 클래스의 가중치를 각각 산출하는 단계

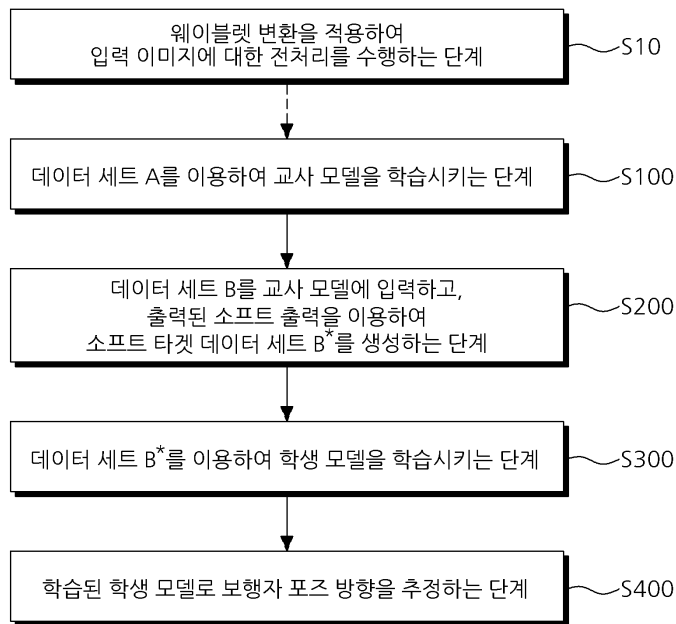
S440: 각 클래스의 가중치로부터 최종 보행자 방향을 추정하는 단계

도면

도면1

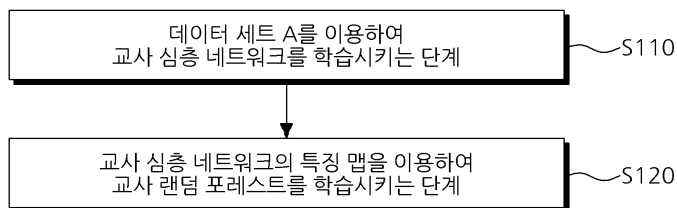


도면2



도면3

S100



도면4

Algorithm 1: Student network training

B^* : Soft target dataset

W_S : Parameters of a student network

1. Initialise W_S
2. Input B^* to pretrained network
3. Compute loss function $\mathcal{L}(W_S)$ using Eq. (4)
4. $W_S^* \leftarrow \arg \min_{W_S} \mathcal{L}(W_S)$
5. Select the best parameter W_S^* for student network

도면5

Algorithm 2: Student RF training

$\mathbf{B}^*$: Soft target dataset

T : The maximum number of decision trees to grow

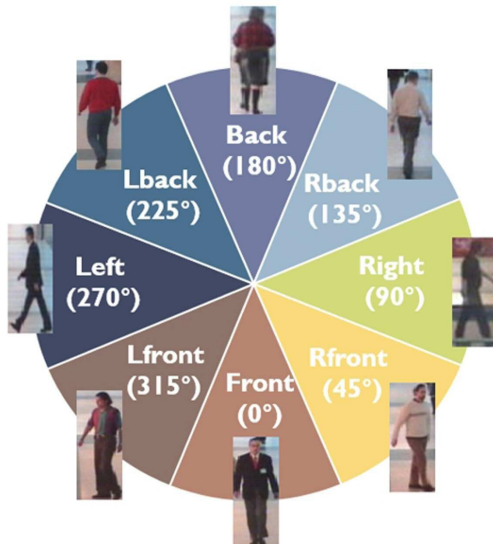
$T - RF_t$: t -th tree structure of teacher RF

$S - RF_t$: t -th tree structure of student RF

θ : The minimum threshold for stopping the boosting

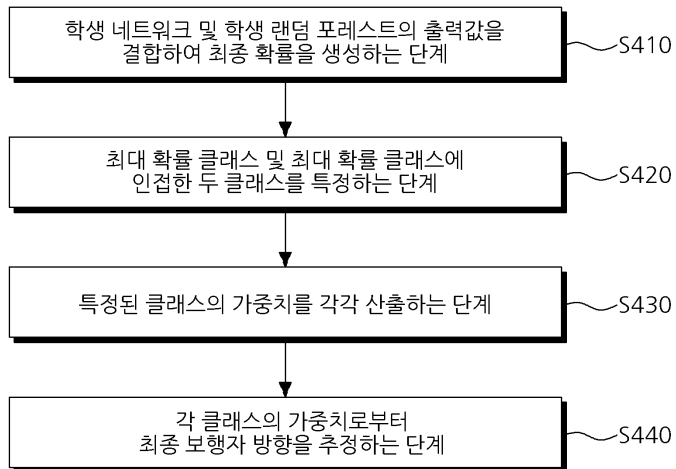
1. Copy $T - RF_t$ to $S - RF_t$ (transfer learning)
2. Input $\mathbf{B}^*$ with v to one of the decision trees of $S - RF_t$
3. Generate split functions $f(v)$ at node O
4. Compute information gain ΔE of node O using Eq. (6)
5. Compute the entropy $\text{Tr}(\text{Te}, S)_t$ using Eq. (7)
6. If the cross-entropy is below θ
 - If $\text{Tr}(\text{Te}, S)_t < \theta$
 - 6-1. then save current $S - RF_t$
 - $t = t + 1$, go to step 1
 - 6-2. otherwise, go to step 2.
7. Continue from step 1 to 5 until T trees are constructed.

도면6

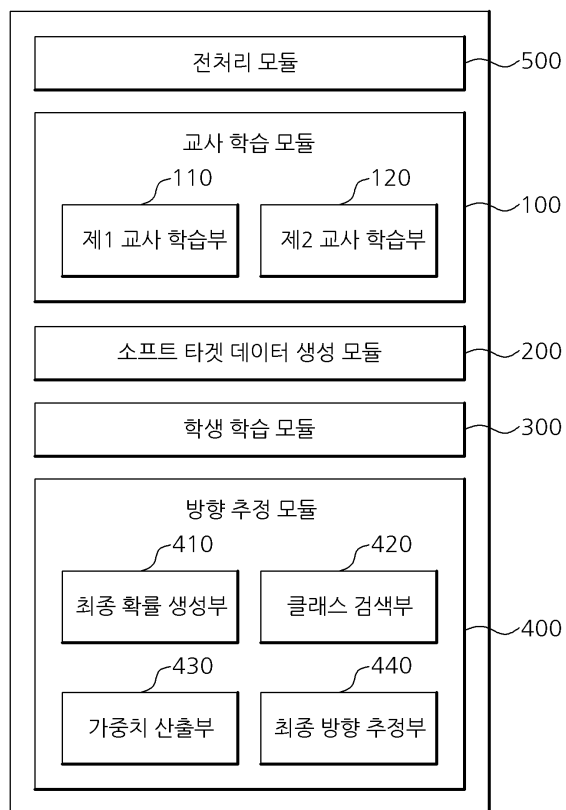


도면7

S400



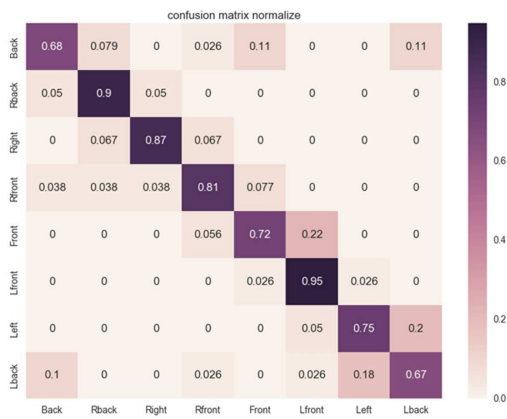
도면8



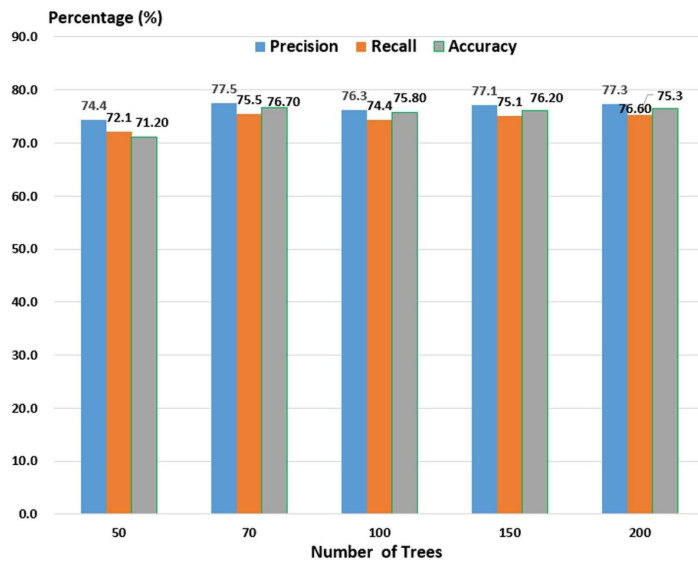
도면9

Methods	AP (%)	AR (%)	AFPR (%)
MoAWG [39]	67.4	65.4	4.9
PLS-RF [40]	66.3	62.3	5.0
MACF [41]	41.1	40.0	8.5
VGG-16 [42]	67.6	68.6	4.3
ResNet101 [23]	73.5	74.6	3.9
Proposed T-Model without handcraft filters	75.4	76.8	3.1
Proposed T-Model	85.6	84.6	2.0
Proposed S-Model	78.3	79.3	2.9

도면10



도면11



도면12

Methods	Accuracy (%)	No. of parameters (M)	No. of operations (M)
MobileNet V2 [43]	60.73	2.2	430
SqueezeNet V1.1 [44]	53.06	0.72	283
Teacher model	85.08	47.2	7564
Proposed student model	78.63	9.5	385

도면13

Methods	Average Orientation Similarity (AOS, %)		
	<i>Easy</i>	<i>Moderate</i>	<i>Hard</i>
DPM-VOC+VP [45]	53.66	39.83	35.73
Mono3D [46]	68.58	58.12	54.94
SubCNN [47]	78.33	66.28	61.37
FRCNN [48]	66.84	52.62	48.72
Proposed student model	81.1	68.61	63.54

도면14

