



**(19) 대한민국특허청(KR)**  
**(12) 공개특허공보(A)**

(11) 공개번호 10-2020-0092848  
(43) 공개일자 2020년08월04일

(51) 국제특허분류(Int. Cl.)  
G06N 3/08 (2006.01) G06N 3/04 (2006.01)  
(52) CPC특허분류  
G06N 3/084 (2013.01)  
G06K 9/00624 (2013.01)  
(21) 출원번호 10-2019-0119513  
(22) 출원일자 2019년09월27일  
심사청구일자 2019년09월27일  
(30) 우선권주장  
16/258,248 2019년01월25일 미국(US)

(71) 출원인  
주식회사 스트라드비전  
경상북도 포항시 남구 지곡로 394, 제5벤처동 304호, 305호, 306호, 307호, 308호(지곡동, 포항테크노파크)  
(72) 발명자  
김계현  
서울특별시 서대문구 서소문로 27, 1004호  
김용중  
경상북도 포항시 남구 청암로 67, 연구4동 4427호  
(뒀면에 계속)  
(74) 대리인  
특허법인 수

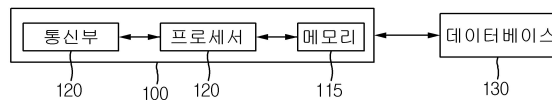
전체 청구항 수 : 총 28 항

(54) 발명의 명칭 객체의 스케일에 따라 모드 전환이 가능한 CNN 기반의 감시용 객체 검출기의 학습 방법 및 학습 장치, 이를 이용한 테스트 방법 및 테스트 장치

**(57) 요약**

본 발명은 객체의 스케일에 따라 모드 전환이 가능한 CNN 기반의 감시용 객체 검출기의 학습 방법에 관한 것으로, 학습 장치가, (a) 컨벌루션 레이어로 하여금 이미지에 대하여 컨벌루션 연산을 적용하도록 하여 특징 맵을 출력하도록 하고, RPN(Region Proposal Network)으로 하여금 이미지 내의 ROI(Region Of Interest)를 출력하도록 하는 단계; (b) 풀링 레이어로 하여금 특징 맵 상의 각각의 ROI영역을 해당 스케일 별로 풀링하도록 하여, 제1 FC 레이어로 하여금 제2 특징 벡터를 출력하도록 하며, 제2 FC 레이어로 하여금 클래스 정보와 리그레션 정보를 출력하도록 하는 단계; 및 (c) 로스 레이어로 하여금 클래스 정보, 리그레션 정보 및 이에 대응하는 GT(Ground Truth)를 참조하여 클래스 로스와 리그레션 로스를 생성하도록 하는 단계를 포함하는 학습 방법을 제공한다.

**대표도 - 도2**



(52) CPC특허분류

*G06N 3/04* (2013.01)

(72) 발명자

**김인수**

경상북도 포항시 남구 청암로 67, 연구4동 4427호

**김학경**

경상북도 포항시 남구 청암로 67, 연구4동 4427호

**남운현**

경상북도 포항시 남구 연일읍 유강길9번길 57, 20  
3동 803호

**부석훈**

경기도 안양시 만안구 만안로55번길 20-9, B02호

**성명철**

경상북도 포항시 북구 장량로174번길 13

**여동훈**

경상북도 포항시 남구 청암로 67, 연구4동 4427호

**유우주**

경상북도 포항시 남구 청암로 67, 연구4동 4427호

**장태웅**

서울특별시 강남구 언주로113길 18-5

**정경중**

경상북도 포항시 남구 지곡로 294, 232동 501호

**제홍모**

경상북도 포항시 남구 지곡로 20, 5동 1805호

**조호진**

경상북도 포항시 남구 청암로 67, 연구4동 4427호

## 명세서

### 청구범위

#### 청구항 1

CNN(Convolutional Neural Network) 기반의 객체 검출기를 학습하는 방법에 있어서,

(a) 적어도 하나의 객체를 포함하는 트레이닝 이미지가 입력되면, 학습 장치가, 적어도 하나의 컨벌루션 레이어(Convolutional Layer)로 하여금 상기 트레이닝 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 특징 맵을 출력하도록 하고, RPN(Region Proposal Network)으로 하여금 상기 특징 맵을 이용하여 상기 트레이닝 이미지 내의 k개의 ROI(Region Of Interest)를 출력하도록 하는 단계;

(b) 상기 학습 장치가, (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 상기 특징 맵 상에서 상기 k개의 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 k개의 ROI 각각에 대응되는 k개의 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 제1 FC 레이어(Fully Connected Layer) 각각으로 하여금 해당 스케일 별로 k개의 상기 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 k개의 제2 특징 벡터를 각각 출력하도록 하며, (iii) 적어도 하나의 제2 FC 레이어로 하여금 해당 스케일 별로 k개의 상기 제2 특징 벡터를 각각 이용하여 상기 k개의 ROI 각각에 대응되는 각각의 클래스(Class) 정보와 각각의 리그레션(Regression) 정보를 출력하도록 하는 단계; 및

(c) 상기 학습 장치가, 제1 로스 레이어(Loss Layer)로 하여금 각각의 상기 클래스 정보, 각각의 상기 리그레션 정보 및 이에 대응하는 제1 GT(Ground Truth)를 참조하여 상기 k개의 ROI에 대응되는 클래스 로스(Class Loss)와 리그레션 로스(Regression Loss)를 생성하도록 함으로써, 상기 클래스 로스와 상기 리그레션 로스를 이용한 백프로파게이션을 통해 상기 제2 FC 레이어, 상기 제1 FC 레이어 및 상기 컨벌루션 레이어를 학습하도록 하는 단계;

를 포함하는 것을 특징으로 하는 학습 방법.

#### 청구항 2

제 1 항에 있어서,

상기 k개의 ROI는 상기 RPN에 의해 기설정된 ROI 후보 중에서 선택된 것이며,

상기 (a) 단계 이후에,

상기 학습 장치는, 제2 로스 레이어로 하여금 상기 ROI 후보 및 그에 대응하는 제2 GT를 참조하여 ROI 로스를 생성하도록 함으로써, 상기 ROI 로스를 이용한 백프로파게이션을 통해 상기 RPN을 학습하도록 하는 것을 특징으로 하는 학습 방법.

#### 청구항 3

제 1 항에 있어서,

상기 제2 FC 레이어는 제2\_1 FC 레이어 및 제2\_2 FC 레이어를 포함하며,

상기 (b) 단계에서,

상기 학습 장치는, (i) 상기 제2\_1 FC 레이어로 하여금 상기 제1 FC 레이어 각각에서 출력되는 해당 스케일 별로 k개의 상기 제2 특징 벡터 각각을 컨캐터네이팅(Concatenating)하도록 하여 제3 특징 벡터를 출력하도록 하며, (ii) 상기 제2\_2 FC 레이어로 하여금 상기 제3 특징 벡터를 이용하여 상기 k개의 ROI에 대응되는 각각의 상기 클래스 정보와 각각의 상기 리그레션 정보를 출력하도록 하는 것을 특징으로 하는 학습 방법.

#### 청구항 4

제 1 항에 있어서,

상기 (c) 단계에서,

상기 학습 장치는, 상기 제1 FC 레이어 각각으로부터 출력된 해당 스케일 별로 k개의 상기 제2 특징 벡터 각각에 대응되는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스의 적어도 일부를 이용하여 상기 제1 FC 레이어 각각을 학습하도록 하는 것을 특징으로 하는 학습 방법.

**청구항 5**

제 4 항에 있어서,

상기 학습 장치는, 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대하여 각각의 가중치를 적용하여 백프로파게이션을 통해 상기 제1 FC 레이어 각각을 학습하되, 각각의 상기 가중치는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대한 각각의 상기 제1 FC 레이어의 기여도에 대한 정보를 참고로 하여 결정되는 것을 특징으로 하는 학습 방법.

**청구항 6**

제 1 항에 있어서,

상기 학습 장치는, 각각의 상기 제1 FC 레이어에서 출력되는 해당 스케일 별로 각각의 상기 제2 특징 벡터의 채널 수가 서로 동일하게 되도록 하는 것을 특징으로 하는 학습 방법.

**청구항 7**

제 1 항에 있어서,

상기 학습 장치는, 각각의 상기 제1 FC 레이어에서 출력되는 상기 제2 특징 벡터의 총 채널 수가 기설정된 수 미만이 되도록 하는 것을 특징으로 하는 학습 방법.

**청구항 8**

CNN(Convolutional Neural Network) 기반의 객체 검출기를 테스트하는 방법에 있어서,

(a) 학습 장치가, (1) 적어도 하나의 컨벌루션 레이어(Convolutional Layer)로 하여금 트레이닝 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 학습용 특징 맵을 출력하도록 하고, RPN(Region Proposal Network)으로 하여금 상기 학습용 특징 맵을 이용하여 상기 트레이닝 이미지 내의 k개의 ROI(Region Of Interest)를 출력하도록 하며, (2) (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 상기 학습용 특징 맵 상에서 상기 k개의 학습용 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 k개의 학습용 ROI 각각에 대응되는 k개의 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 제1 FC 레이어(Fully Connected Layer) 각각으로 하여금 해당 스케일 별로 k개의 상기 학습용 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 k개의 상기 학습용 제2 특징 벡터를 각각 출력하도록 하며, (iii) 적어도 하나의 제2 FC 레이어로 하여금 해당 스케일 별로 k개의 상기 학습용 제2 특징 벡터를 각각 이용하여 상기 k개의 학습용 ROI 각각에 대응되는 각각의 학습용 클래스(Class) 정보와 각각의 학습용 리그레션(Regression) 정보를 출력하도록 하며, (3) 제1 로스 레이어(Loss Layer)로 하여금 각각의 상기 학습용 클래스 정보, 각각의 상기 학습용 리그레션 정보 및 이에 대응하는 제1 GT(Ground Truth)를 참조하여 상기 k개의 학습용 ROI에 대응되는 클래스 로스(Class Loss)와 리그레션 로스(Regression Loss)를 생성하도록 함으로써, 상기 클래스 로스와 상기 리그레션 로스를 이용한 백프로파게이션을 통해 상기 제2 FC 레이어, 상기 제1 FC 레이어 및 상기 컨벌루션 레이어를 학습한 상태에서, 테스트 장치가, (i) 상기 컨벌루션 레이어로 하여금 테스트 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 테스트용 특징 맵을 출력하도록 하고, (ii) 상기 RPN으로 하여금 상기 테스트용 특징 맵을 이용하여 상기 테스트 이미지 내의 m개의 테스트용 ROI를 출력하도록 하는 단계; 및

(b) 상기 테스트 장치가, (i) 서로 다른 풀링 스케일을 가지는 상기 풀링 레이어 각각으로 하여금 해당 스케일 별로 상기 테스트용 특징 맵 상에서 상기 m개의 테스트용 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 m개의 테스트용 ROI 각각에 대응되는 m개의 테스트용 제1 특징 벡터 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 상기 제1 FC 레이어로 하여금 해당 스케일 별로 m개의 상기 테스트용 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 m개의 테스트용 제2 특징 벡터를 각각 출력하도록 하며, (iii) 상기 제2 FC 레이어로 하여금 해당 스케일 별로 m개의 상기 테스트용 제2 특징 벡터를 각각 이용하

여 상기  $m$ 개의 테스트용 ROI 각각에 대응되는 각각의 테스트용 클래스 정보와 각각의 테스트용 리그레션 정보를 출력하도록 하는 단계;

를 포함하는 것을 특징으로 하는 테스트 방법.

### 청구항 9

제 8 항에 있어서,

상기  $k$ 개의 학습용 ROI는 상기 RPN에 의해 기설정된 학습용 ROI 후보 중에서 선택된 것이며,

상기 (1) 프로세스 이후에,

상기 학습 장치는, 제2 로스 레이어로 하여금 상기 학습용 ROI 후보 및 그에 대응하는 제2 GT를 참조하여 ROI 로스를 생성하도록 함으로써, 상기 ROI 로스를 이용한 백프로파게이션을 통해 상기 RPN을 학습하도록 하는 것을 특징으로 하는 테스트 방법.

### 청구항 10

제 8 항에 있어서,

상기 제2 FC 레이어는 제2\_1 FC 레이어 및 제2\_2 FC 레이어를 포함하며,

상기 (b) 단계에서,

상기 테스트 장치는, (i) 상기 제2\_1 FC 레이어로 하여금 상기 제1 FC 레이어 각각에서 출력되는 해당 스케일 별로  $m$ 개의 상기 테스트용 제2 특징 벡터 각각을 컨캐터네이팅(Concatenating)하도록 하여 테스트용 제3 특징 벡터를 출력하도록 하며, (ii) 상기 제2\_2 FC 레이어로 하여금 상기 테스트용 제3 특징 벡터를 이용하여 상기  $m$ 개의 테스트용 ROI에 대응되는 각각의 상기 테스트용 클래스 정보와 각각의 상기 테스트용 리그레션 정보를 출력하도록 하는 것을 특징으로 하는 테스트 방법.

### 청구항 11

제 8 항에 있어서,

상기 (3) 프로세스에서,

상기 학습 장치는, 상기 제1 FC 레이어 각각으로부터 출력된 해당 스케일 별로  $k$ 개의 상기 학습용 제2 특징 벡터 각각에 대응되는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스의 적어도 일부를 이용하여 상기 제1 FC 레이어 각각을 학습하도록 하는 것을 특징으로 하는 테스트 방법.

### 청구항 12

제 11 항에 있어서,

상기 학습 장치는, 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대하여 각각의 가중치를 적용하여 백프로파게이션을 통해 상기 제1 FC 레이어 각각을 학습하되, 각각의 상기 가중치는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대한 각각의 상기 제1 FC 레이어의 기여도에 대한 정보를 참고로 하여 결정되는 것을 특징으로 하는 테스트 방법.

### 청구항 13

제 8 항에 있어서,

상기 테스트 장치는, 각각의 상기 제1 FC 레이어에서 출력되는 해당 스케일 별로 각각의 상기 테스트용 제2 특징 벡터의 채널 수가 서로 동일하게 되도록 하는 것을 특징으로 하는 테스트 방법.

### 청구항 14

제 8 항에 있어서,

상기 테스트 장치는, 각각의 상기 제1 FC 레이어에서 출력되는 상기 테스트용 제2 특징 벡터의 총 채널 수가 기설정된 수 미만인 되도록 하는 것을 특징으로 하는 테스트 방법.

**청구항 15**

CNN(Convolutional Neural Network) 기반의 객체 검출기를 학습하는 학습 장치에 있어서,  
인스트럭션을 저장하는 적어도 하나의 메모리; 및

(I) 적어도 하나의 컨벌루션 레이어(Convolutional Layer)로 하여금 적어도 하나의 객체를 포함하는 트레이닝 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 특징 맵을 출력하도록 하고, RPN(Region Proposal Network)으로 하여금 상기 특징 맵을 이용하여 상기 트레이닝 이미지 내의 k개의 ROI(Region Of Interest)를 출력하도록 하는 프로세스, (II) (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 상기 특징 맵 상에서 상기 k개의 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 k개의 ROI 각각에 대응되는 k개의 제1 특징 벡터 (Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 제1 FC 레이어 (Fully Connected Layer) 각각으로 하여금 해당 스케일 별로 k개의 상기 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 k개의 제2 특징 벡터를 각각 출력하도록 하며, (iii) 적어도 하나의 제2 FC 레이어로 하여금 해당 스케일 별로 k개의 상기 제2 특징 벡터를 각각 이용하여 상기 k개의 ROI 각각에 대응되는 각각의 클래스(Class) 정보와 각각의 리그레션(Regression) 정보를 출력하도록 하는 프로세스 및 (III) 제1 로스 레이어(Loss Layer)로 하여금 각각의 상기 클래스 정보, 각각의 상기 리그레션 정보 및 이에 대응하는 제1 GT(Ground Truth)를 참조하여 상기 k개의 ROI에 대응되는 클래스 로스(Class Loss)와 리그레션 로스(Regression Loss)를 생성하도록 함으로써, 상기 클래스 로스와 상기 리그레션 로스를 이용한 백프로파게이션을 통해 상기 제2 FC 레이어, 상기 제1 FC 레이어 및 상기 컨벌루션 레이어를 학습하도록 하는 프로세스를 수행하기 위한 상기 인스트럭션을 실행하도록 구성된 적어도 하나의 프로세서;

를 포함하는 것을 특징으로 하는 학습 장치.

**청구항 16**

제 15 항에 있어서,

상기 k개의 ROI는 상기 RPN에 의해 기설정된 ROI 후보 중에서 선택된 것이며,

상기 (I) 프로세스 이후에,

상기 프로세서는, 제2 로스 레이어로 하여금 상기 ROI 후보 및 그에 대응하는 제2 GT를 참조하여 ROI 로스를 생성하도록 함으로써, 상기 ROI 로스를 이용한 백프로파게이션을 통해 상기 RPN을 학습하도록 하는 것을 특징으로 하는 학습 장치.

**청구항 17**

제 15 항에 있어서,

상기 제2 FC 레이어는 제2\_1 FC 레이어 및 제2\_2 FC 레이어를 포함하며,

상기 (II) 프로세스에서,

상기 프로세서는, (i) 상기 제2\_1 FC 레이어로 하여금 상기 제1 FC 레이어 각각에서 출력되는 해당 스케일 별로 k개의 상기 제2 특징 벡터 각각을 컨캐터네이팅(Concatenating)하도록 하여 제3 특징 벡터를 출력하도록 하며, (ii) 상기 제2\_2 FC 레이어로 하여금 상기 제3 특징 벡터를 이용하여 상기 k개의 ROI에 대응되는 각각의 상기 클래스 정보와 각각의 상기 리그레션 정보를 출력하도록 하는 것을 특징으로 하는 학습 장치.

**청구항 18**

제 15 항에 있어서,

상기 (III) 프로세스에서,

상기 프로세서는, 상기 제1 FC 레이어 각각으로부터 출력된 해당 스케일 별로 k개의 상기 제2 특징 벡터 각각에 대응되는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스의 적어도 일부를 이용하여 상기 제1 FC 레이어 각각을 학습하도록 하는 것을 특징으로 하는 학습 장치.

**청구항 19**

제 18 항에 있어서,

상기 프로세서는, 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대하여 각각의 가중치를 적용하여 백프로파게이션을 통해 상기 제1 FC 레이어 각각을 학습하되, 각각의 상기 가중치는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대한 각각의 상기 제1 FC 레이어의 기여도에 대한 정보를 참고로 하여 결정되는 것을 특징으로 하는 학습 장치.

**청구항 20**

제 15 항에 있어서,

상기 프로세서는, 각각의 상기 제1 FC 레이어에서 출력되는 해당 스케일 별로 각각의 상기 제2 특징 벡터의 채널 수가 서로 동일하게 되도록 하는 것을 특징으로 하는 학습 장치.

**청구항 21**

제 15 항에 있어서,

상기 프로세서는, 각각의 상기 제1 FC 레이어에서 출력되는 상기 제2 특징 벡터의 총 채널 수가 기설정된 수 미만 되도록 하는 것을 특징으로 하는 학습 장치.

**청구항 22**

CNN(Convolutional Neural Network) 기반의 객체 검출기를 테스트하는 테스트 장치에 있어서, 인스트럭션을 저장하는 적어도 하나의 메모리; 및

학습 장치가, (1) 적어도 하나의 컨벌루션 레이어(Convolutional Layer)로 하여금 트레이닝 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 학습용 특징 맵을 출력하도록 하고, RPN(Region Proposal Network)으로 하여금 상기 학습용 특징 맵을 이용하여 상기 트레이닝 이미지 내의 k개의 ROI(Region Of Interest)를 출력하도록 하며, (2) (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 상기 학습용 특징 맵 상에서 상기 k개의 학습용 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 k개의 학습용 ROI 각각에 대응되는 k개의 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 제1 FC 레이어(Fully Connected Layer) 각각으로 하여금 해당 스케일 별로 k개의 상기 학습용 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 k개의 상기 학습용 제2 특징 벡터를 각각 출력하도록 하며, (iii) 적어도 하나의 제2 FC 레이어로 하여금 해당 스케일 별로 k개의 상기 학습용 제2 특징 벡터를 각각 이용하여 상기 k개의 학습용 ROI 각각에 대응되는 각각의 학습용 클래스(Class) 정보와 각각의 학습용 리그레션(Regression) 정보를 출력하도록 하며, (3) 제1 로스 레이어(Loss Layer)로 하여금 각각의 상기 학습용 클래스 정보, 각각의 상기 학습용 리그레션 정보 및 이에 대응하는 제1 GT(Ground Truth)를 참조하여 상기 k개의 학습용 ROI에 대응되는 클래스 로스(Class Loss)와 리그레션 로스(Regression Loss)를 생성하도록 함으로써, 상기 클래스 로스와 상기 리그레션 로스를 이용한 백프로파게이션을 통해 상기 제2 FC 레이어, 상기 제1 FC 레이어 및 상기 컨벌루션 레이어를 학습한 상태에서, (I) (i) 상기 컨벌루션 레이어로 하여금 테스트 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 테스트용 특징 맵을 출력하도록 하고, (ii) 상기 RPN으로 하여금 상기 테스트용 특징 맵을 이용하여 상기 테스트 이미지 내의 m개의 테스트용 ROI를 출력하도록 하는 프로세스, 및 (II) (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 상기 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 상기 테스트용 특징 맵 상에서 상기 m개의 테스트용 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 m개의 테스트용 ROI 각각에 대응되는 m개의 테스트용 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 상기 제1 FC 레이어(Fully Connected Layer)로 하여금 해당 스케일 별로 m개의 상기 테스트용 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 m개의 테스트용 제2 특징 벡터를 각각 출력하도록 하며, (iii) 상기 제2 FC 레이어로 하여금 해당 스케일 별로 m개의 상기 테스트용 제2 특징 벡터를 각각 이용하여 상기 m개의 테스트용 ROI 각각에 대응되는 각각의 테스트용 클래스(Class) 정보와 각각의 테스트용 리그레션(Regression) 정보를 출력하도록 하는 프로세스를 수행하기 위한 상기 인스트럭션을 실행하도록 구성된 적어도 하나의 프로세서;

를 포함하는 것을 특징으로 하는 테스트 장치.

**청구항 23**

제 22 항에 있어서,

상기 k개의 학습용 ROI는 상기 RPN에 의해 기설정된 학습용 ROI 후보 중에서 선택된 것이며,

상기 (1) 프로세스 이후에,

상기 학습 장치는, 제2 로스 레이어로 하여금 상기 학습용 ROI 후보 및 그에 대응하는 제2 GT를 참조하여 ROI 로스를 생성하도록 함으로써, 상기 ROI 로스를 이용한 백프로파게이션을 통해 상기 RPN을 학습하도록 하는 것을 특징으로 하는 테스트 장치.

**청구항 24**

제 22 항에 있어서,

상기 제2 FC 레이어는 제2\_1 FC 레이어 및 제2\_2 FC 레이어를 포함하며,

상기 (II) 프로세스에서,

상기 프로세서는, (i) 상기 제2\_1 FC 레이어로 하여금 상기 제1 FC 레이어 각각에서 출력되는 해당 스케일 별로 m개의 상기 테스트용 제2 특징 벡터 각각을 컨캐터네이팅(Concatenating)하도록 하여 테스트용 제3 특징 벡터를 출력하도록 하며, (ii) 상기 제2\_2 FC 레이어로 하여금 상기 테스트용 제3 특징 벡터를 이용하여 상기 m개의 테스트용 ROI에 대응되는 각각의 상기 테스트용 클래스 정보와 각각의 상기 테스트용 리그레션 정보를 출력하도록 하는 것을 특징으로 하는 테스트 장치.

**청구항 25**

제22 항에 있어서,

상기 (3) 프로세스에서,

상기 학습 장치는, 상기 제1 FC 레이어 각각으로부터 출력된 해당 스케일 별로 k개의 상기 학습용 제2 특징 벡터 각각에 대응되는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스의 적어도 일부를 이용하여 상기 제1 FC 레이어 각각을 학습하도록 하는 것을 특징으로 하는 테스트 장치.

**청구항 26**

제 25 항에 있어서,

상기 학습 장치는, 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대하여 각각의 가중치를 적용하여 백프로파게이션을 통해 상기 제1 FC 레이어 각각을 학습하되, 각각의 상기 가중치는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대한 각각의 상기 제1 FC 레이어의 기여도에 대한 정보를 참고로 하여 결정되는 것을 특징으로 하는 테스트 장치.

**청구항 27**

제 22 항에 있어서,

상기 프로세서는, 각각의 상기 제1 FC 레이어에서 출력되는 해당 스케일 별로 각각의 상기 테스트용 제2 특징 벡터의 채널 수가 서로 동일하게 되도록 하는 것을 특징으로 하는 테스트 장치.

**청구항 28**

제 22 항에 있어서,

상기 프로세서는, 각각의 상기 제1 FC 레이어에서 출력되는 상기 테스트용 제2 특징 벡터의 총 채널 수가 기설정된 수 미만인 되도록 하는 것을 특징으로 하는 테스트 장치.

**발명의 설명**

**기술분야**

[0001] 본 발명은 객체의 스케일에 따라 모드 전환이 가능한 CNN(Convolutional Neural Network) 기반의 감시용 객체 검출기의 학습 방법 및 학습 장치, 그리고 이를 이용한 테스트 방법 및 테스트 장치에 관한 것으로, 보다 상세하게는, 상기 CNN(Convolutional Neural Network) 기반의 상기 객체 검출기를 학습하는 방법에 있어서, (a) 적어도 하나의 객체를 포함하는 트레이닝 이미지가 입력되면, 학습 장치가, 적어도 하나의 컨벌루션 레이어(Convolutional Layer)로 하여금 상기 트레이닝 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 특징 맵을 출력하도록 하고, RPN(Region Proposal Network)으로 하여금 상기 특징 맵을 이용하여 상기 트레이닝 이미지 내의 k개의 ROI(Region Of Interest)를 출력하도록 하는 단계; (b) 상기 학습 장치가, (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 상기 특징 맵 상에서 상기 k개의 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 k개의 ROI 각각에 대응되는 k개의 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 제1 FC 레이어(Fully Connected Layer) 각각으로 하여금 해당 스케일 별로 k개의 상기 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 k개의 제2 특징 벡터를 각각 출력하도록 하며, (iii) 적어도 하나의 제2 FC 레이어로 하여금 해당 스케일 별로 k개의 상기 제2 특징 벡터를 각각 이용하여 상기 k개의 ROI 각각에 대응되는 각각의 클래스(Class) 정보와 각각의 리그레션(Regression) 정보를 출력하도록 하는 단계; 및 (c) 상기 학습 장치가, 제1 로스 레이어(Loss Layer)로 하여금 각각의 상기 클래스 정보, 각각의 상기 리그레션 정보 및 이에 대응하는 제1 GT(Ground Truth)를 참조하여 상기 k개의 ROI에 대응되는 클래스 로스(Class Loss)와 리그레션 로스(Regression Loss)를 생성하도록 함으로써, 상기 클래스 로스와 상기 리그레션 로스를 이용한 백프로파게이션을 통해 상기 제2 FC 레이어, 상기 제1 FC 레이어 및 상기 컨벌루션 레이어를 학습하는 단계;를 포함하는 학습 방법 및 학습 장치, 그리고, 이를 이용한 테스트 방법 및 테스트 장치에 관한 것이다.

**배경기술**

[0002] 기계 학습(Machine Learning)에서 컨벌루션 뉴럴 네트워크(Convolutional Neural Network; CNN 또는 ConvNet)는 시각적 이미지 분석에 성공적으로 적용되어온 딥 피드포워드 인공 뉴럴 네트워크(Deep, Feed-Forward Artificial Neural Network)의 한 클래스(Class)이다.

[0003] 도 1은 기존의 기술에 따른 CNN을 사용한 학습 프로세스를 나타낸 것이다.

[0004] 도 1은 학습 장치가 예측한 바운딩 박스(Bounding Box)를 이에 대응하는 GT의 바운딩 박스와 비교하는 프로세스를 나타낸 도면이다.

[0005] 도 1을 참조하면, 기존의 학습 장치가 바운딩 박스를 예측하고 이 예측된 바운딩 박스를 GT의 바운딩 박스와 비교하여 적어도 하나의 로스 값을 획득한다. 여기서, 로스 값은 예측된 바운딩 박스와 GT의 바운딩 박스 사이의 차이 값을 의미한다. 예를 들어, 도 1에서 로스 값은  $dx_c, dy_c, dw, dh$  를 포함할 수 있다.

[0006] 먼저, 도 1의 학습 장치는 RGB 이미지를 획득하여 이를 컨벌루션 레이어에 입력시킨다. RGB 이미지가 컨벌루션 레이어를 통과한 후, RGB 이미지의 너비와 높이가 줄어들지만 채널 수는 증가하도록 특징 맵이 생성된다.

[0007] 도 1의 학습 장치는 특징 맵을 RPN(Region Proposal Network)에 입력시켜 프로포잘 박스를 생성하고, 특징 맵 상의 프로포잘 박스에 대응되는 영역에 포함된 픽셀 데이터에 대해 맥스 풀링(Max Pooling) 또는 에버리지 풀링(Average Pooling) 연산 중 어느 하나를 적용하여 풀링된 특징 맵을 생성할 수 있다. 여기서, 맥스 풀링은 각각의 세부 영역마다 이에 대응하는 프로포잘 박스 내의 세부 영역 각각에서 가장 큰 값을 세부 영역 별 대표 값으로 각각 선정하는 방법이고, 에버리지 풀링은 각각의 세부 영역마다 이에 대응하는 프로포잘 박스 내의 세부 영역 각각에 대한 각각의 평균 값을 각각의 대표 값으로서 산출하는 방법이다.

[0008] 다음으로, 도 1의 학습 장치는 풀링된 특징 맵을 FC 레이어(Fully Connected Layer)에 입력한다. 이때, 학습 장치는 FC 레이어로 하여금 분류 연산을 통해 상기 RGB 이미지 상의 객체의 종류가 무엇인지를 확인하도록 할 수 있다. 풀링된 특징 맵은 “특징 벡터(Feature Vector)” 로 칭할 수도 있다.

[0009] 또한, 입력된 RGB 이미지 상의 적어도 하나의 바운딩 박스는 FC 레이어와 로스 값을 통해 예측되며, 상기 로스는 예측된 바운딩 박스를 이에 대응하는 GT(Ground Truth)의 바운딩 박스와 비교하여 획득된다. 여기서, GT의 바운딩 박스는 이미지 상의 객체를 정확하게 포함하는 바운딩 박스로서, 일반적으로는 사람이 직접 생성할 수

있다.

- [0010] 마지막으로, 도 1의 학습 장치는 백프로파게이션을 수행하면서 로스 값을 감소시키기 위해 FC 레이어의 적어도 하나의 파라미터, 및 컨벌루션 레이어의 적어도 하나의 파라미터를 조정할 수 있다. CNN의 파라미터가 조정된 후, 테스트 이미지 상의 새로운 객체에 대응되는 새로운 바운딩 박스가 예측될 수 있다.
- [0011] 다만, 컨벌루션 레이어에서 출력되는 특징 맵 상에서 ROI에 대응되는 영역이 단일 스케일(Single Scale)로 풀링 되면, 단일 스케일로 풀링된 특징 맵만을 이용해서는 정보를 효율적으로 표현하는 것이 제한된다. 그러므로 객체를 보다 정확하게 검출하기 위하여 많은 수의 특징(Feature)을 이용하여야 하며, 그에 따라 객체 검출을 위한 연산 수가 많아질 뿐만 아니라 성능이 저하되는 단점이 있었다.
- [0012] 따라서, 본 발명은 서로 다른 스케일을 가지는 풀링 레이어 각각으로 하여금 각각의 풀링 연산을 수행하도록 하여 연산 수를 줄일 수 있는 학습 방법을 제안하고자 한다.

**발명의 내용**

**해결하려는 과제**

- [0013] 본 발명은 상술한 문제점을 해결하는 것을 목적으로 한다.
- [0014] 또한, 본 발명은 서로 다른 스케일을 가지는 풀링 레이어 각각으로 하여금 각각의 풀링 연산을 수행하도록 하여, 연산량을 줄이는 학습 방법 및 학습 장치, 이를 이용한 테스트 방법 및 테스트 장치를 제공하는 것을 다른 목적으로 한다.

**과제의 해결 수단**

- [0015] 상기한 바와 같은 본 발명의 목적을 달성하고, 후술하는 본 발명의 특징적인 효과를 실현하기 위한, 본 발명의 특징적인 구성은 하기와 같다.
- [0016] 본 발명의 일 태양에 따르면, CNN(Convolutional Neural Network) 기반의 객체 검출기를 학습하는 방법에 있어서, (a) 적어도 하나의 객체를 포함하는 트레이닝 이미지가 입력되면, 학습 장치가, 적어도 하나의 컨벌루션 레이어(Convolutional Layer)로 하여금 상기 트레이닝 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 특징 맵을 출력하도록 하고, RPN(Region Proposal Network)으로 하여금 상기 특징 맵을 이용하여 상기 트레이닝 이미지 내의 k개의 ROI(Region Of Interest)를 출력하도록 하는 단계; (b) 상기 학습 장치가, (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 상기 특징 맵 상에서 상기 k개의 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 k개의 ROI 각각에 대응되는 k개의 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 제1 FC 레이어(Fully Connected Layer) 각각으로 하여금 해당 스케일 별로 k개의 상기 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 k개의 제2 특징 벡터를 각각 출력하도록 하며, (iii) 적어도 하나의 제2 FC 레이어로 하여금 해당 스케일 별로 k개의 상기 제2 특징 벡터를 각각 이용하여 상기 k개의 ROI 각각에 대응되는 각각의 클래스(Class) 정보와 각각의 리그레션(Regression) 정보를 출력하도록 하는 단계; 및 (c) 상기 학습 장치가, 제1 로스 레이어(Loss Layer)로 하여금 각각의 상기 클래스 정보, 각각의 상기 리그레션 정보 및 이에 대응하는 제1 GT(Ground Truth)를 참조하여 상기 k개의 ROI에 대응되는 클래스 로스(Class Loss)와 리그레션 로스(Regression Loss)를 생성하도록 함으로써, 상기 클래스 로스와 상기 리그레션 로스를 이용한 백프로파게이션을 통해 상기 제2 FC 레이어, 상기 제1 FC 레이어 및 상기 컨벌루션 레이어를 학습하는 단계;를 포함하는 것을 특징으로 한다.
- [0017] 일 실시예에서, 상기 k개의 ROI는 상기 RPN에 의해 기설정된 ROI 후보 중에서 선택된 것이며, 상기 (a) 단계 이후에, 상기 학습 장치는 제2 로스 레이어로 하여금 상기 ROI 후보 및 그에 대응하는 제2 GT를 참조하여 ROI 로스를 생성하도록 함으로써, 상기 ROI 로스를 이용한 백프로파게이션을 통해 상기 RPN을 학습하도록 하는 것을 특징으로 한다.
- [0018] 일 실시예에서, 상기 제2 FC 레이어는 제2\_1 FC 레이어 및 제2\_2 FC 레이어를 포함하며, 상기 (b) 단계에서, 상기 학습 장치는, (i) 상기 제2\_1 FC 레이어로 하여금 상기 제1 FC 레이어 각각에서 출력되는 해당 스케일 별로 k개의 상기 제2 특징 벡터 각각을 컨캐터네이팅(Concatenating)하도록 하여 제3 특징 벡터를 출력하도록 하며, (ii) 상기 제2\_2 FC 레이어로 하여금 상기 제3 특징 벡터를 이용하여 상기 k개의 ROI에 대응되는 각각의 상기

클래스 정보와 각각의 상기 리그레션 정보를 출력하도록 하는 것을 특징으로 한다.

- [0019] 일 실시예에서, 상기 (c) 단계에서, 상기 학습 장치는, 상기 제1 FC 레이어 각각으로부터 출력된 해당 스케일 별로 k개의 상기 제2 특징 벡터 각각에 대응되는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스의 적어도 일부를 이용하여 상기 제1 FC 레이어 각각을 학습하도록 하는 것을 특징으로 한다.
- [0020] 일 실시예에서, 상기 학습 장치는, 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대하여 각각의 가중치를 적용하여 백프로파게이션을 통해 상기 제1 FC 레이어 각각을 학습하되, 각각의 상기 가중치는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대한 각각의 상기 제1 FC 레이어의 기여도에 대한 정보를 참고로 하여 결정되는 것을 특징으로 한다.
- [0021] 일 실시예에서, 상기 학습 장치는, 각각의 상기 제1 FC 레이어에서 출력되는 해당 스케일 별로 각각의 상기 제2 특징 벡터의 채널 수가 서로 동일하게 되도록 하는 것을 특징으로 한다.
- [0022] 일 실시예에서, 상기 학습 장치는, 각각의 상기 제1 FC 레이어에서 출력되는 상기 제2 특징 벡터의 총 채널 수가 기설정된 수 미만이 되도록 하는 것을 특징으로 한다.
- [0023] 본 발명의 다른 태양에 따르면, CNN(Convolutional Neural Network) 기반의 객체 검출기를 테스트하는 방법에 있어서, (a) 학습 장치가, (1) 적어도 하나의 컨벌루션 레이어(Convolutional Layer)로 하여금 트레이닝 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 학습용 특징 맵을 출력하도록 하고, RPN(Region Proposal Network)으로 하여금 상기 학습용 특징 맵을 이용하여 상기 트레이닝 이미지 내의 k개의 ROI(Region Of Interest)를 출력하도록 하며, (2) (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 상기 학습용 특징 맵 상에서 상기 k개의 학습용 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 k개의 학습용 ROI 각각에 대응되는 k개의 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 제1 FC 레이어(Fully Connected Layer) 각각으로 하여금 해당 스케일 별로 k개의 상기 학습용 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 k개의 상기 학습용 제2 특징 벡터를 각각 출력하도록 하며, (iii) 적어도 하나의 제2 FC 레이어로 하여금 해당 스케일 별로 k개의 상기 학습용 제2 특징 벡터를 각각 이용하여 상기 k개의 학습용 ROI 각각에 대응되는 각각의 학습용 클래스(Class) 정보와 각각의 학습용 리그레션(Regression) 정보를 출력하도록 하며, (3) 제1 로스 레이어(Loss Layer)로 하여금 각각의 상기 학습용 클래스 정보, 각각의 상기 학습용 리그레션 정보 및 이에 대응하는 제1 GT(Ground Truth)를 참조하여 상기 k개의 학습용 ROI에 대응되는 클래스 로스(Class Loss)와 리그레션 로스(Regression Loss)를 생성하도록 함으로써, 상기 클래스 로스와 상기 리그레션 로스를 이용한 백프로파게이션을 통해 상기 제2 FC 레이어, 상기 제1 FC 레이어 및 상기 컨벌루션 레이어를 학습한 상태에서, 테스트 장치가, (i) 상기 컨벌루션 레이어로 하여금 테스트 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 테스트용 특징 맵을 출력하도록 하고, (ii) 상기 RPN으로 하여금 상기 테스트용 특징 맵을 이용하여 상기 테스트 이미지 내의 m개의 테스트용 ROI를 출력하도록 하는 단계; 및 (b) 상기 테스트 장치가, (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 상기 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 상기 테스트용 특징 맵 상에서 상기 m개의 테스트용 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 m개의 테스트용 ROI 각각에 대응되는 m개의 테스트용 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 상기 제1 FC 레이어(Fully Connected Layer)로 하여금 해당 스케일 별로 m개의 상기 테스트용 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 m개의 테스트용 제2 특징 벡터를 각각 출력하도록 하며, (iii) 상기 제2 FC 레이어로 하여금 해당 스케일 별로 m개의 상기 테스트용 제2 특징 벡터를 각각 이용하여 상기 m개의 테스트용 ROI 각각에 대응되는 각각의 테스트용 클래스(Class) 정보와 각각의 테스트용 리그레션(Regression) 정보를 출력하도록 하는 단계;를 포함하는 것을 특징으로 한다.
- [0024] 일 실시예에서, 상기 k개의 학습용 ROI는 상기 RPN에 의해 기설정된 학습용 ROI 후보 중에서 선택된 것이며, 상기 (1) 프로세스 이후에, 상기 학습 장치는, 제2 로스 레이어로 하여금 상기 학습용 ROI 후보 및 그에 대응하는 제2 GT를 참조하여 ROI 로스를 생성하도록 함으로써, 상기 ROI 로스를 이용한 백프로파게이션을 통해 상기 RPN을 학습하도록 하는 것을 특징으로 한다.
- [0025] 일 실시예에서, 상기 제2 FC 레이어는 제2\_1 FC 레이어 및 제2\_2 FC 레이어를 포함하며, 상기 (b) 단계에서, 상기 테스트 장치는, (i) 상기 제2\_1 FC 레이어로 하여금 상기 제1 FC 레이어 각각에서 출력되는 해당 스케일 별로 m개의 상기 테스트용 제2 특징 벡터 각각을 컨캐테이팅(Concatenating)하도록 하여 테스트용 제3 특징 벡터를 출력하도록 하며, (ii) 상기 제2\_2 FC 레이어로 하여금 상기 테스트용 제3 특징 벡터를 이용하여 상기 m개

의 테스트용 ROI에 대응되는 각각의 상기 테스트용 클래스 정보와 각각의 상기 테스트용 리그레션 정보를 출력하도록 하는 것을 특징으로 한다.

- [0026] 일 실시예에서, 상기 (3) 프로세스에서, 상기 학습 장치는, 상기 제1 FC 레이어 각각으로부터 출력된 해당 스케일 별로 k개의 상기 학습용 제2 특징 벡터 각각에 대응되는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스의 적어도 일부를 이용하여 상기 제1 FC 레이어 각각을 학습하도록 하는 것을 특징으로 한다.
- [0027] 일 실시예에서, 상기 학습 장치는, 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대하여 각각의 가중치를 적용하여 백프로파게이션을 통해 상기 제1 FC 레이어 각각을 학습하되, 각각의 상기 가중치는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대한 각각의 상기 제1 FC 레이어의 기여도에 대한 정보를 참고로 하여 결정되는 것을 특징으로 한다.
- [0028] 일 실시예에서, 상기 테스트 장치는, 각각의 상기 제1 FC 레이어에서 출력되는 해당 스케일 별로 각각의 상기 테스트용 제2 특징 벡터의 채널 수가 서로 동일하게 되도록 하는 것을 특징으로 한다.
- [0029] 일 실시예에서, 상기 테스트 장치는, 각각의 상기 제1 FC 레이어에서 출력되는 상기 테스트용 제2 특징 벡터의 총 채널 수가 기설정된 수 미만이 되도록 하는 것을 특징으로 한다.
- [0030] 본 발명의 또 다른 태양에 따르면, CNN(Convolutional Neural Network) 기반의 객체 검출기를 학습하는 학습 장치에 있어서, 인스트럭션을 저장하는 적어도 하나의 메모리; 및 (I) 적어도 하나의 컨벌루션 레이어(Convolutional Layer)로 하여금 적어도 하나의 객체를 포함하는 트레이닝 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 특징 맵을 출력하도록 하고, RPN(Region Proposal Network)으로 하여금 상기 특징 맵을 이용하여 상기 트레이닝 이미지 내의 k개의 ROI(Region Of Interest)를 출력하도록 하는 프로세스, (II) (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 상기 특징 맵 상에서 상기 k개의 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 k개의 ROI 각각에 대응되는 k개의 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 제1 FC 레이어(Fully Connected Layer) 각각으로 하여금 해당 스케일 별로 k개의 상기 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 k개의 제2 특징 벡터를 각각 출력하도록 하며, (iii) 적어도 하나의 제2 FC 레이어로 하여금 해당 스케일 별로 k개의 상기 제2 특징 벡터를 각각 이용하여 상기 k개의 ROI 각각에 대응되는 각각의 클래스(Class) 정보와 각각의 리그레션(Regression) 정보를 출력하도록 하는 프로세스 및 (III) 제1 로스 레이어(Loss Layer)로 하여금 각각의 상기 클래스 정보, 각각의 상기 리그레션 정보 및 이에 대응하는 제1 GT(Ground Truth)를 참조하여 상기 k개의 ROI에 대응되는 클래스 로스(Class Loss)와 리그레션 로스(Regression Loss)를 생성하도록 함으로써, 상기 클래스 로스와 상기 리그레션 로스를 이용한 백프로파게이션을 통해 상기 제2 FC 레이어, 상기 제1 FC 레이어 및 상기 컨벌루션 레이어를 학습하는 프로세스를 수행하기 위한 상기 인스트럭션을 실행하도록 구성된 적어도 하나의 프로세서;를 포함하는 것을 특징으로 한다.
- [0031] 일 실시예에서, 상기 k개의 ROI는 상기 RPN에 의해 기설정된 ROI 후보 중에서 선택된 것이며, 상기 (I) 프로세스 이후에, 상기 프로세서는, 제2 로스 레이어로 하여금 상기 ROI 후보 및 그에 대응하는 제2 GT를 참조하여 ROI 로스를 생성하도록 함으로써, 상기 ROI 로스를 이용한 백프로파게이션을 통해 상기 RPN을 학습하도록 하는 것을 특징으로 한다.
- [0032] 일 실시예에서, 상기 제2 FC 레이어는 제2\_1 FC 레이어 및 제2\_2 FC 레이어를 포함하며, 상기 (II) 프로세스에서, 상기 프로세서는, (i) 상기 제2\_1 FC 레이어로 하여금 상기 제1 FC 레이어 각각에서 출력되는 해당 스케일 별로 k개의 상기 제2 특징 벡터 각각을 컨캐터네이팅(Concatenating)하도록 하여 제3 특징 벡터를 출력하도록 하며, (ii) 상기 제2\_2 FC 레이어로 하여금 상기 제3 특징 벡터를 이용하여 상기 k개의 ROI에 대응되는 각각의 상기 클래스 정보와 각각의 상기 리그레션 정보를 출력하도록 하는 것을 특징으로 한다.
- [0033] 일 실시예에서, 상기 (III) 프로세스에서, 상기 프로세서는, 상기 제1 FC 레이어 각각으로부터 출력된 해당 스케일 별로 k개의 상기 제2 특징 벡터 각각에 대응되는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스의 적어도 일부를 이용하여 상기 제1 FC 레이어 각각을 학습하도록 하는 것을 특징으로 한다.
- [0034] 일 실시예에서, 상기 프로세서는, 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대하여 각각의 가중치를 적용하여 백프로파게이션을 통해 상기 제1 FC 레이어 각각을 학습하되, 각각의 상기 가중치는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대한 각각의 상기 제1 FC 레이어의 기여도에 대한 정보를 참고로 하여 결정되는 것을 특징으로 한다.

- [0035] 일 실시예에서, 상기 프로세서는, 각각의 상기 제1 FC 레이어에서 출력되는 해당 스케일 별로 각각의 상기 제2 특징 벡터의 채널 수가 서로 동일하게 되도록 하는 것을 특징으로 한다.
- [0036] 일 실시예에서, 상기 프로세서는, 각각의 상기 제1 FC 레이어에서 출력되는 상기 제2 특징 벡터의 총 채널 수가 기설정된 수 미만이 되도록 하는 것을 특징으로 한다.
- [0037] 본 발명의 또 다른 태양에 따르면, CNN(Convolutional Neural Network) 기반의 객체 검출기를 테스트하는 테스트 장치에 있어서, 인스트럭션을 저장하는 적어도 하나의 메모리; 및 학습 장치가, (1) 적어도 하나의 컨벌루션 레이어(Convolutional Layer)로 하여금 트레이닝 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 학습용 특징 맵을 출력하도록 하고, RPN(Region Proposal Network)으로 하여금 상기 학습용 특징 맵을 이용하여 상기 트레이닝 이미지 내의 k개의 ROI(Region Of Interest)를 출력하도록 하며, (2) (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 상기 학습용 특징 맵 상에서 상기 k개의 학습용 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 k개의 학습용 ROI 각각에 대응되는 k개의 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 제1 FC 레이어(Fully Connected Layer) 각각으로 하여금 해당 스케일 별로 k개의 상기 학습용 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 k개의 상기 학습용 제2 특징 벡터를 각각 출력하도록 하며, (iii) 적어도 하나의 제2 FC 레이어로 하여금 해당 스케일 별로 k개의 상기 학습용 제2 특징 벡터를 각각 이용하여 상기 k개의 학습용 ROI 각각에 대응되는 각각의 학습용 클래스(Class) 정보와 각각의 학습용 리그레션(Regression) 정보를 출력하도록 하며, (3) 제1 로스 레이어(Loss Layer)로 하여금 각각의 상기 학습용 클래스 정보, 각각의 상기 학습용 리그레션 정보 및 이에 대응하는 제1 GT(Ground Truth)를 참조하여 상기 k개의 학습용 ROI에 대응되는 클래스 로스(Class Loss)와 리그레션 로스(Regression Loss)를 생성하도록 함으로써, 상기 클래스 로스와 상기 리그레션 로스를 이용한 백프로파게이션을 통해 상기 제2 FC 레이어, 상기 제1 FC 레이어 및 상기 컨벌루션 레이어를 학습한 상태에서, (I) (i) 상기 컨벌루션 레이어로 하여금 테스트 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 테스트용 특징 맵을 출력하도록 하고, (ii) 상기 RPN으로 하여금 상기 테스트용 특징 맵을 이용하여 상기 테스트 이미지 내의 m개의 테스트용 ROI를 출력하도록 하는 프로세스 및 (II) (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 상기 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 상기 테스트용 특징 맵 상에서 상기 m개의 테스트용 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 상기 m개의 테스트용 ROI 각각에 대응되는 m개의 테스트용 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 상기 풀링 레이어에 대응되는 각각의 상기 제1 FC 레이어(Fully Connected Layer)로 하여금 해당 스케일 별로 m개의 상기 테스트용 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 m개의 테스트용 제2 특징 벡터를 각각 출력하도록 하며, (iii) 상기 제2 FC 레이어로 하여금 해당 스케일 별로 m개의 상기 테스트용 제2 특징 벡터를 각각 이용하여 상기 m개의 테스트용 ROI 각각에 대응되는 각각의 테스트용 클래스(Class) 정보와 각각의 테스트용 리그레션(Regression) 정보를 출력하도록 하는 프로세스를 수행하기 위한 상기 인스트럭션을 실행하도록 구성된 적어도 하나의 프로세서;를 포함하는 것을 특징으로 한다.
- [0038] 일 실시예에서, 상기 k개의 학습용 ROI는 상기 RPN에 의해 기설정된 학습용 ROI 후보 중에서 선택된 것이며, 상기 (1) 프로세스 이후에, 상기 학습 장치는, 제2 로스 레이어로 하여금 상기 학습용 ROI 후보 및 그에 대응하는 제2 GT를 참조하여 ROI 로스를 생성하도록 함으로써, 상기 ROI 로스를 이용한 백프로파게이션을 통해 상기 RPN을 학습하도록 하는 것을 특징으로 한다.
- [0039] 일 실시예에서, 상기 제2 FC 레이어는 제2\_1 FC 레이어 및 제2\_2 FC 레이어를 포함하며, 상기 (II) 프로세스에서, 상기 프로세서는, (i) 상기 제2\_1 FC 레이어로 하여금 상기 제1 FC 레이어 각각에서 출력되는 해당 스케일 별로 m개의 상기 테스트용 제2 특징 벡터 각각을 컨캐터네이팅(Concatenating)하도록 하여 테스트용 제3 특징 벡터를 출력하도록 하며, (ii) 상기 제2\_2 FC 레이어로 하여금 상기 테스트용 제3 특징 벡터를 이용하여 상기 m개의 테스트용 ROI에 대응되는 각각의 상기 테스트용 클래스 정보와 각각의 상기 테스트용 리그레션 정보를 출력하도록 하는 것을 특징으로 한다.
- [0040] 일 실시예에서, 상기 (3) 프로세스에서, 상기 학습 장치는, 상기 제1 FC 레이어 각각으로부터 출력된 해당 스케일 별로 k개의 상기 학습용 제2 특징 벡터 각각에 대응되는 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스의 적어도 일부를 이용하여 상기 제1 FC 레이어 각각을 학습하도록 하는 것을 특징으로 한다.
- [0041] 일 실시예에서, 상기 학습 장치는, 각각의 상기 클래스 로스와 각각의 상기 리그레션 로스에 대하여 각각의 가중치를 적용하여 백프로파게이션을 통해 상기 제1 FC 레이어 각각을 학습하되, 각각의 상기 가중치는 각각의 상

기 클래스 로스와 각각의 상기 리그레션 로스에 대한 각각의 상기 제1 FC 레이어의 기여도에 대한 정보를 참고로 하여 결정되는 것을 특징으로 한다.

[0042] 일 실시예에서, 상기 프로세서는, 각각의 상기 제1 FC 레이어에서 출력되는 해당 스케일 별로 각각의 상기 테스트용 제2 특징 벡터의 채널 수가 서로 동일하게 되도록 하는 것을 특징으로 한다.

[0043] 일 실시예에서, 상기 프로세서는, 각각의 상기 제1 FC 레이어에서 출력되는 상기 테스트용 제2 특징 벡터의 총 채널 수가 기설정된 수 미만이 되도록 하는 것을 특징으로 한다.

**발명의 효과**

[0044] 본 발명은, 서로 다른 풀링 스케일을 가지는 풀링 레이어 각각으로 하여금 특징 맵 상의 ROI에 대응되는 영역을 해당 스케일 별로 풀링하도록 함으로써 다양한 스케일에 따라 풀링된 특징 맵 상의 더 많은 특징을 사용할 수 있는 효과가 있다.

[0045] 본 발명은, 더 적은 수의 채널을 사용해 연산을 수행하므로 컴퓨팅 연산 성능을 향상시킬 수 있는 다른 효과가 있다.

**도면의 간단한 설명**

[0046] 본 발명의 실시예의 설명에 이용되기 위하여 첨부된 아래 도면들은 본 발명의 실시예들 중 단지 일부일 뿐이며, 본 발명이 속한 기술분야에서 통상의 지식을 가진 자(이하 “통상의 기술자”)에게 있어서는 발명적 작업이 이루어짐 없이 이 도면들에 기초하여 다른 도면들이 얻어질 수 있다.

도 1은 기존의 기술에 따른 학습 과정에서 학습 장치가 바운딩 박스를 예측하고 이 예측된 바운딩 박스를 이에 대응하는 GT의 바운딩 박스와 비교하는 프로세스를 나타낸 도면이다.

도 2는 본 발명의 일 실시예에 따른 CNN 기반의 객체 검출기를 학습하는 학습 장치의 블록도이다.

도 3은 본 발명의 일 실시예에 따른 CNN 기반의 객체 검출기를 학습하는 프로세스를 나타내는 도면이다.

도 4는 본 발명의 일 실시예에 따른 CNN 기반의 객체 검출기를 테스트하는 테스트 장치의 블록도이다.

도 5는 본 발명의 일 실시예에 따른 CNN 기반의 객체 검출기를 테스트하는 프로세스를 나타내는 도면이다.

**발명을 실시하기 위한 구체적인 내용**

[0047] 후술하는 본 발명에 대한 상세한 설명은, 본 발명이 실시될 수 있는 특정 실시예를 예시로서 도시하는 첨부 도면을 참조한다. 이들 실시예는 당업자가 본 발명을 실시할 수 있기에 충분하도록 상세히 설명된다. 본 발명의 다양한 실시예는 서로 다르지만 상호 배타적일 필요는 없음이 이해되어야 한다. 예를 들어, 여기에 기재되어 있는 특정 형상, 구조 및 특성은 일 실시예에 관련하여 본 발명의 정신 및 범위를 벗어나지 않으면서 다른 실시예로 구현될 수 있다. 또한, 각각의 개시된 실시예 내의 개별 구성요소의 위치 또는 배치는 본 발명의 정신 및 범위를 벗어나지 않으면서 변경될 수 있음이 이해되어야 한다. 따라서, 후술하는 상세한 설명은 한정적인 의미로서 취하려는 것이 아니며, 본 발명의 범위는, 적절하게 설명된다면, 그 청구항들이 주장하는 것과 균등한 모든 범위와 더불어 첨부된 청구항에 의해서만 한정된다. 도면에서 유사한 참조부호는 여러 측면에 걸쳐서 동일하거나 유사한 기능을 지칭한다.

[0048] 또한, 본 발명의 상세한 설명 및 청구항들에 걸쳐, "포함하다"라는 단어 및 그것의 변형은 다른 기술적 특징들, 부가물들, 구성요소들 또는 단계들을 제외하는 것으로 의도된 것이 아니다. 통상의 기술자에게 본 발명의 다른 목적들, 장점들 및 특성들이 일부는 본 설명서로부터, 그리고 일부는 본 발명의 실시로부터 드러날 것이다. 아래의 예시 및 도면은 실례로서 제공되며, 본 발명을 한정하는 것으로 의도된 것이 아니다.

[0049] 본 발명에서 언급하는 각종 이미지는 포장 또는 비포장 도로 관련 이미지를 포함할 수 있으며, 이 경우 도로 환경에서 등장할 수 있는 물체(가령, 자동차, 사람, 동물, 식물, 물건, 건물, 비행기나 드론과 같은 비행체, 기타 장애물)를 상정할 수 있을 것이나, 반드시 이에 한정되는 것은 아니며, 본 발명에서 언급하는 각종 이미지는 도로와 상관 없는 이미지(가령, 비포장도로, 골목길, 공터, 바다, 호수, 강, 산, 숲, 사막, 하늘, 실내와 관련된 이미지)일 수도 있으며, 이 경우, 비포장도로, 골목길, 공터, 바다, 호수, 강, 산, 숲, 사막, 하늘, 실내 환경에서 등장할 수 있는 물체(가령, 자동차, 사람, 동물, 식물, 물건, 건물, 비행기나 드론과 같은 비행체, 기타 장애물)를 상정할 수 있을 것이나, 반드시 이에 한정되는 것은 아니다.

- [0050] 이하, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명을 용이하게 실시할 수 있도록 하기 위하여, 본 발명의 바람직한 실시예들에 관하여 첨부된 도면을 참조하여 상세히 설명하기로 한다.
- [0051] 도 2는 본 발명의 일 실시예에 따른 학습 장치(100)의 블록도이다.
- [0052] 도 2에 도시된 바와 같이, 본 발명의 일 실시예에 따른 학습 장치(100)는 통신부(110)와 프로세서(120)를 포함할 수 있다. 경우에 따라 데이터베이스(130)를 더 포함할 수 있다. 또한, 학습 장치는 다음의 프로세스를 수행하기 위한 컴퓨터 판독 가능한 인스트럭션(Instruction)을 저장할 수 있는 메모리(115)를 더 포함할 수 있다. 일 실시예에 따르면, 프로세서, 메모리, 매체 등은 통합 프로세서(Integrated Processor)로 통합될 수 있다.
- [0053] 여기서, 본 발명의 학습 장치(100)는 디지털 기기로서, 프로세서를 탑재하여 다양한 연산 수행 능력을 갖춘 디지털 기기라면 얼마든지 본 발명에 따른 학습 장치(100)로서 채택될 수 있다.
- [0054] 통신부(110)는 적어도 하나의 객체를 포함하는 트레이닝 이미지를 획득할 수 있다. 이때, 트레이닝 이미지는 데이터베이스(130)로부터 획득될 수 있다.
- [0055] 프로세서(120)는 (a) 적어도 하나의 컨볼루션 레이어(Convolutional Layer)로 하여금 트레이닝 이미지에 대하여 컨볼루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 특징 맵을 출력하도록 하고, RPN(Region Proposal Network)으로 하여금 특징 맵을 이용하여 트레이닝 이미지 내의 k개의 ROI(Region Of Interest)를 출력하도록 하는 프로세스; (b) (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 특징 맵 상에서 k개의 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 k개의 ROI 각각에 대응되는 k개의 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 풀링 레이어에 대응되는 각각의 제1 FC 레이어(Fully Connected Layer) 각각으로 하여금 해당 스케일 별로 k개의 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 k개의 제2 특징 벡터를 각각 출력하도록 하며, (iii) 적어도 하나의 제2 FC 레이어로 하여금 해당 스케일 별로 k개의 제2 특징 벡터를 각각 이용하여 k개의 ROI 각각에 대응되는 각각의 클래스(Class) 정보와 각각의 리그레션(Regression) 정보를 출력하도록 하는 프로세스; 및 (c) 제1 로스 레이어(Loss Layer)로 하여금 각각의 클래스 정보, 각각의 리그레션 정보 및 이에 대응하는 제1 GT(Ground Truth)를 참조하여 k개의 ROI에 대응되는 클래스 로스(Class Loss)와 리그레션 로스(Regression Loss)를 생성하도록 함으로써, 클래스 로스와 리그레션 로스를 이용한 백프로파게이션을 통해 제2 FC 레이어, 제1 FC 레이어 및 컨볼루션 레이어를 학습하는 프로세스;를 수행할 수 있다. 이에 대해서는 뒤에서 상세한 설명을 통해 자세하게 알아보기로 한다.
- [0056] 여기서, 데이터베이스(130)는 학습 장치(100)의 통신부(110)에 의해 접근할 수 있으며, GT(예를 들어, 객체가 자동차인지, 보행자인지 등을 나타내는 클래스 정보) 및 바운딩 박스의 위치에 대한 정보를 나타내는 리그레션 정보 등이 기록될 수 있다.
- [0057] 이하에, 이와 같이 구성된 본 발명의 일 실시예에 따른 학습 장치(100)를 이용하여 서로 다른 스케일을 갖는 풀링 레이어 각각을 사용하여 CNN 기반의 객체 검출기를 학습하는 방법도 도 3을 참조하여 설명하면 다음과 같다.
- [0058] 먼저, 적어도 하나의 객체를 포함하는 트레이닝 이미지가 입력되면, 학습 장치(100)는, 컨볼루션 레이어(10)로 하여금 트레이닝 이미지에 대하여 컨볼루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 특징 맵을 출력하도록 한다.
- [0059] 그리고, 학습 장치(100)는 RPN(20)으로 하여금, 특징 맵을 이용하여 트레이닝 이미지 내의 k개의 ROI를 출력하도록 한다. 여기서 k개의 ROI는 기설정된 ROI 후보 중에서 선택된 것이다.
- [0060] 다음으로, 학습 장치(100)는 서로 다른 스케일을 가지는 풀링 레이어(30) 각각으로 하여금 특징 맵 상에서 RPN(20)으로부터 출력된 k개의 ROI 각각에 대응되는 영역 각각을 해당 스케일 별로 풀링하여 해당 스케일 별로 k개의 ROI 각각에 대응되는 k개의 제1 특징 벡터 각각을 출력하도록 한다. 이때, 도 3의 예에서는 서로 다른 풀링 스케일을 가지는 3개의 풀링 레이어(30) 각각이 도시되었으나, 본 발명이 이에 한정되는 것은 아니다. 또한, 풀링 스케일이 큰 제1 특징 벡터는 풀링 스케일이 작은 제1 특징 벡터에 비하여 더 많은 정보를 가질 수 있다.
- [0061] 그리고, 각각의 풀링 레이어(30) 각각으로부터 각각의 스케일 별로 k개의 ROI에 대응되는 k개의 제1 특징 벡터 각각이 출력되면, 학습 장치(100)는 각각의 풀링 레이어(30) 각각에 대응되는 제1 FC 레이어(40) 각각으로 하여금 해당 스케일 별로 k개의 제1 특징 벡터 각각을 이용하여 k개의 해당 스케일 별 제2 특징 벡터 각각을 출력하도록 한다. 도 3의 예에서는 제1 FC 레이어(40)의 개수가 3개로 나타나 있지만, 이는 풀링 레이어(30)의 개수가 3개인 것에 기인한 것이며, 본 발명이 이에 한정되는 것은 아니다.

- [0062] 또한, 제1 FC 레이어(40) 각각에서 출력되는 해당 스케일 별 제2 특징 벡터 각각의 채널 수 각각은 서로 동일할 수 있다. 일 예로, 도 3에 도시된 바와 같이, 제1 FC 레이어(40) 각각에서 출력되는 제2 특징 벡터의 채널 수는 128일 수 있으나, 이에 한정되는 것은 아니다.
- [0063] 또한, 제1 FC 레이어(40)에서 출력되는 제2 특징 벡터의 총 채널 수가 기설정된 개수 미만이 되도록 할 수 있다. 일 예로, 도 3에 도시된 바와 같이, 제1 FC 레이어(40)에서 출력되는 제2 특징 벡터의 총 채널 수는 384(128+128+128)일 수 있으나, 이에 한정되는 것은 아니다. 이 경우, 일반적으로 4,096 개의 채널을 가지는 특징 벡터를 연산하는 종래 방법에 비하여, 본 발명에 따른 연산을 수행하는 채널 수는 384개로 적어지므로 그에 따라 학습 장치(100)의 연산 능력을 상승시킬 수 있게 된다.
- [0064] 이후, 학습 장치(100)는 k개의 제2 특징 벡터 각각을 제2 FC 레이어(50)에 입력함으로써, 제2 FC 레이어(50)로 하여금 k개의 해당 스케일 별 제2 특징 벡터 각각을 이용하여 k개의 ROI 각각에 대응되는 각각의 클래스 정보와 각각의 리그레션 정보를 출력하도록 한다.
- [0065] 이때, 제2 FC 레이어(50)는 제2\_1 FC 레이어(51) 및 제2\_2 FC 레이어(55)를 포함할 수 있다. 제2\_1 FC 레이어(51)는 제1 FC 레이어(40) 각각에서 출력되는 제2 특징 벡터 각각을 컨캐터네이팅(Concatenating)하여 제3 특징 벡터를 출력할 수 있다. 제2\_2 FC 레이어(55)는 제3 특징 벡터를 이용하여 k개의 ROI에 대응되는 각각의 클래스 정보와 각각의 리그레션 정보를 출력할 수 있다.
- [0066] 다음으로, 학습 장치(100)는 제1 로스 레이어(60)로 하여금 각각의 클래스 정보, 각각의 리그레션 정보 및 이에 대응되는 제1 GT를 참조하여 k개의 ROI에 대응되는 클래스 로스와 리그레션 로스를 생성하도록 하며, 클래스 로스와 리그레션 로스를 이용한 백프로파게이션을 통해 제2 FC 레이어, 제1 FC 레이어 및 컨벌루션 레이어를 학습할 수 있다.
- [0067] 이때, 학습 장치(100)는 제1 FC 레이어(40) 각각으로부터 출력된 k개의 해당 스케일 별 제2 특징 벡터 각각에 대응되는 각각의 클래스 로스와 각각의 리그레션 로스를 이용하여 각각의 제1 FC 레이어(40)를 학습할 수 있다. 바꾸어 말하면, 학습 장치(100)는 각각의 클래스 로스와 각각의 리그레션 로스에 대하여 각각의 가중치를 적용하여 백프로파게이션을 통해 제1 FC 레이어(40) 각각을 학습할 수 있으며, 각각의 가중치는 각각의 클래스 로스와 각각의 리그레션 로스에 대한 각각의 제1 FC 레이어(40)의 기여도에 대한 정보를 참고로 하여 결정된다.
- [0068] 또한, 학습 장치(100)는 제2 로스 레이어(21)로 하여금 ROI 후보 및 이에 대응하는 제2 GT를 참조하여 ROI 로스를 획득하도록 하고, 획득된 ROI 로스를 이용한 백프로파게이션을 통해 RPN을 학습할 수 있다. 이러한 RPN의 학습은 ROI 후보군이 출력되면 언제든지 가능하다. 또한, ROI 로스는 ROI 클래스 로스와 ROI 리그레션 로스를 포함할 수 있다.
- [0069] 도 4는 본 발명의 일 실시예에 따른 CNN 기반의 객체 검출기를 테스트하는 테스트 장치(400)를 개략적으로 도시한 블록도이다.
- [0070] 도 4에 도시된 바와 같이, 본 발명의 일 실시예에 따른 테스트 장치(400)는 통신부(401)와 프로세서(402)를 포함할 수 있다. 도면에서는 테스트 장치(400)와 학습 장치(100)를 구별하여 표현하였으나, 테스트 장치(400)와 학습 장치(100)는 동일한 장치일 수도 있다. 또한, 학습 장치는 다음의 프로세스를 수행하기 위한 컴퓨터 판독 가능한 인스트럭션(Instruction)을 저장할 수 있는 메모리(403)를 더 포함할 수 있다. 일 실시예에 따르면, 프로세서, 메모리, 매체 등은 통합 프로세서(Integrated Processor)로 통합될 수 있다.
- [0071] 참고로 이하의 설명에서 혼동을 피하기 위해 “학습용” 이란 문구는 앞서 설명한 학습 프로세스와 관련된 용어에 대해 추가되고, “테스트용” 이란 문구는 테스트 프로세스와 관련된 용어에 대해 추가된다.
- [0072] (1) 학습 장치가, (i) 컨벌루션 레이어(Convolutional Layer)로 하여금 트레이닝 이미지에 대하여 컨벌루션 연산을 적용하여 학습용 특징 맵을 출력하도록 하고, (ii) RPN(Region Proposal Network)으로 하여금 학습용 특징 맵을 이용하여 트레이닝 이미지 내의 k개의 ROI(Region Of Interest)를 출력하도록 하며, (2) 학습 장치가, (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 학습용 특징 맵 상에서 k개의 학습용 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 k개의 학습용 ROI 각각에 대응되는 k개의 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 풀링 레이어에 대응되는 각각의 제1 FC 레이어(Fully Connected Layer) 각각으로 하여금 해당 스케일 별로 k개의 학습용 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 k개의 학습용 제2 특징 벡터를 각각 출력하도록 하며, (iii) 제2 FC 레이어로 하여금 해당 스케일 별로 학습용 제2 특징 벡터를 각각 이용하여 k개의 학습용 ROI 각각

에 대응되는 각각의 학습용 클래스(Class) 정보와 각각의 학습용 리그레션(Regression) 정보를 출력하도록 하며, (3) 제1 로스 레이어(Loss Layer)로 하여금 각각의 학습용 클래스 정보, 각각의 학습용 리그레션 정보 및 이에 대응하는 제1 GT(Ground Truth)를 참조하여 k개의 학습용 ROI에 대응되는 클래스 로스(Class Loss)와 리그레션 로스(Regression Loss)를 생성하도록 함으로써, 클래스 로스와 리그레션 로스를 이용한 백프로파게이션을 통해 제2 FC 레이어, 제1 FC 레이어 및 컨벌루션 레이어를 학습한 상태에서, 통신부(401)은 적어도 하나의 테스트용 객체를 포함하는 테스트 이미지를 획득할 수 있다.

[0073] 프로세서(420)는 (1) (i) 컨벌루션 레이어로 하여금 테스트 이미지에 대하여 컨벌루션 연산을 적용하도록 하여 적어도 하나의 테스트용 특징 맵을 출력하도록 하고, (ii) RPN으로 하여금 테스트용 특징 맵을 이용하여 테스트 이미지 내의 m개의 테스트용 ROI를 출력하도록 하는 제1 프로세서; 및 (2) (i) 서로 다른 풀링 스케일(Pooling Scale)을 가지는 풀링 레이어(Pooling Layer) 각각으로 하여금 해당 스케일 별로 테스트용 특징 맵 상에서 m개의 테스트용 ROI에 대응되는 각각의 영역을 풀링하도록 하여 해당 스케일 별로 m개의 테스트용 ROI 각각에 대응되는 m개의 테스트용 제1 특징 벡터(Feature Vector) 각각을 출력하도록 하고, (ii) 각각의 풀링 레이어에 대응되는 각각의 제1 FC 레이어(Fully Connected Layer)로 하여금 해당 스케일 별로 m개의 테스트용 제1 특징 벡터를 각각 이용하여 해당 스케일 별로 m개의 테스트용 제2 특징 벡터를 각각 출력하도록 하며, (iii) 제2 FC 레이어로 하여금 해당 스케일 별로 m개의 테스트용 제2 특징 벡터를 각각 이용하여 m개의 테스트용 ROI 각각에 대응되는 각각의 테스트용 클래스(Class) 정보와 각각의 테스트용 리그레션(Regression) 정보를 출력하도록 하는 제2 프로세서를 수행할 수 있다.

[0074] 여기서, 테스트 장치(400)에서 RPN이 출력하는 ROI의 개수인 m개는 학습 장치(100)에서 RPN이 출력하는 ROI 개수인 k개보다 작거나 같을 수 있으나, 이에 한정되는 것은 아니다.

[0075] 이와 같이 구성된 테스트 장치(400)를 이용하여 본 발명의 일 실시예에 따른, 서로 다른 풀링 스케일을 갖는 풀링 레이어 각각을 사용하여 CNN 기반의 객체 검출기를 테스트하는 방법을 도 5를 참조하여 설명하면 다음과 같다. 이하의 설명에서, 학습 방법으로부터 용이하게 이해 가능한 부분에 대해서는 상세한 설명을 생략하기로 한다.

[0076] 먼저, 도 3을 참조하여 설명한 학습 방법에 의해 제2 FC 레이어(450), 제1 FC 레이어(440), 및 컨벌루션 레이어(410)를 학습한 상태에서, 테스트 이미지가 입력되면, 테스트 장치(400)는, 컨벌루션 레이어(410)로 하여금 테스트 이미지에 대하여 컨벌루션 연산을 적어도 한 번 적용하도록 하여 적어도 하나의 테스트용 특징 맵을 출력하도록 한다.

[0077] 컨벌루션 레이어(410)로부터 출력된 테스트용 특징 맵을 RPN(420)으로 입력하여, 테스트 장치(400)는 RPN(420)으로 하여금 테스트용 특징 맵을 이용하여 테스트 이미지에 대응하는 m개의 테스트용 ROI를 출력하도록 한다.

[0078] 다음으로, 테스트 장치(400)는, 서로 다른 풀링 스케일을 가지는 풀링 레이어(430) 각각으로 하여금 테스트용 특징 맵 상에서 m개의 테스트용 ROI에 대응되는 각각의 영역을 해당 스케일 별로 풀링하도록 하여 해당 스케일 별로 m개의 테스트용 ROI에 대응되는 m개의 제1 테스트용 특징 벡터 각각을 출력하도록 한다. 이때, 도 5의 예에서는 서로 다른 풀링 스케일을 가지는 3개의 풀링 레이어(430) 각각을 도시하였으나, 본 발명이 이에 한정되는 것은 아니다. 또한, 풀링 스케일이 큰 제1 테스트용 특징 벡터는 풀링 스케일이 작은 제1 테스트용 특징 벡터에 비하여 더 많은 정보를 가질 수 있다.

[0079] 그리고, 테스트 장치(400)는 각각의 풀링 레이어(430)에 대응되는 각각의 제1 FC 레이어(440)로 하여금 m개의 해당 스케일 별 제1 테스트용 특징 벡터 각각을 이용하여, m개의 해당 스케일 별 제2 테스트용 특징 벡터 각각을 출력하도록 한다. 도 5의 예에서는 제1 FC 레이어(440)의 개수가 3개로 나타나 있지만, 이는 풀링 레이어(430)의 개수가 3개인 것에 기인한 것이며, 본 발명이 이에 한정되는 것은 아니다.

[0080] 또한, 제1 FC 레이어(440) 각각에서 출력되는 해당 스케일 별 제2 테스트용 특징 벡터 각각의 채널 수 각각은 동일할 수 있다. 일 예로, 도 5에 도시된 바와 같이, 제1 FC 레이어(440) 각각에서 출력되는 해당 스케일 별 제2 테스트용 특징 벡터의 채널 수 각각은 128일 수 있으며, 이에 한정되는 것은 아니다.

[0081] 또한, 제1 FC 레이어(440) 각각에서 출력되는 제2 테스트용 특징 벡터의 총 채널 수가 기설정된 수 미만이 되도록 할 수 있다. 일 예로, 도 5에 도시된 바와 같이, 제1 FC 레이어(440) 각각에서 출력되는 제2 테스트용 특징 벡터의 총 채널 수는 384(128+128+128)일 수 있으며, 이에 한정되는 것은 아니다. 이 경우, 일반적으로 4,096 채널을 가지는 특징 벡터를 연산하는 종래 방법에 비하여, 본 발명에 따른 총 연산을 수행하는 채널 수는 384개로 적어지므로 그에 따라 테스트 장치(400)의 연산 능력을 상승시킬 수 있게 된다.

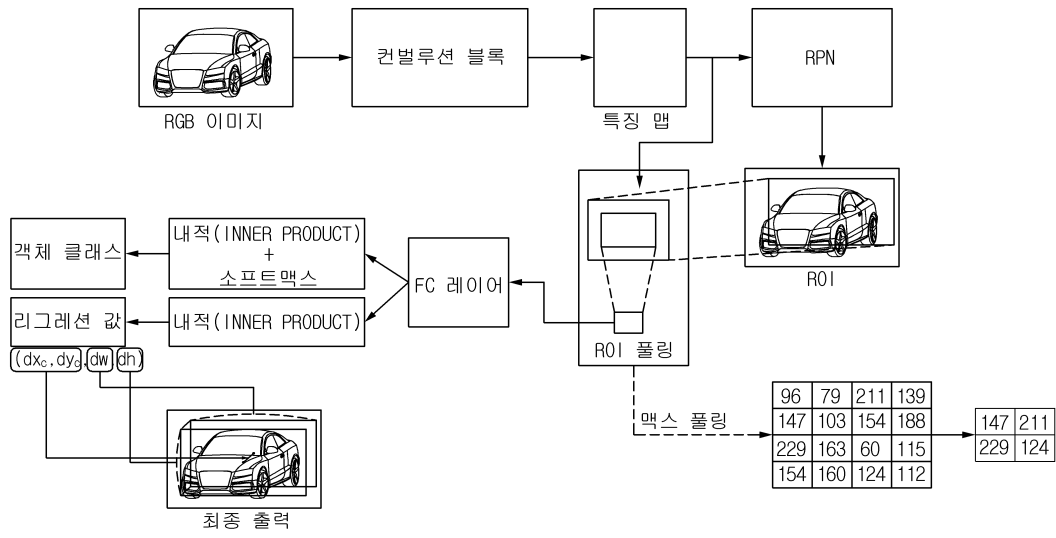
- [0082] 이후, 테스트 장치(400)는 제2 FC 레이어(450)로 하여금 해당 스케일 별로 m개의 제2 테스트용 특징 벡터 각각을 이용하여 m개의 테스트용 ROI 각각에 대응되는 각각의 테스트용 클래스 정보와 각각의 테스트용 리그레션 정보를 출력하도록 한다.
- [0083] 이때, 제2 FC 레이어(450)는 제2\_1 FC 레이어(451) 및 제2\_2 FC 레이어(455)를 포함할 수 있다. 제2\_1 FC 레이어(451)는 제1 FC 레이어(440) 각각에서 출력되는 제2 테스트용 특징 벡터 각각을 컨캐터네이팅하여 제3 테스트용 특징 벡터를 출력할 수 있다. 제2\_2 FC 레이어(455)는 제3 테스트용 특징 벡터를 이용하여 m개의 테스트용 ROI에 대응되는 각각의 테스트용 클래스 정보와 각각의 테스트용 리그레션 정보를 출력할 수 있다.
- [0084] 본 발명에 따른 학습 방법과 테스트 방법은 객체의 스케일에 따라 모드를 전환할 수 있는 CNN을 이용해 감시용으로 사용될 수 있다. 모드에 따라 다양한 레이어의 가중치 및/또는 다양한 로스의 가중치는 조정될 수 있다.
- [0085] 이상 설명된 본 발명에 따른 실시예들은 다양한 컴퓨터 구성요소를 통하여 수행될 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 것들이거나 컴퓨터 소프트웨어 분야의 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능한 기록 매체의 예에는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM, DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 ROM, RAM, 플래시 메모리 등과 같은 프로그램 명령어를 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령어의 예에는, 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드도 포함된다. 상기 하드웨어 장치는 본 발명에 따른 처리를 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0086] 이상에서 본 발명이 구체적인 구성요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나, 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명이 상기 실시예들에 한정되는 것은 아니며, 본 발명이 속하는 기술분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형을 꾀할 수 있다.
- [0087] 따라서, 본 발명의 사상은 상기 설명된 실시예에 국한되어 정해져서는 아니 되며, 후술하는 특허청구범위뿐만 아니라 이 특허청구범위와 균등하게 또는 등가적으로 변형된 모든 것들은 본 발명의 사상의 범주에 속한다고 할 것이다.

**부호의 설명**

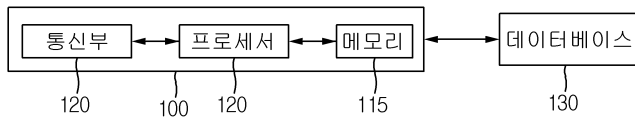
- [0088] 100: 학습 장치
- 110: 통신부
- 120: 프로세서
- 130: 데이터베이스
- 400: 테스트 장치
- 410: 통신부
- 420: 프로세서

도면

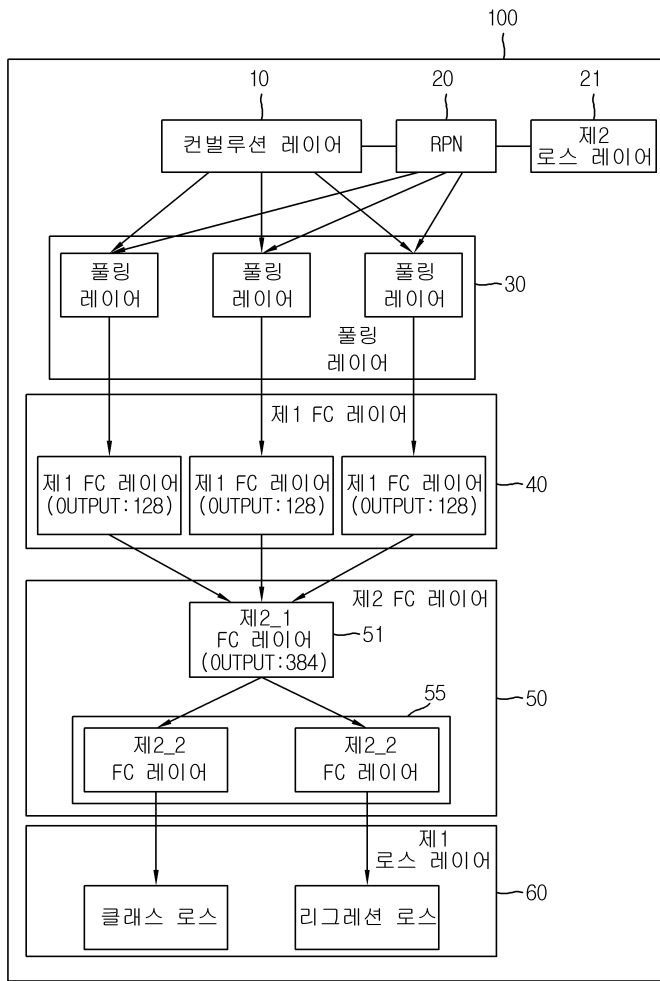
도면1



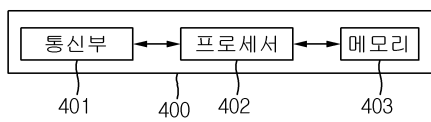
도면2



도면3



도면4



도면5

