



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0153291
(43) 공개일자 2021년12월17일

(51) 국제특허분류(Int. Cl.)
G06T 7/246 (2017.01) G06T 11/40 (2006.01)
G06T 3/40 (2006.01) G06T 7/11 (2017.01)
(52) CPC특허분류
G06T 7/248 (2017.01)
G06T 11/40 (2013.01)
(21) 출원번호 10-2020-0070165
(22) 출원일자 2020년06월10일
심사청구일자 2020년06월10일

(71) 출원인
한국기술교육대학교 산학협력단
충청남도 천안시 동남구 병천면 충절로 1600 (한국기술교육대학교내)
(72) 발명자
오홍선
충청남도 천안시 동남구 병천면 충절로 1600, 공학2관 434호
박주찬
충청남도 천안시 동남구 병천면 가전7길 1-13, 골든빌B동 202호
(뒷면에 계속)
(74) 대리인
조해연

전체 청구항 수 : 총 20 항

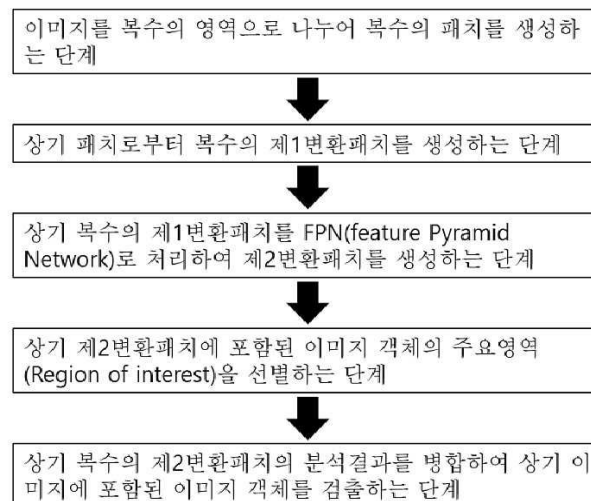
(54) 발명의 명칭 이미지 객체 검출 방법

(57) 요약

본 발명은 이미지 객체 검출 방법에 관한 것이다.

본 발명의 실시 예를 따르는 이미지 객체 검출 방법은, 이미지를 복수의 영역으로 나누어 복수의 패치를 생성하는 단계; 상기 패치로부터 복수의 제1변환패치를 생성하는 단계; 상기 복수의 제1변환패치를 FPN(feature Pyramid Network)로 처리하여 제2변환패치를 생성하는 단계; 상기 제2변환패치에 포함된 이미지 객체의 주요영역(Region of interest)을 선별하는 단계; 및 상기 복수의 제2변환패치의 분석결과를 병합하여 상기 이미지에 포함된 이미지 객체를 검출하는 단계;를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06T 3/40 (2013.01)

G06T 7/11 (2017.01)

(72) 발명자

이선훈

경상남도 창원시 성산구 신촌로 62, 1동 203호

정준욱

대구광역시 달서구 달서대로 67, 유천포스코 더샵
103동 1603호

손성빈

경상북도 구미시 봉곡서로 82, 102동 1005호

명세서

청구범위

청구항 1

이미지를 복수의 영역으로 나누어 복수의 패치를 생성하는 단계;

상기 패치로부터 복수의 제1변환패치를 생성하는 단계;

상기 복수의 제1변환패치를 FPN(feature Pyramid Network)로 처리하여 제2변환패치를 생성하는 단계;

상기 제2변환패치에 포함된 이미지 객체의 주요영역(Region of interest)을 선별하는 단계; 및

상기 복수의 제2변환패치의 분석결과를 병합하여 상기 이미지에 포함된 이미지 객체를 검출하는 단계;를 포함하는,

이미지 객체 검출 방법.

청구항 2

제1항에 있어서,

상기 패치로부터 복수의 제1변환패치를 생성하는 단계는, 상기 패치의 스케일을 변환하여 서로 다른 스케일을 갖는 복수의 제1변환패치를 생성하는 것인,

이미지 객체 검출 방법.

청구항 3

제1항에 있어서,

상기 패치로부터 복수의 제1변환패치를 생성하는 단계는, 상기 패치의 이미지 속성을 변환하여 제1변환패치를 생성하는 것인,

이미지 객체 검출 방법.

청구항 4

제3항에 있어서,

상기 패치의 이미지 속성의 변환은 색감 변환, 밝기 변환, 대비 변환, 가우시안 블러링 적용 및 안개 적용 중 적어도 어느 하나인,

이미지 객체 검출 방법.

청구항 5

제1항에 있어서,

상기 제2변환패치를 생성하는 단계는, 상기 제1변환패치에 대한 다양한 해상도를 갖는 레이어를 생성하는 단계; 및 상기 다양한 해상도를 갖는 레이어를 입력값으로 하여 FPN(feature Pyramid Network)을 수행하는 단계;를 포함하는,

이미지 객체 검출 방법.

청구항 6

제5항에 있어서,

상기 제1변환패치에 대한 다양한 해상도를 갖는 레이어를 생성하는 단계는, 상기 각각의 제1변환패치에 계층 구조의 백본(Backbone)을 형성하는 단계; 및 상기 백본에 대하여 다양한 해상도를 갖는 레이어를 생성하는 단계;를 포함하는,

이미지 객체 검출 방법.

청구항 7

제6항에 있어서,

상기 백본(Backbone)을 형성하는 단계는 ResNetXt를 적용하는 것인,

이미지 객체 검출 방법.

청구항 8

제5항에 있어서,

상기 다양한 해상도를 갖는 레이어를 생성하는 단계는, 각각의 해상도별로 복수의 레이어를 형성하여 Deep features를 생성하는 것인,

이미지 객체 검출 방법.

청구항 9

제1항에 있어서,

상기 제2변환패치를 생성하는 단계는, 상기 FPN 처리로 도출된 각각의 레이어를 Max Pooling 하여 리사이징된 레이어를 생성하는 단계를 더 포함하는,

이미지 객체 검출 방법.

청구항 10

제1항에 있어서,

상기 제2변환패치를 생성하는 단계는, FPN(Feature Pyramid Network)을 수행하는 단계 이후에, 상기 FPN의 결과값을 이용하여 BFN(Balanced feature pyramid network)를 수행하여 제2변환패치를 생성하는 단계를 더 포함하는,

이미지 객체 검출 방법.

청구항 11

제1항에 있어서,

상기 FPN의 결과값은 상기 FPN 처리된 레이어를 Max Pooling한 레이어를 포함하는 것인,

이미지 객체 검출 방법.

청구항 12

제1항에 있어서,

상기 제2변환패치에 포함된 이미지 객체의 주요영역(Region of interest)을 선별하는 단계는, 프로포절 박스(proposal box)를 선정하는 단계; 및 상기 프로포절 박스의 좌표에 RRoI(Rotated Region of Interest)를 적용하여 특징맵(feature map)을 선정하는 단계;를 포함하는,

이미지 객체 검출 방법.

청구항 13

제12항에 있어서,

상기 프로포절 박스의 좌표에 RRoI(Rotated Region of Interest)를 적용하여 특징맵(feature map)을 선정하는 단계는, RoI 변환기(RoI Transformer)를 적용하여 회기층(regression) 및 분류층(classification)으로 각각 공급되는 특징맵을 도출하는 것인,

이미지 객체 검출 방법.

청구항 14

제1항에 있어서,

상기 주요영역(Region of interest)을 선별하는 단계 이후에, 상기 제2변환패치에 포함된 이미지 객체를 카타고리 기준(category-level)에 따라 분류하여 각각의 클래스(class)를 도출하는 단계; 상기 클래스의 불균형을 측정하는 단계; 및 상기 클래스 중 레어 클래스(rare class)에 마진(margin)을 부여하는 단계;를 더 포함하는,

이미지 객체 검출 방법.

청구항 15

아래의 단계를 포함하는 이미지 객체 검출 방법 수행하도록 컴퓨터가 읽을 수 있는 프로그램 코드로 기록한 기록 매체:

이미지를 복수의 영역으로 나누어 복수의 패치를 생성하는 단계;

상기 패치로부터 복수의 제1변환패치를 생성하는 단계;

상기 복수의 제1변환패치를 FPN(feature Pyramid Network)로 처리하여 제2변환패치를 생성하는 단계;

상기 제2변환패치에 포함된 이미지 객체의 주요영역(Region of interest)을 선별하는 단계; 및

상기 복수의 제2변환패치의 분석결과를 병합하여 상기 이미지에 포함된 이미지 객체를 검출하는 단계.

청구항 16

제15항에 있어서,

상기 패치로부터 복수의 제1변환패치를 생성하는 단계는, 상기 패치의 스케일을 변환하여 서로 다른 스케일을 갖는 복수의 제1변환패치를 생성하는 것인,

이미지 객체 검출 방법 수행하도록 컴퓨터가 읽을 수 있는 프로그램 코드로 기록한 기록 매체.

청구항 17

제15항에 있어서,

상기 패치로부터 복수의 제1변환패치를 생성하는 단계는, 상기 패치의 이미지 속성을 변환하여 제1변환패치를 생성하는 것인,

이미지 객체 검출 방법 수행하도록 컴퓨터가 읽을 수 있는 프로그램 코드로 기록한 기록 매체.

청구항 18

제17항에 있어서,

상기 패치의 이미지 속성의 변환은 색감 변환, 밝기 변환, 대비 변환, 가우시안 블러링 적용 및 안개 적용 중 적어도 어느 하나인,

이미지 객체 검출 방법 수행하도록 컴퓨터가 읽을 수 있는 프로그램 코드로 기록한 기록 매체.

청구항 19

제15항에 있어서,

상기 제2변환패치를 생성하는 단계는, 상기 제1변환패치에 대한 다양한 해상도를 갖는 레이어를 생성하는 단계; 및 상기 다양한 해상도를 갖는 레이어를 입력값으로 하여 FPN(feature Pyramid Network)을 수행하는 단계;를 포함하는,

이미지 객체 검출 방법 수행하도록 컴퓨터가 읽을 수 있는 프로그램 코드로 기록한 기록 매체.

청구항 20

제19항에 있어서,

상기 제1변환패치에 대한 다양한 해상도를 갖는 레이어를 생성하는 단계는, 상기 각각의 제1변환패치에 계층 구조의 백본(Backbone)을 형성하는 단계; 및 상기 백본에 대하여 다양한 해상도를 갖는 레이어를 생성하는 단계;를 포함하는,

이미지 객체 검출 방법 수행하도록 컴퓨터가 읽을 수 있는 프로그램 코드로 기록한 기록 매체.

발명의 설명

기술 분야

[0001] 본 발명은 이미지 객체 검출 방법에 관한 것이다.

배경 기술

[0003] 최근 CNN 등의 머신러닝이 시각 영상 분야에서 객체(object) 검출을 위한 용도로 적용되고 있다. 또한, 선행기술문헌인 한국 특허공개공보 제10-2019-0041923호는 높은 정확도의 바운딩 박스를 획득하기 위해 CNN에 포함된 트래킹 네트워크를 이용하는 방법을 개시하고 있다.

[0004] 그러나, 항공촬영물과 같이 높은 해상도를 가진 이미지는 바로 처리하기 어려우며, 일반 사진에 비해 배경 및

물체의 복잡도가 높고, 다양한 물체가 하나의 사진에 포함되어 있는 바 각 물체마다 다양한 회전각도를 가지며, 밀집한 물체를 개별적으로 파악하기 어렵고, 데이터 수가 부족한 문제가 있다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 한국 특허공개공보 제10-2019-0041923호

발명의 내용

해결하려는 과제

[0007] 본 발명은 높은 해상도를 가진 이미지를 효율적으로 처리할 수 있고, 이미지에 포함된 객체의 복잡도가 높고, 데이터 수가 부족한 이미지 처리를 효율적으로 할 수 있는 이미지 객체 검출 방법을 제공함을 목적으로 한다.

[0008] 또한, 데이터 불균형 및 노이즈, 오버피팅(overfitting) 해결에 효과적이다.

과제의 해결 수단

[0010] 본 발명의 실시 예를 따르는 이미지 객체 검출 방법은, 이미지를 복수의 영역으로 나누어 복수의 패치를 생성하는 단계; 상기 패치로부터 복수의 제1변환패치를 생성하는 단계; 상기 복수의 제1변환패치를 FPN(feature Pyramid Network)로 처리하여 제2변환패치를 생성하는 단계; 상기 제2변환패치에 포함된 이미지 객체의 주요영역(Region of interest)을 선별하는 단계; 및 상기 복수의 제2변환패치의 분석결과를 병합하여 상기 이미지에 포함된 이미지 객체를 검출하는 단계;를 포함한다.

[0011] 상기 패치로부터 복수의 제1변환패치를 생성하는 단계는, 상기 패치의 스케일을 변환하여 서로 다른 스케일을 갖는 복수의 제1변환패치를 생성하는 것일 수 있다.

[0012] 상기 패치로부터 복수의 제1변환패치를 생성하는 단계는, 상기 패치의 이미지 속성을 변환하여 제1변환패치를 생성하는 것일 수 있다.

[0013] 상기 제2변환패치를 생성하는 단계는, 상기 제1변환패치에 대한 다양한 해상도를 갖는 레이어를 생성하는 단계; 및 상기 다양한 해상도를 갖는 레이어를 입력값으로 하여 FPN(feature Pyramid Network)을 수행하는 단계;를 포함할 수 있다.

[0014] 상기 제2변환패치를 생성하는 단계는, 상기 FPN 처리로 도출된 각각의 레이어를 Max Pooling 하여 리사이징된 레이어를 생성하는 단계를 더 포함할 수 있다.

[0015] 상기 제2변환패치를 생성하는 단계는, FPN(Feature Pyramid Network)을 수행하는 단계 이후에, 상기 FPN의 결과값을 이용하여 BFN(Balanced feature pyramid network)를 수행하여 제2변환패치를 생성하는 단계를 더 포함할 수 있다.

[0016] 상기 FPN의 결과값은 상기 FPN 처리된 레이어를 Max Pooling한 레이어를 포함할 수 있다.

[0017] 상기 제2변환패치에 포함된 이미지 객체의 주요영역(Region of interest)을 선별하는 단계는, 프로포절 박스(proposal box)를 선정하는 단계; 및 상기 프로포절 박스의 좌표에 RRoI(Rotated Region of Interest)를 적용하여 특징맵(feature map)을 선정하는 단계;를 포함할 수 있다.

[0018] 상기 주요영역(Region of interest)을 선별하는 단계 이후에, 상기 제2변환패치에 포함된 이미지 객체를 카타고리 기준(category-level)에 따라 분류하여 각각의 클래스(class)를 도출하는 단계; 상기 클래스의 불균형을 측정하는 단계; 및 상기 클래스 중 레어 클래스(rare class)에 마진(margin)을 부여하는 단계;를 더 포함할 수 있다.

[0020] 본 발명의 실시 예를 따르는 컴퓨터가 읽을 수 있는 기록 매체는 앞서 설명한 방법을 수행하는 프로그램 코드가 기록된 것이다.

발명의 효과

[0022] 본 발명의 실시 예를 따르는 이미지 객체 검출 방법은, 높은 해상도를 가진 이미지를 효율적으로 처리할 수 있고, 이미지에 포함된 객체의 복잡도가 높고, 데이터 수가 부족한 이미지 처리를 효율적으로 할 수 있다.

[0023] 또한, 데이터 불균형 및 노이즈, 오버피팅(overfitting) 해결에 효과적이다.

도면의 간단한 설명

[0025] 도 1은 본 발명의 실시 예를 따르는 이미지 객체 검출 방법의 순서도이다.

도 2는 본 발명의 실시 예를 따르는 이미지 객체 검출 방법의 진행과정을 도시한 개념도이다.

도 3은 하나의 이미지에서 1024 X 1024 크기의 복수의 패치를 생성하는 단계를 도시한 것이다.

도 4는 이미지 속성 변환을 적용하기 전(a) 및 적용한 후(b)의 이미지를 도시한 것이다.

도 5는 제1변환패치(Image patch)에서 ResNeXt101을 이용하여 백본(Backbone)을 생성한 후 다수의 레이어를 생성하고, FPN 및 BFN을 이용하여 제2변환패치를 생성하는 과정을 도시한 개념도이다.

도 6 및 도 7은 백본(Backbone)으로부터 다양한 해상도를 갖는 복수의 레이어를 생성하는 과정을 도시한 개념도이다.

도 8은 FPN(Feature Pyramid Network)를 수행하는 단계를 도시한 개념도이다.

도 9는 BFN(Balanced Feature Pyramid Network)를 수행하는 단계를 도시한 개념도이다.

도 10은 RPN(Region Proposal Network)의 수행과정을 도시한 개념도이다.

도 11은 NMS를 이용하여 가장 높은 정확도를 갖는 본딩 박스를 선별하는 과정을 도시한 개념도이다.

발명을 실시하기 위한 구체적인 내용

[0026] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시 형태들을 다음과 같이 설명한다. 그러나, 본 발명의 실시 형태는 여러 가지 다른 형태로 변형될 수 있으며, 본 발명의 범위가 이하 설명하는 실시 형태로 한정되는 것은 아니다. 또한, 본 발명의 실시 형태는 당해 기술분야에서 평균적인 지식을 가진 자에게 본 발명을 더욱 완전하게 설명하기 위해서 제공되는 것이다. 따라서, 도면에서의 요소들의 형상 및 크기 등은 보다 명확한 설명을 위해 과장될 수 있으며, 도면 상의 동일한 부호로 표시되는 요소는 동일한 요소이다. 또한, 유사한 기능 및 작용을 하는 부분에 대해서는 도면 전체에 걸쳐 동일한 부호를 사용한다. 덧붙여, 명세서 전체에서 어떤 구성요소를 "포함"한다는 것은 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있다는 것을 의미한다.

[0029] 본 발명의 실시예에 따른 방법의 동작은 컴퓨터, 휴대폰, 태블릿 및 기타 전자기기로 읽을 수 있는 기록매체에 컴퓨터, 휴대폰, 태블릿 및 기타 전자기기가 읽을 수 있는 프로그램 또는 코드로서 구현할 수 있다. 컴퓨터, 휴대폰, 태블릿 및 기타 전자기기가 읽을 수 있는 기록매체는 컴퓨터, 휴대폰, 태블릿 및 기타 전자기기 시스템에 의해 읽힐 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어 분산 방식으로 컴퓨터로 읽을 수 있는 프로그램 또는 코드가 저장되고 실행될 수 있다.

[0030] 또한, 컴퓨터, 휴대폰, 태블릿 및 기타 전자기기가 읽을 수 있는 기록매체는 롬(rom), 램(ram), 플래시 메모리(flash memory) 등과 같이 프로그램 명령 또는 정보를 저장 또는 수행하는 하드웨어 장치를 포함할 수 있다.

[0031] 본 발명의 실시 예에서 각 방법은 앞서 설명한 기기, 장치 및 시스템에서 구현될 수 있는 것으로, 기록매체에 저장된 프로그램 또는 코드를 실행함으로써 수행될 수 있다.

[0033] 도 1은 본 발명의 실시 예를 따르는 이미지 객체 검출 방법의 순서도이다. 도 2는 본 발명의 실시 예를 따르는 이미지 객체 검출 방법의 진행과정을 도시한 개념도이다.

[0035] 도 1 및 도 2를 참조하면, 본 발명의 실시 예를 따르는 이미지 객체 검출 방법은, 이미지를 복수의 영역으로 나누어 복수의 패치를 생성하는 단계; 상기 패치로부터 복수의 제1변환패치를 생성하는 단계; 상기 복수의 제1변환패치를 FPN(feature Pyramid Network)로 처리하여 제2변환패치를 생성하는 단계; 상기 제2변환패치에 포함된 이미지 객체의 주요영역(Region of interest)을 선별하는 단계; 및 상기 복수의 제2변환패치의 분석결과를 병합하여 상기 이미지에 포함된 이미지 객체를 검출하는 단계;를 포함한다.

[0037] 도 3은 하나의 이미지에서 1024 X 1024 크기의 복수의 패치를 생성하는 단계를 도시한 것이다. 이를 참조하면, 3000 X 3000 크기의 이미지에서 1024 X 1024 크기의 패치를 25개 생성할 수 있고(stride 512인 경우), 1500 X 1500 크기의 패치 16개를 생성할 수 있고(stride 700), 3000 X 3000 크기의 패치 1개를 생성할 수 있다. 이를 통해 하나의 이미지에서 총 42개의 패치를 생성할 수 있다. 이를 통해 이미지 데이터 수의 부족을 보완할 수 있다.

[0039] 상기 패치로부터 복수의 제1변환패치를 생성하는 단계는, 상기 패치의 스케일을 변환하여 서로 다른 스케일을 갖는 복수의 제1변환패치를 생성하는 것일 수 있다. 또한, 상기 패치로부터 복수의 제1변환패치를 생성하는 단계는, 상기 패치의 이미지 속성을 변환하여 제1변환패치를 생성하는 것일 수 있으며, 이 경우 상기 패치의 이미지 속성의 변환은 색감 변환, 밝기 변환, 대비 변환, 가우시안 블러링 적용 및 안개 적용 중 적어도 어느 하나 일 수 있다. 도 4는 이미지 속성 변환을 적용하기 전(a) 및 적용한 후(b)의 이미지를 도시한 것이다.

[0040] 또한, 상기 이미지 속성 변환은 서로 중첩적으로 수행될 수 있으며, 이 경우 각 속성에 따라 조합 비율이 결정될 수 있다. 아래의 표 1은 다수의 속성 변환 비율을 적용한 예를 기재한 것이다.

표 1

색상	안개	가우시안 블러링	밝기
일반(65%)Cool(15%) Warm(15%) 적외선(5%)	일반(76%) 1~5 옵션(24%)	일반(76%) Blurring(24%): -필터범위(5,5), 표준편차(1~10)	일반(38%) Bright(62%): -20 ~ 50

[0042] 상기 제2변환패치를 생성하는 단계는, 상기 제1변환패치에 대한 다양한 해상도를 갖는 레이어를 생성하는 단계; 및 상기 다양한 해상도를 갖는 레이어를 입력값으로 하여 FPN(feature Pyramid Network)을 수행하는 단계;를 포함할 수 있다.

[0043] 도 5는 제1변환패치(Image patch)에서 ResNeXt101을 이용하여 백본(Backbone)을 생성한 후 다수의 레이어를 생성하고, FPN 및 BFN을 이용하여 제2변환패치를 생성하는 과정을 도시한 개념도이이고, 도 6 및 도 7은 백본(Backbone)으로부터 다양한 해상도를 갖는 복수의 레이어를 생성하는 과정을 도시한 개념도이다. 도 5 내지 도 7을 참조하면, 상기 제1변환패치에 대한 다양한 해상도를 갖는 레이어를 생성하는 단계는, 상기 각각의 제1변환패치에 계층 구조의 백본(Backbone)을 형성하는 단계; 및 상기 백본에 대하여 다양한 해상도를 갖는 레이어를 생성하는 단계;를 포함할 수 있다. 이 때, 상기 백본(Backbone)을 형성하는 단계는 ResNetXt를 적용하는 것일 수 있다. 상기 다양한 해상도를 갖는 레이어를 생성하는 단계는, 각각의 해상도별로 복수의 레이어를 형성하여 Deep features를 생성하는 것일 수 있다.

- [0045] 상기 제2변환패치를 생성하는 단계는, 상기 FPN 처리로 도출된 각각의 레이어를 Max Pooling 하여 리사이징된 레이어를 생성하는 단계를 더 포함할 수 있다. 도 8을 참조하면, 많은 정보를 가진 상위 레이어가 하위 레이어에 영향을 주는 Top-down 방식으로 특징(features)을 추출할 수 있다. 이를 통해 높은 수준의 정보를 포함하는 레이어를 출력할 수 있다. 이를 위해 딥러닝에서 통상적으로 사용하는 컨볼루션 필터(CONV) 및 액티베이션 함수(ReLU) 등을 이용하여 처리할 수 있다. 도 8을 참조하면, 제1변환패치에서 형성된 Deep features를 컨볼루션 필터를 이용하여 컨볼루션하여 제1차 컨볼루션된 레이어를 형성하고, 상기 제1차컨볼루션된 레이어 중 상위 레이어 및 해당 레이어를 이용하여 제2차 컨볼루션하여 복수의 제2차 컨볼루션된 레이어를 생성할 수 있으며, 이후 다시 컨볼루션을 수행하여 제3차 컨볼루션된 레이어를 생성할 수 있다. 이 때 상위 레이어의 해상도를 해당 레이어의 해상도로 맞추고(업샘플링, Up-sampling), 이를 해당 레이어의 컨볼루션된 레이어와 결합함으로써 컨볼루션된 레이어를 생성할 수 있다. 상기 컨볼루션은 수차례 수행할 수 있다. 또한, FPN처리를 통해 얻어진 각각의 레이어를 Max Pooling함으로써 상기 레이어를 리사이징(resizing) 및 통합하는 단계를 포함할 수 있다. 제2변환패치는 앞서 설명한 n차 컨볼루션을 통해 생성된 레이어 및 Max Pooling으로 생성된 레이어일 수 있다. 이를 통해 제2변환패치가 특징(features)을 보다 명확하게 표현하는 정보를 담도록 할 수 있다.
- [0046] 상기 제2변환패치를 생성하는 단계는, FPN(Feature Pyramid Network)을 수행하는 단계 이후에, 상기 FPN의 결과값을 이용하여 BFN(Balanced feature pyramid network)를 수행하여 제2변환패치를 생성하는 단계를 더 포함할 수 있다. 상기 FPN의 결과값은 상기 FPN 처리된 레이어를 Max Pooling한 레이어를 포함할 수 있다.
- [0048] 본 단계에서는 중간 단의 레이어를 이용하여 균형 잡히고 통합되어 보다 높은 수준의 정보를 갖는 FPN(Feature Pyramid Network)를 얻을 수 있다. 이를 위해 레이어들의 해상도를 중간 단 레이어의 해상도로 맞출 수 있다. 도 9를 참조하면, 중간 단의 레이어보다 해상도가 낮은 레이어는 보간법(interpolation)을 이용하여 해상도를 중간단 레이어와 동일한 수준으로 높이고, 중간 단의 레이어보다 해상도가 높은 경우에는 Max pooling을 이용한 resizing을 통해 중간 단 레이어와 동일하게 해상도를 맞출 수 있다. 그 후 각 레이어의 평균값을 도출하여 신규 레이어를 생성하고, 이를 리파인(refine)하여 Max pooling 및 업샘플링(Up-sampling)을 통해 본래 해상도를 가진 레이어를 생성함으로써 제2변환패치를 생성할 수 있다.
- [0050] 상기 제2변환패치에 포함된 이미지 객체의 주요영역(Region of interest)을 선별하는 단계는, 프로포절 박스(proposal box)를 선정하는 단계; 및 상기 프로포절 박스의 좌표에 RRoI(Rotated Region of Interest)를 적용하여 특징맵(feature map)을 선정하는 단계;를 포함할 수 있다.
- [0051] 본 발명의 실시 예에서는 RRoI(Rotated Region of Interest)를 적용하여 밀집해 있거나 다양한 회전각도로 존재하는 객체(object)를 효과적으로 탐지한다. 이를 위해서는 앵커박스(Anchor box)를 찾아야 한다. 앵커박스는 scale, ratio 및 angle로 표현되는 데, scale은 앵커박스의 크기, ratio는 앵커박스의 가로/세로의 비, angle은 앵커박스의 회전각도를 나타내는 값으로, 앵커박스의 개수는 (num_scale X num_aspect_ratio X num_angles)[1,2,3,4]와 같이 표현될 수 있다. 앵커박스를 찾은 후에는 상기 앵커박스 중에서 프로포절 박스(proposal box)를 선정하게 된다. 이러한 과정을 RPN(Region Proposal Network)이라고 한다. 도 10은 RPN(Region Proposal Network)의 수행과정을 도시한 개념도이다. 일 예에서 제2변환패치의 맵(map)으로부터 앵커박스에 적어도 하나의 객체를 포함하는 지 여부를 나타내는 확률 값인 객체 스코어(object score)를 구하고, 각 맵마다 객체 스코어가 높은 순으로 객체 스코어를 추출하고 이에 대응하는 regression을 추출한다. 이 때, 하나의 맵에서 추출하는 객체 스코어는 특정값, 예를 들면 2000개 이하가 되도록 할 수 있다. 이 중 regression과 앵커박스 간의 IoU가 특정값, 예를 들면 0.7 이상인 것만 선별한다. 이 결과, 각 맵에는 특정값(예를 들면 2000개)을 넘지 않는 regression 및 객체 스코어 쌍이 존재하게 된다. 각 맵의 regression 및 객체 스코어를 합치고, 합쳐진 regression 및 객체 스코어를 다시 객체 스코어에 대하여 내림 차순으로 정렬하여 위에서부터 특정값, 예를 들면 2000개의 상위 regression을 선별할 수 있다. 이렇게 도출된 제안 데이터는 앵커박스의 대각선에 위치한 좌표정보를 포함하고 있다. 도 10에서 Proposals의 값이 2000 x 4인 것은 각 프로포절 박스마다 두 좌표 값 (x1, y1, x2, y2)가 제공되기 때문이다.
- [0053] 상기 프로포절 박스의 좌표에 RRoI(Rotated Region of Interest)를 적용하여 특징맵(feature map)을 선정하는 단계는, RoI 변환기(RoI Transformer)를 적용할 수 있다. RoI 변환기(RoI Transformer)는 HRoI 및 오프셋

(Offset)을 입력값으로 하여 RRoI 학습기(RRoI Learner)를 거쳐 회전각도를 포함하는 RRoI 데이터를 도출할 수 있다. 예를 들어 HRoI 좌표가 (x, y, w, h)인 경우 RRoI 좌표 데이터는 (x1, y1, w1, h1, θ)과 같이 제공될 수 있다. θ 는 HRoI 좌표에 대한 RRoI 좌표의 회전각도이다.

[0054] 프로포절 박스의 정보를 이용하여 RoI 변환기를 통해 특징맵을 도출하는 방법은 다음과 같다. 프로포절 박스의 좌표에는 실수 값이 포함되어 있으며, 이 좌표를 각 맵에 프로젝션(투영)시킨다. 하지만, 각 맵은 정수값을 가지기 때문에 정확한 위치에 프로젝션을 시키기 위해서 RoI Align방법을 사용한다. RoI Align은 bilinear interpolation방법을 이용하여 맵에 실수좌표를 가진 HRoI를 프로젝션 시킬 수 있다. 각 맵에 HRoI좌표와 HRoI에 해당하는 특징맵을 추출하여 RoI 변환기에 들어가게 된다. RoI 변환기에는 RoI 학습기가 존재한다. 앞서 설명한바와 같이 RoI 학습기는 정답 RRoI와 예측한 HRoI의 차이인 오프셋(Offset)을 가지고 학습한다. RoI 학습기의 결과 값으로는 RRoI가 나오며, RRoI의 좌표값을 가지고 다시 각 맵에 프로젝션을 시켜 새로운 RRoI 특징맵을 추출한다. 프로젝션 하는 방법은 RoI Align에서 각도까지 고려하여 발전된 Rotated Position Sensitive RoI Align을 사용할 수 있다. 이때 특징맵의 크기를 1.2배 확대시켜 주변 정보까지 읽어올 수 있다.

[0055] 상기 RoI 변환기를 거쳐 도출된 특징맵은 회귀층(regression) 및 분류층(classification)으로 각각 공급될 수 있다.

[0057] 상기 주요영역(Region of interest)을 선별하는 단계 이후에, 상기 제2변환패치에 포함된 이미지 객체를 카타고리 기준(category-level)에 따라 분류하여 각각의 클래스(class)를 도출하는 단계; 상기 클래스의 불균형을 측정하는 단계; 및 상기 클래스 중 레어 클래스(rare class)에 마진(margin)을 부여하는 단계;를 더 포함할 수 있다. 본 단계는 상기 클래스의 불균형 문제를 해결할 수 있으며, 과적합(overfitting) 등의 문제를 해결할 수 있다.

[0058] 상기 제2변환패치에 포함된 이미지 객체를 카타고리 기준(category-level)에 따라 분류하여 각각의 클래스(class)를 도출하는 단계는 앞서 RoI 변환기를 통해 도출된 분류층(classification)의 데이터에 대하여 softmax 분류기를 실행함으로써 수행될 수 있다.

[0059] 즉, 앞서 RoI 학습기를 통해 도출한 RRoI좌표에 해당하는 특징맵을 이용하여 클래스를 분류하고, 객체의 정확한 위치를 추출할 수 있다.

[0060] 상기 클래스를 도출하는 단계는 구체적으로 아래와 같이 수행될 수 있다. 특징맵에 FC Layer(Fully Connected Layer)를 통과시킬 때 Dropout을 0.5의 확률로 적용한다. FC의 결과 값으로 각 클래스 확률의 스칼라 값(하나의 포인트)을 예측하는 것이 아닌 각 클래스의 분포(평균과 분산)를 예측하도록 한다. 분포를 예측하게 된다면 데이터의 노이즈를 고려할 수 있다. 각 클래스에 대한 평균과 분산을 예측했다면 이것을 가지고 Reparameterization Trick을 사용하여 분포에 데이터 값을 샘플링(Sampling)하여 얻는다. 이와 같은 과정을 총 10번 반복을 한다. Dropout을 적용했기 때문에 샘플링되는 값이 서로 다를 것이며, 샘플링된 데이터의 분산을 구한다. 이 분산이 불균형(Uncertainty) 되는 것이고, Large Margin Softmax에 불균형(Uncertainty)의 크기만큼 클래스(class)에 마진(margin)을 부여한다. 그러므로 rare한 클래스(class)의 영역을 회복하여 두 클래스를 나누는 이상적인 경계선을 형성할 수 있다.

[0061] 불균형(Uncertainty) 측정은 클래스(class)를 분류할 때만 적용하고 본딩박스(bounding box)를 예측할 때에는 단순히 RRoI의 특징맵을 가지고 FC layer를 통과시켜 좌표(x1, y1, w1, h1, θ)를 예측하도록 할 수 있다.

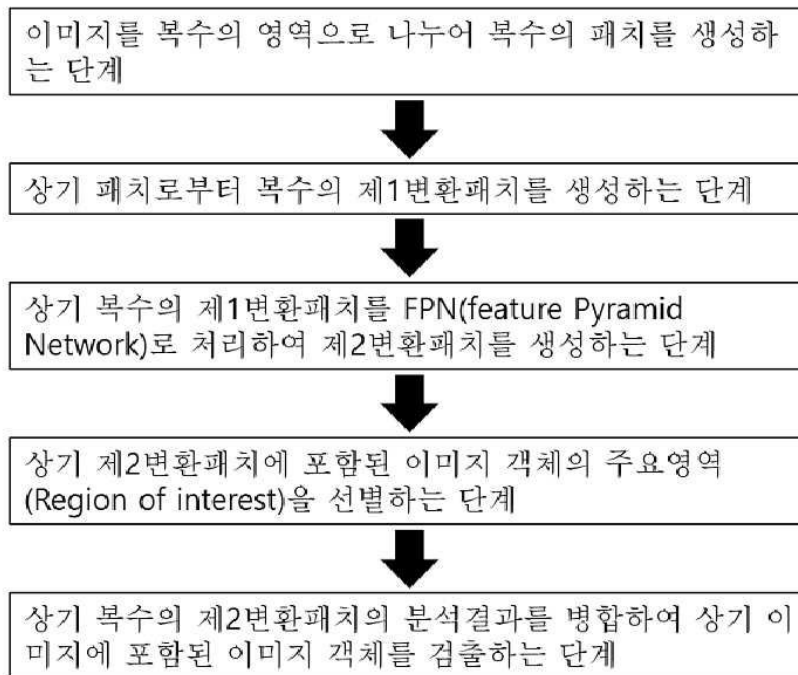
[0063] 상기 복수의 제2변환패치의 분석결과를 병합하여 상기 이미지에 포함된 이미지 객체를 검출하는 단계는 앞서 RoI 변환기를 통해 도출된 회귀층(regression)의 본딩박스(bonding box)를 도출하고, 이를 병합함으로써 수행될 수 있다. 이를 위해 각각의 제2변환패치에 대하여 NMS(Non Maximum Suppression)를 수행할 수 있다. 도 11은 NMS를 이용하여 가장 높은 정확도를 갖는 본딩 박스를 선별하는 과정을 도시한 개념도이다. 도 11에서 보는 바와 같이, 여러 본딩 박스 중에서 가장 높은 스코어를 갖는 본딩 박스를 선택하고 나머지 본딩 박스를 제거함으로써 가장 높은 정확도를 갖는 본딩 박스를 선별할 수 있다.

[0066]

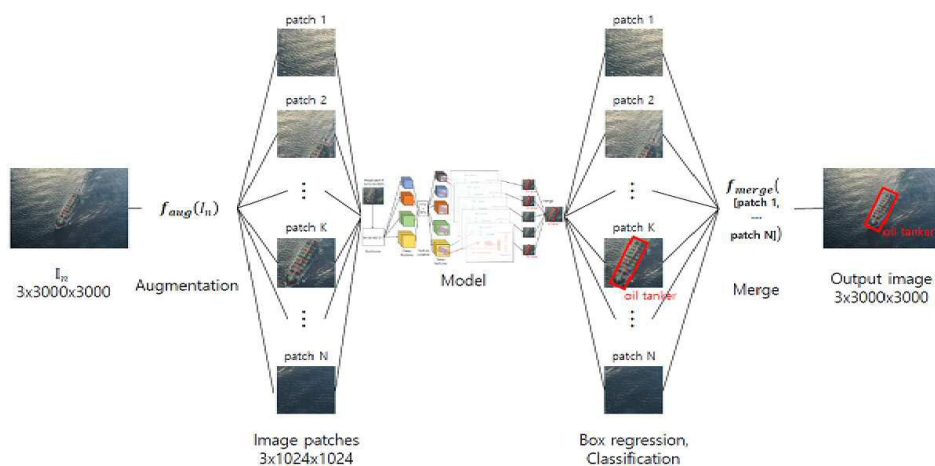
본 발명은 상술한 실시 형태 및 첨부된 도면에 의해 한정되는 것이 아니며 첨부된 청구범위에 의해 한정하고자 한다. 따라서, 청구범위에 기재된 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 당 기술분야의 통상의 지식을 가진 자에 의해 다양한 형태의 치환, 변형 및 변경이 가능할 것이며, 이 또한 본 발명의 범위에 속한다고 할 것이다.

도면

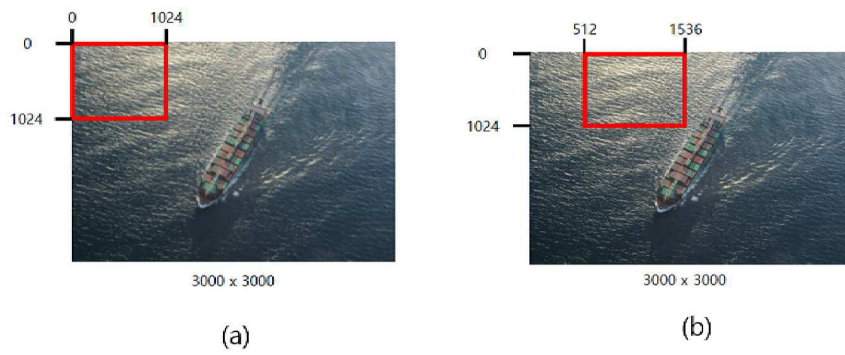
도면1



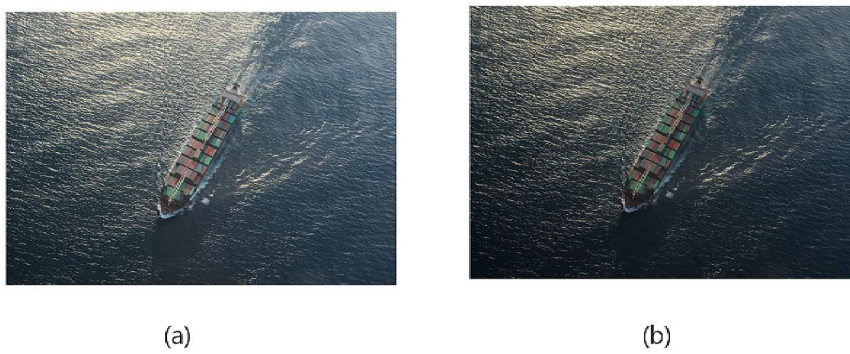
도면2



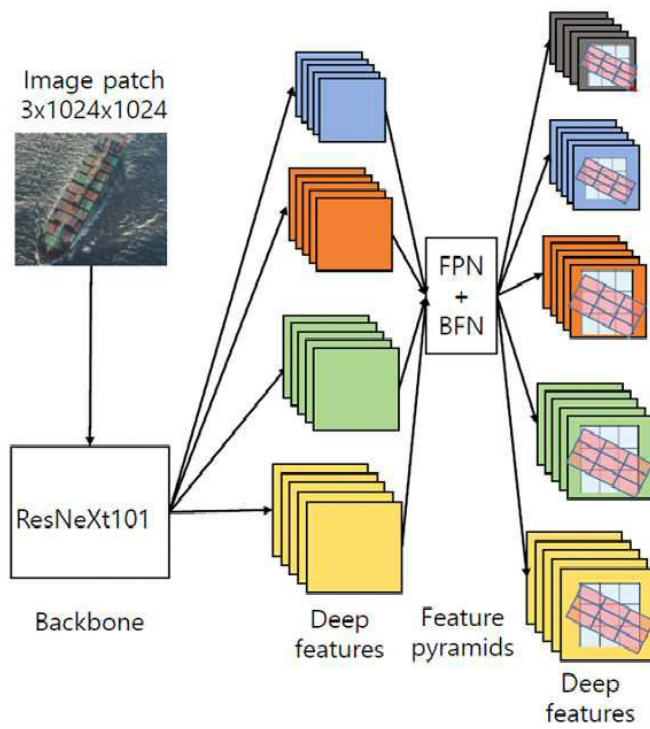
도면3



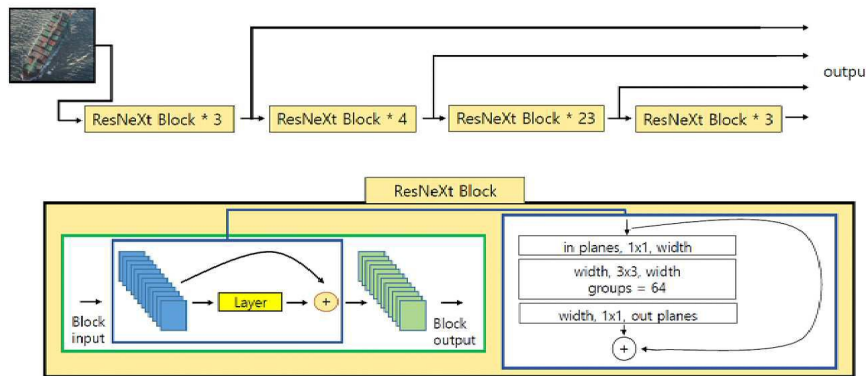
도면4



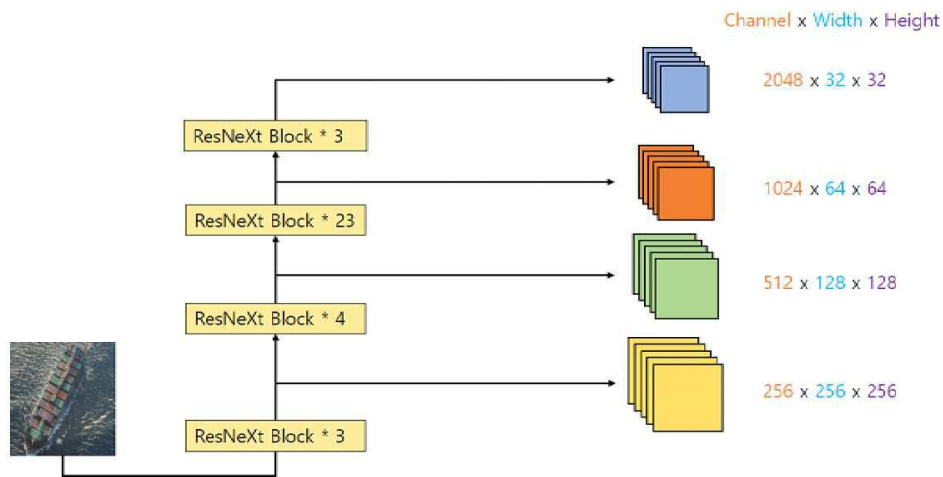
도면5



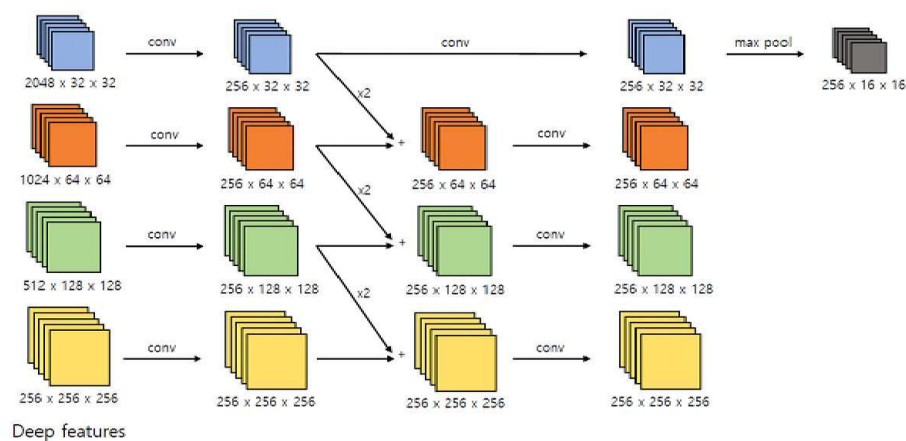
도면6



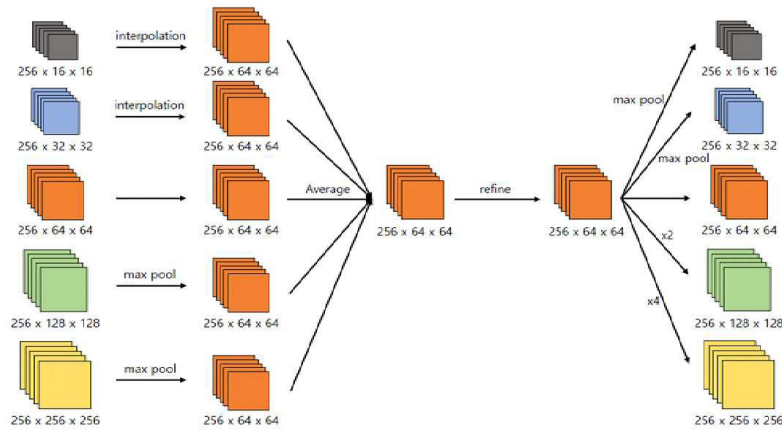
도면7



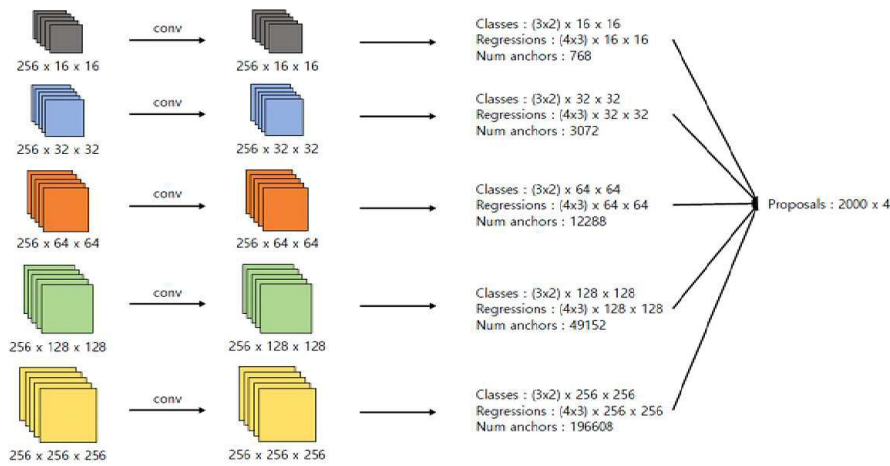
도면8



도면9



도면10



도면11

For each class

