



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0061632  
(43) 공개일자 2022년05월13일

(51) 국제특허분류(Int. Cl.)  
G06N 3/08 (2006.01) G06N 3/04 (2006.01)  
(52) CPC특허분류  
G06N 3/08 (2013.01)  
G06N 3/0472 (2013.01)  
(21) 출원번호 10-2020-0147822  
(22) 출원일자 2020년11월06일  
심사청구일자 2020년11월06일

(71) 출원인  
연세대학교 산학협력단  
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)  
(72) 발명자  
심현정  
인천광역시 연수구 송도문화로28번길 81, 109동 901호(송도동, 송도더샵그린스퀘어)  
이승호  
서울특별시 서초구 언남17길 14, 대원빌딩 5층(양재동)  
(74) 대리인  
민영준

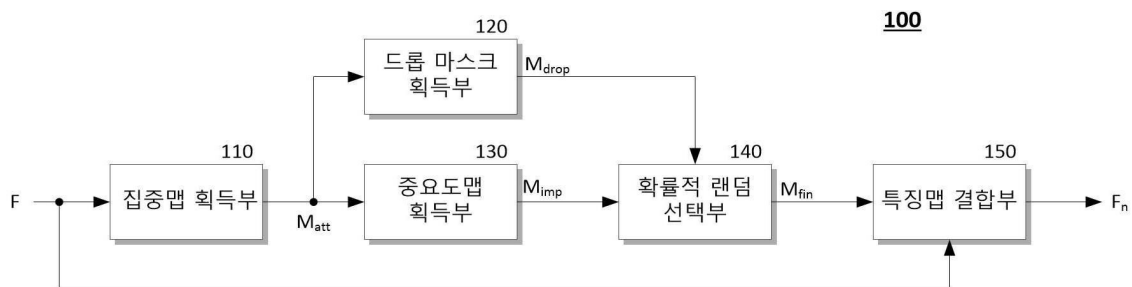
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 자기 집중 모듈 및 이를 이용한 정규화 방법

(57) 요약

본 발명은 다수의 레이어를 포함하는 인공 신경망의 인접한 두 레이어 중 이전 레이어에서 출력되는 3차원의 특징맵을 인가받아 채널 방향의 픽셀값에 대한 평균 풀링을 수행하여 2차원의 집중맵을 획득하는 집중맵 획득부, 집중맵의 각 픽셀의 픽셀값을 기설정된 문턱값과 비교하고, 비교 결과에 따라 각 픽셀에 이진값을 할당하여 드롭 (뒷면에 계속)

대표도



마스크를 생성하는 드롭 마스크 획득부, 집중맵의 각 픽셀의 픽셀값을 기지정된 범위의 값으로 정규화하여 중요도맵을 획득하는 중요도맵 획득부, 드롭 마스크와 중요도맵 중 하나를 기지정된 확률 비율에 따라 랜덤하게 선택하여 최종맵으로 출력하는 확률적 랜덤 선택부 및 특징맵의 각 픽셀에 최종맵의 대응하는 위치의 픽셀값을 가중하여 정규화 특징맵을 획득하고, 획득된 정규화 특징맵을 인접한 두 레이어 중 다음 레이어로 전달하는 특징맵 결합부를 포함하여, 인공 신경망이 특정 부분에만 과도하게 집중하지 않고, 객체 전체 영역에도 주의하여 객체를 식별하도록 하므로, 정확하게 객체와 객체 영역을 정확하게 식별하도록 하는 자기 집중 모듈 및 이를 이용한 정규화 방법을 제공할 수 있다.

이 발명을 지원한 국가연구개발사업

과제고유번호	1711110466
과제번호	2019R1A2C2006123
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	중견연구자지원사업
연구과제명	적대적 생성 신경망의 한계 극복을 위한 기술 개발 및 이를 활용한 영상 변환 및 편집 기술 개발
기 여 율	1/2
과제수행기관명	연세대학교산학협력단
연구기간	2020.03.02 ~ 2021.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1711116255
과제번호	2016-0-00564-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원(한국연구재단부설)
연구사업명	정보통신방송연구개발사업
연구과제명	사용자의 의도와 맥락을 이해하는 지능형 인터랙션 기술 연구개발
기 여 율	1/2
과제수행기관명	경북대학교산학협력단
연구기간	2020.03.01 ~ 2020.12.31

## 명세서

### 청구범위

#### 청구항 1

다수의 레이어를 포함하는 인공 신경망의 인접한 두 레이어 중 이전 레이어에서 출력되는 3차원의 특징맵을 인가받아 채널 방향의 픽셀값에 대한 평균 풀링을 수행하여 2차원의 집중맵을 획득하는 집중맵 획득부;

상기 집중맵의 각 픽셀의 픽셀값을 기설정된 문턱값과 비교하고, 비교 결과에 따라 각 픽셀에 이진값을 할당하여 드롭 마스크를 생성하는 드롭 마스크 획득부;

상기 집중맵의 각 픽셀의 픽셀값을 기설정된 범위의 값으로 정규화하여 중요도맵을 획득하는 중요도맵 획득부;

상기 드롭 마스크와 상기 중요도맵 중 하나를 기설정된 확률 비율에 따라 랜덤하게 선택하여 최종맵으로 출력하는 확률적 랜덤 선택부; 및

상기 특징맵의 각 픽셀에 상기 최종맵의 대응하는 위치의 픽셀값을 가중하여 정규화 특징맵을 획득하고, 획득된 정규화 특징맵을 인접한 두 레이어 중 다음 레이어로 전달하는 특징맵 결합부를 포함하는 자기 집중 모듈.

#### 청구항 2

제1항에 있어서, 상기 드롭 마스크 획득부는

상기 집중맵의 전체 픽셀 수에서 최대 픽셀값의 최상위 자리의 값과 동일한 최상위 자리의 값을 갖는 픽셀의 수의 비에 따라 상기 문턱값을 가변하여 설정하는 자기 집중 모듈.

#### 청구항 3

제1항에 있어서, 상기 중요도맵 획득부는

시그모이드 활성화 함수를 적용하여 상기 집중맵의 각 픽셀의 픽셀값을 정규화하는 자기 집중 모듈.

#### 청구항 4

제1항에 있어서, 상기 자기 집중 모듈은

상기 인공 신경망이 약지도 학습이 수행되는 경우에 다수의 레이어 중 인접한 두 레이어 사이에 삽입 배치되고, 학습 완료 이후에는 제거되는 자기 집중 모듈.

#### 청구항 5

제4항에 있어서, 상기 인공 신경망은

입력되는 이미지에서 학습된 객체의 위치를 탐색하는 객체 위치 탐색 장치 또는 객체별 영역을 식별하여 구분하는 의미적 영역 분할 장치로 이용되도록 학습되는 자기 집중 모듈.

#### 청구항 6

다수의 레이어를 포함하는 인공 신경망의 인접한 두 레이어 중 이전 레이어에서 출력되는 3차원의 특징맵을 인가받아 채널 방향의 픽셀값에 대한 평균 풀링을 수행하여 2차원의 집중맵을 획득하는 단계;

상기 집중맵의 각 픽셀의 픽셀값을 기설정된 문턱값과 비교하고, 비교 결과에 따라 각 픽셀에 이진값을 할당하여 드롭 마스크를 생성하는 단계;

상기 집중맵의 각 픽셀의 픽셀값을 기설정된 범위의 값으로 정규화하여 중요도맵을 획득하는 단계;

상기 드롭 마스크와 상기 중요도맵 중 하나를 기설정된 확률 비율에 따라 랜덤하게 선택하여 최종맵으로 출력하는 단계; 및

상기 특징맵의 각 픽셀에 상기 최종맵의 대응하는 위치의 픽셀값을 가중하여 정규화 특징맵을 획득하고, 획득된 정규화 특징맵을 인접한 두 레이어 중 다음 레이어로 전달하는 단계를 포함하는 자기 집중 모듈을 이용한 정규

화 방법.

## 청구항 7

제6항에 있어서, 상기 드롭 마스크를 생성하는 단계는

상기 집중맵의 전체 픽셀에 대한 픽셀값을 분석하여 최대 픽셀값을 검출하고, 최대 픽셀값의 최상위 자리의 값을 확인하는 단계;

최상위 자리의 값이 동일한 픽셀의 개수를 카운트하는 단계;

상기 집중맵의 전체 픽셀 수와 카운트된 픽셀 수 사이의 비에 따라 문턱값을 설정하는 단계; 및

상기 집중맵의 픽셀값과 설정된 문턱값을 비교하여, 드롭 마스크를 획득하는 단계를 포함하는 자기 집중 모듈을 이용한 정규화 방법.

## 청구항 8

제6항에 있어서, 상기 중요도맵을 획득하는 단계는

시그모이드 활성화 함수를 적용하여 상기 집중맵의 각 픽셀의 픽셀값을 정규화하는 자기 집중 모듈을 이용한 정규화 방법.

## 청구항 9

제6항에 있어서, 상기 자기 집중 모듈은

상기 인공 신경망이 약지도 학습이 수행되는 경우에 다수의 레이어 중 인접한 두 레이어 사이에 삽입 배치되는 자기 집중 모듈을 이용한 정규화 방법.

## 청구항 10

제9항에 있어서, 상기 인공 신경망은

입력되는 이미지에서 학습된 객체의 위치를 탐색하는 객체 위치 탐색 장치 또는 객체별 영역을 식별하여 구분하는 의미적 영역 분할 장치로 이용되도록 학습되는 자기 집중 모듈을 이용한 정규화 방법.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 자기 집중 모듈 및 이를 이용한 정규화 방법에 관한 것으로, 인공 신경망의 약지도 학습을 위한 자기 집중 모듈 및 이를 이용한 정규화 방법에 관한 것이다.

### 배경 기술

[0002] 딥러닝 기술이 고도화됨에 따라 인공 신경망이 이미지 비전 분야에서 높은 성능을 보이고 있다. 이미지 비전 분야에서도 이미지로부터 물체의 위치를 경계 상자(bounding box) 형태로 탐지하는 객체 위치 탐지(object localization) 기술과 의미적 단위에 따라 영상을 분할하는 의미적 분할(semantic segmentation) 기술은 영상 편집과 같은 다양한 용도로 활용될 수 있어 활발하게 연구가 진행되고 있다.

[0003] 도 1은 인공 신경망을 이용한 객체 위치 탐지와 의미적 분할 기술을 설명하기 위한 도면이다.

[0004] 도 1에서 (a)는 입력 이미지이고, (b)는 객체 위치 탐지 기법에 따라 도출된 객체 영역을 나타내며, (c)는 의미적 분할 기법에 따라 이미지에서 서로 다른 의미로 분석된 영역이 분할된 이미지를 나타낸다.

[0005] 도 1의 (b)에 도시된 바와 같이, 객체 위치 탐지 기법은 이미지에서 미리 학습된 객체의 존재 여부를 탐색하여, 탐색된 객체가 포함된 영역을 경계 박스 형태로 추출하는 기술을 나타낸다. 이때 객체 위치 탐지 기법은 경계 박스의 모서리 좌표값을 추출하도록 구성될 수 있다.

[0006] 그리고 의미적 분할 기법은 (c)와 같이, 학습된 방식에 따라 이미지에 포함된 각 객체를 식별하여, 각 객체에 해당하는 영역을 분류하고, 분류된 영역을 구분하여 표시하는 기술로서, (c)에서는 이미지를 고양이, 들판, 산

및 하늘로 의미적으로 구분되어 표시된 예를 나타내었다.

- [0007] 이러한 딥러닝의 성과에도 불구하고 이미지의 객체 위치 탐지 및 의미적 분할 방법을 현실 환경에 적용하기 위해서는 인공 신경망이 이미지에서 각 객체를 식별할 수 있도록 대량의 학습 데이터를 이용하여 미리 지도 학습(supervised learning)되어야 한다. 이때 학습 데이터는 이미지의 각 객체가 미리 구분되고 레이블링되어야 한다. 그러나 학습 데이터는 이미지에서 픽셀 단위 레이블링이 필요하고, 레이블링은 수작업 등으로 이루어지므로 높은 비용으로 인한 학습 데이터 부족 및 클래스의 다양성 부족 문제가 발생한다.
- [0008] 따라서 고비용이 소요되는 학습 데이터를 대량으로 획득하기 어려운 현실 환경에서 딥러닝 모델을 적용하기 위해, 현재는 이미지 수준의 레이블을 활용한 약지도 학습(weakly supervised learning) 방식이 대안으로 제안되었다. 약지도 학습에서는 각 이미지에 포함된 객체의 영역을 정확하게 구분하여 레이블링되지 않고, 단순히 이미지에 포함된 객체의 종류만이 레이블링된 이미지를 학습데이터로 이용할 수 있어, 부족한 픽셀 수준의 레이블 문제를 해결하고 객체 위치 탐지 및 의미적 분할에서 확장성 있는 신경망 도입을 가능하도록 한다. 즉 약지도 학습에서는 각 객체에 대한 식별자인 클래스 정보만을 제공할 뿐, 위치 정보가 제공되지 않은 이미지를 이용하여 학습되도록 한다.
- [0009] 이때 약지도 학습은 클래스 정보만 활용하기 때문에 이미지 분류 학습의 결과로 도출해낼 수 있는 클래스 활성화화맵(class activation map)을 활용하여 객체의 위치 및 형태를 탐지한다. 여기서 인공 신경망에서 클래스 활성화화맵은 객체를 식별하고, 객체 영역을 구분하기 위해 주의를 기울여야하는 영역이 표시되는맵을 의미한다.
- [0010] 도 2는 약지도 학습에서 클래스 활성화화맵의 일 예를 나타낸다.
- [0011] 도 2에서 (a)는 입력 이미지이고, (b)는 클래스 활성화화맵을 나타낸다.
- [0012] 도 2의 (b)를 살펴보면, 약지도 학습에 의해 학습된 클래스 활성화화맵은 객체의 전체 영역에 대해 주의를 집중하기 보다는, 객체를 구분하기에 용이한 희소한 분포를 갖는 특징적(discriminative)인 영역에만 집중하도록 학습된다.
- [0013] (b)와 같은 클래스 활성화화맵의 경우, 객체 식별이 특정 영역에 과도하게 집중되고 나머지 영역은 무시하는 경향이 있다. 이는 객체의 전체적인 윤곽, 즉 객체 영역에 대한 특징이 무시되는 성향이 나타나게 되어, 객체 식별 오류가 발생할 가능성이 존재할 뿐만 아니라, 객체 영역을 식별하기 어렵게 한다. 그러므로 객체가 식별된 특정 영역으로부터 전체적인 객체 영역을 판별하기 위해서 객체 영역을 확장하는 경우에 많은 연산량을 요구하는 문제가 있다.
- [0014] 이에 인공 신경망의 약지도 학습 시에 객체 특징이 강하게 나타나는 특정 영역에 과도하게 집중하지 않고, 객체 영역 전체가 일정 수준 범위에서 고르게 집중될 수 있도록 학습이 수행될 필요가 있다.

## 선행기술문헌

### 특허문헌

- [0015] (특허문헌 0001) 한국 등록 특허 제10-1169340호 (2012.07.23 등록)

## 발명의 내용

### 해결하려는 과제

- [0016] 본 발명의 목적은 인공 신경망의 약지도 학습 시에 추가되어 인공 신경망이 전체 객체 영역에 주의하여 객체를 식별하고 객체 영역을 검출할 수 있도록 하는 자기 집중 모듈 및 이를 이용한 정규화 방법을 제공하는데 있다.
- [0017] 본 발명의 다른 목적은 학습 시에 객체의 가장 특징적인 영역을 기지정된 확률 빈도로 마스킹하여 제거함으로써, 객체의 특정 영역에 대해 과도하게 집중되는 것을 방지할 수 있는 자기 집중 모듈 및 이를 이용한 정규화 방법을 제공하는데 있다.

### 과제의 해결 수단

- [0018] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 자기 집중 모듈은 다수의 레이어를 포함하는 인공 신경

망의 인접한 두 레이어 중 이전 레이어에서 출력되는 3차원의 특징맵을 인가받아 채널 방향의 픽셀값에 대한 평균 풀링을 수행하여 2차원의 집중맵을 획득하는 집중맵 획득부; 상기 집중맵의 각 픽셀의 픽셀값을 기설정된 문턱값과 비교하고, 비교 결과에 따라 각 픽셀에 이진값을 할당하여 드롭 마스크를 생성하는 드롭 마스크 획득부; 상기 집중맵의 각 픽셀의 픽셀값을 기설정된 범위의 값으로 정규화하여 중요도맵을 획득하는 중요도맵 획득부; 상기 드롭 마스크와 상기 중요도맵 중 하나를 기설정된 확률 비율에 따라 랜덤하게 선택하여 최종맵으로 출력하는 확률적 랜덤 선택부; 및 상기 특징맵의 각 픽셀에 상기 최종맵의 대응하는 위치의 픽셀값을 가중하여 정규화 특징맵을 획득하고, 획득된 정규화 특징맵을 인접한 두 레이어 중 다음 레이어로 전달하는 특징맵 결합부를 포함한다.

[0019] 상기 드롭 마스크 획득부는 상기 집중맵의 전체 픽셀 수에서 최대 픽셀값의 최상위 자리의 값과 동일한 최상위 자리의 값을 갖는 픽셀의 수의 비에 따라 상기 문턱값을 가변하여 설정할 수 있다.

[0020] 상기 중요도맵 획득부는 시그모이드 활성화 함수를 적용하여 상기 집중맵의 각 픽셀의 픽셀값을 정규화할 수 있다.

[0021] 상기 자기 집중 모듈은 상기 인공 신경망이 약지도 학습이 수행되는 경우에 다수의 레이어 중 인접한 두 레이어 사이에 삽입 배치되고, 학습 완료 이후에는 제거될 수 있다.

[0022] 상기 인공 신경망은 입력되는 이미지에서 학습된 객체의 위치를 탐색하는 객체 위치 탐색 장치 또는 객체별 영역을 식별하여 구분하는 의미적 영역 분할 장치로 이용될 수 있다.

[0023] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 자기 집중 모듈을 이용한 정규화 방법은 다수의 레이어를 포함하는 인공 신경망의 인접한 두 레이어 중 이전 레이어에서 출력되는 3차원의 특징맵을 인가받아 채널 방향의 픽셀값에 대한 평균 풀링을 수행하여 2차원의 집중맵을 획득하는 단계; 상기 집중맵의 각 픽셀의 픽셀값을 기설정된 문턱값과 비교하고, 비교 결과에 따라 각 픽셀에 이진값을 할당하여 드롭 마스크를 생성하는 단계; 상기 집중맵의 각 픽셀의 픽셀값을 기설정된 범위의 값으로 정규화하여 중요도맵을 획득하는 단계; 상기 드롭 마스크와 상기 중요도맵 중 하나를 기설정된 확률 비율에 따라 랜덤하게 선택하여 최종맵으로 출력하는 단계; 및 상기 특징맵의 각 픽셀에 상기 최종맵의 대응하는 위치의 픽셀값을 가중하여 정규화 특징맵을 획득하고, 획득된 정규화 특징맵을 인접한 두 레이어 중 다음 레이어로 전달하는 단계를 포함한다.

### 발명의 효과

[0024] 따라서, 본 발명의 실시예에 따른 자기 집중 모듈 및 이를 이용한 정규화 방법은 학습 시에 특징맵을 인가받아, 자기 집중맵을 생성하여 객체의 가장 특징적인 영역을 판별하고, 판별된 영역을 기설정된 확률 빈도로 마스킹하여 해당 영역에 대한 집중도를 저하시킴으로써 전체 객체 영역에서 가장 특징적인 영역과 나머지 영역 사이의 집중도 편차가 저감되도록 할 수 있다. 그러므로 특정 영역과 비특정 영역의 집중도 비중이 일정 수준 이내가 되도록 정규화하여 인공 신경망이 특정 부분에만 과도하게 집중하지 않고, 객체 전체 영역에도 주의하여 객체를 식별하도록 하므로, 정확하게 객체와 객체 영역을 정확하게 식별하도록 할 수 있다.

### 도면의 간단한 설명

[0025] 도 1은 인공 신경망을 이용한 객체 위치 탐지와 의미적 분할 기술을 설명하기 위한 도면이다.

도 2는 약지도 학습에서 클래스 활성화맵의 일 예를 나타낸다.

도 3은 본 발명의 일 실시예에 따른 자기 집중 모듈의 개략적 구조를 나타낸다.

도 4는 도 3의 자기 집중 모듈의 구성별 동작을 설명하기 위한 도면이다.

도 5는 도 3의 자기 집중맵 획득부(110)에서 획득된 자기 집중맵의 일 예를 나타낸다.

도 6은 도 3의 드롭 마스크 획득부와 중요도맵 획득부에서 획득되는 드롭 마스크와 중요도맵의 일 예를 나타낸다.

도 7은 본 실시예에 따른 자기 집중 모듈 적용 여부에 따른 클래스 활성화맵과 객체 위치 탐지 결과의 차이를 비교한 도면이다.

도 8은 본 발명의 일 실시예에 따른 자기 집중 모듈을 이용한 정규화 방법을 나타낸다.



## 발명을 실시하기 위한 구체적인 내용

- [0026] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.
- [0027] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.
- [0028] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [0029] 도 3은 본 발명의 일 실시예에 따른 자기 집중 모듈의 개략적 구조를 나타내고, 도 4는 도 3의 자기 집중 모듈의 구성별 동작을 설명하기 위한 도면이다.
- [0030] 본 실시예에 따른 자기 집중 모듈(100)은 약지도 학습이 수행되는 인공 신경망의 학습 시에 결합되는 구성으로, 다수의 레이어를 포함하는 인공 신경망에서 레이어 사이에 배치 결합될 수 있다.
- [0031] 도 3을 참조하면, 본 실시예의 자기 집중 모듈(100)은 집중맵 획득부(110), 드롭 마스크 획득부(120), 중요도맵 획득부(130), 확률적 랜덤 선택부(140) 및 특징맵 결합부(150)를 포함할 수 있다.
- [0032] 인공 신경망의 다수의 레이어 각각은 입력 이미지 또는 이전 배치된 레이어에서 출력되는 특징맵을 인가받아 기 지정된 연산을 수행하도록 구성된 연산 레이어이며, 이미지 처리에 주로 이용되는 대표적인 인공 신경망인 컨볼루션 신경망(Convolutional Neural Network: 이하 CNN)의 경우, 각 레이어는 컨볼루션 연산을 수행한다. 이에 여기서도 자기 집중 모듈(100)이 인공 신경망의 학습시에 컨볼루션 레이어 사이에 배치되는 것으로 가정하여 설명한다.
- [0033] 도 3 및 도 4를 참조하여, 자기 집중 모듈(100)의 각 구성을 설명하면, 집중맵 획득부(110)는 이전 배치된 레이어에서 추출된 특징맵(F)을 입력으로 인가받는다. 여기서 입력되는 특징맵(F)은 이전 배치된 레이어가 입력된 이미지 또는 이전 레이어에서 출력된 특징맵(F)에 대해 현재까지 학습된 방식에 따라 연산을 수행한 결과로서,  $R^{H \times W \times C}$  (여기서 H는 높이, W는 폭, C는 채널 수)의 3차원 크기를 갖는 맵이다.
- [0034] 집중맵 획득부(110)는 특징맵(F)이 인가되면, 특징맵(F)의 채널(C)축 방향으로 채널별 평균 풀링(channelwise average pooling) 연산하여, 3차원의 특징맵(F)을 채널 축 방향으로 압축하여  $R^{H \times W}$  크기의 2차원의 집중맵(attention map)( $M_{att}$ )을 획득한다.
- [0035] 이때, 집중맵( $M_{att}$ )에서 각 픽셀의 세기는 이전 배치되어 특징맵을 출력하는 레이어가 일정 수준이상 학습이 된 상태이라면, 각 픽셀의 세기는 배치된 레이어의 학습 수준에 따른 특징 추출 능력(객체 식별 능력)에 비례하여 나타나는 것으로 볼 수 있다. 따라서 이전 레이어가 충분하게 학습된 상태라면, 특징맵에서 각 픽셀의 세기를 나타내는 픽셀값은 객체 특징을 가장 잘 나타내는 영역에 집중적으로 더 큰값으로 나타내게 된다.
- [0036] 도 5는 도 3의 자기 집중맵 획득부(110)에서 획득된 자기 집중맵의 일 예를 나타낸다.
- [0037] 도 5에 도시된 바와 같이, 인공 신경망이 정상적으로 학습이 완료되어, 새에 대한 이미지로부터 이전 레이어가 정상적으로 특징을 추출한 경우, 추출한 특징 영역을 강조한 집중맵( $M_{att}$ )에서는 새의 특징이 잘 나타나는 일부 영역들에 큰 픽셀값을 갖는 픽셀이 집중되어 나타남을 알 수 있다.
- [0038] 집중맵 획득부(110)는 획득된 집중맵( $M_{att}$ )을 드롭 마스크 획득부(120)와 중요도맵 획득부(130)로 전달한다.
- [0039] 드롭 마스크 획득부(120)는 집중맵( $M_{att}$ )을 인가받고, 인가된 집중맵( $M_{att}$ )의 각 픽셀값이 기지정된 문턱값 이상 인지 여부를 나타내는 드롭 마스크( $M_{drop}$ )를 생성한다. 일 예로 드롭 마스크 획득부(120)는 집중맵( $M_{att}$ )의 각 픽셀값 중 기지정된 문턱값 이상인 픽셀값은 0으로 설정하고, 문턱값 미만인 픽셀값은 1로 설정하여 드롭 마스크

$\text{크}(\text{M}_{\text{drop}})$ 를 생성할 수 있다.

- [0040] 이때, 드롭 마스크 획득부(120)는 집중맵( $\text{M}_{\text{att}}$ )의 전체 픽셀 수에서 최대 픽셀값의 최상위 자리의 값과 동일한 최상위 자리의 값을 갖는 픽셀의 수의 비( $\gamma$ )를 기초로 문턱값을 설정할 수도 있다. 즉 드롭 마스크 획득부(120)는 집중맵( $\text{M}_{\text{att}}$ )의 전체 픽셀에 대한 픽셀값을 우선 분석하고, 분석된 픽셀값 중 최대 픽셀값을 검출하여, 최대 픽셀값의 최상위 자리의 값을 확인한다. 그리고 나머지 픽셀에서 픽셀값의 최상위 자리의 값이 동일한 픽셀의 개수를 카운트한다. 이후 집중맵( $\text{M}_{\text{att}}$ )의 전체 픽셀 수와 카운트된 픽셀 수 사이의 비에 따라 문턱값을 설정하고, 설정된 문턱값에 따라 드롭 마스크( $\text{M}_{\text{drop}}$ )를 생성할 수 있다.
- [0041] 즉 드롭 마스크 획득부(120)는 픽셀값이 큰 픽셀의 수에 대응하여 문턱값을 가변적으로 설정할 수 있다. 이는 집중맵( $\text{M}_{\text{att}}$ )에서 특징이 집중된 영역의 크기에 따라 문턱값을 적응적으로 조절할 수 있도록 하기 위함이다.
- [0042] 한편, 중요도맵 획득부(130)는 집중맵( $\text{M}_{\text{att}}$ )을 인가받고, 인가된 집중맵( $\text{M}_{\text{att}}$ )의 픽셀값의 분포를 기지정된 방식으로 정규화하여 중요도맵( $\text{M}_{\text{imp}}$ )을 생성한다. 집중맵( $\text{M}_{\text{att}}$ )의 경우, 픽셀값의 편차가 매우 크게 나타날 수 있으며, 이는 특징이 강하게 나타나는 특정 영역에 대해 과도하게 집중된 결과를 초래하게 될 수 있다. 따라서 중요도맵 획득부(130)는 특징이 강하게 나타나는 영역과 나머지 영역 사이의 편차가 일정 범위 이내로 제한되도록 함으로써, 특정 영역에 대한 과도하게 집중되는 것을 방지하도록 한다.
- [0043] 중요도맵 획득부(130)는 일 예로 시그모이드(sigmoid) 활성화 함수를 이용하여 집중맵( $\text{M}_{\text{att}}$ )의 각 픽셀값을 0과 1사이의 값으로 정규화함으로써, 중요도맵( $\text{M}_{\text{imp}}$ )을 획득할 수 있다. 즉 중요도맵( $\text{M}_{\text{imp}}$ )에서 특징이 강하게 나타나는 픽셀의 경우 1에 가까운 픽셀값을 갖고, 특징이 나타나지 않는 픽셀의 경우 0에 가까운 픽셀값을 가질 수 있다.
- [0044] 확률적 랜덤 선택부(140)는 드롭 마스크 획득부(120)에서 획득된 드롭 마스크( $\text{M}_{\text{drop}}$ )와 중요도맵 획득부(130)에서 획득된 중요도맵( $\text{M}_{\text{imp}}$ )을 인가받고, 기지정된 마스크 확률 비율에 따라 드롭 마스크( $\text{M}_{\text{drop}}$ )와 중요도맵( $\text{M}_{\text{imp}}$ ) 중 하나를 확률적으로 랜덤하게 선택한다.
- [0045] 여기서 확률적 랜덤 선택부(140)가 마스크 확률 비율에 따라 드롭 마스크( $\text{M}_{\text{drop}}$ )와 중요도맵( $\text{M}_{\text{imp}}$ )을 랜덤하게 선택하는 것은 특정 영역과 나머지 영역 사이에 대한 주의 비율을 조절하기 위해서이다.
- [0046] 중요도맵( $\text{M}_{\text{imp}}$ )은 비록 정규화를 통해 특징이 강하게 나타나는 특정 영역과 나머지 영역 사이의 편차를 감소시키기 위해 획득되는 맵이지만, 결과적으로는 여전히 특정 영역의 픽셀값이 크게 나타나므로 특정 영역에 주의가 집중되는 것을 방지하기 어렵다.
- [0047] 그에 반해, 드롭 마스크( $\text{M}_{\text{drop}}$ )의 경우, 특징이 강하게 나타나는 특정 영역을 마스킹하여 특정 영역 이외의 나머지 영역에 관심을 갖도록 하지만, 드롭 마스크( $\text{M}_{\text{drop}}$ )가 항상 적용되는 경우, 특정 영역에 대해 완전히 주의를 기울이지 않게 되는 결과를 초래하게 된다.
- [0048] 따라서 확률적 랜덤 선택부(140)는 특정 영역에 주의가 집중되지만, 특정 영역에만 과도하게 주의가 집중되지 않도록 마스크 확률 비율에 따라 드롭 마스크( $\text{M}_{\text{drop}}$ )와 중요도맵( $\text{M}_{\text{imp}}$ )을 랜덤하게 선택한다. 일 예로 확률 비율이 0.1로 설정된 경우, 확률적 랜덤 선택부(140)는 중요도맵( $\text{M}_{\text{imp}}$ )과 드롭 마스크( $\text{M}_{\text{drop}}$ )를 9 : 1의 확률 비율로 랜덤하게 선택할 수 있다.
- [0049] 특징맵 결합부(150)는 확률적 랜덤 선택부(140)에서 선택된 최종맵( $\text{M}_{\text{fin}}$ )을 자기 집중 모듈(100)의 입력인 특징맵( $F$ )에 가중하여 정규화 특징맵( $F_n$ )을 출력한다. 특징맵 결합부(150)는  $R^{\text{H} \times \text{W} \times \text{C}}$  크기의 3차원 특징맵( $F$ )에  $R^{\text{H} \times \text{W}}$  크기의 2차원 최종맵( $\text{M}_{\text{fin}}$ )의 대응하는 위치별 픽셀값을 곱하여 정규화 특징맵( $F_n$ )을 획득할 수 있다.
- [0050] 이때, 확률적 랜덤 선택부(140)에서 중요도맵( $\text{M}_{\text{imp}}$ )이 선택된 경우, 특징맵( $F$ )의 픽셀값의 편차는 중요도맵( $\text{M}_{\text{imp}}$ )에 의해 저감되어 정규화 특징맵( $F_n$ )으로 출력된다. 그러나 확률적 랜덤 선택부(140)에서 드롭 마스크( $\text{M}_{\text{drop}}$ )가 선택된 경우, 특징맵( $F$ )에서 픽셀값이 큰 픽셀들은 드롭 마스크( $\text{M}_{\text{drop}}$ )에 의해 마스킹되는 반면, 나머지 픽셀



들은 픽셀값을 그대로 유지하게 된다. 따라서 픽셀값이 큰 픽셀을 제외한 나머지 영역에 대한 중요도를 높일 수 있게 한다.

- [0051] 도 6은 도 3의 드롭 마스크 획득부와 중요도맵 획득부에서 획득되는 드롭 마스크와 중요도맵의 일 예를 나타낸다.
- [0052] 도 6에서 우측 상단이 드롭 마스크( $M_{drop}$ )를 나타내고, 우측 하단이 중요도맵( $M_{imp}$ )을 나타낸다.
- [0053] 도 5에서와 같이, 집중맵 획득부(110)가 새의 특징이 잘 나타나는 일부 영역의 픽셀값이 크게 강조된 집중맵( $M_{att}$ )이 획득되면, 드롭 마스크 획득부(120)는 픽셀값이 문턱값 이상인 영역, 즉 특징이 잘 나타나는 영역이 선택적으로 소거될 수 있도록 드롭 마스크( $M_{drop}$ )를 획득한다. 상기한 바와 같이, 특징맵 결합부(150)에서 드롭 마스크( $M_{drop}$ )가 특징맵(F)에 결합되면, 정규화 특징맵( $F_n$ )은 드롭 마스크( $M_{drop}$ )에 의해 마스크되는 영역(도 5에서는 검은색 영역)의 픽셀값이 모두 0으로 되어 소거되는 반면, 나머지 픽셀의 픽셀값은 그대로 유지된다.
- [0054] 한편, 중요도맵( $M_{imp}$ )의 경우, 집중맵( $M_{att}$ )과 동일한 패턴을 갖지만, 픽셀간 픽셀값의 편차를 줄여준다.
- [0055] 도 7은 본 실시예에 따른 자기 집중 모듈 적용 여부에 따른 클래스 활성화맵과 객체 위치 탐지 결과의 차이를 비교한 도면이다.
- [0056] 도 7에서 빨간색 선은 참값(Groundtruth)을 나타내고, 녹색선은 학습된 인공 신경망의 객체 위치 탐지 결과를 나타낸다.
- [0057] 도 7에서 (a)는 본 실시예에 따른 자기 집중 모듈(100)을 적용하지 않고 학습을 수행한 결과를 나타내고, (b)는 자기 집중 모듈(100)에서 드롭 마스크( $M_{drop}$ )만을 기지정된 확률 비율로 적용하여 학습을 수행한 결과를 나타내며, (c)는 중요도맵( $M_{imp}$ )만을 적용하여 학습을 수행한 결과를 나타낸다. 그리고 (d)는 본 실시예에 따른 자기 집중 모듈(100)이 드롭 마스크( $M_{drop}$ )와 중요도맵( $M_{imp}$ )을 확률 비율에 따라 랜덤하게 선택하여 학습을 수행한 결과를 나타낸다.
- [0058] 도 7에 도시된 바와 같이, 자기 집중 모듈(100)을 적용하지 않고 약지도 학습을 수행하는 경우, 참값과 인공 신경망의 탐지 결과에 매우 큰 오차가 남을 알 수 있다. 그러나 중요도맵( $M_{imp}$ )만을 적용하여 학습을 수행하더라도 참값과의 오차를 크게 줄일 수 있으며, 드롭 마스크( $M_{drop}$ )만을 적용하는 경우, 오차를 매우 줄일 수 있다. 그러나 드롭 마스크( $M_{drop}$ )만을 적용하는 경우, 가운데 이미지와 같이, 경우에 따라서는 여전히 일부 오차를 나타낼 수 있다. 그에 반해, 본 실시예에 따른 자기 집중 모듈(100)을 적용하는 경우, 항시 안정적으로 객체 위치를 정확하게 탐색할 수 있도록 학습되었음을 알 수 있다.
- [0059] 도 8은 본 발명의 일 실시예에 따른 자기 집중 모듈을 이용한 정규화 방법을 나타낸다.
- [0060] 도 3 내지 도 7을 참조하여, 도 8의 자기 집중 모듈을 이용한 정규화 방법을 설명하면, 약지도 학습 시에 이전 레이어에서 획득된 특징맵(F)을 입력받는다(S10). 여기서 특징맵(F)은 이전 레이어가 입력 이미지 또는 그 이전 레이어에서 획득한 특징맵을 인가받고, 현재까지 학습된 방식에 따라 인가된 입력 이미지 또는 특징맵의 특징을 추출하여 획득한 맵으로  $R^{H \times W \times C}$ 의 3차원 크기를 갖는 맵이다.
- [0061] 특징맵(F)이 인가되면, 우선 3차원 크기의 특징맵(F)에서 채널별로 평균 풀링을 수행하여 각 픽셀별로 채널이 집중된  $R^{H \times W}$  크기의 2차원의 집중맵( $M_{att}$ )을 획득한다.
- [0062] 그리고 획득된 집중맵( $M_{att}$ )의 각 픽셀값을 기지정된 문턱값과 비교하여, 문턱값 이상인지 미만인지에 따라 각 픽셀에 서로 다른 이진값을 할당하여 드롭 마스크( $M_{drop}$ )를 생성한다(S30). 이때, 문턱값은 집중맵( $M_{att}$ )의 전체 픽셀 수에서 최대 픽셀값의 최상위 자리의 값과 동일한 최상위 자리의 값을 갖는 픽셀의 수의 비( $\gamma$ )에 따라 가변적으로 설정될 수 있다.
- [0063] 이와 함께 획득된 집중맵( $M_{att}$ )의 각 픽셀값을 기지정된 범위로 정규화하여 중요도맵( $M_{imp}$ )을 생성한다(S40). 이때 중요도맵( $M_{imp}$ )은 집중맵( $M_{att}$ )의 각 픽셀값에 대해 시그모이드 활성화 함수를 적용하여 획득될 수 있다.
- [0064] 드롭 마스크( $M_{drop}$ )와 중요도맵( $M_{imp}$ )이 획득되면, 획득된 드롭 마스크( $M_{drop}$ )와 중요도맵( $M_{imp}$ ) 중 하나를 기설정된

확률 비율에 따라 랜덤하게 선택하여 최종맵( $M_{fin}$ )으로 획득한다(S60).

[0065] 그리고 입력된 특징맵(F)의 각 픽셀에 획득된 최종맵( $M_{fin}$ )의 대응하는 픽셀값을 가중하여 정규화 특징맵( $F_n$ )을 획득한다(S70). 여기서 획득된 정규화 특징맵( $F_n$ )을 인공 신경망의 다음 레이어로 전달된다. 그리고 인공 신경망의 학습 시에 역전파되는 손실은 정규화 특징맵( $F_n$ )이 전달되는 다음 레이어로부터 특징맵(F)이 출력되는 이전 레이어로 그대로 전달되어 기존의 학습 기법이 그대로 유지될 수 있다.

[0066] 결과적으로 본 실시예에 따른 자기 집중 모듈 및 이의 정규화 방법은 인공 신경망의 학습 시에 다수의 레이어 사이에 삽입되어 이전 레이어에서 출력되는 특징맵에서 과도하게 강조된 특징 영역과 나머지 영역 사이에서 주의가 집중되는 차이를 줄여 객체 영역을 더욱 정확하게 구분할 수 있도록 한다.

[0067] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스 될 수 있는 임의의 가용 매체일 수 있고, 또한 컴퓨터 저장 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.

[0068] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다.

[0069] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

## 부호의 설명

[0070] 100: 자기 집중 모듈 110: 집중맵 획득부  
120: 드롭 마스크 획득부 130: 중요도맵 획득부  
140: 확률적 랜덤 선택부 150: 특징맵 결합부

## 도면

### 도면1



(a)



(b)

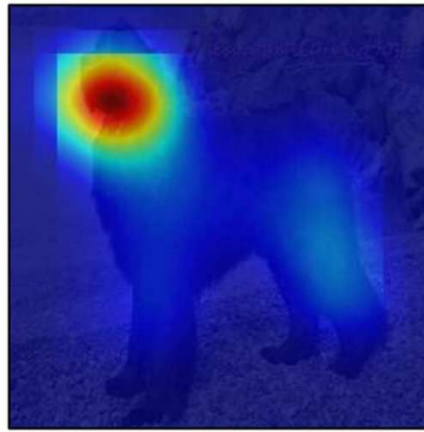


(c)

도면2

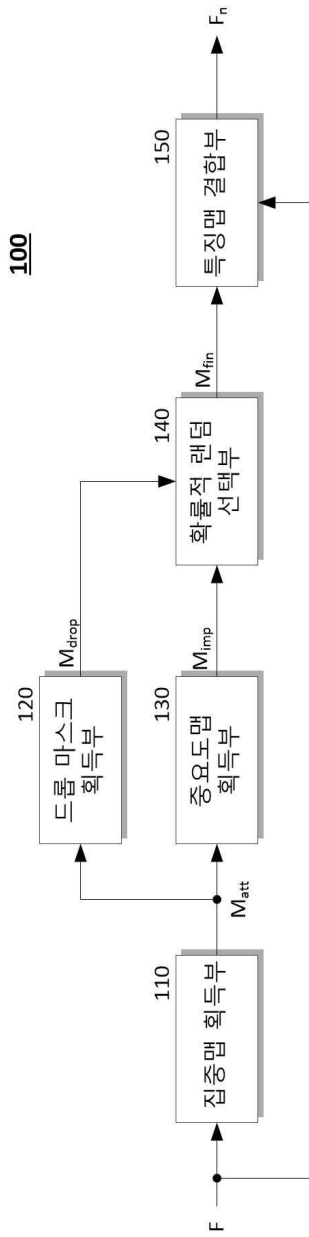


(a)

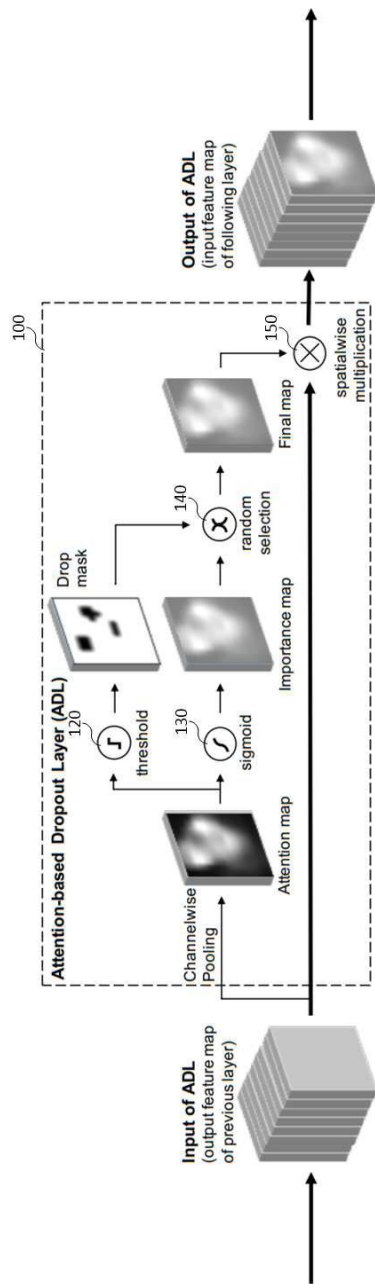


(b)

도면3



도면4



도면5

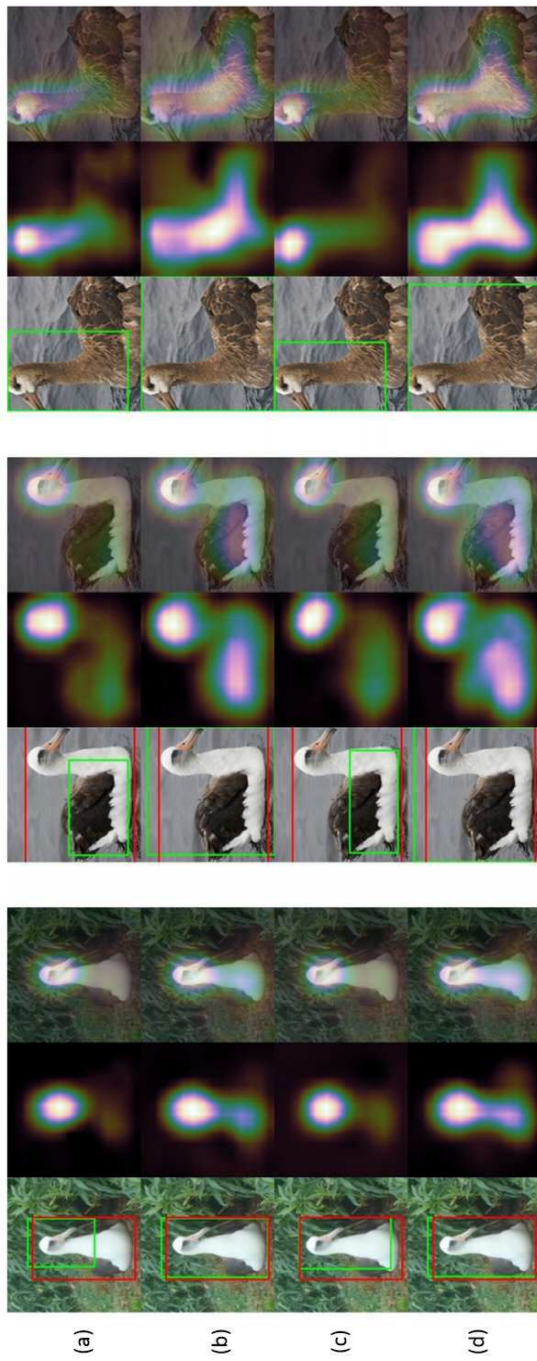


도면6





도면7



도면8

