



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0071378
(43) 공개일자 2025년05월22일

(51) 국제특허분류(Int. Cl.)
G06T 7/11 (2017.01) G06N 3/045 (2023.01)
G06T 5/20 (2024.01) G06V 10/46 (2022.01)
(52) CPC특허분류
G06T 7/11 (2017.01)
G06N 3/045 (2023.01)
(21) 출원번호 10-2023-0157797
(22) 출원일자 2023년11월15일
심사청구일자 2023년11월15일

(71) 출원인
인천대학교 산학협력단
인천광역시 연수구 갯벌로 27, 인천대학교 이노베이션센터(송도동)
(72) 발명자
전광길
인천광역시 연수구 해송로 70, 209동 805호(송도동, 송도웰 카운티2단지)
(74) 대리인
특허법인충정

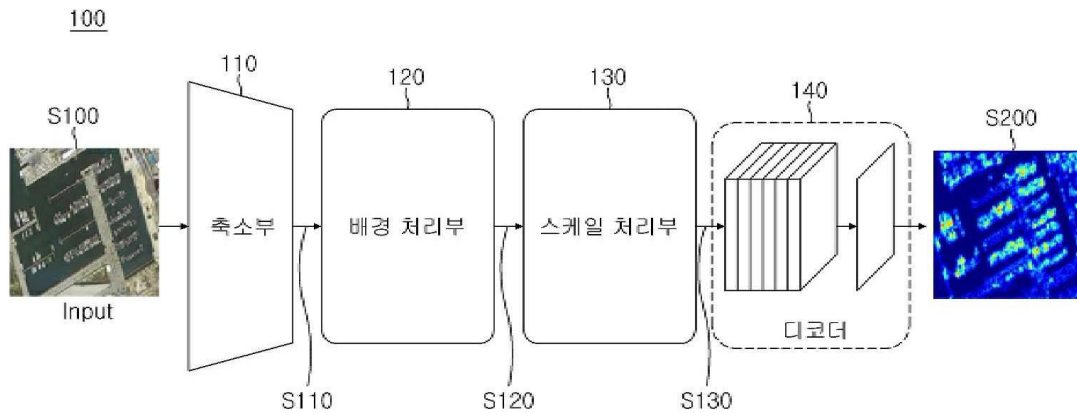
전체 청구항 수 : 총 18 항

(54) 발명의 명칭 원격 감지 객체 계수를 위한 인공지능 이미지 처리 방법

(57) 요약

본 발명은 원격 감지 객체 계수를 위한 인공지능 이미지 처리 방법에 관한 것으로서, 본 발명의 이미지 처리 장치에서 프로세서에서 수행하는 이미지 처리 방법은, 입력 이미지에 대해 다수의 합성곱층과 다수의 풀링층을 이용하여 축소 특징맵을 생성하는 단계, 상기 축소 특징맵으로부터 채널과 폭이 강조되는 제1 차원변경, 높이와 채널이 강조되는 제2 차원변경, 및 높이와 폭이 강조되는 제3 차원변경에 대하여 가중 합성하여, 배경에 대한 특징을 융합하는 배경 특징맵을 생성하는 단계, 및 시계열적으로 수신되는 상기 배경 특징맵에 대하여 가중 합성하여, 스케일에 대한 특징을 융합하는 스케일 인식 특징맵을 생성하는 단계; 및 상기 스케일 인식 특징맵을 기초로 소정의 객체에 대한 인식과 카운팅을 수행하여 해당 객체의 카운팅값을 산출하는 단계를 포함한다.

대표도



(52) CPC특허분류

G06T 5/20 (2024.01)

G06V 10/46 (2023.08)

G06T 2207/20084 (2013.01)

G06T 2207/30242 (2013.01)

명세서

청구범위

청구항 1

이미지 처리 장치에서 프로세서에서 수행하는 이미지 처리 방법에 있어서,

입력 이미지에 대해 다수의 합성곱층과 다수의 풀링층을 이용하여 축소 특징맵을 생성하는 단계;

상기 축소 특징맵으로부터 채널과 폭이 강조되는 제1 차원변경, 높이와 채널이 강조되는 제2 차원변경, 및 높이와 폭이 강조되는 제3 차원변경에 대하여 가중 합성하여, 배경에 대한 특징을 융합하는 배경 특징맵을 생성하는 단계;

시계열적으로 수신되는 상기 배경 특징맵에 대하여 가중 합성하여, 스케일에 대한 특징을 융합하는 스케일 인식 특징맵을 생성하는 단계; 및

상기 스케일 인식 특징맵을 기초로 소정의 객체에 대한 인식과 카운팅을 수행하여 해당 객체의 카운팅값을 산출하는 단계

를 포함하는 이미지 처리 방법.

청구항 2

제1항에 있어서,

상기 배경 특징맵을 생성하는 단계는,

상기 제1 차원변경 특징맵에 대하여, 풀링, 합성곱, 제1 정규화, 제2 정규화를 수행하고, 상기 제1 차원변경 특징맵과 상기 제2 정규화 결과의 원소별 곱셈값을 취하여 제1 중간 특징맵을 생성하며, 상기 제1 중간 특징맵을 채널축으로의 차원변경을 수행하는 제1단계;

상기 제2 차원변경 특징맵에 대하여, 풀링, 합성곱, 제1 정규화, 제2 정규화를 수행하고, 상기 제2 차원변경 특징맵과 여기서의 상기 제2 정규화 결과의 원소별 곱셈값을 취하여 제2 중간 특징맵을 생성하며, 상기 제2 중간 특징맵을 채널축으로의 차원변경을 수행하는 제2단계;

상기 제3 차원변경 특징맵에 대하여, 풀링, 합성곱, 제1 정규화, 제2 정규화를 수행하고, 상기 제3 차원변경 특징맵과 여기서의 상기 제2 정규화 결과의 원소별 곱셈값을 취하여 제3 중간 특징맵을 생성하는 제3단계;

상기 제1 중간 특징맵, 상기 제2 중간 특징맵, 및 상기 제3 중간 특징맵에 대하여, 각 픽셀에 대한 평균값으로 이루어지는 상기 배경 특징맵을 생성하는 제4단계

를 포함하는 이미지 처리 방법.

청구항 3

제2항에 있어서,

상기 제1단계 내지 제3단계에서, 각각의 상기 풀링의 연산은, 소정의 커널을 이용하여 최대값으로 풀링의 결과 및 평균값으로 풀링의 결과를 연결(concatenation)하는 과정을 포함하는 이미지 처리 방법.

청구항 4

제2항에 있어서,

상기 제1단계 내지 제3단계에서, 각각의 상기 합성곱 연산은, 커널 크기가 7*7인 합성곱 연산인 이미지 처리 방법. (Conv7)

청구항 5

제2항에 있어서,

상기 제1단계 내지 제3단계에서, 각각의 상기 제1 정규화 연산은, 배치 정규화 과정을 포함하는 이미지 처리 방법.

청구항 6

제2항에 있어서,

상기 제1단계 내지 제3단계에서, 각각의 상기 제2 정규화 연산은, 시그모이드 함수를 이용한 정규화 과정을 포함하는 이미지 처리 방법.

청구항 7

제1항에 있어서,

상기 스케일 인식 특징맵을 생성하는 단계는,

제1 타임스텝의 상기 배경 특징맵에 대하여 1D 컨볼루션 연산을 수행하여 압축하고 제1커널 간격을 적용하여 합성곱을 수행하는 제1단계;

제N 타임스텝의 (N은 2이상의 자연수)상기 배경 특징맵에 대하여 1D 컨볼루션 연산을 수행하여 압축하고 제N커널 간격을 적용하여 합성곱을 수행한 결과를, 제N-1 단계의 합성곱 결과와 연결하여 제N 합성곱을 수행하는 제N 단계;

상기 제1단계 내지 상기 제N단계의 결과값들을 연결하여 제N+1 합성곱을 수행하여 상기 스케일 인식 특징맵을 생성하는 제N+1 단계

를 포함하는 이미지 처리 방법.

청구항 8

제7항에 있어서,

상기 1D 컨볼루션 연산에 적용하는 합성곱은 커널 크기가 3인 합성곱 연산인 이미지 처리 방법.

청구항 9

제7항에 있어서,

상기 제N 합성곱 및 상기 제N+1 합성곱은 커널 크기가 3X3인 합성곱 연산인 이미지 처리 방법.

청구항 10

이미지 처리 장치에서 이미지 처리 기능을 구현하기 위한 컴퓨터가 판독 가능한 기록매체에 있어서,

입력 이미지에 대해 다수의 합성곱층과 다수의 풀링층을 이용하여 축소 특징맵을 생성하는 기능;

상기 축소 특징맵으로부터 채널과 폭이 강조되는 제1 차원변경, 높이와 채널이 강조되는 제2 차원변경, 및 높이와 폭이 강조되는 제3 차원변경에 대하여 가중 합성하여, 배경에 대한 특징을 융합하는 배경 특징맵을 생성하는 기능;

시계열적으로 수신되는 상기 배경 특징맵에 대하여 가중 합성하여, 스케일에 대한 특징을 융합하는 스케일 인식 특징맵을 생성하는 기능; 및

상기 스케일 인식 특징맵을 기초로 소정의 객체에 대한 인식과 카운팅을 수행하여 해당 객체의 카운팅값을 산출하는 기능

을 수행하는 기록 매체.

청구항 11

제10항에 있어서,

상기 배경 특징맵을 생성하는 기능은,

상기 제1 차원변경 특징맵에 대하여, 풀링, 합성곱, 제1 정규화, 제2 정규화를 수행하고, 상기 제1 차원변경 특

징맵과 상기 제2 정규화 결과의 원소별 곱셈값을 취하여 제1 중간 특징맵을 생성하며, 상기 제1 중간 특징맵을 채널축으로의 차원변경을 수행하는 제1기능;

상기 제2 차원변경 특징맵에 대하여, 풀링, 합성곱, 제1 정규화, 제2 정규화를 수행하고, 상기 제2 차원변경 특징맵과 여기서의 상기 제2 정규화 결과의 원소별 곱셈값을 취하여 제2 중간 특징맵을 생성하며, 상기 제2 중간 특징맵을 채널축으로의 차원변경을 수행하는 제2기능;

상기 제3 차원변경 특징맵에 대하여, 풀링, 합성곱, 제1 정규화, 제2 정규화를 수행하고, 상기 제3 차원변경 특징맵과 여기서의 상기 제2 정규화 결과의 원소별 곱셈값을 취하여 제3 중간 특징맵을 생성하는 제3기능;

상기 제1 중간 특징맵, 상기 제2 중간 특징맵, 및 상기 제3 중간 특징맵에 대하여, 각 픽셀에 대한 평균값으로 이루어지는 상기 배경 특징맵을 생성하는 제4기능

를 수행하는 기록 매체.

청구항 12

제11항에 있어서,

상기 제1기능 내지 제3기능에서, 각각의 상기 풀링의 연산은, 소정의 커널을 이용하여 최대값으로 풀링의 결과 및 평균값으로 풀링의 결과를 연결(concatenation)하는 과정을 포함하는 기록 매체.

청구항 13

제11항에 있어서,

상기 제1기능 내지 제3기능에서, 각각의 상기 합성곱 연산은, 커널 크기가 7*7인 합성곱 연산인 기록 매체.

청구항 14

제11항에 있어서,

상기 제1기능 내지 제3기능에서, 각각의 상기 제1 정규화 연산은, 배치 정규화 과정을 포함하는 기록 매체.

청구항 15

제11항에 있어서,

상기 제1기능 내지 제3기능에서, 각각의 상기 제2 정규화 연산은, 시그모이드 함수를 이용한 정규화 과정을 포함하는 기록 매체.

청구항 16

제10항에 있어서,

상기 스케일 인식 특징맵을 생성하는 기능은,

제1 타임스텝의 상기 배경 특징맵에 대하여 1D 컨볼루션 연산을 수행하여 압축하고 제1커널 간격을 적용하여 합성곱을 수행하는 제1기능;

제N 타임스텝의 (N은 2이상의 자연수)상기 배경 특징맵에 대하여 1D 컨볼루션 연산을 수행하여 압축하고 제N커널 간격을 적용하여 합성곱을 수행한 결과를, 제N-1 기능의 합성곱 결과와 연결하여 제N 합성곱을 수행하는 제N 기능;

상기 제1기능 내지 상기 제N기능의 결과값들을 연결하여 제N+1 합성곱을 수행하여 상기 스케일 인식 특징맵을 생성하는 제N+1 기능

를 포함하는 기록 매체.

청구항 17

제16항에 있어서,

상기 1D 컨볼루션 연산에 적용하는 상기 합성곱은 커널 크기가 3인 합성곱 연산인 기록 매체.

청구항 18

제16항에 있어서,

상기 제N 합성곱 및 상기 제N+1 합성곱은 커널 크기가 3X3인 합성곱 연산인 기록 매체.

발명의 설명

기술 분야

[0001] 본 발명은 이미지 처리 방법에 관한 것으로서, 특히, 트리플 어텐션 (Triple Attention) 및 스케일 인식 (Scale-Aware)을 통한 원격 감지의 객체 계수를 위한 이미지 처리 방법에 관한 것이다.

배경 기술

[0002] 원격 탐색 기술과 위성 플랫폼의 급속한 발전으로 이미지 처리 장치에서 객체감지, 이미지 분류, 이미지 초고해상도와 같은 특정 작업을 구현하기 위한 고품질의 많은 양의 원격 감지 이미지가 제공되고 있다. 이미지에 존재하는 객체 카운팅 기술은 도시 계획, 환경 관리, 모니터링 시스템 및 공공 안전 등 다양한 분야에서 널리 필요한 기술이다. 종래의 객체 계수를 위한 이미지 처리 방법은 검출 기반 방법, 회귀 기반 방법 등 다양한 접근 방식이 시도되고 있다. 최근에는 Convolutional Neural Network(CNN)의 강력한 특징 표현 능력에 힘입어 밀도 추정 기반 방법이 주류의 기술로 자리잡아 가고 있다.

[0003] 도 1은 종래의 원격 감지 이미지에서 객체의 카운팅 기술에 사용된 예시적인 위한 이미지들이다. 도 1과 같이, 위성이나 항공 촬영된 원격 감지 이미지에서와 같이 다양한 스케일의 이미지들이나 배경 클러스터가 있는 이미지들에서, 원호 내의 차량, 박스 내의 선박 등 그 객체의 정확한 식별을 위해 위와 같은 다양한 기술로 접근하고 있다.

[0004] 그러나, 원격 감지 이미지에서의 객체의 카운팅은 혼란스러운 배경 및 다양한 이미지 스케일 등과 같은 다양한 어려운 요인으로 인해 원활하게 수행되기 어려운 문제가 있다. 따라서 원격 감지 이미지에서의 정확한 객체 계수를 위한 이미지 처리 방법을 제안하고자 한다.

[0005] 관련 선행문헌으로서, 특허출원 제 10-2020-0055533호 (2020년05월09일), 특허출원 제10-2022-0090411호 (2022년07월21일), 특허출원 제10-2020-0126326호 (2020년09월28일) 등이 참조될 수 있다.

발명의 내용

해결하려는 과제

[0006] 따라서, 본 발명은 상술한 문제점을 해결하기 위하여 안출된 것으로, 본 발명의 목적은, 딥러닝 기반으로 원격 감지 이미지의 객체 계수를 위한 이미지 처리 성능을 향상시키되, 트리플 어텐션 (Triple Attention) 네트워크를 이용해 3분기 신경망 구조 각각에서 채널/폭, 높이/채널, 높이/폭에 대하여 각각의 가중치로 강조된 특징맵의 처리로 전경과 배경이 구별되도록 하기 위한 특징맵을 생성하고, 스케일 인식 (Scale-Aware) 네트워크를 이용해 1X1 커널을 이용한 합성곱층으로 채널수를 줄인 압축된 4분기 특징맵의 생성과 3X3 커널을 이용해 소정의 가중치로 반영되도록 합성된(가중 합성) 특징맵의 생성으로 왜곡 보상이 이루어져 다양한 스케일에 대한 특징 집계(feature aggregation)가 가능하여, 원격 감지 이미지의 객체 계수 성능을 향상시킬 수 있는 이미지 처리 방법을 제공하는 데 있다.

과제의 해결 수단

[0007] 먼저, 본 발명의 특징을 요약하면, 상기의 목적을 달성하기 위한 본 발명의 일면에 따른 이미지 처리 장치에서 프로세서에서 수행하는 이미지 처리 방법은, 입력 이미지에 대해 다수의 합성곱층과 다수의 풀링층을 이용하여 축소 특징맵을 생성하는 단계; 상기 축소 특징맵으로부터 채널과 폭이 강조되는 제1 차원변경(permute), 높이와 채널이 강조되는 제2 차원변경, 및 높이와 폭이 강조되는 제3 차원변경에 대하여 가중 합성하여, 배경에 대한 특징을 융합(aggregation)하는 배경 특징맵을 생성하는 단계; 시계열적으로 수신되는 상기 배경 특징맵에 대하여 가중 합성하여, 스케일에 대한 특징을 융합(aggregation)하는 스케일 인식 특징맵을 생성하는 단계; 및 상기 스케일 인식 특징맵을 기초로 소정의 객체에 대한 인식과 카운팅을 수행하여 해당 객체의 카운팅값을 산출하는 단계를 포함한다.

- [0008] 상기 배경 특징맵을 생성하는 단계는, 상기 제1 차원변경 특징맵에 대하여, 풀링, 합성곱, 제1 정규화, 제2 정규화를 수행하고, 상기 제1 차원변경 특징맵과 상기 제2 정규화 결과의 원소별 곱셈값을 취하여 제1 중간 특징맵을 생성하며, 상기 제1 중간 특징맵을 채널축으로의 차원변경을 수행하는 제1단계; 상기 제2 차원변경 특징맵에 대하여, 풀링, 합성곱, 제1 정규화, 제2 정규화를 수행하고, 상기 제2 차원변경 특징맵과 여기서의 상기 제2 정규화 결과의 원소별 곱셈값을 취하여 제2 중간 특징맵을 생성하며, 상기 제2 중간 특징맵을 채널축으로의 차원변경을 수행하는 제2단계; 상기 제3 차원변경 특징맵에 대하여, 풀링, 합성곱, 제1 정규화, 제2 정규화를 수행하고, 상기 제3 차원변경 특징맵과 여기서의 상기 제2 정규화 결과의 원소별 곱셈값을 취하여 제3 중간 특징맵을 생성하는 제3단계; 상기 제1 중간 특징맵, 상기 제2 중간 특징맵, 및 상기 제3 중간 특징맵에 대하여, 각 픽셀에 대한 평균값으로 이루어지는 상기 배경 특징맵을 생성하는 제4단계를 포함한다.
- [0009] 상기 제1단계 내지 제3단계에서, 각각의 상기 풀링의 연산은, 소정의 커널을 이용하여 최대값으로 풀링의 결과 및 평균값으로 풀링의 결과를 연결(concatenation)하는 과정을 포함할 수 있다.
- [0010] 상기 제1단계 내지 제3단계에서, 각각의 상기 합성곱 연산은, 커널 크기가 7*7인 합성곱 연산일 수 있다.
- [0011] 상기 제1단계 내지 제3단계에서, 각각의 상기 제1 정규화 연산은, 배치 정규화 과정을 포함할 수 있다.
- [0012] 상기 제1단계 내지 제3단계에서, 각각의 상기 제2 정규화 연산은, 시그모이드 함수를 이용한 정규화 과정을 포함할 수 있다.
- [0013] 상기 스케일 인식 특징맵을 생성하는 단계는, 제1 타임스텝의 상기 배경 특징맵에 대하여 1D 콘볼루션 연산을 수행하여 압축하고 제1커널 간격을 적용하여 합성곱을 수행하는 제1단계; 제N 타임스텝의 (N은 2이상의 자연수)상기 배경 특징맵에 대하여 1D 콘볼루션 연산을 수행하여 압축하고 제N커널 간격을 적용하여 합성곱을 수행한 결과를, 제N-1 단계의 합성곱 결과와 연결하여 제N 합성곱을 수행하는 제N단계; 상기 제1단계 내지 상기 제N 단계의 결과값들을 연결하여 제N+1 합성곱을 수행하여 상기 스케일 인식 특징맵을 생성하는 제N+1 단계를 포함한다.
- [0014] 상기 1D 콘볼루션 연산에 적용하는 합성곱은 커널 크기가 3일 수 있다.
- [0015] 상기 제N 합성곱 및 상기 제N+1 합성곱은 커널 크기가 3X3인 합성곱 연산일 수 있다.
- [0016] 그리고, 본 발명의 다른 일면에 따른 이미지 처리 장치에서 이미지 처리 기능을 구현하기 위한 컴퓨터가 판독 가능한 기록매체는, 입력 이미지에 대해 다수의 합성곱층과 다수의 풀링층을 이용하여 축소 특징맵을 생성하는 기능; 상기 축소 특징맵으로부터 채널과 폭이 강조되는 제1 차원변경, 높이와 채널이 강조되는 제2 차원변경, 및 높이와 폭이 강조되는 제3 차원변경에 대하여 가중 합성하여, 배경에 대한 특징을 융합하는 배경 특징맵을 생성하는 기능; 시계열적으로 수신되는 상기 배경 특징맵에 대하여 가중 합성하여, 스케일에 대한 특징을 융합하는 스케일 인식 특징맵을 생성하는 기능; 및 상기 스케일 인식 특징맵을 기초로 소정의 객체에 대한 인식과 카운팅을 수행하여 해당 객체의 카운팅값을 산출하는 기능을 수행할 수 있다.

발명의 효과

- [0017] 본 발명에 따른 이미지 처리 방법에 따르면, 딥러닝 기반으로 원격 감지 이미지의 객체 계수를 위한 이미지 처리 성능을 향상시키되, 트리플 어텐션 (Triple Attention) 네트워크를 이용해 3분기 신경망 구조 각각에서 채널/폭, 높이/채널, 높이/폭에 대하여 각각의 가중치로 강조된 특징맵의 처리로 전경과 배경이 구별되도록 하기 위한 특징맵을 생성하고, 스케일 인식 (Scale-Aware) 네트워크를 이용해 1D 콘볼루션 연산을 이용한 합성곱층으로 채널수를 줄인 압축된 4분기 특징맵의 생성과 3X3 커널을 이용해 소정의 가중치로 반영되도록 합성된(가중 합성) 특징맵의 생성으로 왜곡 보상이 이루어져 다양한 스케일에 대한 특징 집계(feature aggregation)가 가능하여, 원격 감지 이미지의 객체 계수 성능을 향상시킬 수 있다.

도면의 간단한 설명

- [0018] 본 발명에 관한 이해를 돕기 위해 상세한 설명의 일부로 포함되는 첨부도면은, 본 발명에 대한 실시예를 제공하고 상세한 설명과 함께 본 발명의 기술적 사상을 설명한다.

도 1은 종래의 원격 감지 이미지에서 객체의 카운팅 기술에 사용된 예시적인 위한 이미지들이다.

도 2는 본 발명의 일 실시예에 따른 이미지 처리 장치의 개략적인 블록도이다.

도 3은 도 2의 배경 처리부의 구체적인 블록도이다.

도 4는 도 2의 스케일 처리부의 구체적인 블록도이다.

도 5는 도 4에서의 시계열적 배경 특징맵에 대한 상대적인 시간 차이를 반영하기 위한 시간 확장(dilation)을 위해 커널 간격의 적용을 설명하기 위한 도면이다.

도 6은 본 발명의 일 실시예에 따른 이미지 처리 방법을 처리하는 이미지 처리 장치의 구현 방법의 일례를 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0019] 이하에서는 첨부된 도면들을 참조하여 본 발명에 대해서 자세하게 설명한다. 이때, 각각의 도면에서 동일한 구성 요소는 가능한 동일한 부호로 나타낸다. 또한, 이미 공지된 기능 및/또는 구성에 대한 상세한 설명은 생략한다. 이하에 개시된 내용은, 다양한 실시 예에 따른 동작을 이해하는데 필요한 부분을 중점적으로 설명하며, 그 설명의 요지를 흐릴 수 있는 요소들에 대한 설명은 생략한다. 또한 도면의 일부 구성요소는 과장되거나 생략되거나 또는 개략적으로 도시될 수 있다. 각 구성요소의 크기는 실제 크기를 전적으로 반영하는 것이 아니며, 따라서 각각의 도면에 그려진 구성요소들의 상대적인 크기나 간격에 의해 여기에 기재되는 내용들이 제한되는 것은 아니다.
- [0020] 본 발명의 실시예들을 설명함에 있어서, 본 발명과 관련된 공지기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하기로 한다. 그리고, 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다. 상세한 설명에서 사용되는 용어는 단지 본 발명의 실시 예들을 기술하기 위한 것이며, 결코 제한적이어서는 안 된다. 명확하게 달리 사용되지 않는 한, 단수 형태의 표현은 복수 형태의 의미를 포함한다. 본 설명에서, "포함" 또는 "구비"와 같은 표현은 어떤 특성들, 숫자들, 단계들, 동작들, 요소들, 이들의 일부 또는 조합을 가리키기 위한 것이며, 기술된 것 이외에 하나 또는 그 이상의 다른 특성, 숫자, 단계, 동작, 요소, 이들의 일부 또는 조합의 존재 또는 가능성을 배제하도록 해석되어서는 안 된다.
- [0021] 또한, 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되는 것은 아니며, 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.
- [0022] 도 2는 본 발명의 일 실시예에 따른 이미지 처리 장치(100)의 개략적인 블록도이다.
- [0023] 도 2를 참조하면, 본 발명의 일 실시예에 따른 이미지 처리 장치(100)는, 축소부(110), 배경 처리부(120), 스케일 처리부(130), 및 디코더(140)를 포함할 수 있다.
- [0024] 본 발명의 일 실시예에 따른 이미지 처리 장치(100)를 구성하는 축소부(110), 배경 처리부(120), 스케일 처리부(130), 및 디코더(140) 각각의 장치는, 반도체 프로세서와 같은 하드웨어, 응용 프로그램과 같은 소프트웨어, 또는 이들의 결합으로 구현될 수 있으며, 또한, 본 발명의 이미지 처리 과정의 수행을 위하여 필요한 데이터나 설정 정보 등을 저장하기 위한 메모리를 포함할 수 있다(도 6 참조). 여기서 축소부(110), 배경 처리부(120), 스케일 처리부(130), 및 디코더(140) 각각의 기능에 대하여 설명하지만, 이들 구성 요소들은 2 이상이 결합되어 하나의 블록으로 구현되는 것도 가능하다.
- [0025] 이하, 이미지 처리 장치(100)의 프로세서에서 수행되는 축소부(110), 배경 처리부(120), 스케일 처리부(130), 및 디코더(140)의 각 기능을 자세하게 설명한다.
- [0026] 먼저, 축소부(110)는, 입력 이미지(S100)를 수신하면, 합성곱 신경망(Convolution Neural Network)에서의 다층 퍼셉트론(mlp)에 따른 다수의 합성곱층들과 다수의 풀링층을 이용하여 축소 특징맵(S110)을 생성한다. 상기 합성곱층은 컨볼루션(convolution) 연산과 활성화 함수(예, 시그모이드함수, ReLU 함수 등)를 이용해 정규화할 수 있다. 예를 들어, 소정의 커널을 이용하여 가중 합성하기 위한 10개의 합성곱층과 3개의 맥스(max) 풀링층 또는 평균 풀링층 등이 적절히 배치되어 이용될 수 있다. 축소 특징맵(S110)의 크기는 입력 이미지(S100)의 1/4로 축소된 특징맵일 수 있다.
- [0027] 배경 처리부(120)는, 이와 같은 축소 특징맵(S110)으로부터 채널과 폭이 강조되는 제1 차원변경(permute)(예, 높이축으로의 차원변경), 높이와 채널이 강조되는 제2 차원변경(예, 폭축으로의 차원변경), 및 높이와 폭이 강

조되는 제3 차원변경(예, 채널축으로의 차원변경)에 대하여 가중 합성하여, 전경과 배경이 구분되도록, 배경에 대한 특징을 융합(aggregation)하는 배경 특징맵(S120)을 생성한다.

[0028] 도 3은 도 2의 배경 처리부(120)의 구체적인 블록도이다.

[0029] 도 3을 참조하면, 배경 처리부(120)는, 다수의 전단 차원변경부(121)를 포함한다. 또한, 배경 처리부(120)는, 다수의 전단 차원변경부(121)에서의, 상기 제1 차원변경의 결과인 제1 차원변경 특징맵(S111), 상기 제2 차원변경의 결과인 제2 차원변경 특징맵(S112), 상기 제3 차원변경의 결과인 제3 차원변경 특징맵(S113), 각각에 대하여 중간 특징맵(S114, S115, S116)을 생성하기 위하여, 풀링층(122), 합성곱층(123), 제1 정규화층(124), 제2 정규화층(125), 원소별 곱셈값(element wise product) 연산부(126)를 포함한다. 또한, 배경 처리부(120)는, 복수의 후단 차원변경부(127, 128), 평균 연산부(129)를 포함할 수 있다.

[0030] 축소 특징맵(S110)의 이미지 정보는, 세로 픽셀수 $H * \text{가로 픽셀수 } W * \text{채널 수 } C$ 로 구성될 수 있다. 상기 전단 차원변경부(121)에서 수행하는, 상기 채널과 폭이 강조되는 제1 차원변경(permute)(예, 높이축으로의 차원변경)의 수행 결과인 제1 차원변경 특징맵(S111)에 대하여, 순차적으로 풀링층(122)에서의 풀링, 합성곱층(123)에서의 합성곱(연산), 제1 정규화층(124)에서의 제1 정규화, 제2 정규화층(125)에서의 제2 정규화가 수행된다. 원소별 곱셈값 연산부(126)는 상기 제1 차원변경 특징맵(S111) X_{cw} 과 상기 제2 정규화층(125)의 제2 정규화 결과의 원소별 곱셈값(element wise product)을 취하여 제1 중간 특징맵(X_{top})
$$X_{top} = X_{cw} \odot \text{Sig}(\text{BN}(\text{Conv}(z_{pool}^{th}(X_{cw}))))$$
을 생성한다. 후단 차원변경부(127)는 제1 중간 특징맵(X_{top})에 대하여 채널축으로의 차원변경을 수행한다.

[0031] 여기서, 풀링층(122)에서의 풀링 연산은, 소정의 크기를 갖는 커널을 이용하여 최대값으로 풀링(맥스 풀링)의 결과 및 평균값으로 풀링(평균 풀링)의 결과를 연결(concatenation)한 결과인 $Z\text{-pool}(X) = \text{Cat}(\text{Maxpool}(X), \text{Avgpool}(X))$ 생성하는 과정을 포함할 수 있다. 합성곱층(123)에서의 합성곱 연산은, 예를 들어, 커널 크기가 $7*7$ 인 합성곱 연산(Conv7)일 수 있으며 이에 한정되는 것은 아니다. 제1 정규화층(124)에서의 제1 정규화 연산은 배치 정규화 과정을 포함한다. 상기 배치 정규화(Batch Normalization)는 학습 과정에서 각 배치 단위별로 데이터가 다양한 분포를 가지더라도 각 배치별로 평균과 분산을 이용해 정규화를 수행하는 것이며, 이에 대한 이론은 잘 알려진 바와 같다. 제2 정규화층(125)에서의 제2 정규화 연산은 시그모이드 함수를 이용한 정규화 과정을 포함하며, 다만 이에 한정되는 것은 아니다.

[0032] 상기 전단 차원변경부(121)에서 수행하는, 높이와 채널이 강조되는 제2 차원변경(예, 폭축으로의 차원변경)의 수행 결과인 제2 차원변경 특징맵(S112)에 대하여, 유사한 방식으로 연산이 이루어진다. 제2 차원변경 특징맵(S112)에 대하여, 순차적으로 풀링층(122)에서의 풀링, 합성곱층(123)에서의 합성곱(연산), 제1 정규화층(124)에서의 제1 정규화, 제2 정규화층(125)에서의 제2 정규화가 수행된다. 원소별 곱셈값 연산부(126)는 상기 제2 차원변경 특징맵(S112) X_{ch} 과 상기 제2 정규화층(125)의 제2 정규화 결과의 원소별 곱셈값(element wise product)을 취하여 제2 중간 특징맵(X_{mid})
$$X_{mid} = X_{ch} \odot \text{Sig}(\text{BN}(\text{Conv}(z_{pool}^{th}(X_{ch}))))$$
을 생성한다. 후단 차원변경부(128)는 제2 중간 특징맵(X_{mid})에 대하여 채널축으로의 차원변경을 수행한다.

[0033] 여기서도, 풀링층(122)에서의 풀링 연산은, 소정의 크기를 갖는 커널을 이용하여 최대값으로 풀링(맥스 풀링)의 결과 및 평균값으로 풀링(평균 풀링)의 결과를 연결(concatenation)한 결과인 $Z\text{-pool}(X) = \text{Cat}(\text{Maxpool}(X), \text{Avgpool}(X))$ 생성하는 과정을 포함할 수 있다. 합성곱층(123)에서의 합성곱 연산은, 예를 들어, 커널 크기가 $7*7$ 인 합성곱 연산(Conv7)일 수 있으며 이에 한정되는 것은 아니다. 제1 정규화층(124)에서의 제1 정규화 연산은 배치 정규화 과정을 포함한다. 상기 배치 정규화(Batch Normalization)는 학습 과정에서 각 배치 단위별로 데이터가 다양한 분포를 가지더라도 각 배치별로 평균과 분산을 이용해 정규화를 수행하는 것이며, 이에 대한 이론은 잘 알려진 바와 같다. 제2 정규화층(125)에서의 제2 정규화 연산은 시그모이드 함수를 이용한 정규화 과정을 포함하며, 다만 이에 한정되는 것은 아니다.

[0034] 상기 전단 차원변경부(121)에서 수행하는, 높이와 폭이 강조되는 제3 차원변경(예, 채널축으로의 차원변경)의 수행 결과인 제3 차원변경 특징맵(S113)에 대하여, 유사한 방식으로 연산이 이루어진다. 제3 차원변경 특징맵(S113)에 대하여, 순차적으로 풀링층(122)에서의 풀링, 합성곱층(123)에서의 합성곱(연산), 제1 정규화층(124)에서의 제1 정규화, 제2 정규화층(125)에서의 제2 정규화가 수행된다. 원소별 곱셈값 연산부(126)는 상기 제3 차원변경 특징맵(S113) X_{hw} 과 상기 제2 정규화층(125)의 제2 정규화 결과의 원소별 곱셈값(element wise

product)을 취하여 제3 중간 특징맵(X_{bot})
$$X_{bot} = X_{hw} \odot \text{Sig}(\text{BN}(\text{Conv}(z_{pool}^{th}(X_{hw}))))$$
을 생성한다.

[0035] 여기서도, 풀링층(122)에서의 풀링 연산은, 소정의 크기를 갖는 커널을 이용하여 최대값으로 풀링(맥스 풀링)의 결과 및 평균값으로 풀링(평균 풀링)의 결과를 연결(concatenation)한 결과인

$Z\text{-pool}(X) = \text{Cat}(\text{Maxpool}(X), \text{Avgpool}(X))$ 생성하는 과정을 포함할 수 있다. 합성곱층(123)에서의 합성곱 연산은, 예를 들어, 커널 크기가 7*7인 합성곱 연산(Conv7)일 수 있으며 이에 한정되는 것은 아니다. 제1 정규화층(124)에서의 제1 정규화 연산은 배치 정규화 과정을 포함한다. 상기 배치 정규화(Batch Normalization)는 학습 과정에서 각 배치 단위별로 데이터가 다양한 분포를 가지더라도 각 배치별로 평균과 분산을 이용해 정규화를 수행하는 것이며, 이에 대한 이론은 잘 알려진 바와 같다. 제2 정규화층(125)에서의 제2 정규화 연산은 시그모이드 함수를 이용한 정규화 과정을 포함하며, 다만 이에 한정되는 것은 아니다.

[0036] 이와 같은 제1 중간 특징맵(X_{top}), 제2 중간 특징맵(X_{mid}), 제3 중간 특징맵(X_{bot})에 대하여, 평균 연산부(12

9)는, 각 픽셀에 대한 평균값으로 이루어지는 상기 배경 특징맵(S_{120})
$$X_{tva} = \frac{1}{3}(X_{top} + X_{mid} + X_{bot})$$
을 생성할 수 있다.

[0037] 도 1에서, 스케일 처리부(130)는, 배경 처리부(120)로부터 시계열적으로 수신되는 배경 특징맵(S_{120})에 대하여 가중 합성하여, 스케일에 대한 왜곡 보상이 이루어지도록 스케일에 대한 특징을 융합(aggregation)하는 스케일 인식 특징맵(S_{130})을 생성한다. 이를 위하여 스케일 처리부(130)는 스케일 인식 (Scale-Aware) 네트워크를 이용할 수 있다.

[0038] 도 4는 도 2의 스케일 처리부(130)의 구체적인 블록도이다.

[0039] 도 5는 도 4에서의 시계열적 배경 특징맵에 대한 상대적인 시간 차이를 반영하기 위한 시간 확장(dilation)을 위해 커널 간격(d)의 적용을 설명하기 위한 도면이다.

[0040] 도 4 및 도 5를 참조하면, 스케일 처리부(130)는 다수의 1D 컨볼루션 연산부(131), 다수의 딜레이션(dilation)부(132), 다수의 연결(concatenation)부(133), 다수의 전단 컨볼루션 연산부(134), 후단 연결부(135), 후단 컨볼루션 연산부(136)를 포함할 수 있다.

[0041] 다수의 1D 컨볼루션 연산부(131) 각각은 시계열적인 제1 타임스텝, 제2 타임스텝, ..., 제N 타임스텝(N은 2이상의 자연수) 각각의 배경 특징맵(S_{120})에 대하여 1D 컨볼루션 연산을 수행하여 압축한다. 여기서, 타임스텝 간격은 설계에 따라 미리 정해질 수 있으며, 예를 들어, 수 msec 단위, 수 sec 단위 등 필요한 간격으로 정해질 수 있다. N은 2이상의 자연수로서, 프로세서와 메모리 자원을 고려하여 적절한 수로 정해질 수 있다. 상기 1D 컨볼루션 연산에 적용하는 합성곱은 2, 3, 4 픽셀 등 소정의 커널 크기(1차원으로의 변환이므로 커널 크기는 높이로만 표시됨)를 가질 수 있으며, 바람직하게는 커널 크기가 3인 합성곱 연산일 수 있다.

[0042] 다수의 딜레이션부(132)는, 다수의 1D 컨볼루션 연산부(131)에서의 제1 타임스텝, 제2 타임스텝, ..., 제N 타임스텝 각각의 1D 컨볼루션 연산 결과에 대하여, 각각의 커널 간격($d=1,2,3,4,...,n$)을 가진 커널을 적용하여 합성곱을 수행한다. 상기과 같은 타임스텝에 맞게 그만큼 소정의 픽셀 간격의 이동($d=n$)을 주어 합성곱을 수행한다. 예를 들어, 도 5와 같이, 1x2 커널(세로 1픽셀 x 가로 2픽셀)인 경우, 커널 간격 $d=n$ 은 커널의 가로 제1 픽셀과 제2 픽셀 간의 픽셀 간격이 n 인 경우이다.

[0043] 즉, 상기 제1 타임스텝에 대하여 딜레이션부(132)는 제1 타임스텝의 배경 특징맵(S_{120})에 대하여 1D 컨볼루션 연산을 수행하여 압축하고 제1커널 간격(예, $d=1$)을 적용하여 합성곱을 수행한다(제1단계).

[0044] 또한, 상기 제2 타임스텝에 대하여 딜레이션부(132)는 제2 타임스텝의 배경 특징맵(S_{120})에 대하여 1D 컨볼루션 연산을 수행하여 압축하고 제1커널 간격(예, $d=2$)을 적용하여 합성곱을 수행한 결과를, 상기 제1단계의 합성곱 결과(f_1)와 연결(concatenation)하여 다수의 전단 컨볼루션 연산부(134) 중 하나에서 제2 합성곱(예, 3X3 커널)을 이용한 합성곱 연산)을 수행할 수 있다(제2단계).

[0045] 또한, 상기 제3 타임스텝에 대하여 딜레이션부(132)는 제3 타임스텝의 상기 배경 특징맵(S_{120})에 대하여 1D 컨볼루션 연산을 수행하여 압축하고 제3커널 간격(예, $d=3$)을 적용하여 합성곱을 수행한 결과를, 상기 제2단계의 합성곱 결과(f_2)와 연결하여 다수의 전단 컨볼루션 연산부(134) 중 다른 하나에서 제3 합성곱(예, 3X3 커널)을 이용한 합성곱 연산)을 수행할 수 있다(제3단계).

- [0046] 또한, 상기 제4 타임스텝에 대하여 딜레이션부(132)는 제4 타임스텝의 상기 배경 특징맵(S120)에 대하여 1D 콘볼루션 연산을 수행하여 압축하고 제4커널 간격(예, $d=4$)을 적용하여 합성곱을 수행한 결과를, 상기 제3단계의 합성곱 결과(f3)와 연결하여 다수의 전단 콘볼루션 연산부(134) 중 또 다른 하나에서 제4 합성곱(예, 3×3 커널을 이용한 합성곱 연산)을 수행하고 그 결과(f4)를 출력할 수 있다(제4단계).
- [0047] 이와 같은 상기 제1 타임스텝 이후의 $N=2$ 이상의 제N 타임스텝의 상기 배경 특징맵(S120)에 대하여 유사한 연산이 이루어질 수 있다.
- [0048] 즉, 상기 제1 타임스텝 이후의 제N 타임스텝의 (N은 2이상의 자연수) 상기 배경 특징맵(S120)에 대하여 1D 콘볼루션 연산을 수행하여 압축하고 제N커널 간격(예, $d=N$)을 적용하여 합성곱을 수행한 결과를, 제N-1 단계의 합성곱 결과와 연결하여 다수의 전단 콘볼루션 연산부(134) 각각에서 제N 합성곱(예, 3×3 커널을 이용한 합성곱 연산)을 수행할 수 있다(제N단계).
- [0049] 이에 따라, 후단 연결부(135)는, 상기 제1단계 내지 상기 제N단계의 결과값들을 연결(concatenation)하고, 후단 콘볼루션 연산부(136)는 제N+1 합성곱(예, 3×3 커널을 이용한 합성곱 연산)을 수행하여 상기 스케일 인식 특징맵(S130)을 생성할 수 있다(제N+1 단계).
- [0050] 디코더(140)는 스케일 인식 특징맵(S130)을 기초로 소정의 객체(예, 건물, 선박, 자동차 등)에 대한 인식과 카운팅을 수행하여 해당 객체의 카운팅값을 산출할 수 있다. 이를 위하여 디코더(140)는 상기 다층 퍼셉트론에 따른 다수의 합성곱층들을 이용하여 스케일 인식 특징맵(S130)에 대해 추가적으로 가중합성 처리한 특징맵(S200)으로부터 소정의 객체(예, 건물, 선박, 자동차 등)에 대한 인식과 카운팅을 수행하여 해당 객체의 카운팅값을 산출할 수 있다.
- [0051] 도 6은 본 발명의 일 실시예에 따른 이미지 처리 방법을 처리하는 이미지 처리 장치(100)의 구현 방법의 일례를 설명하기 위한 도면이다.
- [0052] 도 6을 참조하면, 본 발명의 일 실시예에 따른 이미지 처리 방법을 처리하는 이미지 처리 장치(100)는, 하드웨어, 소프트웨어, 또는 이들의 결합으로 이루어질 수 있다. 예를 들어, 본 발명의 이미지 처리 방법을 처리하는 이미지 처리 장치(100)는, 위와 같은 기능/단계/과정들을 수행하기 위한 적어도 하나의 프로세서를 갖는 도 6과 같은 컴퓨팅 시스템(1000) 또는 인터넷 상의 서버 형태로 구현될 수 있다.
- [0053] 컴퓨팅 시스템(1000)은 버스(1200)를 통해 연결되는 적어도 하나의 프로세서(1100), 메모리(1300), 사용자 인터페이스 입력 장치(1400), 사용자 인터페이스 출력 장치(1500), 스토리지(1600), 및 네트워크 인터페이스(1700)를 포함할 수 있다. 프로세서(1100)는 중앙 처리 장치(CPU) 또는 메모리(1300) 및/또는 스토리지(1600)에 저장된 명령어들에 대한 처리를 실행하는 반도체 장치일 수 있다. 메모리(1300) 및 스토리지(1600)는 다양한 종류의 휘발성 또는 비휘발성 저장 매체를 포함할 수 있다. 예를 들어, 메모리(1300)는 ROM(Read Only Memory)(1310) 및 RAM(Random Access Memory)(1320)을 포함할 수 있다. 또한, 네트워크 인터페이스(1700)는 스마트폰, 노트북 PC, 데스크탑 PC 등 해당 장치에서의 유선 인터넷 통신이나 WiFi, WiBro 등 무선 인터넷 통신, WCDMA, LTE 등 이동통신을 지원하는 모뎀 등의 통신 모듈이나, 근거리 무선 통신 방식(예, 블루투스, 지그비, 와이파이 등)의 통신을 지원하는 모뎀 등의 통신모듈을 포함할 수 있다.
- [0054] 따라서, 본 명세서에 개시된 실시예들과 관련하여 설명된 방법 또는 알고리즘의 단계는 프로세서(1100)에 의해 실행되는 하드웨어, 소프트웨어 모듈, 또는 그 2 개의 결합으로 직접 구현될 수 있다. 소프트웨어 모듈은 RAM 메모리, 플래시 메모리, ROM 메모리, EPROM 메모리, EEPROM 메모리, 레지스터, 하드 디스크, 착탈형 디스크, CD-ROM과 같이 컴퓨터 등 장치로 판독 가능한 저장/기록 매체(즉, 메모리(1300) 및/또는 스토리지(1600)) (A non-transitory computer-readable storage medium storing computer executable instructions)에 상주할 수도 있다. 예시적인 저장 매체는 프로세서(1100)에 커플링되며, 그 프로세서(1100)는 저장 매체로부터 정보(코드)를 판독할 수 있고 저장 매체에 정보(코드)를 기입할 수 있다. 다른 방법으로, 저장 매체는 프로세서(1100)와 일체형일 수도 있다. 프로세서 및 저장 매체는 주문형 집적회로(ASIC) 내에 상주할 수도 있다. ASIC는 사용자 단말기 내에 상주할 수도 있다. 다른 방법으로, 프로세서 및 저장 매체는 사용자 단말기 내에 개별 컴포넌트로서 상주할 수도 있다.
- [0055] 상술한 바와 같이, 본 발명에 따른 이미지 처리 장치(100)에 따르면, 딥러닝 기반으로 원격 감지 이미지의 객체 계수를 위한 이미지 처리 성능을 향상시키되, 트리플 어텐션 (Triple Attention) 네트워크를 이용해 3분기 신경망 구조 각각에서 채널/폭, 높이/채널, 높이/폭에 대하여 각각의 가중치로 강조된 특징맵의 처리로 전경과 배경이 구별되도록 하기 위한 특징맵을 생성하고, 스케일 인식 (Scale-Aware) 네트워크를 이용해 1D 콘볼루션 연

산을 이용한 합성곱층으로 채널수를 줄인 압축된 4분기 특징맵의 생성과 3X3 커널을 이용해 소정의 가중치로 반영되도록 합성된(가중 합성) 특징맵의 생성으로 왜곡 보상이 이루어져 다양한 스케일에 대한 특징 집계(feature aggregation)가 가능하여, 원격 감지 이미지의 객체 계수 성능을 향상시킬 수 있도록 하였다.

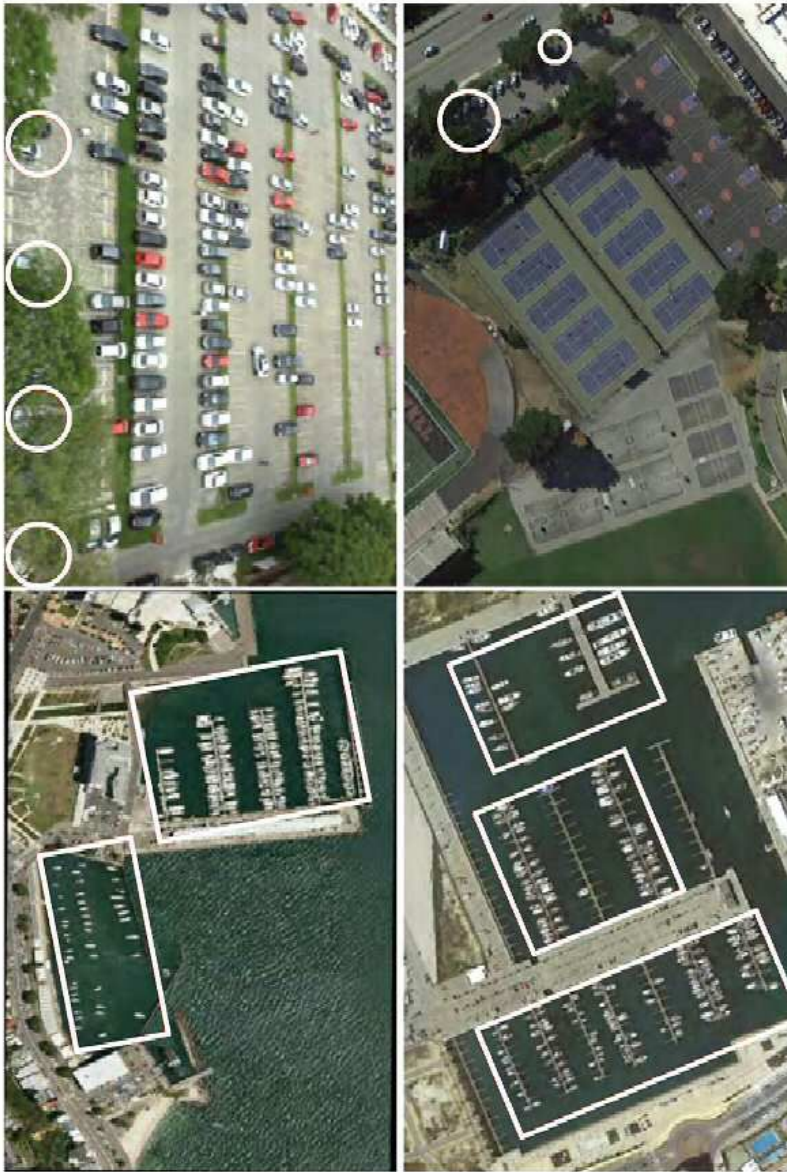
[0056] 이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니 되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

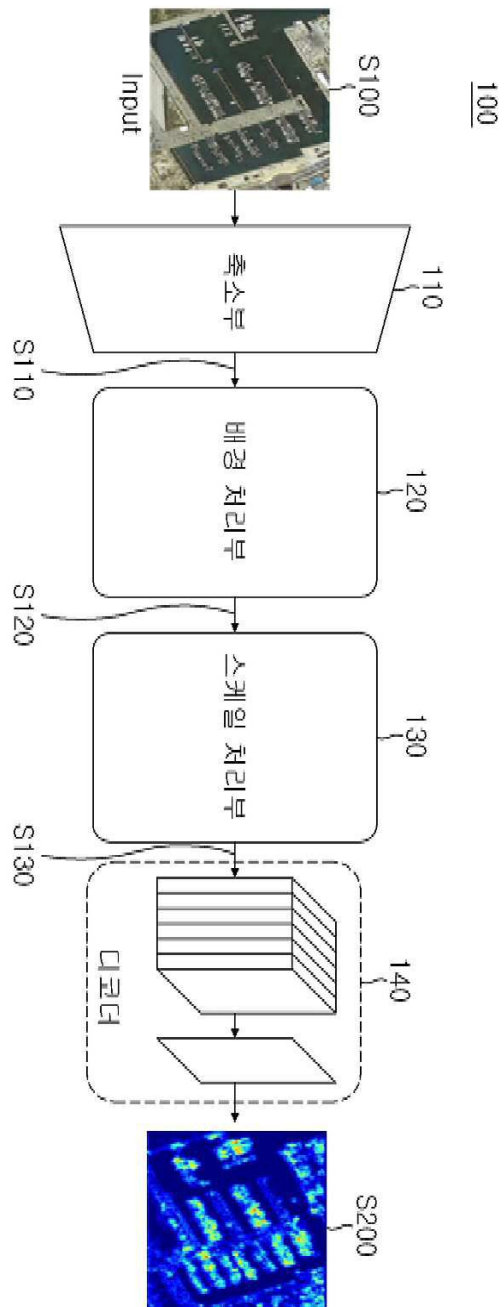
[0057] 축소부(110)
배경 처리부(120)
스케일 처리부 (130)
디코더(140)
다수의 전단 차원변경부(121)
풀링층(122)
합성곱층(123)
제1 정규화층(124)
제2 정규화층(125)
원소별 곱셈값 연산부(126)
복수의 후단 차원변경부(127, 128)
평균 연산부(129)
다수의 1D 컨볼루션 연산부(131)
다수의 딜레이션부(132)
다수의 연결부(133)
다수의 전단 컨볼루션 연산부(134)
후단 연결부(135)
후단 컨볼루션 연산부(136)

도면

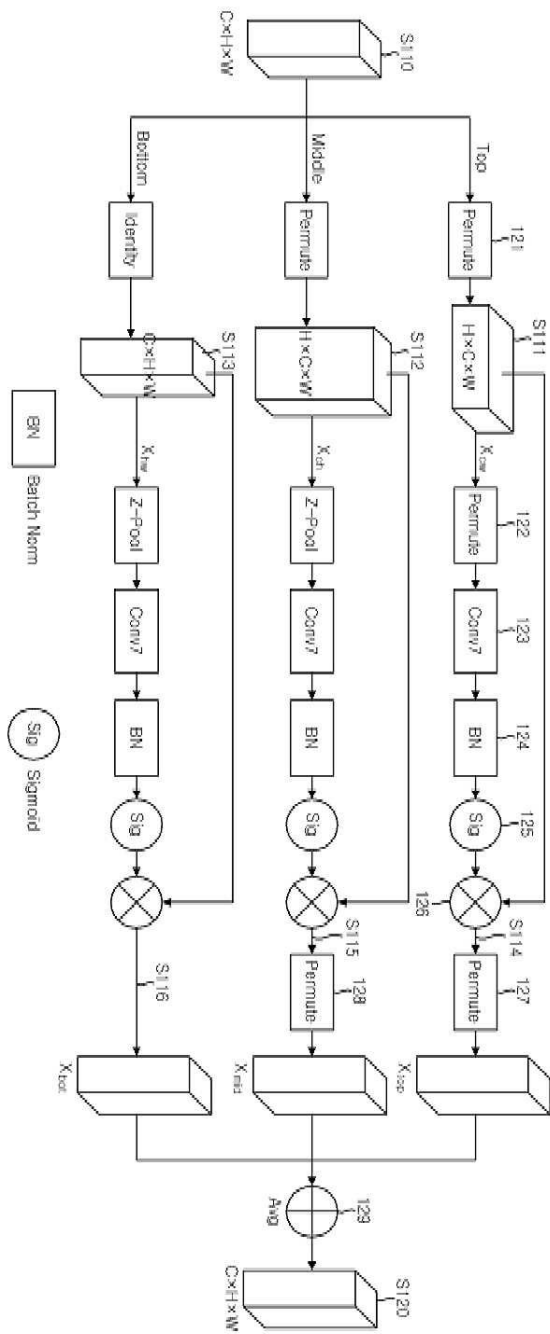
도면1



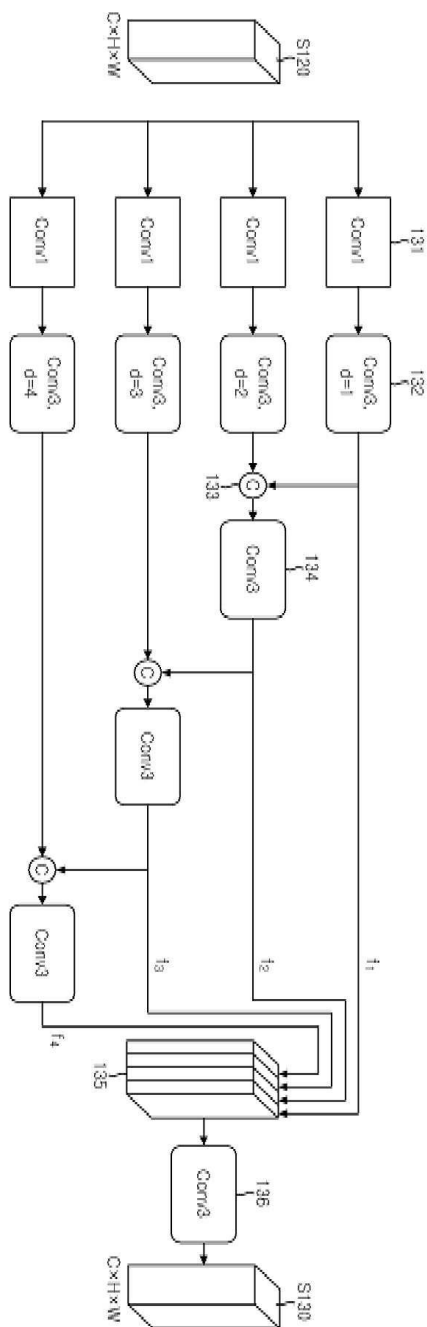
도면2



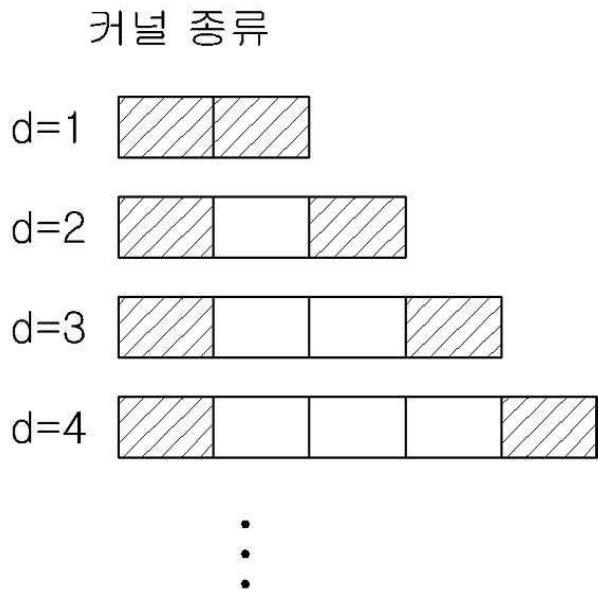
도면3



도면4



도면5



도면6

