



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0085641
(43) 공개일자 2022년06월22일

(51) 국제특허분류(Int. Cl.)

G06N 3/04 (2006.01) G06K 9/62 (2022.01)

G06N 3/08 (2006.01) G06N 5/02 (2006.01)

(52) CPC특허분류

G06N 3/0454 (2013.01)

G06K 9/6267 (2022.01)

(21) 출원번호 10-2020-0175867

(22) 출원일자 2020년12월15일

심사청구일자 2020년12월15일

(71) 출원인

부경대학교 산학협력단

부산광역시 남구 용소로 45, 부경대학교내 (대연동)

(72) 발명자

권기룡

부산광역시 해운대구 좌동순환로 117, 201동 903호 (좌동, 롯데2차아파트)

(74) 대리인

유성원

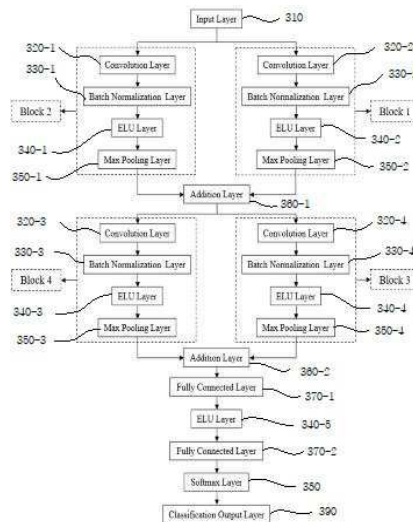
전체 청구항 수 : 총 9 항

(54) 발명의 명칭 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 장치 및 방법

(57) 요약

본 발명은 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 장치 및 방법에 관한 것으로, 더욱 상세하게는 객체를 포함하는 영상을 수집하여 데이터 세트를 생성하는 수집부; 상기 생성된 데이터 세트를 기반으로 병렬 DCNN 모델을 기반으로 학습을 수행하는 학습부; 상기 학습된 병렬 DCNN 모델로부터 병렬 DCNN 모델 프로파일을 추출하는 추출부; 객체가 포함된 영상이 입력되면, 상기 병렬 DCNN 모델 프로파일을 이용하여 상기 입력된 영상에 포함된 객체를 인식하는 인식부; 및 상기 인식된 객체에 대응하는 클래스로 분류하는 분류부를 포함할 수 있다.

대표도 - 도3



(52) CPC특허분류

G06N 3/08 (2013.01)

G06N 5/025 (2019.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711108638
과제번호	2016-0-00318-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	Grand ICT연구센터지원사업
연구과제명	IoT 및 지능정보 기반 동남권 제조 IT 기술 혁신 및 인재양성
기 여 율	1/2
과제수행기관명	부산대학교산학협력단
연구기간	2020.01.01 ~ 2020.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1415166359
과제번호	P0004797
부처명	산업통상자원부
과제관리(전문)기관명	한국산업기술진흥원
연구사업명	혁신도시 공공기관연계 육성사업
연구과제명	첨단 해양산업 오픈랩 구축 및 실감형 융합 콘텐츠 개발
기 여 율	1/2
과제수행기관명	부경대학교 산학협력단
연구기간	2020.01.01 ~ 2020.12.31

명세서

청구범위

청구항 1

객체를 포함하는 영상을 수집하여 데이터 세트를 생성하는 수집부;

상기 생성된 데이터 세트를 기반으로 병렬 DCNN 모델을 기반으로 학습을 수행하는 학습부;

상기 학습된 병렬 DCNN 모델로부터 병렬 DCNN 모델 프로파일을 추출하는 추출부;

객체가 포함된 영상이 입력되면, 상기 병렬 DCNN 모델 프로파일을 이용하여 상기 입력된 영상에 포함된 객체를 인식하는 인식부; 및

상기 인식된 객체에 대응하는 클래스로 분류하는 분류부를 포함하는 것을 특징으로 하는 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 장치.

청구항 2

제1항에 있어서,

상기 병렬 DCNN 모델은,

특징 추출 계층 및 완전 연결 계층을 포함하는 것을 특징으로 하는 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 장치.

청구항 3

제2항에 있어서,

상기 특징 추출 계층은,

네 개의 블록을 포함하되, 각 블록은 컨볼루션 레이어, 배치 정규화 레이어, ELU 레이어 및 최대 풀링 레이어를 각각 포함하는 것을 특징으로 하는 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 장치.

청구항 4

제2항에 있어서,

상기 완전 연결 계층은,

딥샘 레이어, 완전 연결 레이어, 소프트맥스 레이어 및 분류 출력 레이어를 포함하는 것을 특징으로 하는 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 장치.

청구항 5

제1항에 있어서,

상기 병렬 DCNN 모델은,

4개의 컨볼루션 레이어(320-1 내지 320-4), 4개의 배치 정규화 레이어(330-1 내지 330-4), 5개의 ELU 레이어(340-1 내지 340-5), 4개의 최대 풀링 레이어(350-1 내지 350-4), 2개의 딥샘 레이어(360-1 내지 360-2), 2개의 완전 연결(Fully Connected) 레이어(370-1 내지 370-2), 소프트맥스(Softmax) 레이어(380) 및 분류 출력 레

이어(390)를 포함하는 것을 특징으로 하는 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 장치.

청구항 6

객체를 포함하는 영상을 수집하여 데이터 세트를 생성하는 단계;

상기 생성된 데이터 세트를 기반으로 병렬 DCNN 모델을 기반으로 학습을 수행하는 단계;

상기 학습된 병렬 DCNN 모델로부터 병렬 DCNN 모델 프로파일을 추출하는 단계;

객체가 포함된 영상이 입력되면, 상기 병렬 DCNN 모델 프로파일을 이용하여 상기 입력된 영상에 포함된 객체를 인식하는 단계; 및

상기 인식된 객체에 대응하는 클래스로 분류하는 단계를 포함하는 것을 특징으로 하는 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 방법.

청구항 7

제6항에 있어서,

상기 병렬 DCNN 모델은,

특징 추출 계층 및 완전 연결 계층을 포함하는 것을 특징으로 하는 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 방법.

청구항 8

제7항에 있어서,

상기 특징 추출 계층은 네 개의 블록을 포함하되,

각 블록에서는 컨볼루션 레이어, 배치 정규화 레이어, ELU 레이어 및 최대 풀링 레이어를 순차적으로 거쳐 특징이 추출되는 것을 특징으로 하는 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 방법.

청구항 9

제8항에 있어서,

상기 네 개의 블록 중 블록 1 및 블록 2에서는 동일한 영상을 입력으로 하여 특징이 추출된 후 덧셈 레이어를 통해 결합되고, 블록 3 및 블록 4에서는 그 결합된 입력으로 특징이 추출된 후 다른 덧셈 레이어를 통해 결합되는 것을 특징으로 하는 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 방법.

발명의 설명

기술 분야

[0001] 본 발명은 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 장치 및 방법에 관한 것으로, 새로운 병렬 딥 컨볼루션 뉴럴 네트워크를 기반으로 객체를 정확하게 인식 및 분류할 수 있도록 하기 위한 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 장치 및 방법에 관한 것이다.

배경 기술

[0003] 최근 새로운 영상 미디어 서비스 기술이 발전함에 따라, 기존에 방송 콘텐츠를 단순히 제공하던 것에서 벗어나서 해당 콘텐츠의 내용을 분석하여 관련된 광고나 연계 서비스를 고객에게 제공하려는 노력을 하고 있다. 이를

위해서는 영상 콘텐츠에 나오는 인물이나 사물을 인식하는 것이 중요하므로, 영상 기반의 객체 검출 기술이 필수적으로 요구되고 있다.

- [0004] 이를 위해서는 그 객체가 무슨 클래스인지 인식(Recognition)하고 분류(Classification)해야만 한다.
- [0005] 그렇기 때문에 영상 인식 및 분류는 컴퓨터 비전의 최신 주제로 간주된다. 그러나, 자동화 시스템에서는 예를 들어, 동물이 포함된 영상을 통해 인식 및 분류하도록 하는 경우, 동물이 장면에 개별적으로 나타나는 것이 아니라 주변의 모든 자연 물체를 차지한다는 점에서 문제가 존재한다. 따라서, 컴퓨터의 경우 먼저 동물의 모습을 수정한 다음 사람이 동물에만 쉽게 집중할 수 있는 곳으로 분류해야한다. 게다가 동물은 다른 조명 조건, 관점, 스케일 및 모양에서 나타날 수 있다.
- [0006] 인간의 얼굴을 분류하는 방법은 매우 정확하게 확립되었지만 동물 분류의 경우 복잡한 클래스 내 가변성과 클래스 간 유사성을 가진 동물 클래스의 다양성으로 인해 이러한 방법이 정확하지 않은 것으로 입증되었다.
- [0007] 많은 연구자들은 각각의 장점과 단점을 가지고 이러한 문제를 해결하기 위해 여러 가지 접근 방식을 시도하였다. 그러나 최근 몇 년 동안 컨볼루션 뉴럴 네트워크(Convolutional Neural Networks, CNN) 기반 분류 방법이 많은 주목을 받고 있다. CNN 계층이 다수의 계층(multi layer)을 가지고 있으면 딥 컨볼루션 뉴럴 네트워크(Deep Convolutional Neural Networks, DCNN)라고 한다. 이러한 네트워크로는, 예를 들어, Alexnet, GoogleNet, VGG19 등이 있다.
- [0008] 그러나 이러한 네트워크들은 특정 종류에 관계없이 분류 작업에 완전히 집중되어 있으며, 일반적인 형태의 CNN을 따르고 높은 정확도를 달성하였다.
- [0009] 따라서, 영상(비디오)을 기반으로 객체를 인식 및 분류함에 있어서, 이러한 기존의 방법에 비해 더욱 뛰어난 정확도를 보여줄 수 있도록 하는 기술이 개발될 필요가 있다.

발명의 내용

해결하려는 과제

- [0011] 따라서, 본 발명은 상기한 바와 같은 문제점을 해결하기 위하여 제안된 것으로, 배치(batch) 정규화 계층 및 ELU (Exponential Linear Unit) 계층과 같은 업데이트 된 계층으로 구성된 새로운 병렬 DCNN을 기반으로 영상 데이터의 특징을 추출 및 분석하여 객체를 보다 정확하게 인식 및 분류할 수 있도록 하는 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 장치 및 방법을 제공함에 있다.
- [0012] 본 발명의 목적은 이상에서 언급한 것으로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 본 발명이 속하는 기술 분야의 통상의 지식을 가진 자에게 명확히 이해될 수 있을 것이다.

과제의 해결 수단

- [0014] 상기와 같은 목적을 달성하기 위한 본 발명에 따른 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 장치는, 객체를 포함하는 영상을 수집하여 데이터 세트를 생성하는 수집부; 상기 생성된 데이터 세트를 기반으로 병렬 DCNN 모델을 기반으로 학습을 수행하는 학습부; 상기 학습된 병렬 DCNN 모델로부터 병렬 DCNN 모델 프로파일을 추출하는 추출부; 객체가 포함된 영상이 입력되면, 상기 병렬 DCNN 모델 프로파일을 이용하여 상기 입력된 영상에 포함된 객체를 인식하는 인식부; 및 상기 인식된 객체에 대응하는 클래스로 분류하는 분류부를 포함할 수 있다.
- [0015] 또한, 본 발명에 따른 병렬 딥 컨볼루션 뉴럴 네트워크 기반의 영상 분류 방법은, 객체를 포함하는 영상을 수집하여 데이터 세트를 생성하는 단계; 상기 생성된 데이터 세트를 기반으로 병렬 DCNN 모델을 기반으로 학습을 수행하는 단계; 상기 학습된 병렬 DCNN 모델로부터 병렬 DCNN 모델 프로파일을 추출하는 단계; 객체가 포함된 영상이 입력되면, 상기 병렬 DCNN 모델 프로파일을 이용하여 상기 입력된 영상에 포함된 객체를 인식하는 단계; 및 상기 인식된 객체에 대응하는 클래스로 분류하는 단계를 포함할 수 있다.

발명의 효과

- [0017] 본 발명에 의하면, 배치(batch) 정규화 계층 및 ELU (Exponential Linear Unit) 계층과 같은 업데이트 된 계층으로 구성된 새로운 병렬 DCNN을 기반으로 영상(비디오) 데이터의 특징을 추출 및 분석하여 객체를 보다 정확하게 인식 및 분류할 수 있도록 한다.
- [0018] 본 발명의 효과는 이상에서 언급한 것으로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 본 발명이 속하는 기술 분야의 통상의 지식을 가진 자에게 명확히 이해될 수 있을 것이다.

도면의 간단한 설명

- [0020] 도 1은 일반적인 선형 DCNN 모델의 구조를 개략적으로 나타내는 도면이다.
- 도 2는 일반적인 병렬 DCNN 모델의 구조를 개략적으로 나타내는 도면이다.
- 도 3은 본 발명의 실시예에 따른 새로운 병렬 DCNN 모델의 구조를 개략적으로 나타내는 도면이다.
- 도 4는 본 발명의 실시예에 따른 새로운 병렬 DCNN 모델을 기반으로 영상을 분류하는 장치를 나타내는 블록도이다.
- 도 5는 본 발명의 실시예에 따른 새로운 병렬 DCNN 모델을 기반으로 영상을 분류하는 방법을 나타내는 순서도이다.
- 도 6는 동물 얼굴 데이터 세트의 일 예를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0021] 본 발명의 목적 및 효과, 그리고 그것들을 달성하기 위한 기술적 구성들은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 본 발명을 설명함에 있어서 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략할 것이다. 그리고 후술되는 용어들은 본 발명에서의 기증을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다.
- [0022] 그러나, 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있다. 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.
- [0023] 한편, 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 부재를 사이에 두고 "간접적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 구비할 수 있다는 것을 의미한다.
- [0024] 도 1은 일반적인 선형 DCNN 모델의 구조를 개략적으로 나타내는 도면이다.
- [0025] 도 1을 참조하면, 일반적인 선형 DCNN 모델의 구조는 “선형 DCNN + 배치 정규화 + ELU”로 구성될 수 있다. 이 구조는 배치 정규화 레이어와 ELU 레이어로 DCNN의 선형 형태를 구성한다. 이 구조에서는 2 개의 컨볼루션 레이어, 2 개의 배치 정규화 레이어, 3 개의 ELU 레이어, 2 개의 최대 풀링 레이어, 2 개의 완전 연결 레이어 및 각 입력 레이어, 소프트맥스 레이어, 분류 출력 레이어 중 하나를 사용한다.
- [0026] 여기서, 첫 번째 컨볼루션 레이어와 첫 번째 최대 풀링 레이어는 동일한 차원 5×5와 보폭 4×4로 컨볼루션 함수와 풀링 함수를 적용한다. 이후, 두 번째 컨볼루션 레이어와 두 번째 최대 풀링 레이어도 각각 차원이 3×3이고 보폭이 2×2 인 유사한 컨볼루션 함수와 풀링 함수를 적용한다. 이후, 512 개의 뉴런이 있는 완전 연결 계층은 출력을 ELU 레이어, 완전 연결 레이어 및 소프트맥스 레이어를 통해 분류 출력 레이어로 전달한다.
- [0027] 이 선형 DCNN에서 레이어는 특징점 추출을 위해 선형 모양을 형성하는 차례로 스택된다. 하이퍼 매개 변수 설정에 따라 이러한 레이어는 가중치를 생성한다. 단방향 특성 추출이므로 DCNN은 해당 특성만 학습한다. 예를 들어, Guobin Chen은 컨볼루션 레이어와 최대 풀링 레이어의 첫 번째 세트는 컨볼루션 커널 크기 9×9와 풀링

커널 크기 2×2 를 적용하여 60×60 차원의 32 개의 피쳐 맵을 생성한다. 이 선형 구조에서 DCNN에는 하이퍼 매개 변수 설정에 다양성이 없기 때문에 학습할 가중치만 있다.

[0028] 도 2는 일반적인 병렬 DCNN 모델의 구조를 개략적으로 나타내는 도면이다.

[0029] 도 2를 참조하면, 일반적인 병렬 DCNN은 “병렬 DCNN + 교차 채널 정규화(Cross-Channel Normalization) + ReLU” 로 구성될 수 있다. 여기에는 배치 정규화 레이어(Batch Normalization layer)과 ELU 레이어 대신 교차 채널 정규화 레이어와 ReLU 레이어로 구성되어 있다. 이 구조는 선형 DCNN보다 더 효율적이고 정확하다. '병렬'이라는 용어는 이 구조에서 비롯된 것이다. 도 2에서 블록 1과 블록 2는 동일한 입력을 공유한다. 그들은 비슷한 방식으로 특징 추출을 실행하고, 결국 덧셈 레이어에 의해 결합된다. 이 동일한 동작 이후, 블록 3과 블록 4가 이어진다. 이 구조는 4 개의 컨볼루션 레이어, 4 개의 교차 채널 정규화 레이어, 5 개의 ReLU 레이어, 4 개의 최대 풀링 레이어, 2 개의 덧셈 레이어, 2 개의 완전 연결 레이어 및 각 입력 레이어, 소프트맥스 레이어, 분류 출력 레이어 중 하나를 사용하였다. 교차 채널 정규화 레이어의 창 크기는 5로 설정되었다. 이는 이중구조(dualism)와 유사하므로 병렬 DCNN으로 표시하였다. 각 블록 내부 동작은 선형 DCNN과 같이 컨볼루션, 교차 채널 정규화, ReLU 과정을 순차적으로 수행한다.

[0030] 본 발명은 객체 간의 유사도 산출 결과에 따라 데이터 세트(클러스터)를 생성한 후에 그 데이터 세트를 활용하여 인공지능 기반으로 객체 분류 모듈을 학습하고, 영상에서 객체를 식별하면 그 식별된 객체의 정보를 객체 분류 모듈에 입력하여 어떠한 클래스에 귀속하는지를 신속하게 결정하는지에 관한 것이다.

[0031] 도 3은 본 발명의 실시예에 따른 새로운 병렬 DCNN 모델의 구조를 개략적으로 나타내는 도면이다.

[0032] 도 3을 참조하면, 본 발명의 실시예에 따른 새로운 병렬 DCNN 모델은 24개의 레이어들을 포함할 수 있다. 전체 레이어는 입력 레이어(310), 4개의 컨볼루션 레이어(320-1 내지 320-4), 4개의 배치 정규화 레이어(330-1 내지 330-4), 5개의 ELU 레이어(340-1 내지 340-5), 4개의 최대 풀링 레이어(350-1 내지 350-4), 2개의 덧셈 레이어(360-1 내지 360-2), 2개의 완전 연결(Fully Connected) 레이어(370-1 내지 370-2), 소프트맥스(Softmax) 레이어(380) 및 분류 출력 레이어(390)를 포함한다.

[0033] 이 새로운 병렬 DCNN 모델은 크게 특징 추출 계층 및 완전 연결 계층의 두 부분으로 나눌 수 있다. 구체적으로, 블록 1 내지 블록 4는 특징 추출 계층을 나타내고, 나머지는 완전 연결 계층(fully connected layer)를 나타낸다.

[0034] 여기서, 컨볼루션 레이어(320-1 내지 320-4)는 영상을 분석하는데 사용되는 작은 필터를 이미지의 몇몇 픽셀에 매칭하여 곱하는 레이어로 필터의 크기 및 필터가 적용되는 간격을 나타내는 보폭(Stride) 값 등으로 정의된다. ELU 레이어(340-1 내지 340-5)는 중요한 역할을 하는 픽셀에 대하여 그 중요도에 비례하는 값을 곱하여 각 픽셀에 대하여 서로 다른 가중치를 주는 레이어이다. 최대 풀링 레이어(350-1 내지 350-4)는 특정 범위의 픽셀 중 가장 중요한(가중치가 높은) 픽셀을 선택하여 대표값으로 하거나 특정 범위의 픽셀들의 평균을 구하는 것에 의하여 그 특정 범위의 픽셀들을 통합하는 것으로, 특징맵을 감소시켜 네트워크를 경량화하는 레이어이다.

[0035] 도 3를 참조하여 구체적으로 설명하면, 새로운 병렬 DCNN 모델은 227×227 크기의 입력 레이어(310)로 시작한다. 입력 레이어(310) 바로 뒤에서 병렬 DCNN은 서로 병렬로 작동하는 4 가지 특징 추출 계층 집합을 따른다. 도 3에서는 설명을 위해 특징 추출 계층의 네 세트를 블록 1, 블록 2, 블록 3 및 블록 4로 표시했다. 각 블록에는 각각 컨볼루션 레이어, 배치 정규화 레이어, ELU 레이어 및 최대 풀링 레이어가 포함된다.

[0036] 먼저, 블록 1에서는 컨볼루션 계층(320-1)은 차원이 5×5 이고, 보폭(stride)가 4×4 인 컨볼루션 함수를 적용한다. 이것은 $56 \times 56 \times 256$ 의 출력을 생성한다. 이후, 배치 정규화 레이어(330-1)가 그 출력을 정규화하여 그 정규화 된 기능은 최대 유용한 기능 수가 추출되는 ELU 레이어(340-1)로 전달된다. 이후, 최대 풀링 레이어(350-1)는 차원이 5×5 이고 보폭이 4×4 인 ELU 레이어의 출력에서 피쳐를 풀링한다. 이것은 $13 \times 13 \times 256$ 의 출력을 생성한다.

[0037] 블록 2에서는 컨볼루션 레이어(320-2)가 차원이 7×7 이고, 보폭이 4×4 인 컨볼루션 함수를 적용하는 것을 제외하고는 블록 1에서와 동일한 작업이 수행된다. 이는 $57 \times 57 \times 256$ 의 출력을 생성한다. 또한, 블록 1과 마찬가지로 그 출력은 배치 정규화 레이어(330-2)과 ELU 레이어(340-2)를 통과한다. 이후, 최대 풀링 레이어(350-2)는 차원이 13×13 이고 보폭이 4×4 인 풀링 함수를 적용하여 $13 \times 13 \times 256$ 의 출력을 생성한다.

[0038] 다음으로, 덧셈 레이어(360-1)는 블록 1의 최대 풀링 레이어(350-1) 및 블록 2의 최대 풀링 레이어(350-2)의 출력을 요소별로 추가하고 $13 \times 13 \times 256$ 의 출력을 생성한다. 이후 피쳐 맵의 수가 두배가 되는 것을 제외하고는 첫

번째 레이어 세트와 유사한 요소를 포함하는 피처 추출 레이어의 다음 세트로 출력을 전달한다.

- [0039] 블록 3에서 컨볼루션 계층(320-3)은 차원이 3×3 이고 보폭이 2×2 인 컨볼루션 함수를 적용하여 512 개의 특징 맵을 생성한다. 이것은 $6 \times 6 \times 512$ 의 출력을 생성하고 배치 정규화 레이어(330-3)에 의해 정규화된다. 이러한 정규화 된 기능은 ELU 레이어(340-3) 통과한다. 이전과 마찬가지로 이 ELU 레이어(340-3)는 유용한 기능만 추출되어 다음 계층으로 전달되도록 한다. 최대 풀링 레이어는(350-3)은 1×1 의 보폭으로 풀링 함수를 적용하고 $4 \times 4 \times 512$ 의 출력을 생성한다.
- [0040] 반면에, 블록 4에서 컨볼루션 계층(320-4)는 차원이 5×5 이고 보폭이 2×2 인 컨볼루션 함수를 적용하여 유사한 동작을 수행한다. 이 동작은 $7 \times 7 \times 512$ 의 출력을 생성하고, 배치 정규화 레이어(330-4)와 ELU 레이어(340-4)를 통과한다. 최대 풀링 레이어(350-4)는 ELU 레이어(340-4)의 출력을 가져와서 차원이 7×7 이고 보폭이 4×4 인 풀링 함수를 적용하여 $4 \times 4 \times 512$ 의 출력을 생성한다.
- [0041] 다음으로, 덧셈 레이어(360-2)는 블록 3의 최대 풀링 레이어(350-3) 및 블록 4의 최대 풀링 레이어(350-4)의 출력을 요소별로 추가하고, $4 \times 4 \times 512$ 의 출력을 생성한다. 이 시점에서 특징 추출 작업이 완료되고 첫 번째 완전 연결 레이어(에 대해 전달할 준비가 된다.
- [0042] 첫 번째 완전 연결 레이어(370-1)는 두 번째 덧셈 레이어(360-2)의 출력을 가져와 벡터로 바꾼다. 이 벡터는 첫 번째 완전 연결 레이어(370-1)의 512개 뉴런이 차지한 다음 세 번째 및 마지막 ELU 레이어(340-5 및 340-6)로 전달된다. ELU 기능을 적용한 후, 10 개의 최상의 기능만 전달되고, 두 번째 완전 연결 계층(370-2)의 10 개 뉴런이 점유한다. 그런 다음 소프트맥스 레이어(380)는 두 번째 완전 연결 레이어(370-2)의 출력에 소프트맥스 함수를 적용하고, 마지막으로 이를 분류 출력 레이어(390)로 전달한다. 분류 출력 레이어(390)는 가장 활성화 된 기능을 결정하고 해당 클래스 이름으로 레이블을 지정한다.
- [0043] 이러한 새로운 병렬 DCNN 모델을 이용하면, 낮은 컴퓨팅 비용으로 정확도가 높은 네트워크를 만들 수 있다. 선형 DCNN보다 높은 정확도와 효율성을 갖는 병렬 DCNN이며, 또한 배치 정규화 레이어 및 지수 선형 단위(ECL) 레이어를 사용하여 최첨단 레이어로 구축하였다. 대규모 데이터 세트를 다루기 때문에 데이터 세트를 일반화하고 해당 레이어를 사용하여 가장 강력한 기능을 갖도록 한다.
- [0044] 이 새로운 병렬 DCNN 모델의 구조에서는 컨볼루션, 배치 정규화 및 ELU 과정으로 수행되는데, 이 과정에서 순차적으로 수행하는 것이 아니라 배치로서 묶음 단위로 수행하기 때문에 계산시간과 정확도를 향상시킬 수 있다.
- [0045] 도 4는 본 발명의 실시예에 따른 새로운 병렬 DCNN 모델을 기반으로 영상을 분류하는 장치를 나타내는 블록도이다.
- [0046] 도 4를 참조하면, 새로운 병렬 DCNN 모델을 기반으로 영상을 분류하기 위한 영상 분류 장치(100)는 수집부(101), 학습부(103), 추출부(105), 인식부(107) 및 분류부(109)를 포함한다.
- [0047] 수집부(101)는 객체를 포함하는 영상을 수집하여 데이터 세트를 생성한다. 이때, 데이터 세트는 복수개의 영상을 포함한다.
- [0048] 학습부(103)는 수집부(101)를 통해 생성된 데이터 세트를 기반으로 병렬 DCNN 모델을 기반으로 학습을 수행한다. 여기서, 병렬 DCNN 모델은 앞서도 3을 참조하여 설명한 새로운 병렬 DCNN 모델이다.
- [0049] 추출부(105)는 학습부(103)를 통해 학습된 병렬 DCNN 모델로부터 병렬 DCNN 모델 프로파일을 추출한다.
- [0050] 인식부(107)는 객체가 포함된 영상이 입력되면, 학습부(103)에 의해 추출된 병렬 DCNN 모델 프로파일을 이용하여 상기 입력된 영상에 포함된 객체를 인식한다.
- [0051] 분류부(109)는 인식부(107)에 의한 인식 결과에 따라 해당 객체에 대응하는 클래스로 분류한다.
- [0052] 도 5는 본 발명의 실시예에 따른 새로운 병렬 DCNN 모델을 기반으로 영상을 분류하는 방법을 나타내는 순서도이다.
- [0053] 도 5를 참조하면, 먼저, 수집부(101)가 검출하고자 하는 객체를 포함하는 영상을 수집하여 데이터 세트를 생성하고(S201), 학습부(103)가 그 생성된 데이터 세트를 기반으로 앞서 설명한 새로운 병렬 DCNN 모델을 기반으로 학습을 수행한다(S203).
- [0054] 추출부(105)가 그 학습된 병렬 DCNN 모델로부터 병렬 DCNN 모델 프로파일을 추출한다. 이후, 객체가 포함된 영상이 입력되면, 인식부(107)가 상기 병렬 DCNN 모델 프로파일을 이용하여 상기 입력된 영상에 포함된 객체를 인

식하고(S207), 분류부(109)가 그 인식된 객체에 대응하는 클래스로 분류한다(S209).

- [0055] 이와 같이 객체를 인식 및 분류하기 위해서는 먼저 데이터 세트를 생성하는 것으로부터 시작된다. 이하에서는 도 6을 참조하여 하나의 실시예로서 동물의 개체를 인식 및 분류하기 위한 경우에 대해 설명하고, 그 실험 결과에 대해 설명한다. 이러한 실험을 위해 사용한 작업은 Intel Core i7CPU 3.40GHz, 8GB RAM 및 단일 NVIDIA GeForce GTX 1060 3GB GPU가있는 시스템에서 수행하였다.
- [0056] 도 6는 동물 얼굴 데이터 세트의 일 예를 나타내는 도면으로서, 본 발명에 따른 병렬 DCNN 모델을 이용하여 동물 개체를 인식 및 분류하였을 때의 정확성을 확인하기 위한 실험을 진행하기 위한 동물 얼굴 데이터 세트의 일 예이다.
- [0057] 이 데이터 세트는 Google 검색 및 기타 여러 웹 사이트와 같은 다양한 비상업적 출처에서 이미지를 수집하였다. 이 데이터 세트에는 곰, 고양이, 사슴, 개, 코끼리, 말, 사자, 토끼, 너구리, 쥐 등 10 가지 종류의 동물이 포함되어, 각 동물 클래스에는 1000 개의 이미지가 포함되어 있다. 이 수집된 이미지들에는 다양한 배경과 여러 다른 동물 또는 얼굴이 아닌 전신에 집중된 이미지들이 포함되어 있어 이를 수정하고 이미지를 잘라서 얼굴만을 편집하였다. 동물 얼굴에도 변형이 있었다. 예를 들어, 일부 사진은 정면도, 일부는 측면도 등을 보여준다. 본 발명에서는 얼굴을 다루는 것이기 때문에 모든 이미지를 정면 위치에서 촬영하여 약간의 측면 움직임은 쉽게 처리했다. 일부 동물 이미지는 차원이 다르다. 따라서 네트워크는 동일한 크기의 이미지 만 가져 오기 때문에 모든 이미지의 크기가 $227 \times 227 \times 3$ 으로 조정되었다.
- [0058] 동물 인식 시스템의 정확도는 이미지 데이터 세트의 품질에 따라 다르다. 기능이 다양할수록 동물 분류 시스템의 정확도가 높아진다. 특히 색상, 모서리 및 모서리의 차이는 클래스 간에 구분된다. 따라서 데이터 세트의 이미지를 선택하는 동안 정제된 품질을 유지하였다.
- [0059] 이러한 도 6은 이러한 작업을 통해 수집된 동물 데이터 세트를 나타내는 것으로, 이 데이터 세트에는 40 개의 예시 이미지를 포함하여 보여준다.
- [0060] 본 실험을 위해 Tibor Trnovszky 및 Guobin Chen의 작업과 실험하고 비교하였다. 또한 제안한 방법에 도달하기 위해 여러 시뮬레이션을 거쳤다. 이러한 시뮬레이션은 선택할 매개 변수 또는 구조를 결정하는데 도움이 되었다. 이러한 모든 시뮬레이션 중에서 "선형 DCNN + 배치 정규화 + ELU" 및 "병렬 DCNN + 교차 채널 정규화 + ReLU"라는 레이블이 붙은 두 가지 구조와 비교하였다.
- [0061] 이 모든 네트워크의 시뮬레이션을 위해 동일한 원칙을 따랐다. 내장된 데이터 세트를 사용하여 모든 네트워크에서 시뮬레이션을 수행하였다. 데이터 세트는 훈련 데이터 세트(training dataset) 및 검증/테스트 데이터 세트(alidation/testing dataset)를 위해 80 %에서 20 %로 나뉜다. 훈련 데이터 세트에는 총 8000 개의 훈련 이미지가 포함되었고, 유효성 검사 / 테스트 데이터 세트에는 총 2000 개의 검증 / 테스트 이미지가 포함되었다. 분할 후 훈련 데이터 세트에 이미지 확대(image augmentation)를 적용하였다. 이미지 확대는 크기 조정, 회전, 반사, 전단 및 변환 변환의 무작위 조합으로 이미지를 확대하는데 사용되어 결국 네트워크의 성능을 향상시킨다. 이미지 증강은 훈련 과정에서 다양성을 가져오고 DCNN에 더 나은 학습 환경을 제공했다. 모든 시뮬레이션에 대해 SGDM (Stochastic Gradient Descent with Momentum)을 사용하고 초기 학습률을 0.0001로 설정했다. max epoch를 300으로, mini-batch 크기를 32로 설정하였다. 훈련 중 데이터 세트에도 서플링이 사용되었다. 이하의 <표 1>은 앞서 설명한 새로운 병렬 DCNN의 시뮬레이션 결과를 나타낸다.

표 1

Confusion Matrix											
		Target Class									
		Bear	Cat	Deer	Dog	Elephant	Horse	Lion	Rabbit	Raccoon	Rat
Output Class	Bear	179 8.9%	0 0.0%	0 0.0%	2 0.1%	3 0.1%	2 0.1%	1 0.1%	0 0.0%	2 0.1%	0 0.0%
	Cat	1 0.1%	170 8.5%	2 0.1%	5 0.3%	4 0.2%	0 0.0%	2 0.1%	1 0.1%	0 0.0%	2 0.1%
	Deer	0 0.0%	5 0.3%	179 8.9%	2 0.1%	1 0.1%	0 0.0%	0 0.0%	2 0.1%	0 0.0%	5 0.3%
	Dog	2 0.1%	13 0.7%	2 0.1%	182 9.1%	4 0.2%	5 0.3%	2 0.1%	4 0.2%	2 0.1%	2 0.1%
	Elephant	1 0.1%	2 0.1%	5 0.3%	2 0.1%	185 9.3%	4 0.2%	0 0.0%	1 0.1%	0 0.0%	0 0.0%
	Horse	8 0.4%	4 0.2%	6 0.3%	2 0.1%	1 0.1%	189 9.4%	1 0.1%	3 0.1%	1 0.1%	1 0.1%
	Lion	2 0.1%	0 0.0%	1 0.1%	1 0.1%	1 0.1%	0 0.0%	194 9.7%	0 0.0%	0 0.0%	0 0.0%
	Rabbit	1 0.1%	4 0.2%	2 0.1%	1 0.1%	1 0.1%	0 0.0%	0 0.0%	183 9.2%	0 0.0%	5 0.3%
	Raccoon	0 0.0%	0 0.0%	2 0.1%	2 0.1%	0 0.0%	0 0.0%	0 0.0%	1 0.1%	195 9.8%	2 0.1%
	Rat	6 0.3%	2 0.1%	1 0.1%	1 0.1%	0 0.0%	0 0.0%	0 0.0%	5 0.3%	0 0.0%	183 9.2%
Total		89.5% 10.5%	85.0% 15.0%	89.5% 10.5%	91.0% 9.0%	92.5% 7.5%	94.5% 5.5%	97.0% 3.0%	91.5% 8.5%	97.5% 2.5%	91.5% 8.5%

[0062]

[0063]

이 <표 1>은 모든 개별 클래스의 정확도와 함께 총 검증 정확도를 보여준다. 이러한 검증 정확도는 훈련된 DCNN을 20 % 검증/테스트 데이터 세트에 적용했을 때의 결과를 보여준다. 여기서 전체 정확도는 92.0 %, 라쿤 클래스는 97.5 %로 가장 높은 정확도를, Cat 클래스는 85.0 %로 가장 낮은 정확도를 나타냈다.

[0064]

우리는 관련 작업의 동일한 네트워크 구성을 따라 직접 구축했다. 제안한 병렬 DCNN과 같이 모든 네트워크에 대해 이상적인 훈련 조건을 유지했다. 모든 관련 DCNN의 구축 및 훈련을 마친 후 검증/테스트 데이터 세트로 테스트했다. 제안한 병렬 DCNN 방법으로 성공적으로 인식된 동물의 수는 하기 <표 2>와 같다.

표 2

	Bear	Cat	Deer	Dog	Elephant	Horse	Lion	Rabbit	Raccoon	Rat
DCNN by Guobin Chen et al [6].	93	143	164	146	145	109	173	136	192	76
DCNN by Tibor Trnovszky et al [5].	131	145	174	112	164	164	171	148	175	114
Proposed parallel DCNN + Cross Channel Normalization + ReLU	174	169	165	180	171	182	188	181	190	178
Linear DCNN + Batch Normalization + ELU	176	170	178	181	175	187	192	179	191	178
New proposed parallel DCNN + Batch Normalization + ELU Model	187	177	177	177	173	191	194	183	193	176

[0065]

[0066]

<표 2>를 통해 볼 때, 제안한 방법은 기존의 다른 방법보다 더 많은 동물을 인식하였다. Guobin Chen의 방법에서 가장 많이 인식 된 클래스는 Raccoon (200 개 중 192 개)이고 가장 덜 인식 된 클래스는 Rat (200 개 중 76 개 이미지)이다. 반면 Tibor Trnovszky의 방법에서 ??가장 많이 인식 된 클래스는 Raccoon (200 개 중 175 개)이었으며 가장 적게 인식 된 클래스는 Dog (200 개 중 112 개 이미지)였다. “병렬 DCNN + 교차 채널 정규화 + ReLU”에서 가장 많이 인정받은 클래스는 200 개의 이미지 중 190 개를 가진 Raccoon이었고, 가장 덜 인식 된 클래스는 200 개의 이미지 중 165 개를 가진 Deer였다. 마지막으로 “선형 DCNN + 배치 정규화(Batch Normalization) + ELU”에서 가장 많이 인식 된 클래스는 200 개 이미지 중 192 개를 가진 Lion이었고, 가장 적게 인식 된 클래스는 200 개 중 170 개 이미지를 가진 Cat이었다. 새롭게 제안한 “병렬 DCNN + 배치 정규화 + ELU”에서 가장 많이 인식 된 클래스는 200 개 이미지 중 194 개를 가진 Lion이었고, 가장 적게 인식 된 클래스

는 200 개 중 173 개 이미지를 가진 Elephant이었다. 이를 비교했을 때에 제안한 "병렬 DCNN + 배치 정규화 + ELU"가 가장 우수한 인식 구조임을 알 수 있었다.

[0067] 한편, 하기 <표 3>은 이 제안한 방법과 함께 모든 관련 DCNN 방법의 전체 정확도를 보여준다.

표 3

Methods	Accuracy%
DCNN of Guobin Chen et al [6].	68.8%
DCNN of Tibor Trnovszky et al [5].	74.9%
Proposed parallel DCNN + Cross Channel Normalization + ReLU	88.9%
Linear DCNN + Batch Normalization + ELU	90.3%
New proposed parallel DCNN + Batch Normalization + ELU	92.0%

[0068]

[0069] <표 3>을 통해 볼 때, Guobin Chen의 방법은 68.8 %를 달성했고, Tibor Trnovszky의 방법은 74.9 %를 달성했다. 또한 “병렬 DCNN + 교차 채널 정규화 + ReLU” 는 88.9 %, “선형 DCNN + 배치 정규화 + ELU” 는 90.3 %이고, 제안한 “병렬 DCNN + 배치 정규화 + ELU” 방법은 92.0 %를 달성하였다. 기존의 방법보다 정확도가 매우 우수함을 알 수 있었다.

[0070] 앞서 설명한 본 발명은 기존의 선형 DCNN과 비교하여 볼 때, 새로운 계층, 즉 배치 정규화 계층, ELU (Exponential Linear Unit) 계층 중 일부를 포함하는 새로운 병렬 DCNN을 제안하고 있는 것이다. 이들에 포함된 레이어의 효과는 출력 결과를 비교하여 실현할 수 있으며, 교차 채널 정규화보다 배치 정규화의 우월성은 결과에서 확인할 수 있다. 즉, 새로운 병렬 DCNN은 모든 관련 기존 방법보다 더 우수한 성능을 가지고 있으며, 또한 병렬 DCNN 네트워크의 생성은 선형 DCNN으로는 불가능한 더 강력한 기능을 추출하는데 도움이 된다. 즉, 전반적으로 업데이트 된 레이어와 병렬 구조 DCNN을 사용하면 다른 방법에 비해 성능이 더 우수하다는 것을 알 수 있다.

[0071] 본 명세서와 도면에는 본 발명의 바람직한 실시예에 대하여 개시하였으며, 비록 특정 용어들이 사용되었으나, 이는 단지 본 발명의 기술 내용을 쉽게 설명하고 발명의 이해를 돕기 위한 일반적인 의미에서 사용된 것이지, 본 발명의 범위를 한정하고자 하는 것은 아니다. 여기에 개시된 실시예 외에도 본 발명의 기술적 사상에 바탕을 둔 다른 변형예들이 실시 가능하다는 것은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 자명한 것이다.

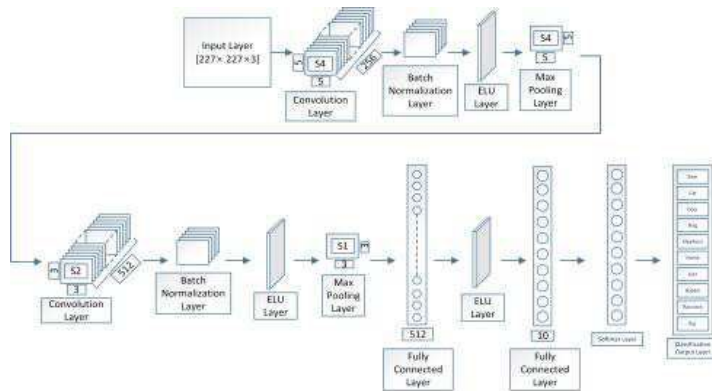
부호의 설명

[0073]

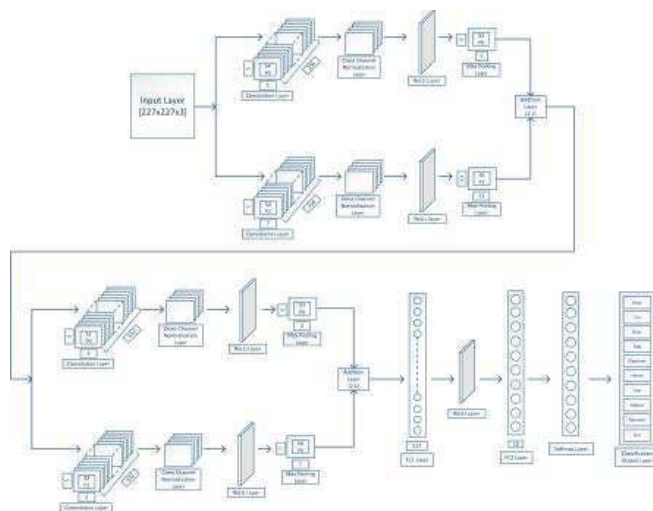
100: 영상 분류 장치	101: 수집부
103: 학습부	105: 추출부
107: 인식부	109: 분류부
310: 입력 레이어	
320-1 내지 320-4: 컨볼루션 레이어	
330-1 내지 330-4: 배치 정규화 레이어	
340-1 내지 340-5: ELU 레이어	
350-1 내지 350-4: 최대 풀링 레이어	
360-1 내지 360-2: 덧셈 레이어	
370-1 내지 370-2: 완전 연결(Fully Connected) 레이어	
380: 소프트맥스(Softmax) 레이어	
390: 분류 출력 레이어	

도면

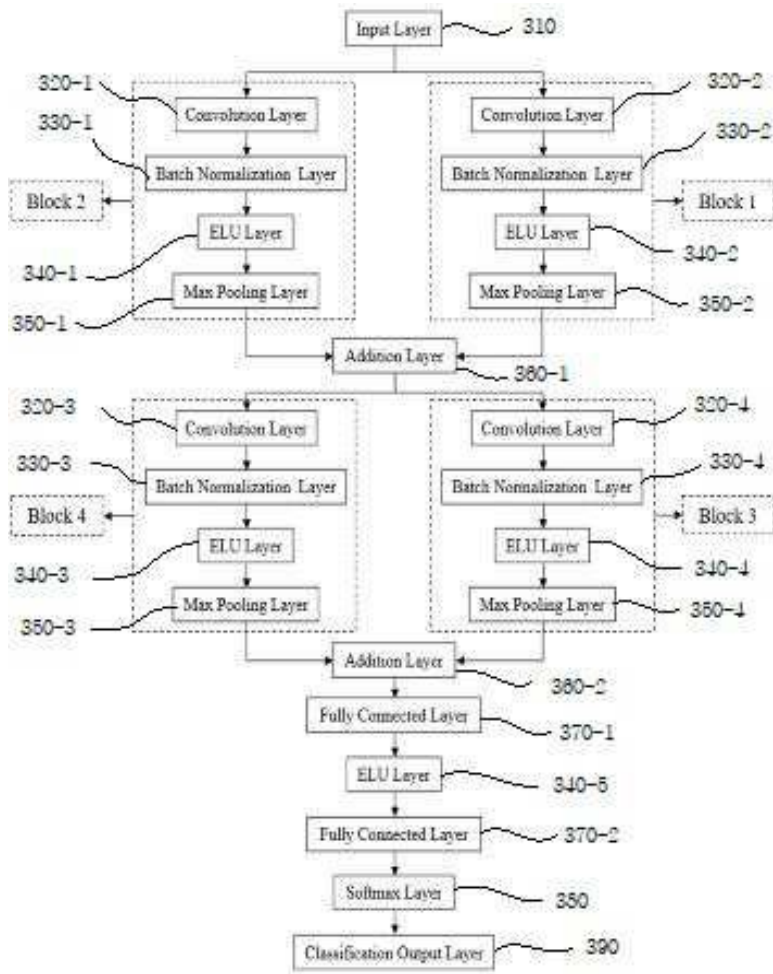
도면1



도면2

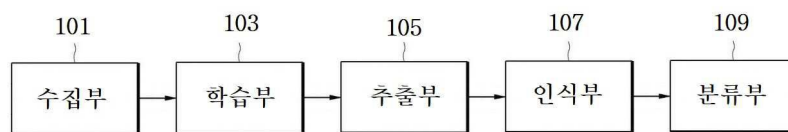


도면3

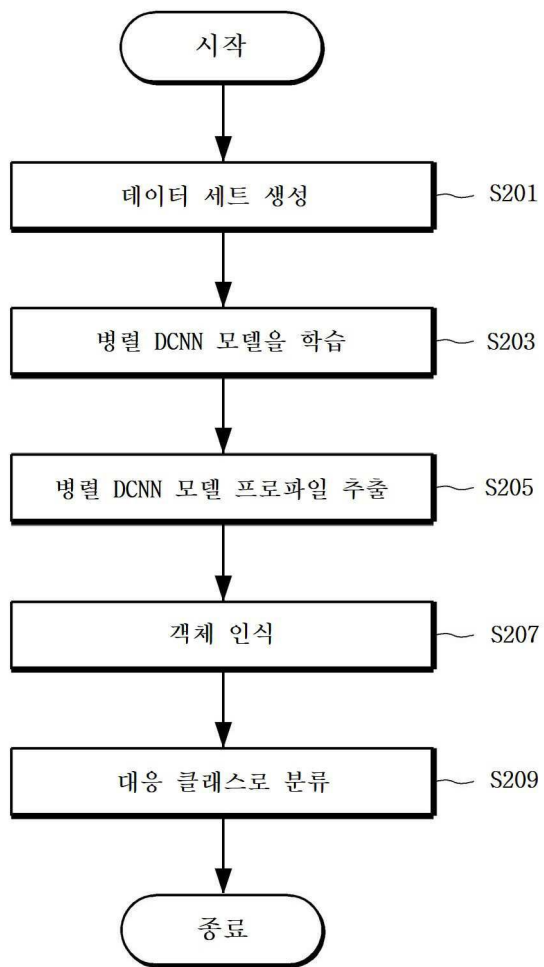


도면4

100



도면5



도면6

