(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0113349
(43) 공개일자 2024년07월22일

(51) 국제특허분류(Int. Cl.)
G06N 3/096 (2023.01) *G06N 3/045* (2023.01)
G06N 3/0464 (2023.01) *G06N 3/048* (2023.01)
G06N 3/0495 (2023.01)
(52) CPC특허분류
G06N 3/096 (2023.01)
G06N 3/045 (2023.01)
(21) 출원번호 10-2023-0039977
(22) 출원일자 2023년03월27일
심사청구일자 2023년03월27일
(30) 우선권주장
1020230004776 2023년01월12일 대한민국(KR)

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
허의남
경기도 용인시 기흥구 동탄기흥로 864, 106동 401호(공세동, 공세 호수마을 공세성원상떼레이크뷰)
조현기
경기도 수원시 영통구 신원로220번길 11-12, 201호(매탄동)
(뒷면에 계속)
(74) 대리인
두호특허법인

전체 청구항 수 : 총 16 항

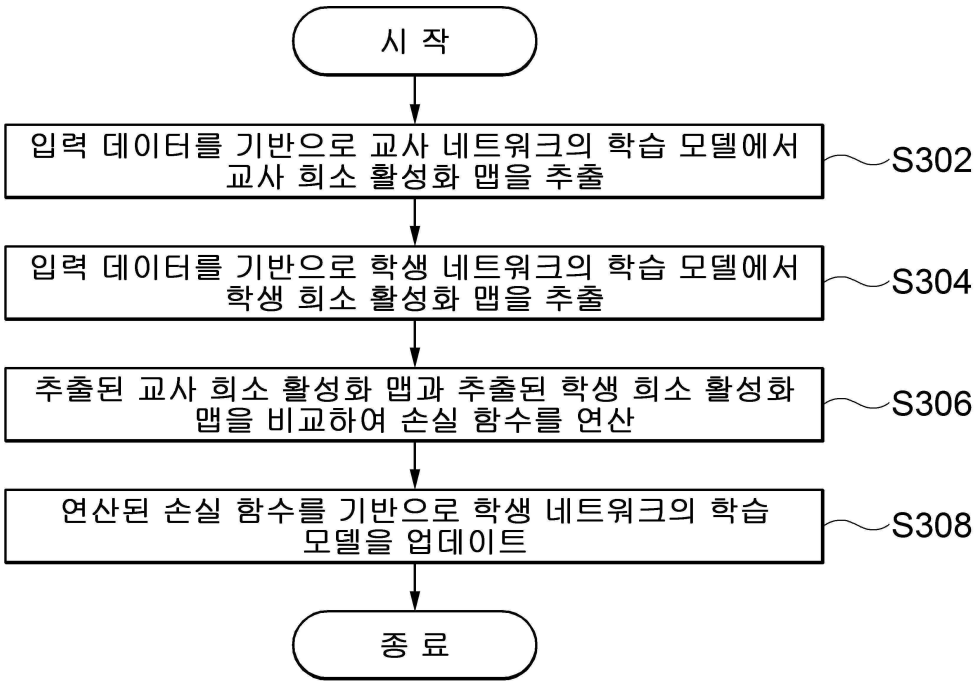
(54) 발명의 명칭 **지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법 및 장치**

(57) 요 약

지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법 및 장치가 개시된다. 본 발명의 일 실시예에 따른 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법 및 장치는 하나 이상의 프로세서들, 및 상기 하나 이상의 프로세서들에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치

(뒷면에 계속)

대 표 도 - 도3



에서 수행되는 지식 증류에서 SAMs(Sparse Activation Maps)를 이용한 활성화된 뉴런 반응 전달의 학습 방법으로서, 입력 데이터를 기반으로 교사 네트워크의 학습 모델에서 특징맵을 추출하고, 상기 추출된 특징맵을 필터링하여 교사 희소 활성화 맵(SAMs)을 추출하는 단계, 상기 입력 데이터를 기반으로 학생 네트워크의 학습 모델에서 특징맵을 추출하고, 상기 추출된 특징맵을 필터링하여 학생 희소 활성화 맵(SAMs)을 추출하는 단계, 상기 추출된 교사 희소 활성화 맵 및 상기 추출된 학생 희소 활성화 맵을 비교하여 손실 함수를 연산하는 단계 및 상기 연산된 손실 함수를 기반으로 상기 학생 네트워크의 학습 모델을 업데이트하는 단계를 포함한다.

(52) CPC특허분류
G06N 3/0464 (2023.01)
G06N 3/048 (2023.01)
G06N 3/0495 (2023.01)

(72) 발명자
샤르만 아카타르
호링가치,라파예푸르,다울랏푸르,쿠시티아

엠디 임티아즈 호씬
경기도 수원시 영통구 청명로59번길 44(영통동)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711152683
과제번호	2021-0-00818-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	ICTR&D혁신바우처지원사업(예타)
연구과제명	스마트 플랜트 이상상태 조기감지를 위한 머신러닝 기반 저대역 영상/통신 Edge Computing시스템
기 여 율	1/1
과제수행기관명	해피콤(주)
연구기간	2022.01.01 ~ 2022.12.31

명 세 서

청구범위

청구항 1

하나 이상의 프로세서들, 및

상기 하나 이상의 프로세서들에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치에서 수행되는 지식 증류에서 SAMs(Sparse Activation Maps)를 이용한 활성화된 뉴런 반응 전달의 학습 방법으로서,

입력 데이터를 기반으로 교사 네트워크의 학습 모델에서 특징맵을 추출하고, 상기 추출된 특징맵을 필터링하여 교사 희소 활성화 맵(SAMs)을 추출하는 단계;

상기 입력 데이터를 기반으로 학생 네트워크의 학습 모델에서 특징맵을 추출하고, 상기 추출된 특징맵을 필터링하여 학생 희소 활성화 맵(SAMs)을 추출하는 단계;

상기 추출된 교사 희소 활성화 맵 및 상기 추출된 학생 희소 활성화 맵을 비교하여 손실 함수를 연산하는 단계; 및

상기 연산된 손실 함수를 기반으로 상기 학생 네트워크의 학습 모델을 업데이트하는 단계를 포함하는, 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법.

청구항 2

청구항 1에 있어서,

상기 교사 희소 활성화 맵을 추출하는 단계는,

상기 입력 데이터를 기반으로 상기 교사 네트워크의 학습 모델에서 합성곱 층(convolution layer)을 이용하여 상기 특징맵을 추출하는 단계;

활성화 함수를 이용하여 상기 추출된 특징맵을 필터링하여 활성화 맵을 추출하는 단계; 및

필터 함수를 이용하여 상기 추출된 활성화 맵을 필터링하여 상기 교사 희소 활성화 맵을 추출하는 단계를 더 포함하는, 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법.

청구항 3

청구항 2에 있어서,

상기 교사 희소 활성화 맵을 추출하는 단계는,

상기 교사 네트워크의 학습 모델에서 추출된 활성화 맵을 기 설정된 범위로 정규화하는 단계; 및

활성화 함수를 통하여 상기 정규화된 활성화 맵을 재필터링하는 단계를 더 포함하는, 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법.

청구항 4

청구항 1에 있어서,

상기 학생 희소 활성화 맵을 추출하는 단계는,

상기 입력 데이터를 기반으로 상기 학생 네트워크의 학습 모델에서 합성곱 층(convolution layer)을 이용하여

상기 특징맵을 추출하는 단계;

활성화 함수를 이용하여 상기 추출된 특징맵을 필터링하여 활성화 맵을 추출하는 단계; 및

필터 함수를 이용하여 상기 추출된 활성화 맵을 필터링하여 상기 학생 희소 활성화 맵을 추출하는 단계를 더 포함하는, 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법.

청구항 5

청구항 4에 있어서,

상기 학생 희소 활성화 맵을 추출하는 단계는,

상기 학생 네트워크의 학습 모델에서 상기 추출된 활성화 맵을 기 설정된 범위로 정규화하는 단계; 및

활성화 함수를 통하여 상기 정규화된 활성화 맵을 재필터링하는 단계를 더 포함하는, 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법.

청구항 6

청구항 3 또는 청구항 5에 있어서,

상기 활성화 함수는, ReLU(Rectified Linear Unit) 함수이며,

상기 필터 함수는, ReLU 함수의 좌표를 이동하고, 통과대역(bandpass)를 조정하여 얻은 함수인, 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법.

청구항 7

청구항 1에 있어서,

상기 손실 함수는, 상기 추출된 교사 희소 활성화 맵과 상기 추출된 학생 희소 활성화 맵 간의 거리 차이에 의한 손실을 비교한 손실함수인, 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법.

청구항 8

청구항 7에 있어서,

상기 손실 함수는, 하기 수학식 2에 의해 산출되는, 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법.

[수학식 2]

$$\mathcal{L}_{res}(\mathcal{S}^S, \mathcal{S}^T) = \sum_{i=1}^{L} ||(\mathcal{S}_i^S - \mathcal{S}_i^T)||_2^2$$

(여기서, $L_{res}(\cdot)$는 손실 함수, S^S는 교사 네트워크의 학습된 모델의 희소 활성화 맵, S^T는 학생 네트워크의 학습된 모델의 희소 활성화 맵, L은 SAMs의 수)

청구항 9

하나 이상의 프로세서들;

메모리; 및

하나 이상의 프로그램들을 포함하고,

상기 하나 이상의 프로그램들은 상기 메모리에 저장되고, 상기 하나 이상의 프로세서들에 의해 실행되도록 구성되며,

상기 하나 이상의 프로그램들은,

입력 데이터를 기반으로 교사 네트워크의 학습 모델에서 특징맵을 추출하고, 상기 추출된 특징맵을 필터링하여 교사 희소 활성화 맵(SAMs)을 추출하기 위한 명령;

상기 입력 데이터를 기반으로 학생 네트워크의 학습 모델에서 특징맵을 추출하고, 상기 추출된 특징맵을 필터링하여 학생 희소 활성화 맵(SAMs)을 추출하기 위한 명령;

상기 추출된 교사 희소 활성화 맵 및 상기 추출된 학생 희소 활성화 맵을 비교하여 손실 함수를 연산하기 위한 명령; 및

상기 연산된 손실 함수를 기반으로 상기 학생 네트워크의 학습 모델을 업데이트하기 위한 명령을 포함하는, 컴퓨팅 장치.

청구항 10

청구항 9에 있어서,

상기 교사 희소 활성화 맵을 추출하기 위한 명령은,

상기 입력 데이터를 기반으로 상기 교사 네트워크의 학습 모델에서 합성곱 층(convolution layer)을 이용하여 상기 특징맵을 추출하기 위한 명령;

활성화 함수를 이용하여 상기 추출된 특징맵을 필터링하여 활성화 맵을 추출하기 위한 명령; 및

필터 함수를 이용하여 상기 추출된 활성화 맵을 필터링하여 상기 교사 희소 활성화 맵을 추출하기 위한 명령을 더 포함하는, 컴퓨팅 장치.

청구항 11

청구항 10에 있어서,

상기 교사 희소 활성화 맵을 추출하기 위한 명령은,

상기 교사 네트워크의 학습 모델에서 추출된 활성화 맵을 기 설정된 범위로 정규화하기 위한 명령; 및

활성화 함수를 통하여 상기 정규화된 활성화 맵을 재필터링하기 위한 명령을 더 포함하는, 컴퓨팅 장치.

청구항 12

청구항 9에 있어서,

상기 학생 희소 활성화 맵을 추출하기 위한 명령은,

상기 입력 데이터를 기반으로 상기 학생 네트워크의 학습 모델에서 합성곱 층(convolution layer)을 이용하여 상기 특징맵을 추출하기 위한 명령;

활성화 함수를 이용하여 상기 추출된 특징맵을 필터링하여 활성화 맵을 추출하기 위한 명령; 및

필터 함수를 이용하여 상기 추출된 활성화 맵을 필터링하여 상기 학생 희소 활성화 맵을 추출하기 위한 명령을 더 포함하는, 컴퓨팅 장치.

청구항 13

청구항 12에 있어서,

상기 학생 희소 활성화 맵을 추출하기 위한 명령은,

상기 학생 네트워크의 학습 모델에서 상기 추출된 활성화 맵을 기 설정된 범위로 정규화하기 위한 명령; 및

활성화 함수를 통하여 상기 정규화된 활성화 맵을 재필터링하기 위한 명령을 더 포함하는, 컴퓨팅 장치.

청구항 14

청구항 11 또는 청구항 13에 있어서,

상기 활성화 함수는, ReLU(Rectified Linear Unit) 함수이며,

상기 필터 함수는, ReLU 함수의 좌표를 이동하고, 통과대역(bandpass)를 조정하여 얻은 함수인, 컴퓨팅 장치.

청구항 15

청구항 9에 있어서,

상기 손실 함수는, 상기 추출된 교사 희소 활성화 맵과 상기 추출된 학생 희소 활성화 맵 간의 거리 차이에 의한 손실을 비교한 손실함수인, 컴퓨팅 장치.

청구항 16

청구항 15에 있어서,

상기 손실 함수는, 하기 수학식 2에 의해 산출되는, 컴퓨팅 장치.

[수학식 2]

$$\mathcal{L}_{res}(\mathcal{S}^S, \mathcal{S}^T) = \sum_{i=1}^{L} ||(\mathcal{S}_i^S - \mathcal{S}_i^T)||_2^2$$

(여기서, $L_{res}(\cdot)$는 손실 함수, S^S는 교사 네트워크의 학습된 모델의 희소 활성화 맵, S^T는 학생 네트워크의 학습된 모델의 희소 활성화 맵, L은 SAMs의 수)

발명의 설명

기 술 분 야

[0001] 본 발명의 실시예들은 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 기술과 관련된다.

배 경 기 술

[0003] 무겁고 복잡한 딥 러닝 아키텍처는 높은 계산 자원과 많은 에너지가 소비된다. 이는 낮은 계산자원을 요구하는 환경과 실시간 응용 프로그램(예: 6G 네트워크, 모바일, IoT, AI 지원 에지, 클라우드 및 포그 컴퓨팅 인프라)에 배포하는 것은 적합하지 않다. 저 성능의 장치에서 실시간 서비스를 지원하기 위해서는 가벼운 AI(Artificial Intelligence) 시스템이 필요하다.

[0004] 그러나, 경량 딥 러닝 아키텍처는 일반화 능력(훈련 데이터가 아닌 실제 데이터를 사용하는 능력)이 부족하기

때문에 무겁고 복잡한 딥 러닝 아키텍처보다 성능이 저하된다.

[0005] 이러한 문제를 해결하기 위해, 계산 자원이 적은 장치에서 최고 성능을 보여주는 경량 딥 러닝 아키텍처가 필요하다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 대한민국 공개특허공보 제10-2020-0053216호(2021.11.12)

발명의 내용

해결하려는 과제

[0008] 본 발명의 실시예들은 교사 네트워크의 학습 모델과 학생 네트워크의 학습 모델을 비교하여 연산한 손실 함수에 기반하여 학생 네트워크가 교사 네트워크와 동일한 뉴런 활성화 경계를 가지도록 활성화된 뉴런 반응을 전달하는 학습 방법을 제공하기 위한 것이다.

과제의 해결 수단

[0010] 본 발명의 예시적인 실시예에 따르면, 하나 이상의 프로세서들, 및 상기 하나 이상의 프로세서들에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치에서 수행되는 지식 증류에서 SAMs(Sparse Activation Maps)를 이용한 활성화된 뉴런 반응 전달의 학습 방법으로서, 입력 데이터를 기반으로 교사 네트워크의 학습 모델에서 특징맵을 추출하고, 상기 추출된 특징맵을 필터링하여 교사 희소 활성화 맵(SAMs)을 추출하는 단계, 상기 입력 데이터를 기반으로 학생 네트워크의 학습 모델에서 특징맵을 추출하고, 상기 추출된 특징맵을 필터링하여 학생 희소 활성화 맵(SAMs)을 추출하는 단계, 상기 추출된 교사 희소 활성화 맵 및 상기 추출된 학생 희소 활성화 맵을 비교하여 손실 함수를 연산하는 단계 및 상기 연산된 손실 함수를 기반으로 상기 학생 네트워크의 학습 모델을 업데이트하는 단계를 포함하는 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법이 제공된다.

[0011] 상기 교사 희소 활성화 맵을 추출하는 단계는 상기 입력 데이터를 기반으로 상기 교사 네트워크의 학습 모델에서 합성곱 층(convolution layer)을 이용하여 상기 특징맵을 추출하는 단계, 활성화 함수를 이용하여 상기 추출된 특징맵을 필터링하여 활성화 맵을 추출하는 단계 및 필터 함수를 이용하여 상기 추출된 활성화 맵을 필터링하여 상기 교사 희소 활성화 맵을 추출하는 단계를 더 포함할 수 있다.

[0012] 상기 교사 희소 활성화 맵을 추출하는 단계는 상기 교사 네트워크의 학습 모델에서 추출된 활성화 맵을 기 설정된 범위로 정규화하는 단계 및 활성화 함수를 통하여 상기 정규화된 활성화 맵을 재필터링하는 단계를 더 포함할 수 있다.

[0013] 상기 학생 희소 활성화 맵을 추출하는 단계는 상기 입력 데이터를 기반으로 상기 학생 네트워크의 학습 모델에서 합성곱 층(convolution layer)을 이용하여 상기 특징맵을 추출하는 단계, 활성화 함수를 이용하여 상기 추출된 특징맵을 필터링하여 활성화 맵을 추출하는 단계 및 필터 함수를 이용하여 상기 추출된 활성화 맵을 필터링하여 상기 학생 희소 활성화 맵을 추출하는 단계를 더 포함할 수 있다.

[0014] 상기 학생 희소 활성화 맵을 추출하는 단계는 상기 학생 네트워크의 학습 모델에서 상기 추출된 활성화 맵을 기 설정된 범위로 정규화하는 단계 및 활성화 함수를 통하여 상기 정규화된 활성화 맵을 재필터링하는 단계를 더 포함할 수 있다.

[0015] 상기 활성화 함수는, ReLU(Rectified Linear Unit) 함수이며, 상기 필터 함수는 ReLU 함수의 좌표를 이동하고, 통과대역(bandpass)를 조정하여 얻은 함수일 수 있다.

[0016] 상기 손실 함수는 상기 추출된 교사 희소 활성화 맵과 상기 추출된 학생 희소 활성화 맵 간의 거리 차이에 의한

손실을 비교한 손실함수일 수 있다.

[0017] 상기 손실 함수는 하기 수학식 2에 의해 산출될 수 있다.

[0018] [수학식 2]

[0019]
$$\mathcal{L}_{res}(\mathcal{S}^S, \mathcal{S}^T) = \sum_{i=1}^{L} ||(\mathcal{S}_i^S - \mathcal{S}_i^T)||_2^2$$

[0020] (여기서, $L_{res}(\cdot)$는 손실 함수, S^S는 교사 네트워크의 학습된 모델의 희소 활성화 맵, S^T는 학생 네트워크의 학습된 모델의 희소 활성화 맵, L은 SAMs의 수)

발명의 효과

[0022] 본 발명의 실시예들에 따르면, 활성화된 뉴런 반응을 전달하는 학습 방법을 통하여 학생 네트워크의 비선형성을 개선하고, 일반화 능력을 향상시키면서 교사 네트워크에 근접한 성능의 학습 모델을 생성할 수 있다.

도면의 간단한 설명

[0024] 도 1는 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 장치를 포함하는 컴퓨팅 환경을 예시하여 설명하기 위한 블록도

도 2는 일반적인 지식 증류 기반 교사-학생 네트워크의 학습 모델을 예시한 블록도

도 3은 본 발명의 일 실시예에 따른 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법을 설명하기 위한 흐름도

도 4는 본 발명의 일 실시예에 따른 활성화 함수 및 필터 함수를 설명하기 위한 도면

도 5는 본 발명의 일 실시예에 따른 희소 활성화 맵(SAMs)을 추출하는 과정을 설명하기 위한 도면

도 6은 본 발명의 일 실시예에 따른 교사 네트워크에서 추출된 학습 모델의 희소 활성화 맵(SAMs)과 학생 네트워크에서 추출된 학습 모델의 희소 활성화 맵을 이용하여 손실 함수를 연산하는 과정을 예시한 도면

발명을 실시하기 위한 구체적인 내용

[0025] 이하, 도면을 참조하여 본 발명의 구체적인 실시형태를 설명하기로 한다. 이하의 상세한 설명은 본 명세서에서 기술된 방법, 장치 및/또는 시스템에 대한 포괄적인 이해를 돕기 위해 제공된다. 그러나 이는 예시에 불과하며 본 발명은 이에 제한되지 않는다.

[0026] 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 실시예들을 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.

[0027] 이하의 설명에 있어서, 신호 또는 정보의 "전송", "통신", "송신", "수신" 기타 이와 유사한 의미의 용어는 일 구성요소에서 다른 구성요소로 신호 또는 정보가 직접 전달되는 것뿐만이 아니라 다른 구성요소를 거쳐 전달되는 것도 포함한다. 특히 신호 또는 정보를 일 구성요소로 "전송" 또는 "송신"한다는 것은 그 신호 또는 정보의 최종 목적지를 지시하는 것이고 직접적인 목적지를 의미하는 것이 아니다. 이는 신호 또는 정보의 "수신"에 있어서도 동일하다. 또한 본 명세서에 있어서, 2 이상의 데이터 또는 정보가 "관련"된다는 것은 하나의 데이터(또

는 정보)를 획득하면, 그에 기초하여 다른 데이터(또는 정보)의 적어도 일부를 획득할 수 있음을 의미한다.

[0028] 한편, 본 발명의 실시예는 본 명세서에서 기술한 방법들을 컴퓨터상에서 수행하기 위한 프로그램, 및 상기 프로그램을 포함하는 컴퓨터 판독 가능 기록매체를 포함할 수 있다. 상기 컴퓨터 판독 가능 기록매체는 프로그램 명령, 로컬 데이터 파일, 로컬 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체는 본 발명을 위하여 특별히 설계되고 구성된 것들이거나, 또는 컴퓨터 소프트웨어 분야에서 통상적으로 사용 가능한 것일 수 있다. 컴퓨터 판독 가능 기록매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM, DVD와 같은 광 기록 매체, 및 롬, 램, 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 상기 프로그램의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함할 수 있다.

[0029] 도 1는 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 장치를 포함하는 컴퓨팅 환경(10)을 예시하여 설명하기 위한 블록도이다. 도시된 실시예에서, 각 컴포넌트들은 이하에 기술된 것 이외에 상이한 기능 및 능력을 가질 수 있고, 이하에 기술되는 것 이외에도 추가적인 컴포넌트를 포함할 수 있다.

[0030] 도시된 컴퓨팅 환경(10)은 컴퓨팅 장치(12)를 포함한다. 일 실시예에서, 컴퓨팅 장치(12)는 본 발명의 실시예에 따른 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습을 수행하기 위한 장치일 수 있다.

[0031] 컴퓨팅 장치(12)는 적어도 하나의 프로세서(14), 컴퓨터 판독 가능 저장 매체(16) 및 통신 버스(18)를 포함한다. 프로세서(14)는 컴퓨팅 장치(12)로 하여금 앞서 언급된 예시적인 실시예에 따라 동작하도록 할 수 있다. 예컨대, 프로세서(14)는 컴퓨터 판독 가능 저장 매체(16)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 상기 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 상기 컴퓨터 실행 가능 명령어는 프로세서(14)에 의해 실행되는 경우 컴퓨팅 장치(12)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.

[0032] 컴퓨터 판독 가능 저장 매체(16)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능 저장 매체(16)에 저장된 프로그램(20)은 프로세서(14)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능 저장 매체(16)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 컴퓨팅 장치(12)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.

[0033] 통신 버스(18)는 프로세서(14), 컴퓨터 판독 가능 저장 매체(16)를 포함하여 컴퓨팅 장치(12)의 다른 다양한 컴포넌트들을 상호 연결한다.

[0034] 컴퓨팅 장치(12)는 또한 하나 이상의 입출력 장치(24)를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(22) 및 하나 이상의 네트워크 통신 인터페이스(26)를 포함할 수 있다. 입출력 인터페이스(22) 및 네트워크 통신 인터페이스(26)는 통신 버스(18)에 연결된다. 입출력 장치(24)는 입출력 인터페이스(22)를 통해 컴퓨팅 장치(12)의 다른 컴포넌트들에 연결될 수 있다. 예시적인 입출력 장치(24)는 포인팅 장치(마우스 또는 트랙패드 등), 키보드, 터치 입력 장치(터치패드 또는 터치스크린 등), 음성 또는 소리 입력 장치, 다양한 종류의 센서 장치 및/또는 촬영 장치와 같은 입력 장치, 및/또는 디스플레이 장치, 프린터, 스피커 및/또는 네트워크 카드와 같은 출력 장치를 포함할 수 있다. 예시적인 입출력 장치(24)는 컴퓨팅 장치(12)를 구성하는 일 컴포넌트로서 컴퓨팅 장치(12)의 내부에 포함될 수도 있고, 컴퓨팅 장치(12)와는 구별되는 별개의 장치로 컴퓨팅 장치(12)와 연결될 수도 있다.

[0035] 도 2는 일반적인 지식 증류 기반 교사-학생 네트워크의 학습 모델을 예시한 블록도이다.

[0036] 도 2와 같이, 지식 증류(Knowledge Distillation) 기반 교사-학생 네트워크의 학습 모델은 교사 네트워크(Teacher Network)와 학생 네트워크(Student Network)를 포함한다. 각각의 네트워크는 인공 뉴런들이 네트워크로 연결된 뉴럴 네트워크로 구성된다. 이러한 지식 증류 기반 교사-학생 네트워크에서는 교사 네트워크이 각 블록들이 관련되고, 학생 네트워크의 각 블록들이 관련되어, 관련된 블록들 전체에 대한 지식 전이(transfer)가 이루어지며, 손실 함수도 전체 블록들에 대하여 결정되게 된다.

[0037] 통상적으로, 교사 네트워크는 이미 유사한 선행 학습을 수행한 네트워크 또는 더욱 방대한 네트워크를 의미하며, 학생 네트워크에 비해 큰 사이즈를 가진다. 여기서, 큰 사이즈라 함은 더욱 많은 레이어(Layer), 커널(Kernel), 노드(Node)를 가진 네트워크, 또는 그러한 조합들로 구성된 네트워크를 의미한다. 또한, 학생 네트

워크는 교사 네트워크에서 추출되는 학습 모델을 추종하며, 교사 네트워크와 유사하지만 그보다는 작은 사이즈의 네트워크 구성요소들을 갖는다.

[0038] 본 발명의 중요한 기술적 특징 중 하나는 학생 네트워크가 교사 네트워크의 활성화 경계를 학습한다는데 있다. 본 발명은 교사 네트워크의 학습 모델과 학생 네트워크의 학습 모델을 비교하여 연산한 손실 함수에 기반하여 학생 네트워크가 교사 네트워크와 동일한 뉴런 활성화 경계를 가지도록 활성화된 뉴런 반응을 전달하는 학습 방법을 제공할 수 있다. 즉, 본 발명은 활성화된 뉴런 반응을 전달하는 학습 방법을 통하여 학생 네트워크의 비선형성을 개선하고, 일반화 능력을 향상시키면서 교사 네트워크에 근접한 성능의 학습 모델을 생성할 수 있다.

[0039] 도 3은 본 발명의 일 실시예에 따른 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법을 설명하기 위한 흐름도이다. 전술한 바와 같이, 본 발명의 일 실시예에 따른 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법은 하나 이상의 프로세서들, 및 상기 하나 이상의 프로세서들에 의해 실행되는 하나 이상의 프로그램들을 저장하는 메모리를 구비한 컴퓨팅 장치(12)에서 수행될 수 있다. 이를 위하여, 상기 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법은 하나 이상의 컴퓨터 실행 가능 명령어를 포함하는 프로그램 내지 소프트웨어의 형태로 구현되어 상기 메모리상에 저장될 수 있다.

[0040] 또한, 도시된 흐름도에서는 상기 방법을 복수 개의 단계로 나누어 기재하였으나, 적어도 일부의 단계들은 순서를 바꾸어 수행되거나, 다른 단계와 결합되어 함께 수행되거나, 생략되거나, 세부 단계들로 나뉘어 수행되거나, 또는 도시되지 않은 하나 이상의 단계가 부가되어 수행될 수 있다.

[0041] 단계 302에서, 컴퓨팅 장치(12)는 입력 데이터를 기반으로 교사 네트워크의 학습 모델에서 특징맵을 추출하고, 추출한 특징맵을 필터링하여 교사 희소 활성화 맵(교사 네트워크의 학습 모델에서 추출된 희소 활성화 맵(Sparse Activation Maps; SAMs)을 추출한다.

[0042] 예시적인 실시예에서, 컴퓨팅 장치(12)는 입력 데이터를 기반으로 교사 네트워크의 학습 모델에서 합성곱 층(convolution layer)을 이용하여 특징맵을 추출할 수 있다. 또한, 컴퓨팅 장치(12)는 추출한 특징맵으로부터 활성화 함수를 이용하여 활성화 맵을 추출하고, 추출한 활성화 맵으로부터 필터 함수를 이용하여 희소 활성화 맵을 추출할 수 있다.

[0043] 이 때, 활성화 함수는 ReLU (Rectified Linear Unit) 함수일 수 있다. ReLU 함수는 입력값이 0보다 작으면 0을 출력(필터링)하고, 0보다 크면 입력값을 그대로 출력하는 함수로써, 본 발명에서는 특징맵에서 0보다 작은 값은 0으로(비활성화), 0보다 큰 값은 그대도 남겨둔(활성화) 행렬인 활성화 맵을 얻을 수 있다.

[0044] 또한, 필터 함수는 ReLU 함수를 이동하고 크기를 조정하여 얻은 함수로써, 본 발명에서는 활성화 맵에서 기 설정된 값보다 작은 값(낮은 활성화 값)은 필터링하고, 기 설정된 값보다 큰 값(높은 활성화 값)만 남겨둔 행렬인 희소 활성화 맵을 얻을 수 있다.

[0045] 도 4는 본 발명의 일 실시예에 따른 활성화 함수 및 필터 함수를 설명하기 위한 도면이고, 도 5는 본 발명의 일 실시예에 따른 희소 활성화 맵(SAMs)을 추출하는 과정을 설명하기 위한 도면이다.

[0046] 도 4를 참조하면, 활성화 함수(a)는 ReLU 함수이며, 필터 함수(b)는 RuLU 함수를 이동(좌표)하고 조정(범위 또는 통과 대역(bandpass))하여 얻은 함수이다.

[0047] 도 5와 같이, 본 발명에서는 활성화 함수를 통하여 특징맵을 필터링하여 활성화 맵을 추출하고, 추출한 활성화 맵을 [-ξ, ξ] 범위로 정규화하고 활성화 함수를 통하여 재필터링하여 희소 활성화 맵(SAMs)를 추출할 수 있다. 여기서, ξ는 분리 초평면의 범위를 조정하기 위한 온도(temperature)일 수 있으며, 예를 들어, ξ = 20 일 때, 가장 좋은 성능을 나타낼 수 있다.

[0048] 즉, 특징맵에서 활성화 함수를 통하여 비활성화된 뉴런(입력 값이 0이하)을 필터링하여 제외하고, 활성화된 뉴런(입력 값이 0 이상)을 가지는 활성화 맵을 생성할 수 있다. 이 때, 활성화 맵은 낮은 값(활성화 값이 기 설정된 값보다 낮은 값)으로 활성화된 뉴런 및 높은 값(활성화 값이 기 설정된 값 보다 높은 값)으로 활성화된 뉴런을 포함하게 된다. 이에 활성화 맵에서 필터 함수를 통하여 낮은 값으로 활성화된 뉴런을 필터링하여 제외하고, 높은 값으로 활성화된 뉴런만 가지는 희소 활성화 맵을 생성할 수 있다. 이를 통해, 높은 활성화 뉴런 반응을 전달하여 학생 네트워크의 비선형성을 개선할 수 있다.

[0049] 이 때, 희소 활성화 맵은 하기 수학식 1과 같이 정의할 수 있다.

[0050] [수학식 1]

[0051] $$S_{i,j,k} = \varphi(A_{i,j,k}) = \emptyset\left(\frac{A_{i,j,k}}{\xi \times \|A\|}\right)\Big|_{i=0,j=0,k=0}^{i=C,j=H,k=W}$$

[0052] (여기서, S는 희소 활성화 맵, A는 활성화 맵, ξ는 범위 조정 값, ϕ(·)는 활성화 함수, $\varphi(\cdot)$는 필터 함수, 아래 첨자 i, j, k는 3차원 배열(특징맵 및 활성화 맵은 3차원 행렬로 이루어짐), W는 너비, H는 높이, C는 채널)

[0053] 단계 304에서, 컴퓨팅 장치(12)는 입력 데이터를 기반으로 학생 네트워크의 학습 모델에서 특징맵을 추출하고, 추출한 특징맵을 필터링하여 학생 희소 활성화 맵(학생 네트워크의 학습 모델에서 추출된 희소 활성화 맵(SAMs)을 추출한다. 여기서, 학생 네트워크의 학습 모델에서 학생 희소 활성화 맵을 추출하는 과정은 상기 교사 네트워크의 학습 모델에서 교사 희소 활성화 맵을 추출하는 과정과 동일하다.

[0054] 단계 306에서, 컴퓨팅 장치(12)는 추출된 교사 희소 활성화 맵과 추출된 학생 희소 활성화 맵을 비교하여 손실 함수($L_{res}(S^S, S^T)$)를 연산한다. 여기서, S^S는 교사 네트워크의 학습된 모델의 희소 활성화 맵, S^T는 학생 네트워크의 학습된 모델의 희소 활성화 맵, $L_{res}(S^S, S^T)$는 두 희소 활성화 맵을 비교한 손실 함수를 의미할 수 있다.

[0055] 예시적인 실시예에서, 손실 함수는 교사 네트워크의 학습된 모델의 희소 활성화 맵과 학생 네트워크의 학습된 모델의 희소 활성화 맵 간의 거리 차이에 의해 연산될 수 있다.

[0056] 손실 함수는 하기 수학식 2에 의해 산출될 수 있다.

[0057] [수학식 2]

[0058] $$\mathcal{L}_{res}(\mathcal{S}^S, \mathcal{S}^T) = \sum_{i=1}^{L} ||(\mathcal{S}_i^S - \mathcal{S}_i^T)||_2^2$$

[0059] (여기서, $L_{res}(\cdot)$는 손실 함수, S^S는 교사 네트워크의 학습된 모델의 희소 활성화 맵, S^T는 학생 네트워크의 학습된 모델의 희소 활성화 맵, L은 SAMs의 수)

[0060] 여기서, 손실 함수는 L2-Norm 방식으로 거리가 계산되는 것으로, i가 1 부터 L까지 일 때 교사 희소 활성화 맵과 학생 희소 활성화 맵 사이의 차이(거리)의 제곱을 모두 더하여 산출될 수 있다

[0061] 도 6은 본 발명의 일 실시예에 따른 교사 네트워크에서 추출된 학습 모델의 희소 활성화 맵과 학생 네트워크에서 추출된 학습 모델의 희소 활성화 맵을 이용하여 손실 함수를 연산하는 과정을 예시한 도면이다.

[0062] 도 6에 도시된 바와 같이, 손실 함수는 각 블록별로(이 때, 블록의 수는 L 개) 희소 활성화 맵을 비교한 손실 함수를 이용하여 전체 네트워크의 손실 함수를 산출할 수 있다. 즉, 컴퓨팅 장치(12)는 교사 네트워크 및 학생 네트워크의 학습 모델의 블록별로 출력된 특징맵(F_L^T, F_L^S)을 활성화 함수(ϕ(·))를 통하여 활성화 맵(A_L^T, A_L^S)을 추출하고, 추출한 활성화 맵을 필터 함수($\varphi(\cdot)$)를 통하여 희소 활성화 맵(S_L^T, S_L^S)추출할 수 있다. 또한, 컴퓨팅 장치(12)는 교사 네트워크의 학습 모델에서 추출한 희소 활성화 맵(S_L^T)과 학생 네트워크의 학습 모델에서 추출한 희소 활성화 맵(S_L^S)을 비교하여 손실 함수($L_{res}(S^S, S^T)$)를 산출할 수 있다.

[0063] 단계 308에서, 컴퓨팅 장치(12)는 연산된 손실 함수를 기반으로 학생 네트워크의 학습 모델을 업데이트한다.

[0064] 즉, 컴퓨팅 장치(12)는 연산된 손실 함수를 최소화하는 가중치를 통하여 학생 네트워크의 학습 모델을 업데이트 할 수 있다. 이렇게 하여 학습 네트워크의 학습 모델이 교사 네트워크의 학습 모델과 동일한 뉴런 활성화 경계를 가질 때까지 훈련이 반복될 수 있다.

[0065] 따라서, 본 발명의 일 실시예에 따른 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법은 활성화된 뉴런 반응을 전달하는 학습 방법을 통하여 학생 네트워크의 비선형성을 개선하고, 일반화 능력을 향상시키면서 교사 네트워크에 근접한 성능의 학습 모델을 생성할 수 있다.

[0066] 한편, 본 발명의 일 실시예에 따른 지식 증류에서 SAMs를 이용한 활성화된 뉴런 반응 전달의 학습 방법은 입력 데이터뿐만 아니라 증강 데이터를 함께 이용하여 학생 네트워크의 학습 모델을 학습할 수 있다. 여기서, 증강 데이터는 입력 데이터를 변환(자르기, 회전, 드롭, 색상변환)하여 생성할 수 있다. 또한, 증강 데이터를 학습하기 위한 교사 보조 네트워크 및 학생 보조 네트워크를 더 포함할 수 있다.

[0067] 컴퓨팅 장치(12)는 입력 데이터를 기반으로 교사 네트워크의 학습 모델을 학습하고, 증강 데이터를 기반으로 교사 보조 네트워크의 학습 모델을 학습할 수 있다.

[0068] 또한, 컴퓨팅 장치(12)는 입력 데이터를 기반으로 학생 네트워크의 학습 모델을 학습하고, 증강 데이터를 기반으로 학생 보조 네트워크의 학습 모델을 학습할 수 있다.

[0069] 또한, 컴퓨팅 장치(12)는 교사 네트워크와 학생 네트워크 간의 지식 증류 손실, 예측 값과 실제 값과의 교차 엔트로피 손실, 교사 보조 네트워크와 학생 보조 네트워크 간의 지식 증류 손실, 및 교사 네트워크의 희소 활성화 맵과 학생 네트워크의 희소 활성화 맵 간의 손실을 이용하여 학생 네트워크를 업데이트할 수 있다.

[0070] 즉, 전체 손실 함수는 하기 수학식 3에 의해 산출될 수 있다.

[0071] [수학식 3]

[0072] $$\mathcal{L}_{NeuRes} = \beta_1 \mathcal{L}_{kd} + \beta_2 \mathcal{L}_{ce} + \mathcal{L}_{ax-st} + \mathcal{L}_{res}$$

[0073] 여기서, L_{kd}는 교사 네트워크와 학생 네트워크 간의 지식 증류 손실, L_{ce}는 예측 값과 실제 값과의 교차 엔트로피 손실, L_{ax-st}는 교사 보조 네트워크와 학생 보조 네트워크 간의 지식 증류 손실, L_{res}는 교사 네트워크의 희소 활성화 맵과 학생 네트워크의 희소 활성화 맵 간의 손실, β1, β2는 밸런싱 하이퍼파라미터이다.

[0074] 이상에서 본 발명의 대표적인 실시예들을 상세하게 설명하였으나, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 상술한 실시예에 대하여 본 발명의 범주에서 벗어나지 않는 한도 내에서 다양한 변형이 가능함을 이해할 것이다. 그러므로 본 발명의 권리범위는 설명된 실시예에 국한되어 정해져서는 안 되며, 후술하는 특허청구범위뿐만 아니라 이 특허청구범위와 균등한 것들에 의해 정해져야 한다.

부호의 설명

[0075] 10: 컴퓨팅 환경

12: 컴퓨팅 장치

14: 프로세서

16: 컴퓨터 판독 가능 저장 매체

18: 통신 버스

20: 프로그램

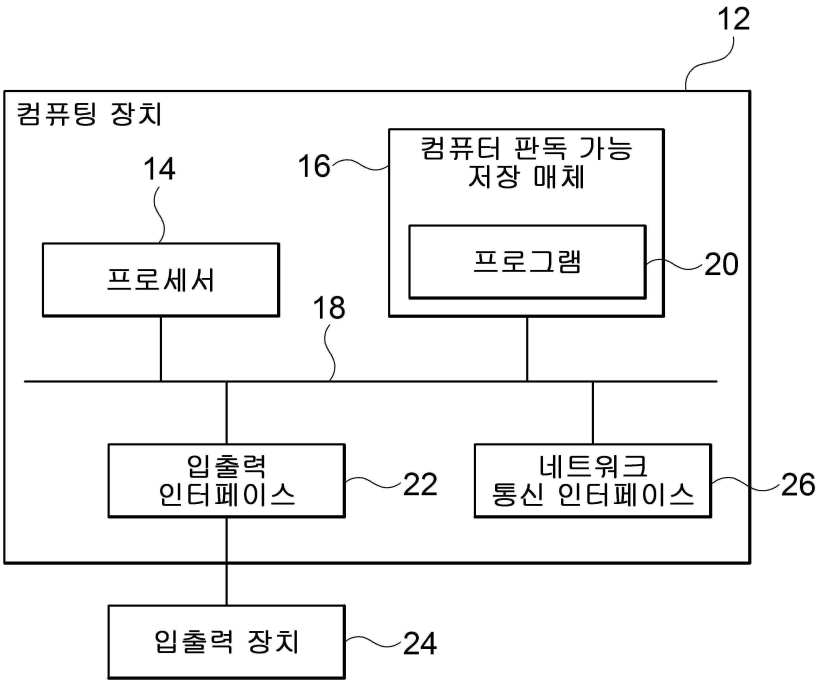
22: 입출력 인터페이스

24: 입출력 장치

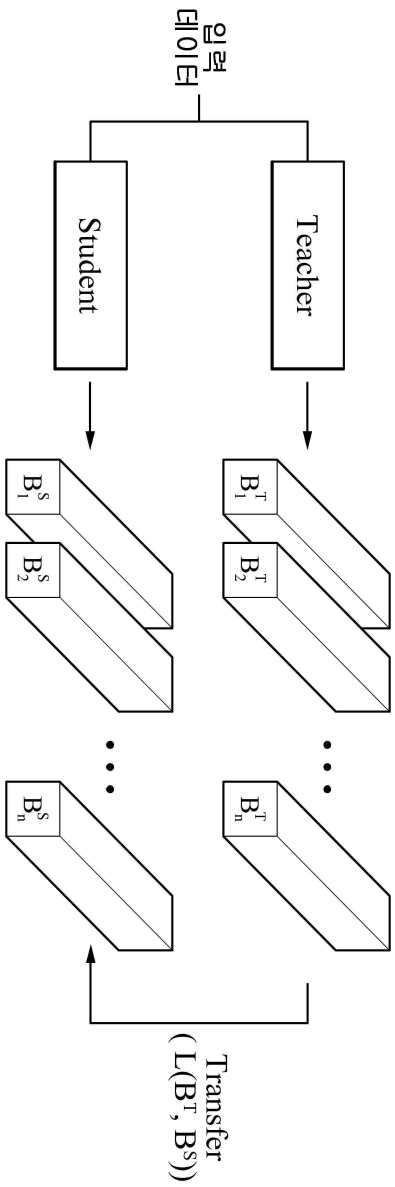
26: 네트워크 통신 인터페이스

도면

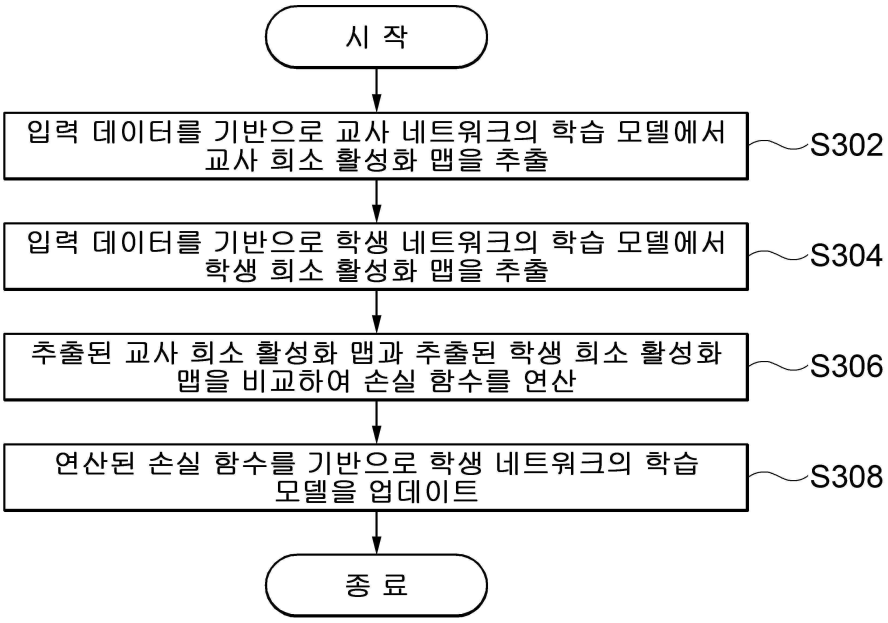
도면1


10
12
컴퓨팅 장치
14
프로세서
16
컴퓨터 판독 가능
저장 매체
프로그램
20
18
입출력
인터페이스
22
네트워크
통신 인터페이스
26
입출력 장치
24

도면2

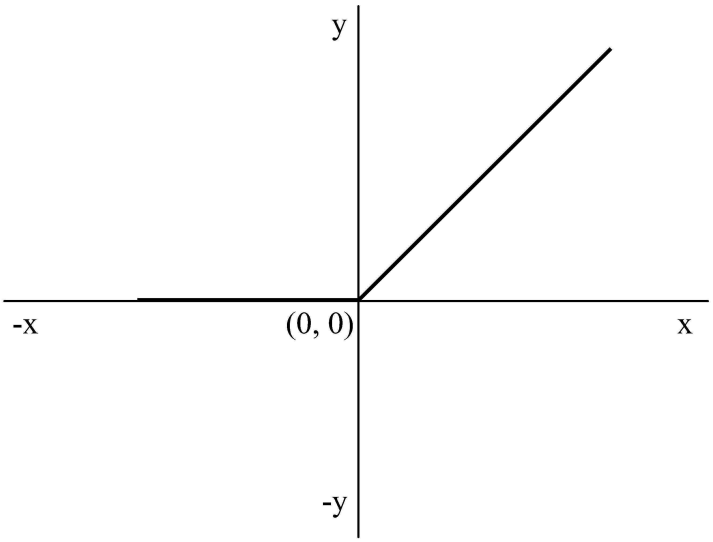

입력 데이터
Student
Teacher
B_1^S
B_2^S
B_n^S
B_1^T
B_2^T
B_n^T
Transfer ($L(B^T, B^S)$)

도면3

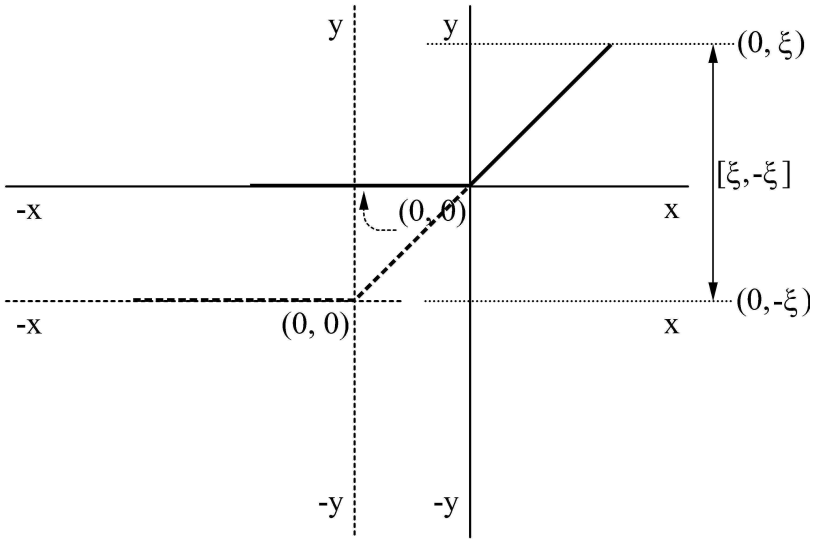

시 작
입력 데이터를 기반으로 교사 네트워크의 학습 모델에서 교사 희소 활성화 맵을 추출
S302
입력 데이터를 기반으로 학생 네트워크의 학습 모델에서 학생 희소 활성화 맵을 추출
S304
추출된 교사 희소 활성화 맵과 추출된 학생 희소 활성화 맵을 비교하여 손실 함수를 연산
S306
연산된 손실 함수를 기반으로 학생 네트워크의 학습 모델을 업데이트
S308
종 료

도면4

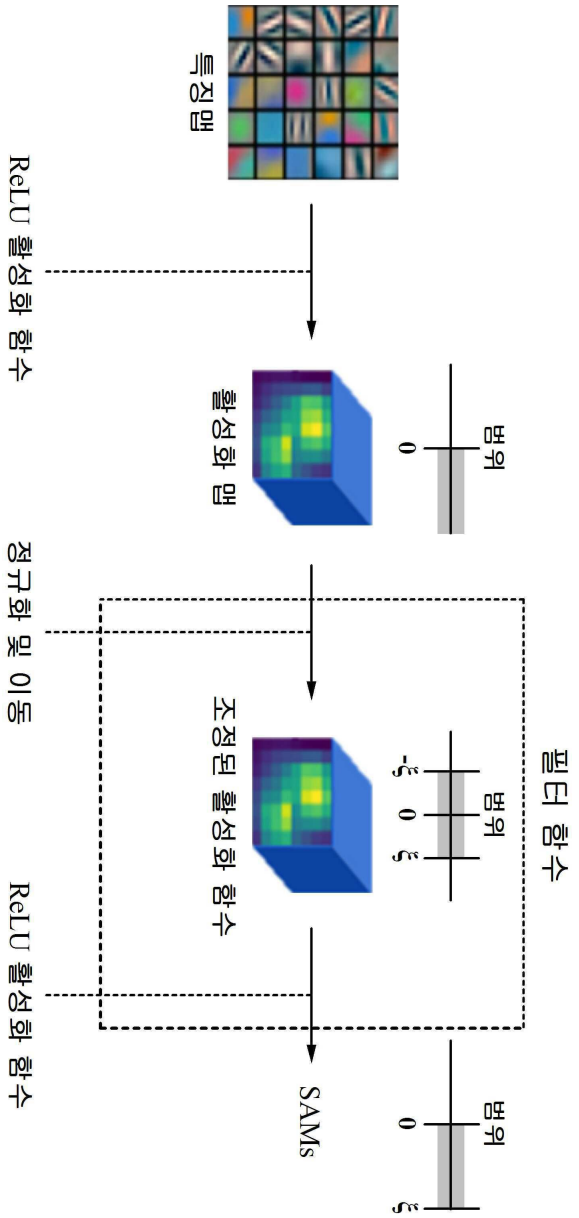
(a)



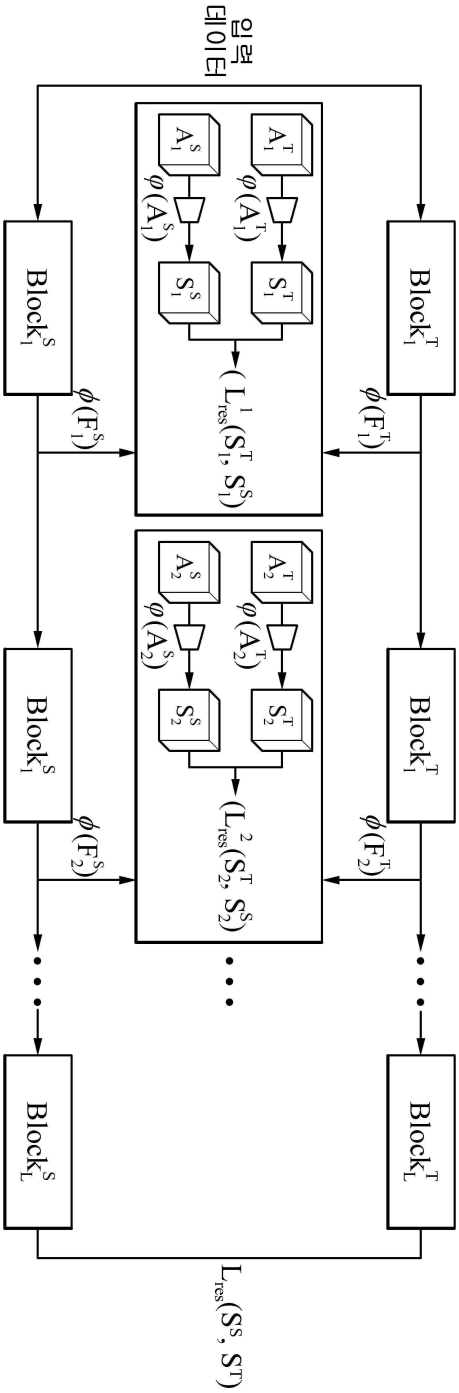
(b)



도면5


특징 맵
ReLU 활성화 함수
활성화 맵
0
범위
정규화 및 이동
조정된 활성화 함수
-ε
0
ε
범위
필터 함수
ReLU 활성화 함수
SAMs
0
ε
범위

도면6


입력 데이터
Block$_1^S$
Block$_1^T$
A_1^S
A_1^T
$\varphi(A_1^S)$
$\varphi(A_1^T)$
S_1^S
S_1^T
$(L_{res}^1(S_1^T, S_1^S)$
$\phi(F_1^S)$
$\phi(F_1^T)$
Block$_1^S$
Block$_1^T$
A_2^S
A_2^T
$\varphi(A_2^S)$
$\varphi(A_2^T)$
S_2^S
S_2^T
$(L_{res}^2(S_2^T, S_2^S)$
$\phi(F_2^S)$
$\phi(F_2^T)$
Block$_L^S$
Block$_L^T$
$L_{res}(S^S, S^T)$