(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0008135
(43) 공개일자 2022년01월20일

(51) 국제특허분류(Int. Cl.)
G06T 3/40 (2006.01) *G06N 3/04* (2006.01)
G06N 3/08 (2006.01)
(52) CPC특허분류
G06T 3/4053 (2013.01)
G06N 3/04 (2013.01)
(21) 출원번호 10-2020-0086360
(22) 출원일자 2020년07월13일
심사청구일자 2021년07월02일

(71) 출원인
서강대학교산학협력단
서울특별시 마포구 백범로 35 (신수동, 서강대학교)
(72) 발명자
강석주
서울특별시 마포구 백범로 35(신수동)
서유림
경기도 성남시 수정구 위례광장로 70, 103동 904호(창곡동, 위례 아트리버 푸르지오 1단지)
(74) 대리인
유미특허법인

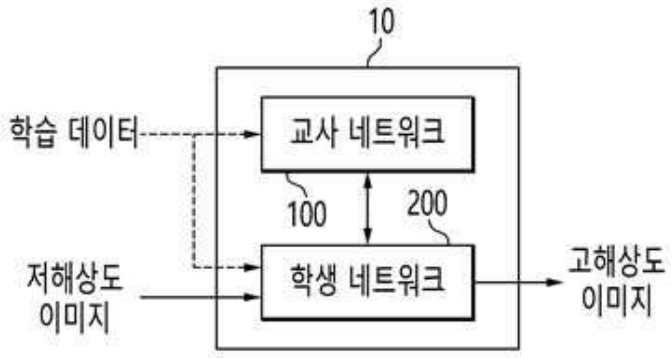
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 **영상 초해상도 처리 방법 및 장치**

(57) 요 약

적어도 하나의 프로세서에 의해 동작하는 컴퓨팅 장치가 동작하는 방법으로서, 저해상도 이미지와 상기 저해상도 이미지에 대응되는 고해상도 이미지가 포함된 학습 데이터를 이용하여 교사 모델을 학습시키고, 상기 학습 데이터로 학생 모델을 초기 학습시키는 단계, 상기 교사 모델의 학습 과정에서 생성된 특징값들 중 중요도가 높은 특징값들에 가중치를 부여하고, 중요 특징값들을 이용하여 상기 학생 모델을 추가 학습시키는 단계, 그리고 임의의 이미지를 상기 학생 모델에 입력하고, 상기 임의의 이미지의 해상도를 높인 이미지를 출력하는 단계를 포함하고, 상기 학생 모델은 상기 교사 모델보다 같거나 작은 크기의 딥러닝 모델이다.

대 표 도 - 도1



(52) CPC특허분류
G06N 3/08 (2013.01)
G06T 3/4046 (2013.01)
G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 1711116161
과제번호 202039017.01
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 대학ICT연구센터지원사업
연구과제명 인공지능 서비스 실현을 위한 지능형 반도체 설계 핵심기술 개발
기 여 율 1/2
과제수행기관명 서강대학교 산학협력단
연구기간 2020.01.01~2020.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 20PQWO-B153359-02
과제번호 202055002
부처명 행정안전부
과제관리(전문)기관명 국토교통과학기술진흥원
연구사업명 스마트도로조명플랫폼개발및실증연구
연구과제명 스마트 도로조명 활용 도시재난안전관리 연계 기술 개발
기 여 율 1/2
과제수행기관명 서강대학교 산학협력단
연구기간 2020.01.01~2020.12.31

명 세 서

청구범위

청구항 1

적어도 하나의 프로세서에 의해 동작하는 컴퓨팅 장치가 동작하는 방법으로서,

저해상도 이미지와 상기 저해상도 이미지에 대응되는 고해상도 이미지가 포함된 학습 데이터를 이용하여 교사 모델을 학습시키고, 상기 학습 데이터로 학생 모델을 초기 학습시키는 단계,

상기 교사 모델의 학습 과정에서 생성된 특징값들 중 중요도가 높은 특징값들에 가중치를 부여하고, 중요 특징값들을 이용하여 상기 학생 모델을 추가 학습시키는 단계, 그리고

임의의 이미지를 상기 학생 모델에 입력하고, 상기 임의의 이미지의 해상도를 높인 이미지를 출력하는 단계를 포함하고,

상기 학생 모델은 상기 교사 모델보다 같거나 작은 크기의 딥러닝 모델인, 동작 방법.

청구항 2

제1항에서,

상기 교사 모델은,

상기 학습 데이터의 특징을 추출하는 복수의 컨벌루션 레이어들, 상기 컨벌루션 레이어들의 결과를 전달하는 활성화 함수, 그리고 상기 활성화 함수의 결과를 스케일링하는 멀티 레이어를 포함하는 잔여 블록(Residual Block)을 적어도 하나 이상 포함하는, 동작 방법.

청구항 3

제2항에서,

상기 학생 모델을 추가 학습시키는 단계는,

상기 교사 모델의 각 잔여 블록의 출력값들을 이용하여 상기 학생 모델의 손실 함수를 수정하는, 동작 방법.

청구항 4

제3항에서,

상기 학생 모델을 추가 학습시키는 단계는,

상기 각 잔여 블록의 출력값들 중 상기 교사 모델이 상기 저해상도 이미지로부터 상기 고해상도 이미지를 생성하기 위한 중요한 정보로 판단되어 가중치가 부여된 잔여 블록의 출력값들을 상기 중요 특징값들로 판단하는, 동작 방법.

청구항 5

컴퓨팅 장치로서,

메모리, 그리고

상기 메모리에 로드된 프로그램의 명령들(instructions)을 실행하는 적어도 하나의 프로세서를 포함하고,

상기 프로그램은

저해상도 이미지와 상기 저해상도 이미지에 대응되는 고해상도 이미지가 포함된 학습 데이터를 이용하여 학생 모델을 초기 학습시키는 단계,

상기 학습 데이터로 학습이 완료된 교사 모델로부터, 상기 교사 모델의 학습 과정에서 생성된 출력값들을 추출하고, 상기 출력값들을 이용하여 상기 학생 모델을 재학습시키는 단계, 그리고

상기 학생 모델에 임의의 저해상도 이미지를 입력하고, 상기 임의의 이미지의 고해상도 이미지를 출력하는 단계를 실행하도록 기술된 명령들을 포함하고,

상기 학생 모델은 상기 교사 모델보다 같거나 작은 크기의 딥러닝 모델인, 컴퓨팅 장치.

청구항 6

제5항에서,

상기 재학습시키는 단계는,

상기 출력값들 중 상기 교사 모델이 상기 저해상도 이미지로부터 상기 고해상도 이미지를 생성하기 위한 중요한 정보로 결정된 중요 출력값들에 가중치를 부여하고, 가중치가 부여된 중요 출력값들을 이용하여 상기 학생 모델의 손실 함수를 수정하는, 컴퓨팅 장치.

청구항 7

제6항에서,

상기 교사 모델과 상기 학생 모델은,

상기 학습 데이터의 특징을 추출하는 복수의 컨벌루션 레이어들과 상기 컨벌루션 레이어들의 결과를 전달하는 활성화 함수를 포함하는 블록(Block)을 적어도 하나 이상 포함하고,

상기 재학습시키는 단계는,

상기 교사 모델에 포함된 각 블록으로부터 출력값들을 추출하는, 컴퓨팅 장치.

청구항 8

제6항에서,

상기 재학습시키는 단계는,

상기 초기 학습에 사용된 손실 함수와, 상기 중요 출력값들과 상기 학생 모델의 초기 학습 과정에서 출력된 출력값들의 차이에 의한 손실 함수를 이용하는, 컴퓨팅 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 영상의 해상도를 높이는 기술에 관한 것이다.

배 경 기 술

[0002] 최근 Ultra-High Definition(UHD) 등의 고해상도 디스플레이가 시장에 등장하고, 이에 따라 높은 해상도에 대한 소비자의 요구가 커지게 되었다. 따라서 기존의 Full-High Definition(FHD)와 같은 저해상도(Low Resolution, LR) 영상을 고해상도(High Resolution, HR) 영상으로 변환할 수 있는 초해상화 (Super-Resolution, SR) 알고리즘에 대한 관심이 커지고 있다.

[0003] 예를 들어, 바이큐빅 보간법(Bicubic Interpolation)과 같은 다항식 기반의 보간법 또는 선형 매핑(Linear Mapping)을 사용하는 로컬 패치(Local Patch) 기반의 초해상화 기법은 낮은 복잡도와 적은 연산량으로 고품질의 고해상도 영상을 생성해낼 수 있다. 따라서 이러한 기법들은 실제 디스플레이 시스템에 적용되는데 문제가 없다.

[0004] 그러나 이러한 기존의 초해상화 기법들은 영상 복원 과정에서 디테일한 부분의 화질 저화 및 열화가 존재한다. 또한 기존 기법들은 비교적 단순한 선형 매핑을 기반으로 하기 때문에, 복잡하고 비선형적인 저해상도-고해상도 모델을 구현하기 힘들다.

[0005] 위의 문제를 해결하기 위해 최근 딥러닝(Deep learning) 기반의 초해상화 알고리즘에 대해 다양한 연구가 진행중이며, 기존 알고리즘 대비 높은 성능을 보여주고 있다.

[0006] 특히 합성곱 신경망(Convolutional Neural Network, CNN) 기반의 초해상화 네트워크는 기존의 알고리즘 대비 높

은 성능을 보여주고 있어, 디스플레이 시스템에 이를 구현하려는 시도가 진행 중이다. 합성곱 신경망은 여러 겹으로 쌓은 다층 네트워크를 사용하여, 저해상도 입력과 고해상도 출력 간의 복잡한 비선형적 관계를 정밀하게 분석한다. 합성곱 필터 파라미터(Filter Parameter)를 학습하기 때문에, 단순한 선형 매핑을 기반으로 하는 기존의 알고리즘보다 우수한 성능을 보여준다.

[0007] 그러나 딥러닝 기반의 초해상화 알고리즘은 많은 수의 레이어와 매개 변수를 수반하여 높은 복잡도를 갖고, 연산량과 메모리 소비량이 증가하는 문제가 발생할 수 있다. 예를 들어, 대표적인 딥러닝 기반의 초해상화 네트워크인 Very Deep Super Resolution(VDSR)은 20겹의 층을 가지고 있고 60만개 이상의 필터 파라미터를 요구한다.

[0008] 이로 인해 딥러닝 기반의 초해상화 알고리즘을 실제 디스플레이에 구현하는 과정에서 문제가 발생된다. 구체적으로, 하드웨어 리소스가 제한된 내장형 시스템 및 모바일 장치와 같은 저전력 장치는 전력 소비와 메모리 크기가 제한적이므로, 딥러닝 기반 초해상화 알고리즘을 배포하는 데 장애가 될 수 있다. 따라서 딥러닝 기반의 초해상화 알고리즘을 경량화하면서도 동시에 성능 감소를 최소화하는 방법이 필요하다.

발명의 내용

해결하려는 과제

[0009] 해결하고자 하는 과제는 학습 이미지를 이용하여 교사 네트워크를 학습시킨 후, 학습된 교사 네트워크의 정보들을 이용하여 경량화된 학생 네트워크를 학습시키고, 학생 네트워크를 이용하여 해상도가 높은 이미지를 출력하는 방법을 제공하는 것이다.

과제의 해결 수단

[0010] 한 실시예에 따른 적어도 하나의 프로세서에 의해 동작하는 컴퓨팅 장치가 동작하는 방법으로서, 저해상도 이미지와 상기 저해상도 이미지에 대응되는 고해상도 이미지가 포함된 학습 데이터를 이용하여 교사 모델을 학습시키고, 상기 학습 데이터로 학생 모델을 초기 학습시키는 단계, 상기 교사 모델의 학습 과정에서 생성된 특징값들 중 중요도가 높은 특징값들에 가중치를 부여하고, 중요 특징값들을 이용하여 상기 학생 모델을 추가 학습시키는 단계, 그리고 임의의 이미지를 상기 학생 모델에 입력하고, 상기 임의의 이미지의 해상도를 높인 이미지를 출력하는 단계를 포함한다. 상기 학생 모델은 상기 교사 모델보다 같거나 작은 크기의 딥러닝 모델이다.

[0011] 상기 교사 모델은, 상기 학습 데이터의 특징을 추출하는 복수의 컨벌루션 레이어들, 상기 컨벌루션 레이어들의 결과를 전달하는 활성화 함수, 그리고 상기 활성화 함수의 결과를 스케일링하는 멀티 레이어를 포함하는 잔여 블록(Residual Block)을 적어도 하나 이상 포함할 수 있다.

[0012] 상기 학생 모델을 추가 학습시키는 단계는, 상기 교사 모델의 각 잔여 블록의 출력값들을 이용하여 상기 학생 모델의 손실 함수를 수정할 수 있다.

[0013] 상기 학생 모델을 추가 학습시키는 단계는, 상기 각 잔여 블록의 출력값들 중 상기 교사 모델이 상기 저해상도 이미지로부터 상기 고해상도 이미지를 생성하기 위한 중요한 정보로 판단되어 가중치가 부여된 잔여 블록의 출력값들을 상기 중요 특징값들로 판단할 수 있다.

[0014] 한 실시예에 따른 컴퓨팅 장치로서, 메모리, 그리고 상기 메모리에 로드된 프로그램의 명령들(instructions)을 실행하는 적어도 하나의 프로세서를 포함한다. 상기 프로그램은 저해상도 이미지와 상기 저해상도 이미지에 대응되는 고해상도 이미지가 포함된 학습 데이터를 이용하여 학생 모델을 초기 학습시키는 단계, 상기 학습 데이터로 학습이 완료된 교사 모델로부터, 상기 교사 모델의 학습 과정에서 생성된 출력값들을 추출하고, 상기 출력값들을 이용하여 상기 학생 모델을 재학습시키는 단계, 그리고 상기 학생 모델에 임의의 저해상도 이미지를 입력하고, 상기 임의의 이미지의 고해상도 이미지를 출력하는 단계를 실행하도록 기술된 명령들을 포함한다. 상기 학생 모델은 상기 교사 모델보다 같거나 작은 크기의 딥러닝 모델이다.

[0015] 상기 재학습시키는 단계는, 상기 출력값들 중 상기 교사 모델이 상기 저해상도 이미지로부터 상기 고해상도 이미지를 생성하기 위한 중요한 정보로 결정된 중요 출력값들에 가중치를 부여하고, 가중치가 부여된 중요 출력값들을 이용하여 상기 학생 모델의 손실 함수를 수정할 수 있다.

[0016] 상기 교사 모델과 상기 학생 모델은, 상기 학습 데이터의 특징을 추출하는 복수의 컨벌루션 레이어들과 상기 컨벌루션 레이어들의 결과를 전달하는 활성화 함수를 포함하는 블록(Block)을 적어도 하나 이상 포함할 수 있다. 상기 재학습시키는 단계는, 상기 교사 모델에 포함된 각 블록으로부터 출력값들을 추출할 수 있다.

[0017] 상기 재학습시키는 단계는, 상기 초기 학습에 사용된 손실 함수와, 상기 중요 출력값들과 상기 학생 모델의 초기 학습 과정에서 출력된 출력값들의 차이에 의한 손실 함수를 이용할 수 있다.

발명의 효과

[0018] 본 발명에 따르면 매개 변수의 수가 적은 학생 네트워크를 이용하여 이미지의 해상도를 높일 수 있으므로, 계산량과 메모리 소비를 줄이고 리소스에 제한이 있는 하드웨어에도 배포할 수 있다.

[0019] 또한 본 발명에 따르면 교사 네트워크의 학습 단계에서 중요성이 높은 특정 정보를 지식 증류 기법으로 학생 네트워크에 전달하므로, 네트워크의 구조를 변경하거나 매개 변수의 수를 늘리지 않고도 학생 네트워크를 학습시킬 수 있다.

도면의 간단한 설명

[0020] 도 1은 한 실시예에 따른 영상 초해상화 장치의 구성도이다.

도 2는 한 실시예에 따른 교사 네트워크의 설명도이다.

도 3은 한 실시예에 따른 학생 네트워크의 설명도이다.

도 4는 한 실시예에 따른 영상 초해상화 장치의 동작 방법의 흐름도이다.

도 5와 도 6은 한 실시예에 따른 지식 전파 방법의 설명도이다.

도 7은 한 실시예에 따른 영상 초해상화 장치의 동작 방법의 예시도이다.

도 8은 한 실시예에 따른 컴퓨팅 장치의 하드웨어 구성도이다.

발명을 실시하기 위한 구체적인 내용

[0021] 아래에서는 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0022] 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" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

[0023] 이하에서는, 본 발명에 적용되는 네트워크 경량화 방법과 지식 증류 기법에 대해 간략히 설명한다.

[0024] 먼저, 경량 아키텍처(Lightweight Architecture)란 최소의 매개 변수로 인공 신경망을 보다 효율적으로 학습시키기 위한 네트워크를 의미한다. 예를 들어 ResNet, DensNet, MobileNet 등이 있다. 특히 SqueezeNet은 AlexNet과 비교할때 50배 적은 매개 변수로도 비슷한 성능을 나타낸다. 구체적으로, SqueezeNet은 3*3 컨볼루션 필터를 1*1과 3*3 필터의 혼합으로 대체하는 Fire 모듈을 사용한다. Fire 모듈은 스퀴즈 레이어와 확장 레이어를 포함한다. 스퀴즈 레이어는 1*1 필터로만 구성되어 있고, 확장 레이어는 1*1 필터와 3*3 필터를 포함한다. 이하 SqueezeNet은 공지된 기술이므로, 자세한 설명은 생략한다.

[0025] 지식 증류(Knowledge Distillation)란, 앙상블(Ensemble) 기법을 통해 학습된 다수의 큰 교사 네트워크들로부터 작은 하나의 학생 네트워크에 지식을 전달하는 방법 중 하나이다. 이때 교사 네트워크는 고성능 신경 네트워크이고, 학생 네트워크는 얕고 성능이 낮은 경우가 일반적이다. 교사 네트워크 및 학생 네트워크의 출력 값만 사용하거나, 숨겨진 특징값(Features)들을 이용하는 연구들이 진행되었다.

[0026] 본 발명은 지식 증류 기법을 활용해 딥러닝 기반의 초해상화 네트워크를 효율적으로 학습시킬 수 있는 방법과, 딥러닝 네트워크 경량화 기법 중 하나인 SqueezeNet을 활용하여 경량화에 따른 성능 감소를 최소화한 네트워크를 제안한다. 이하에서는 본 발명이 제안하는 영상 초해상화 장치에 대해 설명한다.

[0027] 도 1은 한 실시예에 따른 영상 초해상화 장치의 구성도이고, 도 2는 한 실시예에 따른 교사 네트워크의 설명도

이고, 도 3은 한 실시예에 따른 학생 네트워크의 설명도이다.

[0028] 도 1을 참고하면, 영상 초해상화 장치(10)는 저해상도의 이미지를 입력받아 고해상도 이미지로 복원한다. 이때 교사 네트워크(100)를 미리 학습시킨 후, 학습된 교사 네트워크(100)를 이용하여 경량화된 학생 네트워크(200)를 학습시키고, 학생 네트워크(200)를 이용하여 이미지 해상도를 높일 수 있다.

[0029] 지식 증류 기법을 이용하여, 이미 학습이 완료된 교사 네트워크(100)의 정보들을 바탕으로, 학생 네트워크(200)의 가중치 파라미터들을 학습시킬 수 있다.

[0030] 이때 학생 네트워크(200)는 교사 네트워크(100)의 행동을 모사하는 방식으로 학습된다. 이를 통해 학생 네트워크(200)는 일반적인 역전파(Backpropagation) 알고리즘을 사용하여 처음부터 학습한 경우보다 높은 인식률을 얻을 수 있다.

[0031] 교사 네트워크(100)와 학생 네트워크(200)는 하나의 컴퓨팅 장치에 구현되거나, 별도의 컴퓨팅 장치에 분산 구현될 수 있다. 별도의 컴퓨팅 장치에 분산 구현된 경우, 교사 네트워크(100)와 학생 네트워크(200)는 통신 인터페이스를 통해 서로 통신할 수 있다. 컴퓨팅 장치는 본 발명을 수행하도록 작성된 소프트웨어 프로그램을 실행할 수 있는 장치이면 충분하고, 예를 들면, 서버, 랩탑 컴퓨터 등일 수 있다.

[0032] 교사 네트워크(100)와 학생 네트워크(200) 각각은 하나의 인공지능 모델일 수 있고, 복수의 인공지능 모델로 구현될 수도 있다. 그리고 영상 초해상화 장치(10)도 하나의 인공지능 모델일 수 있고, 복수의 인공지능 모델로 구현될 수도 있다. 이에 따라, 상술한 구성들에 대응하는 하나 또는 복수의 인공지능 모델은 하나 또는 복수의 컴퓨팅 장치에 의해 구현될 수 있다.

[0033] 도 2를 참고하면, 교사 네트워크(100)는 컨벌루션 레이어, 잔여 블록(Residual Block) 그리고 업스케일링 모듈을 포함한다. 교사 네트워크(100)는 EDSR(Enhanced Deep Super Resolution)로 구현될 수 있으며, 이하 교사 네트워크(100)를 구성하는 각 요소들에 대해 설명한다.

[0034] 컨벌루션 레이어는 입력되는 저해상도 이미지의 특징을 추출하는 역할을 한다.

[0035] 각 잔여 블록은 9개의 잔여 모듈(Residual Module)을 포함하며, 하나의 잔여 모듈은 2개의 컨벌루션 레이어, 하나의 정류 선형 유닛(Rectified Linear Unit, ReLU) 그리고 멀티 레이어로 구성되어 있다.

[0036] 정류 선형 유닛은 활성화 함수 중 하나이고, 멀티 레이어는 각 모듈의 마지막에 특정 상수값을 곱해주는 역할을 한다. 상수값은 한 예로 0.1일 수 있다.

[0037] 업스케일링 모듈은 이미지의 해상도를 높이는 역할을 하고, Sub-pixel Convolutional Neural Network(SPCNN)으로 구현될 수 있으며, 바이큐빅 보간법(Bicubic Interpolation)과 같은 다항식 보간법(Polynomial Interpolation)보다 좋은 성능을 나타낸다.

[0038] 3개의 잔여 블록 및 9개의 잔여 모듈은 서로 스킵 커넥션(Skip Connection)으로 이어질 수 있다. 스킵 커넥션이란, 하나의 컨벌루션 레이어의 입력값을 다른 컨벌루션 레이어의 출력값에 직접 더하여, 각 컨벌루션 레이어는 입력과 출력의 차이값(Residual)만을 예측하도록 하는 방법이다.

[0039] 도 3을 참고하면, 학생 네트워크(200)는 컨벌루션 레이어, 스퀴즈 블록(Squeeze Block) 그리고 업스케일링 모듈을 포함한다. 학생 네트워크(200)는 SqueezeSR로 구현될 수 있으며, 이하 학생 네트워크(200)를 구성하는 각 요소들에 대해 설명한다.

[0040] 교사 네트워크(100)와 마찬가지로, 처음의 컨벌루션 레이어는 입력되는 저해상도 이미지의 특징들을 추출한다. 이후 3개의 스퀴즈 블록이 연결되며, 각 스퀴즈 블록은 스퀴즈 레이어와 확장 레이어를 포함한다. 한편 각 스퀴즈 블록은 스킵 커넥션으로 연결될 수 있다.

[0041] 스퀴즈 레이어는 1x1 컨벌루션 필터를 의미하고, 확장 레이어는 1x1 컨벌루션 필터와 3x3 컨벌루션 필터가 연결된 것을 의미한다.

[0042] 스퀴즈 레이어는 입력 채널의 수를 줄여서 다음 레이어에 전달한다. 즉 확장 레이어에 전달되는 채널의 크기를 줄일 수 있다. 이는 일반적인 CNN 모델에 비해 파라미터 수가 9배 감소함을 의미한다.

[0043] 한편 공지된 SqueezeNet의 형태와 달리, 본 발명의 학생 네트워크(200)는 배치 정규화(Batch Normalization)를 위한 레이어와 다운 샘플링 계층을 포함하지 않는다.

[0044] 이미지의 해상도를 높이는 업스케일링 모듈은 교사 네트워크(100)와 마찬가지로 SPCNN으로 구현될 수 있다.

[0045] 이하에서는 학생 네트워크(200)가 교사 네트워크(100)의 학습 과정에서 생성된 정보를 받아 추가 학습하는 과정에 대해 설명한다.

[0046] 도 4는 한 실시예에 따른 영상 초해상화 장치의 동작 방법의 흐름도이고, 도 5와 도 6은 한 실시예에 따른 지식 전파 방법의 설명도이다.

[0047] 도 4를 참고하면, 영상 초해상화 장치(10)는 저해상도 이미지와 이에 대응되는 고해상도 이미지를 포함하는 학습 데이터를 이용하여 교사 네트워크(100)와 학생 네트워크(200)를 초기 학습시킨다(S110). 이때 초기 학습에 사용되는 이미지는 컬러 이미지로서, RGB의 3개의 색상 채널을 갖는 이미지일 수 있다.

[0048] 영상 초해상화 장치(10)는 학습된 교사 네트워크(100)와 학생 네트워크(200)를 TAT(Teacher Attention Transfer) 모듈로 연결한다(S120).

[0049] 학생 네트워크(200)의 특징맵과 교사 네트워크(100)의 특징맵이 유사하므로, 교사 네트워크(100)의 학습 과정에서 중요하다고 판단된 특징맵 정보를 학생 네트워크(200)에 전달한다면, 학생 네트워크(200)를 더 효율적으로 학습시킬 수 있다.

[0050] 일반적으로 합성곱 신경망(Convolutional Neural Network)에서는 3개의 색상 채널을 갖는 3차원 입력 데이터의 각 채널마다 필터를 사용하여 합성곱 연산을 수행한다. 이때 필터는 입력된 이미지의 특징을 찾아내는 역할을 한다. 필터의 크기는 이미지의 크기보다 작으므로, 필터는 이미지의 일부분의 특성을 반영하여 특징맵이라는 결과를 출력한다.

[0051] 이때, 특정 필터의 출력값들은 이미지 해상도를 높이는 학습 과정에서 더 중요한 요소로 간주될 수 있다. 따라서 이러한 필터의 출력값에는 가중치를 부여하여 중요도를 반영할 수 있다.

[0052] 한편 가중치를 부여하는 방법으로서 어텐션 메커니즘 기반의 TAT 모듈을 사용할 수 있다.

[0053] 어텐션 메커니즘(Attention Mechanism)이란, 모델로 하여금 중요한 부분에 집중하도록 하는 구조로서, 디코더에서 출력값을 예측하는 매 시점마다, 인코더에서의 전체 입력값을 다시 참고한다. 이때 입력값 전체를 동일한 비율로 참고하는 것이 아니라, 예측해야 할 값과 연관이 있는 입력값의 특정 부분에 집중(Attention)한다. 한편 딥러닝 모델에서, 어텐션은 가중치로 표현되는 중요도 벡터로 구현될 수 있다. 어텐션 메커니즘에 관한 내용은 이미 공지된 기술인바, 자세한 설명은 생략한다.

[0054] 즉, TAT 모듈은 학생 네트워크(200)가 교사 네트워크(100)의 정보를 받을 때, 교사 네트워크(100)의 중요하다고 판단된 필터의 출력값에 초점을 맞추도록 하여, 채널 별로 중요도가 높은 정보에 대한 가중치를 학생 네트워크(200)에 전달해 학생 네트워크(200)가 학습 시 보다 중요한 정보에 집중하여 학습하도록 도와준다.

[0055] 영상 초해상화 장치(10)는 교사 네트워크(100)의 채널별 중요도가 높은 정보에 대한 가중치를 포함한 정보를 학생 네트워크(200)로 전달한다(S130).

[0056] 구체적으로, TAT 모듈은 교사 네트워크(100)를 구성하는 각 잔여 블록으로부터 출력된 특징맵인 t_n을 얻고, 학생 네트워크(200)를 구성하는 각 스퀴즈 블록으로부터 출력 특징맵 s_n을 얻는다. 이때 n은 잔여 블록 또는 스퀴즈 블록의 개수이며, 도 5의 예를 참고하면 1부터 3까지의 자연수이다.

[0057] TAT 모듈은 각 블록의 출력값을 이용하여 손실 함수를 계산하는데, 이때 손실 함수에 입력되는 값 즉 TAT 모듈의 출력값은 수학식 1을 통해 구할 수 있다.

[0058] [수학식 1]

$$T_i = \left(\frac{1}{j}\sum_{j=1}^{C} t_i^j\right)^k, \qquad S_i = \left(\frac{1}{j}\sum_{j=1}^{C\prime} s_i^j\right)^k$$

[0059]

[0060] 수학식 1에서, k는 성능을 높이기 위한 변수를 의미하고, c는 교사 네트워크(100)의 출력 특징맵의 채널 크기를 의미하고, c'는 학생 네트워크(200)의 출력 특징맵의 채널 크기를 의미한다. T_n과 S_n은 최종 손실 함수에 입력되는 값이다.

[0061] k가 1인 경우, T_i는 t_i의 평균 특징맵, S_i는 s_i의 평균 특징맵을 의미할 수 있다.

[0062] 한편 교사 네트워크(100)의 출력 특징맵 t_n은 크기가 C*H*W인 벡터 또는 텐서일 수 있다. 이때 C는 출력 특징맵의 채널 크기, H는 출력 특징맵의 높이, W는 출력 특징맵의 너비를 의미한다.

[0063] 교사 네트워크(100)의 채널별 어텐션 벡터를 얻기 위해, 출력 특징맵의 평균을 추출하는 글로벌 평균 풀링(Global Average Pooling)을 이용할 수 있다. 글로벌 평균 풀링을 각 출력 특징맵 t_n에 적용하고, 이후 활성화 함수를 통해서 결과적으로 크기가 C*1*1인 텐서를 얻을 수 있다.

[0064] 이후 텐서에 학생 네트워크(200)의 출력 특징맵인 s_n을 순서대로 곱한다. 한편, 채널 단위의 텐서를 평균을 계산하여, 최종 텐서의 크기를 1*H*W로 만들 수 있다.

[0065] 영상 초해상화 장치(10)는 TAT 모듈을 통해 계산된 손실 함수를 추가로 이용하여 학생 네트워크(200)의 미세 조정(Fine-tuning)을 진행한다(S140). 이때 미세 조정에 사용되는 손실 함수는 수학식 2와 같을 수 있다.

[0066] [수학식 2]

$$Loss = loss_0(\hat{y}, y) + loss_1(S_1, T_1) + loss_2(S_2, T_2) + loss_3(S_3, T_3)$$

[0067]

[0068] 수학식 2에서, y는 학생 네트워크(200)를 거쳐 출력된 고해상도 이미지를 의미하고, $\hat{y}$는 학습에 사용되는 정답 이미지(Ground Truth, GT)를 의미한다. loss_0는 S110 단계에서 학생 네트워크(200)의 학습 결과에 따른 손실함수로서, 학생 네트워크(200)의 출력 이미지와 학습 데이터에 포함된 정답 이미지 간의 차이에 의해 계산된 손실함수를 의미한다. loss_1, loss_2, loss_3은 학생 네트워크(200)와 교사 네트워크(100)의 각 필터의 출력 특징맵 간 차이에 의해 계산된 손실함수를 의미한다. T_n은 교사 네트워크(100)를 구성하는 잔여 블록의 출력값에 대한 TAT 모듈의 출력값, S_n은 학생 네트워크(200)를 구성하는 스퀴즈 블록의 출력값에 대한 TAT 모듈의 출력값을 의미한다.

[0069] 즉 영상 초해상화 장치(10)의 최종적인 손실 함수는 학생 네트워크(200)의 분류 성능에 대한 손실 함수인 loss_0와 교사 네트워크(100)와 학생 네트워크(200)의 분류 결과의 차이를 나타내는 손실 함수인 loss_1 내지 loss_3의 합으로 계산된다.

[0070] 한편 본 명세서에서는 교사 네트워크(100)의 잔여 블록과 학생 네트워크(200)의 스퀴즈 블록이 각 3개이므로 3개의 손실함수 항이 추가된 것이며, 추가되는 손실함수 항의 개수는 반드시 이에 제한되는 것은 아니다.

[0071] 도 7은 한 실시예에 따른 영상 초해상화 장치의 동작 방법의 예시도이다.

[0072] 도 7의 (a)는 정답 이미지이고, 도 7의 (b)는 해당 이미지를 바이큐빅 보간법으로 복원한 결과이다. 바이큐빅 보간법(Bicubic Interpolation)은 인접한 16개 화소의 화소값과 거리에 따른 가중치의 곱을 이용하여 홀을 복원하는 보간법이다.

[0073] 도 7의 (c)는 학습된 교사 네트워크(100)의 복원 결과이고, 도 7의 (d)는 매개 변수를 추가하지 않고 교사 네트워크(100)로부터 학습된 학생 네트워크(200)의 복원 결과를 나타낸다.

[0074] 학생 네트워크(200)의 파라미터 수는 교사 네트워크(100)의 파라미터 수에 비해 100배 낮고, 알고리즘 처리 속도는 약 10배 정도 빠르기 때문에 학생 네트워크(200)는 실시간 처리가 가능하다.

[0075] 도 8은 한 실시예에 따른 컴퓨팅 장치의 하드웨어 구성도이다.

[0076] 도 8을 참고하면, 영상 초해상화 장치(10)는 적어도 하나의 프로세서에 의해 동작하는 컴퓨팅 장치(300)에서, 본 발명의 동작을 실행하도록 기술된 명령들(instructions)이 포함된 프로그램을 실행한다.

[0077] 컴퓨팅 장치(300)의 하드웨어는 적어도 하나의 프로세서(310), 메모리(320), 스토리지(330), 통신 인터페이스(340)를 포함할 수 있고, 버스를 통해 연결될 수 있다. 이외에도 입력 장치 및 출력 장치 등의 하드웨어가 포함될 수 있다. 컴퓨팅 장치(300)는 프로그램을 구동할 수 있는 운영 체제를 비롯한 각종 소프트웨어가 탑재될 수 있다.

[0078] 프로세서(310)는 컴퓨팅 장치(300)의 동작을 제어하는 장치로서, 프로그램에 포함된 명령들을 처리하는 다양한 형태의 프로세서(310)일 수 있고, 예를 들면, CPU(Central Processing Unit), MPU(Micro Processor Unit),

MCU(Micro Controller Unit), GPU(Graphic Processing Unit) 등 일 수 있다. 메모리(320)는 본 발명의 동작을 실행하도록 기술된 명령들이 프로세서(310)에 의해 처리되도록 해당 프로그램을 로드한다. 메모리(320)는 예를 들면, ROM(read only memory), RAM(random access memory) 등 일 수 있다. 스토리지(330)는 본 발명의 동작을 실행하는데 요구되는 각종 데이터, 프로그램 등을 저장한다. 통신 인터페이스(340)는 유/무선 통신 모듈일 수 있다.

[0079] 본 발명에 따르면 매개 변수의 수가 적은 학생 네트워크를 이용하여 이미지의 해상도를 높일 수 있으므로, 계산량과 메모리 소비를 줄이고 리소스에 제한이 있는 하드웨어에도 배포할 수 있다.

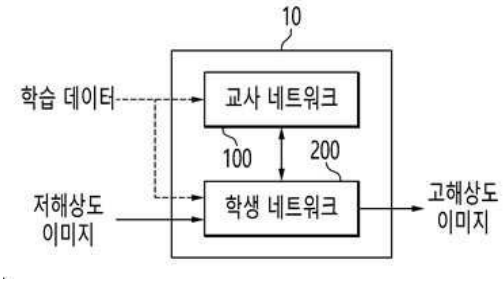
[0080] 또한 본 발명에 따르면 교사 네트워크의 학습 단계에서 중요성이 높은 특정 정보를 지식 증류 기법으로 학생 네트워크에 전달하므로, 네트워크의 구조를 변경하거나 매개 변수의 수를 늘리지 않고도 학생 네트워크를 학습시킬 수 있다.

[0081] 이상에서 설명한 본 발명의 실시예는 장치 및 방법을 통해서만 구현이 되는 것은 아니며, 본 발명의 실시예의 구성에 대응하는 기능을 실현하는 프로그램 또는 그 프로그램이 기록된 기록 매체를 통해 구현될 수도 있다.

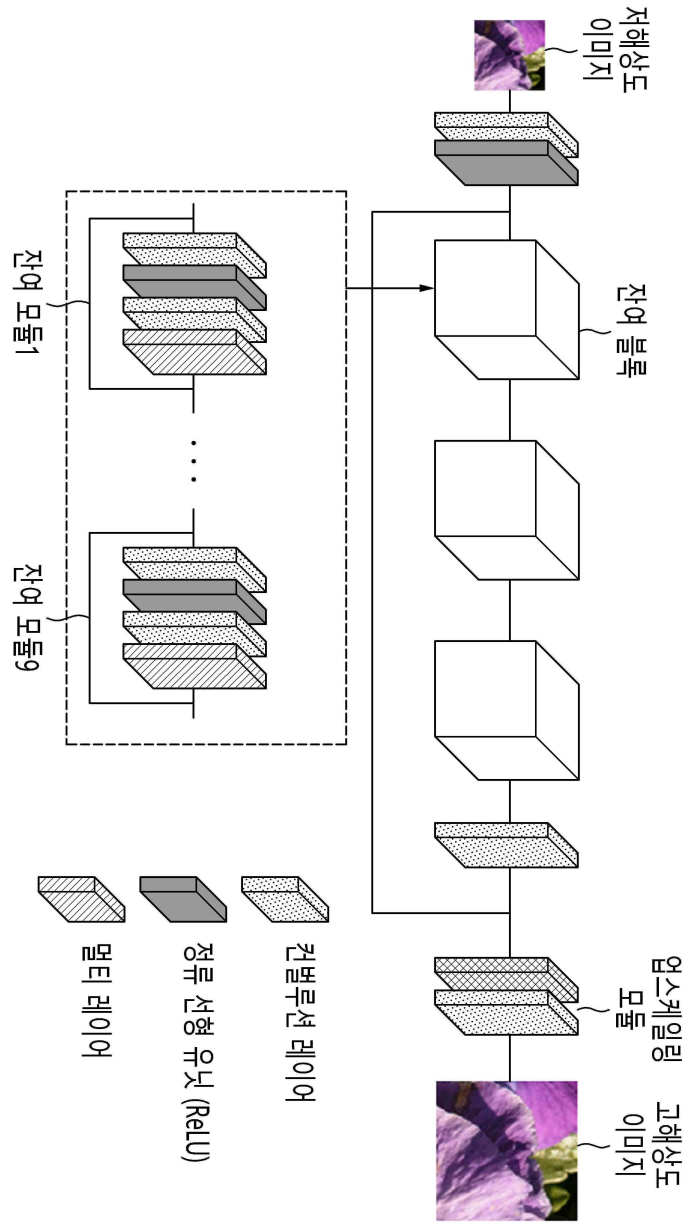
[0082] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

도면

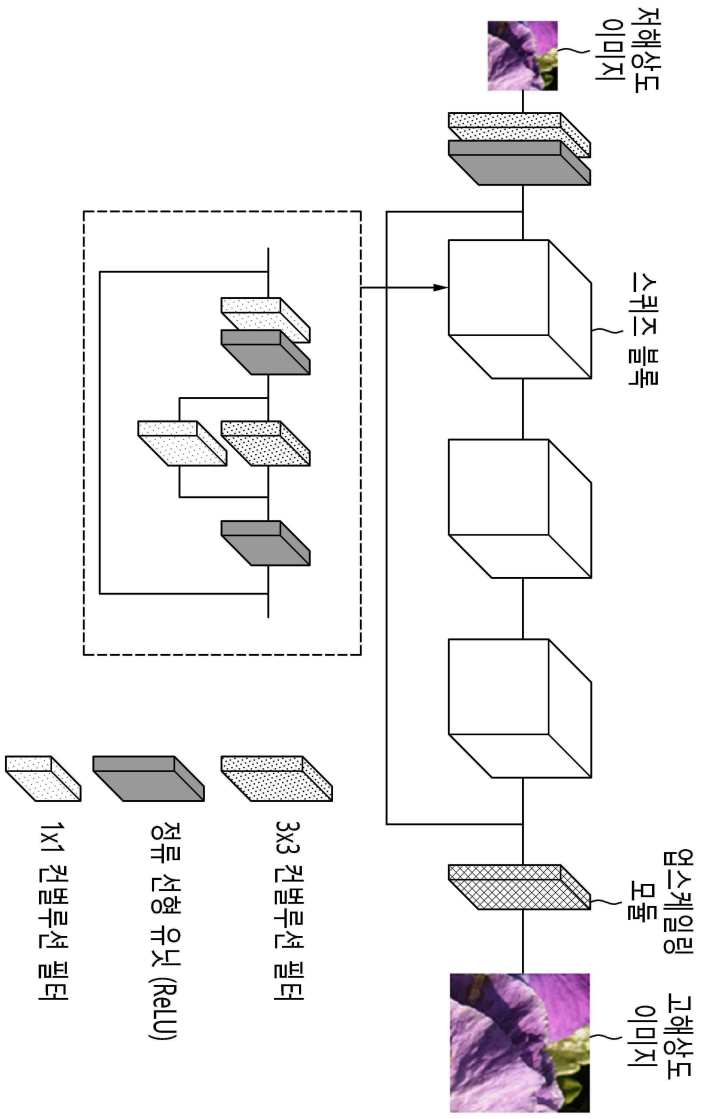
도면1



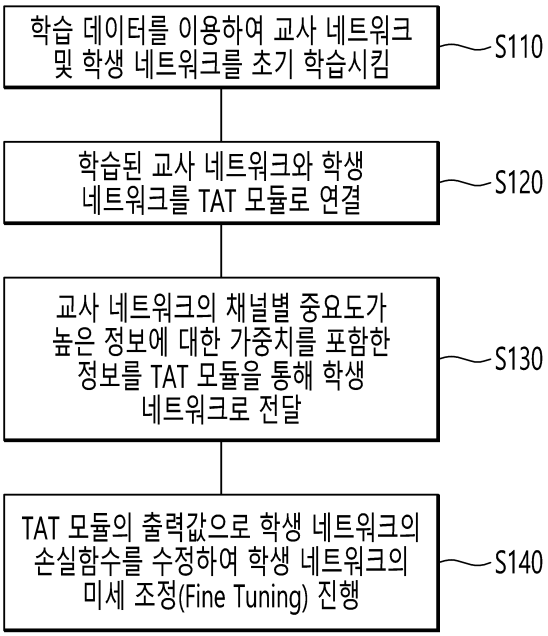
도면2


저해상도 이미지
잔여 블록
업스케일링 모듈
고해상도 이미지
잔여 모듈1
잔여 모듈9
컨벌루션 레이어
정류 선형 유닛 (ReLU)
멀티 레이어

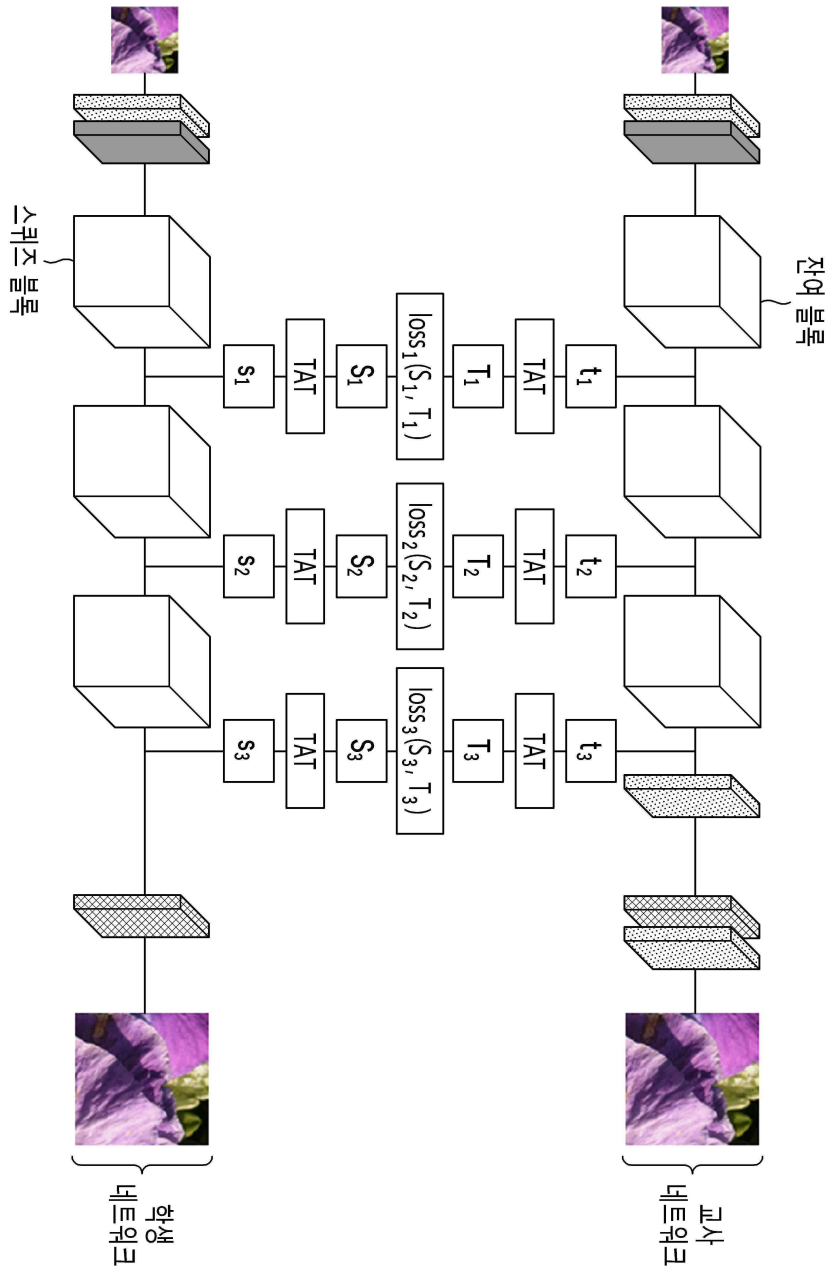
도면3


저해상도 이미지
스퀴즈 블록
업스케일링 모듈
고해상도 이미지
3x3 컨벌루션 필터
정류 선형 유닛 (ReLU)
1x1 컨벌루션 필터

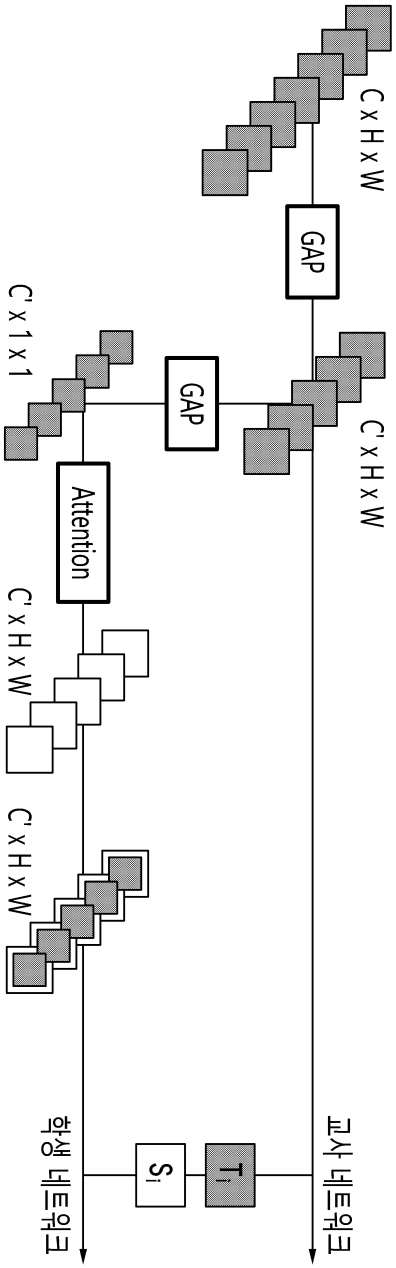
도면4


학습 데이터를 이용하여 교사 네트워크 및 학생 네트워크를 초기 학습시킴
S110
학습된 교사 네트워크와 학생 네트워크를 TAT 모듈로 연결
S120
교사 네트워크의 채널별 중요도가 높은 정보에 대한 가중치를 포함한 정보를 TAT 모듈을 통해 학생 네트워크로 전달
S130
TAT 모듈의 출력값으로 학생 네트워크의 손실함수를 수정하여 학생 네트워크의 미세 조정(Fine Tuning) 진행
S140

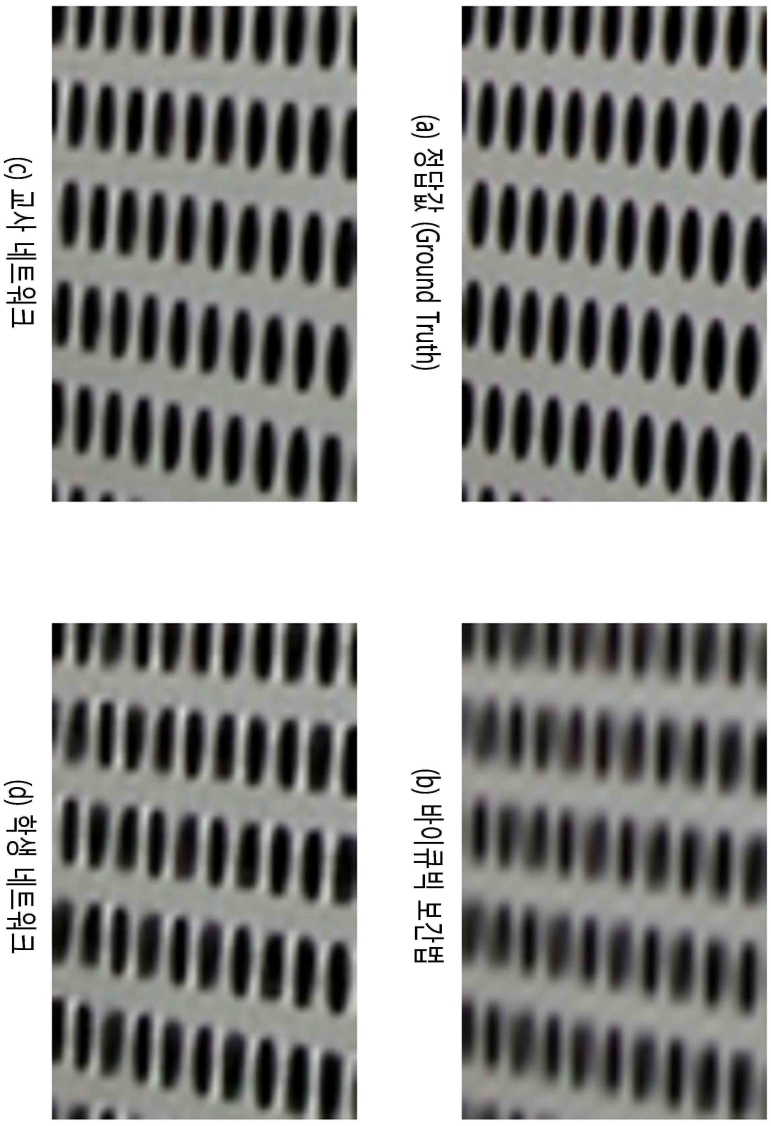
도면5


스퀴즈 블록
잔여 블록
TAT
s_1
s_2
s_3
t_1
t_2
t_3
S_1
S_2
S_3
T_1
T_2
T_3
$loss_1(S_1, T_1)$
$loss_2(S_2, T_2)$
$loss_3(S_3, T_3)$
학생 네트워크
교사 네트워크

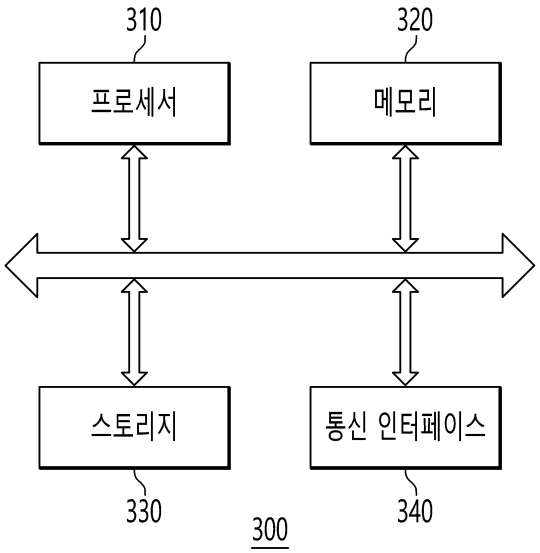
도면6


C x H x W
GAP
C' x 1 x 1
GAP
C' x H x W
Attention
C' x H x W
C' x H x W
학생 네트워크
S_i
T_i
교사 네트워크

도면7


(a) 정답값 (Ground Truth)
(b) 바이큐빅 보간법
(c) 교사 네트워크
(d) 학생 네트워크

도면8


310
320
프로세서
메모리
스토리지
통신 인터페이스
330
340
300