



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0178780
(43) 공개일자 2024년12월31일

- (51) 국제특허분류(Int. Cl.)
G06N 3/096 (2023.01) G06N 3/0464 (2023.01)
G06N 3/047 (2023.01) G06N 3/0475 (2023.01)
G06V 10/764 (2022.01) G06V 10/82 (2022.01)
- (52) CPC특허분류
G06N 3/096 (2023.01)
G06N 3/0464 (2023.01)
- (21) 출원번호 10-2023-0080936
(22) 출원일자 2023년06월23일
심사청구일자 2023년06월23일
- (71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
- (72) 발명자
배승환
인천광역시 미추홀구 인하로 100 (용현동)
천세권
경기도 용인시 수지구 상현로42번길 46, 265동 1003호 (상현동, 상현마을 동일스위트)
- (74) 대리인
특허법인태백

전체 청구항 수 : 총 5 항

(54) 발명의 명칭 교사 모델의 사전 지식을 이용한 지식 증류 시스템

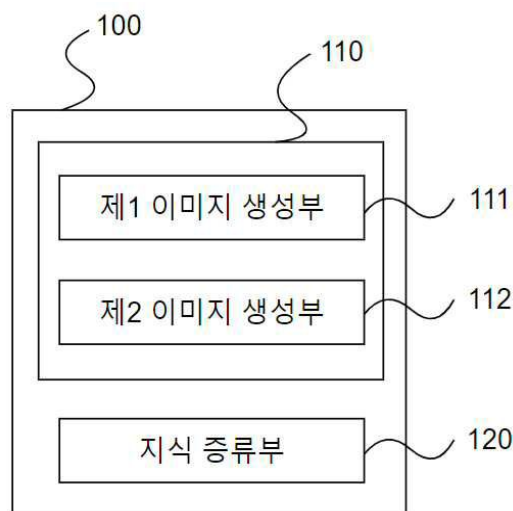
(57) 요약

본 발명은, 교사 모델의 사전 지식을 이용한 지식 증류 시스템에 관한 것이다.

본 발명에 따르면, 교사 모델의 사전 지식을 이용한 지식 증류 시스템은 기 정의된 하드 샘플링 손실(hard sampling loss)을 적용하여 학생 모델이 구분하기 어려운 이미지 데이터를 생성하는 이미지 생성부; 및 상기 이미지 생성부에서 생성한 이미지를 이용하여 지식 증류를 수행하는 지식 증류부를 포함한다.

이와 같이 본 발명의 실시예에 따르면, 데이터를 자동으로 생성할 수 있어, 대규모 학습데이터 수집이 어려운 경우나 데이터 도메인이 계속 변화하는 경우와 같이 학습 데이터를 확보하기 어려운 상황에서도 높은 정확도를 가진 고성능의 학생 모델을 학습시킬 수 있다.

대표도 - 도1



(52) CPC특허분류

G06N 3/047 (2023.01)

G06N 3/0475 (2023.01)

G06V 10/764 (2023.08)

G06V 10/82 (2022.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711190278
과제번호	2022R1C1C1009208
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)
연구과제명	대규모 확장 가능한 On-Vision AI를 위한 E3-AI 네트워크와 협업 학습 연구
기 여 율	60/100
과제수행기관명	인하대학교
연구기간	2023.03.01 ~ 2024.02.29

이 발명을 지원한 국가연구개발사업

과제고유번호	1345362086
과제번호	2022R1A6A1A03051705
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	이공학학술연구기반구축
연구과제명	3D 나노융합소자 연구 센터
기 여 율	20/100
과제수행기관명	인하대학교
연구기간	2023.03.01 ~ 2024.02.29

이 발명을 지원한 국가연구개발사업

과제고유번호	1711160586
과제번호	2022-0-00448-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	사람중심인공지능핵심원천기술개발
연구과제명	인간처럼 회상이 가능한 인공 신경망 지속학습 플랫폼 개발
기 여 율	10/100
과제수행기관명	인하대학교 산학협력단
연구기간	2022.04.01 ~ 2022.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711179347
과제번호	00155915
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	인공지능융합혁신인재양성(R&D)
연구과제명	인공지능융합혁신인재양성(인하대학교)
기 여 율	10/100
과제수행기관명	인하대학교 산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

기 정의된 하드 샘플링 손실(hard sampling loss)을 적용하여 학생 모델이 구분하기 어려운 이미지 데이터를 생성하는 이미지 생성부; 및

상기 이미지 생성부에서 생성한 이미지를 이용하여 지식 증류를 수행하는 지식 증류부를 포함하는 교사 모델의 사전 지식을 이용한 지식 증류 시스템.

청구항 2

제 1 항에 있어서,

상기 이미지 생성부는,

하기의 수학적식

$$L_{inv} = \lambda_t L_{CE}(p_t(\hat{x}), y) + \lambda_i L_i + \alpha_{tv} R_{tv}(\hat{x}) + \alpha_{BN} R_{BN}(\hat{x})$$

을 사용하여 이미지를 생성하고, p_t 는 교사 모델이며, $p_t(\hat{x})$ 은 교사 모델에 이미지 $\hat{x}$ 를 입력했을 때의 출력 분

포이고, $L_{CE}(p_t(\hat{x}), y)$ 는 레이블 y 와 이미지 $\hat{x}$ 사이의 교차 엔트로피(cross entropy)이며, $\lambda_i L_i$ 는 하드

샘플링 손실(hard sampling loss)이고, R_{tv} 는 이미지의 정규화 요소(regularization term)로, 이미지 $\hat{x}$ 의 전체

분산과 L2-노름(L2_norm)으로 구성되어 있으며, α_{tv} 는 R_{tv} 의 척도 인자(scaling factor)인 교사 모델의 사전 지식을 이용한 지식 증류 시스템.

청구항 3

제 2 항에 있어서,

이미지 생성부는

학생 모델의 손실을 최대화하여 학생 모델이 분류하기 어려운 이미지를 생성하는 제1 이미지 생성부; 및

학생 모델이 하나의 클래스로 정하기 어려워하는 이미지를 생성하기 위해 학생 모델이 균일분포 (uniform distribution)를 출력분포로 갖는 이미지를 생성하는 제2 이미지 생성부 중 적어도 하나를 포함하는 교사 모델의 사전 지식을 이용한 지식 증류 시스템.

청구항 4

제 3 항에 있어서,

상기 제1 이미지 생성부는

하기의 수학적식

$$L_{inv} = \lambda_t L_{CE}(p_t(\hat{x}), y) + \lambda_{NSL} L_{NSL} + \alpha_{tv} R_{tv}(\hat{x}) + \alpha_{BN} R_{BN}(\hat{x})$$

을 사용하여 이미지를 생성하고,

L_{NSL}

은 하기의 수학적식

$$L_{NSL} = -L_{CE}(p_s(\hat{x}), y)$$

로 정의되며, p_s 는 학습을 진행 중인 학생 모델이고, L_{NSL} 은 L_{NSL} 의 척도 인자인 교사 모델의 사전 지식을 이용한 지식 증류 시스템.

청구항 5

제 3 항에 있어서,

상기 제2 이미지 생성부는

하기의 수학적식

$$L_{inv} = \lambda_t L_{CE}(p_t(\hat{x}), y) + \lambda_{UDL} L_{UDL} + \alpha_{tv} R_{tv}(\hat{x}) + \alpha_{BN} R_{BN}(\hat{x})$$

을 사용하여 이미지를 생성하고,

L_{UDL}

은 하기의 수학적식

$$L_{UDL} = KL(p_s(\hat{x}), X_U)$$

로 정의되며, $KL(\cdot)$ 은 쿨백-라이블러(Kullback-Leibler) 발산을 의미하고, p_s 는 학습을 진행 중인 학생 모델이

고, 균일분포 X_U 는 클래스의 개수가 c 일 때 $X_U \in \mathbb{R}^c$ 이고, 모든 값을 $\frac{1}{c}$ 로 동일하게 갖는 분포이며, L_{UDL}

의 척도 인자인 교사 모델의 사전 지식을 이용한 지식 증류 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 교사 모델의 사전 지식을 이용한 지식 증류 시스템에 관한 것으로서, 더욱 상세하게는 학생 모델이 정확히 판별하기 어려운 이미지를 생성하여 최종적으로는 테스트 정확도를 향상하는 하드 샘플링(hard sampling) 이미지 생성 기법을 이용하여 지식 증류를 수행하는 교사 모델의 사전 지식을 이용한 지식 증류 시스템에 관한 것이다.

배경 기술

[0002] 지식 증류(knowledge distillation)는 높은 예측 정확도를 가진 교사 모델의 지식을 단순하고 가벼운 학생 모델에 이전하여, 학생 모델이 교사 모델과 유사한 높은 성능을 내도록 하는 딥러닝 방법 중 하나로, 높은 성능을 지녔지만 무거운 모델을 성능은 유지하면서 경량화하는 용도로 사용된다.

[0003] 최근 사생활 보호를 위해 데이터 사용이 법률로 인해 제한을 받고 있는 상황에서 임베디드 기기의 사용 증가로 경량화된 인공지능 모델이 요구됨에 따라 지식 증류 기법의 필요성이 커지고 있다.

[0004] 하지만, 기존의 지식 증류 방법은 이미지 분류를 수행하는 교사 모델과 학습 이미지, 분류 레이블로 이루어진 학습 데이터를 모두 사용하여, 교사 모델을 모방하는 학생 네트워크를 학습시키는데, 이 과정에서 학습에 사용

한 데이터를 필요로 한다.

[0005] 따라서, 기존의 지식 증류 방법은 학습 데이터를 확보하기 어려운 상황에서는 적용하기 어렵다는 한계가 있다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) KR 10-2368064(2022.02.22.)

비특허문헌

[0007] (비특허문헌 0001) Yin, Hongxu, et al. "Dreaming to distill: Data-free knowledge transfer via deepinversion." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020.

발명의 내용

해결하려는 과제

[0008] 본 발명에 따르면, 교사 모델과 학생 모델 간의 지식 증류에서 별도의 학습 데이터 없이도 지식 증류를 수행할 수 있는 교사 모델의 사전 지식을 이용한 지식 증류 시스템을 제공하기 위한 것이다.

과제의 해결 수단

[0009] 이러한 기술적 과제를 이루기 위한 본 발명의 일 실시예에 따르면 교사 모델의 사전 지식을 이용한 지식 증류 시스템은 기 정의된 하드 샘플링 손실(hard sampling loss)을 적용하여 학생 모델이 구분하기 어려운 이미지 데이터를 생성하는 이미지 생성부; 및 상기 이미지 생성부에서 생성한 이미지를 이용하여 지식 증류를 수행하는 지식 증류부를 포함한다.

발명의 효과

[0010] 이와 같이 본 발명에 따르면, 데이터를 자동으로 생성할 수 있어, 대규모 학습데이터 수집이 어려운 경우나 데이터 도메인이 계속 변화하는 경우와 같이 학습 데이터를 확보하기 어려운 상황에서도 높은 정확도를 가진 고성능의 학생 모델을 학습시킬 수 있다.

도면의 간단한 설명

[0011] 도 1은 본 발명의 일 실시예에 따른 교사 모델의 사전 지식을 이용한 지식 증류 시스템을 설명한 도면이다.

도 2 및 도 3은 본 발명의 일 실시예에 따라 지식 증류를 수행한 학생 모델과 종래 기술의 학생 모델의 정확도를 나타낸 도면이다.

도 4는 본 발명의 일 실시예에 따라 생성된 이미지와 종래 기술이 생성한 이미지의 예시를 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

[0012] 이하 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시예를 상세히 설명하기로 한다. 이 과정에서 도면에 도시된 선들의 두께나 구성요소의 크기 등은 설명의 명료성과 편의상 과장되게 도시되어 있을 수 있다.

[0013] 또한 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서, 이는 사용자, 운용자의 의도 또는 관례에 따라 달라질 수 있다. 그러므로 이러한 용어들에 대한 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.

[0014] 도 1은 본 발명의 일 실시예에 따른 교사 모델의 사전 지식을 이용한 지식 증류 시스템을 설명한 도면이다.

[0015] 도 1에 도시된 바와 같이 교사 모델의 사전 지식을 이용한 지식 증류 시스템(100)은 이미지 생성부(110) 및 지

식 증류부(120)를 포함한다.

[0016] 이 때, 이미지 생성부(110)는 제1 이미지 생성부(111) 및 제2 이미지 생성부(112) 중 적어도 하나를 포함한다.

[0017] 이미지 생성부(110)는 하드 샘플링 기법을 사용하여 학생 모델이 구분하기 어려운 이미지 데이터를 생성한다.

[0018] 더욱 구체적으로, 임의로 정한 레이블 y 와 가우시안 랜덤 노이즈로 만든 이미지 $\hat{x}$ 를 사용하여 하기의 수학식 1 을 통해 손실(loss)을 최적화하고 이미지를 생성한다.

수학식 1

$$L_{inv} = \lambda_t L_{CE}(p_t(\hat{x}), y) + \lambda_i L_i + \alpha_{tv} R_{tv}(\hat{x}) + \alpha_{BN} R_{BN}(\hat{x})$$

[0019]

[0020] 이 때, p_t 는 교사 모델이며, $p_t(\hat{x})$ 은 교사 모델에 이미지 $\hat{x}$ 를 입력했을 때의 출력 분포이고, $L_{CE}(p_t(\hat{x}), y)$ 에서 L_{CE} 는 레이블 y 와 이미지 $\hat{x}$ 사이의 교차 엔트로피(cross entropy)이다.

[0021] 또한, R_{tv} 는 이미지의 정규화 요소(regularization term)로, 이미지 $\hat{x}$ 의 전체 분산과 L2-노름(L2_norm)으로 구성되어 있으며, 이미지 $\hat{x}$ 의 전체 분산(total variance)에 페널티를 주어 인접한 픽셀들의 값을 비슷한 값으로 만들어 줌으로써 노이즈와 같은 이미지를 현실적인 이미지로 바꿔주는 역할을 수행하며, a_{tv} 는 R_{tv} 의 척도 인자(scaling factor)이다.

[0022] 또한, R_{BN} 의 경우, 이미지 $\hat{x}$ 가 실제 이미지의 데이터 분포를 더 잘 따라가게 만드는 것으로 하기의 수학식 2를 통해 정의된다.

수학식 2

$$R_{BN}(\hat{x}) = \sum_l \|\mu_l(\hat{x}) - BN_l(\text{learning_mean})\|_2 + \sum_l \|\sigma_l^2(\hat{x}) - BN_l(\text{learning_variance})\|_2,$$

[0023]

[0024] 이 때, $\mu_l(\hat{x})$ 와 $\sigma_l^2(\hat{x})$ 는 모델의 l 번째 컨볼루션 레이어(convolutional layer)에서 추정된 특징맵의 평균과 분산이다.

[0025] 또한, $\lambda_i L_i$ 는 하드 샘플링 손실(hard sampling loss)로, 기존의 이미지 생성 과정에 추가되며, 생성하는 이미지의 목적에 따라 기 설정된 형태 중 하나로 치환된다.

[0026] 제1 이미지 생성부(111)는 수학식 1에서 $\lambda_i L_i$ 를 $\lambda_{NSL} L_{NSL}$ 로 치환하여 하기의 수학식 3과 같이 학생 모델의 손실을 최대화하여 학생 모델이 분류하기 어려운 이미지를 생성한다.

수학식 3

$$L_{inv} = \lambda_t L_{CE}(p_t(\hat{x}), y) + \lambda_{NSL} L_{NSL} + \alpha_{tv} R_{tv}(\hat{x}) + \alpha_{BN} R_{BN}(\hat{x})$$

[0027]

이 때, L_{NSL} 은 하기의 수학식 4로 정의되며, 학습을 진행중인 학생 모델 p_s 에 $\hat{x}$ 를 입력해 나온 출력분포 $p_s(\hat{x})$ 와 레이블 y 간의 교차 엔트로피를 사용하여 학생 모델의 분류 손실(classification loss)을 직접 구하고, 이를 전체 손실에서 차감한다.

[0028]

수학식 4

$$L_{NSL} = -L_{CE}(p_s(\hat{x}), y)$$

[0029]

따라서, 학생 모델의 손실을 최대화하여 학생 모델이 분류하기 어려운 이미지를 생성하는 경우, $L_{CE}(p_s(\hat{x}), y)$ 를 최대화해야 하므로 학생 모델의 손실을 크게 만들기 위해 학생 모델이 맞추기 어려운 이미지를 생성하게 된다.

[0030]

또한, λ_{NSL} 은 L_{NSL} 의 척도 인자이다.

[0031]

반면, 제2 이미지 생성부(112)는 수학식 1에서 $\lambda_i L_i$ 를 $\lambda_{UDL} L_{UDL}$ 로 치환하여 하기의 수학식 5와 같이 학생 모델이 하나의 클래스로 정하기 어려워하는 이미지를 생성하기 위해 학생 모델이 균일분포(uniform distribution)를 출력분포로 갖는 이미지를 생성한다.

[0032]

수학식 5

$$L_{inv} = \lambda_t L_{CE}(p_t(\hat{x}), y) + \lambda_{UDL} L_{UDL} + \alpha_{tv} R_{tv}(\hat{x}) + \alpha_{BN} R_{BN}(\hat{x})$$

[0033]

이 때, L_{UDL} 은 하기의 수학식 6으로 정의되며, λ_{UDL} 는 L_{UDL} 의 척도 인자이다.

[0034]

수학식 6

$$L_{UDL} = KL(p_s(\hat{x}), X_U)$$

[0035]

수학식 6에서 $KL(\cdot)$ 은 쿨백-라이블러(Kullback-Leibler) 발산을 의미한다.

[0036]

제2 이미지 생성부(112)는 쿨백-라이블러를 통해 L_{UDL} 을 사용하여 생성한 이미지 $\hat{x}$ 의 출력분포 값과, 모든 레이블의 값들을 동일한 값으로 갖는 균일분포 X_U 간의 차이를 최소화한다.

[0037]

[0038] 이 때, $p_s(\hat{x})$ 는 이미지 $\hat{x}$ 을 학생 모델 p_s 에 입력했을 때 생성되는 출력분포이며, 균일분포 X_U 는 클래스의 개수 c 가 일 때 $X_U \in \mathbb{R}^c$ 이고, 모든 값을 $\frac{1}{c}$ 로 동일하게 갖는 분포이다.

[0039] 이를 통해, 손실을 줄이는 과정에서 $KL(p_s(\hat{x}), X_U)$ 를 최소화해야 하므로 학생 모델의 출력분포를 균일분포와 비슷하게 만들어주는 이미지를 생성하게 된다.

[0040] 지식 증류부(120)는 이미지 생성부(110)에서 생성된 이미지를 사용하여 지식 증류를 수행한다.

[0041] 이 때, 지식 증류를 수행하는 과정은, 통상의 기술자에게 공지된 방식을 적용할 수 있는 바, 이에 대한 구체적인 설명은 생략한다.

[0042] 본 발명의 일 실시예에 따라 지식 증류를 수행한 학생 모델과 종래 기술의 학생 모델의 성능을 검증하기 위해, 각각 딥인버전(DeepInversion, DI), 적응형 딥인버전(Adaptive DeepInversion, ADI), 부정적 학생 손실(Negative Student Loss, NSL), 균일분포 손실(Uniform Distribution Loss, UDL)을 사용하여 생성한 4개의 모델에 CIFAR-10 데이터셋을 사용한 이미지 분류를 수행시킨 후, 학생 모델의 정확도를 측정하였다.

[0043] 이 때, NSL은 본 발명의 일 실시예에 따라 제1 이미지 생성부(111)에서 설계한 손실이며, UDL은 제2 이미지 생성부(112)에서 설계한 손실이다.

[0044] 또한, CIFAR-10은 학습 데이터셋 50,000장, 테스트 데이터셋 10,000장으로 구성된 데이터셋으로, 10개 클래스 이미지가 포함되어 있다.

[0045] 이때, 교사 모델로 ResNet-34 모델을, 학생 모델로 ResNet-18 모델을 사용하였으며, CIFAR-10 데이터셋을 사용하는 과정에서 교사 모델을 학습할 때에는 학습 데이터셋을 사용해 학습을 진행하였고, 학생 모델을 학습하는 과정에서는 학습 데이터셋을 사용하지 않았다.

[0046] 또한, 모델의 정확도를 측정할 때는 테스트 데이터셋 만을 사용하였다.

[0047] 모든 실험에서 하이퍼파라미터 α_{pr} , α_{tv} , α_{BN} 을 10^{-3} , λ_{NSL} 을 1로 설정하였으며, UDL 실험에서는 수학적 1의 λ_{UDL} 을 10으로 설정하였다.

[0048] 또한, DI (DeepInversion)와 ADI (Adaptive DeepInversion)을 학습할 때는 비특허문헌 1에서 설정한 파라미터를 활용하였다.

[0049] 이 때, 미니배치(minibatch)는 32*32 해상도의 256개의 이미지로 구성 된다. 지식 증류 방법으로는 통상의 기술자에게 공지된 방식을 활용하였다.

[0050] 모델 학습은 총 250epoch에 걸쳐 수행되었다.

[0051] 옵티마이저는 SGD옵티마이저를 사용하였으며 학습률(learning rate)은 0.1로 설정하였고, 이때 학습률 감소를 위해 100epoch마다 0.1씩 곱하였다.

[0052] 지식 증류 과정에서 ADI, NSL, UDL의 학습과정의 경우 미리 생성해둔 DI 이미지 1000 배치를 가지고 학습을 시작하였으며, 1 epoch가 끝날 때 마다 학습을 중단시키고 2개의 배치를 각각의 손실에 따라 생성하여 학습 데이터에 추가하여 최종적으로 1500개의 배치의 이미지를 사용하였다.

[0053] DI방법을 학습할 때에는 처음부터 총 1500개의 배치를 사용하였고 학습 중에 따로 이미지를 생성하지 않았다.

[0054] 도 2 및 도 3는 본 발명의 일 실시예에 따라 지식 증류를 수행한 학생 모델과 종래 기술의 학생 모델의 정확도를 나타낸 도면이다.

[0055] 도 2 및 도 3에 도시된 바에 따르면, UDL 방법이 93.23%의 정확도를 달성해 가장 높은 정확도를 나타냈고, NSL 방법은 92.19%의 정확도로 측정이 되어 ADI의 정확도인 92.74% 보다 떨어지는 것을 확인하였다.

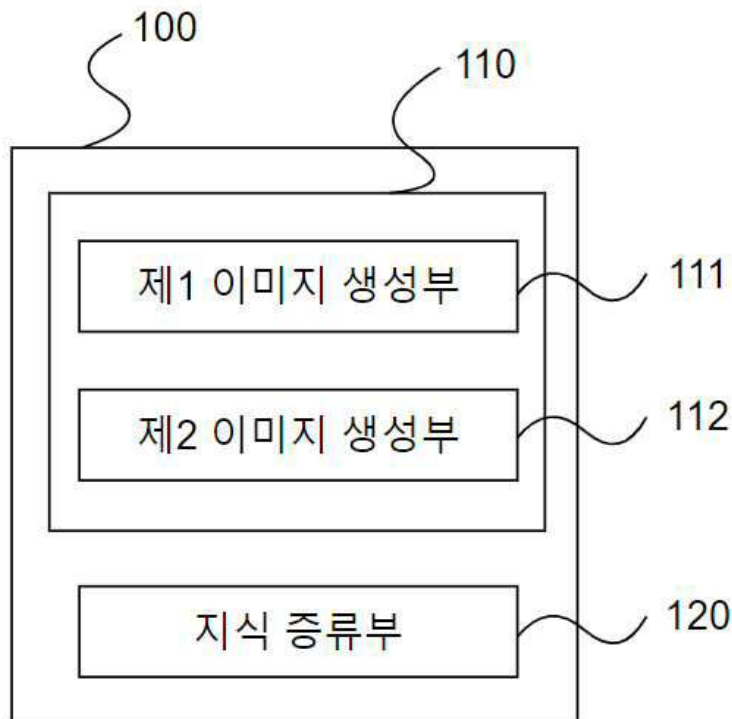
- [0056] 도 4는 본 발명의 일 실시예에 따라 생성된 이미지와 종래 기술이 생성한 이미지의 예시를 나타낸 도면이다.
- [0057] 도 4에 도시된 바에 따르면, NSL 의 경우 학생 모델이 맞추기 힘든 이미지를 무리하게 생성함으로써 전체적으로 사람도 구분하기 힘든 이미지를 생성하는 것을 보였고, UDL은 모든 클래스가 균일분포로 나오게 만드는 이미지를 생성해 냄으로써 여러 개의 사물이 섞인 것과 같은 이미지를 만들어 낸다.
- [0058] 이와 같이 본 발명에 따르면, 데이터를 자동으로 생성할 수 있어, 대규모 학습데이터 수집이 어려운 경우나 데이터 도메인이 계속 변화하는 경우와 같이 학습 데이터를 확보하기 어려운 상황에서도 높은 정확도를 가진 고성능의 학생 모델을 학습시킬 수 있다.
- [0059] 본 발명은 도면에 도시된 실시예를 참고로 하여 설명되었으나 이는 예시적인 것에 불과하며, 당해 기술이 속하는 분야에서 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 따라서 본 발명의 진정한 기술적 보호범위는 아래의 특허청구범위의 기술적 사상에 의하여 정해져야 할 것이다.

부호의 설명

- [0060] 100: 교사 모델의 사전 지식을 이용한 지식 증류 시스템
- 110: 이미지 생성부
- 111: 제1 이미지 생성부
- 112: 제2 이미지 생성부
- 120: 지식 증류부

도면

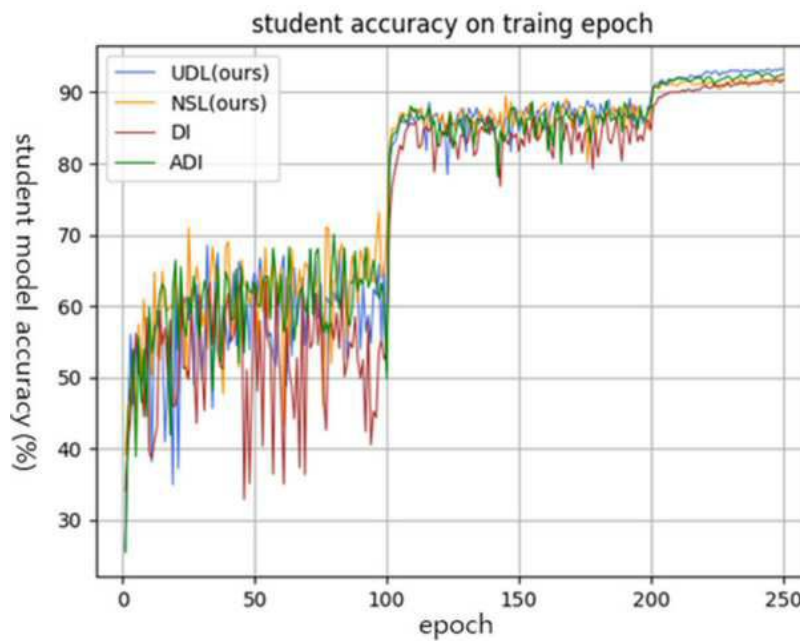
도면1



도면2

DFKD method	ResNet-18 학생 모델 정확도 (↑) (ResNet-34 교사 모델 정확도: 95.32%)
DI	91.69%
ADI	92.74%
NSL(ours)	92.19%
UDL(ours)	93.23%

도면3



도면4

DFKD 방법 클래스	DI	ADI	NSL (Ours)	UDL (Ours)
자동차	 	 	 	 
고양이	 	 	 	 
개구리	 	 	 	