

지식 증류를 이용한 DA-FSL 모델의 학습 수렴 곡선 및 학습 시간 개선 방법

임혜연*¹, 이준목*², 강대성**

A Method for Improving Learning Convergence Curve and Learning Time of DA-FSL Model using Knowledge Distillation

Hye-Youn Lim*¹, Jun-Mock Lee*², and Dae-Seong Kang**

이 논문은 2017 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No.2017R1D1A1B04030870)

요 약

많은 데이터를 확보한 분야에서 머신러닝 모델을 잘 학습시키는 것은 용이하다. 하지만 데이터가 부족한 분야에서 모델을 잘 학습시키는 문제를 해결하기는 쉽지 않다. 이러한 문제를 해결하기 위해 최근 전이학습을 통한 지식 증류가 주목받고 있다. 지식 증류는 앙상블 기법을 통해 학습된 다량의 큰 네트워크로부터 작은 하나의 네트워크에 지식을 전달하는 방법론을 의미한다. 본 논문에서는 DA-FSL 모델에 지식 증류 기법을 적용하여 학습 정확도뿐만 아니라 학습에 소요되는 시간을 개선하는 방법을 제안한다. 세 가지의 지식 증류 방법을 사용해서 DA-FSL 모델에 적용하고 실험을 통해 초기 목표 학습 정확도에 수렴하는 곡선을 평가하였다. 제안된 방법을 통해 지식 증류를 적용한 DA-FSL 모델의 학습 소요 시간과 지식 증류를 적용하지 않은 DA-FSL 모델의 학습에 소요되는 시간을 비교한 결과 최소 3%에서 최대 35%까지 감소되었음을 실험을 통해 확인하였다.

Abstract

It is easy to train a machine learning model well in a field with a lot of data. However, it is not easy to train the learning model well where learning data is deficient. In order to solve this problem, knowledge distillation through transfer learning has recently attracted attention. Knowledge distillation refers to a methodology for transferring knowledge from a large number of large networks learned through ensemble techniques to one small network. In this paper, we propose a method to improve the learning accuracy as well as the learning time by applying the knowledge distillation technique to the DA-FSL model. Three knowledge distillation methods were used to apply it to the DA-FSL model, and through experiments, the curve converging to the initial target learning accuracy was evaluated. We compared the learning time required for model applying knowledge distillation through the proposed method and normal DA-FSL model without knowledge distillation. As a result, it was confirmed through an experiment that it was reduced from a minimum of 3% to a maximum of 35%.

Keywords

few shot learning, knowledge distillation, transfer learning, weight initialization, model optimization

* 동아대학교 전자공학과

- ORCID¹ : <https://orcid.org/0000-0002-5189-461X>

- ORCID² : <https://orcid.org/0000-0002-2454-0672>

** 동아대학교 전자공학과 교수(교신저자)

- ORCID: <https://orcid.org/0000-0003-0186-2430>

· Received: Sep. 04, 2020, Revised: Oct. 14, 2020, Accepted: Oct. 17, 2020

· Corresponding Author: Dae-Seong Kang

Dept. of Dong-A University, 37 NaKdong-Daero 550, beon-gil saha-gu, Busan, Korea,

Tel.: +82-51-200-7710, Email: dskang@dau.ac.kr

I. 서 론

2017년 데미스 하사비스가 알파고를 업그레이드 한 버전인 알파고 제로를 소개하며 선행 학습한 사실에서 인지적인 세부사항을 추상화하고 여기에서 개념적인 지식을 획득하는 것이 전이학습을 할 수 있는 핵심[1]이라고 말하였다. 알파고 제로는 수많은 강화학습[2]을 통해 바둑을 스스로 터득했다고 널리 알려져 있다. 하지만 모든 분야에서 강화학습 모델을 적용하는 것은 인공지능망이 모든 문제에 대해 인간의 인지능력까지 모방할 수 없기 때문에 그 한계점이 존재한다. 많은 데이터를 확보한 분야에서 머신러닝 모델을 잘 학습시키는 것은 용이하나 데이터가 부족한 분야에서는 그렇지 않다. 특히 인간과 기계의 인지 차이로 인해 강화 학습마저 적용하기 어려운 분야라면 더욱더 그렇다. 이러한 문제를 해결하기 위해 주목받는 방법이 바로 전이학습을 통한 지식 증류(Knowledge Distillation)[3]이다.

지식 증류는 앙상블(Ensemble) 기법을 통해 학습된 다량의 큰 네트워크로부터 작은 하나의 네트워크에 지식을 전달할 수 있는 방법론을 의미한다. VGGNet[4], ResNet[5] 등 다수의 큰 모델에서 출력값은 일반적으로 특정 레이블에 대한 하나의 확률값만을 나타내지만, 지식 증류에서는 이를 확률값들의 분포 형태로 변형함으로써, 목표 모델 학습 시 Loss와 기 학습된 모델의 Loss를 동시에 반영하는 형태를 통해 학습에 소요되는 시간과 수렴 정확도를 최적화한다.

본 논문에서는 제한된 학습데이터만을 보유한 상황에서 딥러닝 모델의 정확도를 개선할 수 있는 DA-FSL(Data Augmentation based Few-Shot Learning)[6] 모델에 지식 증류 기법을 적용하여 정확도뿐만 아니라 전반적인 데이터 증강에 소요되는 시간 또한 개선하는 방법을 제안한다. 제안하는 방법을 통해 많은 양의 학습 데이터의 성인 남성의 얼굴 이미지와 소량의 데이터의 성인 여성의 얼굴 이미지로 구성된 학습 데이터 상황에서 일반적인 DA-FSL 모델과 제안하는 방법을 통해 목표 학습 정확도 수렴 곡선과 학습에 걸리는 시간을 비교하는 실험을 수행하여 개선 사항을 평가한다.

II. 관련 연구

목표 모델에 지식 증류를 적용하기 위해 퓨샷 학습(Few-Shot Learning)[7] 모델과 전이 학습에 대한 연구가 필요하다.

2.1 퓨샷 학습

일반적인 딥러닝 모델은 데이터양과 네트워크의 깊이에 비례하여 그 성능이 향상된다. 현재 딥러닝 모델의 데이터 분류 문제 해결 능력은 사람의 인지 능력 수준에 이르렀다. 하지만 모델을 학습할 데이터가 제한되는 경우 특정 클래스에 대한 학습 데이터 불균형 현상이 발생하고 이는 모델 정확도를 저하시키는 큰 요인이다. 퓨샷 학습은 기계가 인간처럼 소량의 데이터만으로 놀라운 학습 능력을 갖출 수 있을까 하는 의문에서 시작된 연구이다. 이러한 적은 양의 학습 데이터만으로도 새로운 데이터까지 잘 분류할 수 있도록 모델을 일반화시키는 방법은 크게 3가지 정도로 압축이 가능하다.

2.1.1 거리 기반 학습

범주별 훈련 데이터의 수가 적은 퓨샷 학습 태스크에서는 일반적인 딥러닝 모델처럼 분류기의 가중치를 훈련하는 방식은 적합하지 않다. 그림 1에서 보듯이 퓨샷 학습 태스크를 구성하는 $N \times K$ 개의 데이터에만 지나치게 적용하는 과적합(Overfitting)[8] 문제가 발생할 수 있기 때문이다.

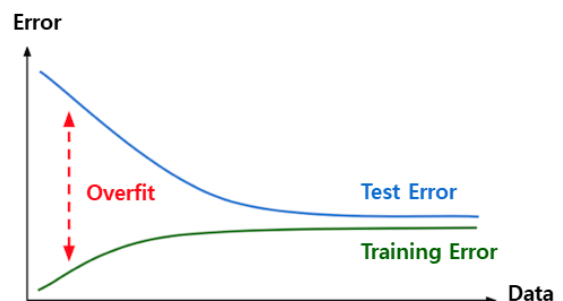


그림 1. 학습 데이터 부족에 따른 과적합 그래프
Fig. 1. Over-fitting graph due to lack of learning data

이러한 문제를 해결하기 위해 그림 2처럼 서포트 데이터와 쿼리 데이터 사이의 유사도 거리를 측정하는 방식을 활용한다. 모델은 주어진 서포트 데이터를 특징 공간(Feature Space)에 나타내고, 특징 공간에서 쿼리 데이터의 범주는 유클리드 거리(Euclidean distance)가 가장 가까운 서포트 데이터로 예측한다. 모델은 두 데이터의 범주가 같으면 거리를 더 가깝게 두고, 다르면 거리를 더 멀게 만드는 방법으로 학습한다.

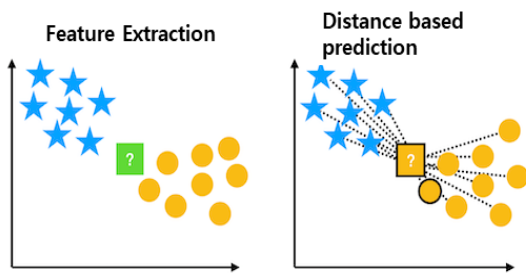


그림 2. 거리 유사성 기반 쿼리 데이터 클래스 예측
Fig. 2. Prediction of query data classes based on distance similarity

2.1.2 그래프 기반 신경망

소량의 데이터만으로 분류 성능을 극대화하고자 데이터 사이의 복잡한 관계 정보를 학습에 활용하는 추세이다. 일반적인 인공 신경망은 입력값을 벡터나 행렬 형태로 활용하는 반면 그래프 기반 신경망[9]은 밀집 그래프(Dense graph) 구조를 활용한다. 그래프에서 노드는 데이터를 노드와 노드를 잇는 Edge는 데이터 사이의 관계 정보를 나타내며 밀집

그래프는 모든 노드가 서로 완전히 연결된 것을 가리킨다. 그래프 기반 신경망은 이러한 그래프 구조와 각 노드에 해당하는 데이터의 특징 벡터를 입력으로 한다. 각 노드는 해당하는 데이터의 특징 벡터로 초기화되며, 특정 노드 V의 인접 노드에 각 노드의 유사 거리를 곱하여 모든 값을 합이 가중 평균이다. 이 가중 평균과 V를 합쳐 새로운 벡터 V를 얻을 수 있다. 이를 차례대로 반복하고 가장 마지막에 쿼리 노드 벡터의 업데이트를 완료하면 N개의 범주와 완전히 연결된 Fully Connected Layer를 통해 쿼리 데이터의 범주를 예측할 수 있다. 그림 3은 그래프 기반 신경망의 구조를 나타낸 것이다.

2.1.3 데이터 증강 기반

딥러닝 모델을 학습시키는 데 필요한 데이터를 확보하는 기법 중 하나인 데이터 증강 기법은 소량의 학습 데이터에 인위적인 변화를 가해 새로운 학습 데이터를 대량 확보하는 방법론을 의미한다. 현재에는 이미지 데이터를 상하좌우로 뒤집거나 자르는 방식으로 새로운 이미지 데이터를 확보하는 데이터 리샘플링 방법이 주를 이루고 있다. 데이터 증강이 실제로 모델 학습에 효과적이라는 사실은 널리 알려져 있다. 소량 학습데이터에 데이터 리샘플링 혹은 적대적 생성 신경망[10]과 같은 딥러닝 모델을 이용해서 유사 학습 데이터를 늘려 부족한 학습데이터 상황을 해소하는 방식이다. 이 데이터 증강 기반 방법은 그래프 기반 신경망 방법과 결합하여 좋은 성과를 보여주고 있다.

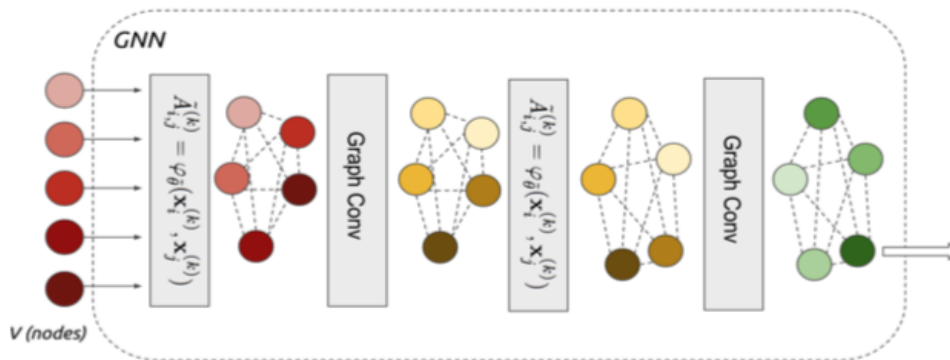


그림 3. 그래프 기반 신경망의 구조
Fig. 3. Architecture of graph based neural network

2.2 DA-FSL

DA-FSL은 학습데이터에서 클래스 불균형이 있는 상황에서 딥러닝 모델의 정확도를 개선하는 방법이다. 데이터셋 사이에서 데이터 불균형을 판단을 통해 D_{enough} 와 D_{lack} 으로 데이터셋을 나누고 D_{enough} 를 통해 객체 검출을 수행하는 BDRP(Base Data Region Proposal)[6] 모듈을 학습한다. BDRP 모듈에서는 테스트 데이터셋에서 데이터 증강을 위한 객체를 검출한다. DA-FSL은 클래스 불균형이 존재하는 소량 데이터인 Dataset_S 와 객체 검출 데이터 Dataset_B 를 통해 적대적 생성 신경망으로 학습 데이터를 증강하여 모델의 정확도를 개선하는 방법론이다.

2.3 전이학습

전이학습[11]은 학습 데이터가 부족한 분야의 모델 구축을 위해 데이터가 풍부한 분야에서 훈련된 모델을 재사용하는 학습 기법이다. 이미 학습된 모델을 사용해서 목표로 하는 모델의 문제를 해결하는 방법을 전이 학습을 통한 지식 증류라고 한다. 복잡한 모델일수록 학습에 소요되는 시간과 자원의 양은 방대하다. 이는 전이학습을 통해 일부분 해결 가능하다. 즉, 이미 잘 학습된 모델이 존재한다면 해당 모델과 유사한 문제를 해결할 때 잘 학습된 모델의 정보를 일부 가져와서 학습 시간과 정확도

를 높일 수 있는 방법이다. 그림 4는 전이학습을 이용할 때의 생략 가능한 층을 나타냈다.

III. 제안하는 방법

본 항목에서 제안하는 바는 그림 5의 DA-FSL 모델에 지식 증류 기법을 적용하여 모델 학습 수렴 속도와 학습에 소요되는 시간을 개선하는 방법이다.

목표 모델에 지식 증류를 적용하기 위해서 세 가지의 방법을 통해 지식 증류를 수행한다.

첫 번째 방법은 합성곱 계층에서 특징 추출기를 고정하여 모델을 전이한다. 미리 구성된 합성곱 신경망을 사용하여 심층 신경망의 마지막 Fully Connected Layer를 제거한 다음에 나머지 부분을 새로운 데이터셋의 고정 특징 추출기로 사용한다.

두 번째 방법은 새로운 데이터셋의 합성곱 계층 상단에서 분류기를 교체하고 재학습하는 것뿐만 아니라 역전파(Back propagation)[12]를 계속함으로써 Pre-defined network의 가중치를 미세 조정(Fine-tune)하는 것이다. 여기에서 합성곱 계층의 모든 계층을 미세 조정하거나 이전 계층 중 일부 가중치만을 고정하고 네트워크의 일부 상위 계층만 미세 조정할 수 있다. 이때 과적합 문제가 발생할 수 있기 때문에 전체 계층에 대해서는 미세 조정을 하지 않는다.

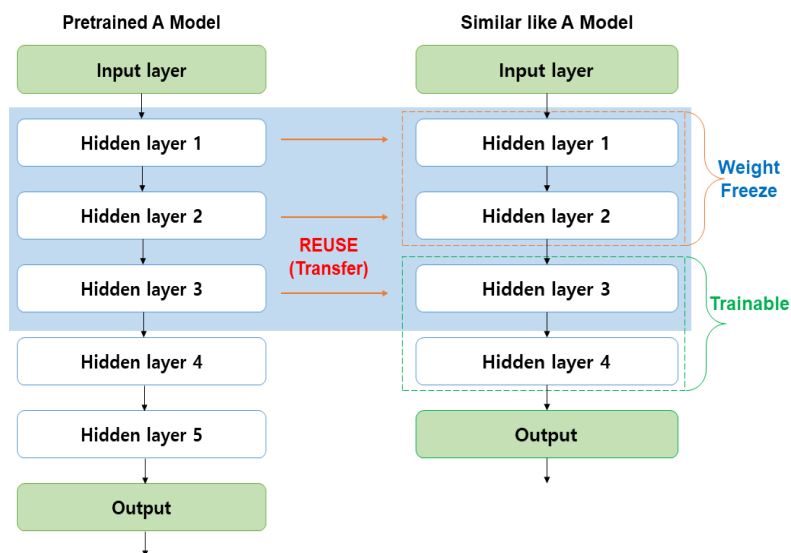


그림 4. 전이학습 시 생략 가능한 층의 예
Fig. 4. Examples of omissible layers in transfer learning

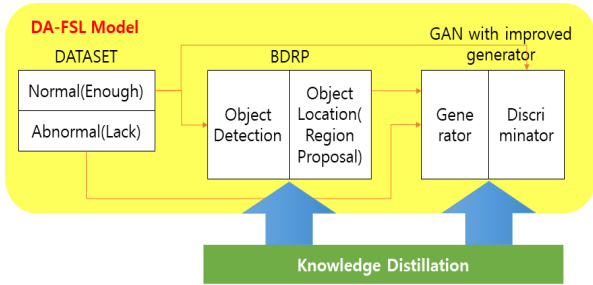


그림 5. 제안하는 방법의 적용 모식도

Fig. 5. Application schematic diagram of the proposed method

마지막은 사전 학습된 모델을 사용하는 방법이다. 심층 신경망 모델을 학습하기 위해서는 다중 GPU를 이용해도 최소 2~3주의 학습 시간이 필요하므로 깊은 네트워크의 체크포인트를 이용하는 방법이 가장 일반적인 지식 증류라 할 수 있다.

새로운 데이터셋에서 수행해야 하는 지식 증류에서 고려해야 하는 방법은 새로운 데이터셋과 원본 데이터셋과의 유사도이다.

식 (1)은 데이터 벡터를 n 차원의 벡터 좌표 (x, y) 로 표기할 때 데이터 벡터 사이의 유사도를 측정하는 방법이다.

$$Distance(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (1)$$

식 (1)을 통해 구한 데이터 유사도를 비교하여 새로운 데이터셋과 기존 데이터셋과의 유사도가 낮다면 합성곱 신경망을 과도하게 미세조정하는 것은 좋지 않다. 이 경우에는 모델의 은닉층에 존재하는 합성곱 신경망에는 지식 증류를 수행하지 않고 Fully Connected Layer의 Softmax나 Sigmoid와 같은 선형 분류기만 학습하는 것이 좋다. 새로운 데이터셋과 기존 데이터셋의 유사도가 높은 경우에만 전체 네트워크에 대해서 지식 증류를 수행하는 것이 좋다.

이와는 별개로 지식 증류를 수행할 때 유의해야 할 사항은 사전 학습 네트워크를 사용하는 경우 새로운 데이터셋에서 사용할 수 있는 구조 측면에서 제한이 존재할 수 있다. 즉 계층의 뉴런 크기 등이 일치해야 지식 증류가 가능하며, 사전 학습된 네트워크의 가중치가 상대적으로 좋다고 가정하고 학습

을 수행하기 때문에 학습 곡선이 너무 많이 왜곡되지 않도록 학습률(Learning Rate)을 조절하는 것에 유의해야 한다. 표 1은 본 논문에서 사용된 DA-FSL의 BDRP와 데이터 증강 모델에 지식 증류를 적용하기 위해서 사용된 사전 학습된 모델을 표시하였다.

표 1. 지식 증류를 위한 사전 학습된 모델 목록
Table 1. List of pre-trained models for knowledge distillation

Base Network	Pretrained Network
Region Proposal Network	ImageNet
Faster R-CNN Classifier	ResNET-50 V1b
Generator	PG-GAN
Discriminator	DC-GAN

IV. 실험 방법 및 결과

실험에서는 성인 남성과 여성의 얼굴 데이터가 제한된 상황에서 정확도를 개선하는 모델의 학습 수렴 곡선 및 소요 시간을 개선하는 방법을 검증하기 위해 DA-FSL 모델의 일부를 이용한다. 지식 증류를 위해 BDRP 모듈과 데이터 증강에 사용된 데이터셋을 그대로 사용한다. BDRP 모듈과 데이터 증강 모델 학습에 이용한 데이터셋은 Flickr-Face-HQ[13]이다. 이 중 남성 10,000장, 여성 10장의 얼굴 이미지를 임의로 추출하여 이용하였다. 모델의 검증을 위한 테스트 데이터셋은 UMDFaces Dataset[14] 1,000장을 이용했다. 모델의 학습 수렴 커브 정확도 평가를 위해서는 식 (2)의 mAP 로 평가한다.

$$mAP = \frac{1}{11} \sum_{r \in \{0, 0.1, \dots, 1\}} p_{\cap p}(r) \quad (2)$$

$$p_{\cap p}(r) = \max_{\tilde{r}: \tilde{r} \geq r} p(\tilde{r})$$

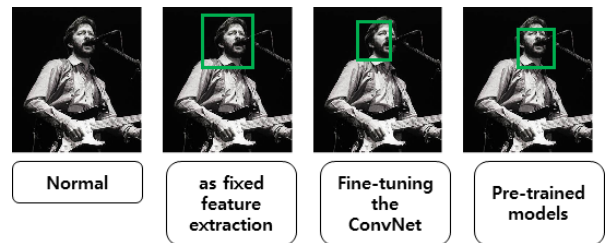


그림 6. 지식 증류 적용에 따른 BDRP의 Object Location
Fig. 6. Results of BDRP object location according to applying knowledge distillation

지식 증류를 통해 BDRP 모듈을 100회 학습한 결과 Object Location 결과는 그림 6에 나타내었다.

표 2는 지식 증류를 적용하기 전과 후의 Object Location 정확도를 나타낸 결과이다. 그림 7은 지식 증류의 적용 전과 후의 BDRP 모듈에서 mAP 0.5 IoU에 수렴하기까지의 학습 곡선을 나타내었다.

표 2. Epoch=100으로 학습했을 때의 mAP 결과
Table 2. mAP results when learning with Epoch=100

Methods	mAP(%)
Normal DA-FSL(Vanilla)	22.38
with ConvNet with as fixed feature extractor	32.32%
with Fine-tuning the ConvNet	31.87%
with Pre-trained models	43.22%

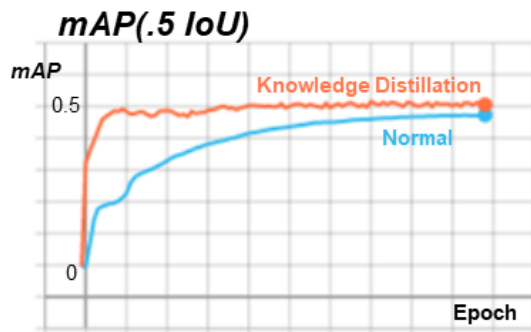


그림 7. IoU 0.5에 수렴하기까지의 학습 곡선
Fig. 7. Learning Convergence curve at IoU 0.5

적대적 생성 신경망을 이용하는 데이터 증강 모델에서는 학습의 수렴은 Generator와 Discriminator Loss의 변화로 평가 할 수 있다. 표 3은 1,000회 Epoch에서 지식 증류를 적용하였을 때와 하지 않았을 때의 Generator와 Discriminator의 Cross Entropy Loss를 나타내었다.

표 3. Epoch=1,000으로 학습했을 때의 Cross Entropy Loss 결과
Table 3. Loss results of generator and discriminator when learning with epoch=1,000

Methods	Cross entropy loss
Normal DA-FSL(Vanilla)	34.32
with ConvNet with as fixed feature extractor	38.32
with Fine-tuning the ConvNet	13.53
with Pre-trained models	4.32

그림 8은 데이터 증강 모델의 지식 증류를 적용하기 전 후의 Cross Entropy Loss 학습 곡선을 나타내었다.

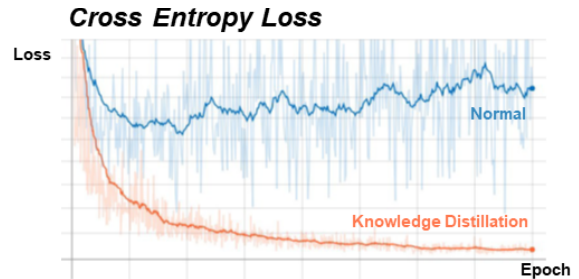


그림 8. 지식 증류 적용 전 후 증강 모델의 Cross Entropy 학습 곡선

Fig. 8. Cross entropy learning curve of data augmentation model applying before and after knowledge distillation

표 4. 지식 증류를 이용할 때 학습에 소요되는 시간
Table 4. Time spent learning when using knowledge distillation

Module	Original learning time	Knowledge distillate time
Faster R-CNN	3d 1h 15m	1d 23h 53m
Generator	4d 12h 48m	4d 5h 15m
Discriminator	2d 18h 22m	2d 16h 4m

두 결과에서 볼 수 있듯이 지식 증류를 통해서 DA-FSL 세부 모듈의 목표 정확도에 수렴하기 위한 학습 곡선이 앞당겨지는 것을 알 수 있다. 표 4는 지식 증류를 적용하였을 때와 적용하지 않았을 때 학습에 소요되는 시간을 나타내었다.

V. 결 론

본 논문에서는 지식 증류를 이용하여 모델의 학습 수렴 곡선과 학습 시간을 개선하는 방법을 제안하였다. 제안된 방법을 DA-FSL 모델에 적용하여 실험하였고, 적은 양의 수행 횟수를 가지고도 모델의 학습에 소요되는 시간을 최소 3%에서 최대 35%까지 개선할 수 있었다. 제안한 지식 증류 방법은 실험에서처럼 DA-FSL 모델에만 적용 가능한 것뿐만 아니라 목표 모델에 어떠한 적용 방법을 제시하는가에 따라 지식 증류가 가능하므로 다양한 분야의 모델에 적용이 가능할 것으로 판단한다. 다만 이

를 위해서 목표와 원 데이터셋에 대한 분석이 필수적이고 모델의 구조 측면에서도 많은 수정이 필요하다[15].

본 논문에서는 국한된 모델에 대해 지식 증류를 적용하는 연구를 진행하였으나 추후 지식 증류를 범용적인 모델에 적용하기 위해 목표 모델의 계층이나 구조를 자동 분석해서 지식 증류를 적용할 수 있는 것으로 확장하는 것이 앞으로의 연구 목표이다.

References

[1] Ford, M., "Architects of Intelligence : The truth about AI from the people building it", Packt Publishing Ltd, pp. 174-180, 2018.

[2] Sutton, R. S. and Barto, A. G., "Reinforcement learning : An Introduction", MIT press, pp. 1-3, 2018.

[3] Chen, G., Choi, W. and Chandraker, M., "Learning efficient object detection models with knowledge distillation", In Advances in Neural Information Processing Systems, pp. 742-751, Jan 2018.

[4] Simonyan, K. and Zisserman, A., "Very deep convolutional networks for large-scale image recognition", arXiv preprint arXiv:1409.1556, Apr 2015.

[5] He, K., Zhang, X., and Sun, J., "Deep residual learning for image recognition", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, Jun 2016.

[6] Lee, J. M., Kim, H. Y. and Kang, D. S., "Improved method for learning data imbalance in gender classification model using DA-FSL", Proceedings of The International Conference on Future Technology, pp. 204-201, Jan 2020.

[7] Ravi, S.. and Larochello, H., "Optimization as a model for few-shot learning", International Conference on Learning Representations, 2017.

[8] Hawkins, D. M., "The Problem of over-fitting", Journal of chemical information and computer sciences, Vol. 44, No. 1, pp. 1-12, Dec 2003.

[9] Yasunaga, M., Zhang, R. and Radev, D., "Graph-based neural multi-document summarization", Proceedings of the 21st Conference on Computational Natural Language Learning, pp. 452-462, Aug 2017.

[10] Goodfellow, I., Pouget-Abadie, J. and Mirza, M., "Generative Adversarial nets", Advances in neural information processing system, pp. 2672-2680, Jun 2014.

[11] Pan, S. J. and Yang, Q., "A survey on transfer learning", IEEE Transactions on knowledge and data engineering, Vol. 22, No. 10, pp. 1345-1359, Oct 2009.

[12] LeCun, Y., Boser, B., and Jackel, L. D., "Back Propagation applied to handwritten zip code recognition", Neural computation, Vol. 1, No. 4, pp. 541-551, Dec 1989.

[13] Flickr-Face-HQ-Dataset, <https://github.com/NVlabs/ffhq-dataset>, [Accessed : May. 7. 2020]

[14] UMDFaces Dataset, <http://umdfaces.io/> [Accessed : May. 13. 2020]

[15] Kim, S. H. and Kim, Y. G., "A Study on Light Weight Authentication Method of Distributed Cluster-based IoT Devices", The Journal of IIBC, Vol. 19, No. 2, pp. 103-109, Apr 2019.

저자소개

임혜연 (Hye-Youn Lim)



2013년 2월 : 동아대학교
전자공학과 (공학박사)
2013년 3월 ~ 2019년 2월 :
동아대학교 전자공학과 조교수
2019년 3월 ~ 현재 : 동아대학교
전자공학과 시간강사
관심분야 : 영상처리, 패턴인식

이 준 목 (Jun-Mock Lee)



2008년 2월 : 동아대학교
컴퓨터공학과 (공학사)
2020년 8월 : 동아대학교
전자공학과 (공학석사)
2020년 9월 ~ 현재 : 동아대학교
전자공학과 박사과정
관심분야 : 영상처리, 인공지능

강 대 성 (Dae-Seong Kang)



1994년 5월 : Texas A&M 대학교
전자공학과(공학박사)
1995년 ~ 현재 : 동아대학교
전자공학과 교수
관심분야 : 영상처리, 패턴인식