파라미터 재정의를 통한 분류 모델의 경량화와 전이 학습을 통한 조밀 예측 모델의 성능 향상에 관한 연구

김영모, 이정우
서울대학교

ymkim@cml.snu.ac.kr, junglee@cml.snu.ac.kr

Reparameterization method for classification model compression and transfer learning for improvement of dense prediction model

Kim Yeong Mo, Lee Jung Woo*
Seoul National Univ.

요 약

본 논문은 기존의 ResNet 등 multi-branch 구조를 가지는 CNN 모델을 single-branch 구조로 단순화하여 분류(classification) 모델의 성능은 유지하면서 총 파라미터 수를 줄여 모델을 경량화하는 reparameterization 기법을 구현하고, 그 reparameterize 된 모델 가중치를 물체 검출(object detection) 및 분할(segmentation) 작업 모델에 전이 학습시켜 해당 모델들이 동 모델을 처음부터 학습시켰을 때에 비하여 더 높은 성능을 갖도록 하였다.

Ⅰ. 서 론

최근 인공지능 기술의 급격한 발전에 따라, 연구뿐만이 아니라 각종 분야에 있어서의 인공지능 기술의 실제 적용에 대한 수요가 증가하고 있다. 그러나 기계학습 기법의 실제 적용에 있어서 가장 큰 문제가 되는 부분이 연구 환경과 사용자 환경에서의 하드웨어 성능 상의 괴리로, 연산용 GPU 를 1 대 이상 준비하고 실험을 수행하는 연구 환경과는 다르게, 실사용자들은 모바일 환경 등 적은 연산량으로도 기계학습을 수행할 수 있는 방법을 필요로 한다.

이러한 문제를 해결하기 위해서 네트워크 경량화 및 최소 파라미터 하에서의 최대 성능을 내기 위한 연구들이 이루어지고 있다. Quantization 과 Pruning 을 통하여 파라미터 자체를 축소하여 정확도 저하를 최소화하면서 네트워크를 경량화 시키거나, Network Architecture Search(NAS) 방법을 사용하여 네트워크 크기에 제한을 두고 그 안에서 최적의 모델을 탐색하는 방법을 사용한다. 그러나 이러한 방법들은 필연적으로 정확도와 네트워크 크기 사이의 trade-off 가 무시할 수 없는 수준으로 발생하게 된다.

본 논문에서는 복잡한 구조에서 학습된 가중치 정보를 reparameterize 하여 단순한 구조에서 복잡한 구조와 비슷한 수준의 성능을 내도록 한다. 그 후에 해당 가중치를 다른 모델에 전이시켜 해당 구조로는 얻기 힘든 수준의 성능을 갖는 모델을 개발하는 연구를 수행한다.

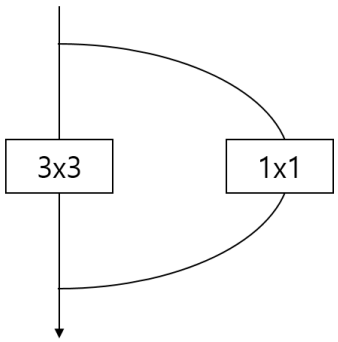
Ⅱ. 본론

RepVGG[1]에서는 ResNet 구조에서 학습을 수행하고, VGGNet 구조에서 inference 를 수행함으로써 정확도 성능은 ResNet 수준을 유지하면서 inference 과정에서 드는 연산 비용은 VGGNet 수준으로 격감시킨다. 구체적으로, ResNet 구조는 3x3 크기의 컨볼루션 레이어에 1x1 크기의 컨볼루션 레이어가 병렬적으로 연결되어 있는 형태이다. 이 네트워트의 특정 레이어에 들어오는 입력은 3x3 커널과 1x1 커널에서 각각 컨볼루션 계산을 실행하고, 두 결과값을 더하여 최종 결과값을 계산한다. 그런데 이 과정에서 컨볼루션과 덧셈 모두 선형적 계산이기 때문에, 컨볼루션 레이어와 배치 정규화 레이어의 값들이 확정된 상황에서는 해당 연산들의 순서를 바꿀 수 있다.

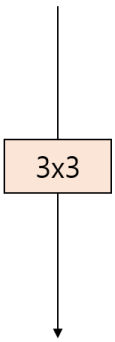
$$f_{i+1} = f_i * w_{3\times3} + f_i * w_{1\times1}$$

$$f_{i+1} = f_i * (w_{3\times3} + w_{1\times1})$$

위 수식에서 f_i 는 i 번째 레이어에서의 입력이고, $w_{3\times3}$와 $w_{1\times1}$는 각각 3x3, 1x1 컨볼루션 레이어에서의 가중치 값이다. 이 때 가중치끼리의 합연산을 먼저 계산해줌으로서, 2 개의 컨볼루션 레이어를 하나로 합칠 수 있다. 1x1 컨볼루션 레이어의 값은 3x3 레이어의 중앙에 더해주고, 최종적으로 3x3 컨볼루션 레이어 하나만이 남게 된다. 모든 residual connection 에 이 작업을 수행한 후에 ResNet 은 VGGNet 과 같은 형태의 네트워크로 단순화된다. 또한 이때 학습된 가중치는 ResNet 에서 학습한 가중치와 같은 값을 가지므로



a) RepVGG training



b) RepVGG inference

정확도 성능은 ResNet 과 비슷한 수준으로 높아지게 된다.

위 방식으로 얻은 가중치는 분류 작업뿐만 아니라 CNN 구조를 활용하는 모델이라면 다른 작업에도 마찬가지로 활용 가능하다. 본 논문에서는 컴퓨터 비전 연구에서 대표적인 주제들인 물체 검출과 분할 모델에 대하여 이 가중치들을 도입하고, 실제 성능 향상이 있는지를 확인하였다.

물체 검출 모델로는 ResNet-50 을 베이스로 하는 RetinaNet[3], 분할 모델로는 마찬가지로 ResNet-50 을 베이스로 하는 DeepLabV3+[4] 모델을 사용하였다. 각 모델의 베이스인 ResNet 부분은 앞에서 만든 분류 모델의 마지막 전연결 레이어만 제거한것과 같기 때문에 해당 부분만 제거하여 그대로 불러온다. 이 때 가중치를 가지지 않는 부분은 RetinaNet 에서는 Feature Pyramid network(FPN), DeepLabV3+ 에서는 decoder 부분이기 때문에 해당 부분을 추가로 학습시켜줄 필요가 있다. 따라서 ResNet 부분의 가중치 값을 고정시킨 상태로, 각 네트워크 전체를 fine-tuning 해주게 된다. 데이터셋은 각각 Microsoft COCO 데이터셋과 Cityscapes 데이터셋을 사용하였고, RetinaNet 에서는 12 epochs, DeepLabV3+ 에서는 40,000 steps 만큼 학습시켰다.

Re-param	RetinaNet	DeepLabV3+
	mAP	mIoU
None	36.2	76.63
RepVGG	**36.4**	**76.94**

표 1. 물체 검출과 분할 모델에 대한 전이 학습 결과

전이 학습의 결과는 표 1 와 같이 나온다. 이때 Re-parameterization 기법을 사용하지 않은 결과값은 오픈소스인 mmdetection, mmsegmentation 에서 제공하는 pretrained model 을 사용한 값이다. 양쪽 모델 모두 전이 학습을 시행하지 않았을 때에 비하여 성능이 향상되었다. 수치가 높아질수록 성능을 더 올리기 힘들어지는 mAP 와 mIoU 의 지표 특성을 고려했을 때, Retinanet 에서의 성능 향상은 미미한 수준이나, DeepLabV3+ 에서의 성능 향상은 주목할 만한 성능 향상이라 할 수 있다.

Ⅲ. 결론

본 논문에서는 네트워크의 경량화를 위해 re-parameterization 을 사용하는 기법인 RepVGG 에 대하여 소개하고, 해당 기법에 전이 학습을 적용하여 분류가 아닌 타 작업을 위한 모델에서도 우수한 성능을 낼 수 있음을 확인하였다. Quantization 이나 pruning 같은 방법 또한 네트워크 경량화에 있어 효율적인 방법이지만, cross-platform 등으로 상황에 따라 하드웨어를 옮겨가며 작업을 수행하는 상황에서는 연산량 감소에 따른 trade-off 에 민감해지게 된다. Re-parameterization 방법은 이러한 문제를 해결할 수 있는 하나의 방법이 될 수 있을 것이다.

ACKNOWLEDGMENT

This work is in part supported by National Research Foundation of Korea (NRF, 2021R1A2C2014504(20%)), Institute of Information & communications Technology Planning & Evaluation (IITP, 2021-0-00106(60%)) grant funded by the Ministry of Science and ICT (MSIT), Center for Applied Research in Artificial Intelligence(CARAI, UD190031RD(20%)) grant Funded by Defense Acquisition Program Administration(DAPA), Agency for Defense Development(ADD), INMAC, and BK21-plus.

참 고 문 헌

[1] Xiaohan D. et al.," RepVGG: Making VGG-style ConvNets Great Again," CVPR 2021

[2] Mu H., et al., "Online Convolutional Re-parameterization," CVPR 2022

[3] Tsung-Yi L., et al., "Focal Loss for Dense Object Detection," ICCV 2017

[4] Liang-Chieh C., et al., "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation," ECCV 2018