

해밍부호 기반 이진인공신경망 학습

이선용, 노종선

서울대학교 전기정보공학부, 뉴미디어통신공동연구소

sonyongyi@snu.ac.kr, jsno@snu.ac.kr

Binary neural network learning strategy via hamming distance

SeonYong Lee, Jong-Seon No

INMC, Dept. of ECE, Seoul National Univ.

요 약

본 논문은 이진인공신경망 네트워크의 성능을 개선하기 위해서 새로운 학습 방법을 제시한다. 새로 제시된 방법을 통해 학습한 이진신공망 네트워크는 별다른 구조적 변화 없이 classification 층의 해밍부호 기반 classification 을 통해 좋은 학습성능을 보인다.

I. 서 론

이진 인공 신경망은 (Binay Neural Network, 이하 BNN) 2016 년 Itay hubara, Yoshua Bengio 에 의해 제안된 메모리 사용을 극단적으로 제한한 인공신경망 구조이다[1]. 인공신경망 네트워크는 Hinton, Bengio, Lecun 에 의해 인공지능 분야의 혁신을 이루었고, 이미지 인식 및 기타 분야에서 사람을 뛰어넘는 성능을 보여주고 있다. GPU 의 발전으로 빠른 다중연산처리가 가능해졌기 때문에 머신러닝 기술이 대두되었으나, 그 계산량과, 전력소모 등에 의해 실제 사용할 때 많은 단점이 따른다. 이에 인공신경망 네트워크를 경량화 하는 방법을 많은 사람들이 연구하고 있다. 그 중 이진인공신경망 네트워크는 기존에 제시된 모델들을 그대로 사용할 수 있어, 기존 모델들이 가진 특성을 반영할 수 있고, 학습 방법, 가중치등을 이용 할 수 있기 때문에 범용성이 높아 네트워크 경량화 분야에서 높은 가능성을 가지고 있다.

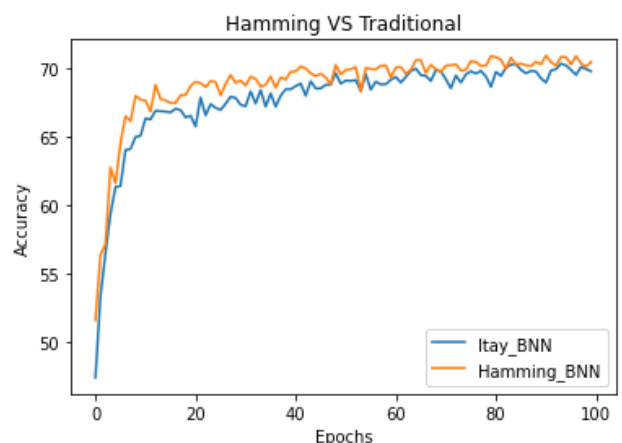
이진 인공 신경망은 모든 가중치를 이진화 하기 때문에 이에 알맞은 방식의 학습방법을 연구해야 한다는 의견이 존재한다. 그에 따라 Binary optimizer 라 불리는 방식이 나왔고[2], 그 외에도 다양하게 연구 되고 있다.

합성곱 인공신경망에서 분류층의 결과를 해밍부호로 나타내었을 때 별다른 조작 없이 adversarial attack 과 학습에서 좋은 결과를 나타냄이 알려져있다[3].

본 논문에서는 위의 결과들에서 착안해 이진 인공신경망의 학습단계에서 분류층을 해밍부호화하여 학습과정을 개선시킬 수 있음을 보였다.

II. 본론

본 논문에서는 BNN 의 분류층을 해밍부호화 하여 성능을 개선하는 방법을 실험했다. 실험데이터셋은 CIFAR-10 데이터셋을 이용했고, BNN 구조는 VGG-small 네트워크를 사용했다. 데이터셋의 정규화 이외에는 다른 전처리를 사용하지 않았다. Classification layer 의 해밍 부호는 (7,4) 해밍부호를 이용했고, 10 개의 classification 을 (7,4) 해밍부호를 통해 나타내 주었다.



III. 결론

본 논문에서는 해밍부호를 기반으로 분류층을 나타내는 방법을 BNN 에서 적용했고, 적용한 결과가 traditional BNN 보다 좋은 성능을 보여줌을 보였다. 추후의 연구로서 cifar-100, imagenet 등 다양한 분류를 요하는 데이터셋에 대해 RS 부호, LDPC 부호등 다양한 부호를 BNN 구조에 적용하는 연구가 수행될 예정이다

참 고 문 헌

- [1] Hubara I, Courbariaux M, Soudry D, El-Yaniv R, and Bengio Y. "Binarized Neural Networks," NeurIPS, pp.4107-4115, 2016.
- [2] Helwegen K, Widdicombe J, Geiger L, Liu Z, Cheng K-T, and Nusselder, R. "Latent Weights do not exist: Rethinking Binarized Neural Network Optimization," NeurIPS, 2019.
- [3] Gunjan V, Ananthram S. "Error Correcting Output Codes Improve Probability Estimation and Adversarial Robustness of Deep Neural Networks, " NeurIPS, 2019