SPRKD: Effective Knowledge Distillation for Deep Neural Networks via Saddle Region Approximation

Dewan, Aditya (School: The Woodlands School - Secondary)

Modern Deep Neural Networks (DNNs) are potent catalysts for scientific-industrial impact. To maximize performance and hypothesis space, however, DNNs involve excessive parameters – impeding strong DNN deployment in low-compute settings such as hospital equipment and energy infrastructure, elevating required resources and hindering impact. Predominant Knowledge Distillation (KD) methods favour replication - smaller students typically mimic teacher output logits – yet empirically yield low basic-task performance, hamper end convergence, inhibit generalization, and act merely as regularization, failing to distill useful teacher traits. This work develops a novel distillation framework to combat these limitations – Saddle Point Recruitment for Knowledge Distillation (SPRKD), a) reorienting KD from replication to employing teachers as optimization curvature and domain proxies and b) characterizing saddle points as strong-further-descent-potential regions through embedding and basin fractal properties. Via Hessian eigenvalue density, SPRKD identifies strong saddle regions for student re-exploration, employing 2-norm transformations, negative Hessian eigensteps, and Gaussian perturbations for teacher-boosted student descent. Testing on NHS-obtained malaria smear data revealed SPRKD outperformed current KD methods by 24.70% (p=6.3E-8) and attained identical accuracy to SGD-based methods (p=1.0) with faster, more complete, and more stable convergence. Further benchmark-dataset analysis demonstrated superior accuracy, wider minima convergence, and smaller Hessian metrics, denoting smoother descent, enhanced noise robustness, and better real-world performance. This work highlights strong SPRKD impact potential, enabling high-performance model deployment across widespread pursuits.

Awards Won:

Fourth Award of $500

National Security Agency Research Directorate : Second Place Award “Cybersecurity”