|  |  |  |
| --- | --- | --- |
|  | |ICAS2025| |  |

**Adaptive Learning Rate Optimization for Efficient Deep Neural Network Training**

Emily J. Parker1\*, Nathan R. Coleman2, Olivia M. Bennett1

1Department of Computer Science, Stanford University, USA
2Department of AI and Data Science, MIT, USA

**Abstract:** The efficiency and stability of training deep neural networks are significantly influenced by the choice of learning rate and its scheduling. Traditional optimizers like Stochastic Gradient Descent (SGD), Adam, and RMSprop offer different approaches to learning rate adaptation; however, they often require manual tuning or rely on global heuristics that may not generalize well across diverse datasets. In this paper, we introduce a novel adaptive optimization technique called Dynamic Thresh old Descent (DTD), which automatically adjusts the learning rate at each training step based on local gradient behavior and batch-level variance.
Unlike fixed schedules or momentum-based methods, DTD leverages real-time statistical analysis of gradient norms to determine whether to increase or decrease the learning rate. This dynamic adjustment reduces sensitivity to initial hyperparameter settings and improves convergence rates, particularly in early training stages. We evaluate DTD on three standard datasets: MNIST, Fashion-MNIST, and CIFAR-10, using convolutional neural networks of moderate depth. Experimental results show that DTD achieves faster convergence while maintaining comparable or better accuracy than existing optimizers. Our approach offers an intuitive mechanism to adapt learning rates based on the training environment’s evolving characteristics, making it suitable for both low-resource and large-scale scenarios. Furthermore, we demonstrate that DTD maintains stability even in the presence of noisy gradients, which is a common challenge in real-world data. Overall, the proposed optimizer represents a step toward more autonomous and context-aware training strategies in deep learning. This work lays the foundation for future explorations in adaptive training, especially in domains that demand robustness, such as healthcare, finance, and autonomous systems.

**Keywords:** adaptive learning rate, deep learning, optimizer, gradient descent, convergence