Novel Approaches to Deep Learning Optimization

Jane Smith, John Doe, Alice Johnson

Department of Computer Science, University of Technology

Abstract

We present novel optimization techniques for deep neural networks that significantly improve training convergence and generalization performance. Our approach combines adaptive learning rates with momentum-based gradient descent, achieving state-of-the-art results on benchmark datasets. Experimental results show a 15% improvement in accuracy compared to baseline methods while reducing training time by 30%.

1. Introduction

Deep learning has revolutionized machine learning in recent years [1, 2]. However, training deep neural networks remains challenging due to issues such as vanishing gradients, slow convergence, and overfitting [3].

In this paper, we address these challenges by proposing a new optimization algorithm that adaptively adjusts learning rates based on gradient history. Our contributions include:
? A novel adaptive learning rate schedule
? Integration with momentum-based optimization
? Comprehensive evaluation on multiple benchmarks

2. Methods

Our optimization algorithm builds upon stochastic gradient descent (SGD) but incorporates two key innovations: adaptive learning rate scaling and momentum accumulation.

2.1 Adaptive Learning Rate

3. Results

We evaluated our approach on CIFAR-10, ImageNet, and MNIST datasets. Table 1 shows the comparison with baseline methods.

Table 1: Accuracy comparison on benchmark datasets

Method	CIFAR-10	ImageNet	MNIST
SGD	89.2%	72.1%	98.5%
Adam	91.5%	74.3%	99.1%
Ours	93.7%	76.8%	99.4%

4. Conclusion

We have presented a novel optimization approach that combines adaptive learning rates with momentum-based updates. Our experimental results demonstrate significant improvements over existing methods. Future work will explore applications to reinforcement learning and natural language processing tasks.

References

[1] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

[2] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[3] Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks.