



Contribution ID: 2

Type: Abstract for Research Paper

## Empirical Evaluation of Adaptive Optimization on the Generalization Performance of Convolutional Neural Networks

Recently, we have witnessed the rise of deep learning with deep neural networks garnering significant interest and popularity in a variety of fields of research due to their effectiveness in search for an optimal solution given a finite amount of data. However, the optimization of these networks has become more challenging as the neural networks become deeper and datasets growing larger. The choice of the algorithm to optimize a neural network is one of the most important steps in model design and training to obtain a model that will generalize well on new, previously unseen data. In machine learning, three main kinds of optimization methods exist. The first one is called batch or deterministic gradient methods that process all training examples simultaneously in a large batch. The second one is the stochastic or online methods that use only one example at a time. The third one is called minibatch, which is a blend of the two whereby during model training, only a part of training set at each epoch is used. In deep learning, minibatch optimization methods are mostly preferred for supervised and unsupervised task. First, they accelerate the training of neural networks and since the minibatches are selected randomly and are independent, an unbiased estimate of the expected gradient can be computed. This paper examines the minibatch-based adaptive algorithms on the generalization performance of convolutional neural networks (CNN) architecture that are extensively used in computer vision tasks. We give a comparative analysis on the behaviour of the minibatch optimization algorithms during model training on three large image datasets, namely, MNIST, Kaggle Flowers and Scene classification.

**Primary authors:** Mr WANJAU, Stephen Kahara (Murang'a University of Technology); Dr WAMBUGU, Geoffrey Mariga (Murang'a University of Technology); Dr OIRERE, Aaron Mogeni (Murang'a University of Technology)

**Track Classification:** Artificial Intelligence