

# On the Relative Impact of Optimizers on Convolutional Neural Networks with Varying Depth and Width for Image Classification

Authors: Eustace M. Dogo, Oluwatobi J. Afolabi and Bhekisipho Twala

**Abstract:** The continued increase in computing resources is one key factor that is allowing deep learning researchers to scale, design and train new and complex convolutional neural network (CNN) architectures in terms of varying width, depth, or both width and depth to improve performance for a variety of problems. The contributions of this study include an uncovering of how different optimization algorithms impact CNN architectural setups with variations in width, depth, and both width/depth. Specifically in this study, three different CNN architectural setups in combination with nine different optimization algorithms—namely SGD vanilla, with momentum, and with Nesterov momentum, RMSProp, ADAM, ADAGrad, ADADelta, ADAMax, and NADAM—are trained and evaluated using three publicly available benchmark image classification datasets. Through extensive experimentation, we analyze the output predictions of the different optimizers with the CNN architectures using accuracy, convergence speed, and loss function as performance metrics. Findings based on the overall results obtained across the three image classification datasets show that ADAM and NADAM achieved superior performances with wider and deeper/wider setups, respectively, while ADADelta was the worst performer, especially with the deeper CNN architectural setup.

**Keywords:** optimization algorithms; neural network; network size; performance analysis; image classification

url: <https://doi.org/10.3390/app122311976>