

Abstract

With the increasing trend of artificial neural networks toward larger structures with more layers, the training characteristics of these structures has become an important area of research. Many of the initial parameters of these networks have an enormous impact on the convergence characteristics during training. Poorly tuned parameters can often lead to orders of magnitude longer training. In this work, a general framework has been created to define the training and structural parameters of convolutional neural networks using a combination of the stochastic diagonal Levenberg-Marquardt method to accelerate training and evolutionary search for structural and training parameter optimization. Our results show that for the task of handwritten digit classification, the networks created by such a system outperform manually defined networks in speed of convergence.