

Hyper Parameters Optimization of Convolutional Neural Networks using a Genetic Algorithm

SILVIU MARC

Computer Science, Middlesex University, London, UK

In the last decade, technological advancements and the increased availability of data have boosted the use of machine learning (ML). An even more recent subfield of ML is deep learning (DL). The latter represents the use of artificial neural networks (ANN) to resolve very hard or even impossible problems for a human being. Applied DL requires the user to decide on hyperparameters such as the number of layers, learning rates, activation functions, etc. The values of these hyperparameters can be both continuous and discrete. Thus, it creates an almost infinite search space, making guessing the importance of these hyperparameters on the first attempt impossible. Hence, applied DL is a highly iterative process. In addition, hyperparameters optimization (HPO) can influence the generalization ability of an ANN. This problem is also known as overfitting. Some solutions to overfitting are collecting more data, regularising the data and data augmentation (DA). Methods for augmenting data have been widely employed in deep learning. The selection of proper data augmentation procedures is more significant than selecting a network structure. However, due to a lack of essential study, data augmentation strategies have long stayed in the intuition and experience stage, and there is no one general decision strategy.

This research aims to automate the hyperparameter optimization task so that a user can get satisfactory results in an inexpensive time frame. Initially, the experiment will be conducted using a genetic algorithm (GA) that imitates the natural selection process or evolution in searching for the optimal DA techniques. The GA chooses the best individuals from the present population to be parents and utilizes them to generate offspring for the next generation at each stage. The population “evolves” toward an optimal solution over successive generations. Based on the findings and performance of the initial experiment, the GA can be then extended to the HPO problem.

The current research results are expected to solve DL-based medical software problems, where data can be scarce. At the same time, it can help any other industry reduce the cost of collecting data and improve the generalization ability of ANNs.

Keywords: neural networks, algorithm, automation