# WEIGHTED PARAMETER AVERAGING IN DEEP NEURAL NETWORK

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

We propose Weighted Parameter Averaging in Deep Neural Network, an algorithm to improve DNN accuracy and generalization. Our method is to average the neural network parameters of multiple local minima post training, weighted by validation loss of respective model checkpoint. The resulting model generalizes better than single best checkpoint based model over unseen data. We present the performance of Weighted Parameter Averaging in Deep Neural Network at multiple experimental setups. For our experiments we use different combinations of Optimizers and Schedulers over multiple datasets and present the resulting performance scores. We demonstrate the reduction in error and good generalization performance of the resulting models over our experimental datasets. We show the accuracy improvements obtained across multiple type of learning rate schedulers like CyclicLR, CosineAnnealingLR and ReduceLROnPlateau. We also present the comparison of Weighted Parameter Averaging with Stochastic Weight Averaging in the form of accuracy.

## 1 INTRODUCTION

Lately, neural network ensembles have been widely applied in the field machine learning. One of the reason being, ensembles of neural networks are known to be much more robust and accurate than individual networks. However, training multiple deep networks as sumbodels and then averaging the predictions is computationally expensive. In this paper we argue that weighted average of multiple local minima of a single network can be as effective, and we acheive this by saving the model parameters of a single neural network at several local minima.