

IMPROVEMENT ON THE QUATERNION-BASED MODELS: EXTENSION TO LARGER DATASETS AND BATCH NORMALIZATION

ADHILSHA AND ARITRA MUKHOPADHYAY (GROUP 2)

Objective: To improve some quaternion models, implement models on larger datasets, and implement batch normalization.

Dataset: MNIST, Cifar10 and Cifar100. (more datasets as we develop the models better)

Baseline models: Lenet-300-100, conv2, conv4, conv6 and later more complex models like mobilenet and resnet.

Relevant Papers:

- ▶ [1] Sahel Mohammad Iqbal and Subhankar Mishra. 2023. Neural Networks at a Fraction with Pruned Quaternions. DOI:<https://doi.org/10.1145/3570991.3570997>
- ▶ Yang You, Igor Gitman, and Boris Ginsburg. 2017. Large Batch Training of Convolutional Networks. arXiv. DOI:<https://doi.org/10.48550/ARXIV.1708.03888>

Work Division

- ▶ **Aritra:** Reading papers, Experimenting with models, tackling coding and data hurdles, and analyzing results.
- ▶ **Adhil:** Reading papers, documenting the project. tuning and experimenting with hyperparameters, and analyzing the results.

Midway Plans:

- ▶ Understanding and Implementing the models, real and Quaternion for each.
- ▶ To understand the practical drawbacks in implementation of Quaternion models and trace it to the source through code analysis and experimentation.
- ▶ Work on the said model through experimentation on hyperparameter and the improvement of underlying functions.
- ▶ To understand *batch normalization* and start working on implementation.

Further Plans:

- ▶ Continue the work from midway plans onto completion.
- ▶ Implement larger models and on larger datasets successfully.
- ▶ If possible, use hyper-parameter tuning, any preprocessing, or alternate implementations of any underlying functions to improve the accuracy or speed further.

Expected Results:

Improve the quaternion models in speed (and accuracy if possible). Enable the implementation on larger datasets with implementation of batch normalization for quaternions.