Seismic Fault Detection By Deep Learning: Scaling To Large Model

Author block: B. Sun, Saudi Aramco

Abstract

Objectives/Scope: Practical application of the ML-based fault detection method is a challenge due to the vulnerability of the neural network. Such vulnerability partly comes from the highly nonlinear nature of the neural network. While most of it is due to its input, i.e., the complex, noisy and uneven seismic waveform signals itself as an end-to-end neural network solution has to replace the denoising, normalization, and data enhancing step in conventional fault detection algorithms. Using state-of-the-art deep learning technology, we address those challenges for a practical application of ML-based fault detection, especially based on a large neural network.

Methods, Procedures, Process: Challenges are addressed from three perspectives: 1) Reliable dataset for training: Training a neural network with synthetic ones can hardly work for field datasets and thus we train the neural network with real datasets from diverse projects. 2) Robust neural networks: we utilize ensembles of 3D U-net equipment with multi-scale attention mechanism. Ensembles of U-net will vote for the best prediction and improve the generality and avoid the imperfectness of a single neural network. The attention mechanism makes the resulting feature-extraction more robust and reliable. 3) Parallel training: the neural network is trained based on hybrid data and model parallelization using hundreds of GPUs. The training procedure is highly optimized with activation checkpoints and mixed-precision training. We split the model into multiple GPUs by domain decomposition. Currently, using 16 V100 GPUs, we can achieve a 3D U-net with input size of 5123 and the channels of each layers (for the encoder part) can be as large as 256, 512, 1024, 1024, 1024.

Results, Observations, Conclusions: With better data, enhanced neural networks, and optimized training procedures, we can perform robust fault detection using machine learning in a real, large scale applications. We demonstrate that a small neural network trained with synthetic datasets can hardly pick reliable faults from noisy field datasets. While the resulting neural network designed and trained with the suggested procedures can deliver much better fault attributes.

Novel/Additive Information: A novel hybrid data and model parallel training scheme based on domain decomposition and halo zone data exchanging are proposed and we shared numerous tricks for reducing the memory usage and improving the training efficiency. To the best of our knowledge, it is the first time to achieve such large neural network for seismic fault detection in a data and model parallel fashion.