JOURNAL OF RESEARCH ON APPLIED GEOPHYSICS

(JRAG)

Shahrood University of Technology

2022, VOL 8, NO 3

(DOI): 10.22044/JRAG.2023.12105.1338

Development of an automated and computationally low-cost method of huge data generation for training deep learning algorithms using direct sequential simulation

Hamid Sabeti[1*] and Vahid Honarbakhsh[2]

1. Department of Mining Engineering, Birjand University of Technology, Birjand, Iran
2. Department of Earth Sciences, Science and Research Branch, Islamic Azad University, Tehran, Iran

Received: 15 July 2022 ; Accepted: 26 February 2023

Corresponding author: sabeti@birjandut.ac.ir

Keywords

Direct sequential simulation
Deep learning
Velocity model
Geostatistics
Secondary image

Extended Abstract

Summary

Recent successes in inverse problems have led to instant demand for available training datasets. Lack of appropriate datasets, in terms of quantity and quality, has become a challenge in geophysical application of deep learning neural networks. Several methods have been developed to overcome this problem. In this paper, a framework for generating training datasets was introduced using a geostatistical simulation method called direct sequential simulation. The main idea was to extract well logs from different available velocity models and using them as input data for simulation and co-simulation algorithms. Using the secondary image for co-simulation with different correlation coefficients of 0.3, 0.5 and 0.7, various models were generated with different continuities. In this paper, various examples of well-known velocity models were selected to generate more suitable models in terms of geological structures.

Introduction

Deep learning methods are becoming more popular for different applications during recent years. These methods were widely used in seismic data processing, interpolation, imaging, noise reduction, and also, retrieving reservoir parameters. These methods, employing supervised learning, use huge amount of training data in order to better train the algorithm resulting in appropriate resolution of output models. There are several approaches for generating huge amount of training datasets, including: 1. random data generation, 2. geostatistical data generation, and 3. deep learning data generation methods. They have all their own pros and cons. In this paper, a novel framework is developed in which huge and diverse training datasets can be generated. The core part of this framework is a geostatistical simulation method called direct sequential simulation.

Methodology and Approaches

There are several geostatistical simulation algorithms developed under a fundamental algorithm called sequential Gaussian simulation (SGS). In these methods, the assumption of Gaussian distribution of primary data is assured by a Gaussian transformation of the primary data before entering the later steps of simulation algorithm. A newer method, called direct sequential simulation (DSS), uses the original primary data. In this paper, several datasets were used to extract artificial velocity well logs in randomly selected locations. These well logs were used in different simulation strategies to generate diverse training datasets. One strategy was simple direct sequential simulation, and also, three different strategies were applied for doing co-simulation using correlation coefficients of 0.3, 0.5 and 0.7.

Results and Conclusions

A simple and low-cost framework was developed to generate huge training datasets for deep learning algorithms. The introduced framework was successfully applied to several well-known datasets. It was shown that using secondary images and different correlation coefficients from 0.1 to 0.9 for doing direct sequential co-simulation, it was possible to generate more sophisticated training models. It was also shown that the framework was able to generate diverse training datasets to be used by different deep learning methods. In addition, this framework can be used to generate reservoir model using different sources of available data.