Deep neural networks for 1D impedance inversion

Vladimir Puzyrev*
Curtin University
Kent St., Bentley
vladimir.puzyrev@curtin.edu.au

Anton Egorov
Curtin University and
Lomonosov Moscow State University
anton21egorov@gmail.com

Anastasia Pirogova
Curtin University and
Lomonosov Moscow State University
anastasya.pir@gmail.com

Chris Elders
Curtin University
Kent St., Bentley
chris.elders@curtin.edu.au

Claus Otto
Curtin University
Kent St., Bentley
claus.otto@curtin.edu.au

SUMMARY

We investigate the applicability of different types of deep neural networks for the estimation of subsurface properties from seismic data. The pre-trained networks can predict velocity models from new data in a few milliseconds, which makes this data-driven approach especially important for multidimensional inversion, where conventional methods inversion methods suffer from large computational cost. At the same time, realistic one-dimensional models such as the 160-layer velocity model used as an example in this study require large synthetic datasets for training, which are not always possible to obtain. Hence, we also study the impact of extending the training data by adding random noise to the modelled examples. We observe that enlarging training datasets by adding synthetic noise to existing samples improves the quality of inversion without a significant increase in computational complexity.

Key words: deep learning, neural networks, seismic wave, impedance inversion.

INTRODUCTION

Deep learning algorithms have shown great potential in many fields of science and technology and recently have become one of the main focuses of attention in the petroleum and mineral industry. Though these methods are now widely used in seismic processing and interpretation (e.g., Araya-Polo et al. 2017; Zhang et al. 2018; Wu et al. 2019), successful attempts to use deep neural networks directly for the estimation of subsurface properties from seismic data are quite limited. The existing approaches employ supervised machine learning techniques, usually convolutional neural networks (CNNs), a class of neural networks with locally connected layers that apply convolution operation between a kernel and the data. Among the recent applications of CNNs in geophysical inverse problems we can mention the reconstruction of subsurface properties in impedance inversion (Das et al. 2018) and in the waveform inversion scenario (Wu et al. 2018).

Since deep neural networks belong to the class of data-driven methods it is critical to use appropriate data for training. In this paper, we investigate several deep learning approaches to inversion of a 1D subsurface velocity model using synthetic surface seismic and vertical seismic profile (VSP) data. Our main goal is to understand how the size of the training dataset impacts the accuracy of the velocity model restoration and whether the accuracy can be improved using data augmentation by adding synthetic noise to the examples used in training.

METHOD

We consider four types of neural networks architectures, namely the fully connected feed-forward network, the CNN with a fully-connected layer at the end, the fully convolutional network, and the recurrent network with gated recurrent units. The latter type of neural networks is significantly more expensive to train compared to the other three networks architectures. Deep CNNs with many stacked layers are highly efficient in processing images and data with a grid-like topology, which made these neural networks widely used in computer vision applications (e.g., Krizhevsky et al. 2012; Szegedy et al. 2015). The fully convolutional architecture can be efficiently used in multi-dimensional inversion (Puzyrev 2018), where the decoder is used to output an estimation of the spatial distribution of the subsurface properties.

Our data-driven inversion based on deep learning has three stages: data generation, network training and prediction of model parameters from new data. The first and second stages can be performed only once for a given scenario. We stress that the networks do not require to see all possible subsurface models during training, which is impossible even for 1D realistic models with several hundreds of parameters. Instead, we create a sufficiently large representative set of models, which allows the networks to learn the mapping from data to model parameter space. We employ max pooling in the CNN-based networks, rectified linear unit (ReLU) activation functions, and batch normalization between convolutional layers. To prevent overfitting, we use dropout (Srivastava et al., 2014) as a regularization technique to ensure the network is able to generalize effectively. The hyperparameters are chosen to minimize the validation set errors. The training loss is defined as the root mean squared error (RMSE) of the model parameters. For training, we use Adam optimization algorithm (Kingma and Ba 2014) and stop the training once the validation errors start to increase (early stopping), thus preventing overfitting to the training data.

The accuracy of the method has been previously validated on a 20-layered velocity model where the network was able to accurately estimate the velocities of all layers from zero-offset surface data (Puzyrev et al. 2019). At the same time, we discovered that more realistic models such as the 160-layer velocity model with variable thicknesses of the layers require training datasets larger than the employed set of 10240 velocity models in order to achieve high accuracy. While the most efficient networks have sufficiently low errors on the training dataset, they exhibit much higher errors on the data that was not used in training, which means poor generalization of the method. One of the possible reasons for that is lack of training data required to learn all the features of realistic models with many thin layers.

Data Generation

Perhaps the main drawback of CNNs, as well as other deep learning methods, is the requirement of a large set of labelled data for training. This dataset should be representative enough and balanced such that the networks do not become over-tuned to specific cases. Obtaining more data is not always possible in real-world applications, and in such cases data augmentation procedure can be used. In the context of image processing data augmentation usually means randomly rotating the image, adding reflections, zooming in, adding a colour filter and other operations that can be done with very little computational effort. For problems when the data results from the simulations of a physical phenomenon, one of the simplest techniques to enlarge the dataset is to add artificial noise to measured data.

The generation of the synthetic data used to train the networks is based on the acoustic matrix propagator technique (Ma et al. 2004). All the interbed multiples are included in the modelling result, while the free surface multiples are not generated. In general, this inversion can be viewed as an alternative to 1D full waveform inversion. In the following example, the density is assumed to be constant and we invert only for velocity. Two acquisition geometries, namely, zero-offset surface seismic and zero-offset VSP are used. The dataset includes 16384 subsurface velocity models; 14336 of them are used in training and the remaining 2048 are split equally between the validation and test sets. The maximum depth of the models is 3500 m; all of them include a 500-m thick water layer on top. Starting from the sea bottom, each model has 160 thin layers, whose thicknesses gradually increase from 10 m to 30 m. The velocity in each layer is generated independently as a random Gaussian quantity. The mean of these quantities lies on the linear velocity trend, which rises from 1500 m/s at the sea bottom to 3200 m/s at the bottom of the model, and the standard deviation is equal to 10% of the trend value. This allows us to account for a wide range of formations including shales, clays, wet sands, chalks, porous and saturated sandstones. Figure 1 shows an example of a surface seismic trace, a VSP gather and two different subsurface models.

NUMERICAL RESULTS

Figure 2 compares the predictions of the CNN trained on the original dataset (16384 models and the corresponding noise-free seismic data) versus the true velocity models with 40 layers of variable thickness. In this relatively simple case, most velocity contrasts are recovered with a sufficient degree of accuracy. The average normalized RMSE of the test dataset is 0.052.

Next, in Figure 3 we compare the predictions of the most successful neural network among those tested with the true velocities of the complex model with 160 thin layers of variable thickness. The dataset is extended by adding 1%-5% random noise to the measured data and now contains 81920 examples. The best achieved average normalized RMSE of the validation and test datasets is at quite large values of 0.12-0.13 and does not decrease after several tens of epochs. The training errors monotonically decrease as expected, but since the validation test accuracy is stable or even decreases, this simply means overfitting to the training data. At the same time, performance is improved compared to the network trained on a smaller (10240 velocity models) dataset with zero noise. Velocities in individual layers of the models from the test dataset shown in Figure 3 are often mismatched but the networks quite accurately capture the general trend. When the training dataset is extended using models with synthetic noise, more velocity contrasts are recovered with a sufficient degree of accuracy.

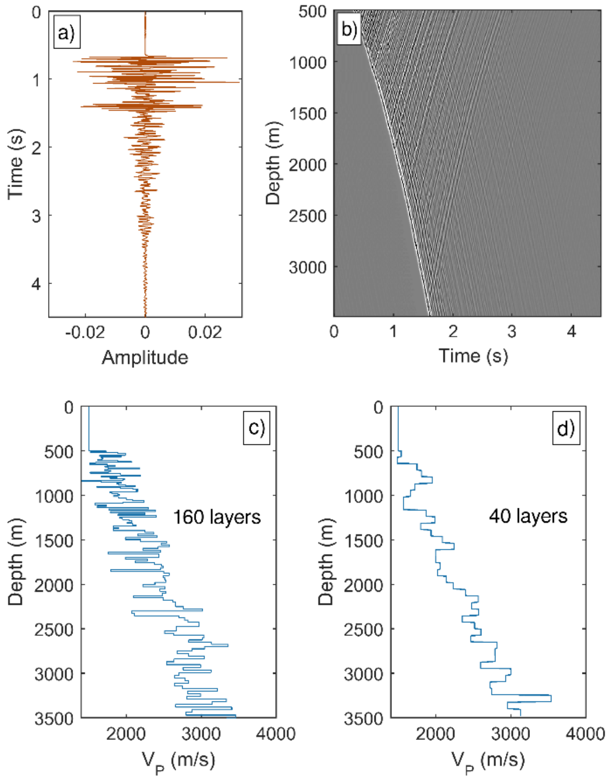

Figure 1. Top row: input data in form of the zero-offset surface seismic trace (a) zero-offset VSP gather (b). Bottom row: output data, namely, the corresponding velocity model with 160 layers (c) and a sample coarse velocity model with 40 layers (d).

CONCLUSIONS

We present an application of deep neural networks in seismic impedance inversion. The method can estimate the unknown 160-layer velocity model from data in a few milliseconds. This will be particularly important for multidimensional inversion, where data-driven methods based on deep learning can deliver results orders of magnitude faster than conventional methods. The CNNs trained on large datasets that include the synthetic noisy examples yield better generalization performance compared to the networks trained on real modelling results only. Using data with synthetic noise overcomes some of the limitations of employing only real field data or noise-free modelling results for training. Future research will target inverting both velocity and density simultaneously.

ACKNOWLEDGEMENTS

The authors acknowledge the Curtin University Oil and Gas Innovation Centre (CUOGIC), the Curtin Institute for Computation (CIC), and the Institute for Geoscience Research (TIGeR). This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia.

REFERENCES

Araya-Polo, M., Dahlke, T., Frogner, C., Zhang, C., Poggio, T. and Hohl, D., 2017, Automated fault detection without seismic processing: The Leading Edge, 36(3), 208-214.

Das, V., Pollack, A., Wollner, U., and Mukerji, T., 2018, Convolutional neural network for seismic impedance inversion: SEG Technical Program Expanded Abstracts, 2071-2075.

Kingma, D.P., and Ba, J., 2014, Adam: A method for stochastic optimization: arXiv preprint arXiv:1412.6980.

Krizhevsky, A., Sutskever, I., and Hinton, G.E., 2012, Imagenet classification with deep convolutional neural networks: Advances in neural information processing systems 1097-1105.

Ma, Y., Loures, L., and Margrave, G. F., 2004, Seismic modeling with the reflectivity method: CREWES Research Report, 16-1.

Puzyrev, V., 2018, Deep learning electromagnetic inversion with convolutional neural networks: arXiv preprint arXiv:1812.10247.

Puzyrev, V., Egorov, A., Pirogova, A., Elders, C., and Otto, C., 2019, Seismic inversion with deep neural networks: a feasibility analysis: 81st EAGE Conference and Exhibition.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R., 2014, Dropout: a simple way to prevent neural networks from overfitting: The Journal of Machine Learning Research, 15(1), 1929-1958.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A., 2015, Going deeper with convolutions: Proceedings of the IEEE conference on computer vision and pattern recognition, 1-9.

Wu, Y., Lin, Y., and Zhou, Z., 2018, InversionNet: Accurate and efficient seismic waveform inversion with convolutional neural networks: SEG Technical Program Expanded Abstracts, 2096-2100.

Wu, X., Liang, L., Shi, Y. and Fomel, S., 2019, FaultSeg3D: using synthetic datasets to train an end-to-end convolutional neural network for 3D seismic fault segmentation: Geophysics, 84(3), 1-36.

Zhang, G., Wang, Z., and Chen, Y., 2018, Deep learning for seismic lithology prediction: Geophysical Journal International, 215(2), 1368-1387.

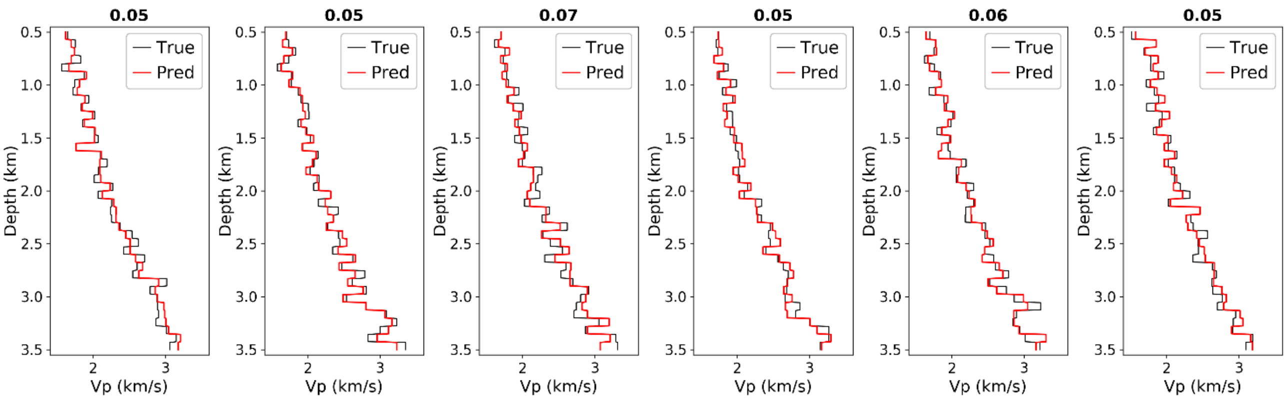

Figure 2. Accuracy of the CNN prediction for the model with 40 layers of variable thickness. The shown examples are randomly chosen from the test dataset. The average normalized RMSE for each example is shown in the title.

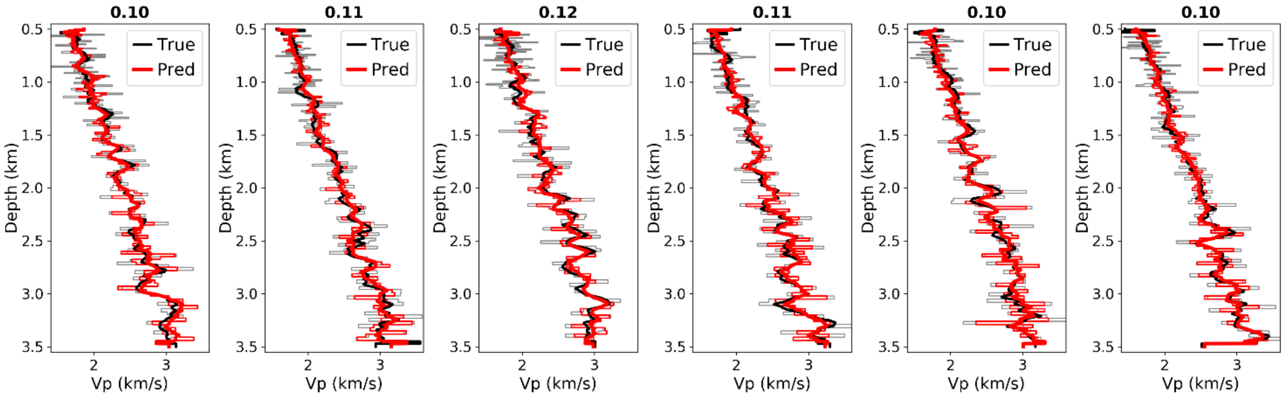

Figure 3. Accuracy of the CNN prediction for the model with 160 layers of variable thickness. The shown examples are randomly chosen from the test dataset. The average normalized RMSE for each example is shown in the title.