Th_R11_12

Transfer Learning For Low Frequency Extrapolation From Shot Gathers For FWI Applications

O. Ovcharenko^{1*}, V. Kazei¹, D. Peter¹, T. Alkhalifah¹

¹ King Abdullah University of Science and Technology

Summary

Low-frequency data proved to be crucial for robust full-waveform inversion (FWI) applications. However, acquiring those data in the field is a challenging and costly task. Deep neural networks can be trained to extrapolate missing low frequencies, but no optimal network configuration exists. Therefore, the search for an acceptable network architecture is a tedious empirical task whose outcome heavily affects the performance of the application. Here, we propose and utilize transfer learning to reduce the computational efforts otherwise spent on an optimal architecture search and an initial network training. We re-train the light-weight MobileNet convolutional network to infer low-frequency data from a frequency-domain representation of the individual shot-gathers, which leads to an efficient, yet accurate inference of low frequencies according to wavenumber theory. In particular, we show that the extrapolated 0.25 - 1 Hz from 2-4.5 Hz data are accurate enough for acoustic FWI on part of the original BP 2004 model and the Marmousi II model of double scale. We bridge the gap between the 1 Hz predicted and the 2 Hz modeled data by the application of a Sobolev space norm regularization.

Introduction

Acoustic waves propagating in the subsurface carry valuable information about velocity anomalies settling on their paths. Seismic full-waveform inversion (FWI) attempts to utilize this information by matching synthetic and observed waveform data through refining a subsurface velocity model (Virieux and Operto, 2009). Despite advances made over the past decades (Gray, 2016; Kazei and Alkhalifah, 2018), inversions of complex geological regions still suffer from stagnating in a local minimum due to the nonlinearity of the optimization problem. However, inversions are likely to converge when initiated from a proximate starting velocity model. Building a plausible initial model for a complex media is a serious problem, especially in the presence of salt bodies. Usually in this case, the challenge of building an initial model is addressed either by salt-flooding (Ovcharenko et al., 2018a; Kalita et al., 2018), or by extracting low-frequency components from the wavefield spectrum and their consecutive inversion to produce a starting model for a high-frequency inversion. Monochromatic components of the wavefield interact with velocity anomalies at different scales, meaning that low-frequency components of the wavefield illuminate the largest bodies and high-frequency components are responsive to fine-scale anomalies. Thus, low-frequency data help to bring us to the area (basin of attraction) of the global minimum.

A number of approaches have been proposed to reconstruct the missing low-frequency content (Hu, 2014; Wang and Herrmann, 2016). Data-driven approaches based on machine learning methods for shot-by-shot (Ovcharenko et al., 2017; Jin et al., 2018) and trace-by-trace (Sun and Demanet, 2018) low-frequency data extrapolation showed promise in particular as they allow less assumptions to be imposed on the task. A feasibility study of this low-frequency extrapolation approach from shot-gather to shot-gather was shown in Ovcharenko et al. (2018b). The structure of the network is crucial as the number and types of the layers and the set of hyper-parameters heavily affect performance of its application. However, manual or semi-automated search for a proper network architecture is a recurrent empirical task which often leads to inconclusive results.

In this study, we use the concept of transfer learning to reduce the ambiguity of finding an appropriate network architecture. We re-train the MobileNet convolutional feature extractor for the low-frequency extrapolation problem and explore benefits and drawbacks of such an approach to seismic data processing.

Method

A generic convolutional neural network consists of two principal parts. A feature extractor, or encoder, which processes input volume to obtain high-level features and a fully-connected part which manipulates the encoder's outputs to address the target task. During training, convolutional kernels in the shallow layers of the encoder are tuned to match general features in the input volume, such as straight lines and color contrasts. More complex features, such as circles, are learned in the deeper layers of the encoder. In real world tasks, objects often share the same low-level features, so an encoder from a pre-trained network might be used to extract features for conceptually similar tasks.

Transfer learning is a machine learning method where a model developed for a certain task is reused as the starting point for another related task (Goodfellow et al., 2016). Such an approach works only when the model features learned from the original task are general and might be found in the new formulation. In transfer learning, a base network is first trained on a general base dataset. Then, the same network is fine-tuned by re-training on a new target dataset which usually is smaller compared to the base dataset.

MobileNets is a class of light-weight deep learning models developed for embedded vision applications on portable devices (Howard et al., 2017). MobileNet's capability is comparable to popular network architectures (Huang et al., 2017), however at a lower computational cost. The key feature of this architecture is that a standard convolution is factorized into depth-wise and point-wise convolutions, which drastically reduces the computational costs of both training and inference. Here, we select the MobileNet as the most suitable architecture for feature extraction which allows for several input data sizes and has an affordable number of trainable parameters to fine-tune them for a new application.

Framework

We run a synthetic experiment mimicking a streamer marine acquisition setup. Fixed-offset data can naturally be used in supervised learning applications where the size of input and target data pairs must be equal throughout training and testing datasets. We arrange each shot-gather into a complex-valued matrix, comprising a number of its mono-frequency representations. Such an arrangement leads to a tensor of $F \times R \times 2$ size for each shot-gather, where F is number of frequencies, R is number of receivers and the third dimension comes from real and imaginary parts of the data in frequency domain. Since the input volume for the MobileNet is a three-dimensional tensor which is usually treated as an RGB image with three channels, we populate the third dimension and add the phase of the complex-valued wavefield as the third channel of the input volume (Figure 1). Moreover, to meet the requirement of the squared input image, we set $F = R$.

Figure 1: The spectral decomposition of a shot-gather, which shapes a three-channel input for the network. From left to right: real and imaginary parts, and angle for a shot-gather. These data shape the RGB image when used as color channels.

The network is trained using pairs of input and target data. The size of the input volume is $F \times R \times 3$, whereas the size of the target data is $1 \times R \times 2$. This way, the network infers a complex-valued mono-frequency shot-gather from its multi-frequency representation. The network architecture consists of an encoder part from the MobileNet with global average pooling applied to its outputs, and two fully-connected layers. There are $F \times R \times 2$ neurons with hyperbolic tangent activation functions in each dense layer. We initiate training from the weights pre-trained on the ImageNet dataset (Deng et al., 2009). Once trained, the network might be applied for mono-frequency data extrapolation in various scenarios, given the same acquisition geometry.

Example

Assuming that frequency data are available in a range from 2 to 4.5 Hz, we use those for training of the network. To generate the training dataset, we solve the Helmholtz acoustic wave equation for 64 frequencies in 200 different velocity models built from 6 km-deep random vertical 1D profiles. We extrapolate data for a single shot-gather at a time, thus we populate the entire data matrix shot-wise and use reciprocity to fill the missing parts, except those which are empty due to the limited offset acquisition. Shot-wise, rather than entire acquisition-wise extrapolations provide a more flexible framework and dramatically reduce computational costs needed for the generation of the training data. This is also natural since the wavefield is computed from a shot, and the Earth response for various frequencies has little to do with acquisition. From a single random velocity model we generate as many training samples as there are sources. The streamer acquisition involves 256 shots and 128 receivers placed at a 40 m spacing, which finally leads to 40,960 data samples used for training and 10,240 samples used for validation.

Multi-scale acoustic FWI (Bunks et al., 1995) reasonably converges when consequently applied to extrapolated mono-frequency data at 0.25, 0.5, 1 Hz and a number of frequencies within the known range, 2-4.5 Hz, for the central part of the BP 2004 (Billette and Brandsberg-Dahl, 2005) and the resized to the same dimensions Marmousi (Bourgeois et al., 1991) benchmark models (Figure 3). In both scenarios, we initiate the inversion from a linear velocity model and use the extrapolated data to build a sufficient background model for inversions at higher frequencies. The network learns the mapping needed for a single, low frequency. We train three identical networks on the same training shot-gather dataset, chang-

Figure 2: Real part for true and extrapolated 0.25 Hz data matrices for the central part of the BP 2004 model. Individual shot-gathers are assembled into a data matrix and using reciprocity to populate the gaps (green). Red stars and blue circles indicate locations of sources and receivers, respectively. Plot in the right bottom demonstrates the extrapolation loss for frequencies in the range 0.25–1 Hz.

ing only the target low-frequency data. The extrapolation accuracy decreases for higher frequencies as the complexity of data increases, exhibiting more fine-scale variations (Figure 2). Thus, we regularize the inversion (Kazei et al., 2017) to mitigate inaccuracies in the extrapolated data, whereas inversion fails when applied directly using the raw network outputs.

Figure 3: Full-waveform inversion of extrapolated 0.25, 0.5 and 1 Hz data, and data from the known range 2–4.5 Hz for the central part of the BP 2004 and Marmousi models. For each extrapolated low frequency we once train a network and then use it for both velocity models. From top to bottom, the initial model build from extrapolated 0.25 Hz data, inverted and true models.

Conclusions

We demonstrated that the encoder part from the MobileNet architecture can serve well to extract features from a frequency-domain representation of a shot-gather. The major benefit of this transfer learning approach is that it allows us to eliminate a tedious search for appropriate neural network architecture. However, the base and target datasets should share similar features for this approach to work. We found that in case of low-frequency extrapolations from a shot-gather, all layers of the network should be fine-tuned from their original weights. This means that our synthetic dataset generated from random initializations of velocity models appears to be significantly different from the base dataset used in the

initial stage of the network pre-training. Furthermore, our results show the benefits of low-frequency extrapolation from shot-gather to shot-gather and demonstrate that extrapolated data are adequate for a noise-free acoustic FWI. However, the extrapolation quality drops at higher frequencies due to the complex subsurface structures considered in this study. Regularized inversions help to partially mitigate inaccuracies in the extrapolated data, but the applicability of such a frequency-bandwidth extension approach for real acquisition applications requires further investigations.

Acknowledgements

The research reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST), Thuwal, 23955-6900, Saudi Arabia.

References

- Billette, F. and Brandsberg-Dahl, S. [2005] The 2004 BP velocity benchmark. In: *67th EAGE Conference & Exhibition*.
- Bourgeois, A., Bourget, M., Lailly, P., Poulet, M., Ricarte, P. and Versteeg, R. [1991] Marmousi, model and data. *The Marmousi Experience*, 5–16.
- Bunks, C., Saleck, F.M., Zaleski, S. and Chavent, G. [1995] Multiscale seismic waveform inversion. *Geophysics*, **60**(5), 1457–1473.
- Deng, J., Dong, W., Socher, R., Li, L.J., Li, K. and Fei-Fei, L. [2009] Imagenet: A large-scale hierarchical image database. In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. Ieee, 248–255.
- Goodfellow, I., Bengio, Y. and Courville, A. [2016] *Deep learning*, 1.
- Gray, S.H. [2016] Seismic imaging. In: *Encyclopedia of Exploration Geophysics*, Society of Exploration Geophysicists, S1–1.
- Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. and Adam, H. [2017] Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.
- Hu, W. [2014] FWI without low frequency data-beat tone inversion. In: *SEG Technical Program Expanded Abstracts 2014*, Society of Exploration Geophysicists, 1116–1120.
- Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer, I., Wojna, Z., Song, Y., Guadarrama, S. et al. [2017] Speed/accuracy trade-offs for modern convolutional object detectors. In: *IEEE CVPR*, 4.
- Jin, Y., Hu, W., Wu, X. and Chen, J. [2018] Learn low wavenumber information in FWI via deep inception based convolutional networks. In: *SEG Technical Program Expanded Abstracts 2018*, Society of Exploration Geophysicists, 2091–2095.
- Kalita, M., Kazei, V., Choi, Y. and Alkhalifah, T. [2018] Regularized full-waveform inversion with automated salt-flooding. *Geophysics*, *submitted*.
- Kazei, V. and Alkhalifah, T. [2018] Waveform inversion for orthorhombic anisotropy with P waves: feasibility and resolution. *Geophysical Journal International*, **213**(2), 963–982.
- Kazei, V., Kalita, M. and Alkhalifah, T. [2017] Salt-body Inversion with Minimum Gradient Support and Sobolev Space Norm Regularizations. In: *79th EAGE Conference and Exhibition 2017*.
- Ovcharenko, O., Kazei, V., Peter, D. and Alkhalifah, T. [2017] Neural network based low-frequency data extrapolation. In: *3rd SEG FWI workshop: What are we getting?*, Society of Exploration Geophysicists.
- Ovcharenko, O., Kazei, V., Peter, D. and Alkhalifah, T. [2018a] Variance-based model interpolation for improved full-waveform inversion in the presence of salt bodies. *Geophysics*, **83**(5), 1–60.
- Ovcharenko, O., Kazei, V., Peter, D., Zhang, X. and Alkhalifah, T. [2018b] Low-Frequency Data Extrapolation Using a Feed-Forward ANN. In: *80th EAGE Conference and Exhibition 2018*.
- Sun, H. and Demanet, L. [2018] Low frequency extrapolation with deep learning. In: *SEG Technical Program Expanded Abstracts 2018*, Society of Exploration Geophysicists, 2011–2015.
- Virieux, J. and Operto, S. [2009] An overview of full-waveform inversion in exploration geophysics. *Geophysics*, **74**, WCC1–WCC26.
- Wang, R. and Herrmann, F. [2016] Frequency down extrapolation with TV norm minimization. In: *SEG Technical Program Expanded Abstracts 2016*, Society of Exploration Geophysicists, 1380–1384.