# Geophysical data and gradient translation using deep neural networks

*Jiashun Yao, Lluís Guasch, Mike Warner (Imperial College London), Tim Lin (S-Cube London) and Elizabeth Percak-Dennett (Amazon Web Services, Houston)*

**Summary**

Image-to-image translation generative adversarial networks (GANs) have been applied successfully to a wide variety of non-geophysical problems, from mundane implementations to turn horses into zebras, to stunning synthetic media deep-fakes. We explore similar applications in geophysics as cost-reduction tools, and demonstrate their potential on 3D field data. We show these networks can be used to learn the mapping between different data types in data processing workflows: acoustic-to-elastic for full-waveform inversion, and hydrophone-to-geophone for wavefield separation.

**Introduction**

We present the application of 2D image-to-image translation algorithms for 3D geophysical data. We demonstrate its potential with three different applications: elastic data synthesis, elastic gradient correction, and geophone data synthesis, where the first and the last are data-domain translations, while the second is formulated in model-space. The objective of elastic translation is to enable practical elastic full-waveform inversion (FWI), by reducing its cost to that of acoustic FWI. The aim of the geophone-data synthesis is to obtain vertical component data without having to deploy dense ocean-bottom arrays.

**Method**

Following state-of-the-art trends in image-to-image translation applications, we use supervised learning based upon fully-convolutional networks for the data generators, and adversarial loss combined with $L_1/L_2$ losses for the network objective functions (Long et al., 2015; Wang et al., 2015; Zhu et al., 2017). In our implementation, the inputs and outputs of the network are paired geophysical datasets from two different domains $A$ and $B$ (acoustic and elastic, or hydrophone and geophone) that are used to train a network $\mathbf{G}$ to perform accurate and efficient translation from domain $A$ to domain $B$, written as $\mathbf{G}(A) \rightarrow B$.

The best cost-efficient results are achieved using a 9-block ResNet for the generator and a 4-layer PatchGAN for the discriminator. To improve the accuracy of the network outputs in the data-domain applications, in addition to the forward matching loss $\|\mathbf{G}(A) - B\|$ and the adversarial loss

$$E_A\left[\log\left\{1-\mathbf{D}\left(\mathbf{G}(A)\right)\right\}\right] + E_B\left[\log\left\{\mathbf{D}(B)\right\}\right],$$

where $E$ is the expected value, we also train an inverse network to perform a reverse translation $\mathbf{F}(B) \rightarrow A$ in order to construct cycle-consistency and identity loss functions (Zhu et al., 2017). In the model-domain, however, due to the large dimensionality of the data, we only train the forward network using conditional-GANs and do not include these extra losses.

For the elastic/acoustic to acoustic/elastic cases we use a shallow water 3D OBC dataset with ~1000 receivers and ~40,000 shots. This dataset contains strong elastic effects caused by a shallow chalk layer that prevents pure-acoustic FWI from converging to a realistic solution because the acoustic wave equation cannot model P-S conversions, which results in post-critical multiples that have much more energy than is observed in the field data. For the hydrophone/geophone application, we use a different shallow water 3D OBC dataset with 400 shots and 8 240-channel receiver cables. In this case, the chalk layer is much deeper and its elastic effects are minimal within the operational time-window. In all cases, we apply reciprocity and divide the 3D gathers into 2D shot lines because the underlying networks are designed to operate in two-dimensions only.

**Acoustic/Elastic data translation**

To avoid the high cost of full elastic FWI, we eliminate the elastic effects from the field data using neural networks to obtain an acoustic version of the field data that will, in principle, allow pure-acoustic FWI to converge to a stable elastic solution. We perform both acoustic and elastic propagation for all the shots using the initial Vp model shown in Figure 1a, and explore if we can train a network based on only a small portion of the data which can then produce the rest of the elastic data accurately. The networks are trained using 50 evenly distributed paired acoustic and elastic shots (~5000 2D gathers), and then applied deployed to the rest of the dataset (~95,000 2D gathers).

Figure 3 summarizes the training process: the forward generator $\mathbf{G}_{A2E}$ (or $\mathbf{G}_{E2A}$) turns a real acoustic (or elastic) gather into a fake elastic (acoustic) gather which accurately matches its paired elastic (acoustic) gather. The results are then used as inputs to generate the original gather, as well as the identity gather (that is, using the network $\mathbf{G}_{A2E}$ with an elastic input and vice-versa), to fine-tune their performance (two right panels in Figure 3). The trained network is then applied on the remaining 950 shots to generate fake elastic data from acoustic synthetics. The average test error is 1.67%, defined as $\text{RMS}(\text{fake}_E, \text{real}_E)$ / $\text{RMS}(0, \text{real}_E)$, compared to the error between the original acoustic and elastic gathers, which is 79.65%.

DOI:10.1190/segam2020-3423766.1

Our approach is twelve times cheaper in terms of the cost of on-demand cloud computing resource when compared to generating all the data using a fully elastic propagator, as shown in Table 1. Furthermore, after removing the elastic effects from the field data, we can run acoustic instead of elastic FWI, which significantly reduces the cost of the model inversion. Our approach results in an FWI model that is more stable than just using field data – compare Figure 1b and c. However, in this example, acoustic FWI still fails to converge in the areas below the chalk at depths around 3 km. As will be shown in the next section, translations in model-space are better suited for this type of problem.

**Acoustic/Elastic model translation**

There are two problems with the data-translation approach above. The first is that, potentially, the network can absorb some of the data mismatch due to model differences, and so prevent FWI from recovering the correct solution. The second problem is that most supervised-learning algorithms rely on the assumption that the target and training data inherently share the same underlying distribution. In the previous example however the synthetic elastic data differs from the observed because it does not contain information associated with geological structures absent in the smooth initial model, Figure 1a. If the initial model deviates far from the real subsurface, this approach will fail.

To mitigate this out-of-distribution problem, we reformulate the acoustic/elastic translation in model-space by training a network that operates on the local FWI gradients. The approach is similar to its data-space counterpart: we initially run both acoustic and elastic FWI using a subset of 50 shots and capture the local single-shot gradients. The gradients generated with elastic FWI use a fixed Vp/Vs ratio that generates a relatively sharp boundary at the top of the chalk to be able to capture the P-S conversion. We then train a conditional GAN to convert acoustic local gradients into elastic local gradients, Figure 2. For the remaining iterations, we only compute acoustic local gradients and convert these to elastic gradients using the trained network. As our FWI implementation is based on frequency continuation, we repeat the same gradient training strategy at every frequency band, starting at 3 Hz and gradually increasing the inversion frequency to 6 Hz.

Figure 1d illustrates how this model-space approach improves the reconstruction of the deep layers below 2 km, and mitigates the artifacts seen on the edges of the model in the data-space approach. This result closely matches the final model obtained running fully elastic FWI, Figure 1d.

**Hydrophone/Geophone data translation**

One other application of data-to-data translation is to train a network to derive vertical-component ocean-bottom geophone data from hydrophone pressure data. The rationale here is to enable the separation of up- and down-going wavefields using P-Z summation using a small number of ocean-bottom geophones in order to reduce acquisition cost. Here we show a brief example using 3,200 paired 2D-gathers of hydrophone and vertical component geophone data from a shallow-water 3D OBC dataset which does not have a shallow chalk layer and, therefore, is not dominated by strong elastic effects.

Due to the small size of the dataset, we split it into two subsets corresponding to alternate in-lines, one for training and one for testing. Despite the small size of the training set, we achieve a good match between the network-generated z-data and its real counterpart, both shown in Figure 4, producing an average cross-correlation coefficient of 0.877. It is important to mention that this evaluation metric includes the effect of noise in the field data, which is absent in the synthesized gathers because the field data contain randomly scattered high frequencies that cannot (and should not) be learned by the network. We also applied a perceptual structural-similarity index metric to evaluate the similarity between the fake and real vertical component gathers, and achieved a value of 0.962.

**On-demand computational resource allocation**

The model-space translation approach involves a relatively heterogeneous and complicated workload which lends itself to finer-grained, on-demand resource allocation. The translation phase happens at the end of every wavefield simulation for a given shot, but typically wave propagation and neural-network application require vastly different hardware profiles. For example the wavefield propagation is typically CPU based, while the neural network translation is typically implemented on GPUs. Furthermore, the retraining step for the neural network at the end of each frequency band will often benefit from having much more GPU memory and high-precision floating-point throughput compared to the inference tasks, which have much lighter memory and precision requirements. It is therefore difficult to design a fixed HPC hardware mix to achieve optimal resource allocation for this kind of hybrid FWI task. Instead of a fixed resource allocation for the entire run, this kind of workflow can be defined as task arrays that have bespoke hardware allocations for the wavefield propagation and neural network tasks.

Evolving this design further leads to an alternate, event-driven approach that is especially suitable for this particular workflow, where individual hardware resources are allocated on an on-demand basis. This is an architecture commonly seen in large cloud environments. The overall FWI algorithm may still be driven by an existing software



10.1190/segam2020-3423766.1

framework, but because in cloud environments hardware allocations are simply function calls (which usually translates to HTTP calls), each worker can perform the assigned computational tasks by delegating to a prescribed set of optimal hardware/software-kernel combinations. This is an extension to the longstanding HPC practice of selecting an optimal computational kernels at runtime.

## Conclusions

Supervised deep neural networks have the potential to impact the way we design processing and imaging algorithms in geophysics. In this paper we demonstrate their potential to reduce costs in two different and unrelated problems: elastic FWI and PZ-summation. With carefully selected network architectures and suitable hyper-parameter selection, this approach is able to synthesize high-fidelity geophysical data at a fraction of its normal operational or processing cost. Although we are still far from generic deep neural networks that are dataset-agnostic, we hypothesize that augmenting the training sets to include data from adjacent areas can result in networks that are region-specific, as opposed to dataset-specific.

## Acknowledgements

We are grateful to the sponsors of the Fullwave Consortium and the Resource Geophysics Academy, Imperial College London, for their support, and to S-Cube and AWS for permission to publish.

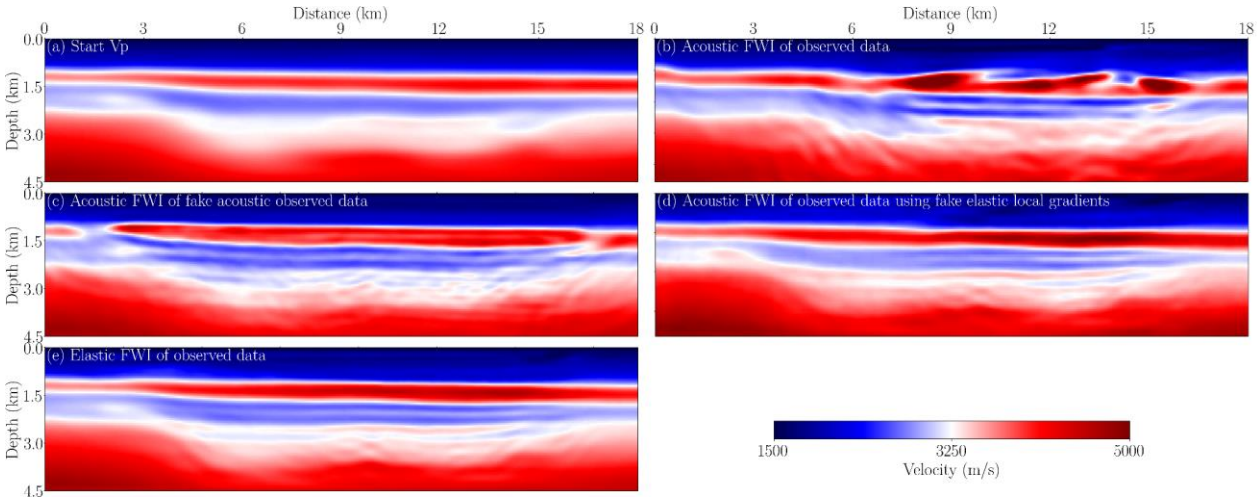


Figure 1: (a) Starting model. (b) Acoustic FWI of observed data. (c) Acoustic FWI of fake acoustic observed data. (d) Acoustic FWI of observed data using fake elastic local gradients. (e) Full-elastic FWI of observed data.

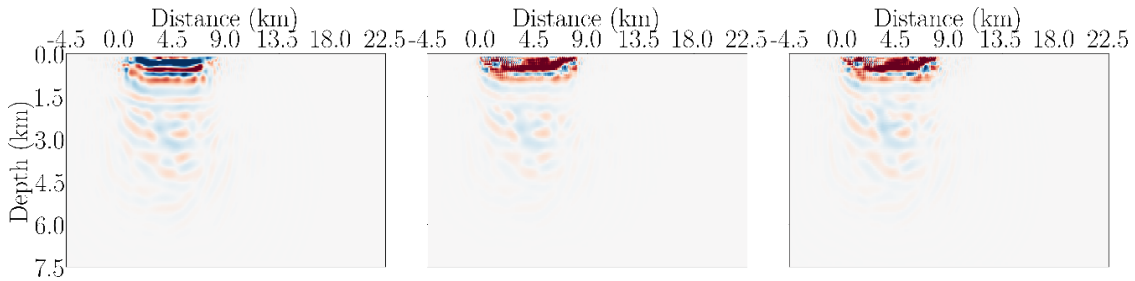


Figure 2: The trained network turns an acoustic single-shot gradient (left) into its elastic analogue (middle), which closely matches the true elastic single-shot gradient (right).


DOI:10.1190/segam2020-3423766.1

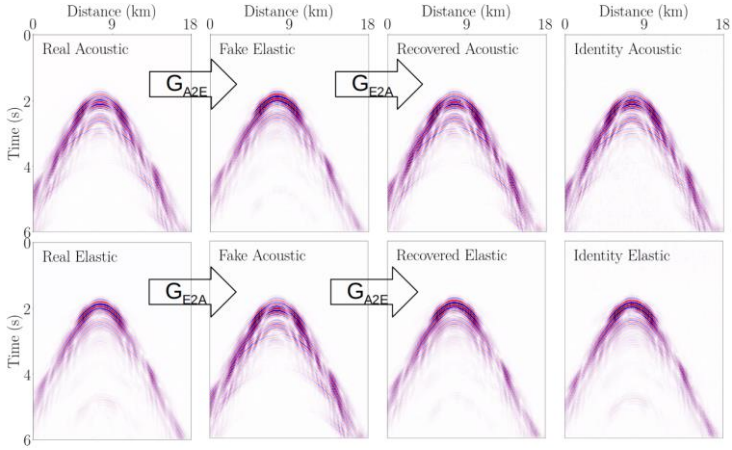


Figure 3: A snapshot of the training process. Note that the networks accurately remove or introduce strong multiples (due to lack of P-S conversions) in the acoustic gathers that are not present in the elastic gathers.

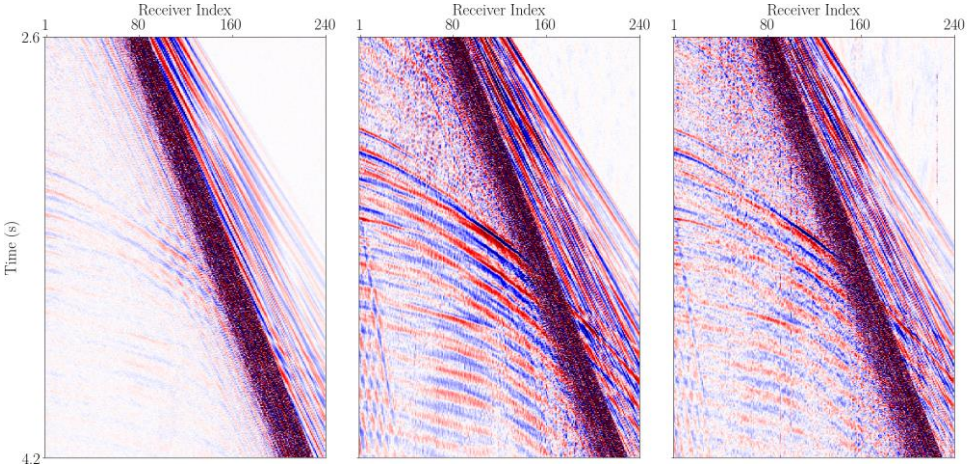


Figure 4: The trained network turns hydrophone (left) into vertical component data (middle), which closely matches the true acquired vertical component data (right).

| Task | Hardware | Time | Approximate Cost on AWS |
|---|---|---|---|
| 3D acoustic propagation of 100k 2D-gathers | 5 HPC nodes | 3.5 hours | $ 17 |
| 3D elastic propagation of 100k 2D-gathers | 32 HPC nodes | 3.5 days | $ 2710 |
| Network training using 5k gathers | Single V100 GPU | 20 hours | $ 66 |
| Network prediction of 100k gathers | Single V100 GPU | 2.5 hours | $ 8 |
| Acoustic propagation of 100k gathers | 5 HPC nodes | 3.5 hours | $ 266 |
| Elastic propagation of 5k gathers | 32 HPC nodes | 4.2 hours | |
| Network training | Single V100 GPU | 22.3 hours | |
| Network prediction 95k gathers | | | |

Table 1: Cost drops by replacing elastic propagation ($ 2710) to neural network elastic data generation ($ 226) for this dataset. Note: this 3D elastic propagation requires at least 72 times more memory than its acoustic analog. Cost estimates are based on the current on-demand price of Amazon Web Services EC2 instances m5.4xlarge (HPC) and p3.2xlarge (GPU) in the Ireland (eu-west-1) region.

## REFERENCES

Long, J., E. Shelhamer, and T. Darrell, 2015, Fully convolutional networks for semantic segmentation: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3431–3440.

Wang, L., W. Ouyang, X. Wang, and H. Lu, 2015, Visual tracking with fully convolutional networks: Proceedings of the IEEE International Conference on Computer Vision, 3119–3127.

Zhu, J.-Y., T. Park, P. Isola, and A. A. Efros, 2017, Unpaired image-to-image translation using cycle-consistent adversarial networks: arXiv preprint.


DOI:10.1190/segam2020-3423766.1