

AI-complemented first-break picking for field low-S/N seismic data

Lionel J. Woog¹, Anthony Vassiliou¹, and Rodney Stromberg¹

<https://doi.org/10.1190/tle40060306.1>

Abstract

In seismic data processing, static corrections for near-surface velocities are derived from first-break picking. The quality of the static corrections is paramount to developing an accurate shallow velocity model, a model that in turn greatly impacts the subsequent seismic processing steps. Because even small errors in first-break picking can greatly impact the seismic velocity model building, it is necessary to pick high-quality traveltimes. Whereas various artificial intelligence-based methods have been proposed to automate the process for data with medium to high signal-to-noise ratio (S/N), these methods are not applicable to low-S/N data, which still require intensive labor from skilled operators. We successfully replace 160 hours of skilled human work with 10 hours of processing by a single NVIDIA Quadro P6000 graphical processing unit by reducing the number of human picks from the usual 5%–10% to 0.19% of available gathers. High-quality inferred picks are generated by convolutional neural network-based machine learning trained from the human picks.

Introduction

Machine learning approaches have been presented to derive first-arrival traveltimes from synthetic data (Yuan et al., 2018; Tsai et al., 2020) or data with high signal-to-noise ratio (S/N) (Veezhinathan et al., 1991; Murat and Rudman, 1992; McCormack et al., 1993; Zhe et al., 2013). We introduce an approach that delivers production-quality results on the kind of real-world low-S/N land data for which first arrivals are most needed.

Due to the rapid progress in processing power of graphical processing units (GPU) and tensor processing units, convolutional neural networks (CNNs) (LeCun et al., 1989) have gained widespread popularity in image and video recognition and classification as well as in other problems that benefit from their translation invariance characteristics.

In this paper, a production tomographic solution was obtained from the inferred results of the neural network (NN) shown in

Figure 1. The NN consisted of both CNN and standard dense NN layers and was trained on human-picked first-arrival traveltimes.

The goal of the chosen NN architecture was to create a NN with good inference capabilities on a very small subset of the seismic data (0.19% of the shot gathers) in a nonstationary problem situation. The first breaks are a function of a ground-filtered source Ricker wavelet and the subsurface impedance characteristics around the first reflector. As ground and reflector characteristics change spatially and suffer from low S/N (Figure 2) and a dearth of training samples, the challenge was to create a NN that did not suffer from either high bias or high variance.

The root-mean-square (rms) errors for training/testing data sets from many NN architectures (number of layers and units) were compared, and the chosen CNN architecture performed vastly better (5 ms training/13 ms testing rms error) than the best performing deep/wide NN architecture (30 ms training/51.2 ms testing rms error).

The training process flow shown in Figure 3 was composed of an outlier rejection step to remove spurious human mispicks, an NN training/testing data set generator, and the NN itself. The outlier rejection step required nonlinear traveltime tomography ray-traced traveltimes computed on the set of human-picked first-arrival traveltimes. Those picks were used exclusively for the outlier rejection step and not directly or indirectly fed into the NN to prevent the introduction of a bias.

As the shot gathers' maximum bandwidth did not exceed 100 Hz they were downsampled from a 2 to a 4 ms sampling interval to decrease needed computational and memory resources. The 4 ms-sampled shot gathers and a scalar traveltime pick for each provided gather trace were used to train the NN.

During training, the NN attempts to learn the patterns of the first-break picks, but as those patterns are obscured by noise and may occur sporadically in other locations, the inference process flow shown in Figure 4 relied on an outlier filtering step to map out such occurrences. The quality of the final tomography attested to the soundness of the process.

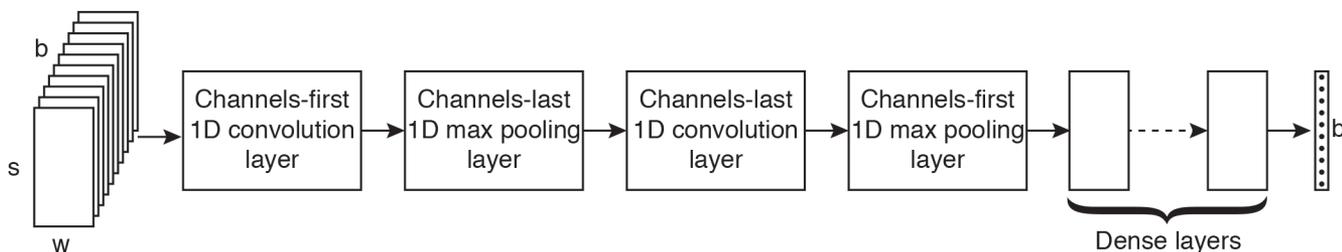


Figure 1. NN architecture. Input is batch (b) matrices of s samples traces grouped in sliding windows of w traces. Output is a vector of b traveltime values.

¹GeoEnergy Inc., Houston, Texas, USA. E-mail: lionel@geoenergycorp.com; anthony@geoenergycorp.com; rod@geoenergycorp.com.

Methods and procedures

Data organization. Seismic shot gathers are composed of sound recordings obtained from a single source and a specific set of receivers. Each source/receiver pair recording yields a data vector (seismic data trace). In land seismic data acquisition, those

vectors often suffer from low S/N. The NN input is derived from shot gather vectors.

For a set of shot gathers, a skilled human operator will pick first arrivals — the NN output, the earliest arrival of source energy to each receiver geophone (as opposed to earlier nonsource energy recorded by said geophones). Such picks contain small errors due to the human mechanics of interpretation and picking. Therefore, the process flow and NN must be able to gracefully handle this input/output uncertainty. We will describe the design, architecture, and flow that allows for such a task to be performed.

Outlier removal. Outlier removal is an important step in both training and inference workflows. In the training workflow, human traveltime picking could mistakenly add spurious misplaced picks in a small number of neighboring traces. In the inference workflow, low-S/N input data can lead to some NN mispicks.

To perform this task, we computed the tomography-based ray-traced traveltimes from the human picks and removed picks that were outside a time window of ± 60 ms around the computed ray-traced traveltime.

Training. Let s be the number of samples in an input trace. Let w be the size in traces of a set of neighboring traces that are (1) sorted by receiver number for a given shot and (2) centered on said input trace. The shot gather was initially padded with $w / 2 - 1$ zero-sampled leading and trailing traces. For each picked travel time value y_i at some trace i , an input matrix X_i was created by collecting the sample values of the w traces centered on trace i .

Shot samples were renormalized by their standard deviation, and picked traveltime values were renormalized by $\Delta t \cdot (s - 1)$, where Δt is the seismic trace sampling interval, thus guaranteeing that the pick values lie between 0 and 1.

Let b_p be the number of human traveltime picks. The previously described outlier removal algorithm was used to discard any $X_i \in \{1 \dots b_p\}$ for which y_i lies beyond the clipping range. The set of b remaining $\{X_i, y_i\}$ pairs was randomly split into training (80%) and testing (20%) subsets.

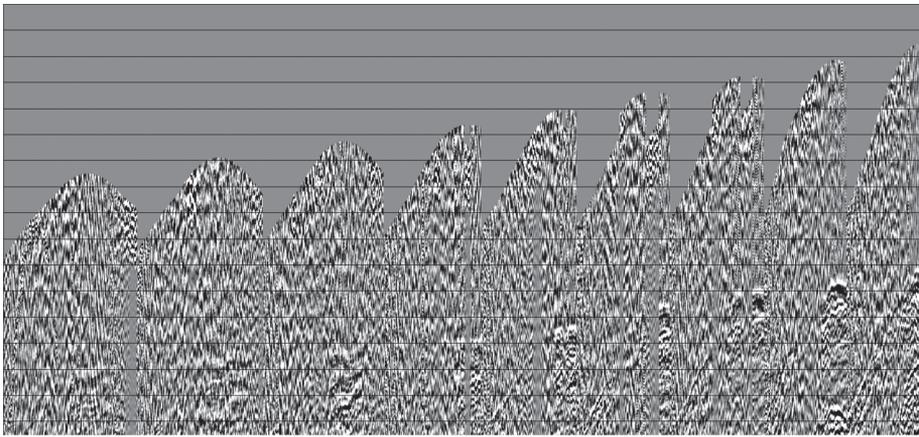


Figure 2. Low S/N gathers from the target data set.

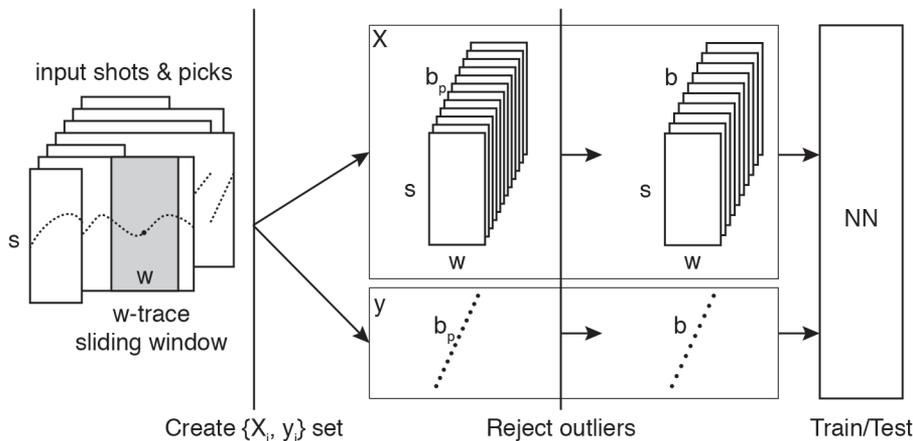


Figure 3. Training/testing flow. Input (X) and output (y) NN training/testing values are extracted from the picked source shot gathers.

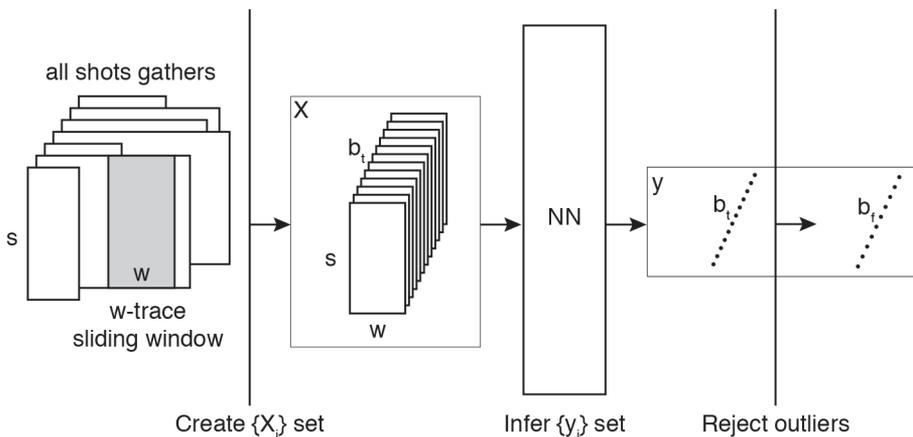


Figure 4. Inference flow. NN inferred values (y) are extracted from the preprocessed complete set of source shot gathers and filtered for outliers.

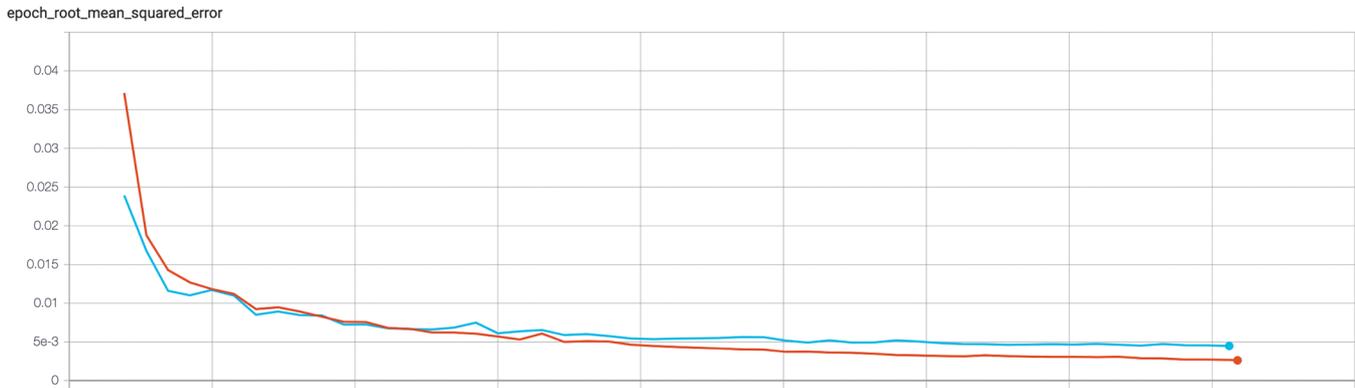


Figure 5. Training (red) and testing (blue) rms errors over training epochs.

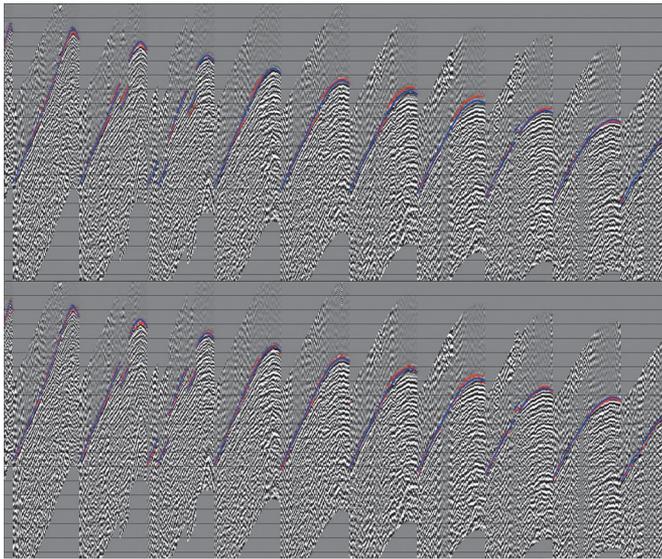


Figure 6. Training (top) and inferred picks (bottom) for a shot gather. The images show the traveltime picks (red) and the ray-traced tomography used for outlier removals (blue).

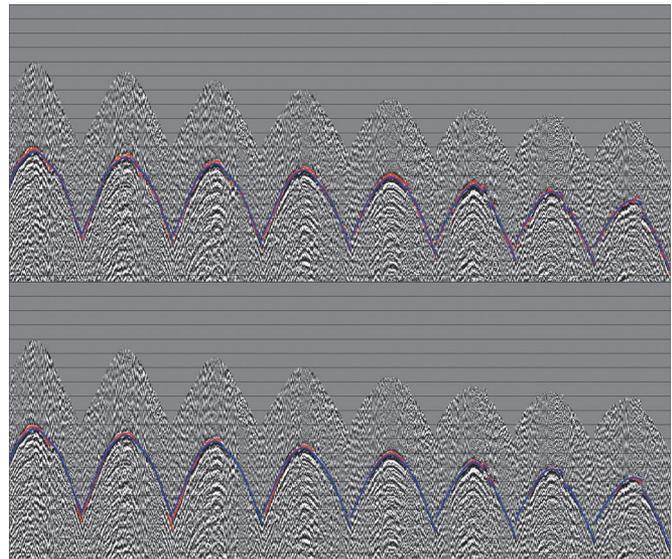


Figure 7. Training (top) and inferred picks (bottom) for a shot gather. The images show the traveltime picks (red) and the ray-traced tomography used for outlier removals (blue).

We used an Adam optimization algorithm (Kingma and Ba, 2015) to train the network described in Figure 1. This algorithm uses both first- and second-order moments and remains invariant to the diagonal rescaling of gradients. The optimizer was set to minimize the rms error cost and use a 0.0001 learning rate. The training converged smoothly, as can be seen in Figure 5.

Examples of results comparing training and NN output picks are shown in Figures 6 and 7.

Training took 7 hours on an NVIDIA Quadro P6000 GPU using the keras API in Google's Tensorflow 2.

Inference. The trained NN was used to process the full set of shot gathers composed of b_i traces using the flow described in Figure 3. Original shot gathers were decomposed into $X_i, i \in \{1 \dots b_i\}$ input matrices, from which the NN inferred $y_i, i \in \{1 \dots b_i\}$ output picks. Those picks were then filtered for outliers with a rejection rate of 6.11% for a final set of b_i usable picks.

Results can be seen in Figure 7; one can recognize the stability of the algorithm in a low-S/N setting.

Inference on the full data set took 3 hours on an NVIDIA Quadro P6000 GPU.

Results

The goal of first-break picking is to create a geologically accurate near-surface velocity field. To that aim, and to quantify and qualify the usefulness of the NN, we computed tomographic solutions for both the set of human picks and the set of NN inferred picks as shown in Figure 8.

The NN results provided a production-quality tomographic solution (shown in Figure 9) of greatly increased resolution and accuracy with 10 hours of GPU compute time substituting for 160 hours of skilled human labor.

Conclusion

We have shown how skilled human labor in the essential process of first-break picking can be reduced by 96.2% when complemented by a supervised NN. These results were reached with real-world land data and no prior knowledge other than that described.

The initial tomography-based outlier rejection step performed extremely well in identifying and removing enough of the NN inference busts to yield a stable final tomographic solution.

While the NN weights obtained from one data set cannot be expected to work on another data set in a nonstationary and

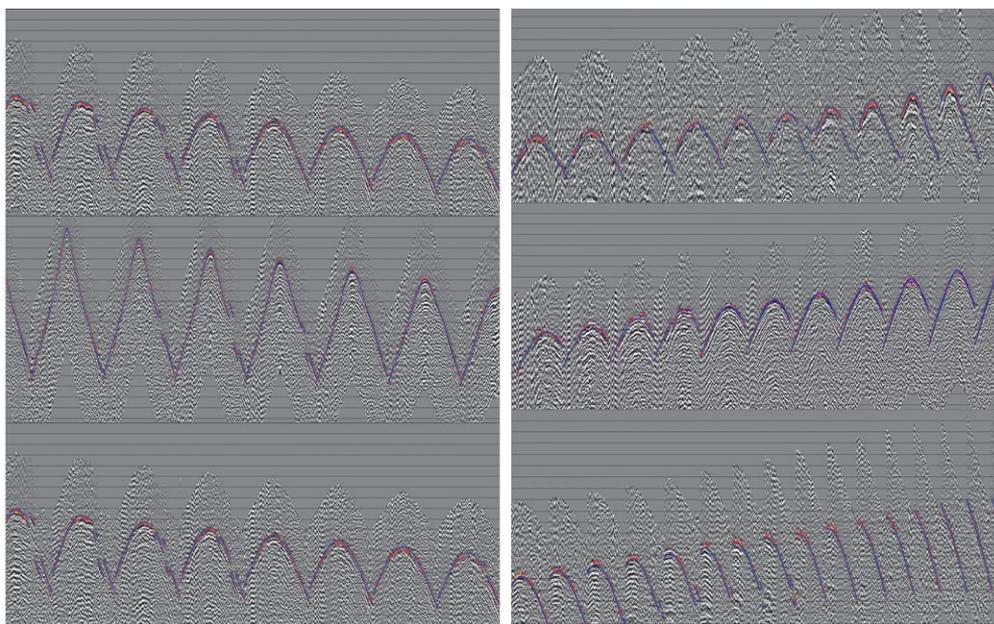


Figure 8. Inferred travel time picks (red) and ray-traced tomography used for outlier removals (blue) for various shot gathers.

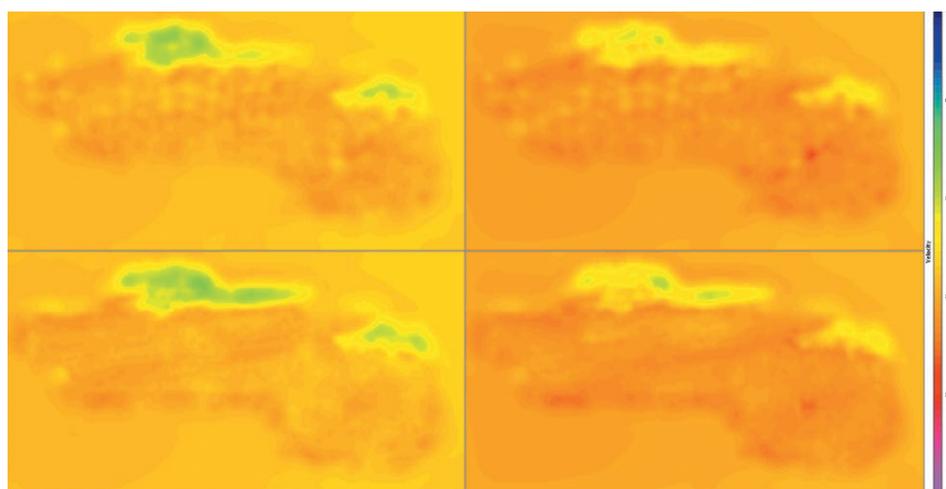


Figure 9. Comparison of two tomography-computed velocity field time slices using human picks (top) or NN picks (bottom)

geology-dependent situation, the described NN architecture did not suffer from high bias or variance and is hence expected to be retrainable unmodified on new data. **TLE**

Data and materials availability

Data associated with this research are confidential and cannot be released.

Corresponding author: lionel@geoenergycorp.com

References

Kingma, D. P., and J. Ba, 2015, Adam: A method for stochastic optimization: Presented at International Conference on Learning Representations.
 LeCun, Y., B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, 1989, Backpropagation applied to handwritten zip code recognition: *Neural Computation*, **1**, no. 4, 541–551, <https://doi.org/10.1162/neco.1989.1.4.541>.
 McCormack, M. D., D. E. Zaucha, and D. W. Dushek, 1993, First-break refraction event picking and seismic data trace editing using

neural networks: *Geophysics*, **58**, no. 1, 67–78, <https://doi.org/10.1190/1.1443352>.
 Murat, M. E., and A. J. Rudman, 1992, Automated first arrival picking: A neural network approach: *Geophysical Prospecting*, **40**, no. 6, 587–604, <https://doi.org/10.1111/j.1365-2478.1992.tb00543.x>.
 Tsai, K. C., W. Hu, X. Wu, J. Chen, and Z. Han, 2020, Automatic first arrival picking via deep learning with human interactive learning: *IEEE Transactions on Geoscience and Remote Sensing*, **58**, no. 2, 1380–1391, <https://doi.org/10.1109/TGRS.2019.2946118>.
 Veezhinathan, J., D. Wagner, and J. Ehlers, 1991, First break picking using a neural network, in F. Aminzadeh and M. Simaan, eds., *Expert systems in exploration: SEG, Geophysical Development Series no. 3*, 179–202, <https://doi.org/10.1190/1.9781560802532.ch8>.
 Yuan, S., J. Liu, S. Wang, T. Wang, and P. Shi, 2018, Seismic waveform classification and first-break picking using convolution neural networks: *IEEE Geoscience Remote Sensing Letters*, **15**, no. 2, 272–276, <http://doi.org/10.1109/LGRS.2017.2785834>.
 Zhe, L., R. Sheng, W. Xu, and J. Yang, 2013, The improvement of neural network cascade-correlation algorithm and its application in picking seismic first break: *Advances in Petroleum Exploration and Development*, **5**, no. 2, 41–51.