7th International Symposium on Three-Dimensional Electromagnetics
Vancouver, BC, Canada, November 13–15, 2023

Reparametrizing the Geophysical Inverse Problem using a Convolutional Neural Network

Anran Xu[1], Lindsey Heagy[1]
[1]Department of Earth, Ocean and Atmospheric Sciences, University of British Columbia

SUMMARY

The recent surge in artificial intelligence has garnered substantial attention among researchers, particularly in the context of incorporating machine learning algorithms into inversion procedures. In the realm of Computer Vision (CV), the Convolutional Neural Network (CNN) architecture has been identified as inherently enforcing prior knowledge, proving advantageous for addressing diverse CV inverse problems, including de-noising and inpainting. This intrinsic regularization effect has shown promise in enhancing models recovered through full waveform inversion of seismic, and it has the potential for application in other geophysical inverse problems. In this study, we examine the applicability to the inversion of DC resistivity data. The CNN maps an arbitrary vector to the model space (e.g., log-conductivity on the simulation mesh). The predicted subsurface model is fed into the SimPEG numerical simulation package to generate corresponding predicted measurements. Subsequently, the data misfit is computed by comparing these predicted measurements with the observed field measurements. This is combined with an L1 smallness term to form the objective function. The backpropagation algorithm is employed to update the trainable parameters of the CNN until convergence. Note that the CNN does not require training prior to the inversion, rather, the CNN weights are estimated in the inversion algorithm. Our preliminary work shows that we can recover models that are comparable to, and even superior to that obtained using a standard inversion. For example, we have found that relying on the implicit regularization of the CNN improves the recovery of the dip of a target when a standard L1 regularization is employed. This method is training-data-free, so it can be adapted to other EM inversion problems.

Keywords: Deep image prior (DIP), Convolutional Neural Network (CNN), Direct Current (DC) inversion

INTRODUCTION

The recent emergence of artificial intelligence has garnered significant attention from researchers, particularly regarding the integration of machine learning algorithms into the inversion algorithm. Researchers in Computer Vision (CV) have discovered that the Neural Network architecture inherently enforces a prior knowledge that is advantageous for addressing diverse CV inverse problems, including de-noising and inpainting (Hattori et al., 2021; Ulyanov et al., 2018). These works show that solving inverse problems in a self-supervised manner is feasible and sometimes has better performance than learning the prior in a supervised manner, which require a large training set, because of the inherent regularization effect in Convolutional Neural Networks (CNN) or Graph Neural Networks (GNN).

In this abstract, we focus on the Direct Current (DC) resistivity survey, which is used for a range of applications including mineral exploration, groundwater studies, and a variety of environmental and geotechnical applications. The goal of the inverse problem is to find a conductivity model of the subsurface that is consistent with the observed data and other prior information or assumptions. There are several avenues through which prior information and assumptions are included in the inversion: the choice reference model, norms on the components of the regularization, and the model parameterization (e.g., using log-conductivity or a parametric model). In this abstract, we explore the use of a CNN to parameterize the model. Our experiments thus far show that we can recover models that are comparable to a standard inversion, and that there may be some advantages in using the CNN. Using an example of a dipping structure, we find that relying on the implicit regularization of the CNN improves the recovery of the dip of a target as compared to standard approaches.

METHODS

In the DC resistivity survey, transmitters on the ground inject a steady state electrical current into the ground and the receivers on the ground observe the resulting distribution of potentials (voltages) on the surface (Fig. 1).

The conventional way of solving this inverse problem is by iteratively updating the conductivity values for each cell in the mesh which discretized the subsurface. At each iteration, we input the current conductivity model into a numerical simulation package (a PDE solver based on Maxwell's equations), whose output would be the predicted measurements. We define the objective function to be the summation of the data misfit term ϕ_d (i.e., the difference between predicted and field measurements) and the regularization term ϕ_m:

$$\min_{\mathrm{m}} \phi_d(m) + \beta\phi_m(m),$$

where m represents the conductivity model (see for example Oldenburg & Li, 2005). In our proposed method, the model, m, is parametrized by a convolutional neural network F_z, where z is the fixed input of the convolutional neural network (i.e., m = F_z(w)), where w is the vector of weights in the CNN. Another update we make is how we "cool" the influence of the regularization. Rather than only reducing the trade-off parameter multiplying the regularization, we simultaneously decrease the trade-off parameter multiplying the regularization and increase the contribution of the data misfit by the same amount. We have found that when using first-order optimization methods employed in deep learning that this leads to better convergence. Thus, the objective function in our proposed method is:

$$\min_{w} \ (1-\beta)\phi_d\big(F_z(w)\big) + \beta|F_z(w) - m_{ref}|^1.$$

Since the inverse problems in geophysics are typically ill-posed, the regularization plays a very important role in inversion.

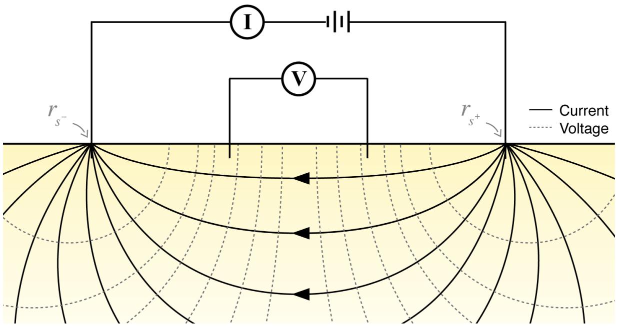

Figure 1: Transmitter and receiver in a DC resistivity survey (Cockett et al., 2016).

To demonstrate, we consider a model that consists of 2 layers with a dike in the second layer, as shown in Figure 2. We consider 3 different dip angles (first, second, and third rows) to test the robustness of the proposed model. We simulate a dipole-dipole survey with 348 data points. We perform several inversions with the conventional methods to serve as the benchmark. The second column used an approximate L0 norm on the smallness and an L1 norm on the smoothness; these inversions were performed without sensitivity weighting. The third column used an L0 norm on the smallness and an L1 norm on the smoothness, and these inversions were performed with sensitivity weightings. Our results using a CNN to parameterize the model are shown in the fourth column. Note that only a smallness term is used in the regularization, and an L1 norm is applied.

All models recover a conductive structure in approximately the correct location, but we see the influence of the first-order smoothness in the standard inversion results, which tends to align structures with the axes along which we are regularizing (horizontal and vertical). The use of the CNN and only the smallness term is better able to recover the dip of the target. Although no explicit smoothness term is used in the regularization, we hypothesize that the CNN provides implicit regularization that promotes recovery of a reasonable target.

DISCUSSION

In this study, we examine the applicability of the inherent regularization effect from the CNN structure to the inversion of DC resistivity data by utilizing trainable weights within the CNN to parameterize the conductivity model. Namely, the CNN maps an arbitrary vector to the mesh space. The predicted subsurface model is then fed into the numerical simulation package to generate corresponding predicted measurements. Subsequently, the objective function value is computed. Compared to the objective function in the conventional methods, the objective function in the proposed methods doesn't have the smoothness term and doesn't use sensitivity weighting. The backpropagation algorithm is employed to update the trainable parameters of the CNN until convergence. Note that the CNN does not require training prior to the inversion, rather, the CNN weights are estimated in the inversion algorithm. Our preliminary work shows that we can recover models that are comparable to, and even superior to that obtained using a standard inversion. For example, we have found that relying on the implicit regularization of the CNN improves the recovery of the dip of a target when a standard L1-smallness regularization is employed (Figure 2). In general, the proposed method can also eliminate the problem that structures are concentrated near the electrodes in the recovered model. This method is training-data-free, so it can be adapted to other EM inversion problems.

CONCLUSION

There are many choices of CNN architecture that can be employed and it's likely that the best choice of network will depend on the nature of the expected model, we conduct inversions for the above examples (3 examples in Figures 2) with the same CNN architecture to illustrate the robustness of the method. The next step of this project would be further exploring this implicit regularization effect by looking at the inversion results using different depths of the CNN architecture, different modes for the up-sampling layers, and different kernel sizes and stride values for the convolutional layers.

ACKNOWLEDGMENTS

We would like to thank Doug Oldenburg, Devin Cowan, Jorge Lopez-Alvis, Joseph Capriotti, John M. Weis, Johnathan C. Kuttai and Santiago Soler for their invaluable guidance and advice on SimPEG usage.

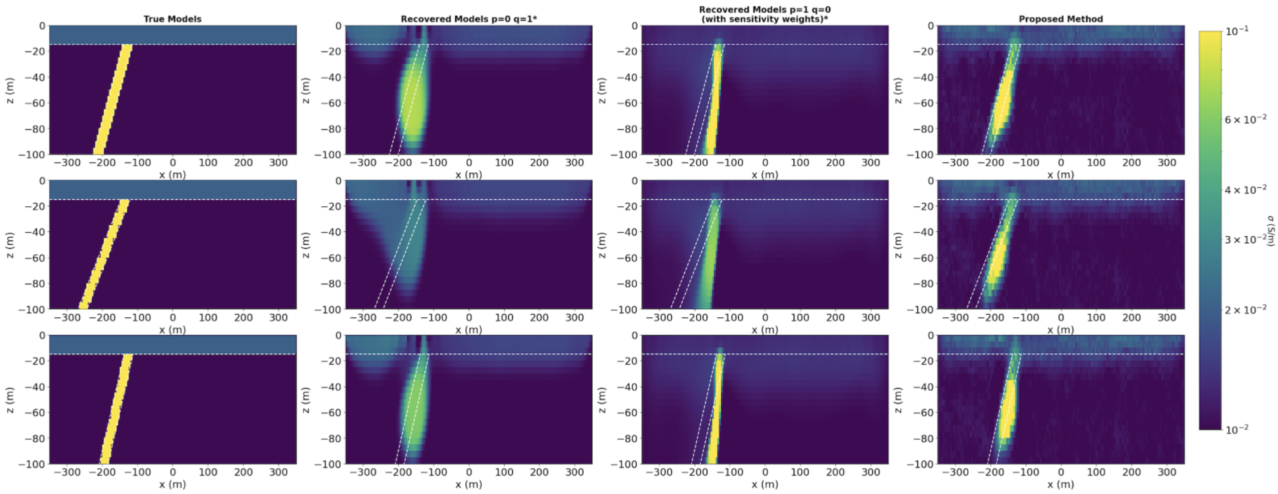

Figure 2: The comparation between the conventional methods (middle two columns) and proposed method (the right most column) in three dike models with the different dip angles. The subplots on the second/third column are from the standard sparse-norms inversions without/with sensitivity weights respectively. The p and q values shown indicate the choice of norm used for the smallness and smoothness terms respectively.

References

Cockett, R., Kang, S., Heagy, L. J., Pidlisecky, A., & Oldenburg, D. W. (2015). SimPEG: An open source framework for simulation and gradient based parameter estimation in geophysical applications. Computers & Geosciences.

Ulyanov, D., Vedaldi, A., and Lempitsky, V., (2018), "Deep Image Prior," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 9446-9454.

Cockett, R., Heagy, L. J., and Oldenburg D. W., (2016). "Pixels and their neighbors: Finite volume." The Leading Edge, 35(8), 703–706.

Zhu, W., Darve, K., Biondi B., and Beroza, G., (2022), "Integrating deep neural networks with full-waveform inversion: Reparameterization, regularization, and uncertainty quantification," *GEOPHYSICS* 87: R93-R109.

Hattori, S., Yatagawa, T., Ohtake, Y., and Suzuki H., (2021), "Deep Mesh Prior: Unsupervised Mesh Restoration using Graph Convolutional Networks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).