Generalized Data-Driven Radio Magnetotelluric Inversion using Convolutional Neural Network

Koustav Ghosal[1], Arun Singh[1] and Deepak Gupta[2]
[1]Department of Applied Geophysics, Indian Institute of Technology (Indian School of Mines), Dhanbad, India
[2]Transmute AI Lab, Indian Institute of Technology (Indian School of Mines), Dhanbad, India

SUMMARY

Machine learning techniques are considered a suitable option for geophysical data inversion. However, machine learning algorithms are very useful for inversion of geophysical data. For a data-driven approach, these algorithms work on the assumption that the statistical properties of the training and testing dataset follow an independent and identical distribution (IID). However, such assumptions are not met in the geophysical field data and the solutions obtained using these algorithms may provide erroneous solutions. We propose a strategy for the inversion of RMT data where the training data was generated for resistivity models based on the Gaussian Random Fields (GRFs). Once the models were created, forward responses in terms of resistivity and phase were generated using a finite-difference based algorithm. A Convolutional neural network was trained in a supervised manner. The trained network was tested for various Out of Distribution (OOD) models including the checkerboard model. The proposed approach overcomes the need to retrain the neural network for different OOD samples. Hence introducing generalization ability in a data-driven supervised learning framework.

Keywords: RadioMagnetotelluric, U-Net, Out-of-Distribution Generalization, Deep Learning, Gaussian Random Field

INTRODUCTION

RMT methods are widely used to investigate the near-surface conductivity variation of the earth. This method has a wide range of applications such as groundwater exploration (Pedersen et al, 2005), waste management (Tezkan et al, 2000), fracture and fault mapping (Linde and Pedersen, 2004; Tezkan et al, 2019), study of porous environment (Turberg et al, 1994) and mineral exploration (Smith, 2014). For estimations of near-surface variations in resistivity from the recorded dataset, the gradient-based optimization methods are the popular choice. However, these methods fail to compute the non-linearity between the model and data due to model regularization. The Deep Learning (DL) based methods provide a suitable alternative to RMT inversion, as it can accurately capture the nonlinear aspects of the inverse problem and provide a rapid resistivity model once it is fully trained. The DL-aided methods for MT inversion have gained significant popularity in recent times (Xie et al, 2023; Liao et al, 2022; Liu et al, 2021). Conventional data-driven neural networks are based on the assumption of IID, which might not come true with field data. Utilizing a pre-trained network directly on field data may lead to inaccurate representations of the resistivity model. Our proposed method is a step toward achieving OOD generalization in data-driven learning. We showed that training neural networks with models generated using GRFs helps achieve OOD generalization. As these resistivity distributions, closely resemble the natural resistivity distribution encountered in nature. In this study, we have used a convolutional neural network to perform the RMT inversion. Through synthetic experiments, we have demonstrated that the predicted resistivity model closely aligns with the true resistivity model, even when considering OOD samples. To validate our proposed method, we have conducted a comparison of the neural network performance on OOD with a gradient-based optimization.

METHODS

Dataset preparation

The diverse set of resistivity models were generated using correlation function.

$$\text{CORR} = \text{EXP}((-1/2)\sum_i((x_1(i)-x_2(i))^2/c_0(i))) \quad (1)$$

where $x_1(i)$ and $x_2(i)$ are the two points in the domain and c_0 is the scaling parameter. A total of 41000 models were generated randomly by varying the c_o with a range of (0.001 to 1.0). The variation in the parameters represents models from single localized anomalies to 4 or more anomalies of random shapes and sizes. These randomly generated GRF resistivity models are then used to create the RMT data which consists of apparent resistivity and phase of TE and TM modes. The forward responses are simulated using a MATLAB-based code (Singh* et al, 2014), employing a finite-difference scheme. RMT data was generated at 13 frequencies logarithmically separated between 1 and 250 KHzover at 21 stations with interstation spacing of 10 m. For TE mode, an 10 air layer was added along with padded grids with increasing spacing was added surrounding the domain which is required to satisfy the boundary conditions. To evaluate the neural network performance over OOD samples we have used checkerboard models with homogeneous backgrounds of 100 Ωm with different geometrically shaped anomalous bodies with resistivity as high as 1000 Ωm and as low as 10Ωm.

Neural Network Architecture and Training

A CNN-based U-Net architecture was chosen as it is efficient for segmentation tasks in image processing. Its U-shaped structure includes a contracting path for feature extraction and an expansive path for precise localization. Skip connections facilitate the flow of detailed spatial information making it efficient to recover small details in the predicted resistivity model. The network comprises convolutional modules with convolutional layers, batch normalization, and ReLU activation. After two such modules and a 2D average pooling layer, input dimensions change from (N, C, H, W) to $(N, 2C, \frac{H}{2}, \frac{W}{2})$. The bottleneck layer has 512 channels and a 16×16 feature map. The Up-convolutional layer (transpose convolution) with a kernel size of 3, stride of 2, padding of 0, and output padding of 1, transforms the feature map from (N, C, H, W) to $(N, \frac{C}{2}, 2H, 2W)$. The loss function used in to trained the network is simple mean square error loss. where, $\rho_a^{i,j}$ and $\rho^{i,j}$ is the predicted and true resistivity respectively at a location i, j in the subsurface. The U-Net model takes input data that includes 4-channel resistive data. The first two channels of the input data contain apparent resistivity for TE and TM mode, while the other two channels have phase data for TE and TM. A single-channel resistive image of the subsurface is generated by the U-Net model. The RMT data and resistivity model were reshaped to 256 $\times$ 256 for a deeper U-Net architecture. The detailed network design is shown in Figure 1

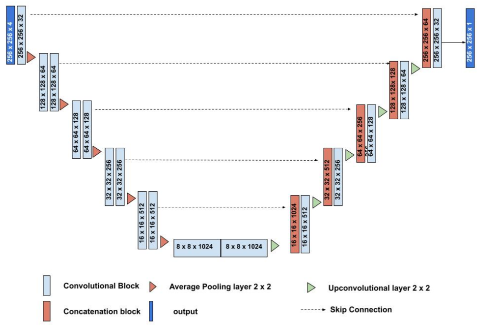

Figure 1: Schematic diagram of U-Net

During the network training phase, we utilized the Adam optimizer with β_1 as 0.9 and β_2 as 0.999. Additionally, a learning rate of 0.0001 was employed, alongside a LambdaLR scheduler which decreases the learning rate by 5% per epoch. We adopted a batch size of 4 and implemented early stopping with the patience of 20 epochs, monitored by the average validation loss and mean absolute error to mitigate overfitting. From the total dataset of 410000, 10% (4100 samples) were allocated for testing, while the remaining 90% were divided between training and validation. Specifically, 76.5% (31,365 samples) are assigned to training, and 13.5% (5535 samples) to validation. The trained weights are used for testing on OOD samples without any further fine-tuning.

RESULTS

The results of both the GRF dataset and OOD samples are shown in Figure 2. The first and second columns represent the true and the predicted resistivity model respectively. The first three rows of

GRF samples include two randomly shaped localized anomalies followed by a model consisting of randomly shaped nearly horizontal anomalies. It is observed that the predicted output of U-Net has a small discrepancy with the true resistivity model. Figure 2(c) shows that reconstruction degrades with higher depth and especially fails to map smaller anomalies at depth. However, large deviations were observed in the case of OOD samples but trained U-Net demonstrated the capability to locate anomalies in close proximity to their original positions, which is evident from figure 2(d) and (e). Two checkerboard models were utilized, and the recovered resistivity model accurately exhibited the position and resistivity of the anomalies. Both checkerboard models are composed of 2 blocks with high resistivity of 10^3 Ωm and and 2 blocks of low resistivity of 10 Ωm. The recovered resistivity model showed 4 anomalies which accurately represent the true resistivity model. However, the models struggled to reconstruct the shapes of anomalies exactly as the true model at depth.

To validate our results, we performed inversion using a gradient-based optimization algorithm WSJointInv2DMT (Amatyakul et al, 2017). We compared the results of both methods for in-distribution as well as OOD samples. In both cases, we observe that the inverse model using U-Net (see Figure 3) is more consistent with the true resistivity model compared to WSJointInv2DMT. In Figure 3(a), the incline conductive body and the resistive body underneath were recovered accurately in the case of U-Net whereas WSJointInv2DMT fails to reconstruct the resistivity body and also fails to map the conductive body at higher depth. In Figure 3(b), U-Net was able to predict four bodies of the checkerboard model whereas WSJointInv2DMT fails to predict the deeper resistive body. The predicted resistivity model not only determined the resistivity values accurately and also the shapes of the rectangular anomalies, but WSJointInv2DMT failed to reconstruct the rectangular anomalies. These experiments showed the efficiency of the proposed method for both in-distribution and OOD scenarios.

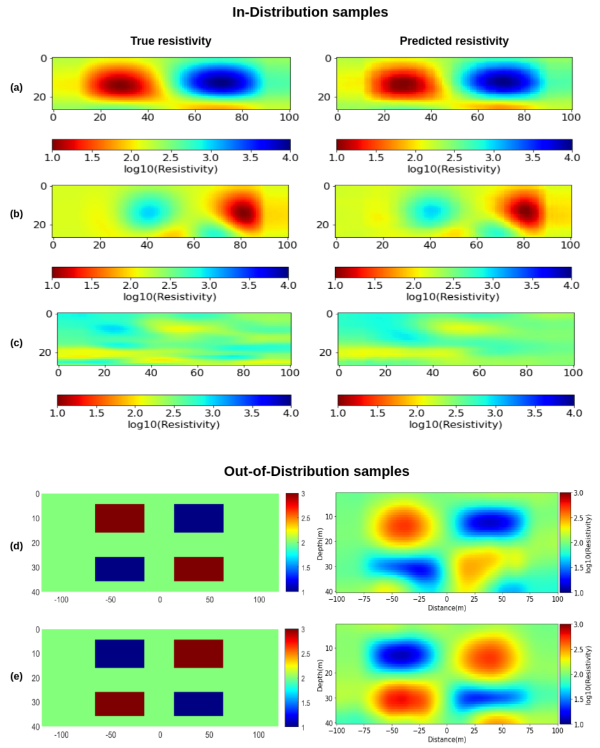

Figure 2: RMT Inversion using U-Net over GRF dataset and OOD samples. The three columns starting from left show the inversion result using U-Net, true resistivity model and misfit between the true and U-Net predictions respectively. The first two rows represent samples from the GRF dataset and the third and fourth represent OOD samples

Conclusions

The assumption that the training and test data have a similar statistical pattern, is invalid when applied to geophysical field data. Consequently, neural networks that are trained in a data-driven manner often fail to yield accurate inversion images for the field data. In this study, we have shown a possible way to enhance the generalization ability of neural networks trained in a data-driven approach. The use of the GRF to generate training dataset induces generalization ability to a CNN-based network like U-Net. Our proposed method was evaluated for various OOD samples, and the predicted solutions accurately captured various anomalies in the true resistivity model. To proposed method, was validated with a traditional gradient-based method, WSJointInv2DMT. Upon comparison, it was concluded that the proposed method yields more consistent results with the true resistivity model. However, the method fails to identify smaller and deeper conductive bodies and also produces many unwanted artefacts in many cases. These issues provide a scope for future development.

Acknowledgments

The authors would like to express their gratitude to the Computer Center(HPC) of the Indian Institute of Technology (Indian School of Mines) Dhanbad for providing the necessary computational resources to conduct this research.

References

Amatyakul P, Vachiratienchai C, Siripunvaraporn W (2017) Wsjointinv2d-mt-dcr: An efficient joint two-dimensional magnetotelluric and direct current resistivity inversion. Computers & Geosciences 102:100–108

Liao X, Shi Z, Zhang Z, Yan Q, Liu P (2022) 2d inversion of magnetotelluric data using deep learning technology. Acta Geophysica 70(3):1047–1060

Linde N, Pedersen LB (2004) Characterization of a fractured granite using radio magnetotelluric (rmt) data. Geophysics 69(5):1155–1165

Liu W, Xi Z, Wang H, Zhang R (2021) Two-dimensional deep learning inversion of magnetotelluric sounding data. Journal of Geophysics and Engineering 18(5):627–641

Pedersen LB, Bastani M, Dynesius L (2005) Groundwater exploration using combined controlled-source and radiomagnetotelluric techniques. Geophysics 70(1):G8–G15

Singh* A, Kharti V, Gupta PK, Israil M (2014) Interpretation of geophysical data using block inversion algorithm: 2d magnetotelluric case. In: SEG Technical Program Expanded Abstracts 2014, Society of Exploration Geophysicists, pp 860–864

Smith R (2014) Electromagnetic induction methods in mining geophysics from 2008 to 2012. Surveys in Geophysics 35:123–156

Tezkan B, Hördt A, Gobashy M (2000) Two-dimensional radiomagnetotelluric investigation of industrial and domestic waste sites in germany. Journal of Applied Geophysics 44(2-3):237–256

Tezkan B, Muttaqien I, Saraev A (2019) Mapping of buried faults using the 2d modelling of far-field controlled source radiomagnetotelluric data. Pure and Applied Geophysics 176:751–766

Turberg P, Müller I, Flury F (1994) Hydrogeological investigation of porous environments by radio magnetotelluric-resistivity (rmt-r 12–240 khz). Journal of Applied Geophysics 31(1-4):133–143

Xie L, Han B, Hu X, Bai N (2023) 2d magnetotelluric inversion based on resnet. Artificial Intelligence in Geosciences 4:119–127

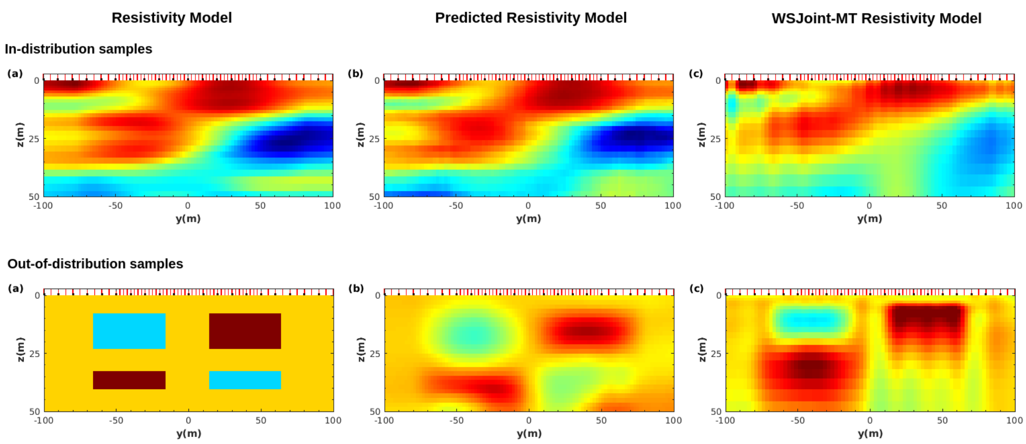

Figure 3: Comparison between the UNet result and the WSJointInv2D-MT-DCR for both in- and out-of-distribution samples