Toward generalized models for machine-learning-assisted salt interpretation in the Gulf of Mexico

Cable Warren[1], Sribharath Kainkaryam[1], Ben Lasscock[1], Altay Sansal[1], Sanath Govindarajan[1], and Alejandro Valenciano[1]

https://doi.org/10.1190/tle42060390.1

Abstract

Interpreting salt bodies in the Gulf of Mexico (GoM) can be complex due to various factors affecting the accuracy of automated techniques. Variability of salt structures, seismic acquisition parameters, and imaging algorithms can impact the resulting seismic image. These differences can result in variations in seismic resolution and texture, making it challenging to develop automated interpretation techniques that are accurate and reliable for identifying salt bodies in the GoM. However, using seismic images with similar acquisition parameters and processing methods minimizes these differences and makes machine-learning (ML) models applicable. Utilizing nine data sets from the eastern GoM, a nine-fold cross-validation technique was applied to measure the generalization performance of the ML model. This method involves using one data set as the test set and the remaining eight for training and repeating the process for all subsets. We further applied an ensemble of the nine models to predict salt on a new unseen survey in Green Canyon. The study aimed to illustrate how salt variability and morphology in the GoM can impact the ability of the ML algorithm to predict salt bodies on unseen data.

Introduction

Exploration for oil and gas in the Gulf of Mexico (GoM) has led to the development of innovative seismic acquisition, processing, and interpretation techniques. One of the main factors driving technological advancements is the region's varied and locally complex nature of the salt structure. Salt bodies in the eastern GoM have complex shapes, roughness, and varying depths due to a combination of factors related to their formation, tectonic setting, and subsequent evolution (Hudec et al., 2013). They typically are formed by the dissolution and evacuation of underlying salt layers. Common features include salt diapirs, walls, domes, plugs, and sheets.

Refining the salt geometry in seismic processing projects is crucial to create a subsalt image. New data acquisition and processing technologies such as ocean-bottom nodes (OBNs), low-frequency sources, and full-waveform inversion make salt model building more data driven (Huang et al., 2020; Mao et al., 2020). When data cannot support it, accurately delineating salt bodies in production workflows requires considerable time and specialized geologic knowledge from interpreters. Due to project constraints, interpretation geologists typically pick salt boundaries on decimated grids. When the image is of good quality, interpretation can be extended using 2D and 3D horizon autopickers. In contrast, interpolation and geologic knowledge are required to fill the gaps in poorly imaged areas. The industry has been experimenting with machine-learning (ML) algorithms that promise automation of salt interpretation and reduced turnaround time for processing projects, but ML algorithms still need improvements to replicate how humans interpret complex salt bodies.

Several elements must be improved for ML algorithms to be widely adopted in production projects. The usual local/interactive approach iteratively trains the ML models on subsets of a specific survey, with active feedback from an interpreter. While this may accelerate the interpretative task, it requires significant manual interpretation and cannot provide a fully automated workflow.

More than 20,000 mi^2 of training data were used in our experiments. This seismic coverage provides the diversity of salt structures required to train a generalizable ML model. Months of personnel investment in data collection, data engineering, and building ML infrastructure made it possible to train the models described in this study. After building a generalizable model, minimal effort should be required to update and improve as new data are incorporated.

Can the complexity of salt allow ML salt segmentation to be generalizable? Using nine data sets with similar processing, vintage, and acquisition geometry, we minimized the irregularity imposed by the variables to create an explicit ML model for eastern GoM salt interpretation. We do not intend to answer whether an ML model can be generalized to the entire GoM. Rather, we want to isolate and analyze how salt bodies and their morphology could affect model generalization.

Data overview

Three-dimensional seismic data from different vintages and acquisition geometries (narrow-azimuth streamers, wide-azimuth streamers, full-azimuth streamers, ocean-bottom cables, and OBNs) cover the GoM. We chose a subset of the data sets in the eastern GoM to sample data with salt bodies. The acquisition types used in this study are narrow-azimuth towed streamers, except for survey 8, which was acquired with ocean-bottom cables. The final processing product includes reverse time migration in a common frequency band (up to 35 Hz), with data binned to common inline, crossline, and depth increments (20, 25, and 10 m).

The labels for image segmentation are represented as a salt mask (the same size as the seismic image). Each pixel is assigned a value, indicating whether that pixel corresponds to salt. Creating these salt labels for supervised learning is a challenging task. Nevertheless, suitable-quality labels (human interpretation) can be obtained with data warehousing and lineage tracking infrastructure. Figure 1 shows the location of the nine data sets used

[1]TGS, Houston, Texas, USA. E-mail: cable.warren@tgs.com; sribharath.kainkaryam@gmail.com; ben.lasscock@tgs.com; altay.sansal@tgs.com; sanath.govindarajan@tgs.com; alejandro.valenciano@tgs.com.

in this study. A subset of each seismic survey has salt delineated by interpreters. These interpreter-picked horizons were converted into a binary mask, indicating the presence of salt in the volume. Figure 2 displays seismic images with superimposed labels from surveys 6 and 3. Note the difference between the morphology of the salt bodies from the two surveys. Figure 2a shows the mother or autochthonous salt, the Louann, which is the salt source for diapirism in the northern GoM (Andrews, 1960).

Software and hardware infrastructure

Training a deep-learning (DL) model with hundreds of thousands of training labels could present a significant computational challenge and cost. Our training and inference pipelines are designed to enable large-scale experimentation by leveraging scalability in both on-premises data centers and cloud resources. We store the seismic surveys in MDIO format (Sansal and Kainkaryam, 2022), a seismic format built using the open-source data storage library Zarr (Miles et al., 2020). MDIO is a chunked compressed file format for storing and manipulating multidimensional energy resource assessment data. MDIO simplifies access to seismic data sets from on-premises and cloud object stores, making the process more efficient.

We also developed a distributed model training library, built on MDIO and PyTorch (Paszke et al., 2019), to perform data augmentations and to iteratively load image and label pairs for training ML models and inference. Our training library ensures that data loading is not a bottleneck for GPU utilization, which is critical for scaling DL models. It also uses optimization techniques for training efficiencies, such as mixed precision training (Micikevicius et al., 2017) and gradient accumulation. Vertex AI, an ML platform on Google Cloud Platform (Google, 2023), enabled distributed training on multiple GPUs. By utilizing these tools in the cloud, we can efficiently train many ML models on large seismic data sets. All of our experiments ran on NVIDIA V100 GPUs with 5120 CUDA Cores, 640 Tensor Cores, 16 GB HBM2 memory, and a memory bandwidth of up to 900 GB/s. We utilize Tensor Cores using automatic mixed precision training.

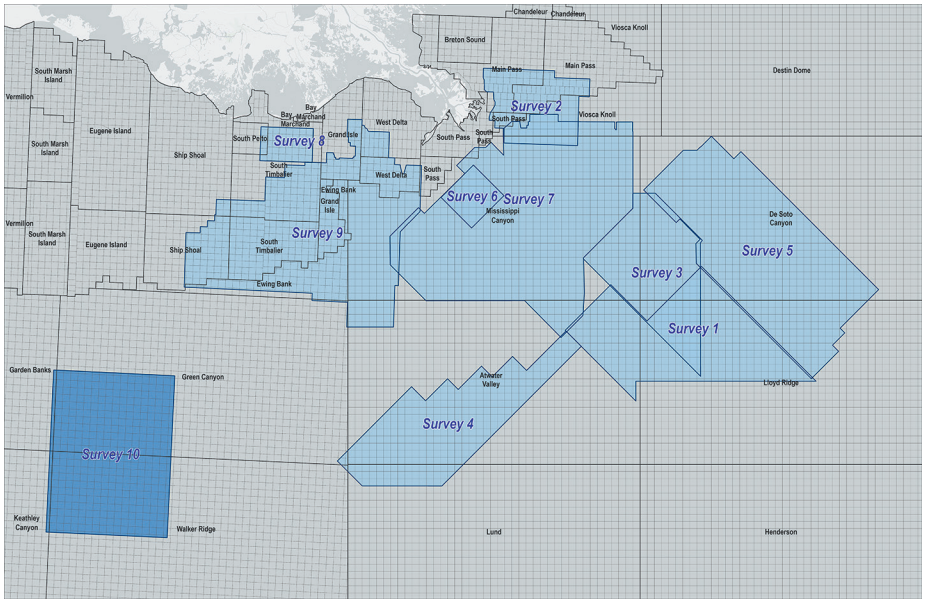

Figure 1. Location of the nine surveys from the eastern GoM used for training in the study. Survey 10 is a test survey that was not used in training.

Model architecture

The experiments described in this paper were run with LinkNet34 (Chaurasia and Culurciello, 2017). LinkNet34 is an efficient convolutional neural network architecture for image segmentation tasks. It is built on the ResNet34 (He et al., 2016) backbone but features unique modifications such as 1 × 1 convolutions and upsampling layers that enhance resolution with fewer parameters. This also makes training and inference cost effective (Chaurasia and Culurciello, 2017).

We use an AdamW optimizer (Loshchilov and Hutter, 2017) with an initial learning rate of 0.001 and a learning rate scheduler that decays for every epoch. To optimize over the nondifferentiable quality metric, intersection over union (IoU), we utilize Lovasz loss (Berman et al., 2018) as a

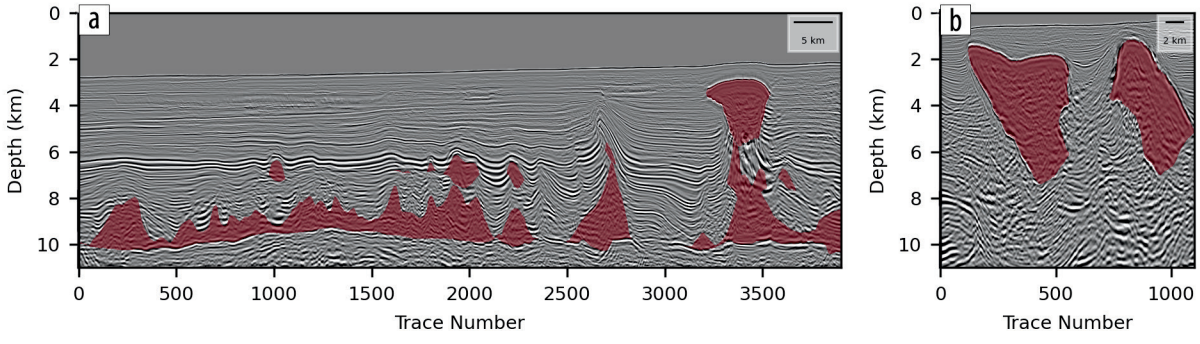

Figure 2. Examples of different salt morphology used in training from two surveys in a deeper to shallow-water setting. Human interpretation in red is superimposed on the seismic image. (a) Survey 3 exhibits complex structural salt displaying deformation and diapirism. (b) Survey 6 only shows shallow canopy salt.

surrogate for IoU. We incorporate various image augmentation techniques in our data preprocessing pipeline including horizontal flipping, small and geologically sensible affine transformations, blurring, and multiscale Gaussian noise addition (Buslaev et al., 2018). We prevent the introduction of artifacts at the edges of the augmentations by sampling larger patches than required for training. After applying augmentations, we crop the image back to the appropriate size expected by the model. Although we ran experiments with other model architectures, we only report results from LinkNet34 in this paper to illustrate that we can obtain generalized salt models with simple architectures.

Methodology

Using nine seismic data sets (Figure 1), we conduct nine-fold cross validation to measure the generalization capability of our salt interpretation model. This method groups eight surveys into a training data set, holding out a single data set (the holdout survey) to test the prediction. We then train a salt interpretation model using the software infrastructure described earlier. Each survey participates as the holdout survey, training from scratch each time to avoid data leakage. The performance quantified by the IoU is evaluated on its corresponding holdout survey.

The IoU metric is employed in computer vision (Everingham et al., 2010) to assess the quality of predictions by computing the extent of overlap between the predicted segmentation mask and the ground-truth mask. In our case, results generated by the ML model are compared to human interpretation. The score ranges between 0 and 1, with 0 indicating no overlap and 1 indicating perfect alignment. Although the IoU metric offers a valuable indication of model performance, the accuracy of the human interpretation benchmark may fluctuate due to data quality, time constraints, and resource availability. A strategy such as label smoothing could alleviate such uncertainty in the labeling task (Müller et al., 2019).

Results in the eastern GoM

Table 1 shows the results of our nine-fold cross validation, including IoU scores for each holdout survey. Seven holdout data sets achieved IoU scores of 0.76 or higher, while the worst performing surveys (surveys 1 and 5) scored 0.26 or lower. Notably, these two surveys are at the eastern end of our study area (as shown in Figure 1), which has different salt morphology and rock properties than the other surveys.

Figure 3 shows a summary of the results of our experiment for a subset of seismic lines. Figures 3b, 3d, 3f, 3h, 3j, and 3l demonstrate good agreement (IoU scores of 0.76 or higher) between the predicted salt bodies and the human interpretation for surveys 3, 9, 7, 6, 4, and 8, respectively. Figure 3a shows the model's accurate prediction of salt (right side), which was not in the labels in survey 3. This survey contains more autochthonous salt with a lower impedance contrast than surveys 6 and 9. The model predicted the base salt well, which is more homogenous but misses a small salt body (middle). This may be explained by the lack of a continuous reflection top salt at that location. Figure 3c shows several salt diapirs that vary in shape and size but have well-defined tops and, in most cases, bases. The model also did well interpreting the salt flanks with no clear boundary, with an overall IoU of 0.921. Figure 3e shows several large salt structures. The mother salt was not accurately imaged; the model suggests adding more salt to the human interpretation. This could provide interpreters with additional scenarios. Lastly, Figure 3g displays two shallow and well-imaged salt diapirs. The model did exceptionally well, with an IoU of 0.971 for the entire survey listed in Table 1, delineating the salt boundary and reasonably approximating where the base could have been better imaged.

Survey 4 in Figure 3i showcases the outboard extent of our nine surveys. The salt body is structurally deformed and has limited diapirism compared to other surveys. Although inclusions are present, it contains vast amounts of well-imaged salt, and the model predicted it very well, with a survey IoU of 0.822. Figure 3k overall has good salt predictions (0.811 for the entire survey) but misses the thin salt in the leftmost structure. It does, however, pick up some of the salt remaining in the weld between the main two bodies, while the prediction on the rightmost salt body closely matches the human interpretation. These results demonstrate the model's ability to provide a reliable automatic salt interpretation, which could be valuable for interpreters.

Figure 4 shows the results where the models did not perform as expected. We observed mixed results. It correctly predicts the shallow and deeper salt, as seen in Figure 4a but not in 4c, although these lines are from the same survey (survey 2). One potential reason for the misprediction in Figure 4d but not in 4b is the low reflectivity of the top salt when compared to Figure 4a. The model predictions failed on survey 5 and need better performance on survey 1. These surveys are from DeSoto Canyon and Lloyd Ridge, located on the eastern portion of the nine surveys. This area has more prominent carbonates than other regions of the eastern GoM. Carbonates have similar velocities to salt, producing equivalent impedance contrasts with sediments. Survey 5 also has thin salt that was not present in the other surveys used for training.

Figure 4e displays a line from survey 1 with a sharp impedance contrast at a depth of approximately 6 km. The top salts near this boundary are predicted, but the deeper salt is not. One possible reason for not predicting deep salt is the abundance of deep reflectivity in this survey, which was absent in the eight surveys used for training. We also observed scaling differences in the data with increasing depth that could affect the model's ability to discern salt. These differences in surveys 1 and 5 highlight the importance of capturing the data variety for training a production-scale model. This can lead to better generalization performance.

Applying ensemble learning to predict salt bodies for unseen data

To better understand the generalization capability of our nine models, we tested each of them on a new data set not seen during any stage of the earlier discussion (see Figure 1, survey 10). The salt in the region covered by survey 10 (Green Canyon/Walker Ridge, southwest of the study area) is more complicated, with multiple salt layers interspersed with other materials.

Ensemble learning is a common technique to increase the predictive power of DL models (Wang et al., 2020). An ensemble prediction for survey 10 was generated by averaging salt

Table 1. The IoU scores for each model predicted on their respective holdout surveys. These scores are calculated based on thresholded binarization of the probability volumes after inference. The thresholds are optimized to maximize IoU on the holdout data and are typically 0.4 across all surveys.

Survey	Survey 1	Survey 2	Survey 3	Survey 4	Survey 5	Survey 6	Survey 7	Survey 8	Survey 9
IoU	0.257	0.764	0.779	0.822	0.010	0.971	0.861	0.811	0.921

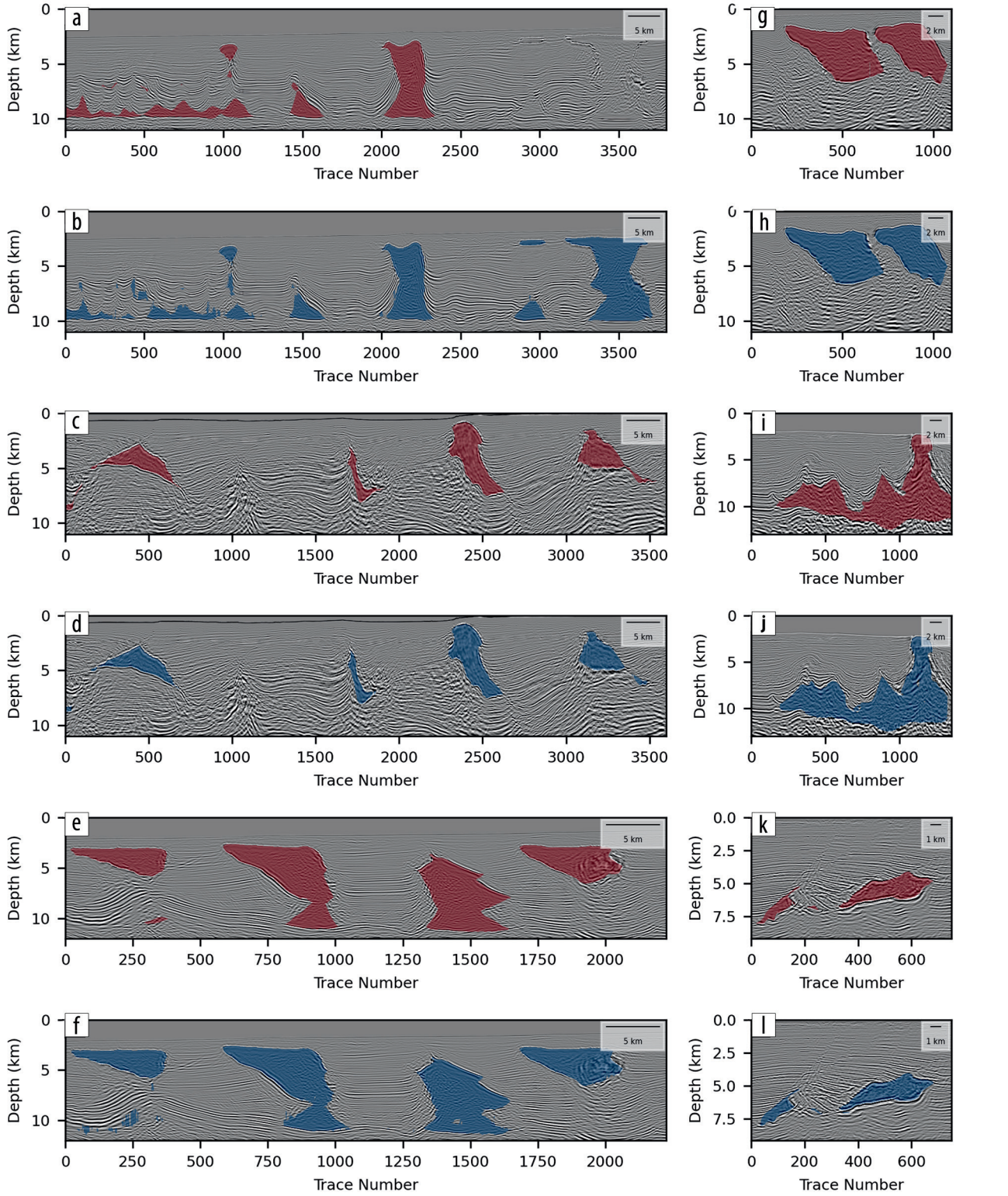

Figure 3. Comparison of labels (red) versus inference (blue) for six surveys (3, 9, 7, 6, 4, and 8) after the nine-fold cross-validation study. In general, the prediction in the holdout's test sets was good. The salt geometries vary dramatically between surveys, ranging from simple well-imaged canopy salt to areas with complex feeder structures, overhangs, and inclusions.

probabilities from nine models. We chose a probability threshold based on our nine-fold analysis to compute the composite salt label mask. The composite threshold was selected by taking the optimized thresholds' median value from each cross-validation fold's holdout data set. Figure 5 compares human interpretation and ML predictions for this survey. The prediction has an exceptionally low rate of false positives in the canopy salt but misses some of the deeper feeder structures. The aggregate model predicted the salt base similar to the interpreter, which was not well defined in the salt body near trace 600. This analysis shows that the ML salt model is close to generalizing throughout the eastern GoM.

Discussion

This paper presents research on building regional-scale models for interpreting salt structure in the eastern GoM. Our goal was to show that a deep neural network trained on multiple surveys has the potential to generalize to unseen surveys. The results are a step toward generalization across the region. The IoU evaluation metric gave us a good idea of how the models perform at this task. However, the IoU score can be affected by the accuracy of the interpretation labels. The labels' accuracy varied between surveys due to data quality, time constraints, and resource availability.

Hyperparameter tuning, modifications to the loss function, and more training data will bring us closer to a generalized model. Although, not explored in this study, additional augmentation techniques, such as AugMix by Hendrycks et al. (2019) and RandAugment by Cubuk et al. (2020), could help improve the accuracy of the salt interpretation. Further, self-supervised representation learning using masked autoencoders (He et al., 2022), contrastive learning (Chen et al., 2020), and knowledge distillation (Hinton et al., 2015) can be used to improve the accuracy of automated salt interpretation.

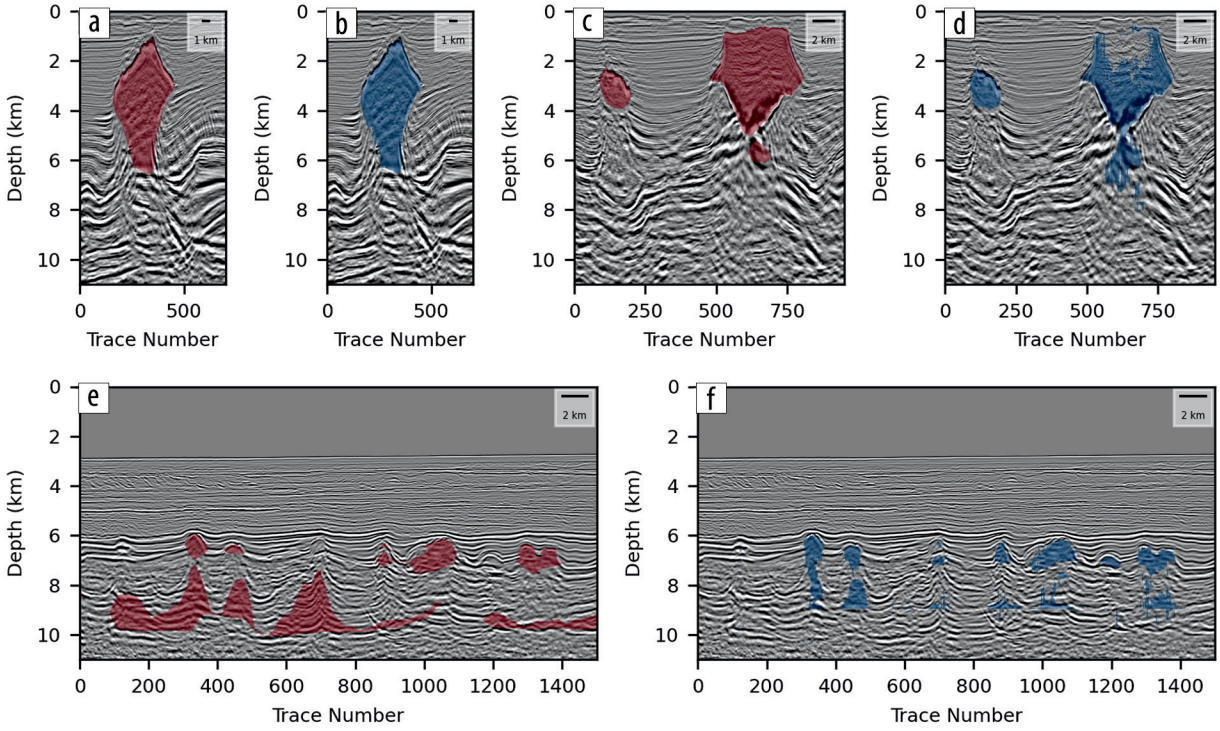

Figure 4. Comparison of labels (red) versus inference (blue) for two surveys after the nine-fold cross-validation study. (a and c) Two lines from survey 2 that highlight the differences in pick quality by the ML model within the same survey. (e) A line from survey 1 highlights the geologic differences within the survey.

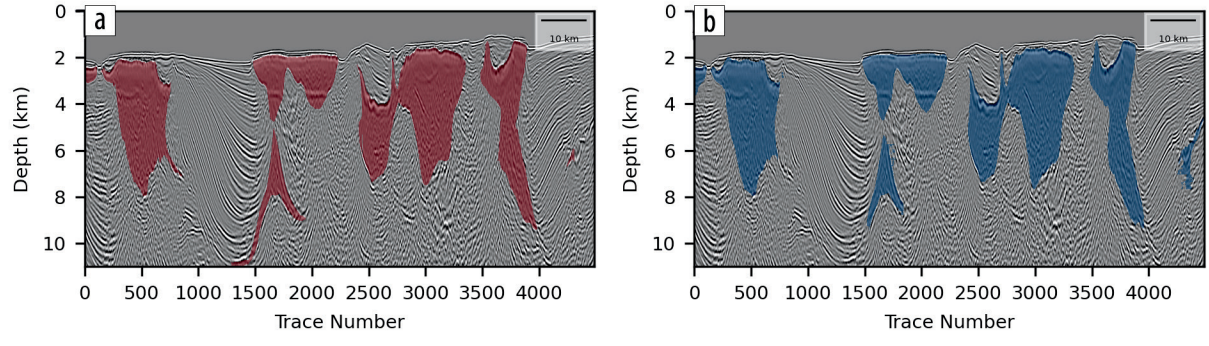

Figure 5. Salt interpretation on a line from survey 10 in Green Canyon. Red is the human interpretation. Blue is the predicted salt made by the ensemble of the nine models.

Conclusions

The complexity of salt bodies in the GoM can make it challenging to automate interpretation techniques. This study used a *k*-fold cross-validation technique to develop nine ML models for salt interpretation, using data sets from the eastern GoM with similar processing, vintage, and acquisition geometries. These models can predict salt with an IoU score of 0.76 or higher on most holdout data sets. Building a pretrained model for inference in new data sets can improve organizational interpretation efficiency. We show that with an extensive set of training data and labels, data engineering, and distributed computing infrastructure, an ensemble of these models can do zero-shot prediction of salt on a tenth data set outside the study area, yielding promising results toward generalization. TLE

Acknowledgments

We thank TGS for allowing the publication of this paper. We thank David Hajovsky for making the GoM data used in the study available.

Data and materials availability

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

Corresponding author: alejandro.valenciano@tgs.com

References

Andrews, D. I., 1960, The Louann salt and its relationship to gulf coast salt domes: Gulf Coast Association of Geological Societies Transactions, **10**, 215–240.

Berman, M., A. Rannen, and M. Blaschko, 2018, The Lovasz-Softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks: Conference on Computer Vision and Pattern Recognition, IEEE, Extended Abstracts, https://10.1109/CVPR.2018.00464.

Buslaev, A., A. Parinov, E. Khvedchenya, V. I. Iglovikov, and A. A. Kalinin, 2018, Albumentations: Fast and flexible image augmentations: arXiv:1809.06839.

Chaurasia, A., and E. Culurciello, 2017, LinkNet: Exploiting encoder representations for efficient semantic segmentation: Visual Communications and Image Processing, IEEE, Extended Abstracts, https://doi.org/10.1109/VCIP.2017.8305148.

Chen, T., S. Kornblith, M. Norouzi, and G. Hinton, 2020, A simple framework for contrastive learning of visual representations: Presented at the 37th International Conference on Machine Learning.

Cubuk, E. D., B. Zoph, J. Shlens, and Q. V. Le, 2020, RandAugment: Practical automated data augmentation with a reduced search space: Conference on Computer Vision and Pattern Recognition, IEEE, Extended Abstracts, https://doi.org/10.48550/arXiv.1909.13719.

Everingham, M., L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, 2010, The PASCAL Visual Object Classes (VOC) challenge: International Journal of Computer Vision, **88**, 303–338.

Google, 2023, Vertex AI documentation: https://cloud.google.com/vertex-ai/docs, accessed 30 January 2023.

He, K., X. Zhang, S. Ren, and J. Sun, 2016, Deep residual learning for image recognition: Conference on Computer Vision and Pattern Recognition, IEEE, Extended Abstracts, https://doi.org/10.1109/CVPR.2016.90.

He, K., X. Chen, S. Xie, Y. Li, P. Dollár, and R. Girshick, 2022, Masked autoencoders are scalable vision learners: Conference on Computer Vision and Pattern Recognition, IEEE, Extended Abstracts.

Hendrycks, D., N. Mu, E. D. Cubuk, B. Zoph, J. Gilmer, and B. Lakshminarayanan, 2019, AugMix: A simple data processing method to improve robustness and uncertainty: arXiv:1912.02781.

Hinton, G., O. Vinyals, and J. Dean, 2015, Distilling the knowledge in a neural network: arXiv:1503.02531.

Huang, Y., J. Mao, H. Xing, and C. Chiang, 2020, Noise strikes, but signal wins in full waveform inversion: 90th Annual International Meeting, SEG, Expanded Abstracts, 805–809, https://doi.org/10.1190/segam2020-3427856.1.

Hudec, M. R., M. P. A. Jackson, and F. J. Peel, 2013, Influence of deep Louann structure on the evolution of the northern Gulf of Mexico: AAPG Bulletin, **97**, no. 10, 1711–1735, https://doi.org/10.1306/04011312074.

Loshchilov, I., and F. Hutter, 2017, Decoupled weight decay regularization: arXiv:1711.05101.

Mao, J., J. Sheng, Y. Huang, F. Hao, and F. Liu, 2020, Multi-channel dynamic matching full-waveform inversion: 90th Annual International Meeting, SEG, Expanded Abstracts, 666–670, https://doi.org/10.1190/segam2020-3427610.1.

Micikevicius, P., S. Narang, J. Alben, G. Diamos, E. Elsen, D. Garcia, B. Ginsburg, et al., 2017, Mixed precision training: arXiv:1710.03740.

Miles, A., J. Kirkham, M. Durant, J. Bourbeau, T. Onalan, J. Hamman, Z. Patel, et al., 2020, Zarr-developers/zarr-python: v2.4.0: Zenodo, https://doi.org/10.5281/zenodo.3773450.

Müller, R., S. Kornblith, and G. E. Hinton, 2019, When does label smoothing help?: Presented at the Advances in Neural Information Processing Systems Annual Conference.

Paszke, A., S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, et al., 2019, PyTorch: An imperative style, high-performance deep learning library: Presented at the Advances in Neural Information Processing Systems Annual Conference.

Sansal, A., and S. Kainkaryam, 2022, MDIO, https://github.com/TGSAI/mdio-python, accessed 30 January 2023.

Wang, X., D. Kondratyuk, E. Christiansen, K. M. Kitani, Y. Alon, and E. Eban, 2020, Wisdom of committees: An overlooked approach to faster and more accurate models: arXiv:2012.01988.

DOI:10.1190/tle42060390.1