Implementation of Denoising Diffusion Probability Model for Seismic Interpretation

Fan Jiang, Konstantin Osypov, and Julianna Toms, Halliburton Landmark*

Summary

In this abstract, we show a novel machine learning-based diffusion model for seismic interpretation. In geophysics, reconstructing the subsurface structure from seismic data is an important inverse problem. Existing supervised machine learning (ML) solutions are to train a model to directly map measurements to seismic images, which are synthesized from images using a fixed velocity model. In this scenario, the generalization capability of models to the unknown measurement process could be hindered and out-of-distribution data could significantly reduce the inference accuracy from the pre-trained model. To address this issue, we implement the diffusion model, as a generative model, for the inverse interpretation problem and it provides a nature way to quantify uncertainty.

Introduction

A generative model becomes more important in exploration since it helps naturally quantify uncertainty. A generative model is designed to learn and capture the underlying probability distribution of the subsurface data (Lamb, 2021). It aims to generate new samples that are similar to the training data it was trained on. In other words, it learns the joint probability distribution of the input features and the corresponding labels or classes, such as Variational Autoencoders (Kingma and Welling, 2013), Generative Adversarial Networks (Goodfellow, 2016), or Diffusion Model (Sohl-Dickstein et al., 2015, Ho et al., 2020).

The advantage of generative models compared to other supervised machine learning models, such as discriminative models, is that they can generate new samples that resemble the training data. This capability allows generative models, especially diffusion models to be used for various tasks, including image and text generation (Rombach et al., 2022), anomaly detection (Wolleb et al., 2022), and super-resolution (Sahak et al., 2023).

Denosing Diffusion Probability Model (DDPM) was first introduced by Ho et al. (2020). DDPM builds upon the concept of diffusion models to learn the distribution of data by modeling the dynamics of a stochastic diffusion process. In traditional diffusion models, a sequence of transformations is applied to the noise vector, gradually refining it to generate realistic samples. However, these models often suffer from a trade-off between the quality of generated samples and the computational efficiency of training and inference. DDPM addresses this challenge by introducing a denoising process during training. Instead of directly modeling the diffusion process, it formulates the problem as a denoising task. The model is trained to remove noise from corrupted images, where the noise is added through a diffusion process. This denoising objective helps in learning a better representation of the data distribution.

In this paper, we implement DDPM for seismic interpretation. We use synthetic fault and salt seismic data as training data to train two separate DDPMs. We demonstrate that DDPM could be considered a posterior uncertainty quantification tool to quantify the data distribution in seismic interpretation.

Methods

The training procedure of DDPM involves alternating between denoising and diffusion steps. In the denoising step, the model is trained to remove noise from corrupted images using an autoregressive network architecture. In the diffusion step, the model is used to generate noisy samples by applying the inverse transformations to the denoised images. The noise added in the diffusion step is annealed gradually, allowing the model to capture the data distribution at different levels of noise.

By combining denoising and diffusion, DDPM achieves high-quality image generation with improved training efficiency. It has been shown to generate realistic images with sharp details and capture complex image distributions. DDPM has also been applied to various image editing tasks, such as inpainting and super-resolution, demonstrating its versatility and effectiveness.

Since its introduction, DDPM has gained attention in the deep learning and generative modeling communities. Researchers continue to explore and extend the capabilities of DDPM and its variants, contributing to advancements in image generation, denoising, and other related areas.

Understand data distribution from prior is a crucial step to build a generalized ML model. However, the probability distribution of data is unknown in most circumstances. Assume each data point x is independent and sampling from this unknown distribution $q(X)$, how can we build a model which could generate similar samples X without knowing $q(X)$? As a generative model, the diffusion model helps to construct model $p_n(X)$, where n is a learning neural network to learn $p_n(X)$ to $q(X)$. Therefore we can sampling the data from $p_n(X)$. Wolleb et al. (2022) present a weakly supervised anomaly detection method based on denoising diffusion implicit models. The forward process to encode an image x_0 into a noisy image x_T by:

10.1190/image2023-3907375.1

$$x_{t+1} = x_T + \sqrt{\bar{\alpha}_{t+1}}\left[\left(\sqrt{\frac{1}{\bar{\alpha}_t}} - \sqrt{\frac{1}{\bar{\alpha}_{t+1}}}\right)x_t + \left(\sqrt{\frac{1}{\bar{\alpha}_{t+1}} - 1} - \sqrt{\frac{1}{\bar{\alpha}_t} - 1}\right)\epsilon_\theta(x_t, t)\right]$$

Where $t \in \{0, \dots, T-1\}$, ϵ_θ is the Unet model. $\bar{\alpha}_t$ is arithmetic multiplication of variances from $t \in \{0, \dots, T\}$ (Ho. et al., 2020).

Then a reverse process can be described as:

$$x_{t-1} = \sqrt{\bar{\alpha}_{t-1}}\left(\frac{x_t - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(x_t, t)}{\sqrt{\bar{\alpha}_t}}\right) + \sqrt{1 - \bar{\alpha}_{t-1} - \sigma_t^2}\epsilon_\theta(x_t, t) + \sigma_t\epsilon$$

where $\sigma_t = \sqrt{(1-\bar{\alpha}_{t-1})/(1-\bar{\alpha}_t)}\sqrt{1-\bar{\alpha}_t/\bar{\alpha}_{t-1}}$.

In this paper, we build a denoising diffusion probability model with seismic as prior for seismic interpretation. During the forward diffusion learning process, we consider interpretation principles, e.g. salt, fault, etc. as a learning object, adding random Gaussian noise at each time step to the principles, and also concatenate seismic as prior to the principles to guide the diffusion process. At each diffusion time step, we update the loss function from a Unet between Gaussian noise and the diffused principle at time t. At the final step *t=T*, the principle will become a complete Gaussian noise with a learned Unet. At inference time, a testing seismic data will be used as input, concatenating with randomly generated Gaussian noise, feed into the pre-trained U-net model. By scheduling the noise removal process at time t, we can predict the diffused principle then subtract it from the input Gaussian noise. At the time t=0, the diffusion process will generate a posterior interpretation distribution based on guided seismic.

Examples

We used the synthetic SEAM data to train the diffusion model. The SEAM data represents the deep-water regions of the Gulf of Mexico, containing sediments with similar amplitude values to the salt body, which brings additional challenges to the interpretation process (Jiang et al., 2020). We divided seismic data to small patches and consider the salt mask as the interpretation principle. We design a linear scheduler to gradually add Gaussian noise to the interpretation principles at each time *t*, then considering seismic as prior to guide the diffusion process. We trained the diffusion model with a Nvidia GPU and saved the weights every 10k time steps.

Figure 2 shows an inference step from *t=T* (completely noise) to *t=0* (sampling result) with different samplings. Since we added a stochastic variable to the forward and reverse process, the sampling results show differences between each other. This will help to estimate the posterior distribution of the data and provide a natural way for uncertainty quantification.

Figure 3 shows a comparison of the same DDPM trained with different time steps. With same testing seismic, A DDPM with different training steps performs differently.

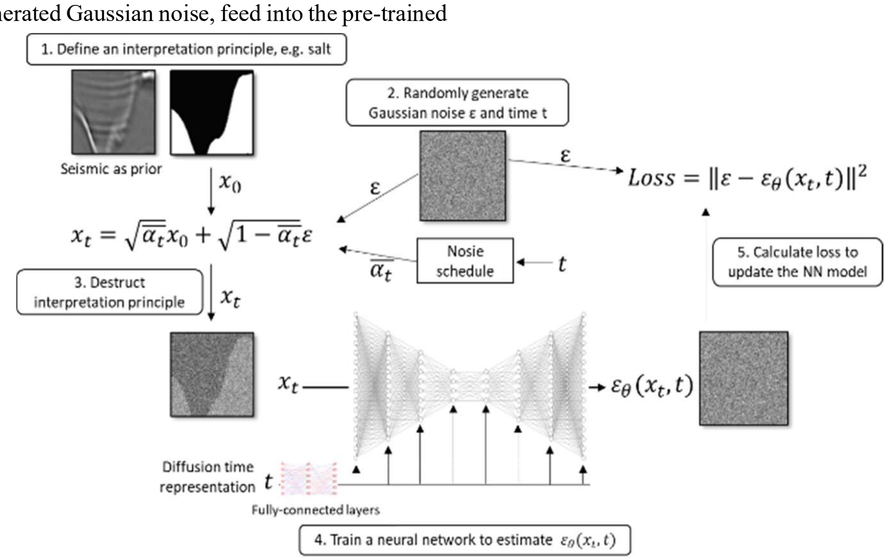

Figure 1: The architecture of the Diffusion Model for seismic interpretation. Seismic image is served as prior to guide the diffusion process on the interpretation principles.

10.1190/image2023-3907375.1

The longer the time T, the better the inference result, quantified by the mean of 100 ensembles. The aleatoric uncertainty and epistemic uncertainty of seismic data (Jiang et al., 2022) decreases with the model trained by longer time T.

Compared with other generative models, such as Generative Adversarial Networks (GANs) and Autoregressive models, the diffusion model is mathematically approved and is relatively easier to train. Currently, there are many the state of the art diffusion models in computer vision for image generation and coloring, there are good possibilities to implement diffusion models to potentially provide broader applications for reservoir characterization and well information enhancement.

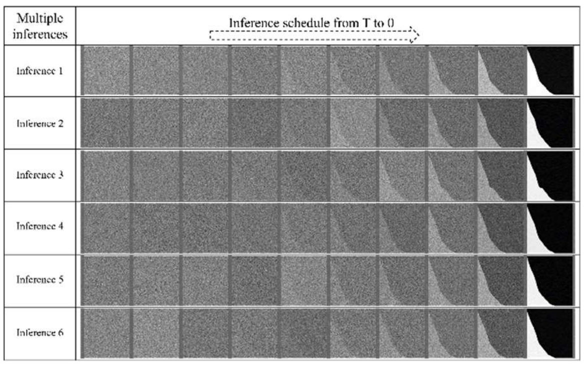

Figure 2: DDPM with salt sampling for multiple inferences.

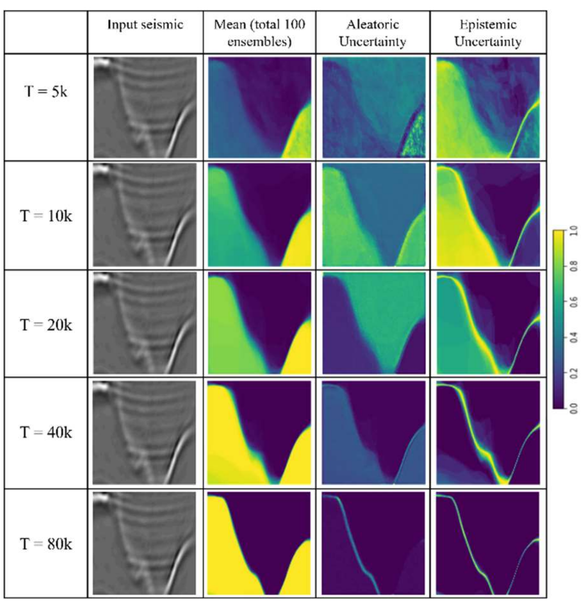

Figure 3: DDPM with different training steps. From left to right, 1: The training step; 2: Input seismic; 3: Mean distribution of 100 ensembles; 4: Aleatoric uncertainty; 5: Epistemic uncertainty.

In a general overview, the supervised machine learning model mimics how people learn seismic data, the diffusion model is an alternative method to build an architecture that how machines can recognize seismic data in a different way than humans.

Figure 4 shows the result by implementing the exact same diffusion process but with different interpretation principles. In this case, we consider fault as the interpretation principle, with diffusion process, we gradually add Gaussian noise to the fault mask, until it becomes completely noisy image. The trained Unet in this process will be different with the Unet trained with Salt images.

By implementing multiple sampling processes, we can build a posterior distribution and quantify the uncertainty in a natural way. Our results show that in some areas there exist high uncertainties because the sampling process could not generate the same interpretation predictions. This could be caused by abnormal or missing data distribution. The proposed implementation of the diffusion process is also suitable for any interpretation objects, e.g. horizon, facies, etc. The Unet used in the diffusion process will be served as an optimization tool to minimize the loss function at each diffusion step t.

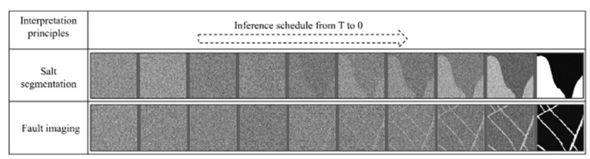

Figure 4: DDPM designed for different seismic interpretation tasks.

Discussions

As generative model, diffusion process gains popularity and recognition in the image generation domain. The rich features and its derivatives bring big potential to the energy industry. Diffusion models can generate new data samples like those which they are trained on. This generative nature led to its rapid adoption for synthetic data generation. The deconstruction and reconstruction process learn the data distribution and could help to inpainting the missing part of a reservoir model to scale up for super resolution.

However, one of the main limitations of diffusion models is the complexity of sample generation. It requires relatively a large number of inference timesteps to capture the data

10.1190/image2023-3907375.1

distribution to denoise the image. For a large dataset, this could become a major limitation to implement for real seismic data. Watson et al (2022) introduced Differentiable Diffusion Sampler Search (DDSS) to optimize fast sampler for any pre-trained diffusion model by differentiating through sample quality scores. Aiello et al. (2023) introduced the ideas of Maximum Mean Discrepancy (MMD) to finetune the learned distribution with a given budget of timesteps. This allows the finetuned model to significantly improve the speed-quality trade-off, by substantially increasing fidelity in inference regimes.

Conclusions

In this paper, we introduce a novel diffusion process with prior to build a new way to teach how the machine can learn the data distribution from the seismic by approximating an unknown data distribution. The proposed method can be used for different interpretation objects with the same diffusion process. It provides an alternative way to build posterior distribution and quantify the uncertainties. The further implementation of the proposed diffusion process could be adapted for reservoir modeling and characterization with some modifications.

Acknowledges

The authors would like to thank the SEG for using the SEAM seismic dataset.

REFERENCES

Aiello, P., M. Dawood, F. Minhas, 2023, Maximum mean discrepancy kernels for predictive and prognostic modeling of whole slide images: arXiv, 2301.09624 [eess.IV].

Goodfellow, I., 2016, NIPS 2016 tutorial: Generative adversarial networks: arXiv:1701.00160 [cs.LG].

Ho, J., A. Jain, and P. Abbeel, 2020, Denoising diffusion probabilistic models: arXiv, 2006.11239 [cs.LG].

Jiang, F., P. Norlund, and Z. Wei, 2020, Analysis of seismic attributes to assist in the classification of salt by multi-channel convolutional neural networks, EAGE Digital, Extended Abstracts, doi: https://doi.org/10.3997/2214-4609.202032049.

Jiang, F., K. Osypov, J. Toms, 2022, Uncertainty analysis for seismic salt interpretation by convolutional neural networks, SPE-211654-MS, doi: https://doi.org/10.2118/211654-MS.

Kingma, D., and M. Welling, 2013, Auto-encoding variational bayes: arXiv, 1312.6114 [stat.ML].

Lamb, A, 2021, A brief introduction to generative models: arXiv, 2103.00265 [cs.LG].

Rombach, R., A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, 2022, High-resolution image synthesis with latent diffusion models: arXiv, 2112.10752 [cs.CV].

Sahak, H., D. Watson, C. Saharia, and D. Fleet, 2023, Denoising diffusion probabilistic models for robust Image super-resolution in the wild: arXiv, 2302.07864 [cs.CV].

Sohl-Dickstein, J., E. Weiss, N. Maheswaranathan, and S. Ganguli, 2015, Deep unsupervised learning using nonequilibrium thermodynamics: arXiv, 1503.03585 [cs.LG].

Watson, D., W. Chan, J. Ho, and M. Norouzi, 2022, Learning fast samplers for diffusion models by differentiating through sample quality: arXiv, 2202.05830 [cs.LG].

Wolleb, J., F. Bieder, R. Sandkuhler, and P. Cattin, 2022, Diffusion models for medical anomaly detection: Medical Image Computing and Computer Assisted Intervention – MICCAI 35–45, doi: https://doi.org/10.1007/978-3-031-16452-1_4.

10.1190/image2023-3907375.1