Mitigating Paucity of Data in Sinusoid Characterization Using Generative Synthetic Noise

Sam Sattarzadeh Shervin Manzuri Shalmani Shervin Azad
Goldspot Discoveries Corporation
{sam.sattarzadeh, shervin.manzurishalmani, shervin}@goldspot.ca

Abstract

Although the remarkable breakthrough offered by Deep Learning (DL) models in numerous computer vision tasks, the need to acquire large amounts of high-quality natural data and fine-grained annotations is a shortcoming that fundamentally increases the cost and time devoted to training these models in real-world applications. Hence, synthetic datasets are considered reliable alternatives that can reduce the data acquisition by replacing or merging with natural data or effective pre-training of the models. To this end, in this work, we propose a novel approach to integrate structural data structures with the synthetic noise structures learned by unsupervised models that mimic the noise structures in natural data. Based on the proposed approach, we introduce the Sinusoid Feature Recognition (SFR) dataset, which contains hard-to-detect fixed-period sinusoid waves. While the previous works in this regard use generative models to sample synthetic data to inflate the training set, we instead apply unsupervised learning models to generate deep synthetic noise which makes training models in the proposed dataset more challenging. We evaluate the segmentation, image reconstruction, and sinusoid characterization models pre-trained or fully trained on the synthetic SFR dataset on a private dataset of grayscale Acoustic Tele-Viewer (ATV) images. Experimental results show that supervision on our proposed synthetic dataset can improve the accuracy of the models by 3-4% via pre-training, and by 17-27% via ad-hoc training while dealing with challenging, realistic real-world images.

1. Introduction

Recognizing sinusoidal structures is a fundamental task with applications in various signal processing and image processing fields such as radar imaging, seismic imaging, underwater acoustics, healthcare, and geology [4, 6, 27]. Each sinusoidal structure can be represented by features determining its frequency, amplitude, and initial phase. More features can be extracted for more complex structures by decomposing the primary signal into a couple of sinusoidal waves, then encoding each wave to its features. In both signal and image data, extracting these sinusoidal features encodes information about the data.

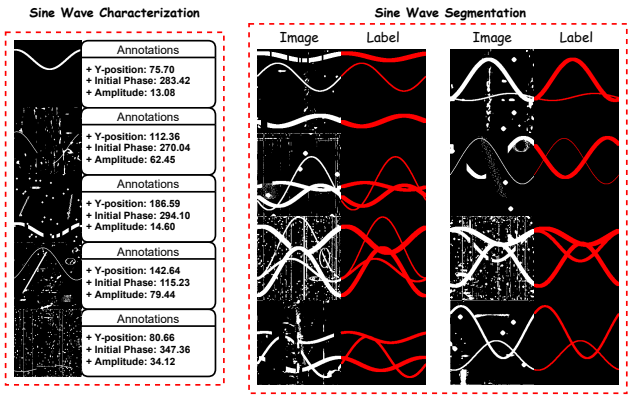

Figure 1. An overview of the proposed Sinusoid Feature Recognition (SFR) dataset for regression (left sub-figure), and segmentation (right sub-figure) tasks.

Automated interpretation of geologic features is one of the most important but often unnoticed applications of sinusoid recognition [27]. In geology sites, drill cores contain valuable information about fractures in subsurface formations that are crucial in several extractive industries. Detecting and interpreting these fractures leads the geology and petroleum experts to assess rock quality, characterize gas and oil reservoirs, and perform geological modeling for subsurfaces. Geologists employ acoustic probes to generate unwrapped grayscale images from boreholes corresponding to the drill cores by transmitting ultrasonic pulses from a transducer and observing the time to receive the signals reflecting from the borehole wall. In these types of images, which are known as Acoustic Tele-viewer (ATV) images, planar fractures are represented with sinusoids whose initial phase and amplitude represent the axial and radial angle of the fractures in the wrapped core [10].

In the past three decades, a considerable research effort was put to inspect automated noise-robust solutions in detecting sinusoidal features from grayscale Acoustic Televiewer images and color Optical Tele-viewer (OTV) images. With the recent advances in Deep-learning (DL), data-driven approaches utilizing object detection and segmentation models have achieved an outstanding accuracy in predicting sine waves and their features from raw ATV images [12, 35, 37]. However, preparing pixel-level or bounding box-level labeled data for training vision models employed in these approaches is a fundamentally costly and time-consuming procedure. Moreover, human errors while providing pixel-level annotations in large-scale data are inevitable, which noticeably hurdles training the network's accuracy during training.

Utilizing synthetic data has increasingly garnered interest in overcoming the above limitations and reducing or negating the need to spend vast amounts of time, money, and human resources. In addition to these advantages, synthetic datasets can also eliminate biases and improve fairness in deep models. For example, some recent approaches integrated raw data with synthetic data to provide clean fuel for training segmentation models [1, 7]. Furthermore, synthetic data can be generated using simulation software or generative models to integrate hand-crafted labeled data with natural information.

This work proposes a novel approach for generating synthetic datasets functionalized for recognizing and characterizing structural shapes. Unlike the prior works that mainly utilize synthetic data collected from a generative model for inflating and oversampling the training set, our approach utilizes unsupervised generative modeling frameworks to derive noise structures close to the ones seen in natural images. Through our proposed approach, we form a dataset containing ideal or partially distorted fixed-period sine waves alongside various types of noisy structures, which make the prediction task more challenging. We show that such a synthetic dataset can facilitate the training of networks specialized for automated sinusoid recognition by reducing the need to provide natural data in large scales or high-level annotations. Our contributions can be summarized as follows:

- We propose an approach to creating synthetic datasets by employing unsupervised vision models for generating complex noise structures.
- We introduce the Sinusoid Feature Recognition dataset (SFR), a publicly available synthetic dataset for fixed-period sinusoid segmentation and characterization tasks.
- We carry out several experiments to verify the ability of the SFR dataset to pre-train or train Deep Learning models that are functional in real-world applications.

2. Related works

2.1. Automated fixed-period sinusoid recognition

The earliest sinusoid characterization operates by generating a binary edge-map from ATV images and then applying Hough transform to obtain the amplitude and initial phase of the existing waves [16]. This framework was developed later by adding noise removal pre-processing steps [15], and replacing Hough transform with generalized Radon transform [34]. In [8], odd-symmetry was employed to detect the regions containing sine waves. Later, [2] used the same idea to estimate the features of sinusoid waves by matching the edge points in the same sine waves.

Instead of proposing image processing solutions, the novel approaches employ deep vision models to perform several sinusoid recognition tasks. In [12], a Fast R-CNN model [14] is trained on ATV images to detect breakouts and sinusoidal fractures. Alzubaidi et al. [1] applied a Mask R-CNN model [18] to segment sinusoidal fractures, and afterward, characterized the sinusoids by applying a two-step fitting using no-linear least squares. Zhang et al. [35] combined a pre-trained ResNet-101 backbone [19] with a U-Net architecture [28] to propose a fracture segmentation model named RUNet.

2.2. Synthetic datasets

In recent years, synthetic datasets have been used in many computer vision applications where providing pixel-level or even bounding-box level annotations is a prohibitive process.These applications include, but are not limited to optical flow estimation [24], autonomous driving [26, 33], and semantic segmentation [29]. More recently, the interest on synthetic dataset generation has been significantly increased in 3D vision tasks such as human shape modelling [36], object shape reconstruction [20], 3D keypoint extraction [32], and human pose estimation [13].

Synthetic data are created in automated ATV image interpretation by adding structural shapes to natural data. Deep learning-based sinusoid recognition solutions such as [1, 12] increase their dataset size by adding synthetic dark sinusoids to smooth image patches. In [12], some synthetic waves become distorted using simple image processing techniques (*e.g.* adding Gaussian noise, decreasing contrast). In these approaches, synthetic data can go as high as 30,000 images [12]. However, unsupervised generative models, which have shown an outstanding ability in creating artificial datasets in applications such as clothing [23] and autonomous driving [21] are not employed in this application yet.

2.3. Unsupervised anomaly localization

Unsupervised anomaly localization is to detect and segment abnormal patterns in test images that differ notice-

ably from the majority of the training data. Anomaly detection and localization methods can be divided into reconstruction-based or embedding similarity-based methods [11]. This work focuses on reconstruction-based anomaly localization approaches. In reconstruction-based methods, an unsupervised model such as Convolutional Autoencoder [5] (CAE), Variational Autoencoder [3] (VAE), or generative adversarial network (GAN) [30,31] is initially learned to reconstruct a set of normal training images. In the evaluation phase, the test images given to the model are compared with their corresponding reconstruction images. Finally, these two images are subtracted to form a residual map where the model's pixels that fail to reconstruct are highlighted as anomalies. This heat map is also known as an *anomaly map*.

3. Sinusoid feature recognition (SFR) dataset

Formerly, deep generative models such as GANs have been used as solid tools for augmenting and generating synthetic data and annotations in various tasks, such as nodule detection in CT lung and breast CT scan images [7, 9]. Unlike this group of approaches, we employ unsupervised models for generating complex noise structures. In this work, we term this type of noise as *deep synthetic noise*. In our proposed dataset, synthetic data are formed by generating random shapes that represent the structure of object instances in natural data, applying random perturbations to the ideal structures, and finally, adding three types of noise to the data, which are incoherent noise, coherent noise, and deep synthetic noise, respectively. The first two types of noise are also called *model-free noise* since they are handcrafted using image processing techniques. Figure 2 shows the workflow for forming structures and noise, as well as samples of the generated synthetic data.

3.1. Structural data

Most structural shapes, such as lines, ellipses, and rectangles, can be represented with a couple of numeric shapes. For example, each nodule can be modelled with an ellipse with five parameters. This work defines its target structural shapes as horizontal fixed-period sine waves. Given a grayscale image $\mathcal{I} = \{I | I : \Lambda \rightarrow \mathbb{R}\}$ with the dimensions $H \times W$ where $\Lambda = \{1, ..., H\} \times \{1, ..., W\}$ represents the set of the locations in its domain, a sine wave can be defined as:

$$y = d + A \sin(\phi + \omega x) \tag{1}$$

where $\omega = \frac{2\pi}{W}$ is the fixed frequency of the waves, and the parameters $d \in [0, H]$, $A \in [0, \frac{H}{2}]$, and $\phi \in [0, 2\pi]$ represent the horizontal offset, amplitude, and initial phase of each wave, respectively. To generate each idea sinusoid, initially, two random values for the initial phase and amplitude are selected, and a parameter t defines the thickness of the wave. Then, the parameter d is selected from the range $[A + t, H - (A + t)]$. Note that all of the parameters above are chosen using a uniform distribution.

To make partial random perturbations on the ideal structural data, we inherit the masking technique used in [25], where random 2-Dimensional masks are generated in low dimensions using Monte-Carlo sampling and then upscaled to mask continuous regions in the image domain instead of individual pixels. The perturbation process can be summarized as follows:

1. Select a random number in range $l = [2, \log_2 \frac{W}{2}]$. Then, sample a binary mask of size 2^l by setting each of its elements independently to 1 with a probability of p_m, or to 0 with a probability of $1 - p_m$.

2. Upsample each array to the size $(W + 2^l)$ via linear interpolation.

3. Randomly crop a continuous subset of the array of the size W, using a uniformly random indent in range $[0, 2^l]$.

4. Vertically Paste H copies of the generated array to extend the mask.

Before perturbation, each sine wave is represented as a binary mask with value 1 in the locations (x, y) and 0 in the other regions. Perturbation is then performed by generating a random mask via the above process and masking a sinusoid image via point-wise multiplication.

3.2. Model-free noise

Adding noisy patterns to the ideal or masked data makes the synthetic dataset more challenging and consequently makes the models pre-trained or trained on this dataset more robust against the causes of distortion. In this work, two model-free types of noise are added to the data, which are described as follows:

3.2.1 Incoherent noise:

This type of noise appears like blobs in different scales. Incoherent noise is generated with a similar process to generating perturbation masks. Initially, a random binary mask of the size $H' \times W'$ is generated, where $H' \in [\frac{H}{8}, \frac{H}{2}]$ and $W' \in [\frac{W}{8}, \frac{W}{2}]$. Each element in this mask is independently set to 1 with a probability of p_i and 0 with a probability of $1 - p_i$. Then, the mask is upsampled to the size (H, W).

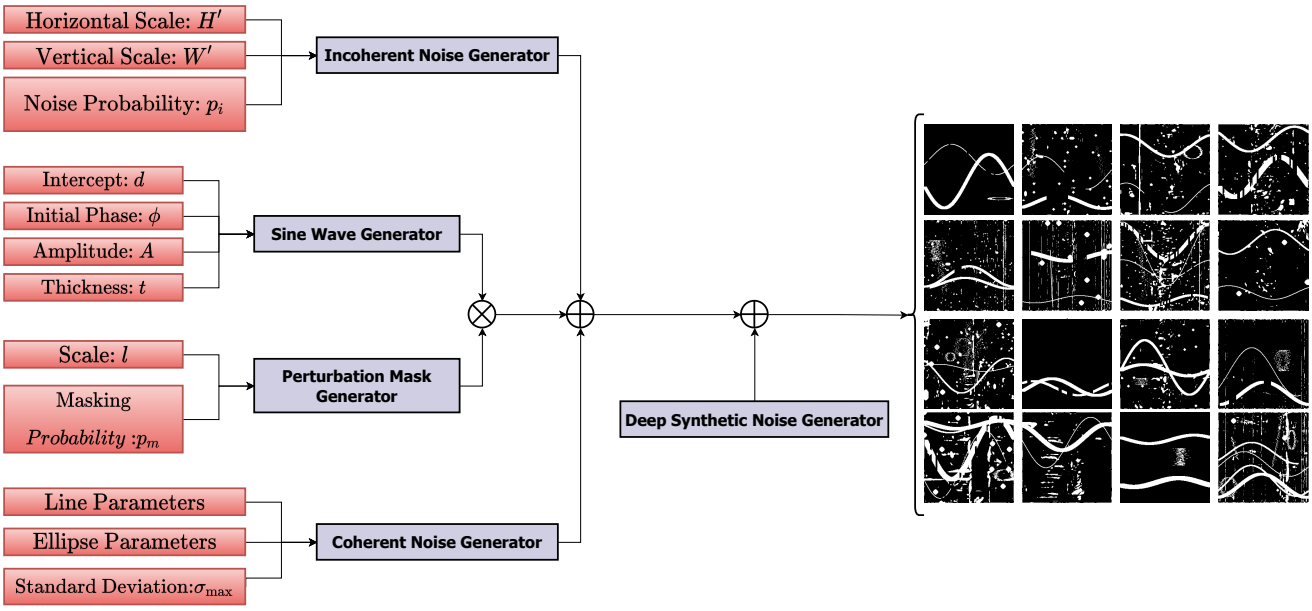

Figure 2. The framework for generating the synthetic SFR dataset. This figure shows the parameters used in each module.

3.2.2 Coherent noise:

This type of noise contains the shapes with a different form rather than that of the structural data. In this work, incoherent noise contains non-sinusoidal shapes. These shapes include random ellipses and lines. Ellipses are defined with two parameters representing the location of their center and two parameters representing their semi-major and semi-minor axis. Lines also can be described with four parameters representing their start and end location. The noisy instances are also deformed with an additive Gaussian noise, which slightly distorts the coherent noisy shapes. Hence, the equation for the coherent noisy lines are formed as:

$$y_{line} = a_{line} + m(x_{line} - b_{line}) + \mathcal{G}(x_{line}, \sigma) \quad (2)$$

where $\mathcal{G}(x, \sigma)$ is a function that outputs a sample from a Gaussian distribution with the standard deviation $\sigma \in [0, \sigma_m]$, independently from its input x. In the same way, given the variable $\theta \in [0, 2\pi]$, the ellipse equations can be formulated as:

$$x_{ellipse}(\theta) = a_{ellipse} + r_x \cos(\theta) + \mathcal{G}(x_{ellipse}, \sigma) \quad (3)$$

$$y_{ellipse}(\theta) = b_{ellipse} + r_y \sin(\theta) \quad (4)$$

3.3. Deep synthetic noise

In addition to the simple noisy patterns discussed above, the most complex type of noise in our synthetic dataset comes from reconstruction-based anomaly segmentation frameworks. Such unsupervised frameworks consist of an encoder-decoder architecture. We employed a Variational Autoencoder (VAE) architecture in the deep synthetic generator [5, 22]. As shown in fig. 4, the VAE's encoder maps each given input $\mathcal{I} = \{I | I : \Lambda \rightarrow \mathbb{R}\}$ a n-dimensional space of numeric representations $z \in \mathbb{R}^n$ known as *latent space*, while the decoder aims to invert the encoder's trained function and form a reconstruction of the input image I'. In the training phase, this framework is trained with a dataset of "normal" images[1] using the following loss:

$$\mathcal{L}_{\Phi,\Theta} = \sum_{\lambda \in \Lambda} ((\mathcal{I}(\lambda) - \mathcal{I}'(\lambda))^2) + \mathcal{D}_{KL}(q_\Phi(z|\mathcal{I}) || p(z)) \quad (5)$$

In the loss function above, the term $\mathcal{D}_{KL}(.)$ minimizes the Kullback–Leibler (KL) distance between the prior distribution of the latent factors $p(z) = \mathcal{N}(0, I)$ and the posterior likelihood distribution $q_\Phi(z|\mathcal{I})$. The VAE's encoder and decoder carry the likelihood distributions $q_{\Phi^*}(z|\mathcal{I})$ and $p_{\Theta^*}(\mathcal{I}'|z)$ respectively, after the VAE parameters are obtained by solving the minimization problem below:

$$\Phi^*, \Theta^* = \text{argmin}_{\Phi,\Theta}(\mathcal{L}_{\Phi,\Theta}) \quad (6)$$

The framework only learns to reconstruct the regular patterns in the image domain and fails to accurately reconstruct anomalous patterns in a minority of the data. In this work, anomalous patterns include sine waves and non-sinusoidal noisy structures. Given each image I, the VAE's encoder samples a set of representations $z_\mathcal{I} \sim q_{\Phi^*}(z|\mathcal{I})$ and passes them to its decoder. The anomaly detection VAE outputs

[1] To generate the SFR dataset, normal images denote the ATV images without sine waves. Thus, they can lead the anomaly detector to learn non-sinusoidal noise structures.

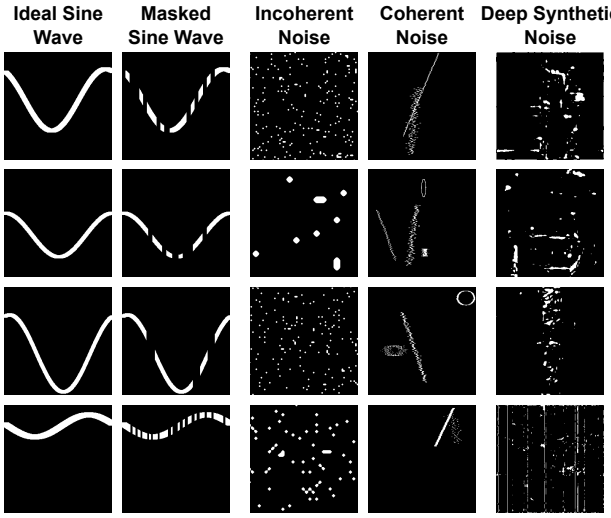

Figure 3. Qualitative samples of the sinusoidal structures and noise generated in the SFR dataset.

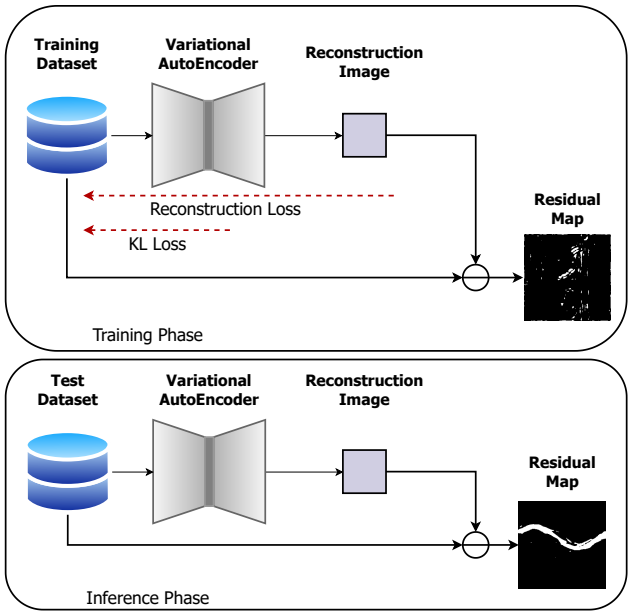

Figure 4. The deep synthetic noise generator module consists of a VAE-Based anomaly detection framework fed with training images in its inference phase, and outputs non-sinusoidal noisy structures that the VAE has not learn to reconstruct during training.

a residual map $\mathcal{R}$ by computing the absolute value of the point-wise difference between the encoder's input and decoder's output.

$$\mathcal{R} = \text{abs}(p_{\Theta^*}(\mathcal{I}'|z_{\mathcal{I}}) - \mathcal{I}) \tag{7}$$

The purpose of training the VAE is to generate non-sinusoidal noisy patterns. This is achieved when the VAE is fed with normal images in the inference phase. Our deep synthetic noise generator employs the anomaly segmentation VAE to form residual maps, given the normal images. Then, it applies a threshold $t_{\mathcal{R}} = 0.1$ on the obtained residual maps to extract non-sinusoidal noise, and finally, adds them to the generated images to form the synthetic data. Figure 3 depicts samples of ideal and masked sinusoids, as well as incoherent, coherent, and deep synthetic noise added to the data.

4. Experiments

In this section, we present the results of experiments conducted to investigate the ability of the proposed synthetic dataset in three sinusoid recognition tasks: 1) pre-training semantic segmentation networks, 2) training image reconstruction models that are operational in the real world, and 3) deep characterization of sinusoids. The following experiments are set up using a private dataset of ATV images containing sinusoid and smooth patches for whom pixel-level annotations are partially available, and patch-level annotations are fully available. All experiments are conducted on an NVIDIA GeForce RTX 3060 GPU with 6GB memory.

To generate the SFR dataset, 9,000 training and 1,000 synthetic validation images containing up to four sinusoids are formed. For each image, Each of the operations masking, adding coherent noise, adding incoherent noise, and adding deep synthetic noise are applied with a probability of 0.5. The masking ratio p_m is randomly selected from the range $[0.7, 1]$. The maximum standard deviation for both coherent and incoherent noise is set to $\sigma_m = 7.0$. Furthermore, the amount of incoherent noise p_i is sampled from the range $[0, 0.04]$.

4.1. Semantic segmentation

This experiment is conducted to evaluate the effect of pre-training a U-Net model [28] on synthetic data on increasing its accuracy and data efficiency. This experiment uses a dataset containing 5,226 training and 581 validation images. Each training or validation data is a 256×256 image containing at least one sine wave. Moreover, pixel-level annotation is provided for all data. The U-Net model is trained with two different strategies as follows:

1. Training with the ATV data in 50 epochs, without pre-training.

2. Pre-training with the SFR data in 20 epochs, and then fine-tuning on the ATV data with 50 epochs.

In both strategies, training is processed using pixel-wise Cross entropy loss and a Stochastic Gradient Descent (SGD) optimizer, setting the learning rate to 0.01 and batch size to 4. The model's sinusoid segmentation maps are

Pre-training	Sample Rate (γ)	Sensitivity	Specificity	F-1 Score
Not Applied	0.33	46.71%	47.94%	20.30
	0.66	47.03%	47.99%	20.53
	1	51.73%	48.88%	22.42%
Applied	0.33	54.21%	48.95%	23.11%
	0.66	**56.75%**	**49.56%**	**24.17%**

Table 1. Evaluation of the UNet trained with limited training data for the states when pre-training with the SFR dataset is applied, and is not applied. The parameter γ denotes the fraction of the training set that the UNet is fed with.

filtered to evaluate the networks based on their median value. Let the validation set contain K images (which is equal to 581 in this experiment), the set of indexes $\{k|k \in \{1, ..., K\}\}$, and the model's binarized outputs and annotations be denoted as S_k and G_k respectively, the evaluation metrics are computed based on the following terms:

- **True Positive:** $\text{TP}_k = ||S_k \odot G_k||_1$
- **False Positive:** $\text{FP}_k = ||S_k \odot G'_k||_1$
- **True Negative:** $\text{TN}_k = ||S'_k \odot G'_k||_1$
- **False Negative:** $\text{FN}_k = ||S'_k \odot G_k||_1$

where S'_k and G'_k represent the complimentary of S_k and G_k, respectively. The evaluation metrics are defined as:

$$\text{Sensitivity} = \frac{1}{K}\sum_{K}\frac{\text{TP}_k}{\text{TP}_k + \text{FN}_k} \tag{8}$$

$$\text{Specificity} = \frac{1}{K}\sum_{K}\frac{\text{TN}_k}{\text{TN}_k + \text{FP}_k} \tag{9}$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \tag{10}$$

where the term *precision* is defined as:

$$\text{Precision} = \frac{1}{K}\sum_{K}\frac{\text{TP}_k}{\text{TP}_k + \text{FP}_k} \tag{11}$$

To observe the data efficiency of the model, we utilize only γ percent of the available training data. The outperforming results of the pre-trained model as shown in Table 1 infer that our synthetic dataset is able to compensate lack of clean or labelled training data for training the model since the coarse non-sinusoidal data are employed to generate representations in the synthetic dataset.

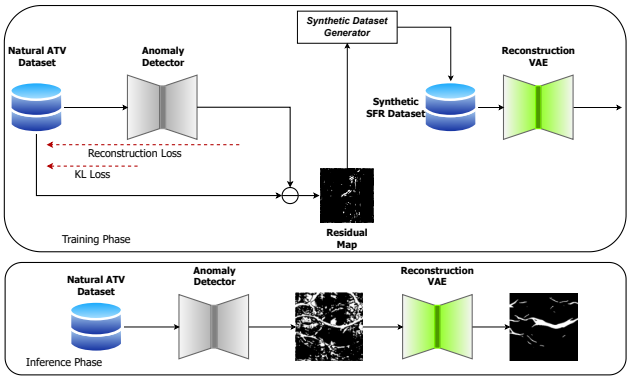

Figure 5. The overview of the image reconstruction experiment. The anomaly detector is trained with the natural data but utilized to generate deep synthetic SFR noise. A reconstruction model is trained with the SFR dataset and applied to remove noise from the anomaly detector's output.

4.2. Image reconstruction

The objective of this experiment is to segment sinusoidal patterns only with a set of normal ATV data. Sinusoid waves are considered anomalies that are not seen in the training data. Hence, reconstruction-based anomaly detection models aim to learn the smooth representations in the training data and output the sinusoidal structures in their residual map. However, their residual map contains various noisy structures in addition to the sine waves, making it very likely for false positives to appear in the residual maps.

In this experiment, we initially train two anomaly detection models with the natural data and then use the trained models as a deep synthetic noise generator for the SFR dataset. Then, we train a VAE that aims to reconstruct the sinusoids, given the noisy residual outputs from the anomaly detection model. The architecture of the VAE is shown in table 2. Figure 5 Shows the overview of the sinusoid reconstruction experiment. Furthermore, the applied anomaly detection models are a VAE trained with Structural Similarity (SSIM) loss with the encoder architecture defined in [5],and referred as *SSIM-VAE*, and the fAnoGAN model proposed in [30].

The utilized training image contains 3,960 training images. The anomaly detectors are trained using the Adam optimizer and an L2 regularizer with a coefficient of 0.01, setting the learning rate to 0.01 and batch size to 32. The image reconstruction model is trained with the same optimizer and configurations as the anomaly detection model, except that the learning rate is set to 10^{-6} and a weight decay with the coefficient of 0.95 is applied. The validation set in the previous section is used to evaluate the quality of residual maps. In the evaluation, we use sensitivity and F1 metrics to measure the amount of noise removed by the reconstruction model. Table 3 shows that the reconstruction

model is able to increase the sensitivity of the segmentation model considerably. This improvement is achieved without extending the training dataset, neither improving the quality of their data nor annotations, but it is reached with full supervision on synthetic data.

Layer	Output Channels	Kernel	Stride	Padding
Input	1			
Conv1	32	4×4	2	1
Conv2	32	4×4	1	2
Conv3	32	3×3	1	1
Conv4	64	4×4	2	1
Conv5	64	3×3	1	1
Conv6	128	4×4	2	1
Conv7	64	3×3	1	1
Conv8	32	3×3	1	1
Flatten	32,768			
Linear	50			
Linear	50			

Table 2. Summary of the architecture of the encoder of the image reconstruction VAE. The last two linear layers output the mean and variance for the probability of the latent factors.

Model	SSIM-VAE		fAnoGAN	
Framework	Original	Proposed	Original	Proposed
Precision	14.12%	**16.95%**	14.45%	**20.75%**
Specificity	50.64%	**77.83%**	60.19%	**87.62%**

Table 3. The results of the image reconstruction experiment. The proposed approach refers to integrating the anomaly detection models with the reconstruction VAE trained with the SFR dataset, as shown in fig. 5

4.3. Sinusoid characterization

Unlike the previous experiments, this experiment focuses on estimating the features describing sinusoid waves rather than segmenting sinusoid pixels. In this section, A two-branch model is designated to detect and characterize sinusoid waves in a stitched image. The goal of this experiment is to estimate three parameters "horizontal depth", "initial phase", and "amplitude" for each of the waves. The model designated for this experiment operates in two steps. In the first step, it provides a binarized segmentation map representing anomalies in the image and the locations where a sinusoid wave is predicted to exist. Next, each segment is fitted to a sine wave with the frequency of $\omega = \frac{2\pi}{W}$.

Figure 7 depicts the sinusoid characterization framework. Initially, a square sliding window is vertically moved throughout the stitched image with a stride of 30 pixels,

Depth Error (pixels)	Phase Error (deg)	Amplitude Error (pixels)
1.150	2.729	0.853

Table 4. Error of the RexNet-100 model on estimating the sinusoid features in the SFR validation set.

and a RexNet-100 classifier [17] is applied to the corresponding patches. Then, the patches where the classifier predicts the existence of a sine wave with the highest confidence score in their neighbourhood are passed to the SSIM-VAE [5] anomaly detector trained in the previous section if the confidence score passes a threshold of 0.5. The anomaly maps binarized by the SSIM-VAE are passed to a sinusoid characterization module. To train the RexNet-100 classifier, 26,970 sinusoid images are added to the non-sinusoid images used in the previous section to train the anomaly detectors. Training this classifier is done using Cross entropy loss and an SGD optimizer with a learning rate of 0.002 and a weight decay of 3×10^{-4}.

In the prior sinusoid characterization literature, sine waves are predicted to the segment locations using non-linear least squares [1]. As a benchmark, we applied this technique on the contours of the large segments in the anomaly maps, as well as the combination of separated segments in the predicted patches. The predicted sine waves by this approach are then enhanced using a non-maximum suppression step to avoid predicting multiple waves for the same locations.

In addition, we train a second RexNet-100 model for a regression task on a recast of the SFR dataset that contains only one sine wave in each of its data. The regression model's output layer is designated with 3 nodes, each of them representing one of the sinusoid features. This model is trained using the Mean Squared Error (MSE) loss between the predicted and correct sinusoid features and a SGD optimizer with a learning rate of 10^{-6} and a momentum of 0.9. Table 4 reports the accuracy measures of this model on the SFR validation set.

The stitched image used for evaluation has a dimension of $115,078 \times 256$ pixels. This image pictures a fraction of a borehole of height of 230.15 meters, containing 678 sine waves. Let the features for each wave denoted as (d_i, ϕ_i, A_i) where $i \in \{1, ..., 678\}$ and the features above represent depth, initial phase, and amplitude, respectively. The features are translated to the features $(d'_i, \alpha_i, \beta_i)$. The first feature represents the depth in meters, and the latter two are named "azimuth" and "dip", respectively. Geology experts are more interested in dip and azimuth values, since these features enable them to plot the orientations in a 360° stereonet. Assume two constant offsets C_α and C_β, these

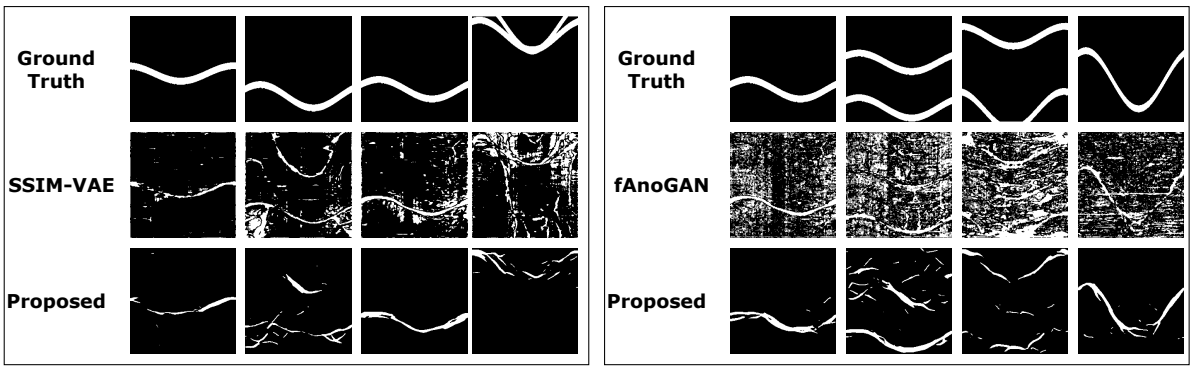

Figure 6. Qualitative comparison of the segmentation maps outputted by anomaly detection models and the segmentation maps enhanced via the model trained on the SFR dataset (proposed).

Reconstruction VAE	Estimator	Precision	Recall
Not Applied	Least Square [1]	0.310	0.194
Not Applied	RexNet-100	**0.379**	**0.301**
Applied	RexNet-100	0.342	0.272

Table 5. The results of the sinusoid characterizaton model using with respect to the configurations according to the Figure 7.

features are calculated as follows:

$$\alpha_i = \phi_i + C_\alpha \tag{12}$$

$$\beta_i = \arctan(\frac{A}{R}) + C_\beta \tag{13}$$

where $R = \frac{W}{2\pi}$ is the radius of the borehole, and the offsets are set to $C_\alpha = 85°$ and $C_\beta = 12°$, respectively. Each predicted sine wave $(d'_{\text{pred}}, \alpha_{\text{pred}}, \beta_{\text{pred}})$ is counted as a True Positive (TP) if there exists a ground truth sine wave with the index i with the conditions below, and a False Positive (FP), otherwise:

1. $|d'_{\text{pred}} - d'_i| < 15$ cm
2. $|\alpha_{\text{pred}} - \alpha_i| < 15°$
3. $|\beta_{\text{pred}} - \beta_i| < 15°$

Each sine wave characterized with $(d'_{\text{pred}}, \alpha_{\text{pred}}, \beta_{\text{pred}})$ is matched with the i-th ground truth wave if the conditions above are met. False Negative (FN) denotes the number of ground truth waves to whom no predicted wave is matched.

5. Conclusion

This work proposed a novel approach to derive synthetic noise by training unsupervised anomaly detectors with natural data and mixing them with structural data. We proposed the SFR dataset to recognize and characterize fixed-period sinusoids in partially obscured and noisy data according to this approach. Our experimental results verified that although our proposed dataset is synthetic, it alleviates the need for large-scale, high-quality natural data and pixel-level labels in real-world problems dealing with natural ATV images, both by efficient pre-training the segmentation models and training image reconstruction and regression models.

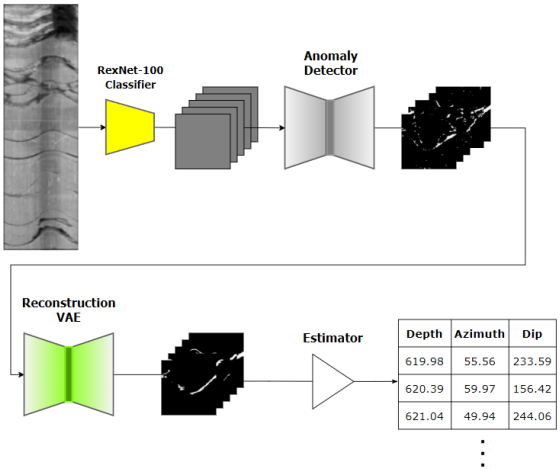

Figure 7. The framework for the sinusoid characterization experiment. The results of this experiment are reported using non-linear least square [1] sinusoid estimator and the RexNet-100 regression model trained with the SFR dataset.

Our experimental results demonstrate that our approach for generating the SFR dataset can be applied to generate more comprehensive datasets used to recognize and characterize more structural shapes. Many target objects such as nodules and polyps in healthcare, underwater imaging, and other fields can be modelled with such structures. Thus, in such fields, synthetic datasets similar to SFR are powerful tools to provide the fuel of Deep Learning vision models while lowering the data acquisition cost and mitigating the paucity of data.

References

[1] Fatimah Alzubaidi, Patrick Makuluni, Stuart R Clark, Jan Erik Lie, Peyman Mostaghimi, and Ryan T Armstrong. Automatic fracture detection and characterization from unwrapped drill-core images using mask r–cnn. *Journal of Petroleum Science and Engineering*, 208:109471, 2022. 2, 7, 8

[2] Said Assous, Peter Elkington, Stuart Clark, and James Whetton. Automated detection of planar geologic features in borehole images. *Geophysics*, 79(1):D11–D19, 2014. 2

[3] Christoph Baur, Benedikt Wiestler, Shadi Albarqouni, and Nassir Navab. Deep autoencoding models for unsupervised anomaly segmentation in brain mr images. In *International MICCAI brainlesion workshop*, pages 161–169. Springer, 2018. 3

[4] Louis G Beatty, James D George, and A Zed Robinson. Use of the complex exponential expansion as a signal representation for underwater acoustic calibration. *The Journal of the Acoustical Society of America*, 63(6):1782–1794, 1978. 1

[5] Paul Bergmann, Sindy Löwe, Michael Fauser, David Sattlegger, and Carsten Steger. Improving unsupervised defect segmentation by applying structural similarity to autoencoders. *arXiv preprint arXiv:1807.02011*, 2018. 3, 4, 6, 7

[6] Liliana Borcea, George Papanicolaou, Chrysoula Tsogka, and James Berryman. Imaging and time reversal in random media. *Inverse Problems*, 18(5):1247, 2002. 1

[7] Marco Caballo, Domenico R Pangallo, Ritse M Mann, and Ioannis Sechopoulos. Deep learning-based segmentation of breast masses in dedicated breast ct imaging: radiomic feature stability between radiologists and artificial intelligence. *Computers in biology and medicine*, 118:103629, 2020. 2, 3

[8] Zou Changchun and Shi Ge. A hough transform-based method for fast detection of fixed period sinusoidal curves in images. In *6th International Conference on Signal Processing, 2002.*, volume 1, pages 909–912. IEEE, 2002. 2

[9] Richard J Chen, Ming Y Lu, Tiffany Y Chen, Drew FK Williamson, and Faisal Mahmood. Synthetic data in machine learning for medicine and healthcare. *Nature Biomedical Engineering*, 5(6):493–497, 2021. 3

[10] Farrin de Fredrick, Ta Nguyen, Clive Seymour, and Gary Dempers. Geotechnical data from optical and acoustic televiewer surveys. *AusIMM Bulletin*, (5), 2014. 1

[11] Thomas Defard, Aleksandr Setkov, Angelique Loesch, and Romaric Audigier. Padim: a patch distribution modeling framework for anomaly detection and localization. In *International Conference on Pattern Recognition*, pages 475–489. Springer, 2021. 3

[12] Luciana Olivia Dias, Clécio R Bom, Elisangela L Faria, Manuel Blanco Valentín, Maury Duarte Correia, P Márcio, P Marcelo, and Juliana M Coelho. Automatic detection of fractures and breakouts patterns in acoustic borehole image logs using fast-region convolutional neural networks. *Journal of Petroleum Science and Engineering*, 191:107099, 2020. 2

[13] Matteo Fabbri, Fabio Lanzi, Simone Calderara, Andrea Palazzi, Roberto Vezzani, and Rita Cucchiara. Learning to detect and track visible and occluded body joints in a virtual world. In *Proceedings of the European conference on computer vision (ECCV)*, pages 430–446, 2018. 2

[14] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015. 2

[15] K Glossop, Paulo JG Lisboa, PC Russell, A Siddans, and GR Jones. An implementation of the hough transformation for the identification and labelling of fixed period sinusoidal curves. *Computer Vision and Image Understanding*, 74(1):96–100, 1999. 2

[16] Jonathan Hall, Marco Ponzi, Mauro Gonfalini, and Giorgio Maletti. Automatic extraction and characterisation of geological features and textures front borehole images and core photographs. In *SPWLA 37th Annual Logging Symposium*. OnePetro, 1996. 2

[17] Dongyoon Han, Sangdoo Yun, Byeongho Heo, and YoungJoon Yoo. Rethinking channel dimensions for efficient model design. In *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*, pages 732–741, 2021. 7

[18] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017. 2

[19] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 2

[20] Yuan-Ting Hu, Jiahong Wang, Raymond A Yeh, and Alexander G Schwing. Sail-vos 3d: A synthetic dataset and baselines for object detection and 3d mesh reconstruction from video data. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1418–1428, 2021. 2

[21] Amlan Kar, Aayush Prakash, Ming-Yu Liu, Eric Cameracci, Justin Yuan, Matt Rusiniak, David Acuna, Antonio Torralba, and Sanja Fidler. Meta-sim: Learning to generate synthetic datasets. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4551–4560, 2019. 2

[22] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013. 4

[23] Christoph Lassner, Gerard Pons-Moll, and Peter V Gehler. A generative model of people in clothing. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 853–862, 2017. 2

[24] Nikolaus Mayer, Eddy Ilg, Philip Hausser, Philipp Fischer, Daniel Cremers, Alexey Dosovitskiy, and Thomas Brox. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4040–4048, 2016. 2

[25] Vitali Petsiuk, Abir Das, and Kate Saenko. Rise: Randomized input sampling for explanation of black-box models. *arXiv preprint arXiv:1806.07421*, 2018. 3

[26] Aayush Prakash, Shaad Boochoon, Mark Brophy, David Acuna, Eric Cameracci, Gavriel State, Omer Shapira, and Stan Birchfield. Structured domain randomization: Bridging the reality gap by context-aware synthetic data. In *2019 In-*

ternational Conference on Robotics and Automation (ICRA), pages 7249–7255. IEEE, 2019. 2

[27] Malcolm H Rider. The geological interpretation of well logs. 1986. 1

[28] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015. 2, 5

[29] German Ros, Laura Sellart, Joanna Materzynska, David Vazquez, and Antonio M Lopez. The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3234–3243, 2016. 2

[30] Thomas Schlegl, Philipp Seeböck, Sebastian M Waldstein, Georg Langs, and Ursula Schmidt-Erfurth. f-anogan: Fast unsupervised anomaly detection with generative adversarial networks. *Medical image analysis*, 54:30–44, 2019. 3, 6

[31] Thomas Schlegl, Philipp Seeböck, Sebastian M Waldstein, Ursula Schmidt-Erfurth, and Georg Langs. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *International conference on information processing in medical imaging*, pages 146–157. Springer, 2017. 3

[32] Supasorn Suwajanakorn, Noah Snavely, Jonathan J Tompson, and Mohammad Norouzi. Discovery of latent 3d keypoints via end-to-end geometric reasoning. *Advances in neural information processing systems*, 31, 2018. 2

[33] Jonathan Tremblay, Aayush Prakash, David Acuna, Mark Brophy, Varun Jampani, Cem Anil, Thang To, Eric Cameracci, Shaad Boochoon, and Stan Birchfield. Training deep networks with synthetic data: Bridging the reality gap by domain randomization. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 969–977, 2018. 2

[34] M Van Ginkel, MA Kraaijveld, LJ Van Vliet, EP Reding, PW Verbeek, and HJ Lammers. Robust curve detection using a radon transform in orientation space applied to fracture detection in borehole images. In *7th Annual Conference of the Advanced School for Computing and Imaging, Heijen, The Netherlands, May 30-June 1, 2000*, 2001. 2

[35] Ye Zhang, Jinqiao Chen, and Yanlong Li. Segmentation and quantitative analysis of geological fracture: a deep transfer learning approach based on borehole televiewer image. *Arabian Journal of Geosciences*, 15(3):1–17, 2022. 2

[36] Luyang Zhu, Konstantinos Rematas, Brian Curless, Steven M Seitz, and Ira Kemelmacher-Shlizerman. Reconstructing nba players. In *European conference on computer vision*, pages 177–194. Springer, 2020. 2

[37] Xianjian Zou, Chuanying Wang, Yiteng Wang, and Huan Song. Morphological feature description method of structural surface in borehole image during in-situ instrumentation. *Rock Mechanics and Rock Engineering*, 53(7):2947–2956, 2020. 2