Bridging Gaps Between Natural and Seismic Image Analysis

Muhammad Amir Shafiq, Mohit Prabhushankar, Haibin Di*, and Ghassan AlRegib**
* Center for Energy and Geo Processing (CeGP),
School of Electrical and Computer Engineering,
Georgia Institute of Technology, Atlanta, GA, 30332

The majority of seismic image processing and interpretation workflows borrow concepts from the relatively more developed field of natural image and video processing. For example, seismic attributes based on edge, texture, frequency distribution etc. adapt the theory and concepts from natural images to be effectively utilized in the seismic domain. However, there have not been many studies carried out in the past to understand the common features between natural and seismic domains. Both natural and seismic images differ significantly from each other and this work is a step forward in understanding common features between these two domains and transfer the knowledge learned from natural images to seismic domain. By developing models based on prevalent features in both domains, we can not only automate the process of seismic interpretation but also develop new seismic attributes that highlight areas of interest in seismic sections and convey the most useful information in a compact manner. A recent trend is to apply machine learning techniques to design computational attention models by learning features labeled by expert geophysicists. However, interpreters must either generate labels themselves to train a model or rely on unsupervised learning techniques. The lack of hand-labeled data for seismic interpretation has been a limitation for developing learning-based models for seismic data.

Unsupervised learning is a branch of learning algorithms that is commonly used for exploratory data analysis when enough labeled data isn't available for finding hidden patterns, distinguishing features, and grouping in data. This work aims at bridging gaps between the broad domain knowledge from natural images and videos and evaluate its applicability in improving seismic attributes, structural automation, and seismic image processing. We propose an unsupervised learning workflow for seismic data analysis based on features learned from natural images such that seismic images were never exposed to training network. Specifically, we propose a novel approach based on a data-driven sparse autoencoder architecture that can automatically extract features from unlabeled 3D seismic volumes. We propose a learning framework as shown in Figure 1 to train a sparse autoencoder using natural images from ImageNet database. One hundred random patches of dimensions 8x8 are sampled from 1000 randomly chosen grayscale images to train an autoencoder. Back propagation is used to train weights $\boldsymbol{W}$ and bias $\boldsymbol{b}$ by setting up the objective function based on l_2-norm reconstruction error, sparsity constraints, and ridge regularizer. During testing phase, each inline section is sampled to extract patches of size 8x8 that are reshaped into a vector $\boldsymbol{V_i x 1}$. These patches are normalized to have zero mean and within the range 0~1. The normalized patches are passed through the sparse model learned during the training stage to extract all hidden layer responses. The responses are passed through the decoder weights acting as the activating functions for the decoder filters. Since in small seismic patches, edges are the most dominating features, decoder filters that exhibit strong edge characteristics are predominantly selected and given a higher weight. These edge filters are selected based on kurtosis

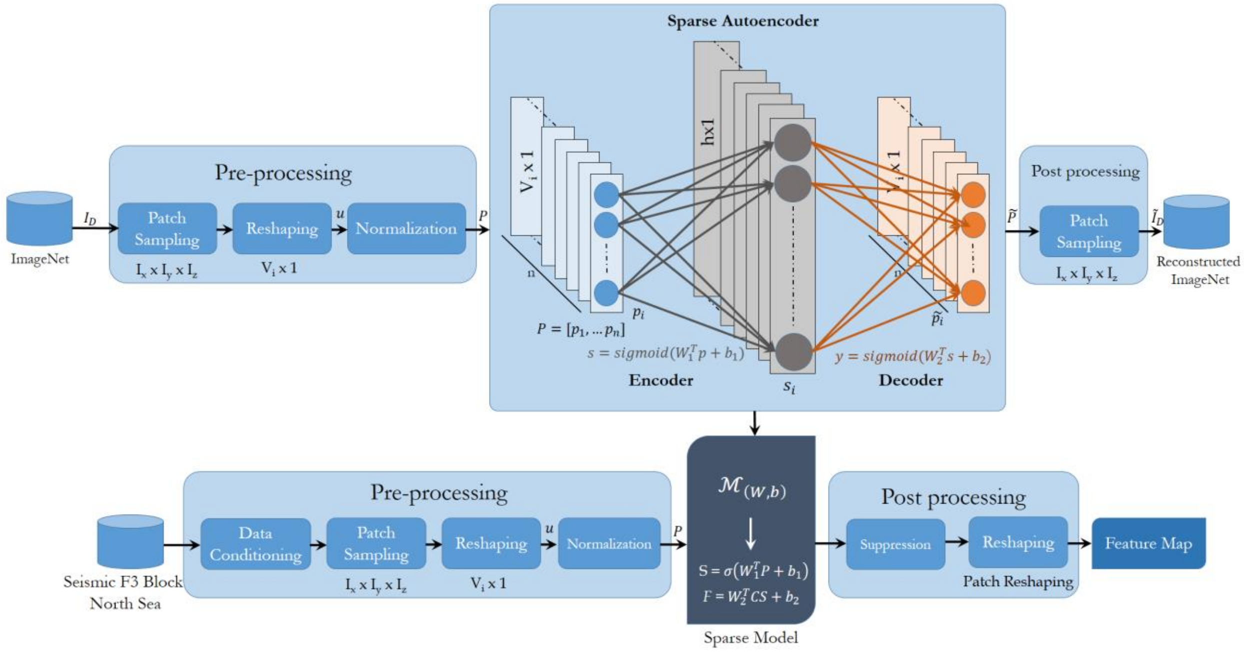

Figure 1: The block diagram of the unsupervised learning framework for seismic data analysis.

measure such that higher kurtosis measure of a filter corresponds to strong edges and vice versa. The decoder reconstructs these weighted seismic patches which are then reshaped to generate seismic feature maps.

We demonstrate that the proposed autoencoder-based approach can effectively estimate salient structures within real seismic datasets such as the F3 block in the North Sea, Netherlands, the Great South Basin, New Zealand, and SEAM dataset, which mimics the details of complex substructures found in Gulf of Mexico. Experimental results show that testing on seismic datasets reveals various features of natural images found in seismic dataset. In addition, by classifying the sparse weights learned by autoencoder based on their orientations, we can highlight different features and structures within seismic volumes. Traditionally, edge-based attributes yield the amplitude anomalies in an Omni-directional fashion; whereas, sparse autoencoder in conjunction with weights classification can help us choose and tune directional features from seismic images as well. This feature of sparse autoencoders is very beneficial in structural interpretation especially for detecting faults, horizons, or chaotic structures. The unsupervised learning combined with seismic training and weak-supervision can also be effectively utilized for seismic labeling and to learn features related to various seismic structures. The preliminary results show that understanding features of natural images and developing intuition of how features in natural images correspond to seismic images can provide valuable insights and cues for developing good attention models and transferring the domain knowledge from natural images to seismic images. Finally, with the addition of seismic labels or weak supervision during training stage, we can create targeted models that highlight only desired structures and prove to be significantly advantageous in the structural interpretation of seismic volumes.