

# **Estimation of DAS microseismic source mechanisms using unsupervised deep learning**

Matt Eaid, Chaoshun Hu, Lin Zhang, Scott Keating, and Kris Innanen

## **ABSTRACT**

Distributed acoustic sensing (DAS) is an increasingly prevalent technology for seismic acquisition, especially in reservoir monitoring settings. Recently, it has attracted interest for microseismic monitoring during hydraulic fracturing and as a complement to broadband seismometers for measuring teleseismic waves generated by earthquakes. A key component of these data is the source mechanism information encoded in the direct arrivals and how to best utilize this data to make inferences about the source mechanism, given DAS measurements of the direct arrivals is an open question. DAS is a relatively new technology, providing very large datasets of different physical aspects of the wavefield than geophones or seismometers. Consequently, conventional moment tensor inversion is challenging to transfer directly to DAS data. Instead, we turn our attention to deep learning algorithms for estimating these source mechanisms. Viewing our seismic data as containing diagnostic features of the source mechanism, we reduce the data to only those features most pertinent to moment tensor inversion by training a convolutional auto-encoder. The extracted features are then analyzed using clustering and generative adversarial networks (GAN). Clustering based on source mechanism is observed, confirming the extracted features contain important source mechanism information. Furthermore, we develop a trained GAN that provides an accurate mapping from feature space to moment tensor estimate, which shows promise when applied to a field DAS-microseismic dataset collected during hydraulic fracturing. Data modeled with the predicted source mechanism shows a strong correlation to the field data event.

## **INTRODUCTION**

Due to the increased frequency and complexity of hydraulic fracturing a need has arisen to monitor treatments and ensure fractures are propagating as planned, remaining in the target formation, propagating deep into the reservoir, and creating permeability for economic production. Microseismic monitoring is the most prevalent technology for realizing this goal (McClain, 1971; Power et al., 1976; Albright and Pearson, 1982; Warpinski et al., 1998; Maxwell et al., 2002; Maxwell, 2010). As hydraulic fractures propagate through the subsurface they generate energy in the form of seismic waves. These waves are conventionally recorded with a combination of 3C geophones in surface arrays, shallow well arrays, and deep downhole arrays (e.g. Eaton, 2018). One of the challenges with this strategy is recording high quality microseismic data in an economic fashion. Surface arrays are an economical way of recording wide aperture microseismic data, a requirement for sufficient source information, but this places sensors far from the sources and in a noisy environment. Deep downhole arrays mitigate these problems by placing receivers closer to the sources, and in a less noisy wellbore environment, however, this typically requires the drilling of dedicated monitoring wells, driving up costs, and reducing the number of sensors that can be placed.

A new technology that has gained significant interest for microseismic monitoring, that also has potential applications in earthquake seismology, is distributed acoustic sensing (DAS). Employing optical fibers, DAS makes measurements of the strain induced by propagating wavefields. A significant benefit of DAS is the noninvasive nature of the fiber used by DAS systems allowing for their placement in active hydraulic fracturing treatment wells. This allows for the dual purposing of wells for both treatment and monitoring, negating the need for dedicated monitor wells. Additionally, fibers generally run from surface to the toe of the well, turning the entire length of the well into a distributed strain sensor, providing the wide aperture recording crucial to estimating microseismic source parameters (Eyre and van der Baan, 2015). However, DAS recording comes with its own challenges. Foremost, the rigidity of the optical fibers results in sensors insensitive to all but the tangential component of strain, meaning that DAS is inherently a single component recording system (Kuvshinov, 2015). DAS also offers very dense spatial sampling on the order of one meter (Daley et al., 2013), which produces very large datasets when coupled with the continuous recording that occurs during multi-day hydraulic fracture treatments, that are cumbersome to process by conventional means. The focus of this paper is the development of efficient methods for extracting source information from DAS-microseismic data.

DAS also has potential applications in the closely related field of earthquake seismology. Broadband seismometers are currently the prevalent technology for recording teleseismic waves produced by earthquakes. However, the quality of the data they can record is limited by their sparsity, and the fact that they tend to exist on continents. Distributed acoustic sensing holds the potential to leverage terrestrial and subsea telecommunications fiber to record earthquake energy. Lindsey et al. (2017) compare teleseismic data from fiber and collocated seismometers and show a strong correlation between the two datasets. They also show DAS is capable of measuring the low-frequency responses crucial in earthquake seismology. Yu et al. (2019) develop similar conclusions, and show that DAS can aid in imaging of deep structures such as the MOHO. The apparent correlation of DAS data to broadband seismometer data, and the affinity of DAS for recording densely sampled, low-frequency data in marine and land settings suggests that DAS could provide a strong complement to seismometer data. In anticipation of this, the developments in this paper produce a workflow for acquiring source-type information from earthquake generated teleseismic DAS data.

Encoded within the seismic waves emitted by propagating fractures or earthquake sources is information about the spatial location of the fracture, and the type of fracturing that is occurring in the subsurface. The spatial distribution of fractures provides information about the extent of the stimulated rock volume, and fracture geometry. Early work on fracture localization focused on inverting the kinematic information carried by the direct arrivals, but has since progressed to stacking-based (Gharti et al., 2010; Grigoli et al., 2013), migration-based (Grandi and Oates, 2009; Mao et al., 2020), and interferometric imaging methods (Schuster et al., 2004; Zhang and Zhang, 2013; Li et al., 2016). Microseismic signals also contain important information about the source that generated the recorded data. An important component of the source information in these signals lies in the moment tensor which provides information about the type of fracturing occurring in the subsurface (e.g. Burridge and Knopoff, 1964; Gilbert, 1971; Aki and Richards, 2002). The conventional method for estimating the moment tensor, moment tensor inversion (MTI), uses phase in-

formation (Eyre and van der Baan, 2015), amplitude information (Eaton et al., 2014; Eyre and van der Baan, 2015), or both (Eyre and van der Baan, 2015; Willacy et al., 2019) to estimate the source mechanism that generated the data. These traditional methods are expensive, especially on large datasets like those supplied by DAS, and do not readily transfer to new technologies. With this in mind, this study sets out to examine alternate methods for extracting the source information embedded in DAS-microseismic data.

Moment tensor inversion relies on polarity information in the first arrivals (Eyre and van der Baan, 2015) and the relative P-wave and S-wave amplitudes (Eaton et al., 2014) to estimate the source mechanism. The problem of determining which data features are most informative about the source mechanism is highly analogous to image-feature extraction problems. Convolutional neural networks (CNN) are a prevalent technology for solving problems of this type. These networks are optimized for processing image data and extracting complex features and structure contained within the image, providing a mapping from input image to its feature space representation (Lecun et al., 1998; Chen et al., 2016). Furthermore convolutional neural networks, like most machine learning algorithms, are data-hungry and thrive when they have access to large datasets from which they can learn these feature mappings. An unsupervised form of CNN, convolutional autoencoders (CAE) are adept at extracting complex features from images in the absence of large labeled datasets (Ghasedi Dizaji et al., 2017; Wang et al., 2020; Song et al., 2020). CAE have already found success in complex geoscience image classification problems including facies classification (Qian et al., 2018), rock image segmentation (Karimpouli and Tahmasebi, 2019) and recently (Sun et al., 2020), which makes them an appealing candidate technology for extracting key features of microseismic data here.

Motivated by (1) the treatment of moment tensor inversion as an image classification problem and (2) the ability of neural networks to efficiently process large datasets like those supplied by the DAS, we set out with the goal of designing a CAE to extract features related to source mechanisms from DAS-microseismic images. We then make use of clustering algorithms and a generative adversarial network (GAN) to make predictions about the source mechanism based on the features extracted from our input images. The theoretical development, network training, and initial tests are performed on synthetic data which allows us to validate the accuracy of our results. An extension to field data shows promise in using deep-CNN architecture for moment tensor estimation.

## **MODELLING DISTRIBUTED ACOUSTIC SENSING DATA**

In the absence of labeled field data, which is especially difficult to obtain for DAS data, it can be challenging to optimize neural network architecture and validate the accuracy of our results. For these reasons, we develop the methods in this paper on synthetic data, allowing for inferences into how network hyperparameters influence our results, and validation of the quality of our results. Once a successful network architecture and workflow is created, these methods will be extended to field data examples. This section discusses the simulation of synthetic DAS-microseismic data that will be used as input to the neural network.

As discussed above, the rigidity of the optical fibers in DAS systems results in dis-

tributed fiber optic sensors which are sensitive only to the component of strain along their tangential direction (Kuvshinov, 2015). Modeling DAS signals then requires a geometric model of the fiber that contains information about the fiber tangents, and the locations where these tangents are samples (Innanen, 2017). The computation of the tangent requires transformation from field coordinates ( $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ ) to a local Frenet-Serret (Serret, 1851; Frenet, 1852) coordinate system in the tangent,  $\hat{\mathbf{t}}(s)$ , its associated normal  $\hat{\mathbf{n}} = \mathbf{n}(s)/|\mathbf{n}(s)|$ , with  $\mathbf{n}(s) = \hat{\mathbf{t}}(s)/ds$ , and the binormal  $\hat{\mathbf{b}}(s) = \hat{\mathbf{t}}(s) \times \hat{\mathbf{n}}(s)$ , where  $s$  is the arc-length along the fiber to the point at which we compute the local Frenet-Serret coordinate system.

The DAS response is computed through the projection of the strain tensor onto the local Frenet-Serret coordinate system at each point along the fiber, for each time sample. For a fixed time  $t$ , the projected strain tensor is,

$$\epsilon_{tnb}(s) = \mathbf{R}(s)\epsilon_{xyz}\mathbf{R}(s)^T \quad (1)$$

where  $\mathbf{R}(s)$  is a rotation operator transforming from field coordinates to Frenet-Serret coordinates,

$$\mathbf{R}(s) = \begin{bmatrix} \hat{\mathbf{t}}(s) \cdot \hat{\mathbf{x}} & \hat{\mathbf{t}}(s) \cdot \hat{\mathbf{y}} & \hat{\mathbf{t}}(s) \cdot \hat{\mathbf{z}} \\ \hat{\mathbf{n}}(s) \cdot \hat{\mathbf{x}} & \hat{\mathbf{n}}(s) \cdot \hat{\mathbf{y}} & \hat{\mathbf{n}}(s) \cdot \hat{\mathbf{z}} \\ \hat{\mathbf{b}}(s) \cdot \hat{\mathbf{x}} & \hat{\mathbf{b}}(s) \cdot \hat{\mathbf{y}} & \hat{\mathbf{b}}(s) \cdot \hat{\mathbf{z}} \end{bmatrix}. \quad (2)$$

The tensor  $\epsilon_{xyz}$  is the field coordinate strain tensor at point  $s$  along the fiber with components,

$$\epsilon_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad (i, j) \text{ ranging over } \mathbf{x}, \mathbf{y}, \mathbf{z} \quad (3)$$

and  $\epsilon_{tnb}$  is the Frenet-Serret coordinate strain tensor at point  $s$ . The DAS response to a wavefield at time  $t$  and point  $s$  is the  $\epsilon_{tt}$  component of  $\epsilon_{tnb}$ ,

$$\begin{aligned} \epsilon_{tt} = & (\hat{\mathbf{t}} \cdot \hat{\mathbf{x}})^2 \epsilon_{xx} + 2(\hat{\mathbf{t}} \cdot \hat{\mathbf{x}})(\hat{\mathbf{t}} \cdot \hat{\mathbf{y}}) \epsilon_{xy} + 2(\hat{\mathbf{t}} \cdot \hat{\mathbf{x}})(\hat{\mathbf{t}} \cdot \hat{\mathbf{z}}) \epsilon_{xz} \\ & + (\hat{\mathbf{t}} \cdot \hat{\mathbf{y}})^2 \epsilon_{yy} + 2(\hat{\mathbf{t}} \cdot \hat{\mathbf{y}})(\hat{\mathbf{t}} \cdot \hat{\mathbf{z}}) \epsilon_{yz} + (\hat{\mathbf{t}} \cdot \hat{\mathbf{z}})^2 \epsilon_{zz}. \end{aligned} \quad (4)$$

In practice, DAS does not supply point measurements of the strain described by equation (4). Instead, in order to improve the signal-to-noise ratio, measurements are averaged over a length of fiber known as the gauge length. This is accomplished interferometric process that delays the signal from one portion of the fiber so that the interference pattern analyzed by the DAS system is constructed of signals from two separated portions of fiber (Masoudi et al., 2013). The spatial resolution achieved through this process is currently on the order of ten meters. Due to the gauge length, DAS supplies an average measurement of the strain over the length ( $L$ ),

$$d(s) = \int_{-L/2}^{L/2} W(s - s', L) \epsilon_{tt}(s') ds' \quad (5)$$

where,

$$W(s, L) = \begin{cases} f(s), & -L/2 < s < L/2 \\ 0, & \text{otherwise} \end{cases}, \quad (6)$$

Here we let  $f(s) = 1/L$  so that the DAS response is simply the mean value of the sensitivities explored by the fiber assuming constant strain over the gauge length.

### Analytic expressions for DAS strain response from moment tensor sources

The direct arrivals in microseismic data are a rich source of information about the source mechanism. Provided we are able to sample a sufficient solid angle of this radiated energy, features encoded in the direct arrivals such as the relative polarity and relative amplitude of the P-wave and S-wave mode can be used to make estimates about the source mechanism, in the form of a moment tensor, that generates the seismic energy. Information about the moment tensor can then be used to make inferences about the in-situ stress state, type of fracturing, occurrence of new fractures and reactivation of existing fractures, and can help optimize future treatments. Analytic modeling methods (e.g. Aki and Richards, 2002) are an efficient way of generating large datasets of the direct arrival data generated by a general moment tensor,

$$\mathbf{M} = \begin{bmatrix} M_{xx} & M_{xy} & M_{xz} \\ M_{yx} & M_{yy} & M_{yz} \\ M_{zx} & M_{zy} & M_{zz} \end{bmatrix} \quad (7)$$

where  $M_{ij} = f_i/\delta x_j$  represents a force couple pointing in the  $i^{th}$ -direction, separated by a small distance  $\delta x_j$  in the  $j^{th}$ -direction.

As stated in the introduction, deep learning algorithms are data-hungry, and generally require large datasets to successfully learn the relationship between input data and network outputs. If we hope to train a deep neural network to extract features associated with moment tensor source mechanics from DAS-microseismic data, then we require a method for efficiently generating large datasets for training. Coupling the ideas that the direct arrivals contain significant information about the source mechanism and that we require an efficient modeling algorithm to generate a sufficiently large dataset, we use an analytic approach for modeling the strain produced by a general moment tensor source, and couple it to a geometric model of the fiber to produce the DAS-microseismic data. While this method is unable to consider the effect of complex geology on our data, it provides a means of efficiently producing the large datasets required to examine the feasibility of using deep learning techniques for estimating properties of moment tensor source mechanisms.

Aki and Richards (2002) developed expressions for the analytic displacement at point  $\mathbf{x} = (x_1, x_2, x_3)$  generated by a general moment tensor source  $\mathbf{M}$ . The far-field displacement, representing the P-wave and S-wave motion excited by a moment tensor source, at a sufficient distance from the source such that the near-field can be ignored (a condition typically met in microseismic monitoring) is,

$$u_i = \frac{1}{4\pi\rho\alpha^3} \frac{m\gamma_i}{r} \dot{s}(t - r/\alpha) - \frac{1}{4\pi\rho\beta^3} \frac{m\gamma_i - \gamma'_i}{r} \dot{s}(t - r/\beta) \quad (8)$$

where  $\gamma_i = x_i/r$  is the directional cosine unit vector,  $r$  is the distance from the source,  $\alpha$  and  $\beta$  are the P-wave and S-wave velocity, and  $\dot{s}(t - r/v)$  is the first time derivative of the source time function propagating with velocity  $v$ . In equation (8) we have made use of the simplifying expressions,  $m = \gamma_p M_{pq} \gamma_q$  and  $\gamma'_i = \delta_{ip} M_{pq} \gamma_q = M_{iq} \gamma_q$ .

To model DAS-microseismic data, we require expressions similar to equation (8) for the strain radiated by moment tensor sources. The strain tensor in equation (3) suggests that this requires derivatives of the form  $\partial u_i / \partial x_j$ . Therefore, taking spatial derivatives of equation (8), and keeping only the resulting far-field terms (i.e.  $\epsilon_{ij} \propto r^{-1}$ ) provides the analytic expression for the far-field strain tensor at point  $\mathbf{x}$  due to a moment tensor source  $\mathbf{M}$ ,

$$\epsilon_{ij} = -\frac{1}{4\pi\rho\alpha^4} \frac{m\gamma_i\gamma_j}{r} \ddot{s}(t - r/\alpha) + \frac{1}{4\pi\rho\beta^4} \frac{m\gamma_i\gamma_j - \Gamma_{ij}}{r} \ddot{s}(t - r/\beta) \quad (9)$$

where  $\Gamma_{ij} = (\gamma_i\gamma'_j + \gamma'_i\gamma_j)/2$ . The DAS response is given by equation (5) with  $\epsilon_{tt}$  given by using equation (9) for the strain components in equation (4).

## Radiation patterns and their influence on the data

The numerators in the two terms of equation (9) are the far-field P-wave and S-wave radiation patterns. They are a function of the directional cosine ( $\gamma_i$ ) and the moment tensor source through the terms  $m$ , and  $\Gamma_{ij}$ . These radiation patterns describe the portion of seismic energy radiated in a given direction  $\gamma_i$ , by a moment tensor source  $\mathbf{M}$ . Figure 1 plots the  $\epsilon_{xx}$  component of these radiation pattern for the P-wave in 1a-1d, and for the S-wave in 1e-1h, for four sources  $\mathbf{M}_{\text{EXP}}$ ,  $\mathbf{M}_{\text{TC}}$ ,  $\mathbf{M}_{\text{CLVD}}$ , and  $\mathbf{M}_{\text{DC}}$ . Let  $\mathbf{M}_{\text{T}} = \{M_{xx}, M_{xy}, M_{xz}, M_{yy}, M_{yz}, M_{zz}\}$  be the six independent components of moment tensor  $\mathbf{M}$  of source type T, then the explosive ( $\mathbf{M}_{\text{EXP}}$ ), tensile crack ( $\mathbf{M}_{\text{TC}}$ ), compensated linear vector dipole ( $\mathbf{M}_{\text{CLVD}}$ ), and double couple ( $\mathbf{M}_{\text{DC}}$ ) sources used to generate the radiation patterns in Figure 1 are  $\mathbf{M}_{\text{EXP}} = \{1, 0, 0, 1, 0, 1\}$ ,  $\mathbf{M}_{\text{TC}} = \{2, 0, 0, 3, 0, 2\}$ ,  $\mathbf{M}_{\text{CLVD}} = \{-1, 0, 0, 2, 0, -1\}$ , and  $\mathbf{M}_{\text{DC}} = \{0, 1, 0, 0, 0, 0\}$ .

Figure 1 suggests that moment tensor sources radiate energy in a manner that is diagnostic of the source type. For example the explosive and tensile crack sources radiate P-wave energy in a similar manner for  $\epsilon_{xx}$  strain as shown by comparing Figure 1a and 1b, but radiate S-wave energy very differently as shown by comparing Figure 1e and 1f. Con-

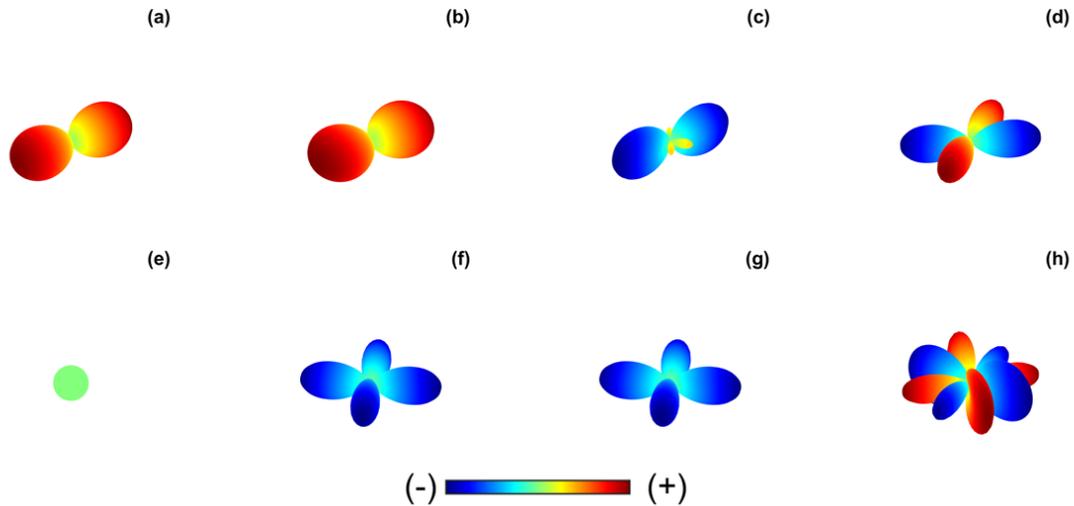


FIG. 1. The  $\epsilon_{xx}$  component of the analytic strain radiation patterns for four source types. P-wave radiation patterns for (a)  $\mathbf{M}_{\text{EXP}}$ , (b)  $\mathbf{M}_{\text{TC}}$ , (c)  $\mathbf{M}_{\text{CLVD}}$ , (d)  $\mathbf{M}_{\text{DC}}$ . S-wave radiation patterns for (e)  $\mathbf{M}_{\text{EXP}}$ , (f)  $\mathbf{M}_{\text{TC}}$ , (g)  $\mathbf{M}_{\text{CLVD}}$ , (h)  $\mathbf{M}_{\text{DC}}$ .

versely tensile crack and CLVD sources radiate S-wave energy in a similar manner for  $\epsilon_{xx}$  strain as shown by comparing Figure 1f and 1g, but radiate P-wave energy very differently as shown by comparing Figure 1b and 1c. Assuming our acquisition geometry records a sufficient aperture of both the P-wave and S-wave radiation pattern, then the recorded DAS-microseismic first arrivals should contain features diagnostic of the source type that generated the.

Consider a horizontal well drilled perfectly on azimuth such that it is parallel to the minimum horizontal stress  $\sigma_h$ . This represents a favorable scenario that allows fractures to propagate deep into the reservoir in a direction transverse to the well, and open parallel to the well (Addagalla et al., 2018). Without loss of generality assume that  $\sigma_h$  is parallel to the x-direction, so that a straight fiber in the horizontal portion of this well is sensitive only to the  $\epsilon_{xx}$  component of strain, represented by the radiation patterns in Figure 1. Suppose a propagating fracture excites seismic energy at the location  $x_s = \{3300, 400, 9300\}$  ft in a homogeneous medium with a P-wave velocity of 15000 ft/s and an S-wave velocity of 9000 ft/s. The energy radiated by this fracture is recorded by a straight DAS fiber in a horizontal well aligned with the minimum stress, at a depth of 9300 ft. Figure 2a plots a schematic representation of the source and fiber locations, and Figure 2b-2e plots the energy recorded by the DAS fiber for the four source types considered in figure 1. Figure 2b-2e shows that the source type is encoded within distinct features in the data. As an example, the tensile crack type source produces balanced P-wave and S-wave amplitude ratios in Figure 2c, a feature diagnostic of these source types (Eaton et al., 2014). Another example is the distinct polarity pattern encoded in the S-wave energy of Figure 2e radiated by double couple types sources. In general, microseismic field data is more complex and events are often generated by closely related moment tensors that represent distinct fracture mechanisms, but result in more subtle variations in the data than the ones shown in Figure 2. Motivated by the idea that diagnostic features in the data provide clues to the fracture mechanism that generated it, but that these features may be complex, the next sections of the paper set out to develop neural networks that learn how to extract these features, and then analyze them to estimate

the source mechanism encoded in the data.

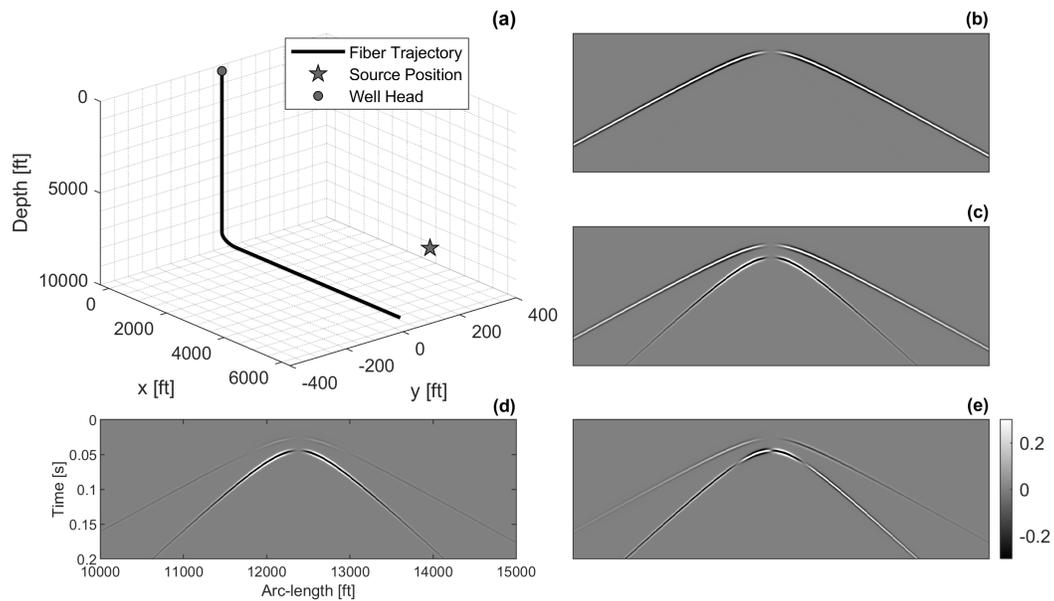


FIG. 2. (a) Schematic representation of the fiber trajectory shown by the black line, and the source location indicated by the red star. Data recorded by the fiber for (b)  $M_{EXP}$ , (c)  $M_{TC}$ , (d)  $M_{CLVD}$ , and (e)  $M_{DC}$ .

## DEEP NEURAL NETWORK FOR FEATURE EXTRACTION

Deep learning algorithms generally fall into two main categories, unsupervised and supervised methods. Supervised methods require access to large labeled datasets in which the output and inputs are known. Their goal is to learn a function that maps the inputs to the outputs. Once trained they are able to label new input data and typically compute the probability of that label being accurate. DAS being a relatively new technology means that we do not have access to the large labeled datasets, such as those provided by conventional MTI, crucial to the development of successful supervised learning networks. In light of this, this paper focuses on the development of unsupervised learning techniques that learn how to extract structure in the data, in the absence of labeled data.

The direct arrivals in microseismic data are rich with information about the source mechanism. This information is encoded in features like the relative P-wave and S-wave amplitudes (Eaton et al., 2014), first arrival polarity (e.g. Hardebeck and Shearer, 2002), and the polarity differences between P-wave and S-wave energy. Making predictions about the presence or absence of features within input microseismic records in the form of images, and then extracting those features offers a method for forming estimates of DAS-microseismic source mechanisms. Convolutional neural networks (CNN) are a proven, and powerful technology for complex and nonlinear feature extraction from image data, finding applications in wide ranging disciplines (e.g. Bin et al., 2012; Hertel et al., 2015; Spanhol et al., 2016; Lopes and Valiati, 2017), with many recent applications in geophysical and geological problems (Qian et al., 2018; Wang et al., 2020; Sun et al., 2020). Deep convolutional neural networks take input data in the form of images, pass them through multiple

layers, known as hidden layers, and compute a feature space representation of the input data. They can be thought of as a method of compressing our generally high dimensional input data, to a low dimensional feature space that represents the input data by only its most salient features.

In convolutional neural networks a filter of a prescribed size ( $N_p \times N_p$ ) is overlain on the input image, where  $N_p$  is the number of  $x$  and  $y$  pixels in the filter. The pixels overlain by the filter are weighted by the filter coefficients,  $w_i$ , summed and passed through a nonlinear activation function  $g$ . The output of this process  $\hat{y}_j$  constitutes the  $j^{th}$  value in the feature map for that filter. Mathematically this process is expressed as,

$$\hat{y}_j = g \left( \sum_i^{N_p^2} w_i y_i \right) \quad (10)$$

where  $y_i$  is the value of the  $i^{th}$  pixel overlain by the filter. This filter slides across the image, in a process akin to convolution, and the feature map value is computed for each position of the filter. Generally, many filters are computed for each layer, and a single output feature map is computed for each filter. Each filter is tasked with detecting the presence of a single feature within the input image, and the feature map describes the spatial presence of features in the input image. Deeper layers follow the same process, but combine all feature maps from the previous layer as their input, allowing them to combine relatively simple features from the previous layer to make predictions of increasingly complex and abstract features as the network depth increases.

During training the filter weights are computed by minimizing a loss function which measures how well the network is mapping the input images to some desired output. For example in classification problems the loss function is typically a cross-entropy loss that measures how well a network correctly labels input data. Minimization creates a network that is tuned for extracting features required for accurate classifications. Many of the early uses of CNN were directed at supervised problems in which the input and outputs were known, and the networks goal was to learn the function, in the form of the filter weights, required to map the inputs to the outputs (e.g. Hasegawa et al., 1991; Lo et al., 1995). In the problem we are considering, we do not have the large labeled dataset required for this, and so we turn our attention towards an augmented form of the CNN tuned for unsupervised learning.

### Convolutional autoencoders (CAE) for unsupervised feature extraction

Convolutional autoencoders (CAE) are an unsupervised learning algorithm, based on the idea of CNN, that are useful in situations where we want to extract features from image data, but do not have access to a large labeled dataset for training. The question in this situation is, how do we extract a meaningful feature space representations of our data, when we are unsure what that representation may look like? CAE networks use an augmented form of standard CNN to solve this challenge. Figure 3 shows the general network architecture for a CAE, which consists of two main components, an encoder and a decoder. The

encoder is a standard deep CNN, and its main goal is to extract features from an input DAS-microseismic image, outputting the salient feature space representation of the image as an  $N$ -dimensional vector. The decoder uses this feature space representation as its input, and passes it through a deconvolutional neural network, which is a mirror image of the encoder, composed of transposed convolutions. Its goal is to form an accurate reconstruction of the original input image, given a feature space representation for that image. With these goals in mind, an appropriate loss function for a CAE architecture should track some measure of the discrepancy in the input images and those generated by the decoder. In this study, the loss function we minimize through training of the CAE is the L2 norm of the difference in the input and reconstructed images,

$$\phi = \sum_i^{N_I} \|\hat{\psi}_i(w) - \psi_i\|_2^2 \quad (11)$$

where  $\hat{\psi}_i(w)$  is the  $i^{\text{th}}$  reconstructed image as a function of the weights  $w$ ,  $\psi_i$  is  $i^{\text{th}}$  input image, and  $N_I$  is the number of images in the training dataset.

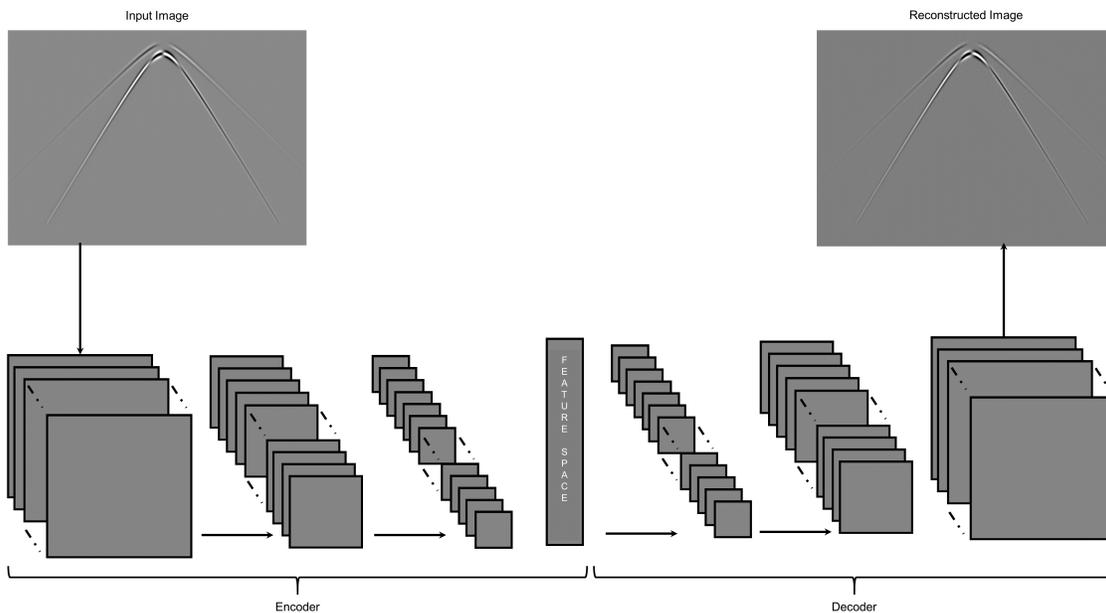


FIG. 3. Overview of network architecture used in convolutional autoencoders. The squares in each layer represent a single feature map.

During training, the loss function in equation 11 is minimized with respect to the model weights  $w$ . As training progresses, the CAE learns the weights that produce an optimal mapping from input images to reconstructed images. For the decoder to be successful in its reconstruction, the feature space representation must contain sufficient information about the important features in the input image. Therefore, for the loss function to be minimized the encoder must learn the weights that allow it to extract the  $N$  most salient features from input images. When training is complete the decoder is detached from the network, resulting in an encoder network that maps input DAS-microseismic images to their salient

features. In the next sections we will discuss methods for using these features to make predictions about the moment tensor source that generated the data.

### Analyzing features for source mechanism information

Training the CAE produces a function that maps a high-dimensional input image containing tens-of-thousands of pixels, to a low-dimensional feature space representation containing the most salient features in the image. The feature space is an  $N$ -dimensional vector that describes the presence and absence of certain features in the image, and is diagnostic of the class of the input image. However, further processing is generally required to understand the information contained in the feature space. In this paper, we take two approaches to realize this goal. The first is a common approach to understanding structure in feature space representations, and is known as clustering. The second approach, is more novel and consists of training a secondary deep neural network, known as a generative adversarial network (GAN), that learns how to map the feature space representation to an estimation of the moment tensor.

#### *Clustering*

Clustering is a commonly used unsupervised learning method for unveiling relationships in complex data. Many algorithms exist for clustering, but all share the same goal of grouping data, such as feature vectors, into groups such that items within the same group are more similar to each other than they are to feature vectors in any other group (e.g. Jain et al., 1999). Many clustering algorithms group points based on some measure of their distance to other points in the feature space. For example, one of the most common methods for clustering is the k-means algorithm. In this algorithm a predetermined number of clusters,  $k$ , and points are then assigned to the cluster associated with their closest centroid  $\mu_i$ . K-means comes with the limitations of requiring a predetermined number of clusters, something we may not know, and its preference for spherical clusters of approximately equal data density (Nagy, 1968; Jain et al., 1999).

In this paper we use a more sophisticated clustering algorithm, introduced by Ester et al. (1996), known as density-based spatial clustering of applications with noise (DBSCAN). Figure 4 highlights the central idea behind this algorithm. A Euclidean search distance, and minimum number of points are specified, and the algorithm cycles through every point in the dataset searching the neighborhood of that point for the presence of other points. If a point has the minimum number of points in its neighborhood then it is deemed a core point for the cluster to which it belongs (the gray star in Figure 4 is a core point). If a point does not include the minimum number of points in its neighborhood but is reachable by a core point (that is, it lies in the neighborhood of a core point), it is deemed a border point of the cluster to which its associated core point belongs (the squares in Figure 4 are border points). The last class of points, deemed noise points, are those points that do not have the minimum number of points in their neighborhood, and are not reachable by a core point (the triangles in Figure 4 are noise points).

One of the main benefits of DBSCAN is that it does not require a predetermined number of clusters. Instead, it decides how many clusters exist in the data based on how many

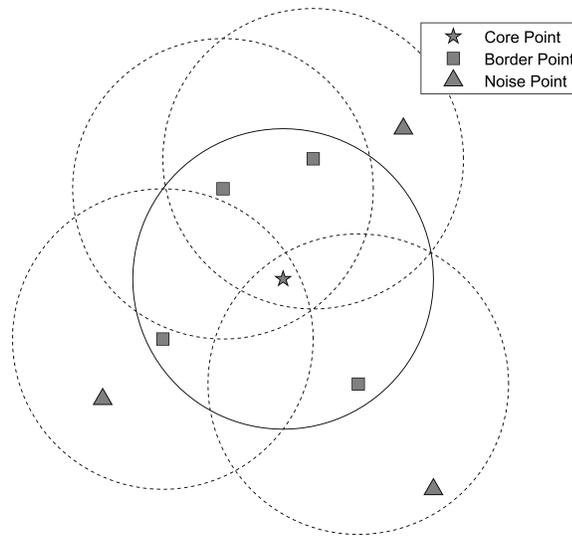


FIG. 4. Schematic overview of DBSCAN clustering algorithm. The gray star represents a core point, the squares border points, and the triangles noise points. The solid circle is the neighborhood of the core point, and the dashed circles are the neighborhood of the border points. The minimum number of points in this example is set to five.

unique core points it finds. It is also not limited to isotropic clusters, but is limited by preferring clusters with similar data density. Another benefit of DBSCAN is that it can cluster nonlinear data, allowing it to be more successful in instances where simpler algorithms like k-means produce poor clusters.

Once the feature space representations are clustered, the data in each cluster can be examined to understand what kind of data each cluster represents. For example, examination of the data may reveal that all of the data in one cluster share the same S-wave polarity patterns that were shown to be diagnostic of double couple type events. Clustering on the feature space representations provides a method for grouping data based on its similarity, and then investigating that similarity for source information. If the features extracted by the CAE carry important source information, then data in the same cluster should share similar source mechanisms.

#### *Generative adversarial networks (GAN) for moment tensor estimation*

The second method we examine for extracting source mechanism information from the feature space representation is a technology known as generative adversarial networks (see figure 5 for a schematic representation). These networks are comprised of two neural networks which compete against each other in a zero sum minimax game (Goodfellow et al., 2014). In our approach, the first network, known as the generator (G), takes feature space representations, output from the CAE, and tries to map them to a good estimate of the true data distribution, consisting of moment tensor labels. Put another way, the generator's goal is to produce believable moment tensor labels for an input feature space representation.

The second network, known as the discriminator ( $D$ ), takes in a feature space representation and moment tensor label pair and assigns a probability the label belongs to the true data distribution. Its goal is to correctly differentiate true labels and those generated by network  $G$ .

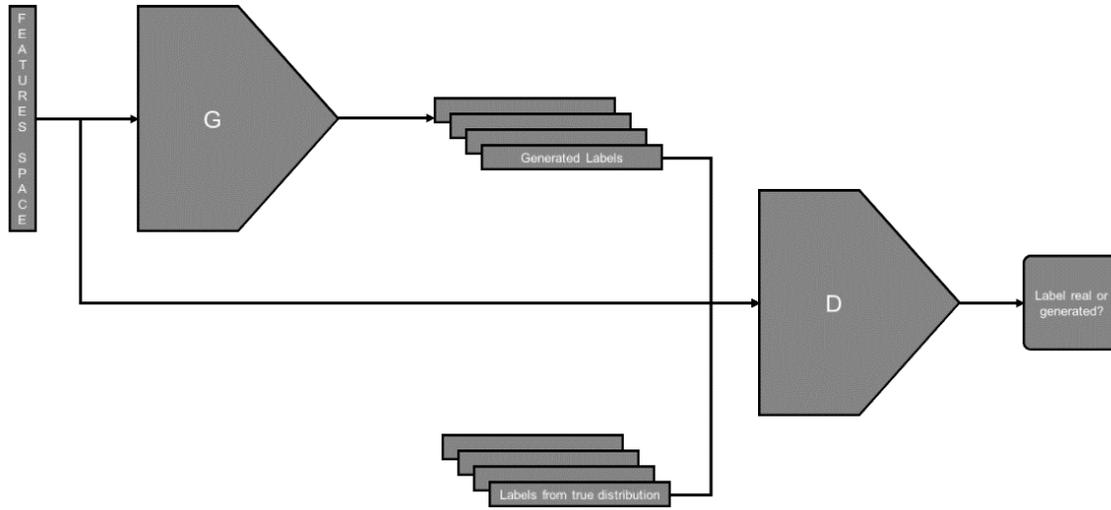


FIG. 5. Schematic representation of a generative adversarial network for predicting moment tensor labels from feature space representations.  $G$  is the generator network, and  $D$  is the discriminator network.

Training progresses by optimizing the cross-entropy loss function (Goodfellow et al., 2014),

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} \log [D(x)] + \mathbb{E}_{z \sim p_z(z)} \log [1 - D(G(z))] \quad (12)$$

where  $D(x)$  is the probability that  $x$  came from the data distribution,  $G(z)$  is the sample generated by the generator with input  $z$ ,  $\mathbb{E}_{x \sim p_{\text{data}}(x)}$  is the expectation over all data samples, and  $\mathbb{E}_{z \sim p_z(z)}$  is the expectation of all inputs to the generator  $z$ . Early in the training process when the generator has not yet learned a mapping of input data  $z$  to a good estimate of the true data distribution, the discriminator is good at identifying true data labels and assigns low probabilities to samples created by the generator network. To minimize equation (12), the generator must learn how to produce labels that better match the true data distribution, convincing the discriminator to assign generated labels a high probability of belonging to the true distribution. As this occurs the maximization of equation (12) requires that the discriminator becomes more adept at detecting generated labels, which in turn forces the generator to improve in its task. This back-and-forth competition between the networks results in a generator that produces labels for the input feature space representations that are a good representation of the true labels.

Once training is complete, the generator network is extracted, providing a network that is capable of forming accurate predictions of the label associated with a feature space representation. In our study we will test two candidates for the type of labels we try to predict,

the Hudson space representation of the moment tensor (Hudson et al., 1989), and the six independent components of the moment tensor. This may seem to be problematic from the perspective of requiring information about the true moment tensors associated with the data. However, the approach we take is to learn the mapping from feature space representation to moment tensor label on synthetic data and then use the synthetically trained network to make predictions of the moment tensor on field data. Given sufficient training data, with a sufficient variation in the moment tensor, this should provide a physics driven catalog for mapping features to moment tensor predictions. Ideally as the use of DAS becomes more prevalent in field applications, a conventional MTI could be computed on small subset of the data, and the supervised techniques we develop here could readily accommodate these field data labels. It is important to note that identification of the clusters is unsupervised but their labeling by moment tensor source was supervised task.

The above problem of mapping data in the form of feature space representations to moment tensor label can also be achieved with a standard feed-forward neural network using supervised learning. However, generative adversarial networks come with some potential advantages. Chief among these is the distinct difference in the way the GAN and standard neural networks learn. Neural networks take data, here CAE extracted features, and predict a model that maps the data to an estimate of the moment tensor. This is akin to an inverse problem in which the network can only leverage the labeled data. Conversely GAN learn in two distinct phases. The generator takes in data and learns how to produce moment tensor estimates, but it does so independently of existing moment tensor labels, instead learning how the data and moment tensor relate by gaining an understanding of the behavior of the GAN discriminator (Goodfellow et al., 2014). Thus the generator training is unsupervised with respect to the moment tensor labels. The discriminator learns how the generator behaves in forming moment tensor predictions, and in doing so learns how the data and moment tensor relate through a verification problem. In other words the discriminator learns by verifying that the input data (CAE features) and the moment tensor its been supplied form a valid pair, which is more akin to a forward problem. This suggests the GAN may be able to leverage smaller datasets since it learns about the physics of the problem through simpler forward problem instead of the inverse problem the neural network must learn through.

Generative adversarial networks come with other advantages as well. It is not always trivial to determine and effective model-space loss function when using neural nets. GAN circumvent this issue by using the data driven loss function in equation (12). Another potential benefit arises if a small subset of labeled field data is used to augment the synthetic data. In this case the discriminator can determine if systematic differences exist between the field and synthetic data. If discrimination occurs on the basis of data alone, and the discriminator assigns a low probability to correct label-field-data pairs, then it suggest a difference exists between field data and modeled data. Thus the discriminator can be used to gain information that can lead to better modeling or feature extraction. The distinct advantages that GAN supply over simpler neural networks motivates their use in this paper.

## NUMERICAL EXAMPLES

### Analytic modeling dataset

To test the feasibility of using the above deep learning tools for extracting meaningful feature space representations from DAS-microseismic data, and then use those representations to form predictions of the underlying source mechanism, we generate an analytically modeled test dataset. This dataset consists of 10,000 events with random moment tensors, constrained by being either compensated linear vector dipole, double couple, or tensile crack dominant. The Hudson space (Hudson et al., 1989) distribution of these sources is plotted in Figure 6. Data were generated using equation (9), where the P-wave velocity ( $\alpha$ ), S-wave velocity ( $\beta$ ), and source position were allowed to vary for each event, with  $\alpha = 15000 \pm 1000$  ft/s,  $\beta = 9000 \pm 600$  ft/s,  $x_s = 3500 \pm 200$  ft, and  $y_s = 400 \pm 200$  ft. A straight DAS fiber with the geometry of the well in Figure 2 is used, that is with a vertical depth of 8800 ft, a total depth of 9300 ft, and a horizontal length of 5700 ft. To simulate the DAS fiber response, the analytically propagated wavefield is projected onto the fiber using equations (4), (5), and (6). Figure 7 plots six randomly selected event records from this dataset, as a representation of the input data.

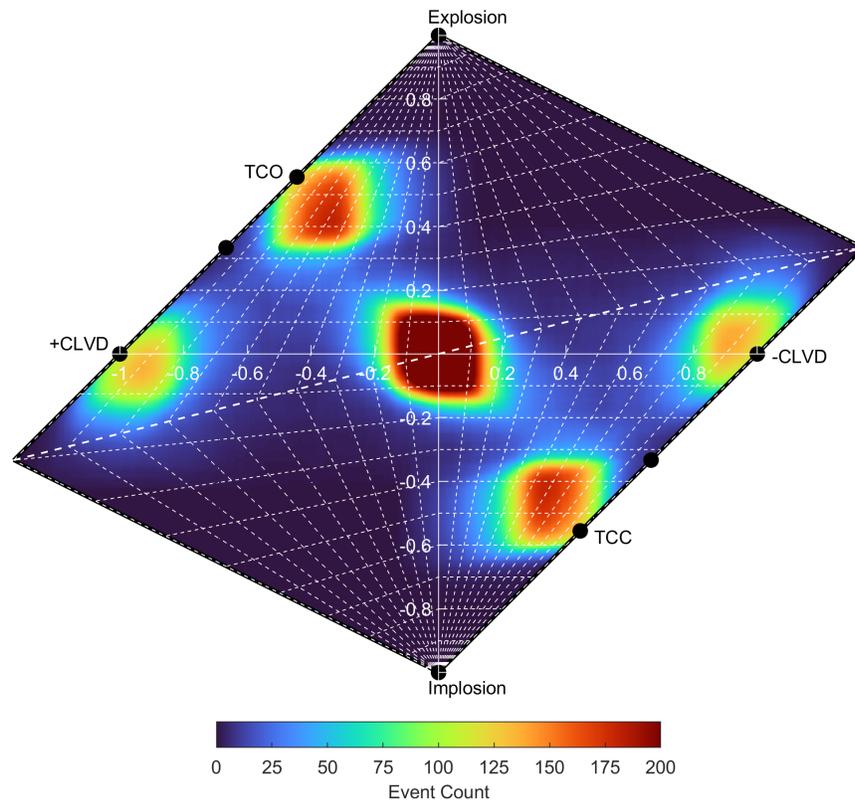


FIG. 6. Hudson space representation of the moment tensor source distribution in the input data set.

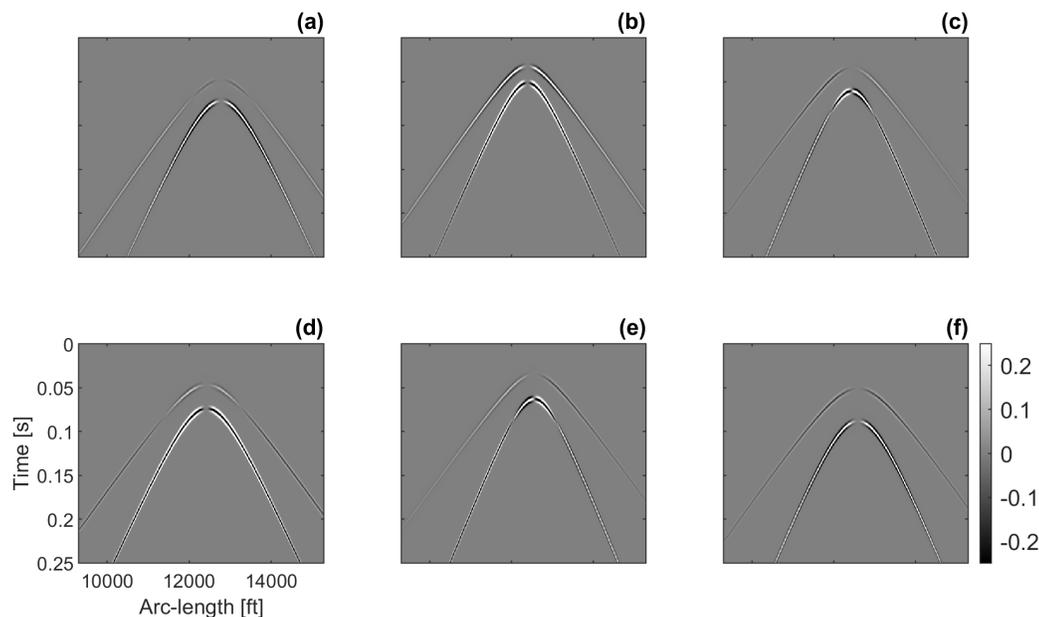


FIG. 7. Images randomly selected from the data distribution in Figure 6. (a) Image 1436, (b) image 1740, (c) image 4120, (d) image 5494, (e) image 6016, (f) image 8002.

### CAE network architecture and design

Successful networks require tuning of what are known as network hyperparameters. These hyperparameters define the network architecture and are typically not trainable. Important hyperparameters for the CAE architecture in Figure 3 are the number of convolutional layers, the number of filters in each layer, and the size of the filters in each layer, respectively influencing the nonlinearity that can be modeled, the number of features that can be learned, and how local those features are. Generally, the optimal hyperparameters are deduced through trial-and-error by running many simulations and monitoring how a change in a hyperparameter affects the loss functions minimization. The goal is to produce the network that is minimally complex while still meaningfully reducing the loss function. Overly complex networks are expensive to train, and run the risk of memorizing the input data, preventing them from generalizing to new data, while overly simple networks cannot learn complex relationships between data and features. The network architecture used in this study is summarized in table 1; the stride dictates how far the filter moves in pixels, strides smaller than the filter size result in overlapping filters.

Table 1. Convolutional autoencoder network architecture

| Layer No.     | Activation | Number of Filters/Nodes | Filter Size | Stride |
|---------------|------------|-------------------------|-------------|--------|
| 1             | ReLU       | 32                      | 5x5         | 2      |
| 2             | ReLU       | 64                      | 5x5         | 2      |
| 3             | ReLU       | 128                     | 5x5         | 2      |
| Feature Space | tanh       | 7                       | N/A         | N/A    |

A batch normalization (Bjorck et al., 2018) is included between each of the convolu-

tional layers in the encoder to make training more stable. The optimizer chosen for this study was Adam with an initial learning rate of  $1 \times 10^{-5}$  due to its success in training deep CNN (Kingma and Ba, 2017). Rectified linear units (ReLU), is the activation function chosen for the convolutional layers, while the output layer in the decoder has a hyperbolic tangent (tanh) activation to force the reconstructed images to the same  $[-1,1]$  range as the input images.

### *Pre-processing of data*

Perhaps the most important hyperparameter is the feature space dimensionality. Similar to the other hyperparameters, we desire the minimally complex feature space that captures sufficient information about the input image, such that the decoder can successfully reconstruct the input images. Higher dimensional feature spaces lead to better image reconstructions but run the risk of learning less meaningful features. For example, the upper bound on the feature space dimensionality is the dimensionality of the input images, but in this case the input images are mapped to themselves, and every pixel (feature) in the input image is learned. Clearly, not every pixel will contain important information about the source mechanism. How do we then decide on the size of the feature space representation?

A reasonable starting range for the feature space dimensionality can be determined through domain specific knowledge and intuition. In our study, we are trying to capture features associated with a moment tensor that can be fully characterized by six independent components,  $M_{xx}$ ,  $M_{xy}$ ,  $M_{xz}$ ,  $M_{yy}$ ,  $M_{yz}$ , and  $M_{zz}$ . Thus the quantity we are trying to estimate contains six components, or features, that can influence the data. It is then hypothesized that the feature space will need to contain approximately six features in order to capture the variability in the input images related to the moment tensor sources. To test this hypothesis we will monitor the loss function over 100 epochs for feature space dimensionality ranging from three features to twenty-five features. The minimally complex feature space that sufficiently minimizes the loss function will be chosen for the feature space dimensionality. If multiple feature space dimensionalities produce similar reduction in the loss function, then the smallest feature space of that group will be selected for further testing.

Training the CAE on raw input data proved to be challenging. Figures 8a-8c plot three randomly selected input images, Figures 8d-8f plot the CAE predictions using ten features in the feature space, and Figures 8g-8i, plot the difference between input and reconstructed images. On the whole these predictions are relatively poor for what is expected from the CAE training process. Large errors exist in the prediction of the P-wave (arrows labelled B in figure 8), and the existence of erroneously predicted moveout (arrows labelled A in figure 8), and nodal locations is evident (arrows labelled C in figure 8). Even more discouraging is that the images in columns 2 and 3 of Figure 8, which represent images held back for validation, have many errors. This suggests that the algorithm is not generalizing well to data that was absent during training. It is possible that our hypothesis was incorrect, and more than ten features are required to characterize the input data. Figure 9 plots the same results as Figure 8 but now for a feature space defined by 25 features. While the predictions improve for all three input images, this example highlights a problem with using the raw data as an input. Comparison of Figures 8h and 9h, shows that the increased number of

features has allowed the CAE to make better predictions about the moveout. On the surface this may seem beneficial, however, we are interested in learning information about the source mechanism. Further increase in the feature space dimensionality would likely result in better predictions, however, it is apparent that the algorithm is now learning features associated with source location, and velocity structure, parameters that while important are not what we are interested in learning for this study. Their inclusion here acts to increase the complexity of the feature space, without providing us improved information about source mechanisms.

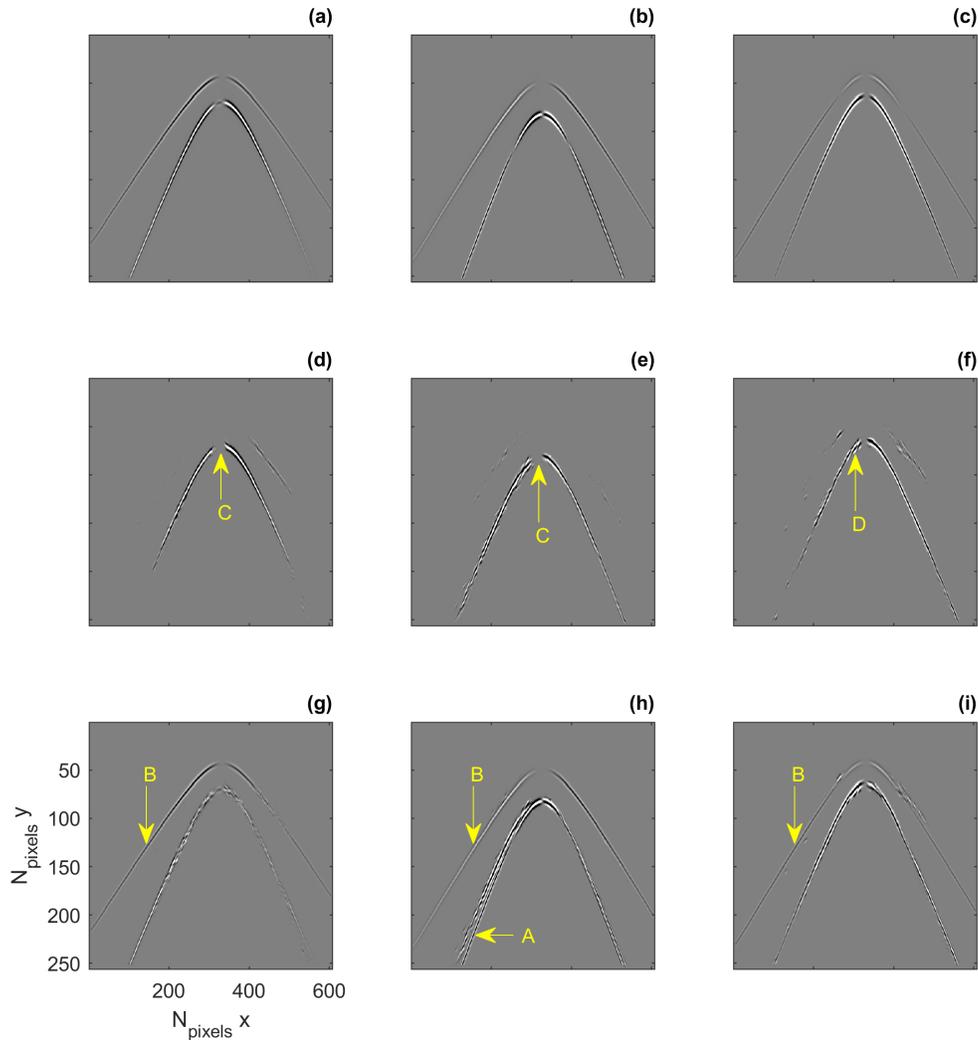


FIG. 8. (a)-(c) Three randomly selected input images, (d)-(f) CAE predictions of these images for ten features in the feature space, (g)-(i) differences between original and predicted images. Arrow A points to incorrectly predicted moveout, arrows B point to poorly predicted P-waves, arrows C point to poorly predicted nodes, and arrow D points to incorrectly predicted polarity.

Motivated by the above discovery that the raw data is unsuitable for our study, because it inherently contains information not related to the source mechanism, we set out to develop a pre-processing workflow that removes velocity and location information, allowing the CAE

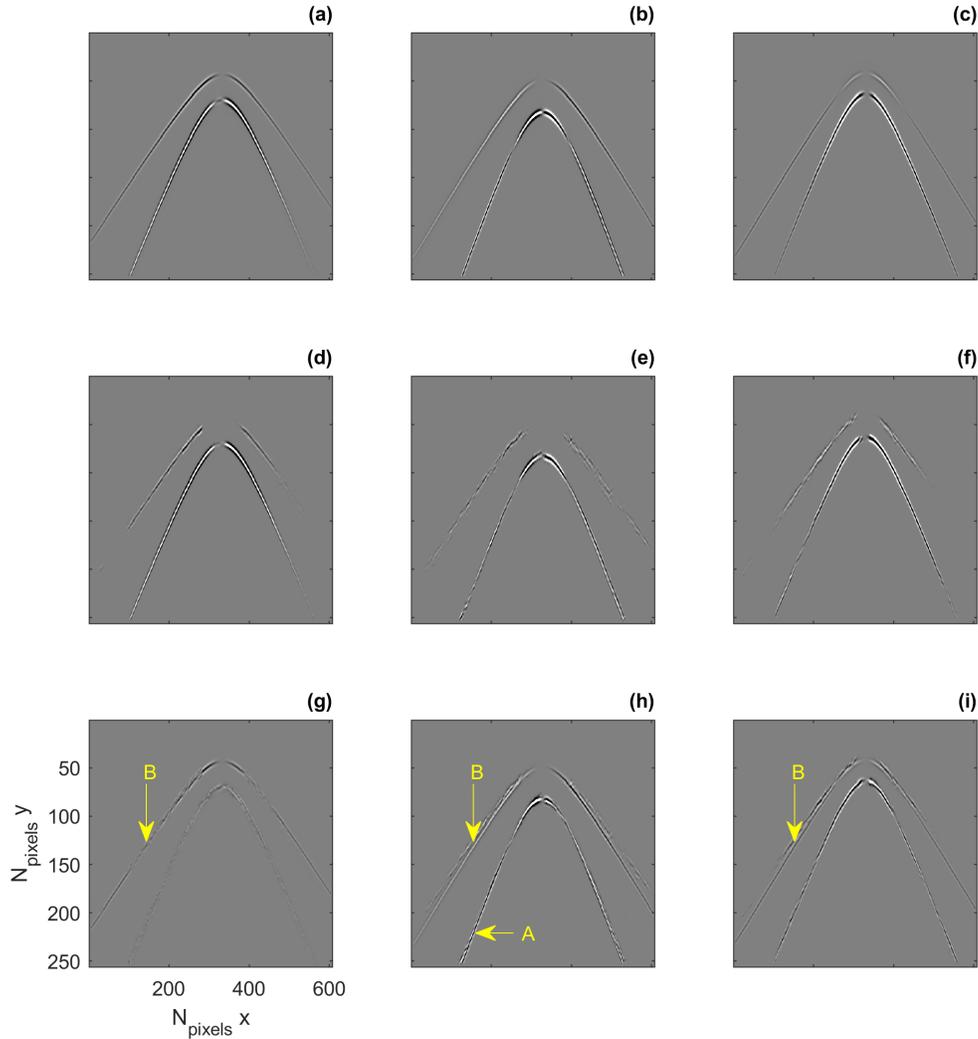


FIG. 9. (a)-(c) The same three input images from Figure 8a-8c, (d)-(f) CAE predictions of these images for twenty-five features in the feature space, (g)-(i) differences between original and predicted images. Arrow A points to the improvement in the predicted moveout compared to Figure 8h, arrows B point to poorly predicted P-waves.

to focus on source mechanism parameters. Figure 10 highlights the workflow we adopt in this paper. First, an apex is picked for the P-wave and S-wave modes. Next, NMO curves are fit to the P-wave and S-wave, and an NMO correction is applied to each separately, producing flattened P-waves and S-waves. When each NMO correction is applied, the unflattened event is filtered, producing the image in Figure 10c. The P-wave and S-wave are then extracted in windows centered on the P-wave and S-wave apexes, that is 32-pixels by 416-pixels in the x-direction and y-direction respectively. The windowed P-wave and S-wave images are then concatenated, producing the data that will be input to the CAE in Figure 10d.

We then proceed by examining the optimal feature space dimensionality on the pro-

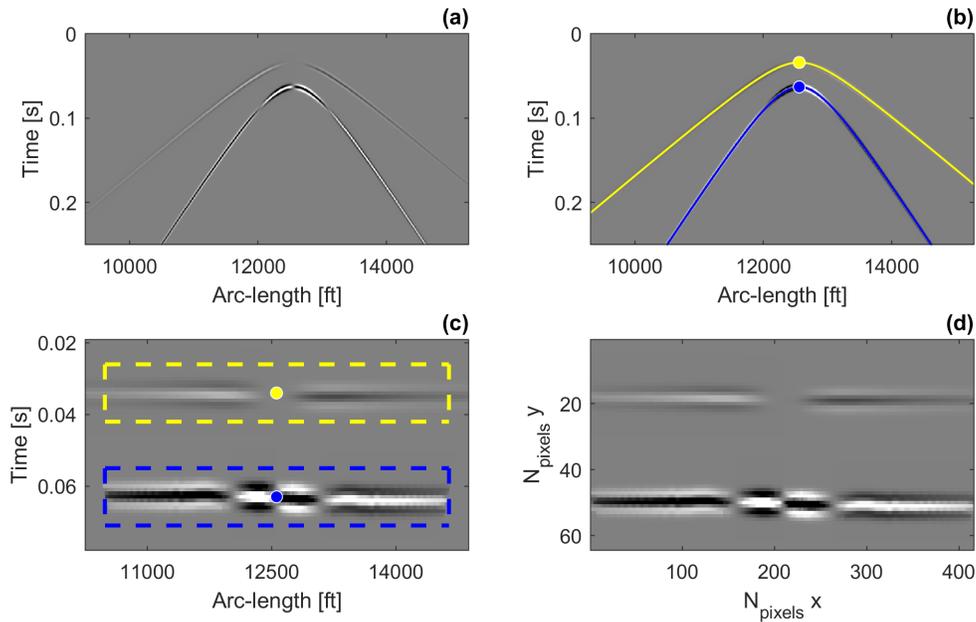


FIG. 10. Processing workflow to produce images that aid the CAE in learning source mechanism features. (a) Raw input data, (b) data with picked apex locations and NMO curves, (c) flattened data with window outlines for P-wave (yellow) and S-wave (blue) extraction, (d) processed image after NMO correction, windowing, and P-wave and S-wave concatenation.

cessed images. Figure 11 plots the loss function for 300 epochs as a function of the feature space dimensionality for three to twenty-five features. The loss function continues to decrease with the number of features in the feature space, until the feature space consists of approximately 6-7 features. After this point the loss function appears to saturate, and the inclusion of more features in the feature space acts to increase the complexity, without producing substantially more accurate predictions. For this reason, the optimal number of features required to characterize the input data is chosen as seven for this study.

In the training phase, 80% of the data (8000 images) are used for training and 20% are used for validation. The validation data is not used for training, but the loss is calculated on reconstructions using the validation data at the end of each epoch. Monitoring of the validation loss is important to minimize the risk of over-fitting our model. A lack of decrease, or an increase in the validation loss signals that the algorithm is no longer generalizing to data that is absent during training, and that the training data is being memorized, leading to an over-fit model. To mitigate this risk, an early stopping criteria is used such that if the validation loss does not decrease for 100 epochs, the training is stopped. Figures 12a-12e plot five randomly selected images from the processed dataset, Figures 12f-12j plot the CAE predictions using seven features, and Figures 12k-12o plot the differences in the processed images and their predictions. It is apparent that the processing has helped the CAE learn how to accurately predict the input images. Furthermore, the number of features required for success is similar to our hypothesis of six features.

While the accurate predictions in Figure 12 are encouraging, we need to analyze the features extracted from the input images, using the clustering and GAN methods discussed previously, to understand if they contain information about the source mechanism encoded

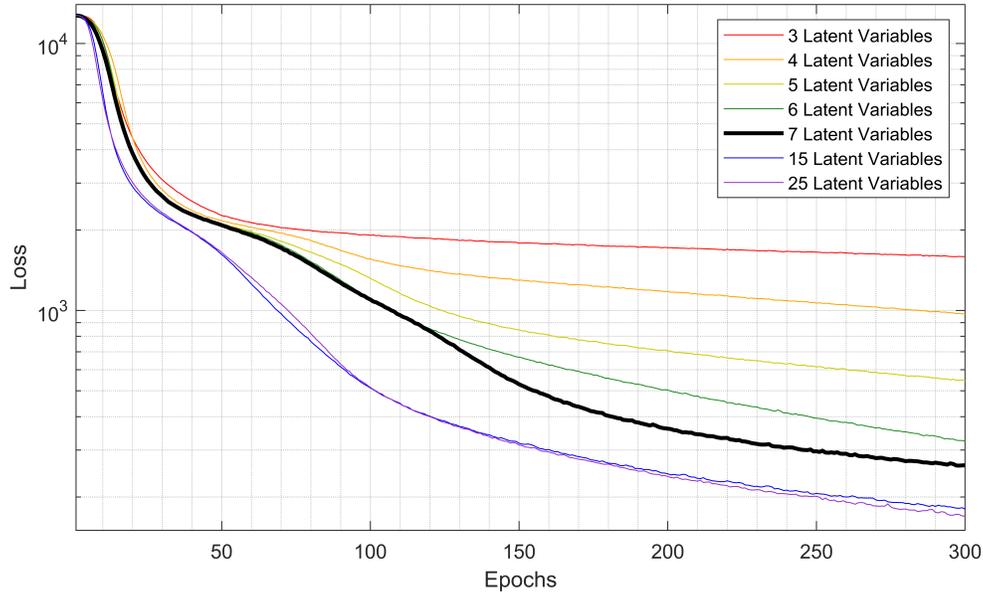


FIG. 11. Training loss versus epoch for the loss function in equation (11), for feature spaces defined by 3 latent variables (red), 4 latent variables (orange), 5 latent variables (yellow), 6 latent variables (green), 7 latent variables (black), 15 latent variables (blue), and 25 latent variables (purple).

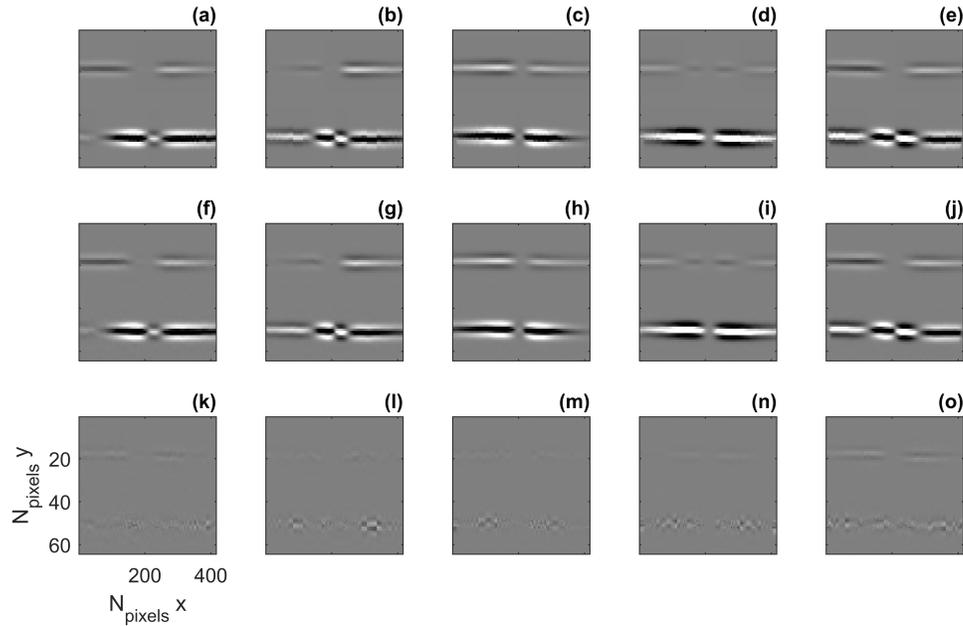


FIG. 12. (a)-(e) Five randomly selected input images from the processed dataset, (f)-(j) CAE predictions of these images for seven features in the feature space, (k)-(o) differences between original and predicted images.

in the data. The first method we will examine is clustering using the DBSCAN algorithm. Prior to clustering, it is common to perform a dimensionality reduction step. As the dimensionality of the feature space increases, the finite dimensional training dataset occupies a smaller subset of the feature space. When this occurs some of the ideas involved in cluster-

ing, such as Euclidean distance and nearest neighbor become less effective discriminators (Domingos, 2012).

Introduced by van der Maaten and Hinton (2012), a powerful method for this dimensionality reduction is t-distributed stochastic neighbor embedding (t-SNE). In the high dimensional space the similarity of two data points  $x_i$  and  $x_j$  is the joint probability  $p_{ij}$  that they would choose each other as neighbors, if neighbors were chosen in proportion to their probability density under a Gaussian centered on  $x_i$  where the probability is (van der Maaten and Hinton, 2012),

$$p_{ij} = \frac{\exp(-\|x_i - x_j\|^2/2\sigma^2)}{\sum_{k \neq l} \exp(-\|x_k - x_l\|^2/2\sigma^2)}. \quad (13)$$

A similar joint probability is computed in the low dimensional space between the counterpart points  $y_i$  and  $y_j$ . To reduce challenges associated with crowding, moderate distances in the high dimensional space are mapped to large distances in the low dimensional space by using a heavier-tailed distribution (van der Maaten and Hinton, 2012) such that moderately dissimilar points in high dimensional space become very dissimilar in low dimensions. A Student t-distribution with one degree of freedom is a common choice for the heavy-tailed distribution resulting in the low dimensional joint probability,

$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|y_k - y_l\|^2)^{-1}}. \quad (14)$$

T-SNE finds a low dimensional representation of the data that minimizes the mismatch between the two probability distributions in equations (13) and (14). A natural measure of this mismatch is the Kullback-Leibler divergence (Kullback and Leibler, 1951),

$$KL(P||Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \quad (15)$$

Figures 13a-13u plot cross-plots of one feature versus another for all seven features in the feature space, where each dot represents one of the 10,000 input images, colored by their assigned cluster from applying DBSCAN to the t-SNE representation of the data as described below. The cross-plotted features in Figure 13a-13u are hard to interpret due to their high-dimensionality and the crowding of points that occurs in the high-dimensional space. Applying t-SNE to this feature space representation, the result of which is plotted in Figure 13v, creates a much more cluster friendly representation of the data. Clearly, t-SNE has separated the natural clusters in the feature space (Figures 13a-13u) resulting in a more successful application of the DBSCAN clustering algorithm.

The t-SNE representation of the feature space is clustered using DBSCAN with a search radius of 5.2, and a minimum number of samples of 120. Using these parameters, DBSCAN detects the seven clusters plotted in Figure 13v. What we are interested in, is whether these

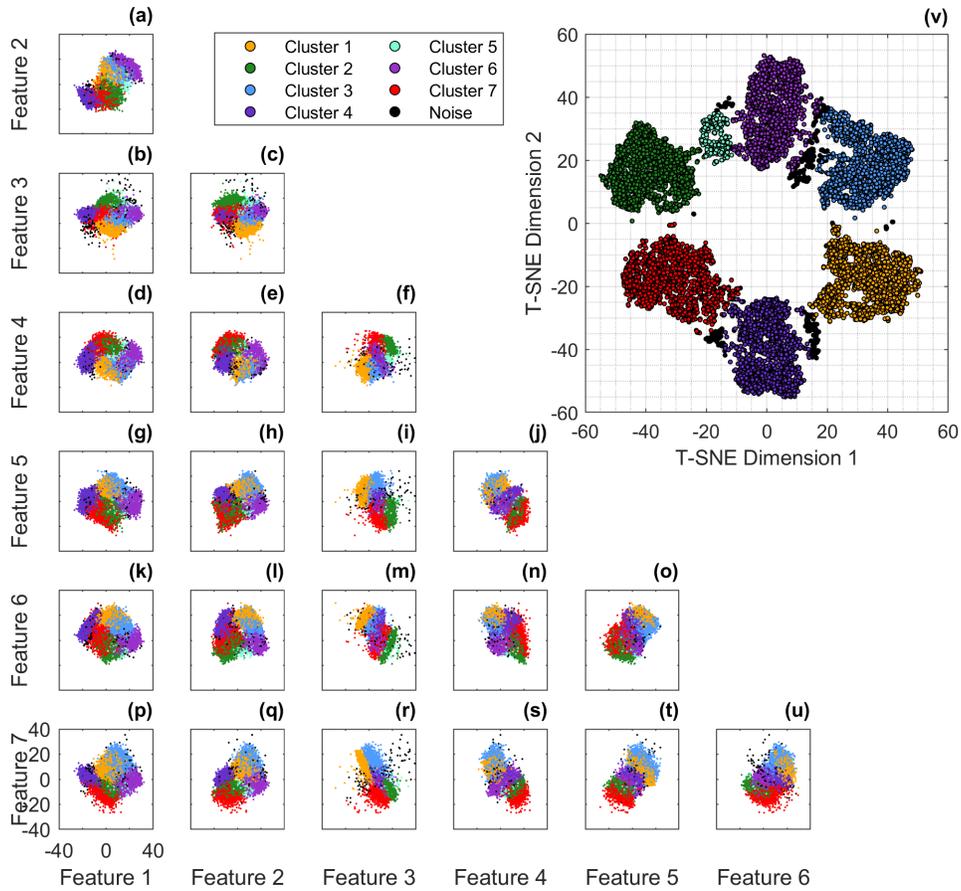


FIG. 13. (a)-(u) Cross-plots of the seven features in the feature space, where each subplot is a cross-plot of one feature versus another, and each dot is one of the 10,000 images. (v) The t-SNE representation of the cross-plotted features in (a)-(u).

clusters contain events that share similar source mechanisms, which would indicate the extracted features contain important source information. To examine this, the 10,000 input images are plotted by the Hudson space representation of their moment tensor source, and colored by the cluster to which they belong. The results of this are plotted in Figure 14. By-and-large, points that fall within the same cluster in Figure 13v, plot in a similar region of Hudson space. For example, points that fall within the red cluster tend to plot in the tensile crack closing (TCC) portion of Hudson space. This indicates that points that are similar in their feature space representation, which results in them being grouped in the same cluster, also share similar source mechanisms. This suggests that the convolutional autoencoder we have trained is extracting meaningful source mechanism information from the data, and that clustering groups images with similar source characteristics.

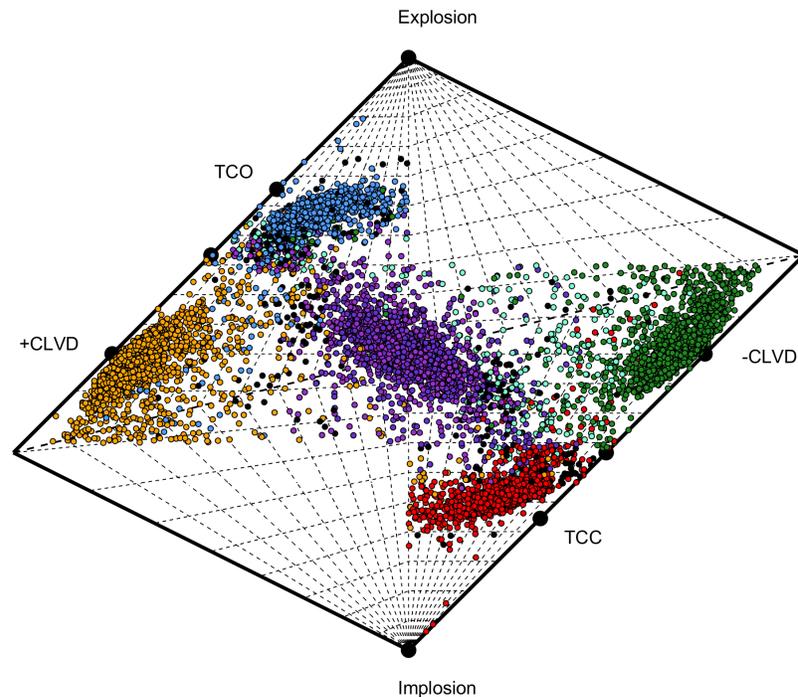


FIG. 14. Images plotted by the Hudson space representation of their moment tensor source, colored by the cluster to which they belong in Figure 13v.

Encouraged by the results obtained through clustering analysis, we now analyze the use of generative adversarial networks for predicting the source mechanism information encoded in the feature space. A GAN is trained for 250 epochs, using 80% of the data for training, and 20% for validation. For more robust results, Goodfellow et al. (2014) suggests training the discriminator for multiple iterations per epoch, and training the generator only once per epoch. The number of times the discriminator is trained is a network hyperparameter, again tuned through trial-and-error. Here we train the discriminator 16 times for each generator training. Two scenarios are tested, one where the goal of the generator is to estimate the two-component Hudson space label, and another where the generator tries to predict the six-independent components of the moment tensor using the feature space representations as inputs. The generator and discriminator network are standard densely connected neural networks, each with three layers. Table 2 summarizes the network architecture for the generator and table 3 summarizes the discriminator architecture.

Figure 15a plots the true Hudson space labels associated with the input data in purple, and the labels predicted by the trained generator in yellow. The distribution of the predicted labels matches the true dataset very well. Figure 15b plots the difference in the true and predicted labels with histograms showing the variance in the error in  $u$  and  $v$  components of the Hudson space labels. On the whole, the generator accurately predicts the Hudson space

Table 2. Generator network architecture

| Layer No. | Number of Neurons | Activation                 |
|-----------|-------------------|----------------------------|
| 1         | 64                | Leaky ReLU, $\alpha = 0.1$ |
| 2         | 128               | Leaky ReLU, $\alpha = 0.1$ |
| 3         | 64                | Leaky ReLU, $\alpha = 0.1$ |
| Output    | 2 or 6            | Linear                     |

Table 3. Discriminator network architecture

| Layer No. | Number of Neurons | Activation                | Dropout |
|-----------|-------------------|---------------------------|---------|
| 1         | 64                | Leaky ReLU $\alpha = 0.1$ | 0.35    |
| 2         | 128               | Leaky ReLU $\alpha = 0.1$ | 0.35    |
| 3         | 64                | Leaky ReLU $\alpha = 0.1$ | 0.35    |
| Output    | 1                 | Sigmoid                   | 0       |

label associated with the feature space representation. This lends further credence to the idea that the feature space representation contains important source information. If it did not then the association between feature space and Hudson label would be approximately random, and the generator would struggle to find a quality mapping between the two.

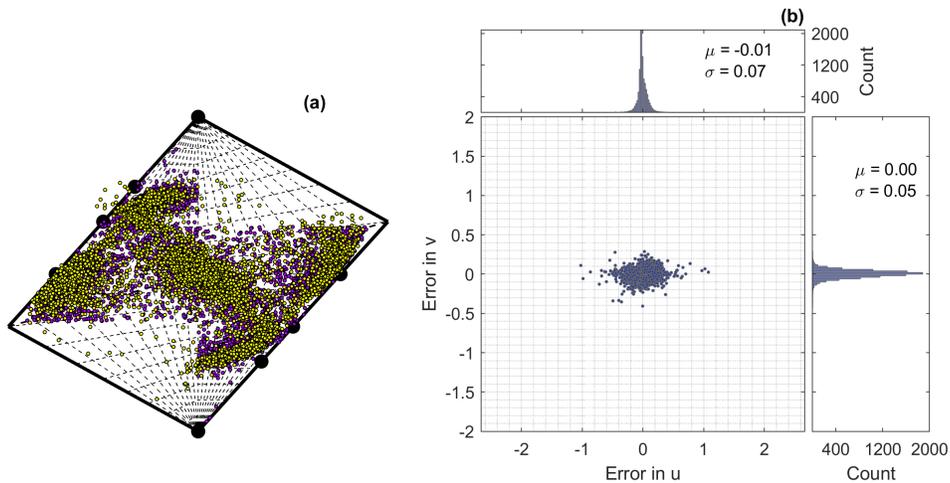


FIG. 15. (a) True labels (purple) and labels predicted by the trained GAN (yellow). (b) The difference between the true and predicted labels in (a) with histograms showing the variance in the error in the Hudson variables  $u$  and  $v$ .

A more challenging problem, that provides superior information about the source mechanism, is estimation of the full moment tensor from the feature space. Solving this problem follows a similar course to estimating the Hudson space label, but the generator is now

trained to predict a six-component vector, that is a good representation of the true moment tensors in the input data. Figure 16a-16f plots histograms for the error in the estimated moment tensor components for  $M_{xx}$ ,  $M_{xy}$ ,  $M_{xz}$ ,  $M_{yy}$ ,  $M_{yz}$ , and  $M_{zz}$ . While the errors are larger than those when predicting the Hudson space labels, they are still reasonable for all six-components. This suggests that the extracted features contain important information about the full moment tensor source that generated the data, and that the GAN can learn this relationship to predict the moment tensor.

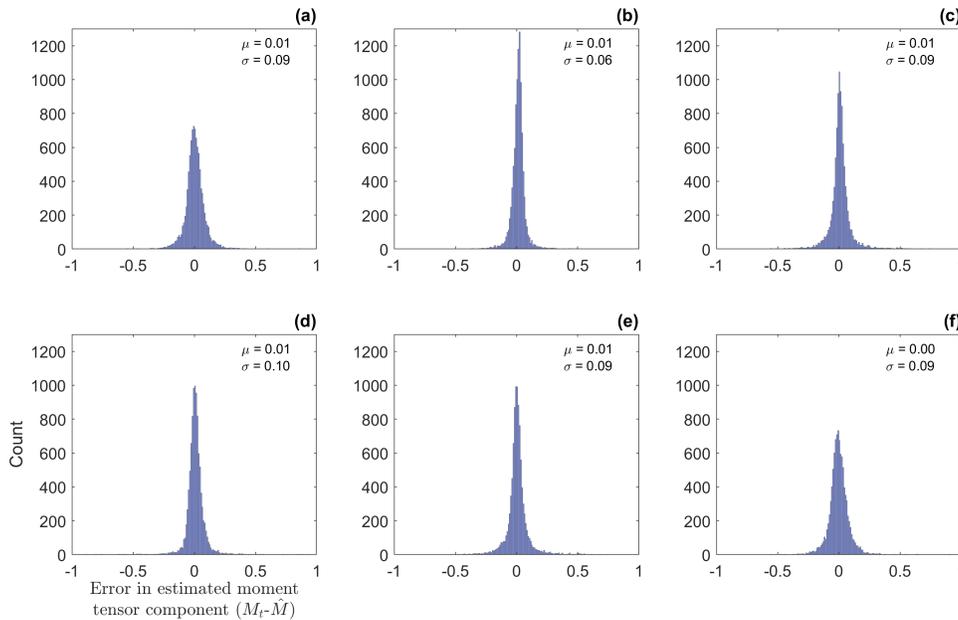


FIG. 16. Histograms for the error in the estimated moment tensor components using the GAN prediction for (a)  $M_{xx}$ , (b)  $M_{xy}$ , (c)  $M_{xz}$ , (d)  $M_{yy}$ , (e)  $M_{yz}$ , and (f)  $M_{zz}$ .

### EXTENSION TO FIELD DATA

The encouraging results of the synthetic examples suggest that the direct arrivals in microseismic data contain important source information, and that CAE architecture can be used to extract them. We are interested in extending this concept to examine whether or not we can extract source information from more complex field data. We follow the same processing workflow from Figure 10, to place the field data in a more CAE friendly domain. Figure 17a-17c plots the raw field data with NMO curve fit through semblance analysis, the extracted P-wave and S-wave, and the concatenated image for the CAE respectively. This concatenated image is then passed into the CAE that was trained on the synthetic data, and the reconstruction is shown in Figure 17d. While this reconstruction is not a perfect representation of the field data, it captures the significant features. For example the field data has attenuation of its P-wave on the toe side of the well (right side of Figure 17c) which is captured by the reconstruction, and the distinct polarity of the S-wave is also captured.

Motivated by this successful reconstruction, the features extracted by the encoder are supplied to the generative adversarial network, and a Hudson space label of  $(u, v) = (-0.1994, 0.3006)$  is predicted. The moment tensor,

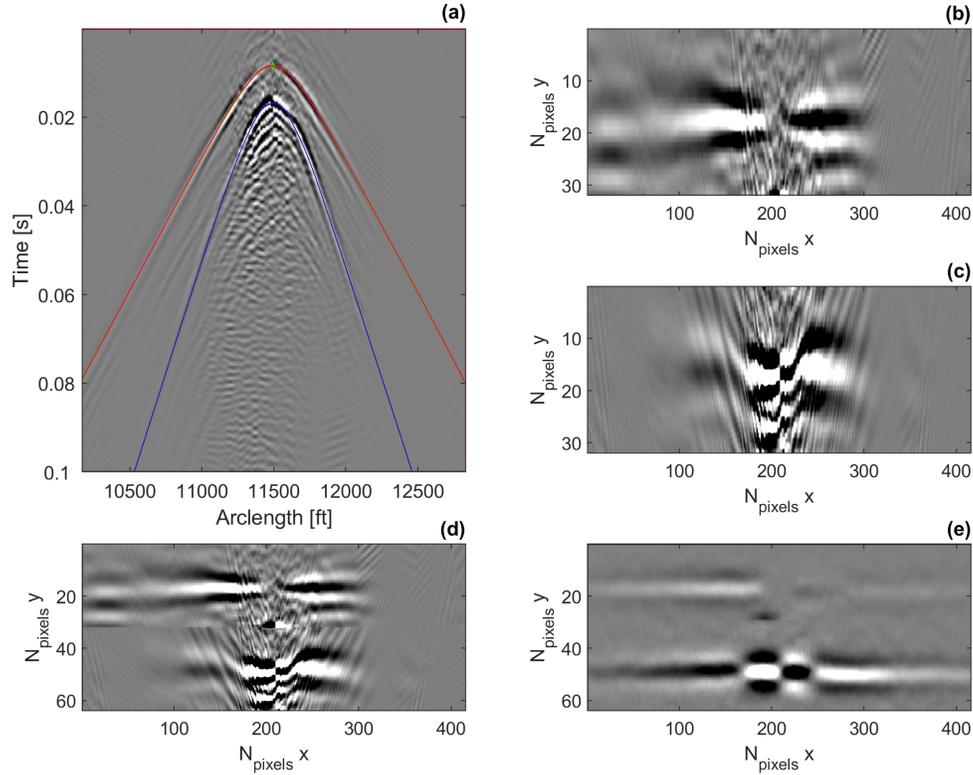


FIG. 17. (a) Field data, with picked apex and NMO for P-wave (red) and S-wave (blue), (b) extracted P-wave and (c) S-wave, (d) concatenated image for input to the CAE, and (e) the CAE reconstruction.

$$M = \begin{bmatrix} 0.69 & 1.00 & -0.69 \\ 1.00 & 0.35 & -0.22 \\ -0.69 & -0.22 & 0.69 \end{bmatrix} \quad (16)$$

producing a Hudson space label  $(u, v) = (-0.2012, 0.3019)$  is then used as input to equation (9) for modeling. If the Hudson space label predicted by the GAN is an accurate one for the input data, then the modeled data should correlate well to the input field data. Figure 18a plots the field data, and Figure 18b the synthetic modeled data with the Hudson space label  $(u, v) = (-0.2012, 0.3019)$ . The modeled data captures the important trends in the field data such as the S-wave polarity pattern, and attenuation of the toe side P-wave energy, suggesting that the predicted Hudson space labels is a good approximation of the true moment tensor generating the field data.

## DISCUSSION

We have developed deep learning techniques for estimating the source mechanism encoded within the direct arrivals of DAS-microseismic data. We are motivated by (1) the growing interest in DAS in reservoir monitoring during hydraulic fracturing, and in seismology studies, (2) the lack of studies focused on extracting earthquake source mechanisms

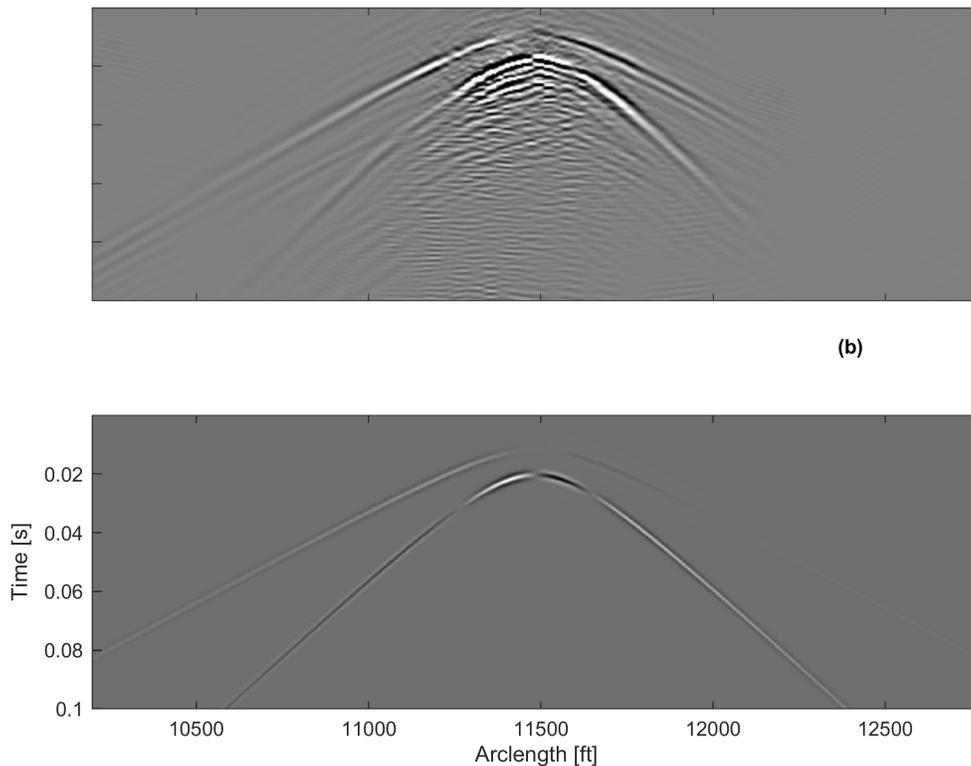


FIG. 18. (a) Field data, (b) data modeled with a moment tensor producing a similar Hudson space label to the one predicted by the generator.

from DAS data, an (3) the desire for efficient data-hungry algorithms that would benefit from the large datasets that DAS supplies. DAS is becoming an increasingly prevalent technology for microseismic monitoring, and has potential applications in seismology. Being a relatively new technology, robust methods do not currently exist for processing these large DAS-microseismic datasets for source information. The methods developed in this paper enable the estimation of source mechanism information from DAS data, increasing the information we can extract from DAS-microseismic data.

DAS technology offers an economic means of acquiring microseismic data during hydraulic fracturing, that allows for the dual purposing of wells for production and monitoring, and reduces the need for dedicated monitoring wells. The direct arrivals in the DAS-microseismic data are encoded with important information about the source mechanism that generated the data. This information is encoded within features of the direct arrivals such as the relative P-wave and S-wave amplitudes and polarity. In our method we leverage feature extracting deep learning techniques to detect the presence of these features, and make predictions about the class of source mechanism that generated the data.

Our results supply evidence that DAS-microseismic data do in fact contain features that provide clues to type of source mechanism that generated the data, and that these features can be extracted and processed to gain knowledge of the source mechanism. Part

of the motivation for this work is to extend the usefulness of DAS-microseismic data that is being acquired with increasing frequency. DAS data has the potential to transform entire unconventional fields into distributed strain sensors, and the methods for extracting source information developed here could expedite its application. Furthermore DAS has the potential to complement broadband seismometers used for seismology studies of large earthquakes, where methods for understanding the source mechanism are crucial.

Many of our conclusions and results derive from numerical simulations in relatively simple media, however, the results presented in this paper appear to transfer to the processing of direct arrivals in field data for source information. For the work presented here to be useful, they must be robust on large field datasets. Future work should focus on the extension of these methods to a full field dataset. Ideally, a dataset with conventional MTI should be compared to the methods developed in this communication, and neural network training should be performed on field data. The advantage with the method presented here, is that once a mapping is learned, it should generalize well to other datasets, especially similar data from perhaps a different well in the same field.

## **CONCLUSIONS**

Distributed acoustic sensing (DAS) is a growing technology for seismic reservoir monitoring, especially in unconventional fields requiring hydraulic fracture treatments. Offering access to unique acquisition geometries, DAS in principle supplies datasets that should contain features necessary for moment tensor estimation. DAS supplies large datasets of seismic strain measurements, preventing them from being directly included in conventional moment tensor inversion. To facilitate the estimation of moment tensors from DAS data we turn our attention to algorithms rooted in deep learning. Using a convolutional autoencoder we compress DAS microseismic images to a vector of only their most salient features, with the expectation that these features contain important information about the moment tensor. These features are then analyzed for the moment tensor information they contain using clustering techniques and through training of a GAN for moment tensor prediction. Feature vectors are shown to form localized clusters, where each cluster contains data with strongly correlated source mechanisms. The trained GAN proves to be accurate in its ability to predict the moment tensor associated with a given feature vector. The GAN trained on synthetic data is then used to predict the moment tensor from an example field data event. Data modeled with the predicted moment tensor correlates strongly with the field data. These results suggest that DAS-microseismic data contain features indicative of the source mechanism, and that deep learning techniques can be used to extract these features, and process them to form predictions of the source mechanism.

## **ACKNOWLEDGMENTS**

The authors would like to thank the sponsors of the CREWES project as well NSERC under the grants CRDPJ 461179-13 and CRDPJ 543578-19 for making this work possible through their financial support. Matt Eaid and Scott Keating were partially supported through scholarships from the SEG foundation.

## REFERENCES

- Addagalla, A., Maley, I., Lawal, I., Jadhav, P., and Luigi, M., 2018, Minimum stress, maximum pressure: A new high performance bridging system facilitates drilling depleted formations at high overbalance in Middle East: Presented at the SPE Asia Pacific Oil and Gas Conference and Exhibition, Brisbane, Australia.
- Aki, K., and Richards, P., 2002, *Quantitative Seismology*: University Science Books, 2 edn.
- Albright, J., and Pearson, C., 1982, Acoustic emissions as a tool for hydraulic fracture location: Experience at the Fenton Hill Hot Dry Rock site: *Society for Petroleum Engineers*, **22**, No. 4, 523–530.
- Bin, G., Gao, J., Li, X., and Dhillon, B., 2012, Early fault diagnosis of rotating machinery based on wavelet packets—Empirical mode decomposition feature extraction and neural network: *Mechanical Systems and Signal Processing*, **27**, 696 – 711.
- Bjorck, N., Gomes, C. P., Selman, B., and Weinberger, K. Q., 2018, Understanding batch normalization, *in* Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R., Eds., *Advances in Neural Information Processing Systems 31*, Curran Associates, Inc., 7694–7705.
- Burridge, R., and Knopoff, L., 1964, Body force equivalents for seismic dislocations: *Bulletin of the Seismological Society of America*, **54**, No. 6A, 1875–1888.
- Chen, Y., Jiang, H., Li, C., Jia, X., and Ghamisi, P., 2016, Deep feature extraction and classification of hyperspectral images based on convolutional neural networks: *IEEE Transactions on Geoscience and Remote Sensing*, **54**, No. 10, 6232–6251.
- Daley, T., Freifeld, B., Ajo-Franklin, J., Dou, S., Pevzner, R., Shulakova, V., Kashiakr, S., Miller, D., Goetz, J., Hennings, J., and Lueth, S., 2013, Field testing of fiber-optic distributed acoustic sensing (DAS) for subsurface seismic monitoring: *The Leading Edge*, **32**, No. 6, 593–724.
- Domingos, P., 2012, A few useful things to know about machine learning: *Communication of the Association for Computing Machinery*, **55**, No. 10, 78–87.
- Eaton, D., 2018, *Passive Seismic Monitoring of Induced Seismicity: Fundamental Principles and Application to Energy Technologies*: Cambridge University Press, 1 edn.
- Eaton, D. W., van der Baan, M., Birkelo, B., and J.B., T., 2014, Scaling relations and spectral characteristics of tensile microseisms: Evidence for opening/closing cracks during hydraulic fracturing: *Geophysics Journal International*, **196**, No. 3, 1844–1857.
- Ester, M., Kriegel, H.-P., Sander, J., and Xu, X., 1996, A density-based algorithm for discovering clusters in large spatial databases with noise, *in* *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, KDD'96*, AAAI Press, 226–231.
- Eyre, T. S., and van der Baan, M., 2015, Overview of moment-tensor inversion of microseismic events: *The Leading Edge*, **34**, No. 8, 882–888.
- Frenet, F., 1852, Sur les courbes à double courbure.: *Journal de Mathématiques Pures et Appliquées*, 437–447.
- Gharti, H. N., Oye, V., Roth, M., and Kijohn, D., 2010, Automated microearthquake location using envelope stacking and robust global optimization: *GEOPHYSICS*, **75**, No. 4, MA27–MA46.
- Ghasedi Dizaji, K., Herandi, A., Deng, C., Cai, W., and Huang, H., 2017, Deep clustering via joint convolutional autoencoder embedding and relative entropy minimization, *in* *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.
- Gilbert, F., 1971, Excitation of normal modes of the Earth by earthquake sources: *Geophysics Journal International*, **22**, No. 2, 223–226.

- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y., 2014, Generative adversarial nets, *in* Advances in Neural Information Processing Systems 27, Curran Associates, Inc., 2672–2680.
- Grandi, S., and Oates, S., 2009, Microseismic event location by cross-correlation migration of surface array data for permanent reservoir monitoring: 71st EAGE Conference and Exhibition, Amsterdam, The Netherlands, Expanded Abstracts.
- Grigoli, F., Cesca, S., Amoroso, O., Emolo, A., Zollo, A., and Dahm, T., 2013, Automated seismic event location by waveform coherence analysis: *Geophysical Journal International*, **196**, No. 3, 1742–1753.
- Hardebeck, J. L., and Shearer, P. M., 2002, A New Method for Determining First-Motion Focal Mechanisms: *Bulletin of the Seismological Society of America*, **92**, No. 6, 2264–2276.
- Hasegawa, A., Zhang, W., Itoh, K., and Ichioka, Y., 1991, Image processing based on supervised learning in neural networks, *in* Mikaelian, A. L., Ed., *Optical Memory and Neural Networks*, vol. 1621, International Society for Optics and Photonics, SPIE, 374 – 379.
- Hertel, L., Barth, E., Käster, T., and Martinetz, T., 2015, Deep convolutional neural networks as generic feature extractors, *in* 2015 International Joint Conference on Neural Networks (IJCNN), 1–4.
- Hudson, J. A., Pearce, R. G., and Rogers, R. M., 1989, Source type plot for inversion of the moment tensor: *Journal of Geophysical Research: Solid Earth*, **94**, No. B1, 765–774.
- Innanen, K., 2017, Determination of seismic tensor strain from HWC-DAS cable with arbitrary and nested-helix winds: SEG Expanded Abstracts.
- Jain, A. K., Murty, M. N., and Flynn, P. J., 1999, Data clustering: A review: *Association for Computing Machinery Computing Surveys*, **31**, No. 3, 264–323.
- Karimpouli, S., and Tahmasebi, P., 2019, Segmentation of digital rock images using deep convolutional autoencoder networks: *Computers and Geosciences*, **126**, 142 – 150.
- Kingma, D. P., and Ba, J., 2017, Adam: A method for stochastic optimization, 1412.6980.
- Kullback, S., and Leibler, R. A., 1951, On information and sufficiency: *Annals of Mathematical Statistics*, **22**, No. 1, 79–86.
- Kuvshinov, B. N., 2015, Interaction of helically wound fibre-optic cables with plane seismic waves: *Geophysical Prospecting*, **64**, No. 3.
- Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P., 1998, Gradient-based learning applied to document recognition: *Proceedings of the IEEE*, **86**, No. 11, 2278–2324.
- Li, L., Chen, H., and Wang, X., 2016, Relative elastic interferometric imaging for microseismic source location: *Journal of Geophysics and Engineering*, **13**, No. 5, 733–744.
- Lindsey, N. J., Martin, E. R., Dreger, D. S., Freifeld, B., Cole, S., James, S. R., Biondi, B. L., and Ajo-Franklin, J. B., 2017, Fiber-optic network observations of earthquake wavefields: *Geophysical Research Letters*, **44**, No. 23, 11,792–11,799.
- Lo, S. B., Lou, S. A., Jyh-Shyan Lin, Freedman, M. T., Chien, M. V., and Mun, S. K., 1995, Artificial convolution neural network techniques and applications for lung nodule detection: *IEEE Transactions on Medical Imaging*, **14**, No. 4, 711–718.
- Lopes, U., and Valiati, J., 2017, Pre-trained convolutional neural networks as feature extractors for tuberculosis detection: *Computers in Biology and Medicine*, **89**, 135 – 143.
- Mao, Q., Azeem, T., and Zhang, X., 2020, A migration-based location method using improved waveform stacking for microseismic events in a borehole system: *Acta Geophysica*.

- Masoudi, A., Belal, M., and Newson, T., 2013, A distributed optical fibre dynamic strain sensor based on phase-OTDR: *Measurement Science and Technology*, **24**.
- Maxwell, S., 2010, Microseismic: Growth born from success: *The Leading Edge*, **29**, No. 3, 338–343.
- Maxwell, S., Urbancic, T., Steinsberger, N., and Zinno, R., 2002, Microseismic imaging of hydraulic fracture complexity in the Barnett shale: Presented at SPE Annual Technical Conference and Exhibition.
- McClain, W., 1971, Seismic mapping of hydraulic fractures: Technical Report ORNL-TM-3502, Oak Ridge National Laboratory.
- Nagy, G., 1968, State of the art in pattern recognition: *Proceedings of the IEEE*, **56**, No. 5, 836–863.
- Power, D., Schuster, C., Hay, R., and Twombly, J., 1976, Detection of hydraulic fracture orientation and dimensions in cased wells: *Journal of Petroleum Technology*, **29**, No. 9, 1–116.
- Qian, F., Yin, M., Liu, X.-Y., Wang, Y.-J., Lu, C., and Hu, G.-M., 2018, Unsupervised seismic facies analysis via deep convolutional autoencoders: *Geophysics*, **83**, No. 3, A39–A43.
- Schuster, G., Yu, J., and Sheng, J., 2004, Interferometric/daylight seismic imaging: *Geophysics Journal International*, **157**, 838–852.
- Serret, J.-A., 1851, Sur quelques formules relatives á la théorie des courbes à double courbure.: *Journal de Mathématiques Pures et Appliquées*, 193–207.
- Song, H., Gao, Y., Chen, W., Juan Xue, Y., Zhang, H., and Zhang, X., 2020, Seismic random noise suppression using deep convolutional autoencoder neural network: *Journal of Applied Geophysics*, **178**, 104,071.
- Spanhol, F. A., Oliveira, L. S., Petitjean, C., and Heutte, L., 2016, Breast cancer histopathological image classification using convolutional neural networks, *in* 2016 International Joint Conference on Neural Networks (IJCNN), 2560–2567.
- Sun, J., Niu, Z., Innanen, K. A., Li, J., and Trad, D. O., 2020, A theory-guided deep-learning formulation and optimization of seismic waveform inversion: *Geophysics*, **85**, No. 2, R87–R99.
- van der Maaten, L., and Hinton, G., 2012, Visualizing data using t-SNE: *Journal of Machine Learning Research*, **9**, 2579–2605.
- Wang, Y., Wang, B., Tu, N., and Geng, J., 2020, Seismic trace interpolation for irregularly spatial sampled data using convolutional autoencoder: *Geophysics*, **85**, No. 2, V119–V130.
- Warpinski, N., Branagan, P., Peterson, R., and Wolhart, S., 1998, An interpretation of m-site hydraulic fracture diagnostic results: Presented at SPE Rocky Mountain Regional/Low-Permeability Reservoirs Symposium.
- Willacy, C., van Dedem, E., Minisini, S., Li, J., Blokland, J.-W., Das, I., and Droujinine, A., 2019, Full-waveform event location and moment tensor inversion for induced seismicity: *GEOPHYSICS*, **84**, No. 2, KS39–KS57.
- Yu, C., Zhan, Z., Lindsey, N. J., Ajo-Franklin, J. B., and Robertson, M., 2019, The potential of das in teleseismic studies: Insights from the goldstone experiment: *Geophysical Research Letters*, **46**, No. 3, 1320–1328.
- Zhang, W., and Zhang, J., 2013, Microseismic migration by semblance weighted stacking and interferometry: 83rd SEG Conference and Exhibition Expanded Abstracts.