Pre-migration diffraction separation using generative adversarial networks

Brydon Lowney1,2*, Ivan Lokmer1,2, Gareth S. O’Brien1,3 and Christopher J. Bean4

1School of Earth Science, University College Dublin, Belfield, Dublin D04 V1W8, Republic of Ireland, 2Irish Centre for Research in Applied Geoscience, University College Dublin, Belfield, Dublin D04 V1W8, Republic of Ireland, 3Applied Geophysics and Technology, Tullow Oil Ltd., Dublin, D18 NH10, Ireland, and 4School of Cosmic Physics, Dublin Institute of Advanced Studies, Dublin, D02 Y006, Republic of Ireland

Received September 2020, revision accepted March 2021

ABSTRACT

Diffraction imaging is the process of separating diffraction events from the seismic wavefield and imaging them independently, highlighting subsurface discontinuities. While there are many analytic-based methods for diffraction imaging which use kinematic, dynamic or both, properties of the diffracted wavefield, they can be slow and require parameterization. Here, we propose an image-to-image generative adversarial network to automatically separate diffraction events on pre-migrated seismic data in a fraction of the time of conventional methods. To train the generative adversarial network, plane-wave destruction was applied to a range of synthetic and real images from field data to create training data. These training data were screened and any areas where the plane-wave destruction did not perform well, such as synclines and areas of complex dip, were removed to prevent bias in the neural network. A total of 14,132 screened images were used to train the final generative adversarial network. The trained network has been applied across several geologically distinct field datasets, including a 3D example. Here, generative adversarial network separation is shown to be comparable to a benchmark separation created with plane-wave destruction, and up to 12 times faster. This demonstrates the clear potential in generative adversarial networks for fast and accurate diffraction separation.

Key words: Data processing, Imaging, Seismics.

INTRODUCTION

The seismic wavefield is formed by interactions of a source wavefield with the subsurface as it travels through it. These interactions can take many forms depending on the physical parameters of the subsurface with the main events being reflections, refractions and diffractions (Kennett, 2001). On top of this, a variety of coherent and incoherent noise is always present in the seismic wavefield which act to further complicate it (Bonnefoy-Claudet et al., 2006). The goal of the geophysicist is to take this raw data and process it to an interpretable form which can be passed on to an interpreter whereby it can be used to locate raw materials (Claerbout, 1985). Conventionally, this processing favours reflections, which form at continuous subsurface boundaries (Moser and Howard, 2008). While this reflection processing is convenient for continuous horizons, there are limitations in reflection images. In a reflection image, faults are inferred through a break in the reflection event, as opposed to directly imaged. On top of this, the resolution of reflection images is limited by the Rayleigh criterion (Khaidukov et al., 2004).

Diffractions are formed when the source wavefield encounters objects or discontinuities in the subsurface which are small compared to the wavelength (Landa, 2007). Where the wavefield encounters these discontinuities, it spreads as a diffraction. The term diffraction here is used to describe both edge diffractions (which occur at the tips of reflections, e.g., at
faults) and point diffractions (which occur because of objects) (Schwarz, 2019a). As diffractions are formed by these discontinuities, by imaging the diffractions independently these discontinuities can be directly imaged, creating a diffraction image (Berkovitch et al., 2009). Pioneering works by Krey (1952) showed the first clear field records containing seismic diffractions and stressed the importance of diffractions in imaging faults. Torey (1970) noted further the subsurface response of diffractions and suggested that when imaging discontinuities one must consider diffractions over reflections. An early attempt at diffraction separation was then performed by Harlan et al. (1984) who used lateral coherency to separate diffraction and reflection energy. This was followed by a method to both image and boost diffractions above reflections to image discontinuities using Normal Moveout (NMO) correction by Kanasewic and Phadke (1988). These early diffraction images demonstrated the wealth of information contained by diffractions which is imperceptible on reflection images. As diffractions are not limited by the Rayleigh criterion, they have the potential to enhance the seismic resolution beyond reflection images (Gelius and Asgedom, 2011). Additionally, as diffractions are treated akin to noise on a conventional seismic image, by independently imaging the diffractions, these can also be removed from reflection images to enhance this image, a technique known as specular imaging (JafarGandomi et al., 2019). Diffractions can also limit imaging of deep structures and thus by imaging and removing these scatterers we can potentially enhance deep structural imaging (Martini et al., 2001).

To create a diffraction image, the diffractions must first be separated from the seismic wavefield. This is a challenging task due to the blending of diffraction and reflection energy and the complexity of the wavefield itself (Schwarz, 2019b). Additionally, diffractions are extremely weak, approximately an order of magnitude lower amplitude than reflections, and can therefore be masked by the stronger reflection energy (Klokov et al., 2010). There exist several analytic-based methods which aim to separate the diffracted portion of the wavefield based on the dynamic, kinematic or both, properties of the diffracted wavefield, and these can broadly be divided into pre-migration techniques and migration gather separation techniques (Popovici et al., 2015).

Pre-migration diffraction imaging techniques involve the separation of diffractions before migration where they can then be processed independently and migrated. This is useful as it allows for separate migration schemes to be applied to diffractions, which are more sensitive to velocity errors than the corresponding reflections (Landa et al., 2006). There are many pre-migration techniques including multifocussing (Berkovitch et al., 2009), anti-stationary phase filtering (Moser and Howard, 2008), common-reflection surface (Dell and Gajewski, 2011) and coherent subtraction (Schwarz, 2019b). Plane-wave destruction (Claerbout, 1985; Fomel et al., 2007) is one of the most common pre-migration diffraction imaging techniques which has been used in this paper to create comparison datasets and will therefore be described in detail later. Diffraction imaging using migration gathers, on the other hand, involves the separation of diffractions after migration has occurred. As migration collapses diffractions to a point on the seismic image, a separate domain is used for separation generated during the migration process (Reshef and Landa, 2009). The dip-angle gather domain is commonly used for this form of separation (Landa et al., 2008), and many methods exist within this domain for separation, including radon filtering (Klokov and Fomel, 2012) and apex destruction (Kozlov et al., 2004). Additionally, there exist machine learning methods for separation including pioneering work from Serfaty et al. (2018) who were the first to demonstrate diffraction separation using deep learning and used the directional image gather domain, separation in the dip-angle gather domain (Tschannen et al., 2020) and diffraction point detection (Maciel and Biloti, 2014).

In this paper, we focus on pre-migration separation using an image-to-image (Pix2Pix) generative adversarial network (GAN) to automatically detect and separate diffractions on pre-migration, common-offset seismic data. This image-conditional GAN has been chosen as it is more suitable for image-to-image translation tasks, such as a seismic separation, than a conventional neural network, especially when data are scarce (Goodfellow, 2016). First, we describe plane-wave destruction (PWD), a common method for diffraction imaging which has been used here as a state-of-the-art comparison to demonstrate the performance of GAN separation as well as used to train the GAN itself. We then show quantitative and qualitative analyses of the GAN on synthetic data. Finally, we demonstrate the ability of GAN to separate diffractions on multiple field datasets, including an open-source dataset for reproducibility and a migrated 3D example.

**METHOD**

On common-offset seismic data, diffractions appear hyperbolic and have low amplitudes, while reflections follow dominant local slopes and have higher amplitudes (Decker et al., 2013). This forms the basis for most pre-migration separation techniques and our generative adversarial network
(GAN) based technique is no different. By training the neural network to recognize hyperbolic, low-amplitude shapes as desirable and anything which does not fit these criteria as undesirable, pre-migrated diffractions can effectively be separated from the wavefield (Lowney et al., 2020).

In this section, we briefly describe plane-wave destruction (PWD) and following this, we describe the neural network architecture of the GAN, before discussing the training regime for the network and quantifying its performance on synthetic data. Subsequently, we analyse the performance of the GAN with PWD on simple synthetic data. Finally, we discuss how this is adapted to field data.

**Plane-wave destruction**

PWD is a common diffraction imaging technique which calculates the slope of reflection energy and removes anything which conforms to this slope:

$$\frac{\partial P}{\partial x} + \sigma \frac{\partial P}{\partial t} = 0,$$

where \( P(x,t) \) is the wavefield and \( \sigma \) is the local slope (Chen et al., 2013). PWD is a common diffraction imaging technique and has seen successes in many areas (Decker et al., 2013). However, there are some pitfalls associated with PWD. PWD requires a continuously variable slope and therefore does not perform well in areas this assumption does not hold true, for example synclines (Decker et al., 2015). In addition, PWD requires a dip field (calculated as the dip of the reflectors) as well as careful parameterization of the algorithm (such as the accuracy order and smoothing parameters) to avoid the introduction of errors (Fomel, 2002). While the dip field can be used for quality control, any errors introduced in the dip field will be carried forward into the final image. Despite these downsides, PWD remains a simple and powerful method for diffraction extraction and as such, we use it here as a benchmark comparison for the GAN separation. For the PWD separations in this paper, we have used Madagascar software (Fomel et al., 2013).

**Network architecture of GAN**

GANs, first outlined in Goodfellow et al. (2014), are a form of adversarial or competitive network, in which there are two neural networks that compete against each other during training (Fig. 1). The first of these networks, the generator, is responsible for creating data based on the input training data (Radford et al., 2016). While the training data can consist of either classes or images, in our case, the training data consist of images of both unprocessed seismic data as the input, and diffraction-only data as the target. The generator attempts to transform the raw data into the diffraction-only data, akin to how a conventional neural network would learn. However, instead of learning to create a single specific image, the generator is learning to recreate diffraction data realistic enough to fool the second neural network, the discriminator. The discriminator network also uses the training data to understand what constitutes diffraction-only data. With this understanding, the discriminator network then provides both real (from the training data) and generated (created by the generator network) samples and outputs the probability that a given sample is real (Nguyen et al., 2017). The discriminator network is trained independently for a while before the generator is trained so that it can label the initial generated images. Both networks are then trained simultaneously, with the generator trying to create data which can fool the discriminator and the discriminator attempting to identify fake data (Goodfellow et al., 2014). As GANs calculate the joint probability distribution, compared to a convolutional neural network (CNN) which maps an input to a class label, GAN may be more suited for image-to-image translation tasks such as seismic event separation (Alotaibi, 2020). Additionally, GANs can perform well with a lack of data and are thus more suitable when data availability may be an issue (Han et al., 2019). However, the adversarial nature also makes this type of network more difficult to train as, unlike a conventional neural network, the loss cannot be used reliably as a measure of the training, an issue which will be further discussed later (Kahng et al., 2018).

There are two main types of GAN: conditional and unconditional. An unconditional GAN transforms a random noise vector into an image with the same data distribution as the training data. This generated image is then fed into the discriminator network alongside the target data (Tran et al., 2019). The discriminator network then outputs a probability that the given image is real or fake (Tiwari et al., 2020). A conditional GAN, such as the Pix2Pix GAN used here, provides both generator and discriminator with the input data in the form of a tuple \([G(x), y] \) and \([y, x] \), where \( x \) is the input data (raw seismic data), \( y \) is the target data (the PWD diffraction separated data) and \( G(x) \) is the generated data (the GAN diffraction separated data) (Isola et al., 2017). By providing the target data as an additional input, a conditional GAN directs the data generation process as it is constrained by the target data. This additional constraint makes conditional GAN more suitable for image translation than an unconditional GAN (Zhang et al., 2019). While the input data for a conditional GAN can be either classes or images, in this paper,
Figure 1 Flowchart showing network architecture of the Pix2Pix GAN (a). Here, the input data (raw seismic data) are input into the generator which creates fake diffraction data. This generator has a U-Net architecture which is displayed larger in (b). This input is then fed through to the discriminator alongside either a real (target) or generated image as a tuple. The discriminator outputs a probability that the diffraction data is either real or fake. The discriminator network is a convolutional network shown in (c).

the focus is on conditional image GANs, specifically the Pix2Pix GAN.

For our GAN, we have used the network architecture outlined in the original Pix2Pix GAN paper (Isola et al., 2017). This network has already seen some successes in geophysics with examples showing ground roll attenuation (Yuan et al., 2020), seismic interpolation (Oliveira et al., 2018) and bandwidth extension (Aharchaou and Baumstein, 2020). The
Pix2Pix GAN uses a U-Net architecture for the generator and a PatchGAN architecture for the discriminator (Isola et al., 2017). The U-Net, first proposed by Ronneberger et al. (2015) and so-called for its U-shaped architecture, is an artificial neural network which is designed to be fully convolutional and is commonly used in image segmentation tasks (Li et al., 2018). Most neural networks are designed to be contracting, with the number of nodes in each successive layer decreasing, allowing for reduction of the spatial information while increasing the feature information, and essentially reducing an image into key features (Basheer and Hajmeer, 2000). After the conventional contracting seen in other network architectures, the U-Net then has an expansive path which combines both feature and spatial information which gives a more precise output than networks with just a contracting layer (Lian et al., 2018). Additionally, the U-Net has skip connections, allowing the network to capture high-resolution detail and prevent the vanishing gradient problem (Sun et al., 2020). This architecture allows the U-Net to map features of the image which can then be used to recreate the image (Sankesara, 2019). In this way, the U-Net can learn features of an image or dataset and apply these same features to new images. The generator is trained using an adversarial loss (the variable loss function from the discriminator network) as well as updated using an $L_1$ loss function (least absolute deviations):  

$$L_1 = \sum_{i=1}^{n} |y_i - \hat{y}_i|.$$  

(2)

where $n$ is the number of values, $y$ is the true value and $\hat{y}$ is the predicted value (Shekhar, 2019). $L_1$ loss functions, as defined above, have been used due to its relative insensitivity to outliers and as $L_2$ loss functions result in blurry images (Cai et al., 2019). The $L_1$ loss is measured between the expected output (i.e. the target image) and the generated output, while the adversarial loss aims to minimize the loss from the discriminator network for images marked as ‘real’ (Brownlee, 2019). This combination of loss functions (applied as a weighted sum with a 100 to 1 weighting towards the $L_1$ loss) allows the network to create more realistic translations of the target images as opposed to the target domain in general. The specific U-Net architecture used here is displayed in Figure 1(b).

The PatchGAN discriminator is also a CNN with several contracting layers, however, unlike the U-Net used for the generator, there is no expansive path (see Fig. 1c). The PatchGAN used here utilizes binary cross-entropy (BCE) as a loss function:  

$$BCE = -\frac{1}{N} \sum_{i=1}^{N} y_i \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i)).$$  

(3)
where $y$ is the label (i.e., whether the data is real or fake) and $p(y)$ is the probability of the point being real for all $N$ points (with $N$ being the total number of points) (Ho and Wookey, 2019). The discriminator attempts to minimize this loss and outputs a probability that a given sample is real or fake. The ‘patch’ in PatchGAN refers to the fact that the discriminator is run in $N \times N$ patches trying to determine if each patch is fake as opposed to the image as a whole (Xu et al., 2017). This is then run convolutionally along the image and the responses are averaged to provide the final output, making it faster than conventional discriminators and with fewer parameters (Brownlee, 2019). Here, we have used $70 \times 70$ patches of the input image to allow for large enough patches that the features are still visible. Here, smaller patches and larger patches were both tested ($35 \times 35$ and $140 \times 140$, respectively) but the $70 \times 70$ patches showed the best trade-off. This $70 \times 70$ patch size is same as the findings of the original Pix2Pix GAN paper (Isola et al., 2017).

**Training and quantification**

As mentioned previously, training a GAN is a complex task due to the adversarial nature of the two networks. In a conventional neural network, the training loss (the loss function on the training data) and validation loss (the loss function on the validation data) are expected to decrease with increasing training until a point where the validation loss begins to increase at which point the model is trained (Knerr et al., 1990).
This does not hold true for GANs as when one network exhibits a low loss this conversely means a high loss on the other network, making judgement of the neural network through the loss function difficult due to the complexity of the loss (O’Brien, 2020) (Fig. 2). As such, other metrics must be used to quantify the performance of the GAN. Here, we have used a simple synthetic toy model for quantification (the data from which was not used in training the GAN). This toy model was created using a finite-difference modelling scheme and mimics field data with several diffractions and reflections. The model was created as a zero-offset model with a 50 Hz Ricker wavelet and 1 m trace spacing. As this is synthetic data, this allowed for the data to be generated twice, once with reflections, diffractions and noise, which is used as an input dataset, and once with diffractions only, which is used as a training dataset.

To calculate these twice, two velocity models have been used. One which contains reflections, and the other which contains diffraction points in place of the diffractions (from the reflection model). In this way, we aim to not only separate diffractions but also reduce background noise and diffracted multiples.

The preliminary training data consisted solely of synthetic data generated using the same method outlined above (i.e. once with reflections, diffractions and noise as an input dataset, and once with diffractions only as the training dataset) but with different models which have different reflectors and diffractors. Overall, 10 synthetic datasets were generated for the original training data for the synthetic tests which were divided into 256 x 256 pixel, greyscale images for a total of 200 images with each pixel representing a seismic time-sample on a trace (i.e. a 1:1 scale conversion between seismic data and image). The network was trained over 100 epochs with each epoch consisting of individual ‘steps’ where a single image is passed through the network. An epoch then refers to when all images have been passed through the network.

The trained network (at each step) can be used to generate GAN-separated diffraction data, which can be compared to the original diffraction synthetic data, to calculate the change in diffraction energy, and with the raw data, to compare the remnant reflection energy and any artefacts created (Figs 3 and 4, Table 1). Calculating the diffraction energy is easily achieved on synthetic data as the true answer is known, therefore the reflections and noise can be perfectly removed, and the remaining energy summed up to give the diffraction energy (while the removed energy can be summed up to give the noise energy). From Table 1 it is evident that the relationship between the training steps and GAN performance is complex. While there is an approximation that with increased training
Table 1 Quantitative analysis of diffraction and noise energy (here classified as anything non-diffraction) for the raw data, PWD data and the GAN

|         | Normalized Diffraction Energy | Normalized Noise Energy | Ratio (%) |
|---------|-------------------------------|-------------------------|-----------|
| GAN     | 1.99 × 10^3                  | 5.97 × 10^4             | 3.340989  |
| Raw Data| 4.46 × 10^3                  | 7.06 × 10^4             | 6.317884  |
| PWD     | 2.87 × 10^3                  | 7.50 × 10^4             | 3.826558  |
| 20      | 2.27 × 10^3                  | 7.72 × 10^4             | 2.93889   |
| 40      | 2.45 × 10^3                  | 7.10 × 10^4             | 3.450413  |
| 60      | 2.08 × 10^3                  | 7.71 × 10^4             | 2.701722  |
| 80      | 2.60 × 10^3                  | 7.76 × 10^4             | 3.345916  |
| 100     | 2.18 × 10^3                  | 7.56 × 10^4             | 2.883725  |
| 120     | 2.47 × 10^3                  | 7.03 × 10^4             | 3.506477  |
| 140     | 2.64 × 10^3                  | 7.49 × 10^4             | 3.52547   |
| 160     | 2.60 × 10^3                  | 7.41 × 10^4             | 3.509198  |
| 180     | 3.97 × 10^3                  | 7.32 × 10^4             | 5.428411  |
| 200     | 3.35 × 10^3                  | 6.20 × 10^4             | 5.399523  |
| 2000    | 3.49 × 10^3                  | 4.93 × 10^4             | 7.062662  |
| 4000    | 3.24 × 10^3                  | 5.97 × 10^4             | 5.425227  |
| 8000    | 3.62 × 10^3                  | 4.63 × 10^4             | 7.830596  |
| 10,000  | 3.12 × 10^3                  | 6.22 × 10^4             | 5.010844  |
| 12,000  | 4.08 × 10^3                  | 5.18 × 10^4             | 7.868074  |
| 14,000  | 4.23 × 10^3                  | 6.76 × 10^4             | 6.258962  |
| 16,000  | 4.36 × 10^3                  | 5.38 × 10^4             | 8.097166  |
| 18,000  | 4.29 × 10^3                  | 5.16 × 10^4             | 8.324282  |
| 20,000  | 4.27 × 10^3                  | 4.78 × 10^4             | 8.934507  |

Note: The numbers for the GAN refer to the number of steps in training (a step refers to when a single image pair has been passed through the network). There is a slight increase in diffraction energy in some of the models and the PWD due to remnant reflection energy in the apexes of the diffractions, while when this decreases it shows some energy has been lost. Any increase in the noise may be a result of artefacts while decreases are generally favourable.

the performance improves, this is far from linear and appears logarithmic. While the quantitative analysis provides a useful figure for analysis, qualitative analysis is also required as artefacts within overlapping diffractions can contribute to overall diffraction energy skewing the result to appear better than it is. Quality control is more difficult with GAN separation as the network performs in a black-box fashion whereas with PWD there are measures for quality control, such as the dip field.

As shown in Table 1, the PWD and GAN at the final step have similar levels of noise (here defined as all energy outside of diffraction energy) with the GAN having slightly less noise, although this noise takes different forms as evident when analysing the images qualitatively (Fig. 5). In the GAN images, noise (i.e. energy outside of diffraction energy) is mostly remnant reflection energy, where this energy has been reduced but not removed. On the PWD image however, noise is from the multiple diffractions which are removed in the GAN image as they have been removed in the training data but are still present in the input data. This is a benefit of the GAN as the PWD operator destroys reflections, leaving a volume with diffractions and noise, while the GAN can be trained to remove noise as well as diffraction multiples (which have the same geometrical expression as diffractions). On top of this, towards the apex of diffraction hyperbolae the tops flatten which causes the PWD to recognize this as a continuous slope and therefore remove some of this energy. While the GAN has reduced some of the energy in the apex, this effect is lessened compared to the PWD example.

With a qualitative measure in place, additional data were collated for real applications which consisted of ~5000 256 × 256 images from field and synthetic data from a variety of geological and geographically distinct datasets (Fig. 6). The simple synthetic data were mostly generated by the authors although the more complex synthetic data, such as the SEAM dataset (Fehler and Keliher, 2011), came from open-source repositories. While synthetic sources are useful as they allow for perfect training data, they can never capture the full complexity of field data. To create training data from field data, diffraction-separated data are required. To do this, we have used plane-wave destruction and then screened the data to attempt to remove any bias from the neural network. Screening
the data is performed manually with areas where the PWD has not performed adequately removed from the training data. These areas, mainly areas without a continuously variable slope such as synclines and complex geologies, have large amounts of remnant reflection energy in the PWD-separated data. By removing these areas before training, the network is not introduced to these areas and does not create any negative associations with them (such as learning to retain the reflection energy which is present in the PWD training data). Therefore, when the network is exposed to these areas during prediction, it can make its own assumptions about them, which can be more favourable than PWD as the network is looking exclusively for hyperbolae as opposed to PWD which is dependent on the dip field (which can have errors of its own, especially in complex geologies). This changes the network from a simple PWD replicator to a more idealized PWD operator which can alleviate the drawbacks of the conventional PWD method. When combined with the synthetic data in the training phase, this is no longer a neural network representation of the PWD method.

With the full training data in place, the network underwent training over 100 epochs. Plots were created for every 25,000 steps to qualitatively assess the performance of the GAN as well as the same qualitative measure in place using the original synthetic dataset. The best models were then separated and tested individually to narrow down to an ideal model. This ideal model was then used to generate the prediction data on data which the network had no previous exposure to. With each new field data example, the GAN was applied blind to test the cross-data applicability of the GAN, before incorporating the data into the training dataset and applying transfer learning. The models are assessed in the same way as previously with the best model taken forward to be applied to another new dataset. When new datasets are incorporated, the GAN improves its prediction capabilities and cross-data applicability. Here, we have continued to train the GAN on additional datasets until the GAN can accurately predict on new data without the need for further training. The final trained model is freely available and can be applied blind to new datasets using the codes provided in this paper.
RESULTS

Using the pre-trained generative adversarial network (GAN) model, we demonstrate the effectiveness of the separation on field data. None of the lines shown in the Results section have been used in training the neural network though in the first field data example, adjacent lines have been incorporated in training. The first field data example shows the initial training of the neural network with field data (using the GAN pre-trained on synthetic data). Following this, the second field dataset shows the effect of additional training and incorporating multiple field datasets. We also analyse the effect of the neural network on the power and F-K spectra before and after additional training. Finally, we demonstrate the effectiveness of additional training data not only on new data but when revisiting data which has already been predicted. In the penultimate field dataset, we demonstrate the effectiveness of the fully trained GAN on new data. Ultimately, we show the performance of the GAN on 3D data which has been migrated to better reflect the diffraction energy.

Field data example 1 – initial training

The first field data example is from a deep-water marine environment on the continental shelf (Figs 7 and 8). The top 900 ms are characterized by heavily faulted sedimentary (sand/mud) layers with occasional channels and slump deposits, all of which cause intense diffractions. Underlying these layers are heavily fractured carbonates which also result in concentrated diffractions within this layer. Couple this with sedimentary features such as sand bodies and pinchouts scattered throughout the data and this provides an exemplary dataset to test diffraction imaging.

First, plane-wave destruction (PWD) was applied to the data both to give a benchmark comparison as well as to provide training data (Figs 7 and 8). Following this, the trained GAN was applied to the data (Figs 7 and 8). As seen here, the GAN effectively diminished the reflection energy leaving a comparable image with PWD, although it leaves more remnant reflection energy. However, there are also some areas where the GAN outperforms the PWD. These tend to correlate with complex areas where there may be difficulties calculating the dip as well as in the synclines which PWD struggles with but have been screened out of the GAN training data. This allows the GAN to make its own decision in these areas (as it has not been previously exposed to them outside of synthetic data). By applying PWD post GAN separation, it is possible to reduce the remnant reflections in the GAN image which can be

Figure 7 Raw data sorted to common offset (top), benchmark PWD data (middle) and the GAN data (bottom) that were trained using adjacent lines from this field dataset (however was not exposed to any other field datasets). As seen here, the GAN and PWD results are comparable, with similar problems as seen in the synthetic data where the GAN helps remove some background noise (compared to PWD) though has slightly more remnant reflection energy and lower frequency content. PWD has also missed some of the weaker, deeper, diffraction energy in the basement (from ≈3250 ms downwards) which is well visible in the GAN image. A close-up of the area shown by the red box is shown in Figure 8.
Pre-migration diffraction separation using a GAN

Figure 8 A zoomed in section from Figure 7 showing raw data sorted to common offset (top), PWD data (middle) and the GAN data (bottom) that was trained using adjacent lines from this field dataset (however was not exposed to any other field datasets). Here, a closer comparison of the GAN with the benchmark data can be seen, with more remnant reflection energy in the GAN (red arrows) although deeper diffraction energy is also visible (green arrows). Additionally, the PWD image has higher frequency.

Figure 9 Raw data sorted to common offset (top), PWD data (middle) and the original GAN data (bottom) from the second field dataset. The GAN here was applied with no additional training (it was not exposed to any data from this area prior to separation). Here, the GAN does not perform as well as the PWD in this area, with more remnant reflection energy and artefacts. However, there are some areas where the GAN outperforms PWD, especially in synclinal reflection energy, much of which is removed. While in the synthetic tests, synclinal energy was present in both PWD and GAN images, after further training the GAN has identified this as undesirable (likely due to the screening of training images which removed synclines from the training data).
Figure 10 Original GAN data (left), and the new GAN data (right) from the first field dataset. The original GAN is the GAN used in Figure 7 which was trained using just the first field dataset. The new GAN has been trained on both field datasets which has improved the separation with far less noise and remnant reflection energy than the original GAN.

Figure 11 Original GAN data (left), and the new GAN data (right) from the first field dataset. The original GAN is the GAN used in Figure 7 which was trained using just the first field dataset. The new GAN has been trained on both field datasets which has improved the separation with far less noise and remnant reflection energy (green arrows) than the original GAN although some diffraction energy in the flanks is also lost and there are some additional artefacts (red arrow).

useful for both interpretation and creating additional training data. A difficulty in the GAN separation comes from the amplitude consistency between 256 × 256 blocks. Occasionally, there is an amplitude disparity between blocks, however this is rare. This can easily be relieved by applying the network in larger blocks of data; however, this would then require more training as it effectively reduces the training data as the training data sizes and prediction data sizes must be consistent.

Field data example 2 – further training

The second field data example is also from deep-water marine environment but is distinct both geographically and geologically (Fig. 9). The data still appear on the continental shelf, however, has shallow dips in comparison to the first example. Additionally, the data are characterized more by channel complexes with minor faulting. The final challenge, for the GAN especially, is the distinct wavelet in comparison to previous data.

Plane-wave destruction was again applied to these data; however, this was not initially used in training. Instead, the pre-trained GAN were applied directly to the data with no additional training to test cross-data applicability (Fig. 9). As seen here, despite not having seen any data from this location previously, the GAN identifies and removes reflection energy well, highlighting underlying diffractions. However, there is more remnant reflection energy than in the previous examples in comparison to PWD and less high frequency. On top of this,
The power spectrum for the first field dataset (calculated as the average of the section shown in Figure 11) showing how PWD and the two GANs affect the spectrum of the data. As seen here, the original GAN has higher amplitudes, however, does not follow the shape of the target spectrum (PWD). With additional training, this issue is somewhat abated with the new GAN showing a spectrum far closer to the target spectra (PWD) though there is still some cut-off in the high frequencies. The GAN spectra are a composite of all the previous spectra input in the training data and thus attempt to replicate what they have seen previously. As such, the frequency is not consistent post-GAN separation which can cause issues. More training with additional wavelets further alleviates this issue and reduces any problems arising from it.

Figure 12

To test the hypothesis and prevent the power spectrum changes, the additional PWD data is incorporated into the training data (except for this line) and transfer learning is applied (i.e. further training on a pre-trained neural network). This improves the separation not only on this example (Fig. 10) but also on the previous data example (Fig. 11). Despite this improved separation, there are still some changes to the wavelet, with the improved separation showing a similar wavelet to the original separation but with lower frequencies reduced (Figs 12 and 13). As such, more field data were incorporated into the neural network from a variety of closed- and open-source data for a total of 14,132 screened training images. As more field data are added, the network improves its separation to the point the network can be used on new data instantly. While the initial training cost is high, this is a one-off cost and the network, when fully trained, can be applied directly on new data with a lower computational cost than PWD.

Field data example 3 (walker ridge) – application of fully trained network

To test the applicability of the fully trained neural network on new data, as well as to provide a reproducible example, open-source data was used from a seismic survey shot in Walker Ridge, Gulf of Mexico (Triezenberg et al., 2016). Walker Ridge is deeper water than the previous examples with water depth between 1500 and 2000 m. The seismic data were shot to image gas hydrates from gas-bearing sediments, however, there are many other areas of interest in the data (Miller et al., 2016). Here, we have focussed on several smaller areas which show diffractions including faults, rugose surfaces and hydrates themselves (Fig. 14).

As with previous examples, plane-wave destruction was applied to provide a comparison dataset (Figs 14 and 15). Here, as with Example 2, we have not incorporated any data from Walker Ridge in training and thus this provides an excellent opportunity to examine the cross-data applicability of the method however, with more data used to train the network than Example 2. Due to the increased data used in training, from a wide variety of field data sources, we can...
Raw data sorted to common offset (top), PWD data (middle), and the new GAN data (bottom) from Walker Ridge. The GAN has now been trained on several field datasets and can now be applied with no additional training to new data. Here, the GAN has also been applied blind (i.e. the GAN has not been trained on any of the data from this area). As seen here, the GAN image is far cleaner, with less noise and more diffraction energy than PWD which appears to have difficulty with diffractions from the seabed (possibly due to the comparatively high amplitudes of these diffractions). This may also have to do with the dip field as Walker Ridge is a relatively complex area and thus the dip field calculation may have difficulty (especially with smoothing which can further blur these lines and is required for PWD stability). While the PWD image may be improved with parameter tweaking, the GAN separation is acceptable with no parameterization. The trained model and data for this figure are available online.

Figure 15 Dip field calculated for PWD in Figure 14. Note that the dip field has captured some of the diffraction flanks which is why these have been destroyed in Figure 14.

anticipate the performance will be better than seen in Example 2 and this is seen to be the case in Figure 14. The broad pool of data used prevents any bias in the wavelet while removing reflection energy efficiently and effectively. As the data are manually screened prior to training, the remnant reflection energy present in synclines in PWD is diminished in the GAN results. Finally, as there was no need for additional training in this example, the processing speed improvement is drastic, with the GAN results obtained in a fraction of the time of the PWD flow (07:47:10 for PWD and 01:09:29 for GAN (MM:SS:ms) on an Intel i7 Quadcore processor). Note that the dip field (Fig. 15) calculation is included in the PWD flow time. While some energy may be missing from the GAN, the separation is reasonable and it can be reasoned that with even more data, the GAN could be further improved giving an even better separation.

Field data example 4 – application in 3D and migrated results

The final field data area is a large 3D dataset showing a complex, hyperextended rift basin, characterized by several highly fractured basin wide chalk layers. The fractured nature of these chalk layers causes them to be rife with diffraction energy and makes an excellent case study to test the 3D applicability of the method. Both GAN and PWD were applied to the data and the data migrated and, as with the previous example, no additional training was applied to the GAN before separation. Again we saw a dramatic speed-up with the GAN example taking 02:06:42 (HH:MM:SS) while the PWD flow took 25:57:57 (HH:MM:SS) to process the same area. The hardware used here is same as the previous section. A small section of the 3D was then analysed, with focus on a fractured carbonate layer at 2472 ms depth (Fig. 16). An inline
from these data was also taken to show the performance of both methods in a vertical slice (Fig. 17).

From both Figures 16 and 17, it is evident that the GAN is apt at locating and separating diffraction energy, with a performance comparable to the benchmark PWD dataset on both time slice and vertical sections. Both images have similar diffraction energy although, as with previous examples, the GAN images suffer from more remnant reflection energy. This further demonstrates the potential of GANs in the separation of diffraction energy on seismic data.

**DISCUSSION**

Diffraction separation is a complex task due to the nature of diffractions. Diffractions are low amplitude and blend with reflection energy, making separation difficult. On top of this, analytic-based methods require additional parameterization and inputs such as a dip field. Although these parameters may be used for quality control, they require additional computation and any errors introduced in these will be carried forward into the final image. While the generative adversarial network (GAN) also requires initial parameterization during training, application of a fully trained GAN requires no additional inputs. In this paper, we have outlined a novel method of separation which uses a GAN to automatically remove diffractions from the wavefield, allowing them to be independently processed. This has been demonstrated on three separate field datasets as well as synthetic data.

To train the GAN we have used a combination of synthetic and field data, with the field data created using plane-wave destruction and then screened to remove any bias (i.e. areas where plane-wave destruction (PWD) did not perform well, such as synclines and complex geologies were manually removed in the training). This improves the training of the neural network as it prevents the network picking up the ‘bad habits’ of the PWD operator and allows the network to make its own assumptions in these areas, free from any negative bias. This screening process also changes the GAN from a simple PWD replicator to an independent separation method which is akin to an idealized PWD image which can cope better with synclines and complex geologies as well as identify diffracted multiples. Based on observations, the GAN appears to be identifying hyperbolic shapes efficiently, recognizing where the shape and amplitude may be a syncline as opposed to a true diffraction or alternatively a multiple. As the GAN is solely identifying diffraction hyperbolae it is not dependent on any dip field and as such can identify hyperbolae in complex geologies without issue (whereas PWD will be dependent on the generated dip field which may be inaccurate in particularly complex areas).

On synthetic data, the GAN method and the benchmark PWD algorithm perform comparably with similar levels of diffraction energy and noise however, the nature of the noise
differs. Here, we define noise as anything which is not diffraction energy, which includes background noise, coherent noise and remnant reflection energy. The noise in PWD is generally background noise and coherent noise such as multiple diffractions (which appear visually akin to real diffractions). On the other hand, the noise in the GAN separation appears as reflection energy, with some of the background noise removed as well as the diffracted multiples reduced. This is due to the training of the GAN which involved synthetic data which were free from noise and multiples. As such, the GAN is more suitable in noisier data where diffractions may be obscured and as a first pass diffraction imaging test (due to its speed), whereas PWD may be used for finer diffraction imaging as it retains more high-frequency content.

The same observations can be made of field data, with the initial examples comparable to PWD but having more remnant reflection energy. However, after extensive training, the GAN separation increasingly improves. GAN separations also appear to alter the spectrum of the data between the desired output and the GAN output, and had more remnant reflection energy, although the benchmark PWD data often has a higher frequency content. Further exposure to a multitude of wavelets may improve this and prevent the changing of the wavelet. However, with the GAN there will come a point of diminishing returns where adding data no longer improves the separation. One of the difficulties with the GAN method is knowing when this point will be achieved as data are often scarce and synthetic data do not have the required complexity to improve the separation drastically. Alternatively, the power spectrums of each image may be able to be used as an additional input, forcing the neural network to satisfy both time–space and power spectrum representations of the data. Despite the flaws, however, the GAN provides a fast and accurate method of separating diffraction and reflection energy.

The current GAN has been directly applied to both 3D dataset and open-source datasets with no additional training allowing it to perform a separation seven times faster than the equivalent plane-wave destruction separation (when including the dip field calculation) and with a qualitatively comparable diffraction image both pre- and post-migration. This decrease in processing time can be substantial when handling large datasets when no further training is required. Overall, we have demonstrated the potential for GAN as a useful tool in diffraction extraction.

ACKNOWLEDGEMENTS

This research has emanated from research supported in part by a research grant from Science Foundation Ireland (SFI) under grant number 13/RC/2092 and is co-funded under the European Regional Development Fund by PIPCO RSG and its member companies. The authors extend their gratitude to Tullow Oil and the Petroleum Affairs Division of Ireland for providing field data used in training. The authors would also like to thank Song Hou, Henning Hoeber and Ewa Kaszycka.
of CGG for their discussions on neural networks and diffractions. Finally, the authors would like to thank Shearwater for providing an academic license for Shearwater Reveal, which was used in this study.

Open access funding provided by IReL.

CODE AND DATA AVAILABILITY

Codes for the neural network (both training and prediction) and the final trained network used for the Walker Ridge and the 3D data example are available at https://github.com/b-lowney/GAN. These codes were written using original Pix2Pix code from https://machinelearningmastery.com/how-to-develop-a-pix2pix-gan-for-image-to-image-translation/. The original Walker Ridge data can be found at https://walrus.wr.usgs.gov/namss/survey/p1-13-la-walker-ridge/. The other field data examples are confidential and cannot be shared. Madagascar software was used for the PWD results and as such these are fully reproducible also (parameters are shared. Madagascar software was used for the PWD results. The other field data examples are confidential and cannot be shared. Madagascar software was used for the PWD results and as such these are fully reproducible also (parameters are shared).

REFERENCES

Aharchaou, M. and Baumstein, A. (2020) Deep learning-based artificial bandwidth extension: training on ultrasonic OBN to enhance towed-streamer FWI. The Leading Edge, 39, 718–726.
Alotaibi, A. (2020) Deep generative adversarial networks for image-to-image translation: a review. Symmetry, 12, 1705–1731.
Basheer, I. and Hajmeer, M. (2000) Artificial neural networks: fundamentals, computing, design, and application. Journal of Microbiological Methods, 43, 3–31.
Berkovitch, A., Beller, I., Hassin, Y. and Landa, E. (2009) Diffraction imaging by multifocusing. Geophysics, 74(6), WCA75–WCA81.
Bonnefoy-Claudet, S., Cotton, F. and Bard, P. (2006) The nature of noise wavefield and its applications for site effects studies: a literature review. Earth-Science Reviews, 79, 205–227.
Brownlee, J. (2019) How to develop a Pix2Pix GAN for image-to-image translation. Accessed December 1, 2020. Available at: http://www.machinelearningmastery.com/how-to-develop-a-pix2pix-gan-for-image-to-image-translation.
Cai, L., Gao, H. and Ji, S. (2019) Multi-stage variational autoencoders for coarse-to-fine image generation. Proceedings of the 2019 SIAM International Conference on Data Mining, Spokane, WA, pp. 630–638.
Chen, Z., Fomel, S. and Lu, W. (2013) Accelerated plane-wave destruction. Geophysics, 78(1), V1-V9.
Claerbout, J. (1985) Fundamentals of Geophysical Data Processing with Applications to Petroleum Prospecting. Hoboken, NJ: Blackwell Scientific Publications.
Decker, L., Janson, X. and Fomel, S. (2015) Carbonate reservoir characterization using seismic diffraction imaging. Interpretation, 3(1), SF21–SF30.
Decker, L., Klokov, A. and Fomel, S. (2013) Comparison of seismic diffraction imaging techniques: plane-wave destruction versus apex destruction. 83rd SEG Annual International Meeting, Houston, TX, Expanded Abstracts, pp. 4054–4059.
Dell, S. and Gajewski, D. (2011) Common-reflection-surface-based workflow for diffraction imaging. Geophysics, 76(5), S187–S195.
Fehler, M. and Keliher, P. (2011) SEAM Phase I: Challenge of Subsalt Imaging in Tertiary Basins, with Emphasis on Deepwater Gulf of Mexico. Tulsa, OK: Society of Exploration Geophysicists.
Fomel, S. (2002) Applications of plane-wave destruction filters. Geophysics, 67(6), 1946–1960.
Fomel, S., Landa, E. and Taner, M. (2007) Poststack velocity analysis by separation and imaging of seismic diffractions. Geophysics, 72(6), U89–U94.
Fomel, S., Sava, P., Vlad, I., Liu, Y. and Bashkardin, V. (2013) Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments. Journal of Open Research Software, 1(1), e8.
Gelius, L. and Asgedom, E. (2011) Diffraction-limited imaging and beyond – the concept of super resolution. Geophysical Prospecting, 59, 400–421.
Goodfellow, I. (2016) NIPS 2016 tutorial: generative adversarial networks. Neural Information Processing Systems 2016, Barcelona, Spain, pp. 1–57.
Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. et al., (2014) Generative adversarial nets. Neural Information Processing Systems 2014, Montreal, Canada, pp. 2672–2680.
Han, C., Murao, K., Sato, S. and Nakayama, H. (2019) Learning more with less: GAN-based medical image augmentation. ACM International Conference on Information and Knowledge Management, Beijing.
Harlan, W., Claerbout, J. and Rocca, F. (1984) Signal/noise separation and velocity estimation. Geophysics, 49(11), 1869–1880.
Ho, Y., Wookey, S. (2019) The real-world-weight cross-entropy loss function: modeling the costs of mislabeling. IEEE Access, 8, 4806–4813.
Isola, P., Zhu, J., Zhou, T. and Efros, A. (2017) Image-to-image translation with conditional adversarial networks. IEEE Conference on Computer Vision and Pattern Recognition, Manchester, pp. 1125–1134.
JafarGandomi, A., Bukola, O., Refaar, R. and Hoeber, H. (2019) Specular imaging of converted wave data and its AVO impact. 81st EAGE Conference and Exhibition, London, Expanded Abstracts, pp. 1–5.
Kahng, M., Thorat, N., Chau, D.H.P., Viégas, F.B. and Wattenberg, M. (2018) GAN lab: understanding complex deep generative

© 2021 The Authors. Geophysical Prospecting published by John Wiley & Sons Ltd on behalf of European Association of Geoscientists & Engineers, Geophysical Prospecting, 69, 949–967
models using interactive visual experimentation. *IEEE Transactions on Visualization and Computer Graphics*, 25, 1–11.

Kanasewich, E. and Phadke, S. (1988) Imaging discontinuities on seismic sections. *Geophysics*, 53(3), 334–345.

Kennett, B. (2001) *The Seismic Wavefield: Introduction and Theoretical Development (Vol. 1)*. Cambridge: Cambridge University Press.

Khaidukov, V., Landa, E. and Moser, T. (2004) Diffraction imaging by focusing-defocusing: an outlook on seismic superresolution. *Geophysics*, 69(6), 1478–1490.

Klokov, A., Baina, R. and Landa, E. (2010) Separation and imaging of seismic diffractions in dip-angle domain. 72nd EAGE Conference and Exhibition, Barcelona, Spain, Expanded Abstracts, p. CP161.

Klokov, A. and Fomel, S. (2012) Separation and imaging of seismic diffractions using migrated dip-angle gathers. *Geophysics*, 77(6), S131–S143.

Knerr, S., Personnaz, L. and Dreyfus, G. (1990) Single-layer learning revisited: a stepwise procedure for building and training a neural network. In: Soulé, F.F. and Hérault, J. (Eds.) *Neurocomputing*. Heidelberg: Springer, pp. 41–50.

Kozlov, E., Barasky, N., Korolev, E., Antonenko, A. and Koshchuk, E. (2004) Imaging scattering objects masked by specular reflections. *74th SEG Annual International Meeting*, Denver, CO, Expanded Abstracts, pp. 1131–1134.

Krey, T. (1952) The significance of diffraction in the investigation of faults. *Geophysics*, 17(4), 843–858.

Landa, E. (2007) *Beyond Conventional Seismic Imaging*. Houten, The Netherlands: European Association of Geoscientists and Engineers. Landa, E., Fomel, S. and Moser, T. (2006) Path-integral seismic imaging. *Geophysical Prospecting*, 54, 491–503.

Landa, E., Fomel, S. and Reshef, M. (2008) Separation, imaging, and velocity analysis of seismic diffractions using migrated dip-angle gathers. *78th SEG Annual International Meeting*, Las Vegas, NV, pp. 2176–2180.

Li, R., Liu, W., Yang, L., Sun, S., Hu, W., Zhang, F. et al., (2018) DeepUNet: a deep fully convolutional network for pixel-level sea-land segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11, 3954–3962.

Lian, S., Luo, Z., Zhong, Z., Lin, X., Su, S. and Li, S. (2018) Attention guided U-Net for accurate iris segmentation. *Journal of Visual Communication*, 56, 296–304.

Lowney, B., Lokmer, I., O’Brien, G.S. and Bean, C.J. (2020) Direct diffraction separation by deep learning on pre-migrated seismic data. *1st EAGE Annual Conference Online*. Houten, The Netherlands: European Association of Geoscientists and Engineers, pp. 1–5.

Maciel, S. and Bilitori, R. (2014) Detection of diffractions in seismic sections using Support Vector Classifiers. *74th SEG Annual International Conference*, Denver, CO, Expanded Abstracts, pp. 4816–4820.

Martini, E., Bean, C.J., Dolan, S. and Marsan, D. (2001) Seismic image quality beneath strongly scattering structures and implications for lower crustal imaging. *Geophysical Journal International*, 145, 423–435.

Miller, J., Agena, W., Haines, S. and Hart, P. (2016) Processing of multi-channel seismic reflection data acquired in 2013 for seismic investigation of gas hydrates in the Gulf of Mexico. Reston, VA: U.S. Geological Survey. Open-File Report.

Moser, T.J. and Howard, C. (2008) Diffraction imaging in depth. *Geophysical Prospecting*, 56, 627–641.

Nguyen, T., Le, T., Vu, H. and Phung, D. (2017) Dual discriminator generative adversarial nets. *Advances in Neural Information Processing Systems*, 1, 2670–2680.

O’Brien, G.S. (2020) Common image gathering conditioning using cycle generative adversarial networks. *Geophysical Prospecting*, 68, 1758–1770.

Oliveira, D., Ferreira, R., Silva, R. and Brazil, E. (2018) Interpolating seismic data with conditional generative adversarial networks. *IEEE Geoscience and Remote Sensing Letters*, 15, 1952–1956.

Popovici, A., Sturzu, I. and Moser, T. (2015) High-resolution diffraction imaging of small-scale fractures in shale and carbonate reservoirs. *Unconventional Resources Technology Conference*, San Antonio, TX, pp. 1121–1129.

Radford, A., Metz, L. and Chintala, S. (2016) Unsupervised representation learning with deep convolutional generative adversarial networks. *International Conference on Learning Representations*, San Juan, Puerto Rico.

Reshef, M. and Landa, E. (2009) Post-stack velocity analysis in the dip-angle domain using diffractions. *Geophysical Prospecting*, 57, 811–821.

Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Boston, MA, pp. 234–241.

Sankesara, H. (2019) U-Net: introducing symmetry in segmentation. Accessed September 16, 2020. Available at: https://towardsdatascience.com/u-net-b229b32b4a71.

Schwarz, B. (2019a) An introduction to seismic diffraction. In: C. Schmelzbach (Ed.) *Advances in Geophysics: Recent Advances in Seismology*. New York: Academic Press, pp. 1–64.

Schwarz, B., (2019b) Coherent wavefield subtraction for diffraction separation. *Geophysics*, 84(3), V157-V168.

Serfati, Y., Itan, L., Levy, R. and Koren, Z. (2018) Application of deep learning along directional image gathers for high-definition classification of subsurface features. *80th EAGE Conference and Exhibition*, Copenhagen, Denmark, Expanded Abstracts, pp. 1–5.

Shekhar, A. (2019) What are L1 and L2 loss functions? Accessed October 3, 2020. Available at: https://afteracademy.com/blog/what-are-l1-and-l2-loss-functions.

Sun, J., Slang, S., Elboth, T., Larsen Greiner, T., McDonald, S. and Gelius, L.J. (2020) Attenuation of marine seismic interference noise employing a customized U-Net. *Geophysical Prospecting*, 68, 845–871.

Tiwari, Y., Rasool, A. and Hajela, G. (2020) Machine learning with generative adversarial networks. *Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, Expanded Abstracts, pp. 543–548.

Tran, N., Tran, V., Nguyen, B. and Yang, L. (2019) Self-supervised GAN: analysis and improvement with multi-class minimax game.
Pre-migration diffraction separation using a GAN

Advances in Neural Information Processing Systems, 32, 13253–13264.

Triezenberg, P. J., Hart, P. E. and Childs, J. R. (2016) National Archive of Marine Seismic Surveys (NAMSS): A USGS data website of marine seismic reflection data within the U.S. Exclusive Economic Zone (EEZ). Reston, VA: U.S. Geological Survey. s.l.: U.S. Geological Survey Data Release.

Trorey, A. W. (1970) A simple theory for seismic diffractions. Geophysics, 35(5), 762–784.

Tschannen, V., Ettrich, N., Delescluse, M. and Keuper, J. (2020) Detection of point scatterers using diffraction imaging and deep learning. Geophysical Prospecting, 68, 830–844.

Xu, R., Zhou, Z., Zhang, W. and Yu, Y. (2017) Face transfer with generative adversarial network. Conference on Computer Vision and Pattern Recognition, Manchester, Expanded Abstracts.

Yuan, Y., Si, X. and Zheng, Y. (2020) Ground roll attenuation using generative adversarial networks. Geophysics, 85(4), IJA–Z18.

Zhang, T., Fu, H., Zhao, Y., Cheng, J., Guo, M., Gu, Z. et al., (2019) SkrGAN: sketching-rendering unconditional generative adversarial networks for medical imaging synthesis. International Conference on Medical Image Computing and Computer-Assisted Intervention, Lima, Peru, Expanded Abstracts, pp. 777–785.

© 2021 The Authors. Geophysical Prospecting published by John Wiley & Sons Ltd on behalf of European Association of Geoscientists & Engineers, Geophysical Prospecting, 69, 949–967