

De-risking geological carbon storage from high resolution time-lapse seismic to explainable leakage detection

Ziyi Yin, Huseyin Tuna Erdinc, Abhinav Prakash Gahlot, Mathias Louboutin, Felix J. Herrmann
Georgia Institute of Technology

Abstract

Geological carbon storage represents one of the few truly scalable technologies capable of reducing the CO₂ concentration in the atmosphere. While this technology has the potential to scale, its success hinges on our ability to mitigate its risks. An important aspect of risk mitigation concerns assurances that the injected CO₂ remains within the storage complex. Amongst the different monitoring modalities, seismic imaging stands out with its ability to attain high resolution and high fidelity images. However, these superior features come, unfortunately, at prohibitive costs and time-intensive efforts potentially rendering extensive seismic monitoring undesirable. To overcome this shortcoming, we present a methodology where time-lapse images are created by inverting non-replicated time-lapse monitoring data jointly. By no longer insisting on replication of the surveys to obtain high fidelity time-lapse images and differences, extreme costs and time-consuming labor are averted. To demonstrate our approach, hundreds of noisy time-lapse seismic datasets are simulated that contain imprints of regular CO₂ plumes and irregular plumes that leak. These time-lapse datasets are subsequently inverted to produce time-lapse difference images used to train a deep neural classifier. The testing results show that the classifier is capable of detecting CO₂ leakage automatically on unseen data and with a reasonable accuracy.

Introduction

For various reasons, seismic monitoring of geological carbon storage (GCS) comes with its own set of unique challenges. Amongst these challenges, the need for low-cost highly repeatable, high resolution, and high fidelity images ranks chiefly. While densely sampled and replicated time-lapse surveys—which rely on permanent reservoir monitoring systems—may be able to provide images conducive to interpretation and reservoir management, these approaches are often too costly and require too much handholding to be of practical use for GCS.

To overcome these challenges, we replace the current paradigm of costly replicated acquisition, cumbersome time-lapse processing, and interpretation, by a joint inversion framework mapping time-lapse data to high fidelity and high resolution images from sparse non-replicated time-lapse surveys. We demonstrate that we arrive at an imaging framework that is suitable for automatic detection of pressure-induced CO₂ leakage. Rather than relying on meticulous 4D workflows where baseline and monitoring surveys are processed separately to yield accurate and artifact-free time-lapse differences, our approach exposes information that is shared amongst the different vintages by formulating the imaging problem in terms of an unknown fictitious common component, and innovations of the baseline and monitor survey(s) with respect to this common component. Because the common component is informed by all time-lapse surveys, its image quality improves when the surveys bring complementary information, which is the case when the surveys are not replicated. In turn, the enhanced common component results in improved images for the baseline, monitor, and their time-lapse difference. Joint inversion also leads to robustness with respect to noise, calibration errors, and time-lapse changes in the background velocity model.

To showcase the achievable imaging gains and how these can be used in a GCS setting where CO₂ leakage is of major consideration, we create hundreds of time-lapse imaging experiments involving CO₂ plumes whose behavior is determined by the two-phase flow equations. To mimic irregular flow due to pressure-induced opening of fractures, we increase the permeability in the seal at random locations and pressure thresholds. The resulting flow simulations are used to generate time-lapse datasets that serve as input to our joint imaging scheme. The produced time-lapse difference images are subsequently used to train and test a neural network that as an explainable classifier determines whether the CO₂ plume behaves regularly or shows signs of leakage.

Our contributions are organized as follows. First, we discuss the time-lapse seismic imaging problem and its practical difficulties. Next, we introduce the joint recovery model that takes explicit advantage of information shared by multiple surveys. By means of a carefully designed synthetic case study involving saline aquifers made of Blunt sandstone in the Southern North Sea, we demonstrate the uplift of the joint recovery model and how its images can be used to train a deep neural network classifier to detect erroneous growth of the CO₂ plume automatically. Aside from determining whether the CO₂ plume behaves regularly or not, our network also provides class activation mappings that visualize areas in the image on which the network is basing its classification.

Seismic monitoring with time-lapse imaging

To keep track of CO₂ plume development during geological carbon storage (GCS) projects, multiple time-lapse surveys are collected. Baseline surveys are acquired before the supercritical CO₂ is injected into the reservoir. These baseline surveys, denoted by the index $j = 1$, are followed by one or more monitor surveys, collected at later times and indexed by $j = 2, \dots, n_v$ with n_v the total number of surveys.

Seismic monitoring of GCS brings its own unique set of challenges that stem from the fact that its main concern is (early) detection of possible leakage of CO₂ from the storage complex. To be successful with this task, monitoring GCS calls for a time-lapse imaging modality that is capable of

- detecting weak time-lapse signals associated with small rock-physics changes induced by CO₂ leakage
- attaining high lateral resolution from active-source surface seismic data to detect vertically moving leakage
- handling an increasing number of seismic surveys collected over long periods of time ($\sim 50 - 100$ years)
- reducing costs drastically by no longer insisting on replication of time-lapse surveys to attain high degrees of repeatability
- lowering the cumulative environmental imprint of active source acquisition

Monitoring with the joint recovery model

To meet these challenges, we choose a linear imaging framework where observed linearized data for each vintage is related to perturbations in the acoustic impedance via

$$\mathbf{b}_j = \mathbf{A}_j \mathbf{x}_j + \mathbf{e}_j \quad \text{for } j = 1, 2, \dots, n_v. \quad (1)$$

In this expression, the matrix \mathbf{A}_j stands for the linearized Born scattering operator for the j th vintage. Observed linearized data, collected for all shots in the vector \mathbf{b}_j , is generated by applying the \mathbf{A}_j 's to the (unknown) impedance perturbations denoted by \mathbf{x}_j for $j = 1, 2, \dots, n_v$. The task of time-lapse imaging is to create high resolution, high fidelity and true amplitude estimates for the time-lapse images, $\{\hat{\mathbf{x}}_j\}_{j=1}^{n_v}$, from non-replicated sparsely sampled noisy time-lapse data.

We argue that our choice for linearized imaging is justified for four reasons. First, CO₂-injection sites undergo extensive baseline studies, which means that accurate information on the background velocity model is generally available. Second, changes in the acoustic parameters induced by CO₂ injection are typically small, so it suffices to work with one and the same background model for the baseline and monitor surveys. Third, when the background model is sufficiently close to the true model, linearized inversion, which corresponds to a single Gauss-Newton iteration of full-waveform inversion, converges quadratically. Fourth, because the forward model is linear, it is conducive to the use of the joint recovery model where inversions are carried out with respect to the common component, which is shared between all vintages, and innovations with respect to the common component. Because the common component represents an average, we expect this joint imaging method to be relatively robust with respect to kinematic changes induced by time-lapse effects or by lack of calibration of the acquisition [Oghenekohwo and Herrmann, 2017b].

By parameterizing time-lapse images, $\{\mathbf{x}_j\}_{j=1}^{n_v}$, in terms of the common component, \mathbf{z}_0 , and innovations with respect to the common component, $\{\mathbf{z}_j\}_{j=1}^{n_v}$, we arrive at the joint recovery model where representations for the images are given by

$$\mathbf{x}_j = \frac{1}{\gamma}\mathbf{z}_0 + \mathbf{z}_j \quad \text{for } j = 1, 2, \dots, n_v. \quad (2)$$

Here, the parameter, γ , controls the balance between the common component, \mathbf{z}_0 , and innovation components, $\{\mathbf{z}_j\}_{j=1}^{n_v}$ [Li, 2015]. Compared to traditional time-lapse approaches, where data are imaged separately or where time-lapse surveys are subtracted, inversions for time-lapse images based on the above parameterization are carried out jointly and involve inverting the following matrix:

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\gamma}\mathbf{A}_1 & \mathbf{A}_1 & & \\ \vdots & & \ddots & \\ \frac{1}{\gamma}\mathbf{A}_{n_v} & & & \mathbf{A}_{n_v} \end{bmatrix}. \quad (3)$$

While traditional time-lapse imaging approaches strive towards maximal replication between the surveys to suppress acquisition related artifacts, imaging with the joint recovery model—which entails inverting the underdetermined system in Equation 3 using structure-promotion techniques (e.g. via ℓ_1 -norm minimization)—improves the image quality of the vintages themselves in situations where the surveys are not replicated. This occurs in cases where $\mathbf{A}_i \neq \mathbf{A}_j$ for $\forall i \neq j$, or in situations where there is significant noise. This remarkable result was shown to hold for sparsity-promoting denoising of time-lapse field data [Wei et al., 2018, Tian et al., 2018], for various wavefield reconstructions of randomized simultaneous-source dynamic (towed-array) and static (OBC/OBN) marine acquisitions [Oghenekohwo and Herrmann, 2017b, Kotsi, 2020, Zhou and Lumley, 2021], and for wave-based inversion, including least-squares reverse-time migration and full-waveform inversion [Oghenekohwo, 2017, Oghenekohwo and Herrmann, 2017a]. The observed quality gains in these applications can be explained by improvements in the common component resulting from complementary information residing in non-replicated time-lapse surveys. This enhanced recovery of the common component in turn improves the recovery of the innovations and therefore the vintages themselves. The time-lapse differences themselves also improve, or at the very least, remain relatively unaffected when the surveys are not replicated. Relaxing replication of surveys obviously leads to reduction in cost and environmental impact. Below, we show how GCS monitoring also benefits from this approach.

Monitoring with curvelet-domain structure promotion

To obtain high resolution and high fidelity time-lapse images, we invert the system in Equation 3 [Yang et al., 2020, Witte et al., 2019b, Yin et al., 2021] with

$$\begin{aligned} \underset{\mathbf{z}}{\text{minimize}} \quad & \lambda\|\mathbf{Cz}\|_1 + \frac{1}{2}\|\mathbf{Cz}\|_2^2 \\ \text{subject to} \quad & \|\mathbf{b} - \mathbf{Az}\|_2^2 \leq \sigma, \end{aligned} \quad (4)$$

where \mathbf{C} is the forward curvelet transform, λ the threshold parameter, and σ the magnitude of the noise. At iteration k and for $\sigma = 0$, solving Equation 4 corresponds to computing the following iterations:

$$\begin{aligned}\mathbf{u}_{k+1} &= \mathbf{u}_k - t_k \mathbf{A}_k^\top (\mathbf{A}_k \mathbf{z}_k - \mathbf{b}_k) \\ \mathbf{z}_{k+1} &= \mathbf{C}^\top S_\lambda(\mathbf{C} \mathbf{u}_{k+1}),\end{aligned}\tag{5}$$

where \mathbf{A}_k , with a slight abuse of notation, represents the matrix in Equation 3 for a subset of shots randomly selected from sources in each survey. The vector \mathbf{b}_k contains the extracted shot records from \mathbf{b} and the symbol $^\top$ refers to the adjoint. Sparsity is promoted via curvelet-domain soft thresholding, $S_\lambda(\cdot) = \max(|\cdot| - \lambda, 0) \text{sign}(\cdot)$, where, λ , is the threshold. The vectors \mathbf{u}_k and \mathbf{z}_k contain the baseline and innovation components.

Numerical case study: Blunt sandstone in the Southern North Sea

Before discussing the impact of high resolution and high fidelity time-lapse imaging with the joint recovery model on the down-stream task of automatic leakage detection with a neural network classifier, we first detail the setup of our numerical experiments using techniques from simulation-based acquisition design as described by Yin et al. [2021]. In order to generate realistic time-lapse data and training sets for the automatic leakage classifier, we follow the workflow summarized in Figure 1. In this approach, use is made of proxy models for seismic properties derived from real 3D imaged seismic and well data [Jones et al., 2012]. With rock physics, these seismic models are converted to fluid-flow models that serve as input to two-phase flow simulations. The resulting datasets, which include pressure-induced leakage, will be used to create time-lapse data used to train our classifier. For more detail, refer to the caption of Figure 1.

Proxy seismic and fluid-flow models

Amongst the various CO_2 injection projects, GCS in offshore saline aquifers has been most successful in reaching scale and in meeting injection targets. For that reason, we consider a proxy model constructed representative for CO_2 injection in the South of the North Sea involving a saline aquifer made of the highly permeable Blunt sandstone. This area, which is actively being considered for GCS [Kolster et al., 2018], consists of the following three geologic sections (see Figure 2 for the permeability and porosity distribution):

- (i) the highly porous (average 33%) and permeable (> 170 mD) Blunt sandstone reservoir of about 300 – 500 m thick. This section, denoted by red colors in Figure 2, corresponds to the saline aquifer and serves as the reservoir for CO_2 injection;
- (ii) the primary seal (permeability $10^{-4} - 10^{-2}$ mD) made of the Rot Halite Member, which is 50 m thick and continuous (black layer in Figure 2);
- (iii) the secondary seal made of the Haisborough group, which is > 300 m thick and consists of low-permeable (permeability 15 – 18 mD) mudstone (purple section in Figure 2).

To arrive at the fluid-flow models, we consider 2D subsets of the 3D Compass model [Jones et al., 2012] and convert these seismic models to fluid-flow properties (see Figure 1 (b)) by assuming a linear relationship between compressional wavespeed and permeability in each stratigraphic section. For further details on the conversion of compressional wavespeed and density to permeability and porosity, we refer to empirical relationships reported in Klimentos [1991]. During conversion, an increase of 1km/s in compressional wavespeed is assumed to correspond to an increase of 1.63mD in permeability. From this, porosity is calculated with the Kozeny-Carman equation [Costa, 2006] $K = \phi^3 \left(\frac{1.527}{0.0314*(1-\phi)} \right)^2$, where K and ϕ denote permeability (mD) and porosity (%) with constants taken from the Strategic UK CCS Storage Appraisal Project report.

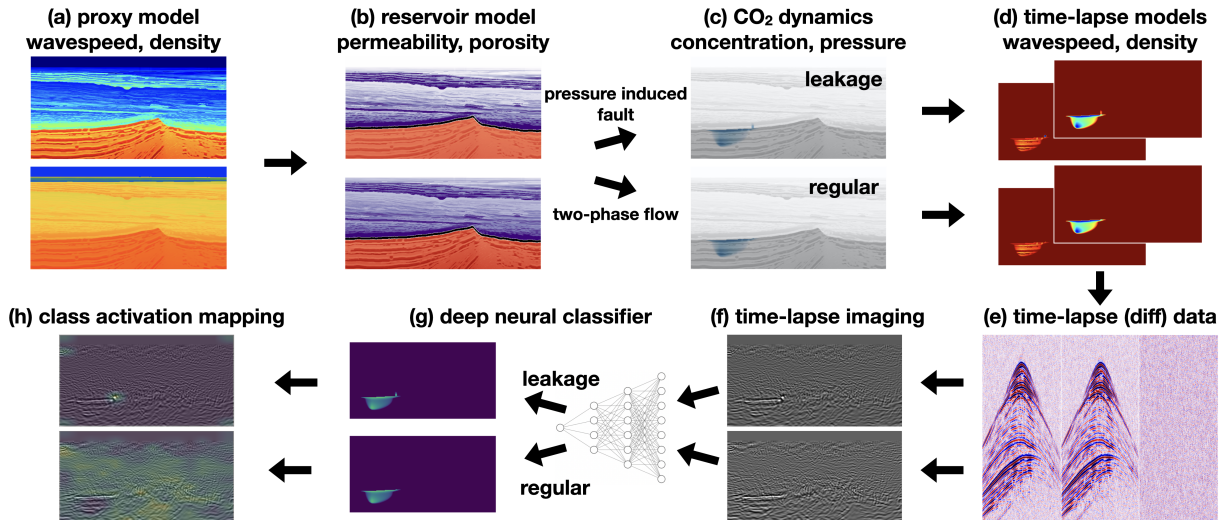


Figure 1: Simulation-based monitoring design framework. Starting with a proxy model for the wavespeed and density (a), the workflow proceeds by converting these seismic properties into permeability and porosity (b). These fluid flow properties are used to simulated CO₂ plumes that behave regularly or exhibit leakage outside the storage complex (c). Induced changes by the CO₂ plume for the wavespeed and density are depicted in (d) and serve as input to simulations of time-lapse seismic data (SNR 8.0 dB) and shot-domain time-lapse differences (SNR -31.4 dB). Imaging results for regular and irregular plume developments are plotted in (f) and serve as input to the deep neural classifier (g), which determines whether the flow behaves regularly or leaks. Activation mappings in (h) show regions on which the network is basing its classification. As expected, the activation mapping is diffusive in case of regular CO₂ plume development and focused on the leakage location when CO₂ plume behaves irregularly.

Fluid-flow simulations

To model CO₂ plumes that behave regularly and irregularly, the latter due to leakage, we solve the two-phase flow equations numerically¹ for both pressure and concentration [Li and Xu, 2021, Li et al., 2020]. To mimic possible pressure-induced CO₂ leakage, we increase the permeability at random distances away from the injection well within the seal from 10^{-4} mD to 500 mD when the pressure exceeds ~ 15 MPa. At that depth, the pressure is below the fracture gradient [Ringrose, 2020]. Since pressure-induced fractures come in different sizes, we also randomly vary the width of the pressure-induced fracture openings from 12.5 m to 62.5 m. Examples of fluid-flow simulations without and with leakage are shown in Figure 1 (c).

Rock-physics conversion

To monitor temporal variations in the plume’s CO₂ concentration seismically, we use the patchy saturation model [Avseth et al., 2010] to convert the CO₂ concentration to decrease in compressional wavespeed and density. These changes are shown in Figure 1 from (d). The fact that these induced changes in the time-lapse differences in seismic properties are relatively small in spatial extent (~ 800 m for the plume and < 62.5 m for the leakage) and amplitude (18% of the impedance perturbations with respect to the background model) calls for a time-lapse imaging modality with small normalized root-mean-square (NRMS) [Kragh and Christie,

¹We used the open-source software [FwiFlow.jl](#) [Li and Xu, 2021, Li et al., 2020] to solve the two-phase flow equations for both the pressure and concentration.

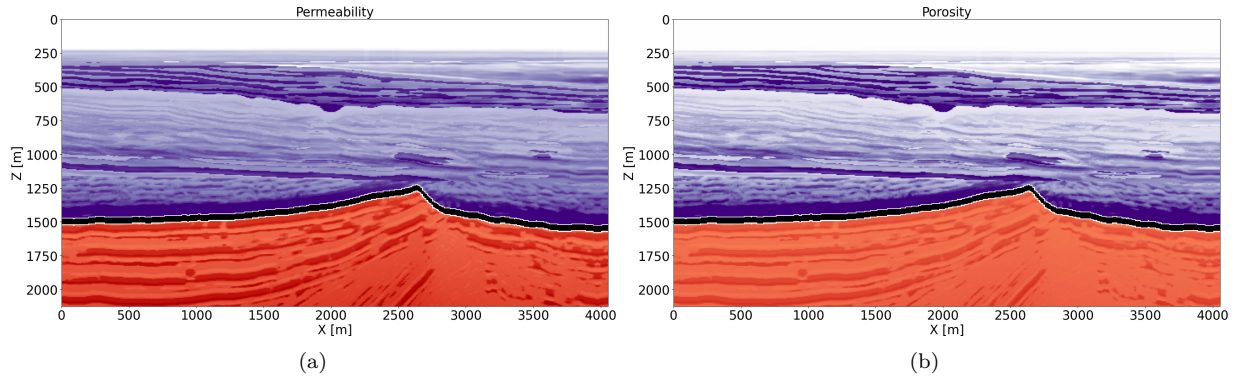


Figure 2: Permeability and porosity derived from a 2D slice of Compass model.

2002] values ($\sim 2.5\%$). This NMRS value is based on the relative amplitude of impedance perturbations with respect to the background model (14%).

Time-lapse seismic simulations

To train and validate automatic detection of CO_2 leakage from the storage complex requires the creation of realistic time-lapse datasets that contain the seismic imprint of regular as well as irregular (leakage) plume development. To this end, baseline surveys are simulated prior to CO_2 injection for different subsets of the Compass model. Monitor surveys are simulated 200 days after leakage occurs to verify that potential leakage can be detected automatically early on. For regular plume development, we shoot monitor surveys for each subset at random times after CO_2 injection. In order to strike a balance between acquisition productivity and time-lapse image quality, use is made of dense permanent acoustic monitoring at the seafloor with 25 m receiver spacing. Time-lapse acquisition costs are reduced by non-replicated coarse shooting with the source towed at 10m depth below the ocean surface. Subsampling artifacts are reduced by using a randomized technique from compressive sensing where 32 sources are located at non-replicated jittered [Herrmann and Hennenfent, 2008] source positions, yielding an average source sampling of 125 m. Given this acquisition geometry, linear data is generated² with Equation 1 for a 25 Hz Ricker wavelet and with the band-limited noise term set so that the data’s signal-to-noise ratio (SNR) is 8.0 dB. This noise level leads to an extremely poor SNR of -31.4 dB for time-lapse differences in the shot domain. See Figure 1 (e).

Imaging with joint recovery model versus reverse-time migration

Given the simulated time-lapse datasets with and without leakage, time-lapse difference images are created according to two different imaging scenarios, namely via independent reverse-time migration (RTM), conducted on the baseline and monitor surveys separately, and via inversion of the joint recovery model (cf. Equations 3 and 4). To limit the computational cost of the Bregman iterations (Equation 5), four shot records are selected per iteration at random from each survey for imaging [Yin et al., 2008, Witte et al., 2019b, Yang et al., 2020, Yin et al., 2021], limiting the cost of the joint inversion to three RTMs. The recovered baseline images are shown in Figures 3a for RTM and 3b for JRM. For the leakage scenario, the time-lapse differences are plotted in Figures 3c and 3d, for RTM and JRM respectively. For the regular plume, the time-lapse differences are plotted in Figures 3e and 3f, for RTM and JRM respectively. From these images, it is clear that joint

²We used the open-source software [JUDI.jl](#) [Witte et al., 2019a, Louboutin et al., 2022] to model the wave propagation. This Julia package calls the highly optimized propagators of [Devito](#) [Louboutin et al., 2019, Luporini et al., 2020, 2022].

inversion leads to relatively artifact-free recovery of the vintages and time-lapse differences. This observation is reflected in the NRMS values, which improve considerably as shown by the histograms in Figure 4 for 1000 imaging experiments. Not only do the NRMS values shift towards the left, their values are also more concentrated when inverting time-lapse data with the joint recovery model. Both features are beneficial to automatic leakage detection.

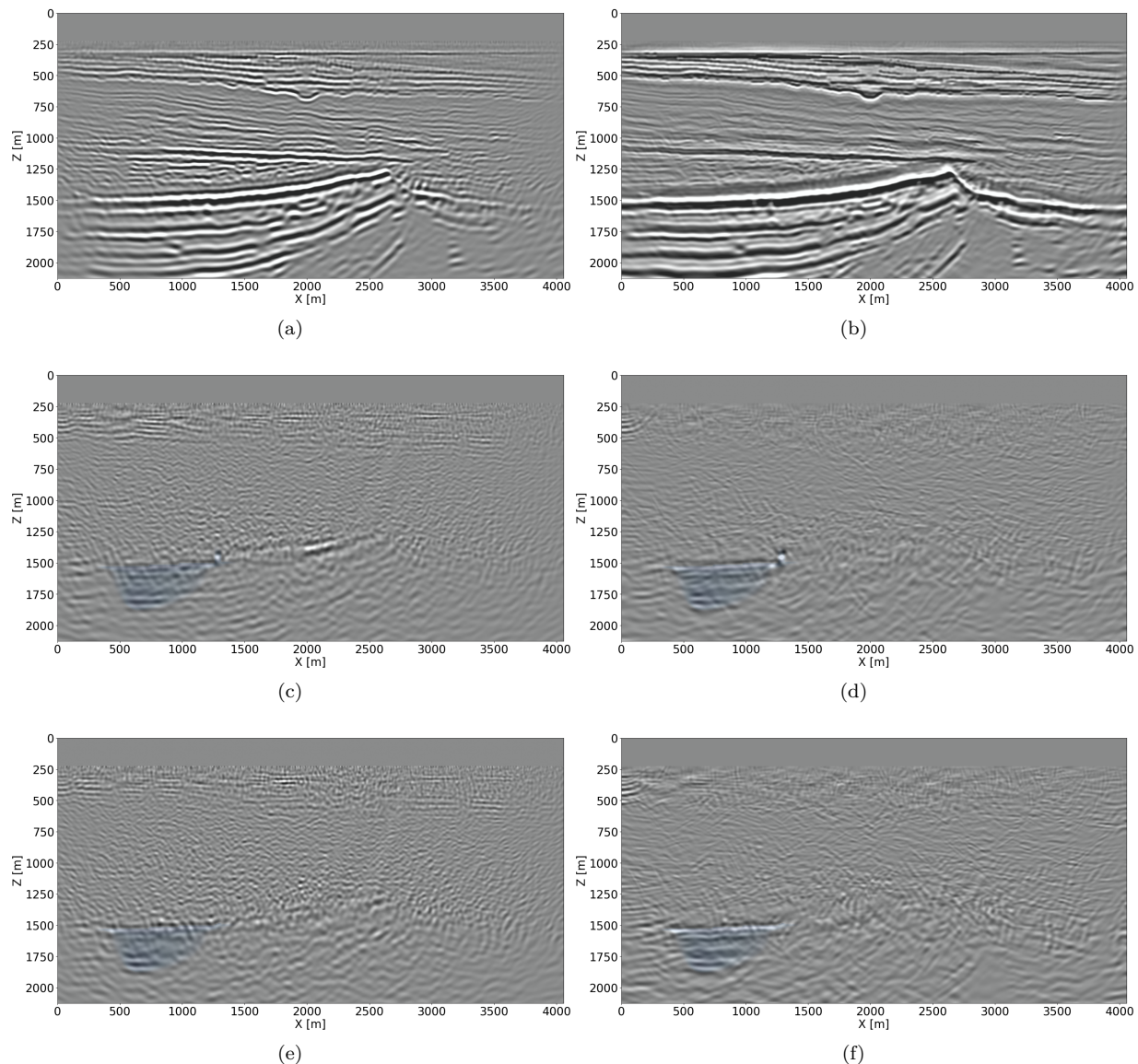


Figure 3: Reverse-time migration (RTM) versus inversion joint recovery model (JRM). (a) RTM image of the baseline; (b) JRM image of the baseline; (c) time-lapse difference and CO₂ plume for independent RTM images with leakage; (d) time-lapse obtained by inverting the time-lapse data jointly with leakage; (e) time-lapse difference and CO₂ plume for independent RTM images without leakage; (f) time-lapse obtained by inverting the time-lapse data jointly without leakage. Notice improvement in the time-lapse image quality. This improvement is reflected in the NRMS values that decrease from 8.48 % for RTM to 3.20 % for JRM.

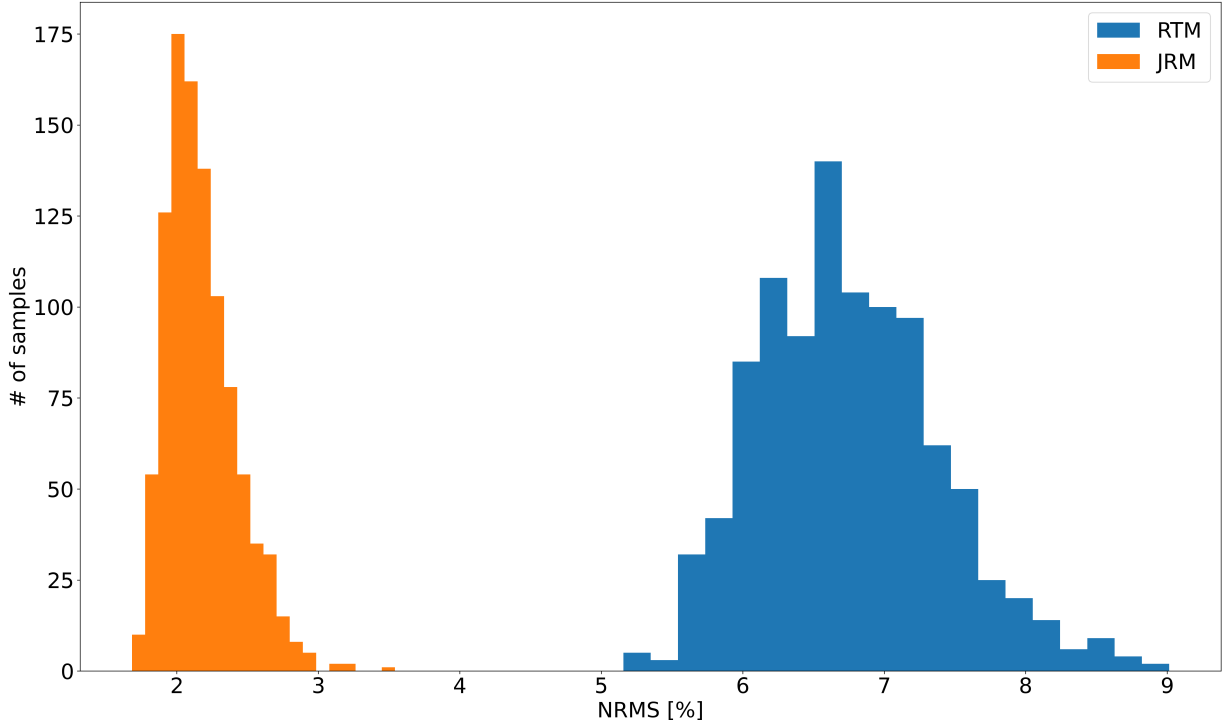


Figure 4: NRMS values for 1000 time-lapse experiments.

Deep neural network classifier for CO₂ leakage detection

The injection of supercritical CO₂ into the storage complex perturbs the physical, chemical and thermal environment of the reservoir [Newell and Ilgen, 2019]. Because CO₂ injection increases the pressure, this process may trigger CO₂ leakage across the seal when the pressure increase induces opening of pre-existing faults or fractures zones [Pruess, 2006, Ringrose, 2020]. To ensure safe operations of CO₂ storage, we develop a quantitative leakage detection tool based on a deep neural classifier. This classifier is trained on time-lapse images that contain the imprint of CO₂ plumes that behave regularly and irregularly. In case of irregular flow, CO₂ escapes the storage complex through a pressure induced opening in the seal, which causes a localized increase in permeability (shown in Figure 3d).

Because time-lapse differences are small in amplitude, and strongly localized laterally when leakage occurs, highly sensitive learned classifiers are needed. For this purpose, we follow Erdinc et al. [2022] and adopt the Vision Transformer (ViT) [Dosovitskiy et al., 2020]. This state-of-the-art classifier originated from the field of natural language processing (NLP) [Vaswani et al., 2017]. Thanks to their attention mechanism, ViTs have been shown to achieve superior performance on image classification tasks where image patches are considered as word tokens by the transformer network. As a result, ViTs have much less image-specific inductive bias compared to convolutional neural networks [Dosovitskiy et al., 2020].

To arrive at a practical and performant ViT classifier, we start from a ViT that is pre-trained on image tasks with 16×16 patches and apply transfer learning [Yosinski et al., 2014] to fine-tune this network on 1576 labeled time-lapse images. Catastrophic forgetting is avoided by freezing the initial layers, which are responsible for feature extraction, during the initial training. After the initial training of the last dense layers, all network weights are updated for several epochs while keeping the learning rate small. The labeled (regular vs. irregular flow) training set itself consists of 1576 time-lapse datasets divided equally between regular and

irregular flow.

After the training is completed, baseline and monitor surveys are simulated for 394 unseen Earth models with regular and irregular plumes. These simulated time-lapse datasets are imaged with JRM by inverting the matrix in Equation 3 via Bregman iterations in Equation 5. The resulting time-lapse difference images (see Figures 3d and 3f for two examples) serve as input to the ViT classifier. Refer to Figure 5 for performance, which corresponds to a two by two confusion matrix. The first row denotes the classification results for samples with regular plume (negative samples), where 193 (true negative) out of 206 samples are classified correctly. The second row denotes the classification results for samples with CO₂ leakage over the seal (positive samples), where 147 (true positive) out of 188 samples are classified correctly. Due to the fact that JRM recovers relatively artifact-free time-lapse differences, the classifier does not pick up too many artifacts related to finite acquisition as CO₂ leakage. This leads to much fewer false alarms for CO₂ leakage.

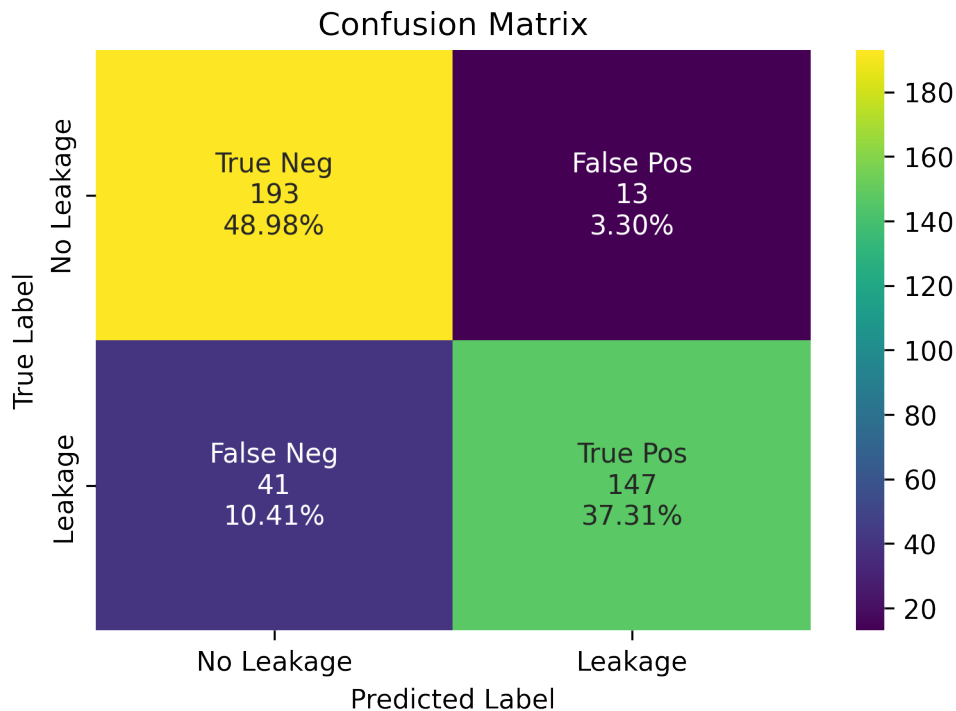


Figure 5: Confusion matrix for classifier trained on recovery images from JRM.

Class activation mapping based saliency map

While our ViT classifier is capable of achieving good performance (see Figure 5), making intervention decisions during GCS projects calls for interpretability and trustworthiness of our classifier [Hooker et al., 2019, Zhang et al., 2021, Mackowiak et al., 2021]. To enhance these features, we take advantage of class activation mappings (CAM) [Zhou et al., 2016]. These saliency maps help us to identify the discriminative spatial regions in each image that support a particular class decision. In our application, these regions correspond to areas where the classifier deems the CO₂ plume to behave irregularly (if the classification result is leakage). By overlaying time-lapse difference images with these maps, interpretation is facilitated, assisting practitioners to make decisions on how to proceed with GCS projects and take associated actions. Figure 6 illustrates how the Score

CAM approach [Wang et al., 2020] serves this purpose³. Figure 6a shows the CAM result for a time-lapse difference image classified as a CO₂ leakage (in Figure 3d). Despite few artifacts around the image, the CAM clearly focuses on the CO₂ leakage over the seal, which could potentially alert the practitioners of GCS. When the plume is detected as growing regularly, the CAM result is diffusive (shown in Figure 6b). This shows that the classification decision is based on the entire image and not only at the plume area. The scripts to reproduce the experiments are available on the SLIM GitHub page <https://github.com/slimgroup/GCS-CAM>.

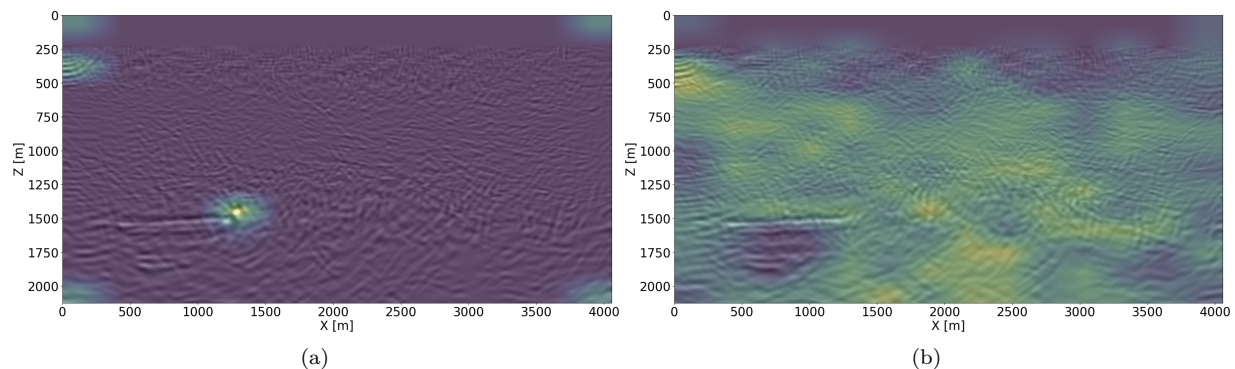


Figure 6: CAM for time-lapse difference images with a leaking plume and with a regular plume.

Conclusions & discussion

By means of carefully designed time-lapse seismic experiments, we have shown that highly repeatable, high resolution and high fidelity images are achievable without insisting on replication of the baseline and monitor surveys. Because our method relies on a joint inversion methodology, it also averts labor-intensive 4D processing. Aside from establishing our claim of relaxing the need for replication empirically, through hundreds of time-lapse experiments yielding significant improvements in NMRS values, we also showed that a deep neural classifier can be trained to detect CO₂ leakage automatically. While the classification results are encouraging, there are still false negatives. We argue that this may be acceptable since decisions to stop injection of CO₂ are also based on other sources of information such as pressure drop at the wellhead. In future work, we plan to extend this methodology to different leakage scenarios and quantification of uncertainty.

Acknowledgement

We would like to thank Charles Jones and Philipp A. Witte for the constructive discussion. The CCS project information is taken from the Strategic UK CCS Storage Appraisal Project, funded by DECC, commissioned by the ETI and delivered by Pale Blue Dot Energy, Axis Well Technology and Costain. The information contains copyright information licensed under ETI Open Licence. This research was carried out with the support of Georgia Research Alliance and partners of the ML4Seismic Center.

³We used the open-source software [PyTorch library for CAM methods](#) [Gildenblat and contributors, 2021] to calculate the CAM images.

References

- Per Avseth, Tapan Mukerji, and Gary Mavko. *Quantitative seismic interpretation: Applying rock physics tools to reduce interpretation risk*. Cambridge university press, 2010.
- Antonio Costa. Permeability-porosity relationship: A reexamination of the kozeny-carman equation based on a fractal pore-space geometry assumption. *Geophysical research letters*, 33(2), 2006.
- Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.
- Huseyin Tuna Erdinc, Abhinav P. Gahlot, Ziyi Yin, Mathias Louboutin, and Felix J. Herrmann. De-risking carbon capture and sequestration with explainable co2 leakage detection in time-lapse seismic monitoring images. In *AAAI 2022 Fall Symposium: The Role of AI in Responding to Climate Challenges*, 08 2022. URL <https://slim.gatech.edu/Publications/Public/Conferences/AAAI/2022/erdinc2022AAAIIdcc/erdinc2022AAAIIdcc.pdf>. Accepted in AAAI 2022 Fall Symposium: The Role of AI in Responding to Climate Challenges.
- Jacob Gildenblat and contributors. Pytorch library for cam methods. <https://github.com/jacobgil/pytorch-grad-cam>, 2021.
- Felix J. Herrmann and Gilles Hennenfent. Non-parametric seismic data recovery with curvelet frames. *Geophysical Journal International*, 173:233–248, 04 2008. doi: 10.1111/j.1365-246X.2007.03698.x. URL <https://slim.gatech.edu/Publications/Public/Journals/GeophysicalJournalInternational/2008/herrmann2008GJInps.pdf>.
- Sara Hooker, Dumitru Erhan, Pieter-Jan Kindermans, and Been Kim. A benchmark for interpretability methods in deep neural networks. *Advances in neural information processing systems*, 32, 2019.
- CE Jones, JA Edgar, JI Selvage, and H Crook. Building complex synthetic models to evaluate acquisition geometries and velocity inversion technologies. In *74th EAGE Conference and Exhibition incorporating EUROPEC 2012*, pages cp–293. European Association of Geoscientists & Engineers, 2012.
- Theodoros Klimentos. The effects of porosity-permeability-clay content on the velocity of compressional waves. *Geophysics*, 56(12):1930–1939, 1991.
- Clea Kolster, Simeon Agada, Niall Mac Dowell, and Samuel Krevor. The impact of time-varying co2 injection rate on large scale storage in the uk bunter sandstone. *International Journal of Greenhouse Gas Control*, 68:77–85, 2018.
- Maria Kotsi. *Time-lapse seismic imaging and uncertainty quantification*. PhD thesis, Memorial University of Newfoundland, 2020.
- ED Kragh and Phil Christie. Seismic repeatability, normalized rms, and predictability. *The Leading Edge*, 21(7):640–647, 2002.
- Dongzhuo Li and Kailai Xu. lidongzh/fwiflow.jl: v0.3.1, September 2021. URL <https://doi.org/10.5281/zenodo.5528428>.
- Dongzhuo Li, Kailai Xu, Jerry M Harris, and Eric Darve. Coupled time-lapse full-waveform inversion for subsurface flow problems using intrusive automatic differentiation. *Water Resources Research*, 56(8): e2019WR027032, 2020.
- Xiaowei Li. *A weighted ℓ_1 -minimization for distributed compressive sensing*. PhD thesis, University of British Columbia, 2015.

- Mathias Louboutin, Fabio Luporini, Michael Lange, Navjot Kukreja, Philipp A. Witte, Felix J. Herrmann, Paulius Velesko, and Gerard J. Gorman. Devito (v3.1.0): an embedded domain-specific language for finite differences and geophysical exploration. *Geoscientific Model Development*, 2019. doi: 10.5194/gmd-12-1165-2019. URL <https://slim.gatech.edu/Publications/Public/Journals/GMD/2019/louboutin2018dae/louboutin2018dae.pdf>. (Geoscientific Model Development).
- Mathias Louboutin, Philipp Witte, Ziyi Yin, Henryk Modzelewski, Carlos da Costa, Kerim, and Peterson Nogueira. slimgroup/judi.jl: v3.1.9, September 2022. URL <https://doi.org/10.5281/zenodo.7086719>.
- Fabio Luporini, Mathias Louboutin, Michael Lange, Navjot Kukreja, Philipp Witte, Jan Hückelheim, Charles Yount, Paul HJ Kelly, Felix J Herrmann, and Gerard J Gorman. Architecture and performance of devito, a system for automated stencil computation. *ACM Transactions on Mathematical Software (TOMS)*, 46(1):1–28, 2020.
- Fabio Luporini, Mathias Louboutin, Michael Lange, Navjot Kukreja, rhodrin, George Bisbas, Vincenzo Pandolfo, Lucas Cavalcante, Tim Burgess, Gerard Gorman, and Ken Hester. devitocodes/devito: v4.7.1, August 2022. URL <https://doi.org/10.5281/zenodo.6958070>.
- Radek Mackowiak, Lynton Ardizzone, Ullrich Kothe, and Carsten Rother. Generative classifiers as a basis for trustworthy image classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2971–2981, 2021.
- Pania Newell and Anastasia G Ilgen. Overview of geological carbon storage (gcs). In *Science of Carbon Storage in Deep Saline Formations*, pages 1–13. Elsevier, 2019.
- Felix Oghenekohwo. *Economic time-lapse seismic acquisition and imaging—Reaping the benefits of randomized sampling with distributed compressive sensing*. PhD thesis, The University of British Columbia, Vancouver, 08 2017. URL <https://slim.gatech.edu/Publications/Public/Thesis/2017/oghenekohwo2017THetl/oghenekohwo2017THetl.pdf>. (PhD).
- Felix Oghenekohwo and Felix J. Herrmann. Improved time-lapse data repeatability with randomized sampling and distributed compressive sensing. In *EAGE Annual Conference Proceedings*, 06 2017a. doi: 10.3997/2214-4609.201701389. URL <https://slim.gatech.edu/Publications/Public/Conferences/EAGE/2017/oghenekohwo2017EAGEitl/oghenekohwo2017EAGEitl.html>. (EAGE, Paris).
- Felix Oghenekohwo and Felix J. Herrmann. Highly repeatable time-lapse seismic with distributed Compressive Sensing—mitigating effects of calibration errors. *The Leading Edge*, 36(8):688–694, 08 2017b. doi: 10.1190/tle36080688.1. URL <https://slim.gatech.edu/Publications/Public/Journals/TheLeadingEdge/2017/oghenekohwo2017hrt/oghenekohwo2017hrt.html>. (The Leading Edge).
- Karsten Pruess. On co2 behavior in the subsurface, following leakage from ageologic storage reservoir. Technical report, Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2006.
- Philip Ringrose. *How to store CO2 underground: Insights from early-mover CCS Projects*. Springer, 2020.
- Yue Tian, Lei Wei, Chang Li, Shauna Oppert, and Gilles Hennenfent. Joint sparsity recovery for noise attenuation. In *SEG Technical Program Expanded Abstracts 2018*, pages 4186–4190. Society of Exploration Geophysicists, 2018.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- Haofan Wang, Zifan Wang, Mengnan Du, Fan Yang, Zijian Zhang, Sirui Ding, Piotr Mardziel, and Xia Hu. Score-cam: Score-weighted visual explanations for convolutional neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pages 24–25, 2020.

- Lei Wei, Yue Tian, Chang Li, Shauna Oppert, and Gilles Hennenfent. Improve 4d seismic interpretability with joint sparsity recovery. In *SEG Technical Program Expanded Abstracts 2018*, pages 5338–5342. Society of Exploration Geophysicists, 2018.
- Philipp A. Witte, Mathias Louboutin, Navjot Kukreja, Fabio Luporini, Michael Lange, Gerard J. Gorman, and Felix J. Herrmann. A large-scale framework for symbolic implementations of seismic inversion algorithms in julia. *Geophysics*, 84(3):F57–F71, 03 2019a. doi: 10.1190/geo2018-0174.1. URL <https://slim.gatech.edu/Publications/Public/Journals/Geophysics/2019/witte2018alf/witte2018alf.pdf>. (Geophysics).
- Philipp A. Witte, Mathias Louboutin, Fabio Luporini, Gerard J. Gorman, and Felix J. Herrmann. Compressive least-squares migration with on-the-fly fourier transforms. *Geophysics*, 84(5):R655–R672, 08 2019b. doi: 10.1190/geo2018-0490.1. URL <https://slim.gatech.edu/Publications/Public/Journals/Geophysics/2019/witte2018cls/witte2018cls.pdf>. (Geophysics).
- Mengmeng Yang, Zhilong Fang, Philipp A. Witte, and Felix J. Herrmann. Time-domain sparsity promoting least-squares reverse time migration with source estimation. *Geophysical Prospecting*, 68(9):2697–2711, 08 2020. doi: 10.1111/1365-2478.13021. URL <https://slim.gatech.edu/Publications/Public/Journals/GeophysicalProspecting/2020/yang2020tdsp/yang2020tdsp.html>. (Geophysical Prospecting).
- Wotao Yin, Stanley Osher, Donald Goldfarb, and Jerome Darbon. Bregman iterative algorithms for l_1 -minimization with applications to compressed sensing. *SIAM Journal on Imaging sciences*, 1(1):143–168, 2008.
- Ziyi Yin, Mathias Louboutin, and Felix J. Herrmann. Compressive time-lapse seismic monitoring of carbon storage and sequestration with the joint recovery model. In *SEG Technical Program Expanded Abstracts*, pages 3434–3438, 09 2021. doi: 10.1190/segam2021-3569087.1. URL <https://slim.gatech.edu/Publications/Public/Conferences/SEG/2021/yin2021SEGcts/yin2021SEGcts.html>. (IMAGE, Denver).
- Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. How transferable are features in deep neural networks? *Advances in neural information processing systems*, 27, 2014.
- Yu Zhang, Peter Tiño, Aleš Leonardis, and Ke Tang. A survey on neural network interpretability. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2021.
- Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. Learning deep features for discriminative localization. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2921–2929, 2016.
- Wei Zhou and David Lumley. Non-repeatability effects on time-lapse 4d seismic full waveform inversion for ocean-bottom node data. *Geophysics*, 86(4):1–60, 2021.