A deep learning model for regional-scale 3D subsurface structure mapping

Zhenjiao Jiang [1,2], Dirk Mallants [2], Lei Gao [2], Gregoire Mariethoz [3], Luk Peeters [4]

1 Key Laboratory of Groundwater Resources and Environment, Ministry of Education, College of Environment and Resources, Jilin University, Changchun, 130021, China
2 CSIRO Land & Water, Locked Bag 2, Glen Osmond, SA 5064, Australia
3 University of Lausanne, Faculty of Geosciences and Environment, Institute of Earth Surface Dynamics, Lausanne, Switzerland
4 CSIRO Mineral Resources, Locked Bag 2, Glen Osmond, SA 5064, Australia

Corresponding author: Zhenjiao Jiang (jiangzhenjiao@hotmail.com)

Abstract. This study introduces an efficient deep learning approach based on convolutional neural networks with joint autoencoder and adversarial structures for 3D subsurface mapping from surface observations. The method was applied to delineate palaeovalleys in an Australian desert landscape. The neural network was trained on a 6,400 km^2 domain by using a land surface tomography as 2D input and an airborne electromagnetic (AEM)-derived probability map of palaeovalley presence as 3D output. The trained neural network has a maximum square error < 0.10 and produces a square error < 0.10 across 93% of the validation areas, demonstrating that it is reliable in reconstructing 3D palaeovalley patterns beyond the training area. Due to its generic structure, the neural network structure designed in this study and the training algorithm have broad application potential to construct 3D geological features (ore bodies, aquifer) from 2D land surface observations.

1 Introduction

Imaging the Earth's subsurface is crucial for the exploration and management of mineral, energy and groundwater resources; its reliability depends on the availability and quality of geological data. Although the amount and quality of geological data obtained from borehole logs, geophysical prospecting and remote sensing has increased ~~dramatically~~ over the past decades, their spatial distribution is highly uneven. Most data exist ~~as land surface observations~~ in a limited number of highly-developed areas such as mining and oil fields. ~~P~~redictive models are much-needed for ~~extracting hidden information from~~ local-scale rich ~~big~~ datasets ~~for predicting 3D regional-scale subsurface structures in data-poor areas~~.

Commonly used methods for modelling complex geological structures include geostatistical approaches such as sequential Gaussian or indicator simulation (Lee et al., 2007), transition probability simulation (Felletti et al., 2006; Weissmann and Fogg, 1999), or multiple-point simulation (MPS) methods (Hu and Chugunova, 2008; Mariethoz and Caers, 2014; Strebelle, 2002). However, they often present drawbacks such as ~~not being able to represent elongated and connected geological structures such as palaeovalleys,~~ being inefficient in capturing essential features and patterns from very large training datasets, or presenting a high computational cost. A fast and reliable tool for high-resolution 3D subsurface imaging based on multiple-support big dataset ~~is still lacking~~. Deep learning approaches specialised in big data mining have the potential to ... this gap (Gu et al., 2017; Hinton and Salakhutdinov, 2006; Marcais and de Dreuzy, 2017). Applications in the geosciences include earthquake detection based on seismic monitoring (Mousavi and Beroza, 2019; Perol et al., 2018), or disaster recognition from remote sensing data (Amit et al., 2016; Längkvist et al., 2016), among others. A recent breakthrough in deep learning is the 2D to 3D image processing (Niu et al., 2018; Sinha et al., 2017; Wu et al.,

Summary of Comments on gmd-2020-106-manuscript-version1_ed.pdf

Page: 1

Number: 1 Author: Subject: Cross Out Date: 12/6/20 8:00:43 PM

Author: Subject: Sticky Note Date: 12/6/20 8:00:57 PM
??

Number: 2 Author: Subject: Cross Out Date: 12/6/20 7:59:57 PM

Number: 3 Author: Subject: Cross Out Date: 12/6/20 8:02:43 PM

Author: Subject: Sticky Note Date: 12/6/20 8:03:25 PM
"data poor" contradicts "rich/big".

Number: 4 Author: Subject: Cross Out Date: 12/6/20 8:01:53 PM

Number: 5 Author: Subject: Cross Out Date: 12/6/20 8:03:32 PM

Number: 6 Author: Subject: nserted Text Date: 12/6/20 8:01:49 PM
Improved

Number: 7 Author: Subject: nserted Text Date: 12/6/20 8:02:22 PM
analyzing

Number: 8 Author: Subject: Cross Out Date: 12/6/20 8:05:46 PM

Author: Subject: Sticky Note Date: 12/6/20 8:07:06 PM
This is most certainly not true of those methods, although one might need to use them more expertly (e.g., through zoning of the model region) when non-stationarities are present.

Number: 9 Author: Subject: nserted Text Date: 12/6/20 8:10:30 PM
would be beneficial (To authors: this is weak as a motivational statement, and distorts the truth about capabilities of existing methods to justify your method. It is sufficient for your method to produce improvements on the other methods, and you can rest your case on that premise rather than asserting the other methods not nearly as functional.)

Number: 10 Author: Subject: Highlight Date: 12/6/20 8:11:50 PM
You have not identified a gap, but rather a potential way of improving upon other methods. Rewrite to better describe the "gap" and what your method potentially does.

2016; Yi et al., 2017). Such approaches point out a novel way to rapidly and automatically identify covered 3D subsurface structures directly from readily-available 2D surface observations (digital elevation models, land cover maps, signals captured by airborne geophysical surveys).

To this end, we designed a deep convolutional neural network (CNN) with joint autoencoder (Kingma and Welling, 2013) and adversarial structures (Goodfellow et al., 2014). The autoencoder component features large input and output images connected by a small latent space. This structure is advantageous for the fusion of complex input data and 3D image reconstruction. Its training involves direct back-propagation according to a voxel-wise independent heuristic criterion, and thus often needs a large training dataset to constrain the model and avoid overfitting (Laloy et al., 2018). The generative adversarial learning tries to generate multiple images inheriting the probability structure of one real image, which relaxes the need for very large training dataset. The proposed approach successfully generates regional-scale 3D palaeovalley patterns from 2D digital terrain information in an Australian desert landscape, demonstrating that the interplay between autoencoder and adversarial components provides a generic tool to exploit more effectively geophysical, land surface and other data to generate realistic regional-scale 3D geological structures.

2 Method

The adversarial neural network for 3D subsurface imaging involves three steps: (1) patch extraction and representation, (2) nonlinear mapping and reconstruction, and (3) statistical expression of the generated image (Fig. 1). The first step is referred to as 'encoder' (Fig. 1a), which is employed to fuse the information contained in the 2D land surface observation images (input data) into a low-dimension layer by successive convolutions (Hinton and Salakhutdinov, 2006):

$$h(\mathbf{x}) = f(\mathbf{W} \cdot \mathbf{x} + \boldsymbol{b}), \quad (1)$$

where f is a nonlinear function referred to as "activation function", $\mathbf{W}$ is a matrix of weights and $\boldsymbol{b}$ is a bias vector in the encoder.

The encoder can be designed to contain multiple layers, where the number of layers is defined as 'depth'. Each layer can contain multiple images, with the number of images defined as 'width'. The images in one layer are convoluted to generate the elements in the image of the next layer by weight filters, and the elements in the low-dimension layer of the encoder (the output) are called 'code'. The process of convolution is illustrated in Fig. 1b, which shows that with a filter size of 2×2 (for a 2D image convolution for example), one element in the output layer is related to 4 elements in the input layer. Thus, the spatial correlation scale addressed by the convolutional neural network can be controlled by the filter size in both vertical and horizontal directions.

The weight and bias in the encoder are trained to ensure that the code follows a standard normal distribution, by minimizing the Kullback–Leibler divergence ($L1$), defined as (Kullback and Leibler, 1951):

$$L1 = \frac{1}{2N}\sum_{i=1}^{N}(\mu^2 + \sigma^2 - \log\sigma^2 - 1)_i, \quad (2)$$

where N is number of codes in the final output layer of the encoder, μ and σ are the mean and standard deviation of the codes, respectively.

Number: 1 Author: Subject: nserted Text Date: 12/6/20 8:12:40 PM

e.g.,

$$L3 = -\frac{1}{V}\log[D(G(\mathbf{z}))], \quad (4)$$

and

$$L4 = -\frac{1}{V}\log[D(\mathbf{Y})] - \frac{1}{V}\log[1 - D(G(\mathbf{z}))], \quad (5)$$

where V is the size of the output vector via the discriminator, and $D(\cdot)$ represents the calculations (Eq. 1) in the discriminator. The weights in the discriminator are trained to minimize $L4$, which attempts to distinguish the vectors generated from the real and simulated 3D images. The weights in the generator are trained to minimize $L3$, which attempts to fool the discriminator to be unable to identify the vector generated from the simulated 3D image. In such a way, the generator can produce images aligned with the real image in terms of probability structure (Goodfellow et al., 2014).

Finally, while the loss function $L4$ is minimized to optimize the weights in the discriminator, a comprehensive loss function combining $L1$, $L2$ and $L3$ is employed to optimize the weights in the generator, which is expressed as (Wu et al., 2016):

$$L_g = \mathrm{a} \cdot L1 + \mathrm{b} \cdot L2 + \mathrm{c} \cdot L3, \quad (6)$$

where a, b, c are the coefficients on each loss function. This loss function makes it convenient to vary the neural network structure between semi-supervised learning with additional adversarial neural network by defining coefficient c as non-zero value, and supervised learning with merely autoencoder neural network with c as zero. The hyperparameters (including the width, depth, filter size and the coefficients in generator loss functions, etc.) defining the neural network structure, are determined by trial-and-error tests (Supplementary materials). Weight and bias in generator and discriminator are trained to minimize L_g and $L4$ using the stochastic gradient descent algorithm, referred to as adaptive moment estimation (ADAM) (Kingma and Ba, 2014). We implemented the above convolution neural network using the Tensorflow Python library (Abadi et al., 2016). Once the neural network is trained, the 'generator' in the network (Fig. 1a) is used independently to generate 3D subsurface structures from the 2D land surface observations.

3 Results

We use a CSIRO dataset to test the effectiveness of our deep-learning approach in predicting 3D palaeovalley patterns in the Anangu Pitjantjatjara Yankunytjatjara (APY) lands of South Australia (Fig. 2a and 2b). The dataset includes a 100-m, 2D Multiple-resolution Valley Bottom Flatness (MrVBF) index (calculated from the digital elevation model) across the entire model domain (Gallant and Dowling, 2003), and a 3D electrical conductivity dataset (400-m horizontal and 10-m vertical resolution) interpreted from an airborne electromagnetic (AEM) survey in the APY Lands (Ley-Cooper and Munday, 2013; Soerensen et al., 2016). Previous hydrogeological characterization indicate that high bulk electrical conductivity values (EC) are a proxy for palaeovalley presence (Jiang et al., 2019; Munday et al., 2013; Taylor et al., 2015). Thus, a palaeovalley aquifer index (PAI) is defined as:

$$\mathrm{PAI} = \frac{\log_{10}(EC) - \log_{10}(EC)_{min}}{\log_{10}(EC)_{max} - \log_{10}(EC)_{min}}, \quad (7)$$

where *max* and *min* represent the maximum and minimum logarithm of EC values over the entire dataset, respectively. PAI ranges from 0.0 to 1.0 and is calculated in the first 100 m depth at the AEM-surveyed area,

Page: 4

Number: 1 Author: Subject: Highlight Date: 12/6/20 9:06:24 PM

This would appear to represent existing fluvial drainage characteristics. For this to be useful for 'training' the DL model there would have to be a mechanistic connection (geologic) between these surface features and the subsurface distribution of paleochannels. There is a big problem with this approach: modern geomorphic surface characteristics seldom represent or correlate to the morphology and distributions of subsurface facies or rock types.

Number: 2 Author: Subject: Highlight Date: 12/6/20 9:10:20 PM

Not clear. Do the valley bottoms in c correspond each to the type of channel and facies depicted in b? If yes, does that mean these are all incised into granite? In that case, the predictive geologic problem would appear to be trivial.

Number: 3 Author: Subject: Highlight Date: 12/6/20 9:08:11 PM

Is the point here to use the AEM results as a ground truth and demonstrate that you could do as good, or almost as good, without the AEM and just using your DL approach based on surficial information? Not clear.

which is considered as a ground-truth 3D probability map of palaeovalley occurrences with a spatial resolution of 400 m×400 m×10 m.

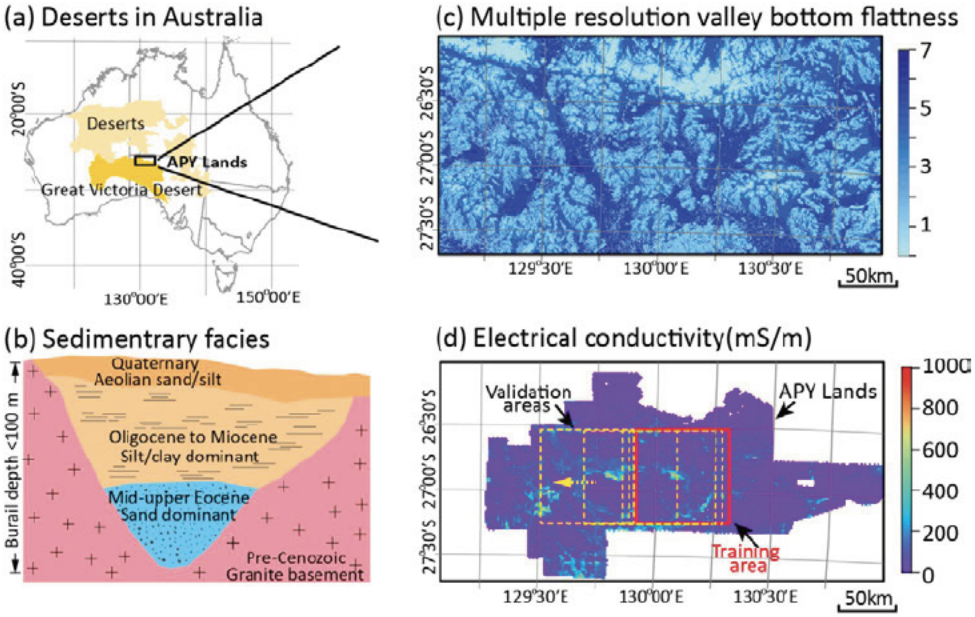

Figure 2. Datasets for delineating 3D palaeovalley in the Anangu Pitjantjatjara Yankunytjatjara (APY) Lands of South Australia: (a) location of the largest deserts in Australia and (b) general conceptual model of palaeovalley sedimentary facies revealed by over 90% borehole logs, (c) multiple resolution valley bottom flatness index, and (d) electrical conductivity (at depths of 30 to 40 m with a horizontal resolution of 400 m) inferred by airborne electromagnetic surveys in the APY Lands, forming an indicator of palaeovalley occurrence.

A neural network simulator is established and trained to relate the AEM-derived PAI (output image) with 2D MrVBF data (input image). The training dataset covers part of the APY Lands (6,400 km^2) (hereafter referred to as 'training area'). Both loss functions for discriminator and generator were monitored when training the model to verify the network being trained sufficiently (Supplementary materials). Training of the network under 10,000 iterations on a high-performance computer (Tesla P-100-SXM2-M-16GB) required 100 to 150 minutes of computation time. Once trained, generating of 3D image from 2D MrVBF required less than five seconds on a desktop computer.

An area 80 km west of the training area is first used to validate the trained neural network in generating 3D PAI. The statistics of square errors between the simulated 3D PAI and real PAI are calculated at all 200×200×10 voxels. As shown in Fig. 3, the squared error in the training dataset is below 0.1 and with a mean value of about 0.03, and the squared error of the predicted 3D PAI is well below <0.1 for 93% of the validation domain, with a mean squared error of about 0.04. The patterns of the generated palaeovalley in both horizontal and vertical directions align with those inferred from the AEM-derived PAI. This indicates that the deep-learning neural network structure developed in this work is capable of incorporating the relationships between the MrVBF and the buried palaeovalley patterns, and allowing for reliable predictions beyond the training area.

Page: 5

Number: 1 Author: Subject: Highlight Date: 12/6/20 9:11:10 PM

But you have not established why these should possibly be related.

Number: 2 Author: Subject: Highlight Date: 12/6/20 9:19:15 PM

The vertical extents appear to be extraordinarily thick, essentially extending over the total 100 m thickness of the modeled region. This does not seem realistic unless these paleo vallies are just incised valley fill deposits produced by long term hard-rock incision, followed by relatively simple aggradation within those vallies. As such, this does not appear to be a difficult predictive test for the method, and the problem you are attempting to solve (or demonstrate) is even less clear.