GC Convolution Neural Networks – If They can Identify an Oncoming Car, can They Identify Lithofacies in Core?*

Rafael Pires de Lima[1], Fnu Suriamin[1], Kurt Marfurt[1], Matthew Pranter[1], and Gerilyn Soreghan[1]

Search and Discovery Article #42312 (2018)
Posted November 5, 2018

*Adapted from the Geophysical Corner column, prepared by the authors, in AAPG Explorer, October, 2018. Editor of Geophysical Corner is Satinder Chopra (schopra@arcis.com). Managing Editor of AAPG Explorer is Brian Ervin. AAPG © 2018. DOI:10.1306/42312deLima2018

[1]University of Oklahoma, Norman, Oklahoma (kmarfurt@ou.edu)

General Statement

Advances in deep learning and artificial intelligence promise to not only drive our cars but also taste our beer. Specifically, recent advances in the architecture of deep-learning convolutional neural networks (CNN) have brought the field of image classification and computer vision to a new level. Very deep convolutional neural networks emerged in 2014 and have achieved new levels of accuracy in several artificial intelligence classification problems. Current CNN models are able to differentiate the image of a leopard from that of a scooter, but moreover can differentiate images of leopards from their biological cousins – cheetahs and snow leopards. Deep convolutional neural network architecture achieved a 3.5-percent top-5 error (how frequent the model fails to predict the correct class as one of the top 5 guesses) and 17.3-percent top-1 error in a visual recognition challenge in 2015. The current benchmark in object category classification and detection consists of hundreds of mixed-object categories and millions of images.

Although machine learning has been significantly used in geoscience fields, the application of this technique in corebased lithofacies identification, a key component to better understand oil and gas reservoirs, remains limited. Machinelearning techniques have been intensely used to aid seismic-facies classification, lithofacies classification from well logs, and even for seismicity studies. Cored wells are important as they provide the only ground-truthing for subsurface reservoirs, including data on lithofacies variations. The goals of core-based rock-type descriptions are to identify key lithofacies and facies associations, evaluate facies stacking and identify stratigraphic surfaces, interpret depositional environments, evaluate relationships between rock properties and lithofacies, and help operators identify optimal zones for designing completions. Traditional core-based lithofacies identification is costly, time consuming and subjective (e.g. different geologists describe the same core with different results).To address some of these challenges, we evaluate whether a CNN can help a specialist on an image-recognition task.

CNN results are directly related to the amount of labeled data used during training. As more and more examples are provided to the CNN, higher accuracy rates are generated, thereby developing improved rules. Imagine how a child can understand complex and difficult-to-grasp entities based on examples, such as “what is a cube?” The first example of a cube can be a six-faced die. In this example, the child may observe that each side of the die exhibits a different number of dots, the object has a color, a size, and so forth. Given this single example, the child may struggle creating a mental model of a cube. Does a cube have something to do with a particular size, or perhaps with different numbers of dots on each side? The same child then learns that a cardboard box might be termed a cube; later, that the shape of the ice in their drink is also a cube. After a sufficient number of examples are provided, the child builds a mental model of a cube even in the absence of a formal definition. Moving forward, the child approaches the world with a set of attributes in mind whenever a new object requires shape classification. Other classes of objects might have completely different characteristics (what is a tree?) or shared characteristics (what is a parallelogram?) and can be added the child's knowledge.

Just like the child who draws upon examples to learn object classification, the CNN needs examples to understand the characteristics of each "class" it tries to differentiate. Our work focuses on transferring the learning of a complex CNN trained on more than one million random images to correctly classify a lithofacies on well core images. The great advantage of "transfer learning" is that layers that have been previously trained with a significant amount of labeled data can be reused to address different objectives without any alteration. Our job then is to use an already trained model (a CNN model that has several rules for several different images) and add an additional lithofacies identification layer.

Preliminary Results

One of the most important factors contributing to the robustness of CNN models is the amount of labeled data that can be used for training. We used well core images captured through modern photographic equipment to generate the set of data to feed our CNN. The particular section used for this project consisted of approximately 50 feet of core from the Mississippian limestone and chert reservoirs in the Anadarko Shelf, Grant County, Oklahoma. The set of core images show four different lithofacies: bedded skeletal peloidal packstone-grainstone, chert breccia in greenish shale matrix, spiculitic mudstone-wackestone and splotchy packstone-grainstone. To ensure images supplied to the CNN are consistent, the first stage of the process consists of careful cropping and selection of the images to be used as input for the training. We used a sliding window technique to extract consistent squared, cropped sections from the original core image (Figure 1). This cropping process further augmented the number of samples of our initially small collection, which helps the CNN. We exclude part of the cropped pictures from the training set to be used as testing data (Figure 1) after the CNN is trained. After the cropping process the bedded skeletal peloidal packstone-grainstone, chert breccia in greenish shale matrix, spiculitic mudstone-wackestone and splotchy packstone-grainstone lithofacies had, respectively, 285, 165, 605 and 285 images that were used for training.

The validation accuracy during training achieved 96 percent. After the training process, we can use the CNN model to predict the lithofacies of a suite of images (extracted from the same core) never used in the training. The CNN classification results are shown in Figure 2 through Figure 5, where the probability of each lithofacies is shown in the accompanying table.

Future Work

Although our CNN classification test was created with very few images and their variations, the preliminary results are quite promising. Overall, the CNN model selected the correct lithofacies with a significantly higher probability when compared to the other options. Use of images of similar sizes and qualities facilitates the lithofacies identification, as sedimentary features will maintain proportion – the CNN will not mistakenly classify conglomerates as sandstone, as size is preserved. Improved training requires incorporation of a larger number of examples to enable generation of a model sufficiently robust to overcome complications such as variances in color, core quality and textures. The move toward digitization of oil and gas data will provide the big-data enrichment to enable evaluation of the CNN methodology for examples from around the world. Changes in sedimentation characteristics, texture, core size and colors will complicate the picture, but hopefully lead to deeper learning.

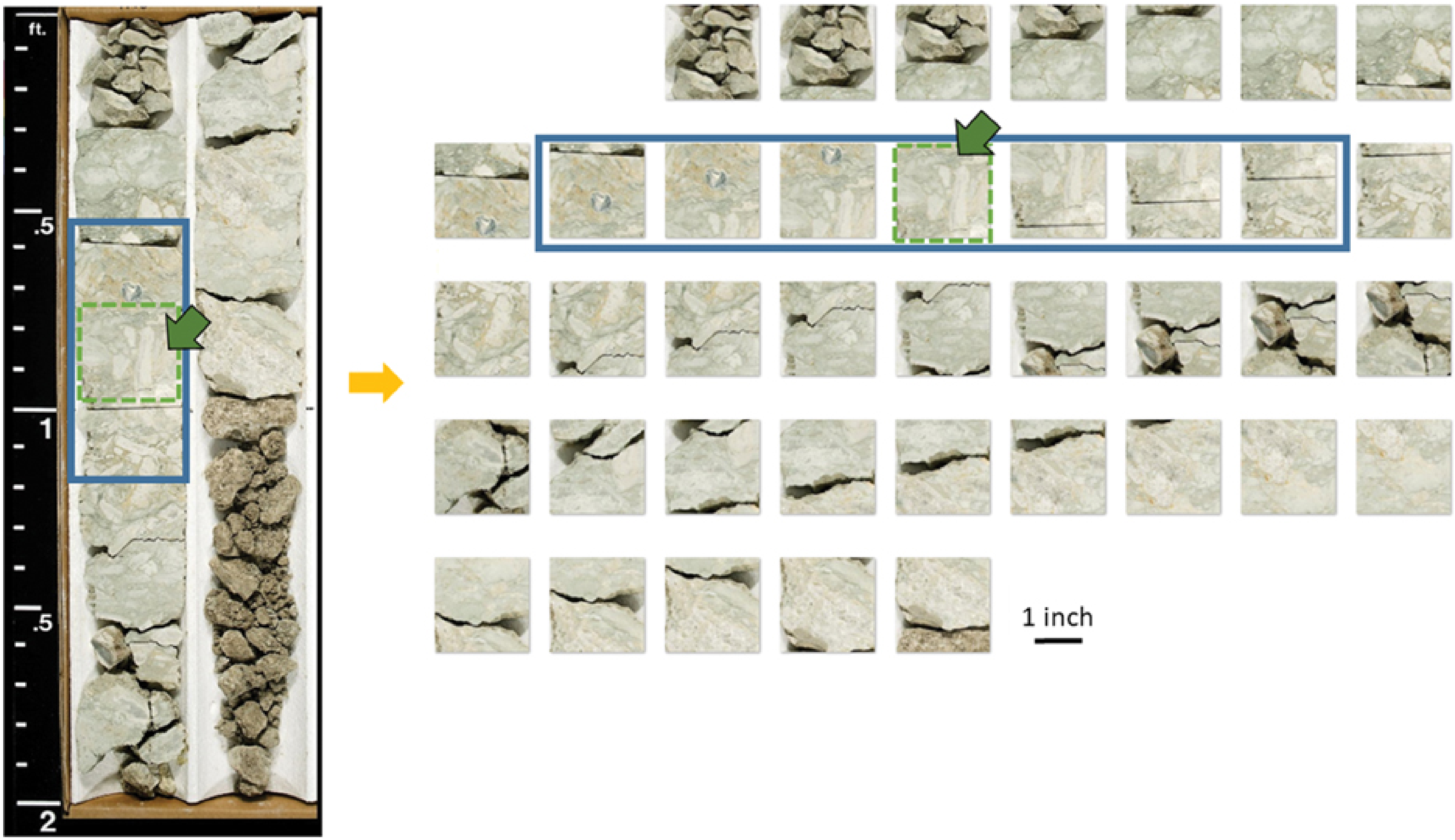


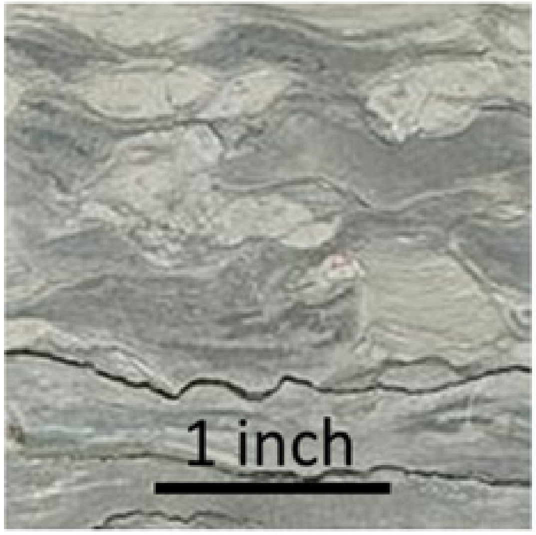
Figure 1. Figure showing how we augmented and standardized the amount of data by splitting the core in squared cropped images using a sliding window. This approach helps the CNN to access more training data and, used with care, will not force the CNN to overfit the data. The blue rectangle shows images that were never used during training (test data). The green arrow indicates the image presented in Figure 3. Separation of testing data was the same for the other three lithofacies used in this project.

Lithofacies	Probability
Bedded skeletal peloidal packstone-grainstone	**0.82**
Chert breccia in greenish shale matrix	0.15
Spiculitic mudstone-wackestone	0.02
Splotchy packstone-grainstone	0.01

Figure 2. A bedded skeletal peloidal packstone-grainstone sample image from the core not used in the CNN training.

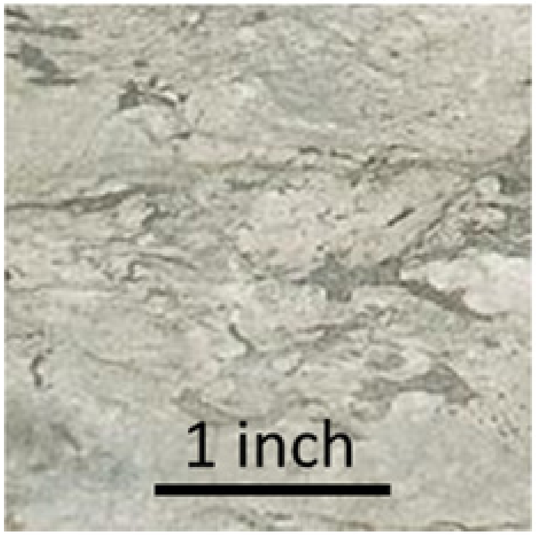
Lithofacies	Probability
Bedded skeletal peloidal packstone-grainstone	0.00
Chert breccia in greenish shale matrix	**0.96**
Spiculitic mudstone-wackestone	0.00
Splotchy packstone-grainstone	0.04

Figure 3. A chert breccia in greenish shale matrix sample image from the core not used in the CNN training.



Lithofacies	Probability
Bedded skeletal peloidal packstone-grainstone	0.01
Chert breccia in greenish shale matrix	0.02
Spiculitic mudstone-wackestone	**0.89**
Splotchy packstone-grainstone	0.08

Figure 4. A spiculitic mudstone-wackestone sample image from the core not used in the CNN training.



Lithofacies	Probability
Bedded skeletal peloidal packstone-grainstone	0.00
Chert breccia in greenish shale matrix	0.07
Spiculitic mudstone-wackestone	0.02
Splotchy packstone-grainstone	**0.91**

Figure 5. A splotchy packstone-grainstone sample image from the core not used in the CNN training.