GeoNeurale

GeoNeurale
Am Nymphenbad 8
D-81245 München
www.GeoNeurale.com

T +49 (0)89 89691118

Dynamics of latent space information in convolutional neural networks and their applications in the seismic attributes research

For the first time GeoNeurale – "Neural-Geophysics Lab" relates interneural activity to seismic attributes introducing a new chapter in the exploration seismic research

Imagine that you can measure the impulses at the interface between two neurons in the human brain and detect which is their correlation between the images you can see with your eyes or the music you can hear with your ears, how the impulses can be stored into memory and how can we subsequently pronounce the words:
"Mona Lisa of Leonardo da Vinci" or "Adagio from Symphony N. 2 of Sergei Rachmaninoff".
How can we classify the impulses that our brain detected ?
Certainly the brain mechanism is extremely complicated and mysterious for us, much more than a neural network.
If we wish to extrapolate our thinking into the future we could ask ourselves how much will we understand about human and animal brain that inspired the development of artificial intelligence in 100 years, with the same philosophical attitude that we could pose ourself the question how many more elements the Mendelejew Table will have and which other interactions will be discovered other that gravitational, electrical and magnetic field in 100 years ?

If the animal brain is so complicated however we can already extrapolate useful algorithms from what we consider its very elementary models. Convolutional neural networks or recurrent neural networks.

In deep-learning convolutional neural networks can already be used to extract a lot of informations in the interneurons data flow which translates images into categorical data.

Before diving into this discussion it is worth to review some historical examples how geophysical research lead to the necessity of more flexible models trying to avoid the classical linear algebra limitations.

GEOPHYSICISTS STRIVING FOR NON-LINEARITY

Why is artificial intelligence so important for the exploration geophysics ?

Up to the beginning of this millennium most algorithms in the processing and interpretation of seismic and petrophysical data were based on linearity.
Take for instance the PreSDM workflow.
Certainly we have the presence of a cost function but in linear multivariate regression.
In the Kirchhoff Prestack Depth Migration (PreSDM) and subsequent Tomographic Inversion of Image Gathers, a first anisotropic velocity model in depth is constructed based on Common Image Gathers volume. Imaged events are processed as diffraction points and each event is weighted and summed along the diffraction surface (Kirchhoff Summation) to converge on the Reflectivity point.
Traveltimes are computed with ray-tracing through the 3D velocity models often with the Eikonal Equation.
The goal is to find the velocity that best flat the gathers. After flatting, the weighted diffraction events are summed and placed on the reflectivity point.
In this flattening process the linear Tomographic Inversion is applied. It is in this process that a Cost function is implemented (Fig. 1).

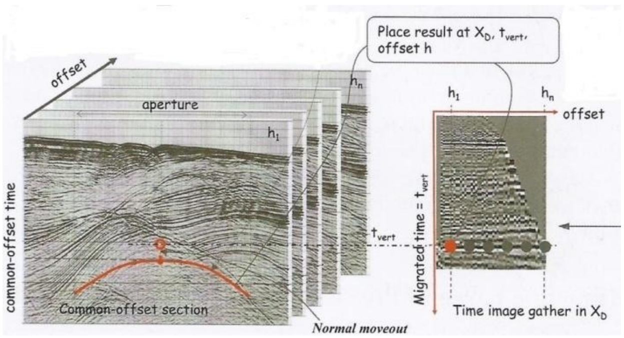

Fig. 1 . Common Image Gathers .
Courtesy: Etienne Roubain – Seismic Imaging – EAGE

The Cost function is minimized through an optimization process where the residual moveout DZ (local depth at sample offset minus depth at the zero offset) is minimized. Well ties, VSP and logs analysis can also be included into the program as further constraints.
The model is therefore a multivariate linear regression optimization process to minimize a Cost Function and flattening the gathers.

A further development after the PreSDM is the Full Wave Inversion.
In the forward problem synthetic shot gathers are generated with the goal to compare them to the actual ones from the seismic survey. This is also a linear optimization process where all elastic parameters are considered into the algorithm.

But we are still in the domain of linear models and it appears from all evolution of the research in the last 30 years that geophysicists were striving looking for non-linear algorithms to solve processing and inversion problems.

Among the first artificial intelligence applications giving optimal results were inversion applications implemented in the multiattribute analysis, where one property in the 3D seismic volume (target)

could be predicted through the normal equation or the multivariate linear regression optimization from other input features properties derived from log measurements or other seismic attributes.
Later linear regression will be substituted by non-linear optimization, and in fact, the real improvement and refinement on the properties prediction has been reached by the introduction of dense connected neural networks implementation (Fig. 2).

This optimization process was applied to the prediction and distribution of secondary target seismic attributes or rock-physics / elastic properties / Porosity and Petrophysical properties within the 3D seismic cube from logs and primary input attributes.

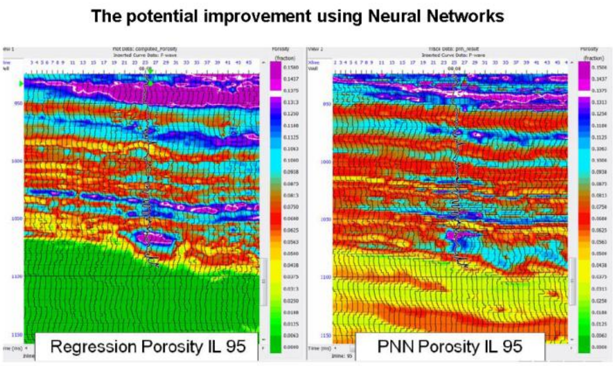

Fig. 2. Improvement in the distribution of Porosity from well logs to the 3D seismic volume. Courtesy: Hampson-Russell

The companies involved in interpretation software development continued their research and came with more sophisticated and precise algorithms for 3D seismic processing, inversion and attributes analysis.

To this point most artificial intelligence applications consider a set of input data and a form of categorical classification output.

GOAL OF THE RESEARCH

Our research lab studied the possible informations deriving from the dynamics of the internal activations in convolutional neural networks to classify 3D seismic amplitude maps in terms of structural attributes.
We can manipulate the probability distributions within the Softmax or exchange the categorical output constraining the backpropagation parameters with feature data of other attributes like images or RNN wave data.
Then the backpropagation will be consequently influenced.

Consider for instance the results of linear algorithms in attributes analysis for pattern recognition and texture mapping like: Energy, Entropy, Homogeneity (K. Marfurt, S. Chopra).
These were based on the statistical properties calculated from the co-occurrence matrix.
Similar attributes for pattern and structural recognition were extracted from the statistical analysis of interneural activations at different deeper level layers information of very deep convolutional neural networks.
These networks are also capable of producing outputs of amplitude maps with enhanced edges resolution.

PROGRAM OPTIMIZATION

For the model implementation Adam optimization was used with 500 Epochs on minibatches. Adam is a very flexible optimization algorithm which speed up the calculations and showed a great flexibility on a various range of neural networks architectures. It combines gradient descent with Momentum and RMSprop, by taking advantage of running averages on the gradients, and optimizing the directivity versus the global minimum of the cost function. It provides smoothing the “orthogonal directivity” components on gradient descent iteration steps, smoothing oscillations and concentrating it to the maximum gradient direction. This also allows to use larger learning rates.

Adam optimization is a combination of Momentum and RMSprop.
First exponentially weighted average of past gradients are computed and stored in a variable ***v*** (momentum), then bias correction is performed to get: ***Vcorrected***

Bias correction is applied to avoid a defect of the running averages algorithmus where the initial part of the function is “biased” in excess with respect to the real average values.

In a second step, the exponentially weighted average of the squares of the past gradients are computed and stored in a variables ***s*** (RMSprop).

After performing bias correction in ***s*** we can get: ***Scorrected***

Momentum →

$$v_{dW^{[l]}} = \beta_1 v_{dW^{[l]}} + (1-\beta_1)\frac{\partial \mathcal{J}}{\partial W^{[l]}}$$

$$v^{corrected}_{dW^{[l]}} = \frac{v_{dW^{[l]}}}{1-(\beta_1)^t}$$

← **Bias correction**

RMSprop →

$$s_{dW^{[l]}} = \beta_2 s_{dW^{[l]}} + (1-\beta_2)(\frac{\partial \mathcal{J}}{\partial W^{[l]}})^2$$

$$s^{corrected}_{dW^{[l]}} = \frac{s_{dW^{[l]}}}{1-(\beta_1)^t}$$

← **Bias correction**

Adam →

$$W^{[l]} = W^{[l]} - \alpha\frac{v^{corrected}_{dW^{[l]}}}{\sqrt{s^{corrected}_{dW^{[l]}}}+\varepsilon}$$

RESULTS

Statistical operations have been performed over shallow and deep layers of a very deep neural network.
Statistics give diagnostic indications on the structural parametrizations and structural similarity of two amplitude attribute maps.
But this method can be even more useful in parametrizing specific structure by comparing different attribute maps of the same or different structures. For instance an amplitude map and a filtered frequency attribute map. Their correlation can be specific for a single reservoir and can have both textural and petrophysical diagnostic significance.

Crossplotted statistical parameters show that points coming from different attributes of the same structure (In-Phase), as well as attribute maps fron similar structures cluster together in the area of minimal Entropy (red points), while the attributes from different structures (out of phase) show sparsity around the crossplot into the field of maximal entropy.

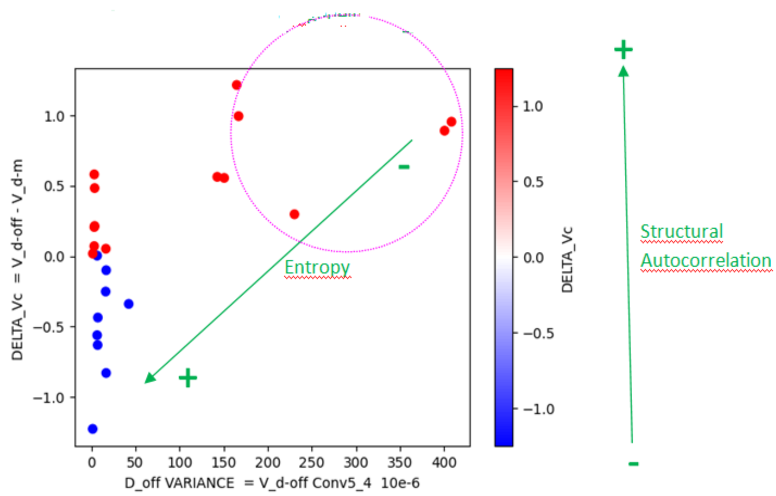

Fig. 16. circled points represent increasing structural autocorrelation

Angelo Piasentin
Scientific Direction

GeoNeurale-Wavefields
Neural-Geophysics Lab, Munich

www.GeoNeurale.com
www.Wavefields.eu