

# Seismic Scattering Wave Field Imaging Method Based on Convolution Neural Network and Equivalent Training Model

Xian XI<sup>1</sup>, Jiangqing HUANG

*China University of Geosciences (Wuhan), School of Mathematics and Physics,  
Wuhan 430074, China*

**Abstract.** The inversion of scattered waves in seismic exploration is a hot and difficult problem in related fields. We propose a hypothesis: the local wave field near each point on the wave field profile has a relationship with the minimum distance between that point and each scatterer, and this relationship can be recognized by a convolutional neural network (CNN). Based on this, the designed CNN can steadily classify and identify the scattered wavefield point by point, and realize the inversion imaging of the scatterer. We propose a new method of inversion imaging of scattered wavefield based on CNN and equivalent training model. The new method transforms the optimal inversion problem of the scatterer into the optimal design problem of the equivalent training model. Through the equivalent training model constructed by a parameterized method, the wavefield inversion and scatterer imaging of two different layered random cave media models are well realized, and the scattered wave imaging of the wavefield of the Marmousi2 model synthesis data is also realized robustly. The use of Bayesian discriminant imaging improves the resolution of the inversion results and helps interpreters quickly and accurately locate the scatterers.

**Keywords.** Equivalent training model, category of the scattering distance field, convolutional neural network, scattered wave field inversion imaging

## 1. Introduction

Studies the inversion imaging of the scattered wavefield in seismic exploration. We hypothesize that the local wavefield near each point on the wave field profile is related to the minimum distance from that point to each scatterer, and this relationship can be identified by the convolution neural network (CNN). The minimum distance is defined as the scattering distance field, and the category of the scattering distance field (ie, the size level) is used as the expected output of the CNN inversion, and the input is the local wave field near the point. Research shows that the above-mentioned CNN can classify and identify the scattered wavefield point by point, and realize the inversion imaging of underground scatterers [1-2]. A scatter wavefield inversion framework based on CNN and equivalent training model has been initially established. The problem of optimal inversion imaging of the scatterer can be transformed into an optimization design

---

<sup>1</sup> Corresponding Author, Xian XI, China University of Geosciences (Wuhan), School of Mathematics and Physics, Wuhan 430074, China; Email: 3080117816@qq.com.

problem of the equivalent training model. The CNN inversion network we designed can adapt to different types of source wavelets and different types and levels of wavefield noise; it can pre-evaluate the inversion effect based on the training model; it can also use Bayesian discrimination and filtering to improve the imaging effect. We use a relatively simple equivalent training model to realize the wavefield inversion and scatterer imaging of the layered random cave media model with large differences. At the same time, we also perform a robust inversion imaging of the scattered waves from the wavefield synthetic data of the Marmousi2 model.

## 2. Methods

### 2.1. Lenet-5 Network

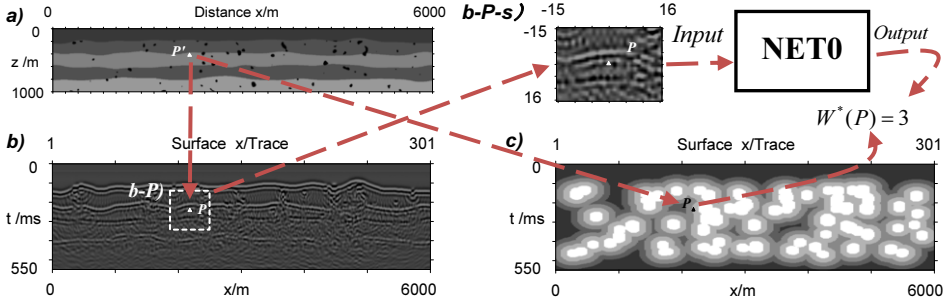
The LeNet-5 network [3] is used by us to study the inversion of scattered wave fields in some layered random cave media. The convolutional neural network is called NET0, as shown in figure 1. Its input is a  $32 \times 32$  sampling matrix (acquisition from wave field profile), and its ideal output is the  $k \times 1$  classification vector (Calculation according to medium model). Where  $k$  is the imaging level ( $k \geq 2$ ).



**Figure 1.** Input and output of the NET0 used for wave field inversion

### 2.2. Local Wave Field and Sampling Matrix

We take each grid point on the seismic wave field profile as an imaging point, and near each imaging point, extract the corresponding "local wave field matrix" from the profile. Resample the local wavefield matrix to get a "sample matrix" that is the input of NET0. The sampling radius of the "local wavefield matrix" is recorded as  $R_x, R_t$ , and the size of the sample matrix is  $32 \times 32$  after re-sampling. Suppose the scattering wave field profile with a total grid number of  $n_x (Trace) \times n_t (ms)$ . A total of  $n_x \times n_t$  sample (inputs) of NET0 can be extracted over the entire profile. For example, figure 2 is a schematic diagram of NET0 training / testing (Related concepts will be introduced in the next section: select the sampling radius  $R_x = 16 Trace, R_t = 100ms$ ; select the imaging radius  $C_x = 20 Trace, C_t = 100ms$  and select imaging level  $k = 5$ ). (b) is the wavefield profile, where (b-P in b) and (b-P-s) are respectively the local wavefield matrix and sampling matrix of an imaging point P on the profile.



**Figure 2.** Schematic diagram of NET0 training/testing. (a) is a scattering medium model of size  $6000 \times 1000$  ( $m^2$ ). (b) is the corresponding wavefield profile of size  $301(Trace) \times 550(ms)$ . (c) is the Category which contains the ideal output of  $301 \times 550$  samples of NET0. (b-P in b) is the local wavefield matrix corresponding to any imaging point P in the wavefield (b). (b-P-s) is the  $32 \times 32$  "sampling matrix" obtained by resampling the local wavefield matrix, which is the input of a training/testing sample of NET0. The value of the Category at the marked point P is  $W^*(P) = 3$

### 2.3. Definition of Scattering Distance Field

It is assumed that the scattering medium model includes a plurality of scatterers. The grid points in the scatterers are defined as "scattering points". Below we define the "scattering distance field" of the scattering medium model in the space-time domain (i.e. wave field domain).

Suppose the total number of scatterers in the model are  $m$ , and the  $j$ -th scatterer occupies  $i_j$  grid points ( $j = 1, 2, \dots, m$ ), then the scatterers in the model occupy a total grid points of  $i_1 + i_2 + \dots + i_m = n$ . These  $n$  grid points are called scattering points. Suppose  $P'_1, P'_2, \dots, P'_n$  are the center points of the  $n$  scattering points in the space-depth domain. Through time-depth conversion, the corresponding points of the  $n$  scattering points are  $P_1, P_2, \dots, P_n$ , in the space-time domain (i.e. wave field domain). Select the imaging radius  $C_x$  (x-space direction),  $C_t$  (t-time direction). The "normalized scattering distance field" at any point P in the wavefield domain is defined as follows:

$$W(P) = W(P, C_x, C_t) \stackrel{\text{def}}{=} \sqrt{\min_{1 \leq i \leq n} \left[ \left( \frac{x(P) - x(P_i)}{C_x} \right)^2 + \left( \frac{t(P) - t(P_i)}{C_t} \right)^2 \right]} \quad (1)$$

Here,  $(x(P), t(P))$  represent the coordinates of any point P on the wave field profile.

Suppose the total number of imaging points in the profile is:  $n_x(Trace) \times n_t(ms)$ . From Equation 1, the "scattering distance field" of size  $n_x \times n_t$  in the space-time domain is obtained.

### 2.4. Definition of Scattering Distance Field Category

The input of NET0 is "sample matrix", and the expected output is the scattering distance field Category, which is defined in the wave field domain as follows:

$$W^*(P) = W^*(k, W(P))$$

$$\stackrel{\text{def}}{=} \begin{cases} k-i, & \text{When } \frac{i-1}{k-1} \leq W(P) < \frac{i}{k-1}, \quad i=1,2,\dots,k-1 \\ 0, & \text{When } W(P) \geq 1 \end{cases} \quad (2)$$

Where,  $W(P)$  is the value of the "scattering distance field" at each imaging point  $P$  obtained by Equation 1;  $k$  ( $k \geq 2$ ) is the number of selected imaging level.

That is, the interval  $[0, 1]$  is equally divided into  $k-1$  small intervals. As the value of  $W(P)$  increases in the interval  $[0, 1]$ , the value of  $W^*(P)$  takes  $k-1, \dots, 2, 1$ , in order. When  $W(P) \geq 1$ , The value of  $W^*(P)$  is 0.

"Scattering distance field Category"  $W^*(P)$  is simply referred to as "Category" or "Category function". The expected output of the NET0 for each sample is a  $k$ -dimensional classification vector.  $W^*(P) = k-1$  corresponds to the imaging point of type 1-th (point  $P$  happens to belong to the Category closest to the scatterers) corresponding output is  $(1, 0, \dots, 0)^T$ , and  $W^*(P) = 0$  corresponds to the imaging point of type  $k$ -th (point  $P$  happens to belong to the Category farthest from the scatterers) corresponding output is  $(0, \dots, 0, 1)^T$ .

### 2.5. Schematic Diagram of Training/Testing of Net0

For the selected training / testing model and its wave field profile, the following system parameters need to be selected before starting the training / testing of NET0: 1. sampling radius, 2. imaging radius, 3. imaging level, and 4. the size of the sample matrix.

First, use the samples from the training model to train NET0. Then, use the trained NET0 to invert the wave field profile of the testing model. Figure 2 is a schematic diagram of NET0 training / testing (select the sampling radius  $R_x = 16 \text{ Trace}$ ,  $R_t = 100\text{ms}$ ; select the imaging radius  $C_x = 20 \text{ Trace}$ ,  $C_t = 100\text{ms}$  and select imaging level  $k=5$ ). In figure 2, (c) is the scattering distance field Category calculated by model (a) according to equations 1 and 2. The value of the Category corresponding to the white imaging point in (c) is 4, and the value of the Category corresponding to the black imaging point is 0.

### 2.6. Establish Equivalent Training Model and Testing Models

To achieve good inversion results, the training model of NET0 should be like the test model. We also hope that the "equivalent training model" has a simple structure, parameterizable representation, easy optimization, and wide adaptability. 9 "equivalent training models" is constructed as follows:

Set the background P-wave velocity to be uniform in the horizontal direction and continuously increase linearly from top to bottom in the vertical direction, from  $3000 \text{ m} \cdot \text{s}^{-1}$  to  $7000 \text{ m} \cdot \text{s}^{-1}$ , namely:

$$V_P^0(x, z) = 3000 + 4 \cdot z, \quad (0 \leq x \leq X, 0 \leq z \leq H) \quad (3)$$

Where,  $X = 6000$ ,  $H = 1000$ . Corresponding to the model of  $6000 \times 1000 \text{ m}^2$ .

Set forty-five square karst caves with randomly distributed locations and the same size of  $28 \times 28 \text{ m}^2$ ;

Set six equal-space parallel low-velocity thin-layer reflective interfaces. The location and shape of the six reflective interfaces are:

$$z = f_i(x) = \frac{i}{7} \cdot H + A \cdot \sin\left(2\pi \cdot B \cdot \frac{x}{X}\right), \quad 0 \leq x \leq X, \quad i = 1, 2, \dots, 6 \quad (4)$$

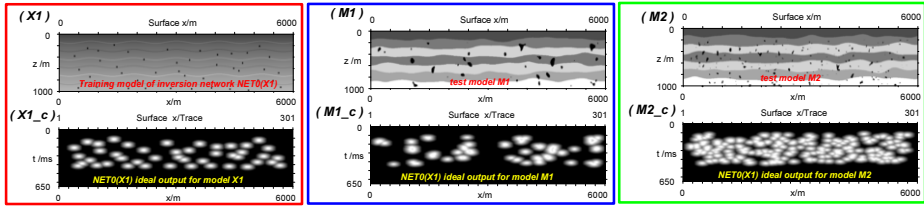
Among them, A and B are called amplitude and frequency parameters, respectively.

According to the equations 3 and 4, the P-wave velocity of the "equivalent training model" is:

$$V_p(x, z) = \begin{cases} 0.6 \times V_p^0(x, z), & \text{When } |z - f_i(x)| \leq \delta, \quad i = 1, 2, \dots, 6 \\ 1500, & \text{When } (x, z) \in \text{Cave area} \\ V_p^0(x, z), & \text{Other} \end{cases} \quad (5)$$

$\delta = 4 \text{ m}$  is the thickness of the thin layer reflective interface.

In Equation 4 and 5, the following model parameters are selected: amplitude parameter  $A = 20 \text{ (m)}$ , frequency parameter  $B = 11$ , thin layer velocity ratio  $\theta_k = \begin{cases} 1.7, & \text{When } k = 1, 3, 5 \\ 0.6, & \text{When } k = 2, 4, 6 \end{cases}$ , the equivalent training model  $XI$  is obtained from Equation 3-5, as ( $XI$ ) shown in figure 3.



**Figure 3.** Corresponding in red, blue and green Baskets: equivalent training model  $XI$ , test model  $M1$  and  $M2$ . Where: the above are the media models; the below are the ideal outputs (scattering range field category) of the inversion network  $NET0(XI)$  for the wave field profiles of the three models.

A random medium modeling method [4-7] was used to construct multi-layer random cave media models, as shown in figure 3 ( $M1-2$ ), call them 2 test models, each test model has 6 random undulating interfaces. The P-wave velocities of the 5 layers from top to bottom are:  $V_p^1 = 3000 \text{ m} \cdot \text{s}^{-1}$ ,  $V_p^2 = 4000 \text{ m} \cdot \text{s}^{-1}$ ,  $V_p^3 = 6000 \text{ m} \cdot \text{s}^{-1}$ ,  $V_p^4 = 4000 \text{ m} \cdot \text{s}^{-1}$ ,  $V_p^5 = 6000 \text{ m} \cdot \text{s}^{-1}$ ,  $V_p^6 = 5000 \text{ m} \cdot \text{s}^{-1}$ , and  $V_p^7 = 7000 \text{ m} \cdot \text{s}^{-1}$ . Because the shape and size of caves are completely random, some caves are too small to be displayed, and it is easier to identify them by combining the corresponding "Category" map. The three models  $XI$ ,  $M1$ ,  $M2$  and their corresponding categories (select the sampling radius  $R_x = 16 \text{ Trace}$ ,  $R_t = 100 \text{ ms}$ ; select the imaging radius  $C_x = 10 \text{ Trace}$ ,  $C_t = 50 \text{ ms}$  and select imaging level  $k = 10$ ) are shown in figure 3.

## 2.7. Wave Field Simulation and Inversion

Using the staggered grid finite difference method to implement the forward simulation of elastic wave [8-11]. When there are more than ten grid points in each wavelength, the simulation result is good [12]. The surrounding boundary adopts the absorption boundary condition of Cerjan [13]. Each single shot record of the elastic wave is simulated, and then obtain a zero-offset wavefield profile. Finally, the 15-degree finite difference method is used to perform reverse time migration to obtain the time migration profile of the vertical component profile of the zero-offset wave field, which is referred to as the "Migration profile of wave field" or "profile" for short. Elastic waves are simulated in the scattering media models  $XI$ ,  $M1$ ,  $M2$  shown in figure 3. We select the source spacing as  $dx = 20 \text{ m}$ , and each wave-field profile contains  $n_x = 301$  traces with a sampling interval of  $dt = 1 \text{ ms}$  and the length of each record is  $n_t = 650 \text{ ms}$ .

In order to increase the anti-noise ability of inversion, three different profile noises are selected: no noise  $P0$ , Gaussian noise with 10% standard deviation  $P1$ , and 5% salt & pepper noise  $P2$ . In order to increase the universality of inversion, the following two different source wavelet are used for wave field simulation:

(1) Rick wavelet  $w(t)$  with frequency  $f = 30\text{Hz}$ :

$$w(t) = (1 - 2 \cdot (\pi \cdot f \cdot t)^2) \cdot \exp(-(\pi \cdot f \cdot t)^2) \quad (6)$$

$3 \times 3$  Rick wavelet profiles of models  $X1$ ,  $M1$ ,  $M2$  are obtained, which are respectively recorded as:  $i\_P0$ ,  $i\_P1$ ,  $i\_P2$ , where model  $i = X1, M1, M2$ .

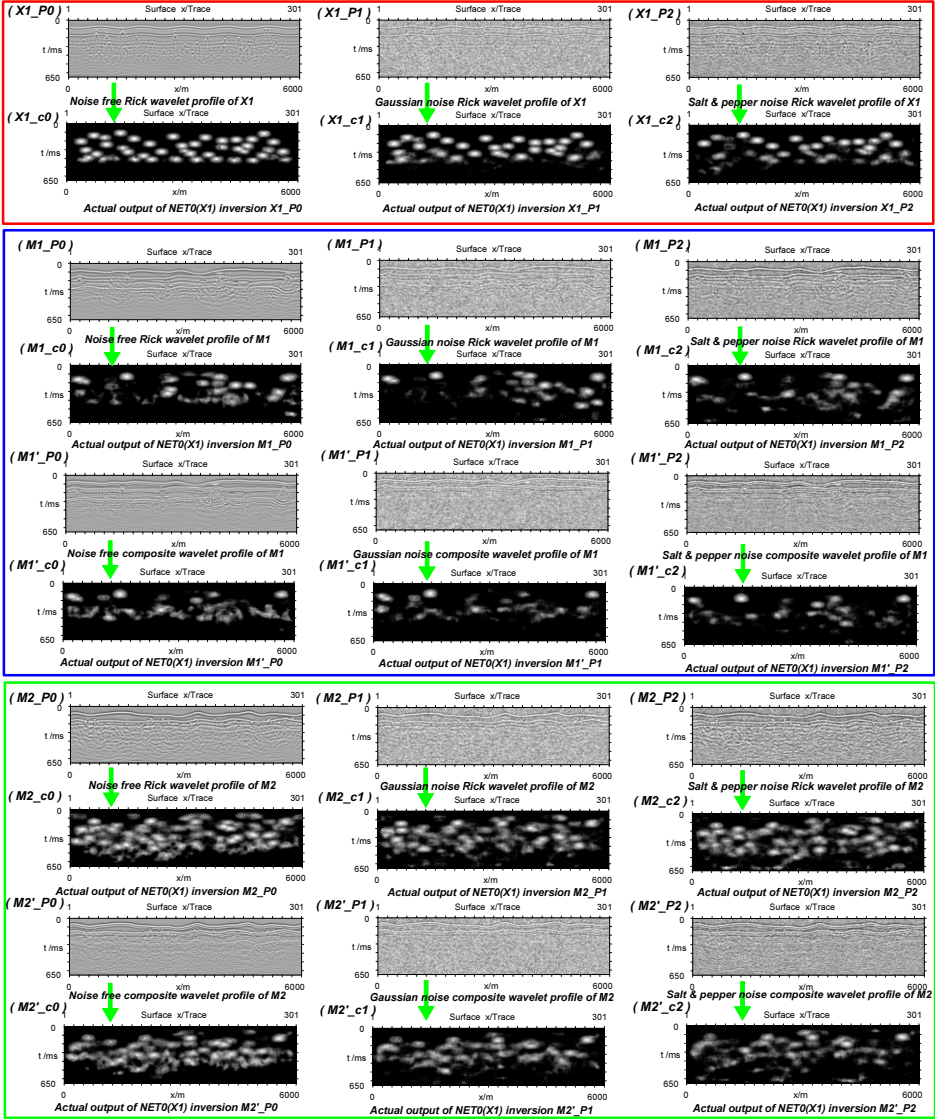
(2) Composite wavelet  $ww(t)$  obtained by combining Rick wavelet  $w(t)$ :

$$ww(t) = w(t + 0.01) + 0.5 \cdot w(t - 0.01) \quad (7)$$

Where: the time unit is seconds, and the function  $w$  is the Rick wavelet defined by Equation 6.

$2 \times 3$  composite wavelet profiles of test models  $M1$  &  $M2$  are obtained by simulation, which are respectively recorded as:  $i'_P0$ ,  $i'_P1$ ,  $i'_P2$ , where model  $i = M1, M2$ .

Models  $X1, M1, M2$  correspond to:  $1 \times 1 \times 3 + 2 \times 2 \times 3$  (model  $\times$  wavelet  $\times$  noise) = 15 wave field profiles and the inversion results of  $NET0(X1)$ , are shown in figure 4.



**Figure 4.**  $NET0(X1)$  inversion of 15 profiles. Corresponding to 2 kinds of source wavelet, 3 kinds of wave field noise levels, The input of  $NET0(X1)$  are profiles (P0-2) and the output are categories (c0-2). The red, blue and green baskets correspond to the training/testing models X1, M1 and M2 respectively.

Among them, the training model X1 only has Rick wavelet profile, and 2 test models also have composite wavelet profiles.

By comparing figure 4 and figure 3, it can be seen that although the training model X1 is quite different from the test models M1 and M2,  $NET0(X1)$  can still automatically adapt to the changes of different wavelet and noise, inverse the seismic profile and image the underground scatterers. There is an obvious correlation between the actual inversion output and the expected ideal output, especially the shallow karst cave imaging is good. It is expected that the inversion network  $NET0(X1)$  has strong anti-noise ability and

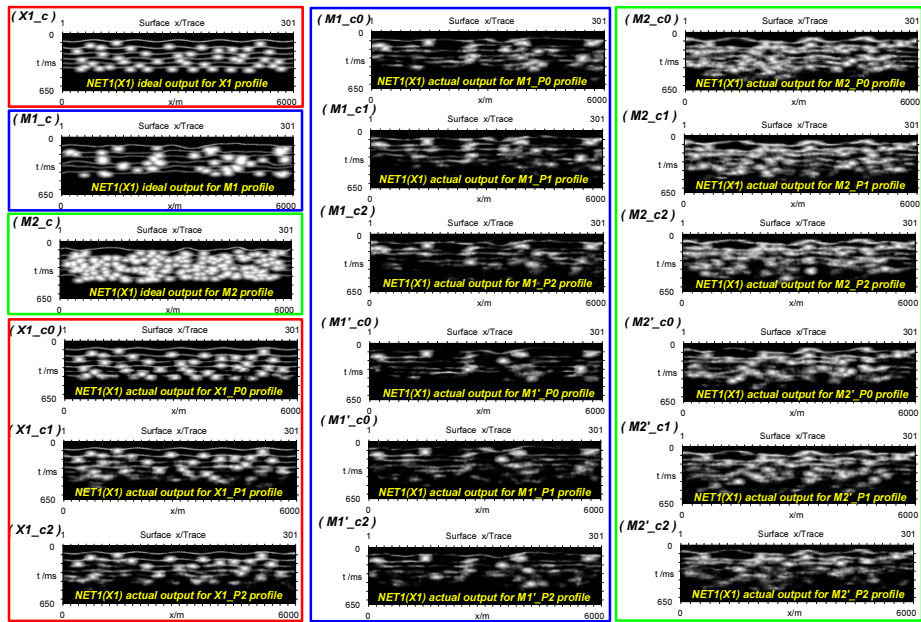
robustness, can adapt to certain wave field noise and source wavelet changes, and inverse complex scattered seismic profiles.

## 2.8. The Extension and Application of the Cnn Inversion Method

The above inversion method can be further popularized and applied. For example, different characteristic inversion networks can be obtained by setting different scattering distance fields; Bayesian discriminant method can further improve the imaging effect and so on.

### 2.8.1 Set the Scattering Distance Field to Invert the Reflection Interfaces and Scatterers at the Same Time.

We treat the interface reflection point as a special scattering point, and set the imaging radius of the interface reflection point as one tenth of the imaging radius of the scattering point in the scatterer. Then the inversion network trained by the equivalent training model  $X1$  and its wave field is recorded as  $NET1(X1)$ . The only difference between the inversion network  $NET1(X1)$  and  $NET0(X1)$  is that the scattering distance field is set differently. For the 15 wave field profiles (shown in figure 4), the 15 inversion results of  $NET1(X1)$  are shown in figure 5.



**Figure 5.** The actual inversion output (c) of 15 scattered wave field profiles (corresponding to 3 models, 2 source wavelets and 3 wave field noise levels) is inverted with  $NET1(X1)$ . The red, blue and green baskets correspond to the equivalent training model  $X1$ , test model  $M1$  and  $M2$  respectively. The 3 figures in the upper left corner are the ideal output of  $NET1(X1)$ , and the other figures are the actual output.

By comparing figure 4 and figure 5, it can be seen that the inversion results of  $NET0(X1)$  and  $NET1(X1)$  are very different, which can be compared and reference complementary. The latter can image the interface at the same time, and the imaging of scatterers can also provide effective additional information.



### 2.8.2 Bayesian Discriminant Imaging.

A Bayesian classifier is designed based on the training model, and then the classifier is used to judge the test model point by point. According to the inversion results of  $NET0(X1)$  and  $NET1(X1)$ , the corresponding Bayesian imaging results can be obtained respectively. The specific steps are as follows:

All  $n$  sample points in the training model  $X1$  are divided into  $k$  categories ( $k = 10$  is selected in this paper), and the corresponding  $k$  populations are recorded as  $G_1, G_2, \dots, G_k$ . Calculate the prior probability  $q_i$  and density function  $f_i(x)$  of the  $i$ -th population as follows:

Replace the probability with frequency to obtain the priori probability:  $q_i = \frac{n_i}{n}$ , where  $n$  is the total number of sample points,  $n_i$  is the number of sample points from the  $i$ -th population in the training model, and  $n_1 + n_2 + \dots + n_k = n$ .

The feature of each sample point is taken as the CNN's recognition result of all  $m$  sample points in the neighborhood of the sample point. Obtain the  $n_i \times m$  sample data matrix of the  $i$ -th population:

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n_i1} & x_{n_i2} & \cdots & x_{n_im} \end{pmatrix}$$

The sample mean vector  $\mu^{(i)}$  and sample covariance matrix  $\Sigma^{(i)}$  of population  $G_i$  in the training model are calculated. It is assumed that the density function of population  $G_i$  follows  $m$ -dimensional normal distribution, and the density function is:

$$f_i(x) = (2\pi)^{-\frac{m}{2}} |\Sigma^{(i)}|^{-\frac{1}{2}} \cdot \exp \left\{ -\frac{1}{2} (x - \mu^{(i)})' \Sigma^{(i)-1} (x - \mu^{(i)}) \right\}$$

Bayesian formula can be used to judge point by point. For the test model sample  $x$  ( $m$ -dimensional vector), the posterior probability from the  $i$ -th population can be calculated:

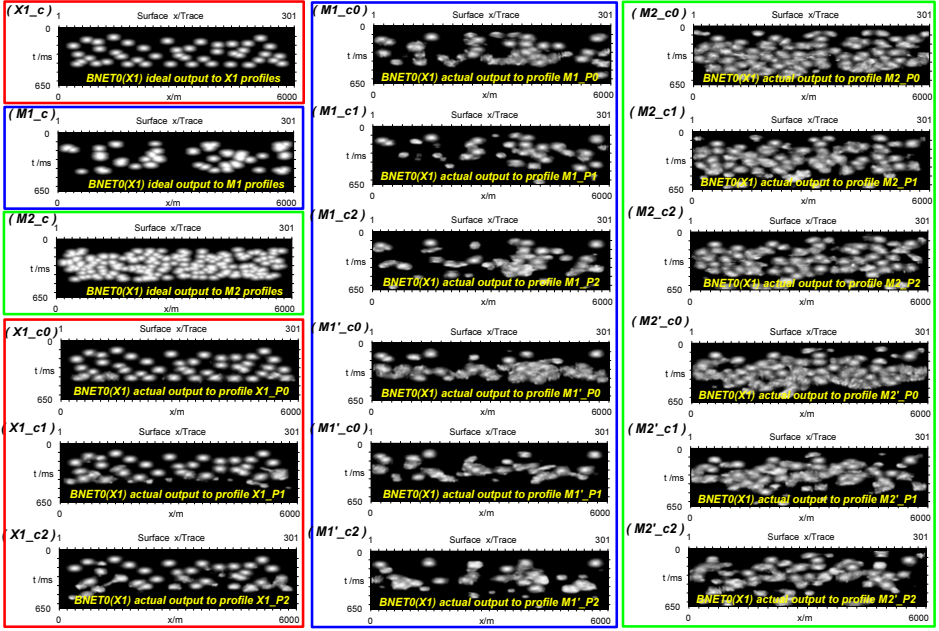
$$p(i/x) = \frac{q_i f_i(x)}{\sum_{i=1}^k q_i f_i(x)}, \quad i = 1, 2, \dots, k$$

In the study, we choose the square loss function:  $L(j/i) = (i - j)^2$ , which represents the loss of the sample  $x$  was originally the  $i$ -th population misjudged as the  $j$ -th population. The average loss of awarding sample  $x$  to the  $j$ -th population is:

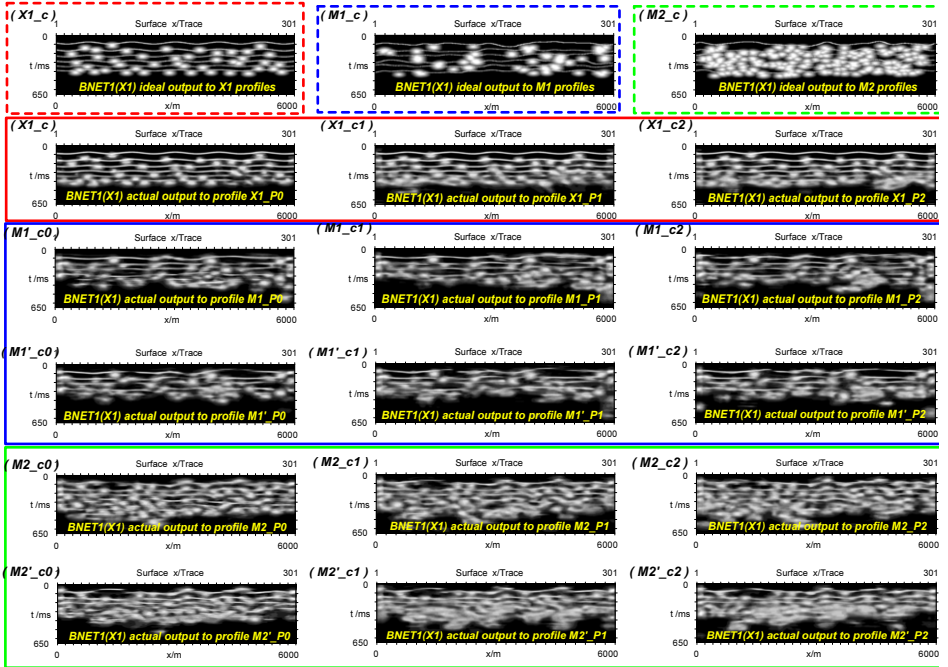
$$E(j/x) = \sum_{i=1}^k \frac{q_i f_i(x)}{\sum_{i=1}^k q_i f_i(x)} \cdot L(j/i)$$

Therefore, the Bayesian criterion is established as follows: When  $E(j/x) = \min_{1 \leq i \leq m} E(i/x)$ , it is determined that  $x$  comes from the  $j$ -th population.

Bayesian classification maps based on the inversion results of  $NET0(X1)$  and  $NET1(X1)$  are called  $BNET0(X1)$  and  $BNET1(X1)$  inversion maps of corresponding wave field profiles, respectively. Bayesian discriminant classification is carried out point by point for the  $2 \times 15$  CNN inversion imaging results of figure 4 and figure 5. The results are shown in figure 6 and figure 7.



**Figure 6.** The ideal output and actual output of 15 scattered wave field profiles are inverted by  $\text{BNETO}(X1)$ . Corresponding to 2 kinds of source wavelet, 3 kinds of wave field noise levels. The red, blue and green baskets correspond to 3 training/testing models  $X1$ ,  $M1$  and  $M2$  respectively. Among them, the training model  $X1$  only has Rick wavelet profile, and 2 test models also have composite wavelet profiles.

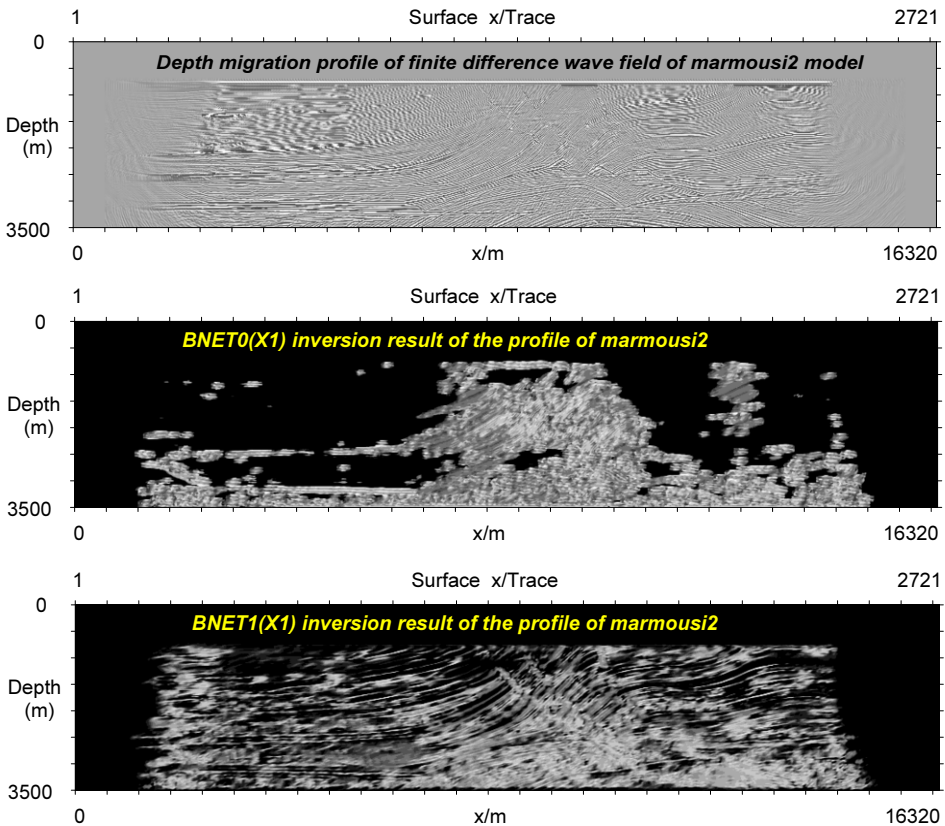


**Figure 7.** The ideal output and actual output of 15 scattered wave field profiles are inverted by  $\text{BNETO1}(X1)$ . Corresponding to 2 kinds of source wavelet, 3 kinds of wave field noise levels. The red, blue and green baskets correspond to 3 training/testing models  $X1$ ,  $M1$  and  $M2$  respectively. Among them, the training model  $X1$  only has Rick wavelet profile, and 2 test models also have composite wavelet profiles.

By comparing figure 4-5 and figure 6-7, it can be seen that Bayesian secondary imaging inversion can further improve the clarity and resolution of inversion results.

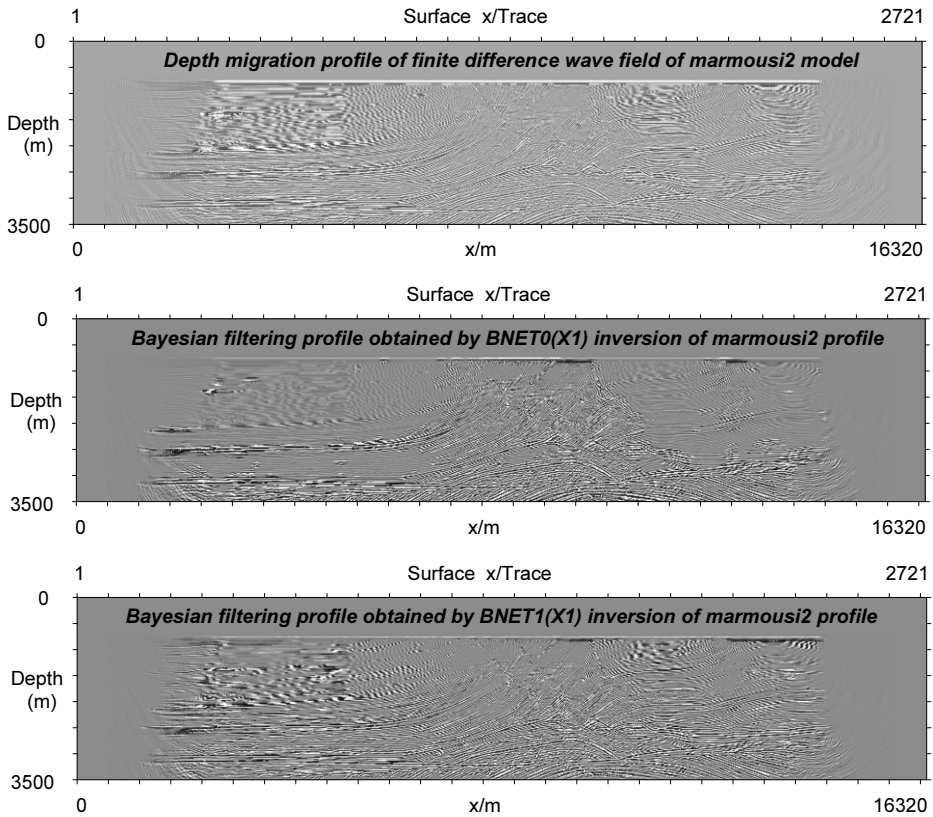
### 2.8.3 Inversion of Marmousi2 Model and the Bayesian Filter Profile.

The finite difference wave field synthesis data of marmousi2 model is often used to verify the practicability of various wave field inversion methods and theories. We use Bayesian inversion networks  $BNET0(X1)$  and  $BNET1(X1)$  based on training model  $X1$  to invert the depth migration profile of marmousi2 model point by point, and the corresponding inversion results are shown in figure 8.



**Figure 8.** From top to bottom is the wave field profile of marmousi2 model and its inversion output of  $BNET0(X1)$  and  $BNET1(X1)$ .

By simply multiplying the results of Bayesian inversion with the test model profile, the Bayesian filter profile can be obtained. The Bayesian filtering profile obtained by  $BNET0(X1)$  and  $BNET1(X1)$  inversion of marmousi2 model profile is shown in figure 9.



**Figure 9.** From top to bottom is the wave field profile of marmousi2 model and profiles obtained by multiplying the Bayesian inversion results of  $BNET0(X1)$  and  $BNET1(X1)$  with marmousi 2 profile.

It is noted that there is a great difference between marmousi2 model and training model  $X1$ ; The time migration profile used by the training model is also different from the depth migration profile of the test model. All kinds of factors will affect the inversion effect. As can be seen from figure 8-9, the inversion network mistakenly considers pinch out and fault as scattering points, but it is still helpful for interpreters to quickly and accurately locate the scattering / diffraction information in the wave field.  $BNET1(X1)$  images both the scatterer and the reflection interface, and the inversion output may have more useful information details. Choosing different equivalent training models and category functions will get different imaging effects, which can help interpreters from different perspectives.

### 3. Conclusion

Through CNN, we can automatically realize end-to-end inversion, so as to improve the efficiency of wave field interpretation. The CNN inversion method proposed by us can be successfully applied to the imaging and positioning of scatterers in various complex scattered wave fields. This method has the following characteristics:

- The inversion method has strong anti-noise, robustness, generalization and adaptability.

- The inversion result map is very intuitive, which can improve the efficiency of wave field interpretation.
- The filtered profile can strengthen the primary scattered wave and suppress other waves at the same time, which can help us understand the specific shape of the scattered wave on the scattered wave field profile identified by the inversion network, and help the interpreter locate the scatterers accurately.

Numerical simulation shows that the very complex scattered wave field in human eyes (a large number of multiple scattered waves, diffracted waves and even multiple waves interfere with each other) has traces to follow in the eyes of computer, so that only through the learning of an equivalent training model X1, the deep learning CNN network can automatically learn some characteristics of the relevant primary wave.

## References

- [1] Xi Xian, Huang JiangQing, (2020), Location and imaging of scatterers in seismic migration profiles based on convolution neural network. *Chinese Journal of Geophysics(in Chinese)*,63(2): 687-714,doi: 10.6038/cjg2020M0490
- [2] Xi Xian, Huang Jiang-qing, (2018), Deep learning inversion imaging method for scattered wavefield. *Progress in Geophysics (in Chinese)*, 33(6): 2483–2489 , doi:10.6038/pg2018BB0531.
- [3] Lecun Y, Bottou L , Bengio Y , et al., (1998), Gradient - based learning applied to document recognition. *Proceedings of the IEEE* , 86 ( 11 ): 2278 - 2324
- [4] Ikelle T, Yung K, Daube F, (1993), 2-D random media with ellipsoidal autocorrelation function. *Geophysics*, 58(9): 1359-1372
- [5] Ergintav S, Canitez N, (1997), Modeling of multi-scale media in discrete form. *Journal of seismic exploration*, 6(1): 77-96
- [6] Xi Xian, Yao Yao, (2002), Simulation of random medium model and mixed type random medium . *Geoscience Journal of China University of Geosciences ( in Chinese)*, 27 (1): 67-71
- [7] Xi Xian, Yao Yao, Gu H M, (2005), Construction of a random cave medium model . *Journal of Huazhong University of Science and Technology (Natural Science Edition) ( in Chinese)*, 33 (9): 105-108
- [8] Virieux J, (1984), SH-wave propagation in heterogeneous media: Velocity-stress finite difference method.*Geophysics*, 49(11): 1933-1957
- [9] Virieux J, (1986), P-SV wave propagation in heterogeneous media: Velocity stress finite difference method.*Geophysics*, 51(4): 889-901
- [10] Xi Xian, Yao Yao, (2001), two-dimensional stochastic medium and wave equation forward modeling . *Oil geophysical prospecting ( in Chinese)*, 36 (5): 546-552
- [11] Xi Xian, Yao Yao, (2004), Analysis of wave field characteristics in two-dimensional viscoelastic random medium . *Progress in Geophysics ( in Chinese)*, 19 (3): 608-615
- [12] Levander A, (1988), Fourth-order finite difference P-SV seismograms. *Geophysics*, 53(11): 1425-1436
- [13] Cerjan, C, Kosloff, D, Kosloff, R, and Reshef, M. (1985), A nonreflecting boundary condition for discrete acoustic wave and elastic-wave equation . *Geophysics*, 50(4): 705~708