

# AI驱动：三维智能断层训练与预测功能

汇报人：熊伟

中国石油东方物探公司物探技术研究中心  
中油油气勘探软件国家工程研究中心

2025年9月29日

GeoEast



## CONTENES 目录

01

02

03

### 研究背景

### 三维智能断层预测功能介绍

三维合成数据断层预测

个性化稀疏三维标签断层预测

个性化三维标签断层预测

### 注意事项及流程建议



# 训练数据是人工智能的“原油”



AI系统= 数据 + 模型/算法

前期训练数据准备 后期数据飞轮迭代

模型构建 训练

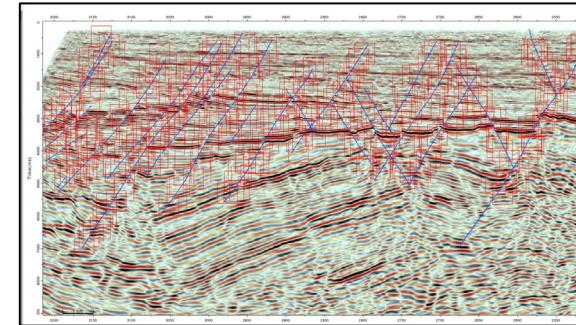
80%

20%

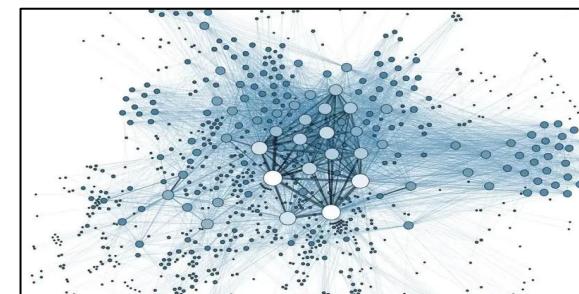
系统的核心是模型与训练数据，但在实际应用中我们发现：

系统落地效果的好坏只有20%取决于算法，80%取决于训练数据的质量

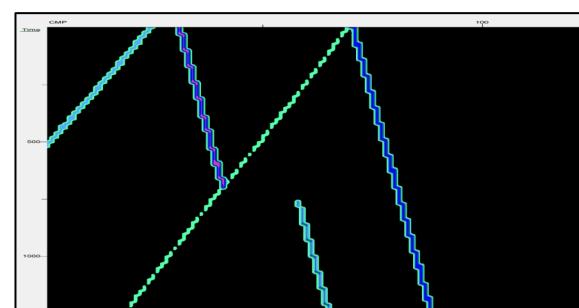
系统的成败80%取决于训练数据是否优质



训练样本集建立



网络训练



断层预测

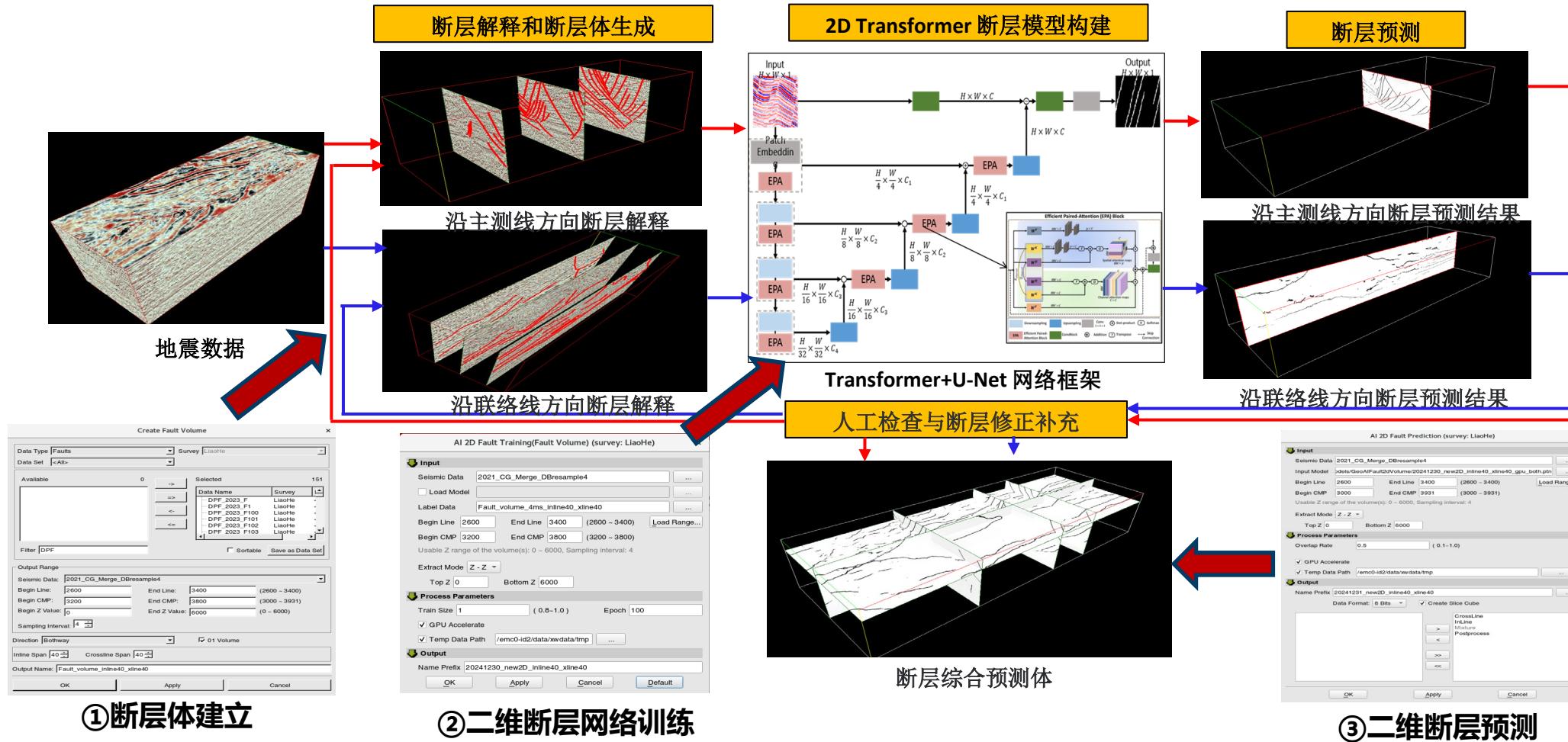
个性化2D标签智能断层预测流程



# 个性化二维标签智能断层预测技术

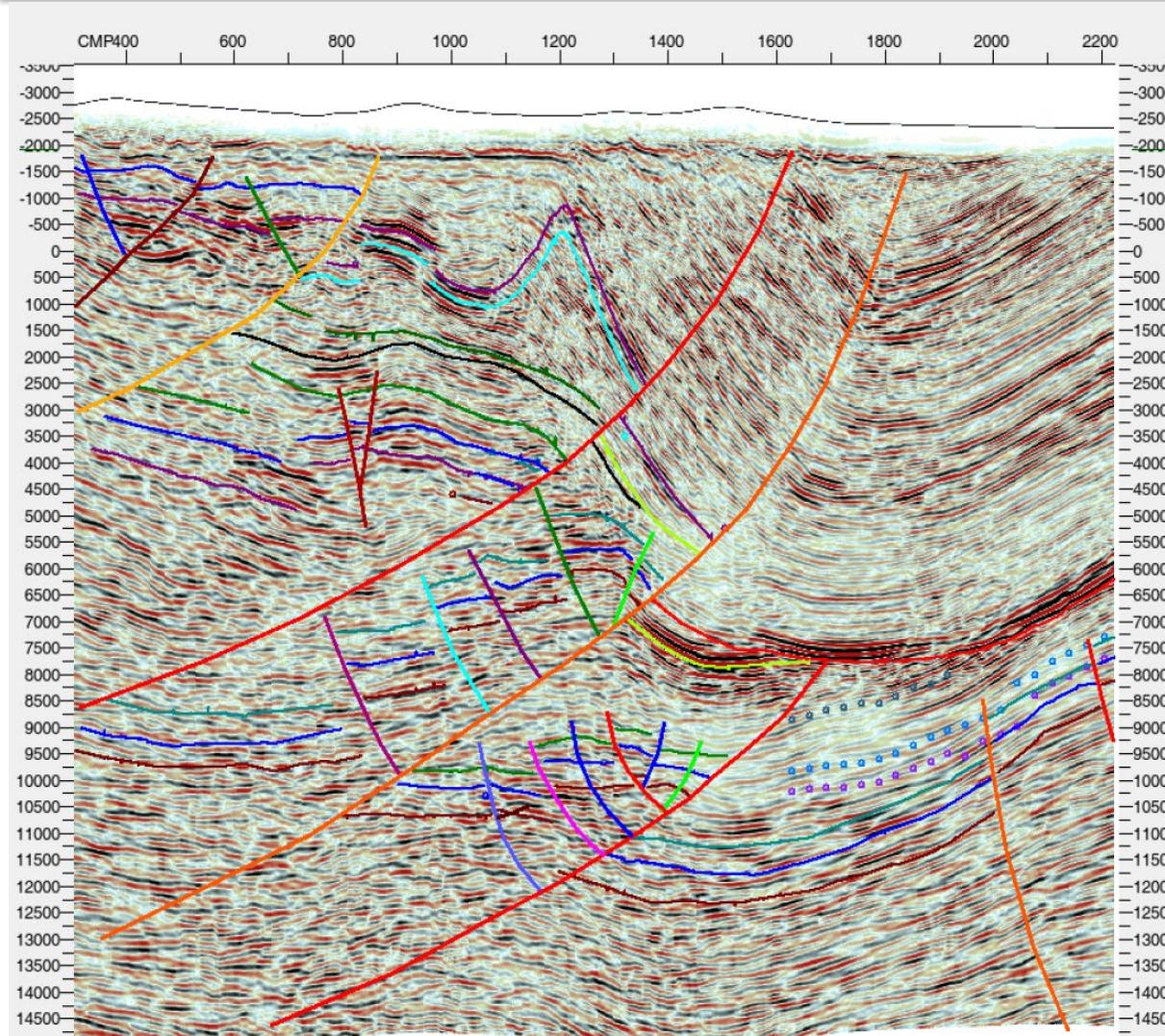


人工断层解释通常在二维剖面上完成，难以获取的实际三维断层标签，因此采用了Transformer网络结合个性化二维标签（专家标签）的智能断层预测。

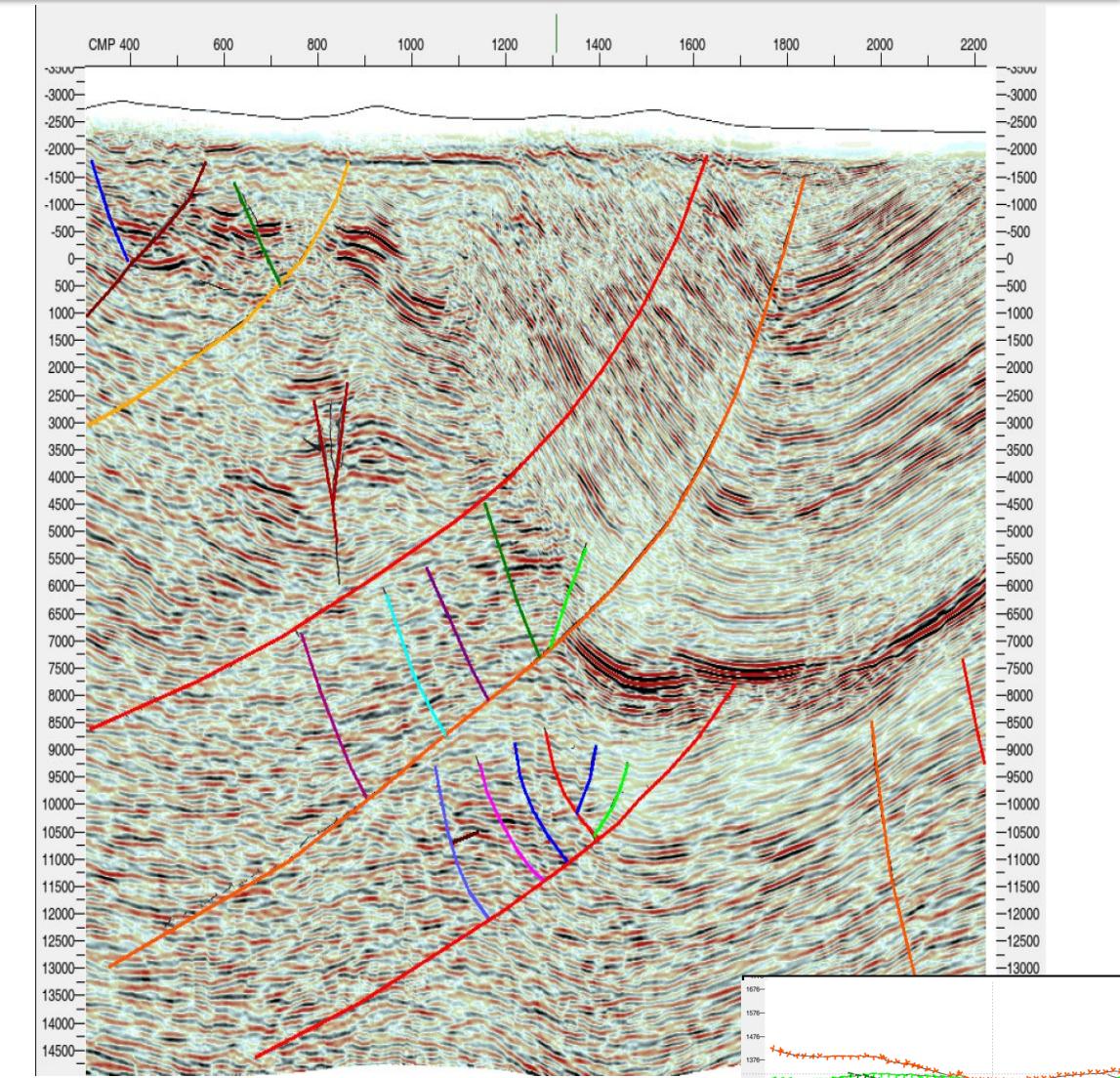




# 个性化二维标签智能断层预测技术



低角度逆掩断裂

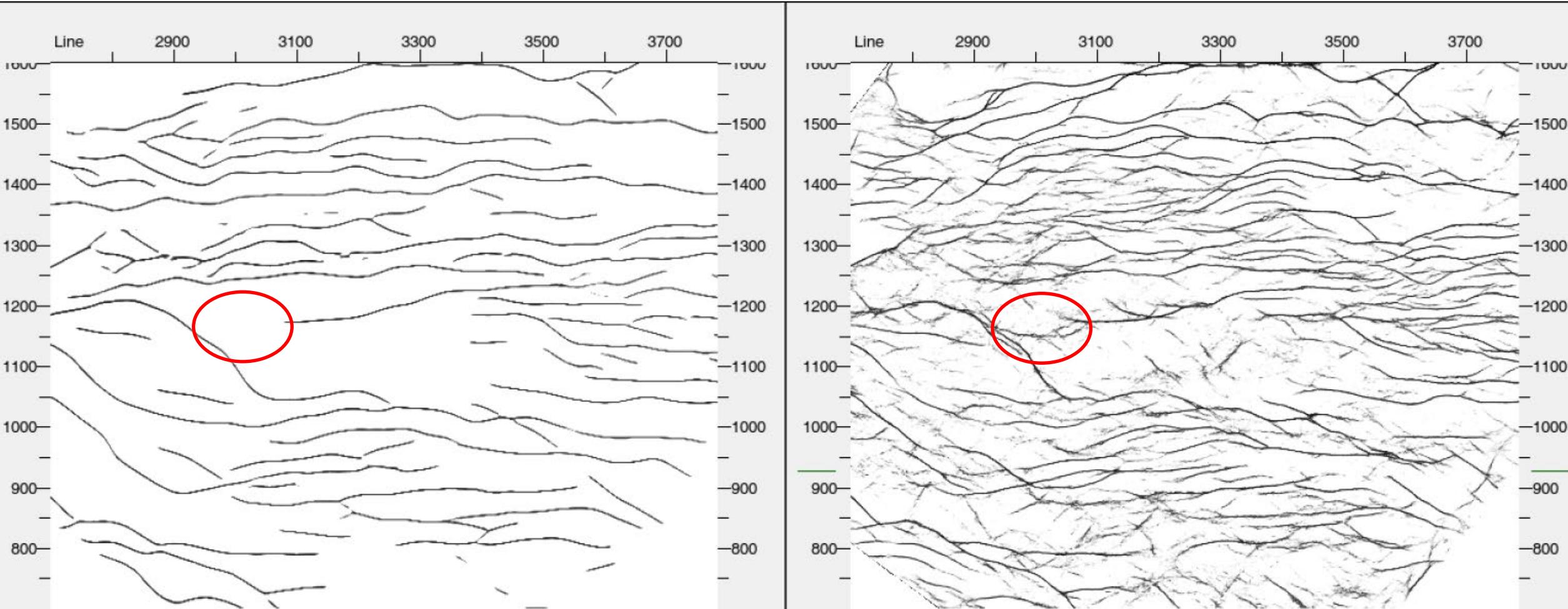


断层预测结果



# 面临的问题

## 个性化2D二维标签与三维合成训练数据断层预测效果对比



个性化2D标签

时间切片

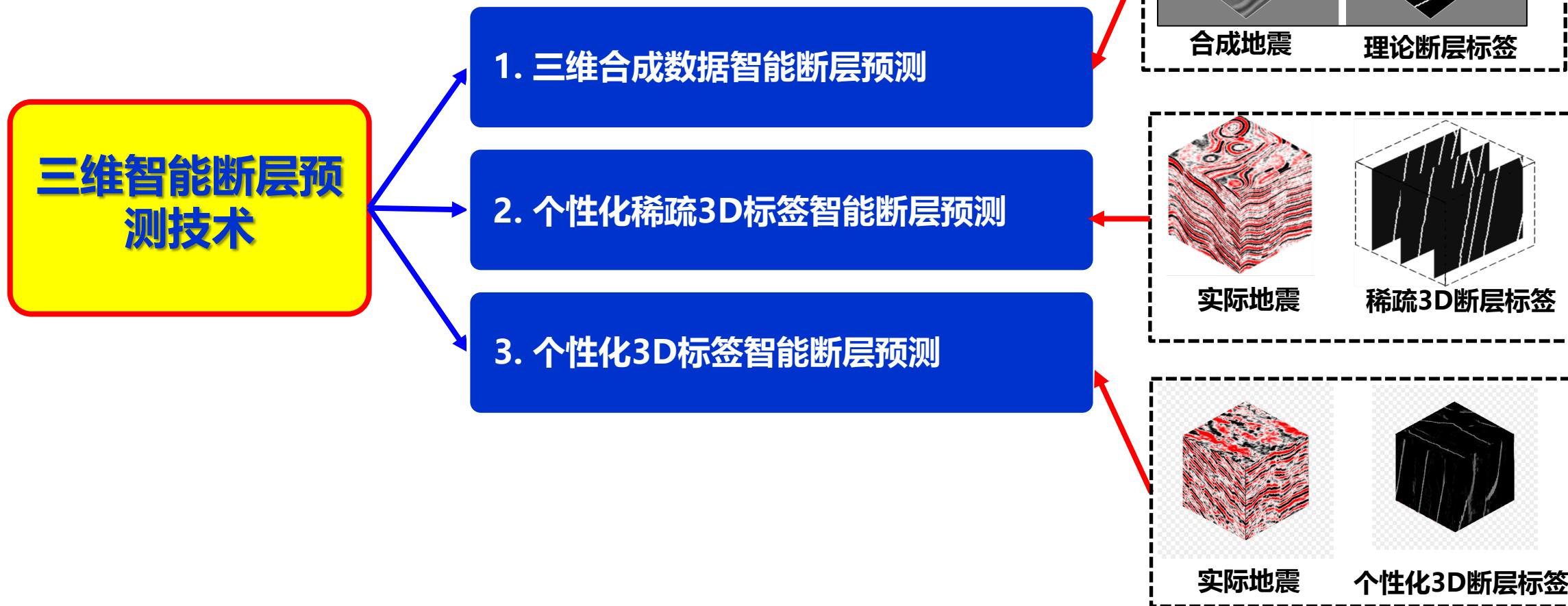
三维合成训练数据



# 解决方案



GeoEast目前提供3种三维训练数据智能断层预测技术。





## CONTENES 目录

01

02

03

研究背景

## 三维智能断层预测功能介绍

三维合成数据断层预测

个性化三维稀疏标签断层预测

个性化三维标签断层预测

注意事项及流程建议



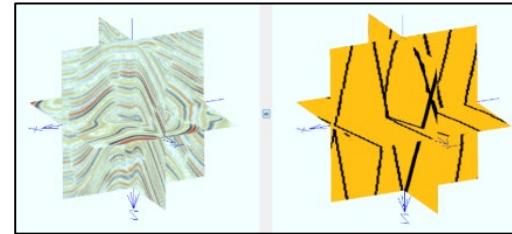
# 1.3D合成数据

## 方法原理



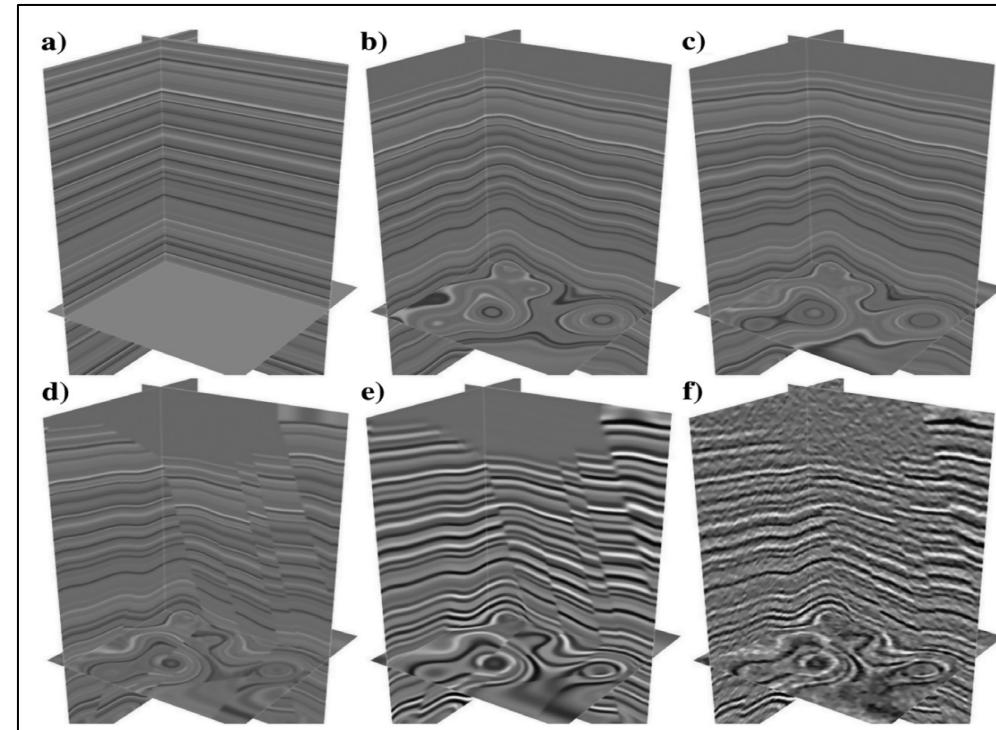
合成数据由200训练集、20验证集组成，每个样本大小为128\*128\*128。

至少含有五个断层，断层越多越有利于训练，但这些断层不应该彼此太近。



合成步骤：

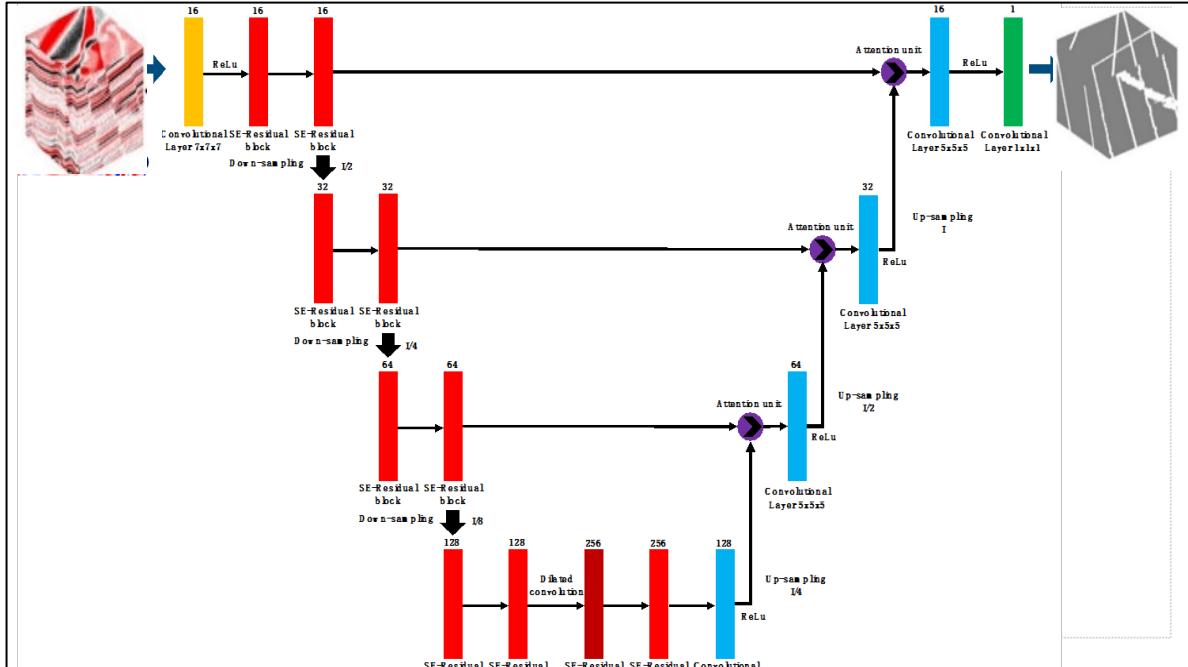
- ①首先生成水平反射率模型；—— (a)
- ②通过[-1,1]的随机数向模型中添加折叠结构，折叠结构利用了二维正态分布函数， $a_0$ 、 $b_k$ 、 $c_k$ 、 $d_k$  描述了一些特定的空间折叠结构；—— (b)
- ③向模型中添加平面剪切， $e_0$ 、 $f$ 、 $g$  这些参数是随机得来；—— (c)
- ④向模型中添加平面断层，每个断层上的倾角和走向都有所不同；—— (d)
- ⑤将 (d) 与Ricker wavelet卷积合成；—— (e)
- ⑥添加随机噪声；—— (f)



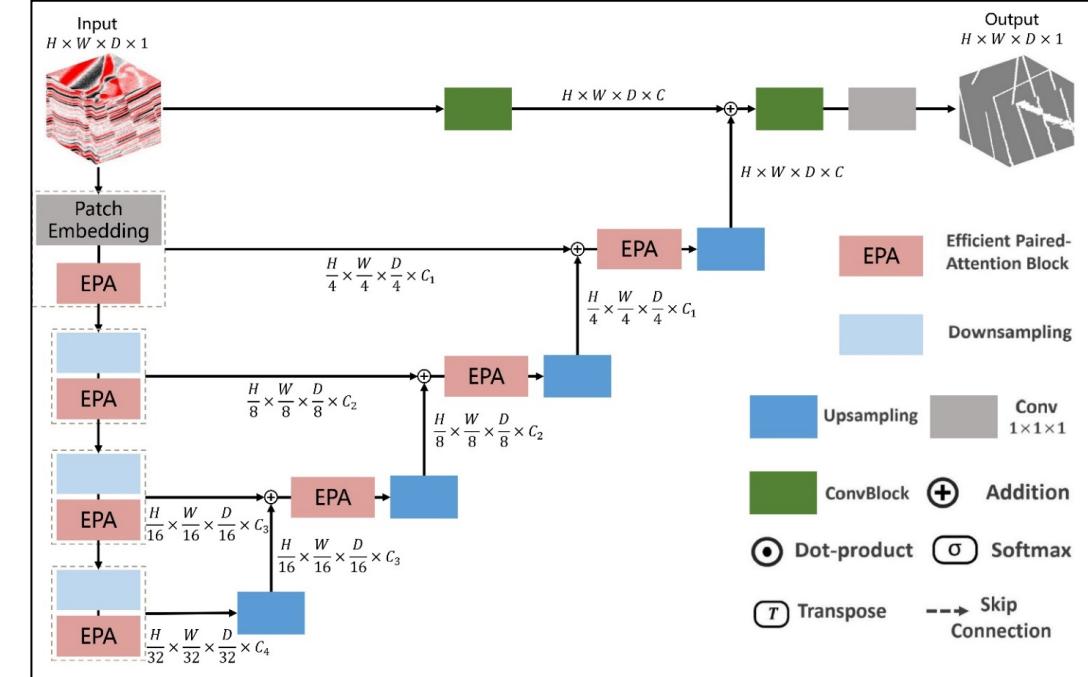
合成断层数据集建立过程



## 两种网络：基于卷积的U-Net网络和Transformer网络



U-Net网络

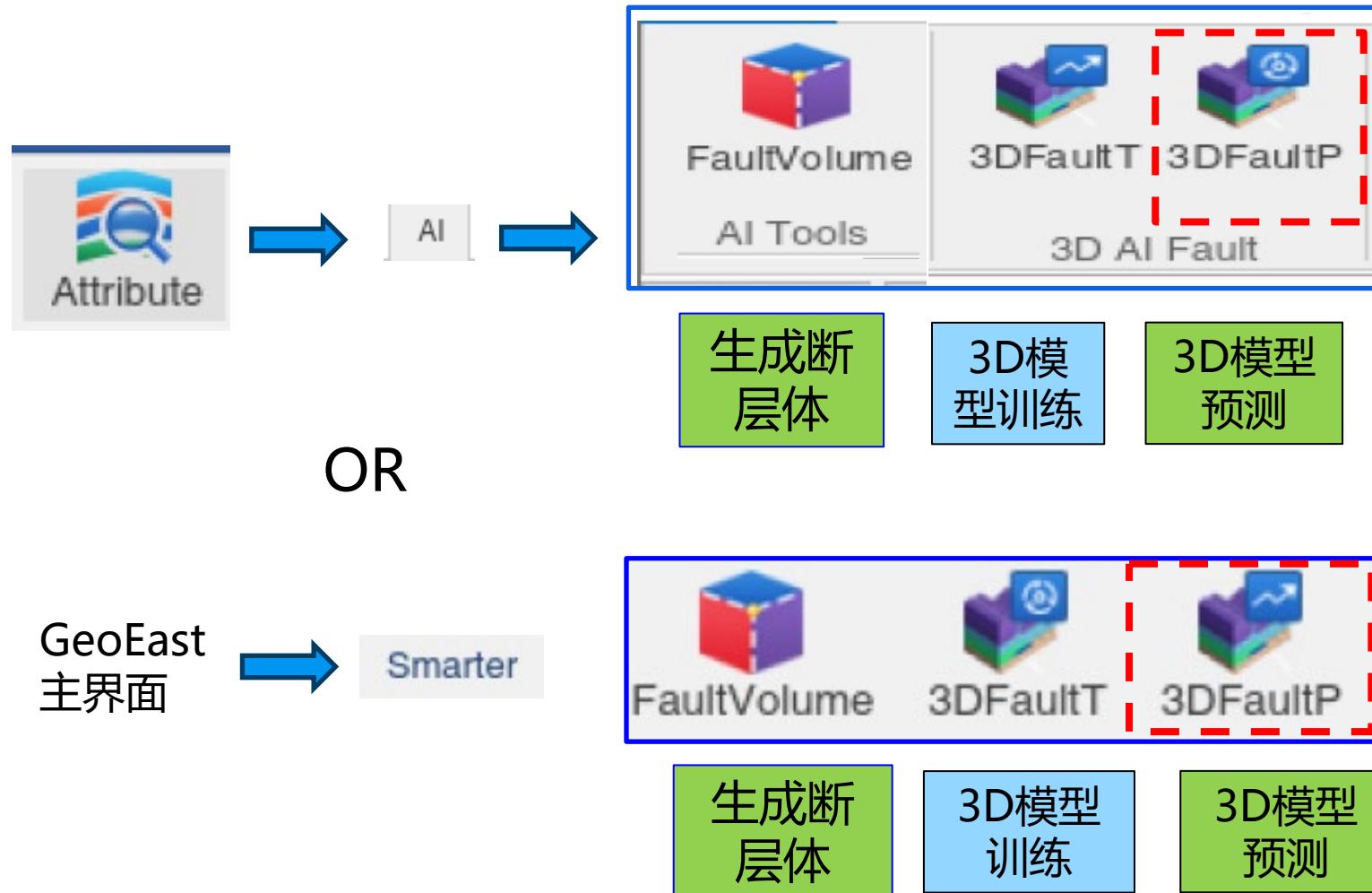


高效成对注意力机制的Transformer网络



# 1.3D合成数据

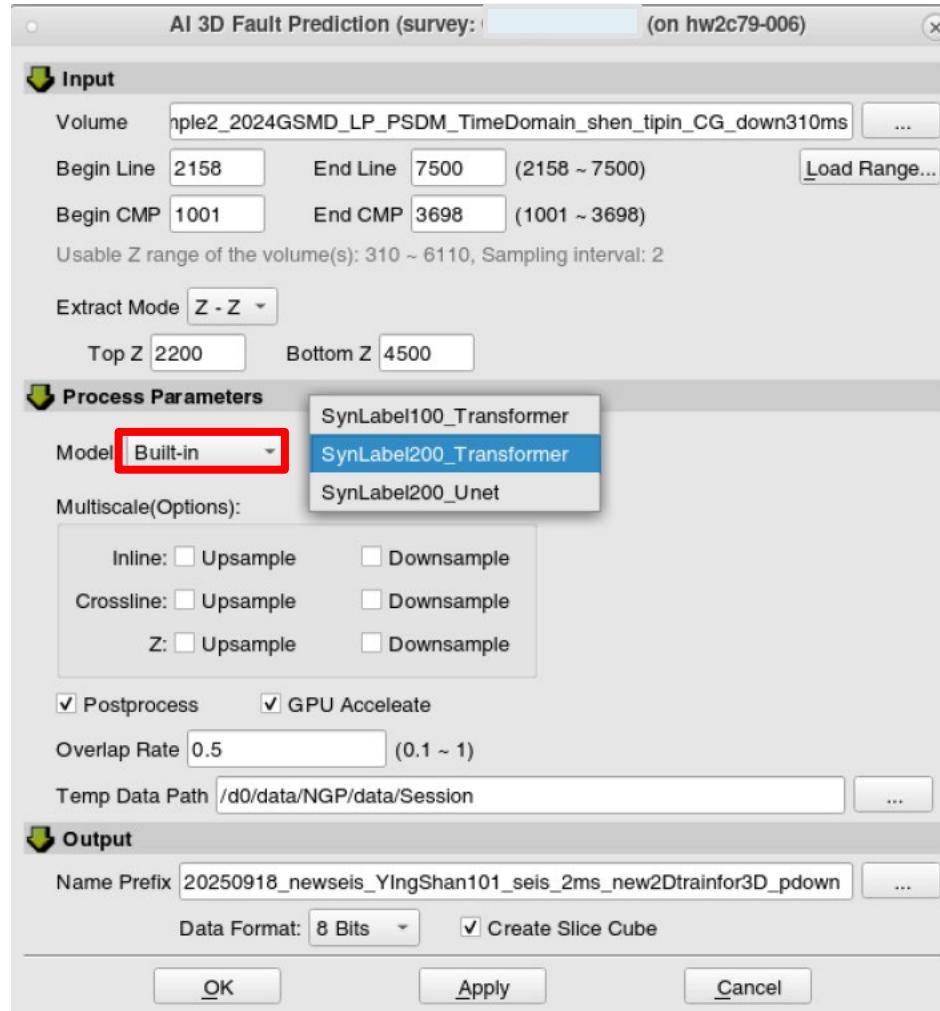
启动位置





# 1.3D合成数据

## 界面及主要参数



输入:

**Seismic Data :** 预测的地震数据,  
程序内部将数据切成128\*128\*128的大小进行预测

主要参数:

**Model:** Built-in 选择内置网络模型

**SynLabel100\_Transformer:** 100个合成数据+Transformer网络

**SynLabel200\_Transformer:** 200个合成数据+Transformer网络

**SynLabel200\_Unet:** 200个合成数据+Unet网络

**Postprocess:** 连续性及细线化处理

**GPU Accelerate :** 是否采用GPU加速训练

**Overlap:** 预测过程块体尺寸重叠率

V4.4



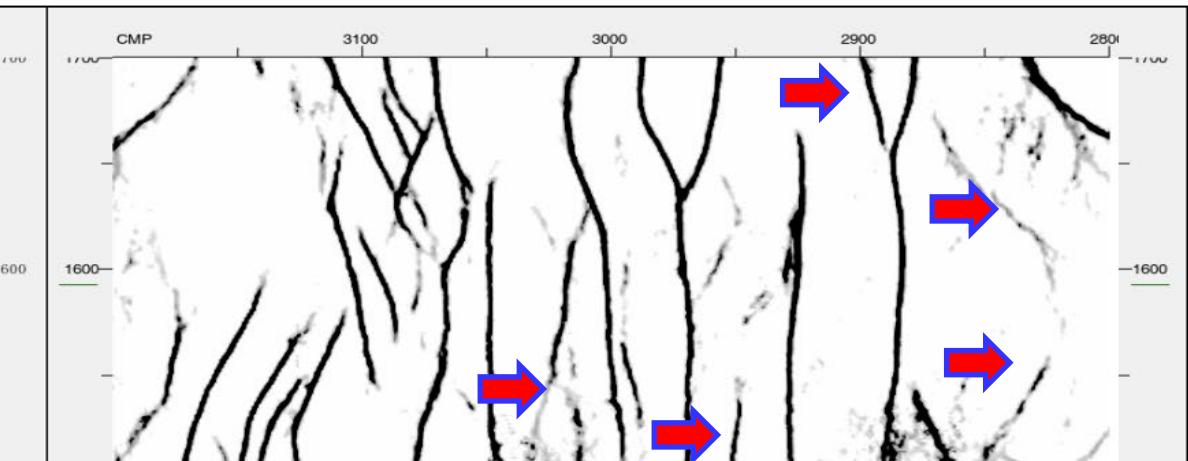
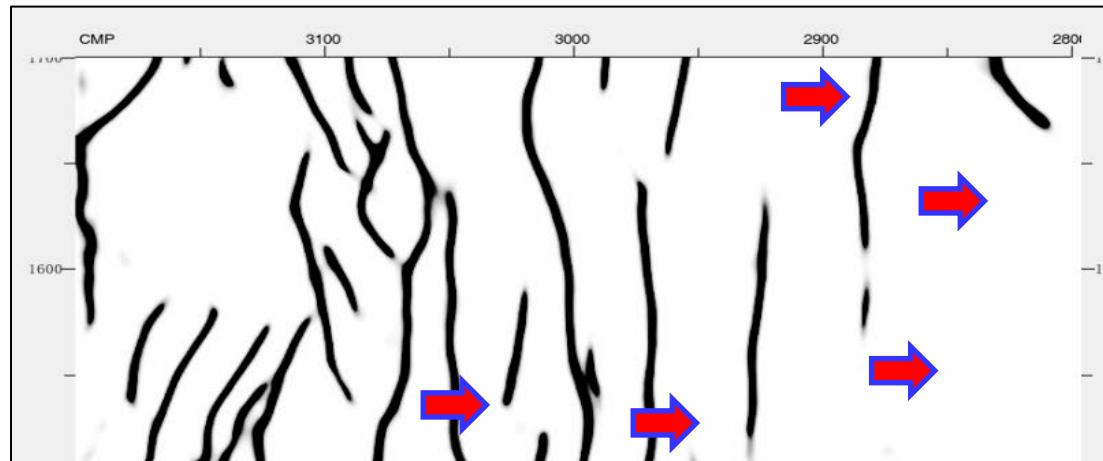
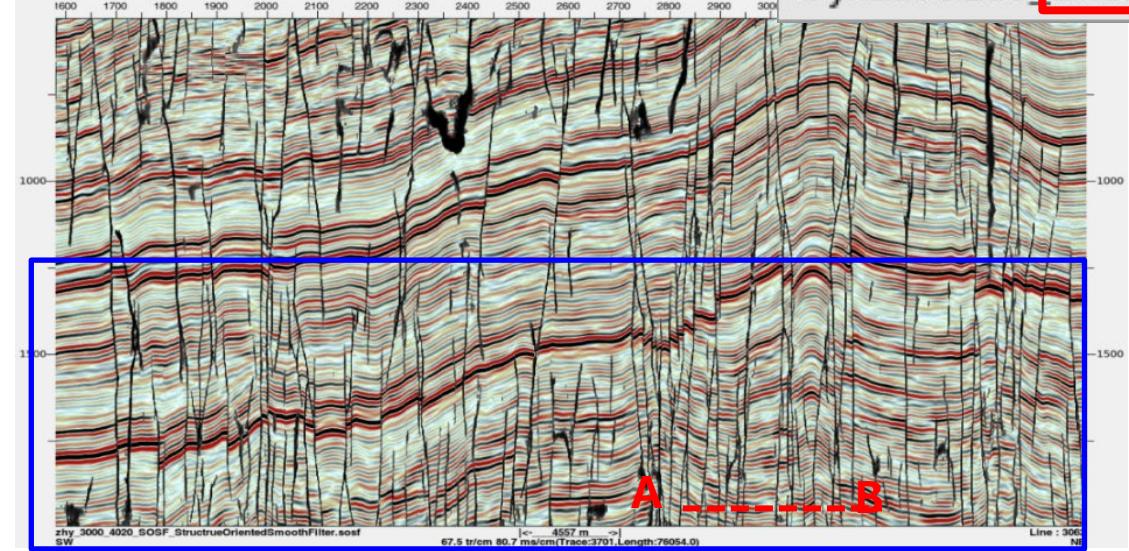
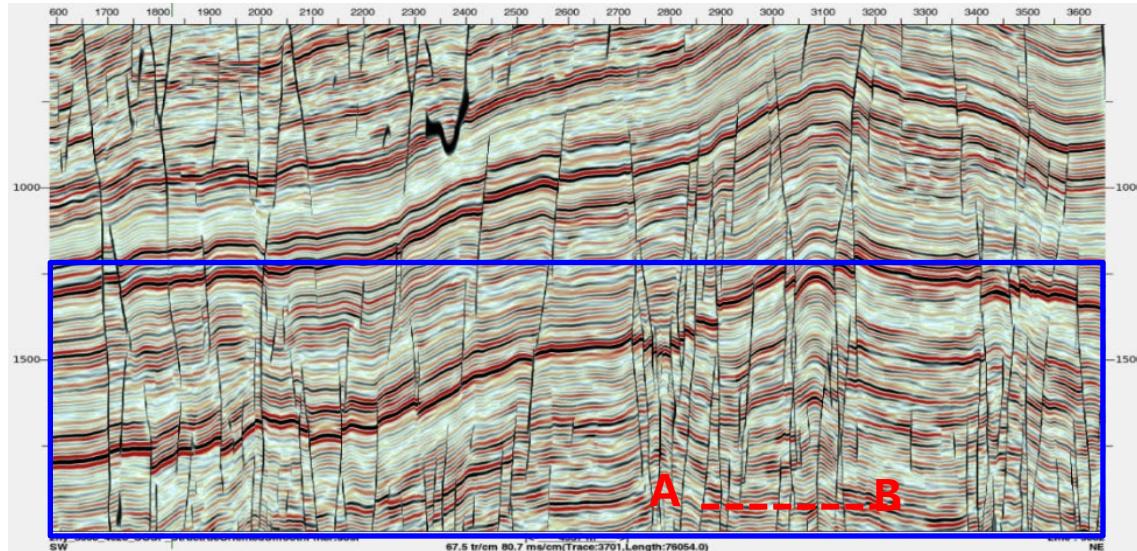
# 1.3D合成数据

## Unet与Transformer效果对比

SynLabel100\_Transformer

SynLabel200\_Transformer

SynLabel200\_Unet



Unet

Transformer



# 1.3D合成数据

## SynLabel100与SynLabel200对比

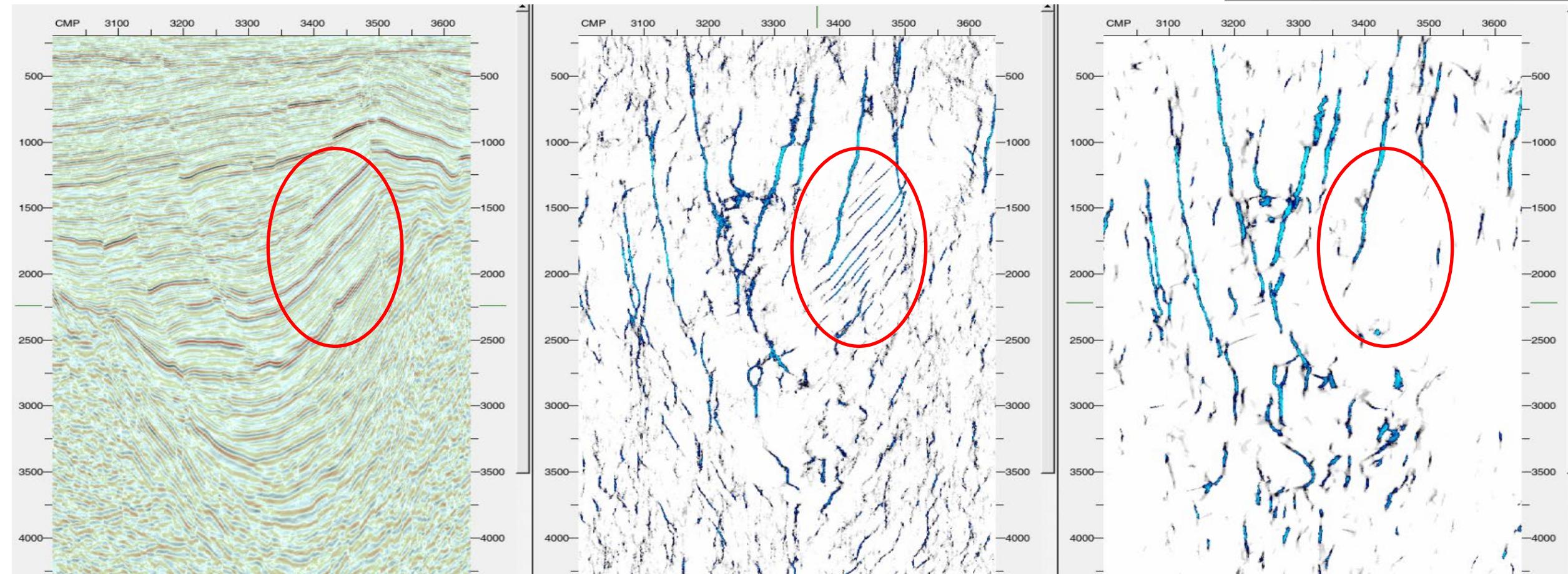


针对将倾斜层位识别为断层，对200对训练样本进行优选，选择100对样本进行训练。

SynLabel100\_Transformer

SynLabel200\_Transformer

SynLabel200\_Unet



原始地震

SynLabe200\_Transformer

SynLabel00\_Transformer



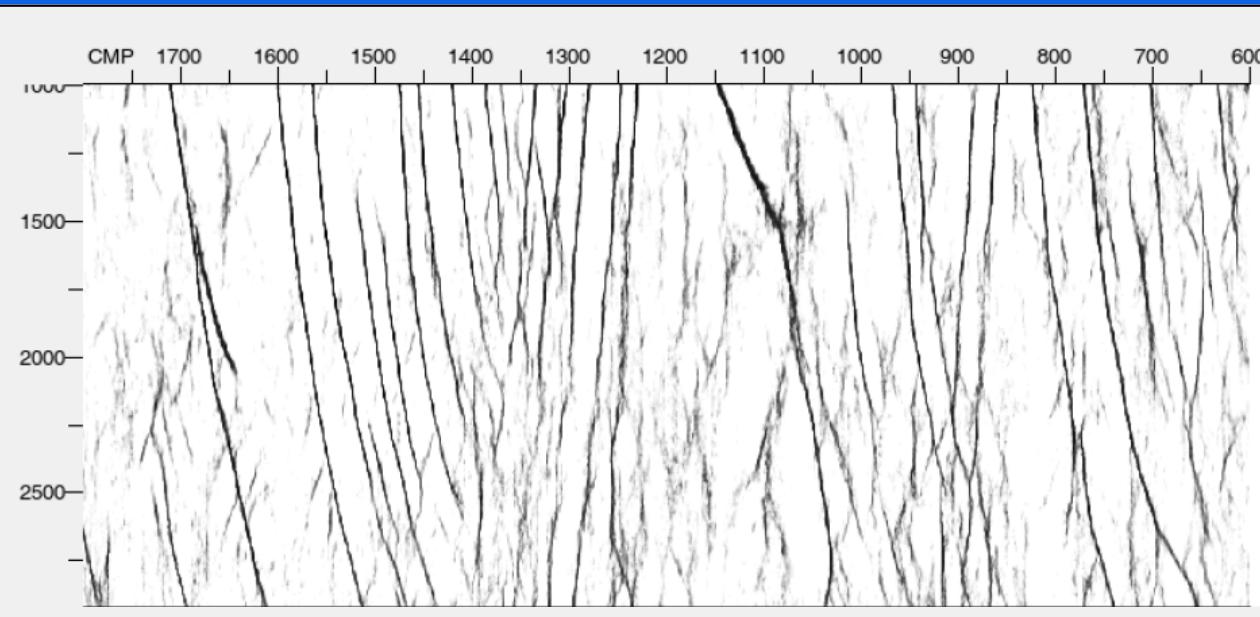
# 1.3D合成数据

## PostProcess的影响

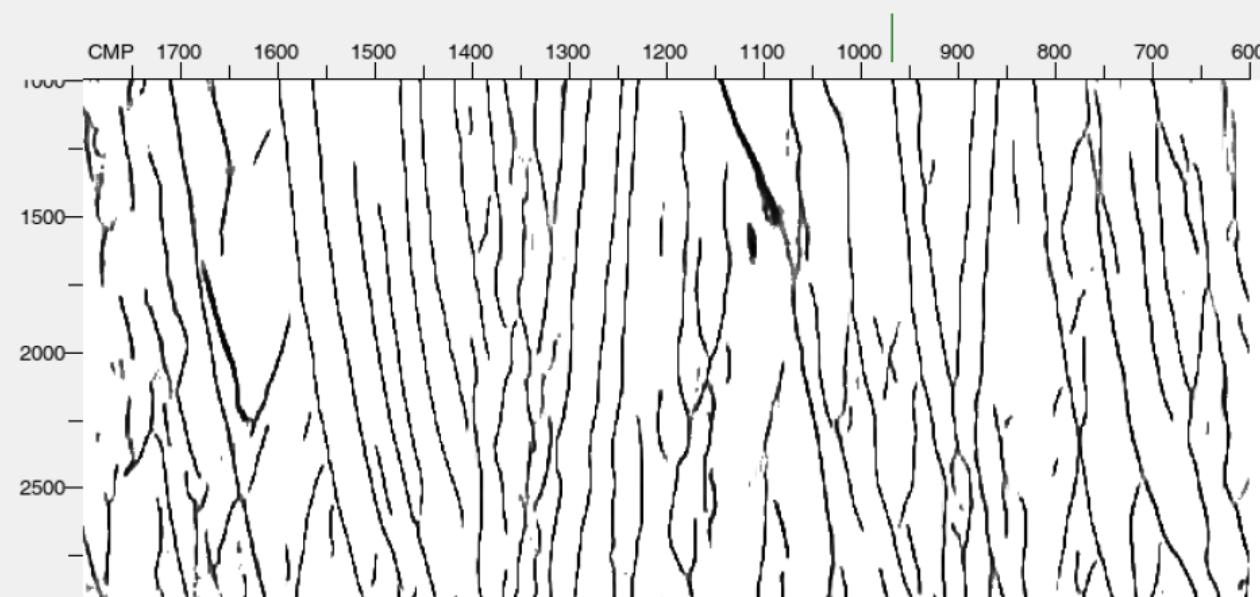


连续性及细线化处理

剖面显示



没有进行后处理结果



进行后处理结果

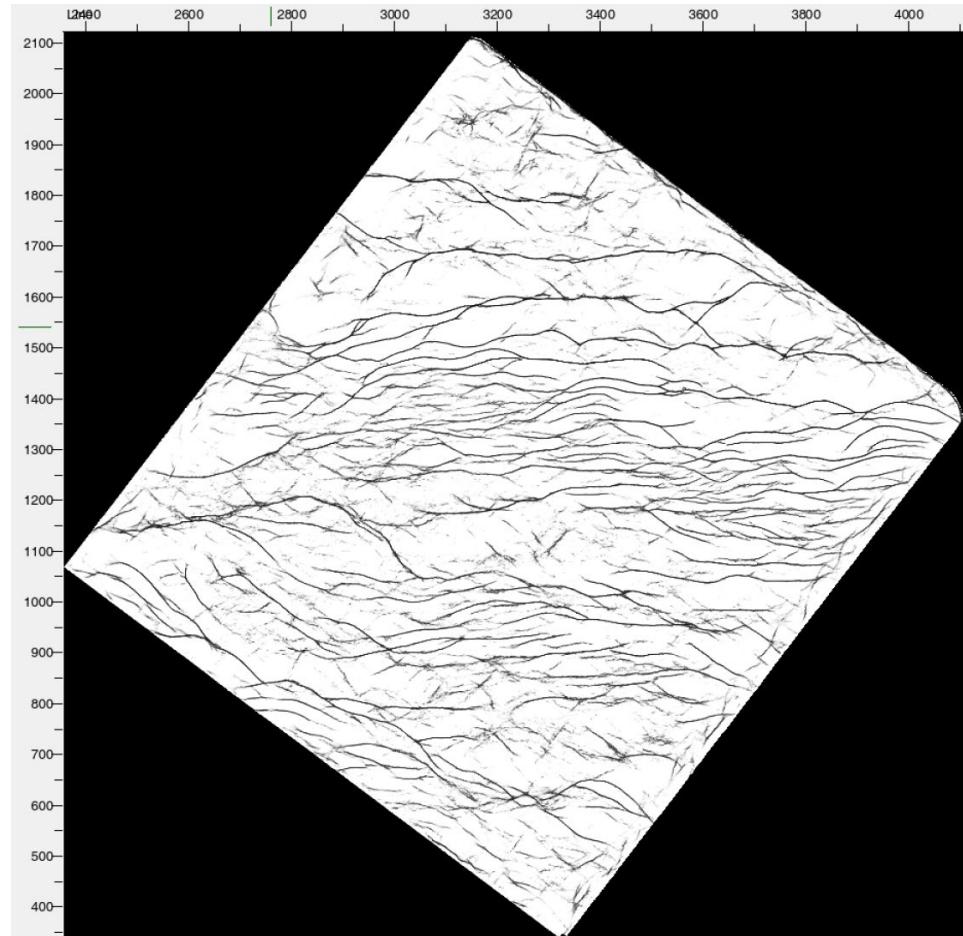


# 1.3D合成数据

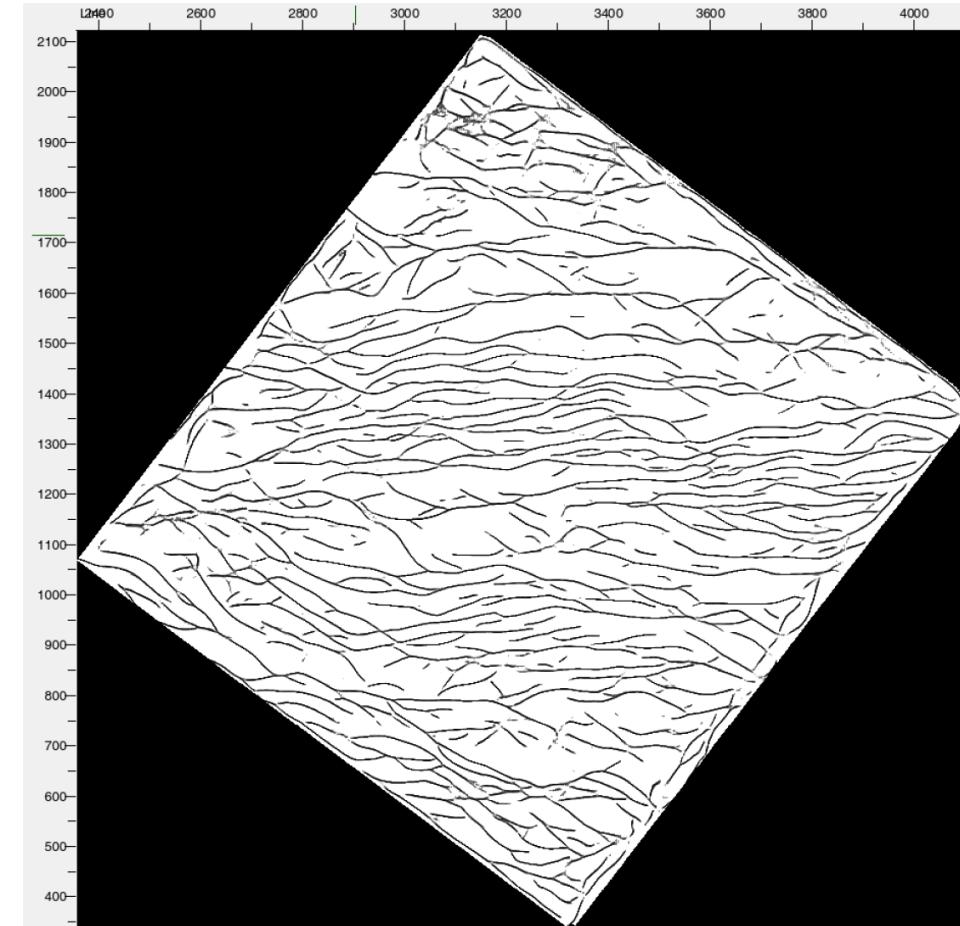
# PostProcess的影响



连续性及细线化处理



没有进行后处理结果



进行后处理结果

时间切片

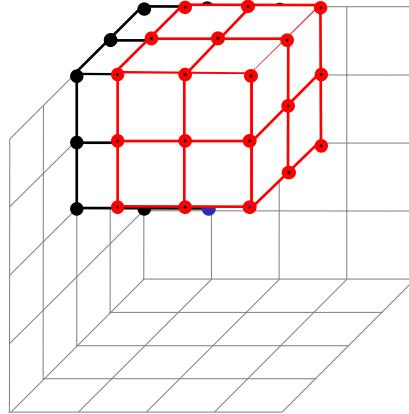


# 1.3D合成数据

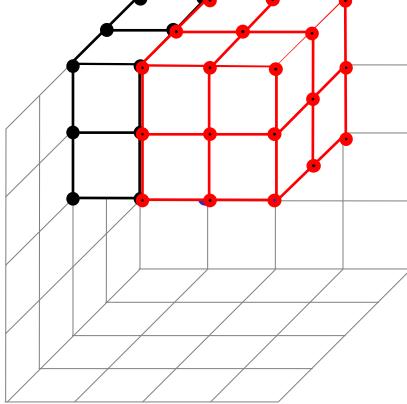
## Overlap Rate的影响



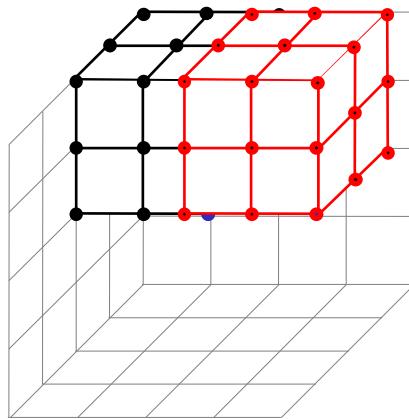
以固定时窗大小（128\*128\*128），按照叠合率滑动进行断层预测



Overlap Rate=0.25

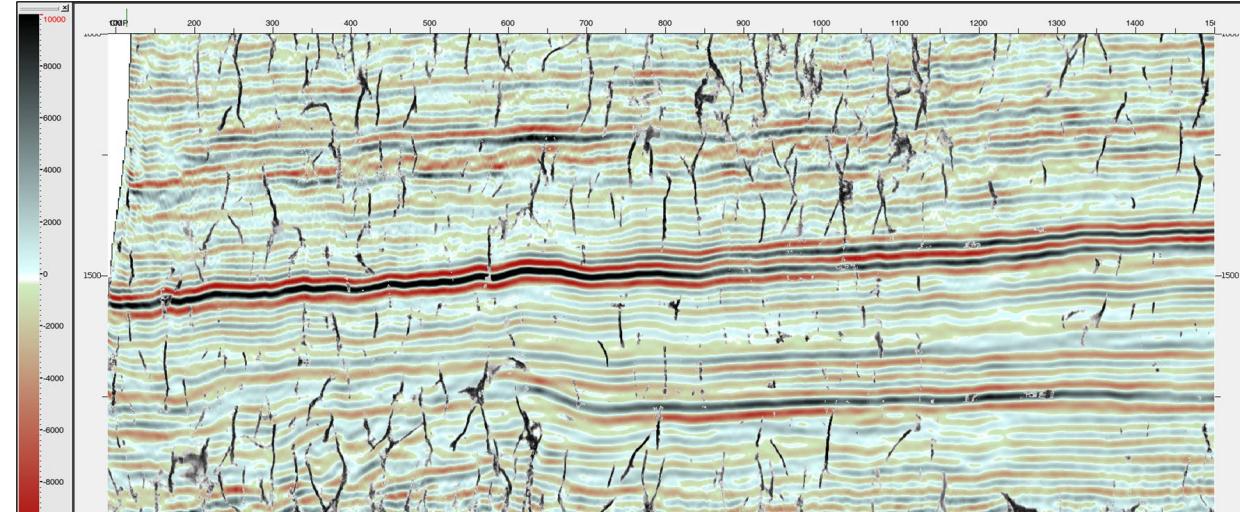


Overlap Rate=0.5

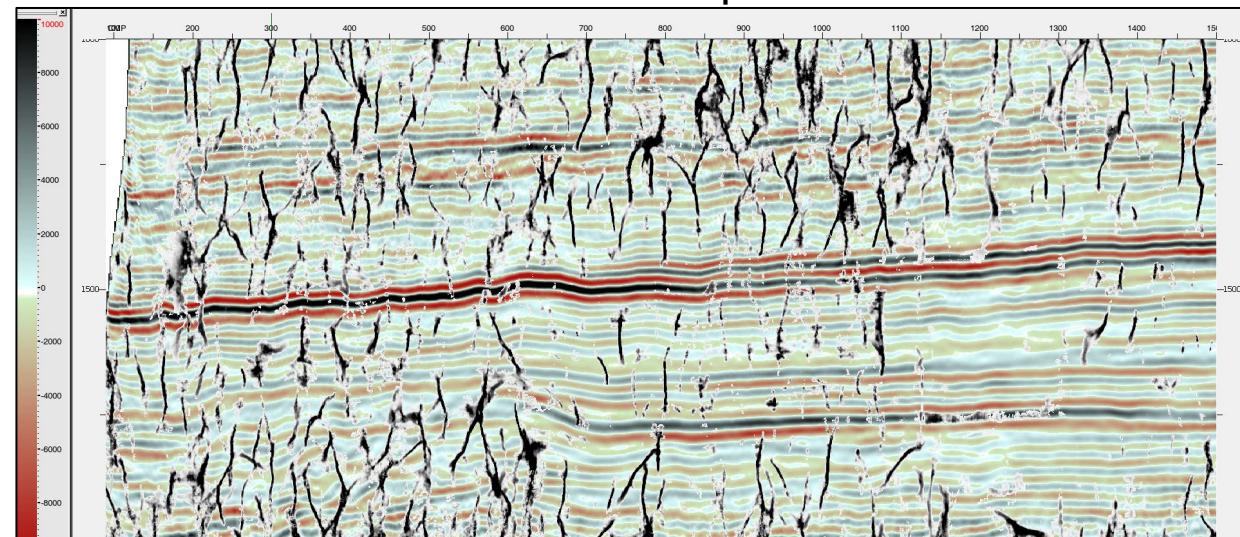


Overlap Rate=0.75

叠合时窗越大，  
识别断层越多，  
运算时间越慢。



Overlap Rate=0.25



Overlap Rate=0.5



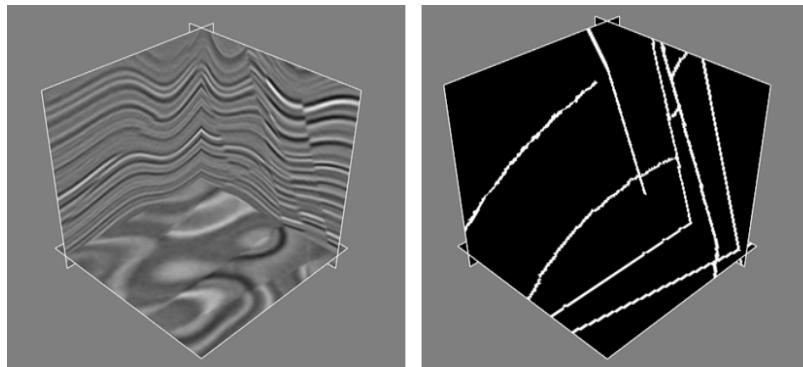
# 1.3D合成数据

内置尺寸 (128\*128\*128)



由于不同工区的线道间距与纵向采样率不同，相同的3D合成数据，对应实际数据尺度不同，会严重影响预测结果。

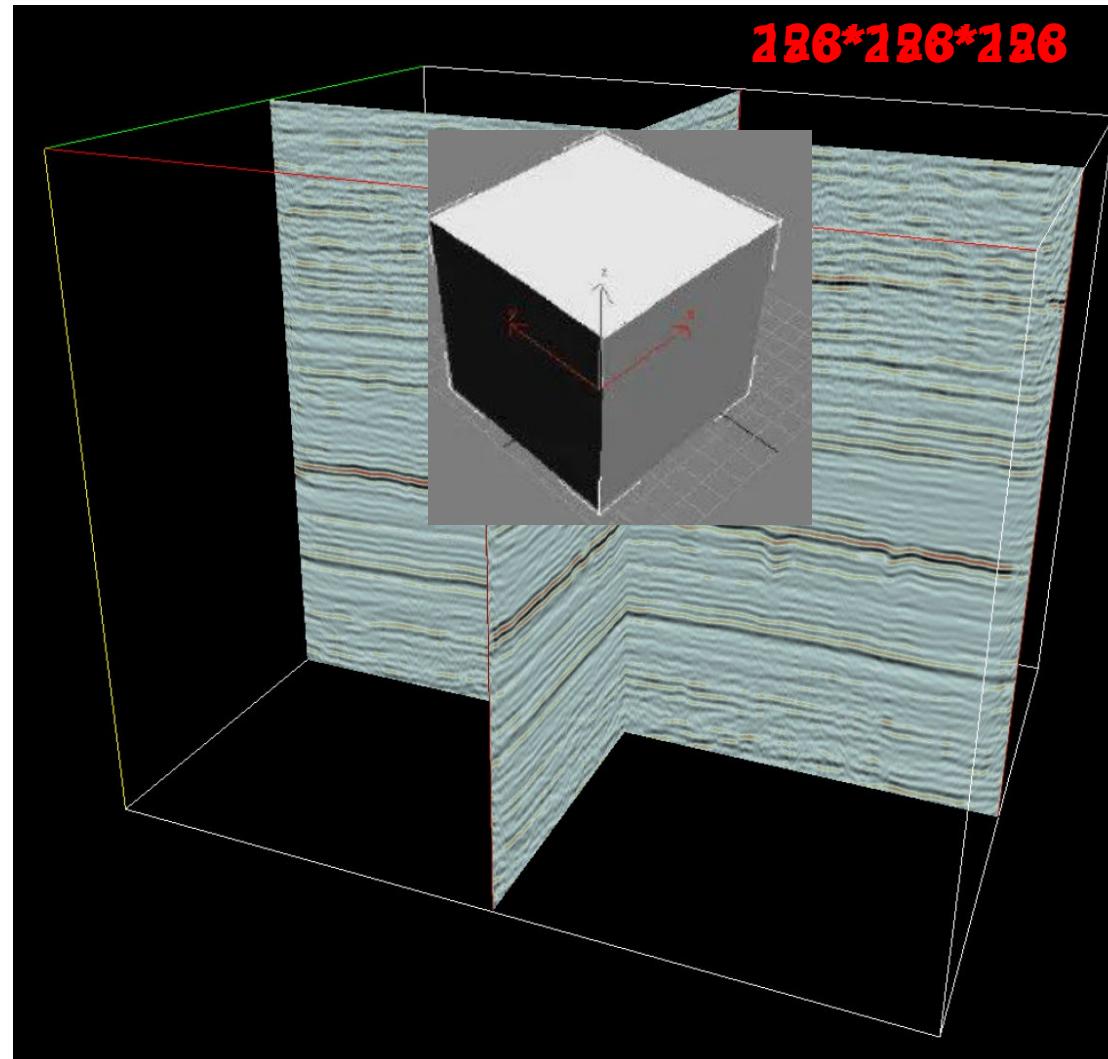
合成训练样本



合成地震

(128\*128\*128)

理论断层标签

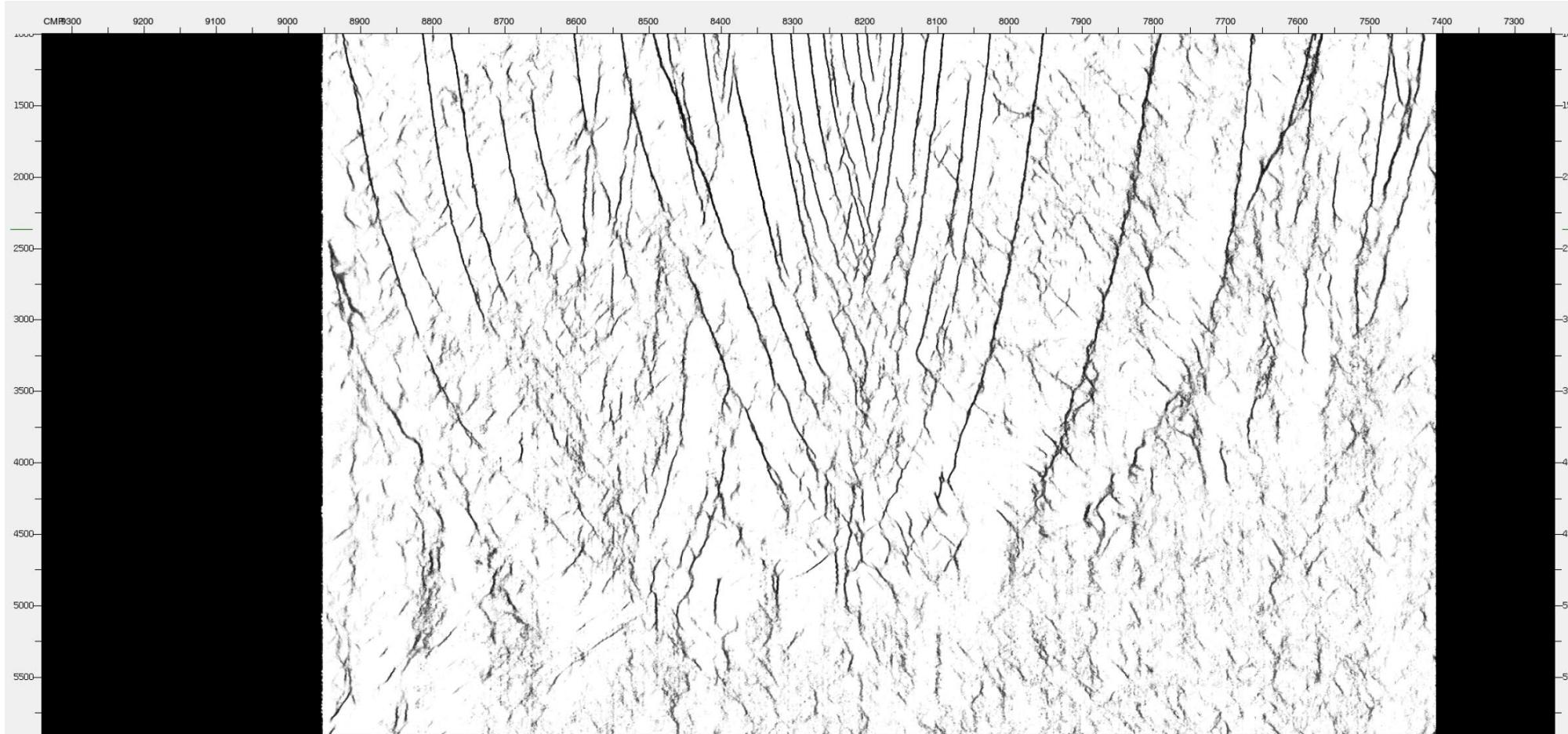


至少含有五个断层，断层越多越有利于训练，但这些断层不应该彼此太近。



# 1.3D合成数据

## 采样率的影响



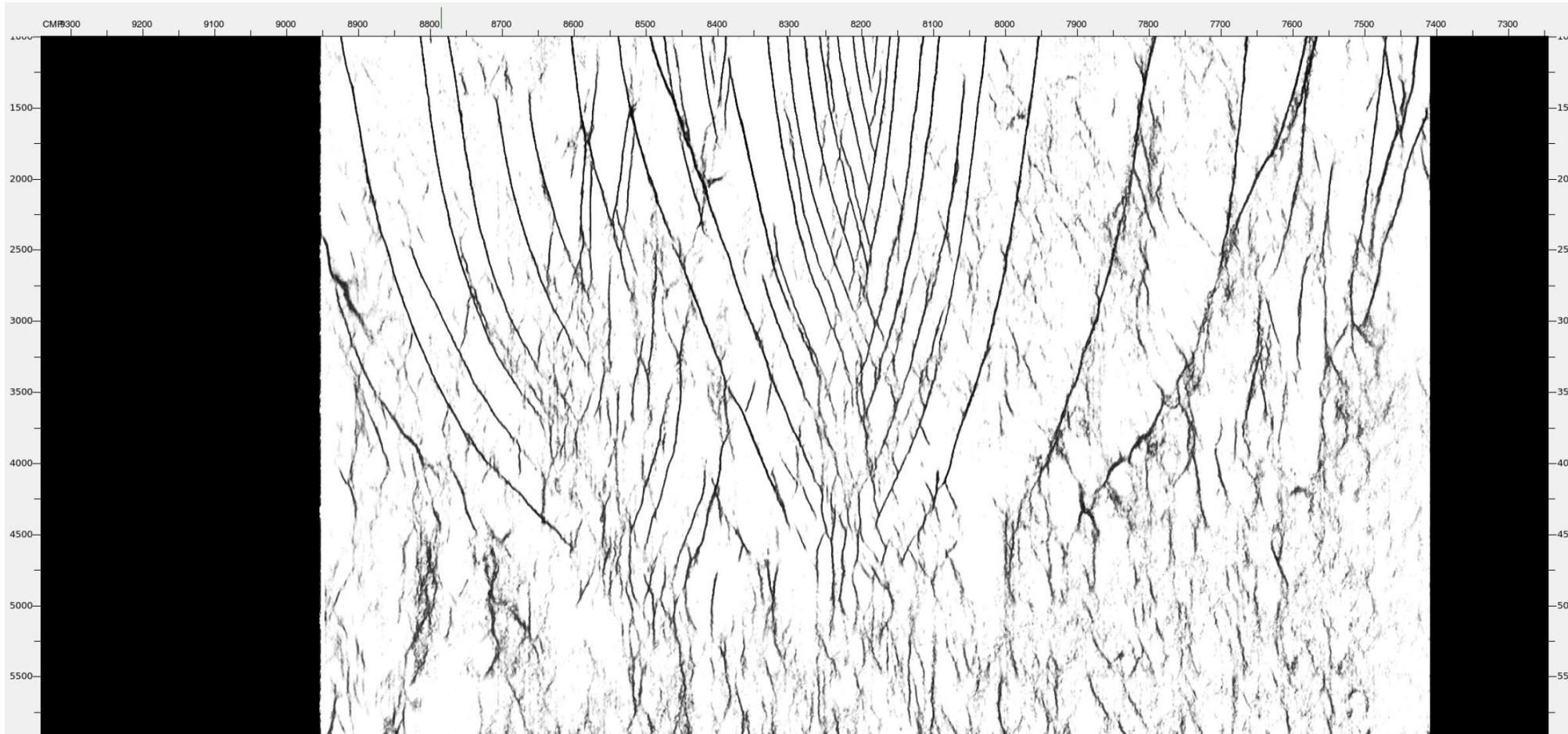
2ms采样

使用技巧：  
选取500线\*500道  
数据进行测试  
至少选择200线  
\*200道



# 1.3D合成数据

## 采样率的影响



4ms采样

使用技巧：  
选取500线\*500道  
数据进行测试  
至少选择200线  
\*200道

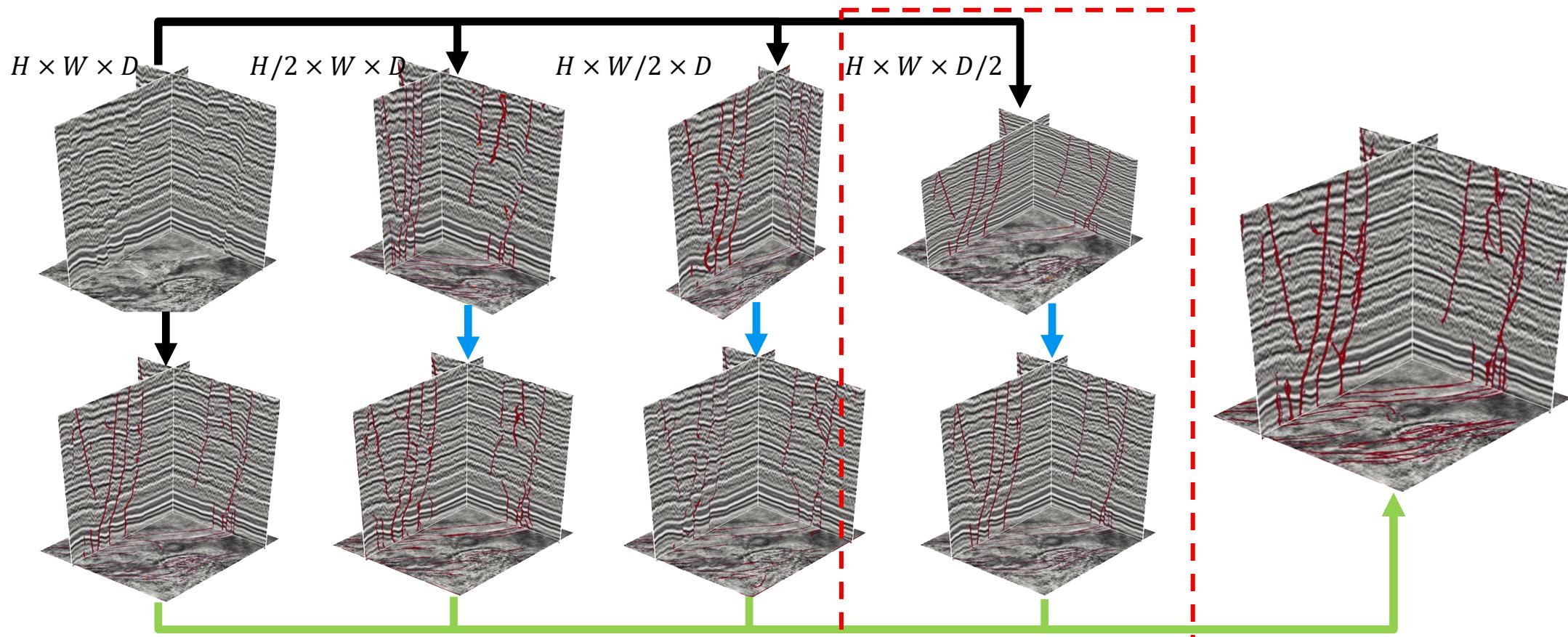


# 1. 三维多尺度

方法原理



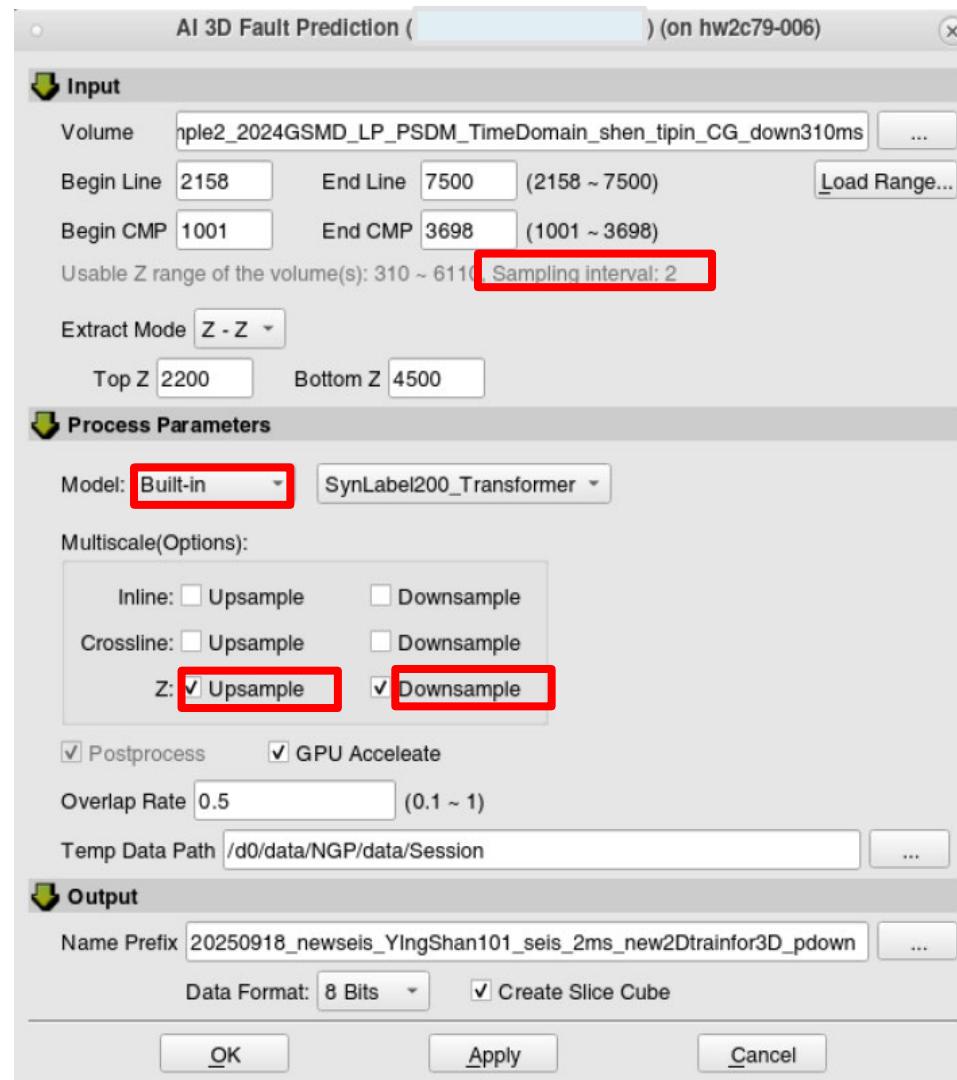
针对目前基于3D合成数据的智能断层预测模块，只对单一尺度进行分析，预测结果受采样率影响严重，研究了**多尺度预测&融合**的预测流程。





# 1.三维多尺度

## 界面及主要参数



输入:

**Seismic Data :** 预测的地震数据,

程序内部将数据切成128\*128\*128的大小进行预测

主要参数:

**Model:** Built-in 选择内置网络模型

**MultiScale (Options)**

**Z: Upsample (采样率减小一倍)**

**Downsample(采样率增大一倍)**

**Postprocess:** 连续性及细线化处理

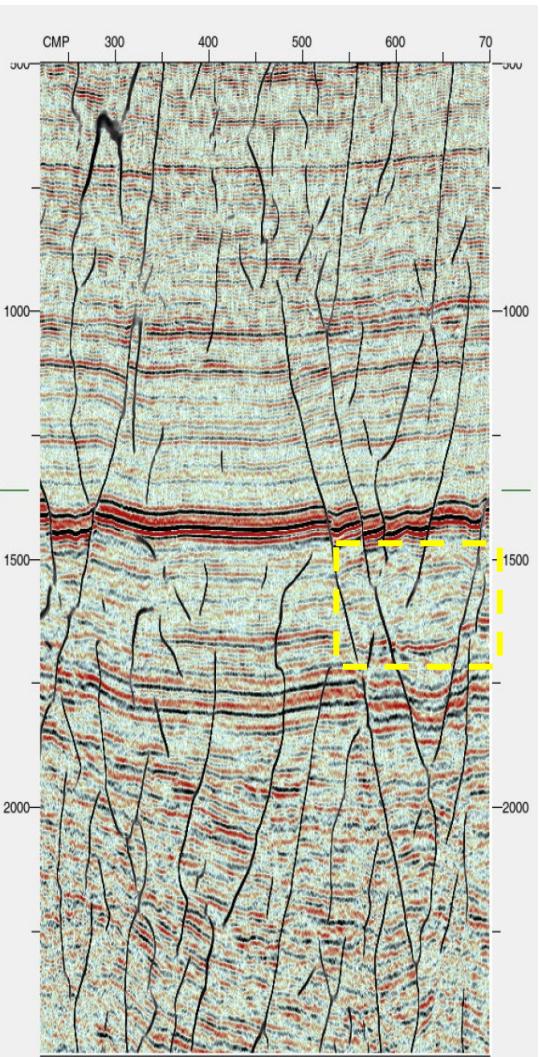
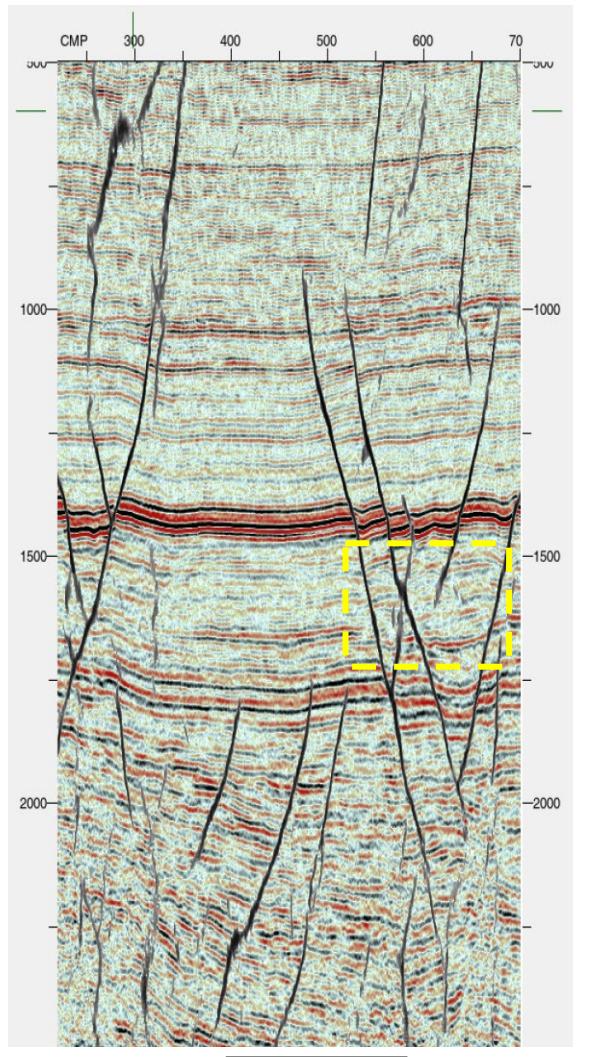
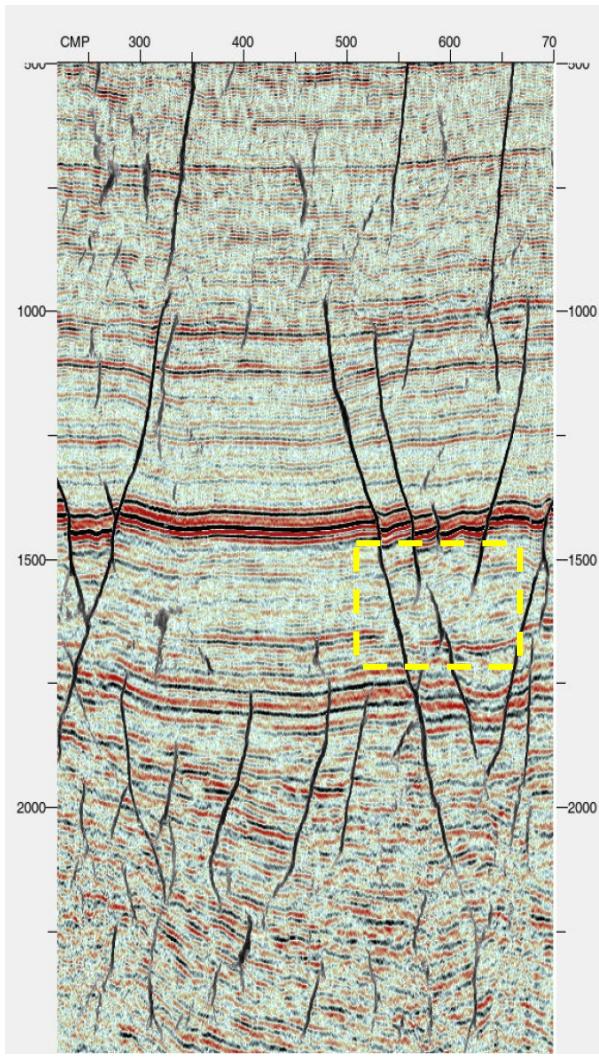
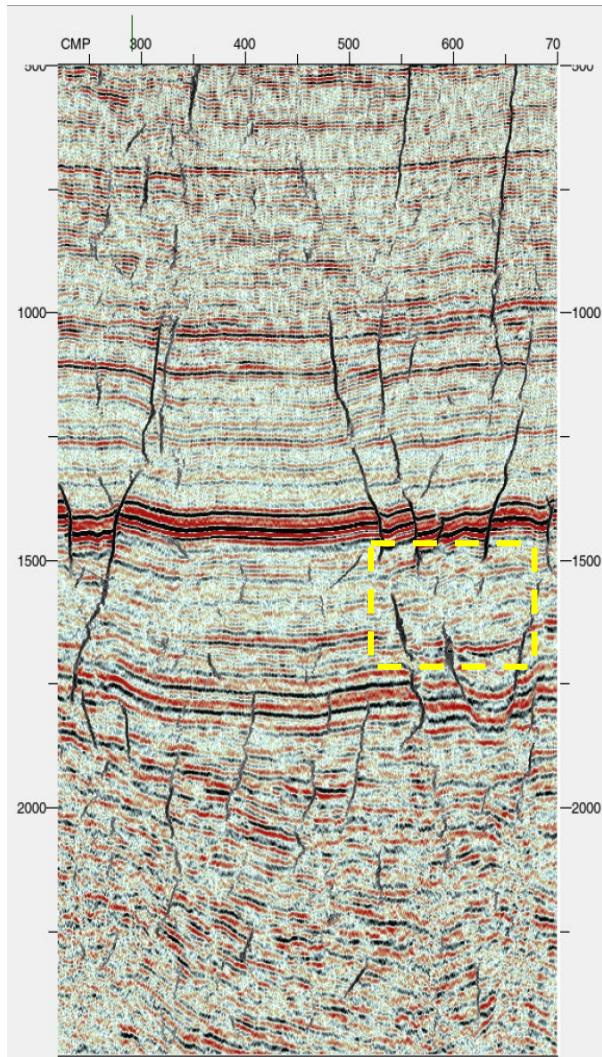
**GPU Accelerate :** 是否采用GPU加速训练

**Overlap:** 预测过程块体尺寸重叠率



# 1. 三维多尺度

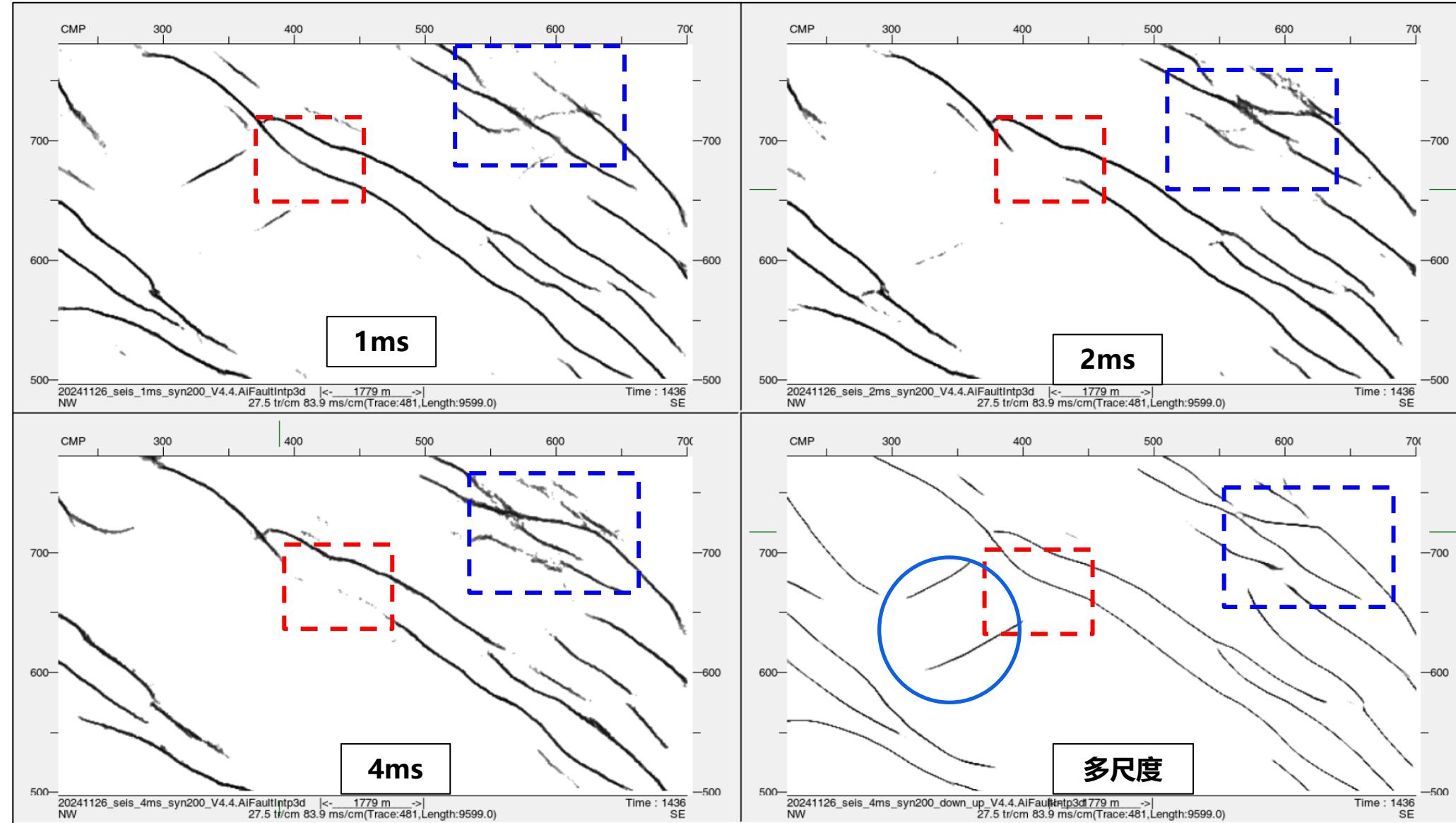
应用实例-东部某工区





# 1.三维多尺度

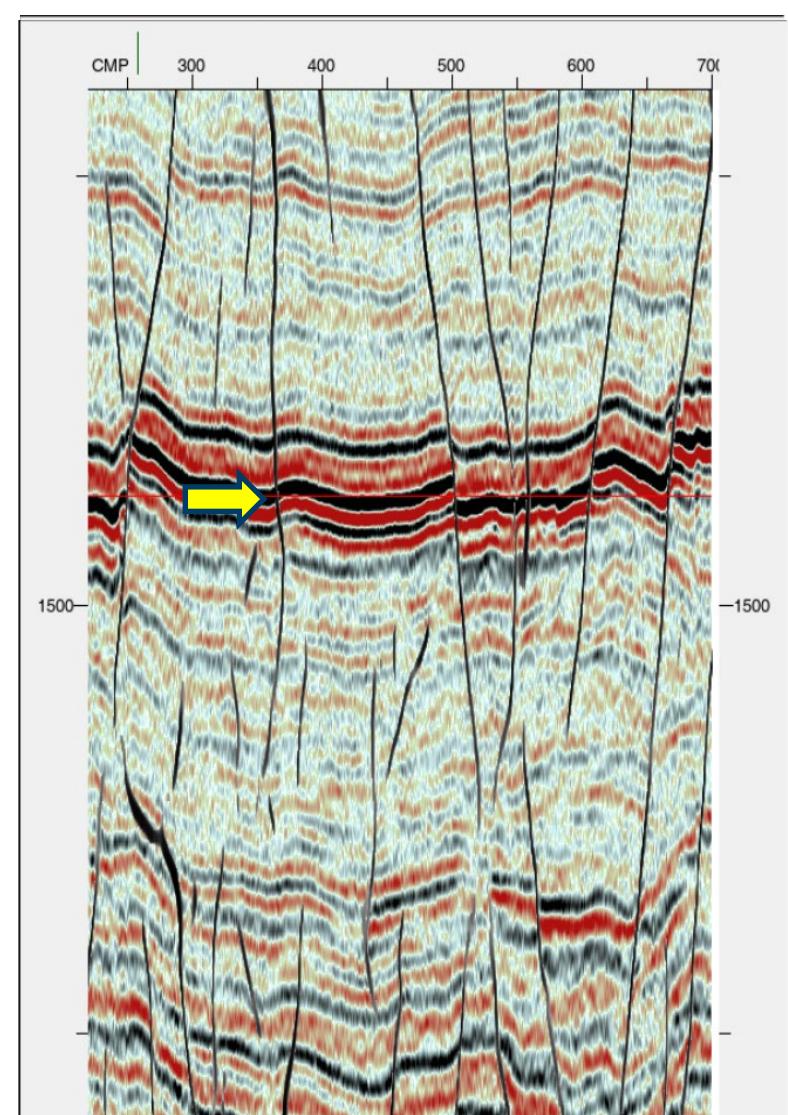
## 应用实例-东部某工区



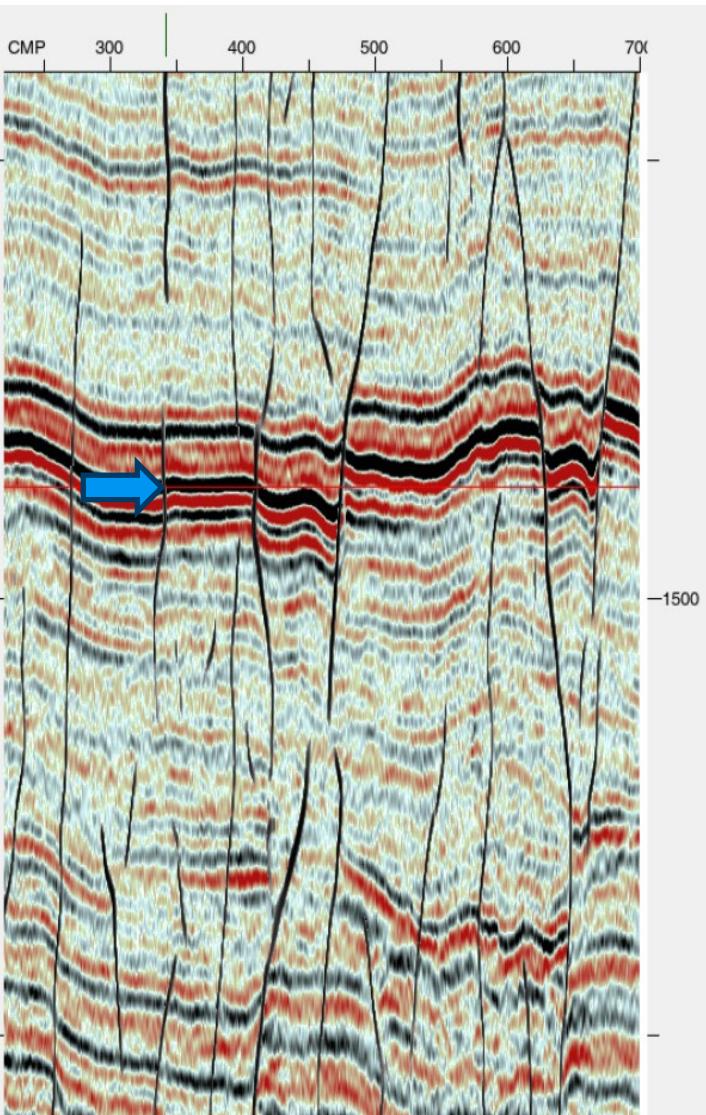


# 1. 三维多尺度

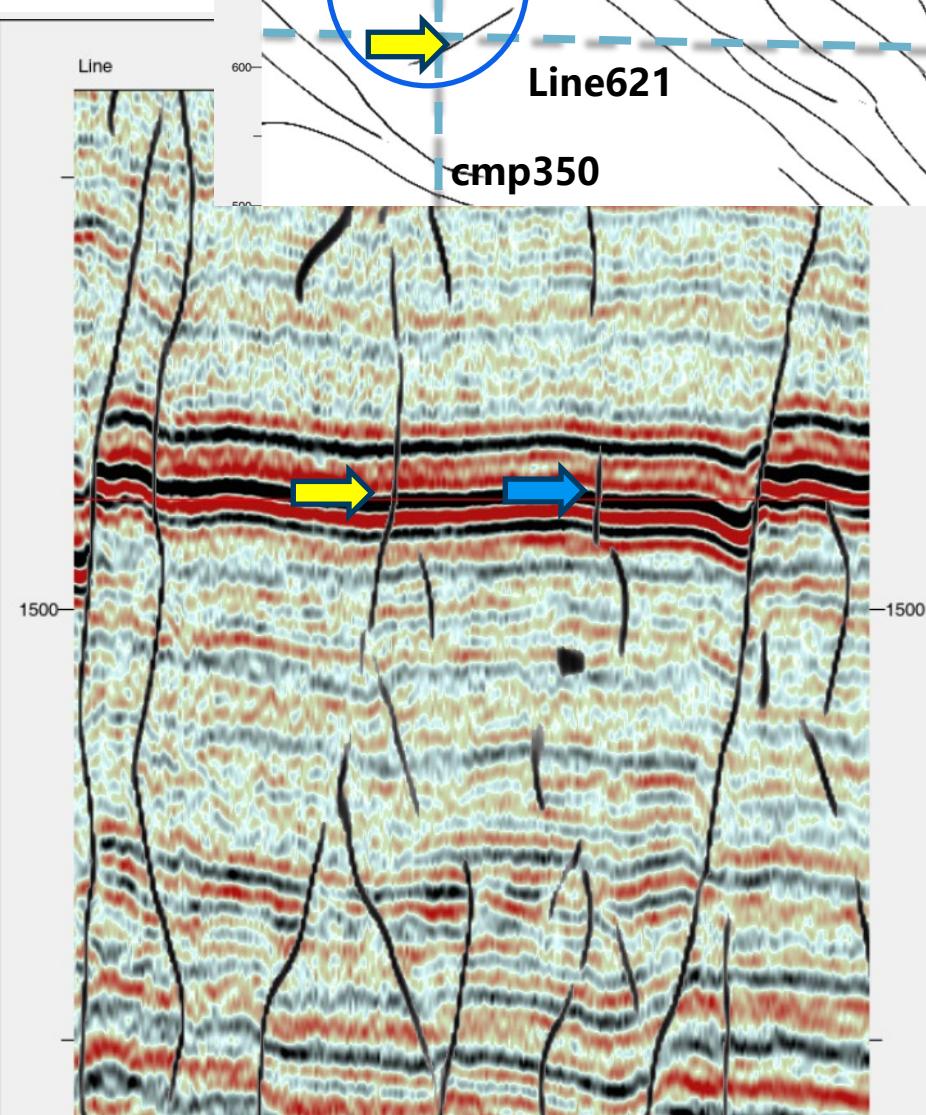
## 应用实例-东部某工区



Line621



Line678



crossline350



## CONTENES 目录

01

02

03

研究背景

## 三维智能断层模块功能介绍

三维合成数据断层预测

个性化三维稀疏标签断层预测

个性化三维标签断层预测

注意事项及流程建议

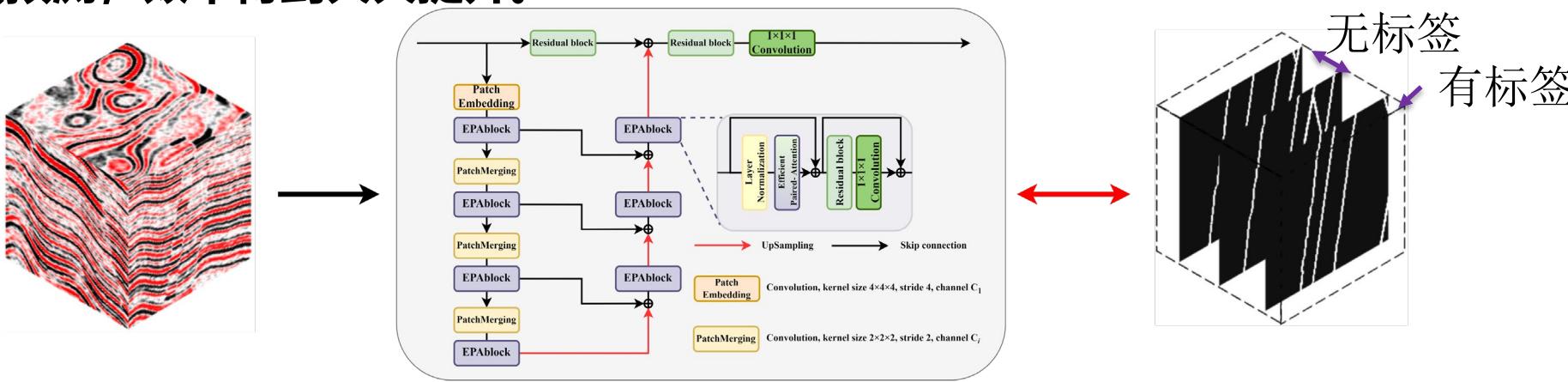


## 2. 个性化三维稀疏标签

方法原理



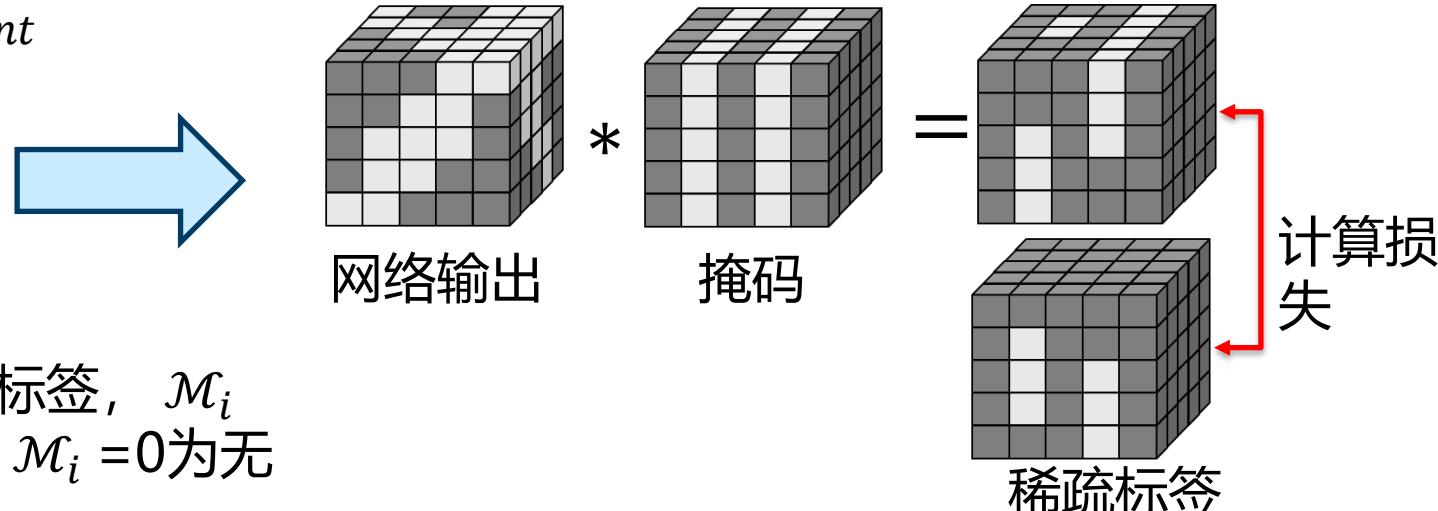
针对个性化3D标签生成程复杂，采用自注意力机制，用稀疏三维标签（伪3D标签）进行方式进行断层预测，效率得到大大提升。



$$\begin{aligned}\mathcal{L}_{md}(\hat{y}, y) &= 1 - \mathcal{M} * DiceCoefficient \\ &= 1 - \frac{\sum_i^n \mathcal{M}_i \hat{y}_i y_i}{\sum_i^n \mathcal{M}_i (\frac{1}{2} \hat{y}_i + \frac{1}{2} y_i)}\end{aligned}$$

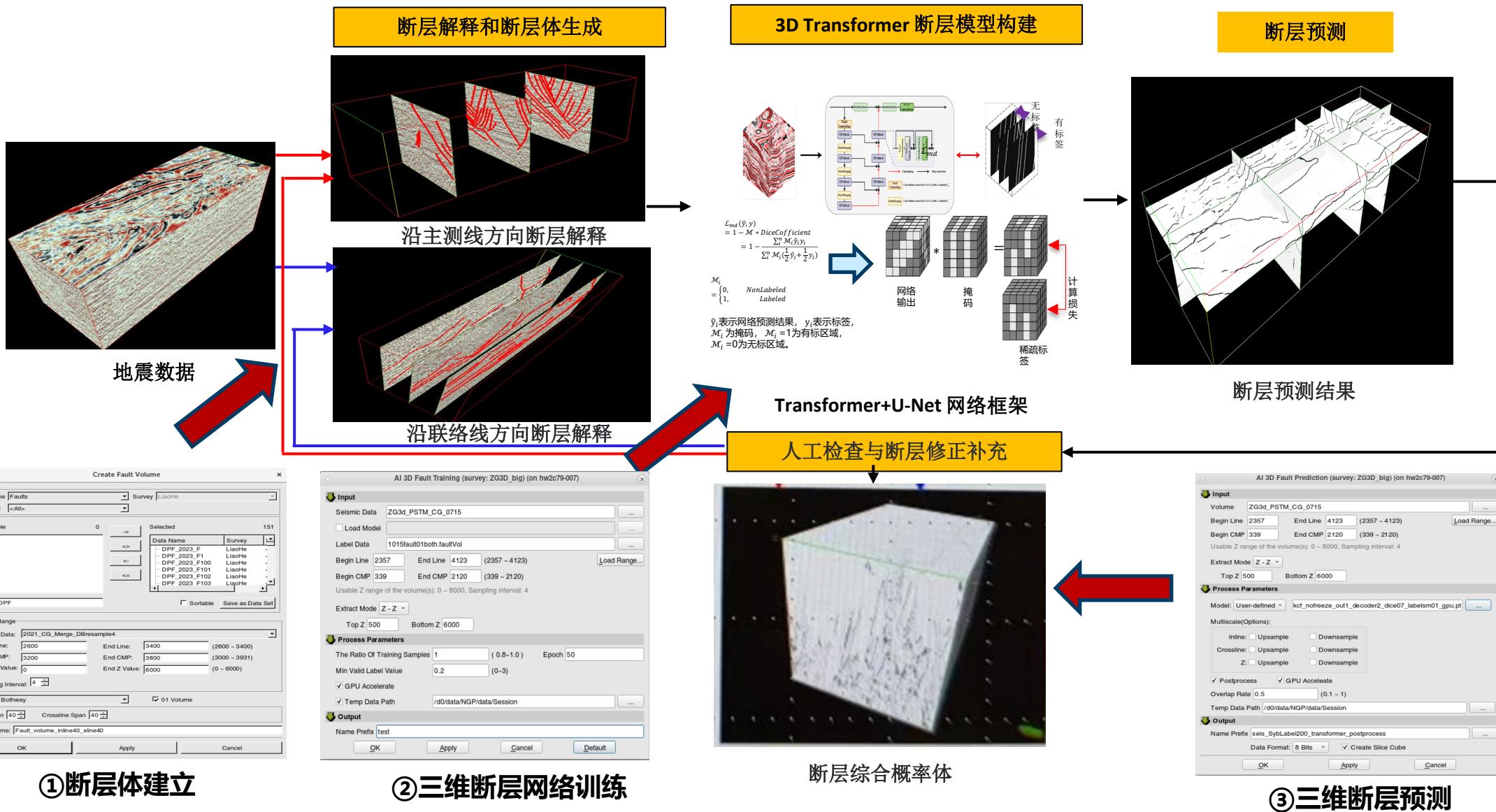
$$\mathcal{M}_i = \begin{cases} 0, & NonLabeled \\ 1, & Labeled \end{cases}$$

$\hat{y}_i$  表示网络预测结果， $y_i$  表示标签， $\mathcal{M}_i$  为掩码， $\mathcal{M}_i = 1$  为有标区域， $\mathcal{M}_i = 0$  为无标区域。





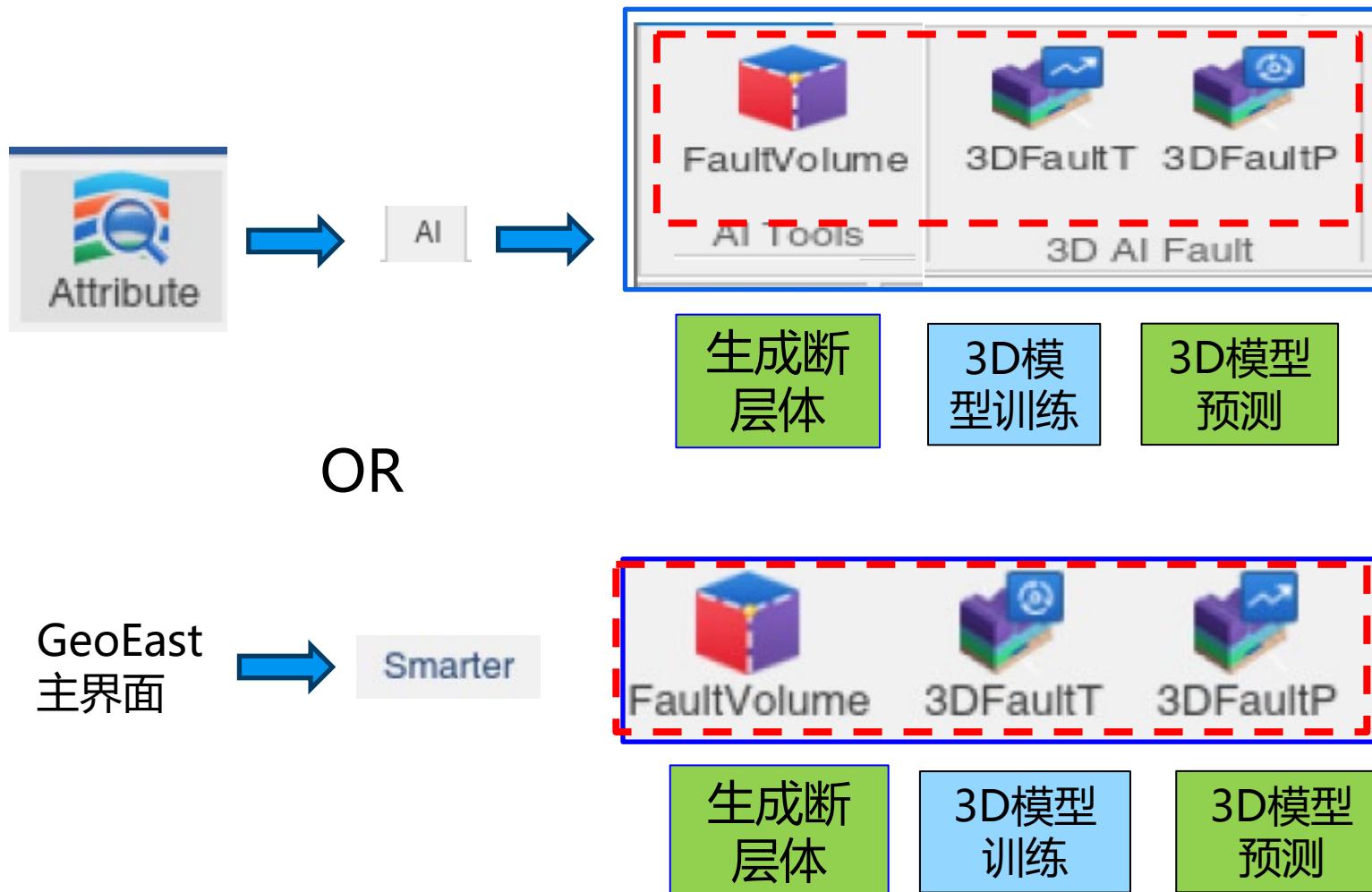
# 个性化三维稀疏标签智能断层预测- 应用流程





## 2. 个性化三维稀疏标签

启动位置



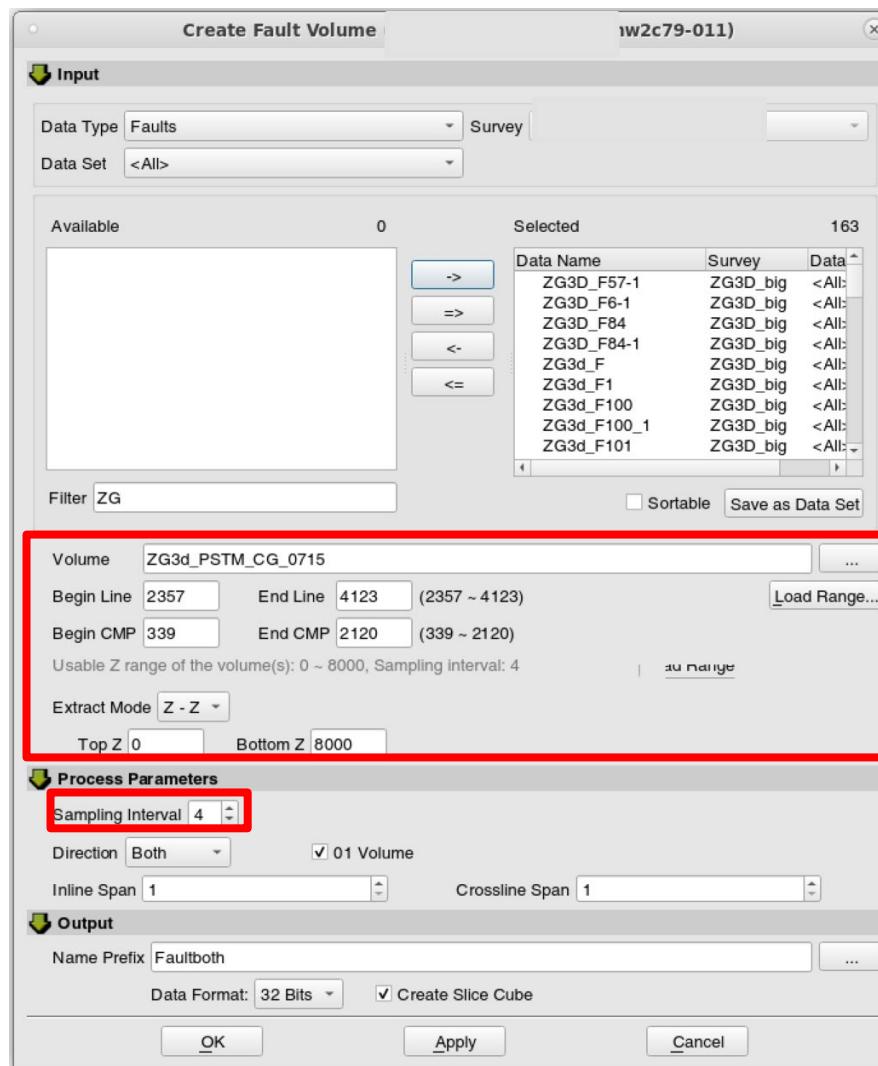


## 2.三维稀疏标签-①生成断层体

界面参数



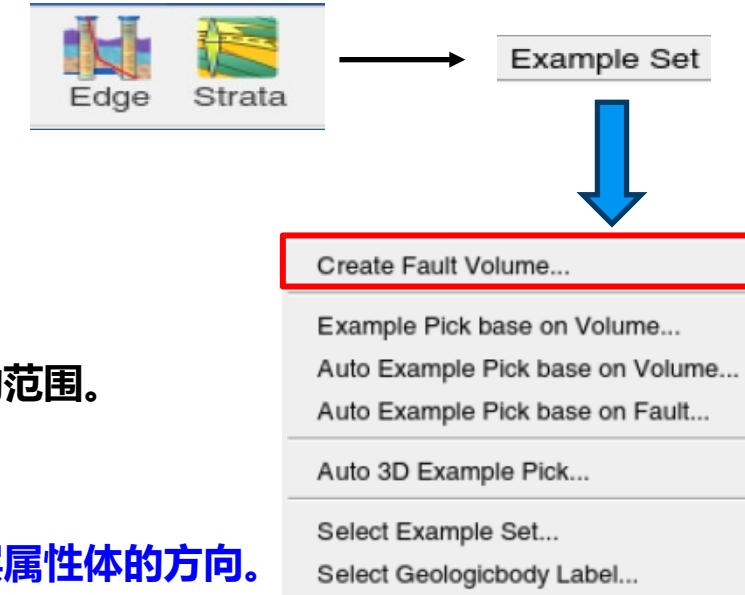
将断层解释方案生成断层体，用于后续的样本建立。



**Selected:**

选择用于生成断层体的断层。

也可以从这里提取



**OutPut Range:**

利用已知数据体控制生成的断层体的范围。

**Direction:**

按照断层解释方案方向选取创建断层属性体的方向。

**Sampling Interval:** 生成断层体的采样间隔 (建议4ms)

**Inline Span:** 生成断层体的Inline间隔

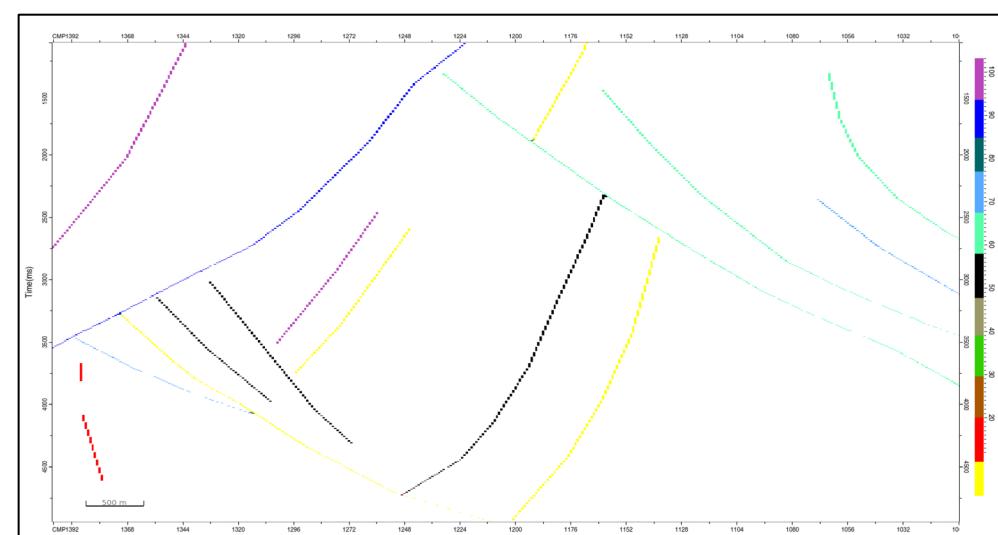
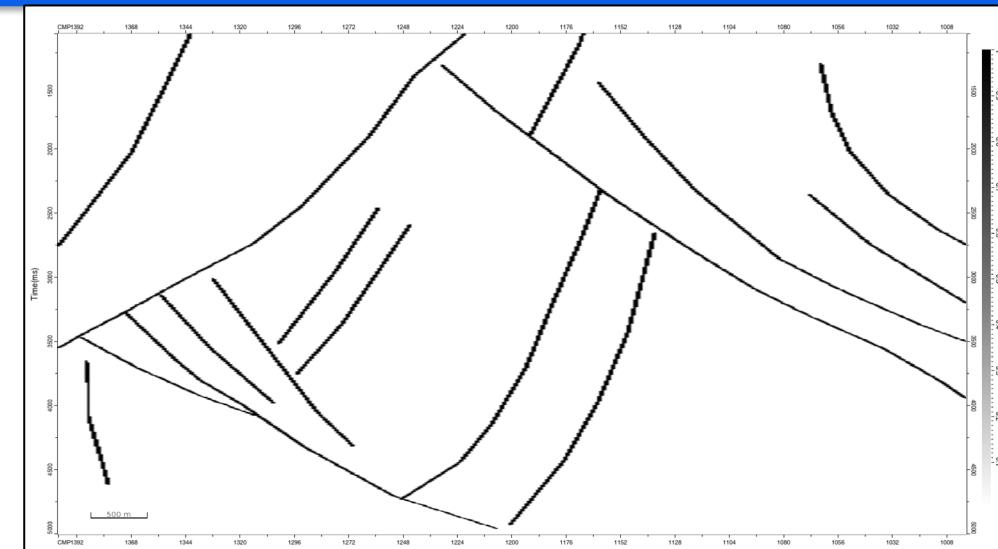
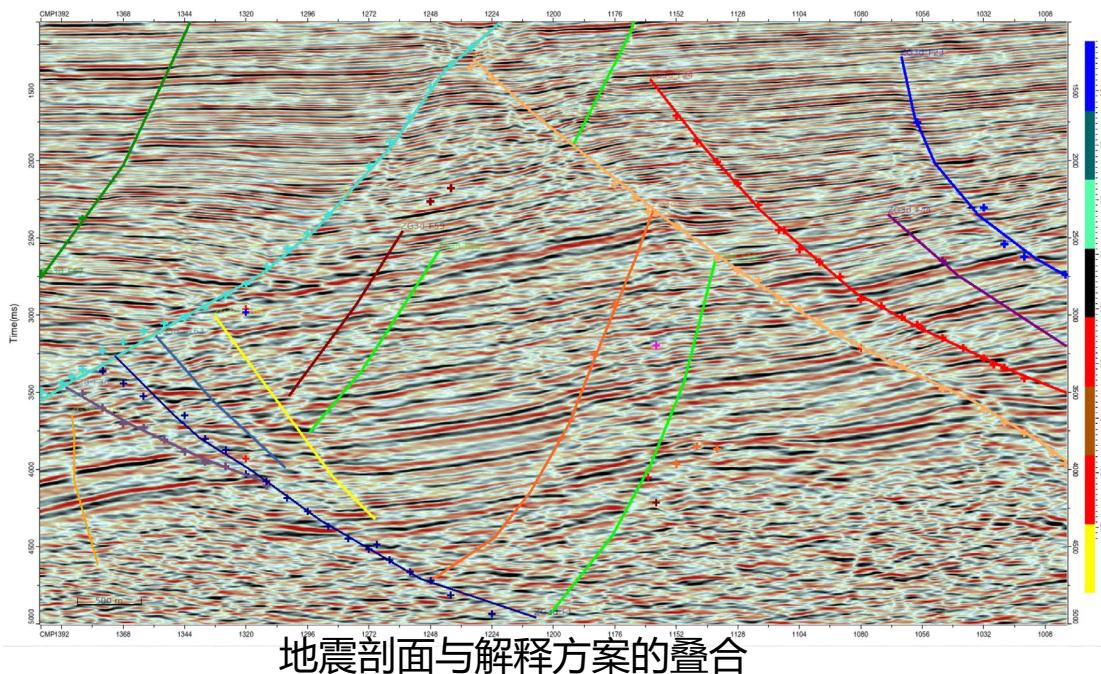
**Crossline Span:** 生成断层体的Crossline间隔

**01体:**

必须勾选01体，生成的断层属性体所有断层赋值为1背景为0。



## 2. 三维稀疏标签-①生成断层体

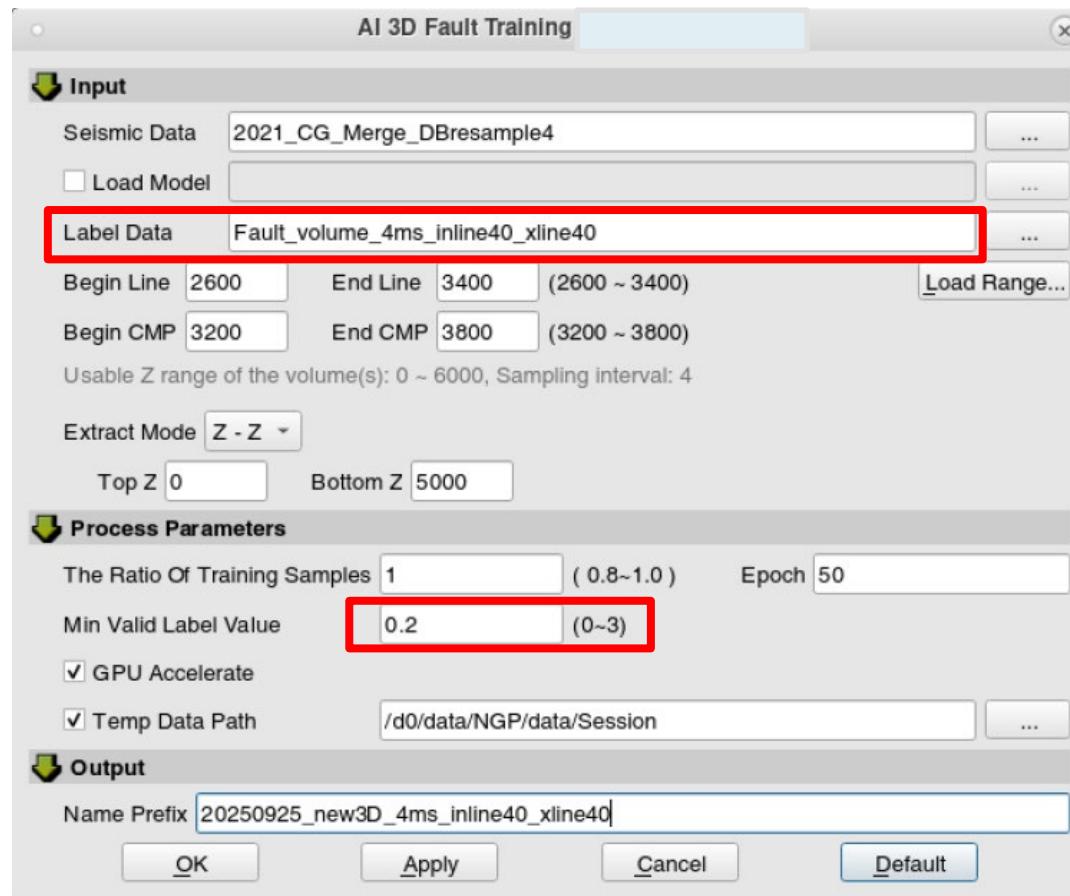


生成的断层体剖面 (不选择01体)



## 2.三维稀疏标签-②训练

界面参数



输入：

**Seismic Data :** 输入地震数据

**Load Model:** 增量学习（若迭代次数50次效果不好，可以再进行一轮学习）

**Label Data :** 生成的断层体（生成01体）

内部自动切割地震与断层标签组成样本，大小128\*128\*128

训练生成的模型位于ai\_models/GeoAIFault3d下

处理参数：

**The Ratio of Training Samples:** 训练样本占训练数据集的比例，建议为1

**Epoch:** 迭代次数（缺省50次）

**Min Valid Label Value** 最小有效标签数，对于Label Data，小于该值的数据都为0。

对于解释方案生成的断层体，由于生成的是01体，是断层的地方为1，所以这里为0.2不影响label体。



## 2.三维稀疏标签-②训练

质控窗口



Python Std Out

Extracting 3D patches  
Total training samples are 142, total validation samples are 0  
Dataset generate successfully!!

抽取的3D样本个数，个数越大，训练时间越长

[[32m[2024-12-31 10:40:47 20241231\_new3D\_4ms\_inline40\_xline40\_gpu)][0m][33m(QC\_widget.py 36)][0m: INFO Running by GPU device!  
[[32m[2024-12-31 10:40:47 20241231\_new3D\_4ms\_inline40\_xline40\_gpu)][0m][33m(QC\_widget.py 36)][0m: INFO Utilize transfer learning strategy!  
[[32m[2024-12-31 10:40:50 20241231\_new3D\_4ms\_inline40\_xline40\_gpu)][0m][33m(QC\_widget.py 36)][0m: INFO Current model resolution is 128/128/128!  
[[32m[2024-12-31 10:40:50 20241231\_new3D\_4ms\_inline40\_xline40\_gpu)][0m][33m(QC\_widget.py 36)][0m: INFO Creating model success!  
[[32m[2024-12-31 10:40:52 20241231\_new3D\_4ms\_inline40\_xline40\_gpu)][0m][33m(QC\_widget.py 36)][0m: INFO number of params: 39048505  
[[32m[2024-12-31 10:40:52 20241231\_new3D\_4ms\_inline40\_xline40\_gpu)][0m][33m(QC\_widget.py 40)][0m: INFO Start training

是否用到了GPU

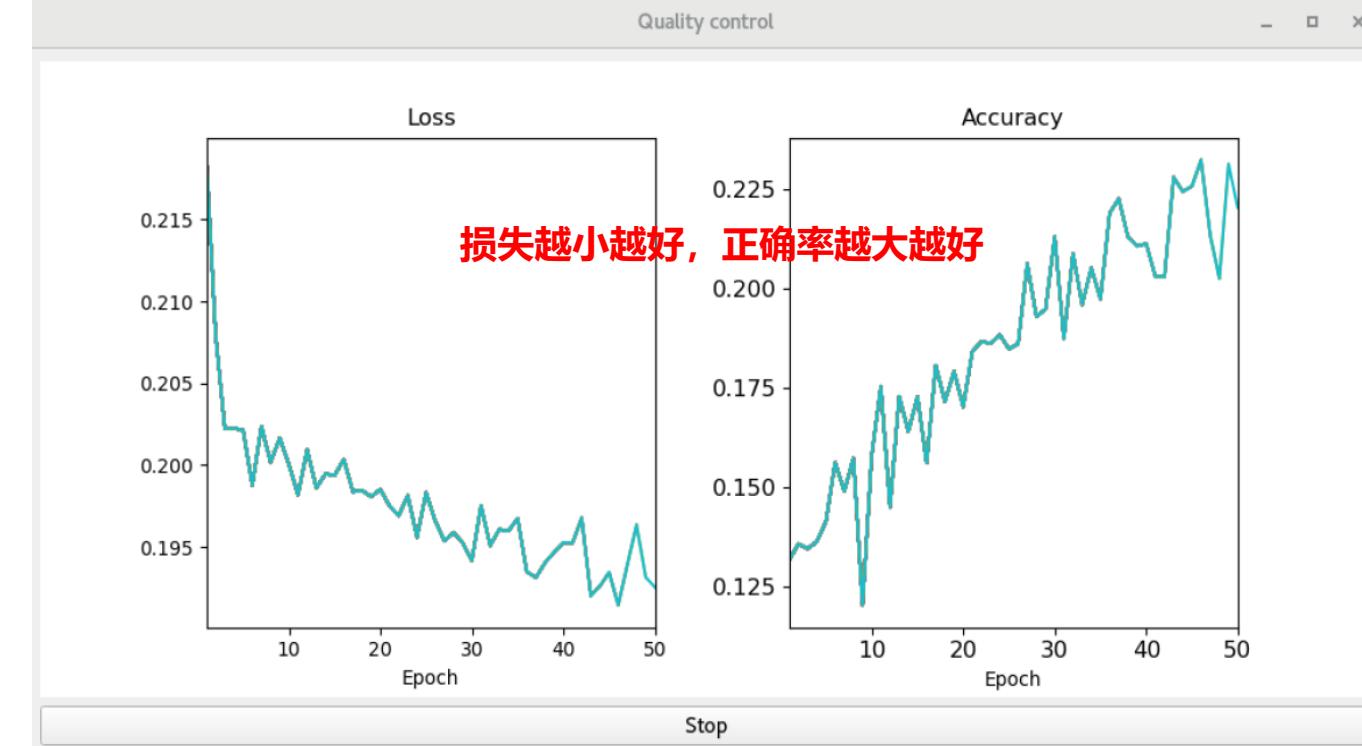
内部运算尺寸大小

Python Std Err

86%|██████████| 43/50 [2:29:27<24:17, 208.16s/it]  
88%|██████████| 44/50 [2:32:54<20:46, 207.82s/it]  
90%|██████████| 45/50 [2:36:22<17:18, 207.77s/it]  
92%|██████████| 46/50 [2:39:52<13:54, 208.56s/it]  
94%|██████████| 47/50 [2:43:19<10:24, 208.15s/it]  
96%|██████████| 48/50 [2:46:46<06:55, 207.82s/it]  
98%|██████████| 49/50 [2:50:14<03:27, 207.72s/it]  
100%|██████████| 50/50 [2:53:41<00:00, 207.56s/it]  
100%|██████████| 50/50 [2:53:41<00:00, 208.43s/it]

信息窗口

Process Information

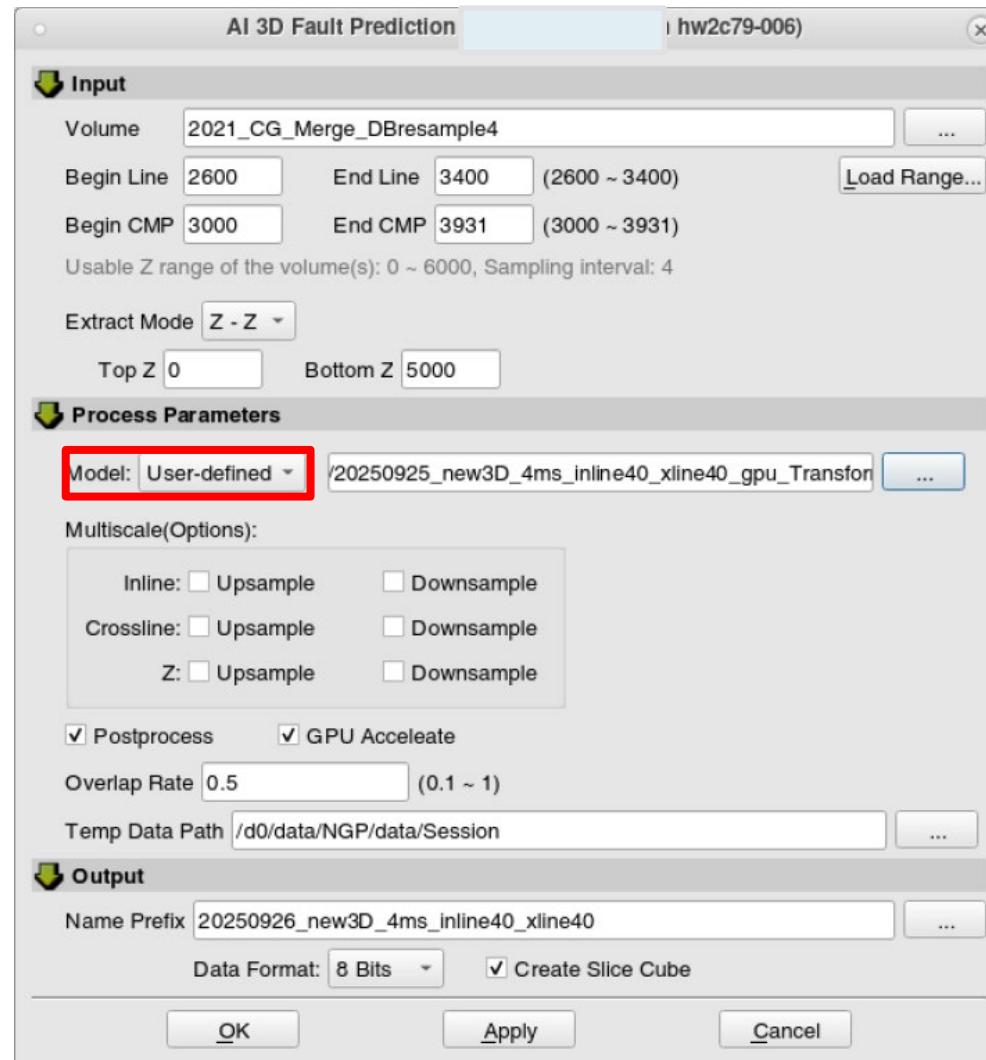


质控窗口



## 2.三维稀疏标签-③预测

界面参数



输入：

**Seismic Data**：预测的地震数据

程序内部将数据切成128\*128\*128的大小进行预测

主要参数：

**Model: User-defined**

选择选择训练好的网络模型，模型位于ai\_models/GeoAIFault3d  
下

**PostProcess**：连续性处理（稀疏标签方法必须进行连续性处理）

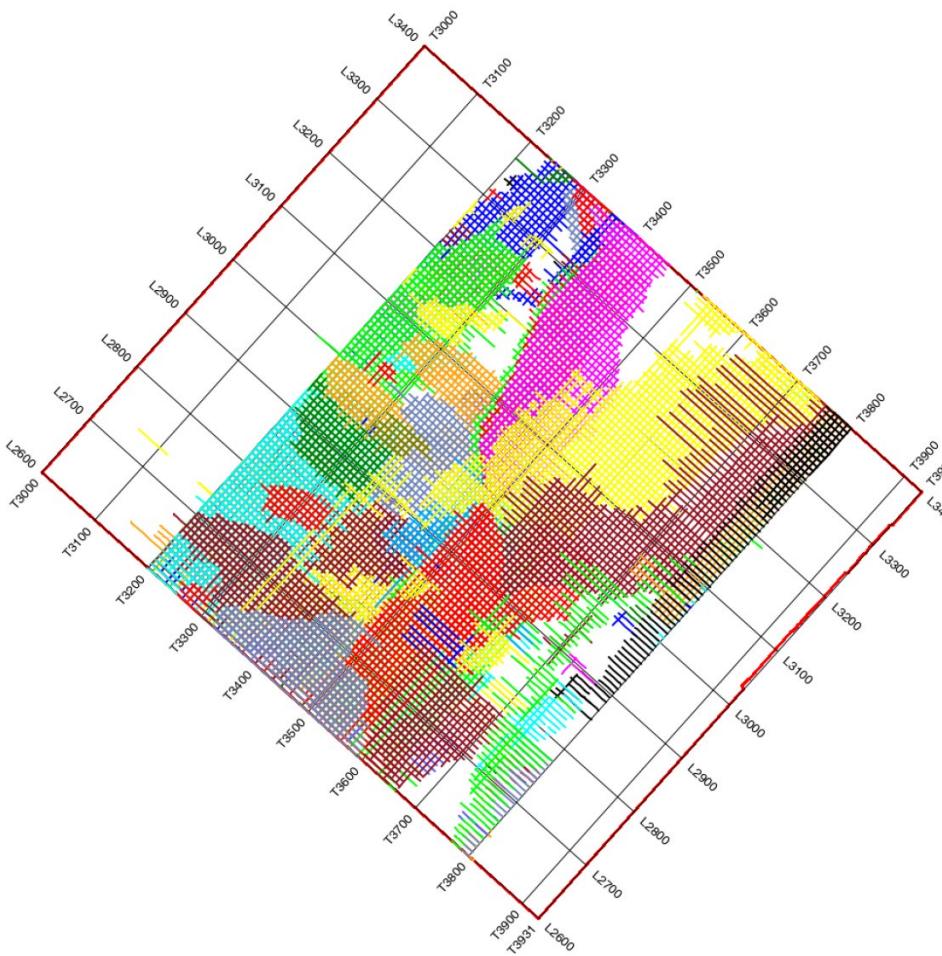
**GPU Acceleration**：是否采用GPU加速训练。

**Overlap**：预测过程的预测尺寸的重叠率

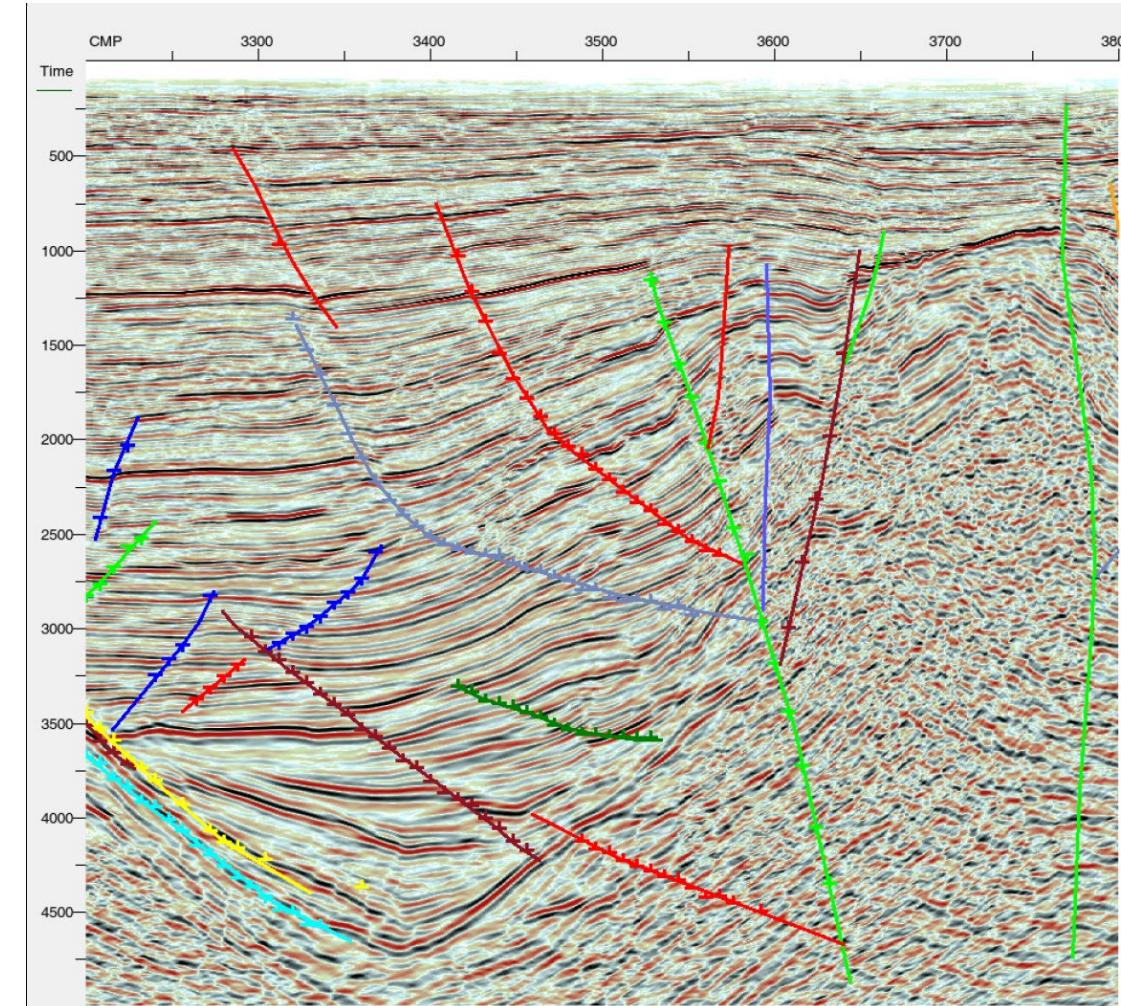


## 2.三维稀疏标签

应用实例-东部某工区



东部工区 (460km<sup>2</sup>)  
解释方案8\*8

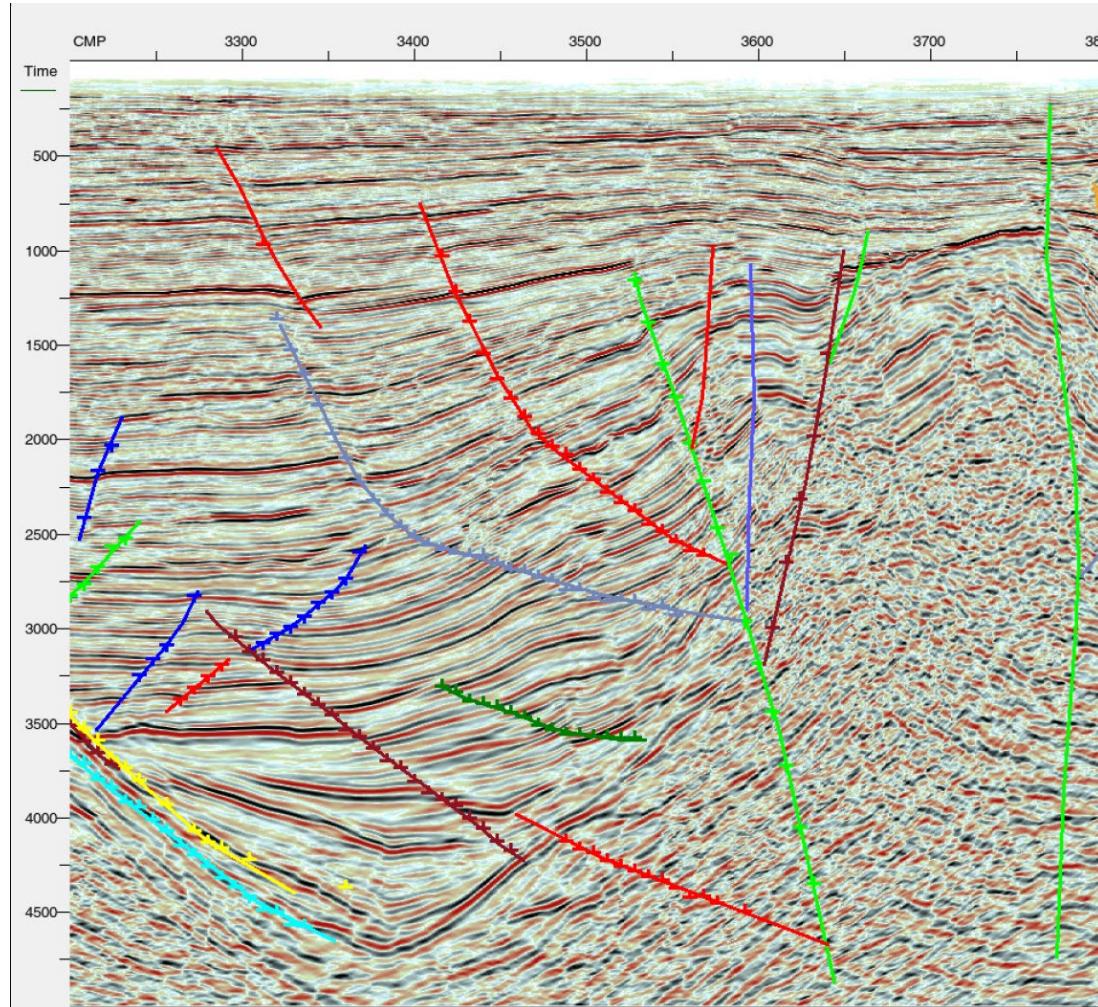


花状构造

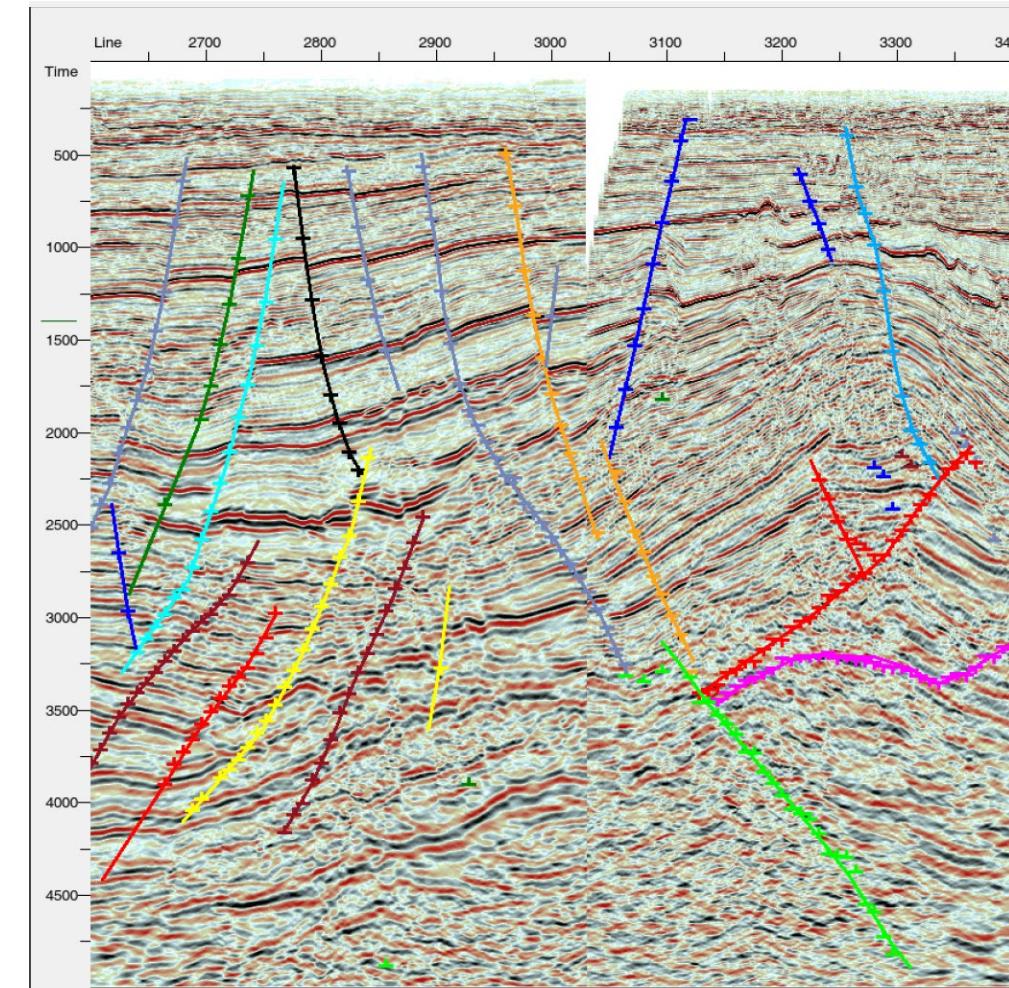


## 2.三维稀疏标签

应用实例-东部某工区



Line方向

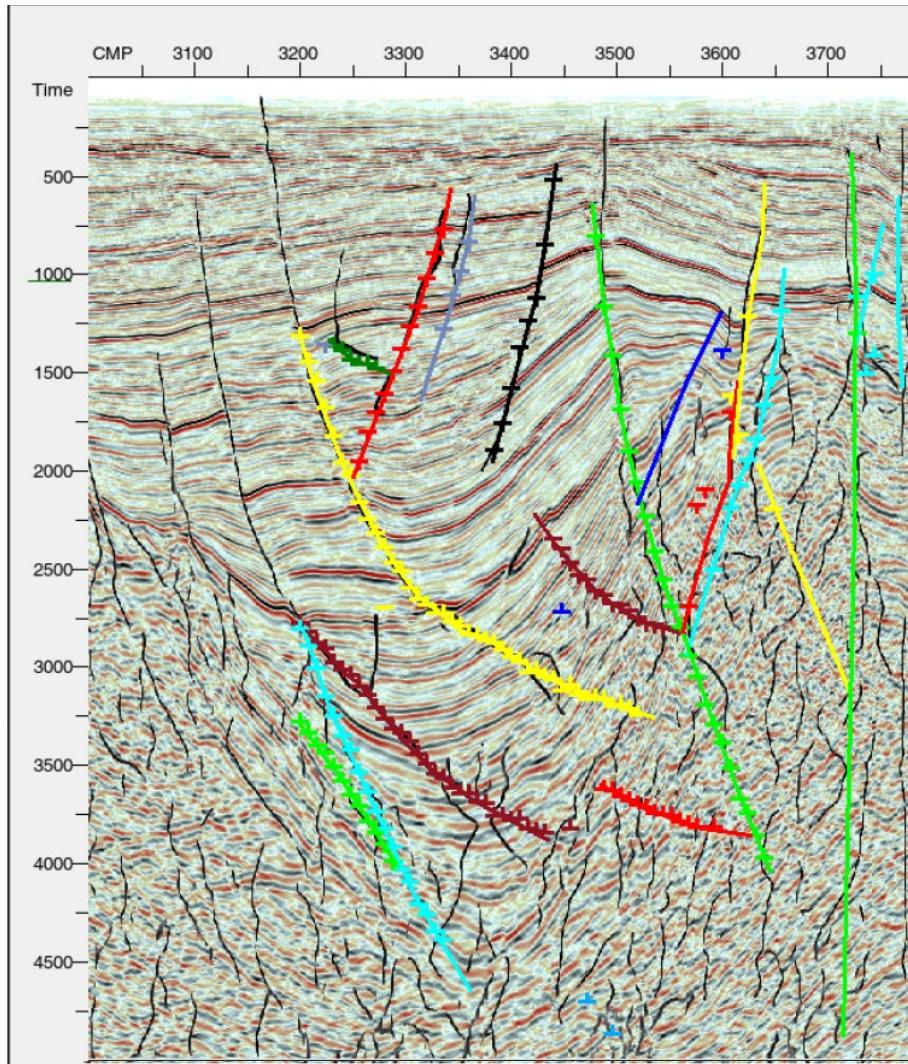


Crossline方向

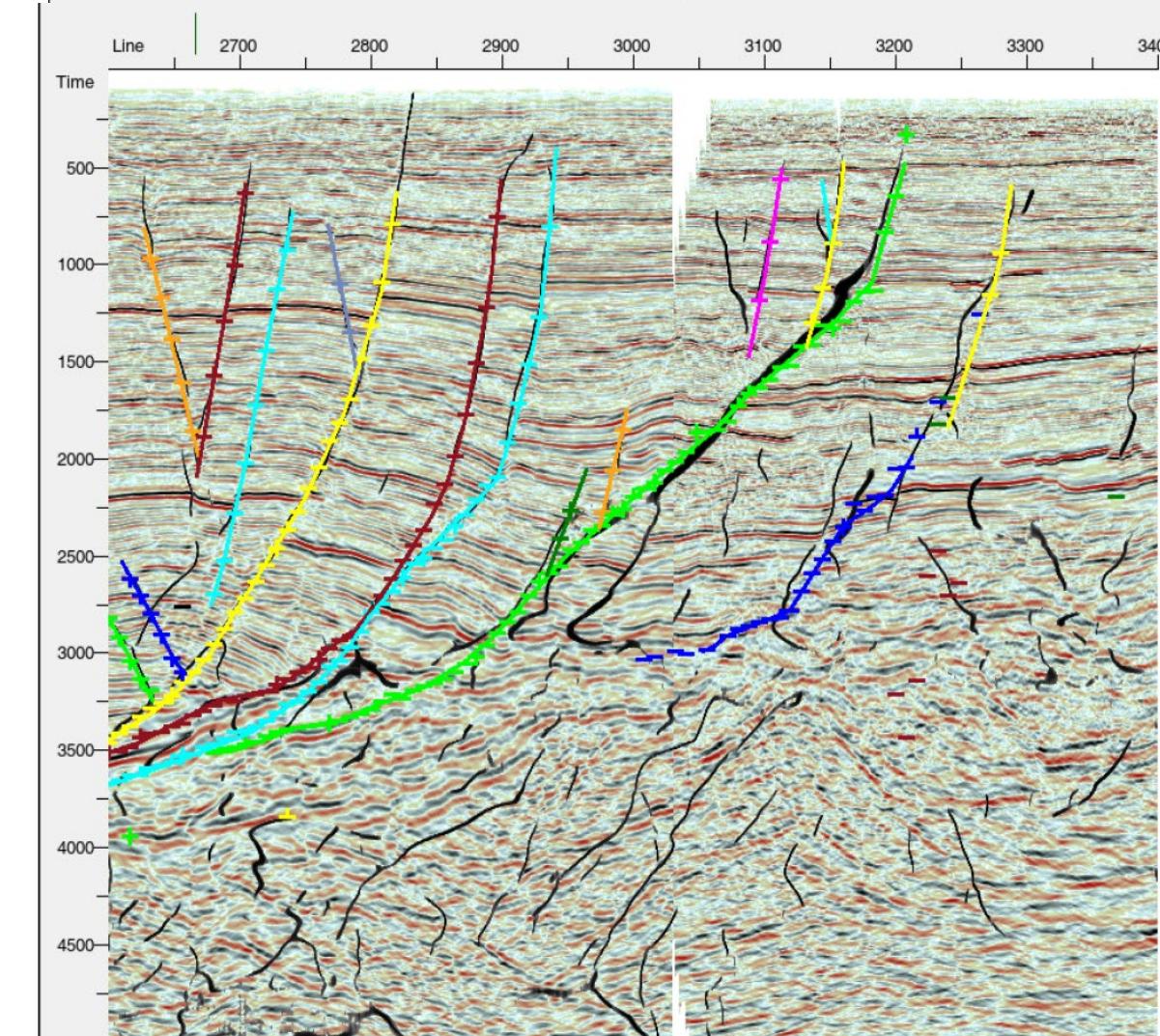


## 2.三维稀疏标签

应用实例-东部某工区



Line方向



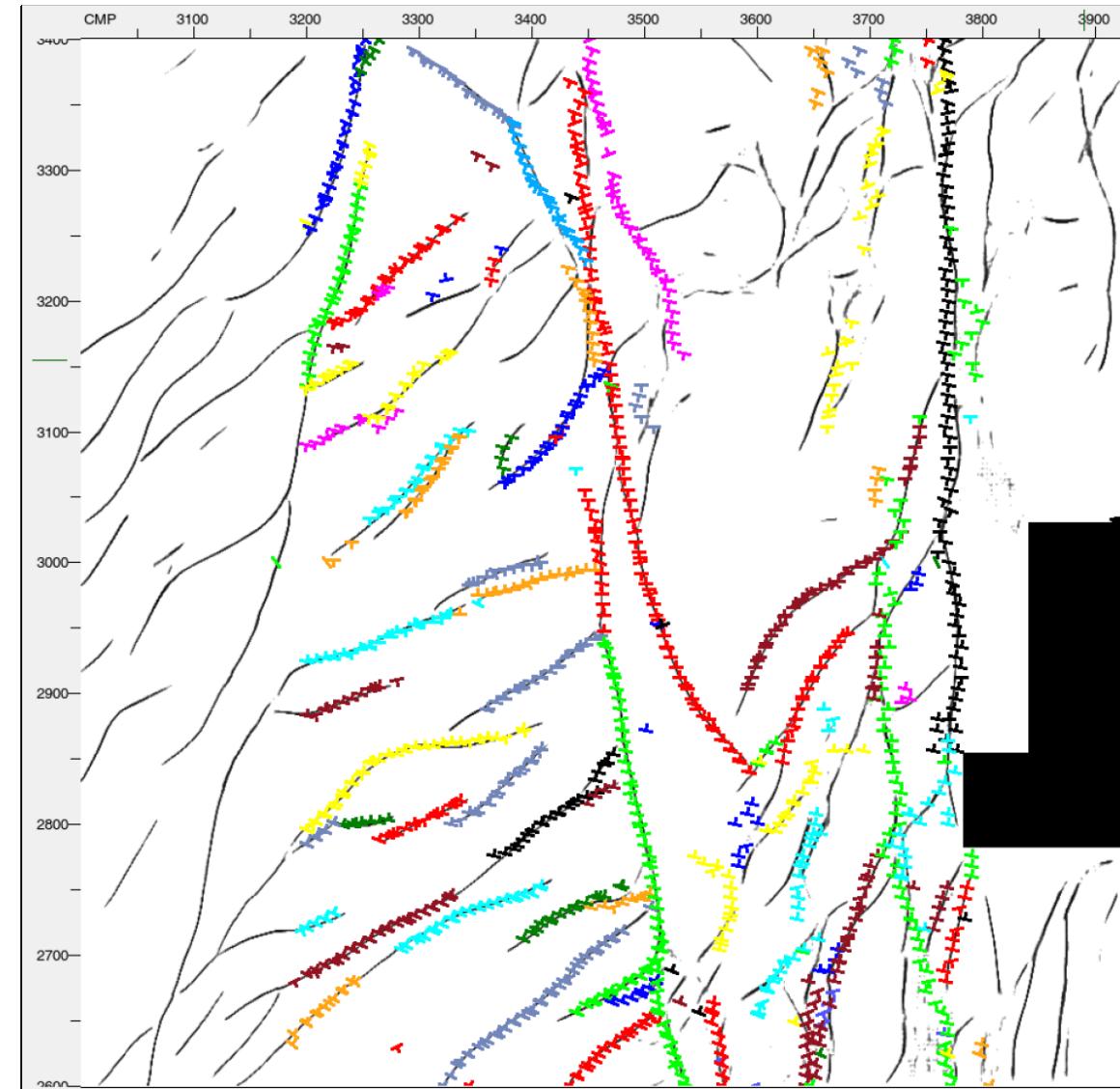
解释方案40\*40

Crossline方向



## 2.三维稀疏标签

应用实例-东部某工区



时间切片 1400ms      解释方案 40\*40

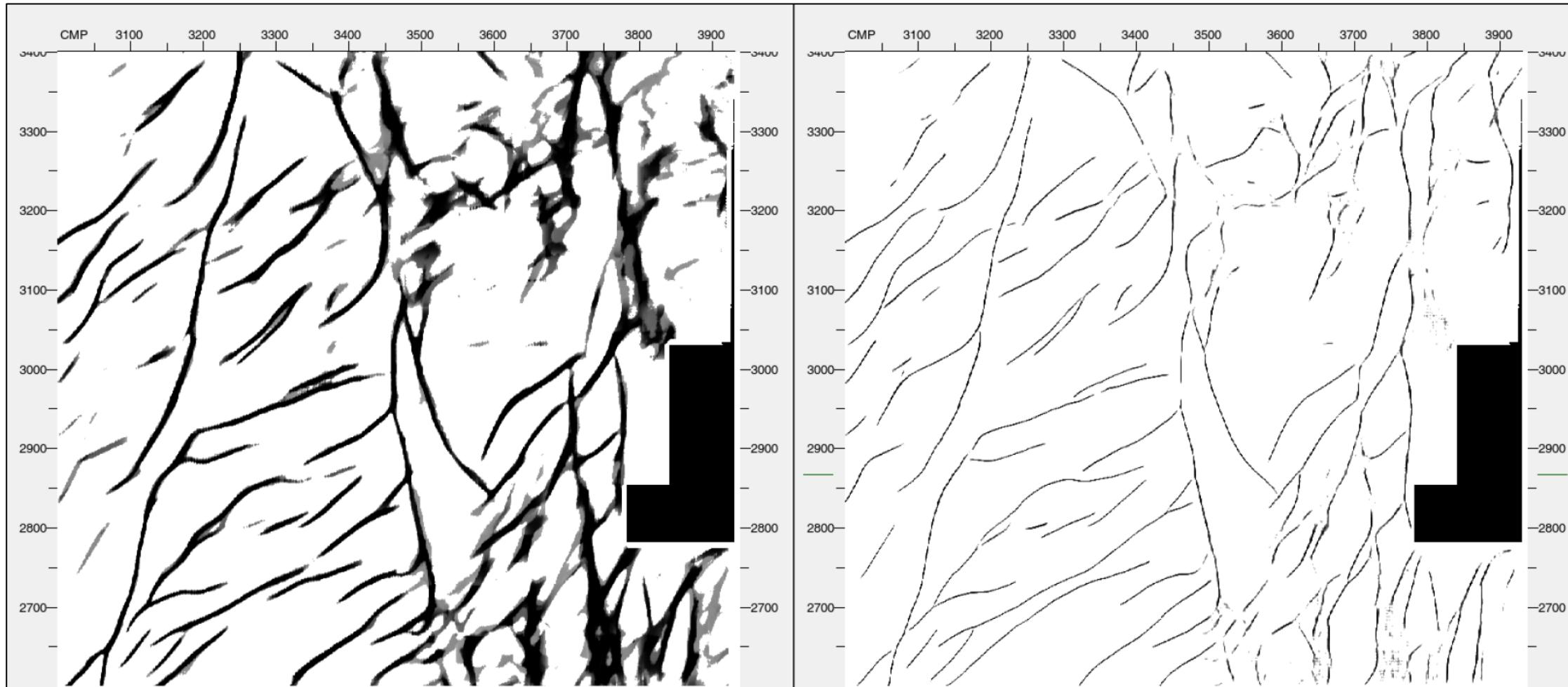


## 2.三维稀疏标签

应用实例-东部某工区



时间切片 1400ms



没有进行后处理结果

进行后处理结果



## CONTENES 目录

01

02

03

研究背景

## 三维智能断层模块功能介绍

三维合成数据断层预测

个性化三维稀疏标签断层预测

个性化三维标签断层预测

注意事项及应用流程

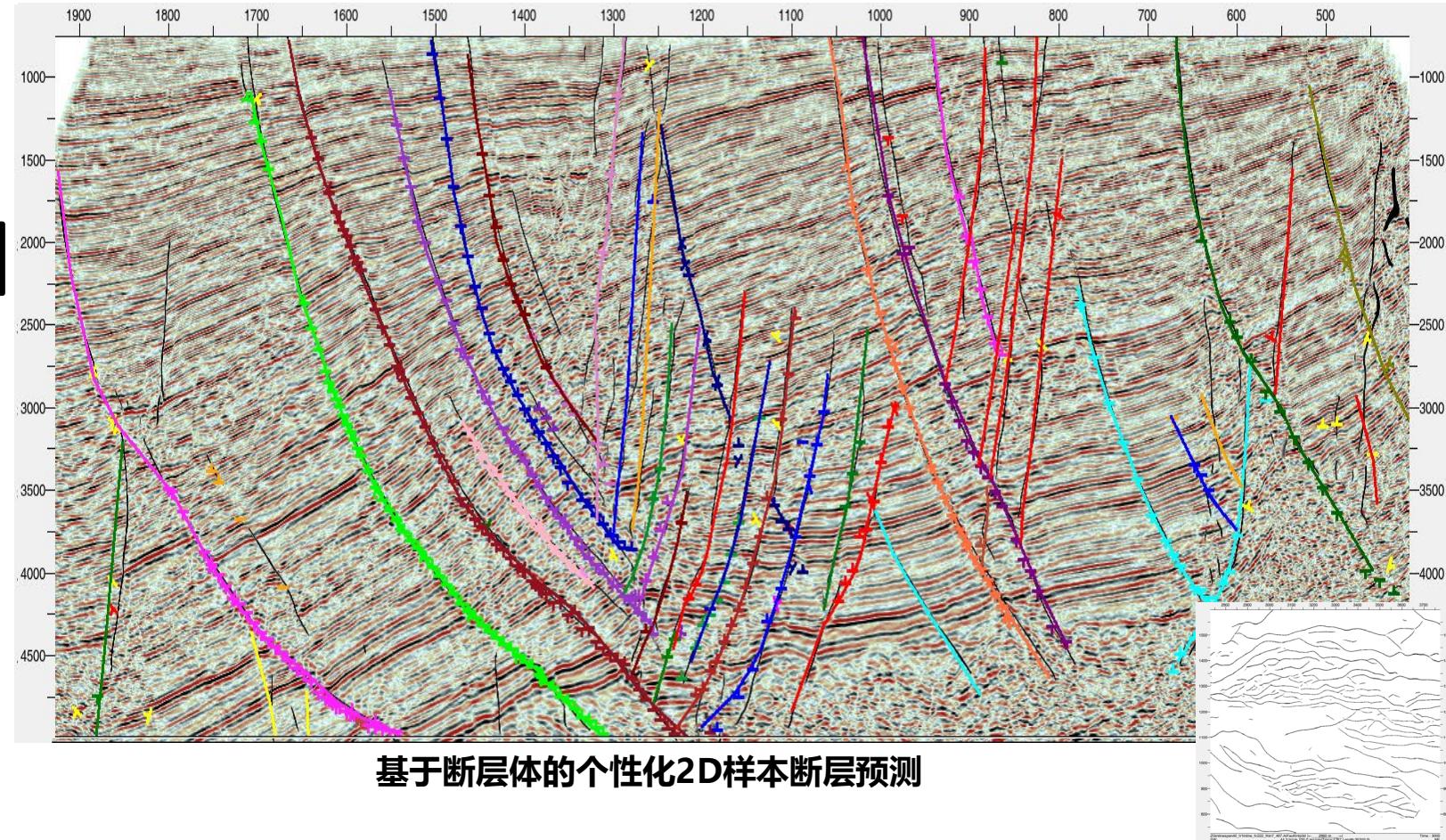


### 3.个性化3D标签

工作流程



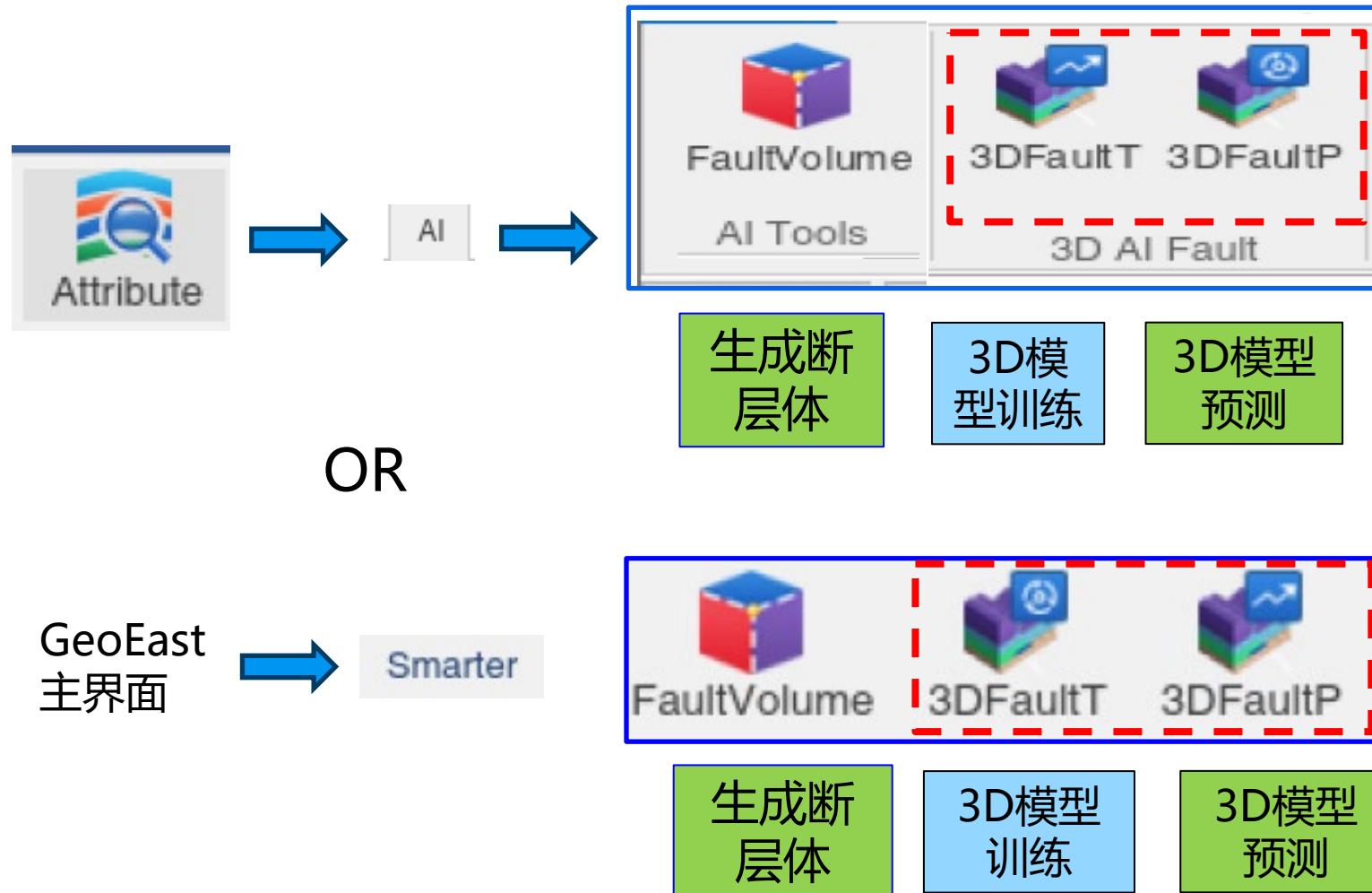
个性化3D标签建立、训练及预测一体化软件及流程，进一步提升智能预测适应性。





## 2.个性化3D训练样本

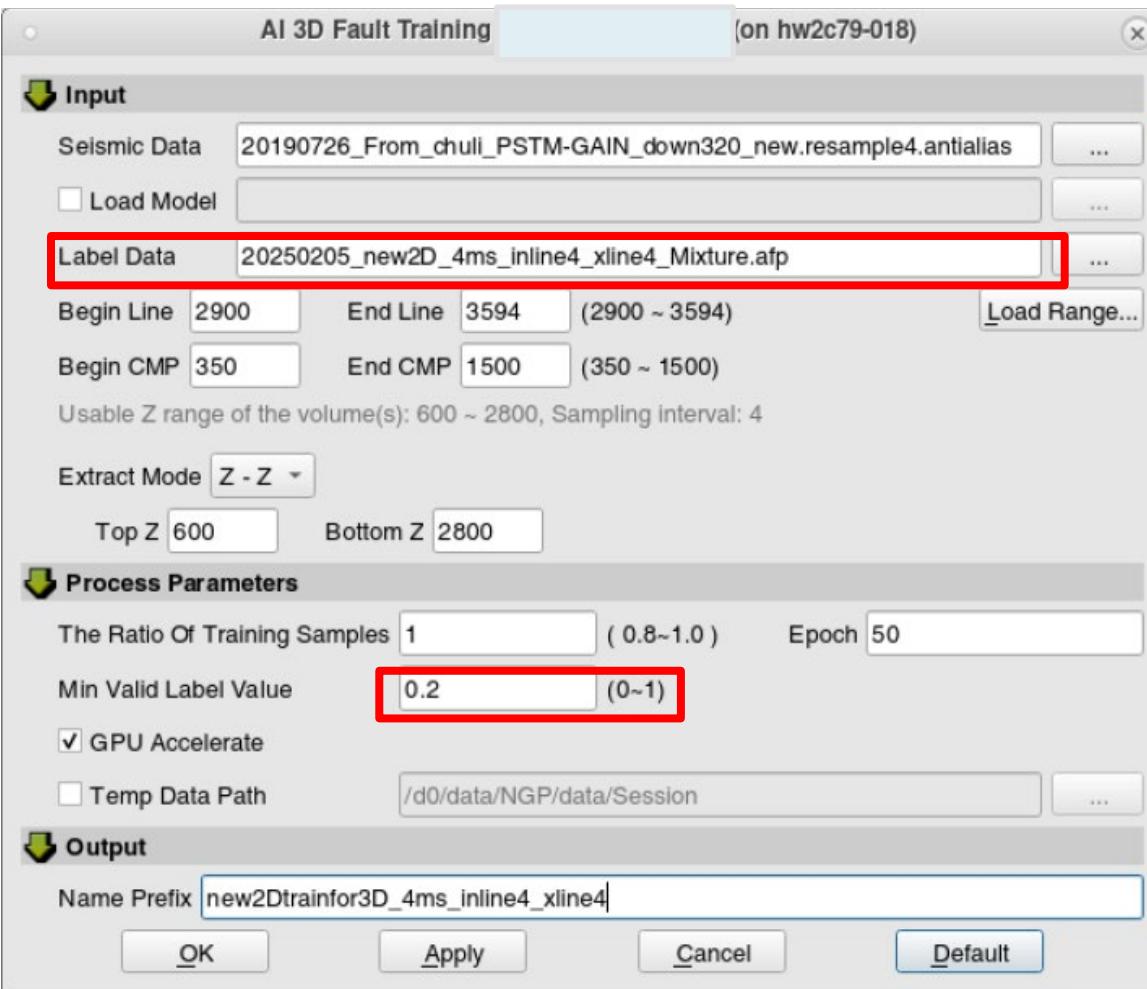
启动位置





### 3.个性化3D标签- ①训练

界面参数



输入:

**Seismic Data :** 输入地震数据

**Label Data :** 个性化2D标签概率体... (0-1之间)

**内部自动切割地震与断层标签组成样本，大小128\*128\*128**

**训练生成的模型位于ai\_models/GeoAIFault3d下**

处理参数:

**The Ratio of Training Samples:** 训练样本占训练数据集的比例，**建议为1**

**Epoch:** 迭代次数，缺省为50

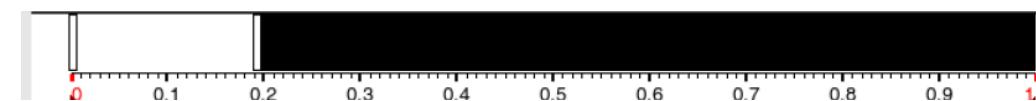
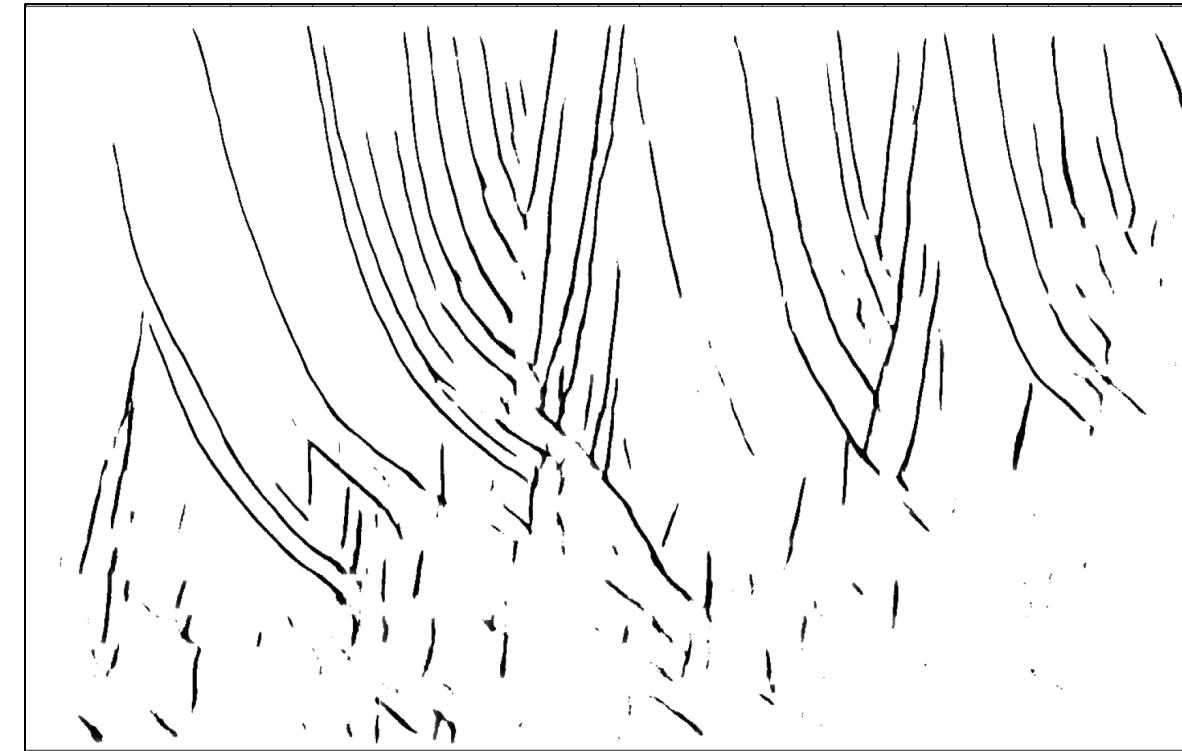
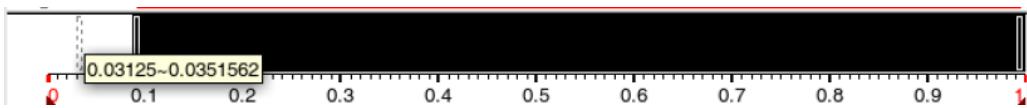
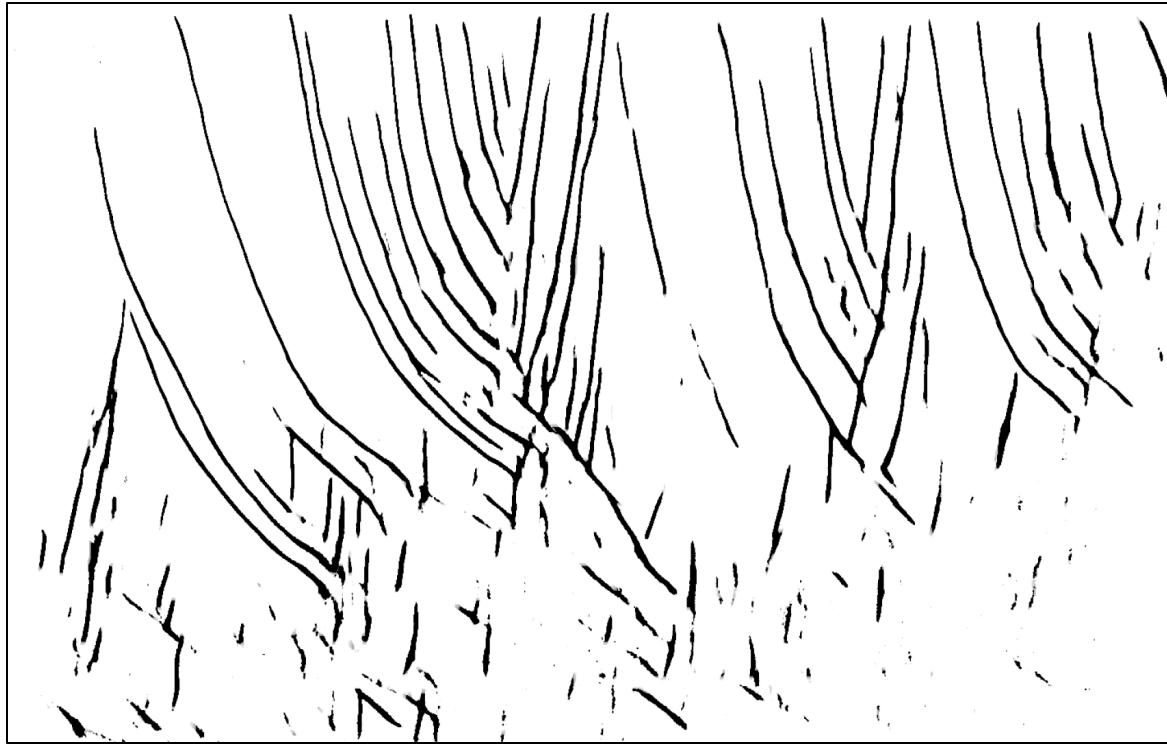
**Min Valid Label Value:** 最小有效标签数，对于Label Data，**小于该值的数据都为0。**

**对于个性化2D标签生成等的概率体，数值范围在0-1之间，需要对概率体进行浏览，确定该值。**



### 3.个性化3D标签

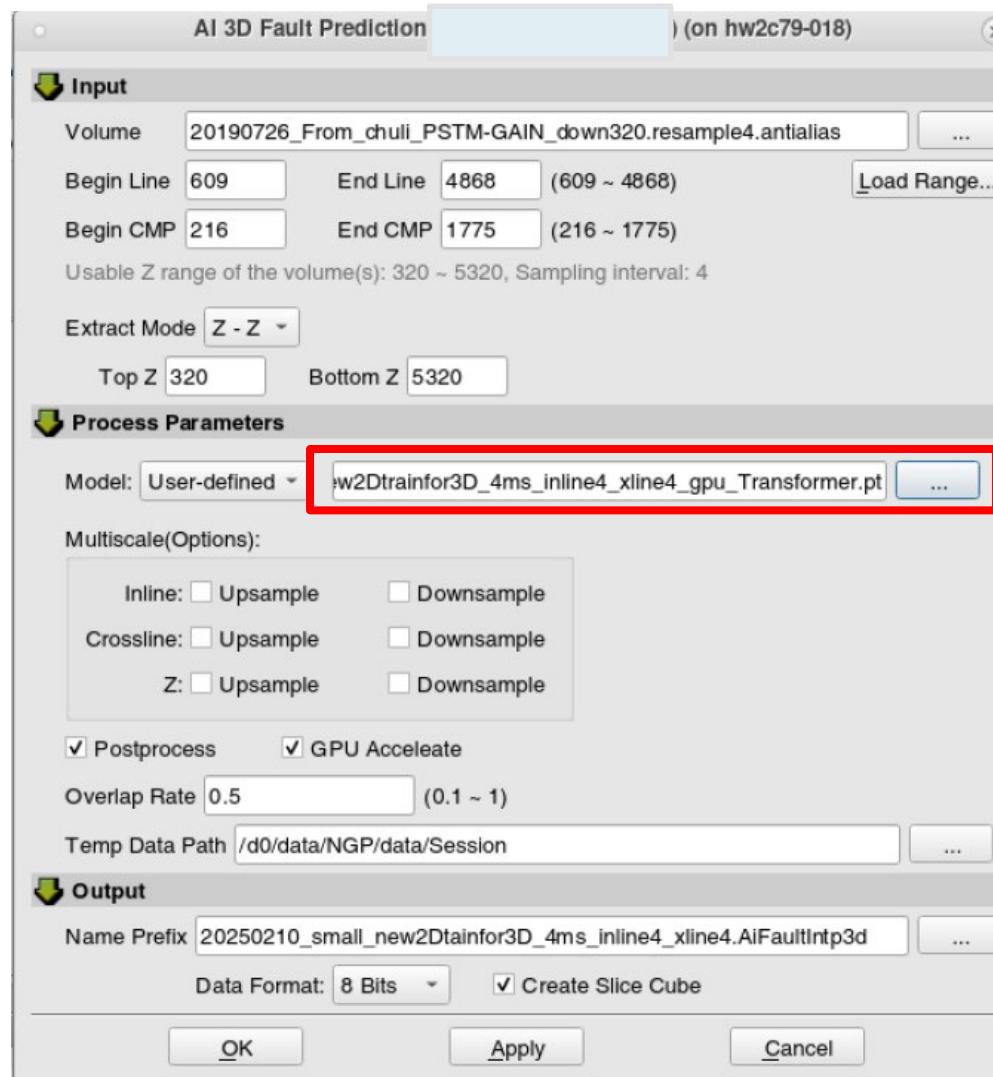
最小有效标签数





### 3.个性化3D标签- ②预测

界面参数



输入:

**Seismic Data : 预测的地震数据**

**程序内部将数据切成128\*128\*128的大小进行预测**

主要参数:

**Model: User-defined**

**选择训练好的网络模型, 模型位于ai\_models/GeoAIFault3d下**

**GPU Acceleration : 是否采用GPU加速训练。**

**PostPocess: 连续性处理 (个性化3D标签预测方法必须进行连续性处理)**

**Overlap Rate: 预测过程的预测尺寸的重叠率**

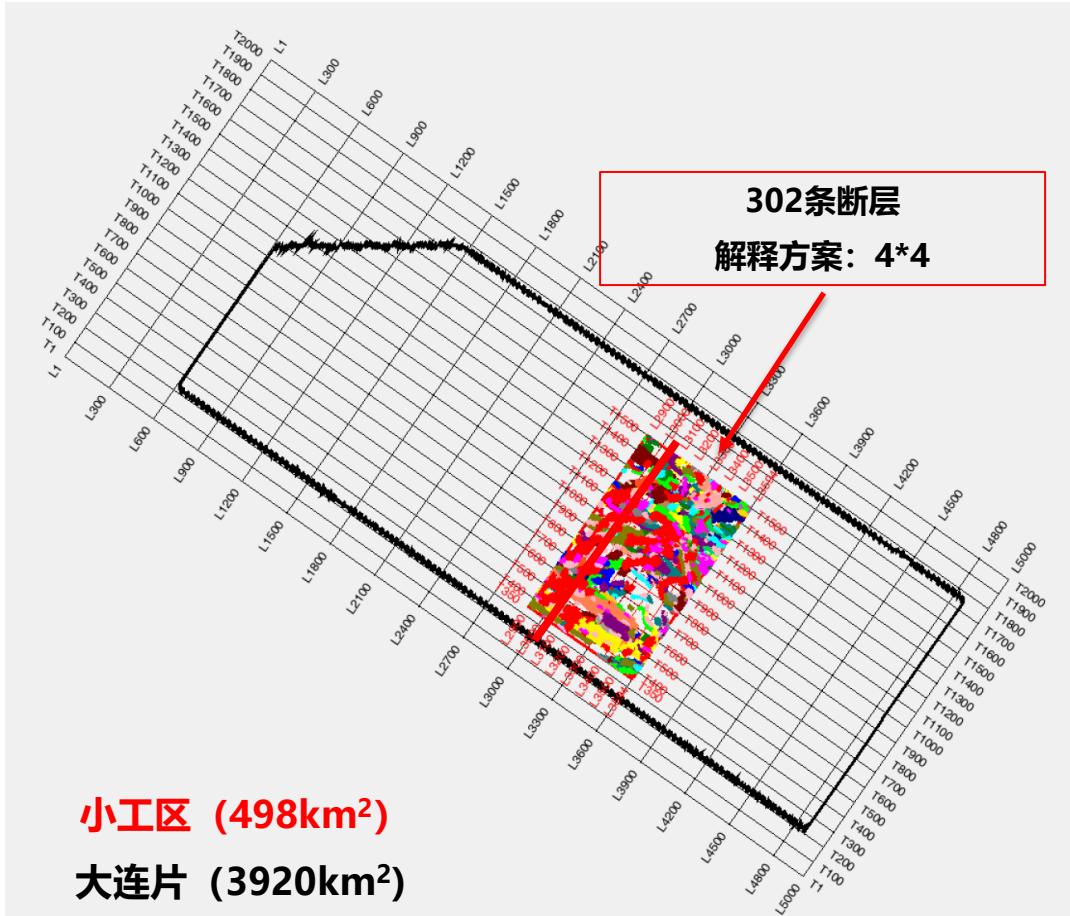


### 3.个性化3D标签

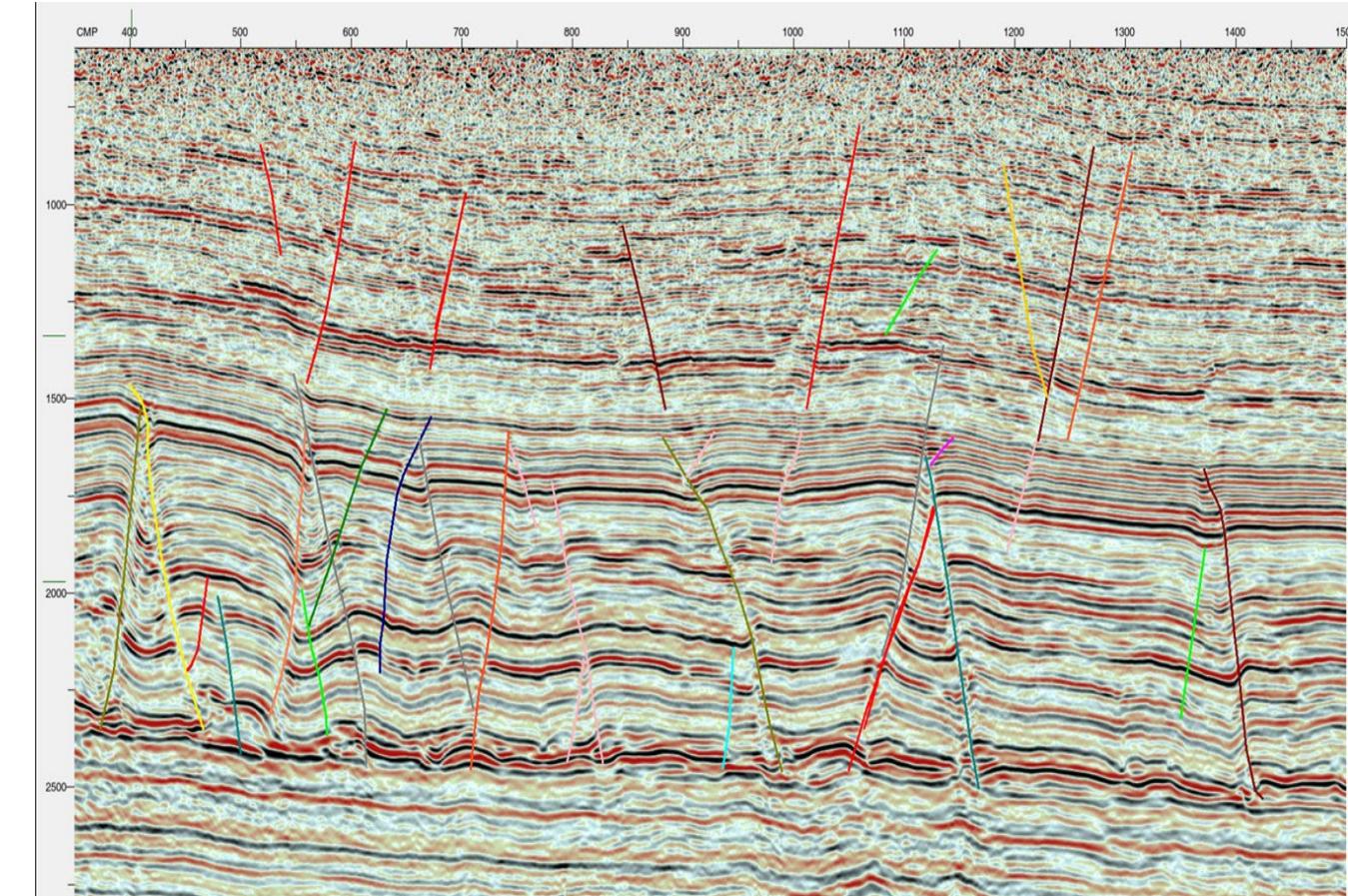
应用实例-西部某工区



小工区个性化3D标签训练模型在大工区/临近工区的智能断层预测



工区情况



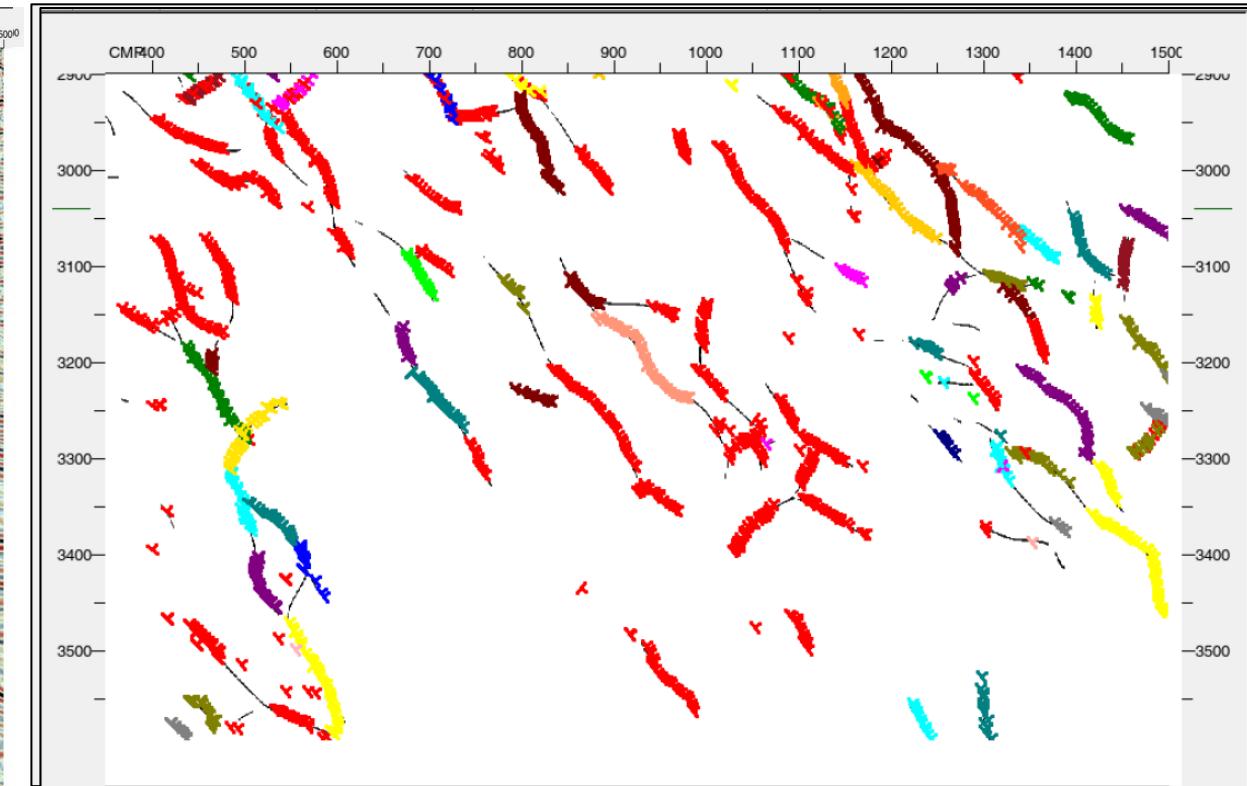
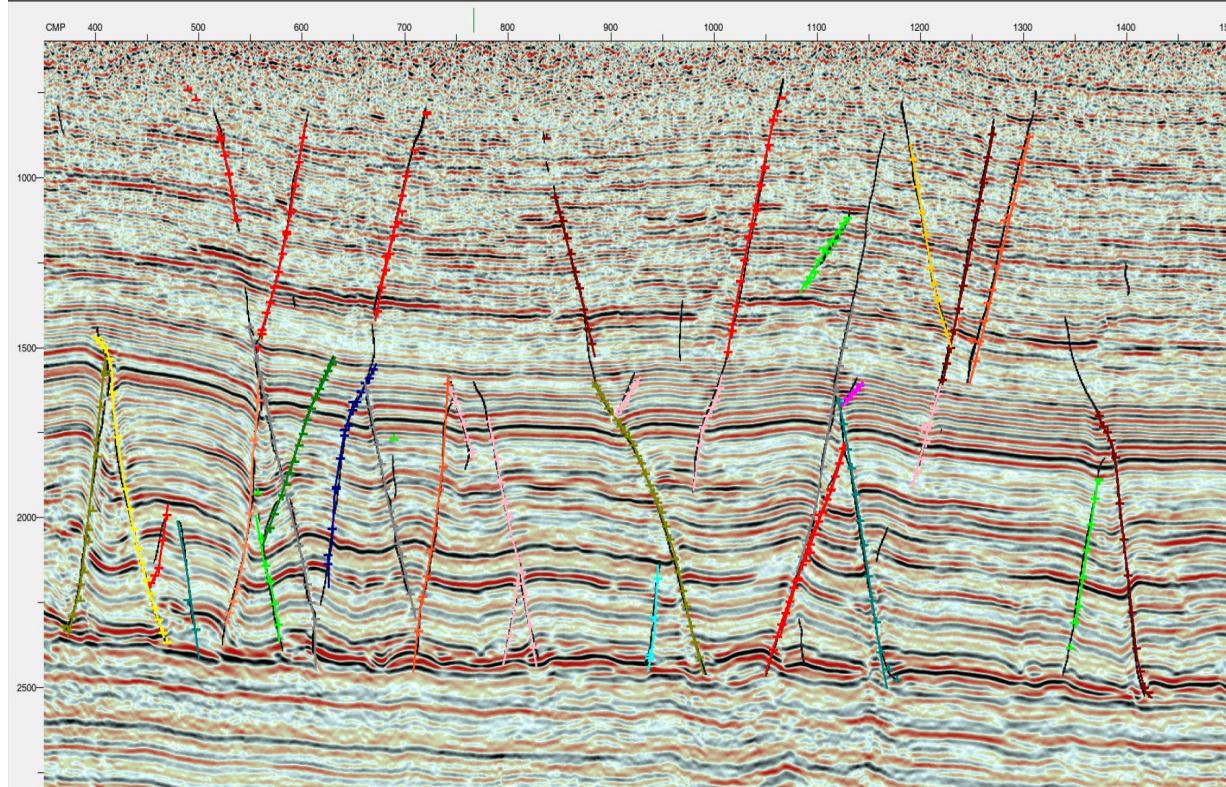


### 3.个性化3D标签

应用实例-西部某工区



#### 小工区基于断层体的个性化2D标签智能断层预测结果



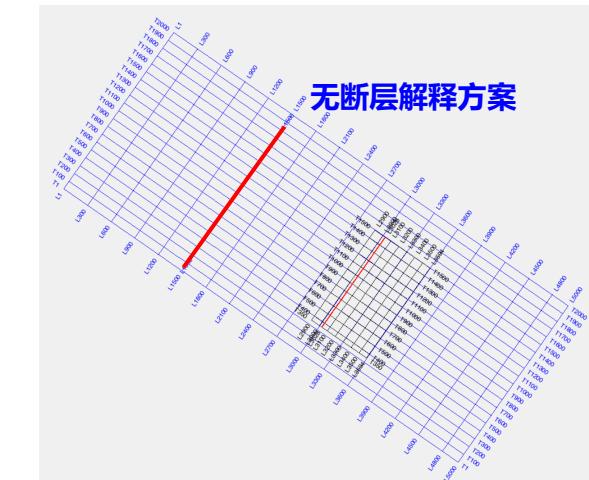
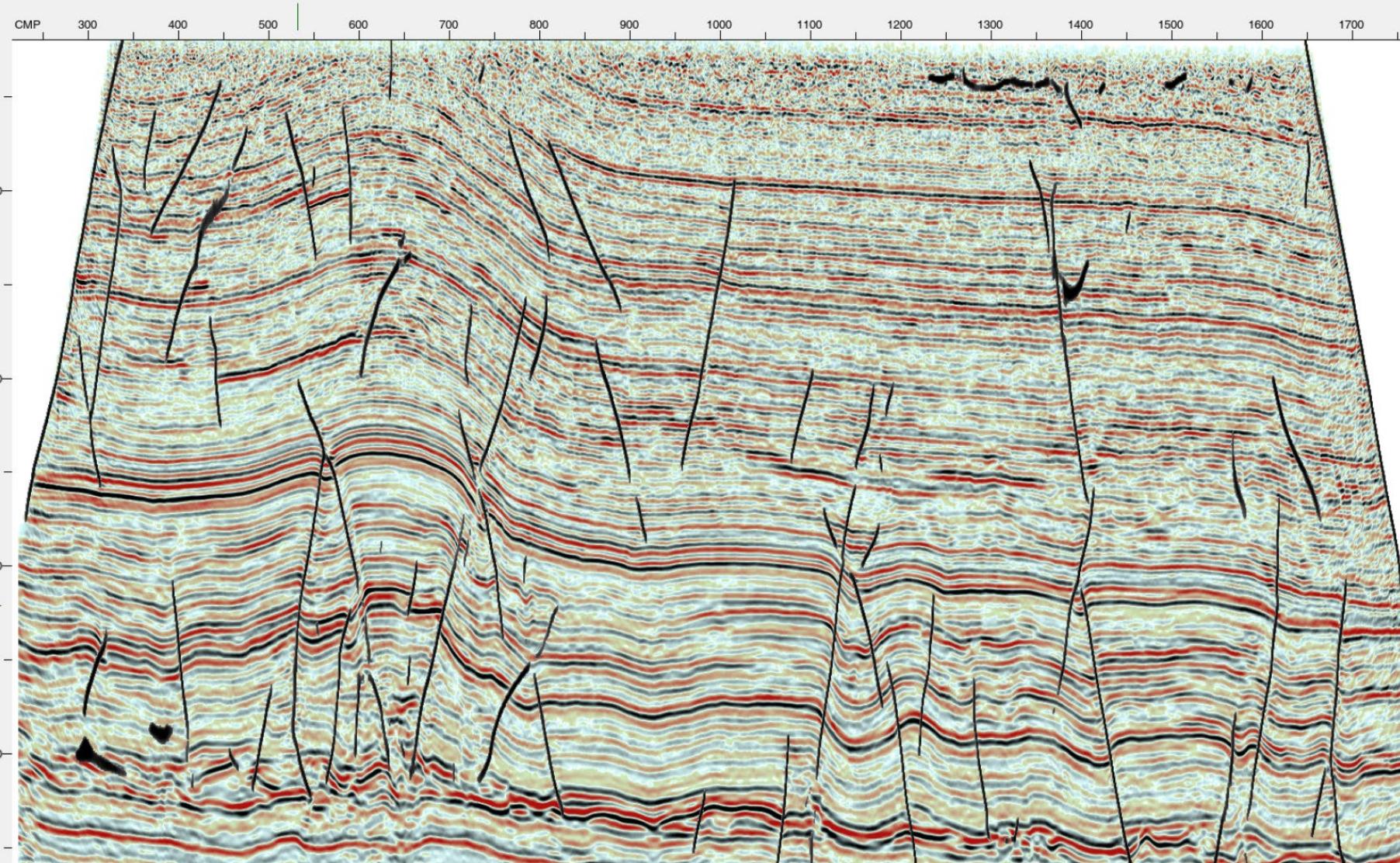
基于断层体的个性化2D标签智能断层预测 (2ms, 4\*4)

完全符合断层解释方案，适用于快速构造建模



### 3.个性化3D标签

应用实例-西部某工区

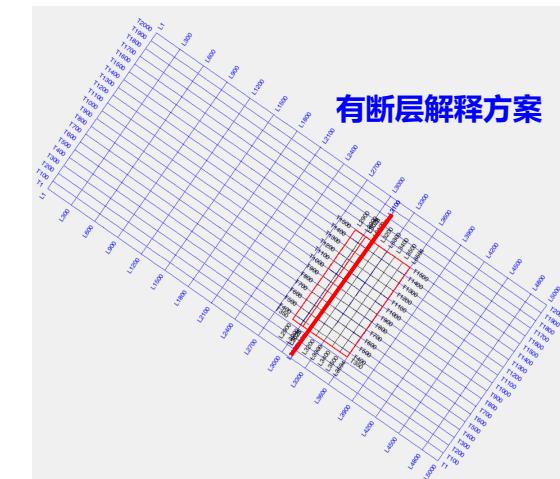
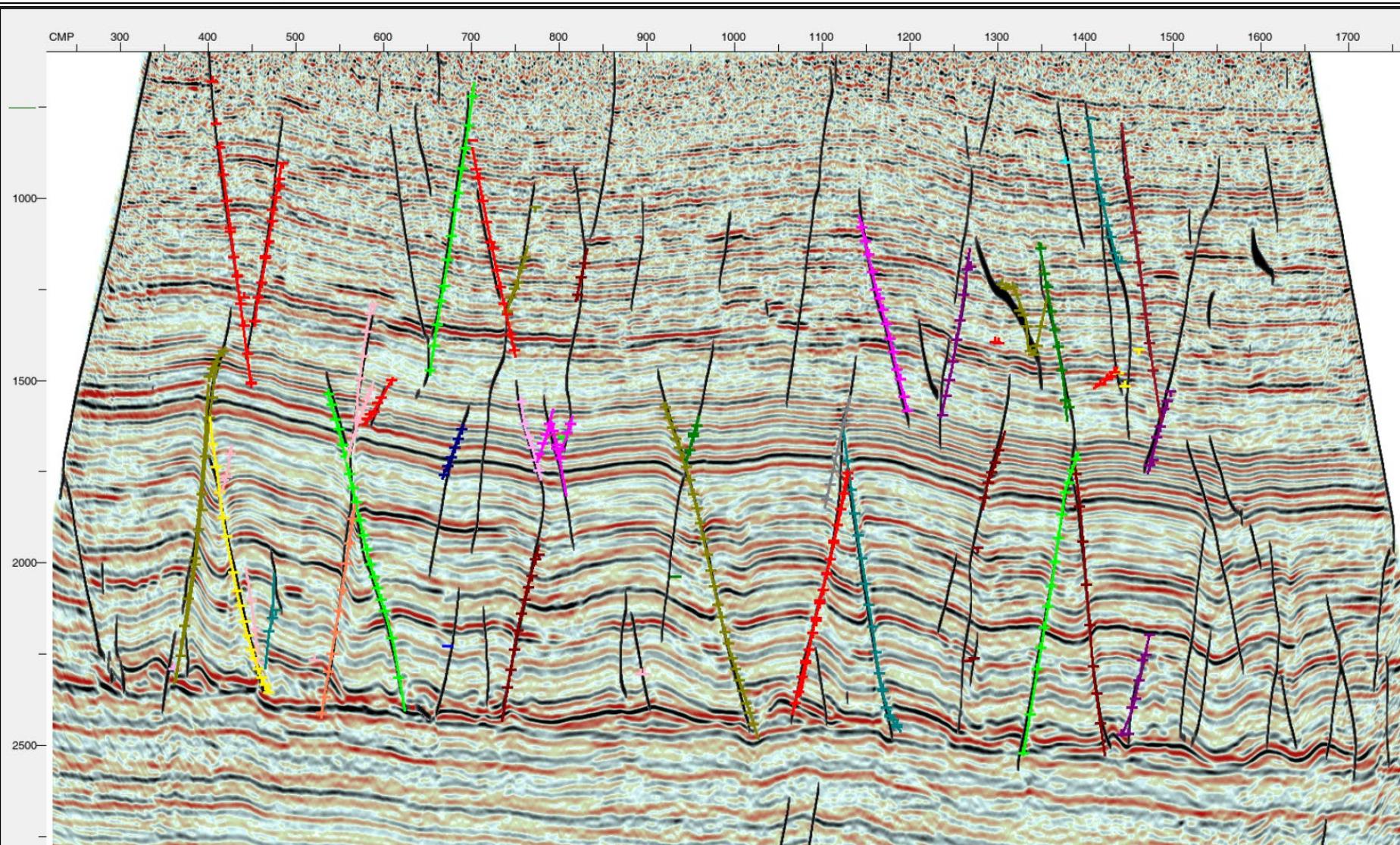


个性化3D标签训练模型在大工区的预测



### 3.个性化3D标签

应用实例-西部某工区

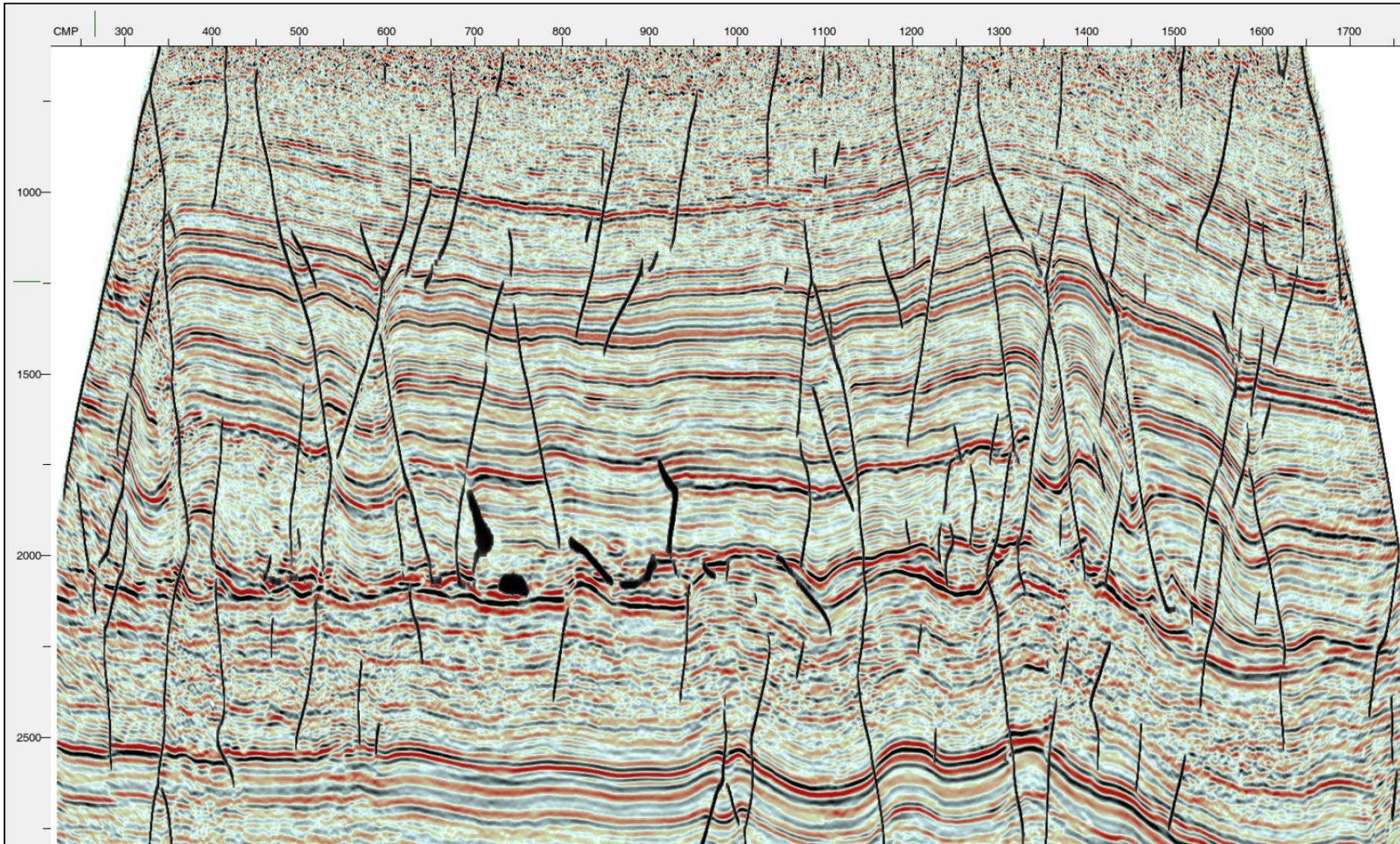


个性化3D标签训练模型在大工区的预测



## 3. 个性化3D标签

## 应用实例-西部某工区



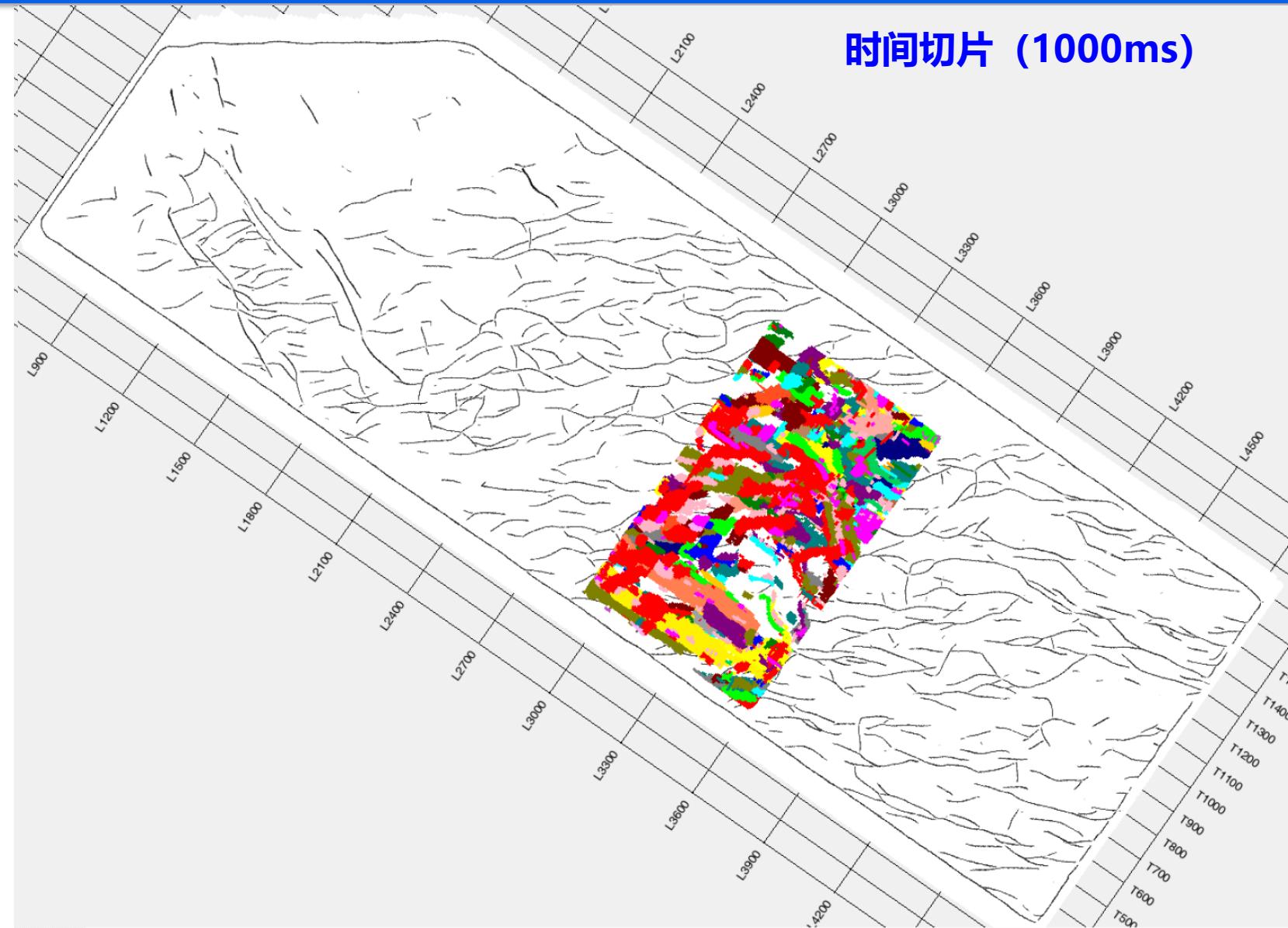
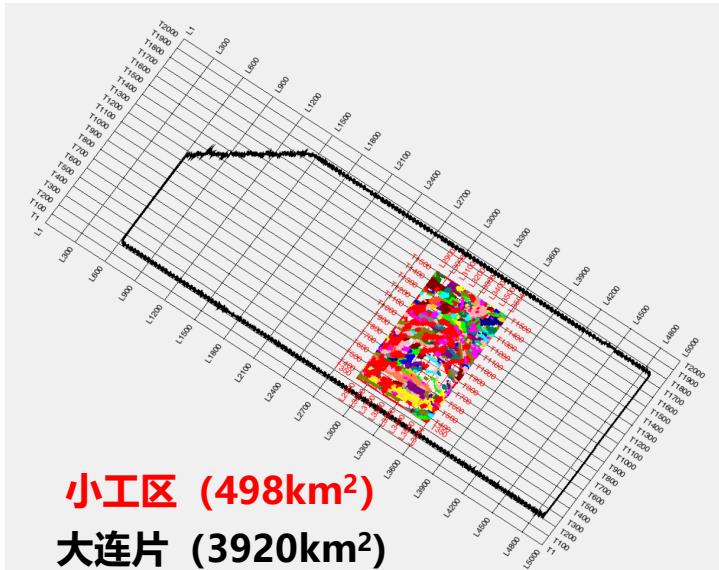
## 个性化3D标签训练模型在大工区的预测

50



### 3.个性化3D标签

应用实例-西部某工区

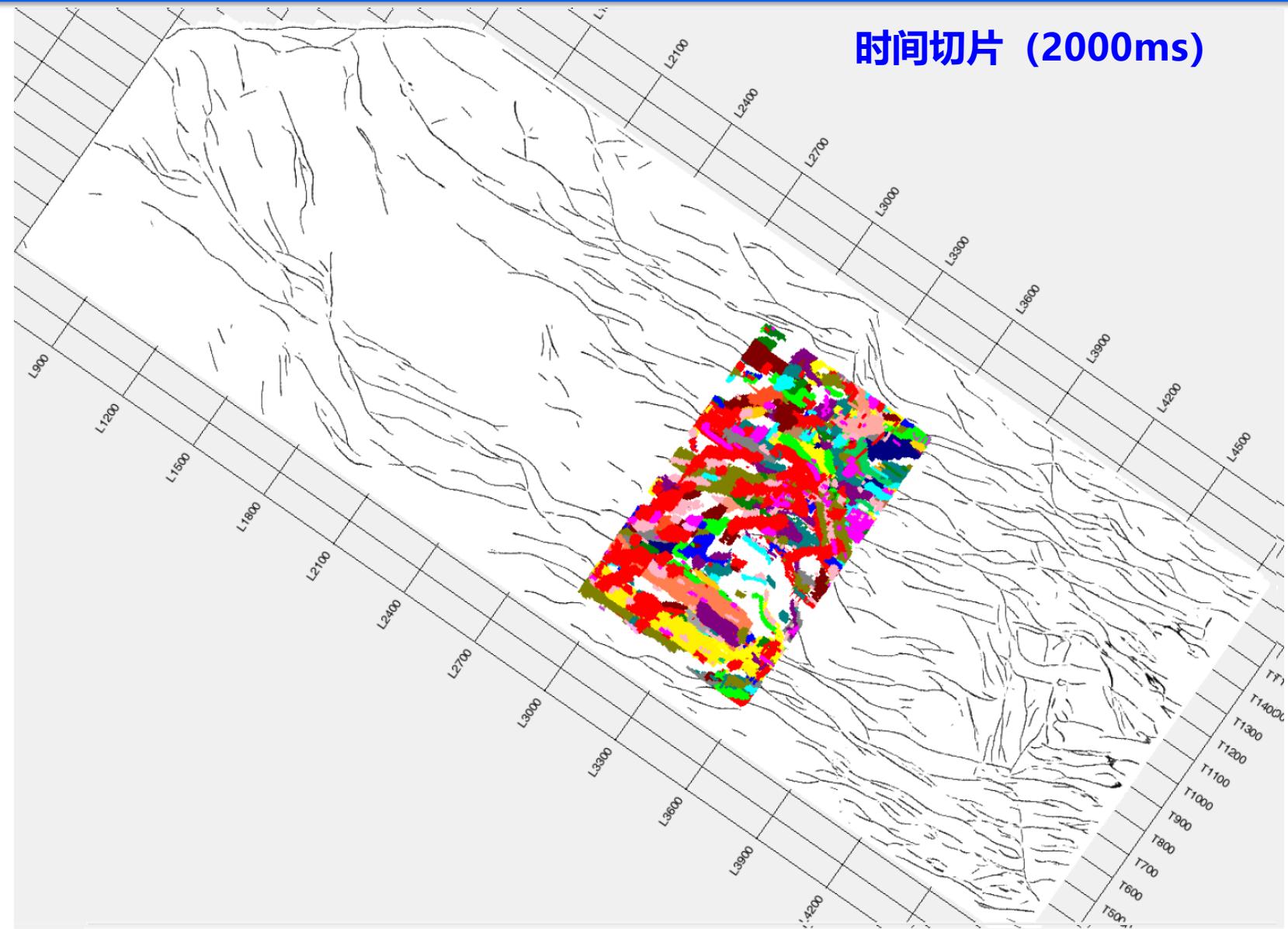
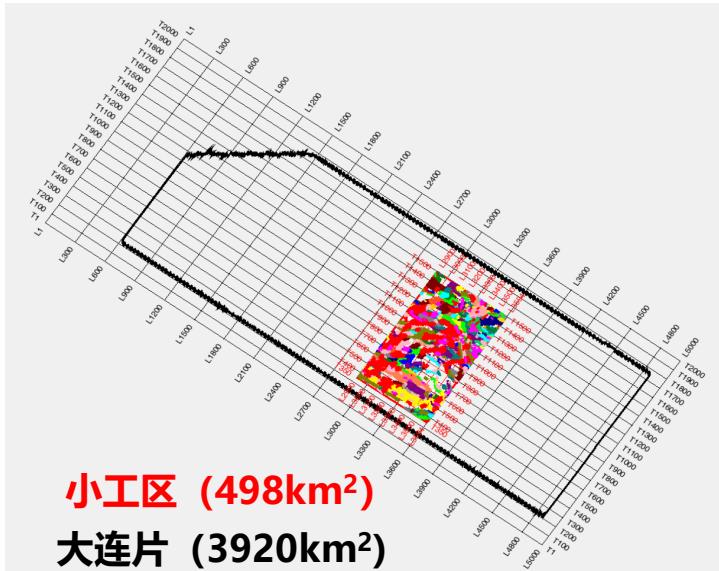


个性化3D标签训练模型在大工区的预测



### 3.个性化3D标签

应用实例-西部某工区



个性化3D标签训练模型在大工区的应用



## CONTENES 目录

01

02

03

研究背景

三维智能断层模块功能介绍

三维合成数据断层预测

个性化三维稀疏标签断层预测

个性化三维标签断层预测

注意事项及建议流程



# 注意事项-GPU问题



## 1、首先用nvidia-smi命令来看GPU使用情况，GPU要预留最少5GB的显存：

```
2 nvidia-smi
```

GPU状态表显示成功代表环境正常，只需等待显卡显存占用率低即可使用人工智能解释模块，如下图：

```
fnode044@V4.2:/bgpssd/home/gs136>nvidia-smi
Tue Oct 10 16:58:37 2023
+-----+
| NVIDIA-SMI 535.104.05      Driver Version: 535.104.05    CUDA Version: 12.2 |
| Persistence-M | Bus-Id      Disp.A  | Volatile Uncorr. ECC | | | | | | |
| GPU  Name      Fan  Temp  Perf  Pwr:Usage/Cap | Memory-Usage | GPU-Util  Compute M. |
|          |             |          |          |          |          |          |          |          |
+-----+
| 0  NVIDIA T1000 8GB     Off | 00000000:27:00.0 Off | N/A      | | | | |
| 38% 58C   PO      N/A / 50W | 3151MiB / 8192MiB | 75%     Default |
|          |             |          |          |          |          |          |
+-----+
               | Processes:                               GPU Memory |
               | GPU  GI  CI      PID  Type  Process name        Usage  |
               | ID  ID          |          |          |          |
+-----+
| 0  N/A  N/A      26185  C  ...rt/miniconda3/envs/nv/bin/python3.9  3148MiB |
+-----+
```

有计算作业正在使用GPU，建议等待显存充足时再使用



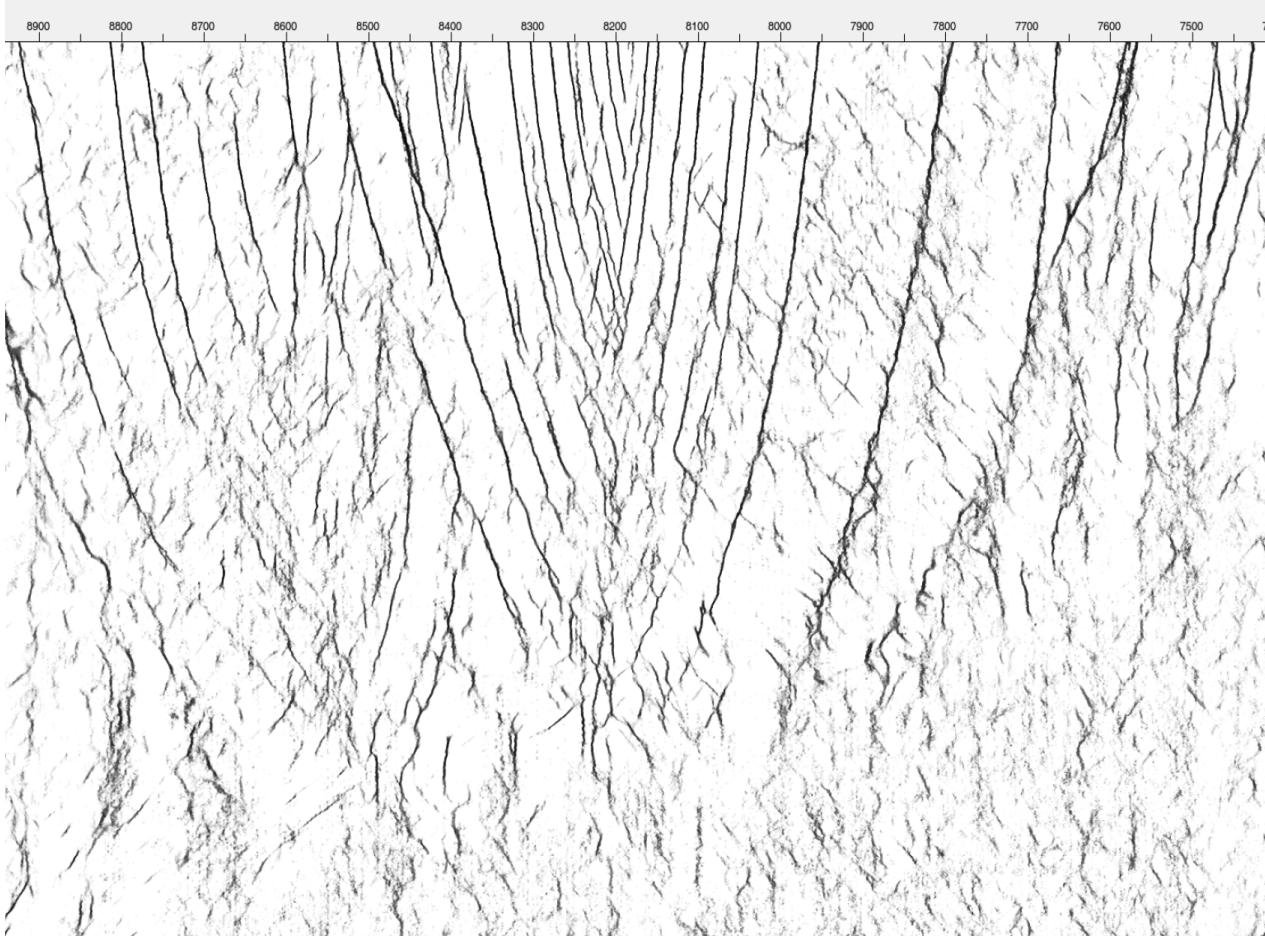
<https://note.youdao.com/s/HYIuOhmJ>



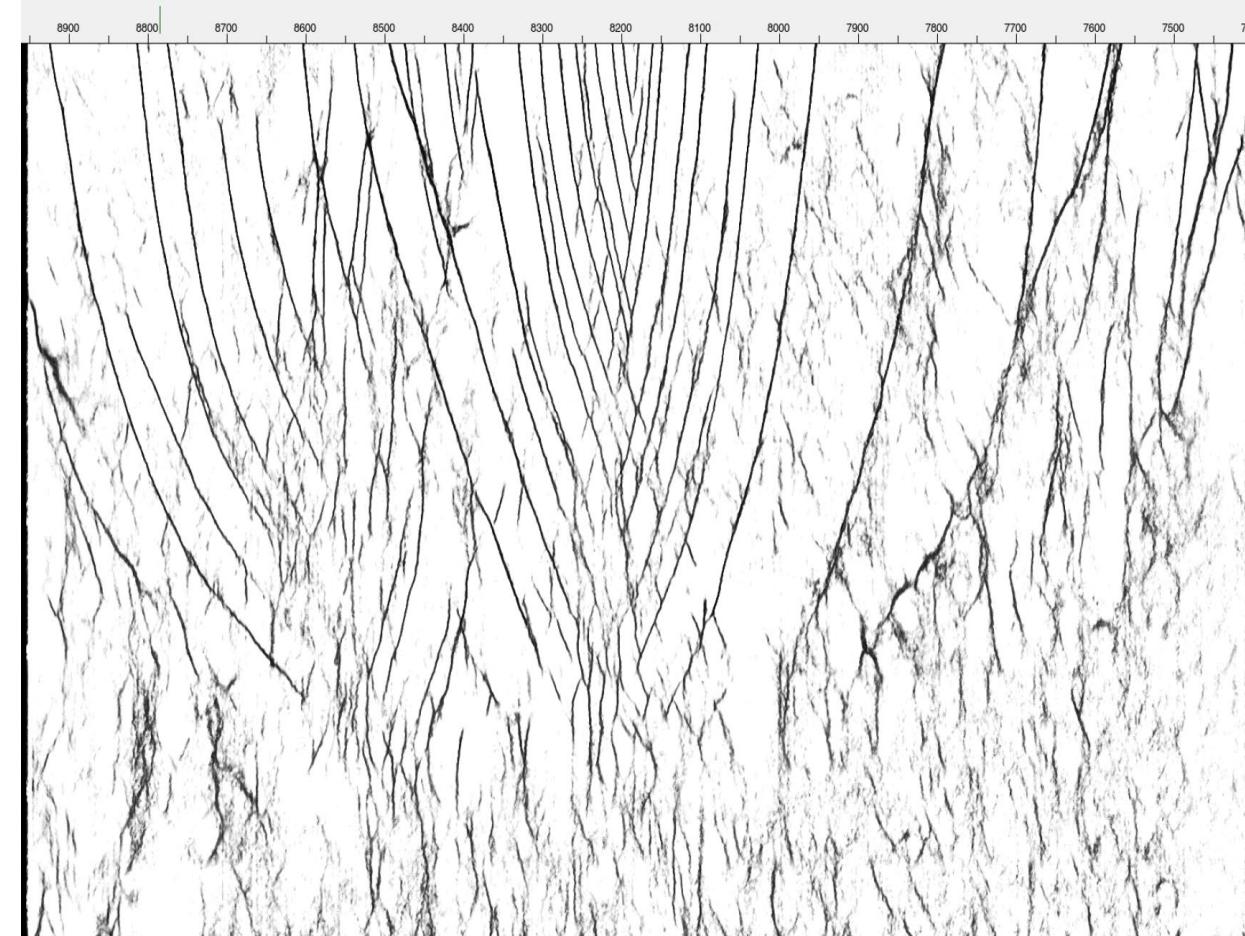
## 注意事项-采样率



2、内部尺寸为128\*128\*128，时间域建议用4ms，深度域为10m



2ms采样

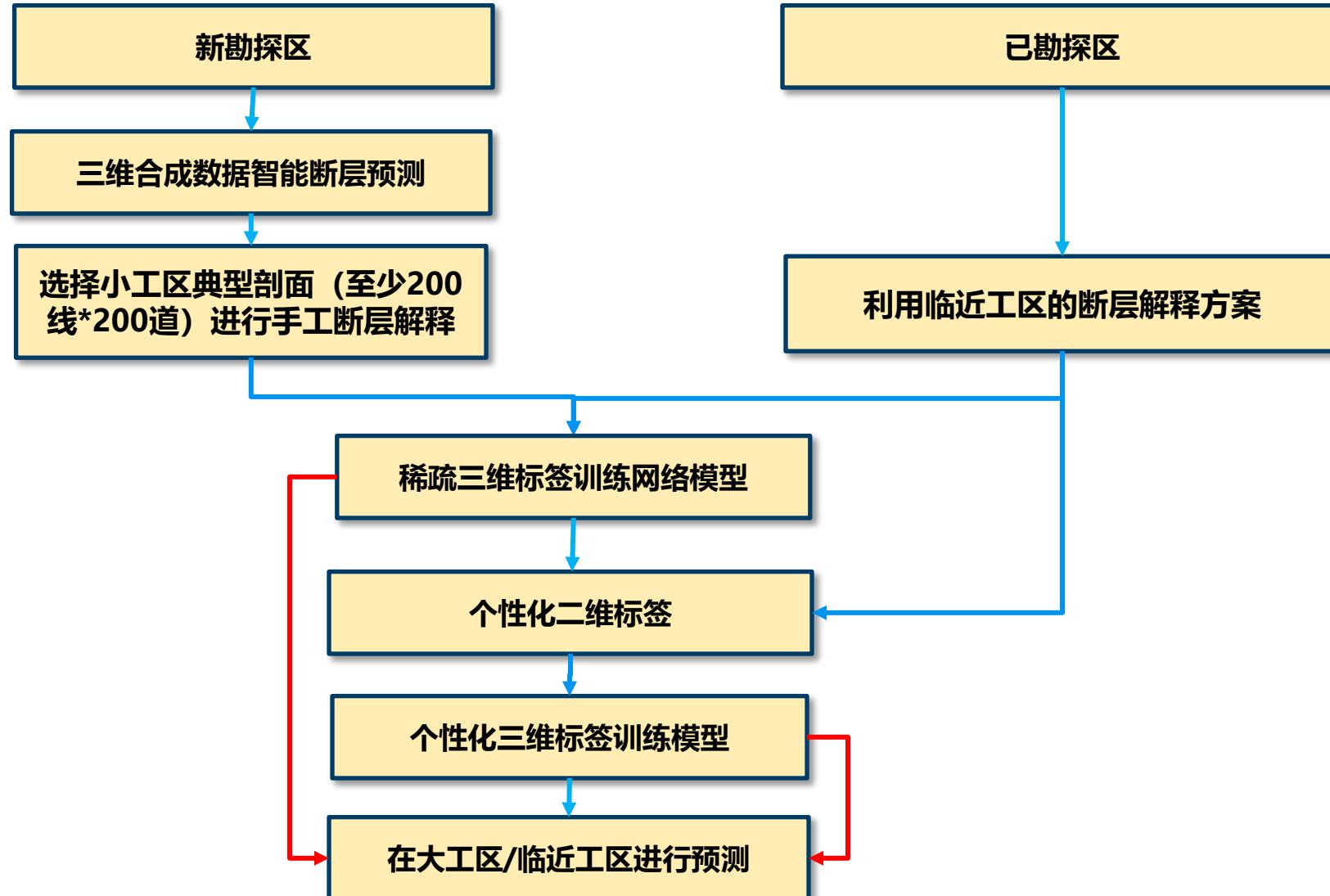


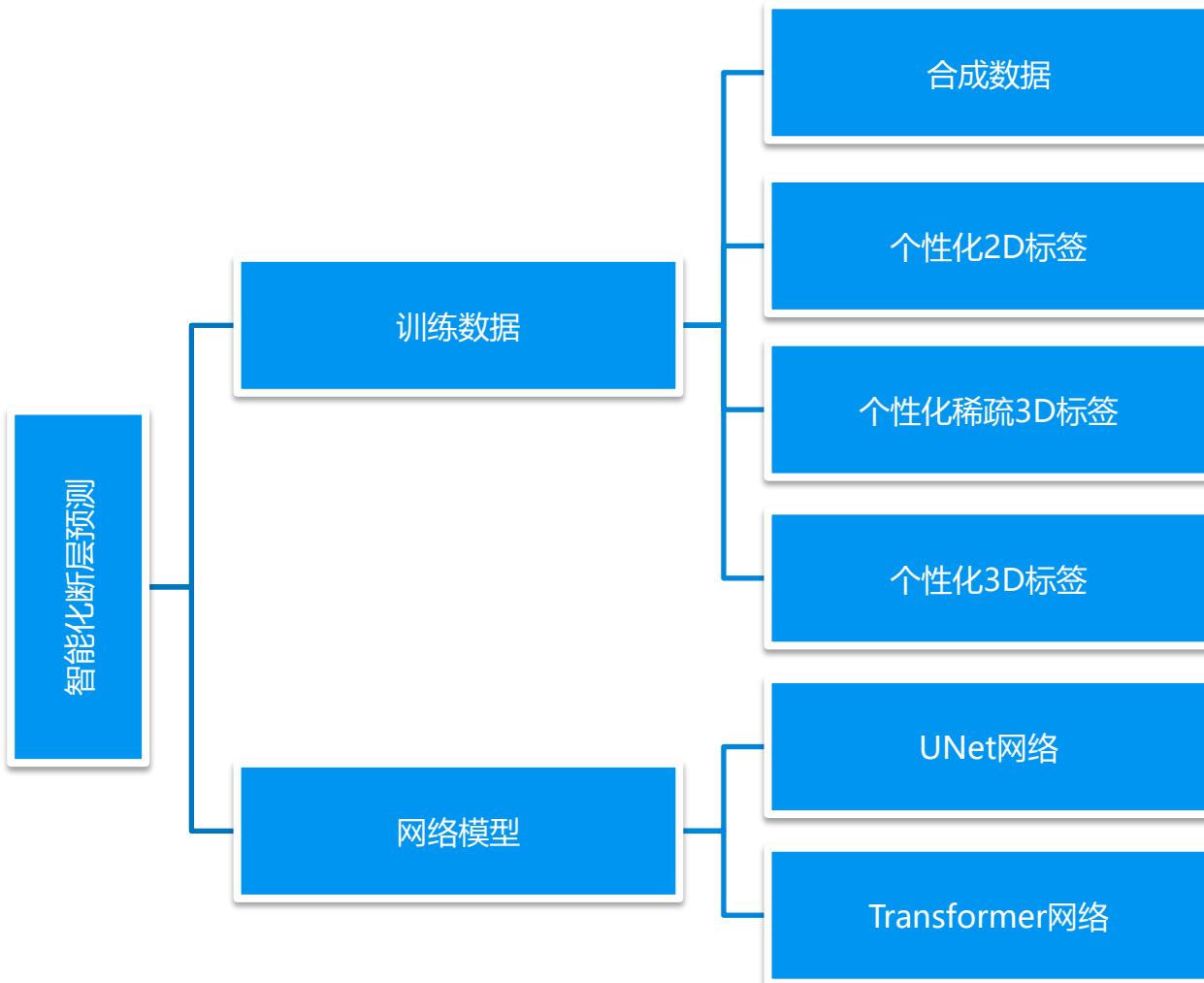
4ms采样

**使用技巧：**  
选取500线\*500道数据进行测试  
至少选择200线\*200道



# 三维智能断层预测建议工作流程





**提升断层预测适应性**

**提高断层预测精度**

**智能断层预测关键技术点**

# 感谢大家对GeoEast软件的信任和支持！

更多详情请关注



GeoEast微信公众号



解释技术支持QQ群



Bilibili视频教程