模式分析与机器智能
工业和信息化部重点实验室
MIIT Key Laboratory of
Pattern Analysis & Machine Intelligence

ParNeC 模式识别与神经计算研究组
PAttern Recognition and NEural Computing

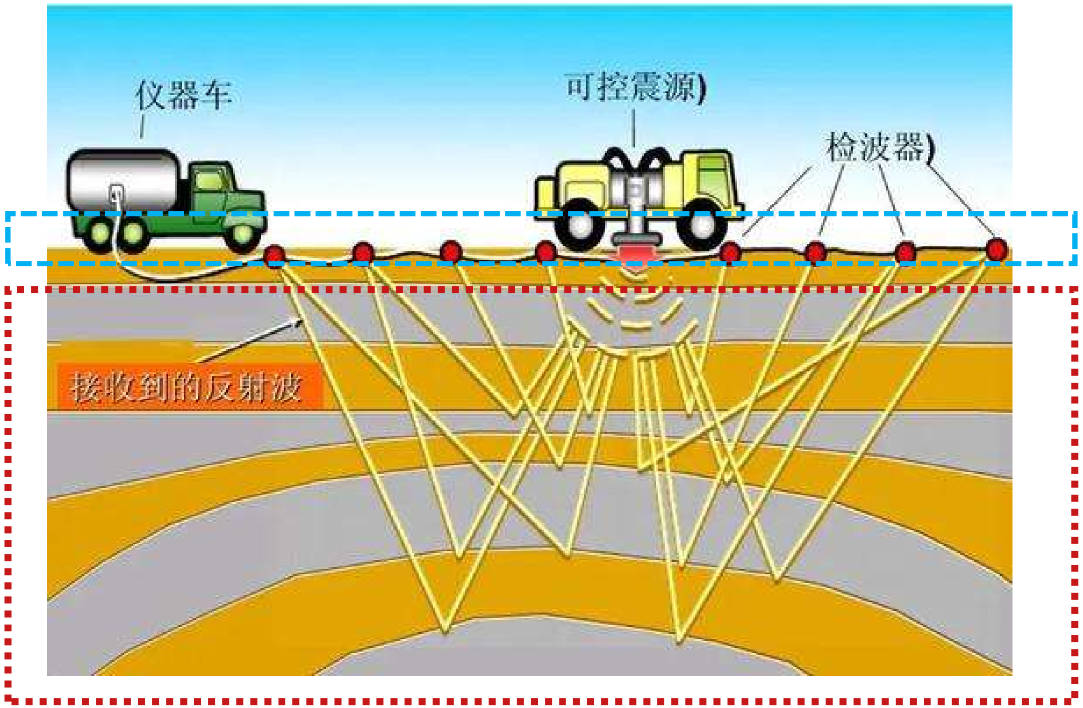
UNSUPERVISED LEARNING OF FULL-WAVEFORM INVERSION: CONNECTING CNN AND PARTIAL DIFFERENTIAL EQUATION IN A LOOP

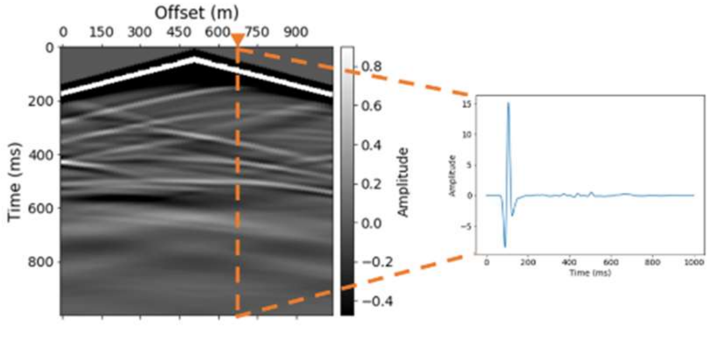
Peng Jin[1,2,*], Xitong Zhang[1,3,*], Yinpeng Chen[4], Sharon Xiaolei Huang[2], Zicheng Liu[4] and Youzuo Lin[1]

ICLR 2022

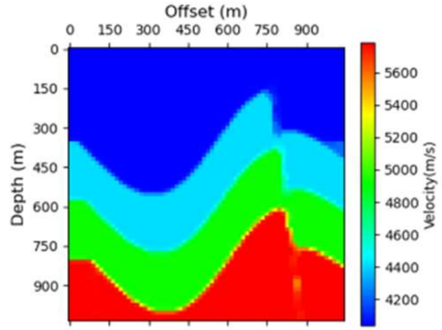
Introduction

Full Waveform Inversion（FWI)





(b)



(a)

Forward Modeling and Data-driven FWI

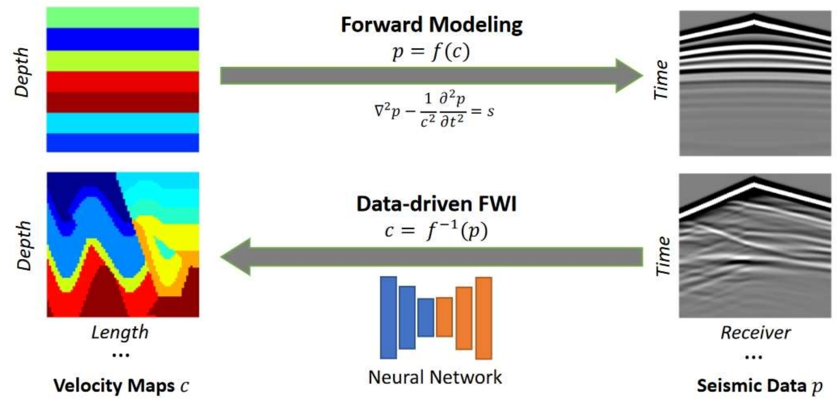


Figure 2: **Schematic illustration of data-driven FWI and forward modeling.** Neural networks are employed to infer velocity maps from seismic data while forward modeling is to calculate the seismic data using governing wave equations with velocity map provided.

acoustic wave equation :

$$\nabla^2 p(x,z,t) - \frac{1}{c(x,z)^2}\frac{\partial^2 p(x,z,t)}{\partial t^2} = S_{xs,f}(x,z,t) \quad (1)$$

P: pressure wavefield
C: velocity map
S: source

These models are prone to making unreasonable or unrealistic predictions that do not conform to physical mechanisms.

Unsupervised Learning FWI

Advantage: It incorporates physical prior knowledge, reducing the demand for labeled data.

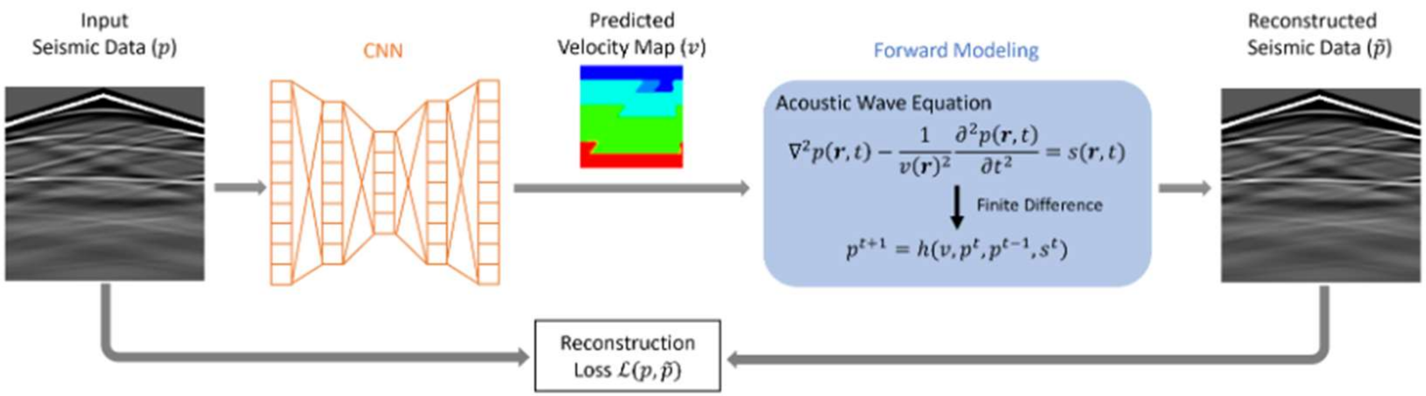


Figure 1: Schematic illustration of our proposed method, which comprises a CNN to learn an inverse mapping and a differentiable operator to approximate the forward modeling of PDE.

Differentiable forward modeling

Apply the standard **finite difference (FD)** in the space domain and time domain to discretize the original wave equation.

acoustic wave equation :

$$\nabla^2 p(\boldsymbol{r}, t) - \frac{1}{v(\boldsymbol{r})^2} \frac{\partial^2 p(\boldsymbol{r}, t)}{\partial t^2} = s(\boldsymbol{r}, t) \ , \tag{1}$$

$$\frac{\partial^2 p(\boldsymbol{r}, t)}{\partial t^2} \approx \frac{1}{(\Delta t)^2}(p_{\boldsymbol{r}}^{t+1} - 2p_{\boldsymbol{r}}^{t} + p_{\boldsymbol{r}}^{t-1}) + O[(\Delta t)^2] \ , \tag{5}$$

where $p_{\boldsymbol{r}}^t$ denotes the pressure wavefields at timestep t, and $p_{\boldsymbol{r}}^{t+1}$ and $p_{\boldsymbol{r}}^{t-1}$ are the wavefields at $t + \Delta t$ and $t - \Delta t$, respectively. The Laplacian of $p(\boldsymbol{r}, t)$ can be estimated in the similar way on the space domain (see Appendix A.2). Therefore, the wave equation can then be written as

$$p_{\boldsymbol{r}}^{t+1} = (2 - v^2 \nabla^2) p_{\boldsymbol{r}}^t - p_{\boldsymbol{r}}^{t-1} - v^2 (\Delta t)^2 s_{\boldsymbol{r}}^t \ , \tag{6}$$

where ∇^2 here denotes the discrete Laplace operator.

$$\begin{aligned}\nabla^2 p(\boldsymbol{r}, t) =& \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2}, \\ \approx& \frac{1}{(\Delta x)^2} \sum_{i=-2}^{2} c_i p_{x+i,z}^t + \frac{1}{(\Delta z)^2} \sum_{i=-2}^{2} c_i p_{x,z+i}^t \\ &+ O[(\Delta x)^4 + (\Delta z)^4] \ ,\end{aligned} \tag{14}$$

where $c_0 = -\frac{5}{2}$, $c_1 = \frac{4}{3}$, $c_2 = -\frac{1}{12}$, $c_i = c_{-i}$

Loss Function

The reconstruction loss of our UPFWI includes a pixel-wise loss and a perceptual loss as follows:

$$\mathcal{L}(\boldsymbol{p}, \tilde{\boldsymbol{p}}) = \mathcal{L}_{pixel}(\boldsymbol{p}, \tilde{\boldsymbol{p}}) + \mathcal{L}_{perceptual}(\boldsymbol{p}, \tilde{\boldsymbol{p}}), \tag{8}$$

where $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$ are input and reconstructed seismic data, respectively. The pixel-wise loss $\mathcal{L}_{pixel}$ combines ℓ_1 and ℓ_2 distance as:

$$\mathcal{L}_{pixel}(\boldsymbol{p}, \tilde{\boldsymbol{p}}) = \lambda_1 \ell_1(\boldsymbol{p}, \tilde{\boldsymbol{p}}) + \lambda_2 \ell_2(\boldsymbol{p}, \tilde{\boldsymbol{p}}), \tag{9}$$

where λ_1 and λ_2 are two hyper-parameters to control the relative importance. For the perceptual loss $\mathcal{L}_{perceptual}$, we extract features from `conv5` in a VGG-16 network (Simonyan & Zisserman, 2015) pretrained on ImageNet (Krizhevsky et al., 2012) and combine the ℓ_1 and ℓ_2 distance as:

$$\mathcal{L}_{perceptual}(\boldsymbol{p}, \tilde{\boldsymbol{p}}) = \lambda_3 \ell_1(\phi(\boldsymbol{p}), \phi(\tilde{\boldsymbol{p}})) + \lambda_4 \ell_2(\phi(\boldsymbol{p}), \phi(\tilde{\boldsymbol{p}})), \tag{10}$$

where $\phi(\cdot)$ represents the output of `conv5` in the VGG-16 network, and λ_3 and λ_4 are two hyper-parameters. Compared to the pixel-wise loss, the perceptual loss is better to capture the region-wise structure, which reflects the waveform coherence. This is crucial to boost the overall accuracy of velocity maps (e.g. the quantitative velocity values and the structural information).

Experiment

Supervision	Method	FlatFault			CurvedFault		
		MAE ↓	MSE ↓	SSIM ↑	MAE ↓	MSE ↓	SSIM ↑
Supervised	InversionNet	15.83	2156.00	0.9832	23.77	5285.38	0.9681
	VelocityGAN	16.15	1770.31	0.9857	25.83	5076.79	0.9699
	H-PGNN+ (our implementation)	**12.91**	1565.02	0.9874	24.19	5139.60	0.9685
Unsupervised	**UPFWI-24K** (ours)	16.27	1705.35	0.9866	29.59	5712.25	0.9652
	UPFWI-48K (ours)	14.60	**1146.09**	**0.9895**	**23.56**	**3639.96**	**0.9756**

Table 1: **Quantitative results evaluated on OpenFWI** in terms of MAE, MSE and SSIM. Our UPFWI yields comparable inversion accuracy comparing to supervised baselines. For H-PGNN+, we use our network architecture to replace the original one reported in their paper, and an additional perceptual loss between seismic data is added during training.

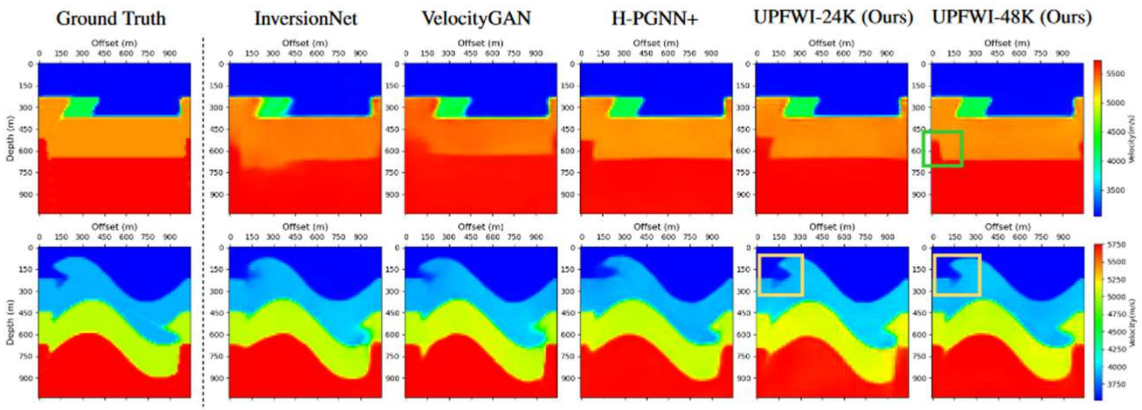


Figure 4: **Comparison of different methods on inverted velocity maps of FlatFault (top) and CurvedFault (bottom)**. For FlatFault, our UPFWI-48K reveals more accurate details at layer boundaries and the slope of the fault in deep region. For CurvedFault, our UPFWI reconstructs the geological anomalies on the surface that best match the ground truth.

Ablation study

Loss			Velocity Error			Seismic Error		
pixel-ℓ_2	pixel-ℓ_1	perceptual	MAE ↓	MSE ↓	SSIM ↑	MAE ↓	MSE ↓	SSIM ↑
✓			32.61	10014.47	0.9735	0.0167	0.0023	0.9978
✓	✓		21.71	2999.55	0.9775	0.0155	0.0025	0.9977
✓	✓	✓	**16.27**	**1705.35**	**0.9866**	**0.0140**	**0.0021**	**0.9984**

Table 2: Quantitative results of our UPFWI with different loss function settings.

Network	MAE↓	MSE↓	SSIM↑
CNN	16.27	1705.35	0.9866
ViT	41.44	11029.01	0.9461
MLP-Mixer	22.32	4177.37	0.9726

Table 4: Quantitative results of our UP-FWI with different architectures.

Thanks