SnT 2023
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE
HOFBURG PALACE - Vienna and Online
19 TO 23 JUNE

Advances in Application of Machine Learning to Seismic Monitoring Data Processing

LI Jian,Huang Lihong, WANG Xiaoming,LIU Zhehan,SHANG Jie, GAI Lei, LI Ming,XIAO Yong
CTBT Beijing National Data Center

CTBT Beijing National Data Center
Beijing Radionuclide Laboratory

INTRODUCTION

- With the dramatic increase in the monitoring data, Problems and challenges are emerging in the quality of automatic processing of the event bulletin.
- Many issues in waveform data processing can be converted into classification and regression problems, and solved with ML.

METHODS/DATA

- Local phase picking Based on multi-task CNN
- Noise signal recognition using GAN_LSTM
- Seismic phase sequence detection based on Transformer
- Event association based on probabilistic model

START

RESULTS

- The recall rate of P and S phases are 98% using Multi-task CNN method.
- GAN_LSTM method classified seismic phases and noise signals with a recall of 95% and 97%.
- Grad-T method achieves the phases sequence detection.

CONCLUSION

- ML has become a research hotspot in the field of waveform monitoring data processing.
- Training data are lacking in some categories and
- regions. The model generalization performance needs to be improved.
- The interpretability and explainable problems of ML methods for verification needs to be addressed.

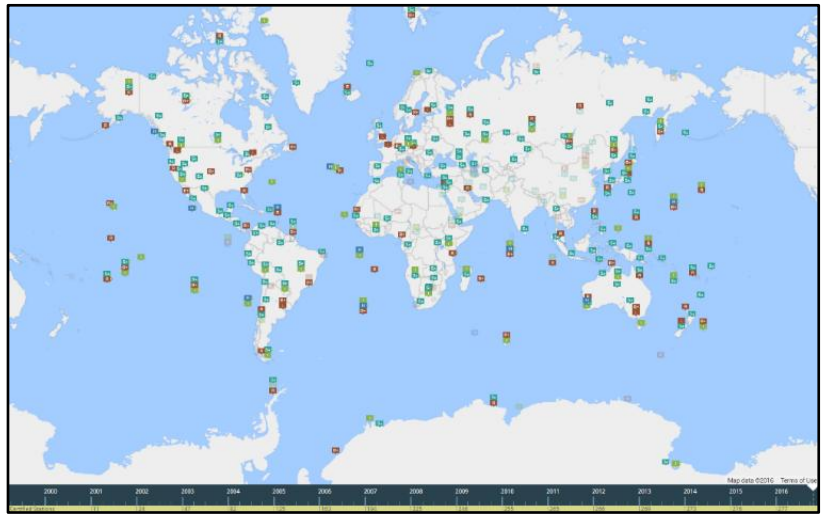
Please do not use this space, a QR code will be automatically overlayed

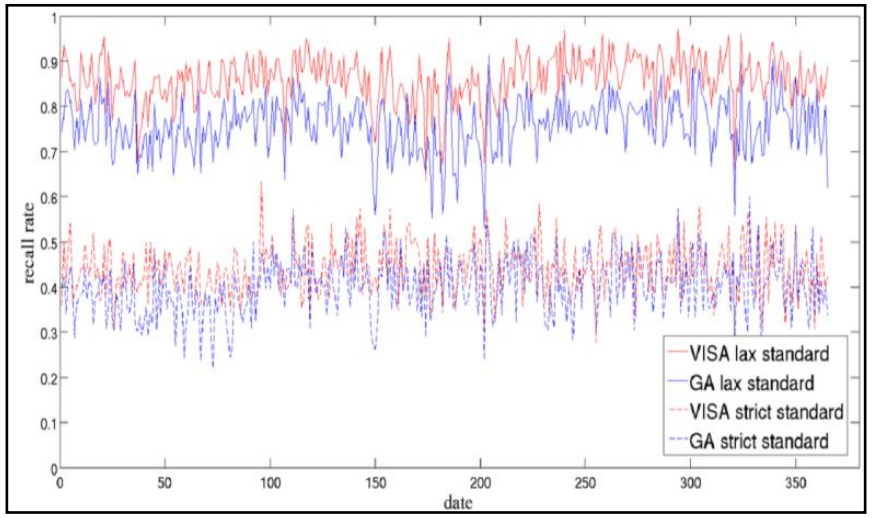
P3.5-449

Introduction

CTBT Beijing National Data Center
Beijing Radionuclide Laboratory

With the dramatic increase in the number of seismic stations and the volume of historical data in the global earthquake monitoring network, the performance of traditional seismic data processing methods is not satisfactory. Problems and challenges are emerging in the quality of automatic processing of the event bulletin and the application of historical data mining.





In recent years, Machine Learning have achieved excellent performance in a wide range of fields.Many issues in waveform data processing can be converted into classification and regression problems, and solved with ML. ML can be used in different stages of seismic data processing, such as event detection, denoising, phase picking, phase association, event location ,and etc.

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Research on intelligent data processing technology in China NDC

CTBT Beijing National Data Center
Beijing Radionuclide Laboratory

In China NDC, we pay more attention to the AI application, especially in waveform data processing. There are several work we have conducted to improve the performance of signal and event detection. These efforts include local events detection based on multi-task CNN, the seismic phase sequence detection based on Transformer, GAN-LSTM joint network applied to seismic noise signal recognition.

Local Phase Picking and Association Based On Deep Learning

- Research on phase picking method using multi-task CNN
- Study on the phases association method based on P/S Wave Velocity Ratio for local events.

Noise Signal Recognition using GAN_LSTM

- Study on noise signal recognition method to reduce seismic phases false association.

Seismic phase sequence detection based on Transformer

- Study on the phases sequence detection method based on transformer for regional events.
- Grad-CAM was used to improving the explainable of the model.

Event association and location based on probabilistic model

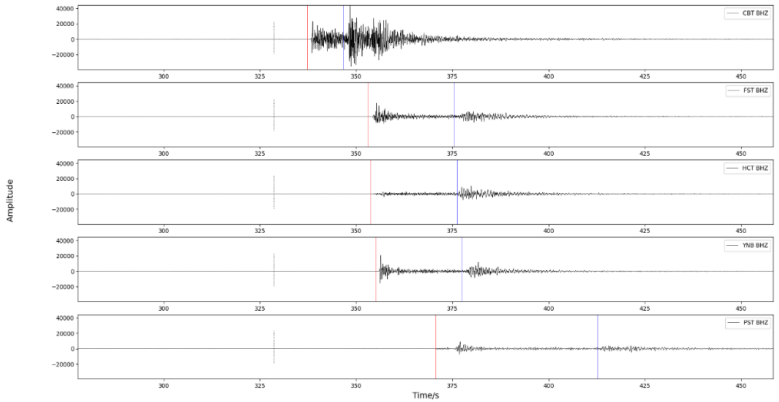
- Evaluation the NET-VISA method , and an association and location method is realized based on the probabilistic model.

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

1.Phase Picking and Association Based On Multi-task CNN for local event

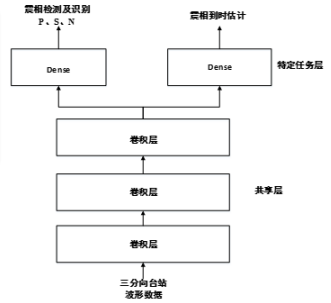
CTBT Beijing National Data Center
Beijing Radionuclide Laboratory

A multi-task convolution neural network was built to detect and identify the direct P and S phases, as well as estimate their arrival time. The network model is trained, verified, and tested using the large-scale Southern California SCSN data set(about 2 million pieces of phase data). Through transferring learning and data enhancement methods, the model was transferred to a small scale data set of our target seismic network. We built the phase picking method for continuous waveform data using transformed model.



Multi-task Convolution Neural Network ：

Realizing seismic phase detection and arrival time estimation simultaneously.



Transfer Learning:

Transferring the model to a small-scale data set of stations in target area.

phases association

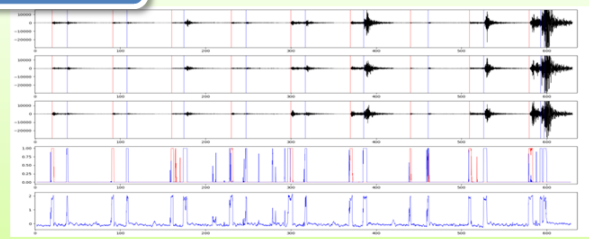
$$\alpha = v_P / v_S = \left| T_{Si} - T_{Sj} \Big/ T_{Pi} - T_{Pj} \right| + \varepsilon_{ij}$$

$$\Rightarrow \Delta T_S = \alpha \Delta T_P + \varepsilon$$

$$\hat{\alpha} = \left(\Delta T_P^T \Delta T_P\right)^{-1} \Delta T_P^T \Delta T_S \qquad \delta\bar{\alpha}_{rms} = \sqrt{\sum_{i=1}^{N} \delta\alpha_i^2 / N}$$

$$T_0 = T_P - \frac{T_S - T_P}{v_P / v_S - 1}$$

Testing Result



For SCSN dataset, the recall rate of P and S 98%; After transfer learning, for small dataset, the recall rate can reached 91%.

We use the ratio of the P arrival time difference to the S arrival time difference between pairs of stations and the least square estimation to get the average wave velocity ratio.Then according to the root mean square residuals between observed value and estimation value of wave velocity ratio should be less than a threshold, we achieved the association of multiple stations P/S phases.

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

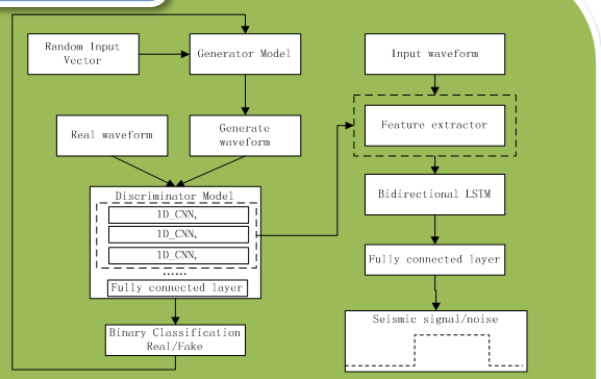
Please do not use this space, a QR code will be automatically overlayed

2.Noise Signal Recognition using GAN_LSTM

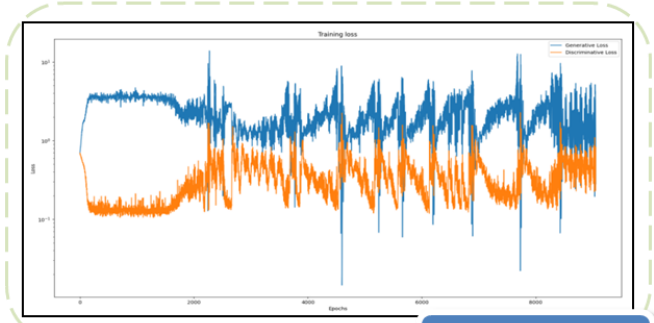
CTBT Beijing National Data Center
Beijing Radionuclide Laboratory

We proposed a method to identify the noise signal using GAN_LSTM joint network. Firstly, we trained the GAN on P 、 S phases (Pn,Sn,Pg,Lg,P,PKP) data. When the generator of GAN can generate these types of phase, the discriminator is considered to have extracted the features of phases. Then we connect the discriminator with the LSTM to form a new network. Fixing GAN network parameters, we only trained the LSTM and fully-connected layer parameters using the mixed phase and noise signals. Finally, we realized the noise signals recognition.

Model building

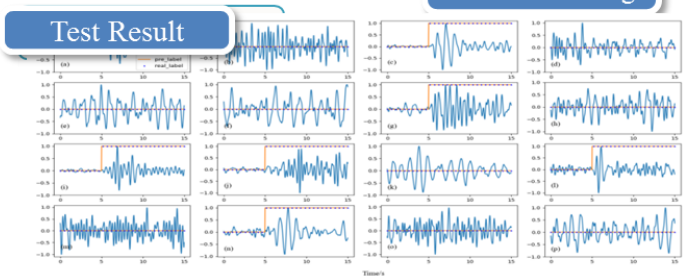


- Training GAN using seismic phase data, the discriminator of GAN can obtain the signal features of seismic phases.
- Training the GAN(discriminator)+LSTM on mixed seismic phases and noise signals to classify 2 type signals.



Model Training

Test Result



- Using beamforming data of MKAR array as the training and testing dataset, our method classified seismic phases and noise signals with a recall of 95.28% and 97.64%.
- Testing on 2019 MKAR 17060 arrivals, IDC method identified 6930 noise signals, GAN_LSTM identified 9836, 9023 among them were not used to form LEB events, account for 96.13%. Out method can reduces the false events in the automatic bulletin.

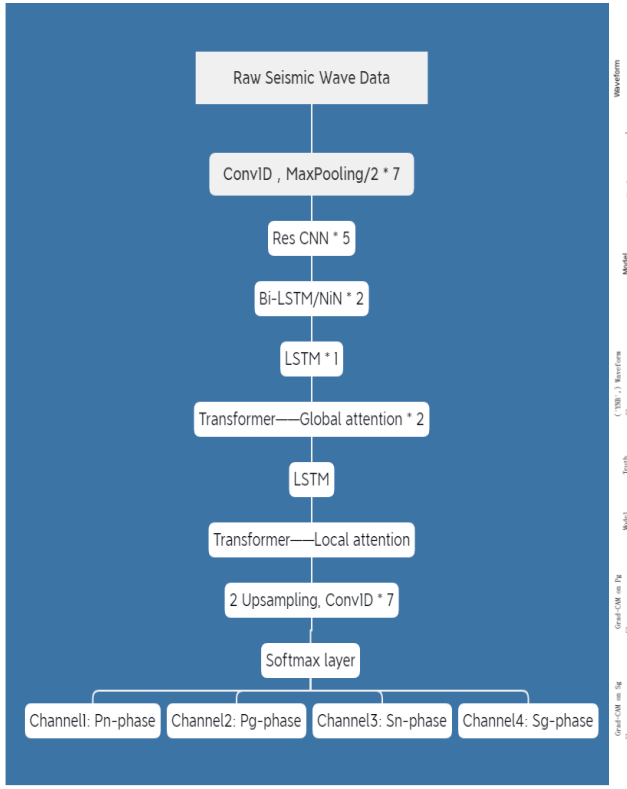
INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

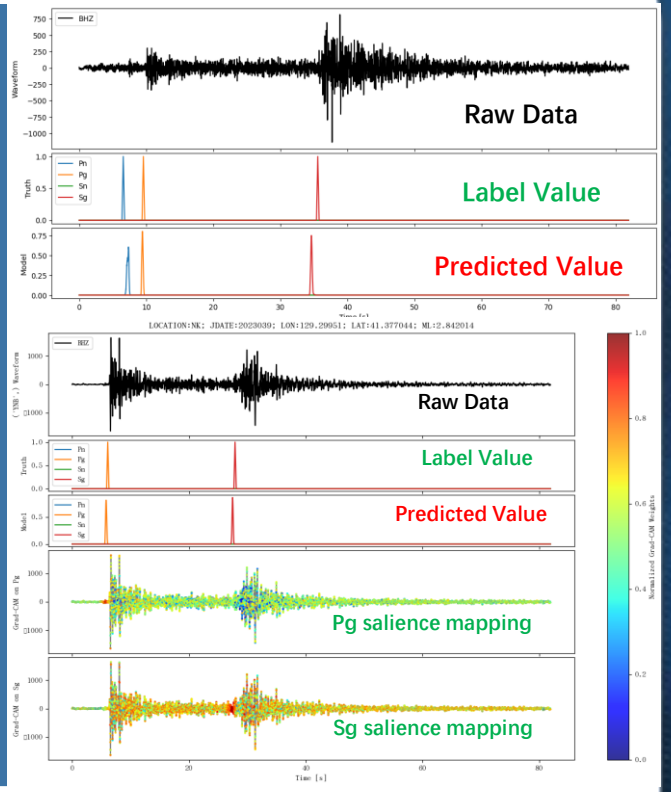
Please do not use this space, a QR code will be automatically overlayed

3.Seismic phase sequence detection based on Transformer

CTBT Beijing National Data Center
Beijing Radionuclide Laboratory

We proposed a seismic phases sequence detection method based on Transformer. The network consists of encoder and decoder. In terms of encoder structure, multi-layers Conv1D, Res CNN and LSTM were used to extracted the waveform data features. Global and local transformer were used to extracted attention information. In terms of decoder structure, the Conv1D deconvolution network was used to upsample the extracted features. Finally, the fully connected and softmax layers output the phase sequence prediction results. Grad-CAM was used to produces a heat map that highlights the important regions of seismogram by using the gradients of the target of the final convolutional layer.





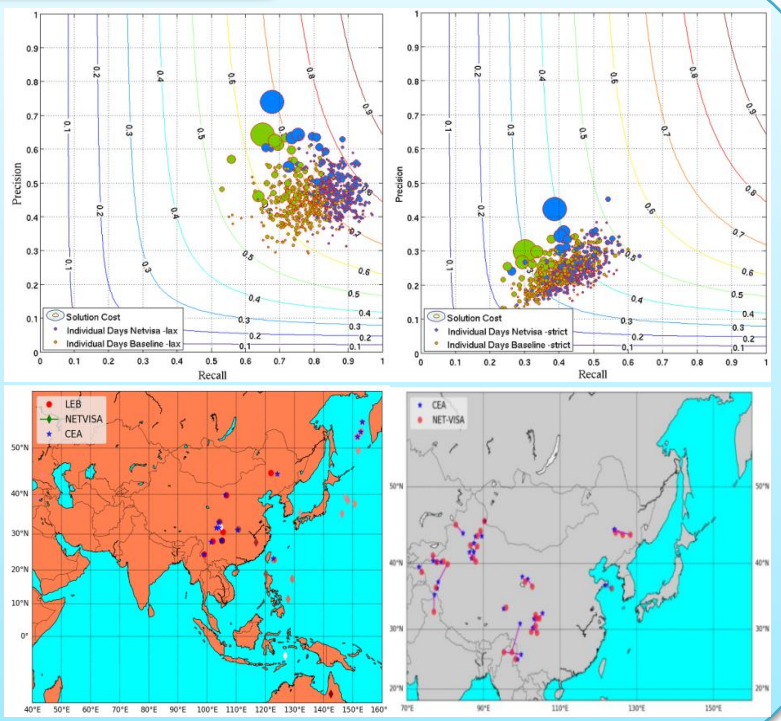
INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Please do not use this space, a QR code will be automatically overlayed

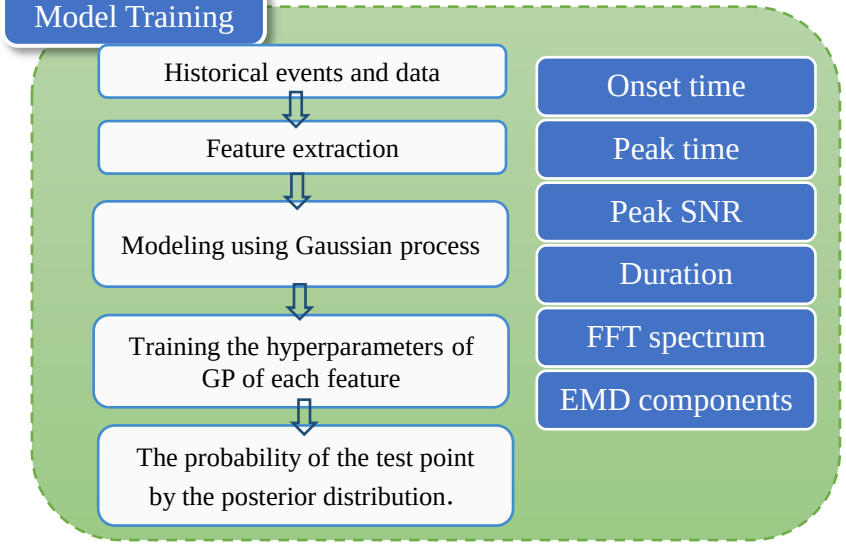
4.Event association and location based on probabilistic model

CTBT Beijing National Data Center
Beijing Radionuclide Laboratory

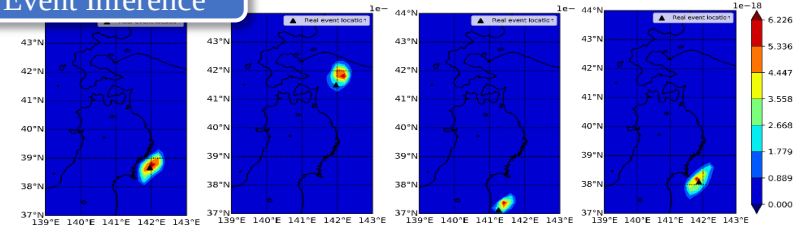
NET-VISA Evaluation



Model Training



Event Inference



- With the same precision, NET-VISA achieves a higher recall rate than GA.NET-VISA ourperforms GA for repeated small-magnitude explosion events.
- An association and location method is realized based on the probabilistic model. The Gaussian process is used to construct the probability model of waveform characteristic in time-frequency domain. An event association scheme based on global grid search is designed.

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Please do not use this space, a QR code will be automatically overlayed

Discussion and Conclusion

CTBT Beijing National Data Center
Beijing Radionuclide Laboratory

1. ML has become a research hotspot in the field of waveform monitoring data processing.
2. Training data are lacking in some categories and regions. The model generalization performance needs to be improved.
3. The interpretability and explainable problems of ML methods for verification needs to be addressed.
4. ML methods Evaluation.

References

CTBT Beijing National Data Center
Beijing Radionuclide Laboratory

[1] Li J, Wang XM, Zhang YH, et al.2020.Research on the seismic phase picking method based on the deep onvolution neural network. Chinese J. Geophys.(in Chinese),63(4):1591-1606,doi:10.6038/cjg2020N0057.

[2] Li,J;Hei,D.;Cui,G.;He,M.;Wang,J.;Liu,Z.;Shang,J.;Wang,X.;Wang,W.GAN-LSTM Joint Network Applied to Seismic Array Noise Signal Recognition. Appl.Sci.2021,11,9987.

[3] Jian Li, JuanWang, XiaomingWang, ChangshengJiang, WeidongWang, Junmin Liu. Evaluation and comparison of the results of the NET-VISA seismic event association method based on Bayesian theory. Natural Hazards, 2021.

[4] Li J, Hei D, Cui G,He M, Wang J, Liu Z, Shang J, Wang X, Wang W.2021. GAN-LSTM joint network applied to seismic array noise signal recognition. Appl Sci 11(21):9987.

INTRODUCTION
OBJECTIVES
METHODS/DATA
RESULTS
CONCLUSION

Please do not use this space, a QR code will be automatically overlayed