P3.5-792

Survey in Deep Learning Approaches for Local to Teleseismic Earthquake Detection

Maya Liu Abdel Rehem, Nimar Arora

Bayesian Logic Inc.

INTRODUCTION AND MAIN RESULTS

Deep learning models excel in earthquake detection and phase picking but struggle to generalize across regions. While they perform well over a wide range of source distances, we find that accuracy drops significantly in unseen geographic areas. This regional dependence limits their use in global applications, such as treaty monitoring.

SnT 2025
8 SEPTEMBER ONLINE DAY
9 TO 12 SEPTEMBER AT HOFBURG PALACE, VIENNA & ONLINE
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

Survey in Deep Learning Approaches for Local to Teleseismic Earthquake Detection

Maya Liu Abdel Rehem, Nimar Arora

P3.5-792

Introduction

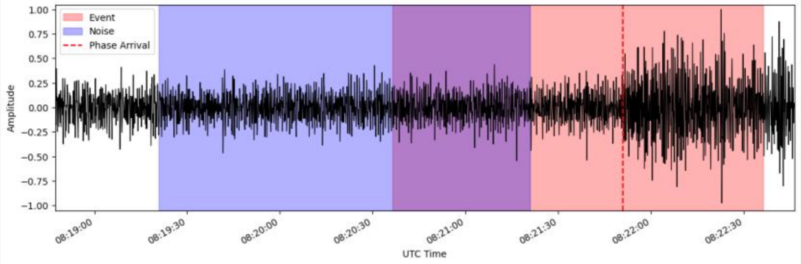
Deep learning models have proven highly effective for earthquake detection and phase picking. Yet their generalization capabilities remain incompletely understood—particularly regarding source distance and geographic transferability.

Our work reveals a critical limitation: while these models demonstrate robust performance across varying source distances, they fail to generalize to previously unseen source regions. This regional specificity severely constrains their applicability for treaty monitoring applications, which must reliably detect events from regions without prior training data.

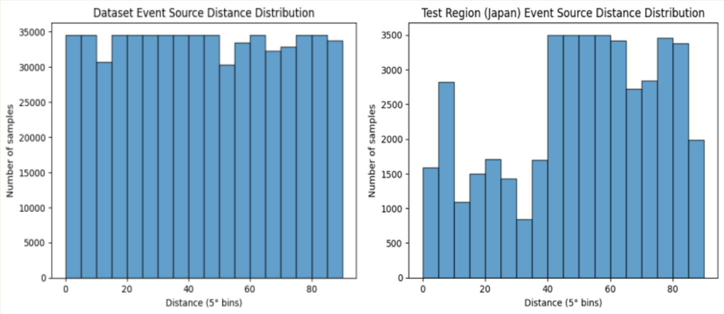
Dataset

Our dataset is derived from the MLAAPDE database [1], with ~800,000 waveforms from global events distributed uniformly from local to teleseismic epicentres.

Half of the samples include phase arrivals—event samples. The other half are noise samples, which are preceding waveform data from known P, Pn, or Pg arrivals that have no catalogued phase arrival within the sample window.



Events originating from Japan were held out from the training data and used for the "test region" set.



Methods/Models

Four benchmark model architecture approaches:

- CNN-RNN, e.g. CRED
- CNN-RNN with Transformer, e.g. EqTransformer
- U-Net, e.g. PhaseNet
- U-Net with Transformer

Trained on 2-minute-long event and noise waveforms with no bandpass filtering.

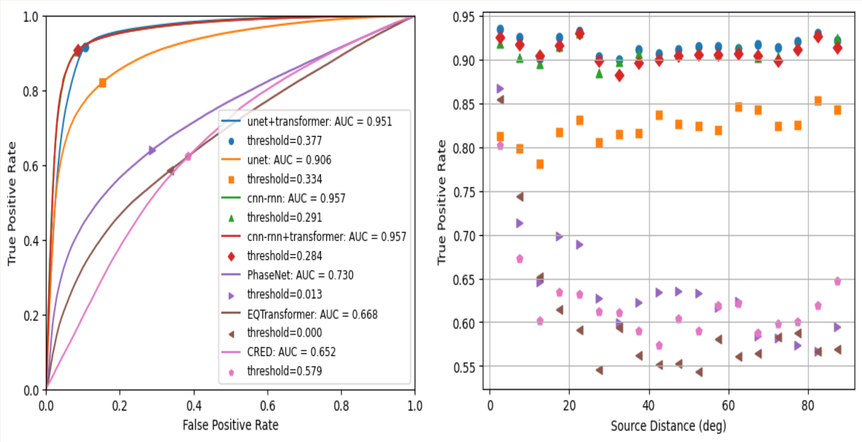
Compared the former with preexisting deep learning models of similar architecture, including EqTransformer [2], PhaseNet [3], and CRED [4], as well as classic STA/LTA methods.

Results

Each model demonstrated uniform detection performance across all distance ranges.

Original versions of EqTransformer, PhaseNet, and CRED, which are all trained on local events, demonstrate poor overall detection performance and a clear drop in performance on distances farther than local.

Hence, Deep Learning methods should be trained on datasets balanced with respect to source distance.



Bayesian Logic Inc.

SnT 2025
8 SEPTEMBER ONLINE DAY
9 TO 12 SEPTEMBER AT HOFBURG PALACE, VIENNA & ONLINE
CTBT: SCIENCE AND TECHNOLOGY CONFERENCE

Survey in Deep Learning Approaches for Local to Teleseismic Earthquake Detection

Maya Liu Abdel Rehem, Nimar Arora

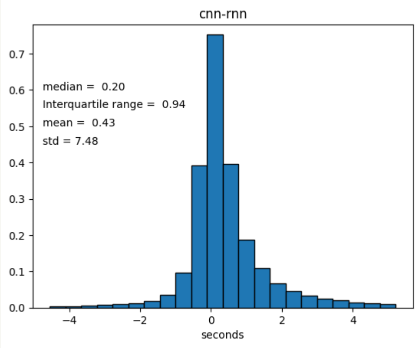
P3.5-792

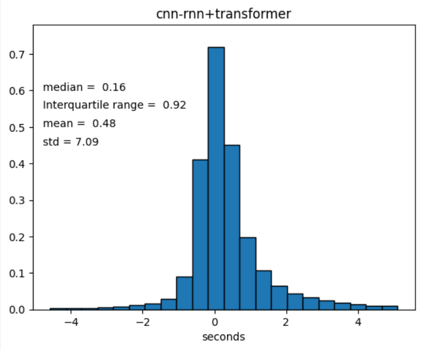
Results cont.

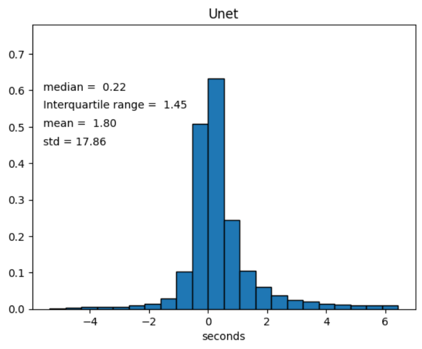
All models demonstrated a significant drop in performance on events from the held-out test region (Japan).
Deep learning methods outperform classic STA/LTA at detection on regions on which they have been trained. However, STA/LTA performance remains consistent for all regions.
For unseen geographical regions, deep learning methods perform worse at detection than classic STA/LTA methods.

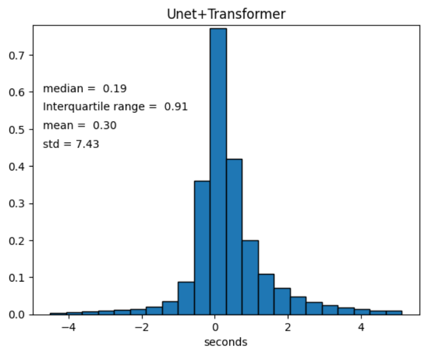
	Test		Test Region
Model/Method	*F1*	*Recall*	*Recall*
U-Net+Transformer	0.91	0.92	0.63
U-Net	0.83	0.82	0.54
CNN-RNN+Transformer	0.91	0.91	0.61
CNN-RNN	0.91	0.91	0.60
STA/LTA	0.69	0.68	0.70

Each benchmark model demonstrates similar overall phase picking residual distributions except for the U-Net-based model, which has the most significant standard deviation.

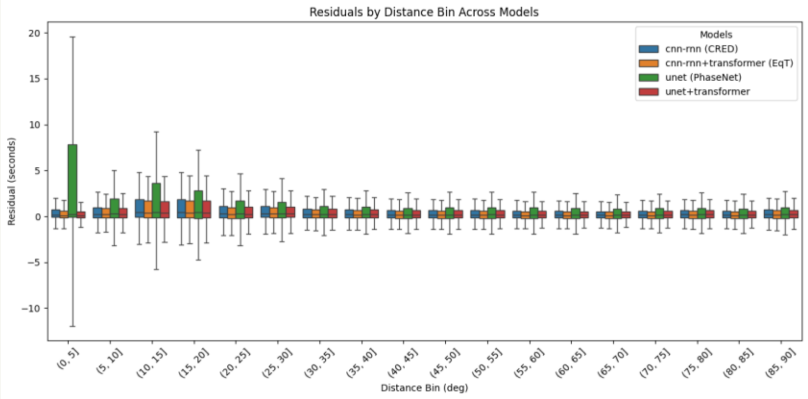








Across different distance ranges, the pick residuals remain uniform, except for the U-Net model, which exhibits the most significant residuals for local events.



Conclusion

Overall, all benchmark architectures maintain robust detection accuracy across distance ranges when trained on local to teleseismic events, demonstrating the importance of diverse training data for distance generalization.

The addition of transformers demonstrated an increase in detection performance for the U-Net-based model. Such an addition did not show any meaningful difference in the performance of the CNN-RNN-based model.
A generalization gap remains, preventing models from accurately predicting earthquakes in unseen geographical locations. Complex models are prone to overfitting and may be learning regional characteristics rather than a general representation of earthquakes. The results suggest that classic STA/LTA methods remain more effective and simpler for the task of earthquake detection in regions where deep learning models cannot or have not been trained. This is of critical importance for treaty monitoring purposes.

References

[1] Hank M. Cole, William L. Yeck, Harley M. Benz; MLAAPDE: A Machine Learning Dataset for Determining Global Earthquake Source Parameters. Seismological Research Letters 2023;; 94 (5): 2489–2499. https://doi.org/10.1785/0220230021

[2] Mousavi, S.M., Ellsworth, W.L., Zhu, W. et al. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. Nat Commun 11, 3952 (2020). https://doi.org/10.1038/s41467-020-17591-w

[3] Weiqiang Zhu, Gregory C Beroza, PhaseNet: a deep-neural-network-based seismic arrival-time picking method, Geophysical Journal International, Volume 216, Issue 1, January 2019, Pages 261–273, https://doi.org/10.1093/gji/ggy423

[4] Mousavi, S.M., Zhu, W., Sheng, Y. et al. CRED: A Deep Residual Network of Convolutional and Recurrent Units for Earthquake Signal Detection. Sci Rep 9, 10267 (2019). https://doi.org/10.1038/s41598-019-45748-1

Bayesian Logic Inc.