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. Training on a wide range of event source distances is key to robust detection performance. Training data that spans a diverse set of geographical regions allows for robust detection performance in unseen geographical regions.

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

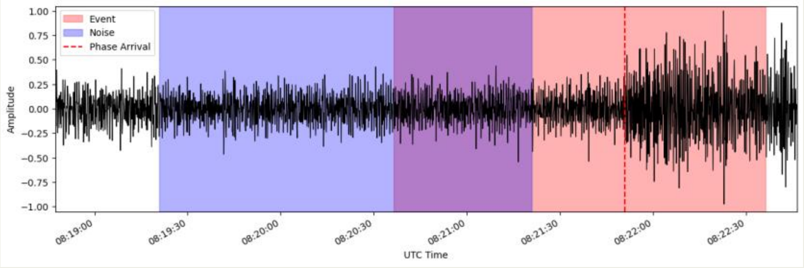
Introduction

Deep learning models have demonstrated strong performance in earthquake detection and phase picking, yet their robustness depends critically on training data composition. By leveraging a large, distance-balanced dataset derived from MLAAPDE, we evaluate multiple benchmark architectures and examine how dataset diversity and model complexity influence detection performance from local to teleseismic distances and events from geographical unseen during model training.

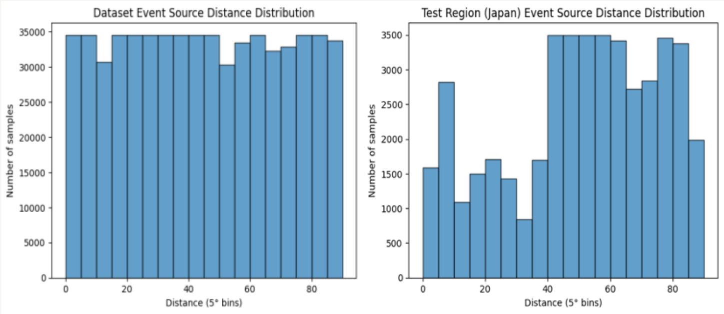
Dataset

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 (M=million):

- CNN-RNN (0.44 M parameters), e.g. CRED
- CNN-RNN with Transformer (1.1 M parameters), e.g. EqTransformer
- U-Net (0.66 M parameters), e.g. PhaseNet
- U-Net with Transformer (1.3 M parameters)

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

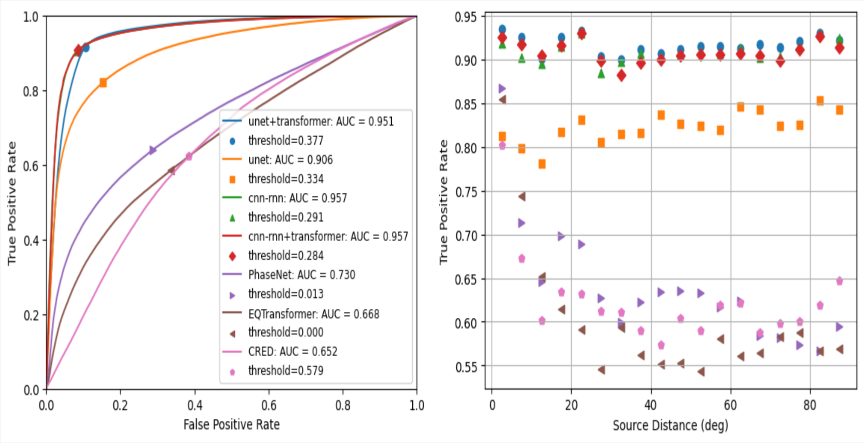
Compared the former with preexisting deep learning models of similar architecture, including EqTransformer [2], PhaseNet [3], and CRED [4], as well as classic STA/LTA methods.

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.

Survey in Deep Learning Approaches for Local to Teleseismic Earthquake Detection

Maya Liu Abdel Rehem, Nimar Arora

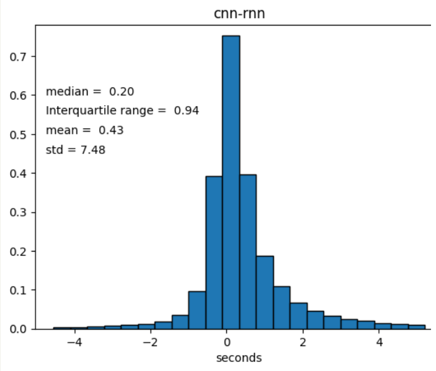
P3.5-792

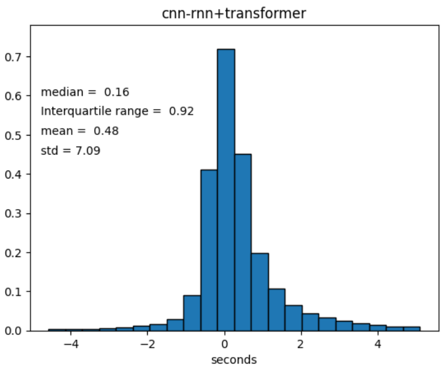
Results cont.

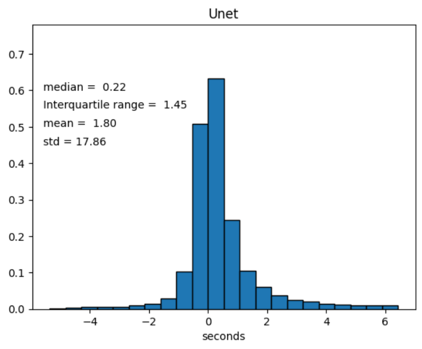
All models demonstrated consistent performance on events from the held-out test region (Japan).
Deep learning methods outperform classic STA/LTA at detection on regions on which they have been trained and have not been trained.

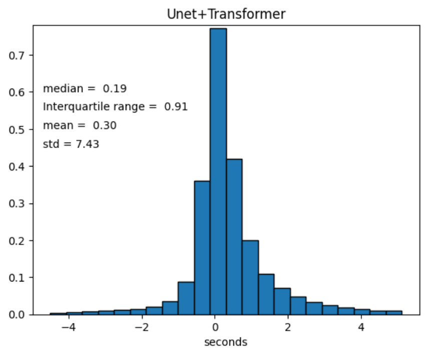
Model / Method	Test F1	Test Recall	Test Region F1	Test Region Recall
U-Net + Transformer	0.91	0.92	0.94	0.94
U-Net	0.83	0.82	0.87	0.81
CNN-RNN + Transformer	0.91	0.91	0.93	0.92
CNN-RNN	0.91	0.91	0.93	0.92
STA/LTA	0.69	0.68	0.80	0.75

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

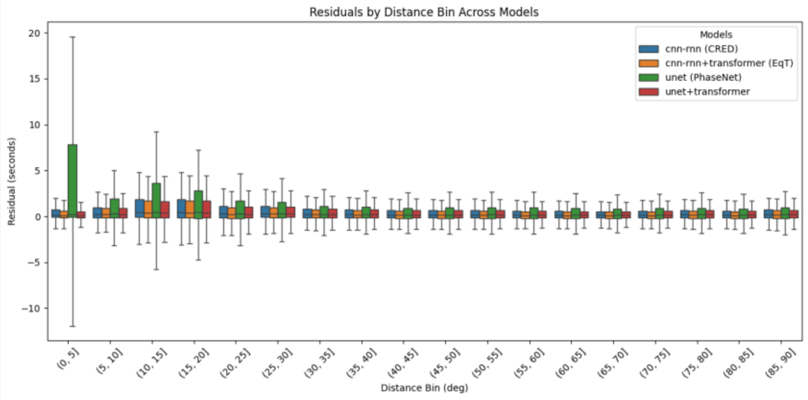








Across different distance ranges, the pick residuals remain uniform, except for the U-Net model, which exhibits the most significant residuals for local events.



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.
When tested on events from unseen geographical regions, detection performance remained robust and still outperforms STA/LTA. Deep learning models across the literature have generally seen a drop in performance in detection for unseen geographical regions. Our results suggest that a diverse set of geographical regions improves the model's capability to generalize to these regions. Across all benchmark architectures, performance gains were not strictly correlated with increasing model complexity or parameter count implying meaningful seismic inference does not require highly overparameterized models when training data are sufficiently diverse.

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.