

一元化震源データを利用した深層学習検測器の再学習の試み

直井誠（北大理）・溜渕功史（気象研）・大柳修慧（京大防災研），加藤慎也（東大地震研）

近年、深層学習を用いた走時検測器が多数作成されており、主に海外のデータを用いた訓練済みモデルが公表されている（例えば、Zhu and Beroza 2019; Mousavi et al. 2020）。これらはしばしば再学習なしで他の地域の地震カタログ作成に用いられ、多くの論文が出版されるようになってきているが、実際に使用するデータセットを用いて再訓練や転移学習を行えば、より高いパフォーマンスが発揮できる場合がある（e.g., Kim et al. 2023）。従って、日本の気象庁一元化震源のような、ユーザー数が多いデータに関しては、それらをもとに学習したモデルが公開・共有されることが望ましく、これによって関連研究が促進されると考えられる。そこで本研究では、広く利用されている深層学習検測器 PhaseNet (Zhu and Beroza 2019) の再訓練を、気象庁一元化検測値をもとに実施し、カリフォルニアのデータで訓練されたオリジナルモデルとの性能を比較した。

訓練には、2018–2020 の 3 年分のデータを使用し、気象庁が公開している検測値情報をもとに訓練を行った。2020 年 9 月 1 日からは S-net 記録の検測値がカタログ作成に利用されはじめており、これらのデータも訓練に含まれている。S-net 波形については Takagi et al. (2019) の方位情報をもとにして、NS, EW, UD 成分に Rotation してから使用した。この期間に得られた 10,546,390 個の 3 成分波形から、1) P と S 両方の読み取り値がある、2) 3 成分揃っている、3) P, S 両方で手動チェックが行われている、の 3 つの条件を満たす 3 成分波形 1,295,195 個を選別して訓練に使用した。このうち、80% をネットワークの訓練データ、20% を検証データとして使用した。

上記の手順で一元化検測値と波形をもとに訓練した PhaseNet モデル、並びに Zhu and Beroza (2019) のオリジナルのモデルを、訓練に使用していない 2021 年のデータに適用して、Precision-recall curve を作成して性能を評価したところ、顕著に性能が向上したことが確認された。向上の度合いは Hi-net 観測点のデータに対しては限定的で、S-net や DONET 点のデータに対して大きかった。また、走時残差分布からは、Zhu and Beroza (2019) のオリジナルモデルを利用した場合には気象庁検測値に比べてやや後ろを読む傾向がある一方で、再訓練したモデルではこのバイアスが軽減したことが確認できた。