

畳み込みニューラルネットワークを用いた 地震波形検測 (分類)

工藤 祥太*1, 下條 賢悟*1, 溜瀧 功史*2
*1気象庁 (大阪管区気象台地震火山課) *2気象研究所

はじめに

気象庁は2016年4月から、自動震源決定手法であるPF法(溜瀧・他, 2016, 駿震時報, Tamaribuchi, 2018, EPS)の運用を開始した。
IPF法は、震源精度や検知能力の大幅な向上をもたらしたものの、依然としてノイズの誤検測によって震源を誤決定する事例が多い(図1)。地震/ノイズの識別によりノイズの誤検測を減らすことができれば、PF法の震源精度をさらに向上させることが期待される。
地震/ノイズの識別手法として、畳み込みニューラルネットワーク(CNN)による方法が提案されている(Ross, 2018, BSSA)。本研究では、一元化震源カタログの検測値、気象庁の保有する地震波形データを用いて(Ross, 2018, BSSA)に準拠したCNNモデルを学習させ、地震/ノイズの識別能力を評価する。また、このモデルを用いてPF法にノイズ識別のプロセスを追加し、PF法の震源がどのように変化するかを検証する。

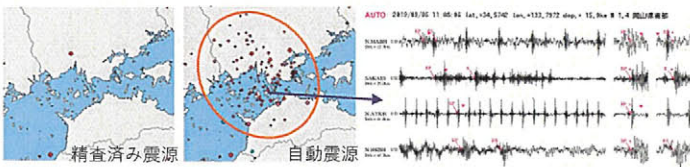


図1. (左) 人手による精査済み震源 (中) PF法による自動震源 (右) ノイズの誤検測事例

データ/前処理

データ

モデルの学習には、一元化震源カタログの検測値、気象庁が保有する地震波形データのうち、以下の条件を満たすデータを用いた。

- 期間: 2013年8月~2018年10月
- 観測点: 一元化震源決定に用いられる全観測点(1800点)
- 震源精度: K, k登録(高精度かつ手動で決定した)震源
- 地震計種別: 速度計
- 成分: N/S, E/W, U/Dの3成分
- 周波数: サンプリング周波数100Hz

前処理

- オフセット・トレンド除去、およびパワー特性のバンドパスフィルタ(1~20Hz)を適用
- P相の4秒前を中心とした4秒間をノイズ、P相・S相を中心とした4秒間をP・Sとして切り出し(図2)。ただし、切り出しの中心位置には±0.5秒の一樣ランダムな振動を加える
- 4秒間の振幅の最大絶対値で全体を割り規格化
- ノイズ/P/Sのラベルを付した3成分・4秒間波形がノイズ: 10万個, P: 10万個, S: 10万個(計30万個)となるようにデータセットを作成。同様に各々5万個(計150万個)のデータセットも作成
- 訓練・検証・テストデータを6:1:1の割合で分配

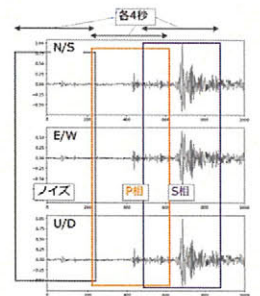


図2. ノイズ・P・Sの切り出し

モデル

層	L1	L2	L3	L4	L5	L6	L7
フィルタ数/ユニット数	32	64	128	256	200	200	3
フィルタサイズ	21	15	11	9			
活性化関数	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	softmax

表1. モデルの構成。C: 1D-CNN, B: Batch Normalization, P: Max Pooling, FC: Fully Connectedを表す。

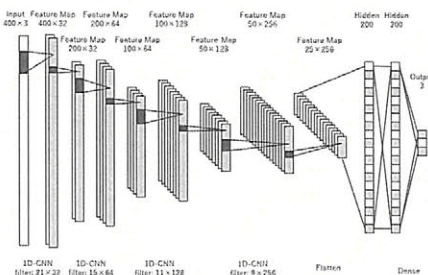


図3. モデルの模式図

損失関数	categorical crossentropy
最適化	Adam (学習率: 0.001)
バッチサイズ	480
モデル採用	5エポック連続で損失が改善がなければ早期終了。検証データに対する損失が最小となるモデルを採用

表2. モデルの学習設定

モデルによる分類例

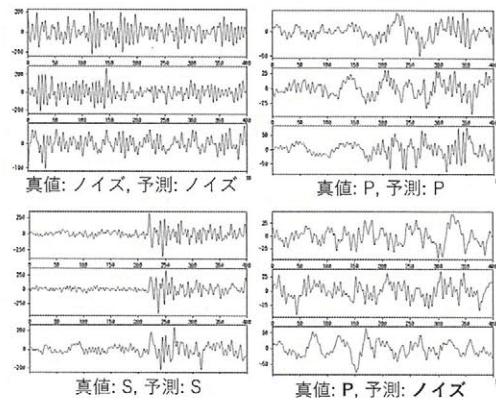


図4. 学習済みのモデルによる波形の分類例。右下は誤分類の例。それ以外は正しく分類できた例。

学習済みのモデルに、テストデータを入力し、分類させた例を示す(図4)。なお、モデルの予測値はノイズ: 0.5, P: 0.2, S: 0.3のように総和が1となる確率値として出力され、3クラスのうち最も確率の高いクラスに分類する。

モデルの性能評価

データ数30万、および150万で学習させたモデルにテストデータを入力し、それぞれ分類性能を比較するための指標を求めた(表3, 図5)。データ数30万, 150万のいずれも各種指標が95%を上回っているが、データ数150万で学習させたモデルの方が性能が良い。

	総データ数30万			総データ数150万		
	ノイズ	P	S	ノイズ	P	S
Accuracy	0.966			0.975		
Precision	0.972	0.959	0.966	0.970	0.979	0.976
Recall	0.960	0.970	0.967	0.978	0.971	0.976

表3. モデルの予測値のうち最も確率の高いクラスに分類した場合の指標。総データ数30万, 150万の場合についてそれぞれ算出。

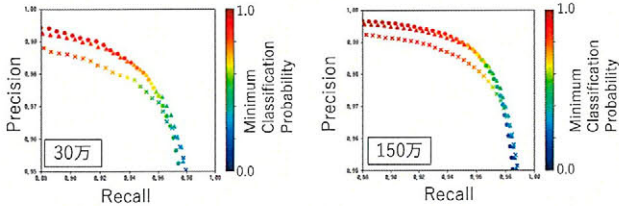


図5. Precision-Recallトレードオフ曲線。×: ノイズ, ●: P, ▲: S. ノイズ, P, Sについてそれぞれ, 分類の閾値を変化させていったときのRecall, Precisionを算出しプロット。

5
6
7
8

PF法への適用(1)

データ数150万での学習済みモデルを、ノイズ除去器として、PF法のプロセスに追加し(図7)、震源の変化を検証した。

方法

AR-AIC等を用いた従来の地震波検測(一次検測)による検測値を中心とした4秒間の波形データをモデルに入力すると、ノイズ/P/Sの確率が出力される。このうち、ノイズの確率が50%以上である検測値については一次検測値から除去する。そして、ノイズ除去済みの検測値から震源決定を行う。

検証

PF法, PF法(ノイズ除去)の両方で震源決定を行った。対象期間は2019年4月1日の0時~24時とした。低周波イベントを除いた一元化震源を真値とし、一元化震源と震源時が±5秒, 震源の緯度・経度がそれぞれ±0.5度以内の場合に震源が一致するとした。また、各種法による震源のうち一元化震源にないものは過剰, 一元化震源にあるがその手法による震源にはないものを不足として扱う。

結果

表4(上)によると、ノイズ除去により検測値は3/5程度に削減できている。震源数は7%程度削減されている。表4(下)によると、ノイズ除去により一致, 不足はほぼ変化なしである一方、過剰を14%程度削減することができている。また、図8から、台湾周辺と秋田県の周辺の震源が、ノイズ除去により削減されていることがわかる。

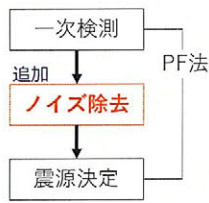


図7. ノイズ除去を加えたPF法の流れ

連続波形への適用

2019年6月18日の山形県沖の地震(M6.7)近傍の地震観測点(TU.AT2)での、本震直前より15分間の連続波形に対し、データ数150万で学習させたモデルによる分類を適用した(図6)。余震が連発しているような事例であるが、おおむね適切にP相, S相を分類できている。

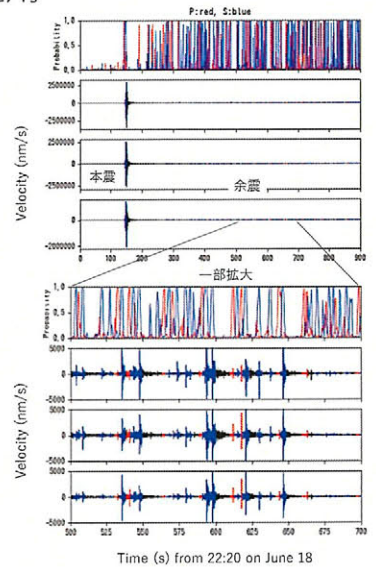


図6. 連続波形に対するモデルの適用例。赤がP相, 青がS相を示す。モデルが95%以上の確率を予測したクラスを相として選択。

PF法への適用(2)

	検測値数	震源数
一元化震源	-	465
PF法	1415736	590
PF法(ノイズ除去)	833671	550

	一致	過剰	不足
PF法	402	188	63
PF法(ノイズ除去)	399	151	66

表4. (上) 一元化震源(低周波イベントを除く), PF法, PF法(ノイズ除去)の検測値数と震源数。(下) 一元化震源と比較したときの、震源の一致, 過剰, 不足の数。

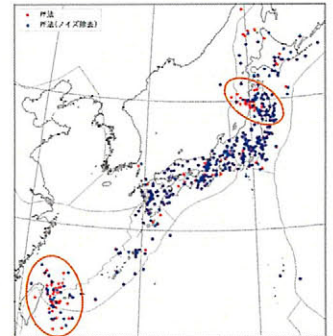


図8. PF法による震源(赤)とPF法(ノイズ除去)による震源(青)のプロット

課題・今後の展望

ノイズ除去によりPF法の過剰な誤震源は削減されたものの、依然として過剰な誤震源が多く残っている。これをさらに削減するためには、モデルの汎化性能と精度を向上させ、ノイズ判定の確率の閾値を調整する必要がある。モデルの性能向上には、人間でも地震/ノイズの判別が難しいようなエッジケースを追加することが必要である。また、一次検測をPF法の手法で行うのではなく、モデルのノイズ/P/S分類のみによって行った場合に震源精度がどのように変化するかを検証したい。さらに、計算時間を大幅に増やすことなくモデルを自動震源決定のルーチンシステムに組み込む方法も検討したい。

