

畳み込みニューラルネットワークによるボーリング基盤深度の回帰予測モデル構築

アジア航測（株）正会員 ○泉田 温人

アジア航測（株）非会員 小宮 佑登

（公財）鉄道総合技術研究所 正会員 田中 浩平 坂井 公俊 名波 健吾

1. 目的

地表から数 10 m 程度の深度に堆積する表層地盤の状況によって、例えば地震時には地震動の局所的な増幅による構造物の被害、地盤の液状化、斜面崩壊等の災害が発生することがある。この表層地盤の状況を把握するために、ボーリング調査を実施するのが一般的であるが、その実施数量が限られていることから、面的な変動までを把握することは困難である。一方で、比較的容易に入手できる面的な地形データ（標高、地形区分等）とその下に広がる表層地盤の関係については、近年十分なデジタルデータが整備されつつあるものの、その関係性を検討した事例は限られる。そこで、本研究では近年急速に発展している畳み込みニューラルネットワークを採用した深層学習による回帰モデルを、画像状の地形情報からその地点の表層地盤における基盤深度を予測するという問題に適用し、その有効性を確認した。

2. ネットワーク構造

畳み込みニューラルネットワーク（CNN）は深層学習における演算方法の一種であり、主に画像処理分野に適用されることが多いネットワークである。CNN では畳み込み演算により入力画像を徐々に縮退させることで、画像等の多次元構造を持つデータから効率的に特徴量を抽出することができる。本研究の基盤深度予測モデルは、画像状に加工した複数種類の入力データから各々抽出した特徴量を、更に一層の全結合層で演算することで基盤深度の予測値を得る構造を持つ（図 1）。特徴量の抽出には ResNet50¹⁾を使用した。

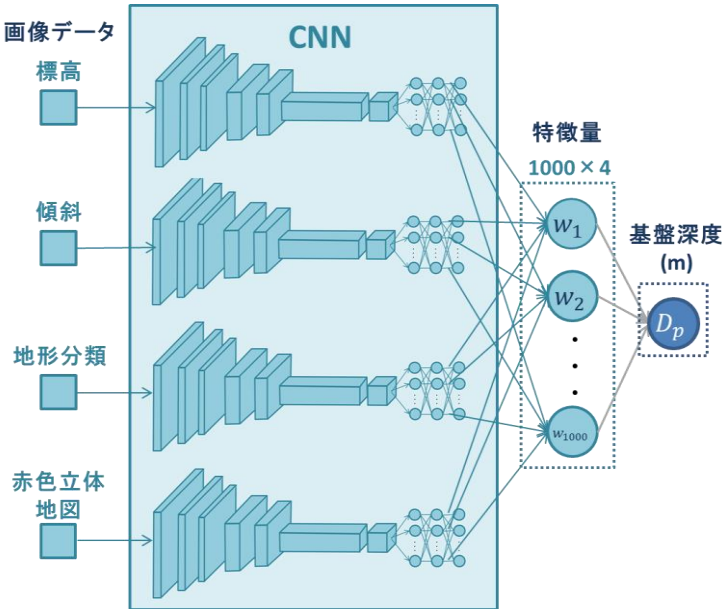


図 1 ネットワーク構造

3. 地域・データ

平野部および盆地部から基盤深度の記載を含むボーリングデータを収集した（表 1）。基盤深度予測モデルの入力データは標高、傾斜、地形分類を表すラスターデータおよび、地形立体表現手法「赤色立体地図」（特許第 3670274 号）による地図画像である。これらの画像データをボーリング地点の周囲 1000 m の範囲で矩形状に切り出し、ボーリング地点の基盤深度と組み合わせたデータセットを作成した。全地点のデータをランダムに二分割し、予測モデルの構築に使用する学習データセットと、未知地点における予測値の評価に使用するテストデータセットに振り分けた。地形分類は国土地理院が公開している「地形分類（自然地形）」²⁾を用い、7 種類の地形に再分類して入力した。地形分類はカテゴリを表す質的データであるため、他の量的データと同時にモデルに入力できるよう、二値のダミー変数による数量変換を行った。赤色立体地図は RGB 画像として学習モデルに入力した。

表 1 ボーリングデータセット

データ	件数
東京都多摩地区～東部	13, 061
福岡県福岡市周辺	2, 089
千葉県柏市周辺	1, 298
山梨県甲府市周辺	130
学習データセット	8, 323
テストデータセット	8, 255
合計	16, 578

キーワード ディープラーニング、回帰分析、ボーリング、地形分類、沖積層

連絡先 〒215-0004 神奈川県川崎市麻生区万福寺 1 丁目 2-2 アジア航測株式会社 TEL 044-967-6458

4. 予測モデルの構築

学習データセットを用い、表2に示す条件で予測モデルを構築した。ネットワークの損失関数として平均絶対パーセンテージ誤差（MAPE）を使用したため、学習データセットにおいてこの指標が最小となるようにモデルのパラメタが最適化される。学習時には入力画像をランダムに回転・反転させる処理を加えることで実質的な学習データ数を増量した。また、重回帰分析を実施し、入力変数のうち独立である11変数を説明変数 X_n とし、偏回帰係数 b_n を求めることで式(1)のように基盤深度 Y を推定した。

$$Y = b_0 + b_1X_1 + b_2X_2 + \cdots + b_nX_n$$

・・・(1)

5. 結果

各々の手法による予測モデルの精度指標として、平均二乗誤差（RMSE）、平均絶対パーセンテージ誤差（MAPE）、標本相関係数（r）を用いた。これらの指標は各予測値 Y_i 、各真値 A_i 、データ件数 n を用いて式(2)、式(3)、式(4)のように表される。

平均二乗誤差

$$RMSE = \frac{1}{n} \sqrt{\sum (Y_i - \hat{Y}_i)^2}$$
・・・(2)

平均絶対パーセンテージ誤差

$$MAPE = \frac{1}{n} \sum \left| \frac{A_i - Y_i}{A_i} \right|$$
・・・(3)

標本相関係数

$$r = \frac{\sum (Y_i - \hat{Y}_i)(A_i - \hat{A}_i)}{\sqrt{\sum (Y_i - \hat{Y}_i)^2} \sqrt{\sum (A_i - \hat{A}_i)^2}}$$
・・・(4)

テストデータセットに対するCNNによる予測結果は、真値との絶対誤差を表すRMSEが10.83 m、相対誤差を表すMAPEが44.4%だった。この結果は、重回帰分析による予測値よりも高い精度だった（表3）。CNNの予測モデルでは学習データセットに対する精度指標がテストデータセットよりも良く、学習データの基盤深度に適合するようにパラメタの最適化が進んだことがわかる。CNNによる予測モデルでは、データの分割をランダムに行った場合でも、どのデータを学習させるかによってモデルのパラメタが大きく変動する特徴が示唆される。これは、重回帰分析において二つのデータセットの精度指標が極めて近いことと対照的である。

テストデータセットの予測値と真値の頻度分布を比較したグラフを図2に示す。CNNによる予測値のグラフはより真値の分布に近い形状を示した。ところが、基盤深度の真値が5 m以下あるいは30 m以上である地点では予測精度が低下した。そのような範囲の基盤深度を持つ学習データが相対的に不足していることが一因であると考えられる。

6. まとめ

画像状データを入力とするCNNの予測モデルにより、工学的基盤深度をRMSE10.83 m、MAPE44.4%程度の精度で推定するモデルを構築した。この結果は、ボーリング地点の入力データを変数とした重回帰分析のモデルよりも、高い精度となっていることを確認した。今後は他地域への適用やデータ加工方法の検討により、本手法の有効性をさらに確認する。

謝辞 本研究の一部は、国土交通省の鉄道技術開発費補助金を受けて実施した。

参考文献

1) Kaiming He ほか, “Deep Residual Learning for Image Recognition”, arXiv:1512.03385 (2015).
2) 国土地理院 ベクトルタイル「地形分類」: https://www.gsi.go.jp/bousaichiri/lfc_index.html

表2 学習条件

項目	設定値
学習回数	500 エポック
最適化関数	Adam
損失関数	MAPE
画像サイズ	224×224 ピクセル

表3 予測モデルの精度指標

データセット	予測モデル	RMSE	MAPE	r
学習	CNN	8.35 m	21.2%	0.82
	重回帰分析	11.53 m	77.0%	0.47
テスト	CNN	10.83 m	44.4%	0.63
	重回帰分析	11.61 m	76.1%	0.47

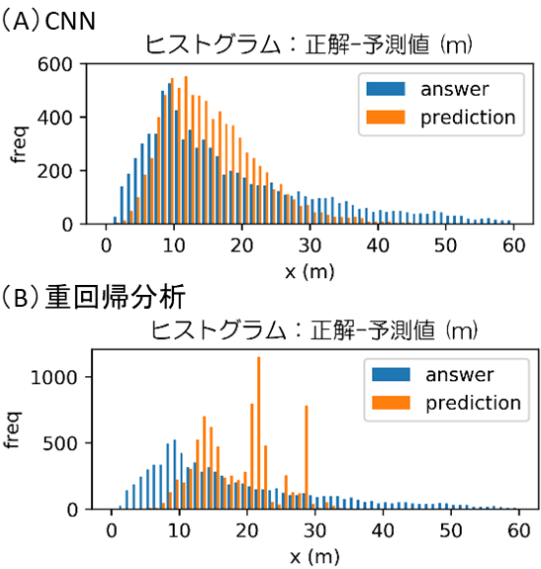


図2 予測値と真値の頻度分布