畳み込みニューラルネットワークによる レーダ雨量を用いた河川水位予測 RIVER WATER LEVEL PREDICTION BY CONVOLUTIONAL NEURAL

NETWORK USING RADAR RAINFALL

荒木健¹・箱石健太²・一言正之³・島本卓三⁴・房前和朋⁴ Takeru ARAKI, Kenta HAKOISHI, Masayuki HITOKOTO, Takuzo SHIMAMOTO and Kazutomo FUSAMAE

¹正会員 工修 日本工営株式会社(〒102-8539 東京都千代田区九段北1-14-6)
²学士 株式会社コーエイシステム(〒102-0083 東京都千代田区麹町4-2)
³正会員 博(工) 日本工営株式会社(〒300-1259 茨城県つくば市稲荷原2304)
⁴正会員 元国土交通省九州地方整備局 九州技術事務所(〒830-8570 福岡県久留米市高野1-3-1)

A river water level prediction model was proposed by applying Convolutional Neural Network using radar rainfall. Spatial distribution of hourly rainfall, accumulated rainfall and time series of water level or change of water level were used as input data. Model output is 60 minutes prediction of water level. 4 flood events out of 30, selected by peak water level, were used as validation target. NS coefficient was above 0.9 in all 4 flood events and it was confirmed that proposed model shows good accuracy. It was suggested that using both rainfall and water level results in better accuracy than using them separately. It was also suggested that setting too wide spatial range as input data causes loss of accuracy and it is better to set input data range with reference to the basin boundary.

Key Words : river water level prediction, convolutional neural network, radar rainfall, real-time prediction

1. はじめに

洪水時の河川水位予測は、住民の円滑な避難などのために重要である.予測手法としては降雨-流出過程を表現する物理型モデルが全国の河川で適用されているが、 予測の過程に様々な誤差が内在する.また、統計的手法の一種として、近年、人工知能(AI)を用いた河川水位 予測の研究が進んでいる.既往研究では、地点雨量や流 域平均雨量などを入力データとして、予測地点の水位や 水位変化を予測するモデルが提案されている^{1).2)}.しか しながら、こうした手法では流域内に十分な密度の観測 データが得られなければ、満足な予測精度が得られない 可能性が高い.

一方,画像分類などのAIでは,畳み込みニューラル ネットワーク(Convolutional Neural Network, CNN)^{3),4} の発展が著しい. CNNは入力画像の空間的な特徴,つま り隣り合うピクセル同士の関係を失うことなく取り扱え る手法であることから,レーダ雨量などの降雨の面的な 分布情報の活用に適していると考えられ、降雨の発生予 測に関する研究⁵⁾などにも適用され始めている.

そこで本研究では、レーダ雨量の面的な分布情報の活 用に着目し、河川水位予測の精度向上・適用性向上に向 けた検討を行った.具体的には、CNNを用いてレーダ雨 量データから河川水位を予測するモデルを構築した.構 築したモデルにより対象流域の降雨強度や累積降雨量の 二次元分布、水位や水位変化を入力データとし、60分先 の水位を予測した.複数の出水における精度検証を行 い、河川水位のリアルタイム予測に向けた検討を行っ た.

2. 畳み込みニューラルネットワークの概要と河

川水位予測への適用方法

(1) 畳み込みニューラルネットワークの概要 畳み込みニューラルネットワークは,畳み込み層と プーリング層という二つの層を含む構造のニューラル ネットワークであり,深層学習の一種である.二次元 データの扱いを得意としており,近年の画像認識AIの発 展を後押ししている.

本検討では、畳み込みニューラルネットワークの特徴 である畳み込み、プーリングに加え、ドロップアウトや 配列変換、全結合といった処理からなるモデルを構築し た. 各処理の概要は以下の通りである.

a) 畳み込み

畳み込みでは、一定領域の情報をひとつに集約して接続し、空間情報(隣り合うセル同士の関連)が維持された特徴マップを作成する(図-1).

畳み込みは次式の積和計算で表わされ、画像認識において画像内から線、輪郭、形状などといった特徴を抽出 する手順にあたる.フィルタにより畳み込む領域が指定 され、フィルタ数分だけ特徴マップが作成される.

$$u_{ij} = \sum_{p=0}^{H-1} \sum_{q=0}^{H-1} x_{i+p,j+q} h_{pq}$$
(1)

ここで、xは画像セルの値、uは畳み込みの計算結果、H はフィルタのセルサイズ、hはフィルタセルの値、i,jは 画像セルのid、p,qはフィルタセルのidである.

b) プーリング

プーリングでは、特徴マップの空間的な特徴を維持し ながらサイズを圧縮する.一定領域内(次式のP_i))の最 大値を取るMAXプーリングや、平均をとるAverageプー リングなどの手順が考案されているが、本研究では広く 用いられているMAXプーリングを用いた.

$$u_{ijk} = \max_{(p,q)\in P_{ij}} z_{pqk} \tag{2}$$

c) 配列の変換

最終的に全結合層に接続して予測値を出力するため, 三次元のデータ(画像サイズ縦×画像サイズ横×フィル タ数)を一次元のデータに並べ替える.

d) 全結合

変換された一次元のデータを各ノードに入力し、最終 的に予測値という一つのノードに接続する.一次元デー タに展開した後は、以下の基礎式で表わされるような通 常のニューラルネットワークと同様の手順で計算を行う.

$$u = \theta_i + \sum_{i=1}^{K} w_i x_i \tag{3}$$

$$z = f(u) \tag{4}$$

ここで、uは各素子の入力和,xは入力値,wは重み係数, θはバイアス,Kは各階層の構成素子数,f(u)は活性化関 数,zは素子の出力である.

e) ドロップアウト

ドロップアウトでは、全結合層を伝播する値を確率的 に伝播させないという処理を行う.一つあるいは少数の



ニューロンに対する過度な依存を防ぎ,過学習による精 度低下を避けるために導入する.

(2) 河川水位予測へのCNNの応用

水位予測のためには,予測対象地点上流の流域内の降 雨分布が重要である.強くて大きな雨域があれば,水位 上昇する可能性が高い.また,予測対象地点上流の水位 データも,水位予測に重要な因子である。これらの降雨 分布や水位や水位変化などの特徴量をCNNによって把握 し,水位予測を行う.

CNNによる河川水位予測の入力データの模式図を図-2に示す.一般的な画像認識では、入力データはRGBの3 チャンネルから構成される二次元データである(図-2 左).今回の水位予測では、複数のレーダ雨量データ (生の数値データ)と水位観測所の二次元化データを利 用するため、複数チャンネルの入力データ構造となる (図-2右).

水位データは、二次元に並べることで降雨分布と同様 の形式でCNNの入力データとすることができる.具体的 には、レーダ雨量の空間範囲(メッシュ数)に合わせ、 横方向には同じ時間における値を並べ、縦方向に時系列 に値を並べることで、水位や水位変化の時系列を示す二 次元データを作成した.なお、入力地点数の変化に対応 できるよう、観測所ごとに1チャンネルのデータとした.

3. 対象流域と対象データ

(1) 対象流域

対象流域は、図-3に示す筑後川水系城原川とした.城 原川では、上流から仁比山、日出来橋、柴尾橋と水位観



図-3 城原川流域および水位観測地点

測所が存在する.本研究では、日出来橋を対象地点とした.最下流の柴尾橋は潮位の影響を受けるため、そうした地点の予測は今後の課題である.

城原川の流域面積は64.4km², 幹線流路延長は31.9km である.上流域は山地に囲まれており,中下流は市街地 に囲まれている.中流域の神崎町仁比山付近は扇状地が 発達し天井川となっており,横流入河川が少ない.下流 域はゼロメートル地帯のはん濫地形で,下流は感潮域と なっている.

(2) 検討対象出水

国交省の同時刻合成レーダ雨量が入手できる2006年以降で、日出来橋において水防団待機水位以上の水位を記録した出水を抽出し、そのうち上位30出水を対象出水とした.対象30出水は学習・評価対象出水(学習データ)と予測対象出水(テストデータ)に分離した.

学習データは26出水とし、学習用と評価用に分けて利 用した.学習においては、学習用データに対し誤差が最 小化するようモデルの調整を行い、学習に用いていない 未知の評価用データにより、そのモデルの汎化能力の確 認を行った.この評価・学習を繰り返し、精度向上が見 られなくなった段階で学習を終了した.

その後、学習にも評価にも用いていないテストデータ によりモデルの検証を行った.本検討では、出水規模に 応じた直近の出水事例を抽出し、2010/7/14、2016/6/22、 2017/4/17、2017/7/6の4出水を予測対象とした.

学習・評価対象および予測出水におけるピーク水位・ 1時間あたりの最大水位変化幅を図-4に示す.図のよう



図-4 学習・評価対象洪水および予測対象洪水のピーク水位お よび1時間あたりの最大水位変化



図-5 雨量データの設定範囲

に、予測対象の4出水はおおむね学習データの幅の中に 収まっている.

4. 予測計算の具体手順

(1) 入力データの設定

雨量データについては7種類,水位データについては2 ~4種類のデータを作成した.水位および水位変化の入 カデータの組合せと,学習データ数の組合せから,合計 15パターンの入力データを作成した.以下にその詳細を 示す.

a) 雨量データ

レーダ雨量は、強雨時の降雨の空間スケール(数km ~数+km)に対応し、検討対象の城原川流域を包含する50×50メッシュ(1メッシュ約1km)の範囲を利用した(図-5).流域外の範囲も包含することで、雨域の移動などの影響を学習させることを期待した.なお、雨量データの入力範囲の影響については後で考察を述べる.

データ種類は時間雨量と累加雨量とし、降雨の短期的

表-1 抽出方法と学習対象データ数

	データ数	
1	±0.1m/hを閾値として水位変化の大 小を分類し、両者が同量になるよ うにランダムにデータを抽出	5258
2	±0.2m/hを閾値として "	3041
3	±0.3m/hを閾値として "	1906
4	±0.4m/hを閾値として "	1149
5	ピーク水位の-12hr~8hrのデータを 抽出(データ数は20hr×6×26出水 =3120個で固定になる)	3120

および長期的な影響を考慮できるよう,10分前,20分前, …,60分前の時間雨量(6チャンネル)と360分~10分前 の累加雨量の合計7チャンネルのデータを作成した. データの組合せは,降雨-水位上昇のタイムラグが60分 程度であることを踏まえて設定した.

b) 水位データ

水位データは、レーダ雨量のサイズに合わせ50×50の データサイズになるように作成した.1観測所で1チャン ネルのデータを作成し、データ利用地点は日出来橋(予 測対象)と仁比山(予測対象上流)とした.なお、観測 ピッチが60分のデータは、線形補間により10分データに 変換したうえで用いた.

横方向には同じ時間の値を50列並べ,予測対象上流水 位観測所からの流下時間(約1時間)を踏まえ,縦方向 には10分前~250分前のデータを並べた.50行になるよ うに,10分,10分,20分,20分,…,250分,250分のよ うに同じ時間の値を2行ずつ並べた.

データ種類は水位および水位変化とし、①水位のみ、 ②水位変化のみ、③水位・水位変化の両方を用いる3パ ターンのデータを作成した.

c) 学習データの抽出

学習対象の出水期間データをそのまま全て用いると、 水位変化が小さいデータ(平常時のデータ)を多く含む ことになる.水位が上昇する期間の精度を向上させるた め、水位変化がある期間の学習対象データの比率を確保 する必要があると考えた.

そこで表-1に示すように、ある閾値で水位変化の大小 を分類し、両者が同量になるようにランダムにデータを 抽出する方法や、ピーク付近の時間帯のデータを抽出す る方法により、5パターンの学習データセットを作成し た.実際の水位データと閾値の関係で抽出できるデータ 数が決まるため、閾値を高く設定すると、データ数が少 なくなるという関係がある.

(2) CNNの設定

設定したパラメータを表-2に、ネットワーク構成図を 図-6に示す.本研究では画像認識で一般的に利用されて

表-2 設定パラメータ

パラメータ種類		数値		
畳み込み	フィルタサイズ	①2×2 ②2×2		
層	フィルタ数	1)32 2)64		
プーリン	ウィンドウサイズ	①2×2 ②2×2		
グ層				
全結合層	ニューロンの個数	9216 (12×12×64)		
ドロップ	ドロップアウト率	①0.25②0.25		
アウト		30.5		
学習設定	バッチサイズ	16		
	エポック数	精度が16エポック変		
		わらなければ終了		
		(本検討のモデルで		
		は, 概ね50程度で収		
		束した.)		

※①, ②, ③は図-6と対応している.

図-6 ネットワーク構成図

表3	本検討て	「用いた	計算環境
----	------	------	------

項目	内容			
CPU	Core i3-4130 3.4GHz			
メモリ	32GB			
GPU	Geforce GTX 1060 6GB			
OS	Windows10Pro			
ミドルウェア	python 3.6.5 / tensorflow 1.8.0			
	/ keras 2.2.0			

いる設定を用いた.今回の検討ケースでは、入力データの設定に比べパラメータが結果に与える影響は少なかったため、詳細な検討は割愛した.

各ノードの入力を出力へ変換する活性化関数には ReLU関数(Rectified Linear Unit)を用いた.また学習に 用いられる指標となる損失関数には、二乗和誤差

(Mean Squared Error, MSE)を用い,損失関数の値を小 さくする最適化手法にはAdam⁶⁾を用いた.

本検討に用いた計算環境を表-3に示す.計算所要時間 は1回の学習に6~7秒程度であった.仮にモデルの修正 の繰り返しを50回としても、5~6分で学習が可能であり、 リアルタイム計算も問題なく実行できるといえる.

表−5 選定ケースにおける誤差指標								
七垣	出水 No.				日本			
1日1示	2	4	19	27	十均			
RMSE (m)	0.15	0.16	0.05	0.10	0.11			
NS係数	0.96	0.96	0.99	0.96	0.97			
ピーク時刻差(h)	-0.33	1.00	1.00	0.67	0.58			
ピーク水位差(m)	0.36	0.18	0.04	0.26	0.16			

5. 水位予測結果と考察

(1) 各計算ケースの予測結果

表-4の通り、15ケースの計算を行い、誤差指標などから最適ケースを選定した.ここで、表-4の①水位のみがケース1~5、②水位変化のみがケース6~10、③水位と水位変化がケース11~15に対応している.

一例として平均二乗誤差 (Root Mean Squared Error, RMSE)の結果を図-7に示す.いずれも極端に精度が低 下するケースはなかったが、ケース4、9などは比較的精 度が低い傾向がみられた.ここでは、いずれの出水にお いても良いケース10を最適ケースとして選定した.

選定ケースにおける各誤差指標を表-5に示す.検証対象の出水ではいずれもNS係数は0.9以上となり、十分な予測精度であることが確認できた.

予測結果のグラフを図-8に示す.予測対象とした4出 水において,全体的な波形は一致した.また,単純な一 山洪水の波形となる2017/4/17出水は,良く一致していた. 一方,ピーク付近の予測状況を詳細にみると,特に 2017/4/17出水や2017/7/6出水などのように,数十分程度 ではあるが水位上昇・下降が遅れる傾向がみられた.60 分予測では早い降雨流出応答を追随できない場合がある と考えられるが,予測雨量も含めた入力データの組合せ 等により改善の可能性があると考えられる.また 2017/7/6出水において,水位ピークより前の時間帯であ る7/5 16時~22時付近にかけて精度低下がみられた.

図-8 予測対象4出水における最適ケースの1時間先予測結果

(2) 入力データに関する考察

本検討では雨量や水位を組合せて入力データに使用したが、雨量のみ、あるいは水位のみで予測ができれば、 より汎用性が高まる.そこで前述のケース10をベースに、 雨量のみ、および水位変化のみを入力データとするケースの精度を検証した.

計算結果の例を図-9に示す.雨量のみを利用した場合 は、低水時のノイズが多く、ピーク時刻の予測遅れが大 きい.一方、水位変化のみを利用した場合、ピークが過 大評価になっている箇所が散見されるが、雨量のみの場 合よりも、ピーク時刻の予測精度が良い.

以上から、上流の水位観測データは精度向上への寄与 が大きいと考えられる.雨量のみを用いた場合でもある 程度の予測が可能であるが、雨量・水位の両方を用いた 場合の方が、予測精度が高くなる結果となった.

なお、今回360分前~10分前と設定した雨量の累加時 間を延ばすことで、低水時の精度を上げられる可能性が あるが、入力データの最適な組み合わせの検討について は今後の課題である.

(3) 入力データサイズの設定に関する考察

2017/76出水において、水位ピークより前の時間帯で ある7/5 16時~22時付近にかけて精度低下がみられた. これは、流域外に降雨がとどまっており、実際には水位 上昇に寄与しない降雨を入力データにしたことが悪影響 を与えたと考えられる(図-10).

そこで、城原川流域を丁度包含するような範囲(24× 8メッシュ)に入力データ領域を変更したモデルで試行 を行った.水位データも雨量データのメッシュ数に合わ せて24×8にするため、考慮する時間は10分前~240分前 までに変更した.50×50メッシュの場合は250分前まで 考慮していたが、降雨流出が早い城原川においては、こ の違いはあまり影響がないと判断した.

図-11に計算結果の例を示す.雨域を24×8メッシュとしたモデルでは、50×50メッシュと比べて流域外降雨の影響による予測精度低下を解消しており、また立ち上がりやピークでの予測精度も維持することができた.

以上のように、本検討では入力データの範囲は流域界 にあわせて設定したケースの方が、流域外まで広く設定 したケースよりも妥当な結果が得られた.

6. おわりに

本研究では、レーダ雨量の降雨分布情報を有効に活用 し、対象流域の降雨強度や累積降雨量の二次元分布、水 位や水位変化を入力データとし、60分先の水位を予測す るモデルを開発した.様々な規模の出水事例で検証を 行ったところ、検証対象の出水ではいずれもNS係数は 0.9以上となり、十分な予測精度を確認した.

開発した手法は、分布型流出解析モデルなどと比べて 計算コストが小さく、モデル構築も容易である. レーダ 雨量は全国各地、世界各地で得られるため、国内の中小

図-10 精度低下がみられた際の降雨分布

河川を始めとして、様々な河川での実用化が見込まれる.

参考文献

- 一言正之, 櫻庭雅明, 清雄一: 深層学習を用いた河川水位 予測手法の開発, 土木学会論文集B1(水工学), Vol.72, No.4, pp.I_187-I_192, 2016.
- 2) 石尾将大,野島和也,一言正之,房前和朋:ニューラルネットワークと一次元不定流を組み合わせた縦断的な河川水位 予測手法,河川技術論文集,Vol.24,2018.
- Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner: Gradient-based learning applied to document recognition. Proceedings of the IEEE Vol.86 (11), pp.2278-2324, 1998.
- 4) A. Krizhevsky, I. Sutskever, and G. E. Hinton: ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger, eds. Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp.1097-1105, 2012.
- 5) 鈴木紹晟, キムスンミン, 立川康人, 市川温, 萬和明:豪 雨の発生予測に対する畳み込みニューラルネットワークの 応用, 土木学会論文集B1(水工学), Vol.74, No.5, pp.I_295-I_300, 2018.
- Diederik Kingma and Jimmy Ba.: Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs], 2014.

(2019.4.2受付)