

深層学習を用いたマルチモーダル学習による降水量予測

林 政行 †

Rafik Hadfi ††

伊藤 孝行 †††

† 名古屋工業大学情報工学科

†† 名古屋工業大学大学院情報工学専攻

††† 名古屋工業大学大学院産業戦略工学専攻

1 はじめに

本研究では降水量予測のために Sequence-to-Sequence 学習フレームワークに基づく時空間データの予測モデルを用いてマルチモーダル学習を行う手法を提案する。

過去5年間に於いて、2012年に発生した九州北部豪雨などの集中豪雨が観測されている。集中豪雨は突発的に発生する為、その発生を予測することは困難であるが、異常気象による一次災害、二次災害を最小限に抑えるために、局所的な地域での5分後、10分後、...、6時間後といった比較的近い将来における高精度な降水量の予測が求められている。

降水量の予測の問題は、時系列性と空間性の両方を持つ時空間データの予測と捉えることができる。近年の研究において、Sutskeverら[1]により Recurrent Neural Network (RNN) を用いた Sequence-to-Sequence (seq2seq) 学習フレームワークが提案され、深層学習を用いて任意長の系列データから任意長の系列データを生成できるようになった。さらに、Shiら[2]により Convolutional Neural Network (CNN) の操作を取り入れた Convolutional LSTM (ConvLSTM) 層を用いて時空間データの予測を行う seq2seq モデルが提案され、降水量予測のタスクにおいて実験が行なわれた。しかし、実際の降水量は雲の厚さ、地形、および風向きといった複数の要因(モダリティ)に依存して変化するが、Shiらの提案した予測モデルでは過去の降水量のみが予測に用いられている。そこで本研究では seq2seq モデルによる時空間データの予測において複数モダリティを予測に用いるマルチモーダル学習を行う。本研究の目的は、降水量の予測において過去の降水量のデータに加え、雲の厚さや地形といった複数の要因を入力データに加えることで、より精度の高い降水量予測を行うことである。

本稿では、2章で ConvLSTM を用いた時空間データの予測モデルについて述べ、3章で提案するマルチモーダル学習の手法を述べる。4章で実データを持ちた評価実験について述べ、5章で本研究のまとめを行う。

2 ConvLSTM による時空間データの予測モデル

Sutskeverら[1]により、RNN を使って系列から系列を生成する基盤となる seq2seq 学習フレームワークが提案された。seq2seq 学習フレームワークでは、RNN の派生である LSTM を何層にもスタックさせた、2つの同じ構成のネットワーク(Encoder と Decoder) を時間軸で結合させ、系列データを入力し、系列データを出力する

モデルが用いられる。seq2seq モデルでは Encoder ネットワークにより入力ベクトルから固定長ベクトルの符号(code)を得る符号化(encode)の操作を行い、Decoder ネットワークにより符号から出力ベクトルを得る復号化(decode)の操作を行う。LSTM は各タイムステップにおいて隠れユニットとセルの2種類の内部状態を持っており、Encoder ネットワークと Decoder ネットワークの結合は、Encoder ネットワークの LSTM の最後のタイムステップでの隠れユニットとセルの状態を Decoder ネットワークの LSTM にコピーすることで成される。

Sutskeverら[1]が提案した seq2seq 学習フレームワークをもとに、Shiら[2]により ConvLSTM を用いて時空間データを予測するモデルが提案された。ConvLSTM は LSTM を拡張し畳み込み演算を取り入れることで、LSTM の状態において空間性を保つことを可能とした層である。時空間データから時空間データを生成する Shiらの予測モデルは、2つの同じ構成の ConvLSTM をスタックさせたネットワークを時間軸で結合した Encoding-Forecasting ネットワークから成る seq2seq モデルである。

3 時空間データのマルチモーダル学習モデル

本章では、Shiら[2]の提案した ConvLSTM を用いた時空間データの予測モデルにおいて、ConvLSTM を拡張し、マルチモーダル学習を行う手法について述べる。

時空間データは時系列性と空間性を持つデータである。各タイムステップにおいて、空間性を持つ1つのデータは画像として扱うことができる。画像を扱う上で複数のモダリティをニューラルネットで用いる手法として、モダリティを全て結合し、複数のモダリティからなる多チャンネル画像として扱う方法がある[3]。提

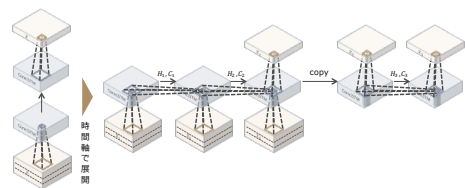


図1: Encoding-Forecasting ネットワーク

案するモデルは図1に示す Encoding-Forecasting ネットワークから成る seq2seq モデルに、複数のモダリティからなる多チャンネル画像を入力することでマルチモーダル学習を行うものである。

ConvLSTM では、LSTM の入力ユニットからの入力と隠れユニットからの入力の両方に畳み込み演算が用いられている。CNN では畳み込み層の直後に局所正規化の操作が取り入れられることが多々ある[3]。従って提案するモデルでは、図2に示すように、隠れユニットからの入力を正規化する操作(図2の①)、入力層からの入力を正規化する操作(図2の②)、および出力を正規化する操作(図2の③)を取り入れたものを検討する。

Short-term Precipitation Forecasting using Multi-modal Deep Learning
Masayuki Hayashi Rafik Hadfi Takayuki Ito

†Department of Computer Science, Nagoya Institute of Technology

††Department of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology

†††School of Techno-Business Administration, Graduate School of Engineering Nagoya Institute of Technology

表 1: 既存手法と4つの提案手法での降水量予測の結果

モデル	予測した全フレーム 交差エントロピー			30分後予測値 平均スキルスコア			60分後予測値 平均スキルスコア			90分後予測値 平均スキルスコア		
	平均	最小	最大	CSI	FAR	POD	CSI	FAR	POD	CSI	FAR	POD
	ConvLSTM-SM (既存手法)	1290.45	1283.78	1295.28	0.334	0.574	0.620	0.281	0.638	0.593	0.247	0.685
ConvLSTM-MM-入力層の局所正規化	1275.79	1262.43	1283.83	0.341	0.545	0.586	0.283	0.624	0.560	0.244	0.663	0.518
ConvLSTM-MM-隠れ層の局所正規化	1276.03	1266.71	1287.34	0.342	0.570	0.614	0.282	0.637	0.588	0.248	0.668	0.543
ConvLSTM-MM-出力層の局所正規化	1276.77	1269.94	1289.29	0.340	0.560	0.595	0.279	0.638	0.581	0.244	0.673	0.526
ConvLSTM-MM-局所正規化なし	1276.63	1265.23	1286.12	0.340	0.560	0.599	0.277	0.637	0.570	0.244	0.683	0.542

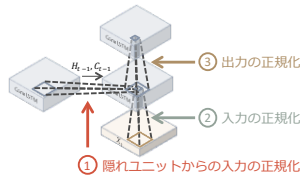


図 2: ConvLSTM 層への局所正規化操作の導入

4 実データを用いた降水量予測の評価実験

4.1 実験設定

実験では、既存手法と提案手法を用いた降水量予測について評価を行う。既存手法は Shi ら [2] により提案されたモデルにて、単一モーダル学習により過去と現在の時空間データから将来の時空間データを予測する ConvLSTM-SM とする。提案手法は、3 章で述べた ConvLSTM を用いたモデルにて、マルチモーダル学習により同様の予測を行う ConvLSTM-MM とする。

本実験では、気象庁が公開している気象レーダーによる降水量の観測データおよび気象衛星ひまわり 7 号による雲の観測画像（赤外画像）を用いる。レーダー観測値は 120 × 120 の格子から成る観測データであり、各格子点を画像のピクセル輝度値として捉えることでグレースケール画像として扱う。衛星画像は「東アジアポーラステレオ投影」の衛星画像からレーダー観測値と同じ位置での観測値を抽出し、得られた大きさ 120 × 120 の画像をグレースケール化したものを用いる。

本実験では 2014 年 8 月 1 日-31 日および 2015 年 8 月 1 日-31 日の期間に観測された 62 日分の観測データを用いる。30 分間隔で観測される同時刻の各観測値を用い、全 2,976 個の観測データから、サイズ 20 のウィンドウをスライドさせ、20 個のデータ（入力データ 10 フレーム、出力データ 10 フレーム）から成る連続データ（シーケンス）を 2,957 個得る。得られた 2,957 個のシーケンスのうち前半 2,365 個を訓練データ、295 個を検証データ、残り 296 個をテストデータとして用いる。

Shi ら [2] による先行研究を参考に、各モデルは 2 層の ConvLSTM 層からなる Encoding-Forecasting ネットワークとし、各 ConvLSTM 層は大きさ 3 × 3 のフィルタ、64 個の特徴マップ数を持つよう設定する。また、パッチサイズを 4 とする。各モデルの学習はミニバッチ学習で行い、ミニバッチの大きさは 16 とする。また、学習には RMSProp の最適化手法を用い、学習率を 0.001 と、減衰率を 0.9 とする。学習にはコスト関数として目標シーケンスと出力シーケンスとの交差エントロピー誤差を用いる。

評価指標として交差エントロピーと Critical Success Index (CSI), False Alarm Rate (FAR), および Probability of Detection (POD) の 3 種類のスキルスコアを用いる。各スキルスコアは次のようにして求める。まず、降水

量の予測値と正解値を閾値 0.5mm/h を境に 0/1(晴天/雨天を示す)の値が入った行列に変換する。次に、変換された行列において、予測値=1 かつ正解値=1 である成分の個数を *hits*(的中数), 予測値=0 かつ正解値=1 である成分の個数を *misses*(ミス数), および予測値=1 かつ正解値=0 である成分の個数を *falsealarms*(誤警報数)として数える。その後、 $CSI = \frac{hits}{hits+misses+falsealarms}$, $FAR = \frac{falsealarms}{hits+falsealarms}$, $POD = \frac{hits}{hits+misses}$ とする。

4.2 実験結果

表 1 に、実験の結果得られた最良のバリデーション評価値、および、予測した 10 フレームの内の先頭の 3 フレーム (30 分後予測値, 60 分後予測値, および 90 分後予測値) それぞれに対する CSI, FAR, および POD のスキルスコアでのテスト評価値を示す。実験結果より、提案手法の全てについて既存手法よりも平均的に低い交差エントロピー値をが得られている。また、スキルスコアでのテスト評価値を見ると、提案手法において既存手法よりも概ね高い CSI 値、および低い FAR 値が得られている。以上の結果から、提案するマルチモーダル学習手法は既存手法である単一モーダル学習手法を上回る精度で予測ができると考えられる。

5 まとめ

本研究では、降水量予測を行うために、Sequence-to-Sequence フレームワークに基づいた時空間データの予測モデルを拡張し、マルチモーダル学習を行う手法を提案した。本研究の目的は、降水量に影響を与える複数の気象要素をもとに予測を行うことで、予測精度を向上させることである。レーダー観測値と衛星画像を用いた評価実験の結果、提案したマルチモーダル学習手法が既存手法である単一モーダル学習手法を上回る精度で予測できることがわかった。ただし、提案モデルとして複数モダリティを結合し多チャンネルの画像として用いる最も簡単なアプローチをとっている。従って、今後の課題として、正準相関分析を取り入れるなど、より洗練された方法により、モダリティ間の相関を大きくするパラメータを学習させることが挙げられる。

参考文献

- [1] Ilya Sutskever, Oriol Vinyals, and Quoc VV Le. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pp. 3104–3112, 2014.
- [2] Xingjian Shi, Zhoung Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. *arXiv preprint arXiv:1506.04214*, 2015.
- [3] Wenlu Zhang, Rongjian Li, Houtao Deng, Li Wang, Weili Lin, Shuiwang Ji, and Dinggang Shen. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. *NeuroImage*, Vol. 108, pp. 214–224, 2015.