

기상레이더 강수 합성데이터를 활용한 심층신경망 기반 초단기 강수예측 기술 연구

안소정 · 최윤 · 손명재 · 김광호 · 정성화* · 박영연
기상청 기상레이더센터 레이더분석과

Short-Term Precipitation Forecasting based on Deep Neural Network with Synthetic Weather Radar Data

Sojung An · Youn Choi · MyoungJae Son · Kwang-Ho Kim · Sung-Hwa Jung* · Young-Youn Park
Weather Radar Center, Korea Meteorological Administration
E-mail : {sojungan · ychoi17 · sonmj88 · khkim777 · shjung95 · youngyounpark}@korea.kr

요 약

초단기 강수예측 시스템은 단시간 발생하는 집중호우와 같은 위험기상에 대응하기 위해 사회·경제적으로 중요하다. 최근 국내·외에서 심층신경망을 활용한 초단기 강수예측 연구가 활발히 진행되고 있다. 심층신경망을 이용한 강수예측 모델은 훈련 데이터를 만들 때 기상데이터의 구조와 종류가 복잡하고 방대하므로 기상학적 이해를 바탕으로 복잡한 전처리 과정이 필요하다. 또한, 비선형적인 패턴의 강수 현상을 예측하기 위하여 기상과의 상호작용에 대한 이해를 바탕으로 입력 데이터를 구성해야 한다. 따라서 본 연구에서는 다음과 같은 접근법을 제안하고자 한다. i) 기상레이더 합성 강수장과 강수발달에 영향을 줄 수 있는 주요 인자(레이더, 지형, 온도, 등)를 훈련 데이터 구축을 위해 패턴 분석에 적합한 형태로 정제하고 이를 구조화하여 통합한다. ii) 합성곱 신경망과 합성곱 장단기 기억 신경망을 접목하여 초단기 예측 강수장을 산출한다. 2020년 강수 사례를 이용하여 제안한 모델의 정확성을 검증하였다. 제안한 모델은 비선형적인 패턴의 강수 현상을 잘 모의하였고, 강수의 규모 및 강도에 대한 예측능력이 향상되었다. 이는 강수를 동반한 초단기 위험기상의 방재에 활용할 수 있을 것으로 기대된다.

ABSTRACT

The short-term quantitative precipitation prediction (QPF) system is important socially and economically to prevent damage from severe weather. Recently, many studies for short-term QPF model applying the Deep Neural Network (DNN) has been conducted. These studies require the sophisticated pre-processing because the mistreatment of various and vast meteorological data sets leads to lower performance of QPF. Especially, for more accurate prediction of the non-linear trends in precipitation, the dataset needs to be carefully handled based on the physical and dynamical understands the data. Thereby, this paper proposes the following approaches: i) refining and combining major factors (weather radar, terrain, air temperature, and so on) related to precipitation development in order to construct training data for pattern analysis of precipitation; ii) producing predicted precipitation fields based on Convolutional with ConvLSTM. The proposed algorithm was evaluated by rainfall events in 2020. It is outperformed in the magnitude and strength of precipitation, and clearly predicted non-linear pattern of precipitation. The algorithm can be useful as a forecasting tool for preventing severe weather.

키워드

Quantitative Precipitation Forecasting (QPF), Feature Extraction, Data Preprocessing, Nowcasting, Weather Radar

1. 서 론

집중호우와 같은 위험기상을 예방하기 위해 레

이더 데이터의 중요성이 대두되고 있다. 레이더는 한반도 전역을 5분 간격으로 관측하여 강수량 정보를 실시간으로 제공하고 있다. 이를 통해 생성되는 정량적 레이더 강수량 데이터는 위험기상을 감시하는데 용이할 뿐아니라 농업, 항공, 수자원 및

* corresponding author

하천을 관리하는 등에 탁월하다.

최근 이러한 레이더 데이터를 활용하여 인공지능 기술을 접목한 강수예측 연구가 활발히 이루어지고 있다[1, 2, 3]. 에코는 과거의 경향성을 바탕으로 이동하기 때문에 이동 벡터를 기반으로 강수를 예측할 수 있다[4]. 하지만 에코 내에서는 갑작스럽게 발생하는 강수의 생성 및 소멸 현상은 예측하기 어렵다. 이러한 현상들은 위험기상으로 발전할 수 있으며, 최근 지구온난화에 따라 뇌우와 같은 위험기상 현상이 더욱이 증가하고 있다. 따라서 본 논문에서는 이러한 위험기상을 사전에 예측하기 위하여, 기상데이터를 합성하고 이를 인공지능 기술을 접목하는 초단기 강수예측 방법을 제안하고자 한다.

II. 관련연구

기존의 강수예측 연구는 광학적 흐름(Optical Flow)을 기반으로 시계열 행렬을 예측하는 시스템으로 구성되어 있었다[4]. 최근 시공간의 비선형 데이터를 분석하기 위해 C-RNN 기술이 등장함에 따라 이를 활용한 강수예측 연구들이 시도되고 있다. 특히, (Shi, et al, 2016)의 ConvLSTM(Convolutional Long Short Term Memory) 모델은 FC(Fully Connected) Layer에서 벗어나 강수의 시공간 패턴을 분석함으로써 비선형적인 강수의 흐름을 효과적으로 예측할 수 있었다[1]. 하지만 ConvLSTM은 구름의 회전과 같은 자연적 이동과 변형을 고려하지 못하는 한계가 존재하였다. 이러한 문제를 보완하기 위하여 TrajGRU 모델이 제안되었다[2]. TrajGRU는 광학적 흐름을 분석하기 위하여 Convolutional 레이어를 우선 적용하고 이를 GRU 유닛과 결합하였다. TrajGRU는 강수 회전을 예측하는데 우수한 성능을 보였다.

그러나 심층신경망 학습을 기반으로 한 대부분의 강수예측 연구들은 고해상도의 강수 영상을 예측하는 데 어려움이 있다. 최근 고해상도 강수예측을 위하여 적대적 생성 신경망(GAN; Generative Adversarial Network)을 적용한 연구가 발표되었고, 고해상도의 강수 영상을 생성할 수 있었다[5]. 그러나 에코의 이동 방향과 강도에 따른 정확도를 개선하지 못했다.

III. 합성데이터 기반 초단기 강수예측

본문에서는 합성데이터 기반 강수예측을 위한 새로운 모델을 제안한다.

3.1 데이터 합성

우리나라의 경우 지형의 특성이 강수 현상에 영향력이 크다. 주로 바람이 불어오는 산맥의 사면에

집중적으로 비가 내리나 이와 반대쪽의 사면에는 잘 내리지 않는다. 이 외에도 온도 차로 인하여 발생하는 계절풍, 호우, 등이 발생한다. 이러한 기상학적 특성을 분석하기 위해 다음 네 가지의 데이터를 합성하여 입력데이터를 구성한다.

- HSR: 지형 클러스터, 빔 차폐, 비 기상 에코 및 밝은 띠의 영향을 받지 않는 하이브리드 고도면의 반사도를 합성한 레이더 데이터
- 지형: 한반도 고도 정보
- 온도: 지상, 1500m, 3000m 온도
- 날짜 및 시간

다음과 같은 Z-R 관계식은 레이더 관측으로 표출한 반사도(dbz)를 강우 강도 R로 바꿔주기 위함이다.

$$Z = aR^b, \quad (1)$$

a, b는 강수 유형에 따라 경험적으로 얻는 파라미터이다. 강수량 추정 오차를 완화하기 위하여 우리나라 특성에 적합한 상수(a=148, b=1.59)[6]를 적용한다. 기상데이터의 경우 자연 관측 값이므로 이상치 값을 벗어나는 값들이 존재한다. 따라서 통계 분석을 바탕으로 데이터 분포와 연속성을 고려하여 데이터의 범위를 지정하고 이를 정규화한다.

3.2 강수예측 네트워크

본 논문에서는 (X. Shi, et al, 2017)의 Convolutional-RNN 모델을 기반으로 확장된 알고리즘을 이용하여 강수 패턴을 추출한다. 강수예측을 위한 네트워크 구조는 [알고리즘 1]과 같다.

알고리즘 1. 강수예측을 위한 ConvLSTM 기반 훈련 알고리즘의 의사 코드

Algorithm: Precipitation Forecasting based on synthetic weather radar data

Input : HSR $H_{i:i+6}$; Altitude $A_{i:i+6}$;
Date $D_{i:i+6}$; TIME $T_{i:i+6}$;
Temperature $\{T_{10,i:i+6}, T_{1500,i:i+6}, T_{3000,i:i+6}\}$;

output: HSR $H_{i+6:i+10}$;

concatenation input data to $X_{i:i+6}$

Initialize the parameter w

$n \leftarrow$ length of network

While $\epsilon < Epoch$ **do**

for $step: 1 \rightarrow n$ **do**

if $step = 1$ **then**

 Compute parameters via Conv

else

 Down-sampling X via Conv

$w_{si} := \sin(w_{si} \circ c_{t-1} + b_i)$

 Compute hidden state via ConvLSTM

for $step: n \rightarrow 1$ **do**

 Generate output hidden state via ConvLSTM

if $step = 1$ **then**

 Compute parameters via Conv

 Compute parameters via 1x1 Conv

else

 Up-sampling X via Deconv

$w_{si} := \sin(w_{si} \circ c_{t-1} + b_i)$

 Update parameter ∇w

입력 및 출력데이터는 각각 7개(1시간 전), 12(2시간)개의 시계열 행렬로 구성되어 있다. 알고리즘의 네트워크는 Convolutional-ConvLSTM 구조로 구성되어 있으며, n 은 해당 네트워크 구조의 개수를 의미한다. 위의 알고리즘과 같이 Convolutional을 적용하고 이를 바탕으로 각 네트워크의 파라미터들은 시공간을 도함수 표현에 적합한 Sine 활성화 함수(Sitzmann, 2020)를 적용한다[11].

IV. 실험

본 연구에서는 제안한 방법론을 검증하기 위하여 다음과 같이 실험 데이터를 구성하였다.

4.1 데이터의 구성 및 검증 방법

모든 데이터는 10분 간격으로 추출하였으며, 두 시간(총 12장) 강수 시계열 행렬을 예측하는 것을 목표로 하고 있다. 훈련 데이터는 2018~2019년도 105,120장, 검증/테스트 데이터는 2020년도 52,704장으로 구성되어 있다. 한반도 규모의 16km 해상도 64X64의 2차원 레이더 데이터로 재구성하였다. 이때 시계열 행렬의 강수 규모가 1% 이상인 경우만을 데이터로 구성하였다. 총 훈련 8,484세트, 검증 2,082세트, 테스트 2,082세트로 구성되어 있다.

4.2 검증 결과

최종적으로 평균 제곱 오차(MSE; Mean Squared Error), 상관계수(CORR; Correlation), F1 Score, 임계성공지수(CSI; Critical Success Index)를 이용하여 예측결과를 평가하였다. 세 강수 임계치 0.2/0.5/1.0(mm)를 기준으로 정확도를 평균하였다.

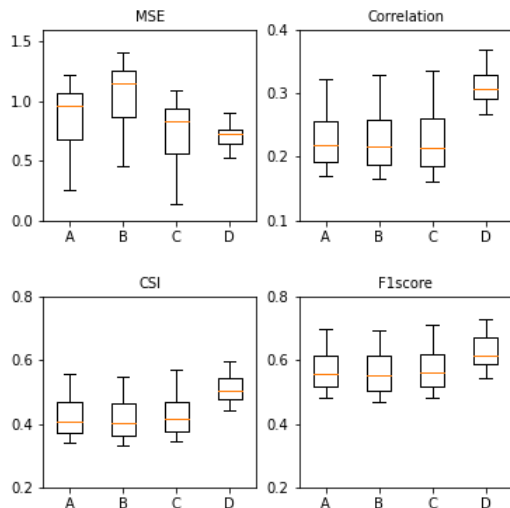


그림 1. 실험 결과에 따른 정확도 비교. 12장의 예측 시계열 행렬에 대하여 정확도 분포를 나타낸 그림으로, A-ConvLSTM, B-ConvGRU, C-TrajGRU, D-제안한 모델을 나타낸다.

V. 결론

위험기상을 방지하기 위하여 인공지능 기술을 접목한 레이더 예측 연구들이 활발히 진행되고 있다. 그러나 복잡한 대기 과정에 대한 불완전한 이해로 인한 불확실성 등 여러 조건으로 인해 신뢰할 수 없는 결과로 나타나는 제한점이 발생하기도 한다. 이에 뇌우, 집중호우와 같은 위험기상을 예방하기 위하여 새로운 접근법을 제안하였다. 예측결과는 대기의 흐름을 잘 이해하였고, 에코의 규모 및 강도 측면에서 보다 더 정확하게 예측할 수 있었다.

Acknowledgement

이 연구는 기상청 기상레이더센터 R&D 연구개발사업 “국가레이더 통합 활용기술 개발 (KMA2021-00220)”의 지원으로 수행되었습니다.

References

- [1] X. Shi, et al, “Convolutional LSTM network: A machine learning approach for precipitation nowcasting,” *The Journal of Neural Information Processing System*, Jun. 2015.
- [2] X. Shi, et al “Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model,” in *Proceeding of the 31st Conference on Neural Information Processing System, CA, USA, May. 2017*.
- [3] C. K. Sønderby, et al., “MetNet: A Neural Weather Model for Precipitation Forecasting,” Available : <http://arxiv.org/abs/2003.12140>.
- [4] H. C. Lee, Y. et al, “McGill Algorithm for Precipitation Nowcasting by Lagrangian Extrapolation (MAPLE) applied to the South Korean radar network,” *The Journal of Asia-Pacific Journal of Atmospheric Sciences*, Vol. 46, No. 3, pp.383-391, Nov. 2010
- [5] J. Jing, et al, “AENN: A Generative Adversarial Neural Network for Weather Radar Echo Extrapolation,” *Remote Sensing, The Journal of Remote Sensing and Spatial Information Science*, Vol. 42, pp.89-94, Jan. 2019.
- [6] Weather Radar Center, Development and Application of Cross Governmental Dual-pol Radar Harmonization, Dec. 2020.
- [7] V. Sitzmann, et al, “Implicit Neural Representations with Periodic Activation Functions,” *The Journal of Advances in Neural Information Processing System*, Vol. 33, Dec. 2020.