

기상 데이터와 기상 위성 영상을 이용한 다중 딥러닝 모델 기반 일사량 예측

Radiation Prediction Based on Multi Deep Learning Model Using Weather Data and Weather Satellites Image

김재정¹ · 유용훈¹ · 김창복²

¹가천대학교 에너지 IT학과

²가천대학교 전기공학과

Jae-Jung Kim¹ · Yong-Hun You¹ · Chang-Bok Kim^{2*}

¹Department of Energy IT, Gachon University, Gyeonggi-do, 13120, Korea

²Department of Electrical Engineering, Gachon University, Gyeonggi-do, 13120, Korea

[요 약]

딥러닝은 데이터의 품질과 모델에 따라 예측 성능에 차이를 보인다. 본 연구는 발전량 예측에 가장 영향을 주는 일사량 예측을 위한 최적의 딥러닝 모델을 구축하기 위해 다양한 입력 데이터와 다중 딥러닝 모델을 사용하였다. 입력 데이터는 기상청의 기상 데이터와 천리안 기상영상을 기상청 지역의 영상을 분할하여 사용하였다. 본 연구는 기본적인 딥러닝 모델인 DNN, LSTM, CNN 모델에 대해 중간층의 깊이와 노드를 변경하여 일사량을 예측하여, 비교 평가하였다. 또한, 각 모델에서 가장 좋은 오차율을 가진 모델을 연결한 다중 딥러닝 모델을 구축하여 일사량을 예측하였다. 실험 결과로서 다중 딥러닝 모델인 모델 A의 RMSE는 0.0637이며, 모델 B의 RMSE는 0.07062이며, 모델 C의 RMSE는 0.06052로서 단일 모델보다 모델 A 그리고 모델 C의 오차율이 줄었다. 본 연구는 실험을 통해 두 개 이상의 모델을 연결한 모델이 향상된 예측률과 안정된 학습 결과를 보였다.

[Abstract]

Deep learning shows differences in prediction performance depending on data quality and model. This study uses various input data and multiple deep learning models to build an optimal deep learning model for predicting solar radiation, which has the most influence on power generation prediction. As the input data, the weather data of the Korea Meteorological Administration and the clairvoyant meteorological image were used by segmenting the image of the Korea Meteorological Agency. , comparative evaluation, and predicting solar radiation by constructing multiple deep learning models connecting the models with the best error rate in each model. As an experimental result, the RMSE of model A, which is a multiple deep learning model, was 0.0637, the RMSE of model B was 0.07062, and the RMSE of model C was 0.06052, so the error rate of model A and model C was better than that of a single model. In this study, the model that connected two or more models through experiments showed improved prediction rates and stable learning results.

Key word : Satellites Image, Weather Data, Multi Deep Learning Model, Radiation Prediction, Artificial Intelligence.

<https://doi.org/10.12673/jant.2021.25.6.569>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 8 November 2021; Revised 1 December 2021

Accepted (Publication) 23 December 2021 (30 December 2021)

*Corresponding Author : Chang-Bok Kim

Tel : +82-10-8908-3946

E-mail : cbkim@gachon.ac.kr

I. 서 론

실생활에서 사용하는 에너지는 대부분 전기 에너지이며, 에너지 발전은 환경오염이 심한 화력과 원자력 발전이 전체의 80% 이상 차지하고 있다. 신재생 에너지는 온실 가스나 오염 물질을 거의 발생시키지 않기 때문에, 지구 온난화와 대기 환경오염을 발생하지 않는 장점이 있다. 특히, 태양광 발전은 정부의 정책적인 지원을 받아 신재생 에너지 중 급속히 성장하고 있으며 상용화가 이루어지고 있다[1].

태양광 발전은 햇빛이 비치는 곳이면 간단히 설치할 수 있으며, 소형으로도 제작할 수 있다. 이러한 장점은 에너지 발전과 소모를 국지적인 지역에서 실시간으로 전력을 생성하고 소모할 수 있다. 태양광 발전은 각종 시설물, 에너지 자립 섬, 자동차 및 비행기와 같은 모바일 등 전기를 사용하는 모든 사물에서 직접 발전 및 소모함으로써 전기 에너지를 효율적으로 사용할 수 있게 된다. 그러나 태양의 빛 에너지를 이용하는 특성상 일사량에 영향을 받는다. 또한, 전기를 생산하는 태양광 패널이 반도체 소자이므로 온도, 습도, 풍속, 전운량, 미세먼지 등 기상 상태에 영향을 받는다. 효율적으로 태양광 에너지를 사용하기 위해서는 발전량을 정확히 예측하는 기술이 필수적이다. 발전량 예측을 위한 기상정보는 기상청의 일기예보를 통해 쉽게 획득할 수 있다. 하지만, 발전량 예측에 많은 영향을 미치는 일사량은 일기예보에 포함되어 있지 않기 때문에, 효율적인 발전량 예측을 위해 정확한 일사량 예측이 필요하다.

일사량 예측은 물리적 방법(physical based method)과 데이터 기반 방법(data driven method)으로 나누어진다. 물리적 방법은 동역학 및 물리 방정식으로 구름, 대기 복사 등의 대류 현상을 계산한 다음 일사량을 추정한다[2][3][4]. 이 방법은 비선형성을 갖는 일사량, 전운량 및 기온 등의 기상정보를 관계식으로 나타내기 어려울 뿐만 아니라 외부 환경이 바뀔 때마다 지면에 대한 경계조건 및 상관계수들을 재설정해야 한다는 단점이 있다. 데이터 기반 방법은 많은 양의 데이터를 이용하여, 다양한 기계학습 모델을 생성하는 방법으로 입출력 사이의 관계가 비선형일 경우에도 적용할 수 있다. 데이터 기반 방법 중 딥러닝은 이전의 기상정보를 학습하여, 미래의 정보를 예측하는 알고리즘으로 기계학습 중 우수한 성능을 제공하고 있다.

딥러닝 알고리즘은 DNN(deep neural network), LSTM(long short term memory), CNN(convolution neural network)이 기반이 되며, 입력 데이터의 특징에 따라 서로 다른 성능을 제공한다. DNN은 독립변수로 기상정보를 입력하며, 종속변수인 일사량을 출력으로 하여 신경망을 학습하는 모델이다. LSTM은 DNN을 병렬로 타임스탬프(time stamp)만큼 연결한 모델이다. 즉, 이전 DNN에서 학습한 결과를 다음 DNN의 입력으로 하여 학습하는 방법으로 시계열 데이터에 적합한 학습모델이다. 기상정보는 시간에 따라 다양하게 변하는 시계열 데이터로서 시계열 데이터를 효과적으로 처리할 수 있는 순환신경망인 LSTM이 가장 우수한 성능을 제공하는 것으로 알려져 있다.

CNN은 영상을 입력으로 하여, 영상에서 특징을 추출하는 합성곱을 통해 영상의 특징을 추출하고 영상의 특징을 유지한 채 크기를 축소하는 풀링을 거친 특성맵을 DNN의 입력으로 하여 학습하는 모델이다[5][6][7].

딥러닝을 이용한 예측분야는 학습 데이터의 질이 우수해야 하며, 학습 데이터의 특성에 적합한 딥러닝 모델을 구축하여야 한다. 본 연구는 전체 기상 영상으로 일사량을 구축하는 것은 실제 일사량을 측정한 기상청의 지역 외 영상까지 학습하는 문제점을 제거하기 위하여 기상청 지역의 영상을 분할하여 일사량을 예측하였다.

딥러닝 모델은 중간층이 깊으면 활성화 함수의 특성과 과적합에 의해 오차율이 떨어짐으로서 모델의 성능이 저하되는 특성이 있다. 본 연구는 이를 보완하기 위해 DNN, LSTM, CNN 등 모든 모델에 대해 중간층과 노드 수를 변경하여 일사량을 예측하였으며, 각 모델에서 가장 좋은 예측률을 거친 모델을 병합한 다중 딥러닝 모델을 구축하여, 일사량을 예측하였다.

II. Backgrounds

딥러닝은 다양한 예측을 위하여 DNN, LSTM, CNN 등의 모델을 사용한다. 각 모델은 입력데이터의 형태와 특성에 따라 성능에 대한 차이가 있다. DNN(deep neural network)은 입력층, 중간층, 출력층으로 나뉘며 정량화된 데이터가 입력층으로 입력되며, 1층 이상의 중간층으로 구성되고, 예측 값이 출력층으로 출력된다. RNN(recurrent neural network)은 DNN을 병렬로 연결한 구조로, 중간층을 은닉층이라 하며, 은닉층의 결과 값을 출력층 방향으로 보내면서 동시에, 다음 DNN의 입력으로 보내는 특징을 갖고 있다. RNN은 연결된 DNN의 개수를 타임스탬프(time stamp)라 하며, 타임스탬프 전체를 학습하는 것이 1회 학습이다. RNN은 이전 타임스탬프에서 학습한 결과를 현재 학습에 이용함에 의해 시계열 데이터에 대한 예측에 효과적인 것이다. LSTM(long short term memory)은 은닉층의 메모리 셀에 입력, 망각, 출력 게이트를 추가하여, 타임스탬프가 길어질 수록 앞의 정보가 뒤로 충분히 전달되지 못하는 문제를 해결하기 하였다[8]. CNN(convolution neural network)은 영상 처리에 탁월한 성능을 보이는 신경망으로 합성곱층과 풀링층으로 구성된다. 합성곱 연산은 영상의 특징을 추출하는 역할을 하며, 풀링은 특성 맵을 다운샘플링하여 크기를 줄이는 연산으로 최대 풀링과 평균 풀링이 있다.

일사량에는 직달 일사량과 산란 일사량 그리고 전천 일사량이 존재한다. 직달 일사량은 복사 에너지가 반사되지 않고 지표면에 도달하는 것이고, 산란 일사량은 복사에너지가 대기 중의 입자나 구름에 의한 산란 광선을 측정한 일사량이다. 또한, 전천 일사량은 직달 일사량과 산란 일사량을 더한 일사량으로 기상청은 전천 일사량을 측정한다[9]. 그림 1에 전천 일사량에 대해서 나타났다.

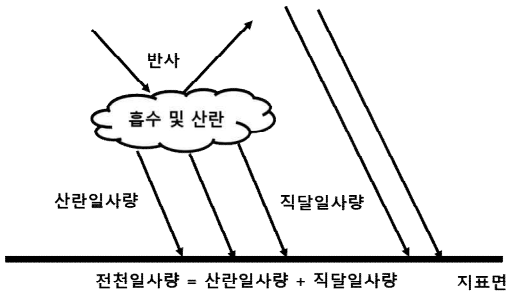


그림 1. 전천 일사량
Fig. 1. Global radiation

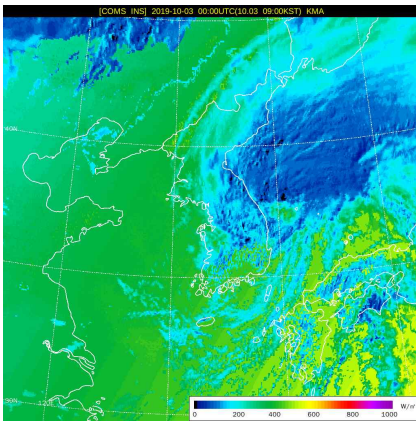


그림 2. 표면 도달 일사량
Fig. 2. Radiation reaching the surface

태양 에너지는 열, 빛, 자외선 등의 다양한 형태로 지구표면에 도달하며, 오존이나 구름 등에 의해 감소 혹은 반사되어 일부 분만 도달한다. 복사 에너지는 적외선이나 가시광선 등이 운반하는 에너지이며, 지표면에 도달하는 복사 에너지를 일사량이라 한다. 일사량은 MJ/m² 단위로 측정되며 농업 생산량, 발전량, 적조 예측 등 다양한 연구 모델의 데이터로 활용된다. 그림 2에 표면 도달 일사량에 대해서 나타냈다. 본 영상은 2019년 10월 3일 9시의 일사량이며, 영상에 나타난 색의 일사량과 실제 기상청에서 측정한 일사량이 일치함을 확인하였다.

III. 제안 시스템

딥러닝 모델은 중간층이 깊으면 활성화 함수의 특성과 과적합에 의해 오차율이 떨어짐으로서 모델의 성능이 저하되는 특성이 있다. 따라서 본 연구는 DNN, LSTM, CNN 등 하나의 모델만을 사용하여 학습하는 것이 아니라 다양한 딥러닝 모델을 병렬로 병합하여 모델의 성능을 높이는 방법을 제안하였다. 연구의 연구 방법과 순서는 다음과 같다.

1. DNN 및 LSTM 모델의 성능 예측을 위해 다양한 중간층과

노드 수로 변경하여 예측하였다.

2. DNN은 현재 시간의 기상요소를 입력으로 하였으며, 실제 데이터는 다음 시간의 일사량이다. 학습모델은 중간층을 2층과 3층으로 구성하였으며, 노드는 16, 32, 64, 128로 하였다.

3. LSTM은 1일 24시간 주기로 학습하기 위하여 타임스탬프를 24로 하였으며, 각 타임스탬프의 입력은 현재시간의 기상요소이며, 실제 데이터는 한 시간 후의 일사량이다. 은닉층을 1층과 2층으로 구성하였으며, 은닉층의 노드 수는 16, 32, 64, 128개로 구성하였다.

4. CNN은 서울 기상청 지역의 일사량을 측정한 곳의 좌표를 이용하여 학습하였다. 학습을 위한 좌표 반경은 기상청 위치를 중심으로 24 픽셀, 48 픽셀, 86 픽셀로 하였다. 기상 영상은 1픽셀당 공간 해상도는 천리안 위성은 4km이다. 따라서 24픽셀을 예로 들면, 기상청 위치로부터 반경 48(12 X 4)km의 영상이다. 각 영상에 대한 CNN 모델은 24 픽셀 영상은 2번의 합성곱과 2번의 풀링, 48 픽셀은 3번의 합성곱과 3번의 풀링, 96 픽셀은 4번의 합성곱과 4번의 풀링으로 하였다. 또한 커널은 3X3으로 하였으며, 합성곱 마다 채널을 2배씩 늘렸으며, 풀링마다 특성 맵의 크기를 두 배로 줄였다.

5. 각 모델에서 가장 좋은 오차율을 가진 네트워크를 이용하여, 병렬로 다중 네트워크를 구성하여 예측을 성능 평가를 하였다.

6. 각 단일 모델 중에서 가장 오차율이 낮은 모델을 이용하여, 본 연구에서 제안하는 다중 딥러닝 모델을 구축하여 모델 평가 및 비교를 한다.

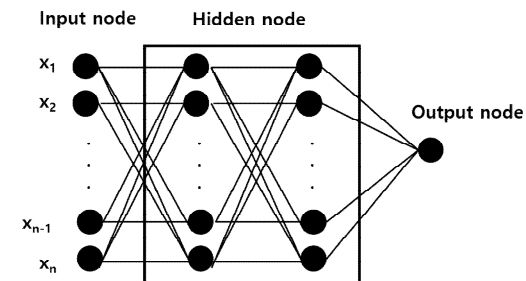


그림 3. DNN 모델
Fig. 3. DNN model

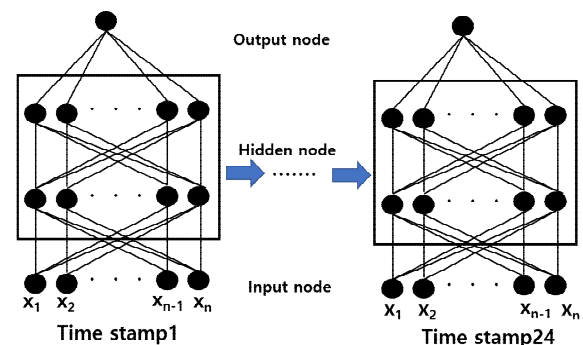


그림 4. LSTM 모델
Fig. 4. LSTM model

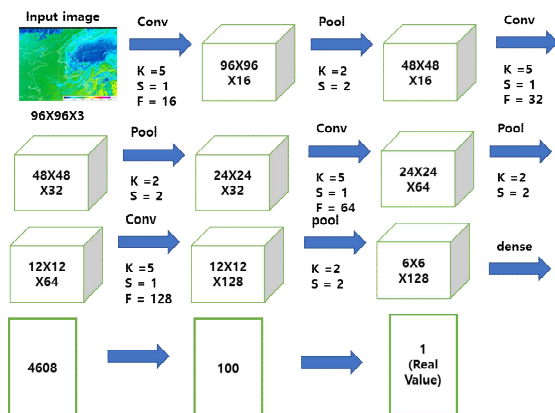


그림 5. CNN 모델
Fig. 5. CNN Model

본 연구에서 제안하는 다중 딥러닝 모델로서, 모델 A는 DNN_LSTM 모델이고, 모델 B는 LSTM_CNN 모델 그리고 모델 C는 DNN_LSTM_CNN 모델이다.

모델 A는 DNN과 LSTM 모델의 출력을 연결하여, 두 개의 출력 노드를 생성하고 이에 대해서 4개의 DNN층으로 연결하여 최종적으로 일사량을 예측하였다.

모델 B는 LSTM과 CNN 모델의 출력을 연결하여, 두 개의 출력 노드를 생성하고 이에 대해서 4개의 DNN층으로 연결하여 최종적으로 일사량을 예측하였다.

모델 C는 DNN과 LSTM 그리고 CNN 모델의 출력을 연결하여, 세 개의 출력 노드를 생성하고 이에 대해서 4개의 DNN층으로 연결하여 최종적으로 일사량을 예측하였다. 그림 6과 7 그리고 8에 제안 모델 A, B, C에 대해서 나타났다.

제안 모델들은 입력 데이터의 형태가 다르기 때문에, 입력데이터의 매칭이 필요하다. 모델의 입력데이터를 구축하기 위해 DNN과 CNN의 학습 입력데이터에 대해 23개의 데이터를 삭제하였다. 이것은 LSTM 데이터 구축 시 24번째 데이터까지가 각 타임스탬프에 입력되는 데이터이며, 25번째 일사량이 실제 데이터이다. 따라서 DNN과 CNN의 23개 데이터를 삭제해야만 LSTM의 실제 데이터와 일치하기 때문이다. 그림 9에 제안 모델의 데이터 형식에 대해서 나타났다.

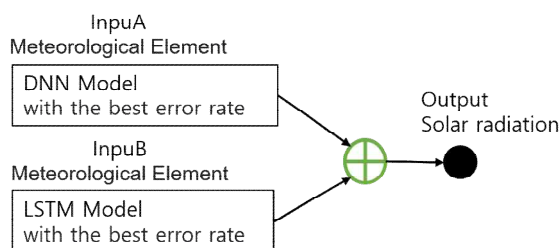


그림 6. 제안 모델 A
Fig. 6. Proposal Model A

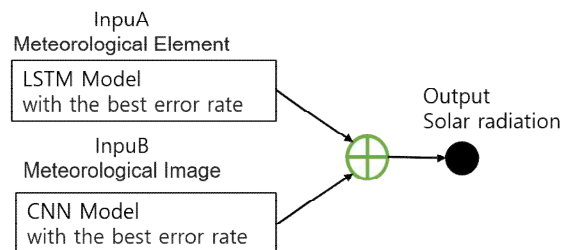


그림 7. 제안 모델 B
Fig. 7. Proposal Model B

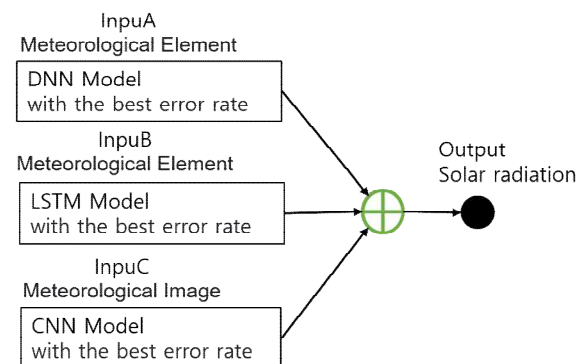


그림 8. 제안 모델 C
Fig. 8. Proposal Model C

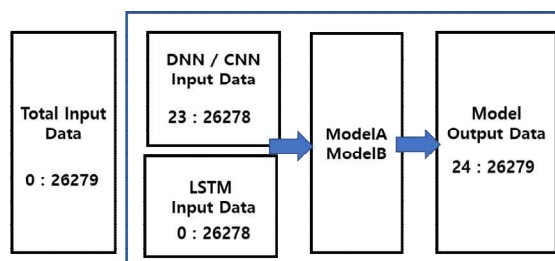


그림 9. 입출력 데이터 형식
Fig. 9. I/O data format

그림 9에서 각 숫자는 입출력 데이터의 샘플을 나타낸다. 총 샘플 수는 26280이며, DNN과 CNN의 입력 샘플 수는 26257개이고, LSTM의 입력 샘플 수는 26280이다. 또한 실제 데이터의 샘플 수는 다음 시간의 일사량으로 26257개이다.

IV. 실험결과 및 비교검토

4-1 실험결과

본 실험은 아나콘다 환경에서 파이썬 3.6, tensor flow, keras 등과 OpenCV 및 numpy 모듈을 사용하였다,

DNN 학습은 SGD(stochastic gradient descent) 알고리즘을 이용하였으며, 학습률은 0.001, 배치 사이즈는 10으로 하여 100번 반복 학습하였다. 학습 데이터의 shape는 (21864, 10),

(21864, 1)이며, 테스트 데이터의 shape는 (4392, 10) (4392, 1)이다. 표 1에 DNN의 실험결과에 대해서 나타냈다. 예측을 측정하는 평균 제곱근 오차인 RMSE((root mean square error)으로 하였다. 실험 결과 RMSE가 0.7686에서 0.083 정도로 나타났다. 각 층마다 32개의 노드로 3층 구조인 모델의 RMSE가 가장 적음을 알 수 있다. 학습 파라미터는 총 1441개의 학습 파라미터를 사용하였다. 좀 더 많은 노드수의 모델은 오히려 과적합에 의해 오차가 증가함을 볼 수 있었다.

LSTM 학습은 SGD 알고리즘을 이용하였으며, 학습률은 0.001, 배치 사이즈는 10으로 하여 100번 반복 학습하였다. 또한, 학습 데이터 shape는 (21864, 24, 10) , (21864, 1)이며, 테스트 데이터의 shape는 (4392, 24, 10) (4392, 1)이다. 표 2에 LSTM 실험 결과에 대해서 나타냈다. 실험 결과 RMSE가 0.064에서 0.075 정도로 나타났다. 32개의 노드로 1층 구조인 모델의 오차가 가장 적음을 알 수 있다. 학습 파라미터는 총 75,329개를 사용하였다. 좀 더 많은 노드수의 모델은 DNN과 마찬가지로 과적합에 의해 오차가 증가하였다.

CNN 모델은 가장 예측률이 좋은 모델로서, 합성곱과 풀링을 3회 반복하였으며, CNN은 서울 기상청을 중심으로 48픽셀셀의 영상이 가장 오차율이 낮았다. 출력을 펼쳐서 DNN 층의 노드수를 100으로 하여, 학습 파라미터는 총 6,590,921개의 학습 파라미터를 사용하였다. DNN과 LSTM은 3회 반복하여, 평균으로 모델 평가를 하였으며, CNN은 3회 반복하여, 평균으로 모델 평가를 하였다.

표 1. DNN 실험 결과

Table 1. DNN experiment results

1 layer	2 layer	3 layer	iter. 1	iter. 2	iter. 3	avg
16	16		0.079	0.078	0.077	0.078
16	16	16	0.0792	0.0777	0.0768	0.079
32	32		0.078	0.077	0.077	0.0773
32	32	32	0.0772	0.077	0.0764	0.7686
64	64		0.077	0.077	0.077	0.077
128			0.0804	0.0799	0.0797	0.08
128	128		0.078	0.087	0.080	0.083

표 2. LSTM 실험 결과

Table 2. LSTM experiment results

1 layer	2 layer	3 layer	iter. 1	iter. 2	iter. 3	avg
16			0.0667	0.0662	0.0648	0.066
16	16		0.0719	0.0704	0.0695	0.071
16	16	16	0.0687	0.0658	0.065	0.067
32			0.064	0.065	0.0638	0.064
32	32		0.0677	0.0688	0.067	0.068
64			0.0663	0.0661	0.065	0.066
64	64		0.0696	0.0701	0.0681	0.069
128			0.0693	0.0694	0.070	0.070
256			0.0676	0.0679	0.0669	0.067
256	256		0.0766	0.0743	0.0746	0.075

표 3. CNN 실험 결과

Table 3. CNN experiment results

size	iter.1	iter.2	iter.3	iter.4	iter.5	avg
24 pixel	0.308	0.313	0.314	0.311	0.307	0.3106
48 pixel	0.306	0.309	0.309	0.310	0.303	0.3074
96 pixel	0.313	0.312	0.314	0.313	0.311	0.3126

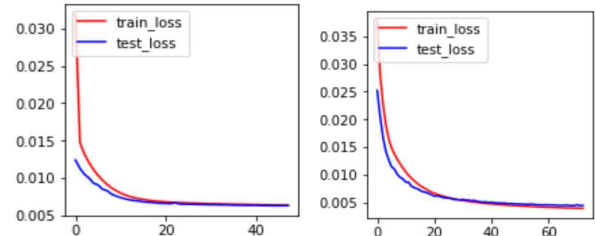


그림 10. DNN과 LSTM의 학습 오차

Fig. 10. Training error of DNN and LSTM

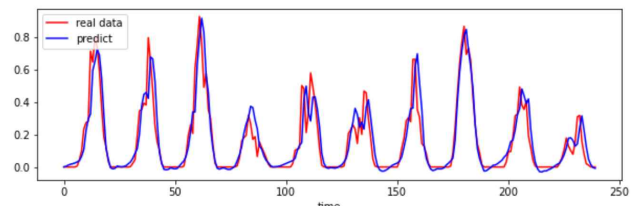


그림 11. 10일간 LSTM 예측결과

Fig. 11. 10-day LSTM prediction results

그림 10에 DNN과 LSTM의 학습 오차에 대해서 나타냈다. 또한, 그림 11에 10일간의 LSTM 예측결과를 나타냈다.

표 4에 각 모델의 학습결과를 나타냈다. 모든 모델은 동일하게 학습은 SGD 알고리즘을 이용하였으며, 학습률은 0.001, 배치 사이즈는 10으로 하여 100번 반복 학습하였다. 또한, 오차가 감소하다가 다시 증가하는 경우가 과적합 신호이기 때문에 오차 감소 후 증가가 3회 반복하면 학습을 중지하도록 하였다. 학습 결과 DNN과 CNN 그리고 모델 B가 다소 학습이 빨리 종료하였으며, LSTM과 모델 A 그리고 모델 C가 안정적으로 오랫동안 학습을 하였다. 또한, LSTM과 모델 A 그리고 모델 C의 학습오차와 테스트 오차가 월등하게 줄었음을 알 수 있다.

표 4. 모델의 학습 결과

Table 4. Training results of model

model	Epoch	train_loss	test_loss
DNN	32/10	0.0063	0.0063
LSTM	94/100	0.0038	0.0044
CNN	7/15	0.0041	0.0048
model A	94/100	0.0032	0.0038
model B	14/100	0.0043	0.0050
model C	111/150	0.0030	0.0036

표 5. 실험 결과

Table 5. Experiment results

mode	iter.1	iter.2	iter.3	iter.4	iter.5	avg
DNN	0.0772	0.077	0.0764	0.081	0.079	0.0781
LSTM	0.064	0.065	0.0638	0.066	0.067	0.0652
CNN	0.306	0.309	0.309	0.310	0.303	0.3074
model A	0.0638	0.0623	0.0638	0.0641	0.0645	0.0637
model B	0.0703	0.0710	0.0703	0.0700	0.0715	0.07062
model C	0.0599	0.0612	0.0607	0.0610	0.060	0.06065

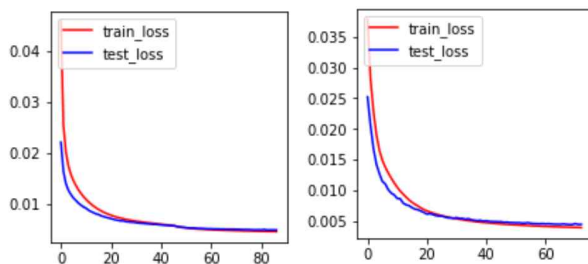


그림 12. 모델 A와 모델 B의 학습 오차

Fig. 12. Training error of model A and model B

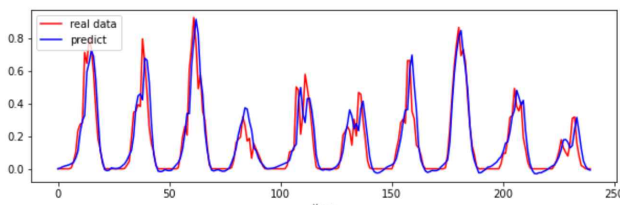


그림 13. 10일간 모델 A 예측결과

Fig. 13. 10-day model A prediction results

표 5에 제안 모델의 실험 결과에 대해서 나타냈다. 모든 모델의 결과는 5회 반복하여, 평균으로 모델 평가를 하였다. 학습결과 모델 A의 RMSE는 0.0637이며, 모델 B의 RMSE는 0.07062이며, 모델 C의 RMSE는 0.06052이다. 실험결과와 같이 모델 A와 모델 C가 다른 모델 보다 RMSE가 좋았음을 알 수 있다.

그림 12에 두 모델의 학습 오차에 대해서 나타냈다. 또한 그림 13에 10일간의 모델 1에 대한 예측결과를 나타냈다. 그림 11와 13을 비교하였을 때, 그림 14가 좀 더 예측을 잘하고 있음을 알 수 있다.

4-2 비교 검토

본 연구는 전체 영상으로 일사량을 구축하는 것은 실제 일사량을 측정된 기상청의 지역 외 영상까지 학습하는 문제점을 제거하여, 기상청 지역의 영상을 분할하여 일사량을 예측하였다. 또한, 모든 딥러닝 모델을 이용하여 일사량을 예측하여, 비교 평가하였으며, 특히, 두 개 이상의 모델을 연결하여, 예측함

로써 일사량 예측에 적합한 다중 딥러닝 모델을 구축하였다. 실험결과 다음과 같은 결과를 얻을 수 있었다.

1. 우수한 예측을 위해서는 딥러닝 모델 뿐 아니라 우수한 데이터셋이 필요하다. CNN 모델의 입력 데이터인 영상의 화질 문제로 오차가 다른 모델에 비해서 증가하였다.
2. CNN 모델에서 전체 영상을 학습하는 것 보다는 관심영역(region of interest)만을 추출하여, 학습한 모델이 예측률이 좋았다. CNN은 보다 우수한 화질을 이용하면 더욱 예측률이 향상될 것으로 예상된다.
3. 모든 모델이 깊은 층과 노드 수를 늘리면 과적합에 의해 학습 오차가 좋지 않았으며, LSTM 모델은 예상처럼 시계열 데이터인 기상 데이터에 대해서 가장 좋은 예측률을 보였다.
4. 두 개 이상의 다중 딥러닝 모델을 이용하여 학습하는 것이 하나의 모델로 학습하는 것이 학습효과가 좋았다. 즉, 두 개 이상의 다중 딥러닝 모델은 중간층의 깊이를 줄이면서도 많은 파라미터로 학습함으로써 좋은 예측률을 보였다.

VI. 결 론

딥러닝 모델을 이용한 예측은 우수한 학습 데이터가 기본이 되어야 하며, 학습 데이터의 특성에 맞는 모델을 선택하는 것이 기본이다. 딥러닝 모델은 입력과 출력 사이의 중간층이 깊으면 활성화 함수의 특성과 과적합에 의해 오차율이 떨어짐으로서 모델의 성능이 저하되는 특성이 있다. 이러한 딥러닝의 문제점을 해결하기 위해 ResNet이나 Inception과 같은 모델들을 사용하기도 한다.

본 연구는 이러한 선행 네트워크를 고려하여 두 개 이상의 다중 딥러닝 모델을 구축하여, 일사량 예측에 적합한 딥러닝 모델을 구축하였다. 실험 결과 두 개 이상의 모델을 연결한 다중 딥러닝 모델이 향상된 예측율과 안정된 학습 결과를 보였다. 즉, 하나의 모델로 깊은 층을 구성하고 많은 노드 수로 학습함으로써 발생하는 문제점인 과적합에 의한 예측률 감소를 방지하는 방법을 제시하였다.

Acknowledgments

본 연구는 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구 과제입니다. (No. 20194030202290)

References

- [1] M. A. Islam, M. Hasanuzzaman, N. A. Rahim, A. Nahar, and M. Hosenuzzaman, "Global renewable energy-based electricity generation and smart grid system for energy security," *The Scientific World Journal*, Vol. 2014, pp. 1-13, Aug. 2014.
- [2] Y. R. Lee, H. J. Kim, D. H. Lee, C. J. Lee and D. H. Lee, "Validation of Forecasting Performance of Two-Stage Probabilistic Solar Irradiation and Solar Power Forecasting Algorithm using XGBoost," *The transactions of The Korean Institute of Electrical Engineers*, Vol. 68, No. 12, pp. 1704-1710, DEC, 2019.
- [3] J. H. Choi, D. W. An, and J. H. Gang, "A Survey on Prognostics and Comparison Study on the Model-Based Prognostics," *Journal of institute of control, robotics and systems*, Vol. 17, No. 11, pp. 1095 - 1100, 2011.
- [4] S. H. Hwan, T. G. Kim, and S. H. Lee and J. Y. Choi, "Trend Analysis of Projected Climate Data based on CMIP5 GCMs for Climate Change Impact Assessment on Agricultural Water Resources," *Journal of The Korean Society of Agricultural Engineers*, Vol. 57, No. 5, pp. 69- 80, 2015.
- [5] J. H. Park, D. H. Shin, and C. B. Kim, "Deep Learning Model for Electric Power Demand Prediction Using Special Day Separation and Prediction Elements Extention," *Journal of Advanced Navigation Technology*, Vol. 21, No. 4, pp. 365-370, 2017.
- [6] Y. S. Kim, and H. J. Park, "Modeling and Predicting South Korea's Daily Electric Demand Using DNN and LSTM," *Journal of Climate Change Research*, Vol. 12, No. 3, pp. 241-253, 2021.
- [7] Y. H. Lee and Y. S. Kim, "Comparison of CNN and YOLO for Object Detection," *Journal of the Semiconductor & Display Technology*, Vol. 19, No. 1, pp. 86-92, 2020.
- [8] Y. S. Kim, S. H. Lee and H. W. Kim, "Prediction Method of Photovoltaic Power Generation Based on LSTM Using Weather Information", Vol. 44, No. 12, pp. 2231-2238, 2019.
- [9] J. S. Sim and D. S. Song, "Effect of Atmospheric Particulate Matter Concentration on Solar Insolation," *The Society Of Air-Conditioning And Refrigerating Engineers Of Korea*, Pyeongchang : Yongpyeong Resort, pp. 167-170, 2019.



김 재 정 (jae-Jung Kim)

2015년 3월 ~ 현재: 가천대학교 에너지IT학과 재학
관심분야: 딥러닝, 빅 데이터, IoT, 스마트 폰 어플리케이션



유 용 훈 (Yong-Hun You)

2015년 3월 ~ 현재: 가천대학교 에너지IT학과 재학
관심분야: 데이터 마이닝, 데이터 베이스, 딥러닝, IoT



김 창 복 (Chang-Bok Kim)

1986 2월 : 단국대학교 전자공학학 졸업(공학사)
1989년 2월 : 단국대학교 전자공학과 (공학석사)
2009년 2월 : 인천대학교 컴퓨터 공학과(공학박사)
1994년 ~ 현재 : 가천대학교 IT대학 전기공학과 교수
※ 관심분야: 데이터 마이닝, 딥러닝, 강화학습, 사물인터넷