

MATLAB EXPO

구름영상예측을 이용한 태양광 발전량 예측의 고도화

최승연 교수, 고려대학교



목차

1. 회사 및 발표자 약력 (Introduction to Organization and Business)
2. 프로젝트 개요 (Project Overview)
3. 기술적인 해결과제 (Project Goals and Challenges)
4. MathWorks 솔루션을 통한 해결 방안 및 결과 (How did we get there and leverage MathWorks)
5. 결과 및 정리 (Achievements and Outlook)
6. 다음 과제 (Further Details on Solutions Adopted)*
7. 결론 (Concluding Remarks)

회사 및 발표자 약력



최 승 연

전기전자공학부 부교수
고려대학교

- 조지아공대 공학박사 (2013)
- 한국전기연구원 스마트전력망연구센터 (2014-2018)
- 고려대학교 전기전자공학부 (2018-현재)

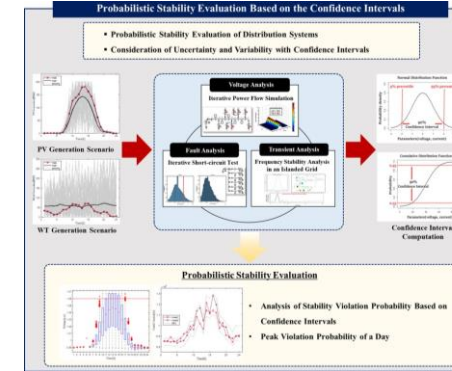
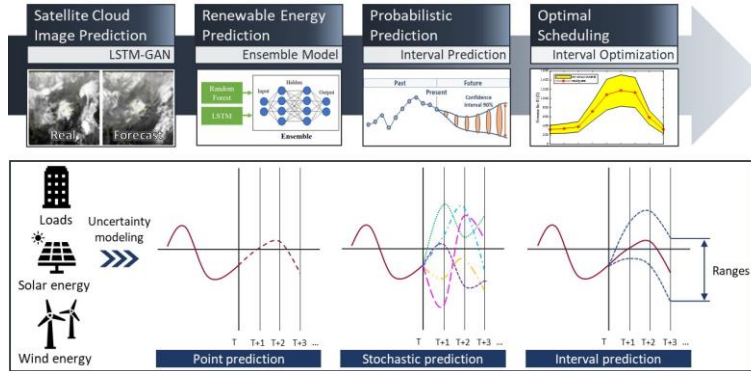
- 미래배전시스템 연구실(Smart Microgrid Lab) 운영
- MATLAB®와 Simulink®를 기반으로 Simscape Electrical, 최적화, 기계학습 등의 Toolbox 활용하여 재생에너지 예측과 전력시스템 해석 및 최적화 연구

미래배전시스템 연구실 (Smart Microgrid Lab)

| Advanced Distribution System Operation (AdvDSO)

신재생에너지 불확실성 대비 실시간 운영 최적화
기법 연구

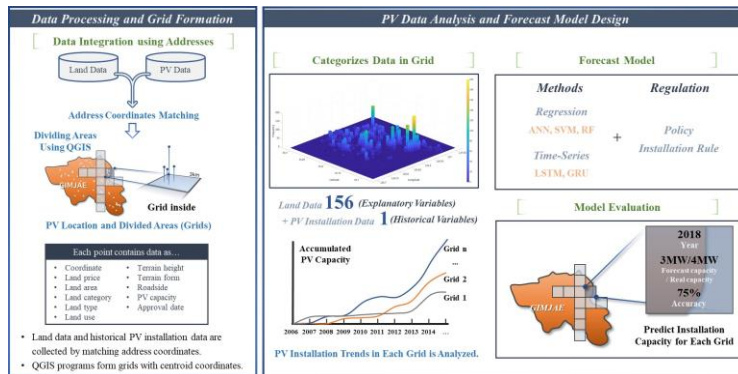
신재생에너지 변동성을 고려한 확률론적 안정도
분석 연구



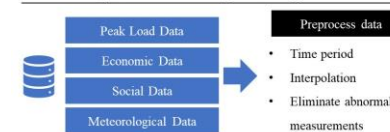
| Advanced Distribution System Planning (AdvDSP)

중장기 배전계통 분산자원 연계 예측기법 연구

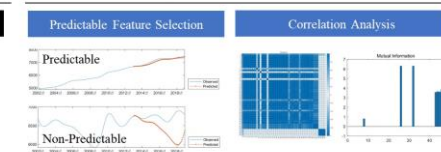
머신러닝 기반 중장기 배전계통 부하 예측기법
연구



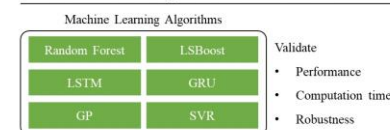
1. Data Acquisition and Preprocessing



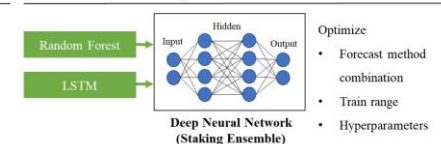
2. Feature Selection



3. Machine Learning Performance Validation



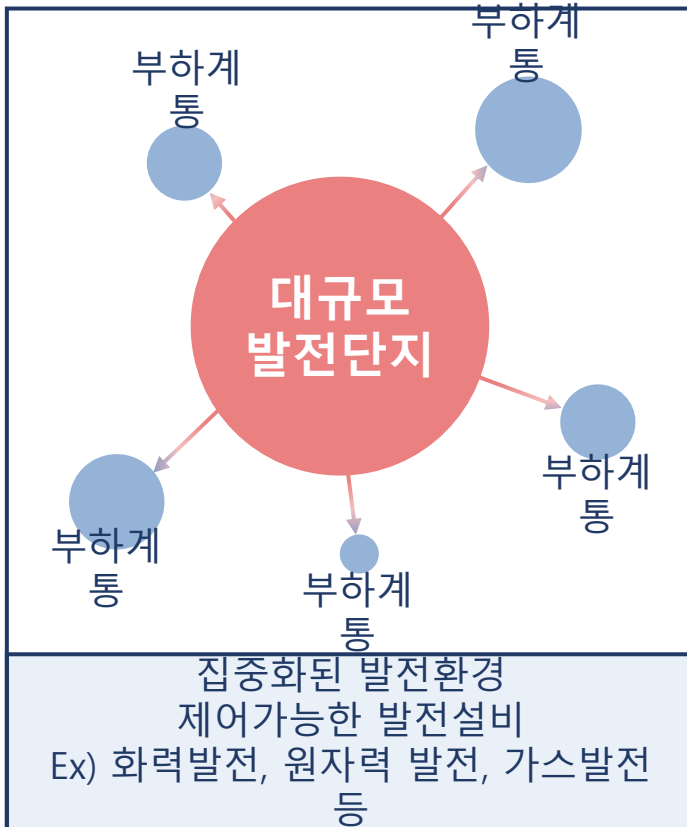
4. Ensemble Model Optimization



프로젝트 개요

| 연구목적

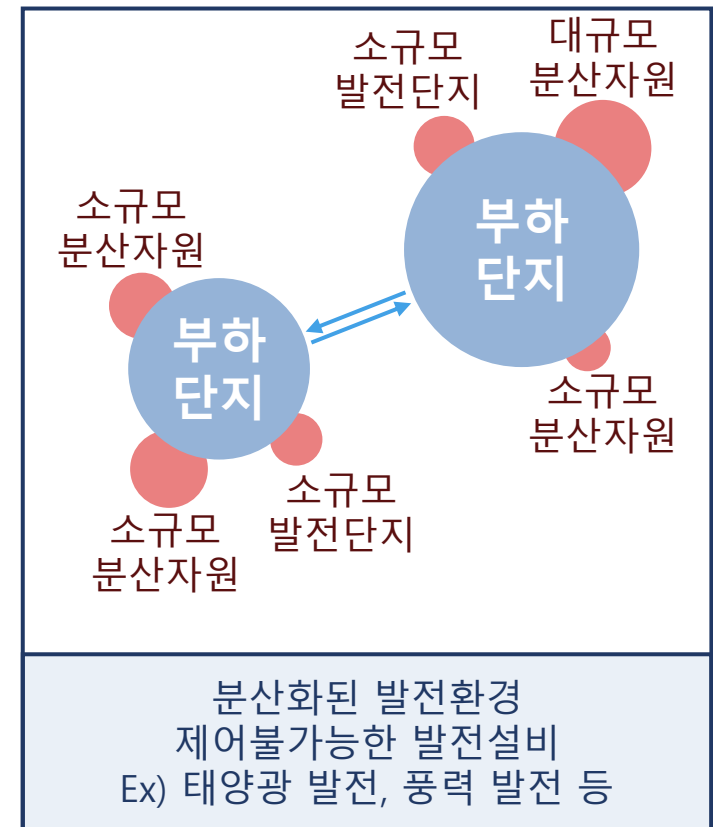
기존 전력계통



신재생에너지 발전용량 비중
전망

20%
2030

미래 전력계통



"전력 공급 및 수요 균형 유지를
위한 신재생에너지 발전량 예측
필요"

프로젝트 개요

| 연구배경

기존연구

신재생에너지 발전량 예측을 위한 데이터

- 기상 정보(일사량, 운량, 온도 등)
- 발전량 정보



신재생에너지 발전량 예측을 위한 방법론

- 회귀기법 : LR
- 통계학적기법 : ARIMA
- 인공지능기법 : ANN, MLP
- 머신러닝기법 : SVR
- 딥러닝기법 : LSTM, CNN

기존연구의 한계점

1 예측정보의 제한성

- 대부분의 기상사이트에서는 일사량과 운량 등 태양광 발전량과 상관성이 큰 예측정보를 제공하고 있지 않음
- 기상사이트의 예측 정확도에 따라 모델의 정확도가 달라짐

2 예측모델의 실용성

- 회귀 모델의 경우, 높은 상관성을 가진 예측정보가 필요하지만, 상관관계가 낮은 예보 데이터만 제공됨
- 시계열 모델의 경우, 외부요인에 의한 발전량 변동을 반영하기 어려움

제안하는 방법론

[분석]

- 일사량과 운량을 대체할 수 있는 새로운 예측 데이터 필요
- 위성 이미지는 기본적으로 구름의 위치, 형상, 양 등 다양한 정보를 가지고 있음
- 국가기상위성센터의 천리안 1호는 15분 단위로 위성 이미지가 제공함

[가정]

시계열 위성 이미지를 학습하여, 미래의 위성 이미지를 얻을 수 있다면, 이미지에서 원하는 지역에 해당하는 픽셀 데이터를 활용하여 운량 정보로 변환가능

[문제점]

고해상도 이미지의 경우, 데이터의 계산 복잡성이 급격하게 증가하여, 자원과 시간 소모가 많음

프로젝트 개요

LSTM-GAN 기반의 위성 이미지 예측 방법론 개요

데이터 처리 과정

데이터 수집 및 정리

☀️ 기상 데이터

🛰️ 위성 데이터

일사량
운량
온도
습도
...

적외선
가시광선
수증기
...

데이터베이스
화

- 현재 기상청에서는 공식적으로 일사량과 운량에 대한 예측정보를 제공하고 있기 때문에, 다른 경로를 통해 데이터를 수집해야 함

→ 구름의 위치나 형상은 위성 이미지에서 추출 가능함

데이터 전처리



Database

❌ Missing Data

❌ Outlier

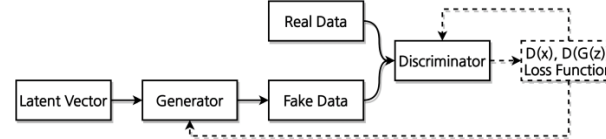
- 측정 데이터는 일반적으로 누락 데이터나 이상치가 존재할 가능성이 크기 때문에, 별도로 처리하는 과정이 필요함

→ 정밀한 상관관계 분석을 위해 데이터 필터링 또는 보간 과정이 필요함

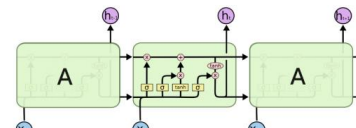
예측 방법론

구름 이미지 예측과정

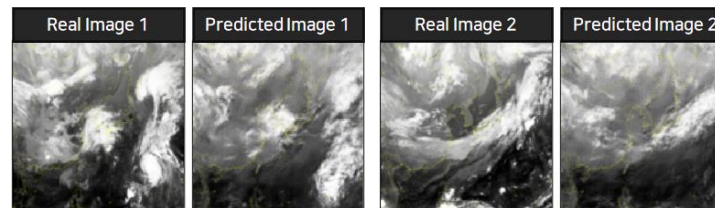
GAN (Generative Adversarial Network)



LSTM (Long Short-Term Memory)



시계열 모델과 생성 모델을 기반으로 구름 이미지를 예측함

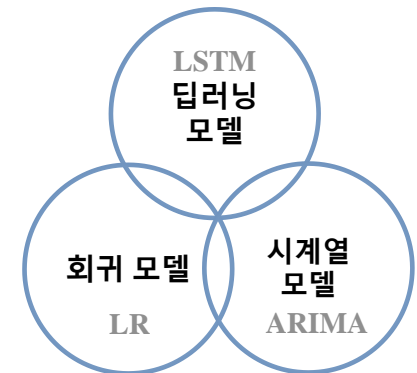


픽셀의 밝기가 구름의 양과 두께를 나타냄

앙상블 기반 예측성능 개선

기상정보 데이터베이스 + 구름 예측 정보

최적의 입력 데이터 조합 테스트



Ensemble Learning

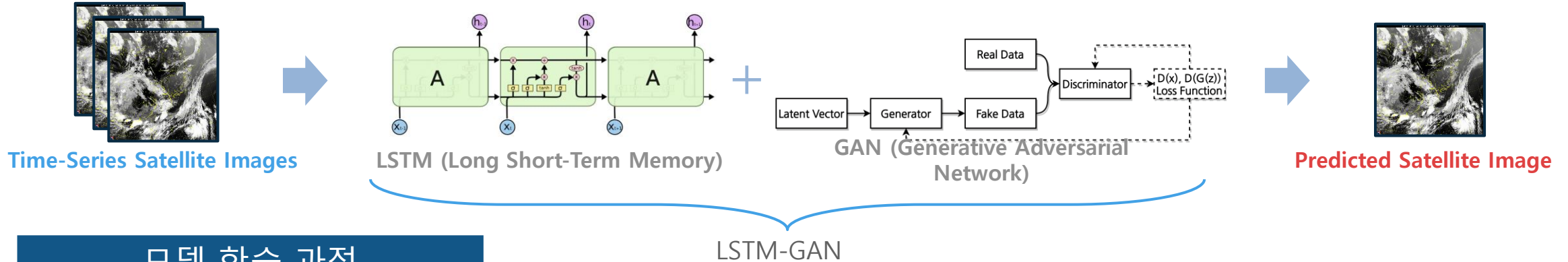
- Voting
- Bagging
- Boosting
- 다양한 딥러닝 모델을 기반으로 하는 앙상블 모델을 이용하여 정확도를 개선함

프로젝트 개요

LSTM-GAN 기반의 위성 이미지 예측 과정

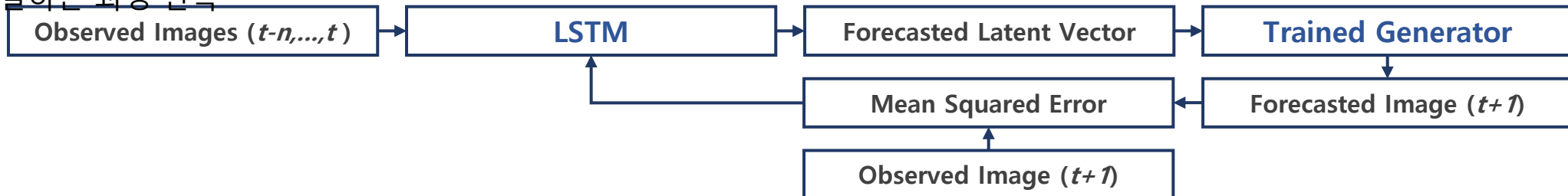
각 모델의 역할

- Deep Convolutional GAN | 고품질의 위성 이미지의 특징을 학습하여, 잠재벡터 기반의 효율적인 이미지 생성 수행
- Convolutional LSTM | 시계열 이미지를 입력 받아 설계자가 의도하는 임의의 출력 생성
- LSTM-GAN | 시계열 이미지를 입력 받아 잠재벡터를 생성하고, 생성 모델에 이를 입력하여 새로운 이미지 생성



모델 학습 과정

- 관측된 이미지를 이용하여, LSTM(ConvLSTM)은 미래 시점의 잠재벡터(Forecasted Latent Vector)를 예측함
- 예측된 잠재벡터는 학습이 완료된 생성자(Trained Generator)에 입력되어 예측 위성 이미지(Forecasted Image)를 생성함
- 예측 위성 이미지와 실제 위성 이미지(Observed Image)의 오차를 평균제곱오차(Mean Squared Error)로 계산한 후, LSTM에 전달하는 과정 반복



기술적인 해결과제

| 위성 이미지 특성

위성 이미지 특성

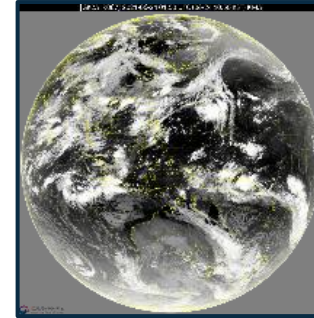
- 데이터 출처 | 국가기상위성센터 천리안 1호
- 데이터 제공기간 | 2011.04.01 ~ 2020.04.01 (9년)
- 데이터 주기 | 15분
- 데이터 범위 | 전구, 동아시아, 한반도
- 데이터 종류 | 적외선, 가시광선, 수증기 등

위성 이미지 사용방안

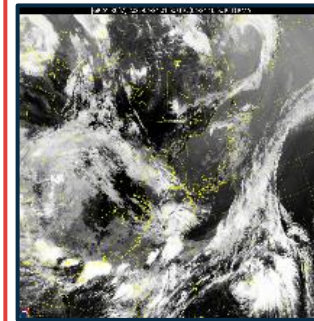
- 구름의 형상과 방향은 시계열 구름 이미지를 통해 관찰 가능함
- 시계열 구름 이미지를 입력하여 딥러닝 모델을 학습함
- 새로운 시계열 이미지를 입력하여 모델이 위성 이미지를 잘 예측하는지는 지상에서 위성 이미지와 비교함

문제점

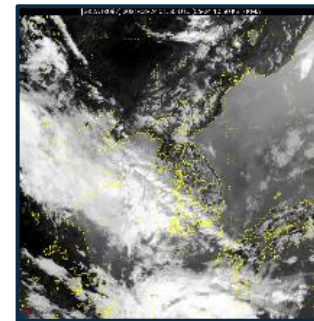
- 데이터 집합의 모든 이미지를 작업공간에 업로드하여 학습할 시,
램 용량부족으로 인한 오류 발생
- 위성 이미지는 1300 * 1500 pixel의 해상도로 구성된 대용량 데이터로 은닉층이 증가할수록 계산 복잡성이 급격하게 증가하여, 제대로 된 학습이 진행되지 않음



전구 범위
 관찰범위가 너무 넓음



동아시아 범위
 구름의 형상 및 이동 반영 가능



한반도 범위
 구름의 형상이 잘 표현되지 않음

MathWorks 솔루션을 통한 해결 방안 및 결과

| MATLAB 라이브러리 및 기술지원 활용

대용량 이미지 처리

- 주요함수
 - ImageDatastore | 데이터의 폴더나 파일 주소를 지정
 - ArgumentedImageDatastore | 원하는 형태로 데이터 일괄 수정
- 모든 이미지 데이터를 한번에 메모리에 업로드 하여 작업할 필요가 없으므로, 컴퓨터 자원소모 감소 및 대용량 자료 활용 가능
- 객체함수를 이용하여 다양한 데이터저장소 설정 가능

객체 함수	
countEachLabel	ImageDatastore 레이블의 파일 개수 계산
hasdata	데이터를 읽을 수 있는지 확인
numpartitions	데이터저장소 파티션 개수
partition	데이터저장소 파티셔닝
preview	데이터저장소에 있는 데이터의 서브셋 미리보기
read	데이터저장소의 데이터 읽어 들이기
readall	데이터저장소의 모든 데이터 읽기
readimage	데이터저장소에서 지정된 이미지 읽기
writeall	Write datastore to files
reset	데이터저장소를 초기 상태로 재설정
shuffle	데이터저장소의 모든 데이터 읽기 섞기
splitEachLabel	비율을 기준으로 ImageDatastore 레이블 분할
subset	Create subset of datastore or file-set
transform	Transform datastore
combine	여러 데이터저장소의 데이터 결합
isPartitionable	Determine whether datastore is partitionable
isShuffleable	Determine whether datastore is shuffleable

기존방법의 한계점



메모리가 허용하는 제한된 양의 데이터 집합만 업로드 가능하므로, 데이터 활용성이 제한됨
(전체 데이터 양 < 램 메모리)

MathWorks 솔루션 적용



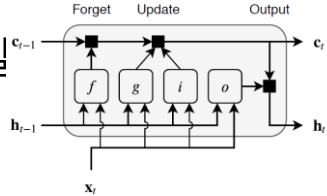
실제 계산을 수행하는 데이터만 불러와서 작업한 후, 새로운 데이터를 업로드하므로 모든 저장데이터에 대한 작업수행 가능
(한 번에 불러오는 데이터 양 < 램 메모리)

MathWorks 솔루션을 통한 해결 방안 및 결과

| MATLAB 라이브러리 및 기술지원 활용

MATLAB LSTM 라이브러리

- 설정된 함수를 통해 신경망의 구조를 쉽게 커스텀 가능
- 하이퍼파라미터 설정을 통한 모델의 성능 개선 가능
- 도움말



제공

이 번역 페이지는 최신 내용을 담고 있지 않습니다. 최신 내용을 영문으로 보려면 여기를 클릭하십시오.

장단기 기억 신경망

이 페이지에서는 장단기 기억(LSTM) 신경망을 사용하여 분류 및 회귀 작업에 대해 시퀀스 및 시계열 데이터로 작업하는 방법을 설명합니다. LSTM 신경망을 사용하여 시퀀스 LSTM 신경망은 시퀀스 데이터의 시간 스텝 간의 장기적인 종속성을 학습할 수 있는 순환 신경망(RNN)의 일종입니다.

LSTM 신경망 아키텍처

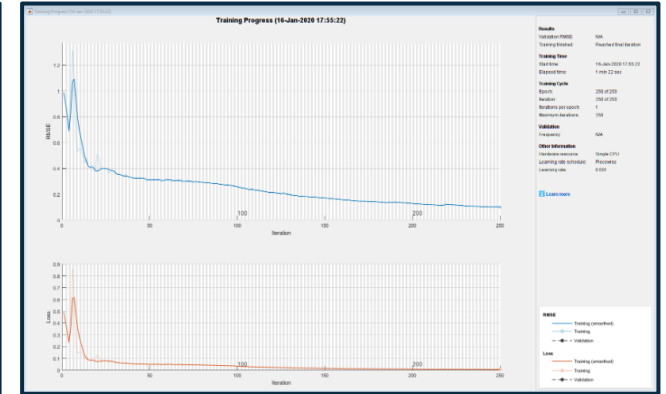
LSTM 신경망의 핵심 구성요소는 시퀀스 입력 계층과 LSTM 계층입니다. 시퀀스 입력 계층은 신경망에 시퀀스 또는 시계열 데이터를 입력합니다. LSTM 계층은 시퀀스 데이터 다음 도식을 분류를 위한 간단한 LSTM 신경망의 아키텍처를 보여줍니다. 이 신경망은 시퀀스 입력 계층으로 시작하고, 그 뒤에 LSTM 계층이 옵니다. 출력 레이어를 예측합니다.

다음 도식은 회귀를 위한 간단한 LSTM 신경망의 아키텍처를 보여줍니다. 이 신경망은 시퀀스 입력 계층으로 시작하고, 그 뒤에 LSTM 계층이 옵니다. 신경망의 끝부분에는 완전 연결 계층이 있습니다. 이 신경망은 시퀀스 입력 계층으로 시작하고, 그 뒤에 LSTM 계층이 옵니다. 신경망의 끝부분에는 완전 연결 계층이 있습니다. 이 신경망은 시퀀스 입력 계층으로 시작하고, 그 뒤에 LSTM 계층이 옵니다. 신경망의 끝부분에는 완전 연결 계층이 있습니다.

다음 도식은 비디오 분류를 위한 신경망의 아키텍처를 보여줍니다. 신경망에 영상 시퀀스를 입력하려면 시퀀스 입력 계층을 사용하십시오. 특징 추출을 위해 컨볼루션 계층으로 구성된 시퀀스에서 학습하도록 LSTM 계층을 사용하려면 평탄화 계층 뒤에 LSTM 계층과 풀링 계층을 사용하십시오.

분류 LSTM 신경망

sequence-to-label 분류를 위한 LSTM 신경망을 만들려면 시퀀스 입력 계층, LSTM 계층, 완전 연결 계층, 소프트맥스 계층, 분류 출력 계층을 포함하는 계층 배열을 만듭니다. 시퀀스 입력 계층의 크기를 입력 데이터의 특징 개수로 설정합니다. 완전 연결 계층의 크기를 클래스 개수로 설정합니다. 시퀀스 길이는 지정할 필요가 없습니다. LSTM 계층의 경우, 은닉 유닛의 개수와 출력 모드 "last"를 지정합니다.

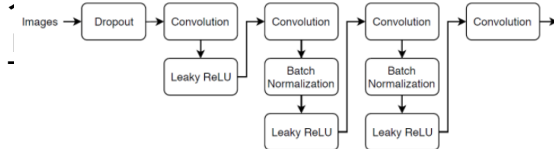


MathWorks 도움말 센터 LSTM 예제

LSTM 학습과정 예시

MATLAB GAN 라이브러리

- 설정된 함수를 통해 신경망의 구조를 쉽게 커스텀 가능
- 하이퍼파라미터 설정을 통한 모델의 성능 개선 가능



생성적 적대 신경망(GAN) 훈련시키기

이 예제에서는 생성적 적대 신경망을 훈련시켜서 영상을 생성하는 방법을 보여줍니다.

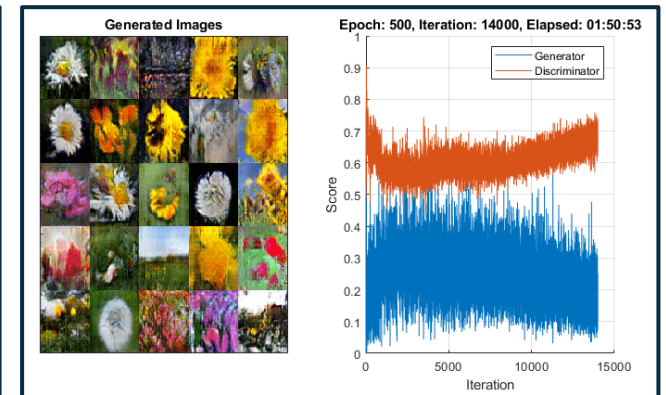
생성적 적대 신경망(GAN)은 입력값인 실제 데이터와 비슷한 특징을 갖는 데이터를 생성할 수 있는 딥러닝 신경망의 한 종류입니다.

GAN은 함께 훈련되는 다음과 같은 두 개의 신경망으로 구성됩니다.

1. 생성기 — 입력값으로 난수로 구성된 벡터(잠재 입력값)가 주어지면 이 신경망은 훈련 데이터와 동일한 구조를 갖는 데이터를 생성합니다.
2. 판별기 — 두 훈련 데이터의 관측값을 포함하는 데이터의 배치(batch)와 생성기에서 생성된 데이터가 주어지면 이 신경망은 관측값을 "real"과 "generated"로 분류하려고 시도합니다.

GAN을 훈련시키려면 두 신경망의 성능을 극대화할 수 있도록 다음과 같이 두 신경망을 동시에 훈련시키십시오.

- 판별기 "속이"는 "속이" 데이터를 생성하도록 생성기를 훈련시킵니다.
- 실제 데이터와 생성된 데이터를 구분하도록 판별기를 훈련시킵니다.



MathWorks 도움말 센터 GAN 예제

GAN 결과 예시

MathWorks 솔루션을 통한 해결 방안 및 결과

| MATLAB 라이브러리 및 기술지원 활용

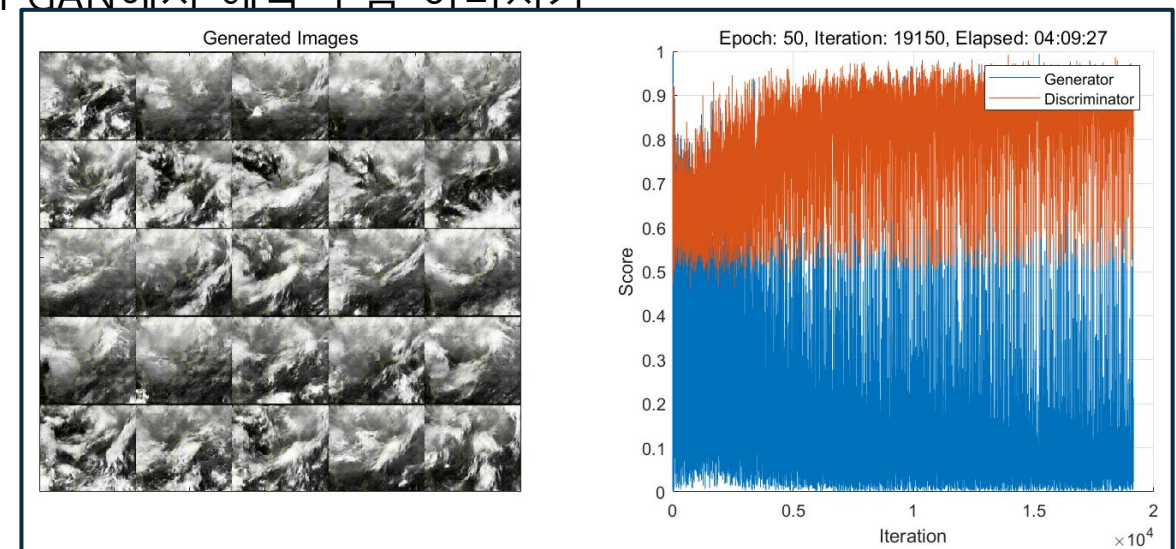
LSTM-GAN 코드지원

- LSTM, GAN은 예제코드가 제공되어 있지만, LSTM-GAN 연계에 대한 예제코드는 제공되어 있지 않음
- MathWorks 기술지원을 통해 LSTM-GAN 연계코드를 제공 받음
- 하이퍼파라미터 조정을 통한 모델 성능 평가
- GAN을 통해 구름 이미지 모사 성능을 확인하고 LSTM-GAN에서 예측 구름 이미지가

```

1 %% Train Generative Adversarial Network (GAN)
2 % This example shows how to train a generative adversarial network to generate
3 % images.
4 %
5 % A generative adversarial network (GAN) is a type of deep learning network
6 % that can generate data with similar characteristics as the input real data.
7 %
8 % A GAN consists of two networks that train together:
9 %%
10 % # Generator — Given a vector of random values (latent inputs) as input, this
11 % network generates data with the same structure as the training data.
12 % # Discriminator — Given batches of data containing observations from both
13 % the training data, and generated data from the generator, this network attempts
14 % to classify the observations as 'real' or 'generated'.
15 %%
16 %
17 %
18 % To train a GAN, train both networks simultaneously to maximize the performance
19 % of both:
20 %%
21 % * Train the generator to generate data that "fools" the discriminator.
22 % * Train the discriminator to distinguish between real and generated data.
23 %%
24 % To optimize the performance of the generator, maximize the loss of the discriminator
25 % when given generated data. That is, the objective of the generator is to generate
26 % data that the discriminator classifies as 'real'.
27 %
28 % To optimize the performance of the discriminator, minimize the loss of the
29 % discriminator when given batches of both real and generated data. That is, the
30 % objective of the discriminator is to not be 'fooled' by the generator.
31 %
32 % Ideally, these strategies result in a generator that generates convincingly
33 % realistic data and a discriminator that has learned strong feature representations
34 % that are characteristic of the training data.
35 %% Load Training Data
36 % Create an image datastore.
    
```

MathWorks 기술지원을 통한 LSTM-GAN 코드 예시



LSTM-GAN 모델을 통한 위성 이미지 학습과정 예시

결과 및 정리

| LSTM-GAN 기반의 위성 이미지 예측 결과

GAN 이미지 생성 결과

- 생성모델의 성능평가는 객관적인 지표가 없음
- 일반적으로 아래 평가기준으로 모델의 성능 판단
 - 실제 이미지와 생성된 이미지의 유사성 판단
 - 이미지의 다양성 확인
 - 잠재벡터의 선형적 변화에 대한 출력 이미지의 선형적 변화

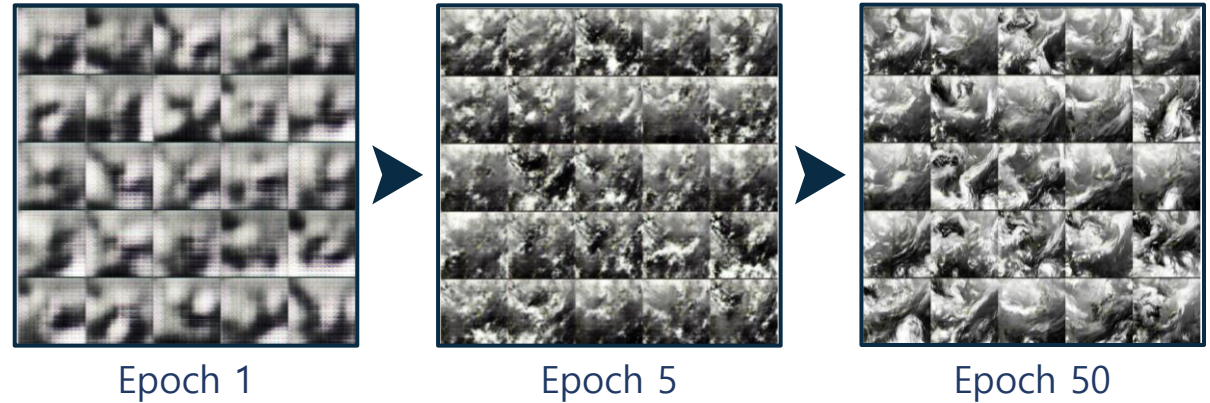
LSTM-GAN 이미지 생성 결과

- 실험환경
 - 6시간(24개 이미지)를 시계열로 입력
 - 3시간 이후의 구름 이미지를 예측
- 이미지 예측 성능 판단
 - 실제 이미지와 생성된 이미지의 유사성 확인
 - 구름의 이동 및 형상 예측 확인

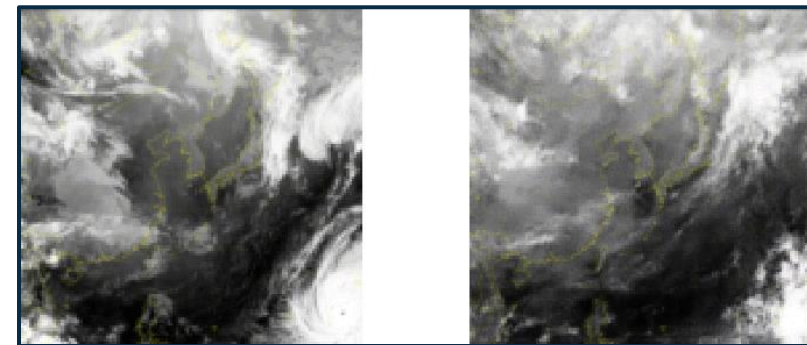
시뮬레이션 환경

- Windows 10 Pro for Workstations 64 bit
- Intel® Xeon® W-2133 CPU @3.60GHz (12 CPUs)
- 64GB RAM
- Nvidia Quadro RTX 5000

GAN Image Training Process



LSTM-GAN Prediction Result



Ground Truth

Predicted Image

결과 및 정리

| LSTM-GAN 모델 기반 태양광 발전량 예측 결과

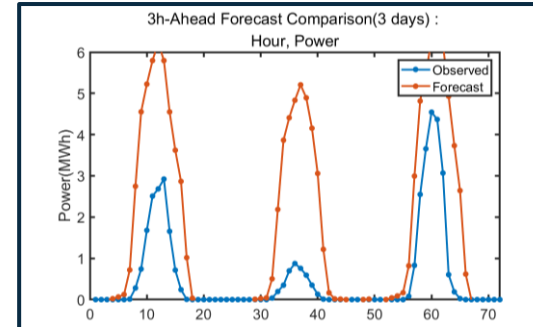
태양광 발전량 예측 결과

- 4가지 입력변수 조합 시험을 통한 위성 이미지 데이터 실효성 검증
 - 전력 데이터, 시간
 - 전력 데이터, 시간, 기상 데이터(온도, 습도)
 - 전력 데이터, 시간, 위성 데이터
 - 전력 데이터, 시간, 기상 데이터, 위성 데이터
- 3시간 후 예측하여 성능 비교
- 3일 동안의 실제 태양광 발전량과 예측 발전량 간 평균오차 산정

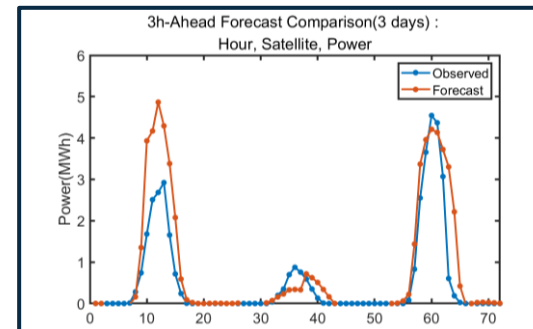
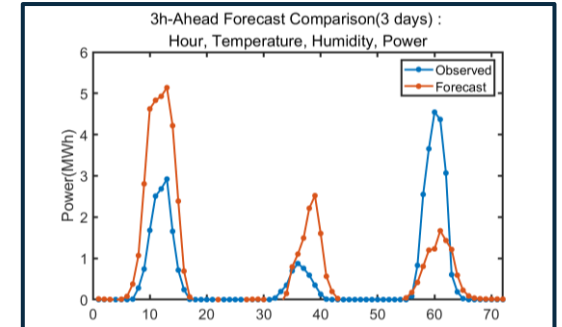
- 예측모형 사용

Input Parameter	1 month MSE
Power	2.05
Power, Meteorological	0.97
Power, Satellite	1.27
Power, Meteorological, Satellite	0.94

MSE Comparison of PV Forecasting



PV Power Forecasting *WITHOUT* Satellite Image Prediction (3 days)



PV Power Forecasting *WITH* Satellite Image Prediction (3 days)

- 위성 이미지 기반 예측 데이터가 포함된 예측모델이 포함하지 않은 예측모델보다 전반적으로 우수한 결과를

다음과제

| LSTM-GAN 모델 성능 개선



최적 하이퍼
파라미터 탐색

예측모델의 최적
하이퍼파라미터 탐색을 통한
위성 이미지
예측성능 개선



최신 Hybrid
Model 적용

LSTM-GAN 외 GRU-GAN,
BiLSTM-GAN 등 다양한 Hybrid
Model에 데이터를 적용해보고
모델의 성능을 평가



다양한 위성
이미지 데이터
실험

적외선 이미지 외 가시광선
이미지, 수증기 이미지 등
다양한 구름영상을 통한 구름
이동 예측 개선



앙상블 모델
적용

다수의 모델을 통해 이미지를
예측하고, 결과를 앙상블하여
이미지 유사도가 개선되는지
검토

결론

- 연구목적

- 위성 이미지 예측 가능성 검증 및 태양광 발전량 예측성능 개선여부 확인

- 연구결론

- MATLAB을 이용하여, 시계열 이미지 생성이 가능한 LSTM-GAN 모델을 쉽게 구현하였음
- MATLAB을 이용하여, GAN 모델이 고품질과 다양한 패턴의 위성 이미지를 생성하는 것을 검증함
- 설계된 LSTM-GAN 모델이 구름의 패턴을 예측할 수 있다는 것을 확인함
- 예측된 이미지 값을 새로운 입력변수로 활용하였을 시, 태양광 예측의 정확도가 개선되는 것을 검증함

- MATLAB 활용성

- Deep Learning Toolbox를 이용하면 다양한 분야에서 쉽고 효과적으로 AI를 적용할 수 있음.
- 본 연구진은 MathWorks의 기술지원을 통해 연구에 필요한 예측모델을 설계할 수 있었고, 해당 모델이 태양광 예측성능 개선에 기여한다는 점을 검증할 수 있었음

MATLAB EXPO

Thank you



© 2022 The MathWorks, Inc. MATLAB and Simulink are registered trademarks of The MathWorks, Inc. See [mathworks.com/trademarks](https://www.mathworks.com/trademarks) for a list of additional trademarks. Other product or brand names may be trademarks or registered trademarks of their respective holders.