Deep Learning for Air Quality Forecasting

Regional Air Quality Planning Advisory Committee

UH AI AQF group
PI: Dr. Yunsoo Choi
Department of Earth and Atmospheric Sciences
University of Houston

January 2020
Numerical model results from UH web
이세돌 vs 알파고 ‘세기의 대국’

- 날짜: 9일, 10일, 12일, 13일, 15일 (각 오후 1시)
- 장소: 서울 광화문 포시즌스호텔 특별 대국장
- 시간: 각 2시간, 1분 초읽기 3회
- 상금: 100만 달러, 알파고 승리 시 상금은 자선단체에 기부
- 규칙: 중국식 룰을 적용해 백을 짐을 기사에게 7집 반 제공
- 대국 방식: 알파고 개발자 아자 황(아마 6단)이 알파고가 둔 수를 바둑판에 놓고, 이세돌 9단이 둔 수를 컴퓨터에 입력
- 생중계: 네이버 · 유튜브 · 바둑TV · 에브리온TV · 아프리카TV 전 경기 중계
Machine Learning & Deep Learning

- How Deep Learning is different from other methods?
Application of machine learning in atmospheric sciences (published or submitted/prepared):

✓ Prediction and model:
  • Weather and hurricane forecasting (Salman et al., 2015; Hsieh et al., 2015; Zhang et al., 2016; Wang et al. 2017; Eslami et al., 2020; Sayeed et al., 2020c).
  • Air quality and pollen predictions (Li et al., 2016 & 2017; Pardo, et al., 2017; Fan et al., 2017; Bui et al., 2018; Eslami et al., 2019a; Eslami et al., 2019b; Lops et al., 2019; Sayeed et al., 2020a; Sayeed et al., 2020b).
  • Indoor air pollution (Ahn et al., 2017)

✓ Approximation and analysis:
  • Identifying significance of predictors in ozone and PM forecasting (Yeo et al., 2020; Salman et al., 2020)

✓ Classification and clustering:
  • Predicting extreme climate pattern and climate events (Iglesias et al., 2015).
  • Discovering spatial and temporal patterns in climate and air pollution data (Anderson et al., 2015; Zhang et al., 2016, Kotsakis et al., 2019; Jeon et al., 2019).
Deep Learning

TRAINING
During the training phase, a neural network is fed thousands of labeled images of various animals, learning to classify them.

INPUT
An unlabeled image is shown to the pretrained network.

FIRST LAYER
The neurons respond to different simple shapes, like edges.

HIGHER LAYER
Neurons respond to more complex structures.

TOP LAYER
Neurons respond to highly complex, abstract concepts that we would identify as different animals.

OUTPUT
The network predicts what the object most likely is, based on its training.

10% WOLF
90% DOG
Deep Neural Network

Deep MultiLayer Perceptrons (MLP)
Deep Neural Network

- **Input variables**
  - Meteorology
  - Air quality
  - Emissions

- **Target output**
  - Hourly prediction of air pollution

- **Characteristics**
  - Seasonal
  - Weekly/bi-weekly
  - Hourly/bi-hourly
Convolutional Neural Network

Regressive Deep Convolutional Neural Network:

- Meteorology
- Air quality
- Emissions

Input variables → Input layer → Feature maps layer 1 → Feature maps layer 2 → ... → Fully-connected layer → Output layer

Hourly prediction of air pollution
UH DNN real-time ozone prediction system for Texas, USA:

TCEQ CAMS
- Temperature
- Precipitation
- Wind field
- RH
- Dew point T
- Pressure
- Solar/Net radiation
- Nitrogen oxides
- Ozone

Inputs
- Meteorology
- Air Pollution

Deep Neural Network
- CNN

Real-time hourly forecasting of ozone

DNN modeling time period:
- Next day prediction: 2017 (training is updated every day)

Houston ozone

Monthly mean comparison between observation and CNN prediction of all CAMS stations:
Houston ozone

Model-measurement comparisons for a daily mean of concentrations
Accuracy comparison

Monthly Index of Agreement of all CAMS stations over Texas:
Houston pollen forecasting

UH DNN real-time pollen prediction system for Houston, TX:

- Inputs:
  - Pollen
  - Meteorology
  - Processed Data

- Deep Neural Network (CNN)

- 1-7 Day Pollen forecasting

TCEQ CAMS:
- Temperature
- Precipitation
- Wind field
- RH
- Solar Radiation
- Pressure

Houston Health Department:
- 24 Tree Species
- 15 Weed Species
- Grass

DNN modeling time period:
- Training data: 2009–2015
- 1-7 day prediction: 2016

✓ Missing data imputation: Butterworth filter method.
Houston pollen forecasting

Total Pollen Concentration Observed vs Predicted

- Observation
- CNN_Forecast

Index of Agreement

Pearson Correlation Coefficient

Days Predicting Ahead

Mean IOA
Mean r
Min-Max Range
CMAQ bias-correction AI over US

UH CMAQ-AI real-time ozone prediction system:

- DNN modeling time period (for continental U.S. – 1081 AQS stations):
  - Next day prediction: 2014 (April – October)
CMAQ bias-correction AI over Korea

• Index of Agreement (IOA) for 255 measurement sites
CMAQ bias-correction AI over Korea

- Daily Maximum ozone for one measurement site

Observation  CMAQ-AI ver2.0  CMAQ-AI ver2.1
CMAQ twin model AI

Modeling configuration:
- Sample Case: Seoul
- Training & Validation data: 2014 – 2017
- Verification: 2018
CMAQ twin model NO$_2$ concentration

![Graph showing NO$_2$ concentration over time](image)
CNN real-time image forecasting

AOD prediction (left) and hurricane tracking (right) are both image forecasting problems…

Hurricane Irma, 2017 (source: GOES, NOAA)
Methodology

Testing Image Forecasting with AI:

Question: Can AI predict basic movements from just receiving previous states with just image as input?

YES IT CAN!
Testing Image Forecasting: Part 2

Can the AI follow two features traveling independently and understand collisions between them?

YES IT CAN!
AOD real-time forecasting

Testing Image Forecasting: Part 3:

Applied 2D-CNN to forecast CMAQ AOD 3-hours ahead with just 3 images as input.

Model Accuracy after 3-month training:

✓ ≈ 0.8 IOA & COR
Deep learning can be used:

1) to forecast air quality forecasting
2) to improve air quality modeling results
3) to build up twin AI model to mimic air quality model
4) to accurately forecast remote sensing AOD images
On-going other studies

✓ A diagnosis tool for AI models
✓ AI inverse modeling
✓ AI data assimilation
✓ AI data augmentation – AI GAN model
✓ AI Reinforcement learning – Self-learning
✓ AI weather, Hurricane, and Climate models
✓ AI database, GUI-interface, and smartphone-app platforms
Thank you for your attention