~~Developing~~ ~~Researching for~~ Building Operational AI Weather Service in Taiwan

Hsuan-Tien Lin
林軒田

Professor, National Taiwan University

June 12, 2024
Institute of Information Science, Academia Sinica

Central Weather Administration in GPU Age

(downloaded from LibertyTimes for educational purpose)

Press Conference, 02/26/2024

- (自由時報)台積電助攻 氣象署第6代超級電腦啓用
- (聯合報)氣象署第6代高速運算電腦啓用 提升天氣預報準確度
- (中國時報)縣市長都笑了 氣象署啓用「超級電腦」 颱風假更精準
- (國語日報)氣象署啓用第6代超級電腦 預報更精準

(新聞稿)氣象署第6代高速運算電腦，採用台積電生產之7奈米ARM架構晶片所建構之富士通FX1000機型高速運算電腦……在世界高速運算電腦排名爲69名……同時，爲因應各種人工智慧發展及高效能運算的需求，搭配192片由輝達(NVIDIA)產製的A100繪圖處理單元(GPU)。

is Taiwan ready for
GPU-powered weather models?

Question from a Meteorologist (1/4)

Could you provide hints for using GPU supercomputing in developing super-deep DL models?

My Polite Answer

good start with the many GPUs already ☺, what problem do we want to start solving?

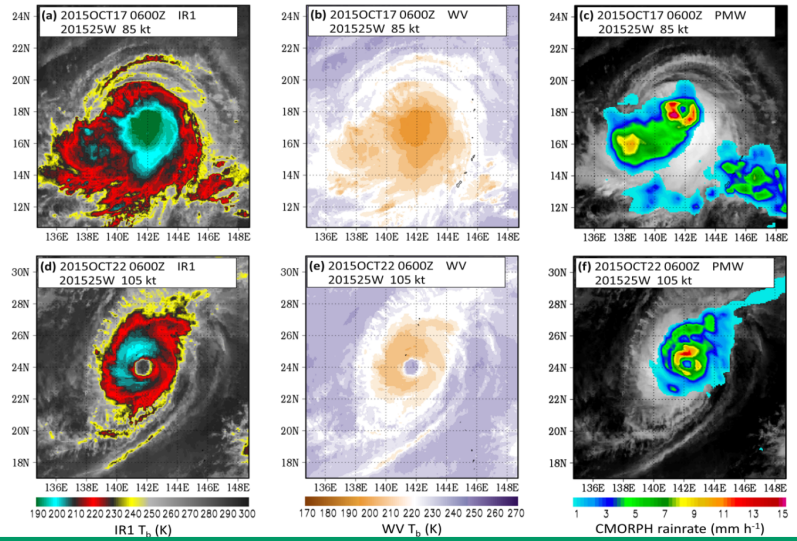
My Honest Answer

I don't know.

wisdom: super-deep models come from trying less deep ones on starting problem

Story 1: Tropical Cyclone Intensity Estimation

experienced meteorologists can
'feel' & estimate TC intensity from image



Can ML/DL do the same/better?

- lack of ML-ready datasets
- lack of model that properly utilizes domain knowledge

issues addressed in our pioneering works

(Chen et al., KDD '18; Chen et al., Weather & Forecasting '19)

How the Story Actually Began

Chen et al. '18

Boyo Chen, Buo-Fu Chen, and Hsuan-Tien Lin. Rotation-blended CNNs on a new open dataset for tropical cyclone image-to-intensity regression. In KDD, pages 90–99, August 2018.

Chen et al. '19

BuoFu Chen, Boyo Chen, Hsuan-Tien Lin, and Russell L. Elsberry. Estimating tropical cyclone intensity by satellite imagery utilizing convolutional neural networks. Weather and Forecasting, 34(2):447–465, April 2019.

- my M.S. student Boyo Chen:
  cannot find a thesis topic that interested him
- Dr. BuoFu Chen:
  the elder brother, a meteorologist that appreciates CS

wisdom: cross-domain collaboration
requires open-mindedness, trust and luck

Question from my Student

Chen: Can I work with my brother on the tropical cyclone intensity estimation problem towards my M.S. thesis?

My Polite Answer

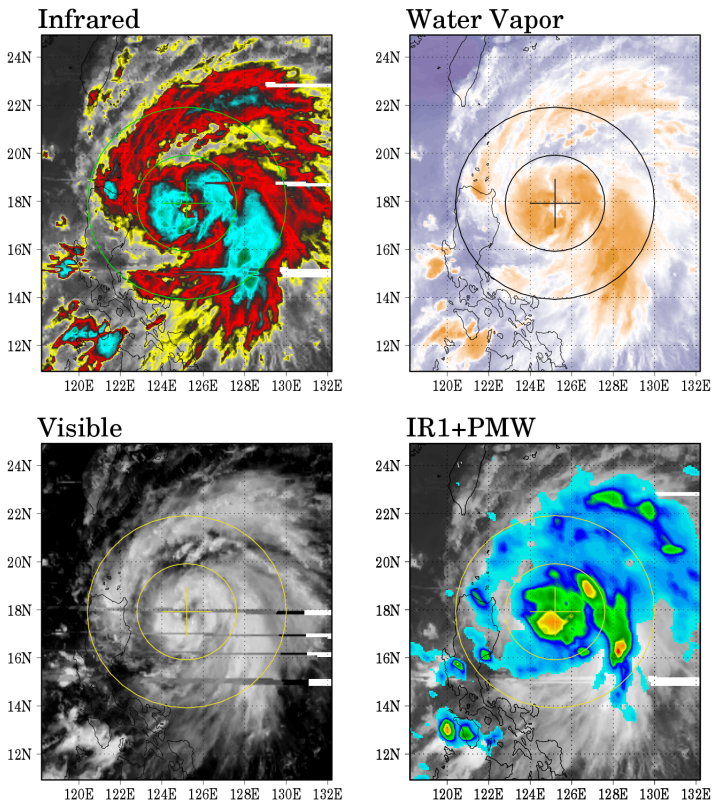
OK (if you do not have any better topics). Have you two discussed the goal?

My Honest Answer (2017)

Isn't this problem easily solvable by a mature Convolutional Neural Network?

observation: everyone thinks
every problem is easily solvable

Collecting Data is NOT Easily Solvable

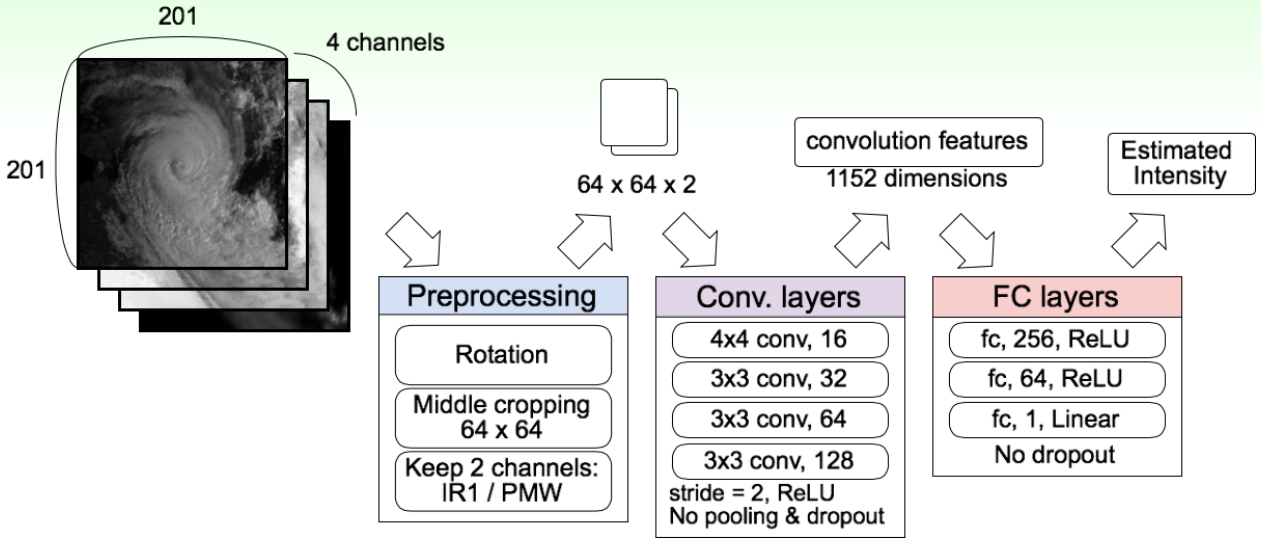


- base: GridSat,
  but we want TC data only!
- extension: CMORPH,
  but needs upscaling to unify with other data!
- label: best-track,
  but they are in different databases!

70501 frames of 1285 TCs, each 201*201*4 + some other features

data collection is tedious but impactful
—evidenced by our 105+71 citations

Designing Regression CNN is NOT Easily Solvable



- dropout under-estimates significantly
- max pooling is not suitable for contour-less target
- average pooling wipes out many important details

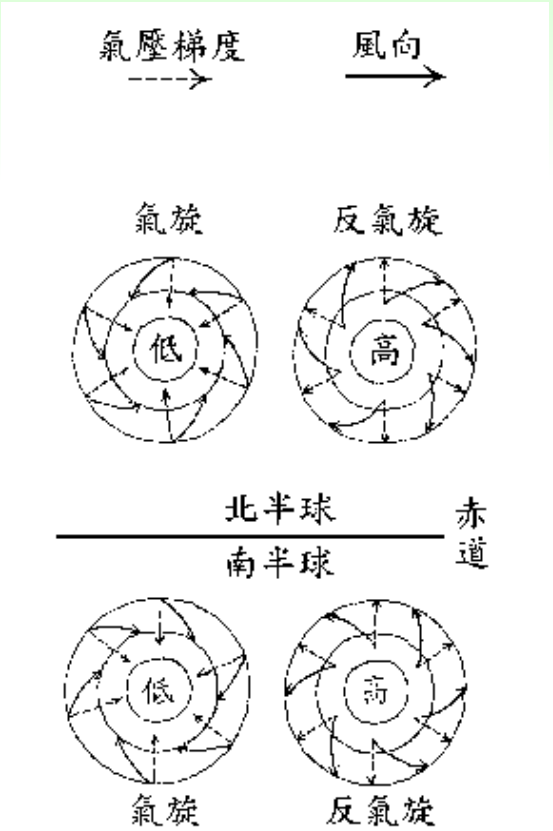
simplified AlexNet after trial-and-error
—some can inspire more general research

Professional Touch: Utilizing Domain Knowledge

Image Invariances

need: some regularization
(to replace dropout)

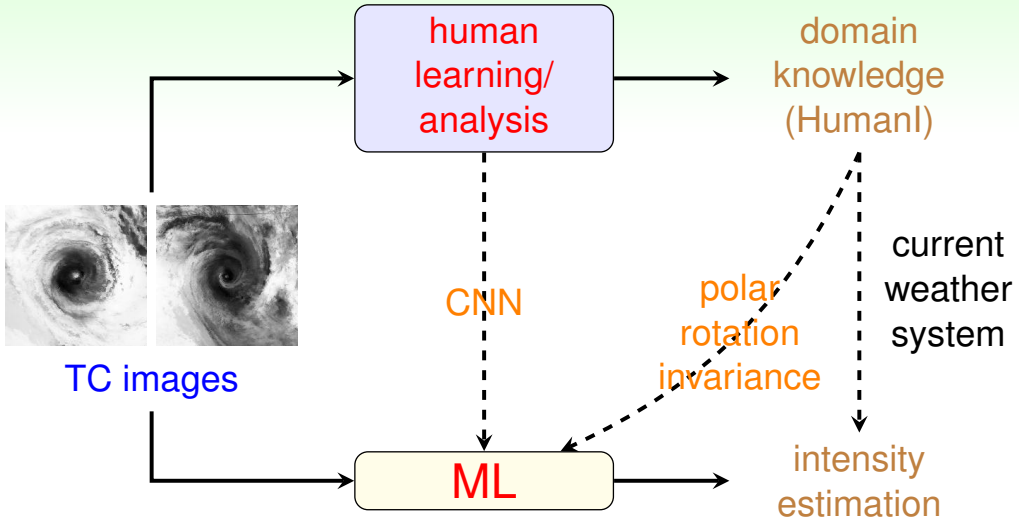
- flipping: possibly not
- shifting: no need (TCs are easily centered)
- rotating: yes, but ...



(downloaded from MOE for educational purpose)

CNN-TC: rotation-blended model with polar rotation invariance

ML for Modern Artificial Intelligence



Human Learning

- subjective
- produce domain knowledge
- fast basic solution

Machine Learning

- objective
- leverage computing power
- continuous improvement

wisdom: usually use humans as much as possible first before going to machines

Results

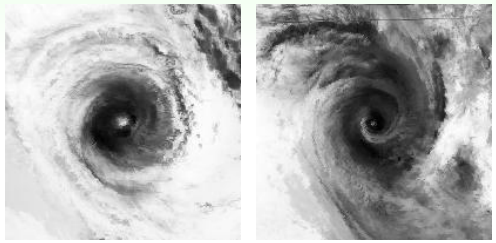
RMS Error

ADT ("automated human learning")	11.75
AMSU	14.40
SATCON (blending of ADT, AMSU, ...)	9.66
CNN-TC	9.03

CNN-TC much better than current weather system (SATCON)

are we using this cool model in production? ☺

Lessons Learned from Research on Tropical Cyclone Intensity Estimation



1. yes, “an upgraded version of CNN-TC is used in production, because it is stable and effectively fulfills what CWA needs.”
2. no, not easy to claim production ready
   —can ML be used for ‘unseenly-strong TC’?
3. cross-domain collaboration important
   e.g. even from ‘organizing data’ to be ML-ready
4. good AI system requires both human and machine learning
   —still an ‘art’ to blend the two
5. hard to do continuous research after production prototyping

Question from THE Meteorologist (2/4)

How can we follow up on rapidly developing DL models and select suitable ones for weather forecasting?

(or similarly) What is the best machine learning model for (my precious big) data and AI?

My Polite Answer

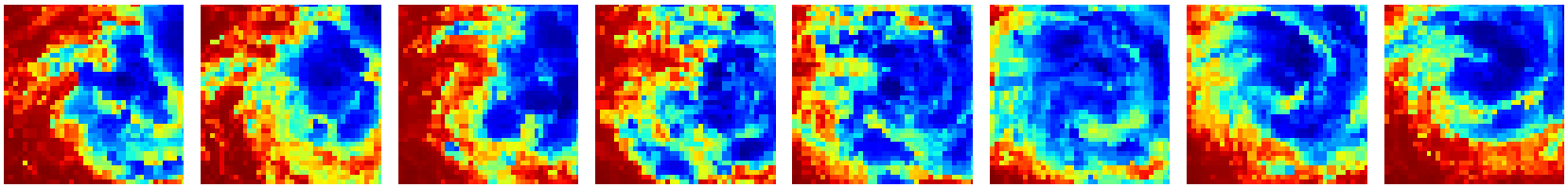
the best model is problem-dependent, let's chat about your problem first

My Honest Answer

I don't know.

wisdom: good models come from interactive modificationS to meet problem goal

Story 2: TC Rapid Intensification Identification



Rapid Intensification: TC intensity ↑ 25 knots within 24 hours
—a rare event (10% of intensity changes)

existing solutions

- feature (predictor) engineering
- linear models (LDA, logistic regression, etc.) on a few features

—can DL do better with raw data?

My Honest Thought (2019)

easily solved by a mature Recurrent Neural Network (RNN)?!

maybe not,
as explained with (Bai et al., ECML/PKDD '20)

TC Rapid Intensification (Selected Stories)

dataset ready? yes!

https://www.csie.ntu.edu.tw/~htlin/program/TCRISI/

Reference Score

Brier Score (MSE) = 0.300

Our Score

BS (w/o attention) = 0.271
BS (w/ attention) = 0.273

slightly better, but not as expected, why?

ah, BS not suited for imbalanced data. How about Heidke Skill Score?

but that's harder to optimize from ML perspective

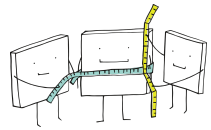
hmm, how about area under precision-recall curve (PR-AUC)?

truth: a paper was written,
but nobody cares about our criteria/results

Question from THE Meteorologist (3/4)

You have successfully cooperated with a team of meteorologists and forecasters. If you went back in time and did that again, what is the first thing to improve and make it further successful?

evaluation criteria!!



(free image by Manfred Steger from Pixabay)

suggest improvement opportunities

⇕

data

hint
preparation steps

⟸

techniques

assist
model/tech. choices

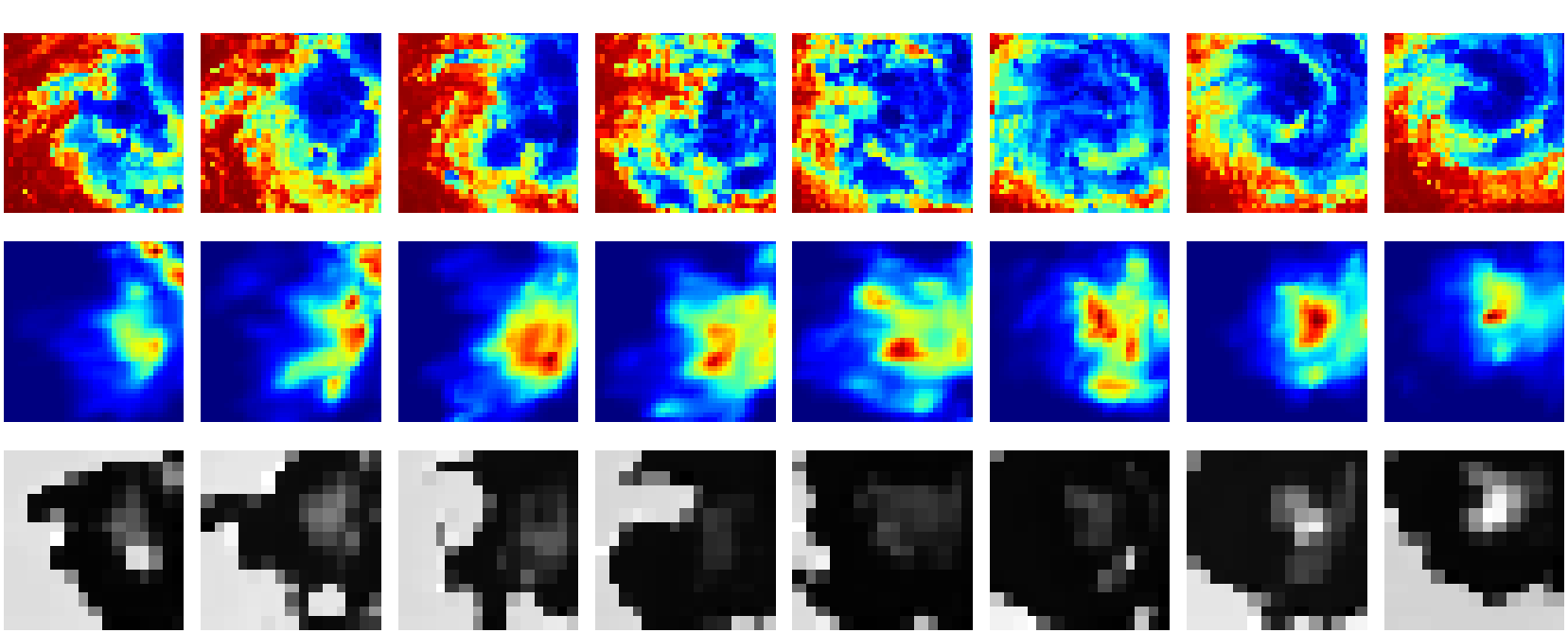
⟸

usage

define
acceptance goals

wisdom: agree on reasonable & measurable criteria to describe prioritized goal

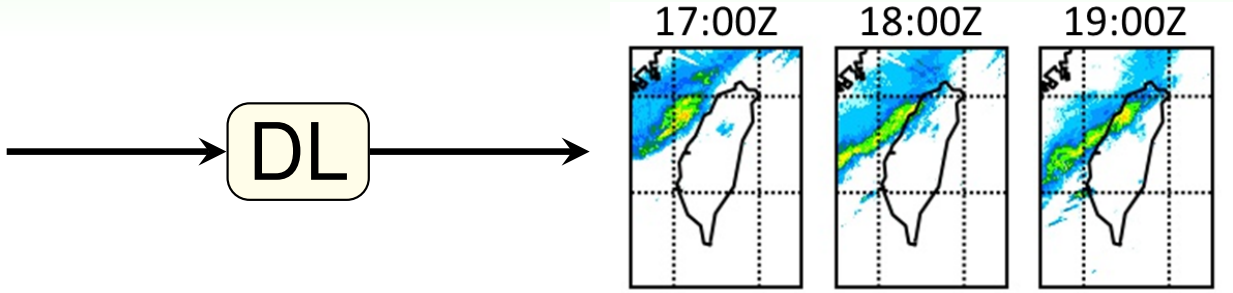
Lessons Learned from Research on TC Rapid Intensification



1. solutions often do not easily work out of the box
2. agreeing on a reasonable evaluation criteria is extremely important
3. easy to produce a paper that no one cares
4. hard to continuously persuade CS students

Story 3: Precipitation Nowcasting

6 (radar, rain) image pairs this hour



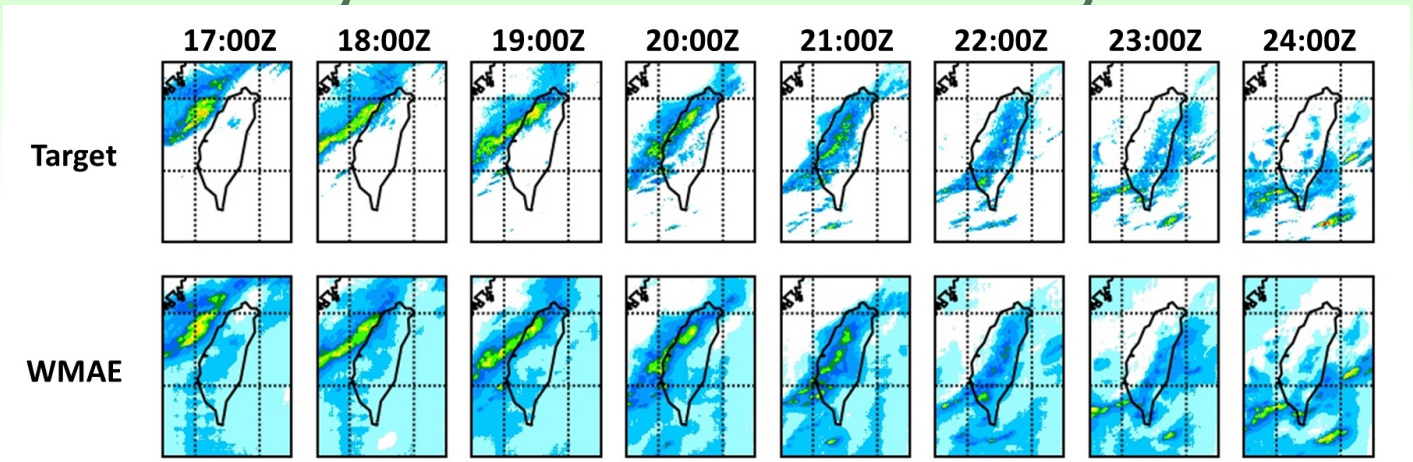
My Honest Thought (2020)

easily solved by a mature Recurrent Neural Network (RNN)?!
(yes? Shi et al., NeurIPS 2015, by ConvLSTM on Radar Data only)

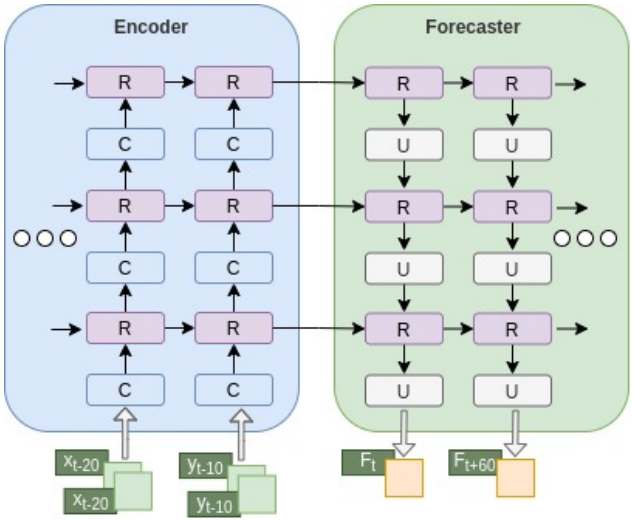
maybe not, as explained with

(Ashesh et al., AI for the Earth Systems '22)

First Try: Taiwan Rains Everywhere

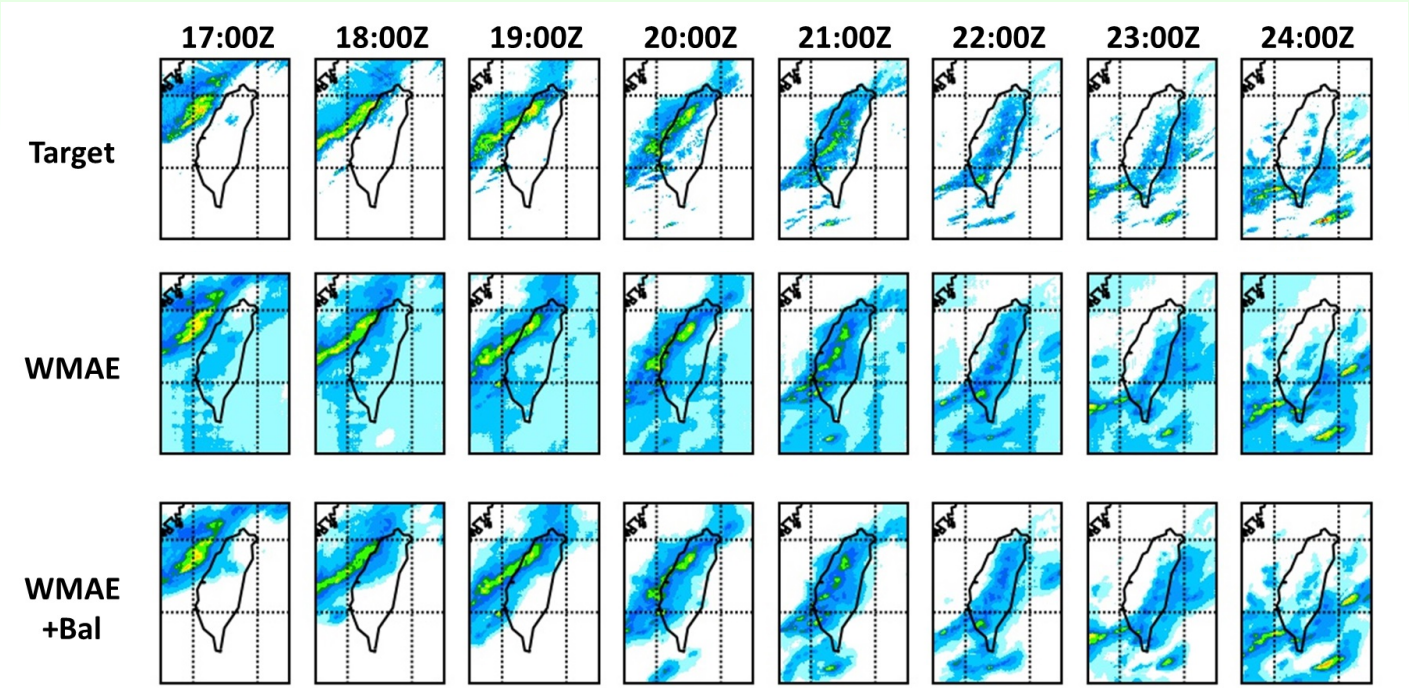


- encoder-forecaster design (ConvGRU), like previous studies
- weighted MAE metric, as the domain demands



- regression model feels "safer" to predict a bit of rain
- the metric (loss) ignores no/low-rain pixels

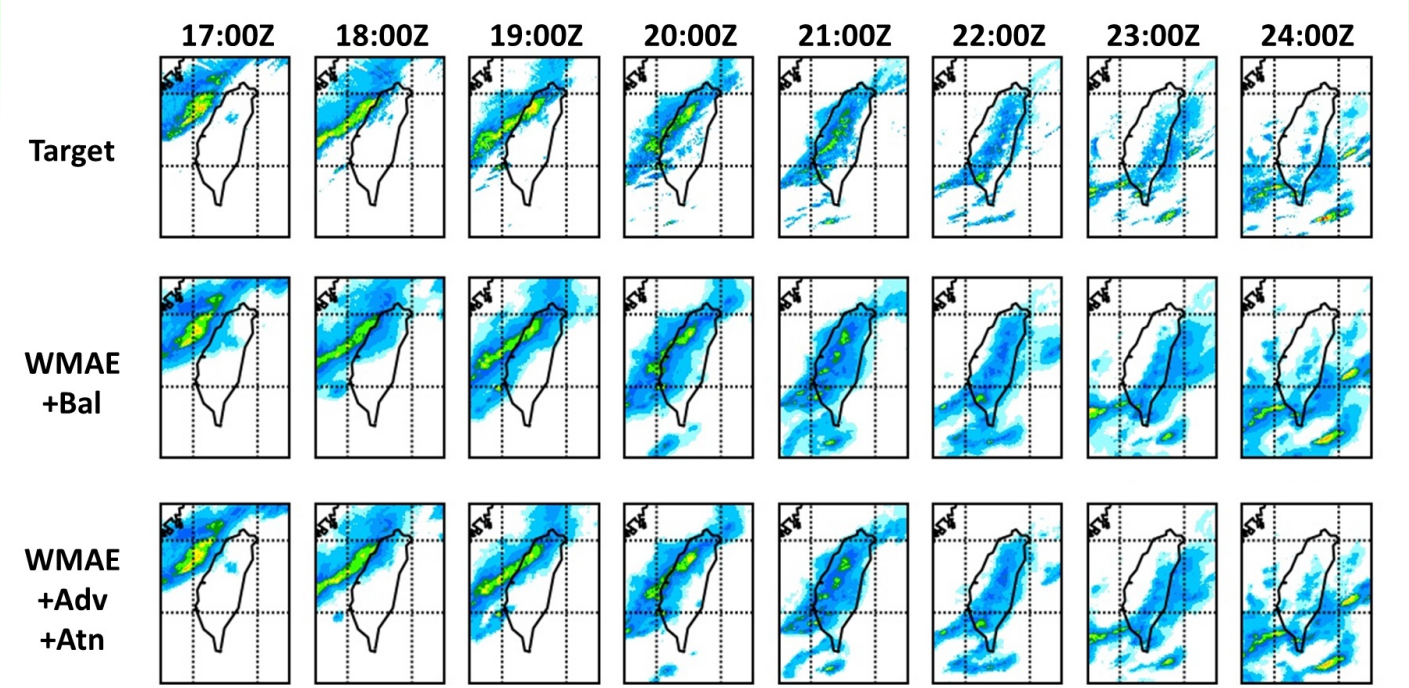
Second Try: Lacking “Details” Visually



- + another loss that focuses on low/no-rain regions

issue: pixel-by-pixel prediction
lacks structure (details)

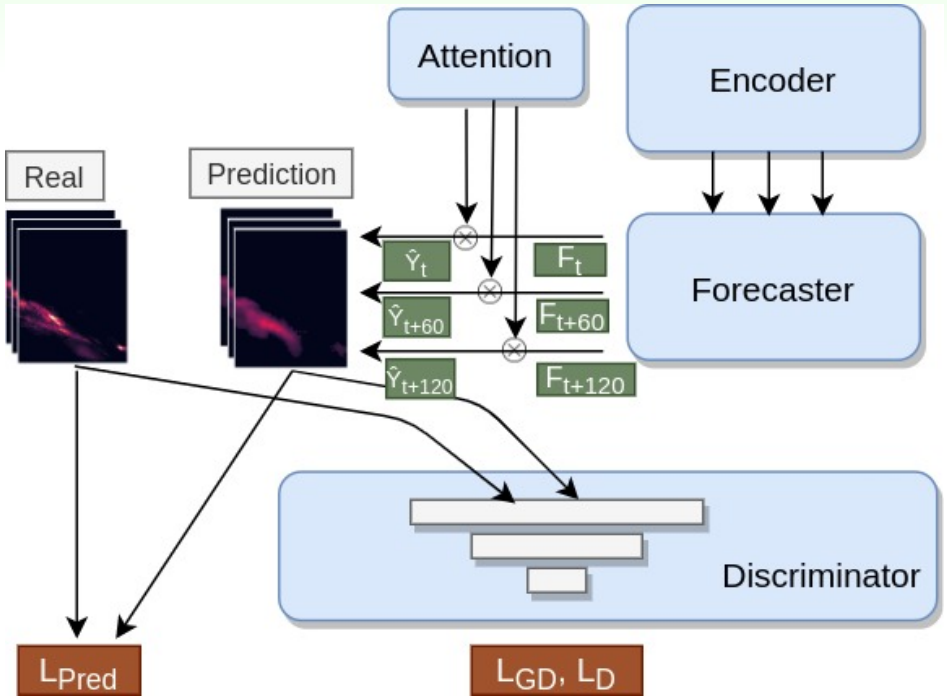
Third Try: Enforcing Structure by Discriminator



- + another loss that discriminates target image and generated image, like GAN design

issue: longer-term performance
—solved by adding sequence attention

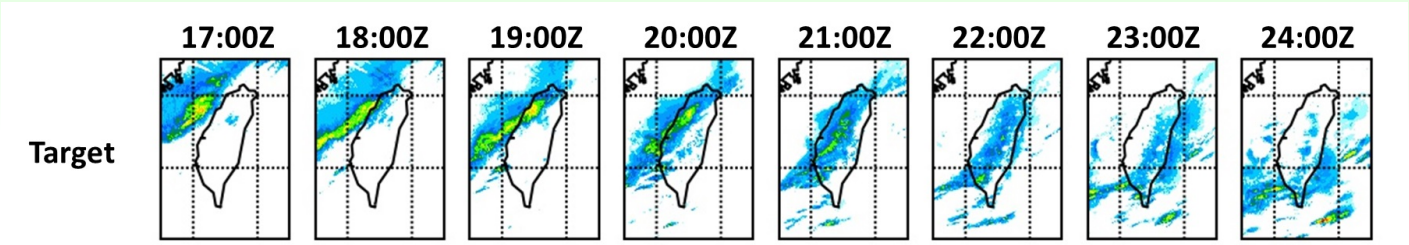
Accurate & Clear Precipitation Nowcasting



- rejected by KDD 2021 Applied Science Track
- rejected by WACV 2022: `the paper is more focused on the data and analysis more than the computer vision contribution, limiting its applicability to this community`

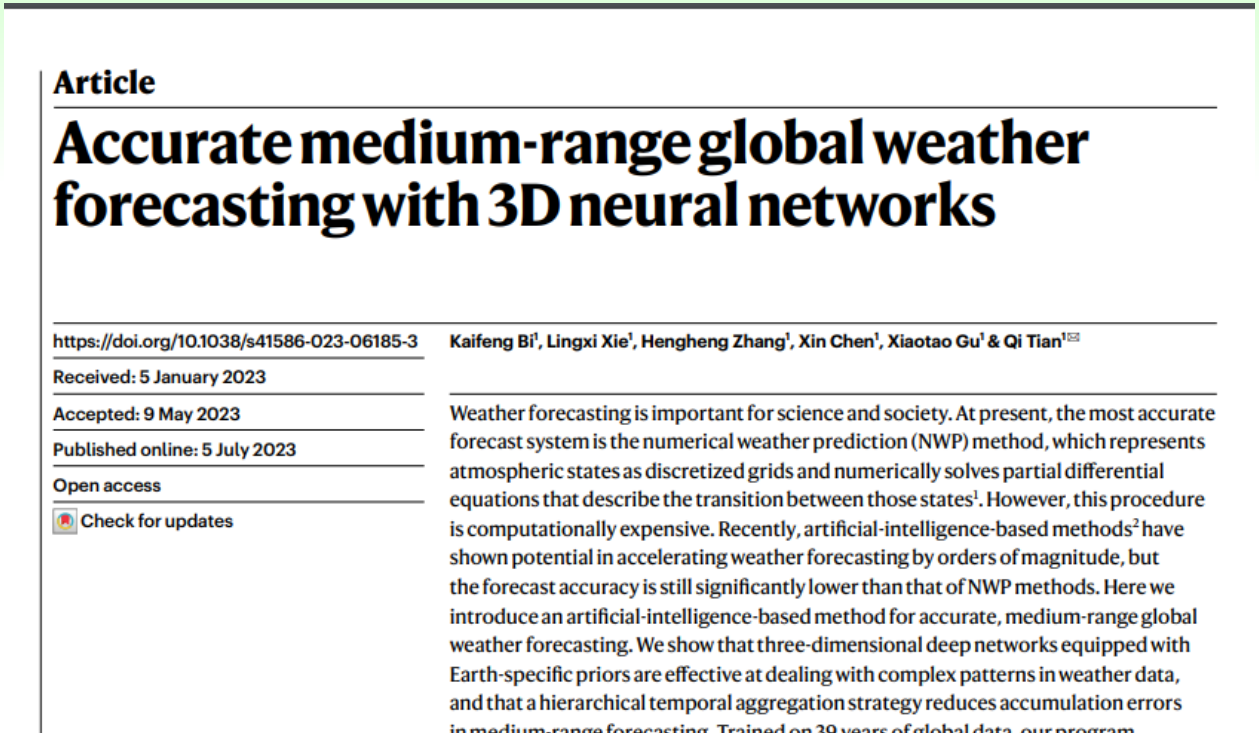
09/29/2021: DeepMind published "Skilful precipitation nowcasting using deep generative models of radar" in Nature

Lessons Learned from Precipitation Nowcasting



1. not hard to work on a Nature-level topic while
   - being rejected by CS conferences
   - falling behind tech. giants
2. many Nature-level topics in meteorology
3. operation ready? Yes!
   - not always easy to take global solution to local use
   - no one else will build local models for Taiwan
4. important to blending human experts with ML with interactive modifications

Speaking of Nature Topic



Article

Accurate medium-range global weather forecasting with 3D neural networks

https://doi.org/10.1038/s41586-023-06185-3

Received: 5 January 2023

Accepted: 9 May 2023

Published online: 5 July 2023

Open access

Check for updates

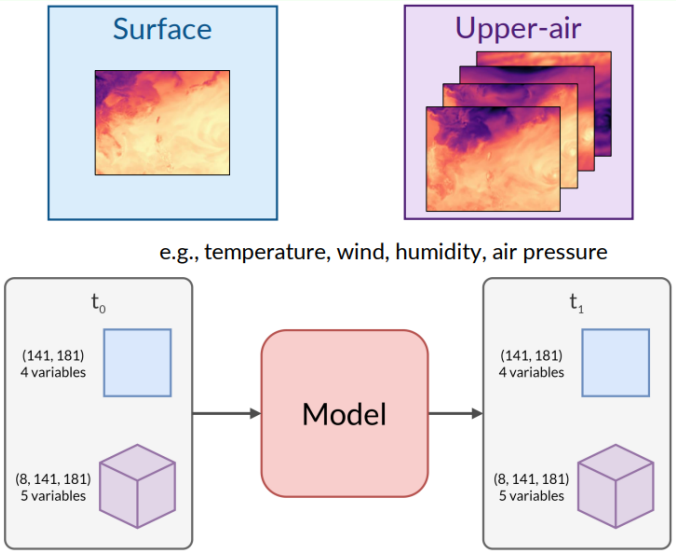
Kaifeng Bi[1], Lingxi Xie[1], Hengheng Zhang[1], Xin Chen[1], Xiaotao Gu[1] & Qi Tian[1]✉

Weather forecasting is important for science and society. At present, the most accurate forecast system is the numerical weather prediction (NWP) method, which represents atmospheric states as discretized grids and numerically solves partial differential equations that describe the transition between those states[1]. However, this procedure is computationally expensive. Recently, artificial-intelligence-based methods[2] have shown potential in accelerating weather forecasting by orders of magnitude, but the forecast accuracy is still significantly lower than that of NWP methods. Here we introduce an artificial-intelligence-based method for accurate, medium-range global weather forecasting. We show that three-dimensional deep networks equipped with Earth-specific priors are effective at dealing with complex patterns in weather data, and that a hierarchical temporal aggregation strategy reduces accumulation errors in medium-range forecasting. Trained on 39 years of global data, our program

(Snapshot from Nature for educational purposes)

07/05/2023: ...trained on 39 years of global data, our program, Pangu-Weather, obtains stronger deterministic forecast results on reanalysis data in all tested variables...

Story 4: Taiwan Weather PredictionS (Ongoing)



(courtesy of my student Huai-Yuan Kuo)

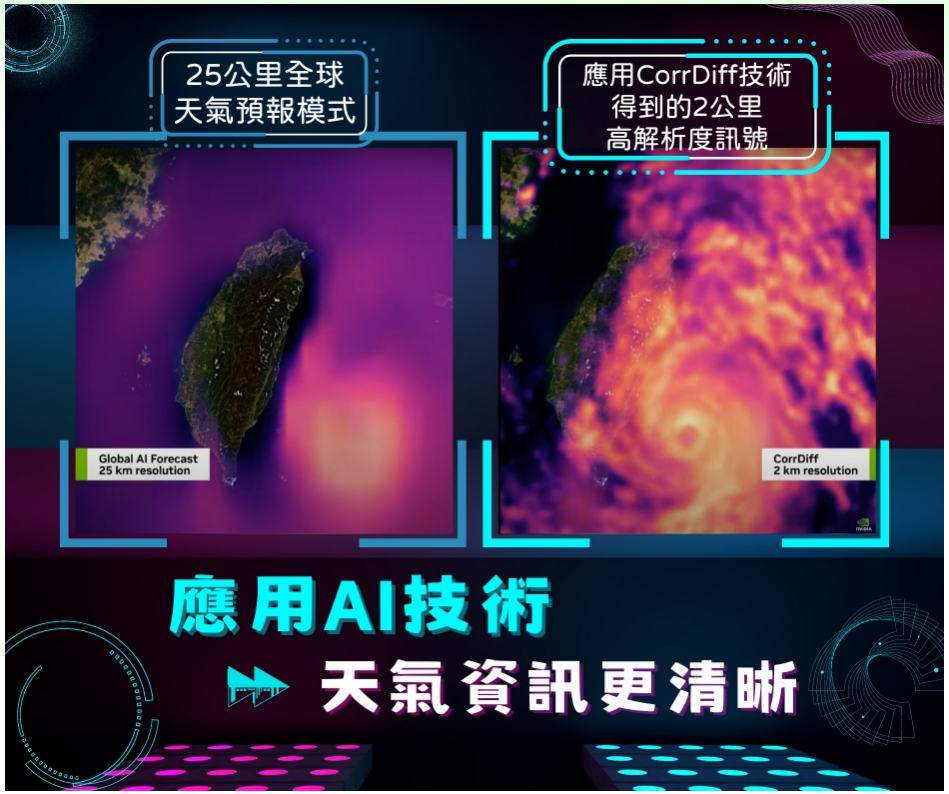
want: robust predictions of

- 5 layers of 8 upper-air measures (40 channels)
- 4 surface measures (4 more channels)
- high resolution of 2 to 5 km
- 72 hours of long-term prediction

and connecting with global models!

solving important society needs for Taiwan

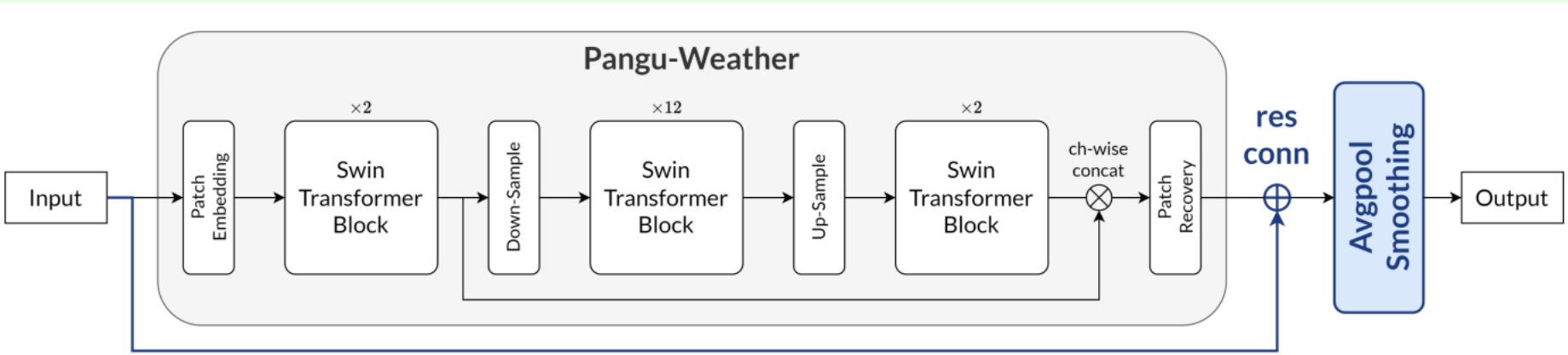
NVIDIA CorrDiff: “Our” Proof-of-concept Demo



(downloaded from CWA for educational purposes)

4 instead of 44 channels on 2018–2022
Taiwan data with diffusion model

Our Model after Interactive Discussions



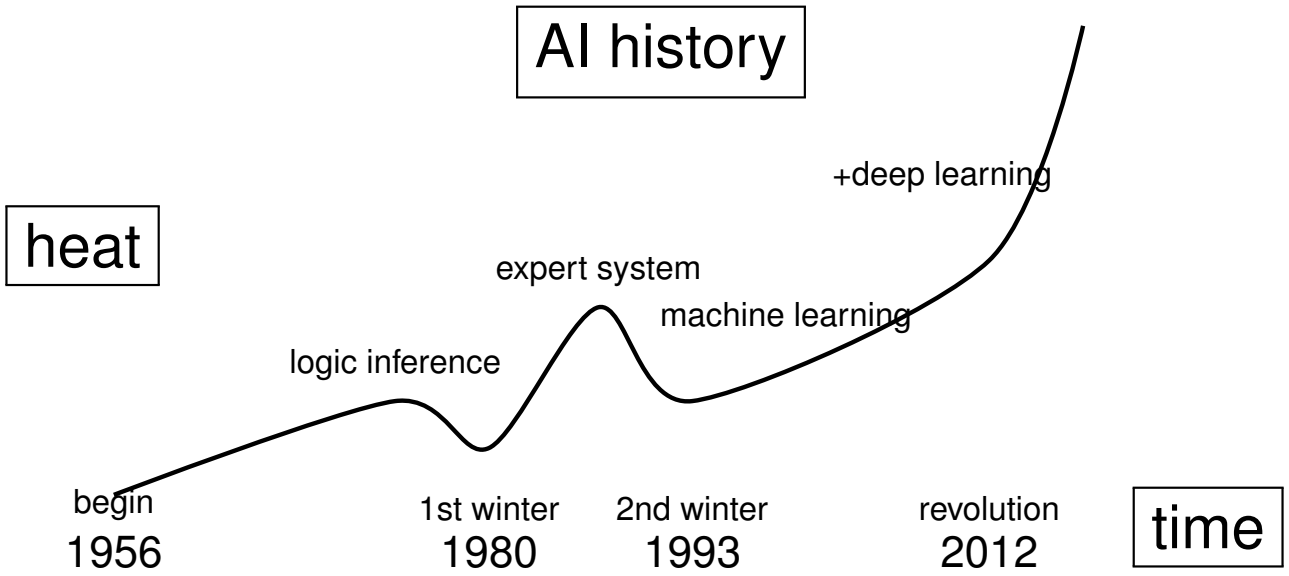
(courtesy of my student Huai-Yuan Kuo)

- across-patch average pooling:
  address mosaic outputs
- residual-in-time:
  achieve longer-term stability by following physics principles better

simple & effective techniques
can make big impacts

Question from THE Meteorologist (4/4)

Do you know some cases of failure in inter-discipline AI applications, and what can we learn from them?



expectation mismatch: the key sin for application intelligence

- first AI winter: AI cannot solve 'combinatorial explosion' problems
- second AI winter: expert system failed to scale

wisdom: open expression of expectations, and respect each other

Thought 1: Meteorology and CS/ML

Meteorology

- lots of data—arguably one of the oldest "data science" applications
- usually some measurable criteria with fast feedback
- some longstanding human knowledge as kickstart (or compare against)

—very ideal ML playground

But...

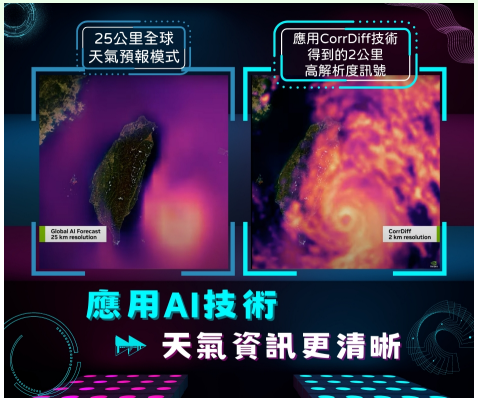
- specialized interests
- stereotype misunderstandings
  - meteorology: CS people must have ready solutions
  - CS: meteorology people only need us as IT developers
- different publication/evaluation systems

need more success stories
to encourage collaborations

Thought 2: Meteorology and Generative AI

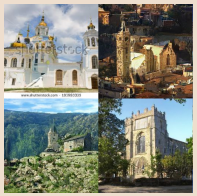
many meteorology needs are generative AI

- CorrDiff (Generative Correction Diffusion Model) is genAI
- our precipitation nowcasting model includes discriminator, an important design in genAI



Two Properties of Generative AI

variation (creativity)



(Extracted from Ho et al.

for educational purposes)

complexity (structure)

(Licensed under

CC0 on Wikipedia)

meteorology needs:
complex outputs with variations, i.e., genAI

Thought 3: Meteorology and ML Research

the million-dollar question: why is this a research problem?

ML

- deficiency in current model, needing model improvements
- hard to know model deficiency before running the first few models
- suggestion: try some baseline models openly, understand needs, and raise research questions from them

Meteorology

- deficiency in solution quality, needing running/analyzing more solutions
- hard to excite ML people if just running more models
- suggestion: describe your goal, understand how ML/DL experts model toward your goal, and then provide feedback

wisdom: communication/research protocol also important

Summary

- wisdom from some success and failure stories
  - leverage human knowledge properly
  - clear evaluation criteria
  - interactive feedback protocols towards improvement
  - open expression of expectations
- meteorology is an ideal playground for (gen-)ML/AI
  - mutual understanding/respect is important
  - more collaborations encouraged with more success stories
- no one will build Taiwan-prioritized models except us!

feel free to discuss more if you have problems/interests

Acknowledgments

- Dr. BuoFu Chen, Prof. Hung-Chi Kuo, and all my meteorology collaborators
- Boyo Chen, Ching-Yuan Bai, Ashesh, Huai-Yuan Kuo, and all my lab students
- my daughter Tiffany Lin, for cooking dinner yesterday with her genHI, allowing me to produce three pages of slides with my genHI, and all my family members

Thank you! Questions?