

Challenges of integrating classical numerical simulations and DL-based image processing

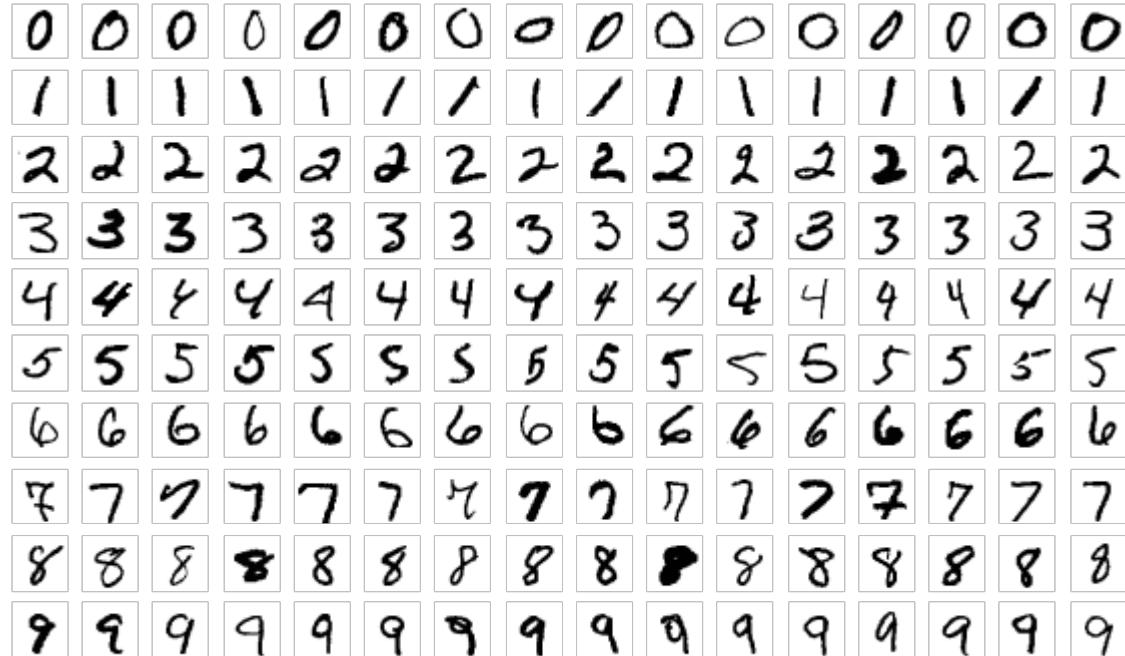
Heuna Kim

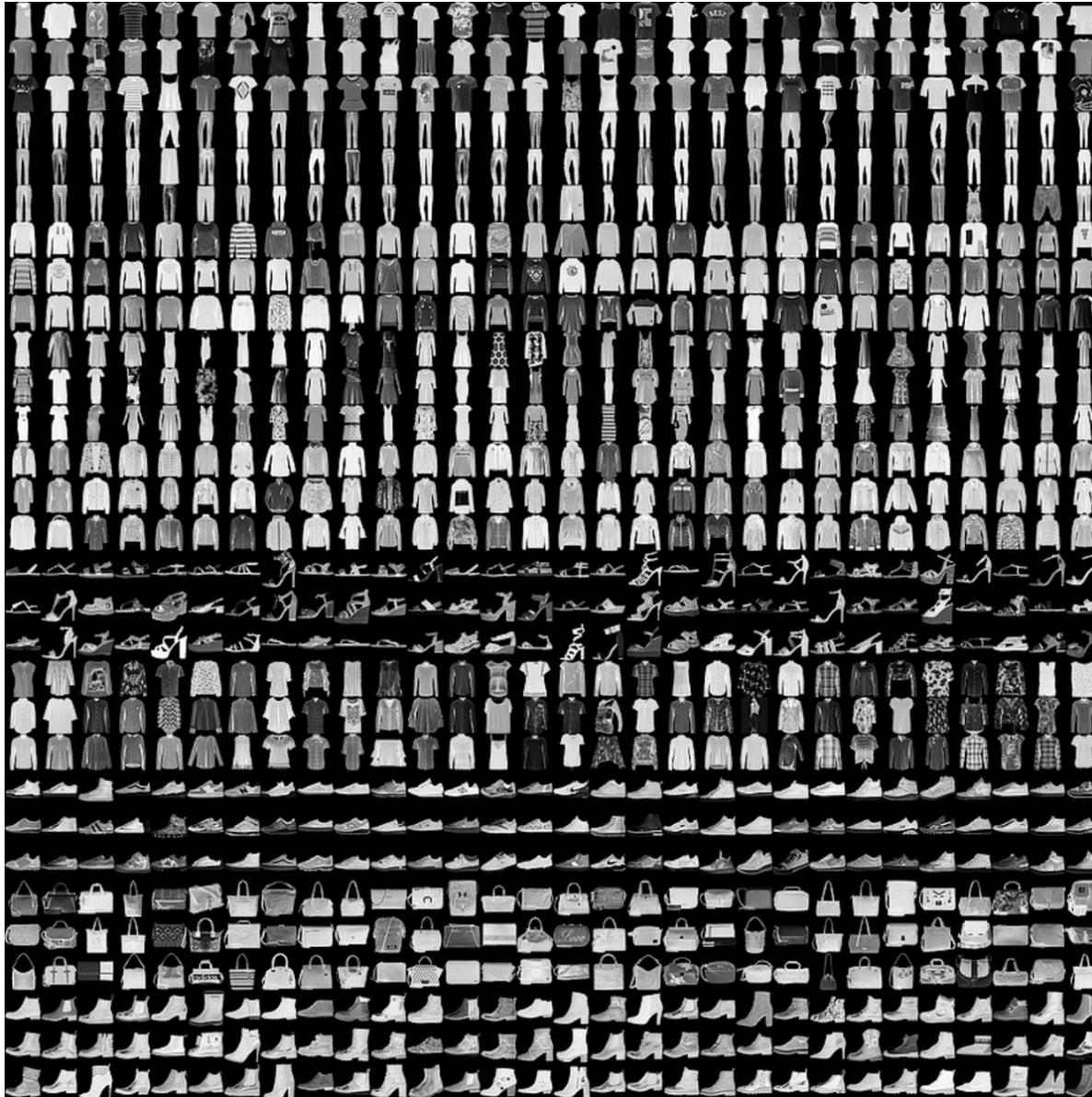
What's up for today?

- A kind archeology tour
- Examples in Meteorology
- Examples in Eye Tracking

Let's start with a little bit of
Archeology

MNIST





The most valid approach in 2021

Import TensorFlow into your program:

```
import tensorflow as tf

from tensorflow.keras.layers import Dense, Flatten, Conv2D
from tensorflow.keras import Model
```

Load and prepare the [MNIST dataset](#).

```
mnist = tf.keras.datasets.mnist
```

```
from torchvision import datasets
from torchvision.transforms import ToTensor

train_data = datasets.MNIST(
    root = 'data',
    train = True,
    transform = ToTensor(),
    download = True,
)

test_data = datasets.MNIST(
    root = 'data',
    train = False,
    transform = ToTensor()
)
```

THE MNIST DATABASE of handwritten digits

[Yann LeCun](#), Courant Institute, NYU

[Corinna Cortes](#), Google Labs, New York

[Christopher J.C. Burges](#), Microsoft Research, Redmond

Please refrain from accessing these files from automated scripts with high frequency. Make copies!

CLASSIFIER	PREPROCESSING	TEST ERROR RATE (%)	Reference
Linear Classifiers			
linear classifier (1-layer NN)	none	12.0	LeCun et al. 1998
linear classifier (1-layer NN)	deskewing	8.4	LeCun et al. 1998
pairwise linear classifier	deskewing	7.6	LeCun et al. 1998
K-Nearest Neighbors			
K-nearest-neighbors, Euclidean (L2)	none	5.0	LeCun et al. 1998
K-nearest-neighbors, Euclidean (L2)	none	3.09	Kenneth Wilder, U. Chicago
K-nearest-neighbors, L3	none	2.83	Kenneth Wilder, U. Chicago
K-nearest-neighbors, Euclidean (L2)	deskewing	2.4	LeCun et al. 1998
K-nearest-neighbors, Euclidean (L2)	deskewing, noise removal, blurring	1.80	Kenneth Wilder, U. Chicago
K-nearest-neighbors, L3	deskewing, noise removal, blurring	1.73	Kenneth Wilder, U. Chicago
K-nearest-neighbors, L3	deskewing, noise removal, blurring, 1 pixel shift	1.33	Kenneth Wilder, U. Chicago
K-nearest-neighbors, L3	deskewing, noise removal, blurring, 2 pixel shift	1.22	Kenneth Wilder, U. Chicago
K-NN with non-linear deformation (IDM)	shiftable edges	0.54	Keysers et al. IEEE PAMI 2007
K-NN with non-linear deformation (P2DHMDM)	shiftable edges	0.52	Keysers et al. IEEE PAMI 2007
K-NN, Tangent Distance	subsampling to 16x16 pixels	1.1	LeCun et al. 1998
K-NN, shape context matching	shape context feature extraction	0.63	Belongie et al. IEEE PAMI 2002
Boosted Stumps			
boosted stumps	none	7.7	Kegl et al., ICML 2009
products of boosted stumps (3 terms)	none	1.26	Kegl et al., ICML 2009
boosted trees (17 leaves)	none	1.53	Kegl et al., ICML 2009
stumps on Haar features	Haar features	1.02	Kegl et al., ICML 2009
product of stumps on Haar f.	Haar features	0.87	Kegl et al., ICML 2009
Non-Linear Classifiers			
40 PCA + quadratic classifier	none	3.3	LeCun et al. 1998
1000 RBF + linear classifier	none	3.6	LeCun et al. 1998
SVMs			

CLASSIFIER	PREPROCESSING	TEST ERROR RATE (%)	Reference
Linear Classifiers			
linear classifier (1-layer NN)	none	12.0	LeCun et al. 1998
linear classifier (1-layer NN)	deskewing	8.4	LeCun et al. 1998
pairwise linear classifier	deskewing	7.6	LeCun et al. 1998
K-Nearest Neighbors			
K-nearest-neighbors, Euclidean (L2)	none	5.0	LeCun et al. 1998
K-nearest-neighbors, Euclidean (L2)	none	3.09	Kenneth Wilder, U. Chicago
K-nearest-neighbors, L3	none	2.83	Kenneth Wilder, U. Chicago
K-nearest-neighbors, Euclidean (L2)	deskewing	2.4	LeCun et al. 1998
K-nearest-neighbors, Euclidean (L2)	deskewing, noise removal, blurring	1.80	Kenneth Wilder, U. Chicago
K-nearest-neighbors, L3	deskewing, noise removal, blurring	1.73	Kenneth Wilder, U. Chicago
K-nearest-neighbors, L3	deskewing, noise removal, blurring, 1 pixel shift	1.33	Kenneth Wilder, U. Chicago
K-nearest-neighbors, L3	deskewing, noise removal, blurring, 2 pixel shift	1.22	Kenneth Wilder, U. Chicago
K-NN with non-linear deformation (IDM)	shiftable edges	0.54	Keysers et al. IEEE PAMI 2007
K-NN with non-linear deformation (P2DHMDM)	shiftable edges	0.52	Keysers et al. IEEE PAMI 2007
K-NN, Tangent Distance	subsampling to 16x16 pixels	1.1	LeCun et al. 1998
K-NN, shape context matching	shape context feature extraction	0.63	Belongie et al. IEEE PAMI 2002
Boosted Stumps			
boosted stumps	none	7.7	Kegl et al., ICML 2009
products of boosted stumps (3 terms)	none	1.26	Kegl et al., ICML 2009
boosted trees (17 leaves)	none	1.53	Kegl et al., ICML 2009
stumps on Haar features	Haar features	1.02	Kegl et al., ICML 2009
product of stumps on Haar f.	Haar features	0.87	Kegl et al., ICML 2009
Non-Linear Classifiers			
40 PCA + quadratic classifier	none	3.3	LeCun et al. 1998
1000 RBF + linear classifier	none	3.6	LeCun et al. 1998
SVMs			

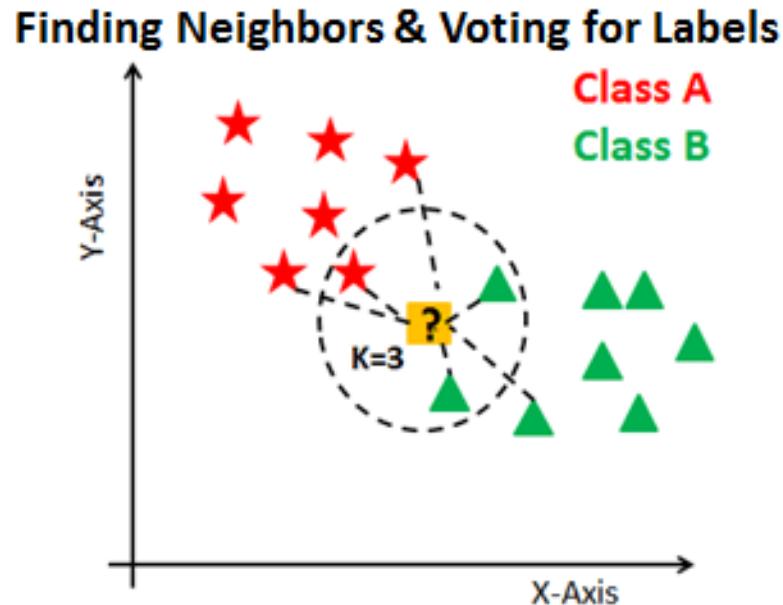
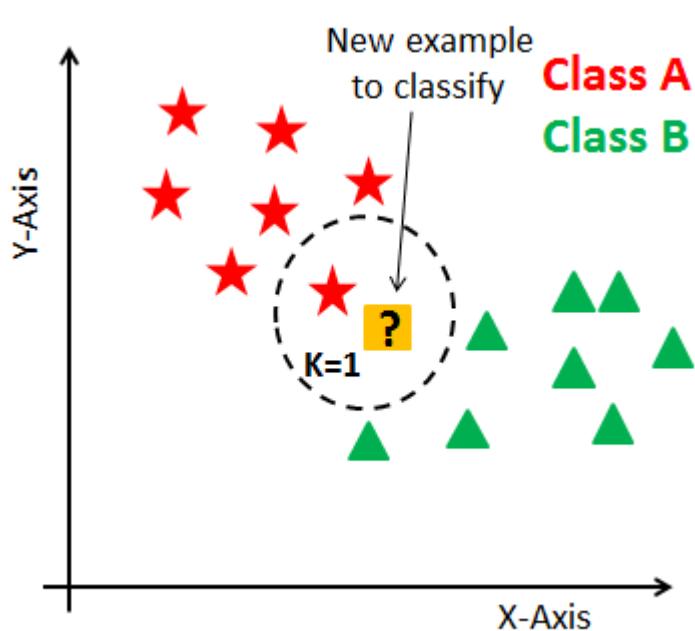
K-NN, Tangent Distance

- Was the MNIST problem “solved” in 1998?
- Is the solution in 1998 valid?

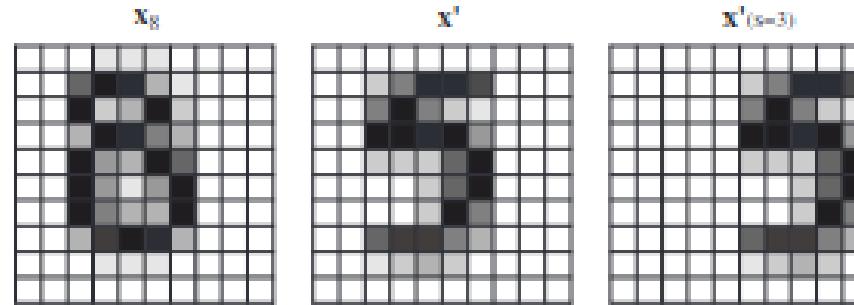
Invariance



Nearest-Neighbor Classification

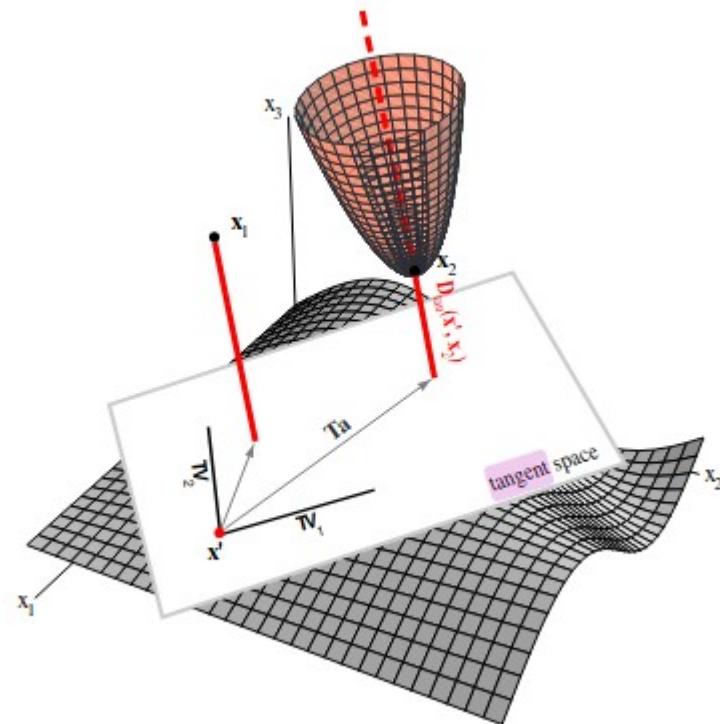


k-NN on MNIST



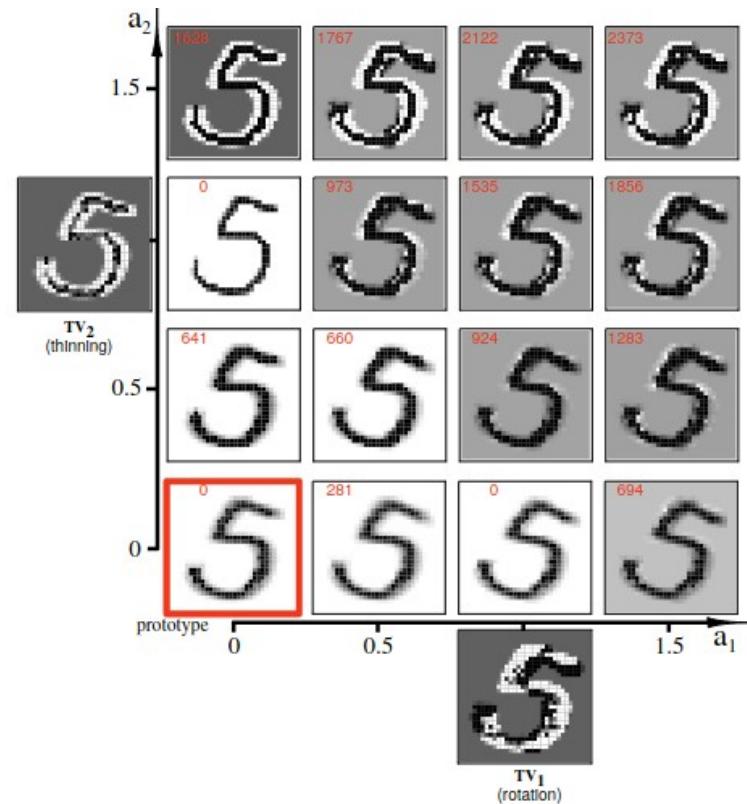
- What would be the problem of using Euclidean distance?

Tangent Distance

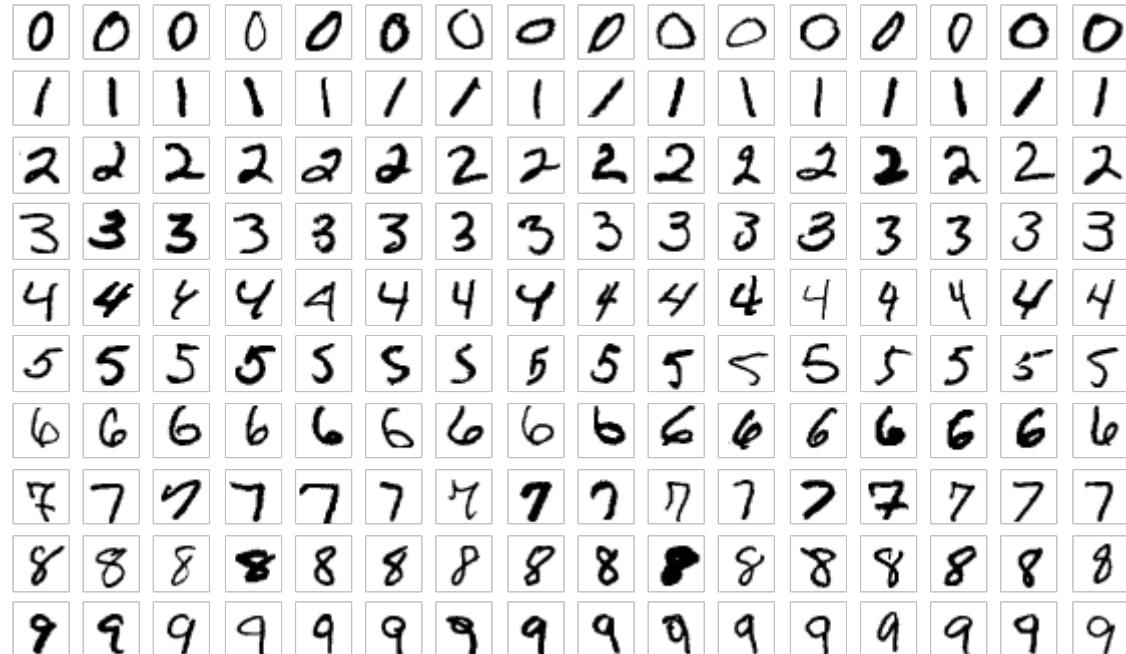


$$D_{tan}(\mathbf{x}', \mathbf{x}) = \min_{\mathbf{a}} [\|(\mathbf{x}' + \mathbf{T}\mathbf{a}) - \mathbf{x}\|]$$

Tangential Vector



K-NN with Tangential Distance in MNIST



K-NN, Tangent Distance

- Was the MNIST problem “solved” in 1998?
- Is the solution in 1998 valid?

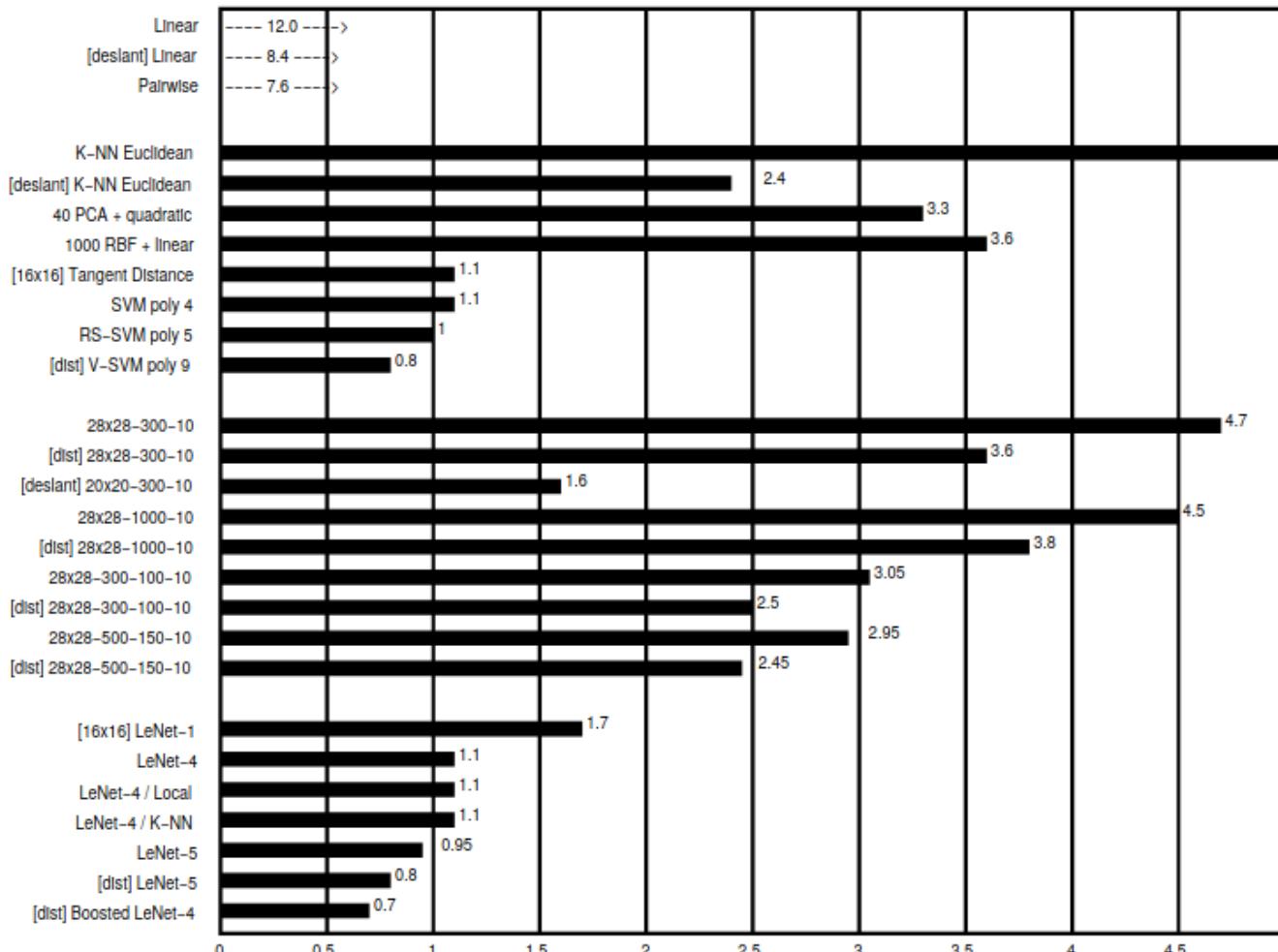


Fig. 9. Error rate on the test set (%) for various classification methods. [deslant] indicates that the classifier was trained and tested on the deslanted version of the database. [dist] indicates that the training set was augmented with artificially distorted examples. [16x16] indicates that the system used the 16x16 pixel images. The uncertainty in the quoted error rates is about 0.1%.

What does matter in the Algorithm?

- Accuracy
- Latency and Throughput
- Training time and inference time
- Memory requirements

What's up for today?

- A kind archeology tour
- **Examples in Meteorology**
- Examples in Eye Tracking

Meteorology



Deep Learning-Winter is coming...

Segmenting and Tracking Extreme Climate Events using Neural Networks

Mayur Mudigonda^{*1}, Sookyung Kim^{*2}, Ankur Mahesh^{*1}, Samira Kahou Karthik Kashinath⁴, Dean Williams², Vincent Michalski⁵, Travis O'Brien⁴ and Mi University of California Berkeley¹, Lawrence Livermore National Lab², Microsoft R Lawrence Berkeley National Lab⁴, Montreal Institute for Learning Algorithms mudigonda@berkeley.edu, kim79@llnl.gov, ankur.mahesh@berkeley.samira.ebrahimi.kahou@gmail.com, kkashinath@lbl.gov, williams13@ll michalskivincent@gmail.com, taobrien@lbl.gov, prabhat@lbl.gc

Abstract

Predicting extreme weather events in a warming world is one of the most pressing and challenging problems that humanity faces today. Deep learning and advance

Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model

Xingjian Shi, Zhihan Gao, Leonard Lausen, Hao Wang, Dit-Yan Yeung
Department of Computer Science and Engineering
Hong Kong University of Science and Technology
{xshiaab,zgaoag,lelausen,hwangaz,dyyeung}@cse.ust.hk

Wai-kin Wong, Wang-chun Woo
Hong Kong Observatory
Hong Kong, China
{wkwong,wcwoo}@hko.gov.hk

Abstract

With the goal of making high-resolution forecasts of regional rainfall, precipitation nowcasting has become an important and fundamental technology underlying various public services ranging from rainstorm warnings to flight safety. Recently,

Deep Learning-Winter is coming...



The latest news from Google AI

A Neural Weather Model for Eight-Hour Precipitation Forecasting

Wednesday, March 25, 2020

Posted by Nal Kalchbrenner and Casper Sønderby, Research Scientists, Google Research, Amsterdam

Predicting weather from minutes to weeks ahead with high accuracy is a fundamental scientific challenge that can have a wide ranging impact on many aspects of society. Current forecasts employed by many meteorological agencies are based on physical models of the atmosphere that, despite improving substantially over the preceding decades, are inherently constrained by their computational requirements and are sensitive to approximations of the physical laws that govern them. An alternative approach to weather prediction that is able to overcome some of these constraints uses *deep neural networks* (DNNs): instead of encoding explicit physical laws, DNNs discover patterns in the data and learn complex transformations from inputs to the desired outputs using parallel computation on powerful specialized hardware such as GPUs and *TPUs*.

Building on our previous research into precipitation nowcasting, we present “[MetNet: A Neural Weather Model for Precipitation Forecasting](#),” a DNN capable of predicting future precipitation at 1 km resolution over 2 minute intervals at timescales up to 8 hours into the future. MetNet outperforms the current state-of-the-art physics-based model in use by [NOAA](#) for prediction times up to 7-8 hours ahead and makes a prediction over the entire US in a matter of seconds as opposed to an hour. The inputs to the network are sourced automatically from radar stations and satellite networks without the need for human annotation. The model output is a probability distribution that we use to infer the most likely precipitation rates with associated uncertainties at each geographical region. The figure below provides an example of the network’s predictions over the continental United States.



Using Machine Learning to “Nowcast” Precipitation in High Resolution

Monday, January 13, 2020

Posted by Jason Hickey, Senior Software Engineer, Google Research

The weather can affect a person’s daily routine in both mundane and serious ways, and the precision of forecasting can strongly influence how they deal with it. Weather predictions can inform people about whether they should take a different route to work, if they should reschedule the picnic planned for the weekend, or even if they need to evacuate their homes due to an approaching storm. But making accurate weather predictions can be particularly challenging for localized storms or events that evolve on hourly timescales, such as thunderstorms.

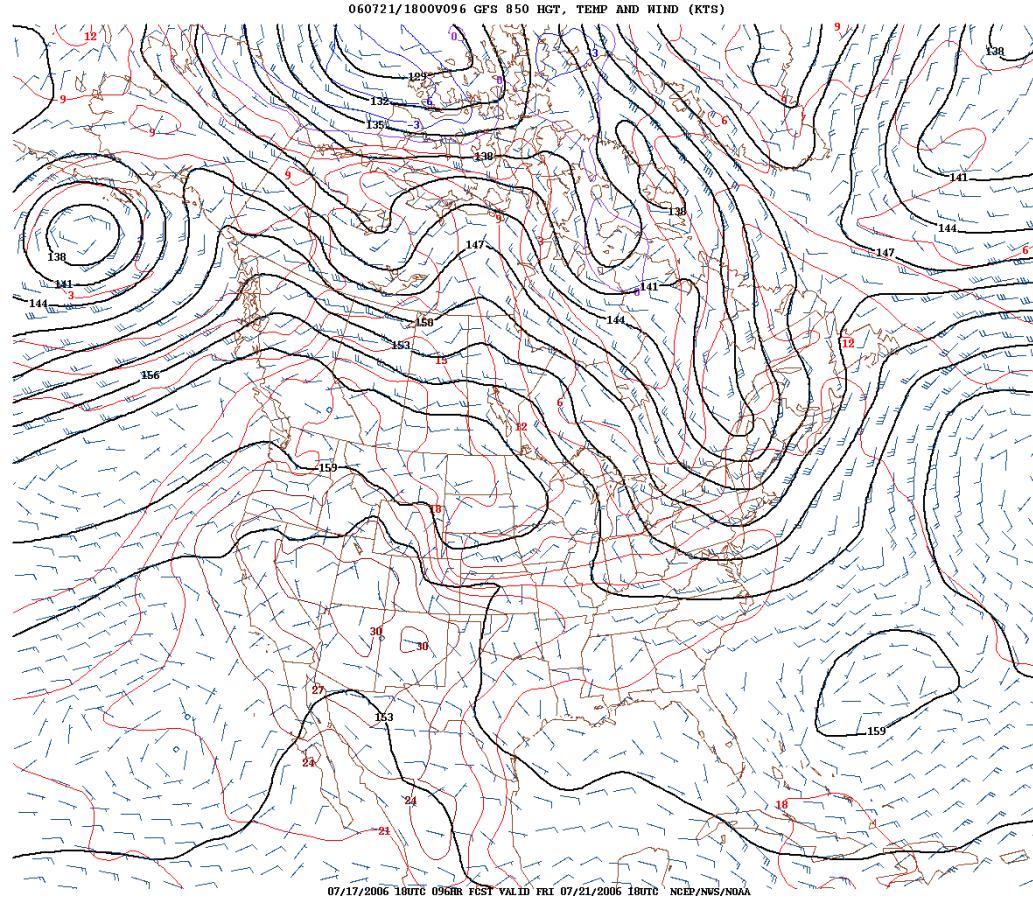
In “[Machine Learning for Precipitation Nowcasting from Radar Images](#),” we are presenting new research into the development of machine learning models for precipitation forecasting that addresses this challenge by making highly localized “physics-free” predictions that apply to the immediate future. A significant advantage of machine learning is that inference is computationally cheap given an already-trained model, allowing forecasts that are nearly instantaneous and in the native high resolution of the input data. This *precipitation nowcasting*, which focuses on 0-6 hour forecasts, can generate forecasts that have a 1km resolution with a total latency of just 5-10 minutes, including data collection delays, outperforming traditional models, even at these early stages of development.

Moving Beyond Traditional Weather Forecasting

Weather agencies around the world have extensive monitoring facilities. For example, Doppler radar measures precipitation in real-time, weather satellites provide multispectral imaging, ground stations measure wind and precipitation directly, etc. The figure below, which compares false-color composite radar imaging of precipitation over the continental US to cloud cover imaged by geosynchronous satellites, illustrates the need for multi-source weather information. The existence of rain is related to, but not perfectly correlated with, the existence of clouds, so inferring precipitation from satellite images alone is challenging.



Numerical Weather Forecasting



Weather Agency



Formerly the National Climatic Data Center (NCDC)... [more about NCEI »](#)

Home Climate Information Data Access Customer Support Contact About

Search

Home > Data Access > Model > Numerical Weather Prediction

Quick Links

Land-Based Station

Satellite

Radar

Model

Datasets

Climate Model Data in the
NOAA Big Data Project

Ocean Models

Reanalysis

Numerical Weather
Prediction

Climate Prediction

Derived / Other Model
Data

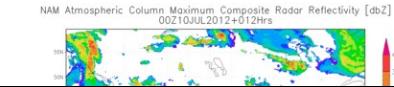
NOMADS

Numerical Weather Prediction

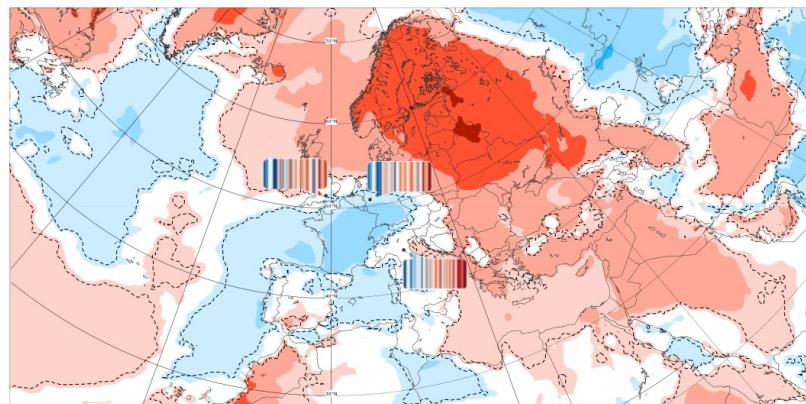
Numerical Weather Prediction (NWP) data are the form of weather model data we are most familiar with on a day-to-day basis. NWP focuses on taking current observations of weather and producing computer models of weather. Known weather is just as computer models. Current weather can be used to run the numerical process known as to produce outputs of precipitation, and meteorological elements at the top of the atmosphere.



Home About Forecasts Computing R



Advancing global NWP through international collaboration



2m Temperature: Weekly mean anomalies & warming stripes for the 3 ECMWF locations (Reading, Bonn, Bologna)

Base Time: Thu 08 Jul 2011 00 UTC T+264
Valid time: Mon 12 Jul 2011 00 UTC - Mon 19 Jul 2011 00 UTC

The anomalies have been calculated relative to a 20-year model climatology. The areas where the ensemble forecast is not significantly different from the climatology are blanked.

A "warming stripes" are a visual representations of the change in temperature. #ShowYourStripes

Plot your own here

[View all charts >](#)



정보공개

참여와소통

지식과 배움

행정과 정책

기상청소개

날씨누리 바로가기

청/차장 소개 미션·비전 조직·직원 주요업무 소속·산하기관소개 홍보실 찾아오는길 관련사이트

주요업무

관측업무

정보통신업무

예보업무

· 업무개요

· 기상예보

홈 > 기상청소개 > 주요업무 > 예보업무 > 수치예보

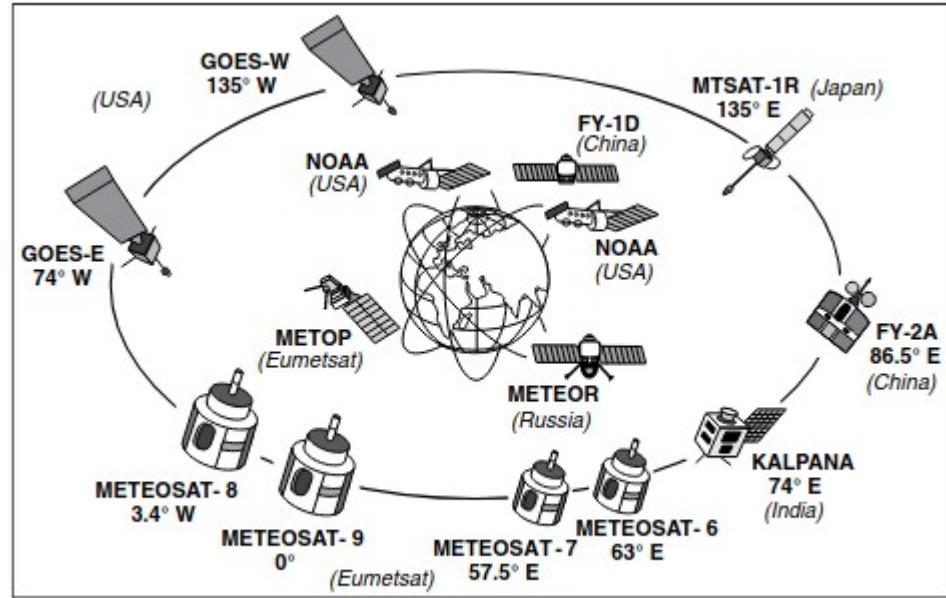
예보업무 | 수치예보

인쇄

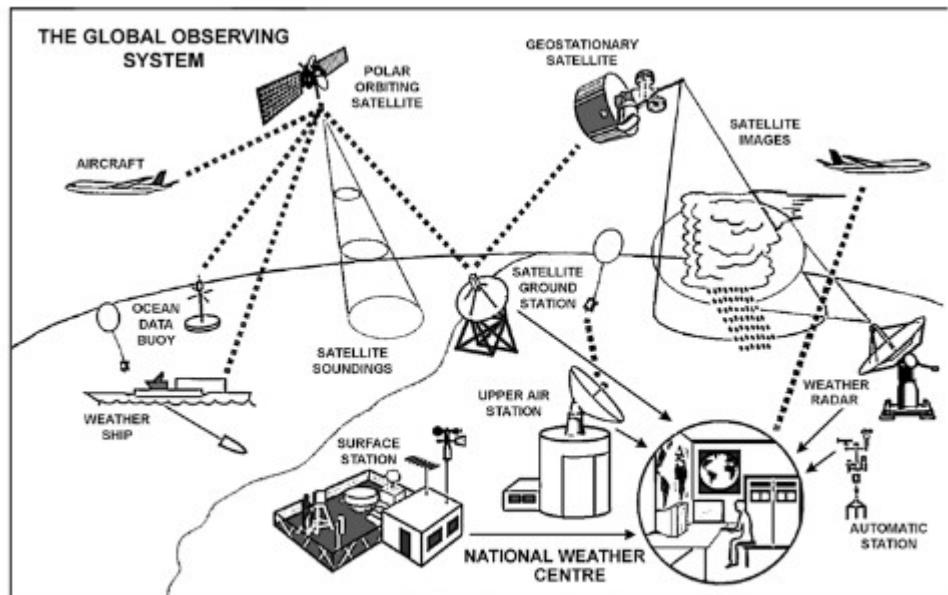
대기의 운동과 변화를 설명하는 역학 및 물리방정식을 슈퍼컴퓨터로 풀어 미래의 대기상태를 예측하는 일련의 과정을 수치예보라 하고 수치예보를 위해 슈퍼컴퓨터 내에서 가동되는 컴퓨터 프로그램들을 수치예보모델이라 한다.

수치예보의 원리를 간단히 설명하면, 사이버 공간 속에서 대기를 양파껍질처럼 여러 층으로 나누고, 각 층을 다시 바둑판처럼 여러 개의 작은 면으로 분할하여 대기 를 수만 개의 사각형으로 쪼개고 있다. 이렇게 쪼개고 나니 대기의 이동과 기압과 같은 기본적인 특성은 그대로 계산된다. 이를 통해 대기의 운동과 기온 등 기후학적 특성을 예측할 수 있다.

Observation



The operational system of meteorological satellites in 2007. (From EUMETSAT)



The various components of the Global Observing System. (After a WMO image)

Data Assimilation

3D-Var [edit]

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + (\mathbf{y} - H[\mathbf{x}])^T \mathbf{R}^{-1} (\mathbf{y} - H[\mathbf{x}]),$$

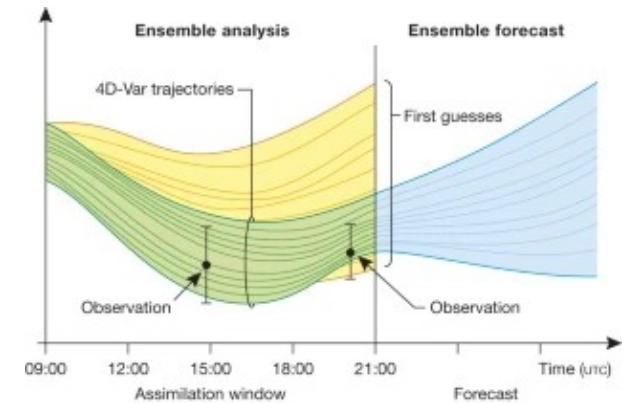
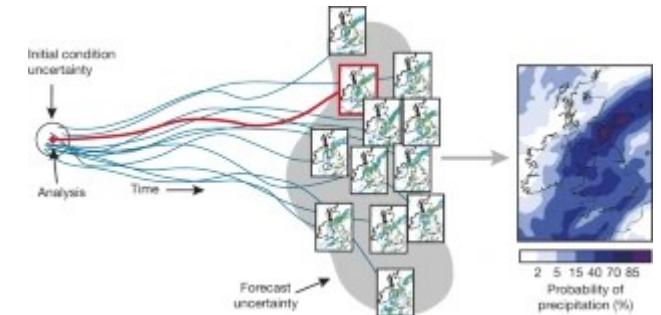
where \mathbf{B} denotes the background error covariance, \mathbf{R} the observational error covariance.

$$\nabla J(\mathbf{x}) = 2\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) - 2H^T \mathbf{R}^{-1}(\mathbf{y} - H[\mathbf{x}])$$

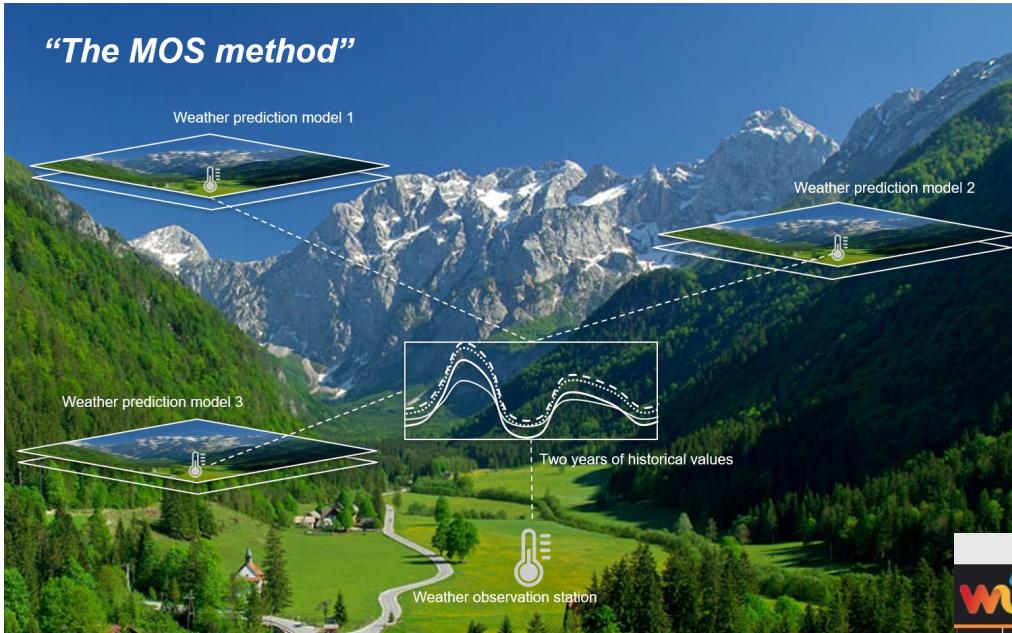
4D-Var [edit]

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \sum_{i=0}^n (\mathbf{y}_i - H_i[\mathbf{x}_i])^T \mathbf{R}_i^{-1} (\mathbf{y}_i - H_i[\mathbf{x}_i])$$

provided that H is a linear operator (matrix).



MOS (Model Output Statistics)



AccuWeather Berlin, Berlin 17°c

NOW HOURLY DAILY RADAR MINUTECAST MONTHLY AIR QUALITY

CURRENT WEATHER
8:08 AM

RealFeel Shade™ 17°
Air Quality Excellent
Wind NNE 10 km/h
Wind Gusts 18 km/h

Light rain

MINUTECAST®
Periods of rain for at least 60 min

MORE DETAILS →

WEATHER UNDERGROUND | Sensor Network Maps & Radar Severe Weather News & Blogs Mobile Apps More ▾

Popular Cities 59 °F San Francisco, CA 73 °F Manhattan, NY 60 °F Schiller Park, IL (60176) 64 °F Boston, MA 77 °F

San Francisco, CA

57 °F Feels like 55 °
--° | 52°
5% /
0.00 in

12AM 57° 54° 53° 58° 63° 65° 62° 57° -/- Jul 9
6AM Jul 10
NOON +0200
6PM
12AM



Can Deep Learning Replace Numerical Weather Prediction?

By Oliver Peckham

March 3, 2021

Numerical weather prediction (NWP) is a mainstay of supercomputing. Some of the first applications of the first supercomputers dealt with climate modeling and even to this day, the largest climate models are heavily constrained by the scale of the supercomputers that run them. While some wait for the exascale era – and beyond – to brute force punishingly accurate and complex climate models into existence, others are looking for a deep learning-powered shortcut to the same results. In a [paper](#) for *Philosophical Transactions of the Royal Society*, eight researchers from the Jülich Supercomputing Center explored whether deep learning could ever actually beat numerical weather prediction at its own game – and if so, how and when that might happen.

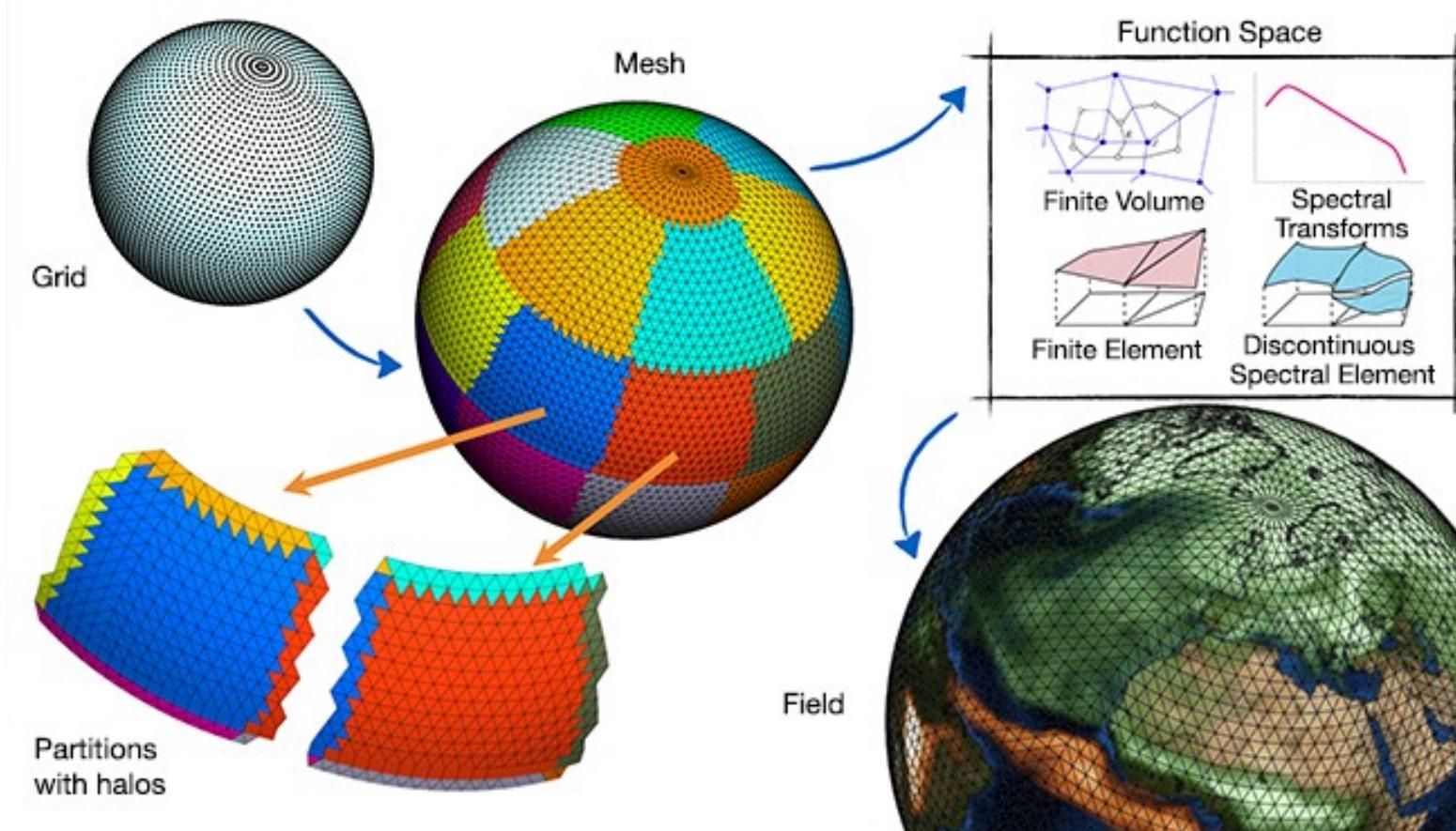
Status quo

The weather and climate supercomputing community is no stranger to deep learning, but it has hitherto mostly been used to augment NWP approaches (e.g. in resolving post-processing issues). These modelers, however, are reticent to incorporate deep learning in more meaningful capacities.

"[There] are still reservations about DL in this community," the authors write. "Two core arguments in this regard are the lack of explainability of deep [neural networks] and the lack of physical constraints. Furthermore, some scepticism prevails due to the fact that researchers have experimented with rather simple [neural networks] which were clearly unsuited to capture the complexity of meteorological data and feedback processes, but then extrapolate these results to discredit any [neural network] application including the much more powerful [deep learning] systems."

Deep Learning ~~Winter is~~ coming...

Spatial Grid and Scalability



Can Deep Learning Replace Numerical Weather Forecasting?



What does matter in the Algorithm?

- Accuracy: **BENCHMARK DATASET**
- Latency and Throughput **Parameter/Model Optimization**
- Training time and inference time **GLOBAL SPATIAL GRID**
- Memory requirements **Model Drift**

Numerical Weather Forecasting

- Is the weather forecasting problem “solved” in 2021?
- Will the current solution in 2021 valid in the future?

What's up for today?

- A kind archeology tour
- Examples in Meteorology
- **Examples in Eye Tracking**

Eye Tracking



- 1 Free examination.
- 2 Estimate material circumstances of the family
- 3 Give the ages of the people.
- 4 Surmise what the family had been doing before the arrival of the unexpected visitor.
- 5 Remember the clothes worn by the people.
- 6 Remember positions of people and objects in the room.
- 7 Estimate how long the visitor had been away from the family.
- 3 min. recordings of the same subject
-
-
-

Neural Network in Eye Tracking

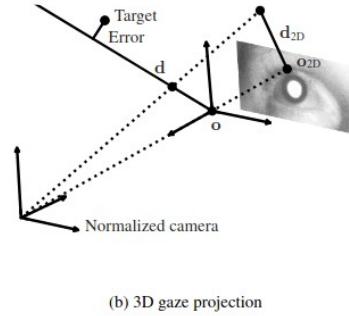
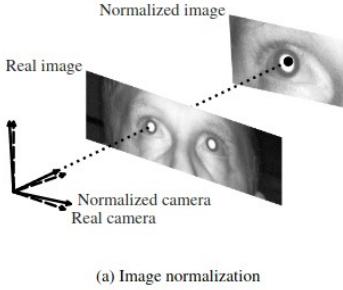


Figure 2: (a) The image captured by the physical camera is warped into a normalized camera looking directly at the reference point, in this case the persons right eye. (b) The 2D gaze origin and gaze direction are combined with the corrected distance to form a gaze ray in 3D space. The miss distance between the gaze ray and the gaze target is the loss used to train the neural network.

camera coordinate system by $\mathbf{g}(t) = \mathbf{R}^{-1}\hat{\mathbf{g}}(t)$.

3.2. 3D gaze projection

Here we describe how the output from the network is translated into a pair of gaze rays. The network has five outputs. For each eye, it predicts a 2D gaze origin \mathbf{o}_{2D} and a 2D gaze direction \mathbf{d}_{2D} . It also generates a distance correction term, c , which is common to both eyes. We assume that the distance from eye to camera is approximately the same for both eyes, and our estimate of that distance will be called ρ . First, given the input image and the eye detections, we find a rough distance ρ_{rough} such that the separation between the eyes becomes 63 mm at that distance, approximately the average human interocular distance [6]. This distance is then

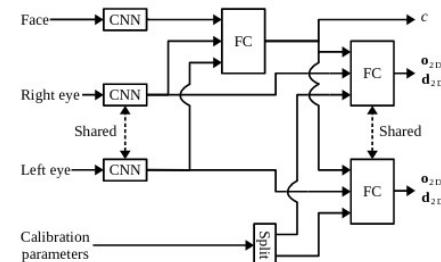


Figure 3: Network architecture.

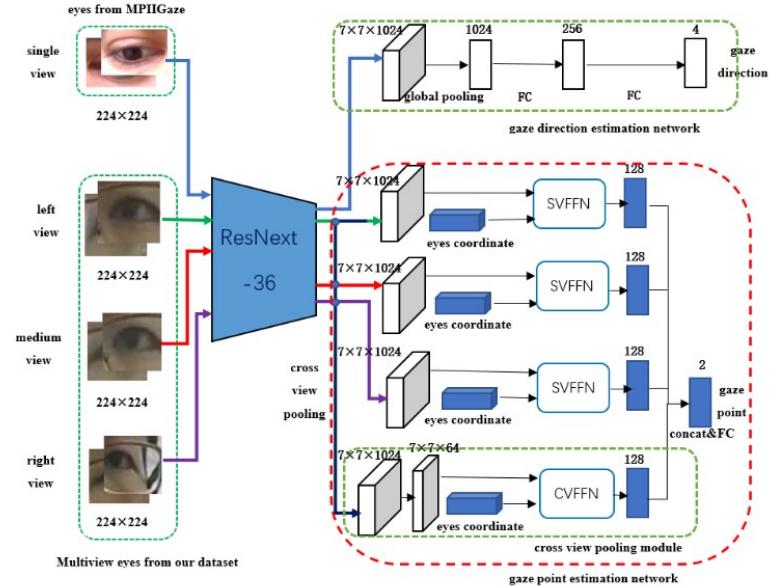
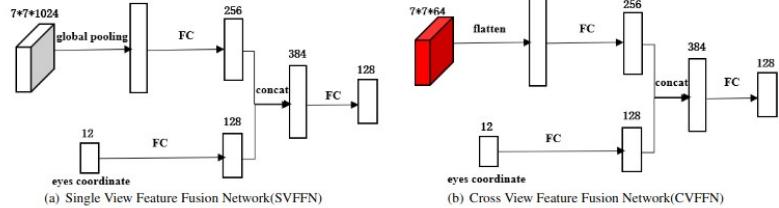
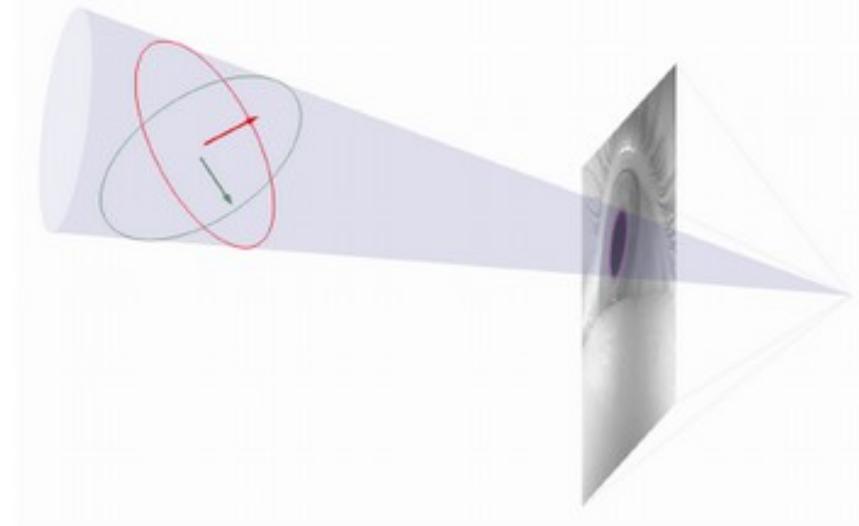
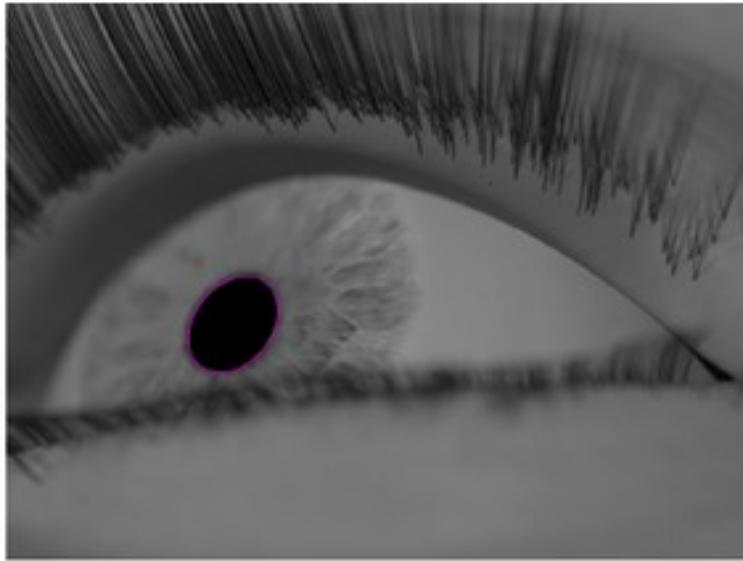


Figure 2: Overall architecture of multiview multitask network for the 3-D gaze direction and 2-D gaze point estimations.



3D Eye Model



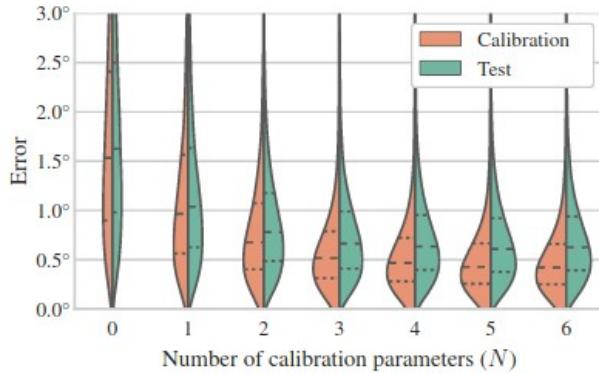
3D Eye Model

	Mean	Std. dev.
Ellipse Unprojection	3.5558	4.2385
Unoptimised Model	2.6890	1.5338
Region Contrast Maximisation	2.2056	0.1629
Edge Distance Minimisation	1.6831	0.3372

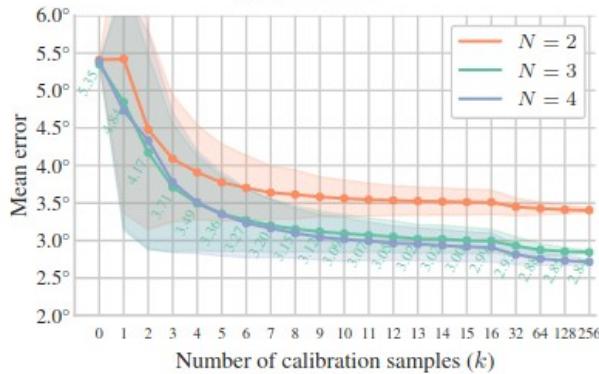
	Mean	Std. dev.
Pupil Tracker	3.9301	5.8056
Unoptimised Model	4.4420	4.9157
Region Contrast Maximisation	2.1194	1.1354
Edge Distance Minimisation	2.7474	2.7615

Based on 3D eye model in 2013

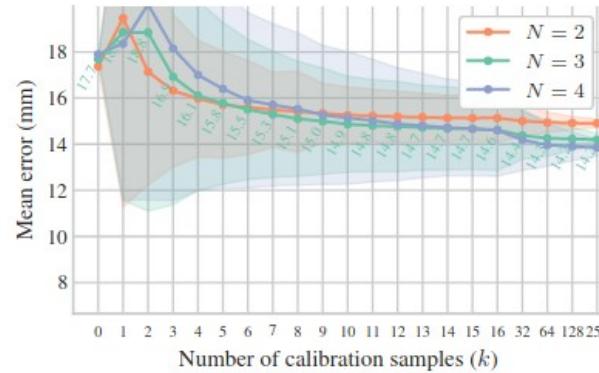
Some Gaze Estimation Model based on Neural Network



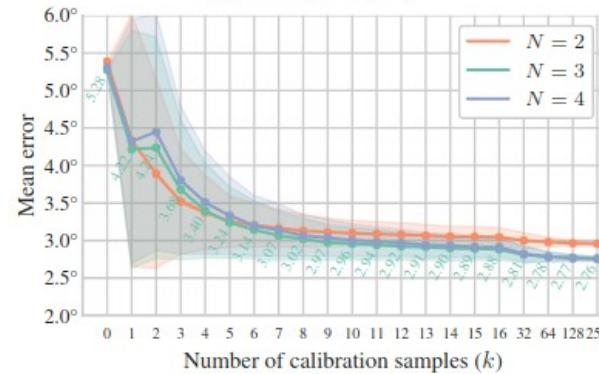
(a) NIR dataset



(c) GazeCapture-to-MPIIGaze



(b) GazeCapture (test)



(d) Within-MPIIGaze

What does matter in the Algorithm?

- Accuracy
- Latency and Throughput
- Training time and inference time
- Memory requirements

Dataset

Calibration Samples

Gaze Estimation

- Is the Gaze Estimation problem “solved” in 2021?
- Will the current solution in 2021 valid in the future?

Conclusion

- What do you think,

it will be **FUNDAMENTAL and CANONICAL**

in our future?

