Link to these slides: http://tiny.cc/nips

arXiv:1711.03121

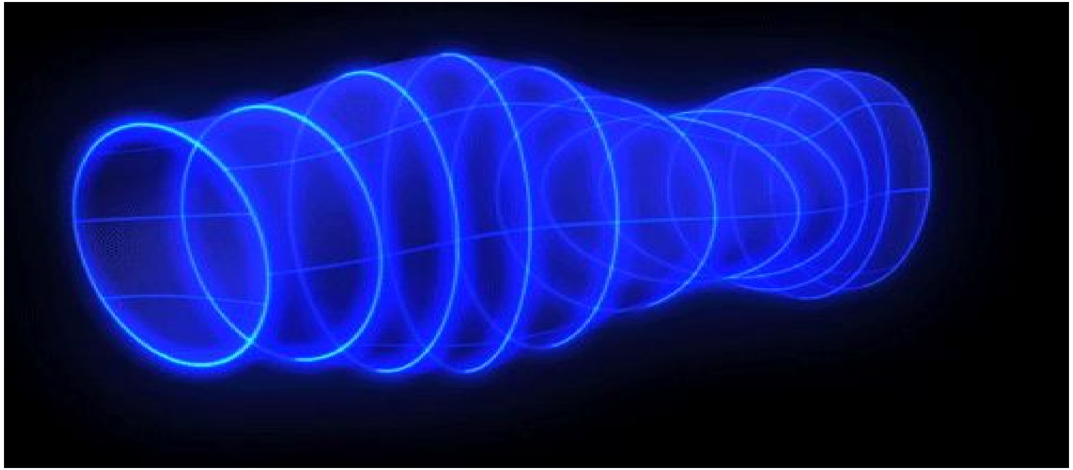
Deep Learning for Gravitational Wave Analysis
Results with LIGO Data

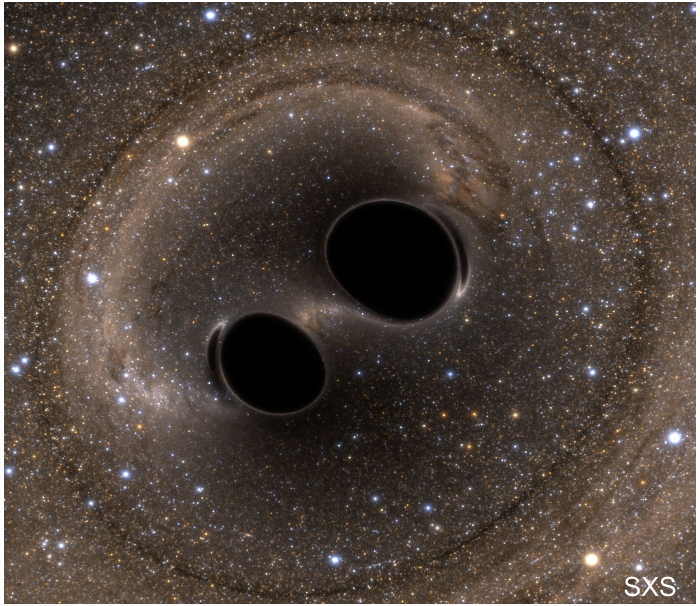
Daniel George & E. A. Huerta

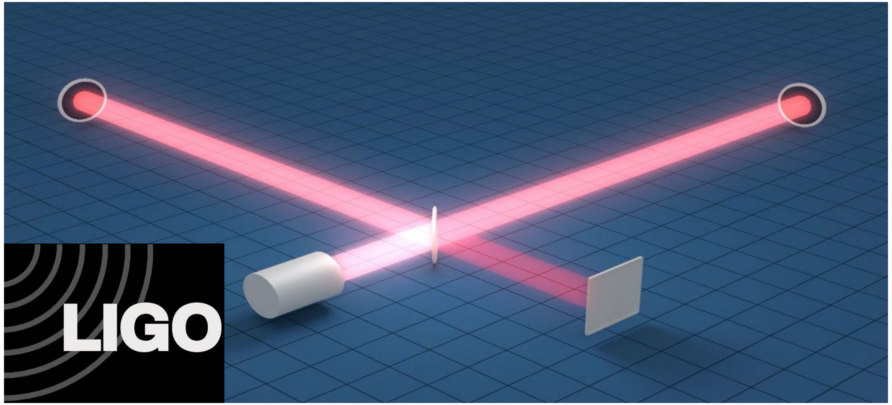
NCSA Gravity Group - http://gravity.ncsa.illinois.edu/
Department of Astronomy
University of Illinois at Urbana-Champaign

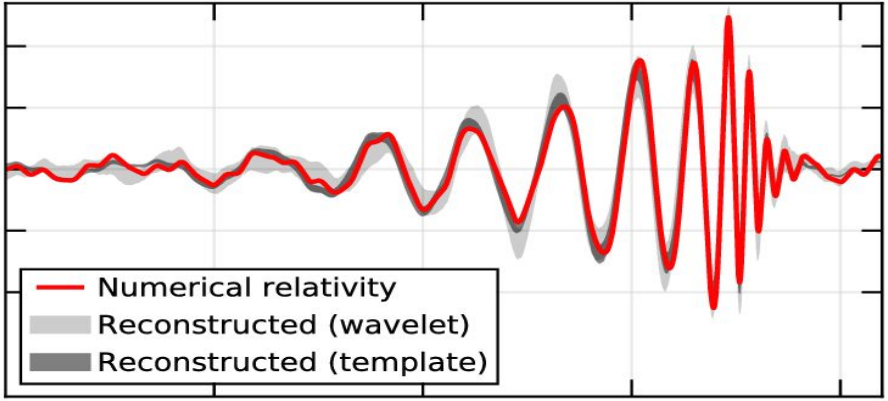
Dec 8, 2017

Gravitational Waves



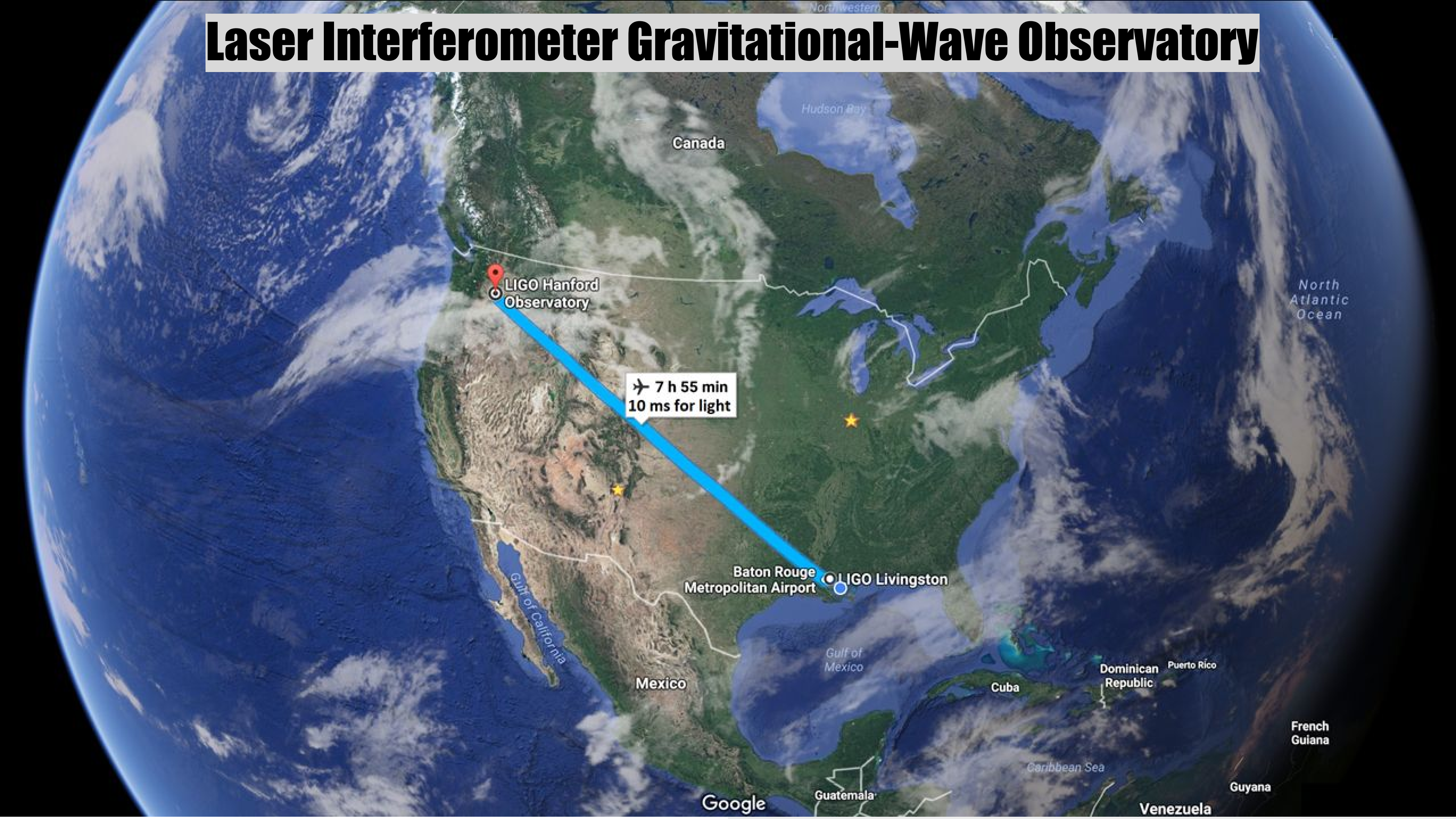


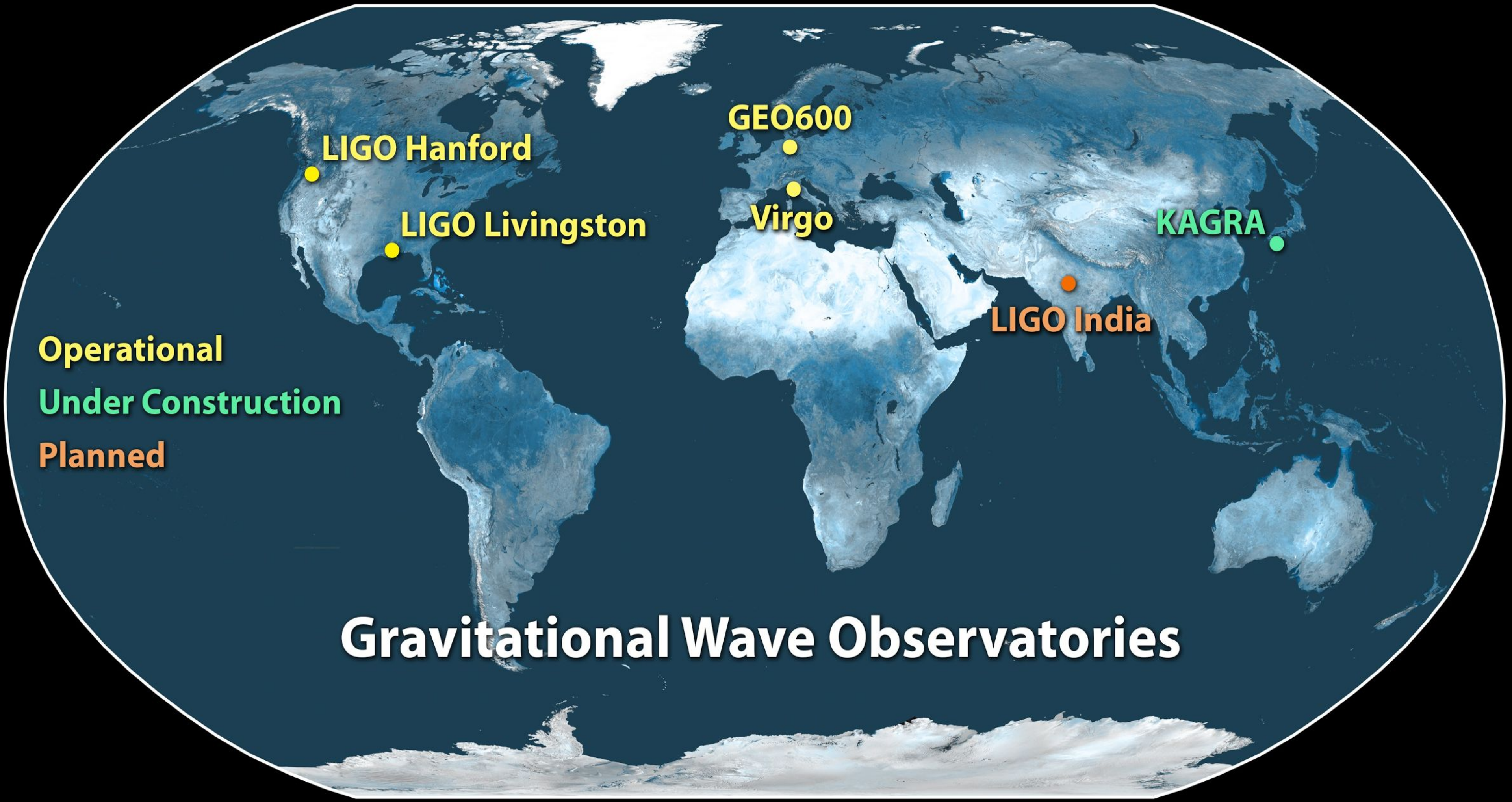




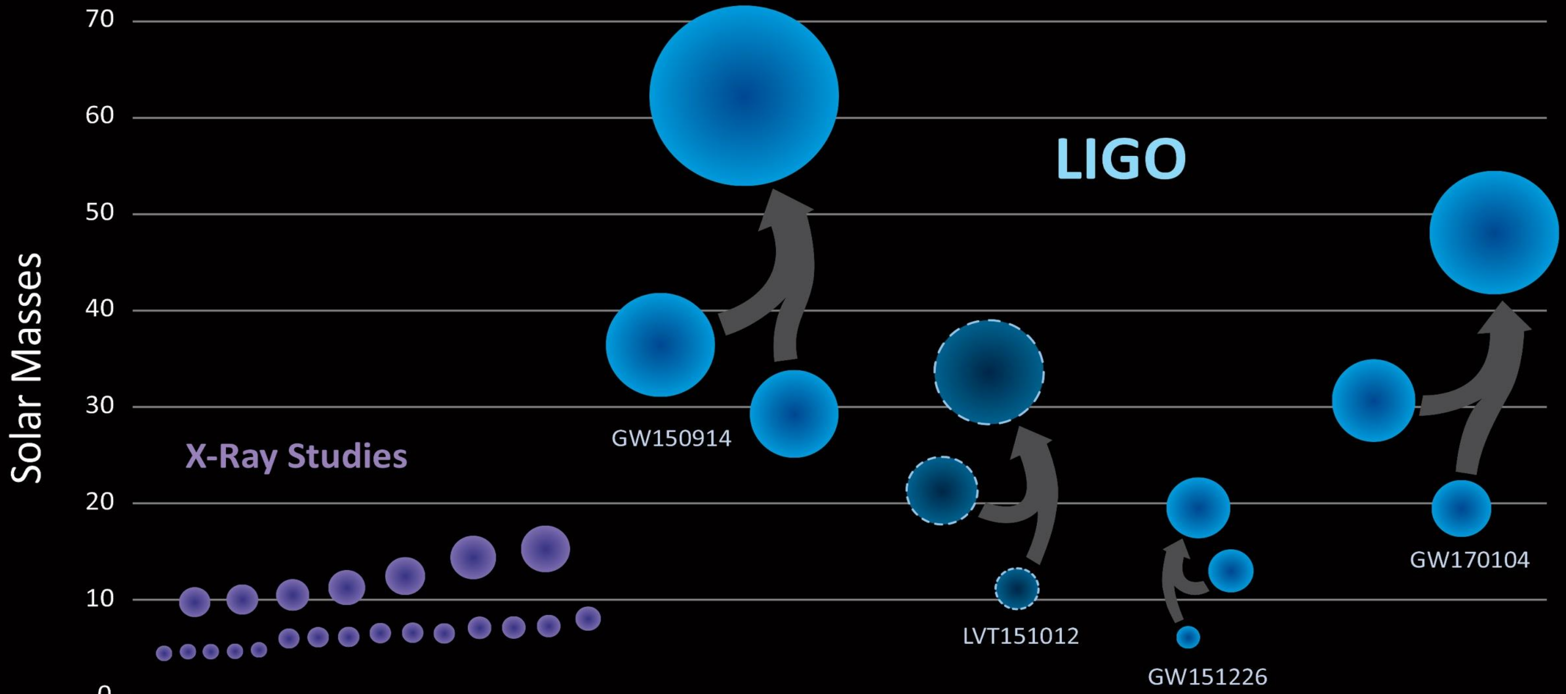
Source: ligo.org

Laser Interferometer Gravitational-Wave Observatory



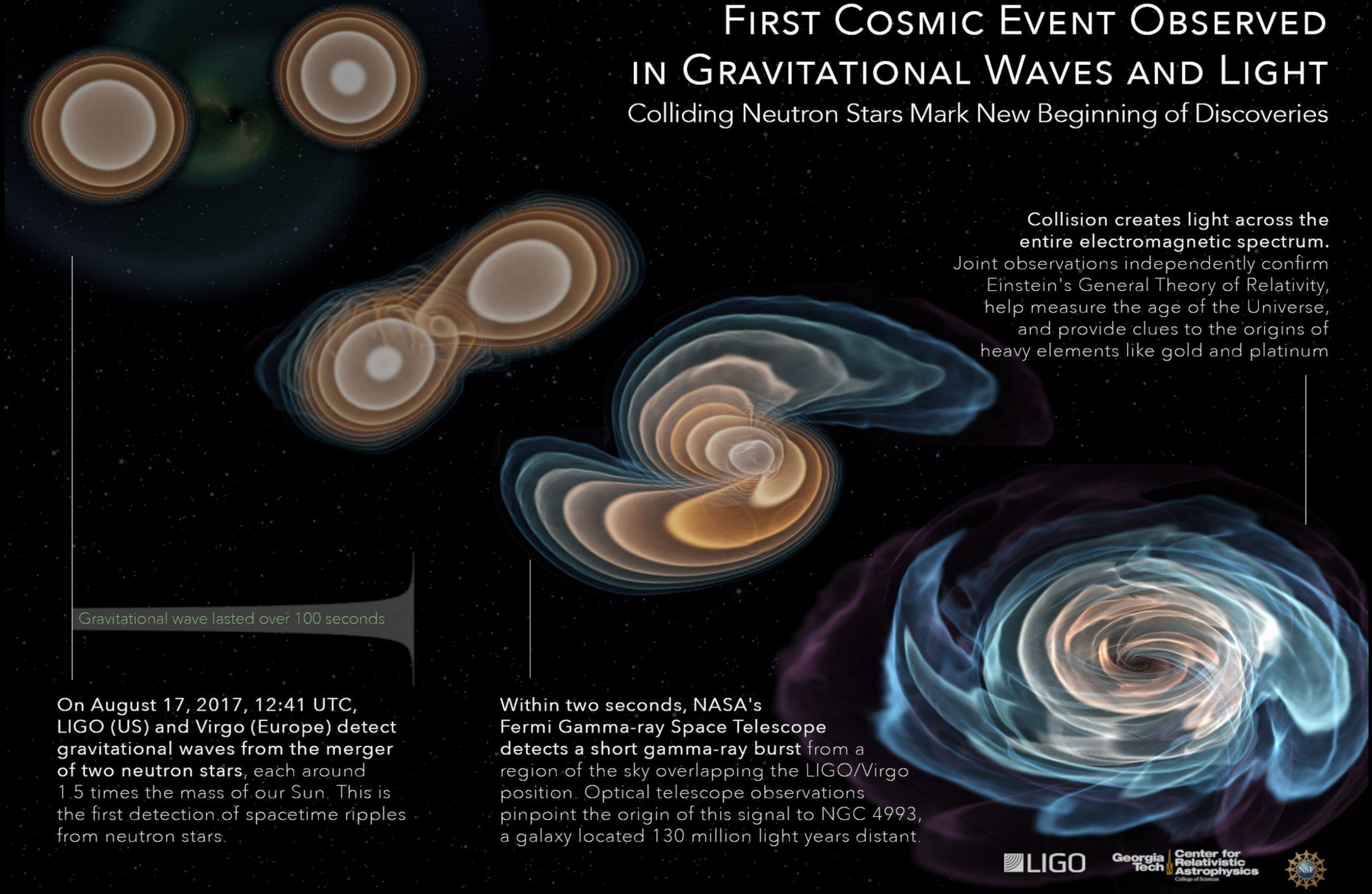

GEO600
LIGO Hanford
Virgo
KAGRA
LIGO Livingston
LIGO India
Operational
Under Construction
Planned
Gravitational Wave Observatories

Black Hole Detections



First Cosmic Event Observed in Gravitational Waves and Light

Colliding Neutron Stars Mark New Beginning of Discoveries



Collision creates light across the entire electromagnetic spectrum. Joint observations independently confirm Einstein's General Theory of Relativity, help measure the age of the Universe, and provide clues to the origins of heavy elements like gold and platinum

On August 17, 2017, 12:41 UTC, LIGO (US) and Virgo (Europe) detect gravitational waves from the merger of two neutron stars, each around 1.5 times the mass of our Sun. This is the first detection of spacetime ripples from neutron stars.

Within two seconds, NASA's Fermi Gamma-ray Space Telescope detects a short gamma-ray burst from a region of the sky overlapping the LIGO/Virgo position. Optical telescope observations pinpoint the origin of this signal to NGC 4993, a galaxy located 130 million light years distant.

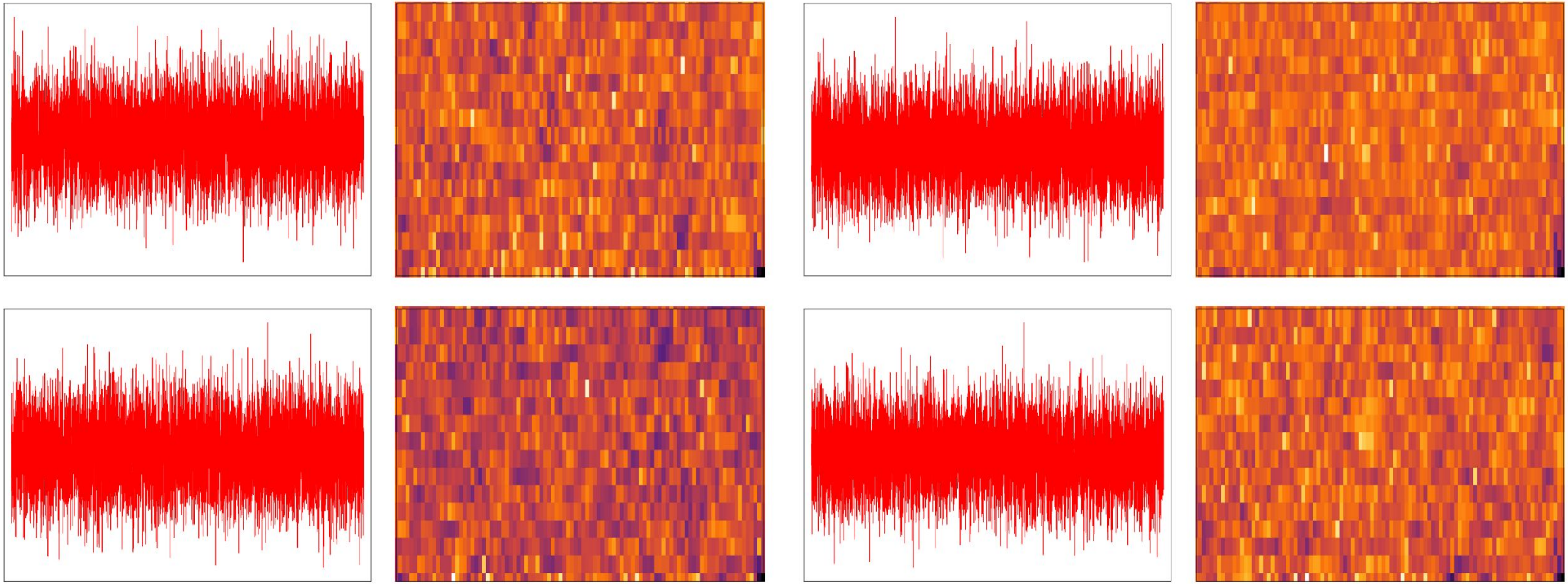
LIGO Georgia Tech Center for Relativistic Astrophysics College of Sciences

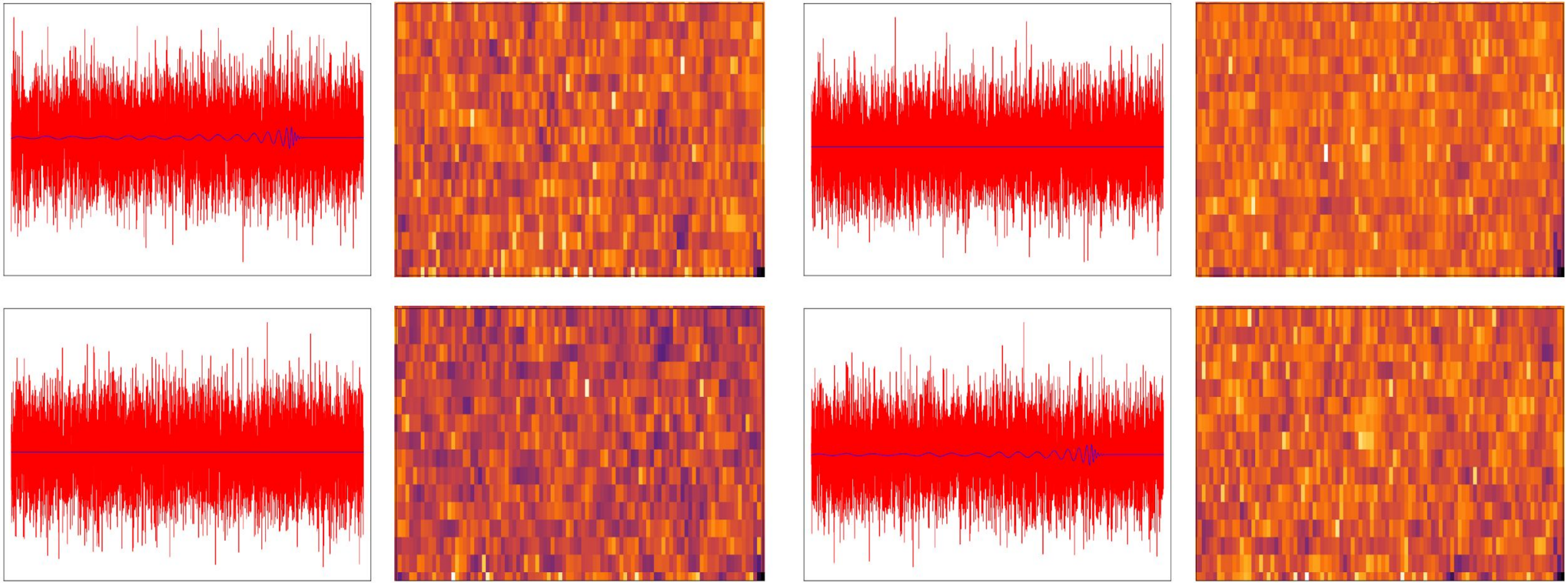
The Royal Swedish Academy of Sciences has decided to award the

2017 NOBEL PRIZE IN PHYSICS

to

Challenge



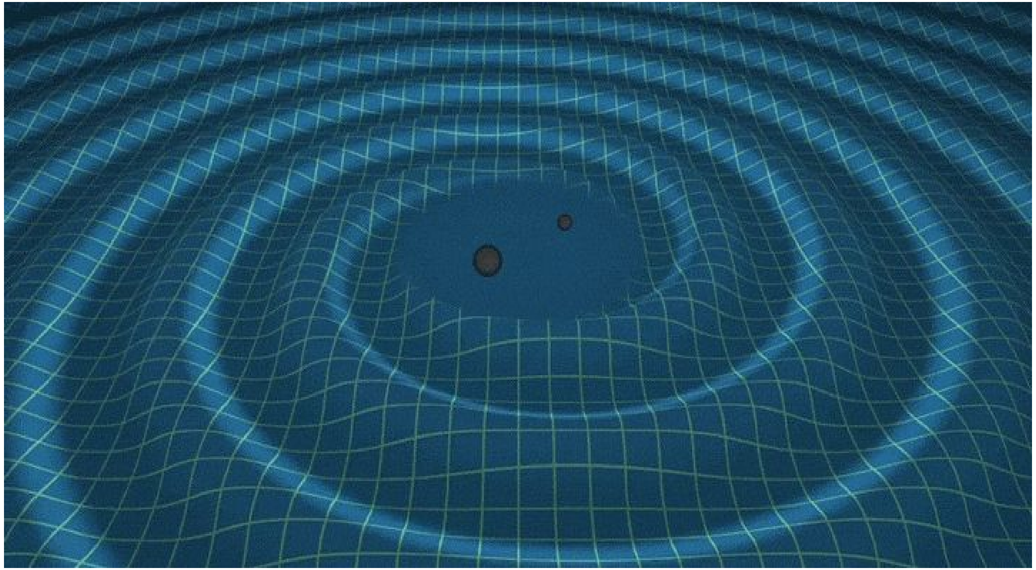


Classification & Regression

Numerical Relativity - Supercomputing

einstein toolkit

NCSA

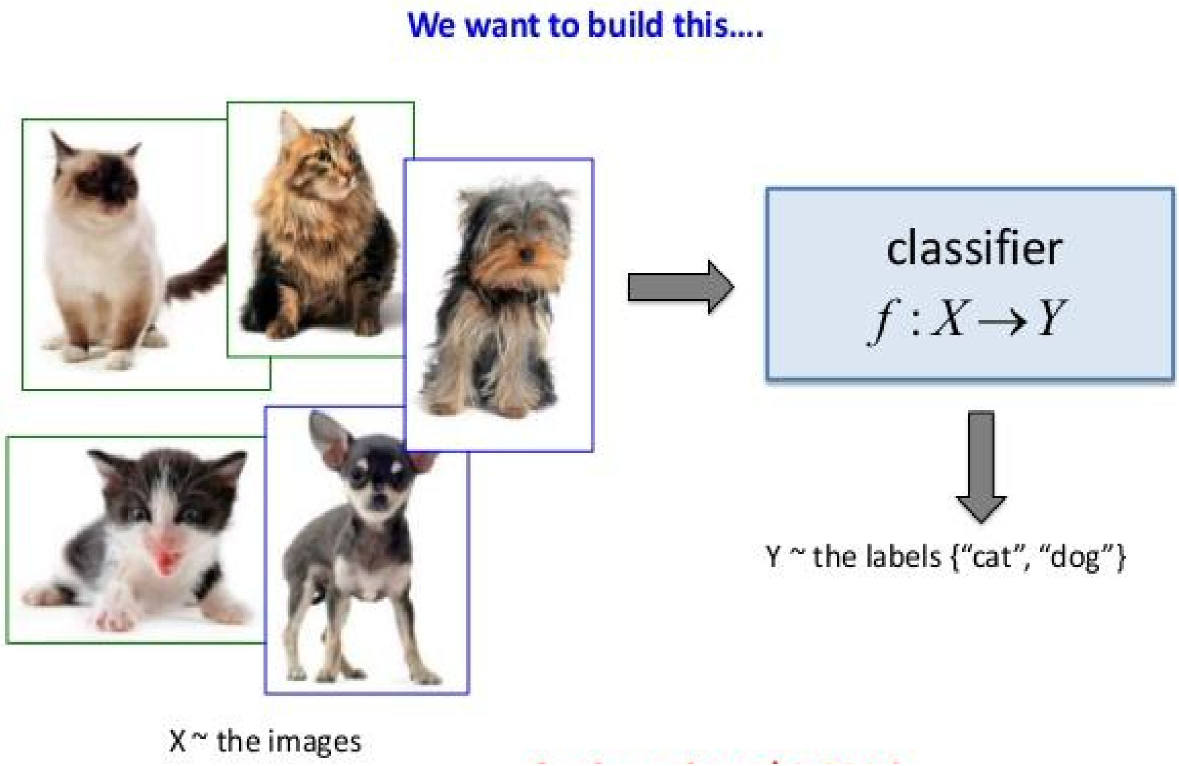


Method?

Matched-Filtering:

Compare an input image with millions of photos of cats and dogs and see which matches best.

Template matching is not scalable



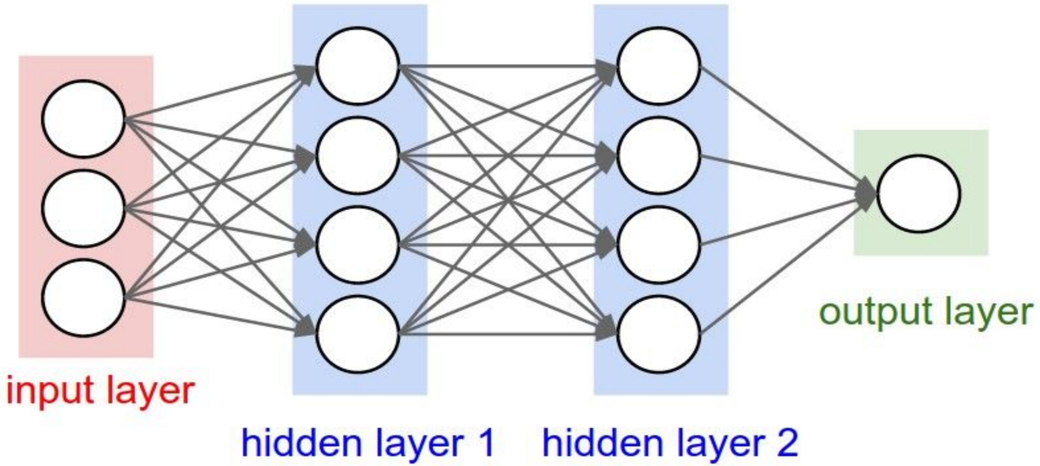
Solution:

Deep Learning!

Source: https://www.slideshare.net/andrewgardner5811/deep-learning-for-data-scientists-dsatl-talk-alpharetta-20140108

Deep Learning

Overview



- Very long networks of artificial neurons (dozens of layers)
- State-of-the-art algorithms for face recognition, object identification, natural language understanding, speech recognition and synthesis, web search engines, self-driving cars, games (Go) etc.

- Does not require hand-crafted features to be extracted first
- Automatic end-to-end learning
- Deeper layers can learn highly abstract functions

Signal Processing with Convolutional Networks

Our method: Deep Filtering

CNNs for directly processing highly noisy time-series data for classification and regression.

Advantages

- Can process raw/whitened data
- Automatically learns optimal strategies
- Train once. Constant-time for evaluation
- Optimized hardware (GPUs, FPGAs)
- Resilient to glitches, non-stationary noise
- Does not perform template matching
- Learns patterns connecting signals
- Interpolates to new templates
- Small and efficient (few MBs)

Designing 1-D CNNs

- Explored only simple designs.

- Up to 4 dilated convolutional layers and 3 fully connected layers.

2 nets: *Classifier* for detection;

Predictor for estimating source parameters

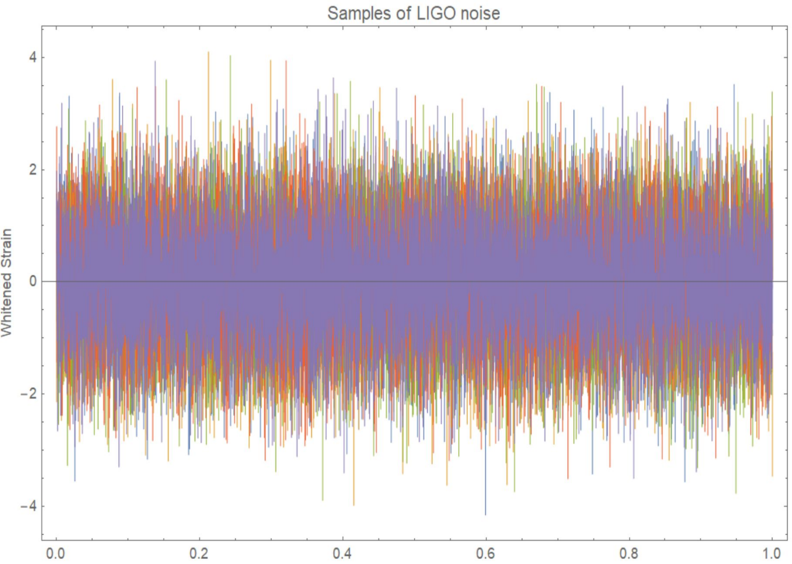
	Input (1s, 8192Hz)	vector (size: 8192)
1	Reshape Layer	tensor (size: 1 × 1 × 8192)
2	Convolution Layer	tensor (size: 16 × 1 × 8177)
3	Pooling Layer	tensor (size: 16 × 1 × 2045)
4	Ramp	tensor (size: 16 × 1 × 2045)
5	Convolution Layer	tensor (size: 32 × 1 × 2017)
6	Pooling Layer	tensor (size: 32 × 1 × 505)
7	Ramp	tensor (size: 32 × 1 × 505)
8	Convolution Layer	tensor (size: 64 × 1 × 477)
9	Pooling Layer	tensor (size: 64 × 1 × 120)
10	Ramp	tensor (size: 64 × 1 × 120)
11	Flatten Layer	vector (size: 7680)
12	Linear Layer	vector (size: 64)
13	Ramp	vector (size: 64)
14	Linear Layer	vector (size: 2)
15	Softmax Layer	vector (size: 2)
	Output	vector (size: 2)

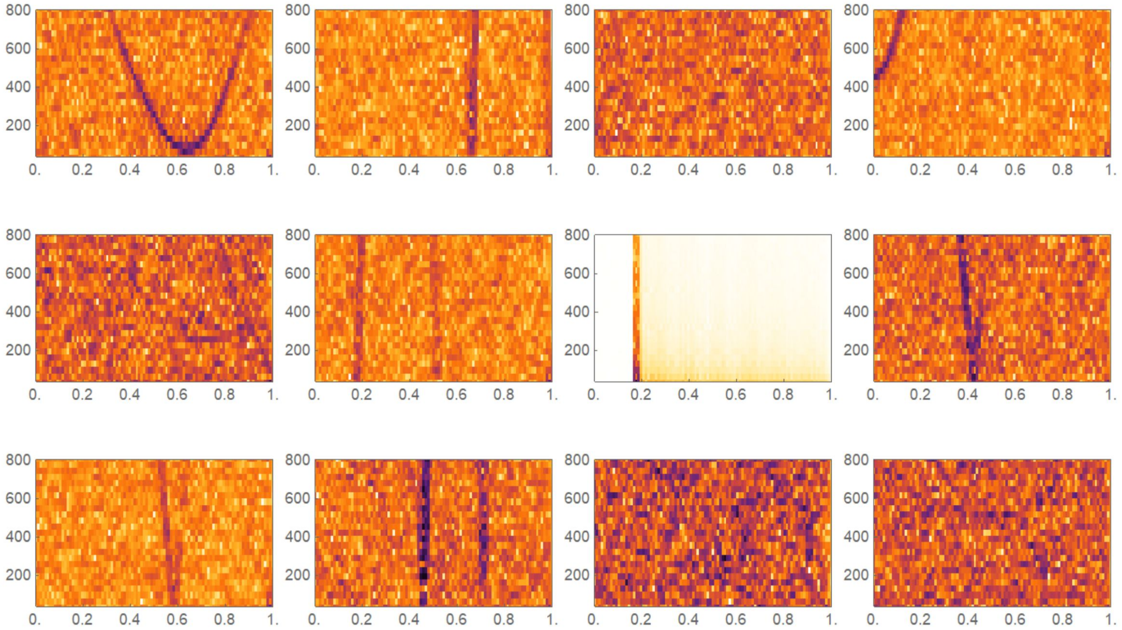
Using Real LIGO Data

1. Added real noise from LVT151210 and GW151226 for training (4096s each).
   Open source, data taken from https://losc.ligo.org/events/GW150914/

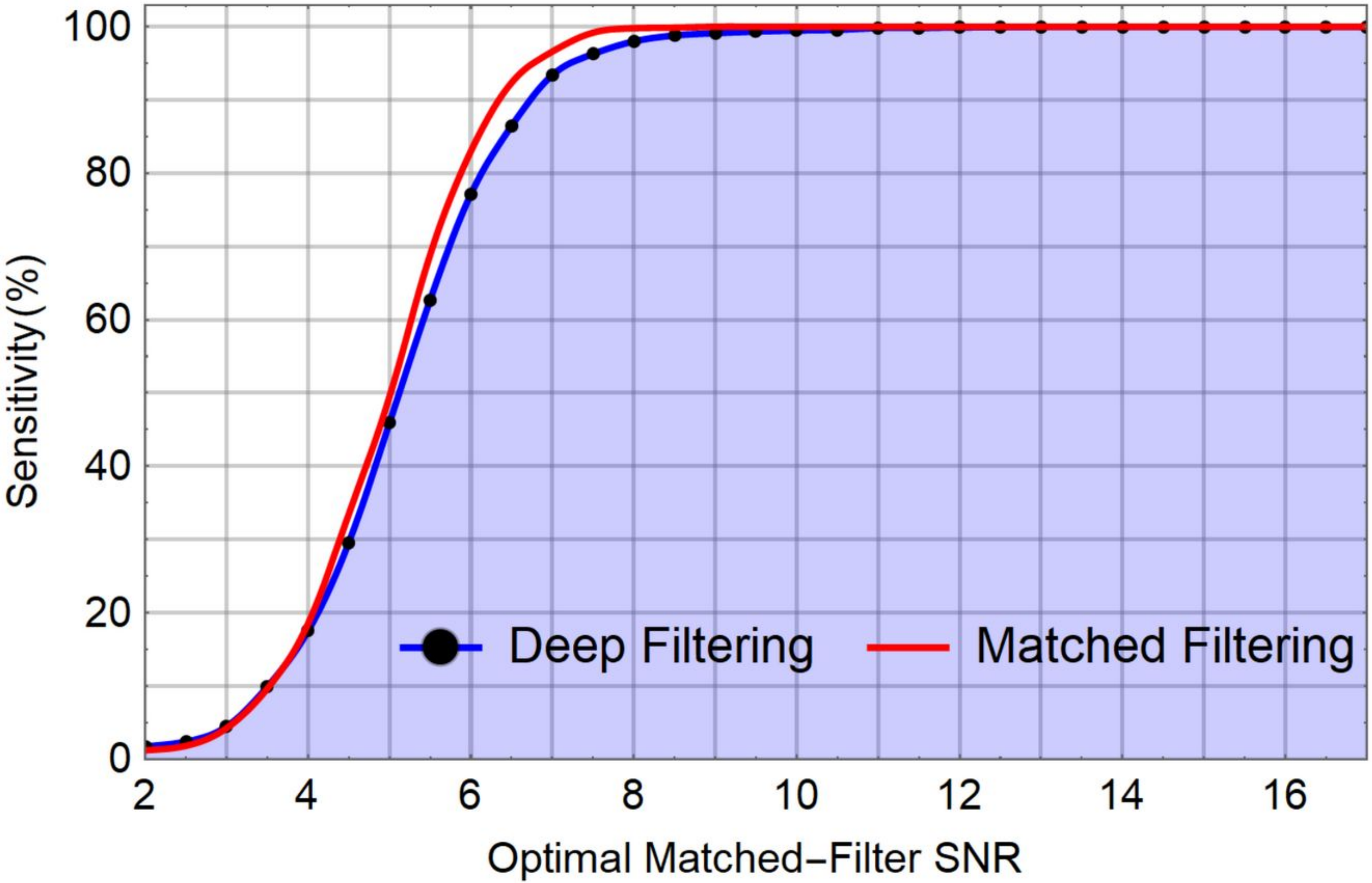
2. Tested on real data from GW150914 (includes many glitches as shown below)

3. Same nets work on non-stationary colored noise with different PSD without re-training.





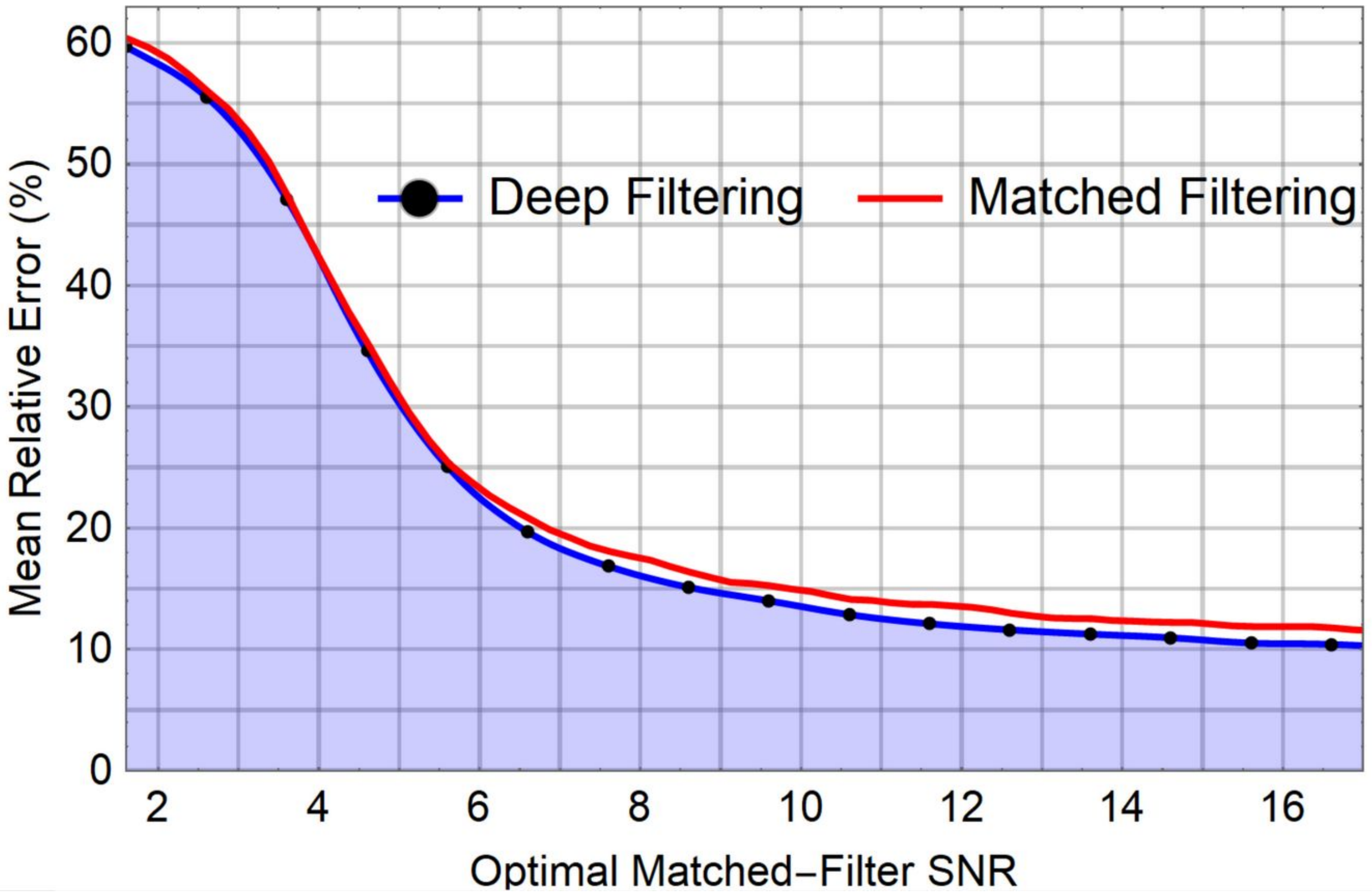
Accuracy of Classifier (Detection)



100% Sensitivity for SNR > 10

False Alarm Rate < 0.6%

Relative Error in Predicting Masses

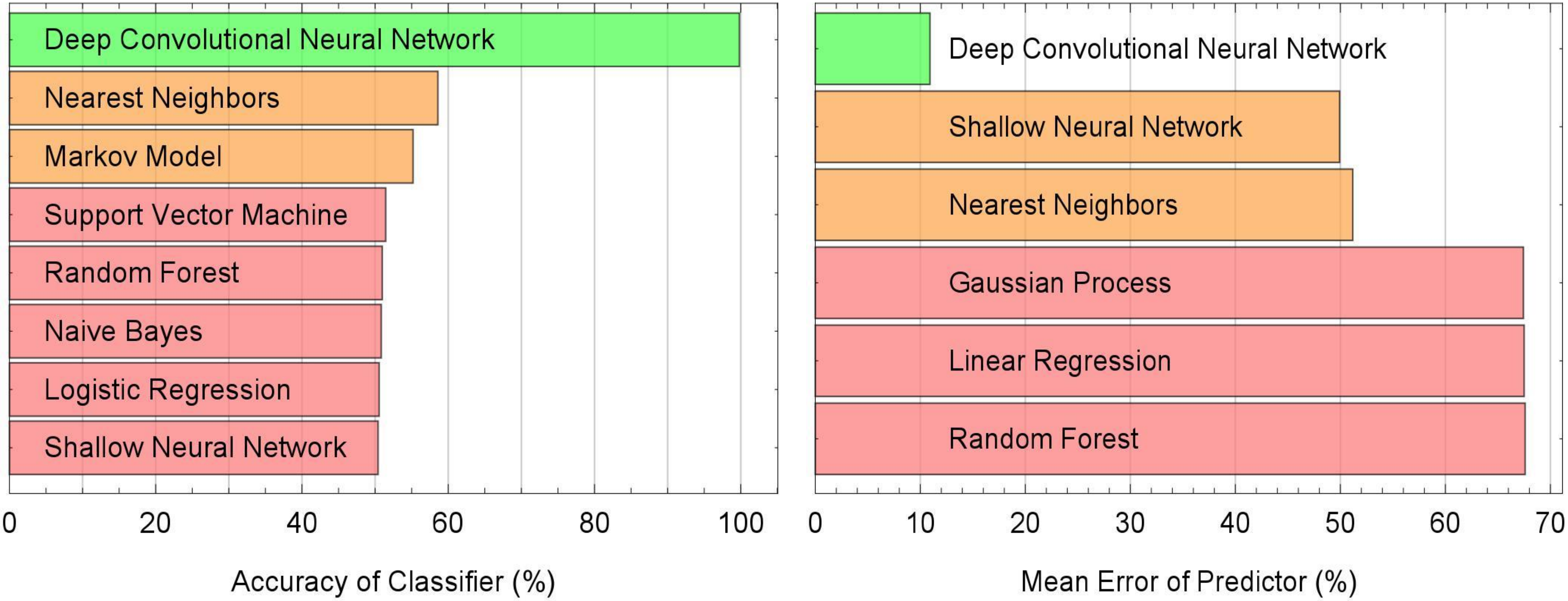


Deep Filtering error < 5% for SNR>50

Can interpolate between templates

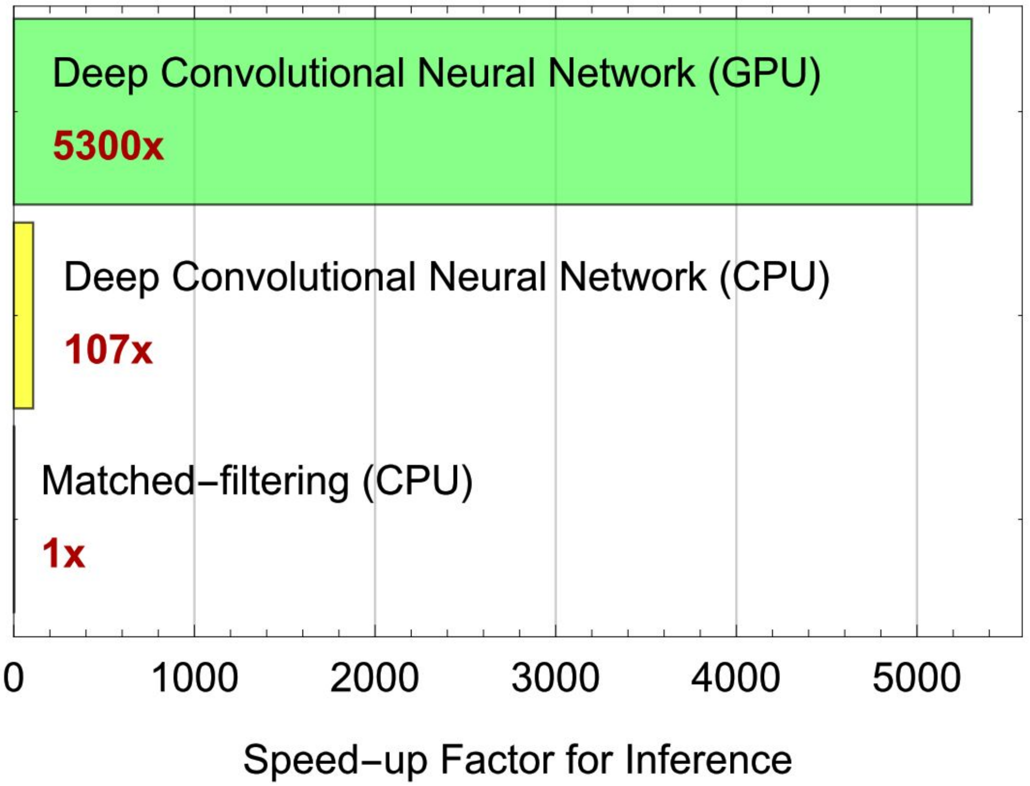
Matched-Filtering error with same template bank is always > 11%

Detection and Parameter Estimation

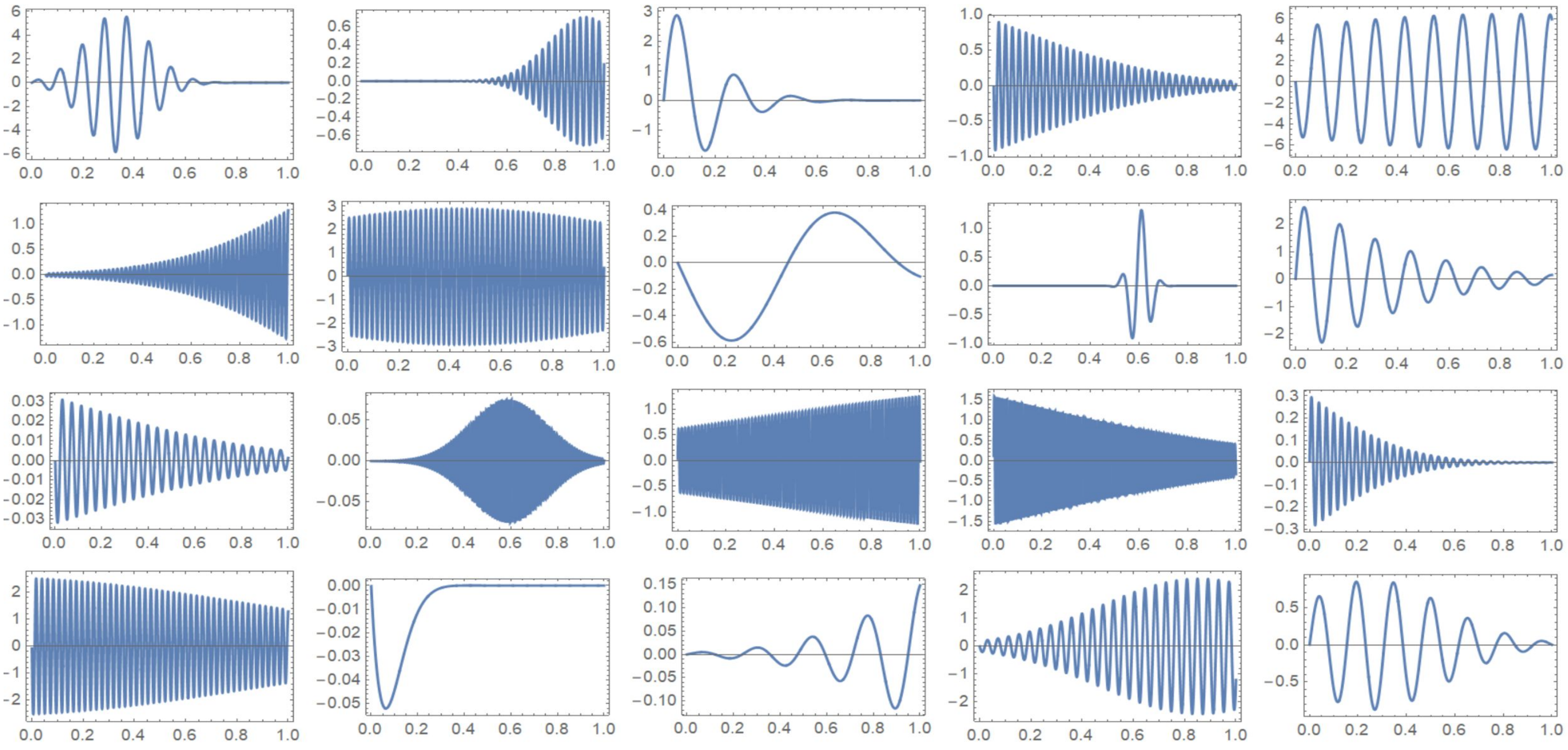


Speed-Up

- Real-time analysis (milliseconds).
- Constant time regardless of number of templates, after training once.
- Thousands of inputs can be processed at once on a cheap GPU.
- Dedicated inference engines can offer additional speed-up.

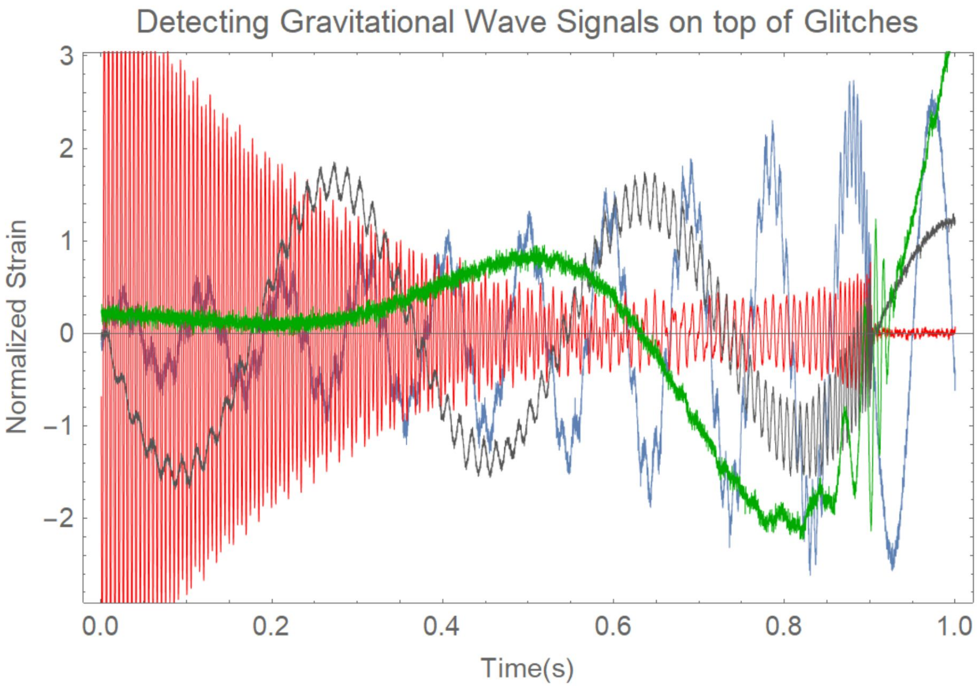


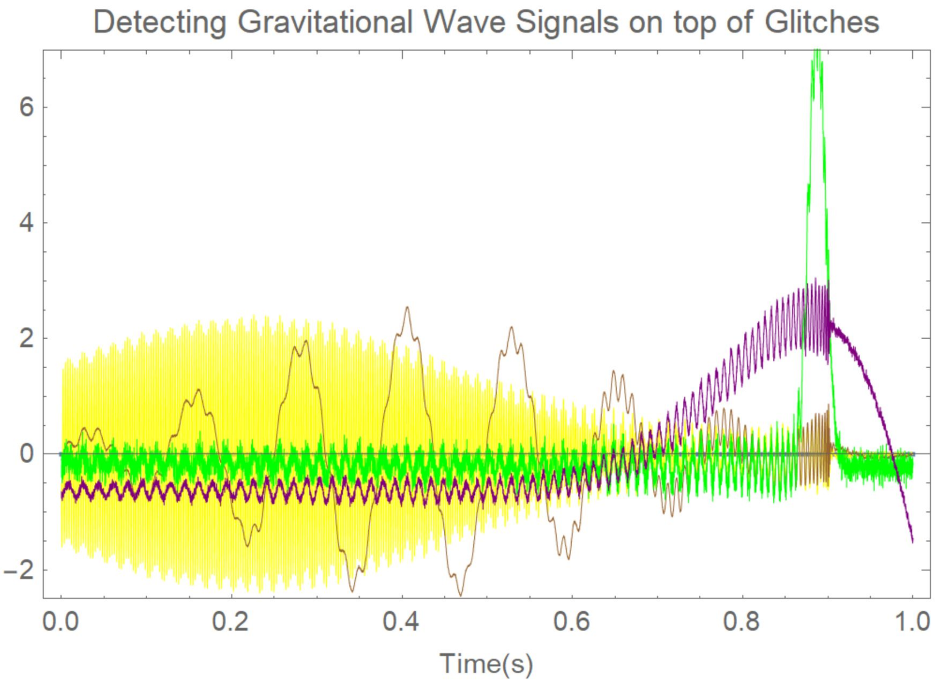
Automatically resilient to Glitches



<u>**False Alarm Rate with sine-gaussian glitch injections:**</u> Matched-Filter = ~**30%**, Deep Filtering = <**1%**

Even works for Events during Glitches!





Successfully recovered ~80% of signals injected in real noise plus sine-gaussian glitches.

Mean relative error of parameter estimation during glitches <30% for SNR>10

Live Demo:
www.tiny.cc/DLGW

Detecting GW150914

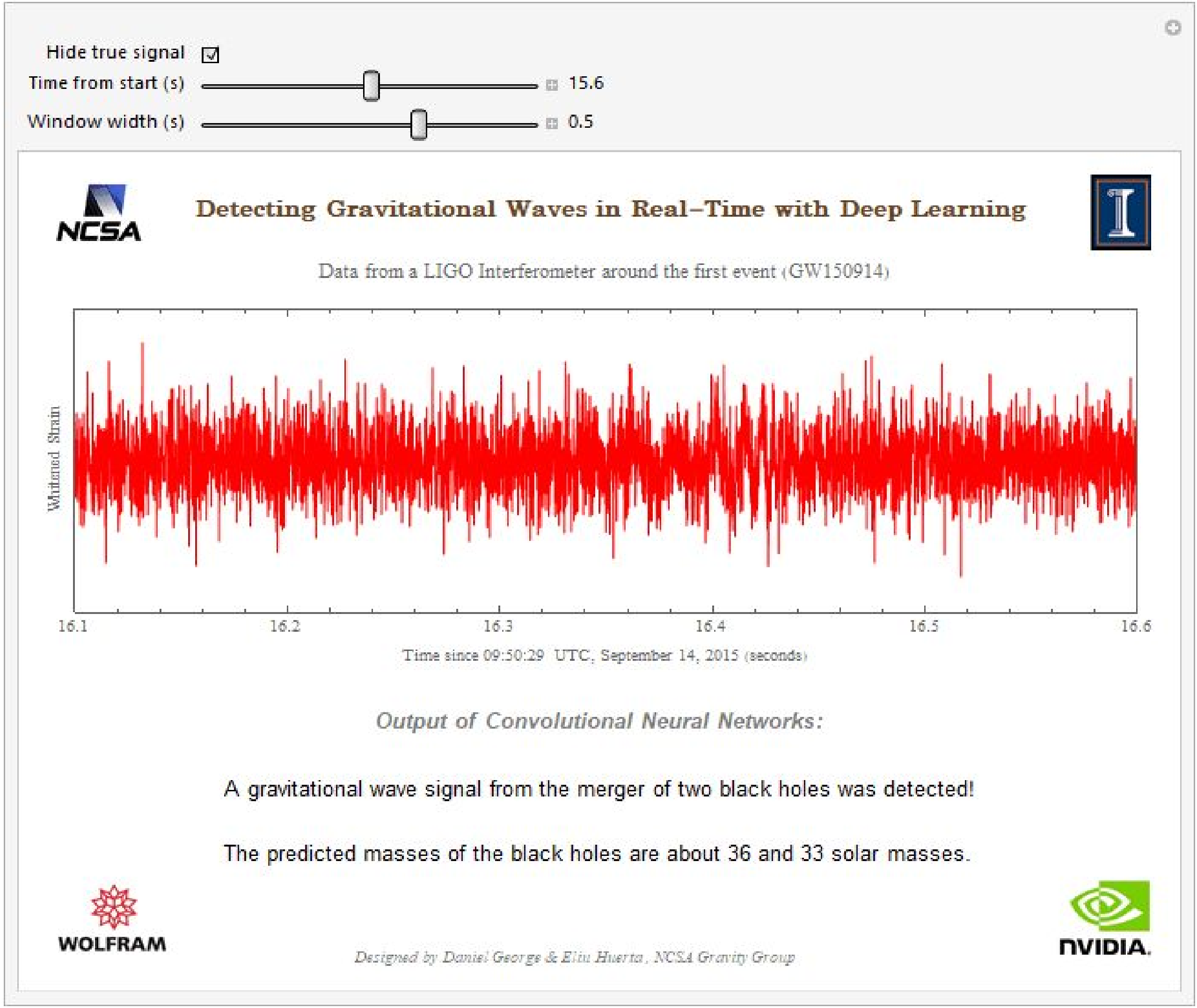
Data not included in training

Trained with only non-spinning, non-eccentric simulations

~1s to analyze 4096s of data.

Masses correct within error bars

No False Alarms with two detectors!



New Types of GWs

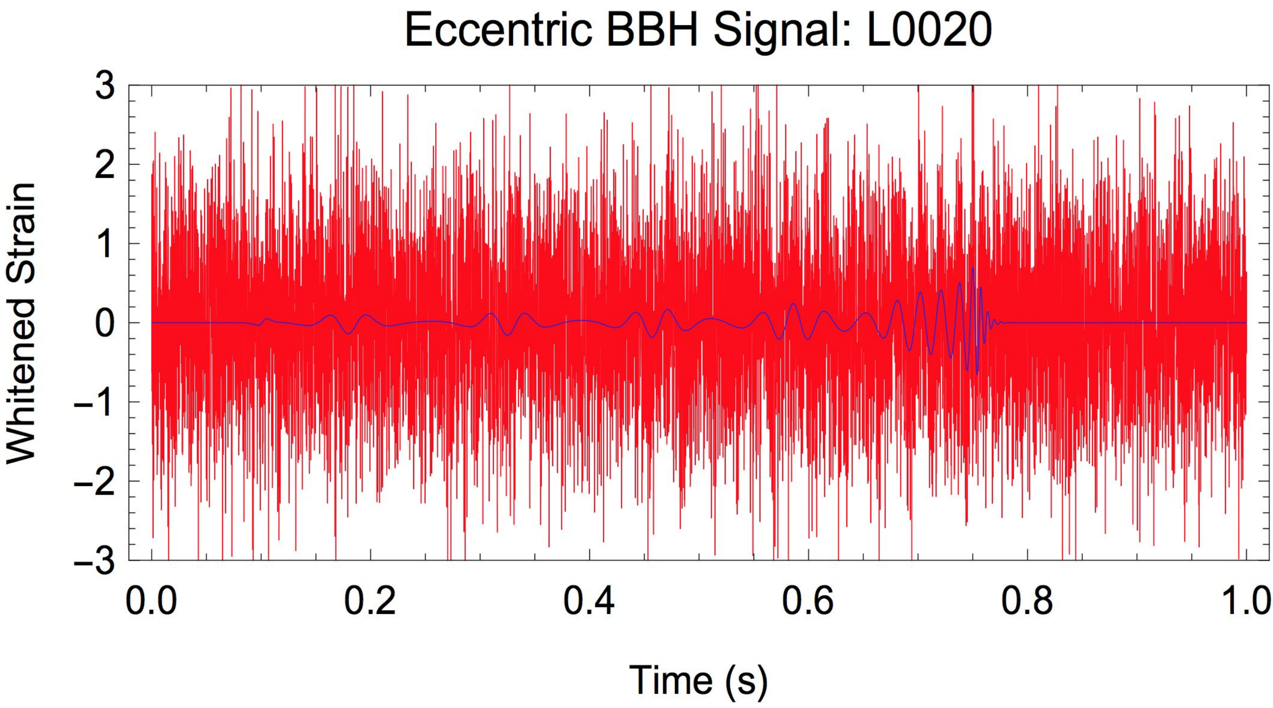
Eccentric, Spinning

Not included in training.

Same accuracy of detection.

DNNs learned to generalize.

Missed by current methods.

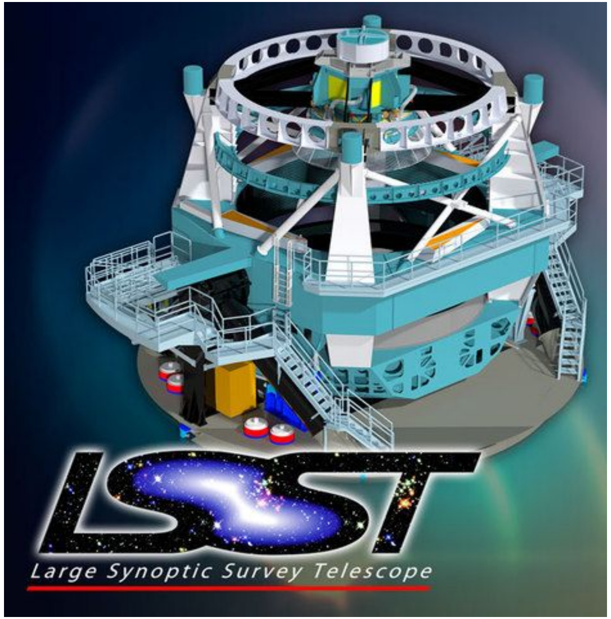


Real-time Multimessenger Astrophysics

Hear **gravitational waves**

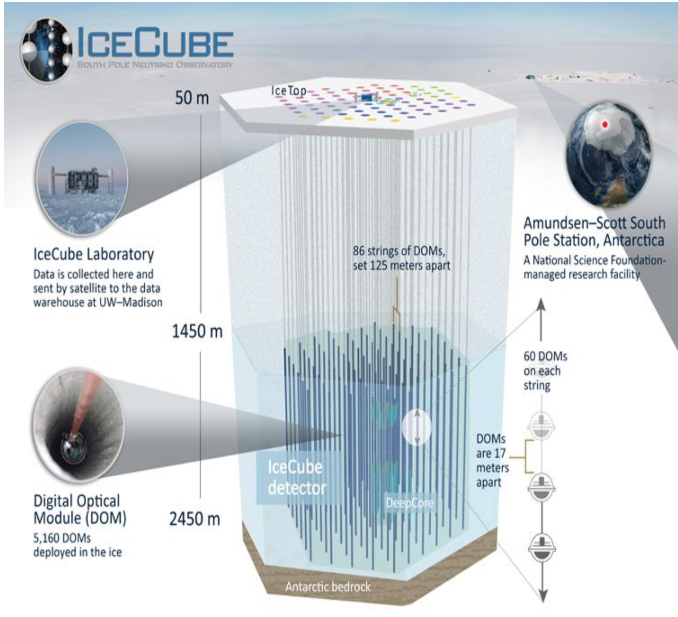
LIGO, VIRGO, KAGRA, eLISA

See **electromagnetic waves**



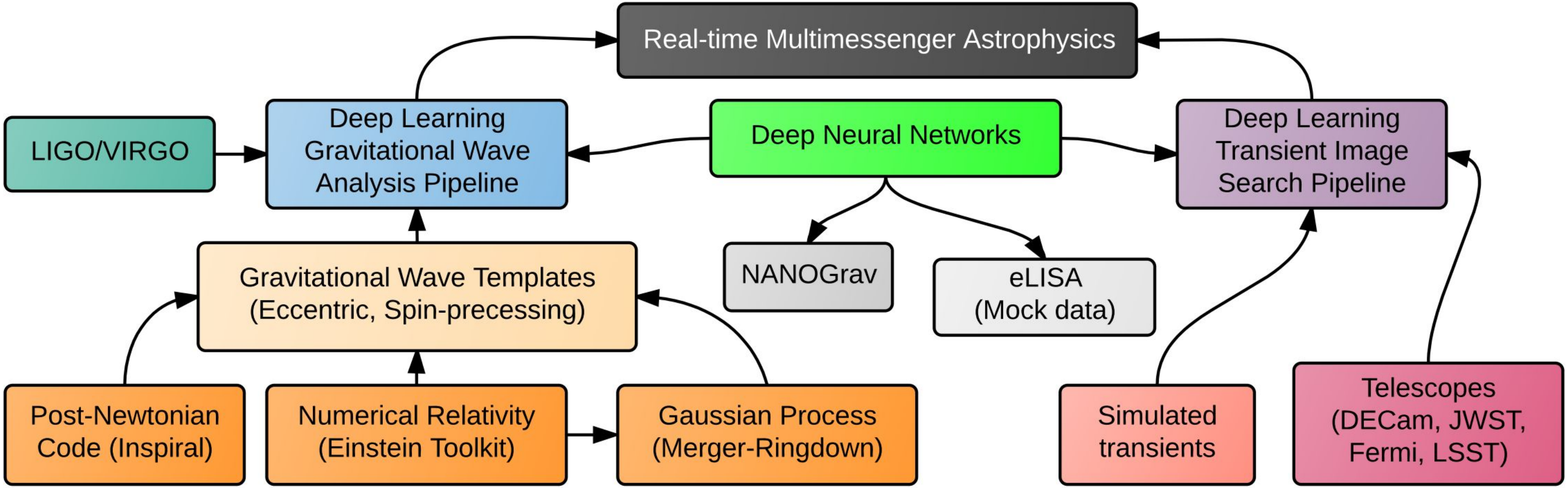
DES, LSST, JWST, WFIRST

Feel **astroparticles**



IceCube (neutrinos)

Enabling Real-Time Multimessenger Astrophysics



Link to these slides: **www.tiny.cc/nips**

Extended article: **arXiv:1711.03121**
Awarded 1st place at the ACM student research competition at SC17

Conclusion

HPC (*Blue Waters*) + AI (*Deep Learning*) + GPUs
= Real-time Big Data Analysis for Science!

Email: **dgeorge5@illinois.edu** *Webpage:* **http://gravity.ncsa.illinois.edu**

Anomaly Detection, Unsupervised Clustering

Glitch Classification and Clustering for LIGO with Deep Transfer Learning

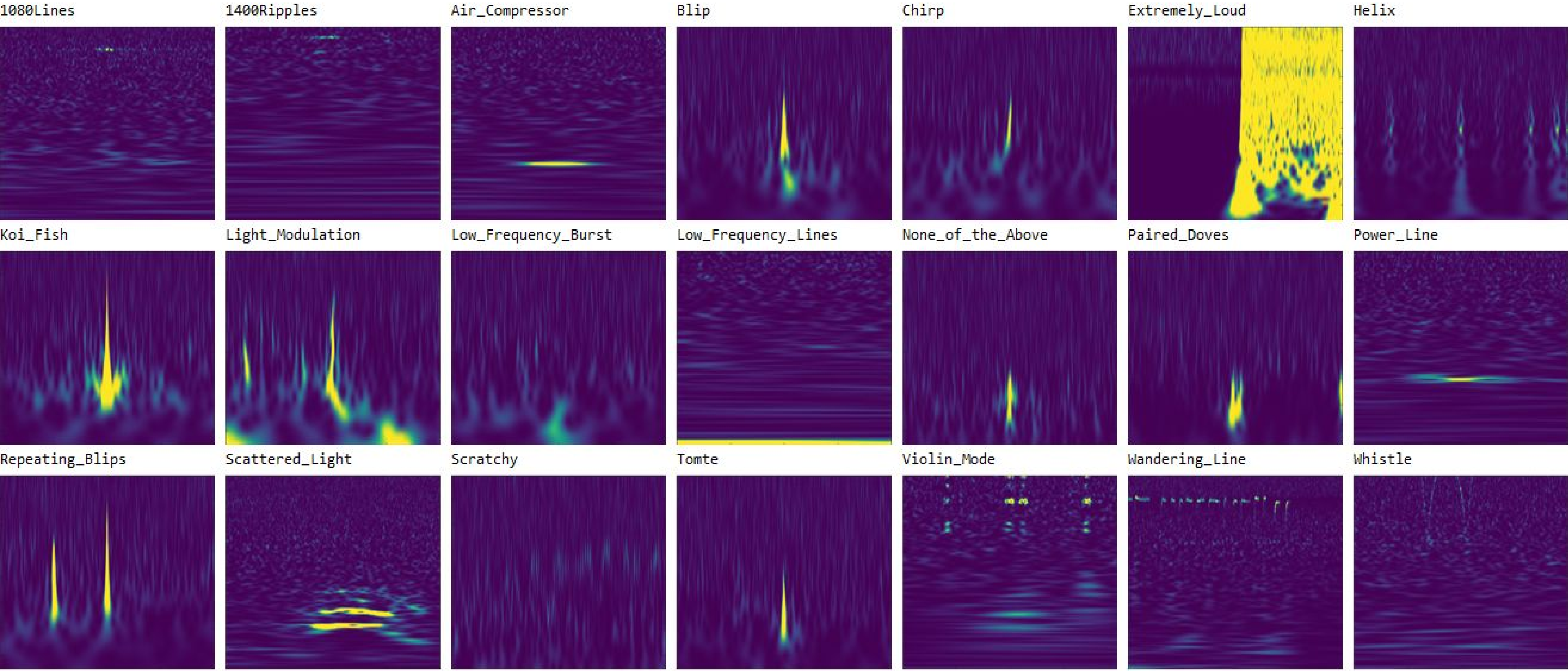
With LIGO O1 Gravity Spy Dataset

arXiv:1711.07468

Daniel George, Hongyu Shen, Eliu Huerta

National Center For Supercomputing Applications (NCSA)
University of Illinois at Urbana-Champaign

Classes in the Gravity Spy dataset



Transfer Learning Approach

Humans aren’t trained on large glitch datasets yet perform well. Why?

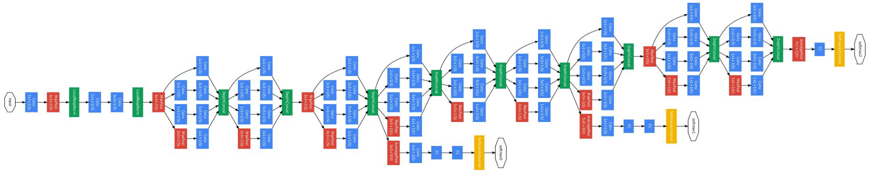
Indicates real-world pattern recognizers are useful even for glitch classification

Idea: Use pre-trained weights from state-of-the-art networks trained on real photos

Provides off-the-shelf features in the initial layers, fine-tuned later

Much faster (few minutes of training > 98.7% accuracy), less data needed

Fine-tuning Inception



98.8+ % (4 rounds / 5 min of training)

5ms to 15ms speed of evaluation

Perfect precision and recall:

1080 Lines, 1440 Ripples, Air Compressor,

Chirp, Helix, *Paired Doves*, Scratchy, Power Line

True Classes / Predictions

True Classes	1080Lines	1400Ripples	Air_Compressor	Blip	Chirp	Extremely_Loud	Helix	Koi_Fish	Light_Modulation	Low_Frequency_Burst	Low_Frequency_Lines	No_Glitch	None_of_the_Above	Paired_Doves	Power_Line	Repeating_Blips	Scattered_Light	Scratchy	Tomte	Violin_Mode	Wandering_Line	Whistle
1080Lines	66																					
1400Ripples		47																				
Air_Compressor			12																			
Blip				373												1						
Chirp					14																	
Extremely_Loud						91																
Helix							56															
Koi_Fish								165	1													
Light_Modulation						1			110	3						1						
Low_Frequency_Burst										131			1									
Low_Frequency_Lines											91											
No_Glitch											1	36										
None_of_the_Above						1		1	1	1			11						3			
Paired_Doves														6								
Power_Line															91							
Repeating_Blips				1												56						
Scattered_Light											1						91					
Scratchy																		71				
Tomte																			24			
Violin_Mode																				95		
Wandering_Line																				1	8	
Whistle																				1		60

CNNs as a Feature-Extractors

Original CNN (VGG-16)

relu4_3	Ramp	3–tensor (size: 512 × 28 × 28)
pool4	PoolingLayer	3–tensor (size: 512 × 14 × 14)
conv5_1	ConvolutionLayer	3–tensor (size: 512 × 14 × 14)
relu5_1	Ramp	3–tensor (size: 512 × 14 × 14)
conv5_2	ConvolutionLayer	3–tensor (size: 512 × 14 × 14)
relu5_2	Ramp	3–tensor (size: 512 × 14 × 14)
conv5_3	ConvolutionLayer	3–tensor (size: 512 × 14 × 14)
relu5_3	Ramp	3–tensor (size: 512 × 14 × 14)
pool5	PoolingLayer	3–tensor (size: 512 × 7 × 7)
flatten_0	FlattenLayer	vector (size: 25088)
fc6	LinearLayer	vector (size: 4096)
relu6	Ramp	vector (size: 4096)
drop6	DropoutLayer	vector (size: 4096)
fc7	LinearLayer	vector (size: 4096)
relu7	Ramp	vector (size: 4096)
drop7	DropoutLayer	vector (size: 4096)
fc8	LinearLayer	vector (size: 22)
prob	SoftmaxLayer	vector (size: 22)
	Output	class

Truncated CNN

relu4_3	Ramp	3–tensor (size: 512 × 28 × 28)
pool4	PoolingLayer	3–tensor (size: 512 × 14 × 14)
conv5_1	ConvolutionLayer	3–tensor (size: 512 × 14 × 14)
relu5_1	Ramp	3–tensor (size: 512 × 14 × 14)
conv5_2	ConvolutionLayer	3–tensor (size: 512 × 14 × 14)
relu5_2	Ramp	3–tensor (size: 512 × 14 × 14)
conv5_3	ConvolutionLayer	3–tensor (size: 512 × 14 × 14)
relu5_3	Ramp	3–tensor (size: 512 × 14 × 14)
pool5	PoolingLayer	3–tensor (size: 512 × 7 × 7)
flatten_0	FlattenLayer	vector (size: 25088)
fc6	LinearLayer	vector (size: 4096)
relu6	Ramp	vector (size: 4096)
drop6	DropoutLayer	vector (size: 4096)
fc7	LinearLayer	vector (size: 4096)
	Output	vector (size: 4096)

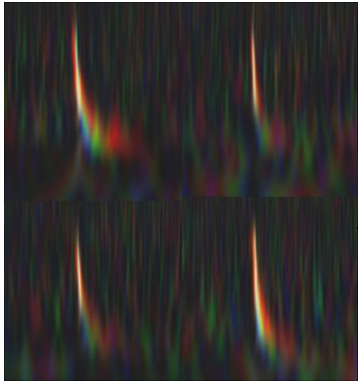
Output: Feature Vector

Visualizing Clusters of New Classes

After transfer learning, our truncated CNNs can automatically cluster new classes not seen before.

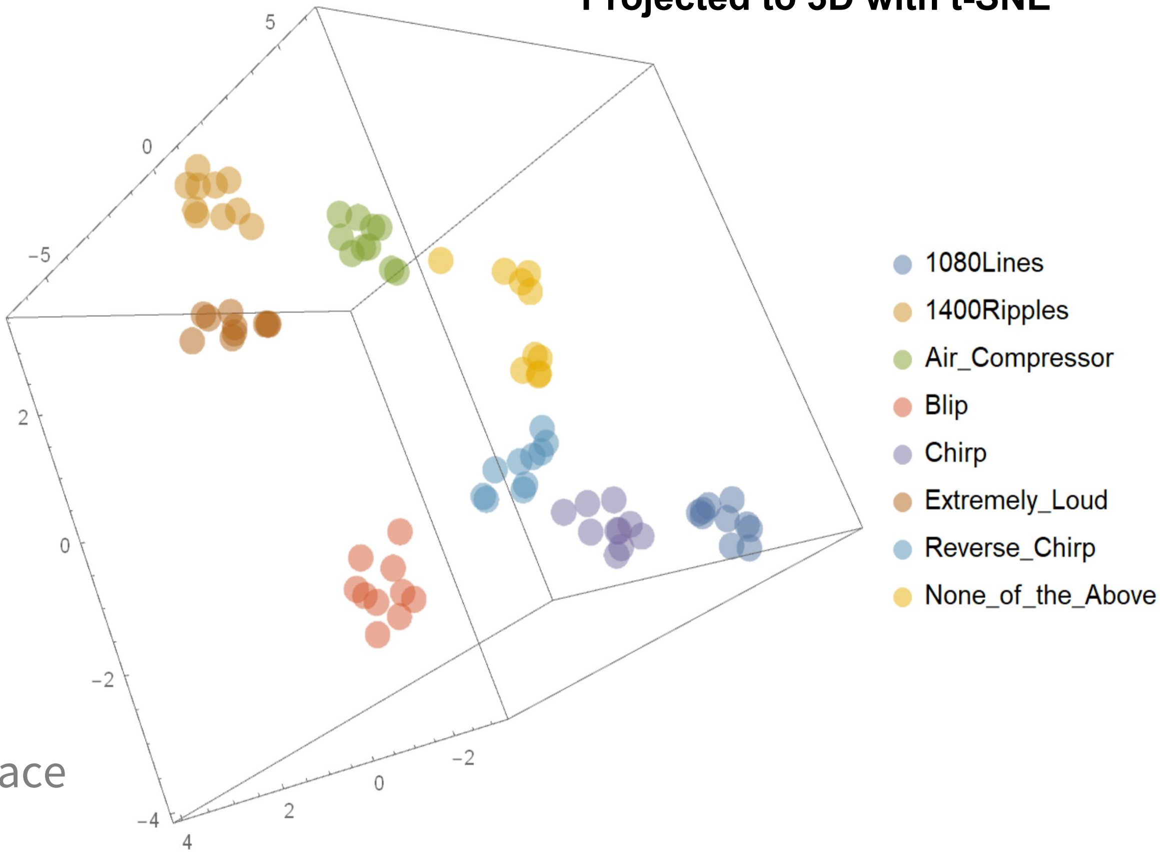
Synthetic class:

Reverse_Chirp



Glitches are organized according to their morphologies in this feature-space

Projected to 3D with t-SNE



Denoising

Denoising Gravitational Waves with Recurrent Neural Networks

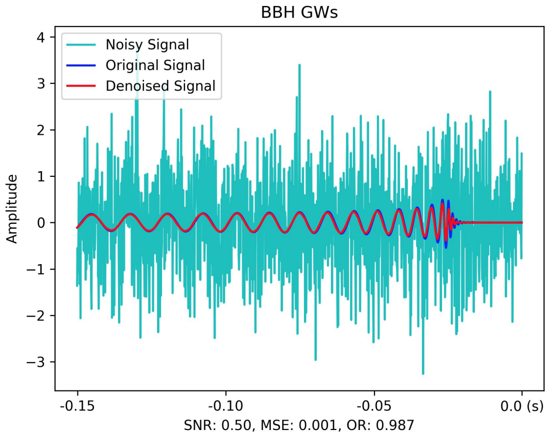
arXiv:1711.09919

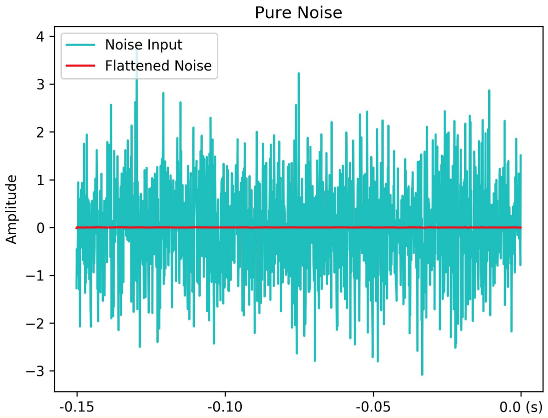
Hongyu Shen, Daniel George, E.A. Huerta, Zhizhen Zhao

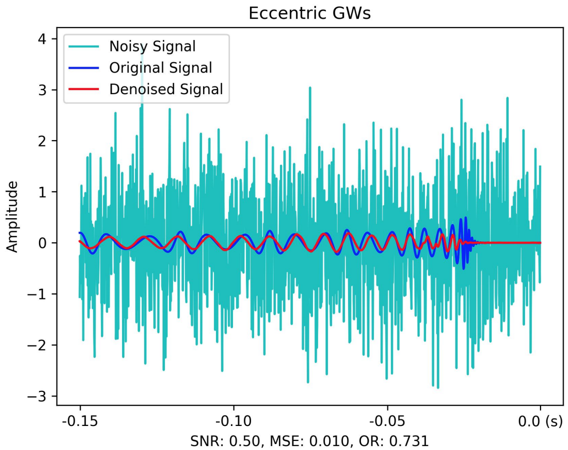
NCSA Gravity Group - http://gravity.ncsa.illinois.edu/
Department of Statistics, Department of Astronomy,
Department of ECE and Coordinated Science Lab, UIUC

Results on Real Noise

1. At very low SNR, the network can still recover the shape of the true signal
2. Pure noise input is denoised to a flat line (red in the middle).
3. The network is capable of denoising new types of waveforms (right).







Questions?