Convolutional Neural Networks for Gravity Wave Detection

Matthew Wise, Dr. Scott Field, and Dr. Gaurav Khanna

University of Massachusetts Dartmouth

UMass | Dartmouth

Introduction

Gravity waves (GW), the small perturbances in the curvature of space caused by quick, regular motion of large masses such as merging neutron stars and black holes provide a new window into astronomy. Just as complimenting visual light astronomy with x-ray and radio telescopes revealed important new details about the cosmos, the observation of gravity waves, the first non-electromagnetic source of information about distant objects, stands to contribute much to our understanding of the universe.

However, detecting gravity waves has proven difficult. Very low signal to noise ratios (SNRs) make traditional means of signal detection either too inaccurate or too slow to handle the data steaming in from GW sensors such as aLIGO. Our hope is that the convolutional neural network (CNN), a structure that has proven quite successful in areas like semantic image processing [2] and speech recognition [3], will be able to quickly and accurately detect these signals. Past work has given some reason to be optimistic about this approach [1].

Objectives

- Apply deep learning principles to the problem of GW detection.
- Design and train a CNN capable of quickly and accurately detecting GW signals.
- Determine the relative strengths and weaknesses of this method relative to traditional signal processing techniques (e.g. Matched Filtering)
- Determine if further information about the binary can be extracted from a trained network.

Methods

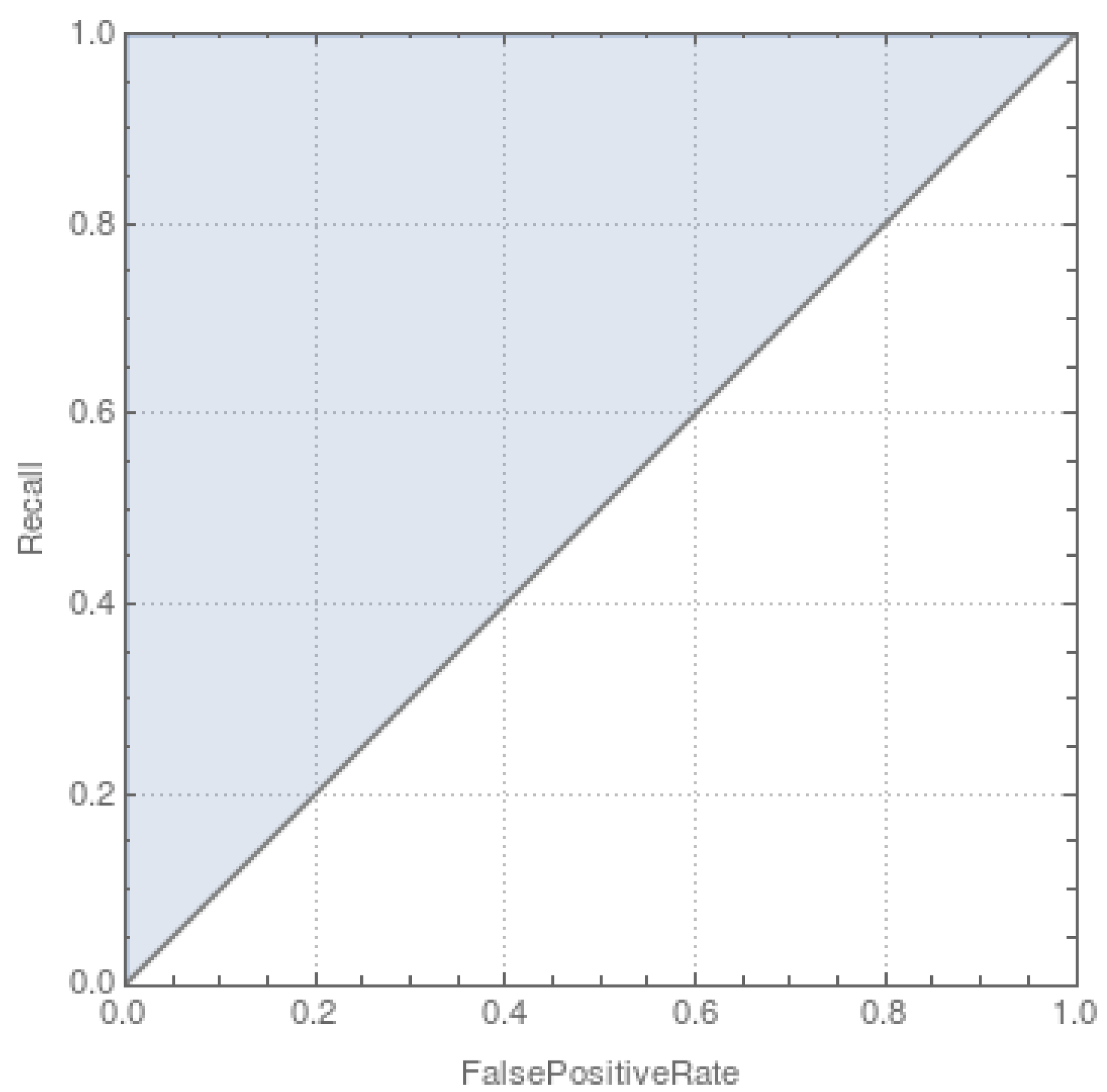
Data approximating potential real sensor data for different binary black hole mergers with Gaussian noise was generated in the form of normalized strain time series data for a period of one second sampled at 8192 Hz. 5000 points were generated, split evenly between genuine signals and random noise. Distance to the merger was set in the model so as to generate an SNR of about 24, similar to expectations for real events. This data was then split into training, verification, and testing sets at an 18:1:1 ratio.

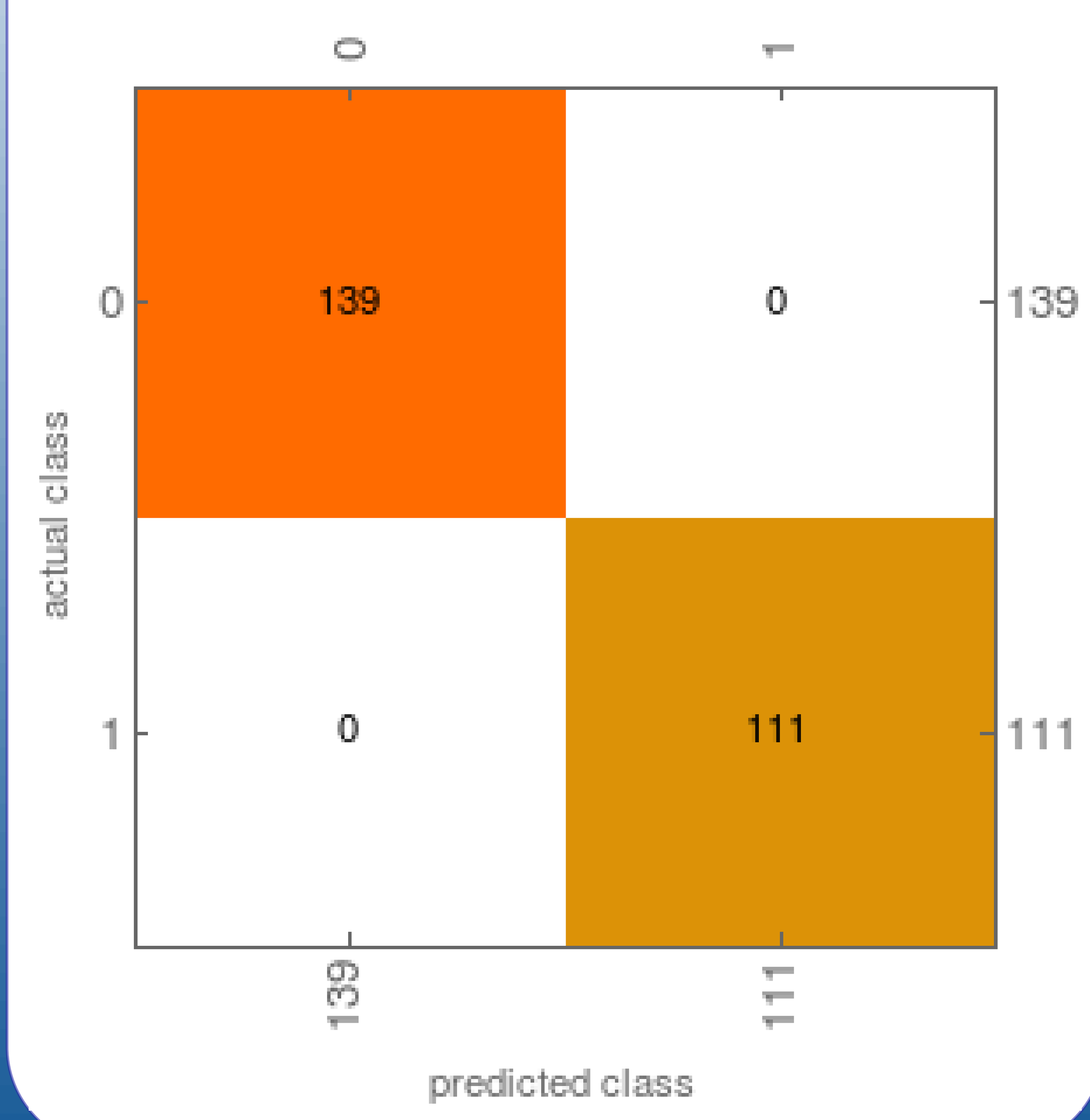
Following in the footsteps of past authors [1], we designed a CNN with three convolutional layers followed by two linear layers. The full architecture is included below. Other hyper parameters, including convolution kernel size, stride, and dilation, pooling layer stride, network learning rate and learning rate decay, and so on were set according to the optimization done in [1].

	Input	vector (size: 8192)
1	ReshapeLayer	3-tensor (size: 1×1×8192)
2	ConvolutionLayer	3-tensor (size: 16×1×8177)
3	PoolingLayer	3-tensor (size: 16×1×2045)
4	Ramp	3-tensor (size: 16×1×2045)
5	ConvolutionLayer	3-tensor (size: 32×1×2017)
6	PoolingLayer	3-tensor (size: 32×1×505)
7	Ramp	3-tensor (size: 32×1×505)
8	ConvolutionLayer	3-tensor (size: 64×1×477)
9	PoolingLayer	3-tensor (size: 64×1×120)
10	Ramp	3-tensor (size: 64×1×120)
11	FlattenLayer	vector (size: 7680)
12	LinearLayer	vector (size: 64)
13	Ramp	vector (size: 64)
14	LinearLayer	vector (size: 2)
15	SoftmaxLayer	vector (size: 2)
	Output	class

Results

For an SNR of 24, after 500 epochs of training on the 4500 point training set (taking about 80 hours), the network was able to correctly classify all 250 test points.





Conclusions

These results suggest CNN detection may be a very fruitful approach to GW detection. While an SNR of 24 is not very low, and would certainly be detected by traditional approaches (e.g. matched filtering), complete success for a comparatively unexplored approach is very encouraging. While present matched filtering approaches have been highly optimized for the problem space, the application of deep learning methods is novel. Further research is warranted to explore whether further optimization of this approach may allow it to exceed the quality of matched filtering.

Even if that proves not to be the case, other means of detection still provide a valuable service. It is likely that by combining multiple templates into one data set for CNN training, the time required to find a signal, or at least tag potentially interesting areas for a more thorough follow up, could be dramatically reduced.

Alternatively, often superior accuracy to the best single approach to a problem may be achieved via the combination of several methods of varying individual quality. This works on the theory that while certain features are better indicators than others, the sum of the features provides a better indicator than any individual one.

References

[1] George, D., & Huerta, E. A. (2016). Deep Neural Networks to Enable Real-time Multimessenger Astrophysics. *arXiv preprint arXiv:1701.00008*.

[2] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[3] Abdel-Hamid, O., Mohamed, A. R., Jiang, H., & Penn, G. (2012, March). Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on* (pp. 4277-4280). IEEE.