Classifying Non-Gaussian Transient Noise in LIGO

By Maanya Malhotra

Mentored by Anthony Cuturrufo

Abstract: Space exploration methods have never been as diverse and thorough as it is in today's time. New techniques that span far beyond the conventional method of using electromagnetic radiation to identify and study astronomical evidence have been discovered and are now amplified using machine learning technology. One such technique is using interferometers to record gravitational waves released from astronomical events. Using computer vision to analyse gravitational waves can provide valuable information about previously obscured information. Also, gravitational waves are fundamentally unrelated to electromagnetic radiation thereby providing us with information that is entirely different from what we previously could access. This has consequentially opened an altogether new avenue for research and further exploration. The contrastive model for machine learning-based detection and classification performed the best.

Introduction:

The measurements made by LIGO operate on a very small order of magnitude. Without elaborating much on the working of LIGO, it can be safely said that the typical variation is about 10^{-18} m, thereby making it one of the smallest detections made. With all other disturbances such as electric storms and seismic events, the detector must be highly sensitive and simultaneously capable of identifying the noise from the actual data. By using machine learning models, the transient non-gaussian noise can be detected and separated to improve our analysis of the gravitational waves recorded as well as the stability of the detector. (LIGO, n.d.)

The story does not end at detection. To eliminate the possibility of deteriorating stability the detections must be classified as well. This is where harnessing machine learning will provide us with an advantage. By using techniques such as CNNs, Vision transformers, and Contrastive learning we can classify the Non-Gaussian noise so that the detector can provide us with a more accurate representation of the gravitational waves.

While this research considers a variety of methods, contrastive learning is predicted to outperform the others. This is primarily because contrastive learning involves "training a model to differentiate between similar and dissimilar pairs of data points by maximising their similarity within the same class and minimising it between classes." (BuiltIn) Since the whole experiment and the subsequent analysis are highly reliant on the detections being accurate, the importance of proper noise classification cannot be overstated. Additionally, as previously stated, we only have access to a very small amount of data and therefore drawing parallels between existing data to improve the performance of the model works as a sufficiently successful strategy.

Background:

It was Albert Einstein who first predicted the existence of gravitational waves in 1916 in his paper on the Theory of Relativity. Described as a ripple in space-time produced by the highly energetic processes of the universe, gravitational waves originate from massive celestial bodies accelerating in a manner that disrupts space-time. They are believed to travel at the speed of light, while simultaneously carrying valuable information about both their origin and source as well as the nature of gravity itself.

Cataclysmic events are known to produce the strongest gravitational waves. While these events are violently energetic themselves, their gravitational waves are considerably weaker by the time they reach Earth. Other predictions about sources of gravitational waves include the rotation of neutron stars and the remaining energy from the Big Bang.

To explore these very gravitational waves, LIGO (Laser Interferometric Gravitational-Wave Observatory) was set up in the United States of America, across 2 cities: Hanford (Washington) and Livingston (Louisiana).

Up until 2015, the evidence for the existence of gravitational waves was only theoretical and mathematical, but then the first detection was made.

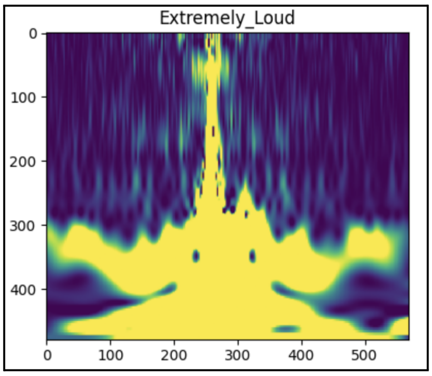
There are several sources of disturbances which can introduce errors such as seismic events, electromagnetic disturbances etc. Some of the ones considered in this dataset are power lines and low-frequency bursts.

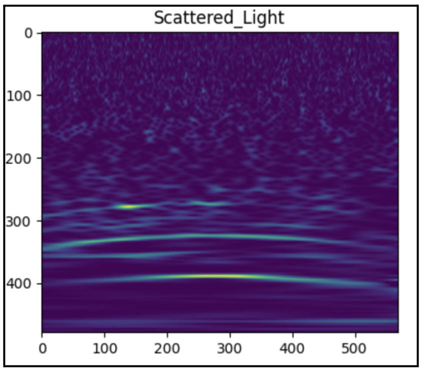
The methods I have used to classify these are as follows:

1. Vision Transformers: A ViT represents images as a sequence with class labels allowing the model to learn image structure individually by splitting it into patches and projecting it into the desired input dimension. (PapersWithCode, n.d.)
2. Convolutional Neural Network: This refers to a deep learning architecture that essentially learns directly from the data. Their specific strength is their ability to identify patterns. (Kundu, 2022)
3. ResNext101: This is a model based on the ResNet101which is trained with mixed precision using Tensor cores. (PyTorch, n.d.)
4. Contrastive Learning: Contrastive Learning is a form of self-supervised learning that has gained popularity due to it not requiring data annotation. Specifically used in computer vision and natural language processing, contrastive learning models learn by comparing two data points. This makes them highly efficient, as well as highly resistant to overfitting. (Reynolds, 2023)

Dataset:

The dataset has 22 classes, and 4800 images in total, which have been annotated by enthusiasts all over the world. The semi-supervised learning model was chosen as the entire dataset is not fully annotated. The dataset was taken from Kaggle. (Harrand, 2023)Some sample images from the dataset are as follows:





The dataset was already split into train and test data with an 80-20 ratio to ensure that the model is well trained without the risk of overfitting.

Methodology:

All models had the same parameters as they did in their respective papers. The contrastive models specifically used the following augmentations: RandGaussianNoise, RandAffine and RandAdjustContrast among others.

This is a summary of my CNN model:

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 149, 149, 32)      896

 max_pooling2d (MaxPooling2  (None, 74, 74, 32)        0
 D)

 conv2d_1 (Conv2D)           (None, 36, 36, 32)        9248

 max_pooling2d_1 (MaxPoolin  (None, 18, 18, 32)        0
 g2D)

 flatten (Flatten)           (None, 10368)             0

 dense (Dense)               (None, 22)                228118

=================================================================
Total params: 238262 (930.71 KB)
Trainable params: 238262 (930.71 KB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

The vision transformer, the image and the patch size used were 32 and 4 respectively.

Results:

Here is the summary of the models:

Model	Batch Size	Best Validation Accuracy	Best Training Accuracy
CNN	12	0.916	0.923
CNN	24	0.909	0.836
CNN	48	0.917	0.951
ViT	32	0.920	0.931
Contrastive	5	0.981	0.961

Discussion:

The contrastive learning model outperformed the other models, as projected previously. One key limitation of this comparative study is the variation in batch size and number of trials primarily due to the shortcomings of the coding environment used, i.e. Google Colab. Herein, the large dataset size, coupled with the complexity of the models made the code execution highly time and memory-consuming. Due to a lack of unlimited access to GPUs and logistical limitations, it was not possible to execute the same batch size for each model, however, the number of epochs remained consistently the same.

Conclusion:

The contrastive learning model is highly efficient but works best with small batch sizes, however, the convolutional network, despite not being nearly as efficient, is more capable of handling larger batch sizes. The research however does not end here. With more in-depth work and access to better coding environments, research for larger batch sizes can be executed. Furthermore, for the CNN model, different activation functions and different numbers of layers can be introduced. In the Contrastive model, more kinds of data augmentations can be added to prevent overfitting.

Acknowledgements:

I would like to thank my mentor Anthony Cuturrufo for all his guidance and support throughout this journey. He has helped me understand concepts such as contrastive learning which formed the backbone for this project.

I would also like to thank Inspirit AI for providing me with the opportunity to work on this interesting project. The Inspirit Team's support was instrumental in the execution and completion of this project.

References

Harrand, R. (2023, June 16). *Gravity Spy (Gravitational Waves)*. Kaggle. https://www.kaggle.com/datasets/tentotheminus9/gravity-spy-gravitational-waves

Kundu, R. (2022, May 22). *The Beginner's Guide to Contrastive Learning*. V7 Labs. Retrieved November 15, 2023, from https://www.v7labs.com/blog/contrastive-learning-guide

LIGO. (n.d.). *Facts | LIGO Lab | Caltech*. LIGO Caltech. Retrieved November 15, 2023, from https://www.ligo.caltech.edu/page/facts

LIGO. (n.d.). *Sources and Types of Gravitational Waves | LIGO Lab | Caltech*. LIGO Caltech. Retrieved November 15, 2023, from https://www.ligo.caltech.edu/page/gw-sources

LIGO. (n.d.). *What are Gravitational Waves? | LIGO Lab | Caltech*. LIGO Caltech. Retrieved November 15, 2023, from https://www.ligo.caltech.edu/page/what-are-gw

PapersWithCode. (n.d.). *Vision Transformer Explained*. Papers With Code. Retrieved November 15, 2023, from https://paperswithcode.com/method/vision-transformer

PyTorch. (n.d.). *ResNeXt101*. PyTorch. Retrieved November 15, 2023, from https://pytorch.org/hub/nvidia_deeplearningexamples_resnext/

Reynolds, M. (2023, April 4). *Contrastive Learning: A Tutorial*. Built In. Retrieved November 15, 2023, from https://builtin.com/machine-learning/contrastive-learning