Applying Machine Learning to Accelerate Cosmological Simulations

Patryk Drozd, Dr John Brennan
Department of Theoretical Physics, Maynooth University

Introduction and aims

Today, physicists heavily rely on cosmological models, which involve modelling how the universe evolves. Ideally these simulations would involve as many parameters as possible but they would take an extraordinary amount of time to run. We instead sacrifice some accuracy by omitting phenomena that we deem negligible in relation to what we want to study which significantly reduce the runtime for our simulations. With the aid of machine learning we can even further decrease the runtime by a meaningful amount for our simulations and also keep a significant amount of accuracy.

Why Machine learning

There is a field in computer science known as image processing. This field involves extracting and transforming images to get useful information out of them. Image processing involves problems like image classification and upscaling. Quite often machine learning is used to aid this task. In cosmological simulations properties of space are stored similarly to values in images. In this project I use this key observation to translate solved problems in image processing into solutions in improving cosmological simulations, where I train models, which serve us as universal approximators, to reconstruct data.

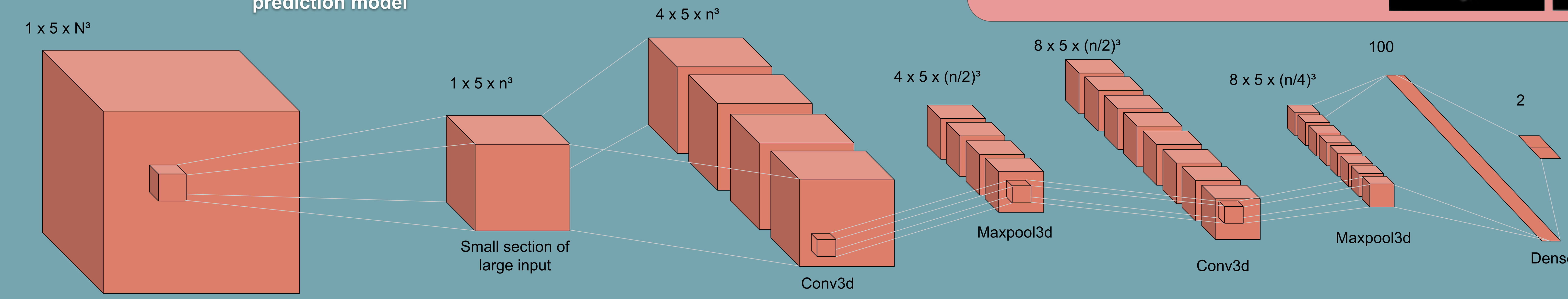
Predicting Star Formation with Machine Learning model trained on random data

A key feature of cosmological simulations is to insert features like stars, or black holes into simulations when certain criteria are met in a region of space. This classically involves checking each subdivision of the space we are considering, one by one. This process is painfully slow for large samples. To reduce the time to check these criteria, I worked on training a model on a randomly generated dataset. With this dataset we could bypass the time to create a realistic dataset with time to generate a randomly generated dataset.

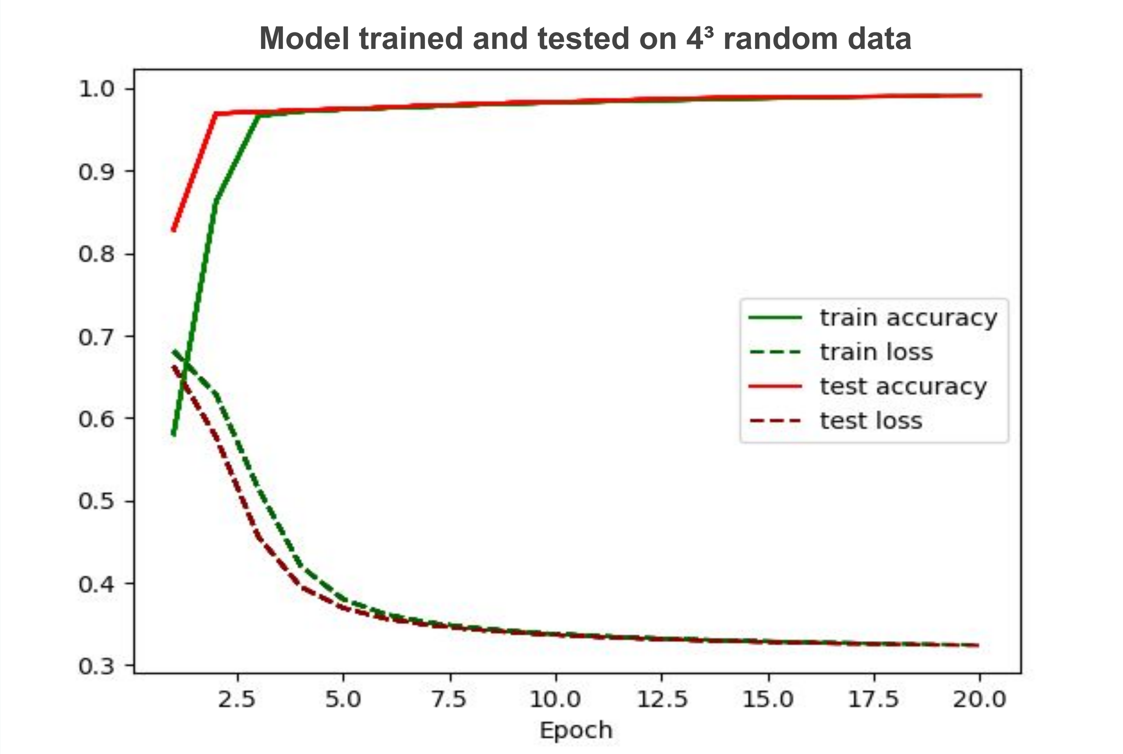
Image classification in computer science

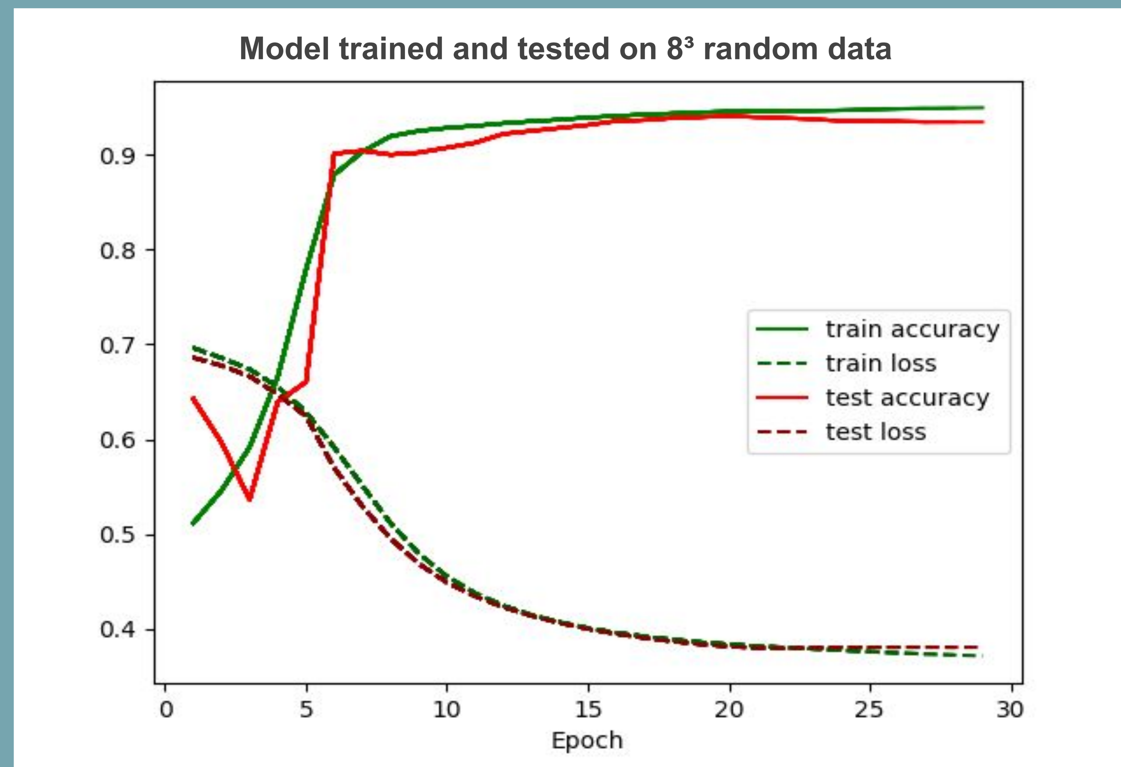
Image classification is a classical application of machine learning. To the right are some hand drawn characters; numbers and hiragana characters. These classification models would learn to tell you what character is drawn.

Architecture for star formation prediction model



I trained my model on small samples of size 4^3 and 8^3 to have a very fast training time. These models were able to classify more than 95% of randomly generated test data correctly. The plot to the right shows the training process and accuracy of each model per epoch, where one epoch represents one loop of training. There are two datasets considered in each graph, the training and test datasets. For each we evaluate the accuracy, and loss which is a function with which we minimise to train the model. To adapt to larger datasets of any size, this model has two parts to it, a sliding window and the neural network itself. Once the neural network is trained I can apply it to small regions of the large input successively and classify the large input.





Reconstruction of simulation data with Machine Learning model

Architecture for reconstruction model

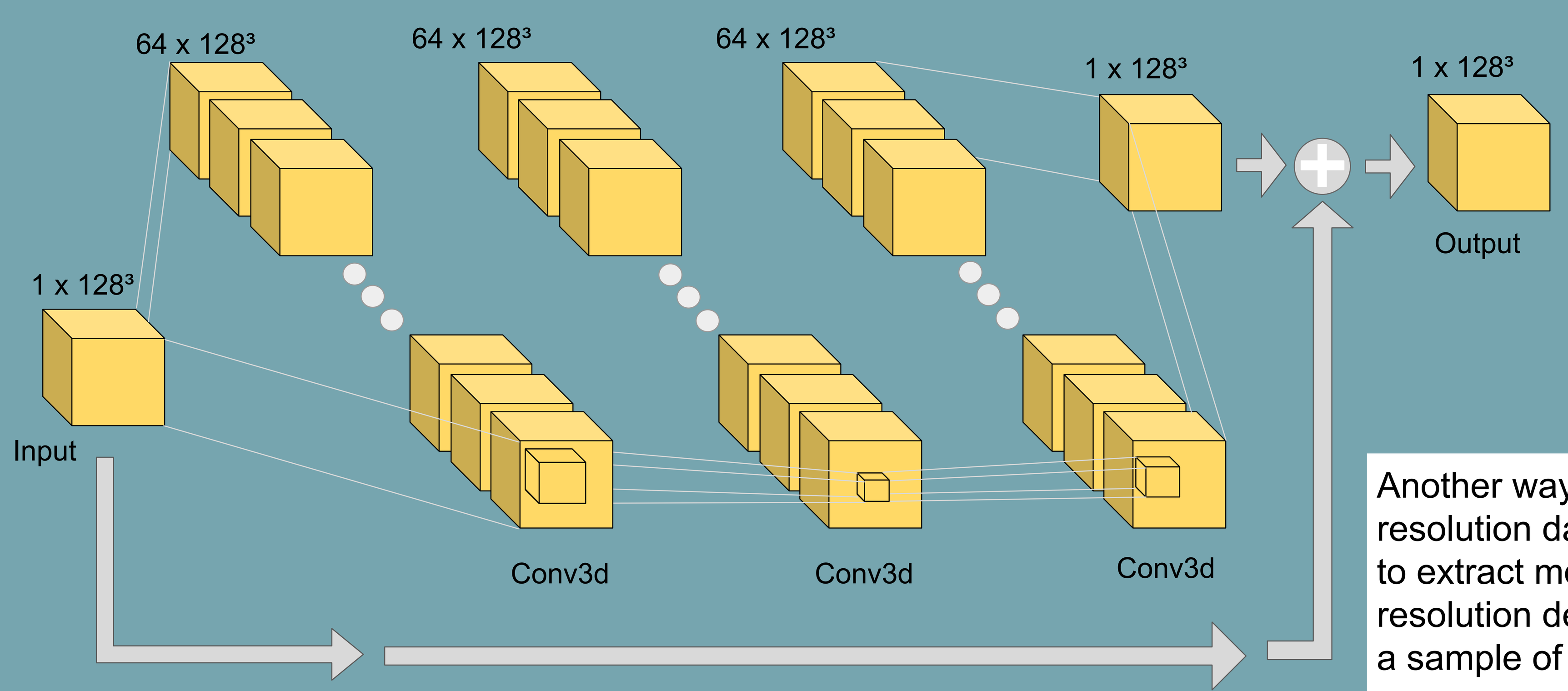
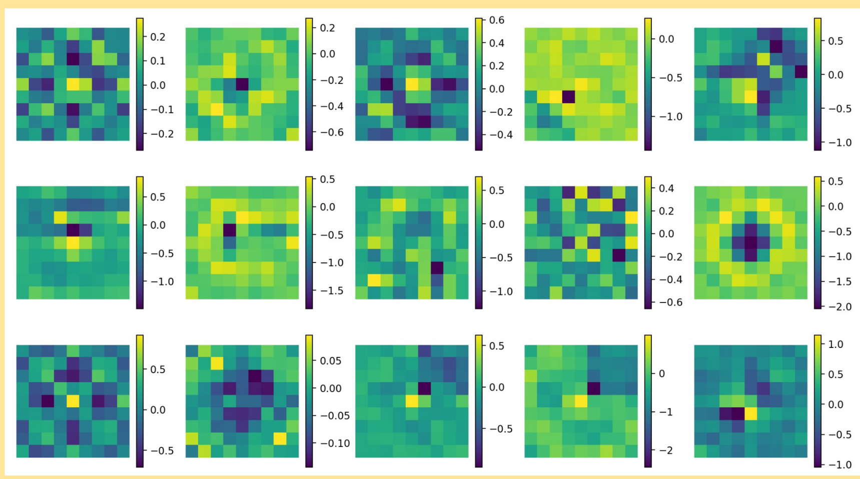
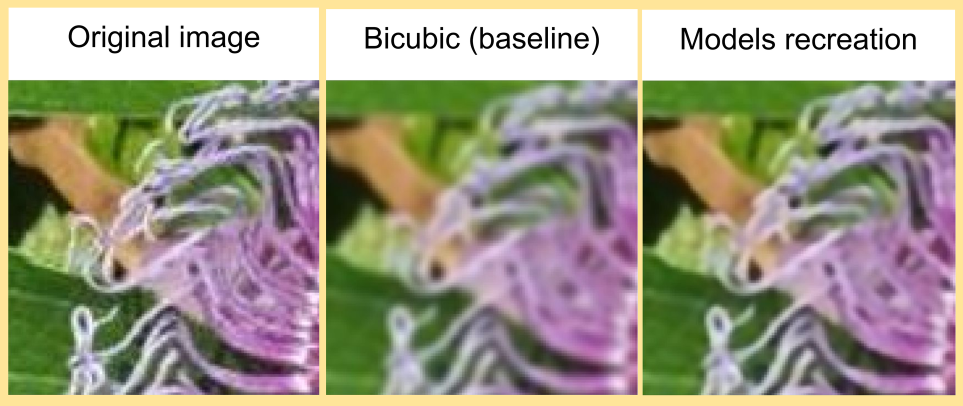


Image reconstruction in computer science

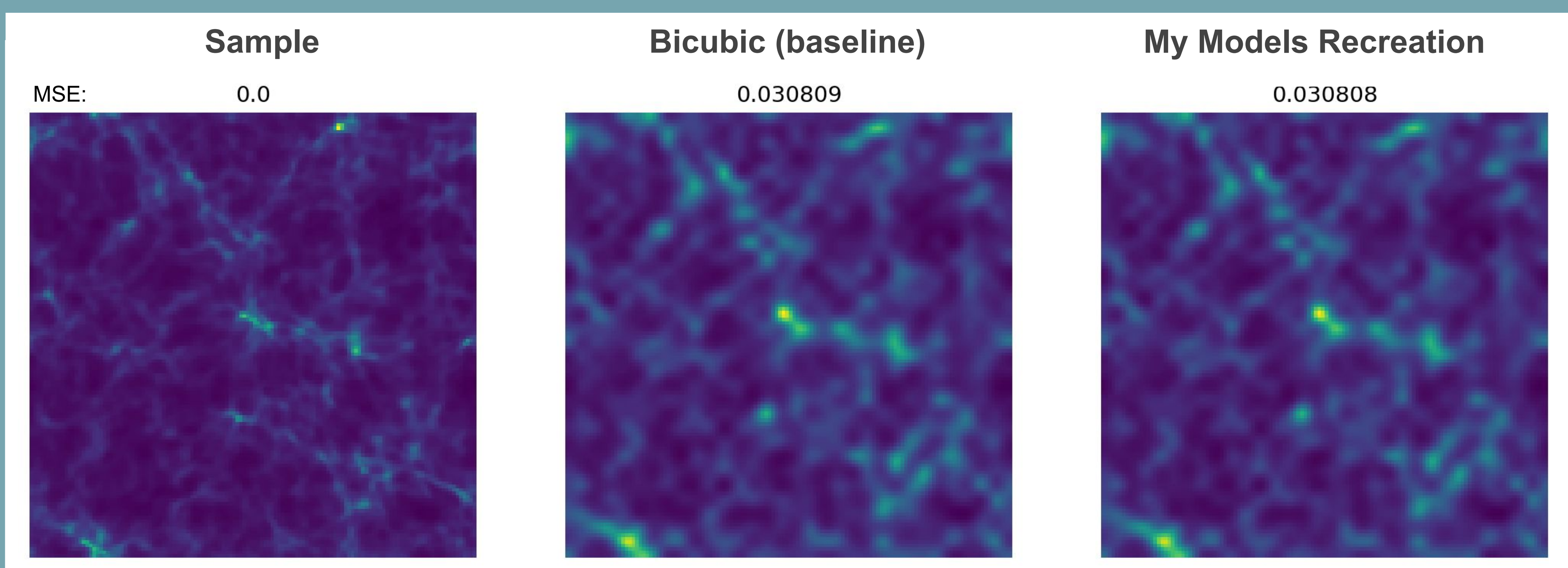


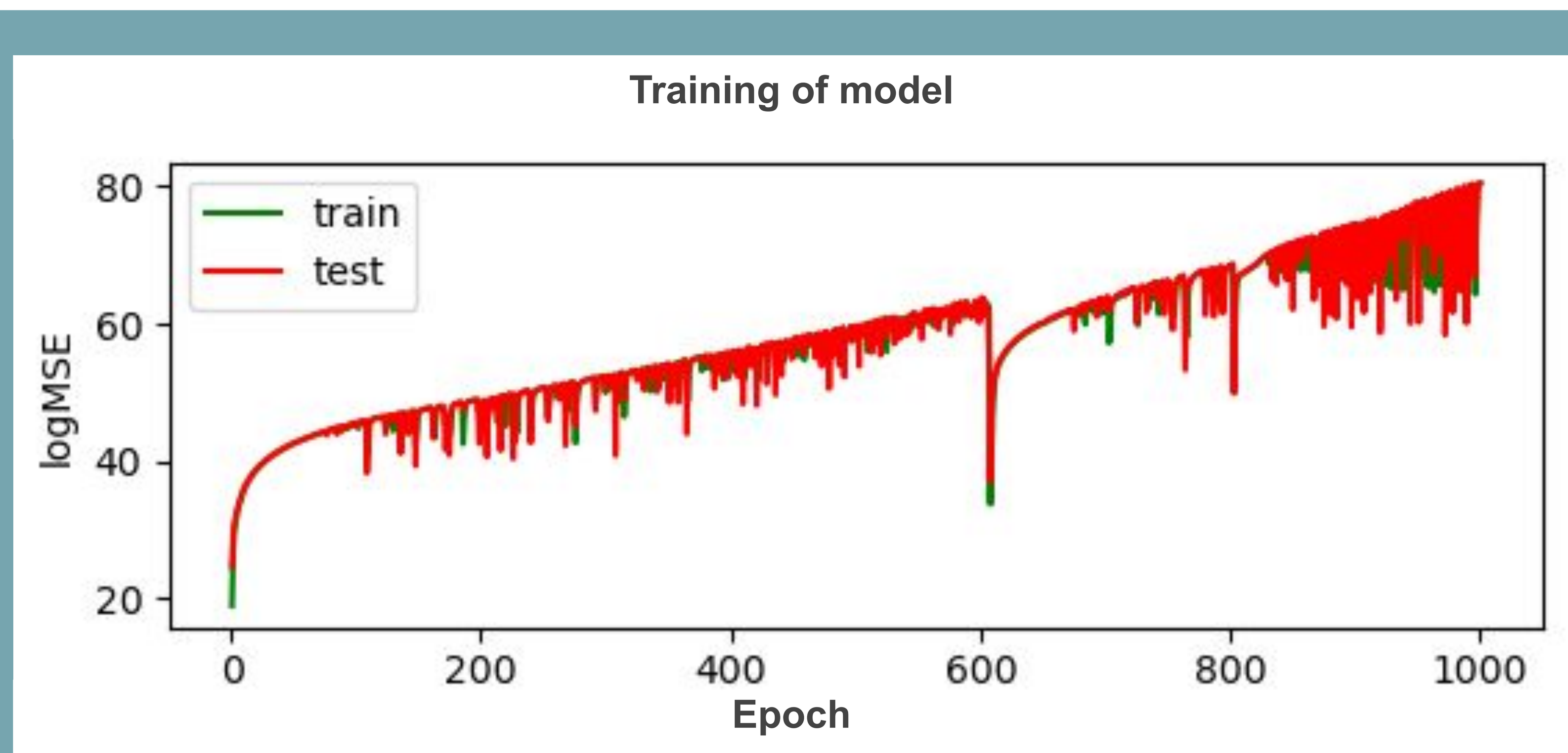
Filters of network trained to reconstruct images of flowers

To apply this technology to physics I first worked on trying to reconstruct images of flowers. Below is a comparison of the same image through two reconstruction methods and to the right are some of the models filters.



Another way we can use machine learning to our advantage is to use it to reconstruct high resolution data from low resolution data. The idea is that we could run low resolution cosmological simulations and use this reconstruction model to extract more detailed information from a volume of space when we need it. I applied this model to reconstruct a high resolution density field from a low resolution density field. The plots below show the training of the model and its result on a sample of space. Increase in logMSE means improvement of model and one epoch represents one loop of training.





Source code on Github

Conclusion

Machine learning has proven to be an excellent candidate in accelerating cosmological simulations. I have shown that image classification and image reconstruction problems can be translated into solutions for classifying star forming regions and reconstructing density fields respectively. This work forms a basis for a proof of concept paper currently being written by Dr John Brennan.

Acknowledgements

I would like to thank Dr John Brennan for being an excellent mentor and supervisor for this SPUR project. His interest in this field has been contagious and his experience invaluable. I would also like to thank Dr John Regan for his advice and guidance during this project.

MU SUMMER PROGRAMME FOR UNDERGRADUATE RESEARCH SPUR

Ollscoil Mhá Nuad
Ollscoil na hÉireann Má Nuad