

MAIS 202 Final Project Deliverable 1

Stephen Fay

October 5, 2019

1 Physics

‘In cosmology reionization is the process that caused the matter in the universe to reionize after the lapse of the “dark ages”’. Today there are several candidate models that could describe the manner in which re-ionization took place at this time in the universe; however we have not yet been able to distinguish between them using the data we currently have. New observations will be made at some point in the future but there is still alot to be done before-hand.

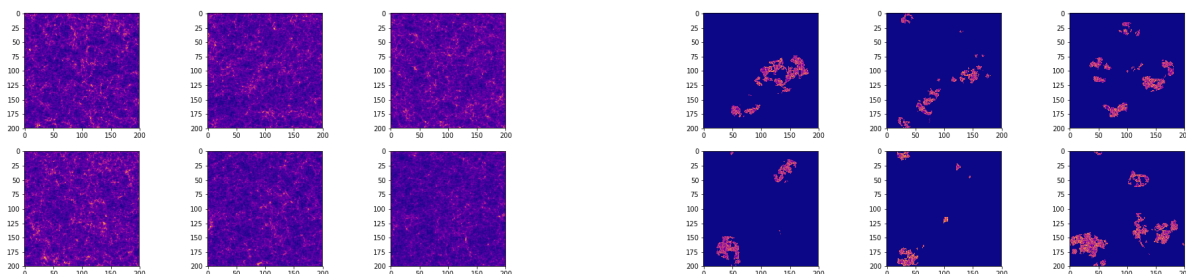
2 Dataset

Build a CNN that maps the mass density field to the ionization bubble maps.

The simulation built in [this Santos et. al. 2010 paper](#) was improved in 2017. A classifier was neural network classifier was built to distinguish between two maps generated by simulations that assumed different models for re-ionisation, [see this paper - Reionization models classifier using 21cm map deep learning](#).

The input data for our model is a 3d array of floating point numbers, and that out-puts are also a 3-d array of floating point numbers. The implementation of this wouldn't be too dissimilar to artistic style transfer.

The goal is to turn the LHS into the RHS.



3 Methodology

3.1 Preprocessing

The data-set I was given is roughly 100 gigabytes of data, but not all of it is useful so my first task will be to free some space on my disk download only what I need. The files are all named and classified according to red-shift (distance). Because we are working mainly with 2D cross sections of 3D arrays I can slice them in 3 different directions to get not just 200 cross sections per block but $200 \cdot 3 = 600$ cross sections per block. We can safely assume that even though the block represent chunks of the universe 300 mega parsecs long the change in the red-shift value is negligible.

3.2 Machine Learning Model

This project will attempt to implement a convolutional neural network based off of [Sam Gagnon's](#) research project. Doing so will make working with the data-set much easier as he uses the same data-set.

3.3 Final Conception

My goal is to make something that can guess with a high degree of accuracy what data will look like, the final product will be a program that, based off an input density mapping, guesses what the ionization map ought to look like like.

4 References

1. <https://academic.oup.com/mnras/article-abstract/483/2/2524/5228756?redirectedFrom=fulltext>
- a free pdf of this article can be found here <https://arxiv.org/pdf/1807.03317.pdf>
2. I plan to make use of Samuel Gagnon's research project as a resource (uses same / similar? data that I use)
<https://github.com/samgagnon/remove-wedge>