Accelerating The Search For Intelligent Life Beyond Earth Using Deep Learning

Laidlaw Proposal

[Peter] Xiangyuan Ma

The Proposal

[Peter] Xiangyuan Ma

May 2021

Introduction

The search for life beyond Earth is a monumental scientific endeavour. The question of "are we alone in the universe" has been asked by generations of curious minds before us. Today, the search for signs of intelligent life beyond Earth, known as SETI, is a field within astronomy with the goal of searching our universe for signs of technological signatures. This international research effort has brought together astronomers and telescopes across the globe from the Green Bank Telescope in the United States to the MeerKAT telescope in South Africa. Despite the impressive advancements in the recent years with the rise of high performance computing, the current leading problem in SETI deals with human-made Radio Frequency Interference (RFI), which often results in tens of thousands of false positive detections. [1]

With this problem in mind, the promise of modern computing algorithms, such as deep learning, have the potential to outperform the current search techniques in rejecting false positive detections. In essence, these algorithms could provide a fundamentally new way to examine signals detected qualitatively and assess their similarity to known interferes or their uniqueness. Few deep learning techniques have yet to be implemented in production [2], sharpening the basic question: ***"can deep learning algorithms improve the search for intelligent life beyond Earth?"*** The proposed research goal is to develop a novel deep learning algorithm and benchmark it against the existing algorithms.

Personally, like all people before me, I have always wondered: are we alone in the universe? Although I grew up being told that this was a uninteresting question to ask, my childhood excitement was reignited when I discovered SETI, a leading scientific effort in answering this very question. During my last year of high school I began working with the UC Berkeley SETI Research Center on building a hybrid cloud infrastructure pipeline to scale SETI searches across multiple compute nodes. Ever since then, this line of research has become exciting and fulfilling for me to work in.

[1] Paper on recent SETI surveys on the nearest 1327 stars in the galaxy https://seti.berkeley.edu/listen2019/BL1327stars.pdf

[2] Previous Deep Learning approaches Zhang et al https://arxiv.org/abs/1901.04636

Research Question: "can deep learning algorithms improve the search for intelligent life beyond Earth?"

Methodology

The methodology for the research proposal follows this tentative plan.

1. We create a testable metric/benchmark. This includes false alarm rate, true positive recovery rate and computational performance/efficiency. We will do so by simulating a range of events from satellites and ground-based RFI since they would characterize a desired detection. [2 weeks]
2. Brainstorm new ML algorithms currently implemented either in the commercial domain or other academic research fields and attempt to adopt them to SETI searches for identification of the events described in (1). [Ideally 1-2 different variants] [1 week]
3. Implement said algorithms. [2 weeks]
4. Take each algorithm (including existing algorithms) and benchmark it on our metrics for performance. [1 week]

End Goal: In completing the proposed project we will have theorized, implemented and tested novel deep learning algorithms used for SETI searches.

Context: The existing algorithms used for radio technosignature searches are called TurboSETI[3] and Energy Detection and are the two that this research project is hoping to compare against. The tentative metrics for each algorithm are, memory, runtime and recovery accuracy of which are determined through a series of simulated events.

Note: currently there is no plans dedicated training, the preparation for the research is planned to take place prior to the 6-8 week research period.

Response to COVID-19: The current plan is to complete the project remotely. Most of the interactions will be on zoom meetings, or communication platforms like Slack and the work is shared on Github. The project involves software solutions and thus is possible to be completed independently without being at the physical lab.

Research Advisors

My project has two co-advisors, Dr. Andrew Siemion who is the Director of the UC Berkeley SETI Research Center and Dr. Cherry Ng who is the Project Sci-

[3] The algorithm currently established in SETI literature https://ui.adsabs.harvard.edu/abs/2019ascl.soft06006E/abstract

entist for the Center's MeerKAT[4] and VLA[5] SETI programs and she is based at the University of Toronto. Dr. Cherry Ng will be working directly with me. Dr. Ng's role is to facilitate the success of this project. This entails weekly meetings, facilitates any logistical questions and mediating connections with other lab members. The advisor would also provide technical feedback regarding radio astronomy, signal processing, SETI search parameters etc. There may be collaborations on potential papers, reviews, posters and presentations as well. In general, the advisor will guide me in producing valuable science.

Besides my immediate supervisor, I will also be collaborating with the rest of the Breakthrough Listen Team which consists of other researchers, advisors, engineers and other interns. Their relation with this project is to help with integrate my work with the rest of the pipeline.

Interdisciplinary and International Focus

This research project helps connect radio astronomy research from observatories around the world. The findings from this research will be highly applicable to the SETI searches carried out around the world, including the MeerKAT and the VLA telescopes. The work I contribute will help the collaboration of multiple international observatories to conduct the largest search for life beyond Earth.

This work also involves many disciplines. This project merges together areas of data science along with radio astronomy and software engineering. Fundamentally this is a data science and computer science solution that solves a radio astronomy problem. The search for life in the universe, particularly communicative life, also has deep connections with anthropology, sociology, biology, philosophy and theology. All aspects will need to come into play to produce the final product and answer the research question

Outcomes

After answering the question of "can deep learning algorithms outperform classical solutions to our problem", if the answer is yes, then the outcome is having a production ready piece of software for SETI searches. The ideal performance benchmark is to have an algorithm run as fast or faster than real time observation. The outcome is to develop an algorithm in such a fashion that it can process approximately 50 terabytes of data per day. Following this, the direct consequence would be to deploy the software on the MeerKAT system to perform a search on 1 million nearby stars in our galaxy. In the end the research

[4]https://www.skatelescope.org/news/meerkat-breakthrough-partnership/
[5]https://www.seti.org/seti-institute-and-national-radio-astronomy-observatory- team-up-for-seti-science-at-very-large-array

would also provide valuable insights as to the weaknesses of existing algorithms and the areas of improvement.

The other potential outcome would be to implement the TurboSETI algorithm into production in the event where deep learning algorithms cannot out perform classical solutions.

Tentative Proposed Start Date:

The revised tentative start date is **July 1, 2021**. The reason being, my advisor and I believe the time in between the start dates would be useful in learning the material to conducting the research.

Location

During the time frame discussed above, I will be living in two locations, one being in **Toronto Ontario Canada**, and the second being **Vancouver, British Columbia Canada** . This will not impact my research in any way as all the work is conducted online and from home.

Work Cited

Margot, Jean-Luc. "The Radio Search for Technosignatures in the Decade 2020–2030." Astro2020 Science White Paper, 19 Mar. 2019, doi:https://arxiv.org/pdf/1903.05544.pdf.

Breakthrough Initiatives. Breakthrough Initiatives, breakthroughinitiatives.org/initiative/1.

Enriquez, Emilio, and Danny Price. "TurboSETI: Python-Based SETI Search Algorithm." NASA/ADS, Astrophysics Source Code Library, ui.adsabs.harvard.edu/abs/ 2019ascl.soft06006E/abstract.

McDonald, Rebecca. "SETI Institute and National Radio Astronomy Observatory Team Up for SETI Science at the Very Large Array." SETI Institute, Feb. 2020, www.seti.org/seti-institute-and-national-radio-astronomy-observatory-team-up-for-seti-science-at-very-large-array.

"New Partnership between SKA Precursor Telescope MeerKAT SETI Programme Breakthrough Listen." Public Website, 2 Oct. 2018, www.skatelescope.org/news/ meerkat-breakthrough-partnership/.

Price, Danny C., et al. "The Breakthrough Listen Search for Intelligent Life: Observations of 1327 Nearby Stars Over 1.10–3.45 GHz." The Astronomical Journal, vol. 159, no. 3, 2020, p. 86., doi:10.3847/1538-3881/ab65f1.

Wright, Jason. "NASA AND THE SEARCH FOR TECHNOSIGNATURES." NASA TECHNOSIGNATURES WORKSHOP REPORT, 28 Nov. 2018, pp. 1–5., doi:https://arxiv.org/ftp/arxiv/papers/1812/1812.08681.pdf.