Visual Generative AI with Selective Attention: A Spiking Neural Network Framework for High Dimensional Satellite Image

About NGA:

National Geospatial-Intelligence Agency (NGA) is a unique combination of intelligence agency and combat support agency. It is the world leader in timely, relevant, accurate and actionable geospatial intelligence (GEOINT). NGA enables the U.S. intelligence community and the Department of Defense to fulfill the president's national security priorities to protect the nation.

RESEARCH LOCATION: Springfield, VA

PROJECT DESCRIPTION:

Investigate the benefit of spiking neural networks in application to large vision-language models by modifying the architecture of the model to show the benefits of a spiking neural network in encoding and decoding multi-modal input, including text and image, and how the model performs when compared to the traditional CLIP algorithm. Furthermore, how well the model can deal with large data inputs, potentially 2000x2000 pixel or larger images, is a question of interest in remote sensing.

This project involves exploring converting an available open source vision-language model to demonstrate how a novel spiking neural network architecture will out-perform currently available models using traditional CPU and GPU architecture. A highly successful internship will leave behind a demonstrable model that can be executed on simulated spiking neuromorphic architecture and a report detailing how well this new architecture performed in comparison to the traditional architecture for remote sensing image feature recognition tasks.

The internship includes building a spiking neural network based visual language model to test electro-optic (EO) and multi-spectral images that are already in HPCMP. Replacing the traditional visual transformer (ViT) with selective attention based on dynamical visual transformer, where the tokenization is designed based on the prioritized task-relevant region. This means the model should omit the regions that don't have interesting objects and focus on interesting, predefined objects.

Fortunately, we already have papers and code to realize this algorithm using a state space model. We can start with an existing model and modify it for better performance. When the model is relatively mature, we will explore different prompts and benchmarks to compare it with other models like ChatGPT and Claude. HPCMP has multiple clusters that provide powerful GPUs which will allow us to test our algorithm on GPUs simulating a Neuromorphic architecture.

Activities:

1. A kick-off meeting regarding background, literature for review, example code, and system explanation for HPCMP.
2. Learn about HPCMP and get familiar with the data that is already made available on multiple clusters.
3. Get familiar with multiple visual language models.
4. Meet multiple times a week with the mentor to touch base and have a Q&A session so that they are familiar with the problem.

5. After three weeks, and under the guidance of a mentor, the intern will research new solutions better than the example code provided.
6. After five weeks, there should be a meeting to demonstrate preliminary results and start preparing the benchmark to compare with ChatGPT and Claude.
7. After eight weeks, there will be final presentation, and the mentor will provide feedback.
8. The intern should have good practice of version control and make the repository of code available all the time, so that I can assist if necessary.

ANTICIPATED START DATE:

May/June 2026 – Exact start dates will be determined at the time of selection and in coordination with the selected candidate.

QUALIFICATIONS:

The ideal candidate should be a high quality PhD student who has a background in both computer science and one of the basic sciences, like physics or mathematical biology. It is desired that the candidate has some basic knowledge of Geospatial information science (GIS), but strong coding and data science knowledge is enough. Skills in software development is also desired.

ACADEMIC LEVEL:

Degree received within the last 60 months or currently pursuing:

- Doctoral

DISCIPLINE NEEDED:

- Computer, Information, and Data Sciences
- Earth and Geosciences
- Mathematics & Statistics
- Physics