

Can you spot the difference?

Instruments like satellites and temperature sensors and rainwater gauges collect a ton of data, so much that often monitoring software turns it into a graphic because there are just too many numbers to spot a pattern. There are so many of these instruments though that collect so much data so quickly that there ends up being too many images for a human too. But, then how can we spot if an image should be flagged for follow up? Well, most of these instruments collect fairly stable data, so the graphs all should look mostly the same. Sudden spikes, even if not wrong, are of interest. In this project, we will use neural nets to compute similarity scores for a database of images that we will then test on a new set of images. The work will primarily be done using the Python programming language will utilize modern technologies like scalable analytics and distributed computing libraries, and possibly cloud-based web services.