

ASTR W3646

Observational Astronomy

Final Project

Experiments in Crowdsourced Astronomy with AstroBin

Rasmi Elasmr

Monday, December 21, 2015

Introduction

The contributions of amateur observers and citizen scientists in astronomy have become more meaningful in doing "real science" through both crowdsourced data analysis and telescope observations (Marshall 2015). There remains huge untapped potential in crowdsourcing amateur observations on the internet and extracting meaningful scientific data from them. Some laborious attempts tailored towards particular objects (e.g. comets) have proven successful (Lang 2012), but there remains work to be done in generalizing the crowdsourcing process and improving the way data is retrieved and consolidated. This project seeks to mine and process images from AstroBin, a website where amateur and professional astronomers can post their observations. AstroBin is likely to provide images of a higher quality and with more accurate metadata than more generalized web searches because of its user base. The API, which is liberal in its terms of use, allows for searching by object, date, and by telescope properties, which can allow for interesting applications in analyzing the data.

This project explores methods of processing images of varying quality and content in hopes of creating time-series visualizations of Solar System objects such as the Moon and Jupiter. This turned out to be a much greater challenge than was initially suspected; however, progress was made and lessons were learned throughout the process.

Process

Data Gathering

AstroBin provides a searchable API. Images can be tagged by Solar System object upon submission, but this isn't guaranteed and may exclude results. To ensure thoroughness, run two searches: one with "jupiter" in the title, and one for "jupiter" in the description. On 12/18/15, these returned 7,532 and 2,047 results, respectively, for a total of 9,579 results. After removing duplicates by ID, 7,875 images remain. Once the IDs and metadata are in place, downloading the images is a simple but time-consuming and storage-intensive process.

On the first attempt, 7,867 images were downloaded and 8 failed. After a second attempt, 5 more images downloaded and 3 failed again, for a total of 7,872 images. Subsequent attempts yielded no new images.

Running similar searches for "moon" on 12/19/15 returned 9,551 and 4,812 results, respectively, for a total of 14,363 results. After removing duplicates by ID, 12,152 images remain. Two image downloads continued to fail after multiple attempts, leaving a total of 12,150 images.

Classification

Analysis

Results

Challenges

Future Ideas