In Search of Long-Period Comets with Deep Learning Tools

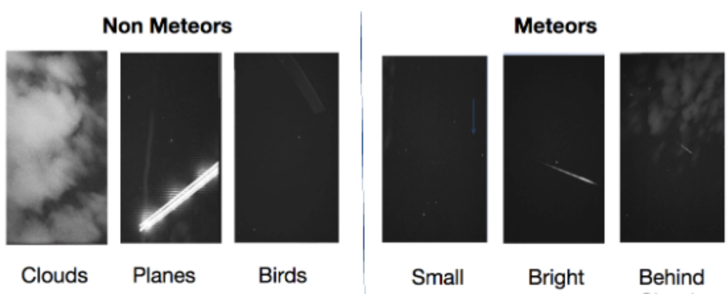
Due to their large size and fast traveling speeds of up to 70 km/s, long-period comets are recognized as potentially the most devastating impact threat to our planet, in the rare but possible event that they were to enter in a collision course with Earth. Evidence indicates that the impact of a comet or asteroid, having a diameter of about 10 km, was responsible for the mass extinction of most species of dinosaurs and marine life. However, any new comet on an impact trajectory with Earth would likely only be discovered about one year before impact, when it becomes visible as the Sun's heat and wind start sublimating its icy surface and ejecting rocky debris.

The aim of this paper is to provide deep learning tools to aid the search for debris of long-period comets. The orbits of such debris can be used by astronomers to guide the search for comets while they are still far out, providing us years of extra warning time in case of a collision path. Most suitable for this are the rare aperiodic meteor showers from debris ejected by a comet in a previous orbit. Our goal is to enable the continuous search for such showers.

The Cameras for Allsky Meteor Surveillance (or CAMS) is a network of low-light video cameras, established by the SETI[1] Institute in different locations across the globe, that monitors the sky to detect meteors. Until now, processing the images collected by CAMS has required large time-consuming human input. On an average night, an astronomer receives -per camera- around 500 detections consisting of images and light intensity curves (a sequence of measurements of how light intensity changes as detected objects move in the sky). A total of 8,000 observations with 16 cameras per site. Figure 1 presents examples of images captured by CAMS. Most of these turn out to be false detections, such as planes, birds, clouds, etc, and only about 15% are actually meteors. Sorting through these every night is not scalable. To alleviate this, we automate this process using deep learning. To the best of our knowledge, this is the first time that deep learning techniques have been applied to this endeavour.

Specifically, we trained a Convolutional Neural Network (CNN) that discerns images of meteors vs. other objects in the sky. We used five convolutional layers followed two fully connected layers and a binary softmax classifier. Dropout and max-pooling layers were used. We also performed standard data augmentation techniques, such as rotation and flipping, to alleviate the effects of imbalanced data. Our CNN achieves precision and recall scores of 88.3% and 90.3%, respectively. In addition, we trained a Long-Short Term Memory (LSTM) network that encodes the light curve tracklets into a latent space, and learns to predict whether or not the tracklet corresponds to a meteor. The LSTM achieves a precision of 90.0% and a recall of 89.1%. One key advantage of using deep learning is that we did not have to hand-engineer the meaningful features from both images and light curves. The models learned these on their own.

We did a qualitative evaluation by inspecting instances where the networks predictions are incorrect. False negatives often happen when meteors are very faint and hard to see. False positives tend to occur when there is an object (like a satellite) that looks very similar to meteors. It takes an average of 1.8s. to perform one forward pass of one image on an off-the shelf laptop[2] with no GPU. This makes it very suitable to be deployed on site, where the cameras are located. Compared to human performance, an expert at peak productivity can annotate about one image per second and achieve 99% precision and recall. However, it is not sustainable for an astronomer to perform this process every day. Ultimately the methods presented here free the astronomer from this low-cognitive task and help the search for debris of potentially dangerous long-period comets.



[1] Search for Extraterrestrial Intelligence

[2] 2.5Ghz Intel Core i7 processor, 16Gb of RAM.