**Introduction:** Lunar swirls have been recognized for many decades as distinct albedo features present within various regions on the Moon [1]. These features are defined as broad bright albedo areas separated by darker off-swirl lanes or terminating against the darker surrounding background. The formation mechanisms behind lunar swirls are numerous and long-debated [2-9]. Because the identification and mapping of swirls is done primarily by observation and is initially defined by albedo contrast, regions within swirls that gradually transition from bright to dark are often difficult to classify with certainty. Furthermore, and more importantly, the mapping is not quantitatively-defined. In order to address these issues, we have employed machine learning techniques to help assess swirl features and to map them based on actual data. Here, we focus on a small area in Reiner Gamma, the archetype example of lunar swirls (Figure 1).

**Data and Methods:** ESRI ArcGIS Desktop® commercial software (ArcGIS) and open-source Python SciKit-Learn [10] were used to apply machine learning techniques on reflectance albedo data from processed Lunar Reconnaissance Orbiter Camera (LROC) images. The first technique utilized the Maximum Likelihood Classification (MLC) algorithm in ArcGIS, which has been applied to multi-band Earth satellite data for many decades [11]. The MLC is a "supervised" classification method, which requires training areas in the image for a defined class of features. Pixels within each area for every class, and for every band, are then used to classify every pixel in N-dimensional space. The MLC also calculates the probability for each pixel, i.e. how likely the assigned class is correct based on the training data.

The second technique utilized a K-means classification method in the SciKit-Learn library. This "unsupervised", clustering-style algorithm is fully automated and only requires user-specified \( k \) number of clusters be identified in N-dimensional space. This algorithm iteratively determines the locations of \( k \) cluster centroids and assigns each pixel to the nearest centroid. As the pixel locations are known, the image data evaluated in albedo parameter space (for every band) uses this information to determine the cluster centroid assignment. K-means also has the capability of calculating the optimal number of clusters using an "elbow" method.

The reflectance albedo, or I/F, was derived from a co-registered Digital Elevation Model (DEM) of the study area produced using stereophotoclinometry (SPC). Details of the SPC process can be found in [12-13]. As the study area is wider than a single LROC Narrow Angle Camera (NAC) strip, the left and right channels of image pairs were combined for full coverage. A search for NAC image pairs produced three sets with incidence angles of 18°, 42° and 66°. These image datasets were then used in each classification algorithm.

Impact craters and other low/high relief features in a given area can produce artificially high albedo values due to slope effects. In order to focus on the albedo differences of swirls, impact craters down to 50 m in diameter were identified and masked using ArcGIS. A large trough-like feature and several high-standing areas were also masked.

For the MLC algorithm, we selected multiple training areas for four defined "classes" of swirls: on-swirl, off-swirl, diffuse-swirl, and high-swirl (see Fig. 1). On-swirl areas are general high-albedo, off-swirl are the dark lanes between high-albedo areas or the dark background surface, diffuse-swirl are the transition areas between on- and off-swirl, and high-swirl are very high-albedo areas within the on-swirl locations. Selected areas are the most representative for each class type. Pixel counts for on-, off-, diffuse-, and high-swirl are 916,322, 401,842, 96,990, and 25,852 respectively.

**Results and Discussion:** The MLC swirl classification map is generally a good match to what is observed in the reflectance albedo (Figure 2). Areas with higher albedo are classified on-swirl, areas that are mostly dark are off-swirl, and areas in between are diffuse-swirl. The probability of a correctly assigned class ranges from 86% (low) to 99% (high). Most on- and off-swirl areas have high probabilities (> ~93%), including the training areas (see Fig. 2 right). Most low probability areas (< 90%) are associated with anomalously high albedos, typically associated with high relief or geologic features such as the edges of crater rims or ejecta deposits. Even so, the algorithm still classified the vast majority of low probability areas to what would be expected from observation. Pixel counts for on-, off- and diffuse-, and high-swirl of the entire study area are 1,757,432, 850,269 and 1,647,026, and 116,651 respectively.

The K-means produced a very similar map to the MLC algorithm using three optimal classes minus the high-swirl area (Figure 3). Any differences appear to
be in locations where albedo values border between two class types. These results suggest that either the training areas for the MLC are good representatives for each class type, that unsupervised and optimal supervised are equally effective, or perhaps both (see Fig. 3). For the MLC algorithm, there is more prior knowledge about the feature classes with user-supplied data. For the K-means algorithm, there is less user interpretation of feature classes, but the optimal number of classes is derived from the data. The similarity in results from both algorithms provides increased confidence in the resulting classifications and also takes advantage of the strengths of each method.

**Future Work:** This is the first quantitative mapping of lunar swirls based on albedo, providing measurable criteria for on-, off-, and diffuse-swirl areas. This work is part of a larger project focused on establishing a better understanding of the various processes acting on the surface at lunar swirl locales. There is currently evidence that on-swirl areas in Mare Ingenii are associated with meter-scale topographic lows which has implications for dust migration at the finest size-fractions [14]. Using these algorithms, we intend to combine topographic, photometric, and mineralogic data for machine learning to determine which are the key criteria for defining the different sub-areas in swirl regions. We will also expand the use of our algorithms to swirls in Mare Ingenii.