# Project Proposal

**Chao-Ming Yen** **Tongkai Shao**

We plan to complete a standalone machine-learning algorithm that is able to accurately classify the image data set of the project. First, we plan to apply boundary/edge recognition techniques on raw data to separate the target object from the background. Next, we extract characteristic features from these images that have no background. The characteristic image features have 3 categories: shape (number of corners, diameters, concavity…etc), color and texture. For shape, we want to try HOG/SIFT methods and Harris corner detector to extract some shape features. For color, we want to try some analysis on RGB frequency and produce some useful features. For texture, we want to try different ways to analyze the spatial distribution of color and generate the texture features. After we process raw image data into these different kinds of features, we then apply machine-learning classifiers on these features. The classifiers that we want to try including Logistic Regression, Neural Network and Support Vector Machine. We will train these different classifiers with training dataset and find one that perform best on the testing dataset to serve as our primary machine-learning algorithm of this project. We sum up the whole process in the following picture: