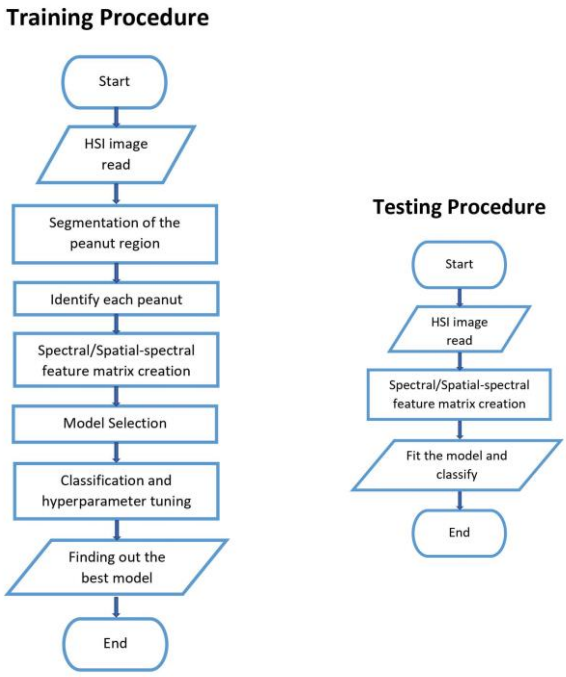
I.METHODOLOGY

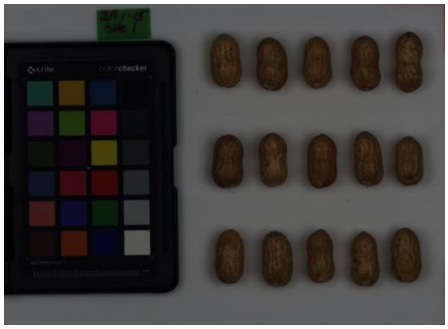
A. Data Preprocessing

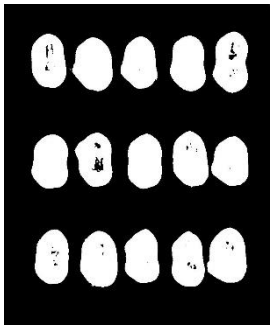
All the data preprocessing steps are shown in figure-1 as a part of the methodology. Spectral Python library (SPy) [7] has been used to read hyperspectral image data. Since each image file contains 15 images total, firstly, the peanut pixels have been segmented from the background. After segmentation, calibration panel and non-uniform background have been removed by cropping the segmentation mask beyond calibration panels' location. Then, all the pixels in the image have been identified and labeled according to which peanut they belong to. For spectral based peanut maturity classification, peanut spectrum has been calculated by averaging the spectrum of all pixels of a peanut. For Spatial-spectral based classification, spectrum at different regions of the peanut has been estimated which has decided the maturity of the peanut.



1. Image segmentation

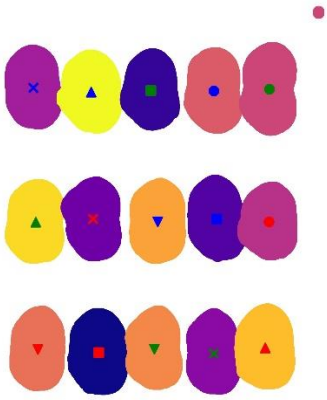
First of all, reflectance values of Red, Green and Blue wavelength from HSI image were merged together to form an RGB image for segmentation purpose. For red, green and blue channel, 450 nm, 550 nm and 650 nm wavelength were considered respectively. Ostu thresholding has shown good peanut segmentation performance for all of the images. An example of ostu thresholded image is shown in figure-2 where peanut pixels are labeled as white and backgrounds are labeled as black. In the segmentation mask, some peanut's region has been labeled as background (figure-2). Implementation of the morphological dilation operation with kernel size (5,5) has shown improvement in filtering those false background pixels in peanut region but it enlarges the peanut mask at the edges (figure-3). Segmentation mask found after all of those mentioned operation has been used for the peanut indentification and feature matrix creation.





2. Indentify each peanut from the image:

The identification of each peanut in the segmentation mask is a tricky task. The pixel position, row and column of that pixel, of one peanut is far away from other peanuts. Therefore, k-means clustering was algorithm used to group peanut pixel based on the distance between the pixel position. Figure-4 is showing the masks of all identified peanut. The convention of reading peanut number is given in the caption of figure-3. In figure-4, it is clear that kmeans algorithm does not label all the peanut squentially. In figure-4, first peanut (top-left) has been labeled as 8. This issue has been handled by re-ordering the label of all the peanut pixel.



B. Create Feature Matrix for classification:

Two types of feature matrixes have been created for classification – spectral feature based, spectral-spatial feature based. Section 2 only returns position of each peanut in the image. For spectral feature creation, average reflectange at all the pixel position of a peanut at certain wavelength is calculated first and the average reflectance of all wavelengths gives the spectrum of that peanut. For spectral-spatial feature matrix, whole peanut image has been divided into (2,8) matrix. Then average spectrum is calculated for all the regions of a peanut. Lastly, the information of all the spectrum has been organised as an image of shape (16, Number of wavelength) with all the regions as a row and reflectance at all wavelength in the column.

C. Classifier selection , Classification and Analysis:

Lastly, a suitable classifier is chosen for maturity classification and hyperparameters are tuned to find out best model. Detailed analysis of classifier selection and performance of the classifier will be discussed in the Section IV.