

ABSTRACT

Papaya is one of the fruits that are widely consumed in Indonesia. This fruit has many nutrients that are good for the human body. The quality of the papaya fruit can be determined by its skin, size, and level of maturity. In determining the level of ripeness of papaya fruit, many people only estimate it by looking at the texture of the skin and the weight of the papaya fruit. However, this cannot be used as a reference because everyone's perception varies, especially when going to export papaya fruit in large quantities. Besides, it takes a long time and exhausts the eyes if done in the usual way. Therefore, a more effective way to determine the level of ripeness of papaya fruit is distinguished into three, namely raw, half-ripe, and ripe.

This final project analyses the level of ripeness and weight prediction of papaya fruit using the machine learning methods. Papaya images are processed using image processing to get the GLCM features to classify the ripeness and predict the weight of papaya. The values of these features are then used as input for calculating the process using the machine learning method. Machine learning methods used for classifying the ripeness of papaya are K-Nearest Neighbor (KNN), Naïve Bayes, Random Forest, and Support Vector Machine (SVM). Meanwhile, linear regression is used to predict the weight of papaya.

The results of the analysis on this final project is expected to help determine the level of ripeness and predict the weight of papaya fruit more effectively. The four machine learning methods used to determine the level of ripeness of papaya fruit, KNN, Naïve Bayes, Random Forest, and SVM are compared and analyzed to know which method has better accuracy. Meanwhile, for papaya weight prediction using linear regression, the accuracy of this method is analyzed. From the research conducted, it is known that for papaya fruit classification the best method is Random Forest with an accuracy of 70.58%, while for papaya fruit weight prediction, linear regression method has an accuracy of 50%.

Keywords : *Image processing, K-Nearest Neighbor, Naive Bayes, Random Forest, SVM*