**Chapter III**

**METHODOLOGY**

**Introduction**

The Philippine Grain Standardization Program is a government program spearheaded by the National Food Authority to provide commercial assessment standards for the determination of the grade and quality of milled rice products. The implementation of the program started on September 21, 2002. From its establishment, the National Grains Standard has been formed. The National Grains Standards defined the characteristics classification of the rice grain samples. Factors for determining grade include dimensional length, degree of milling, percentage by weight of broken kernels, brewers, red kernels, immature kernels, chalky kernels, damaged kernels, yellow kernels, age-related changes, and other characteristics. The grades are based on the percentage by weight of the classified grains to the overall weight of the product. The Grade 5 is the lowest and the Premium grade is the highest grade a milled rice product can be classified to. Moreover, the implementation of these standards in the market is expected to boost the quality of the rice products in the Philippines.

Existing studies are aimed to develop simple, affordable, and accessible grading methodologies. Image processing techniques are the most common methods used to classify and grade rice grains based from different standards all over the world. Thresholding techniques were used to distinguish the chalky region of a grain and ultimately quantify its region percentage (Chandra, Barman, & Ghosh). The amount of the chalky region signifies the breaking capacity of the grain and this degrades the quality of the product. A lot of studies correlate the degree of milling of the rice products to its quality. A study made in 2001 monitors the degree of milling of rice grain samples using the whiteness of the rice grains (Yadav & Jindal). An image of the rice grain samples is obtained using CCD Camera mounted to a platform equipped with image enhancing components. The image is analyzed by a computer running an analysis software. Several studies even use machine learning algorithms to determine the grade of the milled rice. The machine learning program learned how to distinguish between grades when fed with the training data obtained from manual methods (Neelamegam. P, S, & Valantina.S.).

The purposes of this study are to lessen the subjectivity of rice grading assessments based on the National Grain Standards by utilizing the consistency and precision of computer-aided assessments and to speed up the grading time. Using image processing methodologies, the study aims to develop a milled rice grading system that is portable and accessible to people and organizations who are working directly on rice like millers, distributors, and farmers. The main parts of this study are directed towards (1) the gathering of the qualitative reference values set by the NGS and the rice quality assessors and feature extraction; (2) creation of a portable standalone device for classification and image acquisition; (3) development of the classification process using image analysis and machine learning/classifier, specifically support vector machines with adaptive boosting; and (4) the display of the grade report. Activities performed by the image processing application include the counting and dimensioning, color analysis, grain classification, percentage computation, and grading.

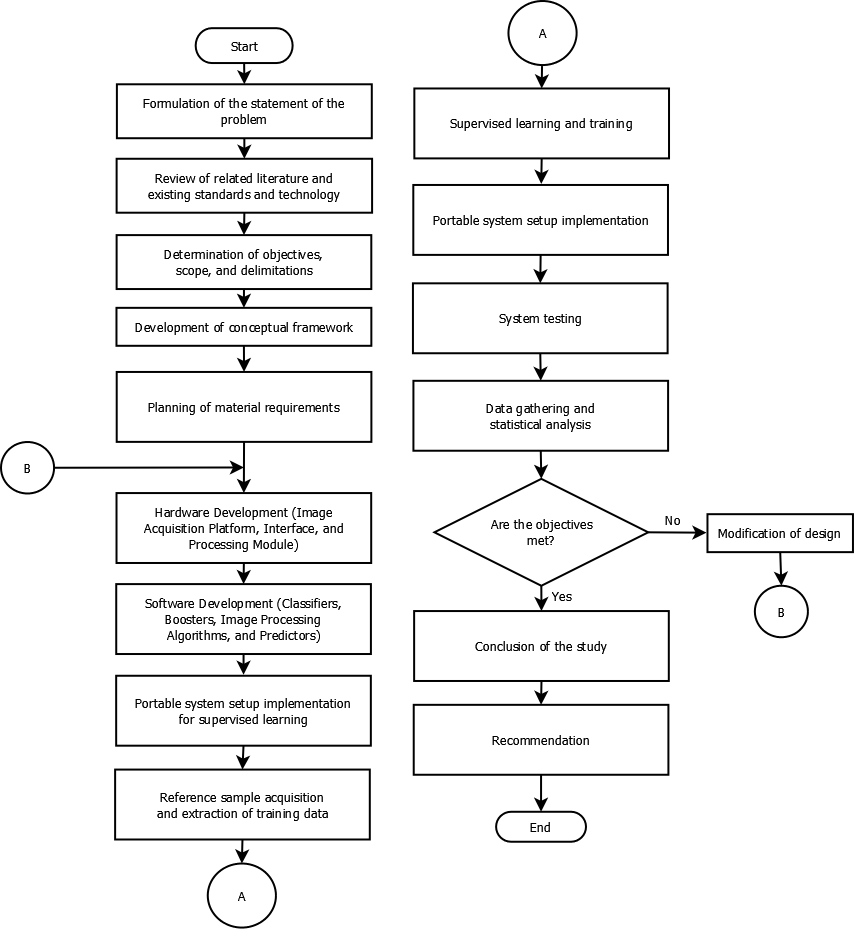


Figure. The Research Flow

The flow of the study is illustrated by the figure \_\_\_\_. The evaluation of the current grading process of rice begins with knowing the processes involved in it. Based from these evidential circumstances, the promise of automation is to be investigated. Reviewing current related technologies and literature involve the analysis of the standards set by the respective governing body and the process of how manual grading is done. A thorough review of related technologies (e.g. machine learning, image processing) is also essential to the study process. Appropriate algorithms and methodologies are needed to be reviewed when it comes to this kind of applications.

The study will have boundaries, as set by scopes and delimitations, and specific objectives that need to be met. Basing from the perceived problem, the specific objectives aim to set the path of the research with the goal of solving it. The scopes are determined by their correlation to the specific objectives. Moreover, the delimitations are to be determined based on the scope itself with the consideration of the objectives and the fact that the problem only requires a certain kind of solution to be investigated.

The objectives’ completion is dependent on the proper formulation of the conceptual framework. The framework summarizes the necessary input to processes involved in meeting the objectives and the corresponding output of these processes. The conceptual framework abstracts the necessary algorithms and modules to be used for the study. Along with the construction of the conceptual framework, the step by step process of the grading system are to be developed with, as mentioned, the objectives in first consideration.

After designing the framework of the system, careful deliberation on the materials procurement will be done. This process includes the selection of necessary hardware and components along with the specifications and compatibility among others. The circuitry and mechanism involved in the process of image acquisition and grading will be considered as part of the hardware-software interaction planning.

The hardware assembly will proceed after the materials planning. In this phase, the image acquisition and computing platform need to be constructed and be ready for the classifier training and supervised learning. The implementation of the hardware comes hand in hand with the software. The software implementation begins with the translation of the conceptual framework processes into codes and modules for the software. The system, composed of the hardware and the software, will be prepared for tuning in a non-production environment.

As the study involves machine learning (i.e. SVM), proper training data must be acquired first. The training data are in a form of compiled images relating to the grain type being classified. Credibility of the data is crucial to the success of the study. Therefore, basing from the standards of the NGS and the accredited personnel, the training data will be labelled properly. Moreover, some of the training data will be reserved and to be used as test data.

After the labeling of the training data, the supervised learning will be performed for the machine learning processes implemented in the system. A certain level of accuracy of the results should be met before implementing it on a production environment. System implementation will start after meeting this level.

The system implementation will be performed after all necessary tunings and training are done. However, this implementation will be subjected to testing measures and the results of the predictions compared with the manual grading testing will undergo statistical analysis to see if the objectives are met. In case some of the objectives are not met, modifications to the conceptual framework and the algorithms involved will be done. Otherwise, the objectives’ completion will be evaluated and necessary conclusions will be reported. Lastly, recommendations for further improvement of the system will be included to facilitate development.

Figure. The Conceptual Framework

The figure \_\_\_\_\_\_\_ illustrates the conceptual framework of the study. An input image of non-overlapping grains will be segmented with a threshold found using Otsu’s method. The background will be subtracted based from this value. A connected-component labelling will be done along with the counting of the total individual labels in the image. The image must have more than \_\_\_\_ objects. Otherwise, the system outputs an error stating that the sample size is below the recommended count and requires the user to increase the sample count. These individual objects may or may not be grains (e.g. foreign materials). Bounding boxes will be overlaid on each of the labelled objects and they will be extracted into new images.

The grain validation process will be performed for every extracted image. In this process, the histogram of gradients (HOG) features will be created. The object from the image will be classified into ‘grain’ or ‘foreign material’ by the support vector machine (SVM) called SVM-VAL which is trained to determine if an input image is a ‘grain’ or ‘foreign material’ with the HOG features as the feature vector. If the object is a grain, then it will be processed further. Otherwise, it will be classified as a ‘foreign material’ and the respective counters will be incremented.

After validating that the object is a grain, the grain size classification will start. In this process, the grain image’s HOG features obtained from the previous HOG determination will be used as an input feature vector for the SVM-BKN that determines if the input image is of a ‘broken’ or ‘unbroken’ grain. Respective counters will be incremented based on the classification. If the grain is ‘unbroken’, the major axis length will be recorded for later averaging process.

The next process is the grain type classification. In this process, different SVMs will be used: SVM-GRN, SVM-YLW, SVM-RED, SVM-DOM, SVM-PAD, and SVM-DAM. The SVM-GRN is used to classify if the grain is immature or not. Using the SVM-YLW, the grain will be classified as fermented or not. The SVM-RED will classify the grain as red kernel or not. Moreover, SVM-DOM is a multiclass classifier that determines the degree of milling of the grain. The SVM-PAD determines if the grain is a paddy or a milled rice grain. Lastly, the SVM-DAM determines if the grain is damaged. The input feature vectors for these SVMs are based on the histogram characteristics of the image. Respective counters will be incremented based on the classification of the image. Based from the average length of the unbroken grains obtained earlier, the grain will be classified into regular, broken, and brewer and the respective counters will be incremented.

The final classification of the grain is the determination if it is chalky. A thresholding process will be performed to separate the area of the chalky region from the non-chalky. The percentage area of the chalky region will be computed and based from this, the grain will be determined if it is chalky or not. Again, the respective counter will be incremented.

After all the processing, the percentage by count of the classifications will be computed. The count of the red, green, yellow, damaged, chalky, broken, brewer, paddy kernels and foreign materials will be compared to the total number of objects in the rice grain sample for percentage. The degree of milling is determined by the majority degree of the grains processed.

Based from these percentages, the grade will be determined. The rice grains will be classified into PREMIUM, or GRADE 1 to GRADE 5 accordingly.

The output of the system is the grade of the sample along with the summary of the characteristics of the sample.

**Hardware Development**