“KAALAMAN”: Development of farmer’s rice doctor kiosk for increase production

R C Muñoz, Jr.1,2,5, W G Valdez3, M I Muñoz4, M P Arnoco1, A C Avito1, V B Baloro1, J N Barrientos1, M T Gervacio1, R E Hipolito1, S P Rodrigo1, C R Salenga1, H P Sanchez1, and C G Zulueta1

1Electronics Engineering Department, College of Engineering and Architecture, Bataan Peninsula State University, Philippines, 2100.

2Research and Development Office, Bataan Peninsula State University, Balanga City, Bataan, 2100 Philippines

3Agricultural Engineering Department, College of Engineering and Architecture, Bataan Peninsula State University, Philippines, 2100.

4Information and Technology Consultancy, Balanga City, Philippines, 2100.

5E-mail: rhodjr@bpsu.edu.ph / rhodjr@outlook.ph

Abstract. Rice is the major source of food here in the Philippines but pests and diseases damaging the crops leave with a lesser quality harvest. The nutrient such as Nitrogen (N), Phosphorus (P), and Potassium (K) contents of the soil is also important as to its right amount. The study is a point-of-service kiosk helping farmers avoid inaccurate leaf color reading. This benefitted all farmers in the province of Bataan and the region where most of the living is obtained from agricultural farms. This is a technological solution to the quality of crops before harvest by analyzing (N) level of deficiency. Here, farmers get close to the accurate quantified level of deficiency and recommend a substantial amount of fertilizer. It implements an RGB color mapping technique in pest and disease identification and Chi-squared distance comparison on RGB histogram in the image detection. Testing and calibration were done by comparing the sample data to standard laboratory results that show a detection accuracy of at least 95%. The production yield of the crops shall improve while decreasing the production cost. This serves as a virtual agricultural technician (VAT) accessible to the farmers. Information is available via internet connectivity with response good enough providing quick answers to queries.

1. Introduction
Rice is the world's second most important cereal crop following corn based on production volume. Nearly 482 million metric tons of husked rice was produced in the last harvesting year of 2017 worldwide [1]. Countries in Asia have the largest share in world rice production. It is the staple food for over half the world's population [2]. A recorded increase in rice production has been observed since the start of the Green Revolution. The United States Department of Agriculture (USDA) estimated that the world rice production in 2017-2018 with 481.04 million metric tons, around 0.26 million tons less than the earlier month's projection of 2017. The rice production by some country for the month of June 2017...
measured in metric tons are: China: 145,000,000; India: 106,000,000; Indonesia: 37,000,000; Bangladesh: 34,700,000; Vietnam: 28,100,000; Thailand: 19,500,000; Burma: 12,300,000 and Philippines: 11,200,000 [1].

Fifteen countries account for 90% of the world's rice harvest [3]. China and India alone account for up to 50% of the rice grown together with Indonesia, Bangladesh, Vietnam, Myanmar, Thailand, the Philippines, Japan, Pakistan, Cambodia, the Republic of Korea, Nepal, and Sri Lanka. Asian countries account for 50% of the world's total rice production [2].

Rice production in the Philippines is important to the food supply in the country and economy [3]. It is the 9th largest rice producer in the world, accounting for 2.8% of global rice production. In 2017, it was included in the annual report that the country produced a total of 19,276,346.63 metric tons of palay [1]. It was also the world's largest rice importer in 2010 [3]. Rice is produced extensively in Luzon, the Western Visayas, Southern Mindanao, and Central Mindanao [4].

Central Luzon contains the largest plain in the country with its agricultural plains account about 40% of the region's area [5]. It produces most of the country's rice supply, earning itself the nickname "Rice Bowl of the Philippines or Rice Granary of the Philippines". It accounts for 18% to 19% of the country's total palay production. The region only needs two million metric tons to sustain its population and the rest is sold as surplus to other regions [6].

The province of Bataan in Central Luzon contributes to the production of palay. Its annual volume of production, which was measured in metric tons, from 2011 up to 2014 increased from 123,511 to 155,056 but dropped from 2015 to 2016 with 151,610.00 to 122,393.00, respectively. It increased the following year with 136,269 metric tons. The province’s annual report of palay production for 2017 was 136,269 metric tons as shown in table 1 [1].

Table 1. Rice volume production by a quarter in the year 2017.

| Quarter | Volume production, (metric ton) | Total (metric ton) |
|---------|---------------------------------|--------------------|
|         | Irrigated                       | Rainfed            |                    |
| 1       | 53,687                          | 2                  | 53,689             |
| 2       | 18,240                          |                    | 18,240             |
| 3       | 12,436                          |                    | 12,436             |
| 4       | 50,424                          | 1,480              | 51,904             |
| Total   | 136,269                         |                    |                    |

Figure 1. Farmlands in Municipalities of Bataan (unit: ha).
Philippine Rice Information System (PRISLM) monitors farmlands in Bataan regarding the occurrence of pests and palay diseases as shown in figure 1 [7]. By far, two towns Hermosa and Dinalupihan contributes the largest farmland in the province and greatly affected by pest and diseases.

Pests have been always one of the main issues of the farmers causing damage to the crops. Farmers lose an average of 37% of their rice crop to pests and diseases every year that greatly reduce yield [8]. They are mainly caused by bacteria, viruses, or fungi, and planting a resistant variety is the simplest and often the most cost-effective management for diseases [8]. However, agricultural people have the nature of spraying pesticides that cause the rice to absorb chemicals not necessary for human intake. The knowledge of when those insects usually attack is the latter's way of avoiding their possible massive occurrence.

Several common insects such as the leaf folder, whorl maggot, and armyworms can cause highly visible damage symptoms in the early stages of the rice crops [9]. Although the damage is rarely enough to reduce yield because the crop can compensate for early damage. Insecticides are applied in rice fields during the early crop stages to control leaf folders or whorl maggots but do not benefit farmers economically. Instead, they can cause an imbalance in the natural insect population that may lead to pest outbreaks [9]. Natural fertilizers have immense benefits for soil and crop production. Organic fertilizers also help improve the structure of the soil including the circulation of air, which sustains beneficial microorganisms that help release nutrients to the soil [10].

One of the most important fertilizers in rice production is Nitrogen (N). It is usually applied during the growing season to ensure that the crop’s nitrogen needs. The national agency suggested the use of the Leaf Colour Chart (LCC) to determine the (N) fertilizer needs of rice crops. It has six green strips, with color ranging from yellow-green to dark green. It determines the greenness of the rice leaf, which indicates its N content [9]. However, due to the environment and visual impairments, misjudgment from farmers occur and they take it for granted relying on the traditional way.

Another concern of the farmer is checking the soil condition if it has sufficient nutrient needs of the crop. It is important to optimize crop production, to protect the environment from contamination by run-off and leaching of excess fertilizers, to aid in the diagnosis of plant culture problems, to improve the nutritional balance of the growing media, and to save money and conserve energy by applying only the amount of fertilizer needed [11].

Nowadays, developing a system that makes use of GSM and SMS technologies making a farmer's work much easier. Less dependence on the conditions present would be greatly beneficial to them. They could contact those in the authority for inquiries by sending a message through different platforms. The problems arising in the production of palay, and information point-of-service would be beneficial to the farmers. It can help in disseminating information and awareness about the diseases caught by the crop and recommends preventive measures the farmers can apply. Pest present in each season can be projected and determined, as well as the nutrients required can be calculated through the color of the leaf using some computer algorithms. This could also progress in populating production information that would be of help to others in the future.

2. Significance of the study

Pest spreading remains a serious concern for Filipino farmers [12]. Its significant increase usually becomes our most food competitors. An estimated average of 37% of the rice crop losses to pests and diseases every year. Farmers are implementing pesticides and chemical fertilizers to battle these factors such as bacteria, viruses, or fungi, and by other living organisms, like rats, or by non-living factors, such as wind, water, temperature, radiation, and soil acidity [9]. But incorrect and inaccurate use of fertilizers and insecticides are expensive, and worst can harm the environment [13]. Too little of it will cause the crops not to produce as much as they should. And excess nutrients at the wrong time will run off to the fields and pollute streams and groundwater. While fertilizers serve as an important purpose, careful and right amounts at the right time can avoid its negative effects on the environment [13]. This especially significant in the case of improper application of (N)-(P)-(K) fertilizers.
3. Methodology

3.1. Crop health and farmland assessment
Surveys to farmers were conducted to profile the production fields' condition as to variety, crop establishment method, nutrient inputs, and pest control methods. It also helps to assess damages caused by diseases, animal pests, and weeds. Results are used to produce information to serve as the basis for prioritizing activities for the study. Over eighty-two (82) respondent farmers from the municipality of Hermosa were assessed based on crop health and farmland as shown in figure 2. The three common problems met by the farmers are pest control, disease control, and the right amount of fertilizer application. This study focuses on giving a solution to these issues in rice production. PRISM’s methodology of crop health assessment was used in identifying requirements to include in the study.

![Figure 2. Common problems met by the farmers of Hermosa, Bataan.](image)

3.2. Implementation of target areas
In this study, the municipality of Hermosa with the biggest contribution in rice production in the province of Bataan set as the pilot site.

3.3. System diagram
Image processing was used to determine what type of pest or disease is of concern. The process also determines the nutrient content of the crop or soil. Database stores and retrieve information processed by the system. Local hosting of the server computer accesses the network services required by the program. The point-of-service kiosk serves as an information hub giving information on pests, diseases, and nutrients of rice crops to the farmers. It is composed of two main components, the raspberry pi3 (Rpi3 or Rpi) and the touch screen monitor. The monitor is the input and output device of the system. It shows historical statistical graphs of pest and disease. The Rpi is connected to a 5V 2.5A power supply where the touch screen monitor is also connected. The system is limited to information distribution, inquiries submission, and image processing.

Farmers may submit an image of possible newly discovered pests by uploading through a USB port or Bluetooth connection to the Rpi. The kiosk will send to the server computer the inquiry of the attached image, while image processing is done in the local unit. It can also manage inquiries and respond through SMS messages. This process was illustrated in figure 3.
3.4. Result collection and validation

3.4.1. RGB technique and Chi-square distance. The study was designed as a stand-alone system that processes the input image of pest, disease, leaf, and soil using an RGB histogram algorithm. The chi-squared distance equation was used to weight the regions based on the importance of the information they contain. This was also used to calculate the percent difference between the two images being compared. The weighted Chi-square distance can be defined as:

$$\chi^2(x, \xi) = \sum_{i,j} \omega_j (x_{ij} - \xi_{ij})^2 / (x_{ij} + \xi_{ij})$$  \hspace{1cm} (1)

where:
- $x$ and $\xi$ = the normalized enhanced histograms of the two images to be compared.
- $i$ and $j$ = indices of the $i^{th}$ bin in histogram corresponding to the $j^{th}$ local region.
- $\omega_j$ = the weight for region $j$.

The result from the image matching computed using the Chi-squared distance equation will be evaluated to identify the optimal threshold, otherwise, the output information is “NO MATCH FOUND” displayed in the monitor.

3.4.2. System recalibration. Input values in the system are calibrated to analyze the accuracy and capability for pest and disease detection. The same process applied to the nutrients based on its leaf and soil color. RGB Histogram was used in the detection and the basis in providing preventive recommendations. Eighty-two (82) trials (one trial per ha) using the image processing of uploaded pest photos to the point-of-service (kiosk) was conducted. It should automatically search images in the database with the set threshold in the algorithm then displaying the result.

The chi-squared distance of images uploaded to the kiosk and the images in the database are calculated and compared by the algorithm. The values having the least difference in the chi-squared distance value were selected and displayed. The same method is implemented in disease detection to provide preventive measures. The value of the threshold was determined as the highest acceptable value of the difference in chi-squared distance to search for a valid matching in the database.
ISO 5725-1:1994 suggests that attaining the 95% accuracy in image detection was recommended. In quantifying precise recommendation for (N) fertilizer, the same methodology was applied using RGB Histogram and the Chi-squared difference to get the best match of the sample photo with the one in the database. Again, eighty-two (82) trials (one trial per ha) using the image processing is done by directly uploading ten (10) leaf photos to the kiosk. The same procedure is used for a precise recommendation in the amount of (K) and (P).

4. Test Results

4.1. Controlled environment testing
To calibrate the algorithm for better and higher accuracy, several trials (one trial per ha) is conducted by directly uploading pest photos for analysis. The system automatically searches images in the database satisfying the set threshold in the algorithm. Results are displayed (figure 4) and the difference in the chi-squared distance is evaluated. It was evident from the figure that the difference in the results falls in the range 4.181 to 4.874 with an average result value of 4.572. The results fall within the 95% objective of the system but can be further improved through system and algorithm calibration.

![Figure 4. Image matching for leaf blight disease.](image)

4.2. System calibration
Cropping of images was used to improve the image matching. Query images were cropped with 400 x 300 pixels resolution. After cropping the image, the difference in the chi-squared distance resulted lowered to a range of 0.007 ~1.416 with an average value of the difference of 0.412 as shown in figure 5. This has a difference of 4.104 from the uncalibrated algorithm. The accuracy of the process in terms of image matching greatly improves. Then after setting the threshold, some results of a “No Match Found” (NMF) which implies the result does not meet the value of the threshold set. However, the result range falls within 0.007 to 0.671 and an average result of 0.257 as shown in figure 6 in this specific disease. This gives a 0.155 improvement in terms of matching.

![Figure 5. Testing for Leaf Blight Disease after Cropping of Image.](image)

![Figure 6. Test result after cropping of images with threshold setting.](image)
Several tests resulted in an optimal threshold specific for each type of image. Table 2 shows the experimental results of a specific threshold for implementations.

**Table 2. Parameters and the Computed Threshold.**

| Parameter         | Threshold |
|-------------------|-----------|
| Disease Matching  | 0.914     |
| Pest Matching     | 0.691     |
| Soil Matching     | 0.478     |

4.3. Field test

4.3.1. Pests detection using RGB Histogram with at least 95% accuracy. The system calculates the chi-squared distance of both the pest image and of the images in the database. The database image with a lesser value compared to the query image will be selected and displayed based from the threshold set. The pest in the rice field was identified first correctly. Illustrating in figure 7, the mean value of the difference was at 0.1825 which interpreted as average accuracy value of 99.8175%. With this confidence in the accuracy, the kiosk was reliable in identifying pests and providing preventive measures correctly.

![Figure 7. Chi-squared Distance Difference for Pests image.](image)

4.3.2. Disease detection using RGB Histogram with at least 95% accuracy. The size of the image and the threshold being set in the algorithm must be taken into consideration. In this process, cropping the uploaded image is done instead of resizing because it may compress or enlarge the pixels. This image is not considered raw data that caused an inaccurate result. Only the needed part will be cropped to preserve pixels for more accurate data. Even though this may give more accurate results it does not guarantee not to output wrong information. As shown in Table 2, the threshold computed is 0.914. The difference between the computed chi-squared distance of the inquiry image and the database image must be less than the threshold to have a valid match. This was observed in the prior discussion on the system calibration. This means cropping the image improves the performance restrictions to the system. Figures 5 and 6 show the results when applying the threshold.

Figures 8(a) and 8(b) shows the frequency of distribution of the results obtained from the cropped images of Dirty Panicle. It can be observed that both the results are ranging from 0.009 to .514 and had a mean of .17756. The results are the same which means that cropping the image is still within the range of the defined threshold.
Figure 8 Dirty Panicle results compared when the image is cropped without setting the threshold (a) and (b) with the threshold set.

As observed in figure 9(a) and 9(b), the results are steeper compared to the initial results. It is due to the restrictions introduced by the defined threshold. This threshold filters out the difference between the computed chi-squared distance of the inquiry message and the database image that is beyond 0.914.

Figure 9 Leaf Blight results compared when the image is cropped without setting the threshold (a) and (b) with the threshold set.

Table 3. Different tests of chi-square difference of the other types of diseases with and without threshold.

| Diseases type       | With threshold | Without threshold |
|---------------------|----------------|-------------------|
|                     | Lower | Mean   | Higher | Lower | Mean   | Higher |
| Leaf Blight         | 0.007 | 0.3912 | 1.416  | 0.007 | 0.2566 | 0.671  |
| Leaf Blast          | 0.004 | 0.14109| 0.479  | 0.004 | 0.14109| 0.479  |
| Sheath Rot          | 0.003 | 0.07533| 0.332  | 0.003 | 0.07533| 0.332  |
| Brown Spot          | 0.003 | 0.17361| 0.656  | 0.003 | 0.17361| 0.656  |
| False Smut          | 0.021 | 0.15002| 0.589  | 0.021 | 0.15002| 0.589  |
| Narrow Brown Spot   | 0.026 | 0.29695| 0.651  | 0.026 | 0.29695| 0.651  |
| Sheath Blight       | 0.004 | 0.08936| 0.356  | 0.004 | 0.08936| 0.356  |
Though it is expected that results must be the same as shown in figure 8(a) and 8(b) whether implementing a threshold or not, it was observed that discrepancy is introduced in some images as shown in table 3. This contributes to the effect in the accuracy of detection which is another avenue to study for further improvement.

4.3.3. Nitrogen fertilizer recommendation using RGB Histogram with at least 95% accuracy. Ten images of rice leaves or strips are uploaded to the system, match with the standard database images, and display results with recommendations. To achieve at least 95% accuracy, the program shall be limited to a maximum error of 5%. Results are presented in figure 10, the highest error that the program identified is 2.925% (97.075% without error) with its mean value of 1.838. This described that the program meets at least 95% (ISO 5725-1:1994) accuracy on determining the LCC category from the ten (10) leaves uploaded for the query.

![Figure 10](image1.png)

**Figure 10** Chi-squared distance difference of the query images from the database images.

4.3.4. Potassium (K) and Phosphorous (P) fertilizer recommendation using RGB Histogram with at least 95% accuracy. Presence for (P) and (K) undergo chemical tests using the Soil-Test-Kit (STK). The system is again comparing the chi-squared distance of the image of interest to that of the images stored in the database. After the process of matching the soil color, output recommends the amount of (P) and (K) the farmers can apply. Again, the program is limited to a maximum of 5% error to achieve a minimum of 95% (ISO 5725-1:1994) accuracy. The results presented in figure 11 shows a maximum error of 0.078% with its mean value of 0.01756. This corresponds to an accuracy much higher than the minimum accuracy requirement of 95%.

![Figure 11](image2.png)

**Figure 11** Chi-squared distance difference of the query images sample soil1 from the database images.
Table 4 illustrated below shows the chi-squared distance difference of other soil samples and their corresponding results to identify its accuracy.

| Soil sample | Lower | Mean   | Higher | Accuracy (%) |
|-------------|-------|--------|--------|--------------|
| Soil 1      | 0.002 | 0.01756| 0.078  | 99.98244     |
| Soil 2      | 0.002 | 0.02724| 0.090  | 99.97276     |
| Soil 3      | 0.192 | 0.35654| 0.519  | 99.64346     |
| Soil 5      | 0.004 | 0.02072| 0.065  | 99.97928     |
| Soil 6      | 0.001 | 0.00864| 0.068  | 99.99136     |

5. Conclusion
The technology developed for determining the pests, disease, and nutrients provides quantified data. The study proved that for pest and disease detection, using RGB Histogram and chi-squared difference satisfies the minimum 95% accuracy in image matching. Cropping of images and threshold settings included in the process improves the accuracy in the image match. LCC and soil comparative detection using RGB Histogram also satisfies the minimum 95% accuracy in image matching. The resulting accuracy can provide reliable recommendations to the farmers in terms of rice production.

Improving the edge detection for disease and pest image achieve more accurate and consistent results. Adding more types of pests and diseases in learning the algorithm will increase the accuracy, since the concept of the RGB histogram algorithm is to compare the inputted image to the images stored in the database pixel by pixel, thus, adding more images in the database will provide better results.

Future work for this study is the population of reliable good quality images to be added in the reference images in the database. This will increase the accuracy of the system. Deployment of this system in the localities where the farmer can be served is one of the priority activities to be planned.

Acknowledgments
This work was supported by the Department of Science and Technology (DOST) Grant-in-Aid (GIA) projects of DOST-Region 3 Satellite Office and the Research and Development Office (RDO) of the Bataan Peninsula State University (BPSU) grant funded by the university through General Appropriation Act (GAA).

References
[1] Shahbandeh M 2017 Rice-Statistics and Facts.
[2] Muthayya S S 2014 An overview of global rice production, supply, trade, and consumption. *Annals of the New York Academy of Sciences*. 1(1324): 7-14.
[3] Thukral N GM 2011 Factbox - Top 10 rice exporting, importing countries.
[4] Muthayya S 1991 Country study: rice and the green revolution. [cited 2014. Available from: https://en.wikipedia.org/wiki/Rice_production_in_the_Philippines.
[5] Cuyco A 2015 Central Luzon contains the largest plain in the country. Philippines:, National Nutrition Council.
[6] Flora I 2016 Rice bowl of the Philippines or rice granary of the Philippines. Sunstar Philippines.
[7] DA 2017 An interview with the department of agriculture, an interview. Balanga City, Bataan, Philippines: Office of the Provincial Agriculture, Agriculture.
[8] (IRRI) 2017 Mobile application for the identification and management of rice crop problems. [Online]: Available from: http://www.knowledgebank.irri.org/decision-tools/rice-doctor.
[9] (IRRI) 2014 Pests cause lesser yield and nuisance to farmers. [Online]: Available from: http://www.knowledgebank.irri.org/step-by-step-production/growth/pests-and-diseases12.
[10] Natural fertilizers for the soil and crop production 2015 The Organic Farmer.
[11] Cox D 2001 Current methods of greenhouse media testing and how they differ. University of Massachusetts, Extension.
[12] Stuecker M T 2018. Climate variability impacts on rice production in the Philippines. *PLoS ONE.* 8(13).
[13] DeJoia A 2015 Why do farmers use fertilizers? [Online]. Available from: https://soilsmatter.wordpress.com/2015/03/18/.

**Appendices**

As part of the validation process as to the performance of the system, the system was presented to the actual farmers to seek their practical advice and suggestions on the content and appropriateness of the system functionality. In regards to the technical validations of the system, it was presented into different forums and exhibits to gauge the impact of the specialists and experts. The pictures below show the activities where the system was opened.

![Demonstration with the Farmers](image1)

![National Science & Technology Week](image2)

![Research & Invention Competition Exhibit](image3)