International Journal of Engineering & Technology, 7 (4) (2018) 2909-2916

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i4.15679
Research paper

Paddy crop and weed classification using color features for computer vision based precision agriculture

Radhika Kamath [1] *, Mamatha Balachanra [2], Srikanth Prabhu [3]

[1,2,3] *Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India*
**Corresponding author E-mail: radhika.kamath@manipal.edu*

Abstract

Weed detection in paddy fields using robotic vision is still a challenging task. The main reason for this being lack of dataset. In this research work, the creation of paddy crop and weeds image dataset has been described. The images were acquired using a digital camera under natu-ral lighting conditions. For every image in the dataset, its annotated image was created using manual annotation. Annotation was carried out after plant segmentation from the soil background. Dataset consists of 300 images. In addition, classification of paddy crop and two types of weeds namely, sedges and grass-type weeds has been done using only color features. Usually, crop and weed discrimination is done using texture, shape and color features. However, using all three features may result in a computationally intensive system. Color features are extracted using a novel approach based on Speeded-up Robust Features (SURF). Random Forest, K-Nearest Neighbors (K-NN) and Least Squares Support Vector Machine (LSSVM) classifiers have been used for the classification of the paddy crop and two types of weeds. An accuracy around 86% was obtained by all the classifiers. This indicates color features can be relied upon in discrimination between paddy crop and sedges and grass-type weeds.

Keywords: *Classifiers; Computer Vision; Color Feature Extraction; Dataset; Precision Agriculture.*

1. Introduction

Lately, there is a big difference in the way the agronomists and farmers can gather and analyze data because of the advances in the technology. Automated livestock management, precision weed control and measurement of phenotypic characteristics of the plants and crops all allow us in attaining good yield and profit with less input. The main concept behind these systems is Computer Vision (CV). Computer Vision is defined as the process of analyzing images and videos automatically to obtain meaningful inference or measurements without human intervention. This is one of the latest technologies that are being used in precision agriculture. Precision Agriculture is defined as 'art and science of enhancing crop production using latest technology' [1]. Since many years, technology being used in agriculture such as mechanical harvester, various sensor networks to know about the current environmental condition and to predict environmental changes that may happen in near future. One of the reasons for using computer vision in agriculture is to eliminate the extensive use of chemical herbicides and to favor the development of environmentally friendly and non-chemical methodologies. In a few decades, it is predicted that most of the manual farming chores will be replaced by robotic farming, which will be based on the computer vision techniques to do the things like preparing the land for cultivation, weed control, monitoring and harvesting [2].

Weeds are unwanted plants growing amongst crops. Weeds compete with the crops for nutrients, water and sunlight thus can cause low yield. India is second largest in the farm output but the yield is very low and one of the main reasons for this is due to weeds [3]. Weed management is very poor India especially in coastal Karnataka state region of India because of non-availability of labors. In addition, weeds are mainly controlled in India by chemical or mechanical weeder. Overuse use chemical herbicide leads to contamination of groundwater. Many farmers in India lack the knowledge of site-specific treatment. That is, for a particular weed, suitable herbicide in right amount has to be used otherwise, it leads herbicide-resistant weeds. Both mechanical and chemical ways of controlling weeds take into account the general condition of the field without considering the spatial or temporal changes that can occur at minute-level. Therefore, this is the best time to harness the power of computer vision technology for precision weed management. Use of technology for weed management not only reduces labor problem but also gives way for chemical-free farming and helping in increasing the yield and reducing loss. India is most populous country next to China. Therefore, India has to use technology in agriculture to increase food production and hence meet growing food demand. According to [4], in future many of tasks in agriculture will be automated using computer vision or robotic vision technology. Common agricultural tasks like land preparation, sowing, weeding… etc. will be carried out by agricultural robots operating under computer vision technology. Even weed detection and removal will be automated with the help of computer vision as shown in fig. 1. These field robots have camera sensors, which will take images of the field, and images will be processed using advanced image processing techniques, which will help robots to take the appropriate action such as weed removal if detected. Also by flying drone on the field, images of the field can be captured. These images can be processed using advanced image processing techniques or computer vision techniques, which will help in identifying the weeds, so that specific type of treatment can be given in controlling or eradicating the weeds.

Fig. 1: Showing Future of the Agriculture. Weed Robots Removing Weeds [4].

Robotic Vision approaches to detect or to classify crop and weeds, which can be called as robotic weeding usually use shape, color and texture feature to discriminate crop and weeds. Many research works have been carried out which are based on these features to discriminate between crop and weeds. In [5], carrot crop, and weed image dataset have been created and using sliding window approach for feature extraction, crop and weed is discriminated using Random forest classifier. In [6], color features, size-dependent object descriptors, size independent shape features, and moment features have been used to discriminate between crop and weeds. In [7], corn crop and weeds were discriminated using a model built with normalized difference vegetation index (NDVI), shape index like area, length and texture index like entropy. In [8], to discriminate between crop and weed, shape, color and texture features are combined to form a common feature space. In [9], a classification algorithm was developed for Cabbage and Carrot crops and weeds. Eight morphological and color features are combined to create a joint feature space and for each object. Then feature selection is used to determine which the features suitable for discriminating weeds and crops are. The fuzzy logic based membership function is used for classification. In [10], texture features based on curvelet transform and Tamura texture feature were used in discriminating crop and weeds. In [11], the weed and crop were discriminated using texture feature extracted using color co-occurrence method. In [12], a review is made of the crop and weed segmentation. The paper very well gives the information about various techniques that have been used in segmenting weed and crops in the previously published works that have been carried out. In [13], Gabor filter was used get the texture features to discriminate between wheat crop and weeds like Bidens and Lolium. Here only Gabor filter was used get the texture information and accuracy obtained was in the range of 80's and 70's. In addition, false alarm rate was also quite high. In [14], the texture features like contrast, energy etc. and 11 shape features was extracted in order to classify individual weed leaves. In [15], texture features were extracted using gray-level co-occurrence matrix for leaf recognition for weed identification. Individual weed leaves of different types were taken and their features were extracted for weed identification. Weed-crop discrimination is nothing but differentiating two different types of plants. In [16], the images acquired using Unmanned Arial Vehicle (UAV) are analyzed for weed types, crop spatial distribution, and weed-crop ratio. Using visual features and geometrical features are extracted from the vegetation segmentation from the background. Random Forest classifier was used to classify the vegetation into crop and weed. The dataset consisted of sugar beet plants and the commonly available weeds on the fields of Germany and Switzerland. In [17], plants are differentiated using shape, color and texture features based on contrast, regularity, energy etc. In many studies, a combination of features based on shape, color, and texture has been done. In [18], many features like region-based features, contour-based features, and skeletal-based features are used for the identification of weed species for precision agriculture for site-specific treatment. [19] reviews the various approaches used for the automation of weed detection for site-specific treatment. Provides a good insight what has been done and what can be explored in the automation of weed detection using sensor technology and image-based methods.

Using shape features is not recommended by [20], because of the various reasons. First being leaves may overlap, secondly, leaves may move because of wind resulting in unclear boundaries. In addition to this, leaf structure/shape may vary significantly at different stages. This leads to having a large database of images captured at every stage of the crop. Moreover, some weeds exactly have the same shape as the crop. These factors also negatively affect texture analysis. Also, texture analysis is a time-consuming approach. Therefore, in these situations, color approaches place a few constraints when compared to shape and texture features. In literature, only a few research works have been carried out which concentrates on only color features for crop and weed discrimination. In [21], weed and crop discrimination are done using the color feature. HIS based Color co-occurrence matrix(CCM) is used to extract Entropy(E), Inertia quadrature (IQ), Inverse difference moment or local homogeneity (IDM) and Angular second moment (ASM). The accuracy of 78% is achieved with the Artificial neural network (ANN). In [22], classification of weed species has been done using HSI based CCM and discriminant analysis is used to discriminate between six types of weed species. In [23], discrimination between crop and weeds has been done using a blob coloring analysis based on region-based segmentation. According to [24], color features significantly improves classification accuracy. In this study discrimination between cabbage crop and weeds and carrot, the crop is done using morphological and color features. In [25], using four types of relative color indices of RGB gray levels, weeds were detected.

1.1. Objectives of the study

The main objectives of this study were as follows:

1) To create digital image dataset of paddy crop and weeds.
2) To extract color features of paddy crop and weeds using a novel approach based on SURF points.
3) To use classification techniques to classify crop and weeds and analyze their performance.

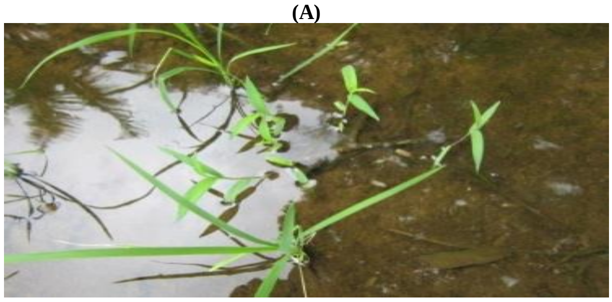
2. Materials and methods

2.1. Data collection

2.1.1. Image acquisition

A digital camera (Canon Powers hot SD3500 IS) was used to acquire images from paddy fields around Manipal, Karnataka State, India. Images were acquired under natural variable lighting conditions with the camera fixed at different heights. The pictures were shot randomly from different locations during July 2016 to August 2016 and July 2017 to August 2017. The weeds and crop of varying canopy size were selected for acquisition so as to increase the difficulty of identification. The images were stored in RGB color space with a resolution of 3456 X 2592 in JPG format. The MATLAB R2017a was used to process the images. While working with MATLAB images were resized to 1296 X 966. The Dataset consists of 300 images. To the best knowledge of the authors, this is the first Paddy crop and Weed image dataset.

(A)



(B)

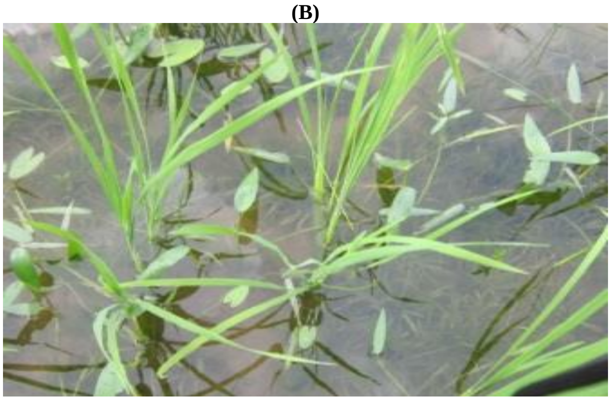


Fig. 2: A) and B) Showing Paddy Fields Images.

The images acquired from the field presented a variety of challenges due to the complex background. Since the paddy fields require standing water problems such as reflections, shadows of plants and other objects near the fields result. In this research work, background subtraction is done using YCbCr color space as shown in the fig. 3.

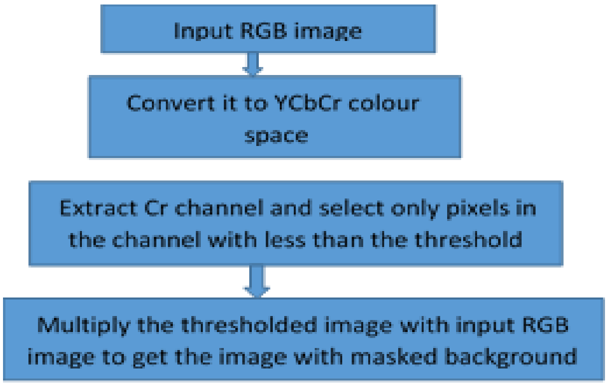
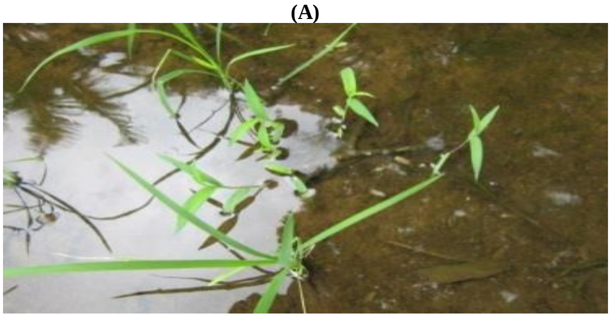


Fig. 3: Showing Steps Involved in Soil Background Subtraction.

(A)



(B)

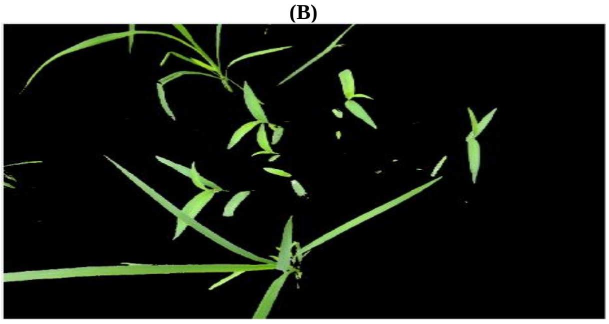
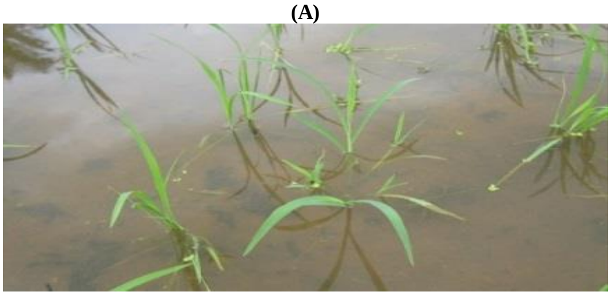


Fig. 4: A) & B) Showing before and after Soil Background Removal Respectively.

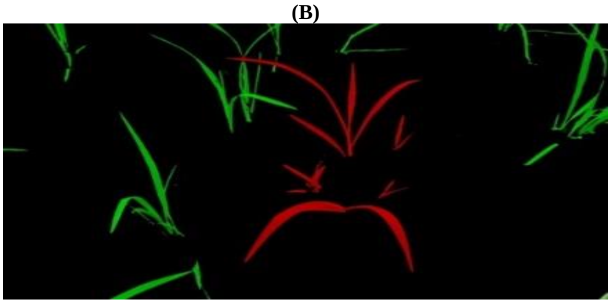
2.1.2. Crop and weed annotations

After green extraction, the images were annotated using GNU image manipulation software (GIMP) [26]. The plants which after the green extraction which resulted in too thin plant objects were labeled with black color in order to make it background. Also, the plants which heavily occluded or has heavy overlap and that cannot be identified as crop or weed are left out from the annotations. The crop was labeled with green color polygons and weeds are labeled with two colors other than green color. The most common types of weeds in paddy fields are grass-type weed and sedges weed. These weeds have their own unique characteristics. Grass-type weeds are represented by red color and all sedges by light blue color. The dataset contains the original images, images after background removal in JPG format and annotated images in JPG format and annotated images in GIMP image format. The dataset creation mainly for the purpose of segmenting crops and weeds and identification of weeds. Therefore, dataset consists of a variety of images ranging from instances of simple images to images with overlapping, containing both types of weeds. Also, it consists of crop and weeds with varying canopy size. This particular dataset throws many challenges since it is captured under natural variable conditions and different types of plants growing close together. Also, it gives us much information about crops and weeds. The individual plant information like leaf count, growth stage and other information that may be useful to agronomists can be obtained.

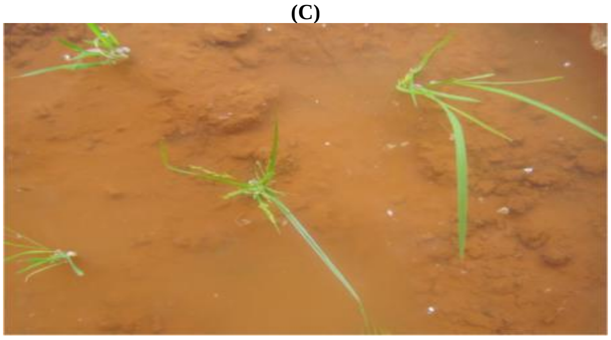
(A)



(B)



(C)



(D)

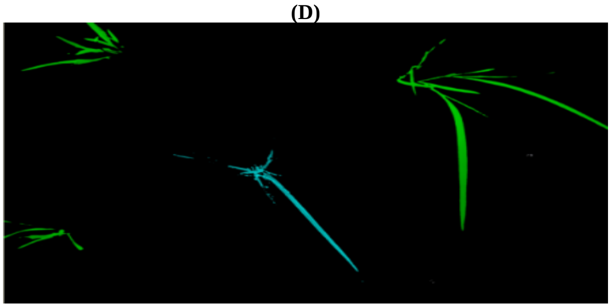


Fig. 5: A) and C) Shows Original Images, B), and D) Its Annotated Image after Green Plant Segmentation.

2.2. Methodology

2.2.1. Color feature extraction

The images with only one type of plant are obtained by using color-based segmentation from the annotated images as shown in the fig.6. This segmentation step helps in extraction of relevant features by focusing on the only one type of plant (that is crop, grass or sedges).

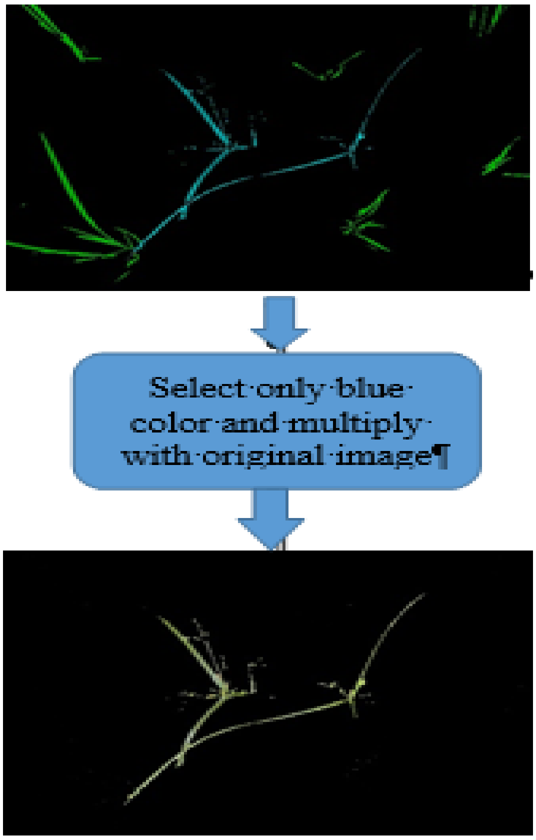
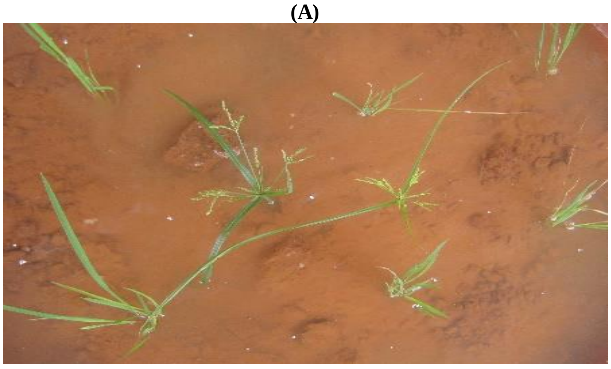
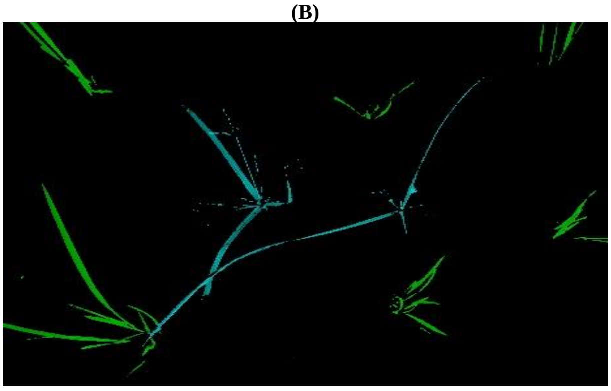


Fig. 6: Showing Steps Involved in the Extraction of Only One Type of Plant from Annotated Image.

(A)



(B)



(C)

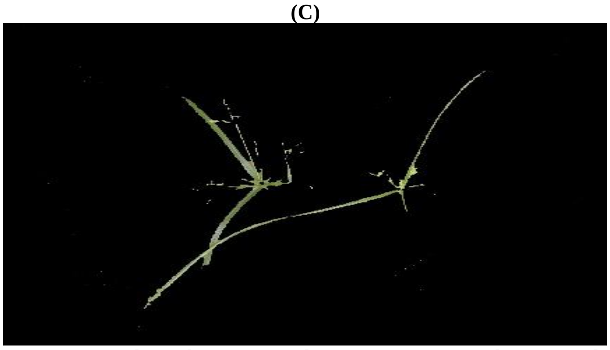


Fig. 7: A) B) and C) Shows Original Image, Its Annotated Image and an Image Containing Only One Type of Plant (Sedges) Respectively.

After obtaining an image with only one type of plant, the connected component algorithm [27] is used to get the connected components. Speeded-up Robust Feature (SURF) [28] is an interest point detector. Based on the approximation of Hessian matrix determinant, interest points are detected. Then, a window is created around each point and is divided into 16 sub-regions and 4 Haar wavelets responses are calculated from each sub-region using the integral images. This results in SURF descriptors in a vector. SURF points are detected on each connected components. From the preceding, step SURF points are obtained. Each SURF point is taken and a window of 65X65 is created with the point as center according to the equation 1.

$$S(p) = \{W(s,t) \in M^2 / s \in \left(x-\frac{N}{2}, x+\frac{N}{2}\right) \; and \; t \in (y-\frac{N}{2}, y+\frac{N}{2})\} \quad (1)$$

Where (x, y) are coordinates of a SURF point p and (s, t) are the coordinates of the neighborhood pixels of pixel p, N=64.

This window is converted into YCbCr color space. The reason for converting the window to YCbCr color space is that it is robust to illumination variations since the luminance or Y component is separated from two Chroma components Cb and Cr. Features like mean, minimum value, maximum value, standard deviation, skewness, and kurtosis are extracted from Cb and Cr component. These features are used to train the classifiers. Connected components less than 5000 pixels are not considered. The idea is to get details of a 65X65 region assuming that it may contain a small young plant or part of the plant. The color of a young plant varies when compared to the color of grown-up plant of the same type. This novel approach is summarized in the following pseudocode.

Pseudocode for Color Feature Extraction
Input: Read an image I consisting of only one type of plant
Find connected components in the inputted image I
For each connected component of size greater than 5000 pixels
Find SURF points on it
For each SURF point
Create a window W of size 65X65 with the SURF point as the center
Convert W to YCbCr color space
Compute Statistical features like mean, maximum, minimum, standard deviation, skewness and Kurtosis of Cb component and Cr component
Store these values in Feature Vector F(v)
End For
End For
Output: F (v)

2.2.2. Classification

Classification can be defined as a process of assigning classes to the given information. The tasks like identifying a given object in the image, classifying a given object in an image etc. involve a thorough understanding of perceived information. Using this understanding, presented information is assigned a class. A classifier is a computer agent which performs this type of classification. These classifiers are divided into two broad categories namely supervised and unsupervised. Supervised techniques learn with the help of training data. That is, they will have some prior knowledge. But unsupervised techniques are not based on prior knowledge. In this research work, K-NN, Random Forest, and LS-SVM classifier are used to check whether the result is consistent with all the classifiers.

2.2.3. K-NN classifier

The K-NN classifier is a supervised classifier, which is very simple and has been used by the data scientists from 1950's to perform pattern classification. K-NN classifier predicts the class of an object by finding the nearest neighbor. The closest neighbor is found by using distance measures such as Euclidean distance, which is most popular and frequently used.
Let (Ai, Ci) where i=1, 2, 3, 4, n data points. Ai denotes feature values and Ci denotes the class to which each Ai belongs. Let us assume Ci ϵ {1, 2, 3... C} for values of i. Now, consider a point A whose class is not known. To find out its class using K-NN classifier, the steps are given as follows

1) Let d (A, Ai) where i=1, 2, 3,…, n. 'd' denotes the distance between the points A and Ai. As a distance measure Euclidean, Minkowski, Spearman and many others can be used. In this study, Minkowski distance has been used as a distance measure. Consider a vector x (x_1, $x_{2,\dots,}x_n$) of length n and another vector y (y_1 ,$y_{2,\dots,}$ $y_{n)}$ having length n, then the Minkowski distance between vector x_s and y_t is given by

$$Minkowskidistance\ d = \sqrt[p]{\sum_{j=1}^{n}\left|x_{sj} + y_{tj}\right|^{p}} \tag{2}$$

2) The distance d is then sorted according to the ascending order
3) Assign a positive integer to K and pick the first K-distances obtained in step 2.
4) Find those K-points from the K-distances and determine their class.
5) The majority measure will be used to assign the class to A.

Choosing the value of K is a most important part. Too small value of K will result in overfitting problem and too large value for K would result in increased cost of computation. So, normally K will be equal to $n^{1/2}$. Cross-Validation can be used to optimize the result. The value of K is varied and one, which gives good accuracy, can be considered as optimal.

2.2.4. Random forest classifier

The tree-based supervised learning algorithm is considered to be one of the best as it provides high accuracy and maps the non-linear relationships effectively. Random Forest [29] is most popular method among data scientists as it can perform both classification and regression. It also performs well in handling outliers, filling missing values and other essential issues in data analytics. It comes under ensemble learning model wherein a group of weak learners comes together to form a strong model. In Random Forest multiple trees are built. If the classification of objects is based on features, multiple trees are built. Each tree gives a classification. The forest goes with the majority vote. The following points summarize the steps involved in Random Forest classifier as follows

1) First randomly m features are chosen from M features where m<M.
2) Using these m features build a node b which will be a root node using the best feature among m features. This is called as best split approach.
3) Make node b to have child nodes by using same best split approach.
4) Repeat the steps from one to three until 'p' numbers of nodes have been reached.
5) Repeat the steps from 1 to 4 until 'n' trees have been built.
6) Test Features are now taken and rules of each tree are applied to predict the class.
7) Final prediction is done by considering the majority vote in the forest.

2.2.5. LSSVM

SVM was improved to provide a solution to multi-class classification problem. Given a training set of N points {x_k, y_k} where k=1, 2, 3… N, $x_k \in \mathbb{R}^n$ is the k-th input data point and $y_k \in \{-1, +1\}$ the corresponding class label, then LSSVM [30] takes the form

$$y_x\ sign[\sum_{k=1}^{N} \alpha_k\ y_k\ K(x, x_k) + b] \tag{3}$$

Where α_k are called support values and b is a constant and are the solution to the linear system in contrast with standard SVM which is a quadratic. K is positive definite kernel and can be linear SVM given by

$$K(x, x_k) = x_k^T\ x \tag{4}$$

It can be polynomial SVM of degree d given by

$$K(x, x_k) = (x_k^T x + c)^{\ d} \tag{5}$$

It can be an RBF (Radial Basis Function) and is given by

$$k(x, x_k) = \exp(-\|x - x_k\|^2 \div 2\sigma^2) \tag{6}$$

3. Evaluation method

The evaluation of K-NN and Random Forest classifiers in classifying the weeds is done quantitatively using Confusion Matrix. The confusion matrix [31], gives us a summary of the prediction done on a classification problem. The number of correct and not correct predictions are presented with a count value and separated by each class in the confusion matrix. This is shown in table number 3. This essentially tells us how the classification model is confused while making the predictions.

Table 1: Confusion Matrix

Actual class	Predicted class	
	Positive	Negative
Positive	True Positive(TP)	False Negative(FN)
Negative	False Positive(FP)	True Negative(TN)

Table 1 shows confusion matrix for binary classification problem or when we have two-class classification problem. If the classifier outcome is the positive and actual case is also positive, then we have a true positive. If classifier outcome is the negative but actual case is positive, then we have a false negative. If classifier outcome is the negative and actual case is also negative, then we have a true negative. If the classifier outcome is positive but the actual case is negative, then we have a false positive. The table 2. shows evaluation parameters for the confusion matrix for a binary classification problem.

Table 2: Shows Evaluation Parameters for Confusion Matrix as Based on the Table No. 1

Evaluation Parameter	Formula	Describes
Accuracy	(TP+TN)/TP+TN+FP+FN	The ability of the classifier to correctly label the class
Recall or Sensitivity	TP/(TP+FN)	Compares positives with ones that should have been positives.
Precision	TP/(TP+FP)	Gives us how many of the positively predicted were indeed positives
Specificity	TN/(FP+TN)	How effective a classifier in identifying the negatives

F1-score	(2*TP) / (2*TP) + FP + FN	Relates the real positives with those given by classifier

Table 3: Shows Confusion Matrix for Multi-Class Classification Problem

Actual Class	Predicted Class		
	Paddy	Weed1(Grass)	Weed2(Sedges)
Paddy	Count1	Count2	Count3
Weed1(Grass)	Count1	Count2	Count3
Weed2(Sedges)	Count1	Count2	Count3

In the table 3., which shows the multi-class confusion matrix, the diagonal element gives you the true positives of respective classes. Here we have three-class confusion matrix. According to [32], to evaluate the performance of the classifier in multi-class classification case, for each separate class C_i, the TP_i, TN_i, FP_i, FN_i, $Accuracy_i$, $Recall_i$, and $Specificity_i$ can be calculated from the counts, $count_i$ from each class C_i. The performance of the classifier is calculated two ways one using macro-averaging and another micro-averaging. In case of macro-averaging, an evaluation parameter is the average of the same parameter but in case of the micro-averaging cumulative sum of counts to get the cumulative values of TP, TN, FP, and FN are obtained and then evaluation parameters are calculated. In this study, macro-averaging is used since this will make all classes treated equally whereas micro-averaging helps bigger classes. The calculation of evaluation parameters for multi-class classification is as shown in table 4.

Table 4: Showing Evaluation Parameters for Multi-Class Classification Problem. L Is Number of Classes

Evaluation Parameter	Formula	Describes
Accuracy	$\frac{\sum_i^l \frac{TP_i + TN_i}{TPi + TN_i + FP_i + FN_i}}{l}$	Average per class effectiveness of the classifier
Recall o	$\frac{\sum_i^l \frac{TP_i}{(TP_i + FN_i)}}{l}$	An average per-class effectiveness of a classifier to identify class labels
Precision	$\frac{\sum_i^l \frac{TP_i}{(TP_i + FP_i)}}{l}$	An average per-class agreement of the data class labels with those of a classifiers
Error Rate	$\frac{\sum_i^l \frac{FP_i + FN_i}{(FN_i + TN_i + FP_i + TP_i)}}{l}$	Average per-class classification error
F1-score	$\frac{\sum_i^l 2*TPi/(2*TPi + FPi + FNi)}{l}$	Relates the real positives with those given by classifier based on per-class average

4. Results and discussions

The extracted features were used to train K-NN, Random Forest, and LSSVM classifiers. All the classifiers were trained with the same data. The classifiers are trained to identify one species of plant among other plants and it does not matter if it is crop or weed. Once trained, images can be processed and test features can be extracted and given to classifiers for classification. The fig. 8. is showing the result obtained by K-NN classifiers, fig. 9. is showing the result obtained by Random Forest classifier, fig. 10. is showing the result obtained by LSSVM classifier Results obtained by the three classifiers are very similar. This consistent result shows that color features are enough to discriminate and classify paddy crop and sedges and grass-type weeds. Weeds in paddy fields can be broadly categorized as grass-type weed sedges and broad-leaved weed. However, if more than two types of weeds are found, then the proposed method is not enough for the discrimination of the crop and weeds. In this scenario, some more features are required for the discrimination. It is very uncommon to find more than two species of above said weeds together in the same field. If there are more than two types of weed, then it would be better to use multi-classifier system wherein each classifier may be trained with different features.

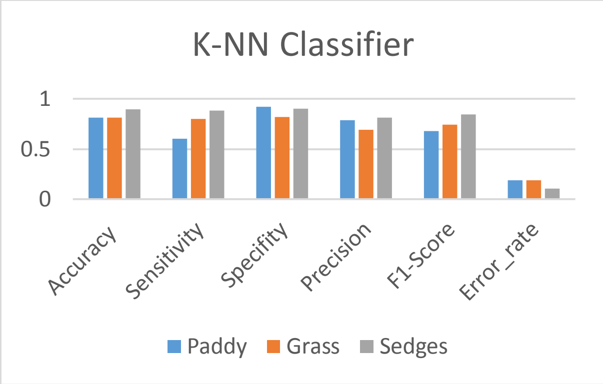


Fig. 8: Shows The Result Obtained by the K-NN Classifier.

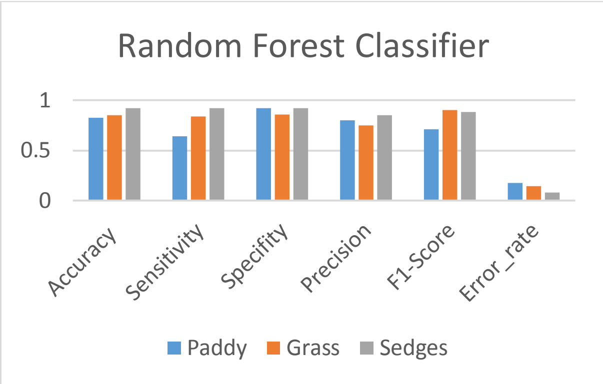


Fig. 9: Shows The Result Obtained by the Random Forest Classifier.

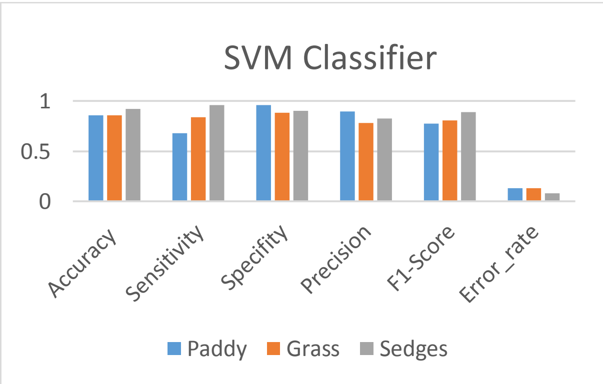


Fig. 10: Shows the Result Obtained by the LSSVM Classifier.

Even though it is difficult to compare the result of this study with other published work in the literature because each research work has been carried out for different crop under different field conditions, different lab conditions and boundary conditions, this research work showed that color features can be an important parameter in classification crop and weed. Color features can be considered in cases where crop and weed have similar shape features. Here, in this study, paddy crop and grass-type weed and sedges have similar shape features but slightly vary in color. In addition, the color feature extraction is based on YCbCr color space and thus makes robust to illumination variation. Therefore, taking color as discrimination criterion helped in classifying paddy crop and sedges and grass-type weed with very good accuracy.

5. Comparative analysis

There are no publicly available datasets for paddy crop and weeds. Therefore, dataset prepared by [5] is used for the comparative analysis of the proposed method. This dataset consists of carrot crop and weed. The result of using the proposed method on this dataset is as shown in the following fig. no. 11. Using the proposed method, carrot crop and weed is discriminated with higher accuracy of 91% when compared accuracy obtained in [5].

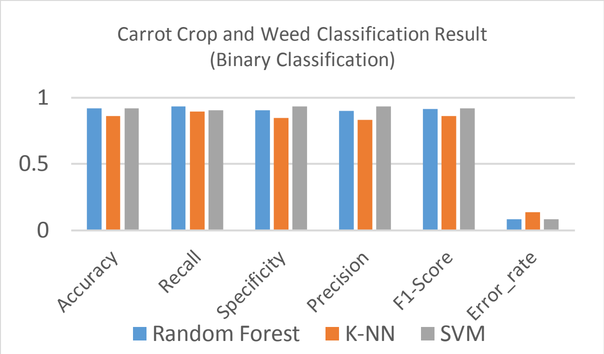


Fig. 11: Shows the Result of Carrot Crop and Weed Classification.

This shows that proposed method of extracting color features indeed gives a good result. This is due to the extracting of features around SURF point. This helps us getting enough information needed for discrimination of plant types. However, if there are more than two types weeds are found, then this method is not sufficient for the discrimination of the crop and weeds. In this scenario, some more features need to be extracted for the discrimination crop and weed.

6. Conclusion

In this research work, the creation of paddy crop and weed image dataset is explained for robotic vision applications in agriculture. Paddy crop and weeds such as sedges and grass-type weeds are classified using only color features. An accuracy of 86% is obtained. Compared to other published research work, the paddy crop and weed classification system presented in this research work is simple and is based on only color features which are robust to illumination variation, still achieving very good accuracy. Three classifiers were trained with the color features. The performance of the classifiers was evaluated based on confusion matrix. All the classifiers gave over 86% accuracy. This indicates that color feature is an important feature and is enough to discriminate paddy crop and sedges and grass type weed. Also, this system is not computationally intensive. As a part of future work, the novel approach of color feature extraction is used for other crop and weed discrimination and analyze the results. In addition, multi-classifier system will be built to classify the three broad categories of the weed found in paddy field, namely grass-type weed, sedges and broad-leaved weed.

References

[1] Adams, M. L., Cook, S., & Corner, R., 'managing uncertainty in site-specific management: what is the best model?' Precision agriculture, (2000), *2*(1), 39-54. https://doi.org/10.1023/A:1009984516714.

[2] linkedin.com, from cows to crops: Computer Vision for Precision Agriculture. Online available at https://www.linkedin.com/pulse/from-cows-crops-computer-vision-precision-agriculture-yulia-ivanov. Accessed on 30 November 2016.

[3] Mandal, S. K., & Maity, A., 'Precision farming for small agricultural farm: Indian scenario'. American Journal of Experimental Agriculture, (2013), 3(1), 200. https://doi.org/10.9734/AJEA/2013/2326.

[4] Young, S. L., Meyer, G. E., & Woldt, W. E.,' Future directions for automated weed management in precision agriculture'. In Automation: The future of weed control in cropping systems (2014), (pp. 249-259). Springer Netherlands. https://doi.org/10.1007/978-94-007-7512-1_15.

[5] Haug, S., & Ostermann, J., 'A Crop/Weed Field Image Dataset for the Evaluation of Computer Vision-Based Precision Agriculture Tasks'. In ECCV Workshops (4), (2014), (pp. 105-116).

[6] Ahmed, F., Bari, A. H., Hossain, E., Al-Mamun, H. A., & Kwan, P. 'Performance analysis of support vector machine and Bayesian classifier for crop and weed classification from digital images'. World Applied Sciences Journal, (2011), 12(4), 432-440.

[7] Lin, F., Zhang, D., Huang, Y., Wang, X., & Chen, X., 'Detection of Corn and Weed Species by the Combination of Spectral, Shape and Textural Features'. Sustainability, (2017), *9*(8), 1335. https://doi.org/10.3390/su9081335.

[8] Kadir, A., Nugroho, L. E., Susanto, A., & Santosa, P. I., 'Leaf classification using shape, color, and texture features'. arXiv preprint, (2014), arXiv:1401.4447.

[9] Hemming, J., & Rath, T., 'PA—Precision agriculture: Computer-vision-based weed identification under field conditions using controlled lighting'. Journal of agricultural engineering research, 78 (3), (2001), 233-243. https://doi.org/10.1006/jaer.2000.0639.

[10] Kumar, D. A., & Prema, P., 'A Novel Approach for Weed Classification using Curvelet Transform and Tamura Texture Feature (CTTTF) with RVM classification'. International Journal of Applied Engineering Research, 11(3), (2016), 1841-1848.

[11] Li, Z., an, Q., & Ji, C., 'Classification of weed species using artificial neural networks based on color leaf texture feature'. In International Conference on Computer and Computing Technologies in Agriculture, (2008), (pp. 1217-1225). Springer, Boston, MA.

[12] Kumar, D. A., & Prema P., 'A review on crop and weed segmentation based on digital images'. In Multimedia Processing, Communication and Computing Applications, (2013), (pp. 279-291). Springer, New Delhi. https://doi.org/10.1007/978-81-322-1143-3_23.

[13] Kodagoda, S., Zhang, Z., Ruiz, D., & Dissanayake, G., 'Weed detection and classification for autonomous farming'. Intelligent Production Machines and Systems, (2008).

[14] Agrawal, K. N., Singh, K., Bora, G. C., & Lin, D.,'Weed recognition using image-processing technique based on leaf parameters'. Journal of Agricultural Science and Technology. B, (2012), 2(8B), 899.

[15] Ehsanirad, A., & Sharath Kumar, Y. H., 'Leaf recognition for plant classification using GLCM and PCA methods'. Oriental Journal of Computer Science and Technology, (2012), 3(1), 31-36.

[16] Lottes, P., Khanna, R., Pfeifer, J., Siegwart, R., & Stachniss, C, 'UAV-based crop and weed classification for smart farming'. IEEE International Conference on Robotics and Automation (ICRA), (2017), (pp. 3024-3031). IEEE.

[17] Slaughter, D. C., Giles, D. K., Fennimore, S. A., & Smith, R. F, 'Multispectral machine vision identification of lettuce and weed seedlings for automated weed control'. Weed Technology, 22(2), (2008), 378-384. https://doi.org/10.1614/WT-07-104.1.

[18] Weis, M., & Gerhards, R., 'Feature extraction for the identification of weed species in digital images for the purpose of site-specific weed control'. Precision agriculture, 7, (2007), 537-545.

[19] Weis, M., & Sökefeld, M, 'Detection and identification of weeds'. In Precision, crop protection-the challenge and use of heterogeneity, (2010), (pp. 119-134). Springer, Dordrecht. https://doi.org/10.1007/978-90-481-9277-9_8.

[20] Guyer, D. E., Miles, G. E., Schreiber, M. M., Mitchell, O. R., & Vanderbilt, V. C. 'Machine vision and image processing for plant identification'. Transactions of the ASAE, (1986), 29(6), 1500-1507. https://doi.org/10.13031/2013.30344.

[21] Li, Z., an, Q., & Ji, C, 'Classification of weed species using artificial neural networks based on color leaf texture feature'. In International Conference on Computer and Computing Technologies in Agriculture, (2008), (pp. 1217-1225). Springer, Boston, MA.

[22] Burks, T. F., Shearer, S. A., & Payne, F. A, 'Classification of weed species using color texture features and discriminant analyses. Transactions of the ASAE, (2000), 43(2), 441. https://doi.org/10.13031/2013.2723.

[23] Gée, C., Bossu, J., Jones, G., & Truchetet, F, 'Crop/weed discrimination in perspective agronomic images'. Computers and Electronics in Agriculture, 60(1), (2008), 49-59. https://doi.org/10.1016/j.compag.2007.06.003.

[24] Hemming, J., & Rath, T. 'PA—precision agriculture: computer-vision-based weed identification under field conditions using controlled lighting'. Journal of agricultural engineering research, 78(3), (2001), 233-243. https://doi.org/10.1006/jaer.2000.0639.

[25] El-Faki, M. S., Zhang, N., & Peterson, D. E. 'Weed detection using color machine vision'. *Transactions of the ASAE*, 43(6), (2000), 1969. https://doi.org/10.13031/2013.3103.

[26] Gimp. G. N. U. Image manipulation program. 'User Manual, Edge-Detect Filters, Sobel'. The GIMP Documentation Team, 8(2), (2008), 8-7.

[27] Gonzalez, R. C., & Woods. R. E. Digital image processing, Edition. 2012.
[28] Bay, H., Tuytelaars, T., & Van Gool. Surf: Speeded up robust features. In European conference on computer vision, (2006), (pp. 404-417). Springer, Berlin, Heidelberg.
[29] Breiman, L. 'Random forests'. Machine learning, 45(1), (2001), 5-32. https://doi.org/10.1023/A:1010933404324.
[30] Suykens, J. A., & Vandewalle, J., 'Least squares support vector machine classifiers'. Neural processing letters, 9(3), (1999), 293-300. https://doi.org/10.1023/A:1018628609742.
[31] Suh, H. K., Hofstee, J. W., & van Henten, E. J. 'Improved vegetation segmentation with ground shadow removal using an HDR camera'. *Precision Agriculture*, (2017), 1-20.
[32] Sokolova, M., & Lapalme, G., 'A systematic analysis of performance measures for classification tasks'. Information Processing & Management, 45(4), (2009), 427-437. https://doi.org/10.1016/j.ipm.2009.03.002.