International Journal of Research in Engineering, Science and Management
Volume-2, Issue-7, July-2019
www.ijresm.com | ISSN (Online): 2581-5792

Performance Analysis of K-Nearest Neighbor Method for the Weed Detection

Gurpreet Khurana[1], Navneet Kaur Bawa[2]
[1]*M.Tech. Student, Department of Computer Science and Engineering, Amritsar College of Engineering and Technology, Amritsar, India*
[2]*Associate Professor, Department of Computer Science and Engineering, Amritsar College of Engineering and Technology, Amritsar, India*

***Abstract*: Weeds are among the major factors that harm crop yield. Image processing technique has become a promising tool for precise real time weed and crop detection in field. The various steps of weed detection are pre-processing, segmentation, feature extraction and classification. In the previous work, technique of linear interpolation is applied for the weed detection which results in the low accuracy. To improve accuracy for the weed detection, technique of region based segmentation, textural feature and KNN classification is applied. The proposed method is implemented in MATLAB and results are analyzed in terms of accuracy, precision, recall and execution time.**

***Keywords*: Linear Interpolation, KNN, K-mean**

1. Introduction

A method that aims to transform an image into a digital form by performing certain operations on it is known as image processing. Through this method an improved image is generated or certain amount of important information is extracted from that image which can be used in further applications [1]. The input of this method is given in the form of a vide frame or an image. An image or properties relevant to that image are given as output. Considering two images as 2-D signals and applying set of signal processing techniques on them are the general tasks of image processing. Commonly known as regions-of-interest, there are sub-images included in an image. Groups of objects, each of which depends on certain region, are included regularly in the images [2]. Generally, the images are available in a digitized form in image processing. Initially, a digital image is converted into an analog signal which is scanned to an output to show a digital image. The computer vision and computer graphics are very closely related to the image processing.

A plant that grows unintentionally along with the intentionally grown agricultural products is known as a weed. Any pant or vegetation which interferes with the objectives of forestry is defined as weed [3]. Thus, the identification and classification of weeds is important. Earlier, some men were employed to detect and remove the weeds from fields. Each and every plant was checked individually to detect the weeds. Then, either hands or spades were used to pluck the weeds manually from crops. To detect the weeds in different parts of world however, still manual power is being used [4]. The plants that grow in wrong place and consume the nutrients, light, water and space given to those plants are called weeds. The yield production can be reduced as a result. Also, disturbance in agriculture can be caused by using the machinery to remove these weeds from certain areas. The pests and diseases which can possibly be spread over the cultivated crops can also be hosted by weeds. Instead of adopting weed management techniques against an individual weed species, they can be designed for a specific group of weeds which is helpful through classification of weeds [5]. The nutrients crop can be extracted by weeds due to which the crops can be impacted by them to a great extent. The weeds that exist in between the rows of crops are detected through images in case of Weed Detection between Rows method. The weeds that exist in between the interval of two crops are detected through the Weed Detection between Crops method. However, image processing is applied if the crops are planted randomly or spread irregularly across the field [6]. An important step of image analysis is image segmentation. It is the process in which an image is divided into regions in correspondence to various objects. One of the various categorizes is assigned to each of the pixel present in an image. All the other stages of image analysis are made simpler when segmentation is performed in appropriate manner. When applying automatic segmentation algorithms, only partial success is achieved. However, these problems can be overcome by performing manual intervention. Depending upon the range of values in which pixels lie, the pixels are assigned to categories through thresholding based segmentation. Categories are divided such that one category includes the pixels that have values less than 128 and another category includes the remaining pixels [7]. Another type of segmentation is region-based segmentation in which the pixels that are neighbors are grouped together and the pixels with dissimilar values are split into groups. After performing preprocessing in character recognition systems, feature extraction is performed. Taking an input pattern and assigning it as one of the possible output classes in an appropriate manner is the major objective of pattern recognition [8]. The low level patterns are

represented as texture. They are categorized among the spectral, structural and statistical techniques. One of the most important approaches through which the tissue is characterized such that the changes in functional properties of certain regions can be defined is known as texture feature analysis. There are several methods through which texture feature can be extracted. The composition of well-defined texture elements is known as a texture as per the structural texture analysis techniques. The statistical distributions of pixels are used to identify the statistical features [9]. Higher-order moments of their gray scale histograms are used for describing texture of regions present in an image in case of statistical texture analysis techniques. Depending upon the weighted average of pixel intensities in their neighborhood, an empirical model of every pixel in the image is described through model based texture analysis techniques. The information is extracted from images, label and pixels of images using image classification. Several images of similar objects are needed for performing classification. The basics for an effective classification are an appropriate classification scheme and adequate amount of training samples. Generally, user requirements are the base of classification system [10]. There are supervised or unsupervised kinds of classification techniques. For image classification, the spectral signatures attained from training samples are used in supervised classification. However, the output relies on machine without any interaction with the user in case of unsupervised classification. The pixels belong to similar category are grouped into one class in this technique.

2. Literature review

S. Umamaheswari, et. al (2018) proposed a new method in which parallel processing was applied in GPU such that the previous systems could be improved in real-time scenarios [11]. To perform classification, the proposed system took real time image of farm as input and the type and location of weed in an image was detected. Under the deep learning framework, the images of crops and weeds were used to train the system along with classification and feature extraction. The automated weed detection system could use the achieved results such that tasks in precision agriculture could be performed.

Sarmad Hameed, et.al (2018) proposed that to capture the high quality of RGB images, UAV was applied through which the data acquisition of wheat crops in various phases can be performed [12]. Background subtraction was applied by the proposed approach to extract weed, wheat, and barren land in wheat crop field. It was still possible to provide certain improvements even though good detection mechanism was provided. Better results could be achieved by applying CNN even in stage where the colors of weed and wheat were similar. The problems arising because of variation in sunlight and shadows can be overcome by deploying better cameras.

Oscar Barrero, Diana Rojas, et.al 2016) presented a research in which the weed plants in rice fields could be detected from aerial images using neural networks [13]. On an autonomous electoral fixed wind plane, the images were collected from a height of 50 meters using the 16.1 megapixels CMOS digital camera. Further, due to the ortho-correction of an image, the pixel in-formation on final map was considered more reliable with the help of an orthomosaic map of field generated by stitching 250 images. For texture classification and Normalized Difference Index (NDI) for color, the GLCM was applied with the Haralicks descriptor. To detect the weeds, 99% of precision was achieved on the test data which indicated that a better performance on weed detection of rice fields was achieved by using neural networks.

Lawrence Charlemagne G. David, et.al (2016) presented a methodology in which the land usage in vegetable farms could be mapped automatically using a very high resolution aerial image which was collected from 100 m altitude [14]. An object-based technique was used to delineate the soil from vegetation along with the color index of vegetation extraction and Otsu's thresholding method. Further, the multi-resolution algorithm was applied to segment the vegetation subsequently. The GLCM was included to improve the classification which thus, resulted in increasing the kappa index of agreement along with overall accuracy. For the farmers and other management agencies, the output map served as guide such that the correct farm interventions could be known.

Inkyu Sa, et.al (2018) proposed a new method that could be applied on multispectral images gathered from a MAV to perform dense semantic weed classification [15]. An encoder-decoder cascaded CNN was designed which was named SegNet. Through this technique, dense semantic classes were inferred and the sugar beet and weed datasets were used with any number of input image channels and class balancing. Six models that included different numbers of input channels were trained and conditioned such that improvements could be achieved. For MAV investigation, an embedded Graphics Processing Unit (GPU) system (Jetson TX2) was tested.

Shubham Lavania, et.al (2015) proposed two techniques such that the crop row orientation of images from agricultural fields that had high weed pressure could be detected [16]. Also, it was possible to distinguish among the crop and weed through this method. Image segmentation, double thresholding based on 3D Otsu's method and crop row detection were the three processes applied initially to detect crop row detection using image processing. Secondly, the PCA was applied to classify the weeds and crops by compressing the 3D vectors of an image to one dimension. Towards the last step, the weed in crop rows could be detected as well as the weed from crop could be classified by combining the Otsu method and PCA. It was thus, concluded that for real-time applications that require weed detection, this approach provided better results.

3. Research methodology

Following are the general steps applied to detect weed using image processing:

a) *Image Acquisition:* To increase the accuracy in RGB format,

the images of weed are collected from online dataset or from the crop fields with the help of a high resolution camera. Within a respective size and in jpg format, every achieved image is stored.

b) *Pre-processing:* There are several factors based on which the obtained images are affected. The poor resolution of an image and unwanted background, the noise, and the lighting variations are few of such factors. To perform RGB to Gray scale conversion, certain pre-processing tools are applied. Further, the noise and unwanted objects from background are removed using the filtering techniques such that the gray scale images can be converted to binary image.

c) *Segmentation and feature extraction:* To detect the weed, the features are extracted after pre-processing. A process in which a set of features are defined such that an efficient representation can be seen for analysis and classification is known as feature extraction. There are various types of features existing in an image. The DWT algorithm is used for the feature extraction. The DWT algorithm uses the "Haar" wavelet for the feature extraction. The region based segmentation approach called k-mean is used for the segmentation. In the k-mean segmentation, the centroid points are selected and from that points Euclidian distance is calculated which can segment the data into certain classes.

d) *Classification:* For weed classification, this step is performed. An input is given to the classifiers in the form of feature vectors. The images of various weeds are used to train, validate and test the classifiers in classification process. Numerous classifiers have been designed by researchers over the years. The KNN classification method is applied for the classification. K-Nearest Neighbor classifier is a simple classifying approach in the machine learning methods where the classification is attained through the identification of the adjacent associates to enquiry illustrations and then utilizes those associates for determining the group or class of the doubt. In K-Nearest Neighbor classifier, the classification i.e. to which group the specified point belongs is relied on the computation of the least remoteness amid the specified point and other points. In the form of a classifier, the closest associate does not comprise any kind of training procedure. It is not appropriate in case of huge amount of training illustrations as it is not vigorous to noisy information. For the plant leaf categorization, the Euclidean distance amid the test patterns and training samples are computed. In this way, it discovers analogous ways and consequently the class for test patterns. A pattern is classified on the basis of maximum amount of votes from the k neighbors, by means of the pattern being allocated to the class mainly frequent amid its k nearest neighbors.

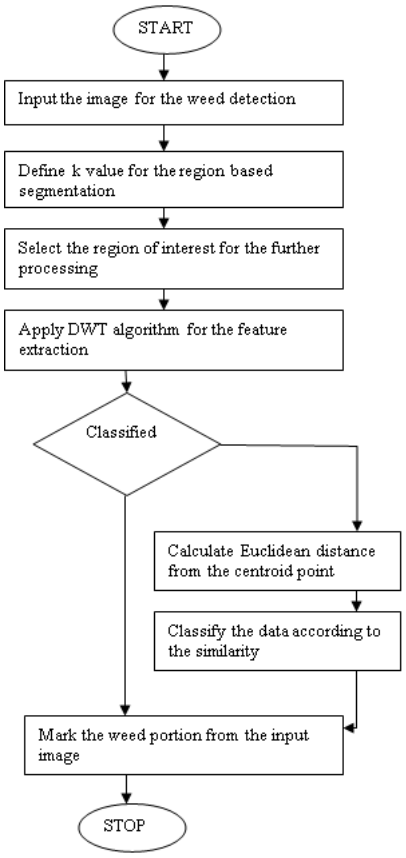

Fig. 1. Proposed methodology

4. Result and discussion

This research work is related to the weed detection, on sugar beet plant. The data set is collated from the different internet sources. The proposed approach has the various phases like pre-processing, segmentation, feature extraction and classification. The proposed methodology is implemented in MATLAB using computer vision and neural network toolbox. The performance of the proposed model is analyzed in terms of accuracy, precision, recall and execution time.

- *Precision:* In pattern recognition, information retrieval and binary classification, precision (also called positive predictive value) is the fraction of relevant instances among the retrieved instances.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

- *Recall:* Recall is the fraction of relevant instances that have been retrieved over the total amount of relevant instances.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

- *Accuracy:* Accuracy is defined as the number of points correctly classified divided by total number of points multiplied by 100.

$$Accuracy = \frac{Number\ of\ points\ correctly\ classified}{Total\ Number\ of\ points} * 100$$

- *Execution Time:* The execution time is the time which is taken by the algorithm to get completed.

The input images and comparison of Linear Interpolation and KNN classification is shown as below in terms of various parameters.

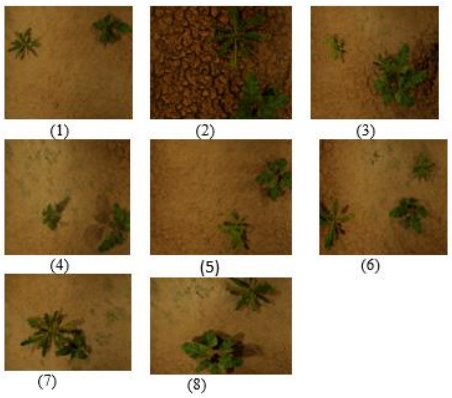

Table 1
Comparison in terms of Accuracy (%)

Image Number	Linear Interpolation	KNN Classification
1	88.26	91.26
2	88.15	93.02
3	89.23	92.94
4	89.21	91.37
5	90.34	91.83
6	88.49	92.54
7	90.37	93.50
8	89.08	93.14

Table 2
Comparison in terms of Precision (%)

Image Number	Linear Interpolation	KNN Classification
1	83.65	92.99
2	78.69	86.89
3	81.72	90.70
4	83.50	92.72
5	82.37	91.97
6	82.69	91.84
7	83.00	92.29
8	82.74	91.77

Table 3
Comparison in terms of Recall (%)

Image Number	Linear Interpolation	KNN Classification
1	72.52	75.38
2	93.47	99.88
3	80.08	83.79
4	73.08	76.34
5	77.43	79.02
6	76.20	79.50
7	74.99	77.89
8	75.99	79.76

Table 4
Comparison in terms of Execution Time (sec)

Image Number	Linear Interpolation	KNN Classification
1	0.4640	0.3876
2	0.2541	0.2515
3	0.2525	0.2325
4	0.2767	0.2182
5	0.2254	0.2075
6	0.2815	0.2363
7	0.2278	0.2221
8	0.2284	0.2115

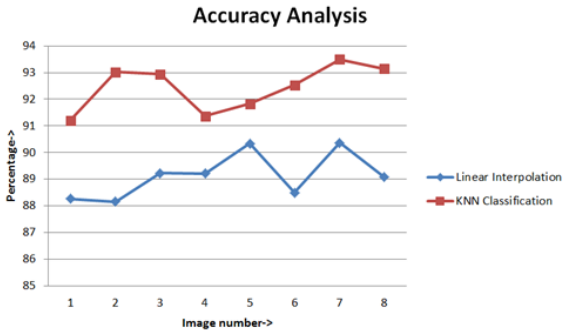

Fig. 2. Accuracy analysis

As shown in figure 2, the two algorithms which are linear interpolation and KNN are applied for the weed detection. It is analyzed that KNN classification method gives high accuracy for the weed detection as compared to linear interpolation.

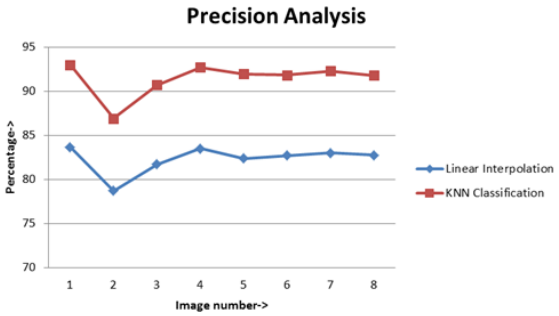

Fig. 3. Precision analysis

As shown in figure 3, the precision value of the linear interpolation method is compared with the KNN classification method. The precision value of KNN classifier is high as compared to linear interpolation.

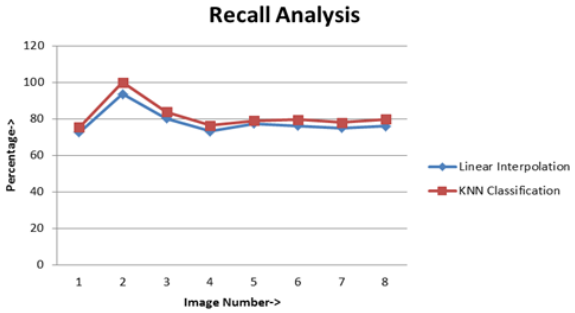

Fig. 4. Recall analysis

As shown in figure 4, the recall value of the linear interpolation method is compared with the KNN classification method. The recall value of KNN classifier is high as compared to linear interpolation

Fig. 5. Execution time

As shown in figure 5, the execution time of linear interpolation is compared with the KNN classification. It is analyzed that KNN classification method take less time for the weed detection as compared to linear interpolation.

5. Conclusion

In this paper, it is concluded that the Linear interpolation method [17] is complex and also has low accuracy for the weed detection. In this research work, the classification approach called KNN classification is applied for the weed detection. The performance of KNN classification is analyzed in terms of accuracy, precision, recall and execution time. It is analyzed that KNN classification method given high accuracy for the weed detection and results are optimized upto 8 percent as compared to linear interpolation method.

References

[1] Latha, A Poojith, B. V. Amarnath Reddy, G. Vittal Kumar; "Image Processing in Agriculture", IJIREEICE; 2014.
[2] Xavier P. Burgos-Artizzu, Angela Ribeiro, Maria Guijarro, Gonzalo Pajares; "Real- time image processing for crop/weed discrimination in maize fields"; Elsevier; 2010.
[3] Hossein Nejati, Zohreh Azimifar, Mohsen Zamani, "Using Fast Fourier Transform for weed detection in corn fields"; IEEE; 2008
[4] Grianggai Samseemoung, Peeyush Soni, Hemantha P. W. Jayasuriya, Vilas M. Salokhe; "Application of low altitude remote sensing (LARS) platform for monitoring crop growth and weed infestation in a soybean plantation"; Springer; 2012.
[5] Anup Vibhute, S. K. Bodhe; "Applications of Image Processing in Agriculture: A survey; International Journal of Computer Applications"; 2012.
[6] Alasdair McAndrew; "Introduction to digital image processing with MATLAB"; Course Technology; 2004
[7] Riya Desai, Kruti Desai, Zinal Solanki, "Removal of weeds using Image Processing: A Technical review" in International Journal of Advanced Computer Technology (IJACT).
[8] Latha, Poojith, Amarnath Reddy, "Image Processing in Agriculture" in International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering, June 2014.
[9] Tsampikos Kounalakis, Georgios A. Triantafyllidis, Lazaros Nalpantidis, "Weed Recognition Framework for Robotic Precision Farming" in IEEE 2016.
[10] Keun Ha Choi, Sang Kwon Han, Kwang-Ho Park, Kyung-Soo Kim and Soohyun Kim, "Vision based Guidance Line Extraction for Autonomous Weed Control Robot in Paddy Field" in IEEE Conference, Zhuhai, China, December 2015.
[11] S. Umamaheswari, R. Arjun, D. Meganathan, "Weed Detection in Farm Crops using Parallel Image Processing", 2018 Conference on Information and Communication Technology (CICT)
[12] Sarmad Hameed, Imran Amin, "Detection of Weed and Wheat Using Image Processing", 2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS).
[13] Oscar Barrero, Diana Rojas, Christian Gonzalez, Sammy Perdomo, "Weed detection in rice fields using aerial images and neural networks", 2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA).
[14] Lawrence Charlemagne G. David, Alejandro H. Ballado, "Vegetation indices and textures in object-based weed detection from UAV imagery", 2016 6th IEEE International Conference on Control System, Computing and Engineering (ICCSCE).
[15] Inkyu Sa, Zetao Chen, Marija Popović, Raghav Khanna, Frank Liebisch, Juan Nieto, Roland Siegwart, "weedNet: Dense Semantic Weed Classification Using Multispectral Images and MAV for Smart Farming", 2018, IEEE Robotics and Automation Letters, Volume 3, Issue 1.
[16] Shubham Lavania, Palash Sushil Matey, "Novel Method for Weed Classification in Maize Field Using Otsu and PCA Implementation", 2015 IEEE International Conference on Computational Intelligence & Communication Technology.
[17] Maurilio Di Cicco, CiroPotena, Giorgio Grisetti and Alberto Pretto, "Automatic Model Based Dataset Generation for Fast and Accurate Crop and Weeds Detection," 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).