A MACHINE LEARNING ALGORITHM FOR DETECTING DISEASE IN PADDY LEAVES

M. Jasmine
School of Agriculture, Ponnaiyah Ramajayam Institute of Science and Technology, India

Abstract

In general, the paddy blight affects all plant parts above ground. The effects of the disease shown in leaves, bark, node, neck, part of rays, and sometimes leaves sheath. The leaves are pale yellow to pale green with sharp tips, with eye-shaped lesions. The margins of these fractures are distorted and its center is gray or white. The extent of degeneration depends on the age of the crop, the time of disease onset and their type. As these lesions grow, the leaves gradually dry out. If the area where the leaves and sheaths join is affected, rot will appear on the neck and the leaves above the junction will die. The ankle will also be affected. It forms brown nodules and the stems may break off, occasionally dying throughout the nursery and young crops. In this paper, a disease detection machine learning approach was demonstrated. In this algorithm analyze the different changes in paddy leavess and link with the existing paddy images. Based on that this will identify the issues and provides the treatment details accordingly.

Keywords:

Paddy Leaves, Disease, Degeneration, Young Crops, Machine Learning

1. INTRODUCTION

The Shrub disease is a major challenge to paddy cultivation. The disease is endemic in more than 80 paddy growing countries [1]. The disease affects all stages of growth of the crop. The disease affects all parts of the leaves, neck and node. Small spots first appear on the upper part of the leaves. The spots then change from white to gray with eye-shaped spots with dry margins with a central center. Many points together form large irregular patches [2]. During a severe attack, the entire crop will look like it has been burned. This is what we call “disorder”. Crops will tilt as soon as the rays are released [3]. If these symptoms appear on the leaves, we call it “leaves blight”. The first gray to brown spots appear on the neck and then turn black, causing the rays to shrink and the rays to break and hang [4]. We call this “neck rot disease” and “neck rot disease”. The nodes turn black and break. We call this “node disorder”. Interstitial attack also occurs at the base of the crop, which causes the rash to appear [5]. Grains do not form if the neck area is infected before the radiation stage. In the case of post-radiation attack, the grains are formed but appear to be of low quality. [6] The spots on the rays and radishes are brown or dark brown. Depending on the variety of paddy, the size and shape of the spots will vary [7]. The disease is caused by the fungus. Airborne fungal spores (conidia) are the primary cause of the disease. It also spreads to other places through irrigation water [8]. It is also spread through crop residues, straw and seeds in the ground. The disease lives in the weeds in the range during the paddy growing season and spreads to the paddy crop during the paddy cultivation [9]-[11]. The disease is most prevalent when the sky is cloudy with moderate showers and humidity of 93 to 99 percent and when the night temperature is below 26°C [12].

The Invasion by viruses can cause the greatest damage, causing the greatest severity of the virus attack and causing great damage to the entire resulting crop, although the number of viral diseases is low [13]. For example, jaundice is a viral disease of the jaw. Jaundice is also caused by a virus. The virus is spread by whiteflies or green parasites. Viral infections are difficult to control after infection [14]. Controlling the spread of the disease and controlling the pests and burying the infected plants in the soil immediately are the main ways to prevent the virus spread [15].

Most plant diseases are caused by fungi. Different types of fungi attack crops in different ways. Although the fungal attack does not cause much damage, it does cause some damage [16]. Fungi cause many diseases such as leaves spot, wilt, rust, stem rot and root rot [17]. For example, rice blight is caused by a fungus. Paddy stalks tilt as they attack the stalk of the stalk. This can cause the most damage during the favorable season for the disease [18]. As the disease progresses, a large number of leaves fall off and cause yield loss. However, there are a number of beneficial fungi [19].

The Leaves spot disease and germs are caused by bacteria. For example, the citrus canker found in lemons is caused by bacteria. The diseased leaves and fruits have brown spots. If the disease is severe, the leaves spreads completely, the leaves lose their greenness and fall off, and the yield decreases due to reduced photosynthesis. Fruits are cheaper because they are cheaper. The nematodes attack the roots and form knots in the roots. Not only the root but also the central stem of the root swells and the root curve becomes wrinkled. The affected roots gradually rot and eventually cause wilt. In banana, nematodes attack the root and then gradually move and affect the tuber. Thus, the growth of the trees is affected and only small pods are produced. It is also likely to be accompanied by a fungal attack at the site of the nematode infestation. Thus, when combined with nematode and fungal infestation the rhizome of the tree is completely rotten and the root is weak. Such trees tend to bend easily when the wind blows.

2. LITERATURE REVIEW

In the paddy Plant diseases are spread by seeds, soil, air, water and insects. Seed diseases can be controlled by selecting disease free seeds and treating the seeds with saline solution. The Seed treatment by fungus controls soil borne diseases [11]. Generally, the diseases are prevented when the bio-film is mixed with water because the numerous beneficial microorganisms in the bio-film compete with the pathogenic microorganisms and destroy them [12]. About a bio-fence is essential to prevent airborne diseases. Life fences can be herbaceous or any other type of tree. It is better to grow trees like Neem, Pungan and Sita which are essential for natural insect repellent. Insects attack and damage crops and is the main cause of the spread of viruses [13]. The Juice-sucking insects transmit viral diseases. Aphids, whiteflies, mealy bugs,

leaves hoppers, aphids, scaly insects are the most important. When an insect drinks the sap of an infected plant, it multiplies the number of viruses that enter the insect body. When insects fly and travel to many places, the virus spreads throughout the crop to which they suck sap [14]. The tungro virus of rice is transmitted by herbivores. Therefore, when viral diseases occur, it is important to identify the pest that can transmit the virus and control the pest. Bacteria, fungi, etc. spread only under favorable conditions. Favorable conditions otherwise will only cause short-term damage [15].

3. PROPOSED METHOD

The proposed disease detection machine learning algorithm (DDMLA) was shown in the Fig.1. In order for the crops to grow healthily and give a timely flowering yield, the nutrients required for the crop must be present in the soil. Those nutrients should be in a position to be easily taken up by the roots. Such essential nutrients can be divided into nutrients and micronutrients. Nitrogen, Phosphorus, Potassium (NPK), Calcium and Magnesium are the five macronutrients (Macro nutrients).

- **Spot Identification:** Eye-shaped spots with dry margins with white to gray center on upper leaves. Many points together form large irregular patches. Gray to brown spots appear on the neck, turn black, the rays are shrunken / partially filled, the rays are broken and hanging. We call this "neck disorder".
- **Disorder Identification:** During a severe attack, the entire crop will look like it has been burned. This is what we call "disorder". Crops will tilt as soon as the rays are released.
- **Disease Identification:** The nodes turn black and break. We call this "knuckle disease". The spots on the rays and ray branches are brown or dark brown. Depending on the variety of paddy, the size and shape of the spots will vary.
- **Attack Identification:** Interstitial attack also occurs at the base of the crop, causing white ray symptoms to appear. Grains do not form if the neck area is infected before the radiation stage. But in the case of post-radiation attack, the grain is formed but found to be of low quality.

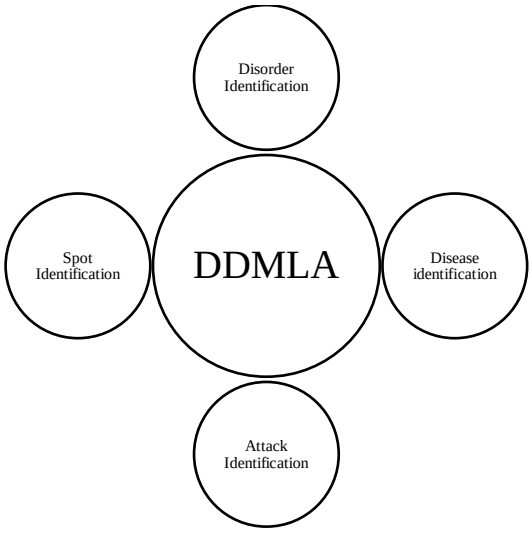

Fig 1: Proposed algorithms block

The Micronutrients include boron, zinc, magnesium, iron, copper, molybdenum, and chlorine. Although the contribution of micronutrients is only one percent, it is very important for essential functions such as crop growth, flowering, and water absorption. If we see signs of disease in crops, we need to look carefully at what caused those symptoms. Security measures should be determined accordingly.

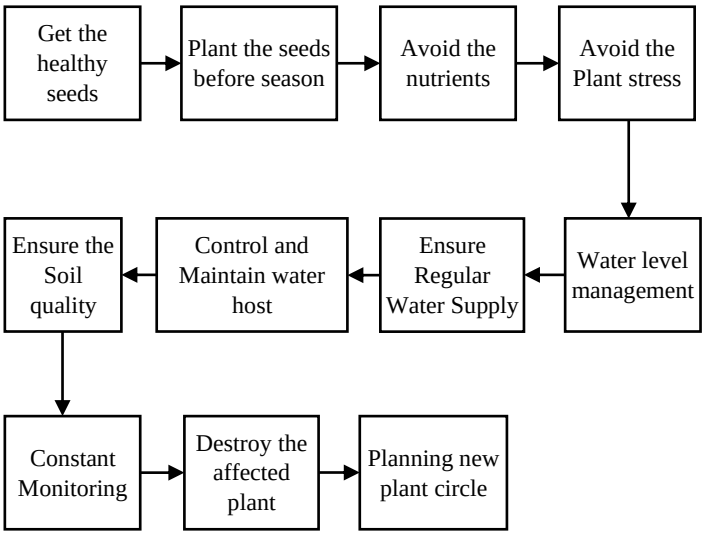

Fig.2. Proposed System

Algorithm 1: Proposed system design

Step 1: Use healthy or certified seed products.

Step 2: Cultivate resistant varieties.

Step 3: Sow the seeds before the monsoon season as soon as the rainy season begins.

Step 4: Avoid overuse of nutrients and use it in two or more portions.

Step 5: Water regularly and avoid stressing the plants.

Step 6: Maintain the amount of water required for the proper growth of the paddy crop.

Step 7: Provide regular watering to avoid drainage occurring too soon.

Step 8: Control and maintain weed transplanting hosts.

Step 9: If the soil is known to be silicon deficient, apply silicone fertilizer.

Step 10: Cheap sources of silicon include straw from rice paddies with high silicon content.

Step 11: Constantly monitor the fields for any signs of disease.

Step 12: Destroy all affected plant wastes to prevent the spread of fungi in the field.

Step 13: Planning the crop cycle is a simple and effective way to reduce the risk of disease.

Fungal reproductive systems, which are susceptible to the disease in stored paddy seeds and affected stalks, will transmit the disease from seed to the next season paddy crop. The Spread the fungus seeds through the air to other paddy fields over long distances. The Clusters of gonads are 2-4 in diameter, basal slightly swollen, and slender at apex. The conidia measure 20-22×10^{-12} micrometers. The Conidia slightly elongate and slender reach towards the apex.

4. RESULTS AND DISCUSSION

The proposed disease detection machine learning algorithm (DDMLA) was compared with the existing K-means-based segmentation (KMBS), Rice disease identification using pattern recognition (RDIPR), The recognition of cucumber disease based on image processing (CDBIP) and Analysis of automatic rice disease classification (AARDC)

There are the 4 parameters are evaluating the paddy leaves diseases. That is the accuracy, precision, recall and F1-score. Before understand the quality rate of the parameters, will know about the following,

- **Positive-T (PT):** It is the perfect predicted correct or above the calibration level.
- **Negative-T (NT):** It is the negative prediction values below the calibration level.
- **Positives-F (PF):** When the exact values are in calibration level and the predicted samples are in same level
- **Negative-F (NF):** When the exact values are in calibration level but the predicted samples are in different level

Accuracy is the parameter which describes the ratio between perfectly predicted paddy leaves input images from the given samples to the total number of collected image samples. When the rate of accuracy is high then the given output image sample getting high quality rate.

$$\text{Accuracy} = (PT+PN)/(PT+NT+PF+NF) \quad (1)$$

Table.1. Comparison of accuracy of different algorithms

Samples	KMBS	RDIPR	CDPIP	AARDC	DDMLA
1000	70.46	70.77	80.68	74.46	94.42
2000	71.75	71.52	85.30	77.86	94.52
3000	71.50	71.49	85.30	77.50	94.59
4000	71.37	70.67	84.83	76.31	94.63
5000	71.45	70.58	84.63	76.44	94.67
6000	71.44	70.45	84.37	76.46	94.70
7000	71.09	70.03	83.74	76.03	94.72

Precision is the ratio between the positive true samples and total true samples. The total true samples are calculated by the sum of positive true samples and false positive samples.

$$Precision = PT/(PT+PF) \quad (2)$$

Table.2. Comparison of precision of different algorithms

Samples	KMBS	RDIPR	CDPIP	AARDC	DDMLA
1000	71.32	71.83	84.01	77.78	94.41
2000	71.82	71.83	85.10	78.04	94.52
3000	72.57	72.66	86.24	78.61	94.58
4000	72.57	71.93	85.88	77.47	94.63
5000	71.52	70.82	84.35	76.45	94.67
6000	71.24	70.42	83.71	76.21	94.70
7000	71.96	70.99	84.29	76.86	94.72

Recall is the ratio between the positive true samples and the sum of positive true samples and false negative true samples.

$$\text{Recall} = PT/(PT+NF) \quad (3)$$

Table.3. Comparison of recall of different algorithms

Samples	KMBS	RDIPR	CDPIP	AARDC	DDMLA
1000	63.03	80.36	59.50	78.94	92.41
2000	61.37	74.50	66.34	72.76	92.52
3000	60.92	75.64	67.63	71.27	92.58
4000	65.61	74.50	69.77	68.03	92.63
5000	66.00	73.62	68.20	68.75	92.67
6000	65.84	72.42	66.58	68.88	92.70
7000	65.10	70.77	64.78	67.61	92.72

F1-Score is measured by the average sample values of precision and recall of the samples.

$$\text{F1-Score} = (2*(Recall*Precision))/((Recall+Precision)) \quad (4)$$

Table.4. Comparison of F1-Score of different algorithms

Samples	KMBS	RDIPR	CDPIP	AARDC	DDMLA
1000	66.97	75.90	74.75	78.36	91.41
2000	66.26	73.14	74.54	75.32	91.52
3000	66.33	74.12	75.80	74.78	91.58
4000	68.94	73.20	76.99	72.48	91.63
5000	68.67	72.20	75.41	72.42	91.67
6000	68.45	71.41	74.16	72.38	91.70
7000	68.39	70.88	73.25	71.97	91.72

The computation duration is nothing but the time taken to calculate the prediction of two different images.

$$Duration = (\text{No. of input samples})/(\text{Computation Speed}) \quad (5)$$

Table.5. Comparison of Computation duration of different algorithms

Samples	KMBS	RDIPR	CDPIP	AARDC	DDMLA
1000	1723	2324	-341	941	229
2000	1824	2525	-109	1074	309
3000	1925	2726	100	1328	411
4000	2027	2928	166	1524	515
5000	2129	3130	433	1684	599
6000	2234	3335	631	2195	708
7000	2337	3538	746	2501	789

5. CONCLUSION

In the later stages of development, severe shrub disease reduces the leaves surface area and consequently reduces the grain and yield. It is one of the most serious diseases that can destroy paddy crops. It is one of the most serious diseases that can destroy paddy crops. It may also affect other important agricultural grains such as wheat, rye, barley and rye. The fungus, which lives on straw after harvest, carries the disease to the next season. Crops are usually less susceptible to disease as they mature.

Cold temperatures, frequent rainfall and low soil moisture are favorable conditions for this disease. Long-term leaves moisture can lead to infection. In upland paddy fields, the soils that are subject to frostbite formation conditions (which vary greatly between night and day temperatures) are at high risk. Finally, crops grown in soils with high nitrogen or low silicon levels are more prone to disease. The proposed DDMLA was getting better results while compared with the existing KMBS, RDIPR, the recognition of CDBIP and AARDC.

REFERENCES

[1] N.V. Kousik, M. Sivaram and R. Mahaveerakannan, "Improved Density-Based Learning to Cluster for User Web Log in Data Mining", *Proceedings of International Conference on Inventive Computation and Information Technologies*, pp. 813-830, 2021.

[2] H. Azath, M. Mohanapriya and S. Rajalakshmi, "Software Effort Estimation using Modified Fuzzy C Means Clustering and Hybrid ABC-MCS Optimization in Neural Network", *Journal of Intelligent Systems*, Vol. 29, No. 1, pp. 251-263, 2018.

[3] H. Azath and R.S.D. Wahidabanu, "Function Point: A Quality Loom for the Effort Assessment of Software Systems", *International Journal of Computer Science and Network Security*, Vol. 8, No. 12, pp. 321-328, 2008.

[4] K. Lee and K. Hong, "An Implementation of Leaves Recognition System using Leaves Vein and Shape", *International Journal of Bio- Science and Bio-Technology*, Vol. 5, No. 2, pp. 57-66, 2013.

[5] T. Munisami, M. Ramsurn, S. Kishnah and S. Pudaruth, "Plant Leaves Recognition using Shape Features and Colour Histogram with K-Nearest Neighbour Classifiers", *Procedia Computer Science*, Vol. 58, pp. 740-747, 2015.

[6] X. Wang, J. Du and G. Zhang, "Recognition of Leaves Images based on Shape Features using a Hypersphere Classifier", *Proceedings of International Conference on Intelligent Computing*, pp. 87-96, 2005.

[7] J.S. Cope, D. Corney,J.Y.Clark, P.Remagnino and P.Wilkin, "Plant Species Identification using Digital Morphometrics: A Review", *Expert Systems with Applications*, Vol. 39, No. 8, pp. 7562-7573, 2012.

[8] J. Chaki and P. Ranjan, "Plant Leaves Recognition using Shape based Features and Neural Network Classifiers", *The Journal of Advanced Computer Science and Applications*, Vol. 2, No. 10, pp. 41-47, 2011.

[9] W. Lu, T. Ai, X. Zhang and Y. He, "An Interactive Web Mapping Visualization of Urban Air Quality Monitoring Data of China", *Atmosphere*, Vol. 8, No. 8, pp. 148-157, 2017.

[10] S. Nusser, L. Miller, K. Clarke and M. Goodchild, "Geospatial IT for Mobile Field Data Collection", *Communications of the ACM*, Vol. 46, No. 1, pp. 45-46, 2003.

[11] A.D. Vibhute, R.K. Dhumal, A.D. Nagne and S.C. Mehrotra, "Analysis, Classification, and Estimation of Pattern for Land of Aurangabad Region using High Resolution Satellite Image", *Proceedings of International Conference on Computer and Communication Technologies*, pp. 413-427, 2016.

[12] R. Olyazadeh, K. Sudmeier Rieu and M. Jaboyedoff, "An Offline Online Web-GIS Android Application for Fast Data Acquisition of Landslide Hazard and Risk", *Natural Hazards and Earth System Sciences*, Vol. 17, No. 4, pp. 549-556, 2017.

[13] A.D. Vibhute, K.V. Kale, S.V. Gaikwad and R.K. Dhumal, "Classification of Complex Environments using Pixel Level Fusion of Satellite Data", *Multimedia Tools and Applications*, Vol. 79, pp. 34737-34769, 2020.

[14] S. Ye, D. Zhu, X. Yao and Z. Zhang, "Development of a Highly Flexible Mobile GIS-Based System for Collecting Arable Land Quality Data", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 7, No. 11, pp. 4432-4441, 2014.

[15] C. Hu, X. Wang and F. Wu, "Motioncast: On the Capacity and Delay Tradeoffs", *Proceedings of ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pp. 18-21, 2009.

[16] Mostaque Md. Morshedur Hassan, "Current Studies on Intrusion Detection System, Genetic Algorithm and Fuzzy Logic", *International Journal of Distributed and Parallel Systems*, Vol. 4, No. 2, pp. 35-47, 2013.

[17] T. Amalraj Victoire and M. Sakthivel, "A Refined Differential Evolution Algorithm Based Fuzzy Classifier for Intrusion Detection", *European Journal of Scientific Research*, Vol. 65, No. 2, pp. 246-259, 2011.

[18] R. Ghadge, J. Kulkarni, M. Pooja, N. Sachee and R.L. Priya, "Prediction of Crop Yield using Machine Learning", *International Research Journal of Engineering and Technology*, Vol. 5, No. 2, pp. 31-37, 2018.

[19] Fabrizio Balducci, Donato Impedovo and Giuseppe.Pirlo, "Machine Learning Applications on Agricultural Datasets for Smart Farm Enhancement", *Machines*, Vol. 6, No. 3, pp. 21-38, 2018.