TEST Engineering & Management

November-December 2019
ISSN: 0193-4120 Page No. 6336 - 6341

Classification and Detection of Plant Diseases using higher order Dynamic Conditional Random Fields Through Spatial and Multitemporalimages

A. Gayathri, Dr. P. Abdul Khayum
A. Gayathri, M.Tech (CSP) student, Department of ECE, G.P.R.E.C, Kurnool, Andhra Pradesh, India.(Email: gayathriavvari24@gmail.com)
Dr. P. Abdul Khayum, M.Tech, Ph.D., Professor, Department of ECE, G.P.R.E.C, Kurnool, Andhra Pradesh, India.(Email: abdkhayum@rediffmail.com)

Article Info
Volume 81
Page Number: 6336 - 6341
Publication Issue:
November-December 2019

Article History
Article Received: *5 March 2019*
Revised: *18 May 2019*
Accepted: *24 September 2019*
Publication: *28 December 2019*

Abstract:
Agricultural areas should be continuously monitored, because they undergo random changes throughout the year. The problem arises when different crops show similar phenology and backscatter. This occurs when crops are classified based on single date remote detecting pictures. We should consider multitemporal images of crops. In this paper for classification of crops, we design an ensemble classifier which combines both spatial and temporal images of crops. To detect the affected areas of crops we implement first order and higher order dynamic conditional random fields (HDCRF). In order to enhance the diseased area of crop, we use k-means segmentation. All the training datasets are stored in multisvm (support vector machine). By the obtained results from this paper, we consider HDCRF as the best technique when compared with MRF (Markov Random Field) and CRF (Conditional Random Field) techniques.

Keywords: *phenology, backscatter, multitemporal, ensemble classifier, MRF, CRF, HDCRF, k-means segmentation.*

I. INTRODUCTION

Demand for food production has been raised enormously because of the increase in population. Farmers and strategy producers should have an idea that, how much amount of food production is required. But the major problem is the detection of affected crops by the farmers. It becomes difficult for farmers in order to classify the healthy and diseased crops, so weprefer remote detecting by radar to capture the images of the entire crop [1].

TerraSAR-X microwave images are used to detect the dynamic changes and phenology, which refers to stages that crops go throughfrom seeding to harvesting [2]. Different crops may show similar phenology, but they can be discriminated on their backscatter values. Some crops will exhibit different backscatter values for same type of crops. These changes can be found out by using radar detectors, which take into account different properties of crops like spatial and temporal images, backscatter values for each crop etc., so it is not sufficient to consider only single date remote detecting pictures, we must consider multitemporal images of crops. All these images are stored in an ensemble classifier. The ensemble classification technique is only used to classify the crops, but it does not detect the diseased crops. This paper fills this gap by detecting the affected crops by using higher order dynamic conditional random fields. By using HDCRF method, we can classify the crops as well as identify the affected crops in the field.

II. EXISTING MECHANISMS

Markov Random Field (MRF)

In the earlier works, they have proposed MRF[3] technique for incorporating spatial data in image classification. Later, this method was extended for ordering of the multitemporal pictures based on the Bayesian theory network. Based on the conditional probability dependencies of the dynamic data, MRF was considered as the useful method for classifying the temporal images rather than spatial images.The figure depicts the MRF concept.

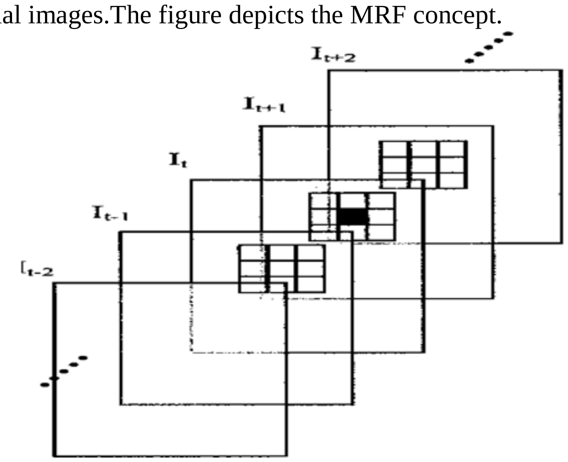


Fig 1 MRF based temporal dependencies of each plant in strawberry crop.

November-December 2019
ISSN: 0193-4120 Page No. 6336 - 6341

It gives the result of an individual crop at a time i.e., it calculates the cumulative distribution function (CDF) of a particular crop and is stored in group transitional matrix. MRF[4] was extended to provide multitemporal images of crops and has ignored spatial information. By skipping the spatial data, it is difficult to know the phenology of crops and also to identify the damaged areas of leaves. Another problem with this technique was, it does not support multiple fields means the entire crop area cannot be shown. It gives only the individual images of crops and which consumes more time while grouping the crops. So, for calculating spatial pixels, CRFs are proposed.

Conditional Random Fields (CRF)

CRFs [5] are proposed for spatial information of crops. Initially it was used for 1-D ordering and later expanded to 2-D image class [6]. This method calculates the probability distribution function (PDF) of the adjacent pixels in the given specimen. So it gives the overall layout of crop field as shown in the below figure.

Fig 2 Original image of entire plantation of strawberries considered as an input image.

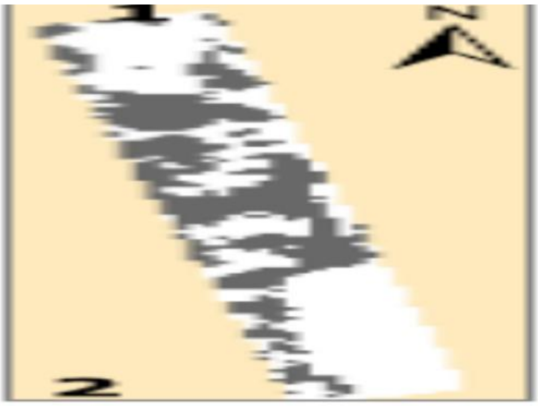


Fig 3 Outline map depicting the above image.

In fig 3 the black colour portion depicts the planted crops area, and the white colour portion shows the emptied area. By using this method, we cannot able to find the diseased plants, so we use Higher order Dynamic Conditional Random Fields. All these datasets are stored in local group transitional grid. Global class modulation matrix contains expert based phenological knowledge of diverse crops. While analyzing the spacial and dynamic pixels of a plant, they are compared with the patterns accumulated in global group grid. If the respective subsets of particular picture don't match with the data kept in predefined database, some error has been occurred during the processing of spatial and random nodes. Therefore categorizing resemblance of crops is decreased and we cannot determine the affected leaves in plants[7].In this paper, we develop a multi-dynamic spacial and temporal dependence model for grouping of crops and also discover the altered leaves present in plants.

Contribution

The main contributions that are considered for designing this model are

A functional interaction model is designed, where the spacial coordinates and temporal datasets specified by the user are compared with the information present in the database.

The remaining partof this paper is sorted out as follows. Section III introduces the proposed work. Section IV displays the results of the above technique. Conclusions for this work are manifested in section V.

III. PROPOSED WORK

A plant or leaf structure, which is piled up in input data points is selected by the client, for the purpose of crop association as well as to examine, whether that given plant is healthy or altered one.

Contrast Enhanced image

The intensity levels of a picture are ameliorated by making the difference in lightness. The purpose of executing the degree of diverse improvement in image is to batten the darker regions, so that we can easily form sets of crops and to expose, whether the leaves of plants are attacked.

Segmentation of Picture

For viewing the leaf outlines and the malfunctioned area, we prefer to divide the given object. The method used for this process is K-means clumping. By using this procedure, wecan extract the features from the region of interest.

K-Means Clustering Algorithm

1. Consider K number of lumps,which contains a set of input data points
2. (X1, X2…..Xn).
3. The algorithm starts by placing;K clusters at random locations in high dimensional vector space.
4. The below steps are computed iteratively.
5. For each input event (Xi), we find the nearest clump center by calculating Euclidean distance between every node and to all cluster centers.
6. That particular data element is assigned to the lump having minimum distance.
7. For each one of the input points, the above process is repeated.
8. The mean is reckoned for all repositioned context, and then divided by the total elements present in linear space.
9. After calculating the average, the clump centers are changed.
10. The above steps are performed recursively, until further division of data sets is not possible.

Segmentation Process

November-December 2019
ISSN: 0193-4120 Page No. 6336 - 6341

In this work, a consumer determined graph is taken from the predefined arguments, which is considered as the original object that is to be segmented. Here we assume three clusters.

- In first lump, the leaf epochs of the user referred file is measured by the Euclidean span formula. To obtain more precise outer layers of leaves, we calculate the mean of all the displaced pixels. The expression is formulated as

$$P_{ij} = \frac{\sqrt{\sum_{i=1}^{s} |g_i(x) - g_j(x)|^2}}{s} \quad (1)$$

- In second cluster, if the given image contains any disordered leaves, then that part is exaggerated, or else there will be no issues with the healthy plants.
- In third clump, the background region of the assumed object is captured.

In this way, the algorithm is applied for dividing the image into clusters.

Extraction of Features from ROISegmented Object

As our aim is to identify the affected crops, we consider the abnormal regions of leaves as our Region of Interest. The following features are processed from the selected cluster.

Affected space

The impacted leaves of assumed clump image can be calculated by using below formula.

Affected part of leaf= (Disease Attacked area) / (Total leaf region).

The below properties are computed from ROI clump.

Homogeneity

This template describes the closeness of distribution, among adjacent pels that are accumulated in the dim level grid. The equation is

Homogeneity = $\Sigma_i \Sigma_j \frac{S[i,j]}{1+|i-j|}$(2)

Energy

It is defined as the sum of squares of all spacial pixels, which are stored in co-occurrence grid. If the given subset has greater homogeneity, then energy is also high. The formula is given as

Energy = $\Sigma_i \Sigma_j \ S_{ij}{}^2$(3)

Contrast

It gives the deviation in intensity levels of a pel with its neighboring piscels over the total picture. Difference in brightness values should be low, so that image can be viewed appropriately. The expression is

Contrast = $\Sigma_i \Sigma_j$ (i-j) ²S [i, j] (4)

Inverse Distinct Moment (IDM)

It tells the homogeneity in the adjacent pels of the local gray array, which covers only a particular region of a file. It is represented by the following form

IDM = $\frac{\sum i \sum j \ S_{ij}}{1+(i-j)^2}$ (5)

Entropy

It calculates the loss of data, during processing of the spatial range. This is specified as

Entropy= $\Sigma_i \Sigma_j -S_{ij} *\log S_{ij}$(6)

Correlation

It describes the similarity between one pel to its side piscels over the entire portray.

Correlation = $\frac{\sum i \sum j \ (i,j) S(i,j) - \mu_x \mu_y}{\sigma_x \ \sigma_y}$ (7)

Mean

It calculates the average of all pixels, gathered in the local gray grid. This is evaluated as

Mean = $\frac{1}{N} \sum_{i=0}^{N-1} x_I$ (8)

Standard Deviation

It measures by how much amount each pel deviates from its center pel in the given image.

$\sigma^2 = \frac{1}{N-1} \Sigma_{i=0}^{N-1} (x_i - \mu)^2$ (9)

Skewness

It measures the symmetrical properties over the entire graph, by considering center pictels as reference point.

Kurtosis

It measures whether the data values are having high or low amplitudes, which are represented in Gaussian distribution.

From all the above features, the spacial interactions of pixels in the given subset are processed. The temporal relations of piscels are computed by using upcoming method.

Dynamic Conditional Random Fields (DCRF)

The goal of DCRF is to classify the given plant, and to specify the disease name if the leaves are imparted. It proposes the temporal function, which gives the graphical model having similar weighing pixels deviations at various timing slots. It defines the program based on the conditional probability dependencies of the multitemporal and spatial nodes. Consider the image (x), where c denotes the set of random variables and K represents the timing levels for the model at Δt = t±1.

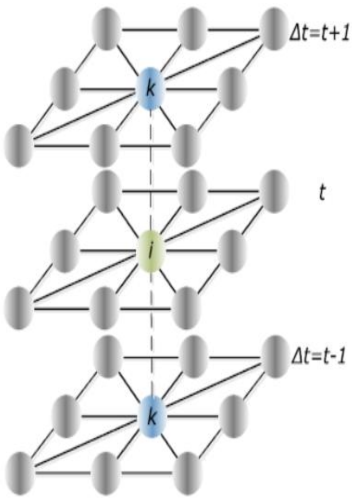


Fig 4: Spacial and temporal node relations at different instants of time.

November-December 2019
ISSN: 0193-4120 Page No. 6336 - 6341

The first order DCRF function is modeled by the following expression.

$$P(y/x) = \frac{1}{z(x)} \exp\{\sum_{i\in S} A(y_i, x) + \Sigma_{i\in S}\Sigma_{j\in N} I(y_i, y_j, x) + \Sigma_{t\in T}\Sigma_{i\in S}\Sigma_{c\in C} DP(y_{i,t,c}, x, x_{\Delta t})\} \quad (10)$$

Where DP represents dynamic potential, used for interpreting the altered plants in run time.By using first order DCRF, we get knowledge about sum of all the stochastic and spacial pixels, which gives the overall field outlay. The main drawback of this method is that, it cannot discriminate the plants with same phenology. So we go for HDCRFs.

Higher order Dynamic Conditional Random Fields(HDCRF)

Different crops with similar backscatter can be discriminated by using a transform technique called Discrete Wavelet form. By using this method, the lower and higher levels in the image are differentiated, that gives more exactness for association of plants. All the categorized groups of varying pel differing values are stocked in optimal classifier, where we can assign corresponding gray level weights to various crops based on expert knowledge. Hence by using HDCRFs, condition probability subordinates of random nodes are improved, so that distinguishing of leaf outlines can be easily identified. The expression is:

$$P(y/x) = \frac{1}{z(x)} \exp\{\Sigma_{i\in S} A(y_i, x)^{\Psi 1} + \Sigma_{i\in S}\Sigma_{j\in N} I(y_i, y_j, x)^{\Psi 2} + \Sigma_{t\in T}\Sigma_{i\in S}\Sigma_{k\in K} DP(y_i, y_k, x_{1:T}, t)^{\Psi 3} \quad (11)$$

Where Ψ1, Ψ2, Ψ3 are weights corresponding to association, interaction, dynamic potentials respectively.The individual or user defined plants can be determined by using the increasing order of DCRFs. The optimized collection classifier can be designed as depicted in the following diagram. Each plantdata is accumulated in an appraised database management system. Themodel and equation for reckoning is as follows:

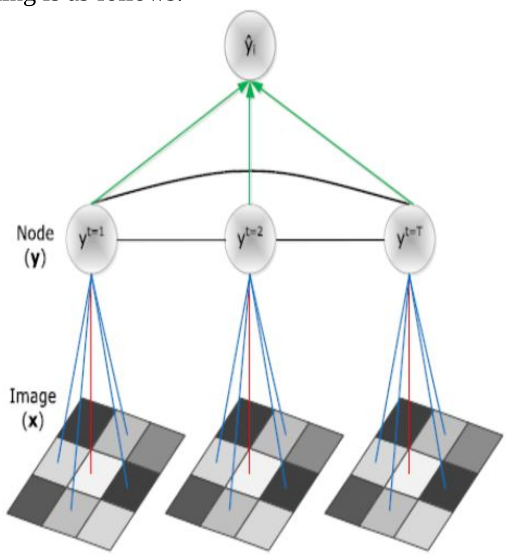


Fig 5: Estimated ensemble classifier of higher order DCRFs from the arrangement of dynamic nodes.

As shown in the above figure, the center pel collects the information from neighboring pels, and transmits it to other temporal images for improving accuracy. The below equation gives the user selected plant picture from the entire field. This overall crop layout was calculated by previous method named CRF.

$$\hat{y_i} = \arg_{l=1}^{m} \max \{\max_{t=1}^{T} (F[P(y_1|x), t]) \,.\, \text{User accuracy}\} \quad (12)$$

Where F[P (yr| x), t] is the conditional probability, representing to group r, having maximum F-score at time t. Here arg denotes atmost number of levels selected from the entire crop field, max gives the user defined plants. F-score is used for accumulating all the processed images, which is calculated as

F-score = 2(Predefined accuracy .User accuracy) / (predefined accuracy + User accuracy) (13)

In the above function, the entire expert based phenology crop images are grouped in predefined databank and User given input pictures are stored in training dataset for plant recognition purpose. Dot product in the above mentioned equation represents the comparison of point to point summation of the images based on corresponding producer and user accuracies.

If the given source picture matches with any one of the frames that are preserved in the database, then classification of plant along with the disease name can be obtained if that particular plant is abnormal, or else it forms a new template in the databank. All these input files, classes of various crops respective testing matrices and the corresponding resultant output attributes are aggregated in multisvm (support vector machine). The percentage of the affected region is calculated by using the linear kernel function. It computes the iterations until we get the exact value of the influenced part of leaf. So, by using segmentation process, HDCRFs and logical core we are able to classify the plants and recognize the disease name from the extracted features of ROI image. The results are shown in the upcoming section.

IV. SIMULATED RESULTS

The input picture is taken from the training dataset, which is shown below

Fig 6: Healthy leaves plant image

November-December 2019
ISSN: 0193-4120 Page No. 6336 - 6341

Using HDCRFs concept, we get to know which plants are healthy and diseased with the help of spatial and temporal images of crops. If the particular crop doesn't match with the normal plant as shown in below figure, then all the parameters are calculated which are defined in the proposed work. Finally the affected area of the leaves is computed.

Fig 7: Diseased leaf Image

Fig 8: Contrast Image

The above image gives the enhanced contrast form for given query picturewhich is in the altered way.

The segmented layouts are depicted below in the form of individual clumps.

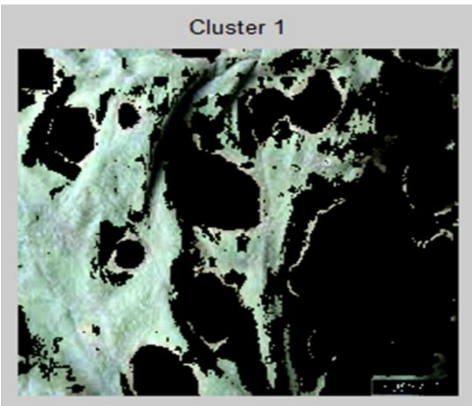


Fig 9: Outline of the given leaf

The above lump shows the leaf epochs, which are calculated by using Euclidean distance metric.

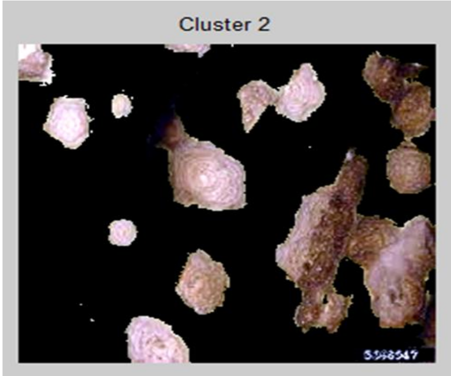


Fig 10 Affectedareas of given image

The infected areas are highlighted in overhead cluster. It is taken as Region of Interest by the user, to extract the features of input picture.

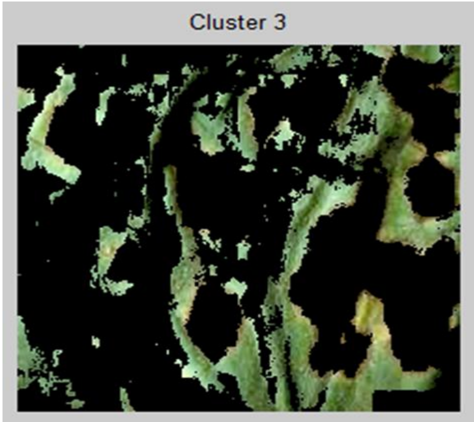


Fig 11

In the above layout, we can view the background scene of the given dataset.Here we took cluster 2 as our ROI image and performed the analysis.

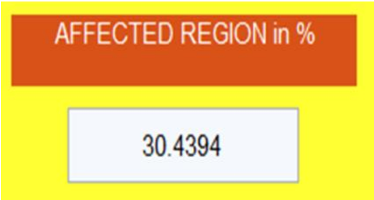
FEATURES	
Mean	71.8145
S.D	90.5264
Entropy	3.89944
RMS	9.44023
Variance	5343.55
Smoothness	1
Kurtosis	1.5368
Skewness	0.583604
IDM	255
Contrast	1.85335
Correlation	0.870152
Energy	0.332283
Homogeneity	0.871057

These are the features that are computed for the selected clump. From the evaluated data, we can find out the disease

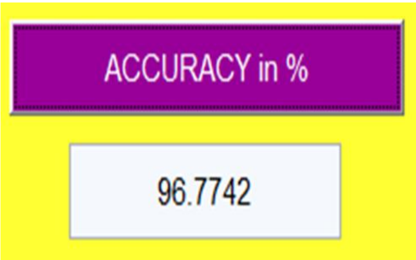
November-December 2019
ISSN: 0193-4120 Page No. 6336 - 6341

name, along with the affected area of the given diseased leaf which can be shown as

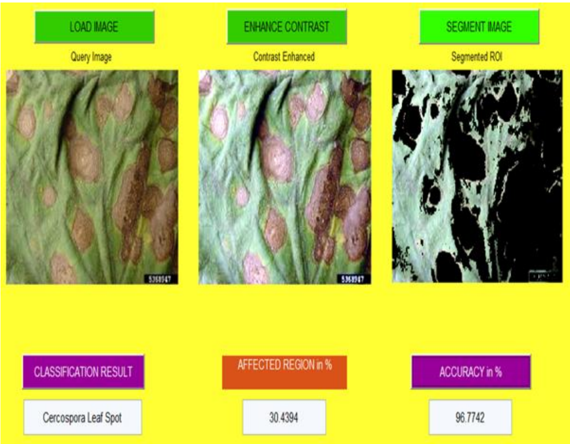
In this way, the given pictures are classified by considering all these aspects. Then we can know the altered part of the segmented area in terms of percentage as displayed beneath.



Linear kernel program is used to determine the accuracy of dysfunction region of the input subset, which can be viewed as



The entire output for the defined input image is displayed as



V. CONCLUSION

The main goal of this paper is to classify the plants based on their computed parameters and to acknowledge the name of the disease, if it has infected leaves. We also estimated the disordered region accuracy, so that it will be treated properly with the suitable pesticides. Therefore we can differentiate the defected crops from healthy ones in a field, which helps the farmers in the cultivation process.

TABLE 1
Comparison with previous methods in terms of accuracy for the above diseased plant.

Method	Accuracy
Markov Random Field	66.5
Conditional Random Field (CRF)	67.06
Dynamic Conditional Random Fields (DCRF)	75.78
Higher order Dynamic Conditional Random Fields max F1-score (HDCRF)	96.77

The accuracy is calculated from the above mentioned equations 10, 11, 12 for proposed work i.e., HDCRF. Accuracies for other methods are computed already from the existing papers. The accuracy can be improved by using Higher orders of Dynamic Conditional Random Fields.

REFERENCES

1. A. Kross, D. Lapen, R. Caves, “Early season monitoring of corn and soya beans with TerraSAR-X and RADARSAT-2,” Int. J. Appl. Earth Observ, May 2016.
2. S. Siachalou, G. Mallinis,, “A Hidden Markov models approach for crop classification: Linking crop phenology to time series of multi-sensor remote sensing data,” 2015.
3. OpenCV. Random Trees, accessed on Nov 20, 2014.
4. R.Sonobe, H. Tani, N.Kobayashi, “Discrimination of crop types with TerraSAR-X derived information.
5. B.K. kenduiywo, Spatial-Temporal Dynamic Conditional Random Fields Crop Type Mapping Using Radar Images, 2016.
6. D. Bargiel, U. Sorgel, “Spatial-temporal conditional random field crop classification from TerraSAR-X images,” 2015.
7. T. Hoberg, F. Rottensteiner, C. Heipke, “Conditional Random Fields for Multitemporal and Multiscale classification of optical satellite imagery,” Feb 2015.