AN OVERVIEW OF IMAGE ANALYSIS APPROACH ON PHENOTYPE FOR PLANT SYSTEM

Karthiga Rani.D [1]

ABSTRACT

In recent years, there has been a rapid advancement in computer vision technology which is much effective in extracting useful information from plant images in the field of plant phenomics. Phenomic approaches are widely used in the identification of relationship between phenotypic traits and genetic diversities among the plant species. The need for automation and precision in phenotyping have been accelerated by significant advancements in genotyping. Regardless of its significance, the shortage of freely available research databases having plant imageries has significantly obstructed the plant image analysis advancement. There were several existing computer vision techniques employed in the analysis of plant phenotypes. Conversely, recent trends in image analysis with the use of machine learning and deep learning based approaches including convolutional neural networks have increased their expansion for providing high-efficiency phenotyping of plant species. Thus, to enhance the efficiency of phenotype analysis, various existing machine learning and deep learning algorithms have been reviewed in this paper along with their methods, advantages and limitations.

***Index Terms*— Phenomics, machine learning, deep learning, phenotyping, convolutional neural networks**

I INTRODUCTION

Plant phenotyping is the analysis of the changes in the performance and appearance of plants caused due to genotypic variations and the surrounding environments to which the plants are exposed. Wide research in plant phenotyping is a crucial factor to meet the agricultural demands in the future and in enabling less usage of land, water and fertilizer. Image analysis is developing as an important technique in plant phnotyping. These phenotypic approaches are useful in the identification of interactions prevailing in in-between phenotypic traits and genetic diversities in plants. This could also be accomplished by analyzing the non-invasive quantitative parameters which reflect physiological states and traits all through the life of plants. The requirement for automation and precision in phenotyping has been accelerated by significant advancements in genotyping. The phenotyping rate is not found to match the growing genotyping rate thereby developing a bottleneck. The integration of effective DNA sequencing platforms and plant phenotyping afford opportunities for the exploration of genetic factors such as environmental tolerance, growth, disease resistance and production. For the mapping of genotype to phenotype ratio with the use of statistical methods like GWASs (genome-wide association studies) machine learning has been employed in different applications that include image analysis. The steps typically followed in image analysis are represented as follows: preprocessing or data augmentation, segmentation of the preprocessed data, feature extraction and classification. Algorithm on the basis of deep learning depicts most accurate results compared with conventional techniques like plant

[1] *Assistant professor and head of the department of computer science and IT, N.M.S.Sermathai Vasan College for Women, Madurai.*

CLEF. Further, machine-learning-based algorithm offers distinct features that help in the dissection of complex traits and in the determination of visual signatures linked with plant traits. The algorithm implied in plant phenotyping must be combined with reliable platforms and should process plenty of data developed in the experiments. So, this review focuses on the development of a robust, accurate and automated analytical algorithm for image analysis for the improvement of crop yield with respect to climate change and population demography.

This survey is structured as follows: Section II offers various state-of-art preprocessing techniques involved in the image analysis of plant phenotype, Section III explains various segmentation algorithms for image analysis of plant phenotype, Section IV provides the overview of Traditional feature extraction techniques employed for plant phenotyping, Section V provides a detailed depiction of classification approaches in plant phenomics, Section VI describles numerous datasets used in the existing approaches and Section VII concludes the study.

II DIFFERENT PREPROCESSING APPROACHES IN PLANT PHENOTYPING

Preprocessing is the initial step in image processing which is involved in the data organization that aids the succeeding steps to derive effective results. In particular, when field images under uncontrolled environment are employed, this preprocessing technique enhances the quality of image processing. Data transformation methods like conversion of gray scale, cropping steps, standardization and contrast improvement are implemented in preprocessing. To generalize the pattern analysis, variation in dataset images is increased by means of data augmentation. Scaling of images, addition of noise and flipping are the methods that are used frequently in data augmentation. This process increases the dataset size and reduces over-fitting.

In this study validation and training were achieved by splitting 285 labelled images into 45 validation and 20 training pairs. It was followed by three steps in which each image was made to undergo a flip a mirror, and a ninety-degree rotation so as to yield final dataset. This suggested work employed both augmented and original dataset correspondingly in order to test the efficiency of the augmentation. Preprocessing has been recognized as a reliable level for effective segmentation. [] This paper implemented an algorithm to reduce the noise of the plant improving the segmentation process. The noise had been reduced by color space conversion and binarization methods. At last this work compared the suggested algorithm with the existing algorithms.

III VARIOUS SEGMENTATION ALGORITHMS FOR IMAGE ANALYSIS OF PLANT PHENOTYPE

This section provides a detailed survey of various existing segmentation approaches employed for image analysis of plant phenotypes. enerally, segmentation represents the most significant and initial step. It is employed for the extraction of target information from the image data that are preprocessed by separating pixel sets in the image. It also helps in the recognition and quantification of specific plant organs automatically. [] presented a work on maize tassels in the wild using local count regression networking approach. It was essential to count the tassels of maize accurately so as to monitor the progress of maize plant growth. This work is done usually with the aid of man power and regarded as the most tedious task. In the modern plant phenotyping context, automation process was essential for facing a huge scale analysis of plant phenotype and genotype. Due to the emergence of increased computational resources and large-scale datasets, a computer vision technology turns out to be the most substantial innovation in recent years. Image-based approaches are gaining more support in studies related to plant systems. In order to count the maize tassels automatically, deep convolutional neural network and Tassel Net have been applied in this work for an effective segmentation process. However, this work proves its superiority it also have some drawbacks like reduced rate of counting performance.

described the segmentation of leaves in the plant phenotype analysis. The segmentation of each and every leaf in the specific image was the major contributing task in this work. Even though, the leaves share their characteristics of shape and appearance, there were some challenges like:

The availability of occlusions and variability at leaf shape and pose,

· The imaging conditions

The intention of this paper is to compare several solutions employed for leaf segmentation thereby conducting some typical phenotypic experiments. For the plant segmentation, a simple color-based approach has been employed from which the results are refined using raster graphics editing software tool. A segmentation process using 3D histogram could be carried through IPK pipeline approach which in turn employs distance amp and unsupervised clustering for segmentation. Also, watershed algorithm is applied to attain individual segmentation of leaves. Moreover, some approaches still lack individual leaf segmentation.

stated the difficulty of segmentation of individual leaves in the plant images. The accurate identification of the areas of individual leaves helps in enhancing the evaluation of biomass and plant performance. In plants, the counting of leaves is also the most substantial one for recognizing the progress of plant growth. A novel method has been developed for the detection of particular borders of leaf and to resolve incorrect segmentation of leaf due to leaf overlapping.

described the approach of in-field segmentation and the recognition of plant structures with the use of 3D image. An automated approach for mapping the 3D images in the outdoor field is presented for the process of segmentation. The segmentation techniques are then evaluated and analyzed effectively for both greenhouse and in-field environment.

These existing techniques attempt to prove segmentation accuracy thereby enabling the identification and quantification of plant organs automatically and estimating the yields of fruits and grains and biomass. This work also aims at enhancing the reproducibility in plant phenotype by means of swapping traditional human-dependent approach of phenotyping, every so often which is a labor-demanding and time-consuming aspect.

IV TRADITIONAL FEATURE EXTRACTION TECHNIQUES FOR PLANT PHENOTYPING

Feature extraction is a process of developing accurate and reliable data that are used for a satisfactory representation of images. Since the efficiency of the pattern analysis depends greatly on the feature quality, various approaches are under investigation for phenotyping.

The features are selected on the basis of object characteristics of images like gradient texture, pixel intensity and shape. []This existing study chooses features like color, shape, homogeneity, correlation and entropy from wheat for the classification. Of the various approaches involved, CNN has been proved to be very effective which is involved in the automatic extraction and classification of the features from the images. Previous research showed feature extraction on the basis of CNN advantages in plant phenotyping. It enables the learning of hierarchical features by means of taxonomical categorization on leaf datasets. [] The feature extraction in this paper provides data regarding leaf shape, leaf overlapping and leaf border. The study detects issues like less sensitivity in case of small leaves. Machine learning associated with IAP software has been employed using a greater number of phenotypic features. These outcomes are implied in the modification of leaf sensitivity in the suggested leaf labelling techniques. [] described a classification approach for plant phenotype using elliptical Fourier descriptor and the texture feature set named Haralick's texture descriptors for characterizing the plant seed for classification based on taxonomy. Using typical tool for the process of feature extraction, Scale Invariant Features Transforms (SIFT) is employed. This in turn performs as an invariant feature descriptor in terms of rotation, illumination, and viewpoint despite the scaling factor.

V STATE-OF-ART CLASSIFICATION APPROACHES IN PLANT PHENOTYPE

This section offers an overview of several existing approaches that are employed in the process of plant phenotype classification.

The outcomes of preprocessing, segmentation, and feature extraction, are attained in the classification stage. There are several existing machine-learning-based classification approaches. These ML-based approaches are applied hugely in phenotypes of plant recently. The classification techniques in turn highlight two main applications like classification based on taxonomy and physiological state of plant classification.

introduced a computer-vision-dependent methodology for the classification of wheat grains using artificial neural networks. A novel artificial neural network based on the multi-layer perceptron approach is employed for the perfect classification of the wheat grains into the durum or bread. The images are taken by camera of high resolution and are preprocessed initially. Then, the feature extraction process is carried out for the extraction of visual features like three color features and five textures. These visual feature datasets are regarded as the classifier's (ANN) input. The classifier model is trained using 180 grains, testing the accuracy level with 20 grains out of 200 wheat grains. These techniques could be integrated with industry easily for classifying the agricultural grains.

presented a computer-vision classification approach for grains with the use of adaptive Neuro-fuzzy inference system. About 200 wheat grain Images taken by high resolution camera are considered as input. The features are extracted on the basis of color (#3), texture (#5), and dimension (#4) and carried as an input for the classifier. The categorization of subset features are made for evaluating the classification effects.

Classification techniques that are based on computer-vision technology have been employed widely for the recognition of disease symptoms in plants. A technique of hyper spectral imaging is useful in the detection of downy mildew signs occurring in grapevine plants by *Plasmopara viticola []*. Nowadays, deep-learning-dependent approaches lead to enhancement in the accuracy and throughput in the plant disease symptoms analysis.

Revealed the probability of applying a deep CNN approach for the identification of 26 diseases appearing in 14 crops. The popular fine-tuned pre-trained deep CNN architectures are applied, like Google Net and Alex Net [], by a widely presented dataset containing 54,306-images of unhealthy and healthy plants from the Plant Village. A Transfer learning is also utilized for training CNN models so as to identify the crop disease symptoms like olives and is presented in

VI DATASETS EMPLOYED IN EXISTING PLANT PHENOTYPING ANALYSIS

Different datasets for phenotyping are available for the development of analytical techniques. A peculiar dataset is presented in Arabidopsis that can be used as a fundamental dataset for the development and growth of plant models in computer vision-based plant phenotyping. Further a plant CV website presents datasets for grass species. The significance of phenotypes, traits and gene functions could enable data integration from phenomics and genomics. [] In this paper standard datasets of annotated and raw rosette images are collected. The material of the plant, and image setup are also described in this paper, and are accessible at []. Mobile app Plantix that implies deep learning and crowd-sourced database facilitates customized preferences and diagnoses for the identification of plant diseases, deficiencies and pests on different crops worldwide.

Different datasets for phenotyping are available for the development of analytical techniques. A peculiar dataset is presented in Arabidopsis that can be used as a fundamental dataset for the development and growth of plant models in computer vision-based plant phenotyping. Further a plant CV website presents datasets for grass species. The significance of phenotypes, traits and gene functions could enable data integration from phenomics and genomics.

In this paper standard datasets of annotated and raw rosette images are collected. The material of the plant, and image setup are also described in this paper, and are accessible at

Mobile app Plantix that implies deep learning and crowd-sourced database facilitates customized preferences and diagnoses for the identification of plant diseases, deficiencies and pests on different crops worldwide.

Author	Technique	Methodology	Algorithm used	Advantage/ disadvantage
[17]	preprocessing	A preprocessing algorithm was presented to eliminate the noise and separate foreground from the 16 background which results the plant image to help the plant image segmentation. The preprocessing 17 is one of important levels has effect on better image segmentation and finally better plant's image 18 labeling and analysis. Our proposed algorithm is focused on removing noise such as converting the 19 color space, applyi ng the filters and local adaptive binarization step such as Niblack's.	Niblack's binarization	Niblack's method is more effective than the 298 other binarization such as Otsu, Sauvola, and Bernsen
[18]	Feature extraction	Deep phenotyping: deep learning for temporal phenotype/genotype classification	Fourier descriptor, and GLCM based feature extraction	Reduced rate of robustness

[6]	Segmentation/Feature extraction/ classification	Utilizing machine learning approaches to improve the prediction of leaf counts and individual leaf segmentation of rosette plant images	Leaf-Count Regression Model	Better Prediction of leaf counts and to segment individual leaves
[19]	Segmentation/classification	Watershed and supervised classification based fully automated method for separate leaf segmentation. Pixel classification based plant seg mentation and watershed based sep arate leave s segmentation.	Watershed algorithm	highest accuracy
[20]	Classification	An application of ANN model trained by artificia l bee colony (ABC) optimization algorithm was presented for classification of the wheat grains into bread and durum. ABC algorithm is used to optimize the weights and biases of three -layer multilayer perceptron(MLP)based ANN	ANN Trained by ABC Algorithm	It can achieve accuracy rate of about 100%
[21]	Classification	A database of images of approximately 960unique plants belonging to 12 species at sev eral growth stages is made publicly available. It comprises annotated RGB images with a physical resolution of roughly 10 pixels per mm.	naive Bayes classifier	It is capable of maintain constant result Disadvantage: it is sensitive to correlated attributes.

VII CONCLUSION

This survey provides the overview of various existing techniques implemented for the analysis of plant phenotypes. Computer-vision-based technology in plant phenotyping is a fast-developing field that combines knowledge of plant science, machine learning, mechanical engineering and spectral sensing. Highly efficient image analysis of plant phenotyping enables the experiments to perform with different genetic resources. Further recent advances in machine learning approaches play a vital role in the forecasting of traits by means of genotypic and phenotypic relationship. Progression in CNN-based algorithm and plant image dataset for plant phenotyping have achieved significant improvement in image analysis and taxonomic classification. This review has focused on the implementation of an efficient, accurate and automated analytical algorithm for image analysis.

REFERENCES

[1] K. Mochida, et al., "Computer vision-based phenotyping for improvement of plant productivity: a machine learning perspective," GigaScience, vol. 8, p. giy153, 2018.

[2] K. Keller, et al., "Soybean Leaf Coverage Estimation with Machine Learning and Thresholding Algorithms for Field Phenotyping," Proceedings of BMVC 2018, p. 0032, 2018.

[3] H. Attari and A. Ghafari-Beranghar, "An Efficient Preprocessing Algorithm for Image-based Plant Phenotyping," 2018.

[4] H. Lu, et al., "TasselNet: counting maize tassels in the wild via local counts regression network," Plant methods, vol. 13, p. 79, 2017.

[5] H. Scharr, et al., "Leaf segmentation in plant phenotyping: a collation study," Machine vision and applications, vol. 27, pp. 585-606, 2016.

[6] J.-M. Pape and C. Klukas, "Utilizing machine learning approaches to improve the prediction of leaf counts and individual leaf segmentation of rosette plant images," Proceedings of the Computer Vision Problems in Plant Phenotyping (CVPPP), pp. 1-12, 2015.

[7] P. Sodhi, et al., "In-field segmentation and identification of plant structures using 3D imaging," in 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017, pp. 5180-5187.

[8] K. Sabanci, et al., "Computer vision?based method for classification of wheat grains using artificial neural network," Journal of the Science of Food and Agriculture, vol. 97, pp. 2588-2593, 2017.

[9] M. Lo Bianco, et al., "Morpho?colorimetric characterisation of Malva alliance taxa by seed image analysis," Plant Biology, vol. 19, pp. 90-98, 2017.

[10] K. Sabanci, et al., "Grain classifier with computer vision using adaptive neuro?fuzzy inference system," Journal of the Science of Food and Agriculture, vol. 97, pp. 3994-4000, 2017.

[11] E.-C. Oerke, et al., "Hyperspectral phenotyping of the reaction of grapevine genotypes to Plasmopara viticola," Journal of experimental botany, vol. 67, pp. 5529-5543, 2016.

[12] S. P. Mohanty, et al., "Using deep learning for image-based plant disease detection," Frontiers in plant science, vol. 7, p. 1419, 2016.

[13] A. Krizhevsky, et al., "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097-1105.

[14] A. C. Cruz, et al., "X-FIDO: An effective application for detecting olive quick decline syndrome with deep learning and data fusion," Frontiers in plant science, vol. 8, p. 1741, 2017.

[15] M. Minervini, et al., "Finely-grained annotated datasets for image-based plant phenotyping," Pattern recognition letters, vol. 81, pp. 80-89, 2016.

[16] "http://www.plant-phenotyping.org/datasets."

[17] M. Henke, et al., "Comparison and extension of three methods for automated registration of multimodal plant images," Plant methods, vol. 15, p. 44, 2019.

[18] S. T. Namin, et al., "Deep phenotyping: deep learning for temporal phenotype/genotype classification," Plant methods, vol. 14, p. 66, 2018.

[19] D. Vukadinovic and G. Polder, "Watershed and supervised classification based fully automated method for separate leaf segmentation," in The Netherland Congress on Computer Vision, 2015, pp. 1-2.

[20] A. Kayabasi, "An Application of ANN Trained by ABC Algorithm for Classification of Wheat Grains," International Journal of Intelligent Systems and Applications in Engineering, vol. 6, pp. 85-91, 2018.

[21] T. M. Giselsson, et al., "A public image database for benchmark of plant seedling classification algorithms," arXiv preprint arXiv:1711.05458, 2017.