International Journal of Research in Circuits, Devices and Systems

E-ISSN: 2708-454X
P-ISSN: 2708-4531
IJRCDS 2021; 2(1): 66-70

www.circuitsjournal.com
Received: 07-11-2020
Accepted: 15-12-2020

Indraneel Parab
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India

Correspondence
Indraneel Parab
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India

Diagnosis of the crop disease based on machine learning

Indraneel Parab

Abstract
India is known for the multiple types of crops the farmers here yield. But out of all these multiple crops Wheat is the most important crop. According to the database the total area under the wheat crop is about 29.8 million hectares in the country. However, India's overall wheat yields per unit area of land are still well below the world average. Analysts suggest that to compete with other agriculture leading countries there is further room for development available for increasing wheat productivity in the country. According to world-grain.com India's wheat consumption is forecast to increase in 2020-21 to 93 million tonnes due to the immense rate at which our population is growing which will result in excessive demand in domestic supplies. Every year the amount of wheat production lost owing to plant diseases varies between 10–25%. To yield and supply such a huge amount of wheat farmers across the country need to enhance their farming techniques as soon as possible. This project is one of the ways through which they can succeed in doing so. Disease diagnose system is a project designed primarily for use in the rural agricultural area where farmers take wheat crops on average sized farm lands. This project will allow farmers to increase the crop yield by reducing the possibility of crops catching a disease. This system also provides the user (farmer) with solutions for given disease. Common diseases and their symptoms are listed out in this project using which we will create an algorithm which helps the farmer to identify whether or not his crop is affected with the disease. The algorithm will help in identifying the pattern and make decisions with minimal human intervention. This project will not only help the farmers in their cultivation but it will also stop the wastage of crops because the farmers can start the appropriate treatment as soon as possible once the correct disease is identified. This will make farmer's lives far easier.

Keywords: Wheat, disease diagnosis, machine learning, artificial intelligence, image processing, moments, hue moments, haralick, color histogram.

1. Introduction

Indian farmers have been using traditional methods to diagnose and cure crop diseases. Traditional methods that Indian farmers have been using mainly rely on specialists, experience, and user manuals, but the majority of them are expensive, time-consuming, and labor-intensive which makes it difficult to detect the disease precisely. It requires a lot of work, expertise in the plant diseases, and also needs a lot of processing time. Therefore, development of a rapid and accurate approach to identify plant diseases is the need of the hour which will be beneficial for business and ecology to agriculture. It's not moderate for wheat crop farmers to go to an agribusiness office and discover what the disease may be. The Principal objective of this system is to distinguish the illness introduced in a crop by watching its morphology by keen observation of pictures taken and analyzing them via machine learning. Many advanced technologies have been developed to reduce the crops' vulnerability towards dangerous diseases and to increase the productivity as much as possible. In this field there are many laboratory-based approaches such as polymerase chain reaction, gas chromatography, mass spectrometry, thermography and hyper spectral techniques have been used for diagnosis of the accurate disease. But still we'll have to consider that these techniques are not expensive because of the involvement of the laboratory apparatus and are therefore high time consuming, due to the involvement of the complex machinery and expertise required in handling such machines.

The next logical step in the field of new age agriculture seems to be developing such a system which can diagnose the crop diseases by just processing on input images. The farmer will just have to install cameras at specific areas in his/her farm. Which will capture the photographs at regular intervals. These cameras need not have much high configurations, any simple camera will do. Post processing will take care of any discrepancies in the images.

2. Background

In 1964-66, whole India was facing two major droughts which eventually led to food shortage and famines among India's increasing population. After such an unfortunate experience, many prominent scientists in the field of agriculture came forward and started contributing towards the betterment of the Indian agriculture scenario. During the early 1970's the 'Green Revolution' in India propelled the Indian agriculture sector to new heights. Indian farmers recorded all-time great productions across all the crops. Wheat was no exception to this. However, food security in India remains threatened by a number of factors including global climate change, the decrease in number of pollinators, plant diseases, increasing pollution and other such factors. Plant diseases at the same time being a threat to food security at the global scale, can also have disastrous consequences for smallholder farmers which constitutes a large chunk of Indian farmers whose livelihoods depend on healthy crops. It is estimated that around 30% of India's wheat production is lost due to diseases caught by the crop. Due to such immense loss Indian farmers might be losing out on a huge chunk of profit.

As of now India is self-sufficient in the case of wheat grain, but considering the population growth rate in the coming years this loss of grains might become a serious potential threat. And it will take no time to return to the era of food scarcity and famine. To avoid such a scenario, we will have to start taking instant measures towards better agriculture methods.

3. Understanding the idea of Image Processing Using Machine Learning

After looking at the complex nature of diseases and the multiple types of symptoms involved in the crops, it seems arduous to develop a technique which can imitate what the experts do. The obvious question arises whether our computer technology is developed enough to perform such advanced tasks which consist of multiple small tasks. This is where Artificial Intelligence comes into the picture.

What is Artificial Intelligence? It has become a major buzz word in tech-industry during the last few years and there has been a huge amount of hype around the term. Artificial intelligence is the newly developed area of computer science that tries to make machines that can help replicate the human brain, to do tasks that humans can do but which are historically very difficult for machines. This doesn't necessarily involve what we humans call as the intelligent things, for example we normally think that person has to be immensely intelligent or should be equally rated to beat a Grandmaster in the game of chess but the computer has already done that in 1990's by beating the then chess World champion Garry Kasparov. On the flipside some basic activities which all humans can do effortlessly in day-to-day life are still arduous challenges for AI like having a conversation about how your day was. Speech recognition in computer systems and responding perfectly during a conversation is something our scientists have been working really hard to achieve. Machine Learning has also created its own niche in the industry. Machine learning is one of the most rapidly growing fields of artificial intelligence concerned with the study and design of computer algorithms for learning good representations of data, at numerous levels of abstraction. Machine Learning uses Statistical algorithms to learn from vast amount of data. For example, if we want our machine to distinguish between a picture of cat and a dog then instead of programming a computer by describing every minute detail of how to differentiate, we teach it by giving it examples of what exactly it needs to do. So, for distinction between cat and dog pictures we can show our system lots of pictures of cats and dogs and it can learn to tell the difference between a cat and a dog. The machine will figure out a few recurring characteristics/properties in both types of pictures and form an algorithm to match these properties with input images. This is one of the naive ways through which we can teach our systems to distinguish between two types of images.

Generally speaking, image processing is altering an image in such a way to enhance it or extract information from it. By implementing Image Processing lots of properties can be extracted from a picture than visible to naked eyes. Machines can be taught to interpret images the same way our brains do and to analyze those images much more thoroughly than our normal human brain can. Like in the case of our current project we are going to extract information from the images of disease infected parts of crops. Generally, there are two ways of image processing:

- **Analog image processing**- It is used for processing physical printouts, digital photographs and other hard copies of images. It is applied on analog signals and it processes only two-dimensional signals. These signals can be periodic or non-periodic. The images are manipulated using special lenses or optical filters. In other words, the processing is done in a physical realm.
- **Digital image processing-** Digital image processing uses many of the same basic algorithms used in analog processing. It is used for altering digital images (a matrix of small pixels and elements) with the help of computer algorithms. It is amongst one of the fastest developing industries. In this process images are manipulated by storing them in a computer and then processing is done by using computational algorithms. Applying digital filters on specific parts of an image without touching other unnecessary parts is one of the most significant properties of digital processing.

Digital image processing is advantageous as compared to analog image processing. Such image processing allows us to use a huge number of algorithms than analog image processing. Noise and distortion are the two factors of an image which might become major obstacles during processing. Digital image processing plays an important role in overcoming these obstacles easily with the help of these available algorithms. Digital Image Processing (DIP) is a software which is used to alter the digital images by the use of a computer system. It is also used to enhance the quality of images, to extract some crucial information out of it. Similarly, it is also used during the conversion of image signals into digital images. Lots of algorithms are used in image processing. It provides a perfect platform to perform various operations like image enhancement, analog and digital signals image signals, voice signals etc. processing. This ability of computers to analyses an image can be leveraged and used in the field of disease diagnosis in plants and humans. Basically, we capture images of crop leaves and stems and all those useful parts of plants which can help by showing symptoms and feed them all to our system. Image processing console of the system then takes over the

whole operation and extracts all the possible information out of these images. This information is then compared with that of the already available database for the diseases and their symptoms. Then the matching algorithm is implemented between input data and database to predict the type of the disease.

4. Literature Review

There's not been enough study done in this field. Almost all the research done in the agricultural field in recent years has been about Satellite and GPS technologies, sensors, smart irrigation, drones, and automation. Although these fields are proved to be beneficial for new age farming, research in the field of disease diagnosis which is one of the key components of profitable farming is very rare. Still few researchers and scientists have initiated the research by taking into consideration a single crop at a time. Some of the earlier works done in this field include-

a) Godliver Owomugisha, John A. Quinn, Ernest Mwebaze and James Lwasa [1], proposed a method for automated vision based diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease. During this study they used nearest neighbors, Decision tree, random forest, extremely randomized tree, Naïve bayes and SV classifier.
b) SS Sannakki and VS Rajpurohit [2], put forth a "Classification of Pomegranate Diseases Based on Back Propagation Neural Network" which mostly focuses on the procedure of segmentation of affected parts and color and texture are used as the main characteristics. Using this technique resulted in a 97.03% accuracy. Disadvantage for this method was it was only used for limited crops.
c) Aakanksha Rastogi, Ritika Arora and Shanu Sharma [3], in their paper about Leaf Disease Detection and Grading using Computer Vision Technology & Fuzzy Logic" used digital image processing and machine vision technology. They proposed a system with two phases, in the first phase using the features of the leaves, the plant species is recognized, which includes pre-processing of leaf images, and feature extraction followed by Artificial Neural Network based training and classification for recognition of leaf. During the second phase of this method disease observed on the leaf is classified using K-Means based segmentation of defected area, feature extraction of defected portion and the ANN based classification of disease.

However, even after looking at all these research papers it is concluded that there's enough space for progress. This paper discusses all of the above studies and research papers and aims to fill within the gaps in previous analysis and errors like, conjointly exploring alternative prospects in fields like attainable medication, by continuing to research further in the subject.

5. Methodology

Like multiple other crops, it is observed that most of the diseases in wheat plants can be diagnosed by mere observation of leaves and stem. For example, Powdery mildew disease found in high humidity regions can be easily diagnosed by the formation of white, powdery patches on the upper surface of leaves and stem. Powdery growth later becomes black lesion and causes drying off of leaves and other parts [6]. Database containing the major wheat crop diseases and their symptoms will be used for comparison. Some of the most common diseases seen in wheat crops are- Powdery mildew, loose smut, Brown rust, Black rust. Most of the time these diseases depend upon various factors such as weather, soil type, water type, types of fertilizers used. Diagnosing a disease at the same time simultaneously keeping in mind all these factors become the most crucial thing. Crop growth is tracked 24/7 with cameras taking the pictures of leaves and stems periodically. These cameras can be placed in such a way that they capture the representative plant of many such plants which are part of a much larger piece of land. One or two such cameras can be installed in an average size farm; the number of cameras can be increased if a larger farming area is concerned. The pictures captured by these cameras are then compared with the already intensely analyzed pictures from the vast database available. The comparisons can be done via following three general features-

a) Shape of the leaves: In statistics, generally the moments have been used to analyze the shape characteristics of probability density or mass functions. Scientist Hu thought of an idea that moments could be used to characterize 2D image functions. Since his discovery many other scientists have come forward and proposed many types of moments and their invariants. The invariant is a function of the image that will not change or just change a little if we transform, i.e., rotate, scale, and blur, etc. the image. There are many invariants and each of them consist of different transformer functions. In image processing the outline (shape) of the subject is determined by image moment. Hu moments assist in outlining the leaves' outside edges. How these moments are calculated is beyond the scope of this paper.
b) Texture of the leaves: Texture segmentation has been an important aspect of image processing. Image textures is one of the many ways used to help in image segmentation or classification. Texture is an innate property of virtually all surfaces. Generally healthy leaves and diseased leaves have different textures. In computer graphics, mainly two approaches are there to analyze an image texture- Structured approach and Statistical approach. We'll use the Haralick texture property to differentiate between the textures of healthy and unhealthy leaves. Texture segmentation is nothing but dividing a picture into regions with different textures containing similar groups of pixels.
c) Colors observed: According to Wikipedia, in image processing and photography, a color histogram is a representation of the distribution of colors in an image. Like other kinds of histograms, the color histogram is a statistical parameter that can be used for measuring continuous distribution of color values. In simple words it can be said that it is a graph which represents intensity distribution.

These color histograms play a major role in image search engines. In such search engines it is assumed that images with similar color distributions have almost similar content. Distance metrics of type Euclidean, correlation, Chi-squared, intersection, and Bhattacharyya are used for comparison between two images. These histograms consist

of three separate histograms for red, green and blue channels. Histograms also come in use as a compulsory step before doing thresholding or edge detection.

6. Procedure

a) **Image acquisition:** Image acquisition is basically photographing the image with our camera device without which image processing cannot be done. Images taken are not processed at all. In simple words image acquisition is nothing but capturing a raw unprocessed image Common object or scene by an optical device into a required form for further processing and analysis. It is a fundamental step in image processing as no other step can be performed without acquisition of an image.

b) **Filtering:** RAW and unfiltered images contain noise. This noise can affect the texture and colour of the image. So, we need to remove the noise using a suitable filter. This is where filtering comes into the picture. Filtering is a modification and enhancement technique for an image. It consists of smoothing, edge enhancement and sharpening. It is used for highlighting specific features or for removing unnecessary features.

c) **Segmentation:** Image segmentation is done to focus our study on a particular part of the photograph or to isolate an image into multiple small images to study them separately. Image is segmented into multiple small parts. In other words, image segmentation is partitioning an image in such a way that the parts formed are meaningful and do not have anything in common between them. In a broader sense segmentation is done via three different ways- threshold, edge and region-based method.

As far as the crops are concerned, perfect features of soil are also crucial for successful cultivation. Important features of soil are colour as well as the texture of soil, power of hydrogen (PH) tested in laboratory, organic matter content. Colour of the soil leads us to very crucial features of soil such as mineral composition, age, existence of moisturized organic matter, crop species which are suitable for that part of land etc. [4]. Unlike in the case of leaf colors, Munsell soil colour notation system is used instead of the image moments system for the analysis of soil images. In this notation system each colour consists of three characteristics hue, value and croma [4, 5]. Hue refers to main wavelengths reflected by soil. Colors such as red, orange, yellow, etc. In other words Hue answers the question- Is a color red, green, blue or something in between? Hue helps us to distinguish between these three colours. Colors which have a hue are called “Chromatic colors”. And the colours which do not have a hue are called “neutral colors”. Value signifies the light or dark shade of the colour in comparison to neutral grey scale. It answers the question- Is color light or dark? The value parameter is applied to the chromatic as well as neutral colors as mentioned above. Chroma alludes to the intensity or brightness of colour, the saturation or brilliance of a color. Each color shares a logical and visual relationship with all other colors in Munsell color notation. It answers - How bright our colour really is? The low croma colors are called “weak”, whereas those of high croma are highly saturated/ strong. The scale of croma is visually uniform. For example- Fluorescent coloured materials sometimes may have a croma value as high as 30.

7. Scope in Future

During the practical use of this system there will be many challenges due to complex environmental conditions [7, 8, 9, and 10]. Obstruction in the field of view of the camera, totally different diseases with similar types of symptoms these are some of the potential problems that might arise during successful processing of this model. The model which successfully overcomes all these problems of environmental complexity will get an accurate identification result in practical application. This model will not only come in handy in adapting to these complex environments, but also increases the accuracy percentage for identification.
Same technique can further be used for various other crops with an available database for the diseases and their symptoms. Even the farmers living in rural areas will not have to keep frequently visiting an agribusiness office or keep waiting for experts to come and diagnose the disease. With adequate and timely resources in hand, these farmers can adopt this system which will result in their own economic gain. Taking in consideration the current scenario of this technology, it is not possible to implement it in human disease diagnosis. Once fully developed and tasted in the field of agriculture, this diagnosis system can be implemented in our bio medical sciences also. Such practice will help in decreasing the workload of doctors. As we have already seen in the last couple of years, the medical system alone is incapable of handling the pandemic efficiently. During the Covid-19 pandemic this incapability became more visible due to the immense amount of workload doctors and medical professionals had to lift. Implementing this system can somehow decrease this intense pressure, because then the patients can diagnose their mild illness at their own house. Which will help in reducing the over crowdedness at the hospitals.
In summary, further researchers are recommended to refer articles [11-20] to know more about machine learning, deep learning/ some other essential technologies and their importance in several sectors in detail.

8. Conclusion

New developments in agricultural technologies today are important for attaining sustainability goals in agriculture. Satellite and GPS technologies, sensors, smart irrigation, drones, and automation - these are some of the upcoming technologies which we can expect to see implemented in the next 10-15 years. Ultimate aim of such a method is to observe and infer about exceptions that are found in plants. Implementing this method of diagnosis will provide a major boost in the field of new age agriculture technologies. Compared to the traditional techniques used earlier, this system of diagnosis will not only increase the accuracy but will also help in improving crop management. Which will ultimately result in rapid development of Indian farmers. Due to industrial advancement and plenty of job opportunities in those similar sectors, the younger generation in India is getting attracted to the industrial field. Due to which, the number of people coming in the agri. fields are seen to be depleting. People have simply left farming and turned to something else for a stable income. Because of such a situation it is observed that one of the many problems faced by our farmers is unavailability of agriculture experts. This system, if used and developed to its fully ultimate capacity, can be easily handled by a common man that is our Indian farmer. This will compensate for the non-availability and scarcity of agricultural experts in rural

areas.

9. References

1. Godliver Owomugisha, John Quinn A, Ernest Mwebaze, James Lwasa. Automated Vision-Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease, Preceding of the 1'st international conference on the use of mobile ICT in Africa, 2014. https://www.researchgate.net/publication/333683045 Automated Vision Based Diagnosis of Banana Bacterial Wilt Disease and Black Sigatoka Disease
2. Sannakki SS, Rajpurohit VS. Classification of Pomegranate Diseases Based on Back Propagation Neural Network, International Research Journal of Engineering and Technology (IRJET). 2015 May;2:02. https://www.semanticscholar.org/paper/Classification-ofPomegranate Diseases Based on Sannakki Rajpurohit/5c52974c84aee3adb6c1eeb6787f4a2808bf681a
3. Aakanksha Rastogi, Ritika Arora, Shanu Sharma. Leaf Disease Detection and Grading using Computer Vision Technology & Fuzzy Logic 2nd International Conference on Signal Processing and Integrated Networks (SPIN), 2015. https://ieeexplore.ieee.org/document/7095350
4. Maniyath SRR, Hebbar AKN, ALS, Subramoniam SR. Soil Color Detection Using Knn Classifier, International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), 2018, 52-55. Doi: 10.1109/ICDI3C.2018.00019. https://ieeexplore.ieee.org/abstract/document/8437087
5. S Ramesh *et al*., Plant Disease Detection Using Machine Learning, 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), 2018, 41-45. Doi: 10.1109/ICDI3C.2018.00017. https://ieeexplore.ieee.org/document/8437085
6. Wheat diseases and symptoms Available at: https://vikaspedia.in/agriculture/crop production/integrated-pest-managment/ipm-for-cerels/ipm-strategies-for-wheat/wheat-diseases-and-symptoms
7. Kumari S, Vani V, Malik S, Tyagi AK, Reddy S, Abraham A, *et al*. Analysis of Text Mining Tools in Disease Prediction. In: (eds) Hybrid Intelligent Systems. HIS. Advances in Intelligent Systems and Computing, 2020, 1375. Springer, Cham. https://doi.org/10.1007/978-3-030-73050-5_55
8. Gudeti B, Mishra S, Malik S, Fernandez TF, Tyagi AK, Kumari S. A Novel Approach to Predict Chronic Kidney Disease using Machine Learning Algorithms, 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, 2020, 1630-1635, doi: 10.1109/ICECA49313.2020.9297392.
9. Amit Kumar Tyagi, Dr. Meenu Gupta, Aswathy SU, Chetanya Ved. Healthcare Solutions for Smart Era: An Useful Explanation from User's Perspective", in the Book "Recent Trends in Blockchain for Information Systems Security and Privacy, CRC Press, 2021.
10. Nair MM, Tyagi AK, Sreenath N. The Future with Industry 4.0 at the Core of Society 5.0: Open Issues, Future Opportunities and Challenges, International Conference on Computer Communication and Informatics (ICCCI), 2021, 1-7. doi: 10.1109/ICCCI50826.2021.9402498.
11. Kute SS, Tyagi AK, Aswathy SU Tyagi AK, Abraham A, Kaklauskas A. Industry 4.0 Challenges in e-Healthcare Applications and Emerging Technologies. In: (eds) Intelligent Interactive Multimedia Systems for e-Healthcare Applications. Springer, Singapore, 2022. https://doi.org/10.1007/978-981-16-6542-4_1
12. Kute SS, Tyagi AK, Aswathy SU. Security, Privacy and Trust Issues in Internet of Things and Machine Learning Based e-Healthcare. In: Tyagi AK, Abraham A, Kaklauskas A. (eds) Intelligent Interactive Multimedia Systems for e-Healthcare Applications. Springer, Singapore, 2022. https://doi.org/10.1007/978-981-16-6542-4_15
13. Madhav AVS, Tyagi AK. The World with Future Technologies (Post-COVID-19): Open Issues, Challenges, and the Road Ahead. In: Tyagi AK, Abraham A, Kaklauskas A. (eds) Intelligent Interactive Multimedia Systems for e-Healthcare Applications. Springer, Singapore, 2022. https://doi.org/10.1007/978-981-16-6542-4_22
14. Nair MM, Kumari S, Tyagi AK, Sravanthi K. Deep Learning for Medical Image Recognition: Open Issues and a Way to Forward. In: Goyal D, Gupta AK, Piuri V, Ganzha M, Paprzycki M. (eds) Proceedings of the Second International Conference on Information Management and Machine Intelligence. Lecture Notes in Networks and Systems, Springer, Singapore, 2021,166. https://doi.org/10.1007/978-981-15-9689-6_38
15. Tyagi AK, Fernandez TF, Mishra S, Kumari S. Intelligent Automation Systems at the Core of Industry 4.0. In: Abraham A, Piuri V, Gandhi N, Siarry P, Kaklauskas A, Madureira A. (eds) Intelligent Systems Design and Applications. ISDA 2020. Advances in Intelligent Systems and Computing, Springer, Cham, 2021,1351. https://doi.org/10.1007/978-3-030-71187-0_1
16. Goyal Deepti, Tyagi Amit. A Look at Top 35 Problems in the Computer Science Field for the Next Decade, 2020. 10.1201/9781003052098-40.
17. Varsha R, Nair SM, Tyagi AK, Aswathy SU, RadhaKrishnan R. The Future with Advanced Analytics: A Sequential Analysis of the Disruptive Technology's Scope. In: Abraham A, Hanne T, Castillo O, Gandhi N, Nogueira Rios T, Hong TP. (eds) Hybrid Intelligent Systems. HIS 2020. Advances in Intelligent Systems and Computing, 2021, 1375. Springer, Cham. https://doi.org/10.1007/978-3-030-73050-5_56
18. Tyagi Amit Kumar, Nair, Meghna Manoj, Niladhuri, Sreenath, Abraham Ajith. Security, Privacy Research issues in Various Computing Platforms: A Survey and the Road Ahead, Journal of Information Assurance & Security. 2020;15(1):1-16.
19. Tyagi Amit Kumar, Rekha G. Machine Learning with Big Data (March 20, 2019). Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur - India, 2019 Feb; 26-28.
20. Gillala Rekha, Amit Kumar Tyagi, V Krishna Reddy. A Wide Scale Classification of Class Imbalance Problem and its Solutions: A Systematic Literature Review, Journal of Computer Science. 2019;15(7)886-929. ISSN Print: 1549-3636.