PROJECT REPORT ON

"Identification of Medicinal plant using Deep Convolutional Neural Networks"

Submitted by:

Brahmbhatt Parita P. [170130111014]

Dixit Vyoma R. [170130111022]

Under The Guidance of:

Prof. Suhas Patel

Department of

Electronics & Communication Engineering

MAY 2021

DEPARTMENT OF ELECTRONICS & COMMUNICATION ENGINEERING

GOVERNMENT ENGINEERING COLLEGE

GANDHINAGAR - 382028

GUJARAT TECHNOLOGICAL UNIVERSITY, AHMEDABAD

GOVERNMENT ENGINEERING COLLEGE, GANDHINAGAR

DEPARTMENTOF ELECTRONICS AND COMMUNICATION

CERTIFICATE

This is to certify that Miss. ***Brahmbhatt Parita Pragneshkumar*** *of Government Engineering College, Gandhinagar, Enrollment no.* ***170130111014*** *has completed the project work titled* ***Identification of Indian Medicinal Plants using Deep Convolutional Neural Networks*** *as prescribed by Gujarat Technological University during the academic term 2020-21.*

Signature:
Name: Prof. Suhas Patel
(Project Guide)

Signature:
Name: Dr. K. G. Maradia
(Head of Department)

GOVERNMENT ENGINEERING COLLEGE, GANDHINAGAR

DEPARTMENTOF ELECTRONICS AND COMMUNICATION

CERTIFICATE

This is to certify that Miss. ***Dixit Vyoma Rakeshkumar*** *of Government Engineering College, Gandhinagar, Enrolment no.* ***170130111022*** *has satisfactorily completed the project work titled* ***Identification of Indian Medicinal Plants using Deep Convolutional Neural Networks*** *as prescribed by Gujarat Technological University during the academic term 2020-21.*

Signature:
Name: Prof. Suhas Patel
(Project Guide)

Signature:
Name: Dr. K. G. Maradia
(Head of Department)

INDEX:

ABSTRACT:

In Atharva Veda, Ayurvedic species and its usages are explained. Different kind of plant species are used as medicine in Ayurveda for various disease. The recognition of medicinal plant is done manually currently. But as Ayurveda is getting popular worldwide it is necessary to develop an accurate system which can recognize the plant automatically. In recent years, identification of plant using Deep Neural Networks has increasingly become mature. The identification of plants using conventional method is time consuming, complex and frustrating for non-experts and according to researchers, 80% Indians use Ayurveda as medicine in various diseases. Hence, it is important for people to know which medicinal plant they are using. The availability of digital cameras and smart phones have allowed the idea of automatic medicinal plant identifier to become a reality. In this system, our trained model can identify six types of Indian Medicinal plant species, which are Neem, Basil, Mango leaf, Lemon Leaf, Curry Leaf and Nyctanthes Arbor – Tristis.

INDEX TERMS: Computer vision, Feature Extraction**,** OpenCV**,** Botany, Ayurveda**,** Leaf classification, Image processing

INTRODUCTION:

Ayurvedic medicinal treatment is a part of India's great history. Ayurveda is developed by hermits in Vedic period or Vedic age. It is originally shared as an oral tradition from one generation to another. Ayurveda was recorded almost 5000 years ago in Sanskrit, in one of the sacred texts called as Vedas in 1200-1000 BCE. Ayurveda mainly consists of healthcare techniques developed using natural ingredients of plants or trees like its flowers, roots, barks and leaves. The benefits of different plant products are explained in detail and how to use that is explained in Ayurveda. In India almost 8500 plants products have medicinal benefits. For making of Ayurvedic medicine almost 1500 plants and its products is used. Currently, large Ayurvedic medicine making firms like Patanjali and Ayush use nearly 500 plants for making their daily products. Almost 80% plants which is used in herbal medicine are found in forest and barren areas. While, rest 20% plants are cultivated in agricultural land and at home.

In early time, Ayurvedic experts used to go to forest and barren places to pick the medicinal plants and used to prepare medicines for their patients. As this Ayurvedic knowledge is transferred orally over generations it is very difficult to recognize different plant species now. And the manual identification methods mainly depend on people's subjective judgement. Botanist use different plant characteristics as identification keys to identify the medicinal plants. However, identification of medicinal plant from field observation necessitate considerable botanical expertise, which put it on the far side of the reach of most aficionados of nature. Manual plant species identification is almost impossible for ordinary people and effortful even for professionals who deal with botanical challenges every day, such as farmers, horticulturists and landscape architects have difficulty identifying plant species. This issue is called the taxonomic crisis in the field of related research. In order to overcome such problems an automated medicinal plant recognition system is needed to be developed.

In this information era, the technology of target recognition based on image processing has become accessible to people due to the development of digital image processing and pattern recognition, etc. The theoretical base and technical preparation for leaf identification provided by these techniques. Thus, in recent years, study on

plant recognition using image processing techniques has become an area of interest and many researches have been conducted by international researches ever since.

In past studies, features of plant identification were elected from plant organs such as flowers, leaves, fruits, and stems. But most of the plants can be identified by its leaves. Because every plant's leaf is different from the other plant's leaf in its shape, size, style or texture. So, we can use these qualities to develop an automated plant classification system. Hence, most of the plant can be identified be performing image processing on the image of plant leaves. This process is done with semi-supervised method in machine learning. However, this method is time consuming and difficult.

At present, we can develop a digital image processing and leaf pattern recognition method. Because in past years, performance of deep learning convolutional neural networks has become excellent, development of an automatic medicinal plant recognizer is possible. As digital cameras and portable devices are cheaper and have higher resolution this process can be used to develop a digital image processing systems.

The application of deep learning methods in plant identification has achieved an excellent performance as compare to most of the manual feature extraction classification method. The process of system development will consist of domains like biology, medicinal, data science and computer science coding. In deep learning you do not have to manually point feature extraction.

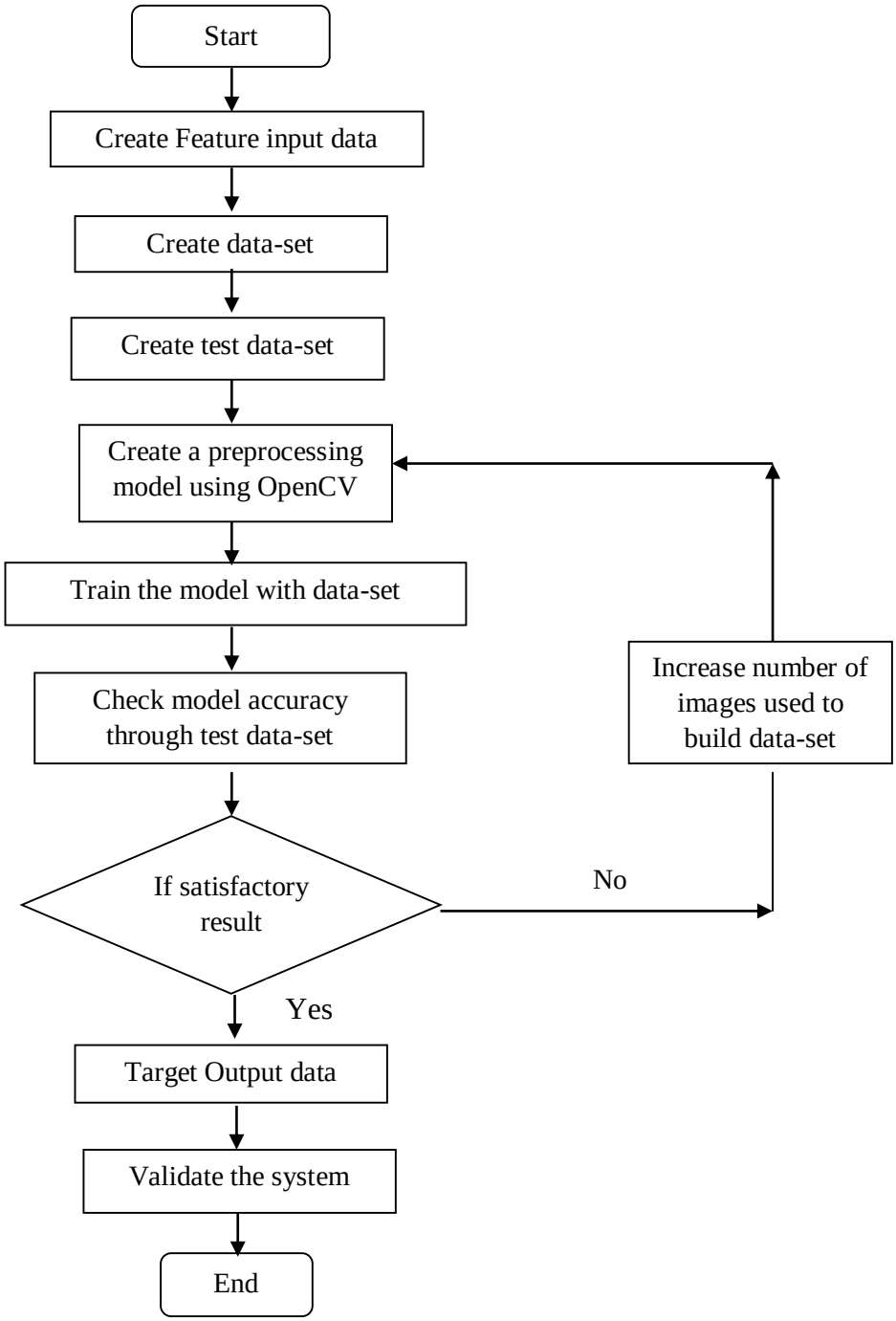
For leaf recognition a data set has to construct for each plant species manually to train the system for recognizing the plant species. The accuracy of the system will depend on the data-set which is created. The accuracy will be mostly directly proportional to the number of images means a greater number of layers and its means network will be deeper. The data-set to be properly structured with perfect data for increasing accuracy.

PROPOSED SYSTEM:

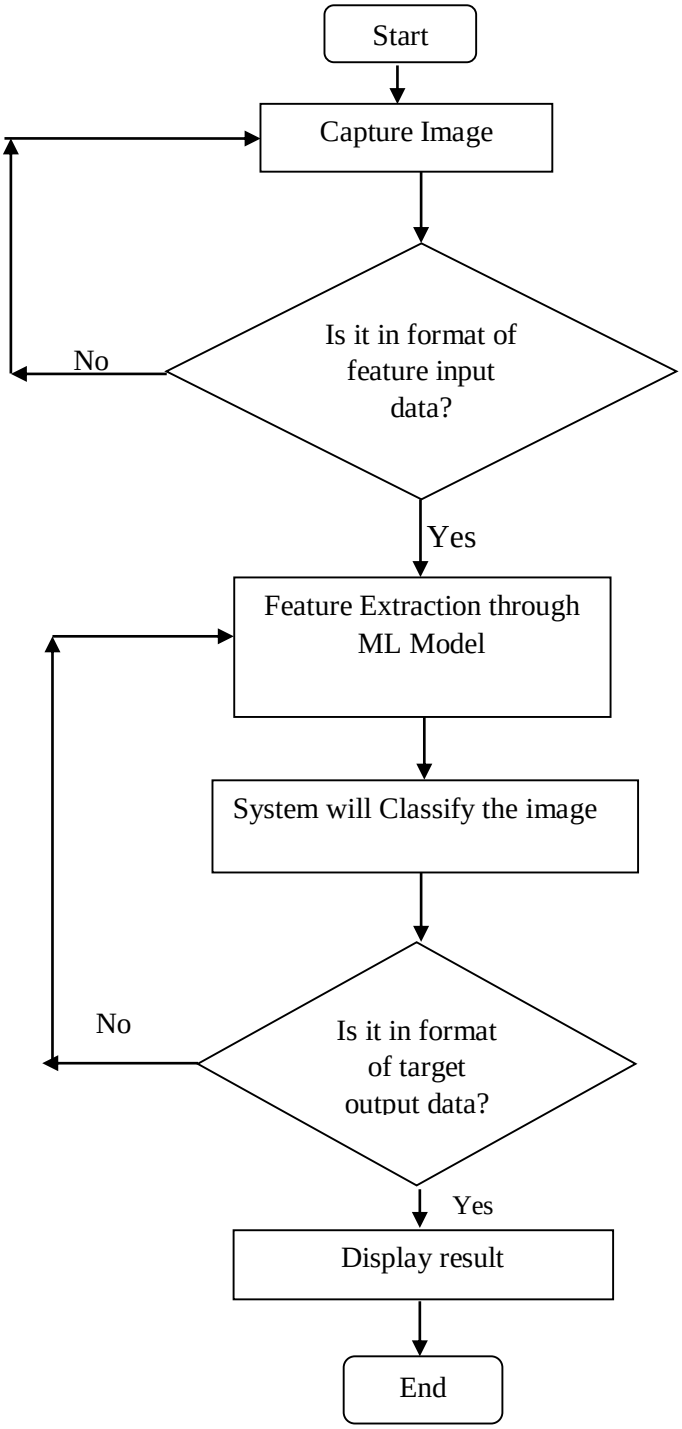
First the model will be train with the help of OpenCV library in the Python. In this process, many image processing techniques will be used and it will create a .csv file with dimensions of each leaf. After this, next comes a classification process. The system will take an image file in either .jpeg or .jpg form, from the user. After taking the image it will manipulate the image and it will match the data of that plant with the data of trained model (.csv file). At last, the model will show the output using a web browser. This html page will show the various information about the plant such as its name and medical uses.

FLOWCHART:

Creating Model Flowchart:



+ Execute Model Flowchart:



DATA-SET:

First need in realising our project was to make a dataset. We tried to find a readymade dataset from the internet but we could not find it. Therefore, to create our dataset, we have used different approaches in order to achieve needed accuracy. The first approach was to download various images from the internet. When we trained out model we came to know that the model cannot distinguish the leaf from many leaves. Hence, with such approach, we could not realize good accuracy. After an abortive effort, we needed a whole new approach. We applied a different approach. We have collected leaves of two medicinal plants (Neem and Basil). Then, we captured images with different background colours, i.e. black and white. After that, we manipulated those images using Python programming and we came to know that the model works more efficiently with white background as compared to that of black background. So, we have collected 412 pictures of six different plant species with the help of mobile phone (Redmi Note 5 Pro). All the training images are in .jpeg form. The size of each training image is between 40 kb to 60 kb and the dimensions are 1280x960. The pictures of these six plants are shown below.

Basil

Neem

Lemon Leaf

Nyctanthes Arbor – Tristis

Mango Leaf

Curry Leaf

TRAING THE MODEL:

For training of our model, we have used 412 images of six different Indian medicinal plant species. The program will take images one by one using for loop and it will perform different image processing operations in Python. The libraries used for image processing are OpenCV and mahotas, and the libraries used for performing different mathematical operations are numpy, pandas and scipy.

After reading a training image the first operation will be performed is converting the image into a grayscale image and it will know the shape of a grayscale image. After this the model will smooth the image using Gaussian filters supported by OpenCV. Next operation is the boundary extraction of the leaf and it is very important operation. This operation will extract the boundaries of the image as shown in figure 3. After the

boundary extraction, the next task is to find the contours. However, the OpenCV library have inbuilt instructions for finding contours. The figure 4 shows the image with contours. With the help of contours, the area and perimeter of the leaf are found and are added to the datasheet.

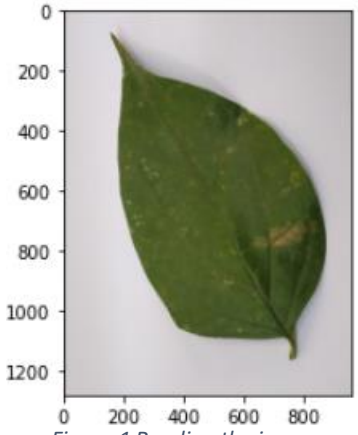


Figure 1 Reading the image

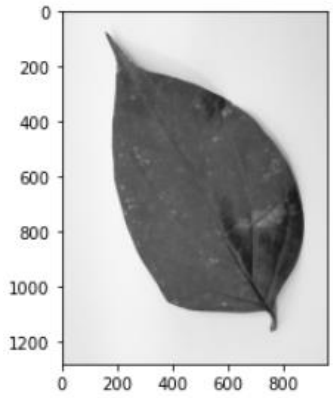


Figure 2 Greayscaling of the image

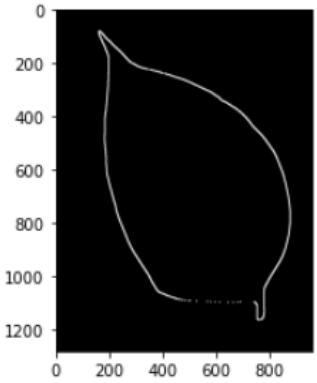


Figure 3 Boundary extraction of the greyscale image

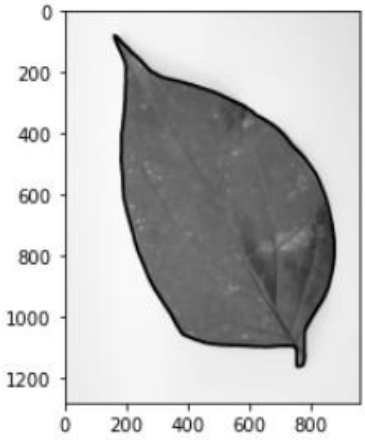


Figure 4 Image with Counters

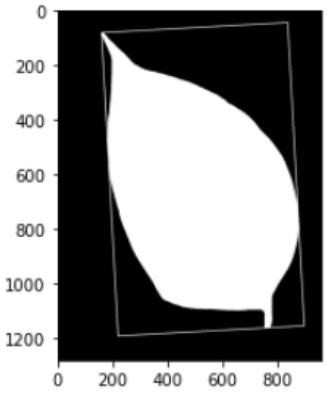


Figure 5 Image with best fitted Rectangle

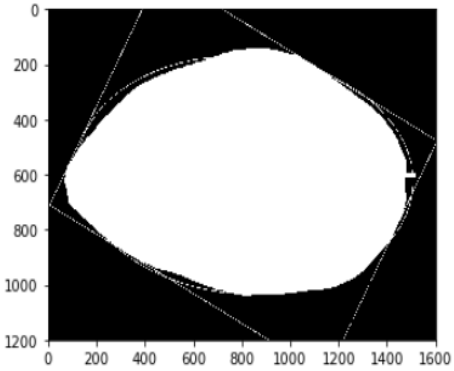
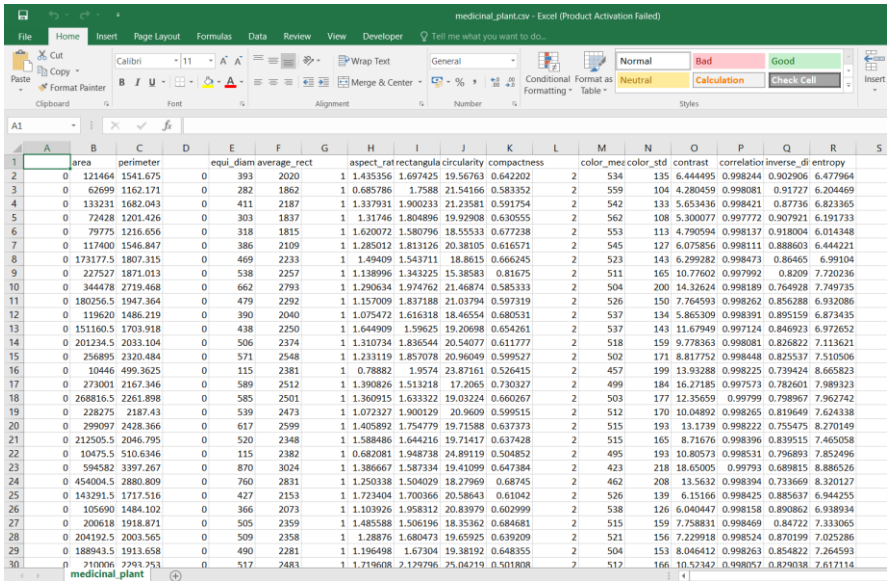


Figure 6 Image with Rectangle and Ellipse

Now, the operations performed on the image will be used to find shape based feature of the leaf. The best rectangle is chosen and not ellipse as removes some portion at the extreme ends if the leaf image. The figure 5 shows the image with the best fitted rectangle and figure 6 shows the image both rectangle and ellipse. Based on the figure 6, different shape based feature such as aspect ratio, rectangularity and circularity is found and added to the datasheet. Next comes the colour based feature of the images. The first instruction will find the mean of the redness of every pixel of the leaf image and this mean is stored in the datasheet. The same operation is performed for finding and storing the value of mean of green and blue pixels.

Next operations are to find the texture based features of the leaf image such as contrast, correlation, inverse difference moments and entropy. For performing this operation, Mahotas library is used. The screenshot of the data sheet created at the end of the program is shown in below figure.



	area	perimeter	equi_diam	average_rect	aspect_rat	rectangula	circularity	compactness	color_mea	color_std	contrast	correlatio	inverse_di	entropy
2	0	121464	1541.675	0	393	2020	1	1.435356	1.697425	19.56763	0.642202	2	534	135	6.444495	0.998244	0.902906	6.477964
3	0	62699	1162.171	0	282	1862	1	0.685786	1.7588	21.54166	0.583352	2	559	104	4.280459	0.998081	0.91727	6.204469
4	0	133231	1682.043	0	411	2187	1	1.337931	1.900233	21.23581	0.591754	2	542	133	5.653436	0.998421	0.87736	6.823365
5	0	72428	1201.426	0	303	1837	1	1.31746	1.804896	19.92908	0.630555	2	562	108	5.300077	0.997772	0.907921	6.191733
6	0	79775	1216.656	0	318	1815	1	1.620072	1.580796	18.55533	0.677238	2	553	113	4.790594	0.998137	0.918004	6.014348
7	0	117400	1546.847	0	386	2109	1	1.285012	1.813126	20.38105	0.616571	2	545	127	6.075856	0.998111	0.888603	6.444221
8	0	173177.5	1807.315	0	469	2233	1	1.49409	1.543711	18.8615	0.666245	2	523	143	6.299282	0.998473	0.86465	6.99104
9	0	227527	1871.013	0	538	2257	1	1.138996	1.343225	15.38583	0.81675	2	511	165	10.77602	0.997992	0.8209	7.720236
10	0	344478	2719.468	0	662	2793	1	1.290634	1.974762	21.46874	0.585333	2	504	200	14.32624	0.998189	0.764928	7.749735
11	0	180256.5	1947.364	0	479	2292	1	1.157009	1.837188	21.03794	0.597319	2	526	150	7.764593	0.998262	0.856288	6.932086
12	0	119620	1486.219	0	390	2040	1	1.075472	1.616318	18.46554	0.680531	2	537	134	5.865309	0.998391	0.895159	6.873435
13	0	151160.5	1703.918	0	438	2250	1	1.644909	1.59625	19.20698	0.654261	2	537	143	11.67949	0.997124	0.846923	6.972652
14	0	201234.5	2033.104	0	506	2374	1	1.310734	1.836544	20.54077	0.611777	2	518	159	9.778363	0.998081	0.826822	7.113621
15	0	256895	2320.484	0	571	2548	1	1.233119	1.857078	20.96049	0.599527	2	502	171	8.817752	0.998448	0.825537	7.510506
16	0	10446	499.3625	0	115	2381	1	0.78882	1.9574	23.87161	0.526415	2	457	199	13.93288	0.998225	0.739424	8.665823
17	0	273001	2167.346	0	589	2512	1	1.390826	1.513218	17.2065	0.730327	2	499	184	16.27185	0.997573	0.782601	7.989323
18	0	268816.5	2261.898	0	585	2501	1	1.360915	1.633322	19.03224	0.660267	2	503	177	12.35659	0.99799	0.798967	7.962742
19	0	228275	2187.43	0	539	2473	1	1.072327	1.900129	20.9609	0.599515	2	512	170	10.04892	0.998265	0.819649	7.624338
20	0	299097	2428.366	0	617	2599	1	1.405892	1.754779	19.71588	0.637373	2	515	193	13.1739	0.998222	0.755475	8.270149
21	0	212505.5	2046.795	0	520	2348	1	1.588486	1.644216	19.71417	0.637428	2	515	165	8.71676	0.998396	0.839515	7.465058
22	0	10475.5	510.6346	0	115	2382	1	0.682081	1.948738	24.89119	0.504852	2	495	193	10.80573	0.998531	0.796893	7.852496
23	0	594582	3397.267	0	870	3024	1	1.386667	1.587334	19.41099	0.647384	2	423	218	18.65005	0.99793	0.689815	8.886526
24	0	454004.5	2880.809	0	760	2831	1	1.250338	1.504029	18.27969	0.68745	2	462	208	13.5632	0.998394	0.733669	8.320127
25	0	143291.5	1717.516	0	427	2153	1	1.723404	1.700366	20.58643	0.61042	2	526	139	6.15166	0.998425	0.885637	6.944255
26	0	105690	1484.102	0	366	2073	1	1.103926	1.958312	20.83979	0.602999	2	538	126	6.040447	0.998158	0.890862	6.938934
27	0	200618	1918.871	0	505	2359	1	1.485588	1.506196	18.35362	0.684681	2	515	159	7.758831	0.998469	0.84722	7.333065
28	0	204192.5	2003.565	0	509	2358	1	1.28876	1.680473	19.65925	0.639209	2	521	156	7.229918	0.998524	0.870199	7.025286
29	0	188943.5	1913.658	0	490	2281	1	1.196498	1.67304	19.38192	0.648355	2	504	153	8.046412	0.998263	0.854822	7.264593
30	0	210006	2293.253	0	517	2483	1	1.719608	2.129796	25.04219	0.501808	2	512	166	10.52342	0.998057	0.829038	7.617114

Figure 7 Datasheet created at the end of the training

TESTING THE MODEL:

After training the model, next task is to test the model. This system provides accuracy of more than 90 percent. To test the model, user has to enter the image on either .jpg or .jpeg form. Next step is to read the .csv file which was created after training. Here, breakpoints are used alongside the image file to create a vector of target labels. For pre-processing and predicting the leaf, instructions supported by sklearn library are used.

Next step is to extract the feature of the leaf image which was entered by the user. Area, perimeter, colour based features, texture based features and shape based features are extracted with the help of OpenCV and Mahotas library. The features of the leaf are stored in variables. Moreover, the feature extraction method used is similar to the method used while training the model. Now, the extracted features of the leaf are compared with the data of the training model which was stored in .csv file. The result is displayed using a html webpage. Below figure shows such html page. This page contains the name of the leaf, a picture of the leaf, uses of the leaf as a medicine, and other helpful website links.

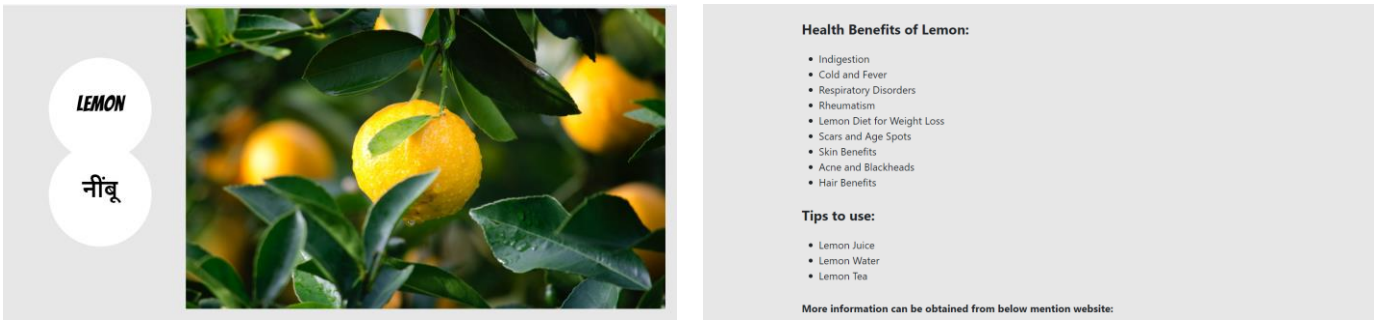


Figure 8 snippet of the output

CONCLUSION:

This system develops a computer vision approach for plant leaf recognition. In this system, computer can automatically recognize 6 different kinds of plants through the leaf images which will be loaded to system through digital cameras, scanner, or mobile phones. The extraction of feature is done by OpenCV. From the different experiments, our model indicates accuracy of 90% if mention format is used for capturing image while keeping in mind the needed distance for image capture.

FUTURE SCOPE:

In most of the papers, plant identification using simple leaves are only considered and compound leaves are not considered. There are 2 types of compound leaves, i.e., palmate and pinnate. From this pinnate compound leaves structure is different from simple leaves. There are millions of plant species available on this globe and here most of the researchers are dealing with very few species. Moreover, the learning time is more for deep learning. Real-time detection of plant species by using images in the real background should be considered in future studies. Object detection by using SSD, RSSD, etc. can also play a big role in this real-time image detection.

REFERENCES:

[1] Plant Leaves Classification: A Few-Shot Learning Method Based on Siamese Network BIN WANG AND DIAN WANG

[2] https://en.wikipedia.org/wiki/Ayurveda

[3] Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review Jana Wäldchen,Patrick Mäder

[4] Review of Plant Identification Based on Image Processing

[5] Gaston KJ, O'Neill MA (2004) Automated species identification: why not? Philos Trans R Soc Lond B Biol Sci 359(1444):655–667. doi:10.1098/rstb.2003.144

[6] V. Narayan and G. Subbarayan, ''An optimal feature subset selection using GA for leaf classification,'' Int. Arab J. Inf. Technol., vol. 11, no. 5, pp. 447–451, Sep. 2014. doi: 10.1109/TIT.2014.2344251.

[7] J. Hu, Z. Chen, M. Yang, R. Zhang, and Y. Cui, ''A multiscale fusion convolutional neural network for plant leaf recognition,'' IEEE Signal Process. Lett., vol. 25, no. 6, pp. 853–857, Jun. 2018. doi: 10.1109/ LSP.2018.2809688

[8] Nayana G. Gavhale1, Dr.A.P.Thakare 2 "Identification of Medicinal Plant Using Machine Learning Approach" by International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 p-ISSN: 2395-0072

[9] https://www.takingcharge.csh.umn.edu/where-ayurveda-come-from

PLAGIARISM SCAN REPORT

Report Generation Date: April 25,2021

Words: 1198

Characters: 8024

Excluded URL :

0% Plagiarism	100% Unique
0 Plagiarized Sentences	61 Unique Sentences

PLAGIARISM SCAN REPORT

Report Generation Date: April 25,2021

Words: 1205

Characters: 7783

Excluded URL :

3% Plagiarism	97% Unique
2 Plagiarized Sentences	62 Unique Sentences

Content Checked for Plagiarism

DEPARTMENT OF ELECTRONICS & COMMUNICATION ENGINEERING GOVERNMENT ENGINEERING COLLEGE

GANDHINAGAR - 382028 GUJARAT TECHNOLOGICAL UNIVERSITY, AHMEDABAD PROJECT REPORT ON

"Identification of Medicinal plant using Deep Convolutional Neural Networks" Submitted by: Brahmbhatt Parita P. [170130111014] Dixit Vyoma R. [170130111022] Under The Guidance of: Prof. Suhas Patel Department of ELECTRONICS & COMMUNICATION ENGINEERING MAY 2021 GOVERNMENT ENGINEERING COLLEGE, GANDHINAGAR DEPARTMENTOF ELECTRONICS AND COMMUNICATION CERTIFICATE This is to certify that Miss. Brahmbhatt Parita Pragneshkumar of Government Engineering College, Gandhinagar, Enrollment no. 170130111014 has completed the project work titled Identification of Indian Medicinal Palnts using Deep Convolutional Neural Networks as prescribed by Gujarat Technological University during the academic term 2020-21. Signature: Name: Prof. Suhas Patel (Project Guide) Signature: Name: Dr. K. G. Maradia (Head of Department) GOVERNMENT ENGINEERING COLLEGE, GANDHINAGAR DEPARTMENTOF

ELECTRONICS AND COMMUNICATION CERTIFICATE This is to certify that Miss. Dixit Vyoma Rakeshkumar of Government Engineering College, Gandhinagar, Enrollment no. 170130111022 has satisfactorily completed the project work titled Identification of Indian Medicinal Palnts using Deep Convolutional Neural Networks as prescribed by Gujarat Technological University during the academic term 2020-21.

Signature: Name: Prof. Suhas Patel (Project Guide) Signature: Name: Dr. K. G. Maradia (Head of

Department) INDEX: • ABSTRACT • INTRODUCTION • PROPOSED SYSTEM • FLOWCHART • DATA-SET •

TRAINING THE MODEL • TESTING THE MODEL • CONCLUSION • FUTURE SCOPE • REFERENCES • PLAGIARISM REPORT ABSTRACT: In Atharva Veda, Ayurvedic species and its usages are explained. Different kind of plant species are used as medicine in Ayurveda for various disease. The recognition of medicinal plant is done manually currently. But as Ayurveda is getting popular worldwide it is necessary to develop an accurate system which can recognize the plant automatically. In recent years,

identification of plant using Deep Neural Networks has increasingly become mature. The identification of plants using conventional method is time consuming, complex and frustrating for non-experts and according to researchers, 80% Indians use Ayurveda as medicine in various diseases. Hence, it is important for people to know which medicinal plant they are using. The availability of digital cameras and smart phones have allowed the idea of automatic medicinal plant identifier to become a reality. In this system, our trained model can identify six types of Indian Medicinal plant species, which are Neem, Basil, Mango leaf, Lemon Leaf, Curry Leaf and Nyctanthes Arbor – Tristis. INDEX TERMS: Computer vision, Feature Extraction, OpenCV, Botany, Ayurveda, Leaf classification, Image processing

INTRODUCTION: Ayurvedic medicinal treatment is a part of India's great history. Ayurveda is developed by hermits in Vedic period or Vedic age. It is originally shared as an oral tradition from one generation to another. Ayurveda was recorded almost 5000 years ago in Sanskrit, in one of the sacred texts called as Vedas in 1200-1000 BCE. Ayurveda mainly consists of healthcare techniques developed using natural ingredients of plants or trees like its flowers, roots, barks and leaves. The benefits of different plant products are explained in detail and how to use that is explained in Ayurveda. In India almost 8500 plants products have medicinal benefits. For making of Ayurvedic medicine almost 1500 plants and its products is used. Currently, large Ayurvedic medicine making firms like Patanjali and Ayush use nearly 500 plants for making their daily products. Almost 80% plants which is used in herbal medicine are found in forest and barren areas. While, rest 20% plants are cultivated in agricultural land and at home. In early time, Ayurvedic experts used to go to forest and barren places to pick the medicinal plants and used to prepare medicines for their patients. As this Ayurvedic knowledge is transferred orally over generations it is very difficult to

recognize different plant species now. And the manual identification methods mainly depend on people's subjective judgement. Botanist use different plant characteristics as identification keys to identify the medicinal plants. However, identification of medicinal plant from field observation necessitate considerable botanical expertise, which put it on the far side of the reach of most aficionados of nature. Manual plant species identification is almost impossible for ordinary people and effortful even for professionals who deal with botanical challenges every day, such as farmers, horticulturists and landscape architects have difficulty identifying plant species. This issue is called the taxonomic crisis in the field of related research. In order to overcome such problems an automated medicinal plant recognition system is needed to be developed. In this information era, the technology of target recognition based on image processing has become accessible to people due to the development of digital image processing and pattern recognition, etc. The theoretical base and technical preparation for leaf identification provided by these techniques. Thus, in recent years, study on plant recognition using image processing techniques has become an area of interest and many researches have been conducted by international researches ever since. In past studies, features of plant identification were elected from plant organs such as flowers, leaves, fruits, and stems. But most of the plants can be identified by its leaves. Because every plant's leaf is different from the other plant's leaf in its shape, size, style or texture. So, we can use these qualities to develop an automated plant classification system. Hence, most of the plant can be identified be performing image processing on the image of plant leaves. This process is done with semi-supervised method in machine learning. However, this method is time consuming and difficult. At present, we can develop a digital image processing and leaf pattern recognition method. Because in past years, performance of deep learning convolutional neural networks has become excellent, development of an automatic medicinal plant recognizer is possible. As digital cameras and portable devices are cheaper and have higher resolution this process can be used to develop a digital image processing systems. The application of deep learning methods in plant identification has achieved an excellent performance as compare to most of the manual feature extraction classification method. The process of system development will consist of domains like biology, medicinal, data science and computer science coding. In deep learning you do not have to manually point feature extraction. For leaf recognition a data set has to construct for each plant species manually to train the system for recognizing the plant species. The accuracy of the system will depend on the data-set which is created. The accuracy will be mostly directly proportional to the number of images means a greater number of layers and its means network will be deeper. The data-set to be properly structured with perfect data for increasing accuracy.
PROPOSED SYSTEM: First the model will be train with the help of OpenCV library in the Python. In this process, many image processing techniques will be used and it will create a .csv file with dimensions of each leaf. After this, next comes a classification process. The system will take an image file in either .jpeg or .jpg form, from the user. After taking the image it will manipulate the image and it will match the data of that plant with the data of trained model (.csv file). At last, the model will show the output using a web browser. This html page will show the various information about the plant such as its name and medical uses.

3

FLOWCHART:

Creating Model Flowchart:

+ Execute Model Flowchart:

DATA-SET:

First need in realising our project was to make a dataset. We tried to find a readymade dataset from the internet but we could not find it. Therefore, to create our dataset, we have used different approaches in order to achieve needed accuracy. The first approach was to download various images from the internet. When we trained out model we came to know that the model cannot distinguish the leaf from many leaves. Hence, with such approach, we could not realize good accuracy. After an abortive effort, we needed a whole new approach. We applied a different approach. We have collected leaves of two medicinal plants (Neem and Basil). Then, we captured images with different background colours, i.e. black and white. After that, we manipulated those images using Python programming and we came to know that the model works more efficiently with white background as compared to that of black background. So, we have collected 412 pictures of six different plant species with the help of mobile phone (Redmi Note 5 Pro). All the training images are in .jpeg form. The size of each training image is between 40 kb to 60 kb and the dimensions are 1280x960. The pictures of these six plants are shown below. Basil Neem Lemon Leaf

Nyctanthes Arbor – Tristis Mango Leaf Curry Leaf

TRAING THE MODEL:

For training of our model, we have used 412 images of six different Indian medicinal plant species. The program will take images one by one using for loop and it will perform different image processing operations in Python. The libraries used for image processing are OpenCV and mahotas, and the libraries used for performing different mathematical operations are numpy, pandas and scipy. After reading a training image the first operation will be performed is converting the image into a grayscale image and it will know the shape of a grayscale image. After this the model will smooth the image using Gaussian filters supported by OpenCV. Next operation is the boundary extraction of the leaf and it is very important operation. This operation will extract the boundaries of the image as shown in figure 3. After the boundary extraction, the next task is to find the contours. However, the OpenCV library have inbuilt instructions for finding contours. The figure 4 shows the image with contours. With the help of contours, the area and perimeter of the leaf are found and are added to the datasheet.

Figure 1 Reading the image

Figure 2 Greayscaling of the image

Figure 3 Boundary extraction of the greyscale image

Figure 4 Image with Counters

Figure 5 Image with best fitted Rectangle

Figure 6 Image with Rectangle and Ellipse

Now, the operations performed on the image will be used to find shape based feature of the leaf. The best rectangle is chosen and not ellipse as removes some portion at the extreme ends if the leaf image. The figure 5 shows the image with the best fitted rectangle and figure 6 shows the image both rectangle and ellipse. Based on the figure 6, different shape based feature such as aspect ratio, rectangularity and circularity is found and added to the datasheet. Next comes the colour based feature of the images. The first instruction will find the mean of the redness of every pixel of the leaf image and this mean is stored in the datasheet. The same operation is performed for finding and storing the value of mean of green and blue pixels.

Next operations are to find the texture based features of the leaf image such as contrast, correlation, inverse difference moments and entropy. For performing this operation, Mahotas library is used. The screenshot of the data sheet created at the end of the program is shown in below figure.

Figure 7 Datasheet created at the end of the training TESTING
THE MODEL:

After training the model, next task is to test the model. This system provides accuracy of more than 90 percent. To test the model, user has to enter the image on either .jpg or .jpeg form. Next step is to read the .csv file which was created after training. Here, breakpoints are used alongside the image file to create a vector of target labels. For pre-processing and predicting the leaf, instructions supported by sklearn library are used.

Next step is to extract the feature of the leaf image which was entered by the user. Area, perimeter, colour based features, texture based features and shape based features are extracted with the help of OpenCV and Mahotas library. The features of the leaf are stored in variables. Moreover, the feature extraction method used is similar to the method used while training the model. Now, the extracted features of the leaf are compared with the data of the training model which was stored in .csv file. The result is displayed using a html webpage. Below figure shows such html page. This page contains the name of the leaf, a picture of the leaf, uses of the leaf as a medicine, and other helpful website links. Figure 8 snippet of the output

CONCLUSION:

This system develops a computer vision approach for plant leaf recognition. In this system, computer can automatically recognize 6 different kinds of plants through the leaf images which will be loaded to system through digital cameras, scanner, or mobile phones. The extraction of feature is done by OpenCV. From the different experiments, our model indicates accuracy of 90% if mention format is used for capturing image while keeping in mind the needed distance for image capture.

FUTURE SCOPE:

In most of the papers, plant identification using simple leaves are only considered and compound leaves are not considered. There are 2 types of compound leaves, i.e., palmate and pinnate. From this pinnate compound leaves structure is different from simple leaves. There are millions of plant species available on this globe and here most

of the researchers are dealing with very few species. Moreover, the learning time is more for deep learning. Real-time detection of plant species by using images in the real background should be considered in future studies. Object detection by using SSD, RSSD, etc. can also play a big role in this real-time image detection.

REFERENCES:

[1] Plant Leaves Classification: A Few-Shot Learning Method Based on Siamese Network BIN WANG AND DIAN WANG

[2] https://en.wikipedia.org/wiki/Ayurveda

[3] Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review Jana Wäldchen,Patrick Mäder

[4] Review of Plant Identification Based on Image Processing

[5] Gaston KJ, O'Neill MA (2004) Automated species identification: why not? Philos Trans R Soc Lond B Biol Sci 359(1444):655–667. doi:10.1098/rstb.2003.144

[6] V. Narayan and G. Subbarayan, ''An optimal feature subset selection using GA for leaf classification,''Int. Arab J. Inf. Technol., vol. 11, no. 5, pp. 447–451, Sep. 2014. doi: 10.1109/TIT.2014.2344251.

[7] J. Hu, Z. Chen, M. Yang, R. Zhang, and Y. Cui, ''A multiscale fusion convolutional neural network forplant leaf recognition,'' IEEE Signal Process. Lett., vol. 25, no. 6, pp. 853–857, Jun. 2018. doi: 10.1109/ LSP.2018.2809688

[8] Nayana G. Gavhale1, Dr.A.P.Thakare 2 "Identification of Medicinal Plant Using Machine Learning Approach" by

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 p-ISSN: 2395-0072

[9] https://www.takingcharge.csh.umn.edu/where-ayurveda-come-from

Matched Sources :

Epidemiology, transmission dynamics and control of SARS ...

Philos Trans R Soc Lond B Biol Sci. 2004 Jul 29;359(1447):1091-105. doi: 10.1098/rstb.2004.149 0.

https://pubmed.ncbi.nlm.nih.gov/15306395/ (https://pubmed.ncbi.nlm.nih.gov/15306395/) 2%

International Research Journal of Engineering and ... 2%

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-005 6 p-ISSN: 2395-0072 Volume: 07 Issue: 08 | Aug 2020 www.irjet.net . International Researc h Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Pollution Prediction [7] [8] Computer Science. ...

https://www.irjet.net/archives/V7/i8/IRJET-V7I8740.pdf (https://www.irjet.net/archives/V7/i8/IRJET-V7I8740.pdf)

Page 5