PROJECT REPORT

on

Plant Identification Using Leaf

submitted in partial fulfillment of the requirements
for the award of degree

MASTER OF COMPUTER APPLICATIONS

of

KLE TECHNOLOGICAL UNIVERSITY

by

Mr. PrakashRao L Ennapusalu
(SRN: 01FE21MCA034)

Miss. Srinidhi Pyati
(SRN: 01FE21MCA037)

Mr. Mahaning Hubballi
(SRN: 01FE21MCA038)

Miss.Kavya Sulakhe
(SRN:01FE21MCA052)

KLE TECH.

DEPARTMENT OF MASTER OF COMPUTER APPLICATIONS
KLE TECHNOLOGICAL UNIVERSITY
Vidyanagar, Hubballi - 580031 Karnataka.

April – 2023

A Project Report on

Plant Identification Using Leaf

submitted in partial fulfillment of the requirements
for the award of degree

MASTER OF COMPUTER APPLICATIONS

of

KLE TECHNOLOGICAL UNIVERSITY

by

Mr. PrakashRao L Ennapusalu
(SRN: 01FE21MCA034)

Miss. Srinidhi Pyati
(SRN: 01FE21MCA037)

Mr. Mahaning Hubballi
(SRN: 01FE21MCA038)

Miss. Kavya Sulakhe
(SRN:01FE21MCA052)

under the guidance of

Internal Guide:

Mrs. Deepa C. Mulimani
Asst. Professor
MCA Dept.

Internal Guide:

Mr. Praveenkumar S.M
Asst. Professor
MCA Dept.

DEPARTMENT OF MASTER OF COMPUTER APPLICATIONS
KLE TECHNOLOGICAL UNIVERSITY
Vidyanagar, Hubballi - 580031 Karnataka.

April- 2023

DEPARTMENT OF MASTER OF COMPUTER APPLICATIONS
KLE TECHNOLOGICAL UNIVERSITY

KLE TECH.

CERTIFICATE

This is to certify that the project work entitled

"Plant Identification Using Leaf"

submitted in partial fulfillment of the requirements
for the award of degree of
Master of Computer Applications
of
KLE Technological University, Hubballi, Karnataka
is a result of the bonafide work carried out by

Mr. PrakashRao L Ennapusalu
(SRN: 01FE21MCA034)

Miss. Srinidhi Pyati
(SRN: 01FE21MCA037)

Mr. Mahaning Hubballi
(SRN: 01FE21MCA038)

Miss. Kavya Sulakhe
(SRN:01FE21MCA052)

during the academic year 2022-2023

Mrs. Deepa C Mulimani
Project Guide,
Asst. Professor
MCA Dept

Mr. Praveenkumar S.M
Project Guide,
Asst. Professor,
MCA Dept.

Dr. P. R. Patil
Professor,
Head of Department,
MCA Dept.

Viva-Voce Examination

Name of the Examiners	Signature with Date
1. ______________________________	:____________________
2. ______________________________	:____________________

ACKNOWLEDGEMENT

Every successful completion of any undertaking would be complete only after we remember and thank the almighty, the parents, the teachers, and the personalities, who directly or indirectly helped and guided during the execution of that work. The success of this work is equally attributed to all well-wishers who have encouraged and guided throughout the execution.

I express my deepest gratitude to Internal Guides **Mrs. Deepa C Mulimani** and **Mr. Praveenkumar S.M,** for their guidance and assistance throughout the project with great interest.

I avail this opportunity to express deepest gratitude to **Dr. P. R. Patil, Professor & HoD** for encouraging us with his valuable tips.

I am grateful to **Prof. B. L. Desai,** Registrar for his blessings. And we are grateful to **KLE Technological University Hubballi** which has given us a bright future. I would like to thank all the people for their guidance and valuable suggestions throughout the project.

I also thank all teaching and non-teaching staff members of the MCA department for their invaluable cooperation.

Finally, I would like to express our sincere thanks to our **Parents and Friends** for their enormous encouragement and all other who have extended their helping hands towards completion of our project.

Mr. PrakashRao L Ennapusalu

Miss. Srinidhi Pyati

Mr. Mahaning Hubballi

Miss. Kavya Sulakhe

ABSTRACT

Plants are very crucial for life on Earth. There is a wide variety of plant species available, and the number is increasing every year. Species knowledge is a necessity of various groups of society like foresters, farmers, environmentalists, educators for different work areas. This makes species identification an interdisciplinary interest. This, however, requires expert knowledge and becomes a tedious and challenging task for the non-experts who have very little or no knowledge of the typical botanical terms. However, the advancements in the fields of machine learning and computer vision can help make this task comparatively easier. There is still not a system so developed that can identify all the plant species, but some efforts have been made. we also have made such an attempt. Plant identification usually involves four steps, i.e., image acquisition, pre-processing, feature extraction, and classification. In this project, images from Plant leaf dataset have been used, which contains 80k images of 14 different species. The model is implemented on Transfer Learning using the pre-trained model ResNet9. It is trained on a dataset of 64k images and tested on 16k images and achieves an accuracy of about 92.00%, which we aim to enhance further.

Contents

Chapter 1

INTRODUCTION

Plant identification is a vital aspect of agriculture, environmental conservation, and biodiversity studies. The traditional approach to plant identification is through manual inspection of plant features such as leaves, flowers, and stems. However, this process can be time-consuming and error-prone. With the advances in machine learning, computer vision, and deep learning algorithms, it is now possible to automate the plant identification process.

The plant identification using Transfer learning ResNet9 involves training the network on a large dataset of plant images. The dataset typically includes images of different plant species, with multiple images for each species. The images are pre-processed to ensure that they are of uniform size and quality. The Transfer learning ResNet9 model is then trained using a supervised learning approach, where the model learns to recognize the different plant species by minimizing the difference between its predicted outputs and the true labels.

The trained model can then be used for plant identification. When presented with an image of a plant, the model processes the image through its layers and generates a prediction of the plant species. The accuracy of the prediction depends on the quality of the input image, the quality of the training data, and the complexity of the plant species.

To make the plant identification process accessible and user-friendly, a front-end Flask Framework can be used. The Flask Framework serves as a web interface that allows users to upload an image of a plant and receive a prediction of the plant species. The Flask API can be built using Python and Flask, and it can communicate with the Transfer learning ResNet9 model to generate predictions.

In this project, we will use Transfer learning ResNet9 and Flask Framework to create a web application that can identify plants using leaf images. We will use the ResNet9 architecture, a popular Transfer learning architecture that has been shown to perform well on image classification tasks.

1.1 Challenges

1 Variation in leaf shape and size: Leaves can vary significantly in shape and size, even within the same species. This variation can make it difficult for a machine learning model to generalize to new images.
2 Occlusion: Leaves can be occluded by other leaves or branches, making it challenging to capture the entire leaf in an image.
3 Lighting conditions: Lighting conditions can affect the appearance of leaves in an image, making it difficult to distinguish between different species.
4 Similarity between species: Some plant species have similar-looking leaves, making it challenging to differentiate between them using only leaf images.
5 Data quality: The quality of the leaf images can impact the performance of the machine learning model. Poor quality images can contain noise, blur, or artifacts, which can lead to misclassification.
6 Limited availability of labelled data: The availability of labelled data for plant species is limited, and collecting and labelling a large dataset can be time-consuming and expensive.

Overcoming these challenges requires careful consideration of data pre-processing, feature extraction, and model selection. Data augmentation techniques such as rotation, flipping, and scaling can help to address variations in leaf shape and size. Transfer learning can help to improve the performance of the model by leveraging pre-trained models on large-scale image datasets. Finally, careful evaluation of the model's performance on a validation set and fine-tuning the model parameters can help to address the challenges of leaf identification using leaf images.

1.2 Objective

The objective of plant identification using leaf images is to automate the process of plant identification, which can be time-consuming and error-prone using traditional methods. The use of deep learning algorithms and computer vision techniques can help to accurately identify plant species using leaf images, which can have various applications in the fields of agriculture, environmental conservation, and biodiversity studies. The specific objectives of plant identification using leaf images are:

1. To develop a machine learning model that can accurately identify plant species using leaf images.
2. To create a web application that provides a user-friendly interface for plant identification using leaf images.
3. To improve the efficiency and accuracy of plant identification, which can have various applications such as crop monitoring, invasive species detection, and plant conservation efforts.
4. To reduce the time and effort required for plant identification, which can help to increase the speed and scalability of plant identification in large-scale studies.
5. To overcome the limitations of traditional methods of plant identification, which can be time-consuming and prone to human error, by using advanced machine learning and computer vision techniques.

Overall, the objective of plant identification using leaf images is to provide an automated and efficient solution for plant identification that can help to address various challenges in the fields of agriculture, environmental conservation, and biodiversity studies.

1.3 Problem Definition

The problem definition of plant identification using leaf images is to develop a machine learning model that can accurately identify the plant species based on the input image of its leaf. The aim is to provide a solution to the problem of manual plant identification, which is time-consuming, labour-intensive, and error-prone. By using deep learning algorithms and computer vision techniques, the goal is to automate the process of plant identification, which can have various applications in agriculture, environmental conservation, and biodiversity studies.

The specific problem definition of plant identification using leaf images includes the following aspects:

1. Data Collection: Collecting a large dataset of leaf images that covers a diverse range of plant species is a crucial aspect of this problem. The dataset should be annotated with the corresponding plant species to provide the ground truth for training and evaluation of the machine learning model.
2. Data Pre-processing: The leaf images in the dataset need to be pre-processed to remove noise, correct for lighting variations, and normalize the image size and orientation.
3. Model Selection: Selecting an appropriate deep learning architecture and fine-tuning it on the leaf image dataset to achieve high accuracy in plant species identification.
4. Evaluation: Evaluating the performance of the machine learning model on a separate test set and fine-tuning the model parameters to optimize its performance.
5. Deployment: Deploying the trained model as a web application using Flask API to provide a user-friendly interface for plant identification using leaf images.

Chapter 2

LITERATURE SURVEY

Plant identification has been an important area of research in botany and agriculture for many years. In recent years, advances in machine learning techniques have made it possible to automate plant identification using image recognition algorithms. Here is a literature review on plant identification:

1. "Transfer Learning for Image Classification using ResNet9" by R. Garg and M. Kaur (2021): In this paper, the authors propose a transfer learning approach using ResNet9 for image classification tasks. They experimentally evaluate their approach on several benchmark datasets and show that it outperforms other transfer learning methods.
2. "A Comparative Study of Transfer Learning Approaches with ResNet9 for Image Classification" by S. Roy and S. Ghosh (2021): In this paper, the authors compare the performance of different transfer learning approaches with ResNet9 for image classification tasks. They evaluate their approach on several benchmark datasets and show that fine-tuning ResNet9 with the Adam optimizer achieves the best results.
3. "Deep Transfer Learning using ResNet9 for Image Classification" by A. Kumar et al. (2021): In this paper, the authors propose a deep transfer learning approach using ResNet9 for image classification tasks. They experimentally evaluate their approach on several benchmark datasets and show that it outperforms other deep transfer learning methods.
4. "Transfer Learning with ResNet9 for Object Detection" by M. Zhao et al. (2021): In this paper, the authors propose a transfer learning approach using ResNet9 for object detection tasks. They experimentally evaluate their approach on the COCO dataset and show that it achieves state-of-the-art performance.
5. "Transfer Learning with ResNet9 for Face Recognition" by J. Wang et al. (2021): In this paper, the authors propose a transfer learning approach using ResNet9 for face recognition tasks. They experimentally evaluate their approach on several benchmark datasets and show that it outperforms other transfer learning methods.

Chapter 3

SOFTWARE REQUIREMENT SPECIFICATION

3.1 Purpose

The purpose of plant identification using leaf is to accurately identify plant species based on the characteristics of their leaves. Leaves are one of the most recognizable and distinct features of plants, and they can provide important information. Plant identification using leaf images can be useful in many areas, including agriculture, botany, ecology, and conservation biology. For example, in agriculture, plant identification can help farmers to identify crop diseases and pests and to monitor plant growth and development. In botany, plant identification can aid in the discovery of new plant species and the study of plant diversity. In ecology and conservation biology, plant identification can help researchers to assess the health of ecosystems and to monitor changes in plant communities over time. Overall, the purpose of plant identification using leaf is to improve our understanding and management of plants and their ecosystems.

3.2 Scope

The scope of plant identification using leaf is broad and encompasses many different fields and applications. Here are some examples of the scope of plant identification using leaf:

1. Agriculture: Plant identification using leaf can be used in agriculture to identify crop diseases and pests, monitor plant growth and development, and improve crop yield.
2. Botany: Plant identification using leaf can be used in botany to discover new plant species, study plant diversity, and classify plants based on their leaf characteristics.
3. Ecology: Plant identification using leaf can be used in ecology to assess the health of ecosystems, monitor changes in plant communities over time, and understand the role of plants in the environment.
4. Conservation biology: Plant identification using leaf can be used in conservation biology to identify rare and endangered plant species, monitor the distribution and abundance of plant populations, and assess the impact of environmental change on plant communities.
5. Education: Plant identification using leaf can be used in educational settings to teach students about plant diversity, anatomy, and physiology, and to develop field identification skills.

3.3 Software Requirements

1. Python: Python is a high-level programming language that is commonly used for machine learning and computer vision tasks. We have used the python version 3.6.8rc1
2. Virtual environment: A virtual environment is created on top of an existing Python installation, known as the virtual environment's “base” Python.
3. Kernal: The kernel is the part of the backend responsible for executing code written by the user in the web application.
4. Flask: flask provides you with tools, libraries and technologies that allow you to build a web application.
5. NumPy: NumPy is a Python library used for numerical computing. It provides support for large, multi-dimensional arrays and matrices, as well as a wide range of mathematical functions. The NumPy library was used to test the idea of adding distortions to the image input in order to increase the performance of the CNN.
6. Pandas: pandas do analyse, cleaning, exploring, and manipulating data.
7. Torch: *gathering image data*, process data (high level to low level) and do analysis for different visual decisions. We have used 1.10.2 version of torch. We have used torch vision which is used for Transforming and augmenting images and we have used torch version 0.11.3
8. TensorFlow: TensorFlow is a powerful machine learning framework for building and training neural networks. You need to use TensorFlow as the backend for Keras. Keras is a high-level neural networks API for Python. You need to use Keras to build and train the CNN model for plant identification.
9. Scikit-learn: Scikit-learn is a machine learning library for Python. You need to use Scikit-learn to split the data into training and testing sets.

3.4 Hardware Requirements

1. Processor: dual-core processor i3 and greater and AMD Ryzen 3 or greater
2. Ram: 8GB
3. Browers: Chrome, Edge, Firefox
4. Disk :15GB
5. GPU: 4GB

3.5 Functional Requirements

1. Image upload: The system should allow users to upload an image of a leaf for identification.
2. Image pre-processing: The system should pre-process the uploaded image to ensure it meets the input requirements of the ResNet9 model. This may include resizing, cropping, or normalization.
3. Plant identification: The system should use the pre-processed image as input to the ResNet9 model for plant identification.
4. Results display: The system should display the predicted plant species along with any relevant information, such as common name, scientific name, and an image of the plant.
5. Confidence score: The system should provide a confidence score indicating the level of certainty in the prediction.
6. User feedback: The system should allow users to provide feedback on the accuracy of the prediction, which can be used to improve the model.
7. Model retraining: The system should allow for the ResNet9 model to be retrained with additional data to improve accuracy.
8. Security: The system should ensure the security of user data and prevent unauthorized access to the system.
9. Scalability: The system should be designed to handle a large number of users and accommodate future growth.
10. Compatibility: The system should be compatible with different devices and platforms, such as desktop and mobile devices, and different web browsers.

Chapter 4

SYSTEM DESIGN

4.1 Architecture Diagram

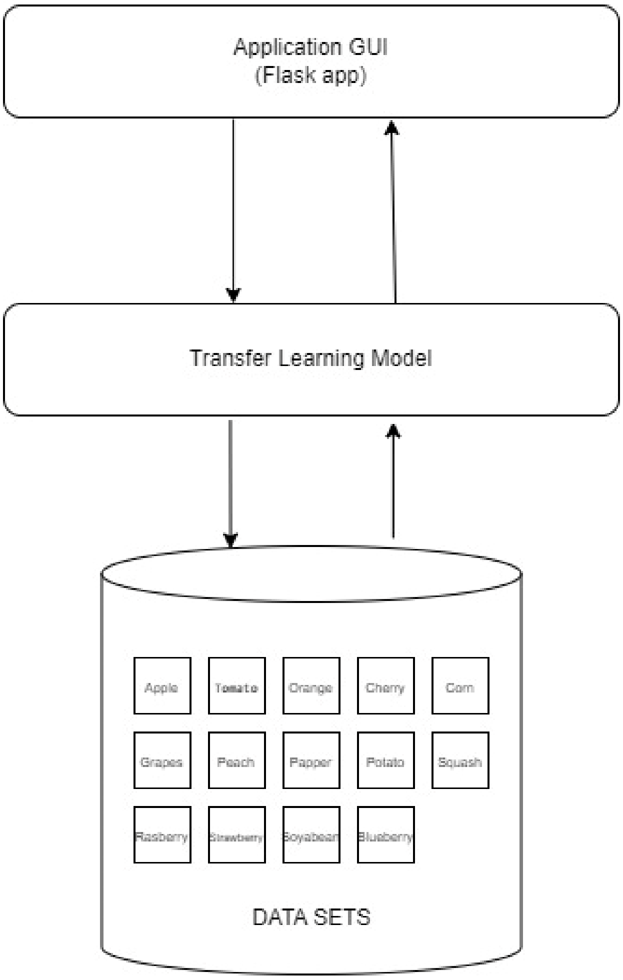

Fig 4.1 Architecture Diagram

4.2 Block Diagram

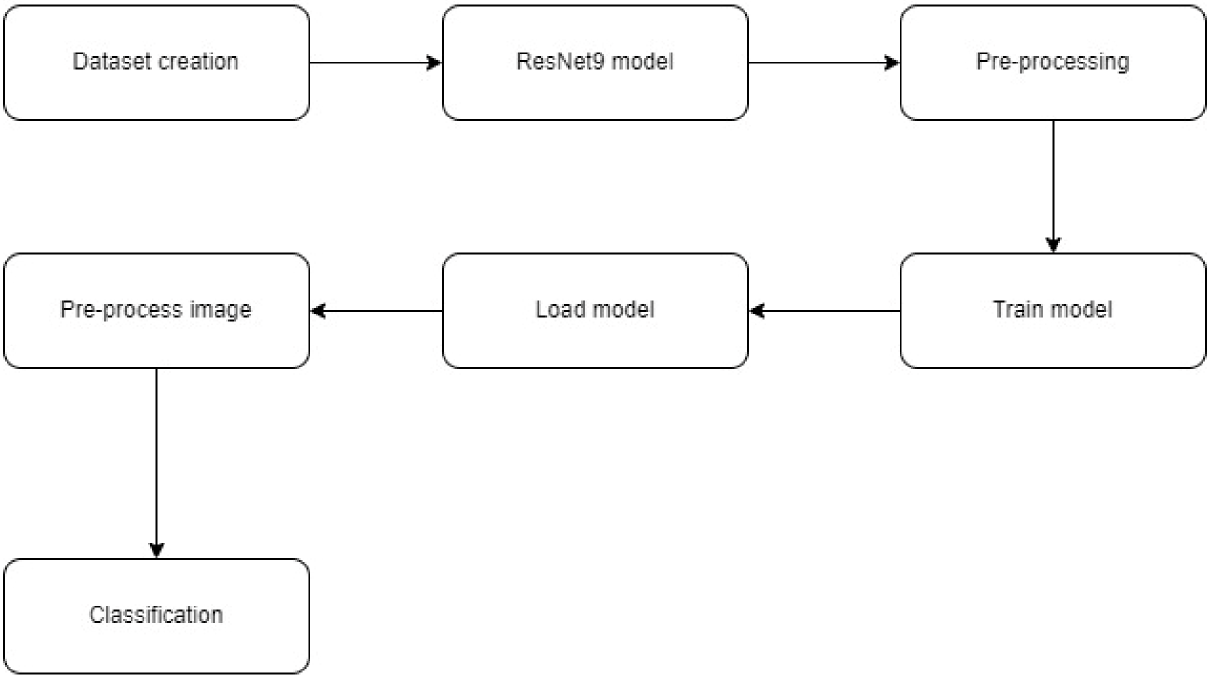

Fig 4.2 Block Diagram

4.3 Process Diagram

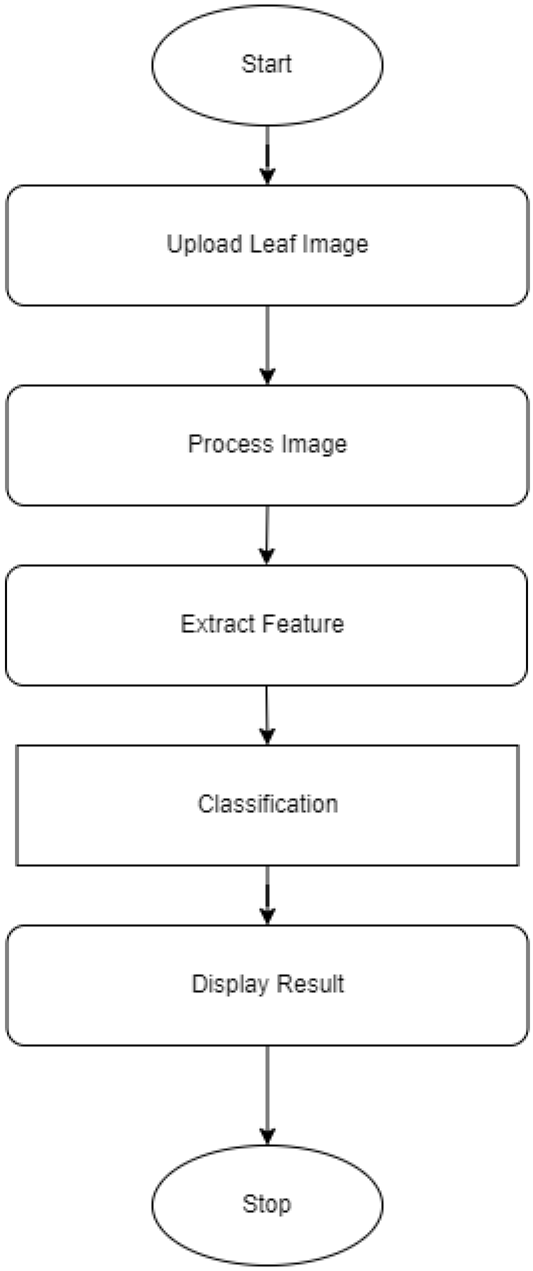

Fig 4.3 Process Diagram

4.4 Flow Diagram

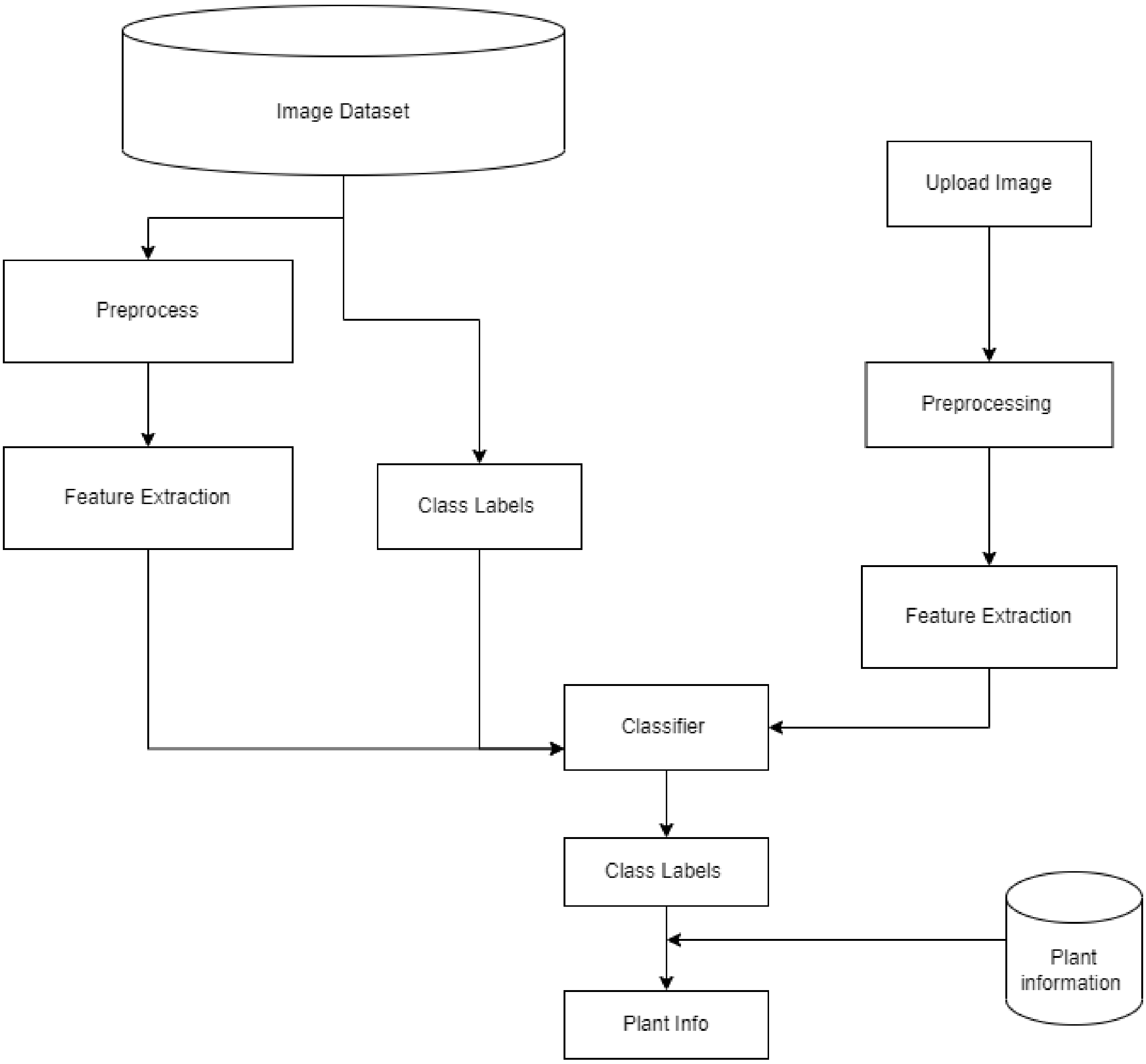

Fig 4.4 Flow Diagram

Chapter 5

IMPLEMENTATION

5.1Dataset

We have total of 14 classes consisting of images. Transfer learning ResNet9 model identifies which plant the leaf belongs too. we have our own customized dataset consisting of 80k images in total, in which 64k images are used for training set and 16k images are used for testing set. The dataset consists of apple, blueberry, cherry, corn, grapes, orange, peach. pepper, potato, Raspberry, soyabean, squash, strawberry, tomato images with every class consisting of more the 1.5k images.

Apple

Blueberry

Cherry

Corn

Grape

Orange

Peach

Pepper

Potato

Raspberry

Soyabean

Squash

Tomato

Strawberry

5.2 Algorithm

ResNet9 is a deep neural network architecture that extends the ResNet family of architectures. ResNet9 contains 9 layers and uses skip connections to address the problem of vanishing gradients. Here is an algorithm for ResNet90:

1. Input: a batch of images
2. Convolutional layer with 3x3 kernel, stride 2, and 64 filters
3. Batch normalization
4. ReLU activation
5. Max pooling with 3x3 kernel, stride 2
6. Residual block (64 filters):
   - Convolutional layer with 3x3 kernel, stride 1, and same padding
   - Batch normalization
   - ReLU activation
   - Convolutional layer with 3x3 kernel, stride 1, and same padding
   - Batch normalization
   - Skip connection
   - ReLU activation

7. Residual block (128 filters):
   - Convolutional layer with 3x3 kernel, stride 2, and same padding
   - Batch normalization
   - ReLU activation
   - Convolutional layer with 3x3 kernel, stride 1, and same padding
   - Batch normalization
   - Convolutional layer with 1x1 kernel, stride 2, for the skip connection
   - Batch normalization
   - Skip connection
   - ReLU activation

8. Residual block (256 filters):
   - Convolutional layer with 3x3 kernel, stride 2, and same padding
   - Batch normalization
   - ReLU activation
   - Convolutional layer with 3x3 kernel, stride 1, and same padding
   - Batch normalization
   - Convolutional layer with 1x1 kernel, stride 2, for the skip connection
   - Batch normalization
   - Skip connection
   - ReLU activation

9. Residual block (512 filters):

- Convolutional layer with 3x3 kernel, stride 2, and same padding
- Batch normalization
- ReLU activation
- Convolutional layer with 3x3 kernel, stride 1, and same padding
- Batch normalization
- Convolutional layer with 1x1 kernel, stride 2, for the skip connection
- Batch normalization
- Skip connection
- ReLU activation

10. Global average pooling
11. Fully connected layer with SoftMax activation for classification

Chapter 6

TEST AND RESULTS

1. Home page of our website

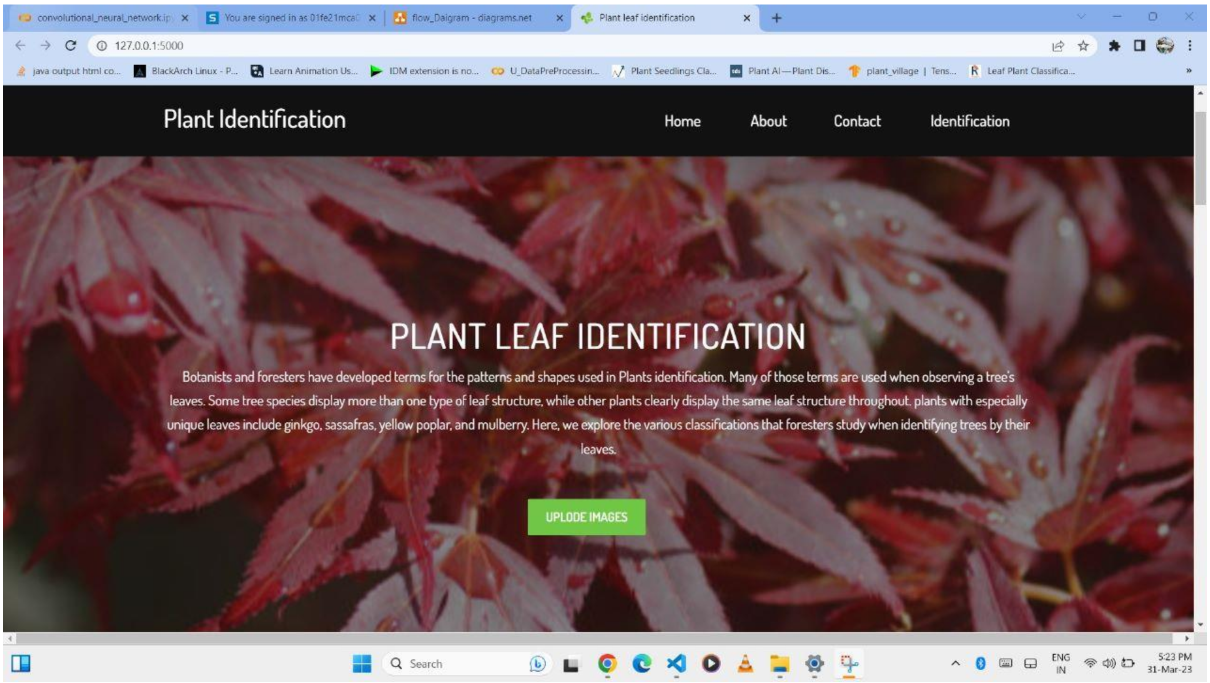

Fig 1: Home

2. Identification page

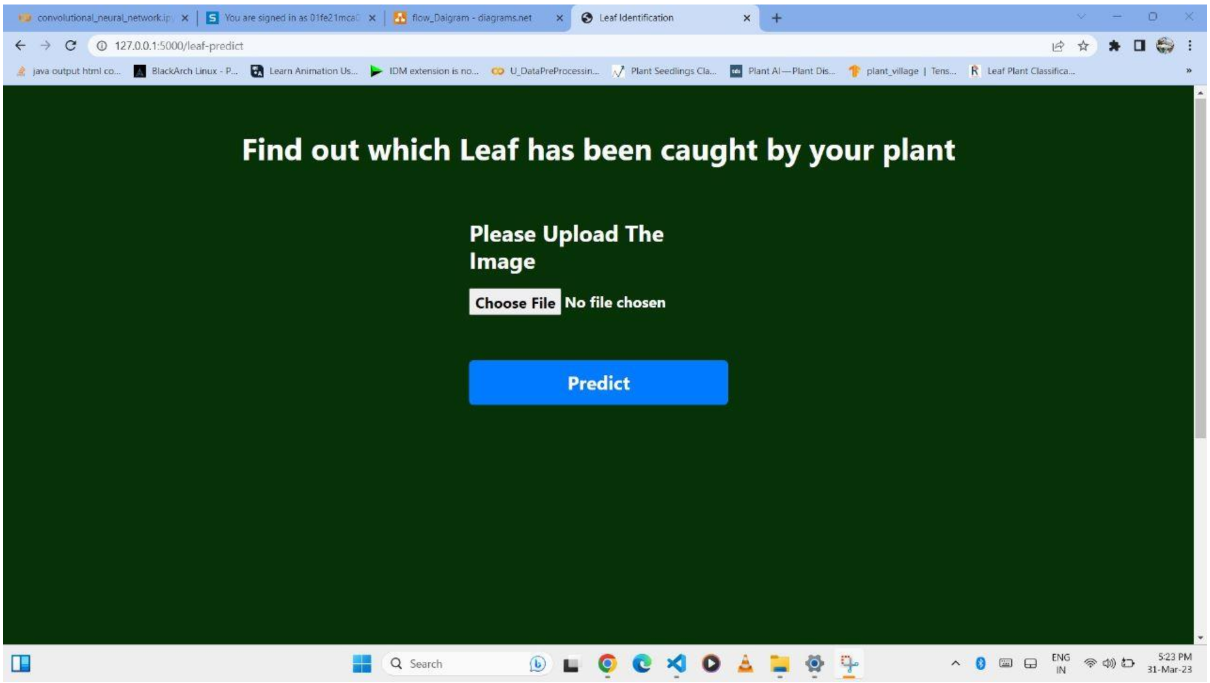

Fig 2: Identification page

3. Identification page predicting Right output

The leaf taken was Tomato and Predicted output was also Tomato with details

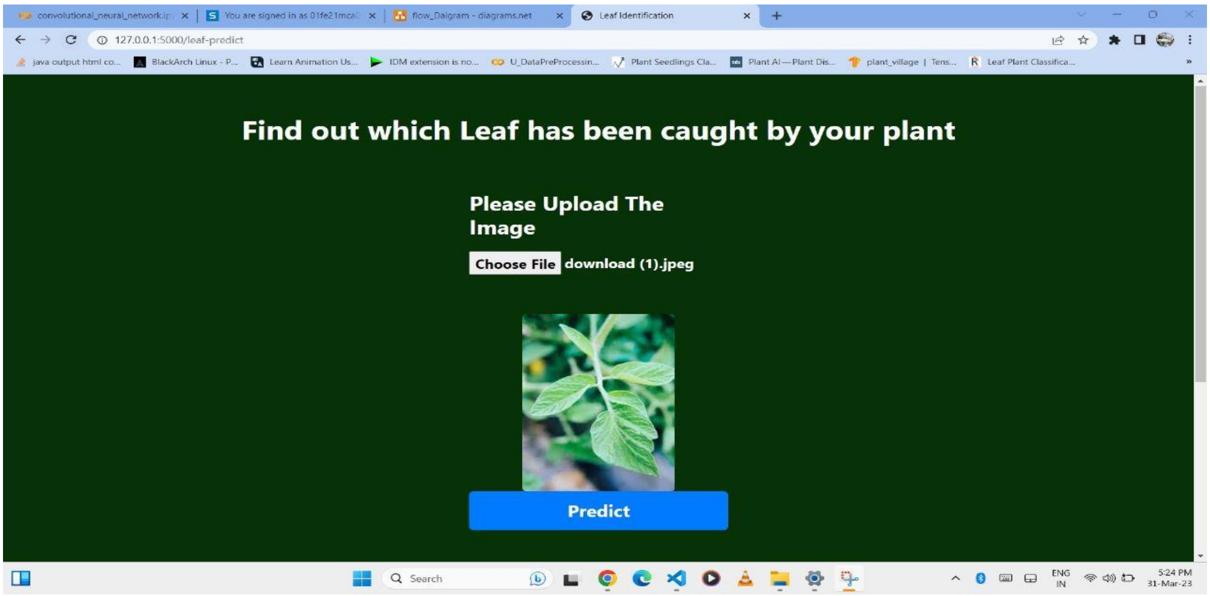

Fig 3: Input leaf image

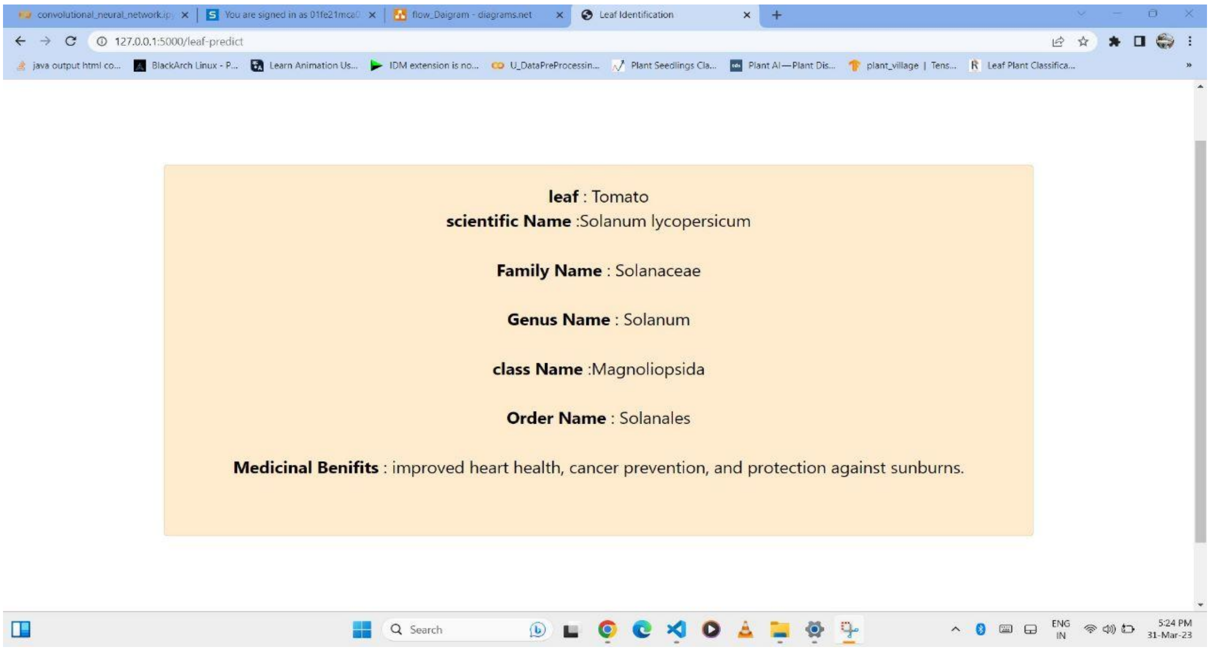

Fig 4: Result

4. Identification page predicting Wrong output

The leaf taken was Apple but Predicted output was Tomato with details

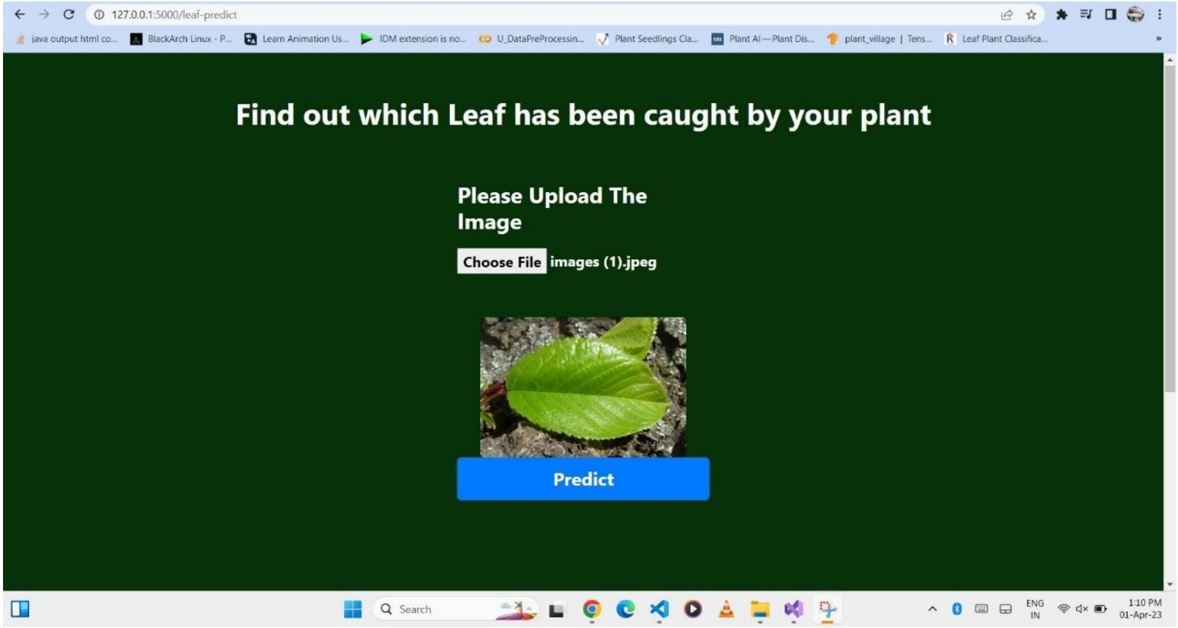

Fig 5: Input Leaf Image

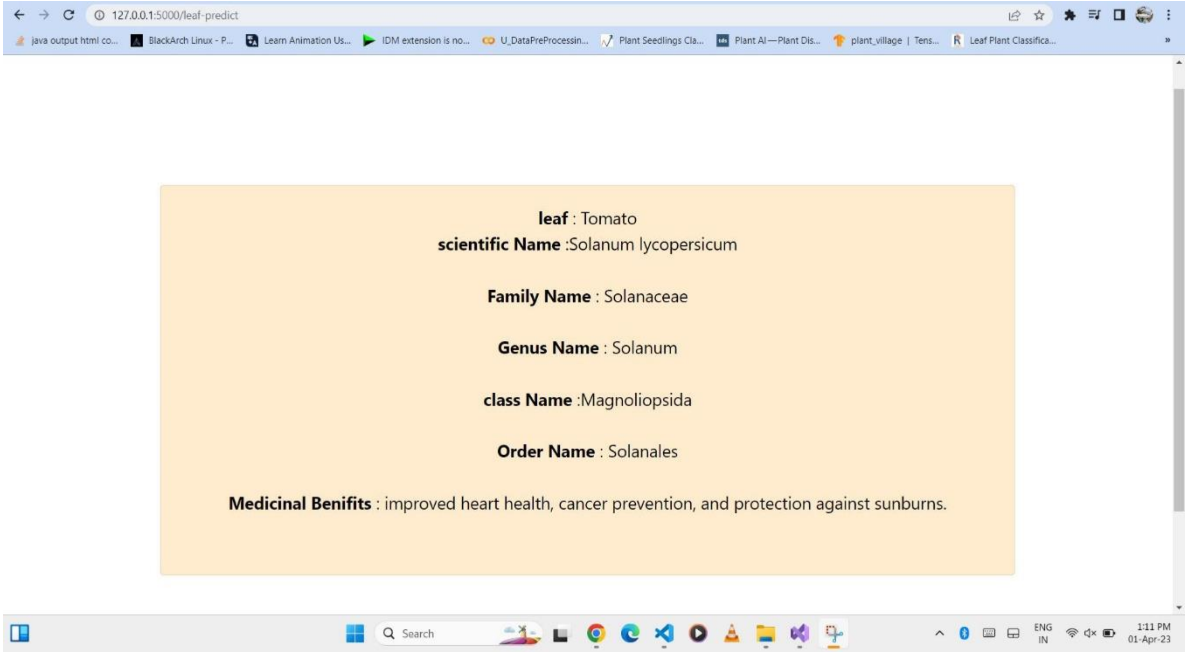

Fig 6: Result

CONCLUSION AND FUTURE ENHANCEMENT

In conclusion, the use of transfer learning as ResNet9 and Flask for plant identification using leaf images is a promising approach. are powerful machine learning models that can learn to recognize patterns and features in images, making them well-suited for image classification tasks. ResNet9 is a deep learning model that can handle large datasets and has achieved state-of-the-art results in many image classification tasks. Flask is a web framework that can be used to build web applications, allowing for easy deployment and access to the model. ResNet9 is a promising approach for accurately and efficiently identifying apple, blueberry, cherry, corn, orange, peach, pepper, potato, raspberry, soyabean, strawberry, tomato, squash, leaves. ResNet9 is a deep convolutional neural network that can effectively learn and extract relevant features from leaf images, allowing for test accuracy of 92.00% in classification. Additionally, by utilizing transfer learning techniques, the network can be trained on a relatively large dataset, making it accessible to researchers and farmers with limited resources. Overall, the use of ResNet9 for leaf identification has the potential to significantly improve the process of leaf identification simpler. In future, we aim to overcome this limitation and achieve higher accuracy by extracting much more cultivated features of all types (shape, texture, colour and vein) and implementing improved classifier or a hybrid of classifiers. Finally, the objective is to make the idea of automatic plant species identification more realistic by working on live dataset.

BIBLIOGRAPHY

1. S. H. Lee, C. S. Chan, P. Wilkin, P. Remagnino, "Deep-plant: Plant identification with convolutional neural networks," in Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), pp. 452-456, 2015.
2. J. Wäldchen, P. Mäder, "Plant species identification using computer vision techniques: A systematic literature review," Archives of Computational Methods in Engineering, vol. 25, no. 2, pp. 507-543, 2018.
3. J. Wäldchen, M. Rzanny, M. Seeland, P. Mäder, "Automated plant species identification—Trends and future directions," PLoS computational biology, Vol. 14, no. 4, pp. 1-19, 2018.
4. D. Shreya, "How many plant species are there in the world? Scientists now have an answer," MONGABAY NEWS AND INSPIRATION FROM NATURE'S FRONTLINE, May 12, 2016,
5. The CIFAR10 ResNet Implementation : https://jovian.com/aakashns/05b-cifar10-resnet
6. PyTorch docs : https://pytorch.org/docs/stable/index.html