A Project Stage-II Report on

"Fruit Classification Using Raspberry PI"

By

Amit Joshi
Seat No: B150233016

Aditya Vardhan Singh
Seat No: B150233005

Karan Ramchandra Kendre
Seat No: B150233108

BE (Electronics and Telecommunication)

Under the guidance of

Dr. Achala M Deshmukh

Sinhgad Institutes

Electronics And Telecommunication
Sinhgad College Of Engineering,
Vadgaon(Bk.),Pune-411041

Affiliated to Savitribai Phule Pune University, Pune.

2021-22

CERTIFICATE

This is to certify that the BE Project Report entitled

"Fruit Classification using Raspberry Pi"

has successfully completed the Project stage-II under the supervision of **DR. Achala Deshmukh** for partial fulfillment of Bachelor of Engineering - Electronics and Telecommunication of Savitribai Phule Pune University. This work has not been submitted elsewhere for the award of any other degree.

Submitted By:

Amit Joshi

Aditya Vardhan Singh

Karan Ramchandra Kendre

Dr. Achala M Deshmukh
Guide

Dr.M.B.Mali
H.O.D. E&TC

Dr.S.D.Lokhande
Principal

ACKNOWLEDGEMENT

I am feeling very humble in expressing my gratitude. It will be unfair to bind the precious help and support which I got from many people in a few words. But words are the only media of expressing one's feelings and my feeling of gratitude is absolutely beyond these words. It would be my pride to take this opportunity to say the thanks.

Firstly, I would thank my beloved guide **Mrs. Achala M Deshmukh** for his valuable guidance, patience and support; He was always there to force me a bit forward to get the work done properly and on time. He has always given me freedom to do dissertation work and the chance to work under his supervision.

It is the love and blessings of my family and friends which drove me to complete this dissertation work.

Thank you all!

Amit Joshi

Aditya Vardhan Singh

Karan Ramchandra Kendre

TABLE OF CONTENTS

ABSTRACT

In this project we attempt to implement a machine learning approach to Fruit Classification using Raspberry Pi. TensorFlow model is effectively implemented in predicting the type of fruit. The objective is to classify the type of fruit in order to make identification of different types of fruits easier with good accuracy. We propose a fruit identification system that go through a TensorFlow model and then through CNN model for the purpose of achieving better accuracy and making classification easier.

Chapter No. 1

Introduction

This Project describes how fruits can be determined using Raspberry Pi. As we know food plays a vital role in human survival, this project has a detector system which detects the variety of fruits using raspberry pi. Whenever the fruit is being picked, it should be placed in a convey or belt, where it is passed through the sensor unit which can detect and display the types of fruits. Images are the important sources of data and information in the agricultural science. The basic concepts and technologies associated with computer vision system and automatic vision-based technology, tool used in image analysis and automated sorting and grading is highlighted. The proposed system starts the process by capturing the fruit image by using Raspberry Pi. Then, the image is transmitted to the processing level where the fruit's characteristics like color,shape,size of fruits samples are extracted. After that by artificial neural network fruit Images are going through the training and testing section. In this neural network is used to detect the shape, size and color of fruit and with the combination of these three features the results are very promising.

1.1 Problem Statement

In order to classify the different varieties of fruits, we have analyzed papers and have given an overview on how these algorithms give precise and accurate future predictions. In this paper, we used several algorithms from which we observed that not all the algorithms implemented can predict data we need.

Finding the closest and getting an accurate proximate result out of such an unpredictability is a problem in itself. Merging of the dataset getting the best output to increase the efficiency alongside considering the different expects of the moderator is tough and we took the same in consideration and implemented with every aspect to generate the best out of the same and get a result that has better efficiency. This is totally based on Machine Learning Algorithm to proceed and provide an effective result. Getting the data and processing it and generating a forecast for three days is the problem statement that we worked on.

1.2 Objective

The objective of image classification is to identify and portray, as a unique gray level, the features occurring in an image in terms of the object or type land cover these features actually represent on the ground. Image classification iOS perhaps the most important part of digital image analysis.

Image classification is the task of categorizing and assigning labels to groups of pixels or vectors within an image dependent on particular rules. The categorization law can be applied through one or multiple spectral or textural characterizations. Image classification techniques are mainly divided into two categories: Supervised and unsupervised image classification techniques.

A computer analyzes an image in the form of pixels. It does it by considering the image as an array of matrices with the size of the matrix reliant on the image resolution. Put simply, image classification in a computer's view is the analysis of this statistical data using algorithms. In digital image processing, image classification is done by automatically grouping pixels into specified categories, so-called "classes."

The algorithms segregate the image into a series of its most prominent features, lowering the workload on the final classifier. These characteristics give the classifier an idea of what the image represents and what class it might be considered into. The characteristic extraction process makes up the most important step in categorizing an image as the rest of the steps depend on it.

Image classification, particularly supervised classification, is also reliant hugely on the data fed to the algorithm. A well-optimized classification dataset works great in comparison to a bad dataset with data imbalance based on class and poor quality of images and annotations.

1.3 Project Features

Features deals with the flexibilities that this model can present. The project was headed with the resource available with help of documentation and all trusted authenticated data on internet. Taking about all the useful information on TensorFlow and CNN model. So, the featuring of the idea came with handling and automating the resource which other agents are making fortune out of it.

Knowledge is a bliss and learning is the curiosity whereas outcome is the expectation so the resource deals with the importation and extraction of multiple machine learning algorithms to learn, process and yield the result to derive and conclude a possible outcome set that is effective and generative in nature.

There is various model that outflows in market which are trying it's best on creating a resource and give the predictability to most of it accurate but everything is not the same and the conclusion of the same are not ideal. The efficiency varies as the variation in the test and train dataset varies and its prediction.

The project was purposed with the intent sole to make and undergo the following way of computing. The first deals with the data extraction that is done with clearing of data and its chunks from the database or the dataset. The second flow is the converting images into array to start computation. Last part is the building a model using customized convolutional and pooling layers which provides the result after computation of the same.

Salient features included are the Adam optimizer instead of Stochastic Gradient descent that gives a boost. Adam takes the advantage of both the momentum and rmsprop by leading us fast to the optimized values. Diving and initializing the expects that needs to be considered. Mitigating the risk factors to bridge and uplift the result.

1.4 Modules Description

1.4.1 Data Set

This is the fundamental module before starting of the project. The dataset is a group of data that are mended together to show the data variations in a time span to undergo further estimation and the source of the resources and its outcome for the later time of evaluation. It generates the result optimization and gives a feasible time period to customize and get the flow to the derivation.

This increases and are used in the level of research and finding the best suitable resource out of the same the resources have to be finely estimated and derived for the best possible outcome and the finest the value become the better is the level of extraction and closure is the best yield values that needs to be considered.

1.4.2 Data Abstraction

Abstraction is the finding of the resource to its best to categorized the above dataset and learning the best out of it. Abstraction of the data is the integral part to the flow. All the data are a huge set of chunks which on processing can lemmatize the yield result and the computational mean too. Thus, with the available resources the data yield had to be derivative.

Abstraction of the dataset is to customize the data set and finding the best suitable constraints to take into consideration and the unwanted resources are the dump which will be dumped and the supreme cluster is created with the valuable constrains and a pattern is needed to be derived from the same.

Data are cleared on this level for the beginning of the process. The valuable data are the set that brings the value to the data set for a better understanding and giving a better yield and production by evaluating the same.

This is a feature abstraction module to extract the featuring of the dataset. This is a feature model process where all the feasible resources are categorized and the same will be in use for the featuring.

1.4.3 Training Dataset

After the abstraction of the data and clustering of the same. The machine had to be trained for which the training data plays the important role. There are thousands of machine learning algorithms that are into place and evolving with the same. The best to the practice of machine learning is to yield the result and the content to derive what's needed with the time frame.

This is a supervised learning form where the input are passed so that the system learns from the same. Various variants of inputs are passed which were stored in the dataset. Every resource is considered and taken into consideration. After considering the whole set of information and the resource the machine tries to learn from the passed dataset. The dataset has to be wide and versatile. After considering the learning it tries to integrate with the same type and flow like the same as the human mind and creates a pattern and the links between the same.

1.4.4 Test Dataset

These are the sets of data that gives the result after learning from the data. This is the test generation with the output result. Results are generated in each phase of testing. This is also termed as the testing phase. Now a new set of datasets are passed which are deliberately like the training dataset and the efficiency of the same is calculated.

Over-Fitting of the dataset. Validation of the same with the effective constraints and hyper parameters are checked. This phase is training and the output is evaluated with the set of training. After each process of computation, the set of data are trained and efficiency of the same is measured and is evaluated with the others.

Various batches of the test are implemented to get to the level of accuracy and derive result to fetch and yield for the best performance and to be true to the effectiveness of the data which is not biased with any constrains available. This determines the efficiency of the system which is must for the predictions.

1.4.5 Evaluation of Report

This is the main part for any implementation of the project. Evaluation of the key point to the success. All the categorization of the work and the best to know the resource fundaments and again establishing the same to check the validity and the work flow and check on the output is must. The evaluation, utilization and implementation undergo a various level of extraction and evaluation.

The main theme is to provide and come up with the output with an accuracy that can be used and implemented. From the starting to the final the process is categorized, supervised and efficiency is check and the working is undergone. Testing is tested and its evaluation are mended.

The process undergoes the same for various time and phase. Testing of the same undergoes sequential iteration for many more to meet up to the constituency. The remarks are to be noted and further work is done on the same with the implementation of the different aligned resources that are integrated with the available resources and its outcome.

Chapter No. 2 Literature Review

2.1 Background Study

Artificial Intelligence (AI) is a multidisciplinary science that includes in making smart machines capable to perform tasks that are analogous to human intelligence. It is an intelligence shown by machines, similar to the natural intelligence displayed by humans and animals to ideally validate and perform actions that will achieve a specific goal in a better way. Deep learning or deep structured learning is a part of a broader family of machine learning techniques based on artificial neural networks. In deep learning, artificial neurons in the networks examine the large dataset and it will inevitably determine the underlying patterns without human interference. In deep learning, a model/network learns to perform classification tasks on images, text, or sound. Deep learning models can achieve state-of-the-art accuracy on image recognition and classification, sometimes exceeding human level performance. Computer vision has various applications like image classification, image recognition and object detection. Among those, image detection is considered as the fundamental application and forms the basis for other computer vision applications. Nowadays internet has an abundance of images, which is very handy for the development of applications and algorithms. There are major advances in image labelling, object detection, image recognition areas across the world. Artificial neural networks have been used extensively for object detection and recognition and this work focuses on suggesting a suitable architecture/network for image recognition and image classification.

A Convolutional Neural Network, a Deep Learning algorithm, can take an input image, assign importance to several aspects/objects in the image and be able to distinguish one from the other. A typical CNN is composed of single or multiple blocks of convolution and sub-sampling layers, after those one or more fully connected layers and an output layer.

The convolutional layer (conv layer) is the central part of a CNN. In a large Input image, a small section of the image is considered and we convolve them into a single output using a filter (Kernel)

Pooling merely means down sampling of an image. It takes small region of the convolutional output as input and sub-samples it to produce a single output. Different pooling techniques like max pooling, mean pooling, average pooling is available. Max pooling takes largest of the pixel values of a region

Neural networks and CNNs in precise rely on a non-linear "trigger" function to signal identification of likely features on every hidden layer. CNNs may use a range of certain functions —such as rectified linear units (ReLU) and continuous trigger (nonlinear) functions—to proficiently implement this non-linear triggering.

A ReLU implements the function y = max(x,0), thus the input and output sizes of this layer will be the same. It rises the nonlinear properties of the decision function and of the overall network without affecting any of the receptive fields of the convolution layer. ReLU layer functionality is demonstrated in, with its transfer function plotted over the arrow.

Fully connected (FC) layers are always used as the last layers of a CNN. This layer takes input from all neurons in the previous layer and performs arithmetic sum of the preceding layer of features with individual neuron in the current layer to generate output

Chapter No. 3

Methodology

Image classification involves the extraction of features from the image to observe some patterns in the dataset. Using an ANN for the purpose of image classification would end up being very costly in terms of computation since the trainable parameters become extremely large.

For example, if we have a 50 X 50 image of a cat, and we want to train our traditional ANN on that image to classify it into a dog or a cat the trainable parameters become –

(50*50) * 100 image pixels multiplied by hidden layer + 100 bias + 2 * 100 output neurons + 2 bias = 2,50,302

We use filters when using CNNs. Filters exist of many different types according to their purpose.

Working of CNN:

CNNs use image recognition and classification in order to detect objects, recognize faces, etc. They are made up of neurons with learnable weights and biases. Each specific neuron receives numerous inputs and then takes a weighted sum over them, where it passes it through an activation function and responds back with an output.

CNNs are primarily used to classify images, cluster them by similarities, and then perform object recognition. Many algorithms using CNNs can identify faces, street signs, animals, etc.

3.1 Block Diagram

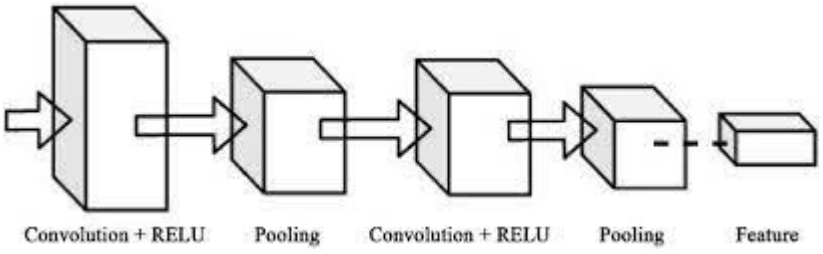

Figure 3.1 Block diagram of method

They are prompt by volume and utilize multi-channeled images. As opposed to flat images that humans can see that only have width and height, CNNs cannot recognize that. Due to digital color images having red-blue-

A convolutional network ingests such images as three separate strata of color stacked one on top of the other. A normal color image is seen as a rectangular box whose width and height are measured by the number of pixels from those dimensions. The depth layers in the three layers of colors (RGB) interpreted by CNNs are referred to as channels.

3.1 Machine Learning

One of the finest words heard in today time is Machine Learning. Either it be at work or different places the machine learning has been an integral part of today's technology. Though its evolutional and developing in a rapid rate and development and deployment of the same is still in progress. The machine learning itself had brought a random change in today worlds because of which automation is in frame which was a mere existence in the past.

It's an aspiring term in today's time. One of the moves that all the firm are interested into. It’s a leading pillar for tomorrow leading the world to a better future of evolution where the customization and labor work can be reduced to half and the safety of the survival can be withheld to stand tall for the better utilization of human mind. Keeping that in picture it’s been a hazard to many more in terms of irrespective field of interest. Since Machine is considered most efficient and the level of mistakes are kept at the minimum the level of work flow can be a work of hazard and further improvement on the same may create a thousand sitting idle in home creating a larger impact on unemployment and livelihood. Which in other way is a threat to the society too.

ML is the abbreviation for Machine Learning. In other word it is making a human mind fitting inside a machine which uses the same to perform the task of thousands. Machine Learning deals with the higher aspects of learning techniques and algorithm which are highly aligned to make the work flow seamlessly effortless with the human tendency of doing work.

Algorithm of such is improvising in nature which learns by themselves and fit themselves in the world of impairment by getting the required data and adjusting with the same giving the effective results out of the same. ML is a subsidiary or the subset of an AI (Artificial Intelligence). It is a mathematical model where computation of the testcases plays the major role in driving of the results.

A wide level of machine learning architecture is implemented today to turn on the yield factor and make people life more efficient in terms of livelihood. Various use of such in Message Filtering like spams, Trash automation are automated and carried out by the same. Since the efficiency is way more than a human tendency. Multi-tasking and processing are also initiated by the same giving a dual output which a human can never ever possibly be able to.

Statistics is the major key role in driving the machine learning in figure. It deals with computation of statistics in a wide range view and processing the same to give a data driven output causing it more sensible and resources able. Not only to the same it optimizes the resources and the efficiency is unbitable and reliable in terms of any means.

Though its being evolutional in nature but it has integrated itself well with the terms of computational and digitalization. Various computational fields like Data Mining, Statistical Analysis, Optimization of resources, Automation are a major part of it. Here the machine has the capacity to process the result on its own as same as the human bring. This process can be initiator as well as the derivable. The statistical flow is mainly reasonable with data driven pattern even the unstructured or the semi-structured data can be processed and approximate answer to the same can be derived. All the equations

3.1.1 Supervised Machine Learning

Supervised learning is the types of machine learning in which machines are trained using well "labelled" training data, and on basis of that data, machines predict the output. The labelled data means some input data is already tagged with the correct output.

In supervised learning, the training data provided to the machines work as the supervisor that teaches the machines to predict the output correctly. It applies the same concept as a student learns in the supervision of the teacher.

Supervised learning is a process of providing input data as well as correct output data to the machine learning model. The aim of a supervised learning algorithm is to find a mapping function to map the input variable(x) with the output variable(y)**.**

In the real-world, supervised learning can be used for Risk Assessment, Image classification, Fraud Detection, spam filtering, etc.

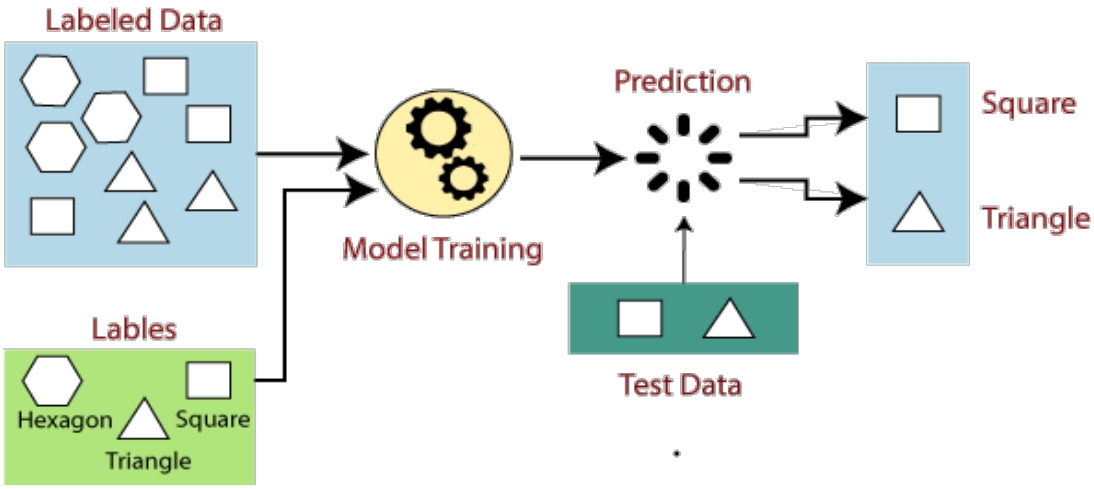

Figure 3.2.1 Supervised machine learning

3.1.2 Unsupervised Machine Learning

Unsupervised learning cannot be directly applied to a regression or classification problem because unlike supervised learning, we have the input data but no corresponding output data. The goal of unsupervised learning is to find the underlying structure of dataset, group that data according to similarities, and represent that dataset in a compressed format.

Example: Suppose the unsupervised learning algorithm is given an input dataset containing images of different types of cats and dogs. The algorithm is never trained upon the given dataset, which means it does not have any idea about the features of the dataset. The task of the unsupervised learning algorithm is to identify the image features on their own. Unsupervised learning algorithm will perform this task by clustering the image dataset into the groups according to similarities between images.

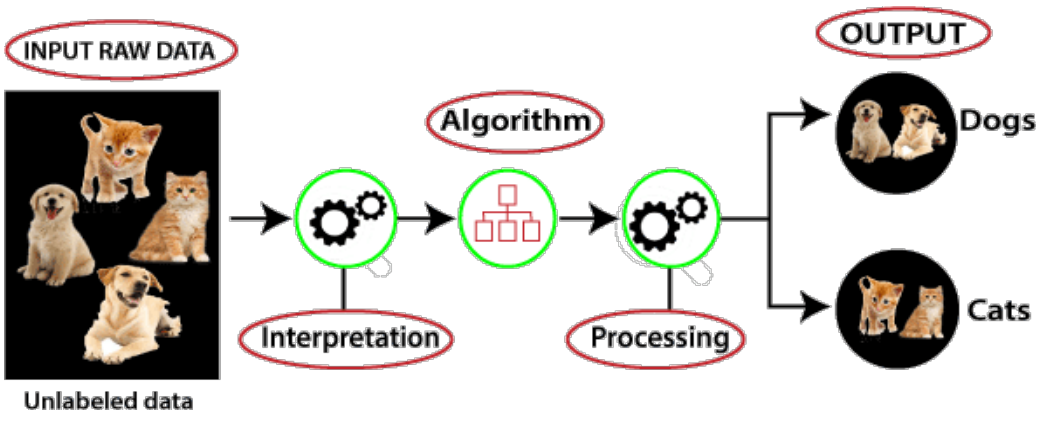

Figure 3.2.2 Un-Supervised machine learning

3.2 Technical Survey

3.2.1 Survey – I

On the surface, teaching a computer to do something like image classification seemed very intriguing to us. Moreover, there are countless real-world applications of this concept. It is in light of these reasons that we decided to work on Image Classification. Thankfully though, this topic has been well-researched by the scientific community and we didn't break a sweat finding resources to learn from. So naturally, we perused a bunch of research papers that dealt with image classification, each from a different perspective. We then decided to implement image classification on a small-scale with the limited hardware we were in possession of. As difficult as it was, we started with SVM and a very small dataset to achieve an accuracy of 93%. Although SVM is a very strong technique, achieving such a high accuracy is still an anomaly. We realized that our results boasted such a high accuracy due to the lack of a large enough dataset. So, using data augmentation, we more than tripled the size of our dataset. On performing SVM now, we achieved an accuracy of 82%, a significant decrease. Unsatisfied with the results, we decided to move to other deep learning techniques. This quest led us to Neural Networks and, CNN. On successfully implementing CNN, we achieved an accuracy of a staggering 93.57% on the very same dataset. This stands as a testimony to the increased potential of deep learning techniques over the more traditional machine learning techniques.

3.2.2 Survey – II

Aiming at the issue of timeliness and lack of partial image data in life, an algorithm, transfer learning which based on convolutional neural network (CNN) is proposed, combining image histogram of oriented gradient (HOG) feature extraction method and support vector machine (SVM) pre-classification method. Firstly, the HOG features of the training sample similar to the attributes of the samples which to be classified are extracted, then the hog features of the training samples are imported into the SVM classifier to get the pre-classification results. Finally, the pre-classification results are used as training samples to train the transfer network of CNN for getting new transfer learning model, this model can be used to classify similar pre-classification samples. The experimental results show that the classification accuracy of the five categories of elephants and dinosaurs used in this paper is effectively improved, and the overall classification accuracy can reach 95%, compared with the traditional classifier algorithm and convolutional neural network algorithm. The classification accuracy has been improved by

about 5%.

3.2.3 Survey – III

With an upsurge in the rate of data production, pervasive usage of cameras for automation and surveillance and the requirement of visual input for artificially intelligent devices all across the globe, there has been a rapid increase in the mass of image data being generated today. This gives rise to the essentiality of automated image processing required to simplify image related tasks. Automated image processing bridges the gap between the human visual system and the pixel level data of images. Deep Convolution Neural Networks are being deployed expansively to analyze, detect and classify images for a diverse number of tasks. These neural networks, similar to the human neural network, contain neurons with learnable weights and biases, which are trained to identify and classify different objects or features across the image. This paper presents a functional implementation of image recognition using a small convolutional neural network, proposing less complexity and yielding good classification accuracy for all tested data sets.

Software Description

4.1.1 Jupyter Notebook

Jupiter Notebook or so called IPython Notebook is an interactive web based computational mean for starting with Jupiter Notebook documents. The term notebook itself is a huge entity to represent the integration with different entity sets. JSON is the main document form from the same for the execution which follows the brief on the schema and the input and output means. It has high integration with several language set and has various flexibilities with the choices.

The extension used for the same is ".ipynb" which runs in this platform. It's an open-source software package with interactive communication means. It has its open standards for the same. It's an open community best for budding programmers. The flexibility of the same is phenomenon and splendidly done the configuration and integration of the same is simplest and easy on hold so that no prior distortion is generated and the efficiency of the same is measured throughout any system of choice. It's the best software sets that been used across cross for designing and developing of the products and support wide help support.

Not only to that, it provides scalability in the code and the deployment of the same. Various Language can be changed and the project can be undertaken on the same. The created notebook files can be shared and stored in various means for further utilization. It supports cultivated and interactive output sets. Easily crossed over for graphing, plotting and visualizing of the elements.

Data Integration of the same is to its best. The integration of big data and it can process chunks of values in an approx. time which gives a better performance and the higher computational means. Various works on data like cleaning, cleansing, transforming modeling and visualizing can be done by the same.

4.2 Requirement Analysis

4.2.1 Functional Requirements

Functional requirements deal with the functionality of the software in the engineering view. The component flow and the structural flow of the same is enhanced and described by it.

The functional statement deals with the raw datasets that are categorized and learning from the same dataset. Later the datasets are categorized into clusters and the impairment of the same is checked for the efficiency purpose. After the dataset cleaning the data are cleansed and the machine learns and finds the pattern set for the same it undergoes various iteration and produce output.

4.2.2 Non-Functional Requirements

Nonfunctional requirement deals with the external factors which are non- functional in nature It is used for analysis purpose. Under the same the judgment of the operations is carried out for its performance. Stock is feasible and is ever changing so these extra effects and the requirements helps it to get the latest updates and integrate in a one goes where the technicians can work on and solve a bug or a draft if any.

The non-functional requirements followed are its efficiency and hit gain ratio. The usability of the code for the further effectiveness and to implement and look for the security console. The System is reliable and the performance is maintained with the support of integration and portability of the same.

4.2.3 HARDWARE REQUIREMENT

- Processor : intel i5 or above
- Ram : Minimum 8GB or More
- Hard Disk : Minimum 2 GB of space
- Input Device : Keyboard
- Output Device: Screen of monitor or Laptop

-

4.2.4 SOFTWARE REQUIREMENT

- Operating System: Windows & Linux
- IDE : Jupiter Notebook
- Dataset : .jpg file
- Visualization : matplot lib, numpy,sklearn,
- Server : Web Server with HTTP process

Chapter 5 Design

• 5.1 Design Goals

To make the project runs smoothly it's required that we make plan and design some accepts like flowcharts and system architecture which are defined below.

• 5.1.1 Data Collection

Data collection is one of the important and basic things in our project. The right dataset must be provided to get robust results. Our data mainly consists of images of fruits of different varieties. We will be taking and analyzing data from Kaggle. After that seeing the accuracy, we will use the data in our model.

• 5..1.2 Data Preprocessing

Human can understand any type of data but machine can't our model will also learn from scratch so it's better to make the data more machine readable. Raw data is usually inconsistent or incomplete . Data preprocessing involves checking missing values, splitting the dataset and training the machine etc.

• 5.1.3 Training Model

Similar to feeding somethings, machine/model should also learn by feeding and learning on data. The data set extracted from Kaggle will be used to train the model. The training model uses a labelled set of data as the undefined dataset which is collected from Kaggle and from the same dataset a refine view is presented which is seen as the desired output. For the refining of the dataset algorithms are implemented to show the desired output.

5.2 System Architecture

The dataset we use for the proposed project is been taken from Kaggle. The data set is a collection of different types of fruits. The initial step is to convert raw data into processed data. Which is done by feature extraction, since the raw data collected have multiple attributes but only some of those attributes are needed for the prediction. Feature extraction is a reduction process.

The structure, behavior and views of a system is given by structural model.

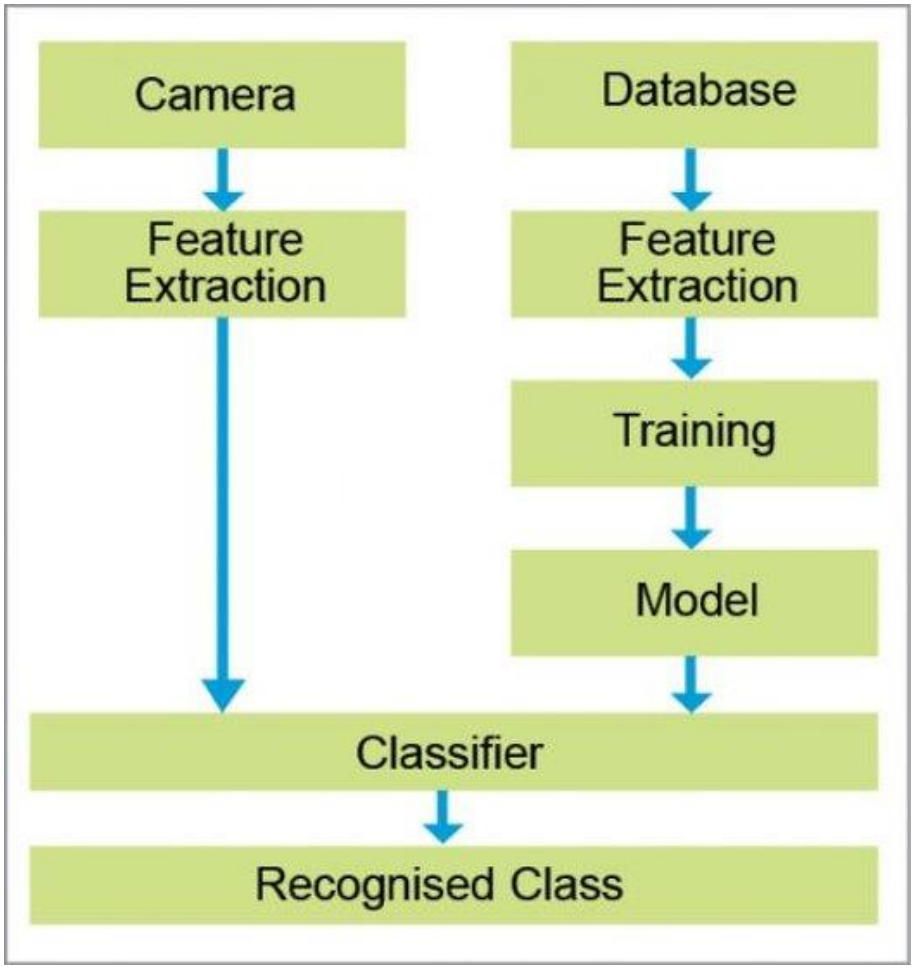

Fig 5 System Architecture

The above figure 5 gives the demonstration on the dataset extraction and refining the dataset by categorizing into two phases of training and testing data. From the given dataset a well modified classification is extracted.

5.3 Data Flow Diagram

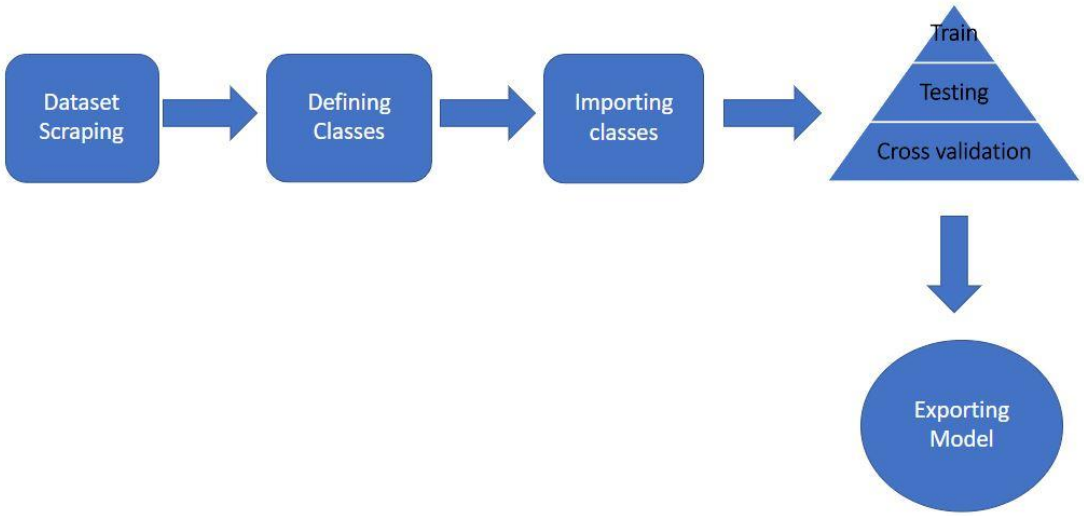

Fig 7Data Flow Diagram

In the above fig7 we are taking a dataset and defining classes for each verity, then we train the data and test it and pass it through cross validation. In this way data is flowing in our system.

Chapter 6

Implementation

These are the Machine Learning Algorithms implemented during the building of the project

6.1 CNN: -

CNN is a powerful algorithm for image processing. These algorithms are currently the best algorithms we have for the automated processing of images. Many companies use these algorithms to do things like identifying the objects in an image.

Images contain data of RGB combination. Matplotlib can be used to import an image into memory from a file. The computer doesn't see an image, all it sees is an array of numbers. Color images are stored in 3-dimensional arrays. The first two dimensions correspond to the height and width of the image (the number of pixels). The last dimension corresponds to the red, green, and blue colors present in each pixel.

Three Layers of CNN

Convolutional Neural Networks specialized for applications in image & video recognition. CNN is mainly used in image analysis tasks like Image recognition, Object detection & Segmentation.

There are three types of layers in Convolutional Neural Networks:

1) Convolutional Layer: In a typical neural network each input neuron is connected to the next hidden layer. In CNN, only a small region of the input layer neurons connects to the neuron hidden layer.

2) Pooling Layer: The pooling layer is used to reduce the dimensionality of the feature map. There will be multiple activation & pooling layers inside the hidden layer of the CNN.

3) Fully-Connected layer: Fully Connected Layers form the last few layers in the network. The input to the fully connected layer is the output from the final Pooling or Convolutional Layer, which is flattened and then fed into the fully connected layer.

6.2 Adam Optimizer

Adaptive Moment Estimation is an algorithm for optimization technique for gradient descent. The method is really efficient when working with large problem involving a lot of data or parameters. It requires less memory and is efficient. Intuitively, it is a combination of the 'gradient descent with momentum' algorithm and the 'RMSP' algorithm.

Momentum:
This algorithm is used to accelerate the gradient descent algorithm by taking into consideration the 'exponentially weighted average' of the gradients. Using averages makes the algorithm converge towards the minima in a faster pace.

Root Mean Square Propagation (RMSP):
Root mean square prop or RMSprop is an adaptive learning algorithm that tries to improve AdaGrad. Instead of taking the cumulative sum of squared gradients like in AdaGrad, it takes the 'exponential moving average'.
Adam Optimizer inherits the strengths or the positive attributes of the above two methods and builds upon them to give a more optimized gradient descent.
Here, we control the rate of gradient descent in such a way that there is minimum oscillation when it reaches the global minimum while taking big enough steps (step-size) so as to pass the local minima hurdles along the way. Hence, combining the features of the above methods to reach the global minimum efficiently.

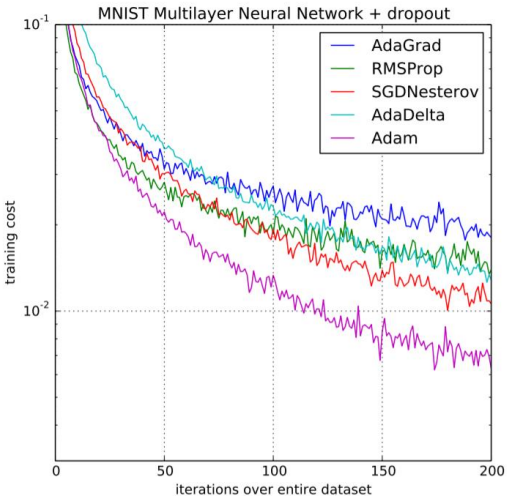

Figure 6.2

6.3 SoftMax Function

The SoftMax function is a function that turns a vector of K real values into a vector of K real values that sum to the input values can be positive, negative, zero, or greater than one, but the SoftMax transforms them into values between 0 and 1, so that they can be interpreted as probabilities. If one of the inputs is small or negative, the SoftMax turns it into a small probability, and if an input is large, then it turns it into a large probability, but it will always remain between 0 and 1.

The SoftMax function is sometimes called the soft argmax function, or multi-class logistic regression. This is because the SoftMax is a generalization of logistic regression that can be used for multi-class classification, and its formula is very similar to the sigmoid function which is used for logistic regression. The SoftMax function can be used in a classifier only when the classes are mutually exclusive.

Many multi-layer neural networks end in a penultimate layer which outputs real-valued scores that are not conveniently scaled and which may be difficult to work with. Here the SoftMax is very useful because it converts the scores to a normalized probability distribution, which can be displayed to a user or used as input to other systems. For this reason, it is usual to append a SoftMax function as the final layer of the neural network.

8.1 Conclusion

Our System strives to be fast, precise for easy adoption and high efficiency. We implement the CNN neural network framework as a fruit detector. We use a sparse, three-layer convolutional classifier as our base classifier, and alter the network in various ways to increase speed and precision. The model trained has the flexibility to adapt and learn with new dataset thus providing multifunctionality across many platforms.

8.2 Future Enhancement

Artificial Intelligence has already reached a very high level, enabling rapid and flawless analysis of data obtained during visits of sales reps in stores and photos taken by them. Image Recognition continuously changes the way manufacturers sell products in point of sales and on-line channels. It will also change the way customers behave when shopping in the store.

Computer vision involves obtaining, describing and producing results according to the field of application. Image recognition can be considered as a component of computer vision software. Computer vision has more capabilities like event detection, learning, image reconstruction and object tracking.

Machine vision is the vision system involving both hardware and computer vision software. Therefore, computer vision and image recognition can be considered as components of machine vision software.

The project can be further continued to gain the effectiveness of the classification with addition implementations of the content that can involve real time scenario and the way of executing and processing the real time scenario. Various constrains has to be added and performance of the same can be acylated in the future time for the effective results.

Chapter 9

References

1) Sai Yeshwanth Chaganti, Ipseeta Nanda, Koteswara Rao Pandi, Niraj Kumar,"Image Classification using SVM and CNN," in 2020 IEEE Xplore.

2) Shyava Tripathi, Rishi Kumar, "Image Classification using small Convolutional Neural Network" in 2019, IEEE.

3) Yunyan Wang, Chongyang Wang, Lengkun Luo, Zhigang Zhou, "Image Classification Based on transfer Learning of Convolutional neural network" in 2019 38th Chinese Control Conference, IEEE, 2019

4) Zhiyong Dong, Sheng Lin, "Research on image classification based on Capsnet," 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference, IEEE, 2019

Thank You!