Food Spoilage Detection System

Minor project report submitted in partial fulfillment of the requirement for the degree of Bachelor of Technology

in

Computer Science and Engineering

By

KULDEEP MISHRA (191263) & GEETANSH GARG (191308)

UNDER THE SUPERVISION OF

PRATEEK THAKRAL

Department of Computer Science & Engineering and Information Technology

Jaypee University of Information Technology, Waknaghat, 173234, Himachal Pradesh, INDIA

TABLE OF CONTENT

DECLARATION

I hereby declare that, this project has been done by me under the supervision of Mr. **Prateek Thakral, Assistant Professor,** Jaypee University of Information Technology. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

(Prateek Thakral)

Assistant Professor

Department of Computer Science & Engineering and Information Technology

Jaypee University of Information Technology

Submitted by:

Kuldeep Mishra 191263

Geetansh Garg 191308

Computer Science & Engineering Department

Jaypee University of Information Technology

CERTIFICATE

This is to certify that the work which is being presented in the project report titled **Food Spoilage Detection System** in partial fulfilment of the requirements for the award of the degree of B.Tech in Computer Science And Engineering and submitted to the Department of Computer Science And Engineering, Jaypee University of Information Technology, Waknaghat is an authentic record of work carried out by **Kuldeep Mishra, 191263 and Geetansh Garg, 191308** during the period from January 2022 to May 2022 under the supervision of Mr. Prateek Thakral, Department of Computer Science and Engineering, Jaypee University of Information Technology, Waknaghat.

Kuldeep Mishra 191263
Geetansh Garg 191308

The above statement made is correct to the best of my knowledge.

(**Prateek Thakral**)
Assistant Professor
Computer Science & Engineering and Information Technology
Jaypee University of Information Technology, Waknaghat

ACKNOWLEDGEMENT

Firstly, I express my heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the project work successfully.

I really grateful and wish my profound my indebtedness to Supervisor **Mr. Prateek Thakral, Assistant Professor**, Department of CSE Jaypee University of Information Technology, Waknaghat. Deep Knowledge & keen interest of my supervisor in the field of "**Deep Learning**" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

I would like to express my heartiest gratitude to **Mr. Prateek Thakral,** Department of CSE, for his kind help to finish my project.

I would also generously welcome each one of those individuals who have helped us straight forwardly or in a roundabout way in making this project a win. In this unique situation, I might want to thank the various staff individuals, both educating and non-instructing, which have developed their convenient help and facilitated my undertaking.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

Kuldeep Mishra 191263

Geetansh Garg 191308

ABSTRACT

Food spoilage is the process where a food product becomes unsuitable to ingest by the consumer. The cause of such a process is due to many outside factors as a side-effect of the type of product it is, as well as how the product is packaged and stored. Bacteria and various fungi are the cause of spoilage and can create serious consequences for the consumers, such as nausea, vomiting, and diarrhea.

Humans can only detect the quality of the food by just checking out the food from the naked eye or by smelling the food product. But, in most cases, humans fail to predict the quality of the food product by just looking at the food item. Food spoilage not only depends on the appearance of food, or the smell of the food item but also on the oxygen level, nitrogen level, and temperature around the food product. Food quality also depends on the pH level i.e., the acid level in the food. Therefore, we are developing a "Food Spoilage Detection System" which will help millions of people from consuming spoiled food.

We are predicting the quality of the food item by first capturing the images of the food and then using our machine learning algorithm to classify the food quality. We have used Convolutional Neural Networks (CNN) for the classification of the quality of the food from the images. We have used Neural Network for image classification of spoiled food because of the ability of the Neural Network to quickly gather useful information from the images. CNN makes the classification of food products both fast and efficient.

Our future aspect will be of working on with oxygen level sensors, nitrogen level sensors and temperature around the food items. We had already gathered information about these mentioned sensors and how they actually affect food products. We will be looking forward to make this system a full-fleshed predicting device with all the required sensors installed.

Chapter 01: INTRODUCTION

1.1 Introduction

Food is any substance consumed to provide nutritional support for an organism. Food plays an important role in the promotion of health and disease prevention. Food provides energy to our body and helps us to regulate our daily activities. But many times, the food we are eating may lead to adverse effects on our body such as nausea, vomiting, and diarrhea. These adverse effects generally rise when we eat spoilt food i.e., the food we are consuming is not good for our health. The determination of the contaminated food by a lay-man becomes difficult as it depends on numerous factors such as oxygen level, odor, nitrogen level, acidity, appearance of the food product. Therefore, we will be focusing on developing a full-fleshed device which help humans to determine the quality of the food item, they are going to consume.

1.2 Objective

The main objective of this project is to help millions of people from consuming spoiled food. Since, most of food industries brands are still checking the quality of the food items manually, which leads to high chance of error in prediction of the quality of the food product. Therefore, we are aiming to provide an automated device to food industries to find out spoiled food products as they directly lead to the loss of the entire batch.

1.3 Motivation

According to WHO – Food Safety, there is an estimate that out of every 10 people, 1 is suffering from food poisoning and around 2.89% of the population of the world dies from food poisoning which is on its own a huge number of deaths. Therefore, to prevent people from getting ill due to food poisoning, we are developing a “Food Spoilage Detection System”.

1.4 Language Used

Python is the language which is mainly used in this project. Since, we are working with image classification between good food and spoiled food which requires machine learning algorithm to successfully predict the result and python is best known for working with machine learning algorithms. The various libraries of python are used to implement the given project. The **TensorFlow** library is the main library used in the project and other libraries like matplotlib, NumPy and cv2 are also used to implement the project.

1.5 Technical Requirements

The requirements for this project are:

- CPU: Intel i5/AMD Ryzen 5 or above
- GPU: NVIDIA GeForce 110/NVIDIA GeForce GTX or above
- RAM: 8GB or more
- Hard Disk Drive (HDD): 500 GB or more
- Solid State Drive (SSD): 128 GB or more
- Operating System: Windows 10 or above/Linux/Unix
- Software: VS CODE with Tensorflow (GPU) Library

1.6 Deliverables/Outcomes

The outcome of the project is a working machine learning model which helps to predict whether the food product is of good quality or not by using the image of the food product. The model is working with 97% accuracy in predicting whether the food is fresh or spoiled.

Chapter 02: Feasibility Study, Requirements Analysis and Design

2.1 Feasibility Study

2.1.0 Literature Review

- **FOOD QUALITY DETECTION AND MONITORING SYSTEM** by Atkare Prajwal, Patil Vaishali, zade payal, Dhapudkar Sumit

  The system is developed for the people so that they can identify the quality of food. The sensors like the PH sensor, odor sensor, temperature sensor can help to detect the food quality. The diseases due to the spoilage of food can be overcome up to great extent.

- **FRESHNESS OF FOOD DETECTION USING IOT AND MACHINE LEARNING** by Nachiketa Hebbar

  The research has led us to conclude that the food industry can be revolutionized by a simple combination of sensors, IoT and machine learning. After integration, this model will create a competition between food manufacturers to sell more healthy food and create awareness among consumer to purchase more healthy food.

2.1.1 Problem Definition

In today's world, even after so much advancement in technology, a lot of death occurs due to food poisoning. Bacteria and other harmful microorganisms can't be detected through naked eyes and by their odor. Most of the people trust their vendor and they simply purchase the contaminated food. Many of us don't have enough time to check for each and every single fruit and food. Therefore, we have come up with an innovative solution for all these problems.

2.1.2 Problem Analysis

By carefully observing the contaminated food and performing various tests on them, we have come up few conclusions:

- Outer appearance of the food completely changes once they are spoilt.
- Spoilt food produces pungent odor.
- Oxygen level decreases around the contaminated food as bacteria consumes more oxygen.
- Concentration of few gases such as Ammonia increases near the spoilt food.

2.1.3 Solution

As per the solution is concerned, we have built a machine learning model that will scan the images of the food and categories them whether they are spoilt or not.

We have used CNN (CONVOLUTION NEURAL NETWORK) for this purpose. The reason why CNN is chosen over ANN is because:

- **Hardware Dependence**:
    - Artificial Neural Networks require processors with parallel processing power, by their structure.
    - For this reason, the realization of the equipment is dependent.
- **Unexplained functioning of the network**:
    - This the most important problem of ANN.
    - When ANN gives a probing solution, it does not give a clue as to why and how.
    - This reduces trust in the network.
- **Assurance of proper network structure**:
    - There is no specific rule for determining the structure of artificial neural networks.
    - The appropriate network structure is achieved through experience and trial and error.
- **The difficulty of showing the problem to the network**:
    - ANNs can work with numerical information.

- Problems have to be translated into numerical values before being introduced to ANN.
- The display mechanism to be determined will directly influence the performance of the network.
- This is dependent on the user's ability.

- **The duration of the network is unknown**:
  - The network is reduced to a certain value of the error on the sample means that the training has been completed.
  - The value does not give us optimum results.

Now if we talk about CNN, it has got some advantages over other neural networks. Such as:
Over the years, research on convolutional neural networks (CNNs) has progressed rapidly, however the real-world deployment of these models is often limited by computing resources and memory constraints. What has also led to extensive research in ConvNets is the accuracy of difficult classification tasks that require understanding abstract concepts in images.

Another reason why CNN are hugely popular is because of their architecture, the best thing is there is no need for feature extraction. The system learns to do feature extraction and the core concept of CNN is, it uses convolution of image and filters to generate invariant features which are passed on to the next layer. The features in next layer are convoluted with different filters to generate more invariant and abstract features and the process continues till one gets final feature / output (let say face of X) which is invariant to occlusions.

Also, another key feature is that deep convolutional networks are flexible and work well on image data. As one researcher points out, convolutional layers exploit the fact that an interesting pattern can occur in any region of the image, and regions are contiguous blocks of pixels. But one of the reasons why researchers are excited about deep learning is the potential for the model to learn useful features from raw data. Now, convolutional neural networks can extract informative features from images, eliminating the need of traditional manual image processing methods.

Our model can detect food spoilage in various fruits such as apples, bananas. We have used more than 12000 images to train our model. As a result, our model has an accuracy of over 97 percentage. You can surely rely on it.

For the future aspect, in order to improve the result, we are planning to include temperature sensor, oxygen level measuring device, ammonia detector as well as odor detector.

We will fit all these devices with our machine learning model and sell it to vendors. The whole system will be fully automated. It will scan the foods at real time and predict whether quality of food.

2.2 Requirements

2.2.1 Functional Requirements

The functional requirements for the Food Spoilage Detection are: -

- The model must be able to handle any kind of errors produce during the model predictions.
- The model must have a good architecture which can handle the food item images having very less quality.
- The model must have a good output producing source.
- The model must be compatible to use with any other model in order to create a new and big project.

2.2.2 Non-Functional Requirements

The non-functional requirements for the Food Spoilage Detection are: -

- The model can run in various machines which support the model environment.
- The model gives a fast response.

2.3 E-R Diagram / Data-Flow Diagram (DFD)

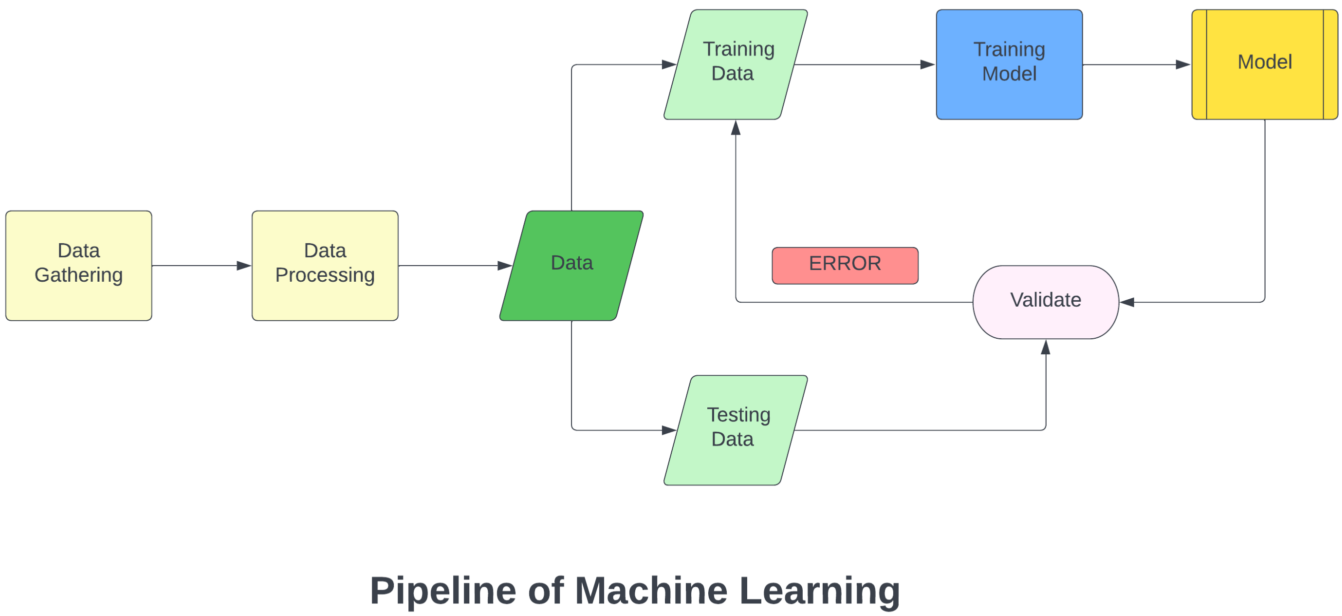

Figure 2.1: Pipeline of Machine Learning

Pipeline of Machine Learning

A machine learning pipeline is used to help automate machine learning workflows. Machine learning (ML) pipelines consist of several steps to train a model. Machine learning pipelines are iterative as every step is repeated to continuously improve the accuracy of the model and achieve a successful algorithm.

Data collection and cleaning are the primary tasks of any machine learning project. But getting data and especially getting the right data is an uphill task in itself. Data quality and its accessibility are the next two main challenges one will come across in the initial stages of building a pipeline. Now, the processed data gets divided into two parts Training data and Testing data. Training data is used for training the model, while the testing data is used for validating the trained model.

With the ML pipeline, each part of the workflow is abstracted into an independent service. Then, each time we design a new workflow, we can pick and choose which elements we need and use them where we need them, while any changes made to that service will be made on a higher level.

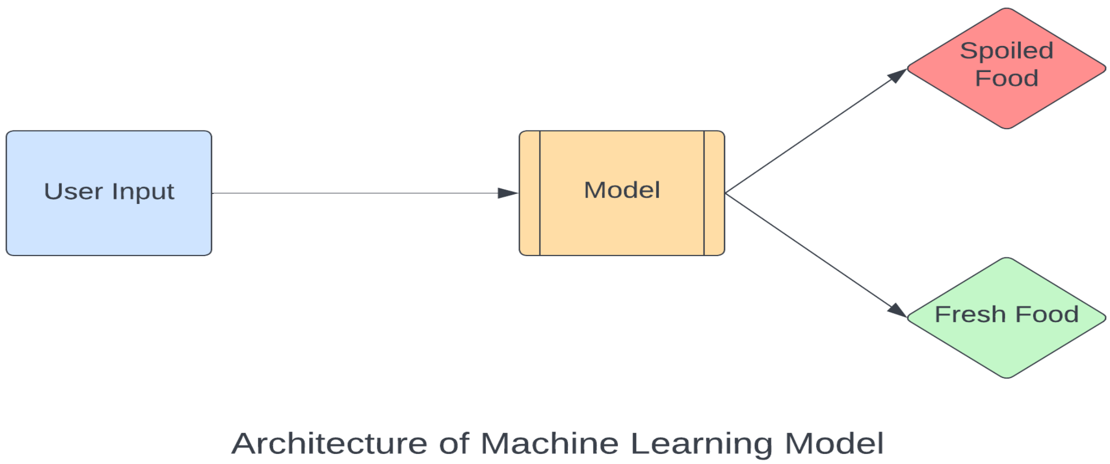

Figure 2.1: Architecture of Machine Learning

Our proposed model has following steps: -

- The model takes food item image as an input.
- The model then computes the image.
- Finally, it predicts whether the food is fresh or spoiled.

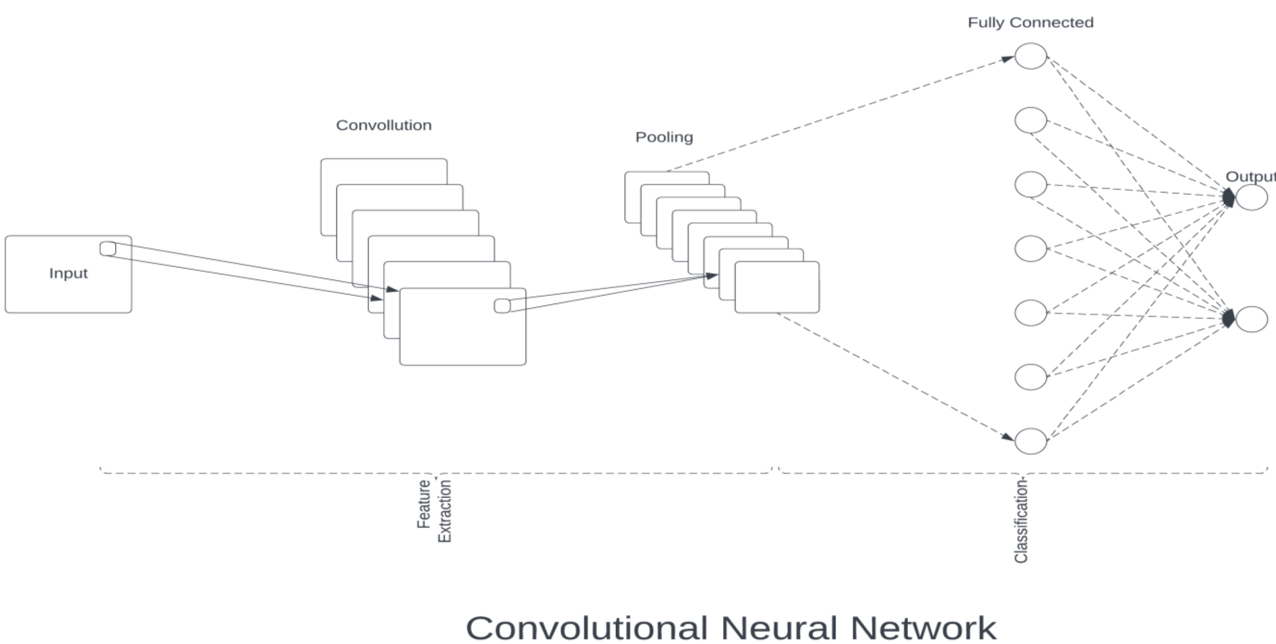

Figure 2.3: CNN Working Model

Convolutional Neural Network

In deep learning, a convolutional neural network (CNN) is a class of deep neural networks, most commonly applied to analyze visual imagery. Now when we think of a neural network, we think about matrix multiplications but that is not the case with CNN. It uses a special technique called Convolution. Now in mathematics convolution is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other. Convolutional neural networks are composed of multiple layers of artificial neurons. Artificial neurons, a rough imitation of their biological counterparts, are mathematical functions that calculate the weighted sum of multiple inputs and outputs an activation value. When you input an image in a CNN, each layer generates several activation functions that are passed on to the next layer. The first layer usually extracts basic features such as horizontal or diagonal edges. This output is passed on to the next layer which detects more complex features such as corners or combinational edges. As we move deeper into the network it can identify even more complex features such as objects, faces, etc. Based on the activation map of the final convolution layer, the classification layer outputs a set of confidence scores (values between 0 and 1) that specify how likely the image is to belong to a "class".

The CNN architecture consists of several kinds of layers; Convolutional layer, pooling layer, fully connected input layer, fully connected layer and fully connected output layer.

- **Convolutional layer:** Convolutional layer is the backbone of any CNN working model. This layer is the one where pixel by pixel scanning takes place of the images and creates a feature map to define future classifications.
- **Pooling layer**: Pooling is also known as the down-sampling of the data by bringing the overall dimensions of the images. The information of each feature from each convolutional layer is limited down to only containing the most necessary data.
- **Fully connected input layer:** This is also known as the flattening of the images. The outputs gained from the last layer are flattened into a single vector so that it can be used as the input data from the upcoming layer.
- **Fully connected layer:** After the feature analysis has been done and it's time for computation, this layer assigns random weights to the inputs and predicts a suitable label.
- **Fully connected Output layer:** This is the final layer of the CNN model which contains the results of the labels determined for the classification and assigns a class to the images.

Chapter 03: IMPLEMENTATION

3.1 Date Set Features

3.1.1 Date Set Used in the Minor Project

The dataset used for the project is an image-based dataset consisting of numerous numbers of images of food items. The dataset consists of real images of various fruits (banana, apple, etc.) in JPG and PNG format. We have used about 12000 images as training dataset.

3.1.2 Description of the data set

The dataset we have chosen consists of around 20000 different images of food items. Approximately 48% of them are rotten classified images and rest 52% are fresh. This equality in the two categories helped us to reduce the chances of “Dumb Model”. The images are either in the format of JPG or PNG.

3.2 Steps of Execution

1. Data Collection
2. Data Cleaning
3. Data Preprocessing
4. Splitting the Data
5. Training the Model
6. Evaluation of Model
7. Predicting the food items

3.3 Algorithm of the Project Problem

- Firstly, the dataset of images is divided into two different parts: Training data and Testing data.
- The images of the dataset are then renamed to avoid chances of mismatch failure as it increases the chances of model being either “Dumb Model” or case of “Underfitting”.
- At the next step, the training set images are mapped according to the corresponding category and stored to a new csv file for further uses. (0 for **Rotten** & 1 for **Fresh)**
- Now, we resize all the images to dimensions of (200, 200) i.e., 200 width and 200 height. It helps to increase the efficiency of the model by making dimensions of all images to the same value.
- Now, we start building the model for predicting the food quality. We have used “Os, Pillow, Tensorflow, NumPy, Pandas, etc.” libraries for building the correct model.
- Keras module of Tensorflow library has been used to execute the CNN (Convolutional Neural Network) algorithm for image classification.
- Now, after the successful training of the model, evaluation of the model is done using the keras.model.evalute() method which takes two required arguments : testing_data, predicted_vals.
- Now, if the accuracy of the model is not good enough than the training of model again takes place until a good accuracy is achieved.
- Finally, we can attempt for the prediction of the food items using our trained model.

3.4 Flow graph of the Minor Project Problem

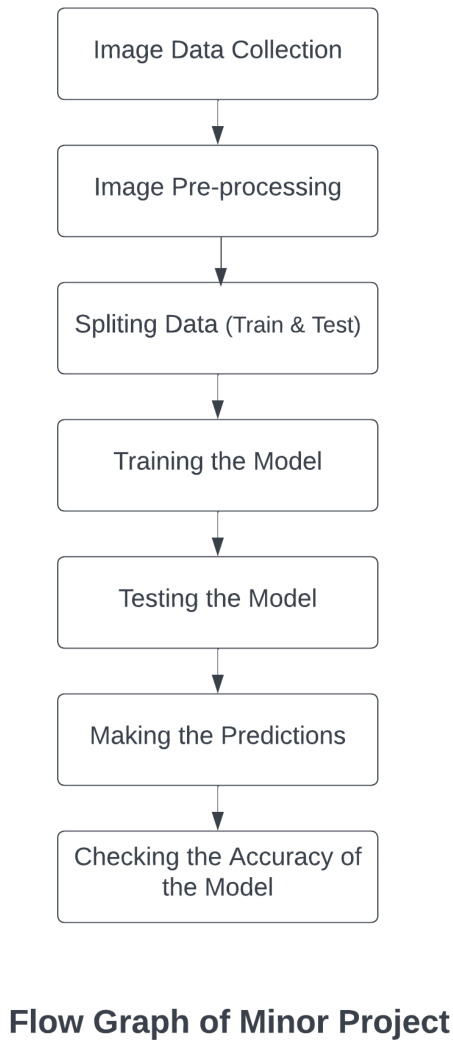

Figure 3.1: Flow Graph of Minor Project

3.6 Screen shots of the various stages of the Project

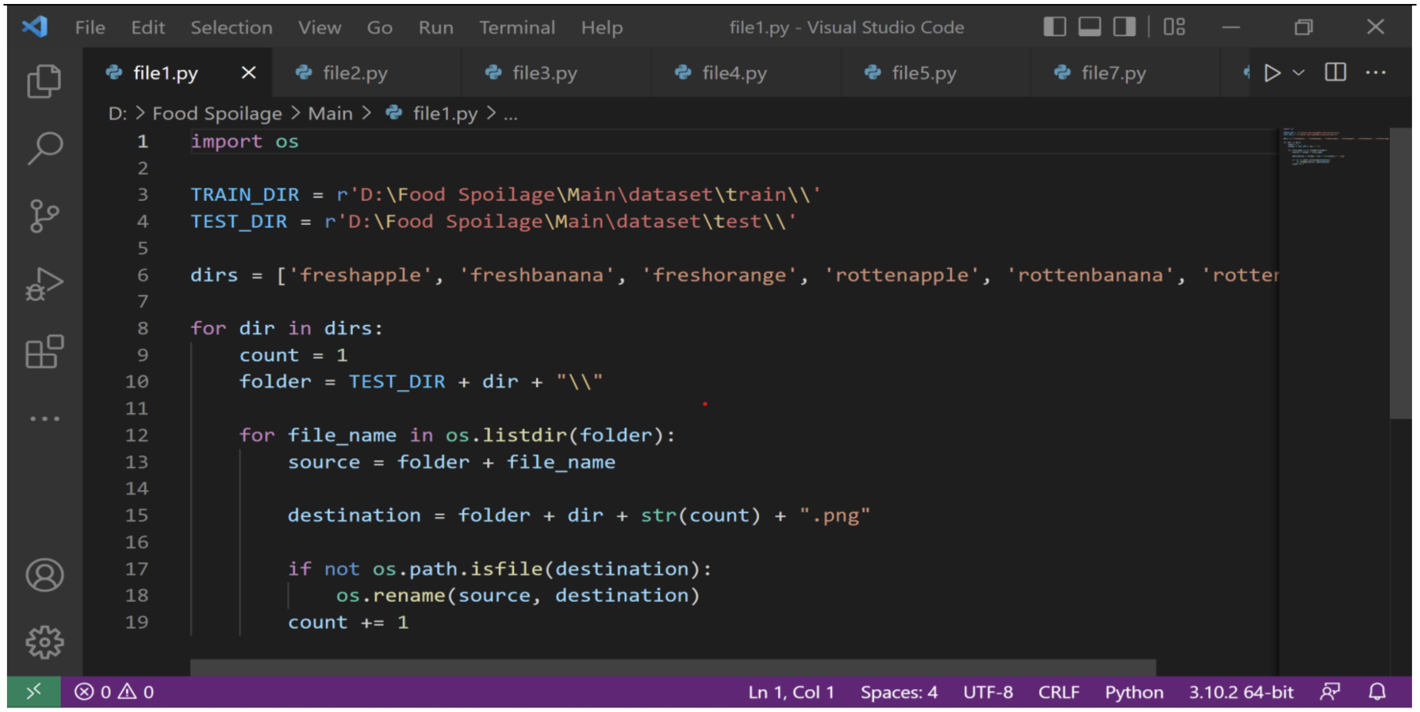

Figure 3.2: Renaming all the image files of the dataset for making csv file

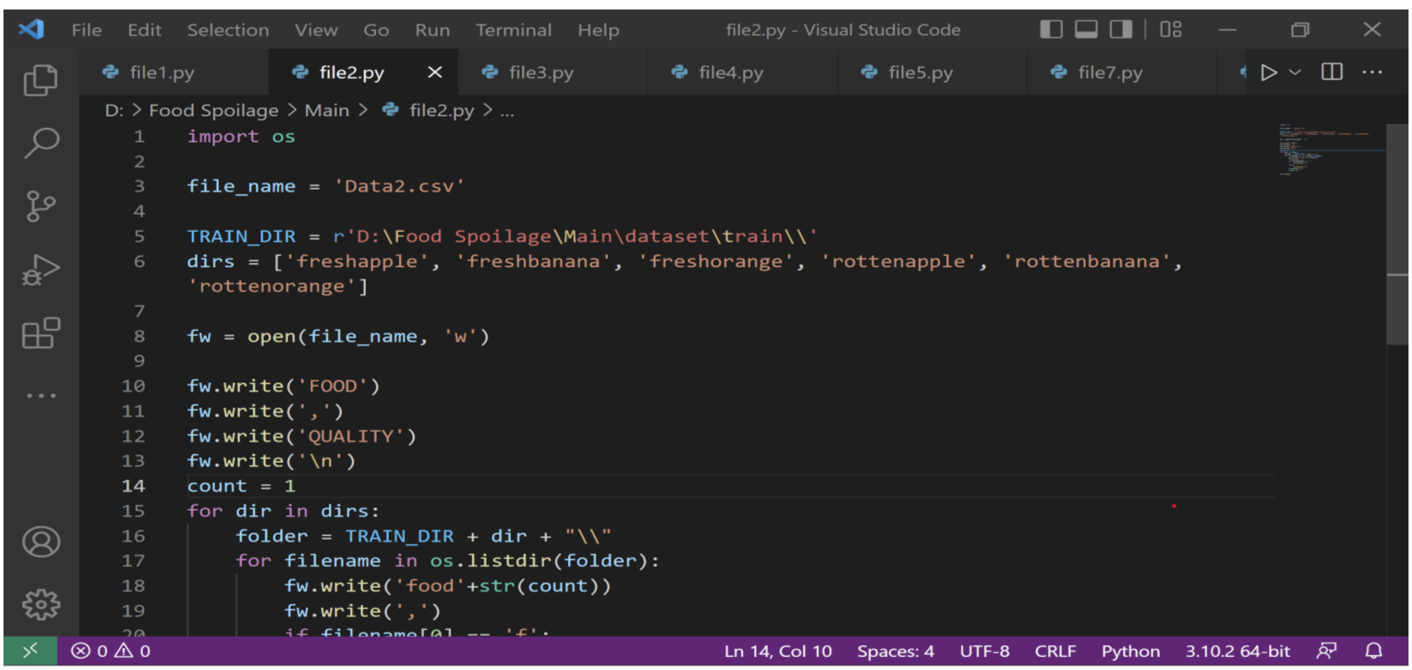

Figure 3.3: Entering of image dataset to csv file

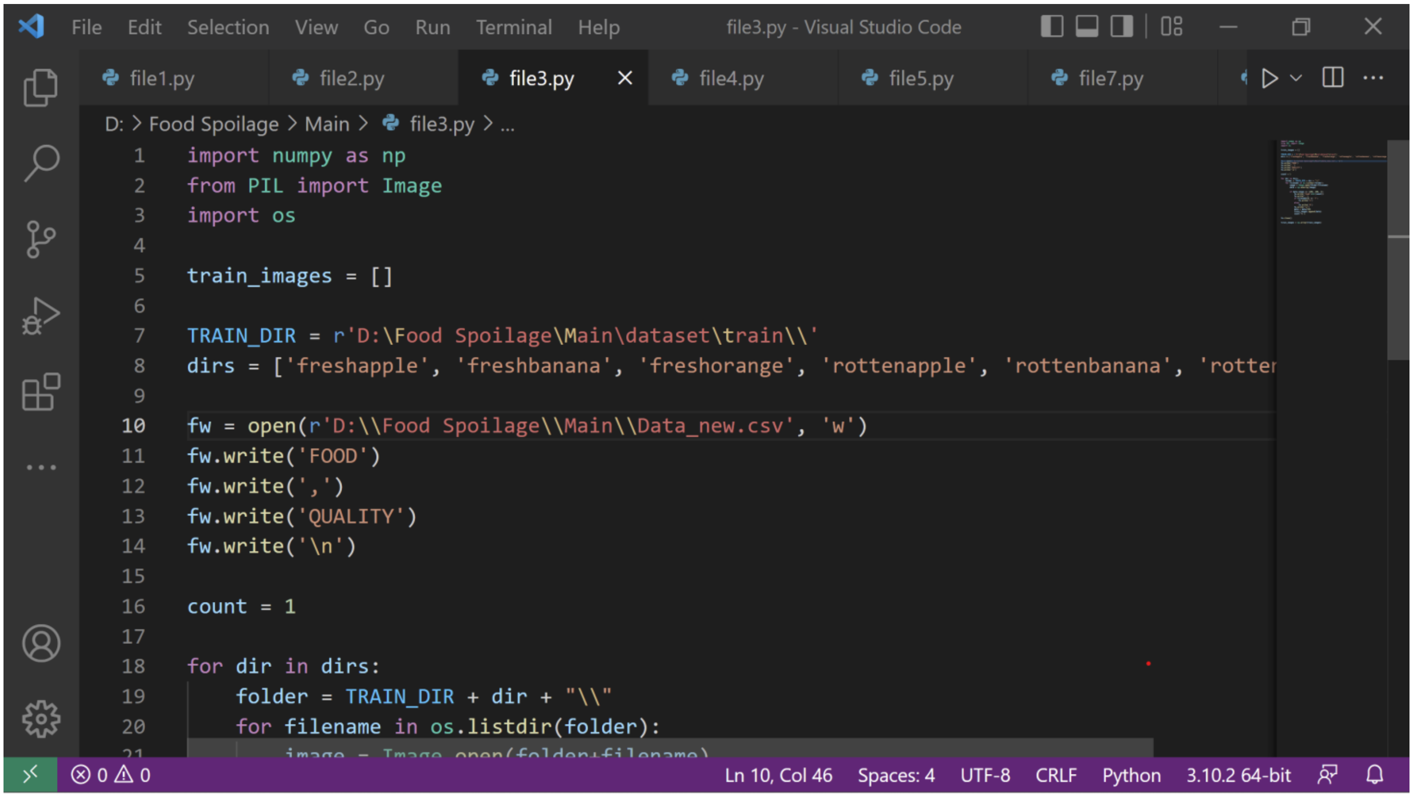

Figure 3.4: Converting the image to np array and storing in 4d np array – (i)

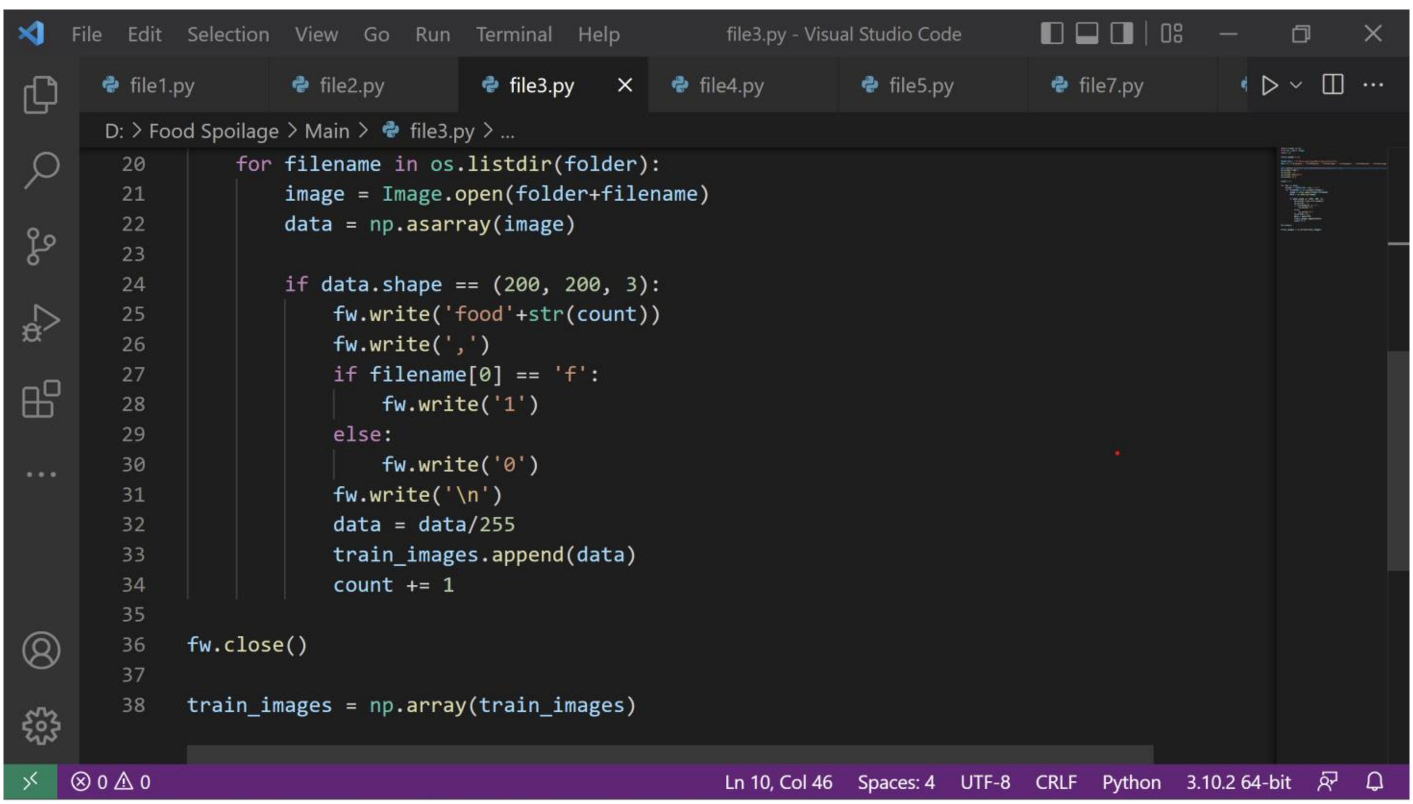

Figure 3.5: Converting the image to np array and storing in 4d np array – (ii)

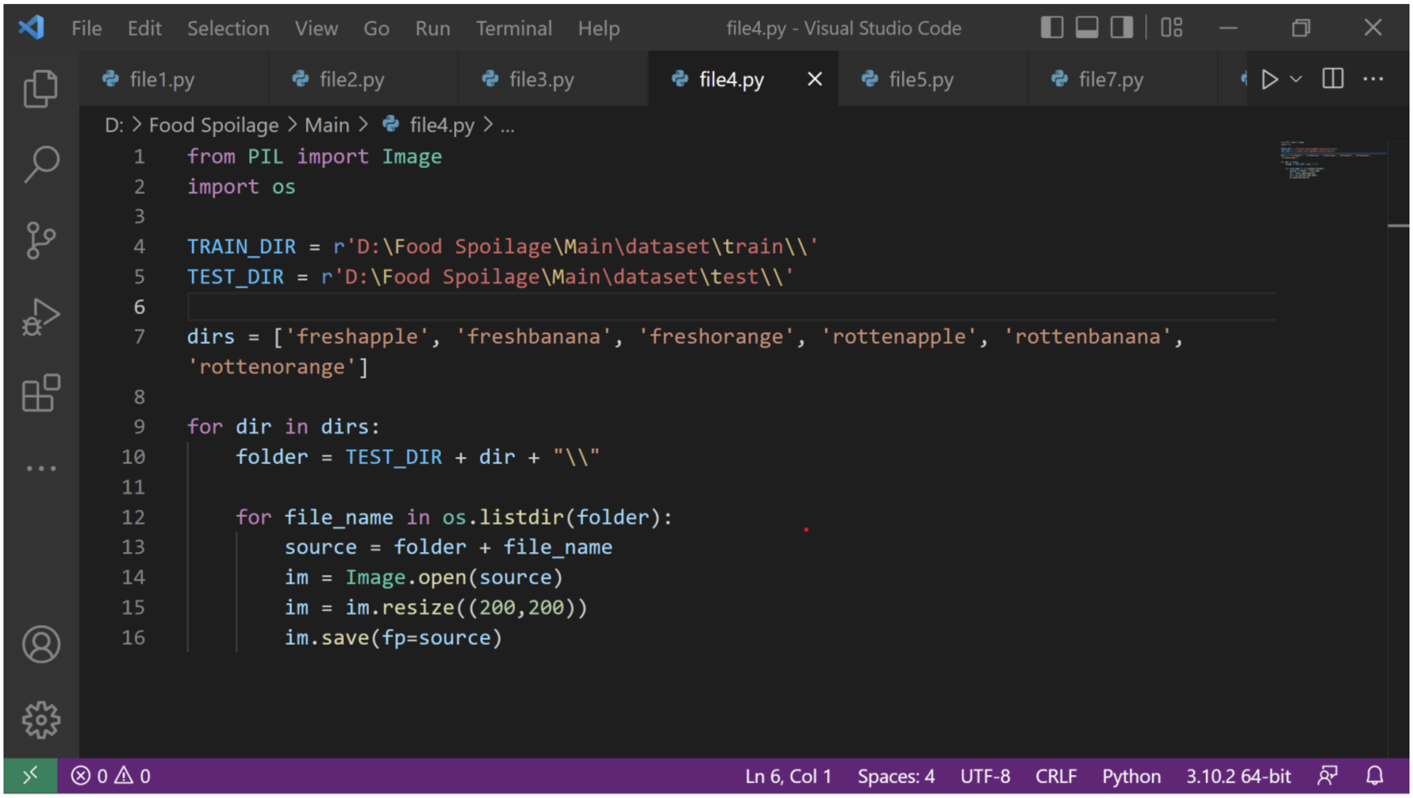

Figure 3.6: Resizing all the images to the dimension 200 x 200 x 3 and saving them

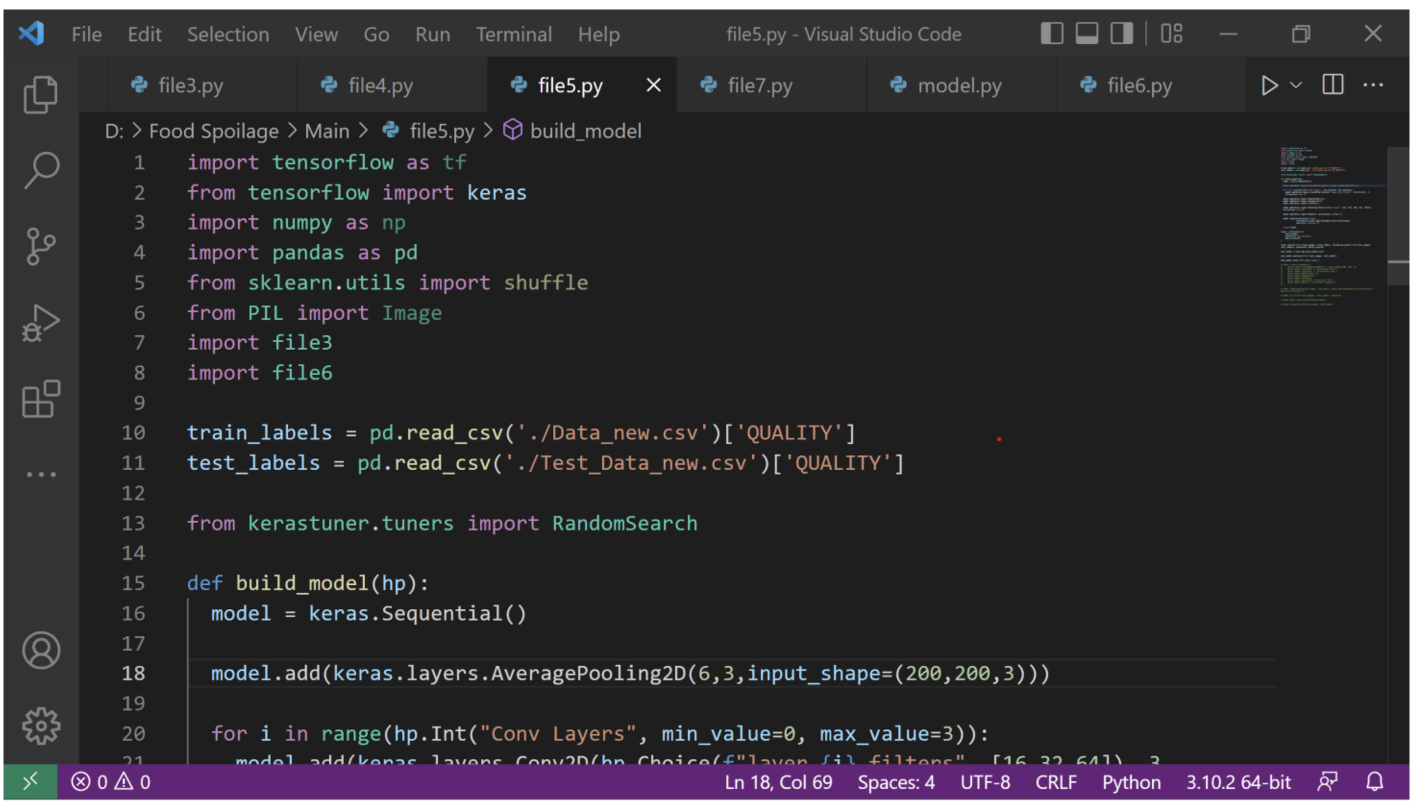

Figure 3.7: Training the model – (i)

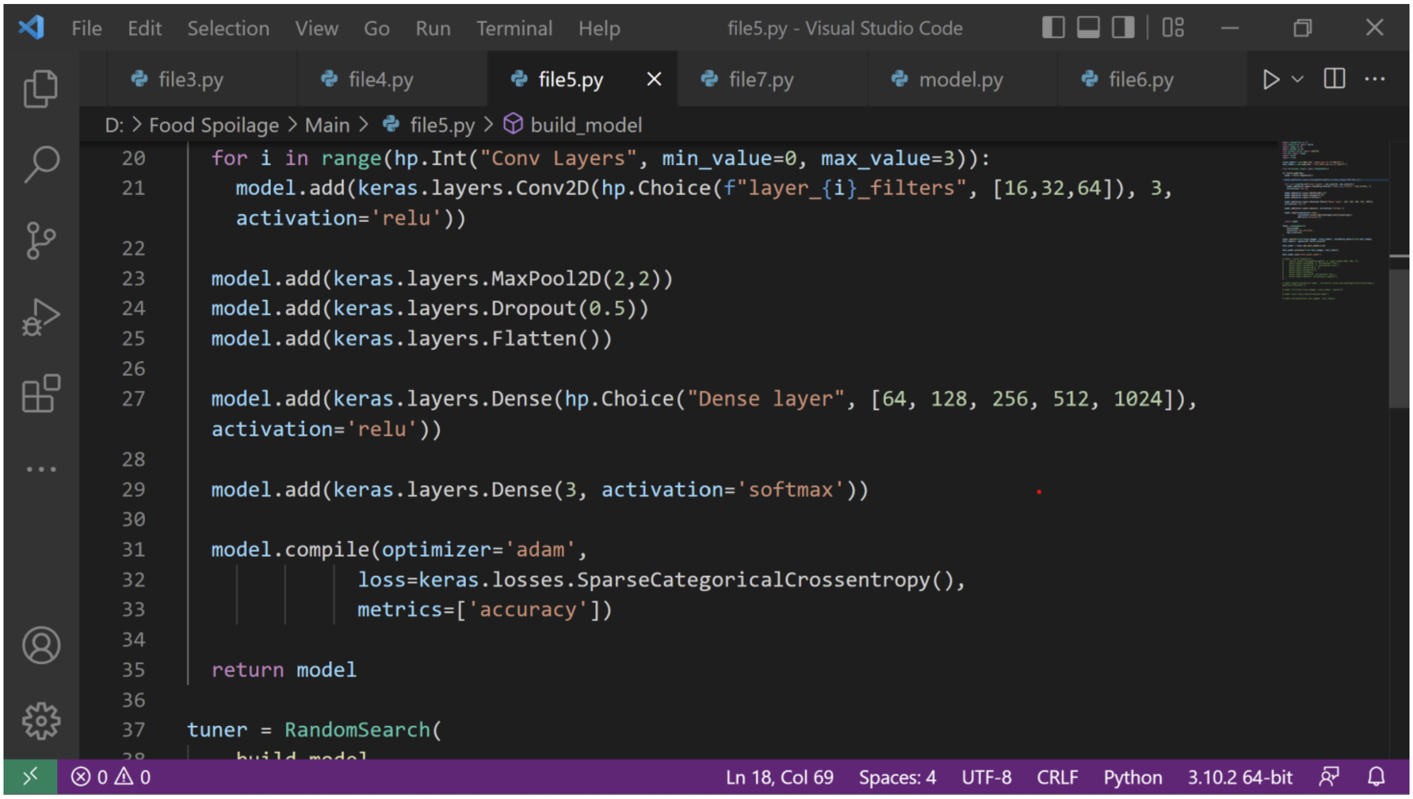

Figure 3.8: Training the model – (ii)

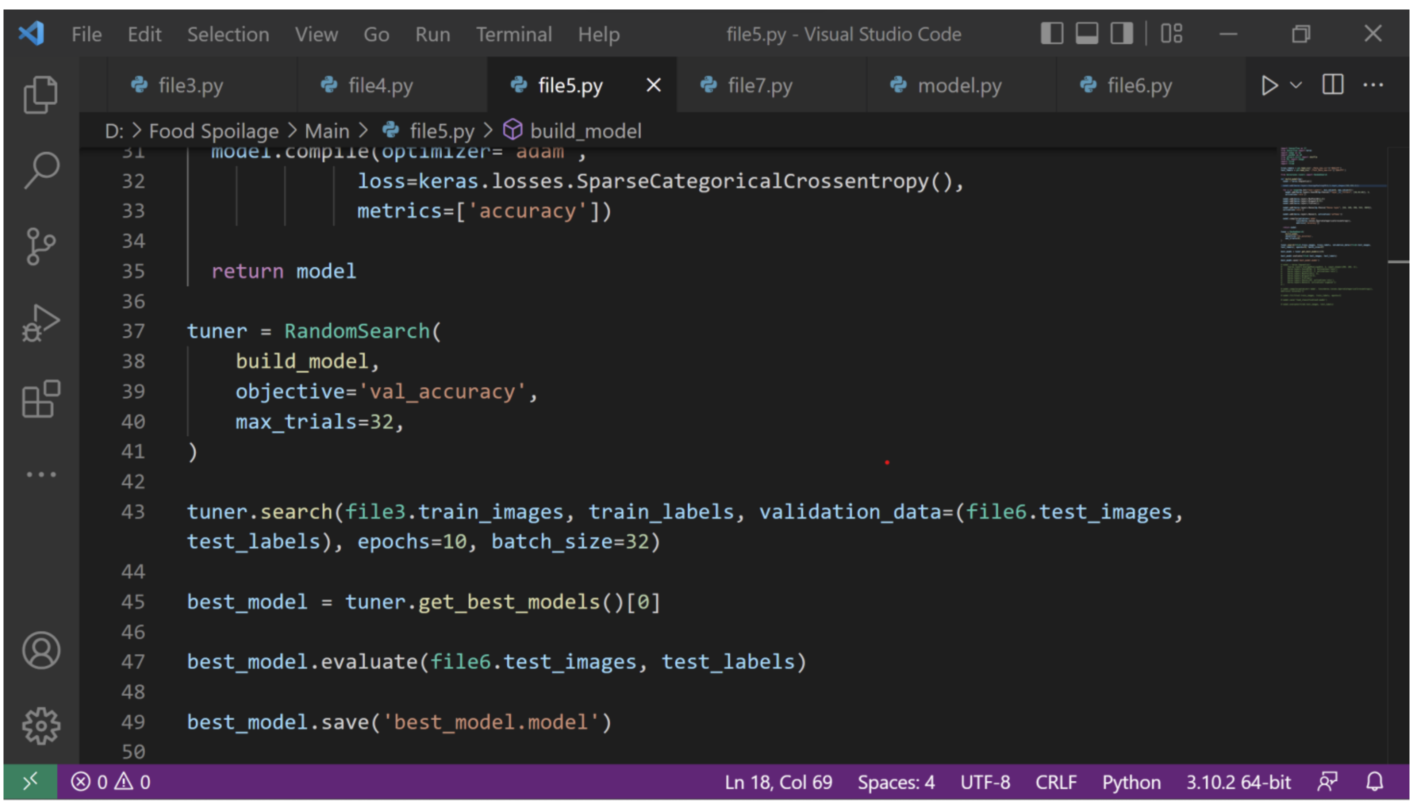

Figure 3.9: Training the model – (iii)

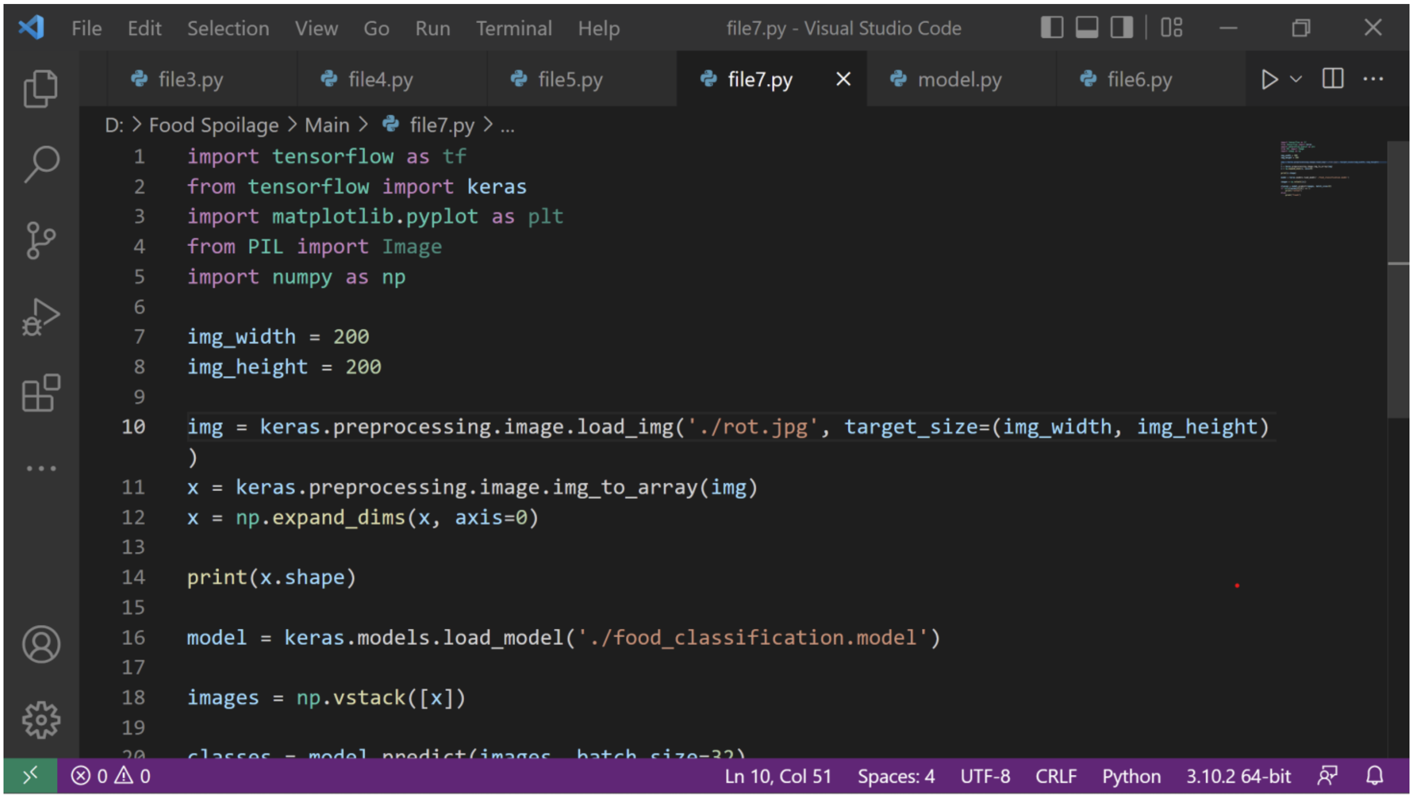

Figure 3.10: Code for predicting the image – (i)

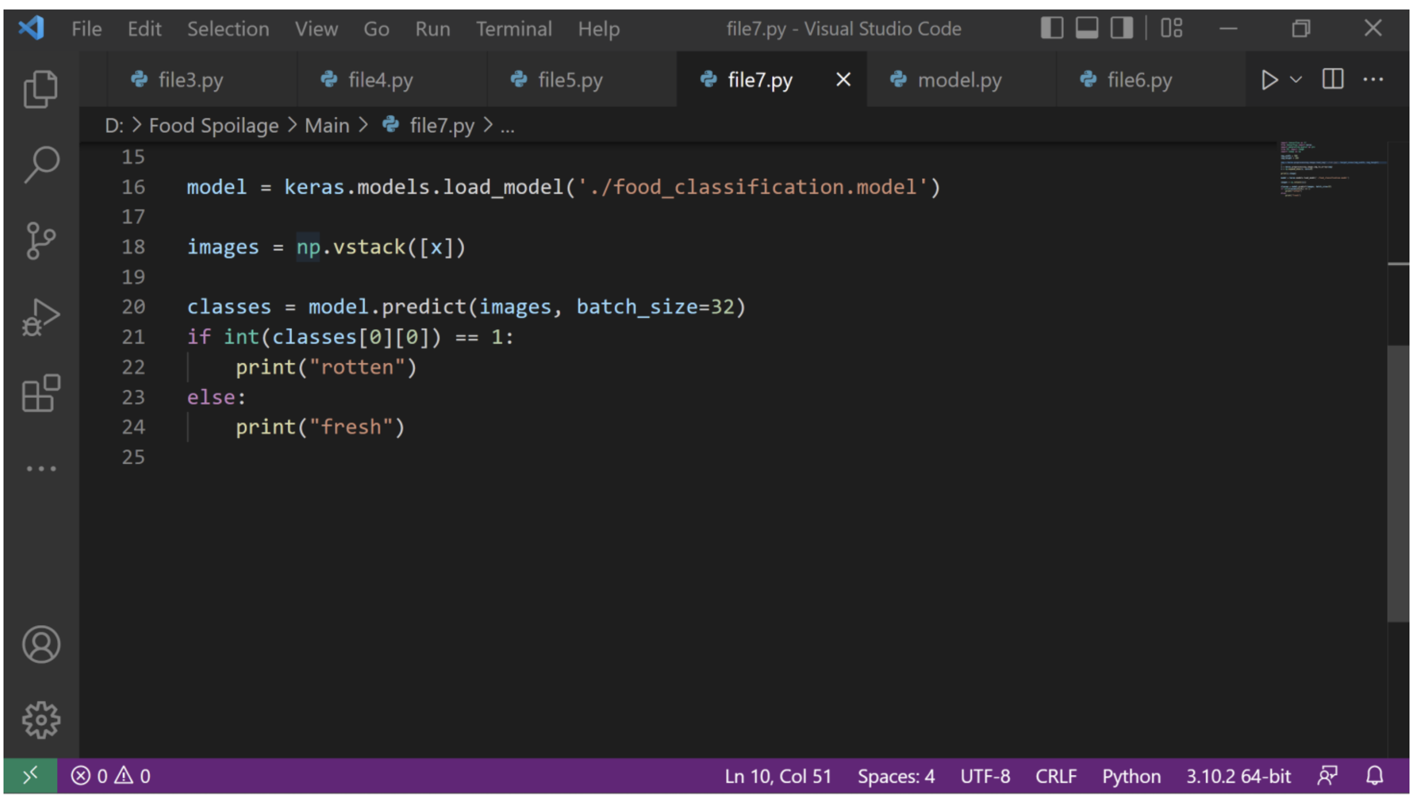

Figure 3.11: Code for predicting the image – (ii)

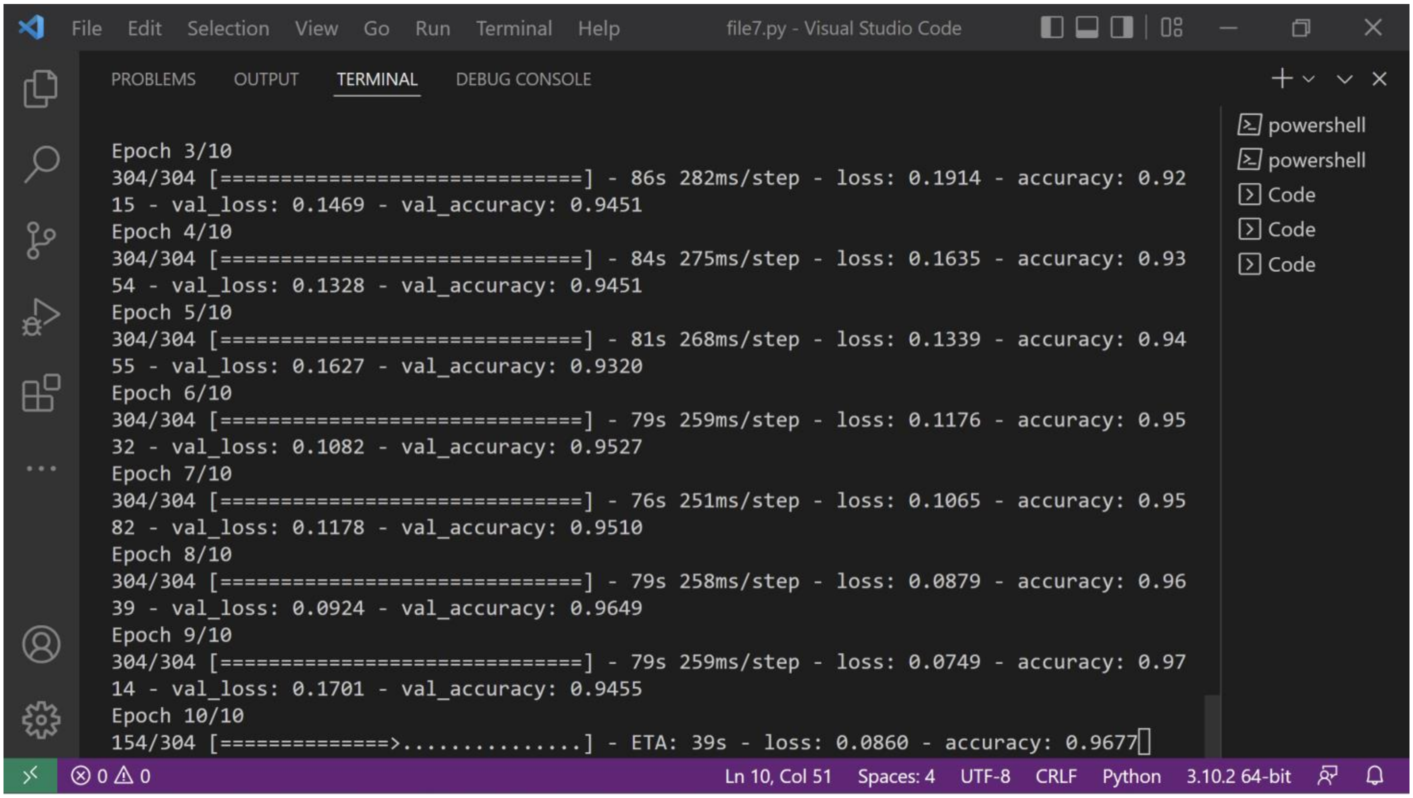

Figure 3.12: Running of code – (i)

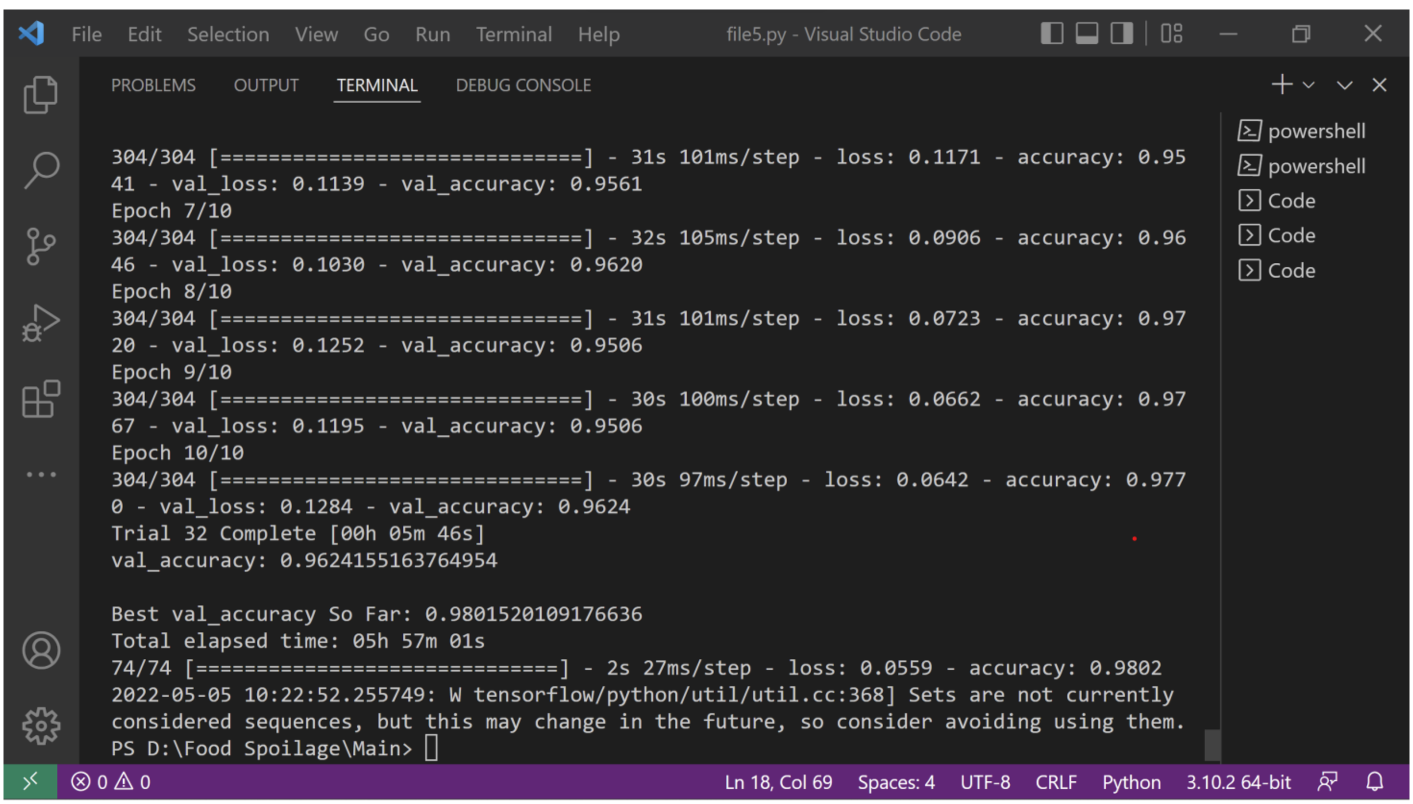

Figure 3.13: Running of code – (ii)

Chapter 04: RESULTS

4.1 Discussion on the Results Achieved

Initially, when we used about 1000 training data our machine learning model was giving an accuracy of 51 percentage due to under fitting of data. So, we increased the quantity of training data to over 15000. We were shocked to observe that this time the accuracy was about 71 percentage. We studied about this problem from the browser and came to know that this occurs due to over fitting of training data. Due to this our ML algorithm was getting confused to come to a proper conclusion. After many trials and errors, we have chosen the amount of training dataset to be around 12000. This time we are getting an accuracy of 95.82 percent.

4.2 Application of the Minor Project

There are various applications of Food Spoilage Detection System:

- It can be used at mega factories where raw fruits are packed and sold to different vendors across the country.
- It can be used at shopping complexes such as Big Bazaar, etc.
- It can even be used at small vendor shops.
- Furthermore, our Food Spoilage Detection System can be installed in refrigerators at homes and can inform us whether the fruits and foods are suitable for consumption or not.

4.3 Limitation of the Minor Project

Some of the limitations of our ML model are:

- The project that we have created requires very high-end computers i.e., a computer needs to have a GPU unit in order to work properly and efficiently.
- Even though our ML model has an accuracy of over 95 percentage, yet it is not 100 percent. So, in few cases our model is going to give wrong output.
- Image pre-processing is one of the major prerequisites for the project.
- Furthermore, our completed Food Spoilage Detection System device with oxygen level and nitrogen level sensors will be a costly device.

4.4 Future Work

Our future aspects:

- Building a machine learning model for oxygen level sensor and nitrogen level sensor.
- Start working with IOT devices (oxygen level sensor and nitrogen level sensor) to get real time measure of both the features
- Connecting all three components – Camera, Oxygen Sensor and Nitrogen Sensor to predict the food quality.

References

[1] Nachiketa Hebbar, Freshness of Food Detection using IoT and Machine Learning, MA, India: VIT Vellore University, 2020

[2] Atkare Prajwal, Patil Vaishali, zade payal and Dhapudkar Sumit(4), Food Quality Detection and Monitoring System, India: Datta Megha Institute of Engineering Technology and Research, Wardha Maharashtra, 2020

[3] Hoo-Chang Shin, Member, IEEE, Holger R. Roth, Mingchen Gao, Le Lu, Senior Member, IEEE, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M. Summers, Deep Convolutional Neural Networks, USA, IEEE Medical Imaging, 2016

[4] D. Lu & Q. Weng (2007) A survey of image classification methods and techniques for improving classification performance, International Journal of Remote Sensing, 28:5, 823-870, 2007

[5] W. Rawat and Z. Wang, "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review," in Neural Computation, vol. 29, no. 9, pp. 2352-2449, Sept. 2017, doi: 10.1162/neco_a_00990.

[6] Lone Gram, Lars Ravn, Maria Rasch, Jesper Bartholin Bruhn, Allan B Christensen, Michael Givskov, Food spoilage—interactions between food spoilage bacteria, International Journal of Food Microbiology, Volume 78, Issues 1–2, 2002