Computer Engineering Department

A.P. Shah Institute of Technology

— G.B.Road,Kasarvadavali, Thane(W), Mumbai-400615

UNIVERSITY OF MUMBAI

Academic Year 2020-2021

A Project Report on

Food Detection and Recognition

Submitted in partial fulfillment of the degree

of

Bachelor of Engineering(Sem-7)

in

Computer Engineering

By

Anshul Khairnar (17102038)
Shreya Chaudhari (17102007)
Amruta Deshpande (17102004)

Under the Guidance of

Prof. Sachin Malave
Prof. Ramya R.B

1.Project Conception and Initiation

1.1 Abstract

In this project, we propose a new deep convolutional neural network (CNN) configuration to detect food images. Various types of food with different color and texture reflect the fact that the food image detection is considered a challenging task. However, deep learning has been widely used as an efficient image detection method, and CNN is the contemporary approach for deep learning to be implemented. CNN has been optimized to the tasks of food detection with few modifications.

For evaluation of detection performance, CNN achieved significantly higher accuracy than traditional approaches with manually extracted features. Additionally, it was found out that convolution masks show that the features of food color dominate the features map. For the process of food detection, CNN also exhibited considerably higher accuracy than other conventional methods.

1.2 Objectives

1. To develop a model to classify different food items from other images (which maybe non-food) using Machine Learning and Deep Learning.

2. To analyze the different food items and predict, recognize the output effectively.

3. To create a frontend GUI application which would be easy for the user to navigate, using Flutter.

1.3 Literature Review

1.Food Detection and Recognition Using Convolutional Neural Network[1]

In this paper, we apply a convolutional neural network (CNN) to the tasks of detecting and recognizing food images. Be-cause of the wide diversity of types of food, image recognition of food items is generally very difficult. However, deep learning has been shown recently to be a very powerful image recognition technique, and CNN is a state-of-the-art approach to deep learning. We applied CNN to the tasks of food detection and recognition through parameter optimization. We constructed a dataset of the most frequent food items in a publicly available food-logging system, and used it to evaluate recognition performance. CNN showed significantly higher accuracy than did traditional support-vector-machine-based methods with handcrafted features. In addition, we found that the convolution kernels show that color dominates the feature extraction process. For food image detection, CNN also showed significantly higher accuracy than a conventional method did.

2.Machine Learning Based Approach on Food Recognition and Nutrition Estimation[2]

Nowadays, standard intake of healthy food is necessary for keeping a balanced diet to avoid obesity in the human body. In this paper, we present a novel system based on machine learning that automatically performs accurate classification of food images and estimates food attributes. This paper proposes a deep learning model consisting of a convolutional neural network that classifies food into specific categories in the training part of the prototype system. The main purpose of the proposed method is to improve the accuracy of the pre-training model. The paper designs a prototype system based on the client server model. The client sends an image detection request and processes it on the server side. The prototype system is designed with three main software components, including a pre-trained CNN model training module for classification purposes, a text data training module for attribute estimation models, and a server-side module. We experimented with a variety of food categories, each containing thousands of images, and through machine learning training to achieve higher classification accuracy.

3.Masked Face Recognition Using Convolutional Neural Network[3]

Recognition from faces is a popular and significant technology in recent years. Face alterations and the presence of different masks make it too much challenging. In the real-world, when a person is uncooperative with the systems such as in video surveillance then masking is further common scenarios. For these masks, current face recognition performance degrades. An abundant number of researches work has been performed for recognizing faces under different conditions like changing pose or illumination, degraded images, etc. Still, difficulties created by masks are usually disregarded. The primary concern to this work is about facial masks, and especially to enhance the recognition accuracy of different masked faces. A feasible approach has been proposed that consists of first detecting the facial regions. The occluded face detection problem has been approached using Multi-Task Cascaded Convolutional Neural Network (MTCNN). Then facial features extraction is performed using the Google FaceNet embedding model. And finally, the classification task has been performed by Support Vector Machine (SVM). Experiments signify that this mentioned approach gives a remarkable performance on masked face recognition. Besides, its performance has been also evaluated within excessive facial masks and found attractive outcomes. Finally, a correlative study also made here for a better understanding.

4.Automated Helmet Detection for Multiple Motorcycle Riders using CNN[4]

Automated detection of traffic rule violators is an essential component of any smart traffic system. In a country like India with high density of population in all big cities, motorcycle is one of the main modes of

transport. It is observed that most of the motorcyclists avoid the use of helmet within the city or even in highways. Use of helmet can reduce the risk of head and severe brain injury of the motorcyclists in most of the motorcycle accident cases. Today violation of most of the traffic and safety rules are detected by analysing the traffic videos captured by surveillance camera. This paper proposes a framework for detection of single or multiple riders travel on a motorcycle without wearing helmets. In the proposed approach, at first stage, motorcycle riders are detected using YOLOv3 model which is an incremental version of YOLO model, the state-of-the-art method for object detection. In the second stage,a Convolutional Neural Network (CNN) based architecture has been proposed for helmet detection of motorcycle riders. The proposed model is evaluated on traffic videos and the obtained results are promising in comparison with other CNN based approaches.

5. Gesture recognition using Integral Imaging[5]

In this paper we present some results summarizing our previous work during the last few years in the analysis of Synthetic Aperture Integral Imaging (SAII) performance for human gesture recognition, with a particular emphasis in its accuracy in the presence of occlusions, also when comparing it with sensors and strategies that are commonly used for these tasks, like the RGB-D sensors. Experiments show that its performance compares favorably against monocular imaging and to RGB-D sensors and that gesture classification results support the increase in performance for SAII particularly under partial occlusions.

1.4 Problem Definition

The analysis of food images is considered a challenging task due to many factors including the identification of multi food classes within single plate or the variance of the food texture for the same type.
Additionally, the number of overall food classes is not determined yet.

1.5 Scope

1.Based on image analysis, a conventional automatic vision based dietary assessment system involves food classification.

2.Recently, the development in image processing and object detection, machine learning approaches, and specifically deep learning and its implementation of convolutional neural networks (CNN), has enhanced the image identification and recognition accuracy.

3.The project model is expected to give an accuracy of 80%.The model contain s 3 layers of CNN

1.6 Technology stack

1.**Python Keras for the CNN model**: **Keras** is a powerful and easy-to-use free open source Python library for developing and evaluating deep learning models**.**It wraps the efficient numerical computation libraries Theano and TensorFlow and allows you to define and train neural network models in just a few lines of code.

2. **Visual Studio Code**: Visual Studio Code is a free source-code editor made by Microsoft for Windows, Linux and macOS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git.

3.**Flutter**: Flutter is an open-source UI software development kit created by Google. It is used to develop cross platform applications for Android, iOS, Linux, Mac, Windows, Google Fuchsia, and the web from a single codebase. Here, it is used to develop an app for frontend GUI for easy interaction with the user.

1.7 Benefits for environment & Society

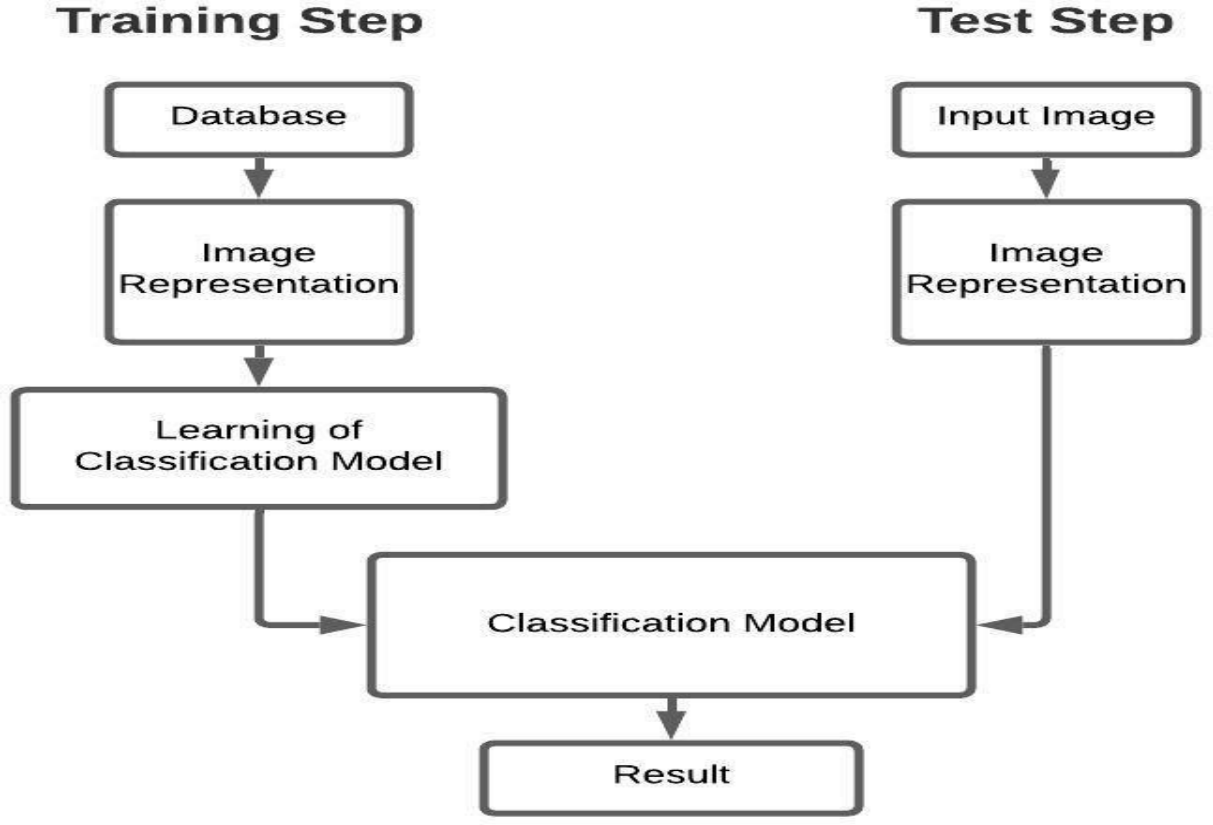
1. Food classification system can help social media platform to identify food.
2. Food classification system can enable an opportunity for social media platform to offer advertisement service for restaurants and beverage companies to their targeted users.
3. Food classification can be used in diet & supplementary applications and websites for the assessment of nutritional information of the food.
4. Food classification is an important task to help human beings record the daily diets.
5. Images of food are one of the most important information to reflect the characteristics of food.

2. Project Design

2.1 Proposed System

We propose a new structure of an enhanced food detection model using CNN.
After a food image is acquired, the first challenge is to detect if a food item exists in the image contents.
A deep CNN model will be developed to identify the different food images from the dataset.
The images will be identified based on various parameters like shape, size, color and texture.
The Food-101 dataset is used for training the model.

2.2 Design(Flow Of Modules)



Training Step :

1. The model will select an input image from the database (the loaded data set).
2. The image will be pre-processed to an image required for the model.
3. The Image classification model will then learn from the set of images to classify between food and non-food images.
4. The classification model will then give an accuracy (the number of images which are correctly identified) for the given data set.
5. If the accuracy is less than the desired, the model would be trained again by changing the features.

Testing Step :

1. In this step, the input image provided will be unknown to the model.
2. The image will be pre-processed to an image required for the model.
3. The trained model will now classify whether the input image is food or non-food.
4. If the actual result and the result predicted by the model is same, then the model is successfully tested and can classify images.

Dataset:

Non Food

Food

2.3 Description Of Use Case

The application will help the user to classify whether the image is a food image or a non-food image and determine the accuracy(confidence of the model).

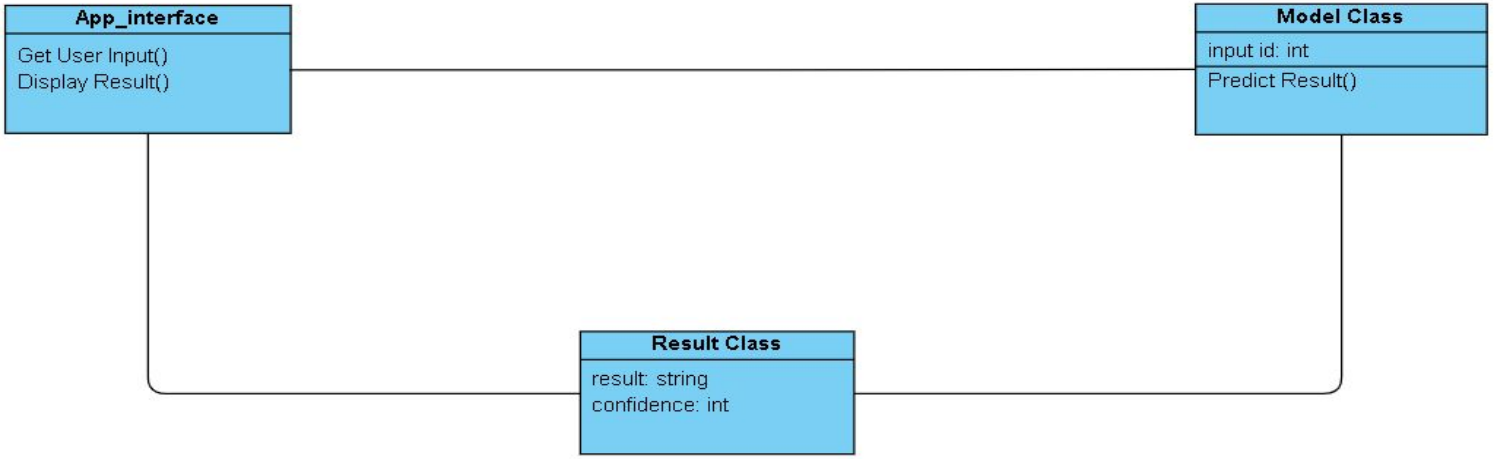
The user can upload images on the application using two ways :

1. Camera : Using real time images
2. Gallery : Using images previously stored in the phone

The application will confirm with the user the image to be uploaded by giving two options : retry and ok.

Once the image is uploaded successfully, the application will predict the image and the accuracy.

2.4 Class Diagram



2.5 Module-1

- The Classification model consists of two models namely: Food Detection and Deployment of this system on Flutter. It is worth noting the distinction between food recognition and food detection.
- Food detection refers to the problem of simply detecting food i.e. it is a binary classification problem. This is the first step of our project where food is detected among other items.

Module-2

- The second module of deployment of this system using Flutter.
- Using flutter, the food classification model is deployed on android Application.
- This application is developed to make using this system easier for the user.
- In this mobile app, on the home screen, there are two ways to provide image to the classification system:
  1. Using camera to capture a new image.
  2. Using an image which is already available in the mobile storage.

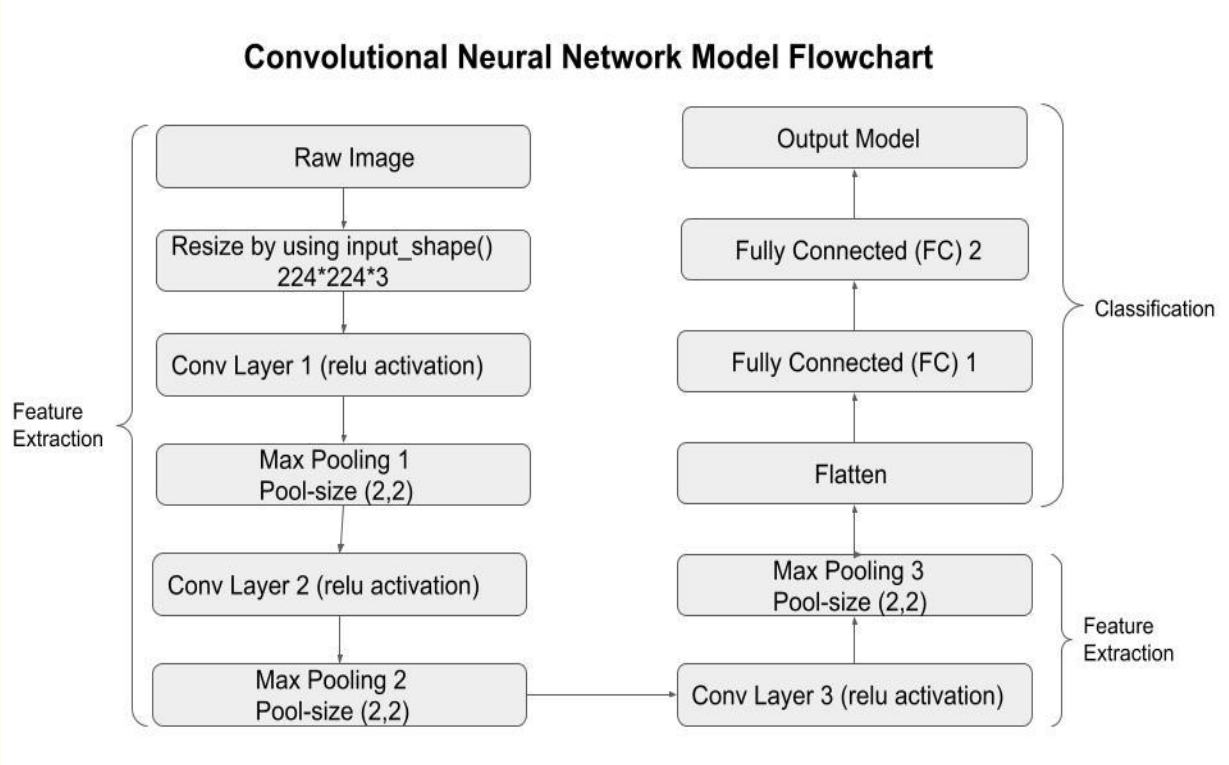
3.Implementation

3.1 Proposed System

In this project, we propose a new structure of an enhanced food detection model using CNN. After a food image is acquired, the challenge is to detect if a food item exists in the image contents. A deep CNN model will be developed to identify the different food images from the dataset. The images will be identified based on various parameters like shape, size, color and texture. The Food-101 dataset is used for training the model. Also, a frontend GUI is developed and deployed using Flutter, where an android app is developed which is capable of performing the specified module of the project.

3.1.1 Algorithms

Recently, the development in image processing and object detection, machine learning approaches, and specifically deep learning and its implementation of convolutional neural networks (CNN), has enhanced the image identification and recognition accuracy. We have created a CNN model to achieve the purpose of detection of food images. Following is the flowchart of the CNN used.



Above flowchart explains all the operations involved in the project i.e., Feature Extraction which includes 3 layers of CNN and the classification stage.

3.1.2 Pseudo code

```
program start
take input image from user
rescale input image
send input image to cnn model
function(cnn-operations)
        input image passed through layer 1
        pool-size(2,2)
                input image passed through layer 2
        pool-size(2,2)
        input image passed through layer 3
        pool-size(2,2)
get results from model
calculate accuracy from model
display results from model
display accuracy from model
program end
```

3.1.3 Platforms for execution

1.**Flutter**: Flutter is an open-source UI software development kit created by Google. It is used to develop cross platform applications for Android, iOS, Linux, Mac, Windows, Google Fuchsia, and the web from a single codebase. Here, it is used to develop an app for frontend GUI for easy interaction with the user.

2.**Visual Studio Code**: Visual Studio Code is a free source-code editor made by Microsoft for Windows, Linux and macOS. Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git. Here, code is executed and CNN model is embedded in the project. Also, it is used to build the gradle of the app using Android Studio which can be later used by the user.

4. Results

This system is successful in predicting whether the given image is a food image or non-food image. Also for every input image, it gives accuracy(confidence) of the model in predicting the image properly .Following are the comparison of some results.

Confidence or accuracy of images can vary from 10% to 99% depending on the image taken as input by the user.

Input Image	Label	Confidence
	Non-Food	99%
	Food	10%

	Food	99%

5.Conclusion

Food Detection system is used to detect and identify if given image is food image or not. This is the first module of the bigger system which will recognize the given food item by their names. Also, the app developed is user friendly as it allows the user to input the image using two ways: take image from phone camera and to take image from phone gallery. In this way, this app is successful in predicting whether the image is food image or non-food image

Future Scope

1. We plan to develop a food recognition training and testing model on the images that are identified as food based on features like color and look of the dish.
2. We also plan to recognize calories of the food which is recognized using our food recognition system.

6. References

- Y. Kawano and K. Yanai, “Food image recognition with deep convolutional features,”[1] in Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, 2014, pp. 589–593
- H. Kagaya and K. Aizawa, “Highly accurate food/non-food image classification based on a deep convolutional neural network,”[2] in International Conference on Image Analysis and Processing, 2015, pp. 350–357.
- https://www.kaggle.com/theimgclist/multiclass-food-classification-using-tensorflow
- http://images.nvidia.com/content/APAC/events/ai-conference/resource/ai-for-research/FoodAI-Food-Image-Recognition-with-Deep-Learning.pd f
- K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” [3] ArXiv Prepr. ArXiv14091556,2014.

7.Bibliography

- Y. Kawano and K. Yanai, “Food image recognition with deep convolutional features,”[1] in Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, 2014, pp. 589–593
- H. Kagaya and K. Aizawa, “Highly accurate food/non-food image classification based on a deep convolutional neural network,”[2] in International Conference on Image Analysis and Processing, 2015, pp. 350–357.
- K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” [3] ArXiv Prepr. ArXiv14091556,2014.

Thank You