A Project on Automated Food Classifier

A PROJECT SUBMITTED TO THE

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

OF

BANGLADESH UNIVERSITY OF BUSINESS AND TECHNOLOGY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF BACHELOR OF SCIENCE

BY

Jannatul Ferdows Modina (16172103078)

Farisa Benta Safir(16172103049)

Zabir Mohammad(16172103203)

Project Title	An Automated Food Classifier
By	Jannatul Ferdows Modina (16172103078)
	Farisa Benta Safir (16172103049)
	Zabir Mohammad (16172103203)
Field of Study	Keras, Tensorflow, mobileNet version2
Project Advisor	Ariful Mailk (Assistant lecturer)
Academic Years	2020

ABSTRACT

This is a simple food classifier project. It is based on supervised learning. For developing the project data for different 10 class food were feed into the model. This project can simply classify 10 different food classes.

ACKNOWLEDGEMENTS

We like to say our gratitude to our creator ALLAH to let us into the world and our parents, who supported us in this whole study and always prayed for our success and good health. We would like to express our sincere gratitude and profound indebtedness to our supervisor Md. Ariful Islam Malik for his guidance, valuable suggestions, commendable support and endless patience towards the completion of this project. We feel very proud to work with him. Without the inspiring enthusiasm and encouragement of our supervisor, this work could not have been completed. We thank all the staffs and graduate students at Bangladesh University of Business and Technology (BUBT) and all the friends for their support and encouragement. We would also like to extend our elder and younger brothers. Finally, I wish to express my gratitude to Bangladesh University of Business and Technology (BUBT) for providing an excellent environment for research and all the other facilities to complete the project successfully.

Jannatun Ferdows Modina

Farisa Bente Safir

Zabir Mohammad

1.Introduction

In this neoteric era of computers, computer vision is an interdisciplinary scientific field.its deal with how computers can gain high-level understanding from digital images or videos. Computer vision is a field of artificial intelligence that trains computers to interpret and understand the visual world. Using digital images from cameras and videos and deep learning models, machines can accurately identify and classify objects and then react to what they "see". An image classification comes under the computer vision project category, which is a fascinating deep learning project.

An image classifier is a subset of pattern recognition in computer vision. It is an approach of classification based on contextual information in images. "Contextual" means this approach is focusing on the relationship of the nearby pixels, which is also called neighbourhood. The goal of this approach is to classify the images by using the contextual information. Some of its applications include systems for factory automation, face recognition, booth monitoring, and security surveillance. Image recognition is embedded in technologies that enable students with learning disabilities to receive the education they need — in a form they can perceive.

During the past decade image classification has attracted significant attention in the computer vision and pattern recognition disciplines. They are widely applied in computer interfaces to identify the feature of the digital image or video. Therefore, in this Neural Network project, we create an automated food classifier that can identify or classify the food items from a digital image. To create our automated food classifier, we use one of the famous machine learning algorithms out there which is used for Image Classification i.e. Convolutional Neural Network (CNN). We also use keras and tensorflow as backends for the Convolutional Neural Network. A convolutional neural network (CNN) is a specific type of artificial neural network that uses perceptrons, a machine learning unit algorithm, for supervised learning, to analyze data.

Convolutional networks perceive images as volumes, i.e., three-dimensional objects, rather than flat canvases to be measured only by width and height. That's because digital color images have a red-blue-green (RGB) encoding, mixing those three colors to produce the color spectrum humans perceive. A convolutional network ingests such im-

ages as three separate strata of color stacked one on top of the other.So a convolutional network receives a normal color image as a rectangular box whose width and height are measured by the number of pixels along those dimensions, and whose depth is three layers deep, one for each letter in RGB.

The implementation of automated food classifier using CNN, Keras, and Tensorflow backend that rescales the image applies shear in some range, zooms the image, and does horizontal flipping with the image. This ImageDataGenerator includes all possible orientation of the image. Beside, we implement our project in google colab. So the dataset has to be in the google drive and the directory should be set to access those dataset that we build on food items. In the dataset, there are 1000 images of food divided into 10 target classes, with each category containing 100 images. Each category has 2 partitions i.e train and test. Test data contains 20 images and the remaining 80 images are used for training the neural network model.

In this project, we use a sequential model to construct a neural network which is appropriate for a plain stack of layers where each layer has exactly one input tensor and one output tensor. Beside we use Mobilenet version 2 as a neural network architecture that runs very efficiently on mobile devices.Mobilenet version 2 uses depthwise separable convolution as efficient building blocks.

Stages of Image Classification

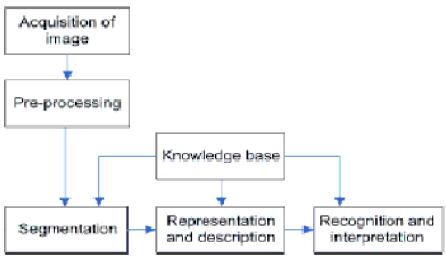

Figure 1: Stages of Image Classifier

2.Existing Work

In this section, a brief introduction about previous and related works of digital image classifier that we have surveyed during this project. Image classification was introduced during the early 1960s. Many techniques of digital image processing were developed in ,1960s at the Jet Propulsion Laboratory, Massachusetts Institute of Technology, Bell Laboratories, University of Maryland. The cost of processing the image was very high. But that changed in the 1970s, when digital image processing grew rapidly as cheaper computers and hardware became available. Images then could be processed in real time. With the development of fast computers available in the 200s,digital image processing has become one of the most common forms of image processing. It is used because it is not only the most versatile method, but also the cheapest.
There are dozens of image classifier project on internet. We will discussed only those similar to our automated food classifier.

2.1 Image classification with python keras on Cifar-10 data base:

In this project, they build a convolution neural network in Keras with python on a CIFAR-10 dataset. First, they explored cifar-10 dataset, and then they train the neural network using python and Keras.CIFAR-10 is a very popular computer vision dataset. This dataset is well studied in many types of deep learning research for object recognition.
This dataset consists of 60,000 images divided into 10 target classes, with each category containing 6000 images of shape 32*32. This dataset contains images of low resolution (32*32), which allows researchers to try new algorithms

2.2 Image classifier for identifying cat-vs-dogs using TFLearn in Python

In this project they created a Image Classifier of their own which can distinguish whether a given pic is of a dog or cat or something else depending upon your fed data. To achieve our goal, we will use one of the famous machine learning algorithms

out there which is used for Image Classification i.e. Convolutional Neural Network(or CNN). For the dataset they used the kaggle dataset of cat-vs-dog.

2.3 An Image Classifier Using Qt, OpenCV and TensorFlow

In this post we're going to learn how to create an image classifier application with a proper GUI that allows the users to choose a camera or a video file as the input and classify the incoming images (video or camera frames) in real time. We'll be using the power of Qt for cross-platform GUI creation and anything related to visualizing the output and allowing a seamless experience for the user. We'll also be using TensorFlow and OpenCV to handle the actual classification task, along with accessing cameras and video files.

3. Motivation

Computer vision is one of the most current and popular research topic present days.Computer vision is achieved with convolutional neural networks that can use images and video to perform segmentation, classification and detection for many applications. And We are very keen to goes in this topic for our further study. Image classification is a subset of pattern recognition in computer vision. Image classifier of 10 different class is the basic to go under research in computer vision.

Beside Image processing has its wide applications in robotics, machine learning, neural networking, signal processing, medical field, graphics and animations and in many other fields.

4.Proposed System

In our proposed system, we implement an efficient and highly accurate food classification and recognition method for 10 food items. Several research approaches have proposed automated image classification that include pattern recognition, machine learning, and deep learning. Most of the techniques are single food items. Hence, multi-label classification is used in this work. Multi-label classifications can output several classes at the same time, from the same input image.

5.Dataset

10 publicly available food datasets are used during the experiments. The datasets were gathered manually. The dataset contains leaf images of 10 different food items. Additionally, the number of images used in each class is 100. From them 80 images are used as training dataset, and rest of the 20 are in test dataset .The food items are Burger, Pizza, Ramen, Momes, French fries, Steak, Fuska, Coffee, Pasta, Sandwich.

6.Methodology

Different CNN architectures are implemented and bench-marked to perform multilabel classification. In the following sub-sections, the general process and layers of various CNN architectures are briefed.

6.1 Image pre-processing

As CNN architectures require input images to be of the same shape, each image data is reshaped into 224 by 224 pixels. Data normalization ensures that each input parameter has an analogous data distribution and results in faster convergence of the CNN.

6.2 Baseline architecture

The concept of computer vision is to enable machines to understand the world as humans do. CNN is a deep learning technique that takes input, adds weights and biases, and classifies images. Efforts have been made to explain the methodologies of CNN architectures . In this section, we focus on implementing the network architecture for food classifier. the CNN architecture,with the input layer (the raw image), convolutional layers, dense layer, and an output layer.

6.3 Input Layer

The inputs of the CNN architecture is the raw leaf images of different plants. Leaf images with different widths and heights are resized into the shape of 224×224×3 before being given to the CNN architecture.

6.4 Convolutional layers

Multiple convolution layers follow the input layer of the model. It is a mathematical operation that takes two inputs, an image matrix and a filter (or kernel), and conserves the relationship between pixels by learning image features. In every layer, there is a bank of m1 kernels. Each layer identifies particular features at every location on the input.A fully connected layer would contain at least the same number of trainable parameters the same as the number of pixels in an input image. Such a large number of parameters often result in overfitting. On the contrary, convolution does not increase the number of parameters. Instead, it only searches for a specific feature matrix K hat is learned through backpropagation. Recent deep neural network (DNN) architectures limit the height and width of the matrix K to a maximum value of 5. Hence the trainable parameters are balanced, and therefore convolution is used in the first layer of the architecture.

6.5 Batch Normalization

Batch normalization changes inter-layer outputs into a standard format. Batch normalization re-calibrates each of the data values based on the mean and variance for a specific data batch. Batch normalization increases the stability of DNN architectures and often leads to faster convergence.

6.6 ReLU6

ReLU function is used in every convolution for a simple calculation that returns the value provided as input directly. The function returns zero if it receives any nonpositive input, but for any positive value of x it returns that value.

7.Future Work

Our project is a simple image classifier of 10 different classes of food items. Beside for the short time during this pandemic situation, we can not complete the data set that we proposed in our system. For this reason the system gives us less accurate predictions. So in future, we are going to complete the dataset and add more classes of food to our system. Besides, we are planning to develop the system that automatically tells the the fat and cholesterol contain in that food.

8.Conclusion

In this project, we implements and tests a classifier based on different image classifier baselines. We practiced a transfer learning scheme to train and test our approach precisely. Further, we evaluate the architecture on a dataset consisting of 10different food items. We observe that spatial convolution, skip connections, and shorter hidden layer connectivity can massively improve the performance of image classification. As no dataset is available for multi-label food classification, we will consider adding more diverse foods to evaluate our methods in future work.