2020 IEEE REGION 10 CONFERENCE (TENCON)
Osaka, Japan, November 16-19, 2020

Personalised Food Classifier and Nutrition Interpreter Multimedia Tool Using Deep Learning

M. Sundarramurthi
Department of Electronics and Telecommunication
Dayananda Sagar College of Engineering (DSCE)
Bangalore, India
sundar_m@icloud.com

Nihar. M
Department of Electronics & Communication
BMS Institute of Technology and Management (BMSIT & M)
Bangalore, India
niharraju4@gmail.com

Anandi Giridharan
Principal Research Scientist
Department of Electrical Communication Engineering
Indian Institute of Science (IISc)
Bangalore, India
anandi@iisc.ac.in

Abstract— **Food plays a vital role in our day-to-day life to get all the required nutrients for a healthy lifestyle. In recent years, obesity has become one of the major concerns among humans. Therefore, it is necessary for each individual to keep track of the nutrition intake in order to have a balanced diet. This has scaled up the implementation of automatic food analysis and semantic food detection using different image classification approaches, among which Deep Learning has brought a series of breakthroughs in this field. We have proposed the Food Classifier and Nutrition Interpreter (FCNI), a user-friendly tool that classifies various food types with a different graphical representation of food nutrients values in terms of calorie estimation along with a multimedia audio response. FCNI improves state-of-the-art food detection by a considerable margin on achieving about 96.81% accuracy.**

Keywords— Convolutional Neural Networks (CNN), Deep Learning, Image Classification, Food-101 Dataset.

I. INTRODUCTION

In today's hectic schedule, diet tracking must be one of the primary practices to have a healthy lifestyle. As machines are increasingly capable of accomplishing various applications using Artificial Intelligence, we can also manage our health and diet in a much simpler way by using AI. Hence AI and food are the important fragments that we come across in our daily life to track the diet. The food we consume on a daily basis will have proteins, fats, carbohydrates and vitamins but a balanced diet is very essential. By keeping this as a primary goal, we developed a tool that performs proper identification of food with specific segmentation of various values of nutrients through which the user can more rely on different foods.

As a part of AI, Image recognition is found to be progressing towards the visual revolution in various applications such as Facial recognition, visual search engines, etc. [1]. In general, various challenges like image occlusion, viewpoint variation, background clutter, scale variation, intra-class variation, illumination conditions, etc., are part of the image classification [2]. The different stages involved in the image classification are preparation of labeled dataset, feature selection, construction of the classification model and model validation using test dataset [3]. Based on the proposed model requirement, an accumulated dataset with the labeled images of various classes are selected for the training purpose. These images are used to train the image classification model and make it learn, how the different classes look like. By observing different emerging patterns of the image, the model can recognize different visual elements of the image and classify them accordingly. Once the model is trained, it is exposed to a new set of images and run the classifier to obtain its quality by making it to predict the labels of the new images [3]. By comparing the predicted labels with verification labels, we can verify whether the classifier has properly classified the image or not. The proposed application FCNI tool is mainly emphasized on real-time food and nutrient analysis with high accurate recognition of food by using the concept of image recognition. There are various approaches for image classification, some of them are explained below.

A. *Machine Learning Approach*

The process of different feature extraction in this category is implemented manually by inspecting the visual features found in the food images, such as colour, texture, and shape [4]. The features are further used to train the prediction model based on existing algorithms such as Support Vector Machines (SVM), Bag of Features (BoF), K-Nearest Neighbors (KNN), Random Forests (RF) and Multiple Kernel Learning (MKL). The traditional classification methods execute three following tasks: Segmentation, Feature Extraction, and Classification. Segmentation is an essential step in identifying different regions of an image and then extracting the locations of the objects [4].

B. *Deep Learning Approach*

Under AI, Deep Learning is a subfield of machine learning. The way a human brain learns from different experiences, a Deep Learning algorithm focuses on repeatedly performing a task and tweaking its model to improve the outcome. There has been an exponential increase in data generation over the few years (more than 2.6 quintillion bytes). Since these algorithms require tons of data, this enormous increase in data creation is the main reason that deep learning is growing rapidly. Deep learning neural network can solve various complex problems by using a diverse, interconnected and unstructured dataset. We come across a lot of practical examples like Virtual Assistants (Siri, Alexa, etc.), Autonomous cars, tumor diagnoses, service chatbots, etc. in which deep learning is implemented.

This approach can be used to train a more effective neural network [4]. For the final classification, various deep learning algorithms adapt the feature extraction inevitably through a series of connected layers followed by a fully connected layer. Compared to many other traditional methods, deep learning has an outstanding classification ability and has recently become popular owing to its marginally exceptional performance with enhanced processing abilities [4].

Convolutional Neural Network (CNN) is one of the most prominent techniques used in deep learning. CNN's are widely preferred in computer vision applications owing to have its level of exceptional ability to learn operations on variant fields, on visual data and obtain high accuracies in challenging tasks with large-scale image data [4]. CNN used in deep learning makes the model to learn from the fed dataset and performs an operation that tends to be used in the classification of tasks directly from variant images, video, text, or sound. By studying the following advantages, we have used this Deep Learning approach along with a CNN model that has been integrated with the FCNI Tool which gives it the ability to classify different food classes efficiently.

II. PROPOSED ARCHITECTURE

The main objective of this project is to design and develop a GUI based interactive tool, which is capable of identifying the type of food with good efficiency. This is achieved by the deep learning approach, where the image classification model is trained by using supervised learning, in order to obtain the highest number of correct identifications. The architecture of the proposed FCNI tool is given in Fig. 1.

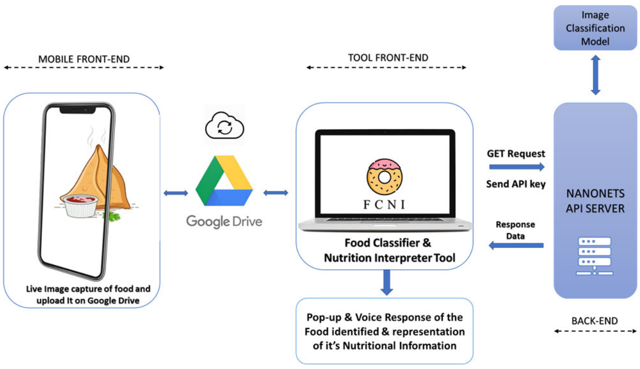

Fig. 1. Proposed FCNI Architecture

A. *FCNI Model Training on Nanonets API Server (Back-End)*

Nanonets is chosen as the platform to train and build our image classification model using the deep learning approach. Nanonets exclusively use deep neural networks and in the backend, their stack uses Mysql, Python, Caffe, Cassandra, TensorFlow, Theano, Docker, Keras, Torch and Elasticsearch. All these are hosted on the Amazon Web Services (AWS). Nanonets have language bindings in Shell, Ruby, Golang, Java, C# and Python which indeed makes it feasible to code in any language. Nanonets gives the flexibility to train models of different weighted data, to test and to validate the results accordingly. An API is created once the FCNI model is built and trained on the nanonets.

Food101 Dataset is used for training the proposed model, out of which, five classes of food, viz., Donuts, French Fries, Samosa, Chocolate Cake and Pizza, were chosen for the proof of concept. Further, it can be extended to any number of food classes based on the requirements [5]. The FCNI model was trained on the Nanonets API server with 1000 images of each class. The dataset is to be convoluted to find the labels of the given input. The training time depends on the number of images provided for each class. Thus, the FCNI model having an accuracy of 96.81%, was made into an API, which was used to integrate with python for the FCNI tool front-end development. The confusion matrix obtained after the training of the FCNI model is shown in Fig. 2.

PREDICTIONS / TRUE	Donuts	French_Fries	Samosa	chocolate_cake	pizza
Donuts	101 / 106	0 / 106	2 / 106	2 / 106	1 / 106
French_Fries	1 / 107	105 / 107	0 / 107	0 / 107	1 / 107
Samosa	1 / 107	0 / 107	105 / 107	1 / 107	0 / 107
chocolate_cake	2 / 108	0 / 108	3 / 108	103 / 108	0 / 108
pizza	1 / 106	0 / 106	1 / 106	1 / 106	103 / 106

Fig. 2. Confusion matrix obtained after FCNI model training

B. *Required Datasets and Software Libraries*

The FCNI tool front-end is designed and developed by using the python scripting language.

- **Dataset:** Food101
- **Software Libraries:** Requests library, PyQt5 & PyQt5.QtWidgets, gTTS (Google Text To Speech).

C. *Requests Library*

The requests library is a human-friendly HTTP library that can be used to send all kinds of HTTP requests to get information from various websites. It is very easy to use and has features like passing parameters in the URL and passing custom headers as well. Some of the features of request libraries which include Keep-alive and connection pooling, Elegant key/value cookies, Automatic decompression, Sessions with cookies persistence, Streaming Downloads, Basic/Digest Authentication, Chunked requests, etc.

The main reason to use python-requests is that the developer need not manually add query strings to URLs or form-encode post data. This library can be used for GET and POST requests as well. With an image of the food as input and by making use of the GET request for retrieving results of the classification from the API (nanonets server), which is stored on to a response object. The response data is converted into JSON type and this data is processed further for the prediction of food type.

D. *PyQt5 & PyQt5.QtWidgets*

PyQt5 is known for its ease and simplicity offered to a developer for the development of interactive Graphical User Interfaces (GUIs) [6][7]. PyQt5 is written in python but, it's a wrapper around the popular Qt GUI toolkit that major companies have been using it. It is a cross-platform GUI toolkit, unlike .NET which works on only Windows. But, PyQt5 is independent of the operating systems and hence it can be used on Mac, Linux or Windows. This can be used for building any kind of application or widget development using python. PyQt5 provides a tool (QtDesigner) for the front-end development of the GUI. Widgets are considered to be the building block of an application. PyQt5 provides a wide range and different types of widgets like checkboxes, list boxes, pushbuttons, etc. [7].

E. *gTTS (Google Text To Speech)*

gTTS is a python library and a command-line interface tool (CMI). This tool is interfaced with Google Translator's TTS (text-to-speech) API [8]. By using this library, it makes it easy to convert text to speech as it provides a variety of languages for audio manipulation. It has various sets of parameters to customize audio such as lang (Language, default: English), slow (Speed of spoken content), etc. This spoken content can be saved in terms of '.mp3' file format. To play sound files within the python we can use the playsound() module or run the file separately. Some of the features of the gTTS library include:

- Automatically all supported languages are retrieved.
- To support pronunciation corrections, it provides customizable text pre-processors.
- It supports customization of speech-specific sentence tokenizer for reading lengthy (or unlimited) texts, all while keeping proper intonation, decimals, abbreviations, and more.

F. *FCNI Tool GUI Development (Front-End)*

Graphical User Interface (GUI) plays a vital role in carrying out the smooth interaction between the user and the digital system. The GUI of the FCNI tool provides maximum user convenience and is self-explanatory for the end-user. This tool can be used by almost everybody, regardless of their technical or programming skills. The interactive design of the tool containing necessary input controls, navigation components and information components makes it an informative architecture. The key advantage of this tool is that it provides the user with immediate audio and visual feedback with the required information about the captured food. The GUI of the FCNI tool is shown in Fig. 3.

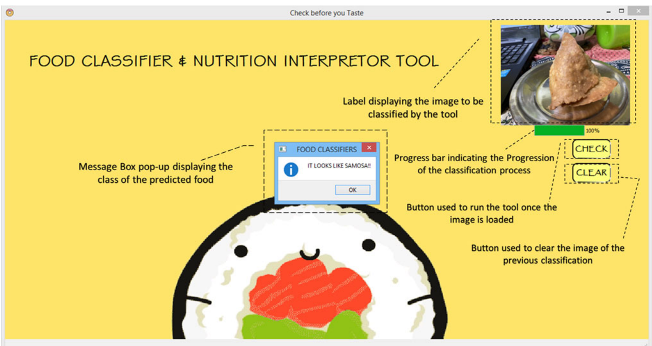

Fig. 3. GUI of FCNI Tool

III. CASE STUDY

FCNI tool plays a major role in diet control and some of the real-time case studies are discussed below.

A. *Case 1: Obese patient scans high-calorie food*

The rate of Obesity and Overweight is continuing to grow in adults and children. It is known to be one of the most significant health concerns among many people. Some changes in lifestyle like a food diet, physical activity, etc. would be effective and prevent chronic diseases. When the user scans high calorie or high fat containing food like French Fries, Cakes, Pizza, Ice-Cream, etc. FCNI gives an alert message saying that the predicted food is to be avoided by obese people. Further, it suggests other healthy food dishes in order to keep the person healthy and fit.

B. *Case 2: Diabetic Patient scans high sugar content food*

Diabetes is found to be one of the most common diseases around the world. Having a healthy diet can control, prevent and also reverse diabetes. FCNI tool also indicates diabetic people to avoid high sugar content dishes and suggest sugar-free food like Greens, Wheat Bread, Sugar-free candy, etc. to have a controlled and healthy diet.

C. *Case 3: Post-Surgery patient scans Fatty foods*

A proper diet plays an important role in a person's recovery process post-surgery. For these users, FCNI suggests few nutritious dishes which contain low fat and less spice.

IV. WORKFLOW OF THE FCNI TOOL

In this section, the different stages of the working process of the FCNI tool are discussed. To start with, the picture of the food is captured on a mobile device and it is directly uploaded to a specific directory on Google drive. The "Food Classifier & Nutrition Interpreter Tool" has been synced with Google drive over the cloud, hence it downloads the captured image and saves it as the input to the tool for the further classification process. When the user runs the tool by clicking on the 'CHECK' button, the input image downloaded is loaded on to the tool.

Once the image is loaded, the tool uses the request library to send the loaded image along with the API key to the Nanonets API server. This server containing the image classification model verifies the API key received. In case the API key is not authentic the response code is returned as 404. If the verification was successful with a valid API key, then the model processes the image, classifies the image and sends back the response code (200) along with a response data.

This response data contains the label followed by the probability of each class. The data obtained is of the type 'response' and cannot be processed directly. So, the response data has to be converted into a JSON file and then it is processed to abstract the class with the highest probability. Thus, we obtain a class of captured food. Based on the class, the tool displays a pop-up message followed by an audio response of the food identified. Following this, the tool also displays the complete nutritional information of the predicted food with a graphical and tabular representation. This kind of representation makes it easier for the user to get a quick view of the various nutritional contents and some facts about the predicted type of food.

The 'CLEAR' button erases and removes the previous input image entirely from the directories. The user can again take another picture from their mobile device and use the tool for the next classification.

V. FCNI TOOL INTEGRATION AND REAL-TIME TESTING

A. *Test Setup*

The FCNI tool is compatible with different versions of Windows operating system (Windows 7, 8 and 10). Using the 'FCNI TOOL v1.0 .exe' setup file, the user can easily install this tool on their systems by providing an authentic license key. Additionally, 'Google Backup and Sync' (Google Drive) is a pre-requisite software to be installed and it is included along with the setup file. After the installation procedure, the user can run the FCNI tool. The system should have an active internet connection and 'Google Backup and Sync' running, in order to experience the smooth working of the FCNI tool. Google Drive App should have been installed in the mobile device to integrate with the FCNI tool.

B. *Image Capture on Mobile Device*

By using the ‘Google Drive’ mobile application, the user can upload the image of the food to be classified. They can either capture a live picture from their camera or upload an existing image from their gallery. Once the picture is clicked/selected, this image will be uploaded to a specific directory. The screenshots are shown in Fig. 4.1 and Fig. 4.2.

Fig. 4.1

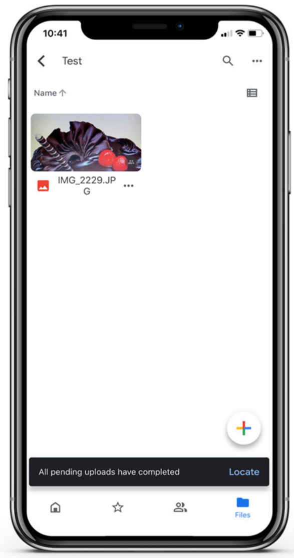

Fig. 4.2

C. *Test Results*

The FCNI Tool can classify various food dishes like Donuts, French Fries, Samosa, Chocolate Cake, Pizza, etc. The Tool was tested thoroughly with different real-time test cases of the above dishes in various scenarios and accurate results were obtained. One of the sample test results of Chocolate Cake is given in Fig. 5.

Fig. 5. Test Result of the Classified Food

D. *Advantages of FCNI Tool*

- The FCNI Tool provides the user with maximum information of the classified food through a multimedia response (Audio Response along with a Virtual and Graphical representation). Therefore, making nutrition analysis simpler.
- Suggests appropriate food dishes based on the user's health conditions.
- Accurate and Quick classification of the captured food.
- It is a Cost-efficient and a user-friendly tool which is suitable for all kinds of users.
- Easy Installation and compatible with various versions of the Windows Operating System.

VI. CONCLUSION

The rapid progressions in technology over recent years can be attributed to how Neural Networks has evolved and implemented in many fields and industries – automation, finance, healthcare, etc. In this proposed model we have addressed the effectiveness of Convolutional Neural Networks (CNNs) in Deep Learning for food image classification and detection. The accuracy of CNNs for these kinds of classification is quite remarkable and can be applied to various real-time applications. The current hot topic among the different levels of people is diet control, to keep our body fit. Hence obesity is influenced by our diet, physical activity and lifestyle. By keeping this in mind as a primary goal and to overcome this, we successfully designed, developed and tested a compact and cost-effective Food Classifier and Nutrition Interpreter Tool (FCNI) using a deep learning approach. An accuracy of 96.81% is achieved. This FCNI tool alleviates the process of food intake estimation and dietary assessment, hence beneficial for the user.

REFERENCES

[1] Narit Hnoohom and Sumeth Yuenyong, “Thai fast food image classification using deep learning”, 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-NCON2018), pp. 116-119, February, 2018.

[2] Murat Taskiran and Nihan Kahraman, “Comparison of CNN Tolerances to Intra Class Variety in Food Recognition”, 2019 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA), July, 2019.

[3] Rajayogi J R, Manjunath G and Shobha G, “Indian Food Image Classification with Transfer Learning”, 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), December, 2019.

[4] Mohammed Ahmed Subhi, Sawal Hamid Ali, Mohammed Abulameer Mohammed. "Vision-Based Approaches for Automatic Food Recognition and Dietary Assessment: A Survey", IEEE Access, Volume 7, 2019.

[5] David J. Attokaren, Ian G. Fernandes, A. Sriram, Y.V. Srinivasa Murthy, and Shashidhar G. Koolagudi, “Food Classification from Images Using Convolutional Neural Networks”, TENCON 2017 - 2017 IEEE Region 10 Conference, pp. 2801-2806, November, 2017.

[6] Iti Shrunkhla, Bramha Swaroop Tripathi and S.R.N. Reddy, “SmartTab: A Design & Implementation of Tablet for Learning Purposes based on PyQT framework”, 2nd IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), February, 2019.

[7] Y. CHABANE, A. HELLAL and AA. LADJICI, “PyPS: An Open Source Large Scale Power System Small Signal Analysis Software Package”, International Conference on Electrical Sciences and Technologies in Maghreb (CISTEM), October, 2018.

[8] Ravivanshikumar Sangpal, Tanvee Gawand, Sahil Vaykar, and Neha Madhavi, “JARVIS: An interpretation of AIML with integration of gTTS and Python”, 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), pp. 486-489, July, 2019.