

# Food Classification Using Deep Learning

Sridevi G M<sup>1</sup>, Raksha Raj K<sup>2</sup>, Roshini B M<sup>3</sup>

<sup>1</sup> Assistant Professor, Information Science & Engineering, SJB Institute of Technology, Karnataka, India

<sup>2</sup> Student, Information Science & Engineering, SJB Institute of Technology, Karnataka, India

<sup>3</sup> Student, Information Science & Engineering, SJB Institute of Technology, Karnataka, India

## ABSTRACT

Image classification has become less complicated with deep learning and availability of larger datasets and computational assets. The Convolution neural network is the most popular and extensively used image classification technique in the latest days. Image classification is performed on diverse food dataset using various transfer learning techniques. The food plays a vital role in human's life as it provides us different nutrients and consequently it is necessary for every individual to maintain a watch on their eating habits. Therefore, food classification is a quintessential thing for a healthier lifestyle. Unlike the traditional methods of building a model from the scratch, pre trained models are used in this project which saves the computation time and cost and also has given better results. The food dataset of many classes with many images in each class is used for training and validating. Using these pre-trained models, the given food will be recognized, and the nutrient content will be predicted based on the colour in the image.

**Keyword** - Calorie estimation, Convolutional Neural Network (CNN), Deep Learning technique, Food Image classification

## 1. INTRODUCTION

Image classification has become easier by research in Computer vision and machine learning. There is availability of huge data and computational resources. Techniques that can be used for Image classification are KNN classifier on local and global features used in [9], artificial neural networks, SVM and Random Forest technique to classify different classes using different set of features. However, these methods fail when the dataset is huge. Because the convolutional neural network can be easily handled with large amount of data and still provide high classification accuracy, it has gained attention in Image classification recently. Therefore, food classification is a quintessential thing for a healthier lifestyle. In fact, the numbers of people across the world that are suffering from obesity are more than 10%. The rate of obesity is increasing in an alarming rate from the past few years. In this growing digital world, it is very important to keep a track of calorie intake in our food. As the world is growing, problems like obesity and weight gains are also equally growing. It becomes inevitable to maintain a track of food intake. It becomes difficult to keep everything on track in a diary.

Statistics show that 95% of the people no longer follow any dietary plan as these restrict the people from consuming their day-to-day food. So, the primary cause for obesity is imbalance of the amount of food intake and energy consumed by the individual, and a healthy meal is necessary. Thus, maintaining a healthy diet is an important goal for many people. The process of tracking the number of calories consumed can be very tedious as it requires the user to keep a food journal and perform little messy calculations to estimate the number of calories consumed in every food item. Through this research we try to classify Indian food images into their respective classes. The proposed software model uses machine learning as the base which recognizes the food image that is uploaded as an input by the user, processes the food image, recognizes it, and estimates the calories from the predicted image. People record, upload, and share food images more willingly than ever on websites like Instagram, Facebook etc. So, it is more convenient to locate more data (images and videos) related to food. Consequently, supporting users in diet management and reducing the need for the manual paper approach.

Food image recognition and calorie estimation can aid in diet management, food blogging and recognizing the Indian foods. The organization of the paper is as follows: The related works are presented in section II. Section

III deals with the proposed methodology followed by the experimental results, conclusion and future work in section IV and section V respectively. To overcome the manual labour and the erroneous data, various applications were developed to calculate the food intake. One of the latest technological advancements to overcome difficulties in pictures of food items is a variety of e-health applications were developed to calculate calories in food that used the concept of image processing.

## 2. RELATED WORK

Introduction Research on all existing techniques and comparison was done for food recognition [1] and the results were recorded. There is much advancement in the food image recognition in the past few years.

The work proposed by Viswanath C, et al [4]., proposed a method to classify Indian food images by adopting a Google Inception-V3 based Convolutional neural networks (CNNs) model. Here they have used convolution layer that is able to create its own convolution kernel to convolve with input layer to generate the tensor outputs. The Max-pooling function is used for features extraction from the data and help to train the CNN model. The dataset contains data from some real time South-Indian food data where some of the training and testing images has some noise, different colour intensity and Images with the wrong-labels. The proposed system uses the custom Inception-V3 weights which are pre-trained using ImageNet and it considers the images after reshaping to the size of 150x150x3 for all images. The average pooling function is performed on the food image dataset to take the average of image features and the dimensionality of space output is defined through the dense function.

ManpreetkourBasantsingh, et al [5]., proposed an algorithm for fruit recognition and its calorie prediction primarily based on shape, colour, and texture. The histogram of gradients and GLCM with neighborhood binary pattern algorithms for texture segmentation scheme acknowledges the fruits and the area, major axis and minor axis are calculated by the way of usage of shape characteristic to get more accurate calorie value. The features are fed to multi SVM classifier for greater accuracy using the dietary look up table. For the dataset, 5 categories of fruit images are captured using Samsung grand prime cellular telephone and the images obtained were 3264x1836 pixels in size. Pre-processing steps to be achieved are rgb to gray conversion, filtering, resizing to 256x256 and adaptive histogram equalization. The histogram of orientated gradients (HOG) is a characteristic descriptor used for the cause of object detection. For acquiring the correct features, appropriate segmentation scheme is used.

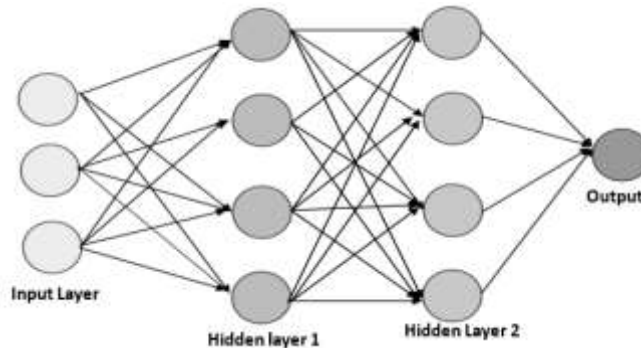
Hemraj Raikwar, et al [6] proposed a model which focused on estimation of number of calories in the food item by just taking its image as input using SVM. The proposed model applies some techniques of image processing followed by feature extraction. The authors designed the dataset, applied this dataset to some image processing techniques, then processed dataset is applied to the feature extraction process. The features extracted from all the images are then applied to the classifier support vector machine (SVM) which classifies the images in different classes as specified in the learning algorithm. The model consists of several intermediate activities which are: a. extracting the feature vector of image, b. identifies the food item in the image; c. predicts the calorie content of the food item in the image. The dataset includes images from PFID (Pittsburgh Fast Food Image Dataset) and calorie information from nutrition. For pre-processing, it includes background subtraction to remove noise and unnecessary information.

The Google Net [7] refers to the Inception architecture developed by Szegedy et al., is a deep convolution neural architecture that was codenamed as Inception. The objective of the inception module is to act like a multi-stage function extractor by using 1x1, 3x3, and 5x5 convolutions inside a single module of the network, then the result of this module is fed as input to the next layer within the network. It scored well with detection and classification in the ImageNet Large Scale Visual Recognition Challenge 2014 (ILSVRC14) and bagged the first place. It was implied by [7] for vision networks and covering the hypothesized outcome by dense, readily available components. With a little tuning of the module, modest profits were seen in comparison to the other reference networks. Inception V3, the latest version has been used to build a classifier in the paper.

Patanjali C, et al [9]., proposed an automatic food detection system that detects and recognizes varieties of Indian food. The proposed food recognition system is developed in such a way that it can classify the Indian food items based on two different classification models i.e., SVM and KNN. The proposed system uses a combined colour and shape features. A comparative study on the performance of both the classification models is performed. Parameters such as food density tables, colour, and shape acknowledgment as a part of image processing, and classification with the SVM and KNN have been considered. The data set contains the feature vector extracted from the sample images. They have considered around 200 image samples with cluttered food and individual food items. They have considered two combined features for these 200 samples and used 80% of the images as training set and 20% of them as the testing set.

### 3. PROPOSED SYSTEM

The proposed system consists of four phases: Image pre-processing, Feature selection and extraction, Classification and Output. Fig. 1 shows a diagram of all the layers of the CNN model that are interconnected



**Fig -1:** A block diagram stating the design of the CNN Model.

#### 3.1 Dataset

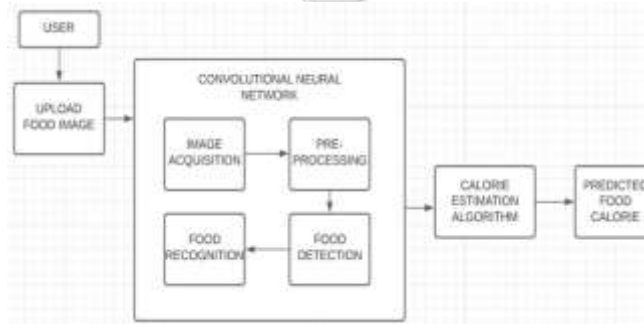
For the study, we have taken the Indian food dataset. It con-trains 12 different classes of food, and each class has around 70 sample images. The dataset inherently comes with a lot of noise since there are images in which there are unwanted things. These sample images also contain a lot of colours. The Fig. 2 shows the sample food images from the dataset which are taken as colour images



**Fig -2:** Sample food images used in the dataset

#### 3.2 Image pre-processing

The dataset contains 12 different classes of food images. Each class of image is divided into training and testing images wherein 80% of images from each class are considered as training and the remaining 20% samples as test samples. Altogether, there are 710 training samples and 142 testing samples. The training set images are fed to the CNN model and validation is made using the test dataset. Fig. 3 below shows the proposed methodology and design.



**Fig -3:** Architecture of the system

### 3.3 Convolutional Neural Network (CNN)

Convolutional neural networks (CNN), a class of deep neural network, are majorly used for the process of image recognition. CNN consists of some basic layers like hidden layers and fully connected layers where hidden layers are used to extract and learn the features of training images and fully connected layers are used for classification of the image. The structure of CNN is inspired by the structure of human body nervous system which consists of neurons. The way each neuron passes messages to the next neuron in the body, the layers in CNN also communicate with each other in the same way for the process of feature extraction.

As the name says CNN performs convolution on the input data using filters (or kernel) in convolutional layer. This is done to extract the features from the input. For the proposed system, we have used four layer which are convolution layer, relu layer, pooling layer, and fully connected layer. Each convolutional layer is followed by pooling layer which is used to reduce the dimension of the images while preserving the spatial invariance. Hence reduces the computation cost in the CNN network. In our architecture, the max-pooling is of the filter size 3. A total of 100 epochs has been taken to train the CNN model and which gets the most prominent features which is selected by taking the maximum value of the features from the prior layers. After gathering all the features and converting them into a matrix, fully connected (FC) layer is used to map features and classify the images into correct categories. Rectified Linear Units, also known as Relu, is accuracy layer and activation function used in Deep Learning models. It helps in speeding up the training process by activating necessary things and giving sparse outputs. Any negative elements in the computational step of Relu are set to 0.2 dropouts out of 0.8 to prevent over fitting. SoftMax function is most used in the final layer of neural network. It is equivalent to the logistic regression over the features extracted from the layer before the final fully connected layer. These are trained under the cross entropy. SoftMax function turns the logits (logical regression) into probabilities and all these probabilities sums up to 1. We have set the epoch value for 100 and each epoch is iterated 120 times as snapshot steps.

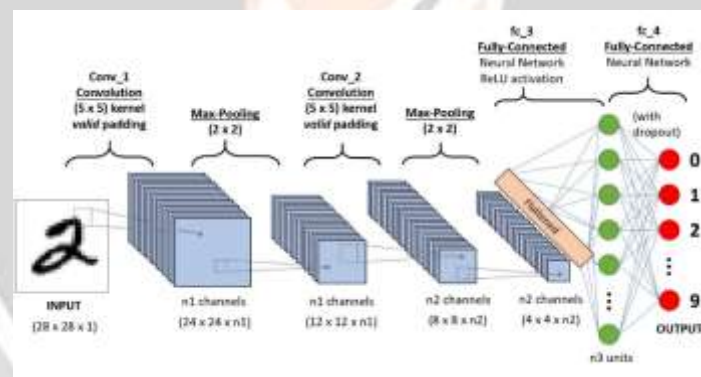


Fig -4: CNN Model

### 3.4 Testing

After the training of the model using the train dataset, validation is done using test dataset. Train data is the sample of data used to fit the model and test data is the sample of data used to provide an unbiased evaluation of the fit on train dataset while tuning of model hyper parameters. Then it is finally tested using the test dataset.

During the classification phase, difficulty arises while considering several different size and variety of dishes in the dataset. The CNN is a better approach to overcome the scaling problem due to its capability to capture the patterns associated with the images. In addition, it is also capable to deal with the noise that already exists in the images.

### 3.5 Calorie Extraction

Once the image of food is analyzed and the food is detected, our classifier can be used to estimate the calorific content of the classified food from the dataset. It also gives the accuracy value based on the detection.

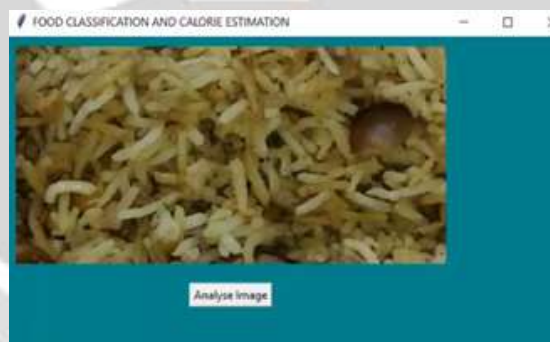
**4. CRESULT**

The result and analysis of our proposed model is given in this section, which shows the performance evaluation of food classification process. The data used for training our model consist of large number of images; The testing has been done in the standalone system under Anaconda prompt, the system is configured with 12 GB RAM, operating system Windows 10 and processor Intel core i5. Here, a total of 12 classes are considered which contain 855 numbers of images, for the testing purpose 20% of the total images are used (i.e., randomly selected from all classes) and 80% remaining image for training purposes. Our work considers 80% of images from dataset such as 710 images are selected for the training phase and got training accuracy of 84%, which means out of 710 images the model has fitted correctly at 596 images. However, 20% of images such as 145 images have kept for testing., miss-classification happened for some images whose colour are very identical; otherwise, it has given optimal results at the unseen dataset. Whereas our proposed model has been tested on self-collected food dataset and manages to get classification accuracy in percentage.



**Fig -5:** Food Images after labeling in the dataset.

After the user has selected the input of a food image, the sys-tem has been used for the analyzing and classifying the food item and it gives the following results as shown below.



**Fig -6(a):** GUI of the home page



**Fig -6(b):** GUI of the final output window

## 5. CONCLUSION AND FUTURE ENHANCEMENTS

In this research study, the Convolutional Neural Network, a Deep learning technique is used to classify the food images into their respective classes. As far as the future enhancement is concerned, the task of classification can be improved by removing noise from the dataset. The same research can be carried out on larger dataset with a greater number of classes and a greater number of images in each class, as larger dataset improves the accuracy by learning more features and reduces the loss rate. The weights of the model can be stored and used to design a web app or mobile app for image classification and in addition calories extraction of the classified food.

The proposed system would be improved when the calorie measurement will be done for multi-food and complex food items which would help people who will be using our application to apprehend deeply the complexities of food. Additionally, the expansion of dataset with more variety of food types which will improve the result and accuracy of the system. As fast response is one of the most crucial elements these days, the computational part of the model which takes longer time can be offloaded to the cloud. With the use of cloud, a model which takes less time to produce the result by performing large computations can be obtained.

The work started with collecting real-time food images from various sources. The convolution neural network-based model is trained over large amount of food images, which enhances your model capability to get required features quickly. In the result analysis, the accuracy of the training dataset of images obtained is about 84%. For the future However, a massive-scale dataset is needed to train the CNN and even a well-trained network cannot have a segmentation accuracy of 100%, which will en-large the error rate of recognition. Therefore, we can build a larger dataset which includes different food images to get a better result.

The need to have a system that measures daily food intake for healthy diet is crucial due to the insufficient knowledge of diet and calorie requirements. In addition, correct food recognition is considered challenge. Hence, we proposed a measurement method to estimate the number of calories from different food images by measuring the features such as colour of the food from the image. We have efficaciously implemented a robust system for the correct identification of the fruit.

## 6. ACKNOWLEDGMENT

The authors wish to sincerely thank Mrs. Sridevi G M, Mrs. Rekha and Mr. Mohan H S for the opportunity and support on behalf of the institution.

## 6. REFERENCES

- [1]. Rajayogi, J.R., Manjunath, G. and Shobha, G., 2019, December. Indian Food Image Classification with Transfer Learning. In 2019 4th Inter-national Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS) (Vol. 4, pp. 1-4). IEEE.
- [2]. Reddy, V.H., Kumari, S., Muralidharan, V., Gigoo, K. and Thakare, B.S., 2019, May. Food Recognition and Calorie Measurement using Image Processing and Convolutional Neural Network. In 2019 4th In-ternational Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT) (pp. 109-115). IEEE.
- [3]. Subhi, M.A. and Ali, S.M., 2018, December. A deep convolutional neural network for food detection and recognition. In 2018 IEEE-EMBS conference on biomedical engineering and sciences (IECBES) (pp. 284-287). IEEE.
- [4]. Burkapalli, V.C. and Patil, P.C., TRANSFER LEARNING: INCEPTION-V3 BASED CUSTOM CLASSIFICATION APPROACH FOR FOOD IMAGES.
- [5]. Fruit Recognition and its Calorie Measurement: An Image Processing Approach ManpreetkourBasantsingh Sardar1, Dr. Sayyad D. Ajj2 (2016)
- [6]. Raikwar, H., Jain, H. and Baghel, A., 2018. Calorie Estimation from Fast Food Images Using Support Vector Machine. International Journal on Future Revolution in Computer Science & Communication Engineering, 4(4), pp.98-102.
- [7]. Subhi, M.A. and Ali, S.M., 2018, December. A deep convolutional neural network for food detection and recognition. In 2018 IEEE-EMBS conference on biomedical engineering and sciences (IECBES) (pp. 284-287). IEEE.
- [8]. Zhang, W., Zhao, D., Gong, W., Li, Z., Lu, Q. and Yang, S., 2015, Au-gust. Food image recognition with convolutional neural networks. In 2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015

IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom) (pp. 690-693). IEEE.

[9] Pathanjali, C., Salis, V.E., Jalaja, G. and Latha, A., 2018. A Comparative Study of Indian Food Image Classification Using K-Nearest-Neighbour and Support-Vector-Machines. *J. Eng. Technol*, 7, pp.521-525.

[10] Christodoulidis, S., Anthimopoulos, M. and Mougiakakou, S., 2015, September. Food recognition for dietary assessment using deep con-volitional neural networks. In *International Conference on Image Analysis and Processing* (pp. 458-465). Springer, Cham

