INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 07 ISSUE: 05 | MAY - 2023 IMPACT FACTOR: 8.176 ISSN: 2582-3930

# Food Recognition and Calorie Estimation Using CNN Algorithm

Prakash Kumar.C[1],Harish.S[2],Ellakiya.R[3],Kamalesh.S[4], Santhosh Kumar.K[5]
Assistant Professor[1],Department of Computer Science Engineering,
UG Scholar[2,3,4,5], Department of Computer Science Engineering.
Tiruppur,TamilNadu, India.

***Abstract*— This paper proposes the relationship between nutritional ingredients identification in food and inspecting Calories through Machine Learning models to perform the data analysis. It identifies the Nutrition that we may get affected by the lack of certain nutritional ingredients in our body and recommend the food that can benefit the rehabilitation of those Weight Groups. To achieve high accuracy and low time complexity, the proposed system was implemented using CNN Machine Learning models. Image of food gets scanned by the CNN machine learning model and it separates the image layer by layer. Separated layers get compressed and converted into single layers and make analysis with the existing data. Real-time nutritional and calorie data get displayed on the webpage. The state of art accuracy of 92.04%**

**Keywords— CNN machine learning model, calorie data, nutritional data , image classification**

## I. INTRODUCTION

A healthier lifestyle requires a record of what a person is consuming. This was done manually via smartphones but was rigged with miscalculations and errors. To overcome manual labor and erroneous data, various applications were developed to calculate food intake. One of the latest technological advancements was that a picture of the food item was taken so as to keep a record.

Suitable nutritional diets play an important role in maintaining health and preventing the occurrence of Diseases. With the gradual recognition of this concept, India has also repositioned the impact of food on health. However, research on nutritional ingredients in food via Machine Learning, which is conducive to the rehabilitation of diseases is still rare in India. At present, India has just begun the IT (Information Technology) construction of smart health care.

In India, studying the relationship between nutritional ingredients and diseases using Machine Learning is immature. Most doctors only recommend specific food to patients suffering from Multiple Diseases, without giving any relevant nutrition information, especially about nutritional ingredients in food

## II . MACHINE LEARNING

Machine learning is a branch of artificial intelligence that involves teaching machines to learn from data without being explicitly programmed. It has the ability to analyze large amounts of data and identify patterns and insights that can be used to make predictions or decisions.
The concept of machine learning has been around for several decades, but recent advancements in technology and the availability of big data have led to significant breakthroughs in the field. Machine learning has found applications in various industries, such as healthcare, finance, and transportation, and has the potential to revolutionize the way we work and live.

There are several types of machine learning, including supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, machines are trained on labeled data to make predictions or decisions. In unsupervised learning, machines learn from unlabeled data and identify patterns and insights on their own. In reinforcement learning, machines learn through trial and error, receiving feedback on their actions to improve their performance.

The applications of machine learning are vast and diverse, ranging from image recognition and natural language processing to predictive analytics and autonomous vehicles. In healthcare, machine learning is being used to improve patient outcomes and reduce costs by predicting and preventing disease. In finance, it is being used to detect fraud and make better investment decisions.

However, the use of machine learning also raises important ethical and social questions, such as concerns about privacy and bias in algorithmic decision-making. As such, researchers and practitioners in the field of machine learning must be mindful of these issues and work to develop ethical and responsible practices.

Machine learning plays a crucial role in image processing and classification. The main idea behind using machine learning in image processing is to provide the computer with the ability to learn patterns and features in the images without the need for explicit programming. This is accomplished through the use of algorithms that can automatically learn from data, such as neural networks and decision trees.

Here's a brief overview of how machine learning works in image processing and classification:

1. Data preparation: The first step in using machine learning for image processing is to prepare the data. This involves collecting, cleaning, and organizing the images in a format that can be used by machine learning algorithms. This includes converting the images into a numerical representation, such as pixels or features, which can be fed into a machine learning model.

2. Training the model: Once the data is prepared, the machine learning model can be trained. This involves feeding the model a large dataset of labeled images, which are images that have already been categorized or classified. The model learns to recognize patterns and features in the images and associates them with the corresponding labels.

3. Testing and evaluation: After the model has been trained, it needs to be tested to evaluate its performance. This involves feeding it a set of images that it has never seen before and evaluating its accuracy in classifying them. The model is then adjusted and refined based on its performance.

4. Deployment: Once the model has been trained and tested, it can be deployed to classify new images. The model takes in an image and produces a classification based on what it has learned from the training data.

In image processing and classification, there are several machine learning techniques that can be used, such as convolutional neural networks (CNNs), support vector machines (SVMs), and decision trees. Each of these techniques has its strengths and weaknesses and is suited for different types of image processing and classification tasks.

## III. RELATED WORKS

B.Arslan, S.Memis,E.B.Sonmez and O.Z.Batur in the paper" Fine-Grained Food Classification Methods on the UEC FOOD-100 Database"[1] explains a model that can accurately classify different types of foods in a fine-grained manner. the major disadvantage of this method is a large amount of labeled training data, which may not always be available or may be expensive to obtain. Additionally, the proposed method may struggle with recognizing food images that are heavily occluded or partially visible, which is a common occurrence in real-world scenarios. Finally, the method may also be prone to bias and misclassification when dealing with foods that have similar visual features or cultural differences in preparation.

D. Zhao, " Food image recognition with CNN " [2] major disadvantage is that the model heavily relies on the availability of large-scale labeled datasets for training, which may not always be available or may be expensive to obtain. Additionally, the model may struggle with recognizing food images that are heavily occluded or partially visible, which is a common occurrence in real-world scenarios. Finally, the model may also be prone to bias and misclassification when dealing with foods that have similar visual features or cultural differences in preparation. The highest accuracy of 89.36%.

## IV. CONVOLUTIONAL NEURAL NETWORK ALOGRITHM.

CNN stands for Convolutional Neural Network, which is a type of deep learning algorithm that is commonly used in image processing applications such as object detection, classification, segmentation, and recognition.

A CNN consists of several layers, each of which performs a specific operation on the input data. The basic layers of a CNN include convolutional, pooling, and fully connected layers.

The convolutional layer is the most important layer of a CNN, as it performs the process of convolution on the input image. In convolution, a filter or kernel is applied to the input image to extract features such as edges, corners, and blobs. The output of the convolutional layer is a feature map that contains these extracted features.

The pooling layer is used to downsampled the feature map, reducing its size and complexity. The most common type of pooling is max pooling, where the maximum value in a particular region of the feature map is retained, and the other values are discarded.

The fully connected layer is the final layer of the CNN, which takes the output of the pooling layer and produces the final output of the network. This layer performs a classification or regression operation depending on the task at hand.

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 07 ISSUE: 05 | MAY - 2023 IMPACT FACTOR: 8.176 ISSN: 2582-3930

In image processing, a CNN is trained using labeled data to recognize specific features in images. During training, the network adjusts its parameters to minimize the difference between the predicted output and the actual output. This process is known as backpropagation, and it is used to update the weights and biases of the network.

Once the CNN is trained, it can be used to perform various image processing tasks, such as object detection, where it can identify the location of objects in an image, or image classification, where it can recognize the content of an image and classify it into one of several categories.

Overall, CNNs are an extremely powerful and widely used algorithm in image processing due to their ability to automatically learn and extract complex features from images without the need for human intervention.

## V. Proposed System

CNN has added a new dimension to machine learning. In terms of image detection, no feature can come close to CNN. That is why we have used convolutional neural network for object or image detection. CNN can take a set of images of an object, give the objects specific weights and biases and with these variables and can differentiate each object or image by following a set of steps in Figure

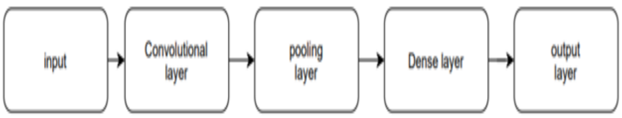


**Fig 1. BLOCK DIAGRAM OF CNN**

CNN mainly works in multiple layers where the first layer takes the shape of the image as we have mentioned earlier. The output of this first layer will work as an input for the second layer and the output of the second layer will work as an input for the third layer and so on. This process will continue to perform until the last and final layer. In this way, the first convolutional neural network is responsible for extracting low level features (edges, gradient orientation etc.) of the images and the deeper layers are responsible for extracting the high-level features from the images. For extracting high level features from complex images, CNN uses a fully connected network. After making the images suitable for multi-layer perceptron, CNN now flattens (Figure 2.2) the image into column vectors. These flattened images are put into a neural network which we call feed forward network and then we use back propagation for backtracking. For analyzing and finding errors, the process is repeated in every iteration while training.

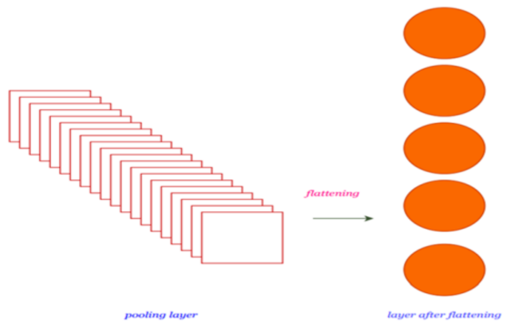


**Fig 2. FLATTING PROCESS**

Feed-forward neural networks are a kind of artificial neural network where the relation between units does not perform a repetitive cycle. Feed forward is one of the first one of its kind for image detection where it is much simpler than recurrent neural network. The name feed-forward is mainly because it only forwards the data to the next layer and since it has no loops it cannot be traversed backwards. Feed forward network mainly takes a function f on a default size of inputs i so that f(i)=j where we can say the training pairs (i, j) which we can see in (Figure 2.3).

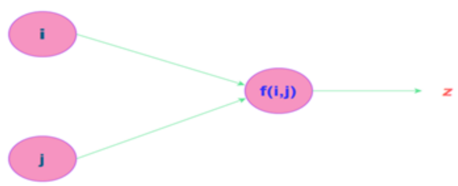


**Fig 3.** FEED FORWARDING NETWORK

After feed forwarding we need back-propagation to traverse backwards to the previous layer. Back propagation is easy and simple to perform. Back propagation mainly uses inputs and weights for calculating every neuron from the input layer and calculates the error from the previous layer. After calculating the error, Back propagation traverses back to the hidden layer or inner layer for adjusting the weight in order to decrease the error. Back propagation keeps repeating this process until the algorithm reaches the potential output. There are mainly two kinds of back propagation and they are Static back propagation and

VOLUME: 07 ISSUE: 05 | MAY - 2023 IMPACT FACTOR: 8.176 ISSN: 2582-3930

Recurrent back propagation Static back propagation mainly works on static inputs and static outputs, so it is mostly used for static image recognitions. On the other hand, recurrent back propagation is mainly used for non-static inputs and outputs. Back propagation takes full use of the chains of fully connected nodes and networks by traversing back to hidden layers and it can work with any number of nodes and inputs which is shown in Figure 2.4.

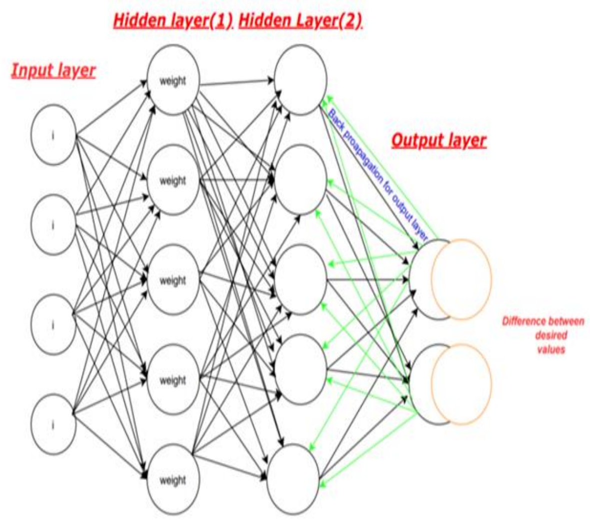


Fig 4. BACK PROPAGATION PROCESS

In this project we estimate the type of food in one single image by training a CNN network on known type for the food. This means that we will be able to predict the type of food by only looking at one single image. The calories can then be calculated by using the weight together with information from a food table as an image file.

**ADVANTAGES:** Our model overcomes the deficiencies by using Inception V3 which takes really less time for processing and it can also detect multiple foods from an image**.** Our CNN based model outperformed them by a big margin by getting an accuracy of 98.48% which was trained from the scratch. Additionally, our framework not only recognizes food items from a given picture but also predicts the calorie of the distinguished food items

## IV. THE DATABASES OF FOOD

There are only a few number of food image databases, which have different characteristics such as the number of food categories, the total number of images, the type of cuisine, i.e. Western (French, Italian, Turkish, ...), Asian (Japanese, Chinese, Thai, ...) or the Fast-Food (as it may be considered a world-wide type of food), quality and type of the images, i.e. single versus multi food images, and different contexts, i.e. the same dish may contain more than one food, or different comestibles are separated in several dishes on a tray.

**TABLE I**

DATABASES OF FOOD IMAGES. "S" STANDS FOR SINGLE-FOOD ITEM AND "M" FOR MULTI-FOOD ITEM IMAGE.

| DATASET | IMAGE CLASS | TYPE | YEAR |
|---|---|---|---|
| PFID | 61-1,098 | S | 2009 |
| FOOD-85 | 85-8500 | S | 2010 |
| UEC Food-100 | 100-12,740 | S&M | 2012 |
| FOOD-101 | 101-101,100 | S | 2014 |
| UNIMIB | 73-1027 | M | 2016 |

The diversity present in the available databases make the comparison a difficult task. This analysis focuses on publicly available databases with more than 60 classes and 1,000 images. The Pittsburgh fast-food image dataset (PFID) [32] is one of the first publicly available database with food pictures. Released in 2009, it collects 1,098 images belonging to 61 classes. The pictures were taken in several popular fast-food restaurants, where the same item was photographed from different angles; additional pictures were collected under the controlled environmental settings in the lab. All the images in the database contains only a single food item.

The Food-85 database has been introduced by Hoashi et al. [33] in 2010. The database stores 8,500 single food images belonging to 85 classes of Western and Asian cuisine.

The UEC Food-100 database [1] has been introduced by Matsuda et al. in 2012. It stores a total of 12,740 images belonging to 100 classes. Since this work challenges the classification experiment on this dataset, UEC Food-100.

In 2014, three interesting databases of food have been released: (1) the Food-101 [34], which stores 101,000 single food images belonging to 101 classes of to the Western cuisine; all pictures were taken in an uncontrolled environment; (2) the UEC Food-256 [2], which is an extension of the UEC Food-100 including 256 classes of Asian and Western comestibles, and (3) the UNICT FD889 dataset [35] with 3,583 single food pictures belonging to 889 dishes of international food items. At the time of writing, this database has the higher number of food classes. However, there are only few instances per class, since the dataset has been designed for “near duplicate food retrieval”. In 2016, the UNIMIB database [36] has been released to target the food classification issue on the pictures taken in the Western canteens. That is, all the pictures are multi-food items, where there is a dish for every comestible, and all the dishes are on the top of a tray

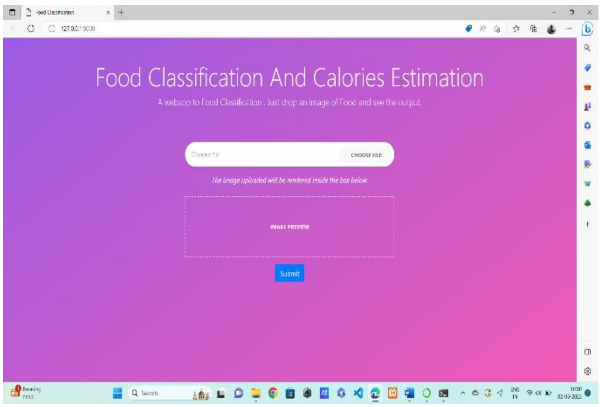


Fig 5. User Interface Website

Fig 5 displays the user interface to select the image from the files and to identify the food and displays the nutritional values for the user

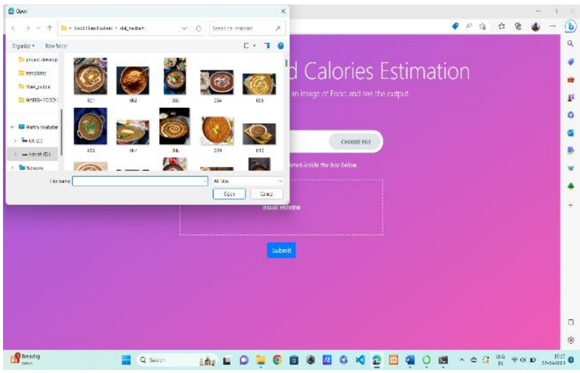


Fig 6. Selection Of Images For Search

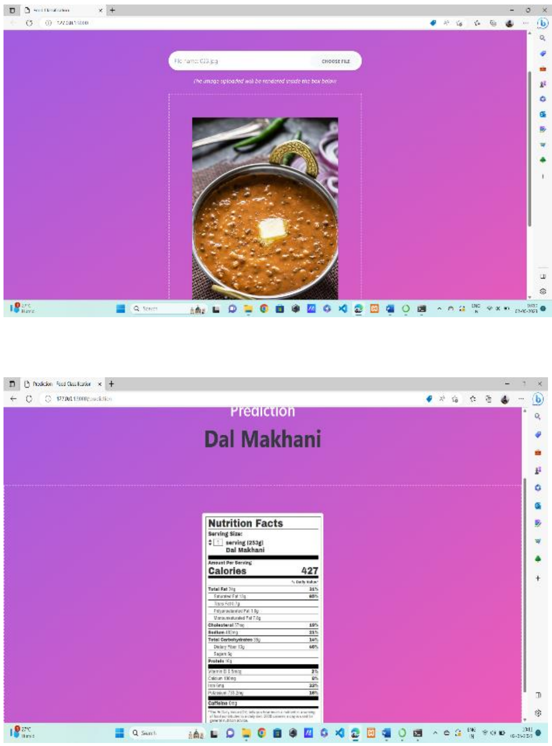


Fig 7. Output Details About The Uploaded Image

## VI. CONCLUSION AND FUTURE WORKS

Our main purpose of this study is to identify the calorie of the food from a given image which will help people overcome diseases like obesity, diabetes, heart problem, kidney failure, high blood pressure and other diseases caused by being overweight. We believe that if people know about the calories of the food, it will help them to lead a healthier life by keeping track of how much calories they are consuming.

We have researched and also looked into various methods for the food recognition process. We have studied deep learning and machine learning and feature extraction techniques such as ResNet, MobileNet, Inception etc. We have also studied Max-Pooling and some api such as keras, tensorflow etc.

CNN is most suitable for image or object detection processes. CNN can provide or give better outputs than other machine learning and deep learning algorithms. In case of performance, CNN outperforms other neural networks and machine learning and deep learning algorithms on conventional 2D or 3D image recognition tasks and other object detection tasks. CNN's are also useful for single dimension problems like time series, and in our case, 3D image classification where the images we used are food. In terms of statistical results, they also have high calculating efficiency in terms of object or image classification system.

At this stage of our research, we focused on building a Webpage based on CNN algorithm which is able to classify different types of foods and show us the nutrition value of these foods. We think this app will be very beneficial for this generation because people are obsessed with junk foods these days and they also do not want to get obese. So, while eating these high calories contained foods, they just have to take a picture using our app and can keep track of how many calories they are consuming in that particular time.

**SCOPE FOR FUTHURE ENHANCEMENT**

In the upcoming stages, we are planning on developing a mobile app that not only identifies the food items with a great accuracy from an image captured on the smartphone but also it will review the medical reports of the user to suggest whether the amount of calorie should be taken or not. Moreover, in the near future we want to work with a bigger dataset for example: a dataset of 100 or more food items using this model we have shown in this project. As such a dataset is not available, we are thinking of making our own dataset with 2000 or more images for each food category

**REFERENCE**


[1] B.Arslan, s.Memis, E.B.Sonmez and O.Z.Batur conference paper on " Fine-Grained Food classification Methods on the UEC FOOD-100 Database" : Raytheon Technologies. Downloaded on September 03,2021 at 09:18:38 UTC from IEEE Xplore.

[2] Weishan Zhang, Wenjuan Gong, Su Yang, D.Zhao ," Food image recognition with convolutional neural network" Published 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).

[3] K.Yani " The real-time mobile food recogmition system " IEEE Conference 2013

[4] "CEP- Calories estimation from food photos", Taylor and Francis, June 2018

[5] "Fuzzy clustering and Whale-based neural network to food recognition and calorie estimation for daily dietary assessment" Indian Academy of Sciences, 2018

[6] "Comparison of convolutional neural network models for food image classification", Taylor and Francis, Mar 2018

[7] "Neural network classifier and multiple hypothesis image segmentation for dietary assessment using calorie calculator", Taylor and Francis, Aug 2017

[8] "Computer vision-based food calorie: Dataset, method, and experiment", School of Information Science and Engineering, may 2017.

[9] "Food recognition and calorie estimation using Bag of surf and facial pyramid matching methods, 2017.

[10] "A deep Learning and Auto-Calibration Approach for Food Recognition And Calorie Estimation in Mobile e-health", School of EE and CS, 2015

[11] Horea Muresan, Mihai Oltean, Fruit recognition from images using deep learning, Acta Univ. Sapientiae, Informatica Vol. 10, Issue 1, pp. 26-42, 2018.

[12] T. Gupta, Deep learning: Feedforward neural network, Dec. 2018. [Online]. Available: https : / / towardsdatascience . com / deep - learning - feedforward -neural-network-26a6705dbdc7.36

[13] N. Martinel, G. L. Foresti, and C. Micheloni, \Wide-slice residual networks for food recognition," in 2018 IEEE Winter
Computer Vision (WACV), IEEE, 2018, pp. 567{576.

[14] S. Saha, A comprehensive guide to convolutional neural networks-the eli5 way, Dec. 2018. [Online]. Available: https://towardsdatascience.com/a-

comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53.
[15] M. A. Subhi and S. M. Ali, A deep convolutional neural network for food detection and recognition," in 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), IEEE, 2018, pp. 284-287.
[16] R. Balsys, Convolutional neural networks (cnn) explained step by step, Feb. 2020. [Online]. Available: https://medium.com/analytics-vidhya/convolutional-neural-networks-cnn-explained-step-by-step-69137a54e5e7.
[17] Feedforward neural network, Dec. 2020. [Online]. Available: https://en.wikipedia.org/wiki/Feedforward neural network.
[18] V. Kurama, A guide to resnet, inception v3, and squeezenet, Jun. 2020. [On-line]. Available: https://blog.paperspace.com/popular-deep-learning-architectures-resnet-inceptionv3-squeezenet.
[19] W. H. Organization, Obesity and overweight, https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight, 2020.
[20] Shashikant, Convolutional neural network: A step by step guide, Jan. 2020. [Online]. Available: https://towardsdatascience.com/convolutional-neural-network-a-step-by-step-guide-a8b4c88d6943.