Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 6 –Issue 2,August 2023
Paper 52

Food Image Detection using Deep Learning

1st G. Anirudhron
Department of ECE)
SRM Institute of Science and Technology
Chennai, India
ag8277@srmist.edu.in

2nd Harshith Ramesh
Department of ECE
SRM Institute of Science and Technology
Chennai, India
hr8285@srmist.edu.in

3rd Dr. P. Vijayakumar
Department of ECE
SRM Institute of Science and Technology
Chennai, India
vijayakp@srmist.edu.in

***Abstract*—Food is the key to a person's body. Therefore, a diet plan must always take the total number of calories to be consumed into account in order to maintain a fit and healthy lifestyle. Unfortunately, people find it challenging to estimate and measure the amount of food they consume due to a lack of nutritional information and the time-consuming process of manually recording this information. Consequently, it will be beneficial to have a system in place to track and control calorie intake. Therefore, accurate calorie prediction in such circumstances is equally crucial.**

In this project, we created a food recognition and calorie estimation system that recognises food using photographs provided by the user before estimating the number of calories included in the same food item. One of the promising uses for visual object recognition in computer vision is food image recognition. For the purpose of identifying food items, the system employs computational intelligence and image processing. To categorise the food images in high-resolution photos in each category, we trained amassive, deep convolutional neural network.

***Index Terms*—CNN, Convolutional, Image processing, Food detection**

I. Introduction

Food image recognition is the process of identifying food items in an image. Food recognition for dietary analysis, meal suggestion program, and even automating order taking in the restaurant industry are just a few of the numerous applications it can be useful for.

Deep learning models are capable of deciphering and analysing the intricate patterns and properties found in food photographs, such as colour, texture, and shape. A deep learning network can accurately classify various food items by being trained on a sizable dataset of labelled food photos.

Some of the open-source image processing and computer vision libraries for Python include OpenCV, TensorFlow, and Keras. These libraries can be used to build models for the detection of food photos using convolutional neural networks (CNNs) and other deep learning approaches.

TensorFlow: A well-known open-source machine learning framework called TensorFlow provides tools for creating and enhancing the models.

Keras: Keras is a high-level neural network API that is based on TensorFlow. It provides a simple user interface for learning and growing.

To develop a system that can recognise the type of food being input. The input has been pre-processed, and it can identify the dish's name and calories.

Make a computer program that can identify the sort of food displayed in an image and calculate its calorie count. Combining image processing and machine learning can accomplish this. The consumer will be able to keep a quality lifestyle and a nutritious diet thanks to this system.

II. Ease of Use

A. Support for non-experts

How effective the suggested method is for categorising food images using CNN depends on the target audience as well as the level of computer vision and machine learning proficiency.If the user is an expert in these disciplines, the suggested solution is reasonably easy to use because it leverages typical picture classification algorithms employing CNN.

To utilise and implement the suggested procedures, however, may require a certain amount of training or knowledge if the user is not familiar with these methods.

When determining the type of food, the model's usability, image clarity, and image angle can all pose challenges. The availability of appropriate datasets for training and testing the suggested approach can potentially hinder usability. Overall, the potential advantages for emergency response make the proposed technique a valuable tool for people working in the field, even though putting it into practise might call for a certain amount of experience and effort.

Large datasets of food images may require a lot of time and resources to gather and tag, which could affect how useful the suggested approach is.

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 6 –Issue 2,August 2023
Paper 52

III. DATA-SET DESCRIPTION

To ensure the model's fit and improve performance during training and testing, a sufficient data set must be acquired and cleaned.We get good accuracy when the data is clear and of a good resolution. When the data is clear and of a good resolution, we gain good accuracy.

A training set, a validation set, and a test set. A portion of the data is often used for training, another for testing, and a third for validation. The model's training effectiveness is evaluated using the validation data subset, and its hyper parameters are adjusted. The model is trained using the subset of training data. After training, the model's performance is evaluated using the test data subset.

IV. MODEL AND METHODOLOGY

A potential technique for categorising food photos requires convolutional neural networks (CNN). The methodology's three main steps are data collection and pre-processing, model training and evaluation, and model testing and validation. The initial step entails gathering a dataset of images featuring diverse culinary items.

A. Methodology

The dataset is pre-processed to eliminate any extraneous images. The next step is to train the model, after which it may be used to recognise food photos. employing CNNs (convolutional neural networks) or other supervised learning techniques for object detection.

Following the identification of the food item, the weight is used as an input to estimate the total number of calories. Images of food can be detected using Convolutional Neural Networks (CNNs), which are extensively employed in image detection applications.

A high-level breakdown of the procedures needed in creating a CNN model for food image detection is provided below: Data Collection: assemble a big library of images pertaining to food. Regarding the various cuisines and food items that the model will need to recognise, the dataset should be accurate and diverse.

Data preparation Pre-process the images by scaling them to a standard size, normalising the pixel values, and applying data augmentation techniques like rotation, flipping, and cropping to enhance the size of the dataset and the generalizability of the model.

Model Architecture: Create a CNN architecture as a model utilising one or more fully linked layers, pooling layers, and numeric convolutional layers as the foundation. The convolutional layers' collected properties from the input photos are used by the fully connected layers to categorise the images.

Training: On the pre-processed dataset, train the model using a loss function like categorical cross-entropy and an optimisation method like stochastic gradient descent (SGD). During training, the model develops the ability to distinguish between the traits of various foods and meals.

Testing: Run tests on various test sets to see how well the final model performs on brand-new, contemporary images.

Deplyoment:Deploy the model in a production environment to recognise food images in real-time applications like restaurant menus, food delivery apps, and nutrition tracking programmes.

In addition to these procedures, the performance of the model can be improved by using a range of complex techniques like transfer learning, ensembling, and fine-tuning.

B. Algorithm Used

1) Convolutional Neural Network: Artificial neural network applications for object detection and image processing frequently use convolutional neural networks (CNNs). The ability of CNNs to analyse high-dimensional input, such photographs, with a comparatively small number of parameters is one of their key advantages. Today, they are able to perform complex image analysis and recognition. CNNs have revolutionised a variety of computer vision applications, such as face identification, object recognition, and self-driving automobiles.

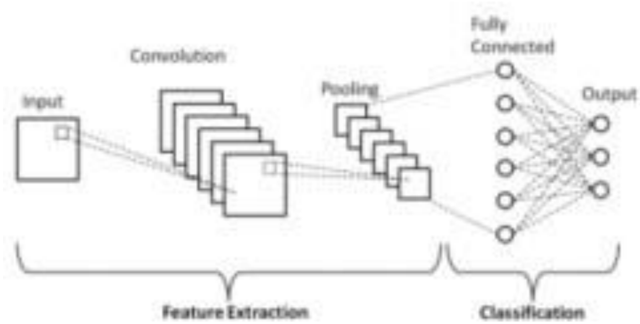


Fig. 1. CNN Architecture.

The structure of a CNN is composed of convolutional layers, pooling layers, and a fully connected layer at the end. The convolutional layers learn feature maps by convolving filters over the input picture, whilst the pooling layers down sample the feature maps to minimise the spatial dimensionality. The fully connected layer combines the retrieved characteristics to produce a classification output.Numerous computer based applications, including face detection, material recognition, and self-capable driving cars, have been transformed by

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 6 –Issue 2,August 2023
Paper 52

CNN.

Residual Networks(ResNet): Residual Networks (ResNet) introduced the concept of a residual block, which is a feature added to the output feature created by processing the input with one or more convolutional layers.That is, a layer can feed layers that are several levels ahead of it in the architecture by skipping connection blocks and using leftover blocks. The word ”residual” in the block’s name refers to the mathematical operation that is used to determine the difference between the input and output signals.

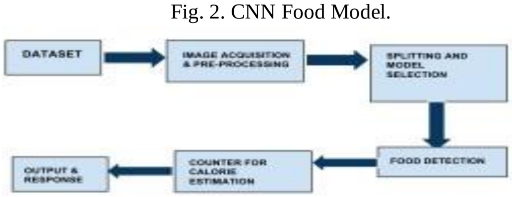
Although CNN-based systems are useful for identifying food images, some issues still need to be corrected. One issue is the model’s accuracy under varied weather and lighting conditions. The performance of the CNN may be affected by glare, shadows, and low light levels, among other factors. Another issue with real-time image processing is that it can be difficult to see rescue vehicles in densely populated areas.

In general, CNN have demonstrated good results in the detection of food images and have been applied in a number of contexts, including nutritional evaluation, food recognition systems, and program to reduce food waste.

C. MODEL

In the image processing project, a CNN model is employed to differentiate between photographs of food.
CNN first layers teach lower-level features like edges and corners, while their deeper layers teach higher-level features like textures and patterns. CNN are made to systematically learn and extract information from photos. Thanks to this, CNN can now accurately capture the fine visual details of food photographs.
No matter where a feature appears in the image, CNN can distinguish it because to the usage of pooling layers and convolutional filters. This is especially useful for recognising food photos because the positioning and orientation of food items may differ. Pre-trained CNN models can be used for food image detection by fine-tuning the model on a specific food image data set. This can save a large amount of time and computational resources because the pre-trained model has already mastered the recognition of a wide variety of features from sizable image data sets.

Fig. 2. CNN Food Model.



CNN can be supplemented by additional training data, such as rotated or flipped versions of the original images, to broaden the diversity of the training data and boost the robustness of the model.
This is especially useful for difficult image identification tasks like the detection of food photos, where it may be difficult to define hand-crafted characteristics.

CNN training can be computationally expensive and requires a lot of memory and computing power. This could make using CNN on devices with limited resources morechallenging.

CNN are prone to overfitting, especially when trained on small data sets or data sets with class imbalance. Strategies for regularisation like dropout and weight decay may be used to address this issue.
CNN are frequently referred to as ”black box” models since it can be difficult to read the learned features and know how the model is creating predictions. As a result, CNN might not be as helpful in circumstances when interpretability is important.

D. FOOD DATABASE

We are utilising the Food, Fruit and Vegetable data set for our research. A test set, a validation set, and a training set make up this data set. The data set has been sliced and filtered with a variety of photographs from various perspectives so that the model can detect and distinguish between the images. The data set is broken up into classes and contains about 300+ photos.

Fig. 3. Sample Datset used.

There are a very few food image databases, and they vary in many ways, such as the number of food groups, the overall number of images, the sort of cuisine such as Western, French, Italian, Turkish, Asian etc., the quality and type of the images, single versus multi-food images, and different contexts.
For deep learning-based food image detection to be successful, a high-quality database must be built. Deep

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 6 –Issue 2,August 2023
Paper 52

learning models that can recognise a variety of food items can be trained using a database with a broad set of food photos that spans many cuisines, dishes, and cooking ways.

The photographs must be accurately labelled with the name of the appropriate food item when creating a custom database for food image detection. To account for real-world events, the photographs should also be of a good quality and taken from several perspectives and lighting situations. Overall, for the successful training of deep learning models for food image detection, a high-quality and diversified database is required.

V. RESULTS AND DISCUSSIONS

As a result, we were able to build our machine learning model utilising the CNN architecture.

Fig. 4. Sample Output.

The model is trained in a way to identify a prticular food item and give the nutritional values or data. The count is been calculated and respective number is provided.

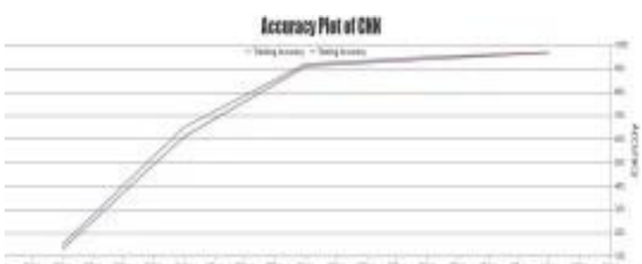


Fig. 5. Sample Output.

The accuracy of training and testing accuracy is been depicted in the graph.

VI. CONCLUSION

Ultimately, the main objective is to determine the type of food that was supplied as input and to estimate the number of calories the food will contain using its volume. The CNN algorithm is used in this operation, and the accuracy is close to 80 percent.The major objective of the work discussed in the paper is to recognise and display the nutritional counts in user- supplied food photos, which will assist them in monitoring and controlling how many calories and other nutrients they consume each day. As a result, chronic conditions including diabetes, renal failure, and heart disease may be prevented. Consequently, it is essential to monitor food intake

REFERENCES

[1] Kundan Kumar, Abhishek Pachpole, Akshay Shelke, Rutuja Mandapmalvi, Prof. Ankita Kotalwar, “Food Calories Estimation Using Image Processing”, May 2022.
[2] Ivars Namatevs, “Deep Convolutional Neural Networks Architecture”, March 2022.
[3] Narayana Darapaneni, Vishal Singh, Yasharth Singh Tarkar, Subhav Kataria, Nayana Bhansal, “Food Image Recognition and Calorie Preiction”, April 2021.
[4] Dipali Dhanwate, Rohini Jadhav, Payal Bhosale, Nishigandha Patil,A.S. Vibhute,” Features Based Fruit Gradation System using Image Processing”, May 2019.
[5] Manal Chokr, Shady Elbassuoni, “Calories Prediction from Food Images”, February 2017..
[6] Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun, “Deep Residual Learning for Image Recognition”, December 2016.