**Food Seeker**

**Mohd. Tabrez, Undergraduate scholar, CSE, MIT Moradabad; Pratham Kumar Singh Rathore, Undergraduate Scholar,CSE, MIT Moradabad, Author, Juhi Rastogi, Undegraduate Scholar, CSE, MIT Moradabad, Gunjan Radhawal, Undergraduate Scholar, CSE, MIT Moradabad, Mohd. Bilal, Undergraduate Scholar, CSE, MIT Moradabad**

**ABSTRACT**

The main focus of this research paper is a web application for food recognition of the recipe with just one camera click. This work has been implemented on two platforms: an Android Application and a Web application. It provides label to the food whose picture is taken of, then searches in its database for recipe. We have applied a Convolutional Neural Network (CNN) for detecting and recognizing food images. Image recognition of food is very difficult as a variety of food is available around us. However, deep learning has show great effect recently to be a very powerful image recognition technique. We used parameter optimization approach in CNN for detection and recognition, by constructing a dataset of the most frequent food items which is in a publicly available food-logging system, and used this data to evaluate recognition performance. Compared to SVM i.e support vector machine based method, CNN showed significantly higher accuracy with handcrafted features.

**1. INTRODUCTION**

The project performs the food recognition and detection of the recipe with just one camera click. Healthy diet is important to human health. Natural products have been widely used as food, and they can also be processed to meet the demand of consumers. Food attributes such as type, compositions, nutrients, and process styles are concerned issues for healthy diet. It is a fact that people from different regions have different eating habits. Knowing the attributes of food is important to inspect food quality and safety for consumers all over the world. Food is one of the most important components of our life. Records of what we eat are essential for our health care. The diets and eating habits can affect health of human beings. Especially for diabetics and allergic people, they should strictly monitor and control their dietary behaviour. Food recognition and classification is an important task to help human beings record the daily diets. Images of food are one of the most important information to reflect the characteristics of food. Moreover, image sensing is a relatively easy and low‐cost information acquisition tool for food appearance analysis. For natural products like food and processed food, the large variations in food shape, volume, texture, color, and compositions make food recognition a challenging task. Various background and layout of food stuffs also introduce variations for food recognition and classification. At present, due to the common use of CNN, image analysis has been the most commonly used pattern in food recognition and classification.

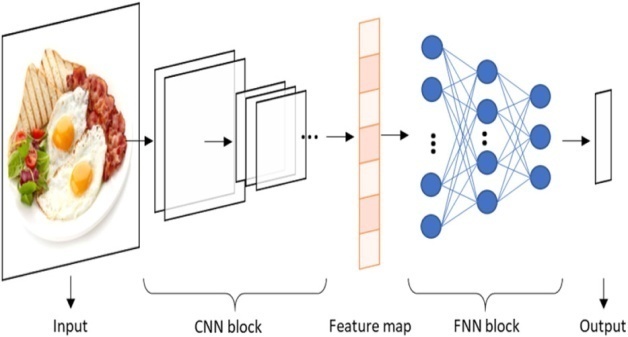
We have implemented the “Food Seeker” system, where the user takes photos of his foods and uploads them to the system, and the system performs image processing to detect food images and estimate food recipe. Kitamura et al. proposed a system which detects food images and estimates the food recipe. They extracted the three kinds of features; colour, circle and SIFT features, and put them into Support Vector Machine (SVM) and they were able to detect food images by 92% accuracy and the food balance by 38%. Joutou and Yanai attempted to estimate the food menu out of 50 items [1]. They used the Multiple Kernel Learning (MKL) method to integrate the Bag of Features (BoF), colour histogram, and Gabor texture features, and achieved the 61.24% classiﬁcation rate. Zhu et al attempted to estimate the food item and the volume of food by using the food images before and after eating. The dataset was limited to 50 images, and they prepared the same white dishes on a checkerboard to segment the food items and estimated. This work is supported by JST CREST. Wu and Yang used videos of eating behaviour to estimate the number of calories in a given meal. Their target was only fast foods and they used key point matching of SIFT features. In our project, we have investigated on how improve the accuracy by using user feedbacks. Our project “Food-Seeker” system enables the user to correct the results of the image processing. Then, making use of the user’s corrections as feedback, we can update the Bayesian network to improve the performance.

* 1. **Motivation**

Nowadays, deep learning has been introduced into food field by analyzing RGB images and spectra images of food. Recent technological advances such as smart-phones equipped with cameras and other rich sensors, pervasive networks and artificial intelligence have powered new uses of technology related with food. In conventional food logging for diet monitoring requires expertise and effort from the user, and is prone to inaccuracies and forgetting. In opposite to this, an automatic food annotation system could perform automatic analysis, annotation and logging with minimum human intervention. For example, photos from smart-phones are convenient and powerful entry points to many applications which involves recognition and retrieval. As a result, food-oriented social networks and restaurant review services have grown, where food enthusiasts connect and share their information regarding recipes, photos or comments about restaurants. So, hereby reliable food analysis from images is essential for these types of applications.

**1.2 Overview of Proposed System**

This online application enables the end users to take picture of the food, read the database and check for the label. The results, after taking food pictures will directly appear on the screen. The app has many areas open to extensibility, from utilizing more advanced machine learning models for food recognition, adding in recommendation systems, integrating better routing and price comparisons. We have applied a convolutional neural network (CNN) to the tasks of detecting and recognizing food images.

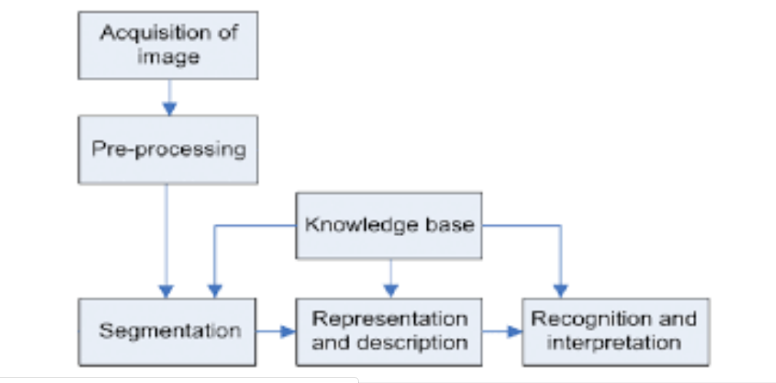


**Fig 1 A typical CNN structure for image classification**

* 1. **Application**

We have focussed on context and knowledge modelling, but there are also interesting applications of automatic food recognition to self-service restaurants and dining halls. For instance, accurate detection and segmentation of the different food items in a food tray can be used for monitoring food intake and nutritional information, and automatic billing to avoid the cashier bottleneck in self-service restaurants.The photos from smart-phones are convenient yet powerful entry points to many applications involving recognition, retrieval or recommendation. In this direction, food-oriented social networks and restaurant review services have bloomed, where food enthusiasts (e.g., foodies, gourmets, cooks) connect and share information (e.g., recipes, photos, comments about restaurants).Thus, reliable food analysis from images is essential for various applications. Despite remarkable advances in computer vision, food recognition in the wild still remains a very challenging problem even for humans as there are variety of dishes available in the surrounding. We largely rely on contextual and prior information. Similarly, context and prior knowledge can be integrated in automatic food analysis systems.

**2. LITERATURE SURVEY**

* 1. ** Stages Of Image Processing**

**Fig 2 Stages of image processing**

**2.1.1 Image Acquisition**

This is the first step of the fundamental steps of digital image processing. To deal with images and before analyzing them the most important thing is to capture the image in a proper manner and suitable angle. This is called Image Acquisition. Image Acquisition is achieved by suitable camera.

**2.1.2 Image Enhancement**

Image enhancement techniques are [mathematical techniques](https://www.sciencedirect.com/topics/engineering/mathematical-technique) that are used to improvt the quality of the selected image. It results in another image which focuses on certain features in a manner that are better as compared to their appearance in the original image. One may also derive or compute multiple processed versions of the original image accordingly, each presenting a selected feature in an enhanced appearance.

**2.1.3 Image Restoration**

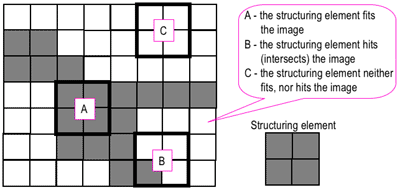
The concerns of the image restoration are the removal or reduction of degradations which are included during the acquisition of images e.g.; Noise, pixel value errors, out of focus blurring or camera motion blurring using prior knowledge of the degradation phenomenon. It deals with the modelling of the degradation and applying the process (inverse) to reconstruct the image. The image restoration has got a wide scope of usage in many different fields at large scale.The purpose of image restoration is to compensate for or undo the effects of the related images.

**2.1.4**  **Compression**

At its core, image compression is whatever you do to reduce the size of the selected image. It is used in certain situations as in [website optimization](https://www.keycdn.com/blog/website-performance-optimization) where many sites take time more than expected when the image is uncompressed which cause difficulty for the users/visitors. Compression is required for sending and uploading images and for reducing the storage impact on hard drive.

* + 1. **Morphological Processing**

Morphological techniques probe or select an image with a small shape or template called a structuring element in general. The structuring element is positioned at all possible locations in the image and it is compared with the corresponding neighbour-hood of pixels. Some operations test whether the element "fits" within the neighbour-hood, while others test whether it "hits" or intersects the neighbour-hood:



**Fig 3 Probing of an image with structuring element**

* + 1. **Segmentation**

Segmentation partitions an image into various regions with having similar pixels attributes of each. For the meaningful and useful for image analysis and interpretation, the regions should strongly relate to depicted objects. The reliability of segmentation is the key source of the success of image analysis, but an accurate partitioning of an image is generally a very challenging problem.

* + 1. **Representation and description**

Representation of regions can be done in 2 ways:

• Based on external characteristics (its boundary): – Shape characteristics

• Based on internal characteristics (its region): – Regional properties: colour, texture. Describes the region based on a selected representation: boundary or textural features

Description includes the concavities in the boundary, statistical measures of region.

* + 1. **Recognition and interpretation**

Image recognition is used to perform a large number of machine-based visual tasks, such as labeling the content of images, performing image content search and guiding autonomous systems. Image interpretation extract meaningful information from raw image data through various techniques. [Statistical Shape Models](https://www.sciencedirect.com/topics/computer-science/statistical-shape-model) (SSMs) provide a means to describe the variation in the shape of an object class across a set of images. It also allows the qualitative and quantitative analysis of image data.

### **Machine Learning**

**2.2.1 CNN (Convolutional Neural Network)**

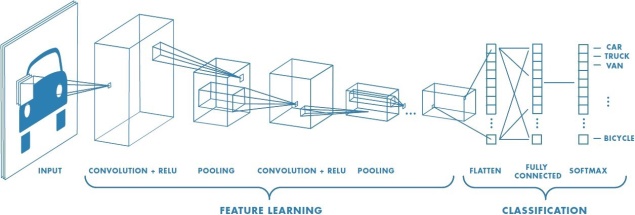
The convolutional neural network (CNN) is used to satisfy the requirements of the deep learning approach. It is also used for image recognition. In this project, we apply CNN to the recognition and detection of food images and evaluate its performance. We have followed following procedure:

(1) Firstly we built a dataset for food recognition experiments by using food-domain images obtained from a food logging system which is available for public use;

(2)Then we have performed optimization of CNN’s hyper parameters, showing that CNN signiﬁcantly improved the food recognition accuracy as compared with a conventional method using a support vector machine (SVM).

(3) After the observation of our trained CNN, we found that colour features dominate the food recognition process vastly;

Each unit of the final layer indicates the class probability in the image classification.



**Fig 4 CNN Architecture**

A CNN has hyper parameters which includes the number of middle layers, the size of the convolution kernels, and the active functions.

**2.2.2 Building of Dataset of Ordinary Food-item**

Here number of images of ordinary meals are required for the evaluation of food item recognition. A meal image involves several food items at large scale. Each food item region of the image need to be identiﬁed and isolated for the dataset for evaluation purpose. People can use the Android app for food learning for their own food recording using both photos and text. The user takes a photo of a meal, and speciﬁes each region involving a food item via the touch panel display of the smart-phone by inputting the name of the food item.

### **2.3 Brief introduction of deep learning**

Machine learning has been active in various fields, which acts as an effective tool for data processing. For the lack of ability to analyse raw natural data, traditional machine learning techniques usually needs to be supplemented by a manual feature extraction method. With the development of hardware computing ability and storage capacity, the abilities of machine learning can be promoted by adding more complex structures to achieve deep representation of the data. Representation learning enables a machine to extract the features from raw data for detection, classification, or regression.

CNN including a set of components (convolutional layers, pooling layers, fully connected layers, and so on) is currently considered as one of the most popular machine intelligence models for big data analysis in various research areas. Convolution operations are implemented by traversing input matrices with convolution kernels that can be understood as filters for feature extraction. Different from filters used in conventional image processing method whose parameters need to be set manually, the parameters inside the kernel can be learned automatically by deep learning method. Convolutional layers are built by a set of convolution kernels, whose parameters (channels, kernel size, strides, padding, activation, and so on) should be set and optimized according to the practical problem. The computed output from convolutional layer is then sub-sampled by pooling layers. A group of chained convolutional layers and pooling layers can learn high-level features representing the original input. The fully connected network (FNN) block, composed by fully connected neural units, is usually placed at the end as the classifier or used to generate numerical output for regression problems exploiting the learned feature map. Besides classification and regression, deep learning also demonstrates strong capabilities to process image segmentation tasks. Each pixel inside the output image represents a category.

**2.3.1 Training and Optimization of Deep Neural Network**

The weights of deep neural networks (DNNs) are initialized randomly or by Xavier method and tuned during the training procedure including forward and back propagation process. In the forward propagation process, the difference between the output value (and predicted value) and the label value (or ground truth) is calculated according to the defined loss function. In the back-propagation process, the weight of the neural network is updated to minimize the loss function via Stochastic Gradient Descent (SGD), AdaGrad, and Adam algorithms. The hyper parameters of the network, such as learning (which controls the pace of weight adjustment), batch size, number of convolution kernels and layers, and so on, could be fitted by evaluating the performance (the output of loss function) on validation set. Besides, the combination of feature graphs, the form of convolution kernels, and the parallel network structure should be considered depending on the specific problems.

**2.4. Naive Bayes**

Classification is a form of data analysis that extracts models describing all the important data classes. Effective and scalable methods have developed in for Decision Tree Induction, Naïve Bayes Classification, rule-based classification and many other classification methods. Advantages are easy to implement and good results obtained in most of the cases.

**2.4.1 Classification Model**

Classification is the processing technique of finding a set of model which distinguish data class so that this model can be used in the prediction of class of objects whose class label is unknown. It is the process to construct a model based on the training set and uses it to classify new data set. It is a supervised learning as observations; measurements are accompanied by known class labels in a large amount of training set and new data is classified based on training set.

**2.4.2 Naive Bayes Model**

Naïve Bayes Classifier classifies with the method probability and statistic, predict opportunities in future based on experience so is named as Bayes Theorem.

**2.4.3 Decision Tree Model**

A decision tree is a predictive modelling technique in machine learning and statistics that builds a simple tree-like structure to classify the data according to their categories. A decision tree is a flowchart like tree structure, where each internal node is denoting a test on an attribute and each branch represents an outcome of the test. The least nodes represent classes or class distributions.

**3. MODULES OF PROJECT**

**3.1 Module 1**

**3.1.1 Web Application for food recognition and detection**

As we all know that web application is a software application that runs on a remote server. So it give the leverage to the user that they can use this application from anywhere and from any place they want. They can use any food image that is available in their system to recognize and detect that food item along with their respective recipe.

**3.1.2 Drag and Drop**

Drag and drop is a very common feature. It is when one "grab" an object from a location and drag it to the different location.

**3.1.3 Process**

The process of building our application comprises of 4 stages:

1) Web-scraping

2) Dataset preparation

3) Models Building

4) Model Testing

* 1. **Module 2**

**3.2.1 Google Image Download Script**

We have used Google Images Downloader Script in Python to download photos from Google Images and creating the dataset of our own and train the data to recognize the food items that we take the picture of. With the help of this script we also take some erroneous images from google images therefore we can also delete those images to improve the quality of the dataset and increase the accuracy to recognize the image.

Next step is to run train script to create the dataset of food items. After performing almost 4 million training steps to increase the accuracy of recognition of the food items, the accuracy of the system is increased to 78.8%.

To further improve the accuracy of the project we collected more photos that will enhance the accuracy of the system. So for a particular food item, nearly 1000 images were collected. After the training process is completed, we export the TensorFlow models that will further integrate into Mobile App (Android App) using the technology of the Java.

Finally we install TensorFlow Lite Library to Mobile Project of the dataset.

**3.2.2 Food Recognition**

This model used to recognize food in a photo which is available in summary. For this we have downloaded numbers of photo per food on Google Image. After that having deleted unnecessary images and just keep correct images for the dataset and further used these images to train the model.

* 1. **Module 3**
     1. **Creating the Dataset**

Generally, food recognition datasets consist of images and their corresponding class labels, which are used to train food classifiers. These datasets have evolved to include progressively more food classes with various subclasses, from early datasets with a few number of cuisine-specific images to larger datasets that include a much larger number of images per class as larger the dataset bigger the accuracy, and cover wider ranges of foods and cuisines[3]. These large datasets are suitable for learning deep CNNs for food recognition.Datasets for recipe analysis and retrieval incorporate ingredients and also possibly other cooking information. The number of ingredients can vary from a few tens to several thousands.

Python is a multi-purpose language, widely used for scripting the fro various purposes. Let’s say we want to download google images with multiple search queries. Instead of doing it manually we can automate the process. We use the dataset with large number of images which are small, clearly labelled and with no noise which make it ideal for the task and take considerably less pre-processing.

**3.3.2 Challenges for creating Dataset**

The first goal is to be able to automatically classify an unknown image using the dataset, but beyond this there are a number of possibilities for looking at what regions / image components are important for making classifications, build object detectors which can find similar objects in a full scene, identify new types of food as combinations of existing tags. The challenge is to train models which can look at images of food items and detect the individual food items present in them. We use dataset of food images collected through the script creating using python and TensorFlow where numerous images of their daily food intake. This growing data set has been annotated - or automatic annotations have been verified - with respect to segmentation, classification, and weight / volume estimation. This is an evolving dataset, where we will release more data for the larger datasets the dataset grows over time.

**3.4 Module 4**

**3.4.1 Android Application**

Simply snap a food photo and get the information of your meal. Food-Seeker App is powered by Food AI API. Food AI API works on deep learning and image classification technology to accurately identify food items.

**3.4.2 Image Uploader**

It is first step of the project or the front end of our project where you click an image and upload it for the further detection process. The results after taking food pictures will directly appear in the screen. The app has many areas open to extensibility, from utilizing more advanced machine learning models for food recognition, adding in recommendation systems, integrating better routing and price comparisons, and allowing users to submitted recipes with voting and reviews whether positive or negative.

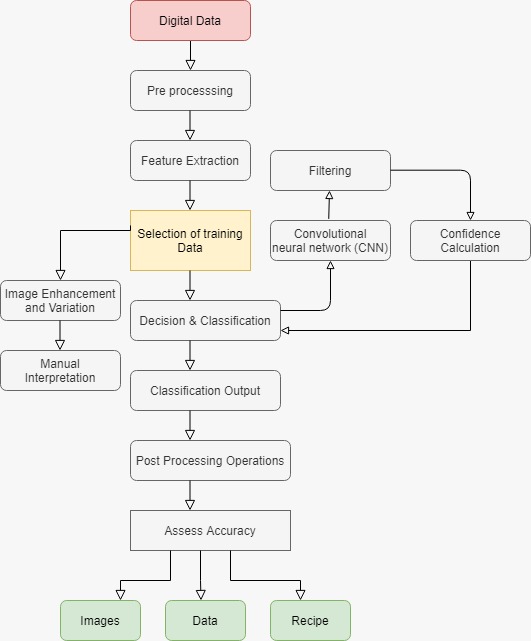
**3.4.3 Crop the Uploading Image**

In the step we crop the uploading image using the android application of the image which we take the picture of. Upload the image you want to crop. After image is loaded, choose the cropping if needed.

**3.4.4 Getting Features From The Image**

In this we get the features or the information about the input images in the application. By this we get information regarding the food, we recognize the food item in the picture and get the respective recipe from it.

* 1. **Flow Chart of the Project**

****

**Fig 3.1 Flow Chart of Complete Project**

**4. METHODOLOGY**

1│── **Data preparation**  
│ └── Clearing data  
│ └── Data augmentation

2│── **Data analysis and visualization, split data (Train, Valid, Test)**

3│── **Topic Modeling**  
│ └── Latent Dirichlet Allocation (LDA)   
│ └── Non-negative Matrix Factorization

4│── **Feature Extraction**  
│ └── k-nearest neighbor’s  
│ └── t-SNE visualization

5│── **Transfer Learning: Training pre-trained CNN (Convolutional Neural Network)**  
│ └── AlexNet, VGG, ResNet, GoogLeNet

**5. CONCLUSION**

In this paper, we have described framework for food detection and analysis. We have used the concept of Convolutional Neural Network for food detection and identification.

5.1. Limitation of the Project

Our project didn’t consider the nutrients

**5.1 Future Scope**

We can improve our system to detect the ingredients that are used for that particular food item and can provide the details of nearby restaurants which can provide it to the user.

More work can be done to restore blurred images given as input by the user.

**6. REFRENCES**

[1] [Yuto Maruyama](https://ieeexplore.ieee.org/author/37594025900); [Gamhewage C. de Silva](https://ieeexplore.ieee.org/author/37297098300); [Toshihiko Yamasaki](https://ieeexplore.ieee.org/author/37270142500); [Kiyoharu Aizawa](https://ieeexplore.ieee.org/author/37335828200), Personalization of food image analysis, 2010 16th International Conference on Virtual Systems and Multimedia

[2] F. Zhu, M. Bosch, I. Woo, S. Kim, C. J. Boushey,D. S. Ebert, and E. J. Delp. The use of mobile device application android application. IEEEJournal of Selected Topics in Signal Processing, 4(4):756–766, 2010.

[3] Luis Herranz, Weiqing Min and Shuqiang Jiang, “Food recognition and recipe analysis: integrating visual content, context and external knowledge”

[4] F. Kong and J. Tan, “Dietcam: Automatic dietary assessment with mobile camera phones,” Pervasive and Mobile Computing, vol. 8, no. 1, pp. 147–163, 2012.

[5] O. Beijbom, N. Joshi, D. Morris, S. Saponas, and S. Khullar, “Menu- match: Restaurant-speciﬁc food logging from images,” in Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on. IEEE,

2015, pp. 844–851.

[6] Y. Kawano and K. Yanai, “Foodcam: A real-time food recognition system on a smartphone,” Multimedia Tools and Applications, pp. 1– 25, 2014.

[7] F. Zhu, M. Bosch, I. Woo, S. Kim, C. Boushey, D. Ebert, and E. Delp, “The use of mobile devices in aiding dietary assessment and evaluation,” Selected Topics in Signal Processing, IEEE Journal of, vol. 4, no. 4, pp. 756–766, 2010.

[8] Z. Ahmad, N. Khanna, D. A. Kerr, C. J. Boushey, and E. J. Delp, “A mobile phone user interface for image-based dietary assessment,” in IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, 2014, pp. 903 007–903 007.