Grenze International Journal of Engineering and Technology, June Issue

Calorific Value Estimation and Cuisine Recommendation using Neural Networks

Vineeth R
Jain University, Bengaluru, India
Email: vineeth.artifintell@gmail.com

***Abstract*—In this modern era, people are more interested in having good health and lifestyle. There are many factors that influence health, such as exercise, sleep, nutrition, pollution, etc. Due to the availability of a variety of food options, people tend to consume more food than the body needs. So, people take in more calories than required by the body (required calories: 2000-3000 Cal per day). This leads to weight gain(Obesity), which further makes them vulnerable to chronic diseases(heart diseases and cancer). So, from this, we can say that people are unaware of the number of calories they consume. So, the proposed system can recognize the food item by using a mobile camera and can give calorific value estimation and nutritional Values of the specific food item, and it also recommends other food items with a lower calorific value similar to the given food item using a Convolutional neural network and K-means Clustering Algorithm. By, this you can keep track of your calorie intake regularly.**

***Index Terms*— Obesity, Deep Learning,, Calorie Estimation, K-means Clustering, Naive Bayes, Recommendation, Convolutional Neural Network, Transfer Learning, Diet.**

I. INTRODUCTION

The main source of energy for all living beings is food.It provides energy to carry out different life functions, maintains the chemical composition of the cells, and also it renders protection from different infections and diseases. Because of changes in lifestyle and evolution of food items, people get drawn to new tastes and varieties, so people consume more and more, leading them to an unhealthy lifestyle. Nowadays, maintaining a healthy lifestyle and diet may be a challenge for many people because of complications in measuring food and dietary calorie intake. More calorie intake leads to obesity, because of the decomposition of unburnt fat because of less physical activity. So, we must always monitor our regular calorie intake and have regular exercise. A healthy and fit lifestyle can remarkably reduce the severity of diabetes and minimize its complications[1]. By regular exercise and consuming a low calorie diet, we can improve the standard of life of type 2 diabetic Mellitus people[2]. Due to advancements in computer vision, deep learning, and neural networks within the past five years, we can manipulate it to recognize food items and estimate their calories and provide us with recommendations. The estimation of the dietary calorie intake of daily foods renders a necessary solution to steer a disease-free healthy and fit lifestyle.

II. LITERATURE SURVEY

Image recognition is the process of detecting and identifying specific features from the given image. The models

Grenze ID: 01.GIJET.8.2.12
© Grenze Scientific Society, 2022

are trained using various-labeled datasets of images using a specific learning algorithm. These days ,food recognition models are extensively being used by food and medical industries to determine the intake of dietary calories and nutrients to help people in maintaining a healthy lifestyle[3]. Nowadays, people's awareness regarding health has significantly increased and they give more intention towards calorie estimation tasks. The dietary calorie estimation may start with identifying the food item and estimating the calorie based on the nutrient datasets available. In this , we propose a healthy food recommendation system, with the core idea to recommend an individual and customized healthy diet to users. So by incorporating the computing capabilities of mobile devices and deep learning algorithms, we present a novel method using deep learning techniques which can recognize food items, estimate their calorific value, and recommend food from different cuisines using the K-means clustering algorithm.

III. Proposed System

Food item recognition is an existing idea that can detect and recognize the food item based on the input image given by the user. We have trained our models with various food image datasets like food-101[3], Indian food 101[4] by transfer learning using inception v3[5]. Inception v3 architecture is widely used for the image classification with greater accuracy and it has 7 × 7 convolutions and is 48 layers deep. For calorie estimation, we have used a dataset from calorie counter[5] and trained using multinomial naive Bayes algorithm and for the recommendation, we have used K-means clustering[6] on the yelp restaurant dataset[7]. So, The user fed image of the food item is recognized by the image classifier and the calorific value estimator gives the estimated calorific value of the food based on the recognized food item,and it also suggests other food alternatives. Both these values are given to the recommender as input and it suggests you with appropriate cuisine.

IV. Architecture

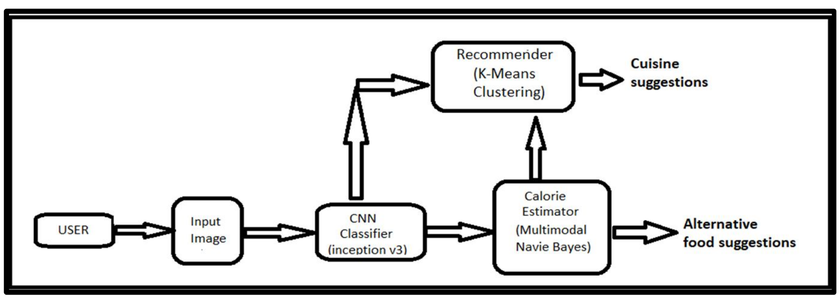

Fig. Model Architecture

V. Methodology

A. Convolutional Neural Network (CNN)

Deep Convolutional Neural Network(D-CNN) is a type of deep neural network that has neurons with specific weights and biases in each of the layers. The common scope of a deep convolutional neural network is in Figure 2.There are five main layers in CNN. These are the convolutional layer, batch normalization layer, pooling layer, and dense layers[9].

Let us consider an image as input . For every patch of input image, the CNN layer calculates output of neurons,when an image is passed through the CNN layer. During this operation, a filter-like matrix moves over each and every pixel of the given image to produce a feature-map. A feature-map is the dot product of the neurons co-efficient and image vector.

Major things to notice is that we should initialize the weight of the matrix correctly. In this approach, we have used kernel weights through which we can reach the global minimum Weight cost functions more efficiently and faster. Depending upon the weights in the previous layer, the values differ,and it is corrected using the backpropagation technique. The regularization method is being used to boost the learning rate by optimizing the algorithm. In this approach, we have used L2 regularization with the aim of reducing the complication of the presented model.Thus, it reduces overfitting by optimizing large weights.

Some activation functions or transfer functions are used in the convolutional layer output to map it to a certain range from 0 to 1 or -1 to 1.Most reliable activation functions that are being used are Sigmoid function, RELU

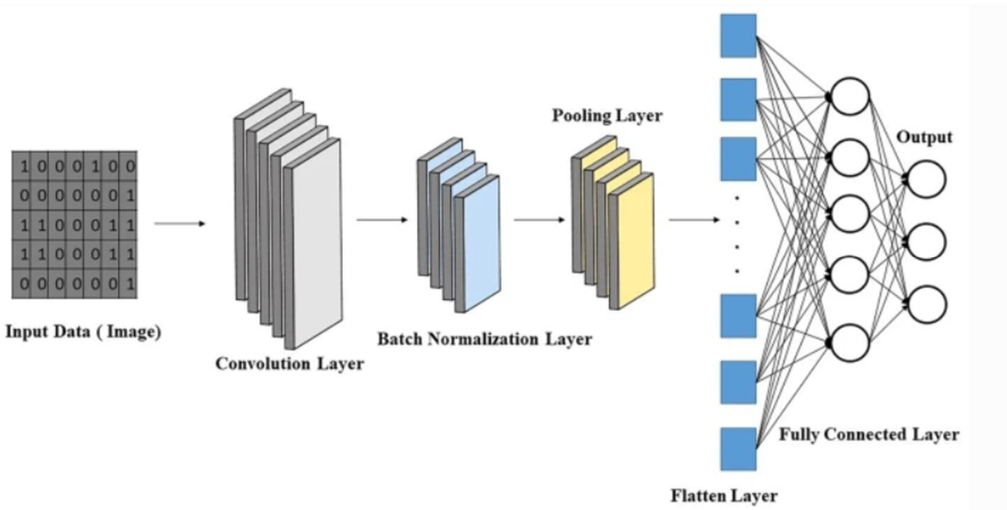

Fig. 2. CNN Basic Architecture

function, linear function, and hyperbolic tangent function(Refer fig 3). Usually, batch normalization techniques are used to steer up the process of learning. Pooling layers are used on feature maps to reduce their spatial size so that they can learn important features in the image. We can make use of a flatten layer to convert multi-dimensional feature maps to one-dimensional feature maps. These flattened layers are passed through dense layers to construct prediction models. Output layer calculates probability score of the image in each and every class and the class that has high score is selected as the predicted class.

In this approach, we have used the ReLu activation function in the convolutional layer to produce a sparse matrix. In comparison to the Sigmoid function, it requires smaller computational power and it also prevents possibilities of vanishing gradients; this results in faster learning. Finally, Applying softmax function to the final layer to find the probability score of each food class.

Here we have used Adam Optimizer because it outperforms classical stochastic gradient descent,and it is used to solve noisy problems or sparse gradients. It has properties of two important algorithms that are AdaGrad and RMSProp.It is very easy to implement and it consumes compact memory.

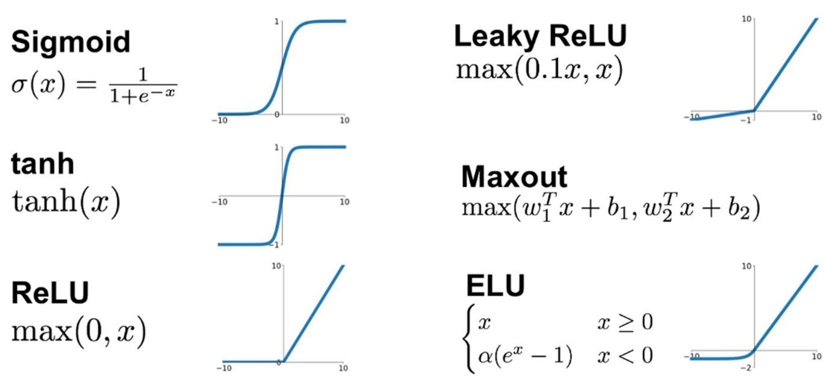

Fig. 3. Activation Functions

B. Transfer Learning

This methodology uses intelligence acquired by the pre-trained model to train convolutional models using a different dataset. We require a huge dataset to train a deep Convolutional model from the start. If we train deep CNN with inadequate data, the model gets worse and sometimes it gets overfitted to training data. Even if huge data is not available,we can use transfer learning techniques to train our deep CNN model.

This can be implemented in two methods by using feature extraction or fine-tuning. Feature extraction approach is done by giving a new dataset to a pre-trained CNN model as input to get feature-vectors in some batches.

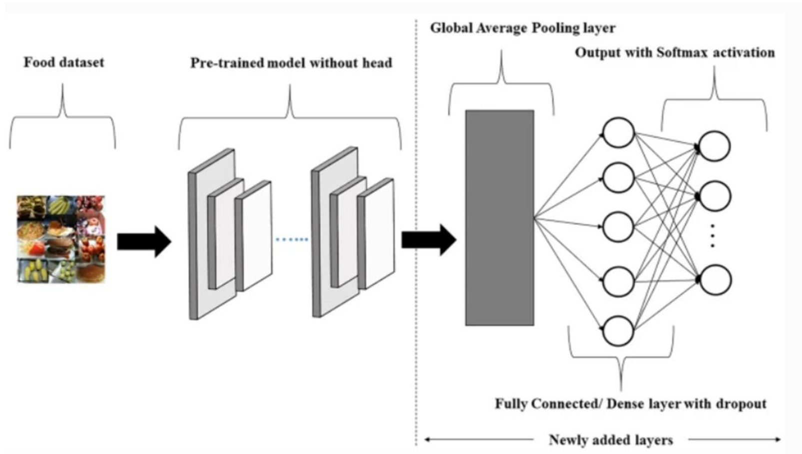

Fig. 4. Transfer Learning Method

These features are used to create prediction model classifiers. In the fine tuning method, we optimize the model by putting back the final layer by a classification layer and optimizing it depending upon the classes present .The presented model is executed without head, means executed without fully-connected layers. In this work transfer learning with fine-tuning is implemented using inception v3 architecture.In this, we have used a sequential API from Keras to create CNN and tensor flow to train our dataset. Fig 4 general setting of the fine-tuning method used in this work. On the top of the head, a global average pooling layer is added and on top of it, a dense layer is added with 512 neurons. A dropout is also applied over a dense layer to prevent overfitting. Finally, on top of it, an activation function softmax is added to classify food items.

C. Calorie Estimation Using Multinomial Naive Bayes

Using the calorie dataset we have created a dictionary with the key value pairs and passed to the TF-IDF vectorizer. The word frequency-inverse document frequency vectorizer is a popularly used algorithm to convert the textual information into numbers,which is used to fit it to a machine learning frequent algorithm for predicting results. TfidfVectorizer has functionalities of both CountVectorizer, TfidfTransformer. TF-IDF Vectorizer considers the overall word weightage of a document [6]. It helps in dealing with almost all the words. TF-IDF Vectorizer weights the word counts by computing how often a word appears in the particular word file. The TF-IDF Vectorization formula is shown in fig 5.

$$\text{tf-idf}(t) = \text{tf}(t, d) \times \text{idf}(t)$$

Fig. 5. TF-IDF Vectorizer Formula

Multinomial Naive Bayes is a learning algorithm that is most popularly used for Natural Language Processing tasks. This algorithm totally depends upon Bayes Theorem (Figure 6) and it helps to find word tags in the given dataset. It computes the probability score of every word tag for a given sample and returns the word tag with the highest score. It is generally applied for textual data analysis and multi-class problems. We have used the output of vectorization to fit the model to predict calories and to suggest similar food alternatives. We have used this algorithm because it is very simple to execute as we only compute the probability score and it can be widely used for large datasets and as it is highly scalable.

$$P(A|B) = P(A) * P(B|A)/P(B)$$

Fig. 6. Bayes Theorem

D. Recommendation using K-Means Clustering Algorithm

This clustering algorithm gets input data and it returns the specific cluster to which the data belongs. Clustering aims to divide a given dataset into subsets on the basis of similarity of attributes and each of these subsets is called a cluster. The K-means algorithm is kind of classical clustering algorithm. It divides the dataset of N foods into K clusters based on the similarity degree attributes[6]. The similarity degree of various foods is measured by using a sum of squared errors(SSE)(refer to figure 7).

$$E = \sum_{i=1}^{k} \sum_{p \in C_i} (p - m_i)^2$$

Fig. 7.Sum Of Squared Errors

The scope of the K-means clustering algorithm is to add some data points(K) to the given data and these points are called centroids. It moves around and it centers itself to the one of the k clusters [8]. when the data points stop moving around the algorithm ends. The steps of the K-means algorithm are shown.(Figure 8). We have trained the K-Means model on 12 clusters and on Food calories, Food name, and Food cuisine. If we give new food and calorie from classifier and calorie estimator the algorithm predicts and suggests us with appropriate cuisines.

K-means clustering algorithm:

1. Input: k (the number of clusters), D (a set of lift ratios)
2. Output: a set of k clusters
3. Method:
4. Arbitrarily choose k objects from D as the initial cluster centers;
5. Repeat:
6. (re)assign each object to the cluster to which the object is the most similar,
7. Based on the mean value of the objects in the cluster;
8. Update the cluster means, i.e., calculate
9. the mean value of the objects for each cluster
10. Until no change;

Figure 8 K-Means Algorithm

VI. Future Scope

The web application can be converted into a mobile application for better user experience and regular use. More categories of food can be trained with huge datasets for higher accuracy and prediction. Image recognition accuracy can also be increased by fine-tuning with appropriate hyperparameters. With this, we are also able to detect different layers present in food items. We can calculate calories using depth and volume estimation. In this, we can estimate volume by voxel grid estimation by superimposing a grid of squares on the image such that every square has an equal number of pixels and equal area [10]. We can try to fine-tune the Gaussian Naive Bayes model for better performance in alternative food prediction. In future enhancements, we have planned to show both calorie and nutritional information of the recognized food by implementing a user interface.

VII. Conclusion

To help people to maintain a healthy lifestyle and to manage their regular calorie intake ,Our paper presented a novel method based on estimating food calorie intake, recommending food alternatives and cuisines based on deep learning techniques. The method firstly uses food datasets to train a deep Convolutional to recognize and classify food items using image and we have also train another model for calorie estimation and food alternative suggestion using Multinomial Naive Bayes, and finally use K-means clustering to recommend appropriate cuisine to the users. The idea behind the presented system is to assist people in finding well individual and customized food spending upon their daily diets and help them to steer a healthy lifestyle , thus reducing the complications developed by eating unhealthy foods. The above presented model has proven to have a good performance and high accuracy and recommendations in real-time interaction with users.

References

[1] Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med. 1993;329(14):977–986.

[2] Snel M. Sledding MA, Peijl ID, et al.” Quality of life in type 2 diabetes mellitus after a very low-calorie diet and exercise”. Eur J Intern Med. 2012;23(2):143–149. doi:10.1016/j.ejim.2011.07.004
[3] Howard AG, Zhu M, Chen B, et al. “MobileNets: efficient convolutional neural networks for mobile vision applications”. 2017.
[4] Youguo Li, Haiyan Wu, “A Clustering Method Based on K-Means Algorithm”.2012
[5] Pouladzadeh P, Shirmohammadi S, Al-Maghrabi R. “Measuring calorie and nutrition from a food image” IEEE Trans Instrum Meas. 2014;63(8):1947–1956.
[6] David Joseph Attokaren, “Food classification from images using convolutional neural networks Conference Paper” November 2017 DOI: 10.1109/TENCON.2017.8228338
[7] Maiyaporn Phanich , Phathrajarin Pholkul , “Food Recommendation System Using Clustering Analysis for Diabetic Patients”, January 2010, 10.1109/ICISA.2010.5480416
[8] Chang Liu, Yu Cao, Yan Luo, Guanling Chen, Vinod Vokkarane, and Yunsheng Ma. “Deepfood: Deep learning-based food image recognition for computer-aided dietary assessment”. In ICOST, 2016.
[9] Chae, R. Maciejewski,, I. Woo, S. Kim, C. J. Boushey, , F. Zhu, E. J. Delp, and D. S. Ebert, “Volume estimation using food specific shape templates in mobile image-based dietary assessment”, SPIE, vol. 7873. NIH Public Access, 2011.
[10] Taehee Lee and Stefano Soatto. Learning and matching multiscale template descriptors for real-time detection, localization and tracking. CVPR 2011, pages 1457–146, 2011.
[11] L. Bossard, M. Guillaumin, and L.Van Gool, “Food-101–mining discriminative components with random forests,” in Computer Vision– ECCV 2014. Springer, 2014, pp. 446–461.