

SURVEY ON APPLICATIONS FOR PREDICTING CALORIES AND NUTRITIONAL VALUES FROM FOOD

Rachoti Biradar¹, Rahul Thiru², Ravi Purwar³
Guided By : Latha N R⁴

Abstract - Recently, people are becoming used to a modern lifestyle since they can be fully consumed by busy schedules at work and at home. Obesity in adults is becoming a common problem. The main cause of obesity is a combination of excessive food consumption and lack of physical activities. Consuming food with a high amount of calories can cause several problems for our health. There is widespread nutritional information and guidelines that are available to users at their fingertips on the internet. However, such information has not prevented diet related illnesses or helped patients to eat healthily. In most cases, people find it difficult to examine all of the information about nutrition and dietary choices. Furthermore, people are obvious about measuring or controlling their daily calorie intake due to the lack of nutritional knowledge, irregular eating patterns or lack of self-control. Recording the amount of calorie intake during each meal is a tedious task. Although people can record their meals and discuss with doctors or experts, it is not very convenient for them to know the number of calories before the meal. Our goal is to empower users by a convenient, intelligent and an accurate system that helps them become aware of their calorie intake and also find the individual nutrients content in the food item.

Keywords - Machine Learning, Data Analytics, Neural Networks, Linear Regression, Multiple Linear Regression, Data Mining, Decision Trees, Food.

1. INTRODUCTION

Food image recognition is one of the crucial applications used these days. It allows smartphone users to know the name of the food. Many people are interested in tracking what they eat to help them achieve weight loss goals or manage their diabetes or food allergies. This is quite important for travellers who travel to foreign countries. However, most current mobile apps (MyFitnessPal, LoseIt, etc) require manual data entry, which is tedious and time-consuming. Consequently, most users do not use such apps for very long. Furthermore, amateur self-reports of calorie intake typically have an error rate that exceeds 400 calories per day. Several previous approaches rely on an expert nutritionist to analyze the image offline (at the end of each day). Other approaches use crowdsourcing to interpret the image, instead of an expert. However, crowdsourcing is both costly and slow, which hinders widespread adoption. We mainly focus on Indian food as no app in the market currently caters to it. Several existing works do use computer vision algorithms to reason about meals but only work in laboratory conditions where the food items are well separated and the recognition, however, has been

example using image segmentation to separate the food from the background image. This technique will increase the effectiveness of food identification. In this paper, we take some initial steps towards such a system. Our approach utilizes several deep learning algorithms, tailored to run on a conventional mobile phone, trained to recognize food items and predict the nutritional contents meals from images taken "in the wild".

2. LITERATURE SURVEY

A. NU-InNet: Thai Food Image Recognition Using Convolutional Neural Networks on Smartphone Since AlexNet[2] and GoogLeNet[3] the storage space required for the model is huge these people proposed their model which is NU-InNet[1] (Naresuan University Inception Network), which is helpful for mobile applications because, in mobile applications, Processing time and storage space should be minimum. so the inception module adopted by GoogLeNet was adopted and further improved by keeping the accuracy at the same level. In that NU-InNet model, they have created two versions I.e., NU-InNet 1.0 and NU-InNet 1.1 so in NUInNet 1.0 model the GoogLeNet inception model is modified by changing its 3×3 max-pooling layer and 1×1 convolutional layer to be 1×1 and 7×7 convolutional layers, respectively, and in NU-InNet 1.1 network the NU-InNet 1.0 is modified by changing [4,5] any 5×5 convolutional layer to be 2 3×3 convolutional layers and changing any 7×7 convolutional layer to be 3 3×3 convolutional layers By doing this the accuracies of the models increased compared to the models which own at ILSVRC, which is helpful for the mobile applications.

B. A New Deep Learning-based Food Recognition System for Dietary Assessment on An Edge Computing Service

1. Rachoti Biradar is currently pursuing bachelor's degree in Computer Science & engg in BMS College of Engg, Bangalore.
Email : 1bm16cs075@bmsce.ac.in
2. Rahul Thiru is currently pursuing bachelor's degree in Computer Science & engg in BMS College of Engg, Bangalore.
Email:1bm16cs077@bmsce.ac.in
3. Ravi Purwar is currently pursuing bachelor's degree in Computer Science & engg in BMS College of Engg, Bangalore.
Email : 1bm16cs080@bmsce.ac.in
4. Latha N R is Assistant Professor at Computer Science department BMS College of Engg, Bangalore. Email : latha.cse@bmsce.ac.in

mainly focused on the correctness of the food name for the given food image. Many techniques are applied, for

Infrastructure This model is also directly inspired by AlexNet[2] and GoogLeNet[3] along with LeNet-5[6].this paper focused on major 2 research efforts, the first effort aims to develop new food image recognition algorithm using deep learning, processing and segmentation algorithm and the second effort aims at designing a real-time food recognition system.in this model, the computing is done two times one at the user mobile phone and the other at the cloud which is generally considered as the edge computing in the mobile application the food image is checked for the blurriness if it is not then the image background information is reduced for accuracy by using watershed segmentation[7]and the cropped image is sent to a server in the cloud for further computation and in the backend, the existing AlexNet[2] is modified by adding 1x1 convolutional layers and also instead of feeding input directly into the convolutional layer, an additional convolutional layer with size 1x1 is added to reduce the input dimension. By doing this it outperformed existing models and also it developed a real-time system whose response time is close to the minimal of existing techniques.

C. Food Image Recognition with Convolutional Neural Network As proposed by authors in [8], mainly this model consists of five layers in which four are convolution-pooling layer and the fifth one is a fully connected layer, the input is passed into the first layer which takes 128 x128 x3 size image and after filtering the image is passed on to the second convolution network and like that the process continues until the fourth convolution layer and the output of this layer is passed on to the fully connected layer which consists of 1000 neurons. Here they have used two sets of datasets UEC-FOOD100 dataset [19] and another dataset for fruit images which they have created and they used techniques called flipping and blurring on images to enlarge the dataset to reduce overfitting problem and for training the model they have gradient descent approach with batch size 100 and initial learning rate of 0.01, and epochs of 200. By doing this they were getting the accuracy of 80.8% for the fruit image dataset and 60.9% for the UECFOOD100 dataset and they also concluded that the accuracy is low because of the small-sized dataset.

D. Inverse Cooking: Recipe Generation From Food Images As the name suggests the authors[9] have created a model that takes food image as input and output sequence of cooking instructions using instruction decoder which takes input as two things one is the visual features extracted from the image and second one is the ingredients extracted from a food image. Cooking instructions are generated using instructions. transformer, which needs two inputs one is image representation which is extracted using ResNet-50[10] and ingredients embedding which is a set of ingredients using these the recipes are generated and for recipe dataset, they have used Recipe1M dataset from which they have used only recipes containing images and excluded recipes with less than two ingredients and two

instructions. Using this we are going to get the recipe of food using the food image.

E. Calories Analysis of Food Intake Using Image Recognition In this model, the authors[11] have developed a model that is used to find calories contained in that food. Here they have used 3 techniques Image segmentation, feature extraction and Classification this model mainly concentrates on Thai food. In image segmentation, the portion without food is made as black and the cropped image which contains only food image is forwarded to the feature extraction where since Thai food has distinct characteristic and color for feature extraction mainly 3 techniques were used Bag- of-Features(BOF), Segmentation based Fractal Texture Analysis(SFTA), and Color Histogram the individual vectors from these three techniques were combined and the resulting feature vector is passed to SVM which is used to recognize the input image. By using this algorithm we can classify the food image into food type and the associated calories using the associated database. This is the feature we are going to use in our project.

F. Food Image Recognition By Personalized Classifier In this model, the authors[12] applied some different and practical approach to food image recognition. since the datasets, Food-101 [13] and UECFOOD-256 [14] are having fixed classes but in the real world the food may differ between nationality and there may be inter class and intra class classification so which constitutes too many varieties of classes because of this they built a personalized classifier which combines Nearest Class Mean(NCM) Classifier and Nearest Neighbour (NN) classifier for each particular user because of the class imbalance problem. In these new classes can be added at nearly zero cost and the problem of food image variation can be avoided. The newly proposed time-dependent food distribution model and weight optimization algorithm to make personalized classifiers learn the user's data and adapt to users eating habits.

G. Food Calorie Measurement Using Deep Learning Neural Network Here authors[15] have proposed the Deep Learning Neural Network method handles the training and testing requests at the top layers, without affecting the central layers. Firstly the segmentation is done using Graph Cut segmentation followed by deep learning method. Here the user's thumb is used for size calibration. In this method, the first step is to generate a pre-trained model file and then the system is trained with a positive set of images. In the second step, they re-train the system with the set of negative images (images that do not contain the relevant object). In this system, once the model file is generated from the training, they load it into the application and test it against the images captured and submitted by the user. The system then performs the image recognition process and generates a list of probabilities against the label name. The label with the highest probability is prompted to the user in the dialog box, to confirm the object name. Once the object

name is confirmed, the system performs the calorie computation part by calculating the size of the food item concerning the thumb in the frame. It finally prints the output to the user with the required calorie.

H. Im2Calories: towards an automated mobile vision food diary In this proposed model the calories are calculated using two methods one method is using restaurant data here location and in which restaurant the user is eating is found out by using Google's Places API and then retrieve the menu of the nearest restaurant into a list, and either ask the user or perform a google search to retrieve the images of all the list items, Here they based their study on "MenuMatch" dataset which consists of 646 images, tagged with 41 food items, taken from 3 restaurants, and used GoogLeNet[2] classifier which has been pre-trained on ImageNet removed the last 1000- softmax and replaced it with 41 logistic nodes, and fine-tuned on the MenuMatch training set this gives an accuracy of test set as 81.4%, and gradually they extended the number of restaurants from MenuMatch dataset to 25 so that it may look like a practical application. Since as in the first approach collecting dataset for all the restaurants is a hard process so, they proposed another approach in which the calories are estimated using the food image not based on the restaurant where he is as mentioned above, so here they mainly used segmentation of the food image, here for segmentation they used 'DeepLab' system this model uses a CNN, after segmentation, the next step would be finding the volume here the model is trained on NYUv2RGBD dataset and then fine-tuned it on a new 3d food dataset they collected which they call the GFood3d dataset which consists of 50 different food videos. after knowing the volume the next step would be to find calorie's for the predicted volume this is done using the USDA National Nutrient Database (NNDDB)[18] after doing all this the calorie is obtained.

I. A Study of Food Recognition Techniques Accurate recognition of food from only an image can be a cumbersome task. The food items may be deformable, vary in the style of cooking or in the way they are garnished; this further makes it more complex to identify the food in the image. Candidate Region Detection is used to distinguish between multiple foods present in the image. The food is distinguished in one of the following ways: The whole Image - we assume that one image contains one food; Circle Detector - Identifies food based on shape, it extracts circular contours from the image; Region Segmentation - JSEG divides an image by color using a color class map. After each food is segmented, the features of the food are extracted using: SIFT and CSIFT; Histograms of Oriented Gradients (HoG); Gabor texture feature.

J. Fast YOLO: A Fast You Only Look Once System for Realtime Embedded Object Detection in Video The goal of object detection is to localize different objects in a scene and assign labels to the objects' bounding boxes. A standard

sliding window approach can be used where a classifier determines the existence of an object and its associated label for all possible windows in the scene. However, this type of approach has significant limitations in terms of not only high computational complexity but also high detection error rate. Recently, deep neural networks (DNNs) have shown superior performance in a range of different applications, with object detection being one of the key areas where DNNs have significantly outperformed existing approaches. In the Region-CNN (R-CNN) approach, a CNN architecture is used to generate bounding box proposals in an image instead of a sliding window approach, and thus a classifier only performs classification on bounding box proposals. Although R-CNN can produce state-of-the-art accuracy, the whole procedure is slow and difficult to optimize since each component must be trained individually. More recently, a You Only Look Once (YOLO) object detection approach was proposed that mitigated the computational complexity issues associated with RCNN by formulating the object detection problem as a single regression problem, where bounding box coordinates and class probabilities are computed at the same time. Although YOLO was demonstrated to provide significant speed advantages over R-CNN. For each video frame, an image stack consisting of the video frame with a reference frame is passed into a 1×1 convolutional layer. The result of the convolutional layer is a motion probability map, which is then fed into a motion adaptive inference module to determine if the deep inference is needed to compute an updated class probability map. YOLO is a neural network capable of detecting what is in an image in one pass. It gives the bounding boxes around the detected objects, and it can detect multiple objects at a time. It works by performing a regression - it predicts the bounding boxes and the class probabilities for each, doing so with a single network pass. Other approaches usually employ a pipeline of tasks, like passing on the image some classifier to detect objects in different locations and/or utilizing some other added methodologies.

3. CONCLUSION

In all the above systems, a method for measuring calories and nutrition of food object is carried out. The systems help people closely controlling their daily food intake. They focused on identifying food and extracting ingredients from the food image and predicting the calorie content from the quantity of food.

REFERENCES

- [1] Termritthikun, Chakrit, Paisarn Muneesawang, and Surachet Kanprachar. "NU- InNet: Thai food image recognition using convolutional neural networks on smartphone." *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* 9.2-6 (2017): 63-67.
- [2] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in*

- neural information processing systems. 2012.
- [3] Szegedy, Christian, et al. "Going deeper with convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
- [4] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [5] Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." arXiv preprint arXiv:1502.03167 (2015).
- [6] LeCun Yann, et al. "Gradient-based learning applied to document recognition." Proceedings of the IEEE 86.11 (1998): 2278-2324.
- [7] R. C. Gonzalez and R. E. Woods, Digital Image Processing: Prentice Hall, 2008.
- [8] Zhang, Weishan, et al. "Food image recognition with convolutional neural networks." 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). IEEE, 2015.
- [9] Salvador, Amaia, et al. "Inverse cooking: Recipe generation from food images." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019.
- [10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In CVPR, 2016.
- [11] Tammachat, Natta, and Natapon Pantuwong. "Calories analysis of food intake using image recognition." 2014 6th International Conference on Information Technology and Electrical Engineering (ICITEE). IEEE, 2014.
- [12] Yu, Qing, et al. "Food Image Recognition by Personalized Classifier." 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE, 2018.
- [13] Bossard, Lukas, Matthieu Guillaumin, and Luc Van Gool. "Food101—mining discriminative components with random forests." European Conference on Computer Vision. Springer, Cham, 2014.
- [14] Kawano, Yoshiyuki, and Keiji Yanai. "Automatic expansion of a food image dataset leveraging existing categories with domain adaptation." European Conference on Computer Vision. Springer, Cham, 2014.
- [15] Pouladzadeh, Parisa, Pallavi Kuhad, Sri Vijay Bharat Peddi, Abdulsalam Yassine, and Shervin Shirmohammadi. "Food Calorie Measurement Using Deep Learning Neural Network." 2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings, 2016, DOI:10.1109/i2mtc.2016.7520547, pp 1-6.
- [16] Meyers, Austin, et al. "Im2Calories: towards an automated mobile vision food diary." Proceedings of the IEEE International Conference on Computer Vision. 2015.
- [17] O. Beijbom, N. Joshi, D. Morris, S. Saponas, and S. Khullar. Menumatch: restaurant-specific food logging from images. In WACV, 2015.
- [18] USDA National Nutrient Database for Standard Reference, Release 27 (revised).
- [19] Y. Matsuda, H. Hoashi, and K. Yanai, "Recognition of multiple-food images by detecting candidate regions," in Multimedia and Expo (ICME), 2012 IEEE International Conference on. IEEE, 2012, pp. 25–30.
- [20] Liu, Chang, et al. "A new deep learning-based food recognition system for dietary assessment on an edge computing service infrastructure." IEEE Transactions on Services Computing 11.2 (2017): 249-261.
- [21] A Study of Food Recognition Techniques Suvarna Pansambal¹, Yamini Tawde², Chetali Surti³, Chhaya Patil⁴, Dhara Patel⁵,^{1,2,3,4,5} Department of Computer Engineering, Atharva College of Engineering, University of Mumbai, Mumbai-400095, India.
- [22] Fast YOLO: A Fast You Only Look Once System for Real-time Embedded Object Detection in Video, University of Waterloo, ON, Canada, Mohammad Javad Shafiee, Brendan Chywl, Francis Li, Alexander Wong.