Analysis & Numerical Simulation of Indian Food Image Classification Using Convolutional Neural Network

Anubhav Sharma[a], Chetan Kumar[b]
[a] *M.Tech. Scholar, Department of Computer Science &Engineering, KITE, Jaipur, India*
[b] *Associate Professor, Department of Computer Science & Engineering, KITE, Jaipur, India*
Anubhav5401@gmail.com[a], chetanmnit@yahoo.in[b]

Article History

Article Submission
26 December 2022
Revised Submission
02 February 2023
Article Accepted
05 March 2023
Article Published
30 March 2023

Abstract

Recognition of Indian food can be assumed to be a fine-grained visual task owing to recognition property of various food classes. It is therefore important to provide an optimized approach to segmentation and classification for different applications based on food recognition. Food computation mainly utilizes a computer science approach which needs food data from various data outlets like real-time images, social flat-forms, food journaling, food datasets etc, for different modalities. In order to consider Indian food images for a number of applications we need a proper analysis of food images with state-of-art-techniques. The appropriate segmentation and classification methods are required to forecast the relevant and upgraded analysis. As accurate segmentation lead to proper recognition and identification, in essence we have considered segmentation of food items from images. Considering the basic convolution neural network (CNN) model, there are edge and shape constraints that influence the outcome of segmentation on the edge side. Approaches that can solve the problem of edges need to be developed; an edge-adaptive As we have solved the problem of food segmentation with CNN, we also have difficulty in classifying food, which has been an important area for various types of applications. Food analysis is the primary component of health-related applications and is needed in our day to day life. It has the proficiency to directly predict the score function from image pixels, input layer to produce the tensor outputs and convolution layer is used for self- learning kernel through back-propagation. In this method, feature extraction and Max-Pooling is considered with multiple layers, and outputs are obtained using softmax functionality. The proposed implementation tests 92.89% accuracy by considering some data from yummly dataset and by own prepared dataset. Consequently, it is seen that some more improvement is needed in food image classification. We therefore consider the segmented feature of EA-CNN and concatenated it with the feature of our custom Inception-V3 to provide an optimized classification. It enhances the capacity of important features for further classification process. In extension we have considered south Indian food classes, with our own collected food image dataset and got 96.27% accuracy. The obtained accuracy for the considered dataset is very well in comparison with our foregoing method and state-of-the-art techniques..

Keywords- *CNN, SVM, Machine Learning, Segmentation, Image Processing.*

ISSN: 2250-0839

I. INTRODUCTION

The speedy development in our culture has given more attention towards the standard life, particularly to the food we eat. Manual classification of food, in some part, is not very appropriate in this speedy culture anymore. It is required to have an automation food recognition and classification system with optimized speed, accuracy and reduced cost of production. In past years, computer vision approaches become very much popular in food recognition. Commonly, there are two types of methods such as: conventional machine learning methodology which includes preprocessing of image, feature selection, feature extraction and classification. Another approach is deep learning that is very much popular in recent image recognition area. Like in [1], proposed an approach of bag-of-features (BoF) model through computation of the dense features, food images classification by linear support-vector-machine (SVM) and system was able to achieve 78% classification accuracy. In [2], proposed a Deep CNN (DCNN) algorithm and for training used a dataset of ImageNet LSVRC2010, they enable to get huge improvement of 84.7% with respect to conventional machine learning approach. Similarly, several efforts have been made in research field of food class prediction, whereas to extract information of food in a particular image is challenging role.

In this research, the proposed work is entirely focused on convolution neural network (CNN), steps are split as follows: image acquisition, dataset planning, network structure, training phase and testing phase. The first main aim is to collect the real time food images, which can be collected via camera, mobile camera, food blog website and other internet resources. Most classes are South Indian cuisine since very less dataset is available on the open source platform as per our study. Subsequently, CNN based Inception-V3 architecture is considered as base of our model because of its capability to get important features that can be utilized in process of food classification.

In food-related research, food computation mainly utilizes computer science approaches. It needs food data from various data outlets (e.g. social platform, internet of things and recipe distribution websites) for different modalities like images of food, food records, recipe, taste and their aroma to be acquired and checked. Such activities need research that rely on computer vision, artificial learning, data mining and other emerging food and human interaction technology in order to facilitate human-centered behaviors, such as nutritional enhancement, directing human actions, and understanding human culture.

In a historically multidisciplinary area, Computer technology meets traditional fields associated with food like food science, medication, genetics, agrology, sociology and culinary science. Therefore, food computation borrows ideas and techniques from other fields, such as neurobiology, perception science and chemistry, in addition to computer science. Food computation primarily consists of five foundational functions, ranging from awareness, identification, retrieval, projection and monitoring recommendations. It also makes different uses, being fitness, tradition, agriculture and medicine, for many fields. Computation methods for the acquisition and analysis of diverse food data from various resources are therefore used in food computing to resolve food-related questions in health, biology, agronomy etc.

The provision of various human-centric facilities is one significant objective of food computing. Gathering food data is considered to be the first step; from various data sources, like cookbooks, social platforms, and web based recipe-sharing, in different forms. Furthermore, other unique food databases, like the aroma threshold database and the Volatile Organic Compounds in the food data base, are also available. On the basis of food database, we use machine learning, computer vision, data mining, and numerous additional food data processing techniques. Food experiences are multimodal with sensory information such as taste, aroma and sensation of touch. The taste and sense perception of that food will control our food appetite and influence what extent we eat and drink.

One specific task is identification and it is primarily used to predict food products from food images, such as the category or ingredients. Single modal and cross modal retrieval, such as visual food extraction and recipe fetching with food-oriented retrieval are more important in applications like the extraction of food image recipes. This advice can not only prescribe foodie people what they can eat, how they can also help them for a healthy diet. Consuming food guidelines often require well nuanced and composite details.

Prediction and monitoring is primarily carried out on the basis of social networking, such as public health surveillance. These various activities are not autonomous, but are closely intertwined and relying on one another. The recognized outcomes will further help the tasks of retrieval and advice. Most of the work is also based on the technology of food perception recognition [3]. When the food image categories are large, it is also possible to use the retrieval- based approach for food recognition. Social media prediction may also be useful for the suggestion task. For instance, important information for personalized food recommendation would be the food preference of users predicted from social media.

II. RESEARCH OBJECTIVE

On the basis of problem formulation our main objective is to segmentation, recognition and analysis of Indian food images for finding solutions to the health applications using state-of- the-art techniques. To achieve the above objectives followings are considered:

1. Creating a dataset of images containing Indian food items.
2. Segmentation of food item from the acquired Indian food image.
3. Classification based approach for recognition of food items using the segmented features.
4. Estimation of total calories information from the recognized food classification for dietary advice application.
5. Analysis of the food items for sufficiency of cooking and baking.

III. LITERATURE REVIEW

Recognition and analysis systems for food items are an attractive and valuable field of research as they allow quantitative food behavior measurements. In many dietary-related contexts, this role is helpful and welcomes, particularly to manage health states or assessing the study of eating habits. Among the most critical activities undertaken for automatic dietary tracking involves food image processing. Following are the brief survey done in the area of food data collection, image segmentation and classification with different approaches:

[1] Proposed standard database to learn the image representation in food applications. The primary food image database with 800 plates of food, UNICT-FD889 dataset, is introduced based on comparing Bag of Textons, PRICoLBP and SIFT label methods. Texture and color properties are considered for obtaining accurate outputs.

[2] Presented extensive food images database, i.e AIFood, for food recognition based on ingredients. Each image is labeled using available information with food name or ingredient list and done manual inspection of food images to find out unknown ingredients information and re- label them.

[3] Presents an effort to construct extensive food image dataset, with food class selection, data collection and clean, and labeling. They shared a slandered method with deep convolutional neural network (CNN) on ChineseFoodNet. Method considers two-step data fusion called ―TastyNet. They achieved top-1 accuracies of 81.43% on the validation set and 81.55% on the test set.

[4] Proposed an image segmentation process to segment food and drinks items from meal images used for diet computation. They used ―Snakes or active contours‖ method to isolate object borders and segmenting images. Different methods of edge initialization and integrated context removal methods were discussed for improving the outputs of segmentation.

[5] Proposed segmentation method found on normalized cutting and super pixels. The methods rely on ―color and texture‖ suggestion with speedy computing and sufficient use of memory. The approached work used for segmenting food imagery with mobile recording system that has built for diet analysis and control. Using the Berkeley Segmentation Dataset, the approach achieves competitive results and outperforms some of the most common food image dataset techniques.

[6] Proposed experimental approaches for automated segmenting and identification of multiple food images for type 1 diabetic patient with glucose and insulin counting. Initially pyramidal―mean-shift filtering‖ and

―region-growing‖ techniques were used to plate segmenting. Then each of resulting segments are characterized by the features and graded into one of six distinct major food groups by a support vector machine. A new method was proposed, and experimental findings demonstrated the efficacy of 88.5% for segmentation accuracy and 87% for recognition rate.

[7] Focused on food image processing and a grouping method with k-nearest neighbors, glossary trees characteristics and their combinations. With 1453 food images dataset, eating occasions of 42 groups were collected by 45 members in normal eating environments and methods were tested. Experimental findings showed that the use of feature and glossary hierarchy for classification, which increases the classification efficiency by around 22% of the Top-1 and 10% of the Top-4 values.

IV. PROPOSED METHODOLOGY

The inactivity of most people may be traced back to their common practice of overeating. Research on accurate food classification is becoming increasingly important since people's increasingly hectic and stressful lifestyles make it harder for them to maintain healthy eating habits. The usage of intelligent apps for accurate food analysis using deep learning is a major development in today's technology.

Additionally, these apps are effective in keeping the user's eating habits in check and providing warnings when unhealthy options are recognized. Training a deep learning model now takes a reasonable amount of time due to advancements in memory and processing technology. In addition, the data these constructed models can handle is current.

Naive Bayes, Support vector machine, Adaboost, and Artificial Neural Networks (ANNs) are just a few of the classifiers that have been employed in the classification process. To further boost the recognition rate during the food identification process, a framework based on pair-wise categorization was presented. The bag of features (BoF'), which is developed from the bag of words, is another widely used classifier in NLP. The program is programmed to learn high-frequency words by ignoring the order in which they often appear. Names and food photos are often paired on social media. Furthermore, the food photographs have a common visual pattern that is utilized to foretell the shape of the meal and reduces process complexity difficulties by employing a direct image-matching approach. As a result, various works have been discovered which employ BoF methodology. In addition, a similar procedure of BoF models has been carried out using texton based histograms. It was determined, however, that the BoF approach has a limited capacity for data transmission and fails under high-resolution imaging conditions.

However, in order to adapt the system to varying lighting circumstances, the checker-board method is used for color capture. As the number of classes grows, the performance accuracy tends to deteriorate. However, the database is seen as a crucial feature for the food classification assignment, and the model should be able to deal with this in a real-world setting. From there, they employed the SIFT (Scale invariant feature transform) function and tested with the seven classes, all while creating a real-time library of food photographs for future trials to serve as a benchmark. There was a clear performance gap between the fake food photographs and the genuine ones; this might be due to the different types of images captured or their larger file sizes. However, when the number of classes is small and there are several photos available, employing the SIFT features derived from the starter-food and foods in shows superior results.

Indian cuisine identification often requires a high degree of fine-grained picture perception, whereas the linked category will have significant variances at the recognition point. Forecasting the relevant and improved analysis then calls for the proper segmentation and classification procedures. No matter what they're called on the menu, the hotel or restaurant's dishes always have a wide range of ingredients. Food product identification can be impacted by factors such as the dish's presentation in different contexts (different restaurants, cooking methods, and methods of delivery). Because of the presenting style and the unusual recipes, the long-tail is strengthened, and the total complexity and recognition mechanism may be created using a very little training sample.

In most cases, spectral data alone is used in the segmentation procedure, leading to unwanted information variability and subpar model output for the same class labels. Given the difficulties encountered in actually

ISSN: 2250-0839

classifying food photos, segmentation is a viable option worth considering. When employing multispectral knowledge to construct a model, segmentation proves to be an efficient method due to the wealth of information it provides on the component, size, and shape, among other crucial features.

Recent advances in multilayer neural networks have the potential to make the training and testing of a wide variety of machine learning tasks, including classification and regression issues, possible on their own. The deep network model has the potential to improve classification accuracy over more traditional classifiers since it can extract more abstract data attributes. Some deep learning approaches often employ a convolutional neural network to identify and flag problems in images at the pixel level. Taking a CNN model into account makes it easier to build up features correctly, which leads to more accurate labeling.

The classification result maps are smoother with CNN than with certain indirect simulation methods because of the efficient means by which the structural properties beneath the neighboring input data patches can be evaluated. In comparison to other methods of machine learning, CNN's segmentation findings are less refined because of the model's inability to account for important factors like form and edge. Noise introduced during the CNN training procedure on the input side might lead to gaps in the classification mapping. It is motivated by the absence of constraints such as shape and edge, and its performance is typically erratic in comparison to that of other machine learning methods. Segmentation in CNN requires a post-processing phase due to this limitation in order to achieve optimal performance. In this study, we propose an Edge Adaptive (EA) method for modeling crucial features such as pixel connectivity, shape, region of interest, and other specifics. Training metrics may take use of any available spatial data relationship frameworks thanks to the direct construction of the EA from the DCNN method. Therefore, the EA can aid in optimizing efficiency once functions have been obtained from a DCNN. For computer programs, it is crucial to train the model with a high projected value or else it would fail to operate on test data due to over-fitting. Over-fitting may be mitigated by increasing the size of the train data set. In our technique, we subdivide all input pictures into smaller pieces we call "patches," which are then used to more accurately and efficiently depict the underlying structure. Initially, the deep CNN model is used to create a preliminary map function, and then the edge adaptive is run to provide the final segmentation results in the implemented EA-based DCNN model. To maximize better food picture segmentation, the EADCNN train model uses "convolution, rectified linear unit (ReLu) and pooling" to extract the relevant extent. We evaluated several assessment aspects in the outcome review section with the ground truth data to ascertain the efficacy of our constructed model.

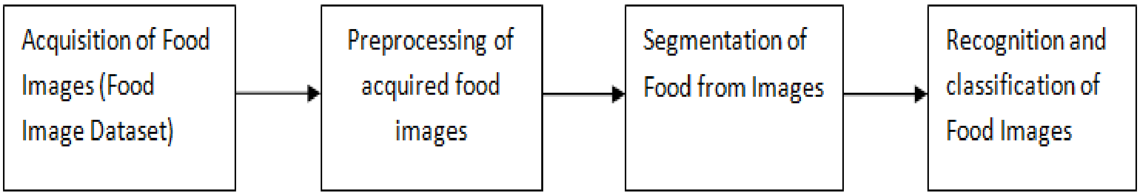
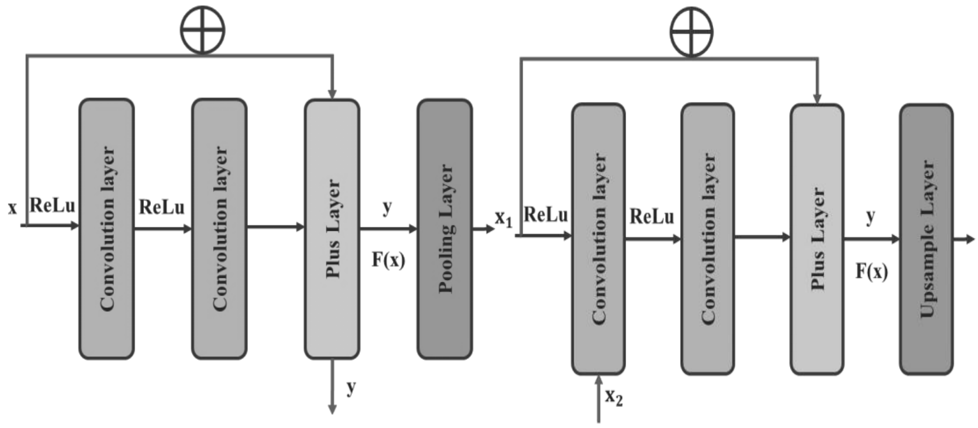


Figure 4.1: Block Diagram of Proposed Methodology

In this context, a full-stack deep convolutional neural network is taken into account, with a three-channel (one channel per color) input picture yielding binary segmentation maps as the final output. Here, we transform the neural network's output into a low-dimensional class issue using a final softmax function and a second convolutional layer. This aids the overlap tile approach in its ability to divide up big pictures without losing sight of the fundamental building blocks.

The outline of both the down and up sampling designs are suggested by the framework. The uniform sampling model is shown in Figure 4.2 for both the down and up sampling blocks. In this case, we combine two convolution layers using a ReLu layer. Adding a plus layer and then several convolution layers helps to streamline the design and tackle the error loss problem that arises when the network goes very deep. The effectiveness of the deep learning algorithm in the dynamic phase of picture segmentation may be attributed to the fact that the strategy used makes it possible to prevent the convergence of the training stage for the best local solution. Since y is the expected output of the plus layer and x the expected input of the down-block, the plus layer will grant permission for x and supply the max pooling layer with the optimal result of layer y.



(a) (b)

Figure 4.2: Design of (a) Down-Sampling and (b) Up-sampling blocks

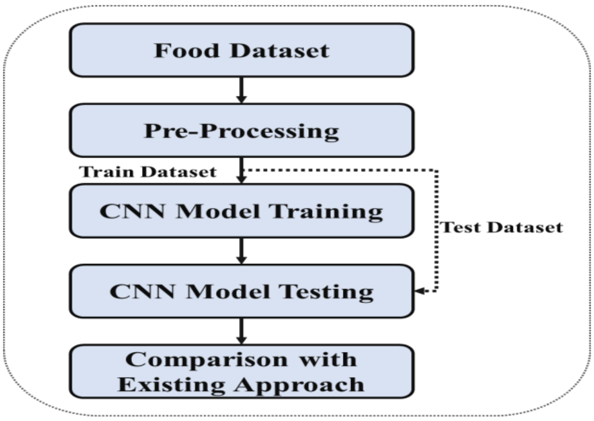


Figure 4.3 Proposed Scheme for Food Images Classification

The suggested method for classifying food photographs is shown in Figure 4.2, and further explanations of each block that was taken into account are provided below. In this article, we discussed methods for gathering, cleaning, and training a CNN on data pertaining to food. The resultant trained model may then be used to the test dataset in order to categorize the photos of food.

We have taken into account data from the Yummly dataset as well as real-time south-Indian cuisine photos. For the training and testing phases, we resized the photos to 299 by 299 pixels and removed noise, variations in color intensity, and incorrect labeling from the images.

Figure 4.4 depicts the suggested categorization model for food images. We've already completed the first stage of the dataset gathering process, and the bulk of the classes we've gathered are related to South Indian food. All of the food photos have been organized by kind and placed in their own folders. Since all of the photos' resolution/size was quite different to each other, the size of the photographs is deemed to be colored images before processing continues.

ISSN: 2250-0839

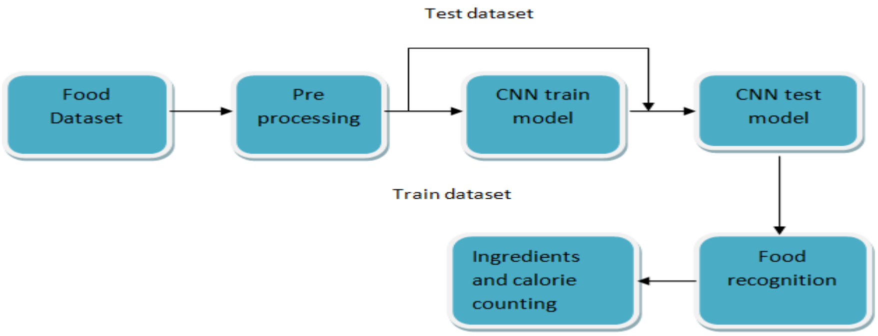


Figure 4.4: Proposed Approach based on Inception V3 Model

Pre-processing involves choosing a specific size of food, and it has been shown that the bulk of data photographs come from online sources, which results in lower image resolution compared to self-taken food images. To ensure that all images contribute to the model-training process, it was necessary to account for the general size of images while yet preserving the relationships between neighboring pixels. If the image size is smaller than the specified size, it must be resized; if it is larger, it must be cropped to the specified size. Additionally, we have a pre-processing stage, a CNN model training stage, and a classification stage, where unseen test data is used in the classification process.

V. RESULTS AND DISCUSSIONS

Multiple models may be saved during an assessment, and then later on, those models can be loaded and compared for best accuracy and least loss. In addition to the results from the CNNs, we also get a confusion matrix.

Each class label may be plotted in the confusion matrix, revealing the proportion of right predictions relative to false ones. Accordingly, numerous crops have been evaluated rather than a single value to validate the test dataset, which improves accuracy when compared to a strategy that only evaluates one crop at a time. The judgment is derived from the test dataset's crops, with the result formed from the top projection for each unique crop, which was then utilized to offer the top five projections.

Table 5.1

Food Classes Considered for Proposed Work

Food class	Food_class_name	Food class	F o o d _class_name
1	aloo_parathas	9	kesribath_upma
2	bhindi_masala	10	onion_rings
3	dal	11	paneer_masala
4	dosa	12	ragi_dosa
5	fried rice	13	samosa
6	idly	14	upma
7	idly Wada	15	Wada
8	kesribath		

Therefore, projections are obtained for each individual picture during the processing step, and the mapping technique is then employed for the map test element index. Table 5.1 lists the fifteen distinct categories that can be applied to photographs of food. Figure 5.1 displays a few example food photographs, and the model was evaluated using 15 randomly selected food images. Many transformations, such as scaling, rotation, and cropping, have been applied to the training photos, greatly increasing their total count.

ISSN: 2250-0839

In this case, we take into account the EA segmentation model's resulting segmented features (SF) as an input to the classifier. Inputs to the model are: - unique in the 'ImageNet' or random in pre-training on the ImageNet, and The include top option specifies whether the network's topmost layer should also include the "fully connected layer." Shape of input is (299 x 299 x 3) (i.e., 'channels_last' format of data) or (3 x 299 x 299) (i.e., 'channels first' format of data) unless the include top is considered as False, in which case the input shape: must be given. In this case, the pooling option may be considered if the top is False parameter is included in the feature extraction, and if the specified 3 input channels have the necessary height and width.

With the pooling function option set to "None," the model's predictions from the previous convolutional layer are output as 4D tensors. The model's output will be a two-dimensional tensor if the pooling function is assumed to be "average," which means that the global average pooling function is selected based on the result of the previous convolution layer. Specifying "max" as the argument triggers the global-max pooling function. The 'classes' function also provides the number of classes that may be used to categorize the photos if the include_top parameter is set to true and no weights are provided. At the end, it provides a new instance of the Keras Model class. Accuracy of the system is determined by:

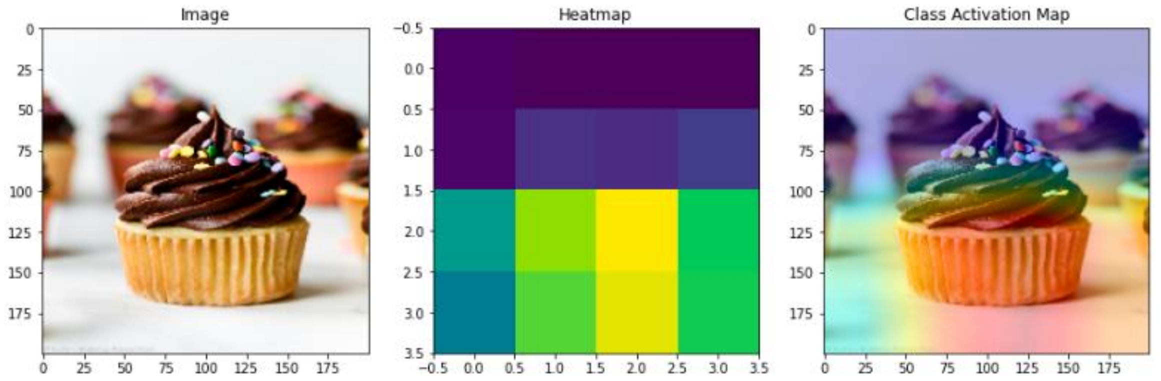


Figure 5.1: Pre-processed Food Image at Different Angle and Position

Our suggested model of idly images from a variety of angles and orientations is used to preprocess (data augment) images of food, as shown in Figure 5.1. Table 5.2 compares the accuracy of the current method with that of many other models, including Naive Bayes, Random Forest, Multinomial Logistic Regression, and Linear SVC, all of which have been evaluated using data from sites like Food.com, Yummly, and Epicurious. We put our suggested model through its paces on a variety of cuisines, and it achieved a 93% success rate in classifying them.

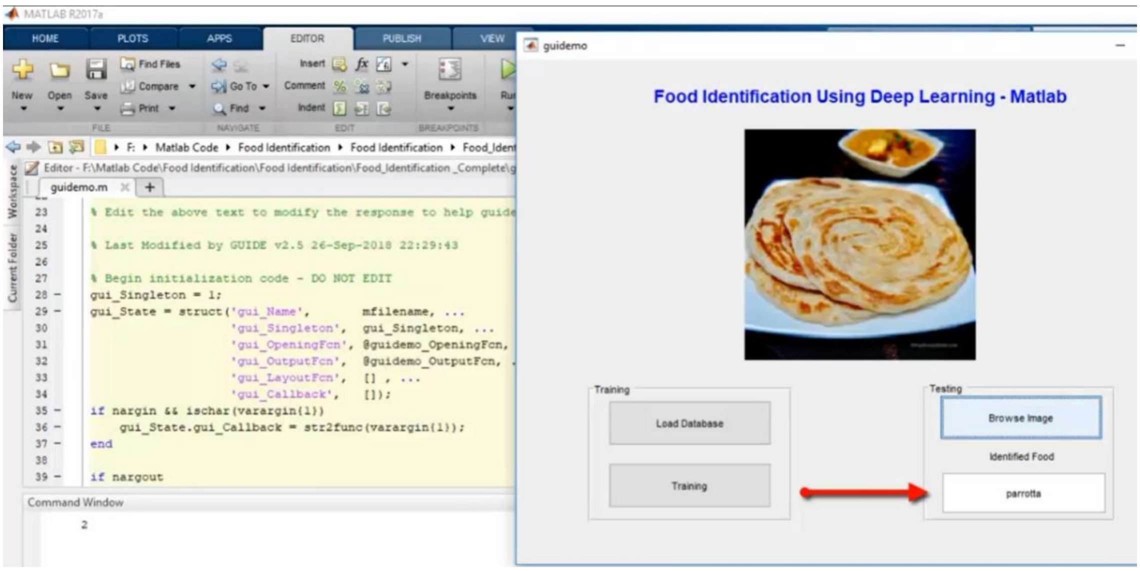


Figure 5.2: Design of Graphical User Interface of the Proposed System

ISSN: 2250-0839

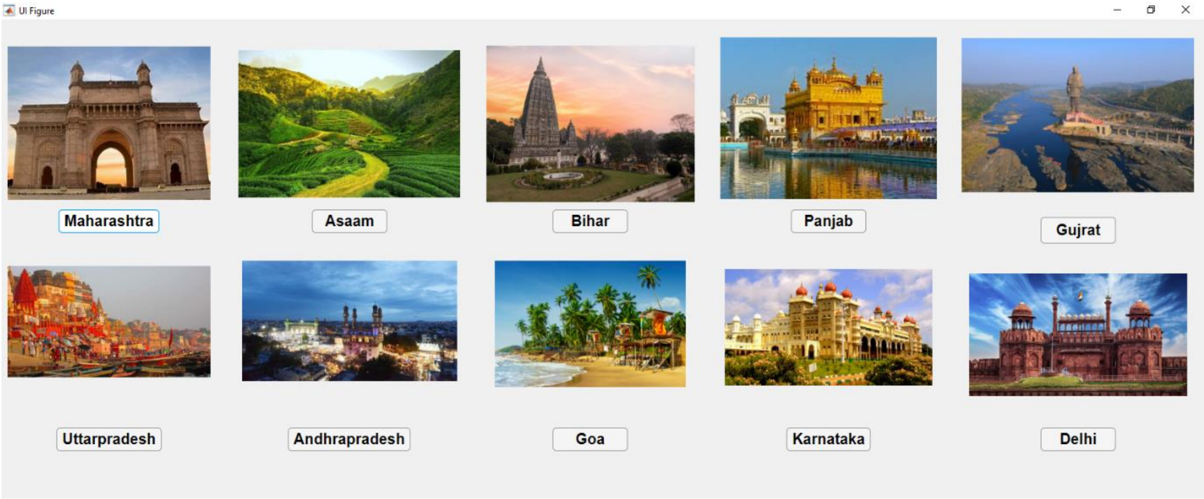


Figure 5.3: Data of Food from Selected State

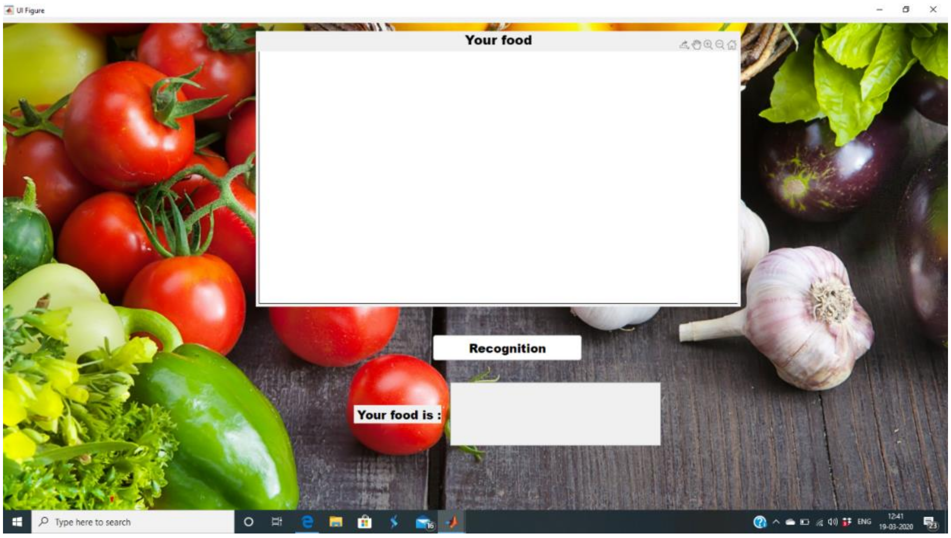


Figure 5.4: Design of Food Identification System

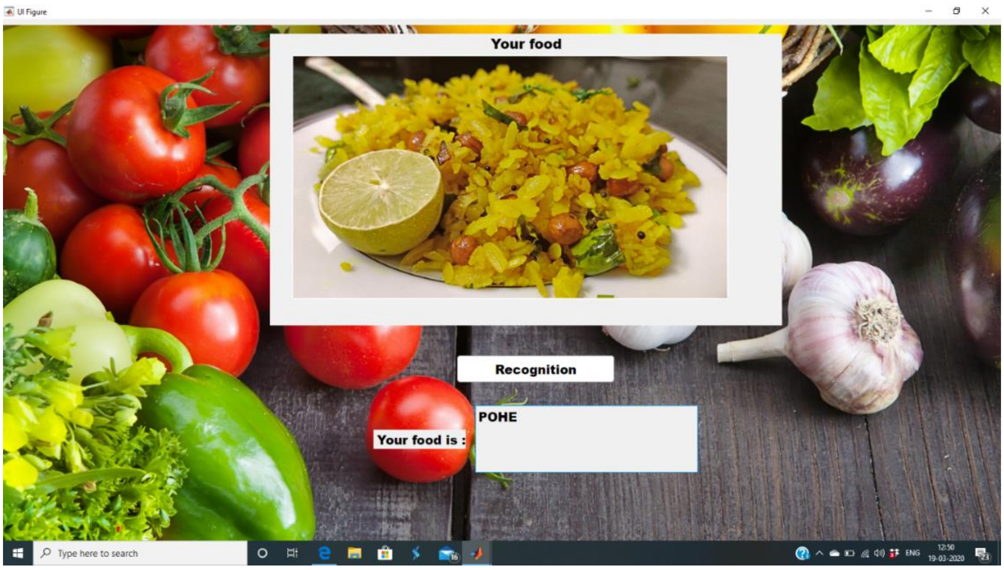


Figure 5.5: Processing of Proposed Methodology

ISSN: 2250-0839

Figure 5.6: Detection of Food Images by Proposed Methodology

Table 5.2

Classification Accuracy Comparison with the Existing Approaches

Different Classification Approaches	**Classification Accuracy (%)**
Naive Bayes	71.4
Random Forest	75.2
Multinomial Logistic Regression	77.7
Linear SVC	78
Proposed CNN without SF	82.18
Proposed CNN with SF	91.79

ISSN: 2250-0839

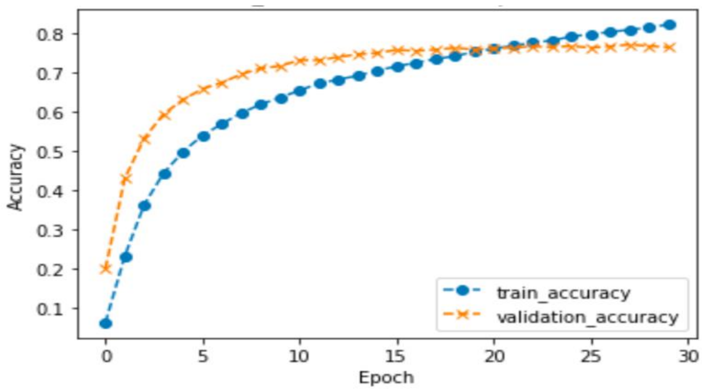


Figure 5.7: Analysis of Accuracy by Proposed Methodology

The detailed information of classification has given through the confusion matrix for 15 classes of foods as shown in figure 5.3. Table 5.2 represents top1 (%), top5 (%) and top10 (%) result of classification accuracy and we have used 32 epochs in training model. Training accuracy with respect to number of epochs is shown.

Table 5.3

Top result of classification accuracy with proposed CNN

Top picks	Accuracy of Proposed CNN
Top-1(%)	88.14
Top-5(%)	90.12
Top-10(%)	91.79

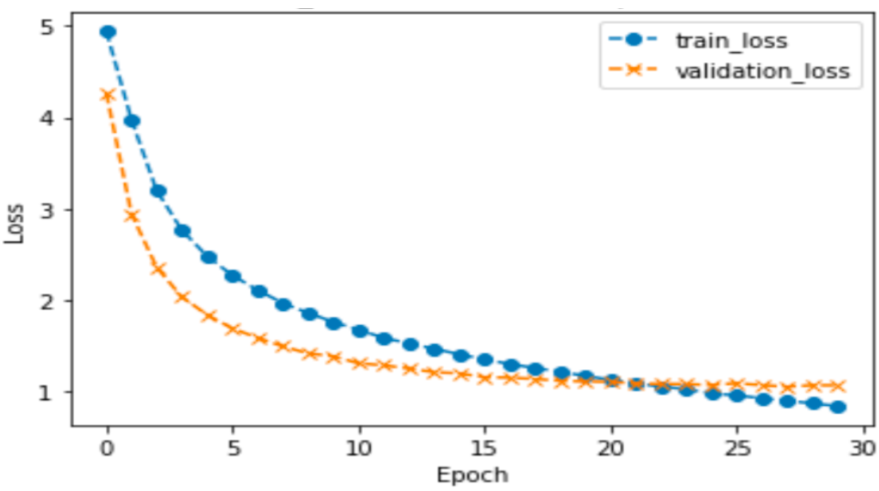


Figure 5.8: Analysis of Loss by Proposed Methodology

We evaluate the effectiveness of our suggested model and talk about its results and analysis in the context of food classification. Our model was trained using thousands of photos, Nvidia Pascal Titan GPU, batch size of 128, and 100 iterations totaling 1719 images. Tensor objects with a float 32 representation were created from the input. In this project, Python has been examined for use in the creation of models. All of our tests have been conducted on a completely separate PC running Anaconda Prompt. Multiple models are saved as part of the assessment process, allowing for fast loading times with little data loss and maximum precision. Using many crops for validation of the test set rather than a single value consistently increases accuracy compared to a single-crop based method. The correct label prediction vs the false label prediction for the various classes are seen in the confusion matrix, which is also created to plot each class label. Here, we assume a total of 16 classes, as shown in Table 5.1, containing tens of thousands of photos; 20% of these images are utilized for testing (i.e., randomly picked from all classes), while the remaining 80% are used for training.

ISSN: 2250-0839

Because of this, the mapping technique is used to map the index of test element to gain the best estimates for each unique image, such as shown in the figure.

After using our suggested model to our dataset, we achieved a classification accuracy of 91.79 percent. The figure presents the percentage accuracy in classifying cooked vs uncooked food items from the artificial dataset.

We first began amassing real-time food photographs from a wide range of online and offline sources, including cameras, mobile phones, food blog websites, and more. The south Indian cuisine is the most represented style in the collection. Then, we improve our model's capacity to swiftly get crucial features by training a bespoke CNN-based model on tens of thousands of food photos using the most robust weights from imageNet. The results show that after 100 training epochs, the model's accuracy is high enough for use in testing. During testing, the mapping technique is used to make the most accurate predictions possible for each particular image by mapping the index of each test element. As a consequence, our testing yielded a classification accuracy of 96.27 percent, which is significantly higher than that of previous methods. We find that the custom CNNs model we provide is the best option for identifying dishes in food photographs. It accurately distinguishes between correctly supported, less backed, and over backed, and it can estimate calories and ingredients for us. More progress can be made in the future by employing various frameworks as TFOD, YOLO, Dectoron2, etc.

VI. CONCLUSION & FUTURE SCOPE

Image segmentation is a common technique used in many programs that rely on vision. There is no agreed-upon approach for choosing a segmentation algorithm or comparing the results of different procedures. Because of this defect, incorrect conclusions may be drawn and/or unintended results may occur. Again, the lack of a concrete interpretation has shown to be troublesome for image segmentation. In computer graphics, segmentation is the process of dividing a large collection of pixels into smaller ones. Some interpretations of the literature may share this idea, although the criteria used to evaluate it might be debated. Humans use a process of interpretation called segmentation, and one example of this is the recognition of patterns. The magnitude of the issue grows dramatically when segmentation is considered.

We started by collecting photos of food taken recently using cameras, smartphones, food blogs, and other online resources. The bulk of the assembled classes favor south Indian food. Next, we leverage the top imageNet weights to train a custom CNN-based model on tens of thousands of food images, enhancing our model's ability to quickly detect essential traits. By analyzing the outcomes, we find that the model performs well in both the training and testing phases. In order to optimize the model's performance during testing, we employ a mapping strategy to execute the separate image predictions by mapping the index of the test element to acquire the best estimates. When compared to prior approaches, our 91.87% accuracy in testing is a significant improvement. It's possible that our proposed custom CNNs model would significantly enhance food image classification. It can assess calorie content, provide a list of components, and determine if a food has sufficient, insufficient, or excessive support. Improvements can be made in the future by employing alternative frameworks like TFOD, YOLO, Dectoron2, and others.

REFERENCES

[1] Giovanni Maria Farinella, Dario Allegra, F. Stanco, A Benchmark Dataset to Study the Representation of Food Images, Workshop on Assistive Computer Vision and Robotics (ACVR 2014) Volume: 8927, September 2015, pp 584-599, DOI: 10.1007/978-3-319-16199-0_41.

[2] GwoGiun Lee; Chin-Wei Huang; Jia-Hong Chen; Shih-Yu Chen; Hsiu-Ling Chen, AIFood: A Large Scale Food Images Dataset for Ingredient Recognition, IEEE Region 10 Conference (TENCON), 2019, DOI: 10.1109/TENCON.2019.8929715

[3] XinChen , Yu Zhu , Hua Zhou, Liang Diao, and Dongyan Wang , ChineseFoodNet: A Large-scale Image Dataset for Chinese Food Recognition, arXiv:1705.02743v3 [cs.CV] 15 Oct 2017.

[4] Y. He, N. Khanna, C. J. Boushey and E. J. Delp, Snakes assisted food image segmentation, 2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP), Banff, AB, 2012, pp. 181-185,

ISSN: 2250-0839

doi: 10.1109/MMSP.2012.6343437.

[5] Y. Wang, C. Liu, F. Zhu, C. J. Boushey and E. J. Delp, Efficient superpixel based segmentation for food image analysis, 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, 2016, pp. 2544-2548, doi: 10.1109/ICIP.2016.7532818.

[6] M. Anthimopoulos, J. Dehais, P. Diem and S. Mougiakakou, Segmentation and recognition of multi-food meal images for carbohydrate counting, 13th IEEE International Conference on BioInformatics and BioEngineering, Chania, 2013, pp. 1-4, doi: 10.1109/BIBE.2013.6701608.

[7] Y. He, C. Xu, N. Khanna, C. J. Boushey and E. J. Delp, Analysis of food images: Features and classification, 2014 IEEE International Conference on Image Processing (ICIP), Paris, 2014, pp. 2744-2748, doi: 10.1109/ICIP.2014.7025555.

[8] Y. He, N. Khanna, C. J. Boushey and E. J. Delp, Image segmentation for image-based dietary assessment: A comparative study, International Symposium on Signals, Circuits and Systems ISSCS2013, Iasi, 2013, pp. 1-4, doi: 10.1109/ISSCS.2013.6651268.

[9] S. Aslan, G. Ciocca and R. Schettini, Semantic segmentation of food images for automatic dietary monitoring, 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp. 1-4, doi: 10.1109/SIU.2018.8404824.

[10] Y. He, C. Xu, N. Khanna, C. J. Boushey and E. J. Delp, Food image analysis: Segmentation, identification and weight estimation, 2013 IEEE International Conference on Multimedia and Expo (ICME), San Jose, CA, 2013, pp. 1-6, doi: 10.1109/ICME.2013.6607548.

[11] M. Bolaños and P. Radeva, Simultaneous food localization and recognition, 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, 2016, pp. 3140-3145, doi: 10.1109/ICPR.2016.7900117.

[12] StergiosChristodoulidis, MariosAnthimopoulos and StavroulaMougiakakou, Food Recognition for Dietary Assessment Using Deep Convolutional Neural Networks, ICIAP 2015: New Trends in Image Analysis and Processing -- ICIAP 2015 Workshops pp 458- 465

[13] Joachim Dehais, MariosAnthimopoulos and StavroulaMougiakakou, Dish Detection and Segmentation for Dietary Assessment on Smartphones, ICIAP 2015: New Trends in Image Analysis and Processing -- ICIAP 2015 Workshops pp 433-440

[14] J. Lee, A. Banerjee and S. K. S. Gupta, MT-Diet: Automated smartphone based diet assessment with infrared images, 2016 IEEE International Conference on Pervasive Computing and Communications (PerCom), Sydney, NSW, 2016, pp. 1-6, doi: 10.1109/PERCOM.2016.7456506.

[15] ElnazJahani, Heravi H.H and Aghdam D.P, An optimized convolutional neural network with bottleneck and spatial pyramid pooling layers for classification of foods Pattern Recognition Letters, Volume 105, 1 April 2018, Pages 50-58

[16] B. S. Anami, Dayanand G. Savakar,Effect of Foreign Bodies on Recognition and Classification of Bulk Food Grains Image Samples, Journal of Applied Computer Science andMathematics, Volume 3(6), Pages: 77-83, 2019.

[17] S. J. Minija and W. R. S. Emmanuel, Food image classification using sphere shaped Support vector machine, 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, 2017, pp. 109-113, doi: 10.1109/ICICI.2017.8365314.

[18] Aslan S., Ciocca G., Schettini R. (2017) On Comparing Color Spaces for Food Segmentation. In: Battiato S., Farinella G., Leo M., Gallo G. (eds) New Trends in Image Analysis and Processing – ICIAP 2017. ICIAP 2017.

[19] Knez S., Šajn L. (2015) Food Object Recognition Using a Mobile Device: State of the Art. In: Murino V., Puppo E., Sona D., Cristani M., Sansone C. (eds) New Trends in Image Analysis and Processing -- ICIAP 2015 Workshops. ICIAP 2015.

[20] R. Dinic and T. Stütz, EatAR Tango: Results on the Accuracy of Portion Estimation, 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR- Adjunct), Nantes, 2017, pp. 284-287, doi: 10.1109/ISMAR-Adjunct.2017.90.

[21] J. Sun, K. Radecka and Z. Zilic, Exploring Better Food Detection via Transfer Learning, 2019 16th

ISSN: 2250-0839

International Conference on Machine Vision Applications (MVA), Tokyo, Japan, 2019, pp. 1-6, doi: 10.23919/MVA.2019.8757886.

[22] G. Su, D. Li, Y. Wang, L. Wang and M. Zhang, Chinese Dish Segmentation Based on Local Variation Driven Superpixel Grouping and Region Analysis, 2018 International Conference on Audio, Language and Image Processing (ICALIP), Shanghai, 2018, pp. 41-45, doi: 10.1109/ICALIP.2018.8455779.

[23] M. Chiang, C. Wu, J. Feng, C. Fang and S. Chen, Food Calorie and Nutrition Analysis System based on Mask R-CNN, 2019 IEEE 5th International Conference on Computer and Communications (ICCC), Chengdu, China, 2019, pp. 1721-1728, doi: 10.1109/ICCC47050.2019.9064257.

[24] F. Konstantakopoulos, E. I. Georga, K. Klampanas, D. Rouvalis, N. Ioannou and D. I. Fotiadis, Automatic Estimation of the Nutritional Composition of Foods as Part of the GlucoseML Type 1 Diabetes Self-Management System, 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), Athens, Greece, 2019, pp. 470-473, doi: 10.1109/BIBE.2019.00091.

[25] Wang X., Rohani N., Manerikar A., Katsagellos A., Cossairt O., Alshurafa N. (2017) Distinguishing Nigerian Food Items and Calorie Content with Hyperspectral Imaging. In: Battiato S., Farinella G., Leo M., Gallo G. (eds) New Trends in Image Analysis and Processing – ICIAP 2017. ICIAP 2017.

[26] Y. Lu, T. Stathopoulou, M. F. Vasiloglou, S. Christodoulidis, Z. Stanga and S. Mougiakakou, An Artificial Intelligence-Based System to Assess Nutrient Intake for Hospitalised Patients, IEEE Transactions on Multimedia, doi: 10.1109/TMM.2020.2993948.

[27] Aslan, S., Ciocca, G., Mazzini, D. et al. Benchmarking algorithms for food localization and semantic segmentation. Int. J. Mach. Learn. & Cyber. (2020).

[28] Y. Wang, F. Zhu, C. J. Boushey and E. J. Delp, Weakly supervised food image segmentation using class activation maps, 2017 IEEE International Conference on Image Processing (ICIP), Beijing, 2017, pp. 1277-1281, doi: 10.1109/ICIP.2017.8296487.

[29] P. Pouladzadeh, S. Shirmohammadi and A. Yassine, Using graph cut segmentation for food calorie measurement, 2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Lisboa, 2014, pp. 1-6, doi: 10.1109/MeMeA.2014.6860137.

[30] P. Pouladzadeh, P. Kuhad, S. V. B. Peddi, A. Yassine and S. Shirmohammadi, Food calorie measurement using deep learning neural network, 2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings, Taipei, 2016, pp. 1-6, doi: 10.1109/I2MTC.2016.7520547.

[31] Sri Vijay Bharat Peddi, PallaviKuhad and AbdulsalamYassine, An intelligent cloud- based data processing broker for mobile e-health multimedia applications, Future Generation Computer Systems, Volume 66, January 2017, Pages 71-86

[32] Pouladzadeh P., Yassine A., Shirmohammadi S. (2015) FooDD: Food Detection Dataset for Calorie Measurement Using Food Images. In: Murino V., Puppo E., Sona D., Cristani M., Sansone C. (eds) New Trends in Image Analysis and Processing -- ICIAP 2015 Workshops. ICIAP 2015.

[33] P. Kuhad, A. Yassine and S. Shimohammadi, Using distance estimation and deep learning to simplify calibration in food calorie measurement, 2015 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), Shenzhen, 2015, pp. 1-6, doi: 10.1109/CIVEMSA.2015.7158594.

[34] M. T. Islam, B. M. N. KarimSiddique, S. Rahman and T. Jabid, Food Image Classification with Convolutional Neural Network, 2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), Bangkok, 2018, pp. 257- 262, doi: 10.1109/ICIIBMS.2018.8550005.

[35] A. Metwalli, W. Shen and C. Q. Wu, Food Image Recognition Based on Densely Connected Convolutional Neural Networks, 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Fukuoka, Japan, 2020, pp. 027-032, doi: 10.1109/ICAIIC48513.2020.9065281.

[36] Q. Yu, M. Anzawa, S. Amano, M. Ogawa and K. Aizawa, Food Image Recognition by Personalized Classifier, 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, 2018, pp. 171-175, doi: 10.1109/ICIP.2018.8451422.

ISSN: 2250-0839

[37] A. Şengür, Y. Akbulut and Ü. Budak, Food Image Classification with Deep Features, 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, Turkey, 2019, pp. 1-6, doi: 10.1109/IDAP.2019.8875946.

[38] J. R. Rajayogi, G. Manjunath and G. Shobha, Indian Food Image Classification with Transfer Learning, 2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), Bengaluru, India, 2019, pp. 1-4, doi: 10.1109/CSITSS47250.2019.9031051.

[39] G. Ciocca, P. Napoletano and R. Schettini, Food Recognition: A New Dataset, Experiments, and Results, in IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 3, pp. 588-598, May 2017, doi: 10.1109/JBHI.2016.2636441.

[40] Giovanni Maria Farinell , Dario Allegra and, Marco Moltisanti, ―Retrieval and classification of food images‖ Computers in Biology and Medicine, Volume 77, 1 October 2016, Pages 23-39

[41] Kagaya H., Aizawa K. (2015) Highly Accurate Food/Non-Food Image Classification Based on a Deep Convolutional Neural Network. In: Murino V., Puppo E., Sona D., Cristani M., Sansone C. (eds) New Trends in Image Analysis and Processing -- ICIAP 2015 Workshops. ICIAP 2015.

[42] L. Herranz, S. Jiang and R. Xu, Modeling Restaurant Context for Food Recognition, in IEEE Transactions on Multimedia, vol. 19, no. 2, pp. 430-440, Feb. 2017, doi: 10.1109/TMM.2016.2614861.

[43] GianluigiCiocca, Paolo Napoletano, and RaimondoSchettini, CNN-based features for retrieval and classification of food images, Computer Vision and Image Understanding, Volumes 176–177, November–December 2018, Pages 70-77

[44] D. J. Attokaren, I. G. Fernandes, A. Sriram, Y. V. S. Murthy and S. G. Koolagudi, Food classification from images using convolutional neural networks, TENCON 2017 - 2017 IEEE Region 10 Conference, Penang, 2017, pp. 2801-2806, doi: 10.1109/TENCON.2017.8228338.

[45] S. Horiguchi, S. Amano, M. Ogawa and K. Aizawa, Personalized Classifier for Food Image Recognition, in IEEE Transactions on Multimedia, vol. 20, no. 10, pp. 2836-2848, Oct. 2018, doi: 10.1109/TMM.2018.2814339.

[46] M. T. Islam, B. M. N. KarimSiddique, S. Rahman and T. Jabid, Image Recognition with Deep Learning, 2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), Bangkok, 2018, pp. 106-110, doi: 10.1109/ICIIBMS.2018.8550021.

[47] K. T. Islam, S. Wijewickrema, M. Pervez and S. O'Leary, An Exploration of Deep Transfer Learning for Food Image Classification, 2018 Digital Image Computing: Techniques and Applications (DICTA), Canberra, Australia, 2018, pp. 1-5, doi: 10.1109/DICTA.2018.8615812.

[48] L. Jiang, B. Qiu, X. Liu, C. Huang and K. Lin, DeepFood: Food Image Analysis and Dietary Assessment via Deep Model, in IEEE Access, vol. 8, pp. 47477-47489, 2020, doi: 10.1109/ACCESS.2020.2973625.

[49] M. A. Subhi and S. Md. Ali, A Deep Convolutional Neural Network for Food Detection and Recognition, 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Sarawak, Malaysia, 2018, pp. 284-287, doi: 10.1109/IECBES.2018.8626720.

[50] TaichiJoutou and KeijiYanai, A food image recognition system with Multiple Kernel Learning, 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, 2009, pp. 285-288, doi: 10.1109/ICIP.2009.5413400.

[51] Alaria, S. K. "A.. Raj, V. Sharma, and V. Kumar.“Simulation and Analysis of Hand Gesture Recognition for Indian Sign Language Using CNN”." *International Journal on Recent and Innovation Trends in Computing and Communication* 10, no. 4 (2022): 10-14.

[52] Dogiwal, Sanwta Ram, Y. S. Shishodia, Abhay Upadhyaya, Hanuman Ram, and Satish Kumar Alaria. "Image Preprocessing Methods in Image Recognition." *International Journal of Computers and Distributed Systems* 1, no. 3 (2012): 96-99.

ISSN: 2250-0839