Jestr

Journal of Engineering Science and Technology Review 14 (2) (2021) 165 - 176

JOURNAL OF
Engineering Science and Technology Review

Research Article

Survey with Bibliometric Analysis of Computer Vision based Automatic Dietary Management for Multifood cuisines to Avert Lifestyle Disease – Obesity

Prachi Kadam* and Shraddha Phansalkar

Symbiosis Institute of Technology (SIT), affiliated to Symbiosis International (Deemed University), Pune, India. Symbiosis Institute of Technology, Symbiosis International University,Near Lupin Research Park, Gram: Lavale,Tal: Mulshi, Pune 412 115

Received 6 August 2020; Accepted 16 December 2020

Abstract

Rising incidences of obesity is a global concern. Practitioners in the field of obesity interventions, work around dietary intake and physical activity. Extensive research is performed in the direction of correct logging of dietary intake, but investigation demonstrates that automated dietary monitoring is more reliable than manual methods of food logging. An image based dietary intake monitoring requires capturing food image, segmentation of food image to get distinct food items, food identification & classification with calorie estimation. In this review, we first establish the popularity of automated dietary monitoring as a strategy to control obesity with exhaustive and conclusive bibliometric analysis. We then carry a comprehensive survey of computer vision-based methodologies for image based automated dietary monitoring in processing the captured food image. A best combination of the food image segmentation, feature extraction and classification methods which yield better accuracy is explored. With diverse multifood cuisines, the methods for food image segmentation and classification are analysed to find the best suitable model for multi-food dishes, dishes with accompaniments and garnishes. The paper gives experimental validations to the findings with Indian food image dataset and shows that feature extraction techniques can leverage classification accuracy of state-of-the-art machine learning and deep learning algorithm. The work thus critically analyses a combination of image segmentation, classification model in multi-food Indian cuisine context.

Keywords: Food image dataset, food image segmentation, feature extraction, food image identification and classification, computer vision, obesity management, image filters, Thali, Indian multifood

1. Introduction

There has been a significant growth in non-communicable chronic diseases worldwide[1], over the past few years. The World Health Organisation (WHO) report on Non Communicable diseases (NCD) in India shows increased risk in premature deaths due to diseases like - cardiovascular disease (27%), cancer (9%), chronic respiratory disease (11%), diabetes (3%) and other NCDs (13%) where obesity is a significant risk factor[2]. Obesity is now looked upon as a disease because it increases the risk of other life-threatening diseases such as hypertension, heart disease, diabetes and cancer[3]. People have become aware of the importance of pursuing a healthy lifestyle which includes right eating habits, physical activity and reducing occupational stress. Obesity is a preventable disease and two major factors to control obesity are - *healthy eating habits and exercise*. If diet is monitored regularly by the patient under the supervision of dietician and doctor, obesity can be controlled and treated[4]. The major challenge in following this is to monitor the eating habits of an individual and calibrate the food intake without perceptions.

Contemporary Dietary surveillance is executed by food journaling with two popular approaches - manual approach and automated food Image based surveillance approach. In manual approach, the user enters a description of his meals eaten in the day as a text data, which is based on perceptions and being subjective, is found to be prone to human error with missed information and time-consuming data entry. Whereas, in Image based surveillance method, the user can upload the food images of meals of that day making this method more convenient and accurate[5]. State of art computer vision based methodologies make further food identification, classification and calorie estimation more accurate.

Web based information systems for automatic dietary management became popular with increase in the use of computer and Internet. There was an efficient replacement of the manual logging methods[6]. The advancement in the smartphone technology has enabled the availability of the ample applications to monitor eating patterns[7]. The analysis has proved that image based automatic method using smartphones to manage dietary patterns are better than paper based dietary management techniques [8] , given the easy and user-friendly apps available for maintaining food logs [9]. Some applications provided overall health management covering diabetes, obesity and mental health [10]. One of the

*E-mail address: prachi.kadam@sitpune.edu.in
ISSN: 1791-2377
doi:10.25103/jestr.142.20

studies shows how an aspiration to control and manage ideal body weight and health increases the use of food tracking apps. These apps maintain a food log as real time images of food consumption by the individual throughout the day [11], provide assessment in terms of calorie intake, amount of physical activity done and suggestions for an ideal diet plan. The smartphone apps provide the methods to capture images of every meal consumed by the user. From the food images, calorie intake and nutrition of every meal can be calculated, thus estimating healthy eating patterns of the user [12]. The datasets used for calorie measurement can be referred from restaurant [13] or collected by crowdsourcing [14]. The use of deep learning algorithm along with contour measurement of food item has shown improvement in calorie estimation from food images [15] [16].

This work highlights the increasing importance of research in the area of dietary surveillance with focus on image based food identification, segmentation and classification using computer vision methods with the state-of-art work done in this area by bibliometric survey as well. The Bibliometric analysis is confined to paper results from Scopus database and is limited to English language only. With bibliometric survey we enlist important phases in processing of food images*: Image segmentation, Image feature extraction, classification*. We then present a detailed description and tabular comparison of different algorithms and their combinations is presented with their stated accuracy in context of food image processing. The work further critically analyses the different image segmentation techniques that are employed in single food as well as multifood images context in terms of classification accuracy. We further investigate the effect of feature extraction techniques for Indian food images to leverage the accuracy of machine learning and deep learning food image classifiers with experimental validations on a synthetic Indian food image data set. Thus, the paper helps in choosing a best combination of image segmentation, feature extraction and classification methods in the context of diverse Indian food images. Summarizing the contribution of the work is :

1. Reviewing Computer Vision based methods of Automated dietary management; state of art
2. Critical Review of Image segmentation techniques and Food classification techniques with feature extraction for food images both in single food as well as Multifood image data sets with the resultant accuracy.
3. Experimental validations on accuracy of classification by choosing feature extraction techniques and segmentation methods with different food image classification methods for synthetic food image data set of Indian cuisines

2. Bibliometric Analysis of Dietary management for avoiding Lifestyle disease -Obesity

Bibliometric Analysis gives an insight of impact of the research and its quality. From this analysis, prominent research areas are unveiled as well as uncovered areas and research gaps are discovered. To understand the emerging areas in the field of - Dietary Management, the analysis is performed with specific keywords related to diet management.

Two significant keywords “Dietary Management” OR “Lifestyle disease: Obesity” were used for bibliometric analysis to identify amount of research as well as the most obvious directions of research executed in this area. From the analysis, the frequency of research documents, patterns of research documents, links between recognised works, the territorial distribution is highlighted. This analysis helps us in localization of the problem, its inter-disciplinary relevance and the research gaps to identify the scope for further research. Analysis of year wise research publications shows that research in the methods applied to avoid or prevent obesity is on the rise. The figure 1 shows the number of publications is increasing over year because of the growing incidences of obesity.

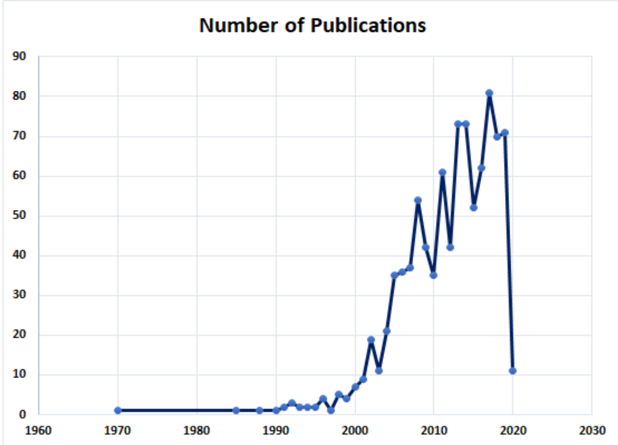

Fig 1. Number of publication per year [17]

The figure 2 shows bibliographic coupling. Bibliographic coupling is a similarity measure used to establish a network between documents when two or more cite a third work. T. Lobstein’s study shows that there is coupling across the rising crisis of obesity in children and youngsters [18]. Whereas, Ve´ronique L. Roger has given a detailed statistical study of heart diseases and stroke, out of which one determinant is obesity [19]. This analysis substantiates the strengthening of obesity control or avoidance methods with the papers that show the catastrophic effects on the health by obesity. There is a coupling between works that are based on obesity with the control methods like dietary control and physical activity etc.

The figure 3 shows the United States, Australia, United Kingdom and Canada are leading in the research related to Obesity. Indian subcontinent shows less but significant research in this area.

For in-depth understanding, keyword analysis carried out to see the coupling between Dietary Management and obesity in literature is discussed as follows. The table 1 shows primary and secondary keywords used for search query.

The keywords were provided as query to the Scopus tool. Total 900+ number of documents were returned. There are 5 clusters formed from the given keywords. The keyword network is shown in the figure. These clusters explain the rising crisis of obesity, risk factors for obesity, lifestyle modifications needed to avoid obesity like controlling diet and including physical activity in day-to-day life, chronic diseases risks from obesity like diabetes, metabolic syndrome, cardiovascular diseases and hypertension. In the keyword network analysis, medical reference of research is more thus establishing the importance of research to curb the obesity pandemic. Our interest is in methods to control obesity effectively with different methods where we found that monitoring the diet was the most popular method. The efficiently we monitor and measure the diet, obesity can be controlled effectively.

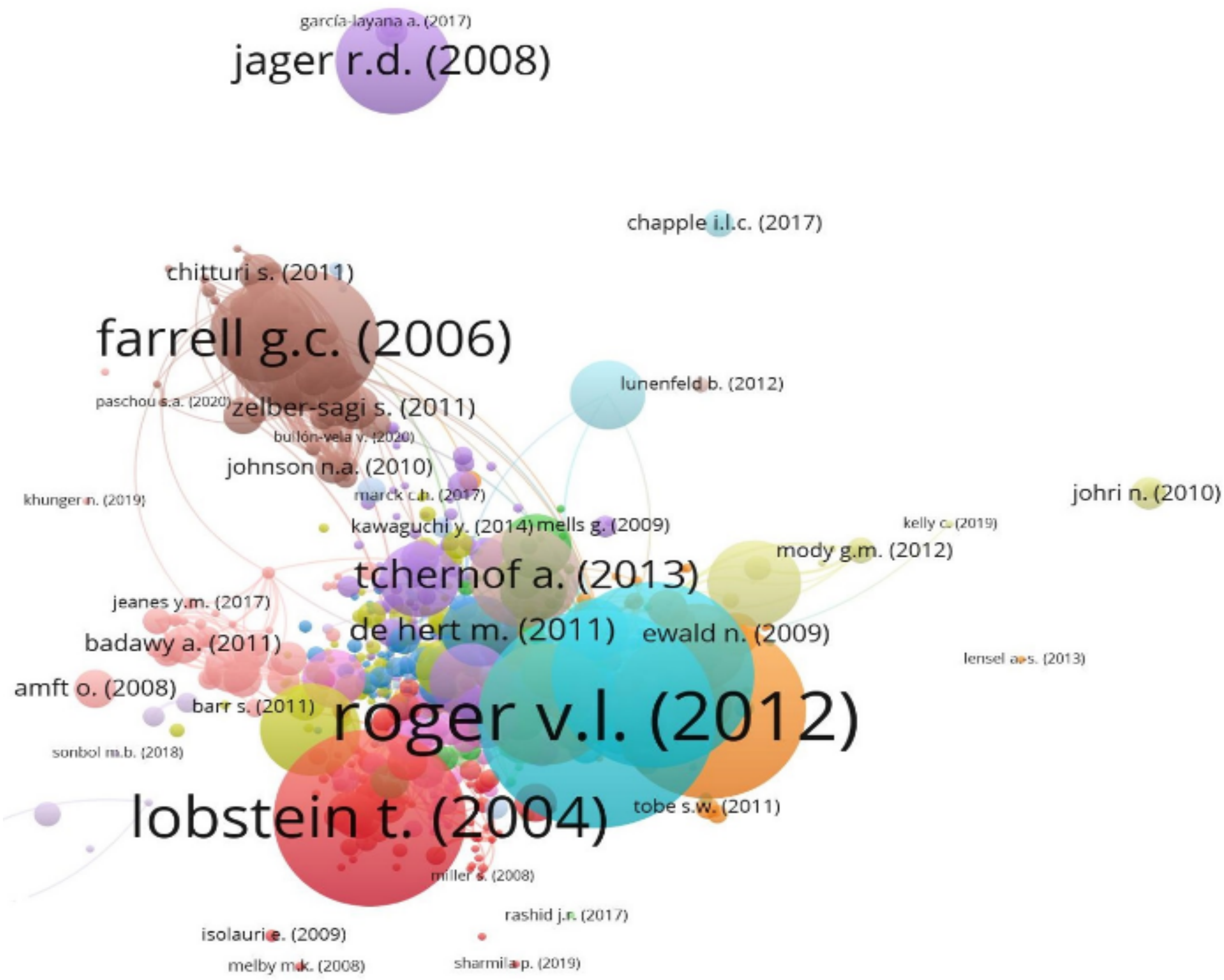

Fig. 2. Bibliographic Coupling [17]

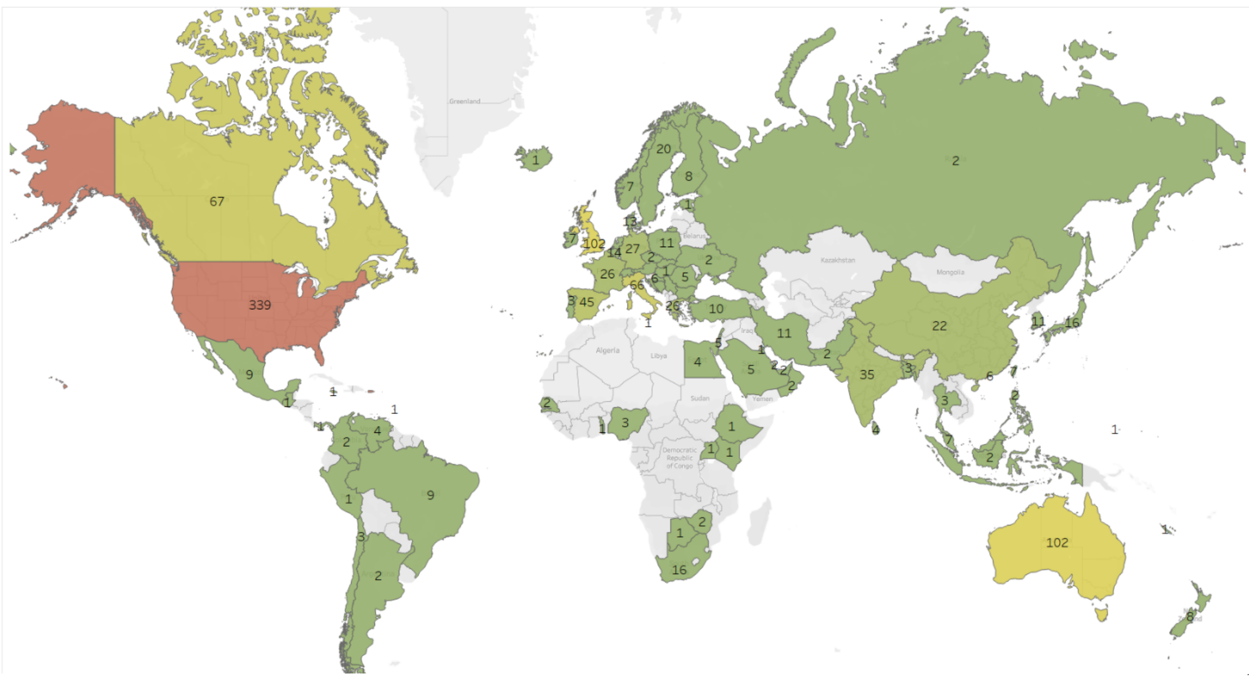

Fig.
3. Country wise Distribution of Publications[17]

Table 1. Primary and Secondary Query Keywords for Dietary Management and Obesity

Primary – Keyword	Dietary Management	
Secondary - Keyword	(AND)	Lifestyle Disease, Obesity
	(OR)	"Lifestyle" OR "human" OR "nutrition" OR "weight reduction" OR "weight loss"OR "dietary intake" OR "dietary assessments" OR "food intake" OR "nutritional assessment" OR "diabetes" OR "cardiovascular diseases" OR "body mass index" OR "Overweight" OR "diet restriction" OR "disease management"

Subject-wise analysis given in the figure 5 for the study of obesity mostly reveals that the study is done predominantly by medicine experts, medical practitioners, genetics and pharmaceutics. There is lesser contribution to this global problem from the stream of Computer Science area; thus, our study will be supplementing the solution space to control obesity with research of computer vision methods.

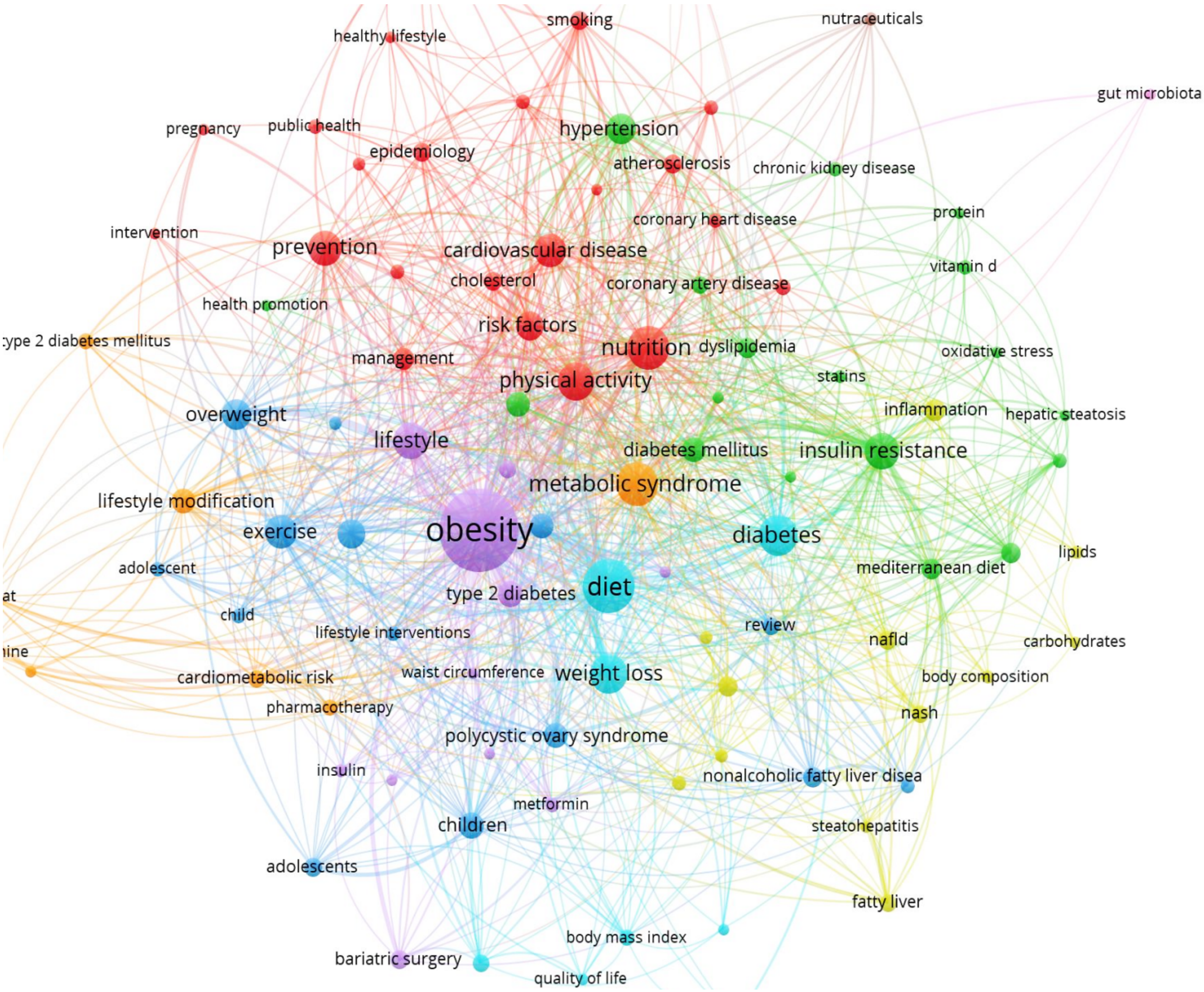

Fig. 4: Keyword network [17]

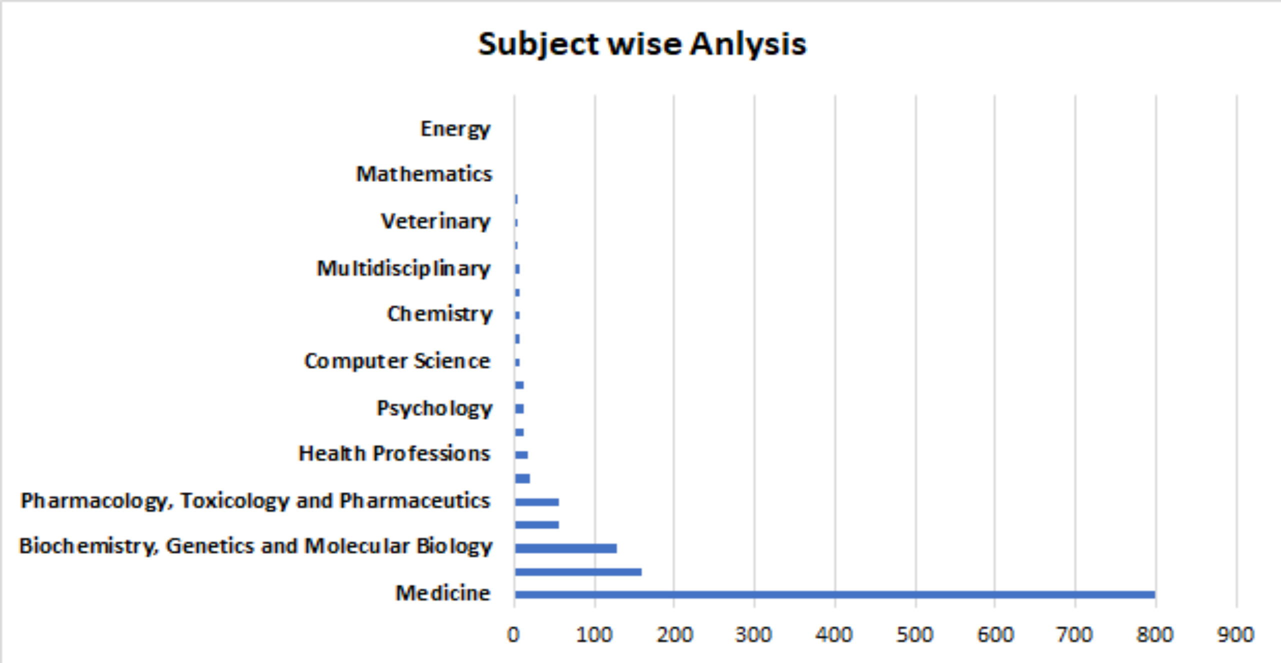

Fig. 5. Publications per Subject Area [17]

3. Bibliometric Analysis of Food Image Processing for Dietary Management

From the study it is vital that diet management is one of the crucial determinants of obesity. For minimum number of human introduced errors, automatic dietary management is suggested. Automatic dietary management is food image based. The image of the meal consumed can be used further to estimate the calorie count. The bibliometric analysis of food image processing in this area is made to understand key areas of research. Steps for food image processing are - segmentation for acquiring individual food item, feature extraction for food item identification and classification. To set a foundation towards realizing the extent of research in this area, basic keywords applied were – "Food image processing" OR "Dietary management".

Table 2. Primary and Secondary Query Keywords for Dietary Management and Food Image Processing

Primary – Keyword	**Food Image Processing**	
Secondary - Keyword	(AND)	Diet Management
	(AND)	Obesity
	(OR)	"Image processing" OR "human" OR "nutrition" OR "image analysis" OR "dietary assessments" OR "food intake" OR "nutritional assessment" OR "computer vision" OR "classification" OR "CNN" OR "image Processing, Computer-Assisted"

One of the prominent clusters of the keyword network shown in the blue region is around dietary surveillance using Image processing. If processing of food images is performed accurately the identification of food item and estimation of its calorific value will be near accurate.

From the bibliometric survey it is clear that the milestones in calculating calories from a food image are - segmentation (a food image may contain multiple food item on one plate, to separate each food item segmentation is used), feature extraction (identifying unique features to classify a food item in the correct category), identification and classification. Following sections are dedicated to detailed survey of publications in the area of food image processing

4. Phases in Image based Automatic Dietary Management

Figure 7 shows various methodologies used in image processing from capturing an image to classifying it.

4.1 Segmentation Techniques

Image segmentation techniques are used to separate individual food items from multifood plate. This helps in classifying the individual items and also for controlling and monitoring the dietary management in terms of calorie estimation. Image segmentation can be divided into basic three categories – thresholding, edge-based and region-based. In thresholding method, a threshold is decided to categorise the pixels of an image in – either within the region or out of the region. Thresholding can be based on calculating the histogram of the image or using multivariate classifiers. In Region based method, contour-based method is popular, where energy functions are used to separate the pixels of interest from the rest of the image. These pixels of interest are bounded by the contour which is interpolation of the points of interest. Interpolation can be done by linear, spline or polynomial method to define the curve in the image around the object of interest. This method forms closed contour around the object in the image which should be segmented from the rest of the image. Another popular region-based method is Graph based method. In this method the image is represented as a weighted graph. The nodes of the graph are the pixels and edge is the weight. The weight gives similarity or dissimilarity of the pixels thus differentiating the object of interest from its background. Normalised cut is a variation in graph-based method which uses pixel intensity similarity and spatial proximity to differentiate pixels of the image. Local variation method is another graph-based segmentation method which forms a graph with pixels as its nodes and weight of the edge is the intensity of the pixel and distance between them[20].

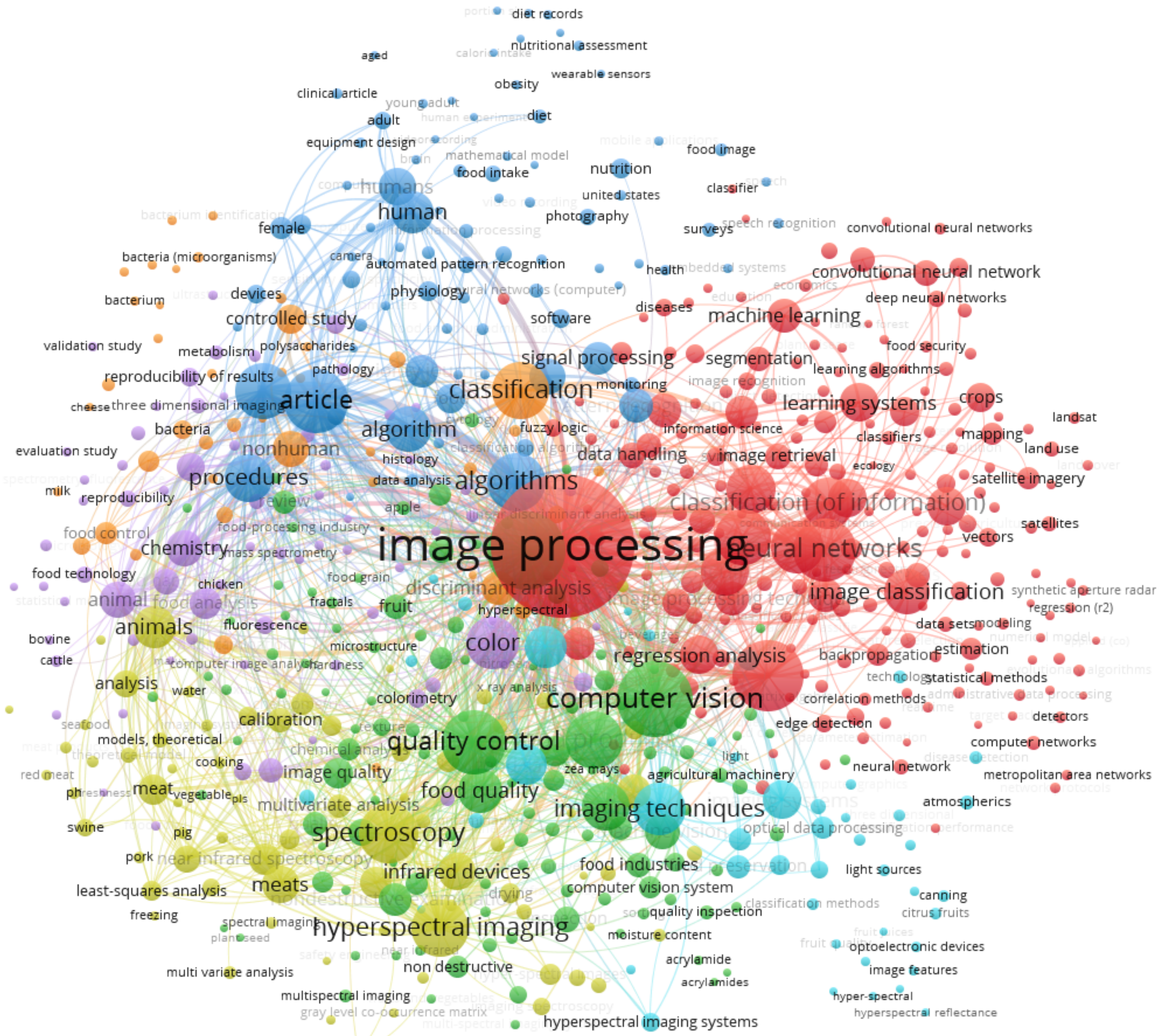

Fig. 6. Keyword network for Food Image Processing and Dietary Management[17]

The authors conclude with a trade-off between segmentation quality and speed of segmentation by comparing algorithms like Spectral Embedding MinCut, Normalized Cuts, Local Variation Algorithm and Mean Shift Segmentation [21]. Integrated image segmentation and classification method was proposed to calculate area and estimate the weight of the food item which in turn was used to calculate the nutritional value of the food. Algorithms like active contours, normalized cuts and local variation were studied for food image segmentation with complex backgrounds. It was found that local variation method was more effective than the other two methods[22]. An interactive segmentation was used to improve segmentation accuracy, where user food item is extracted by the user[23]. With increased accuracy in machine learning algorithms like K-means [24], active contour with K-means [25] and deep learning algorithms like weakly supervised CNN applied on UECFOOD-256 and Food-101food dataset [26], CNN applied on UEC-FOOD100 dataset[27] [28], ResNet used in FoodAI on FoodAI-756 dataset [29] show better accuracy when applied to respective food image datasets.

Jseg segmentation is a robust image segment algorithm. It consists of two steps - colour quantization and spatial segmentation. In colour quantisation, colours are used to identify different regions in an image on the basis of colour classes that form a colour map of the image. In spatial segmentation, on the basis of colour map a colour texture of boundary and interior of the region is found. This is called as a J-image. Now by applying a region growing method image is segmented to find the region of interest [30].

Edge Detection, Morphological Operation, Threshold and Clustering based image segmentation techniques show that Gradient Magnitude method of edge detection gives promising results when applied on different food images

[31]. A supervised learning method that allows the user to manually draw bounding box around the region of interest in the training is the fastest method to train the classifier to identify the food item from multi - dish scenario [32].

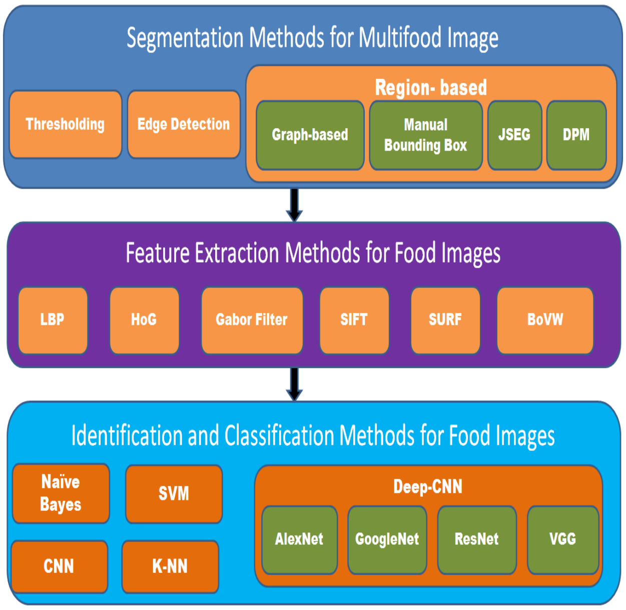

Fig. 7 Methodologies for segmentation, feature extraction and classification of Food Images for Automatic Dietary Management

4.2 Feature extraction methods

The varieties of low-level invariant features are available to describe the object appearances. Two popular methods for feature extraction with the best results are - Scale Invariant Feature Transform (SIFT) and Speeded-Up Robust Feature Transform (SURF). SIFT has used to represent food features as it provides a powerful descriptor due its stability under different scale and orientation changes as well as being robust to occlusion and clutter[33]. From the analytical study of SIFT and SURF, it is found that SIFT is slower than SURF but amount key points found by SIFT is more than SURF [34]. The Local features using SURF is computationally efficient to detect and derive meaningful local descriptors. However, due to complex appearances of the real food images, using a single descriptor in isolation is not sufficient to effectively represent the large variation of foods [35]. Local Binary Pattern (LBP) is known to be a powerful feature for texture classification. In LBP method, the features are calculated by finding LBP values at each pixel which is represented by a histogram. The histogram pattern can be used for classification by applying Euclidean distance [36]. Literature shows combination of local features and global features improves the classification to a large extent like histogram of Oriented Gradients (HOG) and SURF [37]; local color, local entropy color, Tamura perceptual features, Gabor filters, SIFT descriptor, Haar wavelets, Steerable filters, and DAISY descriptor and color and texture; texture features and Gabor filter [38]; color features and Gabor filter [39]; integration of SIFT and SURF using late fusion [40]; local features and BoVW [41]; SIFT and local intensity order pattern (LIOP) descriptors[42]; SIFT and LBP integration method [43]; color & texture features using Gabor filter [44]; texture feature extraction using LBP features [45][46]; CNN based features [47]. Therefore, a review proves that using local features in combination has proved to be more beneficial [48].

An image can be described by its descriptors i.e. unique features called Bag of visual words [49] or Bag of key points. The use of these descriptors, exclusive to that food item only, can be exploited to classify the food items in a specific category. The results are optimised if BoVW are formed by scale-invariant feature transform on the HSV color space and then classified with k-means cluster [50].

4.3 Image Identification and Classification

The feature extraction is a critical contributor to the accuracy of the classification process. The image set is divided into two parts - training and testing images. The classifier with its predefined categories is trained with the images defined by its unique features. Once the classifier is trained, test images are applied to the classifier to be categorised into its class. The accuracy depends on images that are correctly classified into its appropriate class.

The research shows a combination of Feature Extraction techniques and Classifiers to achieve maximum accuracy. As discussed in the previous section, SIFT has shown promising results in extracting exact features. By combining sparse coding with SIFT and applying color histogram and Gabor filter to the food image give concatenated feature vector which is stronger than feature vector applied by only SIFT. The research shows using this feature vector to train an SVM (Support Vector Machine) classifier with 50 categories and 100 food images in each category gives an accuracy of 68.3% [51].

It is necessary to obtain each food item separately, for a food image containing multiple food items. This can be achieved by segmentation for object detection. Multiple food item detection was proposed by Yuji Matsuda, et. al. The method proposed consisted of segmentation, feature extraction and classification. Image segmentation was achieved by combining three techniques - Deformable Part Model, Circle Detector and JSEG region segmentation followed by feature extraction that was done by SIFT, CSIFT, HoG and Gabor texture features. The feature vector was applied to SVM with multi-kernel learning to classify images into 100 categories. 68.9 % accuracy was received with the proposed idea [52].

Due to technological advances in wearable devices and mobile camera, apps have been developed which help capturing food images to maintain food logs and to track dietary intake asked to be maintained by the physicians/ Doctors [53]. These applications capture food images, identify food item, classify into a category and find nutrient contents of the food. Some applications also help to maintain calories consumed in a day. A comparison of such apps shows variety of algorithms used to track food activity - SIFT, SURF, HoG, Deformable part models for segmentation, optical character recognition for food labels marked on the food image, food object identification by supervised learning with region of interest marking, Multi kernel learning and SVM for classification [32]. Uploading the data from all smart devices used by the patient, related to healthcare along with food items logged, on a cloud-based system helps in overall health management [54].

An exclusive idea of identifying food item based on the geographical location of the restaurant, using Weakly supervised learning method(SVM with MKL), is discussed in the paper [55]. Even if a patient is under food activity observation, restaurant eating is not restricted. Thus, a patient can log food consumed in a restaurant with its geographical location information that will help identify the food item from the logged image. This is possible because every restaurant has its unique set of food item on the menu. The restaurant can be identified on the basis of the geographical location, resulting in the identification of the food item from the food image.

Similar idea was proposed by Luis Herranz, et. al but with probabilistic model approach [56]. They use SIFT encoded with bag of visual words to extract feature vector. This feature vector with the geographical location information is used to train the classifier. A probabilistic model to classify the image is adopted by combining location, restaurant, food dish and feature vector. Maximally Stable Extremal Regions (MSER) features with Content-Based Visual Information Retrieval (CBVIR) method proposed by the authors show partially successful [57].

For any kind of classification application, the most important part is a trained classifier. The first step for the food images would be to differentiate the food and non-food items [58]. The training is robust if the classifier is trained with all possible types of images which is tedious and an impossible job. A solution to this problem could be learning by active labelling proposed by Marc Bolaños, et. al. The wearable device was used to capture and log the food consumed in a day by the patient. The images captured are applied to a pre trained classifier with limited images, the image is classified if recognised by the classifier. If the image is unknown, the user is prompted to label the category of the image, thus training the classifier with the new, unknown image. This framework allows the classifier for learning in supervised manner [59]. Supervised method like bounding box method especially for Multifood detection is needed. The introduction of such minimal supervision with deep learning algorithm like CNN, helps improving classification accuracy [60].

The traditional classifiers started getting replaced with the development and enhancement in machine learning algorithms. Most of the literature used methods that included naïve Bayes classifier, support vector machines, K-nearest neighbors, Gaussian mixture model, decision tree and Radial Basis Function (RBF) classifiers. The development in deep learning algorithms in computer vision has not only improved the performance of computational power but also of the algorithm [61]. Sample size of food images is a crucial problem in training neural networks. Using faster RCNN gives better performance in such cases [62].

The combination of three deep CNN methodologies showed an improved accuracy (73.5%) of food classification. There are pre-trained models that use CNN as core technique. The use of pre-trained models allows users a benchmark to start from instead of training a model from scratch. This saves time and we can build the dataset on the existing one. The methodologies used for food image dataset are - AlexNet, GoogleNet and ResNet [63], GoogleNet [58], VGG-16 and ResNet. A study shows the use of multiple classifiers which are trained on features extracted by multiple deep models using Induced Order Weighted Averaging (IOWA) and Particle Swarm Optimization (PSO) based fusion [33]. Such combinations of methodologies are also mentioned as transfer learning [64]. A review by Subhi et. al on automatic food recognition and dietary assessment gives a better insight in different methodologies used for segmentation, feature extraction and classification [65]. Some more studies in the table have been extended and added.

Table 3 . Summary of various feature extraction and classification methods with their performance for Single food dishes

Feature Extraction	Classifier	Performance	Dataset	Ref.
BoF, Gabor, Color, HoG, Texture	MKL	62.50%	Web Images	[49]
CNN	VGG , ResNet and InceptionV3	62.76	Yummly48K dataset	[66]
LBP and color	DCNN-Food	70.40%	UECFOOD100, UEC-FOOD256	[67]
CNN	DCNN, Ensemble Net	72.10%	ETH Foof-101, Own Indian food dataset with 50 classes and 100 images each	[61]
ROOTHoG and colour	DCNN	72.26%	ILSVRC 1000-class dataset	[68]
SIFT, PRICoLBP, and Bag of Textons	SVM	75.74%	UNICT-FD1200	[69]
SIFT and color	SVM	78.00%	Web Images	[50]
PCA, CFS and IG	AlexNet and CaffeNet models, ResNet	80%	Mealcome (MLC dataset), MLC-41 dataset.	[70]
CNN	SVM	82.20%	Food images from restaurants of Bern University hospital	[71]

Table 4 . Summary of various feature extraction and classification methods with their performance for Multifood dishes

Image Segmentation Technique	Feature Extraction	Classifier	Performance	Dataset	Ref.
JSEG segmentation, circle detector and DPM	HoG, SIFT, Gabor, color and texture	MKL-SVM	45%	85 - food categories	[52]
Connected component analysis, active contours, and normalized cuts	DCD, SIFT, MDSFIT, and SCD	KNN	64.50%	Department of Foods and Nutrition at Purdue University, FNDSS	[22]
Texture segmentation	BoF, SFTA and color	SVM	70%	Own- 40 type of Thai street food/ restaurant	[72]

Sobel Operator edge detector, watershed marking	CNN	DCNN+edge computing	77%	UEC-256, UEC-100 Dataset	[73]
Spatial relationships and Semantic Texton forest	Pairwise local features	SVM	78%	PFID	[74]
Deep Lab model	CNN	GoogleNet	79%	Food-101, MenuMatch	[13]
Bounding Box and GraphCut segmentation	SURF and color	SVM	81.60%	Own- 50 category food dataset	[75]
Feature point method	Gaussian region detector and SIFT, color, entropy, Gabor, Tamura, Haar	Multiclass SVM	84%	PFID	[76]
ROI detection, Normalized Cuts framework	Wavelet, Steerable Daisy, Predominant color divided into local and global features	SVM	86.10%	FNDDS, StockFood, gathered by mdFR	[77]
Edge detection	HSV and histogram	KNN and SVM	85% and 75%	Own	[78]
Graph cut segmentation	GraphCut, color, size, shape and texture	SVM	95%	Own dataset with 170,000 images	[79]

5. Discussions

The Section 4 presents a comparative analysis of different methods in food image segmentation, feature extraction and classification techniques applied across different food image data set. This is done in context of single as well as multi-food images context.

For food image segmentation, graph based methods show efficient results when in turn applied to identify and classify food items. Graph cut, ROI (region of interest) detection, normalised cut, bounding box are some of the graph-based methods that have seen more usage in segmentation for food images. [75], [76], (77), [78]. This is because the contour detection using graph cut segmentation methods are popular for being robust and efficient [79]. Edge detection technique (78) shows a significant improvement in classification accuracy of the classification methods and its characteristic to define boundary for adjacent elements makes it appropriate for Multifood images.

Apart from contour or edge detection, it is observed that the extracted color is a predominant and important feature for multi-food images [77] [49] [52] [72] [79]. The performance of any algorithm is dependent on the variety of images (categories) available in the food image dataset, number of diverse images in one category and the quality of image captured - by a digital camera or smartphone. As per the observations from the table, when feature extraction using color, histogram, SIFT or BoVW is chosen to be a step in food image classification and these features are used to train a deep learning algorithm like CNN or ensemble net, the performance is better than classification with only deep learning algorithm in isolation [80]. This is well justified, as every food image has a unique feature vector and features of every aspect of food image is extracted with color, shape, size and texture yielding accurate classification. This is because the classifier is trained with all the features that define an image giving us an enhanced feature vector specific to that food item. Another observation is that, the more the number of food classes included in the dataset, lesser is the accuracy of classification [64]. This is because the amount of misclassification increases when similar looking food items have different categories. For example, vegetable curry and chicken curry. This poses a challenge in the classification of food images. In the next section, we validate the findings by applying a chosen image segmentation technique, feature extraction methods with prominent machine learning and deep learning classifiers on a synthetic Indian image data set.

6. Experimental Validation for Food Image Classification:

For a primary study, we generate a synthetic data set of 30 Indian food categories with 30 images each. Every category also had 40% images from standard Food-101[] dataset and 60% were actually collected from field work. This is the synthetic data set for a primary study. The images were visually checked for sufficient distinctness. Food 101 dataset is a food image dataset compiled from foodspotting.com which is real-world food image dataset. So, the resultant database of single food images consists of the combination of Indian food images and American fast food images. Al the images were first run through edge based segmentation algorithm where 3x3 sobel operator is used to detect edge of the food image. This has resulted in a better smoothening in case of single food image.

Then they were run through the prominent machine learning and deep learning algorithms. Two test results were decided upon – without filters and with filters. All the experiments were run with Python using Tensorflow library. For experimental setup of 'without filter', was prototyped and tested on the dataset with these additional Indian food categories using Convolutional Neural Networks (CNN), Bag of Visual Words(BoVW), RandomForest and SVM individually.

The figure shows flow of classifying the food images that include a combination of American and Indian food categories. The filters used were – colour histogram, edge histogram and Gabor filter. From the various papers discussed till now, some of the image features were highlighted as important for classification. With respect to Indian food images, since Indian food is very colourful, colour histogram collects the distribution of different colours across a dish. The edge histogram is useful in marking the shape of the food item in the image by thresholding its boundaries. There are cuisines which look similar in colour and shape in Indian cuisines, for example – Poha and Lemon rice. But their content, calorie and nutritional values are entirely different. In these cases, gabor filter, which is texture feature filter is useful to identify the correct food item.

The prototype shows accuracy of 71%, 68%, 53.33% and 42.5% respectively when only machine learning algorithms are applied. With application of filters the accuracy enhanced to – 78%, 72.6%, 65.5% and 56.3% respectively. The results are summarised in the table 5 as follows:

From this experimental study, it can be verified that, by applying filters, the efficiency of classification increases in case of Indian food. This verifies the literature studied, that if pre-processing using filter is done by the traditional methods then advanced machine learning algorithms give promising results.

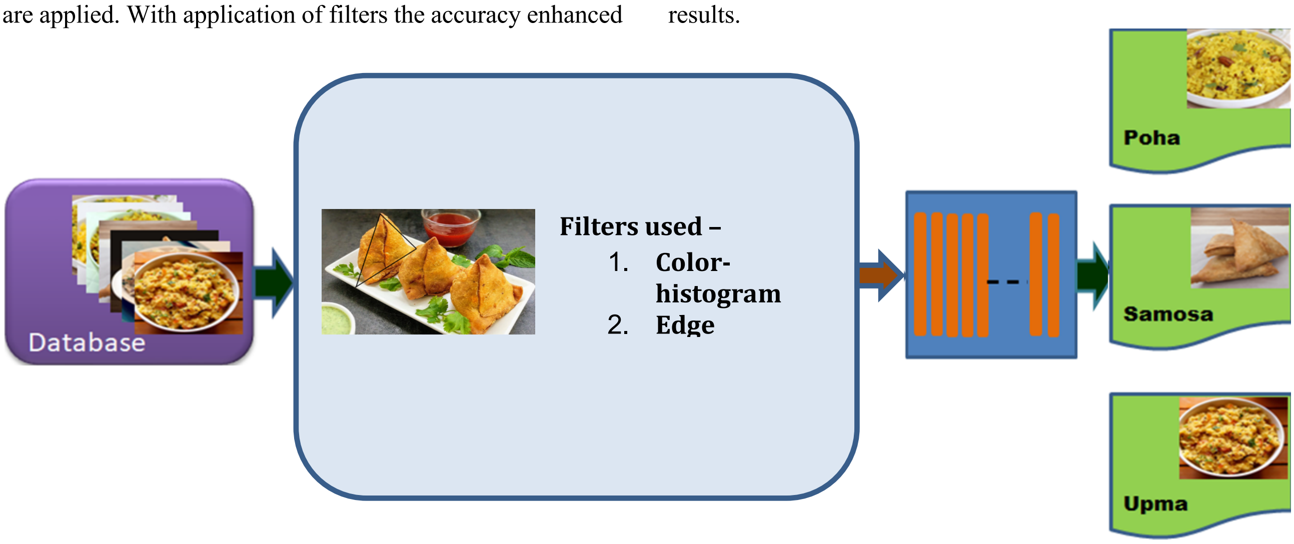

Fig. 8. Classification of food images with Filters

Table 5. Accuracy of various algorithms with/without filters.

	CNN	BoVW	Random Forest	SVM
Without Filters	71	68	53.33	42.5
With Filters (Colour Histogram, Edge Histogram and Gabor filter)	78	72.6	65.5	56.3

7. Conclusion

The area of automated dietary monitoring to control obesity was analysed firstly with bibliometric analysis where the global survey of the different strategies to handle this lifestyle disorder were outlined to find the innovative contribution of this area of work. In this paper, we further investigated computer vision methodologies for image based automatic dietary management research area. Followed by this, various steps in automated food image logging for dietary assessment for multifood dishes have been reviewed in detail. From the keyword network it was clear that computer vision based dietary management should have the following steps: *Image Segmentation, Feature Extraction, Food object Identification and Classification.* The combination of various image segmentation and classification techniques has been reviewed for better results on variety of food image datasets. In food image classification, it is proved with preliminary experimental validations that the use of image filters, followed by the classifiers with machine learning algorithms would yield better results. Although the advanced methods illustrate improved performance in image classification, there exist few challenges in the classification of multifood images. Thus this contribution elevates the importance of: 1. Computer Vision based methods of Automatic dietary management , 2. Phases of Food image processing for dietary monitoring, 3. Critical review of image segmentation and classification methods for multifood and single food images. This will further help in the design of an automated dietary monitoring application which will classify an image, estimate its portion and calibrate calories for users to maintain a healthy lifestyle and keeping obesity and related chronic diseases at bay.

References

1. "World Health Organization - Noncommunicable Diseases (NCD) Country Profiles, 2018."
2. "World Health Organisation." [Online]. Available: https://www.who.int/gho/publications/world_health_statistics/2019/en/.
3. A. Afshin *et al.*, "Health effects of overweight and obesity in 195 countries over 25 years," *N. Engl. J. Med.*, vol. 377, no. 1, pp. 13–27, 2017.
4. P. Pouladzadeh, S. Shirmohammadi, and A. Yassine, "You are what you eat: So measure what you eat!," *IEEE Instrum. Meas. Mag.*, vol. 19, no. 1, pp. 9–15, 2016.
5. C. J. Vieira Bruno, Silva Resende, "HHS Public Access," *J. Heal.*

Med. informatics, vol. 176, no. 5, pp. 139–148, 2017.
6. S. A. Noah *et al.*, "DietPal: A Web-based dietary menu-generating and management system," *J. Med. Internet Res.*, vol. 6, no. 1, pp. 32–48, 2004.
7. S. S. Coughlin, M. Whitehead, J. Q. Sheats, J. Mastromonico, D. Hardy, and S. A. Smith, "Smartphone Applications for Promoting Healthy Diet and Nutrition: A Literature Review.," *Jacobs J. food Nutr.*, vol. 2, no. 3, p. 021, 2016.
8. F. Jimoh *et al.*, "Comparing diet and exercise monitoring using smartphone app and paper diary: A two-phase intervention study," *JMIR mHealth uHealth*, vol. 6, no. 1, pp. 1–14, 2018.
9. J. L. Lemacks, K. Adams, and A. Lovetere, "Dietary intake reporting accuracy of the bridge2u mobile application food log compared to control meal and dietary recall methods," *Nutrients*, vol. 11, no. 1, 2019.
10. O. Byambasuren, S. Sanders, E. Beller, and P. Glasziou, "Prescribable mHealth apps identified from an overview of systematic reviews," *npj Digit. Med.*, vol. 1, no. 1, pp. 1–12, 2018.
11. D. Lupton, "'I Just Want It to Be Done, Done, Done!' Food Tracking Apps, Affects, and Agential Capacities," *Multimodal Technol. Interact.*, vol. 2, no. 2, 2018.
12. P. Pouladzadeh, S. Shirmohammadi, and R. Al-Maghrabi, "Measuring calorie and nutrition from food image," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 8, pp. 1947–1956, 2014.
13. A. Myers *et al.*, "Im2Calories: Towards an automated mobile vision food diary," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2015 Inter, no. December, pp. 1233–1241, 2015.
14. P. author3 Jun Zhou, BE,1 Dane Bell, BA, MA, PhD,2 Sabrina Nusrat, BS, MS, PhD,3 Melanie Hingle, BS, MPH, PhD,4 Mihai Surdeanu, BS, MS, PhD,3 and Stephen Kobourov, BS, MS, "Calorie Estimation From Pictures of Food: Crowdsourcing Study," *Interact. J. Med. Res.*, 2018.
15. Y. Liang and J. Li, "Computer vision-based food calorie estimation: dataset, method, and experiment," 2017.
16. R. Yunus *et al.*, "A Framework to Estimate the Nutritional Value of Food in Real Time Using Deep Learning Techniques," *IEEE Access*, vol. 7, pp. 2643–2652, 2019.
17. "Scopus," 2020. [Online]. Available: https://www.scopus.com/search/. [Accessed: 11-Feb-2020].
18. T. Lobstein, L. Baur, and R. Uauy, "Obesity in children and young people : a crisis in," *Obes. Rev.*, vol. 5, pp. 4–85, 2004.
19. V. L. Roger *et al.*, "Heart disease and stroke statistics-2012 update: A report from the American heart association," *Circulation*, vol. 125, no. 1, pp. 2–220, 2012.
20. "Biomathematics and Statistics Scotland," 2010. . Available: http://www.bioss.ac.uk/people/chris/ch4.pdf. .
21. F. J. Estrada and A. D. Jepson, "Benchmarking image segmentation algorithms," *Int. J. Comput. Vis.*, vol. 85, no. 2, pp. 167–181, 2009.
22. and E. J. D. Ye He, Chang Xu, Nitin Khanna, Carol J. Boushey, "HHS Public Access," in *IEEE international Conference on Multimediaand Expo*, 2013.
23. S. Inunganbi, A. Seal, and P. Khanna, "Classification of Food Images through Interactive Image Segmentation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10752 LNAI, no. January, pp. 519–528, 2018.
24. X. Zheng, Q. Lei, R. Yao, Y. Gong, and Q. Yin, "Image segmentation based on adaptive K-means algorithm," *Eurasip J. Image Video Process.*, vol. 2018, no. 1, 2018.
25. H. Hambali, Hamirul'Aini & Khalid, Salwa & Mahmuddin, Massudi & Harun, Nor & Siraj, Fadzilah & Harun, "Segmentation of Multifood Images Using Integrated Active Contour and K-means.," in *Conference: Progress in Computer Science and Information Technology International Conference*, 2016.
26. Y. Wang, C. Liu, F. Zhu, C. J. Boushey, and E. J. Delp, "EFFICIENT SUPERPIXEL BASED SEGMENTATION FOR FOOD IMAGE ANALYSIS School of Electrical and Computer Engineering , Purdue University Cancer Epidemiology Program , University of Hawaii Cancer Center," pp. 1277–1281, 2017.
27. J. Dehais, M. Anthimopoulos, and S. Mougiakakou, "Food image segmentation for dietary assessment," *MADiMa 2016 - Proc. 2nd Int. Work. Multimed. Assist. Diet. Manag. co-located with ACM Multimed. 2016*, pp. 23–28, 2016.
28. W. Shimoda and K. Yanai, "CNN-based food image segmentation without pixel-wise annotation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9281, pp. 449–457, 2015.
29. D. Sahoo *et al.*, "FoodAI: Food image recognition via deep learning for smart food logging," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, no. May, pp. 2260–2268, 2019.
30. A. Kaur and D. Y. Randhawa, "Image Segmentation with Artificial Neural Networs Alongwith Updated Jseg Algorithm," *IOSR J. Electron. Commun. Eng.*, vol. 9, no. 4, pp. 01–13, 2014.
31. S. V. and S. S., "Study and Analysis of Image Segmentation Techniques for Food Images," *Int. J. Comput. Appl.*, vol. 136, no. 4, pp. 20–24, 2016.
32. S. Knez and L. Šajn, "Food object recognition using a mobile device: State of the art," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9281, pp. 366–374, 2015.
33. S. Khan, K. Ahmad, T. Ahmad, and N. Ahmad, "Food items detection and recognition via multiple deep models," *J. Electron. Imaging*, vol. 28, no. 01, p. 1, 2019.
34. M. N. bin Razali, N. Manshor, A. A. Halin, N. Mustapha, and R. Yaakob, "Analysis of SURF and SIFT Representations to recognize food objects," *J. Telecommun. Electron. Comput. Eng.*, vol. 9, no. 2–12, pp. 81–88, 2017.
35. H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 3951 LNCS, pp. 404–417, 2006.
36. M. Pietikäinen, "Image analysis with local binary patterns," *Lect. Notes Comput. Sci.*, vol. 3540, pp. 115–118, 2005.
37. Z. Mehmood, F. Abbas, T. Mahmood, M. A. Javid, A. Rehman, and T. Nawaz, "Content-Based Image Retrieval Based on Visual Words Fusion Versus Features Fusion of Local and Global Features," *Arab. J. Sci. Eng.*, vol. 43, no. 12, pp. 7265–7284, 2018.
38. M. H. Rahmana, M. R. Pickering, D. Kerr, C. J. Boushey, and E. J. Delp, "A new texture feature for improved food recognition accuracy in a mobile phone based dietary assessment system," *Proc. 2012 IEEE Int. Conf. Multimed. Expo Work. ICMEW 2012*, no. March 2014, pp. 418–423, 2012.
39. F. Zhu, M. Bosch, C. J. Boushey, and E. J. Delp, "An image analysis system for dietary assessment and evaluation," *Proc. - Int. Conf. Image Process. ICIP*, pp. 1853–1856, 2010.
40. M. N. Bin Razali, N. Manshor, A. A. Halin, N. Mustapha, and R. Yakob, "An integration of minimum local feature representation methods to recognize large variation of foods," *AIP Conf. Proc.*, vol. 1891, no. October, 2017.
41. M. Alkhawlani, M. Elmogy, and H. Elbakry, "Content-Based Image Retrieval using Local Features Descriptors and Bag-of-Visual Words," *Int. J. Adv. Comput. Sci. Appl.*, vol. 6, no. 9, pp. 212–219, 2015.
42. M. Yousuf *et al.*, "A Novel Technique Based on Visual Words Fusion Analysis of Sparse Features for Effective Content-Based Image Retrieval," *Math. Probl. Eng.*, vol. 2018, 2018.
43. X. Yuan, J. Yu, Z. Qin, and T. Wan, "A SIFT-LBP image retrieval model based on bag of features," *Int. Conf. Image Process.*, pp. 1061–1064, 2011.
44. Y. Wicaksono and V. Suhartono, "Color and Texture Feature Extraction Using Gabor Filter - Local Binary Patterns for Image Segmentation with Fuzzy C-Means," *J. Intell. Syst.*, vol. 1, no. 1, pp. 15–21, 2015.
45. E. Prakasa, "Ekstraksi Ciri Tekstur dengan Menggunakan Local Binary Pattern Texture Feature Extraction by Using Local Binary Pattern," vol. 9, no. 2, pp. 45–48, 2016.
46. A. Humeau-Heurtier, "Texture feature extraction methods: A survey," *IEEE Access*, vol. 7, pp. 8975–9000, 2019.
47. G. Ciocca, P. Napoletano, and R. Schettini, "Learning CNN-based Features for Retrieval of Food Images," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10590 LNCS, pp. 426–434, 2017.
48. D. P. Tian, "A review on image feature extraction and representation techniques," *Int. J. Multimed. Ubiquitous Eng.*, vol. 8, no. 4, pp. 385–395, 2013.
49. H. Hoashi, T. Joutou, and K. Yanai, "Image recognition of 85 food categories by feature fusion," *Proc. - 2010 IEEE Int. Symp. Multimedia, ISM 2010*, pp. 296–301, 2010.
50. M. M. Anthimopoulos, L. Gianola, L. Scarnato, P. Diem, and S. G. Mougiakakou, "A food recognition system for diabetic patients based on an optimized bag-of-features model," *IEEE J. Biomed. Heal. Informatics*, vol. 18, no. 4, pp. 1261–1271, 2014.
51. M. Y. Chen *et al.*, "Automatic Chinese food identification and quantity estimation," *SIGGRAPH Asia 2012 Tech. Briefs, SA 2012*, no. December 2017, 2012.
52. Y. Matsuda, H. Hoashi, and K. Yanai, "Recognition of multple-food images by detecting candidate regions," *Proc. - IEEE*

Int. Conf. Multimed. Expo, pp. 25–30, 2012.
53. Z. Y. Ming, J. Chen, Y. Cao, C. Forde, C. W. Ngo, and T. S. Chua, "Food photo recognition for dietary tracking: System and experiment," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10705 LNCS, no. January, pp. 129–141, 2018.
54. P. Pouladzadeh, S. V. B. Peddi, P. Kuhad, A. Yassine, and S. Shirmohammadi, "A virtualization mechanism for real-time multimedia-assisted mobile food recognition application in cloud computing," *Cluster Comput.*, vol. 18, no. 3, pp. 1099–1110, 2015.
55. V. Bettadapura, E. Thomaz, A. Parnami, G. D. Abowd, and I. Essa, "Leveraging context to support automated food recognition in restaurants," *Proc. - 2015 IEEE Winter Conf. Appl. Comput. Vision, WACV 2015*, pp. 580–587, 2015.
56. I. Process and C. Academy, "Luis Herranz , Ruihan Xu , Shuqiang Jiang," 2015.
57. A. Śluzek, "Machine Vision in Food Recognition: Attempts to Enhance CBVIR Tools," *Position Pap. 2016 Fed. Conf. Comput. Sci. Inf. Syst.*, vol. 9, pp. 57–61, 2016.
58. A. Singla and L. Yuan, "Food / Non-food Image Classification and Food Categorization using Pre-Trained GoogLeNet Model."
59. M. Bolaños, M. Garolera, and P. Radeva, "Active labeling application applied to food-related object recognition," *CEA 2013 - Proc. 5th Int. Work. Multimed. Cook. Eat. Act.*, no. October, pp. 45–50, 2013.
60. R. Liu, "Food Recognition and Detection with Minimum Supervision," no. March, 2016.
61. P. Pandey, A. Deepthi, B. Mandal, and N. B. Puhan, "FoodNet: Recognizing Foods Using Ensemble of Deep Networks," *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1758–1762, 2017.
62. J. P. Koirala, "Food Object Recognition : An Application of Deep Learning," p. 64, 2018.
63. G. Ciocca, P. Napoletano, and R. Schettini, "Food Recognition: A New Dataset, Experiments, and Results," *IEEE J. Biomed. Heal. Informatics*, vol. 21, no. 3, pp. 588–598, 2017.
64. K. T. Islam, S. Wijewickrema, M. Pervez, and S. O'Leary, "An Exploration of Deep Transfer Learning for Food Image Classification," *2018 Int. Conf. Digit. Image Comput. Tech. Appl. DICTA 2018*, no. December, 2019.
65. M. A. Subhi, S. H. Ali, and M. A. Mohammed, "Vision-Based Approaches for Automatic Food Recognition and Dietary Assessment: A Survey," *IEEE Access*, vol. 7, pp. 35370–35381, 2019.
66. M. M. K. Sarker *et al.*, "CuisineNet: Food Attributes Classification Using Multi-Scale Convolution Network," *Front. Artif. Intell. Appl.*, vol. 308, pp. 365–372, 2018.
67. K. Yanai and Y. Kawano, "FOOD IMAGE RECOGNITION USING DEEP CONVOLUTIONAL NETWORK WITH PRE-TRAINING AND FINE-TUNING Keiji Yanai Yoshiyuki Kawano Department of Informatics , The University of Electro-Communications , Tokyo , Japan," *Int. Conf. Multimed. Expo Work. . IEEE*, pp. 1–6, 2014.
68. S. Jahan, S. R. H, and S. A. Quadri, "Bird ' s Eye Review on Food Image Classification using Supervised Machine Learning," *Int. J. Latest Technol. Eng. Manag. Appl. Sci.*, vol. VII, no. Iii, pp. 153–159, 2018.
69. G. M. Farinella, D. Allegra, M. Moltisanti, F. Stanco, and S. Battiato, "Retrieval and classification of food images," *Comput. Biol. Med.*, vol. 77, pp. 23–39, 2016.
70. L. Pan, S. Pouyanfar, H. Chen, J. Qin, and S. C. Chen, "DeepFood: Automatic Multi-Class Classification of Food Ingredients Using Deep Learning," *Proc. - 2017 IEEE 3rd Int. Conf. Collab. Internet Comput. CIC 2017*, vol. 2017-Janua, pp. 181–189, 2017.
71. M. A. Stergios Christodoulidis and and Stavroula Mougiakakou, "Food Recognition for Dietary Assessment Using Deep Convolutional Neural Networks," in *International Conference on Image Analysis and Processing*, 2015, pp. 458–465.
72. N. Tammachat and N. Pantuwong, "Calories analysis of food intake using image recognition," *Proc. - 2014 6th Int. Conf. Inf. Technol. Electr. Eng. Leveraging Res. Technol. Through Univ. Collab. ICITEE 2014*, 2014.
73. C. Liu *et al.*, "A New Deep Learning-Based Food Recognition System for Dietary Assessment on An Edge Computing Service Infrastructure," *IEEE Trans. Serv. Comput.*, vol. 11, no. 2, pp. 249–261, 2018.
74. S. Yang, M. Chen, D. Pomerleau, and R. Sukthankar, "PSID_Yang2010," *Cvpr*, 2010.
75. Y. Kawano and K. Yanai, "Real-time mobile food recognition system," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, pp. 1–7, 2013.
76. F. Kong and J. Tan, "DietCam: Regular shape food recognition with a camera phone," *Proc. - 2011 Int. Conf. Body Sens. Networks, BSN 2011*, pp. 127–132, 2011.
77. and E. J. D. Marc Bosch, Fengqing Zhu, Nitin Khanna, Carol J. Boushey, "Combining global and local features for food identification in dietary assessment," in *18th IEEE International Conference on Image Processing*, 2011.
78. P. C, V. E Salis, J. G, and L. A, "A Comparative Study of Indian Food Image Classification Using K-Nearest-Neighbour and Support-Vector-Machines," *Int. J. Eng. Technol.*, vol. 7, no. 3.12, p. 521, 2018.
79. P. Pouladzadeh, S. Shirmohammadi, and A. Yassine, "Using graph cut segmentation for food calorie measurement," *IEEE MeMeA 2014 - IEEE Int. Symp. Med. Meas. Appl. Proc.*, no. August, 2014.
80. A. Kumar, "Trends of Occupational Pattern in India : An analysis NSSO," vol. 8, no. 1, pp. 463–476, 2018.