Computer Assisted Methods in Engineering and Science, **26**: 177–189, 2019, doi: 10.24423/cames.239

Image segmentation and classification with application to dietary assessment using BMI-calorie calculator

S. Jasmine Minija[1*], W.R. Sam Emmanuel[2]

Nesamony Memorial Christian College
Affiliated to Manonmaniam Sundaranar University
Abisekapatti, Tirunelveli, Tamil Nadu, India, 627012
[1]*Research Scholar (Reg no: 12487),* [2]*Associate Professor*
[*]*Corresponding Author e-mail: minijakenson@gmail.com*

Nowadays, people are more interested in their health by maintaining a proper diet. Today's lifestyle causes obesity and malnutrition in humans because of an uncontrolled diet. This paper proposes the health monitoring system using the body mass index (BMI) calorie calculator, which guides people to take proper calories from their daily diet. The image processing steps segmentation, features extraction, and recognition are used in the dietary assessment to identify the food items. The improved performance of the multi-hypotheses image segmentation (MHS) and feed-forward neural network (FFNN) classifier for nutritional assessment was evaluated using macro average accuracy (MAA) and standard accuracy (SA) metrics, which provide an enhanced classification rate.

Keywords: segmentation, feature extraction, classification, calorie estimation.

1. Introduction

Recently, the dietary assessment [9] has become the most popular system for assessing individual food consumption. People around the world are very interested in reducing their weight by measuring calories and nutrients values. Therefore, measuring these indicators is very important for observing the body fat. In the estimation of dietary, the determination of calorie measurement's accuracy is conducted using image segmentation, mining of features and food classification [1]. Some image segmentation methodologies are mining the desired objects from an image. However, the performance of segmentation is unsatisfactory due to region feature extraction. In this paper, the MHS along with FFNN) classifier is considered to evaluate the dietary assessment for increasing the performance rate [8]. First, the segmentation process is performed on the input image and is used to determine the regions, where a corresponding food item is placed by salient region detection, multi-scale segmentation, as well as fast rejection. Then, the important features of food items are mined using the global and local feature extraction. Here, the global feature extraction is done using a color and texture descriptor, but the extraction of local features is carried through its local neighborhood pixel. After the features are attained, the classification procedure is performed for every segmented region by an FFNN approach.

Finally, the calorie value is estimated based on the volume of the food area and the measure of calories and nutrition from mass value. Calculating the calorie value of each food item helps to calculate dietary assessment of the person's accurate dietetic consumption. In the end, the evaluation of the experimental outcomes along with the performance analysis is validated, and a better classification performance is achieved with improved accuracy.

2. Literature review

Over the past few years, beverage or food intake has been considered to be an open research problem in the field of nutrition and health. Perfect diet measurement is a challenging task, especially for adolescents. Nevertheless, some researchers have proposed various methods for food segmentation and classification. However, these days, calorie measurement becomes a challenging problem. Thus, the research community is lacking in the area of training and testing of the classification phase. The vision-based measurement (VBM) system has improved the accuracy of food identification performance [12]. By using a smartphone for the dietary assessment system, the user had to capture the food item that was to be consumed to measure calories automatically. In general, the test image was used directly for recognition and classification, whereas the training images were introduced into the pre-processing to remove the unwanted pixels or area from the image. To solve the noise issue, Kawano and Yanai [3] developed the set of food images along with its category of names. The "foodness" classifier and adaptive support vector machine (SVM) were used to suppress the noisy pixels in the image. Thus, the rigid, predefined, and deformable shape are considered as the three appearance properties for food categorization.

The two aspects are usually performed in food classification. First, the features were extracted from the food item or segmented region, which was then used by the classifier for the decision-making module where the class labels of each food item were assigned [13]. The feature vector exhibits the significant representation of the image that could differentiate the objects which belong to the different classes (interclass discrimination). Moreover, the information about two objects of two classes was combined and classified together. Once the feature vector is obtained, a classification system exploits this feature to recognize or classify [10].

Besides, food estimation is also an essential criterion in dietary assessment. The mobile telephone food record system was explained where the correct amount of food intake and nutrients were calculated [14]. Several image segmentation algorithms were developed and applied to quantitative image analysis. Still, the problem of automatic food segmentation is not resolved due to several complex issues in food images, such as a variety of color, texture, and shape, which are often difficult to describe and specify their boundaries.

3. Methodology

The MHS and FFNN classifiers [8] are used here to classify the food items from the input food images and the outputs are compared with the existing methods. This methodology includes four methods: 1) segmentation, 2) feature extraction, 3) FFNN classifier, and 4) calorie estimation. Figure 1 shows the workflow of this paper. Here, the input image is taken from the public dataset

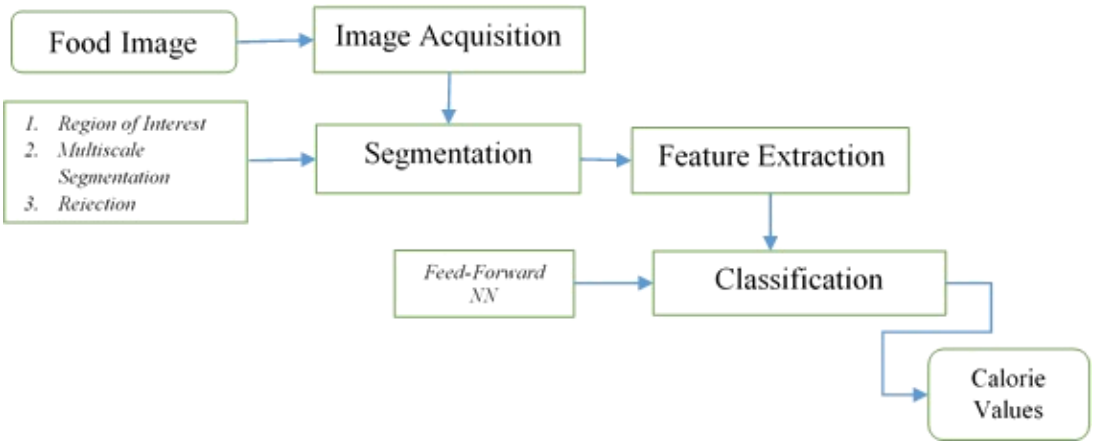

Fig. 1. The workflow of the dietary assessment methodology.

UNIMIB2016 [5]. Once a food image is given as input, the MHS is used to identify the object boundaries for the food items in the input image. This method achieves various segmented regions and then selects the constant segmented areas. The three types of MHS are salient region detection, multi-scale segmentation, and fast rejection. Once the food items of every input image are segmented, the food features of global features and local features are acquired by the feature extraction method based on color, texture and local neighborhood pixel. Then, these features are combined to form the single feature vector. Finally, the FFNN classifier exploits the feature space for the classification performance. Once the food items are classified from the input image, calorie value is computed by the food area and food volume [8].

3.1. Segmentation

The input image undergoes the segmentation process, where the food items get significantly segmented. The method of segmentation takes place through three steps: salient region, multi-scale segmentation, and fast rejection. Further, the MHS is used for segmenting the food image in Zhu *et al.* [13] that comprises of local and global features of food images.

Region of interest detection

Here, the class labels are properly assigned for each image pixel. This is used for reducing the pixel count to avoid trouble by removing the non-food items, including utensils, tablecloths, flowers, etc. Further, the edge regions also get extracted over the Canny edge detector [4], and that is used for providing the normalized edge histogram. Here, the edge pixel's strength is local and gets raised in the gradient direction. Then, both the background and edge pixels are greatly assisted by the Canny operator. This corresponding distance measure is evaluated using the edge histogram a and uniform distribution of regions b. Finally, the threshold value is obtaines which is used for distinguishing the region of interest. Equation (1) defines the Euclidean distance measure:

$$ED = D(a,b) = \sqrt{\sum_{i=1}^{N} (a_i - b_i)^2}. \tag{1}$$

Multi-scale segmentation

In the normalized cut, the image is signified as the weight as well as an undirected graph, in which every pixel's pair generates the pixel nodes. Then the image is subdivided into some count of disjoint sets, which leads to justifying the edge weights. The calculation of a similarity measure among pixel values is done through the normalized cut ratio given in Eq. (2):

$$w_e(x,y) = \exp\left[-\left(\frac{\|e_x - e_y\|_2^2}{\sigma_e^2} + \frac{\|L_x - L_y\|_2^2}{\sigma_L^2}\right)\right], \tag{2}$$

where e_x indicates the pixel intensity of the image and L_x indicates the i-th position of the pixel.

Consequently, in Eq. (3), the edge image's scale is used for the affine assessment of each pixel value, and that leads to providing the potential object boundary:

$$w_A(x,y) = \exp\frac{-\max_{l \in \text{line}(x,y)} \|\text{Edge}(l)\|^2}{\sigma_A^2}, \tag{3}$$

where $\text{Edge}(l)$ designates the edge strength at the l-th position, and joining of two pixels are shown by $\text{line}(x,y)$.

Hence, the two affinity measures get united by α constant. This is evaluated as in Eq. (4):

$$w_{\text{comb}}(x,y) = \sqrt{w_e(x,y) \times w_A(x,y)} + \alpha w_A(x,y). \tag{4}$$

The combination measure specified in Eq. (4) and the spatial separation, the measure of affinity termed $w(x,y)$ provides dissimilar image appearances. Hence, the image gets segmented through various scales of the spatial region, and it is given in Eq. (5):

$$w_{\text{full}} = w_1 + w_2 \approx w_1 + X_{1,2}^T w_2 X_{1,2} = w_{\text{reconstruction}}, \tag{5}$$

where w_i grants the affinity measure between pixels and further utilizes the interpolation matrix of two $X_{1,2}$ scales.

Rejection

The rapid rejection phase is used for solving the issue of redundancy in the segmented area of the food items. In this removal step, some pixel values that have no important data that describes the object classes are rejected. Therefore, the background pixel in the salient region is allocated, which leads to improving the classification enactment.

3.2. Feature extraction

After the segmentation process, the image consequently goes through the process of feature extraction as in [10]. In this process, the technique of global and local feature extraction is applied for the extraction of food features from the segmented image.

Global feature extraction

The texture and color features are extracted to improve the performance of classification. This feature is often used for the demonstration of shape characteristics in the image. Hence, the global feature space consists of the texture and color descriptors.

Color descriptor

The color feature is an important feature that differentiates food items from the input image. The input image from the dataset shows a large color deviation. For example, the color of orange juice is orange, and milk seems to be white color. Therefore, the color characteristic distinguishes the food items, in which the features from segmented food items are extracted significantly. Further, the global color descriptor is achieved by statistics like entropy, global as well as predominant color statistics.

The global color feature is attained from each color space component, and that is RGB, as well as HSV color spaces. The entropy-related feature vector is attained by restraining the statistics of two-moment characteristics. The corresponding features are achieved by Eq. (6) using common representative color in RGB space and it is defined as follows:

$$F^C = \{(C_1 P_1 V_1), ..., (C_p P_p V_p)\}, \tag{6}$$

where C_p specifies the representation of the 3D color vector in the segmented image, P_p denotes the percentage value and V_p is the region's variance measure.

Texture descriptor

This can be obtained by using: 1) gradient orientation spatial dependence matrix (GOSDM), 2) entropy categorization and fractal dimension (EFD) and 3) Gabor-based image decomposition and fractal dimension (GFD) assessment.

In general, the orientation indicates the angular direction of the image. The value of offset defines the gradient orientation of every segmented region. The application of an angular value of 0°, 45°, 90° and 135° is performed. Hence, the offset value includes segmented images' length as well as orientation. In Eq. (7), the feature vector of gradient orientation includes various statistics, including correlation (Cr), angular second moment (A), homogeneity (H), entropy (E) and contrast (Ct). This is indicated as in Eq. (7):

$$F^T = \left[f_{d_1}, f_{d_4}, f_{d_{16}}, ..., f_{d_{\min}(H/2,V/2)}\right], \tag{7}$$

where f_d denotes to the feature vector, and it is determined as given below:

$$f_d = Cr_0,\ A_0E_0, Ct_0, H_0, Cr_{45}, A_{45}E_{45}, Ct_{45}, H_{45}, Cr_{90}, A_{90}E_{90},$$
$$Ct_{90}, H_{90}, Cr_{135}, A_{135}E_{135}, Ct_{135}, H_{135}.$$

Local feature extraction

The local features are mined using local neighborhood pixels of the image points. The benefit of the local feature is its robustness near the image's occlusion and clutter. In addition, local feature information is attained by the scale-invariant features transform (SIFT) [7] and speeded up robust features (SURF) [2] descriptors. Then, the steerable filter [6] is also used for the extraction of the feature. This filter is used for extorting the feature from 1st and 2nd moment statistics through the local neighborhood pixel. Consequently, the SIFT descriptor is concerned with the color component of the RGB image. The extraction of local features from the input image is meant as F^L.

These feature sets are given as the input to the FFNN approach for categorizing the food segments. All the color, texture as well as shape are determined by the global description of food items. Initially, the color features are the discrimination property that is mined through statistics like entropy, global, and predominant color statistics. Additionally, the three global features are attained based on texture descriptors. Then, the EFD, GOSDM as well as GFD are used for computing the texture features. Thirdly, some other six local features are mined through steerable filter, SURF, SIFT, Red-SIFT, Green-SIFT, and Blue-SIFT. This feature vector is given into the developed FFNN classifier that finds the class labels of the segmented food items:

$$F = \left\{F^C \,||\, F^T \,||\, F^L\right\}, \tag{8}$$

where F^C indicates the color feature, F^T refers to the texture feature, and F^L denotes the local feature. This is defined as in Eq. (9):

$$F = \{f_i; \quad 0 \le i \le p\}, \tag{9}$$

where p refers to the count of features.

3.3. The FFNN classifier to identify the food segments

After the features get extracted, the classifier uses this feature for the classification of food items, and then it uses it for the calculation of calorie value. This FFNN [11] is a 3-layered network that incorporates an input layer, hidden layer as well as the output layer. The FFNN classifier comprises of two phases: the training phase and the testing phase. In this, the training algorithm

is the most vital aspect of classification, along with food item recognition. At first, the training images are specified into the system of dietary evaluation, in which the segmentation, along with feature extraction, is done. Then, the features are introduced, that is they are trained by the FFNN classifier. Although in the testing phase of the FFNN, the features attained from testing images are given to the classifier. The testing image's food features are compared to the trained feature set for achieving the desired classified food items.

The FFNN classifier [11] is driven by the neurons that are greatly related to/in the input layer and output layer for solving the distinct problem. The global and local features of the segmented food item are specified as the input to the FFNN approach. Subsequently, the classification of features is done by the hidden neurons that achieve the output for several food item classes of objects. Hence, the FFNN classifier is assessed by Eq. (10):

$$c(j) = \sum_{i=1}^{p} w_i(j) \cdot f_i(j); \quad 1 \le i \le p, \tag{10}$$

where i shows a count of input feature and $c(j)$ denotes the classified output for a j-th food item.

From each and every food image, the feature f is extracted. The hidden layer is also termed as the interconnection layer, in which the input features do multiplication with the values of weight and perform the classification of food items. Similarly, in the phase of training, the food feature's training dataset is already classified for the necessary result of food type. At the testing phase, the result of each food feature is compared to the features of the testing phase. With this comparison, the used FFNN classifies the respective food items. After the completion of food item classification, the segmented food items are processed for the measurement of calories.

3.4. Calorie estimation

After the completion of the FFNN food items classification, the classified outcomes are used for the evaluation of the calorie value. Hence, the calorie value is calculated using the formulation of food area volume and estimation of calorie values.

Food area volume computation

The square grid is used to evaluate the surface area of the food items in the image. Every square grid includes equivalent pixel counts that indicate the equal/same area of food items. The food item volume is formulated through the total area and the depth of food image in Eq. (11):

$$TA = \sum_{i=1}^{n} T_i, \tag{11}$$

where n indicates the total number of square grids in the segmented image. At last, Eq. (12) is used for computing the food area volume:

$$V = TA \times d, \tag{12}$$

where TA represents the total food item area, and$d$ denotes the depth. Also, the volume formulation is a vital aspect for measuring the food item mass.

Calorie value estimation

The measurement of calorie value is done by the mass and volume of the food item. The mass is calculated by the product of density and volume estimated from the food items, which is well defined in Eq. (13):

$$M = \rho V, \tag{13}$$

where M denotes the mass, ρ shows the density value, and V indicates the volume.

At last, the calorie value and nutrition value of food segments are defined as in Eq. (14):

$$\text{Calorie} = \frac{\text{calorie in table} \times \text{mass in image}}{\text{mass from table}}. \tag{14}$$

BMI (body mass index)

The height and weight are calculated to find the BMI, which can be estimated by the Eq. (15):

$$\text{BMI} = \left(\frac{\text{weigh in kilograms}}{\text{height in meters} \times \text{height in meters}}\right). \tag{15}$$

4. Results and discussions

4.1. Experimental setup

The performance of the MHS-FFNN method is analyzed using the UNIMIB2016 dataset [5], which contains 1027 tray images with 73 diversities of food items in a separate tray. In this paper, the enactment of the MHS-FFNN approach was carried out by different evaluation metrics. The input image was exposed to the phase of segmentation. Then it was fed into the processes, namely feature extraction and classification. Furthermore, the accuracy was also a significant measure for the purpose of classification as the greater accuracy level delivers effective classification performance.

The truly positive proportion measure was the percentage of actual positive results which are accurately identified by the classifier. Meanwhile, the false-positive rate was also a similar probability value measure that was falsely rejected by the classifier. The assessments of both TPR and FPR are shown in Eqs (16) and (17):

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{16}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}, \tag{17}$$

where FP indicates the false-positive, TP represents the true-positive, TN refers to the true- negative, and FN denotes the false-negative rate. The segmented image results are given in Fig. 2.

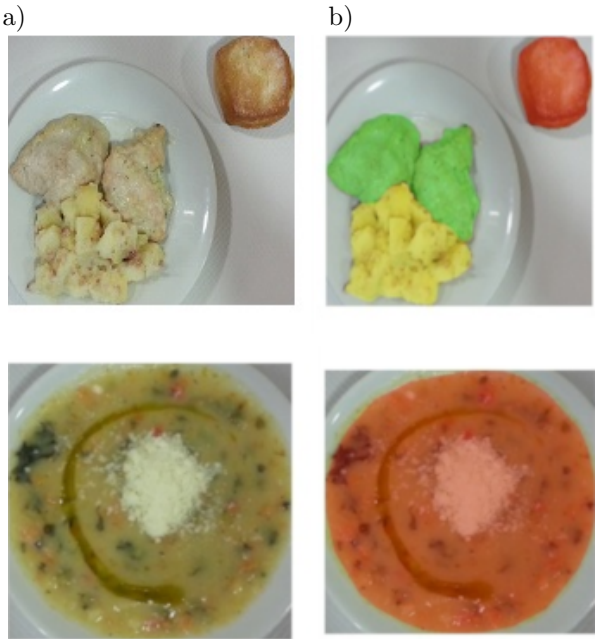

Fig. 2. Image results (a) original sample image and (b) segmented image.

4.2. Performance analysis based on standard accuracy

The segmented accuracy for the sample of 200 images taken from the dataset using the MHS-FFNN model is given in Fig. 3.

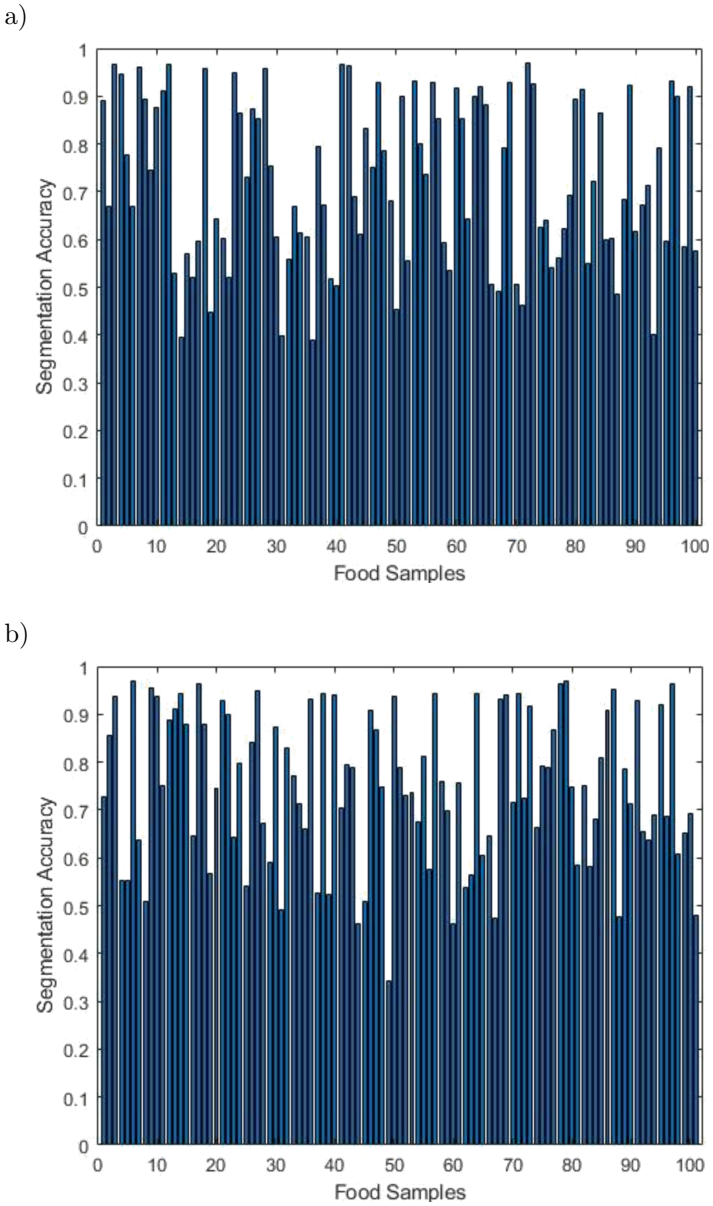

Fig. 3. Segmentation accuracy for tray images: a) $f = 1$ to 100; b) $f = 101$ to 200.

4.3. Performance estimation based on macro average accuracy

Macro average accuracy (MAA) is the proportion of the true positives of a segmented food item to the total number of the true positives in the segmented food item. The performance analysis of MAA is shown in Fig 4, and MAA by varying training percentage is given in Fig. 4a. Here, for training percentage 0.6, the MHS and FFNN technique reaches MAA values of 0.8409, 0.8413, 0.8482, and 0.8669 with the IS of image size 64×64, 128×128, 256×256, and 512×512, respectively. Similarly, for training percentage 0.65, the MHS and FFNN method reaches 0.8640, 0.8712, 0.8784, and 0.8976 values of MAA with image size 64×64, 128×128, 256×256, and 512×512, respectively.

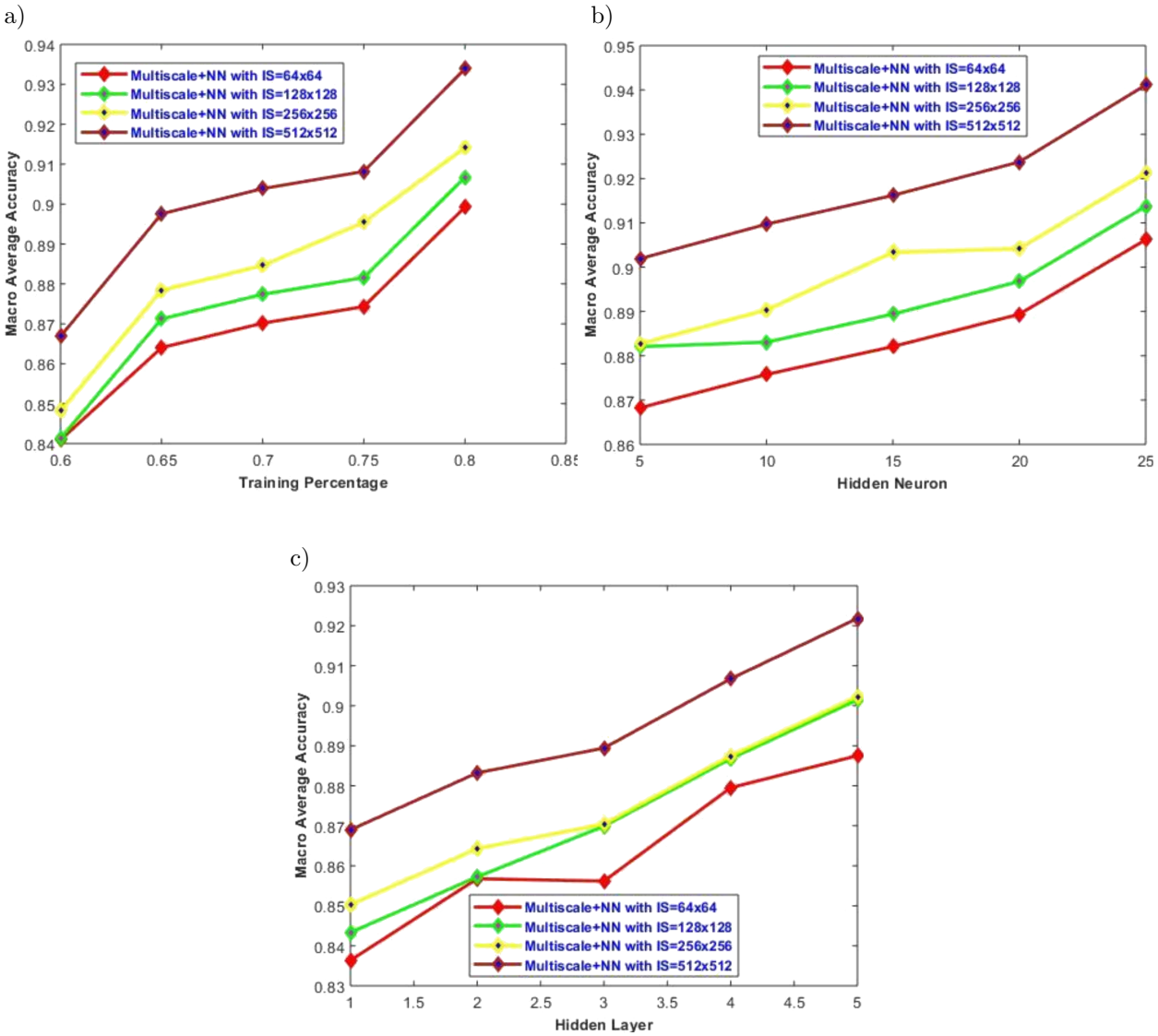

Fig. 4. Performance estimation based on MAA: a) training percentage; b) hidden neuron; c) hidden layer.

The analysis of MAA by varying hidden neuron is given in Fig. 4b. From the graph, it is identified that the MHS and FFNN method for the hidden neuron 5 is 0.8682, 0.8821, 0.8827, and 0.9019 and then, for hidden neuron 10, the MHS and FFNN method attains the MAA values of 0.8758, 0.8831, 0.8903, and 0.9097 with IS of image size 64×64, 128×128, 256×256, and 512×512, respectively. The analysis by varying the hidden layer is shown in Fig. 4c. In this, it is observed that for hidden layer 1, the MHS and FFNN technique achieves MAA values of 0.8364, 0.8433, 0.8503, and 0.8689. Similarly, in the same condition, at hidden layer 2, the MHS and FFNN method attains the MAA values of 0.8568, 0.8572, 0.8643, and 0.8832 with IS of image size 64×64, 128×128, 256×256, and 512×512.

4.4. Performance analysis based on standard accuracy

Standard accuracy (SA) is the ratio of the average of the accurately identified tray images to the entire number of tray images. The results of the MHS and FFNN model in terms of SA are shown in Fig. 5, and, the performance by varying training percentage is given in Fig. 5a. Here, at 0.6

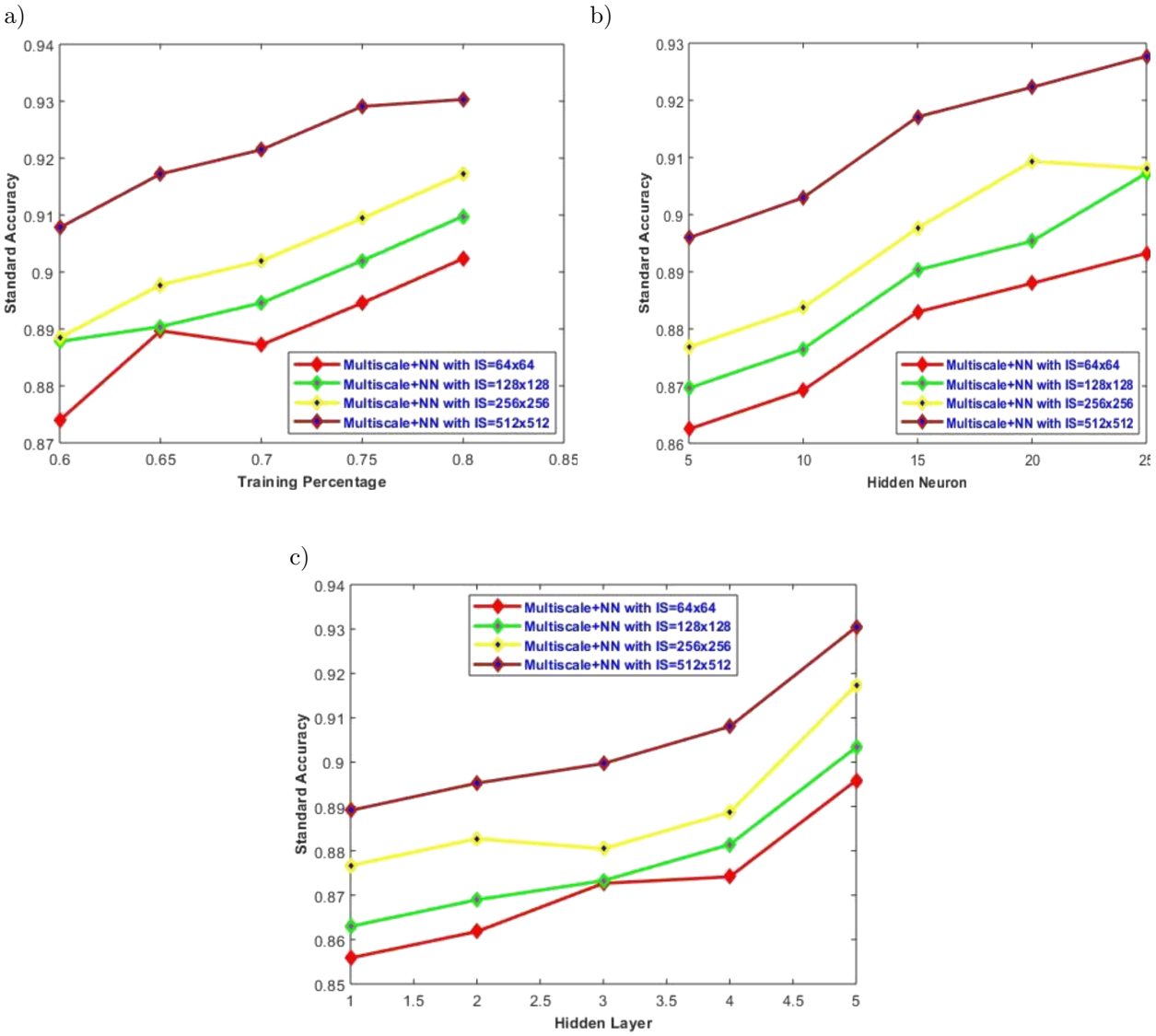

Fig. 5. Performance evaluation based on SA: a) training percentage; b) hidden neuron; c) hidden layer.

training, the MHS and FFNN method achieves the SA values of 0.8739, 0.8878, 0.8885, and 0.9078 and at 0.65 training percentage, the FFNN model attains 0.8897, 0.8904, 0.8977, and 0.9172 values of SA with IS of image size 64×64, 128×128, 256×256, and 512×512, respectively. Similarly, in Fig. 5b for hidden neuron 5, the MHS and FFNN method attains the SA value of 0.8625, 0.8696, 0.8768, and 0.8959. Further, for hidden neuron 10, the developed model attains the SA values of 0.8693, 0.8765, 0.8837, and 0.9029 for IS of image size 64×64, 128×128, 256×256, and 512×512, respectively. Then, Fig. 5c reviews that for one hidden layer, the developed method attains SA values of 0.8559, 0.8629, 0.8767, and 0.8891 with IS of image size 64×64, 128×128, 256×256, and 512×512, respectively. At hidden layer 2, the developed method attains the SA values of 0.8618, 0.8689, 0.8827, and 0.8953 with IS of image size 64×64, 128×128, 256×256, and 512×512.

4.5. Performance analysis based on mean-square error

The performance evaluation of the MHS and FFNN model in terms of mean-square error (MSE) is given in Fig. 6.

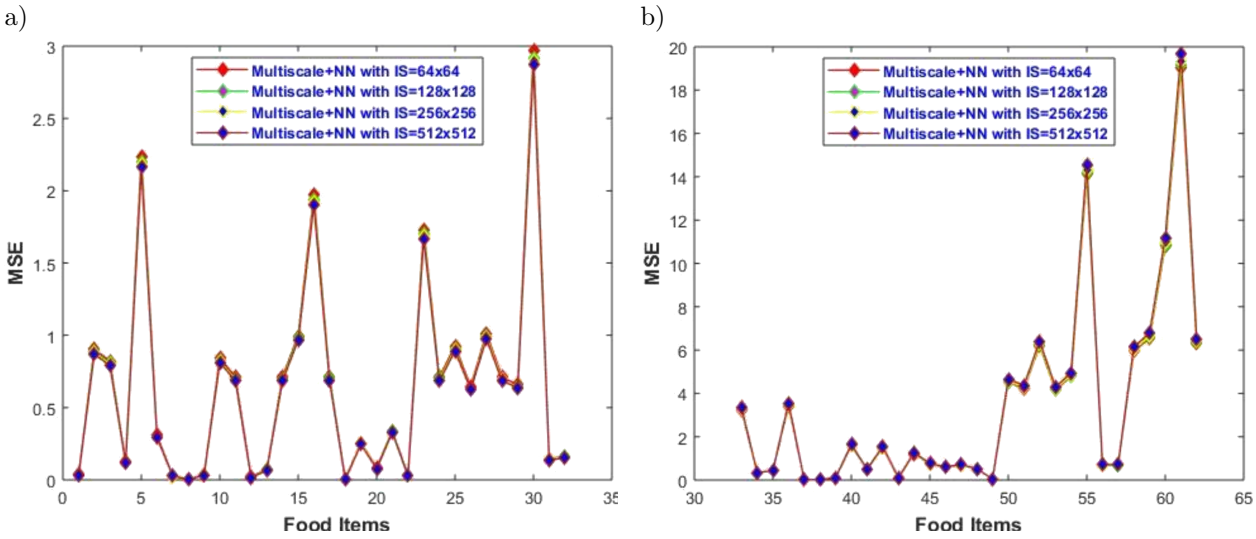

Fig. 6. Performance estimation based on the MSE.

5. Discussion

The comparison was made with the various existing models [5] and [8]. The k-nearest neighbor (k-NN) and SVM classifiers were used for the identification of food items. The multiple segmentation approaches used were the ground truth segmentation (GT), global approach (G), local and patch-based approach, the combination of the sum of the posteriors (G + P), and the combination of the product of the posteriors (G X P). The performance of each model was analyzed with the SA and the MAA metric. The performance of the model varied based on the feature extraction methods. The presented work used the feature extraction techniques, such as LBP (local binary patterns), CEDD (color and edge directivity descriptor), histogram, Gabor, CNN128 (128-dimensional feature), CNN4096 (4096-dimensional feature), and BoCFR (bag of convolutional filter response) [11].

Table 1 shows the comparative analysis of the system with the various existing models based on the SA metric. In comparison with the existing methods, the MHS and FFNN method shows better performance in KNN + GT + P, SVM + GT + G, SVM + GT + P, SVM + GT + (G X P) with the corresponding rate of 0.79, 0.86, 0.852, 0.879. The SA metric for the MHS and FFNN method, combined with the MHS, reaches a better rate of 0.945. Also, the MHS and FFNN classifier with

Table 1. Comparison of the MHS and FFNN method with the existing models based on the SA metric [8].

Methods	SA values for various feature extraction methods							
	LBP	CEDD	Hist	Gabor	CNN128	CNN4096	BoCFR	Proposed feature
k-NN + GT + G	0.394	0.446	0.628	0.427	0.748	0.820	0.761	0.7
k-NN + GT + P	0.543	0.656	0.719	0.682	0.745	0.774	0.734	**0.79**
k-NN + GT + (G + P)	0.504	0.629	0.732	0.641	0.814	0.885	0.811	0.8
k-NN + GT + (G X P)	0.437	0.536	0.705	0.518	0.805	0.858	0.791	0.69
SVM + GT + G	0.480	0.520	0.643	0.456	0.774	0.825	0.756	**0.86**
SVM + GT + P	0.646	0.718	0.759	0.711	0.816	0.857	0.763	**0.852**
SVM + GT + (G + P)	0.672	0.700	0.721	0.698	0.872	**0.891**	0.743	0.87
SVM + GT + (G X P)	0.565	0.619	0.642	0.576	0.816	0.858	0.722	**0.879**
MHS-FFNN approach	**0.694**	0.692	**0.772**	**0.768**	**0.889**	0.879	**0.852**	**0.945**

the existing features of LBP, Hist, Gabor, CNN128 and BoCFR gets a better percentage of 0.694, 0.772, 0.768, 0.889, and 0.852.

Table 2 shows the comparative analysis of the dietary assessment system with the various existing models based on the MAA metric. The comparison with the MHS and FFNN method shows a better performance in KNN + GT + P, SVM + GT + G, SVM + GT + P, SVM + GT + (G + P), with the relevant rate of 0.596, 0.763, 0.637, 0.77, 0.71. The MAA value for FFNN, combined with the MHS, reaches a better rate of 0.937. With these, the FFNN with the existing features of LBP, CEDD, Hist, Gabor, CNN128, CNN4096, and BoCFR produces a better rate of 0.5, 0.565, 0.682, 0.679, 0.758, 0.8, 0.815.

Table 2. Comparison of the MHS and FFNN models based on the MAA metric [8].

Methods	MAA values for various feature extraction methods							
	LBP	CEDD	Hist	Gabor	CNN128	CNN4096	BoCFR	Proposed feature
k-NN + GT + G	0.171	0.219	0.380	0.192	0.555	0.652	0.559	0.603
k-NN + GT + P	0.221	0.312	0.505	0.367	0.464	0.500	0.510	**0.596**
k-NN + GT + (G + P)	0.210	0.313	0.493	0.377	0.586	0.631	0.571	0.624
k-NN + GT + (G X P)	0.222	0.273	0.457	0.291	0.611	0.685	0.614	0.674
SVM + GT + G	0.231	0.249	0.375	0.234	0.552	0.644	0.489	**0.763**
SVM + GT + P	0.346	0.405	0.518	0.388	0.541	0.575	0.505	**0.637**
SVM + GT + (G + P)	0.419	0.444	0.505	0.470	0.677	0.684	0.508	**0.77**
SVM + GT + (G X P)	0.322	0.370	0.434	0.359	0.670	0.687	0.557	**0.71**
MHS-FFNN approach	**0.5**	**0.565**	**0.682**	**0.679**	**0.758**	**0.8**	**0.815**	**0.937**

6. Conclusions

In this paper, the food image was concerned as the input image for the developed system. Initially, the input image was subjected to the segmentation phase. Then, the outcomes of segmentation were attained through the detection of a region of interest, multi-scale segmentation and rejection. Here, the FFNN classifier was used to carry out the classification process. Finally, the investigational results were evaluated based on SA and MAA metrics which attains better classification rate.

References

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

[2] H. Bay, T. Tuytelaars, L. Van Gool. *SURF: Speeded Up Robust Features.* [In:] Leonardis A., Bischof H., Pinz A. [Eds], Computer Vision – ECCV 2006. Lecture Notes in Computer Science, vol. 3951, Springer, Berlin, Heidelberg, 2006, doi: /10.1007/11744023_32.

[3] Y. Kawano, K. Yanai. Automatic Expansion of a Food Image Dataset Leveraging Existing Categories with Domain Adaptation. [In:] *Proceedings of ECCV Workshop on Transferring and Adapting Source Knowledge in Computer Vision* (TASK-CV), pp. 1–16, 2014, doi: 10.1007/978-3-319-16199-0_1.

[4] J. Canny. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, **8**(6): 679–698, 1986, doi: 10.1109/TPAMI.1986.4767851.

[5] G. Ciocca, P. Napoletano, R. Schettini. Food recognition: a new dataset, experiments, and results. *IEEE Journal of Biomedical and Health Informatics*, **21**(3): 588–598, 2017, doi: 10.1109/JBHI.2016.2636441.

[6] W.T. Freeman, E.H. Adelson. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13**(9): 891–906, 1991, doi: 10.1109/34.93808.

[7] D.G. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, **60**(2): 91–110, 2004, doi: 10.1023/B:VISI.0000029664.99615.94.

[8] S.J. Minija, W.R.S. Emmanuel. Neural network classifier and multiple hypothesis image segmentation-for dietary assessment using calorie calculator. *The Imaging Science Journal*, **65**(7): 379–392, 2017, doi: 10.1080/13682199.2017.1356610.

[9] P. Pouladzadeh, G. Villalobos, R. Almaghrabi, S. Shirmohammadi. A novel SVM based food recognition method for calorie measurement applications. [In:] *Proceedings of 2012 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, pp. 495–498, 2012.

[10] M.H. Rahman *et al.* Food volume estimation in a mobile phone based dietary assessment system. [In:] *2012 8th IEEE International Conference on Signal Image Technology and Internet Based Systems*, pp. 988–995, 25–29 Nov. 2012, Naples, Italy, doi: 10.1109/SITIS.2012.146.

[11] K. Saravanan, S. Sasithra. Review on classification based on artificial neural networks. *International Journal of Ambient Systems and Applications*, **2**(4): 11–18, 2014, doi: 10.5121/ijasa.2014.2402.

[12] B.L. Six *et al.* Evidence-based development of a mobile telephone food record. *Journal of the American Dietetic Association*, **110**(1): 74–79, 2010, doi: 10.1016/j.jada.2009.10.010.

[13] F. Zhu, M. Bosch, N. Khanna, C.J. Boushey, E.J. Delp. Multiple hypotheses image segmentation and classification with application to dietary assessment. *IEEE Journal of Biomedical and Health Informatics*, **19**(1): 377–388, 2015, doi: 10.1109/JBHI.2014.2304925.

[14] A. Biem, S. Katagiri. Feature extraction based on minimum classification error/generalized probabilistic descent method. [In:] *Proceedings of IEEE International Conference on Acoustic, Speech, Signal Process*, pp. 275–278, Apr. 1993, doi: 10.1109/ICASSP.1993.319289.