Food recognition improvement by using hyper-spectral imagery

Shirin Nasr Esfahani[1*], Venkatesan Muthukumar[2], Emma E. Regentova[2], Kazem Taghva[1] and Mohamed Trabia[3]
Computer Science Department, University of Nevada, Las Vegas[1]
Electrical and Computer Engineering, University of Nevada, Las Vegas[2]
Mechanical Engineering Department, University of Nevada, Las Vegas[3]

Abstract
In the past decade, dietary assessment has been one of the most popular topics of research in the food industry, which has resulted in developing several automatic or semi-automatic dietary assessment systems using visible spectrum images for food recognition. However, the main shortcoming of visible spectrum image-based systems is its inability to differentiate foods of similar color. Researchers have added additional features such as shape, size and texture to the color model to improve the overall accuracy. However, the shape and size features are rendered inefficient when recognizing food in the mixed or cooked form. The aim of this research is to show the capability of hyperspectral bands for accurate food recognition based on individual spectral bands. In this work we use a hyperspectral imaging system of 240 spectral bands with the wavelength range between 400 nm to 900 nm. The ReliefF and PCA methods select/extract less, but the most informative features which are important to learn Logistic Regression and Support Vector Machine (SVM) as binary and multiple classifiers, respectively. A total of 20 different food samples in various forms (uncut and cut), shapes, and sizes were used in this study. The prediction results indicate that the hyperspectral images have the advantages of being able to recognize different food items from a mixed form with similar color and similar food types with different colors. In our experiments the highest classification accuracy of 0.6874 with 20 different food samples is produced by SVM multi-classification of ReliefF data with the top 110 hyperspectral features. We are able to obtain approximately 0.90 accuracy, using binary classification on a specific subset of food samples.

Keywords
Food recognition, Hyperspectral, Logistic regression, PCA, ReliefF, RGB, SVM.

1.Introduction
The primary objective of this paper, which is an extended version of the work presented in [1] is to determine the feasibility of exploiting spectral features in classification for automatic food recognition and quality/nutrition evaluation systems. Food intake is one of the necessary factors of human health, which means a human body should get enough nutrient and vitamins through one's daily diet in order to stay healthy. Nowadays people are more concerned about what they eat and how that affects their lifestyle. Furthermore, current eating behaviors and range of nutrition consumption has changed.

A tremendous amount of scientific research is published every year to show direct relationship between diet quality and many diseases such as type2 diabetes or/and different types of cancers.

Manual assessment of dietary intake methods and tools are not precise enough to have accurate nutritional measurements. Besides, there should be a nutrition specialist to analyze these measurements. The problem is exacerbated for hospitalized patients who must be screened for nutrition and dietary assessment on admission and discharging to and from the hospital. The assessment then can be used regularly thereafter for the patient's care and treatment. Therefore, various automatic dietary assessment techniques have been developed in the past decade. They significant focus on development in mobile devices and applications [2−12]. In most dietary assessment plans, users self-report food

23

consumed and portion-size from a database to the nutritional assessment system. The user can also take pictures of his/her own eating plate which can be used as a reference for the portion size. Also, the leftover amount of food can be documented.

Use of image processing and machine learning has greatly aided the process of automatic dietary assessment. Using machine learning, the food is classified based on the color features of objects [13−17].

However, identification of some food with visible colors such as black olive and black bean are impossible. To tackle this issue, some systems consider other features such as shape and texture in addition to color [18−20]. Classification accuracy has been improved by applying powerful texture and feature representations. However, researchers have to select a balance between quality and quantity of representation, and efficiency. The texture quality can easily be affected by changes in illumination, rotation, scale, blur, noise, etc. Efficiency related challenges include potentially large number of texture categories and their high dimension. Due to the large inhomogeneities of food materials there are many features to be considered, including shape invariant features such as area, length, width, aspect ratio, and perimeter.

The advantage of using hyperspectral imaging (HSI) systems rather than other conventional digital imaging devices (visible spectrum) is that it acquires spectral response of hundreds of different spectral bands at each spatial location instead of having grayscale or three-color intensities values. Combining optical imaging and hyperspectral technology provides concurrent spatial and 3D spectral data as shown in *Figure 1*.

Spectral reflectance is the amount of reflectance from an object which is measured as a function of wavelength and shows how much energy a surface reflects at a specific wavelength. Thus, more information is made available about objects being analyzed that can be used to uniquely recognize them. Hyperspectral data analysis is becoming increasingly a popular tool in food industry and agriculture [21]. It has been also used in a wide variety of applications ranging from food quality [22], food safety [23], and precision agriculture [24, 25].

Classification using hyperspectral cube data has been studied extensively, but there are still some major problems such as "curse of dimensionality", and limited number of training samples yet to be efficiently solved. Data transformation methods such as PCA is one of pre-processing steps to deal with high dimensionality of features [26]. Band selection is another important pre-processing technique which selects a subset of bands with a low correlation and representative information [27−30]. Some issues that come with high dimensional data occur during analyzing or visualizing the data, and some appear at the stage of training machine learning models. The "curse of dimensionality" is a general term that refers to a set of challenges that arise when working with high-dimensional data. Additionally, there is a problem known as Hughes phenomenon, as the number of features increases, the classifier's performance increases until the optimal number of features is reached. Adding more features to the training set will then degrade the classifier's performance.

In [31], Support vector machine (SVM) was used for classification of hyperspectral data. In addition to showing low sensitivity against "curse of dimensionality", SVM hardly suffers from Hughes phenomenon [32]. Since SVM-based classifiers are able to get higher classification accuracy, they have been the state-of-the-art method for hyperspectral data analysis for a long time.

In this paper, we have made two major improvements of our previous work [1]. First, we have extended our datasets to cover more types of foods with similar color as well as the same food types of different colors. Second, to obtain better classification accuracy, we apply PCA and ReliefF for feature selection and extraction methods to narrow down the feature space and determine which features are essential. The first research effort aimed at showing the potential of hyperspectral imaging for dietary assessment studies. The SVM classification is utilized for multi-classification task. We also use multiple logistic regression analyzer as a binary classifier to differentiate between specific class of foods such as beans, noodles, olives, and rice.

The rest of this paper is organized into 5 sections. Section 2 present the literature on food recognition systems using optical as well as hyperspectral images. Section 3 introduces the dimension reduction, classification methods, and Image and Data Acquisition and Pre-processing. In Section 4,

experimental results of the use of Support Vector Machines (SVMs) and Logistic Regression classification are demonstrated on original datasets, PCA features, and ReliefF data. In section 5, we discuss the advantage and disadvantage of using hyperspectral techniques. The conclusion section summarizes the results of this paper.
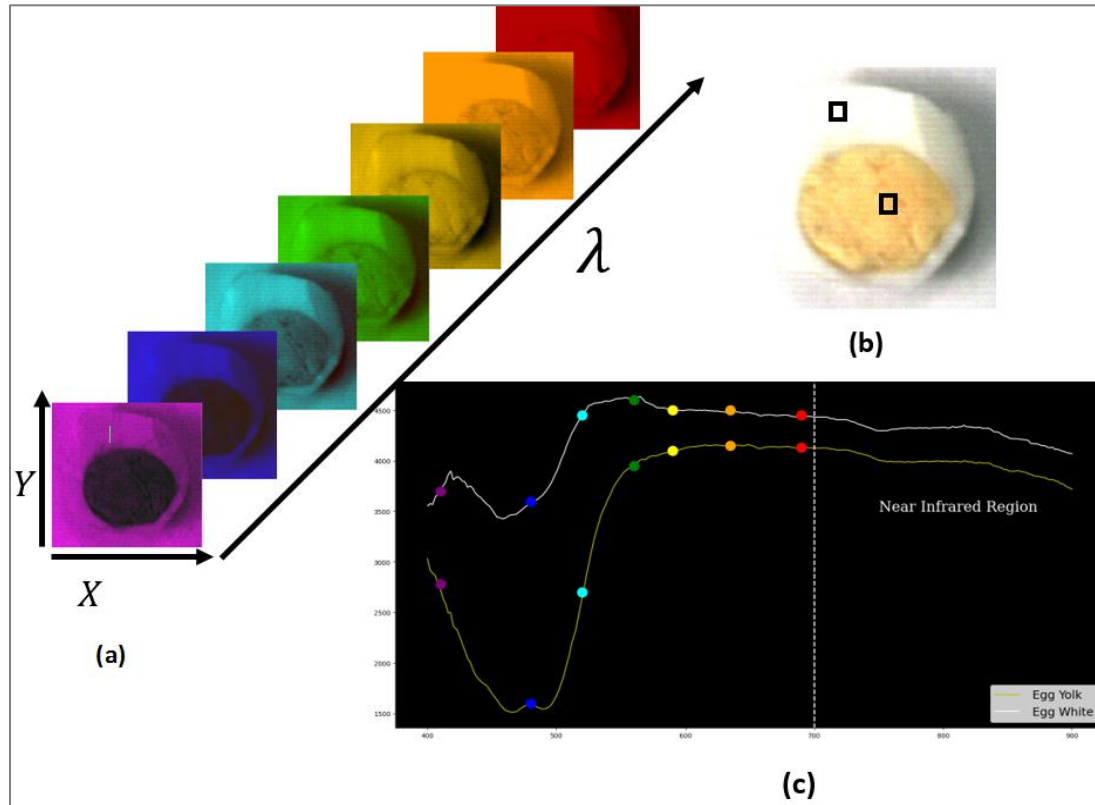

Figure 1 a) Hyperspectral data cube, b) boiled egg cut image, c) spectral reflectance curves of two pixels obtained from white and yolk of an egg

2.Literature review

The treatment of many diseases including heart problems and diabetes type 2 can be efficiently implemented if there is a monitoring system for individual's food consumption in place. Automatic processing is a preferred option whether it has to be done in a medical facility or at home. Collecting images of food is a universal method available for both professionals and individuals, therefore the popularity of using image processing/recognition method to infer about the amount and the nutritional contents of the consumed food. On the other hand, it is challenging to perform measurement accurately in real life experiments. There is ongoing research work on complex food categorization that includes food classification and measurements. Here, we focus on the analysis of the work done in classification of the food using a) optical images and; b) hyperspectral data.

2.1Overview on ingredient recognition task

Table 1 summarized the main characteristics of recent researches in ingredient recognition which use three-component color space such as RGB as well as have pretty similar ingredient dataset comparing to ours. Along with providing the recognition results in different evaluation scores, unequalized number of food categories and different type of ingredients, makes the comparison across the following studies very difficult. Furthermore, the published investigations for the recognition of similar food types using hyperspectral technology are summarized in *Table 2*. Section 3.2 contains the definition of all used evaluation metrics.

Table 1 Brief description of the related work using three component color images

Ref.	Dataset (Food types/# of images)	Method (Classifier/Features)	Evaluation metrics
[33]	150 food types/ 500 images	Multi-SVM (feature selector: Ada boost) /LAB color space texture	Accuracy: 90% (for 6 different food types in a plate)
[34]	209 food type/ 11,704 images	The Canberra distance to determine the similarity between feature vectors/ RGB color space, texture, edge orientation.	Mean average precision: nearly 100% in 5 major food classification (breads, cereal, fruits, vegetables, fast food)
[35]	12 food types/ 1,636 images	Multi-SVM; features in RGB color space, texture, geometry (seize, shape).	Accuracy: 85%
[36]	25 food types/ 545 images	Minimum distance classifiers/ SIFT, Gabor texture.	Accuracy: 60.7%
[37]	12 food types	Class rankers, Borda count for combining outputs from multiple feature matchers/RGB color, SURF.	Precision: 86% Recall: 83%
[38]	100 food types/15,000 images 30 fruit types/ 40,000 images	Convolutional Neural Networks (CNN) / Gray and RGB color.	Accuracy: 60.9% for 100 food categories; 80.8% for 30 food categories
[39]	30 food types/ 10,000 images	Deep Learning Neural Network	Accuracy: 99%
[40]	55 food types/ 12,262 images	Multi-kernel SVM/ SIFT, RGB color, texture, shape	Accuracy: 90%
[41]	257 ingredient labels in 172 Chinese food/ 110,241 images with the average of 3 ingredients per image	Multi-task DCNN, Fine-grained classifiers	Accuracy: 71% for 275 food categories
[42]	10 food types/5,822 images	CNN	Accuracy: 94%
[43]	73 food types/3,616 food instances	CNN; Gabor filters; HSV color	Accuracy: 79%
[44]	101 food types /90,840 images	CNN	Accuracy: 40.44%
[45]	164 food types/ 9,254 images	Contextual Relation networks (CR-Nets)	Precision: 87% Recall: 89.12%
[46]	80 types/ 12,083 images	Spatial Regularization Network (SRN)	Macro Precision: 73.7% Micro Precision: 77.3 % Macro Recall: 70.7 % Micro Recall: 74.7 % Macro-F1: 72% Micro- F1: 76.0%
[47]	353 ingredient labels in 172 types/ 110,241 images	Attention Fusion Network (AFN) and & the double-flow feature extraction network	Macro-F1: 58.80% Micro-F1: 74.10% Accuracy: 34.29%
[48]	CineseFoodNet: 172 ingredient labels; VIREO Food 172: 353 ingredient labels/ 185,628 images in total	Multi-Task Deep Learning Networks	CineseFoodNet: top-1 accuracy: 65.58±0.23 CineseFoodNet: top-5 accuracy: 90.41±0.11 VIREO FOOD-172: top-1 accuracy: 91.08 ± 0.21. VIREO FOOD-172: top-5 accuracy: 98.86 ±0.14
[49]	10 types/1,250 images	Deep Learning (FCN, Segnet, Enet, DeepLapV3+Mask RCNN)	Best result by FCN: Precision: 79% Recall: 81% Specificity: 99%
[50]	406 ingredient labels in 251 food types/169,673 images	Deep CNN (DCNN)	Macro-F1: 61.74 % Micro F1: 75.77 %

Table 2 Brief description of the related work using Hyperspectral images

Ref.	Dataset (Food types/# of samples)	Spectral Range (nm)	Method (Classifier/Feature Extractor)	Evaluation Metrics
[51]	3 different types of pork / 75 samples	900-1700	creation of score images /PCA	Accuracy: 96%
[52]	2 different types of cooked beef / 475 samples	922-1739	Linear Discriminant Analysis (LDA) /PCA	Accuracy: 75.4%
[53]	4 different types of rice / 360 samples	400 - 1000	SVM	Accuracy: 90.2%
[53]	5 different types of Chinese tea leaves / 20 samples	400 - 800	Artificial Neural Networks (ANN) / PCA	Accuracy: 91.44%
[54]	3 different types of rice / 1,152 samples	400 - 1000	Back Propagation Neural Networks (BPNN)/ PCA	Accuracy: 94.45%
[55]	13 muscles types of standard loin and leg chops/140 samples	672 - 958	3D CNN	Accuracy: 95.81%
[56]	4 different types of rice / 5,226 samples	380 - 1030 874 - 1734	CNN	(380 - 1030) Accuracy: 80.9% (874 - 1734) Accuracy: 87%
[57]	3 different types of soybean/ 165 samples	400 - 1000	Extreme Learning Machine (ELM)/ information measure of the FRS theory/IM-FRS	Accuracy: 98.42%
[58]	13 different classes of fruits-vegetables/ 2,700 samples	470 - 430	ImageNet pre-trained CNN	Accuracy: 92.23%
[59]	17 different types of maize seed/ 1,632 samples	400 - 1000	Least Square Support Vector Machine (LS-SVM)/ Multi-Linear Discriminant Analysis (MLDA)	Accuracy: 99.13%
[60]	3 different types of Bacon/ 5,000 samples	400 - 1000	SVM/ CNN	Accuracy: 99.2%
[61]	10 different types of rice/ 4,320 samples	400 - 1000	Principal Component Analysis Network (PCANet)	Accuracy: 98.57%
[62]	90 different types of rice / 8,640 samples	1000 - 1700	Random Forest (RF)/ LDA	F1- Score: 78.27%
[63]	3 different types of maize seed/ 4,800 samples	874 - 1734	Long Short-Term Memory (LSTM)	Accuracy: 90.11%
[64]	11 different types of lettuce/ 1,614 samples	350 - 1000	LDA / STEPWISE	Accuracy: 81.04 %

2.2The impact of hyperspectral imagery on food recognition

Hyperspectral images have the ability to distinguish the same ingredient with different type of cooking, texture, shape and size even color as well different food type with very similar color. Moreover, the number of images that need to train a classification model to achieve high recognition rate is significantly lower compared to other recognition system with similar results. Therefore, using hyperspectral capabilities, makes the data-acquisition and data preparation steps extremely beneficial in terms of time and cost.

3.Methods
3.1Data dimension reduction

Food classification is a complex process which requires consideration of many factors such as selecting/extracting an optimal subset of features, producing an efficient learning model, and accuracy assessment. All these factors directly affect the quality of the food classification result. This section

focuses on the data preprocessing step, which involves a brief explanation of ReliefF and PCA.

The capability and performance of a hyperspectral imaging system is based on the spectral resolution of the camera. Since we use a line scan spectral imager, the spectral resolution is the spectral width that can be achieved in one horizontal scan. Comparing to multispectral images, hyperspectral images can have a nearly contiguous bands with much narrower spectral width. Although hyperspectral images bring new capabilities, they suffer issues related to "curse of dimensionality". Therefore, some preprocessing steps such as optimal feature extraction or selection must be considered. The feature extractor transforms the original features into a reduced feature space by maximizing the variance of projected data then choosing the features that are significant. The feature selection methods determine a subset of non-redundant relevant features while preserving crucial information.

In this paper, discriminating capabilities of both PCA as feature extractor, and ReliefF as feature selector are exploited which, in turn is expected to improve the classification performance. Experiments have been conducted with four datasets: original hyperspectral, RGB dataset, selected PCA bands, and ReliefF data. The results are evaluated and compared.

3.1.1 Principal component analysis

In the spectral dimension of hyperspectral data, neighboring spectral bands are highly correlated. The main purpose of applying PCA to hyperspectral data is to capitalize on this correlation and obtain the optimum linear combination of the original bands with a maximum variation of the projected data. A brief discussion of using PCA for hyperspectral data can be found in [26].

Consider a hyperspectral data cube X with P and K as spatial dimensions and N is the number of spectral bands (*Figure 2*).

X consists of M pixel vectors, X_i, which are N dimensional vectors (M refers to the number of pixels in an image and equals to P × K. In other words, all elements of X_i, have the same spatial locations but in different spectral bands. $X_i = [x_1, x_2, \ldots, x_N]_i^T$.
The covariance matrix is computed by equation (1):

$$Cov(X) = \frac{1}{M}\sum_{i=1}^{M}(X_i - \bar{X})(X_i - \bar{X})^T \qquad (1)$$

were, $\bar{X}$ is the mean vector of all pixel vectors, calculated as $\bar{X} = \frac{1}{M}\sum_{i=1}^{M}X_i$. After the decomposition based on eigen-values of the covariance matrix, which takes the form of:

$$Cov(X) = ADA^T \qquad (2)$$

Where D = diag $(\lambda_1, \lambda_2, \ldots, \lambda_N)$ and A are the diagonal and the orthonormal matrices composed the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ of the Cov(X), and the corresponding N dimension eigenvectors, a_k ($k = 1, 2, \ldots, N$), respectively. The linear transformation is defined by: $Y_i = A^T X_i (i = 1, 2, \ldots, M)$, where Y_i are the PCA- transformed bands of original pixel vectors (X_i). Let the eigenvalues and their corresponding eigenvectors be sorted in decreasing order, then $\lambda_1 \geq\geq \lambda_2 \geq, \ldots, \geq \lambda_N$. The first K (usually K << N) can be eased to get the original data approximately as follows:

$$Z_i = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_K \end{bmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1K} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2K} & \cdots & a_{2N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{K1} & a_{K2} & \cdots & a_{KK} & \cdots & a_{KN} \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix} (1 \leq i \leq M)$$
$$(3)$$

The property of Z_i, the transformed PCA bands, is that the highest variance is in the first few bands, meaning that they contain the major information of the original hyperspectral data. This property leads to reducing the original dataset to the most informative and noise free.
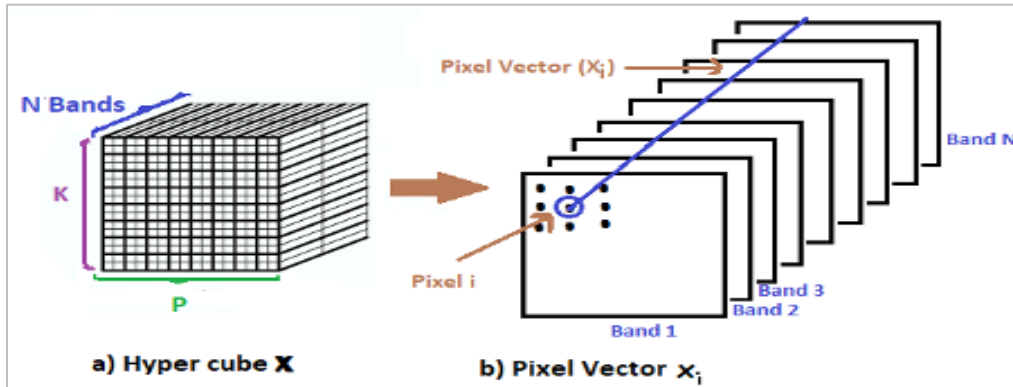

a) Hyper cube **X**   b) Pixel Vector x_i

Figure 2 a) Hypercube X, b) Pixel Vector X_i

3.1.2 ReliefF

Feature selection is a method of selecting of a small subset of features that are necessary and sufficient to describe the object. There are usually realized in two steps a) generating of feature subsets, and b) evaluating them. ReliefF, a filter ranker, is a nearest neighbor feature selection that defines ranks of features based on their relevance to specific task performance such as classification accuracy.

ReliefF [65] is an extension of an algorithm called Relief [66] that was an efficient method to estimate the quality and relevance of all attributes. Comparing to the original Relief, ReliefF has got the ability to deal with noisy and incomplete data, can be used in multi-classification problems, and can estimate the conditional probabilities more reliably without increasing the complexity. The conditional probability that two values should be the same or different, approximated with relative frequencies from the dataset and it can be calculated if one or both features are missing.

The key idea of Relief as a robust filter ranker method was to estimate attributes based on how well their values distinguish among the samples that were near to each other. To achieve this Relief searches for its two nearest neighbors: nearest hit (one from the same class) and nearest miss (the other from a different class). Relief also randomly selects n training instances, where n is defined by the user.

Compared with many heuristic measures, the main advantage of Relief algorithm is that it does not make assumption of conditional independency of the attributes (upon the target variable) for estimating the quality of them that results in the efficiency and awareness of the contextual information. This leads to its ability to correctly estimate the quality of attributes with strong dependencies among them. However, their incapability to find the optimal number of selected features is one of their main disadvantages. As a simple solution it can be done manually by increasing the number of selected features from the top of the rank list, then classifying data based on the current feature set. Then, one with the highest classification accuracy is optimal. The whole algorithm and the extended version of Relief can be explained with the following pseudo-code:

Algorithm 1: Basic Relief
      (only for binary classification)

Result: weight vector for attributers ($\mathbf{W}$)

Initialize all elements in $\mathbf{W}$ to zero.

for *every training instance* R_i ($1 \leq i \leq m$) **do**

29

    Find nearest hits H and miss M;

    for *every attribute (feature)* A ($1 \leq A \leq F$) **do**

$$W[A] = W[A] - \frac{diff(A,R_i,H)}{m} + \frac{diff(A,R_i,M)}{m}$$

    end

end

---

where $diff\left(A, I_i, I_j\right) = \dfrac{|values(A,I_i) - values(A,I_j)|}{\max(A) - \min(A)}$, I_i and I_j are different training samples.

Updating attribute weights is based on a simple idea that if R_i and H that are instances of the same class have a large different value means that attribute separates two instances with the same class is not desirable. Then, its weight should be reduced in contrast if the R_i and M (instances of different classes) have a large different value, the attribute which separates them is desirable and hence its weight should be increased.

Algorithm 2: Relief- Extension (ReliefF)

Result: weight vector for attributers ($\mathbf{W}$)

Initialize all elements in $\mathbf{W}$ to zero.

for *every training instance* R_i ($1 \leq i \leq m$) **do**

    Select K nearest hits H_j.

    for *every class C, $C \neq$ Class (R_i)* **do**

        find K nearest misses $M_j(C)$:

        for *every attribute (feature)* A($1 \leq A \leq F$) **do**

$$W[A] = W[A] - \sum_{j=1}^{K} \frac{diff(A,R_i,H_j)}{m \times K} + \sum_{C \neq Class(R_i)} \frac{diff\left(\left|\frac{P(C)}{(1-P(Class(R_i)))} \sum_{j=1}^{K} diff(A,R_i,M_j(C))\right|\right)}{m \times K};$$

        end

    end

end

---

where P(C) is the prior probability of class C.

At first, ReliefF randomly selects a training instance R_i similar to basic Relief algorithm, but the way of searching for K nearest hits (the same class) and K nearest misses (for each of the other class) is different. Finally, after updating the weight of each attribute, the average of contribution of all hits and all misses is computed.

The basic difference of two above algorithms is in selecting K hits and K misses in the ReliefF, which brings greater robustness to the system when noise is concerned.

Previous studies have indicated that the correctness of ReliefF rank estimation increases when K, the number of nearest neighbors increases. If the features

are conditionally dependent, then the increasing rate is not monotonic, but it reaches a peak and then drop off as K increases. It is essential to know that every specific dataset has a different optimal value for K.

3.2 Classification models

3.2.1 Support vector machine multi-classification

Support Vector Machines (SVM), a supervised machine learning model was originally designed to do binary classification by con-structing a hyperplane in an n-dimensional space that separates the data points to their potential classes having the maximum distance to the hyperplane. Among all data points, those with the minimum distance to the hyperplane are called Support Vectors, which have higher influence on the exact position of the hyperplane than of other points. In this paper, SVM implementation from scikit-learn library was used. Selecting the most appropriate kernel function is essential due to its significant effect on the performance of SVM. Linear kernel and Radial basis function (RBF) are widely used. The RBF defined by the following formula and has the ability to deal with nonlinear data efficiently by mapping data points to a higher dimension spaces that enables them to be linearly separated.

$$K(x, \acute{x}) = exp(-\gamma \|x - \acute{x}\|^2) \qquad (4)$$

3.2.2 Logistic regression

Logistic regression, one of the most popular machine learning models is a special regression problem with discrete output which models the conditional probability $p_w(y_i|X_i)$ in binary classification. It is formulated as follows [67]:

$$log\left(\frac{p_w(y_i|X_i)}{1-p_w(y_i|X_i)}\right) \qquad (5)$$

Then

$$P(y = y_i|X) = \frac{1}{1+exp(-y_iX_iW)} \qquad (6)$$

The maximum likelihood estimation of W is obtaining by:

$$W^* = arg\,max\sum_{i=1}^{N}\{-\log(1 + \exp(-y_iX_iW))\} \qquad (7)$$

Adding the regularization parameter λ, usually improves the numerical stability and the solution also will be biased by improving the generalization performance of the model.

$$W^* = arg\,min_w\,f(W) + \lambda h(W) \qquad (8)$$

In the other words, regularization can be seen as a penalty against complexity. Increasing its strength, large weight coefficients will be penalized which means regularization prevents overfitting by reducing the flexibility of the model. There are two popular types of regularization that prevent overfitting by

shrinking the coefficients(W) a) Lasso Regularization (L1), and b) Ridge Regularization (L2). L1 shrinks some coefficients to zero, while L2 shrinks all by the same proportions without eliminating any of them.

3.3 Evaluation metrics for binary classification

Binary classification metric is a number that measures the performance of a binary classifier such as Logistic Regression and typically is presented with a value between 0 and 1. Score of 1 usually reserved for the perfect model. There are different metrics, that selection of each is based on the problem solved. In this work, four performance metrics are used:

3.3.1 Accuracy

It measures how many samples (both positive and negative) are classified correctly.

$$\frac{number\ of\ samples\ classified\ correctly}{total\ number\ of\ samples} \qquad (9)$$

3.3.2 positive predictive value (precision)

It measures how many samples predicted as positive are really positive. Taking bean classification problem, it means what is the ratio of samples correctly classified as a bean.

$$\frac{True\ Positive}{True\ Positive + False\ Positive} \qquad (10)$$

3.3.3 True positive rate (Recall or sensitivity)

It measures how many samples out of all positive samples are classified as positive. For example, how many beans samples are recalled from all bean's samples.

$$\frac{True\ Positive}{True\ Positive + False\ Negative} \qquad (11)$$

3.3.4 True negative rate (Specificity)

It measures how many samples out of all negative samples are classified as negative. For example, how many non-beans samples are recalled from all non-bean's samples.

$$\frac{True\ Negative}{True\ Negative + False\ Positive} \qquad (12)$$

3.3.5 F1 Score

F1 is the weighted average of precision and recall. In the case of imbalanced dataset, the F1 score is usually more useful than accuracy.

$$\frac{2 \times Recall \times Precision}{Recall + Precision} \qquad (13)$$

3.3.6 Micro and macro averaging scores

All above scores are usually used in binary classification problems. They can be computed for multi-classification problems too: First, they are calculated on each individual class and then be averaged over all classes (macro-averaging). There is also another way to compute them: in the case of having *m* categories and *n* test instances, they are

computed globally over all the m × n binary decisions (micro-averaging). The micro-averaged scores tend to be influenced by the classifier's performance on common categories, while the macro-averaged scores are more dominated by the performance on rare categories. By using both kinds of score, the more informative comparison can be provided [68].

3.3.7AUC- ROC curve
Area under the Curve (AUC) represents degree or measure of separability by using area under Receiver Operating Characteristics (ROC), which is a probability curve. It indicates if the binary classifier is capable of distinguishing between two classes. Higher the AUC, better the classifier at distinguishing between beans and no beans sample. The ROC curve is plotted with True Positive Rate on the y-axis and against the False Positive Rate on the x-axis.

3.4Image acquisition and pre-processing
A Resonon Pika II hyperspectral camera is used to acquire both RGB image and the hyperspectral response data in the spectral range of 400-900 nm, with a band width of 2.1nm covering the visible/near infrared regions (*Figure 3*).

Resonon Pika II is a line-scan imager that collects data, one vertical line at a time. To create a two-dimensional image, multiple lines are assembled from lines to form a complete image, as illustrated in *Figure 4*. Collection from vertical lines in 2D that forms an image in a single spatial band and each pixel in different spectral bands is a spectral signature.

Figure 5 shows the spectral signatures of various food samples. The collection of all 1D signatures from the spatial XY planes makes a 3D hyperspectral cube. Pika II hyperspectral imager used in our work produces, $200 \times 640 \times 3$ color image and hence $200 \times 640 \times 240$ point spectral cube.
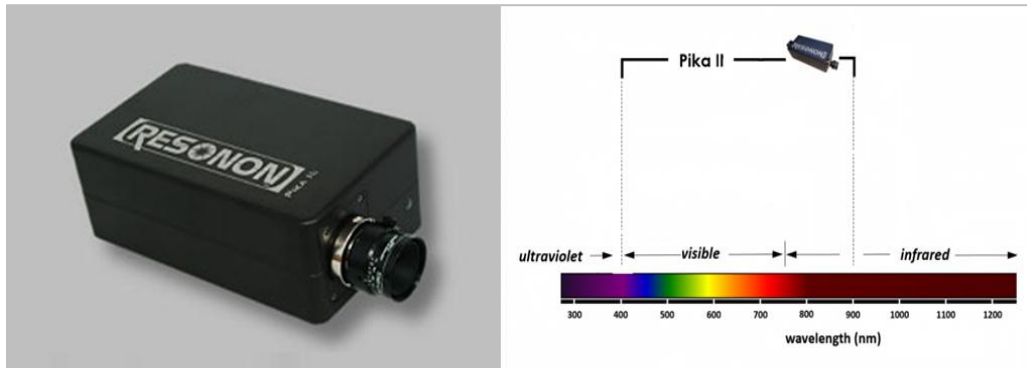

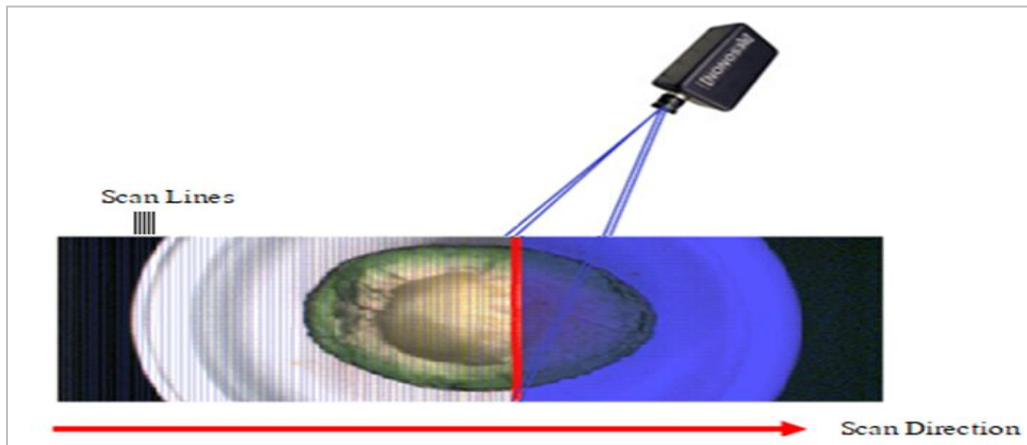**Figure 3** Left: Pika II hyperspectral imaging camera, right: covered spectral range by it

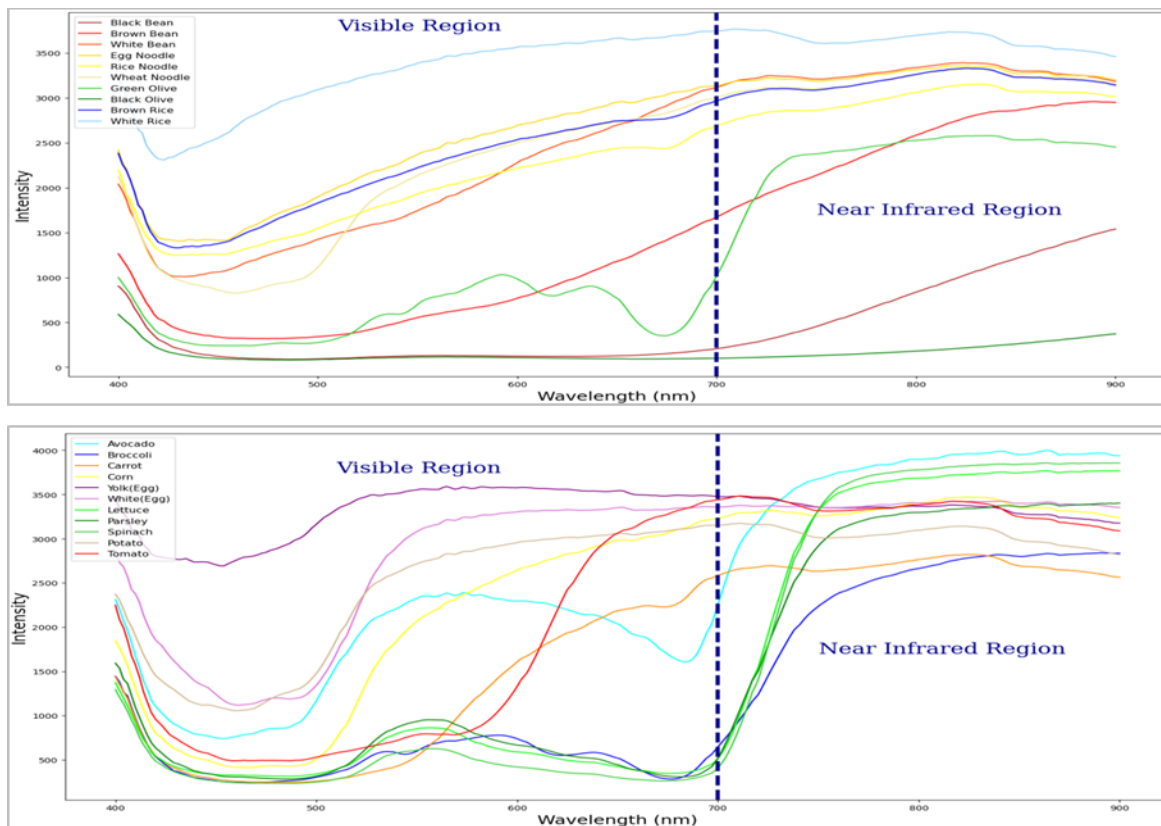**Figure 4** Scanning line by line from left to right

Figure 5 Reflectance spectrum of all different food used in this work, top: 3 type of bean, 3 type of noodle, 2 type of olive and 2 type of rice, bottom: avocado, broccoli, carrot, corn, egg (yolk and white), lettuce, parsley, potato, spinach and tomato

3.4.1 Hyperspectral camera calibration

Prior to analyzing data, the data needs to be calibrated. The dark current characteristics of the hyperspectral camera array detectors known to produce a non-uniform response. The average dark current noise is measured by blocking the light entering the camera, collecting the response, and averaging. Also, to remove the reference reflectance spectrum, a white target was placed in front of camera. The calibration is carried by calculating P_{ij} by equation (14). It normalizes the radiance spectrum and results in the reflectance that excludes spectral non-uniformity of the illumination device and subtracts the dark current.

$$\forall P_{ij} \in I(x, y, \Lambda): \acute{P}_{ij} = \frac{P_{ij} - D_{ij}}{W_{ij} - D_{ij}} \qquad (14)$$

where P_{ij} and $\acute{P}_{ij}$ are pixels at position (i,j) at all wavelengths λ in the original image (I) and normalized image respectively, D_{ij} is the pixel value in the dark reference image, and W_{ij} is the pixel in the white reference image.

3.4.2 Image segmentation

The purpose of image segmentation is to extract an object from the background in 2D and then extract the respective spectral information. Segmentation is performed on the RGB image that produces a boundary around the objects. Based on the spectral values inside the boundary, hyperspectral cube data is extracted for processing. *Figure 6* explains the data extraction.
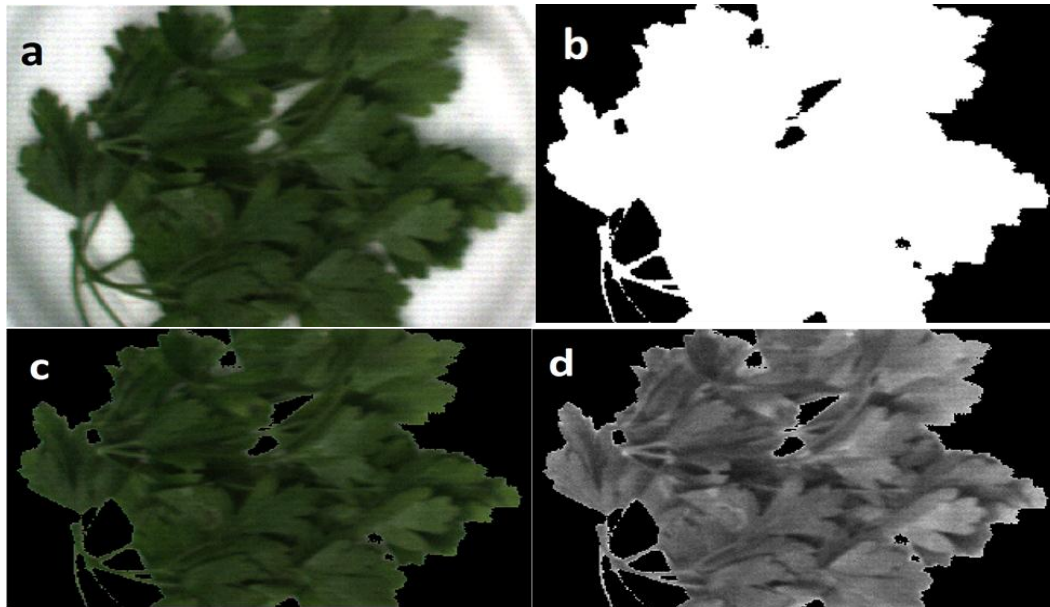
Figure 6 a: Original RGB image, b: Binary image, c: RGB image after applying binary mask, d: Hyperspectral image after applying binary mask (band = 100)

3.4.3 Convert 3D data cube to 2D matrix
Only spectral features are used for classification. For computations, the 3D hyperspectral data cubes are represented by 2D matrices with height × width corresponding to the number of rows, and λ, the number of spectral bands columns in the new matrix. After constructing new matrix, rows with all zero elements are removed.

3.4.4 Normalization
If the range of data values varies widely, it may cause the independent variable with larger scale being arbitrarily weighted more heavily by the learning model. To avoid this issue, data and features are normalized. Here, a standard scalar normalization is used to rescale the distribution of values with zero mean and standard deviation equal to one, respectively.

3.5 Data description
3.5.1 Food samples
RGB and hyperspectral images have been acquired from twenty different fruits, vegetables, and seeds. *Table 3* shows a complete list of experimental samples used in this study. RGB as well as hyperspectral images/data are obtained for each food type to create training dataset. For the test dataset, RGB image and hyperspectral image of completely different sample but of the same type of food was captured.

Table 3 Food names, type of food processing and parts used in experiments

No.	Name	Type (Processing)	Parts
1	avocado	raw	cut, no pit
2	black bean	simmered	whole grain
3	black olive	raw	whole, no pit
4	broccoli	steamed	head & stalk, sep.
5	brown bean	simmered	whole grain
6	brown rice	simmered	whole grain
7	carrot	steamed	whole
8	corn	canned	whole
9	egg	simmered	white & yolk, sep.
10	egg noodle	boiled	bowl
11	green olive	raw	whole w/ pit
12	lettuce	raw	leaves
13	parsley	raw	leave w/ short stem whole or peeled
14	potato	simmered	bowl
15	rice noodle	boiled	leaves

No.	Name	Type (Processing)	Parts
16	spinach	raw	cut
17	tomato	raw	bowl
18	wheat noodle	boiled	whole grain
19	white bean	simmered	whole grain
20	white rice	simmered	whole grain

3.5.2 Data acquisition

We have setup an imaging system for the food evaluation with the hyperspectral Pika II. The system consists of three symmetric light sources, objective lenses, and an imaging camera mounted on a stand (see, *Figure 7*). Food samples are placed on a linear translation stage constructed to move under the camera that allows the capture of hyperspectral data cube of the food sample.

The most common light source used in such systems are halogen lamps. Use of LED as lighting source was also studied. LED emits light in a narrow range of wavelength. However, based on the study in [69], beyond 750 nm (NIR), the LED spectrum is noisy as demonstrated in *Figure 8*. Therefore, we have used incandescent light which brings in the benefit of capturing near infrared information into the current experiment.
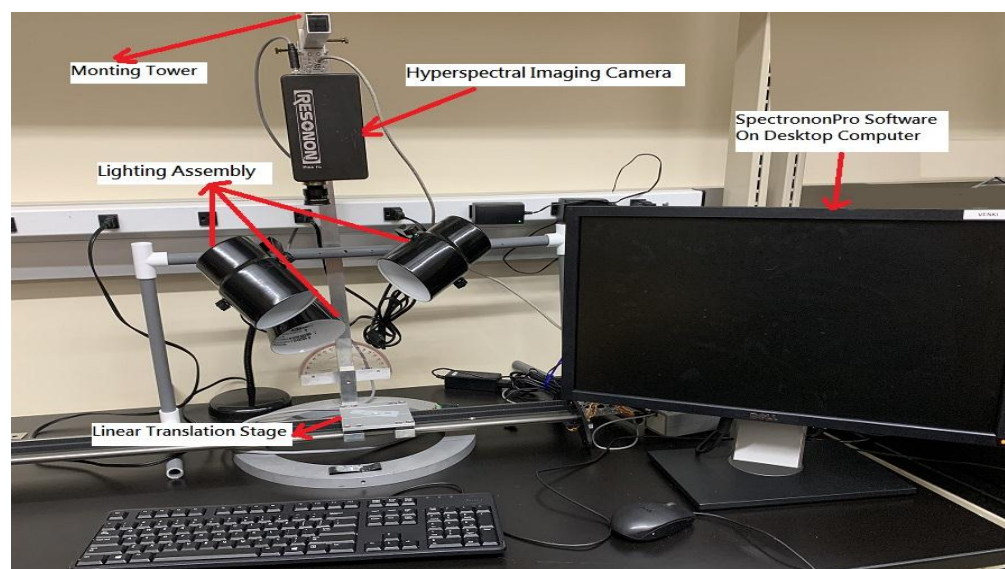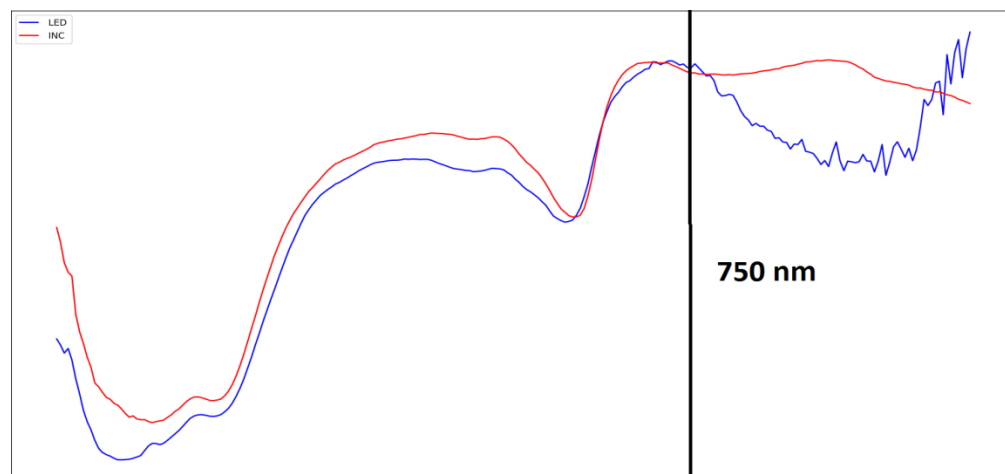

Figure 7 Hyperspectral imaging system

Figure 8 Hyperspectral signature of an apple under two different illumination conditions, blue: LED, red: incandescent

3.5.3 Training sample preparation

Training data samples are collected with twenty individual food items shown in *Figure 9*. After completing the scan of food samples, the region of interests (ROIs) is defined on the reference image for computing the average color and the spectral response. ROI of 8×8-pixel block is considered in this study. There are total of 20 matrices of 1000x 240 and 1000 x 3 of hyperspectral and RGB training dataset, respectively. That means every food item has 1000data points in each of hyperspectral and RGB training data sets.

3.5.4 Test sample preparation

A meal plate was assembled by combining different food samples to collect test dataset (*Figure 10*). N=1000, testing samples for each food item was extracted to create the test dataset.

3.6 Data visualization

Data visualization is an important step in analyzing the data. It provides important insights about the subject under study and plays an essential role in model building, evaluation, and testing. t-Distributed Stochastic Neighbour Embedding (t-SNE) is a probabilistic technique suitable for visualization of high-dimensional datasets. It creates a reduced feature space where similar samples are represented by nearby points while dissimilar ones are modelled by distant points with high probability. The t-SNE defines a similar distribution for the points in the low-dimensional embedding and minimizes the Kullback–Leibler divergence between the two distributions with respect to the location of the points. *Figure 11* shows t-SNE plots for both hyperspectral and RGB data sets. It can be seen that classes that are not separated on RGB images are separated in hyperspectral data. This is particularly seen in foods with similar colors such as black beans and black olives (black arrows → black circles) as well as for lettuce and spinach (green arrows → green circles).
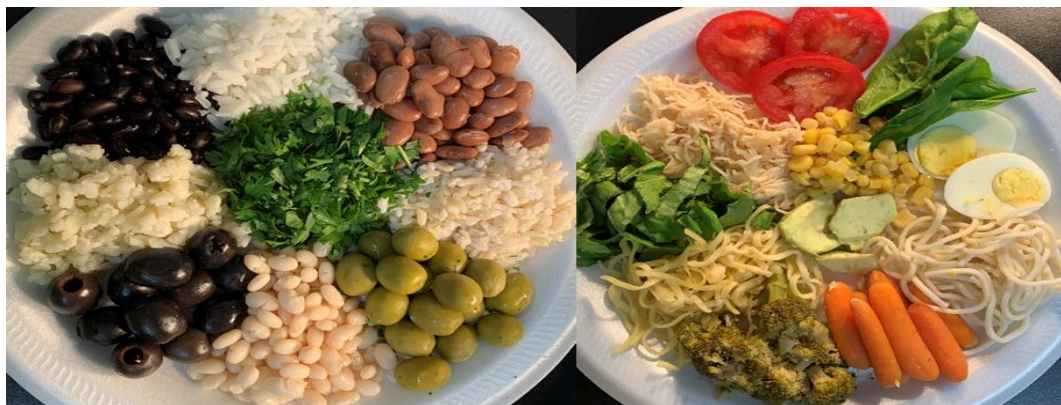

Figure 9 Training samples images

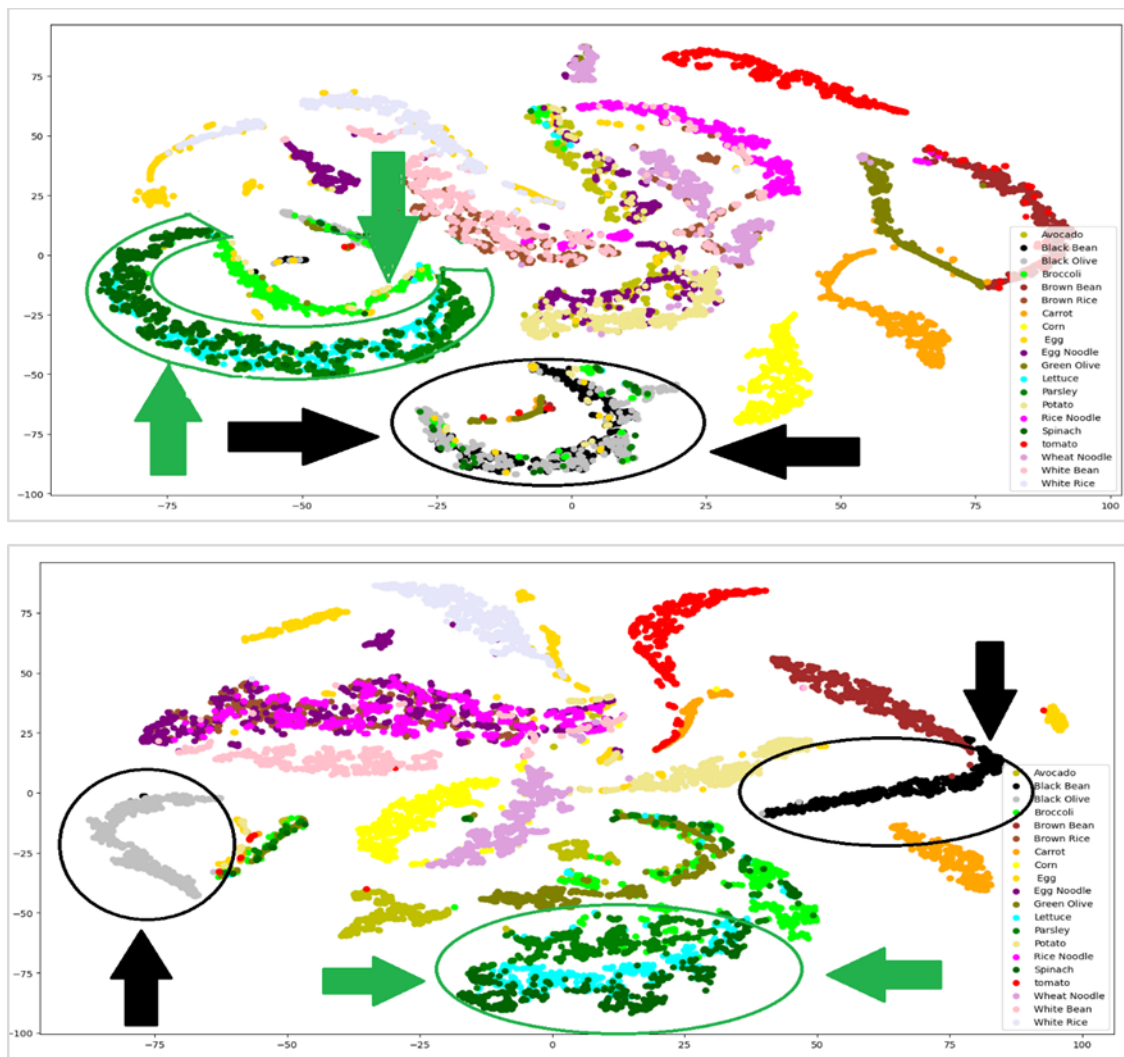**Figure 10** Test samples images

Figure 11 t-SNE plot, top: RGB, bottom: Hyperspectral

4.Results

In this section, we summarize the performance of Logistic Regression and SVM as binary and multi-classifiers on RGB and hyperspectral datasets of 20 different foods. *Figure 12* shows the workflow diagram of the proposed method. The first stage in the diagram, (# 1), is related to hyperspectral and RGB data collection, and preparation steps. The pre-processing block (# 2) includes ReliefF and PCA feature selection, and extraction. To evaluate the effectiveness of the hyperspectral data for food recognition in dietary assessment systems, experiments have been done in two separate phases: a) multi-classification, and b) binary classification (# 3). By using SVM in the first stage, we differentiate the test instance. In the next step, using linear regression, we obtain a hierarchical classifier wherein

each end node corresponds to a food group as shown in *Figure 12* (# 3).

Results of ReliefF for rank estimation, ReliefF is invoked with different K's to find the peak monotonic increase before it starts dropping off. When applied to hyperspectral datasets it yields results demonstrated in *Figure 13*.

Figure 14, shows the accuracy of SVM multi-classification with respect to the size of the feature subset for different values of K. The formation process of feature subsets starts with ten top features, followed by additions of the next ten into the previous subset until the entire set of features is included. The SVM classification scores for training and testing datasets are plot separately.

The average relative weight of each feature, i.e., spectral band for values of K between 300 and 400 is shown in *Figure 15*. Spectral bands with the higher weights, that is high rank features, having relative weights higher than 0.8 are colored in dark green. As the feature's relative weight decreases, the green color, gradually be-come lighter and turns to yellow, then orange and finally bright and dark red, where the

less important features with relative weight is less than 0.2.

Figure 16 shows the mean and the standard deviation of ReliefF testing score of different feature subsets for the same K values as in *Figure 15*.
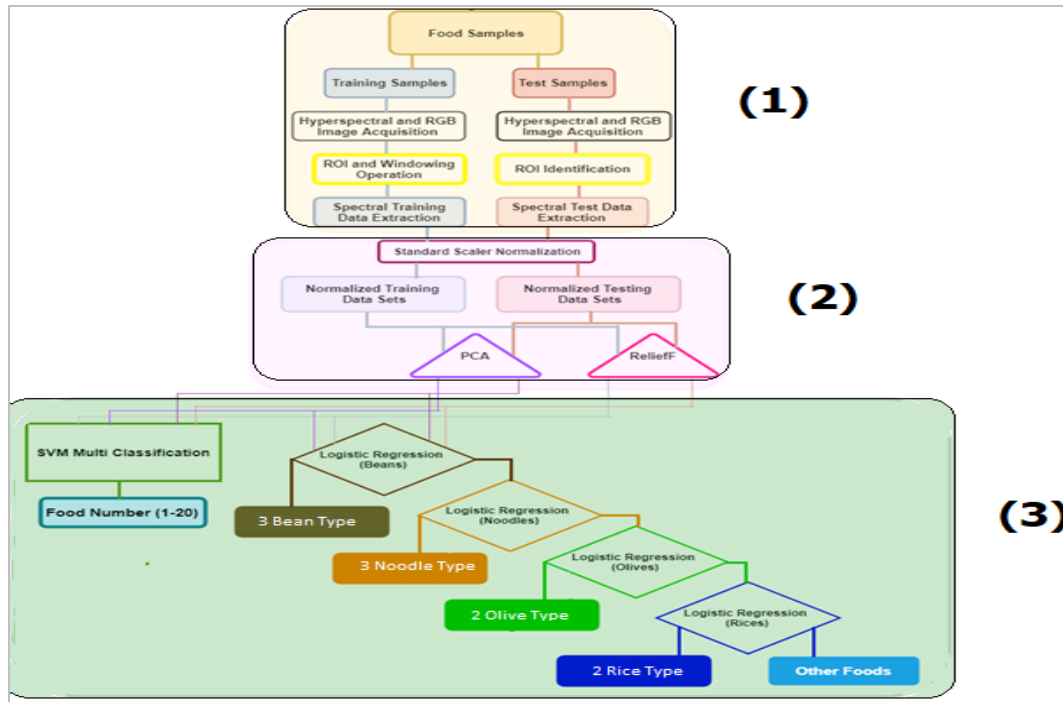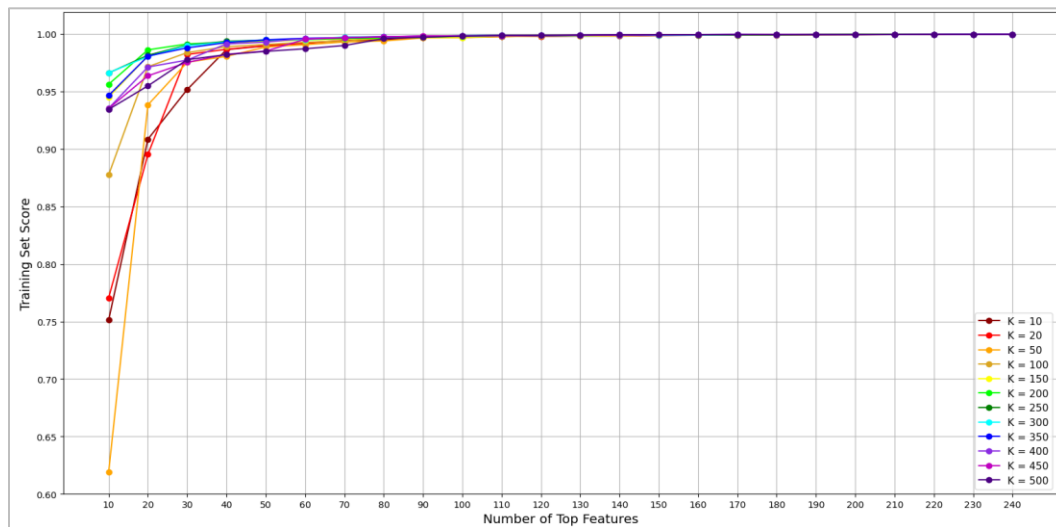

Figure 12 Food Classification Flow, 1) Data Collection, 2) Feature Extraction/Selection, 3) Classification
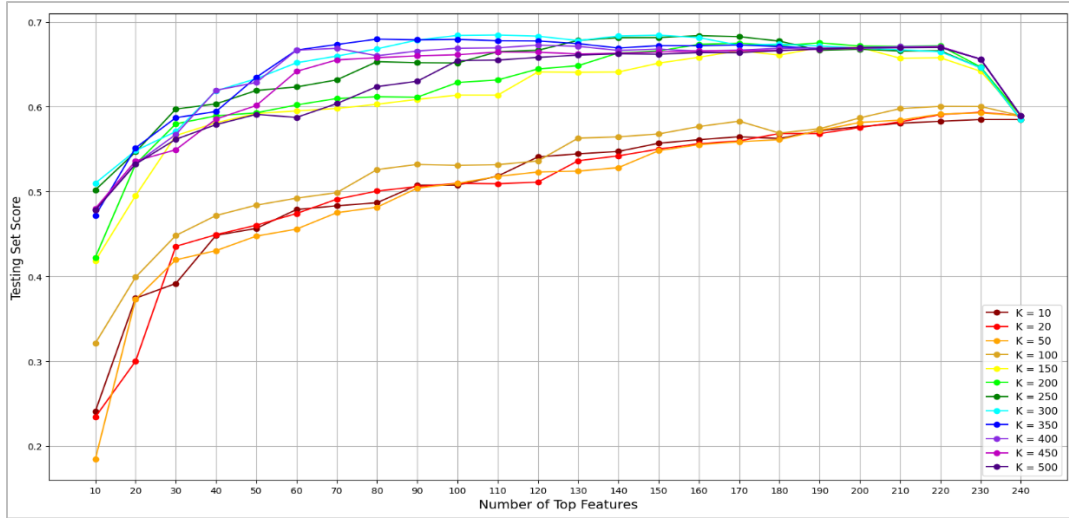
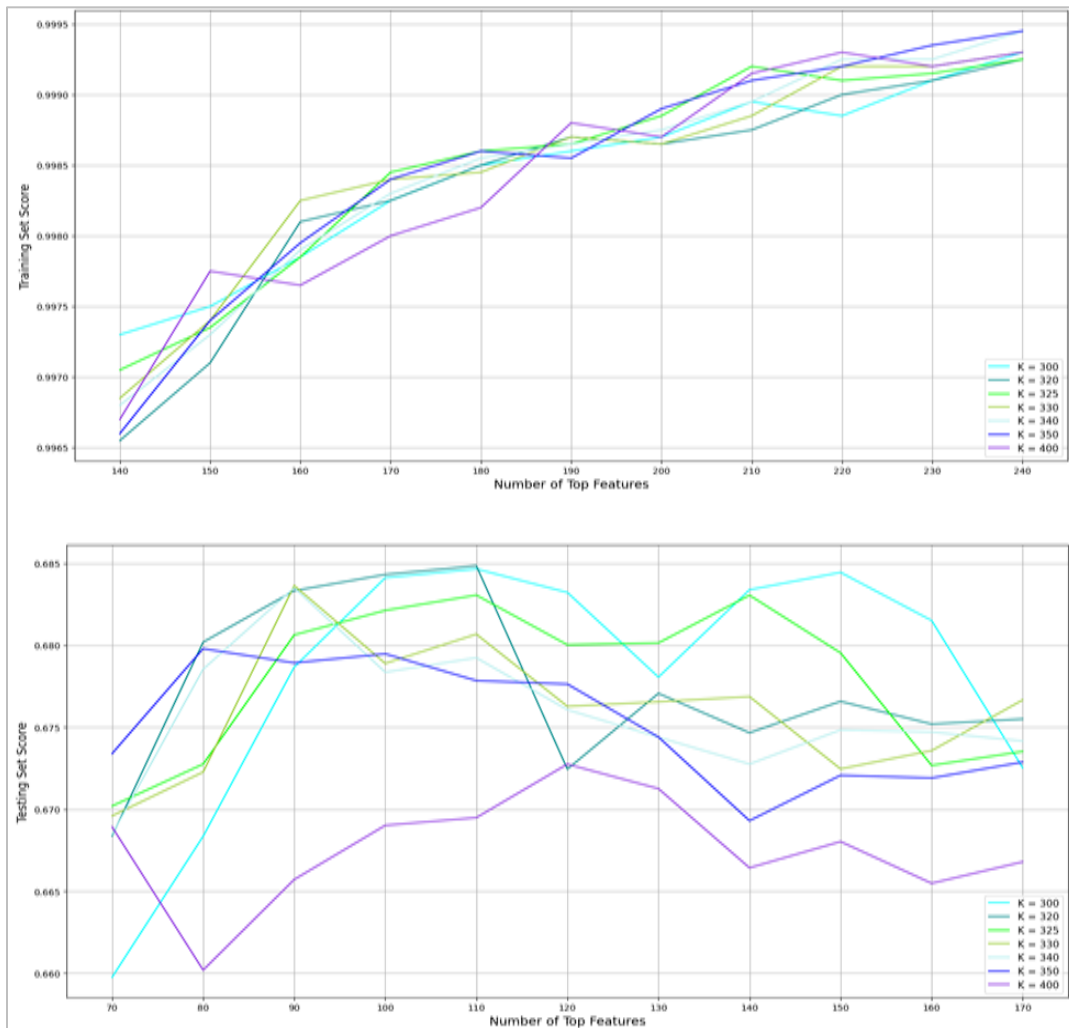**Figure 13** Classification accuracy for different value of K, top: training set, bottom: test set

Figure 14 Classification accuracy for different value of K, top: training set, bottom: test set
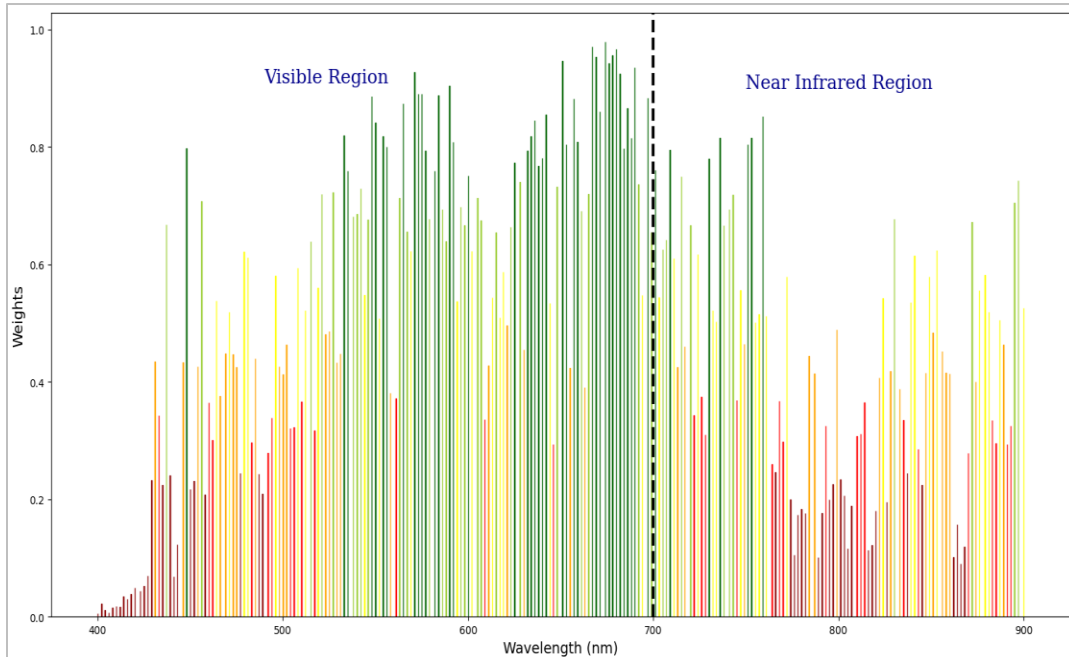
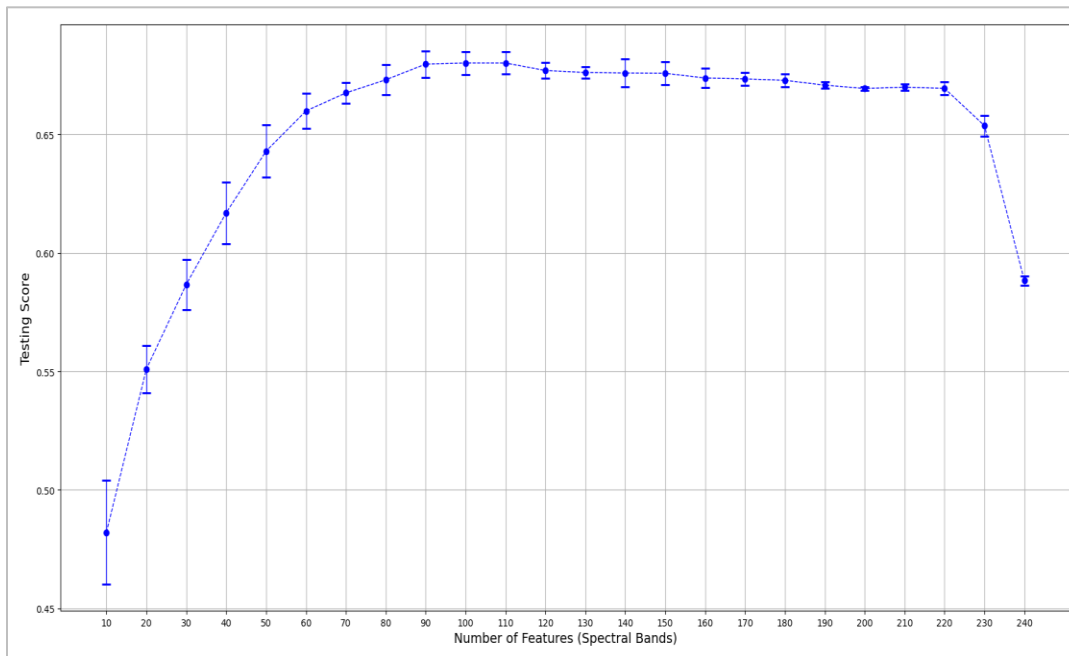**Figure 15** Average relative weights of spectral features for different value of K (between 300 to 400)

Figure 16 Mean and standard deviation of testing score for different ReliefF feature subsets

4.1Principal component analysis

In *Figure 17,* we demonstrate that the first three components account for over 95% of the variance which means that by using the first few components, most of the essential characteristics of the data can be recovered.

Table 4 and *Figure 18* show the performance of SVM multi-classifications with radial-basis function and hyperparameters, $\gamma = 0.001$ and C=1000 on the PCA transformed data. The percentage of pixels that were correctly classified relative to the total number of pixels in training dataset, testing dataset, and total running time in seconds is presented in *Table 4*. From

39

the results, with the first 9 components the highest classification score for the testing dataset is 0.61248 with 0.97975 being the score for the training set.
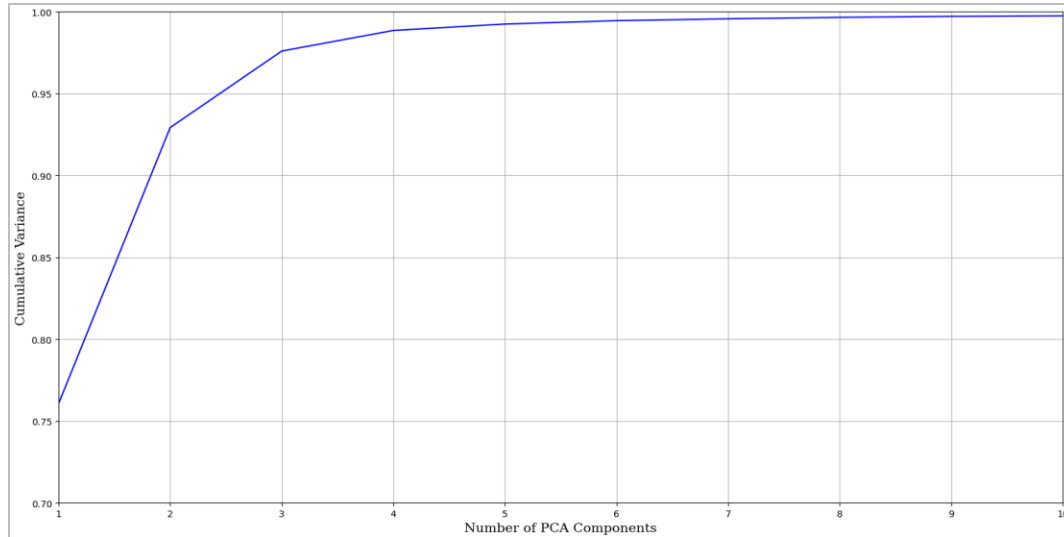

Figure 17 Cumulative variance of the first 10 PCA components

Table 4 Food names, type of food processing and parts used in experiments

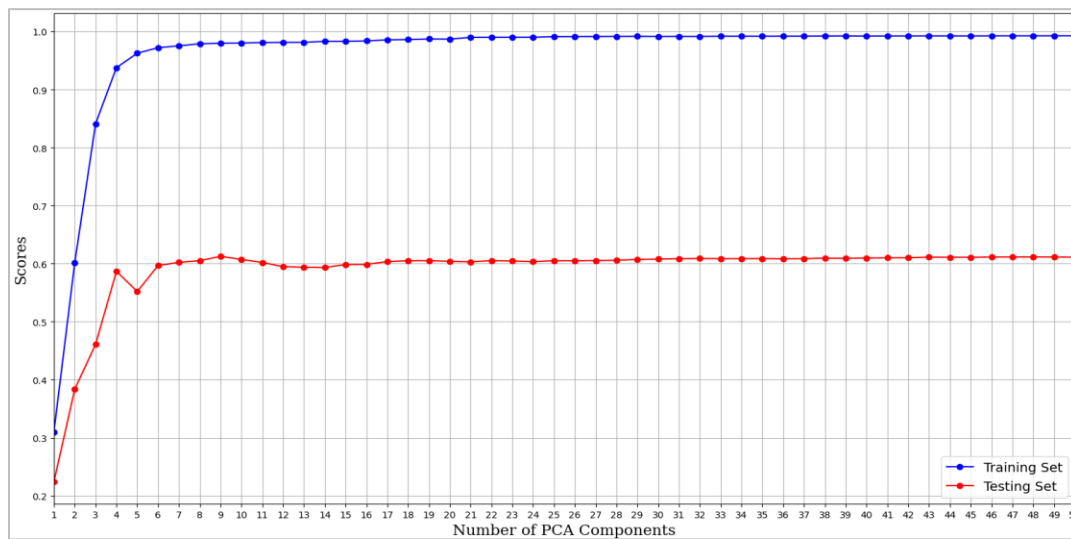| No. of PCs	Training Score	Testing Score	Total time
1 − 5	0.962	0.552	41.105
1 − 10	0.980	0.607	34.279
1 − 15	0.983	0.559	36.027
1 − 20	0.987	0.604	37.970
1 − 30	0.991	0.608	40.343
1 - 40	0.992	0.610	45.746
1 − 50	0.993	0.610	50.590
1 - 100	0.993	0.610	103.362
1 - 150	0.993	0.610	103.362
1 - 200	0.993	0.610	130.588
1− 240 (all)	0.993	0.610	156.060

Figure 18 Effect of the number of PCA components on: classification score, bottom: total running time (bottom)

40

4.2Support vector machine multi-classification

In this experiment, a grid-search procedure is applied on the training dataset to select the optimal kernel function as well as optimal hyperparameters values for soft margin constant (C) and γ values. A smaller value of C ignores points close to the boundary and increases the margin where γ >0 is a parameter that controls the width of RBF kernel function [70]. The SVM multi-classification is applied to hyperspectral and RGB as well as the ReliefF and PCA dimensionally reduced datasets using 10-fold cross validation. The overall classification accuracy values are tabulated in *Table 5* and corresponding confusion matrix are presented in *Figure 19*. Number of samples in validation set and testing set are 2000 and 200,000, respectively. It can be seen that the classification accuracy for 14 foods (out of 20) are improved in ReliefF reduced dataset over others.

Comparing the other 6 foods in all datasets, Potato, Rice Noodle, and Spinach (14, 15, and 16) do not achieve the desired results. RGB model performance is the lowest in recognizing foods with very similar color such as Black Bean and Black Olive (2 and 3). However, its ability to recognize Rice Noodle (15) is the best (50%,) comparing with other models. PCA, having the least number of features (less than 4% of all features) is the second-best model with nearly similar results as ReliefF. From the results, it has been found that the ReliefF which retains less than half of the features yields results better than RGB, PCA, and hyperspectral in terms of classification accuracy. However, PCA with only 9 features out of 240 is reasonable when more complex classifier is used to boost the accuracy (> 90%).
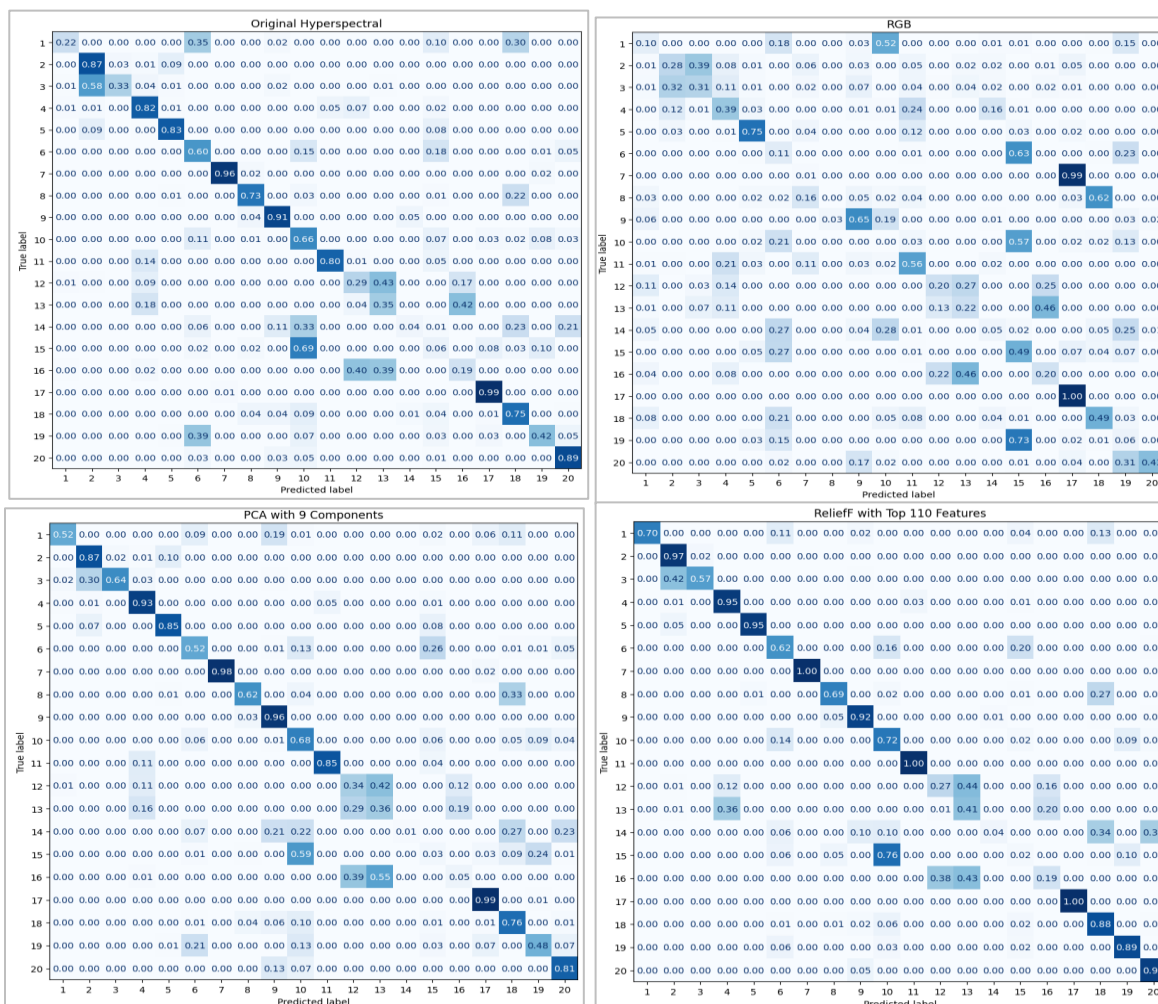

Figure 19 Confusion Matrix, top left: original hyperspectral, top right: RGB, bottom left: PCA, bottom right: ReliefF

Table 5 Multi-Classification Accuracy on RGB, hyperspectral, PCA reduced and ReliefF reduced datasets

Datasets	Model Spec.	Total FTRs	Validation Acc.	Testing Acc.
RGB	rbf, C=1000, γ =100	3	0.8513	0.3141
PCA	rbf, C=1000, γ =0.001	9	0.9713	0.6140
ReliefF	linear, C=1000	110	0.9966	0.6874
Hyperspectral	linear ,C=100	240	0.9948	0.5850

4.3 Logistic regression binary classification
4.3.1 Bean
Bean classification is the first level of the hierarchical binary classification in this work that separate bean samples (black, brown, and white bean) from others. Among 20,000 training samples, 3000 are beans that creates an imbalanced binary classification problem. To deal with unbalanced datasets, the inverse of label distribution is assigned as a weight to samples which means bean class, will have weight of 17 and non-bean class will use weight of 3. So, the penalty of wrong prediction of minority class (bean class) would be $\frac{17}{3}$ times more severe than wrong prediction of majority class (non-bean class).

The Logic Regression classifier is performed on RGB, original hyperspectral as well as PCA and ReliefF reduced datasets. Since logistic regression in sklearn library is used as binary classifier. There are five algorithms that can be selected for optimizations, which are liblinear, newton-cg, sag, saga, and lbfgs. Hyperparameters such as type of regulation method, and C, inverse of λ (regularization strength) variations should also be considered. Again, like the previous section, a grid search approach is used to search for optimal hyperparameters. Validation set results obtained from the ten-fold cross-validation procedure by considering 40% of training samples which are 8000 while there are 200,000 testing samples. The results are reported in *Figures 20* and *Tables 6,* and *Table 7*.

One important reason that bean classification does not have high scores is due to the fact that black bean and white bean samples have very similar color as some other samples such as black olive, white-rice, and egg whites.

4.3.2 Noodle
After separating bean samples, the rest are classified based on be-longing to the Noodle group. (1 vs. 0). There are three noodle types: egg, rice, and wheat. Each has 1000 samples in training dataset. Similar experiments have been done and the results are provided in *Figure 21* and *Tables 8 & 9* with 6800 and 170,000 samples in validation and testing datasets, respectively. Similar reason for nearly low scores of noodle classification especially on RGB dataset is due to color similarity between noodle samples and other samples such as rice noodle with potato or brown rice.

Table 6 Bean classification accuracy on different datasets

Datasets	Model Spec.	Total FTR	Validation Acc.	Testing Acc.
RGB	liblinear, C = 1, λ = L2	3	0.6665	0.5860
PCA	lbfgs, C = 0.01, λ = L2	9	0.9436	0.8102
ReliefF	liblinear, C = 10, λ = L2	110	0.9937	0.8818
Hyperspectral	liblinear, C = 10, λ = L2	240	0.9959	0.8953

Table 7 Evaluation metrics of bean classification on different datasets

Datasets	Precision	Recall	F1- score	AUC-ROC
RGB	0.26	0.96	0.41	0.74
PCA	0.43	0.82	0.56	0.81
ReliefF	0.57	0.92	0.70	0.90
Hyperspectral	0.60	0.94	0.73	0.91

Table 8 Noodle classification accuracy on different datasets

Datasets	Model Spec.	Total FTR	Validation Acc.	Testing Acc.
RGB	liblinear, C = 1, λ = L2	3	0.6531	0.5303
PCA	lbfgs, C = 0.01, λ = L2	9	0.7951	0.6168
ReliefF	liblinear, C = 10, λ = L2	110	0.9155	0.7164
Hyperspectral	liblinear, C = 10, λ = L2	240	0.9554	0.7212

Table 9 Evaluation metrics of noodle classification on different datasets

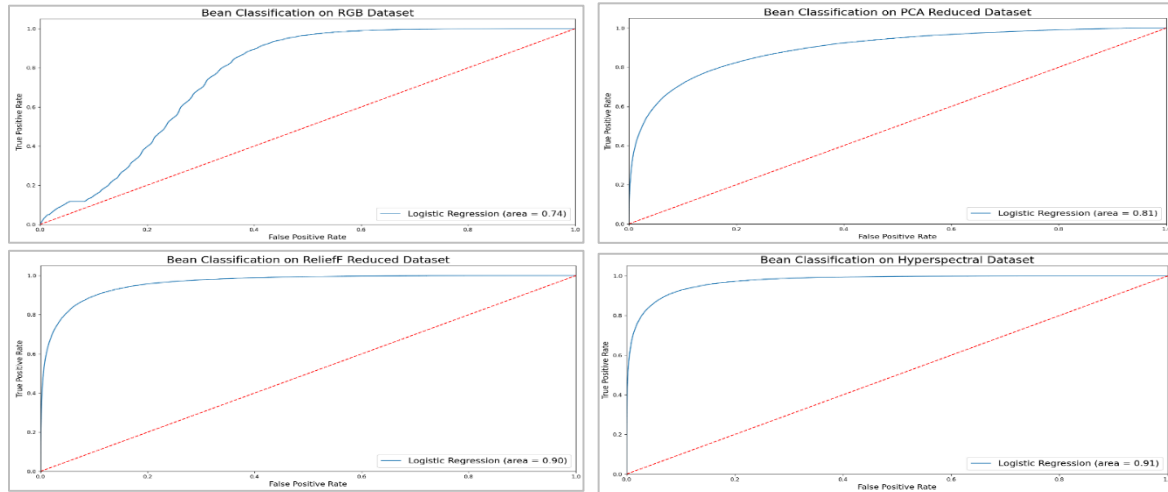| Datasets	Precision	Recall	F1- core	AUC-ROC
RGB	0.26	0.90	0.40	0.67
PCA	0.23	0.52	0.32	0.58
ReliefF	0.35	0.72	0.47	0.72
Hyperspectral	0.35	0.69	0.47	0.71

Figure 20 ROC curve of bean classification, top left: RGB, top right: PCA, bottom left: ReliefF, bottom right: original hyper spectral
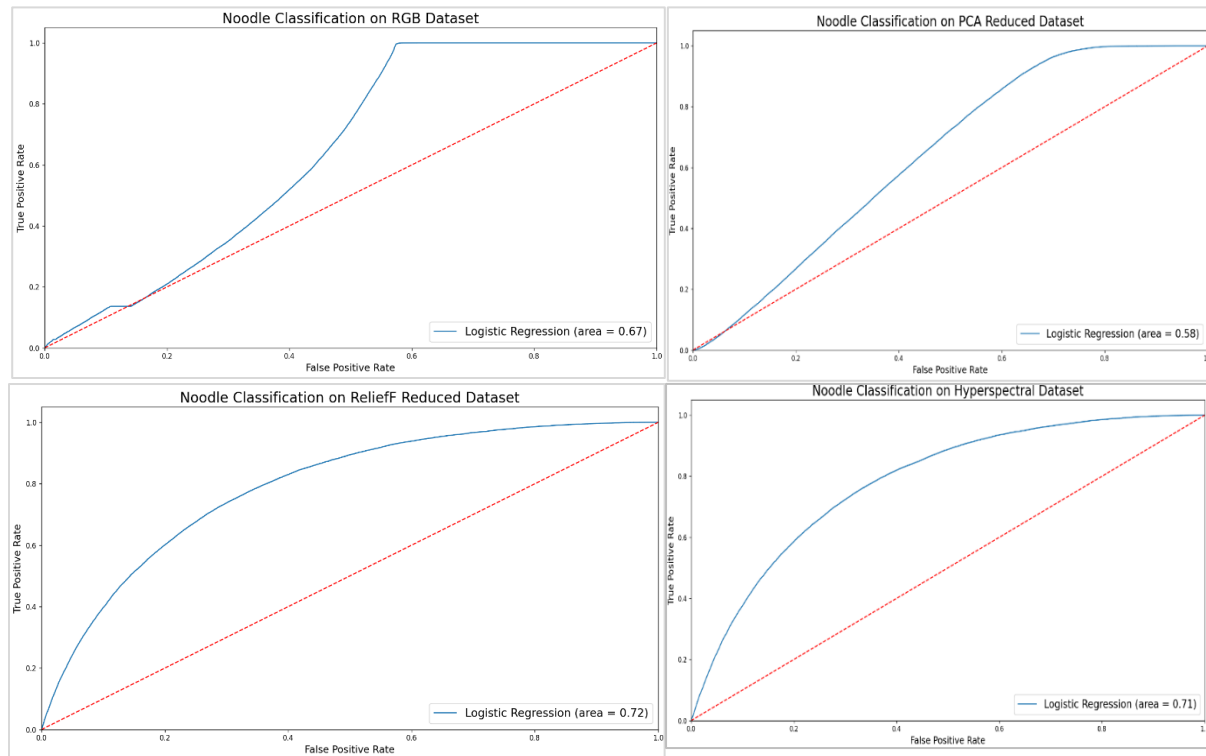

Figure 21 ROC curve of noodle classification, top left: RGB, top right: PCA, bottom left: ReliefF, bottom right: original hyperspectral

43

4.3.3Olive

The third step in the binary classification workflow is separating the olive samples (green and black) from the datasets. After removing the bean and noodle samples, there are 14000 training samples out of which 2000 of them are olive. The results are shown in *Figure 22* and *Tables 10 and Table 11*. The number of samples for validation and testing sets are

5600 and 140,000. Green and black samples do not have high color similarity in the remaining training dataset samples, so the performance of logistic regression on all four datasets is the highest compared with two previous results.

Table 10 Olive classification accuracy on different datasets

Datasets	Model Spec.	Total FTR	Validation Acc.	Testing Acc.
RGB	liblinear, C = 1, λ = L2	3	0.6968	0.6347
PCA	lbfgs, C = 0.01, λ = L2	9	0.9261	0.9747
ReliefF	liblinear, C = 10, λ = L2	110	0.9824	0.9850
Hyperspectral	liblinear, C = 10,λ = L2	240	0.9843	0.9898

Table 11 Evaluation metrics of olive classification on different datasets

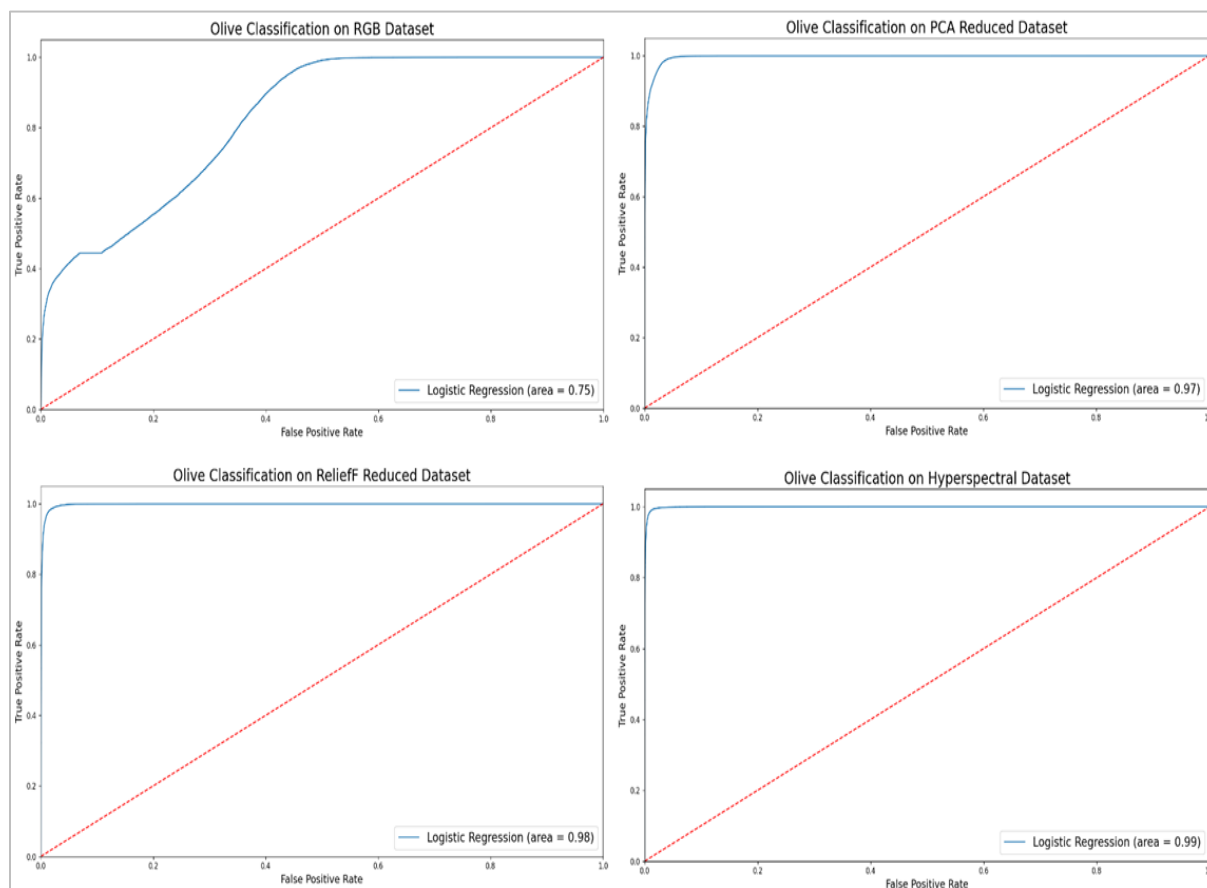| Datasets	Precision	Recall	F1- score	AUC-ROC
RGB	0.27	0.91	0.42	0.75
PCA	0.88	0.96	0.92	0.97
ReliefF	0.92	0.98	0.95	0.98
Hyperspectral	0.95	0.99	0.97	0.99

Figure 22 ROC curve of olive classification, top left: RGB, top right: PCA, bottom left: ReliefF, bottom right: original hyperspectral

4.3.4Rice

The last part of the hierarchical binary classification is on rice samples separation from others (10 different food samples) in the current datasets. Only brown rice has nearly similar color as potato. *Figure 23* and *Tables 12* and *13* show the rice classification results. There are 4800 samples in validation and testing set has 120,000 samples. Logistic Regression has the best results on hyperspectral and the worst on RGB datasets. The ReliefF reduced dataset provides the closer result to hyperspectral dataset and even better for noodle classification while having less than half of all features.

Table 12 Rice classification accuracy on different datasets

Datasets	Model Spec.	Total FTR	Validation Acc.	Testing Acc.
RGB	liblinear, C = 1, λ = L2	3	0.9044	0.7907
PCA	lbfgs, C = 0.01, λ = L2	9	0.9774	0.8354
ReliefF	liblinear, C = 10, λ = L2	110	0.9994	0.8507
Hyperspectral	liblinear, C = 10, λ = L2	240	0.9998	0.9081

Table 13 Evaluation metrics of olive classification on different datasets

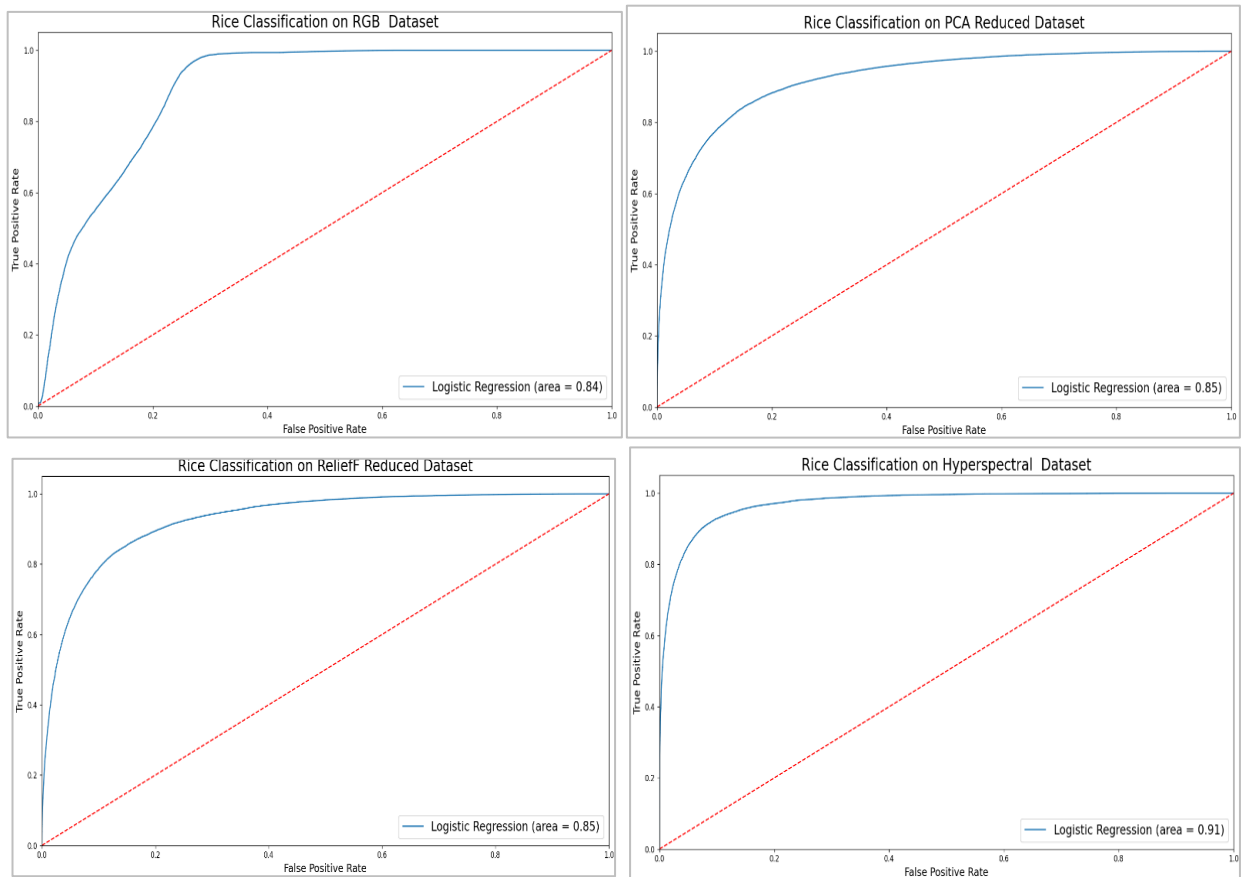| Datasets	Precision	Recall	F1- score	AUC-ROC
RGB	0.44	0.92	0.60	0.84
PCA	0.50	0.86	0.64	0.85
ReliefF	0.53	0.85	0.66	0.85
Hyperspectral	0.66	0.92	0.77	0.91

Figure 23 ROC curve of rice classification, top left: RGB, top right: PCA, bottom left: ReliefF, bottom right: original hyperspectral

5.Discussion

Rapid, objective, and non-destructive ingredients recognition is important for precise nutrition assessments especially for hospitalized patients. The traditional techniques are, however, inaccurate, destructive, and time-consuming which make them unsuitable for automatic nutrition assessment systems. The optical food image recognition techniques are currently among the most common methods in which food images are used to automatically (or semi- automatically) determine the type and the quantity of ingredients, then the nutritional values of the food can be computed. These approaches are widely used in mobile applications.

Although the conventional color images are used in most automatic nutritional assessment systems, utilizing only color information will result in the lower accuracy of ingredient recognition. This problem can be even exacerbated if there are same ingredient type samples with different cooking ways (different complex food may have the same ingredients), or different food types with nearly similar color. Thus, other features such as texture, shape, and size are required to improve the system performance.

In recent years, hyperspectral imaging (HSI), also known as chemical imaging, has shown to be an innovative and promising techniques for precise assessment of food. HSI techniques combine traditional optical spectroscopy and image processing methods into one system to obtain the spatial and spectral information simultaneously.

5.1Hyperspectral imaging advantages

Hyperspectral imaging techniques enable the development of a whole range of narrowband indices that are used for the determination of various characteristics. When analyzing food samples, those indices allow the retrieval of specific information, such as fat and water content in meat products, which is very difficult, even impossible, to achieve in experiments using conventional color images including RGB images. Furthermore, they have potential to increase the accuracy of food recognition techniques without having a large database of food images, especially in classifying different food samples with very similar color or similar food types with different colors.

5.2Hyperspectral imaging limitations

While in general the hyperspectral technology performs very well in many food processing systems, it does have some limitations. First, it is a costly technology, and cannot be used by individuals at home, but can be used at the food laboratories and medical facilities. Second, the amount of data is too large which increases the complexity and the processing time as well as causes *"Curse of dimensionality."* Finally, the spatial resolution is too low and in the case of requiring *high resolution images for specific algorithms such as edge detectors,* it is preferred to use conventional color or gray scale images.

6.Conclusion and future work

In this paper, RGB and hyperspectral features of 20 different types of food are used for food recognition. PCA and ReliefF algorithms are applied on hyperspectral datasets to obtain the determine lesser but most useful features. Two different classification models, SVM multi-classifier and Logistic Regression binary classifier are employed on all four mentioned datasets to provide an explicit comparison between different datasets. Among all, ReliefF with the top110 features has the best results in SVM and logistic regression respectively. Although PCA delivers good results comparing to RGB, ReliefF, and hyperspectral, it does not have satisfactory results in hierarchical binary classification. RGB data with only three-color features shows it's not suitable for recognition of food with similar color. We reviewed some previous studies of similar work in the introduction section. Most of the methods use RGB data for food recognition. Very few works on using HSI for food recognition have been reported. Of the research using HSI for food recognition, we find that the dataset is either proprietary or contains a small set of food types or a specific food type. To the best of our knowledge this work is only work that deals with comprehensive and mixed food types. Hence, comparison to other methods or works is not possible. The results of our research are comparable to the accuracies of other research work in RGB dataset. Though some RGB systems can achieve higher accuracies on average using deep learning methods on large database. Our methods have the potential to attain higher performances on a larger dataset using deep learning. Future work will be to improve the results by exploring deep learning as well as provide more samples for each food type.

Conflicts of interest
The authors have no conflicts of interest to declare.

References

[1] Esfahani SN, Muthukumar V, Regentova EE, Taghva K, Trabia M. Complex food recognition using hyper-spectral imagery. In annual computing and communication workshop and conference 2020 (pp. 662-7). IEEE.

[2] Khanna N, Boushey CJ, Kerr D, Okos M, Ebert DS, Delp EJ. An overview of the technology assisted dietary assessment project at Purdue University. In international symposium on multimedia 2010 (pp. 290-5). IEEE.

[3] Zhu F, Bosch M, Boushey CJ, Delp EJ. An image analysis system for dietary assessment and evaluation. In international conference on image processing 2010 (pp. 1853-6). IEEE.

[4] Zhu F, Bosch M, Woo I, Kim S, Boushey CJ, Ebert DS, Delp EJ. The use of mobile devices in aiding dietary assessment and evaluation. IEEE Journal of Selected Topics in Signal Processing. 2010; 4(4):756-66.

[5] Kong F, Tan J. Dietcam: regular shape food recognition with a camera phone. In international conference on body sensor networks 2011 (pp. 127-32). IEEE.

[6] Bosch M, Zhu F, Khanna N, Boushey CJ, Delp EJ. Combining global and local features for food identification in dietary assessment. In 18th international conference on image processing 2011 (pp. 1789-92). IEEE.

[7] Xu C, Khanna N, Boushey CJ, Delp EJ. Low complexity image quality measures for dietary assessment using mobile devices. In international symposium on multimedia 2011 (pp. 351-6). IEEE.

[8] Rahman MH, Li Q, Pickering M, Frater M, Kerr D, Bouchey C, et al. Food volume estimation in a mobile phone based dietary assessment system. In international conference on signal image technology and internet based systems 2012 (pp. 988-95). IEEE.

[9] Zhu F, Bosch M, Khanna N, Boushey CJ, Delp EJ. Multiple hypotheses image segmentation and classification with application to dietary assessment. IEEE Journal of Biomedical and Health Informatics. 2014; 19(1):377-88.

[10] Liu C, Cao Y, Luo Y, Chen G, Vokkarane V, Yunsheng M, et al. A new deep learning-based food recognition system for dietary assessment on an edge computing service infrastructure. IEEE Transactions on Services Computing. 2017; 11(2):249-61.

[11] Jiang L, Qiu B, Liu X, Huang C, Lin K. DeepFood: food image analysis and dietary assessment via deep model. IEEE Access. 2020; 8:47477-89.

[12] Tan RZ, Chew X, Khaw KW. Quantized deep residual convolutional neural network for image-based dietary assessment. IEEE Access. 2020; 8:111875-88.

[13] Chen M, Dhingra K, Wu W, Yang L, Sukthankar R, Yang J. PFID: pittsburgh fast-food image dataset. In international conference on image processing 2009 (pp. 289-92). IEEE.

[14] Chen HC, Jia W, Li Z, Sun YN, Sun M. 3D/2D model-to-image registration for quantitative dietary assessment. In annual northeast bioengineering conference 2012 (pp. 95-6). IEEE.

[15] He J, Shao Z, Wright J, Kerr D, Boushey C, Zhu F. Multi-task image-based dietary assessment for food recognition and portion size estimation. In conference on multimedia information processing and retrieval 2020 (pp. 49-54). IEEE.

[16] Hafiz R, Islam S, Khanom R, Uddin MS. Image based drinks identification for dietary assessment. In international workshop on computational intelligence 2016 (pp. 192-7). IEEE.

[17] Lo FP, Sun Y, Qiu J, Lo B. A novel vision-based approach for dietary assessment using deep learning view synthesis. In 16th international conference on wearable and implantable body sensor networks 2019 (pp. 1-4). IEEE.

[18] He Y, Xu C, Khanna N, Boushey CJ, Delp EJ. Analysis of food images: Features and classification. In international conference on image processing 2014 (pp. 2744-8). IEEE.

[19] Lo FP, Sun Y, Qiu J, Lo BP. Point2volume: a vision-based dietary assessment approach using view synthesis. IEEE Transactions on Industrial Informatics. 2019; 16(1):577-86.

[20] Subhi MA, Ali SH, Mohammed MA. Vision-based approaches for automatic food recognition and dietary assessment: a survey. IEEE Access. 2019; 7:35370-81.

[21] Park B, Lu R, editors. Hyperspectral imaging technology in food and agriculture. New York: Springer; 2015.

[22] Sun DW. Hyperspectral imaging for food quality analysis and control. Elsevier; 2010.

[23] Bhunia AK, Kim MS, Taitt CR, editors. High throughput screening for food safety assessment: Biosensor technologies, hyperspectral imaging and practical applications. Elsevier; 2014.

[24] Haboudane D, Miller JR, Pattey E, Zarco-Tejada PJ, Strachan IB. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: modeling and validation in the context of precision agriculture. Remote Sensing of Environment. 2004; 90(3):337-52.

[25] Haboudane D, Miller JR, Tremblay N, Zarco-Tejada PJ, Dextraze L. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. Remote Sensing of Environment. 2002; 81(2-3):416-26.

[26] Rodarmel C, Shan J. Principal component analysis for hyperspectral image classification. Surveying and Land Information Science. 2002; 62(2):115-22.

[27] Chang CI, Du Q, Sun TL, Althouse ML. A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing. 1999; 37(6):2631-41.

[28] Guo B, Gunn SR, Damper RI, Nelson JD. Band selection for hyperspectral image classification using mutual information. IEEE Geoscience and Remote Sensing Letters. 2006; 3(4):522-6.

[29] MartÍnez-UsÓMartinez-Uso A, Pla F, Sotoca JM, García-Sevilla P. Clustering-based hyperspectral band selection using information measures. IEEE Transactions on Geoscience and Remote Sensing. 2007; 45(12):4158-71.

[30] Du Q, Yang H. Similarity-based unsupervised band selection for hyperspectral image analysis. IEEE Geoscience and Remote Sensing Letters. 2008; 5(4):564-8.

[31] Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines. IEEE Transactions on Geoscience and Remote Sensing. 2004; 42(8):1778-90.

[32] Gualtieri JA, Chettri S. Support vector machines for classification of hyperspectral data. IEEE 2000 international geoscience and remote sensing symposium. Taking the pulse of the planet: the role of remote sensing in managing the environment. Proceedings 2000 (pp. 813-5). IEEE.

[33] Puri M, Zhu Z, Yu Q, Divakaran A, Sawhney H. Recognition and volume estimation of food intake using a mobile device. In workshop on applications of computer vision 2009 (pp. 1-8). IEEE.

[34] Rahmana MH, Pickering MR, Kerr D, Boushey CJ, Delp EJ. A new texture feature for improved food recognition accuracy in a mobile phone based dietary assessment system. In international conference on multimedia and expo workshops 2012 (pp. 418-23). IEEE.

[35] Pouladzadeh P, Villalobos G, Almaghrabi R, Shirmohammadi S. A novel SVM based food recognition method for calorie measurement applications. In international conference on multimedia and expo workshops 2012 (pp. 495-8). IEEE.

[36] Duan P, Wang W, Zhang W, Gong F, Zhang P, Rao Y. Food image recognition using pervasive cloud computing. In international conference on green computing and communications and IEEE internet of things and IEEE cyber, physical and social computing 2013 (pp. 1631-7). IEEE.

[37] Pham C, Thanh TN. Fresh food recognition using feature fusion. In international conference on advanced technologies for communications 2014 (pp. 298-302). IEEE.

[38] Zhang W, Zhao D, Gong W, Li Z, Lu Q, Yang S. Food image recognition with convolutional neural networks. In international conference on Ubiquitous intelligence and computing and international conference on autonomic and trusted computing and 2015 IEEE 15th international conference on Scalable Computing and communications and its associated workshops (UIC-ATC-ScalCom) 2015 (pp. 690-3). IEEE.

[39] Pouladzadeh P, Kuhad P, Peddi SV, Yassine A, Shirmohammadi S. Food calorie measurement using deep learning neural network. In international instrumentation and measurement technology conference proceedings 2016 (pp. 1-6). IEEE.

[40] He H, Kong F, Tan J. DietCam: multiview food recognition using a multikernel SVM. IEEE Journal of Biomedical and Health Informatics. 2015; 20(3):848-55.

[41] Chen J, Ngo CW. Deep-based ingredient recognition for cooking recipe retrieval. In proceedings of the ACM international conference on Multimedia 2016 (pp. 32-41). ACM.

[42] Lu Y. Food image recognition by using convolutional neural networks (cnns). arXiv preprint arXiv:1612.00983. 2016.

[43] Ciocca G, Napoletano P, Schettini R. Food recognition: a new dataset, experiments, and results. IEEE Journal of Biomedical and Health Informatics. 2016; 21(3):588-98.

[44] Chu WT, Lin JH. Food image description based on deep-based joint food category, ingredient, and cooking method recognition. In international conference on multimedia & expo workshops 2017 (pp. 109-14). IEEE.

[45] Deng L, Chen J, Sun Q, He X, Tang S, Ming Z, et al. Mixed-dish recognition with contextual relation networks. In proceedings of the ACM international conference on multimedia 2019 (pp. 112-20).

[46] Zhang L, Zhao J, Li S, Shi B, Duan LY. From market to dish: multi-ingredient image recognition for personalized recipe recommendation. In international conference on multimedia and expo 2019 (pp. 1252-7). IEEE.

[47] Liu C, Liang Y, Xue Y, Qian X, Fu J. Food and ingredient joint learning for fine-grained recognition. IEEE Transactions on Circuits and Systems for Video Technology. 2020.

[48] Liang H, Wen G, Hu Y, Luo M, Yang P, Xu Y. MVANet: multi-tasks guided multi-view attention network for Chinese food recognition. IEEE Transactions on Multimedia. 2020:1-12.

[49] Freitas CN, Cordeiro FR, Macario V. MyFood: a food segmentation and classification system to Aid nutritional monitoring. In SIBGRAPI conference on graphics, patterns and images 2020 (pp. 234-9). IEEE.

[50] Chen J, Zhu B, Ngo CW, Chua TS, Jiang YG. A study of multi-task and region-wise deep learning for food ingredient recognition. IEEE Transactions on Image Processing. 2020; 30:1514-26.

[51] Barbin D, Elmasry G, Sun DW, Allen P. Near-infrared hyperspectral imaging for grading and classification of pork. Meat Science. 2012; 90(1):259-68.

[52] Cluff K, Naganathan GK, Subbiah J, Samal A, Calkins CR. Optical scattering with hyperspectral imaging to classify longissimus dorsi muscle based on beef tenderness using multivariate modeling. Meat Science. 2013; 95(1):42-50.

[53] Marshall S, Kelman T, Qiao T, Murray P, Zabalza J. Hyperspectral imaging for food applications. In European signal processing conference 2015 (pp. 2854-8). IEEE.

[54] Wang L, Liu D, Pu H, Sun DW, Gao W, Xiong Z. Use of hyperspectral imaging to discriminate the variety

and quality of rice. Food Analytical Methods. 2015; 8(2):515-23.

[55] Al-Sarayreh M, Reis MM, Yan WQ, Klette R. Deep spectral-spatial features of snapshot hyperspectral images for red-meat classification. In international conference on image and vision computing 2018 (pp. 1-6). IEEE.

[56] Qiu Z, Chen J, Zhao Y, Zhu S, He Y, Zhang C. Variety identification of single rice seed using hyperspectral imaging combined with convolutional neural network. Applied Sciences. 2018; 8(2):1-12.

[57] Liu Y, Wu T, Yang J, Tan K, Wang S. Hyperspectral band selection for soybean classification based on information measure in FRS theory. Biosystems Engineering. 2019; 178:219-32.

[58] Steinbrener J, Posch K, Leitner R. Hyperspectral fruit and vegetable classification using convolutional neural networks. Computers and Electronics in Agriculture. 2019; 162:364-72.

[59] Xia C, Yang S, Huang M, Zhu Q, Guo Y, Qin J. Maize seed classification using hyperspectral image coupled with multi-linear discriminant analysis. Infrared Physics & Technology. 2019; 103:103077.

[60] Xiao H, Guo P, Dong X, Xing S, Sun M. Research on the method of hyperspectral and image deep features for bacon classification. In Chinese control and decision conference 2019 (pp. 4682-6). IEEE.

[61] Weng S, Tang P, Yuan H, Guo B, Yu S, Huang L, Xu C. Hyperspectral imaging for accurate determination of rice variety using a deep learning network with multi-feature fusion. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy. 2020; 234:118237.

[62] Fabiyi SD, Vu H, Tachtatzis C, Murray P, Harle D, Dao TK, et al. Varietal classification of rice seeds using RGB and hyperspectral images. IEEE Access. 2020; 8:22493-505.

[63] Zhang C, Zhao Y, Yan T, Bai X, Xiao Q, Gao P, et al. Application of near-infrared hyperspectral imaging for variety identification of coated maize kernels with deep learning. Infrared Physics & Technology. 2020; 111:103550.

[64] Furlanetto RH, Moriwaki T, Falcioni R, Pattaro M, Vollmann A, Junior AC, et al. Hyperspectral reflectance imaging to classify lettuce varieties by optimum selected wavelengths and linear discriminant analysis. Remote Sensing Applications: Society and Environment. 2020; 20:100400.

[65] Kononenko I, Robnik-Sikonja M, Pompe U. ReliefF for estimation and discretization of attributes in classification, regression, and ILP problems. Artificial Intelligence: Methodology, Systems, Applications. 1996:31-40.

[66] Kira K, Rendell LA. A practical approach to feature selection. In machine learning proceedings 1992 (pp. 249-56). Morgan Kaufmann.

[67] Qian Y, Ye M, Zhou J. Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features. IEEE Transactions on Geoscience and Remote Sensing. 2012; 51(4):2276-91.

[68] Yang Y, Liu X. A re-examination of text categorization methods. In proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval 1999 (pp. 42-9).

[69] Zahavi A, Palshin A, Liyanage DC, Tamre M. Influence of illumination sources on hyperspectral imaging. In international conference on research and education in mechatronics 2019 (pp. 1-5). IEEE.

[70] Ben-Hur A, Weston J. A user's guide to support vector machines. In data mining techniques for the life sciences 2010 (pp. 223-39). Humana Press.

Shirin Nasr Esfahani received the B.S. degree from University of Tabriz, Tabriz- Iran and her M.S. degree from Sharif University of Technology, Tehran- Iran, both in Computer Science. She is currently a Ph.D. candidate in Computer Science, University of Nevada, Las Vegas. Her fields of interest include Hyperspectral Image Processing, Medical and Microscopic Image Processing, Machine Learning, and Deep Learning.
Email: nasresfa@unlv.nevada.edu

Venkatesan Muthukumar received the B.E. degree from College of Engineering, Anna University, INDIA, and the M.S. and Ph.D. degrees from Monash University, Australia, in 1996 and 2000, respectively. He is currently a Professor in Electrical and Computer Engineering Department at the University of Nevada Las Vegas. His current research interests include Embedded Systems, Wearable Devices, Network on Chip, High Performance Computing on System on Chips, Multicores and Multiprocessor systems.
Email: venkatesan.muthukumar@unlv.edu

Emma Regentova is a Professor in Electrical and Computer Engineering Department at the University of Nevada, Las Vegas. She has been working and publishing her research performed for numerous funded projects on the topics in Image processing, Computed tomography and compressed sensing, Remote sensing and hyperspectral image analysis, Computer assisted medical diagnostics, Biomedical research, Data compression, and coding, Intelligent Transportation Systems, Neutron and photon imaging for security and inspection, Digital system design, and embedded DSP.
Email: emma.regentova@unlv.edu

Kazem Taghva received his Ph.D in 1980 from the University of Iowa. He is currently the Chair and a Professor in Computer Science Department at the University of Nevada, Las Vegas. Prior to joining UNLV, he was chairman of the Computer Science Department at New Mexico Tech. His research includes Information Retrieval (IR), Machine Learning (ML), and Database Systems (DBMS). His work in IR is focused on retrieval from noisy text (OCR), and in ML covers Named Entity Recognition including proper nouns, addresses, and acronyms, and in DBMS deals with dependency theory. He was supported by Department of Energy for 25 years to construct the largest government database of documents for legal discovery (Licensing Support Network)
Email: kazem.taghva@unlv.edu

Mohamed B. Trabia is an Associate Dean for Research, Graduate Studies, and Computing and a Professor in Mechanical Engineering Department at University of Nevada Las Vegas. He received his B.S. and M.S. degrees from Alexandria University, Alexandria, Egypt, and Ph.D. degree from Arizona State University, all in Mechanical Engineering. His research interests include Optimization Algorithms and their Applications to Mechanical Engineering Design, Finite Element Analysis of Mechanical Components and Systems, Dynamic Analysis and Control of Mechanical Systems with Emphasis on Flexible Robots, Analysis and Design of Robots and Mechanisms, Biomedical applications of mechanical design, Characterization of Biomaterials, and Path Planning and Obstacle Avoidance of Mobile Robots.
Email: mohamed.trabia@unlv.edu