IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 8, 2020, accepted September 23, 2020, date of publication September 25, 2020,
date of current version October 8, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3026925

Prediction of Microbial Spoilage and Shelf-Life of Bakery Products Through Hyperspectral Imaging

ZAINAB SALEEM[1], MUHAMMAD HUSSAIN KHAN[1], MUHAMMAD AHMAD[2], AHMED SOHAIB[1], HAMAIL AYAZ[1], AND MANUEL MAZZARA[3]

[1]Advance Image Processing Research Laboratory (AIPRL), Department of Computer Engineering, Khwaja Freed University of Engineering and Technology (KFUEIT), Rahim Yar Khan 64200, Pakistan
[2]Department of Computer Science, National University of Computer and Emerging Sciences, Chiniot-Faisalabad Campus, Chiniot 35400, Pakistan
[3]Institute of Software Development and Engineering, Innopolis University, 420500 Innopolis, Russia

Corresponding author: Muhammad Ahmad (mahmad00@gmail.com)

ABSTRACT The shelf life of bakery products highly depends on the environment and it may get spoiled earlier than its expiry which results in food-borne diseases and may affect human health or may get wasted beforehand. The traditional spoilage detection methods are time-consuming and destructive in nature due to the time taken to get microbiological results. To the best of the author's knowledge, this work presents a novel method to automatically predict the microbial spoilage and detect its spatial location in baked items using Hyperspectral Imaging (HSI) range from $395 - 1000$ *nm*. A spectral preserve fusion technique has been proposed to spatially enhance the HSI images while preserving the spectral information. Furthermore, to automatically detect the spoilage, Principal Component Analysis (PCA) followed by K-means and SVM has been used. The proposed approach can detect the spoilage almost 24 hours before it started appearing or visible to a naked eye with 98.13% accuracy on test data. Furthermore, the trained model has been validated through external dataset and detected the spoilage almost a day before it started appearing visually.

INDEX TERMS Shelf life of bakery products, fungus detection and prediction, PCA, K-means, SVM, hyper sharpening.

I. INTRODUCTION

Bakery products have been an integral part of the majority of the world population's daily routine [1]. Baked products especially bread and sponge cakes remained popular and high in demand due to its nutritional, taste, and appearance qualities. While on the other hand there are some problems reported regarding baked food quality and shelf-life like food-borne diseases [2]. Every baked product has a pre-defined shelf-life depending upon the type of ingredients used in it or the environment it is placed [3], such as the shelf-life of the baked product may vary due to the variations in basic ingredients (flour, egg, etc), baking process or environmental conditions (temperature, amount of humidity, etc.,) [4]. The variations in the estimated shelf-life of bakery products may result in two ways;

1) **Moderate Environment:** The growth of bacterial microorganisms slow-down in a moderate environment, neither too humid nor dry, which results in enhancing the shelf-life. On the other hand, every product has a pre-defined shelf-life but if it ends before the actual spoilage occurs, the product still got wasted only because it has reached its expiry date.
2) **Extreme Environment:** Extreme humid or drier environment results in affecting the moisture content of the product which may lead to the early spoilage [5]. Therefore, the food-item spoiled before its expiry but kept on consumed by individuals hence results in harmful effects on human health. It has been estimated that 1 out of 10 (600 million per year) humans suffer and around 420, 000 died due to food-borne diseases [6]. World Health Organization (WHO) estimated that almost 33 million years of healthy lives have been lost due to the consumption of unhygienic food and this number is still an underestimation. Food-borne diseases are mostly due to microbial spoilage. Therefore, early detection of food spoilage is of utmost importance to limit the aforementioned concerns [7].

Detection of early spoilage is of utmost importance and being investigated since long but yet not fully explored.

The associate editor coordinating the review of this manuscript and approving it for publication was Qingli Li.

The traditional spoilage detection methods are Molecular-Level Analysis (MLA) such as confocal microscopy [8], immunology-based methods [9], polymerase chain reaction [10] and quartz crystal microbalance [11]. On the other hand, modern techniques include the use of ultrasound [12], electronic nose, digital images, and spectroscopy [13] to detect the spoilage.

MLA detects food contamination to analyze bacterial spoilage and fungus. This process is accurate, however, time-consuming and difficult to conduct as it performs the chemical analysis of the product by using traditional microscopic techniques. It is also destructive and expensive in real-time situations that require skilled labor. For instance, counting the developed bacterial colonies (a common approach for microorganisms identification) is laborious and may often take $2-3$ days for initial results and up to 7 days for confirmation [14].

Confocal microscopy has been used to detect the early growth of bacterial micro-organisms to prevent the negative effects of consuming exposed food items. Confocal microscopy is also time-consuming especially when a large number of samples are being analyzed [8]. Furthermore, some immunology-based methods have been employed for the detection of microbial contamination but these techniques show low sensitivity when the samples contain the excessive amount of microbial microorganisms [9]. Other alternatives like polymerase chain reaction [10], Biofunctional Magnetic Nano-Particles with bio-luminescence [15] and Quartz Crystal Microbalance [11] are fast, however, scientifically complicated as it requires skilled labor as well as it is expensive to get microbiological results.

Modern techniques such as ultrasound technology have been used in the characterization of food quality [12]. However, it is difficult to detect whether the food is contaminated or not at an early stage [16]. The spoilage detection using traditional digital imaging (RGB) is quite a good alternative but it has three major limitations;

1) Traditional digital imaging does not provide information regarding the chemical composition of the materials.
2) Traditional digital imaging covers only limited bands in the visible range ($(400-700\ nm)$) of the electromagnetic spectrum whereas many substances may not reflect in this range.
3) Traditional digital imaging lacks in differentiating the objects with similar visual properties such as shape, color and also lacks in the detection of invisible defects [17].

Spectroscopy is being used to overcome the aforementioned limitations by providing the information regarding the chemical composition of an object and nature of the material [13], [18]. These sensors describe the behavior of light that interacts with an element on different electromagnetic wavelengths (spectral resolution). A limitation of spectroscopy is that it does not provide information about ''where'' and ''how'' the micro-bacterial elements are distributed in the product. The spatial information is important to visualize the disbursement as the measurement of the spoiled area depends upon that portion of the sample [19]. Therefore, to analyze the spoilage as well as to acquire data about the shape and location, Imaging Spectroscopy has been introduced that commonly referred to as Spectral Imaging (SI).

SI is a combination of digital imaging and spectroscopy [20]. In SI, the system that acquires tens or fewer channels/bands is termed as Multispectral Imaging System (MSI). On the contrary, the system that has ten to a few hundred bands referred to as the Hyperspectral Imaging (HSI) [18], [21]. HSI captures data in the form of a hypercube by storing it in two spatial coordinates (x and y) and one spectral dimension (λ) [22]. The spectral dimension of a hypercube describes the capability of an HSI system to store the nature of an object corresponding to each wavelength. Each pixel of a hypercube shows a complete spectrum and the reaction of each wavelength band corresponding to that pixel. HSI has been used for decades in remote sensing [23], [24] while later it became an eminent technology in many other fields such as; food quality and assessment [13], [25], [26], forensic [27], medical [28], defence and home security [29], etc.

HSI has been used to detect fungal, bacterial, and microbial spoilage in several food items. Early detection of spoilage, due to infection or disease in food items/products, helps to take precautionary measures to prevent the disease from spreading on a wider level. To the best of our knowledge, HSI research for bakery items especially cakes and bread is only conducted in a very limited number. For instance, the quality of butter cookies during the baking process has been assessed using Multispectral Imaging by Andresen Mette et. al., [30]. Cookies were accessed on three different levels (under baked, adequate, and over baked) by determining the water content in samples. Adam Polak et. al., [31] predicted the moisture content and hardness of cakes (white and chocolate sponges) using HSI. The authors used reflectance spectra to predict quantitatively the moisture content and hardness simultaneously. The developed model performed well for white sponge cakes as compared to chocolate ones. Noha Morsy et.al., [32] predicted the microbial spoilage in bread using the NIR imaging system. The spoilage was analyzed by PCA and predicted with the help of Partial Least Squares. In this paper, the authors have only predicted the presence of spoilage in the sample.

This work proposes a novel HSI-based technique that not only automatically detects and predicts the microbial spoilage in bakery products but also its spatial location. This work mainly focused on change detection techniques usually deployed in remote sensing [33]. To the best of author's knowledge, this work proposes a novel method to detect and predict the spatial location of microbial spoilage in baked items using HSI. The analysis of microbial change in baked cake sponges has been made spatially, therefore the size of the cake images needed to be the same. While with time, the cake sponges change size due to the enhancement or reduction of moisture content. Therefore, a spectral preserve hyper

sharpening technique has been used to spatially enhance the HSI images while preserving the spectral information. Furthermore, to automatically detect the spoilage PCA followed by K-means has been used.

The rest of the paper is structured as follows. Section II presents the methodologies of sample preparation, data acquisition, spectral pre-processing, hyper sharpening, microbial spoilage detection, and model training using SVM. Section III contains the results and discussion of experimental evaluation and shows the spoilage prediction by using our proposed technique. Finally, section IV summarizes the major contributions of this study and potential future research directions that can be derived from this work.

II. MATERIAL AND METHODS

This section provides information regarding research equipment, sample preparation, data acquisition, and data preprocessing steps used. It further contains the information regarding our two novel contributions, i.e., image fusion and automatic detection of microbial spoilage.

A. HYPERSPECTRAL IMAGING (HSI) SYSTEM

A Specim FX-10 (Specim, Spectral Imaging Ltd, Oulu, Finland, A-line scanner) Hyperspectral camera is used to acquire the samples as shown in Fig. 1. This camera can capture a visible and near infra-red (VNIR) electromagnetic spectrum ranging from $395 - 1000$ *nm* with a spectral sampling of 2.7 *nm* with a total of 224 spectral bands. The camera is coupled with the $1.4/8$ *mm* lens (Scheiner Cinegon). The lab scanner consists of three halogen lamps and a translational platform of size 21 *cm* $\times$ 40 *cm* which is controllable through a computer via the serial communication port though GigE-Vision. The camera is mounted on a lab scanner at 30 *cm* high above the moving tray. The entire HSI system was kept in a dark box to avoid ambient noises.

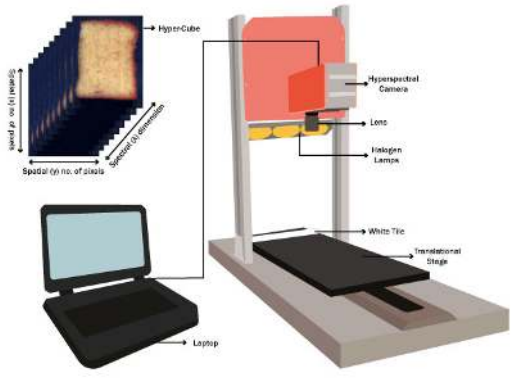


FIGURE 1. Hyperspectral imaging system used in this study.

B. SAMPLE PREPARATION

Simple sponge cake samples were used for analyzing the microbial spoilage growth. These samples were prepared according to the recipe shown in Table 1. The liquid and non-liquid ingredients were mixed separately and the mixtures were blended using an electric blender for 5 minutes. The amalgam was then baked in a pre-heated microwave oven at $180°F$ for 30 minutes.

TABLE 1. Ingredients for baking sample cakes.

Ingredient	Quantity
All purpose flour	120 grams
Sugar	90 grams
Butter	70 grams
Eggs	2
Vanilla Essence	1 tea spoon
Baking Powder	2 tea spoon

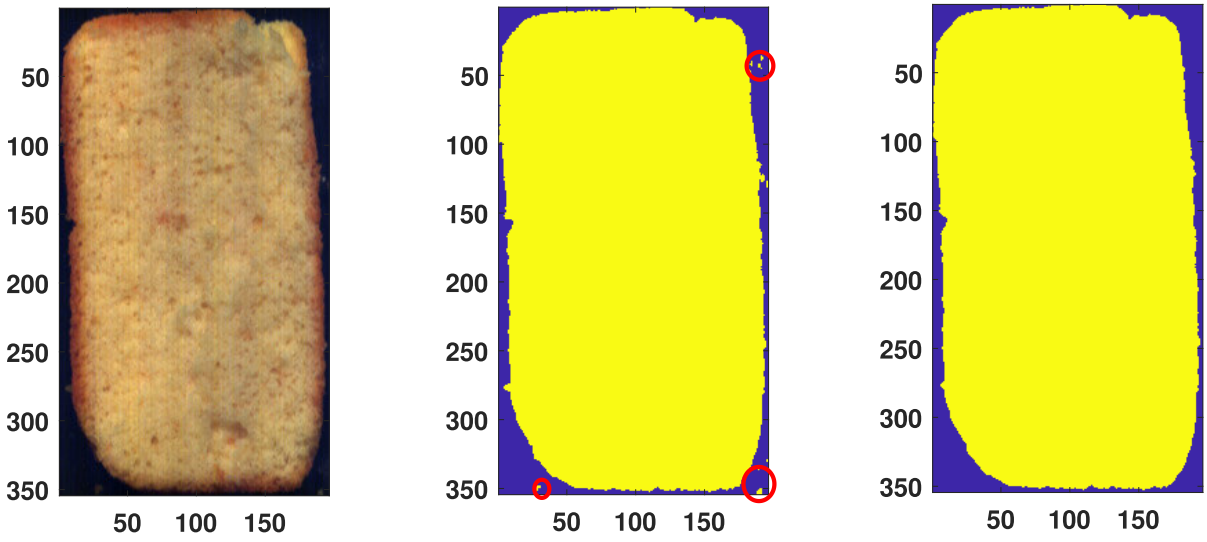
The samples were cooled down at room temperature and wrapped with a plastic sheet to avoid contamination from dust in the air. Samples were then sliced in 10 *cm* in length, 6 *cm* in width, and 2 *cm* depth and kept in the airtight box. Furthermore, a temperature sensor was fixed within the box to determine the current and the average temperature over the span of 8 days which was 13 degree Celsius. Data acquisition was made after almost every 12 hours and the samples were sprayed with distilled water mist after every acquisition to increase the moisture content to grow the microbial spoilage bacteria in cake sponges. Finally, a total of 16 samples were acquired over a span of 8 days.

A glass plate was used as a cake sponge holder for data acquisition and markers were placed on the moving transnational medium to make the sample's place consistent for each acquisition. The exposure time was set as 16 *ms* while the frame rate was 60 *Hz*. The translational platform was moving with a speed of 20 mm/s. The false-color image of the sample was created by choosing the red band at 699.09 *nm*, green at 591.04 *nm*, and blue at 450.16 *nm*. Three halogen lamps were used as a light source. These were placed with a vertical orientation and titled at 45 degrees. The camera was positioned at a distance of 30 *cm* above from the samples.

C. SPECTRAL-SPATIAL PRE-PROCESSING

The HSI system records radiance of sample along with background and reference material and stores it in raw format. To standardize the dataset and to reduce the computational cost, samples were segmented from the background using thresholding and morphological operations. Further to exclude environmental effects, source illumination, and geometry of reflected light, the reflectance of each sample was calculated using an empirical line method. Furthermore, image segmentation followed by thresholding (Fig. 2b) and morphological operations (Fig. 2c) on false-color image (Fig. 2a) has been performed to remove the background.

HSI system acquires a hyper-cube by the information reflected from the sample along with environmental effects. These effects can create false results in the study about the pure nature of the sample. Therefore, the data needs to be corrected and the reflectance to be calculated



(a) False Image (b) Threshold-ed (c) Eroded and Dilated

FIGURE 2. Results of various steps of image segmentation.

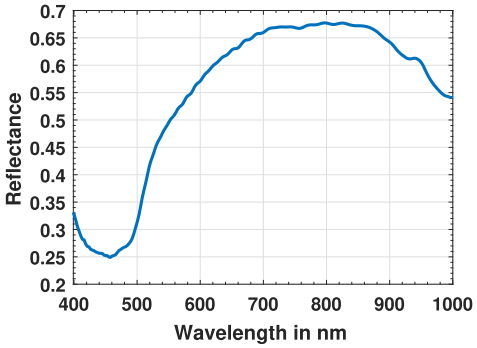


FIGURE 3. Mean reflectance spectra.

before analyzing. Calculating actual reflectance is one of the major steps in data pre-processing. This study caters to this issue by using the darkest pixel (DP) subtraction method which utilizes the external dark and White reference along with a sample object to remove the sensor effect and camera optics from raw data.

The dark reference image (D) is acquired with a closed camera aperture and completely covering the lens with the cap. The white reference image (W) was obtained by placing a white tile which has almost 99.9% reflectance. The calibrated image (IR) was calculated by using the following equation and is shown in Fig. 3.

$$IR = \frac{I_o - D}{W - D} \tag{1}$$

where, I_o is original raw hyper-cube, W is white reference image and D is dark reference image.

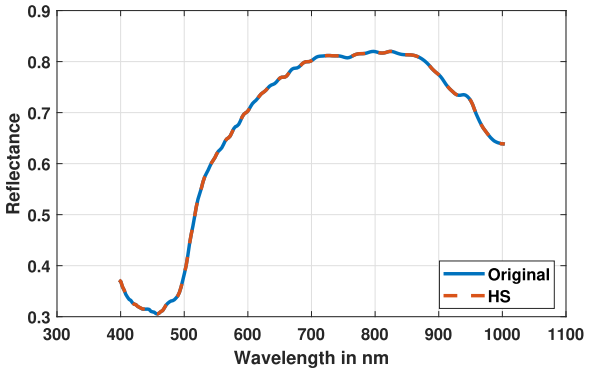
D. IMAGE FUSION TO ENRICH SPATIAL RESOLUTION

The food items tend to change their size with respect to the number of days [34], for instance, if the surrounding is humid, the size may increase due to moisture content whereas, if the surrounding is warmer and drier, it may shrink in size due to the decrease in moisture. Cake sponges used in this study also changed the size with the passage of days. This study detects and predict the spoilage spatially as well as spectrally, therefore, to ensure that each acquired sample is co-registered with the rest of the samples, we proposed an image fusion technique for food items, earlier used for remote sensing applications.

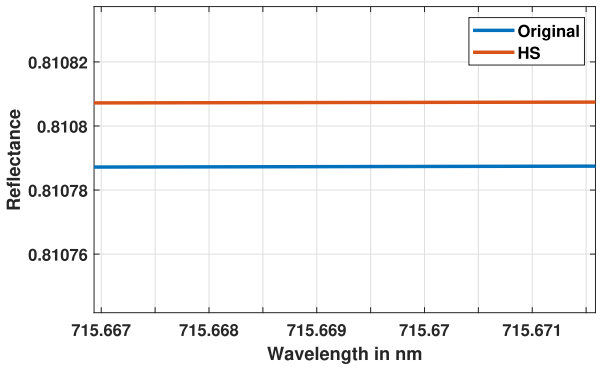
Image fusion techniques have been widely used in remote sensing for various purposes [35], [35], [36]. This work for the very first time explores the use of Hyper Sharpening (HS) [37] method to enrich the spatial resolution of cake samples which varies in size over time. HS enhances the low-resolution image by multiplying its up-scaled version with a ratio between the high-resolution image and its down-scale filtered version. The high-resolution image is obtained by selecting the bands with correlation analysis. Fig. 4a shows the original and enhanced spectra of the sample. It seems like both spectra are overlapped while it is due to minimal difference. On the other hand, Fig. 4b presents the zoomed version of the difference between both original and enhanced spectra of the sample. HS has enhanced the sample image spatially while it has also preserved its spectral resolution with very little error.

E. MICROBIAL SPOILAGE DETECTION

Cakes are usually considered as spoiled if the physical, chemical, or microbiological change occurs [38], which makes these cakes undesirable and unacceptable to consumers [39]. Chemical spoilage occurs when different food items react with each other or some external component is added which alters the food characteristics. A physical change occurs when moistured products are extremely dehydrated or the drier ones



(a) Original and after enhancement spectra.



(b) Difference in original spectra and enhanced spectra

FIGURE 4. Hyper Sharpening (HS).

are moisturized. Microbial spoilage caused by the growth of microorganisms that produce enzymes that leads the food item to become harmful for consumption [40].

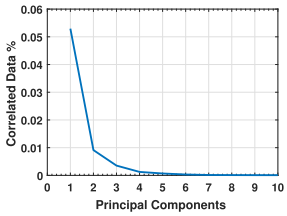
The growth of bacterial organisms can be considered as a change if the microbially exposed sample is compared to the fresh sample. In remote sensing, change detection techniques have been used to monitor the environmental changes and land use [41]. These techniques are useful to measure the natural or man-made phenomena that occurred during an interval of time. Change detection methods can be categorized as supervised or unsupervised depending upon the nature of the data. In supervised methods, ground truths are required to drive a suitable training set whereas, unsupervised methods directly compare the multi-temporal images to compute the desired results without any additional information.

Unsupervised change detection methods primarily consist of the automatic analysis of the change by constructing the differenced data of temporal images. The differenced image is constructed by subtracting (pixel by pixel) two images taken at different times. The computed differenced image is such that it contains higher significantly distinctive values where change occurs and lower to none values where no change occurs. That differenced data is further analyzed to obtain the exact changes that occurred. In this work, the difference images are used to derive a reduced subspace having maximum correlated data by applying PCA and then cluster it into change and no change pixels by using k-means clustering [33].

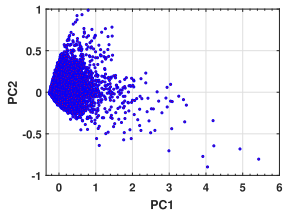
The differenced image contains the data scattered over 224 bands and to process this information effectively, data size must need to be reduced while preserving the maximum information. Therefore, a linear transformation based method i.e., PCA is deployed to reduce the high dimensionality [42]. PCA first calculate the mean of every direction and computes the covariance matrix of the whole dataset as expressed in Equation 2:

$$cov(x, y) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \tag{2}$$

If x and y are the same then the covariance will be the same while it is zero if x and y are uncorrelated [43]. The positive covariance shows that x and y both are large while the negative means that x is large while y is small. For 224 dimensions of a hypercube, the covariance matrix comprises of the size of 224×224. Normally, the first Principal Component (PC) expresses the maximum information as shown in Fig. 5a furthermore, Fig. 5b shows the scatter plot of the first two PCs.



(a) Plot of first 10 PCs



(b) Scatter Plot of PC1 and PC2.

FIGURE 5. Principal component analysis.

The chosen PC contains the maximum information from the differenced image which has both healthy (unchanged) and unhealthy (changed, microbially spoiled) samples data. k-means clustering has been used to differentiate both healthy and unhealthy samples. K-means partitions the data points into a specified k distinct non-overlapping groups (clusters). Each data point belongs to only one cluster. The less variation it has between its clusters, the more homogeneous the data points are.

Clusters are used to annotate the healthy and unhealthy (spoiled) samples for classification [21], [44], [45]. Consequently, the changed pixels have been labeled as spoiled while the unchanged ones as healthy. To classify a sample into a healthy or microbially spoiled, SVM classifier has been used. SVM plots the data in n-dimensions and draws a decision boundary, also known as hyper-plane, to separate classes. The n-dimensions are determined by the number of classes or features exits in data. In this study, there are two classes i.e, spoiled, and healthy. There are multiple numbers of possible boundaries to differentiate the data therefore, SVM finds the

best optimal hyperplane that can maximally divide the data in n-dimensions.

III. RESULTS AND DISCUSSION

Cakes tend to vary in size with the passage of days [46] and the samples acquired for this study changed size in a very minute range with every passing day due to the variations in moisture content. As the spoilage was detected by analyzing the difference image of temporal samples, the acquired data was needed to be spatially co-registered and for that, the HS image fusion technique was used to enhance the spatial information of the sample. In Table 2 the acquired size of cake samples over the span of 8 days, with a difference of almost 12 hours and their spatially enhanced size along with the error is shown. For instance, sample 3 was of the size $305 \times 142 \times 224$ and as sample 1 was a base image, HS was used to enhance it by $309 \times 141 \times 224$. The image at 688.22 *nm* wavelength (randomly chosen) is shown in Fig. 6 displaying the before and after the result of HS and it can be seen that no spatial information has been disturbed.

TABLE 2. HS: Image spatial information enhancement w.r.t. their error rates.

Sample	Sample Size before HS	HS input	Sample Size after HS	Error
Sample1	309x141x224 double	..	..	..
Sample2	307x144x224 double	Sample1,Sample2	309x141x224 double	3.6118e-05
Sample3	305x142x224 double	Sample1,Sample3	309x141x224 double	4.2467e-05
Sample4	304x144x224 double	Sample1,Sample4	309x141x224 double	7.5428e-05
Sample5	306x136x224 double	Sample1,Sample5	309x141x224 double	3.9376e-05
Sample6	306x138x224 double	Sample1,Sample6	309x141x224 double	5.1471e-05
Sample7	309x139x224 double	Sample1,Sample7	309x141x224 double	3.4245e-05
Sample8	312x137x224 double	Sample1,Sample8	309x141x224 double	1.0632e-04
Sample9	312x138x224 double	Sample1,Sample9	309x141x224 double	2.4311e-05
Sample10	311x142x224 double	Sample1,Sample10	309x141x224 double	3.5268e-05
Sample11	313x143x224 double	Sample1,Sample11	309x141x224 double	7.8909e-05
Sample12	313x141x224 double	Sample1,Sample12	309x141x224 double	5.3220e-05
Sample13	312x145x224 double	Sample1,Sample13	309x141x224 double	1.3875e-04
Sample14	312x144x224 double	Sample1,Sample14	309x141x224 double	2.2421e-04
Sample15	312x146x224 double	Sample1,Sample15	309x141x224 double	7.3807e-04
Sample16	313x141x224 double	Sample1,Sample16	309x141x224 double	2.9371e-04

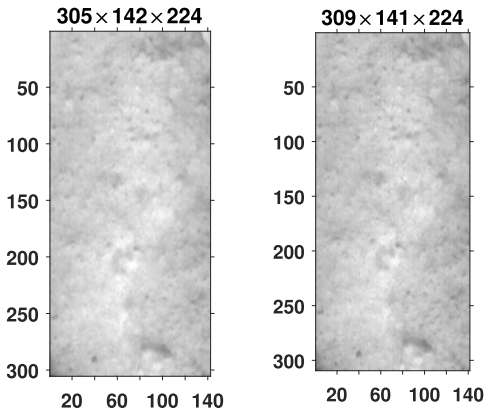


FIGURE 6. HS result of sample 11 at 688.22 *nm* before (left) and after (right) enhancement.

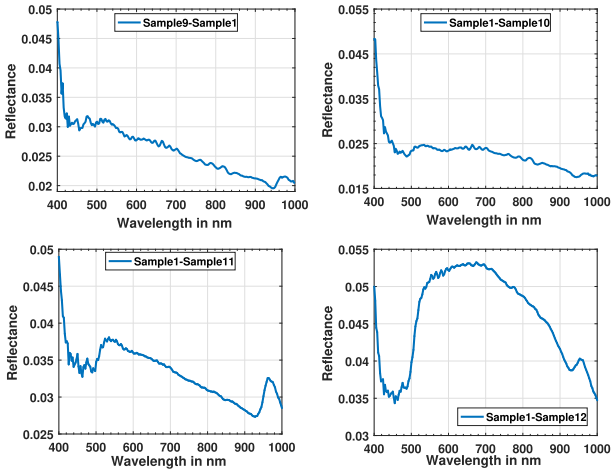


FIGURE 7. Difference images mean spectra (a) Sample 9 - 1 (b) Sample 10 - 1 (c) Sample 11 - 1 (d) Sample 12 - 1.

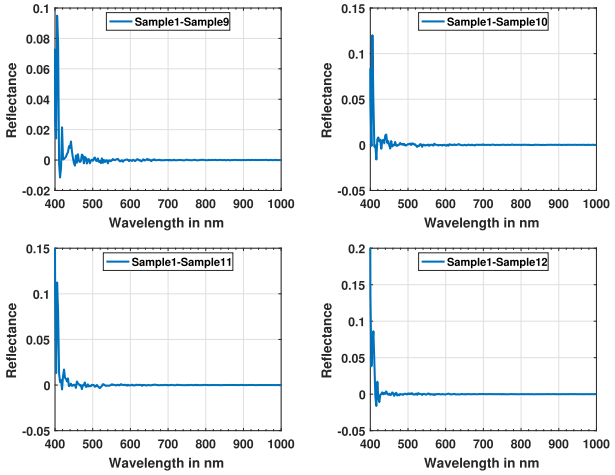
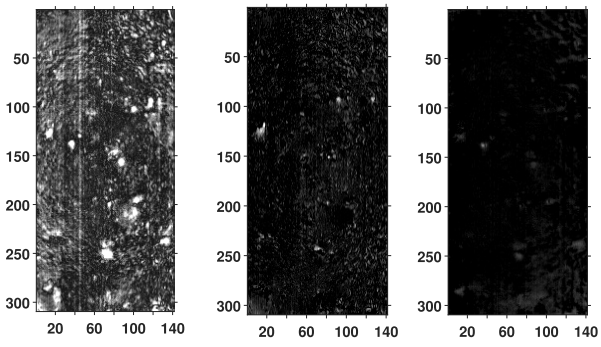


FIGURE 8. Mean spectra of PCA (a) Sample 1 - 9 (b) Sample 1 - 10 (c) Sample 1 - 11 (d) Sample 1 - 12.

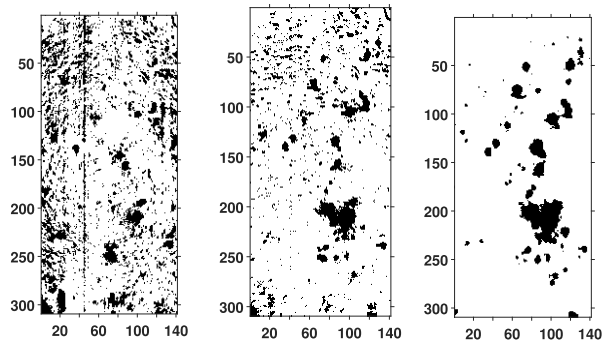


(a) PC 1 (b) PC 2 (c) PC 3

FIGURE 9. First three PC's of difference image (Sample 1 - Sample 10).

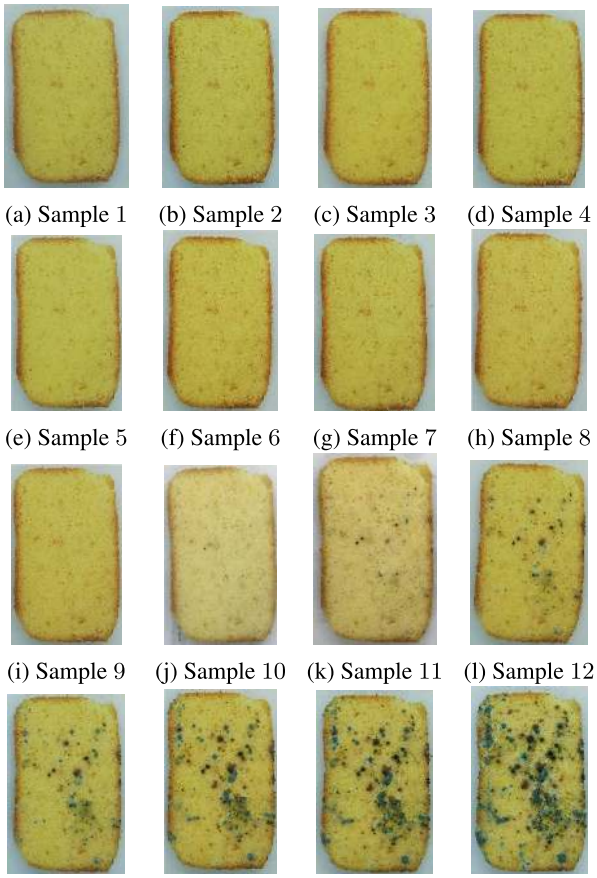
A. SPOILAGE DETECTION AND MODEL TRAINING

The spoiled data has been detected in acquired samples by analyzing the differenced image which were calculated by subtracting every sample (Sample 2 to 16) from sample 1 to obtain the pixels that have got changed over time. Sample 1 is considered as a fresh and healthy piece of the sponge cake. Therefore, subtracting samples (taken on multiple days) from Sample 1 provides information about the changed pixels. The difference of samples acquired in the first four days (Sample 2 to Sample 8) didn't contain any notable spoilage change



(a) Sample 10 and 1 (b) Sample 11 and 1 (c) Sample 12 and 1

FIGURE 10. K-means clustering results for three Sample 10, 11 and 12.



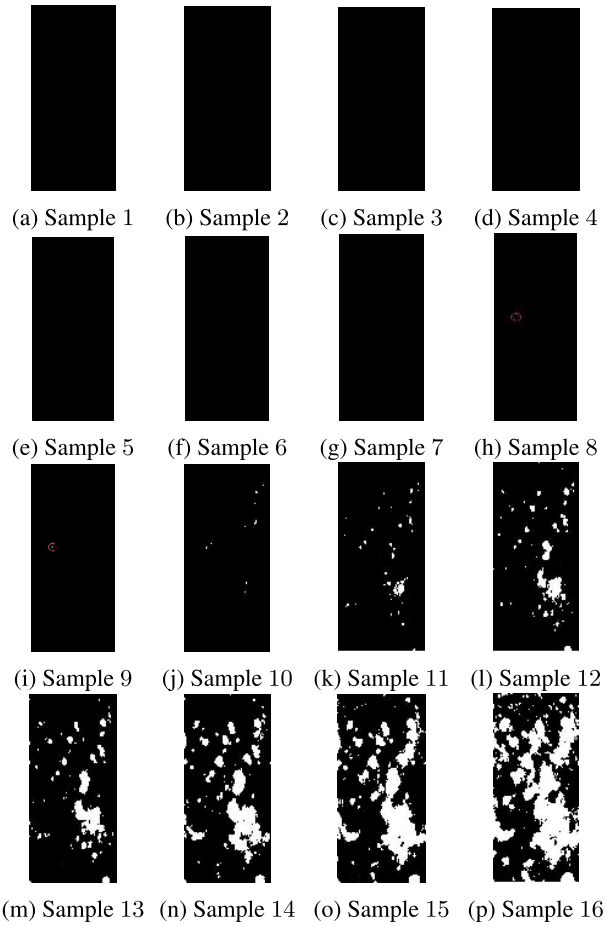
(a) Sample 1 (b) Sample 2 (c) Sample 3 (d) Sample 4

(e) Sample 5 (f) Sample 6 (g) Sample 7 (h) Sample 8

(i) Sample 9 (j) Sample 10 (k) Sample 11 (l) Sample 12

(m) Sample 13 (n) Sample 14 (o) Sample 15 (p) Sample 16

FIGURE 11. True color images of sample 1 to 16. From these images, one can not identify the samples 8 and 9 are spoiled or healthy. Sample 10 to on-wards, the spoilage is visible through naked eyes.



(a) Sample 1 (b) Sample 2 (c) Sample 3 (d) Sample 4

(e) Sample 5 (f) Sample 6 (g) Sample 7 (h) Sample 8

(i) Sample 9 (j) Sample 10 (k) Sample 11 (l) Sample 12

(m) Sample 13 (n) Sample 14 (o) Sample 15 (p) Sample 16

FIGURE 12. Prediction results as compared to the true color images shown in Figure 11 for sample 1 to 16. Spoilage can easily be visualized in Sample 8 and 9 which was not visible by naked eyes though it is visible through HSI.

information except the noise and some constraints of data acquisition. From the fifth day, the difference image started having the required information, that is why Sample 9 to 12 has been used to detect the spoiled information. From sample 13 the spoilage became notable with naked eyes. This finding is also supported by the agreements reported by Latif *et al.* [47] who noticed the spoilage of bread samples after four days of storage under similar conditions. Therefore, the samples from 13 to 16 were not used due to the complete spoilage of the cake. Fig. 7 shows the mean spectra of the differenced image of samples 9 to 12.

After obtaining the differenced images of temporal samples, the data contains the information about spoilage and scattered over the range of 224 bands along with some noise factors as shown in Fig. 7. To reduce the dimensions of related data, PCA has been used. PCA represents the 224 data members as a linear combination and in less number of eigenvectors where each Eigenvector contains the correlated information. Fig. 8 shows the mean spectra of differenced images while Fig. 9 shows spatially the first three PCs. From these results, one can observe that the PC 1 contains the maximum information. Therefore, for this study PC 1 has been used for every sample in further steps. Though PC 1 contains the maximum information of the differenced image, but still, its information needs to be grouped in healthy and spoiled data.

To divide the data into healthy and spoiled segments, K-means has been applied on PC 1. K-means clustered

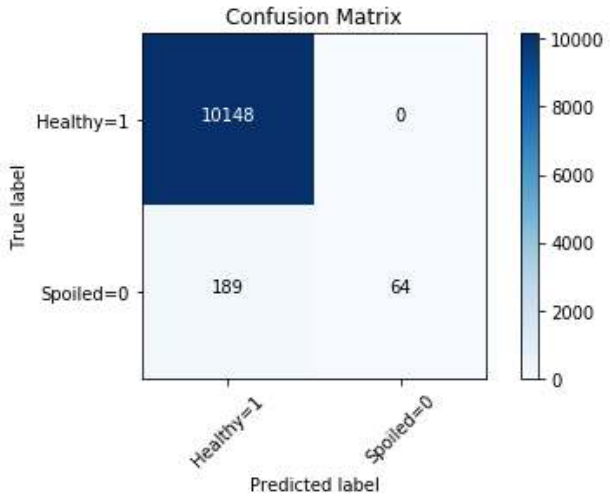


FIGURE 13. Confusion matrix of test data.

the data into two groups which helped in segmenting the hyper-cube pixels in healthy and spoiled. Fig. 10 shows the result of applying k-mean ($k = 2$) on PC 1 of Sample 10 to Sample 12. Sample 1 and Sample 2 (1st-day cake samples) were considered as healthy samples.

After detecting the spoiled pixels, an SVM classifier has been trained. To validate the trained model to distinguish between healthy and spoiled samples, new samples are acquired with the same ingredients and preserved in the same environment. Similarly, new samples have been used to test whether the samples are exposed microbially to spoilage or not without considering the prior knowledge of class information i.e., spoiled or healthy.

To validate the results, predicted spatial locations is shown in the form of predicted maps. After the model has been trained and validated for the spoilage, the final prediction has been made for all the samples from Sample 1 to Sample 16. Fig. 11 and Fig. 12 shows the prediction results and corresponding RGB images of the same samples. It can be seen visibly the spoilage has started from Sample 10, while in prediction results the spoilage has been predicted from Sample 8 almost a day ahead it appeared to naked eyes. This finding is also supported by the agreements reported by Latif *et al.* [47] who noticed the spoilage of bread samples after four days of storage under similar conditions. Furthermore, the proposed model has achieved 98.13% on test data and its confusion matrix is shown in Fig. 13.

Moreover, to strengthen our proposed methodology, we have further performed the external validation of our trained model on a new dataset. The dataset was prepared, acquired and pre-processed in the same way discussed in Section II. After pre-processing, we performed the prediction of our trained model on new dataset and the results are shown in Fig. 14. It is clear from the results that our model is predicting the spoilage almost a day before it starts appearing visually.

Furthermore, in this study, we have deployed some alternative algorithms for image enhancement and spoilage detection but found some limitations. Qu *et al.* [48] proposed an unsupervised encoder-decoder architecture that fuses low spatial HSI with high spatial MSI to generate a high spatial HSI. In this work, the authors have not provided information regarding a transformation matrix that described the relationship between HSI and MSI bases thus it cannot be applied to our dataset. Moreover, we have implemented five anomaly detection algorithms Global Reed-Xiaoli Detector (GRXD) [49], Local Reed-Xiaoli Detector (LRXD) [50], Dual Window-based Eigen Separation Transform (DWEST) [50], Multiple Window Nested Spatial Window-based Target Detection (MW-NSWTD) [50] and Abundance and Dictionary-based Low-Rank decomposition (ADLR) [51]. The Anomaly detection is a single image approach which can be deployed in a real time application as it identifies the data points that deviate from a dataset's normal behavior. However, in our case, those data points can be air holes, foreign particles, the hardness of the surface, cake border, etc rather just spoilage. Moreover, once the model has been trained, from the proposed methodology of this study, it will be able to detect the spoilage in any single image sample baked and resided in a moderate environment as shown in Fig. 14.

Our study proposed a methodology that the difference of temporal images analysed by PCA and K-means can detect

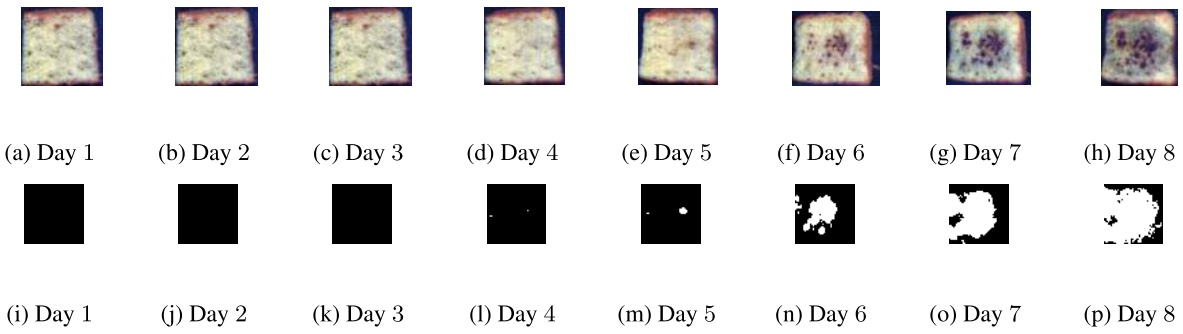


FIGURE 14. True color images and prediction results of external validation dataset from Day 1 to 8. Spoilage was visually appeared from Day 5 while the model predicted it on Day 4.

the spoilage in bakery items. In our scenario for spoilage detection, we are comparing same sample over the passage of days by taking consecutive differences of healthy and spoiled samples. Thus, any substances (like ingredients, air holes etc) except of the spoilage were differenced out in the first step. Therefore, our methodology has only detected and predicted the spoilage. Moreover, the anomaly detection methods were evaluating only a single image at a time for analyzing the anomalies. Thus in a single sample, the methods may considers the substances like presence of any abnormal ingredient, air hole or hard crust, etc as an anomaly thus a spoilage.

IV. CONCLUSION

This study aimed to develop a method that will predict the microbial spoilage in cake sponges. The samples change size over time due to changes in moisture content. Therefore, to co-register, the HS technique has been used. PCA followed by K-means has been used on HSI differenced images to automatically detect the microbial spoilage. Furthermore, multi-class SVM has been trained and tested on spoiled and healthy pixels for prediction. Total of 16 HSI data samples have been acquired over the span of 8 days with a 12 hours difference in each sample. The results revealed that the proposed method not only automatically detected the spoilage but also predicted it a day, with 98.13% accuracy on test data, before it starts appearing visibly notable. The proposed model has further been validated through an external dataset.

DATASET AND CODE AVAILABILITY

The experimental dataset and running demo will be provided upon reasonable request.

REFERENCES

[1] V. Giannou, D. Lebesi, and C. Tzia, *Packaging and Shelf-Life Prediction of Bakery Products*. Hoboken, NJ, USA: Wiley, 2014, ch. 20, pp. 355–371. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118792001.ch20

[2] S. Hoffmann, B. Devleesschauwer, W. Aspinall, R. Cooke, T. Corrigan, A. Havelaar, F. Angulo, H. Gibb, M. Kirk, R. Lake, N. Speybroeck, P. Torgerson, and T. Hald, "Attribution of global foodborne disease to specific foods: Findings from a world health organization structured expert elicitation," *PLoS ONE*, vol. 12, no. 9, Sep. 2017, Art. no. e0183641.

[3] C. Axel, E. Zannini, and E. K. Arendt, "Mold spoilage of bread and its biopreservation: A review of current strategies for bread shelf life extension," *Crit. Rev. Food Sci. Nutrition*, vol. 57, no. 16, pp. 3528–3542, Nov. 2017.

[4] V. Melini and F. Melini, "Strategies to extend bread and GF bread shelf-life: From sourdough to antimicrobial active packaging and nanotechnology," *Fermentation*, vol. 4, no. 1, p. 9, Feb. 2018.

[5] M. M. Gutierrez, M. Meleddu, and A. Piga, "Food losses, shelf life extension and environmental impact of a packaged cheesecake: A life cycle assessment," *Food Res. Int.*, vol. 91, pp. 124–132, Jan. 2017.

[6] *Who Estimates of the Global Burden of Foodborne Diseases: Foodborne Disease Burden Epidemiology Reference Group 2007-2015*, World Health Organization, Geneva, Switzerland, 2015.

[7] A. Majumdar, N. Pradhan, J. Sadasivan, A. Acharya, N. Ojha, S. Babu, and S. Bose, "Food degradation and foodborne diseases: A microbial approach," in *Microbial Contamination and Food Degradation*. Amsterdam, The Netherlands: Elsevier, 2018, pp. 109–148.

[8] D. Vakhare, A. Gupta, and S. R. Chaphalkar, "Confocal microscopy and its implementation in different biological aspects," *Emergent Life Sci. Res.*, vol. 2, pp. 8–12, Dec. 2016.

[9] K. R. Matthews, K. E. Kniel, and T. J. Montville, *Food Microbiology: An Introduction*. Hoboken, NJ, USA: Wiley, 2019.

[10] Y. Zhang, L. Zhu, Y. Zhang, P. He, and Q. Wang, "Simultaneous detection of three foodborne pathogenic bacteria in food samples by microchip capillary electrophoresis in combination with polymerase chain reaction," *J. Chromatography A*, vol. 1555, pp. 100–105, Jun. 2018.

[11] I. Santana Oliveira, A. G. da Silva Junior, C. A. S. de Andrade, and M. D. L. Oliveira, "Biosensors for early detection of fungi spoilage and toxigenic and mycotoxins in food," *Current Opinion Food Sci.*, vol. 29, pp. 64–79, Oct. 2019.

[12] M. Gallo, L. Ferrara, and D. Naviglio, "Application of ultrasound in food science and technology: A perspective," *Foods*, vol. 7, no. 10, p. 164, Oct. 2018.

[13] K. Sendin, P. J. Williams, and M. Manley, "Near infrared hyperspectral imaging in quality and safety evaluation of cereals," *Crit. Rev. Food Sci. Nutrition*, vol. 58, no. 4, pp. 575–590, Mar. 2018.

[14] V. Velusamy, K. Arshak, O. Korostynska, K. Oliwa, and C. Adley, "An overview of foodborne pathogen detection: In the perspective of biosensors," *Biotechnol. Adv.*, vol. 28, no. 2, pp. 232–254, Mar. 2010.

[15] X. Zhang, J. Zhou, C. Zhang, D. Zhang, and X. Su, "Rapid detection of enterobacter cloacae by immunomagnetic separation and a colloidal gold-based immunochromatographic assay," *RSC Adv.*, vol. 6, no. 2, pp. 1279–1287, 2016.

[16] W. M. Ahmed, B. Bayraktar, A. K. Bhunia, E. D. Hirleman, J. P. Robinson, and B. Rajwa, "Classification of bacterial contamination using image processing and distributed computing," *IEEE J. Biomed. Health Informat.*, vol. 17, no. 1, pp. 232–239, Jan. 2013.

[17] Y. Liu, H. Pu, and D.-W. Sun, "Hyperspectral imaging technique for evaluating food quality and safety during various processes: A review of recent applications," *Trends Food Sci. Technol.*, vol. 69, pp. 25–35, Nov. 2017.

[18] L. Ravikanth, D. S. Jayas, N. D. G. White, P. G. Fields, and D.-W. Sun, "Extraction of spectral information from hyperspectral data and application of hyperspectral imaging for food and agricultural products," *Food Bioprocess Technol.*, vol. 10, no. 1, pp. 1–33, Jan. 2017.

[19] R. Vejarano, R. Siche, and W. Tesfaye, "Evaluation of biological contaminants in foods by hyperspectral imaging: A review," *Int. J. Food Properties*, vol. 20, no. 2, pp. 1264–1297, 2017.

[20] M. Kamruzzaman and D.-W. Sun, "Introduction to hyperspectral imaging technology," in *Computer Vision Technology for Food Quality Evaluation*. Amsterdam, The Netherlands: Elsevier, 2016, pp. 111–139.

[21] M. Ahmad, S. Shabbir, D. Oliva, M. Mazzara, and S. Distefano, "Spatial-prior generalized fuzziness extreme learning machine autoencoder-based active learning for hyperspectral image classification," *Optik-Int. J. Light Electron Opt.*, vol. 206, Mar. 2020, Art. no. 163712. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0030402619316109

[22] M. Ahmad, M. Mazzara, R. A. Raza, S. Distefano, M. Asif, M. S. Sarfraz, A. M. Khan, and A. Sohaib, "Multiclass non-randomized Spectral–Spatial active learning for hyperspectral image classification," *Appl. Sci.*, vol. 10, no. 14, p. 4739, Jul. 2020.

[23] H. Aasen, E. Honkavaara, A. Lucieer, and P. Zarco-Tejada, "Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: A review of sensor technology, measurement procedures, and data correction workflows," *Remote Sens.*, vol. 10, no. 7, p. 1091, Jul. 2018.

[24] M. Ahmad, S. Protasov, A. M. Khan, R. Hussain, A. M. Khattak, and W. A. Khan, "Fuzziness-based active learning framework to enhance hyperspectral image classification performance for discriminative and generative classifiers," *PLoS ONE*, vol. 13, no. 1, Jan. 2018, Art. no. e0188996.

[25] H. Pu, L. Lin, and D. Sun, "Principles of hyperspectral microscope imaging techniques and their applications in food quality and safety detection: A review," *Comprehensive Rev. Food Sci. Food Saf.*, vol. 18, no. 4, pp. 853–866, Jul. 2019.

[26] M. Hussain Khan, Z. Saleem, M. Ahmad, A. Sohaib, and H. Ayaz, "Unsupervised adulterated red-chili pepper content transformation for hyperspectral classification," 2019, *arXiv:1911.03711*. [Online]. Available: http://arxiv.org/abs/1911.03711

[27] A. Majda, R. Wietecha-Posluszny, A. Mendys, A. Wójtowicz, and B. Lydzba-Kopczynska, "Hyperspectral imaging and multivariate analysis in the dried blood spots investigations," *Appl. Phys. A, Solids Surf.*, vol. 124, no. 4, p. 312, Apr. 2018.

[28] M. Halicek, G. Lu, J. V. Little, X. Wang, M. Patel, C. C. Griffith, M. W. El-Deiry, A. Y. Chen, and B. Fei, "Deep convolutional neural networks for classifying head and neck cancer using hyperspectral imaging," *J. Biomed. Opt.*, vol. 22, no. 6, Jun. 2017, Art. no. 060503.

[29] M. Shimoni, R. Haelterman, and C. Perneel, "Hypersectral imaging for military and security applications: Combining myriad processing and sensing techniques," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 101–117, Jun. 2019.
[30] M. S. Andresen, B. S. Dissing, and H. Løje, "Quality assessment of butter cookies applying multispectral imaging," *Food Sci. Nutrition*, vol. 1, no. 4, pp. 315–323, Jul. 2013.
[31] A. Polak, F. K. Coutts, P. Murray, and S. Marshall, "Use of hyperspectral imaging for cake moisture and hardness prediction," *IET Image Process.*, vol. 13, no. 7, pp. 1152–1160, May 2019.
[32] N. El-Morsy, S. Mokhtar, and K. Youssef, "Accurate quantification of fungal growth in bread by using spectral analysis," *J. Food Dairy Sci.*, vol. 5, no. 1, pp. 33–44, Jan. 2014.
[33] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and k-means clustering," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 772–776, Oct. 2009.
[34] A. Duralliu, P. Matejtschuk, and D. R. Williams, "Measuring the specific surface area (SSA) of freeze-dried biologics using inverse gas chromatography," *Eur. J. Pharmaceutics Biopharmaceutics*, vol. 142, pp. 216–221, Sep. 2019.
[35] R. Gharbia, A. E. Hassanien, A. H. El-Baz, M. Elhoseny, and M. Gunasekaran, "Multi-spectral and panchromatic image fusion approach using stationary wavelet transform and swarm flower pollination optimization for remote sensing applications," *Future Gener. Comput. Syst.*, vol. 88, pp. 501–511, Nov. 2018.
[36] J. G. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3461–3472, Jan. 2000.
[37] M. Selva, B. Aiazzi, F. Butera, L. Chiarantini, and S. Baronti, "Hyper-sharpening: A first approach on SIM-GA data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 3008–3024, Jun. 2015.
[38] E. Rico-Munoz, R. A. Samson, and J. Houbraken, "Mould spoilage of foods and beverages: Using the right methodology," *Food Microbiol.*, vol. 81, pp. 51–62, Aug. 2019.
[39] M. E. Lopez, M. T. Gontijo, D. M. Boggione, L. A. Albino, L. S. Batalha, and R. C. Mendonça, "Microbiological contamination in foods and beverages: Consequences and alternatives in the era of microbial resistance," in *Microbial Contamination and Food Degradation*. Amsterdam, The Netherlands: Elsevier, 2018, pp. 49–84.
[40] L. Petruzzi, M. R. Corbo, M. Sinigaglia, and A. Bevilacqua, "Microbial spoilage of foods: Fundamentals," in *The Microbiological Quality of Food*. Amsterdam, The Netherlands: Elsevier, 2017, pp. 1–21.
[41] A. Asokan and J. Anitha, "Change detection techniques for remote sensing applications: A survey," *Earth Sci. Informat.*, vol. 12, no. 2, pp. 143–160, Jun. 2019.
[42] Y. Aït-Sahalia and D. Xiu, "Principal component analysis of high-frequency data," *J. Amer. Stat. Assoc.*, vol. 114, no. 525, pp. 287–303, Jun. 2019.
[43] J. Lever, M. Krzywinski, and N. Altman, "Points of significance: Principal component analysis," *Nature Methods*, vol. 14, no. 7, pp. 641–642, Jun. 2017.
[44] W. Di and M. M. Crawford, "Critical class oriented active learning for hyperspectral image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2011, pp. 3899–3902.
[45] M. Ahmad, A. Khan, A. Khan, M. Mazzara, S. Distefano, A. Sohaib, and O. Nibouche, "Spatial prior fuzziness pool-based interactive classification of hyperspectral images," *Remote Sens.*, vol. 11, no. 9, p. 1136, May 2019.
[46] M. Mahiuddin, M. I. H. Khan, C. Kumar, M. M. Rahman, and M. A. Karim, "Shrinkage of food materials during drying: Current status and challenges," *Comprehensive Rev. Food Sci. Food Saf.*, vol. 17, no. 5, pp. 1113–1126, Sep. 2018.
[47] A. Latif, T. Masood, and H. A. Khan, "Quality improvement and shelf life extension of bread," *J. Agricult. Social Sci.*, vol. 1, no. 2, pp. 109–113, 2015.
[48] Y. Qu, H. Qi, and C. Kwan, "Unsupervised sparse Dirichlet-net for hyperspectral image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2511–2520.
[49] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1760–1770, 1990.
[50] T. Marshall and L. N. Perkins, "Color outlier detection for search and rescue," Dept. Elect. Comput. Eng., Boston Univ., Boston, MA, USA, Tech. Rep. ECE-2015-01, 2015.
[51] Y. Qu, W. Wang, R. Guo, B. Ayhan, C. Kwan, S. Vance, and H. Qi, "Hyperspectral anomaly detection through spectral unmixing and dictionary-based low-rank decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4391–4405, Aug. 2018.

ZAINAB SALEEM received the bachelor's degree in computer systems engineering from The Islamia University of Bahawalpur, in 2015, and the master's degree in computer engineering with specialization in machine vision from the Khwaja Fareed University of Engineering and Information Technology (KFUEIT), in 2020. She is currently working as a Lab Engineer with KFUEIT. She is also actively working with the Advance Image Processing Research Laboratory (AIPRL), Research Group, a first Hyperspectral Imaging Laboratory, Pakistan. Her research interests include machine vision, machine learning, and hyperspectral imaging.

MUHAMMAD HUSSAIN KHAN received the bachelor's degree in electronics engineering from the Ghulam Ishaq Khan Institute of Engineering Sciences and Technology and the master's degree in computer engineering with specialization in machine vision from the Khwaja Fareed University of Engineering and Information Technology (KFUEIT), in 2020. He is currently working as a Lab Engineer with KFUEIT. He is also actively working with the Advance Image Processing Research Laboratory (AIPRL), Research Group, a first Hyperspectral Imaging Laboratory, Pakistan. His research interests include machine vision, machine learning, and hyperspectral imaging.

MUHAMMAD AHMAD received the B.S. degree from G.C University, the M.S. degree from International Islamic University, and the Ph.D. degree from Innopolis University, Russia. He is currently working as an Assistant Professor with the Department of Computer Science, National University of Computer and Emerging Sciences, Islamabad, Chiniot-Faisalabad Campus. He is associated with the Advance Image Processing Research Laboratory (AIPRL), Research Group, a first Hyperspectral Imaging Laboratory, Pakistan. He is also associated with the University of Messina, Messina, Italy. He has authored a number of research papers in reputed journals and conferences. His current research interests include machine learning, computer vision, remote sensing, hyperspectral imaging, and wearable computing. He is a Regular Reviewer for a number of top tier journals and conferences.

AHMED SOHAIB received the B.S. degree in computer system engineering from the Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, the degree from the Interactive Multimedia Laboratory, Department of Electrical and Computer Engineering, National University of Singapore (NUS), and the Ph.D. degree from Australian National University (ANU) and National ICT Australia (NICTA). He is currently an Assistant Professor with the Department of Computer Engineering, Khawaja Fareed University of Engineering and Information Technology, where he is also working as the Head. He is leading a Research Group of the Advance Image Processing Research Laboratory (AIPRL), a first Hyperspectral Imaging Laboratory, Pakistan. His major research interests include spectral imaging systems and their exposure control. He received the Gold Medal from the Ghulam Ishaq Khan Institute of Engineering Sciences and Technology.

HAMAIL AYAZ received the Bachelor of Computer Science degree from Comsats Islamabad and the master's degree in computer engineering with specialization in machine vision from the Khwaja Fareed University of Engineering and Information Technology (KFUEIT), in 2020. He is currently working as a Teaching Assistant with KFUEIT. He is also actively working with the Advance Image Processing Research Laboratory (AIPRL), Research Group, a first Hyperspectral Imaging Laboratory, Pakistan. His research interests include machine vision, machine learning, and hyperspectral imaging.

MANUEL MAZZARA received the Ph.D. degree in computing science from the University of Bologna. He is currently a Professor of computer science with the Institute of Software Development and Engineering, Innopolis University, Russia. He cooperated with European and U.S. industries, plus governmental and intergovernmental organizations, such as the United Nations, always at the edge between science and software production. The work conducted by Manuel and his team in recent years focuses on the development of theories, methods, tools and programs covering the two major aspects of software engineering: the process side, describing how we develop software, and the product side, describing the results of this process. His research interests include software engineering, service-oriented architectures and programming, concurrency theory, formal methods, and software verification.