Food tray sealing fault detection using hyperspectral imaging and PCANet

Mohamed Benouis, Leandro D. Medus, Mohamed Saban, Grzegorz Łabiak, Alfredo Rosado-Muñoz

Angaben zur Veröffentlichung / Publication details:

Benouis, Mohamed, Leandro D. Medus, Mohamed Saban, Grzegorz Łabiak, and Alfredo Rosado-Muñoz. 2020. "Food tray sealing fault detection using hyperspectral imaging and PCANet." *IFAC-PapersOnLine* 53 (2): 7845–50. https://doi.org/10.1016/j.ifacol.2020.12.1955.

Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 53-2 (2020) 7845–7850

Food tray sealing fault detection using hyperspectral imaging and PCANet

Mohamed Benouis * **Leandro D. Medus** ** **Mohamed Saban** **
Grzegorz Łabiak, *** **Alfredo Rosado-Muñoz** **

* *University Of M'sila, M'sila, BP 28000 ALGERIA. (e-mail: mohamed.benouis@univ-msila.dz).*
** *University of Valencia, GDDP, Dpt. Electronic Engineering, ETSE School of Engineering. 46100 Burjassot. SPAIN. (e-mail: leandro.d.medus@uv.es).*
*** *University of Zielona Góra, Institute of Electrical Engineering, ul. Podgórna 50, 65-246 Zielona Góra, Poland. (e-mail: G.Labiak@iee.uz.zgora.pl).*

Abstract: Food trays are very common in shops and supermarkets. Fresh food packaged in trays must be correctly sealed to protect the internal atmosphere and avoid contamination or deterioration. Due to the speed of production, it is not possible to have human quality inspection. Thus, automatic fault detection is a must to reach high production volume. This work describes a deep neural network based on Principal Component Analysis Network (PCANet) for food tray sealing fault detection. The input data come from hyperspectral cameras, showing more characteristics than regular industrial cameras or the human eye as they capture the spectral properties for each pixel. The proposed classification algorithm is divided into three main parts. In the first part, a single image is extracted from the hypercube by using pixel-level fusion method: the cube hyperspectral images are transformed into two-dimensional images to use as the input to the PCANet. Second, a PCANet structure is applied to the fused image. The PCANet has two filter bank layers and one binarization layer (three stages), obtaining a feature vector. Finally, a classification algorithm is used, having the feature vector as input data. The SVM and KNN classifiers were used. The database used in this work is provided by food industry professionals, containing eleven types of contamination in the seal area of the food tray and using metallic opaque cover film. Obtained results show that the design of our framework proposed achieves accuracy of 90% (87% F-measure) and 89% (89% F-measure) for SVM and KNN, respectively. Computation time for classification shows that a food tray speed of 65 trays per second could be reached. As a final result, the influence of the dataset size is analyzed, having PCANet a similar behavior for an extended and a reduced dataset.

Keywords: Image processing, PCANet Classifier, Neural networks, Food Industry, Food packaging, quality inspection.

1. INTRODUCTION

Food quality is a very important task in the food industry which includes food supply chain, food contamination, food analysis, and health care [Ryser and Marth, 1999]. Thus, food safety is essential as it might lead to a public health problem. In particular, food packaging is crucial in food manufacturing. Packaging can be affected by different anomalies, either in the production factory or food supply chain. For example, according to the PPP (Waste and Resources Action Programme) in UK, it is estimated that up to 480,000 tons of food is wasted each year due to poor seals in packaging [WRAP, 2000]. Consequently, proper packaging and seals are mandatory to achieve the expected shelf life for many food products. Many techniques have been well established in order to meet both present legislation and consumers' expectations and demands. Traditional inspection for food tray sealing fault detection as human visual inspection, analytical, physical, biological and chemical analyses are time-consuming, destructive, and sometimes environmentally unfriendly [Sun, 2010]. To overcome this, image sensing and spectral analysis techniques have been growing in the food safety field and gained more attention due to the obtained achievements and encouraging results [Li et al., 2017]. In particular, spectral imaging is able to obtain both spatial and spectral information from the object, which is helpful to locate small anomalies, it is fast, and non-destructive. In addition no human intervention is required. Contrary to the conventional image acquisition systems used food quality imaging, hyperspectral imaging combines infrared spectroscopy with machine vision to produce images, having more information about the chemical composition of the objects present in the acquired image. Its ability to identify differences in the chemical composition of organic materials opens up major new possibilities for detecting contamination in food products. Additionally, computing technology allows hyperspectral images to be entered into

10.1016/j.ifacol.2020.12.1955

a classifier and obtaining the classification result in due time to operate in real time, i.e. faster than the production line. Then, they can be used in high speed production lines for food processing and packaging inspection [Huang et al., 2014].

Data cubes generated by hyperspectral imaging (HSI) sensors contain multiple spectral images, generating a large dimensionality. However, the amount of data (spatial and spectral) involved in HSI often needs a deep processing, space reduction, handcrafted feature selection tools in order to retain the important features and ignore the redundant information [Du and Sun, 2006]. Generally, conventional imaging cannot acquire spectral information (at most, RGB images provide information in three spectral points) and spectroscopy measurement cannot cover large sample areas. In recent years, spectral imaging (i.e., hyperspectral and multispectral) has emerged as a better tool for sensing and is being currently applied to multiple applications [Bioucas-Dias et al., 2013].

On the other hand, machine learning-based methods play an important role in food safety and its applications [Du and Sun, 2006]. Some of these methods use standard image processing techniques to learn a set of features which deliver spatial knowledge into a single spectral point (three for RGB); they provide information about color and texture, shapes, etc. and obtain a set of features from the image which are applied to a machine learning classifier such as KNN, SVM, NN, etc. [Tiger and Verma, 2013]. The main drawback of these methods is that they require hand crafted feature extraction from the image and, if not properly selected, the classifier result will not be accurate and would not provide a good accuracy. Methods as low-rank representation are effective in feature extraction and overcome the aforementioned problems. A well known procedure is Principal Component Analysis (PCA) [Ostyn et al., 2007] that, in combination with neural networks [Lee and Seung, 2001] is a common classification algorithm. However, such technique needs to convert the 3D HSI hypercube into a 2D matrix representation and consequently, a loss of important information contained in image cube arises [Erichson et al., 2018]. To tackle this issue, sparse representation is another effective low rank representation, which became widely adopted for HSI [Huang et al., 2017]. The simplest way of sparsely representing these images is a linear combination of atoms from a dictionary, which is constructed or learned from training samples [Chen et al., 2018].

Recently, deep learning-based algorithms have been introduced to food science and engineering and presented promising performance [Zhou et al., 2019]. The aim of deep learning is to extract higher level features which represent more abstract semantics of the original data. Convolutional Neural networks (CNN) are the first kind of deep architecture-based models. To our knowledge, CNN is considered as a successful method obtaining high accuracy in hyperspectral image classification [Haggag et al., 2019], [Kagaya et al., 2014], [Hassannejad et al., 2016], [Liu et al., 2016]. As PCA, the Auto Encoder (AE) can be used as a method of both feature extraction and dimensionality reduction, making it a common technique in HSI classification [Turan and Erzin, 2018]. However, deep learning methods usually suffer from hyperparameter optimization, being time-consuming for the training process. Many techniques have been used in building the network structure to fit the desired application [Zhou et al., 2019].

The PCANet has shown a great success in image classification [Chan et al., 2015] due to its ability for deep feature extraction. Deep structure and filter convolution are two significant parts in PCANet. Deep structure hierarchically extracts deep features. Filter convolution learns multi-scale spatial structure of images. Unfortunately, training these networks end-to-end with fully learnable convolutional kernels is computationally expensive and prone to over-fitting. Thus, many approaches try to replace the process of learning these convolutional kernels with traditional matrix/tensor factorization methods. For example, PCANet learns convolutional kernels by employing PCA instead of back propagation.

Here, we propose a deep network for food tray seal fault detection using a PCA network (PCANet) [Chan et al., 2015]. Despite PCANet is a simple network, it is used in many computer vision tasks [Soon et al., 2018], [Lee et al., 2018]. Motivated by this, we attempt to propose a new PCANet for hyperspectral image classification. However, the spectral data cannot be directly utilized by PCANet and then, this work also proposes different adaptation methods. To deal with this constraint, we proposed to apply the fusion strategy on the image cube which leads to construct a single image and then fed to the PCANet. Pixel-level fusion methods are proposed as they are computationally efficient and easy to implement [Li et al., 2017]. Three pixel-level fusion methods have been developed: spatial image fusion methods and transform image fusion methods.

Recent studies show that PCANet with relatively shallow architectures (i.e. number of stage and filter PCA) are able to learn challenging tasks involving 2D or 1D signals. Although this method losses some spectral information, the networks with shallow architecture is much easier to train and implement, still showing good classification results. The remainder of this work is organized as follows: In section 2 we present detailed description of the proposed PCANet based Food tray sealing fault detection method. Section 3 describes the hyperspectral images used in this study and Section 4 describes the algorithm details for the specific task. Section 5 reports the experimental results and discusses the obtained values. Conclusions are drawn in Section 6.

2. PCANET CLASSIFIER

PCANet is first proposed to handle the task of classification for single image [Chan et al., 2015]. The PCANet aim is to achieve two goals: first, obtain a simple deep learning network; and second, provide a baseline for single image tasks where simpler and end to end learning can be accomplished. Binary quantization (hashing), histogram features and linear support vector machine (SVM) are also adopted to simplify traditional deep learning methods used in other deep learning algorithms. Generally, PCANet contains three processing components: PCA filters, binary hashing, and histogram features. The PCANet was first proposed in [Chan et al., 2015] by cascading PCA filter banks with mean normalization, binary hashing and the

histogram composition step to learn hierarchical representations. The conventional PCANet presented in the literature is designed to exclusively operate in 2D data. This is why we have to apply fusion schemes (average, minimum maximum strategies) on the image cube to made single image 2D instead of multichannel image. In this way, multiple images from different channels are combined into one single image in pixel-wise manner, after which it is fed into the PCANet system to extract the important features. One of the main advantages of the pixel-level fusion is their low computational complexity and easy implementation [Li et al., 2017]. PCANet algorithm consists of many stages of PCA filters to extract higher level feature vectors.

The filter banks are constructed by projecting principal components (PCA) over a set of image vectors (I) where each vector represents the $k1 \times k1$ patch sliding around each pixel. After that, we subtract the patch mean from each vectorized patch (normalization process). Therefore, for all the input training images $\{I, i = 1, 2, \ldots, N\}$, we can obtain the matrix described in equation (1) where $\widehat{m} = m - \left[\frac{k_1}{2}\right]$ and $\widehat{n} = n - \left[\frac{k_2}{2}\right]$.

$$I = \left[\overline{I}_1, \overline{I}_2, \ldots, I_N\right] \in R^{k_1 k_2} \times N\widehat{m}\widehat{n} \tag{1}$$

Assuming that the number of PCA filter banks in i^{th} layer is $F_i^{S_1}$ (L1 convolution filter), output filtered images can be obtained by convolving the training image with PCA eigenvectors, and mathematically given by equation (2).

$$I_{F_1}^{S_1} = I * F_i^{S_1} \tag{2}$$

In this case, $F_i^{S_1}$ principal eigenvectors are selected as PCA filters for convolutional stage as in equation (3).

$$F_i^{S_1} = mat_{k_1 k_2}\left(q_i\left(II^T\right)\right) \in R^{k_1 k_2}, l = 1, 2 \ldots, L_1 \tag{3}$$

To simplify, we depicted an example of three stages PCANet based feature extraction method (see Fig. 1). The PCA is deployed to minimize the reconstruction error within a family of orthonormal filters (equation 4). Two filter banks stages and one filter bank stage conforms the algorithm.

$$min_{V \in R^{k_1 k_2 * L_1}} \parallel I - VV^T \parallel_F^2, stVV^T = I_{L_1} \tag{4}$$

Where $I_{F_i}^{S_1}$ $(i \in [1 \ldots L_1])$ denote the i^{th} filtered image using the $F_i^{S_1}$ filter for the first stage. In the second stage, the same step process used in the first stage is also performed by convolving the output filter images of the pervious stage with the weight filters $F_j^{S_2}$ generated by PCA as in (5) where $I_{F_i}^{S_1}$ $with$ $i \in [1, \ldots, L_1]$ and $j \in [1, \ldots, L_1]$ denote the i^{th} filtered image using the $F_j^{S_2}$ filter for the second stage. If the number of filters in second stage equals L2, the output size filtered images is $L1 \cdot L2$.

$$I_{F_{ij}}^{S_2} = I_{F_i}^{S_1} \times F_j^{S_2} \tag{5}$$

In the last layer, the filter image outputs $I_{F_{ij}}^{S_2}$ are transformed into binary format by using a Heaviside step function (binary hashing) as shown in equation (6).

$$H_{ij}^B = \begin{cases} 1 & if \quad I_{F_{ij}}^{S_2} \geq 0 \\ 0 & otherwise, \end{cases} \tag{6}$$

where H_{ij}^B is a binary image. After that, around each pixel, the vector of L2 binary bits is viewed as a decimal number (equation 7) where O_i^l is an image whose every pixel is an integer in the range $[0, 2^{L_2 - 1}]$.

$$O_i^l = \sum_{i=1}^{L_2} 2^{i-1} H_{ij}^B \tag{7}$$

Then, the histograms of the obtained block images B of size $h_1 h_2$ are extracted either with overlap ratio R or non-overlapping and then concatenated to form a feature vector which represents the input single image (histogram computation step). Thus, the feature vector of the input image I is then defined as in equation (8) where f_i denotes the block-wise histograms of the O_i^l binary image.

$$f_i = [f_1, f_2, \ldots, f_{L_1}]^T \in R^{2^{L_2} L_1 B} \tag{8}$$

As in other deep learning techniques, PCANet requires hyperparameter tuning for optimal classification. In this case, the PCANet parameters which must be tuned are the number of stages (N), the filter sizes in each stage $(K_1, K_2, \ldots, K_N)$, the number of filters in each stage $(L_1, L_2, \ldots, L_N)$, the block size $(h_1, h_2, \ldots, h_N)$, and the overlap ratio R.

Once trained and tuned, the algorithm works in two main phases:

(1) Data acquisition and fusion. After being acquired by the hyperspectral camera, the complete image hypercube is divided in Regions Of Interest (ROI) [Gowen et al., 2007]. Then, ROI hypercubes are fused into a single image by applying a pixel-level fusion method (see section 4).
(2) PCANet-based feature extraction and classifier operation (SVM and KNN in this case).

3. DATASET

Images were captured with a line-scanning imager (pushbroom type) which is an imaging spectrometer. They were analyzed in detail, and after a filtering process, some of them were discarded.

The complete group of images was composed of 210 hyperspectral images (i.e. 210 food trays), divided in 150 tray images with multiple types of contamination or seal faults (11 classes with 15 images each) and 60 images of standard trays in production correctly sealed (control data). In order to generate the dataset, all impurities were manually added and then, trays were sealed with a metallic cover (the human eye cannot detect any impurity in these cases). Selected impurities were those materials that might be present in the sealing process that, due to some production failure, can contaminate the food tray:

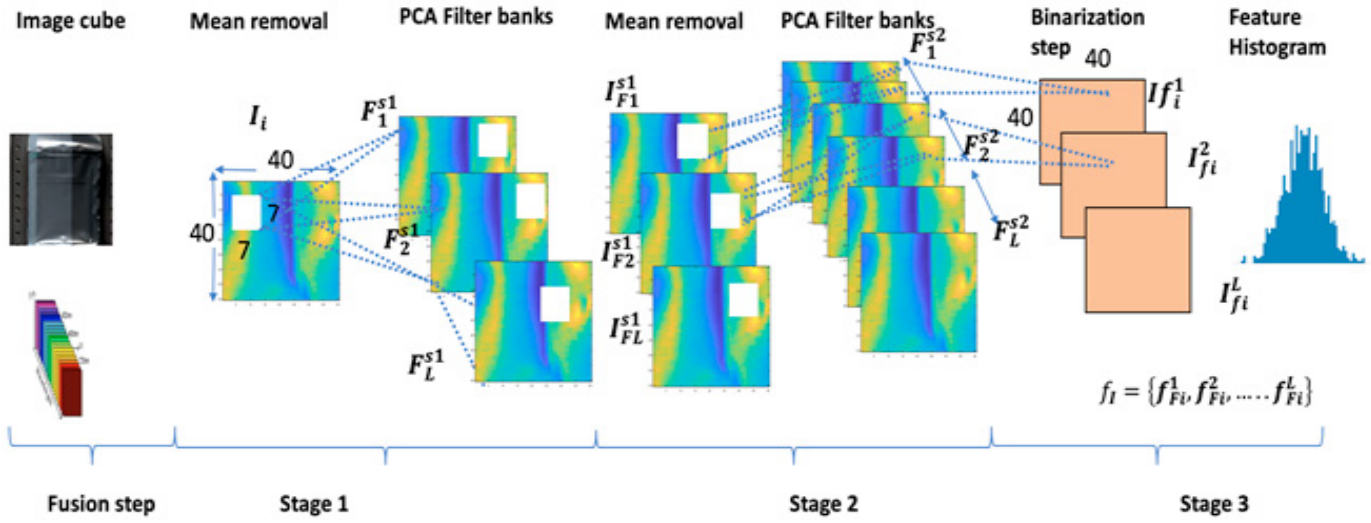

Fig. 1. General diagram of the PCANet classification algorithm. Two filter banks are used, and a binarization stage provides a feature vector which serves as the input to a classifier.

small pieces of plastic or metals, liquids that could be spilled over the edges of the trays, etc. The specific types of contamination (classes) were eleven: washers, sugar, flange, plywood, cork, elastic rubber, wood, paper, silver paper, hair, and polarized plastic. All the images were composed of 410 lines by 320 samples, each one with 168 spectral channels (wavelengths in the range from 891.12 to 1728.45 [nm]). Each data point was coded in a 16 bit signed integer. Thus, each tray hypercube image was 4.204 MiB in size.

In order to generate a dataset from a set of images, a specific labeling software was created, called HypLabTool. The tool was built to be multi-platform, currently running under Linux and Windows. It was developed in Python, using the *PyQt*, *GDAL* and *h5py*packages. The tool is currently available in open source repository at: https://github.com/LeoSf/HypLabTool. A feature to manually define the type of contamination and a Region Of Interest (ROI) around a contaminated area was included in the software. All acquired images were visually inspected and labelled. The ROI is specified by the user indicating the height and width covering the entire area where the fault is located. These dimensions were expressed in pixels. This methodology allows the creation of datasets with complete images or with areas of different dimensions, not following the conventional approach of a pixel by pixel analysis. In this case, a ROI size of 40×40 pixels around the fault, with 130 spectral channels, was chosen, which means that data hypercubes of $40 \times 40 \times 130$ were obtained. After analyzing multiple cases of anomalies, it was observed that the ends of the spectrum did not provide meaningful information. Then, the first and last 19 spectral channels were removed, thus having 130-channel hypercubes with spectral information from 984.2 to 1631.0 [nm]. With their corresponding class designation, data are used for SVM and KNN learning in this case (or any other machine learning process), and classification performance evaluation.

4. PCANET PARAMETER SELECTION

Tuning of the algorithm parameters is required to obtain the best classification ratio. Parameters greatly depends on the application and type of data. A series of experiments was conducted in order to select the best parameters. Table 1 shows the final parameters obtained for this application.

Hyperspectral image fusion was done in order to reduce computation time. The 3D input data cube is reduced to a regular 2D image. This fact also allows to use PCANet 2D image analysis techniques. In order to evaluate how the application of fusion rules affect the algorithm performance, three types of fusion rules were tested:

(1) MEAN fusion rule on the image hypercube and then, PCANet feature extraction and classification.
(2) MAX fusion rule on the image hypercube and then, PCANet feature extraction and classification.
(3) MIN fusion rule on the image hypercube and then, PCANet feature extraction and classification.

Table 1. Optimal parameters for PCANet used in food tray abnormal seal detection.

Parameter	Definition	Values
$m * n$	Size of input image	$40 * 40$
N	Number of stages	2
$K_1 * K_2$	Patch size	$7 * 7$
$L_1 L_2$	Number of two stage filters	$10 * 10$
$h_1 * h_2$	Block size	$5 * 5$
R	Overlap ratio of block R, $(0 < R < 0.9)$	0.6

Classification performance metrics of accuracy (9) and F-measure (11) were obtained. The F-measure can provide a more realistic measure of a test performance by using both precision and recall into a single metric. For equations, T_p=True positive, T_n=True negative, F_p=False positive, P=total number of positive samples, and T=total number of samples. A positive sample (T_p) is considered as a normal tray, and a negative T_n is considered to have an anomaly.

$$accuracy = \frac{T_p + T_n}{T} \tag{9}$$

$$recall = \frac{T_p}{P} \quad ; \quad precision = \frac{T_p}{T_n + F_p} \tag{10}$$

$$F_{measure} = \frac{2 * (precision * recall)}{(precision * recall)} \tag{11}$$

5. RESULTS AND DISCUSSION

After obtaining the feature vector provided by the PCANet, two different classification algorithms were used for comparison purposes. In both cases, a 5-fold cross-validation procedure was used, the dataset was split into ten parts, nine of which were used for training and one for testing. Table 2 shows the classification performance results for the SVM and KNN classifiers using the aforementioned input data fusion methods. As seen in the tables, the best results are obtained in case of MEAN data fusion and both SVM and KNN provide very similar classification values: 90% and 89% in case of MAX data fusion, 87% and 85% for MAX data fusion, and 83% and 82% in case of MIN data fusion, for SVM and KNN, respectively.

Table 2. Performance of the proposed classification technique using SVM and KNN classifiers under three different input data fusion rules

	SVM		KNN	
	Accuracy	F-measure	Accuracy	F-measure
MEAN	0.90	0.87	0.89	0.89
MAX	0.87	0.87	0.85	0.85
MIN	0.83	0.83	0.82	0.82

It is well known that one of the main issues in machine learning and deep learning algorithms is related to the amount of data used for training since effects as overfitting, under-fitting or unbalanced data may arise. In order to verify whether the performance of the proposed PCANet model is affected by this issue, we performed a training of the model with a reduced dataset. Thus, we obtained classification results with the following amount of data:

- 1^{st} Scenario: number of training samples, from 1000 to 1775. The remaining samples were used as test set to verify the performance of the model.
- 2^{nd} Scenario: number of training samples, from 100 to 375. The remaining samples were used as test set to verify the performance of the model.

Table 3 shows the influence of the number of training samples on the performance of the recognition model. From this table, we can see that the PCANet model achieves an accuracy of 80%, i.e. the algorithm still can achieve good results when the training set has less 300 samples. Thus, we do not need to collect a massive number of experimental data when using PCANet to construct the discrimination model for food seal fault detection.

Table 4 shows the computation time for the training dataset (total time in seconds) and the average computation time for each hypercube image, in case of the complete (1^{st} scenario) and reduced dataset (2^{nd} scenario). As expected, a reduced dataset implies that the training time is greatly reduced when compared to a bigger dataset.

Table 3. The performance of SVM classifier on the food seal database

#Exper	accuracy	F-measure
1^{st} Scenario	0.90	0.87
2^{nd} Scenario	0.80	0.79

Table 4. Computation time for PCANet. Total train time and average computation time for one test sample (hypercube).

#Exper	Total train time (s)	Avg. test time (s).
1^{st} Scenario	423.50	0.91
2^{nd} Scenario	43.55	0.76

Concerning the computation time for a single hypercube data, the obtained model for the extended data set requires 0.91 s versus 0.76 s for the model obtained in the reduced dataset. These computation times means that 65 and 78 trays per second could be analyzed in extended and reduced dataset, respectively.

In any case, food industry requires a very strict quality control, which means that classification ratio must be as high as possible. In order to rise the performance of the algorithm, one main issue concerns data. While the obtained dataset was big enough, it is desirable to obtain a dataset with representative anomalies, including all possible contamination products that could happen in the production line. Another future direction concerns the type of anomaly; it was observed that the hyperspectral camera, together with the optical and lighting setup, was capturing some faults poorly. Thus, if acquired images do not have enough quality, the algorithm will not be able to extract enough information for anomaly detection.

Related to algorithms, a future line os research is oriented to individual fault analysis. Typically, having one classifier for each anomaly might lead to better results. On the contrary, this option would require more computation time as multiple algorithms must be running concurrently. Concerning the data fusion techniques, some other improvements apart from the techniques proposed in this work could be used, in order to find the optimal metric for data fusion, applying it in an optimized form.

6. CONCLUSION

In this paper, we propose a novel framework for hyperspectral image classification based on single image data input, PCANet, and SVM/KNN classifiers. Three data fusion techniques for PCANet input data are proposed. The PCANet output generates a binarized histogram (feature vector) obtained by concatenating two PCA filter banks, which shows good invariance to the shape and the illumination of the image. Obtained results are close to 90% accuracy and 90% F-measure.

This work shows that hyperspectral image analysis can detect food tray contamination in cases where the human eye cannot; especially in tray covers using opaque seal such as a metallic cover used in this dataset. Extended and reduced datasets are analyzed, showing that PCANet has excellent performance in learning features from small datasets. The computation time allows a food tray processing speed above 65 trays per minute.

A future research topic to take the most of the spectral information will be the use of other spectral information extraction techniques apart from the fusion rules described in this work. Principal Component Analysis of spectra and other data reduction techniques will be explored.

ACKNOWLEDGEMENTS

We would like to thank the collaboration of PYGSA Company, who provided the hyperspectral camera and prepared all food tray data samples used in this work.

REFERENCES

Bioucas-Dias, J.M., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N., and Chanussot, J. (2013). Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and remote sensing magazine*, 1(2), 6–36.

Chan, T.H., Jia, K., Gao, S., Lu, J., Zeng, Z., and Ma, Y. (2015). Pcanet: A simple deep learning baseline for image classification? *IEEE transactions on image processing*, 24(12), 5017–5032.

Chen, J., Zhang, W., Qian, Y., and Ye, M. (2018). Deep tensor factorization for hyperspectral image classification. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, 4788–4791. IEEE.

Du, C.J. and Sun, D.W. (2006). Learning techniques used in computer vision for food quality evaluation: a review. *Journal of food engineering*, 72(1), 39–55.

Erichson, N.B., Zheng, P., Manohar, K., Brunton, S.L., Kutz, J.N., and Aravkin, A.Y. (2018). Sparse principal component analysis via variable projection. *arXiv preprint arXiv:1804.00341*.

Gowen, A., O'Donnell, C., Cullen, P., Downey, G., and Frias, J. (2007). Hyperspectral imaging–an emerging process analytical tool for food quality and safety control. *Trends in food science & technology*, 18(12), 590–598.

Haggag, M., Abdelhay, S., Mecheter, A., Gowid, S., Musharavati, F., and Ghani, S. (2019). An intelligent hybrid experimental-based deep learning algorithm for tomato-sorting controllers. *IEEE Access*, 7, 106890–106898.

Hassannejad, H., Matrella, G., Ciampolini, P., De Munari, I., Mordonini, M., and Cagnoni, S. (2016). Food image recognition using very deep convolutional networks. In *Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management*, 41–49. ACM.

Huang, H., Liu, L., and Ngadi, M.O. (2014). Recent developments in hyperspectral imaging for assessment of food quality and safety. *Sensors*, 14(4), 7248–7276.

Huang, S., Zhang, H., and Pižurica, A. (2017). A robust sparse representation model for hyperspectral image classification. *Sensors*, 17(9), 2087.

Kagaya, H., Aizawa, K., and Ogawa, M. (2014). Food detection and recognition using convolutional neural network. In *Proceedings of the 22nd ACM international conference on Multimedia*, 1085–1088. ACM.

Lee, D.D. and Seung, H.S. (2001). Algorithms for non-negative matrix factorization. In *Advances in neural information processing systems*, 556–562.

Lee, J.N., Byeon, Y.H., Pan, S.B., and Kwak, K.C. (2018). An eigenecg network approach based on pcanet for personal identification from ecg signal. *Sensors*, 18(11), 4024.

Li, S., Kang, X., Fang, L., Hu, J., and Yin, H. (2017). Pixel-level image fusion: A survey of the state of the art. *Information Fusion*, 33, 100–112.

Liu, C., Cao, Y., Luo, Y., Chen, G., Vokkarane, V., and Ma, Y. (2016). Deepfood: Deep learning-based food image recognition for computer-aided dietary assessment. In *International Conference on Smart Homes and Health Telematics*, 37–48. Springer.

Ostyn, B., Darius, P., De Baerdemaeker, J., and De Ketelaere, B. (2007). Statistical monitoring of a sealing process by means of multivariate accelerometer data. *Quality Engineering*, 19(4), 299–310.

Ryser, E.T. and Marth, E.H. (1999). *Listeria: Listeriosis, and Food Safety*. CRC Press.

Soon, F.C., Khaw, H.Y., Chuah, J.H., and Kanesan, J. (2018). Pcanet-based convolutional neural network architecture for a vehicle model recognition system. *IEEE Transactions on Intelligent Transportation Systems*, 20(2), 749–759.

Sun, D.W. (2010). *Hyperspectral imaging for food quality analysis and control*. Elsevier.

Tiger, B. and Verma, T. (2013). Identification and classification of normal and infected apples using neural network. *International Journal of Science and Research*, 2(3), 160–163.

Turan, M.T. and Erzin, E. (2018). Food intake detection using autoencoder-based deep neural networks. In *2018 26th Signal Processing and Communications Applications Conference (SIU)*, 1–4. IEEE.

WRAP (2000). Example website. `http://www.wrap.org.uk/`. Accessed on 2019-10-28.

Zhou, L., Zhang, C., Liu, F., Qiu, Z., and He, Y. (2019). Application of deep learning in food: A review. *Comprehensive Reviews in Food Science and Food Safety*.