Deep Spectral-spatial Features of Snapshot Hyperspectral Images for Red-meat Classification

Mahmoud Al-Sarayreh*†, Marlon M. Reis †, Wei Qi Yan *, and Reinhard Klette*
*School of Engineering, Computer and Mathematical Sciences
Auckland University of Technology, Auckland, New Zealand. {*malsaray, wyan, rklette*}*@aut.ac.nz*
†AgResearch, Palmerston North, New Zealand. *marlon.m.reis@agresearch.co.nz*

***Abstract*—We investigate the potential and accuracy of snapshot hyperspectral imaging for authentication and classification of red-meat species. Snapshot hyperspectral images are acquired of lamb, beef, and pork samples. We consider 13 muscles types of standard loin and leg chops. We propose a deep 3D convolution neural network (CNN) architecture for extracting and classifying spectral-spatial learned features of red-meat. We present a comparison with state-of-the-art models including partial least-square discriminant analysis and support vector machines. Our results show that the proposed 3D-CNN model outperforms the state-of-the-art models with 95.81% and 96.01% for overall accuracy and average $\mathbf{F}_1$ score, respectively. Visualization results show that the proposed 3D-CNN model is able to convert snapshot hyperspectral image data into an intelligent representation with accurate separation between red-meat types. This study opens the door for more research towards real-time and completely portable hyperspectral imaging systems due to the ability of snapshot hyperspectral cameras to work at video rate.**

***Index Terms*—snapshot hyperspectral imaging, red meat classification, spectral-spatial features, meat processing, adulteration detection, deep learning, 3D-CNN.**

I. Introduction

Meat processing, regarding quality and safety, is gaining more attention in the meat industry and research. Customers pay more attention to the authenticity and safety of meat as they expect a high-quality product compared with the paid price. Meat authenticity is considered to be one of the safety attributes of meat, where accurate labelling of meat products is important from a customer point-of-view due to fair-trading or religious reasons. Detection of meat fraud (e.g. of mislabeled products) is a challenging task in meat processing plants. Practically, red-meat identification processes are usually performed manually (laboratory-based) in the industry, which is time-consuming, and subjected by human errors.

The *spectroscopy technology* [1] is considered as a rapid and non-destructive way to asses and recognize the type of materials by analyzing their spectral properties. Spectroscopy received attention in many application areas like raw material analysis [2], product quality control, and processes monitoring [3]. However, spectroscopy technology does not provide a spatial distribution of the evaluated attributes on the sample surface, where it still requires manual measurements of small parts of the sample, followed by computing the mean of a target measure for the whole sample. In many practical cases, the spatial distribution of quality parameters is also needed for visualization and monitoring purposes.

Recently, *hyperspectral imaging* (HSI) systems are introduced to overcome limitations in spectroscopy technology (i.e., better availability of spatial distribution, processing speed, and accuracy). Moreover, HSI systems are able to merge existing computer-vision technology, by RGB imaging, with *chemometrics analysis* for many tasks such as food safety, quality grading, or classification. Thus, HSI systems provide robust, rapid, and non-destructive solutions in many research areas, for example in agricultural and remote sensing [4], medical imaging [5], and food processing [6]. In fact, the mentioned robustness is possible by providing unique spectral signatures for each material shown in the image, and also by providing spatial attributes of these materials; attributes like graded quality distribution, the shape of objects, texture properties, or the ability of object localization.

HSI systems are a valuable tool for visualizing the chemical components of materials by means of an image, providing detailed information about their types and shapes. An HSI image $\mathcal{H} : \mathcal{D}_x \times \mathcal{D}_y \times \mathcal{D}_\lambda \longrightarrow \mathbb{R}$ defines a 3-dimensional (3D) *hypercube*, where $[\mathcal{D}_x, \mathcal{D}_y]$ represents the space of spatial coordinates $[x, y]$ of an object in the image (describing location and shape of objects), and $\mathcal{D}_\lambda$ is the space of reflectance intensity vectors across a specific range of wavelengths (λ provides a chemical composition of an object inside the image). Figure 1 shows a schematic representation of a hypercube. The hypercube is a stack of grey-scale images at specific wavelengths, each grey-scale image is called a *band* and defined as $\mathcal{B}_\lambda : \mathcal{D}_x \times \mathcal{D}_y \longrightarrow \mathbb{R}$

The rest of this paper is structured as follows. Section II reviews state-of-the-art techniques in HSI classification and analysis. Section III shows the used HSI system and the

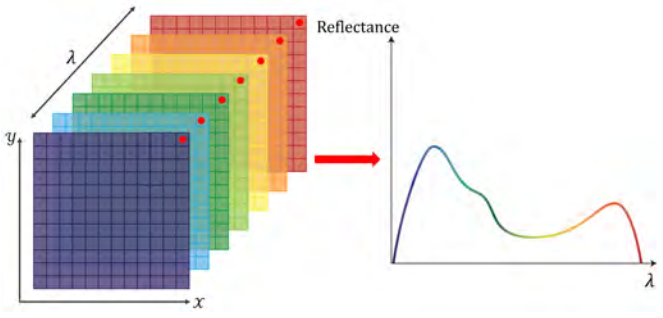

Fig. 1. Schematic representation of a hypercube; red points in this hypercube show one pixel location; its spectral signature is shown on the right

collected dataset. The proposed methodology and models are briefly described in Section IV. Section V provides the experimental setup and the achieved results for the proposed models. An analysis and discussion about the results is given in Section VI. Section VII concludes.

II. Related work

The standard way of collecting the spectral information in HSI systems is by line scanning. In line scanning, the sample moves (on a conveyor belt) and, at the same time, the camera detector detects the reflected light of a particular row (or line) on the sample surface. This way, the whole sample is scanned and then reconstructed in the collected data as one HSI image.

Recently, new hyperspectral sensors [16], called snapshot hyperspectral cameras, are introduced as a solution for limitations (i.e., the speed of hyper-cube acquisition) in line-scanning HSI cameras. A snapshot HSI system is able to collect the images at video rate, which make these sensors more applicable for real-time applications. In the camera's array detector, each pixel is represented as a small sub-array (e.g. 4×4 or 5×5) called micro-pixel [16]. Each cell in the micro-pixel detects the reflectance intensity at a particular wavelength [16]. The resulting image is a high-resolution image containing the spectral pattern (i.e., the micro-pixels), then the image is converted into three dimensions (the hypercube).

Snapshot HSI systems are successfully used in applications like terrain classification for autonomous driving systems [17], [18], vegetation classification using UAVs [19], or for object recognition [20]. Line-scanning HSI systems are used in many (especially indoor) applications areas such as meat [7]–[11], fruit and vegetable [12], [13], or fish processing [14], [15]. The main advantage of the line-scanning method is that it provides deep spectral features (typically covering hundreds of bands with a very fine spectral resolution). However, image acquisition is very slow here, and a large size of the hypercube is computationally expensive.

In [8], a line-scanning HSI system was used to classify the type of lamb muscle into semitendinosus (ST), longissimus dorsi (LD), or psoas major (PM). A combination between *principal component analysis* (PCA) and *linear discrimination analysis* (LDA) was used to fit the data and to classify the type of muscle. In [7], the ability of the HSI system for classifying the types of red-meat was investigated by using a set of LD muscle samples of lamb, beef, and pork. The extracted spectral features were pre-processed by second derivatives for obtaining the optimal wavelengths. *Partial least-squares discriminant analysis* (PLS-DA) was used for classification. Sample-wise results showed that the model performs well on the test set. However, the model produces a lot of misclassified pixel. So, the authors proposed a majority-voting technique to obtain the final classification decision.

Red-meat adulteration was investigated in [9] using HSI. Collected images have challenging conditions (packing, fresh, frozen, and thawed) of meat samples of lamb, beef, or pork. In [9], the main task is to classify one type of meat against the others under the considered conditions. *Support vector machine* (SVM) results showed that a use of only spectral features was insufficient for solving this problem. A new set of texture features was proposed by combining the superpixel technique with Haralick features of the *grey-level co-occurrence matrix* (GLCM) [9], [25], concatenated with spectral features.

Recently, deep learning approaches for supervised learning are considered the state-of-the-art in many computer vision applications. Deep *convolutional neural network* (CNN) models show robustness as novel feature extractor for raw input data. CNN models are applied on HSI images for object recognition [20], remote sensing [22]–[24], plant-disease detection [21], and meat processing [10]. Moreover, CNNs show flexibility to deal with HSI data by introducing the 1D-CNN in [10], [22] (designed for processing spectral inputs), 2D-CNN for single wavelength images or PCA-component images [23], and 3D-CNN for an intelligent combination of spectral and spatial image data [10], [23], [24].

In [10], comprehensive comparisons were provided between a deep learning approach and hand-crafted features (i.e., manually extracted spectral and spatial features as in [9]). In [10], a combination of 1D-CNN and 3D-CNN was proposed for extracting learned features for spectral and spatial domains.

For the same problem as in [9], results showed that the CNN model significantly outperformed the SVM model with hand-crafted features. This study aims at investigating the robustness of the new snapshot hyperspectral imaging sensors for solving the red-meat classification problem. Thus, the main objectives of this paper are as follows: Developing a methodology for data acquisition and sampling of each image into a number of representative samples for modelling, developing a deep-learning model for classifying snapshot HSI image data, and for visualizing the robustness of a CNN model against state-of-the-art models.

III. Dataset and HSI system

The HSI camera, as used here, was introduced by Ximea using on-chip snapshot mosaic filters [16]. The camera's array detector consists of a grid of micro-cells of 216×409, for height and width, respectively. Each cell in the grid is a square of a 5×5 sub-grid (micro-pixel) of 25 spectral pattern filters in the range of $672.74 \sim 957.49$ *nm*. Thus, the dimensions of the resulting HSI images are $216 \times 409 \times 25$. The HSI system consists of a snapshot HSI camera (Ximea MQ022HG-IM-SM5X5-NIR), an illumination unit of two halogen lambs (150 *W*), a controlled movable conveyor belt, and a computer running an image acquisition software (HSImager). The speed of the conveyor belt and the distance between the camera and the conveyor belt were set to be 5.5 *cm/s* and 35.5 *cm*, respectively. These values were empirically adjusted and then synchronized with the camera to capture HSI images with a spatial resolution of 0.27×0.27 *mm/pixel*.

The resulting HSI images (raw spectral irradiance data, with the mentioned HSI system) need to be preprocessed for obtaining reflectance intensities and for a correction of

the illumination distribution. The reflectance and illumination correction is computed as follows:

$$R(x,y) = \frac{R_o(x,y) - D(x,y)}{W(x,y) - D(x,y)} \times N \quad (1)$$

$$N = \frac{\text{mean}\,(R_o)}{\text{mean}\left(\frac{R_o(x,y) - D(x,y)}{W(x,y) - D(x,y)}\right)} \quad (2)$$

where R is a corrected HSI image in reflectance, R_o is the raw snapshot HSI image, and D and W are the dark and white reference images, respectively. The N factor is a normalization constant that is used to recover the original intensities of the camera. The images are normalized to obtain reflectance values in the range of $0 \sim 1.0$ as follows:

$$R_\lambda(x,y) = \frac{R_\lambda(x,y)}{\max(D_\lambda)} \quad (3)$$

A collection of red-meat samples were procured from two local supermarkets. The total number of samples is 140, including lamb (50), beef (55), and pork (35). All of the samples were chosen from the loin and leg chops, which include 13 muscle types based on the meat standards. In [27], more descriptions were provided about muscles types and structures in both loin and leg chops. A set of 105 samples was used for training and processing, and the remainders were used for only testing and evaluation. All samples were scanned by the developed HSI system; each sample was put on a movable conveyor belt while the snapshot HSI camera collected a sequence of images of the sample.

Figure 2 shows an example of these sequences of two lamb samples. The key idea of using the sequences is for obtaining representative images under different illumination conditions (i.e. the same meat portion with different angles to the camera and the light source); in addition, all sample portions are covered in these sequences. The average number of HSI images per sample was 6; the total number of images in the collected data set is 600 which defines a reasonable dataset for deep-learning-based models.

The main task in this work is developing a pixel-based classification model of red-meat products. Thus, considering the mean spectrum of each meat type, for building a model, such as in [7], [8],it is not efficient for the following reasons:

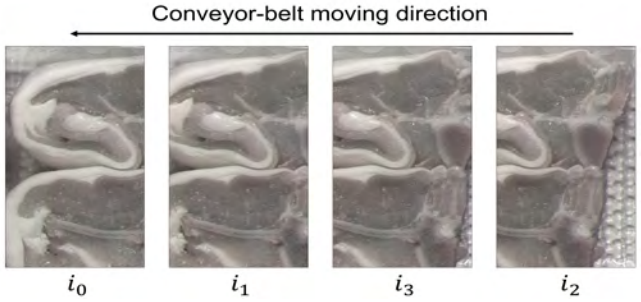

Fig. 2. Example of snapshot HSI images (false-color images), acquired for representing two lamb loin-chop samples. The sequence in the figure represents the time of acquisition during the motion of a conveyor belt

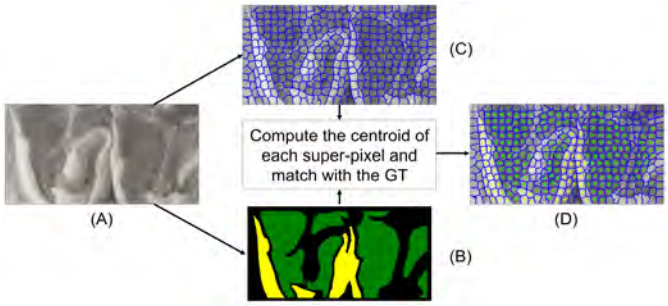

Fig. 3. Methodology for re-sampling an HSI image into a set of representative points. (a) An HSI image of two lamb samples (an extracted false-color image of an input hypercube). (b) GT image manually generated by using a multi-polygon tool. (c) Superpixel segments of the input image. (d) Selected points from each class; colors Green and Yellow denote meat and fat, respectively

(1) Considered muscles are 13; averaging them we lose the spectral patterns. (2) Averaging the spectral response of meat of the whole sample does not represent the variation of light and meat within the sample. (3) Inefficient for complex learning models like deep learning due to a limitation in the number of samples.

In this paper, we develop a methodology for re-sampling HSI image data into a set of representative points (i.e., regions from each muscle shown in the image). These points are then used to extract the spectral features or patches of an $s \times s \times s$ 3D window around each point. As superpixel algorithm we use the SLIC method [26] for segmenting HSI images into a set of labelled segments; one segment (i.e., one superpixel) consists of a group of pixels that share the same spatial and spectral properties. In addition, the images were manually labelled (i.e. for having ground-truth images) into meat and fat classes. Then, the segmentation map and *ground truth* (GT) are matched for generating a ground truth of the super-pixel map. For each super-pixel, the centroid was computed; then its coordinates were used for deciding whether it is meat, fat, or background. Figure 3 sketches the proposed methodology.

The proposed re-sampling methodology was used to extract $86,535$ representative points from the whole training dataset, for all classes; the considered classes are LAMB, BEEF, PORK, and FAT. Table I shows the number of selected points for each class.

IV. Deep 3D-CNN model

Snapshot HSI cameras are designed as on-chip micro-multi-spectral detectors. They usually cover a limited number of wavelengths; for example, the camera that was used here covers only 25 wavelengths. This limitation in spectral features inspired us to investigate complex models for dealing with the

TABLE I
THE NUMBER OF SELECTED POINTS FROM TRAINING IMAGES

Class	Number of selected points
LAMB	23,324
BEEF	30,681
PORK	19,839
FAT	12,691

challenges in spectral information. In this paper, we propose a deeper 3D-CNN model for the extraction of useful features of snapshot HSI images. We evaluate these features on a red-meat classification problem. Moreover, we evaluate and compare the proposed CNN model with a state-of-the-art model of red-meat classification problem as provided in [7], and with an SVM [28] implementation, as the SVM is a common reference for machine-learning models.

The proposed 3D-CNN model consists of a hierarchical structure of 9 layers as follows: An input layer as a 3D window of size $S \times S \times \lambda$ (here, λ is fixed by the camera specs, with $\lambda = 25$), a set of four 3D convolution layers with different kernel representations (structured as two contiguous convolution blocks), and two 3D max-pooling layers for down-sampling the feature maps of each convolution block. The extracted CNN features were then connected with a fully connected dense layer as feature extractor for red-meat types. Finally, we added an output layer as a fully-connected softmax layer for having the probability of being one of the meat types.

3D convolution operations are able to handle and extract learned features among spectral and spatial domains. Multiple 3D-CNN layers extract robust and complex features in a hierarchical way; each layer extracts features from the previous layer. The output of the $i-th$ CNN layer is computed as follows:

$$x_i = \mathcal{F}_i(u_{i-1}) \tag{4}$$

where $u_{i-1} = w_{i-1}^T x_{i-1} + b_{i-1}$ and w_{i-1}^T and b_{i-1} are the weight matrix and the bias vector of the previous layer. $\mathcal{F}_i(\cdot)$ is a non-linear transform applied to the output of each of the 3D-CNN layers.

In the training process, forward-propagation and back-propagation processes were used to compute and update the weights of the CNN and of the fully-connected layers. In the back-propagation process, the weights of the model were adjusted using a stochastic gradient descent approach. In these processes, the main task is to minimize a loss function between the actual output and the model output (i.e., a prediction in feedforward processes). In this work, we used the *adaptive moment estimation* (Adam) optimizer [29] for optimizing the following loss function (categorical cross entropy):

$$\mathcal{L}(\theta) = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{k}\{j = Y^{(i)}\}\log(y_j^{(i)}) \tag{5}$$

where n is the number of samples (batch size), k is the number of classes, Y is the ground-truth vector (encoded as one-hot vector style), y is the model prediction vector, and θ represents the model parameters (i.e., weights and biases).

V. Experiments

The dataset and HSI system, described in Section III, were used for evaluating the proposed deep 3D-CNN model. The training patches (i.e., a window around selected points; see Table I) were extracted using the proposed method as described in Section III. The parameters of SLIC [26], the number of segments and the compactness factor, were empirically chosen to be 300 and 0.4, respectively.

For comparison, we implemented PLS-DA [7] and SVM with *radial basis function* (RBF) kernels [28]. For both models, a window of size $5 \times 5 \times 25$ around each point was extracted; then we compute the mean spectrum of each window as a spectral feature vector of size 25×1. Extracted spectral feature vectors were used as input for the models. In PLS-DA, the spectral features were pre-processed by a 2nd-derivative method [7]. For hyper-parameter tuning, a 10-fold cross-validation approach was used for selecting optimal parameters for the PLS-DA and SVM models.

In the proposed 3D-CNN model, architecture and parameters were chosen empirically and based on numerous experiments for achieving a high accuracy and a light-weight model. Overfitting is one of the challenges in deep learning. To avoid overfitting in our model, we used the dropout technique [30] by adding three dropout layers after the pooling and fully-connected layers. Moreover, we sought to reduce the whole model size by tuning a small number of kernels, which is good for generalization [24], [30]. Table II shows all details of the proposed architecture; in pooling layers, we used a stride of $(2, 2, 2)$ for down-sampling the spatial and spectral domains. In the training phase, we used Adam optimizer [29] with a learning rate of 0.0001 and patch size of 256. The model was trained for $2,000$ epochs until its convergence. The model converged well and showed a stable convergence in both training and validation, as shown in Fig. 4, left.

The spatial size is an important parameter in the proposed model. We evaluated different spatial windows as input of the 3D-CNN model. Results, as shown in Fig. 4, right, show a trade-off in model performance. Thus, the optimal value of spatial size was chosen as a 3D window of size $9 \times 9 \times 25$.

For evaluating the proposed model and comparing with the PLS-DA and SVM-RBF models, we used the standard F_1 measure, overall accuracy, and average accuracy. As F_1 provides a harmonic mean of both recall and precision, we used it for evaluating the accuracy of classes.

As described in Section III, the testing set of samples was used for evaluation. Based on GT images, regions of each class

TABLE II
Architecture and specs of the proposed 3D-CNN model

Layer	Kernel size	Feature maps	Output size	Activation function	Dropout
Input	–	–	9 x 9 x 25	–	–
3D_Conv 1	5 x 5 x 5	4	4 x 9 x 9 x 25	ReLU	–
3D_Conv 2	5 x 5 x 5	4	4 x 9 x 9 x 25	ReLU	–
3D_MaxPooling 1	–	4	4 x 5 x 5 x 13	–	0.25
3D_Conv 3	3 x 3 x 3	8	8 x 5 x 5 x 13	ReLU	–
3D_Conv 4	3 x 3 x 3	8	8 x 5 x 5 x 13	ReLU	–
3D_MaxPooling 1	–	8	8 x 3 x 3 x 7	–	0.25
Flatten	–	–	504	–	–
Fully Connected	–	–	128	ReLU	0.25
Output	–	–	4	Softmax	–
The total number of trainable parameters is 70,272					

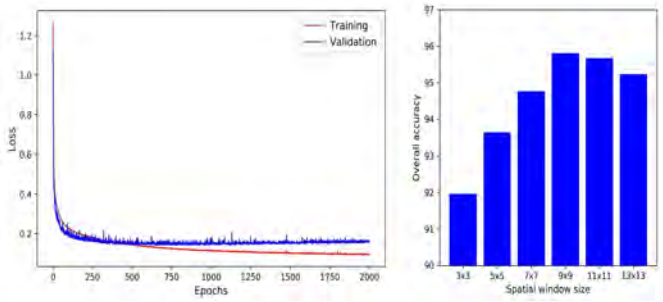

Fig. 4. *Left*: Learning curve of the proposed 3D-CNN model. *Right*: Impact of spatial size on the accuracy of the proposed 3D-CNN.

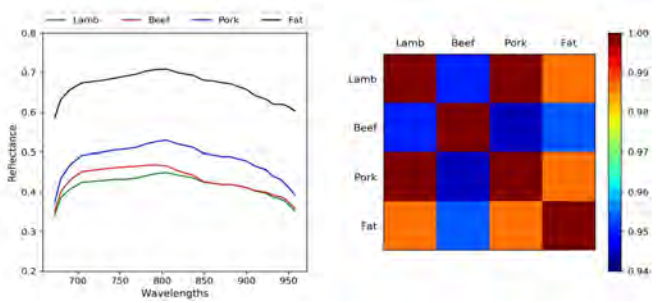

Fig. 5. *Left*: Extracted spectral signatures of red-meat types and their fat. *Right*: Correlation-coefficient matrix showing the similarity (or the dependence) between signatures pairs; see color index on the right for Pearson correlation coefficients

were extracted as patches around each pixel; the patches are $9 \times 9 \times 25$ for the 3D-CNN model, and $5 \times 5 \times 25$ for both PLS-DA and SVM models. Then, we passed the patches to the models for evaluating purposes. Table III shows the evaluating measures for the investigated models.

VI. Results and Discussion

A key benefit of HSI images is that they provide a spectral signature for each material inside the image. For visualizing the signatures of the meat types, we extracted a set of batches (for a set of 2,000 pixels from each class) of size $9 \times 9 \times 25$ from the testing set. Then we compute the mean spectrum of each patch. After that, the mean of all spectra of each class was computed which ends with four 1×25 vectors, one signature for each class. As shown in Fig. 5, left, signatures are highly correlated and have high similarity in shape which shows the challenge in spectral features provided by snapshot HSI sensors. For quantifying this observation, we computed the Pearson correlation coefficient, as shown in Fig. 5, right, where high positive values mean that the linearity between class pairs is high. This observation shows that the only use of spectral features of HSI snapshots is insufficient for achieving high accuracy in material-based classification problems. This result is clearly documented by our experiments; considering both spatial and spectral features by our 3D-CNN outperforms the models using only spectral features, as shown in Table III.

The evaluation results in Table III clearly show that the proposed 3D-CNN model achieved the highest efficiency in terms of per-class F_1 score, average F_1 score of all classes, and overall accuracy. Our 3D-CNN model achieves 95.8% overall accuracy compared with 90.1% and 84.2% for SVM and PLS-DA, respectively. The achieved high F_1 value represents the robustness of our model for a very good discrimination between meat types. Visual results in Fig. 6 show the robustness of the proposed 3D-CNN model in terms of classifying all sample portion correctly and maintaining the edges between meat and fat.

TABLE III
Performance evaluation of the proposed 3D-CNN model, in comparison with PLS-DA and SVM-RBF for red-meat classification

Model	F_1 score				Mean F_1 score	Overall accuracy	Average accuracy
	LAMB	BEEF	PORK	FAT			
PLS-DA	80.5	88.1	87.1	76.5	83.1	84.2	81.2
SVM-RBF	86.1	93.9	89.6	91.8	90.3	90.1	90.9
3D-CNN	93.6	98.3	94.6	97.7	96.1	95.8	96.1

Due to the portability of HSI snapshot cameras (i.e., being a completely portable device), we evaluated the time for image-classification (i.e., for classifying a single image of 216×409) for the investigated models. Results show another efficiency of the 3D-CNN model, where the 3D-CNN model was 4.7 times faster than SVM, and 2.9 times faster than the PLS-DA model; the classification times are 13, 38, 61 seconds for 3D-CNN, PLS-DA, and SVM, respectively, running on the same machine.

We also investigated the features that were learned by the 3D-CNN model. Randomly, we selected a set of patches from the testing set. Then, we projected these patches into the model and extracted the computed output of the fully-connected

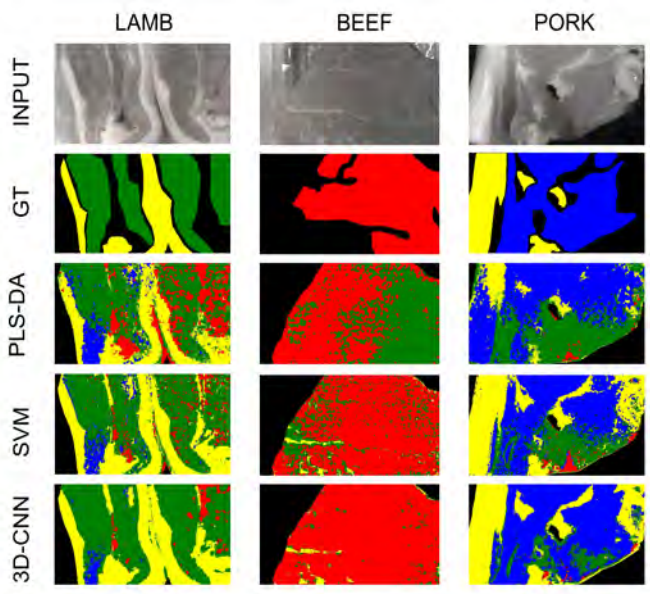

Fig. 6. Classification maps of 3D-CNN, SVM, and PLS-DA provide a visual comparison between the models; colours Green, Red, Blue, and Yellow represent classes LAMB, BEEF, PORK, and FAT, respectively

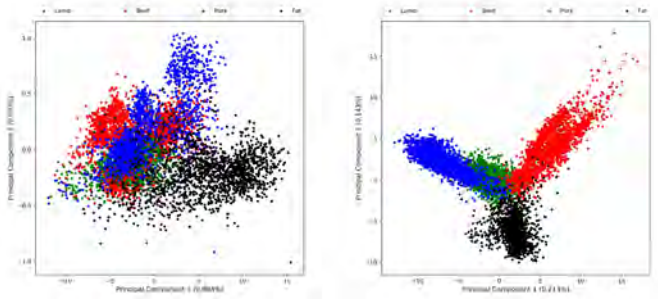

Fig. 7. PCA scatter plots for visualizing the separation between classes. *Left*: First two PCA components of the original spectral features. *Right*: First two PCA components of the learned features extracted by 3D-CNN model.

layer, layer number 8 as shown in Table II. Extracted feature vectors (of size 128) were then fitted on a PCA model for reducing the dimensions. For the same patches, we extracted the mean spectra and applied another PCA model. For visual comparison, we plotted the first two components of both PCA models, as shown in Fig. 7. Clearly, Fig. 7, left, shows that the 3D-CNN model is able to convert raw spectral-spatial data into a useful representation with a very good separation between the classes in the PCA space, while in case of the original spectral data, see Fig. 7, right, the class regions are highly overlapping and look only like two clusters.

VII. Conclusion

Red-meat identification and authentication are important tasks in the meat industry. In this study, we investigated the potential and robustness of snapshot HSI systems, which provide limited spectral information for red-meat identification and authentication. Also, we investigated the robustness of deep learning models (a 3D-CNN architecture) for classifying the type of meat. The quantitative and visual analysis clearly shows that the proposed deep 3D-CNN model outperforms PLS and SVM models by achieving 96.1% and 95.8% for average F_1 and overall accuracy, respectively. In addition, results showed that the 3D-CNN model is much faster than the other two models which supports the use of a completely portable implementation.

References

[1] Cen, H., He, Y.: Theory and application of near-infrared reflectance spectroscopy in determination of food quality. *Trends Food Science Technology*, 2007, 18, 72–83

[2] Wu, D., Nie, P., He, Y., Bao, Y.: Determination of calciumcontent in powdered milk using near and mid-infrared spectroscopy with variable selection and chemometrics. *Food Bioprocessing Technology*, 2012, 5(4), 1402–1410.

[3] Wu, D., He, Y., Feng, S., Sun, D. W.: Study on infrared spectroscopy technique for fast measurement of protein content in milk powder based on LS-SVM. *J. Food Engineering*, 2008, 84(1), 124–131.

[4] Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., Sousa, J.J.: Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote Sensing*, 2017, 9, 1110.

[5] Lu, G., Fei, B.: Medical hyperspectral imaging: A review. *J. Biomedical Optics*, 2014, 19(1), 010901.

[6] ElMasry, G. M., Nakauchi, S.: Image analysis operations applied to hyperspectral images for non-invasive sensing of food quality – A comprehensive review. *Biosystems Engineering*, 2016, 142, 53–82.

[7] Kamruzzaman, M., Barbin, D., Elmasry, G., Sun, D., Allen, P. Potential of hyper spectral imaging and pattern recognition for categorization and authentication of red meat. *Innovative Food Science Emerging Technologies*, 2012, 104, 332–340.

[8] Kamruzzaman, M., Elmasry, G., Sun, D.W., Allen, P. Application of NIR hyperspectral imaging for discrimination of lamb muscles. *J. Food Engineering*, 2012, 104, 332–340.

[9] Al-Sarayreh, M., Reis, M., Yan, W.Y., Klette, R.: Detection of adulteration in red meat species using hyperspectral imaging. Proc. *Pacific-Rim Symp. Image Video Technology*, 2017, LNCS10749, 182–196.

[10] Al-Sarayreh, M., M. Reis, M., Qi Yan, W., Klette, R.: Detection of red-meat adulteration by deep spectral–spatial features in hyperspectral images. *J. Imaging*, 2018, 4, article number 63.

[11] Reis, M., Beers, R., Al-Sarayreh, M., Shorten, P., Yan, W., Saeys, W., Klette, R., Craigie, C.: Chemometrics and hyperspectral imaging applied to assessment of chemical, textural and structural characteristics of meat. *Meat Science*, 2018, 144, 100–109.

[12] Ariana, D. P., Lu, R.: Hyperspectral waveband selection for internal defect detection of pickling cucumbers and whole pickles. *Computers Electronics Agriculture*, 2010, 74(1), 137–144.

[13] Tsuta, M., Takao, T., Sugiyama, J., Wada, Y., Sagara, Y.: Foreign substance detection in blueberry fruits by spectral imaging. *Food Science Technology Research*, 2006, 12(2), 96–100.

[14] Wu, D., Sun, D. W.: Potential of time series-hyperspectral imaging (TS-HSI) for non-invasive determination of microbial spoilage of salmon flesh. *Talanta*, 2013, 111, 39–46.

[15] Wu, D., Sun, D. W.: Application of visible and near infrared hyperspectral imaging for non-invasively measuring distribution of water-holding capacity in salmon flesh. *Talanta*, 2013, 116, 266–276.

[16] Geelen, B., Tack, N., Lambrechts, A. A.: A compact snapshot multispectral imager with a monolithically integrated per-pixel filter mosaic. Proc. SPIE 8974, Advanced Fabrication Technologies Micro/Nano Optics Photonics, 2014, 89740L.

[17] Winkens, C., Sattler, F., Paulus, D.: Hyperspectral terrain classification for ground vehicles. Proc. *VISIGRAPP*, 2017, 417–424.

[18] Winkens, C., Kobelt, V., Paulus, D.: Robust features for snapshot hyperspectral terrain-classification. Proc. *Int. Conf. Computer Analysis Images Patterns*, 2017, LNCS 10424, 16–27.

[19] Ishida, T., Kurihara, J., Viray, F. A., Namuco, S. B., Paringit, E. C., Perez, G. J., Takahashi, Y., Marciano, J. J.: A novel approach for vegetation classification using UAV-based hyperspectral imaging. *Computers Electronics Agriculture*, 2018, 144, 80–85.

[20] Fotiadou, K., Tsagkatakis, G., Tsakalides, P.: Deep convolutional neural networks for the classification of snapshot mosaic hyperspectral imagery. *Electronic Imaging*, 2017, 17, 185–190.

[21] Nagasubramanian, K., Jones, S., Singh, A. K., Singh, A., Ganapathysubramanian, B., Sarkar, S.: Explaining hyperspectral imaging based plant disease identification: 3D-CNN and saliency maps. *arXiv:1804.08831*, 2018.

[22] Hu, W., Huang, Y., Wei, L., Zhang, F., Li, H.: Deep convolutional neural networks for hyperspectral image classification. *J. Sensors*, 2015, article id 258619.

[23] Chen, Y., Jiang, H., Li, C., Jia, X., Ghamisi, P.: Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Trans. Geoscience Remote Sensing*, 2016, 54, 6232–6251.

[24] Li, Y., Zhang, H., Shen, Q.: Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sensing*, 2017, 9, 67.

[25] Klette, R.: *Concise Computer Vision*. Springer: London, UK, 2014.

[26] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Analysis Machine Intelligence*, 2013, 34, 2274–2282.

[27] Swatland, H. J.: *Meat Cuts and Muscle Foods: An International Glossary*, 2nd edition. Nottingham University Press, 2004.

[28] Cortes, C., Vapnik, V.: Support-vector networks. *Machine Learning*, 1995, 20(3), 273–297.

[29] Kingma, D. P., Ba, J.: Adam: A method for stochastic optimization. *arXiv:1412.6980*, 2014.

[30] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: A simple way to prevent neural networks from overfitting. *J. Machine Learning Research*, 2014, 15(1), 1929–1958.