Identification of agricultural quarantine materials in passenger's luggage using ion mobility spectroscopy combined with convolutional neural network

Jixiong Zhang[1, 4, #, *], Jingjing Xia[2, #, *], Qingjun Zhang[3], Nei Yang[3], Guangqin Li[3], Fusuo Zhang[1,4]

[1] College of Resources and Environmental Sciences; National Academy of Agriculture Green Development; Key Laboratory of Plant-Soil Interactions, Ministry of Education, China Agricultural University, Beijing, 100193, China.

[2] Institute of Materia Medica, Xinjiang University, Urumqi, 830017, China

[3] Nucteh Company Limited, Beijing, 100084, China.

[4] National Observation and Research Station of Agriculture Green Development, Quzhou, 057250, China

[#] These authors contributed equally to this work.

[*] Corresponding author: E-mail address: zhangjixiong@cau.edu.cn, jxzhang_28@sina.com (J. Zhang); sumny1210@sina.com (J. Xia)

CONTENTS:

A brief description of CNN layers.

Fig. S1: Scenario of IMS data collection of agricultural quarantine materials in passenger's luggage.

Fig. S2. Schematic of the convolution process for (A) one-dimensional data, (B) two-dimensional data.

Fig. S3: Schematic of the pooling using two different methods.

Fig. S4: Schematic of confusion matrix.

Fig. S5: IMS spectra of eight principal volatile compounds which are the most characteristic or abundant compounds observed in pear, banana, apple, longan and grape from previous study.

Fig. S6: IMS spectra of eight samples randomly selected from (A) pear, (B) banana, (C) apple, (D) longan, (E) blanks and (F) grape group, respectively.

Fig. S7: IMS signal evolution for different types of AQMs：(A)pear, (B) banana, (C) apple, (D) longan, (E) blanks, (F) grape in luggage.

Table S1: Classification accuracy of the four algorithms (over 20 repeated operation).

Reference

A brief description of CNN layers:

Convolution layer: The convolution layer operates the convolution filters on the input and produces the feature maps. The size of the feature map (*OH*, *OW*) can be calculated using the formula (Equ.S1):

$$OH = \frac{H + 2P - FH}{S} + 1$$

$$OW = \frac{W + 2P - FW}{S} + 1$$

(S1)

where *OH* and *OW* are height and width size of the feature map, respectively. *H* and *W* are height and width size of the input, respectively. *FH* and *FW* are height and width size of the filter. *P* is the padding size. S is the stride size.

A simple convolution operation for 1D and 2D IMS data is illustrated in Fig. S2. The convolution filter (yellow rectangle) is sliding over the input (green rectangle). The area where the filter overlaps with the input is referred to a receptive field. The sum of the dot products between the filter and the corresponding receptive field from the input generates the elements of feature map. The filter is then slid by the stride value, and the convolution operation is repeated until the feature map of the given filter is produced (blue rectangle). A bias term is then added to the feature map to generate the output feature map.

Activation function: The feature map that the convolution filter creates is then processed through the activation function. Several activation functions exist, such as the sigmoid function, the tanh function and the ReLU function. In this study, ReLU function was employed[1] (Equ.S2):

$$f(x) = \begin{cases} x & (x > 0) \\ 0 & (x \leq 0) \end{cases}$$

(S2)

Batch Normalization layer: The batch normalization (BN) layer is used to help speed up training of CNN, reduce the sensitivity to network initialization, and prevent the network from overfitting[44]. It is defined as (Equ.S3):

$$BN(x_i) = \gamma \left(\frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \right) + \beta$$

(S3)

where ϵ is a constant for numerical stability, μ_B is the mini-batch mean, σ_B^2 is the mini-batch variance, γ is a scale parameter and β is a shift parameter.

Pooling layer: The pooling layer reduces the size of the input, as it combines neighboring pixels of a certain area of the input into a single representative value. In order to conduct the operations in the pooling layer, we should determine how to select the pooling pixels from the feature map and how to set the representative values. The neighboring pixels are usually selected from square matrix for 2D data and vector for 1D data. The representative value is usually set as the mean or maximum value of the selected pixels. The operation of the pooling layer is surprisingly simple. Fig. S3 shows the resultant cases of pooling using the mean pooling and maximum pooling.

The output feature maps extract from the input using piles of convolutional layer, and pooling layers pairs are flattened into a 1D vector. The vector is then fed into the fully connected layer.

Fully connected layer: The fully connected layer establishes the mapping between extracted features and the output as (Equ.S4):

$$y_j = \sum_i (w_{i,j} x_i) + b_j \tag{S4}$$

where y_j is the output of the j-th neuron, x_i is the i-th extracted feature, $w_{i,j}$ is the weight between extracted feature and the output, and b_j is the bias.

Output layer: The output layer is also a fully connected layer. It has several neurons, which is equal to the number of target variables. Depending on the type of output, the output layer utilizes a different type of activation. For classification, the softmax function is used (Equ.S5):

$$p_i = \frac{exp(v_i)}{\sum_{k=1}^{M} exp(v_k)} \tag{S5}$$

where p_i is the probability of the i-th output node. v_i is the weighted sum of the i-th output node, and M is the number of output nodes.
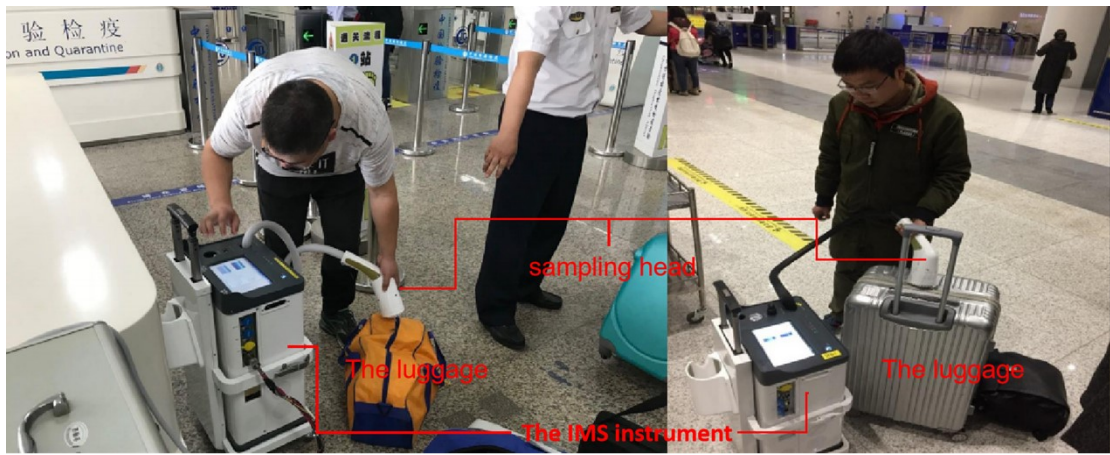
Fig. S1. Scenario of IMS data collection of agricultural quarantine materials in passenger's luggage
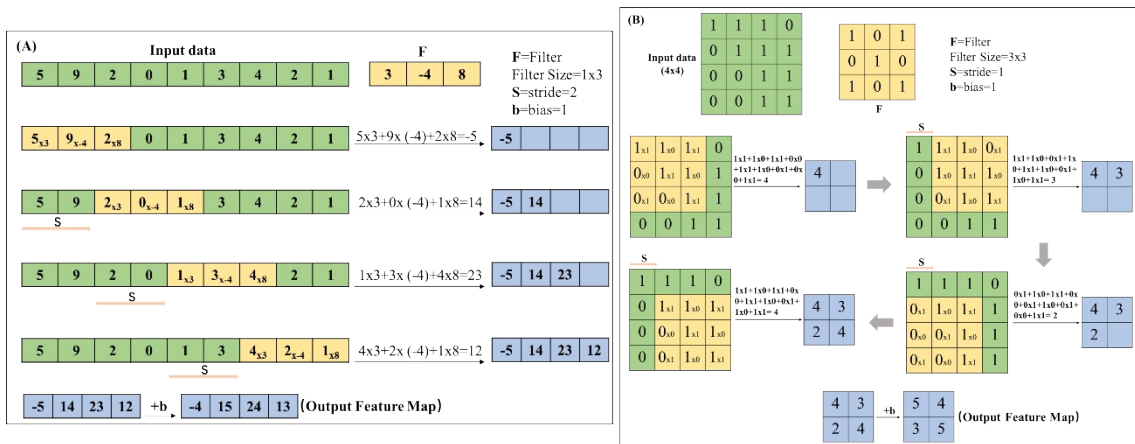

Fig. S2. Schematic of the convolution process for (A) one-dimensional data, (B) two-dimensional data.
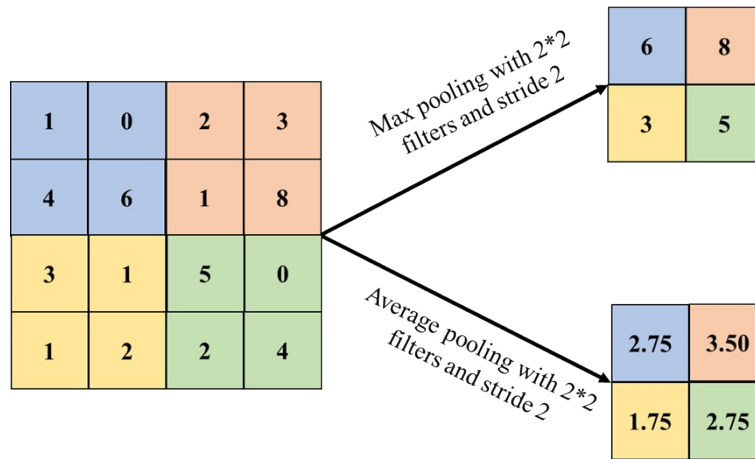
Fig. S3. Schematic of the pooling using two different methods

		Assigned class				
		1	2	3	…	k
True class	1	n_{11}	n_{12}	n_{13}	…	n_{1k}
	2	n_{21}	n_{22}	n_{23}	…	n_{2k}
	3	n_{31}	n_{32}	n_{33}	…	n_{3k}
	…	…	…	…	…	…
	k	n_{k1}	n_{k2}	n_{k3}	…	n_{kk}

Fig. S4. Schematic of confusion matrix.

Fig. S5. IMS spectra of eight principal volatile compounds which are the most characteristic or abundant compounds observed in pear, banana, apple, longan and grape from previous study.
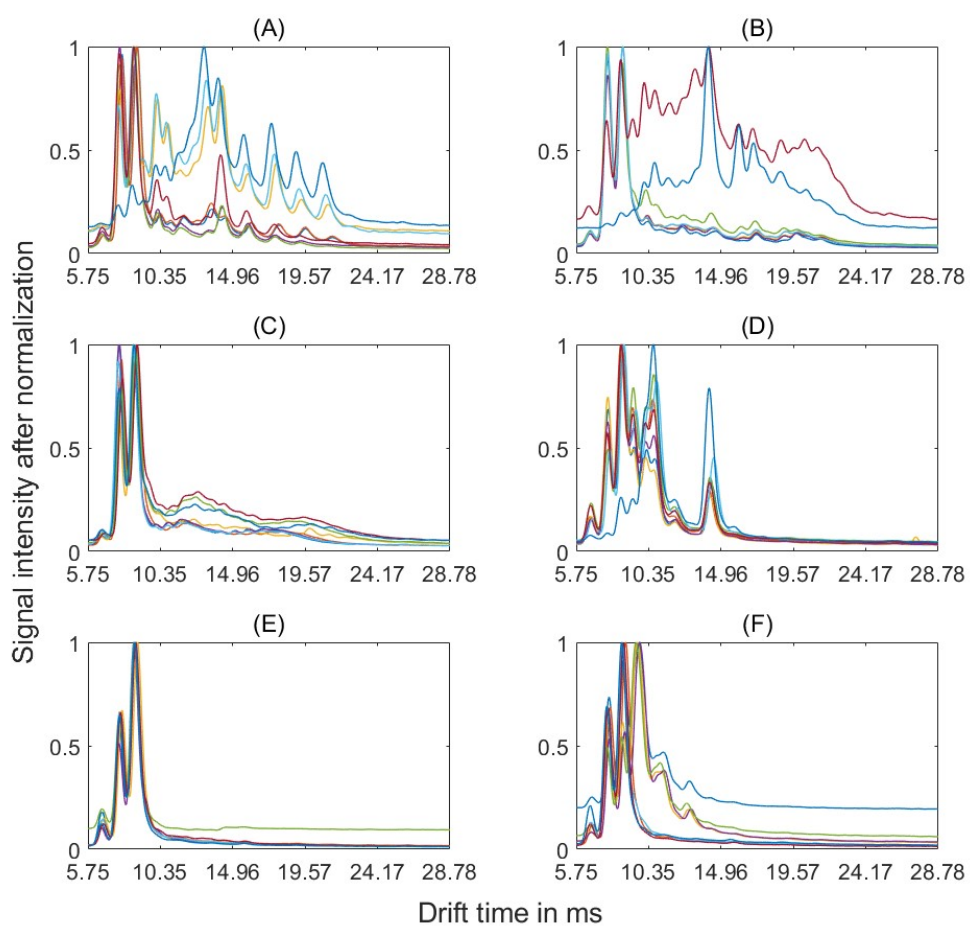
Fig. S6. IMS spectra of eight samples randomly selected from (A) pear, (B) banana, (C) apple, (D) longan, (E) blanks and (F) grape group, respectively.

Fig. S7. IMS signal evolution for different types of AQMs：(A) pear, (B) banana, (C) apple, (D) longan, (E) blanks, (F) grape in luggage.

Table S1. Classification accuracy of the four algorithms (over 20 repeated operation).

	CNN-2D		CNN-1D		PLS-DA		SIMCA	
	Calibration set	Validation set	Calibration set	Validation set	Calibration set	Validation set	Calibration set	Validation set
1	100.00%	98.28%	99.59%	90.52%	90.91%	86.21%	86.36%	77.59%
2	100.00%	94.83%	97.11%	89.66%	90.08%	85.34%	90.08%	83.62%
3	100.00%	98.28%	98.76%	92.24%	88.84%	83.62%	86.36%	81.03%
4	100.00%	96.55%	98.35%	91.37%	90.08%	85.34%	87.60%	87.93%
5	100.00%	97.41%	97.93%	90.52%	90.08%	84.48%	88.84%	85.34%
6	100.00%	98.28%	98.76%	89.66%	89.67%	86.21%	88.84%	83.62%
7	100.00%	98.28%	99.59%	90.52%	89.26%	87.93%	87.19%	82.76%
8	100.00%	98.28%	98.76%	92.24%	88.84%	88.79%	88.43%	84.48%
9	100.00%	98.28%	100.00%	89.66%	88.84%	90.52%	84.71%	84.48%
10	100.00%	98.28%	98.76%	92.24%	89.67%	86.21%	85.95%	83.62%
11	100.00%	96.55%	98.76%	92.24%	92.15%	82.76%	86.36%	76.72%
12	100.00%	98.28%	98.35%	91.37%	90.08%	85.34%	86.36%	83.62%
13	100.00%	98.28%	97.93%	90.52%	90.91%	86.21%	92.56%	87.93%
14	100.00%	98.28%	98.35%	91.37%	90.91%	86.21%	88.84%	79.31%
15	100.00%	98.28%	98.35%	91.37%	91.32%	86.21%	88.84%	84.48%
16	100.00%	98.28%	98.35%	91.37%	90.91%	84.48%	89.26%	75.86%
17	100.00%	99.14%	99.59%	90.52%	92.15%	81.90%	85.95%	77.59%
18	100.00%	98.28%	99.59%	90.52%	90.91%	86.21%	90.08%	87.07%
19	100.00%	100.00%	99.59%	90.52%	90.50%	85.34%	92.56%	82.76%
20	100.00%	99.14%	99.59%	90.52%	88.84%	88.79%	88.84%	85.34%

Reference:

1. A. Krizhevsky, I. Sutskever and G. E. Hinton, *COMMUNICATIONS OF THE ACM*, 2017, **60**, 84-90.