



Kent Academic Repository

Wang, Dayang, Wang, Lijuan and Yan, Yong (2026) *Material identification in bales using multichannel capacitive and weighing sensors with data-driven modelling. Measurement*, 277 . ISSN 0263-2241.

Downloaded from

<https://kar.kent.ac.uk/113994/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.1016/j.measurement.2026.121615>

This document version

Author's Accepted Manuscript

DOI for this version

Licence for this version

CC BY (Attribution)

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

Material identification in bales using multichannel capacitive and weighing sensors with data-driven modelling

Dayang Wang^a, Lijuan Wang^{a,*}, Yong Yan^b

^a School of Engineering, Mathematics and Physics, University of Kent, Canterbury, CT2 7NT, U.K.

^b Hangzhou International Innovation Institute, Beihang University, Hangzhou 311115, China

Corresponding author: l.wang@kent.ac.uk

Abstract

Materials are usually compressed into dense bales for easy storage and transportation in the recycling industry. A range of techniques, such as X-ray, ultrasound, microwave, and optical sensors, have been developed to identify materials distributed on a conveyor belt for sorting purposes. However, identifying materials, especially mixed materials in bales, is more challenging. Currently, there is no effective non-contact measurement system for identifying materials in bales. In this study, a new method is proposed for the first time, based on a purpose-built multichannel capacitive sensor and a weighing sensor with data-driven modelling. The capacitive sensor with three sets of transceiver units is designed for sensing the permittivity information of bales from three directions (down-up, left-right and back-front). It is optimised using the finite element method considering high sensitivity and uniform sensitivity distribution. A signal conditioning circuit is developed to amplify and demodulate the sensor outputs. The weighing sensor is used to obtain the mixture density information of the bales. Subsequently, the support vector machine (SVM) model is established to identify materials in the bales based on the multimodal sensing data. The proposed method was evaluated using model bales containing different materials, including paper, cardboard, polyvinyl chloride (PVC), and their combinations. Results demonstrate that fusing the capacitive sensor with the weighing sensor significantly outperforms using the capacitive sensor alone, validating the advantage of multimodal sensing. In addition, the SVM algorithm was compared with random forest (RF) and multilayer perceptron (MLP) algorithms, and SVM

demonstrated superior performance in data-driven modelling. To further assess general applicability, woodchip was introduced as an additional material to the original bale composition. Despite the increased complexity, the method maintained satisfactory identification performance, confirming its robustness and effectiveness as a novel solution for material identification in bales.

Keywords: Material identification; Baled materials; Multichannel capacitive sensor; Weighing sensor, Data-driven modelling

1. Introduction

Solid waste management represents a critical global issue. Approximately 2.01 billion tons of municipal solid waste are generated annually worldwide, and this figure is expected to rise to 3.4 billion tons by 2050 [1]. To advance a circular economy, reduce resource depletion, and protect the environment, recycling and reuse are the main waste-management strategies [1, 2]. However, the current recycling rate is extremely low, and only 13.5% of global waste is recycled [1]. A key issue causing the low recycling rate is uncertainty related to the quality of recycled waste materials [3]. In recycling industry, waste materials are usually compressed into dense bales for easy storage and transportation [4]. However, the material to be recycled may be mixed with other materials known as contaminations, making effective recycling difficult and reducing the recycling rate. By achieving precise identification of materials in bales, the quality of recycled waste materials can be assessed, thereby improving the recycling rate. Therefore, approaches to achieve precise identification of materials in bales are urgently needed.

Most existing methods for material identification, including X-ray, ultrasound, microwave, capacitive, and optical techniques, focus on inspecting materials spread on conveying systems for sorting purposes. They have not been specifically applied to baled materials, and each has its own set of limitations. Jumanazarov et al. [5] proposed a material classification method based on X-ray attenuation decomposition from spectral X-ray computed tomography. However, X-rays are considered environmentally hazardous due to radiation protection concerns and high energy

consumption, as well as involving substantial capital expenses and administrative complexities. Longo et al. [6] used the ultrasound method to realize identification of material properties. Ultrasound methods have limitations in measuring baled materials due to their limited penetration in high-density or thick bales, dependence on material properties, and sensitivity to environmental conditions. Harrison et al. [7] proposed a material identification method that utilizes a microwave sensor array with five resonating elements, combined with machine learning algorithms. Microwave equipment is typically expensive with complex hardware design. Potential signal attenuation in dense bales also limits its application to measure a large bale. Yunus et al. [8] introduced the capacitive sensing to image the material properties. Ahmad et al. [9] used the capacitance proximity sensor to achieve identification of paper and plastic from recycling materials. The structural design of above capacitive sensors restricts their penetration depth, rendering them inadequate for the identification of baled materials. Rozenstein et al. [10] developed an approach based on near-infrared spectroscopy which provides a non-destructive measurement method to sort black plastics. Zheng et al. [11] used the hyperspectral imaging system to characterize and sort waste plastics, which offers detailed spectral information across a wide range of wavelengths, allowing for more precise material characterization and differentiation. Penumuru et al. [12] used machine vision to extract RGB information to identify and classify flat materials being machined in a typical manufacturing environment. However, the above optical methods, including near-infrared spectroscopy, hyperspectral imaging, and machine vision, have low penetration depth, which limits their application in material identification within bales.

For baled materials, the current inspection method utilizes core sampling, in which a specialized drill is used to extract samples, followed by offline analysis of their properties. However, this method has limitations in accuracy when the extracted samples do not fully represent the material. Additionally, it is a time-consuming process to identify the material compositions. Therefore, automated, non-destructive inspection methods that can accurately identify the baled materials are desirable. The capacitive method offers advantages such as rapid response, low cost, and design

flexibility. By designing capacitive sensors with different structures, researchers have widely applied them to various measurement applications, such as phase fraction measurement, tomography imaging, and solid concentration measurement [13-22]. In our previous work, we designed a capacitive sensor with two measurement channels to identify different bales for the first time [23]. This study focused solely on identifying different bales, rather than identifying the materials within them. Moreover, the two-channel capacitive sensor can only acquire permittivity information from two directions, which limits its effectiveness in addressing the more complex challenge of material identification inside bales. To achieve accurate material identification, it is essential to develop sensor systems capable of capturing multi-dimensional information based on various physical properties of the bales. In addition, it is crucial to establish effective data-driven models to enable precise and reliable material identification based on sensor outputs.

In this study, a new method is proposed for identifying materials in bales, based on a purpose-built multichannel capacitive sensor, and a weighing sensor with data-driven modelling. A novel capacitive sensor with three sets of transceiver units is designed using the finite element method considering high sensitivity and uniform sensitivity distribution. The signal conditioning circuit for the capacitive sensor is designed. The designed capacitive sensor, combined with the weighing sensor, simultaneously acquires the permittivity and density information of the bale. Subsequently, the SVM algorithm is used as the data-driven modelling technique, due to its advantages of wide applicability, strong generalisation ability, and suitability for modelling with a small dataset [24-28]. Pre-made bales composed of different materials, including paper, cardboard, PVC, and their combinations were used to verify the advantage of multimodal sensing. The SVM algorithm was compared with RF and MLP algorithms to demonstrate its superior performance in data-driven modelling. Woodchip was introduced as an additional material to the original bale composition to further assess general applicability of the proposed method. The identification results and the advantages of the multimodal sensing and SVM based modelling are analysed and discussed.

2. Methodology

2.1. Principle of Identification

The principle of material identification in bales using a multichannel capacitive sensor and a weighing sensor with data-driven modelling is shown in Fig. 1. Bales composed of different materials exhibit variations in both permittivity and density. The multichannel capacitive sensor is to obtain the permittivity information of the bale from different directions. A weighing sensor is used to obtain the density information of the bale. The outputs of the capacitive and weighing sensors serve as input $X = (x_1, \dots, x_n, x_{n+1})$ for the data-driven model to identify the materials inside the bale. To address the identification problem with the small sample dataset, SVM algorithm is used to establish the data-driven model. There are m materials which are possibly mixed in the bales during the sorting process. The outputs of the model are $Y = (y_1, \dots, y_m)$, where y_i ($i=1, \dots, m$) represents whether material i is present in the bale. When y_i equals 1, it indicates the presence of the specific material in the bale. When y_i equals 0, it indicates the absence of the material in the bale.

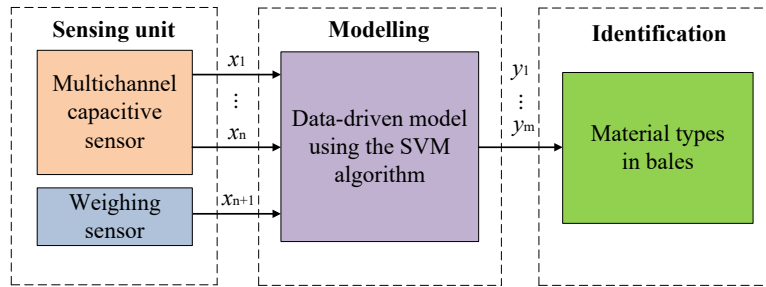


Fig. 1. Principle of materials identification in bales.

2.2. Multichannel capacitive sensor system

In this study, a commercial electronic scale is employed as the weighing sensor, and a capacitive sensor consisting of three sets of transceiver units is designed. The configuration of the capacitive sensor is shown in Fig. 2 (a). Each transceiver unit has a square transmitter unit and a square receiver unit, which form a measurement channel. Therefore, the three sets of transceiver units form three measurement channels to measure the bale from three directions. Channel 1 of the capacitive sensor

system measures the bale using the bottom and top transceiver units (from T1 to R1), channel 2 using the left and right transceiver units (from T2 to R2), while channel 3 using the back and front transceiver units (from T3 to R3). This three-channel sensor configuration enables acquisition of comprehensive information of permittivity distribution of the bale with a small number of electrodes. The transmitter unit features an exciting electrode with a side length of l_1 positioned at the front, and a ground plane with a side length of l_1 located at the back. These two components are situated on opposite sides of a rectangular substrate. The substrate itself has a thickness of 1.6 mm and a side length denoted by l_1 . The receiver unit is designed with a measuring electrode on the front, which has a side length denoted as l_2 . Alongside the measuring electrode, there is a guard electrode. The guard electrode is strategically positioned to encircle the exciting electrode, with both electrodes located on the same side of the substrate. The guard electrode surrounding the exciting electrode is used to reduce the fringe field effect. The guard electrode has a width of w , while the spacing between the exciting electrode and the guard electrode is specified as d . On the back of the receiver unit, a ground plane is present, featuring a side length l_1 . The ground planes located on the back sides of both the transmitter and receiver units combine to form a comprehensive shielding layer. This shielding layer plays a critical role in protecting the sensor from external electromagnetic interference. Since the three sets of transceiver units are identical, we only show the exciting electrode in T2, and the measuring and guard electrodes in R2 in Fig. 2(b), to clearly illustrate the structure and dimensions of the electrodes. To design a capacitive sensor with high sensitivity and uniform sensitivity distribution, parameters l_1 , l_2 , w and d should be optimised and determined.

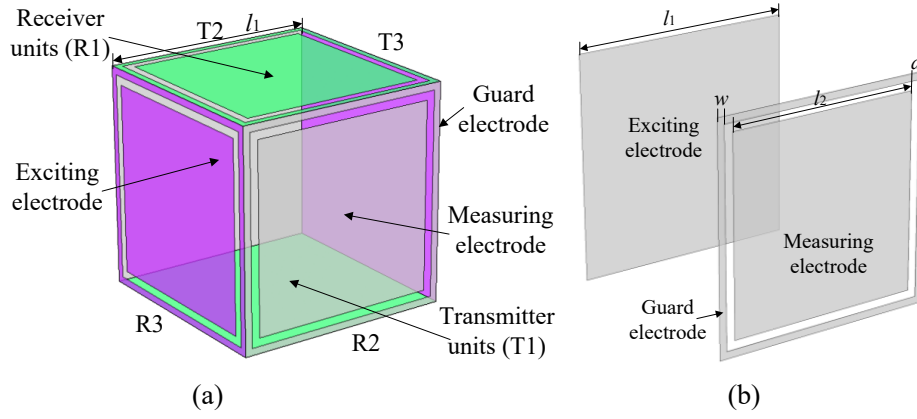


Fig. 2. (a) Structure of the capacitive sensor; (b) Structures of the exciting electrode in T2, and the measuring electrode and the guard electrode in R2.

Taking into account the dimensions of the bales to be tested in this study, the side length l_2 of the measuring electrode is set to vary between 85 mm and 110 mm, with a step size of 5 mm. The width w of the guard electrode is set between 3 mm and 13 mm, with a step size of 2 mm. Additionally, the spacing d between the exciting electrode and the guard electrode is set to range from 1 mm to 6 mm, with a step size of 1 mm. Once l_2 , w and d are optimised and determined, l_1 is determined as well. To optimise and determine these parameters, the finite element method is used. Two parameters named average sensitivity $\bar{S}$ and non-uniformity index SVP are used to evaluate the optimisation result of the sensor configuration. These two parameters $\bar{S}$ and SVP are defined as [21]

$$\bar{S} = \frac{1}{M} \sum_{k=1}^M S_{i,j}(k) \quad (1)$$

$$SVP = \frac{\left(\frac{1}{M} \sum_{k=1}^M (S_{i,j}(k) - \bar{S})^2 \right)^{\frac{1}{2}}}{\bar{S}} \quad (2)$$

where $S_{i,j}(k)$ is the sensitivity of the k th element at point (x, y, z) in region $P(x, y, z)$ and M is the number of the elements. $S_{i,j}(k)$ is defined as [29]

$$S_{i,j}(k) = - \int_{P(x,y,z)} \frac{E_i(x, y, z) E_j(x, y, z)}{U_0^2} dx dy dz \quad (3)$$

where $E_i(x, y, z)$ is the electrical field strength when exciting electrode group i is

excited by voltage U_0 , and $E_j(x, y, z)$ is the electrical field strength when measuring electrode group j is excited by voltage U_0 . In general, a larger $\bar{S}$ means that the sensitivity of the sensor is higher, while a smaller SVP indicates that the sensitivity distribution is more uniform.

In this study, the parameters l_2 , w and d are optimised and determined through orthogonal experiments. Table 1 shows the three structure parameters and their corresponding value ranges, represented using levels A, B, C and D. To optimize three parameters at four levels each, an $L_{16}(4^3)$ orthogonal array is employed. This orthogonal array consists of 16 configurations that ensure balanced, independent, and efficient evaluation of multiple factors using the minimum number of experiments. For each configuration, the values of the three structure parameters are listed in the Table 2. For example, in configuration 1, the values of l_2 , w and d are all set to level A which are 85, 3, and 1, respectively. For all configurations, simulations are performed using the finite element method based on the corresponding values of l_2 , w and d . The $\bar{S}$ and SVP are calculated and recorded in Table 2.

Table 1
Levels of three structure parameters.

Levels	Parameters		
	l_2 (mm)	w (mm)	d (mm)
A	85	3	1
B	90	5	2
C	95	7	3
D	100	9	4

Based on the calculated $\bar{S}$ and SVP under all configurations shown in Table 2, the average of $\bar{S}$ with different parameters of l_2 , w , and d of a certain level is indicated as S_i ($i = A, B, C, D$) and calculated. The V_i ($i = A, B, C, D$) which are the average of SVP with different parameters of l_2 , w , and d of a certain level are also calculated. For the parameter l_2 , its S_A and V_A are obtained by calculating the average of $\bar{S}$ and SVP , respectively, across the four configurations in which l_2 is set to level A in Table 2. The S_i and V_i values of parameter l_2 , w , and d at levels A, B, C, and D can be obtained

using the same calculation method described above, and the results are shown in Table 3. It can be seen in Table 3 that for the parameter l_2 , S_A is larger than S_B , S_C , and S_D , while V_A is smaller than V_B , V_C , and V_D . This means that when l_2 is set to the value of level A, higher sensitivity can be achieved, while when l_2 is set to the value of level D, a more uniform sensitivity distribution can be obtained. For the parameter w and d , S_A is larger than S_B , S_C , and S_D , while V_A is smaller than V_B , V_C , and V_D . This means that when w and d are set to the values of level A, higher sensitivity can be achieved, while when w and d are set to the values of level D, a more uniform sensitivity distribution can be obtained. Since sensitivity is the primary consideration in this study, the final values of l_2 , w , and d are set to level A, which are 85 mm, 3 mm, and 1 mm, respectively. Based on the determined parameters, transmitting and receiving units were fabricated using printed circuit boards (PCBs). For the PCBs of the transmitting and receiving units, their insulating substrates are made of FR4 epoxy resin plate with a thickness of 1.6 mm and a side length l_1 of 93 mm. The conductive copper layer attached to the insulating substrates forms the electrodes and ground planes. For the transmitting unit, the exciting electrode has a side length l_1 of 93 mm and is laid on one side of the insulating substrate, while the ground plane, also with a side length l_1 of 93 mm, is laid on the other side of the insulating substrate. For the receiving unit, the measuring electrode and the guard electrode are laid on one side of the insulating substrate. The side length l_2 of the measuring electrode is 85 mm, the width w of the guard electrode is 3 mm, and the spacing d between the measuring electrode and the guard electrode is 1 mm. The other side of the insulating substrate is a ground plane with a side length l_1 of 93 mm. The transmitting unit, the receiving unit and the formed capacitive sensor are shown in Fig. 3.

Table 2

$L_{16}(4^3)$ Orthogonal array and simulation results.

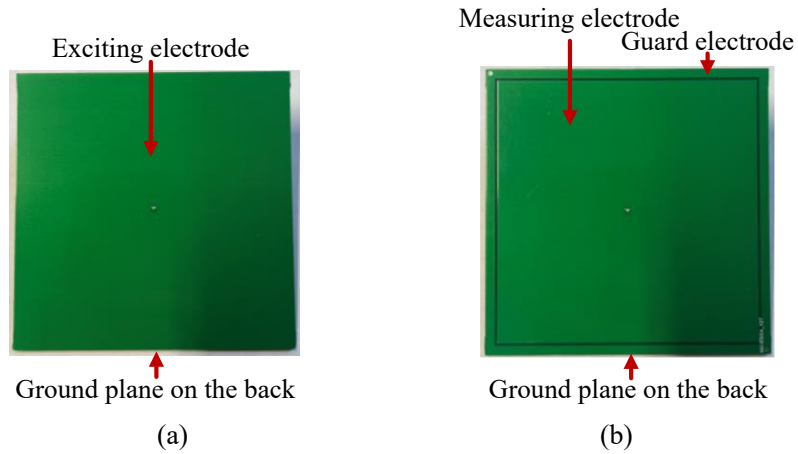
Configuration	Parameters levels			$\bar{S}$	SVP
	l_2	w	d		
1	A	A	A	634.74	2.67
2	A	B	B	385.03	2.41
3	A	C	C	270.28	2.29

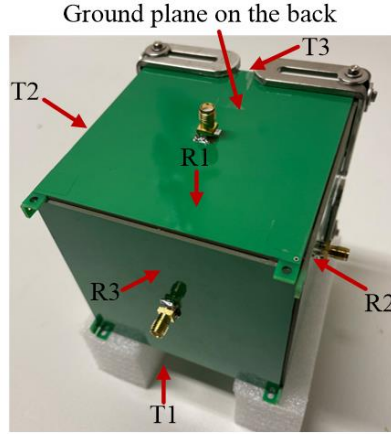
4	A	D	D	203.92	2.21
5	B	A	B	371.18	2.40
6	B	B	A	327.35	2.35
7	B	C	D	203.03	2.21
8	B	D	C	185.89	2.19
9	C	A	C	257.99	2.26
10	C	B	D	199.27	2.20
11	C	C	A	214.06	2.22
12	C	D	B	169.81	2.17
13	D	A	D	194.89	2.19
14	D	B	C	180.57	2.17
15	D	C	B	167.64	2.16
16	D	D	A	156.11	2.15

Table 3

Averages of $\bar{S}$ and SVP with different parameters.

Values	Parameters		
	l_2	w	d
S_A	373.49	364.70	333.07
V_A	2.40	2.38	2.35
S_B	271.86	273.06	273.42
V_B	2.29	2.28	2.29
S_C	210.28	213.75	223.68
V_C	2.21	2.22	2.23
S_D	174.80	178.93	200.28
V_D	2.17	2.18	2.20





(c)

Fig. 3. (a) Transmitter unit; (b) Receiver unit; (c) Capacitive sensor.

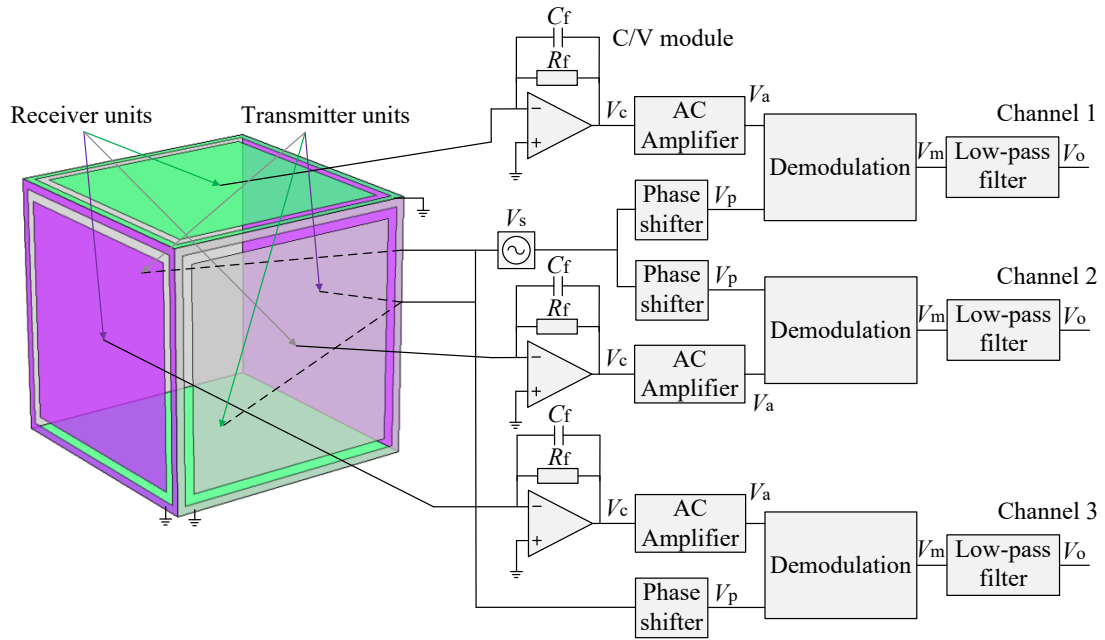


Fig. 4. Circuit diagram of capacitive sensor system with three measurement channels.

Considering that the capacitance to be measured between the electrodes are usually small, an AC-based capacitance measuring circuit is designed to achieve a fast response, high signal-to-noise ratio, and immunity to stray capacitance. Fig. 4 shows the circuit diagram of the capacitive sensor system. The sinusoidal signal V_s with a frequency of 200 kHz and a peak-to-peak voltage of 8 V is generated by the signal source and applied to the exciting electrodes of the transmitter units. The measuring electrodes of the receiver units are connected to the capacitance-to-voltage (C/V)

module, in which the C_f and R_f are set to 2 pF and 51 k Ω , respectively. The resulting voltage V_c is

$$V_c = -V_s \left(\frac{j\omega C_m R_f}{1 + j\omega C_f R_f} \right) \quad (4)$$

By using the C/V module, the measured capacitance C_m is converted into a voltage V_c . The voltage V_c is amplified by an amplifier circuit built around an LF356 and its output V_a is then fed into the demodulation module, which uses an AD734 multiplier. To demodulate the signal, the signal V_p from the phase shifter with the phase adjusted to match that of V_a is fed into the demodulation module. The output V_m of the demodulation module is

$$\begin{aligned} V_m &= V_a V_p = A \sin(\omega t + \varphi_1) B \sin(\omega t + \varphi_2) \\ &= \frac{1}{2} AB [\cos(\varphi_1 - \varphi_2) - \cos(2\omega t + \varphi_1 + \varphi_2)] \end{aligned} \quad (5)$$

The output V_m is filtered by a low-pass filter to remove the high-frequency components from V_m . The remaining part V_o is

$$V_o = \frac{1}{2} AB \cos(\varphi_1 - \varphi_2) \quad (6)$$

The output V_o reflects the measured capacitance which mainly depends on the permittivity of the materials under test. Therefore, the output V_o reflects the permittivity of the materials.

2.3. Data-driven modelling

The SVM is employed to map the complex relationship between the sensor outputs and the material types within the bale. It is a supervised machine learning algorithm used for classification tasks, aiming to find the optimal hyperplane that best separates data points belonging to different classes in the feature space. In this study, the inputs of the model are $X = (x_1, x_2, x_3, x_4)$, where x_1 , x_2 , and x_3 represent the capacitive sensor outputs and x_4 represents the weighing sensor output. The outputs y_i ($i = 1, \dots, m$) of the model indicate the presence or absence of the material y_i . Since this is a multi-label classification problem and SVM is inherently a binary classifier, we decompose the

multi-label classification task into multiple independent binary classification problems. For each label y_i , we train an independent SVM binary classifier SVM_i ,

$$SVM_i; X \rightarrow y_i \quad i=1, \dots, m \quad (7)$$

The structure of a SVM binary classifier is shown in Fig. 5, where $K(X, X_i)$ ($i=1, \dots, L$) is the kernel inner product, u_i ($i=1, \dots, L$) represents the Lagrange multiplier scaled by the corresponding class label, and b is the bias term. The overall multi-label classification is formed by combining the outputs of all individual classifiers:

$$Y = (y_1, \dots, y_m) \quad (8)$$

In the process of using SVM, it is essential to determine the optimal kernel function and its associated hyperparameters to build the most effective SVM model. To find the optimal combination of hyperparameters, grid search with cross-validation is used. This approach systematically evaluates all possible configurations within a predefined search space, using k -fold cross-validation to assess model performance. It helps ensure that the selected hyperparameters promote the model's stability, accuracy, and generalisation ability. In this study, the above approach is applied based on the experimental data to determine the optimal hyperparameter configuration.

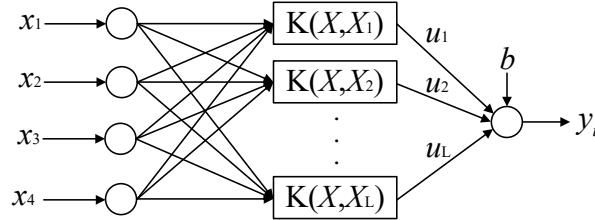


Fig. 5. Structure of a SVM binary classifier.

3. Experimental tests

In this study, all experiments were conducted under controlled laboratory conditions, where environmental temperature and humidity fluctuations were minimal. The model bales with different materials including paper, cardboard, PVC, woodchip and their combinations are made for the experiments by filling the materials into a thin-walled plastic boxes with side length of 80 mm. The paper, cardboard, and PVC materials were placed layer by layer in thin-walled plastic boxes to form bales, while

the woodchip was placed in thin-walled plastic boxes in the form of small chips. Fig. 6 presents eleven typical combinations of materials. By varying the types of materials and their volume fractions, different bales are prepared for experiment. Each bale has 6 faces with the bottom face t1, the top face r1, the left face t2, the right face r2, the back face t3, and the front face r3. The designed capacitive sensor system is shown in Fig. 7 (a), which has three measurement channels. When the bale is placed within the capacitive sensor, the capacitance between the exciting electrode and measuring electrode of each channel is converted into a voltage signal by the hardware. The voltage outputs of the capacitive sensor reflect the permittivity of the bale measured by three different channels. The weighing sensor, as shown in Fig. 7(b), measures the mass of the bale which reflects the mixture density of the bale.

In this study, a total of twenty-one distinct model bales with different volume fractions of materials as listed in Table 4 were measured using the sensor system. For each bale, since it has six faces and the capacitive sensor has three measurement channels, a total of 24 relative positions is formed depending on the placement of the bale within the capacitive sensor. Therefore, a total of 504 different conditions (21×24) is considered. The experimental data is split into 80% for training data-driven models and 20% for testing. To further evaluate the performance of the data-driven models under unseen conditions, four new bales were prepared, and the volume fractions of each material in these bales are presented in Table 5. Since there are 24 relative positions when a bale is placed in the capacitive sensor, a total of 144 different conditions (6×24) are used to validate the generalisation ability of the data-driven model. In this study, we first evaluate the performance of the proposed method, including the advantage of multimodal sensing and the superior performance of SVM in data-driven modelling, using model bales containing paper, cardboard, PVC, and their combinations. This involves bales 1–13 in Table 4, resulting in a total of 312 different conditions (13×24) used for training and testing. Bales 1–4 in Table 5 provide a total of 96 different conditions (4×24), which are used to validate the generalisation ability of the data-driven model. To further assess general applicability, woodchip was introduced as an additional material to the original bale composition.

Model bales containing paper, cardboard, polyvinyl chloride (PVC), woodchip, and their combinations were used, as shown in Table 4, resulting in a total of 504 different conditions (21×24) used for training and testing. All bales in Table 5 provide a total of 144 different conditions (6×24) to validate the generalisation ability of the data-driven model. By comparing with the results obtained without woodchip, the general applicability of the proposed method in handling new scenarios is demonstrated.

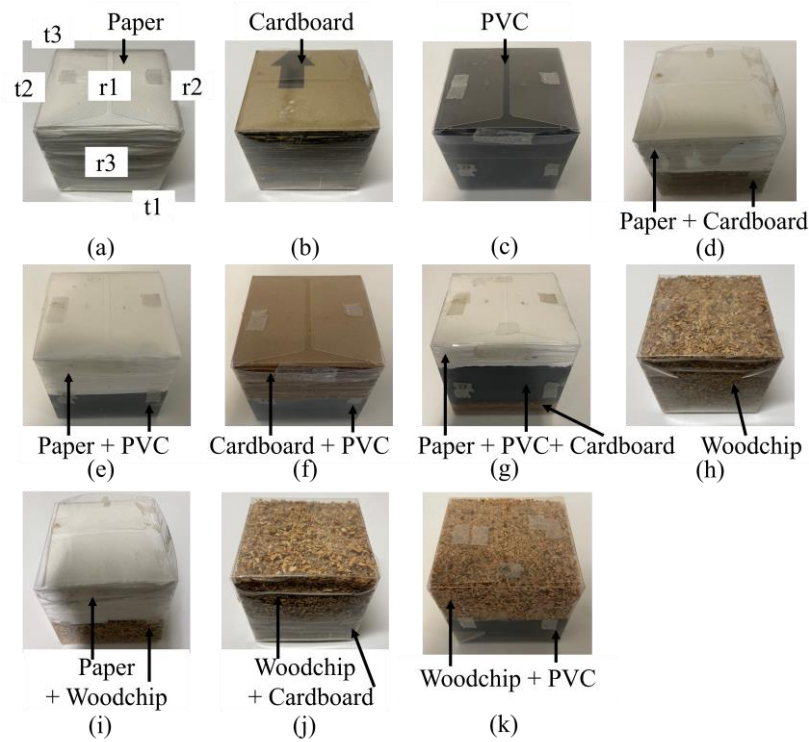


Fig. 6. Typical bale material combinations. (a) paper only; (b) cardboard only; (c) PVC only; (d) paper and cardboard; (e) paper and PVC; (f) cardboard and PVC; (g) paper, PVC and cardboard; (h) woodchip only; (i) paper and woodchip; (j) woodchip and cardboard; (k) woodchip and PVC.

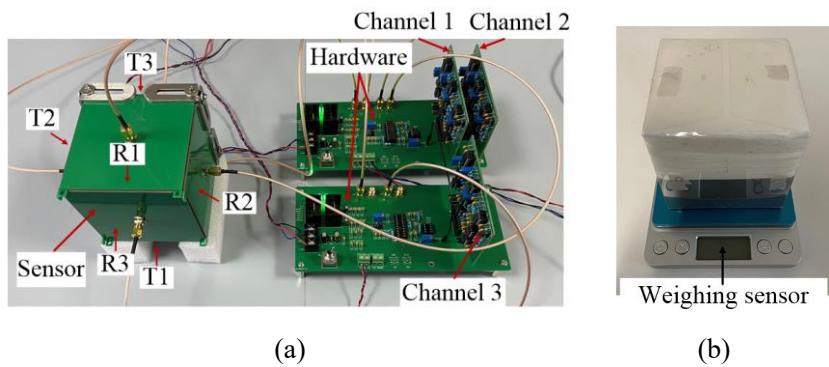


Fig. 7. (a) Capacitive sensor system; (b) Weighing sensor.

Table 4

Material composition of model bales used for training and testing

Bale No.	Material volume fraction			
	Paper	Cardboard	PVC	Woodchip
1	100%	0%	0%	0%
2	0%	100%	0%	0%
3	0%	0%	100%	0%
4	50%	50%	0%	0%
5	50%	0	50%	0%
6	0	50%	50%	0%
7	25%	75%	0%	0%
8	25%	0%	75%	0%
9	0%	25%	75%	0%
10	75%	25%	0%	0%
11	75%	0%	25%	0%
12	0%	75%	25%	0%
13	25%	25%	50%	0%
14	0%	0%	0%	100%
15	50%	0%	0%	50%
16	0%	50%	0%	50%
17	0%	0%	50%	50%
18	75%	0%	0%	25%
19	0%	75%	0%	25%
20	0%	0%	75%	25%
21	25%	0%	0%	75%

Table 5

Material composition of model bales used for generalisation testing

Bale No.	Material volume fraction			
	Paper	Cardboard	PVC	Woodchip
1	37.5%	62.5%	0%	0%

2	62.5%	37.5%	0%	0%
3	25%	50%	25%	0%
4	50%	25%	25%	0%
5	0%	25%	0%	75%
6	0%	0%	25%	75%

4. Results and discussion

4.1 Outputs of capacitive sensor and weighing sensor

Fig. 8 illustrates the three-channel voltage outputs of the capacitive sensor with typical bales under test, respectively. These bales include one-material bales and bales containing different materials. The sensor output without a bale is also shown in Fig. 8 to show the baseline value of the sensor system. Fig. 8 shows the three-channel voltage outputs when the capacitive sensor measures these bales from t1 to r1 by channel 1, from t2 to r2 by channel 2, and from t3 to r3 by channel 3.

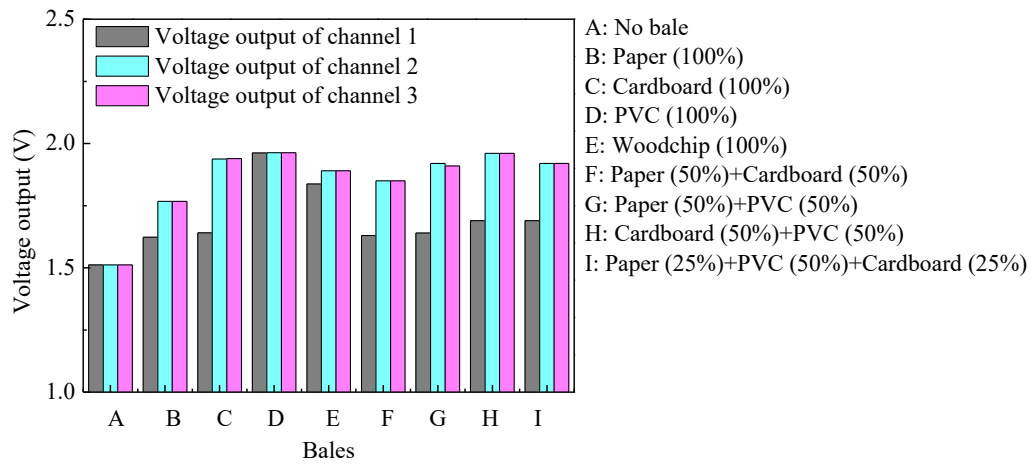


Fig. 8. Three-channel voltage outputs of the capacitive sensor when measuring different bales.

Even without a bale, the sensor system still produces a baseline output, which is determined by the designed measurement system. The sensor outputs for all bales are higher than these baseline values. For the pure paper bale, the capacitive sensor exhibits lower voltage output due to its low permittivity. In contrast, the PVC bale produces the highest voltage output, attributed to its high permittivity. These results

indicate that the sensor outputs effectively reflect the differences in material permittivity. For the PVC bale, the voltage outputs of three channels are almost the same, which is due to fact that the distribution of the PVC material is relatively uniform. Therefore, the three-channel voltage outputs are similar, even though they are measured from three different directions. However, for the paper bale, cardboard bale, and the bale made of mixed materials, the three-channel voltage outputs are different. This is because the distribution of these materials is relatively uneven, leading to an uneven distribution of permittivity. The three-channel voltage outputs contain the information of permittivity distribution in the bale. The three-channel voltage outputs are also influenced by the relative position between the capacitive sensor and the bale, because changing the relative position alters the distribution of the permittivity between the excitation electrode and the measuring electrode. Therefore, in this study, all 24 possible relative positions between the capacitive sensor and each bale were measured, covering all possible scenarios and providing comprehensive data for training and testing the model.

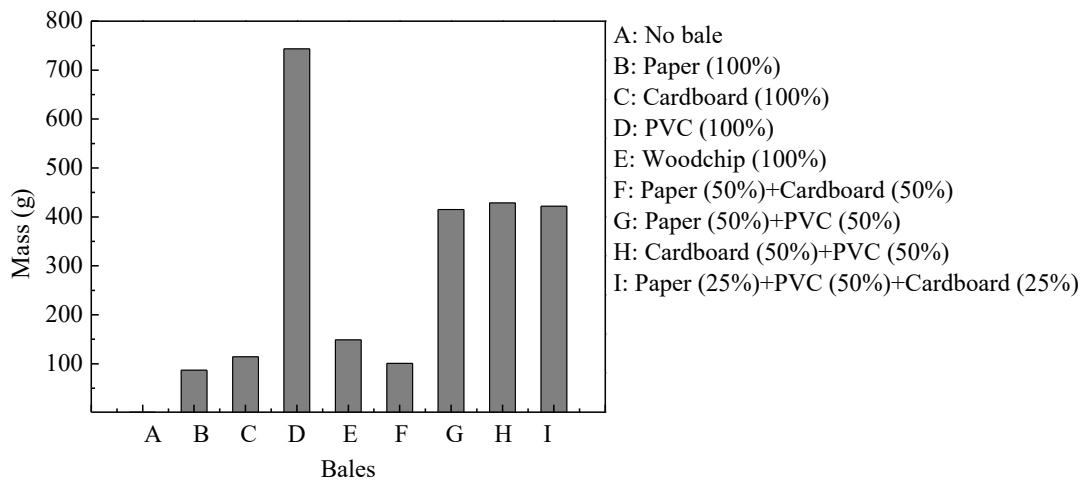


Fig. 9. Outputs of the weighing sensor when measuring different bales.

The seven typical bales are also measured by the weighing sensor and its outputs are presented in Fig. 9. When no bale was present, the output of the weighing sensor was zero. This is attributed to the zero-setting functions of the commercial electronic scale. For the paper bale, its mass is the lowest because of its smallest mixture density. For the PVC bale, its mass is highest because of its largest mixture density. For the

bale contains mixture, the output lies between the outputs of the individual pure materials, which indicates the effectiveness of the weighing sensor. The weighing sensor captures materials information within the bale from the perspective of density characteristics, and its output is solely determined by the mixture density of the bale.

4.2 Data-Driven Modelling and Material Identification

To evaluate the performance of the proposed method, including the advantage of multimodal sensing and the superior performance of SVM in data-driven modelling, the model bales containing paper, cardboard, PVC, and their combinations are used. The three-channel voltage outputs of the capacitive sensor and the output of the weighing sensor are used as the input data for data-driven modelling, while the types of materials are the outputs. For the experimental data obtained from the bales 1-13 as shown Table 4, 80% are used for model training, while the remaining 20% are used for testing. In the SVM modelling process, the RBF kernel function is chosen due to the better performance compared with other kernel functions. The optimal values for the regularization parameter C and the kernel coefficient γ , are determined through 5-fold cross-validation. The average accuracy across all labels for each possible combination of C and γ is calculated. The combination that maximizes this average is selected as the optimal parameter C and the kernel coefficient γ , which are 10 and 10, respectively.

The parameters named accuracy, precision, recall, F1-score, support and prediction time (T) are used to evaluate the performance of the method. The accuracy is the ratio of correct predictions to total predictions. The precision is the proportion of correctly predicted positives among all predicted positives, reflecting the accuracy of positive predictions. The recall is the proportion of correctly predicted positives among all actual positives, showing the ability of model to detect positives. The F1-score is the harmonic mean of precision and recall. The support refers to the number of actual instances for each class in the dataset. The definitions of accuracy, precision, recall, and F1-score are

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (10)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (11)$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (12)$$

In the above formulas, TP (True Positive) refers to the number of instances correctly predicted as positive, and TN (True Negative) refers to the number of instances correctly predicted as negative. FP (False Positive) refers to the number of negative instances incorrectly predicted as positive, and FN (False Negative) refers to the number of positive instances incorrectly predicted as negative.

To compare the performance of the multimodal sensing method using the capacitive and weighing sensors, with the method only based on the capacitive sensor, two SVM models are developed. One takes the three-channel outputs of the capacitive sensor as its inputs, while the other one accepts both the three-channel capacitive outputs and the output of the weighing sensor as its inputs. Fig. 10 shows the confusion matrices of the three-input SVM model, while Fig. 11 shows the confusion matrices of the four-input SVM model. The accuracy, precision, recall, F1-score, and support for each material are calculated and summarised in Tables 6 and 7, respectively. Fig.10 and Table 6 show the classification results of the three-input SVM model based on the capacitive sensor. For each material, the model misclassified 3 out of 63 for paper, 7 out of 63 for cardboard, and 5 out of 63 for PVC, resulting in the accuracy of 0.95, 0.89 and 0.92, respectively. In general, the model performs well for material identification, especially for paper identification which achieves higher accuracy compared with cardboard and PVC. This is due to the fact that the capacitance values of the paper bale are quite different from that of cardboard and PVC (see Fig. 8). By combining the outputs of the three-channel capacitive sensor with the weighing sensor, the four-input SVM model is developed. The performance of the four-input SVM model is summarised in Fig. 11 and Table 7. The misclassified cases for paper, cardboard and PVC have reduced to 1, 1 and 0,

respectively, and hence the accuracy has increased to 0.98, 0.98 and 1, respectively. This demonstrates that the weighing sensor provides useful and complementary information for material identification within bales. The prediction times of the three-input and four-input SVM models are 0.00186 s and 0.00199 s, respectively, indicating that both models achieve high computational efficiency and are suitable for real-time applications.

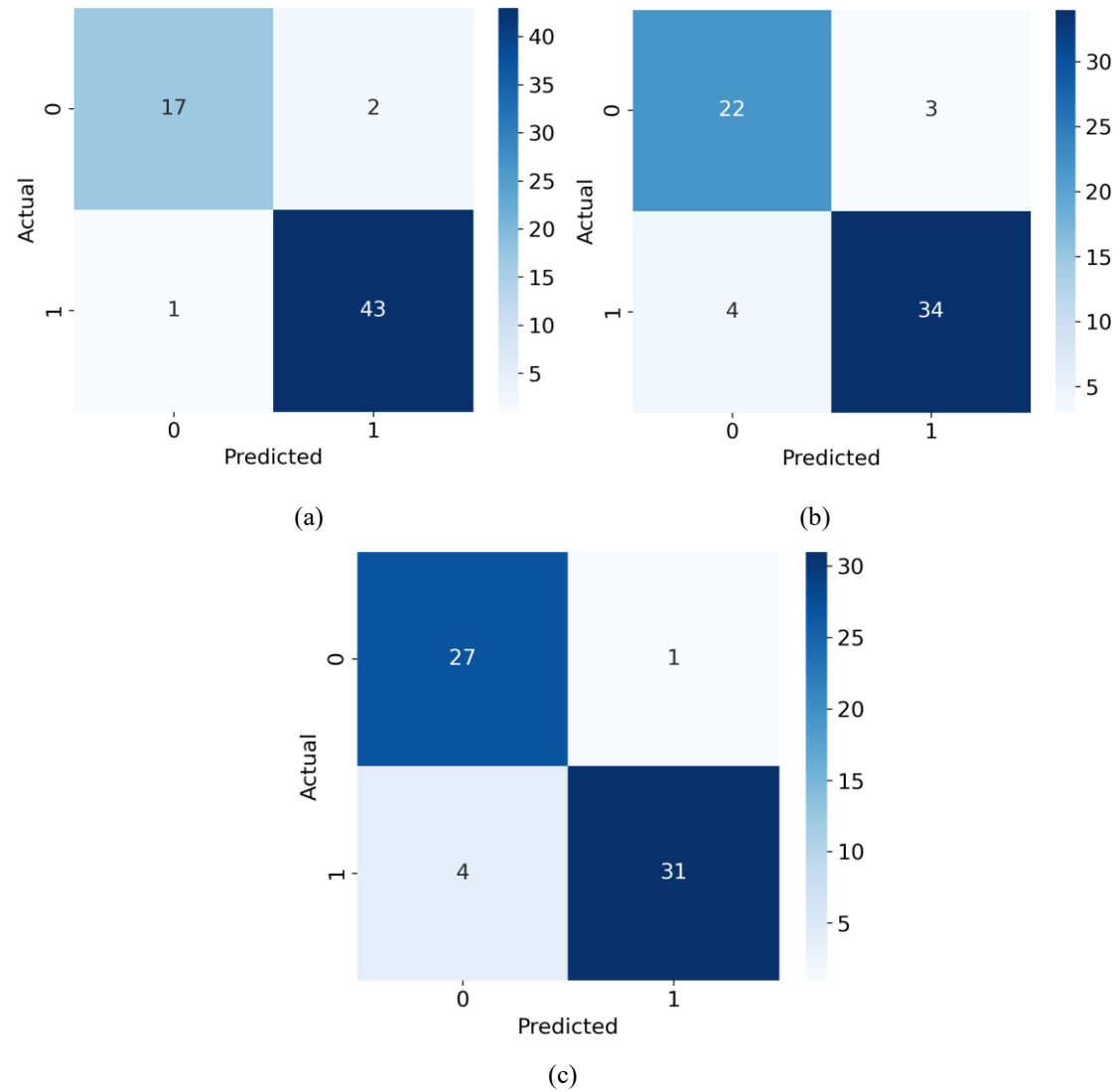


Fig. 10. Confusion matrices of three-input SVM model: (a) y_1 (Paper); (b) y_2 (Cardboard); (c) y_3 (PVC).

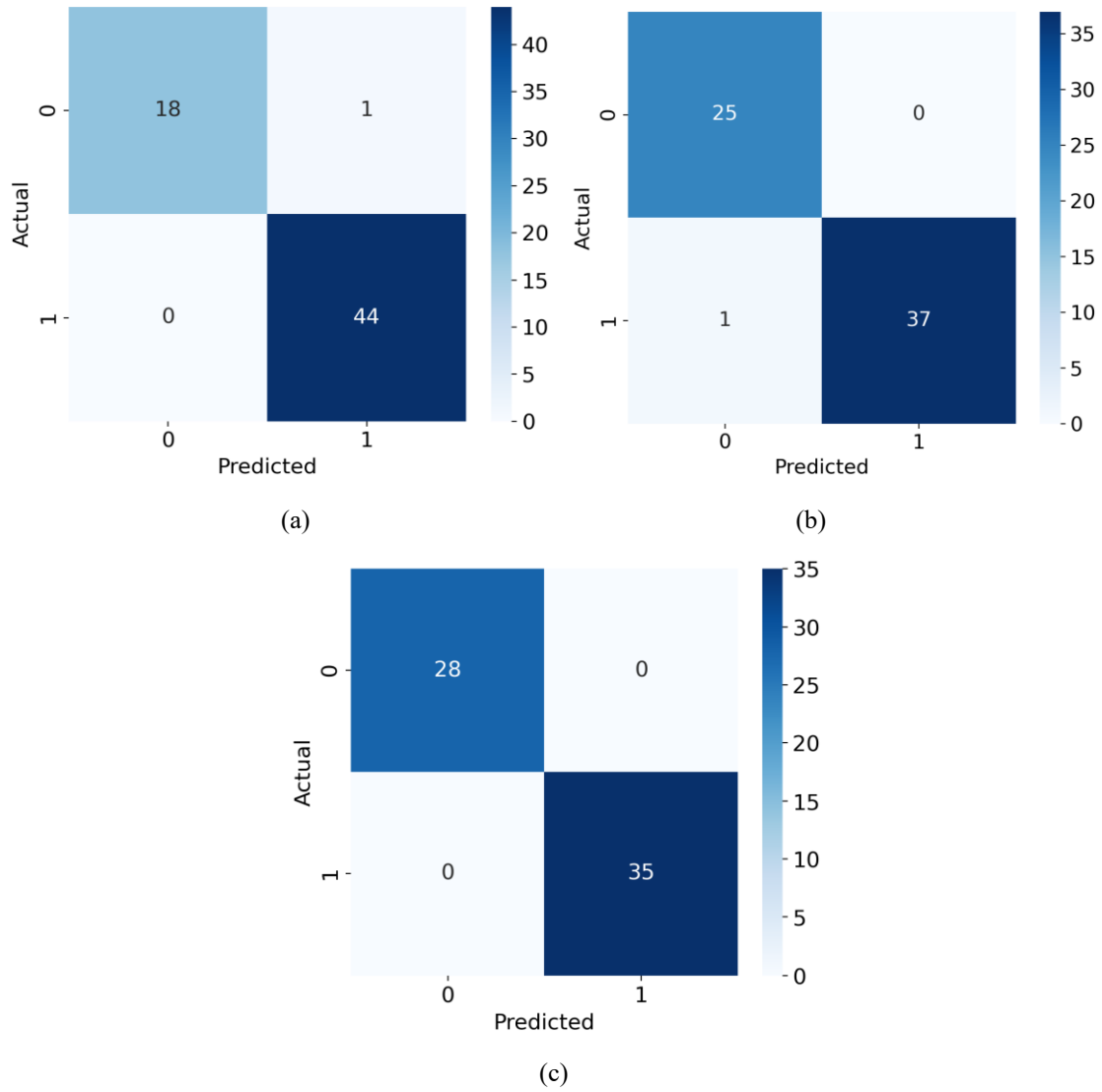


Fig. 11. Confusion matrices of the four-input SVM model: (a) y_1 (Paper); (b) y_2 (Cardboard); (c) y_3 (PVC).

Table 6

Classification results of the three-input SVM model.

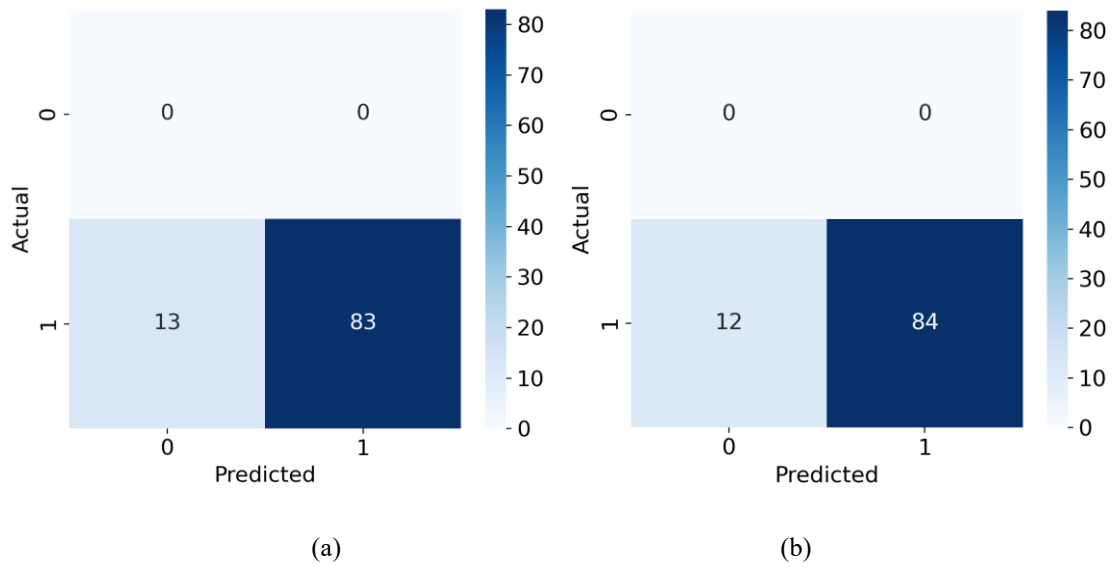
Type of materials	Evaluation metrics				
	Accuracy	Precision	Recall	F1-score	Support
Paper	0.95	0.96	0.98	0.97	44
Cardboard	0.89	0.92	0.89	0.91	38
PVC	0.92	0.97	0.89	0.93	35

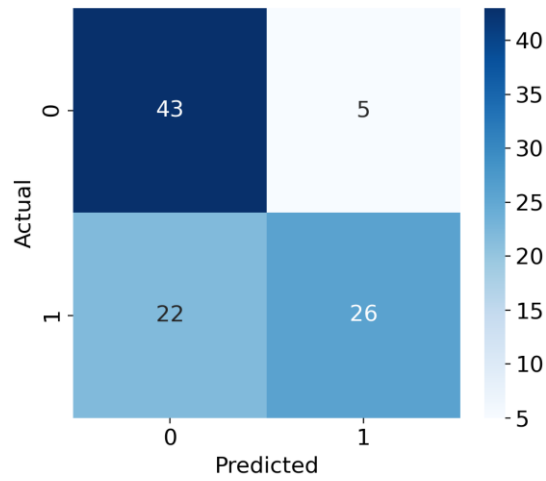
Table 7

Classification results of the four-input SVM model.

Type of materials	Evaluation metrics				
	Accuracy	Precision	Recall	F1-score	Support
Paper	0.98	0.98	1	0.99	44
Cardboard	0.98	1	0.97	0.99	38

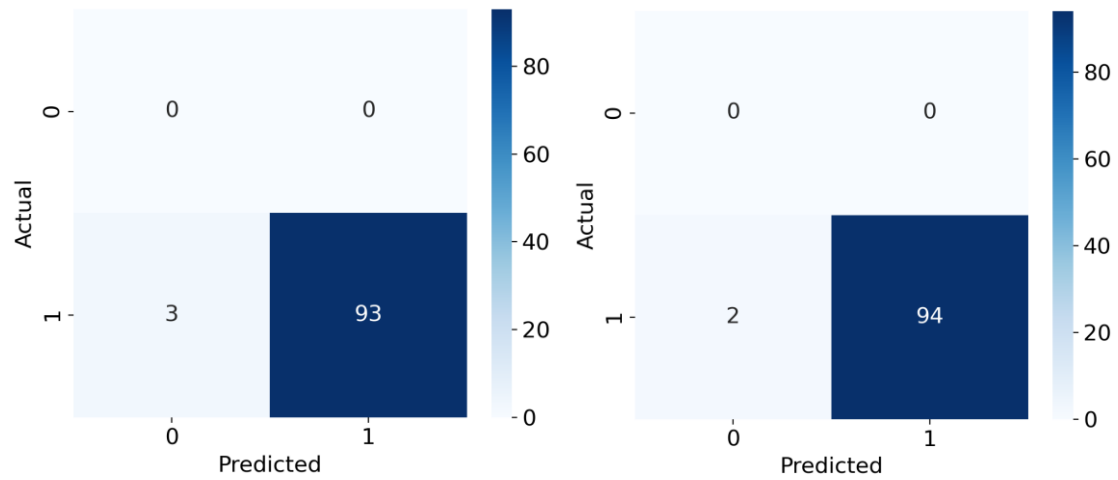
To further evaluate the performance of the three-input and four-input SVM models under unseen conditions, four new model bales (1-4) with different material combinations as described in Table 5, are used for generalisation test. Fig. 12 shows the confusion matrices of the three-input SVM model based on the capacitive sensor. More misclassification cases occur compared with Fig. 10. Accordingly, the accuracy of the model has reduced to 0.86, 0.88 and 0.72 for paper, cardboard and PVC, respectively. The limited generalisation ability is the common problem for data-driven models. By adding the weight information as addition input to the SVM model, the accuracy as shown in Table 9 has improved to 0.97, 0.98 and 0.99 for identification of paper, cardboard and PVC, respectively. The prediction times of the three-input and four-input SVM models are 0.00289 s and 0.00306 s, respectively, indicating high computational efficiency. Even through the accuracy of the four-input SVM for the unseen dataset is slightly lower than the result in Table 7, it significantly outperforms the three-input SVM. This can be attributed to the fact that the capacitive and weighing sensors provide multi-dimensional information including permittivity and density of the materials in bales. Results demonstrate that using multimodal sensing significantly outperforms using the capacitive sensor alone, achieving excellent performance in terms of success rate and generalisation capability.





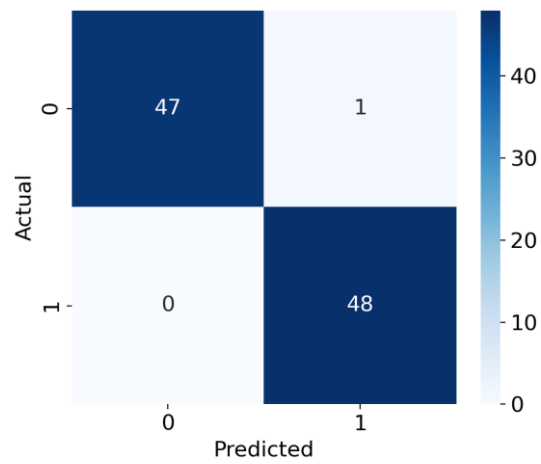
(c)

Fig. 12. Confusion matrices of the three-input SVM model for generalisation test: (a) y_1 (Paper); (b) y_2 (Cardboard); (c) y_3 (PVC).



(a)

(b)



(c)

Fig. 13. Confusion matrices of the four-input SVM model for generalisation test: (a) y_1 (Paper); (b) y_2 (Cardboard); (c) y_3 (PVC).

Table 8

Classification results of the three-input SVM model for generalisation test.

Type of materials	Evaluation metrics				
	Accuracy	Precision	Recall	F1-score	Support
Paper	0.86	1	0.86	0.93	96
Cardboard	0.88	1	0.88	0.93	96
PVC	0.72	0.84	0.54	0.66	48

Table 9

Classification results of the four-input SVM model for generalisation test.

Type of materials	Evaluation metrics				
	Accuracy	Precision	Recall	F1-score	Support
Paper	0.97	1	0.97	0.98	96
Cardboard	0.98	1	0.98	0.99	96
PVC	0.99	0.98	1	0.99	48

In this study, the superior performance of SVM in data-driven modelling compared with other typical models is also verified. We applied the random forest (RF) and multilayer perceptron (MLP) algorithms to the experimental data obtained from bales 1–13 in Table 4 and to the experimental data obtained from bales 1–4 in Table 5. Their performance was compared with the SVM results. Micro Precision, Micro Recall, Micro F1-score and prediction time (T) are used as overall evaluation metrics to evaluate the performance of the data-driven models because they reflect the model's overall performance across all label predictions. These metrics reflect the model's overall performance across all sample-label pairs and are intuitive and stable.

Table 10

Classification results of the three models.

Type of models	Evaluation metrics			
	Micro Precision	Micro Recall	Micro F1-score	T (s)
SVM	0.99	0.99	0.99	0.00199
RF	0.99	0.98	0.99	0.00717
MLP	0.95	0.98	0.97	0.00296

Table 11

Classification results of the three models for generalisation test.

Type of models	Evaluation metrics			
	Micro Precision	Micro Recall	Micro F1-score	T (s)

SVM	1.00	0.98	0.99	0.00306
RF	1.00	0.88	0.93	0.00894
MLP	1.00	0.95	0.97	0.00453

Table 10 presents the results when 80% of the dataset from bales 1–13 in Table 4 is used for model training and the remaining 20% is used for testing. It can be observed that SVM outperforms RF and MLP across all metrics, demonstrating its superior performance. Table 11 presents the results when the bales 1-4 from Table 5 are used for the generalization test. It can be seen that SVM still achieves strong performance on unseen bales. Compared with RF and MLP, SVM exhibits strong generalisation ability. These results demonstrate the superior performance of SVM in data-driven modelling in this study.

Through the identification of paper, cardboard, and PVC materials, it has been demonstrated that the effective identification of materials can be achieved by combining the three-channel output of a capacitive sensor with the output of a weighing sensor, together with a support vector machine (SVM) data-driven algorithm. To enhance the general applicability of the conclusion, and considering that woodchip is also a type of recyclable material, bales containing woodchip were additionally included, as shown in bales 14-21 in Table 4 and bales 5-6 in Table 5. By applying the proposed method to these bales and evaluating its performance, it is further demonstrated that the method has a wide range of applicability.

Table 12

Classification results of the four-input SVM model.

Type of materials	Evaluation metrics				
	Accuracy	Precision	Recall	F1-score	Support
Paper	1	1	1	1	49
Cardboard	0.93	1	0.85	0.92	48
PVC	1	1	1	1	43
Woodchip	0.94	0.88	1	0.94	44

Table 13

Classification results of the four-input SVM model for generalisation test.

Type of materials	Evaluation metrics				
	Accuracy	Precision	Recall	F1-score	Support
Paper	0.99	0.99	1	0.99	96
Cardboard	0.90	0.98	0.9	0.94	120

PVC	1	1	1	1	72
Woodchip	0.89	0.79	0.92	0.85	48

First, 80% of the measurement data from all bales in Table 4 were used for training the data-driven model for four-material identification, and the remaining 20% were used for testing. In the SVM modelling process, the RBF kernel function is chosen. The regularization parameter C and the kernel coefficient γ , were determined through 5-fold cross-validation, with values of 100 and 10, respectively. The classification results for the four materials are shown in Table 12. The Micro Precision, Micro Recall, Micro F1-score and prediction time (T) are 0.97, 0.96, 0.96, and 0.00519 s, respectively. It can be observed that despite the increased complexity, the SVM model trained using the proposed method maintained satisfactory identification performance and high computational efficiency.

To further evaluate the performance of the method under unseen conditions, all bales in Table 5, were used for generalisation testing. Table 13 presents the classification results for the four materials results. The Micro Precision, Micro Recall, Micro F1-score and prediction time (T) are 0.96, 0.95, 0.95, and 0.00819 s, respectively. It can be observed that for the unseen conditions, the SVM model still maintained satisfactory identification performance and high computational efficiency, which exhibit excellent performance in generalisation capability.

The proposed method not only achieves good results in the identification of three materials, but also performs well in the case of four materials. This indicates that when facing new and different recyclable materials, effective identification can be achieved by acquiring data using the designed multimodal sensing system and constructing a SVM based data-driven model. The method thus demonstrates a wide range of applicability, confirming its robustness and effectiveness as a novel solution for material identification in bales.

5. Conclusions

In this study, a new method for material identification in bales is proposed for the first time, based on a purpose-built multichannel capacitive sensor and a weighing

sensor with data-driven modelling. Experiments were conducted to evaluate the proposed method and conclusions are drawn as follows.

The designed capacitive sensor measures the bale from three directions. For different bales, the three-channel voltage outputs of the sensor vary due to differences in their permittivity distributions. For the bale with uniform permittivity distribution like PVC bale, the outputs of three channels are almost the same. For the bales with non-uniform distribution, an obvious difference is seen from the outputs of three channels. The three-channel voltage outputs of the capacitive sensor effectively capture the variations in permittivity distribution. The output of the weighing sensor effectively captures the mixture density of the bale. Combining the capacitive and weighing sensors enable a comprehensive characterization of both the permittivity and density properties of the bales.

SVM models with three inputs and four inputs have been developed to identify materials including paper, cardboard, and PVC in bales. Compared to the model using only the outputs of the capacitive sensor as inputs, incorporating the weight information significantly improves the accuracy of material identification. The results demonstrate that using multimodal sensing has advantages over using the capacitive sensor alone. The SVM algorithm was compared with random forest (RF) and multilayer perceptron (MLP) algorithms, and SVM demonstrated superior performance in data-driven modelling. The Micro F1-score of SVM for identifying paper, cardboard, and PVC is 0.99, which serves as a comprehensive overall metric. For the unseen bales, SVM demonstrates better generalisation ability and its Micro F1-score is 0.99. To further assess general applicability, woodchip was introduced as an additional material to the original bale composition. Despite the increased complexity, the method maintained satisfactory identification performance. The Micro F1-score of SVM for identifying paper, cardboard, PVC, and woodchip is 0.96, while for the unseen bales, the Micro F1-score is 0.95. These results confirm that multimodal sensing with data-driven modelling provides a robust and effective solution for material identification in bales. To address more complex scenarios,

acquiring more information through additional sensors and employing more advanced modelling methods, such as deep learning, represents a promising approach. Related research will be explored in future work.

Acknowledgments

This work was supported by the Engineering and Physical Sciences Research Council, UK (Grant No. EP/W026228/1).

References

- [1] W. B. Group, What a Waste: An Updated Look into the Future of Solid Waste Management, 20 September 2018. [Online]. Available: <https://www.worldbank.org/en/news/immersive-story/2018/09/20/what-a-waste-a-n-updated-look-into-the-future-of-solid-waste-management>.
- [2] S. J. Ijimdiya, M. V. Kumarasamy, J. T. Adu, D. Pandi. Role of the circular economy framework for sustainable waste management and climate change mitigation. *sustainability*. 2026; 18(4):1946.
- [3] The Recycling Partnership. 2024 State of Curbside Recycling Report. 2024. [Online]. Available: <https://www.ptonline.com/news/recycling-report-describes-status-of-material-recovery-from-consumers>.
- [4] J. S. Perera, S. K. Baduge, E. R. K. Chandrathilaka, S. Thilakarathna, T. S. Palle, A. M. Amado, P. Mendis. Enhancing the efficiency of plastic recovery facilities: systematically integrating seasonal and regional variations of municipal solid recyclable waste through infeed management. *Waste Manag.* 193:261-272, 2025.
- [5] D. Jumanazarov, J. Koo, M. Busi, H. F. Poulsen, U. L. Olsen, M. Iovea, System-independent material classification through X-ray attenuation decomposition from spectral X-ray CT, *NDT&E Int.*, vol. 116, 2020, Art. no. 102336.
- [6] R. Longo, S. Vanlanduit, P. Guillaume, Material properties identification using ultrasonic waves and laser Doppler vibrometer measurements: a multi-input multi-output approach, *Meas. Sci. Technol.*, vol. 24, 2013, Art. no. 105206.
- [7] L. Harrsion, M. Ravan, D. Tandel, K. Zhang, T. Patel, R. Amineh, Material identification using a microwave sensor array and machine learning, *Electronics*, vol. 9, 2020, Art. no. 288.
- [8] M. Yunus, S. C. Mukhopadhyay, G. Sen Gupta, A review of material properties estimation using eddy current testing and capacitor imaging, *Sens. Transducers*, vol. 100, 2009, pp. 92-115.
- [9] I. Ahmad, M. Mukhlisin, H. Basri, Application of capacitance proximity sensor for the identification of paper and plastic from recycling materials, *Res. J. Appl. Sci. Eng. Technol.*, vol. 12, 2016, pp. 1221-1228.

- [10] O. Rozenstein, E. Puckrin, J. Adamowski, Development of a new approach based on midwave infrared spectroscopy for post-consumer black plastic waste sorting in the recycling industry, *Waste Manage.*, vol. 68, 2017, pp. 38–44.
- [11] Y. Zheng, J. Bai, J. Xu, X. Li, Y. Zhang, A discrimination model in waste plastics sorting using NIR hyperspectral imaging system, *Waste Manage.*, vol. 72, 2018, pp. 87–98.
- [12] D. P. Penumuru, S. Muthuswamy, and P. Karumbu, Identification and classification of materials using machine vision and machine learning in the context of industry 4.0, *J. Intell. Manuf.*, vol. 31, pp. 1229–1241, 2020.
- [13] M. J. Da Silva, E. Schleicher, and U. Hampel, Capacitance wire-mesh sensor for fast measurement of phase fraction distributions, *Meas. Sci. Technol.*, vol. 18, no. 7, 2007, pp. 2245–2251.
- [14] W. Warsito, Q. Marashdeh, and L.-S. Fan, Electrical capacitance, volume tomography, *IEEE Sensors J.*, vol. 7, no. 4, pp. 525–535, Apr. 2007.
- [15] W. Zhang, C. Wang, W. Yang, and C.-H. Wang, Application of electrical capacitance tomography in particulate process measurement – A review, *Adv. Powder Technol.*, vol. 25, no. 1, 2014, pp. 174–188.
- [16] M. Meribout and I. M. Saied, “Real-time two-dimensional imaging of solid contaminants in gas pipelines using an electrical capacitance tomography system,” *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, 2017, pp. 3989–3996.
- [17] X. Wang, Y. Hu, H. Hu, and L. Li, Evaluation of the performance of capacitance sensor for concentration measurement of gas/solid particles flow by coupled fields, *IEEE Sens. J.*, vol. 17, no. 12, 2017, pp. 3754–3764.
- [18] A. Rivadeneyra and J. A. López-Villanueva, “Recent advances in printed capacitive sensors,” *Micromachines*, vol. 11, no. 4, p. 367, Apr. 2020.
- [19] A. do N. Wrasse, E. N. dos Santos, M. J. da Silva, H. Wu, and C. Tan, Capacitive sensors for multiphase flow measurement: A review, *IEEE Sensors Journal*, vol. 22, no. 22, pp. 21391–21409, Nov. 15, 2022.
- [20] A. Veisi, M. H. Shahsavari, G. H. Roshani, E. Eftekhari-Zadeh, E. Nazemi, Experimental study of void fraction measurement using a capacitance-based sensor and ANN in two-phase annular regimes for different fluids, *Axioms*, 2023, 12, 66.
- [21] D. Wang, J. Sun, Y. Wang, G. Yang, Z. Zhu, Z. Xie, Eliminating the influence of moisture on solid concentration measurement in gas–solid flows using combined sensors, *IEEE T. Instrum. Meas.*, vol. 72, 2023, Art no. 7500310.
- [22] Z. Liu, H. Wang, S. Sun, L. Xu, W. Yang, Investigation of wetting and drying process in a spout-fluid bed using acoustic sensor and electrical capacitance tomography, *Chem. Eng. Sci.*, vol. 281, 2023, Art no. 119160.
- [23] D. Wang, L. Wang, Y. Yan. Identification of baled materials through capacitive sensing and data driven modelling. *Measurement: Sensors*, 2024, 101617.
- [24] C. Cortes, V. Vapnik, Support-vector networks, *Mach. Learn.*, vol. 20, 1995, pp. 273–297.
- [25] Y. Yan, L. Wang, T. Wang, X. Wang, Y. Hu, Q. Duan, Application of soft computing techniques to multiphase flow measurement: A review, *Flow Meas. Instrum.*, vol. 60, 2018, pp. 30–43.

- [26]Y. Yu, M. Rashidi, B. Samali, A. M. Yousefi, and W. Wang, Multi-image-feature-based hierarchical concrete crack identification framework using optimized SVM multi-classifiers and D–S fusion algorithm for bridge structures, *Remote Sensing*, vol. 13, no. 2, p. 240, 2021.
- [27]L. Biddle and S. Fallah, A novel fault detection, identification and prediction approach for autonomous vehicle controllers using SVM, *Automotive Innovation*, vol. 4, pp. 301–314, 2021.
- [28]T. Cuong-Le, T. Nghia-Nguyen, S. Khatir, P. Trong-Nguyen, S. Mirjalili, and K. D. Nguyen, An efficient approach for damage identification based on improved machine learning using PSO-SVM, *Engineering with Computers*, vol. 38, pp. 3069–3084, 2022.
- [29]J. Sun, W. Yang, A dual-modality electrical tomography sensor for measurement of gas–oil–water stratified flows, *Measurement*, vol. 66, 2015, pp. 150-160.