Two-stage machine learning models for bowel lesions characterisation using self-propelled capsule dynamics

Kenneth Omokhagbo Afebu · Jiyuan Tian · Evangelos Papatheou · Yang Liu · Shyam Prasad

Abstract To foster early bowel cancer diagnosis, a non-invasive biomechanical characterisation of bowel lesions is proposed. This method uses the dynamics of a self-propelled capsule and a two-stage machine learning procedure. As the capsule travels and encounters lesions in the bowel, its exhibited dynamics are envisaged to be of biomechanical significance being a highly sensitive nonlinear dynamical system. For this study, measurable capsule dynamics including acceleration and displacement have been analysed for features that may be indicative of biomechanical differences, Young's modulus in this case. The first stage of the machine learning involves the development of supervised regression networks including multi-layer perceptron (MLP) and support vector regression (SVR), that are capable of predicting Young's moduli from dynamic signals features. The second stage involves an unsupervised categorisation of the predicted Young's moduli into clusters of high intra-cluster similarity but low inter-cluster similarity using K-means clustering. Based on the performance metrics including coefficient of determination and normalised mean absolute error, the MLP models showed better performances on the test data compared to the SVR. For situations where both displacement and acceleration were measurable, the displacement-based models outperformed the acceleration-based models. These results thus make capsule displacement and MLP network the first-line choices for the proposed bowel lesion characterisation and early bowel cancer diagnosis.

1 Introduction

Bowel cancer (BC), also referred to as colorectal cancer, affects the large bowel which consists of both colon and rectum. It is widely believed to have emerged from the adenoma-carcinoma sequence during which benign (i.e., adenoma) lesions mutate to become malignant (i.e., adenocarcinoma) and might also spread to other parts of the body, like the liver or lungs. BC ranks as the second most deadliest cancer accounting for about a million deaths globally [22] with new cases expected to reach 3.2 million in 2040 [67]. In the UK, about 268,000 persons are currently living with BC while about 43,000 new cases and 16,500 deaths are recorded every year. 94% of these new cases are diagnosed in people over the age of 50 years while the remaining 6% amounts to about 2,600 cases in people under the age of 50 [9, 12]. In England, the five-year survival rate of BC currently stands at 57.8% [49]. Between 2013-2017, England cancer treatment data showed that patients have 98%, 93%, 89% and 44% survival rate within a year of treatment, if diagnosis is made at Stage I, II, III and IV, respectively [6]. The afore-stated thus makes early diagnosis very crucial to BC treatment and survival.

Several methods are currently employed for BC screening, and these include rectal examination, faecal immunochemical test, complete blood count, computed tomography colonography, magnetic resonance imaging, positron emission tomography scan and endoscopy

K. O. Afebu, J. Tian, E. Papatheou, Y. Liu
Engineering Department, University of Exeter, Exeter, EX4 4QF, UK
E-mail: k.afebu, jt535, e.papatheou, y.liu2@exeter.ac.uk

S. Prasad
Royal Devon University Healthcare NHS Foundation Trust, Barrack Road, Exeter, EX2 5DW, UK
E-mail: shyamprasad@nhs.net

[13]. Results from some screenings can sometimes point to other diseases, hence, endoscopy which could be colonoscopy, sigmoidoscopy or capsule endoscopy is often the gold test for BC diagnosis. Colonoscopy examines the whole inside of large bowel while sigmoidoscopy examines just the lower part, both using a flexible tube with light and camera at its end. The tube is passed into the bowel through the back passage and it is gently pushed through. Inside pictures of the bowel are then viewed on a TV monitor while notable observations are being recorded. With these methods, advanced stage cancer lesions are easily detected by the endoscopist while early stage cancers are quite difficult and burdensome to detect as they often appear as subtle mucosal lesions [19]. Due to increasing demand, associated discomfort, risk of infection, the difficulty in advancing far into the bowel and the vigorous training requirement of colonoscopy, alternative and less invasive capsule endoscopy were developed [73]. Capsule endoscopy (CE) makes use of a wireless pill sized video camera to examine the inside of the bowel. The patient swallows the capsule containing a small disposable camera which takes series of pictures as it travels through the digestive tract. The pictures are either stored on-board or are wirelessly transmitted to a worn external recorder. CE has led to increased patients participation in BC screening [24] and has aided the detection of polyps that are greater than 6 mm [35].

Amongst the current clinically available CEs, we have those propelled by peristalsis and those propelled using external magnetic drag force. The peristalsis based capsule lacks flexibility in their control and take about 10 to 12 hours to traverse the entire gastrointestinal tract. Examples include PillCam from Medtronic Ltd [45], Capsocam plus from CapsoVision Inc· [14], C-Scan cap from Check-Cap Ltd [16] and the Capsules Endoscopy Systems from Jimhans Medical [32]. The magnetic prototypes are propelled by the interaction between an embedded inner magnet and an external magnetic field from a coil or permanent magnet [68]. To a certain degree, they allow for some control, however, it is yet to be ascertained if the external drag force poses danger to the delicate intestinal wall. They are sometimes known to be limited to the patient's body mass index [4]. Examples include the EndoCapsule from Olympus Corp. [50], OMOM robotic capsule from Jinshan Group [33], Mirocam from IntroMedics Co., Ltd [30], Navicam from AnX Robotica Corp. [4], and Dasheng Capsule from JIFU Medical Technology Co., Ltd [31]. Another propelling mechanism is the use of embedded actuators with locomotive legs [57], and this is sometimes combined with magnetic field to form the hybrid capsules [62]. However, a major drawback to commercial locomotive capsules is that the independently moving legs can be endangering to the gastrointestinal tract while for the hybrid capsules, the required power source is difficult to be integrated into a swallowable pill.

To address the drawbacks of existing CEs, the Exeter Small-Scale Robotics Laboratory at the University of Exeter is currently developing a self-propelled vibro-impact capsule (SP-VIC) system [38]. Using internally generated vibro-impacts, the capsule is capable of both forward and backward progressions that allow physicians to revisit areas of interest during endoscopy. The driving and frictional forces are very minimal and thus, pose no treat to the intestinal wall. During a typical procedure, the SP-VIC is operated in two modes, progression and diagnostic modes. In progression mode, the capsule is driven to places of interest within the bowel, while in diagnostic mode, it is stabilised on a lesion of potential concern for further diagnosis. In both modes, the capsule is guided through the bowel using an external coil panel while relying on the live feedback from an affixed camera. At each point, the prevailing mode is dependent on the parameters of the coil (i.e., forcing parameters). The clinical illustration of the proposed early bowel cancer detection is presented in Fig. 1. This novel approach is expected to foster the detection of hard-to-visualise early bowel cancers for the purpose of improving treatment and increasing survival rate.

The rest of this paper is organised as follows. Section 2 discusses the current use of machine learning in BC while Section 3 introduces the mathematical model, operational modes and accompanying dynamics of the capsule. In this section, measurable capsule dynamics and extractable features are also described for the simulation and experimental prototype of the capsule. In Section 4, the preliminaries of the utilised predictive models are discussed. The results and performances of the two-stage machine learning models in characterising the lesion are presented in Section 5 for both simulation and experimental data. Derivable conclusions about the entire study are presented in Section 6.

2 Machine learning and bowel cancer diagnosis

Machine learning (ML) as a sub-field of artificial intelligence, uses computational algorithms to find patterns from large-scale data and build models that can be used for future prediction and forecasting. ML thus provides computers with the ability to learn and improve from data without being explicitly programmed. As earlier stated, most of the currently existing BC screening methods rely on visual examination and often produce
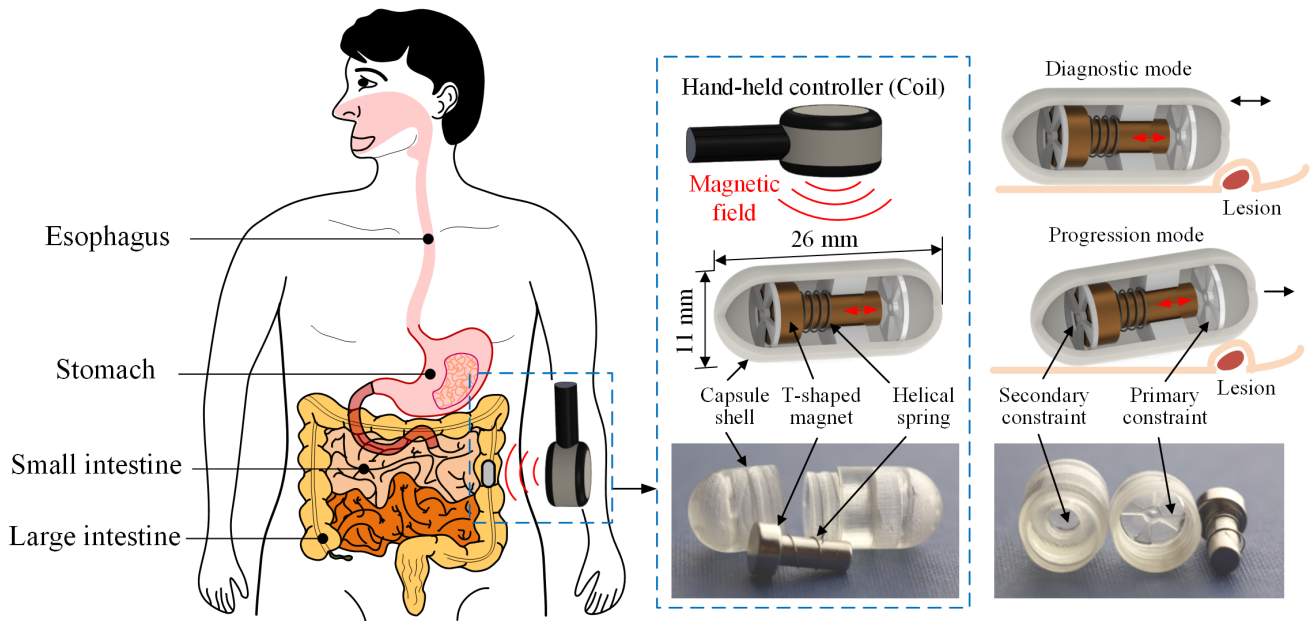
Fig. 1: Clinical illustration of the proposed early bowel cancer detection, where a capsule prototype (26 mm in length and 11 mm in diameter) was designed for experimental testing. The prototype contains a T-shaped permanent magnet for vibration, a helical spring for reverting the magnet's position and a capsule shell with a primary and a secondary constraints for restricting the vibration of the magnet. Once the magnet is excited by an external coil using a square wave signal, it may impact with the constraints, so the prototype can progress either forward or backward. During endoscopic procedure, a clinician will hold the coil panel above the capsule, and the capsule will be excited and travel from the patient's rectum to the cecum for examination.

huge volume of images that can be cumbersome and time consuming for the endoscopists to analyse. With the recent advancement in computer image processing, ML - especially deep learning - has been integrated into BC screening for easy interpretation and improved accuracy. While relying on shape, colour, and textural information, convolutional neural networks have been widely used for the detection and classification of BC from endoscopy images and videos [56, 59, 77]. Other artificial networks such as K-Nearest Neighbour, Support Vector Machine, Random Forest, XGBoost, Multilayer Perceptron, Tree Classifiers and Naive Bayes have been used with extractable image features, vitamin D levels and laboratory data like complete blood counts have been used to do the same [7, 36, 55, 58, 75]. In BC treatment, aside from being used for diagnosis, ML has also found use in treatment response monitoring [21, 60], prepared bowel quality assessment [71] and survival rate prediction [72].

Over the years, the integration of ML into BC diagnosis has evolved into two major components, the computer-aided detection and the computer-aided diagnosis systems [65]. Computer-aided detection systems are useful for locating lesions in the endoscopy images while computer-aided diagnosis systems are useful for characterising the lesions into benign and malignant lesions. A major drawback to the use of existing BC screening methods and their ML-based models for early

BC detection is their reliance on visual examination. Diagnosis is often dependent on post-development or late-stage features, thus impeding efforts to improve BC treatment and survival rate via early detection.

The early stage of BC development has been known to be characterised with changes in the biomechanical properties of affected tissues [51]. These characteristic and intrinsic changes are however obscured and unquantifiable to the endoscopists using visual examination. A typical biomechanical property that tends to increase when bowel lesions become infected with cancer is stiffness [2, 20]. This increase has been attributed to the overproduction of collagens, pathological collagen crosslinking and alignment of fibres [11]. Being a measure of resistance to deformation under an applied force, stiffness has often been denoted as Young's modulus (E) in most literature. Stiffer materials tend to have higher E-values and will only elastically change their shape slightly under an applied load. Studies such as [10, 20, 34] have demonstrated the use of tissue biomechanical changes including stiffness to distinguish between healthy and cancerous bowel tissues. These previous studies however had to carry out rigorous external soft tissue characterisation and invasive endoscopic biopsy. In this current study, an AI-assisted and non-invasive dynamic method of soft tissue characterisation is investigated for the purpose early BC detection. The proposed method explores biomechanical tissue charac-

terisation using the exhibited dynamics of the SP-VIC system [38] travelling and encountering lesions in the bowel. On encountering a lesion, the capsule's responses become greatly influenced by the surrounding tissue's biomechanical properties, including the mucosal friction and elastic modulus. Based on this, the measurable capsule dynamics are envisaged to carry intrinsic information about the biomechanical properties of the encountered lesions. However, with no established mathematical relationship between the measurable dynamics and the tissue's biomechanical properties, the present study utilises a data driven approach to extrapolate the relationship. Measurable dynamic variables of the capsule travelling and encountering lesions in the bowel are analysed for features that may be indicative of changes in their biomechanical properties. Obtained features are used to develop artificial intelligence models that are capable of biomechanical tissue characterisation.

3 The self-propelled vibro-impact capsule

The drawbacks of peristalsis dependent CEs have been stated in Section 1 and these include long travelling hours and lack of flexibility. Also, for those that depend on external magnetic drag, manoeuvring can be difficult and patients under the age of 22 years or with body mass index greater than 38 are often not recommended [4]. The cohesion and dragging of the capsule against the intestinal walls may result in intestinal tears and wears while larger body mass may mask the magnetic interactions. To overcome these limitations while also enabling better manoeuvring of CEs, the SP-VIC [38] is currently being developed by the Small-Scale Robotics Laboratory at the University of Exeter. The capsule navigates the bowel via series of vibro-impacts resulting from an externally excited magnetic inner mass situated in the capsule [38, 39]. The principle of utilising vibro-impacts as driving mechanism in engineering systems has been in existence for a while and was first identified in 1928 at the University of Gottingen in Germany when a vibratory driver was used to install timber piles [63]. Recent applications are seen in ground moling [54, 63], percussive drilling [52] and pipeline inspection [69]. From the nonlinear dynamics point of view, the major differences between the SP-VIC and the ground moling (or percussive drilling) system are twofold. (1) The ground moling system, such as [53, 54], requires a static and a dynamic force to be applied to the system for progression, while the SP-VIC only requires a dynamic force. (2) The moling system has a unidirectional locomotion, while the SP-VIC has a bidirectional locomotion. According to [37], the SP-VIC has more complex phenomena and is more challenging to be controlled. However, both systems achieve their best performances in term of progression rate at their period-one motions [39, 53].

3.1 Physical model of the SP-VIC

The physical model of the SP-VIC system is shown in Fig. 2, having a cylindrical body of length L, radius of R and mass m_c. It is driven by an externally excited magnetic inner mass m_m which vibrates to impose impacts on the primary and secondary constraints of stiffnesses k_1 and k_2, respectively. A damped spring with stiffness k and damping c connects m_m to the capsule shell. k_1 and k_2 are separated from m_m by gaps g_1 and g_2, respectively. The conceptual design shown in Fig. 1 shows a T-shaped permanent magnet acting as the inner mass (m_m) and a helical spring that acts as the damped spring (k) that reverts the inner mass position.
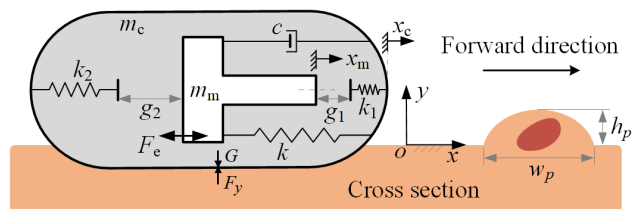

Fig. 2: Physical model of the SP-VIC system [1] and an approached bowel lesion with a width w_p and height h_p.

The magnetic inner mass is subjected to periodic excitations using an external non-sinusoidal force F_e, in this case a square waveform signal, and given as

$$F_e = \begin{cases} P_d, & \mod(t, T) \in [0, DT], \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

where P_d, T and $D \in [0, 100\%]$ are respectively amplitude, period, and duty cycle representing the excitation force parameters, and $\mod(t, T)$ indicates t modulo T.

The capsule is respectively driven forward or backward when the inner mass impacts the forward and backward constraints k_1 and k_2. Based on the capsule's free-body diagram and acting forces, its dynamics while traversing the bowel can be modelled as [70]

$$\begin{cases} m_m \ddot{x}_m = F_e - F_i, \\ m_c \ddot{x}_c = F_i + F_x + F_f, \end{cases} \quad (2)$$

where F_i represents the interactive driving force imposed on the capsule by the impacting inner mass and it is given as

$$F_i = \begin{cases} kx_r + cv_r + k_1(x_r - g_1), & \text{if } x_r > g_1, \\ kx_r + cv_r, & \text{if } -g_2 \leqslant x_r \leqslant g_1, \\ kx_r + cv_r + k_2(x_r - g_2), & \text{if } x_r < -g_2. \end{cases} \quad (3)$$

where $x_r = x_m - x_c$ and $v_r = \dot{x}_m - \dot{x}_c$, and both represent the relative displacement and velocity between the inner mass and the capsule shell, respectively. F_x is the horizontal reaction from the lesion while F_f is Coulomb friction accounting for the tangential contact force between the capsule shell and the lesion. Provided the vertical reaction force, F_y on the capsule from the tissue is cancelled out by gravity G, the overall frictional force is given as

$$F_f = \begin{cases} -\text{sign}(\dot{x}_c)\mu G, & \text{if } \dot{x}_c \neq 0, \\ -\text{sign}(F_i + F_x)\mu G, & \text{if } \dot{x}_c = 0 \ \& \ |F_i + F_x| \geq \mu G, \\ -F_i - F_x, & \text{if } \dot{x}_c = 0 \ \& \ |F_i + F_x| < \mu G, \end{cases} \tag{4}$$

where μ is the frictional coefficient and F_f has been proven to be a good representation of the friction between the capsule and the intestinal walls [25]. A detailed study of F_x is outlined in Yan *et al.* [70].

While operating the SP-VIC, P_d, T and D are often alternated in order to impose either the diagnostic or progression mode. In diagnostic mode, the parameters are carefully selected such that m_m is restricted to forward impacts that causes the capsule to move but sticks and vibrates on any encountered lesion. In progression mode, utilised parameters produce either forward or backward impacts that are strong enough to respectively cause the capsule move either forward or backward without being stuck at lesion points.

3.2 Measurable dynamics and feature extraction

3.2.1 Simulation capsule dynamics

For a dynamically impacting engineering system like the SP-VIC, measurable dynamic variables include displacement (x), velocity (v) and acceleration ($\ddot{x}$) signals. Their measurement is often considered to be valuable for understanding the prevailing dynamics of such system. For simulation, it is possible to measure all the dynamic variables including x, v and $\ddot{x}$. However, this may not be possible during experiment due to restricted size and space or unavailability of necessary sensors. For the SP-VIC, x_c and $\ddot{x}_c$ will be explored during simulation but only x_c will be explored for the experimental validation due to the above mentioned restrictions.

Being a highly sensitive non-smooth dynamical system, resulting x_c and $\ddot{x}_c$ signals are expected to have intrinsic but complex nonlinear correlation with the capsule's operating parameters and the conditions of its surrounding environment. It is on this basis that the dynamic signals are proposed for the biomechanical characterisation of encountered lesion using ML

models which are capable of modeling complex nonlinear problems [5]. The simulation signals were generated by numerically solving Eq. (2) using the Rung-Kutta method of solving differential equations at a constant step-size. During the numerical solution, the Young's modulus (E) of the encountered lesions were varied while the capsule's forcing parameters were kept constant as listed in Table 1.

Table 1: Capsule parameters for the dynamic simulation.

Parameters	Symbols	Units	Values
Capsule length	L	mm	26
Capsule radius	R	mm	5.5
Capsule mass	m_c	kg	0.00167
Inner mass	m_m	kg	0.0018
Primary stiffness	k	N/m	62
Secondary stiffness	k_1	N/m	27900
Tertiary stiffness	k_2	N/m	53500
Gap to k_1	g_1	mm	0.8
Gap to k_2	g_2	mm	0.8
Damping	c	Ns/m	0.0156
Duty cycle ratio	D	-	0.5
Excitation period	T	s	0.05
Lesion height	h_p	mm	2
Lesion width	w_p	mm	8
Excitation amplitude	P_d	mN	12

Table 2: Simulated lesion categories and ranges of E-values.

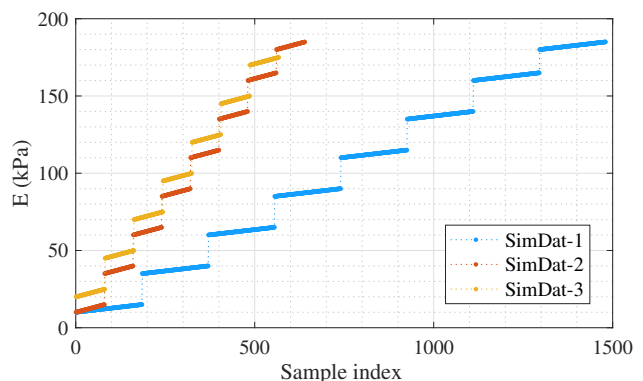| Categories	Ranges (kPa)		
	SimDat-1	SimDat-2	SimDat-3
S1	10 - 15	10 - 15	20 - 25
S2	35 - 40	35 - 40	45 - 50
S3	60 - 65	60 - 65	70 - 75
S4	85 - 90	85 - 90	95 - 100
S5	110 - 115	110 - 115	120 - 125
S6	135 - 140	135 - 140	145 - 150
S7	160 - 165	160 - 165	170 - 175
S8	180 - 185	180 - 185	-
Total samples	1480	640	567

Fig. 3: E-values variation in the three sets of simulation data.

Three sets of data listed in Table 2, SimDat-1, SimDat-2 and SimDat-3 representing dynamic capsule signals from three different patients were simulated. Each patient's data consisted x, v and $\ddot{x}$ measurements resulting from the capsule's interaction with different categories of lesions. SimDat-1, SimDat-2 and SimDat-3 differ based on the ranges of E-values defining each lesion category as well as the step-size at which the ranges were sampled, see Table 2 and Fig. 3. The E-values ranges for SimDat-1 were sampled at a step-size of 0.027 while those of SimDat-2 and SimDat-3 were sampled at 0.062. This way we ensured that no same E-value is repeated in any of the three data set. The remaining parameters of the capsule as defined in Table 1 were carefully selected to ensure that the inner mass only exhibit forward impacts which do not cause the capsule to cross over the lesion. In all, 1480, 640 and 567 signal samples were simulated for SimDat-1, SimDat-2 and SimDat-3, respectively, and the inner mass was seen to relatively exhibit period-one motions with one impact per period of excitation. Typical displacement and acceleration signals of the inner mass and the capsule are shown in Fig. 4 alongside their representative phase portrait on the $x - v$ plane. The impacts of the inner mass on the forward constraint are well obvious from the acceleration signals shown in Fig. 4(b). As expected, the impacts decelerate the inner mass but accelerate the entire capsule, and the two phenomena are well represented as peaks in the acceleration signals. For the simulated data, the imposed capsule acceleration was seen to be almost equal to the inner mass deceleration as the amplitudes of both phenomena showed correlation coefficients, $R \geq 0.99$.

The variation of x_c and $\ddot{x}_c$ signals for different E-values are presented in Fig. 5. The differences in the E-values were more observable from the x_c signals compared to $\ddot{x}_c$ signals, especially for the lower E-values. The capsule showed higher displacement measurements for softer lesions and lesser displacement measurements for the stiffer lesions, see Fig. 5(a). This is likely due to its ability to impose a greater push on the softer lesions compared to the stiffer lesions provided an equal driving force is used. For the parameters defined in Table 1 and Table 2, the relative motion of the capsule was found to exhibit period-one oscillation with one forward-impact.

3.2.2 Experimental capsule dynamics

To validate the use of the SP-VIC for dynamic biomechanical lesion characterisation, a laboratory experiment whose schematics and set-up are described in Fig. 6 was carried out. The prototype capsule described in Fig. 1 is made to move and vibrate on synthetic le-
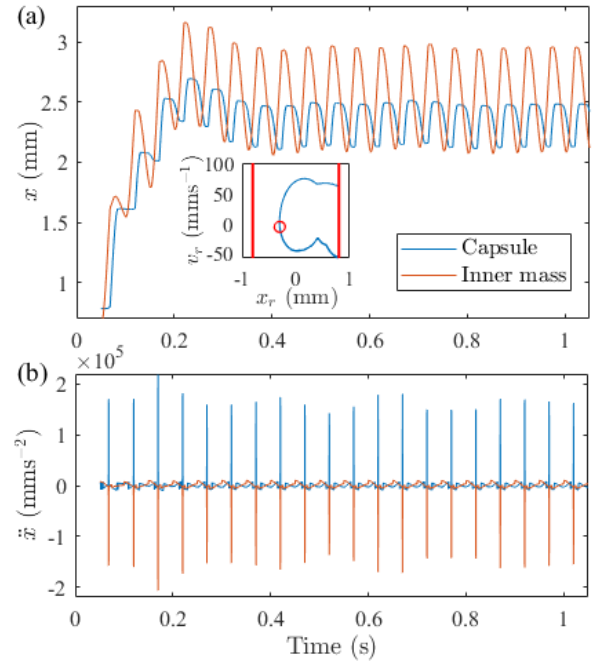

Fig. 4: (a) Typical displacement and (b) acceleration signals of the inner mass and the capsule alongside their representative phase portrait on the $x - v$ plane. The solid red lines are the impact boundary representing the locations of the forward and backward constraints. The red circle indicates the position of the system after each period of oscillation thus implying period-one oscillation.
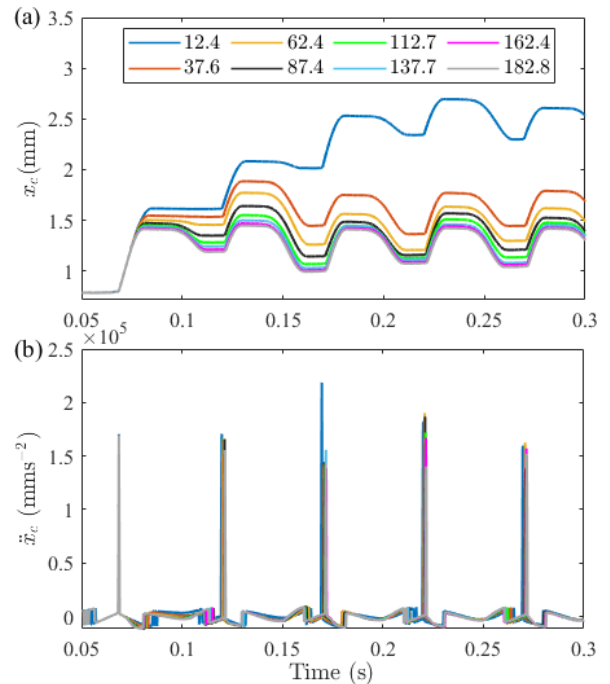

Fig. 5: Variation of (a) x_c and (b) $\ddot{x}_c$ for different E-values.

sions lying along an intestine-like Ecoflex material supported in an half-opened tube. The movement of the capsule is made possible by the impact actions of the T-shaped magnetic inner mass on the forward and backward constraints. The impacts are produced as a result

of the interaction between the excited magnetic inner mass and the restoring helical spring. The inner mass excitation is induced intermittently by an on-again-off-again external electromagnetic field $\vec{B}$ from the coils. A signal generator which produces pulse width modulation signals (in this case square wave signals) via a power amplifier is used to power the coil. The amplifier controls the voltage to the coils producing the external electromagnetic field via the DC power supply unit. The intestine-like Ecoflex material supported in an half-opened tube is placed along the central axis of the coils alongside the prototype SP-VIC. Once the magnetic inner mass is excited by the external electromagnetic field using the square wave signal, it can either impact the forward or backward constraints, causing the capsule to either progress forward or backward. As the capsule traverses the bowel, its displacement (x_c) being the only experimentally measurable dynamics is measured using a laser displacement sensor. Hence, as earlier stated, only x_c signals resulting from the capsule's interaction with the synthetic lesions were utilised for the current experimental validation.

Ecoflex 00-10, 00-30 and 00-50 with manufacturer's E-values of 55.16, 68.95 and 82.74 kPa respectively were utilised in fabricating the synthetic lesions. In an attempt to further increase the possible number of samples and the ranges of experimental E-values, another set of lesions were fabricated by mixing each of the Ecoflex materials with 6% by weight of thinner. The E-values of the fabricated lesions as found via quasi-static deformation test using INSTRON 3367 are presented in Table 3. For each of the Ecoflex materials, lesions resulting from their mixture with thinner showed lowered E-values.

Table 3: Resulting E-values of the Ecoflex materials mixed with different percentages of thinner.

Ecoflex	Thinner (%)	E (kPa)	Categories
$00-10$	0	79.8	E3
	6	68.8	E1
$00-30$	0	83.3	E4
	6	74.1	E2
$00-50$	0	120.4	E6
	6	110.1	E5

The capsule's operating parameters were kept the same as those reported in Table 1 except for P_d. Given that the distance between the endpoint of m_m and the mid-width of the coils is x_{md}, Zhang *et al.* [74] established that the exerted P_d is dependent on $\vec{B}$ and x_{md}. For the coil utilised in this study, they established a nonlinear relationship between P_d and x_{md} and they verified it analytically (using established mathematical formula), numerically (using ANSYS Maxwell) and experimentally (using a weighing scale). Based on this

established relationship, for $x_{md} = 35$ mm as used in this study, a P_d of 21.64 mN is estimated to be exerted on m_m.

By means of repetition, 600 experimental data (ExpDat) were collected for the six synthetic lesions using the set-up in Fig. 6(b). 70% of these data (ExpDat-1) were used for training the experimental MLP and SVR models while the remaining 30% (ExpDat-2) were used as the out-sample testing data. Typical x_c signals resulting from the capsule's interaction with the synthetic lesions of different E-values are shown in Fig. 7.

3.2.3 Dynamic feature extraction and selection

Dynamic systems with real-life application usually produce timestamped data which are initially large and unwieldy. Aside requiring huge computing resources and time, they often produce poor performing ML models. To produce improved and robust ML models, feature extraction and selection are employed as preliminary techniques to compact the initial data into a lower dimensional space. Feature extraction transforms the original raw data into numerical features that are representative of the central information carried in the original data set. Feature selection on the other hand, selects a subset of the extracted features that best map the response variable depending on the utilised selection algorithm. This is because too many features may increase computational cost and as well degrade the performance of the ML model even though they bear relevant information about the response variable. For continuous regression problems, the final features should be quite representative of the response variable(s) and be more discriminating for classification problems. Feature extraction can be manual, requiring the calculation and identification of relevant features for the problem or automated, requiring the use of specialised algorithms or deep networks to automatically extract features from signals or images without human intervention [42]. For this study, a manual feature extraction that yielded forty-one features was carried out. The features include time and frequency domain statistical features, nonlinearity features and waveform features as listed in Table 4.

Each feature in Table 4 is expected to be indicative and representative of the biomechanical variations in the lesions, hence, they were further analysed to identify those that best indicate biomechanical variations. Coefficient of determination (R^2) from correlation analysis and average silhouette value ($AvSlh$) from silhouette analysis were used to evaluate the features. R^2 was used to evaluate the proportion of the variance in E-values that is explained by each feature value [47]. R^2
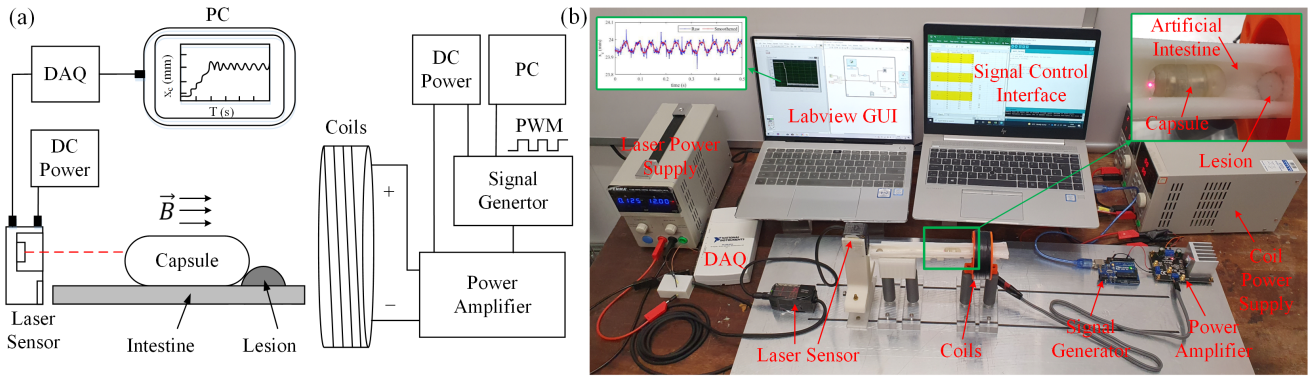
(a)

(b)

Fig. 6: (a) Schematic and (b) photograph of the experimental setup for a prototype SP-VIC. The impacts of the T-shaped magnetic inner mass on the forward and backward constraints aids the capsule's locomotion.

Table 4: Extracted dynamic signal features.

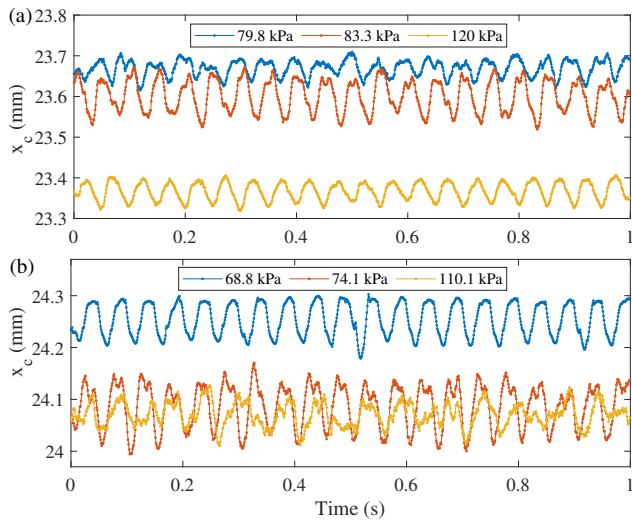| Time domain features:	Root-sum-of-squares (RSSq)°◇	Mean frequency°
1. Mean°	15. Root-sum-of-squares (RSSq)°◇	29. Mean frequency°
2. Minimum°	16. Crest factor°	30. Median frequency°
3. Maximum°	17. Mean absolute value°	31. Band power
4. Standard deviation°◇	18. Form factor°◇	32. Occupied bandwidth
5. Range	19. Impulse factor°	33. Power bandwidth
6. Kurtosis°	20. Mean square root of absolutes°	34. Peak amplitude
7. Variance°◇	21. Kurtosis factor°	35. Power spectral density
8. Skewness°◇	22. Margin factor°	**Nonlinearity features:**
9. Covariance°◇	23. Skew factor°◇	36. Lyapunov exponent
10. Mean normalised frequency°	24. Shape factor°◇	37. Correlation dimension
11. Avg cumulative maximum°	25. Signal-to-noise ratio (SNR) ◇	**Waveform features:**
12. Avg cumulative minimum°	26. Total harmonic distortion (THD)	38. Approximate entropy
13. Maximum absolute value-to-RMS ratio°	27. SNR-to-THD ratio◇	39. Mean of peaks°
14. Root-mean-square (RMS)°	28. Clearance factor°	40. Average peak separations
	Frequency domain features:	41. Area under curve°

(a)

(b)

Fig. 7: Typical experimental x_c signals from the capsule interacting with lesions mixed with (a) 0% and (b) 6% thinner.

ranges between 0 and 1 and features that showed higher R^2 values on correlation with the targeted E-values stand the chance of being better predictors compared to those with lower R^2 values. Silhouette analysis on the other hand, was used to examine the potential of

the features to pair-wisely partition the data into clusters that match their actual categories (see Table 2). The silhouette value for i^{th} data point (S_i) is a measure of how similar the feature value for that data point is to other feature values in the same cluster, compared to feature values in neighbouring clusters. Silhouette values range between -1 and 1 with values close to 1 indicating well separated data point and values close to 0 indicating very close points. -1 indicates data points that could not fit into any of the pre-defined clusters. The averages of the silhouette values ($AvSlh$) resulting from each evaluated feature pairs was used alongside R^2 for the feature selection. Features showing R^2 and $AvSlh \geq 0.6$ were categorised as good indicators and predictors of biomechanical properties. Features with both R^2 and $AvSlh \geq 0.6$ using x_c and $\ddot{x}_c$ signals are indicated with ○ and ◇ respectively in Table 4. The x_c signals were found to show more and better correlated features (indicated with °) compared to $\ddot{x}_c$ signals ((indicated with ◇). For N observed samples, R^2 is given

as

$$R^2 = 1 - \frac{\sum\limits_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum\limits_{i=1}^{N}(y_i - \overline{y})^2} \qquad (5)$$

with y_i being the ith feature variable, $\hat{y}_i$ the ith E-value and $\overline{y}$ the mean of the feature variables. $AvSlh$, on the other hand is given as

$$AvSlh = \frac{1}{N}\sum_{i=1}^{N}(S_i) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{b_i - a_i}{max(b_i, a_i)}\right) \qquad (6)$$

with a_i representing the average distance between the ith feature value(s) and the other feature value(s) in the same cluster as i, while b_i is the minimum average distance between the ith feature value(s) and the other feature value(s) in different clusters, minimised over the clusters. Figure 8 shows the variation of some resulting x_c-based (a-c) and $\ddot{x}_c$-based features (d-f) alongside their R^2 correlation with corresponding $E-$values in an increasing order. Fig. 9 on the other hand, shows the scatter plots and the silhouette plots of feature-pairs exhibiting high and low $AvSlh$-values..
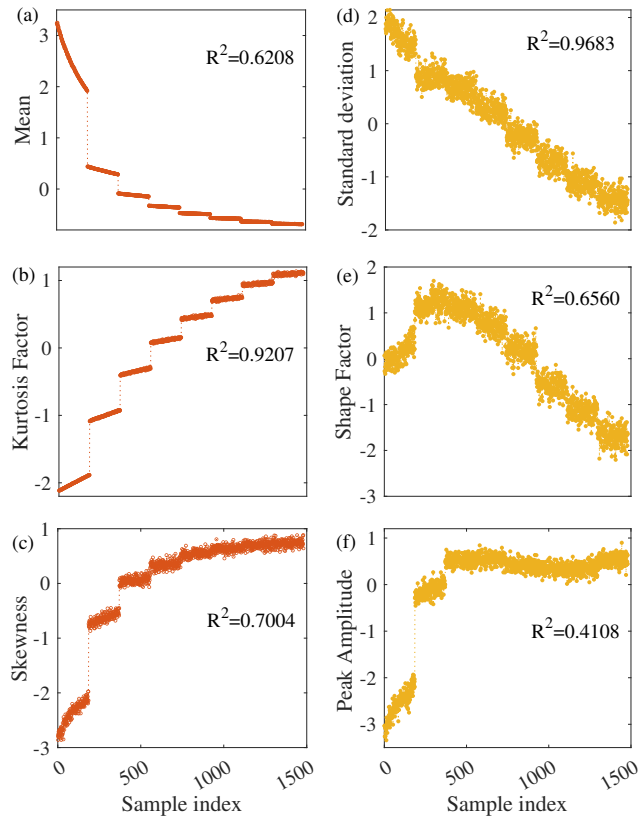

Fig. 8: Variation of some (a-c) x_c-based and (d-f) $\ddot{x}_c$-based features with respect to the increasing E-values of SimDat-1.
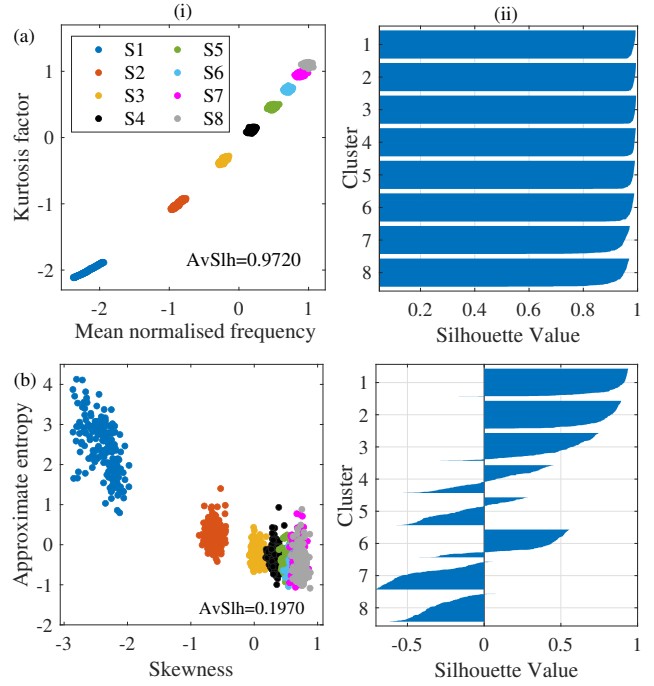

Fig. 9: Typical (i) scatter and (ii) silhouette plots of paired features exhibiting (a) high and (b) low $AvSlh$.

4 Predictive models

For early BC diagnosis, the current study proposes a non-invasive biomechanical evaluation of bowel lesions using machine learning models. To build the models, features that may be indicative of changes in the tissue biomechanical properties (e.g., Young's modulus) are extracted from the dynamic signals. The features are used to build ML models that are capable of detecting early and hard-to-visualise malignant lesions. ML models were adopted as there are no established mathematical models explaining the relationship between the extracted features and the corresponding E-values of their parent signals. The E-values of benign and malignant lesions often tend to vary among patients due to their differences in age, ethnicity and nutritional background. Despite the great ability of ML models to model complex and nonlinear problems [5], the aforementioned makes it challenging to build a single ML model that can generalise for all patients. To circumvent this challenge and to build a more generalised lesion categorisation model, a two-stage prediction ML model has been proposed. The first stage involves a supervised and continuous prediction of lesion E-values from features computed from the dynamics (x_c and $\ddot{x}_c$) of the SP-VIC encountering lesions in the bowel. The second stage involves an unsupervised clustering of the lesions into subsets of different categorical ranges using the predicted E-values. This way, the stiffer and

malignant lesions are well separated from their benign counterpart for each patient irrespective of age, ethnicity and nutritional background. Going by our previous study [1], MLP and SVR were chosen for the supervised continuous regression. On the other hand, K-means clustering using the Squared Euclidean distance metric was utilised for the unsupervised categorisation. The schematic layout of the two-stage machine learning model for the proposed early bowel cancer detection is described in Fig. 10.
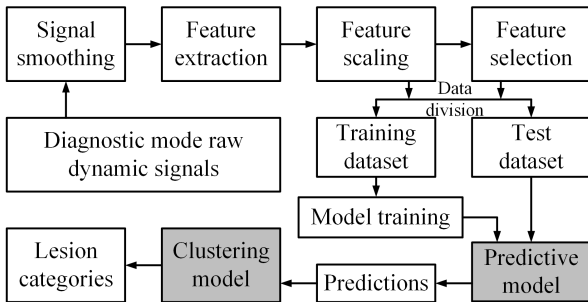

Fig. 10: Schematic of the two-stage machine learning model.

4.1 Supervised regression models

4.1.1 Multi-layer perceptron networks

As feed-forward neural networks, MLP networks tend to pass and evaluate input information in a single forward direction via series of interconnected neurons organised in layers. These include an input layer, an output layer, and a single or multiple hidden layers [46]. The input and output layers usually take up fixed numbers of neurons that are equal to the number of features in the input and output data, respectively. On the other hand, the hidden layer and its number of neurons are usually adjustable in order to make the network perform better. However, excessive numbers of neurons may increase the network training time and also cause it to over-fit while inadequate neurons will cause it to under perform. MLPs are usually fully and hierarchically connected, and the values of the neurons starting with input values propagate to the next layer neuron values via weighted connections. The weights (w), indicates the importance of a neuron value to its next connected neuron value. For a layer, its neuron values are weighted, summed, and transformed using their activation functions, to arrive at new values [23]. Aside w, bias (b), is another element taken into consideration in each layer. The bias elements represent the threshold used in shifting the activation functions in order to condition neuron outputs in both the hidden and output layers [23]. The output variables (y_k) of a two-layered MLP network for input variables x_i can be written as

[8]

$$y_k = f^o \left(\sum_{j=1}^{M} w_{kj}^{(2)} f^h \left(\sum_{i=1}^{N} w_{ij}^{(1)} x_i + b_{j0}^{(1)} \right) + b_{k0}^{(2)} \right), \quad (7)$$

where M and N respectively represent the number of hidden neurons and input data samples (x_i). w_{kj} is the weight parameter between the kth output neuron and j^{th} hidden neuron, w_{ij} is the weight parameter between the ith input data and j^{th} hidden neuron while b_{j0} and b_{k0} respectively denote the jth hidden neuron and k^{th} output neuron bias parameters. f^h and f^o are the hidden and output layers activation functions, were respectively represented as a hyperbolic tangent and linear function.

During training, the output values are compared with the actual target values, to compute the network's performance measure as mean squared errors (mse). The mse is minimised by adjusting the network weights and biases using a gradient descent-based backpropagation algorithm, in this case the Levenberg-Marquardt. By back-propagation of network error, Levenberg-Marquardt updates the network weights and biases in the direction which the performance function decreases most rapidly [26, 44] using

$$X_{k+1} = X_k - \left[H + \mu I \right]^{-1} g, \quad (8)$$

where $H = J^T J$ and $g = J^T E_r$, represent an approximated Hessian matrix and the network gradient, respectively. X_k is a vector of current weights and biases while I is the identity matrix. J is a Jacobian matrix containing the first derivative of the network errors with respect to the weights and biases, and E_r is a vector of network errors. The scalar value μ is step-wisely decreased when an iteration results into a successful reduction in mse and only increased if mse increasesonly increases if there had been an increase in mse. The iteration continues until the network converges.

4.1.2 Support vector regression

SVR, the numeric prediction version of support vector machines, are forms of kernel-based supervised learning models which compared to neural networks are nonparametric, requiring no pre-defined topology or parameters. First discovered by Vapnik [64], support vector machines are assumed to have better generalisation and less susceptible to over-fitting owing to their ability to model complex nonlinear decision boundaries [28]. During training, SVR learns to fix a function $f(x)$ in a higher dimensional feature space such that its outputs are at most, ε deviation from all the actual targets y_i,

and at the same time as flat as possible. This thus permits the flexibility of presenting SVRs with acceptable model error while finding an appropriate hyperplane in high dimensions to fit the data. Again, unlike neural networks training, no attempt is made at reducing the training errors as long as they are less than ε, however, a larger value will not be accepted. For a pair of input and target vectors, $\{(x_i, y_i)\}_i^N$, the generalised nonlinear regression function is given as

$$f(x) = \langle w, \Phi(x_i) \rangle + b, \tag{9}$$

where $\Phi(x)$ is the nonlinear mapping function while w and b are vectors of weight coefficients and bias constants, respectively. Keeping Eq. (9) as flat as possible, means w is kept as small as possible and this can be achieved by minimising the norm thus giving rise to a convex optimisation problem

$$\text{Minimise } \frac{1}{2}\|w\|^2,$$
$$\text{subject to } \begin{cases} y_i - (\langle w, \Phi(x_i) \rangle + b) \leq \varepsilon, \\ (\langle w, \Phi(x_i) \rangle + b) - y_i \leq \varepsilon. \end{cases} \tag{10}$$

The feasibility of the above convex optimisation problem assumes that the function f (i.e., $\langle w, \Phi(x_i) \rangle + b$) actually exists and approximates for all pairs of (x_i, y_i) with ε precision. To cope with possible infeasibility, non-negative loose variables $\xi_i + \xi_i^*$ are introduced into the convex optimisation problem [18] to give

$$\text{Minimise } \left[\frac{1}{2}\|w\|^2 + \mathbb{C}\sum_{i=1}^N \xi_i + \xi_i^* \right],$$
$$\text{subject to } \begin{cases} y_i - (\langle w, \Phi(x_i) \rangle + b) \leq \varepsilon + \xi_i, \\ (\langle w, \Phi(x_i) \rangle + b) - y_i \leq \varepsilon + \xi_i^*, \\ \xi_i, \xi_i^* \geq 0, \end{cases} \tag{11}$$

where $\sum_{i=1}^N \xi_i + \xi_i^*$ represents an error term and $\mathbb{C}$ is its penalty factor that determines the trade-off between the flatness of $f(x)$ and the extent to which deviations larger than ε are tolerated.

The resulting dual convex optimisation problem shown in Eq. (11) can be transformed into a dual Lagrangian problem as

$$f(x) = \sum_{i=1}^N (\Upsilon_i - \Upsilon_i^*) \langle \Phi(x_i), \Phi(x) \rangle + b \tag{12}$$

by replacing the dot product $\langle \Phi(x_i), \Phi(x) \rangle$ with a nonlinear kernel function $\kappa(x_i, x)$, the above can be solved into a regression function as

$$f(x) = \sum_{i=1}^N (\Upsilon_i - \Upsilon_i^*) \kappa(x_i, x) + b. \tag{13}$$

The parameters Υ_i and Υ_i^* represent the Lagrange multipliers while the kernel function $\kappa(x_i, x)$, helps to map the nonlinear separable feature space to a linear separable feature space [15]. For this study, the radial basis kernel function was utilised.

$$\kappa(x_i, x) = \exp\left(-\gamma\|x_i - x\|^2\right), \tag{14}$$

where γ is the kernel function parameter, and $\|x_i - x\|$ is the Euclidean distance between x_i and x.

While training an SVR model, its hyperparameters, including $\mathbb{C}, \gamma$ and ε, are usually varied to optimise its performance [3]. For a set of chosen hyperparameters, the SVR is trained on the training dataset to extrapolate the model parameters Υ_i, Υ_i^* and b. The trained model defined by its extrapolated parameters is evaluated on a set of never seen test data. For an SVR model, the process of hyperparameter optimisation simply involves selecting the best set of hyperparameters using a brute-force search technique. In this study, hyperparameter optimisation via a five-fold cross-validation procedure was carried out on the training data. For a chosen set of hyperparameters, the network is trained on four folds of the training data and tested on the remaining one fold until all folds have been used as the out-of-sample testing fold. The set of hyperparameters that yield the minimum five-fold cross-validation loss (measured as mse) are selected as the optimised parameters. The eligible range of values for $\mathbb{C}$ and γ is given as $[0.001, 1000]$ while that of ε is given as $[0.001, 1000] \times \frac{Y_{iqr}}{1.349}$ with Y_{iqr} being the interquartile range of the response variable. The sequential minimal optimisation solver as available in MATLAB was used as the optimisation routine.

4.2 Unsupervised clustering model

4.2.1 K-means clustering

Data cluster analysis, also referred to as segmentation, is a type of unsupervised ML, and it is often used to draw inferences from unlabelled data. Different types of clustering methods exist in literature, however, they are basically categorised into partitioning, hierarchical and density-based methods [28]. In this study, K-means clustering (or Lloyd's algorithm [40]) belonging to partition clustering method was utilised. First introduced by MacQueen in 1967 [41], K-means clustering has been used in many applications. These include cancer detection and survival prediction [48, 76], image and text pattern recognition [17, 66], delineating zones of mineralisation [61] and food quality inspection [29]. For a predefined k number clustering, K-means randomly selects k subsets of objects from the N observed samples

as the initial clusters (C_i) and computes their centroids. The remaining data are assigned to these clusters based on their distance to the centroids. A new set of clusters are selected and the process is repeated until the resulting clusters attain maximum inter-cluster similarities and minimum intra-cluster similarities. Different distance metrics are available for computing the distance between the centroids and the cluster objects [43], but depending on utilised distance metric, K-means computes centroids differently. For this study, the Squared Euclidean distance which uses the mean of the cluster objects ($\bar{c}_i$) as centroid was used. The distance between an object $\boldsymbol{p} \in C_i$ and $\bar{c}_i$ is thus an Euclidean distance given as ($dist(\boldsymbol{p}, \bar{c}_i)$) and the quality of cluster C_i is measured as [27]

$$E_1 = \sum_{i=1}^{k} \sum_{\boldsymbol{p} \in C_i} dist\left(\boldsymbol{p}, \bar{c}_i\right)^2, \qquad (15)$$

where E_1 is the sum of squared error between the objects in C_i and the centroid $\bar{c}_i$.

Due to the medical importance of current study and the fact that labels are available, aside the use of average silhouette ($AvSlh$) values, the clustering of each data set was further evaluated based on the percentage of rightly clustered samples (i.e. accuracy (%)).

5 Two-stage lesion categorisation and results

In this study, a two-stage ML model has been investigated for the the purpose of bowel lesions characterisation. It involved the training and development of regression models capable of predicting lesion stiffnesses from dynamic signals features and the clustering of the predicted stiffnesses into categorical ranges. MLP and SVR networks have been used for the prediction models while K-means has been used for the clustering models. The MLP and SVR were trained to predict lesion E-values from selected dynamic signal features while K-means was used to unsupervisedly group the predicted E-values into clusters with maximum intra-cluster similarities and inter-cluster dissimilarities. Table 5 shows the analysed capsule signals, the feature selection measures and the notation used to represent the utilised feature data.

5.1 Stage 1: Biomechanical stiffness prediction

The results of the biomechanical stiffness prediction including the training and testing of the MLP and SVR networks using the selected feature data are presented in Table 6 and Table 7, respectively. The networks performances have been reported in relations to the actual target E-values using coefficient of determination (R^2)

Table 5: Notation of the selected feature data.

Signal type	Feature selection	Notation
x_c	All	All–x_c
x_c	Rsq	Rsq–x_c
x_c	$AvSlh$	$AvSlh$–x_c
$\ddot{x}_c$	All	All–$\ddot{x}_c$
$\ddot{x}_c$	Rsq	Rsq–$\ddot{x}_c$
$\ddot{x}_c$	$AvSlh$	$AvSlh$–$\ddot{x}_c$

and normalised mean absolute error (NMAE). These are respectively given in Eqs. (5) and (16).

$$\text{NMAE} = \frac{\text{MAE}}{\overline{y}} = \sum_{i=1}^{N} \mid y_i - \hat{y}_i \mid / \sum_{i=1}^{N} y_i, \qquad (16)$$

where y_i is the ith value of the actual target variables y, $\hat{y}_i$ is its corresponding value in the model predictions $\hat{y}$, $\overline{y}$ is the mean of y, and N denotes number of observations.

In this case, R^2 represents the proportion of the variation in the actual target E-values that is captured in the models predictions. It thus points to how well the trained model predicted E-values match the actual E-values. Values closest to 1 indicate better matching and better network models. NMAE on the other hand, is a measure of the average error between the network predictions and the actual targets normalised with the mean of the actual targets. Normalisation is used to cancel out any probable effect that the diverse scale of the investigated E-values (Fig. 3) may have on the network errors. The closer NMAE is to zero, the better the model and its predictions.

For the simulation data, the performance metrics showed that the x_c-based networks outperformed the $\ddot{x}_c$-based networks (Tables 6 and 7). The MLP networks are seen to slightly outperform the SVR networks both during training and testing despite their shorter training times. Reducing the dimension of the features via feature selection did not always translate to improved network performances. However, it yielded lower NMAE values for all the MLP models and the x_c-based SVR models. It also reduced the training time of the MLP models to about half. On the overall, the average (R^2, NMAE) values were respectively (0.9974, 0.0131) for the MLP networks and (0.9971, 0.0195) for the SVR networks. Based on the utilised dynamic the average (R^2, NMAE) values for using the x_c signals with the MLP and SVR networks were respectively (0.9999, 0.0045) and (0.9993, 0.0085). Using $\ddot{x}_c$ signals, the average (R^2, NMAE) values were (0.9950, 0.0216) and (0.9949, 0.0306) for the MLP and SVR networks, respectively.

The performance plots of using the networks on Rsq–x_c and Rsq–$\ddot{x}_c$ feature data from the test data

Table 6: Performances of MLP network on simulation feature data.

| Simulation feature data | Training | | | | Testing | | | |
| | SimDat-1 | | | | SimDat-2 | | SimDat-3 | |
	R^2	NMAE	Time (s)	Epochs	R^2	NMAE	R^2	NMAE
All–x_c	0.9999	0.0034	4	34	0.9999	0.0038	0.9998	0.0062
Rsq–x_c	0.9999	0.0034	2	37	0.9999	0.0037	0.9999	0.0049
$AvSlh$–x_c	0.9999	0.0034	2	15	0.9999	0.0034	0.9999	0.0050
All–$\ddot{x}_c$	0.9987	0.0143	1	17	0.9974	0.0208	0.9930	0.0357
Rsq–$\ddot{x}_c$	0.9988	0.0140	0	35	0.9972	0.0191	0.9954	0.0290
$AvSlh$–$\ddot{x}_c$	0.9992	0.0115	1	135	0.9989	0.0129	0.9878	0.0123

Table 7: Performances of SVR network on simulation feature data.

| Simulation feature data | Training | | | Testing | | | |
| | SimDat-1 | | | SimDat-2 | | SimDat-3 | |
	R^2	NMAE	Time (s)	R^2	NMAE	R^2	NMAE
All–x_c	0.9999	0.0044	139.4	0.9996	0.0067	0.9978	0.0172
Rsq–x_c	0.9999	0.0032	106.5	0.9999	0.0038	0.9992	0.0104
$AvSlh$–x_c	0.9999	0.0035	73.1	0.9999	0.0036	0.9994	0.0093
All–$\ddot{x}_c$	0.9984	0.0157	71.6	0.9972	0.0214	0.9969	0.0268
Rsq–$\ddot{x}_c$	0.9969	0.0216	34.5	0.9958	0.0268	0.9943	0.0317
$AvSlh$–$\ddot{x}_c$	0.9986	0.0135	68.1	0.9975	0.0193	0.9874	0.0573

sets are presented in Figs. 11 and 12, respectively. Typical distribution of prediction errors from networks built with x_c and $\ddot{x}_c$ feature data are respectively shown in Figs. 13(a) and (b). For both networks, a greater percentage of the errors were found to be distributed around the zero error line. However, it is observed that the $\ddot{x}_c$-based network showed higher and wider range errors ranging between -18.4 to 16.59, compared to the x_c-based network whose errors were lower and ranged between -1.97 to 3.23.

The results of using the MLP and SVR networks on experimental data in the same manner as described above are presented in Table 8 and 9, respectively. As earlier stated, only x_c signals could be measured during the experiment. Similar to simulation results, the MLP network showed better performances on the out-of-sample test data compared to the SVR. However, both networks performed best when used alongside the Rsq–x_c feature data. The actual and predicted E-values for the experimental Rsq–x_c feature data during training and testing are present in Fig. 14 for both the MLP and SVR networks. Their respective histogram of errors over 20 bins for the test data sets are shown in Fig. 15. It can be seen that the SVR once again showed higher and wider range errors ranging between -8.19 to 12.68 compared to the MLP whose errors ranged between -1.84 to 1.36.

5.2 Stage 2: Biomechanical stiffness categorisation

The second stage of the proposed bowel lesions characterisation involves an unsupervised categorisation of the lesions into multiple subsets or clusters such that
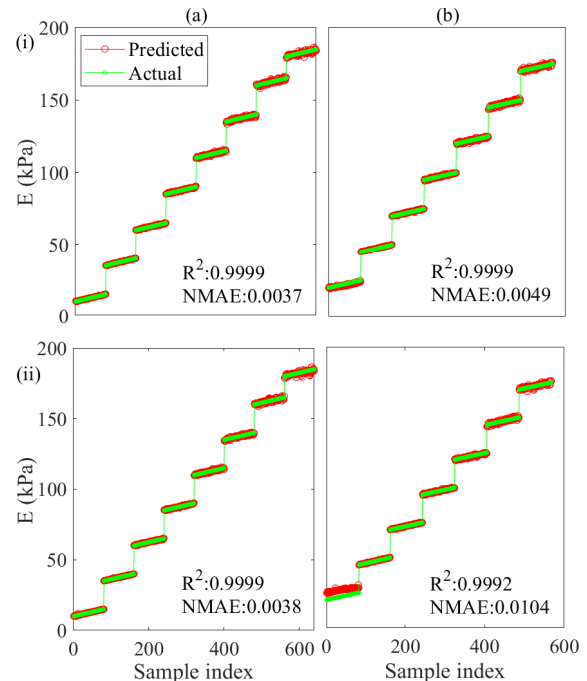

Fig. 11: Actual-vs-predicted plots for (i) MLP and (ii) SVR on Rsq–x_c feature data from (a) SimDat-2 and (b) SimDat-3.

Table 8: MLP performances on experimental feature data.

| Experimental feature data | Training (ExpDat-1) | | Testing (ExpDat-2) | |
	R^2	NMAE	R^2	NMAE
All–x_c	0.9990	0.0021	0.9982	0.0045
Rsq–x_c	0.9999	0.0014	0.9997	0.0021
$AvSlh$–x_c	0.9997	0.0022	0.9995	0.0028

lesions within a cluster have high biomechanical similarity, but very dissimilar to those in other clusters. This
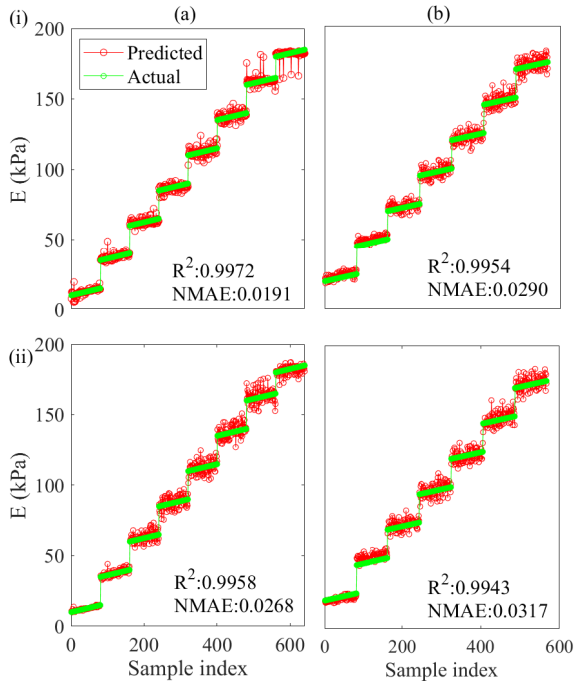
Fig. 12: Actual-vs-predicted plots for (i) MLP and (ii) SVR on Rsq–$\ddot{x}_c$ feature data from (a) SimDat-2 and (b) SimDat-3.
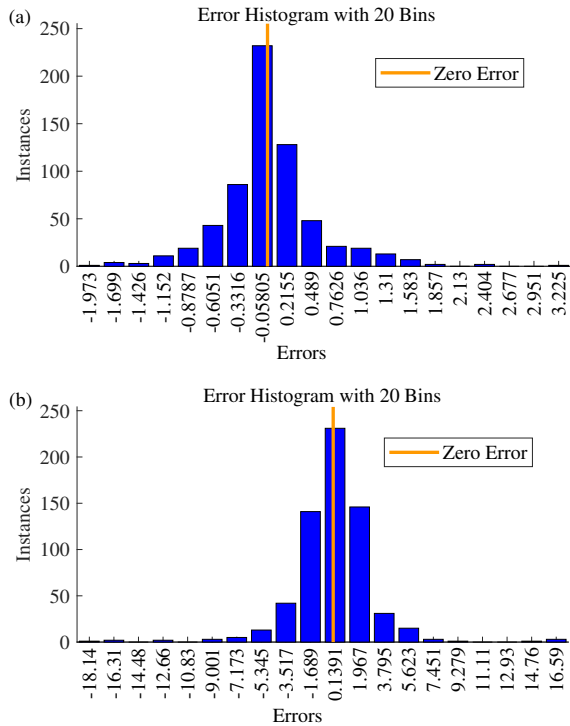
Table 9: SVR performances on experimental feature data.

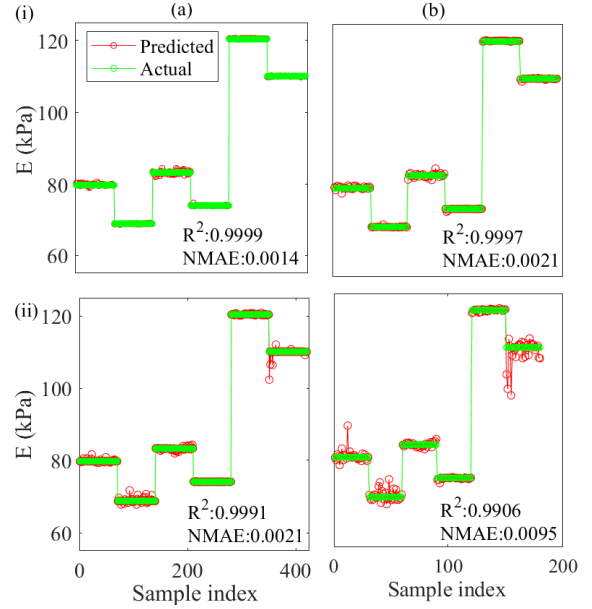| Experimental feature data	Training (ExpDat-1)		Testing (ExpDat-2)	
	R^2	NMAE	R^2	NMAE
All–x_c	0.9991	0.0024	0.9904	0.0117
Rsq–x_c	0.9991	0.0021	0.9906	0.0095
$AvSlh$–x_c	0.9985	0.0072	0.9886	0.0132

Fig. 14: Actual-vs-predicted plots for (i) MLP and (ii) SVR on experimental Rsq–x_c feature data during (a) training and (b) testing.
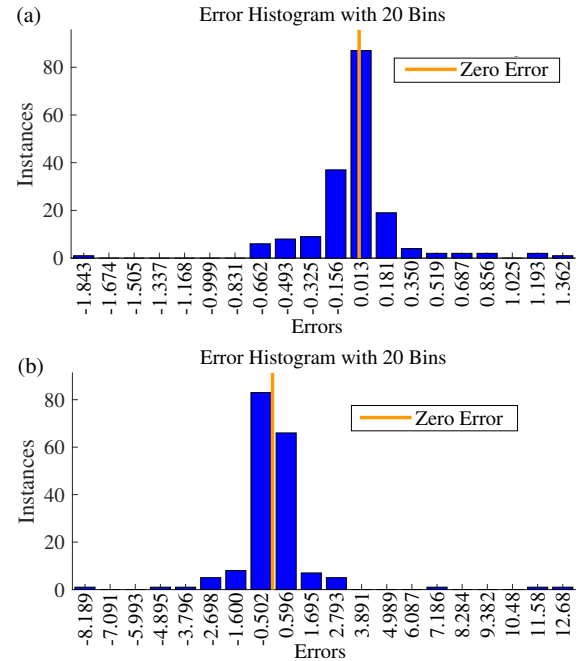

Fig. 13: Typical distribution of prediction errors resulting from networks built with simulation (a) x_c and (b) $\ddot{x}_c$ feature data.

Fig. 15: Histogram of the errors from testing (a) MLP and (b) SVR on experimental Rsq–x_c feature data.

way the less stiffer benign lesions are well differentiated from the stiffer malignant lesions of different cancer stages. An unsupervised categorisation have been adopted for this final stage to allow a more generalised

qualitative characterisation rather than a quantitative characterisation. This thus annuls the problem of vary-

ing background stiffness of the bowel across patients of different age, ethnicity and diet. The similarities or disimilarities between the lesions is assessed based on the regression models' predicted E-values. Each resulting cluster thus comprises E-values with closer interpoint distances compared to those in other clusters. For this study, K-means clustering as available in MATLAB Statistics and Machine Learning Toolbox was utilised for the unsupervised clustering. For N observed data samples, the two sets of N-predicted E-values arising from using MLP and SVR on each selected feature data (see Table 5) are presented to the clustering algorithm as a N-by-2 data matrix. This way, the prediction power of both the MLP and SVR models are combined to cluster the lesions into different categorical ranges. With clustering being a learning by observation approach rather than learning from examples, effort is often made at selecting efficient and effective parameters. Such parameters include the optimal k value and the distance metric that perfectly cluster the data into their correct categorical ranges. For the simulated and experimental test data, including SimDat-2, SimDat-3 and ExpDat-2, their categorical ranges as well as their optimal k values are observably 8, 7 and 6, respectively. However, assuming that these optimal k values are unknown, silhouette evaluation criterion described in Section 3.2.3 is used to determine the optimal k-values. For each of the test data, k is varied from 2 to 10 during silhouette evaluation and the k value that yields the maximum $AvSlh$ is taken as the optimal k value. The maximum $AvSlh$ defined k values were found to be consistent with the known categorical ranges of the data sets as presented in Fig. 16. By trial and error, the Squared Euclidean as a distance metric was found to yield the best clustering results, hence, was adopted for the clustering.

The results of the K-means clustering of the N-by-2 MLP-SVR predicted E-values are shown in Table 10 and 11 for both the simulated and experimental data, respectively. The results have been reported based on the resulting average silhouette (AvSlh) values and the percentage of the rightly clustered samples for each feature data (i.e. accuracy (%)). For the simulated data sets, the E-values predicted from x_c-based features showed optimal clustering with AvSlh values ≥ 0.99 and 100 % accuracies, outperforming those predicted from the $\ddot{x}_c$-based features. It was also observed that most of the simulation data clustering yielded AvSlh values greater than 0.90 and accuracies greater than 99 %. Figure 17 shows the clustering of the MLP-SVR predicted E-values using Rsq–x_c and Rsq–$\ddot{x}_c$ feature data from SimDat-1, SimDat-2 and SimDat-3. The wrongly clustered samples are indicated with red circles and this
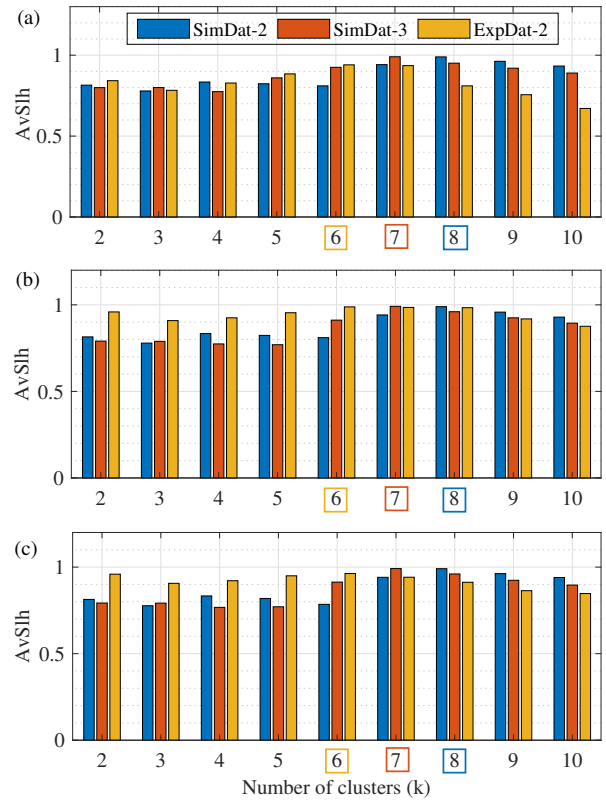

Fig. 16: Resulting $AvSlh$ for different k-values for SimDat-2, SimDat-3 and ExpDat-2 using (a) ALL–x_c (b) Rsq–x_c and (c) $AvSlh$–x_c feature data.

mostly occurred when there is a narrow margin between the network predicted E-values for samples of different categories. The E-values predicted for different sample categories using x_c-based feature data were found to exhibit wider margin compared to those predicted using $\ddot{x}_c$-based feature data.

The results of clustering the MLP-SVR predicted E-values from experimental feature data is shown in Table 11. The best clustering is seen to be obtained using the Rsq–x_c feature data yielding Avslh values of 0.988 and 0.931, and accuracies of 100 % and 99.4 % for the experimental ExpDat-1 and ExpDat-2 data, respectively. For the experimental validation, aside the clustering of the MLP-SVR predicted E-values, the clustering of individual network predicted E-values was also carried. This is reported in Table 12 and 13 for both the MLP and SVR predictions, respectively. The MLP predicted E-values are seen to show better clustering and will probably be the first-line network for the proposed lesion characterisation compared to SVR. Figures 18 and 19 show the clustering results of the MLP-SVR, MLP and SVR predicted E-values using Rsq–x_c feature data. In both cases, the wrongly clustered samples have also been indicated with red circles.

Table 10: Clustering results of the MLP-SVR predicted E-values from simulation data.

Simulation feature data	Training		Testing			
	SimDat-1		SimDat-2		SimDat-3	
	AvSlh value	Accuracy (%)	AvSlh value	Accuracy (%)	AvSlh value	Accuracy (%)
All–x_c	0.990	100	0.990	100	0.990	100
Rsq–x_c	0.990	100	0.989	100	0.991	100
$AvSlh$–x_c	0.990	100	0.991	100	0.991	100
All–$\ddot{x}_c$	0.977	100	0.959	99.8	0.933	99.8
Rsq–$\ddot{x}_c$	0.961	99.8	0.946	99.1	0.936	100
$AvSlh$–$\ddot{x}_c$	0.980	99.9	0.972	99.7	0.858	97.4
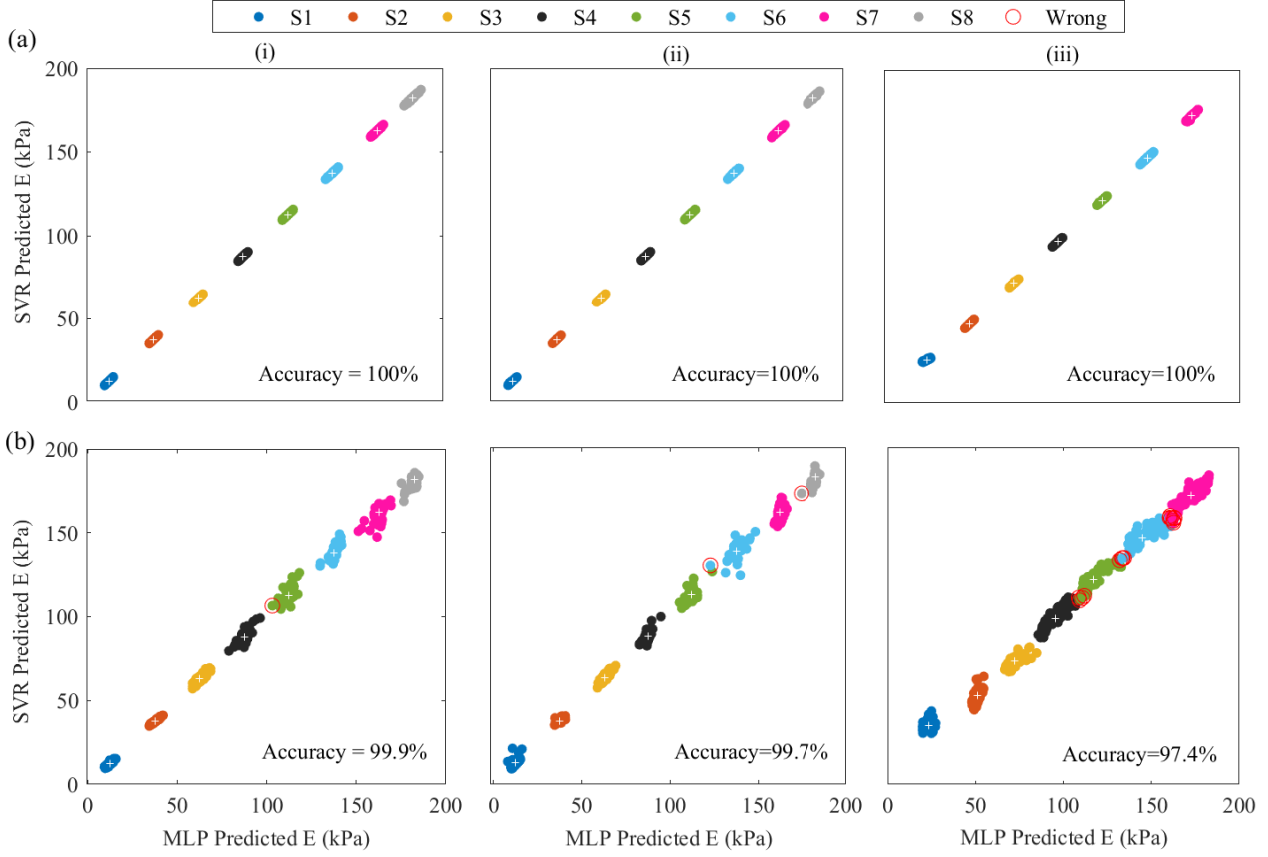
Fig. 17: Clustering results of the MLP-SVR predicted E-values using (a) Rsq–x_c and (b) Rsq–$\ddot{x}_c$ feature data from (i) SimDat-1 (ii) SimDat-2 (iii) SimDat-3.
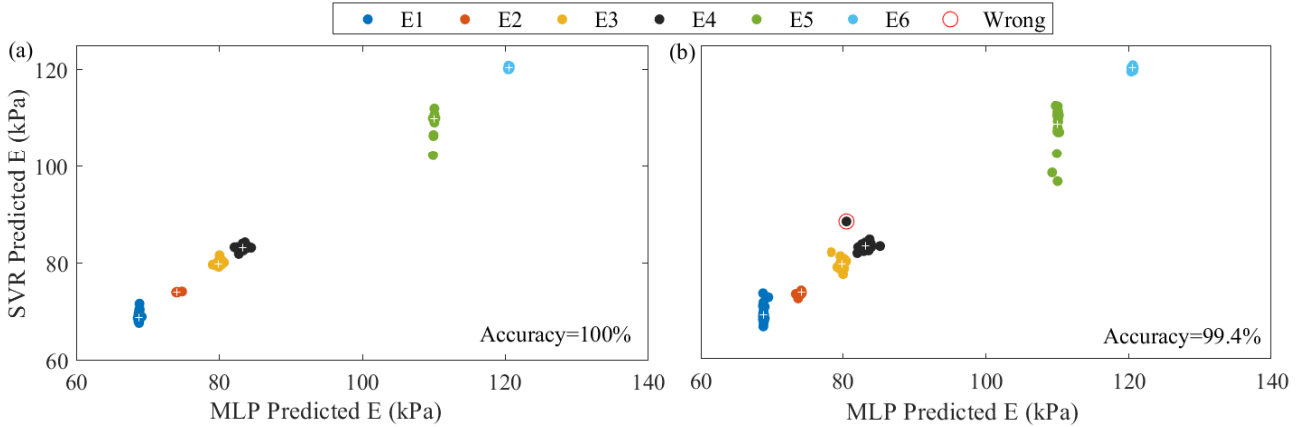
Fig. 18: Clustering of the MLP-SVR predicted E-values using Rsq–x_c feature data from (a) ExpDat-1 and (b) ExpDat-2.
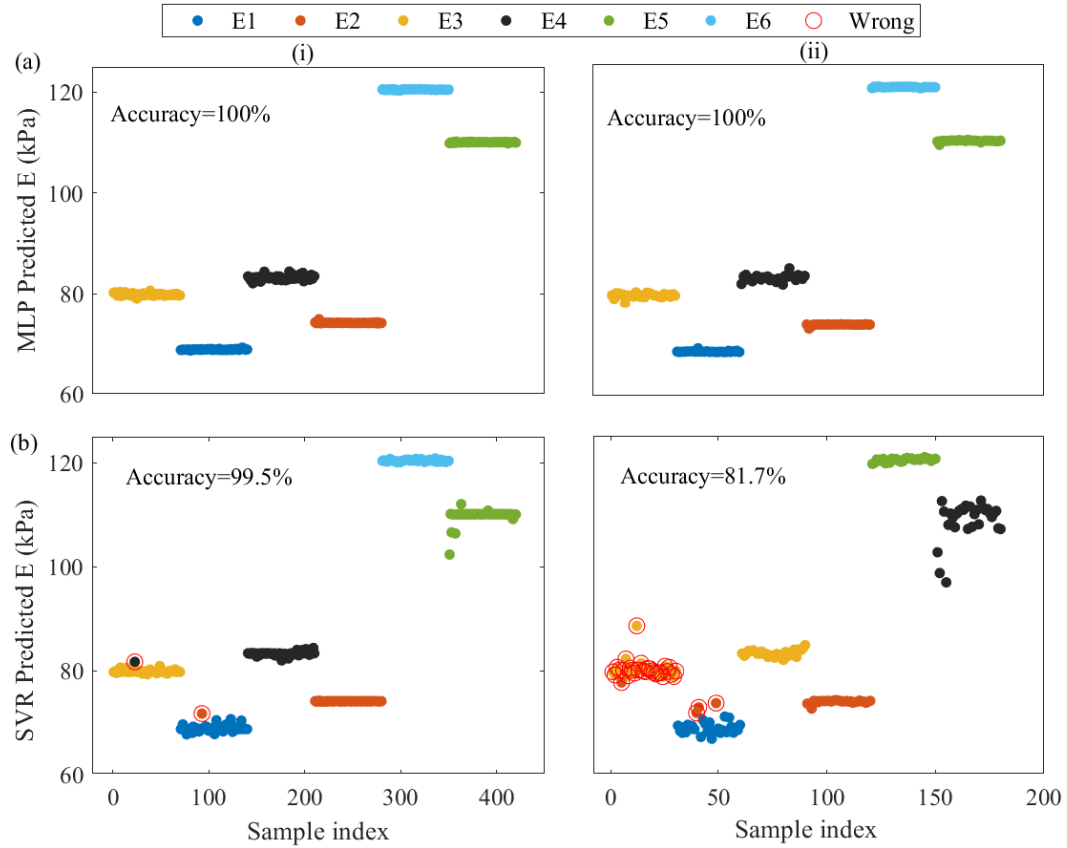
Fig. 19: Clustering of the (a) MLP and (b) SVR predicted E-values using Rsq–x_c feature from (i) ExpDat-1 and (ii) ExpDat-2.

Table 11: Clustering results of the MLP-SVR predicted E-values from experimental feature data.

	Training		Testing	
Experimental	ExpDat-1		ExpDat-2	
feature data	AvSlh value	Accuracy (%)	AvSlh value	Accuracy (%)
All–x_c	0.976	99.5	0.907	98.9
Rsq–x_c	0.988	100	0.931	99.4
$AvSlh$–x_c	0.963	99.8	0.887	97.2

Table 13: Clustering results of the SVR predicted E-values from experimental feature data.

	Training		Testing	
Experimental	ExpDat-1		ExpDat-2	
feature data	AvSlh value	Accuracy (%)	AvSlh value	Accuracy (%)
All–x_c	0.988	99.8	0.870	97.2
Rsq–x_c	0.983	99.5	0.903	81.7
$AvSlh$–x_c	0.950	100	0.881	78.3

Table 12: Clustering results of the MLP predicted E-values from experimental feature data.

	Training		Testing	
Experimental	ExpDat-1		ExpDat-2	
feature data	AvSlh value	Accuracy (%)	AvSlh value	Accuracy (%)
All–x_c	0.961	99.5	0.947	99.4
Rsq–x_c	0.992	100	0.982	100
$AvSlh$–x_c	0.979	99.1	0.968	100

6 Conclusions

For the purpose of early bowel cancer diagnosis using the SP-VIC, a two-stage machine learning (ML) procedure has been explored. As the first stage of the procedure, multi-layer perceptron (MLP) and support vector regression (SVR) networks have been used to predict lesion stiffnesses from measurable capsule dynamics including acceleration and displacement. Using K-means clustering, an unsupervised categorisation of predicted stiffnesses into clusters of high intra-cluster similarity but low inter-cluster similarity has been carried out as the second stage of the procedure. Based on the performance metrics including coefficient of determination (R^2) and normalised mean absolute error (NMAE), the MLP models showed better perfor-

mances on the test data compared to the SVR. For situations where both displacement and acceleration signals are measurable, the displacement-based models outperformed the acceleration-based models. For both the simulation and experimental test data, the displacement-based MLPs were found to achieved R^2 values greater than 0.9980 and NMAE values lesser than 0.0062. During the experimental validation, the predicted E-values from the displacement-based MLPs outperformed the displacement-based SVRs yielding clustering accuracies and average silhouette (AvSlh) values of at least 99.4% and 0.947, respectively. This makes the displacement-based MLPs more preferable for the proposed bowel lesion characterisation. On the overall, both the simulation and experimental validation indicate that the proposed method has the capacity to characterise bowel lesions as low as 2 mm. Thus indicating the huge potential of the method in revealing hard-to-visualise bowel cancers for the purpose of improved treatment and increased survival rate. Based on this current study, the detection of lesions ≤ 2 mm will probably require a smaller sized SP-VIC. It was also noticed that the use of Eco-flex as the intestinal material further subjected the capsule to intense frictional reaction. As future works, we hope to establish a threshold for which the difference between the E-values of resulting clusters raises a concern while also carrying out an in-vivo testing of the proposed method.

Compliance with ethical standards

Conflict of interest

The authors declare that they have no conflict of interest concerning the publication of this manuscript.

Data accessibility

The data sets generated and analysed during the current study are available from the corresponding author on reasonable request.

References

1. Afebu, K.O., Tian, J., Liu, Y., Papatheou, E., Prasad, S.: AI-assisted dynamic tissue evaluation for early bowel cancer diagnosis using a vibrational capsule. IEEE Robotics and Automation Letters (2023). In print

2. Akhtar, R., Sherratt, M.J., Cruickshank, J.K., Derby, B.: Characterizing the elastic properties of tissues. Materials Today **14**(3), 96–105 (2011)

3. Ansari, H.R., Gholami, A.: An improved support vector regression model for estimation of saturation pressure of crude oils. Fluid Phase Equilibria **402**, 124–132 (2015)

4. AnX Robotica: Navicam small bowel capsule-system. `https://www.anxrobotics.com/products/navicam-sb-capsule-system/` (2022). [Accessed: 11.05.2022]

5. Banerjee, A., Pasupuleti, S., Mondal, K., Nezhad, M.M.: Application of data driven machine learning approach for modelling of non-linear filtration through granular porous media. International Journal of Heat and Mass Transfer **179**, 121650 (2021)

6. Bannister, N., Broggio, J.: Cancer survival by stage at diagnosis for England (experimental statistics): adults diagnosed 2012, 2013 and 2014 and followed up to 2015. URL `https://www.gov.uk/search/research-and-statistics`

7. Battista, A., Battista, R.A., Battista, F., Iovane, G., Landi, R.E.: Bh-index: A predictive system based on serum biomarkers and ensemble learning for early colorectal cancer diagnosis in mass screening. Computer Methods and Programs in Biomedicine **212**, 106494 (2021)

8. Bishop, C.M., Nasrabadi, N.M.: Pattern recognition and machine learning, vol. 4. Springer (2006)

9. Bowel Cancer UK: Bowel cancer. URL `https://www.bowelcanceruk.org.uk/about-bowel-cancer/bowel-cancer/`

10. Brás, M.M., Sousa, S.R., Carneiro, F., Radmacher, M., Granja, P.L.: Mechanobiology of colorectal cancer. Cancers **14**(8), 1945 (2022)

11. Brauchle, E., Kasper, J., Daum, R., Schierbaum, N., Falch, C., Kirschniak, A., Schäffer, T.E., Schenke-Layland, K.: Biomechanical and biomolecular characterization of extracellular matrix structures in human colon carcinomas. Matrix Biology **68**, 180–193 (2018)

12. Cancer Research UK: Bowel cancer mortality statistics. URL `https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/mortality#heading-Three`

13. Cancer Research UK: Tests for bowel cancer. URL `https://www.cancerresearchuk.org/about-cancer/bowel-cancer/getting-diagnosed/tests`. Accessed: 11-05-2022

14. Capsovision Inc.: See More - Experience the 360˚ difference with CapsoCam Plus. https://capsovision.com/capsocam-system/ (2022). [Accessed: 11.05.2022]

15. Chang, Y.W., Hsieh, C.J., Chang, K.W., Ringgaard, M., Lin, C.J.: Training and testing low-degree polynomial data mappings via linear svm. Journal of Machine Learning Research **11**(4) (2010)

16. Check-Cap Ltd: The C-scan system. https://check-cap.com/the-c-scan-system/ (2022). [Accessed: 11.05.2022]

17. Cohn, R., Holm, E.: Unsupervised machine learning via transfer learning and k-means clustering to classify materials image data. Integrating Materials and Manufacturing Innovation **10**(2), 231–244 (2021)

18. Cortes, C., Vapnik, V.: Support-vector networks. Machine learning **20**(3), 273–297 (1995)

19. Dekker, E., Tanis, P.J., Vleugels, J.L.A., Kasi, P.M., Wallace, M.B.: Colorectal cancer. The Lancet **394**, 1467–1480 (2019)

20. Deptula, P., Lysik, D., Pogoda, K., Cieśluk, M., Namiot, A., Mystkowska, J., Król, G., Gluszek, S., Janmey, P.A., Bucki, R.: Tissue rheology as a possible complementary procedure to advance histological diagnosis of colon cancer. ACS Biomaterials Science & Engineering **6**(10), 5620–5631 (2020)

21. D'Orazio, M., Murdocca, M., Mencattini, A., Casti, P., Filippi, J., Antonelli, G., Di Giuseppe, D., Comes, M., Di Natale, C., Sangiuolo, F., et al.: Machine learning phenomics (mlp) combining deep learning with time-lapse-microscopy for monitoring colorectal adenocarcinoma cells gene expression and drug-response. Scientific Reports **12**(1), 1–14 (2022)

22. Ferlay, J., Ervik, M., Lam, F., Colombet, M., Mery Land Piñeros, M., Znaor, A., Soerjomataram, I., Bray, F.: Global cancer observatory: Cancer today. URL https://gco.iarc.fr/today

23. Goodfellow, I., Bengio, Y., Courville, A.: Deep learning. MIT press (2016)

24. Groth, S., Krause, H., Behrendt, R., Hill, H., Börner, M., Bastürk, M., Plathner, N., Schütte, F., Gauger, U., Riemann, J.F., et al.: Capsule colonoscopy increases uptake of colorectal cancer screening. BMC Gastroenterology **12**(1), 1–7 (2012)

25. Guo, B., Ley, E., Tian, J., Zhang, J., Liu, Y., Prasad, S.: Experimental and numerical studies of intestinal frictions for propulsive force optimisation of a vibro-impact capsule system. Nonlinear Dynamics **101**, 65–83 (2020)

26. Hagan, M.T., Menhaj, M.B.: Training feedforward networks with the marquardt algorithm. IEEE transactions on Neural Networks **5**(6), 989–993 (1994)

27. Han, J., Kamber, M.: Data mining: concepts and techniques, 2nd. University of Illinois at Urbana Champaign: Morgan Kaufmann pp. 451–452 (2006)

28. Han, J., Pei, J., Tong, H.: Data mining: concepts and techniques. Morgan kaufmann (2022)

29. Hemamalini, V., Rajarajeswari, S., Nachiyappan, S., Sambath, M., Devi, T., Singh, B.K., Raghuvanshi, A.: Food quality inspection and grading using efficient image segmentation and machine learning-based system. Journal of Food Quality **2022** (2022)

30. IntroMedic Co., L.: MiroCam - Capsule endoscope system. http://www.intromedic.com/eng/item/item_010100_view.asp?search_kind=&gotopage=1&no=3 (2022). [Accessed: 11.05.2022]

31. JIFU Medical Technology Co., L.: Dasheng Capsule Endoscopic System. http://www.jifu-tech.com/ (2022). [Accessed: 11.05.2022]

32. Jimhans Medical: Capsule Endoscopic System. http://jimhans.com/?product=capsubot (2022). [Accessed: 11.05.2022]

33. Jinshan Group: Capsule Endoscopy. https://www.jinshangroup.com/solutions/capsule-endoscopy-camera/ (2022). [Accessed: 11.05.2022]

34. Kawano, S., Kojima, M., Higuchi, Y., Sugimoto, M., Ikeda, K., Sakuyama, N., Takahashi, S., Hayashi, R., Ochiai, A., Saito, N.: Assessment of elasticity of colorectal cancer tissue, clinical utility, pathological and phenotypical relevance. Cancer science **106**(9), 1232–1239 (2015)

35. Kjølhede, T., Ølholm, A.M., Kaalby, L., Kidholm, K., Qvist, N., Baatrup, G.: Diagnostic accuracy of capsule endoscopy compared with colonoscopy for polyp detection: systematic review and meta-analyses. Endoscopy **53**(07), 713–721 (2021)

36. Li, H., Lin, J., Xiao, Y., Zheng, W., Zhao, L., Yang, X., Zhong, M., Liu, H.: Colorectal cancer detected by machine learning models using conventional laboratory test data. Technology in Cancer Research & Treatment **20**, 15330338211058352 (2021)

37. Liu, Y., Páez Chávez, J.: Controlling multistability in a vibro-impact capsule system. Nonlinear Dynamics **88**, 1289–1304 (2017)

38. Liu, Y., Paez Chavez, J., Zhang, J., Tian, J., Guo, B., Prasad, S.: The vibro-impact capsule system in millimetre scale: numerical optimisation and experimental verification. Meccanica **55**(10), 1885–1902 (2020)

39. Liu, Y., Wiercigroch, M., Pavlovskaia, E., Yu, H.: Modelling of a vibro-impact capsule system. International Journal of Mechanical Sciences **66**, 2–11 (2013)

40. Lloyd, S.: Least squares quantization in pcm. IEEE transactions on information theory **28**(2), 129–137 (1982)

41. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, vol. 5.1, pp. 281–297. Univ of California Press (1967)

42. Mathworks UK: Feature extraction for machine learning and deep learning. URL https://uk.mathworks.com/discovery/feature-extraction.html

43. Mathworks UK: K-means clustering. URL https://uk.mathworks.com/help/stats/kmeans.html

44. Mathworks UK: Levenberg-marquardt algorithm. URL https://uk.mathworks.com/help/deeplearning/ref/trainlm.html

45. Medtronic Plc: Capsule endoscopy (2022). URL https://www.medtronic.com/covidien/en-us/products/capsule-endoscopy.html

46. Murtagh, F.: Multilayer perceptrons for classification and regression. Neurocomputing **2**(5-6), 183–197 (1991)

47. Nagelkerke, N.J., et al.: A note on a general definition of the coefficient of determination. biometrika **78**(3), 691–692 (1991)

48. Nawaz, M., Mehmood, Z., Nazir, T., Naqvi, R.A., Rehman, A., Iqbal, M., Saba, T.: Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering. Microscopy Research and Technique **85**(1), 339–351 (2022)

49. Nuffield Trust UK: Cancer survival rates. URL https://www.nuffieldtrust.org.uk/resource/cancer-survival-rates#background

50. Olympus Corp.: ENDOCAPSULE 10 System — Small Bowel Capsule Endoscopy System. https://www.olympus.co.uk/medical/en/Products-and-solutions/Products/Capsule-Endoscopy.htm (2022). [Accessed: 11.05.2022]

51. Palmieri, V., Lucchetti, D., Maiorana, A., Papi, M., Maulucci, G., Ciasca, G., Svelto, M., De Spirito, M., Sgambato, A.: Biomechanical investigation of colorectal cancer cells. Applied Physics Letters **105**(12), 123701 (2014)

52. Pavlovskaia, E., Hendry, D.C., Wiercigroch, M.: Modelling of high frequency vibro-impact drilling. International Journal of Mechanical Sciences **91**, 110–119 (2015)

53. Pavlovskaia, E., Wiercigroch, M., Grebogi, C.: Modeling of an impact system with a drift. Physical Review E **64**, 056224 (2001)

54. Pavlovskaia, E., Wiercigroch, M., Woo, K.C., Rodger, A.A.: Modelling of ground moling dynamics by an impact oscillator with a frictional slider. Meccanica **38**, 85–97 (2003)

55. Pratiwi, N.K.C., Magdalena, R., Fuadah, Y.N., Saidah, S.: K-nearest neighbor for colon cancer identification. In: Journal of Physics: Conference Series, vol. 1367, p. 012023. IOP Publishing (2019)

56. Puyal, J.G.B., Brandao, P., Ahmad, O.F., Bhatia, K.K., Toth, D., Kader, R., Lovat, L., Mountney, P., Stoyanov, D.: Polyp detection on video colonoscopy using a hybrid 2d/3d cnn. Medical Image Analysis p. 102625 (2022)

57. Quirini, M., Menciassi, A., Scapellato, S., Dario, P., Rieber, F., Ho, C.N., Schostek, S., Schurr, M.O.: Feasibility proof of a legged locomotion capsule for the GI tract. Gastrointestinal endoscopy **67**(7), 1153–1158 (2008)

58. Ranade, M.: Classification and prediction of severity of inflammatory bowel disease using machine learning. In: 2021 6th International Conference for Convergence in Technology (I2CT), pp. 1–4. IEEE (2021)

59. Sharma, P., Bora, K., Kasugai, K., Balabantaray, B.K.: Two stage classification with cnn for colorectal cancer detection. Oncologie **22**(3) (2020)

60. Shayesteh, S., Nazari, M., Salahshour, A., Sandoughdaran, S., Hajianfar, G., Khateri, M., Yaghobi Joybari, A., Jozian, F., Fatehi Feyzabad, S.H., Arabi, H., et al.: Treatment response prediction using mri-based pre-, post-, and delta-radiomic features and machine learning algorithms in colorectal cancer. Medical Physics **48**(7), 3691–3701 (2021)

61. Shirazy, A., Hezarkhani, A., Shirazi, A., Khakmardan, S., Rooki, R.: K-means clustering and general regression neural network methods for copper mineralization probability in chahar-farsakh, iran. Türkiye Jeoloji Bülteni **65**(1), 79–92 (2022)

62. Simi, M., Valdastri, P., Quaglia, C., Menciassi, A., Dario, P.: Design, fabrication, and testing of a capsule with hybrid locomotion for gastrointestinal tract exploration. IEEe/ASME Transactions on Mechatronics **15**(2), 170–180 (2010)

63. Stevenson, R.G.: An investigation into the use of vibro-impact techniques in the design of a ground moling system. University of Aberdeen (United Kingdom) (1992)

64. Vapnik, V.N.: The nature of statistical learning. Theory (1995)

65. Weigt, J., Repici, A., Antonelli, G., Afifi, A., Kliegis, L., Correale, L., Hassan, C., Neumann, H.: Performance of a new integrated computer-assisted system (cade/cadx) for detection and characterization of colorectal neoplasia. Endoscopy **54**(02), 180–184 (2022)

66. Wu, F., Zhu, C., Xu, J., Bhatt, M.W., Sharma, A.: Research on image text recognition based on canny edge detection algorithm and k-means algorithm. International Journal of System Assurance Engineering and Management **13**(1), 72–80 (2022)

67. Xi, Y., Xu, P.: Global colorectal cancer burden in 2020 and projections to 2040. Translational Oncology **14**(10), 101174 (2021)

68. Xiao, Y.F., Wu, Z.X., He, S., Zhou, Y.Y., Zhao, Y.B., He, J.L., Peng, X., Yang, Z.X., Lv, Q.J., Yang, H., et al.: Fully automated magnetically controlled capsule endoscopy for examination of the stomach and small bowel: a prospective, feasibility, two-centre study. The Lancet Gastroenterology & Hepatology **6**(11), 914–921 (2021)

69. Yan, Y., Liu, Y., Páez Chávez, J., Zonta, F., Yusupov, A.: Proof-of-concept prototype development of the self-propelled capsule system for pipeline inspection. Meccanica **53**, 1997–2012 (2018)

70. Yan, Y., Zhang, B., Liu, Y., Prasad, S.: Dynamics of a vibro-impact self-propelled capsule encountering a circular fold in the small intestine. Meccanica pp. 1–22 (2022)

71. Yang, J.O., Galoosian, A., Peterson, E.K., Soroudi, C., Myint, A., Hsu, W., Maehara, C.K., Kang, Y., Naini, B.V., Muthusamy, V.R., et al.: 542: Validation of a machine learning algorithm to measure bowel preparation quality for screening colonoscopy in a large health system. Gastroenterology **162**(7), S–128 (2022)

72. Yang, M., Yang, H., Ji, L., Hu, X., Tian, G., Wang, B., Yang, J.: A multi-omics machine learning framework in predicting the survival of colorectal cancer patients. Computers in Biology and Medicine **146**, 105516 (2022)

73. Yuce, M.R., Alici, G., Than, T.D.: Wireless endoscopy. Wiley Encyclopedia of Electrical and Electronics Engineering pp. 1–25 (1999)

74. Zhang, J., Liu, Y., Zhu, D., Prasad, S., Liu, C.: Simulation and experimental studies of a vibro-impact capsule system driven by an external magnetic field. Nonlinear Dynamics **109**(3), 1501–1516 (2022)

75. Zhao, D., Liu, H., Zheng, Y., He, Y., Lu, D., Lyu, C.: A reliable method for colorectal cancer prediction based on feature selection and support vector machine. Medical & biological engineering & computing **57**(4), 901–912 (2019)

76. Zhao, M., Tang, Y., Kim, H., Hasegawa, K.: Machine learning with k-means dimensional reduction for predicting survival outcomes in patients with breast cancer. Cancer informatics **17** (2018)

77. Zhou, P., Cao, Y., Li, M., Ma, Y., Chen, C., Gan, X., Wu, J., Lv, X., Chen, C.: Hccanet: histopathological image grading of colorectal cancer using cnn based on multichannel fusion attention mechanism. Scientific reports **12**(1), 1–12 (2022)