In-vitro study on the identification of gastrointestinal stromal tumor tissues using laser-induced breakdown spectroscopy with chemometric methods

BUSHRA SANA IDREES,1,2 QIANQIAN WANG,1,2,3,* M. NOUMAN KHAN,1,2 GEER TENG,1,2,3 XUTAI CUI,1,2 WENTING XIANGLI,1,2 AND KAI WEI1,2

1School of Optics and Photonics, Beijing Institute of Technology, Beijing, 100081, China
2Key Laboratory of Photonic Information Technology, Ministry of Industry and Information Technology, Beijing Institute of Technology, 100081 Beijing, China
3Yangtze Delta Region Academy of Beijing Institute of Technology, Jiaxing, 314033, China

*qqwang@bit.edu.cn

Abstract: Early-stage detection of tumors helps to improve patient survival rate. In this work, we demonstrate a novel discrimination method to diagnose the gastrointestinal stromal tumor (GIST) and its healthy formalin fixed paraffin embedded (FFPE) tissues by combining chemometric algorithms with laser-induced breakdown spectroscopy (LIBS). Chemometric methods which include partial least square discrimination analysis (PLS-DA), k-nearest neighbor (k-NN) and support vector machine (SVM) were used to build the discrimination models. The comparison of PLS-DA, k-NN and SVM classifiers shows an increase in accuracy from 94.44% to 100%. The comparison of LIBS signal between the healthy and infected tissues shows an enhancement of calcium lines which is a signature of the presence of GIST in the FFPE tissues. Our results may provide a complementary method for the rapid detection of tumors for the successful treatment of patients.

© 2021 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

Gastrointestinal stromal tumors (GISTs) or stomach cancer continue to be the second leading cause of cancer-associated deaths all over the world [1]. GIST’s can arise anywhere in the digestive tract and are firstly supposed that they initiate through gastrointestinal stromal cells; therefore, it comes to know that this kind of tumors exist from interstitial cells of Cajal [2]. Current epidemiology displays that every year about 0.70 per 100 000 people in the United States are affected by GIST, and the trend is increasing every year [3]. It has long been thought that gastric cancer is a destructive disease with a five-year existence rate less than 30% [4]. In last few decades, the incidence and mortality rate increased gradually and it becomes the sixth most common cancer disease that cause deaths in the developed countries [5–7].

Early detection of gastric cancer can effectively control the spreading of malignancy and decrease the mortality in the world. Diagnosis of gastrointestinal stromal tumor depends on qualitative histological evaluation of tissue samples. This method is time taking and relies on the histopathology’s interpretation. It is very difficult to identify or diagnose the gastric cancer, especially at early stages using state-of-the-art methods such as computed tomography virtual gastroscopy (CTVG), fibro gastroscopy, three-dimensional imaging by multi slice spiral CT (3DMSCCT) and ultrasonic gastroscopy [8]. Moreover, some emerging optical techniques such as optical coherence tomography (OCT), near-infrared spectroscopy (NIRS), autofluorescence imaging (FAF) and Raman spectroscopy (RS) have been used for diagnoses and classifications of different malignant tissues [9]. The OCT is an imaging technique that probes changes in mark
biological tissue structure for infections [10]. Guo et al. used Raman and surface-enhanced Raman spectroscopies to diagnose the gastric cancer with 94.1% and 98.5% accuracies using principal component analysis and characteristic ratio method, respectively [11]. However, In Raman spectroscopy, Raman signals are still not much robust and it brings exogenic nanoparticles for the surface-enhanced Raman spectroscopy (SERS). Huang et al. used fiber-optic Raman endoscopy to find the primary lesions of stomach for in vivo differentiation of gastric cancer tissue from its healthy tissue [12]. In this technique, it is hard to capture the photos by injecting camera, because of tiny space in stomach. Further, blood loss and cerebrospinal fluid affect the imaging results. Therefore, there is an urgent need to find experimental techniques for the rapid detection of gastrointestinal stromal tumor.

Laser-induced breakdown spectroscopy (LIBS) is an attractive technique due to high resolution and real-time assessment of multi-elements detection [13,14]. The LIBS has been used for the detection and classification of different kinds of biomedical species, such as bio-aerosols, calcified materials and soft tissues [15]. El-Hussein et al. performed an experiment for the discrimination and classification of colorectal and breast cancers using LIBS, and showed that malignant tissues contained calcium and magnesium in higher access compared to its healthy ones [16]. Recently, Teng et al. proposed LIBS as a potential diagnostic tool to discriminate between glioma and infiltrating boundary with k-NN and SVM [17]. LIBS experiment was performed to discriminate the lung tumor and boundary tissues with different classification methods including SVM and Boosting Tree classification models combined with Principal component analysis (PCA) or Random Forest (RF), where RF-Boosting Tree model achieved the higher accuracy of over 98.9% [18]. Therefore, LIBS technique has a potential for rapid detection and discrimination of different biological tissues.

Many researchers presented remarkable researches in the diagnosis and prognosis of different tumor tissues and serum samples by LIBS combined with different machine learning algorithms including PLS-DA, LDA, SVM and k-NN models. k-NN, ANN and SVM classification models were used for the diagnosis and staging of multiple myeloma (MM) with accuracies of over 90%. Area under the curve (AUC) values of these classifiers were 97% with sensitivity of ~93% and specificity of 91% for the discrimination between MM and healthy serum samples. Similarly, AUC values for the different staging of MM were about 97% with 91% sensitivity and 93% specificity respectively. It concluded that diagnosis and staging of serum samples of MM with machine learning methods can achieve the higher accuracy by using LIBS technique [19]. LIBS with chemometric methods used for the discrimination and identification of melanoma and normal FFPE tissues has been performed to achieved their accuracy, sensitivity and specificity all of 100% by using PLS-DA models [20]. Therefore, LIBS with chemometric algorithms is a powerful analytical tool in discrimination and classification of human and animal malignances.

In this work, an effective and rapid method is proposed to identify the Gastric tumor based on elemental components and discriminate it with its healthy tissue by using LIBS combined with chemometric algorithms for the first time. This is possible by comparing LIBS spectra recorded from normal and infected tissues. Classification models including PLS-DA, SVM and k-NN were used for discrimination analysis with two ways of feature line selection, for the efficient detection of GIST. We have found that the PLS-DA gave better accuracy compared to k-NN and SVM algorithms.

2. Materials and methods

2.1. LIBS experimental setup

Gastrointestinal stromal tumor tissue samples were detected using LIBS experimental setup shown in Fig. 1. A flash-pumped Q switched Nd: YAG laser (λ = 1064 nm, pulse frequency 1 Hz, pulse duration τ = 5 ns, beam diameter = Ø6 mm, energy = 40 mJ/pulse) was used for ablation. The laser beam guided through three plane reflectors R1, R2 and R3 and focused on the surface
of sample by a convex lens with 100 mm focal length. A three-dimensional translation stage was used to make a fresh spot of sample for each laser shot. Plasma emission produced on the surface of tissues samples was collected by a convex lens of diameter Ø25 mm and focal length 36 mm into a fiber optic bundle (two fibers, with Ø600 µm aperture each). The optical fiber bundle was connected to a two-channel gated charge-coupled device (CCD) spectrometer (AvaSpec 2048–2-USB2, Avantes). The spectral range of this spectrometer is from 200 nm to 900 nm, with a spectral resolution of 0.20–0.30 nm. External trigger used in the system included a photodetector and a digital delay (SRS- DG535, Stanford Research System). To adjust the time interval between laser and the spectral achievement, the delay time was optimized to 1.29 µs in order to decrease the bremsstrahlung. The integration time of CCD was set to 2 ms. This experiment was performed under ambient air conditions.

![Fig. 1. Schematic of LIBS Experimental Setup for identification of GIST.](image)

### 2.2. Tissue samples

In this work, total twenty samples of GIST and corresponding healthy FFPE tissues of twenty different patients were verified by pathology report, ten GIST and ten healthy controls collected from the Origene Technologies Inc. of America. Each tissue origin was stomach with size of (7cm × 7cm × 6cm), one set thickness was 5µm having size 5 × 5 mm large. Images of GIST and non-GIST tissue obtained from pathology are shown in Fig. 2. There were 10 tissue slides of GIST and 10 tissue slides of healthy class. 90 spectra were collected for each FFPE sample. Total

![Fig. 2. The histopathological images of (a) Healthy tissue, (b) GIST tissue.](image)
900 spectra of each class were taken in which 90 spectra were taken by averaging 10 spectra. Total 180 LIBS spectra were obtained for the discrimination between two classes. After laser ablation FFPE sample of GIST and healthy tissue on glass slides are shown in Fig. 3.

Fig. 3. The FFPE healthy and tumor tissues of GIST after laser ablation.

3. Results and discussion

3.1. LIBS spectra

The emission spectra from the GIST, non-GIST tissue samples and empty glass slide are shown in Fig. 4. In the average spectrum of empty slide, the significant emission lines of Ca, K, Na, N, H, CN and O can be seen. In combined spectra of GIST and its healthy tissues, red line shows average spectrum of tumor tissues and blue line shows average spectrum of healthy tissues. The spectra of GIST and its healthy tissue samples are quite different from that of the empty slide. The NIST atomic emission database was used to label the atomic emission lines in these spectra [21]. Here, the strong spectral lines of Ca, Mg, Fe, Na, Mg, CN, K, N, H and O were observed in the GIST and its healthy tissues. In these spectra, the intensities of several elements in tumor tissue are higher than those in healthy tissue samples.

Fig. 4. LIBS spectra of average intensities of several atomic emission lines of (a) empty slide, (b) GIST and healthy tissues in the spectral range 200–900 nm.

Total 37 emission lines present in LIBS spectra of both GIST and its healthy tissues which included Ca, Mg, Fe, K, Na, H, N, O and CN-band. Biological samples mainly contained Ca,
Mg, Na, K and Fe with different composition and LIBS spectra of the empty slide contained the elements of Ca, Na, K, O, H, N and CN-band. The significant emission lines of Na, Ca, N, H and O were existed in both the spectra of empty slide and malignant samples [22]. Here, Significant emission lines of Ca, Mg, K, Fe and Na with higher intensities were observed in tissue samples. Meanwhile, Ca, K, Na, N, H and O were observed in both spectra of tissue and empty slide. Paraffine wax might also produce difference in intensities of C, H, N and O. Therefore, atomic emission lines emitting from H, N and O were not included in the analysis because of the potential interference produced by the elements in the atmosphere. EL Sherbini et al. concluded in their research that intensities of some specific elements like Ca, K, Mg, Mn, Cu, Na and Fe experienced a dramatic elevation in malignant liver tissues in contrast with healthy once through LIBS [23]. On the basis of strong and significant spectral emission lines existed in GIST tissues, twenty-one atomic lines of Ca, Fe, K, Na and Mg were chosen for the discrimination analysis. These elemental lines are listed in Table 1.

| Elements | Wavelength (nm) |
|----------|-----------------|
| Ca I     | 422.7, 430.2, 443.5, 445.4, 612.2, 616.2, 643.9, 645.7, 714.7 |
| Ca II    | 393.4, 396.8    |
| Mg I     | 516.7, 517.3, 526.06 |
| Fe I     | 308.7, 358.0, 558.8 |
| K I      | 766.5, 769.9    |
| Na I     | 568.3, 588.9    |

Histogram comparing the average normalized intensities of 21 intense LIBS spectral lines of both tumor and healthy samples. The error bars are the standard deviations of the intensities of independent spectra are shown in Fig. 5. Some elements from Fe, Ca, Mg, Na and K in histogram show higher intensities of GIST tissues than its healthy tissues. From these lines, seven specific Ca lines were selected for further improvement in discrimination results as listed in Table 2. These specific Ca lines were selected on the basis of high intensities in tumor tissues as compared to normal tissues for better discrimination results. In this preliminary study, the number of Ca lines were abundant in GIST and its healthy tissues, therefore the classification results by using the Ca lines are different from other elements. This is a similar approach used by Seifalinezhad et al. to distinguish between neoplastic and non-neoplastic gastric tissues by using Spark Discharge LIBS (SD-LIBS). Ca, Mg, K, Na, Fe, Ti, C, O and N elements were identified in the spectrum of gastric tissue. The authors showed that specific elements like Ca and Mg have higher intensities in malignant tissue spectra as compared to its normal ones from the same patient [24].

### 3.2. Discrimination analysis

The principal component analysis (PCA) was performed with the two feature line selections as mentioned above contained 21 feature lines and 7 Ca feature lines. Figure 6(a). 2D plot of scores of PC1 with PC2, PC2 with PC3, PC3 with PC1 and 3D plot of scores of the PC1, PC2 and PC3 with 21 feature line selection and (b) shows the 7 Ca line intensities scatter plots with different combination of PC scores. In both feature line selection, two different classes of tissue are well distinguished. These results show there is no big difference in PCA results of both the feature line selection. The scores of two different classes of tissues are within their respective regions without interfering. In addition, the dispersion of points in each region of the two classes are relatively small and further shows that LIBS combined with PCA can distinguish GIST from its
Fig. 5. Histogram comparing the average normalized intensities of 21 intense LIBS emission lines of both tumor and healthy samples. The error bars are the standard deviations of the intensities of independent spectra.

| Elements | Wavelength (nm) |
|----------|-----------------|
| Ca II    | 393.4           |
| Ca I     | 422.7           |
| Ca I     | 443.5           |
| Ca I     | 445.4           |
| Ca I     | 612.2           |
| Ca I     | 643.9           |
| Ca I     | 714.7           |

healthy tissues. The PCA was performed to prove that elemental components of different GIST are constant.

3.3. Multivariate statistical analysis

3.3.1. SVM and k-NN analysis

For discrimination, two sets of feature lines were used as an input of the PLS-DA, SVM and k-NN classifiers and produced classification accuracies as an output. Firstly, dividing the data into two sets: 72 spectra as test set, another 108 spectra as the training set to build the model. In the process of building the discrimination models by 10-fold cross validation, the data set is divided into ten parts, 9 of which are used for training and one is used for verification, and the average of the results of 10 times is used as an estimate of the accuracy of the algorithm. After this process, the model was built, and we use the test data set to test the model. Training set including 54 spectra of GIST and normal samples (108 spectra in total) and a test set including 36 spectra of gastric and normal samples (72 spectra in total) used for statistical analysis. Among these test data, 36 GIST spectra named as class 1 and 36 healthy tissue spectra named as class 2 were used for discrimination analysis as shown in Fig. 7.
Fig. 6. PCA scores along different combination of PC1, PC2 and PC3 with GIST and healthy tissue samples with (a) 21 feature line and (b) 7 Ca line selection.

Fig. 7. Classification result (a) SVM Model and (b) $k$-NN model based on 21 feature lines.

Support vector machine (SVM) is a supervised machine learning method works on the principle of structure reduction and the vapnik-chervonenkis dimension theory of statistical learning. The working of SVM is to find the best classification hyperplane to fulfil the classification demands which enhanced the blank area on both sides of the hyperplane with confirming the classification accuracy [25]. The kernel function was selected as radial basis function and their parameters
like penalty parameter “c” and kernel function parameter “g” were adjusted by particle swarm optimization (PSO) algorithm.

$k$-NN is used for pattern identification and classification approaches. Euclidean distance measure between particular point “x” and known training data points in the test data, where taking k known training data points by shortest distance from a particular point and calculating the value of k related to which class, calculating the class with the most classification and allocating the “x” point to that class [26]. Euclidean distance was used in selecting nearest samples, where the value of k was 3. This procedure was repeated many times to improve the accuracy, their averaged identification of sensitivity, specificity and accuracy were used to analysis the classification effect. In medical diagnosis, sensitivity is a statistical test which is used to correctly find the malignant samples called as true positive rate, whereas specificity is another statistical test which is used to correctly find the nonmalignant samples called as true negative rate. Meanwhile, Positive predictive value (PPV) is a statistical test’s probability, when returning a positive result, of correctly identifying the malignant samples i.e., identifying true positives rate and at the same instant avoiding false positive rate, where NPV is a statistical test’s probability, when returning a negative result, of correctly identifying nonmalignant samples i.e., identifying true negatives rates and at the same instant avoiding false negative rate.

The final accuracy, sensitivity and specificity of the SVM and $k$-NN model are:

\[
\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)
\]

\[
\text{Sensitivity} = \frac{TP}{TP + FN} \quad (2)
\]

\[
\text{Specificity} = \frac{TN}{TN + FP} \quad (3)
\]

\[
\text{Positive predictive value (PPV)} = \frac{TP}{TP + FP} \quad (4)
\]

\[
\text{Negative predictive value (NPV)} = \frac{TN}{TN + FN} \quad (5)
\]

where TP, FP, TN, and FN represent true positive, false positive, true negative, and false negative, respectively [27,28].

Classification results for SVM and $k$-NN model are shown in Fig. 7. During classification, accuracy decreased by increasing the number of feature lines. In case of 21 feature line selection, SVM accuracy result was 95.83% with sensitivity was 91.67% and specificity was 100% respectively. Its positive predictive value (PPV) was 1 and negative predictive value (NPV) was 0.923 according to Bayes’ formula in eq (4) and (5). On the other hand, $k$-NN classifier achieved the accuracy of about 94.44%, their sensitivity was 88.89% with specificity 100%, respectively. Its PPV was 0.100 and NPV was 0.90. These results specify the difference between the spectra of same kind of tissues because of individual changes between each GIST and its healthy tissues. Classification results of SVM and $k$-NN is not so good because may be the training datasets do not completely occupy the spectral variations from these differences.

Seventeen most important feature lines were of Ca which showed the highest accuracy rate of 100% in both $k$-NN and SVM models. Ca played an important role for classification in contrast to the other elements. Meanwhile, other emission lines of Ca show different classification results in gastric and healthy tissues. Calcium in tumor tissues is abundant with higher intensities than in normal tissues. The results are consistent in the previous studies of emission lines of Ca in gastric tumors [29,30]. Therefore, it is more reliable to choose calcium for discrimination of tumors from normal tissues. Remaining emission lines of Ca, Mg, Fe, K and Na lines from spectral data may be mainly caused by the glass slide as a substrate, paraffine wax and environment so they didn’t show significant effect on the classification and even introduced a decreasing accuracy rate for gastric and healthy tissues.
3.3.2. PLS-DA analysis

PLS-DA is one of the supervised methods which deals with both classification and regression analysis. The PLS-DA model is a multivariate data analysis technique which is very suitable for spectral analysis of biological samples [31,32]. The primary goal of this model is to decrease the number of independent variables to the number of latent variables (LVs) to make the linear groups of dependent variables. For classification, the number of latent variables plays an important role because it acts same as the number of independent variables. If number of LVs are smaller, than it cannot achieve good accuracy but if it is larger than it shows overfitting [33]. In this paper, PLS-DA was used with LIBS technique to identify GIST tissues and its healthy ones. Results of the PLS-DA algorithm are the regression coefficients of the independent variables (LIBS spectra peaks in our analysis). Regression coefficients with high absolute values are usually associated with important peaks for class discrimination and peaks with positive coefficients contribute to the increase of the class calculated response [34]. To make the model, the classification_toolbox_5.0 (Eigenvector Technologies, Inc.) is used in MATLAB. The PLS-DA algorithm showed good classification performance which predicted that all misidentified samples were from diseased tissues and its healthy tissues encountering higher accuracy of about 100%. However, PLS-DA is an effective classification method to discriminate between GIST and its healthy tissues based on atomic and ionic emission lines of LIBS spectra. Total analyzing time were 2 seconds including spectrum calculation and for predictive classification results. All calculations were conducted on a computer with software at Intel Core i7-6600U, CPU 2.60GHz, Windows 10 system, 16 GB RAM. All classification models were used on MATLAB 2016a.

For real-time detection, the number of feature line selection of Ca played an important role in classification but models are unstable for a single element. In order to improve this factor, multiple peaks of this element used to improve the stability of models. Five Ca lines such as 393.4, 396.8, 616.2, 643.9 and 645.7 nm showed 100% accuracy for PLS-DA and SVM, However, 99.01% accuracy were shown by k-NN. Seven selected Ca lines 393.4, 422.7, 443.5, 445.4, 612.2, 643.9 and 714.7 nm showed 100% for all the classification model. In case of all the 11 Ca lines such as 393.4, 396.8, 422.7, 430.2, 443.5, 445.4, 612.2, 616.2, 643.9, 645.7 and 714.7 nm, the accuracy of SVM, PLS-DA and k-NN were 98.6%, 100% and 97.34% respectively. Indeed, for real-time detection, best accuracy result depends on number of feature line selection of Ca. Figure 8 shows the accuracy of SVM, PLS-DA and k-NN classifiers according to different number of feature lines of Ca.

Figure 9 shows the receiver operating characteristic (ROC) curves achieved from 21 lines and 7 selected Ca line intensities by normalization with three discrimination models of GIST. The ROC curve plots sensitivity versus (1-specificity) with different classification models. By using DeLong test, for 21 lines the AUC (Area under the curve) for PLS-DA were 0.100, SVM were 0.9583 and k-NN were 0.9444% respectively. In case of 7 Ca lines AUC for all three discrimination models show 0.100. The working points marked on the upper left corner of ROC curves. Therefore, the tumor tissue slices could be unambiguously discriminated from the healthy ones by using all the three models by leading an ideal ROC property in both cases of 21 feature lines and 7 Ca lines selection. Thus, our results demonstrated that LIBS is a promising analytical method for the detection of GIST and its corresponding healthy tissues.

LIBS technique has a broad application prospect in biomedical, especially in the field of tumor tissues detection [35–37]. In future, it is possible to build in-vitro early diagnosis structure based on LIBS. Classification results of 21 feature lines and 7 selected Ca line intensities of spectral data are listed in Table 3. In our results, PLS-DA model shows the highest accuracy of 100% with latent variable 9 in both cases of 21 lines selection and 7 Ca lines. In case of 21 feature lines, k-NN had accuracy of over 94.44% and the optimal value of k was 3, but for 7 Ca lines k-NN model had accuracy 100% and the optimal value of k was 1. SVM showed 95.83% accuracy and the optimal value of c and g were 0.041 and 0.846. In case of 7 selected Ca lines its SVM...
model had 100% accuracy and the optimal value of c and g were 0 and 1. It was concluded that a smaller number of selected spectral lines of a specific element is important for the best classification results. In multiple verifications, the accuracy difference of the discriminant model is very stable, and variance is less than 4%, so the overall accuracy is used as the comparison standard. Therefore, the accuracy results of PLS-DA, k-NN and SVM models had better results for 7 feature lines selection.

For in-vivo future measurements, the wetness and softness of sample surface is the biggest challenge for LIBS, because it will affect the energy coupling and plasma excitation for biological samples. The inhomogeneity of sample and fluctuating performance of the LIBS system influences the properties of laser-matter interaction. As a preliminary study, this paper demonstrates the
application of LIBS technique for tissue differentiation under in-vitro conditions by examining the atomic emission lines acquired from the samples. It is still necessary to carry out in-depth exploration of soft tissues to improve the better spectral signals of LIBS and its applicability on a large number of human samples. Meanwhile, further studies are required to improve the standards of diagnosis or detection of tumor tissues by using LIBS technique.

4. Conclusion

We introduce the LIBS technique to discriminate between GIST and its healthy tissues with machine learning for the first time. In this work, atomic emission lines of Ca, Mg, Fe, K and Na had been chosen for the discrimination analysis of tumor. Chemometric algorithms, such as PLS-DA, $k$-NN and SVM were used to build the classification models with two feature line selection of elements. From 21 selected lines of Ca, Fe, Mg, K and Na elements, 7 specific Ca lines were selected for further improvement in discrimination results. These specific 7 Ca lines were selected on the bases of high intensities in tumor tissues and also depended on better discrimination results. In the PCA score plots, the dispersion of points in each region of the two classes is relatively small which further shows that LIBS combined with PCA can distinguish GIST and healthy tissues. By comparison of SVM, $k$-NN and PLS-DA classifiers in case of 21 selected lines, their accuracy level increased from 94.44% to 100%. Therefore, both feature selections (21 and 7 emission lines) showed that PLS-DA model have achieved a remarkable discrimination performance for GIST, leading to 100% diagnostic sensitivity as well as specificity and thus representing the potential of LIBS analysis for exploring and screening of soft tissues. SVM and $k$-NN model both gave 100% accuracy for 7 selected Ca lines, and in case of 21 feature lines classification accuracy of SVM was 95.83% and $k$-NN was 94.44%. Therefore, it has been concluded that Ca lines are more important than other lines for identification of GIST. Regarding GIST diagnosis, it is logical to argue that higher peaks of specific Ca lines in a part of a gastric tissues in comparison to the related healthy tissues are associated with the existence of tumor. It can be concluded that, this in-vitro study had high detection sensitivity with highlighting the potential of this tool in clinical applications. On the basis of these results, we will further conduct research for the detection of fresh soft tissues to promote the applications of LIBS in biomedical field.

Funding. National Natural Science Foundation of China (61775017, 62075011).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the corresponding author upon reasonable request.

References

1. A. Rammohan, J. Sathyanesan, K. Rajendran, A. Pitchaimuthu, S. K. Perumal, U. P. Srinivasan, R. Ramasamy, R. Palaniappan, and M. Govindan, “A gist of gastrointestinal stromal tumors: A review,” World Journal of Gastrointestinal Oncology 5(6), 102 (2012).
2. M. Miettinen and J. Lasota, “Gastrointestinal stromal tumors (GISTs): definition, occurrence, pathology, differential diagnosis and molecular genetics,” Pol. J. Pathol. 54(1), 3–24 (2003).
3. N. Patel and B. Benipal, “Incidence of gastrointestinal stromal tumors in the United States from 2001-2015: a United States cancer statistics analysis of 50 states,” Cureus 11(2), e4120 (2019).
4. A. Jemal, E. M. Ward, C. J. Johnson, K. A. Cronin, J. Ma, A. B. Ryerson, and H. K. Weir, “Annual report to the nation on the status of cancer, 1975–2014, featuring survival,” JNCI: Journal of the National Cancer Institute 109(9), djx030 (2017).
5. F. Islami, C. E. DeSantis, and A. Jemal, “Incidence trends of esophageal and gastric cancer subtypes by race, ethnicity, and age in the United States, 1997–2014,” Clin. Gastroenterol. Hepatol. 17(3), 429–439 (2019).
6. J. Stevens, A. A. Botterweck, M. J. Dirx, P. A. van den Brandt, and L. J. Schouten, “Trends in incidence of oesophageal and stomach cancer subtypes in Europe,” Eur. J. Gastroenterol. Hepatol. 22(6), 669–678 (2010).
7. F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, “Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” CA: a cancer journal for clinicians 68(6), 394–424 (2018).
8. Y. Wei, Y. Y. Zhu, and M. L. Wang, “Surface-enhanced Raman spectroscopy of gastric cancer serum with gold nanoparticles/silicon nanowire arrays,” Optik 127(19), 7902–7907 (2016).
9. I. P. Santos, E. M. Barroso, T. C. B. Schut, P. J. Caspers, C. G. van Lantschot, D. H. Choi, M. F. Van Der Kamp, R. W. Smits, R. Van Doorn, R. M. Verdijk, and V. N. Hegt, “Raman spectroscopy for cancer detection and cancer surgery guidance: translation to the clinics,” Analyst 142(17), 3025–3047 (2017).
10. J. G. Fujimoto, M. E. Brezinski, G. J. Tearney, S. A. Boppart, B. Bourma, M. R. Hee, J. F. Southern, and E. A. Swanson, “Optical biopsy and imaging using optical coherence tomography,” Nat. Med. 1(9), 970–972 (1995).
11. L. Guo, Y. Li, F. Huang, J. Dong, F. Li, X. Yang, S. Zhu, and M. Yang, “Identification and analysis of serum samples by surface-enhanced Raman spectroscopy combined with characteristic ratio method and PCA for gastric cancer detection,” J. Innovative Opt. Health Sci. 12(02), 1950003 (2019).
12. Z. Jiang, S. K. Teh, W. Zheng, K. Lin, K. Y. Ho, M. Teh, and K. G. Yeoh, “In vivo detection of epithelial neoplasia in the stomach using image-guided Raman endoscopy,” Biosens. Bioelectron. 26(2), 383–389 (2010).
13. A. Šindelářová, P. Pořízka, P. Modlitbová, L. Vrlíková, K. Kiss, M. Kaška, D. Prochazka, J. Vrábel, M. Buchtová, and J. Kaiser, “Methodology for the Implementation of Internal Standard to Laser-Induced Breakdown Spectroscopy Analysis of Soft Tissues,” Sensors 21(3), 900 (2021).
14. P. Pořízka, P. Prochazková, D. Prochazka, L. Sládková, J. Novotný, M. Petrilak, M. Brada, O. Samek, Z. Přelák, P. Zemánek, and V. Adam, “Algal biomass analysis by laser-based analytical techniques: a review,” Sensors 14(9), 17725–17752 (2014).
15. X. Chen, X. Li, S. Yang, X. Yu, and A. Liu, “Discrimination of lymphoma using laser-induced breakdown spectroscopy conducted on whole blood samples,” Biomed. Opt. Express 9(3), 1057–1068 (2018).
16. A. El-Hussein, A. K. Kassem, H. Ismail, and M. A. Harith, “Exploiting LIBS as a spectrochemical analytical technique in diagnosis of some types of human malignancies,” Talanta 82(2), 495–501 (2010).
17. G. Teng, Q. Wang, H. Zhang, W. Xiangli, H. Yang, X. Qi, X. Cui, B. S. Idrees, K. Wei, and M. N. Khan, “Discrimination of infiltrative glioma boundary based on laser-induced breakdown spectroscopy,” Spectrochimica Acta Part B: Atomic Spectroscopy 165, 105787 (2020).
18. X. Lin, H. Sun, X. Gao, Y. Xu, Z. Wang, and Y. Wang, “Discrimination of lung tumor and boundary tissues based on laser-induced breakdown spectroscopy and machine learning,” Spectrochimica Acta Part B: Atomic Spectroscopy 130, 106200 (2021).
19. X. Chen, Y. Zhang, X. Li, Z. Yang, A. Liu, and X. Yu, “Diagnosis and staging of multiple myeloma using serum-based laser-induced breakdown spectroscopy conducted with machine learning methods,” Biomed. Opt. Express 12(6), 3584–3596 (2021).
20. M. N. Khan, Q. Wang, B. S. Idrees, G. Teng, X. Cui, and K. Wei, “Discrimination of melanoma using laser-induced breakdown spectroscopy conducted on human tissue samples,” Journal of Spectroscopy, 2020.
21. A. Kramida, Y.Ralchenko, and J. Reader, NIST Atomic Spectra Database (ver. 5.0), 2013. National Institute of Standards and Technology, Gaithersburg, MD, USA. (2012), Available online: http://physics.nist.gov/asd (accessed on 3 January 2014).
22. Q. Wang, W. Xiangli, X. Chen, J. Zhang, G. Teng, X. Cui, B. S. Idrees, and K. Wei, “Primary study of identification of parathyroid gland based on laser-induced breakdown spectroscopy,” Biomed. Opt. Express 12(4), 1999–2014 (2021).
23. A. M. E. L. Sherbini, M. M. Hagrah, H. H. Farag, and M. R. M. Rizk, “Diagnosis and classification of liver cancer using libs technique and artificial neural network,” Int J Sci Res 8(4), 1153–1158 (2015).
24. A. Seifalinezhad, M. Bahreini, M. M. H. Matin, and S. H. Tahavassi, “Feasibility study on discrimination of neo-plastic and non-neoplastic gastric tissues using spark discharge assisted laser induced breakdown spectroscopy,” Journal of Lasers in Medical Sciences 10(1), 64–69 (2019).
25. D. Zhang, H. Zhang, Y. Zhao, Y. Chen, C. Ke, T. Xu, and Y. He, “A brief review of new data analysis methods of laser-induced breakdown spectroscopy: machine learning,” Applied Spectroscopy Reviews 1–23 (2020).
26. X. Yan, X. Peng, Y. Qin, Z. Xu, B. Xu, C. Li, N. Zhao, J. Li, Q. Ma, and Q. Zhang, “Classification of plastics using laser-induced breakdown spectroscopy combined with principal component analysis and K nearest neighbor algorithm,” Results in Spectroscopy 4, 100093 (2021).
27. Y. Chu, T. Chen, F. Chen, Y. Tang, S. Tang, H. Jin, L. Guo, Y. Feng Lu, and X. Zeng, “Discrimination of nasopharyngeal carcinoma serum using laser-induced breakdown spectroscopy combined with an extreme learning machine and random forest method,” J. Anal. At. Spectrom. 33(12), 2083–2088 (2018).
28. R. Trevethan, “Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice,” Frontiers in public health 5, 307 (2017).
29. Z. A. Abdel-Salam, A. H. Galmed, E. Tognoni, and M. A. Harith, “Estimation of calcified tissues hardness via calcium and magnesium ionic to atomic line intensity ratio in laser induced breakdown spectra,” Spectrochimica Acta Part B: Atomic Spectroscopy 62(12), 1343–1347 (2007).
30. M. Bahreini, A. Hosseinzadegan, A. Rashidi, S. R. Miri, H. R. Mirzaii, and P. Hadian, “A Raman- based serum constituents’ analysis for gastric cancer diagnosis: In vitro study,” Talanta 204, 826–832 (2019).
31. J. L. Gottfried, F. C. De Lucia, C. A. Munson, and A. W. Miziolek, “Standoff detection of chemical and biological threats using laser-induced breakdown spectroscopy,” Appl. Spectrosc. 62(4), 353–363 (2008).
32. D. M. Haaland, E. V. Thomas, and Edward V. Thomas, “Partial least-squares methods for spectral analyses. 2. Application to simulated and glass spectral data,” Analytical Chemistry 60(11), 1202–1208 (1988).
33. D. M. Hawkins, “The problem of overfitting,” Journal of chemical information and computer sciences 44(1), 1–12 (2004).
34. L. C. Lee, C. Y. Liong, and A. A. Jemain, “Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps,” Analyst 143(15), 3526–3539 (2018).
35. V. K. Singh and A. K. Rai, “Prospects for laser-induced breakdown spectroscopy for biomedical applications: a review,” Lasers Med. Sci. 26(5), 673–687 (2011).
36. X. Y. Liu and W. J. Zhang, “Recent developments in biomedicine fields for laser induced breakdown spectroscopy,” Journal of Biomedical Science and Engineering 01(03), 147–151 (2008).
37. V. K. Singh, V. Kumar, J. Sharma, Y. Khajuria, and K. Kumar, “Importance of laser induced breakdown spectroscopy for biomedical applications: a comprehensive review,” Materials Focus 3(3), 169–182 (2014).