Atomic Spectroscopy

Article

Accuracy Enhancement of Glioma Boundary Tissue Identification by Polarization-resolved LIBS Spectral Fusion

Xiangjun Xu,[a, d, †] Geer Teng,[a, f, †] Qianqian Wang,[a, d, e, *] Haifeng Yang,[b, *] Haiyang Yang,[c] Zhifang Zhao,[a, d] Bingheng Lu,[a, d] Mengyu Bao,[a, d] Yongyue Zheng,[a, d] and Tianzhong Luo [a, d, e]

[a] School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, P. R. China

[b] Department of Neuro-Oncology, Chongqing University Cancer Hospital, Chongqing 400030, P. R. China

[c] Department of Neurosurgery, Kunming Sanbo Brain Hospital, Kunming 650100, P. R. China

[d] Key Laboratory of Photonic Information Technology, Ministry of Industry and Information Technology, Beijing Institute of Technology, Beijing 100081, P. R. China

[e] Yangtze Delta Region Academy of Beijing Institute of Technology, Jiaxing 314033, P. R. China

[f] Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford OX3 7LD, United Kingdom

Received: *April 10, 2024;* ***Revised:*** *May 22, 2024;* ***Accepted:*** *May 22, 2024;* ***Available online:*** *June 18, 2024.*

DOI: 10.46770/AS.2024.101

ABSTRACT: In recent years, laser-induced breakdown spectroscopy (LIBS) combined with machine learning methods has become a hot research topic for detecting malignant tumors. For gliomas with infiltrative features, the tumor boundary is difficult to identify during surgery. To improve the survival time of patients, surgeons often perform extended resection, potentially damaging functional brain areas. Therefore, it is crucial to help surgeons quickly and accurately identify tumor resection boundaries during surgery. In this paper, simulation experiments are conducted using isolated tissues, proposing a polarization-resolved LIBS (PRLIBS) spectral fusion method to boost the accuracy of glioma boundary tissue detection. First, the polarization effect of the plasma emission is analyzed using the Stokes parameters, and it is found that the plasma emission belonged to partially polarized light. To better exploit the polarization information of the plasma, the polarized spectra from the four channels are fused to build a machine learning model. Comparing to classification models using LIBS intensity spectra, polarization parameters, and single-channel polarization spectra, the PRLIBS fusion model exhibits superior classification performance. The correct classification rate (CCR) of support vector machine (SVM) model is 99.05% for the training set and 89% for the test set, respectively. In the future, the PRLIBS spectra fusion method proposed in this research can be used for glioma boundary tissue identification.

INTRODUCTION

According to statistics released to the public by the International Agency for Research on Cancer (IARC), there were 19.3 million newly diagnosed cancer cases and 10 million cancer deaths worldwide in 2020.[1,2] The high mortality rate of cancer poses a great threat to human health, and precision treatment of cancer is important for improving patient outcomes. Glioma is the most common type of brain tumor, which is characterized by aggressive growth.[3,4] Surgical resection currently serves as the primary treatment modality for gliomas. However, during surgery, surgeons relying on the naked eye cannot accurately determine the boundaries of the tumor. Insufficient excision can lead to recurrence of residual cancer cells. Widening the extent of resection during surgery may enhance patient survival, it poses a risk of inadvertently damaging functional brain regions. Therefore, it is important to assist the surgeon in accurately identifying the boundaries of tumor resection during surgery. At present, two methods are used for determining glioma boundaries: rapid

pathology testing of isolated tissue and in-vivo assisted detection techniques, including rapid magnetic resonance imaging, ultrasound and fluorescent labelling.[5-8] The above methods have disadvantages, such as being time-consuming, costly, imprecise, and requiring the injection of fluorescent contrast agents. Hence, it is imperative to investigate new techniques that provide cost-effective detection and higher accuracy.

Laser-induced breakdown spectroscopy (LIBS) is a technique in the field of atomic emission spectroscopy that enables the analysis of a substance by utilizing the interaction between the laser and sample.[9-11] The technique has several advantages, including rapid detection speed, minimal damage, simplified sample preparation requirements, and the capability to simultaneously test numerous elements. Therefore, it has great potential for application in screening of malignant tumors and identification of cancerous tissues. In recent years, the combination of LIBS with machine learning methods for detecting malignant tumors has become a hot research topic.[12-15] Sherbini *et al.*[16] utilized the backpropagation Artificial Neural Network (ANN) to effectively distinguish between 26 malignant liver cancer tissues and 4 normal liver samples, achieving an accuracy rate over 80%. Li *et al.*[17] used the LIBS technique for the diagnostic analysis of pathological sections of cervical cancer tissues and normal cervical tissues. It was found that the normalized peak intensities of Na, Mg, and K in cervical cancer tissues were significantly greater than those in normal tissues, whereas the normalized peak intensity of Ca was lower than that in normal tissues. Han *et al.*[18] used LIBS to distinguish mouse melanoma tissues and dermal tissues, demonstrating that the emission intensity of Mg and Ca spectral lines was notably higher in melanoma than in normal skin tissues. Utilizing the first 15 scores from Principal component analysis (PCA) as inputs to the Linear discriminate analysis (LDA) model, the sensitivity and specificity of the classification model were 96.7% and 99.7%, respectively. The above research results indicate that LIBS technique has great potential for application malignant tumor detection.

Traditional LIBS only collects the intensity information from plasma emission, but since light is a transverse wave, its polarization information is critically important. Consequently, recent years have seen the emergence of new research areas such as the investigation of the polarization effects of plasma emission and the use of polarization-resolved LIBS (PRLIBS) spectroscopy for quantitative and qualitative analysis. Liu *et al.*[19] used a femtosecond double-pulse laser to ablate Si materials. The study observed that continuous emission exhibited a strong polarization effect, displaying a degree of polarization (DOP) above 0.95 in the 265–330 nm, while having a relatively low DOP of 0.2–0.4 in the 350–500 nm for atomic spectral lines. Wubetu[20] investigated the DOP of aluminum's atomic lines and continuous emission using time-resolved LIBS methods. The DOP of the continuous emission was 3-4 times greater than atomic lines during the delay time was 200–700 ns, with the main reason believed to be anisotropic electron velocity distribution. Zhao *et al.*[21] combined PRLIBS and partial least squares (PLS) regression modeling to improve the stability and accuracy of quantitative analysis of elements in alloys. Compared with LIBS measurements, the relative standard deviation (RSD) and root mean square error (RMSE) of PRLIBS were reduced, while the signal-to-background ratio (SBR) and coefficient of determination (R^2) were enhanced. For the PRLIBS dataset, R^2 increased by 3.4%, and root mean square error of calibration (RMSEC) and root mean square error of prediction (RMSEP) decreased by 6.3% and 19.0%, respectively. Presently, analyzing plasma polarization effects necessitates multiple spectral measurements, achieved by rotating the polarizer at different angles to collect spectra in different polarization directions. Consequently, stability and accuracy of analytical results are severely affected by plasma fluctuations.

Since 2020, the utilization of LIBS in glioma boundary detection has been proposed, which combines chemometrics and machine learning methods to achieve better identification results. Teng *et al.*[22] employed LIBS on freshly obtained glioma and infiltrative tissue samples. The classification model, which utilized the Random Forest (RF) feature selection method combined with Support Vector Machine (SVM), achieved an accuracy of 95%. In Ref. 22, although the CCR reached 95%, the sample size was limited, consisting of a single glioma and infiltrative boundary tissue from the same patient. The 200 spectra were collected from the same tissue were divided into a training set and a test set. Thus, it was not possible to demonstrate the robustness of the model when making predictions for multiple unknown patients. Teng *et al.*[23] proposed the method of full Stokes laser-induced breakdown spectroscopy (FSLIBS) and the corresponding experimental system, which required four measurements to demodulate the polarization effect. Then, the elemental composition information and polarization information were combined for glioma boundary detection. A total of 42 fresh tissue samples were analyzed in the experiment, and the results showed that Ca, Na, and K elemental spectral lines and CN, C_2 molecular bands played an important role in the identification of different types of tissues. Meanwhile, compared with traditional LIBS, FSLIBS provides complete polarization information, which reduces the underfitting phenomenon of machine learning models and improves the robustness of the models. In Ref. 23, the demodulation process of the four measurements was highly relied on the stability of the plasma emission, it was difficult to reduce the effects of unevenness in sample composition and fluctuations in pulsed laser energy. Therefore, it is necessary to enhance the experimental setup to enable the demodulation of the polarization effect of plasma emission through a single measurement. In addition, in Ref. 23, although the Stokes parameters show the potential ability of increasing spectral differences, how to effectively exploit the polarization information of the plasma needs to be further investigated.

In this paper, a novel polarization spectral acquisition system is constructed, which can analyze the polarization effect of laser-induced plasma emission with a single measurement. On this basis, to leverage the polarization information of the plasma, a PRLIBS spectral fusion method is proposed to enhance the detection accuracy for glioma boundary tissues. Meanwhile, the method is compared with classification models built using LIBS intensity spectra, polarization parameters, and single-channel polarization spectra.

EXPERIMENTAL

Experiment setup. The experimental setup is shown in Fig. 1(a), which adds a polarization spectrum collection system alongside the acquisition of LIBS intensity spectra compared to the conventional LIBS system. A homemade Q-switched (Nd: YAG) laser with a pulse duration of 10 ns and a maximum repetition frequency of 20 Hz is used as the excitation source. The laser pulse is first reflected by the dielectric film mirrors M_1, M_2 and M_3, and then passes through the dichroic mirror M_4 (GCC-414009, Daheng Optoelectronics), which transmits infrared light at1064nm and reflects the visible light range of 400-872nm. The NIR infinity-corrected objective (working distance 30.5 mm, Plan Apo NIR, Mitutoyu) focuses the laser beam on the sample surface. Here, the diameter of the ablation crater is about ∅100 μm. To observe the focusing point and sample surface morphology, a CMOS camera (INFINITY1-2CB, Teledyne Lumenera, Canada) and imaging lens are mounted for microimaging. In the traditional LIBS collection system on the left side of Fig. 1(a), two silica lenses (L_1, L_2) with a focal length of 75 mm couple the spectral signals into the optical fiber. The spectral data are measured using a dual-channel spectrometer (AvaSpec 2048-2-USB2, Avantes, measurement range 200-1100 nm, resolution 0.2-0.3 nm), and thus, the dual-channel spectrometer collects LIBS intensity spectra.

Figure 1(b) shows an enlarged view of the polarized spectral collection system. The plasma emission is first converted into parallel light through a silica lens L_3 with a focal length of 60 mm, and then divided equally into four beams by three non-polarized beam-splitting prisms (NPBS) with a 50:50 beam-splitting ratio. Here, the use of NPBS as a beam splitter guarantees that the plasma emission from the polarization collection optical paths (1, 2, 3 and 4) maintains consistent intensity and polarization characteristics. For optical path 1, an achromatic 1/4 waveplate (QWP, PO-TWP-L4-25-UVIR, Alphalas, Germany) and a linear thin-film polarizer (LP_1) are mounted in the transmitted optical path of the $NPBS_2$ to collect the circularly polarized component of the plasma emission. For optical path 2, LP_2 (with the transmission axis at an angle of 45° to the horizontal) is installed in the reflected optical path of $NPBS_2$ to collect the 45° linearly polarized component of the plasma emission. For optical path 3, LP_3 (with the transmission axis at an angle of 0° to the horizontal) is installed in the reflected optical path of $NPBS_3$ to collect the 0° linearly polarized component of plasma emission. For optical path 4, LP_4 (with the transmission axis at an angle of 90° to the horizontal) is installed in the transmitted optical path of $NPBS_3$ to collect the 90° linearly polarized component of the plasma emission. L_4, L_5, L_6, and L_7 are silica lenses with a focal length of 50 mm for coupling light into the optical fiber. Four fiber-optic spectrometers of the same model (Ocean Optics, USB2000+, measurement range 350-1000 nm, resolution 0.5 nm) are used to simultaneously detect four sets of polarized spectra. Prior to the measurement, a standard tungsten halogen lamp is utilized for absolute radiation intensity calibration of the four fiber optic spectrometers, ensuring the elimination of any impact from differences in detector response and quantum efficiency on spectral intensity. In addition, each spectrometer requires wavelength calibration to maintain measurement accuracy. Detailed procedures for spectrometer wavelength calibration and absolute radiation calibration are described in in the Supporting Information (Fig. S1).

Fig. 1 Diagram of the experimental setup. (a) Diagram of the overall LIBS system, (b) Polarization spectrum detection system.

In the experiments, the flash lamp and Q-switching of the laser are precisely triggered by a digital delay generator (DG535, Stanford, USA), with a delay of 372 μs. Other experimental parameters are as follows: the laser energy is 100 mJ, the laser repetition frequency is 1 Hz, and the spectrometer integration time and delay time are 1.05 ms and 1 μs, respectively. The dual-channel spectrometer and four spectrometers are synchronized for external triggering using a digital delay generator. Consequently,

Fig. 2 Procedure for simple preparation of tissue samples.

the LIBS intensity spectra and the corresponding four-channel polarization spectra are acquired for each laser pulse.

Glioma samples detection. Tissue samples, including gliomas, boundary tissues, and healthy tissues, were obtained from the hospital's Neuro-Oncology Department following routine tumor surgery. This research has been ethically approved by the Hospital Ethics Committee (Kunming Sanbo Brain Hospital, China, April 2020). As high-grade gliomas are involved, surgery aims to ensure no recurrence, which may result in the collection of a small amount of normal tissue. A total of 56 tissue samples of three types (normal, boundary, and tumor) from 24 patients were provided by the two hospitals. Each tissue sample was histopathological diagnosed to determine its tissue type. All excised tissue samples were cryopreserved at ultra-low temperatures of -80°C.

When the pulsed laser directly ablates fresh tissue samples, the intensity of the excited plasma is low. This is primarily due to the softer surface of isolated fresh tissue, and also due to the residual blood and tissue fluid on the surface of the tissue. Therefore, fresh tissue samples require simple sample preparation. Firstly, the ultra-low temperature frozen tissue samples in Fig. 2(a) are thawed to obtain the samples in Fig. 2(b). Subsequently, take 0.04 g of tissue sample and place it in a sample analysis ring which is made of ABS (∅5 mm in diameter and 2 mm in height). In the final step, position the hairdryer 25 cm away from the tissue sample and blown over the surface for 5 s, as shown in Fig. 2(c). After the blowdown, the surface of the sample is visibly flatter and hardened, with no liquid residue. The inside of the tissue block remained moist, unaffected by the surface treatment. Using the experimental system constructed in the "Experiment setup" section, spectral data from 56 samples are collected. Fifty spectra are acquired for each tissue sample, and the three-dimensional stage is adjusted to ensure that each laser pulse ablates a fresh area of the sample.

Herein, the 56 tissues are divided into a training set and a test set according to the samples, ensuring that the spectra of each tissue existed only in the training or test set. From the total samples, 25% are randomly selected for the test set, and the remaining 75% for the training set. Specifically, the test set consisted of 14 samples, including 4 normal tissue samples, 5 boundary tissue samples, and 5 tumor tissue samples. The training set consists of 42 samples, including 10 normal tissue samples, 16 boundary tissue samples and 16 tumor tissue samples. The detailed description of dividing the training and test sets is presented in the Supporting Information (Table. S1). Meanwhile, Wilcoxon test is performed on the training set spectra and test set spectra. This is to ensure that there is no statistically significant difference in the distribution of the training set and test set spectral data. The performance of model classification is described by the correct recognition rate (CCR), which is calculated as follows.

$$CCR = \frac{1}{N}\sum_{i=1}^{q} \delta_i \times 100\% \quad (1)$$

where δ_i is the number of spectra classified correctly for each type of sample and q is the number of classes. N is the number of all spectra. The classes in this research are three, which are normal tissue, boundary tissue and tumor tissue. The number of spectra in the training set 50×42=2100 and the number of spectra in the test set 14×50=700.

RESULTS AND DISCUSSION

LIBS spectra. The average spectra of normal tissue (No.1), boundary tissue (No.15) and glioma (No.36) are shown in Fig. 3. In the spectral region of 200-950 nm, there are several characteristic spectral lines and molecular bands, specifically including Mg, Ca, Na, H, N, K, O, C, CN and C_2. Mg, a trace element found in the brain, has been proven to be present in

Fig. 3 LIBS spectra of different tissues. (a) tumor tissue, (b) boundary tissue and (c) normal tissue.

Fig. 4 Typical PRLIBS spectra of tumor tissue from No.36. (a) PRLIBS (0°, 0), (b) PRLIBS (45°, 0), (c) PRLIBS (45°, π/2) and (d) PRLIBS (90°, 0).

Fig. 5 The intensities of the polarization spectra of the four channels are compared.

gliomas.[24] Meanwhile, several characteristic spectral lines of Ca were observed at 393.4 nm, 396.8 nm, 421.3 nm, 422.7 nm and 430.8 nm, respectively. Na and K are two common elements in biomedical samples, as they control the balance of physiological regulation. Spectral differences between different tissue types are minimal. Therefore, it is unfeasible to precisely distinguish different tissue types solely based on the intensity of a limited number of characteristic spectral lines.

PRLIBS spectra and polarization effect of plasma emission. The four-channel PRLIBS spectra are shown in Fig. 4, where PRLIBS (0°, 0), PRLIBS (45°, 0), and PRLIBS (90°, 0) represent the 0°, 45°, and 90° linearly polarized light of the plasma emission, respectively, while PRLIBS (45°, π/2) represents the circularly polarized light of the emission. Due to the NPBS's wavelength transmission limitation to the visible wavelength band, the PRLIBS spectrum contains only seven characteristic spectral lines, including CN 387.7 nm, Ca 393.4, Ca 396.8 nm, Ca 422.7 nm, C_2 517.9 nm, Na 589.0 nm and H 656.3 nm. Meanwhile, from Fig. 4, it can be seen that there are disparities in the spectral intensities of the same characteristic spectral line measured in the four channels.

It is well known that the Stokes parameters (S_0, S_1, S_2, S_3) describe not only completely polarized (CP) light, but also unpolarized (UNP) light and partially polarized (PP) light as well[25]. The four parameters satisfy the formula $S_0^2 \geq S_1^2+S_2^2+S_3^2$, where equality applies to CP light and the inequality applies to both UNP and PP light. The Stokes parameters of the spectra can be calculated using equations (2), (3), (4) and (5).

$$S_0 = I\left(0^{\circ},0\right) + I(90^{\circ},0) \tag{2}$$

$$S_1 = I\left(0^{\circ},0\right) - I(90^{\circ},0) \tag{3}$$

$$S_2 = 2 \times I\left(45^{\circ},0\right) - S_0 \tag{4}$$

$$S_3 = S_0 - \times 2 \times I\left(45^{\circ},\frac{\pi}{2}\right) \tag{5}$$

Where I (0°, 0), I (45°, 0), I (90°, 0) and I (45°, π/2) are the intensities of the characteristic spectral lines PRLIBS (0°, 0), PRLIBS (45°, 0), PRLIBS(90°, 0) and PRLIBS(45°, π/2), respectively. Then, the intensity (I_{pol}) of the completely polarized component as well as the degree of polarization (DOP) can be calculated using formulas (6) and (7).

$$I_{pol} = \sqrt{S_1^2 + S_2^2 + S_3^2} \tag{6}$$

$$DOP = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0} \tag{7}$$

Below is an analysis of the plasma's polarization effect, focusing on the intensity and DOP of the characteristic spectral lines. Fig. 5 shows the integral intensities of the characteristic spectral lines of the four-channel polarization spectra. There are differences in the intensities of the characteristic spectral lines between the different

Fig. 6 Prediction results of different models. (a) training set and (b) test set.

Fig. 7 PRLIBS fusion spectra are obtained by merging the characteristic peaks of four-channel polarization spectra.

channels, indicating the presence of polarization effects in the plasma emission. This is because, if the emission is unpolarized light, there is no difference in the intensities of the characteristic spectral lines when collecting spectra at different polarization angles. Meanwhile, specific values of Stokes parameters, I_{pol} and DOP for the characteristic spectral lines are calculated using equations (2)-(7). (Table S2). It is found that for all the characteristic spectral lines, the DOP of the radiated light is between 0 and 1, indicating that the plasma emission is partially polarized light.

Furthermore, the DOPs of the characteristic spectral lines are compared for three different types of tissue samples. For 14 normal tissues, the average DOP values of the three characteristic spectral lines Ca 393.4 nm, Ca 422.7 nm and Na 589.0 nm are 0.349, 0.396 and 0.227, respectively. For the 21 boundary tissues, the average DOP values are 0.348, 0.389 and 0.258, respectively. For the 21 tumor tissues, the average DOP values are 0.387, 0.419 and 0.235, respectively. The DOPs of the characteristic spectral lines for each tissue sample are detailed in Tables S3, S4, and S5. The main characteristic spectral lines exhibit significant polarization effects. However, the differences in average DOPs across tissue types are minimal. Therefore, in the research, methods such as feature selection and polarization spectral fusion are combined with machine learning for model training. This enables better extraction of polarization information for different tissues and achieves better spectral recognition.

Classification results of gliomas based on LIBS spectra. From the "LIBS spectra" section, it can be seen that the elemental compositions of the three different tissues (normal, boundary and tumor) are basically the same, and it is mainly the intensity of the characteristic spectral lines that differs between them. With the rapid development of artificial intelligence in recent years, machine learning methods have been applied to the biomedical field. These models can effectively distinguish subtle differences between spectra of different tissue, significantly enhancing classification accuracy from complex spectral data.

First of all, all the characteristic spectral lines from LIBS spectra are used to construct classification models. Raw spectra undergo preprocessing, which involves background removal, peak finding and total intensity normalization. Variables and labels are input into SVM, PLS-DA, RF and KNN for training model construction, respectively. Specific spectral preprocessing procedures and machine learning model descriptions can be reviewed in our previously published papers.[26-30] Hyperparameters are optimized using 10-fold cross validation (Detailed model parameters are listed in Table S6.). The CCRs of the training and test sets are shown in Fig. 6. The classification models built using all characteristic spectral lines, SVM and RF models perform better on the training set, while PLS-DA has a low CCR of 60.1%. On the test set, all four models' performance is poor, with RF at 53% and SVM, PLS-DA, and KNN having CCRs below 50%. Consequently, there is a requirement to use other methods to improve the classification performance of the models.

In LIBS spectra, there are dozens of representative feature peaks, including numerous atomic spectral lines and molecular bands. However, not all spectral information contained in the feature peaks is helpful for identification and classification. There is also redundant or even interfering information. Therefore, when building a classification model, choosing appropriate feature spectral lines can improve the model's classification performance. Based on spectral properties, seven characteristic spectral lines in the LIBS spectra are selected: Ca 393.4 nm, Ca 396.8 nm, Ca 422.7 nm, Na 589.0 nm, H 656.3 nm, CN 387.7 nm and C_2 517.9 nm. The selection rationale is as follows. Ca plays a crucial role in the human body, including the synthesis and release of neurotransmitters, as well as the production and secretion of hormones. In addition, the greater intensity of the characteristic spectral lines of Ca in tumors may be due to calcification of the tumor tissue. Na is an electrolyte vital for maintaining water balance and acid-base equilibrium. CN and C_2 bands have been used in previous studies to classify biomedical samples, proving that molecular bands as reliable tissue markers. The prediction results of the feature selection model, shown in Fig. 6, exhibit improved classification performance compared to using all feature peaks. The CCR of the SVM model increases from 45% to 53.14%, the PLS-DA model from 47.29% to 54.43%, and the KNN model from 46.57% to 49.86%. The enhancement in prediction accuracy indicates that Ca, Na, H, CN and C_2 play a significant role in identifying different types of tissues.

Classification results of gliomas based on PRLIBS spectra.

The classification model of conventional LIBS utilizes the intensity information of plasma emission, but the results do not achieve high accuracy. While the polarization parameter builds a classification model that utilizes both intensity and polarization information of the emission. Due to the large relative standard deviation (RSD) of the polarization parameters, the model's classification performance is subpar. Results of glioma classification using polarization parameters are detailed in the Supporting Information. Based on the above analysis, the fusion of the 4-channel polarized spectra can ensure low RSD of characteristic spectral lines and fully utilize the polarization characteristics of the plasma. The polarization spectra are not processed using complicated spectral preprocessing so that the polarization information of the plasma could be retained intact. Specifically, pixel intensities are recorded for each spectral peak, and then the four channels are combined. As shown in Fig. 7, the fusion spectra are obtained by merging the characteristic peaks of PRLIBS (0°, 0), PRLIBS (45°, 0), PRLIBS (45°, π/2) and PRLIBS (90°, 0). Here, each set of polarized spectra contains seven characteristic peaks of Ca 393.4 nm, Ca 396.8 nm, Ca 422.7 nm, Na 589.0 nm, H 656.3 nm, CN 387.7 nm and C_2 517.9 nm. Consequently, the number of polarized spectral variables after fusion is four times greater than each channel's individual count. Polarization spectral fusion is performed using the same method for both the entire training and test sets. Following that, total intensity normalization is applied to the polarization fusion spectra, and classification models are trained separately.

The prediction results are shown in Fig. 6, and the PRLIBS fusion model exhibits the best classification performance. The PRLIBS fusion model's CCR is significantly higher than the CCRs of the other three methods, particularly when compared to the SVM and RF models, which had test set CCRs of 89% and 63.43%, respectively. Upon further comparison, the SVM model showed good classification performance, while the RF model showed clear overfitting, and the PLS-DA and KNN models displayed underfitting. This may be due to the following factors: (i) The SVM model has better robustness to noise and anomalous data present in the spectra, which effectively avoids overfitting. SVM also handles high-dimensional data and small sample problems well, making it suitable for LIBS qualitative classification tasks. (ii) RF is an ensemble learning method composed of multiple decision trees, each trained on random subsets of samples and features. The high spectral similarity and insignificant differences between different types of tissues, leading to significant overfitting of the RF model. (iii) The PLS-DA and KNN models are sensitive to noise and outliers in the input data, while these models require a large amount of spectral data for training to obtain better predictions. If the number of spectra is small or the feature dimension is high, the underfitting problem may occur. Last but not least, as a comparison, the polarization spectra of the four channels in Fig. 7 are individually input into SVM, PLS-DA, RF and KNN to build the classification model.

Fig. 8 Confusion matrix for classification results of SVM polarization fusion models.

The predictions for each model are detailed in Fig. S3. Compared to the prediction accuracy of the PRLIBS fusion model in Fig. 6, the single-channel polarization spectral model has poor classification performance. This is mainly due to the fact that the single-channel polarization spectral model does not effectively exploit the polarization effect of the plasma. The dimensionality of the independent variable reduction in these models, to 1/4 of the PRLIBS fusion model, limits their ability to effectively learn the spectral differences between different tissue types.

To evaluate the SVM polarization fusion model's performance in distinguishing among three tissue categories, the confusion matrix is shown in Fig. 8. The CCRs for normal, boundary and tumor tissues in the test set are 73.5%, 97.2% and 93.2%, respectively. The relatively lower accuracy for normal tissues can be attributed to two factors. On the one hand, the number of normal tissues samples is small when compared to the other two tissue types. As a result, the model has limited ability to learn normal tissue spectra. Specifically, there are 16 samples each of boundary and tumor tissues, contrasted with only 10 samples of normal tissue. On the other hand, the tissue samples in the experiments came from clinical surgeries, with a difference in the regions of excised tissues that underwent PRLIBS testing and pathological biopsies.

The variable importance weights of PRLIBS fusion spectra and the corresponding feature selection models are analysed. Here, the importance weights are evaluated by the Relief-F algorithm.[28,29] Fig. 9(a) and (b) show the feature importance weights and the predictions

Fig. 9 Relief-F calculates variable importance weights and corresponding feature selection model prediction accuracies for fusion spectra. (a) Importance weights of variables for fusion spectra and (b) Prediction accuracy of feature selection models.

of the feature selection model, respectively. When compared to the other polarization spectra, PRLIBS (45°, π/2) has the highest significance weight for the molecular spectral band CN 387.7 nm. For the atomic spectral lines Ca 393.4 nm, Ca 396.8 nm and Ca 422.7 nm, the importance weights are higher in each set of polarized spectra, indicating that Ca play an important role in identifying different tissue types. Additionally, the estimated importance weights for the feature spectral lines of all seven features are nonzero, indicating that each spectral line contributes to the classification model. Following this, the spectral variables are ranked from highest to lowest based on their importance weights, and 620 feature selection models are trained sequentially. Specifically, the first model inputs the spectral feature with the largest weight, followed by the top two features with the highest weights, and so on. In the above order, 620 classification models are built separately. As shown in Fig. 9(b), when applying the feature selection model to the test set, the CCR increases as more features are input. The CCR reaches its maximum of 89% when all 620 spectral variables are used to construct the classification model. The PRLIBS spectral fusion method improves the classification accuracy of glioma boundary tissues, primarily due to three key factors: (i) Emission from laser-induced plasma has a polarization effect, which is partially polarized light. Therefore, four-channel polarization spectral fusion can obtain more spectral information. (ii) By fusing the four-channel polarized spectra, more characteristic spectral lines with high importance weights can be obtained. (iii) The prediction results of the feature selection model indicate that the highest classification accuracy is achieved by constructing an SVM model with all the variables of PRLIBS fusion spectra. Consequently, the polarization spectral fusion model efficiently harnesses the plasma's polarization information.

In the end, the advantages of the PRLIBS spectral fusion method and the traditional glioma detection methods are compared. For rapid pathological testing of isolated tissues, although the recognition accuracy exceeds 95%, the testing time exceeds 30 minutes. This method can only analyze a few locations in the lesion area, and it is unable to achieve wide-scale discrimination of tumor resection boundaries. For rapid magnetic resonance imaging and ultrasonography during surgery have detection times of 45 min and 10 min, respectively, and primarily provide a rough estimation of the lesion area without precise tumor boundary identification. For the fluorescent labeling method, the visualization of the tumor margins is enhanced during surgery by injecting a fluorescent agent. It can assist the surgeon in better determining the extent of glioma resection, thereby enhancing the rate of total tumor resection. However, the fluorescent effect decreases significantly with the duration of the procedure. In this paper, simulation experiments have been conducted using isolated tissues, and the proposed PRLIBS spectral fusion method has a

detection time of a few seconds, with a recognition accuracy of up to 89%. Under laboratory conditions, the speed and accuracy of detection obtained better results. Therefore, the method exhibits promising application prospects in glioma tissue detection.

CONCLUSION

In this work, a novel polarization spectral acquisition system is built, which collects the polarization information of the plasma by a single measurement. A simulation experiment for glioma boundary identification is conducted using 56 isolated tissues. By analyzing the spectra of normal, boundary, and tumor tissues, it is observed that there are subtle spectral variations among them. The analysis of laser-induced plasma's polarization using Stokes parameters revealed that the emitted light is partially polarized. To make full use of the polarization information of the plasma, we fused four-channel polarized spectra for machine learning model training. Comparing the classification models built using LIBS intensity spectra, polarization parameters and single-channel polarization spectra, the PRLIBS fusion model has the best classification performance. Notably, the SVM polarization fusion model has excellent CCR with 99.05% and 89% for the training and test sets, respectively. The polarization spectral acquisition and fusion method proposed in this research has promising applications in boundary identification of glioma tissues.

ASSOCIATED CONTENT

Supporting information (Figs. S1−S3 and Tables S1−S6) is available at www.at-spectrosc.com/as/home

AUTHOR INFORMATION

Qianqian Wang received her Ph.D. degree from Beijing Institute of Technology in 2002. She is a professor and doctoral supervisor at Beijing Institute of Technology. She is mainly engaged in the research of optoelectronics and laser technology, and the research covers laser-induced breakdown spectroscopy, machine learning, and optoelectronic imaging and detection. She has presided over more than 40 projects, including the National Natural Science Foundation of China (NSFC), the Nature Foundation of Beijing and the development of optoelectronic equipment. And she has published more than 100 academic papers as the first or corresponding author.

Haifeng Yang received his Ph.D. degree from Capital Medical University in 2020. He is the director of the Department of Neuro-Oncology at Chongqing University Cancer Hospital. He mainly researches on the precise surgical treatment of brain tumors and the mechanism of brain tumor genesis. He has presided various research projects, including the National Natural Science Foundation of China (NSFC), the NSFC Youth Program, and the sub-projects of the National Basic Research Program of China (973 Program).

Corresponding Author

* Q. Q. Wang

Email address: qqwang@bit.edu.cn

* H. F. Yang

Email address: yanghaifeng@cqu.edu.cn

†These authors contributed to the work equally and should be considered co-first authors.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors gratefully thank the National Natural Science Foundation of China (62075011) for financial support of this work.

REFERENCES

1. A. N. Giaquinto, H. Sung, K. D. Miller, J. L. Kramer, L. A. Newman, A. Minihan, A. Jemal, and R. L. Siegel, *CA Cancer J. Clin.*, 2022, **72**, 524−541. https://doi.org/10.3322/caac.21754
2. D. Sun, H. Li, M. Cao, S. He, L. Lei, J. Peng, and W. Chen, *Cancer Biol. Med.*, 2020, **17**, 879−895. https://doi.org/10.20892/j.issn.2095-3941.2020.0387
3. Q. T. Ostrom, N. Patil, G. Cioffi, K. Waite, C. Kruchko, and J. S. Barnholtz-Sloan, *Neuro-Oncology*, 2022, **24**, 1214. https://doi.org/10.1093/neuonc/noaa269
4. S. Pisipati, K. A. Smith, K. Shah, K. Ebersole, R. B. Chamoun, and P. J. Camarata, *Neurosurg. Focus*, 2016, **41**, E13. https://doi.org/10.3171/2016.7.Focus16224
5. L. Cheng, W. He, H. Zhang, Q. Song, B. Ning, H. Li, Y. He, and S. Lin, *Biomed Res. Int.*, 2017, **2017**, 5782619. https://doi.org/10.1155/2017/5782619
6. Y. Liu, Z. Wang, X. Li, X. Ma, S. Wang, F. Kang, W. Yang, W. Ma, and J. Wang, *Mol. Pharm.*, 2019, **16**, 108−117. https://doi.org/10.1021/acs.molpharmaceut.8b00888
7. O. I. Olubiyi, A. Ozdemir, F. Incekara, Y. M. Tie, P. Dolati, L. G. Hsu, S. Santagata, Z. R. Chen, L. Rigolo, and A. J. Golby,

World Neurosurgery, 2015, **84**, 528−536. https://doi.org/10.1016/j.wneu.2015.04.044

8. P. P. Panciani, M. Fontanella, B. Schatlo, D. Garbossa, A. Agnoletti, A. Ducati, and M. Lanotte, *Clin. Neurol. Neurosur.*, 2012, **114**, 37−41. https://doi.org/10.1016/j.clineuro.2011.09.001
9. R. Kanawade, F. Mahari, F. Klampfl, M. Rohde, C. Knipfer, K. Tangermann-Gerk, W. Adler, M. Schmidt, and F. Stelzle, *J. Biophotonics*, 2015, **8**, 153−161. https://doi.org/10.1002/jbio.201300159
10. X. Xu, F. Ma, J. Zhou, and C. Du, *Comput. Electron. Agr.*, 2022, **199**. 107171. https://doi.org/10.1016/j.compag.2022.107171
11. S. A. Davari, S. Masjedi, Z. Ferdous, and D. Mukherjee, *J. Biophotonics*, 2018, **11**, e201600288. https://doi.org/10.1002/jbio.201600288
12. V. Gardette, V. Motto-Ros, C. Alvarez-Llamas, L. Sancey, L. Duponchel, and B. Busser, *Anal. Chem.*, 2023, **95**, 49−69. https://doi.org/10.1021/acs.analchem.2c04910
13. B. S. Idrees, G. Teng, A. Israr, H. Zaib, Y. Jamil, M. Bilal, S. Bashir, M. N. Khan, and Q. Wang, *Biomed. Opt. Express*, 2023, **14**, 2492−2509. https://doi.org/10.1364/boe.489513
14. B. S. Idrees, Q. Wang, M. N. Khan, G. Teng, X. Cui, W. Xiangli, and K. Wei, *Biomed. Opt. Express*, 2022, **13**, 26−38. https://doi.org/10.1364/boe.442489
15. Z. Zhao, W. Ma, G. Teng, X. Xu, K. Wei, G. Chen, Q. Wang, and W. Xu, *Spectrochim. Acta B*, 2023, **202**, 106644. https://doi.org/10.1016/j.sab.2023.106644
16. A. Sherbini, M. M. Hagras, H. H. Farag, and M. R. Rizk, *Int. J. Sci. Res.*, 2015, **4**, 1153−1158. https://www.ijsr.net/getabstract.php?paperid=SUB154404
17. J. Wang, L. Li, P. Yang, Y. Chen, Y. Zhu, M. Tong, Z. Hao, and X. Li, *Laser Med. Sci.*, 2018, **33**, 1381−1386. https://doi.org/10.1007/s10103-018-2500-2
18. J. H. Han, Y. Moon, J. J. Lee, S. Choi, Y. C. Kim, and S. Jeong, *Biomed. Opt. Express*, 2016, **7**, 57−66. https://doi.org/10.1364/boe.7.000057
19. Y. Liu, S. Singha, T. E. Witt, Y. Cheng, and R. J. Gordon, *Appl. Phys. Lett.*, 2008, **93**, 161502. https://doi.org/10.1063/1.3000966
20. G. A. Wubetu, H. Fiedorowicz, J. T. Costello, and T. J. Kelly, *Phys. Plasmas*, 2017, **24**, 013105. https://doi.org/10.1063/1.4973444
21. S. Zhao, Y. Zhao, Z. Hou, and Z. Wang, *Spectrochim. Acta B*, 2023, **203**, 106666. https://doi.org/10.1016/j.sab.2023.106666
22. G. Teng, Q. Wang, H. Zhang, W. Xiangli, H. Yang, X. Qi, X. Cui, B. S. Idrees, K. Wei, and M. N. Khan, *Spectrochim. Acta B*, 2020, **165**, 105787. https://doi.org/10.1016/j.sab.2020.105787
23. G. Teng, Q. Wang, Q. Hao, A. Fan, H. Yang, X. Xu, G. Chen, K. Wei, Z. Zhao, and M. N. Khan, *Biomed. Opt. Express*, 2023, **14**, 3469-3490. https://doi.org/10.1364/BOE.492983
24. E. Andrási, M. Suhajda, I. Sáray, L. Bezúr, L. Ernyei, and A. Réffy, *Sci. Total Environ.*, 1993, **139**, 399−402. https://doi.org/10.1016/0048-9697(93)90036-6
25. B. Schaefer, E. Collett, R. Smyth, D. Barrett, and B. Fraher, *Am. J. Phys.*, 2007, **75**, 163−168. https://doi.org/10.1119/1.2386162
26. X. Xu, Q. Wang, G. Teng, Z. Zhao, and K. Wei, *Spectrochim. Acta B*, 2023, **204**, 106669. https://doi.org/10.1016/j.sab.2023.106669
27. X. Xu, A. Li, X. Wang, C. Ding, S. Qiu, Y. He, T. Lu, F. He, B. Zou, and R. Liu, *J. Anal. At. Spectrom.*, 2020, **35**, 984−992. https://doi.org/10.1039/C9JA00443B
28. X. Xu, G. Teng, Q. Wang, Z. Zhao, K. Wei, M. Bao, Y. Zheng, and T. Luo, *Front. Environ. Sci.*, 2023, **11**, 1175392. https://doi.org/10.3389/fenvs.2023.1175392
29. X. Cui, Q. Wang, K. Wei, G. Teng, and X. Xu, *Plasma Sci. Technol.*, 2021, **23**, 055505. https://doi.org/10.1088/2058-6272/abf1ac
30. Q. Wang, X. Cui, G. Teng, Y. Zhao, and K. Wei, *Opt. Laser Technol.*, 2020, **125**, 106035. https://doi.org/10.1016/j.optlastec.2019.106035