Computers in Biology and Medicine 146 (2022) 105617

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Machine-learning-assisted spontaneous Raman spectroscopy classification and feature extraction for the diagnosis of human laryngeal cancer

Zheng Li [a], Zhongqiang Li [a], Qing Chen [b], Jian Zhang [b], Michael E. Dunham [c], Andrew J. McWhorter [c], Ji-Ming Feng [d], Yanping Li [e], Shaomian Yao [d], Jian Xu [a,*]

[a] Division of Electrical and Computer Engineering, College of Engineering, Louisiana State University, Baton Rouge, LA, 70803, USA
[b] Division of Computer Science & Engineering, College of Engineering, Louisiana State University, Baton Rouge, LA, 70803, USA
[c] Department of Otolaryngology, School of Medicine, Louisiana State University Health Science Center, New Orleans, LA, 70112, USA
[d] Department of Comparative Biomedical Science, School of Veterinary Medicine, Louisiana State University, Baton Rouge, LA, 70803, USA
[e] School of Environment and Sustainability, University of Saskatchewan, Saskatoon, SK S7N 5C9, Canada

ARTICLE INFO

Keywords:
Raman spectroscopy
Laryngeal cancer
Random forest
Convolutional neural network
Principal component analysis

ABSTRACT

The early detection of laryngeal cancer significantly increases the survival rates, permits more conservative larynx sparing treatments, and reduces healthcare costs. A non-invasive optical form of biopsy for laryngeal carcinoma can increase the early detection rate, allow for more accurate monitoring of its recurrence, and improve intraoperative margin control.

In this study, we evaluated a Raman spectroscopy system for the rapid intraoperative detection of human laryngeal carcinoma. The spectral analysis methods included principal component analysis (PCA), random forest (RF), and one-dimensional (1D) convolutional neural network (CNN) methods.

We measured the Raman spectra from 207 normal and 500 tumor sites collected from 10 human laryngeal cancer surgical specimens. Random Forest analysis yielded an overall accuracy of 90.5%, sensitivity of 88.2%, and specificity of 92.8% on average over 10 trials. The 1D CNN demonstrated the highest performance with an accuracy of 96.1%, sensitivity of 95.2%, and specificity of 96.9% on average over 50 trials. In predicting the first three principal components (PCs) of normal and tumor data, both RF and CNN demonstrated high performances, except for the tumor PC2.

This is the first study in which CNN-assisted Raman spectroscopy was used to identify human laryngeal cancer tissue with extracted feature weights. The proposed Raman spectroscopy feature extraction approach has not been previously applied to human cancer diagnosis. Raman spectroscopy, as assisted by machine learning (ML) methods, has the potential to serve as an intraoperative, non-invasive tool for the rapid diagnosis of laryngeal cancer and margin detection.

1. Introduction

The American Cancer Society estimates 13,000 new cases of laryngeal cancer diagnosed annually [1,2]. Approximately 3700 patients die annually from laryngeal cancer [3]. Early diagnosis and the precise control of the disease margins are essential for the surgical management of laryngeal carcinoma. Important visual features related to malignancy may be overlooked in routine laryngoscopy, and multiple biopsies are generally required for diagnosis.

The early detection of laryngeal cancer significantly increases the survival rate and allows for more conservative larynx sparing treatments. It also reduces the length of hospital stay required and associated health care costs [4]. The five-year relative survival rate for Stage I glottic carcinoma is 90%. This decreases to 74%, 56%, and 44% in Stages II, III, and IV, respectively [5]. The lack of locoregional control is the most significant cause of surgical failure.

Over the past 20 years, there has been a trend toward a more conservative, larynx sparing surgery for Stage I and II laryngeal carcinomas, which is performed using a *trans*-oral endoscopic approach [6]. Early-stage laryngeal carcinomas are generally treatable by conducting larynx sparing endoscopic surgery if adequate margins of resection are realized. Persistent positive margins may require a complete total

* Corresponding author.
E-mail address: jianxu1@lsu.edu (J. Xu).

https://doi.org/10.1016/j.compbiomed.2022.105617
Received 13 January 2022; Received in revised form 16 March 2022; Accepted 11 May 2022
Available online 18 May 2022

laryngectomy and/or radiation therapy, which can be avoided by performing a complete initial resection [7,8]. However, realizing adequate margins of resection can be challenging [9].

Optical coherence tomography (OCT) [10] is a near-infrared interferometry technique. It demonstrates a high resolution and can be used to image scales with sizes ranging approximately from 10 μm to 10 mm. It is used widely in ophthalmology for retinal imaging. There are reports on the OCT imaging of the larynx and upper airway to measure airway patency. The main disadvantage is the high cost (approximately $75,000 for a retinal imaging system). Moreover, OCT imaging is deployed with fiberoptic technology to ensure adaptability to existing endoscopes.

Near-infrared (NIR) optical fluorescence spectroscopy [11,12] incorporates a fluorescent dye such as indocyanine green (ICG) to visualize neoplastic vascularization [13]. In 2017, researchers reported the real-time NIR-ICG identification of head and neck mucosal lesions [14]. The positive findings correlated with histological malignancy with an overall accuracy of 89%.

Narrow-band imaging [15,16] uses blue-light (415 nm) and green-light (540 nm) filtering to enhance the visualization of hemoglobin. It is useful in identifying lesions with enhanced microvascular patterns, including neoplasms. The findings are nonspecific (with an accuracy of 88%, sensitivity of 92%, and specificity of 76% [16]); however, the technique is effective in characterizing neo-angiogenesis in pre-cancerous and cancerous aerodigestive lesions. It is inexpensive and can be readily employed deployed in all endoscopes. In particular, it is cost-effective for use in screening endoscopy for colon polyps; however, its effectiveness in identifying head and neck lesions is limited.

There are alternative modern imaging techniques for laryngeal cancer diagnosis, such as dual-energy CT scanning (sensitivity of 86% and specificity of 96%) [17], magnetic resonance imaging (MRI) with a sensitivity of 89–95% and specificity of 74–84% for cartilage invasion [18], and positron emission tomography (PET) - CT scanning (sensitivity of 16.7% and specificity of 97.1%) [19]. Other less frequently use modalities for laryngeal tumor imaging include NIR visual imaging with tumor photosensitizing [20] and in-vivo microscopy [21].

Raman spectroscopy exploits the inelastic scattering of incident monochromatic light exhibited by many substances. Complex biological molecules, including proteins, nucleic acids, and lipids have distinct Raman spectral signatures that have been sufficiently characterized in laboratory experiments [22]. Recent developments in Raman spectroscopy technology have allowed investigators to detect variations in the structure and concentration of biomolecules in tissue, including biochemical markers associated with neoplasia [23].

When monochromic light from a laser source incidents strikes a substance, photons are absorbed by its surface and reemitted. Most of the reemitted light is at the same frequency as the monochromatic source (elastic scattering). Depending on the sample, a small portion of the re-emitted light is radiated at frequencies above and below the incident frequency (inelastic scattering). Inelastic scattering is dependent on the molecular structure of the specimen and is referred to as the Raman effect. The molecular structure of several substances can be determined using their Raman spectra. Raman signals can be assigned to specific molecular chemical groups and chemical bond vibrational modes. Biomolecules with distinct Raman signals include proteins, nucleic acids, and lipids. Molecular interactions can exhibit distinctive spectral features [24].

A typical Raman spectroscopy system consists of a laser light source, an optical filter to define the incident frequency, and a CCD. It is common for a Raman spectroscopy system to incorporate several modifications to improve the frequency resolution and acquisition times. Modular, portable systems were recently developed for in-field implementation [25].

The Raman spectra of tissue specimens represent the weighted sum of their macromolecular species and are highly tissue-specific. Researchers identified cellular differentiation in several epithelial types using Raman spectroscopy. In-situ applications with respect to brain [26], bladder [27], breast [28], colon [29], and skin [30] neoplasms were reported. In previous studies, spectral differences were observed between normal tissue, benign lesions, and carcinoma in ex-vivo laryngeal specimens [31,32].

Principal component analysis (PCA) is a dimension reduction algorithm for data processing. Moreover, PCA is a useful technique for increasing the interpretability of large datasets while minimizing information loss. Previous researchers employed PCA for Raman spectra analysis to reduce complex, multipeak spectra to two or three principal components [33]. Supervised machine learning (ML) algorithms, with or without PCA, can be used to automate Raman spectra interpretation.

Random forests are a type of supervised ML method based on decision tree generation for the solution of classification (majority voting method) or regression (average method) problems. The basic algorithm trains multiple decision trees to predict labeled outcomes. During classification, the model traverses all the decision trees and outputs the class reached by most trees. In 2009, The et al. conducted the RF analysis of laryngeal tissue based on Raman spectra and reported a diagnostic sensitivity of 88.0% and specificity of 91.4% for laryngeal malignancy identification [34].

Neural networks are a form of artificial intelligence (AI) consisting of multiple units (neurons) arranged in layers that calculate the relationship between their inputs and outputs based on a set of parameters. Neural networks are the major focus of current research on ML and AI [35]. CNNs extend the neural network concept with the addition of specialized layers for pattern recognition [36,37]. Two-dimensional CNNs have been extensively studied in the fields of computer vision and image processing. Similarly, 1D CNNs can be implemented to interpret spectral data. In 2019 Dong et al. conducted a 1D CNN analysis of Raman spectra to discriminate human blood from animal blood [38]. There are no reports of CNN-assisted Raman spectroscopy applied to the classification of tissue specimens.

In this study, we employed RF and CNN models to automate the interpretation of the Raman spectra of resected ex-vivo laryngeal cancer specimens. In particular, we applied this approach to the discrimination of laryngeal carcinoma and non-cancerous margins. The proposed CNN-assisted Raman spectroscopy method demonstrated a significantly higher diagnostic performance than previously reported Raman spectroscopy methods for human laryngeal cancer diagnosis and achieved an accuracy of 96.1%, sensitivity of 95.2%, and specificity of 96.9%. The overall performance of the proposed model was superior to other common diagnostic methods (e.g., narrowband imaging (NBI) [16], computed tomography (CT) [17], and intraoperative frozen section analysis [21,39]) for human laryngeal cancer, with the added advantage of rapid intraoperative detection. Moreover, although CNN is effective in various classification applications, it is challenging to extract the features that are critical to the classification [40,41], and such a process has not been reported for the application of Raman spectroscopy for human cancer diagnosis. The proposed scheme can classify the laryngeal tissue and determine the feature importance from the hidden layers by integrating the weights.

2. Materials and methods

2.1. Experimental setup and instrumentation (Fig. 1)

The lab-assembled Raman spectroscopy system consisted of a Raman probe (RPS785, InPhotonics, Inc) connected to a 2 mm optical fiber, 785 nm laser diode source (Turnkey Raman Lasers-785 Series, Ocean Optics, Inc), and spectrometer with a charge-coupled device (CCD) (QE Pro; Ocean Optics, Inc). The system was interfaced with a computer for data collection and assembled at a movable workstation for portability [42]. An enclosure was used to shield the specimen and probe during the ex-vivo tissue study. For the tissue evaluations after following laryngeal cancer resection, the system was deployed in a non-sterile workspace adjacent to the operating room where the procedures were performed.

2.2. Raman spectroscopy evaluation of resected laryngeal cancer specimens

All procedures were performed at Our Lady of the Lake Hospital in Baton Rouge, Louisiana, United States of America. All subjects provided informed consent, and all the procedures were in accordance with human ethical guidelines. In particular, the study was approved by the Louisiana State University Institutional Review Board (IRB #4267).

We collected the Raman spectroscopy data off specimens remove from subjects who underwent resections for histologically confirmed laryngeal carcinoma. Based on prior evaluations of resected tissue (unpublished study) we determined the ideal incident laser parameters. The laser intensity was set at 130 mW, and the Raman spectra were sampled with exposure times of 3 s. The laser power density was a maximum of 1.99 J/cm^2. The distance between probe and sample was ~5 mm. For each sample site, we recorded 7–10 spectra. The Raman spectrum was ranged from 509 cm^{-1} to 3978 cm^{-1}, with 440 wavenumbers; the mean spectral resolution was 7–8 cm^{-1}. The data was collected ex vivo in darkness in this study as the samples were covered by a dark box.

Ten laryngeal cancer specimens were studied, including two total laryngectomy (Fig. 2) cases and eight transoral microsurgical laser resections. All specimens were oriented and labeled by the senior surgeon (Andrew J. McWhorter, MD) in accordance with existing surgical specimen processing protocols. The endoscopically resected specimens were placed in cassettes to maintain orientation with labeled margins (Fig. 3).

For each specimen, multiple Raman spectra were recorded from the labeled margins and central tumor bed. The findings were compared with the final histopathology results. In all the cases, the consulting pathologist was not provided with the spectroscopy findings.

2.3. Data processing and principal component analysis (PCA)

The recorded Raman data were preprocessed prior to analysis [43,44]. The signal processing protocol was coded using MATLAB (R2018a; MathWorks Inc, Natick, Mass) and consisted of 1) autofluorescence removal with baseline subtraction using asymmetric least squares smoothing [45], 2) Savitzky-Golay filter smoothing (the sliding window was set 10 elements; the polynomial order was 6) [46], and 3) signal normalization (the linear scaling method scaled the data to [0, 1]).

The mean Raman spectra were calculated from the labeled, preprocessed data of the cancerous and non-cancerous tissue. PCA was used for dimensional reduction and coded in Python and Sci-kit-learn [47]. The sample size was set as the number of components, and the singular value decomposition (SVD) solver was set to auto. The first three components were calculated (PC1, PC2, and PC3) with specific peak markings. The RF and CNN models were trained and tested using the PCA data to determine the predictive value of ML based on the Raman PCA datasets.

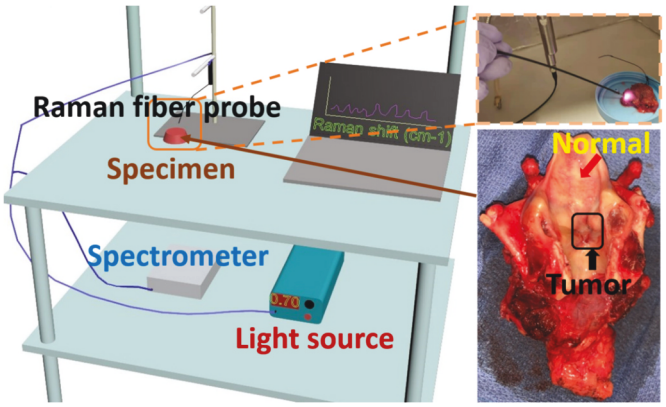

Fig. 1. Experimental setup for the ex-vivo Raman analysis of tissue.

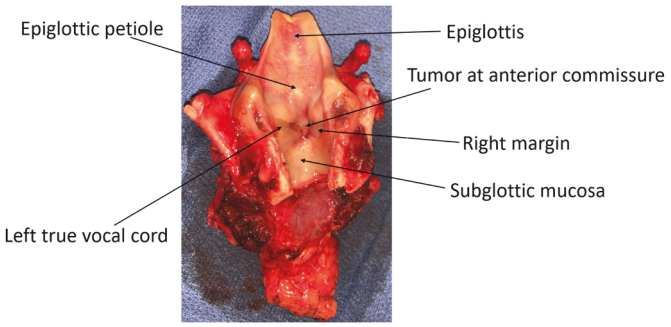

Fig. 2. Representative total laryngectomy specimen.

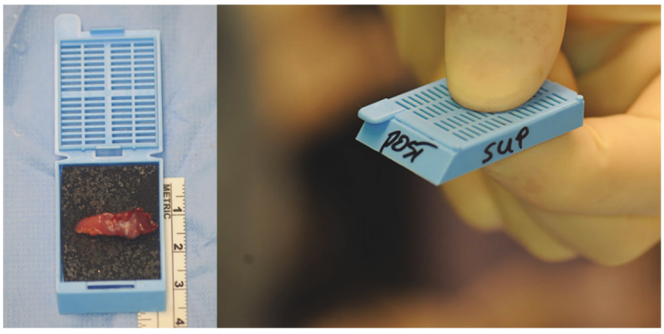

Fig. 3. Representative endoscopic resected laryngeal carcinoma specimen in a cassette with labeled margins.

2.4. Machine learning (ML) methods

The proposed random forest model was developed using Python and Sci-kit-learn. The RF model architecture is shown in Fig. 4a. The CNN model architecture consisted of one convolutional layer (32 channels) and one fully connected layer (Fig. 4b), and it was developed using Python and Pytorch [48]. The CNN model had a kernel size of five and a stride of two; the loss function was binary crossentropy (BCEWithLogitsLoss), with the learning rate of 0.01, momentum of 0.9 and weight decay of 0.00004; the optimizer was Stochastic Gradient Descent (SGD).

For both models, we identified the Raman spectral features corresponding to the cancerous and non-cancerous tissue by examining the trained model weights for individual peaks (RF) and wavenumber ranges (CNN). The features were ranked in the order of decreasing magnitude.

3. Results

A total of 500 Raman spectra were recorded from histologically confirmed cancer-positive biopsy sites and 207 confirmed negative margins. Discrete spectra were recorded from 509 to 3978 cm^{-1} across 440 wavenumbers. For the classifier, we randomly selected 207 of the cancer spectra to generate a balanced dataset with a total of 414 recordings.

3.1. Random forest classification of resected laryngeal carcinoma specimens

RF classification was assessed using 90% of the combined dataset for training and 10% for testing. The average accuracy was 90.5%. The sensitivity and specificity were 88.2% and 92.8%, respectively. The RF-assisted receiver operating characteristic (ROC) curve is shown in Fig. 5. The area under the curve (AOC) was 0.964.

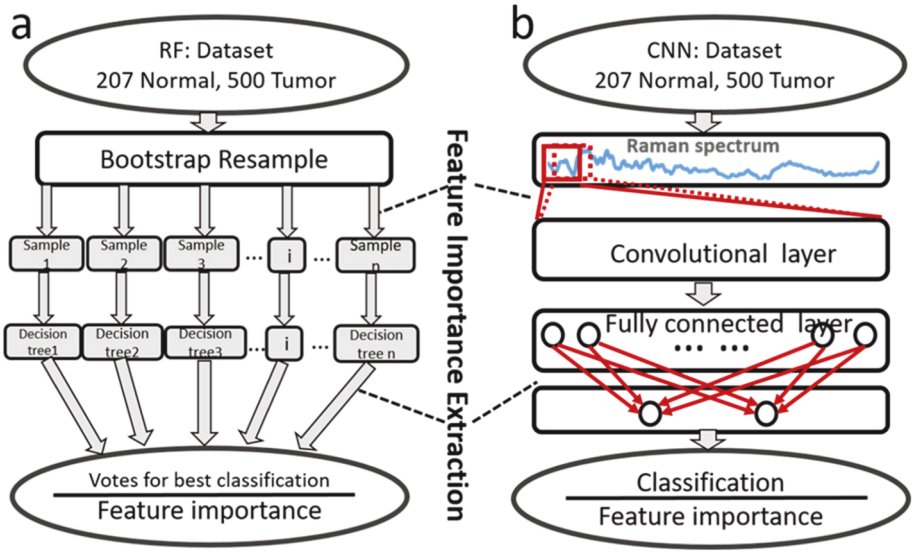

Fig. 4. Schematics of the (a) RF and (b) CNN models for the classification of resected laryngeal tissue.

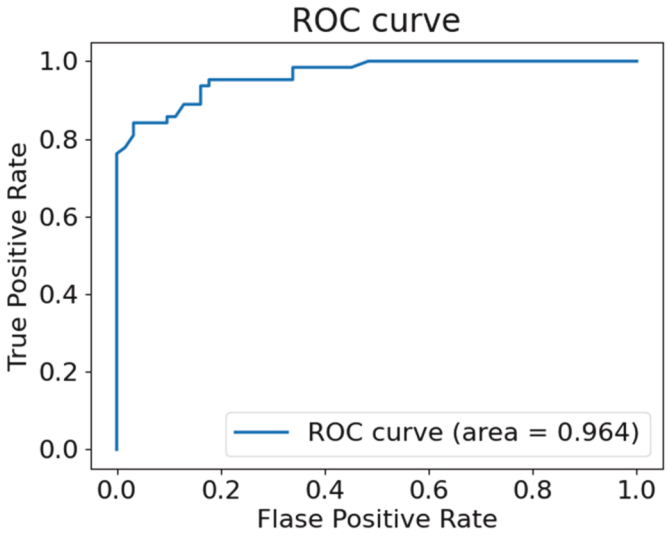

Fig. 5. Receiver operating characteristic (ROC) curve for the random forest classification of resected laryngeal carcinoma specimens.

The averaged weights calculated for the branches in the RF model decision trees corresponding to individual Raman peaks are shown in Table 1, which presents the 24 highest weights.

The 3756 cm $^{-1}$ peak was the most weighted feature. The RF model assigned weights greater than 1% to 12 peaks. The sum of the highest 24 feature weights was 0.259.

Table 1
Random forest model feature weights.

Wavenumber (cm-1)	Weight	Wavenumber (cm-1)	Weight
3756	0.0184	1573	0.0099
1536	0.0155	3570	0.0098
2381	0.0150	1592	0.0095
1601	0.0133	2983	0.0088
1527	0.0129	1564	0.0086
1583	0.0126	2776	0.0086
1545	0.0124	1647	0.0085
1638	0.0117	3509	0.0084
3761	0.0114	1384	0.0084
2783	0.0109	1365	0.0083
1374	0.0108	2990	0.0076
2920	0.0107	911	0.0074

3.2. Convolutional neural network classification of resected laryngeal carcinoma specimens

For CNN training and testing, a 90%/10% random data allocation was used for 50 epochs/run. Convergence was observed after 30 epochs (Fig. 6), and the results were averaged over 50 runs.

The overall accuracy was 96.1%, and the sensitivity and specificity were 95.2% and 96.9%, respectively. The CNN-assisted ROC curve is shown in Fig. 7. The AOC was 0.980.

Table 2 presents the top 26 features by weight in the CNN model, corresponding to spectral wavenumber ranges most important for the binary cancer classifier. The weights were extracted and summed from 32 mappings in the linear layer. There was a wide distribution from 3744 to 3779 cm^{-1} in the high-wavenumber region, which was consistent with the 3756 cm^{-1} and 3761 cm^{-1} peak weights in the RF model. Moreover, the ranges of 921–963 cm^{-1} and 1461 - 1498 cm^{-1} in the fingerprint region were prominent, and both had regression weights approximately equal to 1, which is consistent with existing empirical data [49,50].

The single wavenumber weights were calculated by averaging the range weight, followed by normalization. For example, the average weight corresponding to the first five wavenumbers (w1, w2, w3, w4, and w5), averaged across the 32 convolution mappings, was 0.119; and the average weight corresponding to the second five wavenumbers (w3, w4, w5, w6, and w7, stride = 2) was 0.038. The weight corresponding to the wavenumber w3 was calculated as the mean of the two: (0.119 + 0.038)/2 = 0.0785. The averaged weights were normalized by dividing by the sum of the weights. Table 3 illustrates the 36 largest single wavenumber weights for the CNN model.

3.3. PCA and combined PCA-ML analysis of spontaneous Raman scattering data of human laryngeal cancer

The PCA analysis is summarized in Fig. 8. PC1 plot for laryngeal cancer revealed distinct Raman signals at 727 cm^{-1}, 958 cm^{-1}, and 1553 cm^{-1}. Non-cancerous tissue exhibited stronger signals at 1125 cm^{-1}, 1434 cm^{-1}, and 1655 cm^{-1} for the Raman spectra without dimensional reduction. The PC1 plot was similar to the unreduced data.

The tissue prediction using the RF model and PCA data (PC1, PC2, and PC3) from 10 tests, wherein each test involved 10 runs

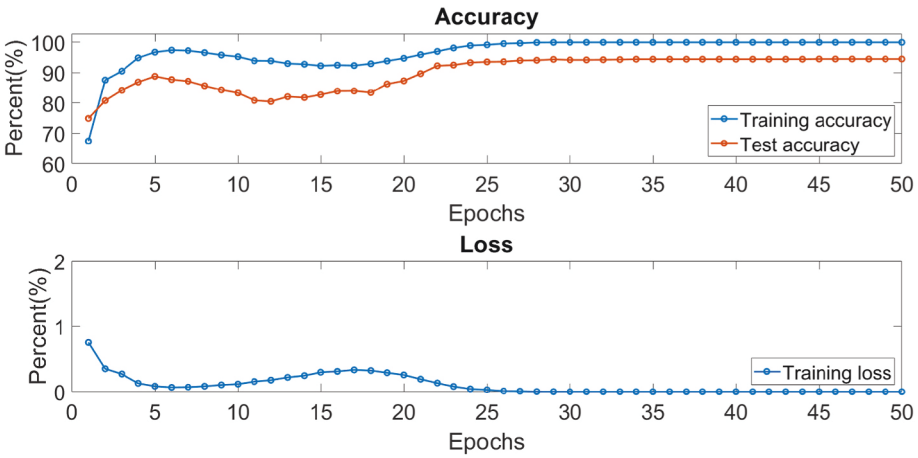

Fig. 6. Training and testing accuracy and training loss convergence plots for the CNN classification of resected laryngeal carcinoma specimens.

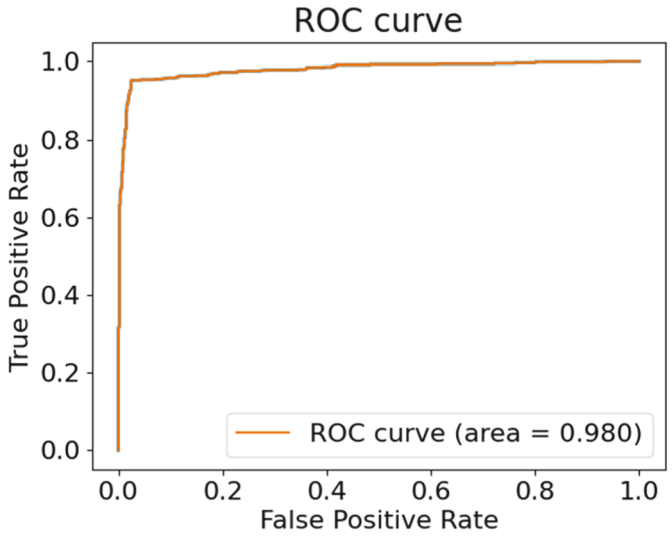

Fig. 7. ROC curve for the CNN classification of resected laryngeal carcinoma specimens.

Table 2
Ranges of feature importance weights for CNN classification of laryngeal cancer.

Wavenumber (cm-1)	Regression weight	Wavenumber (cm-1)	Regression weight
921–963	1.024	2887–2906	0.219
1461–1498	0.939	3552–3577	0.216
3744–3767	0.851	1046–1087	0.214
3756–3779	0.797	1188–1228	0.203
1517–1564	0.451	1647–1684	0.192
2891–2920	0.448	2990–3018	0.175
2613–2643	0.439	1208–1247	0.175
3564–3589	0.421	2962–2990	0.167
1403–1442	0.384	1536–1573	0.167
2776–2805	0.360	575–620	0.165
2805–2834	0.330	2762–2791	0.161
1066–1107	0.293	773–816	0.157
1845–1880	0.250	1555–1592	0.155

demonstrated the specificity of 100% with respect to normal tissue. The tumor prediction sensitivity was 39% using PC1, and no tumor spectra were accurately predicted using PC2. Moreover, the tumor prediction sensitivity was 100% using PC3.

Using the same dataset and the CNN model for 10 tests, the prediction specificity for normal tissue was 100%, 100%, and 95% for PC1, PC2, and PC3, respectively. The tumor tissue sensitivity was 100%, 2%, and 100% for PC1, PC2, and PC3, respectively.

Table 3
Feature importance weights for CNN classification of laryngeal cancer (single wavenumber).

Wavenumber (cm-1)	Weight	Wavenumber (cm-1)	Weight
3761	0.02085	3744	0.00774
3767	0.01422	3779	0.00735
3756	0.01420	1527	0.00681
953	0.01381	2783	0.00659
932	0.01345	1555	0.00653
1470	0.01285	3564	0.00652
1489	0.01283	1077	0.00641
3750	0.01122	2913	0.00628
3773	0.01056	2805	0.00626
963	0.00965	2635	0.00620
942	0.00954	2620	0.00616
921	0.00950	2891	0.00606
1498	0.00929	2906	0.00603
1461	0.00929	1536	0.00595
1480	0.00920	1432	0.00594
2899	0.00844	3583	0.00579
3570	0.00806	1413	0.00571
1545	0.00782	3577	0.00568

4. Discussions

Laryngeal cancer is diagnosed histologically. Frozen section analysis is the current standard method for intraoperative diagnoses and margin assessments [39,51]. This method is expensive and time-consuming. In particular, 20–30 min is required for intraoperative frozen section analysis, and several hours or days are required for postoperative histopathology. Moreover, only a limited number of biopsy points are inspected. Numerous lesions spread in a submucosal plane and multifocal occurrences ("field cancerization") are common. The identification of optimal best sites for biopsy can be difficult and numerous patients undergo multiple biopsy procedures prior to the confirmation of a diagnosis [52]. A non-invasive, rapid (<1 min) optical form of biopsy for laryngeal carcinoma is, therefore, required to improve early detection, allow for more accurate monitoring of its recurrence, and improve intraoperative margin control.

In this work, 207 cancer and 500 normal tissue Raman spectra were collected and analyzed from 10 human laryngeal cancer clinical studies. The ML techniques, including PCA, RF, and CNN, were applied for the Raman spectral analysis to evaluate the classification performance. The accuracy of 90.5%, the sensitivity of 88.2%, and the specificity of 92.8% of the RF model were similar to the previous report on laryngeal cancer

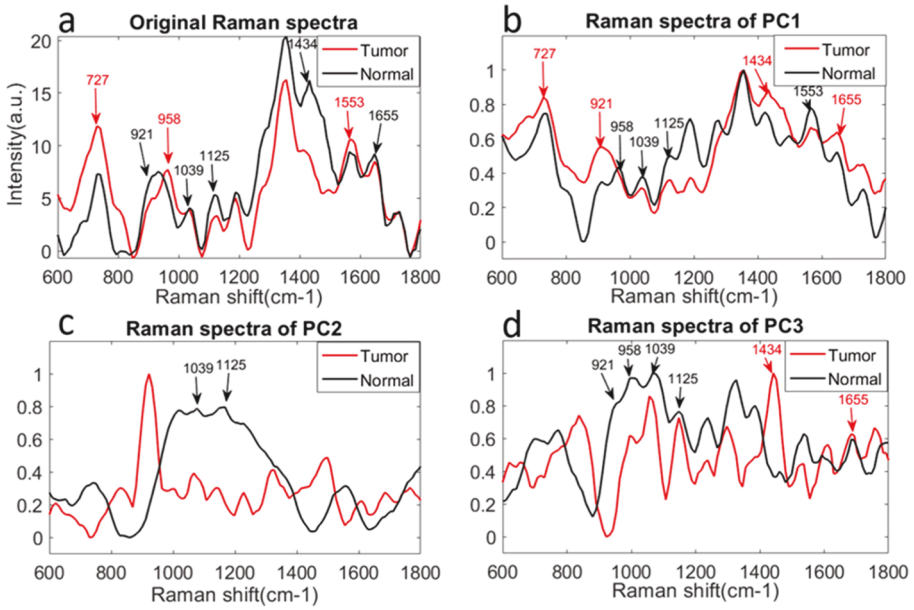

Fig. 8. PCA results of Raman spectra of resected cancerous and non-cancerous laryngeal tissue: a) original Raman spectra; b) PC1; c) PC2; d) PC3. (Key peaks are labeled).

classification by Raman spectroscopy and Random Forest classification algorithm with voting for decision [34]. To the best of our knowledge, no study of human laryngeal cancer Raman spectra analysis by CNN has been reported before, and the classification performances by CNN models, with the accuracy of 96.1%, sensitivity of 95.2%, and specificity of 96.9% are significantly better than the RF model. Our 1D–CNN–aided Raman spectroscopy also outperformed any previous Raman spectroscopy studies for the identification of human laryngeal cancers and most other common diagnostic schemes for laryngeal cancer, with faster diagnosis speed (<1 min) and the potential for intraoperative diagnosis.

Lin et al. published the only report on in-vivo Raman spectroscopy for laryngeal lesions, with reference to the available literature [53]. They documented the spectral discrimination between cancerous and non-cancerous tissue during laryngoscopy. The proposed Raman spectroscopy system conducted simultaneous wave number analysis within the ranges of 800–1800 cm^{-1} and 2800 - 3600 cm^{-1}. Moreover, the system included a near-infrared diode laser source (785 nm) and thermo-electric cooled CCD. The system demonstrated a diagnostic accuracy of 91.1% (sensitivity of 93.3%, specificity of 90.1%) for laryngeal cancer. While our methods hugely improve the performance as the increase of 1% is an excellent work in the machine learning field. In addition, many important feature peaks had been extracted from the RF and CNN models, which has significant influence on cancer tissue component analysis. Another novel approach of our work is using PCs for prediction, which could be applied for new laryngeal Raman data predication. Both PC1 and PC3 showed excellent predictions results. The possible reason of 2nd PCA component did not perform well for prediction was PC2 did not extract the significant peaks for classification. As shown in Fig. 8, the Raman spectra of PC2 contained less similar peaks compared with PC1 and PC3 to the original Raman spectra, which indicated PC1 and PC3 remained more useful peak information.

In addition to the superb classification performance of cancer and normal tissue of Raman spectra, we also extracted critical features by our ML models, which has always been challenging [40]. Fukuhara et al. achieved important regions by extracting spectra from the pooling layer and fully connected layer [54]. The RF model is based on decision trees, a special method of Bootstrap Aggregating: majority voting is for classification, the average is for regression. As our approach is classification, we need to use the voting method. In our RF model, each decision tree contributes the feature importance of classification, generating weight at each node, and 440 weights for 440 features with the sum of 1 of all feature weights. In the CNN model, every five features attribute to the range weight due to the kernel size of 5; both ranges weights and single wavenumber weights are recorded and calculated from CNN layers. The weight reflects how important the feature is to some extent, which means the larger the weight is, the more this feature contributes to the classification of cancer and normal tissues. In our CNN model, only one convolutional layer was applied, which has already resulted in excellent classification performance; the single convolutional layer analysis also facilitated the critical feature importance extraction. Comparing the weights of RF and CNN, we found some important features in both models, such as 1) ~3756 cm-1/3744 cm-1 - 3767 cm-1.2) ~1536 cm-1/1517 cm-1 - 1564 cm-1.3) ~1545 cm-1/1536 cm-1 - 1573 cm-1, and so on.

Different components or chemical bonds result in various peak intensities and wavenumbers of Raman signals. The results revealed prominent laryngeal cancer peaks at 727 cm^{-1} (deoxyribonucleic acid (DNA)-based adenine, bending) [55], and 1553 cm^{-1} (DNA-based Guanine, bending). Non-cancerous laryngeal tissue exhibited stronger Raman signals at 1125 cm^{-1} (proteins, C–N stretching), and 1655 cm^{-1} (amide I, C=O stretching). The findings were consistent with the known biochemical and cellular transformations in cancerous tissue with increased nucleic acids to protein/lipid ratios [24,32].

In previous studies conducted on biochemical, the importance of DNA glutathione (GSH) as a marker in human head and neck cancer was reported [56,57]. Glutathione (GSH) [58] is identified in Raman spectra with typical wavenumbers at ~1365 cm^{-1}, 1536 cm^{-1}, and ~1638 cm^{-1} in the RF model. These key identifiers may play a role in cancerous tissue classification. Additional GSH related peaks were identified by the CNN model, including ~921 cm^{-1}, 932 cm^{-1}, 953 cm^{-1}, ~963 cm^{-1}, ~1077 cm^{-1}, ~1413 cm^{-1}, ~1461 cm^{-1}, and 1536 cm^{-1}. The 1536 cm^{-1} peak was identified by both the RF and CNN models.

In this study, we evaluated the classification capacities of the RF and CNN models with respect to cancerous vs non-cancerous tissue when trained using the original spectra and the first three principal components of the preprocessed data. With respect to the available literature, this is the first study on CNN-assisted Raman spectroscopy for human laryngeal cancer analysis. In previous studies, the overall accuracy of PCA-linear discriminant analysis (LDA) was 84.3% for the detection of colorectal cancer in ex-vivo samples [59]. Teh et al. used NIR Raman

spectroscopy to diagnose laryngeal malignant cancerous tissue and non-cancerous tissue, with the sensitivity of 92.9% and specificity of 83.3% [60]. Lin et al. also achieved a diagnostic 90.3% sensitivity and 90.9% specificity by applying high-wavenumber Raman spectroscopy for laryngeal cancer diagnosis [61]. Our results indicated improved accuracy using ML algorithms, even with a small dataset, when compared with empirical peak assignments. For the binary classification of cancerous and non-cancerous of the laryngeal tissue, the CNN modes outperformed the RF model in terms of specificity, sensitivity, and overall accuracy. The CNN-assisted Raman spectroscopy model demonstrated a significantly higher diagnosis performance than Raman spectroscopy systems proposed in previous studies and most traditional diagnosis methods (e.g., NBI, CT, MRI, and frozen section analysis) for human laryngeal cancer [16–18,39]. Moreover, the important features were extracted to facilitate the analysis of biochemical changes among different classes of tissues. The excellent classification performance and the extracted Raman features with the bio-chemical changes of laryngeal cancer (e.g., GSH biomarkers), we believe that the resolution of the spectrometer and this Raman system is good enough for the laryngeal cancerous and non-cancerous tissue classification.

5. Conclusions

In this work, we employed various machine learning algorithms to effectively classify laryngeal cancer and normal tissues based on their spontaneous Raman scattering signals. The rapid spectrographic classification of cancerous vs non-cancerous tissue during laryngeal carcinoma resection can be realized. Raman spectroscopy integrated with ML allows system designers to automate the classification process, and the CNN model is the most effective ML approach. In particular, the proposed CNN-assisted Raman spectroscopy system demonstrated an accuracy of 96.1%, sensitivity of 95.2%, and specificity of 96.9%, which are significantly superior to those of Raman spectroscopy systems proposed in previous studies conducted on human laryngeal cancer diagnosis and most of the traditional diagnosis schemes in this field. We also extracted the critical features that are important in the classification of laryngeal tissues with RF and CNN models. Both models showed the larger weight in the Raman signal that corresponds to the glutathione signals, which is critical to the classification of laryngeal cancer; this extracted feature is consistent with the previous biochemical finding that glutathione significantly increases in laryngeal cancer, as compared to non-cancerous tissue. Furthermore, the rapid diagnosis and margin delineation of laryngeal carcinoma is within the scope of current Raman spectroscopy technology. Overall, the CNN-assisted Raman spectroscopy showed excellent accuracy, sensitivity, and specificity in classifying laryngeal cancer from normal tissues, which may potentially become a rapid precise diagnostic tool for the early or intraoperative detection of laryngeal carcinoma in the future.

Author contributions

Zheng Li collected data and wrote the initial manuscript. Zhongqiang Li helped to collect data. Qing Chen and Jian Zhang helped to process the data and review the manuscript review. Michael E. Dunham and Andrew J McWhorter helped to collect the specimen and review the manuscript. Ji-Ming Feng, and Yanping Li and Shaomian Yao reviewed the manuscript. Jian Xu conceived the project design, discussed the results, and supervised the project.

Declaration of competing interest

All the authors declare no potential conflicts of interest with respect to the authorship and/or publication in this article.

Acknowledgments

This research was supported by NSF CAREER award [2046929]. Louisiana State University Economic Development Assistantships [000398], LSU Leveraging Innovation for Technology Transfer (LIFT2) Grant [LSU-2021-LIFT-009, LSU-2020-LIFT-008], Health Sciences Center New Orleans, Louisiana State University Grant [HSCNO-2019-LIFT-004], Louisiana Board of Regents Grant [LEQSF (2018-21)-RD-A-09, LEQSF(2020-21)-RD-D-02], Louisiana State University Faculty Research Grants [009875].

References

[1] S.M. Cohen, J. Kim, N. Roy, C. Asche, M. Courey, Direct health care costs of laryngeal diseases and disorders, Laryngoscope 122 (2012) 1582–1588.
[2] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2020, CA A Cancer J. Clin. 70 (2020) 7–30.
[3] A.C. Society, Key Statistics for Laryngeal and Hypopharyngeal Cancers, 2021.
[4] D.J. Arnold, G.F. Funk, L.H. Karnell, A.H. Chen, H.T. Hoffman, J.M. Ricks, M. B. Zimmerman, D.P. Corbae, W. Zhen, T.M.J.T.L. McCulloch, Laryngeal Cancer Cost Analysis: Association of Case-mix and Treatment Characteristics with Medical Charges, vol. 110, 2000, pp. 1–7.
[5] A.C. society, Survival Rates for Laryngeal and Hypopharyngeal Cancers, 2020.
[6] H.T. Hoffman, K. Porter, L.H. Karnell, J.S. Cooper, R.S. Weber, C.J. Langer, K. K. Ang, G. Gay, A. Stewart, R.A.J.T.L. Robinson, Laryngeal cancer in the United States: changes in demographics, patterns of care, Survival 116 (2006) 1–13.
[7] J. Hanna, P.R. Brauer, E. Morse, S.J.O.H. Mehra, N. Surgery, Margins in laryngeal squamous cell carcinoma treated with transoral laser microsurgery, Nat. Database. Stud. 161 (2019) 986–992.
[8] I. Fiz, F. Mazzola, F. Fiz, F. Marchi, M. Filauro, A. Paderno, G. Parrinello, C. Piazza, G.J.F.i.o. Peretti, Impact of Close and Positive Margins in Transoral Laser Microsurgery for Tis–T2 Glottic Cancer, vol. 7, 2017, p. 245.
[9] J. Hamman, C.L. Howe, M. Borgstrom, A. Baker, S.J. Wang, S.J.T.L. Bearelly, Impact of Close Margins in Head and Neck Mucosal Squamous Cell Carcinoma: A Systematic Review, 2021.
[10] M.J. Gora, M.J. Suter, G.J. Tearney, X.J.B.o.e. Li, Endoscopic Optical Coherence Tomography: Technologies and Clinical Applications, vol. 8, 2017, pp. 2405–2444.
[11] Z. Li, Y.V. Holamoge, Z. Li, W. Zaid, M.L. Osborn, A. Ramos, J.T. Miller, Y. Li, S. Yao, J. Xu, Detection and Analysis of Enamel Cracks by ICG-NIR Fluorescence Dental Imaging, vol. 1475, 2020, pp. 52–63.
[12] Z. Li, Z. Li, Y.V. Holamoge, W. Zaid, M.L. Osborn, S. Bhatta, A. Jeyaseelan, Y. Li, S. Yao, J. Xu, Mouthwash to deliver indocyanine green for near infrared dental fluorescence imaging, IEEE J. Sel. Top. Quant. Electron. 27 (2021) 1–8.
[13] E. Haustein, P.J.H.j. Schwille, Trend. Fluroscence Imag. Related. Tech. Untrvel. Biol. Info. 1 (2007) 169–180.
[14] F. Schmidt, A. Dittberner, S. Koscielny, I. Petersen, O.J.H. Guntinas–Lichius, Neck, Feasibility of Real-time Near-infrared Indocyanine Green Fluorescence Endoscopy for the Evaluation of Mucosal Head and Neck Lesions, vol. 39, 2017, pp. 234–240.
[15] C. Piazza, F. Del Bon, G. Peretti, P.J.C. Nicolai, head, n. surgery, Narrow Band Imaging in Endoscopic Evaluation of the Larynx, vol. 20, 2012, pp. 472–476.
[16] Z. Michal, L. Petr, L. Eva, B. Jan, P. Jan, The role of narrow band imaging in the detection of recurrent laryngeal and hypopharyngeal cancer after curative radiotherapy, BioMed Res. Int. (2014) 2014.
[17] H. Kuno, H. Onaya, R. Iwata, T. Kobayashi, S. Fujii, R. Hayashi, K. Otani, H. Ojiri, T. Yamanaka, M. Satake, Evaluation of cartilage invasion by laryngeal and hypopharyngeal squamous cell carcinoma with dual-energy CT, Radiology 265 (2012) 488–496.
[18] V.M. Joshi, V. Wadhwa, S.K. Mukherji, Imaging in laryngeal cancers, Indian J. Radiol. Imag. 22 (2012) 209–226.
[19] A. Rosko, A. Birkeland, A. Shuman, M. Prince, C. Bradford, G. Wolf, F. Worden, A. Eisbruch, A. Srinivasan, K.K. Wong, M.E. Spector, Positron emission tomography-CT prediction of occult nodal metastasis in recurrent laryngeal cancer, Head Neck 39 (2017) 980–987.
[20] X. Zhou, H. Li, C. Shi, F. Xu, Z. Zhang, Q. Yao, H. Ma, W. Sun, K. Shao, J. Du, S. Long, J. Fan, J. Wang, X. Peng, An APN-Activated NIR Photosensitizer for Cancer Photodynamic Therapy and Fluorescence Imaging, Biomaterials, 2020, p. 253.
[21] D. Cikojevic, I. Gluncic, V. Pesutic-Pisac, Comparison of contact endoscopy and frozen section histopathology in the intra-operative diagnosis of laryngeal pathology, J. Laryngol. Otol. 122 (2008) 836–839.
[22] P. Rostron, S. Gaber, D.J.l. Gaber, Raman. Spectroscope. Rev. 21 (2016) 24.
[23] A. Mahadevan-Jansen, R.R. Richards-Kortum, Raman spectroscopy for the detection of cancers and precancers, J. Biomed. Opt. 1 (1996) 31–70.
[24] S.P. Naber, Molecular pathology–detection of neoplasia, N. Engl. J. Med. 331 (1994) 1508–1510.
[25] P. Colomban, The On-Site/remote Raman Analysis with Mobile Instruments: a Review of Drawbacks and Success in Cultural Heritage Studies and Other Associated Fields, 2012, pp. 1529–1535.
[26] M. Jermyn, K. Mok, J. Mercier, J. Desroches, J. Pichette, K. Saint-Arnaud, L. Bernstein, M.-C. Guiot, K. Petrecca, F.J.S.t.m. Leblond, Intraoperative Brain Cancer Detection with Raman Spectroscopy in Humans, vol. 7, 2015, 274ra219-274ra219.

[27] S. Li, L. Li, Q. Zeng, Y. Zhang, Z. Guo, Z. Liu, M. Jin, C. Su, L. Lin, J.J.S.r. Xu, Characterization and Noninvasive Diagnosis of Bladder Cancer with Serum Surface Enhanced Raman Spectroscopy and Genetic Algorithms, vol. 5, 2015, pp. 1–7.
[28] A.S. Haka, Z.I. Volynskaya, J.A. Gardecki, J. Nazemi, R. Shenk, N. Wang, R. R. Dasari, M. Fitzmaurice, M.S. Feld, Diagnosing Breast Cancer Using Raman Spectroscopy: Prospective Analysis, vol. 14, 2009, p. 54023.
[29] L. Wenjing, S. Zhaotian, C. Jinyu, J. Chuanbo, Raman spectroscopy in colorectal cancer diagnostics: comparison of PCA-LDA and PLS-DA models, J. Spectroscope. (2016) 2016.
[30] J. Zhao, H. Lui, D.I. McLean, H. Zeng, Real-time Raman spectroscopy for non-invasive skin cancer detection-preliminary results, in: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2008, pp. 3107–3109.
[31] D.P. Lau, Z. Huang, H. Lui, D.W. Anderson, K. Berean, M.D. Morrison, L. Shen, H.J. L.i.S. Zeng, M.T.O.J.o.t.A.S.f.L. Medicine, Surgery, Raman spectroscopy for optical diagnosis in the larynx, Preliminary Find. 37 (2005) 192–200.
[32] N. Stone, P. Stavroulaki, C. Kendall, M. Birchall, H. Barr, Raman spectroscopy for early detection of laryngeal malignancy: preliminary results, Laryngoscope 110 (2000) 1756–1763.
[33] H. Yao, Z. Tao, M. Ai, L. Peng, G. Wang, B. He, Y.-q. Li, Raman spectroscopic analysis of apoptosis of single human gastric cancer cells, Vib. Spectrosc. 50 (2009) 193–197.
[34] S.K. Teh, W. Zheng, D.P. Lau, Z. Huang, Spectroscopic diagnosis of laryngeal carcinoma using near-infrared Raman spectroscopy and random recursive partitioning ensemble techniques, Analyst 134 (2009) 1232–1239.
[35] O.I. Abiodun, A. Jantan, A.E. Omolara, K.V. Dada, N.A. Mohamed, H. Arshad, State-of-the-art in artificial neural network applications: a survey, Heliyon 4 (2018), e00938.
[36] A. Ajit, K. Acharya, A. Samanta, Engineering, A Review of Convolutional Neural Networks, 2020, pp. 1–5.
[37] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell, Caffe: convolutional architecture for fast feature embedding, in: Proceedings of the 22nd ACM International Conference on Multimedia, Association for Computing Machinery, Orlando, Florida, USA, 2014, pp. 675–678.
[38] J. Dong, M. Hong, Y. Xu, X. Zheng, A Practical Convolutional Neural Network Model for Discriminating Raman Spectra of Human and Animal Blood, vol. 33, 2019, e3184.
[39] L.J. DiNardo, J. Lin, L.S. Karageorge, C.N. Powers, Accuracy, utility, and cost of frozen section margins in head and neck cancer surgery, Laryngoscope 110 (2000) 1773–1776.
[40] Q.-s. Zhang, S.-c. Zhu, Visual interpretability for deep learning: a survey, Front. Info. Tech. Electronic. Eng. 19 (2018) 27.
[41] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-CAM: visual explanations from deep networks via gradient-based localization, Int. J. Comput. Vis. 128 (2020) 336.
[42] Z. Li, W. Zaid, T. Hartzler, A. Ramos, M.L. Osborn, Y. Li, S. Yao, J. Xu, Indocyanine green-assisted dental imaging in the first and second near-infrared windows as compared with X-ray imaging, Ann. N. Y. Acad. Sci. 1448 (2019) 42–51.
[43] V. Mazet, C. Carteret, D. Brie, J. Idier, B.J.C. Humbert, I.L. Systems, Background Removal from Spectra by Designing and Minimising a Non-quadratic Cost Function, vol. 76, 2005, pp. 121–133.
[44] E. Cordero, F. Korinth, C. Stiebing, C. Krafft, I.W. Schie, J. Popp, Evaluation of shifted excitation Raman difference spectroscopy and comparison to computational background correction, Method. Appl. Biochem. Raman Spectra. 17 (2017) 1724.
[45] S. He, W. Zhang, L. Liu, Y. Huang, J. He, W. Xie, P. Wu, C. Du, Baseline correction for Raman spectra using an improved asymmetric least squares method, Anal. Methods 6 (2014) 4402–4407.
[46] S.J. Barton, T.E. Ward, Bryan M. Hennelly, Algorithm for optimal denoising of Raman spectra, Anal. Methods 10 (2018) 3759–3769.
[47] G.V. Fabian Pedregosa, Alexandre Gramfort, Vincent Michel, Thirion Bertrand, Olivier Grisel, Mathieu Blondel, Prettenhofer Peter, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay, Scikit-learn: Machine Learning in Python, 2011.
[48] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L.J.a.p.a. Antiga, Pytorch: an Imperative Style, High-Performance Deep Learning Library, 2019.
[49] B.-C. Chen, J. Sung, S.-H. Lim, Chemical imaging with frequency modulation coherent anti-Stokes Raman scattering microscopy at the vibrational fingerprint region, J. Phys. Chem. B 114 (2010) 16871–16880.
[50] R. Ranjan, M.A. Ferrara, L. Sirleto, Extending Femtosecond Stimulated Raman Microscopy toward Silent and Fingerprint Region of Biomolecules, IEEE, 2020, pp. 1–4.
[51] H. Jaafar, Intra-operative Frozen Section Consultation: Concepts, Applications and Limitations, vol. 13, 2006, p. 4.
[52] B. von Stülpnagel, R. Hagen, B. Olzowy, G. Witt, H.W. Pau, T. Just, Comparative Study between the Surgeon's Intraoperative Evaluation and Histopathology for Diagnosis of Laryngeal Lesions, ISRN Otolaryngology, 2014, pp. 1–6.
[53] K. Lin, W. Zheng, C.M. Lim, Z. Huang, Real-time in vivo diagnosis of laryngeal carcinoma with rapid fiber-optic Raman spectroscopy, Biomed. Opt Express 7 (2016) 3705–3715.
[54] M. Fukuhara, K. Fujiwara, Y. Maruyama, H. Itoh, Feature Visualization of Raman Spectrum Analysis with Deep Convolutional Neural Network, 2020.
[55] L. Haichun, W. Zhengdong, W. Qiong, L. Shuchun, L. Chen, W. Desheng, Surface-enhanced Raman spectroscopy for classification of laryngeal cancer and adjacent tissues, Laser Phys. 29 (2019), 1-1.
[56] Y. Zou, M. Li, Y. Xing, T. Duan, X. Zhou, F. Yu, Bioimaging of glutathione with a two-photon fluorescent probe and its potential application for surgery guide in laryngeal cancer, ACS Sens. 5 (2020) 242–249.
[57] T.P. Mulder, J.J. Manni, H.M. Roelofs, W.H. Peters, A. Wiersma, Glutathione S-transferases and glutathione in human head and neck cancer, Carcinogenesis 16 (1995) 619–624.
[58] J. De Gelder, K. De Gussem, P. Vandenabeele, L. Moens, Reference database of Raman spectra of biological molecules, J. Raman Spectrosc. 38 (2007) 1133–1147.
[59] S. Guo, P. Rösch, J. Popp, T. Bocklitz, Modified PCA and PLS: towards a better classification in Raman spectroscopy–based biological applications, J. Chemometr. 34 (2020) 1–10.
[60] S.K. Teh, W. Zheng, D.P. Lau, Z. Huang, Raman Spectroscopy for Optical Diagnosis of Laryngeal Cancer, Photonic Therapeutics and Diagnostics IV, International Society for Optics and Photonics, 2008, p. 68421S.
[61] K. Lin, D.L.P. Cheng, Z.J.B. Huang, Bioelectronics, Optical Diagnosis of Laryngeal Cancer Using High Wavenumber Raman Spectroscopy, vol. 35, 2012, pp. 213–217.