Neural networks improve brain cancer detection with Raman spectroscopy in the presence of operating room light artifacts

Michael Jermyn
Joannie Desroches
Jeanne Mercier
Marie-Andrée Tremblay
Karl St-Arnaud
Marie-Christine Guiot
Kevin Petrecca
Frederic Leblond

Neural networks improve brain cancer detection with Raman spectroscopy in the presence of operating room light artifacts

Michael Jermyn,a,b,* Joannie Desroches,b Jeanne Mercier,b Marie-Andrée Tremblay,b Karl St-Arnaud,b Marie-Christine Guiot,a Kevin Petrecca,a and Frederic Leblondb,d

aMcGill University, Montreal Neurological Institute and Hospital, Department of Neurology and Neurosurgery, 3801 University Street, Montreal, Quebec H3A 2B4, Canada
bPolytechnique Montreal, Department of Engineering Physics, CP 6079, Succ. Centre-Ville, Montreal, Quebec H3C 3A7, Canada
cCentre de Recherche du Centre Hospitalier de l’Université de Montréal, 900 rue Saint-Denis, H2X 0A9 Quebec, Canada
dUniversité de Montréal, 900 rue Saint-Denis, H2X 0A9 Quebec, Canada

Abstract. Invasive brain cancer cells cannot be visualized during surgery and so they are often not removed. These residual cancer cells give rise to recurrences. In vivo Raman spectroscopy can detect these invasive cancer cells in patients with grade 2 to 4 gliomas. The robustness of this Raman signal can be dampened by spectral artifacts generated by lights in the operating room. We found that artificial neural networks (ANNs) can overcome these spectral artifacts using nonparametric and adaptive models to detect complex nonlinear spectral characteristics. Coupling ANN with Raman spectroscopy simplifies the intraoperative use of Raman spectroscopy by limiting changes required to the standard neurosurgical workflow. The ability to detect invasive brain cancer under these conditions may reduce residual cancer remaining after surgery and improve patient survival. © 2016 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.21.9.094002]

Keywords: Raman spectroscopy; brain cancer; neural networks; neurosurgery; machine learning.

Paper 160196SSPR received Mar. 29, 2016; accepted for publication Aug. 18, 2016; published online Sep. 8, 2016.

1 Introduction

Gliomas are a type of brain tumor that invade the normal brain, making it difficult for a surgeon to precisely identify which tissue is cancer especially when operating close to the margins that are radiographically defined with magnetic resonance imaging (MRI). This is very important clinically because 80% of recurrences originate from residual cancer remaining after surgery. Moreover, due to the eloquence of brain tissue, the removal of normal brain close to the margins could have dire consequences including impairment of cognitive functions. Standard imaging techniques using MRI and intraoperative visual inspection do not detect the full extent of glioma invasion into normal brain, motivating the need for rapid invasive cancer detection during brain tumor resection. Optical techniques are seeing increased use for disease detection and surgical guidance. In particular, Raman spectroscopy gives molecular information at the tissue surface based on the vibrational modes of molecular species and has been used for cancer detection in a number of organs.

We have previously developed a hand-held spontaneous Raman near-infrared (IR) spectroscopy probe system capable of intraoperative brain cancer detection in patients with grade 2 to 4 gliomas with >90% sensitivity and specificity. During brain cancer resection procedures, the level of ambient light can usually be kept to a minimum in order to minimize its impact on optical signals obtained with instruments operating in the near-IR. However, clinical translation and routine use of surgical guidance optical systems could be facilitated by reducing their sensitivity to the effects of ambient light sources in the operating room (OR) such as OR lights, neuronavigation systems, and light from windows. To address this issue, we have used supervised machine learning with artificial neural networks (ANNs) to accurately distinguish cancer from normal brain tissue in vivo from Raman spectroscopy data. ANN uses a nonparametric model, which is suited for handling the inconsistent effects of light artifacts on the Raman spectra. It is able to adapt to nonlinear changes in optical signal caused by sources of lights in the OR. This improves the robustness of the Raman spectroscopy technique for brain cancer detection, making it easier to integrate into the neurosurgical workflow. Reducing the volume of residual cancer improves patient prognosis and radiographically complete resection can help reduce the recurrence rate and increase patient survival.

2 Methods

Our previous 17-patient dataset was used to assess the ability of ANN to distinguish cancer from normal brain tissue under the presence of light artifacts from OR sources. The Raman spectroscopy system was used at the Montreal Neurological Institute and Hospital on patients undergoing brain tumor resection for grade 2 to 4 gliomas. Informed consent was acquired for all patients, with oversight provided by the Montreal Neurological Institute ethics review board. The Raman spectroscopy system is described previously, consisting of a hand-held fiberoptic probe (EMVision LLC, Florida), a 785-nm excitation laser (Innovative Photonic Solutions, New Hampshire), and a high-resolution charge-coupled device spectroscopic detector.
Seven fibers are used for detection and one central fiber is used for laser excitation. The probe samples tissue with a 0.5-mm surface diameter and a sampling depth up to \( \sim \) 1 mm, with the majority of the signal coming from approximately the first 500 microns depth of tissue.\(^{11}\) The probe was used during surgery to measure 5 to 15 tissue locations per patient, with an emphasis placed on areas of invasive cancer toward the periphery of the tumor. Each acquisition consisted of one background measurement with the laser off, and three averaged measurements with the laser on, for a total acquisition time of 0.2 s. Neuronavigation tracking was performed for each measurement using a StealthStation (Medtronic, Minnesota), to colocate the measurements on preoperative MRI. Biopsy samples were taken superficially at each location for neuropathology assessment based on hematoxylin and eosin staining. Sources of light in the OR include the neurosurgical microscope, which provides a white light source as well as a blue light source typically used for fluorescence detection (OPMI Pentero, Zeiss, Germany), the IR source of the tracking unit for the neuronavigation system (StealthStation, Medtronic, Minnesota), and the overhead surgical lights (Model 380, Dr Mach, Germany), hereafter referred to as OR lights. Note however that for all of the acquired in vivo spectra, the light on the neurosurgical microscope was turned off. MATLAB\(^{6}\) (Mathworks, Inc.) was used for analysis. All spectra, with or without ambient OR light sources, were preprocessed using background subtraction (laser off) to minimize the impact of ambient light contributions, for normalization by laser power, and for autofluorescence removal based on an iterative polynomial fit.\(^{17}\) Tissue classification based on Raman spectra was performed using feedforward ANN, a supervised machine learning technique inspired by the function of biological neurons.\(^{18}\) ANN constructs a network consisting of layers of neurons, where each neuron in one layer is connected to each neuron in the next layer (Fig. 1). The neurons in the first layer are also connected to the input data (Raman spectra), and the neurons in the last layer are connected to the output, in this case determining the tissue type as cancer or normal brain. Connections between neurons have weights, which are learned based on training data using backpropagation. Each neuron also applies a transfer function to the weighted sum of inputs, which is what makes ANN able to model complex nonlinear interactions if the transfer function is nonlinear. Once the ANN model is trained using known data, it can be used to predict classes (tissue types) for test data. The Neural Network Toolbox in MATLAB\(^{6}\) 2016a was used for ANN classification, using the “feedforwardnet” functionality. The ANN was created with 30 total neurons in the hidden layers: 20 neurons in the first hidden layer and 10 neurons in the second hidden layer. The number of neurons and layers in a network control the complexity of the model and optimizing these parameters is essential; too many neurons and/or layers may cause overfitting to the training data, while too few may lack sufficient complexity to distinguish spectrally distributed differences between tissues. More than one layer also allows the model to assess different levels of detail in the data (narrow versus wide peaks for instance). In practice, these parameters can be chosen using optimization functions such as grid search (used for this data), random search, or genetic algorithms. The tangent sigmoidal transfer function was used for both hidden layers and a linear transfer function was used for the output layer. Levenberg–Marquardt backpropagation was used to train the network weights. The choice of transfer function and training method determines network training properties (training speed, robustness to noise, and so on) as well as what the network is able to model. We have chosen default parameters that allow us to model nonlinear effects, but there are a variety of different transfer functions (linear, log-sigmoid, tan-sigmoid, hard-limit, rectified linear unit, and so on) and training methods (conjugate gradient, Levenberg–Marquardt, resilient backpropagation, Broyden–Fletcher–Goldfarb–Shanno quasi-Newton, and so on) which can be used.\(^{19\text{-}23}\) Performance metrics of accuracy, sensitivity, and specificity were determined using leave-one-out cross-validation, where each sample is in turn considered the testing set, with the remaining data being the training set.

Classification was used in distinguishing measurements of cancer tissue from normal brain. Under normal Raman spectroscopy probe operation, measurements were made with overhead OR lights pointing away from the surgical cavity and with the IR source of the neuronavigation system turned away. However, despite the use of background subtraction for all measurements, a subset of measurements was nevertheless affected by those sources of ambient light. This can be attributed to the variation of ambient light signal from different sources and changing conditions in the OR (movement, shadows, modulated light, probe position and orientation, and so on), as well as the impact of ambient light on methods for removing background fluorescence.\(^{24,25}\) The effects of ambient light sources can be automatically detected based on assessment of the acquired spectra, leading to the exclusion of the measurement for surgical decision-making. The spectral effects of these light sources on the background and raw spectra are distinct (see Fig. 2 and Sec. 3 for details), and it is based on these spectral signatures in the raw spectra that we are able to easily identify spectra with...
substantial light artifacts, where the effects of light sources are clearly visible based on these spectral signatures. Classification with ANN was performed both on the original dataset ($n = 161$ measurements), where samples with substantial light artifacts were excluded, as well as on the larger dataset with those samples included ($n = 177$ measurements, 16 of which had substantial light artifacts). Classification was also performed using the boosted trees method, as described previously, for comparison with the performance of ANN. Boosted trees uses an ensemble of decision trees to distinguish between classes, with each tree operating on the residual of the previous tree. In this technique, each decision tree operates on the residual of the previous tree, applying threshold decisions to Raman spectral bands. Ambient light artifacts can produce offsets in intensity and nonlinear distortions in the acquired Raman spectra, due to tissue optics and strong spectral signatures produced by these light sources. Boosted trees uses decision trees, which apply threshold values at particular spectral bands. Because these threshold values are fixed, it is unable to adapt to nonlinear spectral distortions, making it unsuitable for handling the effects of ambient light sources. Basically, boosted trees is unable to encode information about variations in intensities at particular spectra bands. This motivates the use of ANN, which is able to model nonlinear processes due to the use of nonlinear transfer functions, and can also encode variations in intensities since each neuron stores this information (weighted sum of its connections).

The analysis of spectra from in vivo brain tissue in humans provided only a limited number of measurements affected by ambient light artifacts. To insure sufficient statistics, further comparative analysis was performed on spectra from ex vivo calf brain tissue to validate the effectiveness of ANN for tissue classification under the presence of light artifacts. Three calf brains were used with a total of 330 measurements taken in the OR (161 from white matter and 169 from gray matter). The probe was placed in contact with the brain tissue for each measurement. Each acquisition consisted of one background measurement with the laser off, and three averaged measurements with the laser on, for a total acquisition time of 0.2 s. Spectra were measured under different light conditions: no lights; high and low intensity OR lights (maximum and minimum Fig. 2 (a) Hand-held Raman spectroscopy probe. (b) Mean preprocessed Raman spectra acquired on calf brain tissue for white matter (light and dark red) and gray matter (light and dark blue), under no lights (dark red and dark blue) and with high intensity OR lights (light red and light blue). (c and d) Raw spectra acquired on (c) white matter and (d) gray matter in calf brain, with no preprocessing applied, under different light conditions: no lights (black), microscope white light on high intensity (purple), microscope blue light on high intensity (blue), IR source of the neuronavigation system (red), and overhead OR surgical lights on high intensity (turquoise).
Spectra were preprocessed using background subtraction, instrument response correction using measurements of an SRM2241 reference standard (National Institute of Standards and Technology, Maryland), autofluorescence removal, and normalization (standard normal variate). ANN was used to classify tissue as either white or gray matter. The ANN was created with 32 total neurons in the hidden layers: 22 neurons in the first hidden layer and 10 neurons in the second hidden layer. These parameters (number of neurons and number of layers) were determined using grid search parameter optimization, iterating over a range of possible values for each parameter. The tangent sigmoidal transfer function was used for both hidden layer and 10 neurons in the second hidden layer. Classification was performed per-brain, where all of the spectra from one of the brains were considered the testing data and the spectra from the other brains were the training data. This was repeated for each of the three brains.

### 3 Results

The hand-held fiberoptic probe was used to take intraoperative measurements of Raman spectra during brain tumor resection for 17 patients at the Montreal Neurological Institute and Hospital [Fig. 2(a)]. Important spectral differences between samples with cancer tissue versus normal brain are associated with variation in cholesterol and phospholipids (700 and 1142 cm⁻¹), proteins (1005 cm⁻¹), and nucleic acid (1540 to 1645 cm⁻¹). Details about the spectral areas affected by different sources of light such as microscope lights, overhead OR lights, LCD screens, and windows, can be found in our previous work. The spectral bands affected by these other light sources can overlap with important Raman signals, impeding the ability to properly use the available information for tissue classification. In the in vivo dataset considered in this work, however, the surgical microscope lights were turned off for all measurements, and so it has been determined from their spectral signatures that the main potential sources of ambient light during neurosurgery were overhead OR lights and the IR source of the neuronavigation system. OR surgical lights contribute mainly in the visible part of the spectrum with a typical decaying intensity in the near-IR while contributions from the IR source of the neuronavigation system produce a signal across the spectral range, though primarily in the near-IR and IR range. Multivariate analysis was performed to make use of the available spectral data. Table 2 shows the classification performance of boosted trees and ANN for distinguishing tissue types with or without the presence of measurements with light artifacts.

All performance metrics are ≥89% for both methods when excluding light artifacts. When measurements with light artifacts are included, boosted trees sees a substantial performance drop in accuracy, sensitivity, and specificity. ANN continues to perform at almost the same level despite the inclusion of light artifacts, with 19% greater sensitivity and 7% greater specificity when compared with boosted trees. This is because ANN can model nonlinear effects and encodes information about variations in intensities at particular spectral bands, which is essential for modeling the offsets and spectral distortions caused by light sources. Boosted trees does not encode this type of information. ANN uses nonlinear transfer functions applied to the weighted sums of network connections at each neuron, allowing it to model complex nonlinear effects. In contrast, boosted trees employs an ensemble of decision trees, in which each decision tree does not transmit any information to subsequent decision trees other than whether the Raman signal at particular spectral bands is above or below thresholds. The actual intensity value at a spectral band is lost in the boosted trees model, making it extremely difficult to account for the contribution of confounding signals in the spectra.

Further validation was performed on ex vivo calf brain tissue (three calf brains with a total of 330 measurements). Multivariate analysis was used to assess the capabilities of

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Number of measurements of calf brain tissue taken under different light conditions, for gray and white matters.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gray matter</td>
</tr>
<tr>
<td>No lights</td>
<td>47</td>
</tr>
<tr>
<td>OR lights (high)</td>
<td>45</td>
</tr>
<tr>
<td>OR lights (low)</td>
<td>45</td>
</tr>
<tr>
<td>Microscope white light (high)</td>
<td>5</td>
</tr>
<tr>
<td>Microscope white light (medium)</td>
<td>5</td>
</tr>
<tr>
<td>Microscope white light (low)</td>
<td>5</td>
</tr>
<tr>
<td>Microscope blue light (high)</td>
<td>5</td>
</tr>
<tr>
<td>Microscope blue light (medium)</td>
<td>5</td>
</tr>
<tr>
<td>Microscope blue light (low)</td>
<td>5</td>
</tr>
<tr>
<td>IR neuronavigation</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>169</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Classification performance of boosted trees and ANNs for distinguishing cancer from normal brain in datasets where those measurements with light artifacts are excluded and included, respectively.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification method</td>
<td>Dataset</td>
</tr>
<tr>
<td>Boosted trees</td>
<td>Excluding light artifacts</td>
</tr>
<tr>
<td></td>
<td>Including light artifacts</td>
</tr>
<tr>
<td>ANNs</td>
<td>Excluding light artifacts</td>
</tr>
<tr>
<td></td>
<td>Including light artifacts</td>
</tr>
</tbody>
</table>
ANN for distinguishing white matter from gray matter in the presence of light artifacts. To illustrate these differences, Fig. 2(b) shows the mean Raman spectra for white and gray matter, with no lights and white high intensity OR lights. Raw spectra (before preprocessing) are provided in Figs. 2(c) and 2(d) for a number of light conditions. The neuronavigation IR source and the OR lights produce spectrally broad contributions to the acquired spectra, whereas the microscope lights contribute a signal that is spectrally sharper at a number of different band ranges. The OR lights produce a notably larger contribution for gray matter than for white matter, which may be because less light is reflected compared with white matter. For the processed Raman spectra, it is evident from Fig. 2(b) that many of the spectral features and peaks remain despite the presence of light artifacts but are offset or deformed by the polluted signal, particularly for gray matter tissue. This motivates the use of ANN for tissue classification, which is able to adapt to nonlinear distortions. Table 3 shows the classification performance of boosted trees and ANN for distinguishing gray matter from white matter in calf brain with or without the presence of measurements with light artifacts.

ANN continues to perform at almost the same level (97% to 100% accuracy) despite the inclusion of light artifacts, whereas boosted trees sees a substantial drop in accuracy. This is consistent with the in vivo findings, suggesting that ANN is better suited to handle the nonlinear effects of ambient light sources on spectra. The ability to detect cancer tissue in a manner that is robust to the effects of light artifacts is critical for improving the ease of clinical use of Raman spectroscopy in vivo. There are often numerous potential sources of ambient light in an OR that may interfere with the acquisition of spectra during surgical operations. Thus, the use of ANN for tissue classification can help facilitate clinical translation of Raman spectroscopy techniques.

### 4 Discussion

Raman spectroscopy is able to detect invasive brain cancer in glioma patients. We have developed a system using a handheld probe that can be used for rapid intraoperative cancer detection. Raman spectra are acquired to provide molecular information that can be used to distinguish between normal brain and cancer tissue, based on supervised machine learning algorithms. Gliomas typically recur at tumor remnants in the resection cavity, and so the ability to detect this cancer during surgery may improve the completeness of resection, and thus impact patient survival. Approaching radiographically complete resection can yield substantial survival advantages, particularly for low-grade gliomas. Sources of light commonly found in the OR can limit the detection capabilities of Raman spectroscopy if their impact is not properly managed either physically by controlling the intraoperative light environment or through automated exclusion of the measurements during surgery for decision-making. Here, we have demonstrated that ANNs can be used to circumvent these issues by better modeling the nonlinear interactions of extraneous light sources. The nonlinear and adaptive nature of ANNs makes it suitable to modeling these kinds of interactions. Moreover, this technique can be used with existing Raman spectroscopy systems, without the need for complex system hardware implementations such as for wavelength-modulated Raman spectroscopy or filtering techniques. Although we have demonstrated in the past that Raman spectroscopy can be achieved in the presence of ambient light background during neurosurgery, some light sources are necessary for other standard operating procedures, and so making the use of Raman spectroscopy more robust to ambient light helps improve clinical integration into the surgical workflow. Here, we have demonstrated that an ANN classification technique can be used to circumvent those limitations associated with overhead OR light levels and other sources such as IR light from a neuronavigation system. We have also presented preliminary results suggesting that by using ANN, the surgical microscope lights did not prevent the detection of gray versus white matter in ex vivo calf brain tissue. Further studies would be needed to assess the ability to distinguish cancer versus normal brain in vivo under the surgical microscope lights. Although the application presented in this study is for brain tumor resection, this methodology could be applied to any clinical use of Raman spectroscopy where ambient light is a consideration.

### Acknowledgments

Eric Marple (EmVision LLC) assisted with system design and validation. Kelvin Mok assisted with neuronavigation techniques. This work was supported by the Fonds de Recherche du Quebec Nature et Technologies, the Natural Sciences and Engineering Research Council of Canada, the Canadian Institute of Health Research, the Groupe de Recherche en Sciences et Technologies, the Natural Sciences and Engineering Research Council of Canada, and the Banque Nationale. K.P. and F.L., are cofounders of ODS Medical Inc., a medical device company that seeks to commercialize the Raman spectroscopy system for real-time detection of tissue abnormalities.

### Table 3 Classification performance of boosted trees and ANNs for distinguishing gray matter from white matter (in ex vivo calf brain tissue) in datasets where measurements with light artifacts are excluded and included, respectively.

<table>
<thead>
<tr>
<th>Classification method</th>
<th>Dataset</th>
<th>Gray matter accuracy (%)</th>
<th>White matter accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boosted trees</td>
<td>Excluding light artifacts</td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Including light artifacts</td>
<td>89</td>
<td>81</td>
</tr>
<tr>
<td>ANNs</td>
<td>Excluding light artifacts</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Including light artifacts</td>
<td>98</td>
<td>97</td>
</tr>
</tbody>
</table>

References

5. A. Mahadevan-Jansen et al., “Development of a fiber optic probe to
measure NIR Raman spectra of cervical tissue in vivo,” Photochem.
breast surgery using Raman spectroscopy,” Cancer Res. 66(6), 3317–
3322 (2006).
7. M. Boncheva et al., “Depth profiling of stratum corneum hydration in
vivo: a comparison between conductance and confocal Raman spectro-
8. A. Mizuno et al., “Near-infrared Fourier transform Raman spectroscopic
study of human brain tissues and tumours,” J. Raman Spectrosc. 25(1),
9. S. Koljenović et al., “Discriminating vital tumor from necrotic tissue in
human glioblastoma tissue samples by Raman spectroscopy,” Lab.
Invest. 82(10), 1265–1277 (2002).
10. S. N. Kalkanis et al., “Raman spectroscopy to distinguish grey matter,
necrosis, and glioblastoma multiforme in frozen tissue sections,” J.
12. J. Desroches et al., “Characterization of a Raman spectroscopy probe
13. J. S. Smith et al., “Role of extent of resection in the long-term outcome
of low-grade hemispheric gliomas,” J. Clin. Oncol. 26(8), 1338–1345
(2008).
15. N. Sanai and M. S. Berger, “Glioma extent of resection and its impact on
16. M. Lacroix et al., “A multivariate analysis of 416 patients with glioblas-
toma multiforme: prognosis, extent of resection, and survival,” J.
17. J. Zhao et al., “Automated autofluorescence background subtraction
algorithm for biomedical Raman spectroscopy,” Appl. Spectrosc.
61(11), 1225–1232 (2007).
18. L. Fausett, Fundamentals of Neural Networks: Architectures,
Algorithms and Applications, Pearson Education, New York City
21. M. Leshno et al., “Multilayer feedforward networks with a nonpolyno-
mial activation function can approximate any function,” Neural Netw.
22. S. Ayat et al., “A comparison of artificial neural networks learning algo-
rithms in predicting tendency for suicide,” Neural Comput. Appl. 23(5),
23. R. Caraana and A. Niculescu-Mizil, “An empirical comparison of super-
vised learning algorithms,” in Proc. of the 23rd Int. Conf. on Machine
24. S. Dochow et al., “Etaloning, fluorescence and ambient light suppres-
sion by modulated wavelength Raman spectroscopy,” Biomed.
25. J. Zhao et al., “Clinical Raman measurements under special ambient
27. E. Bauer and R. Kohavi, “An empirical comparison of voting classifi-
cation algorithms: bagging, boosting, and variants,” Mach. Learn. 36(1–
28. A. Beljebbar et al., “Ex vivo and in vivo diagnosis of C6 glioblastoma
development by Raman spectroscopy coupled to a microprobe,” Anal.
29. W. Stummer et al., “Extent of resection and survival in glioblastoma
multiforme: identification of and adjustment for bias,” Neurosurgery
30. M. J. McGirt et al., “Extent of surgical resection is independently asso-
ciated with survival in patients with hemispheric infiltrating low-grade
31. M. S. Berger et al., “The effect of extent of resection on recurrence in
patients with low grade cerebral hemisphere gliomas,” Cancer 74(6),
gliomas in adults: a critical review of extent of resection as a factor influ-

Biographies for the authors are not available.