FLIm-Based in Vivo Classification of Residual Cancer in the Surgical Cavity During Transoral Robotic Surgery

Mohamed A. Hassan[1], Brent Weyers[1], Julien Bec[1], Jinyi Qi[1], Dorina Gui[2], Arnaud Bewley[3], Marianne Abouyared[3], Gregory Farwell[4], Andrew Birkeland[3], and Laura Marcu[1](✉)

[1] Department of Biomedical Engineering, University of California, Davis, USA
lmarcu@ucdavis.edu
[2] Department of Pathology and Laboratory Medicine, University of California, Davis, USA
[3] Department of Otolaryngology – Head and Neck Surgery, University of California, Davis, USA
[4] Department of Otorhinolaryngology – Head and Neck Surgery, University of Pennsylvania, Philadelphia, USA

Abstract. Incomplete surgical resection with residual cancer left in the surgical cavity is a potential sequelae of Transoral Robotic Surgery (TORS). To minimize such risk, surgeons rely on intraoperative frozen sections analysis (IFSA) to locate and remove the remaining tumor. This process, may lead to false negatives and is time-consuming. Mesoscopic fluorescence lifetime imaging (FLIm) of tissue fluorophores (i.e., collagen and metabolic co-factors NADH and FAD) emission has demonstrated the potential to demarcate the extent of head and neck cancer in patients undergoing surgical procedures of the oral cavity and the oropharynx. Here, we demonstrate the first label-free FLIm-based classification using a novelty detection model to identify residual cancer in the surgical cavity of the oropharynx. Due to highly imbalanced label representation in the surgical cavity, the model employed solely FLIm data from healthy surgical cavity tissue for training and classified the residual tumors as an anomaly. FLIm data from N = 22 patients undergoing upper aerodigestive oncologic surgery were used to train and validate the classification model using leave-one-patient-out cross-validation. Our approach identified all patients with positive surgical margins (N = 3) confirmed by pathology. Furthermore, the proposed method reported a point-level sensitivity of 0.75 and a specificity of 0.78 across optically interrogated tissue surface for all N = 22 patients. The results indicate that the FLIm-based classification model can identify residual cancer by directly imaging the surgical cavity, potentially enabling intraoperative surgical guidance for TORS.

Keywords: TORS · Positive Surgical Margin · FLIm · Head and Neck Oncology

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43996-4_56.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14228, pp. 587–596, 2023.
https://doi.org/10.1007/978-3-031-43996-4_56

1 Introduction

Residual tumor in the cavity after head and neck cancer (HNC) surgery is a significant concern as it increases the risk of cancer recurrence and can negatively impact the patient's prognosis [1]. HNC comprises the third highest positive surgical margins (PSM) rate across all oncology fields [2]. Achieving clear margins can be challenging in some cases, particularly in tumors with involved deep margins [3, 4].

During transoral robotic surgery (TORS), surgeons may assess the surgical margin via visual inspection, palpation of the excised specimen and intraoperative frozen sections analysis (IFSA) [5]. In the surgical cavity, surgeons visually inspect for residual tumors and use specimen driven or defect-driven frozen section analysis to identify any residual tumor [6, 7]. The latter involves slicing a small portion of the tissue at the edge of the cavity and performing a frozen section analysis. These approaches are error-prone and can result in PSMs and a higher risk of cancer recurrence [7]. In an effort to improve these results, recent studies reported the use of exogenous fluorescent markers [8] and wide-field optical coherence tomography [9] to inspect PSMs in the excised specimen. While promising, each modality presents certain limitations (e.g., time-consuming analysis, administration of a contrast agent, controlled lighting environment), which has limited their clinical adoption [10, 11].

Label-free mesoscopic fluorescence lifetime imaging (FLIm) has been demonstrated as an intraoperative imaging guidance technique with high classification performance (AUC = 0.94) in identifying in vivo tumor margins at the epithelial surface prior to tumor excision [12]. FLIm can generate optical contrast using autofluorescence derived from tissue fluorophores such as collagen, NADH, and FAD. Due to the sensitivity of these fluorophores to their microenvironment, the presence of tumor changes their emission properties (i.e., intensity and lifetime characteristics) relative to healthy tissue, thereby enabling the optical detection of cancer [13].

However, ability of label-free FLIm to identify residual tumors in vivo in the surgical cavity (deep margins) has not been reported. One significant challenge in developing a FLIm-based classifier to detect tumor in the surgical cavity is the presence of highly imbalanced labels.

Surgeons aim to perform an en bloc resection, removing the entire tumor and a margin of healthy tissue around it to ensure complete excision. Therefore, in most cases, only healthy tissue in left in the cavity. To address the technical challenge of highly imbalanced label distribution and the need for intraoperative real-time cavity imaging, we developed an intraoperative FLIm guidance model to identify residual tumors by classifying residual cancer as anomalies. Our proposed approach identified all patients with PSM. In contrast, the IFSA reporting a sensitivity of 0.5 [6, 7].

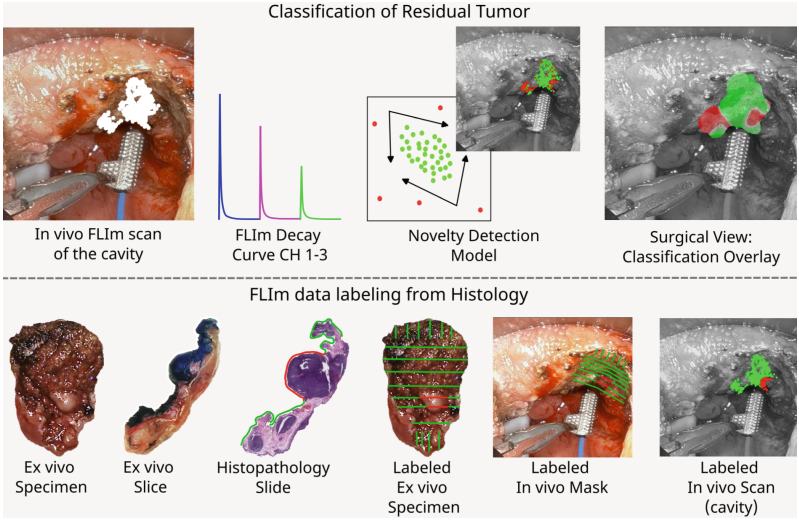

Fig. 1. Overview methodology of the label-free FLIm-based intraoperative surgical guidance, data collection, histopathology registration, and data processing. *Upper panel*: overview of developing the classification method to detect residual HNC in the surgical cavity during the TORS procedure. *Lower panel*: describes the workflow involving generating labels for classifier training and testing. The labels were derived directly from histopathology, evaluated, and annotated by a clinical pathologist (DG). Each annotated H&E section was registered with the ex vivo and in vivo FLIm scan images. The red annotations correspond to cancer labels, and the green annotations correspond to healthy.

2 Method

As illustrated in Fig. 1, the proposed method uses a clinically-compatible FLIm system coupled to the da Vinci SP transoral robotic surgical platform to scan the surgical cavity in vivo and acquire FLIm data. We used the cumulative distribution transform (CDT) of the fluorescence decay curves extracted from the FLIm data as the input feature. The novelty detection model classified FLIm points closer to the healthy distribution as healthy and further from the healthy distribution as a residual tumor. We implemented the image guidance by augmenting the classification map to the surgical view using the predictor output and point locations of the scan.

2.1 FLIm Hardware and Data Acquisition

This study used a multispectral fluorescence lifetime imaging (FLIm) device to acquire data [14]. The FLIm device features a 355 nm UV laser for fluorescence excitation, which is pulsed at a 480 Hz repetition rate. A 365 μm multimode optical fiber (0.22 NA) delivers excitation light to tissue and relays the corresponding fluorescence signal to a set of dichroic mirrors and bandpass filters to spectrally resolve the autofluorescence. Three variable gain UV enhanced Si APD modules with integrated trans-impedance amplifiers receive the autofluorescence, which is spectrally resolved as follows: (1)

390/40 nm attributed to collagen autofluorescence, (2) 470/28 nm to NADH, and (3) 542/50 nm to FAD. The resulting autofluorescence waveform measurements for each channel are averaged four times, thus with a 480 Hz excitation rate, resulting in 120 averaged measurements per second [15].

The FLIm device includes a 440 nm continuous wave laser that serves as an aiming beam; this aiming beam enables real-time visualization of the locations where fluorescence (point measurements) is collected by generating visible blue illumination at the location where data is acquired. Segmentation of the 'aiming beam' allows for FLIm data points to be localized as pixel coordinates within a surgical white light image (see Fig. 1). Localization of these coordinates is essential to link the regions where data is obtained to histopathology, which is used as the ground truth to link FLIm optical data to pathology status [16].

Table 1. Anatomy, Surgical outcome, and Tissue label breakdown for the 22 patients

Surgical Outcome	No. Patients	Tissue Label	No. FLIm Points
Clear margin	19	Healthy	170,535
PSM	3	Cancer	2,451

FLIm data was acquired using the da Vinci SP robotic surgical platform. As part of the approved protocol for this study, the surgeon performed in vivo FLIm scan on the tumor epithelial surface and the surrounding uninvolved benign tissue. Upon completing the scan, the surgeon proceeded with en bloc excision of the tissue suspected of cancer. An ex vivo FLIm scan was then performed on the surgically excised specimen. Finally, the patient's surgical cavity was scanned to check for residual tumor.

2.2 Patient Cohort and FLIm Data Labeling

The research was performed under the approval of the UC Davis Institutional Review Board (IRB) and with the patient's informed consent. All Patients were anesthetized, intubated, and prepared for surgery as part of the standard of care. N = 22 patients are represented in this study, comprising HNC in the palatine tonsil (N = 15) and the base of the tongue (N = 7). For each patient, the operating surgeon conducted an en bloc surgical tumor resection procedure (achieved by TORS-electrocautery instruments), and the resulting excised specimen was sent to a surgical pathology room for grossing. The tissue specimen was serially sectioned to generate tissue slices, which were then formalin-fixed, paraffin-embedded, sectioned, and stained to create Hematoxylin & Eosin (H&E) slides for pathologist interpretation (see Fig. 1).

After the surgical excision of the tumor, an in vivo FLIm scan of approximately 90 s was conducted within the patient's surgical cavity, where the tumor was excised. To validate optical measurements to pathology labels (e.g., benign tissue vs. residual tumor), pathology labels from the excision margins were digitally annotated by a pathologist on each H&E section. The aggregate of H&E sections was correspondingly labeled on the ex vivo specimen at the cut lines where the tissue specimen was serially sectioned.

Thereafter, the labels were spatially registered in vivo within the surgical cavity. This process enables the direct validation of FLIm measurements to the pathology status of the electrocauterized surgical margins (see Table 1).

2.3 FLIm Preprocessing

The raw FLIm waveform contains background noise, instrument artifacts, and other types of interference, which need to be carefully processed and analyzed to extract meaningful information (i.e., the fluorescence signal decay characteristics). To account for background noise, the background signal acquired at the beginning of each clinical case was subtracted from the measured raw FLIm waveform. To retrieve the fluorescence function, we used a non-parametric model based on a Laguerre expansion polynomials and a constrained least-square deconvolution with the instrument impulse response function as previously described [17]. In addition, an SNR threshold of $\geq$50 dB was applied as a filtering criterion to select FLIm points with good signal quality.

2.4 Novelty Detection Model

The state-of-the-art novelty detection models were comprehensively reviewed in the literature [18, 19]. Due to its robust performance, we chose the Generalized One-class Discriminative Subspaces (GODS) classification model [20] to classify healthy FLIm points from the residual tumor. The model trained only on the healthy FLIm points and use a semi-supervised technique to classify residual tumor from healthy. The GODS is a pairwise complimentary classifier defined by two separating hyperplanes to minimize the distance between the two classifiers, limiting the healthy FLIm data within the smallest volume and maximizing the margin between the hyperplanes and the data, thereby avoiding overfitting while improving classification robustness. The first hyperplane (w_1, b_1) projects most of the healthy FLIm points to the positive half of the space, whereas the second hyperplane (w_2, b_2) projects most of the FLIm points in the negative half.

$$\min_{W \in S_d^K, b} F = \frac{1}{2n} \sum_{i=1}^{n} \sum_{j=1}^{2} \| W_j^T x_i + b_j \|_2^2 + \frac{\nu}{2\pi} \sum_i \left[\eta - \min\left(W_1^T x_i + b_1 \right) \right]_+^2 + \left[\eta - \max\left(W_2^T x_i + b_2 \right) \right]_+^2 \quad (1)$$

where W_1, W_2 are the orthonormal frames, $\min_{W \in S_d^K, b}$ is the Stiefel manifold, η is the sensitivity margin, and was set $\eta = 0.4$ for our experiments. ν denote a penalty factor on these soft constraints, and b is the biases. x_i denotes the training set containing CDT of the concatenated FLIm decay curve across channels 1–3 along the time axis. The CDT of the concatenated decay curves is computed as follows: Normalize the decay curves to 0–1. Compute and normalize the cumulative distribution function (CDF). Transforming the normalized CDF into the cumulative distribution transform by taking the inverse cumulative distribution function of the normalized CDF [21].

2.5 Classifier Training and Evaluation

The novelty detection model used for detecting residual cancer is evaluated at the point-measurement level to assess the diagnostic capability of the method over an entire tissue surface. The evaluation followed a leave-one-patient-out cross-validation approach. The study further compared GODS with two other novelty detection models: robust covariance and, one-class support vector machine (OC-SVM) [22]. Novelty detection model solely used healthy labels from the in vivo cavity scan for training. The testing data contained both healthy and residual cancer labels. We used grid search to optimize the hyper-parameters and features used in each model and are tabulated in the supplementary section Table S1. The sensitivity, specificity, and accuracy were used as evaluation metrics to assess the performance of classification models in the context of the study. Results of a binary classification model using SVM are also shown in the supplementary section Table S2.

2.6 Classifier Augmented Display

The classifier augmentation depends on three independent processing steps: aiming beam localization, motion correction, and interpolation of the point measurements. A detailed description of implementing the augmentation process is discussed in [23]. The interpolation consists of fitting a disk to the segmented aiming beam pixel location for each point measurement and applying a color map (e.g., green: healthy and red: cancer) for each point prediction. Individual pixels from overlapping disks are averaged to produce the overall classification map and augmented to the surgical field as a transparent overlay.

3 Results

Table 2 tabulates the classification performance comparison of novelty detection models for classifying residual cancer vs. healthy on in vivo FLIm scans in the cavity. Three novelty detection models were evaluated, and all three models could identify the presence of residual tumors in the cavity for the three patients. However, the extent of the tumor classification over the entire tissue surface varied among the models. The GODS reported the best classification performance with an average sensitivity of 0.75 ± 0.02 (see Fig. 2). The lower standard deviation indicates that the model generalizes well. The OC-SVM and robust covariance reported a high standard deviation, indicating that the performance of the classification model is inconsistent across different patients.

The model's ability to correctly identify negative instances is essential to its reliability. The GODS model reported the highest mean specificity of 0.78 ± 0.14 and the lowest standard deviation. The Robust Covariance model reported the lowest specificity, classifying larger portions of healthy tissue in the cavity as a residual tumor; indicating that the model did not generalize well to the healthy labels. We also observed that changing the hyper-parameter, such as the anomaly factor, biased the model toward a single class indicating overfitting (see supplementary section Fig. S1).

The GODS uses two separating hyperplanes to minimize the distance between the two classifiers by learning a low-dimensional subspace containing FLIm data properties

Table 2. Classification performance comparison of novelty detection models for classifying residual cancer vs. healthy on in vivo FLIm scans in the cavity, mean (sd). B**old font:** Best-performing model.

Novelty Detection Model	Sensitivity (N = 3)	Specificity (N = 22)	Accuracy (N = 3)
OC-SVM	0.68 (0.13)	0.72 (0.18)	0.58 (0.14)
Robust Covariance	0.67 (0.15)	0.63 (0.17)	0.49 (0.11)
GODS	**0.75 (0.02)**	**0.78 (0.14)**	**0.76 (0.02)**

of healthy labels. Residual tumor labels are detected by calculating the distance between the projected data points and the learned subspace. Points that are far from the subspace are classified as residual tumors. We observed that the GODS with the FLIm decay curves in the CDT space achieve the best classification performance compared to other novelty detection models with a mean accuracy of 0.76 ± 0.02. This is mainly due to the robustness of the model, the ability to handle high-dimensional data, and the contrast in the FLIm decay curves. The contrast in the FLIm decay curves was further improved in the CDT space by transforming the FLIm decay curves to a normalized scale and improving linear separability.

4 Discussion

Curent study demonstrates that label-free FLIm parameters-based classification model, using a novelty detection aproach, enables identification of residual tumors in the surgical cavity. The proposed model can resolve residual tumor at the point-measurement level over a tissue surface. The model reported low point-level false negatives and positives. Moreover, the current approach correctly identified all patients with PSMs (see Fig. 2). This enhances surgical precision for TORS procedures otherwise limited to visual inspection of the cavity, palpation of the excised specimen, and IFSA. The FLIm-based classification model could help guide the surgical team in real-time, providing information on the location and extent of cancerous tissue.

In context to the standard of care, the proposed residual tumor detection model exhibits high patient-level sensitivity (sensitivity = 1) in detecting patients with PSMs. In contrast, defect-driven IFSA reports a patient-level sensitivity of 0.5 [6, 7]. Our approach exhibits a low patient-level specificity compared to IFSA. Surgeons aim to achieve negative margins, meaning the absence of cancer cells at the edges of the tissue removed during surgery. The finding of positive margins from final histology would result in additional surgical resection, potentially impacting the quality of life. Combining the proposed approach and IFSA could lead to an image-guided frozen section analysis to help surgeons achieve negative margins in a more precise manner. Therefore, completely resecting cancerous tissue and improving patient outcomes.

The false positive predictions from the classification model presented two trends: false positives in an isolated region and false positives spreading across a larger region. Isolated false positives are often caused by the noise of the FLIm system and are

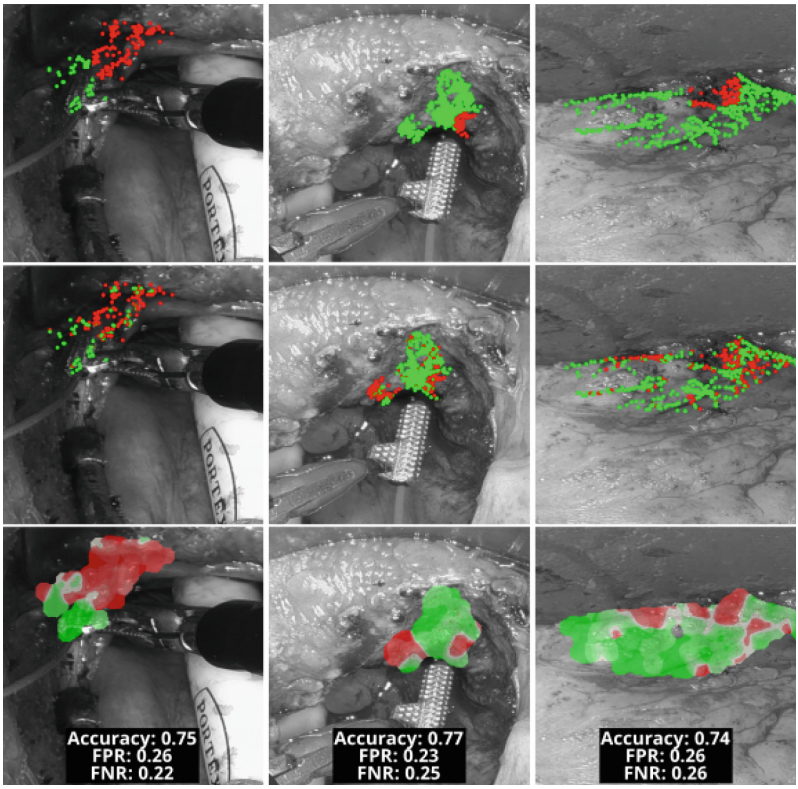

Fig. 2. GODS classification overlay of in vivo cavity scans of three patients presenting with residual tumor. The columns represent each patient, and the rows depict the ground truth labels, the point-prediction overlay, and the augmented surgical view. FPR-False Positive Rate, FNR-False Negative Rate.

accounted for by the interpolation approach used for the classifier augmentation (refer to supplementary section Fig. S2). On the other hand, false positives spreading across a larger region are much more complex to interpret. One insight is that the electrocautery effects on the tissues in the cavity may have influenced them [24]. According to Jackson's burn wound model, the thermal effects caused by electrocautery vary with the different burnt zones. We observed a correlation between a larger spread of false positive predictions associated with a zone of coagulation to a zone of hyperemia.

The novelty detection model generalizes to the healthy labels and considers data falling off the healthy distribution as residual cancer. The FLIm properties associated with the healthy labels in the cavity are heterogeneous due to the electrocautery effects. Electrocautery effects are mainly thermal and can be observed by the levels of charring in the tissue. Refining the training labels based on the levels of charring could lead to a more homogeneous representation of the training set and result in an improved classification model with better generalization.

5 Conclusion

This study demonstrates a novel FLIm-based classification method to identify residual cancer in the surgical cavity of the oropharynx. The preliminary results underscore the significance of the proposed method in detecting PSMs. The model will be validated on a larger patient cohort in future work and address the limitations of the point-level false positive and negative predictions. This work may enhance surgical precision for TORS procedures as an adjunctive technique in combination with IFSA.

Acknowledgment. This work was supported by the National Institutes of Health under Grant 2R01CA187427 in collaboration with Intuitive Surgical, Inc; and P41-EB032840-01. Authors would like to acknowledge Dr. Jonathan Sorger (Intuitive Surgical, Sunnyvale CA) for his support for our ongoing industry collaboration; key areas of his industry support include FLIm visualization aspects and integration of FLIm fiber optic probes into the da Vinci SP TORS platform.

References

1. Gorphe, P., Simon, C.: A systematic review and meta-analysis of margins in transoral surgery for oropharyngeal carcinoma. Oral Oncol. **98**, 69–77 (2019)
2. Orosco, R.K., et al.: Positive surgical margins in the 10 most common solid cancers. Sci. Rep. **8**(1), 1–9 (2018)
3. Li, M.M., Puram, S.V., Silverman, D.A., Old, M.O., Rocco, J.W., Kang, S.Y.: Margin analysis in head and neck cancer: state of the art and future directions. Ann. Surg. Oncol. **26**(12), 4070–4080 (2019)
4. Williams, M.D.: Determining adequate margins in head and neck cancers: practice and continued challenges. Curr. Oncol. Rep. **18**(9), 1–7 (2016). https://doi.org/10.1007/s11912-016-0540-y
5. Poupore, N.S., Chen, T., Nguyen, S.A., Nathan, C.-A.O., Newman, J.G.: Transoral robotic surgery for oropharyngeal squamous cell carcinoma of the tonsil versus base of tongue: a systematic review and meta-analysis. Cancers (Basel) **14**(15), 3837 (2022)
6. Nentwig, K., Unterhuber, T., Wolff, K.-D., Ritschl, L.M., Nieberler, M.: The impact of intraoperative frozen section analysis on final resection margin status, recurrence, and patient outcome with oral squamous cell carcinoma. Clin. Oral Investig. **25**, 6769–6777 (2021)
7. Horwich, P., et al.: Specimen oriented intraoperative margin assessment in oral cavity and oropharyngeal squamous cell carcinoma. J. Otolaryngol. - Head Neck Surg. **50**(1), 1–12 (2021)
8. van Keulen, S., et al.: Rapid, non-invasive fluorescence margin assessment: optical specimen mapping in oral squamous cell carcinoma. Oral Oncol. **88**, 58–65 (2019)
9. Badhey, A.K., et al.: Intraoperative use of wide-field optical coherence tomography to evaluate tissue microstructure in the oral cavity and oropharynx. JAMA Otolaryngol. Head Neck Surg. **149**(1), 71–78 (2023)
10. Zhang, R.R., et al.: Beyond the margins: Real-time detection of cancer using targeted fluorophores. Nat. Rev. Clin. Oncol. **14**(6), 347–364 (2017)
11. Wu, C., Gleysteen, J., Teraphongphom, N.T., Li, Y., Rosenthal, E.: In-vivo optical imaging in head and neck oncology: Basic principles, clinical applications and future directions review-Article. Int. J. Oral Sci. **10**(2), 10 (2018)

12. Hassan, M.A., et al.: Anatomy-specific classification model using label-free FLIm to aid intraoperative surgical guidance of head and neck cancer. IEEE Trans. Biomed. Eng. 1–11 (2023)
13. Marcu, L., French, P.M.W., Elson, D.S.: Fluorescence Lifetime Spectroscopy and Imaging : Principles and Applications in Biomedical Diagnostics. CRC Press, Boca Raton (2014)
14. Gorpas, D., et al.: Autofluorescence lifetime augmented reality as a means for real-time robotic surgery guidance in human patients. Sci. Rep. **9**(1), 1187 (2019)
15. Zhou, X., Bec, J., Yankelevich, D., Marcu, L.: Multispectral fluorescence lifetime imaging device with a silicon avalanche photodetector. Opt. Express **29**(13), 20105 (2021)
16. Weyers, B.W., et al.: Fluorescence lifetime imaging for intraoperative cancer delineation in transoral robotic surgery. Transl. Biophotonics **1**(1–2), e201900017 (2019)
17. Liu, J., Sun, Y., Qi, J., Marcu, L.: A novel method for fast and robust estimation of fluorescence decay dynamics using constrained least-squares deconvolution with Laguerre expansion. Phys. Med. Biol. **57**(4), 843–865 (2012)
18. Perera, P., Oza, P., Patel, V.M.: One-class classification: a survey (2021)
19. Seliya, N., Abdollah Zadeh, A., Khoshgoftaar, T.M.: A literature review on one-class classification and its potential applications in big data. J. Big Data **8**(1), 1–31 (2021)
20. Cherian, A., Wang, J.: Generalized one-class learning using pairs of complementary classifiers. IEEE Trans. Pattern Anal. Mach. Intell. **44**, 6993–7009 (2022)
21. Rubaiyat, A.H.M., Hallam, K.M., Nichols, J.M., Hutchinson, M.N., Li, S., Rohde, G.K.: Parametric signal estimation using the cumulative distribution transform. IEEE Trans. Signal Process. **68**, 3312–3324 (2020)
22. Pedregosa, F., et al.: Scikit-learn: machine learning in python (2011)
23. Gorpas, D., Ma, D., Bec, J., Yankelevich, D.R., Marcu, L.: Real-time visualization of tissue surface biochemical features derived from fluorescence lifetime measurements. IEEE Trans. Med. Imaging **35**(8), 1802–1811 (2016)
24. Lagarto, J.L., et al.: Electrocautery effects on fluorescence lifetime measurements: an in vivo study in the oral cavity. J. Photochem. Photobiol. B Biol. **185**, 90–99 (2018)