PROCEEDINGS OF SPIE

Automated tumor assessment of squamous cell carcinoma on tongue cancer patients with hyperspectral imaging

Francesca Manni, Fons van der Sommen, Sveta Zinger, Esther Kho, Susan Brouwer de Koning, et al.

SPIE.

Automated tumor assessment of squamous cell carcinoma on tongue cancer patients with hyperspectral imaging

Francesca Manni[a], Fons van der Sommen[a], **Sveta** Zinger[a], Esther Kho[b], Susan Brouwer de Koning[b], Theo Ruers[b], Caifeng Shan[c], Jean Schleipen[c], and Peter H.N. de With[a]

[a]Eindhoven University of Technology, P.O. Box 513, Eindhoven, The Netherlands
[b]Netherlands Cancer Institute, P.O. Box 90203, Amsterdam, The Netherlands
[c]Philips Research, P.O. Box 80071, Eindhoven, The Netherlands

ABSTRACT

Head and neck cancer (HNC) includes cancers in the oral/nasal cavity, pharynx, larynx, etc., and it is the sixth most common cancer worldwide. The principal treatment is surgical removal where a complete tumor resection is crucial to reduce the recurrence and mortality rate. Intraoperative tumor imaging enables surgeons to objectively visualize the malignant lesion to maximize the tumor removal with healthy safe margins. Hyperspectral imaging (HSI) is an emerging imaging modality for cancer detection, which can augment surgical tumor inspection, currently limited to subjective visual inspection. In this paper, we aim to investigate HSI for automated cancer detection during image-guided surgery, because it can provide quantitative information about light interaction with biological tissues and exploit the potential for malignant tissue discrimination. The proposed solution forms a novel framework for automated tongue-cancer detection, explicitly exploiting HSI, which particularly uses the spectral variations in specific bands describing the cancerous tissue properties. The method follows a machine-learning based classification, employing linear support vector machine (SVM), and offers a superior sensitivity and a significant decrease in computation time. The model evaluation is on 7 *ex-vivo* specimens of squamous cell carcinoma of the tongue, with known histology. The HSI combined with the proposed classification reaches a sensitivity of 94%, specificity of 68% and area under the curve (AUC) of 92%. This feasibility study paves the way for introducing HSI as a non-invasive imaging aid for cancer detection and increase of the effectiveness of surgical oncology.

Keywords: hyperspectral imaging, tongue cancer, intraoperative tumor detection, support vector machine, image-guided surgery, image classification, cancer detection

1. INTRODUCTION

Head and neck cancer (HNC) is the sixth common cancer by incidence worldwide. About 90% of all head and neck cancers are Squamous Cell Carcinomas (HNSCC).[1] In recent years, besides the elderly people with HNC, an additional growing group of young patients (<45 years) with HNSCC have been reported worldwide,[2] thereby representing a relevant public health problem. The standard treatment, correlated to a good prognosis, is surgical resection. However, 40% of the resected tumors show positive margins, and they are part of a total of 85% of inadequate resections, which includes close positive margins.[3] Positive tumor margins are associated with disease recurrence and a higher rate of second surgical operations. Currently, the surgeon has to rely on palpation and visual inspection for malignant margin assessment, followed by histopathology of the surgical specimen.[4] This method is based on subjective evaluation, which suffers from inter-observer variation. Moreover, histopathology is time-consuming, exposing the patient to a significant risk when residual tumor tissue is remaining after the initial surgery. Hyperspectral imaging holds the potential for non-invasive surgical guidance[5] and may be a helpful imaging tool for the objective assessment of tumor tissue. Recent studies show that hyperspectral imaging (HSI) can be utilized for automated tongue-tumor detection with machine learning-based approaches applied on mice models[6] (sensitivity of 93% and specificity of 85%),[7] as well as on 10 oral cavity squamous cell carcinoma in real patients, where only three tongue-tumor locations are included.[8]

Further author information: (Send correspondence to F.M.)
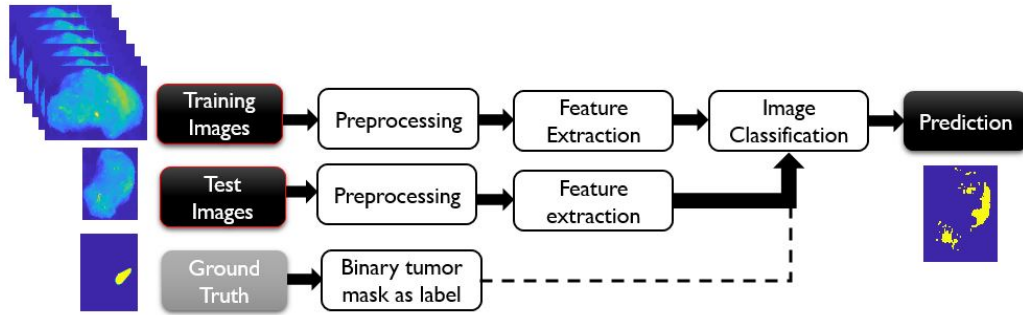E-mail: f.manni@tue.nl

Figure 1. Applied method for training and classification in the system

In this paper, we develop a novel framework for tumor detection on *ex-vivo* surgical specimens of tongue cancer patients with hyperspectral imaging, aiming at an objective disease assessment. We propose two specific aspects: (1) a manual selection of the most informative image bands in a spectral range based on increased discriminative properties in the reflectance spectrum of the tumor and healthy tissue, (2) a block-based feature extraction approach for automated classification purposes. Besides the proposed scheme, we compare this work with an alternative and existing framework that was already applied on mouse cancer models,[6] and evaluated in our study on real patients.

2. METHODS

We apply the general computer vision framework entailing training and testing, as depicted in Figure 1. First, the raw hypercube is preprocessed in order to perform data normalization and noise reduction, then the wavelength range is chosen to extract the best features for malignant and healthy tissue discrimination. The pathology images are outlined by an experienced pathologist and then used as labels for validating the cancer detection with hyperspectral imaging.

2.1 Hyperspectral image acquisition

The hyperspectral images are acquired with a hyperspectral line-scan camera prototype (Philips Research, Eindhoven, The Netherlands). The line-scan camera has a CMOSIS CMV2000 image sensor (IMEC, Leuven, Belgium) with a wavelength range of 430-920 nm, which produces a hypercube of 192 image bands with 2048 pixels per line (y-direction) and a variable number of lines (x-direction) per sample. Seven *ex-vivo* tongue squamous cell carcinoma are gathered at the Netherlands Cancer Institute (NKI, Amsterdam, The Netherlands) and included in this study. After surgical resection, each specimen is cut straight through the tumor, scanned with the hyperspectral camera and sent for histological processing. The tissue is sectioned, stained in hematoxylin and eosin (H&E) and then validated by an experienced pathologist for cancer and healthy tissue contouring. This procedure is the gold standard in this study. All ethical guidelines for *ex-vivo* human studies are followed.

2.2 Hyperspectral data preprocessing

Preprocessing is an essential step for raw data normalization and correction for the instrumentation noise, which can significantly affect the classification accuracy. First, the raw image data and white reference images are normalized, in order to correct for the dark current influence and illumination intensity differences. The data normalization is based on Lu *et al.*[9] and specified by

$$I_{ref} = \frac{I_{raw} - I_{dark}}{I_{white} - I_{dark}} ,\tag{1}$$

where I_{ref} is the normalized reflectance value, I_{raw} the diffuse reflectance value at the given pixel, I_{white} the intensity value for the white reference, acquired by positioning a white reference plate in the field of view, and I_{dark} the dark reference value, acquired by keeping the camera shutter closed. Within the spectral ranges

(a) Pathology annotation   (b) Registration results   (c) Ground truth extraction
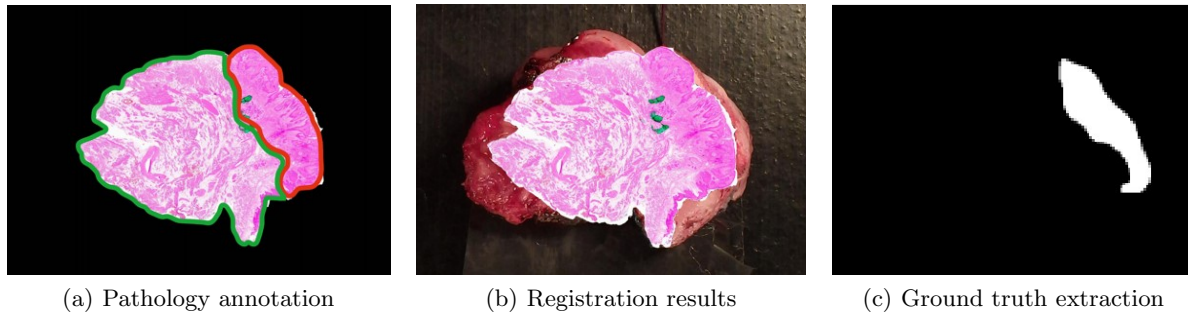
Figure 2. Example of RGB and pathology registration result and ground truth extraction.

430–480 nm and 580–630 nm, the spectrum is strongly affected by the sensor noise, which explains why the image bands within these spectral ranges are discarded for further analysis, resulting in 152 image bands per acquisition. For each image band, a median filter of 6×6 pixels and a two-dimensional (2D) Gaussian smoothing filter are applied to reduce the noise in the spectral signatures.

2.3 Feature extraction

Collecting the spectral signature for each pixel in a 2D array, a three-dimensional (3D) dataset is generated and considered as a hypercube. The pathological information is gathered in the diffuse reflectance of each pixel in the hyperspectral image, for a range of 430–920 nm, which is the quantitative result of light interaction with the biological tissue. To extract the most discriminative information, it is crucial to deal with the high-dimensional data issue. This involves dealing with the large amount of spectral bands and at the same time having the high redundancy within spectral bands that may potentially contain detailed diagnostic information, whereas the redundancy can decrease the classification accuracy.[5] For this reason, 20 image bands showing the highest spectral variation in the reflectance values between the tumor and healthy tissue, are manually selected for classification purposes. This high spectral variation of the most informative wavelengths appears to be consistent among the patients and allows to identify cancerous tissue discrimination. The hypercubes are divided in a grid of 6×6-pixel blocks and within each block, the pixels are averaged to obtain the average spectra. The averaged reflectance value of each block is used for supervised learning. The block-based approach is a powerful solution for increasing the robustness to the spectral noise, while reducing the redundancy between the neighboring pixels and decreasing the computation load.[7] For each patient, the ground truth is extracted from the corresponding annotated H&E section, outlined by an experienced pathologist, and registered with the RGB image of the specimen by selecting control points and using a non-rigid registration algorithm. Figure 2 shows an example of (a) the labeled pathology slide, (b) ground truth registered with the RGB image and (c) the segmented ground-truth mask, used for block labeling. We manually obtain a mask for the region of interest of the tumor area, discarding the image borders affected by histological colorations and motion artifacts due to the line scanning. As an alternative to the manual selection, we apply the implemented method of Chung et al.[6] This method employs Principal Component Analysis (PCA)[10] for automated band extraction, iterative clustering and superpixel segmentation of the first components, using averaged pixel values within each superpixel in order to build discriminative feature vectors. The superpixel algorithm is based on the Simple Linear Iterative Clustering (SLIC) technique,[11] which offers high segmentation performance. Furthermore, the superpixel segmentation has the advantage of pixel clustering in meaningful regions based on statistical decorrelation of spectral image bands.[6]

2.4 Classification of hyperspectral images & validation

As a simple first approach for this feasibility study, linear Support Vector Machine (SVM) is used for supervised learning, to build predictive models for tumor and healthy tissue discrimination from hyperspectral acquisitions. For each patient, one hypercube and one pathology image are collected and employed for classification. Leave-one-patient-out cross-validation is applied, to avoid double patient usage and prevent overfitting. After classification, each block is labeled as tumor or normal tissue. As a result, we obtain binary classified images for each patient.
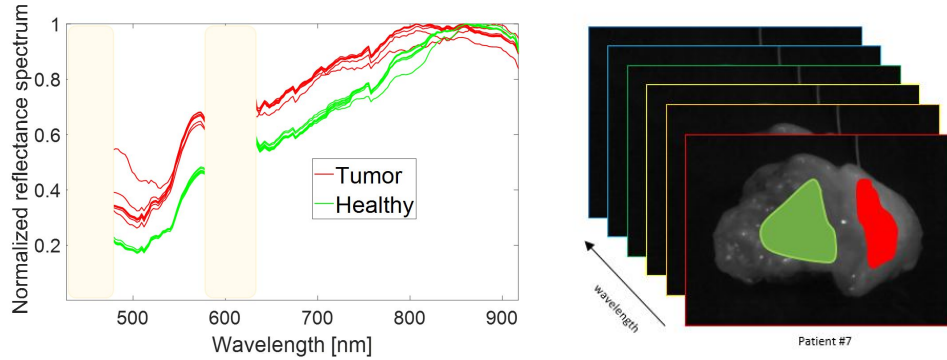
Figure 3. Averaged reflectance spectra for healthy and tumor ROI in 7 tongue-tumor patients.

We compare this result with the pathology ground truth, used as labels for validating the cancer detection with hyperspectral imaging. To evaluate classifier performance, we compute the receiver operating characteristics (ROC) curve, area under the curve (AUC), sensitivity (Se) and specificity (Sp). The last two metrics measure the tumor pixels correctly classified as tumor and as healthy, respectively, as defined in the following equation (TN: true negative, TP: true positive, FP: false positive, FN: false negative). For benchmarking, the same performance metrics are involved during the validation process, and specified by

$$Se = \frac{TP}{TP + FN}, \quad \text{and} \quad Sp = \frac{TN}{TN + FP}. \tag{2}$$

3. EXPERIMENTAL RESULTS

The proposed approach has been trained on 7 patients (20 images per patient at different wavelengths), having cancerous and non-cancerous tissue. During the training phase, for each patient, a region of interest (ROI) is used as ground truth, containing the malignant tissue assessed by an experienced pathologist. We present the qualitative results in Sections 3.1 and evaluate the performance metrics in Section 3.2.

3.1 Cancer detection result evaluation

Figure 3 visualizes the averaged spectra of healthy and tumorous region of interests among the patients. The tumor tissue shows higher reflectance values for the wavelength range of 500–800 nm, where a difference between healthy and tumor spectral fingerprints is revealed. Within this wavelength range, we extract the most significant images for automatic discrimination of tumor and healthy tissue in our classification method. The classification results are visualized in Figure 4. Our method shows less false negatives (FN) compared to the benchmark and, as described in the next paragraph, has the advantage to speed up the computational time. The malignant tissue is clearly detected in both cases. We notice an overestimation of the detected tumor area, that can be due to the number of superpixels or pixel blocks containing both tissues, classified as tumor, as described in ref.[6] As already mentioned in the reference publication, the problem can be overcome by increasing the number of superpixels (which is 1000 in our algorithm) or pixel blocks.

3.2 Benchmarking and performance evaluation

Figure 5 depicts the ROC curves of the detection rates. The AUC values, sensitivity (Se) and specificity (Se) are shown in Table 1. We benchmark our proposal with an existing framework, which represents a reasonable alternative for classifying our hyperspectral dataset, using the same classifier (SVM), and employing a different approach for dimensionality reduction and feature extraction. Table 1 shows that a sensitivity of 94%, specificity of 68% and AUC of 92%, are reached with our proposed method. The balance between specificity and sensitivity can be improved computing a grid search for hyperparameter optimization. However, in this feasibility study, our goal is to aim to a proof-of-concept cancer detection algorithm where the default parameters for the SVM classification model are chosen. By increasing the block size to 10×10 pixels, we reach a higher sensitivity and
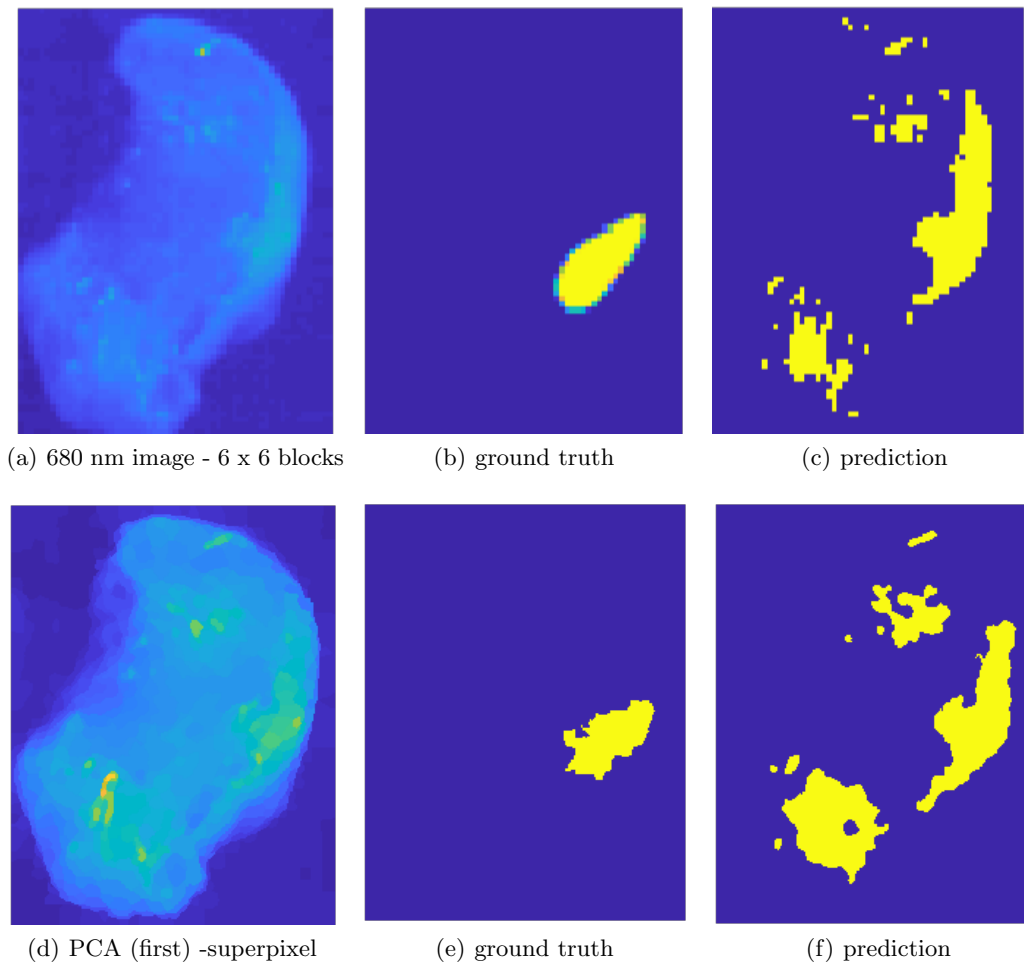
Figure 4. Visualization of the cancer segmentation for Patient 6. (a) 680 nm image band applying 6 x 6 block grid and averaging pixel values within each block, (b) 6 x 6 block grid divided ground truth, (c) classification result -6 x 6 block-based approach-, (d) first principal component from PCA applying superpixel segmentation and averaging pixel values within each superpixel, (e) superpixel ground truth, (f) classification result -PCA-superpixel approach-

specificity compared to the 6×6 pixel block size, while speeding up the averaged training time from 26 minutes and testing time to 2.5 seconds, and accepting the compromise of having a lower image resolution. A significant improvement is achieved in reducing the computation time, when SVM is used for classification, which is essential for intraoperative imaging. Training and testing time are 8 hours and 3 minutes ($\approx$ 2 M data samples), respectively, when the method of Chung *et al.* is applied (E5-1650v4, 3.60 GHz), versus 6 hours and 1 minute ($\approx$ 900.000 data samples) when the 6×6-pixel block-based processing is adopted. Lastly, Lu *et al.*[7] apply the block-based approach for squamous displasia detection on 34 mice, reaching an AUC of 86%, sensitivity of 79%, and specificity of 79%, which are outperformed by our system on human data when employing linear SVM.

4. CONCLUSIONS

Intraoperative tumor assessment facilitates residual cancer removal, which has a significant impact on post-operative prognosis. Currently, tumor-margin assessment depends on the subjective surgeon's capability to visually identify malignant tissue. In this study, we explore the potential of non-invasive and automated cancer discrimination with HSI, extracting spectral features per block and applying a machine learning framework, for squamous carcinoma detection on 7 *ex-vivo* tongue-tumor specimens. The results show that the proposed approach has the ability to find a good classification solution for tongue-tumor detection in real patients with an
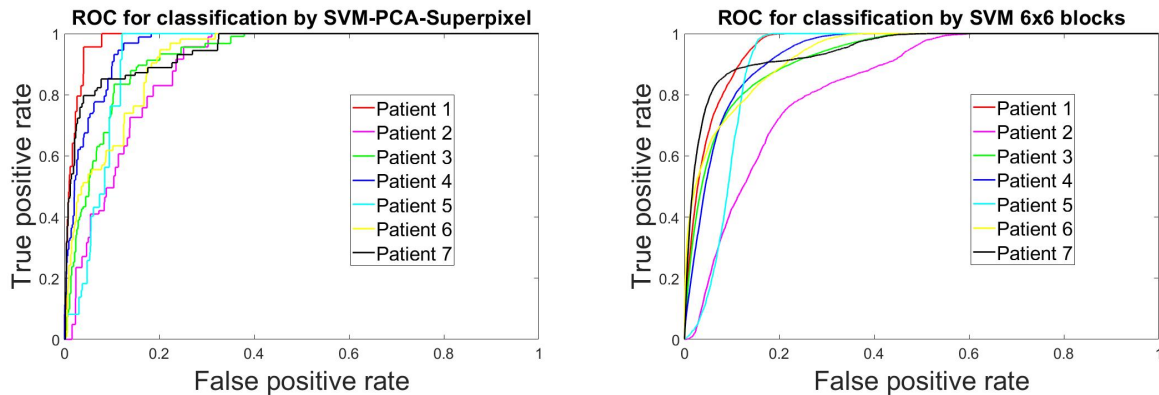
Figure 5. ROC curves for 7 patients with SVM classification.

Table 1. Performance comparison between block- and superpixel-based approach for 7 *ex-vivo* patients with HNSCC.

	Method											
	Proposed method-6x6 blocks				Proposed method-10x10 blocks				Chung *et al.*			
	AUC	Sens.	Spec.	train(min)/test(s)	AUC	Sens.	Spec.	train(min)/test(s)	AUC	Sens.	Spec.	train(min)/test(s)
Patient 1	0.96	1.00	0.67	262/89	0.96	1.00	0.76	55/2.4	0.97	1.00	0.78	360/327
Patient 2	0.83	0.83	0.70	355/96	0.84	0.76	0.14	42/1.5	0.80	0.74	0.64	456/190
Patient 3	0.92	0.97	0.66	410/142	0.93	0.97	0.67	22/6.0	0.92	0.96	0.66	232/347
Patient 4	0.93	1.00	0.58	410/63	0.94	1.00	0.66	55/2.4	0.98	1.00	0.76	520/243
Patient 5	0.91	1.00	0.70	431/55	0.90	1.00	0.79	42/2.0	0.90	1.00	0.73	532/163
Patient 6	0.93	0.75	0.88	326/39	0.94	1.00	0.64	42/2.0	0.93	0.74	0.90	506/138
Patient 7	0.95	0.99	0.59	331/36	0.95	0.96	0.71	44/1.5	0.96	0.93	0.77	548/157
mean	**0.92**	**0.94**	**0.68**	**333/74**	**0.92**	**0.95**	**0.71**	**43/2.5**	**0.92**	**0.91**	**0.74**	**456/224**

AUC of 92%. This proposed solution forms a novel framework for automated tongue-cancer detection, explicitly exploiting HSI, which particularly used the spectral variations in specific bands describing the cancerous tissue properties. The presented method offers a superior sensitivity and a significant decrease in computation time, when compared with an existing approach.[6] Furthermore, our validation was more relevant for patients than in the original publication, where the benchmark approach was tested on 11 cancer mouse models, whereas we perform our validation on 7 real tongue-cancer patients. To the best of our knowledge, the tongue-tumor location is not completely explored in *ex-vivo* studies with machine learning techniques combined with HSI, although it is the most aggressive of all oral squamous carcinoma (OSCC) with higher rate of occult lymph node metastases.[1] Instead, we fully focus on the tongue-tumor location, achieving the feasibility of objective tumor tissue discrimination on real patients. We conclude that the proposed method outperforms the benchmark in speed by a factor 9, when applied on real patients, which is straightforward for real-time intraoperative tumor detection. This feasibility study will serve as a reference for future work when more patient data will be gathered, aiming at higher accuracy and visual aid tool. We conclude that automated tumor assessment with HSI can be successfully employed to aid the surgeon with objective real-time tumor evaluation in tongue-cancer patients.

5. FUTURE RESEARCH

Recent studies show positive results when convolutional neural networks (CNNs) are used for *ex-vivo* surgical tissue classification.[12] Our hyperspectral data collection is an ongoing prospective study, aiming at building a robust dataset. With the current progress of CNNs and their further application in this project, we have the opinion that we can exploit the potential of HSI as a real-time, non-invasive intraoperative imaging tool for classification of cancer tissue, which fits in a straightforward way in the surgical scenario.

ACKNOWLEDGMENTS

REFERENCES

[1] Vigneswaran, N. and Williams, M. D., "Epidemiologic trends in head and neck cancer and aids in diagnosis," *Oral and Maxillofacial Surgery Clinics* **26**(2), 123–141 (2014).

[2] Hussein, A. A., Helder, M. N., de Visscher, J. G., Leemans, C. R., Braakhuis, B. J., de Vet, H. C., and Forouzanfar, T., "Global incidence of oral and oropharynx cancer in patients younger than 45 years versus older patients: A systematic review," *European Journal of Cancer* **82**, 115–127 (2017).

[3] Smits, R. W., Koljenović, S., and Hardillo, J., *et al.*, "Resection margins in oral cancer surgery: room for improvement," *Head & neck* **38**(S1), E2197–E2203 (2016).

[4] Lu, G., Qin, X., Wang, D., Chen, Z. G., and Fei, B., "Quantitative wavelength analysis and image classification for intraoperative cancer diagnosis with hyperspectral imaging," in [*Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling*], **9415**, 94151B, International Society for Optics and Photonics (2015).

[5] Lu, G. and Fei, B., "Medical hyperspectral imaging: a review," *Journal of biomedical optics* **19**(1), 010901 (2014).

[6] Chung, H., Lu, G., Tian, Z., Wang, D., Chen, Z. G., and Fei, B., "Superpixel-based spectral classification for the detection of head and neck cancer with hyperspectral imaging," in [*Medical Imaging 2016: Biomedical Applications in Molecular, Structural, and Functional Imaging*], **9788**, 978813, International Society for Optics and Photonics (2016).

[7] Lu, G., Wang, D., Qin, X., Muller, S., Wang, X., Chen, A. Y., Chen, Z. G., and Fei, B., "Detection and delineation of squamous neoplasia with hyperspectral imaging in a mouse model of tongue carcinogenesis," *Journal of biophotonics* **11**(3), e201700078 (2018).

[8] Fei, B., Lu, G., Wang, X., Zhang, H., Little, J. V., Patel, M. R., Griffith, C. C., El-Diery, M. W., and Chen, A. Y., "Label-free reflectance hyperspectral imaging for tumor margin assessment: a pilot study on surgical specimens of cancer patients," *Journal of biomedical optics* **22**(8), 086009 (2017).

[9] Lu, G., Qin, X., Wang, D., Muller, S., Zhang, H., Chen, A., Chen, Z. G., and Fei, B., "Hyperspectral imaging of neoplastic progression in a mouse model of oral carcinogenesis," in [*Medical Imaging 2016: Biomedical Applications in Molecular, Structural, and Functional Imaging*], **9788**, 978812, International Society for Optics and Photonics (2016).

[10] Rodarmel, C. and Shan, J., "Principal component analysis for hyperspectral image classification," *Surveying and Land Information Science* **62**(2), 115–122 (2002).

[11] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S., and others, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE transactions on pattern analysis and machine intelligence* **34**(11), 2274–2282 (2012).

[12] Halicek, M., Little, J. V., Wang, X., Patel, M., Griffith, C. C., El-Diery, M. W., Chen, A. Y., and Fei, B., "Optical biopsy of head and neck cancer using hyperspectral imaging and convolutional neural networks," in [*Optical Imaging, Therapeutics, and Advanced Technology in Head and Neck Surgery and Otolaryngology 2018*], **10469**, 104690X, International Society for Optics and Photonics (2018).