Classification of hyperspectral endocrine tissue images using support vector machines

Marianne Maktabi | Hannes Köhler | Magarita Ivanova | Thomas Neumuth | Nada Rayes | Lena Seidemann | Robert Sucher | Boris Jansen-Winkeln | Ines Gockel | Manuel Barberio | Claire Chalopin

1Innovation Center Computer Assisted Surgery (ICCAS), University of Leipzig, Leipzig, Germany
2Department of Visceral, Transplant, Thoracic and Vascular Surgery, University Hospital of Leipzig, Leipzig, Germany
3Institute of Image-Guided Surgery (IHU), Strasbourg, France

Correspondence
Marianne Maktabi, Innovation Center Computer Assisted Surgery (ICCAS), University of Leipzig, Semmelweisstraße, Leipzig 04103, Germany.
Email: marianne.maktabi@medizin.uni-leipzig.de

Funding information
Federal Ministry of Education and Research, Grant/Award Number: 13GW0248B

Abstract
Background: Thyroidectomy is one of the most commonly performed surgical procedures. The region of the neck has a very complex structural organization. It would be beneficial to introduce a tool that can assist the surgeon in tissue discrimination during the procedure. One such solution is the noninvasive and contactless technique, called hyperspectral imaging (HSI).

Methods: To interpret the HSI data, we implemented a supervised classification method to automatically discriminate the parathyroid, the thyroid, and the recurrent laryngeal nerve from surrounding tissue (muscle, skin) and materials (instruments, gauze). A leave-one-patient-out cross-validation was performed.

Results: The best performance was obtained using support vector machine (SVM) with a classification and visualization in less than 1.4 seconds. A mean patient accuracy of 68% ± 23% was obtained for all tissues and material types.

Conclusions: The proposed method showed promising results and have to be confirmed on a larger cohort of patient data.

Keywords
computer assisted surgery, head and neck, imaged guided surgery, intraoperative imaging, surgery, thyroidectomy

1 INTRODUCTION

Thyroidectomy is one of the most commonly performed surgical procedures.¹ Benign and malignant thyroid diseases, such as hyperthyroidism, suspicious thyroid nodules, and thyroid cancer are being treated by this procedure. The region of the neck has a complicated structural organization with a lot of nerves and blood vessels well integrated into and around the thyroid gland. Damage of these structures can lead to severe complications and consequences. Parathyroid glands are difficult to detect intraoperatively and to be preserved because of variations of their location and their small size. In addition, they are difficult to distinguish from the surrounding tissue.²⁻⁴ Accidental removal of parathyroid glands is reported to occur in 15%⁵ to 19% of patients.² The most frequent complication following thyroidectomy is temporary (up to 46%) or permanent (up to 2%) hypoparathyroidism.²⁻⁵ A second complication results from the damage of the recurrent laryngeal nerve (RLN) which can lead to palsy. Consequences of RLN palsy are dysphonia, weakened cough, predisposition to aspiration, and life-threatening airway obstruction.⁶ Therefore, the key aim of the surgeon is the precise identification of parathyroid glands and nerve tissue.⁷

Imaging supports the physician in the identification of structures. The standard preoperative imaging technique used for the diagnosis of thyroid diseases is ultrasonography (USG).⁸⁻¹⁰ However, the use of USG is limited due to significant interobserver variability, and its ability to provide three-dimensional (3D) information is limited.¹¹⁻¹³ Other diagnostic tools, such as computerized tomography (CT) and magnetic resonance imaging (MRI), are invasive and costly.¹⁴⁻¹⁶ Thus, it would be beneficial to introduce a noninvasive and contactless tool for the identification of structures during the procedure. Hyperspectral imaging (HSI), a technique that provides detailed information about the chemical composition of an object based on the wavelength and intensity of the radiation emitted, reflects tissue characteristics and can be used as a noninvasive and contactless tool to discriminate different tissue types.¹⁷⁻¹⁹

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2020 The Authors. The International Journal of Medical Robotics and Computer Assisted Surgery published by John Wiley & Sons Ltd.
of diseases of the parathyroid glands is ultrasound imaging, but this is only useful in enlarged glands. In the research context, fluorescence near-infrared (NIR) imaging using external contrast agents has been evaluated as an intraoperative technique for the in vivo identification of normal and pathological parathyroid glands using the administration of low-dose Methylene Blue (MB) and indocyanine green (ICG). The parathyroid glands were intraoperatively detected in 93% of patients with ICG fluorescence imaging. Moreover, the autofluorescence of the parathyroid glands using NIR imaging has been investigated. These studies have shown that this noninvasive technique, which does not require any external contrast agent, enables the localization of the parathyroid glands. The standard intraoperative method to identify the RLN is its meticulous preparation and visualization combined with the neuromonitoring technique. Also, fluorescence imaging techniques were evaluated in the research context. However, it has the same limitations than previously described. Albeit, none of these alternative fluorescence methods have so far been accepted for routine use. This might have several reasons. The contrast agents can lead to neurotoxic reactions and the imaging procedure requires preparation time for injecting the contrast agent into the patient prior to or during the operation. Furthermore, the autofluorescence signal of the parathyroid is weak and the measurement system is very expensive. Also, this technique has limitations in the case of abnormal parathyroid tissue. For example, it has been shown that the autofluorescent signal emission was lower for patients with primary hyperparathyroidism.

To achieve a tool supporting the identification of the parathyroid glands and the RLN during open thyroid dissections, we developed and evaluated a system combing hyperspectral imaging (HSI) with a supervised classification approach. HSI is a noninvasive contactless technique, which has recently been introduced into the medical field. It combines imaging and spectroscopy and provides both, spectral and spatial information. Each acquired hypercube is a collection of 2D images with specified x and y dimensions, taken at different wavelengths usually in the range of ultraviolet to NIR. Each tissue has a unique spectral signature which corresponds to the optical properties of the tissue content (eg, fat and water).

Medical applications of the combination of both technologies are the identification of brain tumor margins during neurosurgery, resection of head and neck tumor tissue, and tongue tumors. However, few works concern the identification of the parathyroid and the RLN. Barberio et al showed that a decision tree based on spectral threshold values enabled to differentiate the parathyroid from the thyroid gland. In Schols et al, a differentiation of the spectra of nerve and fatty tissue was demonstrated. Moreover, machine-learning approaches based on HSI data have been evaluated to automatically classify tissues. For example, k-nearest neighbors (k-NN), random forest (RF), support vector machines (SVMs), artificial neuronal networks (ANN), and band selection have been applied. Furthermore, studies have pointed out that the SVM is well adapted to the processing of HSI data.

In this work, an SVM-based approach was evaluated to highlight the parathyroid and thyroid glands as well as the RLN from surrounding structures (eg, instruments) and tissues (eg, skin, connective tissue of the neck, muscles), using HSI data. HSI recordings were done during nine open thyroidectomies. In order to enable an intraoperative use of the tool, an important requirement was that the HSI data of the region of interest (ROI) are analyzed in real-time.

2 | METHODS

2.1 | HSI system

The commercial TIVITA Tissue system (Diaspective Vision GmbH, Am Salzhaft, Germany) was used in this study. It has a push broom scanner providing hypercubes in the range from 500 to 1000 nm with a spectral resolution of 5 nm. The size of the images is 640 × 480 pixels (x, y-axis). The distance between the camera and the in vivo tissue is 30 cm providing a spatial resolution of 0.1 mm/pixel. The integrated illumination unit consists of six halogen spots (20 W each). The camera is mounted on a mobile medical cart system. A black/white balancing of the system was performed for calibration only once initially to all acquisitions in the operation room. Further technical details about the system are given in and the HSI device used during the surgical procedures is depicted in Reference 30. To reduce interferences of external light sources, the lights in the operating room were switched off during image recording (approximately 10 seconds).

In vivo human tissue measurements were performed with the HSI system at the Department of Visceral, Transplant, Thoracic and Vascular Surgery of the University Hospital of Leipzig (Leipzig, Germany) during thyroid and parathyroid surgery. The study was approved by the Ethics Committee of the University Hospital of Leipzig, Germany (registration number 393/16-ek).

2.2 | Patient database

The data set includes the HSI data of nine patients with different histological findings (Table 1). The surgeon annotated intraoperatively, mainly focusing on the thyroid and parathyroid glands, the nerves and the surrounding tissue, that is, the skin tissue, connective tissue, and muscle. Additional labels were positioned in the operational gauze and instruments in order to identify the background accurately later (post-surgery). The structures presented in each patient are depicted in Table 2. Parathyroid glands were visible in seven patients and the RLN in four patients.

2.3 | Classification framework

The aim of the classification was the automatic identification of the parathyroid gland and the RLN using the HSI data. The approach described in Figure 1 was performed on the classes reported in Table 2 by using Python with the toolbox scikit-learn. The different steps are described in more detail in the following subsections.
The data was preprocessed in two steps. The spectra were initially smoothed with a Savitzky-Golay filter with a polynomial order of three and a window length of nine wavelengths. Additionally, spectra were normalized using the Standard Normal Variate (SNV) algorithm by setting the reflectance mean and SD values to zero and one, respectively.

The classes include different numbers of samples (Table 1). This can influence the classification results and this issue has to be addressed by balancing the data. Balancing means that the number of spectra in each class is equalized. Three setups were considered. First, the original number of spectra was used (unbalanced data). Secondly, the number of spectra in the parathyroid class was defined as a reference and the same number of spectra in the other classes was randomly selected or duplicated. This setup was called underbalanced data. Finally, all classes were provided with the same number of spectra corresponding to the largest class (the thyroid class). This was performed using the method SMOTE. This setup was called overbalanced data.

### Preprocessing and class balancing

The data was preprocessed in two steps. The spectra were initially smoothed with a Savitzky-Golay filter with a polynomial order of three and a window length of nine wavelengths. Additionally, spectra were normalized using the Standard Normal Variate (SNV) algorithm by setting the reflectance mean and SD values to zero and one, respectively.

The classes include different numbers of samples (Table 1). This can influence the classification results and this issue has to be addressed by balancing the data. Balancing means that the number of spectra in each class is equalized. Three setups were considered. First, the original number of spectra was used (unbalanced data). Secondly, the number of spectra in the parathyroid class was defined as a reference and the same number of spectra in the other classes was randomly selected or duplicated. This setup was called underbalanced data. Finally, all classes were provided with the same number of spectra corresponding to the largest class (the thyroid class). This was performed using the method SMOTE. This setup was called overbalanced data.

### Classification

Six classification approaches were tested using the 10-fold cross-validation: k-NN, RF, logistic regression (LR), SVM with linear, and
radial basis function (rbf) kernels, as well as four-layer perceptron network (MLP) with 10 perceptrons in the first hidden layer and 4 perceptrons in the second hidden layer. The SVM and k-NN showed the best F1-Score. Afterward, the parameters of both methods (eg, gamma of the SVM or number of neighbors of k-NN) were optimized. Finally, the best models and their parameters were used to an end test.

Both supervised classification algorithms were trained and tested using a leave-one-patient-out cross-validation. This means, that for each step the annotated spectra of eight patients were used for the training of the classification algorithm, while the entire HSI data of the last patient was used for testing. This step was repeated nine times by each time a different patient was used for testing. Afterward, the configuration providing the best leave-one-patient-out cross-validation results of nerve, parathyroid, thyroid, and other tissues were taken for visualization.

2.4 Visualization tool and background detection

A visualization tool was implemented to enhance the pixels of the test HSI data, which were classified as parathyroid, thyroid, and nerve.

In addition to the classification, the background (eg, surgical instruments, gloves, and gauze) was detected based on the following tree decision algorithm using threshold values which were empirically defined based on observations (Algorithm Background Detection).

Algorithm Background Detection

Input: Set of reflectance data as XXYxM matrix

\[
R = (r_{x,y,m})\text{ with } X \text{ is the width, } Y \text{ is the height, and } M \text{ is the number of wavelengths of the HSI Cube}
\]

Output: Set of pixel data as XXYxM matrix \( P = \{p_{x,y}\} \)

1: For \( x \) in range (\( X \))
2: For \( y \) in range (\( Y \))
3: \( A = \text{Mean}(r_{x,y,m}) \text{ with } 500 \text{ nm} \leq m \leq 1000 \text{ nm} \)
4: \( B = \text{Mean}(r_{x,y,m}) \text{ with } 510 \text{ nm} \leq m \leq 570 \text{ nm} \)
5: \( C = -\log(B/\text{mean}(r_{x,y,m}) \text{ with } 650 \text{ nm} \leq m \leq 710 \text{ nm} - 0.1) / 1.6 \)
6: If \( A < 0.1 \) and \( B > 0.7 \) and \( C < 0 \)
7: \( p_{x,y} \) is background

Afterward, the pixels classified as parathyroid, thyroid, and nerve and which did not belong to the background were visualized.

2.5 Validation of performance

The classification results were evaluated based on several statistical measures: the F1-Score (the quotient of two times of the product of precision and recall divided by the sum of precision and recall), Matthews Correlation Coefficient (MCC), specificity (the quotient of the sum of the true negative and the sum of condition negative), and sensitivity (the quotient of the sum of the true positive and the sum of condition positive).

3 RESULTS

Figure 2 shows the mean reflectance spectra of the different tissues and materials. The reflectances of the skin, gauze, and instruments were larger than the other tissues, especially in the range 500 to 600 nm. This justified the approach to remove the gauze and instruments based on threshold values of the spectra. The instruments were well detected with condition A, the gauze with B. Moreover, it was observed that blood was present on the gauze and the instruments reflected the signal of surrounding tissue. The condition C enabled to remove most of these signals. Furthermore, from 500 to 600 nm, which represents the spectral area of hemoglobin, the mean spectra of the biological tissues were similar. Up to 600 nm, larger differences of the reflectance were observable. At the wavelength of 650 nm, the mean reflectance of the RLN reached a peak, while the mean spectra of the parathyroid glands showed lower reflectance values. Additional large differences between the mean spectra of parathyroid and nerve tissue were observable in the NIR range (800-1000 nm). The nerve mean spectra revealed lower reflectance values than the other tissues from 800 to 960 nm. On the other hand, the mean reflectance values of the parathyroid were larger than for the muscle and the thyroid gland in this range. In Figures 3 and 4, the mean reflectance spectra with standard deviations of the thyroid, parathyroid, and nerve tissue as well as the result of a linear discriminant analysis are depicted. They show a clear differentiation of the parathyroid and thyroid. However, the discrimination of the RLN is not so obvious. The spectra of the RLN and parathyroid overlap and a linear separability of the nerve is not clear. In summary, the mean spectra of the different tissues and surgical materials showed significant differences, which could be detected by machine-learning algorithms in order to be capable to differentiate them automatically. The discrimination of the RLN seems more complex.

The first test with a 10-fold cross-validation showed that the k-NN, MLP, RF, and SVM with a rbf kernel obtained the best results with an averaged precision from 96% (MLP) to 99% (k-NN). Especially the classification of the parathyroid and the RLN was better with the RF, k-NN, and the SVM as compared with the other algorithms (F1-Score over 70% for nerve tissue and over 90% for parathyroid glands). However, the RF demonstrated overfitting due to high variations between the 10 validations for nerve tissue. The use of unbalanced data lead to a lower performance than the under- and overbalanced data sets. Due to the best results of k-NN and SVM with rbf kernel, they were selected for the end test.

The best results of the leave-one-patient-out cross-validation are reported in Table 3. They were achieved with the SVM using the underbalanced data set. The value of the parameter Gamma was calculated with the reciprocal of the count of the used features in training and the parameter C as the correct classification of training examples against maximization of the decision function’s margin with 1000. The classification was performed in the whole range (500-1000 nm). Patients 1, 2, 3, 4, and 6 had high accuracy above 81%. The parathyroid was classified with a specificity above 93% and a sensitivity up to 90% (patient 3). Moreover, the thyroid was classified with a specificity and a sensitivity higher than 99% and 82%,
respectively. The scores of the nerve, only present in patient 1, were also high (specificity = 96% and sensitivity = 100%). Patients 5, 7, 8, and 9 obtained lower accuracy values between 29% (patient 7) and 66% (patient 5). The classification of the parathyroid remained good with a specificity of more than 84% (for patient 9). In addition, the sensitivity in patients 5 and 9 was higher than 73%, but lower for patients 7 and 8 (35% and 61%). Similarly, the F1-Scores of the thyroid classification were 77% and 65% for patients 7 and 8, but only 34% for patient 5. The F1-Scores of the nerve were 2% and 5% for patients 7 and 8.

Figure 5 depicts the visualization of the classification for all patients. The circles represent the annotations performed by the surgeon. The colors blue, yellow, and green correspond to the pixels classified as thyroid, nerve, and parathyroid. In general, parathyroid and thyroid tissues are identified correctly. Furthermore, these results clearly show that the nerve tissue is difficult to detect. Finally, the background could be largely removed. In Appendix S1 further classification results are shown.

4 | DISCUSSION

In this study, the potential of HSI combined with automatic fast tissue classification and visualization was analyzed in endocrine tissue for the first time. The patient recordings were acquired with a commercial HSI system working in the range from 500 to 1000 nm during open-neck surgeries, such as parathyroidectomies and thyroidectomies. An accurate discrimination and visualization of parathyroid
and thyroid tissue was obtained in less than 1.4 seconds with a standard hardware (i7-7500 CPU@ 2.70 GHz-2.90 GHz) with a sensitivity of 65% ± 17% and specificity of 94% ± 6% for the parathyroid; sensitivity of 75% ± 29% and specificity of 96% ± 3% for the thyroid. The performance of the nerve classification was lower with an average sensitivity of 35% ± 56% and an average specificity of 97% ± 1%. The computing time was therefore compatible with intraoperative use.

In other studies, the spectra of nerve tissue, parathyroid, and thyroid glands have been described. Schols et al recorded these
### Table 3: Results of the best-fitted classification (here: SVM with rbf kernel) using a leave-one-patient-out cross-validation

| Class       | Patient 1 | Patient 2 | Patient 3 | Patient 4 | Patient 5 | Patient 6 | Patient 7 |
|-------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
|             | F1-Score  | Sensitivity | Specificity | MCC | Accuracy | F1-Score  | Sensitivity | Specificity | MCC | Accuracy | F1-Score  | Sensitivity | Specificity | MCC | Accuracy | F1-Score  | Sensitivity | Specificity | MCC | Accuracy | F1-Score  | Sensitivity | Specificity | MCC | Accuracy | F1-Score  | Sensitivity | Specificity | MCC | Accuracy |
| Thyroid     | 0.89      | 0.82      | 0.99      | 0.83 | 0.81     | 0.97      | 0.96      | 0.99      | 0.96 | 0.92     | 0.74      | 0.82      | 0.94      | 0.70 | 0.90     | 0.34      | 0.22      | 0.98      | 0.33 | 0.66     | 0.69      | 0.59      | 1.00      | 0.70 | 0.83     | 0.17      | 0.73      | 0.85      | 0.23 | 0.33     | 0.02      | 0.02      | 0.97      | −0.01|
| Parathyroid | Nan       | Nan       | 0.99      | Nan | Nan      | 0.97      | 0.93      | 0.93      | 0.69 | Nan      | 0.68      | 0.62      | 0.97      | 0.64 | Nan      | 0.32      | 0.35      | 0.93      | 0.27 | Nan      | 0.69      | 0.59      | 1.00      | 0.70 | Nan      | 0.02      | 0.02      | 0.97      | −0.01|
| Nerve       | 0.27      | 1.00      | 0.96      | 0.39 | 0.04     | 0.71      | 0.90      | 0.93      | 0.69 | 0.80     | 0.86      | 0.86      | 0.99      | 0.88 | 0.95     | 0.96      | 0.94      | 0.99      | 0.95 | 0.92     | 0.69      | 0.64      | 0.99      | 0.67 | 0.67     | 0.69      | 0.59      | 1.00      | 0.70 |
| Skin        | 1.00      | 1.00      | 1.00      | 0.99 | 0.99     | 1.00      | 1.00      | 1.00      | 0.99 | 0.99     | 1.00      | 1.00      | 1.00      | 0.99 | 0.99     | 0.96      | 0.94      | 0.94      | 0.94 | 0.92     | 0.67      | 0.64      | 0.98      | 0.67 | 0.67     | 0.69      | 0.59      | 1.00      | 0.70 |
| Muscle      | 0.09      | 0.08      | 0.96      | 0.04 | 0.04     | 0.92      | 0.96      | 0.99      | 0.99 | 0.99     | 0.69      | 0.64      | 0.98      | 0.67 | 0.95     | 0.96      | 0.92      | 1.00      | 0.92 | 0.92     | 0.67      | 0.64      | 0.98      | 0.67 | 0.67     | 0.69      | 0.59      | 1.00      | 0.70 |
| Gauze       | Nan       | Nan       | 0.93      | Nan | Nan      | 0.98      | 0.99      | 0.99      | 0.99 | 0.99     | 0.96      | 0.96      | 0.99      | 0.99 | 0.99     | 0.96      | 0.94      | 0.94      | 0.94 | 0.92     | 0.96      | 0.92      | 1.00      | 0.92 | 0.92     | 0.67      | 0.64      | 0.98      | 0.67 |
| Instrument  | 0.81      | 0.72      | 0.98      | 0.77 | 0.77     | 0.59      | 1.00      | 0.88      | 0.60 | 0.58     | 0.98      | 1.00      | 0.88      | 0.99 | 0.97     | 0.96      | 0.94      | 0.94      | 0.94 | 0.92     | 0.96      | 0.92      | 1.00      | 0.92 | 0.92     | 0.67      | 0.64      | 0.98      | 0.67 |

(Continues)
structures during (para-)thyroid surgery using diffuse spectroscopy based on a fiber probe. In Reference 23 as well as in Reference 7, the differences between the spectra of these distinct structures and tissues were pointed out. We observed similar features for the spectra of the thyroid and parathyroid. Especially, the reflectance spectrum valley at 760 nm (deoxygenated hemoglobin) was more pronounced for the parathyroid than for the thyroid. Also, the water content (960 nm) was larger in the thyroid (lower reflectance value) than in the parathyroid. In Reference 33, a hyperspectral camera setup in the visible spectrum from 400 to 700 nm was developed to monitor in vivo the optical properties of different tissue types during mastoidectomy, parotidectomy, and neck dissection. Our recorded spectra of the parathyroid glands and the RLN in the range from 500 to 1000 nm were similar to those in Reference 33. Furthermore, in Reference 33, the authors showed that the variability of the spectrum of the nerve data was higher than for the other tissues.

The optical characterization of organs and tissues in neck surgery based on HSI has been examined in a couple of previous works. In Reference 34, several systems (fiber probe, Overlay Tissue Imaging System [OTIS], and OTIS with NIR autofluorescence) for the label-free real-time identification of the parathyroid gland were presented. The probabilities of identification of the gland were established based on the measured signal intensities. The systems could guide the surgeon by real-time mode and the successful detection of the parathyroid gland was proven by the surgeon and by histological findings. Our recorded spectra of the parathyroid glands and the RLN in the range from 500 to 1000 nm were similar to those in Reference 33. Furthermore, in Reference 33, the authors showed that the variability of the spectrum of the nerve data was higher than for the other tissues.

### TABLE 3 (Continued)

| Class      | F1-Score | Sensitivity | Specificity | MCC  | Accuracy |
|------------|----------|-------------|-------------|------|----------|
| Muscle     | Nan      | Nan         | 0.77        | Nan  |          |
| Gauze      | 0.00     | 0.00        | 0.91        | 0.70 | 0.48     |
| Instrument | 0.64     | 0.83        | 0.70        | 0.66 |          |
| Patient 8  | Thyroid  | 0.77        | 0.97        | 0.91 | 0.75     | 0.38     |
| Parathyroid| 0.36     | 0.61        | 0.93        | 0.36 |          |
| Nerve      | 0.05     | 0.03        | 0.98        | 0.03 |          |
| Skin       | Nan      | Nan         | 0.76        | Nan  |          |
| Muscle     | Nan      | Nan         | 1.00        | Nan  |          |
| Gauze      | 0.46     | 0.30        | 1.00        | 0.32 |          |
| Instrument | Nan      | Nan         | 0.78        | Nan  |          |
| Patient 9  | Thyroid  | 0.65        | 0.51        | 0.94 | 0.47     | 0.50     |
| Parathyroid| 0.36     | 0.79        | 0.84        | 0.37 |          |
| Nerve      | Nan      | Nan         | 0.83        | Nan  |          |
| Skin       | 0.03     | 0.01        | 1.00        | 0.10 |          |
| Muscle     | Nan      | Nan         | 0.99        | Nan  |          |
| Gauze      | 0.61     | 0.59        | 0.89        | 0.48 |          |
| Instrument | 0.23     | 0.44        | 0.93        | 0.23 |          |

Abbreviations: MCC, Matthews Correlation Coefficient; SVM, support vector machine.

OTIS systems do not record any spatial information, but the combination with autofluorescence enables the visualization of the parathyroid. Moreover, supervised classification approaches for the recognition of tissue in HSI have already been reported. In Reference 23, an SVM with polynomial kernel classification without visualization between fat and nerve tissue was achieved with cross-validations. An accuracy up to 78% was obtained. The sensitivity and specificity values varied in a big range from 66% to 100%, and from 46% to 100% for fat and nerve, respectively. These values cannot be compared to ours since we performed another validation method (leave-one-patient-out cross-validation). Furthermore, Schols et al presented an approach to automatically distinguish the parathyroid from thyroid glands and fat from parathyroid tissue with an accuracy of 82%, based on SVM in leave-one-out cross-validation. Similar and higher accuracy values were obtained in this study for five out of nine patients. The advantages of our method compared to the mentioned studies before are the acquisition of spatial information as well as the identification and visualization of the surrounding tissues. The visualization tool is crucial to assist the surgeon. We have achieved accuracies up to 92% to classify more tissues than in Schols et al (only fat, parathyroid, and thyroid gland) and visualized the classification results. A further advantage of our study in comparison to Schols et al is that we performed a leave-one-patient-out cross-validation, which reflects the real case in the operation, when unknown data are classified.

Compared to previous works, the originality of our approach in thyroid and parathyroid surgery is the additional classification of the RLN. A permanent control of the nerve is important to avoid possible injuries during surgery. The spectra of the nerve showed relevant
differences to the other tissues in the NIR range (Figure 3). Automatic discrimination should therefore be possible in theory. However, the classification scores remain unsatisfactory with a mean F1-Score of 41%. Several reasons could explain this fact. Nerves are thin structures that are complex to annotate in the HSI data. Also, the higher variability in the reflectance spectra of nerve tissue (observed in Wisotzky et al and here in Figures 4 and 5) combined with the small database (four patients) used in our study, may reduce the performance of the classification method for nerve tissue. Another clinical interest could be the identification of lymph nodes and their differentiation with regard to benign and malignant cell content, which is hardly impossible by the surgeon’s optical evaluation only.

Patients 5, 7, 8, and 9 had lower accuracies than the other patients. A possible reason for this is their final histology. In Reference 36, the impact of hyperthyroidism on the detection of the parathyroid using NIR autofluorescence was evaluated. It was shown that the pathology decreases the parathyroid gland’s autofluorescence and increases the autofluorescence of the surrounding tissue.

In general, a limitation of the current study is the small sample size (n = 9). Further studies with larger patient numbers are needed to achieve more training dataset, which would have improved the classification’s accuracy. Additionally, a larger number could enable to correlate the classification with the different pathologies.

As already pointed out by previous studies, the SVM algorithms provided a good classification performance in comparison with classical approaches. Another possible improvement would be a more complex neural network, such as a convolutional neuronal network, which would take into account the spatial as well as the spectral information. However, the computing time of such approaches has to be compatible with time-efficient application in the operating room.

Finally, the future benefit of our method is that all relevant tissues (lymph nodes, nerves, and thyroid/parathyroid tissues) might be detected using one single imaging technique assisted by artificial intelligence methods in a very quick and easy fashion without disturbing the surgical procedure relevantly.

5 | CONCLUSION

There is an urgent need for an efficient imaging device, which can guide surgeons during open-neck endocrine surgery. We presented an approach which combines HSI with automatic tissue classification. This represents the first method for the identification, differentiation, and visualization of the parathyroid and thyroid glands, skin, connective tissue, muscle, gauze, and surgical instruments. A mean patient accuracy of 68% ± 23% was obtained for all tissues and material types. The sensitivity and specificity values were 65% ± 17% and 94% ± 6% for the parathyroid, 75% ± 29% and 96% ± 3% for the thyroid, and 35% ± 56% and 97% ± 1% for the nerve. Future studies should also address the differentiation of malignant and benign parathyroid and thyroid tissue as well as the lymph nodes in the neck with the intent of improved operation outcomes in the context of endocrine surgery.

ACKNOWLEDGMENTS

This study was funded by the Federal Ministry of Education and Research 13GW0248B. The authors declare that they have no conflict of interest. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Experimental hyperspectral measurements from patients for the evaluation of the new technology have obtained the ethics approval by the Ethics Committee of the University Hospital of Leipzig under 393/16-ek. Informed consent was obtained from all individual participants included in the study.

CONFLICT OF INTEREST

The authors have no relevant financial interests in the manuscript and no other potential conflicts of interest to disclose.

REFERENCES

1. Adam MA, Thomas S, Youngwirth L, et al. Is there a minimum number of thyroidectomies a surgeon should perform to optimize patient outcomes? Ann Surg. 2017;265(2):402-407. https://doi.org/10.1097/SLA.S000000000001688.
2. Falco J, Dip F, Quadri P, de la Fuente M, Brunello M, Rosenthal RJ. Increased identification of parathyroid glands using near infrared light during thyroid and parathyroid surgery. Surg Endosc. 2017;31(9):3737-3742. https://doi.org/10.1007/s00464-017-5424-1.
3. McWade MA, Paras C, White LM, et al. Label-free intraoperative parathyroid localization with near-infrared autofluorescence imaging. J Clin Endocrinol Metab. 2014;99(12):4574-4580. https://doi.org/10.1210/jc.2014-2503.
4. De Leeuw F, Breuskin I, Abbaci M, et al. Intraoperative near-infrared imaging for parathyroid gland identification by auto-fluorescence: a feasibility study. World J Surg. 2016;40(9):2131-2138. https://doi.org/10.1007/s00268-016-3571-5.
5. Kim SW, Lee HS, Lee KD. Intraoperative real-time localization of parathyroid gland with near infrared fluorescence imaging. Gland Surg. 2017;6(5):516-524. https://doi.org/10.21037/gs.2017.05.08.
6. Serpell JW, Lee JC, Yeung MJ, Grodski S, Johnson W, Bailey M. Differential recurrent laryngeal nerve palsy rates after thyroidectomy. Surgery. 2014;156(5):1157-1166. https://doi.org/10.1016/j.surg.2014.07.018.
7. Barberio M, Matkabi M, Gockel I, et al. Hyperspectral based discrimination of thyroid and parathyroid during surgery. Curr Direct Blinded Eng. 2018;4(1):399-402. https://doi.org/10.1515/cdbme-2018-0095.
8. Ahmed R, Khan N, Ellemdin S, Gayaparsad K. Ultrasound—first imaging modality in the detection of parathyroid adenomas. South African J Radiol. 2008;12(3):64. https://doi.org/10.4102/sajr.v12i3.558.
9. Tummers QRG, Schepers A, Hamming JF, et al. Intraoperative guidance in parathyroid surgery using near-infrared fluorescence imaging and low-dose Methylene Blue. Surgery. 2015;158(5):1323-1330. https://doi.org/10.1016/j.surg.2015.03.027.
10. Yu HW, Chung JW, Yi JW, et al. Intraoperative localization of the parathyroid glands with indocyanine green and Firefly(R) technology during BABA robotic thyroidectomy. Surg Endosc. 2017;31(7):3020-3027. https://doi.org/10.1007/s00464-016-5330-y.
11. Zaldi N, Bucak E, Okoh A, Yazici P, Yigitbas H, Berber E. The utility of indocyanine green near infrared fluorescent imaging in the identification of parathyroid glands during surgery for primary
hyperparathyroidism: ICG Fluorescent Imaging During Parathyroidectomy. J Surg Oncol. 2016;113(7):771-774. https://doi.org/10.1002/jso.24240.

12. Falco J, Dip F, Quadri P, de la Fuente M, Rosenthal R. Cutting edge in thyroid surgery: autofluorescence of parathyroid glands. J Am Coll Surg. 2016;223(2):374-380. https://doi.org/10.1016/j.jamcollsurg.2016.04.049.

13. Ladurner R, Al Arabi N, Guendogar U, Halfeldt K, Stepp H, Gallwjs J. Near-infrared autofluorescence imaging to detect parathyroid glands in thyroid surgery. Ann R Coll Surg Engl. 2018;100(1):33-36. https://doi.org/10.1308/rcsann.2017.0102.

14. Benmiloud F, Rebaudet S, Varoquaux A, Penaranda G, Bannier M, Denizot A. Impact of autofluorescence-based identification of parathyroids during total thyroidectomy on postoperative hypocalcemia: a before and after controlled study. Surgery. 2018;163(1):23-30. https://doi.org/10.1016/j.surg.2017.06.022.

15. Kahramangil B, Dip F, Benmiloud F, et al. Detection of parathyroid autofluorescence using near-infrared imaging: a multicenter analysis of concordance between different surgeons. Ann Surg Oncol. 2018;25(4):957-962. https://doi.org/10.1245/s10434-018-6634-2.

16. Koch M, Ntziachristos V. Advancing surgical vision with fluorescence imaging. Annu Rev Med. 2016;67(1):153-164. https://doi.org/10.1146/annurev-med-051914-020243.

17. Pollack G, Pollack A, Delfner J, Fernandez J. Parathyroid surgery and methylene blue: a review with guidelines for safe intraoperative use. Laryngoscope. 2009;119(10):1941-1946. https://doi.org/10.1002/lary.20581.

18. McWade MA, Sanders ME, Broome JT, Solórzano CC, Mahadevan-Jansen A. Establishing the clinical utility of autofluorescence spectroscopy for parathyroid detection. Surgery. 2016;159(1):193-203. https://doi.org/10.1016/j.surg.2015.06.047.

19. Halicke M, Fabelo H, Ortega S, Callicó GM, Fei B. In-vivo and ex-vivo tissue analysis through hyperspectral imaging techniques: revealing the invisible features of cancer. Cancer. 2019;11(6):756. https://doi.org/10.3390/cancers11060756.

20. Fabelo H, Ortega S, Ravi D, et al. Spatio-spectral classification of hyperspectral images for brain cancer detection during surgical operations. PLOS ONE. 2018;13(3):e0193721. https://doi.org/10.1371/journal.pone.0193721.

21. Fei B, Lu G, Wang X, et al. Label-free reflectance hyperspectral imaging for tumor margin assessment: a pilot study on surgical specimens of cancer patients. J Biomed Opt. 2017;22(8):1. https://doi.org/10.1117/1.JBO.22.8.086009.

22. Liu Z, Wang H, Li Q. Tongue tumor detection in medical hyperspectral images. Sensors. 2011;12(1):162-174. https://doi.org/10.3390/s120100162.

23. Schols RM, Bouvy ND, Amelink A, Wieringa FP, Alic L. Differentiation between nerve and adipose tissue using wideband (350-1,830 nm) in vivo diffuse reflectance spectroscopy. Lasers Surg Med. 2014;46(7):538-545.

24. Hosking A-M, Coakley BJ, Chang D, et al. Hyperspectral imaging in automated digital dermoscopy screening for melanoma: hyperspectral dermoscopy. Lasers Surg Med. 2019;17:214-222. https://doi.org/10.1002/lsm.23055.

25. Ortega S, Fabelo H, Camacho R, de la Luz PM, Callicó GM, Sarmiento R. Detecting brain tumor in pathological slides using hyperspectral imaging. Biomed Opt Express. 2018;9(2):818-831. https://doi.org/10.1364/BOE.9.000818.

26. Akbari H, Halig LV, Schuster DM, et al. Hyperspectral imaging and quantitative analysis for prostate cancer detection. J Biomed Opt. 2012;17(7):0760051. https://doi.org/10.1117/1.JBO.17.7.076005.

27. Nouri D, Lucas Y, Treuillet S. Hyperspectral interventional imaging for enhanced tissue visualization and discrimination combining band selection methods. Int J Comput Assist Radiol Surg. 2016;11(12):2185-2197. https://doi.org/10.1007/s11548-016-1449-5.

28. Ghamisi P, Plaza J, Chen Y, Li J, Plaza AJ. Advanced spectral classifiers for hyperspectral images: a review. IEEE Geosci Remote Sens Mag. 2017;5(1):8-32. https://doi.org/10.1109/MGRS.2016.2616418.

29. Kulcke A, Holmer A, Wahl P, Siemers F, Wild T, Daeschlein G. A compact hyperspectral camera for measurement of perfusion parameters in medicine. Biomed Tech (Berl). 2018;63(5):519-527.

30. Moulla Y, Reifenrath M, Rehmet K, et al. Hybridösoophagektomie mit intraoperativem Hyperspektral-Imaging: Videobeitrag. Chirurg. 2020;17. https://doi.org/10.1007/s00104-020-01139-1.

31. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002;16:321-357. https://doi.org/10.1613/jair.953.

32. Boughorbel S, Jarray F, El-Anbari M. Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. PLOS ONE. 2012;12(6):e0177678. https://doi.org/10.1371/journal.pone.0177678.

33. Wisotzky EL, Uecker FC, Arens P, Domerich S, Hilsmann A, Eisert P. Intraoperative hyperspectral determination of human tissue properties. J Biomed Opt. 2018;23(09):1. https://doi.org/10.1117/1.JBO.23.9.091409.

34. Thomas G, McWade MA, Nguyen JQ, et al. Innovative surgical guidance for label-free real-time parathyroid identification. Surgery. 2019;165(1):114-123. https://doi.org/10.1016/j.surg.2018.04.079.

35. Schols RM, Alic L, Wieringa FP, Bouvy ND, Stassen LPS. Towards automated spectroscopic tissue classification in thyroid and parathyroid surgery: Automated spectroscopic tissue classification. Int J Med Robot. 2017;13(1):e1748. https://doi.org/10.1002/rcs.1748.

36. Squires MH, Shirley LA, Shen C, Jarvis R, Phay JE. Intraoperative hyperspectral camera for measurement of perfusion parameters in medicine. Biomed Tech (Berl). 2018;63(5):519-527.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Maktabi M, Köhler H, Ivanova M, et al. Classification of hyperspectral endocrine tissue images using support vector machines. Int J Med Robotics Comput Assist Surg. 2020;16:e2121. https://doi.org/10.1002/rcs.2121