Micro-FTIR Hyperspectral Imaging Classification for Oral Cavity Histopathology Analysis

Matheus de Freitas Oliveira Baffa
Dept. of Computing and Mathematics
University of São Paulo
Ribeirao Preto - SP, Brazil
mbaffa@usp.br

Luciano Bachmann
Dept. of Physics
University of São Paulo
Ribeirao Preto - SP, Brazil
l.b@usp.br

Thiago Martini Pereira
Dept. of Science and Technology
Federal University of São Paulo
São José dos Campos - SP, Brazil
t.pereira@unifesp.br

Leandro Luongo Matos
State Cancer Institute of São Paulo
University of São Paulo
São Paulo - SP, Brazil
l.matos@fm.usp.br

Denise Maria Zezell
Center for Lasers and Aplications
Nuclear and Energy Research Institute
São Paulo - SP, Brazil
zezell@usp.br

Daniella Lúmara P. M. O. Peres
Center for Lasers and Aplications
Nuclear and Energy Research Institute
São Paulo - SP, Brazil
daniellalumara@usp.br

Joaquim Cezar Felipe
Dept. of Computing and Mathematics
University of São Paulo
Ribeirao Preto - SP, Brazil
jfelipe@ffclrp.usp.br

***Abstract*—Hyperspectral imaging (HSI) has emerged as a promising tool for integrating spatial and biochemical information in computational pathology analysis. While most studies on oral cavity cancer have employed reflectance-based HSI or Raman spectroscopy in the visible and near-infrared ranges, the potential of mid-infrared micro–Fourier transform infrared (micro-FTIR) spectroscopy remains largely unexplored. This study investigates the feasibility of using micro-FTIR hyperspectral data for the classification of oral cavity tissues. Mid-IR spectra provide detailed biochemical information, including protein, lipid, and nucleic acid signatures, which may be clinically relevant for early diagnosis and characterization of tumor margins. Tissue samples were imaged using micro-FTIR spectroscopy, and voxel-level spectra were preprocessed and classified using a fully-connected neural network. The proposed model achieved an accuracy of 88.41%, sensitivity of 87.64%, and area under the receiver operating characteristic curve (AUC) of 96.51%, demonstrating that micro-FTIR–based HSI can successfully differentiate between healthy and malignant oral cavity tissues. These findings provide the first systematic evidence supporting the clinical potential of mid-infrared spectroscopy in oral oncology.**

I. Introduction

Hyperspectral imaging (HSI) is an optical technique that combines high-resolution spatial data with detailed spectral information at each voxel [1]. Unlike conventional histopathological imaging, which primarily conveys morphological details, HSI captures the full spectral signature of a biological sample, encoding biochemical and histochemical composition [2]. This enables detection of subtle molecular variations such as changes in proteins, lipids, and nucleic acids, not visible with standard staining or brightfield microscopy [3].

These images can be acquired using various spectroscopy modalities. Absorbance spectroscopy, for instance, measures the fraction of incident light absorbed at specific wavelengths, linking attenuation to molecular concentrations [2]. Fourier transform infrared (FTIR) spectroscopy applies this principle to detect vibrational modes of molecular bonds, producing spectra whose peaks correspond to functional groups and chemical structures. In the mid-infrared range, micro-FTIR combines absorbance-based molecular specificity with microscopic spatial resolution, enabling biochemical characterization of tissues through their unique vibrational signatures [3].

Traditional computational histopathological analysis using hematoxylin and eosin (H&E) stained images achieves high accuracy in binary classification tasks such as distinguishing healthy from malignant tissue [4], [5]. However, performance often declines in complex cases with multiple coexisting pathologies or heterogeneous tissue composition. In oral cavity histology, a single biopsy may contain squamous cell carcinoma, inflammatory lesions, dysplasia, and benign proliferative changes within overlapping regions [6]. Under binary assumptions, algorithms may struggle to separate tumor margins from reactive or dysplastic epithelium, leading to misclassification.

Oral cavity cancer, which includes malignancies of the lip, other parts of the mouth, and the oropharynx, represents a significant public health concern due to its prevalence and potential for severe outcomes if not detected early. According to the World Cancer Research Fund, it is the 16^{th} most common cancer globally, ranking 12^{th} in incidence among men and 18^{th} among women, with an estimated 389,846 new cases

reported in 2022 [7]. Many oral health conditions, including early-stage oral cancers, are largely preventable and treatable if detected promptly, as highlighted by the World Health Organization [8].

In the literature, HSI has been applied to oral cavity assessment mainly through reflectance-based acquisition in the visible and near-infrared ranges, often focusing on regions such as the tongue, laryngeal tissues, or oral mucosa [6], [9]–[12]. Reflectance spectral signatures have shown to capture tissue-specific optical properties, enabling the discrimination between healthy and malignant areas. In addition to reflectance techniques, Raman spectroscopy has also been investigated for oral cancer diagnosis [13]. Raman spectroscopy measures the inelastic scattering of monochromatic light, providing molecular fingerprints based on vibrational modes of chemical bonds.

Despite the potential of HSI in oral oncology, most studies are restricted to reflectance-based acquisition or Raman spectroscopy, predominantly within the visible and near-infrared spectral ranges. This narrow methodological scope limits exploration of other spectroscopy modalities, such as micro–Fourier transform infrared (micro-FTIR) spectroscopy, which operates in the mid-infrared region and offers direct access to the molecular vibrational information of biological tissues. Therefore, micro-FTIR remains largely unexplored for oral cavity tissues, and its ability to provide high spatial and spectral resolution data has not been systematically validated in this clinical context.

Therefore, this study aims to explore and validate the feasibility of applying micro-FTIR spectroscopy to the analysis of oral cavity tissues, classifying each voxel (individual spectrum) as either cancerous or normal. Unlike most HSI studies, which are predominantly based on reflectance or Raman spectroscopy in the visible and near-infrared ranges, this work investigates a distinct optical approach in the mid-infrared domain, where molecular vibrational modes provide direct biochemical information.

II. Materials and Methods

The proposed method is summarized in the overview flowchart (Figure 1), which outlines the main stages from data acquisition to final classification. This section is organized into three parts. The first part describes the process of obtaining oral cavity tissue samples (Figure 1 - Step 1) and extracting their corresponding micro-FTIR spectra (Figure 1) - Step 2. The second part details the signal processing and data augmentation procedures applied to the spectra to ensure quality and representativeness (Figure 1 - Step 3). The final part presents the deep neural network architecture employed to detect patterns and classify each voxel as either cancerous or normal, enabling the precise mapping of pathological regions (Figure 1 - Step 4).

A. Data Acquisition

A total of 104 oral cavity biopsy samples were collected at Cancer Institute of Sao Paulo State (ICESP), on a previous

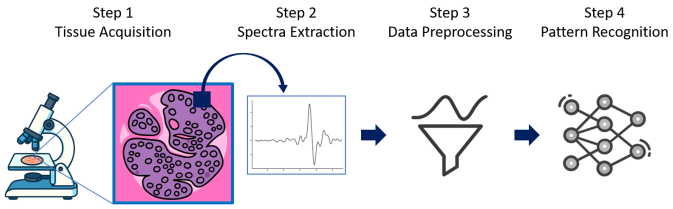

Fig. 1. Overview of the methodological workflow.

studies [14], [15] approved the study under protocol 228/14 (CAAE 32884214.5.0000.0065). Among these, 58 samples were histopathologically diagnosed as squamous cell carcinoma, while 46 were classified as healthy controls. For each patient, three distinct biopsy specimens were obtained to ensure representative sampling. To avoid data leakage during model evaluation, a patient-level grouping strategy was applied in the cross-validation procedure, ensuring that samples from the same patient were never simultaneously present in both the training and testing subsets.

Spectral data acquisition was performed using a FTIR Cary 660 spectrometer equipped with a microscope. The instrument measured absorbance intensities across the mid-infrared range of 866–1800cm^{-1}, with a spectral resolution of 4cm^{-1}, a spatial resolution of 5.5 μm, and an interval of collection of 2cm^{-1}. To minimize atmospheric interference, the relative humidity inside the equipment chamber was maintained below 5% by continuously purging with dry air. In total, 1,253,390 spectra, each with 467 frequencies, were extracted from all 104 samples.

B. Preprocessing and Augmentation

The spectral dataset comprised individual voxels, each representing a single micro-FTIR spectrum obtained from a small portion of an oral cavity tissue samples (voxel). These spectra, encompassing both cancerous and normal tissue signatures, were consolidated into a unified database, enabling a voxel-level classification framework that preserves fine-grained biochemical information.

Given the inherently high dimensionality of spectral data, a feature selection stage was introduced to enhance discriminative capacity and computational efficiency. The *SelectKBest* algorithm, coupled with the Chi-squared statistical test, was employed to rank spectral bands according to their relevance for class separation. The 100 most informative frequencies were retained, thereby reducing noise and redundancy while maintaining the core diagnostic content of the spectra.

To mitigate bias arising from class imbalance, where the number of spectra from cancerous and normal tissues differed, random undersampling was performed. This procedure equalized the representation of both classes by selectively removing samples from the majority class, ensuring that subsequent learning stages were not skewed toward more prevalent patterns.

In order to improve model robustness and generalization to unseen data, a controlled data augmentation strategy was

applied exclusively to the training set. Additive Gaussian noise was injected into the spectra, simulating natural variability introduced by biological heterogeneity, instrumental fluctuations, and measurement noise. Finally, all spectral features underwent z-score standardization via the *StandardScaler* method, producing zero-mean, unit-variance representations. This normalization step ensured consistent feature scaling, facilitating stable optimization and efficient convergence during neural network training.

C. Classification Methodology

Voxel-level classification was performed using a fully-connected neural network (FCNN) designed to operate on the 100-dimensional feature vectors obtained after preprocessing. The architecture was implemented as a sequential model composed of seven layers: one input layer, five hidden layers, and an output layer. The input layer contained 100 neurons to match the 100 frequencies selected from the spectra. The first hidden layer consisted of 512 neurons with ReLU activation, followed by three consecutive hidden layers with 256 neurons each, and a final hidden layer with 128 neurons. The output layer comprised a single neuron with a sigmoid activation function to produce binary predictions. All hidden layers employed the ReLU activation function to introduce non-linearity and promote efficient gradient propagation, while dropout was fixed at 40% after each hidden layer to enhance generalization. This architecture was selected through an automatic fine-tuning process based on a grid search, which systematically evaluated multiple configurations to identify the best-performing model.

Model optimization was performed using the Adam optimizer, which combines adaptive learning rates with momentum to achieve stable and efficient convergence. The binary cross-entropy loss function was used to quantify the classification error between predicted probabilities and ground-truth labels. Training was carried out for a maximum of 1,000 epochs with a batch size of 40,000 spectra, ensuring that each update was computed from a large and representative subset of the data. To prevent overfitting and reduce unnecessary computation, an early stopping strategy was employed by monitoring the validation loss, halting training if no improvement was observed over 50 consecutive epochs, and restoring the model weights to those obtained at the epoch with the best validation performance.

III. Experiments and Results

The proposed method was implemented in Python 3.12, utilizing TensorFlow 2.19.0 with Keras 3.10.0 for deep learning, and Pandas 2.3 with NumPy 2.1.2 for data manipulation and numerical operations. All experiments were executed on a high-performance workstation equipped with two Intel Xeon Silver processors, two NVIDIA RTX A4000 GPUs, 192 GB of RAM, and 2 TB of SSD storage, running Linux Ubuntu 22.04.

To assess model performance in a clinically scenario, we adopted a grouped 10-fold cross-validation protocol, in which the grouping variable corresponded to patient identifiers. This design guaranteed that all spectra (voxels) from the same patient were allocated exclusively to either the training or the testing set within each fold, thereby eliminating any possibility of patient-level data leakage.

Within this evaluation framework, model performance was measured using five metrics to capture different aspects of diagnostic quality. Accuracy (ACC) quantified the overall proportion of correctly classified samples, while precision (PRE) measured the proportion of predicted positive cases that were truly cancerous, reflecting the reliability of positive predictions. Sensitivity (SEN) assessed the proportion of actual cancerous spectra correctly identified, which is important for minimizing false negatives in a medical context. Specificity (SPEC) evaluated the proportion of normal spectra correctly classified, reducing false positives that could lead to unnecessary interventions. Finally, the area under the receiver operating characteristic curve (AUC) summarized the model's discrimination ability across all decision thresholds.

A. Results

The FCNN achieved a mean accuracy of 88.41% ($\pm$7.23%), indicating a high overall capability to correctly classify spectra as cancerous or normal. Precision and specificity were particularly strong, averaging 89.49% and 91.47%, respectively, which reflects the model's reliability in minimizing false positives and avoiding misclassification of healthy tissue. Sensitivity averaged 87.64%, demonstrating the model's effectiveness in detecting cancerous spectra, though with some variability across folds (standard deviation of 12.58%), suggesting that certain patient subsets may present more challenging spectral patterns. The AUC value of 96.51% ($\pm$4.00%) confirms an excellent discriminative performance across decision thresholds.

The results, summarized in Table I, show that while folds such as 5, 7, and 9 achieved the highest performance across all metrics, others like folds 2 and 4 exhibited comparatively lower sensitivity. This variation likely reflects inter-patient heterogeneity and subtle differences in spectral signatures that make certain cases more challenging to classify on some data separation. These findings reinforce the robustness of the proposed approach but also indicate that its sensitivity stability across patients could be improved. It is expected that incorporating a larger and more diverse dataset, capturing a broader range of spectral variations, could enhance generalization and lead to consistently high performance in all folds.

IV. Discussion

The use of HSI based on micro-FTIR spectroscopy for histopathological analysis offers significant advantages over traditional H&E staining images. While H&E provides valuable morphological information to the pathologist, HSI captures the complete spectral signature of tissue samples, reflecting their biochemical composition. This allows the detection of subtle molecular alterations that may precede morphological changes, enabling potentially earlier and more sensitive disease detection. Consequently, HSI can serve as a complemen-

TABLE I
RESULTS OF FCNN FOR ORAL CAVITY CLASSIFICATION

Fold	ACC	PRE	SEN	SPEC	AUC
0	80.17%	97.41%	73.40%	95.56%	94.73%
1	85.35%	68.74%	97.92%	79.60%	95.37%
2	77.58%	65.23%	91.93%	68.31%	86.77%
3	91.84%	94.20%	90.69%	93.24%	97.83%
4	80.69%	87.81%	60.83%	94.24%	92.89%
5	96.97%	99.91%	95.75%	99.80%	99.91%
6	84.23%	99.91%	74.31%	99.90%	99.52%
7	97.12%	98.66%	97.39%	96.38%	99.60%
8	94.18%	92.10%	94.90%	93.61%	98.68%
9	96.00%	90.88%	99.28%	94.03%	99.76%
Mean	**88.41%**	**89.49%**	**87.64%**	**91.47%**	**96.51%**
Std. Dev.	±7.23%	±11.90%	±12.58%	±9.39%	±4.00%

tary modality to conventional histology, expanding diagnostic capabilities beyond purely structural evaluation.

In this study, FCNN was trained on spectral features reduced via SelectKBest with a chi-square criterion. The analysis of mean spectra for each class revealed subtle differences between healthy and cancerous tissues. The spectral regions highlighted in yellow corresponded precisely to the frequencies selected by the feature selection process, indicating that the most discriminative features were not only statistically significant but also physically and chemically meaningful.

Despite these advantages, clinical implementation of micro-FTIR-based systems remains challenging. High-resolution spectrometers are costly, require specialized maintenance, and are not as widely available as H&E imaging systems, which are already integrated into pathology workflows. These practical constraints currently limit the method's adoption to research laboratories and specialized diagnostic centers. However, as technology advances and acquisition costs decrease, the clinical integration of such systems may become increasingly feasible.

Importantly, the flexibility of the proposed approach opens possibilities well beyond conventional histopathology. Potential applications include real-time margin assessment during surgery, rapid biopsy screening, and intraoperative tissue characterization, all leveraging the non-destructive and chemically sensitive nature of hyperspectral imaging. By providing rich biochemical information and high discriminative power, HSI combined with deep learning holds promise as a transformative tool for precision pathology.

V. CONCLUSION

This study addressed the voxel-level classification of micro-FTIR spectra from oral cavity tissue samples using an FCNN. By combining feature selection, class balancing, data augmentation, and normalization, the proposed approach achieved robust performance across multiple evaluation metrics, demonstrating its potential for discriminating between cancerous and normal spectra. Nevertheless, some variability in sensitivity across patients suggests room for improvement in generalization. Future work will investigate alternative architectures such as recurrent neural networks and convolutional neural networks, which can better exploit the sequential nature of spectral data, as well as attention-based mechanisms to enhance the modeling of discriminative spectral patterns.

ACKNOWLEDGMENT

This work was supported in part by the National Council for Scientific and Technological Development (CNPq) through the National Institutes of Science, Technology and Innovation Program under grant CNPq/INCT 406761/2022-1, and by the São Paulo Research Foundation (FAPESP) under grant 2021/00633-0. M.F.O.B. is supported in part by the Coordination for the Improvement of Higher Education Personnel (CAPES) under grant number 88887.498626/2020-00.

REFERENCES

[1] G. Jaiswal, R. Rani, H. Mangotra, and A. Sharma, "Integration of hyperspectral imaging and autoencoders: Benefits, applications, hyperparameter tunning and challenges," *Computer Science Review*, vol. 50, p. 100584, 2023.

[2] L. Gao and R. T. Smith, "Optical hyperspectral imaging in microscopy and spectroscopy–a review of data acquisition," *Journal of biophotonics*, vol. 8, no. 6, pp. 441–456, 2015.

[3] Z. Movasaghi, S. Rehman, and D. I. ur Rehman, "Fourier transform infrared (ftir) spectroscopy of biological tissues," *Applied Spectroscopy Reviews*, vol. 43, no. 2, pp. 134–179, 2008.

[4] J. Folmsbee, X. Liu, M. Brandwein-Weber, and S. Doyle, "Active deep learning: Improved training efficiency of convolutional neural networks for tissue classification in oral cavity cancer," in *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. IEEE, 2018, pp. 770–773.

[5] M. Aubreville *et al.*, "Automatic classification of cancerous tissue in laserendomicroscopy images of the oral cavity using deep learning," *Scientific reports*, vol. 7, no. 1, p. 11979, 2017.

[6] G. Lu *et al.*, "Histopathology feature mining and association with hyperspectral imaging for the detection of squamous neoplasia," *Scientific reports*, vol. 9, no. 1, p. 17863, 2019.

[7] World Cancer Research Fund International, "Mouth and oral cancer statistics," 2022, accessed: 2025-08-14. [Online]. Available: https://www.wcrf.org/preventing-cancer/cancer-statistics/mouth-and-oral-cancer-statistics/

[8] World Health Organization, "Oral health – fact sheet," 2024, accessed: 2025-08-14. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/oral-health

[9] M. Halicek *et al.*, "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*, vol. 10469. SPIE, 2018, pp. 8–16.

[10] G. Lu *et al.*, "Detection of head and neck cancer in surgical specimens using quantitative hyperspectral imaging," *Clinical Cancer Research*, vol. 23, no. 18, pp. 5426–5436, 2017.

[11] A. O. Gerstner *et al.*, "Hyperspectral imaging of mucosal surfaces in patients," *Journal of biophotonics*, vol. 5, no. 3, pp. 255–262, 2012.

[12] R. Martin, B. Thies, and A. O. Gerstner, "Hyperspectral hybrid method classification for detecting altered mucosa of the human larynx," *International journal of health geographics*, vol. 11, no. 1, p. 21, 2012.

[13] H. Krishna, S. K. Majumder, P. Chaturvedi, and P. K. Gupta, "Anatomical variability of in vivo raman spectra of normal oral cavity and its effect on oral tissue classification," *Biomedical Spectroscopy and Imaging*, vol. 2, no. 3, pp. 199–217, 2013.

[14] L. L. Matos *et al.*, "Cancer-associated fibroblast regulation by micrornas promotes invasion of oral squamous cell carcinoma," *Oral Oncology*, vol. 110, p. 104909, 2020.

[15] G. Mendes Menderico Junior *et al.*, "Microrna-mediated extracellular matrix remodeling in squamous cell carcinoma of the oral cavity," *Head & Neck*, vol. 43, no. 8, pp. 2364–2376, 2021.