Article

Raman Spectroscopy Analysis for Optical Diagnosis of Oral Cancer Detection

Ming-Jer Jeng [1,2], **Mukta Sharma** [1], **Lokesh Sharma** [3], **Ting-Yu Chao** [1], **Shiang-Fu Huang** [2,4,*], **Liann-Be Chang** [2,5,*], **Shih-Lin Wu** [3,6,7] and **Lee Chow** [8]

[1] Department of Electronic Engineering, Chang Gung University, Taoyuan 333, Taiwan
[2] Department of Otolaryngology-Head and Neck Surgery, Chang Gung Memorial Hospital, Linkou 244, Taiwan
[3] AI Innovation Research Center, Chang Gung University, Taoyuan 333, Taiwan
[4] Department of Public Health, Chang Gung University, Taoyuan 333, Taiwan
[5] Green Technology Research Center, Chang Gung University, Taoyuan 333, Taiwan
[6] Department of Cardiology, Chang Gung Memorial Hospital, Taoyuan 333, Taiwan
[7] Department of Computer Science and Information Engineering, Chang Gung University, Taoyuan 333, Taiwan
[8] Department of Physics, University of Central Florida, Orlando, FL 32816, USA
[*] Correspondence: shiangfu.huang@gmail.com (S.-F.H.); liann@mail.cgu.edu.tw (L.-B.C.)

check for updates

Abstract: Raman spectroscopy (RS) is widely used as a non-invasive technique in screening for the diagnosis of oral cancer. The potential of this optical technique for several biomedical applications has been proved. This work studies the efficacy of RS in detecting oral cancer using sub-site-wise differentiation. A total of 80 samples (44 tumor and 36 normal) were cryopreserved from three different sub-sites: The tongue, the buccal mucosa, and the gingiva of the oral mucosa during surgery. Linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) were used with principal component analysis (PCA) to classify the samples and the classifications were validated by leave-one-out-cross-validation (LOOCV) and k-fold cross-validation methods. The normal and tumor tissues were differentiated under the PCA-LDA model with an accuracy of 81.25% (sensitivity: 77.27%, specificity: 86.11%). The PCA-QDA classifier model differentiated these tissues with an accuracy of 87.5% (sensitivity: 90.90%, specificity: 83.33%). The PCA-QDA classifier model outperformed the PCA-LDA-based classifier. The model studies revealed that protein, amino acid, and beta-carotene variations are the main biomolecular difference markers for detecting oral cancer.

Keywords: oral cancer; Raman spectroscopy; PCA-LDA; PCA-QDA; cryopreserved tissue

1. Introduction

Oral cancer is one of the most common cancers globally and is the sixth most common malignancy, being closely associated with smoking, alcohol drinking, chewing tobacco, and consuming betel quid. The most common histology of oral cancer is squamous cell carcinoma (SCC) [1]. In males, it is the most common type of oral cancer and is found in various parts of the head and neck. It accounts for 90% of oral malignancies in 300,000 annually diagnosed cases [2]. According to Stewart [3], about 60% of new cases of oral cancer and 68% of deaths related to oral cancer reportedly occur in Asia. Oral squamous cell carcinoma (OSCC) is usually diagnosed late, resulting in an overall five-year survival rate of 50% [4]. The early detection and timely treatment of pre-malignancy may prevent oral potentially malignant disorders (OPMDs), which transform into oral cancer [5]. Although biopsies are the gold standard for diagnosing oral cancer, they are invasive and therefore painful, requiring an incision in

the tissue. Since biopsies are time-consuming and invasive, clinicians increasingly favor non-invasive techniques such as vital staining, light-based detection, and other optical diagnostic technologies [6].

Raman spectroscopy (RS) is an optical technique that is both fast and simple. It is one of the most widely-used techniques for the non-destructive characterization of molecules and materials [7]. RS probes the vibrational modes of a molecule, which are sensitive to its chemical bonds and provides a unique "fingerprint" that enable the identification of chemicals [8]. Most non-invasive techniques, such as light-based detection and optical diagnostic techniques, have great potential for screening and monitoring OPMDs [6]. However, no standalone method can accurately identify OPMDs. RS is a rapidly emerging technique with medical applications in the early diagnosis of various types of cancer [9–11]. Schut et al. [12] studied the effectiveness of RS for the in vivo classification of normal and dysplastic tissue by measuring the palatal tissues of rats. Sundar et al. [13] studied the application of RS to oral tissues in both normal and malignant stages. Changes in the ratio of relative intensities can be useful in analyzing oral tissues to detect oral malignancy. Many studies have demonstrated the use of RS to distinguish between normal and malignant, or among normal, pre-malignant, and malignant forms of oral mucosa using various preservation techniques and analytical methods [14–20]. In Lau et al. [21] studied the use of RS to differentiate normal from cancerous nasopharyngeal tissues. Lau [22] differentiated three stages of cancers in larynx tissue. Some authors have used exfoliated [23] and serum-based samples [24,25] to differentiate between normal and abnormal tissues. RS also has the potential to provide an objective intra-operative evaluation of the cancer surgical margins, favoring the detection of any residual tumor after surgery [26–28].

In this investigation, cryopreservation with fresh excision was used to study the efficacy of RS for classifying normal and tumor tissues. Cryopreservation is a process that preserves organelles, cells, tissues, and any other biological structures. In this work, samples of tissues were extracted after surgery and immediately preserved in liquid nitrogen. This method helped to prevent the alteration of the structure and morphology of tissue. Linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) classifiers with principal component analysis (PCA) were used to distinguish tumor from the healthy tissue structures of oral mucosa. Sub-site-wise differentiation of tongue cancer versus normal, buccal cancer versus normal, and gingiva cancer versus normal were performed. Generally, the QDA model is a better classifier than the LDA model. However, in some cases, the LDA model is better than the QDA model. Therefore, both the LDA and QDA models were studied. This work is the first to attempt the sub-site-wise differentiation of cryopreserved tissue samples using RS.

2. Materials and Methods

2.1. Patients and Samples

This study was approved by the Institutional Review Board (IRB) of Chang Gung Medical Foundation (IRB No: 201800420B0), Taiwan. The study was conducted in the Department of Otolaryngology Head and Neck Surgery with the written and informed consent of the enrolled participants. All specimen and pathological reports were collected at Chang Gung Memorial Hospital for analysis. A total of 36 normal and 44 tumor cryopreserved tissue samples with histologically-proven malignancies and normal oral mucosa were collected, as shown in Table 1. The sub-sites were identified and recorded at the time of surgery and specimen acquisition. Normal tissues were taken from a site adjacent to the tumor. The study period was from February 2017 to December 2018. A total of 80 tissue samples were collected (from the tongue, buccal mucosa, and gingiva), which included 36 normal and 36 tumor samples from the same patients and eight tumor samples from different patients. All samples were at least 3×3 mm in size. Surgical resection specimens from normally appearing mucosa adjacent to the tumor were taken 15–30 min following surgical excision whereas the tumor samples were obtained immediately after surgery. The distance from tumor border to the adjacent tissue or resection margin was 1.5 to 2 cm. The tested cryopreserved samples were freshly cut and kept in liquid nitrogen (N_2) at $-80\ ^\circ$C to prevent alteration to its morphology until its use. Each tissue was

analyzed by RS. Glass was used as the substrate to test each tissue sample because it is the most widely available substrate and yields good results at lower source wavelength [29]. A total of 44 samples of tumor cells and 36 normal cells were obtained. Five spectra of each tissue sample were obtained at different locations. Since five spectra were obtained from each tissue sample, 220 and 180 spectra of tumor and normal tissues were respectively obtained, yielding 400 spectra for analysis by point-wise approach and 80 by patient-wise approach (five spectra per tissue averaged to yield one spectrum: $80 = 400/5$).

Table 1. Description of all tested cryopreserved tissue samples under Raman spectroscopy.

Sub-Sites	Tongue	Buccal Mucosa	Gingiva	Total
Tumor	13	19	12	44
Normal	11	14	11	36

2.2. Pre-Processing and Data Analysis

Data were processed and analyzed in MATLAB (R2018a, MathWorks, USA). First, a Savitsky–Golay filter (with order = 3) was used to smooth the recorded spectra so as to remove interference. After baseline correction, the area under the curve (AUC) technique was used to normalize all spectra to eliminate the data redundancy. The AUC is a function in MATLAB software. It normalizes a group of spectra with peaks by standardizing the area under the curve to the group median. Each value in a sample is divided by the sum over the sample. Unsupervised PCA was applied to the normalized spectrum from 700 to 2000 cm^{-1}. Normalized spectra were fed to multivariate the supervised classifier models PCA-LDA and PCA-QDA.

The LDA and QDA classifier models were used to study the boundary between classes and probabilities of classification. These models maximize the ratio of "between class variance" to "within class variance". This results in diminishing the data variation in the same class and detachment between classes. The LDA classifier assumes a common co-variance matrix and generates a linear boundary, while the QDA classifier assumes that each class has its own co-variance and produces a quadratic boundary. QDA optimally discriminates between the classes in the dataset [30], and requires large computation and data. Therefore, LDA is a good classifier for equal class samples and QDA is a good classifier for unequal class samples [31]. However, in some cases, they perform worse than expected [32]. To evaluate the classify results, the classifier models were optimized using a training dataset and their performance was evaluated using a test dataset.

2.3. Raman Spectroscopy (RS)

A RS instrument (ProTrusTech Co., Ltd., Taiwan) that comprised of a laser with a wavelength of 532 nm as an excitation source and a laser power of 126 mW was used. Spectral acquisition proceeded as follows. The excitation wavelength was 532 nm, the laser power was 6.3 mW~12.6 mW, the integration time was 5 s, the acquisition time was 15 s, and the average value of spectrum was 3 (meaning that the display spectrum averaged from three scanning spectra). The spectra resolution, specified by the manufacturer, was 1 cm^{-1}. The laser spot size was 6~8 micron and the penetration depth was 10~20 micron.

2.4. Multivariate Analysis

The mean normalized spectrum was analyzed using two supervised classifier models, PCA-LDA and PCA-QDA. PCA is a statistical procedure that reduces the number of dimensions and provides principal components (co-ordinates) based on new dimensions. The number of PCA components was less than half of the minimal sample classes to avoid over-fitting [25]. The first three principal components (PC1, PC2, and PC3) accounted for up to 97% variance, as evaluated by PCA. They were fed into the LDA and QDA classifiers. For PCA-LDA and PCA-QDA, scores of factor 1, 2, and 3

were chosen to obtain a three-dimensional scatter plot with a decision boundary. The analysis broadly categorized normal and tumor tissues under point-wise and patient-wise methods. In the point-wise approach, five spectra of a sample were analyzed. Owing to the heterogeneity of the tissue, the measured spectra at the various point varied greatly in intensity and Raman shift. In the patient-wise approach, one average spectrum of each sample was analyzed to eliminate heterogeneity.

3. Results and Discussion

A total of 80 tissue samples from three sub-sites were analyzed. The data were to be distinguished in the following two-class systems: Normal versus tumor, tongue tumor versus normal, buccal mucosa tumor versus normal, and gingiva tumor versus normal. The spectral features, vibrational molecules, and analysis of the three sub-sites will be described below.

3.1. Finger Print Region

Figure 1 presents normalized mean spectra of healthy or normal and tumor or OSCC tissues of the oral mucosa. The fingerprint region (700 cm^{-1} to 1800 cm^{-1}) in biological tissues is rich in proteins, nucleic acids, amino acid, carbohydrates, and lipids. The literature has shown that the normal tissue spectral peaks are lipid-dominated peaks while the malignant tissue peaks are protein-dominated peaks [14,15,19,33,34]. Malignant or tumor samples had higher peaks than normal tissues. Peaks at 1004, 1156, 1339, 1450, 1523, and 1656 cm^{-1} dominated the spectra of normal tissues, whereas peaks at 754, 1064, 1168, and 1220 cm^{-1} dominated those of malignant or tumor tissue samples. The sharp and high peak at 1004 cm^{-1} is attributed to the symmetric ring breathing mode of phenylalanine, which is an amino acid and is observed in protein-enriched malignant tissue spectra [14]. A sharp and intense peak at 1155$\sim$56 cm^{-1} arose from the proteins and was dominated by the protein signal in the tumor tissues [27,34]. The peak at 1220$\sim$1240 cm^{-1} is associated with lipids and =CH bending. The high peak at 1449$\sim$50 cm^{-1} is associated with CH$_2$ bending, and is associated with a protein [33,34]. The peaks at 1339 cm^{-1} in the average tumor spectrum was associated with the adenine feature of nucleic acid [23]. At 1518$\sim$1524 cm^{-1}, a sharp and more intense peak is observed and is associated with the beta-carotene or porphyrin feature and was obtained in both normal and tumor samples [35]. The lower intensity from normal tissues may be a discharge by secretion. In one study [36], this peak was absent from the spectra of normal tissues. A broad and strong peak at 1650$\sim$1655 cm^{-1} is a characteristic of proteins in the alpha-helix structure of amide I, which yields a strong signal in the spectral of tumor tissues. In normal tissues, the small peak at 1655 cm^{-1} is generated by the $C = C$ bond in lipids or phospholipids, and not amide I [14,34]. Normal tissues yielded a small peak at 1123 cm^{-1}, which is attributable to the $C - C$ skeletal stretch in lipids, while tumor tissues had a $(C - N)$ stretching mode of protein. Tumor tissues yielded a high peak at 750 cm^{-1} and a small peak at 823 cm^{-1} due to the Tryptophan and Tyrosine in protein, respectively [35]. Normal tissues yielded Raman peaks at 754, 1064, 1168, and 1220–84 cm^{-1} ($= CH$ bending) that are associated with the lipid. All of the above peaks were obtained from both normal and tumor tissues, with the strongest signals at 750, 1004, 1155, 1449, 1522, and 1656 cm^{-1}. In tumor tissue samples, protein, amide I, greater CH$_2$ bending, amide III, and amino acid (Tryptophan or phenylalanine) yielded signals that enabled such tissue to be distinguished from normal tissues.

Figure 2a–c present the normalized mean Raman spectra of tumor and normal samples from three sub-sites (buccal, gingiva, and tongue). Tumor/normal samples at all sub-sites yielded almost identical mean Raman spectra, but the three sub-sites did not yield the same intensity in corresponding regions. According to one study [37], different sub-site of oral mucosa (tongue, buccal, gingiva, hard palate, and soft palate) have different percentages of collagen and elastin. Carvalho et al. [38] demonstrated the biochemistry associated with healthy oral tissues at each sub-site and differentiated them using the basis of specific biochemical components. Figure 2 reveals that amide I and amide III bands at 1655 and 1250 cm^{-1}, respectively, were more prominent at the buccal tumor sub-site than at the gingiva and tongue tumor sub-sites. Protein/lipids bands at 1155 and 1523 cm^{-1} were more intense at the tongue

and gingival sub-site than at the buccal sub-site. All three sub-sites were known to vary with respect to prognosis, metastasis to lymph nodes, aggressiveness, and overall survival rate. These different genetic alterations and biological differences will be responsible for the basis of classification among buccal, tongue, and gingiva cancer [39–41].
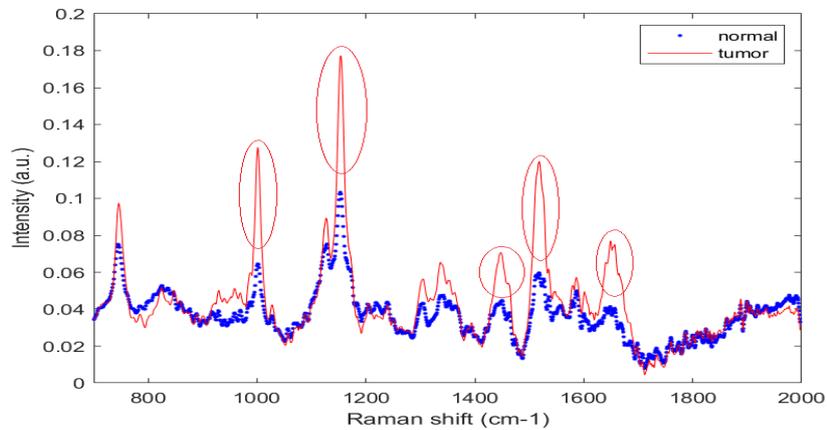

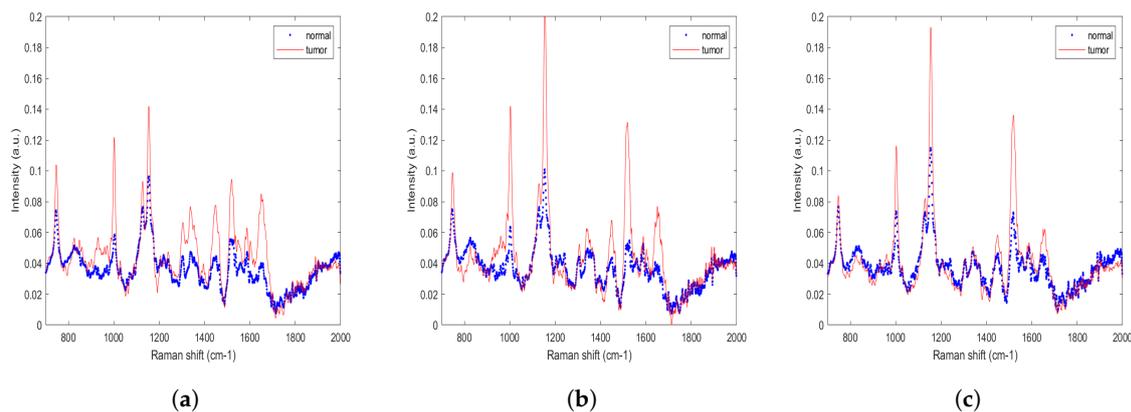**Figure 1.** Mean spectra of oral normal and tumor cryopreserved tissues.

Figure 2. Mean spectra of: (**a**) Buccal mucosa, (**b**) gingiva, and (**c**) tongue.

3.2. Analysis of Normal and Tumor Sample Data

The confusion tables for PCA-LDA and PCA-QDA classifiers were generated. The performance parameters were calculated from the confusion tables (correct and incorrect predictions). Point-wise and patient-wise approaches were analyzed using PCA-LDA and PCA-QDA classification models. To evaluate their performance, their accuracy, sensitivity, and specificity in identifying normal and tumor tissues were calculated. All of the spectra were subjected to PCA. Three PCA components were used for classification using both LDA and QDA models.

Table 2 shows the confusion and performance tables of normal and tumor sample data, analyzed using the PCA-LDA and PCA-QDA model for the point-wise approach. The PCA-LDA model correctly classified 177/220 and 121/180 tumor and normal sample data, respectively. However, the PCA-QDA model correctly classified 184/220 and 143/180 tumor and normal sample data, respectively. The accuracy, sensitivity, and specificity of the PCA-LDA model in differentiating normal and tumor tissues were 74.5%, 80.45%, and 67.22%, respectively. However, the accuracy, sensitivity, and specificity of the PCA-QDA model in distinguishing normal and tumor tissues were 81.75%, 83.63%, and 79.44%, respectively. Figure 3a,b plot the 3D decision boundary curves for normal and tumor sample data using point-wise approach for the PCA-LDA and the PCA-QDA classifier

model, respectively. The decision boundary classified the tumor and normal sample data. The solid red and blue dots represent tumor and normal classes, respectively. Table 3 shows the confusion and performance tables of normal and tumor sample data, analyzed using the PCA-LDA and PCA-QDA model for the patient-wise approach. The PCA-LDA model correctly classified 34/44 and 31/36 tumor and normal sample data, respectively. The PCA-QDA model correctly classified 40/44 and 30/36 tumor and normal sample data, respectively. The accuracy, sensitivity, and specificity of the PCA-LDA model in differentiating normal and tumor tissues were 81.25%, 77.27%, and 86.11%, respectively and those of PCA-QDA model were 87.5%, 90.90%, and 83.33%, respectively. The PCA-QDA model therefore exhibited a better classification performance than the PCA-LDA model. Figure 4a,b plot the 3D decision boundary curve of the patient-wise approach for the PCA-LDA and the PCA-QDA classifier model, respectively. The patient-wise approach is seen as a better classifier than the point-wise approach. Therefore, only the patient-wise approach was used in the following three sub-site analyses.

Table 2. Confusion and performance tables for point-wise approach.

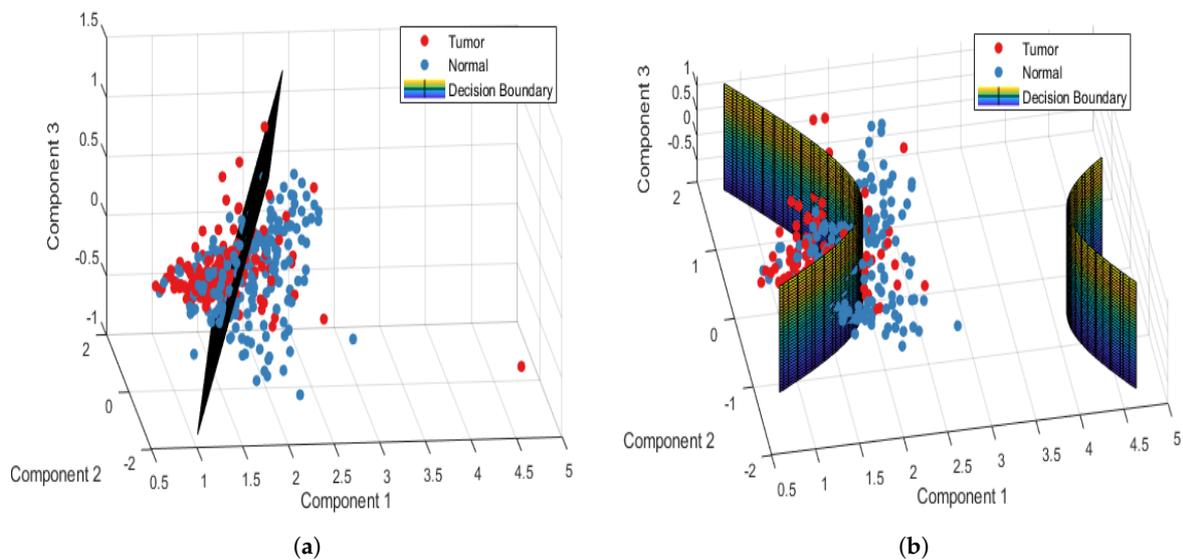| Dataset	Confusion Table			Performance Parameters		
PCA-LDA	Tumor	Normal	Total	Accuracy (%)	Sensitivity (%)	Specificity (%)
Tumor	177	43	220	74.50	80.45	67.22
Normal	59	121	180			
PCA-QDA	Tumor	Normal	Total	Accuracy (%)	Sensitivity (%)	Specificity (%)
Tumor	184	36	220	81.75	83.63	79.44
Normal	37	143	180			

(a)

(b)

Figure 3. Point-wise 3D decision boundary curve for (**a**) PCA-LDA and (**b**) PCA- QDA classifier model.

Table 3. Confusion and performance tables for patient-wise approach.

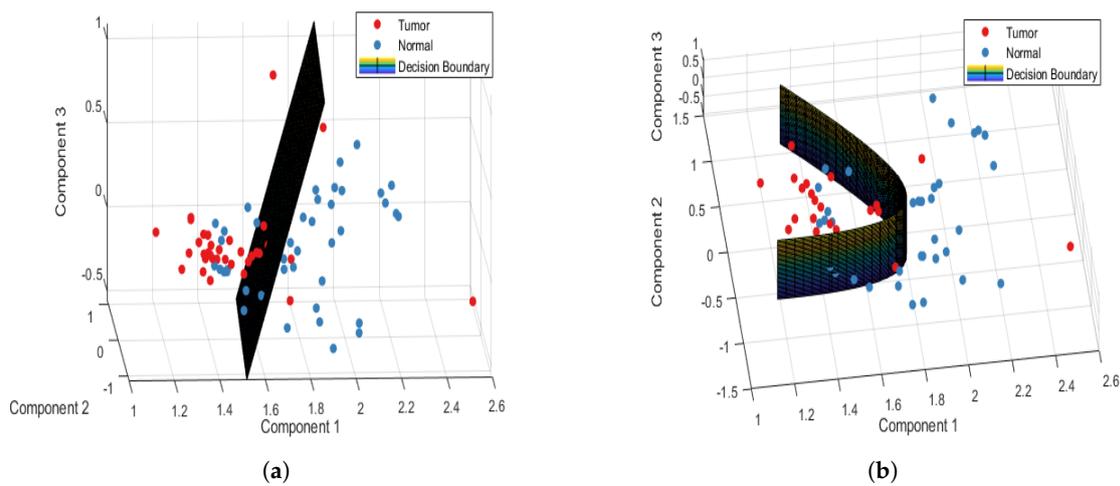| Dataset	Confusion Table			Performance Parameters		
PCA-LDA	Tumor	Normal	Total	Accuracy (%)	Sensitivity (%)	Specificity (%)
Tumor	34	10	44	81.25	77.27	86.11
Normal	5	31	36			
PCA-QDA	Tumor	Normal	Total	Accuracy (%)	Sensitivity (%)	Specificity (%)
Tumor	40	4	44	87.50	90.90	83.33
Normal	6	30	36			

(**a**) (**b**)

Figure 4. Patient-wise 3D decision boundary curve for (**a**) PCA-LDA and (**b**) PCA-QDA classifier model.

3.3. Analysis of Data from Normal and Tumor Samples from Tongue, Buccal Mucosa, and Gingiva Sub-Sites

Spectra of 13 tumor samples and 11 normal samples at the tongue sub-site were collected. Table 4 shows the performance table of the patient-wise approach at the tongue sub-site. The PCA-LDA model differentiates tumor and normal tissues had an accuracy, sensitivity, and specificity of 79.16%, 92.3%, and 63.63%, respectively, and the PCA-QDA model did so with corresponding values of 87.5%, 100%, and 72.72%, respectively. Figure 5a,b plot the 3D decision boundary curve of the PCA-LDA and PCA-QDA classifier models with the tongue sample data, respectively. In Table 4, specificity quantifies the extent to which persons without a disease undesirably screen as positive. The accuracy (1-error rate) is the proportion of correct predictions, including correct positive and negative predications based on the selected samples [42]. The low specificity of the tongue indicates miss-classification between healthy tissue sublayers (surface squamous epithelium, muscle, and gland) and the OSCC structure, which has been discussed in an earlier study [43] and the section on validation methods.

Table 4. Performance table of patient-wise tongue analysis.

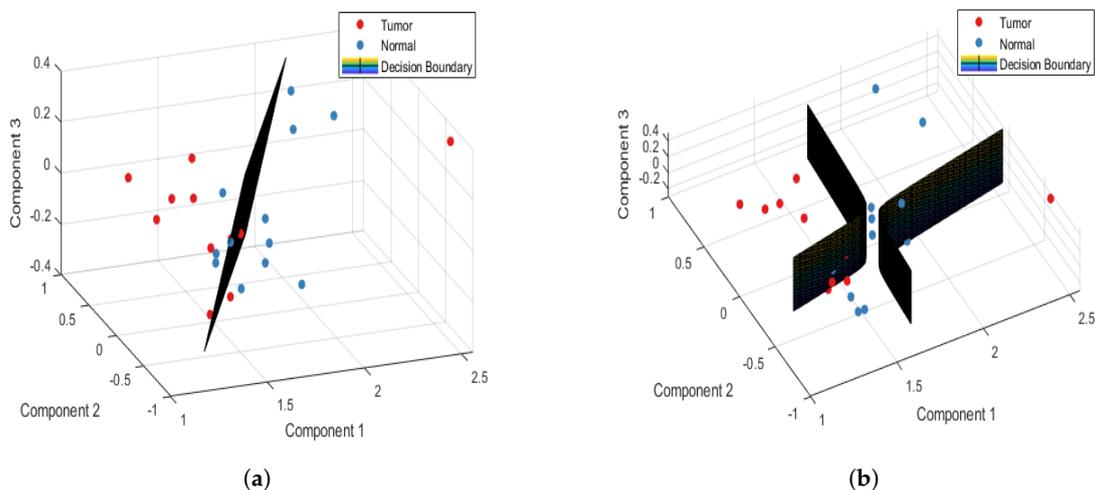| Patient-Wise: Tongue	Accuracy (%)	Sensitivity (%)	Specificity (%)
PCA-LDA	79.16	92.30	63.63
PCA-QDA	87.50	100.00	72.72

(**a**) (**b**)

Figure 5. Tongue patient-wise 3D decision boundary curve for (**a**) PCA-LDA and (**b**) PCA-QDA classifier model.

At the buccal mucosa sub-site, 19 tumor samples and 14 normal samples were collected. Table 5 shows the performance table of the patient-wise approach at the buccal mucosa sub-site. The PCA-LDA model differentiates the buccal mucosa tumor and normal tissues with an accuracy, sensitivity, and specificity of 84.84%, 78.94%, and 92.85%, respectively, and the PCA-QDA model differentiates them with 87.87%, 84.21%, and 92.85%, respectively. Figure 6a,b plot the 3D decision boundary curve of the PCA-LDA and PCA-QDA classifier models, respectively with buccal mucosa sample data. For the PCA-QDA model, the increasing sensitivity results with increasing accuracy are because the true positive cases are more correctly identified in this model. The PCA-QDA model outperformed the PCA-LDA model.

Table 5. Performance table of patient-wise buccal mucosa analysis.

Patient-Wise: Buccal	Accuracy (%)	Sensitivity (%)	Specificity (%)
PCA-LDA	84.84	78.94	92.85
PCA-QDA	87.87	84.21	92.85

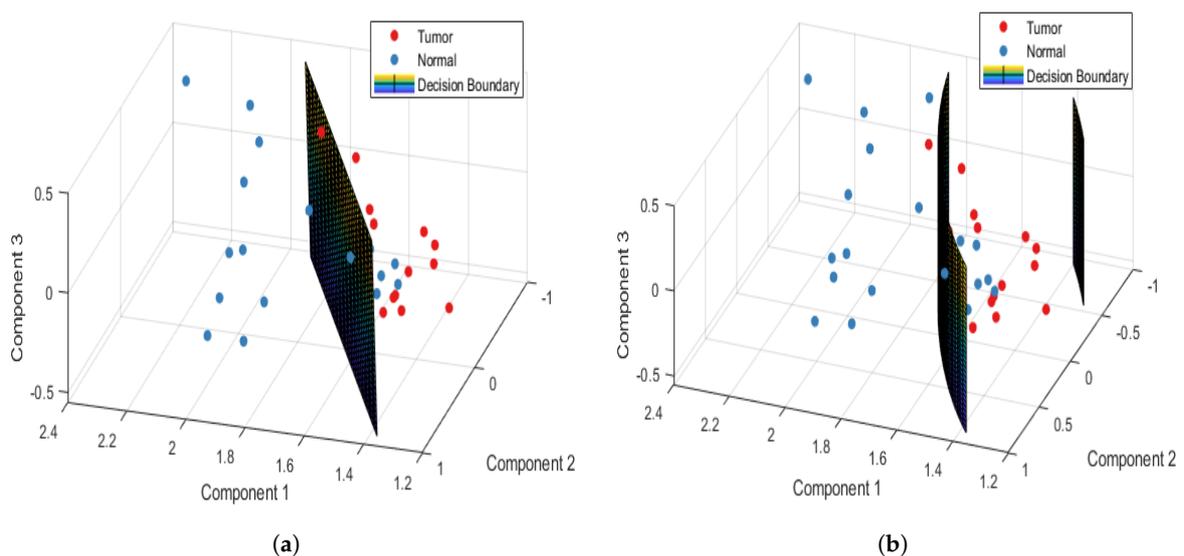(**a**)   (**b**)

Figure 6. Buccal mucosa patient-wise 3D decision boundary curve for (**a**) PCA-LDA and (**b**) PCA-QDA classifier model.

From the gingiva sub-site, 12 tumor samples and 11 normal samples were collected. Table 6 shows the performance table of the patient-wise approach at the gingiva sub-site. The PCA-LDA model differentiates gingiva tumor from normal tissue with an accuracy, sensitivity, and specificity of 91.30%, 91.66% and 90.90%, respectively, and the PCA-QDA model differentiates them with 87.12%, 75%, and 100%, respectively. Figure 7a,b plot the 3D decision boundary curve of the PCA-LDA and the PCA-QDA classifier models with the gingiva sample, respectively. The PCA-LDA model assumed linearity and variance-covariance homogeneity, whereas the PCA-QDA model had different feature covariance matrices for different classes and the consistency of a PCA-QDA model could not be predicted from a few samples, reducing the accuracy of the PCA-QDA model.

Table 6. Performance table of patient-wise gingiva analysis.

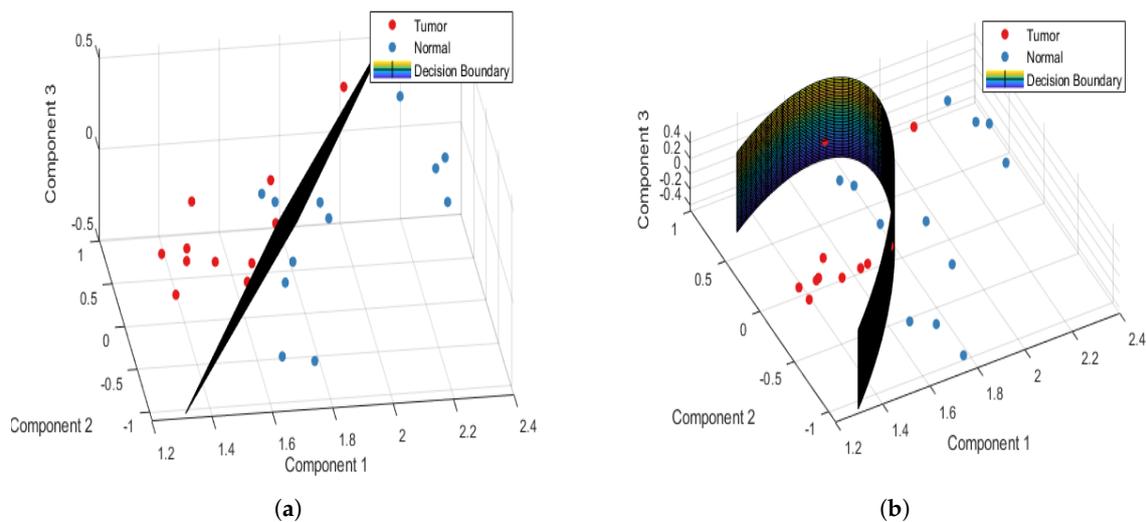| Patient-Wise: Gingiva	Accuracy (%)	Sensitivity (%)	Specificity (%)
PCA-LDA	91.30	91.66	90.90
PCA-QDA	87.12	75.00	100.00

Figure 7. Gingiva patient-wise 3D decision boundary curve for (**a**) PCA-LDA and (**b**) PCA-QDA classifier model.

4. Validation Methods

Cross-validation is a process by which the performance of a model is estimated using a limited number of data sample. It estimates the effectiveness of machine learning models (PCA-LDA and PCA-QDA models herein) with unseen data. The sample dataset is randomly partitioned into two disjoint subsets (training and validation data sets). The validation dataset is used to evaluate the performance of the model [44]. The cross-validation methods herein were k-fold and leave-one-out-cross-validation (LOOCV). In the k-fold method, the parameter K is the number of groups. A given dataset is randomly split into K equal subsets. Each subset is called as a fold and is unique. One fold was used as the test data set and the other K-1 folds were used as a training data set. The machine learning classification model was trained using the training data set and the performance parameters were evaluated. This process was iterated for the K folds and the performance parameters were aggregated. In the LOOCV method, a single sample within a given dataset was used as the test data and the other samples within a given dataset were used as training data. This process was iterated until each sample in a given dataset was used once as the test data. Therefore, each sample held out from the training data set. This method requires a large computation time because many iterations are required for training. The LOOCV method aggregates the estimated error rate by the number of samples in a given dataset. The k-fold method repeatedly randomizes a sub-sampling that can be used for training and testing data set in all the samples. However, the LOOCV method does not require a random process. Hence, in the LOOCV method, the estimation has less in bias but high variance [45]. However, in the k-fold method, the reduction of variance increases the value of K. Therefore, the bias remains low.

Table 7 shows estimates of the error rates of the model for point-wise and patient-wise approaches. The error rates of the PCA-LDA and PCA-QDA models for the point-wise approach were 25.5% and 16.27%, respectively. However, the error rates of the PCA-LDA and PCA-QDA models for patient-wise approach were only 18.75% and 12.5%, respectively. These results were confirmed by the k-fold and LOOCV methods, which yielded error rates of 18.25% and 17% for the point-wise approach, and 16.25% and 11.25% for the patient-wise approach, respectively. These results support the conclusion that the patient-wise approach is better than the point-wise approach. Therefore, only the patient wise approach was used in the following validation.

Table 7. Error Rate of PCA-LDA, PCA-QDA, and validation methods for normal versus tumor.

Error Rate	PCA-LDA (%)	PCA-QDA (%)	Validation: K-fold (%)	Validation: LOOCV (%)
Point-wise	25.50	16.27	18.25	17.00
Patient-wise	18.75	12.50	16.25	11.25

Table 8 provides estimates of the models' error rates at the tongue, buccal mucosa, and gingiva sub-sites. At the tongue sub-sites, both PCA-LDA and PCA-QDA models had low specificity because the validation methods generated higher error rates than other sub-sites. Since the oral tissue is heterogeneous, it comprises of different structures and layers. Cals [43] reported that the sub-layers of a healthy tissue structure (surface squamous epithelium, muscle) have the same protein, lipids, and nucleic acid as OSCC or tumor tissues of the tongue. Therefore, the miss-classification between healthy tissues sublayers (surface squamous epithelium, muscle, and gland) and the OSCC structures was greater than other sub-sites. In earlier studies [27,28], the same miss-classification observed between OSCC and surface squamous epithelium was observed owing to the low specificity of the classification model when applied to tongue tissues. The PCA-LDA and PCA-QDA models reveal that most of the biomolecular information from tissues and cells are critical in discriminating tumorous tissues from healthy tissues. This diagnostic model can also differentiate subgroups using the different components of Raman biochemical and biomolecular features and thus sub-site oral cancer can be distinguished from normal tissue.

Table 8. Error Rate of PCA-LDA, PCA-QDA, and validation methods for each sub-site.

Error Rate	PCA-LDA (%)	PCA-QDA (%)	Validation: K-fold (%)	Validation: LOOCV (%)
Tongue	20.83	12.50	16.67	16.67
Buccal	15.16	12.13	18.18	21.21
Gingiva	8.60	12.88	13.04	19.04

The limitations of this work include the limited number of samples at each sub-site (tongue: 24, buccal mucosa: 33, and gingiva: 23). Our future investigations will target a maximum number of samples for each sub-site in the oral cavity to enhance the classification rate and use other approaches that involve meta-learning and neural networks for classification.

5. Conclusions

This work studied the application of RS to oral cryopreserved freshly-excised tissue samples. This method had several advantages, including its rapidity, lack of need for labeling, and inexpensiveness. It has the potential to improve the efficiency of screening procedures for oral cancers and to identify the boundary for tumor-free resection margin during surgery. The PCA-QDA model for the patient-wise approach had greater classification efficiency than the PCA-LDA model. In the future, we will develop artificial intelligence algorithms to classify data and reduce the error rate.

Author Contributions: Conceptualization, M.-J.J., L.-B.C. and M.S.; Methodology, M.S. and M.-J.J.; Software, L.S. and M.S.; Validation, M.S., L.S. and S.-L.W.; Formal analysis, M.S.; Investigation, M.-J.J., S.-F.H. and L.-B.C.; Project administration, L.-B.C., S.-F.H., M.-J.J. and S.-L.W.; Resources, S.-F.H. and L.-B.C.; Data curation, M.S. and T.Y.-C.; Writing–original draft preparation, M.S.; Writing–review and editing, M.-J.J., M.S., L.S. and S.F.-H.; Visualization, M.-J.J., M.S.; Supervision, S.-L.W., L.S. and L.C.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Reddy, S.S.; Sharma, S.; Mysorekar, V. Expression of Epstein–barr virus among oral potentially malignant disorders and oral squamous cell carcinomas in the South Indian tobacco-chewing population. *J. Oral Pathol. Med.* **2017**, *46*, 454–459. [CrossRef] [PubMed]

2. Kumar, R.; Samal, S.K.; Routray, S.; Dash, R.; Dixit, A. Identification of oral cancer related candidate genes by integrating protein-protein interactions, gene ontology, pathway analysis and immunohistochemistry. *Sci. Rep.* **2017**, *7*, 2472. [CrossRef] [PubMed]

3. Stewart, B.; Wild, C.P. World Cancer Report 2014. Available online: https://www.drugsandalcohol.ie/28525/1/World%20Cancer%20Report.pdf (accessed on 23 August 2019).

4. Allen, C.T.; Law, J.H.; Dunn, G.P.; Uppaluri, R. Emerging insights into head and neck cancer metastasis. *Head Neck* **2013**, *35*, 1669–1678. [CrossRef] [PubMed]

5. Mortazavi, H.; Baharvand, M.; Mehdipour, M. Oral potentially malignant disorders: An overview of more than 20 entities. *J. Dent. Res. Dent. Clin. Dent. Prospect.* **2014**, *8*, 6. [CrossRef]

6. Liu, D.; Zhao, X.; Zeng, X.; Dan, H.; Chen, Q. Non-invasive techniques for detection and diagnosis of oral potentially malignant disorders. *Tohoku J. Exp. Med.* **2016**, *238*, 165–177. [CrossRef] [PubMed]

7. Cordero, E.; Latka, I.; Matthäus, C.; Schie, I.W.; Popp, J. In-vivo Raman spectroscopy: From basics to applications. *J. Biomed. Opt.* **2018**, *23*, 071210. [CrossRef] [PubMed]

8. Vašková, H. A powerful tool for material identification: Raman spectroscopy. *Int. J. Math. Model. Methods Appl. Sci.* **2011**, *5*, 1205–1212.

9. Pence, I.; Mahadevan-Jansen, A. Clinical instrumentation and applications of Raman spectroscopy. *Chem. Soc. Rev.* **2016**, *45*, 1958–1979. [CrossRef]

10. Cui, S.; Zhang, S.; Yue, S. Raman Spectroscopy and Imaging for Cancer Diagnosis. *J. Healthc. Eng.* **2018**, *2018*. [CrossRef]

11. Notingher, I. Raman spectroscopy cell-based biosensors. *Sensors* **2007**, *7*, 1343–1358. [CrossRef]

12. Bakker Schut, T.; Witjes, M.; Sterenborg, H.; Speelman, O.; Roodenburg, J.; Marple, E.; Bruining, H.; Puppels, G. In vivo detection of dysplastic tissue by Raman spectroscopy. *Anal. Chem.* **2000**, *72*, 6010–6018. [CrossRef]

13. Sunder, N.S.; Rao, N.N.; Kartha, V.; Ullas, G.; Kurien, J. Laser Raman spectroscopy: A novel diagnostic tool for oral cancer. *J. Orofac. Sci.* **2011**, *3*, 15–19.

14. Malini, R.; Venkatakrishna, K.; Kurien, J.M.; Pai, K.; Rao, L.; Kartha, V.; Krishna, C.M. Discrimination of normal, inflammatory, premalignant, and malignant oral tissue: A Raman spectroscopy study. *Biopolym. Orig. Res. Biomol.* **2006**, *81*, 179–193. [CrossRef]

15. Singh, S.; Deshmukh, A.; Chaturvedi, P.; Krishna, C.M. In vivo Raman spectroscopic identification of premalignant lesions in oral buccal mucosa. *J. Biomed. Opt.* **2012**, *17*, 105002. [CrossRef]

16. Dai, W.Y.; Lee, S.; Hsu, Y.C. Discrimination between oral cancer and healthy cells based on the adenine signature detected by using Raman spectroscopy. *J. Raman Spectrosc.* **2018**, *49*, 336–342. [CrossRef]

17. Chen, P.H.; Shimada, R.; Yabumoto, S.; Okajima, H.; Ando, M.; Chang, C.T.; Lee, L.T.; Wong, Y.K.; Chiou, A.; Hamaguchi, H.O. Automatic and objective oral cancer diagnosis by Raman spectroscopic detection of keratin with multivariate curve resolution analysis. *Sci. Rep.* **2016**, *6*, 20097. [CrossRef]

18. Cals, F.L.; Bakker Schut, T.C.; Koljenović, S.; Puppels, G.J.; de Jong, R.J.B. Method development: Raman spectroscopy-based histopathology of oral mucosa. *J. Raman Spectrosc.* **2013**, *44*, 963–972. [CrossRef]

19. Guze, K.; Pawluk, H.C.; Short, M.; Zeng, H.; Lorch, J.; Norris, C.; Sonis, S. Pilot study: Raman spectroscopy in differentiating premalignant and malignant oral lesions from normal mucosa and benign lesions in humans. *Head Neck* **2015**, *37*, 511–517. [CrossRef]

20. Knipfer, C.; Motz, J.; Adler, W.; Brunner, K.; Gebrekidan, M.T.; Hankel, R.; Agaimy, A.; Will, S.; Braeuer, A.; Neukam, F.W.; et al. Raman difference spectroscopy: A non-invasive method for identification of oral squamous cell carcinoma. *Biomed. Opt. Express* **2014**, *5*, 3252–3265. [CrossRef]

21. Lau, D.P.; Huang, Z.; Lui, H.; Man, C.S.; Berean, K.; Morrison, M.D.; Zeng, H. Raman spectroscopy for optical diagnosis in normal and cancerous tissue of the nasopharynx—Preliminary findings. *Lasers Surg. Med.* **2003**, *32*, 210–214. [CrossRef]

22. Lau, D.P.; Huang, Z.; Lui, H.; Anderson, D.W.; Berean, K.; Morrison, M.D.; Shen, L.; Zeng, H. Raman spectroscopy for optical diagnosis in the larynx: Preliminary findings. *Lasers Surg. Med. Off. J. Am. Soc. Laser Med. Surg.* **2005**, *37*, 192–200. [CrossRef]

23. Sahu, A.; Tawde, S.; Pai, V.; Gera, P.; Chaturvedi, P.; Nair, S.; Krishna, C.M. Raman spectroscopy and cytopathology of oral exfoliated cells for oral cancer diagnosis. *Anal. Methods* **2015**, *7*, 7548–7559. [CrossRef]

24. Sahu, A.; Sawant, S.; Mamgain, H.; Krishna, C.M. Raman spectroscopy of serum: An exploratory study for detection of oral cancers. *Analyst* **2013**, *138*, 4161–4174. [CrossRef]

25. Sahu, A.; Sawant, S.; Talathi-Desai, S.; Murali Krishna, C. Raman spectroscopy of serum: A study on oral cancers. *Biomed. Spectrosc. Imaging* **2015**, *4*, 171–187.

26. Barroso, E.; Smits, R.; Bakker Schut, T.; Ten Hove, I.; Hardillo, J.; Wolvius, E.; Baatenburg de Jong, R.; Koljenovic, S.; Puppels, G. Discrimination between oral cancer and healthy tissue based on water content determined by Raman spectroscopy. *Anal. Chem.* **2015**, *87*, 2419–2426. [CrossRef]

27. Cals, F.L.; Schut, T.C.B.; Hardillo, J.A.; De Jong, R.J.B.; Koljenović, S.; Puppels, G.J. Investigation of the potential of Raman spectroscopy for oral cancer detection in surgical margins. *Lab. Investig.* **2015**, *95*, 1186–1196. [CrossRef]

28. Cals, F.L.; Koljenović, S.; Hardillo, J.A.; de Jong, R.J.B.; Schut, T.C.B.; Puppels, G.J. Development and validation of Raman spectroscopic classification models to discriminate tongue squamous cell carcinoma from non-tumorous tissue. *Oral Oncol.* **2016**, *60*, 41–47. [CrossRef]

29. Kerr, L.T.; Byrne, H.J.; Hennelly, B.M. Optimal choice of sample substrate and laser wavelength for Raman spectroscopic analysis of biological specimen. *Anal. Methods* **2015**, *7*, 5041–5052. [CrossRef]

30. Siqueira, L.F.; Júnior, R.F.A.; de Araújo, A.A.; Morais, C.L.; Lima, K.M. LDA vs. QDA for FT-MIR prostate cancer tissue classification. *Chemom. Intell. Lab. Syst.* **2017**, *162*, 123–129. [CrossRef]

31. Xue, J.H.; Titterington, D.M. Do unbalanced data have a negative effect on LDA? *Pattern Recog.* **2008**, *41*, 1558–1571. [CrossRef]

32. Eisenbeis, R.A. Pitfalls in the application of discriminant analysis in business, finance, and economics. *J. Financ.* **1977**, *32*, 875–900. [CrossRef]

33. Parker, F.S. *Applications of Infrared, Raman, and Resonance Raman Spectroscopy in Biochemistry*; Springer Science & Business Media: Berlin, Germany, 1983.

34. Movasaghi, Z.; Rehman, S.; Rehman, I.U. Raman spectroscopy of biological tissues. *Appl. Spectrosc. Rev.* **2007**, *42*, 493–541. [CrossRef]

35. Huang, Z.; McWilliams, A.; Lui, H.; McLean, D.I.; Lam, S.; Zeng, H. Near-infrared Raman spectroscopy for optical diagnosis of lung cancer. *Int. J. Cancer* **2003**, *107*, 1047–1052. [CrossRef]

36. Mahadevan-Jansen, A.; Richards-Kortum, R. Raman spectroscopy for cancer detection: A review. In Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Magnificent Milestones and Emerging Opportunities in Medical Engineering (Cat. No. 97CH36136), Chicago, IL, USA, 30 October–2 November 1997; Volume 6, pp. 2722–2728.

37. Ciano, J.; Beatty, B.L. Regional quantitative histological variations in human oral mucosa. *Anat. Rec.* **2015**, *298*, 562–578. [CrossRef]

38. Carvalho, L.F.C.; Nogueira, M.S.; Bhattacharjee, T.; Neto, L.P.; Daun, L.; Mendes, T.O.; Rajasekaran, R.; Chagas, M.; Martin, A.A.; Soares, L.E.S. In vivo Raman spectroscopic characteristics of different sites of the oral mucosa in healthy volunteers. *Clin. Oral Investig.* **2018**, *23*, 3021–3031. [CrossRef]

39. Liu, S.A.; Wang, C.C.; Jiang, R.S.; Lee, F.Y.; Lin, W.J.; Lin, J.C. Pathological features and their prognostic impacts on oral cavity cancer patients among different subsites–A singe institute's experience in Taiwan. *Sci. Rep.* **2017**, *7*, 7451. [CrossRef]

40. Freier, K.; Joos, S.; Flechtenmacher, C.; Devens, F.; Benner, A.; Bosch, F.X.; Lichter, P.; Hofele, C. Tissue microarray analysis reveals site-specific prevalence of oncogene amplifications in head and neck squamous cell carcinoma. *Cancer Res.* **2003**, *63*, 1179–1182.

41. Sathyan, K.; Sailasree, R.; Jayasurya, R.; Lakshminarayanan, K.; Abraham, T.; Nalinakumari, K.; Abraham, E.K.; Kannan, S. Carcinoma of tongue and the buccal mucosa represent different biological subentities of the oral carcinoma. *J. Cancer Res. Clin. Oncol.* **2006**, *132*, 601–609. [CrossRef]

42. Zhu, W.; Zeng, N.; Wang, N. Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. *NESUG Proc. Health Care Life Sci. Balt. Md.* **2010**, *19*, 67.

43. Cals, F.L.; Schut, T.B.; Caspers, P.; de Jong, R.B.; Koljenović, S.; Puppels, G.J. Raman spectroscopic analysis of the molecular composition of oral cavity squamous cell carcinoma and healthy tongue tissue. *Analyst* **2018**, *143*, 4090–4102. [CrossRef]

44. Wang, H.; Zheng, H. Model Validation, Machine Learning. In *Encyclopedia of Systems Biology*; Dubitzky, W., Wolkenhauer, O., Cho, K.H., Yokota, H., Eds.; Springer: New York, NY, USA, 2013; pp. 1406–1407. [CrossRef].

45. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer: New York, NY, USA, 2013; Volume 112.