Electronic Supplementary Material:

Supplementary Table 1: Clinical Diagnostic Features of oral lesions to be included in the study

Disease	Diagnostic Features Used for Biopsy Recommendation	Reference
OLK	• white single localized, multiple, or diffuse widespread lesions with ○ surface nodularity ○ erythema ○ ulceration ○ increased firmness and induration ○ • unexplained hemorrhage	[1, 2]
OSF	• White Patch • Trismus • Soreness of mucosa • Increased sensitivity to chilies • Nodule	[3]
OSCC	Oral: • Any solitary lump, ulcer, white or red lesion retaining for more than three weeks or non-healing socket • Numbness • Unexplained loose tooth' Extraoral: • Cervical lymphadenopathy may be detectable. • Synchronous and metachronous second primary tumors may be found in the upper aerodigestive tract (pharynx, larynx, and esophagus).	[1, 2]

Supplementary Table 2: Details of the intensity features used to differentiate OCT images [4, 5]

Feature	Mathematical Formula	Description
$Mean_{Gray}\,(\mu)$	$\frac{\sum_{k=1}^{N} I_k}{N}$	N is the number of pixels and I_k is the k^{th} intensity level of the image.
$Median_{Gray}$ (m)	$L + \frac{\left\lfloor \frac{N+1}{2} \right\rfloor - F}{f_m}$	Computes the middle value of the pixel intensities. This formula computes the median for ungrouped data. L : lower boundary of the median class, F : cumulative frequency of the classes before the median class, f_m : frequency of the median class, C : class size. For ungrouped sorted data, median is the value of the $\left\lfloor \frac{N+1}{2} \right\rfloor^{th}$ data point.

$Variance_{Gray}(\sigma^2)$	$\frac{\sum_{k=1}^{N}(I_k-\mu)^2}{N}$	The average of the squared differences of the intensities from the mean intensity.
$StandardDeviation_{Gray}(\sigma)$	$\sqrt{\sigma^2}$	A measure of how spread out the intensities are.
$CoefficientofVariance_{Gray}(C_v)$	$\frac{\sigma}{\mu}$	The ratio of the standard deviation to the mean.
$Entropy_{Gray}$	$-\sum_{j=0}^{L-1}p_j\log_2 p_j$	A measure to describe the busyness of the image. L : number of pixel intensity levels in the image, p_j : probability of occurrence of a pixel with intensity value j .
$Skewness_{Gray}$	$\frac{1}{\sigma^3}\frac{\sum_{k=1}^{N}(I_k-\mu)^3}{N}$	A measure of symmetry, or more precisely, the lack of symmetry.
$Kurtosis_{Gray}$	$(\frac{1}{\sigma^4}\frac{\sum_{k=1}^{N}(I_k-\mu)^4}{N})-3$	A measure of any peakedness of the distribution of the data

Supplementary Table 3: Details of the textural features used to differentiate OCT images[4, 6]

Feature	Mathematical Formula	Description
Co-occurrence Probability (C_{ij})	$\frac{P_{ij}}{\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}P_{ij}}$	P_{ij} is the number of co-occurrences of gray levels i and j .
$Inertia_{GLCM}$	$\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}(i-j)^2 C_{ij}$	A measure of the distribution of gray-scales in the image. A higher value indicates presence of higher magnitude elements away from the diagonal ($i \neq j$) in the GLCM.
$Correlation_{GLCM}$	$\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}C_{ij}(\frac{(i-\mu_i)(j-\mu_j)}{\sigma_i\sigma_j})$	A measure of gray level linear dependence between the pixels at the specified positions relative to each other.
$Energy_{GLCM}$	$\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}C_{ij}^2$	High value of energy indicates orderliness in the image window.
$Entropy_{GLCM}$	$\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}C_{ij}\log_2(C_{ij}+\varepsilon)$	Higher entropy value indicates that an image is more homogenous. ε is an arbitrarily small constant equal to the floating point

		accuracy= 2^{-52}
$Homogeneity_{GLCM}$	$\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\frac{C_{ij}}{1+\|i-j\|}$	A high homogeneity measure indicates the presence of only a few gray levels in the image.
Cluster Shade	$\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}C_{ij}(i+j-\mu_i-\mu_j)^3$	A measure of skewness or asymmetry of the GLCM. High value indicates lack of symmetry. $\mu_i=\sum_{i=0}^{L-1}i\sum_{j=0}^{L-1}C_{ij}$; $\mu_j=\sum_{j=0}^{L-1}j\sum_{i=0}^{L-1}C_{ij}$.
Cluster Prominence	$\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}C_{ij}(i+j-\mu_i-\mu_j)^4$	A similar measure of skewness or asymmetry of the GLCM.A low value indicates a peak in the GLCM around the mean values thus implying little variation in the gray scales in the original image.
Information Measures of Correlation	$\sqrt{1-e^{-2(h_2-h_1)}}$	$h_1=Entropy_{GLCM}$; $h2=-\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}C_iC_j\log_2(C_iC_j+\varepsilon)$; $C_i=\sum_{i=0}^{L-1}C_{ij}, C_j=\sum_{j=0}^{L-1}C_{ij}$.
Maximum Probability	$max(C_{ij})$	Maximum value of the co-occurrence matrix elements.
Sum of Entropy(S_E)	$-\sum_{k=2}^{2L-2}p_{i+j}(k)log_2(p_{i+j}(k)+\varepsilon$	$p_{i+j}(k)=\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}C_{ij}$; $i+j=k;k=0,1,2,...,2(L-1)$.
Sum of Variance(S_{σ^2})	$-\sum_{k=2}^{2L-2}(k-S_E)^2p_{i+j}(k)$	A high value indicates equal concentration of the frequency of occurrence in the highest and lowest cells of the GLCM.
Difference Entropy(D_E)	$-\sum_{k=0}^{L-1}p_{\|i-j\|}(k)log_2(p_{\|i-j\|}(k)+\varepsilon$	$p_{i-j}(k)=\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}C_{ij}$; $\|i-j\|=k;k=0,1,2,...,L-1$.
$Mean_{LBP}(\mu_{LBP})$	$\frac{\sum_{k=1}^{M}ILBP}{M}$	Mean value of the intensity levels of the LBP (Local Binary Pattern) image. $ILBP_k$ is the k^{th} level intensity value of the LBP image. M is the number of intensity levels in the LBP image.

$StandardDeviation_{LBP}$	$\sqrt{\frac{\sum_{k=1}^{M}(ILBP_k - \mu_{LBP})^2}{M}}$	The average of the squared differences of the intensities from the mean intensity in the LBP image.

References:

1. B. W. Neville and T. A. Day, *CA: a cancer journal for clinicians*, 2002, **52**, 195-215.
2. C. Scully, O. P. De Almeida, J. Bagan, P. D. Dios and A. M. Taylor, *Oral medicine and pathology at a glance*, John Wiley & Sons, 2013.
3. P. N. Wahi, V. L. Kapur, U. K. Luthra and M. C. Srivastava, *Bulletin of the World Health Organization*, 1966, **35**, 789-792.
4. R. M. Haralick, *Proceedings of the IEEE*, 1979, **67**, 786-804.
5. B. Park and Y. Chen, *Journal of agricultural engineering research*, 2001, **78**, 127-139.
6. G. J. Ughi, T. Adriaenssens, P. Sinnaeve, W. Desmet and J. D'Hooge, *Biomedical optics express*, 2013, **4**, 1014-1030.