











# Oral Dysplasia Classification by Using Fractal Representation Images and Convolutional Neural Networks

Rafael H. O. Carvalho<sup>1</sup><sup>a</sup>, Adriano B. Silva<sup>1</sup><sup>b</sup>, Alessandro S. Martins<sup>2</sup><sup>c</sup>, Sérgio V. Cardoso<sup>3</sup><sup>d</sup>,  
Guilherme R. Freire<sup>4</sup><sup>e</sup>, Paulo R. de Faria<sup>5</sup><sup>f</sup>, Adriano M. Loyola<sup>3</sup><sup>g</sup>, Thaína A. A. Tosta<sup>6</sup><sup>h</sup>,  
Leandro A. Neves<sup>7</sup><sup>i</sup> and Marcelo Z. do Nascimento<sup>1</sup><sup>j</sup>

<sup>1</sup>Faculty of Computer Science, Federal University of Uberlândia, Brazil

<sup>2</sup>Federal Institute of Triângulo Mineiro, Brazil

<sup>3</sup>Area of Oral Pathology, School of Dentistry, Federal University of Uberlândia, Brazil

<sup>4</sup>Department of Informatics Engineering, Faculty of Engineering, University of Porto, Portugal

<sup>5</sup>Department of Histology and Morphology, Institute of Biomedical Science, Federal University of Uberlândia, Brazil

<sup>6</sup>Science and Technology Institute, Federal University of São Paulo, Brazil

<sup>7</sup>Department of Computer Science and Statistics (DCCE), São Paulo State University, Brazil

**Keywords:** Dysplasia, Fractal Geometry, Reshape, Convolutional Neural Network, Ensemble, Histological Image.


**Abstract:** Oral cavity lesions can be graded by specialists, a task that is both difficult and subjective. The challenges in defining patterns can lead to inconsistencies in the diagnosis, often due to the color variations on the histological images. The development of computational systems has emerged as an effective approach for aiding specialists in the diagnosis process, with color normalization techniques proving to enhance diagnostic accuracy. There remains an open challenge in understanding the impact of color normalization on the classification of histological tissues representing dysplasia groups. This study presents an approach to classify dysplasia lesions based on ensemble models, fractal representations, and convolutional neural networks (CNN). Additionally, this work evaluates the influence of color normalization in the preprocessing stage. The results obtained with the proposed methodology were analyzed with and without the preprocessing stage. This approach was applied in a dataset composed of 296 histological images categorized into healthy, mild, moderate, and severe oral epithelial dysplasia tissues. The proposed approaches based on the ensemble were evaluated with the cross-validation technique resulting in accuracy rates ranging from 96.1% to 98.5% with the non-normalized dataset. This approach can be employed as a supplementary tool for clinical applications, aiding specialists in decision-making regarding lesion classification.


## 1 INTRODUCTION


Oral epithelial dysplasia is a benign precancerous abnormality with the potential to progress into malignant


oral cancer (Warnakulasuriya et al., 2008). It is a relatively common precursor of oral cancer and, if not treated in its early stages, it may progress to oral squamous cell carcinoma (Smith et al., 2009). Tabagism, alcoholism, nutritional deficiencies and genetic predispositions are the main etiological and predisposing factors that contribute to the lesion malignant progression (Pires et al., 2013). These lesions are locally invasive and are associated with a 50% survival rate (Sagheer et al., 2021).


In the diagnostic process, these lesions may be categorized into different types by the specialist, rendering the diagnosis subjective and presenting challenges in distinguishing the lesions (Warnakulasuriya et al., 2008). The diagnosis is often performed by


<sup>a</sup> <https://orcid.org/0000-0003-2673-3125>


<sup>b</sup> <https://orcid.org/0000-0001-8999-1135>


<sup>c</sup> <https://orcid.org/0000-0003-4635-5037>


<sup>d</sup> <https://orcid.org/0000-0003-1809-0617>


<sup>e</sup> <https://orcid.org/0000-0001-5883-2983>

<sup>f</sup> <https://orcid.org/0000-0003-2650-3960>

<sup>g</sup> <https://orcid.org/0000-0001-9707-9365>

<sup>h</sup> <https://orcid.org/0000-0002-9291-8892>

<sup>i</sup> <https://orcid.org/0000-0001-8580-7054>

<sup>j</sup> <https://orcid.org/0000-0003-3537-0178>

microscopical analysis of the lesion size and the intensity of morphological alterations in tissue nuclei. These histological structures are commonly assessed using the hematoxylin-eosin (H&E) stain, in which the cell nuclei are dyed purple and the cytoplasm and other structures are dyed in pink color. The difficulty in defining patterns can lead to inconsistencies in the diagnosis, causing a divergence rate ranging between 19% and 38% (Müller, 2018). The development of computer-aided diagnosis (CAD) systems can help pathologists in tissue analysis and reduce errors caused by subjectivity (He et al., 2022).

In CAD systems, different computational algorithms are employed for image classification. These algorithms are employed at various stages, including preprocessing, feature extraction, and classification (Ribeiro et al., 2018). One of the approaches applied during the preprocessing step is color normalization. This technique is relevant to reduce variations in H&E stain colors caused by differences in digitization systems or variations in the concentration of a staining solution (Tosta et al., 2023). These color variations can significantly reduce the performance of CAD techniques. Thus, investigating the impact of H&E normalization on the classification of dysplasia images remains an ongoing challenge.

In the feature extraction stage, numerous methods can be found in the literature, including fractal dimensions, entropy, co-occurrence matrix and deep learning models. However, it is not possible to define a universal approach that provides relevant results for any type of image. As a result, it is necessary to investigate the performance of these methods in different types of applications. In recent years, the use of fractal features, derived from the concepts of fractal geometry, has shown relevant results in the classification of histological images (Roberto et al., 2021a; Martins et al., 2021; Ribeiro et al., 2019).

Fractal geometry (FG) is a way to describe the texture of a surface (Morrison, 1975). From approaches based on FG, it is possible to explore descriptors that are invariant to the scale and rotation of visual elements present in different image types. The main descriptors that can be extracted through FG are the fractal dimension (FD), lacunarity (LAC) and percolation (PERC). The DF descriptor is used to measure the irregularity and complexity of a fractal. The LAC metric indicates how the space is filled, while PERC describes properties related to the presence, size and amount of clusters present within the images. These descriptors have presented relevant results in the analysis of histopathological images (Ivanovici et al., 2009). FG descriptors provide a feature vector of high dimensionality, as they encompass local

attributes for image regions and global attributes for the entire image. These features can be investigated via CNN classification, transforming the 1D descriptor sets into 2D matrices (images).

Among these strategies, the sequential representation (SR) (Lumini and Nanni, 2018) and the recurrence plot (RP) (Eckmann et al., 1995) have been used for histological images (Roberto et al., 2017). In recent years, many studies have investigated histological images through convolutional neural networks (CNN) (Roberto et al., 2017). Other studies have investigated the use of engineering-based descriptors, such as coefficients and heatmaps obtained by wavelet transforms along with CNNs for lesion classification (Hu et al., 2023). The use of strategies based on 2D matrices from FG attributes is still a challenge for the classification of dysplasia levels in oral cavity histological tissues.

This paper presents an ensemble-based model to classify histological dysplasia lesions of the oral cavity using FG images. The FG features were reshaped to generate 2D matrices, which were given as input to CNN models composed of the ResNet-50 and EfficientNet models. The results obtained from these CNN models were employed in the ensemble model for lesion classification. Thus, this study provides the following contributions:

- An investigation of the H&E stain normalization stage for the classification of oral epithelial dysplasia tissues;
- Evaluation of the use of fractal attributes in a computational method for classifying oral cavity histological tissue images;
- Development of an ensemble approach based on FG with reshaping techniques and CNN models to classify histological tissues from the oral cavity;

## 2 METHODOLOGY

The proposed method is represented in Figure 1 where the first step (2D Texture Image Stage) employed the FG feature extraction and the reshaping techniques to compute the 2D representation matrices. For the classification step, the 2D representation matrix was employed as an input of CNN models with 10-fold cross-validation. Finally, an ensemble method was used to combine the classification of the CNN with different image inputs.

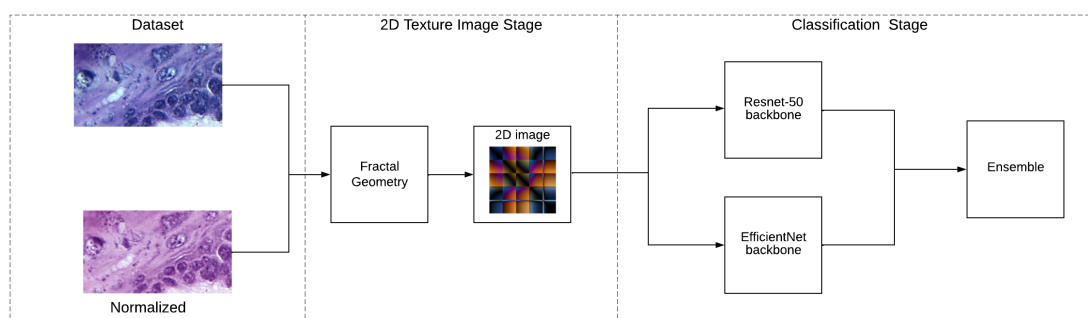


Figure 1: Overview of the proposed approach based on CNN ensemble, FG features, and reshaping procedures for classification of dysplasia of oral cavity tissues.

## 2.1 Image Dataset

This study used an image dataset that consists of microscopic images of 30 H&E-stained mice tongue tissue sections previously submitted to a carcinogen (Silva et al., 2022a). The tongues were processed and embedded in paraffin to create the blocks, which were sectioned and stained with hematoxylin and eosin (H&E) for histopathological study. Using the methodology described by (Lumerman et al., 1995), the images were classified among healthy mucosa and the mild, moderate and severe dysplasias by one specialist.

The histological slides were digitized using the Leica DM500 optical microscope with  $400\times$  magnification, resulting in 66 images stored in TIFF file format, using the RGB color model with an 8-byte quantization. From these images, 74 regions of interest (ROIs) sized at  $450\times 250$  pixels were obtained for each class, totaling 296 ROI images. Examples of these ROIs are displayed in Figure 2.

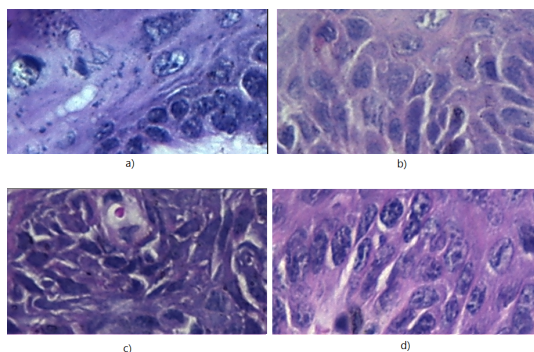


Figure 2: Examples of histological tissues from the oral cavity: (a) healthy tissue, (b) mild dysplasia, (c) moderate dysplasia, and (d) severe dysplasia.

## 2.2 Normalization

The process employed in the staining of tissue histological has a significant influence on color variations or the use of different scanners (Ribeiro et al., 2018). Then, the influence of these variations on color and texture features can reduce the effectiveness of processing image techniques employed during the classification of the histological tissues.

In this work, the normalization technique described by Tosta et al. (Tosta et al., 2023) was applied to compare the influence of color normalization in the process of classification in dysplasia images. Thus, this investigation considered a stain color normalization with a robust dictionary learning method where the estimation of color appearance matrices and stain density maps were applied to the normalization of the H&E histological images. The method considered pixel selection and weight definition to improve the color estimation of histological images.

## 2.3 2D Texture Image Stage

In this step, the gliding box algorithm was considered with three different distances to extract the feature vectors. Thus, the attribute vector was reshaped from a 1D to a 2D matrix using representation models based on a sequential image and a recurrence plot image (2D texture image).

### 2.3.1 Fractal Feature

FG allows quantifying an object about the invariance of its shape when its scale is changed, keeping its structure identical to the original. In digital images, this concept is used to observe self-similarity properties, when a portion of the image can be seen as a replica of everything, on a smaller scale. Among the different algorithms for calculating the self-similarity properties, due to its efficiency, the gliding-box algorithm stands out (Tolle et al., 2008). This algorithm

proposes to divide the images into different scales and then extract information from each sub-image, generating local and global descriptors. Based on this approach, features obtained with the FD, LAC, and PERC metrics have been successfully used in histological tissue analysis (Roberto et al., 2017).

The gliding box algorithm consists of a multiscale analysis where a square box of size  $L \times L$  in the upper left corner of the image, where  $L$  represents the size of the square box under pixels, in order to slide from left to right to the bottom region of the image, passing through all pixels. The box moves through the whole image, pixel by pixel, and the value of  $L$  is increased after reaching the end line or column. After sliding across the entire image, the box is repositioned at the starting point and the value of  $L$  is incremented by two units.

In each box, a color similarity analysis was performed for each pixel in the square box. This analysis was performed by fixing the central pixel and assigning it to a vector  $f_c = (r_c, g_c, b_c)$ , which  $r_c$ ,  $g_c$  and  $b_c$  correspond to the intensity of each of the color channels considering images in RGB color model. The remaining pixels of the box,  $f_i = (r_i, g_i, b_i)$  were compared to the central by calculating a color distance  $\Delta$ , which allows verifying which pixels belong to the RGB hyperspace obtained by the central pixel of the box (Ivanovici et al., 2009). In this work, the Minkowski ( $\Delta_{mink}$ ), Euclidean ( $\Delta_{eucl}$ ) and Manhattan ( $\Delta_{manh}$ ) distances were computed according to the Equations 1-3:

$$\Delta_{mink} = \max(|f_i(k_i) - f_c(k_c)|), k \in r, g, b. \quad (1)$$

$$\Delta_{eucl} = \sqrt{\sum_k (f_i(k_i) - f_c(k_c))^2}, k \in r, g, b. \quad (2)$$

$$\Delta_{manh} = \sum_k |f_i(k_i) - f_c(k_c)|, k \in r, g, b. \quad (3)$$

When the  $\Delta$  value is less than or equal to the scale  $L$ , then  $f_i$  is labeled 1, otherwise, it is labeled 0. After calculating the number of pixels that satisfy  $\Delta$ , for each box and different values of  $L$ , the information obtained was used to build a probability matrix. In this matrix, each element corresponds to the probability that pixels on a  $L$  scale are labeled 1. Finally, this matrix was normalized so that the sum of all elements in a column equals 1. This matrix resulted in the local and global values of the FD, LAC, and PERC metrics as described in (Roberto et al., 2021a) and (Ribeiro et al., 2019).

In this work, the use of the gliding box algorithm in multi-resolution allowed obtaining information in boxes of size 3 to 41, because these values have demonstrated relevant results in classification as presented in (Ribeiro et al., 2019). Then, a local fractal feature vector was obtained with 100 features for each  $\Delta$  parameter.

### 2.3.2 Reshaping Procedure

In order to generate a 2D representation matrix to be given as input to a CNN, two different reshaping procedures, the sequential reshape (SR) and recurrence plot (RP) methods were applied to the set of the 300 local fractal features obtained with 100 values of each  $\Delta$ . The SR reshaping procedure is similar to the model employed in the study of (Roberto et al., 2021b), where a simple reshape was obtained of a  $10 \times 10 \times 3$  matrix. The features were extracted and organized in a sequential arrangement obtained of  $\Delta_{mink}$ ,  $\Delta_{eucl}$  and  $\Delta_{manh}$  correspond to the red, green, and blue color channels, respectively. The RP reshaping procedure is a technique proposed in (Eckmann et al., 1995) for the projection of repeated events into two or three-dimensional spaces. The application of this technique on a feature vector containing  $N$  feature output a squared matrix  $N \times N$  wherein each element  $R_{i,j}$  of the matrix was obtained by the Equation 4:

$$R_{i,j} = \|x_i - x_j\| \quad \forall i, j = 1 \dots N. \quad (4)$$

where  $x_i$  and  $x_j$  are the  $i^{th}$  and  $j^{th}$  features in the vector, and  $\|x_i - x_j\|$  is the norm of the euclidean distance.

Figure 3 shows the images obtained of the reshaping procedures of a healthy histological tissue digital image.

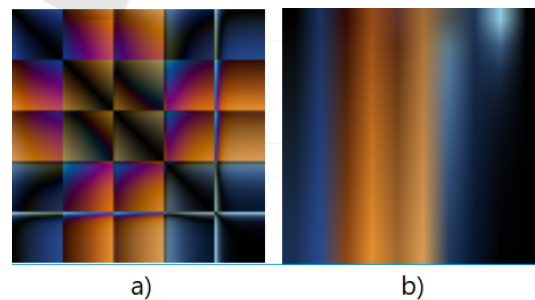


Figure 3: The reshaping methods from a healthy tissue image: (a) RP and (b) SR.

## 2.4 Classification Stage

For the classification step, the 2D representations were employed as an input of CNN models with 10-fold cross-validation where 90% of the dataset was



used for model training and 10% for test. An ensemble method was used to combine the classification of the CNN with different images.

#### 2.4.1 CNN Models

The CNN models are among the most powerful tools for the analysis of digital images. The proposed model combined strategies capable of extracting features and classifying data based on a supervised method (LeCun et al., 1998). In this work, the ResNet-50 (He et al., 2015) and EfficientNet (Tan and Le, 2019) models were applied to classifier the representation model. In order to use both datasets in the CNNs models, the input images were resized to  $224 \times 224$  pixels to fit them to both models. In the training stage of each network, the representation images obtained after reshaping the local fractal features were used with a 10-fold cross-validation technique. In each training fold (90%), 30% of the set was used as the validation set and the remaining 10% was used as a test set. The training was made with 10 epochs and 5 iterations per epoch with an initial learn rate of 0.01, the piecewise function as the learn rate schedule and with 2 as the learn rate drop period, which updates the learning rate every 2 epochs.

In order to reduce the training time of CNN and obtain promising results with a smaller number of epochs, the transfer learning strategy was adopted to decrease the evaluation time of a model without pre-setting the weights. The transfer learning model employed in this work was the network-based transfer, in which the parameters of the first CNN layers are set from a model-trained network. This approach explored a specific weight adjustment strategy, selecting partial instances from the source domain as supplements to the defined training nested in the target domain. In this work, the last fully connected layer and the classification layer for both CNN models were adapted based on the dataset groups.

#### 2.4.2 Ensemble Strategy

With the results obtained from the probabilities of occurrence of each class, using the output of the softmax layer of each CNN, an ensemble model was employed to classify the data. The use of an ensemble model is among the most successful approaches in machine learning applications in recent years, due to its ability to train different models and combine their predictions, with relevant results obtained in histopathology (Kassani et al., 2019).

Among the several ensemble techniques, simple averaging is a widely used combination technique in classification problems with CNN models. In this

strategy, the final decision was obtained by averaging the results of individual models (Zhou, 2012). Then, the sample is classified according to the highest class probability after calculating the average of the probability curve of each model (see Equation 5).

$$A(x) = \frac{1}{N} \sum_{i=1}^N a_i(x), \quad (5)$$

where  $A(x)$  is the final result,  $N$  is the number of models and  $a_i(x)$  is the output generated by the  $i$ -th model.

### 2.5 Experimental Evaluation

For the evaluation of the proposed approach, the following strategies were investigated:

- Strategy 1 - (named CNN baseline): each CNN model was evaluated with the images from the dataset;
- Strategy 2 - (named CNN SR); each CNN model applied with the *SR* representation images;
- Strategy 3 - (named CNN RP): each CNN model applied with the *RP* representation images;
- Strategy 4 - (named SR + RP); an ensemble based in the *SR* and *RP* representation images;
- Strategy 5 (named CNN baseline + SR + RP): an ensemble based in the *SR* and *RP* representation images with the CNN baseline;

The evaluation of the proposed methods was carried out by comparing the results obtained with the gold standard classification carried out by the specialist, in which the metrics of accuracy ( $A_{CC}$ ) and F1-score were obtained.

## 3 EXPERIMENTAL RESULTS

Tables 1 and 2 show the average  $A_{CC}$  and F1-score values obtained with the investigated strategies using both CNN models over the non-normalized and the normalized datasets, respectively. From the results presented in Table 1, it is observed that the ResNet-50 baseline achieved higher  $A_{CC}$  and F1-score values than the EfficientNet baseline, but it also shows a higher standard deviation. For both CNN models, the reshaping strategies allowed an improvement on the results, showing higher metric values and lower standard deviation than the baselines, with the ResNet-50 achieving 97.3% and 94.5% for the metrics of  $A_{CC}$  and F1-score, respectively. When using the ensemble methods, Strategies 4 and 5 allowed an increase on result values for the ResNet-50, but, for the EfficientNet, only Strategy 5 showed the same behavior. On

this experimental test, the highest values were 98.5% and 96.8% for the  $A_{CC}$  and F1-score, respectively, using the ResNet-50 in combination with Strategy 5.

Table 1: Results obtained with proposed classification approaches on the dataset with the healthy, mild, moderate, and severe images.

Model	Strategy	$A_{CC}$	F1-score
ResNet50	Baseline	90.5 ± 4.4	80.6 ± 9.0
	SR	96.9 ± 2.7	93.8 ± 5.5
	RP	97.3 ± 2.7	94.5 ± 5.5
	SR+RP	98.1 ± 2.5	96.1 ± 5.2
	Baseline+SR+RP	98.5 ± 2.2	96.8 ± 4.6
EfficientNet	Baseline	89.5 ± 3.5	78.6 ± 7.5
	SR	96.6 ± 2.7	92.2 ± 5.1
	RP	96.1 ± 2.5	93.1 ± 5.6
	SR+RP	96.1 ± 2.4	92.1 ± 5.1
	Baseline+SR+RP	97.3 ± 2.4	94.5 ± 4.9

The classification results on the normalized dataset are presented in Table 2. The normalization showed no significant difference on the results from the ResNet-50 baseline but showed lower values for the reshaping and ensemble strategies. The same behavior was observed for the EfficientNet model, which showed decrease rates ranging from 0.8% to 4.5% and 2.1% to 8.5% for the metrics of  $A_{CC}$  and F1-score, respectively. From the results presented on this table, it was observed that the normalization stage played no significant role on the classification stage. However, for both CNN models and datasets, the use of reshaping techniques allowed higher classification results than the baseline and even further improvement when using ensemble methods.

Table 2: Performance of proposed classification methods on the normalized image dataset with the healthy, mild, moderate, and severe images.

Model	Strategy	$A_{CC}$	F1-score
ResNet50	Baseline	90.5 ± 4.0	80.3 ± 8.9
	SR	94.8 ± 3.8	89.5 ± 7.6
	RP	95.9 ± 3.3	91.7 ± 6.6
	SR+RP	96.6 ± 2.9	93.2 ± 5.8
	Baseline+SR+RP	97.5 ± 2.6	94.9 ± 5.2
EfficientNet	Baseline	88.7 ± 5.0	76.5 ± 10.6
	SR	92.1 ± 3.3	83.7 ± 7.0
	RP	93.9 ± 4.6	87.7 ± 9.0
	SR+RP	93.8 ± 3.3	87.2 ± 6.7
	Baseline+SR+RP	95.8 ± 2.5	91.4 ± 4.0

Since the results presented on Tables 1 and 2 indicates that the best result were obtained by ResNet-50 on the non-normalized dataset, the Wilcoxon test was performed to assess its statistical relevance compared to the EfficientNet model and the normalized dataset. The test was employed with a significance level of 5% and the obtained  $p$ -values are shown in Table 3. All comparisons showed  $p$ -values higher than 0.05, meaning that there's no statistically significant differ-

Table 3:  $P$ -values obtained using the Wilcoxon test for comparing the non-normalized ResNet-50 model with the other approaches.

Metric	Original dataset		Normalized dataset	
	ResNet	EfficientNet	ResNet	EfficientNet
$A_{CC}$	-	0.17	0.23	0.09
F1-score	-	0.16	0.22	0.09

ence between the methods. It is noted that the EfficientNet model presented  $p$ -values of 0.09, representing the most different results compared to the non-normalized ResNet-50 approach.

Table 4 depicts the obtained  $A_{CC}$  values for each dysplasia class using both models and strategies on the original image data since it presented the best results. The healthy and severe classes presented higher values than the Mild and Moderate. However, this difference is minimized with the use of reshaping and ensemble strategies. The ResNet-50 with strategies 4 and 5 showed  $A_{CC}$  of 99% and 100% for the Severe Moderate classes, respectively. The values shown by EfficientNet using the same strategies were slightly lower.

Table 5 presents an overview of the obtained results about those provided by relevant computer vision methods developed to study histopathological images of the oral cavity. It is important to highlight that the studies by (Amin et al., 2021) and (Deif et al., 2022) were developed to distinguish healthy tissues from lesions in advanced stages. It is also noted that the study presented by (Adel et al., 2018) investigated OED images, but no assessment regarding their grades was proposed. The (Silva et al., 2022b) and (Silva et al., 2022a) studies investigated the classification of lesions and the degree of dysplasia and the proposed solution is a new approach to classification. Considering these results, the proposed methodology provided values compatible with those available in the literature, achieving significant results for the automatic grading task, especially via association of fractal features, reshape approaches and CNN models.

## 4 CONCLUSIONS

This study presented an approach for oral dysplasia grading based on fractal features in combination with reshaping techniques and CNN models. Ensemble strategies were also employed to further improve the classification results. The proposed reshaping and ensemble methodology allowed an increase in the classification results, with the ResNet-50 using the ensemble of all reshaping strategies providing an  $A_{CC}$  value of 98.5.

Table 4: Accuracy results for each dysplasia grade on the non-normalized dataset for both CNN models.

CNN	Tissue	Baseline	SR	RP	SR + RP	Baseline + SR + RP
ResNet50	Healthy	91.5 ± 6.9	95.6 ± 4.3	96.6 ± 3.6	96.9 ± 3.8	97.9 ± 2.9
	Mild	87.2 ± 4.9	94.9 ± 5.7	94.5 ± 5.5	96.2 ± 5.0	96.9 ± 4.4
	Moderate	85.8 ± 8.1	98.3 ± 2.9	97.9 ± 2.4	99.3 ± 1.5	99.0 ± 2.3
	Severe	97.6 ± 4.3	99.0 ± 2.3	100.0 ± 0.0	100.0 ± 0.0	100.0 ± 0.0
EfficientNet	Healthy	90.9 ± 7.2	93.9 ± 4.4	94.9 ± 4.6	94.3 ± 4.3	95.9 ± 3.8
	Mild	83.5 ± 6.0	92.2 ± 5.1	93.6 ± 5.7	92.6 ± 5.3	94.6 ± 4.9
	Moderate	86.5 ± 5.0	98.6 ± 2.4	98.6 ± 2.4	98.3 ± 2.4	99.0 ± 2.3
	Severe	97.3 ± 4.2	99.7 ± 1.1	99.3 ± 1.5	99.3 ± 1.5	99.7 ± 1.1

Table 5: Assessment of the proposed methodology in relation to existing approaches for oral image classification present in the literature.

Study	Lesion Type	Feature Extraction	Classifier	$A_{CC}$
(Adel et al., 2018)	Oral Dysplasia	ORB	SVM	92.6
(Amin et al., 2021)	Squamous Cell Carcinoma	Deep features	CNN	96.6
(Deif et al., 2022)	Squamous Cell Carcinoma	Deep features	XGBoost	96.3
(Silva et al., 2022b)	Oral Dysplasia	Texture and morphology features	HOP	92.4
(Silva et al., 2022a)	Oral Dysplasia	Deep features	HOP	98.0
Proposed Method	Oral Dysplasia	Fractal features and reshape	Ensemble CNN	98.5

Experimental tests indicated that the use of normalization does not bring any effective gains to the classification task. However, regardless of the dataset or CNN model, the reshaping strategies increased the result values and the ensemble techniques allowed further improvements of them.

The obtained results were comparable to those present in the literature, showing  $A_{CC}$  significantly higher than four of the presented studies. The proposed methodology can be used as an automated tool to aid specialists during the histological analysis of oral dysplasia lesions. In future works, other reshaping techniques will be investigated, such as the Self-similarity matrix and Markov transition field methods, and other CNN models will be explored for the classification task.

## ACKNOWLEDGMENT

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. The authors gratefully acknowledge the financial support of National Council for Scientific and Technological Development CNPq (Grants #313643/2021-0, #311404/2021-9 and #307318/2022-2), the State of Minas Gerais Research Foundation - FAPEMIG (Grant #APQ-00578-18 and Grant #APQ-01129-21) and São Paulo Research Foundation - FAPESP (Grant #2022/03020-1).

## REFERENCES

- Adel, D., Mounir, J., El-Shafey, M., Eldin, Y. A., El Masry, N., AbdelRaouf, A., and Abd Elhamid, I. S. (2018). Oral epithelial dysplasia computer aided diagnostic approach. In *2018 13th International Conference on Computer Engineering and Systems (ICCES)*, pages 313–318. IEEE.
- Amin, I., Zamir, H., and Khan, F. F. (2021). Histopathological image analysis for oral squamous cell carcinoma classification using concatenated deep learning models. *medRxiv*, pages 2021–05.
- Deif, M. A., Attar, H., Amer, A., Elhady, I. A., Khosravi, M. R., Solyman, A. A., et al. (2022). Diagnosis of oral squamous cell carcinoma using deep neural networks and binary particle swarm optimization on histopathological images: an aiomt approach. *Computational Intelligence and Neuroscience*, 2022.
- Eckmann, J.-P., Kamphorst, S. O., Ruelle, D., et al. (1995). Recurrence plots of dynamical systems. *World Scientific Series on Nonlinear Science Series A*, 16:441–446.
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition. *CoRR*, abs/1512.03385.
- He, W., Han, Y., Ming, W., Du, J., Liu, Y., Yang, Y., Wang, L., Wang, Y., Jiang, Z., Cao, C., et al. (2022). Progress of machine vision in the detection of cancer cells in histopathology. *IEEE Access*, 10:46753–46771.
- Hu, K., Tan, H., Zhang, Y., Huang, W., and Gao, X. (2023). Mwg-net: Multi-scale wavelet guidance network for covid-19 lung infection segmentation from ct images. *IEEE Transactions on Instrumentation and Measurement*.

- Ivanovici, M., Richard, N., and Decean, H. (2009). Fractal dimension and lacunarity of psoriatic lesions—a colour approach. *medicine*, 6(4):7.
- Kassani, S. H., Kassani, P. H., Wesolowski, M. J., Schneider, K. A., and Deters, R. (2019). Classification of histopathological biopsy images using ensemble of deep learning networks. In *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering*, pages 92–99.
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.
- Lumerman, H., Freedman, P., and Kerpel, S. (1995). Oral epithelial dysplasia and the development of invasive squamous cell carcinoma. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 79(3):321–329.
- Lumini, A. and Nanni, L. (2018). Convolutional neural networks for atc classification. *Current pharmaceutical design*, 24(34):4007–4012.
- Martins, A. S., Neves, L. A., de Faria, P. R., Tosta, T. A., Longo, L. C., Silva, A. B., Roberto, G. F., and do Nascimento, M. Z. (2021). A hermite polynomial algorithm for detection of lesions in lymphoma images. *Pattern Analysis and Applications*, 24:523–535.
- Morrison, P. (1975). Les objets fractals: Forme, hasard et dimension.
- Müller, S. (2018). Oral epithelial dysplasia, atypical verrucous lesions and oral potentially malignant disorders: focus on histopathology. *Oral surgery, oral medicine, oral pathology and oral radiology*, 125(6):591–602.
- Pires, F. R., Ramos, A. B., Oliveira, J. B. C. d., Tavares, A. S., Luz, P. S. R. d., and Santos, T. C. R. B. d. (2013). Oral squamous cell carcinoma: clinicopathological features from 346 cases from a single oral pathology service during an 8-year period. *Journal of Applied Oral Science*, 21:460–467.
- Ribeiro, M. G., Neves, L. A., do Nascimento, M. Z., Roberto, G. F., Martins, A. S., and Tosta, T. A. A. (2019). Classification of colorectal cancer based on the association of multidimensional and multiresolution features. *Expert Systems with Applications*, 120:262–278.
- Ribeiro, M. G., Neves, L. A., Roberto, G. F., Tosta, T. A., Martins, A. S., and Do Nascimento, M. Z. (2018). Analysis of the influence of color normalization in the classification of non-hodgkin lymphoma images. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 369–376. IEEE.
- Roberto, G. F. et al. (2021a). Associação entre atributos manuais e aprendizado profundo baseada em geometria fractal para classificação de imagens histológicas.
- Roberto, G. F., Lumini, A., Neves, L. A., and do Nascimento, M. Z. (2021b). Fractal neural network: A new ensemble of fractal geometry and convolutional neural networks for the classification of histology images. *Expert Systems with Applications*, 166:114103.
- Roberto, G. F., Neves, L. A., Nascimento, M. Z., Tosta, T. A., Longo, L. C., Martins, A. S., and Faria, P. R. (2017). Features based on the percolation theory for quantification of non-hodgkin lymphomas. *Computers in biology and medicine*, 91:135–147.
- Sagheer, S. H., Whitaker-Menezes, D., Han, J. Y., Curry, J. M., Martinez-Outschoorn, U., and Philp, N. J. (2021). Chapter 6 - 4nq induced carcinogenesis: A mouse model for oral squamous cell carcinoma. In Galluzzi, L. and Buqué, A., editors, *Carcinogen-driven mouse models of oncogenesis*, volume 163 of *Methods in Cell Biology*, pages 93–111. Academic Press.
- Silva, A. B., De Oliveira, C. I., Pereira, D. C., Tosta, T. A., Martins, A. S., Loyola, A. M., Cardoso, S. V., De Faria, P. R., Neves, L. A., and Do Nascimento, M. Z. (2022a). Assessment of the association of deep features with a polynomial algorithm for automated oral epithelial dysplasia grading. In *2022 35th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, volume 1, pages 264–269. IEEE.
- Silva, A. B., Martins, A. S., Tosta, T. A., Neves, L. A., Servato, J. P. S., de Araújo, M. S., de Faria, P. R., and do Nascimento, M. Z. (2022b). Computational analysis of histological images from hematoxylin and eosin-stained oral epithelial dysplasia tissue sections. *Expert Systems with Applications*, 193:116456.
- Smith, J., Rattay, T., McConkey, C., Helliwell, T., and Mehanna, H. (2009). Biomarkers in dysplasia of the oral cavity: a systematic review. *Oral oncology*, 45(8):647–653.
- Tan, M. and Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR.
- Tolle, C. R., McJunkin, T. R., and Gorsich, D. J. (2008). An efficient implementation of the gliding box lacunarity algorithm. *Physica D: Nonlinear Phenomena*, 237(3):306–315.
- Tosta, T. A. A., Freitas, A. D., de Faria, P. R., Neves, L. A., Martins, A. S., and do Nascimento, M. Z. (2023). A stain color normalization with robust dictionary learning for breast cancer histological images processing. *Biomedical Signal Processing and Control*, 85:104978.
- Warnakulasuriya, S., Reibel, J., Bouquot, J., and Dabelsteen, E. (2008). Oral epithelial dysplasia classification systems: predictive value, utility, weaknesses and scope for improvement. *Journal of Oral Pathology & Medicine*, 37(3):127–133.
- Zhou, Z.-H. (2012). *Ensemble methods: foundations and algorithms*. CRC press.