



Article

An Intelligent System to Improve Diagnostic Support for Oral Squamous Cell Carcinoma

Afonso U. Fonseca ^{1,*}, Juliana P. Felix ¹, Hedenir Pinheiro ¹, Gabriel S. Vieira ^{1,2}, Yleris C. Mourão ³, Juliana C. G. Monteiro ⁴ and Fabrizzio Soares ¹

¹ Institute of Informatics, Federal University of Goiás, Goiânia 74690-900, GO, Brazil; julianafelix@discente.ufg.br (J.P.F.); hedenir@discente.ufg.br (H.P.); gabriel.vieira@ifgoiano.edu.br (G.S.V.); fabrizzio@ufg.br (F.S.)

² Federal Institute Goiano, Computer Vision Lab, Urutai 75790-000, GO, Brazil

³ General Hospital of Goiânia, Goiânia 74110-010, GO, Brazil; hgg.tutoria.fono@idtech.org.br

⁴ Araújo Jorge Cancer Hospital, Goiânia 74605-070, GO, Brazil; juliana.monteiro@haj.org.br

* Correspondence: afonsoueslei@ufg.br

Abstract: Oral squamous cell carcinoma (OSCC) is one of the most-prevalent cancer types worldwide, and it poses a serious threat to public health due to its high mortality and morbidity rates. OSCC typically has a poor prognosis, significantly reducing the chances of patient survival. Therefore, early detection is crucial to achieving a favorable prognosis by providing prompt treatment and increasing the chances of remission. Salivary biomarkers have been established in numerous studies to be a trustworthy and non-invasive alternative for early cancer detection. In this sense, we propose an intelligent system that utilizes feed-forward artificial neural networks to classify carcinoma with salivary biomarkers extracted from control and OSCC patient samples. We conducted experiments using various salivary biomarkers, ranging from 1 to 51, to train the model, and we achieved excellent results with precision, sensitivity, and specificity values of 98.53%, 96.30%, and 97.56%, respectively. Our system effectively classified the initial cases of OSCC with different amounts of biomarkers, aiding medical professionals in decision-making and providing a more-accurate diagnosis. This could contribute to a higher chance of treatment success and patient survival. Furthermore, the minimalist configuration of our model presents the potential for incorporation into resource-limited devices or environments.

Keywords: oral squamous cell carcinoma; metabolites' salivary biomarkers; intelligent diagnostic support system



Citation: Fonseca, A.U.; Felix, J.P.; Pinheiro, H.; Vieira, G.S.; Mourao, Y.C.; Monteiro, J.C.G.; Soares, F. An Intelligent System to Improve Diagnostic Support for Oral Squamous Cell Carcinoma. *Healthcare* **2023**, *11*, 2675. <https://doi.org/10.3390/healthcare11192675>

Academic Editor: Yoshiharu Motoo

Received: 18 August 2023

Revised: 12 September 2023

Accepted: 25 September 2023

Published: 3 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Oral squamous cell carcinoma (OSCC) is among the ten most-prevalent cancer categories worldwide, associated with high mortality and morbidity rates, representing a global public health problem. In 2020, according to statistics from the World Health Organization's Global Cancer Observatory, lip and oral cavity cancers accounted for more than 377,700 cases worldwide [1]. About 90% of oral malignancies correspond to squamous cell carcinoma [2,3].

OSCC is a malignant tumor that develops from the squamous epithelial cells lining the oral cavity, encompassing the lips, tongue, buccal mucosa, upper and lower gums, and palate, among other areas. Understanding the pathophysiology of OSCC is crucial for early detection, prevention, and effective management of this aggressive cancer. The main risk factors for OSCC include tobacco and alcohol consumption and human *papillomavirus* (HPV) infection present in most patients diagnosed with oral and oropharyngeal cancer, making them critical etiological factors [4]. Moreover, chronic irritation, such as from sharp teeth, ill-fitting dentures, and so forth, and a poor diet in terms of fruits and vegetables are

factors. Additionally, socioeconomic factors, such as poor oral hygiene and limited access to healthcare services, also impact early detection and treatment outcomes [5].

The pathophysiology involves a multi-step process of genetic and molecular changes that lead to uncontrolled cell growth and the formation of malignant tumors. As the tumor advances, it can infiltrate nearby tissues, extend to regional lymph nodes, and even metastasize to distant organs. This progression may lead to the temporary or permanent compromise of crucial functions such as speech, voice, swallowing, and mastication and may also be influenced by the chosen treatment approach [6]. In summary, OSCC involves the accumulation of genetic mutations, dysregulation of signaling pathways, invasion, and metastasis. Early detection and intervention are crucial for improving the prognosis and quality of life for individuals with OSCC. Preventive measures, such as avoiding tobacco and alcohol and receiving the HPV vaccine, can also reduce the risk of developing this devastating disease.

Management of all oral cavity cancers must occur in a Multidisciplinary Head and Neck Oncology Team due to the functional and aesthetic implications of the treatment, and the diagnosis of OSCC is traditionally performed by biopsy. OSCC's 5-year survival rates are around 50%, and most of these patients survive a short time after diagnosis [7]. This scenario is because most tumors are identified late, compromising treatment, prognosis, and patient survival [8]. Given its high mortality, early and accurate diagnosis is extremely important [9].

To better understand the disease and the contribution of early diagnosis, disseminating OSCC signs and symptoms and implementing actions for early detection are essential. In this sense, tools and solutions that help diagnosis and prognosis in the early stages of the disease can be of great value, especially for patients. With early identification or an early prognosis, an increase in survival rates due to the initiation of treatment in the early stages is expected. In addition, a reduction in the mutilating impact of the treatment can also be cited, benefiting the patient and ensuring a better quality of life. These tools help medical professionals decide on the best treatment [10].

Studies indicate that salivary biomarkers can be used for OSCC's early recognition [11]. Furthermore, salivary collection can be performed in a simple and non-invasive way, and its potential to aid in OSCC diagnosing has been studied by several authors [12–18].

Similarly, Artificial-Intelligence-guided approaches have shown considerable benefits for early diagnosis and prognosis of oral cancers [19], presenting computational machine learning solutions with improved accuracy of cancer sensitivity, recurrence, and survival predictions [20] and providing a consequent improvement in patient care and survival [21–25].

These approaches have been gaining more and more space in the medical field, with models based on artificial neural networks successfully used in several applications. For example, classification, prognosis, diagnosis, or identification of lung diseases [26,27], chronic kidney diseases [28], macular degeneration and diabetic retinopathy [29], Parkinson's disease [30], and oral cancer and periodontitis [31–33] are among these applications, to name a few.

It is important to note that OSCC is hard to diagnose in the early stages, especially because patients are not aware of the symptoms and signs, which mimic regular oral irritations, ulcerations, leukoplakia, and gingivitis. Therefore, small injuries can be misinterpreted, allowing the cancer to evolve. Classical OSCC diagnosis relies on a biopsy being conducted, required only after clinical observation; however, injuries are already visible and likely too advanced. Therefore, a simple process for collecting salivary samples to gather biomarkers is not only quick, less expensive, and non-invasive, but also can become a general clinical protocol for prevention. Finally, analyzing the biomarkers is a hard task, so our study aimed to evaluate their impact on classifiers for the early diagnosis of OSCC. Therefore, this work aimed to develop a method with a shallow artificial neural network to distinguish data from people with OSCC and healthy people from salivary biomarker data. The proposal was evaluated using the cross-validation and data augmentation strategies,

and the calculated accuracy, sensitivity, and specificity metrics, combined with the simplicity and portability of the proposed method, point to a way to use the proposal in practice. Some of our main contributions were:

- Evaluating the use and performance of a feed-forward shallow artificial neural network in the classification of control and OSCC cases through salivary biomarkers;
- Evaluating the impact of using data augmentation to class balancing in the model building;
- Classifying salivary biomarkers according to their importance values based on the quadratic weights assigned by a support vector machine (SVM) classifier;
- Selecting a minimum set of salivary biomarkers to preserve high precision for the classification task.

This work is organized as follows: First, the related work is presented in Section 2, and Section 3 presents the materials and methods used in this work, including information from the database and the characteristics of the implemented neural network. Next, a comparative analysis and discussion of the results obtained are presented in Section 4. Finally, our conclusions and future work follow in Section 5.

2. Related Work

Artificial Intelligence has been successfully used to assist in the diagnosis and prognosis of OSCC cases by analyzing hyperspectral clinical, pathological, and radiographic images. Halicek et al. [34,35] used hyperspectral imaging and convolutional neural networks to diagnose aerodigestive tract tumors with a better sensitivity of 81%. Ariji et al. [36] used computerized tomography (CT) images to detect cervical lymph node metastasis and infer disease-free survival. Fujima et al. [37] similarly used F-fluorodeoxyglucose (FDG) uptake by positron emission tomography (PET/CT) and achieved 80% accuracy in predicting survival. Pathological images were used by Folmsbee et al. [38] to identify oral cavity cancerous tissues among six other tissue classes with an accuracy of 96.37%, while Das et al. [39] classified biopsy images with similar performance to the pathological classification system in [40], achieving 97.51% accuracy.

Although the importance of clinical images in treating OSCC and of Artificial Intelligence solutions to aid in the diagnosis and prognosis of this disease is undeniable, image acquisition is not trivial, requiring highly complex human and material resources, while Artificial Intelligence applications that model the training steps need proper data preparation and labeling in an expensive and time-consuming process. Therefore, we must consider other data types such as gene expression data, spectral data, autofluorescence, and saliva metabolites.

Thus, specifically concerning chemical methods, salivary biomarkers are an excellent alternative for analyzing and monitoring OSCC cases. Salivary biomarkers offer good accuracy in the early recognition of this disease. In addition, collecting salivary samples is a simple, non-invasive process with lower costs than clinical imaging.

Hu et al. [18] demonstrated that protein biomarkers of the human salivary proteome have a high capacity to discriminate OSCC patients from healthy patients, indicating that the combination of some salivary proteins reaches a sensitivity of up to 90% in this task and showing a promising approach for the search for OSCC biomarkers.

Using autofluorescence spectroscopy, Yuvaraj et al. [15] performed a pilot study demonstrating marked differences in the spectral signatures between healthy and OSCC patients, obtaining 85.7% sensitivity when using the fluorescence emission spectrum at a 405nm excitation. The authors further suggested that the reasons are linked to endogenous porphyrin, nicotinamide adenine dinucleotide (phosphate), reduced (NAD(P)H), and flavin adenine dinucleotide (FAD) in exfoliated saliva cells.

Sridharan et al. [41] used quadrupole time-of-flight (Q-TOF) liquid chromatography-mass spectrometry to assess the metabolomic profile of OSCC, oral leukoplakia (OLK), and control groups. The authors found a significant increase in 1-methylhistidine, inositol 1,3,4-triphosphate, d-glycerate-2-phosphate, 4-nitroquinoline-1-oxide, 2-oxoarginine, norcocaine

nitroxide, sphinganine-1-phosphate, and pseudouridine in OLK and OSCC in comparison to control groups.

Song et al. [12] used conductive polymer spray–mass spectrometry (CPSI–MS) to quantify saliva samples from OSCC, patients with premalignant lesions, and healthy individuals. The authors showed that using CPSI–MS with machine learning (ML) can be a viable tool to distinguish OSCC and premalignant lesions from healthy conditions in real-time, with an achieved accuracy of 86.7%.

Ishikawa et al. [42] developed a multiple logistic regression model to differentiate OSCC from oral lichen planus. They realized a comprehensive analysis of fourteen hydrophilic metabolites in the saliva samples by capillary electrophoresis–mass spectrometry, showing high discrimination of the combination of indole-3-acetate and ethanolamine phosphate metabolites, resulting in 85.6% of the area under the curve (AUC).

Recently, Kouznetsova et al. [32] studied metabolic pathways related to oral cancer and periodontitis and, using a deep learning neural network, achieved the best accuracy of 79.54% in the classification of these cases.

Finally, Costa et al. [43] analyzed saliva samples using gas chromatography–mass spectrometry with a data-mining approach to find the best biomarkers to diagnose OSCC. Using feature selection classification and random forest classification algorithms, they reported an AUC of 91% with only three metabolites versus 76% when using all 51 metabolites analyzed.

As we have seen, OSCC can be approached differently from gene expression data, spectral data, autofluorescence, salivary metabolites, cytology images, CT images, and clinicopathological images [33]. Thus, some studies that used clinical images and salivary biomarkers were described to illustrate the OSCC research scenario, which, despite being diversified, still has limitations in chemical and computational advances. Therefore, to overcome some of the limitations of this field and contribute to the research, we combined Artificial Intelligence techniques to deliver a low-cost computational model and real-time response. Our proposal uses the same base of salivary biomarkers used by [43] and presents accuracy, sensitivity, and specificity values superior to previous works for several validation schemes analyzed.

3. Material and Methods

We developed a minimalist and effective oral squamous cell carcinoma diagnosis support method. Our proposal consists of two more significant cohesive steps, structuring a pipeline that begins with the biomarker processing and, then, follows with artificial neural network construction. Figure 1 provides an overview of the proposed method, described in the following subsections.

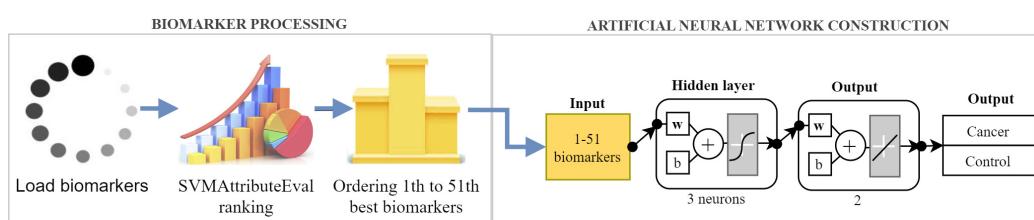


Figure 1. Overview of the proposed method. The diagram is divided into two main steps. In the first step, the biomarkers are classified using the SVMAttributeEval algorithm. In the second step, different numbers of these biomarkers are utilized to build an optimized artificial neural network for the diagnosis of OSCC.

3.1. Biomarker Salivary Database

The dataset used in this study was the same used by and publicly available from [43]. The research volunteers were adults over 18 years old, treated at the Oral Medicine Outpatient Clinic of the Department of Biosciences and Oral Diagnosis of the Institute of Science and Technology of São José dos Campos (ICT-UNESP). All volunteers signed an informed

consent form (ICF) to participate in the research approved by the ethics committee under Protocol Number 1.033.312/2015 PH/CEP.

The database comprises data from chemical analyses of salivary samples obtained from 68 people, of which 27 were in the group with a positive diagnosis for OSCC and 41 people comprised the control group. Among the two groups of people, information on smokers, ex-smokers, and non-smokers was also collected and available. Table 1 presents a summarized description of the demographic and smoking profile of the 68 study participants of this database.

For the composition of the group of individuals with OSCC, age over 18 years and concomitant diagnosis of OSCC were considered as the inclusion criteria, disregarding patients who had already undergone oncological treatment, such as surgery, radiotherapy, or chemotherapy for any type of cancer, including oral cancer. For patients in the control group, the criterion was age over 18 years, without a diagnosis of malignant neoplasm in life, and volunteering for research.

Table 1. Concise demographic and smoking profile description of the 68 study participants.

| Samples | Smoker | | | $\mu \pm \sigma$ | Male | Total |
|---------|--------|--------|----|-------------------|----------|-------|
| | Yes | Former | No | | | |
| Control | 8 | 13 | 20 | 57.34 ± 11.66 | 21 (51%) | 41 |
| OSCC | 10 | 10 | 7 | 57.00 ± 13.87 | 19 (70%) | 27 |
| Total | 18 | 23 | 27 | 57.20 ± 12.54 | 40 (59%) | 68 |

From each participant, a salivary sample of 300 μL was collected. The sample was mixed with 300 μL of ethanol, shaken for 30 s, and then, centrifuged. Over each salivary sample, 51 metabolic biomarkers were extracted from gas chromatography–mass spectrometry.

The biomarkers were as follows: Acetoacetic.acid, Adenine, Arabitol, Batyl.alcohol, Cadaverine, Caproic.acid, Citramalic.acid, Cysteine, Dopamine, Eicosapentaenoic.acid, Elai-dic.acid, Galactosamine, Galactose, Glucono.1.5.lactone, Glucosamine, Glucose, Glucuronic.acid, Glutamic.acid, Glycine, Glycolic.acid, Hydroxylamine, Inositol, Isoleucine, Linoleic.acid, Maleic.acid, Mannitol, Mannose, Margaric.acid, Meso.Erythritol, Methionine.sulfone, Norepinephrine, Octadecanol, Oleic.acid, O.Phosphoethanolamine, Ornithine, Palmitic.acid, Palmitoleic.acid, Phenylalanine, Putrescine, Ribitol, Sarcosine, Threonine.acid, Threonine, Tryptamine, Tyramine, Uridine, X2.Aminoethanol, X2.Deoxy.gluco-se, X2.Hydroxyisovaleric. acid, X3.Aminoglutaric.acid, X3.Aminopropanoic.acid.

3.2. Data Augmentation

As described in Section 3.1, our original database (Orig-DB) contained 68 samples of control and OSCC cases, with more prevalence of samples of control cases (41) in contrast to OSCC cases (27). Thus, to overcome this imbalance between these classes, we also evaluated a more-extensive dataset in which the OSCC minority class was synthetically increased by the synthetic minority oversampling technique (SMOTE) to promote the balance between the classes. As a result, in this new dataset (Aug-DB), the total number of samples was 82, with an equal number of control and OSCC cases.

SMOTE is a data-augmentation technique that performs resampling by taking the sample neighborhood as a reference and creating a new instance connecting to this neighborhood. He et al. [44] gave the following definition for this technique:

“SMOTE first selects a minority class instance ‘a’ at random and finds its ‘k’ nearest minority class neighbors. The synthetic instance is then created by choosing one of the ‘k’ nearest neighbors ‘b’ at random and connecting ‘a’ and ‘b’ to form a line segment in the feature space. The synthetic instances are generated as a convex combination of the two chosen instances ‘a’ and ‘b’.”

When generating our synthetic samples, we used a value of $k = 5$ in SMOTE. This value was defined empirically, considering the number of original samples available (27), in order to balance the important information and have better sample diversity. Thus, we sought to improve the generalization of the model, avoiding both overfitting and underfitting.

To understand the effect of the resampling by neighborhood performed by the SMOTE technique, in Figure 2, we provide the plots of the t-distributed stochastic neighbor embedding (t-SNE) [45] of the original Biomarker Salivary Database and after the data augmentation by SMOTE was performed.

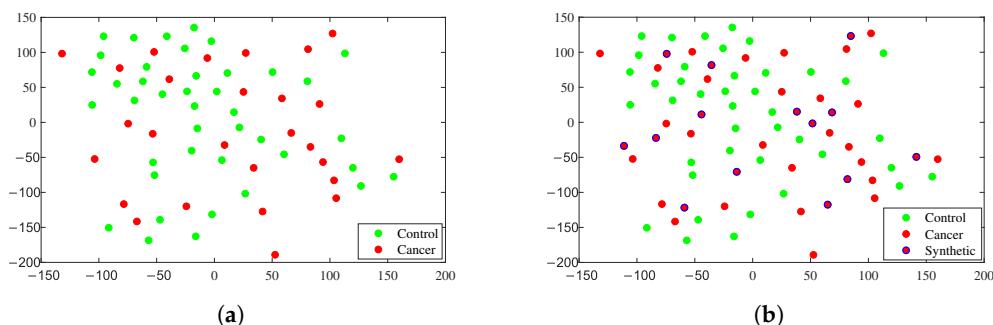


Figure 2. t-SNE plots showing the distribution and dispersion of samples between classes used for OSCC diagnosis. In (a), the resulting t-SNE plot with Euclidean distance over the original data distribution samples (Orig-DB). In (b), the resulting t-SNE plot with Euclidean distance over the SMOTE-augmented synthetic data samples (Aug-DB). Synthetic samples were obtained by SMOTE ($k = 5$) from the original samples and are highlighted with a blue border.

These t-SNE plots were generated from normalized data by the Z-score using the well-known Euclidean distance [46] and show the sample distribution in the Orig-DB and in the Aug-DB, highlighting how the new synthetic samples generated by SMOTE were positioned. These plots also help explain the reason for evaluating the balance effect between classes, taking the similarity of the samples between the classes. Furthermore, for better understanding, we invite the reader to see the original work by Nitesh Chawla et al. [47] with a detailed description of the SMOTE technique.

3.3. Ranking and Feature Selection

Feature selection aims to reduce the number of input variables in a learning model and optimize model performance using those with the most-significant relevance. Some benefits include less model development and training time, more-efficient use of computational resources, and the avoidance of model overfitting. In addition, keeping irrelevant features can also result in models with low generalization power.

In this work, the SVMAttributeEval algorithm was used as a solution for this task, which evaluates the value of a feature using an SVM classifier, which then classifies it by the square of the weight assigned by the SVM. If we have a multiple-class model, the algorithm ranks each class's features separately using a one-vs.-all (OvA) strategy and, then, “negotiating” from the top of each stack to obtain a final ranking. A more-detailed explanation of this algorithm is given in [48]. The algorithm parameter values that we used in this work were as follows:

- Constant elimination threshold = 0;
- Filtering used by SVM = normalization of training data;
- Tolerance $T = 1.0^{-10}$;
- Complexity $C = 1.0$;
- Epsilon $P = 1.0^{-25}$;
- Constant rate of feature elimination = 0;
- Feature elimination percentage rate = 0;
- Evaluator check = yes.

Notably, in both the Orig-DB and the one Aug-DB by SMOTE, we used the *SVMAtributeEval* algorithm. Table 2 shows the ranking position of each biomarker in the databases. Figure 3 shows the violin plot (normalized values by the Z-score technique) of the 15 highest-ranked features from the Orig-DB, while Figure 4 is their correlation heatmap, where darker colors indicate less dependency between features.

Table 2. Ranking of biomarkers using the *SVMAtributeEval* algorithm in the original database (Orig-DB) and SMOTE-augmented database (Aug-DB). The last column is the ranking difference between the Orig-DB and Aug-DB and shows the impact of synthetic samples by SMOTE on the ranking.

| Rank | Orig-DB | Aug-DB (SMOTE) | Difference |
|------|---------------------------|---------------------------|------------|
| 1 | X2.Hydroxyisovaleric.acid | Baty.alcohol | -4 |
| 2 | O.Phosphoethanolamine | Threonic.acid | -5 |
| 3 | Baty.alcohol | Maleic.acid | 2 |
| 4 | Threonic.acid | Threonine | 2 |
| 5 | Margaric.acid | X2.Hydroxyisovaleric.acid | -5 |
| 6 | X2.Deoxy.glucose | X2.Deoxy.glucose | 0 |
| 7 | Threonine | O.Phosphoethanolamine | 3 |
| 8 | Adenine | Adenine | 0 |
| 9 | Tyramine | Oleic.acid | -6 |
| 10 | Maleic.acid | Margaric.acid | 7 |
| 11 | Tryptamine | Phenylalanine | -20 |
| 12 | Putrescine | Ornithine | -9 |
| 13 | Norepinephrine | X2.Aminoethanol | -1 |
| 14 | Hydroxylamine | Norepinephrine | -9 |
| 15 | Dopamine | Tyramine | -35 |
| 16 | Caproic.acid | Cadaverine | -30 |
| 17 | Mannitol | Citramalic.acid | -7 |
| 18 | X2.Aminoethanol | Palmitoleic.acid | 5 |
| 19 | Cadaverine | Sarcosine | 3 |
| 20 | Eicosapentaenoic.acid | X3.Aminopropanoic.acid | -9 |
| 21 | Isoleucine | Putrescine | -13 |
| 22 | Palmitoleic.acid | Elaidic.acid | 4 |
| 23 | Methionine.sulfone | Hydroxylamine | -5 |
| 24 | Oleic.acid | Mannitol | 15 |
| 25 | Elaidic.acid | Glutamic.acid | 3 |
| 26 | Phenylalanine | Glycine | 15 |
| 27 | Citramalic.acid | Galactose | 10 |
| 28 | Sarcosine | Methionine.sulfone | 9 |
| 29 | X3.Aminopropanoic.acid | Eicosapentaenoic.acid | 9 |
| 30 | Arabitol | Inositol | -19 |
| 31 | Glucose | Tryptamine | -4 |
| 32 | Uridine | Mannose | -9 |
| 33 | Mannose | Linoleic.acid | 1 |
| 34 | Galactosamine | Isoleucine | -4 |
| 35 | Linoleic.acid | Glucose | 2 |
| 36 | X3.Aminoglutaric.acid | Palmitic.acid | -4 |
| 37 | Ornithine | Ribitol | 25 |
| 38 | Ribitol | Galactosamine | 1 |
| 39 | Glycine | Acetoacetic.acid | 13 |
| 40 | Glutamic.acid | X3.Aminoglutaric.acid | 15 |
| 41 | Acetoacetic.acid | Uridine | 2 |
| 42 | Glucosamine | Octadecanol | -9 |
| 43 | Glucono.1.5.lactone | Glucono.1.5.lactone | 0 |
| 44 | Galactose | Glucuronic.acid | 17 |
| 45 | meso.Erythritol | meso.Erythritol | 0 |
| 46 | Inositol | Caproic.acid | 16 |
| 47 | Glycolic.acid | Cysteine | -1 |
| 48 | Glucuronic.acid | Glycolic.acid | 4 |
| 49 | Palmitic.acid | Arabitol | 13 |
| 50 | Cysteine | Dopamine | 3 |
| 51 | Octadecanol | Glucosamine | 9 |

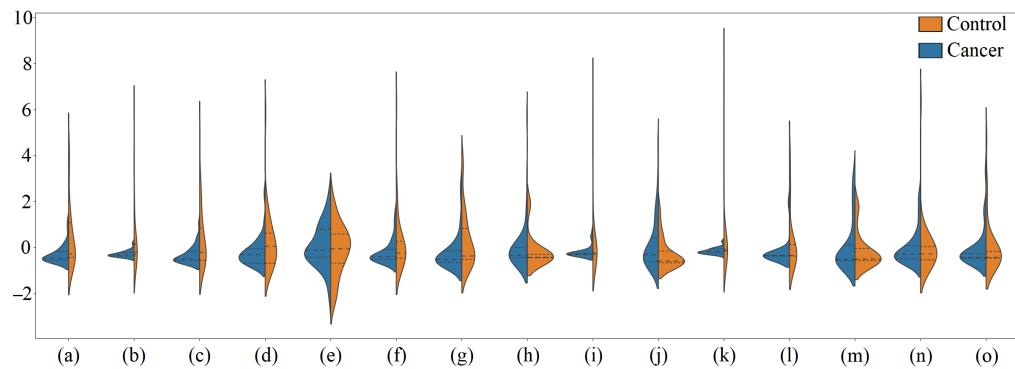


Figure 3. Violin plot illustrating the normalized distribution (by Z-score) of the values of the top 15 features in each class: (a) X2.Hydroxyisovaleric.acid, (b) O.Phosphoethanolamine, (c) Batyl.alcohol, (d) Threonic.acid, (e) Margaric.acid, (f) X2.Deoxy.glucose, (g) Threonine, (h) Adenine, (i) Tyramine, (j) Maleic.acid, (k) Tryptamine, (l) Putrescine, (m) Norepinephrine, (n) Hydroxylamine, (o) Dopamine.

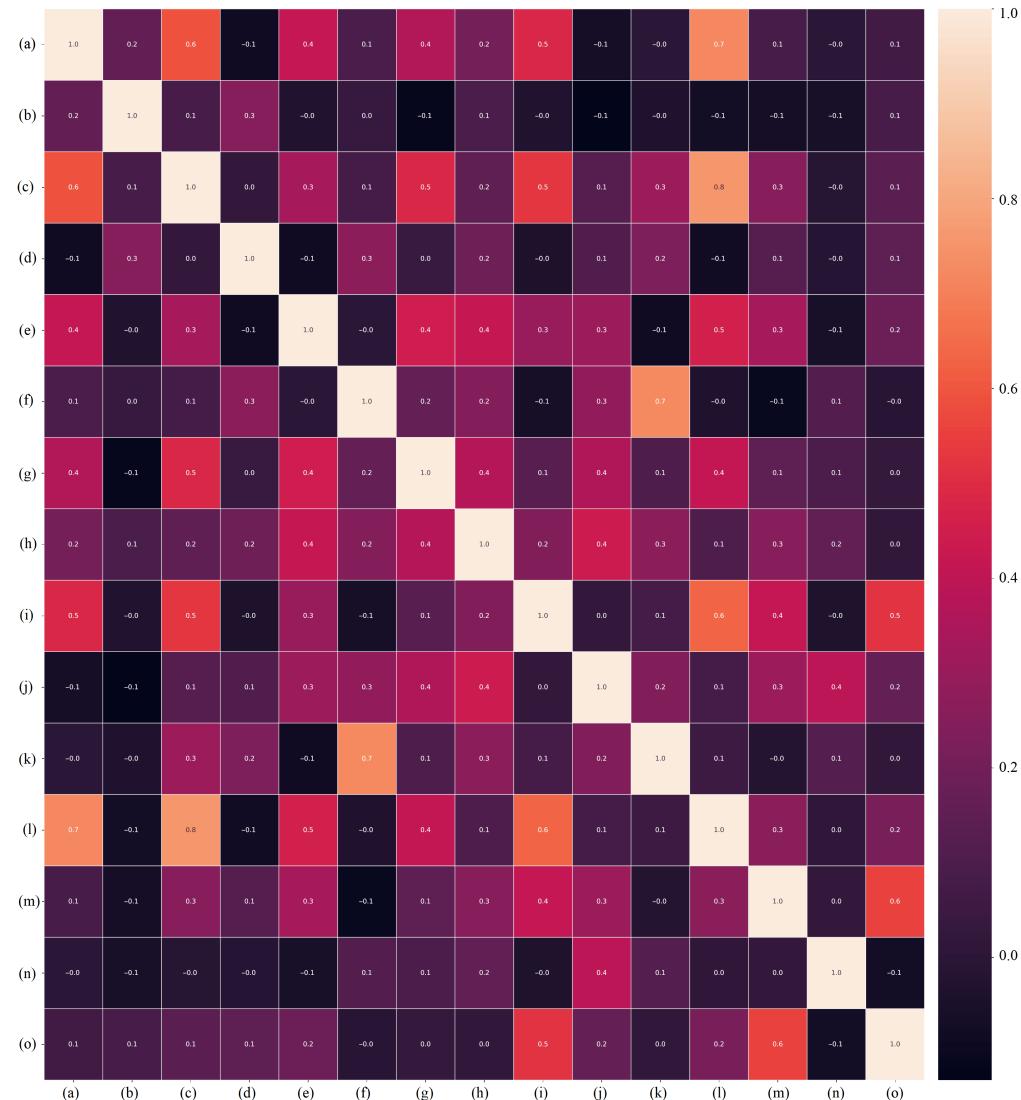


Figure 4. The correlation heatmap between the 15 highest-ranked biomarkers: (a) X2.Hydroxyisovaleric.acid, (b) O.Phosphoethanolamine, (c) Batyl.alcohol, (d) Threonic.acid, (e) Margaric.acid, (f) X2.Deoxy.glucose, (g) Threonine, (h) Adenine, (i) Tyramine, (j) Maleic.acid, (k) Tryptamine, (l) Putrescine, (m) Norepinephrine, (n) Hydroxylamine, (o) Dopamine. Light colors indicate high correlation, while darker colors indicate less dependence.

3.4. Artificial Neural Network

The success of neural networks in different fields of knowledge has placed them as the first option for solving several [49] problems. However, despite their success, most neural networks, especially the so-called deep neural networks, present characteristics of requiring equipment with high computational power, the need for a large volume of data, and much time for training. Thus, these characteristics make their use in solutions designed for environments and equipment with limited resources unfeasible. In this sense, feed-forward neural networks (FFNNs), the first and simplest category of artificial neural networks developed [50], represent a good alternative because they present excellent performance and efficiency in balance with the required resources. In these networks, connections between nodes do not form a loop, so information only moves in one direction—fowards—from input layer nodes, through hidden layer nodes, to output data layer nodes [51].

In this study, we chose these networks for their great predictive capacity, independent of the probability distribution information of the data. Furthermore, these networks have been successfully applied in a wide range of medical applications, including Alzheimer's disease [52], chronic kidney disease [28], lung cancer detection [53], and the prediction of outbreaks of COVID-19 [54,55].

The parallel structure and the ability to improve performance through experience explain the efficiency of these networks. Therefore, we built a shallow FFNN to support diagnosing oral squamous cell carcinoma cases among normal (healthy/control) cases. The proposed network was built with a minimalist structure, delivering good accuracy and requiring as few computational resources as possible. It consists of fully connected layers. At the network entrance, a set of biomarkers is passed through two hidden layers, each with 3 neurons, and then arrives at the output layer, which determines two possible classes: cancer and control. The schematic of the constructed network is shown in Figure 1.

Other network parameters include the Levenberg–Marquardt backpropagation training function [56,57], which updates the weights (w) and bias (b); the Softmax transfer function in the output layer nodes; a maximum number of epochs equal to 50; and a maximum number of validation failures equal to 5 to prevent overfitting. Other model optimization parameters include the mean-squared error (MSE) loss function with the regularization rate equal to 1×10^{-5} , the hyperbolic tangent sigmoid (Tansig) on the hidden layer nodes to produce faster output rates [58], and cross-validation for the batch size. Table 3 summarizes the optimizing parameters of our models.

Table 3. Overview of optimized parameters of our FFNN model.

| Parameter | Value | Explanation |
|---------------------|-----------------------|---|
| Input | From 1 to 51 features | Our biomarker database ranked by the <i>SVMAttributeEval</i> algorithm |
| Learning function | Levenberg–Marquardt | Used to update weight and bias values according to Levenberg–Marquardt optimization |
| Loss function | Mean-squared error | Measures model performance by calculating the mean-squared error between estimated values and actual values |
| Regularization rate | 1×10^{-5} | For faster model convergence while preserving data representativeness |
| Number of neurons | 3 | Optimized with only 3 neurons in the hidden layer |
| Training process | 10-CV/LOOCV | Cross-validation to improve model generalizability and estimate performance in practice |
| Transfer function | Tansig Softmax | Used in the hidden layer nodes to produce faster output rates Used in the output layer nodes to provide better correlation coefficients for the processed hidden layer output data |
| Validation failures | 5 | Avoiding overfitting of the model |

3.5. Data Splitting and Model Validation

In a neural network performance validation, a model is usually trained with available data, called a training set. Then, the model's inference power is evaluated using newly collected data or a set of previously separated data unseen by the model. The latter is called a test set.

In this study, we used the cross-validation (CV) technique to measure the generalization power of the proposed model. This technique provides a good measure of generalization by splitting samples from a dataset into mutually exclusive subsets, which are then used in the training and testing rounds. Cross-validation, thus, assesses how the results of a statistical analysis of the model generalize to an independent dataset. In problems where the modeling objective is prediction, this technique is widely used [59].

In the validation process of this study, two different cross-validation schemes were used, the first scheme with a dataset divided into k exclusive subsets— k -fold cross-validation ($k = 10$)—with nine subsets being used for training and the other used for testing, in a process that was repeated ten times, alternating each time the training and test sets. In the second scheme, a particular case of k -fold cross-validation called leave-one-out cross-validation (LOOCV), we divided the dataset by the number of available samples ($k = n$). Thus, considering n available samples, each evaluation run uses a training set with $n - 1$ samples, and the sample left out is used for testing, repeating the process n times. For both schemes, the reported results always correspond to the average obtained in each test case.

3.6. Evaluation Metrics

Public databases related to OSCC are rare in the literature, making it difficult to apply more-current techniques, such as deep learning, which require a large volume of data. Therefore, comparing these more-current and classic techniques becomes a challenge in evaluating the significance and weight of the results. Thus, the cross-validation schemes used in this study and described in Section 3.5 are viable and valuable alternatives for evaluating the results.

In this study, we considered the following evaluation metrics: *accuracy* (ACC), which represents the rate of correct results among the total number of cases examined; *sensitivity* (TPR), which represents the rate of individuals belonging to the positive class (individuals with OSCC) who were correctly identified as OSCC, and *specificity* (TNR), which measures the percentage of people belonging to the negative class (healthy control patients) who were correctly identified.

Thus, after the completion of all test cases for both cross-validation schemes, the average results of the accuracy, sensitivity, and specificity were calculated as follows:

$$\text{Accuracy (ACC)} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Sensitivity (TPR)} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Specificity (TNR)} = \frac{TN}{TN + FP} \quad (3)$$

where true positive (TP) is the number of OSCC cases correctly identified as such, true negative (TN) is the number of correctly identified control cases, false positive (FP) is the number of control cases incorrectly labeled as OSCC cases, and false negative (FN) is the number of OSCC cases that were labeled as control cases.

Other metrics considered in this study included the confusion matrix, ROC curves, regression analysis, and area under the ROC curve (AUC).

3.7. Setup Configuration

The experiments were carried out on a desktop computer with a Core™ i5-7500 CPU @3.4GHz, 8 GB of RAM, Windows 10—64-bit, without GPU acceleration. We used

the Matlab R2014a program with the packages Deep Learning Toolbox and Signal Processing Toolbox installed to build the network. The other settings adopted were the environment's default.

4. Discussion and Analysis of Results

4.1. Feature Ablation

To define the minimum amount of biomarkers/features needed to build intelligent artificial neural network models to assist in OSCC diagnosis, we first used the *SVMAtributeEval* algorithm both in the Orig-DB and in the one Aug-DB by SMOTE, ranking the biomarkers as described in Section 3.3.

Thus, after ranking the features, we evaluated the model's accuracy in each validation scheme, starting from a subset with only the first best-ranked feature for incremental subsets to the subset with all 51 biomarkers.

Figures 5 and 6 show, respectively, the model accuracy achieved for each feature subset under the 10-CV and LOOCV validation schemes in the Orig-DB and Aug-DB databases, from which we can arrive at the following observations:

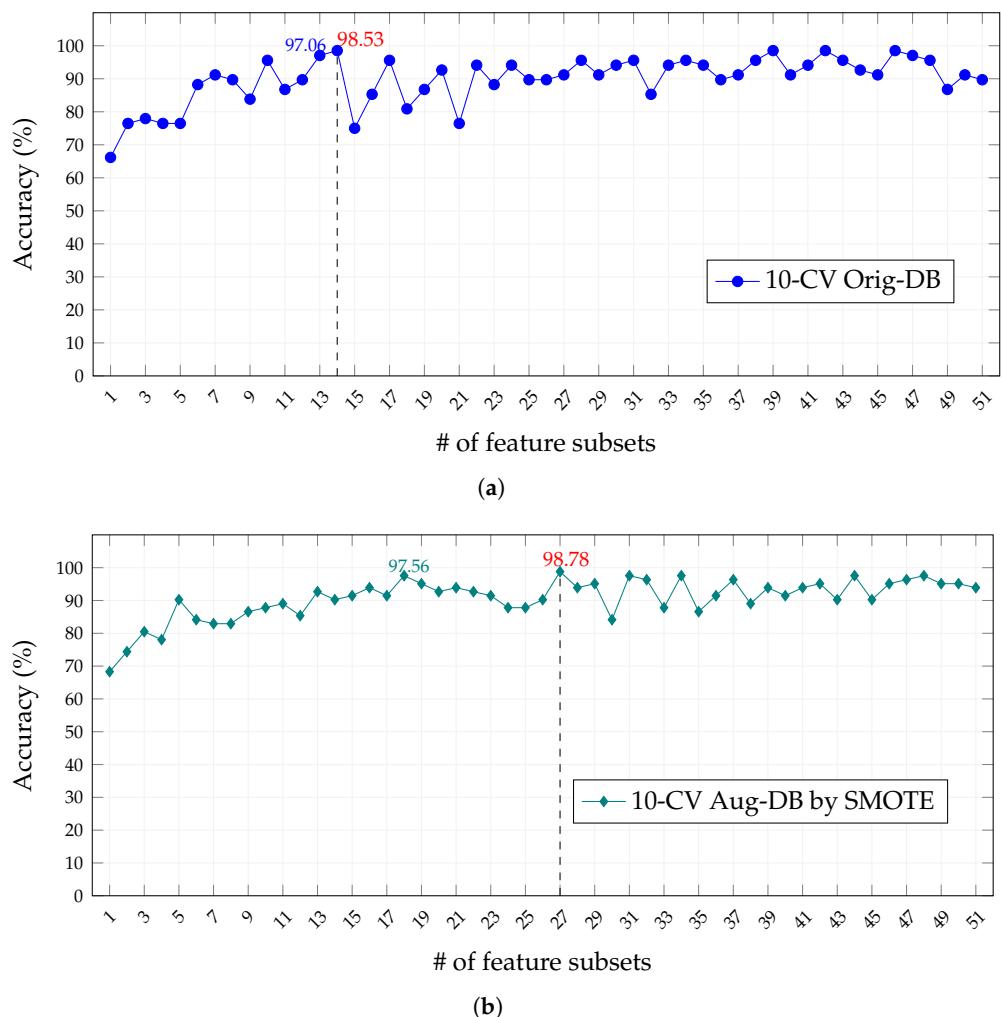


Figure 5. Accuracy performance as a function of selected features for the 10-fold cross-validation (10-CV) scheme in (a) Results of classification by 10-CV with samples from Orig-DB and in (b) Results of classification by 10-CV with samples from Aug-DB. Red values indicate the best accuracy achieved.

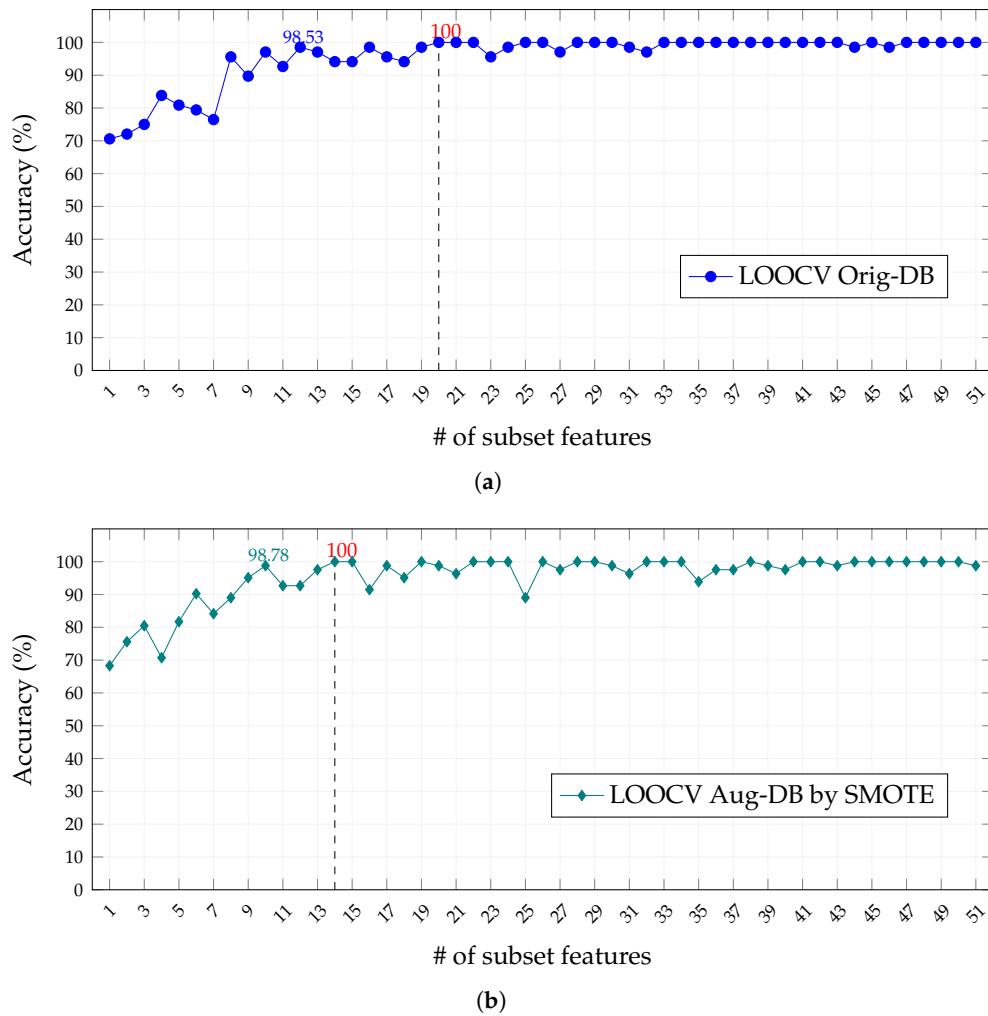


Figure 6. Accuracy performance as a function of selected features for the leave-one-out (LOOCV) scheme in (a) LOOCV classification results with Orig-DB samples and in (b) LOOCV classification results with Aug-DB samples. Red values indicate the best accuracy achieved.

- (i) The classification of features by the SVMAttributeEval algorithm does not guarantee better accuracy for models built with subsets with a more-significant number of features, and the oscillations of the accuracy values, more evident in the 10-CV validation scheme of Figure 5, may be related to the fact that this ranking considers the importance of the characteristics for an SVM classifier and not specifically for the FFNN used in the models;
- (ii) The balancing of cancer and control classes in the Aug-DB by the SMOTE resampling technique does not result in an improvement in accuracy and, when using the 10-CV validation scheme, requires a more-significant number of features to achieve similar accuracy to the Orig-DB with unbalanced classes, while for the LOOCV validation scheme, the opposite occurs;
- (iii) The best performance, least number of features, and highest accuracy were obtained by the LOOCV scheme on the Aug-DB by SMOTE (Figure 6b), explained by the lower selection bias, favored even more by the balance of classes and by the lower tendency to overestimate test errors, as it was trained $n-1$ times;
- (iv) Finally, the FFNN models were efficient for all schemes, reaching expressive values of accuracy with a reduced number of features (above 90% with less than ten features).

Furthermore, with regard to the validation schemes, although the reduced dataset size and performance justify the use of the LOOCV scheme, we must keep in mind that it is subject to greater variance and overfitting, high computational cost, and reduced interpretability.

Therefore, we reinforce that the choice of the best training scheme must carefully consider the advantages and disadvantages and the suitability for the task to be performed.

That being stated, after evaluating the models for each training scheme and dataset and considering the different feature subsets, we selected the feature subset with the best accuracy result in each configuration. This means 14 and 27 features for the Orig-DB and 20 and 14 features for the Aug-DB, respectively, within the context of the 10-CV and LOOCV schemes. In the following sections of this work, we present additional results from each of these configurations.

4.2. Performance Model for Different Validation Schemes

Two different cross-validation schemes were used to assess the model's generalization, as described in Section 3.5. These schemes were tested on the Orig-DB and Aug-DB by SMOTE. This evaluation of the different validation schemes and datasets aimed to point out the impact of their choices and the advantages or not of adopting one option over another, as already initially discussed in Section 4.1.

Accordingly, we now report more-detailed results for each of these configurations, considering only the subset of biomarkers/traits that performed better in the evaluated configuration (cross-validation scheme and dataset).

In Figure 7, we present the confusion matrices for each evaluated configuration, including the cross-validation schemes (10-CV and LOOCV) and the databases (Orig-DB and Aug-DB). These matrices illustrate that the model maintained stable performance across different configurations, consistently achieving accuracy above 98.5% and a sensitivity (TPR) of at least 96.3%.

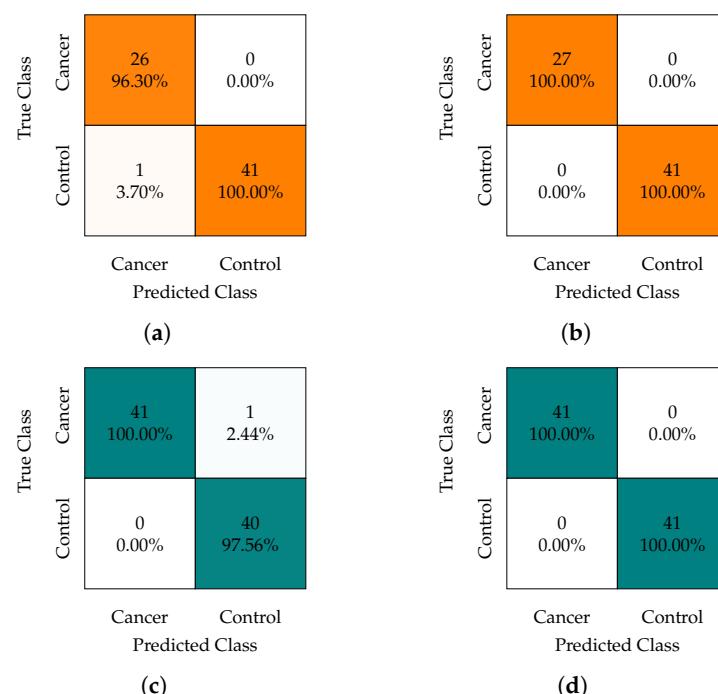


Figure 7. Confusion matrices obtained from the best result achieved by our artificial neural network model in the 10-fold cross-validation (10-CV) and leave-one-out cross-validation (LOOCV) schemes. Subfigures (a,b) of the original database (Orig-DB) in the 10-CV and LOOCV schemes, with 14 and 20 features, respectively. Subfigures (c,d) of the SMOTE-augmented database (Aug-DB) in the 10-CV and LOOCV schemes with 27 and 14 features, respectively.

With regard to the impact of the sample balancing promoted by SMOTE, it led to an improvement in the TPR (from 96.30% to 100% in the 10-CV scheme) and can reduce the number of features required for the model to achieve its best accuracy (from 20 to 14 in the LOOCV scheme). In terms of cross-validation schemes, the LOOCV technique

demonstrated superior accuracy performance compared to 10-CV, albeit with the trade-off of longer training times, which becomes impractical for larger datasets.

Some of these observations can be attributed to the imbalance ratio of 1:1.5 (27:41) in our original database and the utilization of the SMOTE approach. As discussed in [60], the SMOTE approach mitigates the variance of the minority class, leading to an improved TPR, but with a relatively modest impact on the overall classifier performance.

Thus, in light of these findings, we can conclude that the choice of the cross-validation scheme and the use of the data augmentation in our study had minimal influence on the results of the confusion matrices. However, other factors may be considered, as we will see below.

The ROC curves, displayed in Figure 8, illustrate the model's diagnostic capability as the discriminant threshold varies. Notably, the ROC curves exhibited a consistent behavior, echoing the confusion matrices' results. For instance, in the 10-CV scheme, we observed a higher AUC value of 0.993 on the Aug-DB (Figure 8c) compared to the 0.965 value on the Orig-DB (Figure 8a), while for the LOOCV scheme, there was no difference in the value of the AUC for the database used.

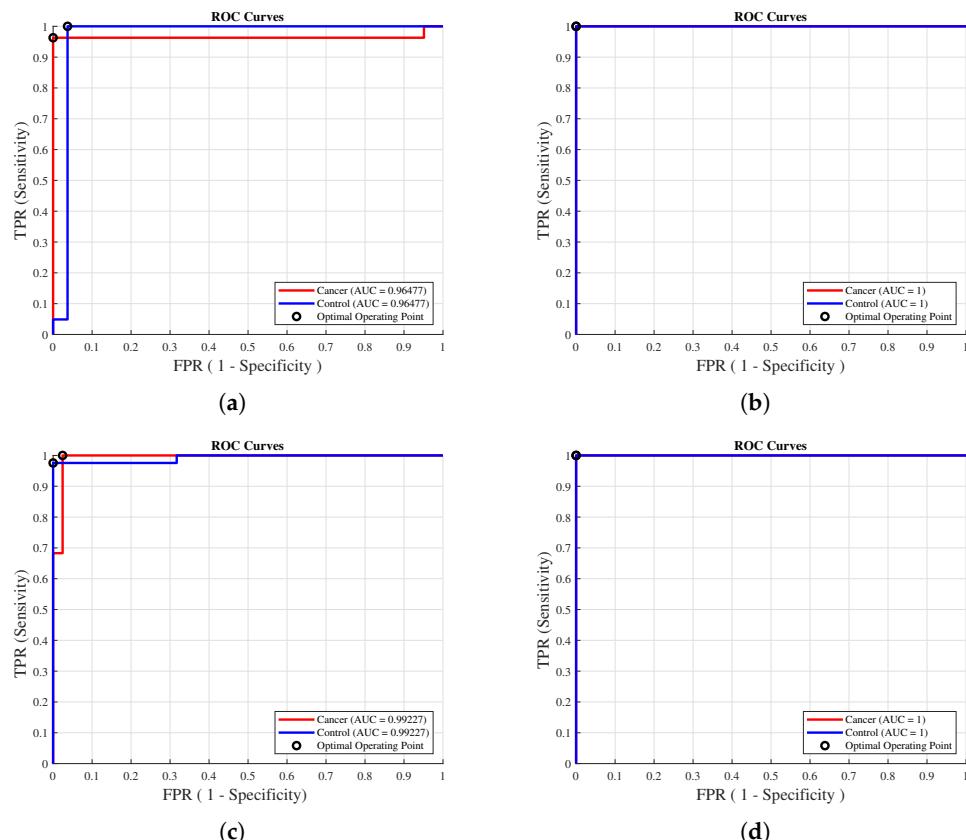


Figure 8. ROC curves of the model with the best subset of features in each training scheme. Subfigures (a,b) of the original database (Orig-DB) in 10-CV and LOOCV schemes, with 14 and 20 features, respectively. Subfigures (c,d) of the SMOTE-augmented database (Aug-DB) in the 10-CV and LOOCV schemes, with 27 and 14 features, respectively.

These ROC curves are of fundamental importance in the interpretation of the results, since their resilience against the potential influence of the class distribution and balance makes them robust indicators of the performance of our model. Furthermore, they have a direct and intrinsic connection with the cost–benefit analysis, a fundamental aspect of diagnostic decision-making.

In this regard, the consistency demonstrated by the curves in Figure 8 with optimal operating points close to 1 across the various configurations we assessed serves as compelling evidence of our model's exceptional discriminative power in the presented context.

Figure 9 brings a view of the model's linear regression analysis. Scatter plots between the actual and predicted classes show a strong model correlation coefficient, indicating that the model prediction accurately follows the actual class and achieves a correlation coefficient of at least 0.953 (Figure 9a). Another highlight in these plots is the fit line, which practically overlaps the reference line ($Y = T$) for all configurations, pointing to a model with an excellent fit to the data, regardless of the validation scheme or dataset used.

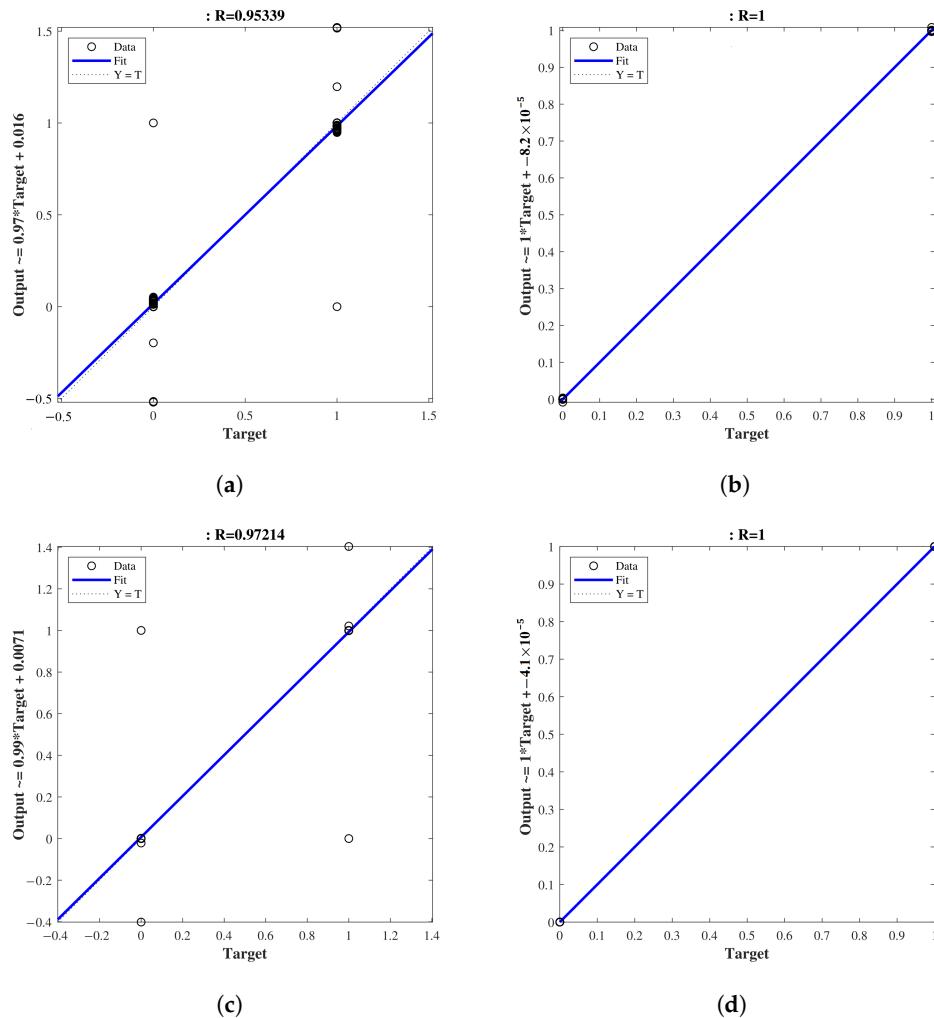


Figure 9. Scatter plots between the actual and predicted classes in the linear regression analysis of the model. In each configuration, the results show a strong correlation coefficient. (a) Orig-DB (10-CV). (b) Orig-DB (LOOCV). (c) Aug-DB (10-CV). (d) Aug-DB (LOOCV).

Figure 10 presents sample error curves of the training process for 10-CV and LOOCV with the original data (Orig-DB) and balanced data with SMOTE (Aug-DB). The training was set up to stop when the MSE increased or reached 50 epochs. As can be seen, models trained without augmentation stopped at 50 epochs, reaching errors of 3.1298×10^{-8} and 2.6377×10^{-5} for 10-CV and LOOCV, respectively, while models trained with data augmentation stopped at 21 and 26 epochs, reaching errors of 1.7284×10^{-12} and 3.7319×10^{-13} , for 10-CV and LOOCV, respectively. Therefore, the data augmentation was a great improvement in terms of training, since significantly fewer MSEs were reached in fewer epochs, reducing the risk of overfitting.

Finally, Figure 11 shows that the gradient, which represents the slope of the tangent of the graph of the function or simply the network error rate, converged more quickly in

the Aug-DB, as well as presenting a lower μ (mean error), which represents the control parameter for the backpropagation neural network that we modeled. Finally, the validation check reflected the error in each cycle pass and was zero for all schemes, indicating good network performance, as seen in the figures.

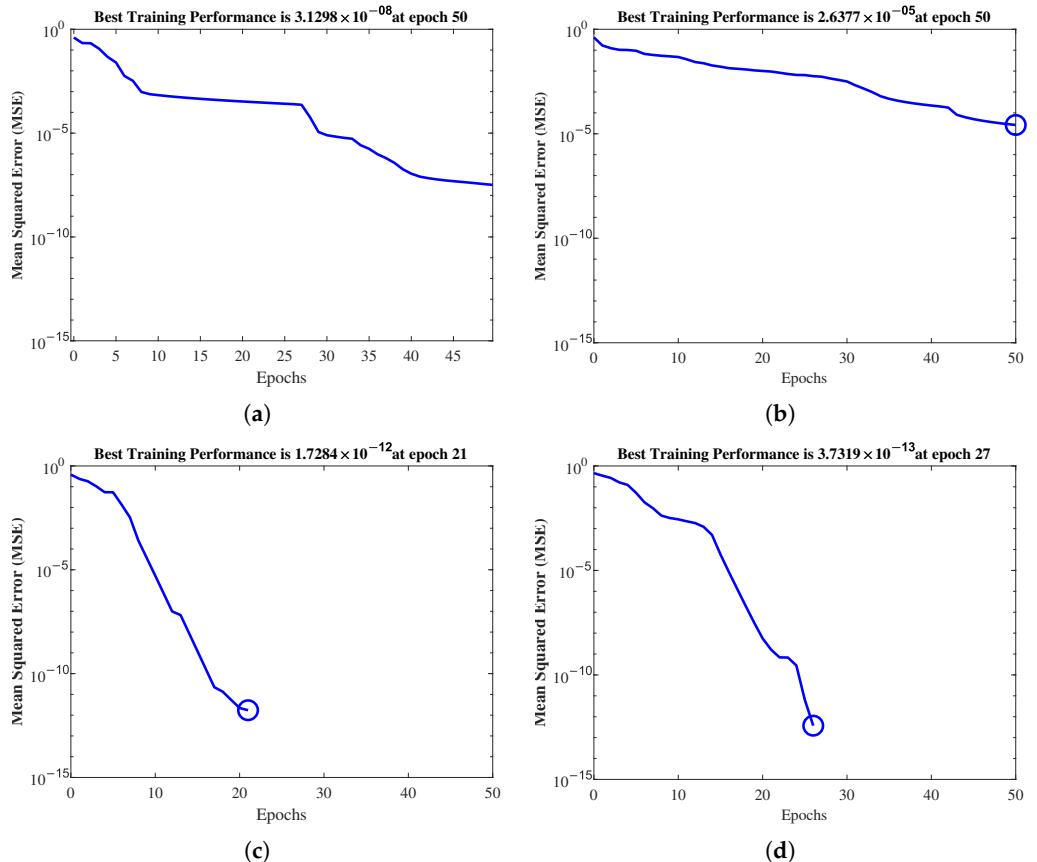


Figure 10. Model error curves in the training process in each configuration. The lower MSE when using the SMOTE-augmented database (Aug-db) in (c,d) suggests better-fitting models. (a) Orig-DB (10-CV). (b) Orig-DB (LOOCV). (c) Aug-DB (10-CV). (d) Aug-DB (LOOCV).

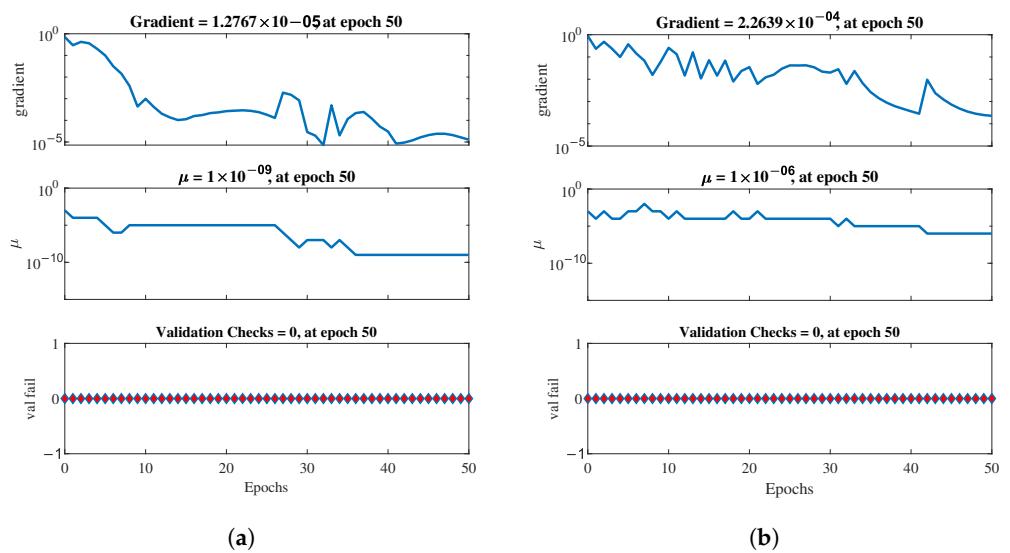


Figure 11. Cont.

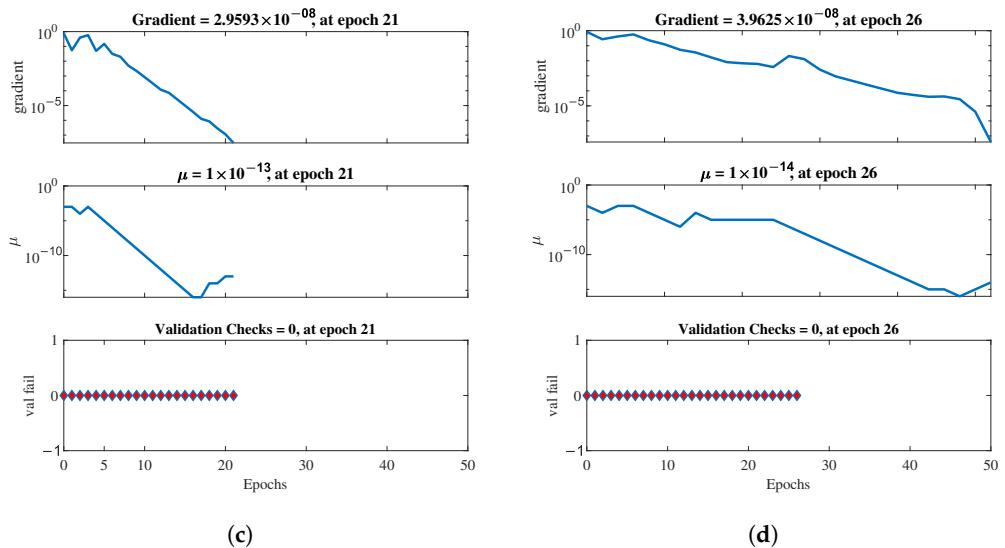


Figure 11. Plots representing a snapshot of the training state of our FFNN for each of the different configurations evaluated. (a) Orig-DB (10-CV). (b) Orig-DB (LOOCV). (c) Aug-DB (10-CV). (d) Aug-DB (LOOCV).

4.3. Comparison with the Literature

Table 4 presents a brief quantitative comparison considering our proposal and some literature studies on diagnosing OSCC cases. Although some works have used different configurations (cross-validation scheme and database) and experimental evaluations, they also have common characteristics, such as measuring the AUC performance and using salivary biomarkers. This comparison is interesting to position our work concerning peers and reveals the principal methodologies used in approaching the diagnosis of OSCC cases. In this sense, our proposal presents, in general terms, better results than the compared studies, both for those that used the same database and those that used other databases. For example, in Costa et al. [43] and Fonseca et al. [61], the same public database was used in a 10-CV validation scheme, while Song et al. [12] used another public database but on a 20-CV validation scheme. The other studies used private databases combining LOOCV validation schemes or statistical analyses.

Thus, in contrast to the compared studies, our proposal was evaluated under two different validation schemes while analyzing the impact of data balancing by SMOTE resampling. In addition, our proposal still ranked the biomarkers by taking the weights of a classified SVM and evaluated the smallest subset necessary to achieve the best performance and accuracy of the model, reaching a minimum ACC and AUC of 98.53% and 96.48% respectively, when using the unbalanced Orig-DB (10-CV) validation scheme.

Finally, we highlight the computational effort required for our model proposed here. The total time spent to perform all model training in the 10-CV validation scheme was only 1.07 s on both the Orig-DB and Aug-DB, while in the LOOCV validation scheme, these times were 5.7 s and 8.2 s, respectively. Furthermore, the average inference time of a sample for all configurations was approximately 2.85×10^{-4} s, and the amount of memory required for the entire model was approximately 6 MB for each analyzed configuration. These characteristics make it possible to use our proposal, regardless of the configuration adopted (database and validation scheme), in real-time solutions, environments, and equipment with limited resources.

Table 4. Comparative analysis of our OSCC diagnosis proposal with related literature studies.

| Year | Study | Method | Eval. | Test Results (%) | | | |
|--------------|--------------------------|--------------------------|-------|------------------|-------|-------|-------|
| | | | | ACC | TPR | TNR | AUC |
| 2008 | Hu et al. [18] | Logistic regression | LOOCV | - | 90.0 | 83.0 | 93.0 |
| 2014 | Wang et al. [62] | Logistic regression | LOOCV | - | 92.3 | 91.7 | 87.1 |
| 2014 | Yuvaraj et al. [15] | Statistical analysis | - | - | 85.7 | 93.3 | 94.2 |
| 2016 | Gleber-Netto et al. [63] | Fractional polynomial | - | - | - | - | 87.2 |
| 2019 | Deepthi et al. [64] | Statistical analysis | - | - | - | - | 98.9 |
| 2020 | Song et al. [12] | FS (Lasso regression) | 20-CV | 86.7 | - | - | - |
| 2020 | Ishikawa et al. [42] | FS (logistic regression) | - | - | - | - | 85.6 |
| 2021 | Kouznetsova et al. [32] | FS (InfoGain) + NN | 10-CV | 79.5 | - | - | - |
| 2022 | da Costa et al. [43] | FS (random forest) | 10-CV | 86.8 | 88.9 | 85.4 | 91.0 |
| 2022 | Fonseca et al. [61] | FFNN multi-layer | 10-CV | 94.1 | 92.6 | 95.1 | 97.6 |
| Our proposal | | | | | | | |
| 2023 | This work | FS (SVM) + FFNN | 10-CV | 98.5 | 96.3 | 100 | 96.5 |
| 2023 | This work | FS (SVM) + FFNN | LOOCV | 100 | 100 | 100 | 100 |
| 2023 | This work | FS(SVM) + FFNN + DA | 10-CV | 98.8 | 100.0 | 97.6 | 99.2 |
| 2023 | This work | FS(SVM) + FFNN + DA | LOOCV | 100.0 | 100.0 | 100.0 | 100.0 |

Eval.: classification evaluation scheme, ACC: accuracy, TPR: sensitivity, TNR: specificity, AUC: area under the ROC curve, LOOCV: leave-one-out cross-validation, k-CV: k-fold cross-validation, NN: neural network, FFNN: feed-forward neural network, FS: feature-selection process, DA data augmentation by SMOTE.

5. Conclusions

This study presented a shallow neural network model of the FFNN-type with low computational cost to aid in diagnosing OSCC cases. Using 51 biomarkers extracted from salivary samples from people diagnosed with OSCC and healthy people, we showed that we can effectively contribute to the early detection of OSCC cases with different numbers of these biomarkers. The analysis of a biomarker-selection process by the SVMAttributeEval algorithm was presented for four different configurations, combining two cross-validation schemes (10-CV and LOOCV) applied to an original dataset (Orig-DB) and an augmented dataset by the SMOTE technique (Aug-DB), which was used for balancing between classes. The feature-selection process yielded different subsets for each configuration, being 14 and 27 features for the Orig-DB and 20 and 14 features for the Aug-DB, respectively, within the context of 10-CV and LOOCV. The results for all configurations showed high values of the model's precision, sensitivity, and specificity. The proposal significantly surpassed previous studies in the literature both for those using the same database and those using other databases.

As a novelty, our model combines the low use of computational resources with high accuracy for the early diagnosis of OSCC with different numbers of salivary biomarkers. Furthermore, the validation in different configurations indicated that our model design, with fewer degrees of freedom, was effectively less prone to bias and overfitting. Thus, the minimalist characteristics make our proposal a viable and inexpensive solution for embedded systems, environments, and equipment with limited processing and storage capacity. Hence, it can effectively aid in early OSCC case detection and be used as a tool in the active search for OSCC cases in health control programs.

In this sense, we concluded that the use of salivary biomarkers combined with a shallow neural network model presents a great potential to assist in the decision-making of medical professionals, providing a more-assertive and early prognosis of OSCC cases and, with this, contributes to a better chance of cure and patient survival. As future advances, we point out tests in other bases of salivary biomarkers and the identification of the stage

of the disease. Furthermore, different optimization parameters can be experimented with, as well as other validation schemes and feature-selection techniques.

Author Contributions: Conceptualization, A.U.F. and J.P.F.; methodology, A.U.F.; software, A.U.F. and G.S.V.; validation, H.P., Y.C.M. and J.C.G.M.; formal analysis, Y.C.M. and J.C.G.M.; investigation, A.U.F.; resources, F.S.; data curation, Y.C.M. and J.C.G.M.; writing—original draft preparation, A.U.F.; writing—review and editing, A.U.F., J.P.F. and G.S.V.; visualization, A.U.F. and H.P.; supervision, F.S.; project administration, F.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Institute of Science and Technology of São José dos Campos (ICT-UNESP) (1.033.312/2015 PH/CEP).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study; since the manuscript was based on publicly available data from the study authors [43], which have been openly shared and published, we did not specifically require a consent form for their inclusion in our work.

Data Availability Statement: The dataset used in this study was the same as that available in [43].

Acknowledgments: This study was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code #001. We would like to thank and acknowledge partial support for this research from Federal University of Goiás (INF/UFG), Federal Institute Goiano (IF Goiano), General Hospital of Goiânia (HGG), and Araújo Jorge Cancer Hospital (HAJ).

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

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---------|---|
| Aug-DB | Augmented database by SMOTE resampling |
| ACC | Accuracy |
| AUC | Area under the curve of the receiver operating characteristic |
| CPSI-MS | Conductive polymer spray–mass spectrometry |
| CT | Computerized tomography |
| CV | Cross-validation |
| DA | Data augmentation process |
| FAD | Flavin adenine dinucleotide |
| FDG | F-fluorodeoxyglucose |
| FFNN | Feed-forward neural network |
| FN | False negative |
| FP | False positive |
| FS | Feature-selection process |
| HAJ | Araújo Jorge Cancer Hospital |
| HPV | Human <i>papillomavirus</i> |
| ICF | Informed consent form |
| LOOCV | Leave-one-out cross-validation |
| ML | Machine learning |
| MSE | Mean-squared error |
| NADH | Nicotinamide adenine dinucleotide, reduced |
| NADPH | Nicotinamide adenine dinucleotide phosphate, reduced |
| NN | Neural network |
| OLK | Oral leukoplakia |
| Orig-DB | Original database |
| OSCC | Oral squamous cell carcinoma |
| OvA | One-vs.-all classification strategy |
| PET/CT | Positron emission tomography |

| | |
|--------|---|
| Q-TOF | Quadrupole time-of-flight |
| ROC | Receiver operating characteristic |
| SVM | Support vector machine |
| SMOTE | Synthetic minority oversampling technique |
| Tansig | Hyperbolic tangent sigmoid |
| t-SNE | t-distributed stochastic neighbor embedding |
| TN | True negative |
| TNR | True negative rate or specificity |
| TP | True positive |
| TPR | True positive rate or sensitivity |
| UFG | Federal University of Goias |

References

1. Ferlay, J.; Colombet, M.; Soerjomataram, I.; Parkin, D.M.; Piñeros, M.; Znaor, A.; Bray, F. Cancer statistics for the year 2020: An overview. *Int. J. Cancer* **2021**, *149*, 778–789. [CrossRef] [PubMed]
2. Dewan, A.K.; Dabas, S.K.; Pradhan, T.; Mehta, S.; Dewan, A.; Sinha, R. Squamous cell carcinoma of the superior gingivobuccal sulcus: An 11-year institutional experience of 203 cases. *Jpn. J. Clin. Oncol.* **2014**, *44*, 807–811. [CrossRef] [PubMed]
3. Jovanovic, A.; Schulten, E.A.; Kostense, P.J.; Snow, G.B.; van der Waal, I. Tobacco and alcohol related to the anatomical site of oral squamous cell carcinoma. *J. Oral Pathol. Med.* **1993**, *22*, 459–462. [CrossRef]
4. Bruschini, R.; Maffini, F.; Chiesa, F.; Lepanto, D.; De Berardinis, R.; Chu, F.; Tagliabue, M.; Giugliano, G.; Ansarin, M. Oral cancer: Changing the aim of the biopsy in the age of precision medicine. A review. *Acta Otorhinolaryngol. Ital.* **2021**, *41*, 108. [CrossRef]
5. Adeyemi, B.F.; Olusanya, A.A.; Lawoyin, J.O. Oral squamous cell carcinoma, socioeconomic status and history of exposure to alcohol and tobacco. *J. Natl. Med. Assoc.* **2011**, *103*, 498–502. [CrossRef]
6. Wong, T.; Wiesenfeld, D. Oral cancer. *Aust. Dent. J.* **2018**, *63*, S91–S99. [CrossRef]
7. Brock, G. The diagnosis of early and premalignant lesions of oral squamous cell carcinoma. *Dent. Nurs.* **2010**, *6*, 132–134. [CrossRef]
8. Warnakulasuriya, S. Living with oral cancer: Epidemiology with particular reference to prevalence and life-style changes that influence survival. *Oral Oncol.* **2010**, *46*, 407–410. [CrossRef]
9. Abati, S.; Bramati, C.; Bondi, S.; Lissoni, A.; Trimarchi, M. Oral cancer and precancer: A narrative review on the relevance of early diagnosis. *Int. J. Environ. Res. Public Health* **2020**, *17*, 9160. [CrossRef]
10. Alabi, R.O.; Elmusrati, M.; Sawazaki-Calone, I.; Kowalski, L.P.; Haglund, C.; Coletta, R.D.; Mäkitie, A.A.; Salo, T.; Leivo, I.; Almangush, A. Machine learning application for prediction of locoregional recurrences in early oral tongue cancer: A Web-based prognostic tool. *Virchows Arch.* **2019**, *475*, 489–497. [CrossRef]
11. Bugshan, A.; Farooq, I. Oral squamous cell carcinoma: Metastasis, potentially associated malignant disorders, etiology and recent advancements in diagnosis. *F1000Research* **2020**, *9*, 229
12. Song, X.; Yang, X.; Narayanan, R.; Shankar, V.; Ethiraj, S.; Wang, X.; Duan, N.; Ni, Y.H.; Hu, Q.; Zare, R.N. Oral squamous cell carcinoma diagnosed from saliva metabolic profiling. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 16167–16173. [CrossRef] [PubMed]
13. Mafessoni, T.P.; Mazur, C.E.; Amenábar, J.M. Salivary lactate dehydrogenase (LDH) as a tool for early diagnosis of oral cancer in individuals with Fanconi anemia. *Med. Hypotheses* **2018**, *119*, 29–31. [CrossRef]
14. Connolly, J.M.; Davies, K.; Kazakeviciute, A.; Wheatley, A.M.; Dockery, P.; Keogh, I.; Olivo, M. Non-invasive and label-free detection of oral squamous cell carcinoma using saliva surface-enhanced Raman spectroscopy and multivariate analysis. *Nanomed. Nanotechnol. Biol. Med.* **2016**, *12*, 1593–1601. [CrossRef] [PubMed]
15. Yuvaraj, M.; Udayakumar, K.; Jayanth, V.; Rao, A.P.; Bharanidharan, G.; Koteeswaran, D.; Munusamy, B.D.; Murali krishna, C.; Ganeshan, S. Fluorescence spectroscopic characterization of salivary metabolites of oral cancer patients. *J. Photochem. Photobiol. B Biol.* **2014**, *130*, 153–160. [CrossRef] [PubMed]
16. Wang, Q.; Gao, P.; Wang, X.; Duan, Y. The early diagnosis and monitoring of squamous cell carcinoma via saliva metabolomics. *Sci. Rep.* **2014**, *4*, 6802. [CrossRef]
17. Wei, J.; Xie, G.; Zhou, Z.; Shi, P.; Qiu, Y.; Zheng, X.; Chen, T.; Su, M.; Zhao, A.; Jia, W. Salivary metabolite signatures of oral cancer and leukoplakia. *Int. J. Cancer* **2011**, *129*, 2207–2217. [CrossRef]
18. Hu, S.; Arellano, M.; Boontheung, P.; Wang, J.; Zhou, H.; Jiang, J.; Elashoff, D.; Wei, R.; Loo, J.A.; Wong, D.T. Salivary proteomics for oral cancer biomarker discovery. *Clin. Cancer Res.* **2008**, *14*, 6246–6252. [CrossRef]
19. Alabi, R.O.; Youssef, O.; Pirinen, M.; Elmusrati, M.; Mäkitie, A.A.; Leivo, I.; Almangush, A. Machine learning in oral squamous cell carcinoma: Current status, clinical concerns and prospects for future—A systematic review. *Artif. Intell. Med.* **2021**, *115*, 102060. [CrossRef]
20. Huang, S.; Yang, J.; Fong, S.; Zhao, Q. Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. *Cancer Lett.* **2020**, *471*, 61–71. [CrossRef]
21. Aubreville, M.; Knipfer, C.; Oetter, N.; Jaremenko, C.; Rodner, E.; Denzler, J.; Bohr, C.; Neumann, H.; Stelzle, F.; Maier, A. Automatic classification of cancerous tissue in laserendomicroscopy images of the oral cavity using deep learning. *Sci. Rep.* **2017**, *7*, 11979. [CrossRef] [PubMed]

22. Uthoff, R.D.; Song, B.; Sunny, S.; Patrick, S.; Suresh, A.; Kolur, T.; Keerthi, G.; Spires, O.; Anbarani, A.; Wilder-Smith, P.; et al. Point-of-care, smartphone-based, dual-modality, dual-view, oral cancer screening device with neural network classification for low-resource communities. *PLoS ONE* **2018**, *13*, e0207493. [[CrossRef](#)] [[PubMed](#)]
23. Bur, A.M.; Holcomb, A.; Goodwin, S.; Woodroof, J.; Karadaghy, O.; Shnayder, Y.; Kakarala, K.; Brant, J.; Shew, M. Machine learning to predict occult nodal metastasis in early oral squamous cell carcinoma. *Oral Oncol.* **2019**, *92*, 20–25. [[CrossRef](#)]
24. Qian, Z.; Li, Y.; Wang, Y.; Li, L.; Li, R.; Wang, K.; Li, S.; Tang, K.; Zhang, C.; Fan, X.; et al. Differentiation of glioblastoma from solitary brain metastases using radiomic machine-learning classifiers. *Cancer Lett.* **2019**, *451*, 128–135. [[CrossRef](#)] [[PubMed](#)]
25. Mintz, Y.; Brodie, R. Introduction to Artificial Intelligence in medicine. *Minim. Invasive Ther. Allied Technol.* **2019**, *28*, 73–81. [[CrossRef](#)]
26. Tabik, S.; Gómez-Ríos, A.; Martín-Rodríguez, J.L.; Sevillano-García, I.; Rey-Area, M.; Charte, D.; Guirado, E.; Suárez, J.L.; Luengo, J.; Valero-González, M.A.; et al. COVIDGR Dataset and COVID-SDNet Methodology for Predicting COVID-19 Based on Chest X-Ray Images. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 3595–3605. [[CrossRef](#)]
27. Lakhani, P.; Sundaram, B. Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology* **2017**, *284*, 574–582. [[CrossRef](#)]
28. Almansour, N.A.; Syed, H.F.; Khayat, N.R.; Altheeb, R.K.; Juri, R.E.; Alhiyafi, J.; Alrashed, S.; Olatunji, S.O. Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. *Comput. Biol. Med.* **2019**, *109*, 101–111. [[CrossRef](#)]
29. Kerney, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* **2018**, *172*, 1122–1131. [[CrossRef](#)]
30. Vásquez-Correa, J.C.; Arias-Vergara, T.; Orozco-Arroyave, J.R.; Eskofier, B.; Klucken, J.; Nöth, E. Multimodal assessment of Parkinson’s disease: A deep learning approach. *IEEE J. Biomed. Health Inform.* **2018**, *23*, 1618–1630. [[CrossRef](#)]
31. Tan, A.; Huang, H.; Zhang, P.; Li, S. Network-based cancer precision medicine: A new emerging paradigm. *Cancer Lett.* **2019**, *458*, 39–45. [[CrossRef](#)] [[PubMed](#)]
32. Kouznetsova, V.L.; Li, J.; Romm, E.; Tsigelny, I.F. Finding distinctions between oral cancer and periodontitis using saliva metabolites and machine learning. *Oral Dis.* **2021**, *27*, 484–493. [[CrossRef](#)] [[PubMed](#)]
33. Alabi, R.O.; Bello, I.O.; Youssef, O.; Elmusrati, M.; Mäkitie, A.A.; Almangush, A. Utilizing deep machine learning for prognostication of oral squamous cell carcinoma—A systematic review. *Front. Oral Health* **2021**, *2*, 686863. [[CrossRef](#)] [[PubMed](#)]
34. Halicek, M.; Shahedi, M.; Little, J.V.; Chen, A.Y.; Myers, L.L.; Sumer, B.D.; Fei, B. Head and neck cancer detection in digitized whole-slide histology using convolutional neural networks. *Sci. Rep.* **2019**, *9*, 14043. [[CrossRef](#)] [[PubMed](#)]
35. Halicek, M.; Little, J.V.; Wang, X.; Patel, M.; Griffith, C.C.; Chen, A.Y.; Fei, B. Tumor margin classification of head and neck cancer using hyperspectral imaging and convolutional neural networks. In Proceedings of the Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling, Houston, TX, USA, 12–15 February 2018; Volume 10576, pp. 17–27.
36. Ariji, Y.; Fukuda, M.; Kise, Y.; Nozawa, M.; Yanashita, Y.; Fujita, H.; Katsumata, A.; Ariji, E. Contrast-enhanced computed tomography image assessment of cervical lymph node metastasis in patients with oral cancer by using a deep learning system of Artificial Intelligence. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol.* **2019**, *127*, 458–463. [[CrossRef](#)] [[PubMed](#)]
37. Fujima, N.; Andreu-Arasa, V.C.; Meibom, S.K.; Mercier, G.A.; Salama, A.R.; Truong, M.T.; Sakai, O. Deep learning analysis using FDG-PET to predict treatment outcome in patients with oral cavity squamous cell carcinoma. *Eur. Radiol.* **2020**, *30*, 6322–6330. [[CrossRef](#)]
38. Folmsbee, J.; Liu, X.; Brandwein-Weber, M.; Doyle, S. Active deep learning: Improved training efficiency of convolutional neural networks for tissue classification in oral cavity cancer. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 770–773. [[CrossRef](#)]
39. Das, N.; Hussain, E.; Mahanta, L.B. Automated classification of cells into multiple classes in epithelial tissue of oral squamous cell carcinoma using transfer learning and convolutional neural network. *Neural Netw.* **2020**, *128*, 47–60. [[CrossRef](#)]
40. Schwab, M. Broder histological classification. In *Encyclopedia of Cancer*; Springer: Berlin/Heidelberg, Germany, 2011; p. 571.
41. Sridharan, G.; Ramani, P.; Patankar, S.; Vijayaraghavan, R. Evaluation of salivary metabolomics in oral leukoplakia and oral squamous cell carcinoma. *J. Oral Pathol. Med.* **2019**, *48*, 299–306. [[CrossRef](#)]
42. Ishikawa, S.; Sugimoto, M.; Edamatsu, K.; Sugano, A.; Kitabatake, K.; Iino, M. Discrimination of oral squamous cell carcinoma from oral lichen planus by salivary metabolomics. *Oral Dis.* **2020**, *26*, 35–42. [[CrossRef](#)]
43. da Costa, N.L.; de Sá Alves, M.; de Sá Rodrigues, N.; Bandeira, C.M.; Alves, M.G.O.; Mendes, M.A.; Alves, L.A.C.; Almeida, J.D.; Barbosa, R. Finding the combination of multiple biomarkers to diagnose oral squamous cell carcinoma—A data mining approach. *Comput. Biol. Med.* **2022**, *143*, 105296. [[CrossRef](#)]
44. He, H.; Ma, Y. *Imbalanced Learning: Foundations, Algorithms, and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2013; p. 47.
45. Van der Maaten, L.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
46. Danielsson, P.E. Euclidean distance mapping. *Comput. Graph. Image Process.* **1980**, *14*, 227–248. [[CrossRef](#)]
47. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [[CrossRef](#)]
48. Guyon, I.; Weston, J.; Barnhill, S.; Vapnik, V. Gene selection for cancer classification using support vector machines. *Mach. Learn.* **2002**, *46*, 389–422. [[CrossRef](#)]

49. Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Mohamed, N.A.; Arshad, H. State-of-the-art in artificial neural network applications: A survey. *Heliyon* **2018**, *4*, e00938. [[CrossRef](#)]
50. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [[CrossRef](#)]
51. Zell, A. *Simulation Neuronaler Netze*; Addison-Wesley Bonn: Boston, MA, USA, 1994; Volume 1.
52. Jha, D.; Kim, J.I.; Kwon, G.R. Diagnosis of Alzheimer’s disease using dual-tree complex wavelet transform, PCA, and feed-forward neural network. *J. Healthc. Eng.* **2017**, *2017*, 9060124. [[CrossRef](#)]
53. Arulmurugan, R.; Anandakumar, H. Early detection of lung cancer using wavelet feature descriptor and feed forward back propagation neural networks classifier. In *Computational Vision and Bio Inspired Computing*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 103–110.
54. Aljaaf, A.J.; Mohsin, T.M.; Al-Jumeily, D.; Alloghani, M. A fusion of data science and feed-forward neural network-based modelling of COVID-19 outbreak forecasting in IRAQ. *J. Biomed. Inform.* **2021**, *118*, 103766. [[CrossRef](#)]
55. Fonseca, A.U.; Felix, J.P.; Vieira, G.S.; Fernandes, D.S.A.; Soares, F. Artificial Neural Networks and BPPC Features for Detecting COVID-19 and Severity Level. In Proceedings of the 2022 IEEE International Conference on Systems, Man and Cybernetics (SMC), Prague, Czech Republic, 9–12 October 2022.
56. Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.* **1944**, *2*, 164–168. [[CrossRef](#)]
57. Marquardt, D.W. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431–441. [[CrossRef](#)]
58. Dorofki, M.; Elshafie, A.H.; Jaafar, O.; Karim, O.A.; Mastura, S. Comparison of artificial neural network transfer functions abilities to simulate extreme runoff data. *Int. Proc. Chem. Biol. Environ. Eng.* **2012**, *33*, 39–44.
59. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In Proceedings of the IJCAI, Montreal, QC, Canada, 20–25 August 1995; Volume 14.2 , pp. 1137–1145.
60. Blagus, R.; Lusa, L. SMOTE for high-dimensional class-imbalanced data. *BMC Bioinform.* **2013**, *14*, 106. [[CrossRef](#)] [[PubMed](#)]
61. Fonseca, A.U.; Felix, J.P.; Vieira, G.S.; Mourão, Y.C.; Monteiro, J.C.G.; Soares, F. Uma Rede Neural Artificial para Suporte ao Diagnóstico de Carcinoma Espinocelular Oral. In Proceedings of the IX Congresso Latino-Americano de Engenharia Biomédica (CLAIB 2022) and XXVIII Congresso Brasileiro de Engenharia de Engenharia Biomédica (CBEB 2022), Virtual, 24–28 October 2022.
62. Wang, Q.; Gao, P.; Cheng, F.; Wang, X.; Duan, Y. Measurement of salivary metabolite biomarkers for early monitoring of oral cancer with ultra performance liquid chromatography–mass spectrometry. *Talanta* **2014**, *119*, 299–305. [[CrossRef](#)] [[PubMed](#)]
63. Gleber-Netto, F.O.; Yakob, M.; Li, F.; Feng, Z.; Dai, J.; Kao, H.K.; Chang, Y.L.; Chang, K.P.; Wong, D.T. Salivary Biomarkers for Detection of Oral Squamous Cell Carcinoma in a Taiwanese PopulationSalivary Biomarkers for Oral Cancer Detection. *Clin. Cancer Res.* **2016**, *22*, 3340–3347. [[CrossRef](#)]
64. Deepthi, G.; Nandan, S.; Kulkarni, P.G. Salivary tumour necrosis factor- α as a biomarker in oral leukoplakia and oral squamous cell carcinoma. *Asian Pac. J. Cancer Prev. APJCP* **2019**, *20*, 2087.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.