

Article

Model Checking-Based Radiomics for Diagnosis and Prognosis of Oral Cavity Tumors: A Two-Tier Approach

Maria Paola Belfiore ^{1,*} , Maria Rita Cristiano ¹, Valeria Sorgente ², Giulia Varriano ^{2,*}, Vittoria Nardone ², Maria Chiara Brunese ², Salvatore Cappabianca ¹, Antonella Santone ² and Luca Brunese ²

¹ Department of Precision Medicine, University of Campania “L. Vanvitelli”, 80138 Naples, Italy; mariaritacristiano@unicampania.it (M.R.C.); salvatorecappabianca@unicampania.it (S.C.)

² Department of Medicine and Health Sciences, University of Molise, 86100 Campobasso, Italy; valeria.sorgente@unimol.it (V.S.); vittoria.nardone@unimol.it (V.N.); mariachiarabrunese@gmail.com (M.C.B.); antonella.santone@unimol.it (A.S.); luca.brunese@unimol.it (L.B.)

* Correspondence: mariapaolabelfiore@unicampania.it (M.P.B.); giulia.varriano@unimol.it (G.V.)

Abstract

Background: Oral cavity tumors are often diagnosed at advanced stages due to non-specific early symptoms and limitations in imaging sensitivity. This study proposes a novel methodology combining Radiomics and Model Checking to improve diagnosis and prognosis. **Methods:** A retrospective dataset of 18 patients (12 with oral squamous cell carcinoma and 6 healthy controls) who underwent contrast-enhanced MRI was analyzed. Radiomic features were extracted and selected, then encoded into formal models. A two-tier Model Checking approach was applied: (i) classification of healthy vs. pathological patients and (ii) prediction of treatment response. **Results:** The proposed method achieved a diagnostic accuracy of 93% and a prognostic accuracy of 75%. The approach demonstrated robustness even with a limited dataset, outperforming traditional data-driven methods in small-sample settings. **Conclusions:** The integration of Radiomics with Model Checking provides an explainable and effective tool for early detection and prognosis of oral cavity tumors. This approach may support clinicians as a decision-making aid, particularly in data-scarce scenarios.

Keywords: Radiomics; Model Checking; oral cancer; MRI; prognosis; Formal Methods



Academic Editors: Jaroslaw Krzywanski, Marcin Sosnowski, Karolina Grabowska, Dorian Skrobek, Anna Zylka, Agnieszka Kijo-Kleczkowska, Bashar Shboul and Tomasz Czakiert

Received: 15 May 2026

Revised: 6 July 2026

Accepted: 8 July 2026

Published: 10 July 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Cancer of the oral cavity is a widespread and significant health issue, prevalent not only in developing nations but also in developed countries [1]. The most frequent manifestation of this cancer is oral squamous cell carcinoma (OSCC), predominantly driven by tobacco and alcohol use [2]. The pathogenesis of OSCC involves a typical pattern of initiation and progression, characterized by the gradual accumulation of genetic mutations that ultimately lead to overt neoplastic transformation [3]. This accumulation of mutations is often facilitated by age, with the adult elderly population being particularly susceptible due to the longer period available for genetic alterations to occur [4,5]. OSCCs commonly originate in the oral tongue, which is not only the most prevalent sub-site but also carries the worst prognosis, followed by the floor of the mouth, buccal mucosa, alveolar mucosa, and hard palate [6,7].

Accurate prognosis prediction for oral cancer is crucial for optimizing treatment strategies [8]. By evaluating the tumor’s aggressiveness and recurrence potential, healthcare

providers can tailor interventions to enhance patient outcomes. Moreover, precise prognostication facilitates vigilant disease monitoring and timely interventions, contributing to more effective oral cancer management and potentially higher survival rates. Overall, predicting oral cancer prognosis offers substantial benefits, including personalized treatment plans, improved patient care, and enhanced clinical outcomes [8].

Medical imaging plays a crucial role in providing details of the tumors that complement clinical examinations and biopsies, identifying lymph node involvement, and detecting distant metastases or synchronous tumors, thereby aiding in effective staging and treatment planning. Commonly used imaging techniques determining the loco-regional extension of the cancer are contrast-enhanced Computed Tomography (CT), contrast-enhanced Magnetic Resonance Imaging (MRI), and neck Ultrasound (US) [9]. In selected cases, Positron Emission Tomography (PET) combined with CT or MRI and performed with dedicated coils and high-field machines (1.5 Tesla) is considered the first-choice examination in the study of tumors of the oropharynx, oral cavity, nasopharynx, hypopharynx, and laryngeal tumors [9]. The current standard for treating oral cancer involves surgical tumor removal, often followed by radiation or chemoradiation therapy. Recent advancements in these techniques have improved survival rates. Effective treatment relies on a multidisciplinary approach to optimize disease control while preserving oral function and aesthetics [2]. Diagnosing oral squamous cell carcinomas (OSCCs) often faces delays due to their generic symptoms, which can be mistaken for benign conditions. Additionally, predicting treatment response is challenging because the tumor's location and extent vary widely. To address these issues, we propose a Model Checking-based approach [10] to facilitate early and automatic diagnosis of oral cavity tumors and predict treatment response or prognosis. The absence of a training phase represents an important methodological difference between the proposed framework and conventional Machine-Learning (ML) or Deep-Learning (DL) approaches. In particular, the Model Checking-based strategy does not estimate parameters from the available cohort and does not depend on train/test splitting or iterative optimization. Therefore, it avoids some limitations typically associated with training data-driven models on small datasets, such as unstable parameter estimation and classical overfitting.

Our approach is evaluated through two tasks: first, we assess its diagnostic performance by measuring its effectiveness in distinguishing between patients with oral cavity tumors (classified as 'sick') and healthy individuals. Second, we validate its accuracy in predicting treatment response using data from patients identified as sick in the initial task. This innovative approach aims to enhance early diagnosis and improve prognosis prediction for OSCCs. The results obtained from the methodology are promising, although more validation will be carried out in future work. Nevertheless, thanks to the Formal Methods [10], we are able to obtain models that can be readable even by medical specialists, who can get more information on the correlation between radiomic features and clinical parameters.

2. Materials and Methods

In this section, we will discuss the proposed methodology. Broadly speaking, after the dataset is collected, the images are manually segmented, and radiomic features are extracted from them. First, we performed a feature selection step through an automatic but human-supervised procedure, looking at the distribution curves of the two relevant classes: healthy vs. pathological patients in the diagnostic tier, and treatment-positive vs. treatment-negative patients in the prognostic tier. Second, we added the exact discretization criterion of each feature into three bins for the patient-specific formal model. Third, each MRI slice was encoded as a state of a finite-state transition system, and transitions represented the

ordered progression from one slice to the next. Diagnostic and prognostic patterns were expressed as modal μ -calculus formulae, and the model checker verified whether each patient-specific model satisfied the corresponding property. Satisfaction of the diagnostic property was interpreted as the presence of radiomic patterns associated with OSCC, whereas satisfaction of the prognostic property was used to identify patterns associated with treatment response.

2.1. Dataset

For this retrospective study, we selected 18 patients, including 12 with histologically confirmed diagnoses of OSCC, and six healthy subjects were used as controls. All participants underwent head and neck MRI at the Radiology Department of the University of Campania “Luigi Vanvitelli” between 2020 and 2023. The study group comprised 11 males and 5 females, all of white ethnicity, with an average age of 63 years. The inclusion criteria were:

- The presence of risk factors such as cigarette smoking and alcohol consumption;
- The location of the tumor in the palate, hemipalate, hard palate, soft palate, tongue, or gingival arch;
- Patients who, after receiving a positive pathology result, continued the treatment at the same healthcare department where the initial MRI screening was performed.

We chose to acquire only the axial plane selected from the MRI dataset, and the T1 series was acquired after the administration of the paramagnetic contrast agent.

Of the patients who were diagnosed as sick, four underwent surgery. The remaining part (8 patients) received radiotherapy treatment: four had a complete response to treatment, one a partial response, and three, unfortunately, progressed with the disease. Therefore, it was necessary to clinically monitor individual patients throughout the period between diagnosis and complete recovery to diagnose any recurrence of the disease or the onset of another pathology or harmful effects resulting from the treatment.

2.2. Model Checking

Model Checking [10] belongs to formal verification techniques, guaranteeing that a system or model meets established properties. It involves examining whether a model aligns with specified behavioral requirements or constraints. Its typical workflow includes the following operations:

- **Model Building:** The system is represented as a formal model (e.g., finite state machines or transition systems);
- **Properties Specification:** A formal specification defines desired properties or requirements, e.g., properties expressed with temporal logic formulas;
- **Model Verification:** These algorithms explore all possible states of the model and verify whether specified properties hold;
- **Counterexample Generation:** If a property is violated, the tool produces a counterexample, aiding debugging.

In particular, in this work the verification was performed using a μ -calculus-based Model Checking framework, with temporal properties and the specification of recursive temporal and modal properties over finite-state transition systems. Transitions encoded the ordered progression from one slice to the next within the radiological examination. Thus, each complete radiological study was represented as a patient-specific formal model describing the spatial evolution of radiomic patterns across slices. The diagnostic or prognostic rule was encoded as a μ -calculus formula. The model checker verified whether the formal model associated with each patient satisfied the corresponding formula. In the first tier, satisfaction of the diagnostic formula was used to classify the patient as

pathological, whereas non-satisfaction indicated absence of the target tumor-associated radiomic pattern. In the second tier, satisfaction of the prognostic formula was used to identify radiomic patterns associated with treatment response.

2.3. Radiomics and Methodology

Radiomics [11] is a quantitative analysis of images that allows us to exploit what we do not see through the naked eye. The radiomic workflow is composed of different steps, such as Image Acquisition, definition of the Region of Interest, and Feature Extraction.

Specifically, starting from diagnostic images, the region of interest is defined, and the radiomic features are extracted [11]. Some of these concern various statistical values such as mean, median, standard deviation, maximum, and minimum, obtained from the histogram of grey levels contained in the slices. These features are quantitative patterns that can help in identifying disease marks. They are used to derive prognostic and predictive factors of therapy response and to discriminate against malignant tissues from benign ones.

As Figure 1 shows, the proposed methodology combines Radiomics and Model Checking in a two-tier approach:

1. In the first tier, the goal is to classify healthy versus sick patients;
2. In the second tier, the goal is to use the patients recognized as sick in the first step to predict the treatment response.

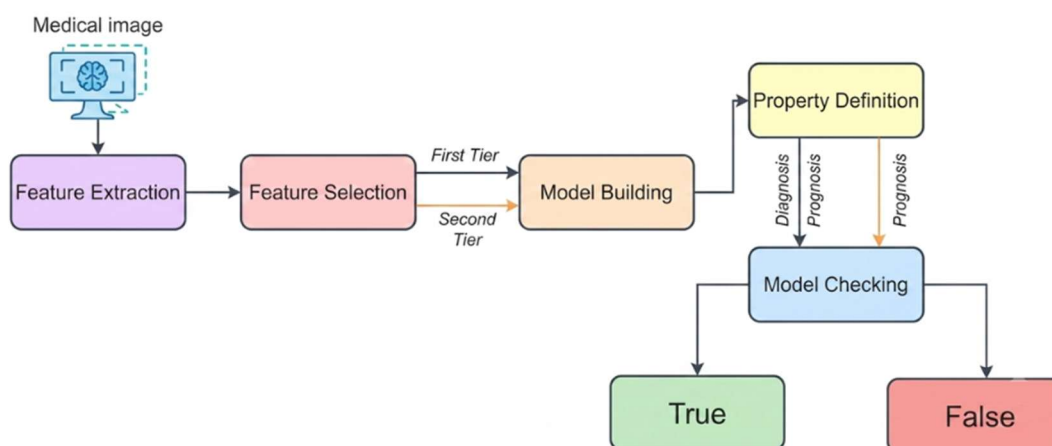


Figure 1. The proposed methodology is divided in first tier for the classification of healthy vs. patients affected by OSCCs, and in a second tier, for the classification of patients who had a positive response to treatment.

According to the IBSI standard [12], from a medical image it is possible to extract many different features, i.e., 107 features [12]. However, only a small subset can represent the sample, while some of them might be unnecessary, introducing additional complexity and noise to the analysis. For the above reasons, we perform the Feature Selection process to collect the most representative features and to reduce useless noise and complexity in image processing by analyzing the feature distribution curves for each category of patient. For each radiomic feature, the value distributions were automatically computed and compared between the two classes considered in each tier. Features showing clearer separation between the distributions were selected as candidates for the formal models. The final selection was supervised by domain experts of radiology and computer science, who retained the most discriminative and interpretable features for the μ -calculus formulae. No fixed universal threshold was used, because the radiomic features have heterogeneous scales and distributions. Given the small cohort, the selected features were not claimed

as statistically validated biomarkers; rather, they were considered discriminative features within this exploratory dataset.

At the end of the Feature Selection process, only four features are selected for each radiomic class to eliminate excessive and redundant data that do not contribute to improving the accuracy of predictive models but tend to decrease it.

Feature Selection is specialized for each tier, producing different discriminant features for the diagnostic phase compared to the predictive one.

Once the most significant features are obtained, to build formal models, it is necessary to feed these values into a discretization algorithm that divides them into intervals. In this specific case, a discretization on three equidistributed levels (low, basal, and up) was performed, all with the same size and computed as the difference between the maximum and minimum values of the features, divided by the number of intervals. This discretization process is done for each slice of the exam and for each chosen feature, as in the example below for the Median, Energy, and Variance features.

$$\begin{aligned} \text{slice 1} &= \text{low_median; low_energy; high variance} \\ \text{slice 2} &= \text{basal_median; low_energy; low_variance} \\ \text{slice 3} &= \dots \end{aligned}$$

After the discretization process, the chosen features are combined to build the formal model, which mirrors the diagnostic exam. More in detail, the formal model is an abstraction of the medical image that can follow how radiomic features vary through the complete diagnostic exam. To sum up, the formal model encloses the information in the medical image. Thus, checking whether a specific behavior is exhibited by the formal model, e.g., a particular disease arises in the image, is feasible through the formal property specification. Formal properties express how the specific disease manifests itself in the medical image. Precisely, a formal property reflects the combination of features in a particular sequence as manifested in the diagnostic examination. Note that a formal property's strength lies in its inherited capacity to be expressed in temporal logic, allowing for accurately delineating and capturing temporal and spatial relationships and behaviors within a system [13]. The formal property enables the precise description of dynamic aspects, temporal ordering, and the evolution of features or events over time. It provides a structured and rigorous method to analyze and verify the behavior of a system leveraging temporal aspects. For our purposes, different properties are defined, one for each tier of the methodology. We will also test whether the property created for the diagnosis can be helpful in the treatment response. The Model Checker tool achieves the clinical results by automatically verifying if the knowledge included in the model exposes the behavior expressed by logical-temporal properties [10]. Roughly speaking, its task is to check whether the formal models, created from the transformation of the most significant features into a formal language, satisfy certain properties related to a specific health state. Results of the first tier are shown in Table 1.

Table 1. Diagnostic results reported for each considered feature class.

Feature Class	Accuracy	Precision	Recall	Specificity
FIRST	93%	89%	100%	83%
GLCM	86%	100%	75%	100%
GLDM	93%	100%	88%	100%
GLRLM	86%	80%	100%	67%
GLSZM	93%	89%	100%	83%

In the second tier of the methodology, the analysis is carried out only on the diseased patients, so the methodology ends for all patients classified as healthy in the first step by the Model Checker. Conversely, if the patient shows typical patterns of oral cavity tumors, the formal verification continues in order to predict the course of the disease and the response to a specific therapeutic treatment. At this point, the 12 diseased subjects were divided into patients who could eradicate the tumor following appropriate treatment, and patients in whom the pathology reoccurred. Some subjects underwent only surgical therapy through resection, while others also underwent systemic treatment.

2.3.1. Patient's Treatment Response Identification Using the First Tier Formula

In the second step, we tried to use the diagnostic property (results are shown in Table 2) to discover information about the patient's treatment response. The aim is to verify whether the same diagnosis formula contains information about the treatment, and hence, it could correctly classify the response to treatment as well. Furthermore, this analysis was designed to discover new treatment information from time 0 (t_0) onwards, i.e., when the patient has just been diagnosed. This could be advantageous in terms of reducing analysis time. We created formal models with the features of FIRST, GLCM, GLRLM, and GLSZM. The GLDM class was the better class, including Dependence Variance, Large Dependence Emphasis, Large Dependence Low Gray Level Emphasis, and Small Dependence Low Gray Level Emphasis.

Table 2. Results of the same diagnostic formula used to predict the response to treatment.

Feature Class	Accuracy	Precision	Recall	Specificity
FIRST	72%	62%	100%	50%
GLCM	67%	67%	67%	67%
GLDM	72%	64%	88%	60%
GLRLM	67%	57%	100%	40%
GLSZM	72%	62%	100%	50%

2.3.2. Patient's Treatment Response Identification Using a New Prognostic Formula

Unfortunately, the property used for diagnosis needs to be more specific for classifying treatment response. Therefore, we performed a new feature selection and, in collaboration with radiologists, formulated a new formula to distinguish patients who responded well to treatment from those who did not benefit. In particular, we exploited the FIRST, GLCM, GLDM, GLRLM, GLSZM, and NGTDM classes; the results are shown in Table 3. In general, the best performance was achieved by the GLSZM class, whose most important features were Gray Level Non Uniformity Normalized, Gray Level Variance, Large Area Emphasis, and Zone Variance.

Table 3. Results for formulae identifying the group of patients that exhibited positive responses to the therapy.

Feature Class	Accuracy	Precision	Recall	Specificity
FIRST	67%	67%	100%	20%
GLCM	75%	78%	88%	50%
GLDM	50%	60%	75%	20%
GLRLM	75%	78%	88%	25%
GLSZM	75%	86%	75%	75%

3. Results

Patient Diagnosis Classification. The first evaluation aims to assess the classification performance of the proposed approach to distinguish healthy patients from those with oral cavity neoplasms. Our approach aims to identify a set of discriminative and interpretable radiomic features whose value distributions differ across predefined patient categories. These features are then discretized and encoded into formal models specific to each patient, allowing the properties of μ -calculus to verify whether the radiomic model associated with a given clinical class is present. The categories are healthy patients, composed of 6 patients, and patients with OSCC neoplasm, consisting of 12 patients.

For the Feature Extraction step, the ROI coincides with the entire slice under examination. Thanks to the feature selection results, we decided not to include the feature classes of NGTDM and SHAPE in the analysis.

An example of the type of μ -calculus rule used in the Model Checking step. After discretization, each selected radiomic feature is represented as a symbolic state variable. For example, a simplified diagnostic property can be expressed as:

$$\text{prop1} = \mu X.((f_{1_up} \wedge f_{2_basal} \wedge f_{3_low}) \vee \langle \text{prop2} \rangle X)$$

where f_1 , f_2 , and f_3 represent selected radiomic features and $\langle \text{prop2} \rangle$ denotes the transition from the first MRI slice to the second and so on. This formula verifies whether, along the ordered sequence of slices, there exists at least one state showing the specified radiomic pattern (up, basal and low). If the patient-specific model satisfies the formula, the corresponding diagnostic pattern is considered present and the patient is classified as affected by oral cancer. The actual formulae used in the analysis were defined using the selected feature combinations for each tier and feature class.

As displayed in Table 1, we reported the rule with the highest Accuracy for each feature class. The best performance was achieved by the FIRST-order, GLDM, and GLSZM features with an Accuracy of 93%. Precision was also considered, as it should be as close to 100% as possible, ensuring no false positives. In this way, the classifier will only identify a patient as sick if it is sure that they are. Therefore, if the neoplasm is not yet diagnosable, it is preferable not to transmit a falsely alarming message, thinking that the disease is present when it is not. This way, healthy patients will only undergo necessary screening and therapies to detect and combat the pathology, which would be useless and avoidable.

In Table 4, there is the Confusion Matrix with True and False Negatives, True and False Positives of this first and second step.

Table 4. Diagnostic and Prognostic Confusion Matrix.

Tier	Diagnosis (GLCM)	Prognosis (GLSZM)
True Positives	9	6
True Negatives	6	3
False Positives	0	1
False Negatives	3	2
Accuracy %	83	75
Precision %	100	86
Recall %	75	75
Specificity %	100	75

This methodology aims to achieve the lowest possible number of false positives because, in such cases, the tool would erroneously indicate a positive treatment outcome, which could lead to a falsely reassuring message suggesting the absence of pathology when it is present. A false negative test result could prevent further diagnostic investigations and appropriate therapies to fight a specific pathology. Nowadays, it is also known that a delay in diagnosis or treatment can make a difference between life and death for the patient. Therefore, as shown in Table 2, the GLCM features are selected as the most relevant, with an Accuracy of 67%, compared to other classes that, although more accurate, showed a higher number of FPs.

Patients' Treatment Response identification using a New Prognostic Formula.

Results of Table 2 suggest we create a property specifically designed to accurately identify and classify the classification of treatment response of patients affected by oral cavity neoplasm in the database. Of these 12 patients, eight are labelled 'Treatment Positive', i.e., the treatment had a beneficial effect on the oral cavity neoplasm. The other four patients are labelled 'Treatment Negative', i.e., the treatment was ineffective.

As seen from the results in Table 3, the feature classes with the highest Accuracy are GLCM, GLRLM, and GLSZM, all reaching an Accuracy of 75%. However, these results are slightly lower than those obtained in the first diagnosis step, where the Accuracy reached 93%. However, the available dataset was more populated, so including additional subjects with oral cavity tumors could help achieve higher performance in this second tier.

Here, the property is modified in the pattern values:

$$\text{prop1} = \mu X.((f_{1_up} \wedge f_{2_basal} \wedge f_{3_up}) \vee \langle \text{prop2} \rangle X)$$

This formula verifies whether, along the ordered sequence of slices, there exists at least one state showing the specified radiomic pattern (up, basal and low). If the patient-specific model satisfies the formula, the corresponding prognostic pattern is considered present and the patient is classified as Treatment Positive.

In Table 4, we reported the true positives, true negatives, false positives, and false negatives, along with precision, recall/sensitivity, and specificity for the best-performing feature class at each level. For the diagnostic level, the positive class corresponds to patients with OSCC, while the negative class corresponds to healthy controls. For the prognostic level, the positive class corresponds to patients with a positive response to treatment, while the negative class corresponds to patients with a negative response to treatment.

These metrics are intended to provide a descriptive interpretation of the results rather than formal inferential evidence, given the limited sample size of the cohort. In particular, the diagnostic model based on GLCM features performed well, as it showed no false positives, achieving 100% precision and specificity, although three patients with the disease were incorrectly classified as healthy. At the prognostic level, the GLSZM-based model achieved more balanced performance, with accuracy, recall, and specificity all at 75% and one false positive.

4. Discussion

Oral cavity cancer is a malignant tumor that develops in the mouth, including the lips, tongue, gums, palate, cheek mucosa, and floor of the mouth [14]. It can manifest as a lump or an ulcerated lesion that does not heal, causing pain and difficulty speaking, swallowing, or chewing. OSCC is the most common histology, and the main etiological factors include smoking, tobacco use, excessive alcohol consumption, radiation, and human papilloma virus [3]. Other malignant tumors can arise from the epithelium, connective tissue, minor salivary glands, lymphoid tissue, and melanocytes [2]. Diagnosing oral cavity tumors can be complex due to several factors [15]. First, mild initial symptoms, such as oral

cavity tumors, can present with mild or asymptomatic initial symptoms, such as small lesions, ulcers, or changes in the color or texture of the tissues [16]. This makes it difficult to recognize early signs and make an early diagnosis. Overlap of symptoms with other conditions is possible because the visible clinical signs of oral cavity cancer can overlap with other benign conditions, such as oral ulcers, traumatic lesions, or infections, making accurate diagnosis challenging [17].

When the diagnostic delay exceeds one month, the risk of having an advanced-stage oral cancer is significantly higher [18]. In most cases, the patient is responsible for a large part of the diagnostic delay; however, delay can also be the result of an incorrect medical approach by not suspecting an oral malignancy and not diagnosing and treating it promptly and adequately [18–20]. Furthermore, the oral cavity is a complex anatomical area that includes various structures such as lips, tongue, gums, palate, cheeks, and mouth floor. There are many different types of oral cavity tumors, including squamous cell carcinomas, melanoma, lymphomas, and sarcomas. Oral cancers belong to the type of OSCCs in more than 90% of cases [3]; other tumors of the oral cavity include those of the salivary minor glands, melanomas, and lymphomas [21]. Each type of tumor can present with distinctive symptoms and characteristics, requiring specific diagnostic approaches.

To confirm the diagnosis of an oral cavity tumor, complementary diagnostic tests such as biopsy, histopathological examinations, diagnostic imaging such as CT or MRI (as shown in Figure 2), and in some cases, specific laboratory tests such as DNA testing or immunological examinations, may be necessary [22]. These tests may require time and specialized skills to perform and interpret correctly. Due to the complexity and diverse clinical presentations of oral cavity tumors, it is crucial to involve a multidisciplinary team of specialists, such as oncologists, pathologists, radiologists, surgeons, and dentists specialized in oral surgery, to obtain an accurate diagnosis and appropriate treatment plan. Early diagnosis of oral cavity tumors is crucial to increase the chances of successful treatment and improve patient prognosis.

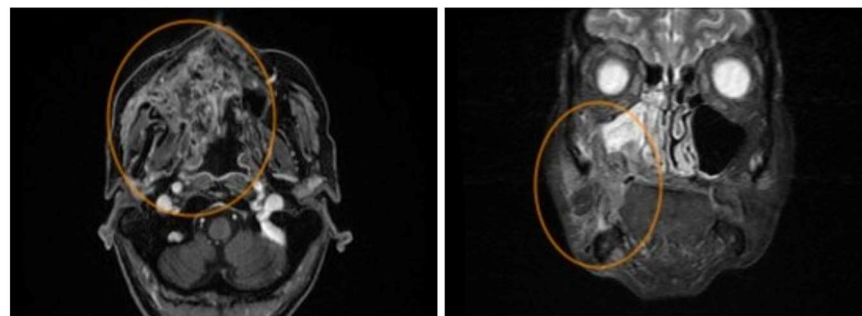


Figure 2. Images show example of the oral cancer inside yellow circles. The axial view (**left image**) shows the heteroplastic formation of the right hard palate which extends to the floor of the maxillary sinus and the nasopharynx; coronal view (**right image**) shows the absence of cleavage plane with the pterygoid muscles and the muscles of the upper lip.

In recent years, Artificial Intelligence (AI) approaches have gained attention as a means of improving image-based diagnosis [23]. Machine Learning (ML) and Deep Learning (DL) are two subsets of AI, which, although the terms are sometimes used interchangeably, have some crucial differences. Briefly, ML algorithms typically require an accurately categorized data input. In contrast, DL networks rely on layers of artificial neural networks to generate their categories by identifying edges (differences) within layers of neural networks when exposed to many data points. Therefore, although both subsets of AI are “intelligent,” DL requires much more data than a traditional ML algorithm. ML performs better with fewer data sets clearly labeled or structured about a gold standard or specific interest

criteria [24]. As a result, researchers have integrated image processing, pattern recognition, ML, and DL methods to develop CAD for oral cancer diagnosis. The literature suggests that automated quantification of an oral cancer diagnosis reduces grading conflicts between pathologists [25]. The major advantage of such technologies is the time reduction in the load of manual visualization of slices. They also assist pathologists in fast decision-making with better accuracy. Computerized image analysis of tissue slices can obtain information that may be missed with traditional viewing [26]. Precise and accurate histological findings are necessary for early diagnosis, classification, prediction, and specific treatment planning for OC [27]. Various reports have been published describing the application of AI in the early diagnosis, prognosis, and classification of OC [28–31].

Recent ML approaches have reported promising results for oral cancer classification. For example, Al-Batah et al. [32] reviewed and compared several data mining and ML methods, including Neural Networks, K-Nearest Neighbors, Support Vector Machines, and ensemble learning techniques, for oral cancer diagnosis and prognosis. In their analysis, Neural Networks achieved the best reported classification performance, with an accuracy of 93.6%. In comparison, our Model Checking-based radiomic framework achieved a diagnostic accuracy of 93%, which is comparable to the best ML-based performance reported by Al-Batah et al. [32]. However, the methodological assumptions of the two approaches are substantially different. Their approach relies on ML models, feature selection, dimensionality reduction, and data-driven classification, whereas our method does not require a training phase or parameter optimization. Instead, radiomic features extracted from MRI are discretized and encoded into patient-specific formal models, and the classification is obtained by verifying μ -calculus formulae. Therefore, our contribution lies not only in achieving comparable diagnostic performance, but also in providing an explainable, training-free framework suitable for small-sample radiomic studies.

Even Deep Learning has recently been proposed as a promising tool for improving the diagnosis of oral cancer. Varalakshmi et al. [33] discussed its potential to transform the care of oral cancer by supporting early diagnosis and clinical decision-making, reporting accuracy values ranging from 85% to 100%, although the available studies are highly heterogeneous in terms of imaging modalities and framework structure. Our approach differs substantially from these DL- and ML-based methods. Furthermore, while DL systems often require additional explainability techniques, the proposed framework of Formal Methods provides explicit logical rules that can be examined and discussed with clinical experts. In this context, our diagnostic accuracy of 93% falls within the performance range reported by DL studies in oral cancer, while our prognostic accuracy of 75% should be considered preliminary.

The Formal Methods-based methodology offers several advantages, including the absence of an initial training phase, which is often time-consuming and resource-intensive in conventional models. Our approach can also be applied to small datasets due to the lack of a training phase, making it particularly useful for analyzing rare diseases where the dataset consists of a small number of patients. The software does not undergo a learning process from a dataset, as is typical in many AI techniques. Instead, all the domain-specific knowledge and expertise of the radiologist is explicitly encoded into the properties and logical rules used by the formal verification environments to classify patients. Thus, our approach eliminates the need for a separate training phase, where models learn patterns from a labelled dataset. By explicitly defining the representative characteristics and properties that constitute the condition of interest, the expertise of radiologists becomes an integral part of the reasoning process of the software. This tight integration of human knowledge into the core logic of the software ensures that the analysis is closely aligned with established clinical practices and standards and makes the tool more

reliable. However, our approach is only partially automatic, as it requires the involvement of a domain expert to define the representative pattern of the OSCC to be found in the formal model. It is important to note that AI tools cannot replace the clinical expertise of healthcare professionals. Instead, AI techniques should be seen as complementary tools that aid in early diagnosis and treatment decision-making [25].

Limitations

The reported performances should be interpreted as preliminary results. The cohort included only 18 subjects in the diagnostic tier and 12 OSCC patients in the prognostic tier, all recruited from a single Italian institution. Therefore, the selected radiomic features and formal properties may be influenced by local acquisition protocols, scanner-related factors, segmentation procedures, and population characteristics. Consequently, the present results should be considered preliminary feasibility findings. Larger, multicenter, and externally validated cohorts are required to assess the reproducibility and transferability of the proposed approach across different clinical settings.

However, we would emphasize that the proposed methodology differs from conventional ML approaches by not involving a training phase. The formal models and temporal-logic properties are not learned from the dataset via parameter optimization; rather, they are explicitly defined via radiomic feature selection, discretization, and domain-driven formal specification. Therefore, the small sample size does not affect the method with respect to training instability or overfitting.

Regarding the involvement of experts in defining the clinically meaningful properties of μ -calculus, this represents a strength in terms of interpretability, since the final rules can be reviewed and discussed with clinicians. Second, the construction of patient-specific formal models and logical properties introduces modeling complexity, especially for those unfamiliar with Formal Methods. Conversely, however, they have a mathematical structure, which makes them extremely straightforward for the computer science expert to manage.

5. Conclusions

Radiomics, which involves extracting and analyzing quantitative data from medical images, can provide valuable insights into the characteristics of oral neoplastic lesions. By combining this approach with Formal Methods, which involve rigorous mathematical analysis of systems, researchers can enhance the accuracy of diagnostic processes and enable early detection of abnormalities.

One of the key advantages of this combined approach is its ability to produce reliable results even when working with small datasets. Unlike ML-based systems, which often require large amounts of training data to achieve high levels of accuracy, the MRI Radiomics and Formal Methods approach can yield meaningful insights from limited sample sizes. By leveraging the strengths of both MRI Radiomics and Formal Methods, healthcare professionals can improve the efficiency and effectiveness of diagnostic and prognostic processes for oral neoplastic lesions with a respective accuracy of 93% and 75%. This integrated approach can help identify potential abnormalities at an earlier stage, leading to better patient outcomes and potentially saving lives.

Therefore, the reported prognostic accuracy should be considered preliminary and not yet sufficient to support definitive clinical conclusions. Further studies involving larger, multicenter, and independent cohorts are necessary to validate the reproducibility, generalizability, and clinical utility of the proposed prognostic formal properties. Nevertheless, the present results suggest that combining MRI Radiomics with Formal Methods may provide an explainable and training-free methodological framework that is particularly suitable for small-sample settings and worthy of further investigation.

Author Contributions: Conceptualization, M.P.B., A.S. and L.B.; Methodology, M.P.B., V.N., G.V., V.S. and A.S.; Software, V.S., G.V. and V.N.; Validation, S.C. and M.P.B. Formal analysis, G.V., V.N. and V.S.; Data preparation, M.C.B. and M.R.C.; Writing—original draft preparation, G.V., M.P.B. and V.N.; Writing—editing draft, G.V. and M.P.B., Supervision, A.S., L.B. and M.P.B. All authors have read and agreed to the published version of the manuscript.

Funding: Thanks to National Plan for NRRP Complementary Investments D3 4 Health: Digital Driven Diagnostics, prognostics and therapeutics for sustainable Health care, by Funder “Ministero dell’Università e della Ricerca”. Thanks to Project e-DAI (Digital ecosystem for integrated analysis of heterogeneous health data related to high-impact diseases: innovative model of care and research), by Funder “Ministero dell’Università e della Ricerca”.

Informed Consent Statement: For each patient we collected informed consent and data following the guidelines of the Ethical Committee established in project no. 24997/2020.

Data Availability Statement: The dataset used and analyzed during this study are available from the corresponding author upon reasonable request.

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

References

1. Jemal, A.; Bray, F.; Center, M.M.; Ferlay, J.; Ward, E.; Forman, D. Global cancer statistics. *CA Cancer J. Clin.* **2011**, *61*, 69–90. [[CrossRef](#)] [[PubMed](#)]
2. Montero, P.H.; Patel, S.G. Cancer of the oral cavity. *Surg. Oncol. Clin.* **2015**, *24*, 491–508. [[CrossRef](#)]
3. Warnakulasuriya, S. Global epidemiology of oral and oropharyngeal cancer. *Oral Oncol.* **2009**, *45*, 309–316. [[CrossRef](#)] [[PubMed](#)]
4. Van Monsjou, H.S.; Wreesmann, V.B.; Van Den Brekel, M.W.; Balm, A.J. Head and neck squamous cell carcinoma in young patients. *Oral Oncol.* **2013**, *49*, 1097–1102. [[CrossRef](#)] [[PubMed](#)]
5. Pai, S.I.; Westra, W.H. Molecular pathology of head and neck cancer: Implications for diagnosis, prognosis, and treatment. *Annu. Rev. Pathol. Mech. Dis.* **2009**, *4*, 49–70. [[CrossRef](#)]
6. Stadler, M.E.; Patel, M.R.; Couch, M.E.; Hayes, D.N. Molecular biology of head and neck cancer: Risks and pathways. *Hematol./Oncol. Clin. N. Am.* **2008**, *22*, 1099–1124. [[CrossRef](#)]
7. Wu, K.; Wei, J.; Liu, Z.; Yu, B.; Yang, X.; Zhang, C.; Abdelrehem, A.; Zhang, C.; Li, S. Can pattern and depth of invasion predict lymph node relapse and prognosis in tongue squamous cell carcinoma. *BMC Cancer* **2019**, *19*, 714. [[CrossRef](#)] [[PubMed](#)]
8. Tralongo, V.; Rodolico, V.; Luciani, A.; Marra, G.; Daniele, E. Prognostic factors in oral squamous cell carcinoma. A review of the literature. *Anticancer Res.* **1999**, *19*, 3503–3510. [[PubMed](#)]
9. Chen, S.-W.; Zhang, Q.; Guo, Z.-M.; Chen, W.-K.; Liu, W.-W.; Chen, Y.-F.; Li, Q.-L.; Liu, X.-K.; Li, H.; Ou-Yang, D.; et al. Trends in clinical features and survival of oral cavity cancer: Fifty years of experience with 3,362 consecutive cases from a single institution. *Cancer Manag. Res.* **2018**, *10*, 4523–4535. [[CrossRef](#)] [[PubMed](#)]
10. Clarke, E.M. Model checking. In *Foundations of Software Technology and Theoretical Computer Science: 17th Conference, Kharagpur, India, 18–20 December 1997, Proceedings 17*; Springer: Berlin/Heidelberg, Germany, 1997; pp. 54–56.
11. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images are more than pictures, they are data. *Radiology* **2016**, *278*, 563–577. [[CrossRef](#)] [[PubMed](#)]
12. Van Griethuysen, J.J.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Beets-Tan, R.G.; Fillion-Robin, J.-C.; Pieper, S.; Aerts, H.J. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* **2017**, *77*, e104–e107. [[CrossRef](#)] [[PubMed](#)]
13. Stirling, C. An introduction to modal and temporal logics for ccs. In *Concurrency: Theory, Language, and Architecture: UK/Japan Workshop Oxford, UK, September 25–27, 1989 Proceedings*; Springer: Berlin/Heidelberg, Germany, 1991; pp. 1–20.
14. Sarode, G.S.; Sarode, S.C.; Maniyar, N.; Anand, R.; Patil, S. Oral cancer databases: A comprehensive review. *J. Oral Pathol. Med.* **2018**, *47*, 547–556. [[PubMed](#)]
15. Abati, S.; Bramati, C.; Bondi, S.; Lissoni, A. Trimarchi, Oral cancer and precancer: A narrative review on the relevance of early diagnosis. *Int. J. Environ. Res. Public Health* **2020**, *17*, 9160. [[CrossRef](#)] [[PubMed](#)]
16. Brands, M.; Brennan, P.; Verbeek, A.; Merks, M.; Geurts, S. Follow-up after curative treatment for oral squamous cell carcinoma. a critical appraisal of the guidelines and a review of the literature. *Eur. J. Surg. Oncol.* **2018**, *44*, 559–565. [[CrossRef](#)] [[PubMed](#)]
17. McCullough, M.; Prasad, G.; Farah, C. Oral mucosal malignancy and potentially malignant lesions: An update on the epidemiology, risk factors, diagnosis and management. *Aust. Dent. J.* **2010**, *55*, 61–65. [[CrossRef](#)] [[PubMed](#)]

18. Go, I.; Seoane, J.; Varela-Centelles, P.; Diz, P.; Takkouche, B. Isdiagnosticdelay related to advanced-stage oral cancer? A meta-analysis. *Eur. J. Oral Sci.* **2009**, *117*, 541–546.
19. McGurk, M.; Chan, C.; Jones, J.; O'regan, E.; Sherriff, M. Delay in diagnosis and its effect on outcome in head and neck cancer. *Br. J. Oral Max-Illofacial Surg.* **2005**, *43*, 281–284. [[CrossRef](#)]
20. Groome, P.A.; Rohland, S.L.; Hall, S.F.; Irish, J.; Mackillop, W.J.; O'Sullivan, B. A population-based study of factors associated with early versus late stage oral cavity cancer diagnoses. *Oral Oncol.* **2011**, *47*, 642–647. [[CrossRef](#)] [[PubMed](#)]
21. Dhanuthai, K.; Rojanawatsirivej, S.; Thosaporn, W.; Kintarak, S.; Subarnbhesaj, A.; Darling, M.; Kryshalskyj, E.; Chiang, C.-P.; Shin, H.-I.; Choi, S.-Y.; et al. Oral cancer: A multicenter study. *Med. Oral Patol. Oral Y Cir. Bucal* **2018**, *23*, e23.
22. Pałasz, P.; Adamski, Ł.; Go, M.; Starzyn, A.; Studniarek, M. Contemporary diagnostic imaging of oral squamous cell carcinoma—A review of literature. *Pol. J. Radiol.* **2017**, *82*, 193. [[PubMed](#)]
23. Munir, K.; Elahi, H.; Ayub, A.; Frezza, F.; Rizzi, A. Cancer diagnosis using deep learning: A bibliographic review. *Cancers* **2019**, *11*, 1235. [[CrossRef](#)] [[PubMed](#)]
24. İlhan, B.; Lin, K.; Guneri, P.; Wilder-Smith, P. Improving oral cancer outcomes with imaging and artificial intelligence. *J. Dent. Res.* **2020**, *99*, 241–248. [[CrossRef](#)] [[PubMed](#)]
25. Khanagar, S.B.; Alkadi, L.; Alghilan, M.A.; Kalagi, S.; Awawdeh, M.; Bijai, L.K.; Vishwanathaiah, S.; Aldhebaib, A.; Singh, O.G. Application and performance of artificial intelligence (ai) in oral cancer diagnosis and prediction using histopathological images: A systematic review. *Biomedicines* **2023**, *11*, 1612. [[CrossRef](#)] [[PubMed](#)]
26. Eid, R.A.; Landini, G. Quantification of the global and local complexity of the epithelial-connective tissue interface of normal, dysplastic, and neoplastic oral mucosae using digital imaging. *Pathol. Res. Pract.* **2003**, *199*, 475. [[CrossRef](#)] [[PubMed](#)]
27. Kayser, K.; GÄsrtler, J.Ä.; Bogovac, M.; Bogovac, A.; Goldmann, T.; Vollmer, E.; Kayser, G. Ai (artificial intelligence) in histopathology—from image analysis to automated diagnosis. *Folia Histochem. Cytobiol.* **2009**, *47*, 355–361. [[PubMed](#)]
28. Lo, X.A.; Matamala, F.; Venegas, B.; Rivera, C. Machine-learning applications in oral cancer: A systematic review. *Appl. Sci.* **2022**, *12*, 5715. [[CrossRef](#)]
29. Jubair, F.; Al-karadsheh, O.; Malamos, D.; Al Mahdi, S.; Saad, Y.; Hassona, Y. A novel lightweight deep convolutional neural network for early detection of oral cancer. *Oral Dis.* **2022**, *28*, 1123–1130. [[PubMed](#)]
30. Ghosh, A.; Chaudhuri, D.; Adhikary, S.; Chatterjee, K.; Roychowdhury, A.; Das, A.K.; Barui, A. Deep reinforced neural network model for cyto-spectroscopic analysis of epigenetic markers for automated oral cancer risk prediction. *Chemomet-Rics Intell. Lab. Syst.* **2022**, *224*, 104548. [[CrossRef](#)]
31. Adeoye, J.; Tan, J.Y.; Choi, S.W.; Thomson, P. Prediction models applying machine learning to oral cavity cancer outcomes: A systematic review. *Int. J. Med. Inform.* **2021**, *154*, 104557. [[CrossRef](#)] [[PubMed](#)]
32. Al-Batah, M.S.; Alqaraleh, M.; Alzboon, M.S. Improving oral cancer outcomes through machine learning and dimensionality reduction. *arXiv* **2025**, arXiv:2506.10189.
33. Varalakshmi, D.; Tharahaswari, M.; Anand, T.; Saravanan, K.M. Transforming oral cancer care: The promise of deep learning in diagnosis. *Oral Oncol. Rep.* **2024**, *10*, 100482. [[CrossRef](#)]

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.