Automatic Machine-Learning Classification of the Mode of Invasion of Oral Squamous Cell Carcinoma Using Digital Microscopic Images: A Retrospective Study

Kunio Yoshizawa (✉ yoshizawak@yamanashi.ac.jp)

University of Yamanashi <https://orcid.org/0000-0003-4779-9940>

Hidetoshi Ando

Yamanashi Daigaku Kogakubu Daigakuin Ikonogaku Sogo Kyoikubu Kogaku Senko

Yujiro Kimura

Yamanashi Daigaku Igakubu Daigakuin Sogo Kenkyubu Igakuiki

Shuichi Kawashiri

Kanazawa Daigaku Daigakuin Iyaku Hokengaku Sogo Kenkyuka Iyaku Hoken Gakuiki Igakurui

Akinori Moroi

Yamanashi Daigaku Igakubu Daigakuin Sogo Kenkyubu Igakuiki

Koichiro Ueki

Yamanashi Daigaku Igakubu Daigakuin Sogo Kenkyubu Igakuiki

Research

Keywords: mode of invasion, oral squamous cell carcinoma, random forest, medical image processing, machine learning

Posted Date: August 6th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-52099/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: The Yamamoto-Kohama criteria (YK), which are used to classify the morphology of the infiltrating protrusions of an oral squamous cell carcinoma (OSCC), are clinically useful for determining the mode of tumor invasion, especially in Japan. However, evaluations of the mode of OSCC invasion are based on subjective visual observations, and this approach has created considerable inter-evaluator and inter-facility differences. In this retrospective study, we aimed to develop an automatic method of determining the mode of invasion of OSCC based on the processing of digital medical images of the invasion front.

Methods: Using 101 digitized photographic images of anonymized stained specimen slides from consecutive patients with OSCC at Kanazawa University, we created a classifier that allowed clinicians to introduce feature values and subjected the cases to machine learning using a random forest approach. We then compared the YK grades (1, 2, 3, 4C, 4D) determined by a human oral and maxillofacial surgeon with those determined using the machine learning approach.

Results: The input of multiple test images into the newly created classifier yielded an overall F-measure value of 87%, (Grade 1: 93%, Grade 2: 67%, Grade 3: 89%, Grade 4C: 83%, Grade 4D: 94%). These results suggest that the output of the classifier was very similar to the judgments of the clinician.

Conclusions: We successfully developed an automatic machine-learning method for discriminating the mode of invasion of OSCC. Our results suggest that a medical diagnostic imaging system could feasibly be used to provide an accurate determination of the mode of OSCC invasion.

1. Background

Oral squamous cell carcinoma (OSCC) accounts for approximately 90% of all cases of oral cancer. Despite improvements in treatment options over the past few decades, the 5-year survival rates have remained fairly low (50–55%) among OSCC patients [1, 2]. Treatment failure in a case of OSCC is mainly ascribed to the highly invasive nature of the tumor [3, 4]. As the tumor becomes more invasive, the invasion front progresses from the epithelium through the stroma to infiltrate the lymphatic and blood vasculature [5]. This phenomenon is associated with a poor survival prognosis.

The degree of tumor invasion and metastasis can be evaluated using high-performance imaging modalities such as computed tomography (CT) and magnetic resonance imaging (MRI). However, a histopathological analysis of tumor specimens remains the most important component of an accurate determination of invasiveness [6]. Particularly, an accurate diagnosis of the invasiveness of an OSCC is a very important component of treatment planning and prognostic predictions.

Yamamoto and colleagues initially proposed a method to distinguish and classify the invasive ability of an OSCC [7]. The resulting Yamamoto–Kohama (YK) classification was developed for the microscopic evaluation of the invasion front in a pathological tissue section of OSCC and is now frequently used in

Japan [7]. This classification system appears to be a powerful predictor of regional metastasis in a patient with clinically node-negative OSCC. A YK classification-based evaluation mainly involves biopsy and excised tissues, and the results are used to determine the prognosis and select treatments. However, this evaluation method is based on subjective visual findings and has led to significant differences in determinations between evaluators and facilities. Therefore, the YK classification is not a sufficiently objective index. Furthermore, no report has described the relationship between the visual aspects of infiltration on images of pathological tissues and the results of an objective image evaluation based on samples from patients with OSCC.

Many recent studies have demonstrated the effectiveness of pathological image analysis methods that incorporate artificial intelligence (AI). One experimental study that compared the diagnostic accuracy of a pathologist and AI with respect to pathological images of breast cancer lymph node metastasis determined that the latter was more time-efficient [8]. In another example, reports from various countries have reported that the Gleason score, an index of prostate cancer malignancy, is poorly reproducible among pathologists. In that context, Arvantii et al. demonstrated the use of AI to match the accuracy rate of the Gleason index with its reproducibility among pathologists [9].

Several recent reports have described various approaches that have used machine learning to detect various types of carcinomas from photographic images of lesions, radiological images and pathological specimens [10–18]. However, no published report has described an automatic method for determining the invasion activity based on the computer processing of a digital image of the invasion front. Therefore, in this study, we aimed to develop a method in which medical image processing could be used to determine the mode of invasion automatically based on digital images of the invasion front of an OSCC.

2. Results

2.1 Binary image processing

Figure 1 depicts representative binarized images used to perform the YK classifications in this research. The borderlines between the tumor and stromal tissues could be distinguished clearly up to a fine point and are clearly binarized on the IHC image corresponding to each YK classification.

2.2 Distributional observation of the number of epithelial areas (feature value 1)

The vertical axis of the graph in Fig. 2 indicates the numerical value of the feature amount and the horizontal axis indicates the serial number of the data. In Grades 1 and 2, most of the feature value 1 results were distributed near 1, which was consistent with the single tumor masses observed on the images. In contrast, most of the data sets for Grade 3, 4C and 4D cases yielded values greater than 1, which was consistent with the appearance of multiple tumor masses on the images. Particularly, the feature value 1 for Grade 4D specimens was 15 or higher at a half ratio. This phenomenon was not

observed in the other Grades (Fig. 1). Feature value 1 provided a good distinction of Grades 1 and 2 from Grades 3, 4C and 4D (Fig. 2).

2.3 Distributional observation of the disturbance of borderline” (feature values 2 and 3)

Most Grade 4D specimens yielded a feature value 2 of 0.08 or greater. Accordingly, this feature value could effectively discriminate Grade 4D tumors (Fig. 2). In contrast, feature value 3 was difficult to classify. For example, one Grade 4C case received a uniquely high numerical value of 30 (Fig. 3).

2.4 Distributional observation of a cord-shaped epithelial area (feature value 4)

Notably, this parameter yielded large values for Grade 4C, intermediate values for Grades 2, 3 and 4D and small values for Grade 1 tumors. Accordingly, feature value 4 could successfully distinguish the cord-like Grade 4C tumors from the other tumors (Fig. 4).

2.5 Distributional observation of the size of the epithelial area (feature value 5)

Grade 4C and 4D tumors accounted for most cases with a feature value 5 of 2 or greater. However, only Grade 4D tumors received values greater than 13. The data suggest that feature value 5 is very effective for discriminating Grade 4D (Fig. 5).

2.6 Distributional observation of the borderline length (feature value 6)

Grade 1 tumors tended to yield low values for feature value 6, and most cases with values less than 1000 met the criteria for this grade. Therefore, feature value 6 can effectively discriminate Grade 1 tumors (Fig. 6).

2.7 Confusion matrix-based performance evaluation

As shown in Table 1, the test data of Grades 1 and 4D yielded high classification accuracy values, whereas the data of Grade 2 yielded a low value. Among Grade 1 cases, only 2 of 23 specimens were misjudged as Grade 2. Among Grade 4D cases, only 1 of 18 specimens was misjudged as Grade 4C. However, 4 of 12 Grade 2 specimens were misjudged as Grade 1.

Table 1
Discrimination result with confusion matrix

		Discrimination result				
		Grade 1	Grade 2	Grade 3	Grade 4C	Grade 4D
Test date	Grade 1	21	2	0	0	0
	Grade 2	1	8	2	1	0
	Grade 3	0	1	25	1	0
	Grade 4C	0	1	2	17	1
	Grade 4D	0	0	0	1	17

The test data show the original correct results of YK classification determined by a clinician; on other hand, the discrimination data are the results of the invasion mode determined by machine learning.

2.8 Precision–recall

The precision–recall was calculated using a confusion matrix and reported using F values, as shown in Table 2. The overall F value was 0.87. In an analysis stratified by classification, Grade 2 received the lowest F value of 0.67, whereas Grades 1 and 4D received the highest F values of 0.93 and 0.94, respectively.

Table 2
Precision-Recall

	Number of correct answers	Discrimination number	Number of matches	Precision	Recall	F-measure
Grade1	23	22	21	0.95	0.91	0.93
Grade2	12	12	8	0.67	0.67	0.67
Grade3	27	29	25	0.86	0.93	0.89
Grade4C	21	20	17	0.85	0.81	0.83
Grade 4D	18	18	17	0.94	0.94	0.94

2.9 Comparison of survival rates according to the YK classifications assigned by a clinician or the machine learning method

A comparison of the Kaplan–Meier survival curves calculated for each YK classification revealed a significant difference between the rates associated with the machine learning and clinician classifications only in Grade 2 cases (Fig. 7). Specifically, a Grade 2 classification via machine learning

was associated with a lower survival rate, compared to the same classification when assigned by a clinician ($p < 0.05$). No other significant differences in classification accuracy were observed for the other YK grades (Fig. 7).

3. Discussion

OSCC is characterized by a high degree of invasion into the surrounding tissues, as well as a high incidence of lymph node metastasis [19]. Therefore, it is important to determine the invasive ability of the tumor in each case, as this will inform the establishment of a treatment strategy. In this study, we developed an automatic machine-learning based method for differentiating OSCC cases according to the YK classification criteria. Overall, this system yielded relatively accurate results, as indicated by a high F value of 0.87. However, a further analysis of individual grades yielded a relatively low F value for Grade 2, suggesting that our method may not accurately distinguish these tumors. When we analyzed the survival rates according to the YK grade, the survival rate decreased as the grade determined by the clinician increased. In contrast, however, the machine learning-determined YK Grade 2 cases had the second-worst survival rate after Grade 4D. Moreover, only two-thirds of Grade 2 cases were correctly assigned by the machine learning system, and three-quarters of the mismatched cases actually met the criteria of a higher grade.

Grade 2 may be particularly easy to misjudge via machine learning because these lesions have an unclear borderline and a cord-like shape and are easily misclassified as more invasive tumors (e.g., Grade 4C), even during a subjective clinician-based analysis. Grade 2 cases also comprised the smallest subpopulation in this study. Consequently, machine learning became inadequate, and many cases were misinterpreted. We further consider that Grade 2 was associated with the second-worst survival rate. Although machine learning and AI are being promoted in the medical field, particularly as diagnostic approaches to head and neck cancers, our findings suggest that clinicians should consider the risk of misjudgment when using machine learning, which is instructed using human-determined features [13].

This study was complicated by the fact that that hematoxylin- and eosin-stained (HE) images were largely not used, despite the desirability of such an approach from the perspectives of cost and convenience. However, as this research involved the challenge of a first approach to this technology, we performed IHC to detect claudin-7, which specifically stains OSCC tumor cells, to further clarify the borderline between the tumor and the stroma and ensure clear binary images [20, 21]. The use of HE specimens alone would have made it particularly difficult to capture the sparsely scattered tumor cells in the stroma tissue of Grade 4D specimens. However the inclusion of a claudin-7 IHC analysis better facilitated the detection of tumor cells even in these Grade 4D cases [20]. In the future, advance are needed to ensure that machine learning can detect bivalence using simpler and more useful HE samples.

To improve the classification accuracy using deep learning, it is necessary to include a substantially high number of cases; however, we did not have the required number of cases.

Therefore, a classifier can be created with a limited number of cases by providing a clinician with a minimal amount of learning of capturing supervised image data. Therefore, we initially aimed to facilitate the creation of classifiers via machine learning by setting the features used by clinicians to determine the YK classification. The good overall F value suggests that good feature values were extracted. Moreover, this approach might also be useful for constructing an automatic YK classification discrimination method, although the accuracy must be improved.

The accuracy of machine learning could potentially be improved by dramatically increasing the number of cases. Although many pathological image findings and clinical information can be obtained from The Cancer Genome Atlas database, this information is provided in a pathological image format and the contents are not uniform [22]. Consequently, it is difficult to apply these data in a machine learning setting. In the future, it will be necessary to collect a larger number of cases through a multi-center collaboration. The deep learning and inter-pathologist reproducibility, including the YK classification, encouraged by these efforts will lead to a breakthrough in the field. Furthermore, increasing numbers of patients will benefit when clinicians and pathologists use a more effective AI system. We should continue to cooperate with the field of AI analysis to develop diagnostic tools.

4. Conclusion

In this study, we developed an automatic machine learning-based classifier system to discriminate the mode of invasion of OSCC. Notably, this classifier was confirmed to generate decisions similar to those made by a clinician. Our results suggest that an automatic medical diagnostic imaging system could feasibly and accurately determine the mode of OCC invasion.

5. Methods

5.1. Specimens

Sixty-seven primary OSCC biopsy specimens were obtained from patients who underwent surgical resection at the Department of Oral and Maxillofacial Surgery, Kanazawa University Hospital between 1989 and 2009. The patients (38 male and 29 female subjects) ranged in age from 32 to 91 years (mean age, 60 years). Informed consent for the experimental use of the samples was obtained from the patients according to the hospital's ethical guidelines.

The engineering department of Yamanashi University performed the imaging analysis of the pathological specimens as a third-party assessment organization to eliminate evaluator bias. A total of 101 specimens were evaluated and assigned the following YK grades: Grade 1, 23 specimens; Grade 2, 12 specimens; Grade 3, 27 specimens; Grade 4C, 21 specimens and Grade 4D, 18 specimens (Table 3).

Table 3
Number of images used in this experiment

YK-classification	Grade 1	Grade 2	Grade 3	Grade 4C	Grade 4D
Number of images	23	12	27	21	18

The retrospective study protocol was approved by the ethics committees of Yamanashi University (approval number: 1267) and Kanazawa University (approval number: 1647-1). This study was conducted in accordance with the Declaration of Helsinki.

5.2 Staining methods

Immunohistochemistry (IHC) of deparaffinized and rehydrated sections was performed according to the labeled streptavidin-biotin (LSAB) method as described by Nozaki et al. [23]. To clearly detect tumor cells at the borderline, the sections were reacted overnight at 4 °C with a primary monoclonal antibody specific for claudin-7 (Invitrogen Corp., Camarillo, CA, USA; 200-fold dilution in phosphate-buffered saline [PBS]). This tight junction component was proven to distinguish OCC in an immunohistochemical analysis of pathological tissue specimens according to the YK classification [20]. The sections were then reacted with a secondary antibody (biotin-labeled goat anti-rabbit immunoglobulin polyclonal antibody; Dako Japan, Kyoto, Japan) at room temperature for 60 min. A section of normal oral epithelium previously identified to stain strongly for claudin-7 was used as a positive control in each batch of stained samples. Sections treated with PBS instead of the primary antibody were used as the negative controls.

5.3 Yamamoto–Kohama (YK) classification

In Japan, the departments of oral–maxillofacial surgery at many institutions use the YK classification. This method is used for the histological evaluation of malignant tissues and is focused on the invasion pattern at the tumor–host tissue border. The YK classification was previously shown to be strongly correlated with the risk of lymph node metastasis and prognosis [7]. The YK evaluation criteria are presented in Table 4. In this study, the clinician observed the histopathological specimens of OSCC under a microscope and observed the demarcation and deformation of the border between the tumor and stroma. Figure 8 presents images of typical OSCCs that were stained to detect claudin-7 and classified according to the YK system.

Table 4
Yamamoto-Kohama classification.

Grade	Histologic grading
1	Well-defined borderline
2	Cords, less marked borderline
3	Groups of cells, no distinct borderline
4C	Diffuse invasion, Cord-like type
4D	Diffuse invasion, Widespread type
^a Yamamoto-Kohama classification	

6. Proposed Machine Learning Methods

6.1. Overview

Two approaches to the automatic determination of the OSCC invasion pattern were applied in this study. IHC images of OSCCs were obtained at 100 × magnification. Second, machine learning was applied to cases for which a clinician had previously evaluated the invasion patterns based on the YK classification, and the images were classified by a computer. Here, we considered a shape characterization of the invasive front in the image to be effective for discriminating the invasion mode. The characterized shape was then extracted to create a feature vector.

The proposed processing method was performed as described by Inoue et al. [24] and is summarized in Fig. 9. The invasion mode (i.e., YK classification) was determined automatically using machine learning according to the following methods, which are presented in order [24]: 1. extraction of color features for binarization; 2. creation of classifiers for binarization; 3. binarization of unknown color data; 4. extraction of shape features for the discrimination of YK classification; 5. creation of classifiers for YK classification and 6. discrimination of the YK classification of the binary image.

6.2. Binarization

The histopathological image of each tumor was divided into epithelial and stromal regions to extract the invasion front from the image. First, a borderline was created to divide the tumor epithelium and stromal regions. Binarization was then performed to distinguish the epithelial and stromal sides according to the color image pixels. In this process, the color pixels on the tumor side were converted to black and those on the stromal side to white. Initially, the binarization processing series was performed by a clinical expert, and the resulting human analyst-generated images were used as a training dataset for machine learning. The binarization process is summarized in the upper panel of Fig. 9.

Furthermore, a color image and corresponding binary image were obtained by a clinical expert, prepared as training data and used to create a classifier that would binarize each pixel of an unknown color image [25]. Next, the classifier was used to extract the color data (RGB) values from 49 pixels within a square area of which one side comprised 7 pixels centered on the pixel of interest. We created a classifier that could use machine learning to determine whether a training data image should be classified as black or white based on the color pattern of the pixel of interest and the neighboring pixels [26].

6.3. Features associated with the YK classification and binary images of OSCC

6.3.1. Features associated with YK Grade 1

As shown in Fig. 10, Grade 1 is characterized by a low level of borderline curvature, which is thought to correspond to the YK criterion of a “well-defined borderline.” The representative image also contains many large epithelial masses, with no more than one island-like nest of tumor cells.

6.3.2. Features associated with YK Grade 2

Grade 2 is characterized by a high level of borderline curvature. In many images, the epithelium exhibits protrusions that may be sharp (Fig. 10). These features are thought to correspond to the YK Grade 2 criteria of “cords and a less well-demarcated borderline, indicating that the borderline is slightly disturbed.” Also, as in Grade 1, the tumor is not divided into islands and often appears as a large dendritic mass.

6.3.3. Features associated with YK Grade 3

This grade is distinguished from Grades 1 and 2 by the presence of multiple epithelial tumor islands, as shown in Fig. 10. In addition, Grade 3 OSCC is less likely to exhibit the sharp protrusions observed in Grades 2, 4C and 4D, and contains a generally round-shaped nest of cancer cells.

These features are thought to correspond to the YK Grade 3 criterion of “groups of cells with no distinct borderline.”

6.3.4. Features associated with YK Grade 4C

Grade 4C is characterized by a cord-like epithelial region, as shown in Fig. 10. Morphologically, the tumors tend to have elongated epithelial regions. This appearance is thought to correspond with the YK Grade 4C criterion of “cord-like diffuse invasion.”

6.3.5. Features associated with YK Grade 4D

Grade 4D is characterized by a smaller epithelial region, as shown in Fig. 10. In these lesions, the cancer nests are often too small to identify tumor masses containing only a few cells. This pattern is thought to correspond to the YK Grade 4 criterion of “diffuse invasion (diffuse type),” in which the borderline is unclear and cancerous nests are created.

6.4. Design of the feature extractor

The clinician applied the following five features to automatically classify the binarized image into YK-classification by machine learning: 1. number of epithelial areas, 2. borderline disturbance, 3. cord-shaped epithelial area, 4. size of the epithelial area and 5. borderline length. These five features were determined by an oral surgeon who received instruction about the classification method directly from Professor Yamamoto [7], the proponent of the mode of invasion.

6.5. Features and extraction

6.5.1. Number of epithelial areas

The inputted binary image data were subjected to labeling on the tumor side (i.e., black-colored side). The areas surrounded by continuous black lines were counted to determine the number of epithelial islands. Figure 10 demonstrates that the number of tumor islands, which was defined as feature value 1 [24], increased as the YK grade increased. The data serial numbers were then ordered from YK Grade 1 to 4D, such that Grade 4D data sets had the highest serial numbers.

6.5.2. Disturbance of the borderline

The inputted binary image was then vectorized with respect to the pixels that represented the tumor side of the borderline, which corresponded to the basal cell layer of OSCC. For a labeled object, i , if the length of the borderline is L_i and the number of division points used for vectorization is N_i , the curvature factor (R) of the average borderline in the image can be expressed as Formula (1) [24]:

$$R = \frac{1}{n} \sum_{i=1}^n \frac{N_i}{L_i} \quad (1)$$

Consequently, the curvature of the borderline increases as the number of division points increases, even when the lengths of the borderlines are identical. This curvature is defined as feature quantity 2. Grade 2, which is characterized by more strongly disordered borderlines relative to Grade 1, contains some borderline areas with a high curvature and many others with a low curvature, as shown in Fig. 10. Therefore, it is difficult to distinguish between these grades when using only the average curvature of the borderline. To correct for this non-specificity, the borderline disorder was calculated by combining feature value 2 with feature value 3. This latter value was extracted from the sharp protrusion that occurred when the angle formed by a vector from the middle division point to the front and rear division points was smaller than the threshold value (Fig. 11).

6.5.3. Cord shape of the epithelial area

For an inputted binary image, the highest numerical value yielded by dividing the square root of the labeled object size by the length of the corresponding contour line was extracted and set as feature value 4 [24].

6.5.4. Size of the epithelial area

The number of labeled objects with a size below a certain required threshold was extracted and set as feature value 5 [24].

6.5.5. Length of the borderline

The length of the borderline was set as feature value 6 because this parameter was expected to facilitate the distinction between YK Grades 1 and 2.

6.6 Performance evaluation test using a random forest approach

Next, we experimentally analyzed the resulting discriminant performance when we performed an evaluation based on the YK classification and the extracted features. Here, we used the leave-one-out evaluation method [27, 28] and the random forest machine learning algorithm to create a classifier [29]. Table 5 summarizes the main parameters of the random forests used in the experimental analysis of the image data subjected to machine learning. The F-measure was used as an indicator of precision–recall and was calculated from a confusion matrix that summarized the discrimination analysis of each YK classification [25].

Table 5
Parameters of Random Forests

Items of parameter	Numercal Value
Number of Tree	300
Maximum depth	10
Feature-number of by Random-selection	3
Number of minimum sample	2
Minimum information-gain	0.01

6.7. Comparison of survival rates determined by the machine learning approach and a clinician in accordance with the YK classification

The Kaplan–Meyer method was used to calculate the 5-year overall survival rates associated with each YK classification. The estimates generated by the machine learning approach and a clinician were compared, and the log-rank test was used to evaluate the statistical significance of the inter-group difference. A p-value of < 0.05 was considered to indicate a significant difference.

7. Statistical Analysis

Data analyses were performed using the statistical software SPSS 23.0 for Windows (SPSS, Inc., Chicago, IL, USA).

List Of Abbreviations

OSCC: oral squamous cell carcinoma, IHC: Immunohistochemistry, LSAB: Labeled streptavidin-biotin, Yamamoto–Kohama classification: YK classification, AI: Artificial intelligence.

Declarations

Author's contributions:

KY has planned the study with the help of KU and AM. HA has executed the experimental work with the help of SK and YK. All authors read and approved the final manuscript.

Acknowledgements

We would like to thank all the members of each of departments for their helpful suggestions and support.

The authors would like to thank Enago (www.enago.jp) for the English language review.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets used and /or analyzed during the current study are available from the corresponding author on reasonable request.

Consent for publication

Not applicable.

Ethics approval and consent to participate

This study was approved by the ethics committees of Yamanashi University (approval number: 1267) and Kanazawa University (approval number: 1647-1). This study was conducted in accordance with the

Funding

We are grateful for the fund for the Promotion of Interdisciplinary Research in Yamanashi University. This work was supported by a Grant-in-Aid for Scientific Research (# 18K09807) from the Ministry of Education, Science, Sports and Culture of Japan.

References

1. Sano D, Myers JN: **Metastasis of squamous cell carcinoma of the oral tongue.** *Cancer Metastasis Rev* 2007, **26**(3-4):645-662.
2. Silverman S, Jr.: **Demographics and occurrence of oral and pharyngeal cancers. The outcomes, the trends, the challenge.** *J Am Dent Assoc* 2001, **132** Suppl:7s-11s.
3. Gao X, Guo Q, Wang S, Gao C, Chen J, Zhang L, Zhao Y, Wang J: **Silencing of uPAR via RNA interference inhibits invasion and migration of oral tongue squamous cell carcinoma.** *Oncol Lett* 2018, **16**(3):3983-3991.
4. Yoshizawa K, Nozaki S, Kitahara H, Kato K, Noguchi N, Kawashiri S, Yamamoto E: **Expression of urokinase-type plasminogen activator/urokinase-type plasminogen activator receptor and maspin in oral squamous cell carcinoma: Association with mode of invasion and clinicopathological factors.** *Oncol Rep* 2011, **26**(6):1555-1560.
5. Michikawa C, Uzawa N, Kayamori K, Sonoda I, Ohyama Y, Okada N, Yamaguchi A, Amagasa T: **Clinical significance of lymphatic and blood vessel invasion in oral tongue squamous cell carcinomas.** *Oral Oncol* 2012, **48**(4):320-324.
6. Steens SCA, Bekers EM, Weijs WLJ, Litjens GJS, Veltien A, Maat A, van den Broek GB, van der Laak J, Futterer JJ, van der Kaa CAH, et al: **Evaluation of tongue squamous cell carcinoma resection margins using ex-vivo MR.** *Int J Comput Assist Radiol Surg* 2017, **12**(5):821-828.
7. Yamamoto E, Kohama G, Sunakawa H, Iwai M, Hiratsuka H: **Mode of invasion, bleomycin sensitivity, and clinical course in squamous cell carcinoma of the oral cavity.** *Cancer* 1983, **51**(12):2175-2180.
8. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, van Ginneken B, Karssemeijer N, Litjens G, van der Laak J, Hermesen M, Manson QF, Balkenhol M, et al: **Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer.** *Jama* 2017, **318**(22):2199-2210.
9. Arvaniti E, Fricker KS, Moret M, Rupp N, Hermanns T, Fankhauser C, Wey N, Wild PJ, Rüschoff JH, Claassen M: **Automated Gleason grading of prostate cancer tissue microarrays via deep learning.** *Sci Rep* 2018, **8**(1):12054.
10. Kleppe A, Albregtsen F, Vlatkovic L, Pradhan M, Nielsen B, Hveem TS, Askautrud HA, Kristensen GB, Nesbakken A, Trovik J, et al: **Chromatin organisation and cancer prognosis: a pan-cancer study.** *Lancet Oncol* 2018, **19**(3):356-369.

11. Linder N, Taylor JC, Colling R, Pell R, Alvey E, Joseph J, Protheroe A, Lundin M, Lundin J, Verrill C: **Deep learning for detecting tumour-infiltrating lymphocytes in testicular germ cell tumours.** *J Clin Pathol* 2019, **72**(2):157-164.
12. Sapienza LG, Ning MS, Taguchi S, Calsavara VF, Pellizzon ACA, Gomes MJL, Kowalski LP, Baiocchi G: **Altered-fractionation radiotherapy improves local control in early-stage glottic carcinoma: A systematic review and meta-analysis of 1762 patients.** *Oral Oncol* 2019, **93**:8-14.
13. Marka A, Carter JB, Toto E, Hassanpour S: **Automated detection of nonmelanoma skin cancer using digital images: a systematic review.** *BMC Med Imaging* 2019, **19**(1):21.
14. Basavanahally AN, Ganesan S, Agner S, Monaco JP, Feldman MD, Tomaszewski JE, Bhanot G, Madabhushi A: **Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology.** *IEEE Trans Biomed Eng* 2010, **57**(3):642-653.
15. Doyle S, Monaco J, Feldman M, Tomaszewski J, Madabhushi A: **An active learning based classification strategy for the minority class problem: application to histopathology annotation.** *BMC Bioinformatics* 2011, **12**:424.
16. Yu E, Monaco JP, Tomaszewski J, Shih N, Feldman M, Madabhushi A: **Detection of prostate cancer on histopathology using color fractals and Probabilistic Pairwise Markov models.** *Conf Proc IEEE Eng Med Biol Soc* 2011, **2011**:3427-3430.
17. Ciaccio EJ, Tennyson CA, Bhagat G, Lewis SK, Green PH: **Classification of videocapsule endoscopy image patterns: comparative analysis between patients with celiac disease and normal individuals.** *Biomed Eng Online* 2010, **9**:44.
18. William W, Ware A, Basaza-Ejiri AH, Obungoloch J: **A pap-smear analysis tool (PAT) for detection of cervical cancer from pap-smear images.** *Biomed Eng Online* 2019, **18**(1):16.
19. Klein Nulent TJW, Noorlag R, Van Cann EM, Pameijer FA, Willems SM, Yesuratnam A, Rosenberg A, de Bree R, van Es RJJ: **Intraoral ultrasonography to measure tumor thickness of oral cancer: A systematic review and meta-analysis.** *Oral Oncol* 2018, **77**:29-36.
20. Yoshizawa K, Nozaki S, Kato A, Hirai M, Yanase M, Yoshimoto T, Kimura I, Sugiura S, Okamune A, Kitahara H, et al: **Loss of claudin-7 is a negative prognostic factor for invasion and metastasis in oral squamous cell carcinoma.** *Oncol Rep* 2013, **29**(2):445-450.
21. Lourenço SV, Coutinho-Camillo CM, Buim ME, de Carvalho AC, Lessa RC, Pereira CM, Vettore AL, Carvalho AL, Fregnani JH, Kowalski LP, Soares FA: **Claudin-7 down-regulation is an important feature in oral squamous cell carcinoma.** *Histopathology* 2010, **57**(5):689-698.
22. Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM: **The Cancer Genome Atlas Pan-Cancer analysis project.** *Nat Genet* 2013, **45**(10):1113-1120.
23. Nozaki S, Endo Y, Kawashiri S, Nakagawa K, Yamamoto E, Yonemura Y, Sasaki T: **Immunohistochemical localization of a urokinase-type plasminogen activator system in squamous cell carcinoma of the oral cavity: association with mode of invasion and lymph node metastasis.** *Oral Oncol* 1998, **34**(1):58-62.

24. Inoue A, Ando H, Ueki K, Yoshizawa K, Kato K, Noguchi N, Kawashiri S: **Automatic classification of mode of invasion (YK-criteria) from digital images using machine learning with pathological specimens of oral squamous cell carcinoma.** *MEDICAL IMAGING TECHNOLOGY* 2016, **34**(5):279-285.
25. Martin DR, Fowlkes CC, Malik J: **Learning to detect natural image boundaries using local brightness, color, and texture cues.** *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2004, **26**(5):530-549.
26. Sa I, Ge ZY, Dayoub F, Upcroft B, Perez T, McCool C: **DeepFruits: A Fruit Detection System Using Deep Neural Networks.** *Sensors* 2016, **16**(8).
27. Guyon I, Weston J, Barnhill S, Vapnik V: **Gene selection for cancer classification using support vector machines.** *Machine Learning* 2002, **46**(1-3):389-422.
28. Gavin C. Cawley NLCT: **Efficient leave-one-out cross-validation of kernel Fisher discriminant classifiers.** *Pattern Recognition* 2003, **36**:2585-2592.
29. Breiman L: **Random forests.** *Machine Learning* 2001, **45**(1):5-32.

Figures

Fig.1

Figure 1

Typical binarized images of each Yamamoto-Kohama (YK) classification used in this research. The border lines between the tumor and stromal tissues can be distinguished clearly up to a fine point and are clearly binarized on the immunohistochemistry images corresponding to tumors in each YK class.

Fig.2

Figure 2

Correlation between feature value1 and YK criteria. The horizontal axis indicates the data number, and the vertical axis indicates the amount of the feature value 1. Symbols: -, cases of Grade 1; ◊, date case of Grade 2; ◻, date case of Grade 3; ×, date case of Grade 4C; ○, date case of Grade 4D.

Fig.3

Figure 3

Correlation between feature value2 and feature value3. The horizontal axis shows the amount of the feature 2, and the vertical axis shows the amount of the feature value 3.

Fig.4

Figure 4

Correlation between feature value4 and YK criteria. The horizontal axis shows the data number, and the vertical axis shows the amount of the feature value 4.

Fig.5

Figure 5

Correlation between feature value5 and YK criteria The horizontal axis shows the data number, and the vertical axis shows the amount of the feature value 5.

Fig.6

Figure 6

Correlation between feature value6 and YK criteria The horizontal axis shows the data number, and the vertical axis shows the amount of the feature value 6.

Fig. 7

Figure 7

Kaplan–Meier analysis of survival based on YK criteria. The survival rates based on the modes of invasion determined by the clinician and by machine learning are compared. * $P < 0.05$

Grade 1

Grade 2

Grade 3

Grade 4C

Grade 4D

Fig. 8

Figure 8

Images of mode of invasion (Yamamoto–Kohama [YK] classification). a. Grade 1 has a well-defined borderline. This is equivalent to carcinoma in situ. b. Grade 2 has a less demarcated borderline with cord. c. Grade 3 is characterized by a group of cells, with no distinct borderline. d. Grade 4C appears as cord-like tumor cells with diffuse invasion. e. Grade 4D is characterized by widespread tumor cells with very diffuse invasion.

Binary image processing

Discrimination of mode of invasion

Fig. 9

Figure 9

Image processing to differentiate the mode of invasion a. The binary image processing procedure. b. The procedure used to discriminate the mode of invasion (Yamamoto–Kohama [YK] classification).

Fig. 10

Figure 10

Representative binary image of each of mode of invasion. Black-colored regions indicate the spread of tumor cells, whereas white-colored regions represent stromal tissue that is not infiltrated by tumor cells.

Fig.11

Figure 11

Sharp protrusion of images In this case, the angle formed by a vector from the middle division point to the front and rear division points is smaller than the threshold value. Accordingly, this is detected as a sharp protrusion, and the resulting extracted number is set as the feature value 3.