Received: June 5, 2025 | Revise: November 1, 2025 | Accepted: November 5, 2025

JOURNAL OF APPLIED RESEARCH ON SCIENCE AND TECHNOLOGY
JARST
ISSN ONLINE: 2773-9473
https://ph01.tci-thaijo.org/index.php/rmutt-journal
INSTITUTE OF RESEARCH AND DEVELOPMENT
RAJAMANGALA UNIVERSITY OF TECHNOLOGY THANYABURI

Detection and classification of human teeth in photographs using convolution mask and the watershed algorithm

Sakdinee Rattana and Jessada Tanthanuch*
[1]School of Mathematical Sciences and Geoinformatics, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima, THAILAND
*Corresponding author: jessada@g.sut.ac.th

ABSTRACT

This research presents a novel image-processing algorithm for classifying human teeth in digital photographs, utilizing a combination of convolution masks and the watershed segmentation algorithm. The primary objective is to accurately identify and classify different types of lower teeth in simulated images derived from the SimKit model, a standardized dental simulation framework. The study focuses on digital images of lower teeth captured using a conventional digital camera, simulating real-world photographic conditions. These images are preprocessed through a series of color space transformations and morphological operations designed to enhance the visibility and separation of dental structures from surrounding elements, such as gums or background noise. Since previous studies on photographic images of human teeth have been limited, this work addresses the gap by improving posterior-tooth detection, where conventional watershed methods are less effective. Building on prior use of features such as cusps, grooves, and ridges, the study further employed convolution masks to detect posterior teeth, while watershed segmentation remained effective for the anterior teeth. To implement the classification system, custom software was developed using MATLAB R2020b. This software applies convolution masks to enhance image features, followed by the watershed algorithm, which segments individual teeth and facilitates their classification based on morphological characteristics. The performance of the classification algorithm was quantitatively evaluated using error rate ratios, with the benchmark set at an acceptable error ratio not exceeding 1.00. The evaluation results indicate promising classification accuracy across different tooth categories: incisors exhibited an error ratio of less than or equal to 0.12; canines showed an error ratio of less than or equal to 0.36; premolars maintained a ratio of up to 1.00; and molars demonstrated error ratios not exceeding 0.75. The results confirm that the proposed method is capable of effectively identifying lower teeth in photographic images with high accuracy. This research contributes to the development of automated dental analysis systems and supports the efficient creation of comprehensive databases of tooth types derived from oral photographs, potentially aiding both clinical diagnostics and educational tools in dentistry.

Keywords: Teeth classification, Image Processing, Convolution masks, Watershed algorithm

INTRODUCTION

Teeth are anatomical structures that are part of the digestive system, playing a crucial role as one of the initial stages in the intake of nutrients into the body. These can deteriorate due to various causes and factors. As a result, the general public is commonly advised to get oral health check-ups every six months. However, access to dentistry remains inconvenient for some groups of people. This inconvenience may occur from an excessive number of patients compared to the capacity of public healthcare facilities, difficulties in traveling for individuals in remote areas, and a shortage of dental professionals [1]. In practice, barriers to dental care access not only affect the ability to undergo regular oral health check-ups but also affect the entire treatment process. Misdiagnosis of dental conditions and delays in receiving appropriate treatment often result in severe dental damage.

To reduce the workload of medical personnel and improve the efficiency of healthcare systems, researchers in computer science and information technology have developed applications to support and enhance operational systems, for improving accessibility to medical services. These applications help alleviate the burden on physicians in diagnosing conditions, enhance the accuracy of treatments, reduce the time patients spend in hospitals, and facilitate better data management for healthcare providers [2].

At present, artificial intelligence (AI) is increasingly being integrated into various applications to assist in the analysis of medical images. AI serves as a highly capable tool for enhancing diagnostic

DOI: 10.60101/jarst.2025.262100

efficiency. However, the fundamental knowledge supporting these advancements lies in the field of medical digital image processing. The knowledge in medical image processing is usually applied to major organs, including the brain, lungs, heart, liver, lymph nodes, bones, head, neck, and breasts. Notably, research related to dental image processing remains relatively limited in comparison [3, 4].

Dental imaging refers to images of the human oral cavity captured using various imaging techniques. The most used type of dental image is the X-ray image, which provides detailed visualization of both external and internal structures, including oral tissues, gums, teeth, and adjacent organs. However, this technique requires highly sensitive equipment, which is expensive and requires specialized expertise to operate.

With advancements in digital imaging technology, modern imaging devices have become more widely available, affordable, high-quality, and user-friendly. Consequently, research on digital image processing for dental analysis has gained increasing attention. Moreover, automated image analysis systems have been developed to interpret oral and dental images for accommodating the diversity of dental images and their distinct analytical purposes. Therefore, research in this field has expanded in multiple directions. Common applications of dental image analysis include detecting tooth position, classifying teeth based on type or location, and identifying oral lesions.

In 2010, Zheng, Zhang, and Ding developed the Hybrid Differential Method (HDM) to detect molars in grayscale intraoral images [5]. Subsequently, Kumar, Jarnadan and Larson utilized watershed-based 3D image segmentation to detect cusps and grooves before isolating individual teeth in 3D oral scans [6]. Thereafter, Rad, Rahim and Norouzi applied the integral projection technique to X-ray images to detect tooth edges along a dental row for individual tooth identification [7].

In 2015, Kang et al. applied global thresholding to CT scan images of the lower dental arch to distinguish teeth from the jawbone and other surrounding structures [8]. Yadollahi et al. employed region growing and convex hull algorithms to segment 3D scanned plaster dental models, and then applied the Hough transform to identify the center of each tooth [9]. Amer and Aqel analyzed relative position calculations between bones and teeth to divide X-ray images into smaller segments for the detection of impacted teeth [10].

Moreover, Kang et al. applied thresholding on grayscale images to detect implants in dental photographs taken with digital cameras [11]. In the same year, Lins et al. employed the watershed algorithm to segment individual teeth in intraoral images captured with specialized dental cameras and used the support vector machine (SVM) algorithm to classify the teeth [12]. Finally, Rattana and Tanthanuch devised a segmentation method to isolate teeth from other components in occlusal-view dental images taken with standard digital cameras, utilizing color components, brightness ratios, and morphological operations [13].

As prior works specifically on photographic images of human teeth were scarce. The closest research that this work can be compared with is Lins et al. [12]. However, the angle of the photos taken is different, and hence, the watershed method to separate individual teeth stops working on the posterior teeth part. This is where our work tends to be most fulfilling. Overall, segmentation methods performed poorly on the posterior teeth part of our photos because the lines separating adjacent teeth (tooth edges) were easily confused with other kinds of edges present in the photos, which were mostly cusps/grooves/ridges. It became necessary to consider using other image features. As reviewed, in the work of Kumar, Jarnadan, and Larson [6] and a few uncredited works with 3D images of human teeth, apart from lines/edges, they also benefited from cusps/grooves/ridges, which were specific image features of their uses. However, these features could be attributed to the "pattern" of the teeth when applied to the 2D images. Therefore, we considered using patterns that consisted of cusps/grooves, and ridges to detect each individual posterior tooth in our photos, which led to the creation of convolutional masks to detect these patterns. Water segmentation was still used with the anterior teeth part in our photos, as it still worked well.

MATERIALS AND METHODS

Background knowledge

1. Oral cavity and teeth

Teeth are structures used for chewing food and speaking [14]. Throughout a human lifetime, there are only 2 sets of teeth. The first set, known as primary teeth or deciduous teeth, emerges in children between the ages of approximately 2 and 12 years. Then, these teeth are replaced by permanent teeth.

The permanent teeth consist of 2 corresponding arches: the maxillary (upper) teeth and the mandibular (lower) teeth. Each arch includes four types of teeth, as illustrated in Figure 1, each serving a different function [15]:

- Incisor: Function to cut food into bite-sized pieces.
- Canine: Serve to grip and tear food into smaller pieces.
- Premolar: Function to tear and grind food.
- Molar: Responsible for grinding food into finer pieces

DOI: 10.60101/jarst.2025.262100

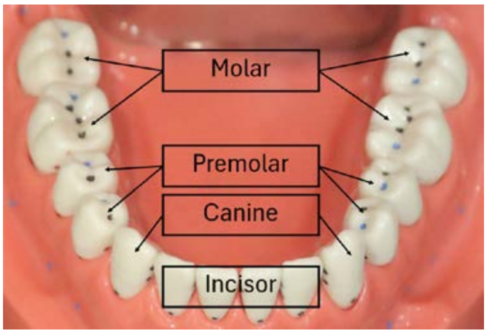


Figure 1 The SimKit lower teeth model of permanent teeth and each type of teeth.

The entire set of teeth and other oral structures, including the tongue, gums, and salivary glands, work together to break down food into smaller pieces, making it suitable for the next stage of digestion.

2. Digital Image and Color Representation

A digital image is composed of small image units called pixels, where each pixel has a different color intensity value. There are various color representations for storing color values in each pixel, such as the HSI (Hue, Saturation, and Intensity), YIQ (Luminance, In-phase, Quadrature), CMYK (Cyan, Magenta, Yellow, and Key (Black)), and RGB (Red, Green, Blue).

For this research, the RGB color representation will be utilized. In this representation, each pixel contains 3 color values: Red (R), Green (G), and Blue (B). The intensity of each color ranges from 0 (darkest) to 255 (brightest).

3. Morphological Operators

Morphological operations are a tool used in computer-based image analysis. They consist of methods that can separate an object in a binary image-composed only of black and white pixels-from its background. Basic morphological transformations include erosion, dilation, opening, closing, edge detection, filling, pixel counting, finding connected components of the object, etc. [16]

This work tends to develop an algorithm for detecting and classifying teeth. To facilitate the evaluation of the algorithm's performance, a permanent teeth model, "SimKit," was utilized, courtesy of the School of Dentistry at Suranaree University of Technology, as a test model. Additionally, software was developed using MATLAB R2020b [17] to support the research process, including both image processing and tooth classification. The research framework follows the workflow diagram, as shown in Figure 2.

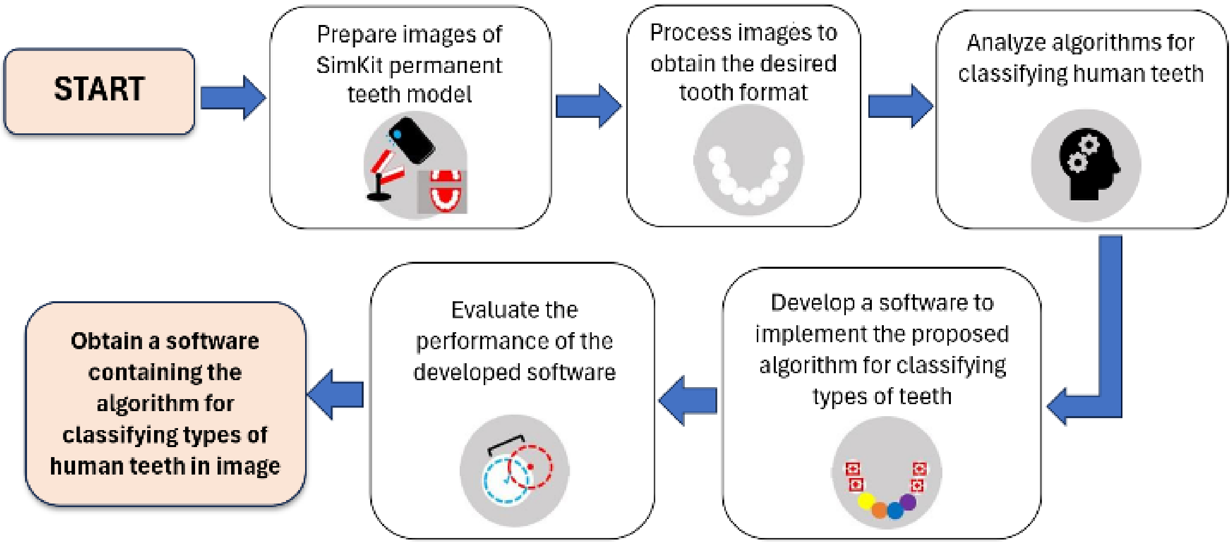


Figure 2 The workflow diagram of this research.

Figure 3 The photograph of the SimKit human permanent dentition model taken with a Canon EOS 80D camera.

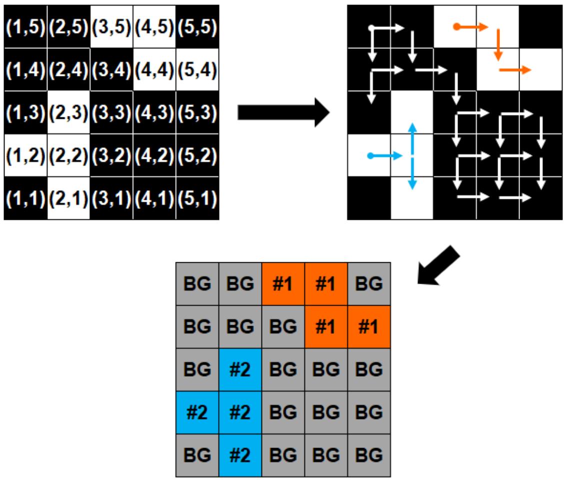
Dental imaging

The images used in this paper are color digital photographs obtained from capturing a model of human permanent teeth. The dental imaging equipment consists of two items: the SimKit human permanent dentition model and a Canon EOS 80D digital camera with an EF-S 18-135mm USM IS lens. The camera's imaging function was set to Auto mode, and the image size was set to 4032 x 3023 pixels with a resolution of 72 dpi in JPEG format. The Shooting angle is the lower occlusal view, captured from the front of the dental model inward, as shown in Figure 3, with a black velvet backdrop.

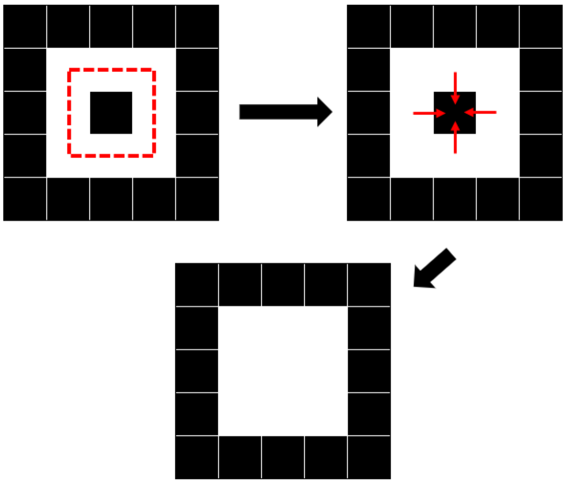
DOI: 10.60101/jarst.2025.262100

Extracting the part of the teeth from the background

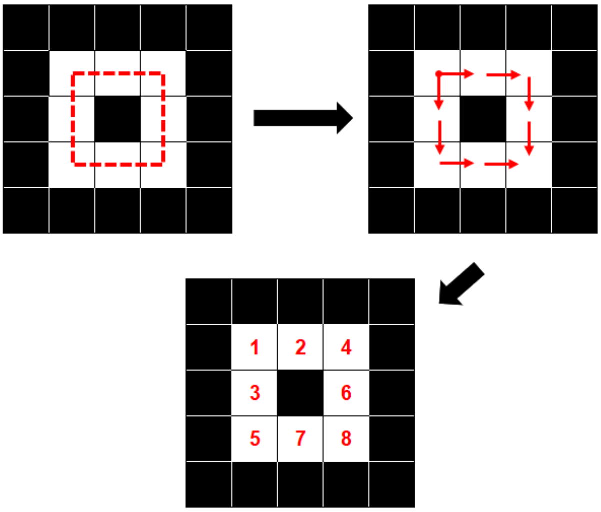
This phase applies image processing techniques inspired by Rattana and Tanthanuch to extract the tooth region from the background in digital images [13]. The research uses colored photographs, and the first step involves applying the RGB color model to isolate the tooth area from other components such as the background and gingiva in the images captured from the SimKit permanent teeth model. Initially, the background is removed by emphasizing the red color component, followed by the removal of the gingival area by emphasizing the green color component.



(a) Defining and labeling the object (adjacent same-color pixels)



(b) Filling holes in the object



(c) Counting image pixels

Figure 4 The morphological operations used include object labeling, filling internal gaps within the object, and pixel counting. (a) Defining and labeling the object (adjacent same-color pixels), (b) Filling holes in the object, and (c) Counting image pixels.

Once these elements are removed, the resulting image is converted into a binary format. Morphological operations are then applied to identify the specific area representing the teeth. The process involves the following preliminary steps:

1. Identifying connected components with similar color characteristics,
2. Labeling the detected objects,
3. Counting the image pixels to identify the largest object,
4. Extracting the largest object group from the remaining parts of the image, presumed to represent the target tooth region, and
5. Filling internal gaps within the object to reduce noise caused by light reflections on the tooth surface in the photograph.

These steps are depicted in Figure 4.

Following the completion of the tooth region identification process, the extracted information is utilized to isolate the tooth area from the original image. The procedure for segmenting the image to extract the tooth region is illustrated in Figure 5.

DOI: 10.60101/jarst.2025.262100

Journal of Applied Research on Science and Technology

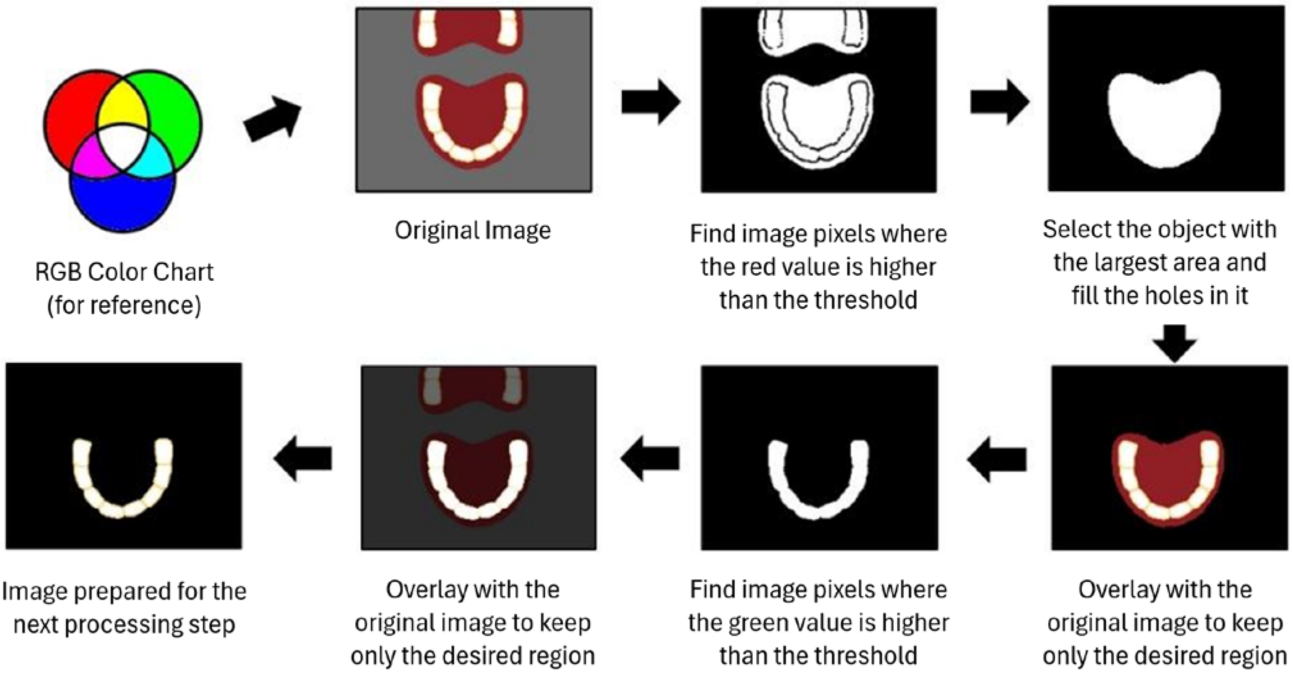


Figure 5 The process of segmenting the image to extract the tooth area from the image.

Application of convolution masks for tooth detection

Based on the observation that posterior teeth-such as premolars and molars-exhibit unique groove patterns characteristic of each tooth position, image convolution is employed to detect these features. Convolution is a process that involves applying a matrix of numerical values (known as a convolution mask) to a digital image, which is composed of pixel intensity values. This operation is commonly used in image processing to detect specific features or for other analytical purposes.

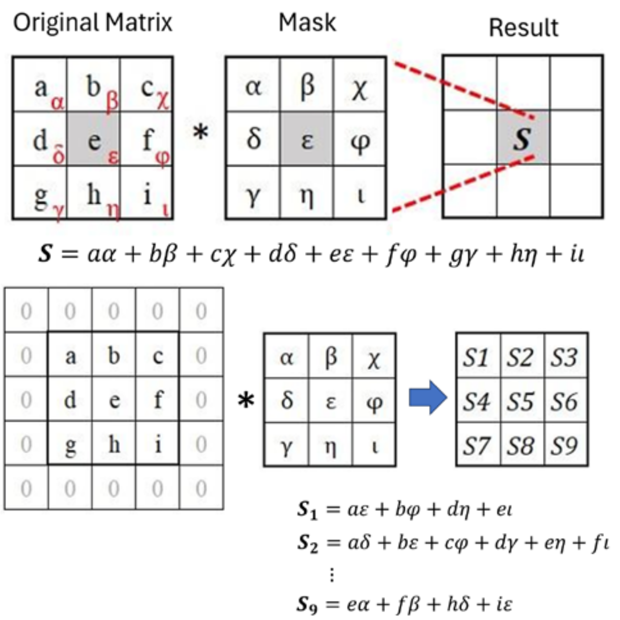


Figure 6 Convolution mask operation.

The convolution mask, or kernel, serves as a filter that highlights targeted patterns within an image, typically designed for grayscale or binary images. In this research, custom convolution masks, an example of which is shown in Figure 6, were developed specifically to detect the characteristic groove patterns of the premolars and molars. The weights within the convolution matrix were carefully configured to enhance the effectiveness of feature detection on grayscale images, following principles outlined by Gonzalez [16].

The creation of a convolution mask for detecting shapes or patterns can be achieved by replicating the form or structure of the target features. After performing convolution between the image and the mask representing the target pattern, it is expected that the maximum value in the resulting output will correspond to the location of the target object within the image, as illustrated in Figure 7.

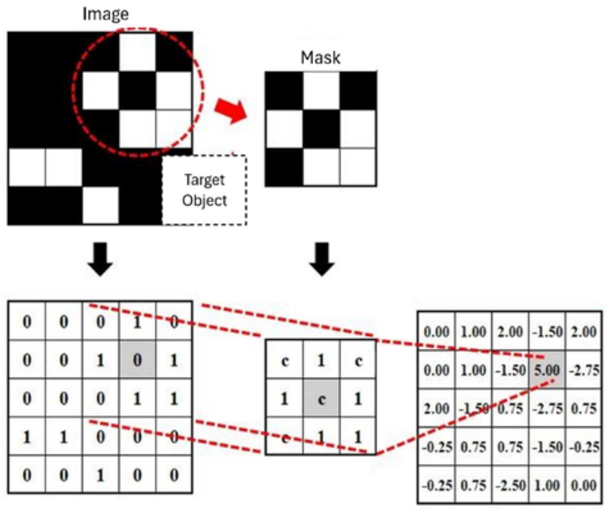


Figure 7 The process of segmenting the image to extract the tooth area from the image.

As shown in Figure 7, this represents the

DOI: 10.60101/jarst.2025.262100

calculation of parameter c,

$$c = -\frac{N_m - N_1}{N_1},$$

where N_m represents the number of pixels in the mask, which equals the sizes of the mask multiplied (height × width), and N_1 is the number of pixels in the mask containing a value of 1. This parameter is used to construct a convolution mask with enhanced robustness against variations in lighting and shading within the image. Parameter c is computed after the mask has been generated and can be automatically calculated using the developed software.

Application of the watershed algorithm in tooth detection

The watershed algorithm is a method used for image segmentation, particularly for grayscale images. It is based on the concept that a watershed line serves as a boundary separating two distinct objects. The fundamental principle of the watershed algorithm is as follows: an edge detection operator such as Sobel or Prewitt is used to generate a gradient map from the intensity values of the image pixels, highlighting potential object boundaries. This gradient map is then interpreted as a topographic surface, where high-intensity pixels correspond to mountain peaks and low-intensity pixels correspond to valleys. Water begins to fill from the lowest points, gradually rising with increasing intensity. When water from different sources merges, watershed lines are formed at their convergence, delineating the boundaries of objects. This method is particularly suitable for detecting and separating multiple convex or semi-convex objects in a single image, especially when there is little or no overlap. As such, it is well-suited for classifying certain types of teeth. An example of applying the watershed algorithm to a tooth-like image is shown in Figure 8.

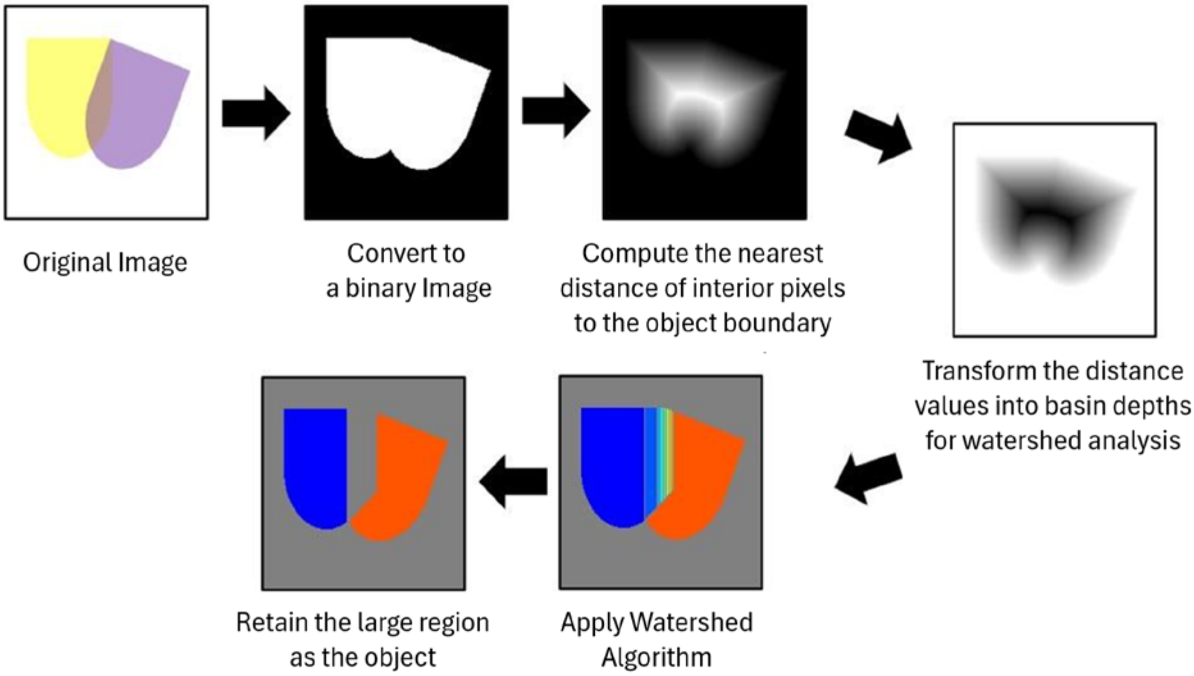


Figure 8 Watershed Algorithm.

Figure 9 Example images of the lower occlusal view of the 'SimKit' dental model.

Data Analysis Tool

Due to the potential for inaccuracies in the proposed process of tooth detection and classification, this study introduces a performance evaluation metric called the error ratio. This metric assesses the accuracy of tooth localization by calculating the ratio between the distance from the predicted

DOI: 10.60101/jarst.2025.262100

position to the detected position of each tooth and the radius of the smallest enclosing circle of that tooth. A lower error ratio indicates higher detection and classification accuracy, with values approaching zero representing optimal performance. Detection is considered successful if the error ratio does not exceed 1.00. This metric effectively reflects the degree to which the detected results match the expected outcomes, relative to the size of the target object.

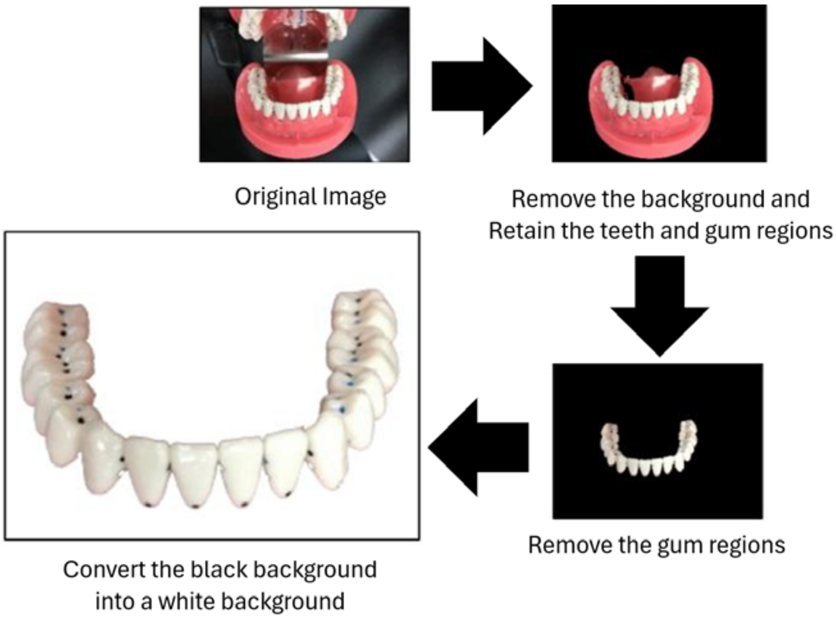
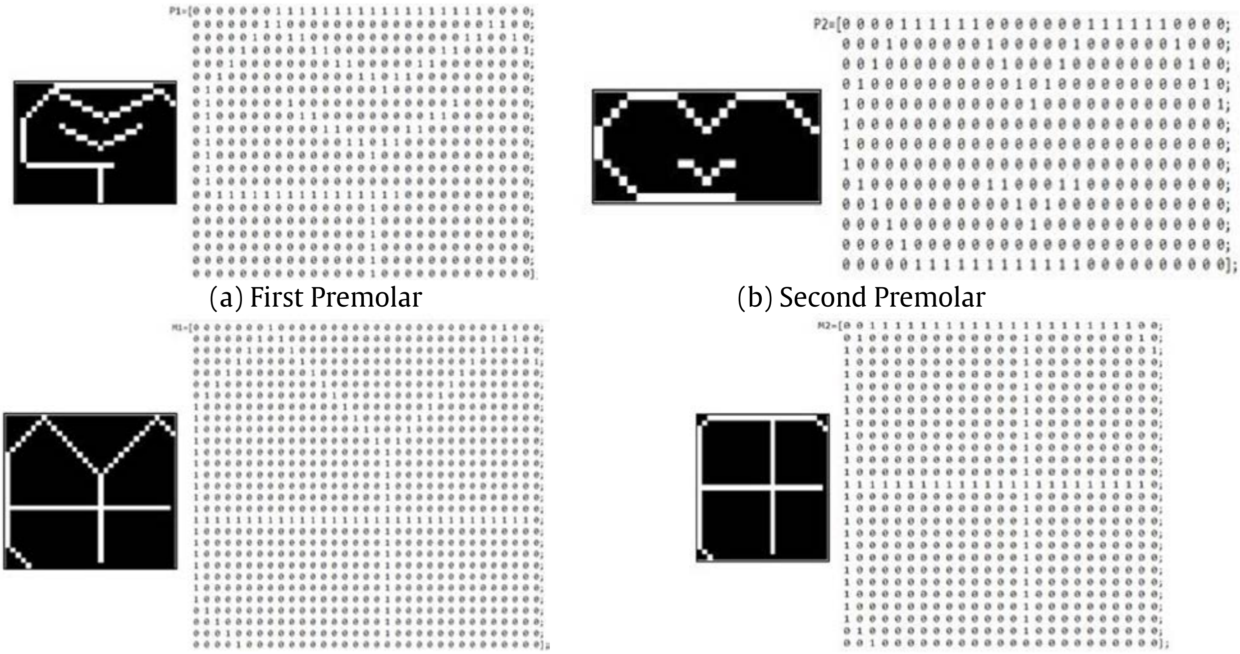


Figure 10 The procedure for extracting the tooth image from the background.



(a) First Premolar (b) Second Premolar

(c) First Molar (d) Second Molar

Figure 11 The convolution mask applied for the detection of premolars and molars (a) First Premolar (b) Second Molar (c) First Molar (d) Second Molar.

RESULTS AND DISCUSSION

Photographic images used in the study

Sample images of occlusal views of lower teeth from the permanent dentition model set "SimKit" are shown in Figure 9, prior to selection for further use in the research.

Preprocessing of images prior to tooth detection and classification analysis

This stage of processing yielded images containing only the tooth regions, as illustrated in Figure 10.

DOI: 10.60101/jarst.2025.262100

Construction of convolution masks for tooth detection

The results of this phase of the study yielded convolution masks specifically created for four distinct tooth patterns. These convolution masks were created by directly copying the groove patterns of each tooth type -first premolar, second premolar, first molar, and second molar - on the left side of the dental model. They were then converted into binary masks and simplified into composites of straight lines to demonstrate robustness. Each of the generated masks was then horizontally flipped to produce an additional four convolution masks for detecting the same types of teeth on the right side of the model. Consequently, a total of eight convolution masks were obtained for detecting all eight tooth types in the images, as illustrated in Figure 11. After obtaining each binary mask, the value contained in the background pixels, which was 0, was changed to the parameter c, and then recalculated using the equation mentioned above for the parameter. As it required only the information on the total number of pixels in the mask and the pixels that contained the value 1 (in the mask) as inputs to the equation, the parameter c was calculated automatically by the software developed for this research. After that, the value 0 in the background pixels changed to the calculated parameter c, before the mask was run on the photo to detect the tooth.

Tooth detection and classification

1. Application of the Watershed Algorithm for Incisors and Canines Detection

In this phase of the study, the watershed algorithm was applied to detect the anterior teeth, which generally exhibit a convex shape. The algorithm successfully detected all four incisors. Subsequently, the centroids of the two outermost detected incisors (one on the far left and one on the far right) were used to estimate the positions of the two canine teeth by applying a linear equation. The steps and the result are shown in Figure 12.

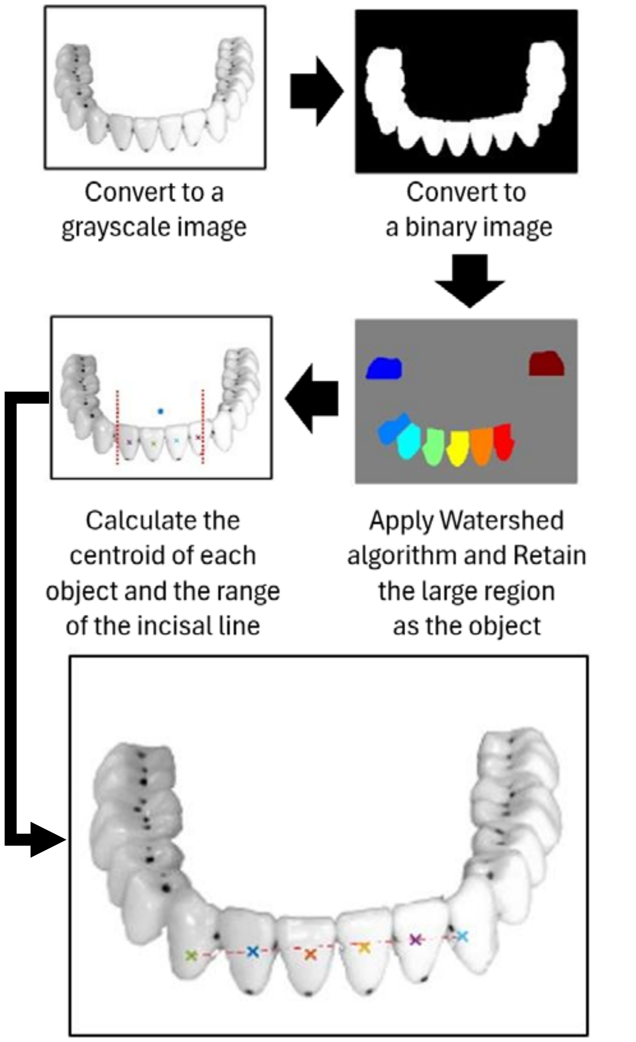


Figure 12 The watershed algorithm to detect incisors combined with the use of a linear equation to calculate the position of the canine teeth.

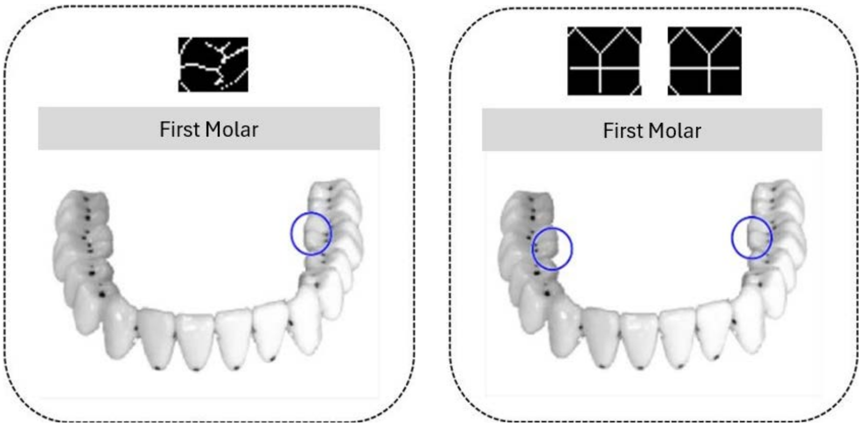


Figure 13 Example results of using a convolution mask to detect molars. (a) The convolution mask designed in the first step (b) The convolution mask designed in the first step but then simplified and combined with the inverted convolution mask.

DOI: 10.60101/jarst.2025.262100

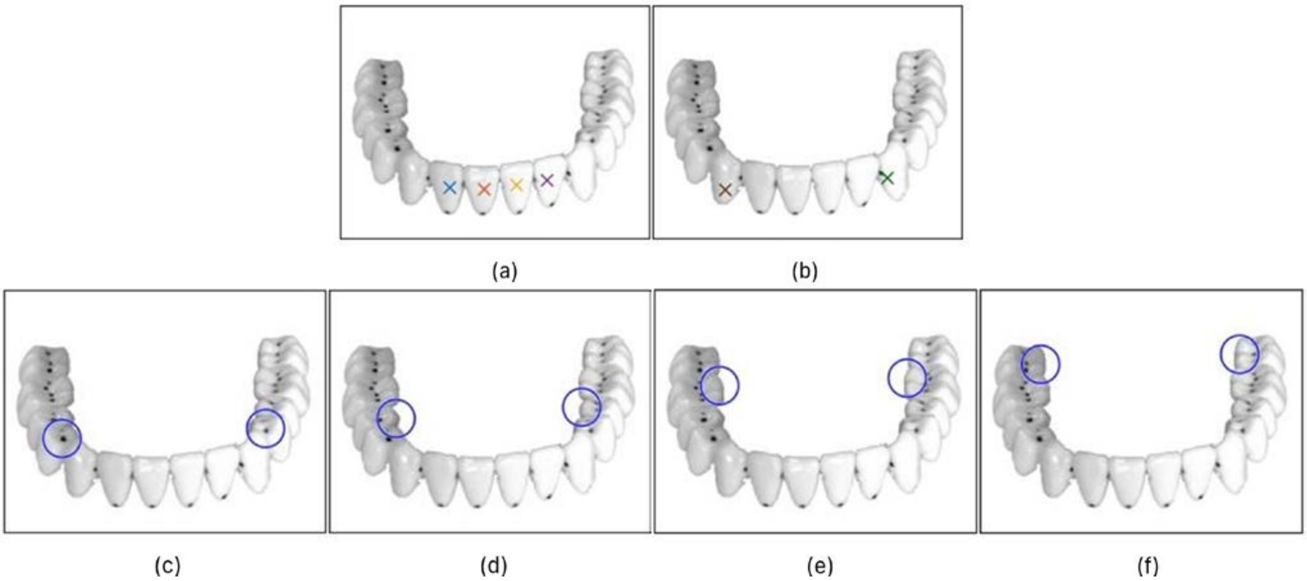


Figure 14 The results of tooth detection and classification: (a) incisors, (b) canines, (c) first premolars, (d) second premolars, (e) first molars, and (f) second molars.

2. Application of Convolution Masks for Premolars and Molars Detection

In this stage, the developed convolution masks were successfully applied to detect and classify the types of molar teeth. The convolution masks in Figure 11 were applied, and some of the results are shown in Figure 13.

Based on the aforementioned process, most types of teeth could be successfully detected and classified, as illustrated in Figure 14.

Evaluation of tooth classification performance

The results of the performance evaluation of tooth classification using the "Error Ratio" are detailed in Table 1.

Table 1 Accuracy of the Proposed Process.

Tooth Type	Error Ratio
Second molar (left)	0.66
Second molar (right)	0.60
First molar (left)	0.75
First molar (right)	0.70
Second premolar (left)	0.88
Second premolar (right)	1.00
First premolar (left)	0.12
First premolar (right)	0.17
Central incisor	0.09
Lateral incisor	0.09
Canine incisor (left)	0.31
Canine incisor (right)	0.36
Third incisor	0.09
Fourth incisor	0.12

CONCLUSION

The results presented in Table 1 demonstrate that the proposed methodology is capable of detecting and classifying all 12 types of lower teeth. Certain types, such as the incisors and first premolars, exhibit relatively low deviation in detection and classification. In contrast, canines and molars present higher deviation levels, with the second premolars showing the lowest detection and classification accuracy. This variation is primarily attributable to the image acquisition process, which most clearly reveals the anterior teeth. The application of the watershed algorithm enables accurate localization and classification of these teeth. However, the convolution mask approach, applied subsequently, must handle posterior teeth, where photographic capture may inadequately reveal the tooth structure and unique morphological patterns necessary for accurate classification. Consequently, a significant difference in deviation ratios between the first and second premolars is observed, despite their anatomical proximity. This suggests that a more appropriately designed convolution mask could mitigate this issue.

There are some limitations to applying the proposed methods to real-world use, as this research was conducted only on photos of a teeth model, not yet on photos of real human teeth. To enhance the robustness of this research, it is essential to utilize a more diverse and human-derived image dataset, especially one that includes convolution masks and various types of teeth, including damaged ones. This study serves as a preliminary effort to facilitate the documentation and classification of teeth prior to the data engineering stage, offering a faster and more efficient alternative compared to processes that lack such foundational support. The next step involves integrating the proposed techniques with

DOI: 10.60101/jarst.2025.262100

machine learning to refine the convolution masks further using real human data. Moreover, the concept of dental arch curves could be incorporated to improve the precision of segmentation by narrowing the effective working area in dental photographs. Various interpolation and curve-fitting techniques can be applied to model these arch curves more accurately. However, the most representative or fitted curve parameters can only be determined from real-human dental data, which are often population-specific and require empirical calibration.

ACKNOWLEDGEMENT

This research was supported by a Royal Scholarship graciously granted by Her Royal Highness Princess Maha Chakri Sirindhorn. The author also wishes to express sincere gratitude to the School of Dentistry and the School of Science, Suranaree University of Technology, for their invaluable support in various aspects of this study.

REFERENCES

1. Estai M, Kanagasingam Y, Huang B, Shiikha J, Kruger E, et al. Comparison of a smartphone-based photographic method with face-to-face caries assessment: A mobile teledentistry model. Telemed J E Health. 2017;23(5):435-40. doi:10.1089/tmj.2016.0122.
2. Silapavitayatorn B, Chitpakdee B. The use of health information technology in nursing for patient safety. J Nurs Health Care. 2020;38(2):6-14.
3. Tang X. The role of artificial intelligence in medical imaging research. BJR Open. 2019 ;2(1):20190031. doi:10.1259/bjro.20190031.
4. Pinto-Coelho L. How artificial intelligence is shaping medical imaging technology: A survey of innovations and applications. Bioengineering. 2023;10(12):1435. doi:10.3390/bioengineering 10121435.
5. Zheng X, Zhang Y, Ding H. The teeth image recognition using HDM and the turn point. In: 2010 Int Conf Optoelectron Image Process; 2010 Nov 11–12; Hainan, China. IEEE Computer Society; 2010. p. 645-8. doi:10.1109/ICOIP.2010. 353.
6. Kumar Y, Janardan R, Larson B. Automatic feature identification in dental meshes. Comput Aided Des Appl. 2012;9(6):747-69. doi:10.3722/ cadaps.2012.
7. Rad A, Rahim M, Norouzi A. Level set and morphological operation techniques in application of dental image segmentation. Int J Comput Inf Eng. 2014;8(4):182-5. doi:10.5281/ zenodo.1091800.
8. Kang HC, Choi C, Shin J, Lee J, Shin YG. Fast and accurate semiautomatic segmentation of individual teeth from dental CT images. Comput Math Methods Med. 2015;2015:810796. doi:10.1155/ 2015/810796.
9. Yadollahi M, Procházka A, Kašparová M, Vyšata O, Marík V. Separation of overlapping dental arch objects using digital records of illuminated plaster casts. Biomed Eng Online. 2015;14:67. doi:10.1186/s12938-015-0066-9.
10. Amer YY, Aqel MJ. An efficient segmentation algorithm for panoramic dental images. Procedia Comput Sci. 2015;65:718-25. doi:10.1016/j.procs. 2015.09.016.
11. Kang X, Xu J, Zou D, Wang X, Wan L, Hu G, et al. Dental biometrics information used for binary image processing, segmentation and matching. J Clin Med Case Stud. 2017;2:22-30.
12. Lins RAS, dos Santos KEA, Neto ADD, Noro L, Roncalli AG, dos Santos Medeiros MC, et al. Use of support vector machine for teeth recognition from occlusal intraoral digital photographic images. In: XIII Simpósio Brasileiro de Automação Inteligente; 2017 Oct 1-4; Porto Alegre - RS, Brazil. p. 1747-52.
13. Rattana S, Tanthanuch J. Teeth region determination in oral cavity images using digital image processing. J Sci Ladkrabang. 2019;28(2):84-93.
14. Rickne CS. Woelfel's Dental Anatomy. 8th ed. Philadelphia: Wolters Kluwer Lippincott Williams & Wilkins; 2011.
15. Rideau Dental Centre. The Four Types of Teeth and Their Functions in Your Mouth [Internet]. [cited n.d.]. Available from: https://rideaudental.ca/blog/education/the-four-types-of-teeth-and-their-functions-in-your-mouth.html.
16. Gonzalez RC, Woods RE. Digital Image Processing. United States of America: Addison-Wesley; 1992.
17. MathWorks. Watershed transform - MATLAB watershed (R2021a) [Internet]. [cited n.d.]. Available from: https://www.mathworks.com/help/images/ref/watershed.html.

DOI: 10.60101/jarst.2025.262100