Comparison of Dental Caries Level Images Classification Performance using KNN and SVM Methods

Yessi Jusman*
Department of Electrical Engineering, Faculty of Engineering Universitas Muhammadiyah Yogyakarta
Yogyakarta, 55183 Indonesia
*Email:yjusman@umy.ac.id

Muhammad Khoirul Anam
Department of Electrical Engineering, Faculty of Engineering Universitas Muhammadiyah Yogyakarta
Yogyakarta, 55183 Indonesia
Email:m.khoirul.ft17@mail.umy.ac.id

Sartika Puspita
Department of Oral Biology, Dental School, Universitas Muhammadiyah Yogyakarta
Yogyakarta, 55183 Indonesia
Email:sartika.puspita@umy.ac.id

Edwyn Saleh
Department of Oral Biology, Dental School, Universitas Muhammadiyah Yogyakarta
Yogyakarta, 55183 Indonesia
Email:edwyn.saleh@umy.ac.id

Siti Nurul Aqmariah Mohd Kanafiah
School of Mechatronics Engineering University Malaysia Perlis
Pauh Putra Campus, 02600 Perlis, Malaysia
Email:aqmariah@unimap.edu.my

Rhesezia Intan Tamarena
Department of Electrical Engineering, Faculty of Engineering Universitas Muhammadiyah Yogyakarta
Yogyakarta, 55183 Indonesia
Email:rhesezia.intan.2016@umy.ac.id

2021 IEEE International Conference on Signal and Image Processing Applications (ICSIPA) | 978-1-6654-3592-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICSIPA52582.2021.9576774

***Abstract*—This study aims to build a dental caries level classification system based on image processing (i.e. to extract texture features) and machine learning methods. The first step was to analyze and discover the extraction results from Gray Level Co-Occurrence Matrix algorithm. After successfully extracting the features, the classification was carried out using a Support Vector Machine (SVM) and K-Nearest Neighbors (KNN). Both machine learnings are analyzed and used to obtain the better alternatives of the classification results. This study employed radiographic images of four dental caries classes consisting of Class 1, 2, 3, and 4. Total of images used after pre-processing are 396 images. Training data is 90% of total images then the rest is the testing data. The classification obtained accuracy value of the SVM and KNN. The SVM classification method revealed the highest accuracy value generated by the Fine Gaussian SVM model was 95.7%. Conversely, the lowest accuracy value generated was 83.3%, derived from the Quadratic SVM model. Meanwhile, the highest accuracy by using KNN is 94.9% of accuracy using Fine and lowest accuracy value generated was 91.4%, derived from Weighted KNN models. The KNN classification results are better than the SVM results.**

Keywords—caries images, X-ray images, Hu's moment invariants, classification, analysis.

I. Introduction

Curvature of the teeth often occur due to the lack of dental care, causing them to not be optimal in processing food. Moreover, teeth can also be a nest of dirt that causes plaque and erodes the teeth, thus causing cavities or caries in the teeth. Patric Kiel Navarro et al. conducted a study on dental images detected using the SVM method, and the decision tree image was processed with Histogram Equalization and augmented to 10x10. The results obtained 84% and 78% accuracy [1]. Dental caries generally occurs due to consuming foods containing carbohydrates, such as sucrose, and rarely brushing teeth, thereby gradually damaging the layers and structures of the teeth [2].

The Ministry of Health of the Republic of Indonesia published on its website, www.kemkes.go.id that in 2016, the global burden of disease study stated that dental and oral health problems, especially dental caries, were one of the diseases experienced by almost half of the world's population, totaling 3.58 billion people. In addition, gum disease is the 11th most common disease in the world. The basic health research also declared that 45.3% of dental problems in Indonesia in 2018 were damaged or perforated teeth.

All doctors in Indonesia establish the diagnosis of dental caries using radiographs or X-Ray rays. However, sometimes the results of radiographs or X-Ray rays are unclear and cause obstacles to diagnosing dental caries. Therefore, all doctors must follow the protocol in reading the results of radiographs or X-Ray rays. Hence, the examination of the patient's teeth can detect whether he has caries on the teeth or not [3].

A new segmentation method based on level set (LS) was proposed in two phases: IC generation using morphological information of image and intelligent level set segmentation utilizing motion filtering and backpropagation neural network. The segmentation results were efficient and accurate as compared to other studies [2]. Otherwise, the research uses segmentation, cropping, and enhancement for the pre-processing process. feature extraction using edge detection. The results obtained are maxillary accuracy of 81.14%, and mandibular accuracy of 73.63%. By using a genetic-based approach to tooth extraction can increase accuracy [4]. The other research used the same preprocessing process. The results obtained were 98.9% sensitivity and 99.6% precision for tooth detection [5].

Review of segmentation process for diagnosis of dental caries were conducted [6], [7]. Comparisons of dental radiography analysis algorithms were carried out based on several research groups [8], [9], [10], [11]. Some deep learning research on caries levels has been presented [12], [13], [14]. Other research's of recognition in dental radiographs by Hu's moment invariants was discussed [15]. The implementation of Hu's moment were explained [16], [17], [18], [19], [20], [21], [22]. For texture features, there are researchs of dental caries employing it [8], [10], [23].

Based on the implementation of Gray Level Co-Occurrence Matrix (GLCM) algorithm, this study is limited for dental caries images. The texture feature is the important features to differentiate the caries levels. Thus, in this study, the authors performed GLCM algorithm and artificial intelligence methods to detect dental caries correctly. The GLCM as features extraction method for dental images extracted the texture features. Then the classification process was carried out by the Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) to obtain dental image results diagnosed according to each class of dental caries.

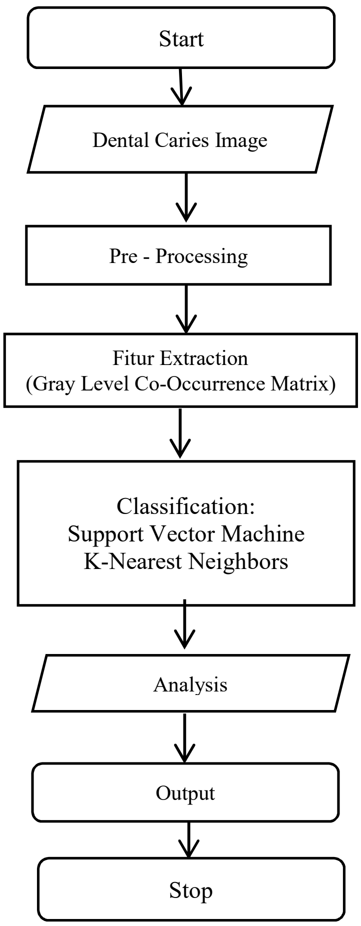


Fig. 1. Research flow chart

II. METHODOLOGY

This study implemented various stages, which started with inputting dental caries images, and then the pre-processing stage was carried out by varying the number of images. The next stage was the feature extraction process by GLCM algorithm, and the feature extraction results were continued to the classification process by the SVM and KNN. These various stages were processed using the hardware of Intel Core i5 9400f, 16.00 GB memory, and 6 GB Nvidia RTX 2060 graphics. The stages in this study are presented in Figure 1.

A. Data and Tools

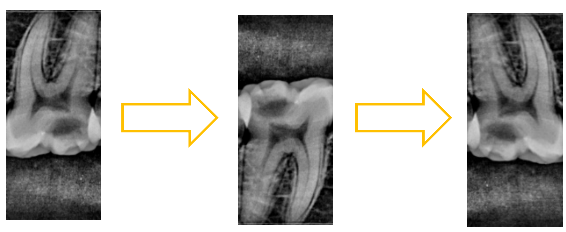
This study utilized radiographic results of four classes of dental caries images, namely Class 1, 2, 3, and 4. Class 1 consisted of 81 images, Class 2 comprised 63 images, Class 3 contained 36 images, and Class 4 involved 18 images. The radiographic results of dental caries images were obtained from the Dental and Oral Hospital (RSGM) of Universitas Muhammadiyah Yogyakarta. The hardware specifications are displayed in Table 1.

TABLE I. HARDWARE SPECIFICATIONS

Hadware Memory	**Characteristic**
Processor	16 Gb
Graphics	Intel Core i5-9400 CPU @ 2.90 GHz
Hadware	GeForce GTX 970 4Gb

B. Pre–Processing Stage

In the pre-processing stage, augmentation process are performed to vary the number of images. This augmentation process consisted of rotating 180 degrees and flipping horizontally to the original image. These processes were chosen because the possible direction of the teeth is only vertical. Total of images used after pre-processing are 396 images. Training data is 90% of total images then the rest is the testing data. The image resolution used was 445 x 1169 pixels. The results of pre-processing are demonstrated in Fig. 2. The rotated 180 degrees results are presented in Fig. 2b, and the horizontal flip results are presented in Fig. 3c.



(a) (b) (c)

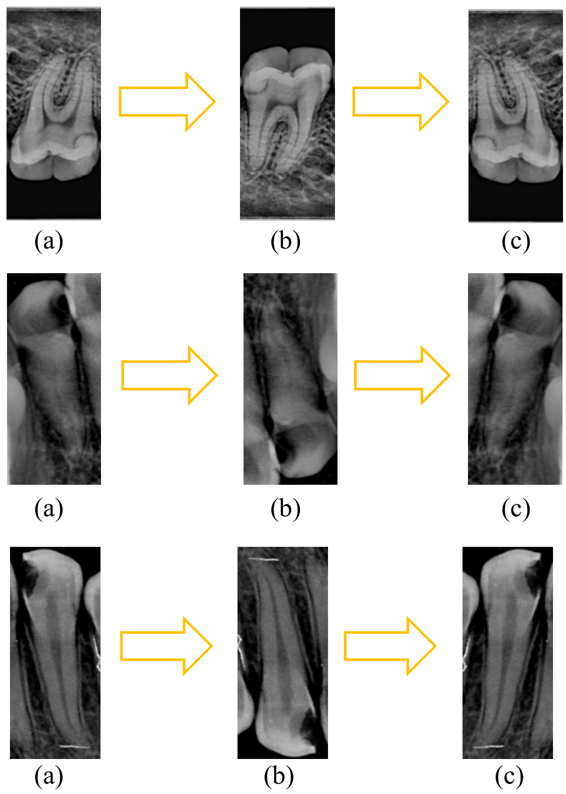


Fig. 2. (a) Initial condition of the image before processing b) Image result after being rotated 180 degrees (c) Image result after the horizontal flip.

C. *Feature Extraction Stage with GLCM Algorithm*

The GLCM algorithm was applied as the feature extraction method in this study. The pre-processing process was carried out first, followed by the extraction with GLCM algorithm, resulting in 16 features in each extracted image.

D. *Classification and Analysis Processes*

Ten folds cross validation is used in this study to arrange the dataset. The feature extraction results obtained by using the GLCM algorithm were then classified using the SVM. The SVM method aimed to discover the best hyperplane functioning as a separator of two data groups in the input space.

The first step in the SVM algorithm process was to input the dataset, proceed with the calculation process with the SVM kernel function, and then was followed by the SVM training process. The SVM process was the testing process on the dataset. Finally, the last stage was the classification evaluation. The following is an explanation of the SVM algorithm process flow. The KNN also was used to train and test to show the accuracy results.

III. RESULTS AND DISCUSSIONS

Features extraction and classification are the main steps in this study to process four classes of dental caries images. The feature extraction process with GLCM algorithm produced feature values owned by each image. Hence, the feature extraction results could be presented in a table and used for the classification process. This classification process utilized three SVM models and two KNN models: Cubic SVM, Quadratic SVM, and Fine Gaussian SVM, Fine KNN, and Weight KNN to produce accuracy values.

A. *Feature Extraction Results*

The feature extraction process using GLCM algorithm produced 16 features. The resulting feature extraction values were used to distinguish one image from another. Thus, each image had a feature extraction value.

Table 2 exhibits feature extraction results, the average value, and standard deviation of each class, namely Class 1, 2, 3, and 4. Hence, the system could distinguish each dental caries image based on its class.

B. *Classification Results*

The classification process employed the three SVM models and two KKN models: Cubic SVM, Quadratic SVM, Fine Gaussian SVM Fine KNN, and Weight KNN, to obtain an accuracy value by running ten times. Table 3 depicts the accuracy values and classification times.

Based on Table 3, the highest accuracy value is in the Fine Gaussian SVM model, obtaining an accuracy value of 95.7% with a time of 2.8 s. In contrast, the lowest accuracy value is discovered in the Cubic SVM model with a value of 83.3% and a time of 2.1 s. Meanwhile, the highest accuracy by using KNN is 94.9% of accuracy using Fine and the lowest accuracy by Weighted KNN models is 91.4%.

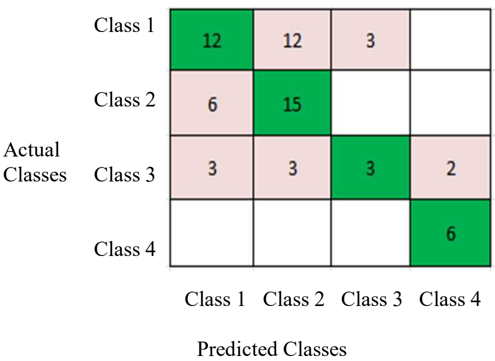
The duration times are 0.8 and 0.4 seconds for fine and weighted KNN models, respectively. The averages of duration times are 1.3 and 1.2 seconds for fine and weighted KNN models, respectively. Based on the Table 2, the feature values are significantly different, so it can be grouped the images data features well. Thus, the basic theory for KNN classification is realized in this case. Because of that, the KNN accuracy values can achieved higher than SVM for this study.

Based on Figure 3, the results of the classification of dental caries images on testing data in the form of a confusion matrix are classified with 3 SVM models. For testing data on the classification of the Cubic SVM model detected as Class 1 which is 12 images, Class 2 is 15 images, Class 3 is 3 images and Class 4 is 6 images. The Quadratic SVM model detected as Class 1 is 18 images, Class 2 is 15 images, Class 3 is 8 images and Class 4 is 6 images. For the Fine Gaussian SVM model detected as Class 1 which is 18 images, Class 2 is 15 images, Class 3 is 8 images and Class 4 is 6 images.

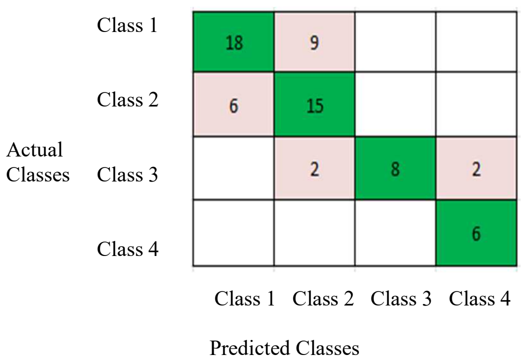
Meanwhile, for the results of the classification of testing data based on the KNN Fine model, it is detected as Class 1, which is 24 images. Class 2 is 19 images. Class 3 is 12 images. Class 4 is 6 images. For the KNN with Weighted model, the results of the classification of testing data, it is detected as Class 1, which is 24 images. Class 2 is 21 images. Class 3 is 7 images. Class 4 is 6 images.

TABLE II. FEATURE EXTRACTION RESULTS

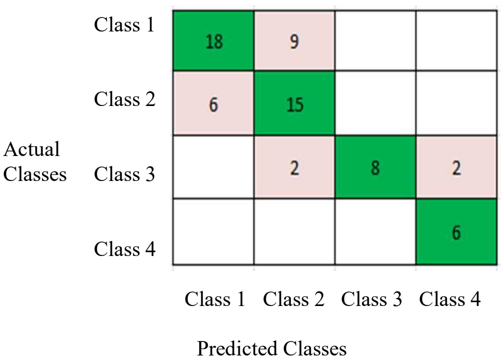
No	GLCM Features		Average ± Standard Deviation			
			Class 1	Class 2	Class 3	Class 4
1	*Contrast*	0°	39.21 ± 17.090	53.53 ± 22.240	56.77 ± 17.700	64.329 ± 19.844
		45°	70.33 ± 23.220	79.52 ± 29.770	54.87 ± 19.660	61.17 ± 20.907
		90°	72.09 ± 21.340	68.77 ± 18.220	45.56 ± 13.080	43.14 ± 11.950
		135°	72.08 ± 23.250	82.13 ± 30.590	54.42 ± 18.210	63.24 ± 22.350
2	*Correlation*	0°	0.46 ± 0.210	0.27 ± 0.210	0.0146 ± 0.142	- 0.065 ± 0.109
		45°	0.031 ± 0.200	- 0.13 ± 0.190	- 0.028 ± 0.153	- 0.084 ± 0.160
		90°	0.018 ± 0.160	0.015 ± 0.160	0.142 ± 0.163	0.24 ± 0.106
		135°	0.015 ± 0.190	- 0.15 ± 0.180	- 0.034 ± 0.151	- 0.119 ± 0.160
3	*Energy*	0°	0.015 ± 0.018	0.009 ± 0.008	0.0056 ± 0.002	0.0057 ± 0.002
		45°	0.007 ± 0.006	0.005 ± 0.002	0.0060 ± 0.002	0.0062 ± 0.002
		90°	0.007 ± 0.008	0.0048 ± 0.002	0.0057 ± 0.002	0.0063 ± 0.002
		135°	0.008 ± 0.007	0.005 ± 0.002	0.006 ± 0.002	0.0063 ± 0.002
4	*Homogeneity*	0°	0.36 ± 0.073	0.316±0.046	0.27±0.035	0.268±0.039
		45°	0.25±0.044	0.24±0.038	0.26±0.365	0.25±0.380
		90°	0.26±0.041	0.27±0.027	0.31±0.029	0.321±0.299
		135°	0.25±0.046	0.24±0.039	0.26±0.033	0.249±0.041



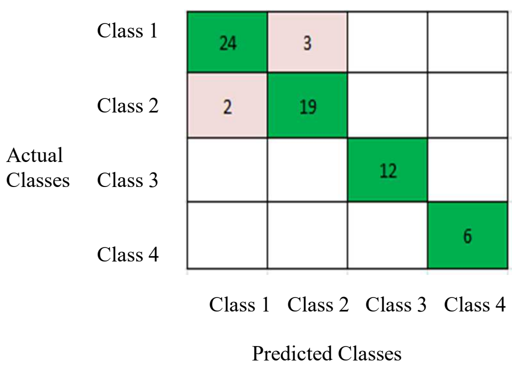
(a)



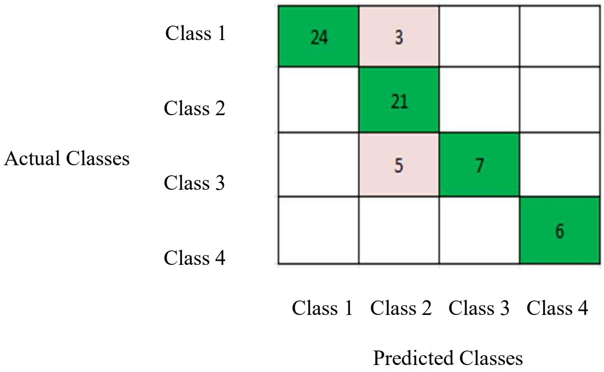
(b)



(c)



(d)



(e)

Fig. 3. Confusion matrix of (a) Cubic SVM, b) Quadratic SVM, (c) Fine Gaussian SVM, (d) Fine KNN, (e) Weighted KNN.

TABLE III. SUPPORT VECTOR MACHINE (SVM) AND K-NEAREST NEIGHBORS (KNN) CLASSIFICATION RESULTS

Datasets	*Cubic* SVM		*Quadratic* SVM		*Fine Gaussian* SVM		*Fine* KNN		*Weighted* KNN	
	Acc (%)	Time (s)	Acc (%)	Time (s)	Acc (%)	Time (s)	Acc (%)	Time (s)	Acc (%)	Time (s)
Run 1	94.4	6.3	84.8	6.4	92.1	6.2	94.1	4.8	92.9	4.5
Run 2	94.4	2.1	**83.3**	2.1	90.7	2.0	93.9	0.8	92.3	1.0
Run 3	93.1	1.5	85.2	1.6	90.4	1.5	94.1	0.8	**91.4**	0.4
Run 4	93.3	1.6	85.7	1.6	90.6	1.5	93.8	0.8	92.4	0.9
Run 5	94.6	1.5	85.4	1.5	90.4	1.5	94.3	1.9	93.1	0.9
Run 6	93.6	2.0	85.7	2.0	90.6	1.9	93.4	0.7	91.9	0.9
Run 7	94.1	5.0	85.5	2.0	91.9	2.6	**94.9**	0.8	93.4	0.4
Run 8	93.3	2.8	85.9	2.9	**95.7**	2.8	93.3	0.8	92.6	0.9
Run 9	94.3	1.0	85	1.0	90.1	1.0	93.3	1.0	92.8	1.2
Run 10	94.6	1.5	84.8	1.6	90.6	1.5	93.6	0.8	92.1	0.4
Averages	**93.97**	**2.5**	**85.13**	**2.3**	**91.31**	**2.2**	**93.87**	**1.3**	**92.49**	**1.2**
Standard Deviation of Samples	**0.59**	**1.7**	**0.75**	**1.5**	**1.67**	**1.49**	**0.51**	**1.2**	**0.60**	**1.2**

IV. CONCLUSIONS

Classification of dental caries level images using texture features could be developed using Support Vector Machine (SVM) and K-Nearest Neighbors (KNN). The features of the gray level co-occurrence matrix (GLCM) algorithm have significantly value to differentiate the caries levels. The averages of cubic, quadratic, and fine gaussian SVM models are sequentially 93.97%, 85.13%, and 91.31% of accuracy values and duration times are longer than the KNN models results. The averages of fine and weighted KNN models are 93.87% and 92.49% of accuracy values and duration times are shorter than the SVM models results. The SVM classification method revealed the highest accuracy value generated by the Fine Gaussian SVM model was 95.7%. Conversely, the lowest accuracy value generated was 83.3%, derived from the Quadratic SVM model. Meanwhile, the highest accuracy by using KNN is 94.9% of accuracy using Fine and lowest accuracy value generated was 91.4%, derived from Weighted KNN models. The KNN classification results are better than the SVM results.

ACKNOWLEDGMENT

This research is supported by Universitas Muhammadiyah Yogyakarta and a research project grant from the Ministry of Research and Technology of the Republic of Indonesia.

REFERENCES

[1] Navarro, P.K., et al., Detecting Smooth Surface Dental Caries in Frontal Teeth Using Image Processing, in Proceedings of the 2019 3rd High Performance Computing and Cluster Technologies Conference2019, Association for Computing Machinery: Guangzhou, China. p. 167–171.

[2] Rad, A.E., et al., Automatic computer-aided caries detection from dental x-ray images using intelligent level set. Multimedia Tools Appl., 2018. 77(21): p. 28843–28862.

[3] Krithiga, R., A Survey: Segmentation in Dental X-ray Images for Diagnosis of Dental Caries, 2016, January.

[4] Leite, A.F., et al., Artificial intelligence-driven novel tool for tooth detection and segmentation on panoramic radiographs. Clinical oral investigations, 2021. 25(4): p. 2257-2267

[5] Haghanifar, A., M.M. Majdabadi, and S.-B. Ko. Automated Teeth Extraction from Dental Panoramic X-Ray Images using Genetic Algorithm. in 2020 IEEE International Symposium on Circuits and Systems (ISCAS). 2020. IEEE.

[6] Agrawal, A. and R.K. Bhogal. A Review—Edge Detection Techniques in Dental Images. in International Conference on ISMAC in Computational Vision and Bio-Engineering. 2018. Springer.

[7] Wang, C.-W., et al., A benchmark for comparison of dental radiography analysis algorithms. Medical image analysis, 2016. 31: p. 63-76.

[8] Obuchowicz, R., et al., Caries detection enhancement using texture feature maps of intraoral radiographs, 2018, Springer.

[9] Agrawal, A. and R.K. Bhogal. Edge Detection Techniques in Dental Radiographs (Sobel, T1FLS & IT2FLS). in International Conference on Communication, Networks and Computing. 2018. Springer.

[10] Jusman, Y., et al. Analysis of Features Extraction Performance to Differentiate of Dental Caries Types Using Gray Level Co-occurrence Matrix Algorithm. in 2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE). 2020. IEEE.

[11] Duong, D.L., M.H. Kabir, and R.F. Kuo, Automated caries detection with smartphone color photography using machine learning. Health Informatics Journal, 2021. 27(2): p. 14604582211007530.

[12] Miki, Y., et al., Classification of teeth in cone-beam CT using deep convolutional neural network. Computers in biology and medicine, 2017. 80: p. 24-29.

[13] Lee, J.-H., et al., Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. Journal of Dentistry, 2018. 77: p. 106-111.

[14] Casalegno, F., et al. (2019). "Caries detection with near-infrared transillumination using deep learning." Journal of dental research 98(11): 1227-1233.

[15] Pattanachai, N., N. Covavisaruch, and C. Sinthanayothin. Tooth recognition in dental radiographs via Hu's moment invariants in 2012 9th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology. 2012.

[16] Liu, Y., Y. Yin, and S. Zhang. Hand Gesture Recognition Based on Hu Moments in Interaction of Virtual Reality. in 2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics. 2012.

[17] Zhang, Y., et al., Pathological brain detection based on wavelet entropy and Hu moment invariants. Bio-Medical Materials and Engineering, 2015. 26: p. S1283-S1290.

[18] Zhang, Y.-D., et al., Alcoholism detection by medical robots based on Hu moment invariants and predator-prey adaptive-inertia chaotic particle swarm optimization. Computers & Electrical Engineering, 2017. 63: p. 126-138.

[19] Salleh, M.A.M., et al., Features Extraction to Differentiate of Spinal Curvature Types using Hue Moment Algorithm. Journal of Physics: Conference Series, 2020. 1471: p. 012060.

[20] Zhang, Y., et al., Pathological brain detection in MRI scanning via Hu moment invariants and machine learning. Journal of Experimental & Theoretical Artificial Intelligence, 2017. 29(2): p. 299-312.

[21] Gornale, S.S., P.U. Patravali, and P.S. Hiremath, Automatic Detection and Classification of Knee Osteoarthritis Using Hu's Invariant Moments. Frontiers in robotics and AI, 2020. 7: p. 591827-591827.

[22] Sari, B.P. and Y. Jusman. Classification System for Cervical Cell Images based on Hu Moment Invariants Methods and Support Vector Machine. in 2021 International Conference on Intelligent Technologies (CONIT). 2021.

[23] Rad, A.E., M.S.M. Rahim, and A. Norouzi, Digital dental X-ray image segmentation and feature extraction. TELKOMNIKA Indonesian Journal of Electrical Engineering, 2013. 11(6): p. 3109-3114.