Machine Learning Classification with GLCM and Watershed for Detecting Pulpitis

Mareska Pratiwi Maharani*, Philip Tobianto Daely†, Jae Min Lee‡, and Dong-Seong Kim§
*†‡§Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea
†Department of Information Technology, Institut Teknologi Telkom Surabaya, Surabaya, Indonesia
*mareskapm@kumoh.ac.kr, †philip.daely@kumoh.ac.kr, ‡ljmpaul@kumoh.ac.kr, §dskim@kumoh.ac.kr

***Abstract*—Pulpitis also has known as a disease that causes painful feelings because of the inflammation in teeth pulp that contains a lot of nerves and blood vessels. The usual method to detect this disease is with a periapical radiograph, but this method can only produce low-quality images. In this paper, pulpitis detection using GLCM and Watershed is proposed. The dataset used contains x-ray images of reversible pulpitis, irreversible pulpitis, and normal tooth.**

***Index Terms*—Grey Level Co-occurrence Method (GLCM), K-Nearest-Neighbour(K-NN), Pulpitis, Watershed**

I. Introduction

Massive technical advances these days have helped the health world to solve some of the problems such as to diagnose internal body conditions in the health world is x-Ray that the dentist uses to see the entire layer of teeth is periapical radiography. Dentists use the periapical x-ray to see the condition of the teeth and underlying tissue inside of the mouth [1]. In order to detect and classify efficiently, machine learning-based techniques are proposed by comparing the performance of machine learning algorithms. The contributions of this paper are summarized as follows:

- Pre-processed the raw periapical radiographs images and characterized it so can help the doctors in the field of radiology
- Applied and compared machine learning algorithms using the Gray Level Co-occurrence Method (GLCM) and Watershed method, as well as feature extraction then classification using the K-Nearest Neighbor (K-NN)
- Improved the accuracy and computing time by compared differents dimension and K value.

II. Related Work

A. Watershed

By using the concept of a water drop, can be identified clearly the terms of the catchment basins in the algorithm established [2]. Watershed segmentation results in a more stable segmentation when the gradient of the image is applied and this gradient picture can be viewed as a topography with boundaries as watershed lines between regions or ridges [3].

B. Grey Level Co-Occurrence Matrix (GLCM)

Grey Level Co-Occurrence Matrix (GLCM) is defined as a tabulation of image pixel data where how often different combinations of grey values appear in the image [4]. Based on this [5] presented feature extraction of an image by using GLCM. In this paper, we will extract all of those features, energy, homogeneity, correlation, contrast, and entropy.

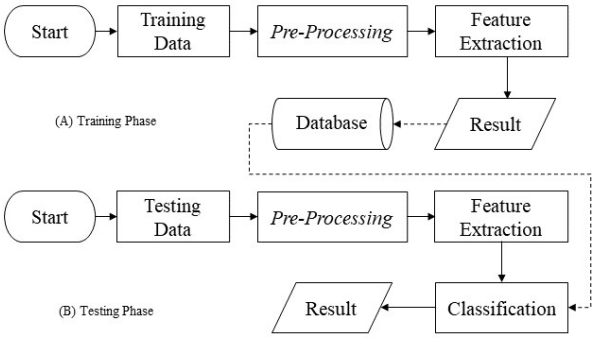

Fig. 1: Proposed systems

C. K-Nearest Neighbor

The K-Nearest Neighbor (KNN) method is one method for classifying data that uses neighboring classification as a prediction of new data [6]. This algorithm only stores feature vectors in the training process and classifies the examples of training data.

III. Proposed System

In this paper, the design of the pulpitis disease detection system has been proposed by using uses the Gray Level Co-occurrence Method (GLCM) and Watershed image segmentation methods, as well as the classification with K-NN. The first process in image acquisitionis to provide a periapical radiograph film, then the film is recorded using a scanner. The recording results then converted into the format. * Jpg. The next process is image identification, this process consists of two processes, changing the color format of the image and increasing the image contrast so that it's processed easily. The whole systems which showed in Fig.1, consists of two phases, the training phase, and the testing phase. The training process is taking the pixel value that will be used as a reference, where the value matched with the test image. While testing has been done by two methods, Watershed and GLCM. The test results will produce the extraction of each characteristic which later processed using K-Nearest Neighbor (K-NN). After all stages

TABLE I: Computation time of compared algorithms

Image Dimension	Watershed	GLCM
128x128	0.3746 s	0.1629 s
256x256	0.378 s	0.378 s
512x512	0.598 s	0.598 s

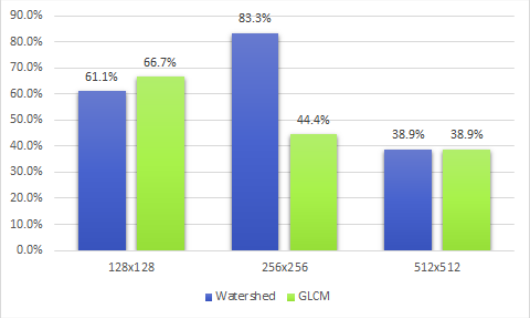

Fig. 2: Varying image dimension accuracy

TABLE II: Computation time of compared algorithms

K Value	Watershed	GLCM
1	0.378 s	0.1629 s
3	0.4813 s	0.2336 s
5	0.3702 s	0.21338 s

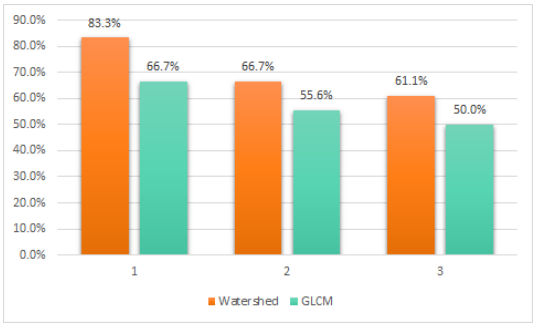

Fig. 3: Varying K Value accuracy

of the testing image and training image process, an evaluation of the system performance has been made.

IV. Performance Evaluation

The proposed system's performance was evaluated using simulations on MATLAB. The images used for training and testing were periapical radiographs stored in JPG format. The 21 training images were consisting of 7 reversible pulpitis diagnostic images, 7 irreversible pulpitis diagnostic images, 7 normal dental images, and 18 testing images consisting of 6 reversible pulpitis diagnostic images, 6 irreversible pulpitis diagnostic images, and 6 normal dental images.

1) *Image Dimension Comparison Testing*: This simulation shows the difference in accuracy and computation time produced by changing the dimensions of the image size 128 x 128, 256x 256, and 512 x 512 pixels with details K = 1 both in Watershed and GLCM. Table I represent the computing time that concludes that when the dimension set up a 128x128 by using GLCM is the fastest but will get the highest accuracy if the image dimension is 256x256 by using Watershed which showed in Fig.2.
2) *K Value*: System performance testing performed on both types of methods. Tests showed differences in accuracy and computational time obtained at k values 1, 3, and 5 as shown in Fig.3 and Table II. The first test uses Watershed feature extraction and 256 x 256 image dimensions, and in the second condition it uses GLCM feature extraction with offset details = 90 °, quantization level = 16 and image dimensions 128x128.

V. Conclusion

In this paper, the comparison of performance evaluation with the machine learning approach has been done by comparing with different dimensions and K Value in two algorithms, GLCM and Watershed, then after that classified with K-NN to classify images into 3 categories, pulpitis reversible, pulpitis irreversible and normal tooth. The results show that the best dimension to extract periapical radiograph is 256x256 which can achieve the highest accuracy compared to other dimensions. While in the classification phase, showed that the less K value can achieve higher accuracy. These methods not only can be applied in the Industrial field but also applied in a military field such as for intrusion detection systems in military base networks as the military needs to have more secure systems for security issues.

Acknowledgement

This work was supported by Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2018R1A6A1A03024003).

References

[1] M. F. Zuandi, B. Hidayat, and S. Sitam, "Granuloma image detection through periapical radiograph by using gabor wavelet method and support vector machine classification," in *2018 International Conference on Information and Communications Technology (ICOIACT)*, 2018, pp. 617–622.

[2] A. P. Chakkaravarthy and A. Chandrasekar, "An automatic segmentation of skin lesion from dermoscopy images using watershed segmentation," in *2018 International Conference on Recent Trends in Electrical, Control and Communication (RTECC)*, 2018, pp. 15–18.

[3] R. Embong and S. R. Anuar, "Structuring elements in the watershed algorithm for the segmentation of mammography images," in *TENCON 2018 - 2018 IEEE Region 10 Conference*, 2018, pp. 2144–2147.

[4] M. A. Ben Atitallah, R. Kachouri, M. Kammoun, and H. Mnif, "An efficient implementation of glcm algorithm in fpga," in *2018 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC)*, 2018, pp. 147–152.

[5] K. N. N. Hlaing and A. K. Gopalakrishnan, "Myanmar paper currency recognition using glcm and k-nn," in *2016 Second Asian Conference on Defence Technology (ACDT)*, 2016, pp. 67–72.

[6] S. Zhang, X. Li, M. Zong, X. Zhu, and R. Wang, "Efficient knn classification with different numbers of nearest neighbors," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 5, pp. 1774–1785, 2018.