Subjective-Objective Matching Evaluation Approach for Enhanced Dental Images

Siti Arpah Ahmad¹, Noor Elaiza Abd Khalid², Mohd Nasir Taib³, Haslina Taib⁴
¹²Faculty of Computer and Mathematical Sciences, University Teknology MARA (UiTM), Shah Alam, Malaysia
³Faculty of Electrical Engineering, University Teknology MARA (UiTM), Shah Alam, Malaysia
⁴School of Dental Sciences, Universiti Sains Malaysia, Health Campus, Kubang Kerian, Kelantan, Malaysia

ABSTRACT
Cases of misdiagnosis and variability evaluation among the dentists do happen. The complexity of anatomical structures and the low contrast of the original images are factors that contribute to the problems. Image enhancement is often used to enhance medical images. However, currently limited work has been done in enhancing the dental pathological features. Dentists come from different background in terms of experience, place of study, method of practices and emotional quotient. These are some of the factors that may cause differences in subjective evaluation among dentists. Therefore, this research focused on identifying objective measurements based on dentists’ subjective evaluation on abnormalities’ detection in jaw area. Objective measurement is based on contrast improvement index (CII) and subjective evaluation is derived from dentists’ questionnaire answering. This paper contributes to new knowledge in the initial phase of identifying dental disease characteristics by means of correlation between the subjective and objective evaluation.

Keywords:
Image enhancement
Objective evaluation
Subjective evaluation
Intra-oral dental images

1. INTRODUCTION
Diagnosing a dental image radiograph is an important first step in identifying disease and determining suitable course of treatment. However, diagnosis is difficult due to image quality problem and the anatomical complexity of the lesion. The regulation of low dosage used has resulted in radiograph images suffering from low contrast [1-3]. In addition, X-ray machine translate original 3-dimensional anatomy of the skull to 2-dimensional results in anatomy overlapping effect on the radiograph images [4]. These limitations have the tendency to introduce variability between dentist’s diagnosis [5-6]. Furthermore, there are variability issues in subjective evaluation among medical officers [7-8]. For example, there was a variability reported for cardiac ultrasound measurement in cats. The variability was evaluated based on different time point for inter-variability between two board of certified echocardiographers in veterinary cardiology [7]. Another example is the evaluation of inter-observer variation in reporting the nuclear features related to malignant neoplasm of the thyroid [8].

Initiative to overcome the above problems had been done by applying image processing methods such as image enhancement [9-11]. Image enhancement is the process of modifying images that enhance visual content to improve human or machine perception [12]. Enhancement methods help to increase contrast and detail information of the image, thus improve the image quality [13]. Good quality images made it possible to analyse structural and functional information of human physiology [14]. However, deciding on
which enhancement method that is suitable in enhancing dental images is still under research due to the complexity of the radiograph images, the uniqueness of dental diseases and abnormalities presentation.

Application of image processing techniques in dentistry involves assessing the influence of variables such as noise, sharpness and contrast with and without the image processing application in detecting carious lesion and enhancement of vertical root fractures [15-17]. Comparison between digital and conventional film based on intra-oral images assess the diagnostic elements of periapical disease lesion also has been done [18]. There are also an investigation of the stability of subjective evaluation by dentists in assessing the disagreement and borderline cases in periapical disease lesions [19]. Subjective evaluation is defined as assessment by human based on their judgments [20]. It usually includes human visual interpretation and used statistical methods such as ANOVA and kappa statistics [15], [21-22]. Objective evaluation on the other hand used methods such as CII, SNR, RMSE and MSE as the means of measuring the image enhancement algorithms performance [23-24]. This work used CII as objective evaluation.

The problem of low contrast can be overcome by applying image enhancement algorithms (IEAs). Previous works shows that IEAs have been applied to dental images. Example of IEAs algorithms applied are smoothing filters, sharpening filter, negative filter, histogram equalization (HE) and adaptive histogram equalization (AHE) [10], [22], [25]. However, current investigation is limited in sense of; comparing between filmed based and digitized dental images and comparing with and without enhancement algorithms between original and enhanced images only. Limited work is done to utilize image enhancement towards characterization of digital intra-oral dental radiographs images. Characterizations of dental images are important due to cases of early cancer detection come from routine dental check-up [26-27] and the relation of dental disease to cancer [28-30]. Dental images also are underutilized; there are a huge collection of dental medical images, but limited usage done towards diseases characterization.

This work initiates the evaluation of subjective-objective matching evaluation approach to enhance the pathological features of digital intra-oral dental radiographs. Abnormalities of interest were periapical radiolucency (PA), widen periodontal ligament space (widen PDLs) and loss of Lamina Dura (Loss of LD). The enhanced image and the questionnaire viewing approach during the subjective evaluation has potential to reduce the variability among the dentists’ subjective evaluation.

2. RESEARCH METHOD

The methodology consists of initial phase which consist of image collection, image enhancement application and objective measurement. The second phase is subjective evaluation and the last phase is the rule to eliminate the variability of the dentists’ score.

2.1. Initial Phase

This work collected 132 original intra-oral dental images. Prior to that, the ethic application was done. The approval was made by Universiti Teknologi MARA Ethical Committee (reference No: 600-RMI (5/1/6)). The raw X-ray radiographs film are taken from Faculty of Dentistry, UiTM Shah Alam using Planmeca Intra-oral machine. The digital forms are produced using the Scanmaker 1000XL Microtek scanner with resolution of 6400 X3200 dpi, with 48bit color. 60kV and 8mA X-ray machine setting are used by X-ray operator to acquire the intra-oral dental radiographs. These images are then enhanced by three IEAs algorithm namely; SAHE, SMAHE and SCLAHE; producing 528 enhanced images. The enhanced images than went through the objective evaluation. The objective evaluation used is Contrast Improvement Index (CII). CII is the popular index used by radiologist to check the visibility of lesions in radiographs [31]. It is calculated by $C_{\text{processes}}/C_{\text{Original}}$ where both are the contrast values for the region of interest in the enhanced images and original images respectively [31]. C is defined as in the following equation [32–33];

$$C = \frac{(f-b)}{(f+b)}$$  \hspace{1cm} (1)

where $f$ is the mean gray-level value of the image and $b$ is the mean gray-level value of the background. Each of the enhanced images will have the CII values.

2.2. Subjective Evaluation

The enhanced images were then subjectively evaluated by 14 dentists. Thus, the total subjective evaluation produced is 2032 evaluations. Table 1 shows the summary of the images used in the subjective evaluations.

| Indonesian J Elec Eng & Comp Sci, Vol. 12, No. 1, October 2018 : 239 – 245 |
The subjective evaluation is done using questionnaire. The questionnaire’s questions are divided into four categories. The first category accesses the image quality (ImgQ) with 5-point scale rating score; 1=the image is very poor; 2=the image is poor; 3=the image is acceptable; 4= the image is good and 5= the image is very good. The evaluation of the PA, widen PDLs and Loss of LD abnormalities uses a 3-point scale. The presence or absence of the periapical pathologies is scored using 3-point scale. For example, for the detection of PA abnormality the scale is as follows; score=1; PA detected; score =2; no PA abnormality detected but other abnormality detected. score =3; no PA detected, and no abnormality detected. Third and fourth categories also use the same 3-point scales similar to the second category, but the detection focuses on widen PDLs and Loss of LD respectively. The rating detail scale can be found in our previous work [9][18].

2.2.1 Questionnaire Sets
Four questionnaires have been used to do the evaluation. Two sets of questionnaire term as 10 images and 30 images have been used for supervised twin-view approach. One set of questionnaires, term as 56 images had been used for unsupervised random approach. The last set is termed as 36 images were used for supervised random approach. This section elaborates the details of the four sets of questionnaires.

2.2.2 Supervised Twin-view Approach
The first set term as 10 images and 30 images used supervised twin-view approach. Supervised means, there is an assistance besides the dentists during the evaluation process. The twin-view approach referred the arrangement between the images during the subjective evaluation process. The original and enhanced images were put side by side as in Table 2. The answering session is about 2 hours. The softcopy of the dental radiograph images is viewed using Windows Photo Gallery software. The Acer Aspire 4720 notebook computer has been used to view the images. The monitor resolution is 1280X 800 pixel and 32-bit color depth with screen size of 14.1 inch. The dentist is free to zoom in and out of the images as he/she wishes. After confirming about the evaluation, the dentist will tick the printed questionnaire for record purposes. There are six dentists involved in this investigation. Three dentists are from Shah Alam and three are from Kubang Kerian.

2.2.3 Unsupervised Random-view Approach
The 56 images set uses unsupervised random view approach. Unsupervised means the dentists answer the questionnaire at his/her own free time without the researcher’s assistance. Random approach is where the dentist does not know which image original and which image is is the enhanced images. Firstly, the images were identified and named as original listings. Then, using the random sequence generator from Random.org, the random numbers were generated. The images on the questionnaires were identified using the random sequence. The random and original sequence were matched back during the analysis of results. Each of the images was placed one image per page on the printed questionnaire. A duration of three-months was given to the dentists to answer the questionnaires. The softcopy of the images is viewed using a 14” computer monitor with a resolution of 1280 X 1024 in the ordinary lecturer room. Five dentists are involved in answering the questionnaires and all of them are from Kubang Kerian, Kelantan, Malaysia.

Subjective-Objective Matching Evaluation Approach for Enhance… (Siti Arpah Ahmad)
2.2.4 Supervised Random-view Approach

The 36 images used supervised random view approach. The viewing approach is the same as the 56 images questionnaire, but it is conducted in two hours with an assistance. On the computer monitor, one image is viewed at a time. The dentist is free to zoom in or out as she/he wishes. The computer Laptop monitor (14” Toshiba Satellite L745) with a screen resolution of 1280 X 720 pixel and 32-bit color depth is used in the study.

2.3 The \( \text{std}=0 \) Rule

After getting the subjective evaluation, new rule term as standard deviation equal zero \( (\text{std}=0) \), was applied to filter the subjective evaluation. Standard deviation \( (\text{std}) \) defines the way data is centered about the mean. Subjective evaluation from the questionnaire answered by the dentists is the main data. It comes from the 5-point scale for image quality grading score and the 3-point scale for abnormality detection score. The values are calculated to determine the agreement between the dentists’ evaluation \[38\]. \( \text{std}=0 \) mean that all the dentists have the same opinion that the abnormalities exist in the image. The values are calculated using Microsoft Excel 2007 formula of \( =\text{stdev}(\text{number1},[\text{number2}]...) \). The rule was used to eliminate images that have score’s variability. Only images that have the same exact score were selected.

3. RESULTS AND ANALYSIS

Results are presented in term of correlation between subjective and objective evaluation. Subjective evaluation is illustrated based on image quality grade. Only score=3 and score=4 is taken as final good image quality. The objective evaluation is presented in term of CII values.

3.1. Image Quality

Table 3 displays the subjective and objective matching evaluation for image quality assessment that are able to meet the \( \text{std}=0 \) criteria. There were only 59 images. This means that the variability in the subjective evaluation is huge for image quality. From the 2032 subjective evaluation only 59 images (2.9%) able to meet \( \text{std}=0 \) rule. Looking at the questionnaire sets, majority of images come from 36 images set (25 images) with random view viewing approach and AHE. However, the image quality score=1. This means that the image quality is very poor. Focusing on the good image quality score (score=4); the result come from 30 images questionnaire set, with SCLAHE under the twin-view viewing approach, with seven images. Average CII values for good image quality is 1.49.

3.2. Periapical Radiolucency (PA) abnormality

Table 4 illustrates the subjective and objective matching evaluation for PA abnormality. detection that was able to meet the \( \text{std}=0 \) criteria. 137 images were able to meet the criteria. From the 2032 subjective evaluations, 6.7% was able to meet \( \text{std}=0 \) rule. Looking at the questionnaire sets, majority of images come from 36 images set (26 images) with supervised random viewing approach and AHE. However, the image quality score=3; which means, “No periapical radiolucency detected, and no abnormality detected”. IEAs was able to get the highest score=1 (Periapical radiolucency detected) is SCLAHE which came from 30 images questionnaire set, with twin-view viewing approach. The number of image is 15, with average CII values is 2.03.
Table 4. Subjective-objective matching evaluation for Periapical Radiolucency (PA)

| No. of image (std=0) | Questionnaire Set | IEAs   | Viewing Approach | Score | No. of images | Average CII |
|----------------------|-------------------|--------|------------------|-------|---------------|-------------|
| 1                    | 10 images         | SAHE   | Supervised Twin-view | Score=3 | 22.40         |
| 6                    | 10 images         | SMAHE  | Supervised Twin-view | Score=1 | 3.32          |
| 6                    | 10 images         | SCLAHE | Supervised Twin-view | Score=1 | 2.82          |
| 9                    | 30 images         | SAHE   | Supervised Twin-view | Score=3; 2 | 12.04         |
| 14                   | 30 images         | SMAHE  | Supervised Twin-view | Score=1 | 5.80          |
| **15**               | **30 images**     | SCLAHE | Supervised Twin-view | Score=1 | **2.03**      |
| 20                   | 56 images         | AHE    | Unsupervised Random | Score=1;4 | 5.08          |
| 14                   | 56 images         | CLAHE  | Unsupervised Random | Score=1;9;5 | 1.39          |
| 9                    | 56 images         | SCLAHE | Unsupervised Random | Score=1;2 | 1.55          |
| 26                   | 36 images         | AHE    | Unsupervised Random | Score=2;2 | 1.63          |
| 9                    | 36 images         | CLAHE  | Unsupervised Random | Score=1;7 | 1.51          |
| 8                    | 36 images         | SCLAHE | Supervised Random  | Score=1;7 | 1.51          |
|                       |                   |        |                  |       |               |             |
| **Total image=137**   |                   |        |                  |       |               |             |

3.3. Widen Periodontal Ligament Space (widen PDLs) abnormality
The result for Widen PDLs abnormality shows that 100 images (4.9%) were able to meet the \( std=0 \) criteria as in Table 5. Majority no. of the images is similar to the previous sections, with 36 images questionnaire sets, with 26 images for AHE and the score is 3. SCLAHE was able to get the highest no. of images (15) with score=1, which means “Widen periodontal ligament space detected “. It came from SCLAHE, with supervised twin-view approach. The average CII values is for this is 3.11.

Table 5. Subjective-objective matching evaluation for Widen PDLs abnormality

| No. of image (std=0) | Questionnaire set | IEAs   | Viewing Approach | Score | No. Of Images | Average CII |
|----------------------|-------------------|--------|------------------|-------|---------------|-------------|
| 1                    | 10 images         | SMAHE  | Supervised Twin-view | Score=1 | 2.16          |
| 5                    | 10 images         | SCLAHE | Supervised Twin-view | Score=1 | 3.39          |
| 7                    | 30 images         | SAHE   | Supervised Twin-view | Score=1 | 9.09          |
| 7                    | 30 images         | SMAHE  | Supervised Twin-view | Score=1;5 | 5.41          |
| **15**               | **30 images**     | SCLAHE | Supervised Twin-view | Score=1 | **3.11**      |
| 12                   | 56 images         | AHE    | Unsupervised Random | Score=3 | 6.49          |
| 10                   | 56 images         | CLAHE  | Unsupervised Random | Score=3;9 | 1.40          |
| 7                    | 56 images         | SCLAHE | Unsupervised Random | Score=3 | 1.31          |
| 26                   | 36 images         | AHE    | Supervised Random | Score=3 | 6.47          |
| 6                    | 36 images         | CLAHE  | Supervised Random | Score=1;5 | 1.49          |
| 4                    | 36 images         | SCLAHE | Supervised Random | Score=1;3 | 1.65          |
|                       |                   |        |                  |       |               |             |
| **Total image=100**   |                   |        |                  |       |               |             |

3.4. Loss of Lamina Dura abnormality
The result of Loss of lamina dura in Table 6 displays that, 84 images (4.13%) were able to meet the \( std=0 \) criteria. Majority of images came from 36 images questionnaire sets, with AHE for the IEAs and under random supervised viewing approach. Score=1, came from 30 images questionnaire with SAHE (2 images), Schedule Objective Matching Evaluation Approach for Enhance... (Siti Arpah Ahmad)
SMAHE (9 images) and SCLAHE (10 images) under Twin-view supervised viewing approach. There was one score=1 for SCLAHE under 36 images questionnaire with random supervised viewing approach. However, majority for score=1, still came under SCLAHE, with 10 images. The average CII score is 2.23.

| No. of image (std=0) | Questionnaire set | IEAs     | Viewing Approach       | Score | Average CII |
|----------------------|-------------------|----------|------------------------|-------|-------------|
| 2                    | 30 images         | SAHE     | Twin-view Supervised   | Score=1 | 3.93        |
| 9                    | 30 images         | SMAHE    | Twin-view Supervised   | Score=1 | 4.12        |
| 10                   | 30 images         | SCLAHE   | Twin-view Supervised   | Score=1 | 2.23        |
| 18                   | 56 images         | AHE      | Random                 | Score=3 | 5.33        |
| 10                   | 56 images         | CLAHE    | Random Un-Supervised   | Score=3 | 1.33        |
| 8                    | 56 images         | SCLAHE   | Random Un-Supervised   | Score=3 | 1.33        |
| 26                   | 36 images         | AHE      | Random Supervised      | Score=3 | 5.49        |
| 0                    | 36 images         | CLAHE    | Random Supervised      | none   | -           |
| 1                    | 36 images         | SCLAHE   | Random Supervised      | Score=1 | 1.27        |

Total image=84

4. CONCLUSION

This work exhibits the investigation on suitable evaluation approach for medical images, with case study on digital dental images. The results show that enhanced image by SCLAHE and supervised twin-view questionnaire viewing approach able to get the lowest variability in subjective evaluation. The advancement of medical images today, has resultant massive of images. However, the utilization of these images towards better understanding towards new knowledge is limited. Reason being is the gap in the evaluation approaches. The importance and sensitivity of these images had made current approach of subjective evaluation compulsory however, since subjective evaluation is expensive (medical officers participate in answering the questionnaire are paid certain amount of monetary honorarium) and time consuming, objective evaluation is initiated. Thus, introducing the subjective-objective matching evaluation is an initial step towards reliable sole objective evaluation in the future.

ACKNOWLEDGEMENT

The authors acknowledge with gratitude the great support from the management of Faculty of Computer and Mathematical Sciences, UiTM Shah Alam, Faculty of Dentistry, UiTM Shah Alam, Faculty of Electrical Engineering, UiTM Shah Alam and Universiti Sains Malaysia, Kubang Kerian. This research previously received E-science fund under the Ministry of Science Technology and Innovation (MOSTI) Malaysia.

REFERENCES
[1] Katsoni E et al. Exploring the useful exposure range limits of three intraoral image receptors for various tube potential, tube current and exposure time setting. Heal. J. 2011; 3: 292–299.
[2] Okano T, Sur J. Radiation dose and protection in dentistry. Japanese Dental Science Review. 2010; 46:112-121.
[3] Praveen BN, Shubhasini AR, Bhanushree R, Sumsum PS, Sushma CN. Radiation in Dental Practice: Awareness, Protection and Recommendations. The Journal of Contemporary Dental Practice. 2013;14(1):143-148.
[4] Krupinski E. Current perspectives in medical image perception. Atten. Percept. Psychophys. 2010; 72(5): 1205–1217.
[5] Baseri H. Introducing a Dental Caries Marking Software and Evaluate Radiologists’ Disagreement in Caries Detection Using this Software. Journal of Dental Biomaterials, 2015; 2(1): 10-17.
[6] Ahmad S.A, Taib M.N, Khalid N.E, Taib H. Sharpening based image enhancement algorithms in reducing the disagreement of medical images subjective evaluation. IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS), 2016.
Subjective-Objective Matching Evaluation Approach for Enhancement... (Siti Arpah Ahmad)