Dental Impression Tray Selection From Maxillary Arch Images Using Multi-Feature Fusion and Ensemble Classifier

MUHAMMAD ASIF HASAN[1], **NORLI ANIDA ABDULLAH**[2], **MOHAMMAD MUSTANEER RAHMAN**[3], **MOHD YAMANI IDNA BIN IDRIS**[1], (Member, IEEE), AND **OMAR F. TAWFIQ**[4]

[1]Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia

[2]Mathematics Division, Center for Foundation Studies in Science, University of Malaya, Kuala Lumpur 50603, Malaysia

[3]School of Information and Communication Technology, Syndicate of Technology, Environments and Design, University of Tasmania, Hobart, TAS 7000, Australia

[4]Faculty of Dentistry, SEGi University, Petaling Jaya 47810, Malaysia

Corresponding author: Norli Anida Abdullah (norlie@um.edu.my)

ABSTRACT Dental impression tray is frequently used in dentistry to record the patient's oral structure for clinical oral diagnosis and treatment planning. Manual procedure of taking impressions is costly, time-consuming, and additionally, no research has been done on selecting dental impression tray from dental arch images using computer vision in real-life scenarios. In this spirit, an intelligent model is proposed based on computer vision and machine learning to select appropriate dental impression trays from maxillary arch images. A dataset of 52 patients' maxillary arch images have been acquired and various sets of features such as colors, textures, and shapes of the images were extracted to better characterize the maxillary arch images. Considering the importance of the features in describing the maxillary arch object and to improve the classification performance, a method based on multi-feature fusion with ensemble classifier is proposed. Besides, the performance of a deep learning based multilayer perceptron neural network is also investigated. The proposed multi-feature fusion with ensemble classifier attained 92.31% precision, 91.75% recall, 91.75% accuracy, respectively, on the dataset, which clearly establishes the feasibility of the proposed model. An illustration of a real-life application of the proposed model is also provided.

INDEX TERMS Dental impression tray, dental arch image, automation in dentistry, computer vision, multi-feature fusion, ensemble classifier.

I. INTRODUCTION

Dental casts (FIGURE 1-a) are accurate, three-dimensional replicas of a patient's teeth, which are made by pouring dental plaster or acrylic into the impressions of the teeth. Dental casts are widely used in dentistry to provide fundamental diagnostic information where most of the advanced diagnosis and treatment planning in prosthodontics, indirect restorations, cosmetic and orthodontic treatments are relying on impression. Dental casts are mainly used especially for the further casting of dental prostheses such as bridges, dentures and partial dentures. According to Gupta *et al.* [1], it is necessary to select an appropriate impression tray (FIGURE 1-b) for making an accurate, definitive cast. Improper use of

impression tray size may fail in recording the tissue correctly. This kind of error may be corrected by adjusting the tray or choosing a better fitting tray and repeating the impression procedure. This procedure involves manual interactions by a human operator, which makes it burdensome, tedious, user-dependent, and sometimes creates a waste of impression materials too, which a professional, well-run dental business should avoid. Furthermore, automation makes it possible to eliminate user dependency and permits evidence-based analysis [2]. Thus, automatic identification of the dental impression tray has important scientific significance and can be a potential application to enable workflow automation in dentistry [3].

Motivated by this, to facilitate the workflow with minimal cost and time, a computer-vision based dental impression tray classification model is proposed in this study.

The associate editor coordinating the review of this manuscript and approving it for publication was Huiling Chen.

FIGURE 1. Example of a patient's cast (a); an impression tray (b).

The vision-based impression tray classification model is proposed to select an appropriate impression tray aimed at working in a real-life scenario. For instance, an intraoral camera can be used to capture the occlusal view of the maxillary arch image and based on the image, the model will select the appropriate tray. In this proposed study, computer-vision based features such as statistical color features, texture feature, morphological shape features, local binary pattern, color histograms are analyzed to design the vision-based impression tray classification model. These features were chosen because of their potential in describing the object and proven wide application in medical image analysis, e.g., classification of skin lesion types [4]. However, in real-life scenarios, the maxillary arch images can be affected by different light sources, rotation and with the variation of the different types of mouth shapes due to the uncontrolled environment in dentistry. On this account, for this proposed study, the detection and extraction of suitable features are important. It is possible that extracting only one single set of features may not be enough to identify the maxillary arches properly by the classifier. The performance of the classification algorithms greatly relied on suitable features. Inability to detect suitable features may even decrease the classification performance as well. In fact, distinct discriminating capabilities are observed in different features. In this case, feature fusion approaches have the high dimensionality and uneven distribution which as a result recently drew great attention in the medical image field because they can solve complex linear and non-linear recognition problems effectively [5].

Therefore, this study focuses on identifying appropriate set of features to improve the classification performance in a limited dataset using fusion method. A dataset of maxillary arches (52 images) was utilized in this study due to the limited number of participants. Collecting and labeling a large amount of ground truth medical data is a challenge, especially in the scholarly area. Large-scale datasets or big data are collected mainly through grand challenge, taking screening initiatives by countries, crowdsource or application programming interface [6]. For this study, it is difficult to fulfil the large-scale datasets requirement, such as the ImageNet dataset which consists of 3.2 million cleanly labeled images [7]. Literature shows that several studies utilized limited dataset in their study for experimental purposes, e.g., Yilmaz *et al.* [8] used 50 patients' cone beam computed tomography (CBCT) between the years 2013 and 2016 for designing computer-aided diagnosis of the periapical cyst. On this ground, through this study, the authors mainly sampled this maxillary arch dataset as a proof-of-

concept. To resolve the overfitting issues which may arise because of this limited sample dataset, artificial image data augmentation has been performed on the dataset such as adding different illumination, scaling, and rotation. After identifying the features, five machine learning classification algorithms were evaluated separately on the set of features. Finally, to suit the problem stated previously and to better categorize the features of maxillary arches properly according to the tray, an attempt to investigate the fusion of multiple sets of features with ensemble classifier was done. In addition, a deep learning based multilayer perceptron (MLP) neural network (NN) was also analyzed as well. The performance of the proposed fusion-features with ensemble classifier method is evaluated in terms of precision, accuracy, recall, F1 score, Jaccard similarity score (JAC), Matthews correlation coefficient (MCC), Zero one loss function, and computational time.

This study is based on the hypothesis that computer vision and machine learning based classification methods are capable of selecting appropriate dental impression tray from maxillary arch images with sufficient accuracy and can be of potential use in dentistry to automate the digital workflow. The contribution of this study is given below:

   i Proposing a first study to select a dental impression tray from the maxillary arch image in a real-life scenario.

   ii A novel method based on multi-feature fusion with ensemble classifier is proposed to elevate the classification performance from a limited dataset.

   iii Illustrating the utilization of the proposed model for dental practitioners.

The rest of the article is organized as follows: Section II exhibits the associated works based on image recognition techniques. Section III depicts the proposed method in detail. Section IV reports the experimental results, followed by the discussion in Section V. The conclusions and some lines of potential future works are finally made in Section VI.

II. RELATED WORKS

Attaining the capability of human recognition remains a challenge in computer vision [9]. However, advancement in image recognition technology is progressing because of the continuous development in this field. This unlocks the door of opportunity for real-world implementation, for instance, implementation of image recognition in the medical application field. In computer vision, image recognition is used to identify the instances of the object in digital images. To identify the object, defining appropriate features is necessary, and then using machine learning classifier techniques to classify the objects. The features can be defined as the measurable characteristics of the object, and the classifier categorizes the features into classes based on their similarity. Properly described image features are capable of representing the object. According to the literature, feature extraction methods are broadly classified into two categories: (i) hand-crafted feature extraction; (ii) deep-learning based feature extraction. Each method is described in this section. Certainly, identifying the target image is not only restricted to medical image

analysis but also related to other fields of image analysis as well. Hence, further related image recognition studies that fit the scope of this study are included in the following section.

The autonomous selection of dental impression trays based on the dental arch images is rarely implemented in the literature. Yergin *et al.* [10] used distance-based image alignment for assessing dental trays. According to their description, this work can only be performed onto a small program or chart which is not broad enough to describe the feasibility of their approach. Also, they did not provide any evaluation metrics to show how efficient their work was in a real-life scenario. On the other hand, Rijal *et al.* [11] proposed a unique technique that represents the dental arch shape in relation to stock tray design. Another investigation was also done by Rijal *et al.* [12] on the homogeneity of 3 groups of arch shape (representation, clustering, and evaluation changes) to suggest 3 groups of multivariate (MV) normal distribution of these groups that may be used to probe the arch shape variation issue. However, these two studies are limited to statistical analysis using stone cast images and did not use any real dental arch images. An essential point is that these studies did not propose any system or workflow to deploy their work in real-life application.

A recent study by Litjens *et al.* [6] describes that promising results have been achieved in medical image analysis using computer vision.For example, at the 12th IEEE International Symposium on Biomedical Imaging (ISBI 2015), computer automated dental caries detection and X-ray image analysis of cephalometric challenge was defined as one of the grand challenges in the dental community. Wang *et al.* [13] presented datasets, methods and results of the challenge and laid down the principles for future uses of this benchmark. Studies assured that texture-based features are considered successful in medical image analysis [14]. For instance, while diagnosing skin cancers, the texture of melanoma plays an important role because it contains essential information about differential structures [15]. Most common texture-based methods are Haralick features derived from grey-level co-occurrence matrix (GLCM) and local binary pattern (LBP) [16]. However, the convenience and effectiveness of the GLCM method in distinguishing objects relied on texture information [16]; while on large-scale structures, LBP is unable to capture dominant features [17]. In several scenarios, shape-based extracted features also showed potentiality in medical image analysis [4]. Generally, shape-based method basically employs these two techniques: region-based and contour-based techniques. However, sometimes information of the object boundary may be unavailable, which makes this method difficult to exploit. Also, in greyscale images, some object shapes may be recognized the same (e.g., lemon and ball). In this situation, shape with color features together can distinguish the objects more specifically [18]. Since the human eye can distinctly recognize the color of an object, therefore in computer vision, proper color analysis is necessary to visualize the object for extracting fundamental information and to reveal the spatial information.

Hence, in medical image analysis color features are also analyzed; for example, Ali *et al.* [19] used color features for detecting gastric abnormalities. For the visual perception of the object, excessive color information is necessary, and this drawback means sometimes using only the color features makes it difficult to discriminate the objects. An interesting point to note is that in certain scenarios, fusion of shape with color features [18], fusion of shape with texture features [20], fusion of color with texture [21] and fusion of color, texture alongside shape features proved to be very efficient in elevating the classification accuracy [22]. For example, Tyas *et al.* [23] used morphological shape, texture and color features with multilayer perceptron neural network (MLP NN) for erythrocyte classification in thalassemia cases; Liu *et al.* [24] utilized color and shape features for detection of apple fruits which even outperformed the region-based convolutional neural network (RCNN). Therefore, it is necessary to find which set of features to choose and combine to increase the classification performance.

In recent years, some scholars have turned to deep learning (DL) algorithms because of their high accuracy and robustness. Deep learning algorithms for instance, convolutional neural network (CNN) can extract features from sample images automatically and do not depend on hand crafted feature extraction methods. In medical sectors, various deep learning algorithms show potential outcomes; Avuçlu *et al.* [25] used image processing techniques with MLP NN to determine age and gender by examining teeth and bone structures. CNN has been used in medical image analysis for automatic classification of peripheral blood cells [26], region extraction and classification of skin cancer [27], detecting of medical text semantic similarity [28] and so on. However, training of these algorithms requires massive, labeled image data. Collecting and labeling these massive images require a great amount of time and labor cost. To train and run these massive data also requires better hardware as well. These issues hinder deploying of these deep learning algorithms. For this reason, some scholars still relied on traditional methods made up of hand-crafted image features and machine learning classifiers.

Classification algorithms play a major role in computer vision, pattern recognition and data mining [29], [30]. Various classifiers such as support vector machine (SVM) [31], K-nearest neighbor (K-NN) [32], random forest (RF) [33], and so on are used in image recognition problems. Each classifier has its own merits and demerits. For example, SVM could simplify calculations and avoid the problems of overfitting on nonlinear and high-dimensional features [34], whereas K-NN is based on the majority vote that classifies through calculating the distance between data. The distance of the data from different classes may be similar, which may increase misclassification [35]. In contrast with the single algorithms, ensemble learning methods attain better predictive performance by utilizing multiple learning algorithms than what could be obtained from any of the individual learning algorithms alone (classification or regression) [36].
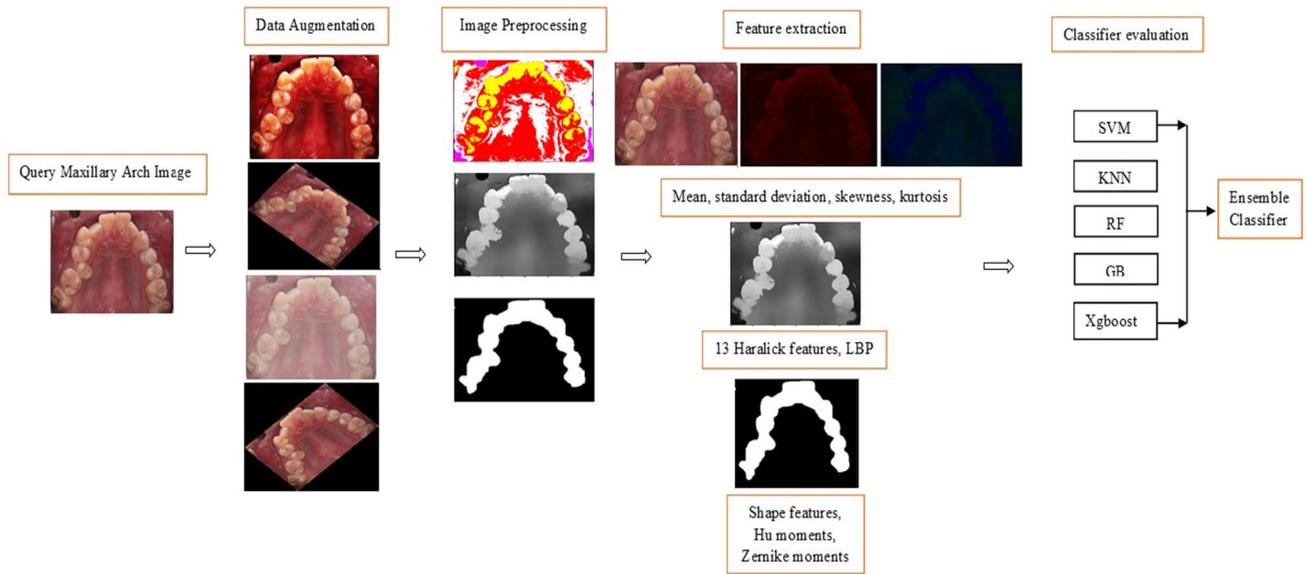
FIGURE 2. Block representation of the proposed method.

This approach allows better predictive performance compared to a single algorithm to a certain extent. Because of their effectiveness, ensemble methods are favored in various medical image analysis too [37].

As presented in the above discussion, to the best extent of the authors' knowledge, none of the studies focused on selecting an appropriate dental impression tray in a real-life scenario using computer vision. Moreover, selection of impression tray from maxillary arches is hindered by the changeability of the objects, background interference, and shortage of prior knowledge. It is observable that color, texture, and shape features show promising results in different fields of computer vision applications along with medical image analysis, as previously discussed. Thus, in this study, these features are chosen to describe the maxillary arch images. The limited number of features cannot express these maxillary arch images precisely, and the efficiency and stability of the classification algorithms should be preferentially considered as well. Therefore, an investigation of fusing multiple sets of features was done to elevate the classification performance for selecting appropriate impression tray. Also, to decrease the variance, bias error and to improve classification rate, an ensemble classifier is leveraged combining the best classifiers using the soft voting approach. The detailed process is described in Section III.

III. PROPOSED METHODOLOGY

The detailed research methodology to design the proposed vision-based dental impression tray classification model is described in this section. Generally, the proposed study comprises six main phases, namely, image data acquisition, image data augmentation, image pre-processing, feature extraction, classification algorithms evaluation, and finally designing

the ensemble classifier. In FIGURE 2 a block representation is presented of the above detail. For data acquisition, 52 patients' maxillary arches image data have been analyzed after obtaining ethical approval from SEGi university. To improve the classification performance and avoid overfitting issue image data augmentation was performed. In the image pre-processing phase, some essential pre-processing tasks were done to get the region of interest (ROI) from the maxillary arch images. Several discriminative and informative features were extracted from the arches through various feature extraction methods. In the fifth phase, five classification algorithms were analyzed, and finally, an ensemble classifier was developed from the classification models. In the subsequent sub-sections, all the phases are described in detail.

A. IMAGE DATA ACQUISITION

Ethical approval for patients' maxillary dental arch impression acquisition was given by the SEGi University Oral Health Center Medical Ethics Committee (SEGi EC/SR/FoD/2019-20/12). Participants received information related to the study and were asked to sign a consent form to take part in this study.

Male and female participants were randomly selected among walk-in patients of the SEGi oral health center. Participants were selected from different age groups ranging from 16-40. Some of the participants indicated or had a history of orthodontic treatment. According to dentist and manufacturer advice, four different sizes of impression trays (Daniel Kurten, Germany) were chosen: extra-large (XL), large (L), medium (M), and small (S). The selected trays were compared in dimensions with several brands of stock impression trays; but, they had no or only slight difference
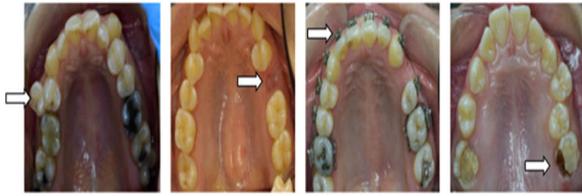
FIGURE 3. Selected occlusal images for various maxillary arches.

Tray	Original Images	Augmented images
Tray 1 (XL)	4	40
Tray 2 (L)	8	80
Tray 3 (M)	14	140
Tray 4 (S)	10	100
Total	36	360

in dimensions. Therefore, the authors decided to use Kurten with the four sizes of trays only for the experiment. Each of the four impression trays was matched to fit the maxillary image of each of the participants. No modification in dimensions or scale was made on the images. The fitting was checked by superimposing the raw images of both the tray and arch obtained using standardized image setting with calibration reference. In this process, 52 maxillary occlusal images were collected, using intraoral dental photography mirrors (OEM rhodium coated glass photography mirror, China) and DSLR camera kit with dental photography equipment (Nikon D5600, 85mm micro lens and R1C1 speedlight system from Nikon Japan). The selected cases included with crooked teeth, buccal defect, braces, and dental caries. Examples of these images are shown in FIGURE 3. The selected 52 patients' images are split into training set (70%) and test set (30%) by random sampling method. In this way, 36 images were used for training and 16 images for testing. The images have various resolution, and the format is JPG.

B. IMAGE AUGMENTATION

The original training dataset consists of 36 images which may not be enough to create a proper model. There may be a possibility of overfitting issue (i.e., the model will learn too much from the limited training dataset). Also, training deep learning algorithms, for example, training CNN from scratch requires a large amount of labeled data together with artificially added data to avoid overfitting issue [38]. One possible solution is to use data augmentation technique to enrich the training data, hence improving the model generalization ability and robustness. Various image augmentation techniques were used, such as different angle rotation, sigmoid correction, adjusting gamma, and blurring together with their random combination. In a natural environment, depending on the light source, different illumination may be created in the images. To make the model more robust to this scenario, different illumination was added, and rotated images were included considering the rotation of the images taken by the dentists. In Table 1, the number of augmented images is described.

C. IMAGE PRE-PROCESSING

This stage includes getting the region of interest (ROI) from the maxillary arch image. A thorough investigation has been performed to get the ROI. Dhivyaa *et al.* [39] used the K-means clustering algorithm for coarse segmentation, which

resulted in better accuracy. Hence, K-means clustering is utilized for segmentation purpose. The other phases are detailed in Table 2. Anyway, morphological shape features from contour show a better visual representation of an image [40]. The shape features such as area, perimeter, aspect ratio and so on are extracted from the binary images. However, the texture properties are extracted from the greyscale images. Finally, color properties are extracted from the color images. In the case of color and texture feature extraction, both the direct input image and the pre-processed ROI images are used.

D. FEATURE EXTRACTION
In this stage, several sets of discriminative and informative features are extracted. In the following subsection, the details of the feature extraction phases are described.

1) COLOR FEATURES
The surface of the object in an image can be described by the global feature. Color feature is one kind of global feature which is used by researchers for this purpose [41]. The statistical calculations of mean (μ) and standard deviation (σ), kurtosis (γ), and skewness (θ) are extracted from each plane of these three different color spaces such as RGB; which is an additive color model, HSV; which is an alternate representation of RGB to align the way human vision perceives, and L*a*b; which describes all the colors visible to the human eye.

The statistical formulas are defined as follows:

$$\mu = 1/(LM) \sum_{x=1}^{L} \sum_{y=1}^{M} P_{xy} \tag{1}$$

$$\sigma = \left(1 \middle/ (LM) \sum_{x=1}^{L} \sum_{y=1}^{M} \left(P_{xy} - \mu \right)^2 \right)^{1/2} \tag{2}$$

$$\theta = \left(\sum_{x=1}^{L} \sum_{y=1}^{M} \left(P_{xy} - \mu \right)^3 \right) \middle/ \left(LM\sigma^3 \right) \tag{3}$$

$$\gamma = \left(\left(\sum_{x=1}^{L} \sum_{y=1}^{M} \left(P_{xy} - \mu \right)^4 \right) \middle/ \left(LM\sigma^4 \right) \right) - 3 \tag{4}$$

here, P_{xy} is the value of color on row y and column x and, L and M are the dimension of the image.

In addition, color histogram feature from the three color spaces was also taken into consideration [42]. Color histogram is the distribution of colors in an image. The color histogram is a statistical calculation representing an estimate of an underlying continuous distribution of color values.

TABLE 2. Proposed Algorithm for Getting ROI.

Input color image (rgb) from file	
# *K-means clustering for image segmentation* Compute the number of clusters for K_i [i=1 to n] for the image matrix until convergence criteria is met. Calculate the within sum of squares(wss) Plot the curve wss. The bend in the plot is the appropriate number of clusters. # *for better visualization it is better to convert into L*a*b color space*	
Enhance image using mean shift filtering. Initialize m=1 and $y_m = x_j$ Compute $y_{m+1} = 1/n_m \sum_{x_i \in s_1(y_m)} x_i$, m← $m + 1$ till convergence Assign $x_j = x_j^s, y_{conv}^r$	
Convert the rgb image to greyscale (gs) gs = r*0.299 + g*0.587 + b*0.114	
Obtain the threshold image $< I_T$ # I_T is the best thresholding obtained from Otsu's method Dilate and erode the binary image by morphological close operation Select the ROI from the binary image	
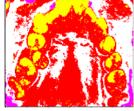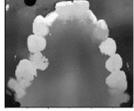
2) TEXTURE FEATURES

A statistical approach of investigating and extracting textual features from the image is grey-level co-occurrence matrix (GLCM). This method is based on the relation between two neighboring pixels of the grey tones within an image. In the 1970s, Haralick *et al.* [43] proposed GLCM for texture descriptor, and it has been popular thanks to its performance to date. Haralick identified 14 grain texture features; however, in this current study 13 Haralick features are used. For simplicity, only the five major Haralick features are described below. To learn more on the other set of features, one can refer to [44].

$$a)\ Energy \sum_{x,y} p(x, y)^2 \tag{5}$$

$$b)\ Correlation \sum_{x,y} ((x - \mu_x)(y - \mu_y) p(x, y) / \sigma_x \sigma_y) \tag{6}$$

$$c)\ Constrast \sum_{x,y} |\text{x-y}|^2 p(x, y) \tag{7}$$

$$d)\ Entropy \sum_{x,y} p(x, y) log_2 p(x, y) \tag{8}$$

$$e)\ Homogeneity \sum_{x,y} p(x, y) / 1 + |x - y| \tag{9}$$

here, *p(x,y)* is the probability density function, and $\mu_x, \mu_y, \sigma_x, \sigma_y$ are the values of the average and standard deviation of the rows and columns of *p(x,y)*.

In addition, local binary pattern (LBP) features are extracted, which is another excellent technique to describe texture information of an image by estimating their neighborhood pixels [45]. Let $I_g(x, y)$ be a greyscale image with a dimension of L × M and (x, y) denotes the position of pixels in the image. Given a central pixel P_c, and the correspondence neighboring pixel P_n; then, the LBP is calculated as follows [17]:

$$\Psi (IF)_{n,r} = \sum_{n=0}^{n-1} i (P_n - P_c) 2^n \tag{10}$$

here, $i(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$, and, n and r denote the neighboring pixel and radius of the neighborhood, respectively.

3) SHAPE FEATURES

Contour and morphological shape features give a meaningful insight of the object's shape [40]. Numerous studies used contour and morphological shape features in medical image analysis. Contour can be defined as a curve joining all the points along the boundary of a certain shape. From the

contour, the image moment was identified which is the certain weighted average of the image pixel intensities. Let I be a binary image and the pixel intensity at location (x,y) is given by I(x,y):

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \tag{11}$$

here, M is the moment which summarizes the shape of that binary image. From the equation the following was calculated:

$$Area\ M_{00} = \sum_x \sum_y I(x, y) \tag{12}$$

$$Centroid\ \bar{x}\bar{y} = \frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}} \tag{13}$$

From the minimum circumscribed rectangular of the sample object, Width (W) and Height (H) is taken. From these, the following was calculated:

$$Aspect\ ratio\ \frac{W}{H} \tag{14}$$

$$Rectangular\ area\ W * H \tag{15}$$

$$Equivalent\ diameter\ \sqrt{\frac{4 * M_{00}}{\pi}} \tag{16}$$

In addition, the Hu invariant moment [46] is used too. In the present study, M1-M6 (six orthogonal absolute invariants) and M7 (one skew orthogonal invariant) are used to describe the shape features of the samples. The first 6 moments have been proved to be invariant to translation, scale, and rotation and the 7th moment's sign changes for image reflection. In addition to Hu moment, Zernike moments [47] are also used to describe the object shape. The magnitudes of Zernike moments are independent of the object rotation, which is a nice property when working with shape descriptors.

E. DATA NORMALIZATION

Normalization operation is used to increase the classifier performance. Frequent normalization methods are used after feature extraction. However, in this study *z*-score normalization process is used. The standard score of a sample x_i is calculated as:

$$z = \frac{x_i - \mu_i}{\sigma_i} \tag{17}$$

here, μ_i is the mean of the samples or zero and σ_i is the standard deviation of the samples or one.

F. PROPOSED ENSEMBLE CLASSIFIER

Ensemble learning is a popular paradigm employed to leverage the strength of individual algorithms and mitigate their weakness [48], [49]. To solve a given problem, an ensemble technique combines a set of given single techniques (classification or regression) using an aggregation rule [50] as illustrated in FIGURE. 4. The main objective is to achieve a high level of accuracy from this model that at least exceeds the performance accuracy of a single model [51]. As majority voting is effective among other aggregation techniques [52] thus, in this study the voting classifier scheme (i.e., soft voting) is
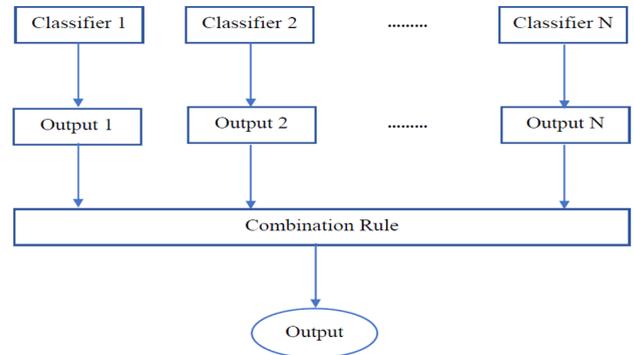

FIGURE 4. Ensemble classifier architecture.

used to create an ensemble classifier. Five machine learning classifiers have been chosen for this study, namely, support vector machine (SVM) [31], K-nearest neighbor (KNN) [53], random forest (RF) [33], gradient boosting (GB) [54] and eXtreme gradient boosting (XGBoost) [55]. These classifiers have been chosen because of their established performance and efficiency. Each classifier has its own advantage and disadvantage; for example, SVM is useful when the dataset observation is bigger than the features. First, the five classifiers will be evaluated based on their accuracy. After that, the three highest accuracy provided classifiers will be chosen to reduce memory consumption.

To find the optimal parameters for a given classifier, grid search method will be employed to perform hyperparameter tuning based on low mean squared error (MSE) rate. It is guided by a performance metric and typically measured by cross validation on the training set [56]. For each observation (i) and classifier (n) the MSE is calculated as follows:

$$MSE_{i,n} = \sum_{c=1}^{n\_c} \left(A_c - P_{c,n}\right)^2 \Big/ n\_c \tag{18}$$

here, A_c is the actual prediction of the class and P_c is the probability of the predicted output of the being class, and n_c is the total number of classes (e.g., two in binary classification).

To create an ensemble model, this study uses the soft voting approach. In soft voting, the class labels are predicted based on the probabilities p for each classifier. Soft voting returns the class label as argmax of the sum of predicted probabilities. The equation for soft voting is:

$$\hat{y} = arg\ max \sum_{j=1}^{m} w_j p_{ij} \tag{19}$$

For the j_{th} classifier, w_j is the assigned weight. In contrast with hard voting, soft voting can improve the accuracy because it considers more information by using each classifier's uncertainty in the final decision.

IV. EXPERIMENTAL SETUP AND RESULTS

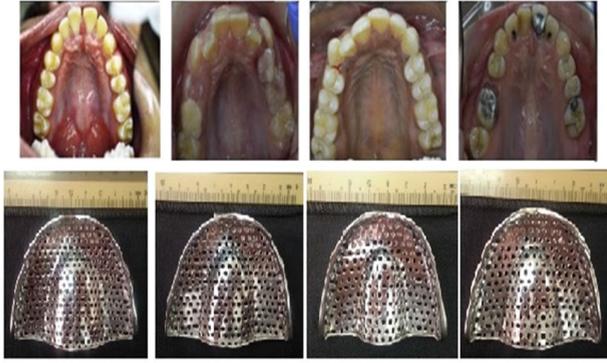To perform all the related experiments, Anaconda release on 64-bit Intel® Core™ with 8 GB of RAM was used.

FIGURE 5. Example of maxillary arches and the trays.

Generally, the dataset consists of 52 real images collected from different patients with age ranging from 16 to 40 for four different types of trays as suggested by the dentist. For the experimental purpose and for faster calculation, the resolution has been set to 300 * 300.

The experimental program is set such that it will automatically convert any image to 300 * 300. Sample of the maxillary dental arch images and impression trays is shown in FIGURE 5. Before identifying the appropriate tray from the maxillary arches, several features were extracted as described previously. In detail, 72 statistical color (SC) features, 1024 color histogram (CH) features, 26 GLCM features, 52 LBP descriptor features and 47 morphological shape (MS) features were extracted. To assess the performance of the proposed method, three sets of experiments were carried out. In the first experiment, the performance of different features set is compared separately and by their fusion using the five classification algorithms to identify which set of features to fuse and to choose from the highest accuracy provider classifiers. In the second experiment, the performance of multi-feature fusion with ensemble classifier is assessed. Finally, deep learning based MLP NN classification performance is evaluated in the third experiment. Statistical tests to investigate the classification performance are carried out as follows (20)–(25), as shown at the bottom of the page.

The terminologies are defined as follows:

True Positives (TP): Number of images correctly classified by the classifier and they indeed belong to the tray.

False Positives (FP): Number of images correctly classified by the classifier, but they do not belong to the tray.

False Negatives (FN): Number of tray images classified falsely, but they belong to that tray.

True Negatives (TN): Number of tray images falsely classified by the classifier and indeed not belong to the tray.

1) PERFORMANCE OF SINGLE FEATURE AND MULTI-FEATURE FUSION

By utilizing more information, multi-feature fusion can accomplish better identification accuracy and to verify this a comparative experiment is implemented. Since color, texture and shape are different feature descriptors, i.e., have distinct dimensions, so each set of features is evaluated separately. At first, the identification accuracy of the individual set of features are evaluated separately. To further validate the effectiveness, other sets of experiment were done on the fusion of best individual features. The identification results (accuracy)% of the individual feature set and their concatenation are shown in Table 3.

According to Table 3, the features set shows different performance of recognition under the five classification algorithms. In accordance with the comparative identification accuracy presented in Table 3, MS features show the best performance on SVM classifier compared to the other single set of features. GLCM features show the worst identification accuracy compared to others. To further validate the results, five different experiments were performed on the fusion of the set of features. It is observed that the fusion of SC, MS and LBP achieves 87.50% on SVM classifier. An interesting finding from the experiment is that the fusion of multiple set of features showed better identification accuracy almost in all cases compared to the single set of features. Thus, it is evident that the multi-feature fusion shows better accuracy compared to the single set of features. Moreover, adding inappropriate features also decreases the classification performance. Two conclusions can be drawn from the analysis: 1) Feature fusion

$$Recall = \frac{TP}{TP + FN} \tag{20}$$

$$Precision = \frac{TP}{TP + FP} \tag{21}$$

$$Accuracy = \frac{TN + TP}{TP + FN + FP + TN} \tag{22}$$

$$JAC = \frac{TP}{TP + FN + FP} \tag{23}$$

$$F1 = \frac{2 * (Precision * Recall)}{(Precision + Recall)} \tag{24}$$

$$MCC = \frac{(TP * TN) - (FP.FN)}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}} \tag{25}$$

TABLE 3. Comparative Performance of Single Set of Features and Multi-Feature Fusion.

Classifiers	Features									
	SC	GLCM	MS	CH	LBP	SC + GLCM + MS + CH + LBP	SC + MS + CH + LBP	SC + MS + LBP	MS + LBP	MS + SC
SVM	50.00	43.75	75.00	56.25	47.50	68.75	84.25	**87.50**	77.50	81.25
KNN (K=3)	68.75	50.00	62.50	62.50	68.75	75.00	75.00	**75.00**	68.75	62.50
RF	43.75	68.75	61.50	62.50	68.75	52.50	68.75	**75.00**	68.75	61.50
GB	56.25	56.25	56.25	56.25	56.25	68.75	68.75	68.75	62.50	56.25
Xgboost	68.75	56.25	50.00	68.75	62.50	81.25	81.25	68.75	68.75	68.75

* Bolded data indicates the best results

TABLE 4. Comparative Performance of Multi-Feature Fusion With Ensemble Classifier.

Classifier	Fusion features (SC + MS + LBP) %								
	Precision	Accuracy	Recall	F1 score	JAC	MCC	Zero one loss	Training time (s)	Testing time (s)
SVM	90.62	87.50	87.89	87.66	78.13	84.04	0.125	0.19	0.03
KNN (K=3)	79.68	75.00	73.46	73.66	67.45	71.01	0.3125	0.11	0.02
RF	62.20	75.00	76.52	67.59	68.85	72.40	0.25	5.10	0.06
GB	68.43	68.75	69.68	67.51	61.67	68.07	0.3252	6.94	0.09
Xgboost	69.91	68.75	67.73	68.01	63.13	66.69	0.25	13.9	0.17
Ensemble	**92.31**	**91.75**	**91.75**	**90.69**	**88.06**	**88.93**	**0.082**	17.4	0.20

* Bolded data indicates the best results

of SC, MS and LBP shows better identification accuracy than single set of feature accuracy; 2) Among the classifiers SVM, KNN and RF show better performance.

2) PERFORMANCE BASED ON ENSEMBLE CLASSIFIER

In this study, a method based on multi-feature-fusion with ensemble classifier is proposed based on soft margin technique. SVM, KNN, and RF are chosen for making the final ensemble classifier. Two significant factors have great influence in ensemble classification performance: hyper parameter tuning and assigning weights. As three classification algorithms were finalized, and the algorithms have different parameters. For example, in SVM the penalty factor C and gamma (γ) have important influence on classification performance [57]. To overcome the first issue grid-search method was adopted, which is a classical way of finding optimal parameters of a given classification algorithm[56]. A range of predefined parameter values was evaluated using 10-fold cross validation to avoid over-fitting. The low MSE rate (Equation (18)) of 10-fold cross-validations is used to evaluate classifier accuracy. At the same time, an exhaustive search was performed to find the optimal weights for each classifier according to the soft voting requirement (Equation (19)). The final weights were chosen [0.5, 0.2, 0.3] for SVM, KNN and RF, respectively. Furthermore, the loss function was calculated for each classification algorithm to estimate the loss of the classifier model. A common loss

function used with classification is Zero-one loss. It assigns 0 to loss for a correct classification and 1 for an incorrect classification. If $\hat{y}_i$ is the predicted value of the i-th sample and the corresponding true value is y_i, then the 0-1 loss L_{0-1} is defined as:

$$L_{0-1}(y_i, \hat{y}_i) = 1(\hat{y}_i \neq y_i) \qquad (26)$$

here, the indication function is defined by 1(x). It is desired that a proper model should have a smaller Zero-one loss value. To further validate the proposed multi-feature fusion with ensemble classifier various statistical experimental results are provided in Table 4. On the basis of, Table 4, it can be observed that ensemble classifier with multi-feature fusion achieves potential results in all sets of features to a certain extent. The multi-feature fusion of color, shape and LBP with ensemble classifier shows 91.75% accuracy with a Zero-one loss value of 0.082, which shows a significant result on this dataset.

3) RESULTS ON MULTILAYER PERCEPTRON NEURAL NETWORK

In addition, deep learning based MLP NN algorithm was employed, and its performance analyzed on the feature fusion (SC, MS, and LBP) of the dataset. The MLP NN was chosen over deep neural network (e.g., CNN) because in case of limited data MLP NN may overcome overfitting issue and showed better classification performance over CNN [58].
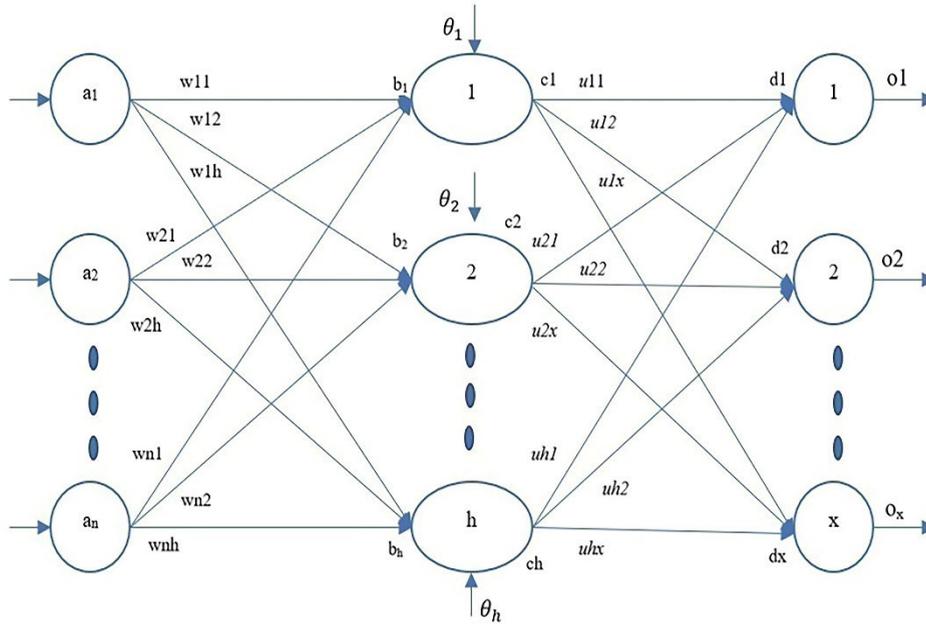
FIGURE 6. One hidden layer MLP NN.

TABLE 5. Experimental Results of MLP NN on Multi-Feature Fusion.

Precision	Accuracy	Recall	F1 score	JAC	MCC	Zero one loss	Training time (s)	Testing time (s)
91.75	88.93	88.82	88.56	79.38	83.42	0.113	8.59	0.13

It is also called second-generation neural network or shallow neural network and commonly has 1 hidden layer (or ≤ 2 hidden layers). In simple terms, an MLP NN can be considered as successively connected series of layers of neurons by weights, which are iteratively adjusted through an optimization process. Back-propagation algorithm is the base for training these neural networks. A three-layer MLP NN is shown in FIGURE 6. From FIGURE 6 it can be seen that one directional connection exists among the nodes which is a special type of the Feed-Forward (FF) neural network family. Here, n defined the number of input nodes, h defined the number of hidden layers and m defined the number of output nodes. The architecture of MLP NN is usually determined by the trial-and-error method. However, in this study, 39 hidden layers, and 'lbfgs' optimizer (which belongs to the family of quasi-Newton methods) was used.

The MLP NN is able to achieve 88.93% accuracy based on Table 5, which is slightly lower than the ensemble classification result. It can be inferred that the poor performance of MLP NN is due to the small dataset size. It may be possible that due to the limited samples and with the increment of hidden layer number it got trapped into poor local minimums [59]. However, various researchers leverage different deep learning algorithms, e.g., transfer learning on a limited dataset and the results showed better accuracy, which outperformed the traditional machine learning

classifiers [60]. Likewise, it is also possible that the specific MLP NN architecture is unsuitable for the selected dataset. Hence, a conclusive result cannot be stated based on the preceding analysis.

V. DISCUSSION

Overall, the rigorous evaluation of multi-feature fusion of statistical color, morphological shape, and local binary pattern with ensemble classifier attained a maximum classification accuracy of 91.75% on this dataset. Additionally, a graphical comparison between the classification accuracy is also presented (FIGURE 7). It can be inferred that single set of color, texture and shape features are unable to describe the arches properly due to the complex lighting changes, absence of texture, and variation of the mouth shapes present in the dataset. In this scenario, the multi-feature fusion with ensemble classifier is able to achieve more than 90% score in precision, recall, and F1 score. The dataset consists of various kinds of low resolution, rotated images but these do not influence the precision, recall, F1 score and accuracy of the proposed method. Also, the Zero-one loss score of ensemble classifier is 0.082, which is lower compared to single machine learning classification algorithms and multilayer perceptron neural network which suggests that the proposed model is worthwhile. This demonstrates that the proposed model based on multi-feature fusion with ensemble classifier

Classification Report

	SVM	KNN(K=3)	RF	GB	XGBoost	MLP NN	Ensemble
Precision	90.62	79.68	62.2	68.43	69.91	91.75	92.31
Accuracy	87.5	75	75	68.75	68.75	88.93	91.75
Recall	87.89	73.46	76.52	69.68	67.73	88.82	91.75
F1 Score	87.66	73.66	67.59	67.51	68.01	88.56	90.69
JAC	78.13	67.45	68.85	61.67	63.13	79.38	88.06
MCC	84.04	71.01	72.4	68.07	66.69	83.42	88.93

FIGURE 7. ML algorithms classification report (model wise).

plays a significant role in discriminating between the maxillary arch images and can be used in real-time systems. The reason that reduced the classification accuracy of multilayer perceptron neural network is due to the small sample dataset utilized in this experiment and also the possibility of getting trapped into poor local minimums. Generally, deep neural networks perform better if they are fed to large numbers of labeled samples. It is difficult, costly, and time-consuming to gather a large number of medical ground truth data; such data has to be medically validated by the practitioners as well. However, as previously discussed, using non-deep learning algorithms performs well on limited training samples and is widely reported to be accurate. In various sectors of image classification tasks, it is still arguable how well these deep learning algorithms perform against machine learning algorithms [61]. Anyway, in this case the ensemble classifier obtained satisfactory performance over multilayer perceptron neural network and the traditional machine learning classification algorithms. The efficacy of the proposed method can be validated by the experiments and gave the following conclusion: 1) The identification performance of multi-feature fusion is significantly superior compared to single set of features; 2) The ensemble classifier method used in this study is more effective for identifying the proper tray compared to the single machine learning classifier. Furthermore, an illustration of the proposed dental tray selection model to use in real-life scenario is shown in FIGURE 8. To illustrate, a dentist can take the occlusal view image of the maxillary

arch by an intraoral camera. The proposed model will process the image and determine which tray may best fit the arch without modifying the tray or use trial and error method even without the need of an expert. It is evident that even expert human observers may miss the proper tray due to different illumination and low resolution of pixels. However, computational algorithms have the potential to tackle these problems with lower error rate, less equipment demand, and less time. This allows instantaneous integration into a clinical computer-assisted diagnostic system. This novel model with computer vision and machine learning presented in this study proved to be capable of selecting appropriate dental impression tray with a potential success rate that signifies its potential for application in clinical practice as part of an automated process chain. Therefore, the hypothesis that computer vision and machine learning can be used to select an appropriate tray from the maxillary arch is confirmed. Thanks to the automation, it is expected to get a fully user-independent process in the future which will eventually eliminate relying on manual inspection.

It is crucial to mention the limitation of the current study. False classification has often been observed on maxillary images with crooked teeth, braces, and images having less illumination. Examples of these images are shown in FIGURE 9. Faults may be corrected by employing additional data in the training sets. In addition, in this current study all the participants are adult patients. The data from infant patients have been omitted. The results will differ if infant
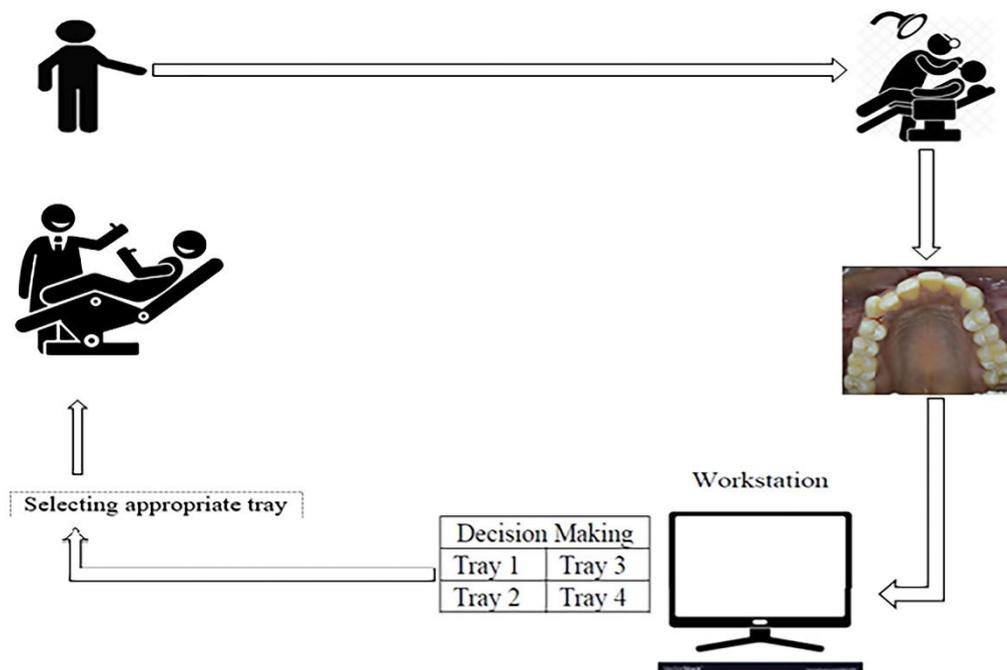
FIGURE 8. Illustration of selecting an impression tray from the proposed model in real-life scenario.

FIGURE 9. Example of identification failures on maxillary arches.

mouth impression images are used. Based on the discussions, a guarantee cannot be given in what scope the results can be directly employed into computer-assisted diagnostics. The authors strongly encourage further studies to strengthen the robustness of the current study.

VI. CONCLUSION AND FUTURE WORK
In this study, multi-feature fusion with ensemble classifier is proposed to select appropriate dental impression tray from maxillary arch images. The proposed method fused statistical color, morphological shape, and local binary pattern features with ensemble classifier to improve its ability in selecting appropriate dental impression tray. Experimental results show that on a limited dataset, the proposed method attains excellent classification results in identifying appropriate dental impression tray with precision (92.31%), recall (91.75%), F1 (90.69%) and accuracy (91.75%). Thus, the hypothesis that computer vision and machine learning are capable of

selecting dental impression tray is confirmed. Although feature fusion with ensemble classifier shows excellent performance, it is crucial to mention that efficiency is the hindrance in the proposed method. Since the features extracted in this study are very large, it is obvious that the data occupies large memory space. Second, in terms of time, the ensemble classification method takes a little bit of computational time. As previously mentioned, deep learning algorithms (e.g., convolutional neural network) show promising results in medical sectors and have a large potential for further improvement. Through convolutional neural network algorithm pre-processing step is not needed; therefore, this step can be skipped to reduce computational costs. In future work, a further study will be done, and a deeper investigation will be made on wide-ranging data to address the efficiency of this method in terms of time and space.

REFERENCES
[1] S. Gupta, A. I. Narayan, and D. Balakrishnan, "*In vitro* comparative evaluation of different types of impression trays and impression materials on the accuracy of open tray implant impressions: A pilot study," *Int. J. Dentistry*, vol. 2017, pp. 1–8, Feb. 2017.

[2] S. Raith, E. P. Vogel, N. Anees, C. Keul, J.-F. Güth, D. Edelhoff, and H. Fischer, "Artificial neural networks as a powerful numerical tool to classify specific features of a tooth based on 3D scan data," *Comput. Biol. Med.*, vol. 80, pp. 65–76, Jan. 2017.

[3] G. J. Christensen, "Will digital impressions eliminate the current problems with conventional impressions?" *J. Amer. Dental Assoc.*, vol. 139, no. 6, pp. 761–763, Jun. 2008.

[4] S. Chatterjee, D. Dey, and S. Munshi, "Integration of morphological preprocessing and fractal based feature extraction with recursive feature elimination for skin lesion types classification," *Comput. Methods Programs Biomed.*, vol. 178, pp. 201–218, Sep. 2019.

[5] L. Yong-Lian, "Multi-feature data mining for CT image recognition," *Concurrency Comput., Pract. Exper.*, vol. 32, no. 1, Jan. 2020, Art. no. e4885.

[6] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[7] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.

[8] E. Yilmaz, T. Kayikcioglu, and S. Kayipmaz, "Computer-aided diagnosis of periapical cyst and keratocystic odontogenic tumor on cone beam computed tomography," *Comput. Methods Programs Biomed.*, vol. 146, pp. 91–100, Jul. 2017.

[9] R. Ramli, M. Y. I. Idris, K. Hasikin, N. K. A. Karim, A. W. A. Wahab, I. Ahmedy, F. Ahmedy, N. A. Kadri, and H. Arof, "Feature-based retinal image registration using D-saddle feature," *J. Healthcare Eng.*, vol. 2017, pp. 1–15, Jan. 2017.

[10] E. Yergin, C. Ozturk, and B. Sermet, "Image processing techniques for assessment of dental trays," in *Proc. 23rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 3, Oct. 2001, pp. 2571–2573.

[11] O. M. Rijal, N. A. Abdullah, Z. M. Isa, F. A. Davaei, N. M. Noor, and O. F. Tawfiq, "A novel shape representation of the dental arch and its applications in some dentistry problems," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2011, pp. 5092–5095.

[12] O. M. Rijal, N. A. Abdullah, Z. M. Isa, N. M. Noor, and O. F. Tawfiq, "A probability distribution of shape for the dental maxillary arch using digital images," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 5420–5423.

[13] C.-W. Wang, C.-T. Huang, J.-H. Lee, C.-H. Li, S.-W. Chang, M.-J. Siao, T.-M. Lai, B. Ibragimov, T. Vrtovec, O. Ronneberger, P. Fischer, T. F. Cootes, and C. Lindner, "A benchmark for comparison of dental radiography analysis algorithms," *Med. Image Anal.*, vol. 31, pp. 63–76, Jul. 2016.

[14] V. Murugappan and R. S. Sabeenian, "Texture based medical image classification by using multi-scale Gabor rotation-invariant local binary pattern (MGRLBP)," *Cluster Comput.*, vol. 22, no. S5, pp. 10979–10992, Sep. 2019.

[15] L. Talavera-Martínez, P. Bibiloni, and M. González-Hidalgo, "Computational texture features of dermoscopic images and their link to the descriptive terminology: A survey," *Comput. Methods Program Biomed.*, vol. 182, Dec. 2019, Art. no. 105049.

[16] Y. Li, S. Wang, Q. Tian, and X. Ding, "Feature representation for statistical-learning-based object detection: A review," *Pattern Recognit.*, vol. 48, no. 11, pp. 3542–3559, Nov. 2015.

[17] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[18] K. T. Ahmed, S. Ummesafi, and A. Iqbal, "Content based image retrieval using image features information fusion," *Inf. Fusion*, vol. 51, pp. 76–99, Nov. 2019.

[19] H. Ali, M. Sharif, M. Yasmin, and M. H. Rehmani, "Color-based template selection for detection of gastric abnormalities in video endoscopy," *Biomed. Signal Process. Control*, vol. 56, Feb. 2020, Art. no. 101668.

[20] N. Bagri and P. K. Johari, "A comparative study on feature extraction using texture and shape for content based image retrieval," *Int. J. Adv. Sci. Technol.*, vol. 80, pp. 41–52, Jul. 2015.

[21] C. Barata, M. Ruela, M. Francisco, T. Mendonça, and J. S. Marques, "Two systems for the detection of melanomas in dermoscopy images using texture and color features," *IEEE Syst. J.*, vol. 8, no. 3, pp. 965–979, Sep. 2014.

[22] N. Z. Tajeddin and B. M. Asl, "Melanoma recognition in dermoscopy images using lesion's peripheral region information," *Comput. Methods Programs Biomed.*, vol. 163, pp. 143–153, Sep. 2018.

[23] D. A. Tyas, S. Hartati, A. Harjoko, and T. Ratnaningsih, "Morphological, texture, and color feature analysis for erythrocyte classification in thalassemia cases," *IEEE Access*, vol. 8, pp. 69849–69860, Mar. 2020.

[24] X. Liu, D. Zhao, W. Jia, W. Ji, and Y. Sun, "A detection method for apple fruits based on color and shape features," *IEEE Access*, vol. 7, pp. 67923–67933, May 2019.

[25] E. Avuçlu and F. Başçiftçi, "Novel approaches to determine age and gender from dental X-ray images by using multiplayer perceptron neural networks and image processing techniques," *Chaos, Solitons Fractals*, vol. 120, pp. 127–138, Mar. 2019.

[26] A. Acevedo, S. Alférez, A. Merino, L. Puigví, and J. Rodellar, "Recognition of peripheral blood cell images using convolutional neural networks," *Comput. Methods Programs Biomed.*, vol. 180, Oct. 2019, Art. no. 105020.

[27] T. Saba, M. A. Khan, A. Rehman, and S. L. Marie-Sainte, "Region extraction and classification of skin cancer: A heterogeneous framework of deep CNN features fusion and reduction," *J. Med. Syst.*, vol. 43, no. 9, p. 289, Sep. 2019.

[28] T. Zheng, Y. Gao, F. Wang, C. Fan, X. Fu, M. Li, Y. Zhang, S. Zhang, and H. Ma, "Detection of medical text semantic similarity based on convolutional neural network," *BMC Med. Inform. Decis. Making*, vol. 19, no. 1, pp. 1–11, Dec. 2019.

[29] A. Onan, "A fuzzy-rough nearest neighbor classifier combined with consistency-based subset evaluation and instance selection for automated diagnosis of breast cancer," *Expert Syst. Appl.*, vol. 42, no. 20, pp. 6844–6852, Nov. 2015.

[30] O. Aytuğ, "Sentiment analysis on Twitter based on ensemble of psychological and linguistic feature sets," *Balkan J. Electr. Comput. Eng.*, vol. 6, no. 2, pp. 69–77, 2018.

[31] A. Onan, S. Korukoğlu, and H. Bulut, "Ensemble of keyword extraction methods and classifiers in text classification," *Expert Syst. Appl.*, vol. 57, pp. 232–247, Sep. 2016.

[32] A. Onan, "Classifier and feature set ensembles for Web page classification," *J. Inf. Sci.*, vol. 42, no. 2, pp. 150–165, Apr. 2016.

[33] B. He, T. Ji, H. Zhang, Y. Zhu, R. Shu, W. Zhao, and K. Wang, "MRI-based radiomics signature for tumor grading of rectal carcinoma using random forest model," *J. Cellular Physiol.*, vol. 234, no. 11, pp. 20501–20509, Nov. 2019.

[34] X. Qiao, J. Bao, H. Zhang, F. Wan, and D. Li, "fvUnderwater sea cucumber identification based on principal component analysis and support vector machine," *Meas. J. Int. Meas. Confederation*, vol. 133, pp. 444–455, Feb. 2019.

[35] A. Duneja and T. Puyalnithi, "Enhancing classification accuracy of k-nearest neighbours algorithm using gain ratio," *Int. Res. J. Eng. Technol*, vol. 4, no. 9, pp. 1385–1388, 2017.

[36] L. Rokach, "Ensemble-based classifiers," *Artif. Intell. Rev.*, vol. 33, nos. 1–2, pp. 1–39, Feb. 2010.

[37] M. Abdar, M. Zomorodi-Moghadam, X. Zhou, R. Gururajan, X. Tao, P. D. Barua, and R. Gururajan, "A new nested ensemble technique for automated diagnosis of breast cancer," *Pattern Recognit. Lett.*, vol. 132, pp. 123–131, Apr. 2020.

[38] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[39] C. R. Dhivyaa and M. Vijayakumar, "An effective detection mechanism for localizing macular region and grading maculopathy," *J. Med. Syst.*, vol. 43, no. 3, p. 53, Mar. 2019.

[40] N. Madian, K. B. Jayanthi, D. Somasundaram, and S. Suresh, "Identifying centromere position of human chromosome images using contour and shape based analysis," *Measurement*, vol. 144, pp. 243–259, Oct. 2019.

[41] M. Souaidi and M. E. Ansari, "Multi-scale analysis of ulcer disease detection from WCE images," *IET Image Process.*, vol. 13, no. 12, pp. 2233–2244, Oct. 2019.

[42] H. Chen, L. Yang, L. Li, M. Li, and Z. Chen, "An efficient cervical disease diagnosis approach using segmented images and cytology reporting," *Cogn. Syst. Res.*, vol. 58, pp. 265–277, Dec. 2019.

[43] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.

[44] B. Dhruv, N. Mittal, and M. Modi, "Study of Haralick's and GLCM texture analysis on 3D medical images," *Int. J. Neurosci.*, vol. 129, no. 4, pp. 350–362, Apr. 2019.

[45] M. A. Khan, M. I. U. Lali, M. Sharif, K. Javed, K. Aurangzeb, S. I. Haider, A. S. Altamrah, and T. Akram, "An optimized method for segmentation and classification of apple diseases based on strong correlation and genetic algorithm based feature selection," *IEEE Access*, vol. 7, pp. 46261–46277, Mar. 2019.

[46] M.-K. Hu, "Visual pattern recognition by moment invariants," *IEEE Trans. Inf. Theory*, vol. IT-8, no. 2, pp. 179–187, Feb. 1962.

[47] S. Abbas, S. Farhan, M. A. Fahiem, and H. Tauseef, "Efficient shape classification using Zernike moments and geometrical features on MPEG-7 dataset," *Adv. Electr. Comput. Eng.*, vol. 19, no. 1, pp. 45–51, Feb. 2019.

[48] A. Onan, "An ensemble scheme based on language function analysis and feature engineering for text genre classification," *J. Inf. Sci.*, vol. 44, no. 1, pp. 28–47, Feb. 2018.

[49] A. Onan, S. Korukoğlu, and H. Bulut, "A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification," *Expert Syst. Appl.*, vol. 62, pp. 1–16, Nov. 2016.

[50] A. Idri, M. Hosni, and A. Abran, "Systematic mapping study of ensemble effort estimation," in *Proc. ENASE*, 2016, pp. 132–139.

[51] A. Onan, "Hybrid supervised clustering based ensemble scheme for text classification," *Kybernetes*, vol. 46, no. 2, pp. 330–348, Feb. 2017.

[52] M. Hosni, I. Abnane, A. Idri, J. M. C. de Gea, and J. L. F. Alemán, "Reviewing ensemble classification methods in breast cancer," *Comput. Methods Programs Biomed.*, vol. 177, pp. 89–112, Aug. 2019.

[53] A. Onan and S. Korukoğlu, "A feature selection model based on genetic rank aggregation for text sentiment classification," *J. Inf. Sci.*, vol. 43, no. 1, pp. 25–38, Feb. 2017.

[54] A. Selvapandian and K. Manivannan, "Performance analysis of meningioma brain tumor classifications based on gradient boosting classifier," *Int. J. Imag. Syst. Technol.*, vol. 28, no. 4, pp. 295–301, Dec. 2018.

[55] A. Tahmassebi, G. J. Wengert, T. H. Helbich, Z. Bago-Horvath, S. Alaei, R. Bartsch, P. Dubsky, P. Baltzer, P. Clauser, P. Kapetas, E. A. Morris, A. Meyer-Baese, and K. Pinker, "Impact of machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy and survival outcomes in breast cancer patients," *Invest. Radiol.*, vol. 54, no. 2, pp. 110–117, Feb. 2019.

[56] C. W. Hsu, C. C. Chang, and C. J. Lin, "A practical guide to support vector classification," Dept. Comput. Sci. Inf. Eng., National Taiwan Univ., Taipei, Taiwan, Tech. Rep., 2003.

[57] Y. Du, J. Liu, W. Song, Q. He, and D. Huang, "Ocean eddy recognition in SAR images with adaptive weighted feature fusion," *IEEE Access*, vol. 7, pp. 152023–152033, Oct. 2019.

[58] J. Yun, J. E. Park, H. Lee, S. Ham, N. Kim, and H. S. Kim, "Radiomic features and multilayer perceptron network classifier: A robust MRI classification strategy for distinguishing glioblastoma from primary central nervous system lymphoma," *Sci. Rep.*, vol. 9, no. 1, pp. 1–10, Apr. 2019.

[59] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.

[60] X. Zhao, Y. Liang, A. J. X. Guo, and F. Zhu, "Classification of small-scale hyperspectral images with multi-source deep transfer learning," *Remote Sens. Lett.*, vol. 11, no. 4, pp. 303–312, Apr. 2020.

[61] S. E. Jozdani, B. A. Johnson, and D. Chen, "Comparing deep neural networks, ensemble classifiers, and support vector machine algorithms for object-based urban land use/land cover classification," *Remote Sens.*, vol. 11, no. 14, p. 1713, Jul. 2019.

MUHAMMAD ASIF HASAN received the B.Sc. degree in computer science and engineering from International Islamic University Chittagong. He is currently pursuing the master's degree in computer science (applied computing) with the University of Malaya. He is also a Graduate Research Assistant with the VLSI Research Lab. His research interests include computer vision, natural language processing, machine learning, and deep learning algorithms.

NORLI ANIDA ABDULLAH received the Ph.D. degree from the University of Malaya, where she modeled the shape of the dental arch to aid dental practitioners in their treatment planning. She is currently a Statistician and a Senior Lecturer with the Center for Foundation Studies in Science, University of Malaya. She has collaborations with cranial surgeons, dentists, and ophthalmologists. Her research interests include biomedical statistics, imaging, and circular data.

MOHAMMAD MUSTANEER RAHMAN received the B.Sc. degree in computer science and engineering from Khulna University, Bangladesh, and the master's degree in computer science from the University of Malaya, Malaysia. He is currently pursuing the Ph.D. degree (Higher Degree by Research) in information technology with the School of Information and Communication Technology, University of Tasmania, Hobart, Australia. He has nine years of industry experience in research and development. He has authored or coauthored in several publications in international journals and conferences. His research interests include e-Learning, affective tutoring systems, recommender systems, adaptive learning, educational technology, spatial crowdsourcing, human-computer interaction, image processing, decision support systems, machine learning, big data, distributed computing, and cloud computing.

MOHD YAMANI IDNA BIN IDRIS (Member, IEEE) received the B.E., M.Sc., and Ph.D. degrees in electrical engineering from the University of Malaya, Kuala Lumpur, Malaysia.

He is currently an Associate Professor with the Department of Computer Systems, Faculty of Computer Science and Information Technology, University of Malaya. His research interests include image processing and computer vision, digital forensics, surveillance systems, and digital signal processing (speech processing and bio-signals).

OMAR F. TAWFIQ was born in Baghdad, Iraq, in 1980. He received the B.D.S. degree in dentistry from Baghdad University, Iraq, in 2004, and the master's degree in prosthodontics from the University of Malaya, Kuala Lumpur, Malaysia, in 2009, where he is currently pursuing the Ph.D. degree in restorative dentistry.

From 2008 to 2011, he was a Research Assistant with the research team in the University of Malaya in a research project sponsored by the Ministry of Science, Technology and Innovation (MOSTI), Malaysia. Since 2011, he has been serving as a Lecturer in the Prosthodontic Department, MAHSA University for two years. Since 2013, he has also been a Lecturer in prosthodontics with the Faculty of Dentistry, SEGi University. He has been co-researcher in five research grants, and authored more than ten articles. His research interests include complete denture rehabilitation, 3D dental cast analysis, and aesthetic dentistry.

Dr. Tawfiq was a recipient of the 2nd Runner Up Award in oral presentation in the International Conference on Innovative Dentistry (ICID 2015) and the 1st Prize for Oral Presentation in MAHSA, in 2013. He is an associate member in several dental professional bodies such as Malaysian and Iraqi Dental Associations and the International Association of Dental Research (IADR).

• • •