Foundation University Journal of Engineering and Applied Sciences

FUJEAS
Vol. 4, Issue 2, 2023.
DOI:10.33897/fujeas.v4i2.794

Research Article

Strategic Analysis of Feature Selection Methods for Enhanced Dental Therapy Recognition in Machine Learning Applications

Shaheena Noor [a, *], Muhammad Imran Saleem [b], Najma Ismat [a], Humera Noor Minhas [c], Rukaiya Rukaiya [a], Aneeta Siddiqui [a]

[a] Department of Computer Engineering, Sir Syed University of Engineering & Technology, Karachi, Pakistan

[b] Department of Software Engineering, Sir Syed University of Engineering & Technology, Karachi, Pakistan

[c] Director of Engineering, Eyeo GmbH, Cologne, Germany

* **Corresponding author**: shanoor@ssuet.edu.pk

Article Citation:
Noor et al. (2023). "Strategic Analysis of Feature Selection Methods for Enhanced Dental Therapy Recognition in Machine Learning Applications". *Foundation University Journal of Engineering and Applied Sciences*
DOI:10.33897/fujeas.v4i2.794

Published by
Foundation University
Islamabad.
Web: https://fui.edu.pk/

Abstract:

The popularity of smart clinics has increased significantly as a result of technical developments in fields like computer vision. At the heart of such systems is the ability to recognize objects and activities as well as perceive the environment as a whole. This is crucial for both eco-independent systems and human-machine interaction, especially in settings with constrained workspaces, like dental care. Our study delves into an extensive analysis of multiple machine learning models designed to robustly predict dental treatments. These models encompass Lazy Instance-based Learning, Sequential Minimal Optimization, Hoeffding Tree, and Random Tree. Leveraging object-oriented input sourced from gaze-guided wearable cameras, we scrutinize intricate attributes such as material properties, patients' dental conditions, and the array of instruments in use. Notably, we exploit the insight that identifying visual cues during an activity holds the potential to address the specific therapy identification challenge. Utilizing a dental data set that we gathered in the real world, we conducted our experiments and discovered that combining multiple criteria enhances accuracy in comparison to using each one alone. We did see, nevertheless, that in certain circumstances employing the symptoms alone produced superior outcomes. Additionally, symptoms demonstrated to have lesser error than combination in terms of RMS error convergence. Finally, we observed that the machine learning models' build and test durations increased as a result of the combined method. This demonstrates that adding additional parameters does not necessarily result in better outcomes in machine learning applications generally and in medical/dental applications in particular. Instead, it relies on the machine learning tool used, the settings taken into account and the input data provided. The versatility of our approach extends beyond dental contexts. It has been systematically validated across diverse domains, including the recognition of kitchen activities within smart home environments. This methodology holds relevance for various outdoor scenarios where the focal point of attention guides ongoing activities.

Keywords: Hoeffding Tree; Inside-out Vision; Lazy Instance-based Learning; Machine Learning; Multinomial Logistic Regression; Random Tree; Sequential Minimal Optimization.

1. Introduction

Recognizing objects and activities are foundational hurdles within the realm of computer vision and machine learning [1]. Object recognition involves the identification and classification of objects or entities within an image or video stream. It plays a crucial role in applications such as autonomous vehicles, robotics, surveillance, and augmented reality.

Activity recognition, on the other hand, focuses on understanding and categorizing human activities based on visual data. This can include actions performed by individuals, interactions between multiple entities, and the temporal sequencing of events. Activity recognition finds applications in healthcare, sports analysis, video surveillance, and human-computer interaction.

Smart systems that incorporate computer vision and image processing have revolutionized various industries and domains and it is making our lives easier [2]. These systems leverage advanced technologies to analyze visual data, extract meaningful information, and make intelligent decisions. We have intelligent systems in indoor environments like smart homes & cities, entertainment & gaming, healthcare, etc., and outdoor scenarios e.g. automotive & transportation. The crux of these systems lies in their profound reliance on perception techniques, prominently encompassing the realms of computer vision and image processing. They need high-performance Machine Learning (ML) solutions on the one hand and advanced image capture systems on the other. One or more humans are at the center of these systems, interacting with the environment to carry out tasks that vary depending on the domain and applications, such as performing surgery in healthcare or cooking in a smart kitchen. The need to experience the world from the perspective of the actor has grown a lot in smart environments. This has significantly altered the working paradigm since the scenario may now be seen from the actor's perspective, providing clues about his cognitive development and removing the need for fixed camera sets surrounding the subject. This kind of visual input, often known as first-person vision, enables the interpretation of complex situations by decomposing them into the individual items that the actor sees and handles. As a result, pertinent data will be recorded and irrelevant data will be ignored. The conventional outside-in perspective is ideal for recognizing people, tracking actions, forecasting activity, and monitoring human behavior in the overall scenario. However, there are difficulties if we try to capture small objects, especially in the presence of self-occlusion. As a result, the accuracy of the overall performance is decreased when certain environmental elements are lacking.

In this research paper, we introduce an inclusive scenario recognition framework that employs objects within the human gaze as contextual cues, coupled with the training of multiple machine learning models for predictive tasks. Within the scope of this study, our focus centers on a specific application, namely dental therapy recognition. The context in this context-aware system is ascertained through the assessment of teeth conditions (referred to as symptoms), the composition of materials, and the array of tools used during the procedure. In pursuit of our objectives, we undertake a comprehensive examination of the efficacy of utilizing these three distinctive features, both in stand-alone and in combination. The machine learning models, encompassing the Hoeffding Tree, Lazy Instance-based Learning (IBk), Multinomial Logistic Regression, Random Trees, and Sequential Minimal Optimization (SMO) techniques, are diligently trained to capture the intricate relationships inherent in the data.

This research paper introduces several novel aspects of the prediction of dental treatments. Firstly, it pioneers the utilization of gaze-guided wearable cameras, an innovative approach to capturing visual data for dental care prediction, an area that has seen limited exploration. Secondly, the methodology incorporates visual cues from wearable cameras for activity recognition, a unique focus within dental care settings. Additionally, it evaluates the effectiveness of combining multiple criteria versus using individual parameters alone, shedding light on the importance of feature combinations in improving prediction accuracy. The versatility of the approach is demonstrated by its applicability beyond dental care, extending to recognizing activities in smart home environments and outdoor scenarios. Lastly, the study outlines a future direction toward deep learning models for treatment recognition, indicating a progressive approach to addressing challenges and improving methodology effectiveness.

The structure of this research paper is outlined as follows: A comprehensive overview of existing literature pertaining to inside-out vision and machine learning algorithms for computer vision applications is presented in Section 2. We explain our approach based on inside-out vision-based identification in Section 3 and present the ML models we utilized in our research. We encompass the provision of our dataset, a detailed exposition of the experimental setup, and a thorough analysis of the obtained outcomes in Section 4. Finally, we sum up our findings in Section 5.

2. Related Work

Inside-out vision is often used in smart homes to detect items [3], and locations, observe human behavior [4], study the nervous system [5], and diagnose & anticipate diseases of the mind. Inside-out vision perspectives are employed, according to Furnari et al. in [6], to predict future behaviors based on interactions with the objects. Inside-out vision cameras have been used in studies on scenario identification to record objects, actions, and activities. The significance of eye movement while interpreting a situation was examined in [7]. Similar to this, in [8] gaze patterns are seen when a scenario is being executed and their relationship to the work being done. In [9], an inside-out user body pose is captured in a virtual reality experience. The system integrates multiple cameras into a handheld controller enabling to perform 3D body pose estimation that offers a colossal VR experience. In [10], a two-stage implantation of voice prosthesis is described using an inside-out approach. Researchers working on machine learning are putting a lot of effort into computer vision and image processing for diagnosing dental diseases and their treatments. In addition to detecting the low-level features that may be inferred from the visual input, such as identifying objects and actions, this also entails analyzing those aspects for scene interpretation and behavior analysis. In [11], Saric et al, utilize a regression algorithm to estimate the dental age based on buccal bone level. The algorithm is based on support vector machines and random forest algorithms. In [12], an in-depth work for predicting dental diseases using a machine learning algorithm is presented. The prediction is performed with six visible classifiers and measured in terms of performance metrics for different machine learning algorithms. In [13], Verma et al. used a hybrid CNN-SVM based machine and deep learning algorithm to detect dental caries infection, changes in periodontal bone height, and third molar impactions. The system extracts the features using CNN and trains the baseline SVM model to classify the

panoramic dental x-rays. In [14], Juneja et al presented a study that investigated the capabilities of deep learning for classifying dental occlusion using 3D stereo lithography (STL) image files which were then converted to 2D histograms using Absolute Angle Shape Distribution (AAD) technique and trained over the machine and deep learning algorithms. An automatic detection approach is provided in [15], to identify and categorize permanent teeth on orthopantomogram (OPG) images. The study uses CCN to evaluate the automated system and achieved high performance for tooth detection and numbering form images. In [16], a home dental care system is presented that classifies tooth diseases and determines the required professional dental treatment. The technology employs deep learning models on images of the maxillary and mandibular teeth to identify early dental issues. It can also integrate patient profiles with mobile applications.

3. The Research Problem

3.1. Inside-out Vision Used for Treatment Identification

3.1.1. Gaze-based Information

The objects visible and the information extracted from FPV images are the focal points of attention. In comparison to outside-in-view, FPV images have much more information as there are no obstruction issues. For example, in Figure 1 [17], the picture was obtained using a wall-mounted camera and the dentist is giving treatment to the patient. The issue is that it is difficult to identify the treatment accurately because the dentist's tools and hands are blocking the image. In addition to this, the condition of the patient's teeth and what the patient is suffering from are also not visible. Figure 2 [18], [19], is the visual taken using an inside-out camera, and it can be seen that the view contains information that the dentist is looking directly at the mouth. This provides much detailed information about the objects and the condition of the teeth, which helps in understanding the disease the patient is suffering from.

Figure 1: Outside-in view: dental patient undergoing treatment [17]. no information: state of the teeth, instruments, or materials not visible in the Image.

Figure 2: Inside-out view: dental treatment being administered [18] [19]. analysis of instruments in use and tooth condition.

Dental treatments of diverse types need a variety of dental instruments, techniques, and materials. These components better help in understanding and predicting the type of treatment the dentist gives to the patient. Pictures shown in Figures 3 to 4 depict different dental procedures, state of teeth, materials, and dental instruments used in it.

Figure 3: Dental treatment (L to R): crowning [20], filling [21], image capture [22], and sealant [23]

Figure 4: Material, state of the teeth, and instruments

The type of dental treatments and the objects like material, symptoms, and instruments are shown in Table 1.

Table 1: Links between treatment and material, state of the teeth & instruments for a dental environment

Treatment	Material	State of the teeth	Instruments
Filling	Ceramics, amalgam, glass ionomer, zinc oxide, resin composites	cracked, decomposed, split	air abrasion or laser, condenser, drill, excavator, carver
Crown	metals, porcelain fused to metal, resin composites, zirconia, ceramics, stainless steel	Split, severely worn down, weakened	burs, cement spatula, retraction cord, handpiece, impression trays, probe
Scaling	chlorhexidine gluconate, scalar tips	plaque, calculus	Sonic, ultra-sonic
Whitening	carbamide peroxide, hydrogen peroxide	pellicle film	slow handpiece, profy
Sealant	acid solution	molar, premolar	plastic instrument, curing light
Root Canal	metals, porcelain fused to metal, resin composites, zirconia, ceramics, stainless steel, guttapercha filling	Split, severely worn down, weakened	burs, cement spatula, retraction cord, handpiece, impression trays, probe

4. Proposed Solution

4.1. Algorithms in Machine Learning for Categorizing Dental Treatments

This section covers the machine learning techniques for treatment identification in the dental arrangement. The study's selection criteria for machine learning algorithms are carefully tailored to meet the specific challenges of predicting dental treatments using visual data from wearable cameras. These criteria encompass various factors crucial for effective algorithm selection. Firstly, accuracy is paramount, as algorithms must reliably predict treatments for precise dental care applications. Additionally, robustness is essential to ensure consistent performance across diverse scenarios, handling variations in lighting, camera angles, and patient demographics. Efficiency plays a vital role in real-world applications, where computational resources are often limited, emphasizing the importance of minimizing resource usage while maintaining accuracy. Interpretability is valued for fostering trust and aiding decision-making among clinicians, necessitating algorithms with some level of interpretability. Scalability ensures that algorithms can handle large datasets and accommodate future expansions without sacrificing performance. Moreover, the ability to handle sequential data is crucial due to the sequential nature of data captured by wearable cameras during dental procedures, suggesting consideration of algorithms like recurrent neural networks. Flexibility and adaptability are sought after to accommodate changing data distributions or the introduction of new techniques without extensive retraining. Finally, previous performance in similar tasks is weighed to leverage algorithms with proven success in related domains. By meticulously considering these criteria, researchers can identify the most suitable algorithms for predicting dental treatments, ensuring optimal performance and relevance in real-world applications. A brief comparative analysis of techniques like Sequential Minimal Operation (SMO), Lazy Instance-based Learning (IBL), Hoeffding Tree, and Random Tree along with the performance, latency, and error metrics. The section also covers the theoretical and mathematical synopsis of these methods.

4.1.1. Multinomial Regression

Multinomial regression serves as a classification method that predicts a category of dependent variables given multiple independent variables. These independent variables encompass both binary (i.e. dichotomous) and continuous (i.e. interval-based or ratios). Distinguishing itself from binomial regression, multinomial regression extends its scope to encompass multiple classes, not confined to a binary scenario. A major advantage of multinomial regression also referred to as multinomial logistic regression, is that it handles outliers very well. Hence, the sample size doesn't have to be set up very carefully. It doesn't assume linearity or normal distribution of data. Finally, the independent variables do not need to be statistically independent of each other, which makes it a very attractive technique for analysis. In order to classify the data using multinomial regression, a mathematical model is defined for which scores are calculated as in Eq. 1.

$$Score(X_i, k) = W_k . X_i \tag{1}$$

where X_i is a vector describing instance i , W_k indicates weights corresponding to category k and score (X_i, k) is the score to assign instance i to class k . The score is converted to a probability and is interpreted as the probability of observation i falling in category k . In multinomial logistic regression, a Ridge estimator is one of the techniques to calculate parameters and minimize errors [24]. Here, out of the k classes, one of them (e.g. the last one) is chosen as the pivot, and logistic regression models are run over the rest $k-1$ outcomes. Hence for all but the last class, the score i.e. probability is calculated as in Eq. 2:

$$P_j(X_i) = \frac{\exp^{X_i . W_k}}{1 + \sum_{j=1}^{k-1} \exp^{X_i . W_k}} \tag{2}$$

An alternative representation of the model can be realized through a latent variable model instead of independent log-linear models. In contrast to several existing methodologies, this approach entails lower computational costs per iteration to estimate the local minima.

4.1.2. Lazy Instance-based Learning (IBk)

Instance-Based Learning (IBk) provides a straightforward implementation of the k-nearest neighbors (k-NN) classifier, as described in [25]. This is achieved by applying distance weighting and choosing the closest value of k based on cross-validation. An evaluation of the Bayes and Lazy classification techniques in [25] reveals the effectiveness of the Lazy classifier in comparison to the Bayesian classifier.

4.1.3. Sequential Minimal Optimization (SMO)

Sequential machine optimization (SMO) [26] is a technique developed to deal with the training support of vector machines. It reduces the computational cost in terms of the complexity of very large Quadratic Programming (QP) with increased accuracy. The LIBSVM tool [27] implemented SMO with some variations as proposed by Platt.

SVM training is an optimization issue that needs a very large quadratic program to be solved. Consider a binary classification problem [28] dataset with an input vector $(x_1, y_1), \ldots, (x_n, y_n)$, and y_i is the corresponding binary label. Eq. (3) expresses a soft-margin SVM.

$$\max_\alpha \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j K(x_i, x_j, \alpha_i \alpha_j) \tag{3}$$

subject to:

$$0 \le \alpha i \le C, \qquad for \qquad i = 1, 2, \ldots n \tag{4}$$

$$\sum_{i=1}^{n} y_i \alpha_i = 0 \tag{5}$$

where *C* is the hyper parameter of SVM, *K(xi, xj)* is the kernel function, and α_i are Lagrange multipliers. SMO is an iterative solution that divides the large QP into smaller tiny QPs and solves it analytically. This allows the limitation on any two Lagrange multipliers α_1 and α_2 to be simplified to:

$$0 \le \alpha_1, \alpha_2 \le C,$$
$$y_1\alpha_1 + y_2\alpha_2 = k$$

The resulting simplified problem can be effectively addressed through analytical means, by finding a minimum of a $1D$ function over the remaining terms in the equality constraint.

4.1.4. Hoeffding Tree

The Hoeffding Tree [29] is a machine-learning solution for data stream processing within the realm of data mining. An enhanced version of the Hoeffding Tree is the Very Fast Decision Tree (VFDT), a decision tree characterized by accelerated speed and optimized memory utilization. VFDT functions as an incremental tree induction algorithm, operating under the assumption that the underlying sample distribution within extensive data streams remains unchanged over time. It assumes that a small sample set is enough to choose the optimal splitting attribute. The tree growth is based mathematically on the Hoeffding bound [30]. The Hoeffding bound calculates the number of observations or examples needed to estimate statistics within a prescribed precision. It is different in that it guarantees that its output is identical to non-incremental trees.

4.1.5. Random Tree

In essence, the outcome of a stochastic process yields a random tree, characterized as a decision tree. Constructed without any form of cropping, this tree is established by selecting k attributes at each node in a random manner. Furthermore, the tree is configured in such a manner that facilitates the estimation of class probabilities through a back-fitting approach.

5. Experiments

This section covers the experimentation and discussion of the results obtained.

5.1. Dataset

Due to the novelty of the approach, very little data is available online that covers the first-person perspective of the dentist. Some smaller datasets were found e.g. on kaggle [31], however, they were (1) limited to teeth conditions only and didn't have any information on material or tools used and (2) captured via x-ray machine hence needed specialized image capturing setup. Keeping in view the limitations of existing datasets, we collected our data in the wild. The data is extracted from the real and synthetic videos of dental procedures and utilized for constructing and testing models. This is a more naturalistic but inherently complicated dataset due to the randomness and unstructured nature of the input. It is also considered a more reliable test set for evaluating ML models. A combination of images depicting dental procedures and symptoms along with some dental tools is shown in Figure 5.

This section also encompasses the setup of the machine learning models employed to predict the treatment outcomes. Preprocessing steps before training the model involve data cleaning, scaling, encoding, and transformation. Data cleaning rectifies errors, missing values, and duplicates, improving accuracy and reliability. Scaling normalizes variable values for consistent model performance.

Figure 5: Illustrative image samples for dental treatment prediction

Encoding converts categorical variables into numerical ones, essential for regression models. Transformation adjusts variable distributions to meet model assumptions. These steps collectively enhance data quality and model performance. The implementation of these models and simulations is facilitated through the utilization of the ML tool WEKA [32].

5.2. Multinomial Logistic Regression

In this study, a logistic model is trained and operated using the logistic regression class in the tool WEKA with a ridge estimator. Any missing values are replaced with their respective filters, and nominal classes are changed to binary if there are any. Values were calculated with up to four decimal places, in batches of 100, and with no limit on the number of repetitions. The initial value of the ridge is 1.0E-8.

5.3. Lazy Instance-based Learning (IBk)

Euclidean distance is the foundation of the WEKA implementation of Lazy IBk for nearest neighbor search. In IBk, parameter settings and preprocessing steps are defined as follows: The parameter k determines the number of nearest neighbors considered for predictions, with the default value set to 1, implying reliance solely on the closest neighbor. While distance weighting is a common practice, this instance opts not to employ it, attributing equal weight to all neighbors. Euclidean distance serves as the metric for computing distances between instances. The dataset is processed in batches of 100 instances for efficient resource management. Computations maintain a precision of two decimal places for accurate distance calculations. Moreover, Lazy IBk trains the model without limiting the maximum number of examples, enabling comprehensive learning from all available data. These settings ensure effective operation for classification tasks.

5.4. Sequential Minimal Optimization (SMO)

The WEKA implementation of the SMO function is the original algorithm proposed by Platt for training the SVM classifier. The nominal attributes are converted into binary and any missing values are replaced globally. The solution for multi-class issues employs pairwise classification, or 1-on-1. A random seed for the cross-validation is utilized for computations with a batch size of 100 and is accurate to two decimal places. We use a logistic calibrator, a polynomial kernel, and normalized training data. The default setting for the round-off error's epsilon is 1.0E-12.

5.5. Hoeffding Tree

Implementing the Hoeffding Tree involves employing the Adaptive Naive Bayes as the strategy for leaf prediction. The leaf's required number of instances to observe before attempting a split is configured at 200. The split decision's permissible error is 1.0E-7, and it is based on the info gain split criteria. For the computations, we set the batch size to 100. The lowest fraction of weight necessary to decrease is set to 0.01 and the hoeffding threshold is set to 0.05; below these thresholds, a split is required to break ties. Hoeffding Trees are chosen for their efficiency in incremental learning, scalability, and support for online learning. They can adapt to new data without retraining on the entire dataset, making them ideal for scenarios with data streams or large datasets. Additionally, they handle concept drift effectively, adjusting their structure over time. With good performance and interpretability, they are well-suited for various classification tasks, particularly in domains where transparent models are preferred. Overall, these advantages make Hoeffding Trees a suitable choice for the given task's requirements and constraints.

5.6. Random Tree

To train a random tree-based model in WEKA, we set the random seed value for attribute selection to 1. We also discarded the unclassified instances before training. For calculations, a minimum total weight of the leaf instances is set at 1.0, and no back-fitting is applied. A batch size of 100 with two decimal places is used. When several attributes look equally good for calculations, breaking off ties randomly is disabled and the tree is allowed to grow randomly for unlimited depth. A minimum variance proportion of 0.001 is established, and the calculation for determining the number of attributes chosen randomly is calculated as given:

$$\mathrm{int}(\log_2(No.\_Of\_\mathrm{Pr}edictors)+1)$$

6. Results and Discussion

The outcomes of the experiments are shown in Table 2. Conforming to our initial hypothesis, objects in focus came out as a powerful predictor of treatment recognition, even if only one of the three conditions is met i.e. material, state of the teeth, and instruments are considered. Moreover, experiments show that combining multiple parameters improves the overall performance metrics. Combining multiple criteria in dental therapy recognition enhances accuracy through various mechanisms, including leveraging complementary information, mitigating bias and variance, improving discriminative power, increasing robustness to noisy data, enhancing generalization, capturing complex decision boundaries, and mitigating overfitting. By considering different aspects of dental treatments simultaneously, models perform better and it discriminate between treatment types, balance biases, reduces prediction variability, and compensates for noisy or incomplete data. This approach leads to more accurate predictions and better generalization to unseen cases, while also promoting the learning of meaningful patterns and preventing the memorization of noise in the training data. Overall, the integration of multiple criteria optimizes model performance and reliability in dental therapy recognition tasks.

The ML models were recreated and trained using the retrieved object information as a focal point. As expected, when taking into account the combined information instead of just the individual factors, the precision and recall are usually increased. When the state of the teeth parameter is employed individually, the highest precision and recall are obtained for the Multinomial Logistic Regression and Tree Hoeffding techniques.

Table 2: Performance metrics for treatment identification

Method	Material		State of the teeth		Tools		Combined	
	Precision %	Recall	Precision %	Recall	Precision %	Recall	Precision %	Recall
Logistic	68.7	68.4	97.5	97.1	94.1	89.8	88.7	78.6
LazyIBk	80.4	73.7	81.7	61.7	90.4	76.3	1	1
SMO	81.9	68.4	88.8	85.3	87.7	74.6	95	83.3
Hoeffding Tree	81.3	71.1	97.5	97.1	90.4	76.3	92.9	85.7
Random Tree	84.9	55.3	90.5	79.4	80.4	54.2	94	61.9

The training and testing times are also measured in seconds for all the models used for the experimentation as shown in Table 3. The time smaller than 0.01 seconds is automatically rounded off to zero. It is noted that the combined models are the slowest to build and test, which is a disadvantage for computationally intensive applications or very large datasets. It is also observed that SMO requires the longest time to construct and Hoeffding Tree requires the longest to test.

Table 3: Time (in seconds) needed to construct and test the models

Method	Material		State of the teeth		Tools		Combined	
	Construct	Test	Construct	Test	Construct	Test	Construct	Test
Logistic	0.01	0	0.01	0	0	0	0.02	0.03
LazyIBk	0	0.1	0	0	0	0	0	0.02
SMO	0.06	0	0.06	0	0.06	0	0.14	0.03
Hoeffding Tree	0	0	0	0	0	0	0.01	0.06
Random Tree	0	0	0	0	0	0	0	0.02

In Table 3, combining multiple parameters generally leads to a slight increase in both construction and testing times across methods like Logistic Regression, SMO, and Hoeffding Tree. The magnitude of this increase varies depending on the method and the nature of the combined parameters. For instance, while Logistic Regression shows a small increase, methods like SMO and Hoeffding Tree exhibit more noticeable increases in construction time. However, testing time may not necessarily see the same increase. Overall, the impact of combining parameters on processing times varies based on the method and the complexity of the parameters.

The Root Mean Squared Error (RMS) of the various models is also examined (see Table 4), and it is shown that the state of the teeth parameter gave the lowest error in the majority of situations.

In this paper, two different dimensions are considered. The first one is the parameters and the second are ML approaches. For the first case, we are using three parameters: material, state of the teeth, and dental instruments. We used them individually for treatment recognition and then combined them together. Here our analysis showed that the combined approach results in higher precision. However, the combined approach is slower comparatively. Also, the RMS (root mean squared) error is higher in the case of the combined approach compared to using "state of the teeth" alone.

For the second case, five different machine learning algorithms have been analyzed, which are logistic, lazy IBK, SMO, Hoeffding tree, and random tree. As mentioned above, combining parameters is better than training them individually. But if we closely look at the values of precision, "state of the teeth" also

Table 4: Root mean squared error

Method	Material	State of the teeth	Tools	Combined
Logistic	0.3088	0.1064	0.1743	0.2673
LazyIBk	0.3059	0.2564	0.1989	0.2113
SMO	0.3397	0.3165	0.3227	0.3277
Hoeffding Tree	0.318	0.0762	0.2026	0.2182
Random Tree	0.389	0.2039	0.3906	0.3563

results in good performance for some ML approaches. Hence, it cannot be generalized that one ML algorithm always performs better than others. Rather, it is a combination of the input data and the ML approach.

The study introduces a methodology for predicting dental treatments through machine learning models and visual data from wearable cameras, presenting notable strengths alongside acknowledged limitations. Firstly, there is a limitation of a real-world dental dataset. Additionally, classic machine learning methods' reliance on pre-extracted features poses challenges, given the complexity and resource-intensiveness of feature extraction from visual data. The methodology's sole dependence on visually perceivable parameters for feature extraction might overlook other influential factors like patient history or non-visual cues. The study evaluates model performance based on metrics like accuracy and speed, additional clinical relevance metrics could provide a more comprehensive assessment. Addressing these limitations could bolster the proposed methodology's robustness, generalizability, and usability in real-world healthcare settings.

7. Conclusions and Future Direction

In this study, we employed an inside-out perspective to identify dental treatments through the application of machine learning algorithms. We investigated various traditional machine learning approaches and provided them with diverse input features, including the condition of the teeth, types of materials used, and dental instruments utilized. The models were trained using these individual parameters as well as in combination.

The fundamental constraint of this approach lies in the fact that classic ML methods require features to be pre-extracted and fed for training. We were restricted to use the visually perceivable parameters as features, which limited the type of data used for training. Furthermore, the process of feature extraction is both time-intensive and resource-demanding. Moreover, based on the performance, speed, and error convergence factors, it was determined that combining numerous parameters sometimes improved performance. It also resulted in low error convergence and poor speed. Additionally, not all ML models perform well with the combined approach; the optimal trade-off is achieved with the inclusion of simply the state of the teeth parameter. This demonstrates that the inclusion of parameters may not provide better outcomes for ML applications generally or especially in the area of the dental environment. Instead, it relies on the ML tool used, the settings taken into account and the input data provided.

Going forward, we aim to delve into treatment recognition using deep learning models. This approach will enable us to inherently extract features, thereby broadening the scope to encompass additional non-semantic visual cues.

Acknowledgment

We are thankful to the research reviewers who helped in writing and improving these guidelines.

Conflict of Interests

Publication of this research article has no conflict of interest.

Funding Statement

This research article is not a product of a funded research project.

8. References

[1] M. S. S. Reddy, P. R. Khatravath, N. K. Surineni and K. R. Mulinti, "Object Detection and Action Recognition using Computer Vision," In *International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, Coimbatore, India, 2023, pp. 874–879, doi: 10.1109/ICSCSS57650.2023.10169620.

[2] E. Dilek, M. Dener, "Computer Vision Applications in Intelligent Transportation Systems: A Survey," *Sensors*, vol. 23(6), p.2938, 2023, https://doi.org/10.3390/s23062938.

[3] H. Kang, A. A. Efros, M. Herbert and T. Kanade. "Image Matching in Large Scale Indoor Environment," In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshop on Egocentric Vision*, IEEE, 2009, pp. 33–40.

[4] L. Sun, U. Klank and M. Beetz, "EYEWATCHME 3D Hand and object tracking for inside out activity analysis," In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, IEEE, 2009, pp. 9–16.

[5] M. Hayhoe and D.Ballard, "Eye movements in natural behavior," *Trends in Cognitive Sciences*, vol. 9(4), pp. 188–194, 2005.

[6] A. Furnari and G. M. Farinella, "What Would You Expect? Anticipating Egocentric Actions with Rolling-Unrolling LSTMs and Modality Attention," In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 6252-6261.

[7] J. M. Henderson and A. Hollingworth, "High-level scene perception," *Annual Review of Psychology*, vol. 50(1), pp. 243–271, 1999.

[8] K. Rayner, T. J. Smith, G. L. Malcolm and J. M. Henderson, "Eye movements and visual encoding during scene perception," *Psychological Science*, vol. 20(1), pp. 6–10, 2009.

[9] K. Ahuja, V. Shen, C. M. Fang, N. Riopelle, A. Kong and C. Harrison, "Controller Pose: Inside-Out Body Capture with VR Controller Cameras," In *CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–13.

[10] N. Hanai, S. Beppu, D. Nishikawa, H. Terada and M. Sawabe, "A novel procedure of secondary voice prosthesis insertion from the inside out: The modified Fukuhara method," *Auris Nasus Larynx*, vol. 49(4), pp. 658–662, 2022.

[11] R. Saric, J. Kevric, N. Hadziabdic, A. Osmanovic, M. Kadic, M. Saracevic, D. Jokic and V. Rajs, "Dental age assessment based on CBCT images using machine learning algorithms," *Forensic Science International*, vol. 334, p.111245, 2022.

[12] I. F. Rimi, M. Arif, A. Islam, S. Akter, M. Rahman, A. H. M. Islam and M. Habib, "Machine learning techniques for dental disease prediction," *Iran Journal of Computer Science*, vol. 5(3), pp. 187–195, 2022.

[13] D. Verma, S. Puri, S. Prabhu and K. Smriti, "Anomaly detection in panoramic dental x-rays using a hybrid Deep Learning and Machine Learning approach," In *2020 IEEE Region 10 Conference (TENCON)*, IEEE, 2020, pp. 263–268.

[14] M. Juneja, R. Singla, S. K. Saini, R. Kaur, D. Bajaj and P. Jindal, "OCLU-NET for occlusal classification of 3D dental models," *Machine Vision and Applications*, vol. 31(6), pp. 1–12, 2020.

[15] M. Estai, M. Tennant, D. Gebauer, A. Brostek, J. Vignarajan, M. Mehdizadeh and S. Saha, "Deep learning for automated detection and numbering of permanent teeth on panoramic images," *Dentomaxillofacial Radiology*, vol. 51(2), p.20210296, 2022.

[16] D. Kim, J. Choi, S. Ahn and E. Park, "A smart home dental care system: integration of deep learning, image sensors, and mobile controller," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–9, 2021.

[17] https://en.wikipedia.org/wiki/Dentist [Accessed: 31-Jan-2020]

[18] [https://karedentalvancouver.com/ourservices/dental-cleaning/ [Accessed: 17-Aug-2023]

[19] https://www.topdoctors.co.uk/medicalarticles/detecting-removing-and-preventing-tooth-decay-this-is-how-itworks# [Accessed: 17-Aug-2023]

[20] https://www.newsmedical.net/health/Dental-Crown-Tooth-Cap.aspx [Accessed: 17-Aug-2023]

[21] https://greendentalalexandria.com/choosingthe-right-type-of-dental-filling/ [Accessed: 17-Aug-2023]

[22] https://www.pinecovedental.com/intra-oralcamera [Accessed: 17-Aug-2023]

[23] https://sandimassedation.com/blog/whatmaterials-are-used-to-make-dental-sealants/ [Accessed: 17-Aug-2023]

[24] S. L. Cessie and J. V. Houwelingen, "Ridge Estimators in Logistic Regression," *Applied Statistics*, vol. 41(1), pp. 191–201, 1992.

[25] S. Vijayarani and M. Muthulakshmi, "Comparative Analysis of Bayes and Lazy Classification Algorithms," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2(8), 2013.

[26] C. J. Platt, "Fast Training of Support Vector Machines using Sequential Minimal Optimization," Advances in Kernel Methods - Support Vector Learning, Published by MIT Press, 1998.

[27] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2(3), pp. 1–27, 2011.

[28] https://en.wikipedia.org/wiki/Sequential minimal optimization#cite note-2 [Accessed: 17-May-2021]

[29] G. Hulten, L. Spencer, P. Domingos, "Mining time-changing data streams," In *Proceeding of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2001, pp. 97–106.

[30] https://en.wikipedia.org/wiki/Hoeffding%27s inequality [Accessed: 24-Nov-2022]

[31] https://www.kaggle.com/datasets/daverattan/dentalxrary-tfrecords [Accessed: 24-Nov-2022]

[32] https://www.cs.waikato.ac.nz/ml/weka/ [Accessed: 17-May-2021]