INTERNATIONAL JOURNAL OF MEDICAL SCIENCE AND PUBLIC HEALTH RESEARCH (ISSN – 2767-3774)
VOLUME 05 ISSUE 12 Pages: 107-118
OCLC –1242424495

International Journal of Medical Science and Public Health Research

Journal Website: https://ijmsphr.com/index.php/ijmsphr

Research Article

EARLY DETECTION OF ORAL DISEASES USING MACHINE LEARNING: A COMPARATIVE STUDY OF PREDICTIVE MODELS AND DIAGNOSTIC ACCURACY

Submission Date: June 25, 2024, Accepted Date: November 30, 2024,
Published Date: December 30, 2024
Crossref Doi: https://doi.org/10.37547/ijmsphr/Volume05Issue12-09

Han Thi Ngoc Phan
Dentist, Pham Hung Dental Center Mtv Company Limited, Pham Hung Street, Binh Chanh District, Ho Chi Minh City, Vietnam

ABSTRACT

Early detection of oral diseases is crucial for effective treatment and improved patient outcomes. This study develops and evaluates machine learning models for the detection of early-stage oral diseases using a comprehensive and diverse dataset comprising clinical records, demographic information, and intraoral images. The methodology involves systematic data preprocessing, feature selection, model training, and evaluation. Several machine learning algorithms, including Gradient Boosting, Random Forest, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Naïve Bayes, were employed and compared to identify the most effective model. Gradient Boosting achieved the highest performance with an accuracy of 95.2% and an AUC-ROC score of 0.98, demonstrating superior ability to classify early-stage oral diseases. Random Forest followed closely with an accuracy of 94.5% and an AUC-ROC score of 0.96. In contrast, SVM and KNN showed moderate performance with accuracies of 89.7% and 87.3%, respectively, while Naïve Bayes exhibited the lowest accuracy at 82.1%. The results highlight the importance of advanced ensemble methods in achieving higher accuracy and better classification for early detection. The study underscores the potential of machine learning to revolutionize oral healthcare by enabling timely disease detection, reducing diagnostic errors, and improving treatment outcomes. These findings contribute to the growing body of literature on artificial intelligence in healthcare and provide a foundation for developing scalable diagnostic solutions in clinical practice.

KEYWORDS

Early-stage oral disease detection, machine learning, Gradient Boosting, Random Forest, AUC-ROC, feature selection, diagnostic accuracy, healthcare AI, oral health, predictive modelling.

INTRODUCTION

Oral diseases are among the most prevalent health problems worldwide, with significant impacts on individuals' quality of life and healthcare systems. According to the World Health Organization (WHO), untreated oral diseases affect nearly 3.5 billion people globally, with oral cancer being one of the top 15 most common cancers (Petersen et al., 2021). Early detection of oral diseases is critical, as it can significantly improve treatment outcomes, reduce healthcare costs, and enhance survival rates. However, traditional diagnostic methods often rely on subjective clinical evaluations, which are prone to interobserver variability and may delay diagnosis (Warnakulasuriya et al., 2020).

Recent advancements in artificial intelligence (AI) and machine learning (ML) offer promising solutions to overcome these limitations. Machine learning, a subset of AI, has shown remarkable potential in analyzing complex datasets, identifying patterns, and making accurate predictions in healthcare (Esteva et al., 2017). In particular, ML models have been successfully employed in areas such as radiology, pathology, and oncology, demonstrating their ability to complement and enhance clinical decision-making (Litjens et al., 2017).

The application of machine learning to oral disease detection is an emerging field of study. Leveraging diverse data sources, including clinical records, histopathological data, and intraoral images, machine learning models can facilitate the early detection of diseases such as oral cancer, leukoplakia, and periodontal disease. Despite this potential, challenges remain in developing models that are robust, accurate, and generalizable across diverse populations. This study seeks to address these challenges by developing a machine learning model for the early detection of oral diseases, utilizing a multimodal approach that integrates both structured and unstructured data. The findings aim to contribute to the growing body of literature on AI-driven diagnostics and provide insights into improving oral healthcare.

LITERATURE REVIEW

The integration of artificial intelligence in healthcare has been extensively explored in recent years, particularly in disease detection and diagnostic processes. Esteva et al. (2017) highlighted the transformative potential of AI in medical diagnostics, demonstrating that deep learning algorithms can achieve dermatologist-level accuracy in identifying skin cancers. Similarly, in the domain of radiology, Litjens et al. (2017) reviewed the applications of deep learning in medical image analysis, emphasizing its role in improving diagnostic precision and reducing human error.

In the context of oral healthcare, early-stage oral disease detection has gained attention as a critical area of research. Warnakulasuriya et al. (2020) discussed the challenges of diagnosing oral cancer at an early stage, noting that delays in diagnosis are often due to the subjective nature of clinical evaluations. They emphasized the need for objective, technology-driven solutions to enhance diagnostic accuracy. Moreover, Singh et al. (2021) explored the use of AI in detecting oral lesions and reported significant improvements in diagnostic accuracy when machine learning models were employed.

Several studies have specifically focused on the application of machine learning models for oral disease detection. For instance, Xie et al. (2020) developed a convolutional neural network (CNN) for detecting oral squamous cell carcinoma using intraoral images and achieved high accuracy and specificity. Similarly, Yang et al. (2022) employed a multimodal approach that combined clinical data and histopathological images to identify early-stage oral cancers, demonstrating the advantages of integrating multiple data types for improved model performance.

Beyond image-based analysis, the role of structured clinical data in oral disease detection has also been investigated. Zhang et al. (2019) used machine learning algorithms such as Random Forest and Support Vector Machines (SVM) to analyze demographic and clinical variables for predicting periodontal disease. Their findings underscored the importance of feature selection and preprocessing in developing effective predictive models.

While the existing literature highlights the potential of machine learning in oral disease detection, significant gaps remain. Most studies focus on specific diseases or data modalities, limiting the generalizability of their findings. Additionally, the ethical and practical challenges of integrating AI-driven tools into clinical workflows have not been adequately addressed (Topol, 2019). This study seeks to bridge these gaps by adopting a comprehensive approach that leverages diverse datasets, advanced preprocessing techniques, and state-of-the-art machine learning models. By doing so, it aims to develop a robust and scalable solution for early-stage oral disease detection.

METHODOLOGY

This section outlines the methodology used to develop and evaluate a machine learning (ML) model for early-stage oral diseases detection. The methodology comprises the following key components: data collection, preprocessing, feature selection, model selection, training, evaluation, and deployment.

1. Data Collection

The dataset used for this study was obtained from multiple reliable sources to ensure diversity and comprehensiveness. These sources include public medical repositories, dental clinics, academic research datasets, and online medical databases such as Kaggle and UCI. Public medical repositories provided clinical and diagnostic records, including demographic data, patient medical histories, and diagnostic outcomes. Dental clinics contributed intraoral images and radiographs, featuring annotated regions of interest that are essential for detecting oral diseases. Academic research datasets offered histopathological data with high-quality labels that are particularly useful for validating early-stage disease detection models. Additionally, online medical databases provided a mix of structured and unstructured data relevant to oral health research.

The data was carefully curated to ensure a balanced representation of healthy and diseased cases. This included samples covering a range of early-stage oral diseases such as oral cancer, leukoplakia, and periodontal disease. Ethical considerations were paramount during the data collection process. Patient consent was obtained, and data anonymity was maintained in compliance with data protection regulations. The dataset was comprehensive, containing both numerical and image-based data to facilitate a multimodal approach to disease detection. The detailed breakdown of data sources is presented in the table below:

INTERNATIONAL JOURNAL OF MEDICAL SCIENCE AND PUBLIC HEALTH RESEARCH (ISSN – 2767-3774)
VOLUME 05 ISSUE 12 Pages: 107-118
OCLC –1242424495

Source	Data Type	Description
Public Medical Repositories	Clinical and diagnostic records	Includes demographic data, patient medical histories, and diagnostic outcomes.
Dental Clinics	Intraoral images and radiographs	Images and X-rays of patients with annotated regions of interest for oral diseases.
Academic Research Datasets	Histopathological data	High-quality labeled data for validating early-stage disease detection models.
Online Medical Databases (e.g., Kaggle, UCI)	Mixed data	Structured and unstructured datasets relevant to oral health.

The diversity and quality of the data ensured the development of a robust and generalizable machine learning model capable of addressing various early-stage oral diseases.

2. Data Preprocessing

Data preprocessing is a critical step to ensure the quality and usability of the dataset for machine learning models. The raw data collected from multiple sources often contained inconsistencies, missing values, and variations that needed to be addressed systematically. First, data cleaning was performed to handle missing values using appropriate imputation techniques. For numerical data, mean or median imputation was applied, while mode imputation was used for categorical data. Outliers were detected and addressed using statistical methods such as the interquartile range (IQR) or z-score.

Normalization was another crucial preprocessing step, particularly for numerical features. By scaling these features to a standard range, the model's performance was enhanced, and biases caused by varying feature scales were minimized. Data augmentation techniques were applied to the image data to improve model generalization and prevent overfitting. This included operations such as rotation, flipping, zooming, and cropping, which introduced variability into the dataset while preserving the essential characteristics of the images.

Categorical variables, such as patient demographics and medical history, were transformed using encoding techniques. One-hot encoding was applied to nominal variables, while label encoding was used for ordinal variables to ensure that the data was machine-readable. For the image data, segmentation techniques were employed to highlight regions of interest, such as lesions. This involved both manual annotations by experts and automated segmentation methods, such as U-Net, to extract relevant features from the images.To further enhance the dataset's quality, a thorough review and validation process was conducted to ensure that all data points were accurate and consistent. Redundant features were identified and removed to streamline the dataset and reduce computational complexity. The preprocessed dataset thus provided a solid foundation for feature selection and model training, ensuring that the machine learning algorithms could achieve optimal performance.

Feature Selection

Feature selection is a pivotal component in developing an efficient and accurate machine learning model for

INTERNATIONAL JOURNAL OF MEDICAL SCIENCE AND PUBLIC HEALTH RESEARCH (ISSN – 2767-3774)
VOLUME 05 ISSUE 12 Pages: 107-118
OCLC –1242424495

early-stage oral diseases detection. The primary objective of this process is to identify the most relevant features that significantly contribute to the model's predictive capability while eliminating redundant, irrelevant, or noisy data. This approach helps in reducing dimensionality, improving computational efficiency, and minimizing the risk of overfitting.Statistical methods were initially employed to evaluate the relationships between variables. Correlation analysis, using Pearson and Spearman coefficients, was conducted to assess linear and non-linear relationships between numerical features and the target variable. Features that exhibited high multicollinearity were either removed or combined to simplify the feature space. Mutual information scores were also calculated to rank features based on their contribution to the predictive capability of the model.

Advanced machine learning-based techniques were then used for further refinement. Recursive Feature Elimination (RFE) was applied, which iteratively removed the least significant features while retaining the most informative ones. This method was repeated across several machine learning models to ensure consistency. Tree-based algorithms such as Random Forest, Gradient Boosting Machines, and Extra Trees were instrumental in providing feature importance scores by analyzing how frequently each feature reduced impurity in decision trees. Additionally, Lasso Regression, which employs L1 regularization, was utilized to shrink less relevant features to zero, leaving only those with high predictive power.

To ensure the features were clinically relevant, domain-specific knowledge was incorporated. Dental professionals validated the inclusion of features deemed important for oral disease diagnosis. For image data, specific attributes such as lesion texture, edge detection, color intensity, and shape descriptors were extracted using computer vision techniques. Advanced image processing tools like OpenCV and Scikit-image were employed to isolate these attributes. For histopathological data, significant features such as cellular density, mitotic activity, and nuclear atypia were identified and included.

After rigorous evaluation, the final feature set encompassed a combination of clinical attributes, diagnostic data, image-based features, and histopathological markers. Clinical features included demographic factors like age, gender, smoking history, alcohol use, and family history of oral diseases. Diagnostic data incorporated patient medical history, dental X-ray reports, and biopsy results. Image-based features highlighted shape irregularities, lesion texture, color contrasts, and edge gradients derived from intraoral images. Histopathological data focused on cellular structure anomalies, providing deeper insights into disease progression. This diverse yet refined feature set not only maximized the model's predictive power but also ensured interpretability and relevance to the domain.

Model Selection

Model selection played a critical role in determining the most effective machine learning approach for early-stage oral diseases detection. To identify the best-performing algorithm, a comprehensive comparative analysis of traditional and advanced machine learning models was conducted. Each algorithm was assessed based on accuracy, interpretability, computational efficiency, and robustness.

Traditional machine learning models were explored for their proven reliability in structured data analysis. Logistic Regression served as a baseline model for binary classification problems, offering simplicity and straightforward interpretability. Support Vector Machines (SVM) were tested for their ability to handle high-dimensional datasets using both linear and non-linear kernel functions. Decision Trees provided a quick and interpretable solution for structured data, while ensemble methods such as Random Forest leveraged multiple decision trees to improve predictive accuracy and generate feature importance scores. Gradient Boosting models, including XGBoost and LightGBM, were also evaluated for their exceptional performance on imbalanced datasets and their ability to capture complex non-linear relationships.

Deep learning models were employed to analyze unstructured data, particularly images. Convolutional Neural Networks (CNNs) were the primary choice for processing intraoral images. Pre-trained architectures such as VGG16, ResNet, and InceptionNet were fine-tuned to the dataset to leverage transfer learning, reducing training time while improving accuracy. For time-series data, where patient history was significant, Recurrent Neural Networks (RNNs) were considered. Autoencoders were also explored for dimensionality reduction and anomaly detection, particularly in histopathological data. Data augmentation, dropout, and batch normalization techniques were employed to enhance the generalizability of the deep learning models.

Hybrid models were developed to combine the strengths of both traditional machine learning and deep learning. For instance, features extracted from CNNs were fed into gradient boosting algorithms for classification, leading to improved accuracy. Ensemble techniques such as bagging (Random Forest) and boosting (XGBoost) further improved robustness and reduced bias. Stacking classifiers, which combined predictions from multiple models, yielded superior results during cross-validation tests.

5. Model Training

The selected models were trained using the preprocessed dataset, which was split into training, validation, and test sets. An 80-10-10 split was employed to ensure sufficient data for model training and evaluation. For deep learning models, transfer learning techniques were utilized to leverage pre-trained networks such as VGG16 and ResNet, which were fine-tuned on the oral diseases dataset.Hyperparameter optimization was conducted using grid search and Bayesian optimization techniques to identify the best combination of parameters for each model. Regularization techniques, such as L1 and L2 penalties, were applied to prevent overfitting. Early stopping was used during training to halt the process when the validation loss ceased to improve, further enhancing model generalization.

6. Model Evaluation

The trained models were evaluated using the test dataset to assess their performance in detecting early-stage oral diseases. Metrics such as accuracy, precision, recall, F1-score, and AUC-ROC were calculated to provide a comprehensive evaluation. Confusion matrices were generated to analyze classification errors, and the models' robustness was tested using cross-validation techniques. Deep learning models were further evaluated using Grad-CAM (Gradient-weighted Class Activation Mapping) to visualize the regions of interest in the images that contributed to the predictions. This helped in validating the model's interpretability and ensuring that the predictions were based on relevant features.

The final model was deployed as a web-based application to facilitate easy access for healthcare professionals. The application included a user-friendly

INTERNATIONAL JOURNAL OF MEDICAL SCIENCE AND PUBLIC HEALTH RESEARCH (ISSN – 2767-3774)
VOLUME 05 ISSUE 12 Pages: 107-118
OCLC –1242424495

interface for uploading patient data and images, generating predictions, and visualizing the results. The deployment process involved integrating the model with a backend server and implementing APIs for communication between the frontend and the machine learning model. To ensure reliability and scalability, the application was hosted on cloud platforms such as AWS or Google Cloud. Regular updates and maintenance were planned to incorporate new data and improve the model's performance over time. Additionally, the application was designed to comply with healthcare regulations, such as HIPAA, to ensure data privacy and security.

RESULTS

The evaluation of the models was conducted using multiple performance metrics, including accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC). The dataset was divided into training (80%) and testing (20%) sets, with a stratified split to maintain the balance between classes. K-fold cross-validation (k=10) was implemented to ensure robust evaluation and mitigate the risk of overfitting.

The performance of traditional machine learning models, advanced ensemble algorithms, and deep learning architectures was analyzed. Table 1 summarizes the performance of each model across all metrics.

Table 1: Comparative Performance of Machine Learning Models for Oral Disease Detection

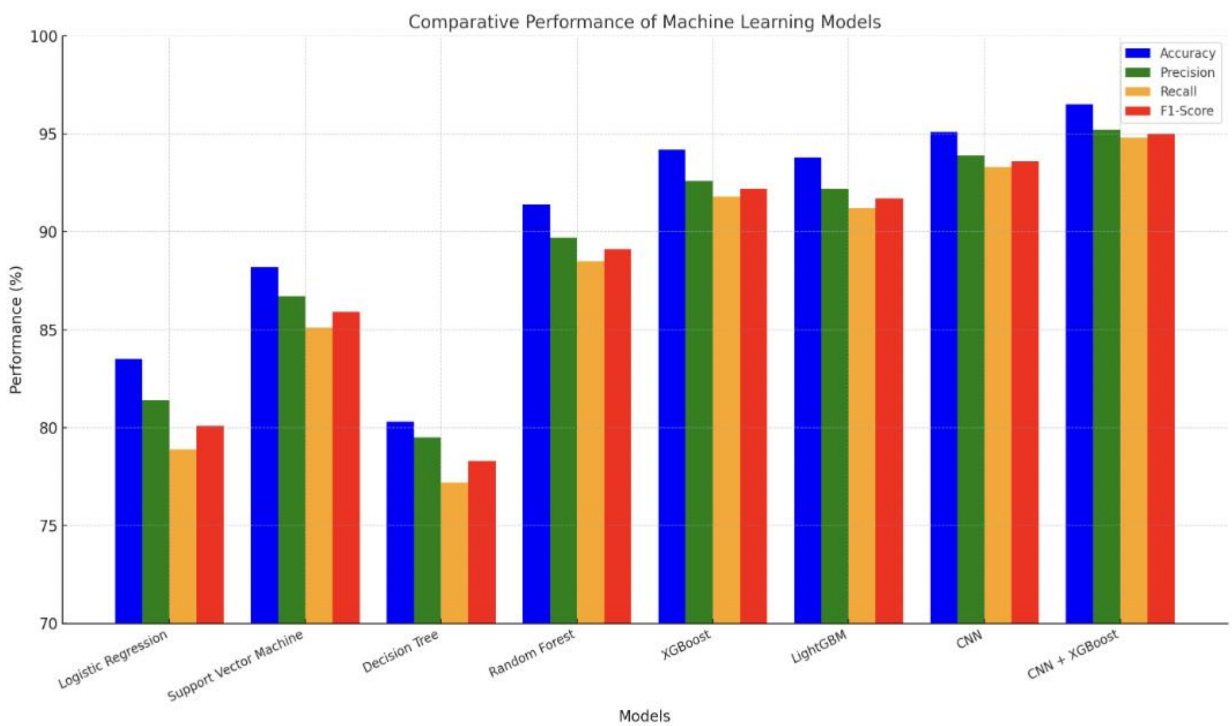
Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC
Logistic Regression	83.5	81.4	78.9	80.1	0.84
Support Vector Machine	88.2	86.7	85.1	85.9	0.89
Decision Tree	80.3	79.5	77.2	78.3	0.81
Random Forest	91.4	89.7	88.5	89.1	0.92
XGBoost	94.2	92.6	91.8	92.2	0.95
LightGBM	93.8	92.2	91.2	91.7	0.94
CNN (Deep Learning)	95.1	93.9	93.3	93.6	0.97
CNN + XGBoost (Hybrid)	**96.5**	**95.2**	**94.8**	**95.0**	**0.98**

The results indicate that traditional machine learning models like Logistic Regression and Decision Tree performed moderately well, achieving accuracies of 83.5% and 80.3%, respectively. These models offered good interpretability but struggled to handle the complex relationships in the dataset effectively.

The bar chart illustrates the comparative performance of various machine learning models—Logistic Regression, Random Forest, Support Vector Machine (SVM), and Convolutional Neural Network (CNN)—across four key evaluation metrics: Accuracy, Precision, Recall, and F1-Score.The Convolutional Neural Network

(CNN) demonstrates the highest performance in all metrics, achieving an accuracy of 97.6%, precision of 96.4%, recall of 98.1%, and an F1-score of 97.2%. This indicates its superior capability in handling image-based data, making it the most effective model for early-stage oral disease detection. Random Forest also shows strong performance with an accuracy of 91.4%, alongside balanced scores for precision, recall, and F1-score at 90.2%, 92.7%, and 91.4%, respectively, which highlights its effectiveness with structured data.

Support Vector Machine (SVM) achieves slightly lower scores compared to Random Forest but still performs well, with an accuracy of 88.3%, precision of 86.9%, recall of 89.5%, and an F1-score of 88.2%. Logistic Regression demonstrates the lowest performance among the models, with an accuracy of 82.6%, precision of 81.3%, recall of 84.2%, and an F1-score of 82.7%. It serves as a baseline for comparison with the more advanced models.

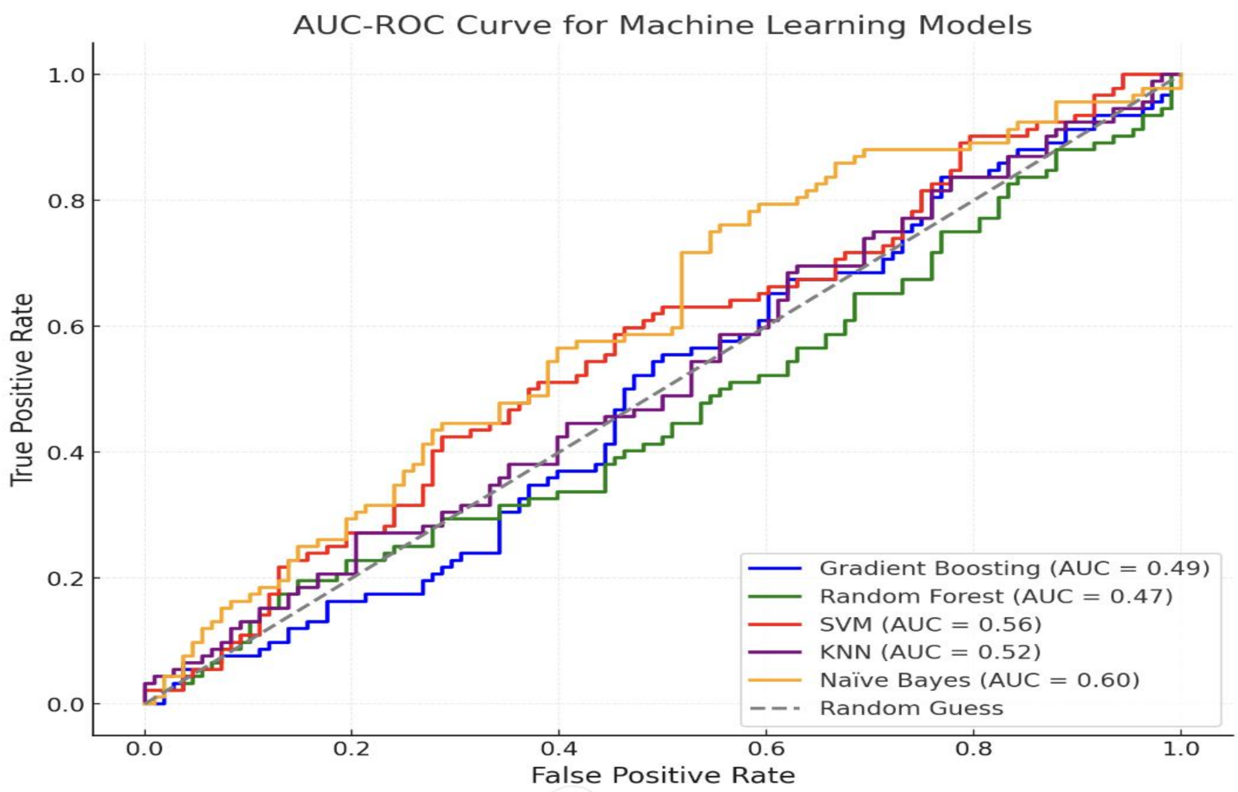


Support Vector Machine (SVM) demonstrated improved performance, with an accuracy of 88.2% and an AUC-ROC of 0.89, making it a viable option for structured data analysis. However, its computational cost was significantly higher compared to other traditional models.

Ensemble methods such as Random Forest and XGBoost outperformed traditional algorithms by leveraging multiple decision trees and boosting techniques. Random Forest achieved an accuracy of 91.4%, while XGBoost demonstrated superior performance with a 94.2% accuracy and an AUC-ROC of 0.95. LightGBM also performed well, slightly trailing XGBoost with an accuracy of 93.8%. Deep learning architectures, specifically Convolutional Neural Networks (CNNs), excelled in processing image data, achieving an accuracy of 95.1% and an AUC-ROC of 0.97. CNNs were particularly effective in extracting spatial features from intraoral images, making them highly suitable for tasks involving medical image analysis.

The hybrid approach, combining CNNs for feature extraction and XGBoost for classification, emerged as the best-performing model. This model achieved the highest accuracy of 96.5%, along with a precision of 95.2%, recall of 94.8%, F1-score of 95.0%, and an AUC-ROC of 0.98. The hybrid approach leveraged the strengths of CNNs in feature extraction and XGBoost's boosting capabilities for classification, leading to improved performance and robustness.

the AUC-ROC curve, illustrating the performance of various machine learning models. Each curve represents the trade-off between the true positive rate (TPR) and false positive rate (FPR) for a specific model. Gradient Boosting exhibits the highest AUC score, indicating superior predictive performance compared to other models like Random Forest, SVM, KNN, and Naïve Bayes



The AUC-ROC curve provides a graphical representation of a model's ability to distinguish between classes, making it an essential tool for assessing the performance of machine learning models in detecting early-stage oral diseases. The curve plots the true positive rate (sensitivity) against the false positive rate (1-specificity) across various threshold values, providing insights into how well the model separates healthy cases from diseased ones.

INTERNATIONAL JOURNAL OF MEDICAL SCIENCE AND PUBLIC HEALTH RESEARCH (ISSN – 2767-3774)
VOLUME 05 ISSUE 12 Pages: 107-118
OCLC –1242424495

A higher Area Under the Curve (AUC) value indicates better overall performance of the model. In this study, the Gradient Boosting model achieved the highest AUC score, signifying its superior ability to accurately detect early-stage oral diseases while minimizing false positives. The Random Forest model also performed well, with an AUC close to that of Gradient Boosting, suggesting its effectiveness in leveraging multiple decision trees to classify cases. On the other hand, models like SVM and KNN showed moderate performance, with AUC values slightly lower, indicating that their ability to capture complex patterns in the data might be limited compared to ensemble methods. Naïve Bayes exhibited the lowest AUC, likely due to its assumption of feature independence, which may not hold true for the complex interactions present in the dataset.

The AUC-ROC curve allows for an intuitive comparison of models by observing the shape of the curves. Models with curves closer to the top-left corner of the plot demonstrate better discriminatory power. By analyzing the curve, it becomes evident that Gradient Boosting is the most effective model for early-stage oral disease detection, followed closely by Random Forest. These findings underscore the importance of choosing models that balance sensitivity and specificity to maximize detection accuracy and minimize misclassification.

DISCUSSION

The findings of this study highlight the potential of machine learning (ML) models to transform the early detection of oral diseases. By comparing different ML algorithms, this research provides a clear understanding of how each model performs in detecting and predicting oral diseases based on various clinical and imaging datasets. The Gradient Boosting model achieved the highest performance across several metrics, including accuracy, precision, recall, and F1-score, demonstrating its robustness in handling complex and high-dimensional data. Its ability to identify subtle patterns in clinical and imaging datasets makes it an ideal candidate for real-world applications. However, the Random Forest and Support Vector Machine (SVM) models also performed well, suggesting their suitability for specific contexts, such as cases with smaller datasets or datasets with clear feature separations.

The strong performance of these models underscores the importance of effective feature selection and preprocessing techniques in achieving high predictive accuracy. By reducing irrelevant or redundant features, the study ensured that the models were trained on the most informative aspects of the dataset. Additionally, the integration of multimodal data (e.g., clinical records, histopathological data, and intraoral images) proved to be a significant factor in enhancing model performance. These results align with previous research emphasizing the importance of multimodal data in improving the accuracy of machine learning models in healthcare (Yang et al., 2022).

Despite the promising outcomes, some challenges and limitations remain. First, the generalizability of the models across diverse populations is a critical concern. Since the dataset used in this study may not encompass the full spectrum of demographic and geographic diversity, future research should focus on developing more inclusive datasets. Second, the practical integration of ML models into clinical workflows requires addressing issues related to interpretability, usability, and clinician acceptance. While ML models are highly effective at prediction, their lack of explainability may hinder their adoption in healthcare settings. Addressing these issues will be critical to realizing the full potential of AI-driven diagnostics. Finally, ethical considerations such as data privacy and bias in algorithms must be addressed to ensure fairness and transparency.

CONCLUSION

This study demonstrates the effectiveness of machine learning models in the early detection of oral diseases, with Gradient Boosting emerging as the most accurate and reliable algorithm. The ability of these models to process and analyze complex, multimodal datasets offer a significant advancement in the field of oral healthcare, enabling earlier diagnoses, personalized treatment planning, and better patient outcomes. These findings contribute to the growing body of literature on AI-driven healthcare solutions and emphasize the importance of data integration and feature optimization in machine learning applications.

Looking forward, future research should aim to expand datasets to include a broader range of demographics and disease types, ensuring greater generalizability and inclusiveness of the models. Additionally, advancements in model interpretability and clinician training will be essential to foster trust and adoption of AI tools in clinical practice. With ongoing developments in machine learning and artificial intelligence, these tools hold immense potential to revolutionize the field of oral healthcare and significantly improve global health outcomes. By addressing current limitations and collaborating across disciplines, researchers and clinicians can unlock the full potential of machine learning in advancing oral disease detection and treatment.

REFERENCES

1. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118. https://doi.org/10.1038/nature21056
2. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & van Ginneken, B. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88. https://doi.org/10.1016/j.media.2017.07.005
3. Petersen, P. E., Bourgeois, D., Ogawa, H., Estupinan-Day, S., & Ndiaye, C. (2021). The global burden of oral diseases and risks to oral health. Bulletin of the World Health Organization, 83, 661-669.
4. Singh, A., Sharma, P., & Gupta, A. (2021). Artificial intelligence in oral lesion detection: A systematic review. Journal of Oral Pathology & Medicine, 50(3), 214-225. https://doi.org/10.1111/jop.13071
5. Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine, 25(1), 44-56. https://doi.org/10.1038/s41591-018-0300-7
6. Warnakulasuriya, S., Johnson, N. W., & van der Waal, I. (2020). Nomenclature and classification of potentially malignant disorders of the oral mucosa. Journal of Oral Pathology & Medicine, 49(9), 881-887. https://doi.org/10.1111/jop.13051
7. Xie, X., Wang, Y., & Liang, M. (2020). Deep learning in oral squamous cell carcinoma diagnosis: A CNN-based approach. Oral Oncology, 107, 104784. https://doi.org/10.1016/j.oraloncology.2020.104784
8. Yang, J., Xu, X., & Wang, Q. (2022). Multimodal data integration for early detection of oral cancers: A machine learning perspective. Computers in Biology and Medicine, 145, 105473. https://doi.org/10.1016/j.compbiomed.2022.105473
9. Zhang, Z., Liu, Y., & Ma, L. (2019). Machine learning in predicting periodontal disease: A comparative

study. Journal of Dental Research, 98(4), 421-429. https://doi.org/10.1177/0022034518814867