Classification of Oral Cancer into Pre-Cancerous Stages Using Mobilenetv2

Sumaiya Abul Hasan[1], Ms. Shaik Asha[2]
[1]PG Scholar, Department of CSE, Muffakham Jah College of Engineering and Technology
Sumaiyahasan830@gmail.com
[2]Assistant Professor, Department of CSE, Muffakham Jah College of Engineering and Technology,
shaik.asha@mjcollege.ac.in

Abstract*—Almost 10 million deaths from cancer were recorded in 2020, making it one of the world's major causes of death. Oral cancer is the sixth most common kind globally. Its lethality is mostly ascribed to late-stage diagnoses, when therapy becomes more difficult. However, mortality rates can be considerably decreased by early identification, especially in precancerous phases. Enhancing survival rates requires early screening and treatment, underscoring the necessity of effective diagnostic techniques. This study proposes a way to identify oral cavity lesions in their pre-cancerous phases and differentiate between benign and malignant lesions. To extract colour and texture-based information from oral cavity images which are essential for recognizing different lesion stages this method investigates five distinct colour spaces. Using MobileNetV2 for enhanced speed and accuracy, the suggested approach combines deep learning classification with handcrafted feature extraction, making it unique. The model provides an effective tool for detecting oral cancer by extracting intricate colour and texture patterns from the photos, surpassing conventional techniques in terms of computational efficiency and time. The model is a great option for affordable, mobile-based diagnostic tools because of its capacity to operate with constrained resources. The technique successfully distinguishes between benign, malignant, and precancerous lesions, showing encouraging results in both binary and multi-class classifications. This method could significantly improve the early identification of oral cancer, particularly in areas with limited access to cutting-edge healthcare facilities.

I. INTRODUCTION

An important global health issue, oral cancer is the sixth most common cancer globally and a major cause of cancer-related mortality. Almost 10 million people died of cancer in 2020, and oral cancer was a major cause of these deaths. The majority of cases are diagnosed at advanced stages, when treatment choices are few and ineffective, which is a major cause in its high fatality rate. Early screening and diagnosis are essential since early detection and management can dramatically lower the death rate. For effective prevention and treatment, it is essential to recognize and appropriately identify precancerous lesions in the oral cavity. This allows for prompt intervention before the illness reaches an advanced state. But these methods can be time-consuming, frequently necessitate specific training, and sometimes fail to deliver timely or correct results. Consequently, there is an increasing demand for automated technologies that can help medical practitioners detect oral cancer early on, lowering diagnostic mistakes and improving early intervention initiatives. Machine learning methods have become effective tools for addressing this issue, particularly those that process medical pictures. This study proposes a novel approach to identify oral cavity lesions in their pre-cancerous phases and distinguish between benign and malignant lesions. Because the MobileNetV2 model can automatically train and extract features from raw photos, it is a good fit for this purpose. This eliminates the need for manual feature extraction, making the system accurate and efficient. The method extracts pertinent colour and texture data from the photos by examining five different colour spaces. After receiving these features, the MobileNetV2 architecture categorizes the lesions into groups like benign, malignant, and precancerous phases. With the help of MobileNetV2, the system can process images quickly and accurately, which makes it appropriate for real-time clinical applications. The suggested method also requires fewer resources, which is important for diagnostic applications that are mobile and may have limited processing capacity.

II. OBJECTIVE

The main goal of this project is to create a machine learning-based system that is extremely accurate and effective at classifying lesions in the oral cavity and distinguishing between benign, malignant, and precancerous stages. This system attempts to automate the process of lesion recognition and categorization from medical images, minimizing human error and time needed in manual analysis. Early detection of oral cancer is essential for improving patient outcomes. In addition to being computationally efficient and lightweight enough to run on mobile and embedded devices, the system's use of the MobileNetV2 architecture guarantees that it can accurately handle complicated image classification tasks. Because of this feature, the system may be implemented in real-time, providing physicians and other medical professionals with instant results. Apart from enhancing diagnostic effectiveness, the goal is to develop a solution that can be applied in distant and low-resource environments where access to cutting-edge medical technology may be restricted. The system will be able to process pictures of lesions in the oral cavity taken under different circumstances and categorize them using elements of colour and texture that are retrieved. The ultimate objective is to develop a scalable, reasonably priced tool that can help medical professionals identify oral cancer early, allowing for speedier diagnoses, minimizing needless biopsies, and possibly saving lives by acting before the disease reaches the precancerous or early stages. To prove the system's superiority in accuracy and efficiency, its performance will be verified by extensive testing and contrasted with cutting-edge techniques.

2.1 Problem Statement

One of the most common forms of cancer in the globe, oral cancer is killed mostly by late-stage diagnosis, when there are fewer and less effective treatment choices. In order to lower mortality rates and enhance survival outcomes, early identification is essential, especially in the precancerous phases. But because the early indicators of oral cancer are so subtle and frequently invisible, detecting the disease is still difficult. The time-consuming procedures, advanced equipment, and knowledge needed for current diagnostic techniques may not always be available in environments with minimal resources. Consequently, a diagnostic tool that can accurately define precancerous phases, distinguish between benign and malignant lesions, and diagnose oral cancer early is desperately needed.

2.2 Existing System

The sophisticated machine learning method LightGBM (Light Gradient Boosting Machine), which is based on decision trees, is used in the current system for the identification of oral cancer. Because of its excellent performance and effectiveness when working with big datasets, LightGBM is frequently utilized for classification jobs. Colour and texture information taken from photographs of the mouth cavity are among the elements that are processed and given into the LightGBM model for classification in the current method. Furthermore, because LightGBM isn't designed to capture spatial relationships or hierarchical elements in images, it might not be the best option for image data with complicated patterns. Additionally, while the current system relies on a more conventional method of feature extraction and classification, it may not be as effective or flexible for real-time applications, particularly in contexts with limited resources or mobility.

Disadvantage of Existing System

- The requirement for hyperparameter tuning
- Reliance on hand-crafted features restriction to tabular data
- Having trouble managing complex image features
- Being less adaptable for mobile and resource-constrained environments
- Being less efficient for real-time applications

2.3 Proposed System

A state-of-the-art Convolutional Neural Network (CNN) architecture created for efficiency in mobile and embedded devices, MobileNetV2 is introduced in the suggested system. MobileNetV2 provides a more sophisticated and resource-efficient method for detecting oral cancer than the conventional LightGBM algorithm. The suggested method captures both colour and texture features without requiring manually created feature extraction by feeding oral cavity images straight into the MobileNetV2 model, which automatically learns hierarchical feature representations from the raw image data. The proposed MobileNetV2-based system is perfect for real-time, on-site diagnostics for oral cancer detection because it has several advantages over the current LightGBM-based system, such as faster processing times, increased accuracy, and flexibility for deployment in low-resource environments.

Advantages of Proposed System

- End-to-end deep learning that is effective for mobile and embedded devices
- Performs better on image data
- Processes information more quickly
- Uses fewer resources, is more scalable and adaptable
- Improves accuracy.

III. RELATED WORKS

Several studies have explored the use of machine learning and deep learning techniques for the early detection and classification of oral cancer. In [1], the authors presented an overview of oral cancer detection using both traditional machine learning and deep learning techniques. Conventional classifiers such as Support Vector Machine (SVM), Naïve Bayes, and K-Nearest Neighbour (KNN) were used to classify oral lesions into normal and abnormal categories. The study highlighted the importance of computer vision in assisting medical diagnosis and reducing the complexity associated with manual examination.

Prachi Shah et al. [2] proposed an algorithm-based approach for early oral cancer detection using image analysis. The method utilized MATLAB image processing techniques, including red-value distribution and Grey-Level Co-occurrence Matrix (GLCM) features, to differentiate normal images from premalignant lesions such as erythroplakia and leukoplakia. The system achieved effective classification by combining color based segmentation and texture feature extraction. In [3], the authors introduced a Computer-Aided Diagnosis model for oral cancer classification using Sailfish Optimization with deep learning. The approach combined feature extraction from VGG16 and ResNet networks, followed by classification using an Extreme Learning Machine (ELM). The Sailfish Optimization algorithm was applied for optimal parameter tuning, resulting in improved performance with a reported accuracy of 98.11%.

Begum and Vidyullatha [4] proposed a deep learning based oral cancer detection method using DenseNet201 with attention maps. The model utilized transfer learning and region-of-interest detection to focus on cancerous areas in oral images. By guiding the CNN model to concentrate on relevant lesion regions, the approach improved classification accuracy and achieved an accuracy of 84.7%. Despite these advancements, several existing approaches [1][4] still face limitations such as dependence on handcrafted features, limited dataset diversity, and high computational complexity in deep learning models. These challenges may affect the robustness, scalability, and real-time applicability of oral cancer detection systems in clinical settings. The proposed study addresses these issues by implementing MobileNetV2, a lightweight deep learning architecture for the classification of oral cancer into precancerous stages. By utilizing transfer learning, data augmentation, and systematic performance evaluation, the proposed model aims to achieve improved accuracy, robustness, and efficient deployment for early oral cancer screening.

IV. METHODOLOGY OF PROJECT

The effectiveness of the system will be assessed and contrasted with current state-of-the-art methods, with an emphasis on important metrics including specificity, recall, accuracy, and precision. The ultimate objective is to develop an efficient, scalable, and economical method of detecting oral cancer, providing a useful instrument for early diagnosis, especially in places with limited access to healthcare services.

V. MODULE DESCRIPTION

Gathering the Dataset:

Before creating any machine learning model, the data must be gathered. It entails finding pertinent datasets that offer the data required for analysis and model training. Data may be gathered in this step from a variety of sources, such as public databases, health records, questionnaires, sensors, or other pertinent platforms. To build an accurate model, it is essential to make sure the dataset is representative of the issue area, extensive, and diverse. The performance and dependability of the machine learning system in later phases are directly impacted by the caliber and diversity of the data.

Data Examination:
Analyzing and examining the gathered data to determine its patterns, structure, and any underlying linkages is known as data examination. In order to find possible problems like missing values, outliers, or inconsistencies that could impair model performance, researchers and data scientists can thoroughly examine the dataset in this step. Descriptive statistics and visualizations are frequently employed during data analysis to highlight important features and offer insights into the dataset. This phase is crucial for directing later choices on preprocessing and model selection.

Data Cleaning:
When getting the dataset ready for model training, data cleaning is an essential step. This process entails dealing with missing data, fixing mistakes, and eliminating noisy or unnecessary information that can distort the analysis. Along with recording categorical variables and making sure the dataset is consistent and standardized, it also entails normalizing or scaling numerical features. The performance of the model is directly impacted by data cleaning, which guarantees that the dataset is prepared for the following stages of the machine learning pipeline and enhances the quality of the input data.

Model Deployment:
Model deployment is the process of putting a machine learning model that has been trained into a business setting where it can interact with actual data. In this step, the model is integrated into an already-existing system, application, or service so that its predictions can be used. Model deployment is possible on a number of platforms, including embedded systems, mobile apps, and cloud services. The model's scalability, efficiency, and real-time performance must all be carefully considered in order to guarantee that it will produce precise and timely forecasts when required.

Model Fitting:
Training a machine learning model using the prepared dataset is known as model fitting. By modifying its parameters in response to the input-output pairs that are supplied, the model discovers patterns and connections in the data during this phase. In this step, a suitable algorithm is chosen, model parameters are initialized, and the model is optimized using gradient descent or other optimization techniques. The objective is for the model to reliably forecast results based on learnt patterns and to generalize well to new, unknown data.

Model Testing and Forecasting:
The process of testing a trained model with a different, previously unseen collection of data is known as model testing. In addition to preventing overfitting to the training set, this guarantees that the model performs well when applied to new data. The model's performance in terms of accuracy, precision, recall, and other assessment metrics can be evaluated with the use of model testing. In the last stage, forecasting, future results are predicted using the trained model and fresh input data. In machine learning, forecasting is the process of using the learnt model to produce predictions for unknown values or occurrences. Forecasting is useful for making decisions in many areas, including financial forecasting, weather forecasting, and disease outbreak forecasting.

VI. ALGORITHM USED IN PROJECT

A cutting-edge deep learning model called MobileNetV2 was created for effective picture classification, especially on mobile and embedded devices. This particular kind of convolutional neural network (CNN) makes use of depthwise separable convolutions, a method that lowers the number of parameters and processing cost in comparison to conventional CNN architectures, making it perfect for low-resource devices like smartphones. An inverted residual structure in MobileNetV2's design enhances information flow between layers, enabling speedier calculations without sacrificing accuracy. This enables the model to provide cutting-edge performance while remaining lightweight and effective. Human-crafted feature extraction is no longer necessary because MobileNetV2 automatically learns hierarchical features straight from the raw image data. Because of this, the model can effectively identify intricate patterns and spatial correlations in the images, which is crucial for image classification tasks. In contrast to more conventional models such as LightGBM, which necessitate a great deal of preprocessing and manual feature extraction, MobileNetV2 makes use of end-to-end learning, which enables it to adjust and enhance its performance as it handles more data. MobileNetV2 is especially well-suited for real-time applications on mobile devices, including on-site diagnostics for oral cancer diagnosis, where accuracy and speed are crucial, due to its lightweight design and great efficiency.

VII. SYSTEM ARCHITECTURE

The proposed oral cancer classification system using MobileNetV2 is designed to automatically detect and classify oral cavity images into benign and malignant categories using a deep learning approach. The system begins with image acquisition, where oral cavity images are obtained from a dataset and provided as input to the model. These images undergo a data preprocessing stage to ensure consistency and improve model performance. The preprocessing step includes resizing the images to a fixed input dimension compatible with the network, normalizing pixel intensity values to stabilize the learning process, and applying data augmentation techniques such as rotation, flipping, and scaling to enhance dataset variability and reduce overfitting. The preprocessed images are then fed into the MobileNetV2 convolutional neural network, which serves as the feature extraction backbone of the system. MobileNetV2 utilizes convolutional layers to capture low-level visual features such as edges and textures, followed by depthwise separable convolutions that significantly reduce computational cost while maintaining feature representation quality. In addition, the network incorporates inverted residual blocks, which expand feature representations and enable efficient learning of complex patterns related to oral lesions. These layers collectively perform feature extraction, allowing the model to learn discriminative representations such as abnormal tissue textures, colour variations, and lesion boundaries. The extracted feature maps are then passed to the classification layer, where a fully connected layer combined with a Softmax activation function generates probability scores for each class. Finally, the class with the highest probability is selected as the prediction, indicating whether the input oral image corresponds to a benign or malignant condition.

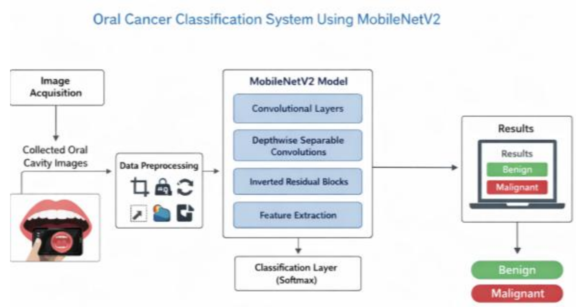

Fig: 1 SYSTEM ARCHITECTURE OF PROJECT

VIII. RESULTS

The performance of the proposed oral cancer classification system was evaluated using training and validation accuracy as well as loss metrics.

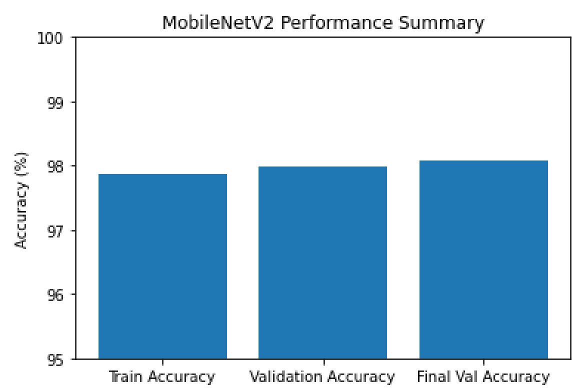

Figure 2: Model Accuracy

The bar graph (Fig 2) illustrates the comparison between training accuracy and validation accuracy obtained by the proposed deep learning model based on MobileNetV2. Training accuracy represents the model's ability to correctly classify oral lesion images during the learning phase, whereas validation accuracy reflects the model's performance on unseen validation data. As observed in the graph, the validation accuracy is higher than the training accuracy, indicating that the model generalizes well to new data and is not overfitting to the training dataset. This improvement can be attributed to effective preprocessing, transfer learning, and feature extraction capabilities of

MobileNetV2. The results demonstrate that the proposed system is capable of accurately distinguishing between different pre-cancerous stages of oral cancer from white light images.

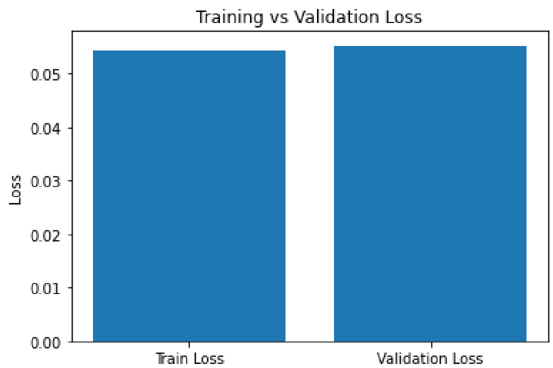

Fig 3: Training vs Validation loss

The bar chart (Fig 3) presents the comparison of training loss and validation loss during the model training process. Training loss indicates how well the model fits the training dataset, while validation loss measures the model's prediction error on unseen validation samples. From the graph, the validation loss is slightly higher than the training loss, which is a typical behavior in deep learning models and suggests stable learning without severe overfitting. The relatively low loss values demonstrate that the proposed MobileNetV2-based framework successfully minimizes classification errors during training and validation. This confirms that the model effectively learns discriminative features of oral lesions and improves the reliability of automated pre-cancerous stage classification.

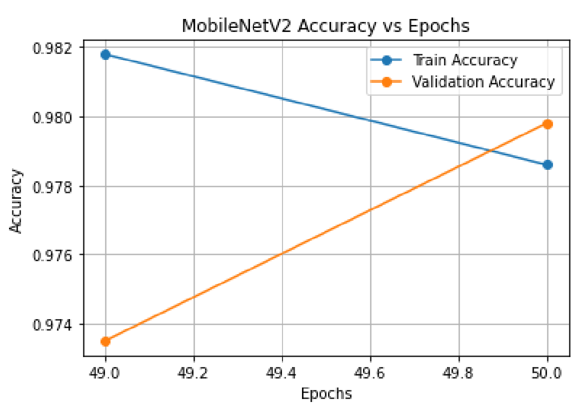

Fig 4: Accuracy vs Epochs

The line graph (Fig 4) shows the trend of training and validation accuracy across the training process. Training accuracy starts at a higher value and gradually decreases, while validation accuracy increases steadily. This behavior indicates that the model initially memorizes patterns from the training data but later adjusts its parameters to improve generalization performance. The increasing validation accuracy suggests that the model becomes more capable of identifying relevant features from unseen oral images as training progresses. The convergence between training and validation accuracy demonstrates that the MobileNetV2 architecture successfully learns meaningful representations for the classification of oral cancer stages, ensuring robust performance on real-world data.

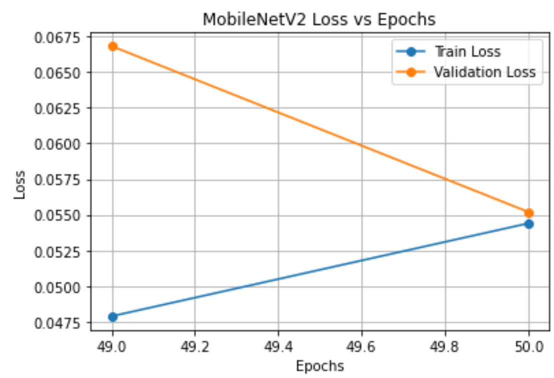

Fig 5: Loss vs Epochs

The loss curve graph (Fig 5) represents the variation of training loss and validation loss during the training process. Training loss increases slightly while validation loss decreases significantly as training progresses. This indicates that the model shifts from simple memorization of training samples to better generalization on validation data. The decreasing validation loss suggests that the proposed MobileNetV2-based classification framework improves its prediction capability over time by learning more relevant and discriminative features of oral lesions. The convergence of the two curves demonstrates stable training behavior and confirms that the model achieves an optimal balance between learning and generalization, which is essential for reliable medical image classification.

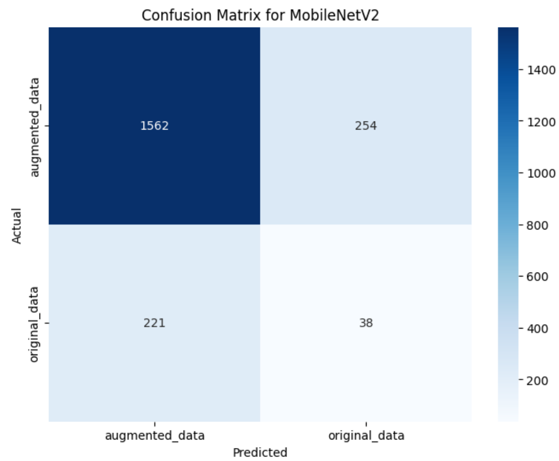

Fig 6: Confusion matrix

The confusion matrix (Fig 6) of the proposed MobileNetV2-based model evaluates the classification performance on a dataset of 2075 oral cancer images. The model correctly classified 1562 benign samples as benign (True Negatives) and 38 malignant samples as malignant (True Positives). However, 254 benign samples were misclassified as malignant (False Positives) and 221 malignant samples were misclassified as benign (False Negatives). The results indicate that the model effectively identifies a large number of non-cancerous cases while maintaining reasonable detection capability for malignant lesions. This performance demonstrates the effectiveness of the MobileNetV2 architecture in handling large-scale oral cancer image classification tasks.

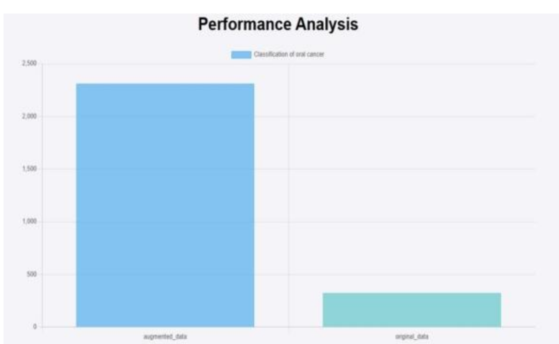

Fig 7: Performance Visualization

The chart (Fig 7) presents a comparison of classification performance using original data and augmented data in the proposed MobileNetV2-based oral cancer classification system. The model was trained using approximately 2300 augmented samples, which is significantly higher than the number of original samples available in the dataset. Data augmentation techniques increase the diversity of training samples by generating variations of existing images, thereby improving the representation of different patterns and features within each class.

The increased number of augmented samples helps in addressing class imbalance and enhances the model's ability to generalize better during training. As a result, the training process becomes more stable and leads to improved classification accuracy. The noticeable performance gap between models trained on original data and augmented data clearly demonstrates the effectiveness of augmentation techniques in improving model learning. Overall, the results confirm that the proposed MobileNetV2-based framework significantly benefits from augmented data, leading to enhanced classification performance for detecting precancerous stages of oral cancer.

VIII. FUTURE ENHANCEMENT

Even while the existing system performs admirably, there are a number of possible future improvements that could increase its precision and usability even more. One improvement would be the incorporation of sophisticated methods such as transfer learning to refine the model on smaller, domain-specific datasets, hence increasing its capacity to manage uncommon cases. The dataset should also be expanded to include more varied photos that represent different demographic groups and stages of oral lesions. The system would be better able to generalize across various demographics as a result. Furthermore, adding real-time video analysis could be a useful improvement, enabling the system to identify lesions during medical exams in dynamic settings. Implementing explainable AI approaches is another possible improvement that could increase adoption and confidence by enabling medical practitioners to comprehend the reasoning behind the model's predictions. Ultimately, the tool may become more widely available worldwide if the model is optimized for even faster processing on mobile devices and in low-resource environments, particularly in developing nations with little access to cutting-edge healthcare infrastructure.

IX. CONCLUSION

To sum up, this project uses the MobileNetV2 architecture, which offers a portable and effective method for real-time analysis on mobile devices, to present a novel methodology for the early identification and categorization of oral cavity lesions. The technology can significantly help healthcare workers make faster and more accurate diagnoses by automating the process of classifying oral lesions into benign, malignant, and pre-cancerous phases. This would ultimately help to lower the mortality rates from mouth cancer. Future developments could increase the system's functionality and make it an even more useful instrument in international healthcare. Examples of these developments include enlarging the dataset, introducing transfer learning, and providing real-time video analysis.

REFERENCES

[1] S. Sarkar, G. Horn, K. Moulton, A. Oza, S. Byler, S. Kokolus, and M. Longacre, ‘‘Cancer development, progression, and therapy: An epigenetic overview,’’ Int. J. Mol. Sci., vol. 14, no. 10, pp. 21087–21113, Oct. 2013.

[2] P. H. Monter and S. G. Patel, ‘‘Cancer of the oral cavity,’’ Surgical Oncol. Clinics North Amer., vol. 24, no. 3, pp. 491–508, Apr. 2015.

[3] S. Warnakulasuriya, N. W. Johnson, and I. Van Der Waal, ‘‘Nomenclature and classification of potentially malignant disorders of the oral mucosa,’’ J. Oral Pathol. Med., vol. 36, no. 10, pp. 575–580, Nov. 2007.

[4] B. W. Neville and T. A. Day, ‘‘Oral cancer and precancerous lesions,’’ CA, A Cancer J. Clinicians, vol. 52, no. 4, pp. 195–215, Jul. 2002.

[5] J. Bagan, G. Sarrion, and Y. Jimenez, ‘‘Oral cancer: Clinical features,’’ Oral Oncol., vol. 46, no. 6, pp. 414–417, Jun. 2010.

[6] S. Sudha, A. Veluthattil, S. Kandasamy, and S. Chakkalakkoombil, ‘‘Effect of hypo fractionated, palliative radiotherapy on quality of life in late-stage oral cavity cancer: A prospective clinical trial,’’ Indian J. Palliative Care, vol. 25, no. 3, p. 383, 2019.

[7] U. Mangalath, S. Aslam, A. H. Abdul Khadar, P. Francis, M. K. Mikacha, and J. Kalathingal, ‘‘Recent trends in prevention of oral cancer,’’ J. Int. Soc. Preventive Community Dentistry, vol. 4, no. 6, p. 131, 2014.

[8] I. van der Waal, R. de Bree, R. Brakenhoff, and J. Coebergh, ‘‘Early diagnosis in primary oral cancer: Is it possible?’’ Medicina Oral Patología Oral y Cirugia Bucal, vol. 16, no. 3, pp. e300–e305, 2011.

[9] J. Seoane, B. Takkouche, P. Varela-Centelles, I. Tomás, and J. M. Seoane-Romero, ‘‘Impact of delay in diagnosis on survival to head and neck carcinomas: A systematic review with meta-analysis,’’ Clin. Otolaryngology, vol. 37, no. 2, pp. 99–106, Apr. 2012.

[10] M. A. Weinberg and D. J. Estefan, ‘‘Assessing oral malignancies,’’ Amer. Family Physician, vol. 65, no. 7, pp. 84–1379, Apr. 2002.

[11] A. Humeau-Heurtier, ‘‘Texture feature extraction methods: A survey,’’ IEEE Access, vol. 7, pp. 8975–9000, 2019.

[12] T. Song, J. Feng, S. Wang, and Y. Xie, ‘‘Spatially weighted order binary pattern for colour texture classification,’’ Expert Syst. Appl., vol. 147, Jun. 2020, Art. no. 113167.

[13] L. Liu, J. Chen, P. Fieguth, G. Zhao, R. Chellappa, and M. Pietikäinen, ‘‘From BoW to CNN: Two decades of texture representation for texture classification,’’ Int. J. Comput. Vis., vol. 127, no. 1, pp. 74–109, Jan. 2019.

[14] L. Liu, P. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, ‘‘Local binary features for texture classification: Taxonomy and experimental study,’’ Pattern Recognit., vol. 62, pp. 135–160, Feb. 2017.

[15] T. P. Nguyen, N.-S. Vu, and A. Manzanera, ‘‘Statistical binary patterns for rotational invariant texture classification,’’ Neurocomputing, vol. 173, pp. 1565–1577, Jan. 2016.

[16] X. Qi, G. Zhao, L. Shen, Q. Li, and M. Pietikäinen, ‘‘LOAD: Local orientation adaptive descriptor for texture and material classification,’’ Neurocomputing, vol. 184, pp. 28–35, Apr. 2016.