Journal of Global Research in Electronics and Communication
Volume 1, No. 8, August 2025
Available Online at: www.jgrec.info
RESEARCH PAPER

A Hybrid Deep Learning Method for Early Recognition of Oral Malignancy Cell Carcinoma in Image Data

Sandeep Gupta
SATI, Vidisha
Sandeepguptabashu@gmail.com

***Abstract*—The seventh most prevalent cancer in the world is oral squamous cell carcinoma (OSCC), often diagnosed through histopathological images, which require expert interpretation due to tumor heterogeneity. This work tackles the urgent need for automated and precise early identification of OSCC using histopathology images. Utilizing a publicly accessible dataset of 5192 biopsy images (2494 normal, 2698 malignant OSCC) captured at 100x magnification, the research proposes a novel hybrid deep learning (DL) framework. The methodology involves comprehensive data preprocessing, including resizing images to 224x224x3, followed by feature extraction and dimensionality reduction using PCA. To combat overfitting and class imbalance, data augmentation, notably SMOTE applied. A Support Vector Machine (SVM) classifier and a ResNet-18 model for robust feature learning form the system's foundation. With a remarkable accuracy of 98.1%, the hybrid model also had good precision (98.22%), recall (98.61%), and F1-score (98.44%). These results highlight the substantial potential of the ResNet-18+SVM model to serve as a reliable and efficient computer-aided diagnostic tool for OSCC, contributing to improved patient prognosis through timely intervention.**

Keywords*—Oral Squamous Cell Carcinoma (OSCC), Histopathological Images, Image Classification, SMOTE, Adaptive Harmony Search (AHA), Early Cancer Detection, Medical Image Processing, Cancer Diagnosis.

I. INTRODUCTION

The development of OSCC can be effectively halted by early identification and prompt treatment intervention, resulting in a 80% increase in patient survival [1]. A clinician's examination is necessary for the diagnosis of OSCC. Some OSCC symptoms, however, might be confused with those of precancerous lesions or mouth ulcers. This occurrence implies that a great deal of expertise is necessary for early OSCC screening [2][3]. Pathological diagnosis, the gold standard for conclusive OSCC diagnosis, is intrusive, painful, and results in poor wound healing.

The unregulated proliferation of abnormal cells in mouth cavity is a hallmark of advanced oral cancer. Numerous genetic defects that accumulate inside the cells and damage their surface cause oral squamous cell carcinoma (OSCC), the most common kind of oral cancer that begins in the oral cavity. Even so, it starts to show up in the oral epithelium. Consequently, the size and shape of these cells and the related nucleus alter [4][5][6]. In particular, OSCC may be divided into three classes: well-differentiated OSCC, moderately well-differentiated OSCC, and weakly differentiated OSCC [7]. OSCC is the seventh most frequent kind of cancer globally, affecting 657,000 individuals year and resulting in over 330,000 deaths, according to statistics from the Public Health Organization. Human papillomavirus (HPV) infection, poor oral hygiene, alcohol and tobacco use, ethnicity, geography, and family history are some of the risk factors linked to OSCC [8]. The risk of OSCC stems from the lack of particular clinical vital signs that would enable professionals to make an accurate OSCC prediction.

In head and neck pathology, oral and oropharyngeal carcinomas are a significant issue because of their prevalence and the grave repercussions that continue to follow their involvement [9][10]. According to data from the most pertinent organizations and institutions, these neoplasms cause 177,757 and 48,143 fatalities annually and 377,713 and 98,412 new cases globally, respectively (GLOBOCAN, IARC, WHO) [11].

The majority of oral cancers, ranging from 84% to 97%, are squamous cell carcinomas (OSCCs), which can develop from either highly malignant lesions or normal epithelial linings [12][13]. Proliferative verrucous leukoplakia (PVL), fibrosis, erythroplakia, leukoplakia, candidal leukoplakia, congenital dyskeratosis, lichen planus, and an inflammatory oral submucosa are preclinical signs of OC. These conditions are regarded as potential malignant disorders. Potentially malignant oral lesions, or PMOLs, often precede OSCC and may be the focus of early diagnosis.

More important details on the local immune response, tumour cell morphology, stromal component, and necrosis—all of which may have significant prognostic significance—can be obtained by histopathological analysis. These histological characteristics are now employed in clinical practice after demonstrating their predictive significance in various cancer types. Squamous cell carcinoma (SCC) is the second most prevalent kind of cancer worldwide [14][15]. It originates in the epithelium's squamous layer. The majority of SCCs are caused by the mucosal layers and the skin of the head and neck. The head and neck regions account for almost 70% of cutaneous SCC cases.

The requirement for time-consuming feature engineering is eliminated by DL, an advanced subset of ML, which is excellent at extracting high-dimensional features from data [16][17]. It provides improved performance in a range of applications, outperforming statistical, physical, and conventional ML techniques. As a result, DL is being used more and more to identify oral squamous cell carcinoma (OSCC) early [18]. Models like RNN and CNN are widely used because of how well they analyze complicated time series data.

A. Motivation and Contribution of Paper

This study is being motivated by the pressing need to get the right and early oral squamous cell carcinoma diagnosis. OSCC is a serious medical condition, and patient survival and effective treatment depend on early discovery. Conventional diagnostic procedures, which involve manual slide review of the biopsies by the pathologist, may be slow and have inter-observer differences. Thus, it is of great interest to design automated, predictable, and efficient tools that may be utilized to preliminarily diagnose OSCC based on the histopathological photos. This study seeks to combine the ability to improve the sensitivity and effectiveness of OSCC detection using a hybrid ResNet-18+SVM model, a state-of-the-art DL method, which in turn has the potential to assist the work of medical workers by allowing them to deliver timely and accurate diagnoses. The study's primary contributions are outlined below:

- The study begins with the collection of an OSCC dataset.
- Employed a painstaking pre-processing including image processing, classification, segmentation, and size reduction to condition the data to analyze.
- The structure integrates the feature extraction into finding the patterns that are relevant to the data, and the data augmentation approaches like SMOTE which solve the category imbalance issue to improve the representation of the minority classes and develop better generalization of the model.
- The main achievement includes the creation and use of a ResNet+SVM model to distinguish OSCC that made use of the architectural superiority of the model to learn features and represent them hierarchically.
- The implementation assesses the efficacy of the suggested model using crucial measures such as F1-Score, Accuracy, Precision, and Recall, which provide a comprehensive examination of the model's diagnostic capabilities.

B. Novelty & Justification of the Study

The presented study uses the ResNet+SVM architecture exclusively to propose a new, comprehensive DL framework designed to identify OSCC in its early stages. Their main innovation is the implementation of the systematic and integrated manner addressing the key challenges in the medical image analysis, This starts with the data's thorough pre-treatment, i.e. the use of the image enhancement, classification, segmentation, size reduction, and augmentation techniques, as well as the utilization of SMOTE to address the points of the class imbalance. The quality and the representativeness of input data provided in this holistic pipeline of pre-processing are essential due to imprecision and diversity of medical imaging datasets. Selection of ResNet+SVM is quite reasonable, since ResNet is very efficient in addressing the vanishing gradient issue in the neural networks and allows reuse of the features due to residual connections and has less parameters hence applicable in small or imbalanced datasets. The SVM classifier also enhances the discriminative power of the model, thus giving a better classification. Judging overall, this framework shows a good performance in comparison with traditional models and bridges a big gap by providing An accurate, dependable, and efficient response to automated OSCC diagnostic.

C. Structure of Paper

The remnant of this investigation are arranged as follows: In section II, A review of a related study on the early identification of OSCC. In section III, The dataset's description and the suggested technique are provided. section IV reports and discusses. Section V is a conclusion that also outlines restrictions and potential paths.

II. Literature Review

The literature review section provides a detailed overview of recent studies conducted to explore the possibility of early detection of OSCC. Table I presents key aspects of the reviewed studies, including the methodologies employed, the performance measures obtained, the significant findings, the research limitations, and the recommended areas for future research.

Kumar and Nelson (2025) work has used a dataset that comprises 5,192 images, divided into OSCC and Normal classes. This dataset is used to train and evaluate the EfficientNetB3 model, with 70:15:15 separating training, validation, and testing. Data augmentation techniques are used to provide key metrics including accuracy, precision, recall, and F1-score, which are used to evaluate the model's efficacy. As demonstrated by its accuracy, recall, and F1-score of 0.98, the model correctly categorized OSCC and healthy tissues in the test set, and the final result showed 98% accuracy. These findings suggest that the newly implemented process offers a workable substitute for early, automated OSCC identification, which enhances patient outcomes and radiological diagnosing of cancer in a clinic facility [19].

Kaur, Sharma and Gupta (2024) developed a paradigm for image classification that uses OSCC in histopathology pictures using a special convolutional neural network (CNN). The 4,946 images used to train the dataset consisted of normal and OSCC classes, as well as 126 test and 120 validation images. The model's resistance was increased by employing several data augmentation techniques, including zooming, rescaling, shearing, and horizontal flipping. The Adam optimizer and a category cross-entropy loss function were used to fine-tune the CNN model. It consists of many convolutional, pooling, and dropout layers. The accuracy of the model after 30 epochs was 86.67% for validation and 94.64% for training. The model yielded an overall accuracy of 72%, a recall of 92%, and a precision of 76% following further testing on the test set [20].

Kaur and Sharma (2024) work attempts to use transfer learning models, namely EfficientNetB3 and DenseNet201, to identify Oral Squamous Cell Carcinoma (OSCC) using histopathology pictures. There are 5,192 histological pictures in the dataset. From the dataset, the DenseNet201 model split a 120-image test set, a 126-image validation set, and a 4,946-image training set. The EfficientNetB3 model likewise used 779 photographs for testing, 779 images for validation, and 3,634 shots for training. The model parameters were adjusted using the Adam optimizer following 50 epochs of training with 32 and 128 batch sizes, respectively. The DenseNet201 model had high accuracy in distinguishing between malignant and non-malignant tissue samples, as evidenced by its test accuracy of 91.27, validation accuracy of 99.083, and Area Under the Curve (AUC) score of 86.6. The EfficientNetB3 model shown remarkable performance with training loss of 50.42, training accuracy of 94.51, validation loss of 50.30,

validation accuracy of 0.9474, test loss of 68.93, and test accuracy of 92.17 [21].

Benagi et al. (2024) work was done in collaboration with Ramaiah Institute of Technology and Ramaiah Dental College to enhance OSCC detection using pipeline-based computer-assisted screening methodology that involves classification, feature detection and risk stratification. In the first experiment, it was found that the classification module with Efficient Net provided an accuracy of 99.25 percent during the process of classifying cancerous and non-cancerous lesions. Using YOLO-NAS, the feature detection module was able to determine some important signs of aggressiveness, such as keratin pearls and appearance of multiple nucleoli, whereas detection of vascular invasion was performed by the Efficient Net classifier [22].

Devindi et al. (2024) in order to close this gap, the study proposes a multimodal DL pipeline that uses a range of data, such as patient information, and approximates the diagnostic approach used by physicians to identify oral cancer early on. The most recent image encoders are used in the study to classify oral lesions as either benign or possibly malignant cells. Inception_v3, MobileNetV3-Large, Mix Net-S, ResNet-50, HRNet-W18-C, and DenseNet-121 are the six pre-trained DL models whose performance is compared. The performance of the proposed pipeline, which uses the MobileNetV3-Large as the image encoder, is 81% overall, with 79% precision, 79% recall, 78% F1-score, and a 0.57 Matthews Correlation Coefficient (MCC) [23].

Ramya, Minu and Magesh (2023) study suggests a novel method for OC classification by applying the 2D-ICNN technology. In order to train the model, their approach focusses on identifying the best characteristics from OSCC biopsy pictures. Utilize GLCM to accomplish effective feature extraction. The suggested model obtains a very high classification accuracy of 98.29%, according to their simulation findings. It is crucial to identify OC as soon as possible, especially in its early phases, in order to enable the best possible care and appropriate treatments. There is potential for bettering OC diagnosis and treatment with the use of DL algorithms, such as their suggested model [24].

Blessy and Sornam (2022) suggested that because of their tiny size, nanoparticles can effectively prevent cancer from spreading to other bodily areas. One of the most crucial stages of treating oral squamous cell carcinoma is accurate identification. Several ML techniques, including SVM, CNN, and NB, are used to extract features for OSCC classification. DL has shown remarkable performance in early-stage cancer diagnosis with a large dataset, producing high-accuracy results in OSCC early-stage identification. It aims to identify mouth cancer faster and more accurately. Early identification may help prevent future fatalities from mouth cancer. Among these algorithms, CNN has been refined in each study and achieved up to 96.6% accuracy [25].

Subhija and Reju (2022) approach, Haar wavelets are used to break down the oral squamous cell cancer and normal histology images into three detail coefficients and one approximation. The high-frequency and low-frequency components' characteristics are then extracted using the Grey Level Co-occurrence Matrix. A second tree classifier is used to choose the pertinent features once the features have been fused to create a feature vector. They used RF, KNN, and voting classifiers to arrive at the final binary categorization of normal and malignant. ROC curves, sensitivity, AUC values of characteristic curves, specificity, F1 score, quality index, error rate, accuracy, precision, confusion matrix, and other efficacy metrics are used to assess algorithm's potentiality. RF has highest accuracy of 97.59% and the highest precision of 95.98% among the classifiers. The suggested approach to oral cancer detection is accurate and efficient. For this reason, it can serve as an accurate and dependable support tool for oral pathologists [26].

TABLE I. SUMMARY OF LITERATURE OVERVIEW AND REVIEW ON EARLY DETECTION OF ORAL SQUAMOUS CELL CARCINOMA

Author	Methodology	Dataset	Key Findings	Challenges	Future Approach
Kumar & Nelson (2025)	EfficientNetB3 with 70:15:15 data split and augmentation	5,192 OSCC & Normal images	F1-score: 0.98, Accuracy: 98%, Precision: 0.98, Recall: 0.98	Generalizability to diverse datasets not discussed	Enhance real-time deployment and integrate clinical metadata
Kaur, Sharma & Gupta (2024)	Custom CNN with augmentation (rescaling, shearing, zooming, flipping)	4,946 train, 120 val, 126 test (OSCC & Normal)	Accuracy: 72%, Precision: 76%, Recall: 92%, Training Accuracy: 94.64%	Low test accuracy despite high training/validation accuracy	Explore deeper models and advanced regularization techniques
Kaur & Sharma (2024)	Transfer Learning with EfficientNetB3 and DenseNet201, Adam optimizer	5,192 images split into train/val/test	EfficientNetB3: Test Accuracy 92.17%; DenseNet201: Test Accuracy 91.27%, AUC: 86.6	High training loss for DenseNet201; needs better optimization	Optimize learning rates and test across varied histological sources
Benagi et al. (2024)	EfficientNet + YOLO-NAS pipeline for classification, feature detection, risk stratification	Collaborative clinical dataset	Classification Accuracy: 99.25%, successful detection of keratin pearls, nucleoli	Complexity in integrating multiple modules	Apply pipeline to multi-center datasets and deploy in clinics
Devindi et al. (2024)	Multimodal DL pipeline with metadata + six pre-trained models (MobileNetV3 best)	Diverse sources including image and patient metadata	Accuracy: 81%, Precision/Recall: 79%, F1-score: 78%, MCC: 0.57	Integration of multimodal data poses complexity	Enhance fusion mechanisms and address metadata inconsistencies
Ramya, Minu & Magesh (2023)	2D-ICNN with GLCM-based feature extraction	Biopsy images	Accuracy: 98.29%, high feature discrimination	Details on dataset size/split not clear	Integrate temporal biopsy changes and hybrid models
Blessy and Sornam (2022)	CNN, SVM, Naïve Bayes; focus on nanoparticles in diagnosis	Large dataset	CNN achieved accuracy up to 96.6%	Limited comparison metrics	Explore hybrid ML-DL models and real-time detection with nanoparticles
Subhija and Reju (2022)	Haar wavelet decomposition + GLCM + Extra Trees + RF/KNN/Voting	Histopathology images (Normal vs OSCC)	RF: Accuracy 97.59%, Precision 95.98%; Effective binary classification	Preprocessing and wavelet tuning can be intensive	Combine with DL models and explore ensemble stacking

III. Methodology

To enhance patient outcomes and survival rates, early detection of OSCC is crucial. The provided image illustrates a workflow for classifying an OSCC dataset using a ResNet+SVM hybrid model. The process begins with collecting the OSCC dataset, followed by a data pre-processing phase that includes image processing, classification, segmentation, and size reduction. Feature extraction is then carried out, which results in data augmentation. Using the SMOTE, class imbalance is addressed. The data is then segregated into training (80%) and testing (20%) sets and classified using a hybrid model of ResNet+SVM. To assess the model's efficacy and determine outcome, key metrics including precision, accuracy, recall, and F1-score are employed. Figure 1 shows a flowchart for the early detection of OSCC.

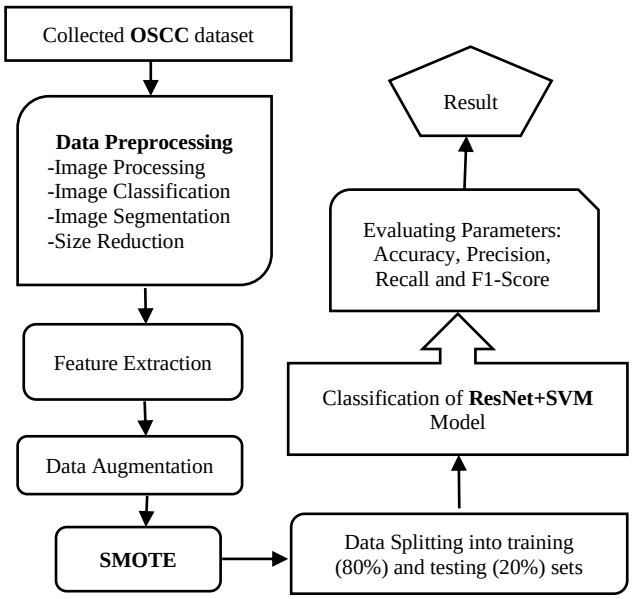

Fig. 1. Flowchart for Early Detection of OSCC

Here is a comprehensive, step-by-step breakdown of the processes illustrated in the flowchart:

A. Data Collection

This study's systems were assessed utilising a publicly accessible collection of OSCC histopathology imaging data. Biopsy slides were gathered from two reputable medical facilities. There are two binary classifications in this oral cancer dataset: OSCC and normal. The Leica DM750 microscope was utilized to take pictures with a camera type ICC50 HD at 100 (10 objective lens 10 eyepiece) magnifications (size 2048 1536 pixels). The collection included 5192 photos taken at a 100x magnification from biopsy slides. Under local anaesthesia, a biopsy was performed to capture the pictures used in this data collection. These biopsies were used by a pathologist to make the diagnosis. Of the 5192 histological pictures, 2698 (52%) were identified as malignant instances of OSCC and 2494 (48%) as normal. These pictures were the main focus of investigation. The heatmap is given below in this section:

Figure 2 illustrates a feature correlation heatmap, which visualizes the pairwise correlation between different extracted features from the pre-processed oral images. The correlation coefficient between two characteristics is represented by each cell in the heatmap, and the color scale shows the association's intensity and direction. White indicates a weak or non-existent link, blue a significant positive correlation and a large negative correlation. The hierarchical clustering on both axes groups similar features together based on their correlation patterns, helping to identify clusters of highly interrelated features.

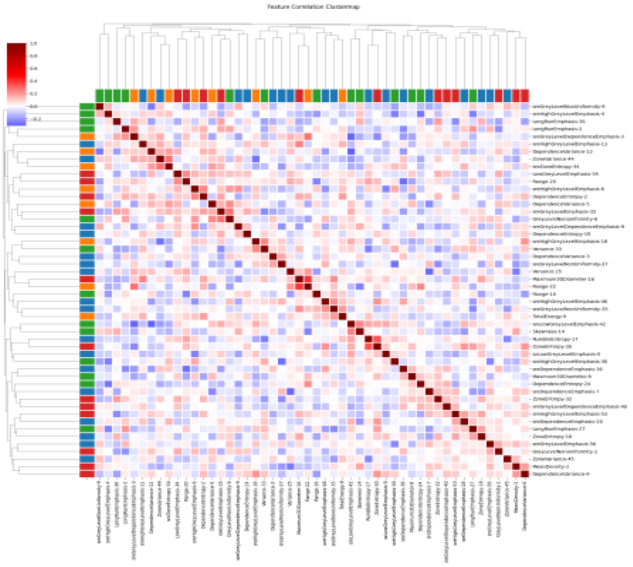

Fig. 2. Correlation Features Heatmap

B. Data Preparation

The data preparation stage of the data analysis and DL process is crucial, which entails converting unstructured data into a clean, organized format appropriate for modelling. To significantly enhance data quality, the model must perform better and produce more dependable results, the research includes image processing, image classification, image segmentation and size reduction. Stages of pre-processing pipeline are as follows:

- **Image Processing:** The collected images exhibit variability in terms of sizes, zoom levels, angles, lighting conditions, and orientations. In particular, they were cropped by hand and resized with the dimensions 224 224 3 and changed to the JPG file format [27]. This pre-processing makes it possible to have uniformity in the data structure of all the pictures' input data, making it compatible with the preferred DL model architecture.
- **Image Classification:** The model assigns appropriate class labels to images throughout the categorization phase [28]. A stack of backpropagation (BP) and the probabilistic generation model is composed of neural networks with restricted Boltzmann machines (RBMs). It includes the output, hidden, and visible layers. During the learning process, several hidden layers pass the features via the input/visible layer, located at the model's end.
- **Image Segmentation:** The development of digital diagnostic, prognostic, and therapeutic tools heavily relies on the automatic segmentation of digital histology images into sections that correspond to various tissue types [29]. Digital pictures are processed using the segmentation approach, which groups pixels with similar colorimetric characteristics in regions that most likely belong to items of interest.

- **Size Reduction:** At this stage, pooling is used to achieve the grid size reduction, and then the convolution operation is done when further convolutional layers are added.

C. Feature Extraction

An essential stage in image-based illness detection is feature extraction, involving the identification of meaningful characteristics from pre-processed oral images. In this study, the most informative parts of the greyscale pictures were preserved while dimensionality was reduced using Principal Component Analysis (PCA), a feature reduction approach. The primary components, or axes, of the new coordinate system created by PCA are organized in decreasing variance order [30]. A cumulative variance limit is used to keep only the most influential parts so as to ensure that key features are not lost and unnecessary data is thrown away. This change helps to reduce the computational complexity and overfitting risk while preparing the model. Training and testing of the classification algorithm that is to be used in automatic disease detection are carried out using the extracted PCA-based features as the inputs. The model's applicability is assessed through the use of the 3-fold cross-validation technique and resilience and demonstrating once more how well PCA works to enhance the computer-aided diagnostic system's performance.

D. Handling Class Imbalance using SMOTE

A data-augmentation method, SMOTE, resamples the data, since it uses the neighborhood of a neighborhood sample as a point of reference and it generates a new example associated with this neighborhood [31]. In SMOTE, k = 5 is used when creating the synthetic samples. This was determined empirically based on the number of original samples so as to strike a balance between valuable information and to achieve a better sample diversity.

E. Data Splitting

The critical stage of DL is data splitting, which means that Training and testing portions of the data are separated, typically accounting for 80% and 20% of the total. The model can learn using the training data and be tested on unseen data, which helps avoid overfitting and guarantees good performance in practice under real-life conditions.

F. Classification Algorithm of ResNet-18+SVM

Three blocks make up ResNet-18: a fully connected layer, a 3 x 3 filter size max pooling layer, and 17 convolution layers. The residual building component The residual building component is the basis of the ResNet-18 network [32]. The result may then be obtained by simply concatenating the input and output vectors via the convolutional layer using the rectified linear unit (ReLU) activation function. The issue of a DNN disappearing and expanding gradient may likely be resolved in several methods, including this one. Equation (1) uses the input and output, the residual function F, x, and y, to improve performance on subjects or tasks that are similar yet unrelated.

$$y = F(x) + x \tag{1}$$

The supervised ML technique known as a SVM is used to solve regression and classification issues. To divide the data into discrete groups, A decision surface, or N-dimensional hyperplane, is produced using the SVM algorithm. The dataset's number of features is its N dimension. The algorithm optimize the margin to achieve this, the separation of the data points from the hyperplane. The ideal hyperplane that divides the data is sought, and the distance is often optimized on the hinge loss. The data points that have the most influence on the location of the hyperplane are the support vectors since they are the ones closest to the decision boundary.

In high-dimensional spaces, support vector machines operate effectively. Even when the dimensions are larger than the samples, the method still performs well. To find the cost function, hinge loss Equation (2) is utilized. To get the best fit model the cost function is minimized.

$$c(x, y, f(x) = (1 - y * f(x))_+ \tag{2}$$

There are several kernel functions that may be used to alter the final input data $f(x)$. The conversion of non-separable information to separable information can be aided by these kernel functions.

G. Evaluation Parameters

The confusion measures were recall, accuracy, precision, F1-score, and the loss function. These measurements crucial for assessing how effectively the model captures the connections and trends seen in the data. The evaluation was conducted using OSCC data to investigate the reliability and robustness of the model. A thorough rundown of the confusion matrix parameters given in the next section:

- **True Positives (TP):** Instances where OSCC is accurately predicted by the model.
- **True Negatives (TN):** Instances where OSCC is accurately predicted by the model.
- **False Positives (FP):** Instances where a healthy patient's OSCC is mis predicted by the model.
- **False Negatives (FN):** Situations in which a patient with OSCC is not identified by the model.

The assessment of individual classes was done by calculating the standard classification formulae:

1) Accuracy

The model's accuracy indicates how frequently, out of all predictions, it accurately detects both malignant and healthy instances. It is more formally defined as in Equation (3):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

2) Precision

Precision indicates how many of the patients predicted to have OSCC truly have the disease. High precision reduces false positives, and precision is determined using the following Equation (4):

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

3) Recall

The model's potentiality to identify real cancer cases is measured by recall, which is essential for reducing missed diagnoses. Mathematically, define it as in Equation (5):

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

4) F_1-Score

The F1-score is useful when there is a difference in the distribution of classes; it is biased towards both the FP and the FN and is calculated as shown in Equation (6) below:

$$F - Score = 2 \times \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

IV. Results Analysis and Discussions

The PyTorch language is used to build the experiment of this article and is written in Python 3.10, and the experiment environment is as follows: Operating System Windows 10, GPU, NVIDIA RTX 2080, DL framework PyTorch. The effectiveness of the hybrid model of ResNet-18 + SVM for early detection of OSCC is demonstrated in Table II. The table also includes key measurement metrics, precision, accuracy, recall, and F1-score, which are used to categorize it. Each of these measures shows how reliable the model is in detecting OSCC situations and lowering FP and FN. The findings prove how the joint approach of deep feature extraction and an SVM classifier can produce high-level predictive performance of OSCC detection.

TABLE II. Evaluation of RESNET-18+SVM Results for Early Detection of Oral Squamous Cell Carcinoma

Metrics	ResNet-18+SVM
Accuracy	98.1
Precision	98.22
Recall	98.61
F1-Score	98.44

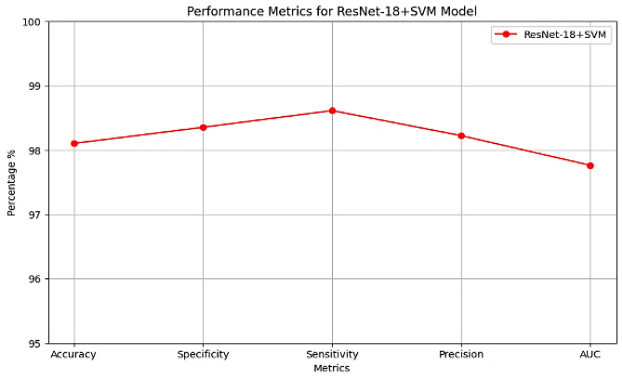

Fig. 3. Percentage graph of Hybrid Model

As Figure 3 shows, the ResNet-18+SVM model graph (the percentage model) indicates that the results are always high in all five evaluation criteria. The model had 98.1% Accuracy, 98.35% Specificity, 98.61% Sensitivity, 98.22% Precision and an 97.76 AUC values which imply good and balanced classification performance.

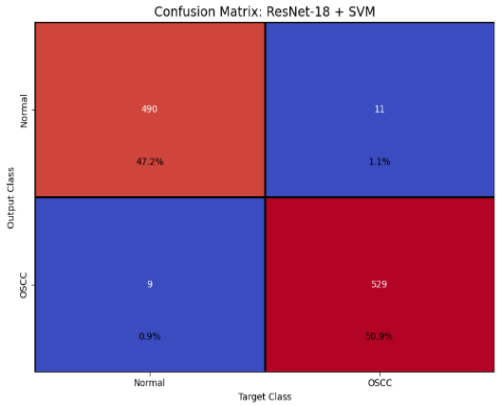

Fig. 4. Confusion Matrix of Hybrid Model

The hybrid ResNet-18+SVM model's confusion matrix for identifying normal or OSCC in oral histopathology images is displayed in Figure 4. The matrix shows a good performance of the model because the TP and TN are high showing that the model performs well by correctly classifying cancerous and non-cancerous cases. Low FP and FN indicate that the model is very accurate in reducing misclassification that is of the essence during clinical diagnosis.

A. *Comparative Analysis*

This part gives a clear comparative study of the proposed models ResNet-18+SVM against other well-known DL models that include Multilayer Perceptron (MLP) [33], Extreme Inception (Xception) [34], Artificial Neural Network (ANN) [35], and Visual Geometry Group 16-layer model (VGG16) [36]. To ensure consistency and fairness, in all the models, the same experiment conditions were applied to train and test the provided OSCC dataset. Accuracy performance of each model was summarized in Table III, and the best one was ResNet-18+SVM which showed accuracy of 98.1%. Comparatively, MLP model attained 90.83%, Xception scored 87.77%, ANN scored 78.95% and VGG16 scored the lowest 74%. These results form a strong testimony of the increased effectiveness of the ResNet-18+SVM model in The correct way to classify OSCC.

TABLE III. Existing DL Model's Comparison for Early Detection of OSCC

Models	Accuracy
MLP	90.83
Xception	87.77
ANN	78.95
VGG16	74
ResNet-18+SVM	98.1

The provided ResNet-18 +SVM model is much effective in the early identification of OSCC due to ResNet-18's potent deep feature extraction and SVM's accurate classification capabilities. This hybrid model contributes to a considerable increase in the accuracy of diagnostics, with a balanced score due to a low rate of FP and FN. The model is highly robust to overfitting and can generalize very well across histopathological data, which makes it a dominant tool to clinically apply to the real world. The ResNet-18+SVM model is an effective option to facilitate the timely and accurate diagnosis of OSCC and finally assist pathologists to intervene and plan treatment early.

V. Conclusion and Future Work

To enhance patient outcomes and survival, OSCC must be identified at an early stage. The use of salivary proteomics in the oral cancer research field is fast emerging and very dynamic in identifying and utilizing new and developed biomarkers in the early detection and prognostication procedures such as the metabolic, proteomics, genomics, and associated bioinformatics procedures. This paper has managed to propose a very efficient hybrid DL model, namely, ResNet-18+SVM to attain early identification of histological pictures of OSCC with improved performance metrics, which makes this model a potential useful automated diagnostic model. In the future, efforts need to be made to improve the utility and reliability of such systems to clinical practice. In order to boost medical professionals' confidence, explainable AI (XAI) will be taken into consideration to provide justifications for the decisions made by the models. Also, possible future studies could look at the incorporation of multi-modal data, and a mix of histopathology with other data about a patient or imaging modalities to form more integrated diagnostic systems. Additionally, it's critical to evaluate the model on bigger, more varied, and multi-centric datasets to confirm its generalizability and then implement a rigorous

prospective study in actual clinical practice to prepare the way to its actual implementation and broad contribution to the treatment of patients and early cancer detection.

REFERENCES

[1] M. G. Madhura, R. S. Rao, S. Patil, H. N. Fageeh, A. Alhazmi, and K. H. Awan, "Advanced diagnostic aids for oral cancer," *Disease-a-Month*, vol. 66, no. 12, Dec. 2020, doi: 10.1016/j.disamonth.2020.101034.

[2] C. Lajolo *et al.*, "Patient perception after oral biopsies: an observational outpatient study," *Clin. Oral Investig.*, vol. 25, no. 10, pp. 5687–5697, Oct. 2021, doi: 10.1007/s00784-021-03870-3.

[3] V. K. Singh, D. Pathak, and P. Gupta, "Integrating Artificial Intelligence and Machine Learning into Healthcare ERP Systems: A Framework for Oracle Cloud and Beyond," *ESP J. Eng. Technol. Adv.*, vol. 3, no. 2, pp. 171–178, 2023, doi: 10.56472/25832646/JETA-V3I6P114.

[4] S. M. Fati, E. M. Senan, and Y. Javed, "Early Diagnosis of Oral Squamous Cell Carcinoma Based on Histopathological Images Using Deep and Hybrid Learning Approaches," *Diagnostics*, vol. 12, no. 8, Aug. 2022, doi: 10.3390/diagnostics12081899.

[5] S. Singamsetty, "Neurofusion: Advancing Alzheimer's Diagnosis with Deep Learning and Multimodal Feature Integration," *Int. J. Educ. Appl. Sci. Res.*, vol. 8, no. 1, pp. 23–32, 2021.

[6] S. B. Shah, "Artificial Intelligence (AI) for Brain Tumor Detection: Automating MRI Image Analysis for Enhanced Accuracy," *Int. J. Curr. Eng. Technol.*, vol. 14, no. 06, Dec. 2024, doi: 10.14741/ijcet/v.14.5.5.

[7] R. P. Mahajan, "Improvised Diabetic Retinopathy Detection Accuracy in Retinal Images Using Machine Learning Algorithms," *TIJER*, vol. 12, no. 3, pp. 155–161, 2025.

[8] S. R. Sagili, S. Chidambaranathan, N. Nallametti, H. M. Bodele, L. Raja, and P. G. Gayathri, "NeuroPCA: Enhancing Alzheimer's disorder Disease Detection through Optimized Feature Reduction and Machine Learning," in *2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, IEEE, Jul. 2024, pp. 1–9. doi: 10.1109/ICEEICT61591.2024.10718628.

[9] I. González-Ruiz, P. Ramos-García, I. Ruiz-Ávila, and M. Á. González-Moles, "Early Diagnosis of Oral Cancer: A Complex Polyhedral Problem with a Difficult Solution," *Cancers (Basel).*, vol. 15, no. 13, Jun. 2023, doi: 10.3390/cancers15133270.

[10] M. A. Mostafiz, "Machine Learning for Early Cancer Detection and Classification: AIBased Medical Imaging Analysis in Healthcare," *Int. J. Curr. Eng. Technol.*, vol. 15, no. 03, pp. 1–10, 2025.

[11] P. P. Jalaja, D. Kommineni, A. Mishra, R. Tumati, C. A. Joseph, and R. V. S. S. B. Rupavath, "Predictors of Mortality in Acute Myocardial Infarction: Insights From the Healthcare Cost and Utilization Project (HCUP) Nationwide Readmission Database," *Cureus*, vol. 17, no. 5, May 2025, doi: 10.7759/cureus.83675.

[12] P. Ajay *et al.*, "Oral cancer prevalence in Western population of Maharashtra, India, for a period of 5 years," *J. Oral Res. Rev.*, vol. 10, no. 1, 2018, doi: 10.4103/jorr.jorr_23_17.

[13] S. Pandya, "Predicting Diabetes Mellitus in Healthcare: A Comparative Analysis of Machine Learning Algorithms," *Int. J. Curr. Eng. Technol.*, vol. 13, no. 06, 2023, doi: 10.14741/ijcet/v.13.6.5.

[14] N. Nasser, N. Nasser Filho, and R. L. Lehmkuhl, "Squamous cell cancer – 31-year epidemiological study in a city of south Brazil," *An. Bras. Dermatol.*, vol. 90, no. 1, pp. 21–26, Feb. 2015, doi: 10.1590/abd1806-4841.20153465.

[15] S. R. Sagili, C. Goswami, V. C. Bharathi, S. Ananthi, K. Rani, and R. Sathya, "Identification of Diabetic Retinopathy by Transfer Learning Based Retinal Images," in *2024 9th International Conference on Communication and Electronics Systems (ICCES)*, IEEE, Dec. 2024, pp. 1149–1154. doi: 10.1109/ICCES63552.2024.10859381.

[16] X. Chen, X. Zhang, M. Dong, L. Huang, Y. Guo, and S. He, "Deep Learning-Based Prediction of Wind Power for Multi-turbines in a Wind Farm," *Front. Energy Res.*, vol. 9, Jul. 2021, doi: 10.3389/fenrg.2021.723775.

[17] R. P. Mahajan, "Transfer Learning for MRI image reconstruction: Enhancing model performance with pretrained networks," *Int. J. Sci. Res. Arch.*, vol. 15, no. 1, pp. 298–309, Apr. 2025, doi: 10.30574/ijsra.2025.15.1.0939.

[18] R. Dattangire, D. Biradar, R. Vaidya, and A. Joon, "A Comprehensive Analysis of Cholera Disease Prediction Using Machine Learning," in *Congress on Intelligent Systems*, 2025, pp. 555–568. doi: 10.1007/978-981-96-2697-7_39.

[19] A. Kumar and L. Nelson, "Enhancing Oral Squamous Cell Carcinoma Detection Using EfficientNetB3 from Histopathologic Images," in *2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI)*, 2025, pp. 950–956. doi: 10.1109/ICMSCI62561.2025.10894480.

[20] G. Kaur, N. Sharma, and R. Gupta, "Enhanced Detection of Oral Squamous Cell Carcinoma from Histopathological Images Using Convolutional Neural Network Architectures," in *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*, 2024, pp. 1525–1530. doi: 10.1109/ICoICI62503.2024.10696099.

[21] G. Kaur and N. Sharma, "Automated Detection of Oral Squamous Cell Carcinoma using Transfer Learning Models from Histopathological Images," in *2024 3rd International Conference for Advancement in Technology (ICONAT)*, 2024, pp. 1–6. doi: 10.1109/ICONAT61936.2024.10775135.

[22] A. Benagi *et al.*, "Predictive Cloud-Based Application for the Early Detection and Risk Stratification of Oral Cancer using Keratin Pearls, Multiple Nucleoli and Vascular Invasion," in *2024 28th International Computer Science and Engineering Conference (ICSEC)*, 2024, pp. 1–6. doi: 10.1109/ICSEC62781.2024.10770727.

[23] G. A. I. Devindi *et al.*, "Multimodal Deep Convolutional Neural Network Pipeline for AI-Assisted Early Detection of Oral Cancer," *IEEE Access*, vol. 12, pp. 124375–124390, 2024, doi: 10.1109/ACCESS.2024.3454338.

[24] S. Ramya, R. I. Minu, and K. T. Magesh, "An Automated Approach for Identification of Oral Squamous Cell Carcinoma based on 2D-ICNN," in *Proceedings of the 2023 2nd International Conference on Augmented Intelligence and Sustainable Systems, ICAISS 2023*, 2023. doi: 10.1109/ICAISS58487.2023.10250540.

[25] J. J. Blessy and M. Sornam, "An Examen of Oral Carcinoma using Machine Learning Approaches," in *2022 International Conference on Automation, Computing and Renewable Systems (ICACRS)*, IEEE, Dec. 2022, pp. 657–661. doi: 10.1109/ICACRS55517.2022.10029007.

[26] E. N. Subhija and V. G. Reju, "Computer Assisted Recognition of Oral Squamous Cell Carcinoma Using GLCM Features in Haar Wavelet Transformed Space and Random Forest Classifier," in *Proceedings - 2nd International Conference on Next Generation Intelligent Systems, ICNGIS 2022*, 2022. doi: 10.1109/ICNGIS54955.2022.10091441.

[27] S. K. D. Sharma, "Oral squamous cell detection using deep learning," pp. 1–13, 2024.

[28] A. A. Alanazi, M. M. Khayyat, M. M. Khayyat, B. M. Elamin Elnaim, and S. Abdel-Khalek, "Intelligent Deep Learning Enabled Oral Squamous Cell Carcinoma Detection and Classification Using Biomedical Images," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–11, Jun. 2022, doi: 10.1155/2022/7643967.

[29] V. Pereira-Prado *et al.*, "Artificial Intelligence for Image Analysis in Oral Squamous Cell Carcinoma: A Review," *Diagnostics*, vol. 13, no. 14, Jul. 2023, doi: 10.3390/diagnostics13142416.

[30] A. Kumari and A. Mehta, "The Effectiveness of Feature Extraction and Machine Learning based approach for Early Detection of Oral Cancer." 2023. doi: 10.21203/rs.3.rs-2704739/v1.

[31] A. U. Fonseca *et al.*, "An Intelligent System to Improve Diagnostic Support for Oral Squamous Cell Carcinoma," *Healthcare*, vol. 11, no. 19, Oct. 2023, doi: 10.3390/healthcare11192675.

[32] S. B. Francis and J. Prakash Verma, "Deep CNN ResNet-18 based model with attention and transfer learning for Alzheimer's disease detection," *Front. Neuroinform.*, vol. 18, Jan. 2025, doi: 10.3389/fninf.2024.1507217.

[33] B. Ananthakrishnan, A. Shaik, S. Kumar, S. O. Narendran, K. Mattu, and M. S. Kavitha, "Automated Detection and Classification of Oral Squamous Cell Carcinoma Using Deep

Neural Networks," *Diagnostics*, vol. 13, no. 5, p. 918, Feb. 2023, doi: 10.3390/diagnostics13050918.

[34] A. Zafar, M. Khalid, M. Farrash, T. M. Qadah, H. F. M. Lahza, and S.-H. Kim, "Enhancing Oral Squamous Cell Carcinoma Detection Using Histopathological Images: A Deep Feature Fusion and Improved Haris Hawks Optimization-Based Framework," *Bioengineering*, vol. 11, no. 9, 2024, doi: 10.3390/bioengineering11090913.

[35] A. U. Rahman *et al.*, "Histopathologic Oral Cancer Prediction Using Oral Squamous Cell Carcinoma Biopsy Empowered with Transfer Learning," *Sensors*, vol. 22, no. 10, pp. 1–14, 2022, doi: 10.3390/s22103833.

[36] M. Das, R. Dash, and S. K. Mishra, "Automatic Detection of Oral Squamous Cell Carcinoma from Histopathological Images of Oral Mucosa Using Deep Convolutional Neural Network," *Int. J. Environ. Res. Public Health*, vol. 20, no. 3, 2023, doi: 10.3390/ijerph20032131.

