

Research Article

Analysis of Histopathological Images for Early Diagnosis of Oral Squamous Cell Carcinoma by Hybrid Systems Based on CNN Fusion Features

Ibrahim Abdulrab Ahmed,¹ Ebrahim Mohammed Senan ,²
and Hamzeh Salameh Ahmad Shatnawi¹

¹Computer Department, Applied College, Najran University, Najran 66462, Saudi Arabia

²Department of Artificial Intelligence, Faculty of Computer Science and Information Technology, Alrazi University, Sana'a, Yemen

Correspondence should be addressed to Ebrahim Mohammed Senan; senan1710@gmail.com

Received 17 April 2023; Revised 11 May 2023; Accepted 18 May 2023; Published 30 May 2023

Academic Editor: Riccardo Ortale

Copyright © 2023 Ibrahim Abdulrab Ahmed et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Oral squamous cell carcinoma (OSCC) is one of the deadliest and most common types of cancer. The incidence of OSCC is increasing annually, which requires early diagnosis to receive appropriate treatment. The biopsy technique is one of the most important techniques for analyzing samples, but it takes a long time to get results. Manual diagnosis is still subject to errors and differences in doctors' opinions, especially in the early stages. Thus, automated techniques can help doctors and patients to receive appropriate treatment. This study developed several hybrid models based on the fused CNN features for diagnosing OSCC-100x and OSCC-400x datasets for oral cancer, which have the ability to analyze medical images with a high level of precision and accuracy. They can detect subtle patterns, abnormalities, or indicators of diseases that may be difficult to recognize with the naked eye. The systems have the potential to significantly reduce human error and provide more consistent and reliable results, resulting in improved diagnostic accuracy. The systems also have the potential for early detection of OSCC for treatment success and improved patient outcomes. By detecting diseases at an early stage, clinicians can initiate interventions in a timely manner, potentially preventing OSCC progression and improving the chances of successful treatment. The first strategy was based on GoogLeNet, ResNet101, and VGG16 models pretrained, which did not achieve satisfactory results. The second strategy was based on GoogLeNet, ResNet101, and VGG16 models based on the adaptive region growing (ARG) segmentation algorithm. The third strategy is based on a mixed technique between GoogLeNet, ResNet101, and VGG16 models and ANN and XGBoost networks based on the ARG hashing algorithm. The fourth strategy for oral cancer diagnosis by ANN and XGBoost is based on features fused between CNN models. The ANN with fusion features of GoogLeNet-ResNet101-VGG16 yielded an AUC of 98.85%, accuracy of 99.3%, sensitivity of 98.2%, precision of 99.5%, and specificity of 98.35%.

1. Introduction

Oral squamous cell carcinoma (OSCC) is one of the most common types of cancer worldwide that arise in the oral cavity. The incidence and death rate of OSCC are increasing annually [1]. According to the International Agency on Cancer, the number of infections in 2020 reached 377,000 new cases and more than 177,000 deaths worldwide [2]. The increasing number of patients compared to doctors is

a health problem; the inability of oral cancer patients to achieve a diagnosis in rural areas and receive treatment promptly means the survival rate for five years is 15%. In contrast, the five-year survival rate [3] in developed countries is about 65%. The OSCC is aggressive, and multiple chemotherapy, radiotherapy, or surgical intervention treatments are necessary [4]. OSCC cancer represents 90% of all types of oral cancer and is the most aggressive, and late diagnosis leads to death [5]. OSCC arises as epithelial

dysplasia from precursor lesions termed potentially malignant oral disorders (PMODs). Leukoplakia, oral lichen, and erythroplakia lead to their development into malignant tumors if neglected without a diagnosis [6]. Studies have shown that PMOD changes into homogeneous and heterogeneous symptoms; heterogeneous symptoms are likely to turn into malignant lesions if not diagnosed early [7]. Thus, distinguishing between malignant clinical features of PMOD is a concern. The OSCC must therefore be diagnosed early to diagnose malignancies at the PMOD stage. There are many risk factors for OSCC, such as alcohol and tobacco use, infection with HPV, age, gender, and family history [8]. There are many clinical indicators of the possibility of infection with OSCC that must be taken into account; a doctor should be consulted immediately, the most important of which are mouth ulcers that do not heal, red and white sores in the tongue, lips, or mouth that do not heal, swelling of the jaw, and difficulty swallowing or speaking [9]. At present, manual assessment examination through visual and tactile analysis followed by tissue biopsy is the gold standard for diagnosing the type of lesion [10]. A biopsy is to take part of the tissue of the suspected area and complete its analysis with hematoxylin-eosin and analyze it with special microscopic devices [11]. However, the process is boring, tedious, takes a long time to analyze pathological tissues, and is subject to the differing opinions of experts and specialists. Also, the wide gap between the number of doctors and patients, especially in developing countries and rural areas in developed countries, poses a challenge to the early diagnosis of OSCC. Therefore, early diagnosis of OSCC is necessary to receive treatment and avoid its development to dangerous stages. Computer-assisted systems improve patients' chances of survival through early diagnosis of OSCC. Artificial intelligence (AI) technologies have participated in various healthcare fields, such as analyzing medical images for early detection of tumors and diseases. AI techniques train models with a large part of the data set so that they gain experience and knowledge and store them, and then, their performance is tested through new images whose features are extracted and compared with the stored features and then classified based on the similarity between the features of the new image with the stored features (data trained). AI technologies have worked to identify biomarkers for OSCC prediction, reduce clinicians' burden, and interpret complex data in histopathological images. In recent years, deep learning techniques have emerged with their superior ability to analyze medical images compared to the performance of human experts. In this work, artificial intelligence techniques were developed that combine deep learning networks and machine learning algorithms along with histopathological image optimization techniques. Clinical indicators that are not visible to the naked eye are similar to benign and malignant tumors in the early stages, which constitutes a challenge for doctors, so features were extracted by many deep learning models and converted into high-level vector features by global average pooling, then combined high-level feature maps to produce new feature maps, and then send them to machine learning algorithms for classification.

The novelty of this paper lies in the combination of specific AI techniques and features for the early detection of OSCC using histopathological images. Here are the key aspects that contribute to the novelty of this paper: The paper proposes the use of a combined set of CNN features for histopathological analysis of OSCC. CNNs have proven effective in image classification tasks, including medical imaging. However, the novelty lies in combining multiple CNN features to capture diverse aspects of the histopathological images. This combination of features enhances the representation and discriminative power of the hybrid model, leading to improved performance in detecting early signs of OSCC lesions. The paper focuses on the early detection of OSCC lesions using histopathological images. Early detection is crucial for successful treatment and improved patient outcomes in OSCC cases. By employing hybrid systems and combining CNN features, the proposed approach aims to enhance the accuracy and efficiency of early OSCC lesion detection, contributing to timely interventions and potentially saving lives. Histopathological images require specialized analysis techniques due to their unique characteristics, such as color variations, structural patterns, and asymmetries. The paper specifically addresses the challenges associated with histopathological image analysis and proposes hybrid systems tailored for this domain. The novelty lies in the adaptation of hybrid systems and CNN features to effectively extract relevant information from histopathological images, enabling accurate diagnosis and detection of OSCC lesions. Overall, the novelty of the paper lies in the specific application of hybrid systems, the combination of CNN features for histopathological image analysis, the focus on early detection of OSCC lesions, and the adaptation of hybrid systems to address the challenges of analyzing histopathological images. These contributions can potentially advance the field of histopathology and improve the accuracy and efficiency of OSCC lesion detection, aiding in early diagnosis and intervention.

Using artificial intelligence (AI) techniques for diagnosing medical images offers several benefits, but it also comes with certain risks. Here are some of the potential benefits and risks associated with the use of AI in medical image diagnosis. The benefits are AI algorithms can analyze medical images with high precision, potentially leading to more accurate and consistent diagnoses. They can detect subtle patterns, variations, or abnormalities that might be difficult for human observers to identify. AI algorithms can process medical images quickly, allowing for faster diagnoses and reducing the time required for analysis. This can lead to more efficient workflows and improved patient outcomes, particularly in time-sensitive cases. AI techniques can help overcome geographical barriers by providing access to expert-level diagnoses in regions with limited healthcare resources. This can bridge the gap in specialized medical expertise and ensure patients receive timely and accurate assessments regardless of their location. AI algorithms can serve as decision support tools for healthcare professionals. By providing additional insights and highlighting potential areas of concern, they can aid physicians in making more

informed decisions and developing appropriate treatment plans. AI can process and analyze vast amounts of medical imaging data, leading to valuable insights and discoveries. It can uncover patterns, correlations, and trends that may not be readily apparent to human observers, potentially advancing medical research and improving patient care.

The risks are AI algorithms heavily rely on the training data they are exposed to. If the training data are limited, biased, or not representative of the target population, the algorithm's performance may be compromised. This can lead to inaccurate diagnoses or the potential for algorithmic bias, especially in underrepresented groups. AI algorithms often work as black boxes, making it challenging to understand how they arrive at their diagnoses. This lack of transparency can raise concerns regarding the interpretability of results and may hinder the trust and acceptance of AI-based diagnoses among healthcare professionals and patients. Medical images contain sensitive patient information. When using AI, there is a risk of unauthorized access, data breaches, or the potential for the misuse of patient data. Implementing robust security measures and ensuring compliance with privacy regulations are crucial to mitigate these risks. AI raises ethical concerns related to responsibility, accountability, and liability. Determining who is responsible for errors or adverse outcomes when AI is involved can be complex. Ensuring proper oversight, regulation, and adherence to ethical guidelines are essential to address these concerns. Integrating AI into medical practice requires effective collaboration between healthcare professionals and AI systems. Striking the right balance between the expertise of healthcare professionals and the capabilities of AI algorithms is crucial to maximize the benefits and minimize potential risks.

It is important to note that while AI techniques have shown promising results in medical image diagnosis, they are not intended to replace healthcare professionals. AI should be viewed as a supportive tool that complements the expertise and clinical judgment of physicians, aiding in more accurate diagnoses and personalized treatment decisions.

The main contributions to this study are as follows:

- (1) Increase the contrast of low-contrast areas and remove artifacts by two successive techniques
- (2) Development of an ARG algorithm to segment regions of interest in OSCC-100x and OSCC-400x images and isolate them from healthy tissues
- (3) Development of hybrid techniques between CNN models, ANN networks, and XGBoost based on the ARG algorithm for effective diagnosis of histopathological images of oral cancer
- (4) Developing effective hybrid techniques to combine features of CNN models and diagnosis by ANN and XGBoost networks for accurate diagnosis of histopathological images of oral cancer

The rest of the study is arranged as follows: Section 2 discusses the techniques and findings of relevant previous studies. Section 3 explains the tools and methodologies applied for the histopathological analysis of oral cancer.

Section 4 summarizes the most critical performance results of the proposed methodologies. Section 5 compares the performance of the methodologies. Section 6 concludes the work.

2. Related Work

Alabi et al. [12] model pipeline based on the CNN for OSCC detection, encompassing precise diagnosis and precision medicine. It highlights the potential of CNN in analyzing oral cancer data, such as imaging and genomic information, to improve diagnostic accuracy. The paper emphasizes the importance of integrating these techniques into clinical practice to enable personalized treatment strategies for oral cancer patients. Lin et al. [13] obtained images of oral cavities and employed the HRNet model for diagnosing them. The HRNet model outperformed the ResNet50 and DenseNet169 models, demonstrating an accuracy of 84.3% and a sensitivity of 83%. Rahman et al. [14] proposed a methodology for prognosticating histopathologic oral cancer through the utilization of biopsy samples from patients with OSCC. The method incorporates transfer learning, which leverages pretrained deep learning models to extract meaningful features from the biopsy images. By training the model on a large dataset, the proposed approach achieves accurate prediction of histopathologic oral cancer. The results demonstrate the potential of this technique in assisting clinicians in the diagnosis and treatment planning of OSCC. The approach achieved an accuracy of 90.06%. Fati et al. [15] employed a combination of deep learning techniques, including SVM and ANN, integrated with both deep learning-based features and manually crafted features, to diagnose OSCC in biopsy images. AlexNet + SVM achieved an accuracy of 97.1% and a sensitivity of 97.81%. Warin et al. [16] used the DenseNet121 model and R-CNN to train 490 oral biopsy images of OSCC and compared the performance of the two systems on 140 test images. The DenseNet121 model achieved better accuracy than R-CNN, with the DenseNet121 model achieving an accuracy of 99%, while R-CNN yielded an accuracy of 76.67%. Camalan et al. [17] developed the inception-ResNet-V2 model and created maps to focus on the affected area to make a classification decision, which achieved an accuracy of 73.6% and an F1 score of 97.9%. Musulin et al. [18] integrated the Xception and SWT approach that works in two phases for multiclass classification and segmentation of stroma and epithelial histology to help clinicians to classify OSCC. The DeepLabv3 approach with Xception_65 achieved an accuracy of 96.3% and semantic segmentation of 87.8%. Lin et al. [19] developed a Genetic and Epigenetic Network (GWGEN) utilizing the GRN and PPIN systems to identify samples of OSCC and non-OSCC. Das et al. [20] employed a 10-layer CNN for the purpose of automatic identification of OSCC. The study involved evaluating the performance of their CNN model by comparing it to pretrained CNN models. The network performed better than the CNN models for evaluating a dataset of 1224 histological images, which achieved an accuracy of 97.82%. Jing et al. [21] utilized machine learning algorithms to forecast the progression from

leukoplakia to oral cancer. Their approach involved examining the shared weighted gene network and differential expression patterns to identify seven genes that are correlated with the development of leukoplakia into oral cancer. Yang et al. [22] developed a deep learning algorithm to detect OSCC. The algorithm underwent training using a dataset consisting of 1925 images and was subsequently tested on 100 images. The results of the evaluation demonstrated a sensitivity of 98%, specificity of 92%, and an F1 score of 95.1%. Additionally, the same set of 100 images was assessed by a pathologist, who achieved an F1 score of 92.21%. However, when utilizing the deep learning model, the pathologist's F1 score improved to 95.66%. Amin et al. [23] adapt three pretrained CNN models to extract features individually and then extract the features sequentially. Sequential models achieved better results than individual models, which reached an accuracy of 96.66% and a recall of 98.3%. Deo et al. [24] extract features by two-dimensional wavelet transform and then apply two pretrained CNN models for OSCC detection; the model achieved an accuracy of 92%. Ghosh et al. [25] utilized Fourier Transform Infrared Spectroscopy and Raman spectroscopy techniques to analyze the spectral characteristics of oral cancer samples, aiming to detect and diagnose epi-genetic alterations associated with the disease. The DRNN layer is used to detect and classify peak epigenetic features that achieve an AUC of 88%. Deif et al. [26] used four DL models to extract and identify important features through the BPSO tool. The selected features were rated by XGBoost, which achieved an accuracy of 96.3%.

Panigrahi et al. [27] investigated the effectiveness of capsule networks in identifying and classifying cancerous cells in these images. The study utilizes a dataset of OSCC images and applies a capsule network architecture for feature extraction and classification. The paper concludes that the capsule network shows promising potential for the analysis of OSCC histopathological images, suggesting its usefulness as a tool for assisting in the diagnosis and treatment of oral cancer. The system reached an accuracy of 97.35%, a sensitivity of 96.92%, and a specificity of 97.78%. Wu et al. [28] developed an automated approach for accurate and efficient region segmentation. The proposed model employs various machine learning techniques and a dataset of H&E-stained histology images from multiple centres. The approach achieved an accuracy of 95.8%, a sensitivity of 79.1%, and a precision of 85.7%. Das et al. [29] developed an efficient and accurate method for identifying cancerous cells in these images. The proposed CNN model is trained on a dataset of histopathological images and employs various layers for feature extraction and classification. ResNet101 achieved an accuracy of 89%, sensitivity of 93%, precision of 89%, and specificity of 88%. Myriam et al. [30] introduced a novel meta-heuristic algorithm that combines particle swarm optimization (PSO) and Al-Biruni Earth Radius Optimization (ABERO) methods for the detection of oral cancer. The algorithm enhanced the accuracy and efficiency of oral cancer detection by leveraging the strengths of both PSO and ABERO. PSO-ABERO achieved an accuracy of 97.3%, a sensitivity of 94.3%, and a precision of 96.3%. Panigrahi

et al. [27] explored the use of capsule networks for analyzing histopathological images of oral squamous cell carcinoma (OSCC). The authors propose a novel approach that leverages the advantages of capsule networks to improve the accuracy of OSCC analysis through the dynamic routing consensus of capsule network. The capsule networks achieved an accuracy of 97.35%, a sensitivity of 97.78%, and a precision of 96.92%.

From the above, the gaps in previous studies that have been addressed in this study can be summarized as follows:

- (1) Limited Feature Extraction: Previous studies have used traditional feature extraction methods that may not capture complex patterns and textures in histopathological images effectively. This research can address this gap by utilizing several models of CNNs, which are known for their ability to automatically learn relevant features directly from images and integrate them into feature vectors.
- (2) Insufficient Classification Accuracy: Previous studies have encountered challenges in achieving high accuracy in classifying OSCC based on histopathological images. This research can address this gap by combining multiple CNN features and taking advantage of complementary information to improve classification accuracy.
- (3) Lack of Robustness: Previous studies have struggled with decreased robustness when dealing with differences in image quality, lighting conditions, and different acquisition devices. This research has addressed this gap by integrating two overlapping techniques to enhance the robustness of the proposed AI techniques, by removing noise and increasing the low contrast of the edges of cancer cells.
- (4) Limited Focus on Early Diagnosis: Previous studies focused mainly on the general classification of OSCC without focusing specifically on early diagnosis. This research can address this gap by specifically targeting early diagnosis of OSCC, bearing in mind that early intervention can significantly improve patient outcomes.
- (5) This study was distinguished from previous studies by applying hybrid techniques based on the extraction of hidden features that are not visible to the naked eye through the extraction of hybrid features of the features of deep learning models.

3. Materials and Methods

3.1. Description of OSCC Dataset. This work evaluated the proposed systems on histopathological images taken as a biopsy from the oral cancer histology dataset's oral cavity. The data set was obtained from the Institute of Cancer Research Dr. B Borooah and Ayursundra Healthcare Pvt. Lt. Images of the dataset were acquired by a Leica ICC50 HD microscope at magnification factors size of 100x and 400x [31]. The dataset contains 1224 histopathological images divided into two groups, with a magnification factor of 100x and 400x. Each

magnification factor contains images of normal epithelium and histopathology of OSCC. First, a dataset with a magnification factor of 100x contains 528 images divided into 89 normal oral histological images and 439 histological images of OSCC. Second, a dataset with a magnification factor of 400x contains 696 images divided into 201 images of normal oral epithelium and 495 histological images of OSCC. All dataset images are at 100x and 400x magnifications with a size of 2048×1536 pixels. Figure 1(a) shows sample histopathological images of an OSCC dataset at a magnification of 100x, while Figure 2(a) shows histopathological samples of an OSCC dataset at a magnification of 400x.

3.2. Enhancing Histopathological Images of OSCC. Dataset images contain artifacts due to the preparation of biopsies, such as different staining and luminance at image acquisition. Hence, these artifacts impair the performance of CNN models and machine learning algorithms. The image optimization step is crucial to eliminate artifacts, standardize the colors of all images, and make systems robust for image classification. A mean of RGB primary colors is calculated, and then, the scale is set to calculate color consistency. Images were passed to an averaging filter to eliminate artifacts and a contrast-limited adaptive histogram equalization (CLAHE) technique to increase the contrast of ROI pathological [32].

First, the average filter works by setting its operator to 4×4 , which means the filter will cover 16 pixels of the image to be processed each time. The filter selects one central pixel and calculates the average of the 15 adjacent pixels [33]. The process is repeated with the number of pixels in the image until all pixels have been processed, as shown in the following equation:

$$z(y) = \frac{1}{m} \sum_{i=0}^{m-1} x(y-i), \quad (1)$$

where $z(y)$ is the output, m is the number of pixels, $a(x)$ is the input, and $x(y-i)$ is the earlier input.

Second, to show the contrast of the affected pathological tissues by increasing the contrast of the edges using the CLAHE technique, this technology works on processing each pixel in the image based on neighboring pixels by distributing the bright pixels to the dark areas of the image [34]. A pixel is compared to its neighbors, and contrast is increased or decreased by comparison. If the value of the target pixel is greater than the value of the neighboring pixels, the contrast decreases whereas the contrast increases if the value of the target pixel is smaller than the value of neighboring pixels.

Figure 1(b) shows histopathological images of the OSCC dataset at 100x magnification after improvement, while Figure 2(b) shows histopathological samples of the OSCC dataset at 400x magnification after improvement.

3.3. Adaptive Region Growing. Images of pathological tissues consist of two parts: regions of interest (ROI) named the affected regions and unimportant regions named healthy regions [35]. Therefore, extracting features from all the images leads to inaccurate diagnostic results. Therefore, it is necessary to separate the regions with interest from other

regions by segmentation algorithms [36]. The segmentation algorithms separate the affected pixels from health. In this study, the adaptive region growing (ARG) algorithm collects similar pixels in ROI. For the sake of the success of the algorithm, it must meet the following conditions:

- (1) $\bigcup_{i=1}^n x_i = x$
- (2) $x_i = 1, 2, \dots, N$ is connected
- (3) $P(x_i) = \text{TRUE}$ for $1, 2, \dots, N$
- (4) $P(x_i \cup x_j) \text{FALSE}$ for $i \neq j$ Where x_i and x_j are neighboring regions

First, the segmentation process must be complete. Second, all the like pixels should be in one area, and the union of all regions leads to the full image. Pixels of the same area should be connected. Third, similar pixels should be correct when applying full pixels of the image. Fourth, there are no similar pixels in different regions. The algorithm work mechanism from bottom to up starts with pixels, and similar pixels are collected to form regions. The algorithm uses local information to create the region. The fundamental concept of the ARG algorithm involves initiating a single pixel for each region, and then allowing the regions to expand progressively, ensuring that neighboring units with similarities are grouped together within the same area. The edges regions grow with similar pixels; the process continues until each pixel is assigned to its area. The ARG algorithm works to collect adjacent pixels x_i to x_j inside ROI based on $|I(x_i) - I(x_j)| \leq T$, where T is the value of the fixed threshold and $I(\cdot)$ is the value of the pixel density. The aim of applying this algorithm is to extract regions of interest only and feed them to the following stages for analysis. This method is one of the essential contributions of this study, as the interest regions of the OSCC data group have been extracted and saved in a new folder called OSCC-ROI. In all literary studies, CNN models receive the OSCC dataset and work to analyze and classify while distinguishing our study is that CNN models receive the OSCC-ROI dataset and work to analyze and extract their advantages. Figure 3 shows samples of the pathological tissue of OSCC after ROI extract.

3.4. Extracting CNN Maps. One of the main aspects of CNN models is their superior ability to extract features from input images. CNN models can automatically interpret and analyze the most representative and relevant features. Thus, CNN models are successive layers for feature discovery: the first layers that follow the input layer extract low-level features while other layers extract higher-level features. Low-level features, such as points and edges, are fed back to the next layers to form the high-level features. The layers of CNN models consist of convolutional, pooling, fully connected, and auxiliary layers [37].

First, convolutional layers receive images from the input layer, analyze them, and extract low-level features, followed by high-level features. Each CNN model consists of convolutional layers different from the other model, among the most important CNN layers. Three main parameters controlling convolutional layers' action are filter size, zero padding, and p-step [38]. Each layer has a filter of a different size than the other filter. The filter $f(t)$ of a given size wraps around an area of the

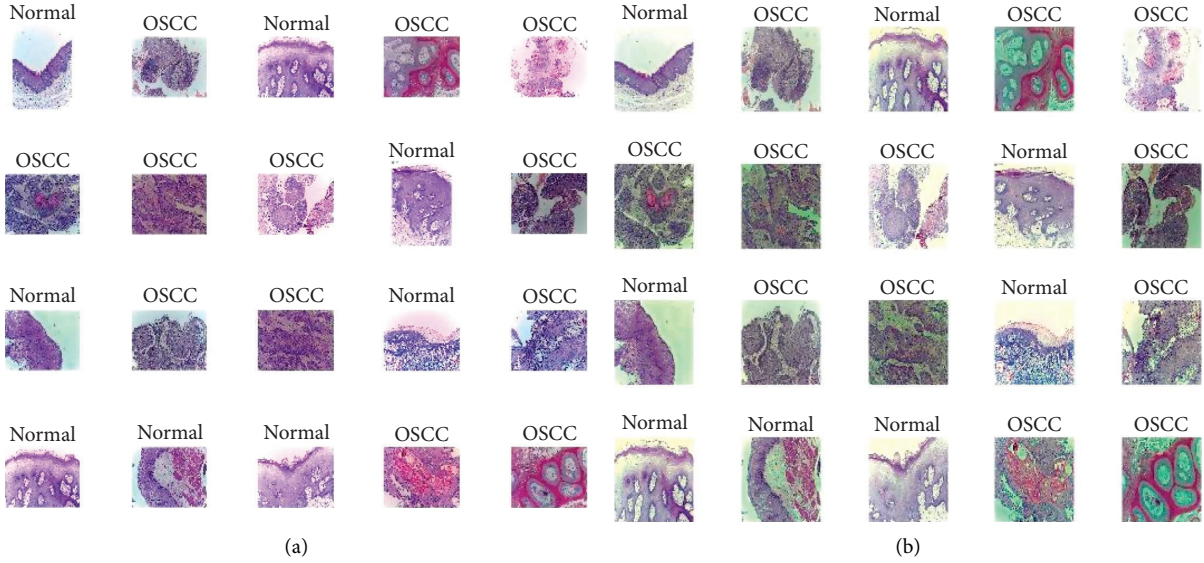


FIGURE 1: Histopathological images of the oral cavity of the OSCC-100x dataset: (a) original images and (b) after improvement.

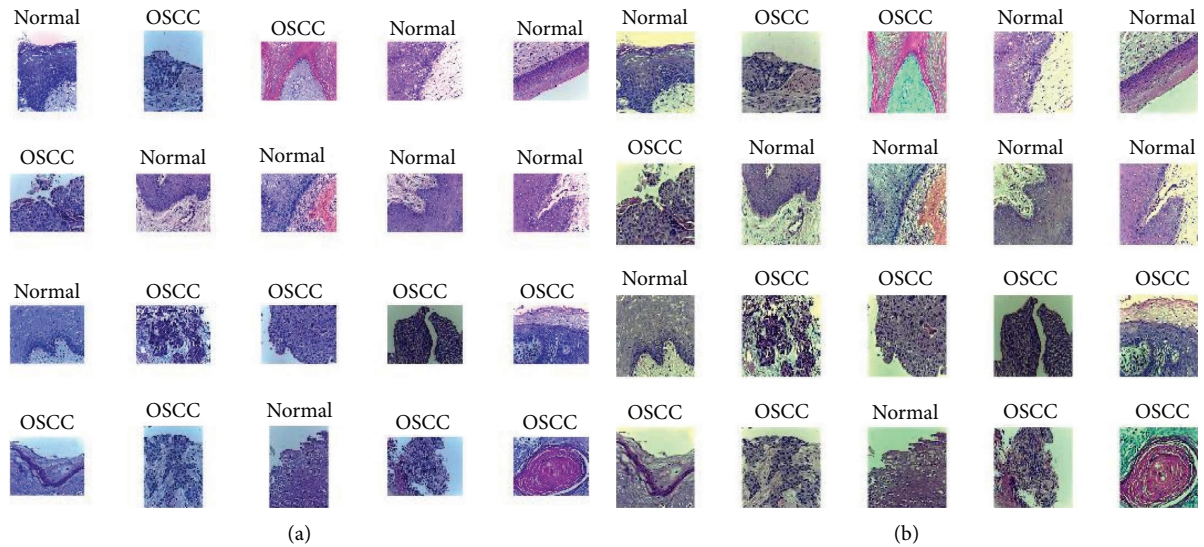


FIGURE 2: Histopathological images of the oral cavity of the OSCC-400x dataset: (a) original images and (b) after improvement.

target image $x(t)$ of the same size as the filter as in equation (2). The benefit of zero padding is preserving the original image's size by inserting zeros around the image edges. P-step specifies the number of steps the filter jumps on the image each time; when p-step = 2, the filter jumps 2 steps each time [39].

$$y(t) = (x * f)(t) = \int x(a)f(t-a)da, \quad (2)$$

where $f(t)$ refers to the filter, and $x(t)$ and $y(t)$ refer to the input and output image.

The output of each neuron in the CNN by calculating the input of the previous neuron W , the filter size K , the number of P-steps S , and the zero padding P is shown in the following equation:

$$\text{Output Neurons} = \frac{W - K + 2P}{S} + 1, \quad (3)$$

where W indicates the size of the input neurons, K is the size of the filter, P is the size of the padding, and S indicates the stride.

Pooling layers reduce the millions of neurons, parameters, and connections generated by convolutional layers, which requires complex and time-consuming calculations. There are two-layer pooling methods, max pooling and average pooling [40]. With max pooling, the largest value of the selected pixels will be taken and represented in one cell. In the case of average pooling, the average value of the selected pixels will be calculated and represented in one cell. Thus, pooling layers reduce the size

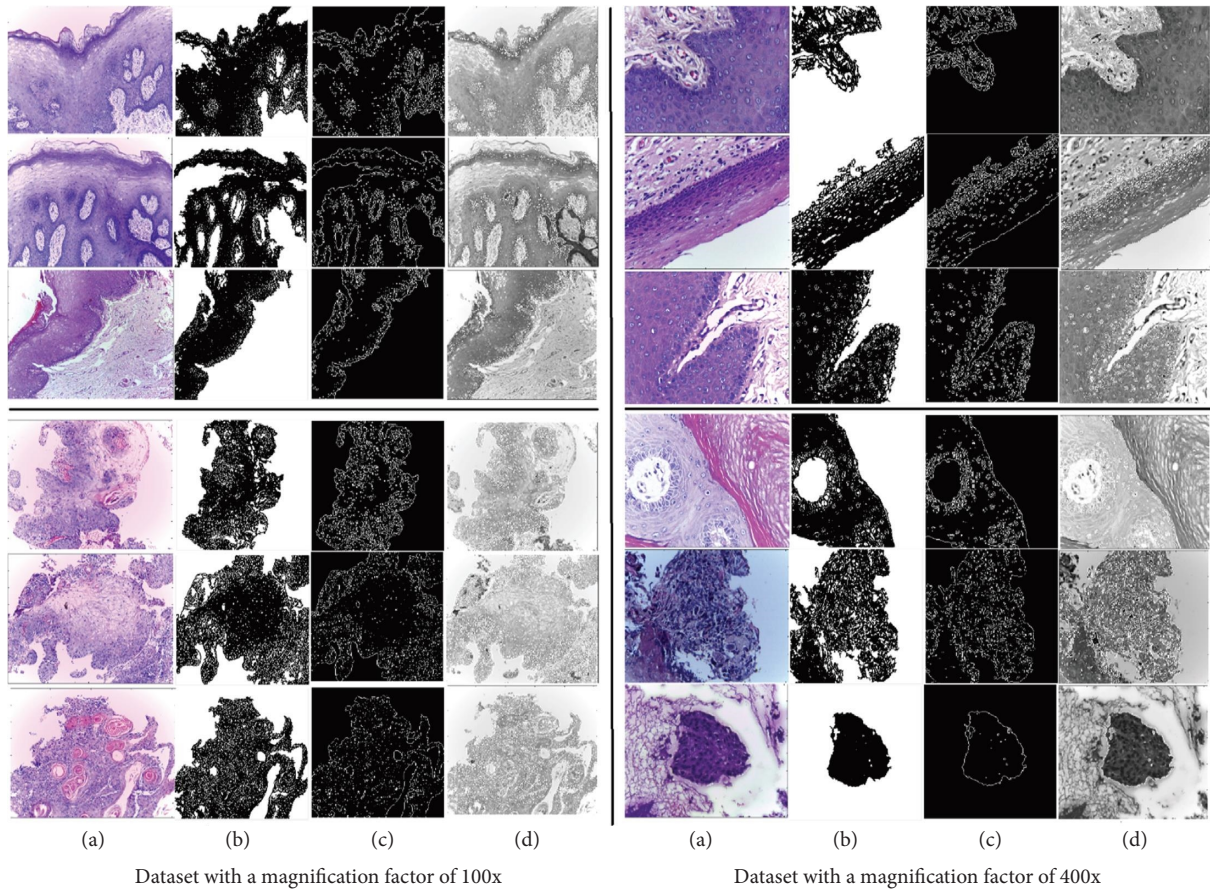


FIGURE 3: Samples from the OSCC dataset at magnifications of 100x and 400x: (a) original images, (b) segmentation of infected histopathological, and (c) and (d) regions of interest.

of the underlying feature maps, making calculations more efficient [41].

Some convolutional layers follow the ReLU activation layer, which receives feature maps and passes positive values, while converting negative values in the maps to zero, as shown by the following equation:

$$\text{ReLU}(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0. \end{cases} \quad (4)$$

Fully connected layers are among the most critical layers of CNN models for figuring out the hierarchy and classification of features. Each neuron in the FCL is connected to all neurons of the previous layer. FCL converts features from high-level 2D to flat (vectors) [42].

Finally, the SoftMax activation function is used, which divides the output into multiple classes according to the dataset classes and labels each feature vector to a specific class, as in the equation.

In this approach, the deep feature maps were extracted using GoogLeNet, ResNet101, and VGG16 models and classified by machine learning algorithms. The last convolutional layers produce higher-level features as follows: (7, 7, 512), (3, 3, 512), and (7, 7, 512), respectively. The last layer of the models is the global average pooling layer after the high-level convolutional layers, which converts high-

level features into low-level features stored in feature vectors of sizes 4096, 2048, and 4096 for GoogLeNet, ResNet101, and VGG16, respectively. Thus, the data set becomes represented by a features matrix: first, a data-set-100x with sizes 4069×528 , 2048×528 , and 4096×528 for GoogLeNet, ResNet101, and VGG16, respectively. Second, a data-set-400x with sizes 4069×696 , 2048×696 , and 4096×696 for GoogLeNet, ResNet101, and VGG16, respectively.

3.5. Classification. Classification is the last stage in biomedical image processing, which depends on the efficiency of the previous stages. After extracting the features of the histopathological images of OSCC cancer from the ROI by CNN models, each image feature was placed into a vector, and all the image features were represented in the feature matrix, which is input to the PCA algorithm. PCA selects highly representative features, deletes redundant and unimportant features, and outputs them as input to ANN and XGBoost. After feature reduction using PCA, the feature matrices for two datasets become: first, the data-set-100x with sizes 660×528 , 530×528 , and 680×528 for GoogLeNet, ResNet101, and VGG16, respectively, and second, the data-set-400x with sizes 660×696 , 530×696 , and 680×696 for GoogLeNet, ResNet101, and VGG16, respectively.

3.5.1. ANN Network. ANN network is a mathematical model inspired by how the nervous system processes data or images. ANN works through interconnected processing units; neurons receive data from previous neurons with weights associated with each neuron [43]. The most important component of ANNs is neurons, which convert input data into output. The data are modified by the weight factor that connects the neurons. ANN has the ability to adjust weights based on trial and error, which makes ANN adaptable to different types of inputs and has the ability to learn. An ANN has three layers: the input, which receives data from outside the network. Hidden layers are layers within the network that receive data from the input layer and process the data and update the weights each time [44]. ANN measures the minimum square error (MSE) between the actual values x_j and the output y_j , and through it, the weights are calculated as in equation (5). In this study, the number of hidden layers was set to 15 through training and trial and error. Output layers contain an activation function to label the data and sort it into two classes, either normal or oral cancer.

$$\text{MSE} = \frac{1}{m} \sum_{j=1}^m (x_j - y_j)^2, \quad (5)$$

where m is the data points.

3.5.2. XGBoost Algorithm. XGBoost is a powerful and efficient algorithm based on the ensemble learning method. Due to the poor performance of the single-learning model algorithms, algorithms based on ensemble learning were developed to obtain a powerful predictive model [45]. The network produces decision trees sequentially, each with a different weight. The network produces an effective model by integrating weak learners with strong ones, called boosting. The idea of boosting is that each subsequent tree performs better than the previous tree so that each subsequent tree takes advantage of the weaknesses of the previous tree and updates the parameters [46]. Boosting results in a strong learner model by treating weak learners with successor trees. The network continues until a strong learner model with high prediction efficiency is generated.

3.6. Training of Proposed Systems

3.6.1. Training of Pretrained Models. The first strategy for classifying histopathological images of oral cancer is used by applying pretrained GoogLeNet, ResNet101, and VGG16 models [47]. The OSCC dataset is fed into models, and CNN layers analyze images to extract features by convolutional layers and pooling and classification through fully connected layers [48].

The second strategy is to diagnose histopathological images of oral cancer through the following sequence of steps: first is applying the averaging filter and CLAHE method to improve the images. The second is the application of the ARG method to extract ROI and isolate them from other tissues and save them in a new OSCC-ROI dataset

folder. Third is feeding the new data set to GoogLeNet, ResNet101, and VGG16 models to extract and classify the features.

3.6.2. Training of Hybrid Method. The third strategy is to analyze histopathological images for the diagnosis of oral cancer by hybrid systems through the sequence of the following implementation steps as shown in Figure 4.

First, removing the artifacts and increasing the ROI region contrast through the average filter and CLAHE methods. The second is obtaining ROI regions only and isolating them from other tissues by the ARG method. Third, the GoogLeNet, ResNet101, and VGG16 models receive ROI regions for analysis and feature map extraction. The GoogLeNet, ResNet101, and VGG16 models produce features of size 4069×528 , 2048×528 , and 4096×528 for data-set-100x and 4069×696 , 2048×696 , and 4096×696 for data-set-400x. Fourth is reducing the characteristics of GoogLeNet, ResNet101, and VGG16 models by PCA to select the most important features. Fifth, ANN and XGBoost receive feature vectors and allocate 80% for system training and weight control and 20% for the testing phase.

3.6.3. Training of Hybrid Method Based on Fusion Features CNN. The fourth strategy for analyzing histopathological images for diagnosing oral cancer by hybrid systems with features fused between CNN models is through [49] the sequence of the following implementation steps as shown in Figure 5. The first four sequential steps in the third strategy are the same as in the fourth. Fifth, a series fusion of the features of the GoogLeNet, ResNet101, and VGG16 models is as follows: GoogLeNet-ResNet101, ResNet101-VGG16, GoogLeNet-VGG16, and GoogLeNet-ResNet101-VGG16. Sixth is saving the fused features in new feature matrices and sending them to ANN and XGBoost networks to train 80% of them and keep 20% for systems performance testing [50].

4. Results of Systems Execution

4.1. Split of OSCC-100-400X Datasets. In this study, two OSCC datasets were used for the 100x and 400x magnification factors to measure the performance of several systems. The OSCC-100x dataset contains 528 histological images divided into 439 OSCC histopathological and 89 normal oral cavity histopathological while the OSCC-400x dataset contains 696 histopathological images divided into 495 OSCC histopathological and 201 normal oral cavity histopathological. The two datasets were split into 80% for systems training and validation and 20% for testing and benchmarking, as shown in Table 1.

4.2. System Performance Metrics. The systems generate a confusion matrix through which the performance of the techniques is evaluated through equations (6)–(10). The confusion matrix contains rows and columns, where all systems performance test samples are represented in the cells of the matrix. The rows of the confusion matrix represent the

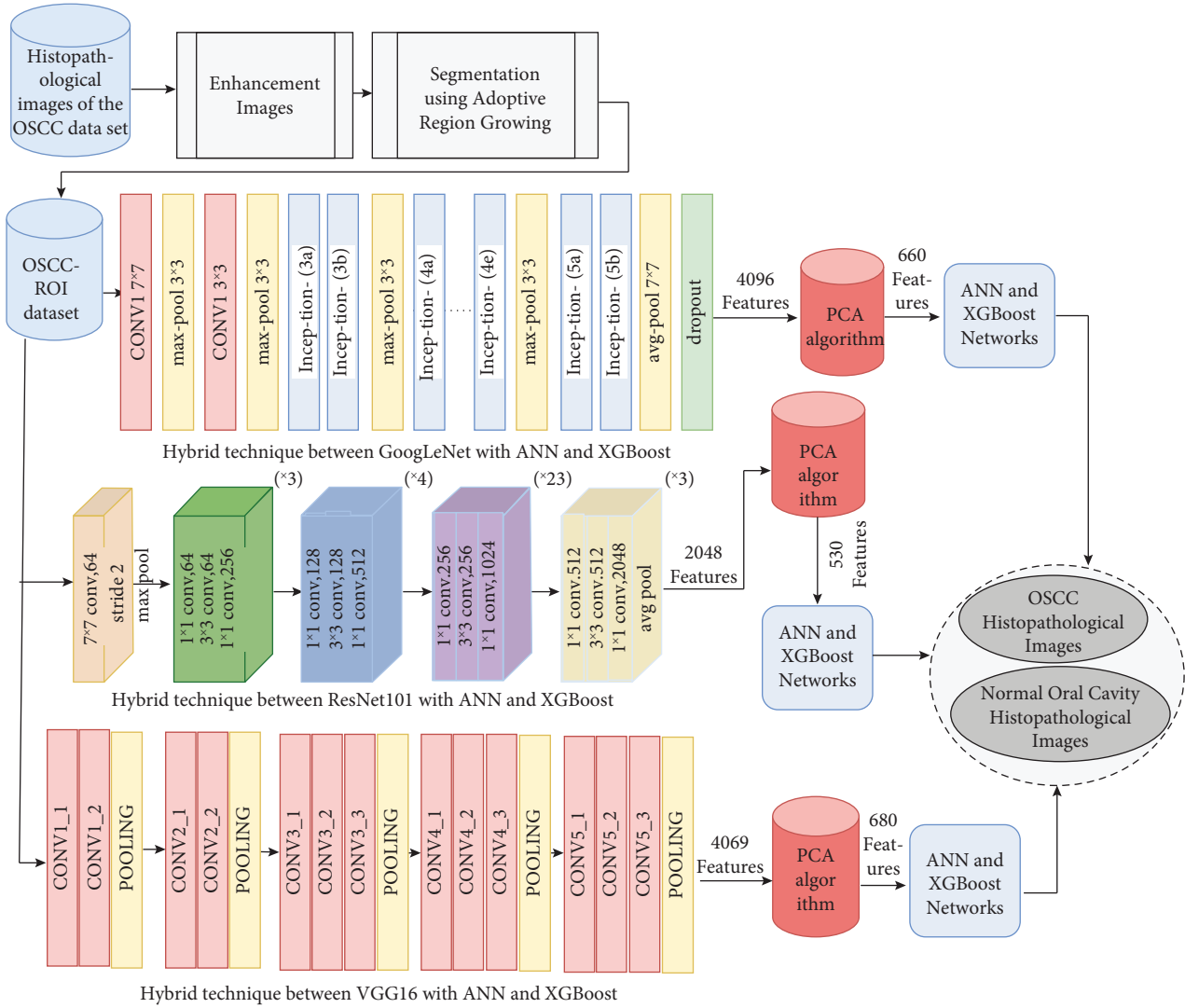


FIGURE 4: Histopathological image analysis strategy for diagnosing an oral cancer dataset by a hybrid approach based on the ARG algorithm.

output classes, while the columns represent the target classes. The confusion matrix contains correctly classified samples in the main diagonal called true positive (TP). At the same time, other cells represent samples incorrectly classified as either false positive (FP) or false negative (FN) [51].

$$\text{AUC} = \frac{\text{TP Rate}}{\text{FP Rate}} * 100\%, \quad (6)$$

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}} * 100\%, \quad (7)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} * 100\%, \quad (8)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} * 100\%, \quad (9)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} * 100. \quad (10)$$

4.3. Augmentation Data Method. Artificial intelligence techniques require huge images to train, especially deep learning models. When AI techniques are trained on a dataset with few images, they perform poorly when tested with new samples. Also, the unbalanced data set represents a challenge for evaluating systems because the accuracy will be large with the majority of classes. Therefore, the data augmentation method was applied since the OSCC data set lacks sufficient images to train the systems and is unbalanced [52]. This method solves both challenges in parallel, where the images of the dataset's classes are increased for the systems to train well, and the images between the classes are increased unevenly. The method increases minority classes by more than majority classes. The method works through many operations, such as rotating images in multiple directions, shifting them up and down, flipping, and others [53]. It is noted from Table 2 that the OSCC-100x dataset has increased by 18 artificial images for normal class images, while the OSCC class has artificially increased by three artificial images for each image. In the OSCC-400x dataset,

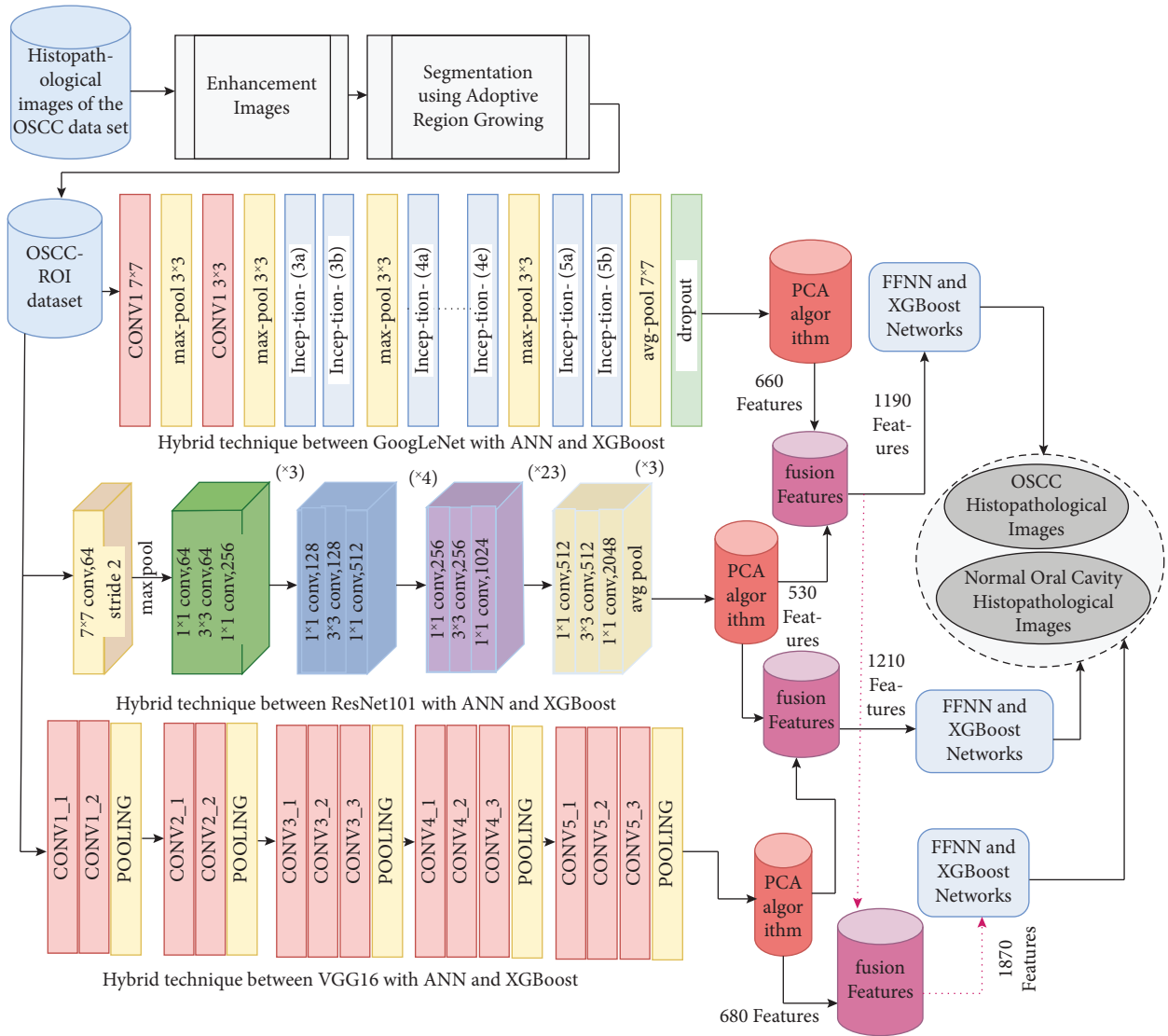


FIGURE 5: Histopathological image analysis strategy for diagnosing an oral cancer dataset by a hybrid approach based on a fusion of features of CNN models.

TABLE 1: Dividing two OSCC datasets into training and testing phases.

Datasets	OSCC-100x			OSCC-400x		
	80% (80 : 20)		Testing 20%	80% (80 : 20)		Testing 20%
Phase	Training (80%)	Validation (20%)		Training (80%)	Validation (20%)	
Normal oral cavity histopathological	57	14	18	129	32	40
OSCC histopathological	281	70	88	317	79	99

TABLE 2: Data augmentation for the OSCC-100x and OSCC-400x datasets.

Datasets	OSCC-100x		OSCC-400x	
	Training dataset		Training dataset	
Phase	Normal oral cavity histopathological	OSCC histopathological	Normal oral cavity histopathological	OSCC histopathological
Before augmentation	57	281	129	317
After augmentation	1,083	1124	903	951

the quantity of normal class images was artificially amplified sixfold for each image, whereas the OSCC class images were artificially amplified twofold.

4.4. Strength, Limitation, and Impact/Significance

4.4.1. Strengths. The use of hybrid systems is based on the features of fused CNN models: the research demonstrates the application of hybrid systems based on the features of fused CNN models to analyze histopathological images in the early detection of OSCC. This approach uses the power of deep learning to extract meaningful features from images and combine them into feature vectors so that feature vectors become highly representative of each image, enabling accurate classification and diagnosis. Improved early detection: by taking advantage of the hybrid systems based on the advantages of fused CNN models, the proposed method can improve the early detection of oral cancer. Early detection is crucial for timely intervention and treatment, leading to better patient outcomes and saving lives. Integrated CNN features are that incorporating CNN features allows for a more comprehensive analysis of histopathological images, capturing diverse aspects and features of OSCC lesions. This approach may enhance the accuracy and robustness of the detection and classification process.

4.4.2. Limitations. Data set limitations: The performance of the proposed method was affected by the quality and size of the available data set. If the dataset used for training and evaluation is limited in size and unbalanced between classes of the dataset, it may affect the generalizability and reliability of the results. This limitation was addressed by the data augmentation method to balance the data set and to increase its size to overcome the overfitting problem. Interpretation of results: Deep learning models, such as CNNs, are often considered black-box models, meaning that their decision-making cannot be easily explained. This lack of interpretation may present challenges in understanding the underlying factors that contribute to the classification or prognosis of OSCC. This limitation has been overcome by hybrid systems between CNNs to extract fused features and classify them by machine learning algorithms.

4.4.3. Real-World Implementation Challenges. While the paper focuses on the development of AI techniques for pathological image analysis, real-world implementation of these systems may face challenges related to integration into existing healthcare infrastructure, regulatory considerations, and acceptance by professionals.

4.4.4. Impact/Significance in Real-Life Scenarios. Improved diagnostic accuracy: The use of hybrid systems based on the features of the built-in CNN models has the potential to enhance the diagnostic accuracy of OSCC. This can help histopathologists, oncologists, and healthcare

professionals to make more informed decisions, reduce diagnostic errors, and improve patient outcomes. Early intervention and treatment: By enabling early detection of OSCC, the proposed method can facilitate timely intervention and treatment. This can improve the outlook, increase survival rates, and improve the quality of life for patients.

4.4.5. Supporting Histopathologists and Oncologists. AI technologies can be valuable tools to support histopathologists and oncologists in their clinical practice. By automating certain aspects of analysis, such as feature extraction and classification, AI systems can help reduce the workload of oncologists and histologists, allowing them to focus on more complex cases and provide more personalized care to patients.

4.5. Results of Pretrained CNN Models. This section discusses the results of pretrained GoogLeNet, ResNet101, and VGG16 models with the ImageNet dataset having 1.2 million images from more than 1000 classes. Unfortunately, ImageNet does not include medical images, so the pretrained CNN models will transfer the experience gained to perform a new OSCC dataset classification task. The input layers of GoogLeNet, ResNet101, and VGG16 models receive histological images of oral cancer and transfer them to convolutional layers for analysis. The fully connected layers classify the feature maps, and the SoftMax activation function labels each image to a specific class.

Table 3 and Figure 6 discuss the results of GoogLeNet, ResNet101, and VGG16 pretrained models for classifying two OSCC-100x and OSCC-400x oral cancer datasets. First, with the OSCC-100x dataset, GoogLeNet achieved an AUC of 69.1%, accuracy of 82.1%, sensitivity of 72.25%, precision of 69.1%, and specificity of 72.35%. The ResNet101 yielded an AUC of 70.65%, accuracy of 81.1%, sensitivity of 69.15%, precision of 67.25%, and specificity of 69.15%. The VGG16 yielded an AUC of 65.45%, accuracy of 78.3%, sensitivity of 62.7%, precision of 62.1%, and specificity of 62.55%.

Second, with the OSCC-400x dataset, GoogLeNet achieved an AUC of 74.2%, accuracy of 77.7%, sensitivity of 71.1%, precision of 72.8%, and specificity of 71.2%. The ResNet101 yielded an AUC of 74%, accuracy of 77%, sensitivity of 69.95%, precision of 71.85%, and specificity of 69.8%. The VGG16 yielded an AUC of 77.5%, accuracy of 79.9%, sensitivity of 72.05%, precision of 76%, and specificity of 72.55%.

4.6. Results of CNN Based on ARG Algorithm. This section discusses the results of pretrained GoogLeNet, ResNet101, and VGG16 models based on the ARG algorithm. In this strategy, the images are first improved, extracted ROI, and saved in new folders OSCC-100x-ROI and OSCC-400x-ROI. The input layers of the GoogLeNet, ResNet101, and VGG16 models receive the ROI of oral cancer and pass them to the

TABLE 3: Results of pretrained CNN models for classification of the two oral cancer datasets.

Datasets	Models	Classes	AUC (%)	Accuracy (%)	Sensitivity (%)	Precision (%)	Specificity (%)
OSCC-100x	GoogLeNet	Normal oral cavity histopathological	61.3	55.6	56.2	47.6	88.3
		OSCC histopathological	76.9	87.5	88.3	90.6	56.4
		Average ratio	69.1	82.1	72.25	69.1	72.35
OSCC-400x	ResNet101	Normal oral cavity histopathological	65.5	50	50.2	45	87.9
		OSCC histopathological	75.8	87.5	88.1	89.5	50.4
		Average ratio	70.65	81.1	69.15	67.25	69.15
OSCC-100x	VGG16	Normal oral cavity histopathological	55.6	38.9	39.3	36.8	85.8
		OSCC histopathological	75.3	86.4	86.1	87.4	39.3
		Average ratio	65.45	78.3	62.7	62.1	62.55
OSCC-400x	GoogLeNet	Normal oral cavity histopathological	69.5	55	55.4	62.9	87.2
		OSCC histopathological	78.9	86.9	86.8	82.7	55.2
		Average ratio	74.2	77.7	71.1	72.8	71.2
OSCC-400x	ResNet101	Normal oral cavity histopathological	68.2	52.5	52.8	61.8	86.7
		OSCC histopathological	79.8	86.9	87.1	81.9	52.9
		Average ratio	74	77	69.95	71.85	69.8
OSCC-400x	VGG16	Normal oral cavity histopathological	71.8	55	54.6	68.8	90.4
		OSCC histopathological	83.2	89.9	89.5	83.2	54.7
		Average ratio	77.5	79.9	72.05	76	72.55

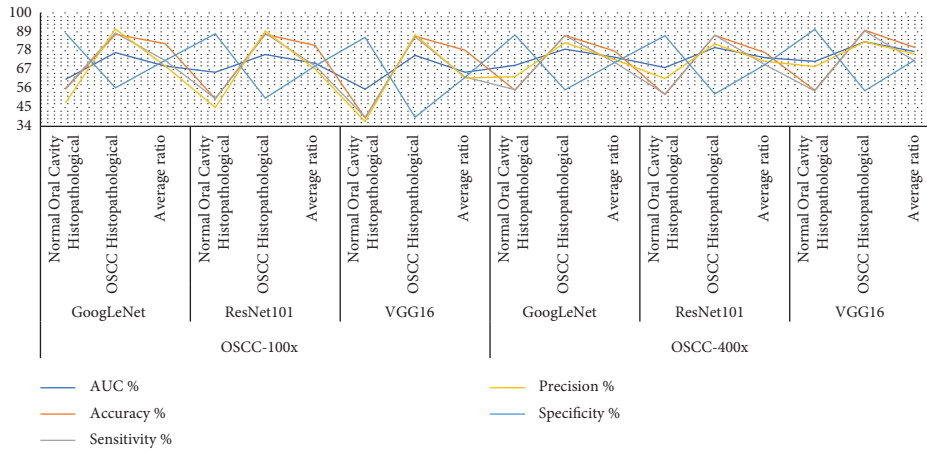


FIGURE 6: Display of results of pretrained CNN models for classification of the two oral cancer datasets.

convolutional layers for analysis. Fully connected layers classify feature maps, and the SoftMax activation function labels each image into a specific class.

Table 4 discusses the results of GoogLeNet, ResNet101, and VGG16 pretrained models based on the ARG algorithm for classifying OSCC-100x and OSCC-400x oral cancer datasets. First, with the OSCC-100x dataset, GoogLeNet achieved an AUC of 77.95%, accuracy of 90.6%, sensitivity of 76.75%, precision of 87.4%, and specificity of 77.15%. The ResNet101 yielded an AUC of 77.6%, accuracy of 89.6%, sensitivity of 76.45%, precision of 84.15%, and specificity of 76.85%. The VGG16 yielded an AUC of 81.9%, accuracy of 88.7%, sensitivity of 86.65%, precision of 79.4%, and specificity of 86.5%.

Second, with the OSCC-400x dataset, GoogLeNet achieved an AUC of 82.4%, accuracy of 91.4%, sensitivity of 88.45%, precision of 89.95%, and specificity of 88.4%. The ResNet101 yielded an AUC of 83.4%, accuracy of 92.8%, sensitivity of 90.55%, precision of 91.8%, and specificity of 90.75%. The VGG16 yielded an AUC of 81.6%, accuracy of 92.1%, sensitivity of 89.25%, precision of 91.15%, and specificity of 88.75%.

Figure 7 presents the implementation performance of GoogLeNet, ResNet101, and VGG16 pretrained models based on the ARG algorithm for classifying two OSCC-100x and OSCC-400x oral cancer datasets.

4.7. Results of Hybrid Systems. The section presents the findings of hybrid strategies between CNN (GoogLeNet, ResNet101, and VGG16) with ANN and XGBoost algorithms for classifying histopathological images of oral cancer with magnification factors of 100x and 400x. The histopathological images were improved, then ROI regions were segmented, and the ROI was sent to CNN networks for analysis and obtaining features. PCA worked after CNN models to reduce the high dimensions and sent the features to ANN and XGBoost algorithms to classify them and distinguish OSCC tissues from normal.

Table 5 and Figure 8 show the results of the hybrid strategy of CNN models with ANN and XGBoost algorithms for histopathological analysis for diagnosing the OSCC-100x dataset.

The GoogLeNet with ANN achieved an AUC of 88.05%, accuracy of 92.5%, sensitivity of 86.5%, precision of 92.5%, and specificity of 86.5%. ResNet101 with ANN yielded an AUC of 85.1%, accuracy of 91.5%, sensitivity of 82.3%, precision of 86.7%, and specificity of 82%. VGG16 with ANN yielded an AUC of 90.35%, accuracy of 93.4%, sensitivity of 91.7%, precision of 86.9%, and specificity of 91.55%.

The GoogLeNet with XGBoost achieved an AUC of 92.2%, accuracy of 92.5%, sensitivity of 92.85%, precision of 84.8%, and specificity of 93.15%. ResNet101 with XGBoost yielded an AUC of 92.7%, accuracy of 94.3%, sensitivity of 91.9%, precision of 88.85%, and specificity of 92.25%. VGG16 with XGBoost yielded an AUC of 91.1%, accuracy of 91.5%, sensitivity of 86.05%, precision of 84.55%, and specificity of 86.25%.

Table 6 and Figure 9 show the results of the hybrid strategy of CNN models with ANN and XGBoost algorithms for histopathological analysis for diagnosing the OSCC-400x dataset.

The GoogLeNet with ANN achieved an AUC of 91.7%, accuracy of 92.1%, sensitivity of 90.1%, precision of 92.1%, and specificity of 90%. ResNet101 with ANN yielded an AUC of 93.85%, accuracy of 94.2%, sensitivity of 94.15%, precision of 92.5%, and specificity of 94.15%. VGG16 with ANN yielded an AUC of 93.15%, accuracy of 93.5%, sensitivity of 92.15%, precision of 92.35%, and specificity of 92%.

The GoogLeNet with XGBoost achieved an AUC of 93.2%, accuracy of 95.7%, sensitivity of 93.55%, precision of 96.15%, and specificity of 93.5%. ResNet101 with XGBoost yielded an AUC of 93%, accuracy of 92.8%, sensitivity of 91.55%, precision of 91.2%, and specificity of 91.6%. VGG16 with XGBoost yielded an AUC of 94.75%, accuracy of 95.7%, sensitivity of 95%, precision of 94.75%, and specificity of 95%.

TABLE 4: Results of the pretrained CNN models based on the ARG algorithm for classifying the two oral cancer datasets.

Datasets	Models	Classes	AUC (%)	Accuracy (%)	Sensitivity (%)	Precision (%)	Specificity (%)
OSCC-100x	GoogLeNet	Normal oral cavity histopathological	72.4	55.6	97.7	83.3	98.1
		OSCC histopathological	83.5	97.7	55.8	91.5	56.2
		Average ratio	77.95	90.6	76.75	87.4	77.15
	ResNet101	Normal oral cavity histopathological	71.8	55.6	56.1	76.9	97.4
		OSCC histopathological	83.4	96.6	96.8	91.4	56.3
		Average ratio	77.6	89.6	76.45	84.15	76.85
VGG16	Normal oral cavity histopathological	78.6	83.3	83.1	62.5	89.6	
	OSCC histopathological	85.2	89.8	90.2	96.3	83.4	
	Average ratio	81.9	88.7	86.65	79.4	86.5	
OSCC-400x	GoogLeNet	Normal oral cavity histopathological	77.9	82.5	82.1	86.8	95.2
		OSCC histopathological	86.9	94.9	94.8	93.1	81.6
		Average ratio	82.4	91.4	88.45	89.95	88.4
	ResNet101	Normal oral cavity histopathological	79.2	85	84.9	89.5	96.2
		OSCC histopathological	87.6	96	96.2	94.1	85.3
		Average ratio	83.4	92.8	90.55	91.8	90.75
VGG16	Normal oral cavity histopathological	76.8	82.5	82.2	89.2	95.7	
	OSCC histopathological	86.4	96	96.3	93.1	81.8	
	Average ratio	81.6	92.1	89.25	91.15	88.75	

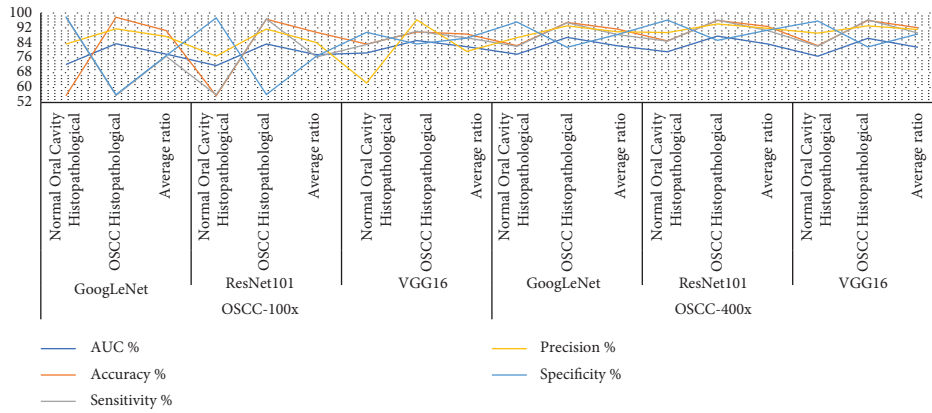


FIGURE 7: Display of results of the pretrained CNN models based on the ARG algorithm for classifying the two oral cancer datasets.

The hybrid strategy between CNN and ANN for classifying histology images for the OSCC-100x data set produces the confusion matrix shown in Figure 10. CNN with ANN yielded promising results for class-level diagnosis: GoogLeNet with ANN yielded an accuracy of 77.8% for normal oral cavity histology and 95.5% for malignant OSCC. ResNet101 with ANN yielded an accuracy of 66.7% for normal oral cavity histology and 96.6% for malignant OSCC. VGG16 with ANN yielded an accuracy of 88.9% for normal oral cavity histology and 94.3% for malignant OSCC.

The hybrid strategy between CNN and XGBoost for classifying histology images for the OSCC-100x data set produces the confusion matrix as shown in Figure 11. CNN with XGBoost yielded promising results for class-level diagnosis: GoogLeNet with XGBoost yielded an accuracy of 94.4% for normal oral cavity histology and 92% for malignant OSCC. ResNet101 with XGBoost yielded an accuracy of 88.9% for normal oral cavity histology and 95.5% for malignant OSCC. VGG16 with XGBoost yielded an accuracy of 77.8% for normal oral cavity histology and 94.3% for malignant OSCC.

The hybrid strategy between CNN and ANN for classifying histology images for the OSCC-400x data set produces the confusion matrix as shown in Figure 12. CNN with ANN yielded promising results for class-level diagnosis: GoogLeNet with ANN yielded an accuracy of 85% for normal oral cavity histology and 94.9% for malignant OSCC. ResNet101 with ANN yielded an accuracy of 92.5% for normal oral cavity histology and 94.9% for malignant OSCC. VGG16 with ANN yielded an accuracy of 87.5% for normal oral cavity histology and 96% for malignant OSCC.

The hybrid strategy between CNN and XGBoost for classifying histology images for the OSCC-400x data set produces the confusion matrix as shown in Figure 13. CNN with XGBoost yielded promising results for class-level diagnosis: GoogLeNet with XGBoost yielded an accuracy of 87.5% for normal oral cavity histology and 99% for malignant OSCC. ResNet101 with XGBoost yielded an accuracy of 87.5% for normal oral cavity histology and 94.9% for malignant OSCC. VGG16 with XGBoost yielded an accuracy of 92.5% for normal oral cavity histology and 97% for malignant OSCC.

4.8. Results of Hybrid Systems Based on Fusion Features of CNN. The section presents the results of hybrid strategies based on fused CNN features and their classification by ANN and XGBoost networks to classify histopathological images of oral cancer with magnification factors of 100x and 400x. The histopathological images were optimized, then the ROI were sectioned, and the ROI were sent to CNNs for analysis and obtaining features. PCA worked after CNN models to reduce the high dimensionality. The features of the CNN models were serially combined to obtain the most efficient and robust feature vectors as follows: GoogLeNet-ResNet101, ResNet101-VGG16, GoogLeNet-VGG16, and GoogLeNet-ResNet101-VGG16. The fused features were sent to ANN and XGBoost algorithms to classify and differentiate OSCC tissues from normal.

Table 7 and Figure 14 show the results of the hybrid strategy of CNN models with ANN and XGBoost algorithms based on CNN fusion features for histopathological analysis for diagnosing the OSCC-100x dataset.

The GoogLeNet-ResNet101 with ANN achieved an AUC of 95.3%, accuracy of 96.2%, sensitivity of 95.5%, precision of 91.9%, and specificity of 95.5%. ResNet101-VGG16 with ANN yielded an AUC of 94.45%, accuracy of 96.2%, sensitivity of 93.45%, precision of 93.3%, and specificity of 93.65%. GoogLeNet-VGG16 with ANN yielded an AUC of 94.2%, accuracy of 95.3%, sensitivity of 90.5%, precision of 92.4%, and specificity of 90.45%. GoogLeNet-ResNet101-VGG16 with ANN yielded an AUC of 98.05%, accuracy of 99.1%, sensitivity of 99.25%, precision of 97.35%, and specificity of 99.5%.

The GoogLeNet-ResNet101 with XGBoost achieved an AUC of 96.05%, accuracy of 95.3%, sensitivity of 94.065%, precision of 89.9%, and specificity of 94.65%. ResNet101-VGG16 with XGBoost yielded an AUC of 97.9%, accuracy of 98.1%, sensitivity of 98.95%, precision of 95%, and specificity of 98.65%. GoogLeNet-VGG16 with XGBoost yielded an AUC of 96.7%, accuracy of 96.2%, sensitivity of 95.4%, precision of 91.9%, and specificity of 95.5%. GoogLeNet-ResNet101-VGG16 with XGBoost yielded an AUC of 99.1%, accuracy of 98.1%, sensitivity of 98.95%, precision of 95%, and specificity of 98.95%.

TABLE 5: Results of hybrid systems classification of an oral cancer dataset for a magnification factor of 100x.

Classifiers	Models for features	Classes	AUC (%)	Accuracy (%)	Sensitivity (%)	Precision (%)	Specificity (%)
ANN	GoogLeNet	Normal oral cavity histopathological	84.6	77.8	78.1	77.8	95.2
		OSCC histopathological	91.5	95.5	94.9	95.5	77.8
		Average ratio	88.05	92.5	86.5	92.5	86.5
	ResNet101	Normal oral cavity histopathological	77.6	66.7	67.4	80	96.7
		OSCC histopathological	92.6	96.6	97.2	93.4	67.3
		Average ratio	85.1	91.5	82.3	86.7	82
VGG16	Normal oral cavity histopathological	86.9	88.9	89.3	76.2	94.2	
	OSCC histopathological	93.8	94.3	94.1	97.6	88.9	
	Average ratio	90.35	93.4	91.7	86.9	91.55	
XGBoost	GoogLeNet	Normal oral cavity histopathological	92.8	94.4	93.8	70.8	92.2
		OSCC histopathological	91.6	92	91.9	98.8	94.1
		Average ratio	92.2	92.5	92.85	84.8	93.15
	ResNet101	Normal oral cavity histopathological	90.8	88.9	89.1	80	95.2
		OSCC histopathological	94.6	95.5	94.7	97.7	89.3
		Average ratio	92.7	94.3	91.9	88.85	92.25
VGG16	Normal oral cavity histopathological	88.6	77.8	77.8	73.7	94.2	
	OSCC histopathological	93.6	94.3	94.3	95.4	78.3	
	Average ratio	91.1	91.5	86.05	84.55	86.25	

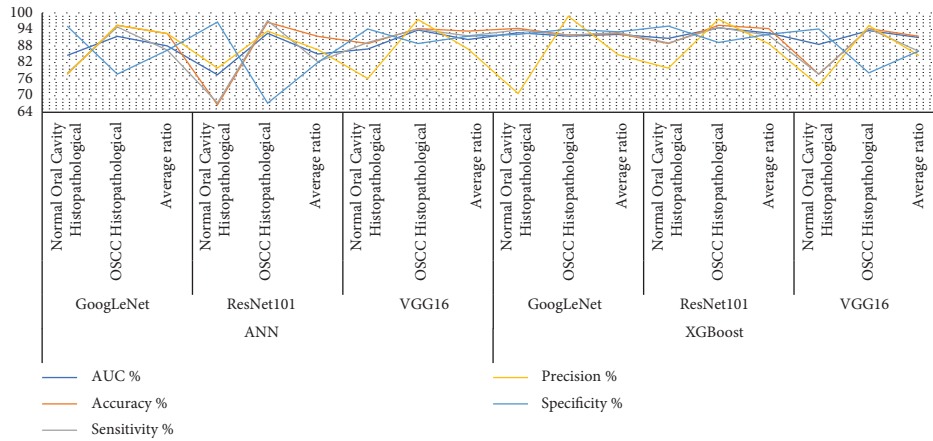


FIGURE 8: Display of results of hybrid systems classification of an oral cancer dataset for a magnification factor of 100x.

Table 8 and Figure 15 show the results of the hybrid strategy of CNN models with ANN and XGBoost algorithms based on CNN fusion features for histopathological analysis for diagnosing the OSCC-400x dataset. The GoogLeNet-ResNet101 with ANN achieved an AUC of 97.95%, accuracy of 98.6%, sensitivity of 98.1%, precision of 98.25%, and specificity of 98%. ResNet101-VGG16 with ANN yielded an AUC of 97.55%, accuracy of 97.8%, sensitivity of 97.65%, precision of 97.05%, and specificity of 97.3%. GoogLeNet-VGG16 with ANN yielded an AUC of 97.45%, accuracy of 97.1%, sensitivity of 97%, precision of 95.95%, and specificity of 97%. GoogLeNet-ResNet101-VGG16 with ANN yielded an AUC of 98.85%, accuracy of 99.3%, sensitivity of 98.2%, precision of 99.5%, and specificity of 98.35%.

The GoogLeNet-ResNet101 with XGBoost achieved an AUC of 95.25%, accuracy of 95.7%, sensitivity of 94%, precision of 95.35%, and specificity of 94.2%. ResNet101-VGG16 with XGBoost yielded an AUC of 97.45%, accuracy of 96.4%, sensitivity of 95.6%, precision of 95.8%, and specificity of 95.5%. GoogLeNet-VGG16 with XGBoost yielded an AUC of 97.2%, accuracy of 97.1%, sensitivity of 95.9%, precision of 97.2%, and specificity of 95.85%. GoogLeNet-ResNet101-VGG16 with XGBoost yielded an AUC of 98.25%, accuracy of 97.8%, sensitivity of 97.15%, precision of 97.7%, and specificity of 96.85%.

The hybrid strategy between CNN and ANN based on CNN fusion features for classifying histology images for the OSCC-100x data set produces the confusion matrix as shown in Figure 16. CNN fusion features with ANN yielded promising results for class-level diagnosis: GoogLeNet-ResNet101 with ANN yielded an accuracy of 94.4% for normal oral cavity histology and 96.6% for malignant OSCC. ResNet101-VGG16 with ANN yielded an accuracy of 88.9% for normal oral cavity histology and 97.7% for malignant OSCC. GoogLeNet-VGG16 with ANN yielded an accuracy of 83.3% for normal oral cavity histology and 97.7% for malignant OSCC. GoogLeNet-ResNet101-VGG16 with ANN yielded an accuracy of 100% for normal oral cavity histology and 98.9% for malignant OSCC.

A hybrid strategy between CNN and XGBoost based on CNN fusion features for classifying histology images for the OSCC-100x data set produces the confusion matrix as shown in Figure 17. CNN fusion features with XGBoost yielded promising results for class-level diagnosis: GoogLeNet-ResNet101 with XGBoost yielded an accuracy of 94.4% for normal oral cavity histology and 95.5% for malignant OSCC. ResNet101-VGG16 with XGBoost yielded an accuracy of 100% for normal oral cavity histology and 97.7% for malignant OSCC. GoogLeNet-VGG16 with XGBoost yielded an accuracy of 94.9% for normal oral cavity histology and 96.6% for malignant OSCC. GoogLeNet-ResNet101-VGG16 with XGBoost yielded an accuracy of 94.4% for normal oral cavity histology and 98.9% for malignant OSCC.

The hybrid strategy between CNN and ANN based on CNN fusion features for classifying histology images for the OSCC-400x data set produces the confusion matrix as shown in Figure 18. CNN fusion features with ANN yielded promising results for class-level diagnosis: GoogLeNet-ResNet101 with ANN yielded an accuracy of 97.5% for normal oral cavity histology and 99% for malignant OSCC. ResNet101-VGG16 with ANN yielded an accuracy of 97.5% for normal oral cavity histology and 98% for malignant OSCC. GoogLeNet-VGG16 with ANN yielded an accuracy of 97.5% for normal oral cavity histology and 97% for malignant OSCC. GoogLeNet-ResNet101-VGG16 with ANN yielded an accuracy of 97.5% for normal oral cavity histology and 100% for malignant OSCC.

A hybrid strategy between CNN and XGBoost based on CNN fusion features for classifying histology images for the OSCC-400x data set produces the confusion matrix as shown in Figure 19. CNN fusion features with XGBoost yielded promising results for class-level diagnosis: GoogLeNet-ResNet101 with XGBoost yielded an accuracy of 90% for normal oral cavity histology and 98% for malignant OSCC. ResNet101-VGG16 with XGBoost yielded an accuracy of 92.5% for normal oral cavity histology and 98% for malignant OSCC. GoogLeNet-VGG16 with XGBoost yielded an accuracy of 92.5% for normal oral cavity histology

TABLE 6: Results of hybrid systems classification of an oral cancer dataset for a magnification factor of 400x.

Classifiers	Models for features	Classes	AUC (%)	Accuracy (%)	Sensitivity (%)	Precision (%)	Specificity (%)
ANN	GoogLeNet	Normal oral cavity histopathological	90.2	85	84.9	87.2	95.2
		OSCC histopathological	93.2	94.9	95.3	94	84.8
		Average ratio	91.7	92.1	90.1	92.1	90
	ResNet101	Normal oral cavity histopathological	93.6	92.5	93.2	88.1	95.2
		OSCC histopathological	94.1	94.9	95.1	96.9	93.1
		Average ratio	93.85	94.2	94.15	92.5	94.15
VGG16	Normal oral cavity histopathological	91.6	87.5	88.2	89.7	96.2	
	OSCC histopathological	94.6	96	96.1	95	87.8	
	Average ratio	93.1	93.5	92.15	92.35	92	
XGBoost	GoogLeNet	Normal oral cavity histopathological	90.8	87.5	87.9	97.2	98.8
		OSCC histopathological	95.6	99	99.2	95.1	88.2
		Average ratio	93.2	95.7	93.55	96.15	93.5
	ResNet101	Normal oral cavity histopathological	91.2	87.5	88.3	87.5	95.1
		OSCC histopathological	94.8	94.9	94.8	94.9	88.1
		Average ratio	93	92.8	91.55	91.2	91.6
VGG16	Normal oral cavity histopathological	93.4	92.5	93.1	92.5	97.1	
	OSCC histopathological	96.1	97	96.9	97	92.9	
	Average ratio	94.75	95.7	95	94.75	95	

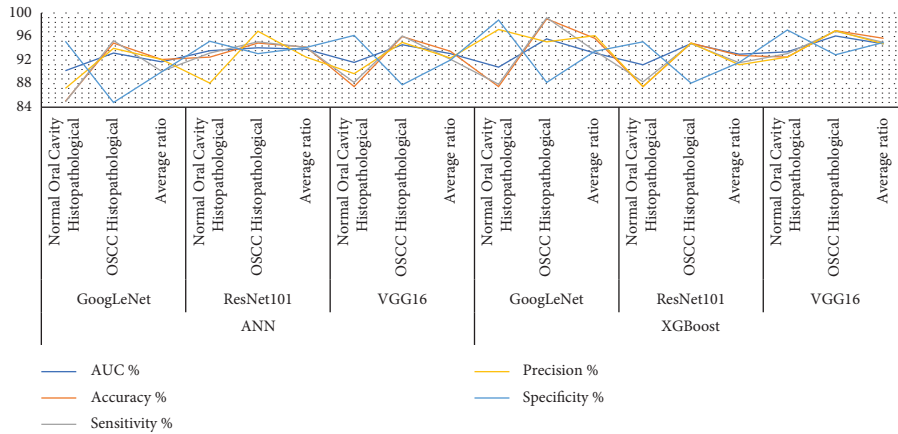


FIGURE 9: Display of results of hybrid systems classification of an oral cancer dataset for a magnification factor of 400x.

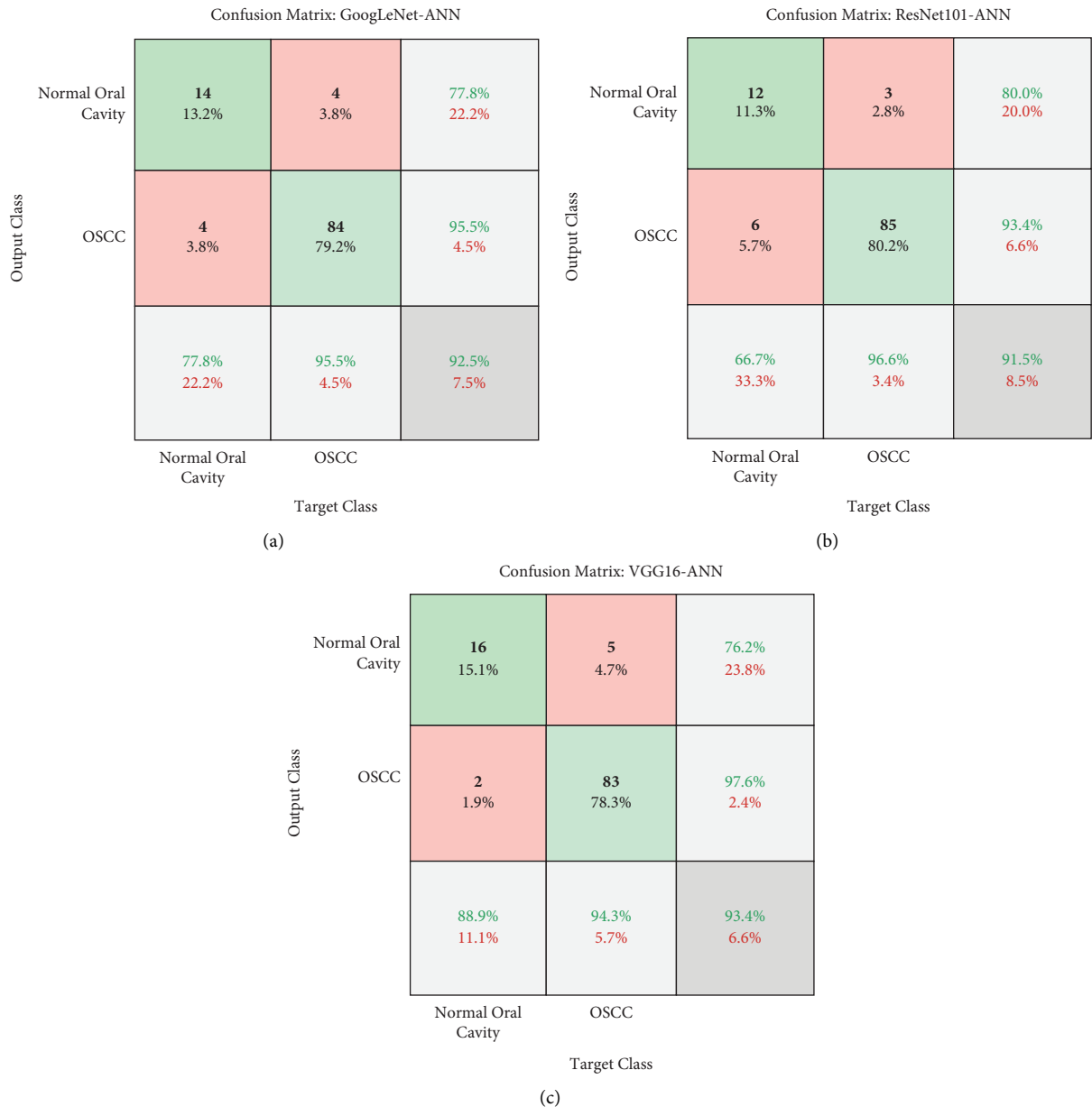


FIGURE 10: Confusion matrix for displaying the results of the classification of hybrid systems for an oral cancer dataset for a magnification factor of 100x by techniques: (a) GoogLeNet-ANN, (b) ResNet101-ANN, and (c) VGG16-ANN.

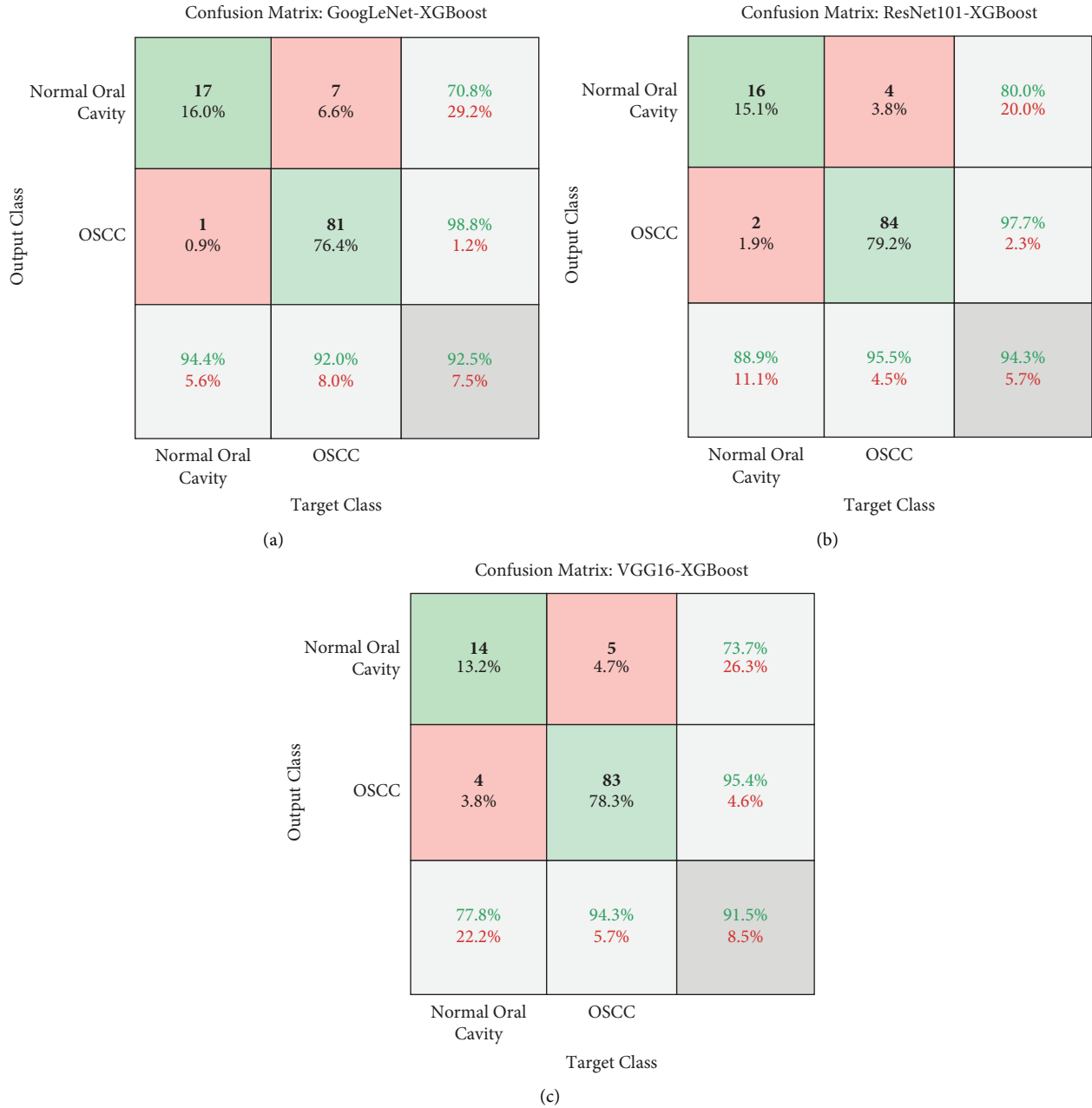


FIGURE 11: Confusion matrix for displaying the results of the classification of hybrid systems for an oral cancer dataset for a magnification factor of 100x by techniques: (a) GoogLeNet-XGBoost, (b) ResNet101-XGBoost, and (c) VGG16-XGBoost.

and 99% for malignant OSCC. GoogLeNet-ResNet101-VGG16 with XGBoost yielded an accuracy of 95% for normal oral cavity histology and 99% for malignant OSCC.

5. Discussion of the Results of the Systems

The OSCC is one of the most common types of cancer and leads to death in the late stages. Normal oral cavity tissues are similar to malignant tissues in the early stages, which makes it difficult to distinguish between them by manual diagnosis. Automatic systems help doctors in early diagnosis to receive appropriate treatment. Several hybrid models

based on the ARG segmentation algorithm and fused CNN features have been developed in this work.

The first model is for diagnosing histopathological images of two OSCC-100x and OSCC-400x datasets by pre-trained GoogLeNet, ResNet101, and VGG16 models. First, with the OSCC-100x data set, the GoogLeNet, ResNet101, and VGG16 models achieved an accuracy of 82.1%, 81.1%, and 78.3%, respectively. Second, with the OSCC-400x data set, the GoogLeNet, ResNet101, and VGG16 models reached an accuracy of 77.7%, 77%, and 79.9%, respectively.

The second model is for diagnosing histopathological images of two OSCC-100x and OSCC-400x datasets by

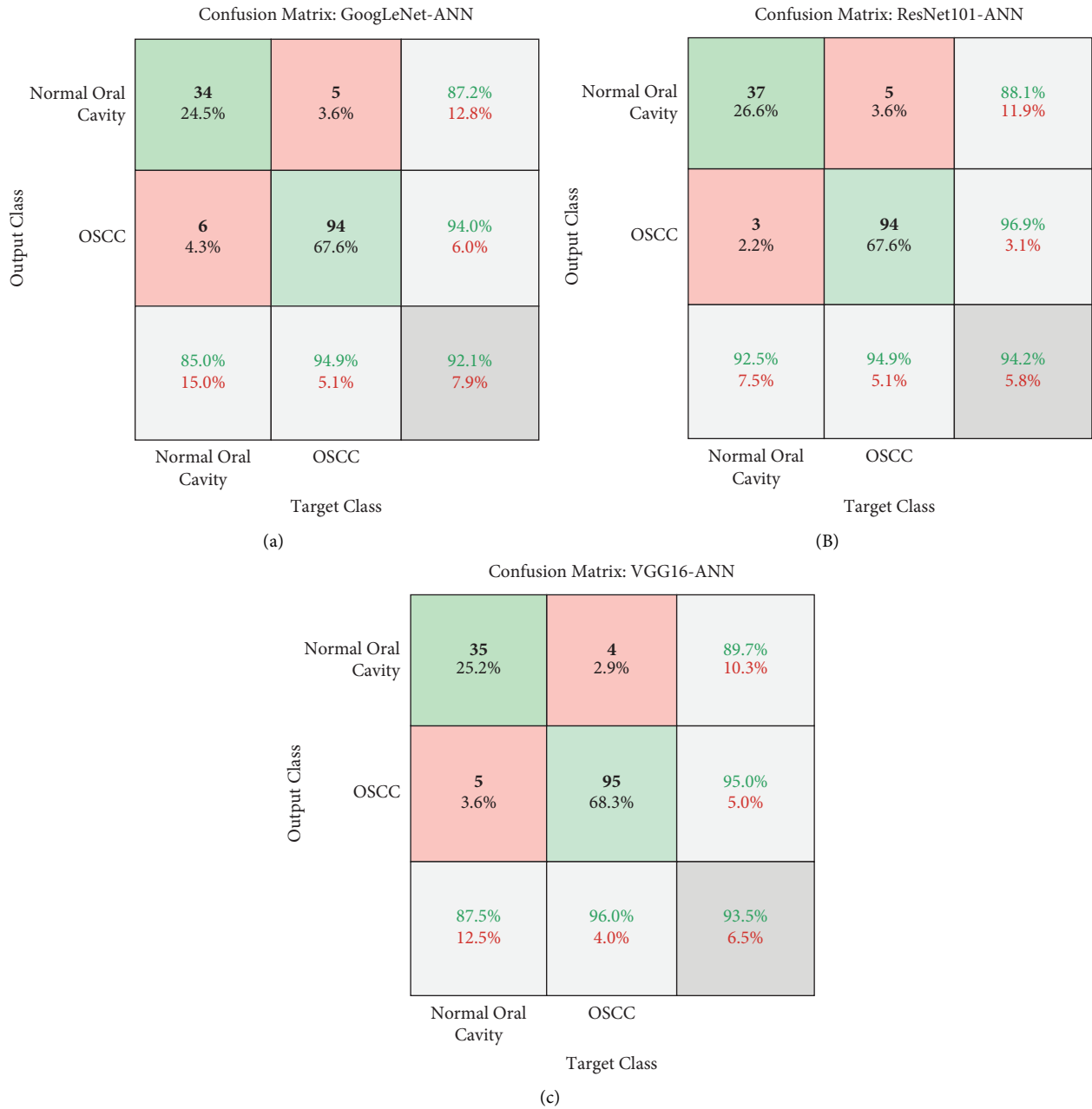


FIGURE 12: Confusion matrix for displaying the results of the classification of hybrid systems for an oral cancer dataset for a magnification factor of 400x by techniques: (a) GoogLeNet-ANN, (B) ResNet101-ANN, and (c) VGG16-ANN.

pretrained GoogLeNet, ResNet101, and VGG16 models based on the ARG segmentation algorithm. First, with the OSCC-100x dataset, the GoogLeNet, ResNet101, and VGG16 models achieved an accuracy of 90.9%, 89.6%, and 88.7%, respectively. Second, with the OSCC-400x data set, the GoogLeNet, ResNet101, and VGG16 models reached an accuracy of 91.4%, 92.8%, and 92.1%, respectively.

The third model is for diagnosing histopathological images of two OSCC-100x and OSCC-400x datasets by hybrid models between CNN (GoogLeNet, ResNet101, and VGG16) and ANN and XGBoost based on the ARG segmentation algorithm. First, with the OSCC-100x dataset, the CNN-ANN reached an accuracy of 92.5%, 91.5%, and 93.4%

for the models GoogLeNet-ANN, ResNet101-ANN, and VGG16-ANN, respectively, while GoogLeNet-XGBoost, ResNet101-XGBoost, and VGG16-XGBoost achieved an accuracy of 92.5%, 94.3%, and 91.5%, respectively. Second, with the OSCC-400x data set, the CNN with ANN reached an accuracy of 92.1%, 94.2%, and 93.5% for the models GoogLeNet-ANN, ResNet101-ANN, and VGG16-ANN, respectively, while GoogLeNet-XGBoost, ResNet101-XGBoost, and VGG16-XGBoost achieved an accuracy of 95.7%, 92.8%, and 95.7%, respectively.

The fourth model is for diagnosing histopathological images of two OSCC-100x and OSCC-400x datasets by hybrid models between CNN (GoogLeNet, ResNet101, and

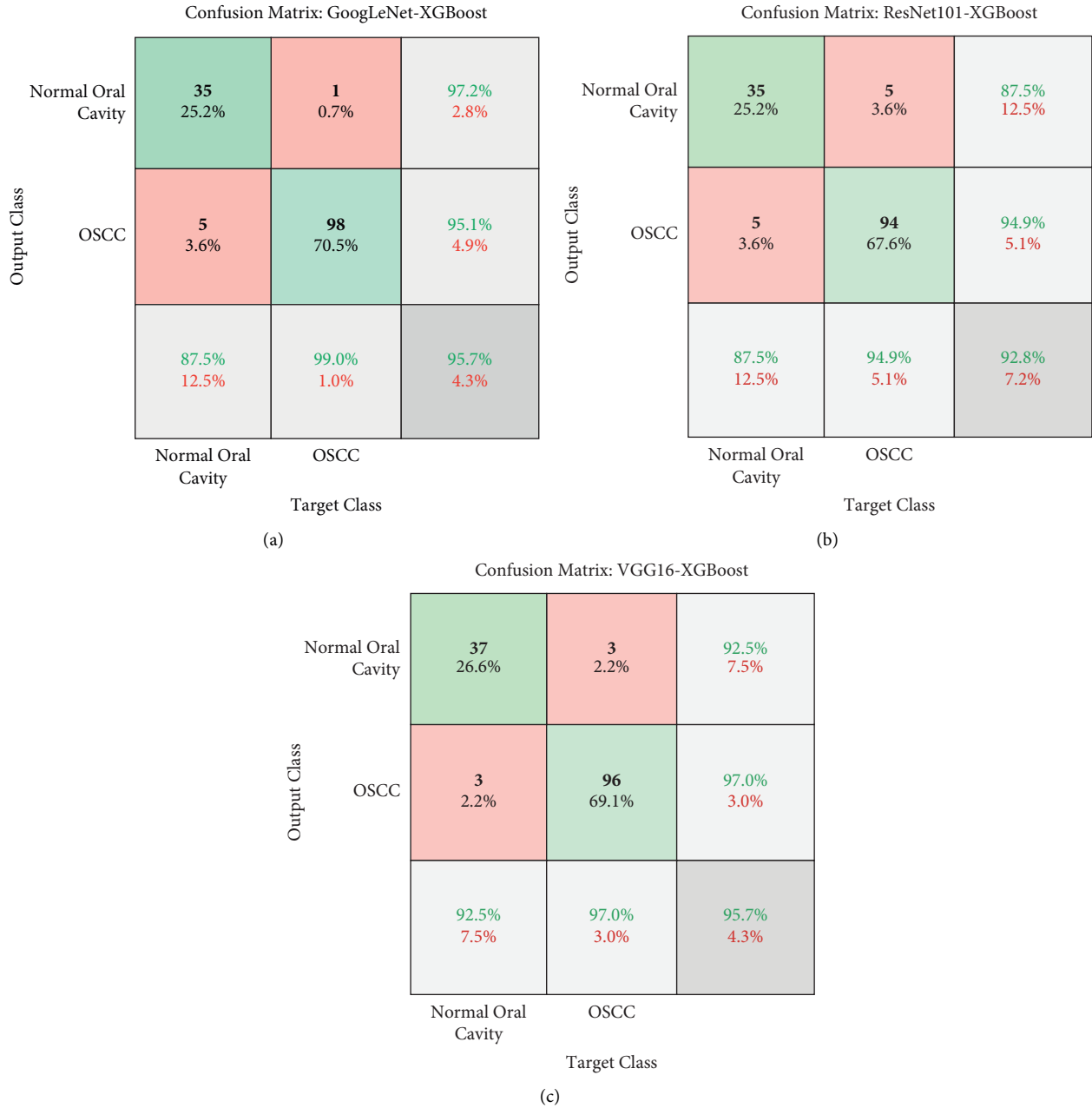


FIGURE 13: Confusion matrix for displaying the results of the classification of hybrid systems for an oral cancer dataset for a magnification factor of 400x by techniques: (a) GoogLeNet-XGBoost, (b) ResNet101-XGBoost, and (c) VGG16-XGBoost.

VGG16) and ANN and XGBoost based on feature extraction from regions of interest and fusion between CNN models. First, with the OSCC-100x dataset, ANN with fused features of CNN achieved an accuracy of 96.2%, 96.2%, 95.3%, and 99.1% for models GoogLeNet-ResNet101-ANN, ResNet101-VGG16-ANN, GoogLeNet-VGG16-ANN, and GoogLeNet-ResNet101-VGG16-ANN whereas XGBoost with fused features of CNN achieved an accuracy of 95.3%, 98.1%, 96.2%, and 98.1% for GoogLeNet-ResNet101-XGBoost, ResNet101-VGG16-XGBoost, GoogLeNet-VGG16-XGBoost, and GoogLeNet-ResNet101-VGG16-XGBoost.

Second, with the OSCC-400x dataset, ANN with fused features of CNN achieved an accuracy of 98.6%, 97.8%,

97.1%, and 99.3% for the models GoogLeNet-ResNet101-ANN, ResNet101-VGG16-ANN, GoogLeNet-VGG16-ANN, and GoogLeNet-ResNet101-VGG16-ANN whereas XGBoost with fused features of CNN achieved an accuracy of 95.7%, 96.4%, 97.1%, and 97.8% for GoogLeNet-ResNet101-XGBoost, ResNet101-VGG16-XGBoost, GoogLeNet-VGG16-XGBoost, and GoogLeNet-ResNet101-VGG16-XGBoost.

Table 9 and Figure 20 summarize the performance of the models for histopathology diagnostics for the OSCC-100x and OSCC-400x datasets of oral cancer. The table describes the accuracy for each model and the accuracy achieved by each model for each class diagnosis. It is noted that the

TABLE 7: Results of hybrid systems with fused features to classify an oral cancer dataset for a 100x magnification factor.

Classifiers	Hybrid features	Classes	AUC (%)	Accuracy (%)	Sensitivity (%)	Precision (%)	Specificity (%)
ANN	GoogLeNet-ResNet101	Normal oral cavity histopathological	93.8	94.4	96.8	85	97.2
		OSCC histopathological	96.8	96.6	94.2	98.8	93.8
	ResNet101-VGG16	Average ratio	95.3	96.2	95.5	91.9	95.5
		Normal oral cavity histopathological	91.8	88.9	89.3	88.9	98.1
	GoogLeNet-VGG16	OSCC histopathological	97.1	97.7	97.6	97.7	89.2
		Average ratio	94.45	96.2	93.45	93.3	93.65
	GoogLeNet-VGG16	Normal oral cavity histopathological	90.6	83.3	82.9	88.2	97.8
		OSCC histopathological	97.8	97.7	98.1	96.6	83.1
	GoogLeNet-ResNet101-VGG16	Average ratio	94.2	95.3	90.5	92.4	90.45
		Normal oral cavity histopathological	98.2	100	99.6	94.7	99.2
GoogLeNet-ResNet101-VGG16	OSCC histopathological	97.9	98.9	98.9	100	99.8	
	Average ratio	98.05	99.1	99.25	97.35	99.5	
XGBoost	GoogLeNet-ResNet101	Normal oral cavity histopathological	95.9	94.4	93.9	81	95.2
		OSCC histopathological	96.2	95.5	95.4	98.8	94.1
	ResNet101-VGG16	Average ratio	96.05	95.3	94.65	89.9	94.65
		Normal oral cavity histopathological	98.2	100	99.8	90	97.7
	GoogLeNet-VGG16	OSCC histopathological	97.6	97.7	98.1	100	99.6
		Average ratio	97.9	98.1	98.95	95	98.65
	GoogLeNet-VGG16	Normal oral cavity histopathological	95.8	94.4	93.7	85	96.8
		OSCC histopathological	97.6	96.6	97.1	98.8	94.2
	GoogLeNet-ResNet101-VGG16	Average ratio	96.7	96.2	95.4	91.9	95.5
		Normal oral cavity histopathological	99.6	100	99.8	90	98.2
GoogLeNet-ResNet101-VGG16	OSCC histopathological	98.6	97.7	98.1	100	99.7	
	Average ratio	99.1	98.1	98.95	95	98.95	

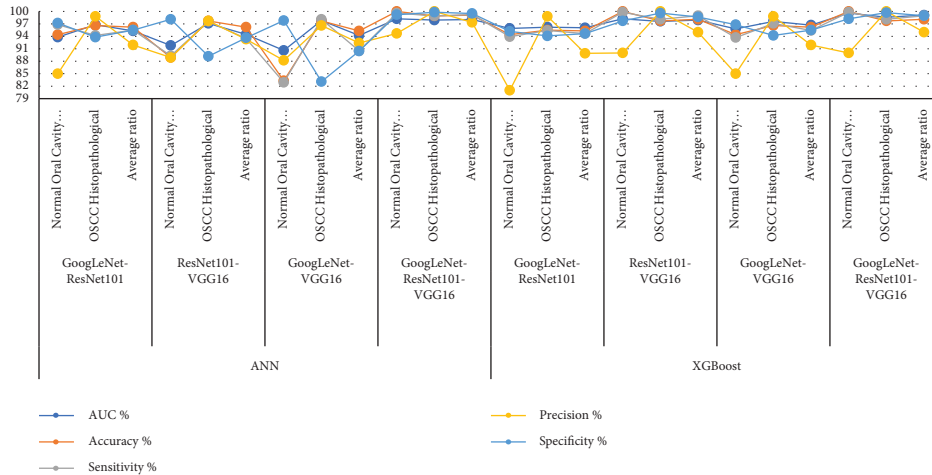


FIGURE 14: Display of results of hybrid systems with fused features to classify an oral cancer dataset for a 100x magnification factor.

TABLE 8: Results of hybrid systems with fused features to classify an oral cancer dataset for a 400x magnification factor.

Classifiers	Hybrid features	Classes	AUC (%)	Accuracy (%)	Sensitivity (%)	Precision (%)	Specificity (%)
ANN	GoogLeNet-ResNet101	Normal oral cavity histopathological	97.3	97.5	97.3	97.5	99.2
		OSCC histopathological	98.6	99	98.9	99	96.8
		Average ratio	97.95	98.6	98.1	98.25	98
	ResNet101-VGG16	Normal oral cavity histopathological	97.3	97.5	97.1	95.1	97.9
		OSCC histopathological	97.8	98	98.2	99	96.7
		Average ratio	97.55	97.8	97.65	97.05	97.3
	GoogLeNet-VGG16	Normal oral cavity histopathological	96.8	97.5	96.9	92.9	97.2
		OSCC histopathological	98.1	97	97.1	99	96.8
		Average ratio	97.45	97.1	97	95.95	97
	GoogLeNet-ResNet101-VGG16	Normal oral cavity histopathological	98.2	97.5	96.9	100	99.6
		OSCC histopathological	99.5	100	99.5	99	97.1
		Average ratio	98.85	99.3	98.2	99.5	98.35
XGBoost	GoogLeNet-ResNet101	Normal oral cavity histopathological	92.9	90	90.1	94.7	98.1
		OSCC histopathological	97.6	98	97.9	96	90.3
		Average ratio	95.25	95.7	94	95.35	94.2
	ResNet101-VGG16	Normal oral cavity histopathological	96.1	92.5	93.1	94.9	97.8
		OSCC histopathological	98.8	98	98.1	97	93.2
		Average ratio	97.45	96.4	95.6	95.95	95.5
	GoogLeNet-VGG16	Normal oral cavity histopathological	95.3	92.5	92.8	97.4	98.7
		OSCC histopathological	99.1	99	99.1	97	93.2
		Average ratio	97.2	97.1	95.95	97.2	95.95
	GoogLeNet-ResNet101-VGG16	Normal oral cavity histopathological	97.2	95	95.1	97.4	98.8
		OSCC histopathological	99.3	99	99.2	98	94.9
		Average ratio	98.25	97.8	97.15	97.7	96.85

pretrained models did not achieve satisfactory results. Still, the results of the pretrained models improved when applied to the data set after segmenting the regions of interest. Their performance on the same two datasets has improved

significantly when using hybrid models between CNN models and ANN and XGBoost networks. Due to the similarity of features of normal and malignant tissues in the early stages, a hybrid technique was applied based on the

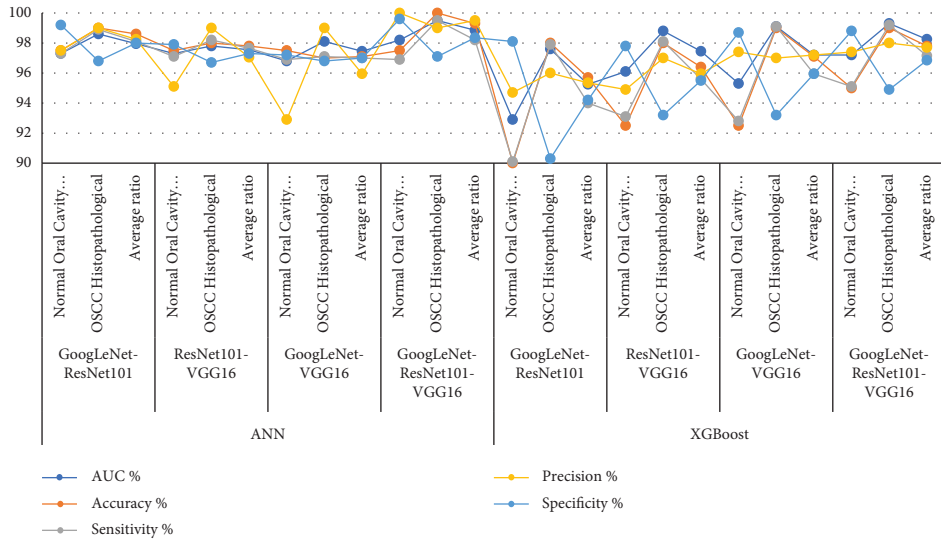


FIGURE 15: Display of results of hybrid systems with fused features to classify an oral cancer dataset for a 400x magnification factor.

Confusion Matrix: GoogLeNet-ResNet101-ANN

Output Class	Normal Oral Cavity	17 16.0%	3 2.8%	85.0% 15.0%
	OSCC	1 0.9%	85 80.2%	98.8% 1.2%
		94.4% 5.6%	96.6% 3.4%	96.2% 3.8%
	Normal Oral Cavity	OSCC		
			Target Class	

(a)

Confusion Matrix: ResNet101-VGG16-ANN

Output Class	Normal Oral Cavity	16 15.1%	2 1.9%	88.9% 11.1%
	OSCC	2 1.9%	86 81.1%	97.7% 2.3%
		88.9% 11.1%	97.7% 2.3%	96.2% 3.8%
	Normal Oral Cavity	OSCC		
			Target Class	

(b)

FIGURE 16: Continued.

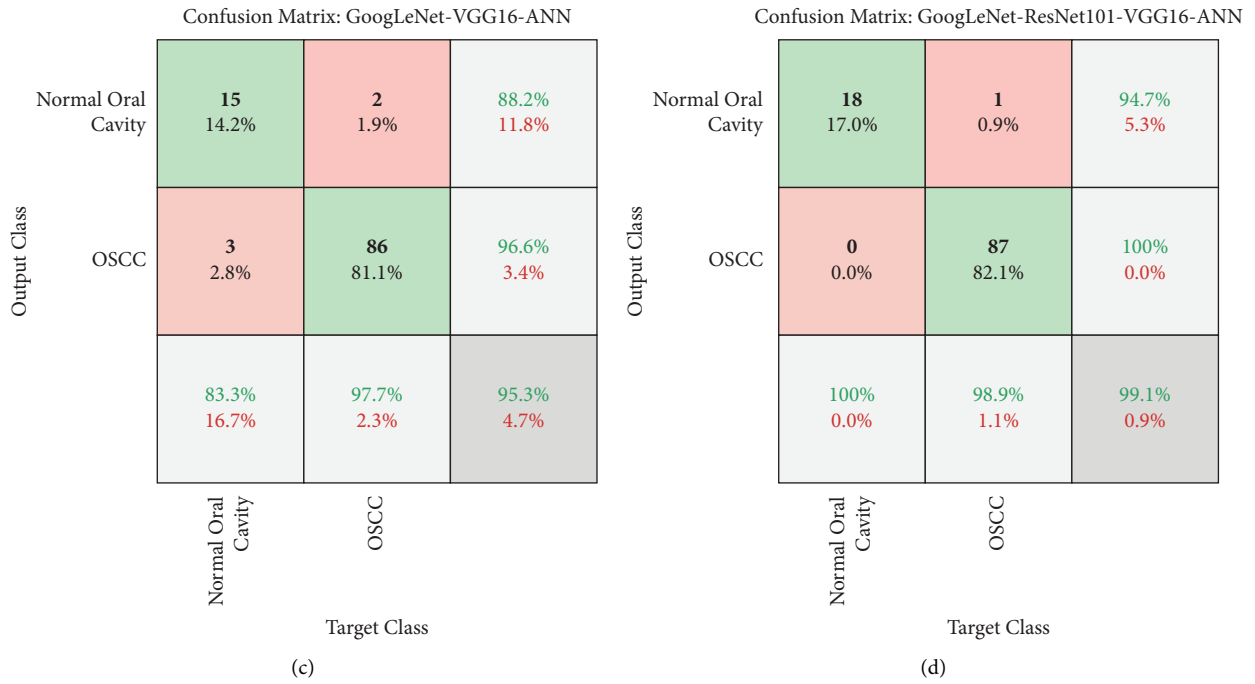


FIGURE 16: Confusion matrix for displaying the results of the classification of hybrid systems with fused CNN features for an oral cancer dataset for a magnification factor of 100x by techniques: (a) GoogLeNet-ResNet101-ANN, (b) ResNet101-VGG16-ANN, (c) GoogLeNet-VGG16-ANN, and (d) GoogLeNet-ResNet101-VGG16-ANN.

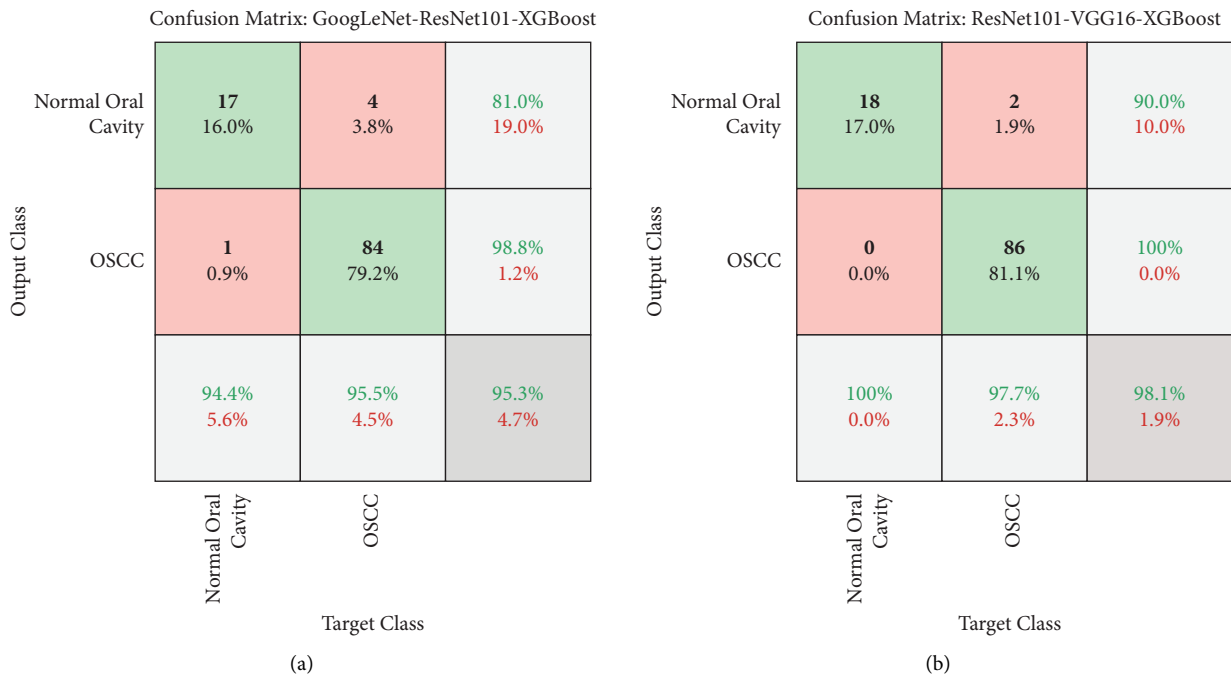


FIGURE 17: Continued.

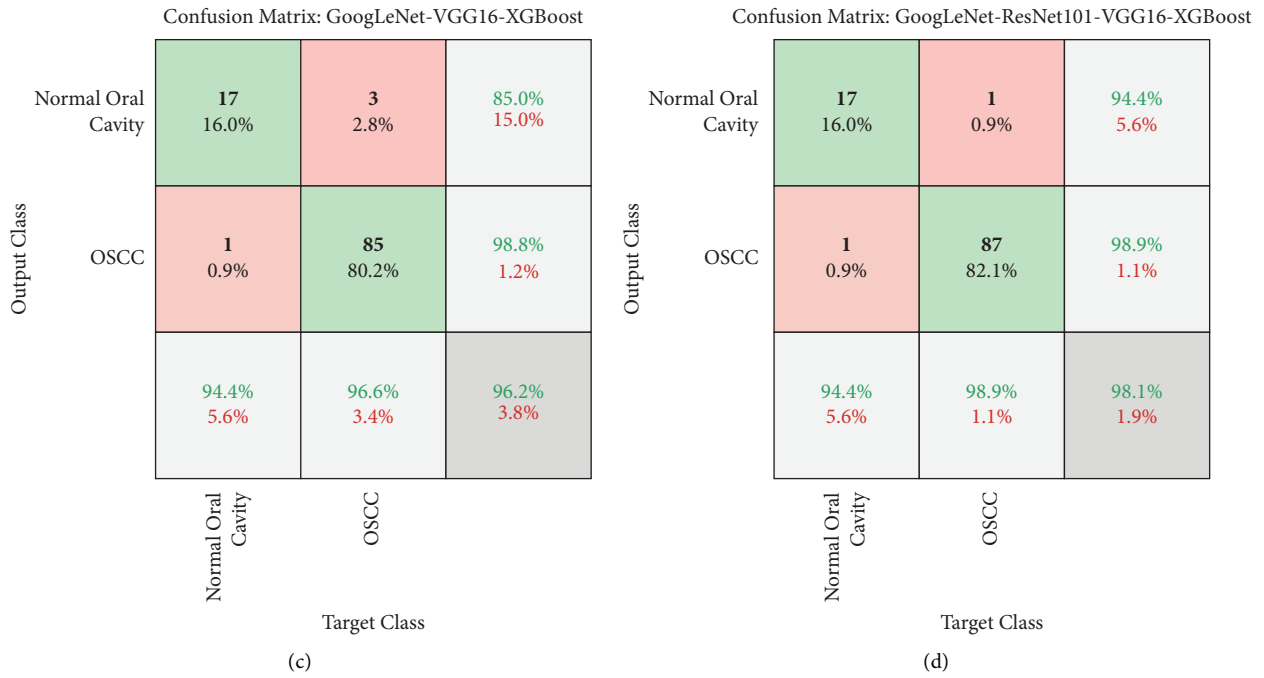


FIGURE 17: Confusion matrix for displaying the results of the classification of hybrid systems with fused CNN features for an oral cancer dataset for a magnification factor of 100x by techniques: (a) GoogLeNet-ResNet101-XGBoost, (b) ResNet101-VGG16-XGBoost, (c) GoogLeNet-VGG16-XGBoost, and (d) GoogLeNet-ResNet101-VGG16- XGBoost.

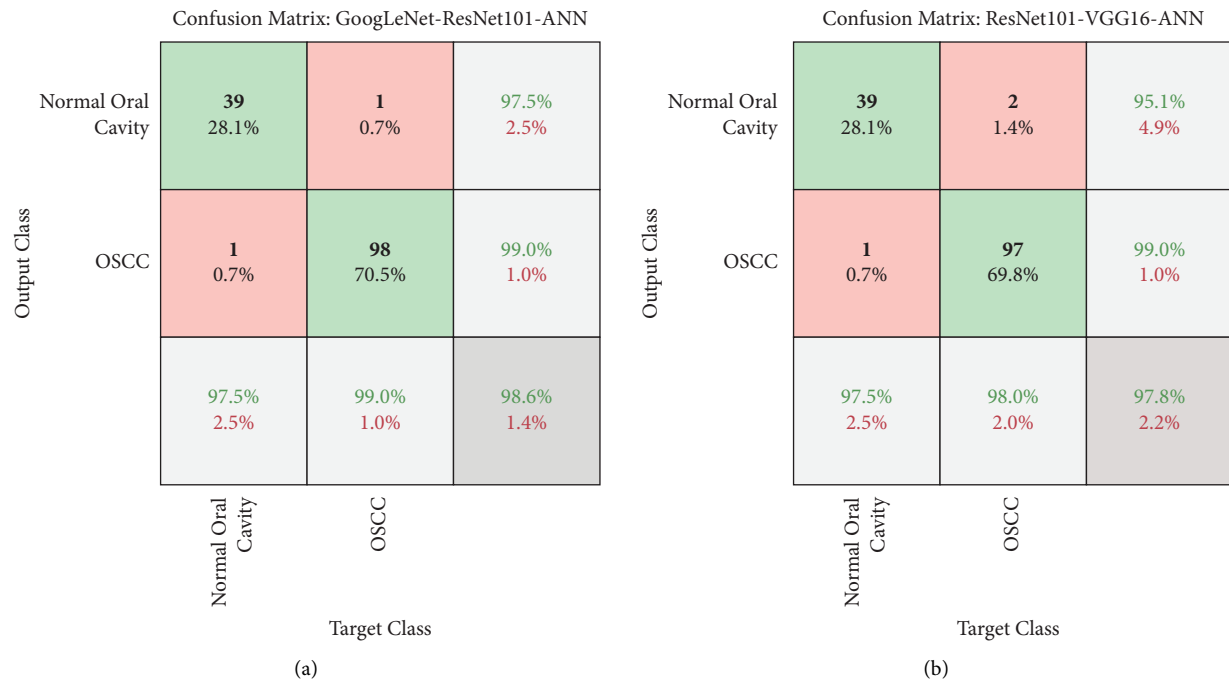


FIGURE 18: Continued.

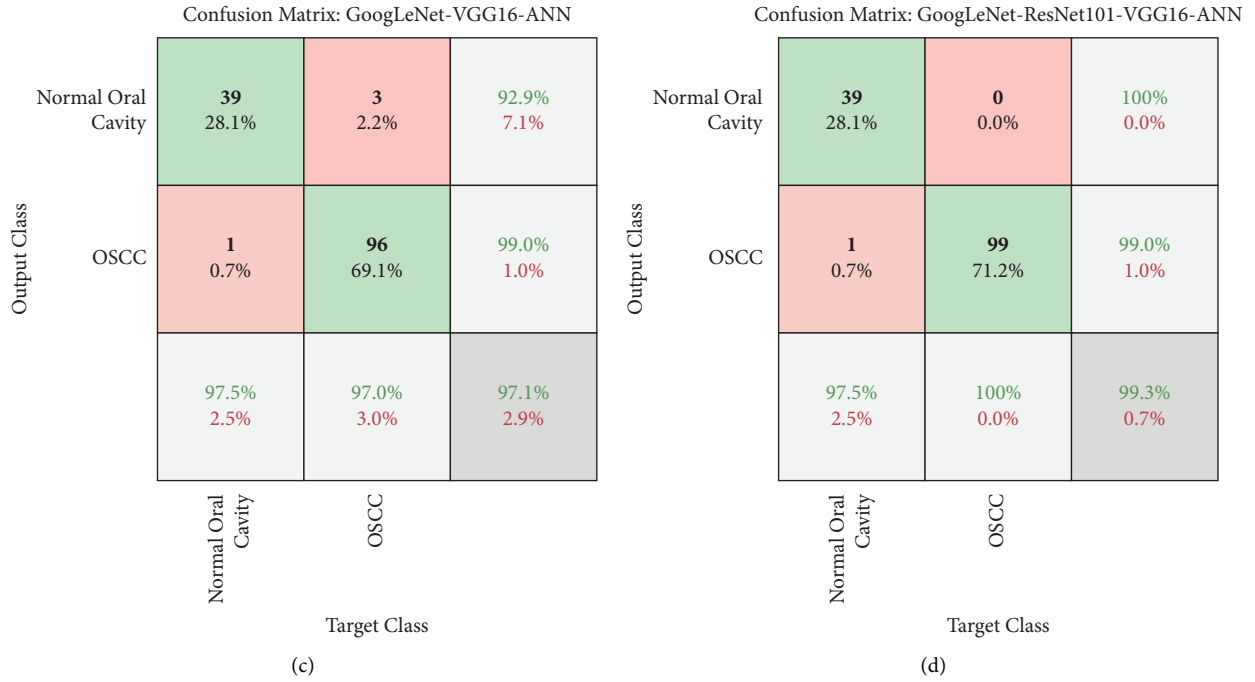


FIGURE 18: Confusion matrix for displaying the results of the classification of hybrid systems with fused CNN features for an oral cancer dataset for a magnification factor of 400x by techniques: (a) GoogLeNet-ResNet101-ANN, (b) ResNet101-VGG16-ANN, (c) GoogLeNet-VGG16-ANN, and (d) GoogLeNet-ResNet101-VGG16-ANN.

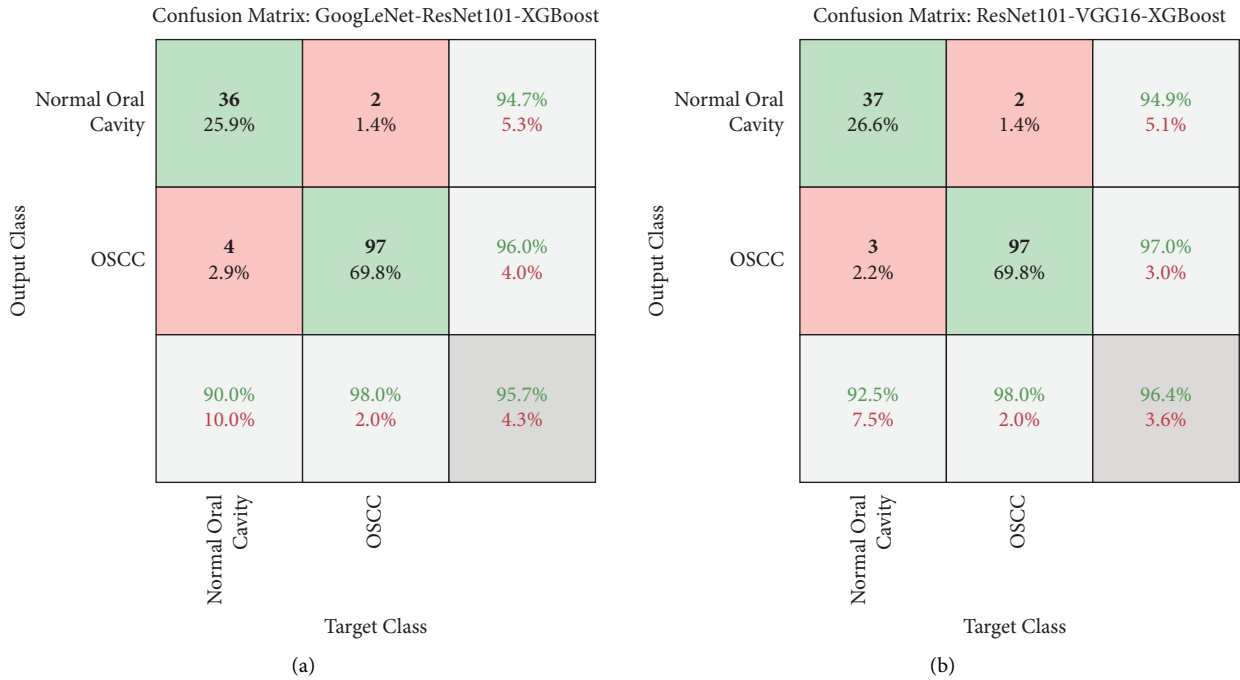


FIGURE 19: Continued.

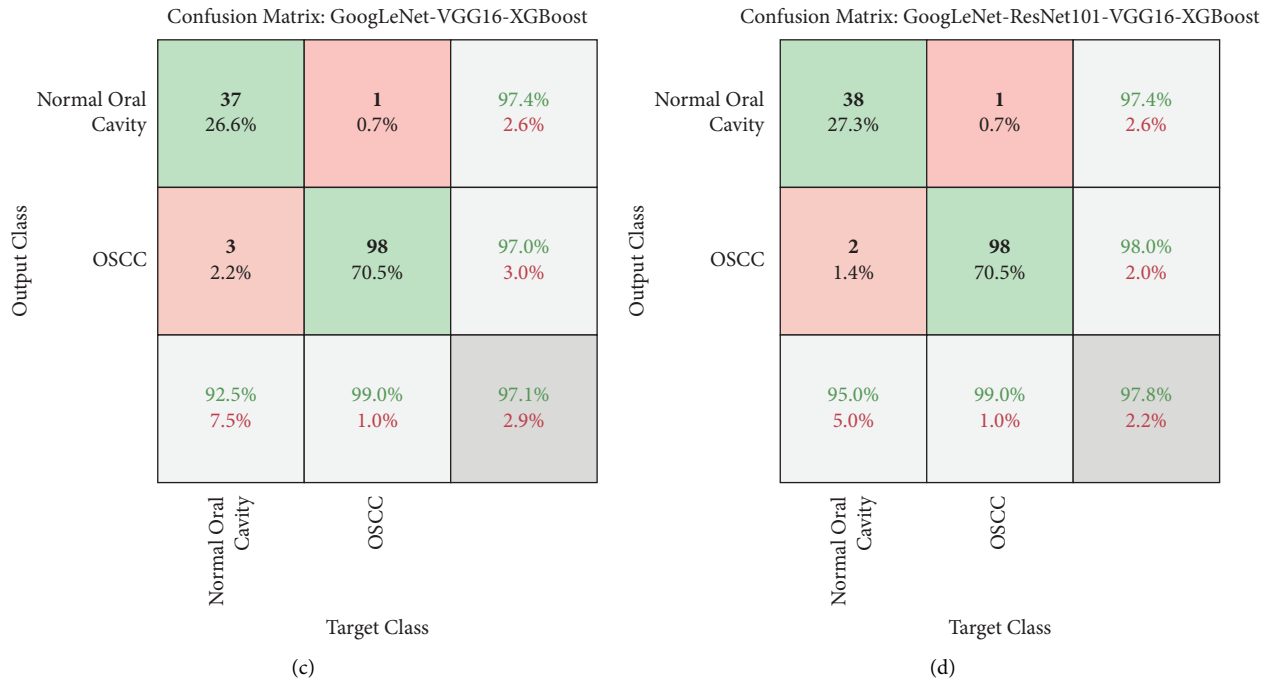


FIGURE 19: Confusion matrix for displaying the results of the classification of hybrid systems with fused CNN features for an oral cancer dataset for a magnification factor of 400x by techniques: (a) GoogLeNet-ResNet101-XGBoost, (b) ResNet101-VGG16- XGBoost, (c) GoogLeNet-VGG16- XGBoost, and (d) GoogLeNet-ResNet101-VGG16- XGBoost.

combined features of more than one CNN model based on the ARG algorithm.

First, with the OSCC-100x data set, the ANN with GoogLeNet-ResNet101-VGG16 features and the XGBoost network with ResNet101-VGG16 features reached 100% accuracy for classifying normal oral cavity tissue class. ANN and XGBoost with GoogLeNet-ResNet101-VGG16 achieved an accuracy of 98.9% for classifying OSCC class. Second, with the OSCC-400x data set, the XGBoost network with GoogLeNet-ResNet101 features reached an accuracy of 99% for classifying normal oral cavity tissue. The ANN with GoogLeNet-ResNet101-VGG16 features achieved an accuracy equal to 100% for classifying OSCC class.

It is important to note that not all studies provided specific accuracy, sensitivity, or specificity values. However, from the available results, the studies by Rasheed et al. [13] highlighted the better performance of the HRNet model compared to ResNet50 and DenseNet169 models, which reached an accuracy of 84.3% and sensitivity of 83%. Rahman et al. [14] employed a transfer learning methodology to make predictions regarding histopathologic oral cancer. Their approach demonstrated a level of accuracy reaching 90.06%. Fati et al. [15] and Warin et al. [16] achieved high accuracy in OSCC detection. The Inception-ResNet-V2 model, developed by Camalan et al. [17] was utilized to generate maps that specifically target the impacted region for classification purposes. This approach yielded an accuracy rate of 73.6% and an F1 score of 97.9%. Musulin et al. [18] an

Integrated Xception and SWT approach for multiclass classification and segmentation. The approach got an accuracy of 96.3% and Semantic segmentation of 87.8%. Das et al. [20] demonstrated the effectiveness of their 10-layer CNN model, achieving a high accuracy rate. Yang et al. [22] showed improvement in F1 score with the assistance of their deep learning model. Amin et al. [23] achieved good results with their sequential models.

In the OSCC-400x dataset, the quantity of normal class images was artificially amplified sixfold for each image, whereas the OSCC class images were artificially amplified twofold. Das et al. [29] achieved an accuracy of 89% with relatively high sensitivity, precision, and specificity. Myriam et al. [30] achieved a high accuracy of 97.3% with good sensitivity and precision. Our proposed system achieved exceptional results with high accuracy of 99.3, AUC of 98.85%, sensitivity of 98.2%, precision of 99.5%, and specificity of 98.35%. The studies utilizing capsule networks (Panigrahi et al. [27]) showed promising potential for the analysis of OSCC histopathological images, with high accuracy and improved sensitivity compared to other methods. Das et al. [29] and Wu et al. [28] focused on specific tasks, achieving good results in identifying cancerous cells and region segmentation. Myriam et al. [30] introduced a novel meta-heuristic algorithm that enhanced the accuracy and efficiency of oral cancer detection. Our proposed system integrated multiple models and achieved

TABLE 9: Results of proposed models for histopathological diagnosis of the two datasets OSCC-100x and OSCC-400x.

Techniques	Datasets			OSCC-100x			OSCC-400x			
	Features	Normal	OSCC	Accuracy (%)	Normal	OSCC	Accuracy (%)	Normal	OSCC	Accuracy (%)
Pretrained	GoogLeNet	55.6	87.5	82.1	55	86.9	77.7	55	86.9	77.7
	ResNet101	50	87.5	81.1	52.5	86.9	77	52.5	86.9	77
	VGG16	38.9	86.4	78.3	55	89.9	79.9	55	89.9	79.9
Based on ARG algorithm	GoogLeNet	55.6	97.7	90.6	82.5	94.9	91.4	82.5	94.9	91.4
	ResNet101	55.6	96.6	89.6	85	96	92.8	85	96	92.8
	VGG16	83.3	89.8	88.7	82.5	96	92.1	82.5	96	92.1
ANN	GoogLeNet	77.8	95.5	92.5	85	94.9	92.1	85	94.9	92.1
	ResNet101	66.7	96.6	91.5	92.5	94.9	94.2	92.5	94.9	94.2
	VGG16	88.9	94.3	93.4	87.5	96	93.5	87.5	96	93.5
XGBoost	GoogLeNet	94.4	92	92.5	87.5	99	95.7	87.5	99	95.7
	ResNet101	88.9	95.5	94.3	87.5	94.9	92.8	87.5	94.9	92.8
	VGG16	77.8	94.3	91.5	92.5	97	95.7	92.5	97	95.7
Hybrid	GoogLeNet-ResNet101	94.4	96.6	96.2	97.5	99	98.6	97.5	99	98.6
	ResNet101-VGG16	88.9	97.7	96.2	97.5	98	97.8	97.5	98	97.8
	GoogLeNet-VGG16	83.3	97.7	95.3	97.5	97	97.1	97.5	97	97.1
	GoogLeNet-ResNet101-VGG16	100	98.9	99.1	97.5	100	99.3	97.5	100	99.3
Hybrid	GoogLeNet-ResNet101	94.4	95.5	95.3	99	98	95.7	99	98	95.7
	ResNet101-VGG16	100	97.7	98.1	92.5	98	96.4	92.5	98	96.4
	GoogLeNet-VGG16	94.4	96.6	96.2	92.5	99	97.1	92.5	99	97.1
	GoogLeNet-ResNet101-VGG16	94.4	98.9	98.1	95	99	97.8	95	99	97.8

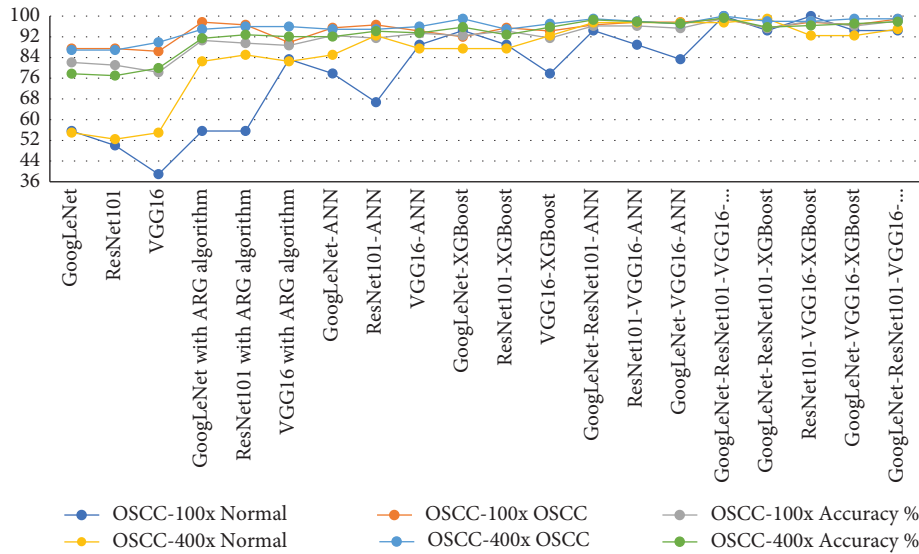


FIGURE 20: Display of results of proposed models for histopathological diagnosis of the two datasets OSCC-100x and OSCC-400x.

superior results in distinguishing normal and malignant oral cavity tissues.

Each study presented different approaches and achieved varying levels of performance. Further research and comparative evaluations are needed to assess these methods' overall effectiveness and suitability in real-world scenarios.

6. Conclusions

Distinguishing between normal and malignant oral cavity histology is difficult. In general, the proposed models in this work showed their superior ability to analyze biopsy images of pathological oral cavity tissues to discover the clinical characteristics of oral cancer and distinguish them from normal clinical characteristics. The first strategy uses pre-trained GoogLeNet, ResNet101, and VGG16 models. The OSCC-100x image and OSCC-400x datasets have been optimized to increase contrast in low-light areas, remove artifacts, and standardize color. The AGR algorithm was applied to separate the tissues of interest for further analysis through CNN models. The second strategy for diagnosing the oral cancer dataset by GoogLeNet, ResNet101, and VGG16 models is based on the AGR segmentation algorithm. The third strategy for diagnosing the oral cancer dataset is based on a hybrid technique between CNN models and ANN and XGBoost networks based on the AGR segmentation algorithm. The fourth strategy depends on integrating the features of CNN models based on the AGR segmentation algorithm and its classification by ANN and XGBoost networks. The models achieved superior results in distinguishing the pathological tissues of the normal from the malignant oral cavity. ANN with features of GoogLeNet-ResNet101-VGG16 reached an AUC of 98.85%, accuracy of 99.3%, sensitivity of 98.2%, precision of 99.5%, and specificity of 98.35%. The future work of this study is to develop the proposed systems in which the handcrafted features are combined with the features of CNN models extracted from several models and classified by several machine learning

algorithms. The proposed hybrid systems can diagnose a new data set because this study was applied to two data sets with magnification factors of 100x and 400x. Thus, the systems can be generalized to diagnose a new data set.

Data Availability

The data supporting the proposed models were collected from the OSCC Oral Cancer Dataset which is available to the public at <https://data.mendeley.com/datasets/ftmp4cvtmb/1>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors are thankful to the Deanship of Scientific Research at Najran University for funding this work, under the General Research Funding program grant code (NU/DRP/SERC/12/8).

References

- [1] P. Ahmad, R. Nawaz, M. Qurban et al., "Risk factors associated with the mortality rate of oral squamous cell carcinoma patients: a 10-year retrospective study," *Medicine*, vol. 100, no. 36, 2021.
- [2] K. Bansal, R. K. Bathla, and Y. Kumar, "Deep transfer learning techniques with hybrid optimization in early prediction and diagnosis of different types of oral cancer," *Soft Computing*, vol. 26, no. 21, pp. 11153–11184, 2022.
- [3] M. Hertel, L. Hagedorn, A. M. Schmidt-Westhausen et al., "Comparison of five-year survival rates among patients with oral squamous cell carcinoma with and without association with syphilis: a retrospective case-control study," *BMC Cancer*, vol. 22, no. 1, pp. 1–7, 2021.
- [4] A. Thomas, S. Manchella, K. Koo, A. Tiong, A. Nastri, and D. Wiesenfeld, "The impact of delayed diagnosis on the outcomes of oral cancer patients: a retrospective cohort

- study,” *International Journal of Oral and Maxillofacial Surgery*, vol. 50, no. 5, pp. 585–590, 2021.
- [5] Z. Ling, B. Cheng, and X. Tao, “Epithelial-to-mesenchymal transition in oral squamous cell carcinoma: challenges and opportunities,” *International Journal of Cancer*, vol. 148, no. 7, pp. 1548–1561, 2021.
 - [6] A. L. O. C. Roza, L. P. Kowalski, W. N. William, G. de Castro, A. L. F. Chaves, and A. L. D. Araújo, “Oral leukoplakia and erythroplakia in young patients: a systematic review,” *Oral surgery, oral medicine, oral pathology and oral radiology*, vol. 131, no. 1, pp. 73–84, 2021.
 - [7] M. Korytowska, G. Schwab, D. Giglio et al., “Patient-reported pain after surgical removal of leukoplakia—an observational 1-year follow-up study,” *Acta Odontologica Scandinavica*, vol. 79, no. 5, pp. 383–389, 2021.
 - [8] S. T. Nethan, S. Gupta, and S. Warnakulasuriya, “Risk factors for oral squamous cell carcinoma in the Indian population,” in *Microbes and Oral Squamous Cell Carcinoma: A Network Spanning Infection and Inflammation*, pp. 9–40, Springer Nature Singapore, Singapore, 2022.
 - [9] Y. F. Ossa and C. F. Putri, “Delayed diagnosis of oral malignancy and effect to prognosis; case series,” in *2nd Aceh International Dental Meeting 2021*, pp. 1–8, Atlantis Press, Amsterdam, Netherlands, 2022.
 - [10] L. F. Sun, C. X. Wang, Z. Y. Cao et al., “Evaluation of autofluorescence visualization system in the delineation of oral squamous cell carcinoma surgical margins. Photodiagnosis and Photodynamic,” *Therapy*, vol. 36, Article ID 102487, 2021.
 - [11] E. Nakamichi, H. Sakakura, S. Mii et al., “Detection of serum/salivary exosomal Alix in patients with oral squamous cell carcinoma,” *Oral Diseases*, vol. 27, no. 3, pp. 439–447, 2021.
 - [12] R. O. Alabi, A. Almangush, M. Elmusrati, and A. A. Mäkitie, “Deep machine learning for oral cancer: from precise diagnosis to precision medicine,” *Frontiers in Oral Health*, vol. 2, p. 97, 2022.
 - [13] H. Lin, H. Chen, L. Weng, J. Shao, and J. Lin, “Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis,” *Journal of Biomedical Optics*, vol. 26, no. 8, Article ID 086007, 2021.
 - [14] A. U. Rahman, A. Alqahtani, N. Aldhafferi et al., “Histopathologic oral cancer prediction using oral squamous cell carcinoma biopsy empowered with transfer learning,” *Sensors*, vol. 22, no. 10, p. 3833, 2022.
 - [15] 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, p. 1899, 2022.
 - [16] K. Warin, W. Limprasert, S. Suebnukarn, S. Jinaporntham, and P. Jantana, “Automatic classification and detection of oral cancer in photographic images using deep learning algorithms,” *Journal of Oral Pathology and Medicine*, vol. 50, no. 9, pp. 911–918, 2021.
 - [17] S. Camalan, H. Mahmood, H. Binol et al., “Convolutional neural network-based clinical predictors of oral dysplasia: class activation map analysis of deep learning results,” *Cancers*, vol. 13, p. 1291, 2021.
 - [18] J. Musulin, D. Štifić, A. Zulijani, T. Čabov, A. Dekanić, and Z. Car, “An enhanced histopathology analysis: an AI-based system for multiclass grading of oral squamous cell carcinoma and segmenting of epithelial and stromal tissue,” *Cancers*, vol. 13, p. 1784, 2021.
 - [19] Y.-C. Lin and B.-S. Chen, “Identifying drug targets of oral squamous cell carcinoma through a systems biology method and genome-wide microarray data for drug discovery by deep learning and drug design specifications,” *International Journal of Molecular Sciences*, vol. 23, Article ID 10409, 2022.
 - [20] 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,” *International Journal of Environmental Research and Public Health*, vol. 20, p. 2131, 2023.
 - [21] F. Jing, J. Zhang, X. Cai et al., “Screening for biomarkers for progression from oral leukoplakia to oral squamous cell carcinoma and evaluation of diagnostic efficacy by multiple machine learning algorithms,” *Cancers*, vol. 14, p. 5808, 2022.
 - [22] S. Y. Yang, S. H. Li, J. L. Liu et al., “Histopathology-based diagnosis of oral squamous cell carcinoma using deep learning,” *Journal of Dental Research*, vol. 101, no. 11, pp. 1321–1327, 2022.
 - [23] I. Amin, H. Zamir, and F. F. Khan, “Histopathological image analysis for oral squamous cell carcinoma classification using concatenated deep learning models,” *medRxiv*, pp. 2021–2105, 2021.
 - [24] M. Pal, P. K. Panigrahi, and A. Pradhan, “An ensemble deep learning model with empirical wavelet transform feature for oral cancer histopathological image classification,” 2022, <https://www.medrxiv.org/content/10.1101/2022.11.13.22282266v1>.
 - [25] A. Ghosh, D. Chaudhuri, S. Adhikary et al., “Deep reinforced neural network model for cyto-spectroscopic analysis of epigenetic markers for automated oral cancer risk prediction,” *Chemometrics and Intelligent Laboratory Systems*, vol. 224, Article ID 104548, 2022.
 - [26] M. A. Deif, H. Attar, A. Amer, I. A. Elhady, M. R. Khosravi, and A. A. Solymán, “Diagnosis of oral squamous cell carcinoma using deep neural networks and binary Particle Swarm optimization on histopathological images: an AIOMT approach,” *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6364102, 13 pages, 2022.
 - [27] S. Panigrahi, J. Das, and T. Swarnkar, “Capsule network based analysis of histopathological images of oral squamous cell carcinoma,” *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 7, pp. 4546–4553, 2022.
 - [28] Y. Wu, C. F. Koyuncu, P. Toro et al., “A machine learning model for separating epithelial and stromal regions in oral cavity squamous cell carcinomas using H&E-stained histology images: a multi-center, retrospective study,” *Oral Oncology*, vol. 131, Article ID 105942, 2022.
 - [29] 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,” *International Journal of Environmental Research and Public Health*, vol. 20, no. 3, p. 2131, 2023.
 - [30] H. Myriam, A. A. Abdelhamid, E. S. M. El-Kenawy et al., “Advanced meta-heuristic algorithm based on Particle Swarm and Al-biruni Earth Radius optimization methods for oral cancer detection,” *IEEE Access*, vol. 11, pp. 23681–23700, 2023.
 - [31] T. Y. Rahman, L. B. Mahanta, A. K. Das, and J. D. Sarma, “Histopathological imaging database for oral cancer analysis,” *Data in Brief*, vol. 29, Article ID 105114, 2020.
 - [32] I. A. Ahmed, E. M. Senan, T. H. Rassem et al., “Eye tracking-based diagnosis and early detection of autism spectrum disorder using machine learning and deep learning techniques,” *Electronics*, vol. 11, no. 4, p. 530, 2022.
 - [33] A. Patel, S. Patel, P. Patel, D. Mandlik, K. Patel, and V. Tanavde, “Salivary exosomal miRNA-1307-5p predicts disease aggressiveness and poor prognosis in oral squamous

- cell carcinoma patients,” *International Journal of Molecular Sciences*, vol. 23, Article ID 10639, 2022.
- [34] J. Kim, L. Tran, T. Peto, and E. Y. Chew, “Identifying those at risk of glaucoma: a deep learning approach for optic disc and cup segmentation and their boundary analysis,” *Diagnostics*, vol. 12, p. 1063, 2022.
- [35] S. Kadry, V. Rajinikanth, D. Taniar, R. Damaševičius, and X. P. B. Valencia, “Automated segmentation of leukocyte from hematological images—a study using various CNN schemes,” *The Journal of Supercomputing*, vol. 78, pp. 1–21, 2022.
- [36] Y. J. Seol, S. H. Park, Y. J. Kim, Y.-T. Park, H. Y. Lee, and K. G. Kim, “The development of an automatic rib sequence labeling system on axial computed tomography images with 3-dimensional region growing,” *Sensors*, vol. 22, Article ID 4530, 2022.
- [37] S. M. Fati, E. M. Senan, and N. ElHakim, “Deep and hybrid learning technique for early detection of tuberculosis based on X-ray images using feature fusion,” *Applied Sciences*, vol. 12, p. 7092, 2022.
- [38] S. M. Fati, E. M. Senan, and A. T. Azar, “Hybrid and deep learning approach for early diagnosis of lower gastrointestinal diseases,” *Sensors*, vol. 22, Article ID 4079, 2022.
- [39] A. M. Al-Hejri, R. M. Al-Tam, M. Fazea, A. H. Sable, S. Lee, and M. A. Al-antari, “ETECADx: Ensemble self-attention transformer encoder for breast cancer diagnosis using full-field digital X-ray breast images,” *Diagnostics*, vol. 13, p. 89, 2023.
- [40] R. M. Al-Tam, A. M. Al-Hejri, S. M. Narangale et al., “A hybrid workflow of residual convolutional transformer encoder for breast cancer classification using digital X-ray mammograms,” *Biomedicines*, vol. 10, p. 2971, 2022.
- [41] Z. G. Al-Mekhlafi, E. M. Senan, B. A. Mohammed et al., “Diagnosis of histopathological images to distinguish types of malignant lymphomas using hybrid techniques based on fusion features,” *Electronics*, vol. 11, p. 2865, 2022.
- [42] E. M. Senan, M. E. Jadhav, T. H. Rassem, B. A. Mohammed, and Z. G. Al-Mekhlafi, “Early diagnosis of brain tumour MRI images using hybrid techniques between deep and machine learning,” *Computational and Mathematical Methods in Medicine*, vol. 2022, Article ID 8330833, 2022.
- [43] S. Silvestrini and M. Lavagna, “Deep learning and artificial neural networks for spacecraft dynamics, navigation and control,” *Drones*, vol. 6, p. 270, 2022.
- [44] H. Xu, R. Chang, M. Pan et al., “Application of artificial neural networks in construction management: a scientometric review,” *Buildings*, vol. 12, p. 952, 2022.
- [45] T.-T.-H. Le, Y. E. Oktian, and H. Kim, “XGBoost for imbalanced multiclass classification-based industrial internet of things intrusion detection systems,” *Sustainability*, vol. 14, p. 8707, 2022.
- [46] M. Amjad, I. Ahmad, M. Ahmad, P. Wróblewski, P. Kamiński, and U. Amjad, “Prediction of pile bearing capacity using XGBoost algorithm: modeling and performance evaluation,” *Applied Sciences*, vol. 12, p. 2126, 2022.
- [47] M. Jamshidi, S. Roshani, F. Daneshfar et al., “Hybrid deep learning techniques for predicting complex phenomena: a review on COVID-19,” *AIDS*, vol. 3, pp. 416–433, 2022.
- [48] A. K. Devarakonda, N. Karuppiyah, T. Selvaraj, P. K. Balachandran, R. Shanmugasundaram, and T. Senjyu, “A comparative analysis of maximum power point techniques for solar photovoltaic systems,” *Energies*, vol. 15, p. 8776, 2022.
- [49] D. Quan, Z. Tang, X. Wang, W. Zhai, and C. Qu, “LPI radar signal recognition based on dual-channel CNN and feature fusion,” *Symmetry Plus*, vol. 14, p. 570, 2022.
- [50] F. Cui, R. Wang, W. Ding, Y. Chen, and L. Huang, “A novel DE-CNN-BiLSTM multi-fusion model for EEG emotion recognition,” *Mathematics*, vol. 10, p. 582, 2022.
- [51] E. M. Senan, I. Abunadi, M. E. Jadhav, and S. M. Fati, “Score and correlation coefficient-based feature selection for predicting heart failure diagnosis by using machine learning algorithms,” *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 8500314, 16 pages, 2021.
- [52] F. Olayah, E. M. Senan, I. A. Ahmed, and B. Awaji, “AI techniques of dermoscopy image analysis for the early detection of skin lesions based on combined CNN features,” *Diagnostics*, vol. 13, no. 7, p. 1314, 2023.
- [53] D. O. Oyewola, E. G. Dada, S. Misra, and R. Damaševičius, “A novel data augmentation convolutional neural network for detecting malaria parasite in blood smear images,” *Applied Artificial Intelligence*, vol. 36, no. 1, Article ID 2033473, 2022.