Automatic Detection and Classification of Oral Cancer from Photographic Images Using Genetic Algorithm Optimized Deep Learning CNN Model

Sayyada Hajera Begum * and P. Vidyullatha

Computer Science Engineering Department, Koneru Lakshmaiah Education Foundation, Vijayawada, India
Email: sayyada.hajera07@gmail.com (S.H.B.); pvidyullatha@kluniversity.in (P.V.)
*Corresponding author

***Abstract*—Oral cancer deals with the cancerous lesions that appear in the mouth, lips, tongue and cheeks and poses a significant health challenge, and its detection in early stages is pivotal for better patient outcomes. Latest Deep Learning Techniques have provided a lot of opportunities for automatic detection and classification of oral cancers with the accuracy better than that of human experts by providing a non-invasive and cost-effective method to detect oral cancer at early-stages, thus enabling early treatment. Convolutional Neural Networks (CNNs) have shown immense potential in the field of medical image analysis, including the automatic detection of oral cancer from histopathological and photographic images. However, optimizing CNNs for this specific task remains intricate and time-consuming due to the numerous hyper parameters involved. The paper proposes a novel approach to optimize CNN hyper parameters that affects the performance and speed of CNN architecture for oral cancer detection using photographic images. The method involves applying Genetic Algorithm (GA) to efficiently tune the hyper parameter space and enhance the CNN's performance. The experimental findings reveal that the GA-optimized CNN outperforms the base CNN model, achieving higher accuracy, sensitivity, and specificity. The proposed approach was able to achieve an accuracy of 95% after GA optimization on DenseNet201 model compared to the performance before optimization. Thus, the trained CNN model holds promising prospects for ensuing early oral cancer detection and subsequently improving patient prognosis and survival rates.**

***Keywords*—deep learning techniques, oral cancer, Convolutional Neural Network (CNN), Genetic Algorithm (GA), medical image analysis**

I. Introduction

CNNs have transformed the healthcare industry by combining data-driven insights with clinical expertise, making it more accurate, efficient, and accessible. Medical images like X-rays, CT scans, MRIs, can be accurately diagnosed for minute patterns, aiding in the early detection of multiple diseases. Segmentation of regions of interest in the medical images can support personalized treatment planning. By analyzing patient-specific data, CNNs can act as diagnostic aids to the clinical experts thus reducing manual workload.

Among various cancers prevalent Oral cancer is a life-threatening disease that affects millions of people globally. Its early detection is crucial to improve the chances of its successful treatment. Oral cancer is a part of head-neck cancers encompassing malignant and benign tumors that manifest in the mouth covering areas like tongue, inner side of the cheeks, top of the mouth and below the tongue.

Among various oral cancers found in head and neck, the most widespread is OSCC which covers over 90% of cancer cases [1]. Another type of oral cancer, Verrucous carcinoma is a rare, slow-growing cancer that looks like a grey lump or wart. Another variant of oral cancer, Minor salivary gland carcinoma originates in the salivary glands and is scattered throughout the mouth and throat, with the hard and soft palates being a common location. Lymphoma oral cancer arises from the immune system, and manifests in the form of oral lesions. And lastly the Melanoma cancer originates in the cells that produces pigments in the skin or in mucous membranes and affects the mouth and lips. Lastly oral cancer can also develop as secondary cancer by originating in different parts of the body.

Oral cancer now ranks as one of the most common cancers globally, with India contributing significantly to the total number of cases. While tobacco use is the primary causative factor for oral cancer, the consumption of alcohol significantly elevates its risk. Roughly 90% of oral cancer patients reported to be tobacco users in various forms [2].

HPV infection is now emerging as a noteworthy source of oral cancer, specifically among youngsters. Surviving of oral cancer patients varies depending on the cancer's stage at diagnosis. For localized cancers, it stands at approximately 84%, but drops significantly for cancers which have already spread to other parts of the body.

Oral cancer, most commonly referred to as mouth cancer, can impact different parts of the oral cavity and

Manuscript received December 5, 2024; revised December 31, 2024; accepted February 26, 2025; published May 15, 2025.

its surroundings. Fig. 1 shows oral cancers on different parts of the mouth. Following are the various common oral cancer types found in mouth:

1. Oral Tongue Cancer: This variant specifically targets the front two-thirds of the tongue, usually a subtype of SCC associated with tobacco and alcohol use.
2. Floor of the Mouth Cancer: This cancer emerges on the mouth's floor, beneath the tongue, and is another form of SCC linked to tobacco and alcohol use.
3. Palate Cancer: Palate cancer can develop on the hard palate (roof of the mouth) or the soft palate (the rear fleshy part). It typically manifests as SCC.
4. Gum Cancer (Gingival Cancer): Gingival cancer affects the gums and can arise from squamous cells of the gums, often linked to factors like smoking and poor oral hygiene.
5. Lip Cancer: Lip cancer primarily targets the lower lip and is frequently associated with excessive sun exposure, more common in fair-skinned individuals.
6. Oropharyngeal Cancer: Although not strictly oral cancer, this type affects the throat and tonsils and is often caused by human papillomavirus (HPV) infection.
7. Nasopharyngeal Cancer: This cancer affects the area behind the nasal passages, extending into the upper throat, linked to viral infections like Epstein-Barr virus (EBV).
8. Sinonasal Cancer: Sinonasal cancer targets the nasal cavity and sinuses, relatively rare but exhibiting symptoms akin to other nasal and sinus conditions.

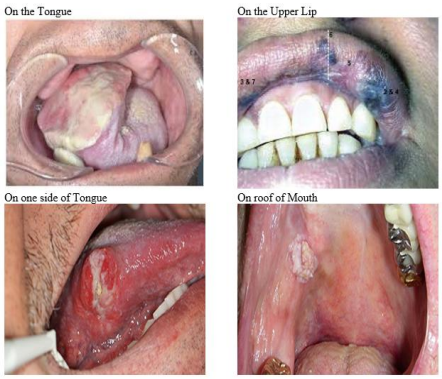


Fig. 1. Sample images of oral lesions in the mouth.

It's crucial to emphasize that these statistics provide a broad overview, and individual cases can exhibit significant variability depending on factors such as cancer subtype, stage of diagnosis, health condition, and the effect of treatment being provided. One notable challenge in addressing oral cancer is the difficulty in early diagnosis, particularly for asymptomatic early-stage cases. The utilization of screening tests, including visual examinations and biopsy, is crucial for early detection, leading to improved survival rates. Current diagnostic methods, primarily reliant on clinical examinations, exhibit limitations such as low sensitivity and a propensity to miss malignant lesions in concealed locations. In contrast to many other internal organs, the oral cavity permits visual examination without requiring dedicated apparatuses. In clinical tests, experts often rely on their individual experience and knowledge with the color and look of the cancer lesions to possibly diagnose oral cancer during visual inspections [3].

To address limitations of manual analysis, CAD systems are indispensable in the identification of oral cancers, enhancing diagnosis and survival rates. Classification is the pivotal aspect of CAD systems, and various ML and DL techniques are employed for this. The accuracy of these techniques' is based on feature quality. For extracting features, input can be provided in the form of Histopathological images, CT images, MRI images or Photographic images. Photographic images are chosen over other modalities as photographic images are rich in color diversity and complexity, can be captured non-invasively, do not require expensive equipment, can easily capture discolorations and irregular textures in the oral lesions and can benefit from learned techniques like DL models. Photographic imaging coupled with DL models can provide a practical solution for oral cancer diagnosis. Despite CAD system advancements, early oral cancer detection remains critical, prompting on-going research to enhance accuracy.

Increasingly, there is compelling evidence indicating that CNNs have demonstrated comparable or even better performance than human experts in discerning subtle graphical patterns from medical images.

II. Related Work

In the context of oral cancer, Deep Learning Techniques (DLTs) offer several advantages, including the assurance of early detection, decreased detection errors, a lighter burden on lab physicians, and improved diagnostic precision [4]. Various deep learning applications in the field of oral cancer include the identifying of infiltration of lymphocytes tumor [5], grading oral cancers [6], estimating cancer free survival rates [7], and categorizing cancer cases into various risk levels [8]. Aubreville *et al.* [9] were able to attain a classification accuracy of 87.02% by applying the Inception v3 CNN model, the researchers aimed to assess various methods for automatic diagnosing OSCC, from images and CNN models. Their primary focus encompassed aspects such as classification, learning, data analysis, and image retrieval. Furthermore, the researchers utilized genetic data and ML models to identify the growth of OC in patients.

Welikala *et al.* [10], considered two DLT namely ResNet-101 and Faster-RCNN for classification and detection, aiming to enhance the performance results. Their proposed model demonstrated significant improvement, achieving an impressive F1-Score of 87%. Shamim *et al.* [11] applied a CNN model based on Resnet-50 to effectively classify cancerous tongue lesions. The proposed CNN model achieved outstanding

results, accurately categorizing tongue lesions into multiple classes, including location, texture, color etc. with an impressive accuracy rate of 97%. Nanditha *et al*. [12] fused Resnet50 and VGG16 models for the analysis of photographic images depicting oral lesions, yielding an impressive accuracy rate of 96.2% from 332 images, encompassing 63 benign and 269 malignant lesions. Begum *et al*. [13] introduced a model that utilized DenseNet201 as its foundational architecture for the classification of cancerous and non-cancerous lesions from histopathological images, and reported a notable accuracy rate of 91.0%.

Song *et al*. [14], presented a dual-mode VGG-CNN classification model using auto fluorescence and white light images. Their model not only attained an accuracy of 86.9% but also demonstrated robust performance by applying a 4-fold cross-validation process. Heba *et al*. [15] introduced a novel approach using transfer learning approach to predict OSCC in histopathological images. The study utilized Grad-CAM mapping to locate lesions. Their model used a dataset of 1224 oral histopathological images. Notably, ResNet-101 achieved 100% accuracy, and EfficientNet-b0 reached 95.65% accuracy demonstrating superior predictive abilities in oral cancer. Zhou *et al*. [16] analyzed 785 clinical oral images, categorizing them into three groups: 251 RAU images, 271normal oral mucosa, and 263 other common conditions. Notably, the pretrained ResNet50 model outperformed others, achieving a precision rate of 92.86%. These impressive results highlight the strong potential of DLTs in the classification process using non-invasive oral lesion images, offering promising prospects for clinical applications in this domain.

The CNN is composed of several hidden layers, including convolution, pooling, and Fully Connected (FC) layers, positioned between the input and output layers [16]. The quality of the features extracted by the CNN is heavily reliant on the network's weights, which encompass the weights of convolution filters and connecting edges within the fc layer. These network weights play a pivotal role in determining the accuracy of classification. Throughout the training period of the model, network weights are continuously adjusted to minimize the errors. The most widely used method for modifying weights in neural networks is back-propagation, which employs a Gradient Descent optimization technique. However, the limitation of the Back-Propagation (BP) algorithm lies in its requirement for extended convergence time and the possibility of becoming trapped in local optima. Local optima occur when, in the pursuit of weights that minimize classification errors, the back-propagation algorithm may settle for weights with lesser error values at neighboring points, which may not essentially be the absolute minimum across all possible points.

To address these drawbacks associated with back-propagation researchers are applying other well-known optimization methods such as the Genetic Algorithm (GA), PSO algorithm and Ant Colony Optimization (ACO). GA is a well-established global optimization technique inspired by the principles of natural selection, offering an alternative approach to optimizing neural network weights [17]. PSO is a probabilistic optimization method that utilizes the notion of social interaction as a means to address search problems. It systematically employs multiple entities in an iterative manner to accomplish its optimization objective [18]. ACO emulates ant colonies' searching behavior, excelling in global and local searches [19]. Nevertheless, the pursuit of novel CNN optimization methods continues. Despite numerous existing algorithms, ongoing research aims to develop innovative approaches that can surpass limitations and elevate CNN training performance.

In this research we exploited the GA for CNN optimization. GA-based optimization in CNN excels over BP gradient descent by offering global search, automated hyperparameter tuning, robustness to noise, and versatility for diverse objectives, parallelization ease, exploratory power, and gradient-independence. Backpropagation is sensitive to local optima, necessitating manual hyperparameter tuning, and relies on gradients, making GAs advantageous for complex CNN optimization tasks. In principle, a CNN performance relies on hyperparameter settings, including layer count, filter numbers, and sizes. Loussaief *et al*. [20] created a CNN model for sign stop image classification. To tackle the vast solution space, authors developed an E-CNN-MP framework using GA which generates an Elite CNN model with ~90% accuracy without transfer learning. This GA applied CNN model produced an accuracy of 98.94%.

Mohanad *et al.* [21] proposed a CAD methodology for oral cancer detection. The dataset applied consist of two distinct categories: normal oral cavity epithelium (NEOR) and OSCC images. The extraction of features from the dataset was executed using 4 DL models, namely VGG-16, AlexNet, ResNet-50, and InceptionV3 and in order to identify the most valuable features, Binary PSO is applied. Pan *et al.* [22] considered three neural network models namely BP, GA-BP, and Probabilistic GA-BP to simulate radiomics data for oral tongue cancer patients. The evaluation involved 59 patients having cancerous lesions in the tongue area. The study showcased that the neural network models optimized using the suggested probabilistic genetic algorithm, in conjunction with the t-SNE dimensionality reduction technique, successfully achieved precise predictions of patient survival.

Huang *et al.* [23] introduced a novel method, combining DLT and a metaheuristic approach. The optimal selection of network weights is accomplished using ISSA (an advanced squirrel search algorithm), aiming for improved accuracy for precise cancer diagnosis. Results outperform existing methods in oral cancer diagnosis. Hadjouni *et al.* [24] introduced an improved approach using a CNN and an optimized DBN. Hyperparameters were optimized with a novel hybrid algorithm, PSO-Al-Biruni Earth Radius (BER) outperforming other methods with 97.35% accuracy on a Kaggle dataset. Statistical tests confirm the method's

significance and stability, making it a promising tool for specialists, though further research on larger datasets is needed.

Radwa *et al*. [25] introduced an DL model for oral cancer detection and classification by applying fuzzy-based contrast enhancement for data pre-processing, DenseNet-169 for feature extraction and COA with AE for detection and classification. Experimental results demonstrate their proposed method outperforms other methods with 90.08% accuracy. Warin *et al.* [26] This study assesses deep CNN algorithms' performance in classifying and detecting cancer lesions from oral images. Four CNN models (DenseNet169, ResNet101, SqueezeNet, and Swin-S) were used for multiclass image classification, and four object detection models (RCNN, YOLO, RetinaNet, and CenterNet2) were employed. Results indicate AUC values ranging from 0.71 to 1.00 for OSCC and 0.80 to 0.98 for OPMDs in image classification, while detection models achieved AUCs of 0.81 to 0.91 for OSCC and 0.34 to 0.64 for OPMDs.

Thus, considering the benefit of the GA for CNN optimization, this study delves into the implementation of a novel method of applying the benefits of GA for optimization of the CNN model. For the problem of oral cancer detection from photographic images, applying GA offers substantial advantages. GAs automates the optimization of critical CNN hyperparameters, such as kernel sizes and layer configurations, reducing the need for manual fine-tuning. Their global search capability ensures a thorough exploration of solution space, mitigating the risk of converging to suboptimal outcomes often faced in gradient-based methods. GAs exhibits robustness in handling noisy medical image data, leading to more stable and reliable CNN models for oral cancer diagnosis. Additionally, GA is versatile, adaptable for various objectives like architecture design and feature selection. Their inherent parallelizability facilitates efficient distributed computing, while their exploratory power allows for the discovery of novel network structures. Importantly, GAs do not rely on gradients, making them suitable for non-differentiable loss functions and complex scenarios typical in medical image analysis, further enhancing their applicability in oral cancer detection.

III. Materials and Methods

The proposed work introduces a fully automated system aimed at improving the identification and classification of oral cancer lesions from photographic images by extending the capabilities of CNN models. CNN models have demonstrated remarkable ability in image classification tasks, but optimizing their hyperparameters can be an arduous task due to the extensive search space involved. The research seeks to tackle this issue by applying the GA to fine-tune the hyperparameters of CNNs specifically for oral cancer detection.

The proposed CNN-based model, optimized via the genetic algorithm, undergoes training and testing using an oral cancer photographic image dataset. The study includes an evaluation of its performance using various metrics like Accuracy, F1-Score, Recall and Precision. The proposed work also explores McNemar's test for correctness of the proposed model.

A. Convolutional Neural Network (CNN)

CNN is a prominent deep learning model employed widely for tasks related to image analysis and computer vision. CNNs excel at recognizing visual patterns by processing images through different convolutional and pooling layers. CNN layers employ filters that convolve with input images, generating activation maps to highlight specific features. Subsequently, pooling layers down sample these maps, reducing data size while preserving vital information. CNNs often integrate fully connected layers at the end of the architecture, utilizing the output from convolution and max pool layers for classifying images into various categories. CNNs have consistently delivered cutting-edge results across image classification, segmentation, object detection etc. Additionally, CNNs have also found applications in domains such as NLP and speech recognition.

A generalized CNN as shown in Fig. 2 refers to a CNN architecture that is designed to be adaptable and applicable for a wide range of image recognition and classification tasks. Instead of specifying a fixed architecture, a Generalized CNN typically incorporates flexible components and modules that can be customized or adjusted based on the specific task at hand.

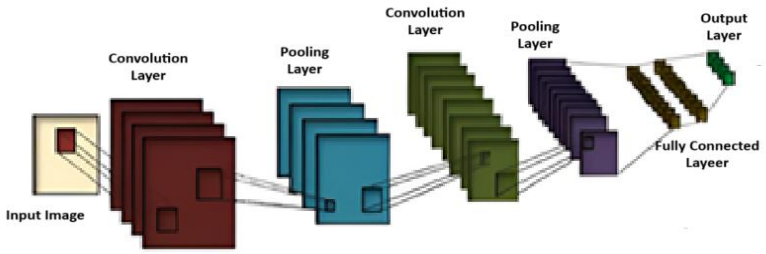


Fig. 2. Generalized CNN model.

An outline of the structural details and common components found in a generalized CNN are as follows:

1. Input Layer takes the input in the form of an image or numerical data.
2. Convolutional Layers comprise of multiple filters (kernels) that move over the input data to extract features. The number of filters and their

dimensions can be adjusted based on task complexity.
3. Pooling Layers are added to down sample feature maps and reduce the spatial dimensions by applying max-pooling or average pooling based on the application.
4. Normalization Layers are included to stabilize and speed up training. These layers adapt to the statistics of each mini-batch during training, improving convergence.
5. Dropout Layers are employed for regularization by arbitrarily dropping a fraction of neurons in training. The dropout rate can be adjusted as needed to control overfitting.
6. Fully Connected (FC) Layers are the final layers added to aggregate extracted features.
7. Output Layer typically has as many neurons as there are classes and uses SoftMax activation.

The classification process consists of following steps:

1. Input images are transferred to pretrained convolution layer with filter size: 32, Kernel size: 4×4.
2. ReLU activation function is applied to introduce non-linearity.
3. Result is then given to MaxPooling layer to down sample the image.
4. A flatten layer is added that converts the image to 1D array,
5. Dropout rate of 0.5 is introduced through dropout layer.
6. SoftMax activation function is used in final layer for classification.

The success of CNN largely depends on their network structure. Traditionally, designing CNN architectures has been a manual process, which becomes challenging for achieving the best designs in classifying oral cancer images due to a lack of expertise in this field.

B. Genetic Algorithm

A Genetic Algorithm (GA) is a search heuristic inspired by the process of natural evolution. It is used to optimize the problems to find the approximate solution. Genetic Algorithm (GA) greatly enhances the performance of CNN by efficiently searching for the best hyperparameters and configurations. This optimization maximizes accuracy, sensitivity, specificity, and F1-Score while also reducing training time, making the model more efficient.

The GA process typically consists of six stages: initialization, fitness function, selection, crossover, mutation, and termination condition check, as illustrated in Fig. 3. In the initialization step, an initial population of candidate architectures are created on which genetic operations will be applied. Generated solution is then represented as a chromosome, often a binary string or an array of values. Size of the population is dependent on the target problem and can go up to a few hundreds and thousands.

After the genetic population is initialized, their fitness scores are computed and a predefined fitness function is employed to identify all the favorable solutions. Determining the fitness function is a crucial step in the GA process, as it can significantly impact the model's performance. An inappropriate fitness function definition for the target problem can lead the GA to converge on the wrong solution.

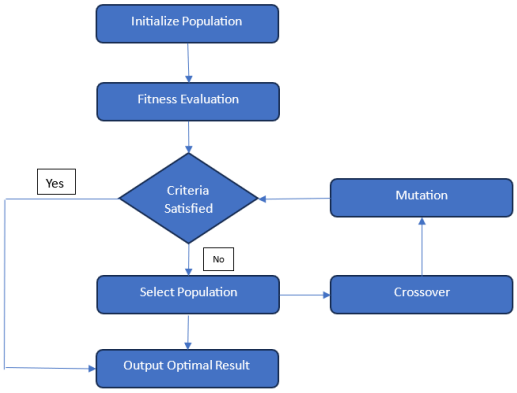


Fig. 3. Stages of genetic algorithm.

After that a fitness value is assigned to each candidate solution in the population. The fitness function enumerates how fit is the solution for the problem and directs the search towards improved solutions.

The selection step chooses individuals from the current candidates depending on their fitness value to form a mating pool. During the fitness calculation, candidates with greater fitness value have a greater chance of getting selected. The core concept is to favor better solutions and allow fittest individuals to pass their genetic information to the next generation. Crossover represents a pivotal stage in GA operations. Two pairs of individuals are taken from the generation pool and a crossover breeding is performed by swapping or mixing parts of the chromosomes of two parents to create new individuals. Two chosen chromosomes exchange corresponding sections of their genetic strings, altering gene combinations to generate new candidate solutions. Fig. 4 provides an illustration of the crossover process. Crossover generates offspring or new individuals that inherit traits from their parents.

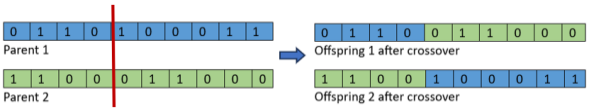


Fig. 4. Crossover of 2 parents to generate 2 new offspring.

The crossover stage has limitations in generating entirely new information, which can be addressed by mutation stage by flipping individual bits of selected chromosomes to a different value. The whole procedure is aimed to preserve genetic variety in the population and to avoid falling into local optima. Mutation step introduces small, arbitrary changes to individual chromosomes, diversifying the population and preventing premature convergence. The old population is then replaced with the new one formed by the offspring and potentially some surviving individuals from the previous

generation. The replacement strategy can be generational or steady-state.

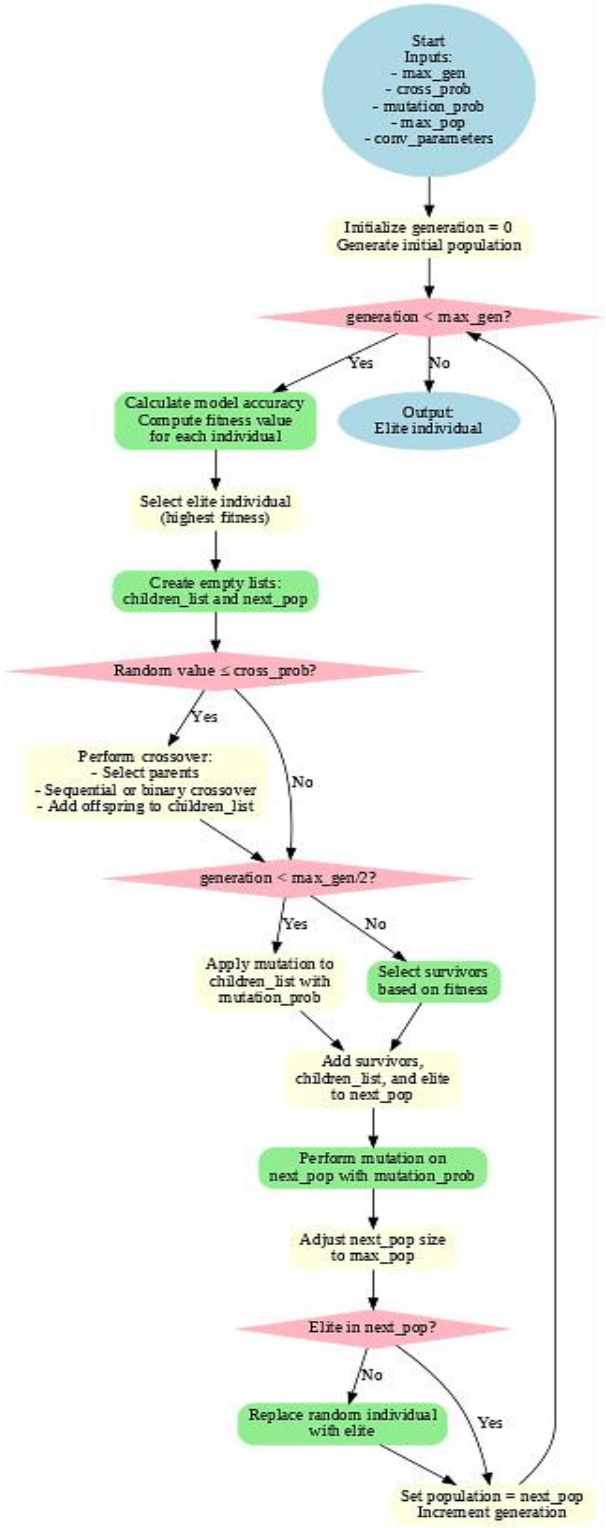


Fig. 5. Outline of genetic algorithm.

The newly generated chromosomes through these genetic operations are then assessed to determine if they have met any of the termination condition. The GA terminates when a acceptable fitness level is achieved. Otherwise, the entire process is iterated from stage 2 until the termination condition is met. Termination of the process might be attainment of a solution that meets the criteria of completing a defined number of generations or reaching a stage where better results are no longer being produced.

Fig. 5 outlines a Genetic Algorithm (GA) for hyperparameter optimization of a CNN with the goal of obtaining an elite individual that represents the best suitable hyper-parameters for a given task.

Input:

1. max_gen: Maximum number of generations.
2. cross_prob: Probability of crossover
3. mutation_prob: Probability of mutation.
4. max_pop: Maximum population size.
5. conv_parameters: CNN parameters.

Output:

elite: The best-performing individual found during the optimization process.

Our proposed approach seeks to overcome this challenge by using Genetic Algorithms (GAs) to automatically design CNN models by tuning hyperparameters. GAs have shown exceptional effectiveness, especially in tasks like medical image classification where there is a shortage of data samples. By harnessing GAs, we streamline the process, making it possible to discover the most optimal CNN model without requiring manual intervention. This, in turn, improves efficiency and effectiveness in addressing the specific problem of oral cancer image classification.

C. Pre-trained Models Used with CNN

1) DenseNet-201

DenseNet-201 is a CNN architecture known for its dense connectivity design, efficient parameter sharing, and excellent performance on various computer vision tasks. DenseNet-201's key innovation lies in its dense connectivity pattern, where feature maps from all earlier layers are combined and used as inputs to each layer within a dense block. This pattern enhances feature reuse, mitigates vanishing gradient problems, and allows for very deep networks with lesser parameters. DenseNet-201 works effectively for image classification and object detection. DenseNet-201 takes a unique approach by connecting every layer to other layers in a forward-feeding style. In this design, the feature-maps from all preceding layers serve as inputs for subsequent layer, and their feature-maps continue to be utilized in all subsequent layers.

2) EfficientNetB0

EfficientNetB0 is from the family of EfficientNet CNN model. This architecture is known for its efficiency and high performance and are designed to achieve better results while maintaining computational efficiency. EfficientNetB0, with its efficient architecture and scaling mechanisms, is capable of achieving high accuracy on various computer vision tasks while minimizing computational resources. Its structural details, including depth wise separable convolution, width and depth scaling, and efficient block repetitions, contribute to its effectiveness and efficiency.

3) ResNet50

Residual Network, is a deep CNN architecture developed to address the challenge of training very deep neural networks. ResNet employs a unique building block called the residual block, which enables the training of tremendously deep networks by allowing skip connections, also known as shortcuts, allow the direct flow of data from one layer to a deeper layer in the network, thus deleting the vanishing gradient problem. Each residual block consists of a set of convolutional layers and the addition of the actual input to the generated output, creating a residual mapping.

4) MobilenetV2

MobileNetV2 is a lightweight CNN architecture developed for mobile and edge devices with resource constraints. Introduced by Google, Key features of MobileNetV2 include the use of inverted residuals and linear bottlenecks, which help capture more complex features while preserving computational efficacy. The model incorporates a blend of depth wise separable convolutions, linear bottlenecks, and shortcut connections, allowing for faster training and reduced computational demands. MobileNetV2 is particularly well-suited for real-time image classification tasks and serves as a feature extractor for applications like object detection and semantic segmentation in situations where computational resources are limited. Its versatility and balanced performance make it a prevalent choice for on-device vision applications. Table I represents the hyperparameters of the 4 pretrained models used in the research.

TABLE I. HYPERPARAMETERS OF THE PRETRAINED MODELS

Pretrained Models	Layers Count	Parameters (millions)
MobileNetV2 [27]	55	3.40
EfficientNetB0 [28]	128	5.30
ResNet50 [29]	50	23.0
DenseNet-201 [30]	201	20.0

These well-established CNN architectures have demonstrated impressive performance in various deep learning classification task. However, the models do come with increased parameter counts and computation time.

D. Synthetic Minority Over-Sampling Technique

SMOTE is a technique for addressing disproportion in datasets. Class unevenness occurs when the minority class has significant lesser instances than the other class leading to biased and less accurate predictions [31]. SMOTE works by creating artificial samples in the minority class to balance the dataset as shown in Fig. 6.

For each minority instance, it creates synthetic instances by linearly interpolating between the instance and its selected nearest neighbors. This involves randomly selecting a neighbor, calculating the difference between features, and applying this difference to the minority instance. The artificial instances are then added to the original dataset, successfully oversampling the minority class. SMOTE helps to alleviate the class imbalance problem by introducing diversity to the minority class, making the dataset more balanced.

This can lead to better generalization and performance of learning models, especially in situations where the imbalance is severe. It's important to note that while SMOTE is a powerful technique, it may not be suitable for all scenarios, and its effectiveness can depend on the characteristics of the data.

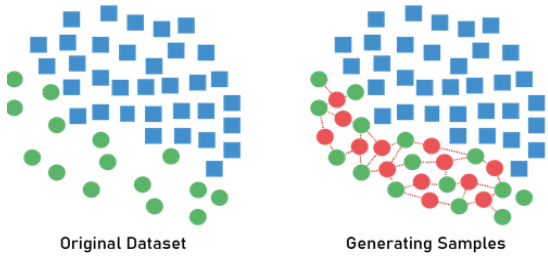


Fig. 6. SMOTE sampling.

E. McNemar's Test

McNemar's Test is a statistical test for paired proportions data typically in a binary classification setting. It is commonly used to determine if the two classification models significantly differ in their predictions, especially when working with matched pairs of observations. In the context of Machine and Deep Learning, McNemar's test can be applied to assess whether the two models' predictions are statistically significant or not. In other words, McNemar's Test can be used to compare the predictions of two models [32].

McNemar's Test, is based on the criteria that the probabilities of two models are same if none of them performs better than the other. Vice Versa, the alternate hypothesis says that the two models are not equal. McNemar's test generates a 2 by 2 contingency matrix from the proposed model's predictions. The steps required for creating a contingency table are:

1. Train both CNN models on the same dataset.
2. Generate predictions for each model on the same set of test data.
3. Create a Contingency Table based on the predictions of the two models as shown in Fig. 7.

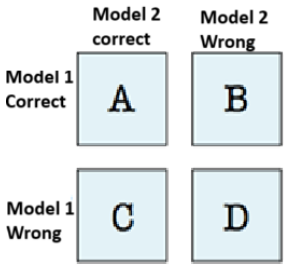


Fig. 7. Contingency matrix for comparing predictions of 2 models.

A: Instances correctly predicted by both models (True Positives, TP).

B: Instances correctly predicted by first Model but incorrectly by second Model (False Positives, FP).

C: Instances incorrectly predicted by first Model but correctly by second Model (False Negatives, FN).

D: Instances incorrectly predicted by both models (True Negatives, TN).

The McNemar statistic test can then be calculated as:

$$\chi^2 = (B - C)^2 / (B + C)^2$$

The above equation represents if the summation of cell B and C is large enough, the χ^2 value works as a chi-squared test with a single degree freedom. If specifying a significant threshold, e.g., $\alpha = 0.05$, p-value, the probability of detecting the chi-squared value is computed. If the p-value is below the chosen significant threshold, the null hypothesis can be rejected representing that the models do not differ significantly.

IV. Model Formulation

A. Data Description

The effectiveness of any DL-CNN model is reliant upon the quality of the dataset used for its training. In the paper, we compiled an oral photographic image dataset through a combination of images from Refs. [33] and [34]. These images exhibit inherent diversity, varying in terms of size, lighting conditions, and capturing different regions of the oral cavity, as perceived by physicians. In Fig. 1, a selection of noncancerous and cancerous image samples is presented, showcasing the dataset's diversity. Additionally, Fig. 8 illustrates sample images from the dataset, highlighting variations in the position of the oral cancer region.

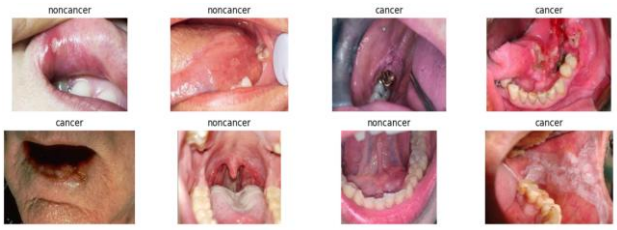


Fig. 8. Sample images from the dataset.

B. Data Preprocessing

Before being incorporated into the model, each image underwent rigorous manual verification and annotation by a certified medical practitioner. All the images were standardized to the JPEG format. The dataset was organized into two distinct directories, categorized as "cancerous" and "non-cancerous". In total, there are 1000 images depicting various aspects of the mouth, lips, and tongue, with 700 belonging to the cancerous category and 300 to the non-cancerous category.

To ensure uniformity, all images were resized to dimensions of 224×224×3 pixels and normalized within the 0–1 range. For the purpose of training and testing, the dataset was randomly divided in 80:20 ratio.

C. Balancing and Augmenting Dataset

CNN models demonstrate improved performance with balanced datasets. To achieve this, the dataset is expanded by generating synthetic samples for the existing minor class. This practice, known as sampling, is done using SMOTE. Thus, the original dataset of 700 cancerous and 300 non-cancerous images is expanded by increasing the minority cases to 700 thus balancing the dataset using SMOTE. The results of sampling are illustrated in Fig. 9. CNNs excel in handling variations in images, even when objects are positioned in different orientations, which forms the foundation of data augmentation. Training CNNs effectively requires ample data for learning and feature extraction.

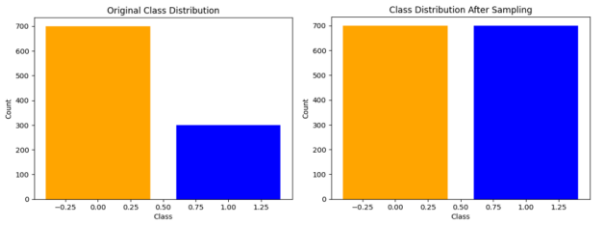


Fig. 9. Class distribution before and after sampling

Thus, augmenting existing data through minor image modifications like scaling, rotation, translation, etc., becomes a valuable approach. Augmenting the dataset by generating more training examples enhances the model's robustness and generalization, consequently improving accuracy. All the input images underwent multiple augmentation stages, like flipping, rotating, scaling, zooming etc with specific ranges as detailed in Table II. Sample images after augmentation is shown in Fig. 10.

TABLE II. Ranges of Various Augmentation Techniques

Augmentation Techniques	Range
Scaling	1/255
Rotation	45°
Breadth Shift	2%
Length Shift	2%
Shearing	4%
Zooming	3%

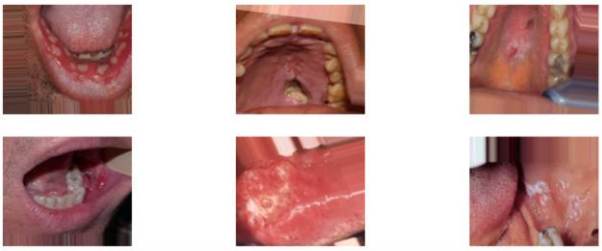


Fig. 10. Sample images after augmentation.

D. Proposed Model

Fig. 11 illustrates the model's chronological steps. Initially, image pre-processing is applied on the input dataset, incorporating oversampling and image augmentation to balance the imbalanced dataset and to improve the dataset size. Subsequently, a GA is utilized to select the optimal hyperparameters, resulting in an optimized CNN model.

The formulation of our proposed DL-CNN is illustrated in the following steps:

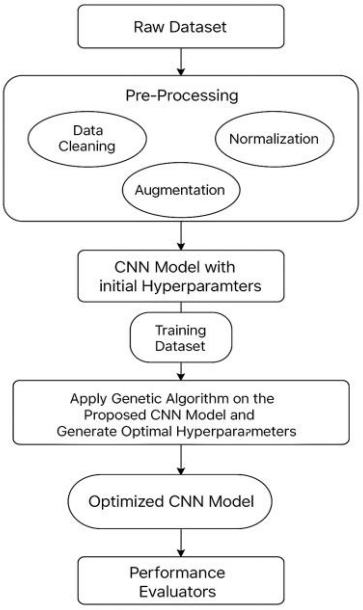


Fig. 11. Proposed model.

1. Load the oral cancer dataset.
2. Perform pre-processing on the dataset.
   i. Apply data sampling using SMOTE.
   ii. Images are resized to 224×224×3 and normalized within 0−1 range.
   iii. Data augmentation such as rotation, flip, shear, zoom etc is applied.
3. Create training and testing datasets in 80:20 ratio.
4. Initialize CNN Hyperparameters to the values as mentioned in Table III.
5. Select four pre-trained models namely, DenseNet201, MobileNetV2, Resnet50 and EfficientNetB0 and train them by applying transfer learning.
6. The top 2 best performing models (DenseNet201 and MobileNetV2) as shown in Table IV are selected for optimization.
7. Apply Genetic Algorithm
   i. Initialize GA hyperparameters to the values as mentioned in Table V.
   ii. Initialize Population.
   iii. Evaluate the fitness value of each candidate from the population using the fitness function.
   iv. Select the best performing individual (elite) based on their fitness scores.
   v. Create a new population through crossover and mutation operations:
      a. Randomly select parents from the elite population.
      b. Perform crossover to create children with characteristics from both parents.
      c. Apply mutation to some individuals with a certain probability.
      d. Construct the new population by combining elite individuals, children, and mutated individuals.
   vi. Repeat the process till all the generations are trained.
8. Evaluate the models on the test dataset to assess its performance.

V. Experimental Results

To evaluate the proposed work, the oral cancer dataset is divided into cancerous and noncancerous directories. The original dataset consists of 700 photographic images under cancerous label and 300 images under noncancerous label, forming a total of 1000 photographic images. The dataset is then balanced using SMOTE technique thus generating 1400 images for the final dataset to work with. The complete data is then divided for training and testing ratios of 80:20 containing equal images of cancerous and noncancerous lesions.

The research work is carried out using Keras and TensorFlow. The Python programming language, in conjunction with Google Colab, serves as the platform for training and testing the entire model, all accomplished using a laptop. Google Colab is a freely accessible Jupyter notebook environment that operates seamlessly in the cloud, requiring no configuration. This environment comes pre-equipped with various ML, DL and AI libraries essential for the implementation. As Google Colab offers free GPU support, the training time for oral cancer classification using CNNs for the provided dataset, can be completed within a few hours using Colab's GPUs. However, for genetic algorithm optimization, high-performance GPUs or TPUs are required to achieve optimal performance.

A. Performance Evaluators

The proposed model is evaluated using various evaluators like Accuracy, Precision, Recall, and F1-Score. Precision is defined as in Eq. (1),

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \quad (1)$$

Where TP, FP, TN, and FN values represent True-positives, false-positives, true-negatives and false-negatives, respectively.

The recall is then calculated as in Eq. (2),

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (2)$$

The correctness of the model is computed from F1-Score. F1-Score is defined as in Eq. (3)

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

The accuracy is of the proposed model is computed as in Eq. (4)

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} \times 100\% \quad (4)$$

B. Results of Pre-trained Models Using Transfer Learning

The hyperparameters of the proposed model used are set to the values as mentioned in the Table III. The results of the four pretrained candidate models on the oral cancer dataset is shown in Table IV.

TABLE III. HYPERPARAMETERS OF THE CNN MODEL

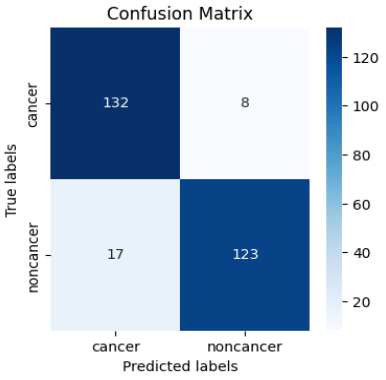
Initial Hyperparameters	Value	Description
Input size	224×224	Input image Size
Batch	32	No. of images taken at a time
Epochs	20	No. of iterations in one cycle
Learning rate	Le–3	Rate at which algorithm learns
inimum_delta	0.01	Minimum change required for the upgrading
Patience	3	No. of epochs without any difference which leads to the halt of training phase
Optimizer	Adam	Optimization algorithm

TABLE IV. PERFORMANCE OF THE PRETRAINED CANDIDATE MODELS

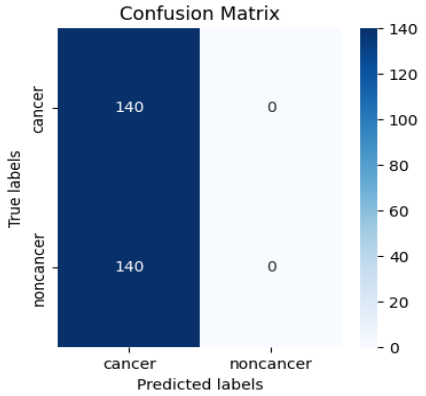
Model	Precision	Recall	F1-Measure	Accuracy
Pretrained MobileNetV2	0.86	0.84	0.90	0.91
Pretrained EfficientNetB0	0.50	0.71	0.67	0.51
Pretrained DenseNet201	0.89	0,87	0.93	0.94
Pretrained ResNet50	0.65	0.64	0.64	0.65

As observed EfficientNetB0 and ResNet50 does not perform well compared to DenseNet201 and MobilNetV2 giving an accuracy of 51% and 65% respectively. EfficientNetB0 and Resnet50 models does not align well with the fine-grained patterns required for medical imaging. Thus, effecting their overall performance.

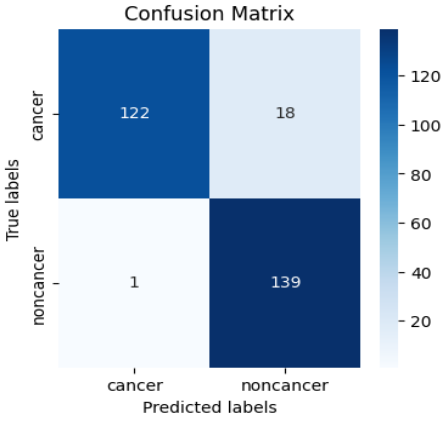
Figs. 12 and 13 show the results of the 4 pre-trained models using confusion matrices and accuracy and loss graphs on the dataset. Confusion matrix represents the predicted labels and the true labels on the testing data.



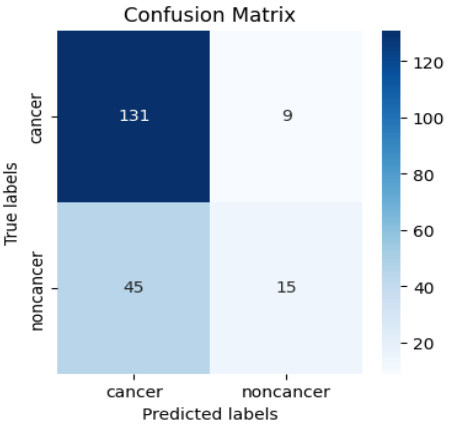
(a)



(b)

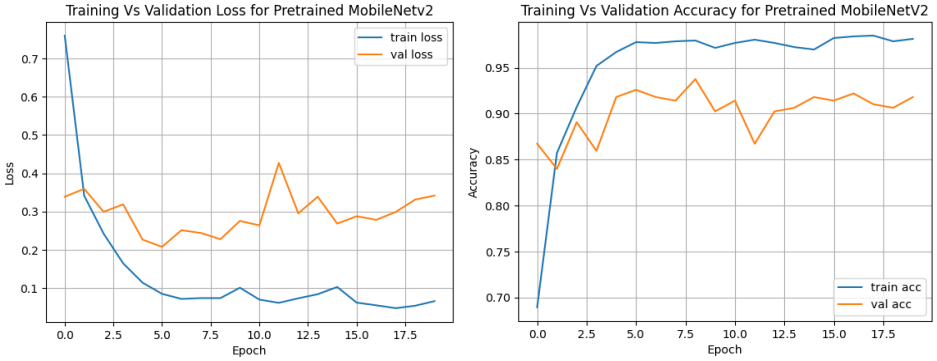


(c)

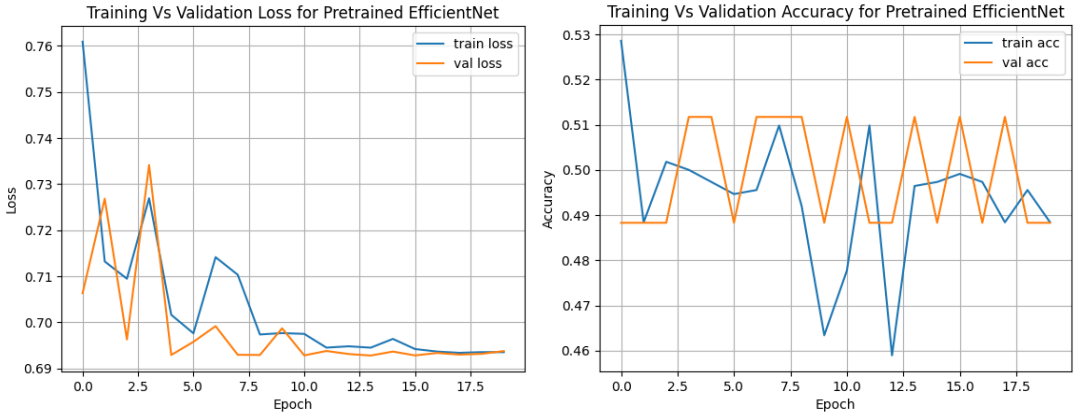


(d)

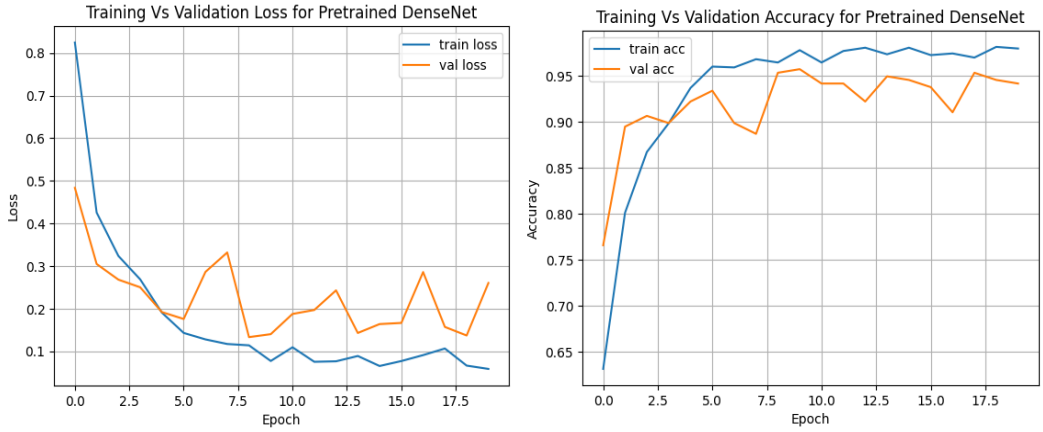
Fig. 12. Confusion matrix for pre-trained CNN models. (a) MobileNetV2 (b) EfficientNetB0 (c) DenseNet201 (d) ResNet50.



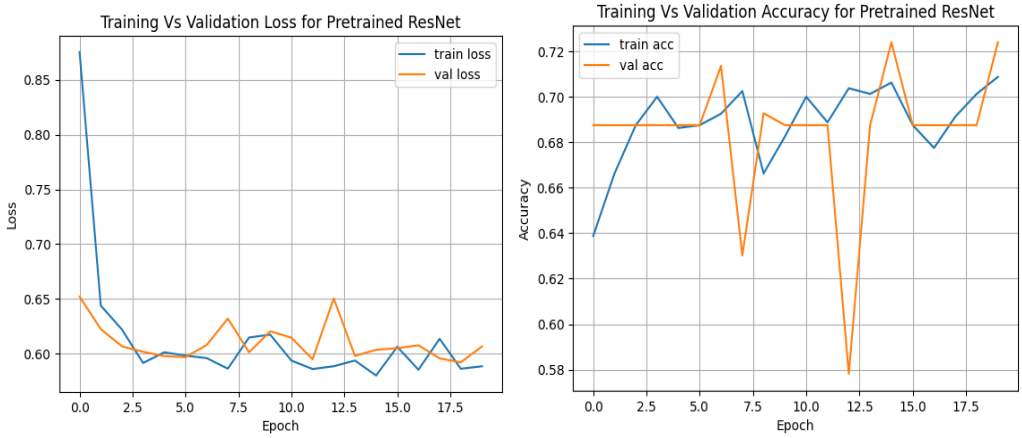
(a)



(b)



(c)



(d)

Fig. 13. Accuracy and Loss Graphs of the pre-Trained CNN Models. (a) MobileNetV2 (b)EfficientNetB0 (c) DenseNet201 (d) ResNet50.

C. *Results of Optimized Models Using Genetic Algorithm*

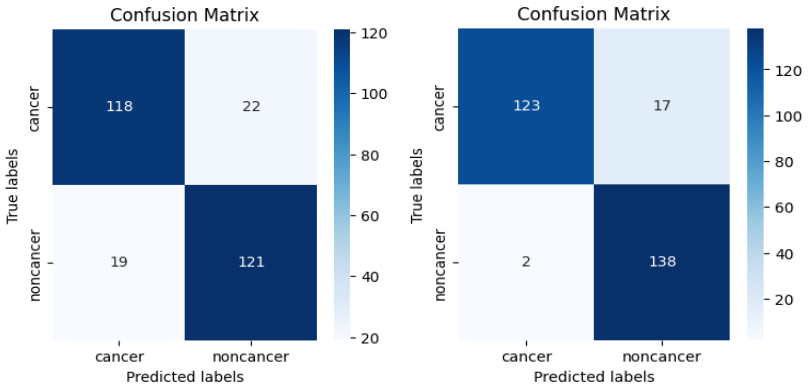
Hyperparameters of the Genetic Algorithm required for tuning of the algorithm is shown in Table V. The values used in GA hyperparameters are often based on empirical data or experience. The authors have considered various combinations experimentally to find parameters which could work best for the problem. The hyperparameters that significantly impacted the performance and efficiency of the genetic algorithm were chosen. After applying Genetic Algorithm for optimizing the pre-trained DenseNet201 and MobileNetV2 CNN models, it was found that DenseNet201 performed extremely well on the dataset provided with an accuracy of 94% compared to MobileNetv2. Table VI Shows the performance of GA optimized MobileNetV2 and GA optimized DenseNet201 on the dataset. Figs. 14 and 15 shows the confusion matrices and Accuracy and Loss graphs for optimized MobileNetV2 and DenseNet20. Fig. 16 Shows the ROC curve and the comparison of performance evaluators for the optimized models. Fig. 17 Depicts the final result of the proposed model which is Optimized DenseNet201 with Genetic Algorithm on the oral cancer dataset.

TABLE V. HYPERPARAMETER TUNING USING GENETIC ALGORITHM

GA HyperParameters	Value	Description
Individual size	3	No. of architecture parameters
Population size	10	No. of different CNN configurations evaluated in each generation
Crossover probability	0.7	Likelihood that two parent CNN configurations will undergo crossover to produce offspring configurations
Mutation rate	0.2	The probability that random changes will be introduced to an individual's genes
Number of generations	10	The total number of iterations the genetic algorithm will run.

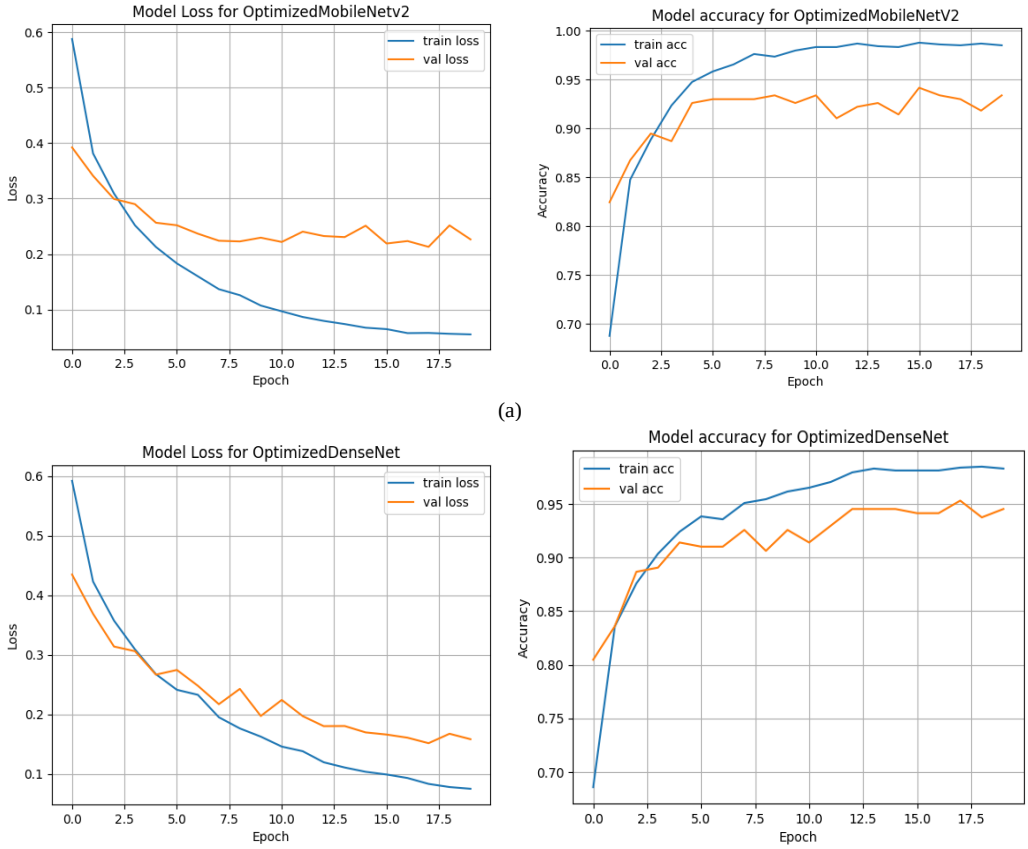
TABLE VI. PERFORMANCE EVALUATORS OF GENETIC ALGORITHM OPTIMIZED MODELS

GA Optimized Model	Precision	Recall	F1-Score	Accuracy
Optimized MobileNetV2 with GA	0.91	0.91	0.92	0.92
Optimized DenseNet201 with GA	0.91	0.90	0.93	0.95



(a) (b)

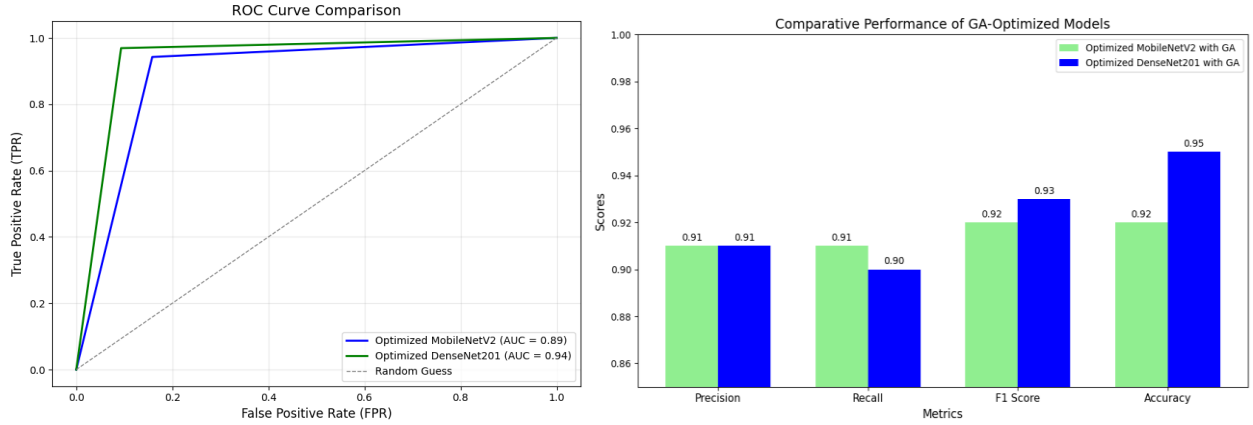
Fig. 14. Confusion matrices of optimized models. (a) Optimized MobileNetV2 with GA. (b) Optimized DenseNet201 with GA.



(a)

(b)

Fig. 15. Accuracy and loss graphs of the optimized models. (a) Optimized MobileNetV2 with GA. (b) Optimized DenseNet201 with GA.



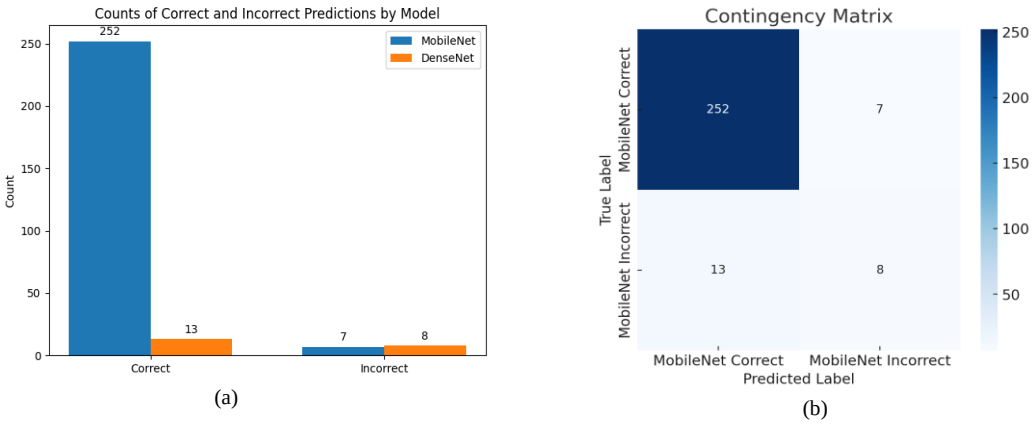
(a) (b)

Fig. 16. Performance evaluators of optimized models. (a) ROC curve (b) Bar graph.

Fig. 17. Detection of cancerous and non-cancerous lesions by DenseNet201 on the dataset.

D. *Model Evaluation with McNemar's Test*

McNemar's Test was applied on the Optimized MobileNetV2 and Optimized DenseNet201 to observe if the two models vary statistically in their performance while classifying the oral cancers as benign and malignant. McNemar's test calculates a *p*-value based on the false positives and false negatives and creates a contingency table showing the correct and incorrect predictions made by the 2 models. A *p*-value greater than 0.05 shows that the models doesn't differ statistically. The McNemar's test on Optimized MobileNetV2 vs. Optimized DenseNet201 generated a *p*-value of 0.1686 which shows that there is no statistically significant difference in the performance of the two models. Fig. 18 illustrates the results of McNemar's Test.



(a) (b)

Fig. 18. Evaluation of McNemar's test (a) Bar graph representing the correct and incorrect predictions (b) Contingency matrix.

E. *Proposed Work Compared to Other Notable Work*

We compared the proposed GA optimized DenseNet201 CNN Model's performance with other notable works using photographic images and histopathological images. It was observed that the proposed work performed better than many notable works in oral cancer detection on photographic images using deep learning models.

Table VII shows that many of the notable works done on photographic images has good accuracy but with a smaller dataset. Whereas our proposed model achieves better accuracy with a bigger dataset of 1400 oral cancer photographic images.

TABLE VII. COMPARISON OF NOTABLE WORKS WITH PROPOSED MODEL

Paper	Dataset	Dataset Size	Accuracy
Huang *et al*. (2023) [23]	photographic images	131	95.20%
Myriam *et al*. (2023) [24]	photographic images	131	97.00%
Marzouk *et al*. (2022) [25]	photographic images	131	87.00%
Warin *et al*. (2022) [26]	photographic images	980	98.00%
Huang *et al*. (2024) [35]	photographic images	131	96.94%
Lin *et al*. (2021) [36]	Smartphone-based oral cancer images	1500	83.00%
Proposed Model	Photographic images	1400	95.00%

Work presented in [37–40] were able to achieve a greater accuracy using various optimization techniques but used histopathological image dataset. Whereas the proposed work focusses on photographic image dataset and achieved an accuracy of 95%.

VI. CONCLUSION

The proposed work is aimed to perform classification and detection and binary classification of cancerous and non-cancerous oral lesions from photographic images by implementing a GA optimized CNN model. For this, initially four pretrained CNN models MobileNetV2, EfficientNetB0, ResNet50, and DenseNet201 were selected by applying transfer learning. The pre-trained models are then modified by adding additional layers. The best performing models MobileNetV2 and DenseNet201 are then selected and optimized using Genetic Algorithm on the oral cancer dataset containing photographic images of mouth lesions. The proposed work showed that GA optimized DenseNet201 CNN model outperformed other modified models to achieve an accuracy of 94%. The proposed model is also found to be better than many other deep learning models used for oral cancer detection. Thus, the proposed GA optimized CNN has performed significantly better than many other CNN models for the detection and classification of tumours from oral photographic images.

CONFLICT OF INTEREST

Authors declare that there is no conflict of interest.

AUTHOR CONTRIBUTIONS

Ms. Sayyada Hajera Begum contributed to the experimental results and paper writing. Dr. P. Vidyullatha contributed to the paper writing. All authors had approved the final version.

REFERENCES

[1] Q. Fu *et al*., "A deep learning algorithm for detection of oral cavity squamous cell carcinoma from photographic images: A retrospective study," *EClinicalMedicine*, vol. 27, 100558, 2020.

[2] S. H. Begum and P. Vidyullatha, "A review of deep learning application in oral cancer prognosis," in *Proc. 2022 IEEE Pune Section International Conference (PuneCon)*, IEEE, 2022.

[3] S. Dixit, "A current review of machine learning and deep learning models in oral cancer diagnosis: Recent technologies, open challenges, and future research directions," *Diagnostics*, vol. 13, 1353, 2023. https://doi.org/10.3390/diagnostics13071353

[4] R. O. Alabi, A. Almangush, M. Elmusrati, and A. A. Makitie, "Deep machine learning for oral cancer: From precise diagnosis to precision medicine," *Front. Oral. Health*, vol. 2, 794248, 2022.

[5] M. Shaban, S. A. Khurram, M. M. Fraz, N. Alsubaie, I. Masood, S. Mushtaq *et al*., "A novel digital score for abundance of tumour infiltrating lymphocytes predicts disease free survival in oral squamous cell carcinoma," *Sci Rep*., vol. 9, 13341, 2019.

[6] J. Musulin, D. Štifanić, A. Zulijani, T. Ćabov, 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, 1784, 2021.

[7] N. Fujima, V. C. Andreu-Arasa, S. K. Meibom, G. A. Mercier, A. R. Salama, M. T. Truong *et al*., "Deep learning analysis using FDG-PET to predict treatment outcome in patients with oral cavity squamous cell carcinoma," *Eur Radiol*., vol. 30, pp. 6322–6330, 2020.

[8] Y. Kim, J. W. Kang, J. Kang, E. J. Kwon, M. Ha, Y. K. Kim *et al*., "Novel deep learning-based survival prediction for oral cancer by analyzing tumor infiltrating lymphocyte profiles through CIBERSORT," *OncoImmunology*, vol. 10, 1904573, 2021.

[9] M. Aubreville, C. Knipfer, N. Oetter, C. Jaremenko, E. Rodner, J. Denzler, C. Bohr, H. Neumann, F. Stelzle, and A. Maier, "Automatic classification of cancerous tissue in laser endomicroscopy images of the oral cavity using deep learning," *Sci. Rep*., vol. 7, 11979, 2017.

[10] R. A. Welikala, P. Remagnino, J. H. Lim, C. S. Chan, S. Rajendran, T. G. Kallarakkal, R. B. Zain, and R. D. Jayasinghe, "Automated detection and classification of oral lesions using deep learning for early detection of oral cancer," *IEEE Access*, vol. 8, pp. 132677–132693, 2020. https://doi.org/10.1109/access.2020.3010180

[11] M. Z. M. Shamim, S. Syed, M. Shiblee, M. Usman, S. J. Ali, H. S. Hussein, and M. Farrag, "Automated detection of oral pre-cancerous tongue lesions using deep learning for early diagnosis of oral cavity cancer," *The Computer Journal*, vol. 65, issue 1, pp. 91–104, January 2022. https://doi.org/10.1093/comjnl/bxaa136

[12] B. R. Nanditha, A. Geetha, H. S. Chandrashekar, M. S. Dinesh, and S. Murali, "An ensemble deep neural network approach for oral cancer screening," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 17, no. 02, pp. 121–134, 2021. https://doi.org/10.3991/ijoe.v17i02.19207

[13] S. H. Begum and P. Vidyullatha, "Deep learning model for automatic detection of Oral Squamous Cell Carcinoma (OSCC) using histopathological images," *International Journal of Computing and Digital Systems*, vol. 13, no. 1, 2023.

[14] B. Song, S. Sunny, R. D. Uthoff *et al*., "Automatic classification of dual-modality, smartphone-based oral dysplasia and malignancy images using deep learning," *Biomed. Opt. Express*, vol. 9, no. 11, pp. 5318–5329, 2018.

[15] H. M. Afify, K. K. Mohammed, and A. E. Hassanien, "Novel prediction model on OSCC histopathological images via deep transfer learning combined with Grad-CAM interpretation," *Biomedical Signal Processing and Control*, vol. 83, 2023.

[16] M. Zhou, W. Jie, F. Tang *et al*., "Deep learning algorithms for classification and detection of recurrent aphthous ulcerations using oral clinical photographic images," *Journal of Dental Sciences*, vol. 19, no. 1, pp. 254–260, 2023. https://doi.org/10.1016/j.jds.2023.04.022

[17] J. McCall, "Genetic algorithms for modelling and optimization," *Journal of Computational and Applied Mathematics*, vol. 184, pp. 205–222, 2005.

[18] J. Kennedy and R. Eberhart, "Particle Swarm Optimization (PSO)," in *Proc. IEEE International Conference on Neural Networks*, Perth, Australia, 1995, pp. 1942–1948.

[19] W. Ma, X. Zhou, H. Zhu, L. Li, L. Jiao, "A two-stage hybrid ant colony optimization for high-dimensional feature selection," *Pattern Recognit.*, vol. 116, 107933, 2021.

[20] S. Loussaief and A. Abdelkrim, "Convolutional neural network hyper-parameters optimization based on genetic algorithms," *Int. J.Adv. Comput. Sci. Appl.*, vol. 9, pp. 252–266, 2018.
[21] M. A. Deif, H. Attar, A. Amer, I. A. Elhaty, M. R. Khosravi, and A. A. A. Solyman, "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*, 2022. doi: 10.1155/2022/6364102
[22] X. Pan, T. Zhang, Q. Yang, D. Yang, J.-C. Rwigema, and X. S. Qi, "Survival prediction for oral tongue cancer patients via probabilistic genetic algorithm optimized neural network models," *Br. J. Radiol*., vol. 93, 20190825, 2020.
[23] Q. Huang, H. Ding, and N. Razmjooy, "Optimal deep learning neural network using ISSA for diagnosing the oral cancer," *Biomedical Signal Processing and Control*, vol, 84, 104749, 2023. https://doi.org/10.1016/j.bspc.2023.104749
[24] H. Myriam *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. doi: 10.1109/ACCESS.2023.3253430
[25] R. Marzouk, E. Alabdulkreem, S. Dhahbi, M. K. Nour, M. Al Duhayyim *et al.*, "Deep transfer learning driven oral cancer detection and classification model," *Computers, Materials & Continua*, vol. 73, no. 2, pp. 3905–3920, 2022.
[26] K. Warin, W. Limprasert, S. Suebnukarn, S. Jinaporntham, P. Jantana *et al*., "AI-based analysis of oral lesions using novel deep convolutional neural networks for early detection of oral cancer," *PLOS ONE*, vol. 17, no. 8, e0273508, 2022. https://doi.org/10.1371/journal.pone.0273508
[27] K. Dong *et al*., "MobileNetV2 model for image classification," in *Proc. 2020 2nd International Conference on Information Technology and Computer Application (ITCA)*, IEEE, 2020.
[28] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. International Conference on Machine Learning*, 2019.
[29] M. B. Sahaai, G. R. Jothilakshmi, and D. Ravikumar, "ResNet-50 based deep neural network using transfer learning for brain tumor classification," in *Proc. AIP Conference*, AIP Publishing, 2022, vol. 2463, no. 1.
[30] G. Huang *et al*., "Densely connected convolutional networks," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2261–2269.
[31] S. H. Begum *et al*., "A lightweight deep learning model for automatic diagnosis of lung cancer," in *Proc. 2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC)*, IEEE, 2022.
[32] B. Şener, K. Acici, and E. Sümer. "Categorization of Alzheimer's disease stages using deep learning approaches with McNemar's test," *PeerJ Computer Science*, voil. 10, e1877, 2024.
[33] Oral Cancer (Lips and Tongue) images. [Online]. Available: https://www.kaggle.com/shivam17299/oral-cancer-lips-and-tongue-images
[34] Oral cancer dataset. [Online]. Available: https://www.kaggle.com/datasets/zaidpy/oral-cancer-dataset
[35] Q. Huang, H. Ding, and N. Razmjooy, "Oral cancer detection using convolutional neural network optimized by combined seagull optimization algorithm," *Biomedical Signal Processing and Control*, vol. 87, 105546, 2024.
[36] H. Lin *et al*., "Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis," *Journal of Biomedical Optics*, vol. 26, no. 8, pp. 086007–086007, 2021.
[37] M. Badawy, H. M. Balaha, A. S. Maklad, A. M. Almars, and M. A. Elhosseini, "Revolutionizing oral cancer detection: an approach using aquila and gorilla algorithms optimized transfer learning-based CNNs," *Biomimetics*, vol. 8, no. 6, 499, 2023. https://doi.org/10.3390/biomimetics8060499
[38] I. A. Ahmed, E. M. Senan, and H. S. A. Shatnawi, "Analysis of histopathological images for early diagnosis of oral squamous cell carcinoma by hybrid systems based on CNN fusion features," *International Journal of Intelligent Systems*, 2662719, 2023.
[39] R. Dharani and S. Revathy, "DEEPORCD: Detection of oral cancer using deep learning," *Journal of Physics: Conference Series*, vol. 1911, no. 1, 2021.
[40] P. Agarwal, A. Yadav, and P. Mathur, "OCS-Net: A novel approach for oral carcinoma stage classification using generative adversarial networks and genetic algorithm-based convolutional neural networks," *International Journal of Information Technology*, vol. 15, no. 8, pp. 4097–4110, 2023.