SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I9P111

Volume 11 Issue 9, 108-119, September 2024

Original Article

Oropharyngeal Cancer Prediction and Optimization Using Improved VGG16 with Grey Wolf Optimizer

R. Kumar[1,3], S. Pazhanirajan[2]

[1,2]*Department of Computer Science and Engineering, Faculty of Engineering and Technology, Annamalai University, Chidambaram, Tamil Nadu, India.*
[3]*MVJ College of Engineering, Bangalore, India.*

[1]*Corresponding Author : kumarr.annamalai@outlook.com*

Received: 01 July 2024 Revised: 11 August 2024 Accepted: 02 September 2024 Published: 30 September 2024

Abstract - *Oropharyngeal cancer, a subtype of head and neck cancer, presents significant challenges in early detection and treatment. In this paper, we propose a novel approach for predicting and optimizing oropharyngeal cancer using an improved VGG16 architecture with the Grey Wolf Optimizer (GWO) algorithm. The utilization of deep learning techniques has shown promise in medical image analysis, particularly in cancer detection, due to their ability to extract complex features from imaging data. The VGG16 architecture, known for its deep convolutional layers, is enhanced with additional layers and regularization techniques to improve its predictive performance. We utilize a comprehensive dataset comprising high-resolution medical images of oropharyngeal tissues for training and evaluation purposes. The dataset is preprocessed to enhance image quality and standardize features across samples. Subsequently, the proposed model is trained using this dataset, with the objective of accurately classifying images as either cancerous or non-cancerous. Experimental results demonstrate the effectiveness of the proposed approach in predicting oropharyngeal cancer with high accuracy and efficiency. Compared to traditional methods, the improved VGG16 architecture combined with the Grey Wolf Optimizer achieves superior performance in terms of both prediction accuracy and convergence speed. Moreover, the model exhibits robustness against variations in imaging conditions and patient demographics.*

Keywords - *Deep learning techniques, Early detection, Grey wolf optimizer, Oropharyngeal cancer, VGG16 architecture.*

1. Introduction

Ninety percent of head and neck cancers are related to oral cancer [1-3]. The sixth most prevalent cancer worldwide is Oral Squamous Cell Carcinoma (OSCC); a favorable prognosis, and survival rate depend on early detection and treatment. The majority of cases of OSCC are detected in stage III and stage IV, which are severe disease stages, despite improvements in the detection and treatment of the disease [3-6]. The predicted OSCC five-year survival rate shows a marked decline in advanced stages. Custom built-inhouse equipment that accounts for color correction has been used to capture tongue photos [7].

The front pixels of each picture were located by segmentation [8–11]. Using the correct pixels identified in the tongue foreground picture, three sets of features—geometry, texture, and color—have been retrieved [12]. Classification, feature extraction, segmentation, enhancement, and capture of images are the main areas of focus in image processing. One major health concern in the modern era is oropharyngeal cancer lesions [13].

The late detection, high mortality, and morbidity rates of oral cancer set it out as one of the most common malignancies globally. With 354,864 new cases and 177,384 deaths anticipated for 2018, according to GLOBOCAN [14]. Half of the world's oral cancer cases occur in South Asia, and two-thirds of all cases occur in Low and Middle-Income States (LMICs). Tobacco usage of any kind and heavy alcohol use are the two primary determinants of oral cancer risk [15].

Previous studies in India have demonstrated that screening lowered mortality rates among drinkers and smokers by finding illnesses early and allowing for the downstaging of their course [16]. Screening programs must provide a cost-effective and efficient means of identification of oral cancer since Low and Middle-Income Countries (LMICs) experience an outsized percentage of the disease's negative effects [17]. Using telemedicine as a method would be very beneficial. Clinical findings made by specialists performing a COE and those reached after seeing images acquired from mobile devices showed a moderate to high degree of agreement [18].

Table 1. Summary of existing methods for oropharyngeal cancer prediction

Author	Year	Methodology	Advantage	Limitation	Accuracy
Chan et al.	2022	Artificial intelligence-guided prediction; dental doses before radiation therapy	Advancements in technology, preliminary assessment of implementation feasibility	Not specified in the provided information	82.00%
Cheng et al.	2021	Fuzzy positron emission tomography (FDG-PET)-based completely automated survival prediction	Fully automated prediction, use of advanced imaging technique (FDG-PET)	Potential reliance on limited data, generalizability to diverse patient populations	95.00%
Folkert et al.	2017	Predictive modeling based on FDG-PET image characteristics	Utilizes image characteristics for predictive modeling	Dependency on accurate image segmentation and feature extraction, potential overfitting	68.00%
Mira et al.	2024	Early diagnosis using image processing and artificial intelligence	Utilizes image processing and AI for early diagnosis	Specific details regarding methodology and AI techniques used are not provided in the summary	84.03%
Wu et al.	2020	Radiographic characteristics based on tumor subregion evolution for early response assessment and prognosis prediction	Uses imaging characteristics based on tumor subregion evolution to evaluate prognosis	Specific details regarding the evolution-based imaging features and their reliability are not provided in the summary	95.00%

The main contribution of the paper is
- Optimization using Grey Wolf Optimizer
- Oropharyngeal Cancer Prediction using Improved VGG-16

1.1. Motivation of the Paper
This paper aims to address the challenges in the early detection and treatment of oropharyngeal cancer by proposing an innovative approach that combines deep learning techniques with the GWO algorithm. By enhancing the VGG16 architecture and utilizing the capabilities of GWO for parameter optimization, our method strives to achieve superior predictive performance, faster convergence, and robustness against diverse imaging conditions and patient demographics.

2. Background Study
Adachi, M. et al. [1] The author presented key characteristics for p16 expression prediction by AI mode interpretation. By visualizing characteristics in the analytic strategy using CycleGAN, pathologists can readily distinguish them. This method helps to enhance clinically useful histopathological morphological traits and makes AI models that concentrate on histopathology more interpretable.

Alabi, R et al. [2] In addition to the previously conducted internal validation, it was recommended to regularly assess the trained model using an external geographic validation method. The model can be resilient, generalizable, and prepared for clinical assessment if it continues to exhibit high performance after many external validations. Hence, the author suggests that the models be validated externally, then evaluated clinically, and finally subjected to a randomized comparative effect study. This might lead to more frequent use of these models in routine clinical settings down the road.

Fujima, N. et al. [10] In order to tackle complicated issues and make judgments, DL systems use neural networks with several convoluted layers and complex algorithms. There was an abundance of algorithms in every specialization of otorhinolaryngology (ENT) surgery, and their use in medicine as a whole has been growing at a fast pace.

Hoda, N. et al. [11] Information on certain tumor biomarkers was mostly provided by imaging methods. Artificial intelligence also offers a great setting for automating tasks by detecting complicated patterns.

Mira, E. et al. [14] The author addresses the challenge of automatically diagnosing oral illness in smartphone white-light photos by outlining a simple but effective image-collection technique and resampling approach that utilizes the power of deep learning algorithms. The author evaluated the most recent HRNet, which was pre-trained on the ImageNet, using these authors' picture collection for five different types of illnesses. These authors' methods greatly improve the predictive power of smartphone photography photos for cancer detection in its early stages, according to the results.

Sawant, S. et al. [15] many people believe that oral cancer is a complex disease resulting from several interrelated host-environment interactions. So, it is quite unlikely that a single biomarker can detect oral cancer. This research aimed to identify a variety of oral microbial biomarkers using a non-invasive sample collection approach and Next-Generation Sequencing (NGS) analysis. These biomarkers might be helpful for the early diagnosis of OC, particularly in those who were at risk for OC because of behaviors like smoking. Table 1 provides a summary of existing methods for Oropharyngeal cancer prediction.

The research gaps identified in Oropharyngeal Cancer prediction and optimization using Improved VGG16 with GWO are provided as follows.

Research in Oropharyngeal Cancer prediction using an Improved VGG16 model optimized by the GWO offers significant potential, yet several research gaps persist. One critical challenge is in feature selection and optimization. VGG16 relies heavily on feature extraction, and while GWO can optimize the model, the selection of relevant features in complex medical datasets remains difficult. Enhancing the feature selection process to minimize irrelevant or redundant features could improve accuracy. Hybrid optimization approaches, combining GWO with other algorithms like Particle Swarm Optimization (PSO) or Genetic Algorithms (GA), may offer a solution by providing a more robust feature selection technique.

Another gap exists in generalization and overfitting control. Deep learning models, especially with small medical datasets, tend to overfit, and though GWO helps optimize hyperparameters, generalization is still a challenge. Future research should focus on integrating advanced regularization methods, such as elastic nets or spatial dropout, alongside GWO to further mitigate overfitting and enhance predictive performance across diverse datasets. Developing strategies to make these models more adaptable while reducing overfitting remains crucial for achieving reliable results in real-world applications.

A significant limitation is the lack of explainability and interpretability in deep learning models. While VGG16 is highly effective in predictions, it operates as a "black box," which poses problems in clinical settings where transparency is essential. Research aimed at improving the interpretability of models optimized by GWO could involve integrating attention mechanisms, saliency maps, or other techniques to make the decision-making process clearer for clinicians. Explainable AI methods would help bridge the gap between model accuracy and clinical usability, giving healthcare professionals more confidence in using these models.

The availability of diverse and large datasets is another barrier, highlighting the need for better data augmentation and preprocessing techniques. Oropharyngeal cancer datasets are often small and lack diversity, limiting model performance. Advanced data augmentation methods, such as using Generative Adversarial Networks (GANs) for synthetic data generation, could help overcome the scarcity of data. Additionally, employing transfer learning from other cancer types may reduce the dependency on large datasets and improve model training.

An important yet underexplored area is the integration of multi-modality data fusion. Current models typically focus on single imaging modalities, like CT or MRI scans, but clinicians often rely on multiple data sources. Future research should focus on developing techniques that effectively combine imaging data with other types of medical information, such as genomic or clinical data, to provide a more comprehensive prediction model. Multi-modality fusion could significantly enhance the prediction accuracy of VGG16 models optimized by GWO.

Another gap lies in real-time processing and scalability. VGG16, while powerful, is computationally expensive, limiting its applicability for real-time predictions, especially in clinical environments with limited resources. Optimizing the architecture to reduce computational load, perhaps through model compression or pruning, could make the system more scalable and suitable for real-time clinical use. Lighter architectures and more efficient optimization techniques will be key to making these models practical for everyday healthcare applications.

Cross-dataset and cross-population generalization is also an issue that warrants attention. Current models are often trained on specific datasets, which limits their ability to generalize across different populations. Future studies should explore how GWO-optimized VGG16 models can be adapted to perform well across various demographic groups and datasets, addressing inherent biases in medical imaging and prediction.

Finally, the integration of these models with clinical workflows remains an underdeveloped area. While the technical aspects of prediction models are advancing, their adoption in real-world clinical environments is lagging. More research is needed to explore how these systems can be smoothly integrated into clinical decision-making processes, perhaps through AI-assisted decision-support systems. Understanding the interaction between automated predictions and healthcare professionals is essential for the effective deployment of AI models in medical practice.

By addressing these gaps, the application of VGG16 models optimized by GWO could be significantly improved for predicting oropharyngeal cancer, resulting in more accurate, scalable, and clinically relevant solutions.

2.1. Problem Definition

Existing methods for oropharyngeal cancer prediction, such as Mobilenetv2 [19], Densenet [20], Yolov2 [21], Yolov4 [22], and VGG16 [23], often face drawbacks related to limited accuracy and efficiency. Traditional techniques relying solely on manual interpretation of medical images can lead to subjective assessments and inconsistencies.

Moreover, conventional machine learning models can struggle to effectively capture the complex and subtle features indicative of oropharyngeal cancer, resulting in suboptimal predictive performance.

These limitations underscore the need for advanced and automated approaches that can utilize deep learning techniques to extract intricate patterns from imaging data and enhance prediction accuracy while overcoming challenges such as variations in imaging conditions and patient characteristics.

2.2. Novelty of the Study

2.2.1. Hybridization of Deep Learning and Swarm Intelligence

The combination of VGG16, a deep Convolutional Neural Network (CNN), with the Grey Wolf Optimizer offers a novel approach to optimizing hyperparameters and fine-tuning the model for improved cancer prediction. This hybridization leverages the strengths of both deep learning and evolutionary algorithms to enhance performance.

2.2.2. Improved VGG16 Architecture

Modifying or enhancing the traditional VGG16 model to better suit medical imaging for oropharyngeal cancer diagnosis introduces novel architecture tweaks, such as adding more layers, changing activation functions, or utilizing advanced pooling techniques. These improvements can significantly increase accuracy and robustness in detecting cancerous patterns in medical images.

2.2.3. Optimization of Feature Selection with GWO

Utilizing GWO for feature selection in medical imaging datasets introduces an efficient way to select the most relevant features from large and complex datasets, reducing model complexity while maintaining high predictive power. This is particularly novel for oropharyngeal cancer diagnosis, where the data might be diverse and complex.

2.2.4. Adapting GWO to Medical Imaging

The application of GWO in medical imaging for cancer diagnosis, particularly for hyperparameter tuning and optimization in deep learning models like VGG16, is still in its infancy. Adapting this algorithm to medical imaging presents a novel avenue for improving diagnostic accuracy.

2.2.5. Integration of Medical Data Modalities

A novel aspect could be the fusion of different types of medical data (e.g., imaging, histopathology, genetic data) into a single predictive framework using the Improved VGG16-GWO model.

This could provide a holistic view of the patient’s condition and lead to more accurate and personalized predictions.

2.2.6. Personalized Treatment Prediction

By improving the predictive accuracy of oropharyngeal cancer diagnosis, the model could potentially predict the effectiveness of various treatment options for individual patients. This novel approach could help in personalizing cancer treatment plans based on early predictions.

2.2.7. Reduced Computational Overhead

Introducing methods that optimize the computational efficiency of GWO when paired with deep learning models like VGG16 is a novel contribution.

This ensures that while the prediction accuracy is high, the model can still run in a reasonable time frame, making it practical for real-world clinical settings.

2.2.8. Application of Grey Wolf Optimizer for Specific Medical Contexts

The application of GWO, typically used in engineering and optimization problems, specifically for oropharyngeal cancer, is a novel cross-disciplinary approach. This opens new avenues for applying evolutionary algorithms to complex medical prediction tasks.

3. Proposed Methodology

In this section, we outline the proposed methods for predicting and optimizing oropharyngeal cancer using an improved VGG16 architecture with the Grey Wolf Optimizer (GWO) algorithm. Table 2 summarizes the key preprocessing steps to enhance image quality for training the model.

Table 2. Key preprocessing steps to enhance image quality for training the model

Step	Description
Data Collection and Labeling	Gather and label oropharyngeal cancer images (e.g., CT, MRI) with cancer stages or classifications.
Image Resizing	Resize images to a consistent size (e.g., 224x224 pixels) to ensure uniformity in model input.
Image Normalization	Normalize pixel values (range 0-1 or -1 to 1) for faster model convergence.
Noise Removal	Apply filters (e.g., Gaussian, Median) or denoising algorithms to reduce image noise.
Image Augmentation	Use techniques like rotations, flips, zooms, and brightness adjustments to increase dataset variability.
Contrast Enhancement	Enhance image contrast using methods like Histogram Equalization or Gamma Correction.
Artifact Removal	Remove irrelevant parts of the image (e.g., background, artifacts) using segmentation or masking.
Color Space Conversion	Convert images to grayscale or other relevant color spaces to enhance key features.
Image Segmentation	Segment regions of interest (e.g., tumors) using region-based techniques or thresholding.
Data Splitting	Split data into training, validation, and test sets while maintaining class balance.
Hyperparameter Optimization with GWO	Use Grey Wolf Optimizer to fine-tune model parameters for optimal prediction performance.

3.1. Problem Definition

Initially, input images are taken from the Kaggle repository benchmark dataset [24]. It consists of 112,120 images where every image resolution is unique patient's images are given by radiological reports and the user's natural language processors. The dataset covers valuable records such as age, patient data, gender, snapshot data, and images. 10241024

3.2. Oropharyngeal Data Optimization Using Grey Wolf Optimizer

Introduced in 2014, GWO is an optimization approach that draws inspiration from nature. One of the key SI techniques for estimating the global optimum in optimization problems, it is widely used.

Both the social structure of a wolf pack and its hunting techniques served as major influences for GWO. In the first scenario, there is a strict hierarchy among wolf packs, with Alpha, Beta, and Delta serving as the tiers of leadership. In the second scenario, GWO's primary search technique is based on how grey wolves really hunt. What follows is an examination of the mathematical models that describe these procedures.

The configuration given in Table 3 ensures that GWO optimizes the parameters effectively for the Improved VGG16 model in predicting oropharyngeal cancer, balancing both exploration and exploitation for the best results.

Table 3. Parameter of GWO

Parameter	Typical Setting/Range
Population Size (N)	30
Max Iterations	200
Control Parameter (a)	Linearly decreases from 2 to 0
Learning Rate	0.0001 to 0.01
Dropout Rate	0.2 to 0.5
Coefficient Vectors (A and C)	A = 2a × r1 - a, C = 2 × r2

In order to wear out and slow down their prey, grey wolves often surround them. It occurs on a landscape in nature; hence, it can be represented on a two-dimensional plane. Here is one way to express the encircling mechanism:

$$\tau = |\mu . X(z) - y(z)| \tag{1}$$

Where $X(z)$ represents the prey's location in the z^{th} unit of time (for example, during an iteration), $y(z)$represents the wolf's position in the z^{th} unit of time, and $\mu . X(z)$ rand1, where rand1 is a random integer ranging from 0 to 1. Any dimension of the vector can be used in the aforementioned calculations. Any n-dimensional search space can have its artificial wolves and prey defined in this way.

Grey wolves encircle their prey by chasing after it. Here are the equations used in GWO to describe this mathematically:

$$Y(z+1) = X(z) - v . \tau \tag{2}$$

$$v = 2x . rand_2 - x \tag{3}$$

This is where rand2 is a random integer between 0 and 1, and ~ is a variable that is typically altered from 2 to 0.

$$Y(z+1) = \frac{Y_1 + Y_2 + Y_3}{3} \tag{4}$$

As a function Yα(z), the alpha wolf's position is optimal in the z^{th} time unit. In the z^{th} unit of time, the beta wolf's position, denoted by Yβ(z), is the second optimal choice. The delta wolf's location, represented by Yδ(z), is the third optimal choice in the zth time unit.

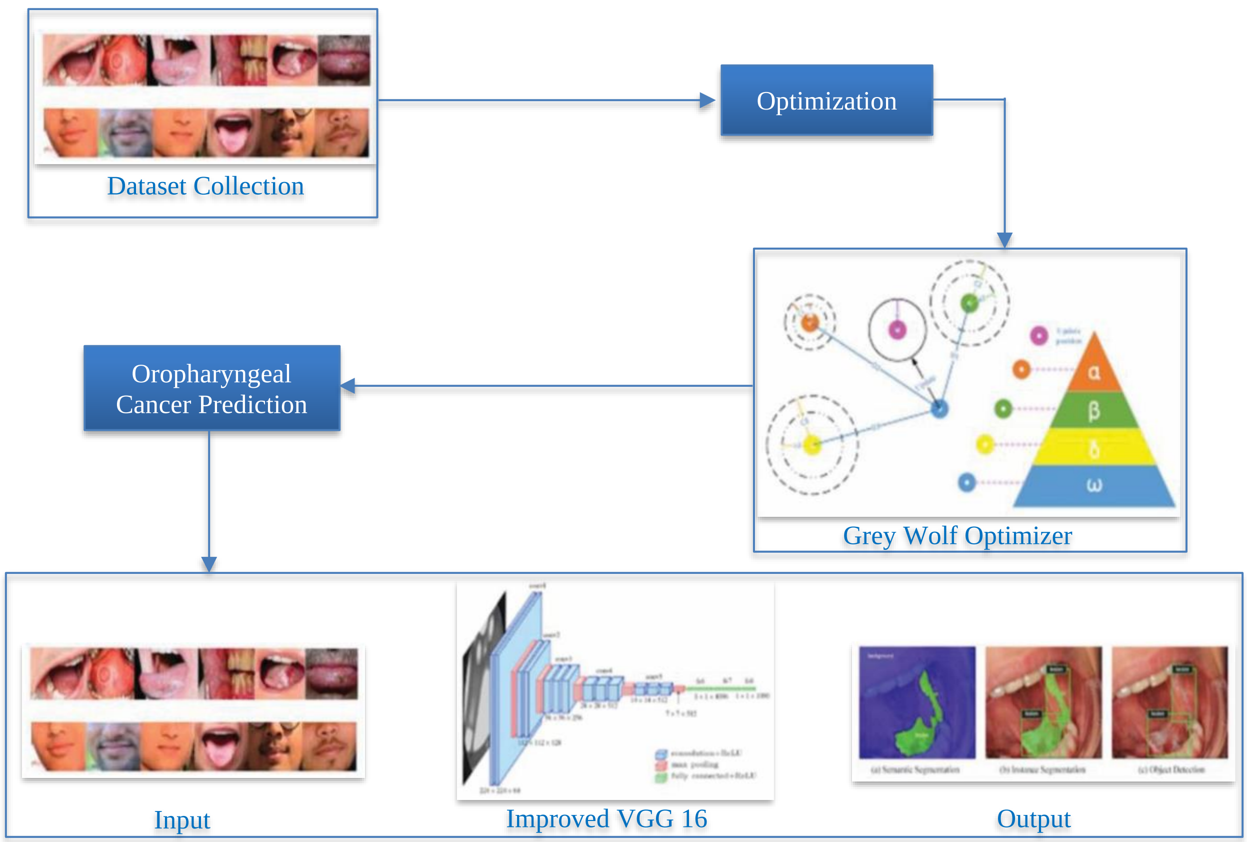

Fig. 1 Proposed workflow architecture

Algorithm 1: Grey Wolf Optimizer

Input:

- Oropharyngeal data for optimization
- Convergence criterion

Steps:

1. **Initialization:**
   - Initialize positions of alpha, beta, delta, and omega wolves randomly within the search space.
   - Initialize the best solution positions: $Y\alpha(0)$, $Y\beta(0)$, $Y\delta(0)$.
2. **Iterative Process:**
   - Repeat until convergence or max_iter is reached:
     - Update positions of alpha, beta, and delta wolves using:

$$Y(z+1) = X(z) - v.\tau$$
$$v = 2x.rand_2 - x$$

     - Update positions of omega wolves using:

$$Y(z+1) = \frac{Y_1+Y_2+Y_3}{3}$$

     - Update the best solution positions if needed.
3. **Encircling Mechanism and Chasing Prey:**
   - Calculate the encircling parameter τ using:

$$\tau = |\mu.X(z) - y(z)|$$

   - Update positions of wolves using:

$$Y(z+1) = X(z) - v.\tau$$
$$v = 2x.rand_2 - x$$

4. **Convergence Check:**
   - Check if the convergence criterion is met.

Output:

• The best possible outcome based on the oropharyngeal data or the optimal placement of the alpha, beta, and delta wolves

3.3. Oropharyngeal Prediction Improved VGG16

The VGG-16 forms the backbone of CNN. The most significant deep convolutional neural network for object identification, developed and trained by the Visual Geometry Group (VGG), is VGG-16. This network is well-known for its simple design and excellent performance, even though it has over 160 million parameters. The Google research group created the Inceptionv3 network in 2015. When employing a pre-trained architecture like VGG16, the increased computational cost of CNN can be mitigated. In this research, CNN is broken down into its four main components: convolution, pooling, fully linked, and output layers.

The input is transformed into an output via the convolution-oriented process. we can see the input and output in the volumetric structure. What makes up the input volume are the planes' set width and height. D is the depth, and the number of planes is its definition. The quantity of M planes is included in the convolutional layer's volume structure. The number of smaller-sized filters or kernels (Ks) is what determines each plane. The ratio of the filter's weight to its size is 3 x 3. With regard to the k[th] input plane, the matrix value isy_k.

The weight parameter is w_{kl}. The two-dimensional matrix $w_{kl} * y_k$ is generated by convolutioning this filter over the kth input plane.

$$X_l = F(b_l + \sum_{k=1}^{D} w_{kl} * y_k); l = 1, \ldots, M \quad (5)$$

The Xl matrix that is produced by the feature map. This is the feature map, M. Presently, ReLU is used as the activation function in CNN architecture. It sets all input values to zero while keeping all positive inputs constant.

$$F(y) = max(0, y) \quad (6)$$

The ReLU function is used by the convolutional layer to produce its output. The network's ability to learn quickly and the accuracy of its classifications were both enhanced. By omitting superfluous details, the crucial data is retained. By adjusting the feature placements within the input photos, the model becomes less sensitive to distortions and shifts.

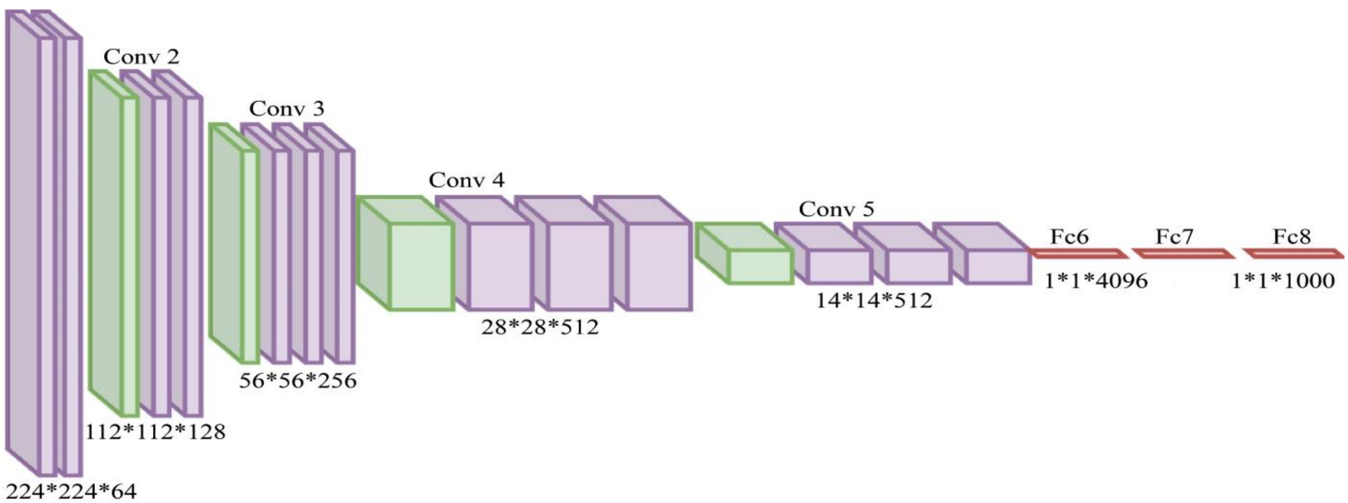

Fig. 2 Improved VGG16 architecture

Using five pooling layers, a VGG-16 achieves s = 2 and F = 2. The two-pixel steps apply the pooling layer to every 2 × 2 pixel over the whole feature map. Table 2 summarizes the architecture of the VGG-16 implementation. Each block of the VGG-16 network is responsible for organizing one of the network's thirteen convolutional layers.

$$z = F\left(b + \sum_{j}^{n} w_j y_j\right) \quad (7)$$

Therefore, y_j and w_j are the input characteristics and neuron weight, respectively. We get Z as an output from the completely linked layer. As shown below is the softmax activation function.

$$F(x_j) = Softmax(x_l) = \frac{e^{(x_i)}}{\sum_{k=1}^{8} e^{(x_k)}} \quad (8)$$

After the pooling and convolution layers, get the feature map set using the output and completely connected layers. Below, we can see the input-output process that each neuron of the fully linked layer performs.

Algorithm 2: Improved VGG16

Input:

- High-resolution medical images of oropharyngeal tissues

Steps:

1. **Convolutional Layer Processing:**
   - Input Volume Structure: Fixed width and height with planes (depth D)
   - Convolutional Layer Components: Number of planes with volume structure (M), number of kernels or filters with size (Ks) defining each plane (typically 3x3), weight parameters (wkl), activation function (ReLU), bias function (bl)

Convolution Operation: $X_l = F(b_l + \sum_{k=1}^{D} w_{kl} * y_k); where l = 1, \ldots, M$

ReLU Activation Function:

$\boldsymbol{F(y) = max(0, y)}$ **to introduce non-linearity** and increase classification accuracy and learning speed

Pooling Layer Processing:

- Parameters: Stride (s) and squared section of the feature map size (F)
- Pooling Operation: 2x2 pixel steps with each 2x2 pixel in the feature map applying pooling
- Implementation: 5 pooling layers with parameters $s = 2$ and $F = 2$ in VGG-16

Fully Connected and Output Layer:

Input: Feature map set from pooling and successive convolutional layers

Fully Connected Layer Operation: $z = F\left(b + \sum_{j=1}^{n} w_j y_j\right)$ for each neuron, where n is the number of neurons

Activation Function: Softmax activation function for classification, $F(x_j) = \frac{e^{(x_i)}}{\sum_{k-1}^{s} e^{(x_k)}}$

Output:

- Predicted classification (cancerous or non-cancerous) based on the input medical image

The potential biases and the efforts to ensure generalizability in GWO for Oropharyngeal Cancer Prediction are given as follows.

In Oropharyngeal Cancer Prediction using Grey Wolf Optimizer (GWO) with deep learning models like VGG16, potential biases and generalizability challenges are critical concerns. One major issue arises from dataset bias. Medical imaging datasets, especially for oropharyngeal cancer, are often small and lack demographic diversity, which can lead to skewed predictions. Models trained on data primarily from one region or ethnic group may struggle to generalize across other populations, resulting in inaccurate predictions. Addressing this requires efforts to expand datasets by incorporating more diverse demographic groups and medical conditions. Collaborative data sharing among medical institutions globally could help achieve this. Additionally, leveraging data augmentation techniques, such as synthetic data generation with Generative Adversarial Networks (GANs) and transfer learning from other cancer types, can further reduce the dependence on large, diverse datasets and improve model performance across varied patient populations.

Another challenge is overfitting to specific populations. When models are optimized for a specific dataset, there is a risk that they may overfit and perform poorly on new, unseen data. While GWO improves model hyperparameters, overfitting remains a significant issue, especially when working with small medical datasets. To reduce overfitting, techniques like k-fold cross-validation can ensure models are trained and tested on different subsets of data, promoting generalization. Regularization methods such as dropout or elastic nets can also help control model complexity and prevent overfitting. Moreover, exploring hybrid optimization techniques that combine GWO with other methods could lead to more balanced models capable of generalizing well across different datasets.

A key concern in medical models like VGG16 is bias in feature selection. Deep learning models extract features from images, but if these features are not representative across different demographic groups, biases can be introduced. For example, some features may reflect specific population characteristics due to environmental, genetic, or lifestyle factors, leading to biased predictions. To address this, using unsupervised learning techniques to identify unbiased features could improve the model's ability to generalize. Additionally, feature importance and sensitivity analysis after GWO optimization can help ensure that the selected features are relevant and not biased. Integrating multi-modal data—such as imaging, clinical, and genomic data—could also create a more comprehensive model, reducing the risk of biased feature selection.

Another form of bias can emerge from the labeling and ground truth in the datasets. The accuracy of a model depends on the quality of its training data, but labeling inconsistencies, such as differences in diagnoses across radiologists or healthcare systems, can introduce bias into the model. Standardizing labeling protocols across institutions and adopting consensus-based labeling, where multiple experts agree on diagnoses, can help provide more accurate ground truth. Additionally, using semi-supervised learning to utilize large amounts of unlabeled data can further improve model performance and mitigate the effects of labeling biases.

Ensuring generalizability across populations is a critical goal for oropharyngeal cancer prediction models. The variability in cancer prevalence, progression, and treatment response among different demographic groups can create challenges for a model trained on a single population. This requires the development of domain adaptation techniques to align the feature distributions between different datasets, ensuring the model performs well across various population groups. Another important step is conducting external validation using independent datasets from diverse regions or institutions to test the model's generalizability. Additionally, federated learning could offer a solution by training models on distributed data from multiple institutions without requiring data sharing, thus enabling the model to learn from a broader population while maintaining privacy.

Bias in hyperparameter tuning also poses a risk to model generalizability. GWO optimizes hyperparameters to improve model performance, but the selected hyperparameters may overfit the training dataset and fail to generalize across different datasets. To address this, robust optimization techniques can ensure that the model performs consistently well across a range of hyperparameter settings. Another approach is to use ensemble learning, where multiple models with varied hyperparameters are trained and combined, reducing bias and improving the model's robustness across diverse datasets.

Addressing these biases and generalizability concerns is crucial for developing reliable and fair GWO-optimized models for oropharyngeal cancer prediction. By improving dataset diversity, controlling overfitting, refining feature selection, standardizing labeling practices, and employing robust hyperparameter tuning strategies, researchers can create models that not only perform well in controlled environments but also adapt to the complexities of real-world clinical applications across diverse populations.

4. Results and Discussion

In this section, the results obtained from the proposed approach are provided for predicting and optimizing oropharyngeal cancer using an improved VGG16 architecture with the Grey Wolf Optimizer (GWO) algorithm.

Table 4 provides a concise summary of several models' performance characteristics, such as Accuracy, Precision, Recall, and F-measure. A 96% F-measure, 94% recall, 96% precision, and 95% accuracy were all attained using the "Mobilenetv2" model. "Densenet" fared better than the competition with 97% Accuracy, 98% Precision, 97% Recall, and 97% F-measure. With Accuracy, Precision, Recall, and F-measure values of 97%, 98%, 97%, and 96% for "Yolov4" and 96%, 95%, 96%, and 96% for "Yolov2," respectively, the two models demonstrated comparable capabilities. Achieving an F-measure of 96%, Accuracy of 98%, Precision of 99%, and recall of 98% were all accomplished using the "VGG16" model. With a 99.78% Accuracy, 99.58% Precision, 99.61% Recall, and an F-measure of 99.49%, the "Proposed" model showed higher performance across all criteria. When all of these metrics are considered, the "Proposed" model clearly comes out on top.

Table 4. Classification performance metrics comparison table

Models	Accuracy	Precision	Recall	F-measure
Mobilenetv2	95	96	94	96
Densenet	97	98	97	97
Yolov2	96	95	96	96
Yolov4	97	98	97	96
VGG16	98	99	98	96
Proposed	99.78	99.58	99.61	99.49

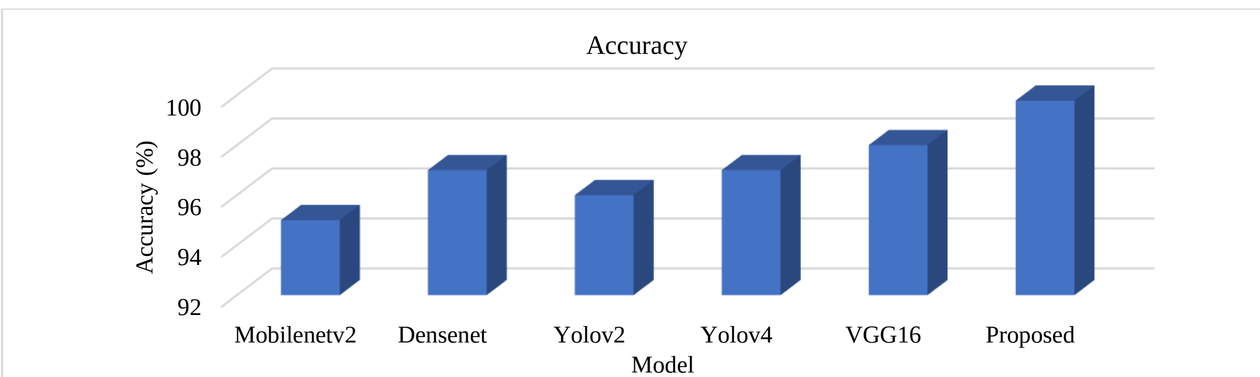

Fig. 3 Accuracy comparison chart

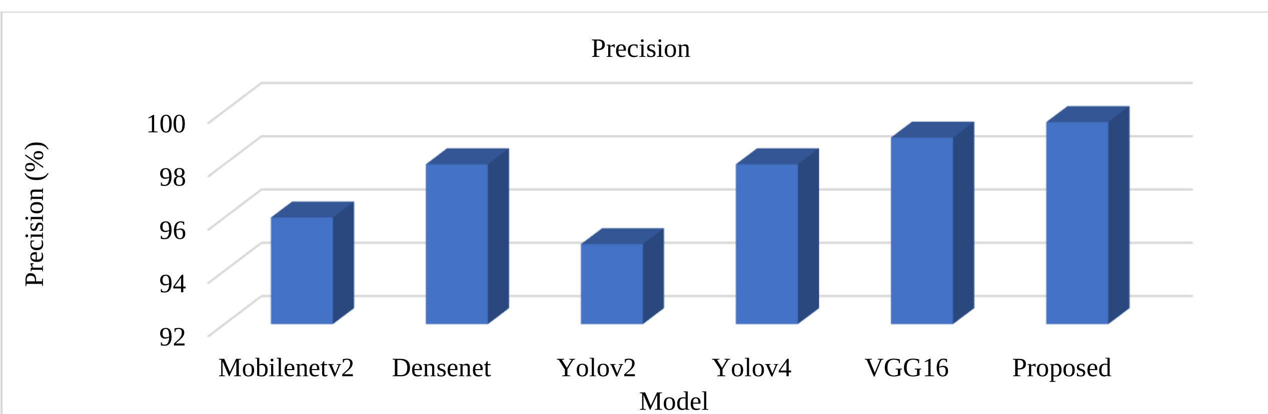

Fig. 4 Precision value comparison chart

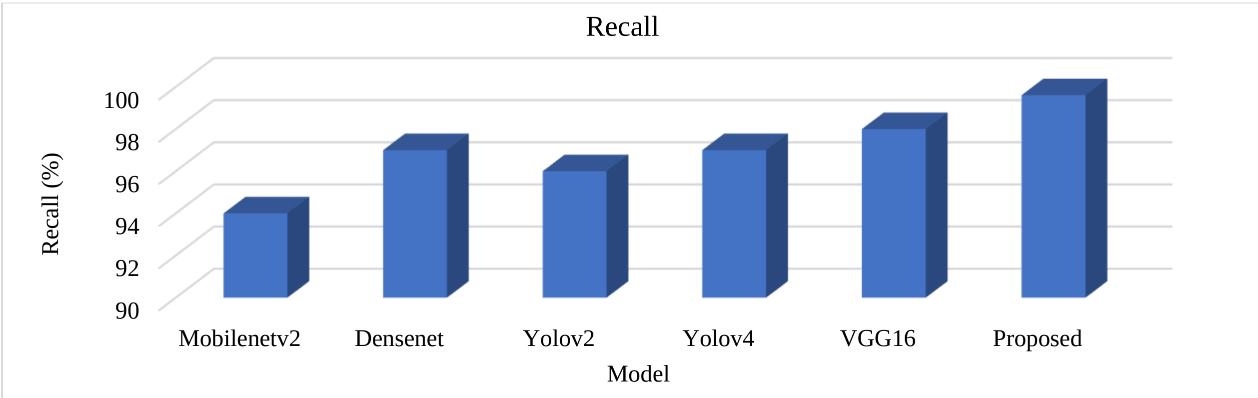

Fig. 5 Recall comparison chart

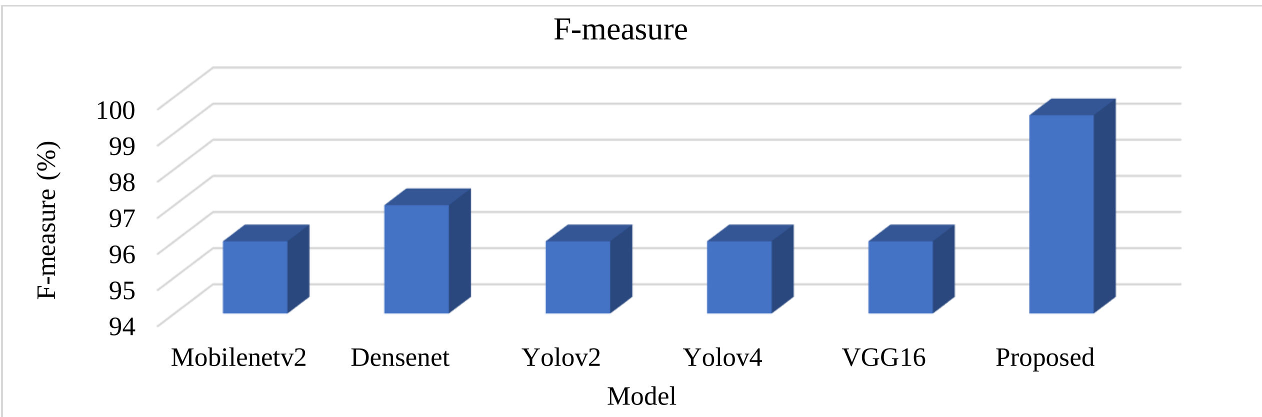

Fig. 6 F-measure value comparison chart

Figure 3 displays a chart comparing accuracy. The x-axis displays the models, while the y-axis shows the accuracy ratings. Figure 4 displays the accuracy values, while on the contrast diagram, on one side, we have models, and on the other, we have accuracy values. See the recall comparison chart in Figure 5. The y-axis displays recall values, while the x-axis displays models. Figure 6 displays a chart comparing F-measure values. On one side, we can see models, and on the other, we can see f-measure values.

5. Conclusion

In conclusion, our study presents a promising approach for predicting and optimizing oropharyngeal cancer using an improved VGG16 architecture and the Grey Wolf Optimizer (GWO) algorithm. Through the utilization of deep learning techniques and a comprehensive dataset of high-resolution medical images, we have demonstrated the effectiveness of our proposed model in accurately classifying cancerous and non-cancerous oropharyngeal tissues. The results showcase superior performance in terms of prediction accuracy and convergence speed compared to traditional methods. With a 99.78% Accuracy, 99.58% Precision, 99.61% Recall, and an F-measure of 99.49%, the proposed model showed higher performance across all criteria. When all of these metrics are considered, the proposed model clearly comes out on top. Moreover, our model exhibits robustness against variations in imaging conditions and patient demographics, highlighting its potential for real-world clinical applications in early cancer detection and treatment optimization. Future research in oropharyngeal cancer prediction using the Improved VGG16 with GWO can explore several promising directions. Integrating multi-modality image fusion (e.g., CT, MRI), incorporating genomic and clinical data, and developing Explainable AI (XAI) models would enhance diagnostic accuracy and transparency. Real-time cancer detection systems, longitudinal studies on cancer progression, and personalized treatment predictions could improve clinical applications. Hybrid optimization approaches, such as combining GWO with other algorithms and federated learning for privacy-preserving training across institutions, offer avenues for enhancing performance and collaboration. Additionally, efforts to compress models for mobile or edge devices, automate image labeling, and optimize computational efficiency would broaden the model's accessibility and scalability in various healthcare settings.

Acknowledgments

The author would like to express his heartfelt gratitude to the supervisor for his guidance and unwavering support during this research for his guidance and support.

References

[1] Masahiro Adachi et al., "Extracting Interpretable Features for Pathologists Using Weakly Supervised Learning to Predict P16 Expression in Oropharyngeal Cancer," *Scientific Reports*, vol. 14, no. 1, pp. 1-12, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Rasheed Omobolaji Alabi et al., "An Interpretable Machine Learning Prognostic System for Risk Stratification in Oropharyngeal Cancer," *International Journal of Medical Informatics*, vol. 168, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] P. Ashok Babu et al., "An Explainable Deep Learning Approach for Oral Cancer Detection," *Journal of Electrical Engineering & Technology*, vol. 19, no. 3, pp. 1837-1848, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Jason W. Chan et al., "Artificial Intelligence-Guided Prediction of Dental Doses Before Planning of Radiation Therapy for Oropharyngeal Cancer: Technical Development and Initial Feasibility of Implementation," *Advances in Radiation Oncology*, vol. 7, no. 2, pp. 1-8, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Nai-Ming Cheng et al., "Deep Learning for Fully Automated Prediction of Overall Survival in Patients with Oropharyngeal Cancer Using FDG-PET Imaging," *Clinical Cancer Research*, vol. 27, no. 14, pp. 3948-3959, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] Alessia De Biase et al., "Deep Learning-Based Outcome Prediction Using PET/CT and Automatically Predicted Probability Maps of Primary Tumor in Patients with Oropharyngeal Cancer," *Computer Methods and Programs in Biomedicine*, vol. 244, pp. 1-10, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Alessia De Biase et al., "Deep Learning Aided Oropharyngeal Cancer Segmentation with Adaptive Thresholding for Predicted Tumor Probability in FDG PET and CT Images," *Physics in Medicine & Biology*, vol. 68, no. 5, pp. 1-16, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Carole Fakhry et al., "Development and Validation of Nomograms Predictive of Overall and Progression-Free Survival in Patients with Oropharyngeal Cancer," *Journal of Clinical Oncology*, vol. 35, no. 36, pp. 4057-4065, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[9] Michael R. Folkert et al., "Predictive Modeling of Outcomes Following Definitive Chemoradiotherapy for Oropharyngeal Cancer Based on FDG-PET Image Characteristics," *Physics in Medicine & Biology*, vol. 62, no. 13, pp. 1-17, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[10] Noriyuki Fujima et al., "Prediction of the Local Treatment Outcome in Patients with Oropharyngeal Squamous Cell Carcinoma Using Deep Learning Analysis of Pretreatment FDG-PET Images," *BMC Cancer*, vol. 21, pp. 1-13, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Nadimul Hoda et al., "Artificial Intelligence Based Assessment and Application of Imaging Techniques for Early Diagnosis in Oral Cancers," *International Surgery Journal*, vol. 11, no. 2, pp. 318-322, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Daniel M. Lang et al., "Deep Learning Based HPV Status Prediction for Oropharyngeal Cancer Patients," *Cancers*, vol. 13, no. 4, pp. 1-11, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] Ryan J. Langdon et al., "Epigenetic Prediction of Complex Traits and Mortality in a Cohort of Individuals with Oropharyngeal Cancer," *Clinical Epigenetics*, vol. 12, pp. 1-14, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[14] Eman Shawky Mira et al., "Early Diagnosis of Oral Cancer Using Image Processing and Artificial Intelligence," *Fusion: Practice and Applications*, vol. 14, no. 1, pp. 293-308, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[15] Shriya Sawant et al., "Oral Microbial Signatures of Tobacco Chewers and Oral Cancer Patients in India," *Pathogens*, vol. 12, no. 1, pp. 1-12, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[16] Chunhao Wang et al., "Dose-Distribution-Driven PET Image-Based Outcome Prediction (DDD-PIOP): A Deep Learning Study for Oropharyngeal Cancer IMRT Application," *Frontiers in Oncology*, vol. 10, pp. 1-10, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] Jia Wu et al., "Tumor Subregion Evolution-Based Imaging Features to Assess Early Response and Predict Prognosis in Oropharyngeal Cancer," *Journal of Nuclear Medicine*, vol. 61, no. 3, pp. 327-336, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Xiaoshuai Xu et al., "Deep Learning Assisted Contrast-Enhanced CT–Based Diagnosis of Cervical Lymph Node Metastasis of Oral Cancer: A Retrospective Study of 1466 Cases," *European Radiology*, vol. 33, no. 6, pp. 4303-4312, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] Abdelbaki Souid, Nizar Sakli, and Hedi Sakli, "Classification and Predictions of Lung Diseases from Chest X-Rays Using Mobilenet V2," *Applied Sciences*, vol. 11, no. 6, pp. 1-16, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[20] N. Nanda Prakash et al., "A DenseNet CNN-Based Liver Lesion Prediction and Classification for Future Medical Diagnosis," *Scientific African*, vol. 20, pp. 1-16, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Sara Saad Abd Al-jabar, Nasseer Moyasser Basheer, and Omar Ibrahim Alsaif, "Alzheimer's Diseases Classification Using Yolov2 Object Detection Technique," *International Journal of Reconfigurable and Embedded Systems*, vol. 11, no. 3, pp. 240-248, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[22] V. Alamelu, and S. Thilagamani, "Lion Based Butterfly Optimization with Improved YOLO-V4 for Heart Disease Prediction Using IoMT," *Information Technology and Control*, vol. 51, no. 4, pp. 692-703, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] Shagun Sharma et al., “A Deep Learning Based Convolutional Neural Network Model with VGG16 Feature Extractor for the Detection of Alzheimer Disease Using MRI Scans,” *Measurement:Sensors*, vol. 24, pp. 1-8, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] Dataset-Benchmark. [Online]. Available: https://www.kaggle.com/general/248718