

AN EFFECTIVE APPROACH FOR HISTOPATHOLOGIC ORAL CANCER PREDICTION USING BITTERLING FISH OPTIMIZATION TECHNIQUE

PULLAIAH PINNIKA¹, VENKATA RAO KASUKURTHI²

^{1,2} Department of Computer Science and Systems Engineering, Andhra University, Visakhapatnam, Andhra Pradesh, India

E-mail: ¹pullaiah.531@gmail.com

ABSTRACT

Oral cancer poses a substantial global health concern, necessitating immediate action to mitigate its severe consequences. However, existing methods used for oral cancer classification exhibit limitations such as getting trapped in the local optima, high sensitivity to noisy data, and poor convergence speed, resulting in lower prediction reliability and clinical applicability. Hence, this research proposes a Bitterling Fish Optimization with Probability Entropy (BFO-PE) algorithm for feature selection during oral cancer classification. The PE enables the BFO to select optimal feature subsets by simulating the spawning behavior of bitterling fish, updating iteratively toward an optimal feature set, thereby enhancing performance. Initially, the Oral cancer (Tips and Tongue) Image (OCI) dataset and Kvasir Video and Image Repository (KVASIR) dataset are used to validate the model's performance. The next step, pre-processing involves resizing images to standardized dimensions using interpolation. Then, feature extraction is performed using the ResNet 152 architecture, followed by which feature selection is carried out. Finally, the ResNet layer is deployed for classification of oral cancer into binary classes. The proposed approach attains a superior accuracy of 99.04% on OCI and 99.75% Kvasir datasets, outperforming existing methods like Convolutional Neural Networks - Improved Tunicate Swarm Algorithm (CNN-ITSA). These results demonstrate the effectiveness, generalizability, and potential clinical applicability of the developed system for oral cancer prediction.

Keywords: *Bitterling Fish Optimization; Convolutional Neural Networks; Improved Tunicate Swarm Algorithm; Oral Cancer And Probability Entropy.*

1. INTRODUCTION

Oral cancer is a malignancy that develops in areas such as the inner lip, tongue, floor of the mouth, palate and other unspecified parts of the oral cavity [1]. The primary causes of oral cancer include heavy alcohol consumption and smoking. Survival rates remain low, particularly in low and middle-income countries, where two-thirds of individuals present with oral lesions at an advanced stage [2]. In complex cases, such as large tumors involving the mandible or those located at the base of the tongue, surgical resection with sufficient free surgical margins is often required [3]. Early diagnosis is crucial for effective management, and recent trends involve leveraging Deep Learning (DL) techniques for Oral Cancer Image (OCI) classification using lightweight methods [4]. DL methods have shown success in classifying oral cancer lesions using data collected from the OCI and KVASIR datasets, which are employed to identify tumor regions. These

datasets support medical image classification through a process-enhanced DL approach to improve overall performance [5]. Both the avoidance of risk factors and the reinforcement of protective factors are crucial components of cancer prevention. The initial stage involves in identifying tumor regions in both cancerous and non-cancerous images, which are then used to extract relevant feature patterns [6] [7].

Oral cancer is characterized by the presence of cancerous tissue within a specific organ system, with subcellular changes likely occurring within the same system [8]. Relevant features are selected to support the classification process and to identify contiguous tissues of the aerodigestive tract while excluding oral cancer [9]. Additionally, the burden of oral cancer disproportionately affects image quality. Therefore, selecting relevant features and refining parameters is essential for identifying key characteristics that differentiate benign from cancerous lesions. After feature selection, the processed input data should support accurate

classification and improve the overall performance of the model [10]. The concept implies that the presence of cancerous tissue in one part of an organ system may indicate subcellular changes occurring elsewhere within the same system [8]. Relevant features are selected to enable efficient classification and identify contiguous tissues of the aerodigestive tract that are not affected by oral cancer [9]. The complexity of oral cancer often leads to variability in image quality. Therefore, selecting relevant features and fine-tuning model parameters helps recognize key characteristics that differentiate distinguish between benign and cancerous lesions. After feature selection, the input data should provide sufficient classification capability while optimized performance [10]. However, accurate diagnosis remains challenging even for experienced pathologists, and neural networks assist in making the process more manageable and improving diagnostic accuracy [11]. Histopathology slides containing cells and tissue are digitized using a whole-slide digital scanner and stored as histopathological images [12]. Classification using a pre-trained network on large-scale data improves the overall process, enabling efficient identification of cancerous and non-cancerous cases from the OCI dataset [13].

Relevant features are selected to represent the classification process and identify contiguous tissues of the aero digestive tract, excluding oral cancer [9]. Additionally, the burden of oral cancer disproportionately affects image quality. Therefore, selecting relevant features and fine-tuning parameters help recognize key characteristics that differentiate benign and cancerous lesions. After feature selection, the input data should provide sufficient classification capability while optimizing overall performance [10]. However, accurate diagnosis is challenging even for experienced theologists, and neural networks are an asset to making the process more manageable and achieving better accuracy [11]. The histopathology slides containing cells and tissue digitized by a whole slide digital scanner and stored as histopathological images [12]. Classification using a pre-trained network on sizable data improves the overall process to efficiently identify cancer and non-cancer from the OCI [13].

This research study addresses key limitations associated with using a single method for classification and feature extraction [14]. Existing classification models, including conventional CNN-based frameworks, face challenges such as slow convergence, overfitting, high-dimensional feature

spaces, inconsistent image quality, and inefficient feature selection, all of which impair classification performance and limit diagnostic accuracy and robustness. This study proposes the integration of Bitterling Fish Optimization with Probability Entropy (BFO-PE) and ResNet-152. In this framework, BFO-PE fine-tunes the feature space to provide optimized input for ResNet-152, while the residual blocks in ResNet-152 maximize the utility of the selected features. The probability entropy mechanism guides the spawning decision in BFO by scoring and selecting feature subsets through the simulated spawning behavior of bitterling fish, and iteratively updates toward optimal features. This hypothesis lays the foundation for the development of an intelligent, automated screening system capable of enhancing diagnostic confidence and reducing clinical workload.

The key contributions of this research study are considered as follows:

- The BFO-PE approach is proposed for the selection of informative and relevant features through simulating adaptive optimization guided by entropy, improving model robustness.
- The ResNet-152 architecture is employed for the extraction of relevant features. It effectively models the complex patterns and fine-grained details in histopathological images using its deep residual architecture, resulting in optimal performance.
- The Light ResNet layer is deployed for the categorization of oral cancer into binary classes of cancerous and non-cancerous cases using a refined feature subset, contributing to higher classification accuracy.

The research paper is further organized as follows: Section 2 provides an account of related, existing models designed for oral cancer detection using DL approaches. Section 3 details the proposed methodology and its functioning. Section 4 presents the results and discussion of the implemented method. Finally, Section 5 concludes the research.

2. RELATED WORK

The section summarizes existing research models introduced for oral cancer classification, along with their advantages and limitations.

Wei et al. [15] presented Convolutional Neural Network and improved Tunicate Swarm Algorithm (CNN-ITSA) for medical image classification as cancerous and non-cancerous. The CNN extracted spatial feature hierarchies for identifying subtle patterns in the image, while ITSA

efficiently explored and exploited the search space using optimal hyperparameters to ensure better performance. However, CNNs combined with ITSA required significant computational resources when processing large data and faced overfitting issues.

Huang et al. [16] suggested a CNN model with Improved Squirrel Search Algorithm (CNN-ISSA) for efficiently processing high-level features and the selection of relevant features from medical images. The ISSA enhanced features by emphasizing on the most relevant, high-level features within the CNN, thereby reducing irrelevant data combinations and improving the CNN's ability to discriminate between classes. However, while the CNN-ISSA model attained high classification accuracy, it did not resize images into a uniform size, resulting in challenges in modelling the texture and edges in oral images.

Song et al. [17] developed a pipeline to provide an effective oral cancer detection system using images for feature selection. The Support Vector Machine (SVM) was employed for classification. Both phases were optimized using a modified version of a cost-effective search optimizer. The introduced approach demonstrated significant improvements in accuracy and overall performance. However, this approach exhibited issues of overfitting, which degraded the classification performance.

Hossain et al. [18] suggested polyp's segmentation and classification based on deep learning for autonomous colonoscopy examination. The developed Double U-NET and Vision Transformer models were employed for the segmentation and classification of polyps, while considering their risks. An advantage of the developed vision transformer-based classifier model was that it classified the polyps by efficiently differentiating between benign and malignant regions by a careful analysis of the features. However, the double-UNet-based segmentation model faced challenges with segmenting the boundaries of polyps.

Demirbas et al. [19] explored a Spatial Attention ConvMixer (SAC) based classification model to detect and classify gastrointestinal diseases. The explored SAC model was used to detect and classify gastrointestinal diseases according to the class labels. Its primary advantage was the use of a patch extraction technique, which extracted significant features by leveraging spatial attention to focus on more informative regions. However, the model struggled to classify gastrointestinal diseases accurately due to high

similarity between classes, which limited its overall classification effectiveness.

Liu et al. [20] designed a semi-supervised medical image classification based on Star-Adversarial Generative Neural Network for diagnosis. This method performed image-to-image translation across multiple classes using a single model, reducing the need to train separate models for each class, thereby enhancing efficiency. However, the Star-GAN model faced complexity when handling multiple images simultaneously, while balancing training across different domains required tuning, and dealing with high-resolution images posed various challenges.

Despite notable advancements in deep learning and medical image analysis, the accurate classification of oral cancer remains a persistent challenge in real-world clinical environments. One of the primary reasons is the high intra-class similarity and inter-class variability present in histopathological images, where cancerous and non-cancerous tissues often exhibit overlapping visual features. Existing methods, such as CNN-ITSA and CNN-ISSA, are prone to overfitting, computational inefficiency, and convergence at local minima due to suboptimal feature subset selection and the curse of dimensionality in medical data. Furthermore, these methods often suffer from issues such as local optimal solutions and sensitivity to noisy data, resulting in slow convergence during the classification process. This research proposes BFO-PE to address the aforementioned challenges by adaptively selecting the most informative features through entropy-guided, bio-inspired optimization. This approach reduces dimensionality and enhances classifier efficiency by using entropy-based evaluation to retain only the most discriminative features. While ResNet-152 effectively extracts deep, high-level features from images, these features can sometimes be redundant or noisy. BFO-PE refines them by selecting only the most relevant features based on entropy scores, thereby improving the classifier's focus on critical patterns. This integration improves classification accuracy, enhances robustness, and ensures better generalization across diverse datasets such as OCI and Kvasir.

3. PROPOSED METHODOLOGY

This research proposes a BFO-PE approach for feature selection to enhance oral cancer classification performance. This model's functions in multiple stages: pre-processing using resizing to standardize images with uniform dimensions using interpolation. Then, feature extraction is performed

using ResNet152, while feature selection is employed using BFO-PE. Finally, classification is carried out through a light ResNet layer for the categorization of oral cancer into binary classes. Figure 1 demonstrates the pipeline of the implemented method.

3.1 Dataset

This section analyzes oral cancer prediction using the Oral Cancer (Lips and Tongue) Image (OCI) and KVASIR datasets. These two datasets on cancer prediction in different organs are standard, benchmark datasets, containing annotated images relevant to cancer detection.

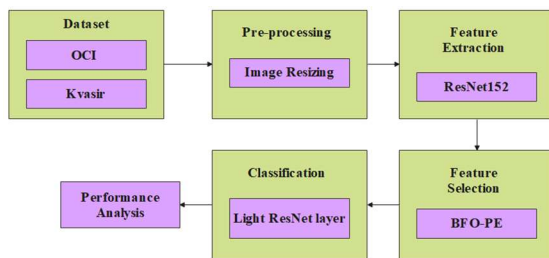


Figure 1: Pipeline of Implemented Method

3.1.1 OCI dataset

The OCI dataset contains images of affected lips and tongue for oral cancer prediction analysis [21]. It is split into three categories: 87 series of oral cancer images and 44 series of non-cancerous lesions, including normal oral images. For this research, the dataset is divided into binary classes, with 70% of the data used for training and 30% for testing.

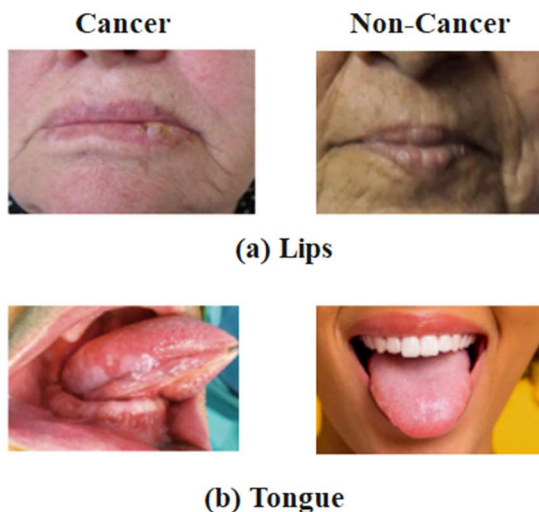


Figure 2: Sample Image of OCI Dataset (a) Lips (b) Tongue

3.1.2 KVASIR dataset

The KVASIR dataset [22] is a large collection of high-quality gastrointestinal (GI) endoscopy images, developed to support research in computer-aided diagnosis and automated disease detection within the GI tract. This dataset comprises over 8,000 annotated images, covering various types of lesions and diseases.

3.2 Pre-processing

Image resizing is performed to ensure uniformity and consistency across images, which are generally of different sizes and resolutions. Resizing procedure reduces executional complexity by limiting input size to 224 × 224, rendering them effective, while maintaining the image integrity by preserving its edges and textures.

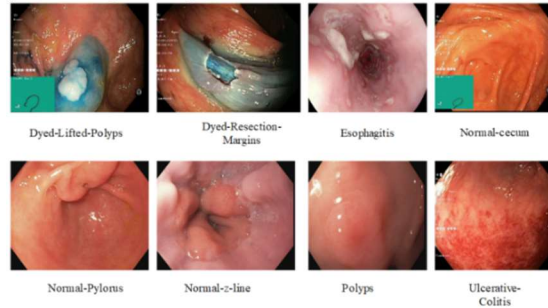


Figure 3: Sample Image of Kvasir Dataset

3.3 Feature Extraction

After pre-processing, feature extraction is performed using pre-trained ResNet152 model, known for its deep residual architecture and robust learning capability. The features are extracted from average pooling layer, which outputs a 2048-dimensional vector, representing high-level image characteristics. This layer integrates spatial data and minimizes feature map dimensions effectively. The extracted features model large patterns crucial for distinguishing cancerous and non-cancerous images. The training is performed over 1000 iteration steps using a consistent training model, after which the outcomes are evaluated. Based on the accuracy and validation error, ResNet-152 demonstrates superior performance compared to other ResNet variants, outperforming in both accuracy and loss metrics. Despite its large model size, higher number of parameters, and decreased computational speed, ResNet is selected as the base model in this research due to its ability to improve classification accuracy. The resulting feature vectors are then passed to the BFO-PE algorithm for optimal feature selection, before classification.

3.4 Feature Selection

Feature selection is performed using the BFO-PE algorithm [23], which refines the 2048-

dimensional feature vectors extracted from ResNet152. This approach simulates the spawning behavior of bitterling fish, directed through probability entropy to determine the most informative features. Thus, the high-dimensional input is minimized to a subset of 100 to 300 selected features, based on the dataset characteristics. This dimensionality reduction helps eliminate redundant data and enhances classification effectiveness. The BFO approach simulates the behavioral dynamics of bitterling fish during the search process. Each initial solution is represented as a bitterling fish, as defined by Equations (1) to (3). This involves generating multiple bitterling populations, initialized as a randomly distributed population.

$$F_i = [F_i^1, F_i^2, F_i^3, \dots, F_i^D] \quad (1)$$

$$F = \begin{bmatrix} F_1^1 & F_1^2 & \dots & F_1^D \\ F_2^1 & F_2^2 & \dots & F_2^D \\ \vdots & \vdots & \vdots & \vdots \\ F_n^1 & F_n^2 & \dots & F_n^D \end{bmatrix} \quad (2)$$

$$F_i^j = l + (u - l) \cdot r \quad (3)$$

Where, the variable D denotes the number of dimensions of every value i , and F denotes the matrix of the main populations of size n . The F_i^j denotes the dimension of i th values in the population in a range $[0,1]$.

3.4.1 Evaluate solution

Each solution in BFO is represented as a fish and is evaluated based on its ability to search and select an optimal oyster. Each fish's merit is determined by the suitability of the oyster it selects, as well as its interactions with other female fish. The objective function, denoted as f , is used to evaluate each fish solution and is defined in Equation (4). The fitness function primarily consists of two components: the total number of selected features and the classification accuracy. It evaluates classification error, where lower error values correspond to better fitness scores, as expressed in Equation (5).

$$Fitness = \begin{bmatrix} f(F_1^1, F_1^2, \dots, F_1^D) \\ f(F_2^1, F_2^2, \dots, F_2^D) \\ \vdots \\ f(F_n^1, F_n^2, \dots, F_n^D) \end{bmatrix} \quad (4)$$

$$Fitness\ Function = w \times \alpha (1 - w) \times \frac{|s|}{|d|} \quad (5)$$

Where, $|d|$ indicates the total number of features in the dataset, $|s|$ denotes the number of features in the selected subset, α is a classification error associated with the utilized feature subset, and $w \in [0,1]$ signifies relative weight values assigned to classification error and the number of features.

3.4.2 The search for the seize oysters

The search strategy faces challenges in effectively navigating the search space and avoiding the selection of unsuitable mating oysters. Each fish evaluates its current location to identify locations containing more optimal oysters. When a fish identifies a suitable oyster, it targets and moves toward it. However, if the oyster has already been claimed by another fish, it is disregarded. This process is represented in Equation (6).

$$F_i^{t+1} = \begin{cases} J \cdot F_i^t + (F^+ - J \cdot F_i^t) \cdot \delta & r \leq P \\ J \cdot F_i^t + (F^* - J \cdot F_i^t) \cdot \delta & r > P \end{cases} \quad (6)$$

Where, F_i^t and F_i^{t+1} denote the position of the i th fish at iterations t and $t + 1$, respectively. F^* indicates the best oyster value discovered so far, while F^+ denotes an arbitrarily selected oyster of good quality in the population, δ and r are arbitrarily selected and lie between the range $[0,1]$, and J is the number of cases in which fish moves toward the eliminated oyster, contributing to reduced iterations in this algorithm. The parameter J denotes a gradual decrease performance of male fish after successful mating. This decrease limits local search and progressively enhances the algorithm's ability to converge toward the global optimum as formulated in equation (7).

$$J(t) = J(t) - \frac{J(1) \cdot t}{Maxt} \cdot U(t) \quad (7)$$

Here, $J(t)$ denotes the phase and jump of each fish as it searches for oysters in iteration t and the arbitrary function is represented by U , as given in equation (8), which generated an arbitrary sequence with $U(1)$ being set to one, as described in Equations (8) and (9).

$$U(t + 1) = \cos(t \times \cos^{-1}(U(t))) \quad (8)$$

$$J(t + 1) = J(1) - \frac{J(1) \cdot t}{Maxt} \cdot \cos(t \times \cos^{-1}(U(t))) \quad (9)$$

The HFO seeks optimal values over time, focusing on exploring the search space and refining the solution through repeated iteration around optimal points. This is achieved by ensuring the condition of $r > P$, which facilitates a more efficient processing of the second criterion. The function is similar to arctangent behaviour in equation (8), where t denotes the repetition counter. Each feature undergoes the process using P values of 0.1, 0.5 and 0.9. These two types of search decreases as the difference between these values becomes smaller.

3.4.3 Escape and not seize the oyster

The escape strategy in oyster behavior involves situations where an oyster is already guarded by another fish, prompting the fish to avoid impending it, and instead, select a different position.

Equation (10) describes the escape behavior for fish that have not succeeded in capturing oysters. In this context, M represents the average position of the fish swarm during escape. A bitterling fish searches within the region defined between this average position and a randomly selected point in the optimal search space. Equation (11) is used to evaluate the population's gravity point.

$$F_i^{t+1} = \begin{cases} J \cdot F_i^t + (F^* - J \cdot M) \cdot \delta r \leq 0.5 \\ l + (u - l) \cdot \delta r > 0.5 \end{cases} \quad (10)$$

$$M = \frac{\sum_{i=1}^n F_i^t}{n} \quad (11)$$

After locating the oyster, the male fish engages with a female fish and lays eggs inside the oyster, following which fertilization occurs, resulting in the generation of new fish, as described in Equation (12).

$$F_i^{t+1} = F_i^t + R * rand(0,1) \quad (12)$$

Where, R denotes the radius of distribution of fish around the shell with the initial value ranging between 0 and 2. The parameter solution is initialized as described in equation (12), which decreases over time, based on the iteration process. The probability of losing values is considered inversely proportional to its advantage. The minimization of the probability of losing a fish equation is given by Equation (13). The probability of eliminating a solution F_i^t is denoted by $d(F_i^t)$ and $f(F_i^t)$ indicates the function value of the solution.

$$d(F_i^t) = \frac{f(F_i^t)}{\sum_{i=1}^n f(F_i^t)} \quad (13)$$

3.4.4 Probability entropy

After applying the BFO algorithm with adaptive PE, it is essential to choose an optimal model component from the decomposed component. The PE metric estimates the uncertainty and data richness of extracted features, commonly utilizes in time series but here adapted to allows the variation and relevance of features obtained from medical images. This enables the BPO approach to prioritize features which displays an effective discriminatory behavior, even in noisy or overlapping classes. Unlike basic fitness functions, PE improves the BPO's capability to detect subtle variations, designing it well appropriates for histopathologic image data. This combination directly improves a precision of feature selection and benefits the classifier in attaining greater performance. The PE method is used to evaluate uncertainty of one-dimensional time series and when compared to other strategies like fractal dimension, it demonstrates a strong anti-interference effect and captures small changes in the sequence.

The time series is represented as $\{x(t), t=1,3,\dots,N\}$ through length of N phase space redeveloped is taken Equation (14) to (17) pursues.

$$\begin{cases} X(1) = \{x(1), x(1+\lambda), \dots, x(1+m-1)\lambda\} \\ X(i) = \{x(i), x(i+\lambda), \dots, x(i+(m-1)\lambda)\} \\ X(N-(m-1)\lambda) = \{x(N-(m-1)\lambda), x(N-(m-2)\lambda), \dots, x(N)\} \end{cases} \quad (14)$$

$$S(g) = \{j_1, j_2, \dots, j_m\} \quad (15)$$

$$H_p(m) = -\sum_{i=1}^k p_i \ln p_i \quad (16)$$

$$H_p = \frac{H_p(m, \tau)}{IN(m!)} \quad (17)$$

Where, embedding dimension is denoted by m and delay time denoted λ , reorganize all elements in $X(t)$ are rearranged in ascending order $x(t + (j_1 - 1)\lambda) \leq x(t + (j_2 - 1)\lambda) \leq \dots \leq x(t + (j_m - 1)\lambda)$. The sequence $X(t)$ is then mapped to a symbol sequence $S(g)$ in m dimensional phase space mapping. The P_t probability rate of i the symbol sequence and entropy of time series $X(t)$. The $0 < H_p \leq 1$ PE utilized an uncertainty as well as complexity of the time series. An entropy of harmonic and modulated signals was significantly low, whereas an entropy of permutation is arbitrarily handled by noisy PE.

Table 1: Hyperparameters of Classifier.

Hyper parameter	Values
Epochs	30 -50
Batch size	30-35
Dropout layer	0.1-0.5
Num Agents	50
Max iterations	1000
Search Space_Dim	10
Step Size Max	0.5
Step Size Min	0.01
Loss	Sparse category entropy
Optimizer	Adam
Activation	SoftMax

3.5 Classification

The selected relevant features are fed into a classifier to distinguish between cancerous and non-cancerous features. After feature selection using BFO-PE, the reduced feature set is fed into a fully connected Dense Neural Network classifier for final classification. ResNet152 is widely used for extracting high-level image features and is not reused during classification. This separation ensures that the model operates effectively on vectorized data, as traditional CNNs are not designed to process non-image inputs. ResNet152, employed for feature extraction, is a residual network that addresses the vanishing gradient problem through skip connections.

In this network, instead of learning the underlying mapping directly, the layers learn the residual mapping. The main goal of ResNet is to use skip connections, which provide inputs to deeper layers directly, facilitating a seamless flow of information. This approach helps eliminate data loss and addresses issues such as vanishing gradients. The primary objective of ResNet is to use skip connections that directly feed inputs to deeper layers, enabling seamless information flow. This approach helps mitigate data loss and addresses the vanishing gradient problem. ResNet-152 is trained over 30 epochs with a batch size of 32 and a dropout rate of 0.4 to prevent overfitting. Validation loss is monitored throughout the training process. Table 1 summarizes the classifier’s hyperparameters.

The ResNet architecture is highly effective in facilitating the training of very deep neural networks and improving accuracy. However, increasing the depth of a CNN often leads to degradation issues. The core concept of ResNet is to pass the input directly to subsequent layers through identity mappings. In ResNet, each residual block

(RB) adds the input to the block’s output. The residual function is expressed in Equation (18).

$$y = F(x, W) + x \tag{18}$$

Where, x describes the input of the residual block, W indicates weight of RB, and y denotes the output of residual block. ResNet network is made up of multiple RBs, each with varying convolutional kernel sizes. To identify the shortcuts, (x) is applied directly to both the input and output, ensuring that they stay within uniform dimensions, as defined by Equation (19) and (20).

$$y = F(x, W_i) + x \tag{19}$$

$$y = F(x, W_i) + W_s x \tag{20}$$

When there is a change in dimension, a shortcut mapping is applied by adding zero-padding to match the increased dimension. The projection shortcut is then used to align dimensions through a specified function. This process does not introduce additional parameters, but the projection is represented by W_s . Each layer in the network forms a subspace within the overall network.

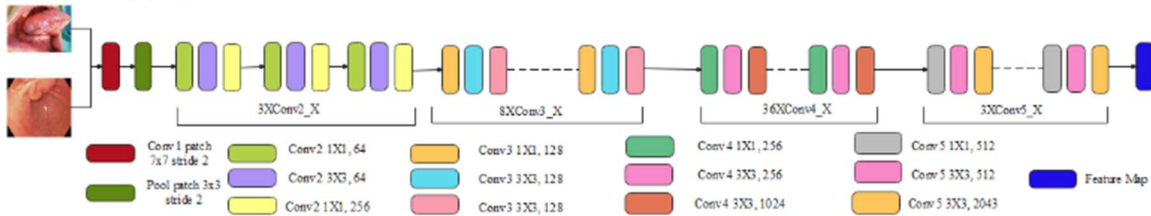


Figure 4: Basic Architecture of ResNet 152

This approach remains effective in performance, with each block consisting of two or three layers, depending on the model’s configuration. The extracted features are then passed to the feature selection stage to identify the most relevant ones. To overcome this complexity, ResNet architecture is introduced, which simplifies the stacked convolutional layers by learning residual mappings, rather than primary mappings in a traditional stacked architecture. The network consists of eight residual blocks, each containing two convolutional layers. Three types of convolutional kernels with varying sizes are used in each Conv layer for adaptation in various scales. The number of kernels for each size follows the pattern 4k, where k initiates at 1 and increases every two convolutional layers. Figure 4 illustrates the architecture of the ResNet-152 model.

4. EXPERIMENTAL ANALYSIS

A comparative experimental framework is employed wherein the BFO-PE with ResNet-152 model is evaluated against benchmark methods like

CNN-ITSA, CNN-ISSA, and SVM using two distinct datasets, OCI and Kvasir, representing diverse imaging scenarios (oral cancer and gastrointestinal lesions). The design includes preprocessing, feature extraction, optimization, and classification stages validated through standard performance metrics. This multi-dataset and multi-domain design ensures model robustness, promotes cross-domain applicability, and enhances confidence in the model’s scalability for other medical image classification tasks. The performance of the ResNet-152 approach is simulated in a Python environment on system configuration being 8 GB RAM, an i5 processor, and Windows 10 (64-bit). The metrics considered to evaluate the proposed ResNet-152 method are accuracy, recall, precision, and F1-score, and their mathematical formulas are expressed in Equations (21) to (24).

$$Accuracy = \frac{TN+T}{TP+TN+FP+F} \times 100 \tag{21}$$

$$Recall = \frac{TP}{TP+F} \times 100 \tag{22}$$

$$Precision = \frac{TP}{TP+F} \times 100 \tag{23}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (24)$$

where, FP shows false positive, TP denotes true positive, FN denotes false negative, and TN shows true negative.

4.1 Performance Analysis

The performance of ResNet-152 based feature extraction algorithm is evaluated based on the metrics of accuracy, recall, f1-score and precision, as represented in Table 2. The existing feature extraction algorithms considered to evaluate the ResNet-152 are MobileNet, VGG-19 and Inception V3. The ResNet-152-based feature extraction algorithm obtains 98.90% accuracy, 98.53% precision, 99.14% recall and 98.82% f1-score on the Oral cancer dataset.

The performance analysis of feature selection using BFO-PE demonstrates its efficiency in selecting relevant features and reducing high dimensionality. The BFO-PE is compared with existing methods namely, WAO, PSO and DMO, as represented in Table 3.

The analysis of classification performance using ResNet-152 reveals that it efficiently captures the complex data and connections, enabling the network to bypass certain layers and preserving information from earlier layers when preventing gradient degradation. Table 4 represents the Resnet-152 method compared with existing methods like CNN, DNN and EfficientNet.

Table 2: Performance Analysis of Feature Extraction on Dataset.

Classifiers	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1_Score (%)
Mobile Net	OCI	96.14	95.32	95.36	95.81
	Kvasir	95.45	95.82	95.3	95.52
VGG16	OCI	97.85	96.05	96.8	96.42
	Kvasir	96.89	97.3	96.7	96.93
Efficient Net	OCI	98.45	97.65	97.65	97.14
	Kvasir	97.62	97.95	97.41	97.65
ResNet152	OCI	99.04	98.68	99.25	98.96
	Kvasir	99.75	99.75	99.75	99.76

Table 3: Performance Analysis of Feature Selection on Dataset.

Method	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1_Score (%)
WAO	OCI	96.75	95.89	95.55	95.21
	Kvasir	96.15	96.45	96.01	96.2
PSO	OCI	97.65	96.92	96.6	96.25
	Kvasir	96.88	97.14	96.71	96.91
DMO	OCI	98.25	97.48	97.45	97.91
	Kvasir	97.59	97.89	97.4	97.62
BFO-PE	OCI	99.04	98.68	99.25	98.96
	Kvasir	99.75	99.75	99.75	99.76

Table 4: Performance Analysis of Classification on Dataset.

Classifiers	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1_Score (%)
CNN	OCI	97.5	95.6	96.72	95.15
	Kvasir	95.9	96.15	95.87	96.01
DNN	OCI	98.12	96.35	97.55	96.9
	Kvasir	96.88	97.12	96.72	96.91
Efficient Net	OCI	98.76	97.9	98.88	97.32
	Kvasir	97.95	98.21	97.85	98.03
ResNet152	OCI	99.04	98.68	99.25	98.96
	Kvasir	99.75	99.75	99.75	99.76

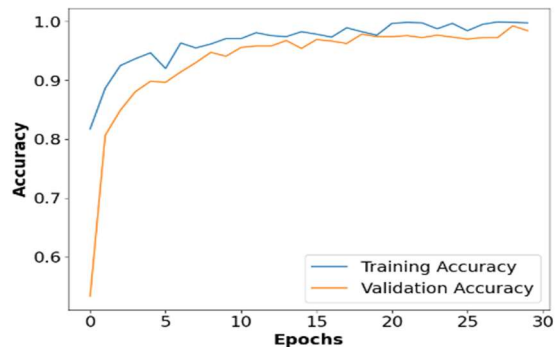


Figure 5: Epoch v/s Accuracy on OCI Dataset

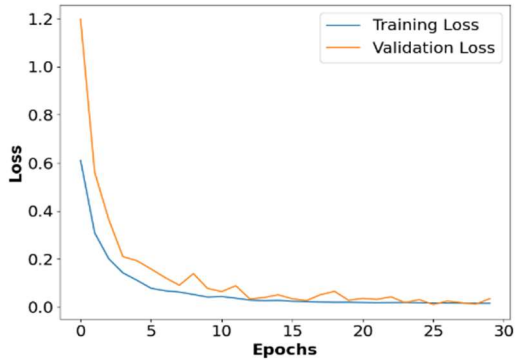


Figure 6: Epoch v/s Loss on OCI Dataset

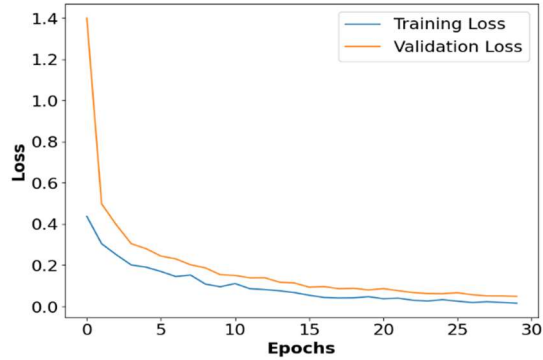


Figure 10: Epoch v/s Loss on Kvasir Dataset

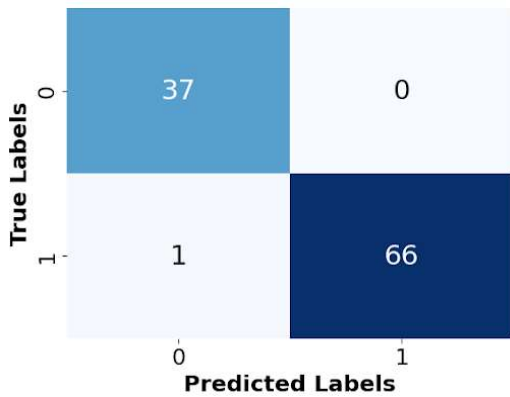


Figure 7: Confusion Matrix on OCI Dataset

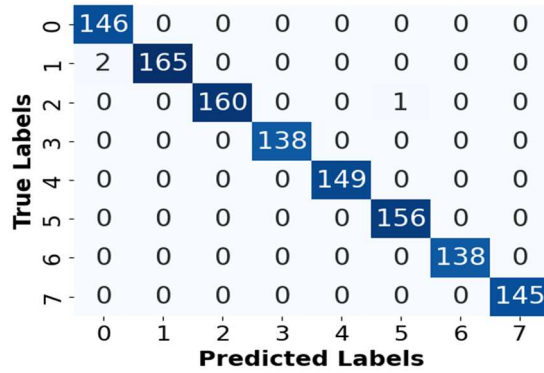


Figure 11: Confusion Matrix on Kvasir Dataset

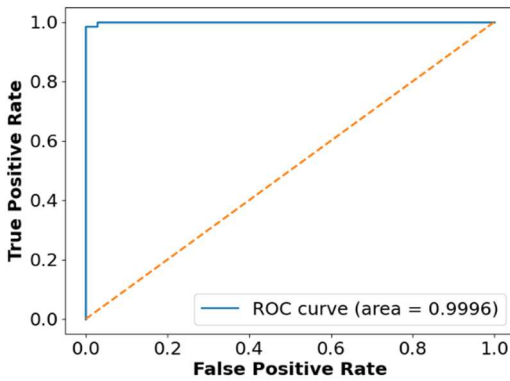


Figure 8: ROC Curve on OCI Dataset

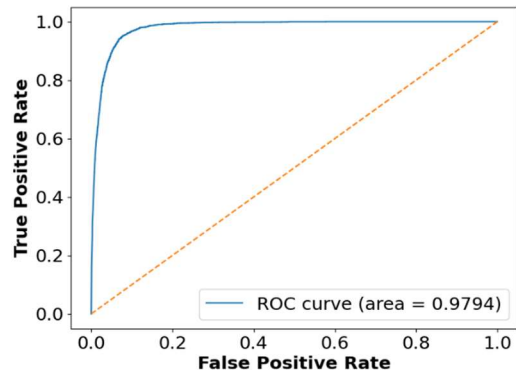


Figure 12: ROC Curve on Kvasir Dataset

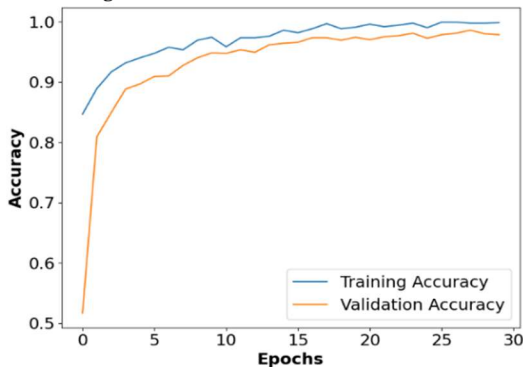


Figure 9: Epoch v/s Accuracy on Kvasir Dataset

Figures 5 to 12 show the accuracy vs epochs and loss vs epochs graphs. The validation loss remains lesser than the training loss, remaining consistent with no increase after a certain number of epochs, demonstrating that the proposed approach is not overfitting. Additionally, validation loss reaches its peak while training loss continues to decrease, indicating superior performance with both training and validation losses minimized and stabilized in low values. This represents effective generalization on unseen data.

Table 5: Comparative Analysis of Proposed Method on Datasets.

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1_Score (%)
<i>OCI dataset</i>				
CNN-ITSA [16]	98.70	96.42	93.71	90.08
CNN-ISSA [17]	97.33	92.66	NA	89.37
SVM [18]	94.65	93.89	82.44	86.27
Proposed BFO-PE with ResNet-152 Method	99.04	98.68	99.25	98.96
<i>Kvasir dataset</i>				
ViT [19]	99	NA	NA	NA
SAC [20]	93.37	93.66	93.37	93.42
SL-CL-NM [21]	93.13	NA	NA	NA
Proposed BFO-PE with ResNet-152 method	99.75	99.75	99.75	99.76

4.2 Comparative Analysis

In this section, the BFO-PE and ResNet-152 models are compared with existing methods like CLAHE-GLCM-ICNN [16], CNN-ISSA [17], SVM [18], ViT [19], SAC [20] and SL-CL-NM [21], as shown in table 5. The proposed method attains an optimal accuracy of 99.04% on the OCI dataset and 99.75% on the Kvasir dataset. It effectively captures complex data and selects relevant features, while the deep layers of ResNet-152 extract detailed patterns from the data, resulting in improved accuracy.

4.3 Discussion

The performance of BFO-PE with ResNet-152 method is evaluated using the Oral cancer dataset comparing it with various algorithms. Moreover, the developed technique is compared to existing algorithms like CLAHE-GLCM-ICNN [16], CNN-ISSA [17], SVM [18], ViT [19], SAC [20] and SL-CL-NM [21] on the OCI and Kvasir datasets. Existing algorithms have certain limitations. For instance, CNNs combined with ITSA require significant computational resources due to large data volumes and are prone to overfitting. CNN-ISSA attained high classification accuracy; however, it does not resize images to a uniform size, which makes it difficult to capture textures and edges effectively in oral images. To address these limitations, image processing is

standardized during the preprocessing phase. In the resizing stage, all images are resized before being used for feature extraction. High-level features are then extracted using the ResNet-152-based feature extraction approach, which helps differentiate cancerous and non-cancerous regions effectively, thereby improving oral cancer detection and classification performance. Compared to existing approaches like CNN-ITSA and CNN-ISSA, which rely solely on deep learning or basic heuristic optimization, the proposed BFO-PE with ResNet-152 incorporates an entropy-driven, bio-inspired feature selection mechanism that reduces dimensionality and improves feature relevance. Existing models such as SVM and CNN-ISSA suffer from overfitting and inefficiency due to the inclusion of irrelevant features. In contrast, BFO-PE selects features based on their information richness, enhancing generalization. Its integration with ResNet-152 effectively manages high-dimensional data from deep layers, ensuring strong robustness and high classification accuracy on both OCI and Kvasir datasets.

The experimental outcomes validate the initial objectives by confirming that BFO-PE significantly enhances the relevance of selected features, while ResNet-152 ensures comprehensive representation learning. The approach effectively minimizes feature redundancy and improves classifier performance, outperforming existing methods across all performance metrics. This demonstrates the successful achievement of the research goals related to accuracy improvement, noise resistance, and feature dimensionality reduction. The proposed BFO-PE with ResNet-152 framework achieves superior classification performance when benchmarked against existing models such as CNN-ITSA and CNN-ISSA, which are affected by resource demands, overfitting, and limited handling of high-dimensional features. The proposed method improves classification accuracy to 99.04% and 99.75% for the OCI and Kvasir datasets, respectively, highlighting its ability to select highly relevant features and strengthen decision boundaries. Moreover, BFO-PE's entropy-guided optimization avoids local minima, a common limitation in swarm intelligence-based models. This methodological advancement offers a viable solution to overcoming noise and convergence issues in medical imaging, ensuring accurate and reliable clinical diagnosis support.

5. CONCLUSION

This research proposes the BFO-PE for the selection of important features and ResNet152 for

feature extraction to enhance the effectiveness of oral cancer classification. This research introduces the integration of BFO-PE for entropy-driven feature selection with ResNet-152-based deep feature extraction in the domain of histopathological cancer classification. Through optimizing the feature space and leveraging residual learning, the approach provides high accuracy and generalizability while addressing basic challenges like feature redundancy and overfitting. The probability entropy scores the spawning decision in BFO; it selects subsets of features by simulating the spawning behavior of bitterling fish and updates the iterations to an optimal feature. This hypothesis lays the groundwork for developing an intelligent, automated screening system capable of improving diagnostic confidence and reducing clinical workload. The pre-processing stage standardizes image dimensions through interpolation to preserve important visual patterns. ResNet152 extracts high-level, discriminative features through its deep residual blocks, enabling the model to capture complex visual cues. The BFO-PE algorithm evaluates and selects the most relevant features by simulating fish spawning behavior directed through entropy-based scoring, resulting in an optimized feature set. Combined, these components improve classification effectiveness while minimizing the dimensionality and computational overhead, making the approach appropriate for reliable and efficient oral cancer detection. The residual block enables the network to learn hierarchical and complex features to classify cancer and non-cancer. The proposed method achieves better accuracy of 99.04 percent on the OCI dataset and 99.75 percent on the Kvasir dataset when compared with existing methods like SVM.

Although the proposed approach attains superior accuracy, its evaluation is limited to binary classification on specific datasets focused on oral and gastrointestinal images. Generalizing the approach to other cancer types or different medical imaging modalities remains an unexplored area. Additionally, while the model demonstrates high accuracy, the computational overhead during the optimization and feature selection stages could be further reduced. Future work can explore the classification of various tumor images using a hybrid approach to enhance overall performance.

REFERENCES

- [1] G. A. I. Devindi *et al.*, “Multimodal deep convolutional neural network pipeline for AI-assisted early detection of oral cancer,” *IEEE Access*, vol. 12, pp. 124375–124390, 2024.
- [2] 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.
- [3] Y.-H. Kim *et al.*, “Retrospective analysis on prognosis of oral cancer patients according to surgical approaches for effective cancer ablation: swing approach versus visor approach,” *Maxillofac. Plast. Reconstr. Surg.*, vol. 46, no. 1, p. 15, Apr. 2024.
- [4] N. S. Piyarathne *et al.*, “A comprehensive dataset of annotated oral cavity images for diagnosis of oral cancer and oral potentially malignant disorders,” *Oral Oncol.*, vol. 156, no. 106946, p. 106946, Sep. 2024.
- [5] B. Goswami, M. K. Bhuyan, S. Alfarhood, and M. Safran, “Classification of oral cancer into pre-cancerous stages from white light images using LightGBM algorithm,” *IEEE Access*, vol. 12, pp. 31626–31639, 2024.
- [6] J. B. Poell *et al.*, “Oral cancer prediction by noninvasive genetic screening,” *Int. J. Cancer*, vol. 152, no. 2, pp. 227–238, Jan. 2023.
- [7] B. S. Deo, M. Pal, P. K. Panigrahi, and A. Pradhan, “An ensemble deep learning model with empirical wavelet transform feature for oral cancer histopathological image classification,” *Int. J. Data Sci. Anal.*, Feb. 2024.
- [8] M. Sakharkar *et al.*, “Non-invasive screening for laryngeal cancer using the oral cavity as a proxy for differentiation of laryngeal cancer versus leukoplakia: A novel application of ESS technology and artificial intelligence supported statistical modeling,” *Am. J. Otolaryngol.*, vol. 46, no. 1, p. 104581, Jan. 2025.
- [9] M. Al Duhayyim *et al.*, “Sailfish optimization with deep learning based oral cancer classification model,” *Comput. Syst. Sci. Eng.*, vol. 45, no. 1, pp. 753–767, 2023.
- [10] H. Lin, H. Chen, and J. Lin, “Deep neural network uncertainty estimation for early oral cancer diagnosis,” *J. Oral Pathol. Med.*, vol. 53, no. 5, pp. 294–302, May 2024.
- [11] Q. Huang, H. Ding, and N. Razmjoooy, “Oral cancer detection using convolutional neural network optimized by combined seagull optimization algorithm,” *Biomed. Signal Process. Control*, vol. 87, no. 105546, p. 105546, Jan. 2024.
- [12] M. Das, R. Dash, S. Kumar Mishra, and A. Kumar Dalai, “An ensemble deep learning model for oral squamous cell carcinoma detection using histopathological image

- analysis,” *IEEE Access*, vol. 12, pp. 127185–127197, 2024.
- [13] V. Sharmila and S. Geetha, “Gastro Intestinal Disease Classification Using Hierarchical Spatio Pyramid TranfoNet With PitTree Fusion and Efficient-CondConv SwishNet,” *IEEE Access*, vol. 12, pp. 113972–113987, 2024.
- [14] A.-U. Rahman *et al.*, “Histopathologic oral cancer prediction using oral squamous cell carcinoma biopsy empowered with transfer learning,” *Sensors (Basel)*, vol. 22, no. 10, p. 3833, May 2022.
- [15] X. Wei, L. Chanjuan, J. Ke, Y. Linyun, G. Jinxing, and W. Quanbing, “Convolutional neural network for oral cancer detection combined with improved tunicate swarm algorithm to detect oral cancer,” *Sci. Rep.*, vol. 14, no. 1, p. 28675, Nov. 2024.
- [16] Q. Huang, H. Ding, and N. Razmjoo, “Optimal deep learning neural network using ISSA for diagnosing the oral cancer,” *Biomed. Signal Process. Control*, vol. 84, no. 104749, p. 104749, Jul. 2023.
- [17] S. Song, X. Ren, J. He, M. Gao, J. Wang, and B. Wang, “An optimal hierarchical approach for Oral Cancer diagnosis using rough set theory and an amended version of the competitive search algorithm,” *Diagnostics (Basel)*, vol. 13, no. 14, Jul. 2023.
- [18] M. S. Hossain *et al.*, “DeepPoly: Deep learning-based polyps segmentation and classification for autonomous colonoscopy examination,” *IEEE Access*, vol. 11, pp. 95889–95902, 2023.
- [19] A. A. Demirbaş, H. Üzen, and H. Firat, “Spatial-attention ConvMixer architecture for classification and detection of gastrointestinal diseases using the Kvasir dataset,” *Health Inf. Sci. Syst.*, vol. 12, no. 1, p. 32, Dec. 2024.
- [20] P. Liu, W. Qian, J. Cao, and D. Xu, “Semi-supervised medical image classification via increasing prediction diversity,” *Appl. Intell.*, vol. 53, no. 9, pp. 10162–10175, May 2023.
- [21] OCI dataset:
<https://www.kaggle.com/datasets/shivam17299/oral-cancer-lips-and-tongue-images>.
[Accessed: 01-Mar-2025].
- [22] Kvasir dataset:
<https://www.kaggle.com/datasets/meetnagadia/kvasir-dataset>. [Accessed: 05-Mar-2025].
- [23] L. Zareian, J. Rahebi, and M. J. Shayegan, “Bitterling fish optimization (BFO) algorithm,” *Multimed. Tools Appl.*, vol. 83, no. 31, pp. 75893–75926, Feb. 2024.