SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES
http://saucis.sakarya.edu.tr/
e-ISSN: 2636-8129
Publisher: Sakarya University
Vol. 8, No. 1, 58-75, 2025
DOI: 10.35377/saucis...1579003

RESEARCH ARTICLE

Enhanced Classification of Ear Disease Images Using Metaheuristic Feature Selection

Furkancan Demircan[1*], Murat Ekinci[2], Zafer Cömert[1], Eyüp Gedikli[3]

[1*]Software Engineering, Faculty of Engineering and Natural Sciences, Samsun University, Türkiye, ror.org/02brte405
[2] Computer Engineering, Faculty of Engineering, Karadeniz Technical University, Trabzon, Türkiye, ror.org/03z8fyr40
[3] Software Engineering, of Faculty of Technology, Karadeniz Technical University, Trabzon, Türkiye, ror.org/03z8fyr40

Corresponding author:
Furkancan Demircan, Software Engineering, Faculty of Engineering and Natural Sciences, Samsun University, Samsun, Türkiye
furkancan.demircan@samsun.edu.tr

Article History:
Received: 04.11.2024
Revised: 22.02.2025
Accepted: 24.02.2025
Published Online: 27.03.2025

ABSTRACT
Ear diseases are characterized by various symptoms, including balance disturbances, delayed speech development in children, headaches, fever, and hearing loss. To prevent further complications, these conditions must be diagnosed and treated promptly. The traditional diagnostic method has been an otoscope examination by otolaryngologists. However, the accuracy of this approach is contingent upon the clinician's expertise and the quality of the equipment used, which can render it susceptible to misdiagnosis. Incorrect diagnoses may result in the administration of antibiotics unnecessarily, disease progression, and other adverse consequences. This study aims to evaluate the efficacy of computationally efficient machine learning models in classifying ear disease images. To enhance classification accuracy, a Histogram of Oriented Gradients (HOG) was employed for feature extraction and optimization algorithms were utilized for feature selection. The Whale Optimization Algorithm (WOA) effectively selected informative features for the k-Nearest Neighbors (kNN) model, achieving a classification accuracy of 92.6%. Furthermore, the Support Vector Machine (SVM) model achieved an accuracy of 92% using a feature map comprising features selected by a range of optimization algorithms. The experimental findings emphasize the potential of strategic feature selection in enhancing the performance of classical machine learning models for ear disease classification. By employing computationally efficient techniques such as HOG and optimization algorithms, these models can attain classification accuracies that are on par with those of more resource-intensive deep learning approaches. Such developments facilitate the creation of accessible and efficient diagnostic tools, particularly beneficial in resource-constrained clinical settings. The findings of this study provide a basis for further research to enhance the diagnostic precision of machine learning-based techniques in medical imaging.

Keywords: Machine learning, Optimization, Feature extraction, Feature selection

1. Introduction

The human ear is a highly sophisticated organ consisting of three sections—the outer, middle, and inner ear—that work in harmony to capture, amplify, and convert sound waves into neural signals, which are then transmitted to the brain for auditory perception. Despite its advanced structure and functionality, the ear is prone to various disorders that can obstruct the transmission of sound waves and lead to hearing loss. Diseases affecting the eardrum, such as otitis media (OM), cerumen impaction, and myringosclerosis, can reduce the eardrum's elasticity or even cause it to rupture, resulting in diminished hearing capabilities. Additionally, certain conditions can affect the ear canal, distorting or reducing the sound waves transmitted to the eardrum[1].

Beyond hearing impairments, ear disorders can also lead to systemic effects on the body, such as fever, nausea, and itching in the ear. Common ailments like otitis externa (swimmer's ear) and otitis media, if left untreated, may escalate into serious complications, including balance disorders, developmental delays in speech among children, and a significant reduction in the quality of life. These challenges highlight the importance of timely and accurate diagnosis and treatment to prevent health complications and improve overall well-being[1].

Traditionally, the diagnosis of ear diseases has relied on examining the ear canal and eardrum using a medical instrument known as an otoscope performed by trained otolaryngologists. This process typically involves waiting for the disease to reach a certain stage before a diagnosis can be confirmed. The treatment approach is implemented following the diagnosis, and its effectiveness is monitored based on the patient's response. Depending on the observed outcomes, the treatment may be adjusted or continued[1].

Cite as: F. Demircan, M. Ekinci, Z. Cömert, and E. Gedikli, "Enhanced Classification of Ear Disease Images Using Metaheuristic Feature Selection," *Sakarya University Journal of Computer and Information Sciences*, vol. 8, no. 1, pp. 58-75, 2025, doi: 10.35377/saucis...1579003

While this method is widely used, its accuracy depends on the clinician's expertise and the quality of the otoscope used. This reliance introduces significant subjectivity into the diagnostic process, increasing the potential for human error. Misdiagnoses can lead to delayed or inappropriate treatments, sometimes resulting in irreversible complications, such as disease progression or unnecessary interventions. These limitations highlight the need for more objective, accurate, and reliable diagnostic tools to mitigate human error risks and improve patient outcomes[1].

Integrating autonomous decision support systems into the diagnostic process has demonstrated significant potential for improving the accuracy and efficiency of identifying ear diseases. By leveraging advanced artificial intelligence (AI) techniques, these systems can analyze otological images with precision, identifying patterns and abnormalities that might be overlooked by human observers. This technology provides objective, data-driven insights, reducing clinicians' reliance on subjective interpretations. Moreover, AI-powered systems can incorporate patient history and other contextual data to support more nuanced and individualized diagnoses[1].

A key advantage of this approach is its ability to minimize the overuse of antibiotics, a common issue arising from misdiagnoses or precautionary prescriptions. By providing clinicians with reliable diagnostic support, these systems can help differentiate between conditions requiring antibiotic treatment and those not, promoting responsible antibiotic usage. This benefits individual patient care and contributes to broader public health efforts to combat antibiotic resistance. Adopting AI-driven decision support systems represents a step toward more accurate, consistent, and efficient diagnostics, particularly in resource-limited clinical settings where access to specialist expertise may be constrained[2].

Previous studies have focused on classifying ear diseases using machine learning and deep learning techniques. However, deep learning often demands significant computational resources and prolonged processing times. This study adopts a more efficient approach to address these challenges by utilizing the Histogram of Oriented Gradients (HOG) algorithm for feature extraction, a traditional yet effective method. Conventional machine learning algorithms are employed for the classification step, providing a less resource-intensive alternative to deep learning. Additionally, feature selection techniques are integrated into the workflow to enhance the model's performance by ensuring that only the most relevant features contribute to the classification process.

2. Literature Reviews

Artificial intelligence and deep learning have led to significant advances in diagnosing ear diseases. Many studies have employed convolutional neural networks (CNNs) to enhance the accuracy of diagnoses across a range of ear disease types.

Wu et al. proposed a mobile application for efficient monitoring and classification of pediatric otitis media using Xception and MobileNet-V2 convolutional neural network (CNN) architectures pre-trained on ImageNet for background classification. The proposed method's training and testing were applied to 12203 images containing three types of otitis media diseases obtained from Shenzhen Children's Hospital between January 2016 and December 2019. As a result of the study, Xception and MobileNet-V2 architectures achieved 97.82% and 96.04% accuracy in the test images on the system and 90.85% and 88.89% accuracy in the test images on the mobile application, respectively[3].

Sundgaard et al. proposed that Inception V3 architecture be pre-trained on ImageNet to diagnose otitis media in otoscope images automatically. The proposed method was trained and tested on 1136 images of three types of otitis media from 519 patients at the Kamide ENT clinic in Shizuoka, Japan. Five different loss functions were compared in the study, and the best success was obtained from the triple loss function with an accuracy of 85%[4].

Alhudhaif et al. proposed a hybrid model based on CNN architecture to support the diagnosis of five types of otitis media diseases. In the training and testing of the proposed architecture, an open-access dataset of 956 otoscope images from Private Van Akdamar Hospital between October 2018 and June 2019, which were classified into eight types of diseases by experts, was used. The decision mechanism of the CNN architecture was developed using a combination of channel and spatial model (CBAM), residual blocks and hyper-column technique. On the same dataset, AlexNet, VGG-16, VGG-19, GoogleNet, ResNet-50 and the proposed method showed 81.40%, 81.40%, 84.88%, 80.23%, 80.23%, 80.23% and 98.26% accuracy for four classes respectively[5].

Tran et al. proposed an automatic diagnostic system for pediatric otitis media images using a level set-based active counter method for segmentation, color, HOG and SSIM features for feature extraction and joint sparse representation algorithm for disease classification. For training and testing the proposed method, 1230 digital otoscope images of children in General Cathay Hospital were used, and experts under two types of diseases diagnosed them. The proposed method showed 91.41% accuracy on the test dataset[6].

Myburgh et al. proposed a mobile application using cloud-based decision trees and neural network structure to diagnose five types of otitis media diseases automatically. In the training and testing of the proposed method, 562 video otoscope images containing five types of otitis media diseases obtained from different sources were used. The proposed decision tree and neural network showed 81.58% and 86.84% accuracy, respectively[7].

Mohammed et al. proposed a system that can automatically diagnose otitis media using CNN-EfficientNetB0 feature extractor, Bayesian optimization hyperparameter selector and CNN-BiLSTM classifier. The combination of CNN-BiLSTM architectures showed higher performance and worked faster than the methods based on CNN architecture alone. In the training and testing of the proposed method, 880 otoscope images with four types of diseases were obtained from the Department of Otorhinolaryngology of the Clinical Hospital of Universidad de Chile. The proposed architecture showed 100% accuracy[8].

Wang et al. proposed the MESIC architecture for classifying intra-ear diseases. The dataset consists of 998 images of three classes obtained from CT scans of patients in Xiangya Hospital. Preprocessing is performed before using the images in the proposed architecture. In the preprocessing, a region of interest with a size of 100x100 pixels is extracted. The proposed architecture extracts features from the images with the VGG-16 model. Three new feature maps are obtained from the extracted features with the attention block. A full-connected layer and softmax are used to classify the feature maps. The proposed architecture showed 90.1% accuracy on the dataset[9].

Cha et al. compared the classification performance of different models to diagnose ear diseases automatically. The dataset was 10544 otoendoscopic images of six classes obtained from patients at Severance Hospital. The images are divided into 80% and 20% for training and testing, respectively. AlexNet, GoogLeNet, ResNet, Inception-V3, Inception-ResNet-V2, SqueezeNet and MobileNet-V2 were selected from open deep learning models. As a result of the application of the dataset on the models, the two models with the highest success were selected. A probability selector is proposed that combines the outputs of the two models. The probability selector classifies the image according to the label with the highest score from the two models. As a result of the study, the proposed model shows 93.67% accuracy[10].

Chen et al. proposed an AI-assisted mobile algorithm that can classify diseases inside the ear. 2161 images of 10 classes obtained from Taipei Veteran General Hospital were used as a dataset. The images are divided into 80% and 20% for training and testing, respectively. "Augmentation" was applied to the images as preprocessing. VGG19, VGG19, Xception, InceptionV3, NASNetLarge and ResNet50 were selected as deep-learning models. The feature maps of the models were extracted using CAM (Class Activation Map). The InceptionV3 model showed the highest accuracy, with 98%. The weight values obtained from the InceptionV3 model were given to the NasNetMobile, MobileNetV2 and DenseNet201 models for use in the mobile system. At this study's end, the mobile system's highest success rate was obtained from the MobileNetV2 model at 97.6%. The proposed study shows that the mobile system can diagnose as successfully as experts[11].

Sundgaard et al. proposed an architecture that uses broadband tympanometry (WBT) signal values to classify otitis media. Broadband tympanometry measures the acoustic properties of the ear canal using frequency signals. The healthy ear canal showed higher absorption values than the diseased one. The study used 1014 WBT measurements obtained from patients aged 2 months to 12 years at the Kamide ENT clinic in Japan as the data set. The measurements were performed using the Titan system. Three data classes were obtained: 488 healthy (NOE), 372 OME and 154 AOM. The data set was divided into 80% for training and 20% for testing; the training data set was again divided into 20% and used for validation. First, a model inspired by the AlexNet architecture is proposed for two classes (healthy and OM). The proposed model is compared with the pre-trained VGG, ResNet and InceptionV3 available in the PyTorch library. Since the pre-trained models are trained image-based, their accuracy is low. The accuracy of the proposed model is 92.6%. In the second approach, OM disease was tried to differentiate but could not be differentiated[12].

Tsutsumi et al. proposed a web-based deep-learning model for automatically classifying ear diseases. The dataset used in the study was a combination of images available on the internet and Google Images images. The dataset comprises 400 images, including 116 AOM, 44 otitis externa, 23 CSOM, and 21 earwax impactions. The dataset is divided into two different classes: categorical and dichotomous. The dataset was divided into 75% for training and 25% for testing. CNN architectures ResNet50, InceptionV3, InceptionResNetV2 and MobileNetV2 were selected for classification. As a result of the study, the highest accuracy success for the binary class was obtained from the MobileNetV2 model with 90%. For the categorical class, the MobileNetV2 model achieved 91% accuracy[13].

Wang et al. proposed a deep-learning model using CT images for COM disease detection. Images from 672 CT scans of 562 patients were used as the dataset. The images were obtained from a 128-channel multi-detector SOMATOM Definition Edge CT scanner. The axial view of the temporal bone allowed both ears to be included in the images, resulting in 1344 ear images. After experts examined the two ears, a final dataset of 1147 images belonging to 3 classes, CSOM, normal and cholesteatoma, was extracted. In the dataset, 85% of the images were used for training and validation and 15% for testing. Faster R-CNN model was used to extract the ear region in the dataset. The number of images increased by applying augmentation to the images in the dataset. The inceptionV3 model was used for feature extraction, and the softmax function was used for classification. The performances of the proposed model and six experts on the test data are compared. The average accuracy of the proposed model on two classes (CSOM and Normal) is 86%. The average accuracy of the experts is 82%[14].

Zeng et al. investigated the potential of deep learning (DL) for identifying key features of otitis media with effusion (OME), a middle ear infection. Within the scope of the study, a data set consisting of 6393 images was used. Their study focused on detecting atelectasis (collapsed tissue) and attic retraction pocket (a specific indentation) in otoscopic images (ear canal images) collected from multiple centers. The researchers developed and validated a DL model (InceptionV3 with CAM) that achieved high accuracy in detecting both features, with 89% accuracy for attic retraction and 79% for atelectasis. The model highlighted the most relevant image regions for identifying these features[15].

In a study exploring the generalizability of deep learning for ear disease detection, Habib et al. investigated how well these algorithms performed on images from different locations. They used a dataset of over 1800 otoscopic images collected across diverse geographic regions: Van (Turkey), Santiago (Chile), and Ohio (USA). The images were classified as normal or abnormal using four deep-learning models: ResNet-50, DenseNet-161, VGG16, and Vision Transformer. Five-fold cross-validation assessed the models' performance on the original dataset and unseen external images. The Vision Transformer model obtained the highest accuracy value of 92%. While the algorithms achieved high accuracy when tested on the initial dataset (internal images), their performance dropped significantly when evaluated on external images from new locations[16].

In a 2021 study, Pham et al. proposed a novel deep-learning approach called EAR-UNet. This model tackles the challenge of segmenting tympanic membranes from otoscopic images. Built upon the foundation of the UNet architecture, EAR-UNet leverages three key components: EfficientNet for efficient feature extraction in the encoder, an Attention gate for improved information flow in the skip connections, and Residual blocks within the decoder to enhance feature learning. To assess its performance, the researchers evaluated EAR-UNet on a dataset containing otoscopic images from 1024 patients, both diagnosed with and without otitis media. Notably, the model achieved an impressive average Dice Similarity Coefficient (DSC) of 0.929, demonstrating its effectiveness in segmentation without requiring additional pre-processing or post-processing steps[17].

Cömert et al. proposed a computerized otoscopy image-based AI model for diagnosing otitis media (OM). The dataset consisted of 880 otoscope images from a publicly available Ear Imagery dataset, categorized into four classes: chronic otitis media, myringosclerosis, earwax plug, and normal conditions. The method involved using Vision Transformers (ViT) for deep feature extraction, combined with Support Vector Machines (SVM) optimized through grid search. The best-performing model, Optimizable SVM, achieved an accuracy of 99.37%, showcasing exceptional classification performance[18].

Demircan et al. proposed a computer-aided diagnosis system for classifying ear diseases using otoscope images. The dataset comprised 880 eardrum images categorized into four classes: chronic suppurative otitis media (CSOM), earwax, myringosclerosis, and normal conditions. The method utilized a Vision Transformer (ViT) as a feature extractor, followed by classification using machine learning algorithms, including k-Nearest Neighbors (kNN), Support Vector Machine (SVM), and Random Forest. The best-performing model, ViT with kNN, achieved an accuracy of 99.00%, demonstrating high classification performance[19].

3. Material and Methods

The study consists of three phases, each addressing a specific component of the proposed methodology. In the first phase, features extracted using the Histogram of Oriented Gradients (HOG) algorithm were classified using traditional machine learning algorithms to establish baseline performance metrics. The second phase investigated the influence of feature selection on classification accuracy by employing metaheuristic optimization algorithms. These algorithms were applied to identify the most relevant features, reducing the dimensionality of the feature space while preserving essential information. In the final phase, the features selected in the previous steps were combined to analyze their collective impact on classification performance. To ensure the reliability of the integration process, duplicate features were eliminated. The results from all three phases were systematically compared, highlighting improvements in classification performance achieved through the proposed methodology.

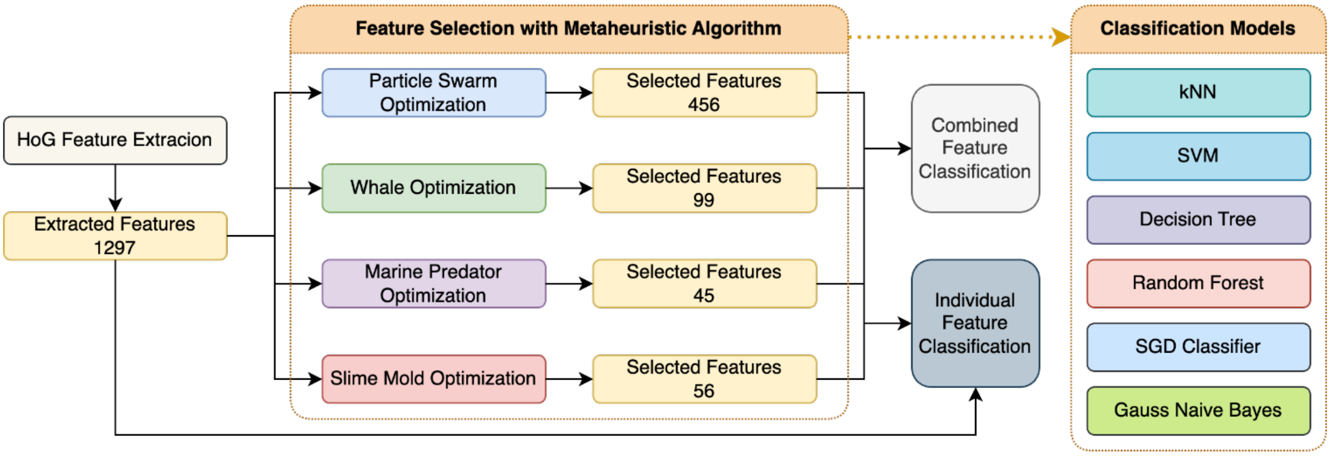

Figure 1. Overview of Feature Extraction, Metaheuristic Optimization, And Machine Learning Classifiers for Automated Ear Disease Diagnosis

Metaheuristic optimization algorithms, including the Whale Optimization Algorithm (WOA), Particle Swarm Optimization (PSO), Marine Predators Algorithm (MPA), and Slime Mold Algorithm (SMA), were integral to the feature selection process. These algorithms iteratively explore the feature space to identify subsets of features that maximize classification accuracy while minimizing redundancy. The selection process is guided by a well-defined fitness function incorporating classification error and the number of selected features, ensuring a balance between accuracy and parsimony. By optimizing feature sets, metaheuristic algorithms significantly enhance the efficiency and effectiveness of the classification process. This structured approach underscores feature selection's critical role in improving the proposed methodology's overall performance for ear disease classification.

3.1. Dataset

The experiment uses a dataset of 880 eardrum images from the Hospital Clínico of Universidad de Chile, categorized into four classes: CSOM, earwax blockage, myringosclerosis, and normal. Images are in JPG format with a 420x380 resolution [7]. Figure 2 presents sample images from each category in the dataset.

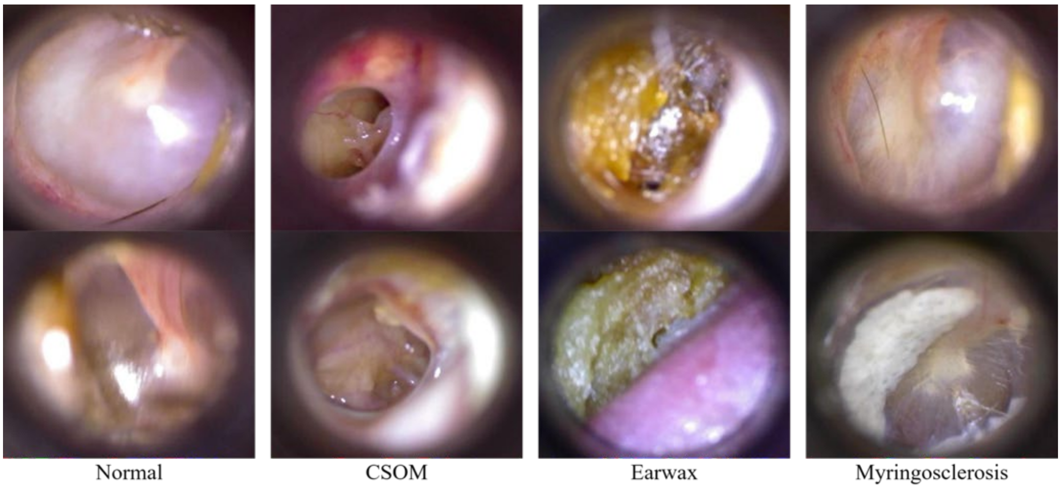

Figure 2. Representative Samples from the Eardrum Image Dataset, Categorized by Disease Type

3.2. Histograms of Oriented Gradients (HOG) Feature Extraction Algorithm

The Histogram of Oriented Gradients (HOG) feature extraction process captures local shape information by effectively analyzing pixel gradients to represent an image’s structure through a feature vector. This widely used technique involves four main steps: gradient computation, orientation binning, descriptor block formation, and feature vector construction. First, image intensity gradients are calculated using convolution to highlight edges and intensity changes, resulting in gradient magnitudes and orientations for each pixel. The image is then divided into small cells (e.g., 8×8 pixels), and a histogram of gradient orientations is constructed for each cell, where each pixel contributes to a bin based on its gradient magnitude and orientation. To enhance robustness against illumination changes, these histograms are normalized within larger blocks composed of multiple cells (e.g., 2×2 cells). Finally, the normalized histograms from all blocks are concatenated to form a comprehensive feature vector [20].

HOG’s success is attributed to its ability to capture detailed local intensity gradients and edge orientations while maintaining resilience to geometric and photometric transformations. This makes it a crucial component in computer vision applications, particularly object detection [20].

3.3. Metaheuristic Optimization Algorithms

Applying metaheuristic optimization techniques enables effective resolution of complex problems through a balanced approach that integrates exploration and exploitation. In contrast to deterministic algorithms, which guarantee optimal solutions but at significant computational expense, metaheuristic methods employ strategies inspired by natural processes to escape local optima and identify high-quality solutions. The versatility of these methods renders them valuable across various disciplines, including engineering, logistics, and machine learning.

3.3.1. Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) is a metaheuristic algorithm inspired by the collective behavior of swarms, such as flocks of birds or schools of fish. It addresses optimization problems by iteratively refining a group of candidate solutions, called particles, within a defined search space. Each particle represents a potential solution and possesses a position and velocity that determine its movement through the search space [21].

The particles adjust their trajectories by considering both their best-known positions and the best-known positions of the entire swarm. This collective sharing of information enables the swarm to converge toward optimal solutions over time. An objective function evaluates each particle’s position's fitness, updating personal and global best positions based on these evaluations.

The algorithm continues this process until a termination criterion is met, such as reaching a maximum number of iterations or achieving a desired fitness level. PSO effectively balances search space exploration with the exploitation of the best-found solutions. Three key parameters influence this balance: the inertia weight, which affects a particle's tendency to continue in its current direction; and two acceleration coefficients, which control the influence of the particle's own best position (cognitive component) and the swarm's global best position (social component) [21].

3.3.2. Whale Optimization Algorithm (WOA)

The Whale Optimization Algorithm (WOA) is a metaheuristic optimization technique inspired by the unique hunting behavior of humpback whales, particularly their bubble-net feeding strategy. This algorithm seeks optimal solutions by simulating the social and cooperative behaviors observed in whale pods during hunting. It operates through three main phases: encircling prey, bubble-net attacking, and searching for prey [22].

In the encircling prey phase, the algorithm assumes that the current best candidate solution represents the prey, and all other candidate solutions (whales) adjust their positions relative to this best solution. This mimics how whales surround their prey in nature. Each whale updates its position by moving closer to the best-known position, effectively narrowing the search area around promising solutions. The bubble-net attacking phase simulates the bubble-net feeding mechanism of humpback whales using two strategies: shrinking encircling and spiral updating. The shrinking encircling strategy involves gradually reducing the distance between the whales and the prey, akin to whales tightening their circle around the prey. The spiral updating strategy models the helical movement of whales as they spiral upward toward the surface while creating bubbles to trap the prey. This combination allows the algorithm to exploit the search space around the current best solution more effectively, refining candidate solutions and enhancing optimization. In the search for prey phase, the algorithm emphasizes exploration by moving whales toward randomly selected candidate solutions rather than the current best one. This simulates the whales' behavior when searching for new prey and helps the algorithm avoid local optima by diversifying the search process. By considering random positions in the search space, the algorithm increases the likelihood of finding the global optimum [22].

3.3.3. Marine Predator Algorithm (MPA)

The Marine Predator Algorithm (MPA) is a modern metaheuristic optimization technique inspired by the hunting strategies of marine animals like fish, sharks, and whales. It efficiently explores and exploits the solution space to find optimal solutions. The MPA process consists of three phases: exploration, transition, and exploitation, which emulate the adaptive behaviors of marine predators and balance global exploration with local exploitation [23].

The algorithm begins by generating a population of candidate solutions, termed "prey," randomly distributed in a multidimensional search space. In the exploration phase, the algorithm mimics the wide-ranging search for prey, using a probabilistic movement pattern known as Lévy flight to explore large areas of the search space. This phase ensures thorough investigation and prevents premature convergence on suboptimal solutions. The transition phase serves as a bridge between exploration and exploitation, fine-tuning the movement of candidate solutions based on their proximity to the best-known solution, referred to as the "best predator." This phase gradually shifts the focus from broad exploration to more targeted exploitation of promising regions while introducing random factors to maintain diversity and avoid local optima. In the exploitation phase, the algorithm intensifies the search for the best solutions, simulating the focused hunting behavior of predators when prey is nearby. Candidate solutions move based on Brownian motion, allowing for precise refinement of solutions in promising regions. The fitness of all candidate solutions is then evaluated, and the best-known solution is updated if a superior solution is found [23].

3.3.4. Slime Mold Optimization Algorithm (SMA)

The Slime Mold Optimization Algorithm (SMA) is a bio-inspired metaheuristic algorithm modeled after the foraging behavior of slime molds. It efficiently explores and exploits the solution space to solve complex optimization problems. The process begins by initializing a population of candidate solutions, or "slime molds," randomly distributed within a multidimensional search space. Each solution is evaluated using a fitness function that determines its quality based on the specific problem [24].

SMA updates the positions of the slime molds by simulating their movement toward better solutions, much like how slime molds move toward nutrient-rich areas. It also incorporates a mechanism to avoid poorer solutions, guiding the slime molds away from suboptimal regions of the search space. To prevent the algorithm from getting stuck in local optima, a random movement component is added to encourage broader exploration [24].

Additionally, SMA mimics the collective behavior of slime molds aggregating around nutrient sources by considering both the best-known solution and the average position of the population. This behavior helps refine the search process and directs the slime molds toward more promising areas. The algorithm iteratively updates the positions and evaluates the fitness of the slime molds until a termination condition, such as a set number of iterations or a desired fitness level, is reached. This balance of exploration and exploitation allows SMA to navigate complex optimization problems and find optimal or near-optimal solutions effectively [24].

3.4. Machine Learning Algorithms

Machine learning, a branch of artificial intelligence (AI), transforms many fields by enabling computers to learn from data and improve their performance without explicit instruction. It develops algorithms that autonomously recognize patterns and make data-driven decisions. Its impact ranges from personalized recommendations and search engine improvements to medical predictions and industrial automation. As the volume of data grows, machine learning's potential to drive innovation and efficiency increases, fundamentally changing how we interact with technology. This integration offers unprecedented opportunities for advances in scientific discovery, business optimization, and societal progress.

3.4.1. k-Nearest Neighbors (kNN)

The k-Nearest Neighbors (kNN) algorithm is a nonparametric classification method based on proximity. It stores the entire data set during training and calculates the distance from a new data point to all training instances to find the k nearest neighbors. The class label for the new point is assigned based on the majority class of these neighbors. The effectiveness of kNN depends on the chosen distance metric, the value of k, and the dimensionality of the feature space[25].

3.4.2. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised learning model for classification tasks. Its goal is to find the optimal hyperplane that maximally separates data points of different classes. The data points closest to the hyperplane, called support vectors, are crucial for defining the decision boundary. To handle non-linearly separable data, SVM uses kernel functions to project the data into a higher-dimensional space[25].

3.4.3. Stochastic Gradient Descent (SGD) Classifier

Stochastic Gradient Descent (SGD) is an iterative optimization algorithm that trains linear classifiers. It updates model parameters by processing small batches of data, making it efficient for large datasets. By minimizing a loss function, such as hinge or logistic loss, SGD adjusts model weights to improve predictive accuracy and can incorporate regularization techniques to prevent overfitting[26].

3.4.4. Naive Bayes

Naive Bayes is a probabilistic classifier based on Bayes' theorem that assumes features are conditionally independent given the class label for efficient computation. It computes the posterior probability for each class based on the observed features and assigns the class with the highest probability. While this independence assumption simplifies the model, it can hinder performance when significant feature dependencies exist[26].

3.4.5. Random Forest

Random Forest is an ensemble learning method that combines multiple decision trees to improve predictive accuracy. Each tree is constructed on a random subset of the data, and feature selection at each node is randomized. By aggregating the predictions of individual trees through majority voting, random forests effectively reduce overfitting and enhance generalization performance [25].

3.4.6. Decision Tree

A decision tree is a supervised learning algorithm that creates a tree-like model of decisions and their possible consequences. It partitions the data into subsets based on feature values, aiming to maximize information gain or minimize impurity at each split. Decision trees can handle both categorical and numerical data and are interpretable, but they can be prone to overfitting [25].

3.5. Metaheuristic Feature Selection

The application of metaheuristic algorithms for feature selection in this study was designed to optimize the classification performance by reducing the dimensionality of the feature space and selecting the most informative attributes. This process was guided by a mathematical fitness function that balanced classification accuracy and model simplicity. The fitness function, denoted as $F(X)$, was formulated as a weighted combination of the classification error rate and the cardinality of the selected feature set, as shown in Equation (1):

$$F(X) = \alpha \cdot Error(X) + \beta \cdot \left(\frac{\|X\|}{\|X_{total}\|}\right) \tag{1}$$

In this equation, α and β are weighting factors, $Error(X)$ represents the classification error rate, $\|X\|$ is the number of selected features, and $\|X_{total}\|$ is the total number of available features. The optimization aimed to minimize $F(X)$, thus achieving a trade-off between maximizing classification accuracy and minimizing the number of selected features to enhance computational efficiency. The error value, denoted as $Error(X)$, is calculated using the formulation provided in Equation (2). This equation measures the proportion of incorrect predictions made by the model relative to the total number of predictions.

Table 1. Selected Parameters and Values for Machine Learning Algorithms

Machine Learning Algorithm	Parameter	Value	Explanation
kNN	Metric	Minkowski	The distance metric is used for neighboring computation. Minkowski is a generalization of Euclidean and Manhattan distances.
	Neighbors	Labels Size (4)	Several nearest neighbors used for classification correspond to the number of class labels.
	Weights	Uniform	Uniform weight means all points in the neighborhood are equally weighted.
SVM/SVC	Regularization Parameter	1	Controls the trade-off between achieving a low error and a large margin.
	Kernel	Radial Basis Function (RBF)	The kernel function is used to transform data into a higher-dimensional space.
	Gamma	1 / (Features * X)	Kernel coefficient for 'rbf' kernel; smaller values mean a wider influence.
SGD Classifier	Loss	Hinge	Specifies the loss function; hinge is used for linear Support Vector Machines.
	Penalty	L2	L2 regularization adds squared penalty terms to prevent overfitting.
	Max Iterations	10	Maximum number of iterations for the algorithm to converge.
	Learning Rate	Optimal	The learning rate schedule adjusts step size based on data and iterations.
Gaussian Naïve Bayes	Variance Smoothing	1,00E-09	A portion of the largest variance was added for numerical stability.
	Priors	None	Prior probabilities of classes; if None, they are inferred from data.
Random Forest Classifier	Number of Trees	Labels Size (4)	Number of decision trees built in the forest. More trees improve accuracy but increase computation.
	Criterion	Gini	Function to measure split quality: Gini impurity is a common measure.
	Bootstrap	TRUE	Whether bootstrap samples are used to build each tree.
Decision Tree Classifier	Strategy	Best	Choosing the best-split strategy based on impurity.
	Criterion	Gini	The function is to measure the quality of a split; Gini impurity is used.
	Max Depth	None	Maximum depth of the tree; if None, nodes are expanded until all leaves are pure.

$$Error(X) = 1 - \frac{\sum_{i=1}^{N} g(\hat{y}_i, y_i)}{N} \tag{2}$$

$$g(\hat{y}_i, y_i) = \begin{cases} 1, & \hat{y}_i = y_i \\ 0, & \hat{y}_i \neq y_i \end{cases} \tag{3}$$

Where, $\boldsymbol{N}$ represents the total number of samples, $\boldsymbol{\hat{y}_i}$ denotes the predicted value for the i-th sample, and $\boldsymbol{y_i}$ corresponds to the true value. The Equation (3), $\boldsymbol{g(\hat{y}_i, y_i)}$ acts as an indicator function that returns a value of 1 when the predicted value matches the true value (indicating a correct prediction) and 0 otherwise (indicating an incorrect prediction).

4. Results

The experimental setup comprised a desktop computer with an Intel i9 13900k processor and an NVIDIA RTX 4080 graphics card with 16GB of GDDR6X memory. The CUDA toolkit enhanced GPU performance, and the PyTorch library was the main framework. The scikit-learn library was also utilized for data preprocessing and implementing baseline machine-learning models.

The initial parameters and their corresponding values for the machine learning algorithms employed in this study are presented in Table 1. These parameters were carefully chosen based on their default settings, which are widely accepted in the literature and serve as a reliable baseline for performance evaluation. Adjustments to these parameters can be considered in future work to optimize results for specific datasets or tasks.

This study extracted HOG features from in-ear images, resulting in a 1297-dimensional feature matrix using a 32x32 filter. The initial experiment classified these features using a variety of algorithms, including kNN, SVM, DT, RF, SGD, and GNB. The performance metrics from the experiments are presented in Table 2.

Table 2. Performance Metrics of the First Experimental Study (The Best-Performing Model Highlighted in Bold)

Models	Accuracy	Precision	Recall
kNN	0,886	0,889	0,886
SVM	**0,914**	**0,914**	**0,914**
SGDC	0,869	0,878	0,869
GNB	0,767	0,768	0,767
RF	0,812	0,813	0,812
DT	0,761	0,768	0,761

The SVM outperformed all other models in classifying ear diseases, achieving the highest accuracy, precision, and recall with minimal false positives and negatives. The kNN algorithm closely followed, demonstrating strong classification capabilities. The SGDC showed moderate performance, while GNB and RF models exhibited balanced but lower effectiveness. The DT model was the least effective among those evaluated.

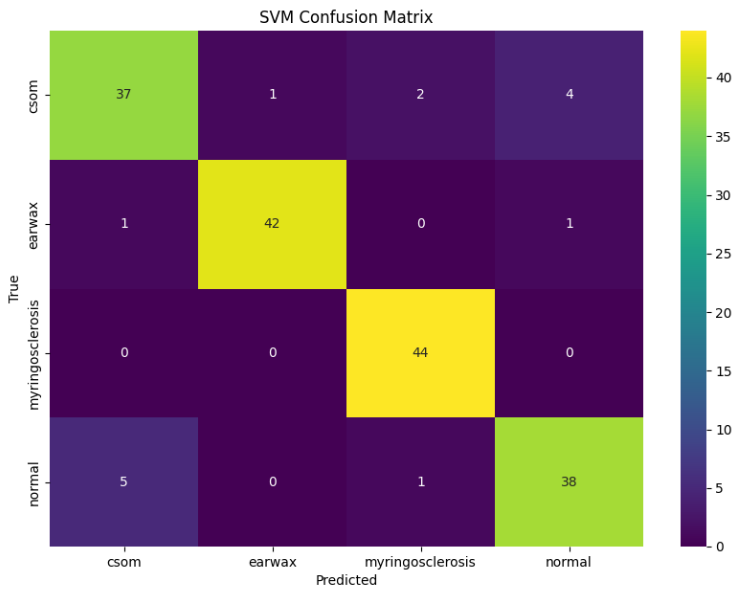

Figure 3. Confusion Matrix of SVM Algorithm

Table 3. Initial Parameters for Metaheuristic Optimization Algorithms

Optimization Algorithm	Parameters	Values
Whale Optimization Algorithm	C Constant	1
	Lower Bond	0
	Upper Bond	1
	Threshold	0,5
	B constant	1
	Population	100
Slime Mold Optimization Algorithm	Lower Bond	0
	Upper Bond	1
	Threshold	0,5
	Control local and global	0,03
	Population	100
Marine Predator Algorithm	Lower Bond	0
	Upper Bond	1
	Threshold	0,5
	Levy Component	1,5
	Constant (P)	0,5
	Fish Aggregating Devices Effect	0,2
	Population	100
Particle Swarm Optimization	Lower Bond	0
	Upper Bond	1
	Threshold	0,5
	Cognitive Factor	2
	Social Weight	2
	Inertia Weight	0,9
	Maximum Velocity	$\frac{(Upper\ Bond\ -\ Lower\ Bond)}{2}$
	Population	100

These findings highlight the reliability and accuracy of the SVM and kNN models for ear disease classification. This analysis can guide the selection of machine learning algorithms based on specific classification needs and resource considerations. The confusion matrix for the SVM algorithm, which demonstrated the highest accuracy, is presented in Figure 3.

The subsequent experimental investigation used metaheuristic optimization algorithms to select pertinent features from the HOG-derived feature maps. The algorithms employed were MPA, SMA, PSO, and WOA. Each algorithm was executed 50 times over 1,000 iterations with consistent parameter settings to facilitate comparison.

The initial parameters selected for the metaheuristic optimization algorithms utilized in this study are presented in Table 3. These parameters have been chosen based on their default values, as commonly recommended in the literature, to ensure a standardized and unbiased evaluation of the algorithms' performance.

The generated feature maps were evaluated using the kNN classifier to identify features with the highest classification accuracy. Figure 4 illustrates the mean fitness values attained by each algorithm.

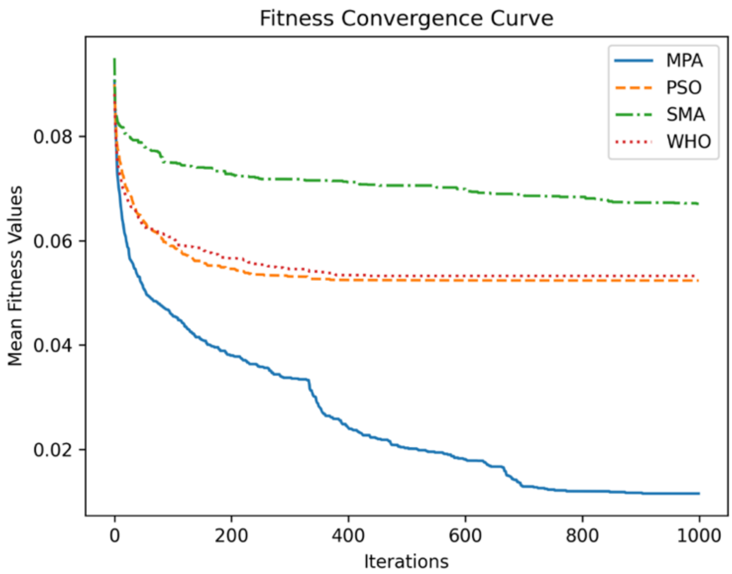

Figure 4. Average Fitness Values of the MPA, PSO, SMA, and WOA Algorithms Over 50 Runs with 1000 Iterations

Figure 4 displays the convergence behaviors of four optimization algorithms: WOA, SMA, PSO, and MPA. The x-axis denotes the number of iterations, while the y-axis represents the fitness value, indicating each algorithm's performance over time. WOA (red line) and SMA (green line) show rapid initial decreases in fitness value but plateau after approximately 20 iterations, limiting further improvement. PSO (orange line) also converges quickly but stabilizes at a higher fitness level, suggesting less optimal solutions. In contrast, MPA (blue line) exhibits a slower initial improvement with continuous enhancement throughout the iterations, ultimately achieving the lowest fitness value among all algorithms.

These findings highlight the superior optimization performance of the MPA algorithm due to its sustained improvement and ability to reach the most optimal solution. The results underscore the importance of considering initial convergence rates and long-term performance when selecting optimization algorithms for scenarios requiring optimal solutions.

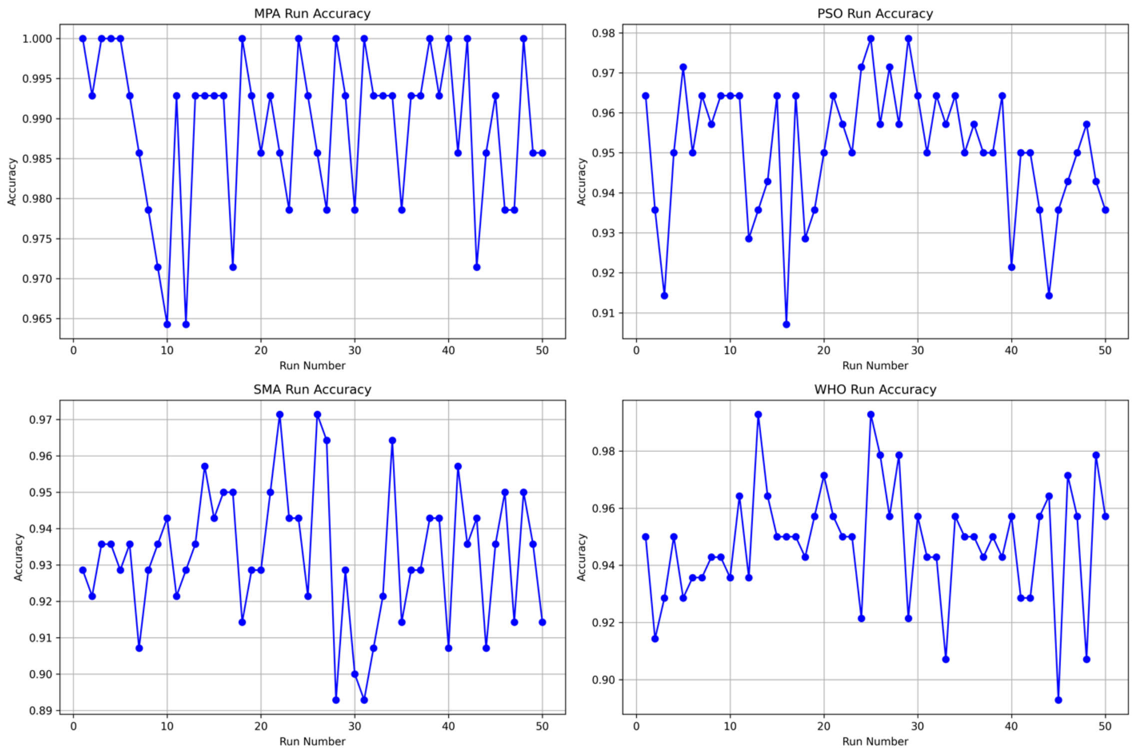

Figure 5. The Accuracy of 50 Iterations for the Metaheuristic Optimization Algorithm

Figure 5 displays the classification accuracy achieved by the kNN algorithm when trained and tested on feature sets selected using various metaheuristic optimization algorithms. The metrics presented are based on 50 iterations and reflect the kNN classifier's performance in evaluating different feature maps to identify those that yield the highest accuracy. By employing the features selected by each optimization algorithm to calculate performance metrics, the study aims to determine the most effective feature selection strategy and to identify the feature set with the highest overall classification accuracy.

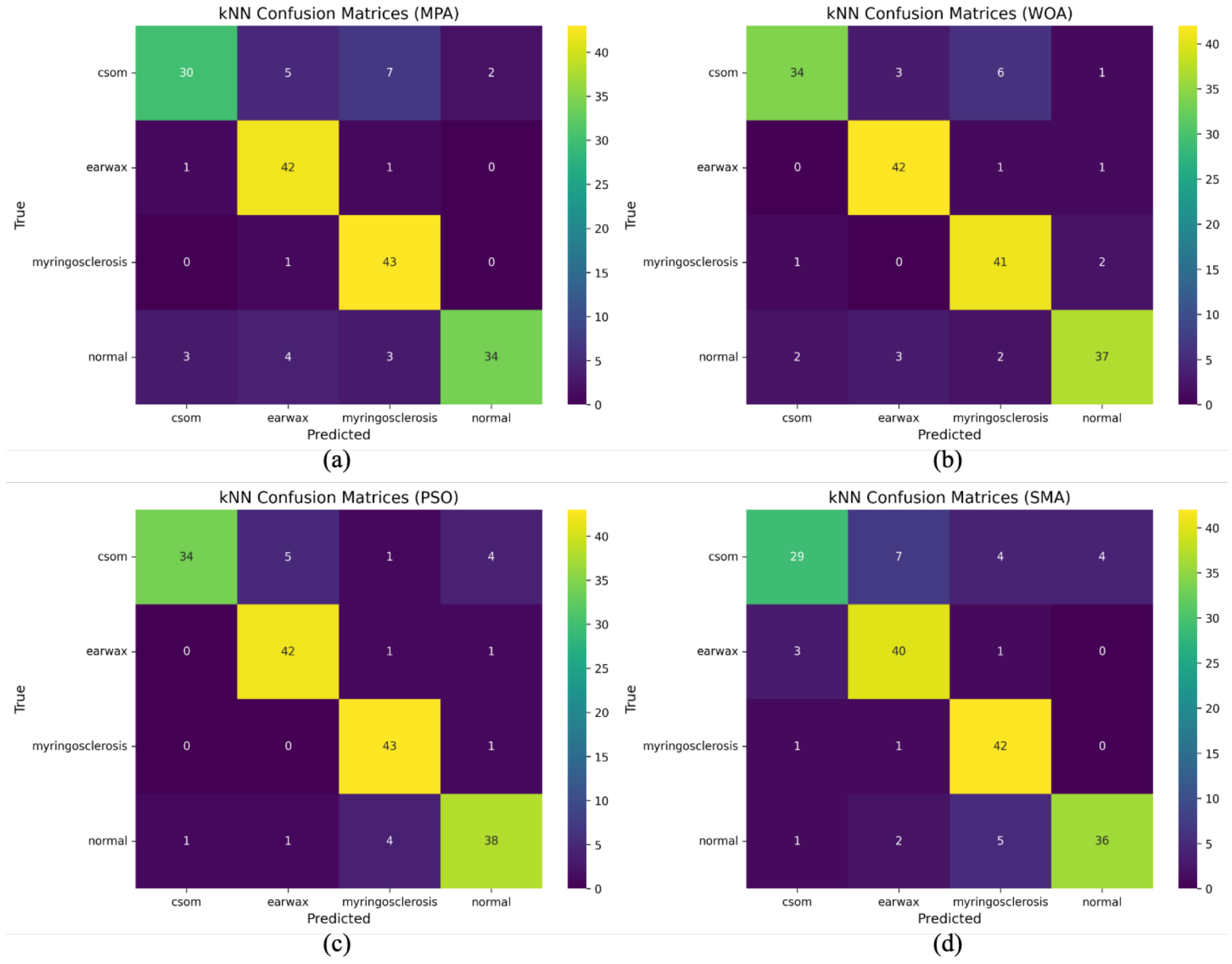

Figure 6. Confusion Matrices Obtained from the First Run with Corresponding Classification Accuracies: MPA - 84.66% (a), WOA - 87.50% (b), PSO - 89.20% (c), and SMA - 83.52% (d)

This study selected feature maps from HOG extraction using four metaheuristic optimization algorithms: WOA, MPA, SMA, and PSO. Table 4 shows the number of features selected by each. Figure 6 presents the confusion matrices obtained from the first run, where the kNN algorithm achieved a classification accuracy of 89.20% using features selected by the PSO algorithm. This initial run provides insight into the model's classification performance based on the selected feature subset. The experiment was repeated for 50 independent runs to ensure a robust evaluation. The highest accuracy achieved by the kNN algorithm across these runs, along with the corresponding run number and the number of selected features, is summarized in Table 4. The "Best Run" column in Table 4 represents the highest accuracy obtained from the classification results of the kNN algorithm using the features selected by the optimization algorithms. The selected features from this best-performing run were subsequently used to generate feature maps for the later stages of the study.

Table 4. Number of Features Selected by Metaheuristic Algorithms and Corresponding Iterations

Metaheuristic Algorithm	Best Run (From 50)	Number of Selected Feature
PSO	29	456
MPA	3	45
SMA	22	56
WOA	25	99

Table 5 displays the performance metrics from the feature selection process for each optimization algorithm, showing the classification accuracy of machine learning models trained on their selected feature sets.

Table 5. Summary of Classification Algorithm Performance (The Best-Performing Model is Highlighted in Bold)

Models	MPA			SMA			PSO			WOA		
	Acc	Prec	Rec	Acc	Prec	Rec	Acc	Prec	Rec	Acc	Prec	Rec
kNN	0,88	0,887	0,88	0,863	0,878	0,863	0,88	0,889	0,88	**0,926**	**0,929**	**0,926**
SVM	0,892	0,895	0,892	0,875	0,88	0,875	0,914	0,914	0,914	0,897	0,899	0,897
SGDC	0,426	0,771	0,426	0,607	0,721	0,607	0,806	0,835	0,806	0,778	0,794	0,778
GNB	0,642	0,668	0,642	0,596	0,63	0,596	0,778	0,781	0,778	0,715	0,724	0,715
RF	0,789	0,79	0,789	0,801	0,803	0,801	0,795	0,793	0,795	0,778	0,785	0,778
DT	0,801	0,811	0,801	0,772	0,779	0,772	0,772	0,785	0,772	0,778	0,785	0,778

In this study, we evaluated the effectiveness of several metaheuristic optimization algorithms for feature selection to enhance machine learning model performance. The results indicated that the WOA algorithm was the most effective feature selection method, achieving the highest performance metrics, particularly when used with the kNN and RF classifiers. Additionally, the PSO algorithm demonstrated significant efficacy, especially in improving the performance of the SVM and SGDC classifiers. These findings underscore the crucial role of selecting appropriate feature selection algorithms to optimize the efficacy of machine learning models. The confusion matrix for the SVM model exhibiting the greatest performance is presented in Figure 7.

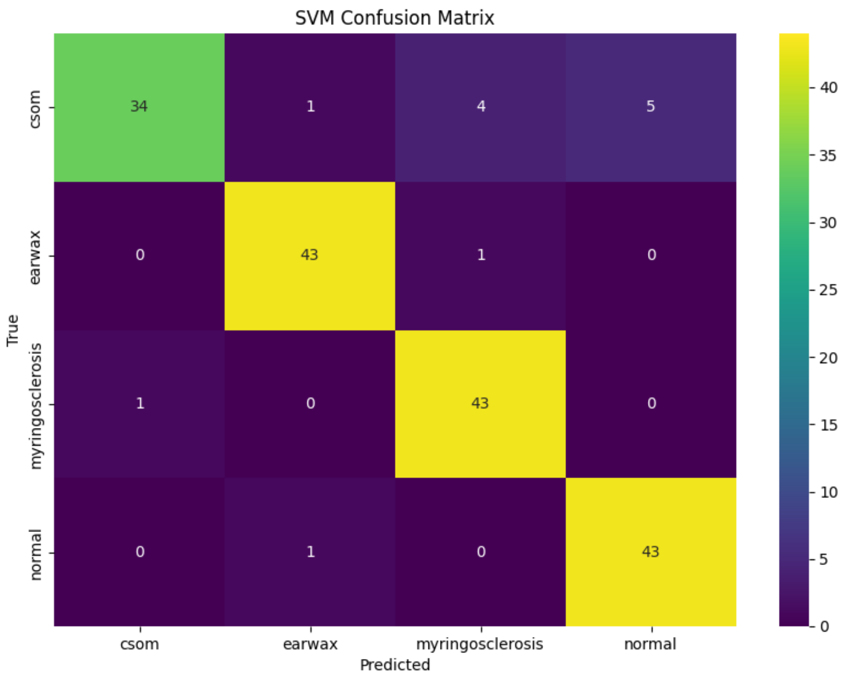

Figure 7. Confusion Matrix of the kNN Algorithm with Feature Selection

In the final experiment, a feature map was created by combining selected features from the optimization algorithms, resulting in 578 unique features. Performance metrics from this study are shown in Table 6.

Table 6. Performance Metrics of Algorithm for Combined Features (The Best-Performing Model Highlighted in Bold)

Models	Accuracy	Precision	Recall
kNN	0,892	0,895	0,892
SVM	**0,92**	**0,919**	**0,92**
SGDC	0,892	0,894	0,892
GNB	0,778	0,784	0,778
RF	0,789	0,79	0,789
DT	0,812	0,815	0,812

To better illustrate the comparative performance of the models, a graphical representation of accuracy values from Table 6 is provided in Figure 8.

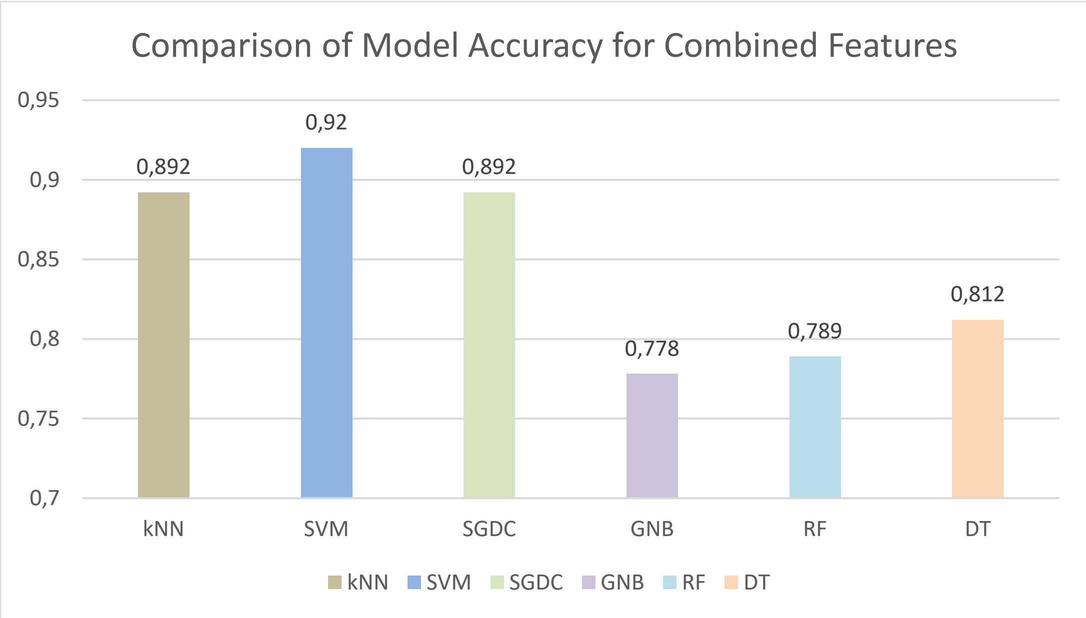

Figure 8. Accuracy Values of Different Models Based on the Combined Features, as Presented in Table 6

A comparative analysis was conducted on six machine learning algorithms to evaluate their performance in classifying ear diseases based on accuracy, precision, and recall metrics. The SVM model demonstrated superior performance, achieving an accuracy of 0.92, precision of 0.919, and recall of 0.92, indicating its effectiveness in minimizing false positives and false negatives. Both kNN and SGD performed well with accuracy scores of 0.892; however, kNN exhibited slightly higher precision at 0.895. GNB and RF showed moderate performance levels, while the DT model had the lowest accuracy at 0.812, highlighting its limited effectiveness for applications requiring high precision. In conclusion, the SVM model was the most effective classifier, outperforming the other algorithms evaluated.

In conclusion, the SVM model was the most effective classifier, outperforming other algorithms. Model selection should consider specific classification requirements. The confusion matrix for SVM is shown in Figure 9.

The model accurately classified 37 out of 44 cases of Chronic Suppurative Otitis Media (CSOM). Still, some misclassifications occurred, primarily labeling CSOM cases as normal or other pathologies, suggesting feature overlap between CSOM and normal cases. It demonstrated high proficiency in identifying earwax impaction, correctly classifying 43 out of 44 instances, and achieving perfect accuracy for all 44 instances of myringosclerosis. For normal cases, 38 out of 44 were correctly identified, with some misclassified as CSOM or myringosclerosis. These findings highlight the need to improve the model's ability to distinguish between CSOM and normal cases to enhance classification accuracy and diagnostic reliability.

A comprehensive investigation assessed the efficacy of diverse feature selection techniques and machine learning algorithms for ear disease classification to identify optimal combinations for enhanced accuracy. Combining WOA with the kNN method yielded the highest classification accuracy of 92.6%, thereby demonstrating the efficacy of the WOA in improving the selection of features. Furthermore, the SVM demonstrated a noteworthy performance, attaining a 92% accuracy rate by utilizing features selected by multiple optimization algorithms. The SVM achieved 91.4% accuracy without explicit feature selection, illustrating its robustness. These findings highlight the crucial importance of strategic feature selection in enhancing the efficacy of machine learning for otitis media diagnosis.

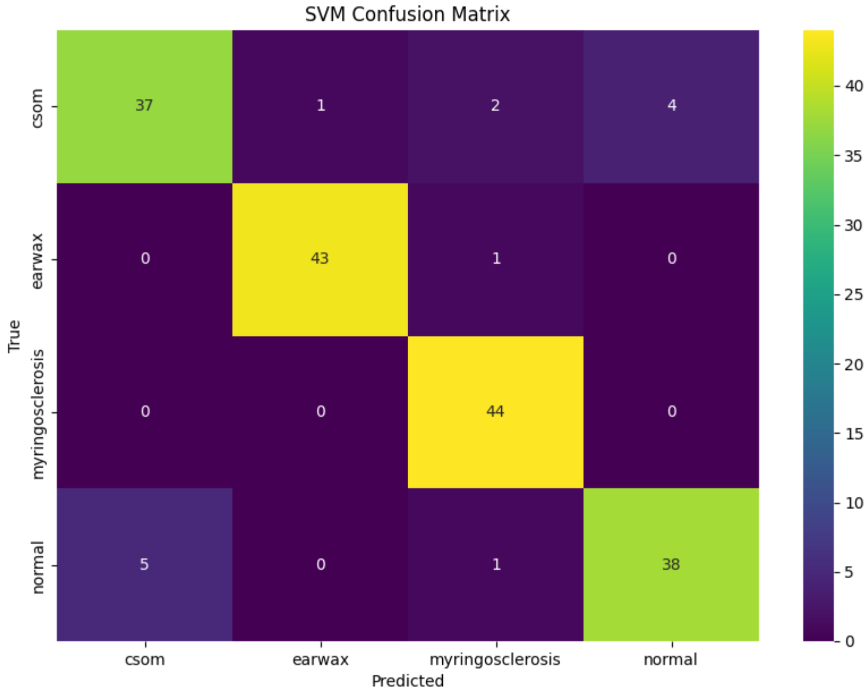

Figure 9. The Confusion Matrix of the SVM Algorithm

5. Discussion

The classification of visual representations of ear diseases has been the subject of investigation in several studies. Table 7 summarizes the datasets utilized, the number of diseases included, and the classification models developed in these studies.

The table highlights various approaches to ear disease classification using different machine learning models, datasets, and feature extraction techniques, showing significant variation in accuracy. Deep learning models generally performed well, with Wu et al. achieving 97.82% accuracy using Xception on a three-class dataset and Alhudhaif et al. improving on this with 98.26% accuracy by combining CNN, CBAM and other techniques on a four-class dataset. However, Sundgaard et al. reported a lower accuracy of 85% using InceptionV3, indicating that model performance can vary based on architecture and dataset. Mohammed et al. reached 100% accuracy by combining CNN with BiLSTM on a four-class dataset, showing hybrid models' potential for capturing spatial and temporal features. Similarly, Uçar et al. achieved 99.06% accuracy using SURF with VGG16, demonstrating that traditional feature extraction methods can enhance deep learning models. Other studies also showed strong results with lightweight architectures. Chen et al. achieved 97.6% accuracy with MobileNetV2 on a ten-class dataset, and Byun et al. reached 97.18% using ResNet18 with a Shuffle Attention Module, showing the benefit of attention mechanisms. However, Wang et al. reported lower accuracies of 86%, using the Faster R-CNN model, suggesting that model selection and dataset complexity can influence outcomes. Cloud-based platforms like Google Cloud AutoML, used by Livingstone et al., achieved 90.9% accuracy on a fourteen-class dataset, demonstrating the effectiveness of automated tools. Tsutsumi et al. achieved 90-91% accuracy using MobileNetV2 on datasets with varying class sizes, further showcasing the adaptability of lightweight models.

The study in this paper presents a computationally efficient approach combining HOG feature extraction with WOA for feature selection and kNN for classification, achieving 92.6% accuracy on a four-class ear imagery dataset. While slightly lower than some deep learning models, this result demonstrates that traditional models, combined with effective feature engineering, can offer competitive performance with lower computational requirements, making them valuable for resource-limited applications.

Table 7. The Previous Studies on Ear Disease Classification

Authors	Datasets	Class Size	Models	Accuracy (%)
Wu et al.[3]	Private	3	Xception	97,82
Sundgaard et al.[4]	Private	3	InceptionV3	85
Alhudhaif et al.[5]	Tympanic Membrane	4	CNN, CBAM, Residual block, and Hypercolumn Technique	98,26
Tran et al.[6]	Private	2	Multitask Joint Sparse Representation Algorithm	91,41
Myburgh et al.[7]	Private	5	NN	86,84
Mohammed et al.[8]	Ear Imagery	4	CNN-BiLSTM	100
Uçar et al.[27]	Ear Imagery	4	SURF + VGG16	99,06
Wang et al.[14]	Private	3	MESIC	90,1
Cha et al.[10]	Private	6	Models + Probability Selectors	93,67
Chen et al.[11]	Private	10	MobileNetV2	97,6
Sundgaard et al.[12]	Private	2	AlexNet Base Model	92,6
Tsutsumi et al.[13]	Mixed Dataset	2, 4	MobileNetV2	90, 91
Wang et al.[9]	Private	2	FasterRCNN + InceptionV3	86
Byun et al.[26]	Private	4	ResNet18+Shuffle Attention Module	97,18
Livingstone et al.[28]	Mixed Dataset	14	Google Cloud AutoML	90,9
Zeng et al.[15]	Private	2	DL model	89
Habib et al.[16]	Private	2	Vision Transformer	92
Pham et al.[17]	Private	2	EAR-UNet	92.9
This paper	Ear Imagery	4	HOG Feature Extraction + WOA Feature Selection + kNN Classification	92.6

6. Conclusions

This study introduces a novel approach to ear disease classification by employing computationally efficient machine learning models in conjunction with advanced feature engineering techniques. The research aims to achieve high classification accuracy by optimizing feature extraction and selection while reducing computational resource demands, facilitating practical clinical implementation and integrating the HOG for feature extraction and the WOA for feature selection to enhance model performance significantly. Notably, the kNN model combined with WOA-selected features achieved an impressive accuracy of 92.6%. The SVM classifier also performed well, attaining 92% accuracy with optimized features and 91.4% without explicit feature selection. These findings demonstrate that traditional machine learning models can perform comparably to computationally intensive deep learning architectures when paired with effective feature engineering. The results contribute to developing efficient diagnostic systems for ear diseases, paving the way for future enhancements in accuracy and accessibility.

References

[1] W. H. Organization, *Primary ear and hearing care training manual*. Genève, Switzerland: World Health Organization, 2023.

[2] Institute of Electrical and Electronics Engineers and Manav Rachna International Institute of Research and Studies, *Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing: trends, perspectives and prospects: COMITCON-2019: 14th-16th February 2019*.

[3] Z. Wu *et al.*, "Deep Learning for Classification of Pediatric Otitis Media," *Laryngoscope*, vol. 131, no. 7, pp. E2344–E2351, Jul. 2021, doi: 10.1002/lary.29302.

[4] J. V. Sundgaard *et al.*, "Deep metric learning for otitis media classification," *Med Image Anal*, vol. 71, Jul. 2021, doi:

10.1016/j.media.2021.102034.
[5] A. Alhudhaif, Z. Cömert, and K. Polat, “Otitis media detection using tympanic membrane images with a novel multi-class machine learning algorithm,” *PeerJ Comput Sci*, vol. 7, pp. 1–22, 2021, doi: 10.7717/PEERJ-CS.405.
[6] T. T. Tran, T. Y. Fang, V. T. Pham, C. Lin, P. C. Wang, and M. T. Lo, “Development of an automatic diagnostic algorithm for pediatric otitis media,” *Otology and Neurotology*, vol. 39, no. 8, pp. 1060–1065, 2018, doi: 10.1097/MAO.0000000000001897.
[7] H. C. Myburgh, S. Jose, D. W. Swanepoel, and C. Laurent, “Towards low cost automated smartphone- and cloud-based otitis media diagnosis,” *Biomed Signal Process Control*, vol. 39, pp. 34–52, Jan. 2018, doi: 10.1016/j.bspc.2017.07.015.
[8] K. K. Mohammed, A. E. Hassanien, and H. M. Afify, “Classification of Ear Imagery Database using Bayesian Optimization based on CNN-LSTM Architecture,” *J Digit Imaging*, vol. 35, no. 4, pp. 947–961, Aug. 2022, doi: 10.1007/s10278-022-00617-8.
[9] Y. M. Wang *et al.*, “Deep Learning in Automated Region Proposal and Diagnosis of Chronic Otitis Media Based on Computed Tomography,” *Ear Hear*, pp. 669–677, 2020, doi: 10.1097/AUD.0000000000000794.
[10] D. Cha, C. Pae, S. B. Seong, J. Y. Choi, and H. J. Park, “Automated diagnosis of ear disease using ensemble deep learning with a big otoendoscopy image database,” *EBioMedicine*, vol. 45, pp. 606–614, Jul. 2019, doi: 10.1016/j.ebiom.2019.06.050.
[11] Y.-C. Chen *et al.*, “Smartphone-based artificial intelligence using a transfer learning algorithm for the detection and diagnosis of middle ear diseases: A retrospective deep learning study,” *EClinicalMedicine*, vol. 51, no. 201, p. 101543, 2022, doi: 10.1016/j.
[12] J. V. Sundgaard *et al.*, “A Deep Learning Approach for Detecting Otitis Media from Wideband Tympanometry Measurements,” *IEEE J Biomed Health Inform*, vol. 26, no. 7, pp. 2974–2982, Jul. 2022, doi: 10.1109/JBHI.2022.3159263.
[13] K. Tsutsumi *et al.*, “A Web-Based Deep Learning Model for Automated Diagnosis of Otoscopic Images,” *Otol Neurotol*, vol. 42, no. 9, pp. e1382–e1388, Oct. 2021, doi: 10.1097/MAO.0000000000003210.
[14] Z. Wang *et al.*, “Structure-aware deep learning for chronic middle ear disease,” *Expert Syst Appl*, vol. 194, May 2022, doi: 10.1016/j.eswa.2022.116519.
[15] J. Zeng *et al.*, “A deep learning approach to the diagnosis of atelectasis and attic retraction pocket in otitis media with effusion using otoscopic images,” *European Archives of Oto-Rhino-Laryngology*, vol. 280, no. 4, pp. 1621–1627, Apr. 2023, doi: 10.1007/s00405-022-07632-z.
[16] A. R. Habib *et al.*, “Evaluating the generalizability of deep learning image classification algorithms to detect middle ear disease using otoscopy,” *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-31921-0.
[17] V. T. Pham, T. T. Tran, P. C. Wang, P. Y. Chen, and M. T. Lo, “EAR-UNet: A deep learning-based approach for segmentation of tympanic membranes from otoscopic images,” *Artif Intell Med*, vol. 115, May 2021, doi: 10.1016/j.artmed.2021.102065.
[18] Z. Cömert, A. Sbrollini, F. Demircan, and L. Burattini, “Computerized otoscopy image-based artificial intelligence model utilizing deep features provided by vision transformer, grid search optimization, and support vector machine for otitis media diagnosis,” *Neural Comput Appl*, vol. 36, no. 36, pp. 23113–23129, Dec. 2024, doi: 10.1007/s00521-024-10457-y.
[19] F. Demircan, M. Ekinci, and Z. Cömert, “Enhancing intra-aural disease classification with attention-based deep learning models,” *Neural Comput Appl*, Jan. 2025, doi: 10.1007/s00521-025-10990-4.
[20] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, 2005, pp. 886–893 vol. 1. doi: 10.1109/CVPR.2005.177.
[21] F. Marini and B. Walczak, “Particle swarm optimization (PSO). A tutorial,” *Chemometrics and Intelligent Laboratory Systems*, vol. 149, pp. 153–165, 2015, doi: https://doi.org/10.1016/j.chemolab.2015.08.020.
[22] S. Mirjalili and A. Lewis, “The Whale Optimization Algorithm,” *Advances in Engineering Software*, vol. 95, pp. 51–67, 2016, doi: https://doi.org/10.1016/j.advengsoft.2016.01.008.
[23] A. Faramarzi, M. Heidarinejad, S. Mirjalili, and A. H. Gandomi, “Marine Predators Algorithm: A nature-inspired metaheuristic,” *Expert Syst Appl*, vol. 152, p. 113377, 2020, doi: https://doi.org/10.1016/j.eswa.2020.113377.
[24] T. T. Nguyen, H. J. Wang, T. K. Dao, J. S. Pan, J. H. Liu, and S. Weng, “An Improved Slime Mold Algorithm and its Application for Optimal Operation of Cascade Hydropower Stations,” *IEEE Access*, vol. 8, pp. 226754–226772, 2020, doi: 10.1109/ACCESS.2020.3045975.
[25] B. ERGEN and M. E. SERTKAYA, “Alzheimer Hastalığının Erken Teşhisinin Çoklu Değişken Kullanarak Tespiti,” *European Journal of Science and Technology*, Mar. 2022, doi: 10.31590/ejosat.1082297.
[26] H. Byun *et al.*, “An assistive role of a machine learning network in diagnosing middle ear diseases,” *J Clin Med*, vol. 10, no. 15, Aug. 2021, doi: 10.3390/jcm10153198.
[27] M. Uçar, K. Akyol, Atila, and E. Uçar, “Classification of Different Tympanic Membrane Conditions Using Fused Deep Hypercolumn Features and Bidirectional LSTM,” *IRBM*, vol. 43, no. 3, pp. 187–197, Jun. 2022, doi: 10.1016/j.irbm.2021.01.001.
[28] D. Livingstone and J. Chau, “Otoscopic diagnosis using computer vision: An automated machine learning approach,” *Laryngoscope*, vol. 130, no. 6, pp. 1408–1413, Jun. 2020, doi: 10.1002/lary.28292.

Authors Contributions
Furkancan Demircan: Conceptualization, Methodology, Writing – Original Draft, Software
Eyüp Gedikli: Conceptualization, Methodology, Writing – Original Draft
Murat Ekinci: Methodology, Writing – review & editing
Zafer Cömert: Methodology, Writing – review & editing

Conflict of Interest Notice
Authors declare no conflict of interest related to this paper.

Ethical Approval and Informed Consent
This article contains no data or other information from studies or experimentation involving human or animal subjects.

Availability of data and material
Availability of data and materials: Data used in this study is openly available and free for research.

Plagiarism Statement
This article has been scanned by iThenticate™.