TÜRKİYE CUMHURİYETİ
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY UNIVERSITY

GRADUATE SCHOOL
COMPUTER ENGINEERING DEPARTMENT

A QUANTITATIVE COMPARISON OF NATURE INSPIRED FEATURE SELECTION ALGORITHMS FOR DIAGNOSIS OF LIVER HCC

BATUHAN ÜNLÜ

M.Sc.

ADANA 2025

TÜRKİYE CUMHURİYETİ
ADANA ALPARSLAN TÜRKEŞ SCIENCE AND TECHNOLOGY UNIVERSITY

GRADUATE SCHOOL
COMPUTER ENGINEERING DEPARTMENT

A QUANTITATIVE COMPARISON OF NATURE INSPIRED FEATURE SELECTION ALGORITHMS FOR DIAGNOSIS OF LIVER HCC

BATUHAN ÜNLÜ

M.Sc.

THESIS ADVISOR
Prof. Dr. Funda R. MERT

ADANA, 2025

DECLARATION OF CONFORMITY

In this thesis study, which was prepared following the thesis writing rules of Adana Alparslan Türkeş Science and Technology University Institute of Graduate School, I declare that I provide all the information, documents, evaluations and results in accordance with scientific ethics and moral codes without resorting to any means or assistance that would be contrary to scientific ethics and traditions. I also declare that I refer to all of the articles I used in this study with appropriate references and accept all moral and legal consequences if a situation is found contrary to my statement regarding my work.

15/08/2025

[Signature]

Batuhan ÜNLÜ

ÖZET

A QUANTITATIVE COMPARISON OF NATURE INSPIRED FEATURE SELECTION ALGORITHMS FOR DIAGNOSIS OF LIVER HCC

Batuhan ÜNLÜ
Yüksek Lisans, Bilgisayar Mühendisliği Anabilim Dalı
Danışman: Prof. Dr. Funda R. MERT
Ağustos 2025, 49 sayfa

Hepatosit karsinomu (HCC), karaciğer kanserinin en yaygın şekli ve dünya çapında kanserle ilişkili ölümlerin dördüncü önde gelen nedenidir. Yüksek ölüm oranının büyük ölçüde sebebi etkili tarama ve erken teşhisteki sınırlamalardır. Bu çalışma, yalnızca tümörlü alanlara odaklanmak yerine tüm karaciğeri içeren ilgili bölgeleri(ROI) kullanarak geleneksel yaklaşımlardan ayrılan radyomik tabanlı bir sınıflandırma çerçevesi sunmaktadır. Bu yaklaşım, HCC'nin karaciğer yapısı üzerinde genel bir etkiye sahip olup olmadığını, genel morfolojisini değiştirip değiştirmediğini incelememizi ve böylelikle sağlıklı bir karaciğerden ayırt edilip edilemediğini anlamamızı sağlar. Radyomik özellikler, hem ham MRI verileri hem de sağlanan yedi filtre kullanılarak karaciğer ROI'si üzerinden çıkarılır. Bu filtreler, çeşitli doku desenlerini yakalamak ve ham MRI verilerinde görünmeyebilecek ince görüntü özelliklerini ortaya çıkarmak için kullanılır. Sınıflandırma performansını iyileştirmek ve alakasız özellikleri ortadan kaldırmak için, doğadan ilham alan dört optimizasyon algoritmasının ikili versiyonları -Afrika Akbabası Optimizasyon Algoritması (AVOA), Yarasa Algoritması (BA), Harris Şahinleri Optimizasyonu (HHO) ve Serçe Arama Algoritması (SSA)- özellik seçimi için uygulanmıştır. Ayrıca segmentasyon görevleri için yaygın olarak kullanılan, herkese açık bir veri setinin sınıflandırma problemleri için de etkili bir şekilde kullanılabileceğini gösteriyoruz. Bu çalışma, tümör maskeleri olmadan bile yüksek sınıflandırma performansına ulaşılabileceğini ve HCC'nin varlığının tüm karaciğerden çıkarılan özellikler kullanılarak doğru bir şekilde tespit edilebileceğini göstermektedir.

Anahtar Kelimeler: Özellik Seçimi, Hepatosit Karsinom, MRI, Doğadan Esinlenen Metasezgisel Algoritmalar, Radyomik

ABSTRACT

A QUANTITATIVE COMPARISON OF NATURE INSPIRED FEATURE SELECTION ALGORITHMS FOR DIAGNOSIS OF LIVER HCC

Batuhan ÜNLÜ
M.Sc., Department of Computer Engineering

Supervisor: Prof. Dr. Funda R. MERT
August 2025, 49 pages

Hepatocellular carcinoma (HCC) is the most prevalent form of liver cancer and the fourth leading cause of cancer-related deaths worldwide. The high mortality rate is largely attributed to limitations in effective screening and early diagnosis. This study presents a radiomics-based classification framework that diverges from conventional approaches by utilizing the entire liver Region Of Interest (ROI) rather than focusing solely on tumor-localized areas. This approach enables us to examine whether HCC has a global effect on the liver structure and alters its overall morphology, thereby allowing it to be distinguished from a healthy liver. Radiomic features are extracted from the liver ROI using both raw MRI data and the seven enhancement filters provided. These filters are employed to capture diverse texture patterns and enhance subtle image characteristics that may not be apparent in raw MRI data. To improve classification performance and eliminate irrelevant features, binary versions of four nature-inspired optimization algorithms—African Vulture Optimization Algorithm (AVOA), Bat Algorithm (BA), Harris Hawks Optimization (HHO), and Sparrow Search Algorithm (SSA)—are applied for feature selection. We demonstrate that a publicly available dataset, commonly used for segmentation tasks, can also be effectively utilized for classification problems. This study demonstrates that high classification performance can be achieved even without tumor masks and the presence of HCC can be accurately detected using features extracted from the whole liver.

Keywords: Feature Selection, Hepatocellular Carcinoma, MRI, Nature-inspired Metaheuristic Algorithms, Radiomics

ACKNOWLEDGEMENTS

This thesis was supported by the Scientific Research Coordination Unit of Adana Alparslan Türkeş Science and Technology University with Project Number: 24303004. It also partially supported by the Scientific and Technological Research Council of Turkey under Project Number 122S588 within the scope of the 1001 Scientific and Technological Research Projects Funding Program.

I would like to express my profound gratitude to my supervisor, Prof. Dr. Funda Raziye MERT, for her invaluable guidance, insightful feedback, and continuous encouragement throughout the preparation of this thesis. Her expertise and support have been instrumental in shaping the direction and quality of this work.

I am also deeply grateful to Prof. Dr. Turgay İBRİKÇİ for his constructive comments, valuable suggestions, and academic guidance, which have significantly contributed to the improvement of this study.

Furthermore, I would like to extend my sincere appreciation to my friends Samet ORAN and Ömer AYANA for their kind assistance, motivation, and unwavering support during the course of this research.

Finally, I would like to express my heartfelt thanks to my beloved parents, Derya ÜNLÜ and Ali ÜNLÜ, for their endless love, patience, and encouragement, which have always been my greatest source of strength and inspiration.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

HCC	: Hepatocellular Carcinoma
CT	: Computed Tomography
MRI	: Magnetic Resonance Imaging
PET	: Positron Emission Tomography
DL	: Deep Learning
CNN	: Convolutional Neural Network
ROI	: Region of Interest
ML	: Machine Learning
AVOA	: African Vulture Optimization Algorithm
BA	: Bat Algorithm
HHO	: Harris Hawk Optimization
SSA	: Sparrow Search algorithm
HH	: Hepatic Haemangioma
DT	: Decision Tree
RF	: Random Forest
k-NN	: k-Nearest Neighbors
LR	: Logistic Regression
AUC	: Area Under the Curve
MI	: Mutual information
LASSO	: Least Absolute Shrinkage and Selection Operation
SVM	: Support Vector Machine
MCC	: Matthews Correlation Coefficient
MVI	: Microvascular Invasion
MLP	: Multi-Layer Perceptron
mRMR	: maximum Relevance-Minimum Redundancy

CDN	: Convolutional Dense Networks
WDBC	: Wisconsin Diagnostic Breast Cancer
RBAVODE	: Relief Binary African Vultures Optimization based on Differential Evolution
EEG	: Electroencephalogram
GA-NN	: Genetic Algorithm-optimized Neural Network
NB	: Naïve Bayes
IBSSA	: Improved Binary Sparrow Search Algorithm
TFSSA	: Tent Lévy Flying Sparrow Search Algorithm
LBP	: Local Binary Pattern
GLCM	: Gray Level Co-occurrence Matrix
3D-CNN	:A 3D convolutional neural network
AOA	:Archimedes Optimization Algorithm ()
LSTM	:Long Short-Term Memory

1. INTRODUCTION

Hepatocellular carcinoma (HCC) is the most common liver cancer, the fourth most common cause of cancer-associated death worldwide and the fastest-growing cause of cancer death in the United States [1], [2]. The main reason for the increased mortality rates associated with HCC is limitations in screening and early diagnosis of HCC and because of these reasons, it is often diagnosed at advanced stages when curative treatment is no longer possible [1], [2]. Also, the weak symptomatic behavior of HCC in the early stages is another reason that makes early diagnosis difficult [1]. The absence of distinct symptoms in the early stages of HCC, combined with the limitations of current diagnostic methods, poses significant challenges for timely detection. This challenge adversely affects patients' quality of life and increases the cost of treatment.

Imaging techniques such as Computed Tomography(CT), Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) are used for liver diagnosis. However, each imaging technique has its own disadvantages, both financially and psychologically. In order to make the tumor area more visible in tomography and PET, patients are injected with a drug. The effect of this substance affects the human body for approximately 2 days to 1 week and shows negative symptoms such as fever, vomiting and kidney problems. In addition, imaging techniques such as tomography and PET are relatively expensive compared to other techniques. However, MRI offers imaging without the need for an additional substance injection [3]. Therefore, MRI is generally used for preliminary and initial diagnosis.

In addition, with the popularity of Deep Learning studies, especially Convolutional Neural Network(CNN) models, the classification of tumors has become easier by extracting features from images. Although CNN obtained remarkable results, it requires big data for feature extraction and model training. However, it is quite difficult to provide such big data in the healthcare field. To overcome this difficulty, a radiomics approach has been proposed, which allows us to obtain features from the Region of Interest(ROI) on the MRI using only mathematical formulas[4]. Radiomics is an approach that allows us to obtain high-dimensional feature sets associated with clinical endpoints from medical images with the help of various mathematical calculations and thus it provides the opportunity to make a quantitative assessment. Moreover, it quantifies subtle changes that cannot be seen. By the way, statistical

analysis and Machine Learning (ML) can be performed on these numbers. Tissue changes can be detected before tumor formation. It helps in early diagnosis and allows us to find hidden patterns in the image [5 – 9].

When we extract features using radiomics, a features set with a very high dimensionality emerges. In this feature set, many features may occur that are irrelevant to the result and therefore negatively affect the classification performance by making it difficult for algorithms to detect existing patterns. It can also increase learning time [10], [11].To overcome this problem, both traditional and nature-inspired feature selection algorithms have been proposed.

Traditional feature selection algorithms select the features with the highest score by measuring only the relationship between the class and the feature with a mathematical formula instead of criteria such as the success rate of the classifier and the relationship between the features[12]. However, nature-inspired algorithms create various subsets of features and obtain a classification result from these sets through classifiers, allowing the selection of the subset of features that provides the highest classification success. In this way, the relationships between the features are also taken into account[13]. This allows nature-inspired algorithms to select more qualified features than traditional algorithms. In this study, we conduct feature selection using binary forms of nature-inspired optimization algorithms to improve classification performance and identify the most relevant subsets.

In addition, we aim to extract the most feasible feature subset using the feature selection coding of nature-inspired optimization algorithms to improve the diagnostic success rate. Four different nature-inspired metaheuristic algorithms such as African Vulture Optimization Algorithm (AVOA), Bat Algorithm (BA), Harris Hawk Optimization (HHO), and Sparrow Search algorithm (SSA) are used to improve the classification performance on the obtained features set and to discard features that do not contribute to the result. At this stage, we will test the feature selection application of powerful nature-inspired optimization algorithms and make a performance comparison. As a result, the best optimization algorithm will be determined and a model that will enable early diagnosis of liver cancer will be proposed using the features produced by this algorithm as output. Among the tested algorithms, AVOA achieved the highest F-score of 0.98. Consequently, it was observed that the study using the entire liver achieved remarkable results. Finally, it has been proven that it is possible to classify whether a liver contains HCC with a high classification performance even without a tumor mask.

As summaries of our contributions, we perform radiomics-based feature extraction on liver MRI slices according to the corresponding liver masks for the detection of HCC in the liver. When both ML, Deap Learning (DL), and radiomics studies are examined, a large number of studies in the literature focus on tumoral lesions. Unlike other studies, a study is performed on features that are derived from ROIs covering the entire liver instead of the tumor ROI for patients with tumors. In this way, we test whether HCC has a general effect on the liver, changing the general structure of the liver and, as a result, whether it becomes distinguishable from a healthy liver.

2. LITERATURE REVIEW

In recent years, numerous studies have been conducted to enhance the detection, classification, and characterization of liver tumors, particularly HCC, through medical imaging techniques such as MRI and CT. These studies have leveraged advanced computational methods, including both conventional ML algorithms and state-of-the-art DL architectures, to improve diagnostic performance and clinical decision-making. The following sections provide a comprehensive overview of the literature, categorized under ML, DL and Radiomics-based approaches.

2.1. Machine Learning Studies

Jingjun Wu et al. [6] studied on 369 patients with 446 lesions, 222 HCC and 224 hepatic haemangioma (HH). They extracted 1029 features for each lesion. Decision Tree (DT), k-Nearest Neighbors (k-NN), Random Forest (RF), and Logistic Regression (LR) were used for the classification of HCC and HH. Based on the obtained results, LR produced the best results, with an area under the curve (AUC) of 0.89 (sensitivity: 0.822; specificity: 0.714).

Xuehu Wang et al. [14] extracted 1316 features for each 196 liver images including 33 cHCC-CC, 88 HCC and 75 CC. They used Mutual information (MI), Chi2-test, F-test, and Least Absolute Shrinkage and Selection Operation (LASSO) methods for feature selection. Classification was performed using Support Vector Machine (SVM) with the selected features. The best results were obtained for accuracy and F-score as 0.89 and 0.88, respectively.

Huai-Wen Zhang et al. [15] extracted 1395 radiomic features from 104 patients with HCC. Feature selection applied with the LASSO. Multiple classifiers such as LR, SVM, RF, AdaBoost, and XGBoost were trained using 5-fold cross-validation. XGBoost achieved the best results with an AUC of 0.9975 (95 % CI: 0.9973–0.9978) and a mean Matthews Correlation Coefficient (MCC) of 0.9369. SVM followed closely with Area Under Curve (AUC) = 0.9846, MCC = 0.9105. Other algorithms also demonstrated strong discrimination (mean AUC range: 0.9644–0.9861).

2.2. Deep Learning Studies

In addition, with the popularity of DL studies, particularly CNN models, the classification of tumors has become easier by extracting features from images [7-9] [16-18].

Zhang et al. [9] conducted a study on 237 patients with HCC. A 3D convolutional neural network (3D-CNN) was developed, including three models trained on single MRI sequences and a fusion model combining all three. The fusion model achieved the best performance, with AUC of 0.81 (sensitivity 69%, specificity 79%) in the training and 0.72 (sensitivity 55%, specificity 81%) in the validation, for predicting Microvascular Invasion (MVI).

Chen Chen et al. [16] studied 444 livers with cancer that included 144 poorly, 168 moderately, and 132 highly differentiated histopathological images. They used five different DL models such as SENet, VGG16, ResNet50, ResNet CBAM, and SKNet for classification. The experimental results showed that the SENet model has achieved the best result with an accuracy of 0.95.

Paula M. Oestmann et al. [17] used 150 lesions (93 HCC and 57 nonHCC) of 118 patients. 3D-CNN was trained on 140 lesions and tested on 10 lesions (5 HCC, 5 non-HCC) for classification. The CNN achieved an overall accuracy of 87.3%, with sensitivities/specificities of 92.7%/82.0% for HCC and 82.0%/92.7% for non-HCC lesions. The AUC was 0.912 .

2.3. Radiomics Approach

Although CNN obtained remarkable results, it requires big data for feature extraction and model training. To address the identified challenge, a radiomic-based feature extraction approach has been proposed to obtain features from the ROI on the MRI [4, 10, 19, 20].

As a sample of radiomics studies, Bing Mao et al. [10] extracted 944 radiomics features for each 114 patients with liver cancer. Feature selection was applied with LASSO. For classification, k-NN, LR, Multi-Layer Perceptron (MLP), RF, and SVM were used. Finally, LR achieved the best result with an accuracy of 0.843 ± 0.078 (AUC, 0.816 ± 0.088; specificity, 0.880 ± 0.117; sensitivity, 0.768 ± 0.232). Huai-wen Zhang et al. [11] extracted 1395 radiomics features for each 104 patients with HCC. LASSO regression was used for feature selection. Classification performance was evaluated by five ML algorithms such as LR, SVM, RF, Xgboost, and Adaboost. The best result was obtained by Xgboost with an AUC of 0.998.

Liu et al. [19] developed a radiomics-based model to predict the histological differentiation of HCC using MRI images from 265 patients and a total of 1,015 radiomic features were extracted. Spearman's rank correlation(correlation coefficient > 0.9), least absolute shrinkage and selection operator (LASSO) and the maximum Relevance-Minimum Redundancy (mRMR)

approach were applied for feature selection. Classification was conducted via SVM algorithms. Among all models, the fused "IntraPeri" model, which combined intra and peritumoral features, achieved the best results with an AUC of 0.95 in the training phase and 0.86 in the validation phase.

Wang et al. [20] studied 200 patients with HCC. Radiomics features were extracted from MR images and reduced via LASSO, yielding the 20 most discriminative features combined with clinical variables to build three predictive models: intratumoral radiomics, clinical-peritumoral radiomics, and clinical-intratumoral radiomics. In the validation set, the clinical-peritumoral radiomics model achieved the highest performance with an AUC of 0.85 (95% CI 0.74–0.95), outperforming the radiomics-only model (AUC = 0.82, 95% CI 0.69–0.93), and the clinical-intratumoral radiomics model (AUC = 0.82, 95% CI 0.68–0.93).

When both ML, DL, and radiomics studies are examined, a large number of studies in the literature focus on tumoral lesions [17, 21, 22, 23]

Jiangfa Li et al. [21] studied 93 patients with histologically confirmed HCC using both contrast-enhanced CT and MRI. A total of 2,153 radiomics features were extracted from intratumoral ROIs. After feature selection using LASSO, selected features combined with clinical variables in a logistic regression model. The model achieved an AUC of 0.916 in the training and 0.816 in the validation for predicting MVI.

Shi et al. [22] studied 342 patients with 449 focal liver lesions using contrast-enhanced CT to differentiate HCC from non-HCC lesions. Three deep learning models based on Convolutional Dense Networks (CDN) were trained using different combinations of CT phases: four-phase (pre-contrast, arterial, portal venous, delayed), and two variations of three-phase protocols (excluding either portal venous or pre-contrast). Model A, which utilized all four CT phases, achieved the best performance with an AUC of 0.925 and an accuracy of 83.3%. Model C (without pre-contrast) achieved similar performance with an AUC of 0.920 and accuracy of 85.6%, showing no statistically significant difference from Model A ($p = 0.765$). In contrast, Model B (without portal venous phase) performed significantly worse with an AUC of 0.862 ($p < 0.05$).

Sato et al. [23] studied 1,582 patients (539 with HCC, 1,043 without HCC) using routinely collected clinical variables and tumor markers. They implemented machine learning such as LR, SVM, RF, gradient boosting, neural network, DL with grid-search hyperparameter tuning.

The gradient boosting model achieved the highest predictive performance with an accuracy of 87.34% and an AUC of 0.94.

2.4. Feature Selection

When we extract features using radiomics, a features set with a very high dimensionality emerges. In this feature set, many features may occur that are irrelevant to the result and therefore negatively affect the classification performance by making it difficult for algorithms to detect existing patterns. It can also increase learning time [24], [25]. Furthermore, recent studies have shown that nature-inspired optimization algorithms combined with wrapper methods significantly enhance classification performance in radiomics-based approaches and these algorithms effectively handle high-dimensional radiomic datasets by selecting the most relevant features and thus, improving diagnostic accuracy and computational efficiency [26-28].

Kamyab Karimi et al. [26] studied the Wisconsin Diagnostic Breast Cancer (WDBC) dataset, which includes 569 samples and 30 features. They proposed two wrapper-based feature selection methods: WFSIC (using Imperialist Competitive Algorithm) and WFSB (using Bat Algorithm). Nine different classifiers were tested: KNN, SVM, DT, Naive Bayes, AdaBoost, LDA, RF, LR, and ANN. Among them, the WFSB + RF combination achieved the best result, with an accuracy of 99.12%, while reducing feature dimensionality by 90%.

Tang et al. [27] studied 253 patients with HCC and other focal liver lesions, using multiphase CT scans. They extracted 688 radiomics features from both original CT and wavelet-transformed domains. A logistic sparsity regression model with Bayesian optimization was used for feature selection and evaluated with sparsity-based LR, MLP, and SVM classifiers. The combined feature set (wavelet + original) with LR achieved the best performance: AUC = 0.96, compared to original-only (AUC = 0.92) and wavelet-only (AUC = 0.90) feature sets.

Dai et al. [28] studied 69 patients with HCC using gadoxetic acid-enhanced MRI to predict MVI. They extracted 167 radiomics features from each MRI sequence and evaluated different gray-level quantization levels, quantization algorithms, classifiers , and feature selection methods. The best performance was achieved using hepatobiliary phase images with a uniform quantization algorithm (64 gray levels), and a two-stage feature selection combining LASSO and GBDT-based recursive feature elimination. The resulting model using the GBDT classifier

achieved the highest performance with an AUC of 0.895, specificity of 82.5%, sensitivity of 93.1%, and accuracy of 87.0%.

Four different nature-inspired metaheuristic algorithms, namely AVOA, BA, HHO and SSA, which have been previously used especially in the health field, were used to improve the classification performance and eliminate features that do not contribute to the result. [29-40]

Gafar et al. [29] addressed the challenge of gene selection for cancer classification in high-dimensional RNA-Seq data by introducing Relief Binary African Vultures Optimization based on Differential Evolution (RBAVO-DE). The method was validated on a dataset with over 20,000 genes across 22 cancer types, including HCC. An initial Relief-based filter reduced the feature space to 500 genes; RBAVO-DE then further selected optimal subsets. The selected genes were evaluated using SVM and k-NN classifiers. RBAVO-DE achieved up to 100% classification accuracy and reduced feature size by ~98% in many datasets, outperforming other binary metaheuristic methods with statistical significance confirmed via Wilcoxon's rank-sum test ($p < 0.05$).

Mostafa et al. [31] developed a novel hybrid optimization algorithm AVOA-AOA that integrates the Archimedes Optimization Algorithm (AOA) into the AVOA to enhance exploration-exploitation balance for both global optimization and feature selection. The hybrid was evaluated on 29 benchmark functions and 19 real-world feature selection datasets. In feature selection using the K-Nearest Neighbors classifier, AVOA-AOA achieved an average accuracy of 97.64% and outperformed state-of-the-art metaheuristics on 15 of 19 datasets.

Pouryosef et al. [32] proposed a novel framework for automated epileptic seizure detection based on Electroencephalogram (EEG) signals, integrating wavelet-based signal decomposition with an advanced feature selection strategy. The study utilized a publicly available EEG dataset comprising 1,200 segments from both healthy individuals and epileptic patients. Initially, multiresolution wavelet decomposition was applied to extract time–frequency domain features. BA was employed for feature selection. The selected feature subset was evaluated using a Genetic Algorithm-optimized Neural Network (GA-NN), alongside benchmark classifiers including k-NN and Naïve Bayes (NB). Experimental results demonstrated that the proposed BA-based model achieved a classification accuracy of 100% on balanced datasets and 75.9% on imbalanced datasets.

Zadsafar et al. [33] studied 324 patients to diagnose malignant mesothelioma. They applied a binary HHO algorithm for wrapper-based feature selection. The CART decision tree classifier was trained using the feature subset selected by HHO algorithm. Model performance evaluated using k-fold cross-validation on a balanced dataset. The model achieved an accuracy greater than 85%, with feature importance validated via Gini impurity analysis.

Mahdi & Yuhaniz [35] conducted a study on two COVID-19 clinical-text datasets comprising 1,446 and 3,053 patient records to develop an intelligent feature selection framework. They proposed Improved Binary Sparrow Search Algorithm (IBSSA), a wrapper-based FS method enhanced with a two-stage approach: in the first stage, RTF-C-IEF term-weighting selected key textual features; in the second stage, IBSSA optimized the feature subset. Selected features were classified using SVM. IBSSA achieved superior results, reducing features by ≈78–84% on average and F1-scores 84.95% and 95.94%, outperforming nine other metaheuristic algorithms.

Yang et al. [36] studied a COVID-19 clinical dataset using wrapper-based feature selection enhanced by a novel Tent Lévy Flying Sparrow Search Algorithm (TFSSA). TFSSA integrates an improved Tent chaotic map, Lévy flight mechanism, and self-adaptive hyperparameters into the Sparrow Search Algorithm to better navigate high-dimensional feature spaces. TFSSA outperforms nine other metaheuristic algorithms on 21 UCI benchmark datasets. Applied to a COVID-19 dataset, TFSSA selected an average of 2.1 features and achieved a classification accuracy of 93.47%.

Sasikala et al. [38] implemented a hybrid BA combined with an Optimum Path Forest classifier for radiomic feature selection in breast ultrasound imaging. The study included 113 patients (62 benign, 51 malignant), from which Local Binary Pattern (LBP) texture features were extracted. The BBA–OPF applied for feature selection. Classification was conducted using SVM. The best result achieved with 93.2% accuracy, 92.3% specificity, 94.4% sensitivity, and F1 score of 91.9%.

Ghadami & Rahebi [39] studied Alzheimer's disease-related MRI images and extracting features using the Gray Level Co-occurrence Matrix (GLCM) and CNNs. They extracted a total of 40 features via GLCM and ResNet from 810 patients. For select most appropriate features, HHO algorithm was applied. The selected features was used in the training phase of MLP neural network and Long Short-Term Memory(LSTM) was used for classify tumors as malignant or

benign. The best result was achieved with accuracy of 97.59%, precision of 97.25%, and sensitivity of 97.41%.

Gad et al. [40] proposed an iBSSA with local search enhancements for feature selection. They tested it on 18 UCI benchmark datasets using k-NN, RF, and SVM classifiers. Compared to 12 other meta-heuristic algorithms, iBSSA consistently achieved up to 92% feature reduction and 100% classification accuracy on some datasets.

3. MATERIAL AND METHOD

This section presents the materials and methods used in a study to classify HCC based on liver MRI images. The proposed approach integrates radiomics-based feature extraction with nature-inspired optimization and machine learning classifiers. First, the characteristics of the dataset used are introduced, followed by a detailed description of the image preprocessing, feature extraction, hyperparameter optimization, feature selection, and classification stages.

3.1. Dataset

In this study, the LiverHccSeg dataset [41] was utilized. The dataset was compiled for research purposes and includes multiphasic MRI scans from 14 patients with HCC and 3 patients without liver tumors. Two experienced radiologists annotated tumor and liver regions. Comprehensive information about the dataset is provided in Table 3.1.

Table 3.1. Dataset information

Parameter	**Tumor Segmentation Cohort**	**Liver Segmentation Cohort**
# of Male	9	11
#of Female	5	6
Age (mean)	60	61
Number of Lesions	16	-
Maximum Tumor Diameter (cm), median	6.34	-
Cumulative Tumor Diameter (cm)	6.34	-
Liver Volume (ccm)	2091.52	1968.94
Total Tumor Volume (ccm)	107.08	-

3.2 Prosed Model

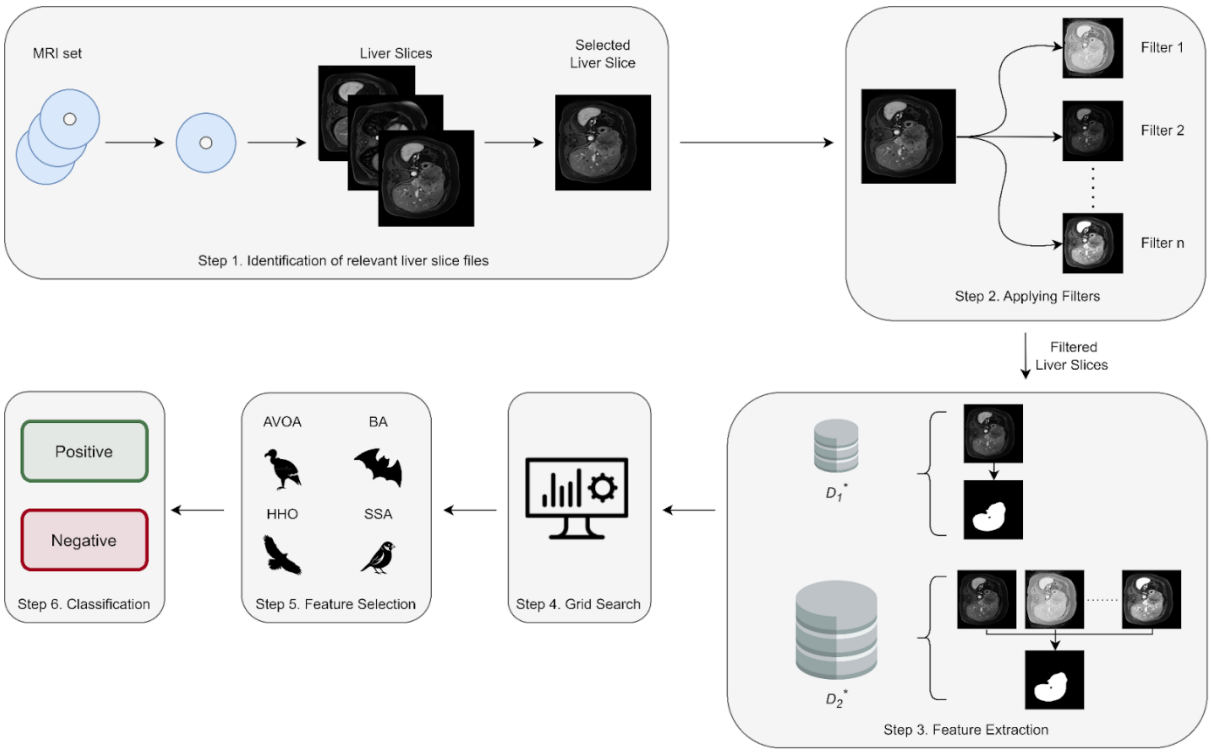


Figure 3.1. Diagram of proposed model

In this section, the detailed steps of our model for classifying whether a liver contains evidence of HCC based on MRI imaging data are explained. As outlined in the introduction section, most existing studies focus solely on tumor regions, restricting the analysis to predefined areas of interest. In contrast, we propose a new model that performs a comprehensive evaluation of the liver MRI image to detect potential evidence of HCC using ROI that encompasses the whole liver rather than just the tumor area. The proposed model demonstrates that a liver affected by HCC can be clearly distinguished from a healthy liver, indicating that HCC alters the fundamental structure of the liver. Our method consists of six main steps: (1) Identification of relevant liver slices, (2) Applying filters to the selected slices, (3) Feature extraction from both raw and filtered images, (4) Hyperparameter optimization using Grid Search, (5) Feature selection using nature-inspired metaheuristic algorithms, (6) Classification of liver images. The

visual representation of the method is given in Figure 3.1.1 Detailed explanations are given in the following sections.

3.2.1. Identification of relevant liver slices files

For HCC classification, liver slices with liver masks from the dataset mentioned in Section 3.1 are used. For each healthy individual, we use liver slices, and the corresponding liver masks are used. For patients with tumors, it is verified whether the ROI is marked for each liver slice. If a marked tumor ROI is present, the liver slice corresponding to the tumor location and the liver mask associated with that slice are used.

3.2.2. Applying filters

In addition to the raw images, seven enhancement filters are applied to the liver slice images to improve the visibility of tumor regions and enhance the performance of the classification algorithms. Filters apply different methods to increase visibility. For example, square multiplies each pixel value by itself. It makes high-density areas brighter. Exponential increases pixel values exponentially. Small differences are greatly magnified. Low values increase rapidly. Even subtle details become more distinct. All the filters used are as follows: Exponential, Gradient, Logarithm, Square, Square Root, Wavelet, and LBP2D. After the filtering process, we generate two distinct datasets: D_1 and D_2, where D_1 contains only the original liver slices, and D_2 comprises both the original and filtered images with their corresponding masks. This approach aims to evaluate whether the application of additional filters to MRI improves the classification performance of HCC detection.

3.2.3. Feature extraction

Radiomics is a field that focuses on extracting numerous quantitative features from medical images [42]. These features capture information related to shape, texture, intensity, and spatial relationships within the images. By transforming images into high-dimensional data, radiomics enables more detailed characterization of tissues and lesions. In the feature extraction stage, separate radiomic features are computed for each image in the datasets D_1 and D_2, resulting in two new feature datasets, D_1^* and D_2^*. Specifically, D_1^* consists of radiomic features extracted exclusively from raw images, whereas D_2^* includes features obtained from both raw and seven

filtered images. Consequently, D_1^* includes 107 features per liver slice image, while D_2^* yields a total of 944 features due to the incorporation of filtered images.

3.2.4. Grid search

In this study, we employ four ML algorithms with distinct working principles to classify liver MRIs:
k-NN, SVM, RF, and NB. In this way, we determine which principle of ML algorithm is more suitable for this study. We chose k-NN according to the distance principle, NB according to the statistical principle, SVM, which allows classifying data with the help of vectors in a spatial plane, and finally RF, which performs classification by looking at the output of more than one model. These algorithms are widely applied in healthcare-related studies [6], [14], [43]. The performances of k-NN, SVM and RF algorithms, which are non-parametric algorithms except NB, depend on the parameters used in the training phase. To optimize classification performance, the grid search approach is used to determine the optimal values of these parameters. The search space and the selected optimal values are presented in Table II within the experimental results section.

3.2.5. Feature selection

In this section, feature extraction is performed on two distinct datasets obtained in the study. Dataset D_1^* contains 107 features, whereas dataset D_2^* comprises 944 features. Features that carry more informative content within these datasets are selected using nature inspired algorithms. Thus, we aim to enhance the classification performance of the algorithm. The reason for preferring nature-inspired algorithms is their ability to overcome certain limitations associated with traditional methods.
Traditional feature selection algorithms select the features with the highest score by measuring only the relationship between the class and the feature with a mathematical formula, instead of criteria such as the success rate of the classifier and the relationship between the features [12]. However, nature inspired algorithms create various subsets of features and obtain a classification result from these sets through classifiers, allowing the selection of the subset of features that provides the highest classification success. In this way, the relationships between

the features are also taken into account [13]. This allows nature-inspired algorithms to select more qualified features than traditional algorithms.

Thereby, AVOA [44], BA [45], HHO [46], and SSA [47], all of which are nature-inspired optimization methods, are employed in this study. Metaheuristic optimization algorithms work with continuous values. These values are converted to the binary solution space using S-shape and V-shape transfer functions to apply feature selection. Detailed information regarding the employed algorithms is provided in Section 4.

3.2.6. Classification of liver images

In this study, a comprehensive classification framework was established to evaluate the diagnostic performance of feature subsets which selected by nature-inspired metaheuristic algorithms. For this reason, four widely used classification algorithms with distinct mathematical foundations—k-NN, SVM, RF, and NB—were employed to classify radiomic features extracted from liver images. Each of these classifiers adopts a different approach to make decision and allows the classification power of the selected features to be assessed from multiple perspectives. For instance, the k-NN algorithm is an instance-based method that performs classification based on neighborhood relationships in the feature space. Random Forest is an ensemble method combined of multiple decision trees and is known for its robustness against overfitting. SVM aims to identify the hyperplane that maximizes the margin between classes, thereby achieving high discriminative power, and is suitable for both linear and non-linear problems. Naive Bayes, on the other hand, is a probabilistic method based on Bayes' Theorem and yields fast and effective results, especially when the assumption of conditional independence among features is reasonably satisfied. By comparatively evaluating the performance metrics such as accuracy, sensitivity, specificity, and F1 score of these four classifiers on feature sets selected by nature-inspired algorithms, it was revealed which classifier is more compatible with which algorithm and the overall performance of the model was analyzed in detail.

3.2.6.1. k-Nearest neighbour

k-NN algorithm is one of the simplest and most widely used classification methods in supervised learning. Introduced by Thomas Cover and Peter Hart [48] in 1967, this algorithm

offers an instance-based approach that does not require training. Its basic principle is to assign a class to an example to be classified using the majority rule, based on the class labels of its k nearest neighbors in the training data. While distance is generally measured using the Euclidean metric, other metrics such as Manhattan and Minkowski can also be used. This method is frequently preferred in fields such as pattern recognition, image analysis, and biomedical classification.

One of the strengths of the k-NN algorithm is that the model requires no training process and can make predictions directly from the training data. However, this also creates a disadvantage: because all training examples must be accessed during testing, computational costs are quite high for large datasets. Furthermore, the algorithm's performance is highly dependent on the selected k value and the distance metric. Smaller k values yield high variance to the model, while larger k values may produce more generalizable but low-precision results. Furthermore, in cases where the class distribution is unbalanced, majority voting may mislead classification. k-NN is preferred due to its ease of application and interpretability, particularly in radiomics-based medical classification studies. Such studies demonstrate that the k-NN algorithm, when used in conjunction with appropriate feature selection and data preprocessing steps, can be an effective tool for complex classification problems in the healthcare field.

3.2.6.2. Support vector machine

SVM is an important classification algorithm, both theoretically and practically, due to its high accuracy and strong generalization capabilities. Developed in the 1990s by Vladimir Vapnik and colleagues [49], SVM primarily aims to find the optimal hyperplane separating examples from different classes in a high-dimensional feature space. This hyperplane is positioned to provide the most reliable separation between classes; that is, the distance between the limiting examples, called support vectors, and the hyperplane is maximized. This approach improves the model's classification performance and reduces the risk of overfitting.

While SVM can be directly applied to linearly separable data, for more complex and nonlinear distributions, the data is transformed into higher-dimensional spaces using kernel functions to make them separable. The most commonly used kernel functions include radial basis function (RBF), linear, sigmoid and polynomial. The RBF kernel is frequently preferred, especially in biomedical applications, and can effectively model inter-class boundaries by calculating the similarity between samples using a Gaussian-like measure.

One of the key hyperparameters that directly affects SVM performance is the C value. This parameter determines the balance between model accuracy and margin width. A small C value provides a wide margin, leading to softer decision boundaries, while a large C value produces a model that fits the training data more tightly (with less error tolerance) and this can lead to overfitting. The second important parameter is the gamma (γ) parameter, specifically used for the RBF kernel. Gamma determines how far an example's influence extends. Lower gamma values produce more general decision boundaries, while higher gamma values can increase the potential risk of overfitting by fitting the training data points more tightly. Therefore, appropriate combinations of the C and gamma parameters are generally determined using methods such as grid search, random search, or Bayesian optimization.

When a polynomial kernel function is used, nonlinear relationships between classes can be modeled with higher-degree polynomials. In this case, one of the prominent hyperparameters is the degree parameter. degree determines the degree of the polynomial kernel; in other words, it controls how complex the interactions between data features can be modeled. A small degree value (e.g., 2 or 3) reflects low-degree interactions, while higher degree values create more complex classification surfaces. However, this complexity can introduce high variance into the model and increase the risk of overfitting. Therefore, the degree parameter is a critical setting that directly affects both the performance and generalization ability of polynomial-kernel SVM models and should generally be optimized through cross-validation. Furthermore, degree is only meaningful for a polynomial kernel; this parameter is ineffective when a linear or RBF kernel is used.

SVM stands out as a powerful classifier, especially in high-dimensional medical datasets with limited samples. Therefore, it is frequently used in radiomics-based hepatocellular carcinoma (HCC) diagnosis studies.

3.2.6.3. Random forest

RF, developed by Leo Breiman [50], is a powerful ensemble learning algorithm based on an ensemble of decision trees. This method is widely used in both classification and regression problems and yields particularly successful results on high-dimensional, complex biomedical data. Random Forest essentially constructs a forest model consisting of many decision trees and trains each tree with a randomly sampled subset of the training data. A randomly selected subset

of features is considered in each node partition. These two fundamental randomization mechanisms reduce the risk of overfitting the model and ensure diversity in predictions.

One of the most important advantages of the RF algorithm is its ability to produce more stable and generalizable results by averaging the errors of individual trees. In classification problems, the model makes decisions based on the majority vote of the trees, while in regression problems, it uses the average value. Furthermore, model accuracy can be estimated internally using observations not included in each tree (out-of-bag data) during training. This feature saves time by eliminating the need for cross-validation. RF offers practical advantages such as the ability to handle missing data, work with categorical and continuous variables, and be tolerant to outliers.

An important feature of the model is its ability to assess the relative importance of features. Metrics such as "mean decrease impurity" based on the Gini index and "permutation importance" obtained by randomly shuffling variable values are used to identify decision-making variables, particularly in radiomics analyses. However, RF models are often considered "black boxes" because explaining and interpreting the decision processes of multiple trees is quite difficult compared to a single decision tree. Therefore, methods such as SHAP values can be used to increase the model's explainability.

The RF algorithm is frequently preferred in medical imaging-based diagnostic systems, providing high accuracy, especially when used in conjunction with radiomics features, in the detection of malignancies such as hepatocellular carcinoma (HCC).

3.2.6.4. Naive bayes

NB is a supervised classification algorithm based on statistical learning and probability theory. This algorithm is widely used in text classification, medical diagnosis, document filtering, and biological data analysis. The foundation of Naive Bayes is Bayes' Theorem. This theorem allows us to calculate the probability that an observation belongs to a particular class based on the observation's features and the prior probability of the class.

The most distinctive aspect of the Naive Bayes algorithm is that it assumes that each feature is conditionally independent of the class. While this is often not perfectly true in practice, the algorithm performs quite well on a large number of real-world datasets. This approach makes the algorithm both simple and computationally efficient. This efficiency, especially in high-dimensional datasets, makes Naive Bayes a valuable tool.

The algorithm calculates the conditional probability of each class for a sample of observations and selects the highest of these probabilities as the predicted class. The basic formula of Bayes' Theorem is as follows:

$$P(C_i \mid X) = (P(X \mid C_i) * P(C_i)) / P(X)$$

- $P(C_i \mid X)$ is the probability that an observation vector X belongs to class C_i.
- $P(X \mid C_i)$ is the probability that X is observed under class C_i.
- $P(C_i)$ is the prior probability of the class.
- P(X) is the total probability of an observation belonging to all classes (constant when comparing across classes).

There are three main variants of the Naive Bayes algorithm:

- Gaussian Naive Bayes – Continuous features are assumed to be normally distributed.
- Multinomial Naive Bayes – Works primarily with word frequencies in text data.
- Bernoulli Naive Bayes – Used when features are binary (0 or 1).

This flexibility allows the NB algorithm to be applied to a wide variety of data types. In medical imaging and radiomics studies, the NB classifier is preferred, particularly for small sample sizes and high-dimensional feature spaces. Its simple structure allows for rapid model training and high interpretability.

4. Optimization Algorithms

This section introduces the optimization algorithms inspired by various natural behaviors, highlighting their underlying mechanisms and position update strategies. The binary versions of these algorithms, utilized for feature selection, are described in detail in Section 4.5.

4.1 African Vulture Optimization Algorithm (AVOA)

AVOA is a nature-inspired metaheuristic optimization approach that imitates the cooperative hunting and scavenging behavior of African vultures [51]. These birds rely on sharp visual perception and strong environmental awareness to efficiently locate carrion. By exploiting rising air currents, vultures can cover vast territories in search of food sources.

4.1.1. Determining the best vultures

First Stage: Identification of Leading Vultures In the initial phase, a population of candidate solutions (vultures) is randomly generated. The fitness of each individual is evaluated, and the two best vultures—one representing the top performer in each subgroup—are selected. Using Equation (1), the remaining vultures update their positions toward these leaders. The fitness values are recalculated after each iteration to determine progress toward the optimal solution.

$$R(i) = \begin{cases} \text{BestVulture}_1 & \text{if } p_i = L_1 \\ \text{BestVulture}_2 & \text{if } p_i = L_2 \end{cases} \quad (1)$$

The probability that each group's vultures move toward the best individuals is determined using Equation (1), while the roulette wheel mechanism in Equation (2) is applied to probabilistically select one of the top vultures in each group.

$$p_i = \frac{F_i}{\sum_{i=1}^{n} F_i} \quad (2)$$

4.1.2. Vulture starvation rate

In nature, well-fed vultures possess high energy levels, enabling them to travel long distances, whereas starving vultures have limited energy and become more aggressive. This behavioral pattern is mathematically modeled using Equation (4), which also governs the transition between exploration and exploitation based on satiety and hunger levels.

$$t = h * \left(\sin^{w}\left(\frac{\pi}{2} * \frac{\text{iteration}_i}{\text{maxiterations}}\right) + \cos\left(\frac{\pi}{2} * \frac{\text{iteration}_i}{\text{maxiterations}}\right) - 1 \right) \quad (4)$$

$$F = (2 * \text{rand}_1 + 1) * z * \left(1 - \frac{\text{iteration}_i}{\text{maxiterations}}\right) + t$$

The overall proportion of vultures involved in exploration decreases with each iteration. When $|F| > 1$, the vultures explore distinct regions (exploration phase). Conversely, when $|F| < 1$, they enter the exploitation phase, focusing on foraging around promising areas.

4.1.3. Exploration

Vultures possess an exceptional ability to detect dying animals even from afar. In the algorithm, this translates into the exploration of new solution spaces. AVOA models exploration through two distinct search strategies, selected probabilistically via the parameter P1, which ranges between 0 and 1. Before each iteration, a random value randP1 $\in [0,1]$ is generated. If randP1 $\geq$ P1, Equation (6) is employed; otherwise, Equation (8) is used. Each vulture searches randomly according to its current energy or saturation level, as expressed in Equation (5).

$$P(i+1) = \begin{cases} \text{Equation}(6) & \text{if } P_1 \geq \text{rand}_{P_1} \\ \text{Equation}(8) & \text{if } P_1 < \text{rand}_{P_1} \end{cases} \quad (5)$$

$$P(i+1) = R(i) - D(i) \times F \quad (6)$$

$$D(i) = |X \times R(i) - P(i)| \quad (7)$$

Equation (6) defines the vultures' random foraging within the vicinity of the two best solutions. Here, P(i+1) denotes the vulture's updated position vector, R(i) is one of the best current vultures, and X is a coefficient vector controlling the intensity of random movements. The coefficient is calculated as $X = 2 \times \text{rand}$, where rand $\in [0,1]$.

$$P(i+1) = R(i) - F + \text{rand}_2 \times \left((ub - lb) \times \text{rand}_3 + lb\right) \quad (8)$$

Equation (8) ensures random generation of solutions within the upper (ub) and lower (lb) variable bounds, while rand_3 is used to increase the randomness coefficient.

4.1.4. Exploitation

If the |F| value is less than 1, AVOA enters the exploitation stage, which has two stages where two different strategies are used in each stage. P_2 and P_3 are used to select strategy. P_2 is used to select the strategies that can be used in the first stage, and P_3 is used to select the strategies that can be used in the second stage. Both parameters must be valued between 0 and 1. When the |F| value is between 1 and 0.5, AVOA enters the first stage of the exploitation stage. In the first stage, two different strategies of roundabout escape and siege-warfare are implemented. P_2 is used to determine the selection of each strategy. If $rand_{P2}$ is greater than or equal to the parameter P_2, the Siege-War strategy is applied slowly. However, if this random number is less than the parameter P_2, the turn-and-escape strategy is applied. This process is shown in Equation (9).

$$P(i+1) = \begin{cases} \text{Equation}(10) & \text{if } P_2 \geq \text{rand}_{P_2} \\ \text{Equation}(13) & \text{if } P_2 < \text{rand}_{P_2} \end{cases} \quad (9)$$

4.1.5. Food competition

When $|F| \geq 0.5$, vultures are moderately full and have sufficient energy to compete for food. As multiple vultures gather at a single carcass, competition emerges—dominant individuals monopolize the food while weaker vultures attempt to exhaust stronger ones and steal portions. This dynamic is described by Equations (10) and (11).

$$P(i+1) = D(i) \times (F + \text{rand}_4) - d(t) \quad (10)$$

$$d(t) = R(i) - P(i) \quad (11)$$

4.1.6. Vulture circular flight

Vultures frequently engage in circular flight patterns, mimicking spiral trajectories during search operations. AVOA models this spiral behavior mathematically through Equations (12) and (13), where sine and cosine functions determine the curvature of movement. The vultures' updated positions are finally computed using Equation (13).

$$S_1 = R(i) \times \left(\frac{\text{rand}_5 \times P(i)}{2\pi} \right) \times \cos\big(P(i)\big) \quad (12)$$

$$S_2 = R(i) \times \left(\frac{rand_6 \times P(i)}{2\pi}\right) \times \sin\left(P(i)\right) \quad (13)$$

$$P(i+1) = R(i) - (S_1 + S_2)$$

4.1.7. Exploitation (Second Stage)

If |F| < 0.5, vultures exhibit intensified competition, aggregating around the best food source. Depending on $rand_{P3}$, if $rand_{P3} \geq P_3$, the vultures gather cooperatively at the food source; otherwise, an aggressive siege-fighting strategy is used, as represented in Equation (14).

$$P(i+1) = \begin{cases} Eq16 & \text{if } P_3 \geq rand_{P_3} \\ Eq17 & \text{if } P_3 < rand_{P_3} \end{cases} \quad (14)$$

Multiple vulture species congregate on a food source:

When multiple vulture species compete for limited resources, their aggregation behavior is expressed by Equations (15) and (16),

$$A_1 = \frac{BestVulture_1(i) - BestVulture_1(i) \times P(i)}{BestVulture_1(i) - P(i)^2} \times F \quad (15)$$

$$A_2 = \frac{BestVulture_2(i) - BestVulture_2(i) \times P(i)}{BestVulture_2(i) - P(i)^2} \times F$$

$$P(i+1) = \frac{A_1 + A_2}{2} \quad (16)$$

where $BestVulture_1(i)$ and $BestVulture_2(i)$ represent the leading vultures from each group. The overall collection and coordination among vultures are defined in Equation (16), involving parameters A_1, A_2, and the updated position vector P(i+1).

4.1.8. Aggresive competition for food

At extremely low energy levels (|F| < 0.5), both leading and follower vultures become desperate and aggressive. They move unpredictably toward the best vulture in each group, as formulated in Equation (17).

$$P(i+1) = R(i) - |d(t)| \times F \times Levy(d) \quad (17)$$

Here, d(t) denotes the distance to the top vulture from one of the best vultures. To improve search efficiency, Levy flight (LF) mechanisms are incorporated into AVOA using Equations (17) and (18),

$$LF(x) = 0.01 \times \frac{u \times \sigma}{|v|^{\frac{1}{\beta}}}, \ \sigma = \left(\frac{\Gamma(1+\beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma(1+\beta)2 \times \beta \times 2^{\frac{\beta-1}{2}}} \right)^{\frac{1}{\beta}} \quad (18)$$

$$LF(x) = 0.01 \times \frac{u \times \sigma}{|v|^{\frac{1}{\beta}}}, \ \sigma = \left(\frac{\Gamma(1+\beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma(1+\beta)2 \times \beta \times 2^{\frac{\beta-1}{2}}} \right)^{\frac{1}{\beta}}$$

where u and v are random variables in [0,1], $\beta = 1.5$ is a fixed constant, and d represents the problem dimension.

4.2 Harris Hawk Optimization Algorithm (HHO)

In this section, the exploration and exploitation mechanisms of the proposed HHO algorithm are modeled based on the behaviors of Harris hawks in detecting, surprising, and attacking their prey [52]. HHO is a population-based and gradient-free metaheuristic, which makes it suitable for solving a wide range of optimization problems once they are properly formulated. Figure 3 illustrates the complete workflow of HHO, while the following subsections detail each stage.

4.2.1. Exploration phase

This phase defines how the exploration process of HHO operates. In the natural environment, Harris hawks rely on their sharp vision to locate prey; however, sometimes the prey may not be visible. In such cases, hawks patiently observe and monitor their surroundings for long periods until they detect potential movement.

In the HHO framework, each hawk represents a candidate solution, and the best individual in each iteration is treated as the target prey (the approximate optimum). The hawks are randomly distributed across the search space and use two distinct perching strategies to detect the prey. Assuming an equal probability q, hawks either perch close to other members of the group and

the prey (for q<0.) or settle on random elevated points within their home range (for q≥0.5), as modeled in Eq. (1).

$$X(t+1) = \begin{cases} X_{rand}(t) - r_1|X_{rand}(t) - 2r_2X(t)| & q \geq 0.5 \\ \left(X_{rabbit}(t) - X_m(t)\right) - r_3\left(LB + r_4(UB - LB)\right) & q < 0.5 \end{cases} \quad (1)$$

Here, X(t+1) denotes the hawk's position in the next iteration, $X_{rabbit}(t)$ is the prey's position, X(t) is the current hawk's location, and $r_1, r_2, r_3, r_4, q \in (0,1)$ are random coefficients updated at each step. LB and UB indicate the lower and upper search bounds, $X_{rand}(t)$ is a randomly chosen hawk, and X_m denotes the mean position of the current population.

A simple model was designed to generate random locations within the range [LB,UB]. The first rule generates new candidate positions relative to random members of the population, while the second rule incorporates the difference between the best-so-far and the population's mean position, scaled by a random factor to enhance stochasticity. The parameter r_3 acts as a scaling coefficient to amplify randomness, particularly when r_4 approaches 1, leading to more diversified exploration. The population's average position is computed using Eq. (2):

$$X_m(t) = \frac{1}{N}\sum_{i=1}^{N} X_i(t) \quad (2)$$

where $X_i(t)$ is the position of hawk i at iteration t, and N is the total number of hawks. This simple averaging approach effectively represents the group's collective position.

4.2.2. Transition from exploration to exploitation

HHO transitions between exploration and exploitation based on the prey's escaping energy. As in nature, prey gradually loses energy while trying to flee. This dynamic is mathematically expressed as:

$$E = 2E_0\left(1 - \frac{t}{T}\right) \quad (3)$$

where E is the prey's escaping energy, T is the maximum number of iterations, and E_0 represents the initial energy state, randomly set within the range (−1, 1). When E_0 decreases toward −1, the prey becomes weaker, whereas an increase toward +1 indicates recovery or strength.

During the optimization process, when |E|≥1, hawks explore new regions in search of potential prey (exploration phase). Conversely, when |E|<1, they exploit the neighborhood of the current

best solutions (exploitation phase). Thus, exploration dominates early iterations, while exploitation gradually becomes prominent as iterations progress.

4.2.3. Exploitation phase

In this phase, Harris hawks perform the so-called "surprise pounce" behavior to capture the prey identified during exploration. However, prey often exhibit escape attempts, resulting in different chasing patterns. Depending on the prey's escape probability r and its remaining energy E, HHO defines four distinct attacking strategies: soft besiege, hard besiege, soft besiege with progressive rapid dives, hard besiege with progressive rapid dives.

4.2.3.1. Soft besiege

When $r \geq 0.5$ and $|E| \geq 0.5$, the prey retains moderate energy and attempts random evasive jumps. During these attempts, hawks softly encircle the prey, exhausting it before executing the final attack. This process is modeled as:

$$X(t+1) = \Delta X(t) - E|X_{rabbit}(t) - X(t)| \quad (4)$$

$$\Delta X(t) = X_{rabbit}(t) - X(t)$$

where $\Delta X(t)$ is the difference between the prey and hawk positions, $r_5 \in (0,1)$ is a random variable, and $J=2(1-r_5)$ represents the random jump strength of the prey.

4.2.3.2. Hard besiege

When $r \geq 0.5$ and $|E| < 0.5$, the prey becomes too weak to escape. Consequently, hawks aggressively tighten the encirclement, updating their positions as in Eq. (6). This behavior enables the algorithm to refine solutions in the vicinity of the current best position.

$$X(t+1) = X_{rabbit}(t) - E|\Delta X(t)| \quad (5)$$

4.2.3.3. Soft besiege with progressive rapid dives

If $|E| \geq 0.5$ but $r < 0.5$, the prey still has sufficient energy to escape, prompting a more adaptive chasing strategy. In this scenario, the algorithm incorporates Levy Flight (LF) patterns to mimic the erratic zigzag motions of prey and the rapid diving maneuvers of hawks.

Levy-based flight patterns are known to represent optimal search strategies in many natural foraging behaviors and have been observed in species such as monkeys and sharks. In HHO,

the LF mechanism allows hawks to perform abrupt and irregular dives, dynamically adjusting their trajectories in response to the prey's unpredictable movements. The LF function is computed as:

$$\mathrm{LF}(x) = 0.01 \times \frac{u \times \sigma}{|v|^{\frac{1}{\beta}}}, \ \sigma = \left(\frac{\Gamma(1+\beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\left(\frac{1+\beta}{2}\right) \times \beta \times 2^{\frac{\beta-1}{2}}} \right)^{\frac{1}{\beta}} \quad (6)$$

where u and v are random values, and β=1.5 is a constant parameter. The final position update in this phase is determined by comparing two potential moves, Y and Z, and selecting the better one.

4.2.3.4. Hard besiege with progressive rapid dives

Finally, when |E|<0.5 and r<0.5, the prey is nearly exhausted, and a hard besiege with rapid dives is initiated. Here, hawks move closer to the prey by reducing the distance between their mean position $X_m(t)$ and the prey's location.

$$X(t+1) = \begin{cases} Y & \text{if } F(Y) < F\big(X(t)\big) \\ Z & \text{if } F(Z) < F\big(X(t)\big) \end{cases} \quad (7)$$

$$Y = X_{\mathrm{rabbit}}(t) - E|X_{\mathrm{rabbit}}(t) - X(t)| \quad (8)$$

$$Z = Y + S \times \mathrm{LF}(D) \quad (9)$$

This process represents the final stage of pursuit, where hawks cooperatively close in on the prey and secure the capture. The sequence of movements and position updates in this step is visually illustrated in Figure 7.

4.3 Bat Algorithm (BAT)

BAT is an optimization technique inspired by the echolocation behavior of microbats. Bats emit ultrasonic sound pulses and interpret the returning echoes to detect objects and prey in their surroundings [53]. The algorithm mimics this echolocation process to balance exploration and exploitation. The algorithm models this behavior using frequency tuning, velocity updates, and adaptive loudness mechanisms. Bats emit sonar pulses and adjust their frequencies to explore the search space effectively. The position update for bats is defined as:

$$X_i^{t+1} = X_i^t + V_i^t \quad (1)$$

$$V_i^{t+1} = V_i^t + (X_{best}^t - X_i^t) * f_i \quad (2)$$

where X_i^t represents the position of the i^{th} bat at iteration t, V_i^t is the velocity of the bat, X_{best}^t is the best-known solution at iteration t, f_i is the frequency controlling the step size. The pulse emission rate and loudness decrease as bats approach optimal solutions, enhancing local search capability. The local refinement is given by:

$$X_i^{t+1} = X_{best}^t + \epsilon * A_i^t \quad (3)$$

where A_i^t is the bat's loudness, is a random number in [-1,1], X_{best}^t is the best solution at iteration t. The loudness A_i^t and pulse rate r_i are updated dynamically using:

$$A_i^{t+1} = \alpha * A_i^t \quad (4)$$

$$r_i^{t+1} = r_i^0 * (1 - e^{-\gamma * t}) \quad (5)$$

where and are control parameters influencing the convergence behavior. Random walks with Levy flights are integrated to improve diversification and prevent stagnation. If a bat is trapped in a local optimum, it performs a random jump based on Levy flights:

$$X_i^{t+1} = X_i^t + \beta * Levy(\lambda) \quad (6)$$

where Levy is a Levy flight distribution, is a scaling factor controlling the step size.

4.4 Sparrow Search Algorithm (SSA)

Sparrows are generally social creatures with numerous species distributed across most regions of the world. They tend to inhabit environments associated with human activity and are omnivorous, primarily consuming grains and weed seeds [54]. These birds are recognized as common residents and, unlike many other small avian species, possess notable intelligence and

a strong memory. Interestingly, house sparrows can be categorized into two behavioral types as producers and scroungers. Producers search for food sources, while scroungers rely on the findings of producers to feed. Several studies have demonstrated that sparrows can flexibly switch between these two strategies depending on environmental conditions. Consequently, it can be inferred that sparrows employ both producing and scrounging tactics in their foraging behavior.

Producers typically possess higher energy reserves and are responsible for locating and guiding others to areas rich in food. Their energy level reflects their fitness value within the population. Sparrows emit alarm calls when predators are detected. If the alarm intensity surpasses a defined safety threshold, producers lead the flock to a secure location. Any sparrow may become a producer if it begins to explore superior food sources, although the overall ratio between producers and scroungers remains constant in the population. Sparrows with greater energy are more likely to act as producers, while hungry scroungers often relocate in search of food to regain energy. Scroungers tend to follow the producer that identifies the most promising feeding area; however, some may compete directly with producers to increase their foraging success. In the event of danger, sparrows on the periphery of the group rapidly move toward safer zones, while those in the center adjust their positions randomly to stay close to others.

In the Sparrow Search Algorithm (SSA), producers with superior fitness values have priority during the search process. Because they are responsible for both exploration and guiding population movement, they can explore a wider search space than scroungers. According to the behavioral rules, during each iteration, the position of a producer is updated using Equation (1),

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^{t} \cdot \exp\left(\dfrac{-i}{\alpha \cdot \text{iter}_{max}}\right) & \text{if } R_2 < ST \\ X_{i,j}^{t} + Q \cdot L & \text{if } R_2 \geq ST \end{cases}$$

where t denotes the current iteration, j = 1, 2, …, d represents the dimension, and $X^{t}_{i,j}$ is the jth dimension value of the ith sparrow. Here, $iter_{max}$ is the maximum number of iterations, $\alpha \in (0, 1]$ is a random coefficient, R_2 and ST denote the alarm value and safety threshold respectively, Q follows a normal distribution, and L is a 1 × d matrix of ones. When $R_2 < ST$, indicating no predators nearby, producers perform global exploration; otherwise, when $R_2 \geq ST$, the sparrows swiftly relocate to a safe region.

Scroungers follow behavioral rules (4) and (5). Some of them continually observe producers, moving toward them immediately after a profitable food source is detected. If they succeed in competition, they obtain the food; otherwise, they keep following producers. The scrounger's position update is represented by Equation (2),

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\dfrac{X_{worst}^{t} - X_{i,j}^{t}}{i^2}\right) & \text{if } i > \dfrac{n}{2} \\ X_{p}^{t+1} + \left|X_{i,j}^{t} - X_{p}^{t+1}\right| \cdot A^{+} \cdot L & \text{otherwise} \end{cases}$$

where X_P denotes the producer's optimal position, Xworst is the global worst position, and A is a $1 \times d$ random matrix with elements of ± 1, while $A^{+} = A^{T}(AA^{T})^{-1}$. When $i > n/2$, the scrounger has a poorer fitness value and is likely starving. Simulation studies assume that 10%–20% of sparrows in the population detect potential threats. Their initial positions are randomly distributed. Following behavioral rule (6), the model is defined by Equation (3).

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^{t} + \beta \cdot \left|X_{i,j}^{t} - X_{best}^{t}\right| & \text{if } f_i > f_g \\ X_{i,j}^{t} + K \cdot \left(\dfrac{\left|X_{i,j}^{t} - X_{worst}^{t}\right|}{(f_i - f_w) + \varepsilon}\right) & \text{if } f_i = f_g \end{cases}$$

In Equation (5), X_{best} represents the global optimal location, β is a normally distributed step-size parameter with a mean of 0 and variance of 1, and $K \in [-1, 1]$ is a random coefficient determining movement direction. f_i, f_g, and f_w refer to the fitness of the current sparrow, the global best, and the global worst individuals, respectively, while ε is a small constant to prevent division by zero. When $f_i > f_g$, the sparrow is assumed to be at the boundary of the group, whereas $f_i = f_g$ indicates that the sparrow resides near the center and should move closer to others for safety. The parameter K controls both the movement direction and step size.

4.5 General Implementation Steps of Optimization Algorithms

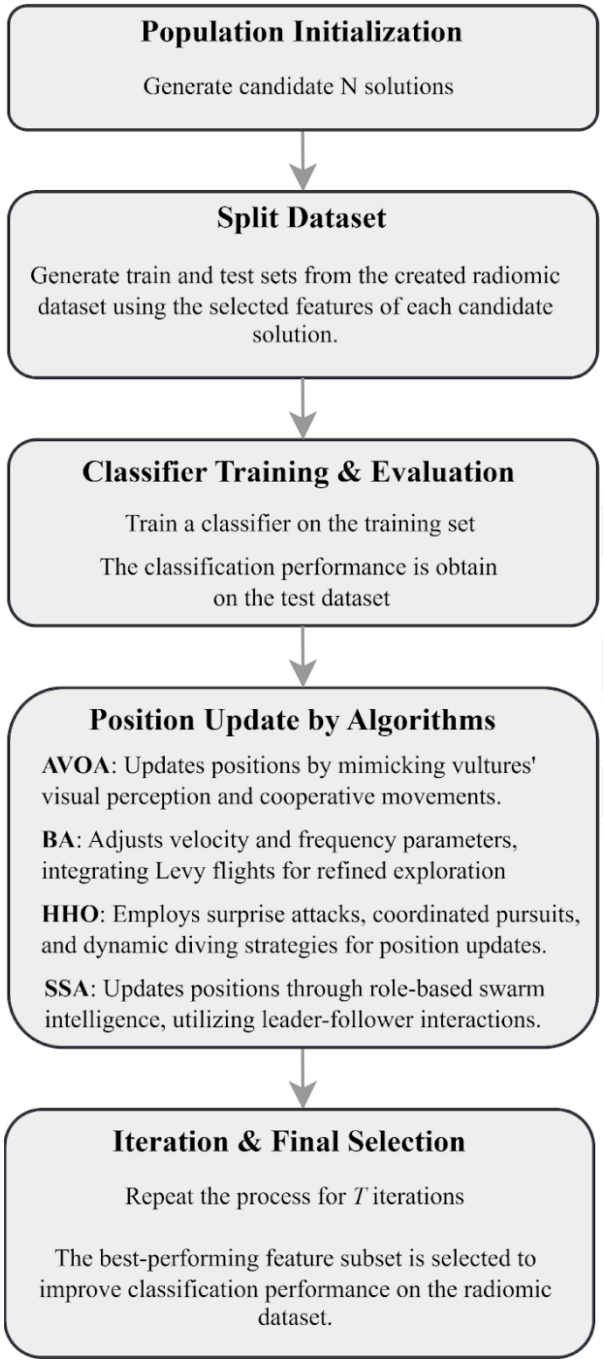


Figure 4.1. Diagram of proposed model

In this section, the steps for applying the AVOA, BA, HHO, and SSA as feature selection methods for the liver HCC dataset are presented. The binary forms of these algorithms are employed in this study for feature selection, as proposed in [40][55 – 57], respectively. The

complete workflow of the proposed feature selection procedure is illustrated in Figure 4.1. The process begins with the initialization of a population comprising N candidate solutions, where each solution corresponds to a subset of features selected from the dataset. The dataset is then partitioned into training and test sets based on the features of each candidate solution. A classifier is trained using the training set, and classification performance on the test set is used to compute the fitness value of each solution. The algorithms iteratively and dynamically update solutions: AVOA exploits elite vulture strategies for selection, BA adjusts velocity and frequency to refine feature subsets, HHO performs surprise pounces or rapid dives to enhance solutions, and SSA employs adaptive movement strategies. This process is repeated for T iterations to obtain the most relevant feature subset, thereby improving the classification performance on the liver HCC dataset.

5. RESULTS AND DISCUSSION

This section presents the experimental results obtained using the proposed model for classifying liver images. All experiments are conducted on the dataset described in Section 3.1. Radiomic features are extracted using PyRadiomics [58]. Consequently, we generate two datasets: D_1^*, consisting of 107 original radiomic features, and D_2^*, consisting of 944 features obtained by applying seven filters in addition to the original features. We compare feature sets to determine the impact of the filter on the classification performance. Four different nature-inspired optimization algorithms for feature selection are applied to D_2^* to reduce the high number of features, improve the classification performance of HCC diagnosis, and shorten the training time.

5.1. Evaluation of Filter Performance

To evaluate the effect of filtered features on classification performance, NB, RF, k-NN, and SVM algorithms are utilized for classification. The hyperparameters of each algorithm are optimized specifically on the D_2^* dataset using a grid search strategy, and the corresponding search spaces along with the selected optimal parameter values are provided in Table II. Using these optimal parameters determined on D_2^*, as shown in Table 5.1, each classifier is subsequently trained and evaluated.

Table 5.1. Search spaces and selected parameter values for each classifier

Classifier	Search Space	Selected Values
k-NN	n neighbors: (1–10)	n neighbors = 4
SVM	C: (1–10) kernel: {linear, poly, rbf, sigmoid, precomputed} degree: (1–5) gamma: {scale, auto}	C = 1 kernel = poly degree = 2 gamma = scale
RF	n estimators: (10–100) max depth: (1–10)	n estimators = 10 max depth = 5

K-fold cross-validation is employed for all machine learning algorithms, with the parameter set to # of fold = 4. The corresponding classification performance results are presented in Table 5.2. Upon examining Table 5.2, it can be observed that combining original features with radiomic features extracted from filtered images yields improved classification performance. As a result, it has been shown that applying filters achieves better results by making the tumor area more visible. Moreover, the RF algorithm outperformed the other classifiers on this combined feature set, achieving an accuracy of 0.93 and an F-score of 0.94. We use the D_2^* dataset and the RF algorithm as the classifier to evaluate each candidate solution generated by the nature-inspired metaheuristic algorithms. The results obtained from these evaluations are presented in the subsequent sections.

Table 5.2. Classification performance of different models using original (D1∗) and filtered (D2∗) feature sets.

Number of Features	Classifier	Accuracy	F-score
Original(107)	RF	0.91	0.93
	SVM	0.70	0.80
	k-NN	0.57	0.62
	NB	0.77	0.83
Original & Filters (944)	RF	0.93	0.94
	SVM	0.72	0.81
	k-NN	0.69	0.74
	NB	0.64	0.62

5.2. Examination of the Impact of Feature Selection

D_2^* contains much more features than D_1^*. Binary forms of nature-inspired metaheuristic algorithms such as AVOA, BA, HHO, and SSA are applied to reduce the number of features, select the best subset, and improve the classification performance. Each dataset (D_1^* and D_2^*) is split into 75% training and 25% testing sets. During training, RF serves as the classifier, and 4-fold cross-validation is used to validate the model's performance. Table 5.3 presents control parameters used for the binary versions of the AVOA, BA, SSA, and HHO algorithms applied to D_2^*.

Table 5.3. Control parameters of the binary variants of AVOA, BA, SSA, and HHO algorithms

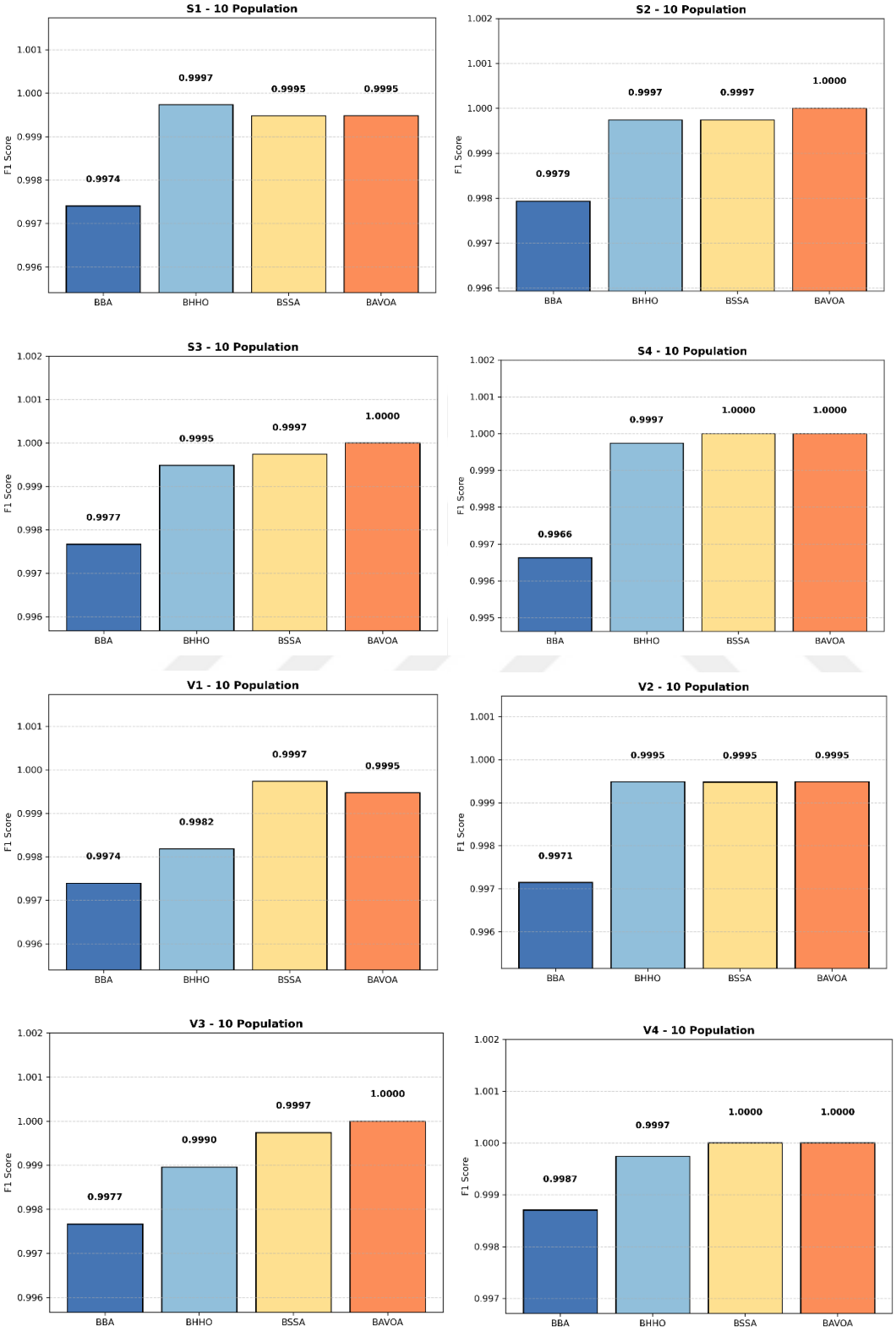
Algorithm	Parameter	Value
AVOA	P1, P2, P3	0.6, 0.4, 0.6
	Alpha (α)	0.8
	Beta (β)	0.2
	Gamma (γ)	2.5
BA	Frequency Range (fmin, fmax)	[0, 10]
	Pulse Rate (r)	0.95
	Loudness (A)	0.8
SSA	Warning Value (ST)	0.8
	Discoverer Proportion (DP)	0.2
	Awareness of Danger Rate (AD)	0.1
HHO	Escape Energy (E)	[−2, 2]
	Jump Strength (J)	[0, 2]

5.3. Examination of Different Population Sizes

Each optimization algorithm was run for 30 independent runs and 100 iterations with populations of 10, 20, and 30, respectively. S-shape and V-shape were used as transfer functions in these runs, and performance comparisons were performed. F-score was used to guide the optimization process and evaluate candidate solutions. The average performance of 30 independent runs is presented in this part.

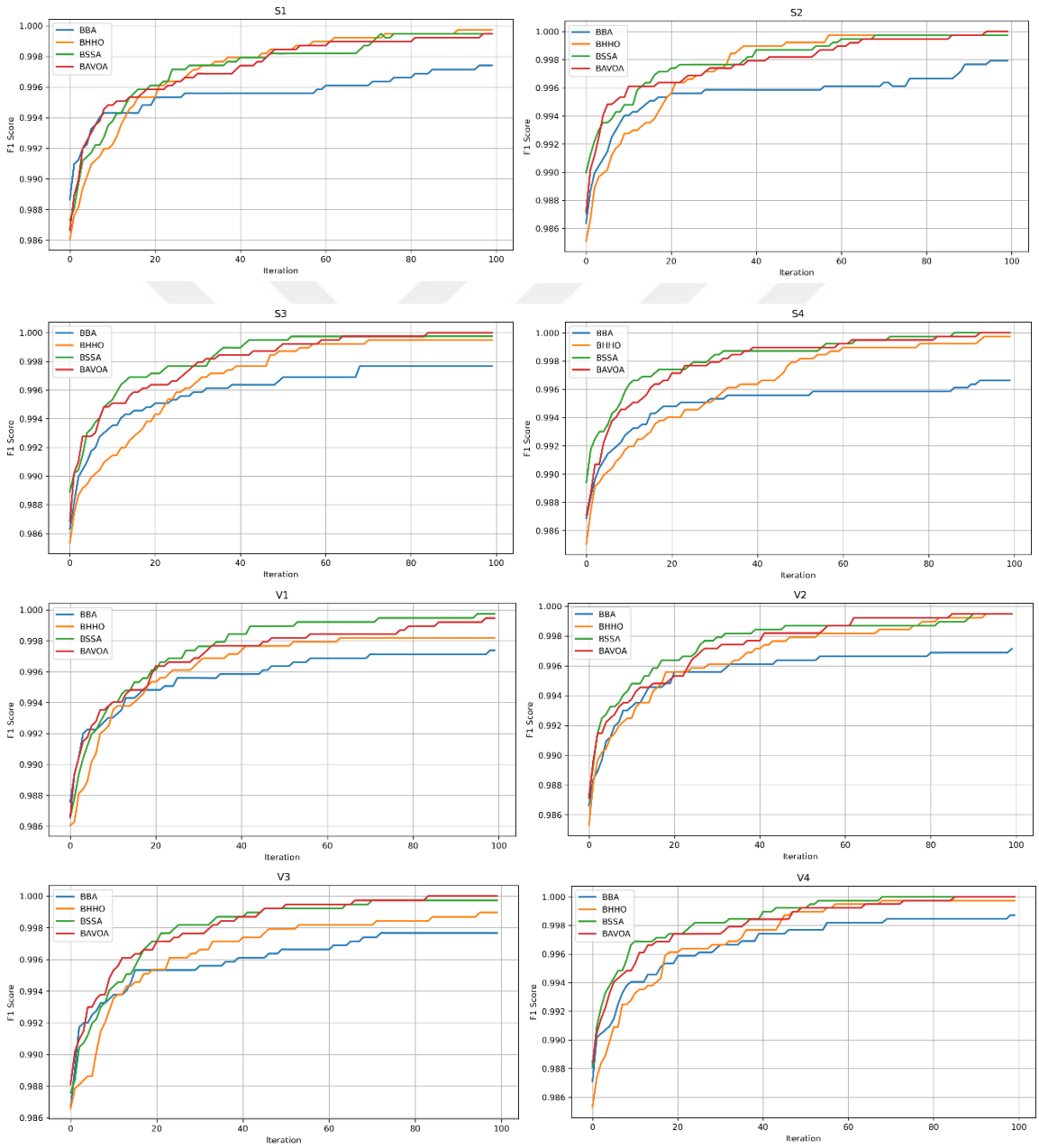
5.3.1. Average performance for 10 population

Figure 5.1. S-Shape and V-Shape F-Score comparison for 10 population



An improvement in F-score values is observed throughout the training process for all optimization algorithms. When examining the F-scores, it is observed that the best results are obtained with BSSA and BAVOA in the S4 and V4 functions. BBA achieved relatively lower results than the other algorithms.

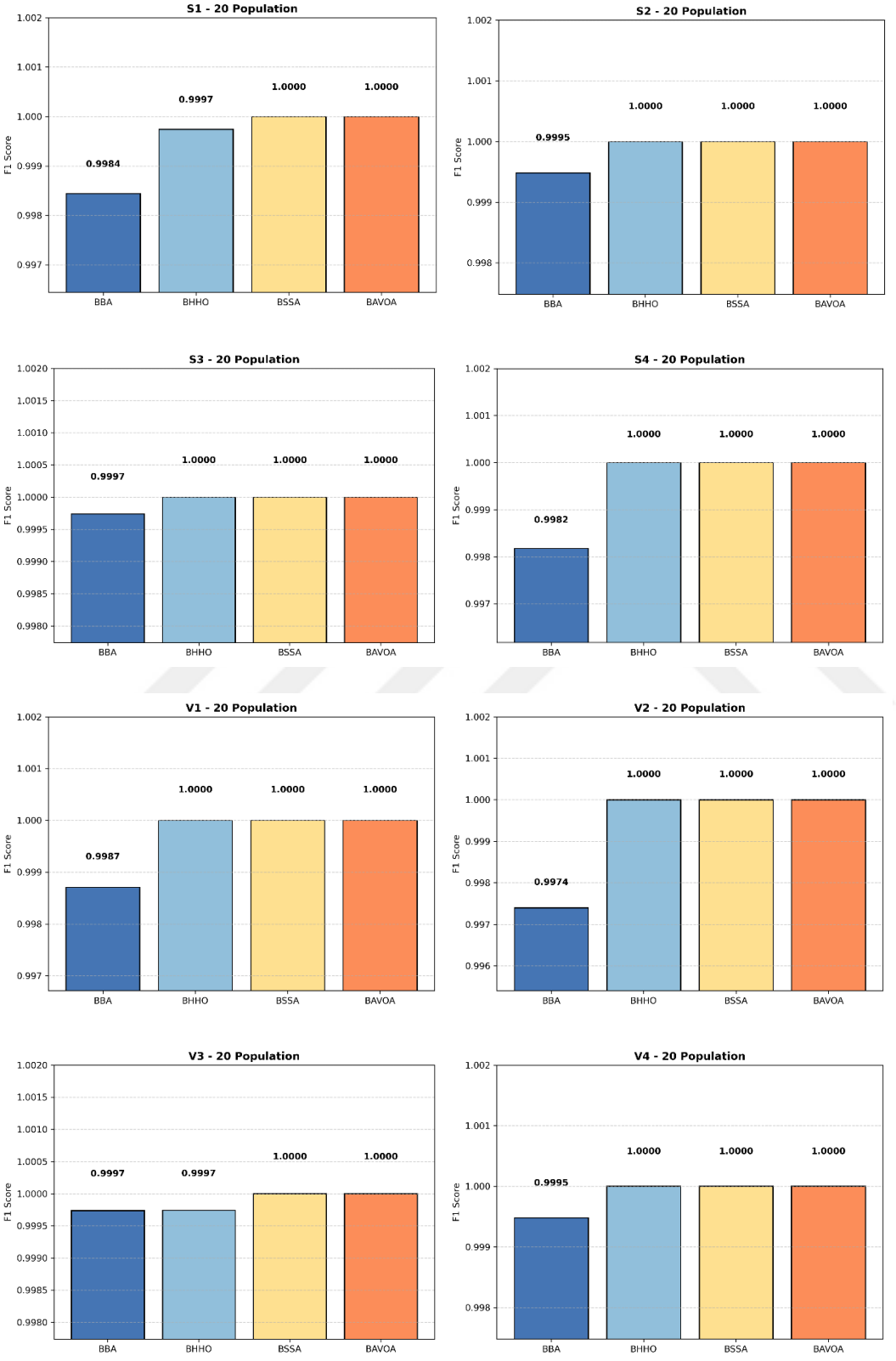
Figure 5.2. Convergence Graph for S-Shape and V-Shape comparison for 10 population



Among these algorithms, AVOA and SSA achieved the highest F-score values. SSA showed a more stable and higher increasing trend compared to AVOA. In contrast, the BA algorithm performed relatively poorly and reached early convergence, indicating limited improvement. These results indicate that AVOA and SSA algorithms are more effective to improve classification performance.

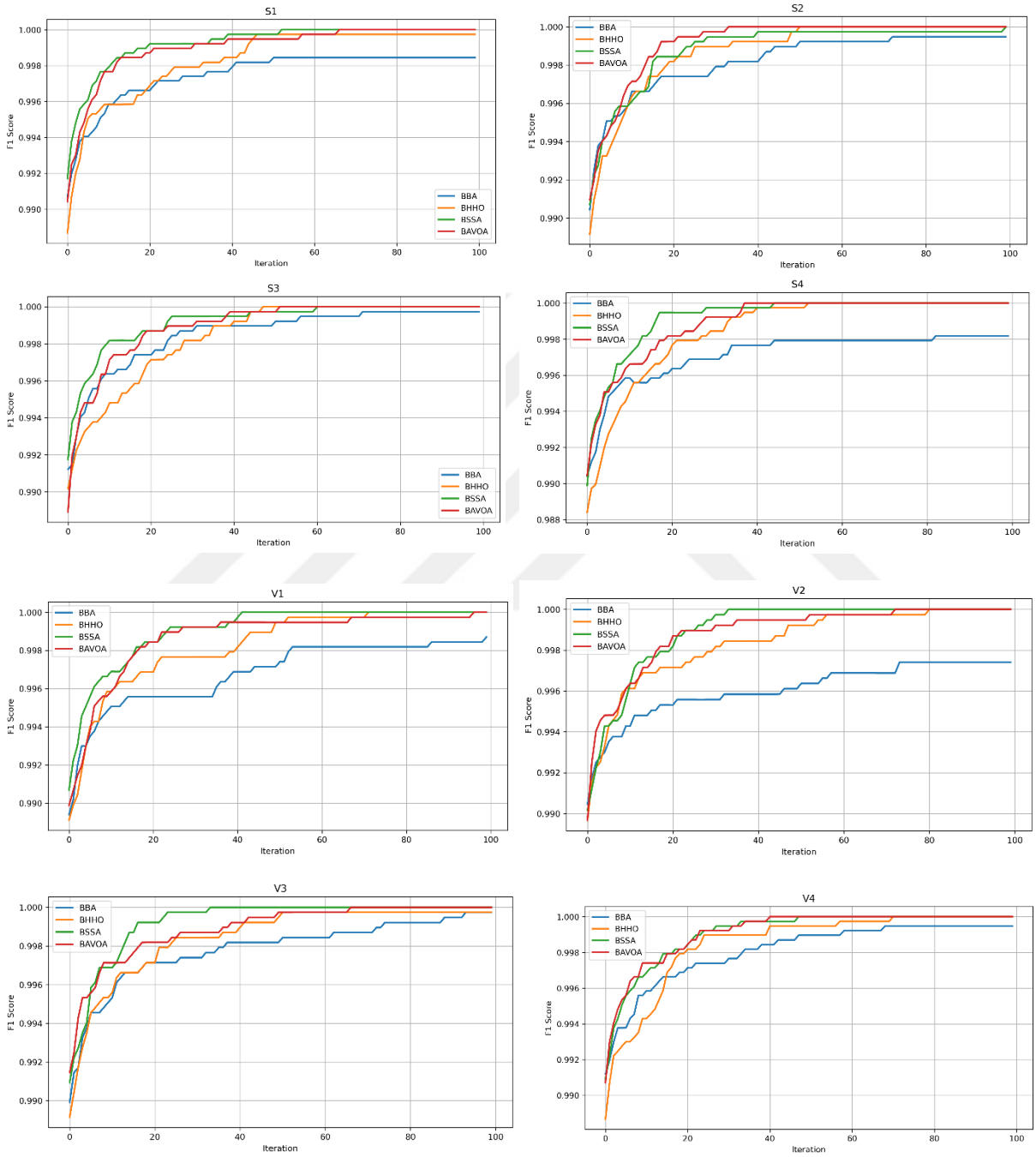
5.3.2. Average performance for 20 population

Figure 5.3. S-Shape and V-Shape F-Score comparison for 20 population



BHHO, BSSA, and BAVOA achieved values of 1 for all transfer functions except S1 and V3. Although BBA yielded relatively lower results, all algorithms generally yielded good results.

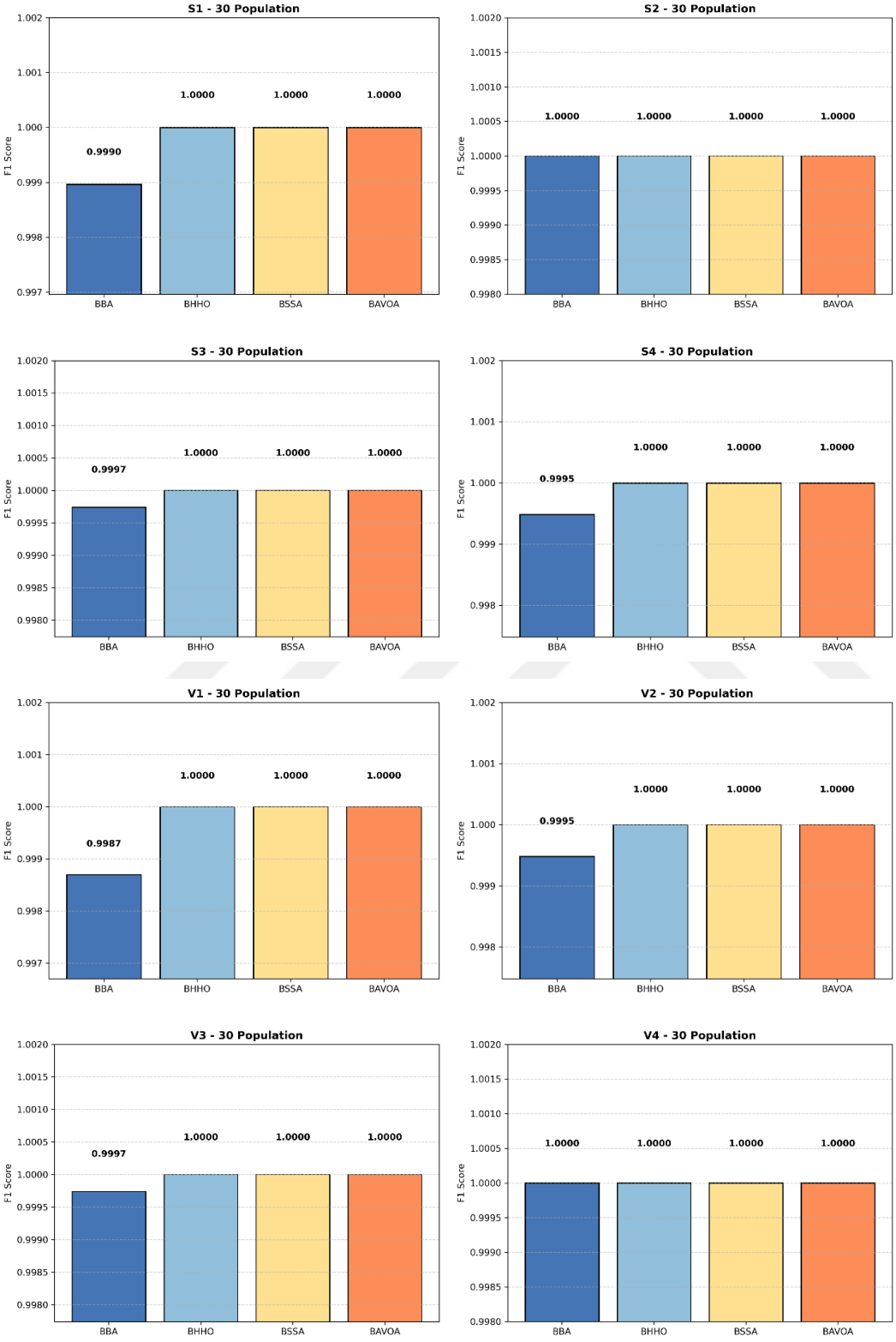
Figure 5.4. Convergence Graph for S-Shape and V-Shape comparion for 20 population



An examination of the results reveals an improvement in F-score values across all optimization algorithms. Among these algorithms, SSA generally achieved the highest F-score values. The HHO and AVOA algorithms showed results similar to SSA. While the BA algorithm showed similar results for some transfer functions, it exhibited relatively lower results. The earliest convergence was achieved with S2 and AVOA. These results demonstrate that SSA algorithms are more effective in improving classification performance.

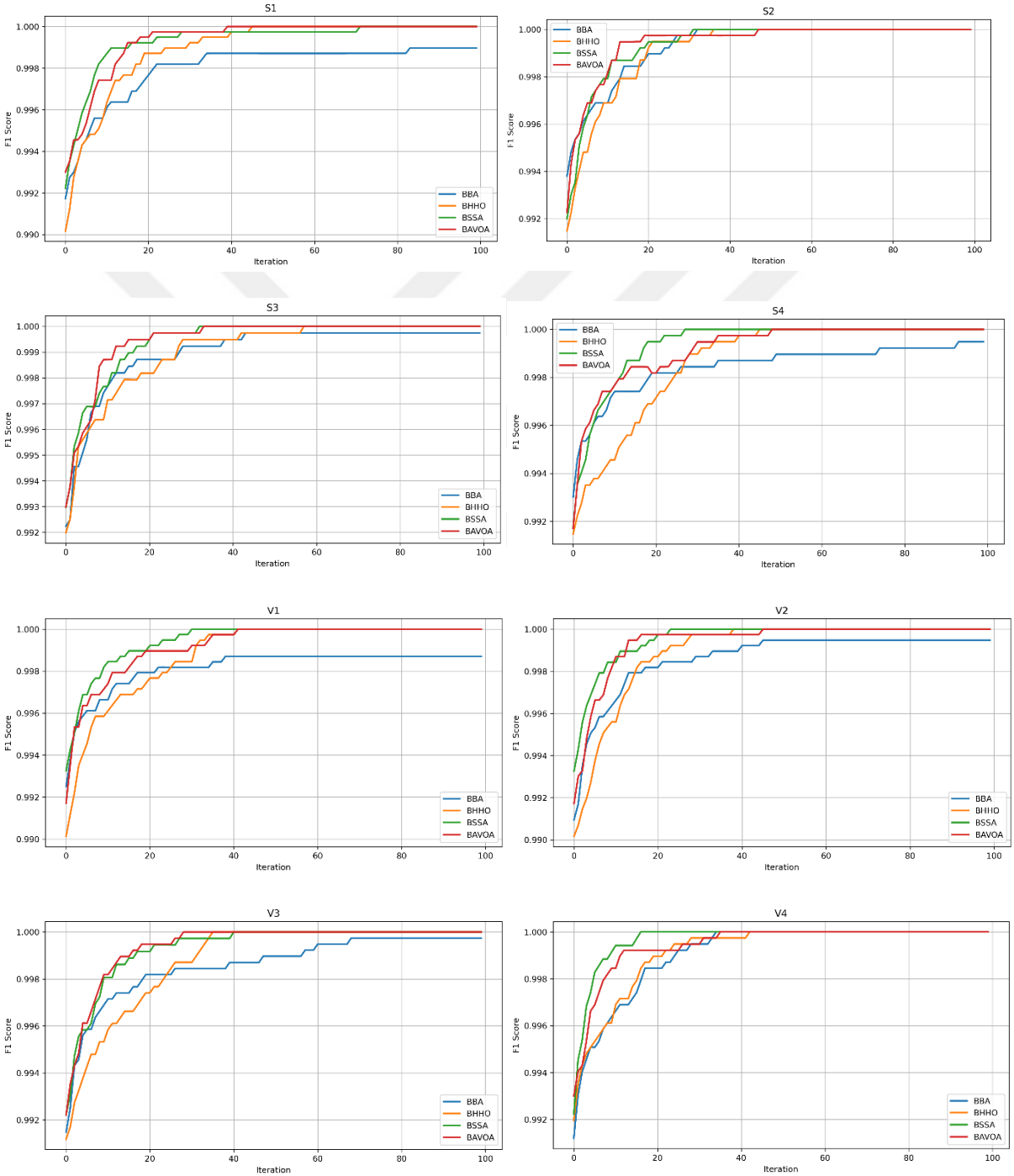
5.3.3. Average performance for 30 population

Figure 5.5. S-Shape and V-Shape F-Score comparison for 30 population



When the results obtained with 30 populations are examined, it is observed that all algorithms with S2 and V4 transfer functions achieve a result of 1. For the other transfer functions, except for Bat, a value of 1 is achieved in all algorithms, and values close to 1 are obtained in BAT.

Figure 5.6. SConvergence Graph for S-Shape and V-Shape comparison for 30 population



With a population of 30, the F-score converged to 1 for all optimization algorithms except BA. Among these algorithms, SSA generally converges earlier. The HHO and AVOA algorithms showed results close to SSA. While all algorithms converged to 1, BA yielded lower results for some transfer functions. These results indicate that, while all algorithms achieved good classification results, the SSA algorithm converged earlier and performed faster.

5.4. Test Results of Selected Features

This part presents the test results of feature selection using nature-inspired metaheuristic algorithms on the D2∗ feature set, which initially contains 944 features.

5.4.1. Test Results for 10 Population

Table 5.4. Test results of optimization algorithms for 10 populations on D_2* feature set using S1 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	67	584
HHO	0.987	0.987	94	551
BAT	0.993	0.993	88	477
SSA	0.987	0.987	49	658

Table 5.5. Test results of optimization algorithms for 10 populations on D_2* feature set using S2 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	50	584
HHO	1	1	113	552
BAT	0.993	0.993	69	476
SSA	0.993	0.993	29	663

Table 5.6. Test results of optimization algorithms for 10 populations on D_2* feature set using S3 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.987	0.987	46	592
HHO	0.993	0.993	367	551
BAT	0.993	0.993	28	478
SSA	0.980	0.980	17	676

Table 5.7. Test results of optimization algorithms for 10 populations on D_2* feature set using S4 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.987	0.987	21	608
HHO	0.980	0.980	103	544
BAT	0.993	0.993	219	468
SSA	0.987	0.987	13	667

Table 5.8. Test results of optimization algorithms for 10 populations on D_2* feature set using V1 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	400	685
HHO	0.993	0.993	435	588
BAT	0.987	0.987	354	480
SSA	0.993	0.993	395	753

Table 5.9. Test results of optimization algorithms for 10 populations on D_2* feature set using V2 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	370	667
HHO	1	1	345	584
BAT	0.993	0.993	397	471
SSA	0.993	0.993	418	743

Table 5.10. Test results of optimization algorithms for 10 populations on D_2* feature set using V3 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	394	675
HHO	0.987	0.987	395	591
BAT	0.993	0.993	368	478
SSA	0.987	0.987	393	745

Table 5.11. Test results of optimization algorithms for 10 populations on D_2* feature set using V4 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	403	652
HHO	0.993	0.993	433	593
BAT	0.993	0.993	355	487
SSA	1	1	418	754

When the test results obtained based on the selected features for the 10 populations were examined, HHO fell slightly behind the other algorithms in terms of the number of selected features, but overall, it achieved better accuracy and F-score values. The fastest algorithm was

BAT, while the slowest was SSA. While the other algorithms perform more aggressive feature selection in some transfer functions, HHO generally achieved better results. It has been observed that the S-shape is more successful than the V-shape in feature selection.

5.4.2. Test Results for 20 Population

Table 5.12. Test results of optimization algorithms for *20* populations on D_{2*} feature set using S1 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.987	0.987	27	1288
HHO	0.993	0.993	130	1131
BAT	0.993	0.993	150	959
SSA	0.987	0.987	33	1378

Table 5.13. Test results of optimization algorithms for *20* populations on D_{2*} feature set using S*2* transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.987	0.987	56	1306
HHO	0.980	0.980	133	1135
BAT	0.993	0.993	58	959
SSA	0.987	0.987	15	1386

Table 5.14. Test results of optimization algorithms for 20 populations on D_2* feature set using S3 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	86	1301
HHO	0.993	0.993	121	1127
BAT	0.980	0.980	118	948
SSA	0.909	0.909	4	1344

Table 5.15. Test results of optimization algorithms for 20 populations on D_2* feature set using S4 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.987	0.987	10	1264
HHO	0.993	0.993	47	1117
BAT	0.974	0.974	22	997
SSA	0.980	0.980	14	1369

Table 5.16. Test results of optimization algorithms for 20 populations on D_2* feature set using V1 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	359	1403
HHO	0.993	0.993	393	1189
BAT	0.993	0.993	364	971
SSA	0.993	0.993	392	1541

Table 5.17. Test results of optimization algorithms for 20 populations on D_2* feature set using V2 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	383	1369
HHO	0.993	0.993	398	1164
BAT	0.993	0.993	399	954
SSA	0.993	0.993	397	1522

Table 5.18. Test results of optimization algorithms for 20 populations on D_2* feature set using V3 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	396	1389
HHO	0.993	0.993	395	1174
BAT	0.980	0.980	342	963
SSA	0.993	0.993	412	1529

Table 5.19. Test results of optimization algorithms for 20 populations on D_2* feature set using V4 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	373	1356
HHO	0.987	0.987	399	1183
BAT	0.987	0.987	342	976
SSA	1	1	399	1541

Generally, all algorithms achieved similar accuracy and F-score values. Although the number of features and the run time is highest, F-score and accuracy values of 1 is achieved using the V4 transfer function and SSA. S-shape's aggressive feature selection did not reduce

performance, on the contrary, similar results were obtained with the other algorithms. In terms of run time, BAT was the fastest algorithm for the 20-population model, as it was for the 10-population model, and SSA was the slowest. While S-shape significantly reduced the number of features by applying more aggressive feature selection, V-shape achieved relatively better accuracy and F-score results.

5.4.3. Test Results for 30 Population

Table 5.20. Test results of optimization algorithms for *3*0 populations on D_2* feature set using S1 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.935	0.935	15	1916
HHO	0.987	0.987	51	1663
BAT	0.987	0.987	22	1459
SSA	0.967	0.967	11	1928

Table 5.21. Test results of optimization algorithms for 30 populations on D_2* feature set using S2 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	1	1	36	1938
HHO	0.980	0.980	190	1685
BAT	0.993	0.993	149	1496
SSA	0.993	0.993	49	1957

Table 5.22. Test results of optimization algorithms for 30 populations on D_2* feature set using S3 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.980	0.980	9	1869
HHO	0.993	0.993	210	1672
BAT	0.877	0.876	7	1493
SSA	0.980	0.980	16	1973

Table 5.23. Test results of optimization algorithms for 30 populations on D_2* feature set using S4 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	18	1945
HHO	0.974	0.974	27	1685
BAT	0.987	0.987	82	1448
SSA	0.961	0.961	7	1925

Table 5.24. Test results of optimization algorithms for 30 populations on D_2* feature set using V1 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.987	0.987	403	2044
HHO	0.993	0.993	407	1782
BAT	0.993	0.993	362	1534
SSA	0.993	0.993	406	2200

Table 5.25. Test results of optimization algorithms for 30 populations on D_2* feature set using V2 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.987	0.987	347	2023
HHO	0.993	0.993	414	1745
BAT	0.993	0.993	376	1514
SSA	0.987	0.987	363	2153

Table 5.26. Test results of optimization algorithms for 30 populations on D_2* feature set using V3 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.987	0.987	386	2046
HHO	1	1	384	1762
BAT	1	1	363	1537
SSA	0.986	0.986	378	2174

Table 5.27. Test results of optimization algorithms for 30 populations on D_2* feature set using V4 transfer function in terms of accuracy, F-score, selected feature size and run time.

Algorithm	Accuracy	F-score	Feature size	Run Time (s)
AVOA	0.993	0.993	373	2115
HHO	0.993	0.993	381	1793
BAT	0.987	0.987	354	1555
SSA	0.995	0.995	380	2245

Algorithms that performed aggressive feature selection by selecting 7 out of 944 features in the S3 and S4 transfer functions yielded poorer results than the other algorithms. The best results were obtained with the S2 transfer function and AVOA, while the worst results were obtained

with the S3 transfer function and the BAT algorithm. HHO performed worse than the other algorithms when reducing the number of features. For the V-shape, the HHO and BAT algorithms generally yielded better results than the others.

When examining studies conducted with all populations, the S-shape exhibited more aggressive feature selection than the V-shape. In some cases, this aggressive feature selection slightly decreased the performance compared to others. This demonstrates that aggressive feature selection is not superior in all cases. The accuracy and F-score values obtained with the V-shape are relatively better. Overall, the slowest algorithm in all studies was SSA, and the fastest was BAT. While it was the fastest algorithm, BAT achieved the lowest overall results. Although studies conducted with 30 populations minimized the number of features for the S-shape, its running time was significantly higher than the other populations.

6. CONCLUSION

In this study, a novel radiomics-based classification approach was proposed to detect HCC using liver MRI scans. By extracting features from the entire liver region rather than focusing solely on tumor-localized areas, the method enables the identification of global structural changes associated with HCC. The use of image filtering techniques contributed to capturing more diverse and subtle texture features, which in turn enhanced classification performance. To optimize the model further, four nature-inspired binary optimization algorithms were employed for feature selection. Binary versions of these algorithms were used to perform feature selection using S-shape and V-shape transfer functions. The results demonstrate that accurate classification of HCC can be achieved without relying on tumor masks, underscoring the potential of full-liver radiomic analysis in non-invasive liver cancer diagnosis.

REFERENCES

[1] T. Wakabayashi, F. Ouhmich, C. Gonzalez-Cabrera, E. Felli, A. Saviano, V. Agnus, P. Savadjiev, T. F. Baumert, P. Pessaux, and J. Marescaux, "Radiomics in hepatocellular carcinoma: a quantitative review," *Hepatology International*, vol. 13, pp. 546–559, 2019.

[2] E. Harding-Theobald, J. Louissaint, B. Maraj, E. Cuaresma, W. Townsend, M. Mendiratta-Lala, A. G. Singal, G. L. Su, A. S. Lok, and N. D. Parikh, "Systematic review: radiomics for the diagnosis and prognosis of hepatocellular carcinoma," *Alimentary Pharmacology & Therapeutics*, vol. 54, no. 7, pp. 890–901, 2021.

[3] S. Hussain, I. Mubeen, N. Ullah, S. S. U. D. Shah, B. A. Khan, M. Zahoor, and M. A. Sultan, "Modern diagnostic imaging technique applications and risk factors in the medical field: a review," *BioMed Research International*, vol. 2022, Art. no. 5164970, 2022.

[4] R. J. Gillies and M. B. Schabath, "Radiomics improves cancer screening and early detection," *Cancer Epidemiology, Biomarkers & Prevention*, vol. 29, no. 12, pp. 2556–2567, 2020.

[5] M. Illimoottil and D. Ginat, "Recent advances in deep learning and medical imaging for head and neck cancer treatment: MRI, CT, and PET scans," *Cancers*, vol. 15, no. 13, Art. no. 3267, 2023.

[6] J. Wu, A. Liu, J. Cui, A. Chen, Q. Song, and L. Xie, "Radiomics-based classification of hepatocellular carcinoma and hepatic haemangioma on precontrast magnetic resonance images," *BMC Medical Imaging*, vol. 19, no. 1, pp. 1–11, 2019.

[7] J. C. Ahn, T. A. Qureshi, A. G. Singal, D. Li, and J. D. Yang, "Deep learning in hepatocellular carcinoma: current status and future perspectives," *World Journal of Hepatology*, vol. 13, no. 12, pp. 2039–2049, 2021.

[8] G. N. Ioannou, W. Tang, L. A. Beste, M. A. Tincopa, G. L. Su, T. Van, E. B. Tapper, A. G. Singal, J. Zhu, and A. K. Waljee, "Assessment of a deep learning model to predict hepatocellular carcinoma in patients with hepatitis C cirrhosis," *JAMA Network Open*, vol. 3, no. 9, e2015626, 2020.

[9] Y. Zhang, X. Lv, J. Qiu, B. Zhang, L. Zhang, J. Fang, *et al.*, "Deep learning with 3D convolutional neural network for non-invasive prediction of microvascular invasion in hepatocellular carcinoma," *Journal of Magnetic Resonance Imaging*, vol. 54, no. 1, pp. 134–143, 2021, doi: 10.1002/jmri.27538.

[10] B. Mao, J. Ma, S. Duan, Y. Xia, Y. Tao, and L. Zhang, "Preoperative classification of primary and metastatic liver cancer via machine learning-based ultrasound radiomics," *European Radiology*, vol. 31, pp. 4576–4586, 2021.

[11] H. Zhang, D. Huang, Y. Wang, H. Zhong, and H. Pang, "CT radiomics based on different machine learning models for classifying gross tumor volume and normal liver tissue in hepatocellular carcinoma," *Cancer Imaging*, vol. 24, no. 1, Art. no. 20, 2024.

[12] S. Solorio-Fernández, J. A. Carrasco-Ochoa, and J. F. Martínez-Trinidad, "A review of unsupervised feature selection methods," *Artificial Intelligence Review*, vol. 53, no. 2, pp. 907–948, 2020.

[13] M. Sharma and P. Kaur, "A comprehensive analysis of nature-inspired meta-heuristic techniques for feature selection problem," *Archives of Computational Methods in Engineering*, vol. 28, pp. 1103–1127, 2021.

[14] X. Wang, S. Wang, X. Yin, and Y. Zheng, "MRI-based radiomics distinguish different pathological types of hepatocellular carcinoma," *Computers in Biology and Medicine*, vol. 141, Art. no. 105058, 2022.

[15] H.-W. Zhang, D.-L. Huang, Y.-R. Wang, H.-S. Zhong, H.-W. Pang, *et al.*, "CT radiomics based on different machine learning models for classifying gross tumor volume and normal liver tissue in hepatocellular carcinoma," *Cancer Imaging*, vol. 24, no. 1, Art. no. 20, 2024, doi: 10.1186/s40644-024-00652-4.

[16] K. A. Tran, O. Kondrashova, A. Bradley, E. D. Williams, J. V. Pearson, and N. Waddell, "Deep learning in cancer diagnosis, prognosis and treatment selection," *Genome Medicine*, vol. 13, pp. 1–17, 2021.

[17] P. M. Oestmann, C. J. Wang, L. J. Savic, C. A. Hamm, S. Stark, I. Schobert, B. Gebauer, T. Schlachter, M. Lin, and J. C. Weinreb, "Deep learning–assisted differentiation of pathologically proven atypical and typical hepatocellular carcinoma (HCC) versus non-HCC on contrast-enhanced MRI of the liver," *European Radiology*, vol. 31, pp. 4981–4990, 2021.

[18] C. Chen, C. Chen, M. Ma, X. Ma, X. Lv, X. Dong, Z. Yan, M. Zhu, and J. Chen, "Classification of multi-differentiated liver cancer pathological images based on deep learning attention mechanism," *BMC Medical Informatics and Decision Making*, vol. 22, no. 1, Art. no. 176, 2022.

[19] H. F. Liu, M. Wang, Q. Wang, Y. Lu, Y. J. Lu, Y. Sheng, *et al.*, "Multiparametric MRI-based intratumoral and peritumoral radiomics for predicting the pathological differentiation of hepatocellular carcinoma," *Insights into Imaging*, vol. 15, no. 1, Art. no. 97, 2024.

[20] K. D. Wang, M. J. Guan, Z. Y. Bao, Z. J. Shi, H. H. Tong, Z. Q. Xiao, *et al.*, "Radiomics analysis based on dynamic contrast-enhanced MRI for predicting early recurrence after hepatectomy in hepatocellular carcinoma patients," *Scientific Reports*, vol. 15, no. 1, Art. no. 22240, 2025.

[21] J. Li, W. Song, J. Li, L. Cai, Z. Jiang, M. Wei, *et al.*, "A clinical study exploring the prediction of microvascular invasion in hepatocellular carcinoma through the use of combined enhanced CT and MRI radiomics," *PloS One*, vol. 20, no. 1, e0318232, 2025.

[22] W. Shi, S. Kuang, S. Cao, B. Hu, S. Xie, S. Chen, *et al.*, "Deep learning assisted differentiation of hepatocellular carcinoma from focal liver lesions: choice of four-phase and three-phase CT imaging protocol," *Abdominal Radiology*, vol. 45, pp. 2688–2697, 2020.

[23] M. Sato, K. Morimoto, S. Kajihara, R. Tateishi, S. Shiina, K. Koike, and Y. Yatomi, "Machine-learning approach for the development of a novel predictive model for the diagnosis of hepatocellular carcinoma," *Scientific Reports*, vol. 9, no. 1, Art. no. 7704, 2019.

[24] W. Zhang, Y. Guo, and Q. Jin, "Radiomics and its feature selection: a review," *Symmetry*, vol. 15, no. 10, Art. no. 1834, 2023.

[25] G. Ge and J. Zhang, "Feature selection methods and predictive models in CT lung cancer radiomics," *Journal of Applied Clinical Medical Physics*, vol. 24, no. 1, e13869, 2023.

[26] K. Karimi, A. Ghodratnama, and R. Tavakkoli-Moghaddam, "Two new feature selection methods based on learn-heuristic techniques for breast cancer prediction: a comprehensive analysis," *Annals of Operations Research*, vol. 328, no. 1, pp. 665–700, 2023.

[27] V. H. Tang, S. T. Duong, C. D. T. Nguyen, T. M. Huynh, V. T. Duc, C. Phan, et al., "Wavelet radiomics features from multiphase CT images for screening hepatocellular carcinoma: analysis and comparison," Scientific Reports, vol. 13, no. 1, Art. no. 19559, 2023.

[28] H. Dai, M. Lu, B. Huang, M. Tang, T. Pang, B. Liao, et al., "Considerable effects of imaging sequences, feature extraction, feature selection, and classifiers on radiomics-based prediction of microvascular invasion in hepatocellular carcinoma using magnetic resonance imaging," Quantitative Imaging in Medicine and Surgery, vol. 11, no. 5, pp. 1836–1853, 2021.

[29] M. G. Gafar, A. A. Abohany, A. E. Elkhouli, and A. A. A. El-Mageed, "Optimization of gene selection for cancer classification in high-dimensional data using an improved African vultures algorithm," Algorithms, vol. 17, no. 8, Art. no. 342, 2024.

[30] W. L. Cheng, K. M. Ang, S. S. Tiang, K. Y. Yap, L. Pan, C. H. Wong, et al., "Feature selection of medical dataset using African vultures optimization algorithm," in Advances in Intelligent Manufacturing and Mechatronics, Singapore: Springer Nature, 2023, pp. 175–185.

[31] R. R. Mostafa, F. A. Hashim, N. E. El-Attar, and A. M. Khedr, "Empowering African vultures optimizer using Archimedes optimization algorithm for maximum efficiency for global optimization and feature selection," Evolving Systems, vol. 15, no. 5, pp. 1701–1731, 2024.

[32] M. Pouryosef, R. Abedini-Nassab, and S. M. R. Akrami, "A novel framework for epileptic seizure detection using electroencephalogram signals based on the bat feature selection algorithm," Neuroscience, vol. 541, pp. 35–49, 2024.

[33] F. Zadsafar, H. Tabrizchi, S. Parvizpour, J. Razmara, and S. Lotfi, "A model for mesothelioma cancer diagnosis based on feature selection using Harris hawk optimization

algorithm," Computer Methods and Programs in Biomedicine Update, vol. 2, Art. no. 100078, 2022.

[34] H. Al-Safi, J. Munilla, and J. Rahebi, "Patient privacy in smart cities by blockchain technology and feature selection with Harris Hawks Optimization (HHO) algorithm and machine learning," Multimedia Tools and Applications, vol. 81, no. 6, pp. 8719–8743, 2022.

[35] A. Y. Mahdi and S. S. Yuhaniz, "An intelligent feature selection approach based on a novel improve binary Sparrow Search Algorithm for COVID-19 classification," International Journal of Intelligent Engineering & Systems, vol. 16, no. 4, 2023.

[36] Q. Yang, Y. Gao, and Y. Song, "A tent Lévy flying sparrow search algorithm for wrapper-based feature selection: A COVID-19 case study," Symmetry, vol. 15, no. 2, Art. no. 316, 2023.

[37] C. Shanthini and S. S. Silvia Priscila, "Sparrow Search Algorithm (SSA) based feature selection and optimized feed-forward neural network (OFFNN) for semen quality analysis," in International Conference on Computing and Co*mmunication Networks*, Singapore: Springer, 2025, pp. 669–682.

[38] S. Sasikala, M. Ezhilarasi, and S. Arunkumar, "Feature selection algorithm based on binary BAT algorithm and optimum path forest classifier for breast cancer detection using both echographic and elastographic mode ultrasound images," Journal of Cancer Research and Therapeutics, vol. 19, no. 2, pp. 191–197, 2023.

[39] R. Ghadami and J. Rahebi, "Alzheimer's prediction methods with Harris Hawks Optimization (HHO) and deep learning-based approach using an MLP-LSTM hybrid network," Diagnostics, vol. 15, no. 3, Art. no. 377, 2025.

[40] A. G. Gad, K. M. Sallam, R. K. Chakrabortty, M. J. Ryan, and A. A. Abohany, "An improved binary sparrow search algorithm for feature selection in data classification," Neural Computing and Applications, vol. 34, no. 18, pp. 15705–15752, 2022.

[41] M. Gross, S. Arora, S. Huber, A. S. Kücükkaya, and J. A. Onofrey, "LiverHccSeg: A publicly available multiphasic MRI dataset with liver and HCC tumor segmentations and inter-rater agreement analysis," Data in Brief, vol. 51, Art. no. 109662, 2023.

[42] P. Lambin, R. T. Leijenaar, T. M. Deist, J. Peerlings, E. E. De Jong, J. Van Timmeren, et al., "Radiomics: the bridge between medical imaging and personalized medicine," Nature Reviews Clinical Oncology, vol. 14, no. 12, pp. 749–762, 2017.

[43] S. Gil-Rojas, M. Suárez, P. Martínez-Blanco, A. M. Torres, N. Martínez-García, P. Blasco, et al., "Prognostic impact of metabolic syndrome and steatotic liver disease in hepatocellular carcinoma using machine learning techniques," Metabolites, vol. 14, no. 6, Art. no. 305, 2024.

[44] B. Abdollahzadeh, F. S. Gharehchopogh, and S. Mirjalili, "African vultures optimization algorithm: A new nature-inspired metaheuristic algorithm for global optimization problems," Computers & Industrial Engineering, vol. 158, Art. no. 107408, 2021.

[45] X. S. Yang, "A new metaheuristic bat-inspired algorithm," in Nature Inspired Cooperative Strategies for Optimization (NICSO 2010), Berlin, Heidelberg: Springer, 2010, pp. 65–74.

[46] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," Future Generation Computer Systems, vol. 97, pp. 849–872, 2019.

[47] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: Sparrow Search Algorithm," Systems Science & Control Engineering, vol. 8, no. 1, pp. 22–34, 2020.

[48] T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21–27, 1967, doi: 10.1109/TIT.1967.1053964.

[49] V. N. Vapnik, The Nature of Statistical Learning Theory, Springer, 1995, doi: 10.1007/978-1-4757-2440-0.

[50] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[51] B. Abdollahzadeh, F. S. Gharehchopogh, and S. Mirjalili, “African vultures optimization algorithm: A new nature-inspired metaheuristic algorithm for global optimization problems,” Computers & Industrial Engineering, vol. 158, p. 107408, 2021.

[52] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, “Harris hawks optimization: Algorithm and applications,” Future Generation Computer Systems, vol. 97, pp. 849–872, 2019.

[53] S. Mirjalili, S. M. Mirjalili, and X. S. Yang, “Binary bat algorithm,” Neural Computing and Applications, vol. 25, no. 3, pp. 663–681, 2014.

[54] J. Xue and B. Shen, “A novel swarm intelligence optimization approach: Sparrow search algorithm,” Systems Science & Control Engineering, vol. 8, no. 1, pp. 22–34, 2020.

[55] A. Atef, "BAVO: Binary African Vulture Optimization Implementation," GitHub repository, 2025. [Online]. Available: https://github.com/D-Amr-Atef/Gene-Selection-RBAVO-DE/blob/main/models/swarm%20based/BAVO.py

[56] A. Atef, "BBA: Binary Bat Algorithm Implementation," GitHub repository, 2025. [Online]. Available: https://github.com/D-Amr-Atef/Gene-Selection-RBAVO-DE/blob/main/models/swarm%20based/BBA.py

[57] A. Atef, "BHHO: Binary Harris Hawk Optimization Implementation," GitHub repository, 2025. [Online]. Available: https://github.com/D-Amr-Atef/Gene-Selection-RBAVO-DE/blob/main/models/swarm%20based/BHHO.py

[58] J. J. M. Van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R. G. H. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, and H. J. W. L. Aerts, "Computational radiomics system to decode the radiographic phenotype," Cancer Research, vol. 77, no. 21, pp. e104–e107, 2017.