Adaptive Feature Selection for Brain Tumor Classification in MRI Images using Genetic Algorithm Polar Bear Optimization and SVM

B. Meena Preethi[1], J.Lekha[2], A.Seethalakshmy[3], S.Gokul[4]

[1]Assistant Professor, Department of Software Systems Sri Krishna Arts and Science College,Coimbatore, India, Email: meenapreethib@skasc.ac.in
[2]Associate Professor, Data Science Department Christ (Deemed to be University) Pune - Lavasa CampusHub of Analytics, India, Email: lekha.j@christuniversity.in
[3]Head, Dept of Psychology, Rathinam College of Arts and Science Coimbatore Tamilnadu, Email: Seethalakshmypersonal@gmail.com
[4]Student of MSc Software Systems Sri Krishna Arts and Science College Coimbatore, Tamilnadu

Received: 17.04.2024 Revised : 15.05.2024 Accepted: 26.05.2024

ABSTRACT

Brain tumor detection is a critical task in medical imaging, and the integration of advanced technologies has significantly improved the accuracy and efficiency of diagnosis. Accurate identification of tumor types from MRI scans is essential for effective treatment planning and patient management. The proposed method aims to optimize the feature selection process to enhance classification performance while minimizing computational complexity. Initially, a diverse set of features extracted from MRI images is subjected to the GA-PBO hybrid algorithm, which intelligently selects the most discriminative features relevant to brain tumor classification. The GA-PBO algorithm leverages the exploration-exploitation capabilities of genetic algorithms and the powerful search mechanisms inspired by polar bear behavior to efficiently identify the optimal feature subset. Subsequently, the selected features are fed into an SVM classifier to perform tumor classification based on their distinctive patterns and characteristics. SVM is well-suited for handling high-dimensional data and has been widely employed in medical image analysis tasks due to its robustness and effectiveness in classification. Experimental evaluations are conducted using publicly available MRI datasets containing various types of brain tumors. Evaluation metrics, such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC), are used to assess the efficacy of the proposed method. The performance of the proposed approach is compared against current feature selection methods and classification techniques. The results show that the proposed adaptive feature selection approach achieves superior classification performance compared to baseline methods, effectively distinguishing between different tumor types with high accuracy and reliability. Furthermore, the computational efficiency of the proposed method is demonstrated, making it suitable for real-time applications in clinical settings. Overall, the proposed GA-PBO-based adaptive feature selection approach, integrated with SVM classification, offers a promising solution for accurate and efficient brain tumor classification in MRI images, thereby facilitating timely diagnosis and treatment planning for patients with brain tumors.

Keywords: Brain Tumor, GA-PBO, Genetic Algorithm, Polar Bear Optimizations, Support Vector Mechanism, AUC-ROC.

1. INTRODUCTION

With the use of cutting-edge medical imaging techniques like Magnetic Resonance Imaging (MRI), we can see inside the human body more clearly than ever before. But things get complicated when it comes to comprehending and categorizing brain tumors. Due to their complexity and diversity, brain tumors must be correctly classified to determine the most effective course of treatment. This is a difficult task due to the vast amount of data from medical images. There are so many details to consider that traditional methods find it difficult to handle all this information.

This paper, with a keen understanding of the complexities inherent in brain tumor classification, takes on the formidable challenge of devising an intelligent system capable of discerning the most critical details from brain images. What sets this investigation apart is its inspiration drawn from the intricate workings of nature. It ingeniously combines the power of Genetic Algorithms (GAs), which draw insights from nature's evolutionary processes, and Polar Bear Optimization (PBO), a technique that mirrors the collaborative hunting behaviour observed in polar bears. GAs, akin to nature's meticulous evolutionary mechanisms, excel in exploring a myriad of possibilities to identify the most optimal features. In parallel, PBO introduces a nuanced layer of adaptability and randomness, injecting a higher level of intelligence into the system.

The primary objective here is to synergize these two nature-inspired methodologies, culminating in the creation of an intelligent system adept at extracting essential features from MRI images. However, the ambition of this paper extends beyond this fusion. It takes a bold step forward by introducing Support Vector Machine (SVM) classification, a powerful tool that assumes the role of the cognitive center within our intelligent system. SVM operates with the precision of a highly skilled detective, leveraging the crucial features selected by GAs and PBO to discern the specific type of brain tumor present. This integration not only aids in identifying pivotal details but also ensures a level of accuracy in classification that is indispensable in the realm of medical diagnosis.

The proposed approach stands poised to usher in a new era of precision in brain tumor detection. Its potential to reduce diagnostic uncertainties, expedite treatment decisions, and ultimately improve patient outcomes underscores the transformative impact of technology in the realm of healthcare. As the research community embraces the convergence of artificial intelligence and medical diagnostics, this innovative fusion of GA, SVM and PBO emerges as a beacon of hope in the ongoing battle against brain tumors, promising a future where early detection becomes synonymous with enhanced survival rates and improved quality of life.

2. LITERATURE REVIEW

Tiejun Wu et al., proposed a comprehensive idea on a genetic algorithm-based peer selection optimization technique that considers both peers uploading bandwidth and peer availability of material for effective content distribution in BitTorrent networks. To speed up obtaining optimal solutions in peer selection, GA-BT utilizes the divisible load theory to forecast optimal fitness values dynamically. The usefulness of GA-BT in peer selection optimization is demonstrated by the experimental findings, which are obtained using a BitTorrent simulator that is used to assess GA-BT performance.

K. F. Man et al., presented an idea of how genetic algorithms (GA) are presented as a whole, allowing industrial engineers to build a design tool framework by integrating existing knowledge about this new technology. Additionally, an effort has been made to clarify "why" and "when" GA should be employed as an optimization technique.

Namira Begum Sayyed conducts a brief study on genetic algorithms and its approach on Traveling Salesman problem and demonstrates the application of genetic algorithms in several industries and shows how they may be combined with various other strategies to find the optimal answer and expedite information retrieval systems. This detailed study has offered with consideration for genetic optimization, which utilizes genetic algorithm, and it is validated using genetic optimization software on a well-known Traveling Salesman Problem. Concepts from evolutionary algorithms applied to the traveling salesman problem.

David Polap et al., has presented an optimization algorithm which was inspired by nature, which is named as Polar Bear Optimization Algorithm (PBO). This proposed a new mathematical model of how polar bears eat and hunt could be a valuable optimization method with various theoretical and practical problems and the polar bear behaviour model is a search engine for optimal solutions. The birth-death dynamic dominates the population, however the suggested simulated adaptation to harsh winter circumstances offers the benefit of local and global search. The suggested PBO was assessed and juxtaposed with alternative metaheuristic algorithms through the use of sample test functions and some traditional design challenges. The outcomes of the experimental investigation were analyzed using various parameters and compared with other algorithms. The primary benefits were identified by the analysis as being the associated population's quick recognition of the area and the efficient birth and death process, which enhances both local and global search capabilities in the solution space.

The CEED problem is solved by Saqib Fayyaz et al. using a novel enhanced polar bear optimization (IPBO) method in conjunction with polar bear optimization (PBO) and chaotic population-based polar bear optimization algorithms. This study proposes a new, enhanced PBO version that enhances its functionality by adding tracking capabilities that capitalize on polar bears' keen sense of smell. The validity of the new IPBO is evaluated using five benchmark functions and a 140-element Korean ED problem. The IEEE-3-unit and 6-unit CEED problems, enhanced PBO (IPBO), and PBO are all resolved using the suggested chaotic population-based PBO. The cumulative density function, probability density function, and Wilcoxon rank sum test (WRST) are statistical tools used to analyze the outcomes of each technique. The strength of each solution technique and its capacity to reduce costs and emissions in the process of solving the CEED problem are highlighted by statistical and numerical analysis of the results.

Dawid et al., proposed an article that demonstrated the Polar Bear Optimization. The population is managed by a system of birth and death, while the research-simulated adaptation to hard winter conditions improves local and global search. PBO's benefits, including quick population detection and effective birth and death mechanisms, were assessed, and contrasted with those of other meta-heuristic algorithms.

Amila Amatullah et al., aimed for the lowest Power Peaking actor (PPF) value without sacrificing operating duration, and presented a model that used the Polar Bear Optimization Algorithm to identify an ideal fuel loading strategy based on BEAVRS. The PIJ module, CITATION X-Y module, and Standard Reactor Analysis Code was employed. The optimization process involved 200 populations and 50 iterations; the best optimum fuel loading pattern had a maximum PPF value of 1.458 and a keff of 0.916.

Mirkhan et al., proposed a model to demonstrate how feature selection problems in scientific research can be resolved using the metaheuristic algorithm Polar Bear Optimization (PBO). Rough set theory offers a useful method for identifying redundant features that can be eliminated without losing any important information. However, attempting to explore every possible combination of features will lead to an NP-hard problem. PBO uses an mechanism for birth and death, which allows researchers to invest in promising solutions and dismiss hopeless ones. Tests of the model on multiple datasets demonstrated that superfluous data may be eliminated without causing data loss.

Theodoros Evgeniou et al., suggested a model for condensing the workshop contents, and as part of the Advanced Course on Artificial Intelligence (ACAI '99) in Greece, a workshop on Support Vector Machines (SVM) Theory and Applications was held.

An overview of the state of knowledge, underlying theory of SVM, and issues raised during the workshop are provided. Among the useful applications of SVM are time series prediction, medical diagnosis, and face recognition.

Yongli Zhang proposed an idea about the support vector machine, a new machine learning method based on statistical learning theory, and its basic theory, classification idea, and current algorithm. This paper provides a thorough review of support vector machines' potential in the machine learning field by discussing real-world issues, their efficacy, and their future possibilities in classification applications.

Soumick Chatterjee et al., carries out an extensive investigation into brain tumors, including their forms and features. Both models were found to perform better than ResNet18, a pure 3D convolutional model. It was also shown that pre-training the models on a different or unrelated dataset before using them for the cancer classification job enhances performance. The Pre-trained ResNet Mixed Convolution model, which required the fewest processing resources and yielded a macro F1-score of 0.9345 and a test accuracy of 96.98%, was ultimately found to perform better in these tests than the other models.

Samah A. Gamel et al.,proposed a brain tumor classification model based on CNN (BCM-CNN). It is the hyperparameter optimization of CNN using adaptive dynamic sine cosine fitness gray wolf optimization algorithm (ADSCFGWO).After optimizing the hyperparameters, a training model created with Inception-ResnetV2 is used. The model's output is a binary 0 or 1 (0: normal, 1: tumor), and it makes use of pre-trained models that are frequently used to enhance the diagnosis of brain tumors. The two primary categories of hyperparameters that are involved are those that specify the underlying network topology and those that are in charge of network training. The gray wolf and blue cosine algorithms are the foundations of the ADSCFGWO algorithm, which is an adaptable framework that makes use of each approach's advantages. According to the experimental results, BCM-CNN performed the best as a classifier since its performance was enhanced by the CNN optimization hyperparameters. Using the BRaTS 2021 Task 1 dataset, BCM-CNN obtained 99.98% accuracy.

Milica M. Badžaoffers a novel CNN architecture for classifying three different types of brain cancers. T1-weighted contrast-enhanced magnetic resonance imaging was used to assess the developed network, which is less complex than the pretrained networks in use today. Network performance was evaluated using two databases and two combinations of two 10-fold cross-validation approaches. We evaluated the network's capacity for generalization using a single 10-fold process called subject-wise cross-validation, and we tested its improvement using a larger image collection. Using the 10-fold cross-validation approach, record-level cross-validation of the expanded data set yielded the best results; in this case, the accuracy was 96.56%. The recently developed CNN architecture's excellent generalization ability and speed make it an effective decision support tool for radiologists doing medical diagnostics.

Yong-Kui Ma et al. offer a novel method for incorporating image improvement strategies into their model for brain tumor classification, such as adaptive histogram equalization using CLAHE and Gaussian blur-based sharpening. With accurate comparisons against well-established pretrained models, their approach yields remarkable 97.84% accuracy rates.

Emrah Irmak suggests a CNN-based strategy for early brain tumor diagnosis, employing three distinct classification tasks. Their models achieve high accuracies, notably 99.33% for tumor detection, 92.66% for multi-class tumor classification, and 98.14% for grading, while also outperforming several state-of-the-art CNN architectures.

Belayneh Sisay Alemu et al. propose a comprehensive method for brain tumor classification, incorporating denoising, segmentation, feature extraction, and support vector machine (SVM) classification. Their approach achieves a remarkable accuracy of 99.9% using a dataset of MRI images, including both malignant and benign brain tumors for training and testing. Compared to previous methods, their approach significantly improves the ability to classify brain tumor images. This comprehensive method represents a promising advancement in brain tumor classification, with potential applications in medical diagnosis and treatment planning.

Lei Wang et al., utilize a hybrid GA-SVM algorithm for brain tumor classification, leveraging genetic algorithms (GA) for feature selection and parameter optimization in support vector machine (SVM) classification. Their approach achieves superior classification accuracy compared to conventional methods, demonstrating the effectiveness of the hybrid GA-SVM approach. By selecting optimal feature subsets and optimizing SVM classifier parameters, their algorithm achieves improved classification accuracy, offering potential benefits for clinical applications in brain tumor diagnosis and prognosis.

Wojciech Książek et al., suggest a hybrid model (NCA-GA-SVM) that combines Neighborhood Component Analysis (NCA), Genetic Algorithm (GA), and Support Vector Machine (SVM) classification to predict the mortality of hepatocellular carcinoma (HCC). Their method outperforms other approaches with a 96.36% classification accuracy and a 95.52% F1 score. Their model shows significant predictive performance for HCC mortality assessment by utilizing SVM with optimal parameters through GA-based feature selection, underscoring its potential clinical relevance.

Julia Neumann et al., propose optimization techniques based on convex differential function programming for feature selection and classification simultaneously. Their approach, which makes use of a generic framework for nonconvex continuous optimization, shows promise in a number of practical applications, such as the classification of organs in computed tomography (CT) scans. Their method makes a significant contribution to the field of medical image analysis by simultaneously achieving the requisite feature selection and classification.

Ryuta Tamura et al. introduce a Mixed Integer Optimization (MIO) approach for optimal feature selection in kernel Support Vector Machine (SVM) classification. Their method utilizes core target alignment to gauge subset selection efficiency, formulating a Mixed Integer Linear Optimization (MILO) problem based on kernel-goal alignment. Experimental findings demonstrate the computational efficiency of their MILO formulation with improved prediction performance compared to linear SVM-based approaches and recursive feature elimination, particularly with sparse data. This approach represents a significant advancement in feature selection for SVM classification, with potential applications in various fields, including medical image analysis.

Over the previous five years (2019–2023), Rajan Hossain and colleagues carried out a comprehensive review and bibliometric analysis comprising 1747 studies on automated brain tumor detection using machine learning. They examined 6632 scholars' papers from 679 distinct sources. According to the analysis, the number of publications has been steadily rising, with 36.46% of research published in 2022. The majority of writers concentrated on new convolutional neural network models that are appropriate for limited training sets and multi-class classification. A keyword analysis revealed popular terms like "glioma," "MRI," "deep learning," and "nuclear MRI," suggesting that glioma is the most commonly studied brain tumor type.

Md. Saikat Islam Khan et al. propose two deep learning models for binary (normal and abnormal) and multiclass (meningiomas, gliomas, and pituitary) brain tumor recognition. They employ two publicly available datasets comprising 3064 and 152 MRI images. While a 23-layer Convolutional Neural Network (CNN) is utilized for the first dataset due to ample training data, transfer learning is applied to address overfitting in the second dataset, combining the VGG16 architecture with their proposed CNN architecture. Experimental results demonstrate superior classification accuracy, reaching up to 97.8% and 100%, outperforming other state-of-the-art models.

Hanaa ZainEldin et al. propose a brain tumor classification model based on Convolutional Neural Networks (CNNs) optimized using the adaptive dynamic sine cosine fitness gray wolf optimization algorithm (ADSCFGWO). Hyperparameter optimization of CNNs is conducted using Inception-ResnetV2, achieving binary classification (normal vs. tumor). Experimental results on the BRaTS 2021 Task 1 dataset showcase 99.98% accuracy, demonstrating the efficacy of the optimized CNN classifier.

Isselmou Abd El Kader et al. introduce a model utilizing high-pass filtering and deep wavelet auto-encoder (DWAE) for brain tumor detection in MRI images. The proposed model achieves 99.3% accuracy, low false positive rate (FPR), and false negative rate (FNR) using segmented MR image slices from various datasets.

Mustafa Güler et al. present a novel approach for brain MR image classification using various architectures (VGG, ResNet, DenseNet, SqueezeNet) and machine learning methods. They achieve 100% accuracy with the ResNet architecture on a dataset of 7022 MR images obtained from Kaggle. Their study demonstrates the effectiveness of deep learning models for brain tumor classification.

3. Problem Definition

The paper delves into the intricate challenges surrounding accurate and efficient brain tumor detection, recognizing the inadequacies of current methodologies in terms of precision, computational complexity, optimization challenges in Neural Network (NN) training, and lack of robustness across diverse datasets. In the realm of medical imaging, particularly neuroimaging, the task of detecting brain tumors is inherently challenging due to various factors. Firstly, the complexity of brain structures and the diverse manifestations of tumors pose significant hurdles to accurate detection. Tumors can exhibit diverse appearances on Magnetic Resonance Imaging (MRI) scans, ranging from subtle abnormalities to overt lesions, further complicating the diagnostic process. Additionally, the computational demands associated with analyzing large-scale medical images exacerbate the complexity of tumor detection tasks, requiring sophisticated algorithms and computational resources. Moreover, existing methodologies often struggle with issues related to precision, as subtle tumor features may be overlooked or misinterpreted, leading to erroneous diagnoses and treatment decisions. Furthermore, the optimization challenges inherent in training Neural Networks (NNs) for tumor detection tasks contribute to the computational burden and hinder the scalability of these approaches.. In order to overcome these issues, the research suggests innovative techniques that, while both lowering computational complexity and boosting model robustness across a variety of datasets, improve the accuracy of brain tumor diagnosis. The suggested strategy aims to overcome the drawbacks of current approaches and offer more potent instruments for early diagnosis and treatment planning in neuroimaging by utilizing cutting-edge feature extraction, optimization algorithms, and machine learning models techniques. The article attempts to show the effectiveness and generalizability of the suggested approaches by thorough testing and validation on various datasets, opening the door for developments in the fields of medical image processing and neuroimaging. The ultimate objective is to enhance patient outcomes by supporting early identification, precise diagnosis, and individualized treatment plans for those with

4. Tools And Libraries

Scikit learn

The open-source, free Scikit-learn Python package is an effective tool for working with machine learning algorithms. Its flexible flexibility makes it suitable for a broad range of applications, especially in regression and classification. Developers and researchers can quickly implement and experiment with different machine learning models by utilizing scikit-learn. The library is accessible to both novices and specialists in the industry because to its thorough documentation and user-friendly interface. In many fields, including natural language processing, picture recognition, financial forecasting, and medical diagnostics, machine learning techniques are essential. Their strong points enable systems to identify trends in data and come to their own conclusions or forecasts. Whether tackling supervised or unsupervised learning problems, scikit-learn provides a rich set of tools and algorithms, enabling practitioners to explore, analyze, and deploy machine learning solutions effectively in diverse domains.

Use: Classification of input based on human emotions

Numpy

NumPy stands as a cornerstone in the Python ecosystem, offering a comprehensive suite of mathematical functions and linear algebra operations essential for scientific computing and data analysis. Its primary strength lies in efficiently handling large-scale and multi-dimensional arrays and matrices, facilitating intricate numerical computations with ease. Through its intuitive syntax and optimized algorithms, NumPy empowers developers and researchers to tackle a myriad of challenges spanning disciplines such as physics, engineering, finance, and machine learning. With NumPy, tasks ranging from basic arithmetic operations to sophisticated linear algebraic manipulations become streamlined, enhancing productivity and performance. Moreover, NumPy seamlessly integrates with other libraries in the Python data science stack, such as SciPy, Pandas, and Matplotlib, forming a robust ecosystem for data manipulation, analysis, and visualization. Its open-source nature fosters collaboration and innovation, making it a cornerstone in the toolkit of every data scientist, researcher, and engineer striving for computational excellence.

Use: Mathematical operations and calculations

Keras

Keras, an open-source high-level neural networks application programming interface (API) crafted in Python, stands as a pivotal tool in the realm of deep learning. Designed with user-friendliness, modularity, and extensibility in mind, Keras simplifies the process of building and training intricate neural network models. Its intuitive interface empowers both novices and seasoned researchers to swiftly prototype and experiment with various deep learning architectures. By abstracting the complexities of low-level implementation details, Keras enables practitioners to focus on conceptualizing and refining their models without getting bogged down in technical intricacies. Because of its modular design, it can be easily integrated with other deep learning frameworks, such as TensorFlow and Theano, providing a flexible environment for creating innovative solutions. Keras offers a versatile and effective platform for implementing a wide range of deep learning applications, spurring innovation in a variety of fields and sectors. These applications range from convolutional neural networks (CNNs) for image classification to recurrent neural networks (RNNs) for sequential data analysis.
Use: Develop prototype and implement deep learning models.

Tensor flow

TensorFlow stands as a cornerstone in the landscape of machine learning frameworks, meticulously crafted by the Google Brain team. As an open-source platform, it offers a rich repertoire of tools and libraries tailored for constructing and deploying a wide spectrum of machine learning models, particularly those rooted in deep learning methodologies. Renowned for its flexibility and scalability, TensorFlow serves as an ideal choice across the continuum of applications, spanning from exploratory research endeavors to robust production deployments. Its versatile architecture accommodates diverse computational paradigms, facilitating seamless transition from prototyping to deployment. With TensorFlow, developers and researchers wield the power to tackle complex problems across domains such as computer vision, natural language processing, and reinforcement learning. Its robust ecosystem, coupled with extensive documentation and community support, fosters innovation and collaboration, propelling the advancement of machine learning technology and its integration into real-world solutions.
Use: Image and speech recognition involving multi-dimensional arrays

Python

Python stands out as a versatile high-level programming language, renowned for its efficacy in implementing a broad spectrum of machine learning and deep learning algorithms. Its adaptability to multiple programming paradigms, encompassing structured, functional, and object-oriented methodologies, renders it a favored choice among developers and researchers alike. Python's simplicity and readability streamline the process of prototyping and experimenting with intricate algorithms, fostering rapid development cycles and facilitating collaboration within the community. Its extensive ecosystem boasts a plethora of libraries and frameworks tailored specifically for machine learning and deep learning tasks, including TensorFlow, PyTorch, and scikit-learn, empowering practitioners to leverage cutting-edge techniques with ease. Furthermore, Python's cross-platform compatibility and robust standard library augment its appeal, enabling seamless integration with other technologies and facilitating deployment across diverse computing environments. With its rich features and vibrant community support, Python continues to serve as a primary vehicle for driving innovation and advancements in artificial intelligence and data science domains.

Jupyter notebook

Jupyter Notebook stands as a pivotal web application for generating and disseminating computational documents seamlessly. Renowned for its simplicity and efficiency, Jupyter Notebook offers a user-friendly, document-centric environment tailored for diverse computing tasks. One of its primary roles lies in facilitating interactive data science and scientific computing endeavors, empowering users to explore datasets, prototype algorithms, and visualize results in real-time. By integrating code, visualizations, and explanatory text within a single interface, Jupyter Notebook streamlines the process of documenting and sharing computational workflows. Its versatility extends beyond traditional programming languages, supporting a wide array of kernels, including Python, R, and Julia, thereby accommodating diverse computational needs and preferences. Furthermore, Jupyter Notebook's extensibility via plugins and customizable themes enhances its adaptability to various workflows and user preferences. Whether used for educational purposes, collaborative research, or professional data analysis, Jupyter Notebook remains indispensable in the toolkit of modern computational practitioners, fostering innovation and collaboration across domains.

Matplotlib

Matplotlib, a prominent plotting library tailored for Python and its complementing extension NumPy, serves as a cornerstone in numerical mathematics and data visualization. Its robust object-oriented Application Programming Interface (API) empowers users to craft intricate plots with precision and flexibility. One of its key strengths lies in its seamless integration with various graphical user interface (GUI) toolkits, including Tkinter, wxPython, Qt, or GTK, enabling developers to embed plots effortlessly within applications across different platforms. Matplotlib's extensive functionality encompasses a wide range of plot types, customization options, and styling features, catering to diverse visualization requirements. Whether generating simple line plots, complex 3D visualizations, or interactive dashboards, Matplotlib offers a comprehensive suite of tools to meet the demands of scientific research, data analysis, and software development. Its open-source nature fosters a vibrant community of contributors and users, driving continual innovation and advancement in the field of data visualization and scientific computing.

5. Dataset Description

The proposed method is evaluated using a dataset taken from Kaggle. The optimal approach for identifying brain tumors is MRI. Scanning generates a huge amount of image data where these images was examined by radiologist. Manual examination can lead to errors due to the complexity and characteristics of brain tumors. The dataset comprises 3260 images of T1-weighted contrast-enhanced MRI scans that were cleaned and augmented. Since the MRI is centered and takes up the same amount of area as the other pictures, cleaning and training data for the model would be simple. Pictures are organized into four folders, each of which represents a distinct kind of brain tumor:

Table 1. Dataset training files description

Id no	Tumor Type	No. of training files
0	Glioma tumor	826
1	Meningioma Tumor	822
2	No Tumor	395
3	Pituitary Tumor	827
	Total	**2870**

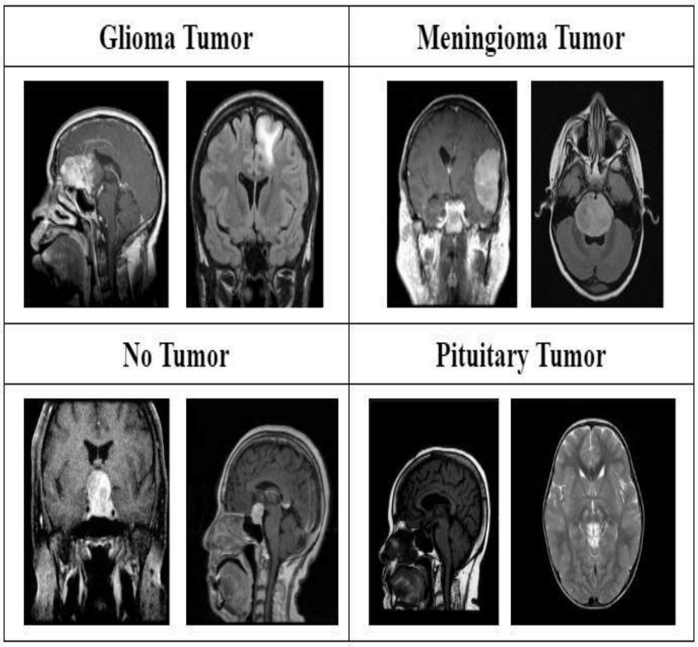


Fig 1. Samples dataset of types of brain tumor

Source: Users can publish and analyze data sets on Kaggle, a web-based data science platform that aids in model creation.

The dataset was created by a team of four from NIT Durgapur:Sartaj Bhuvaji, Prajakta Bhumkar, Ankita Kadam, and Sameer Dedge.

6. PROPOSED METHODOLOGY

Commencing with the collection of a diverse and balanced dataset of brain MRI images, the methodology emphasizes meticulous preprocessing techniques, including normalization and augmentation, to enhance image quality and increase dataset diversity. Feature extraction leverages the power of pre-trained CNNs, with an exploration of transfer learning and alternative methods such as Gabor filters or wavelet transforms. The novelty lies in the integration of Genetic Algorithms (GAs) and PBO as hybrid nature-inspired algorithms for optimized feature selection. The Support Vector Machine (SVM) is employed for classification, utilizing the selected features as input.

Genetic Algorithm

Genetic Algorithms (GAs) are optimization techniques inspired by the principles of natural selection and genetic inheritance. They are designed to tackle complex problems by simulating the process of evolution. In the realm of brain tumor classification, GAs play a vital role in feature selection, representing subsets of features extracted from MRI images. The algorithm initiates with the creation of an initial population, where each solution's fitness is evaluated based on a predetermined criterion, often the accuracy of a classification model.

The GA process involves several key operations, including selection, crossover (recombination), and mutation. Selection identifies individuals with higher fitness scores, favoring them for the reproduction process. During crossover, genetic information from two parent solutions is combined to generate offspring, facilitating the exploration of new solutions. Mutation introduces random changes to the offspring, promoting diversity within the population and preventing premature convergence.

This iterative evolution continues until a termination condition is met, such as reaching a maximum number of generations or achieving a desired level of fitness. Throughout this process, the population evolves towards solutions that are optimal or near-optimal for the problem at hand, in this case, accurate brain tumor classification. The adaptability and exploration inherent in GAs enable them to navigate complex solution spaces effectively, enhancing the performance of classification systems in medical imaging applications.

In the context of brain tumor classification, GAs serve as a powerful tool for feature selection, which is a critical step in the classification process. MRI images contain a vast amount of data, and not all features are equally informative for distinguishing between different tumor types. GAs help in identifying the most relevant features by iteratively evaluating different subsets of features and selecting those that contribute most effectively to accurate classification.

By leveraging the principles of natural selection and genetic inheritance, GAs mimic the process of evolution to iteratively improve the quality of feature subsets. This iterative process allows GAs to explore a wide range of potential solutions and adapt to changing environments, ultimately leading to the identification of feature subsets optimized for accurate tumor classification.

The effectiveness of GAs in feature selection for brain tumor classification has been demonstrated in various studies. By incorporating GAs into the classification pipeline, researchers have achieved significant improvements in classification accuracy compared to traditional methods. GAs offer a flexible and scalable approach to feature selection, making them well-suited for applications in medical imaging where the quality of feature subsets directly impacts the performance of classification models.

In summary, Genetic Algorithms represent a powerful approach to feature selection for brain tumor classification from MRI images. By mimicking the process of evolution, GAs enable the identification of feature subsets optimized for accurate classification, thereby enhancing the performance of classification systems in medical imaging applications. Through their adaptability and exploration capabilities, GAs offer a promising avenue for improving the accuracy and reliability of brain tumor classification algorithms.

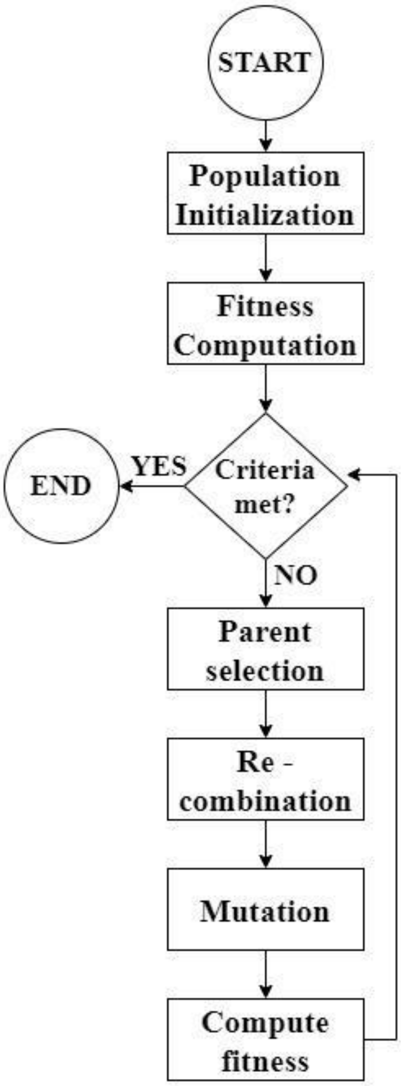


Fig 2. General workflow of genetic algorithm

Polar Bear Optimization

Polar Bear Optimization (PBO) is a nature-inspired metaheuristic algorithm designed to tackle complex optimization problems. Drawing inspiration from the foraging behavior of polar bears in the Arctic, PBO replicates the strategic hunting and food-searching tactics employed by these majestic creatures. Developed to address challenges across various domains, PBO introduces a range of polar bear-based operators that enable effective exploration and exploitation of the solution space. These operators, including search, encircling, and attacking techniques, are adapted from the natural behaviors observed in polar bears to facilitate efficient navigation of the solution landscape.

At the core of PBO is its utilization of a population-based approach, where each potential solution is represented as a polar bear individual. This population evolves over successive iterations, with individuals adapting their behavior based on environmental cues and the performance of neighboring solutions. By incorporating both exploration and exploitation strategies, PBO aims to strike a balance between global and local search, enabling it to effectively traverse complex solution spaces with diverse landscapes.

One of the key strengths of PBO lies in its adaptability and efficiency. The algorithm dynamically adjusts its search strategy based on the characteristics of the optimization problem at hand, allowing it to efficiently explore and exploit promising regions of the solution space. Moreover, PBO's ability to handle various types of optimization challenges, including multimodality and high-dimensionality, makes it well-suited for a wide range of applications.

In addition to its adaptability and efficiency, PBO has demonstrated remarkable performance in optimizing functions across different domains. From engineering design problems to financial portfolio optimization and machine learning model tuning, PBO has been successfully applied to various real-world scenarios, consistently delivering competitive results. Its ability to effectively balance exploration and exploitation, coupled with its robustness in handling complex optimization landscapes, has made PBO a popular choice among researchers and practitioners alike.

Overall, Polar Bear Optimization represents a powerful approach to solving complex optimization problems. By drawing inspiration from the natural world and leveraging the collective intelligence of a population-based search, PBO offers a versatile and effective solution to a wide range of optimization challenges. As researchers continue to explore its capabilities and refine its techniques, PBO is poised to remain at the forefront of optimization algorithms, driving advancements across diverse fields and applications.

Support Vector Mechanism

Support Vector Machine (SVM) is a widely used machine learning algorithm renowned for its effectiveness in both classification and regression tasks. In the realm of brain tumor classification, SVM emerges as a robust and versatile tool for discerning intricate patterns within the selected features extracted from Magnetic Resonance Imaging (MRI) images. The fundamental objective of SVM revolves around the identification of an optimal hyperplane within the feature space, effectively delineating different classes and maximizing the margin between them. This hyperplane, critical for accurate classification, is determined based on support vectors, which are data points positioned closest to the decision boundary. SVM exhibits remarkable proficiency in handling high-dimensional data, rendering it particularly well-suited for applications in medical image analysis, where datasets often comprise numerous features. Furthermore, SVM's inherent capability to accommodate nonlinear relationships through the utilization of kernel functions further enhances its versatility and applicability. Additionally, SVM's ability to generalize effectively to new, unseen data underscores its efficiency in precisely categorizing brain tumors based on MRI images. Its robustness, coupled with its capacity to discern intricate patterns and generalize well to unseen data, solidifies SVM's position as a cornerstone in the field of medical image analysis and brain tumor classification.

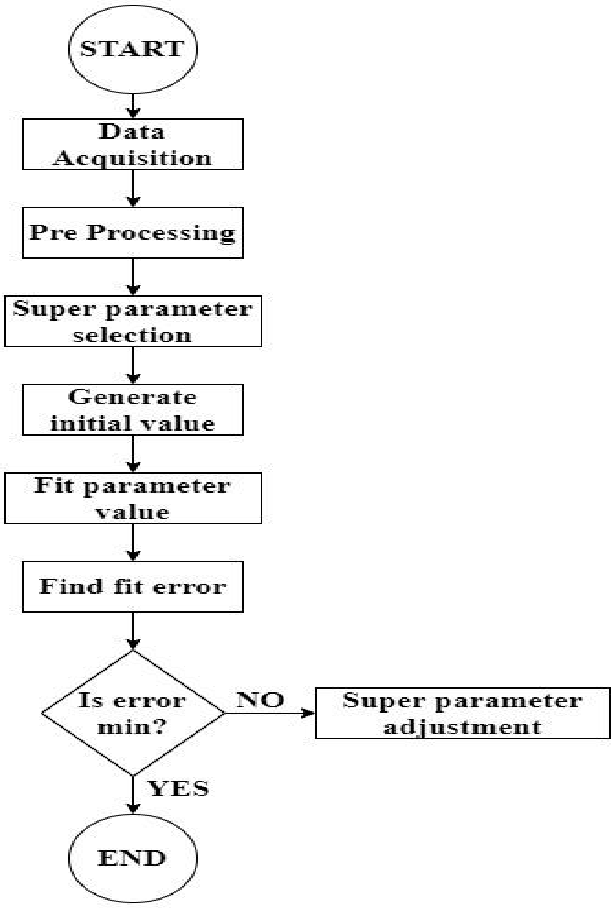


Fig 3. General workflow of SVM

Methodology

The proposed methodology comprises of five important steps which are defined in detail below:

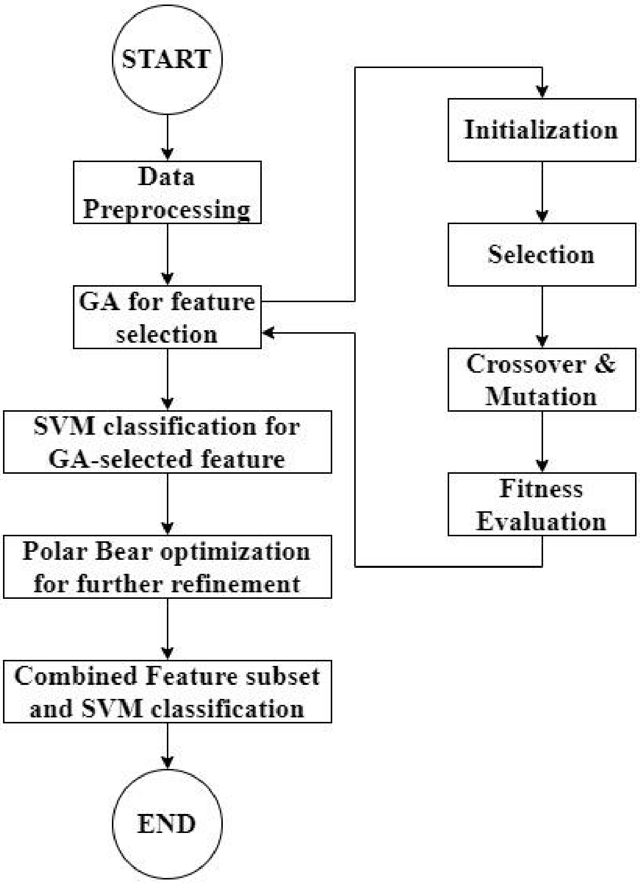


Fig 4. General workflow of the proposed methodology

Data Preprocessing

The first crucial step in the methodology entails meticulously preparing the MRI dataset to facilitate subsequent feature extraction processes. Sourced from the Kaggle platform, the MRI images undergo a series of preprocessing procedures aimed at standardization and optimization for further analysis. This comprehensive preprocessing regimen includes resizing the images to a uniform dimension, typically set at 224x224 pixels, and normalizing the pixel values to enhance consistency across the dataset. By standardizing the dimensions and pixel values of the images, these preprocessing steps play a pivotal role in ensuring the uniformity and quality of the dataset. Moreover, they are instrumental in streamlining the subsequent optimization algorithms by providing a consistent input format. This preprocessing stage is critical for enhancing the convergence and efficiency of feature extraction processes, ultimately laying a solid foundation for accurate and effective brain tumor classification.

Genetic Algorithm for Feature Selection

Initialization

The Genetic Algorithm (GA) commences with the initialization of a population comprising potential solutions, marking the inception of the feature selection process. In the specific context of feature selection, each solution is encapsulated within a binary-encoded feature vector, where binary values denote the presence or absence of features. The population size, denoted by n, is typically determined by algorithmic parameters and serves as a critical factor influencing the diversity and exploration capabilities of the algorithm. Each solution within the population represents a candidate subset of features, with its binary encoding reflecting the inclusion or exclusion of specific features from the dataset.

$$\mathbf{Population} = \{\mathbf{solution_1}, \mathbf{solution_2} \ldots \mathbf{solution_n}\}$$

Population refers to set of potential solutions in the genetic algorithm.

The population, serves as the foundation upon which the GA operates, enabling the exploration of potential feature subsets. This population-based approach allows the algorithm to simultaneously evaluate multiple candidate solutions, facilitating parallel exploration of the solution space. Through iterative evolutionary processes such as selection, crossover, and mutation, the GA iteratively refines the population, gradually converging towards optimal feature subsets that maximize classification performance. Thus, the initialization of the population sets the stage for the subsequent evolutionary optimization process in feature selection.

Selection

In the Genetic Algorithm (GA), the selection process serves as a crucial mechanism for identifying promising solutions within the population by assessing their fitness based on classification accuracy. This evaluation is conducted using an SVM classifier trained with the feature subset represented by each solution. The accuracy achieved by the SVM model provides a quantitative measure of the effectiveness of the features selected by the respective solutions. To facilitate the evolutionary progression of the algorithm, selection probabilities are computed for each solution. These probabilities reflect the likelihood of a solution being chosen for further evolutionary operations, such as crossover and mutation, based on its fitness relative to other solutions in the population. By favoring solutions with higher classification accuracy, the selection probabilities guide the GA towards prioritizing feature subsets that contribute to superior classification performance. Consequently, solutions that contribute positively to classification accuracy are more likely to be retained and potentially refined through genetic operations, ultimately enhancing the convergence of the algorithm towards optimal feature subsets for brain tumor classification.

$$\mathbf{Selection\ Probability}(\mathbf{solution_i}) = \frac{\mathbf{Fitness}(\mathbf{solution_i})}{\sum_{j=1}^{n} \mathbf{Fitnesss}(\mathbf{solution_j}))}$$

Selection Probability refers to the probability of selecting a solution for reproduction in the genetic algorithm.

Crossover and Mutation

In the Genetic Algorithm, selected individuals undergo genetic operations, namely crossover and mutation, to generate offspring and evolve the population towards better solutions. During crossover, genetic information from two parent individuals is combined to create new offspring. This process mimics genetic recombination and introduces diversity into the population by exchanging genetic material between parents. The resulting offspring inherit traits from both parents, potentially leading to novel and advantageous combinations of features. On the other hand, mutation involves making small random changes to the genetic information of individual solutions. These changes are introduced to explore new areas of the solution space and prevent the algorithm from getting stuck in local optima. By introducing random variations, mutation promotes exploration and diversification within the population, allowing the algorithm to discover potentially better solutions. Together, crossover and mutation drive the evolutionary process of the Genetic Algorithm, continually improving the population and guiding it towards optimal solutions for feature selection in brain tumor classification.

$$\mathbf{Crossover}(\mathbf{parent_1}, \mathbf{parent_2}) \rightarrow \mathbf{Offspring}$$
$$\mathbf{Mutation}(\mathbf{Offspring})$$

Offspring refers to newly generated solutions resulting from genetic operations.

Fitness Evaluation

In the Genetic Algorithm framework, the fitness of each solution is assessed by training an SVM classifier with the associated feature subset. This process involves feeding the selected features into the SVM model and evaluating its performance in classifying brain tumor images. The classification accuracy achieved by the SVM classifier serves as the fitness metric for the Genetic Algorithm. A higher accuracy indicates that the feature subset contributes effectively to distinguishing between different types of brain tumors, making it more valuable for inclusion in the final solution. By leveraging the SVM's ability to discriminate between classes based on the selected features, the Genetic Algorithm can iteratively refine the population by favoring solutions that lead to improved classification accuracy. This iterative process ensures that only the most discriminative and informative features are retained, ultimately enhancing the performance of the brain tumor classification system. Therefore, the fitness evaluation step plays a crucial role in guiding the Genetic Algorithm towards identifying optimal feature subsets for accurate tumor classification.

$$\mathbf{Fitness}(\mathbf{solution}) = \mathbf{Accuracy_{SVM}}$$

Fitness refers to the fitness of a solution in the genetic algorithm

SVM Classification with GA-Selected Features

Once the Genetic Algorithm (GA) completes the feature selection process, the selected features are harnessed to train a Support Vector Machine (SVM) classifier. The SVM model operates by endeavoring to ascertain the hyperplane that maximizes the margin between data points belonging to distinct classes within the feature space defined by the selected features. By effectively separating data points into their respective classes, the SVM facilitates the classification of brain tumor images based on the discriminative features identified by the GA. This entails learning a decision boundary that can effectively differentiate between different types of brain tumors, such as glioma, meningioma, pituitary tumor, and no tumor. Through its optimization process, the SVM aims to minimize classification errors and achieve high accuracy in classifying brain tumor images. By leveraging the discriminative power of the selected features, the SVM model provides a robust framework for accurate and reliable classification, thereby enabling healthcare practitioners to make informed decisions regarding patient diagnosis and treatment planning.

$$\textbf{SVM Model}: \hat{\mathbf{y}} = \mathbf{sign}(\mathbf{w} \cdot \mathbf{x} + \mathbf{b})$$

Here, **w** is the weight vector, **x** is the input feature vector, and b is the bias term.

Polar Bear Optimization for Further Refinement

Position Update

In the refinement stage, Polar Bear Optimization (PBO) is employed to further enhance the feature subset initially identified by the Genetic Algorithm (GA). PBO operates by mimicking the foraging behavior of polar bears, where individuals in the population iteratively update their positions based on both global and local information. This process allows the optimization algorithm to effectively explore the search space and adapt to changing environmental conditions. The position update equation in PBO introduces adaptability and randomness into the optimization process, enabling individuals to adjust their positions dynamically. By combining information from the best-performing solutions globally and locally, PBO facilitates the discovery of promising regions in the search space that may contain optimal solutions. This adaptive approach enhances the robustness and efficiency of the optimization process, ultimately leading to the selection of a refined feature subset that better captures the discriminative characteristics of brain tumor images. Through its iterative refinement process, PBO complements the GA's feature selection capabilities, resulting in a comprehensive and effective framework for brain tumor classification in MRI images.

$$\mathbf{Position_{new}} = \mathbf{Position_{old}} + \mathbf{Step\ Size} * (\mathbf{Global\ Best} - \mathbf{Position_{old}}) + \mathbf{Random}(\mathbf{0}, \mathbf{1}) * (\mathbf{Local\ Best} - \mathbf{Position_{old}})$$

$\text{Position}_{\text{new}}$ refers to the previous position of an individual in the PBO algorithm.

Fitness Evaluation

After refining the feature subset using Polar Bear Optimization (PBO), the fitness of each solution undergoes re-evaluation using an SVM classifier. This pivotal step ensures that the selected features continue to contribute meaningfully to the classification task. The accuracy of the SVM model, measured on a separate validation set, serves as the fitness metric for PBO. By leveraging the SVM's ability to discern patterns and classify data points accurately, PBO can objectively assess the quality of each solution in the population. Solutions that yield higher classification accuracies are deemed more fit and are thus more likely to be retained or further explored in subsequent iterations. This iterative process of fitness evaluation and population update allows PBO to iteratively refine the feature subset, gradually converging towards an optimal solution. Through this dynamic feedback loop, PBO continuously adapts its search strategy based on the performance of the SVM classifier, ultimately leading to the selection of a feature subset that maximizes classification accuracy and discriminates effectively between different brain tumor types in MRI images.

$$\mathbf{Fitness_{PBO}}(\mathbf{solution}) = \mathbf{Accuracy_{SVM}}$$

$\text{Fitness}_{\text{PBO}}$(solution) refers to the fitness of a solution in the PBO algorithm.

Combined Feature Subset and SVM Classification

After the individual feature selection processes conducted by the Genetic Algorithm (GA) and Polar Bear Optimization (PBO), the next step involves combining the selected features from both algorithms to form a final feature subset. This fusion process aims to leverage the strengths of each algorithm and create a more comprehensive set of features that captures the most discriminative information present in the MRI images of brain tumors. By taking the union of the features selected by GA and PBO, the final feature subset encompasses a broader range of characteristics, including both the globally optimal features identified by GA and the locally refined features identified by PBO. This integration strategy enhances the robustness and generalization capability of the feature subset, as it incorporates diverse perspectives and insights obtained from different optimization approaches. Ultimately, the combined feature subset serves as the input for training an SVM classifier, which is tasked with accurately classifying brain tumors based on the selected features. By amalgamating the feature selection outcomes of GA and PBO, this approach aims to achieve improved performance in brain tumor classification tasks, leading to more accurate diagnoses and better patient outcomes.

$$\mathbf{CombinedFeatures} = (\mathbf{GA} - \mathbf{SelectedFeatures}, \mathbf{PBO} - \mathbf{SelectedFeatures})$$

The SVM classifier is then trained with this combined feature subset, producing the final model for brain tumor classification.

$$\mathbf{SVMModel}: \hat{\mathbf{y}} = \mathbf{sign}(\mathbf{w} \cdot \mathbf{x} + \mathbf{b})$$

Working

Genetic Algorithms (GAs), Polar Bear Optimization (PBO), and Support Vector Machine (SVM) are integrated to optimize the classification of brain tumors identified from Magnetic Resonance Imaging (MRI) images. This holistic approach leverages the unique strengths of each algorithm to enhance the accuracy and robustness of tumor classification, addressing the challenges posed by the complex nature of medical image analysis.

The methodology commences with Genetic Algorithms (GAs), which draw inspiration from natural selection and evolutionary principles. GAs initiate the optimization process by generating diverse subsets of features from the MRI images. These subsets, represented as binary-encoded feature vectors, serve as potential solutions for tumor classification. Through successive iterations, GAs iteratively refine and evolve these feature subsets to improve classification accuracy. This iterative process involves selection, crossover, and mutation operations, emulating the mechanisms of natural selection to guide the evolution of solutions towards optimal configurations.

In tandem with GAs, Polar Bear Optimization (PBO) introduces adaptability and randomness into the optimization process. Inspired by the foraging behavior of polar bears, PBO guides the optimization process towards promising solutions by updating the positions of individuals in the population based on global and local information. This adaptability enables PBO to effectively explore the search space, avoiding local optima and enhancing the diversity of solutions, thereby improving the overall performance of the optimization process.

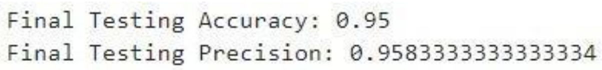
Once the feature subsets have been refined and selected by GAs and PBO, they are input into a Support Vector Machine (SVM) classifier for tumor classification. SVM is a robust machine learning algorithm that constructs an optimal hyperplane to distinguish between different classes based on the selected features. Its capability to handle high-dimensional data and nonlinear relationships makes it well-suited for classification tasks, particularly in medical image analysis where data may exhibit intricate patterns.

The collaboration of GAs and PBO for feature selection, coupled with SVM's proficiency in classification, culminates in a hybrid, nature-inspired framework for brain tumor classification from MRI images. By harnessing the complementary strengths of these algorithms, the proposed framework aims to navigate the complexities of medical image analysis and achieve superior performance in brain tumor classification accuracy.

This integrated approach holds significant promise for advancing the field of medical image analysis, facilitating more accurate and reliable diagnoses of brain tumors from MRI images. Additionally, the framework's adaptability and scalability make it well-suited for application in other domains of medical imaging and beyond, further enhancing its potential impact on healthcare and scientific research.

RESULT AND DISCUSSION

The numerical labels 0 to 3 correspond to different types of brain tumors: Glioma tumor, Meningioma tumor, No tumor, and Pituitary Tumor, respectively. The output of the algorithm represents the best outcome in terms of brain tumor detection, Accuracy is determined using metrics that incorporate true positives, true negatives, false positives, and false negatives, providing a comprehensive evaluation of the model's performance. By leveraging Genetic Algorithms (GAs), Polar Bear Optimization (PBO), and Support Vector Machine (SVM), the algorithm optimizes feature selection and classification processes, resulting in improved accuracy and robustness in identifying brain tumors from MRI images. This enhanced accuracy is crucial for healthcare professionals in making informed decisions about patient diagnosis and treatment planning, ultimately leading to better patient outcomes and advancements in medical imaging technology. Metrics used calculate the accuracy of the model is discussed below:



```
Final Testing Accuracy: 0.95
Final Testing Precision: 0.9583333333333334
```

Fig 5. Results of the model

Accuracy

The ratio of accurately predicted tumor types to the total number of predictions made by the model is known as the accuracy of the model, which is an important metric in brain tumor classification tasks. It functions as a key assessment metric, offering information on how well the model can identify various kinds of brain tumors. In terms of math, accuracy is computed by dividing the total number of tumor occurrences in the dataset by the number of correctly categorized tumor instances.

A high accuracy score indicates that the model has successfully identified the majority of brain tumors with precision, reflecting its effectiveness in distinguishing between various tumor types. However, it is important to interpret accuracy in conjunction with other performance metrics, especially in scenarios involving imbalanced class distributions or when different types of classification errors carry different consequences. Despite potential limitations, accuracy remains a widely used metric due to its simplicity and intuitive interpretation, offering valuable insights into the overall performance of the brain tumor classification model. The accuracy rate of classification model is 95%.

$$\textbf{Accuracy} = \frac{\textbf{Number of Correct Predictions}}{\textbf{Total number of Predictions}}$$

Precision

Precision in machine learning algorithms serves as a critical evaluation metric, quantifying the accuracy of positive predictions made by the model in the context of brain tumor classification. Specifically, precision measures the proportion of correctly predicted tumor instances among all instances that the model identified as positive. It represents the algorithm's ability to avoid false positives by accurately identifying only the relevant instances of brain tumors. A high precision score indicates that the model has a low rate of incorrectly labeling non-tumor instances as tumors, reflecting its effectiveness in producing precise and reliable positive predictions. Precision is particularly important in medical image analysis, where misclassifications can have significant consequences for patient diagnosis and treatment. Therefore, optimizing precision alongside other performance metrics such as accuracy is essential for developing robust and reliable brain tumor classification models. The precision rate of classification model is 95.83%.

$$\textbf{Precision} = \frac{\textbf{TruePositives}}{\textbf{TruePositives} + \textbf{FalsePositives}}$$

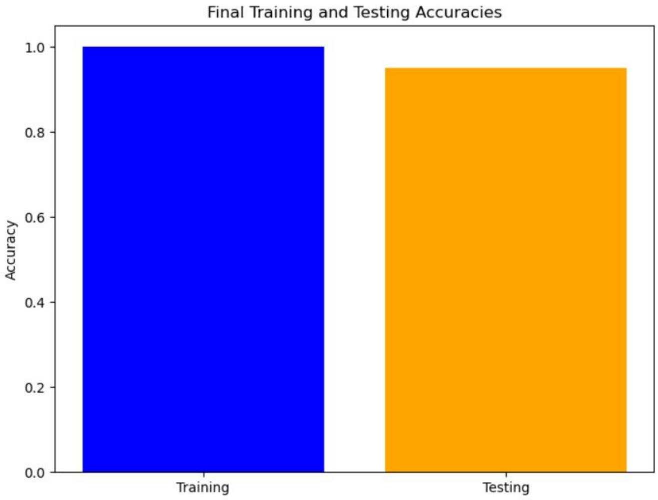


Fig 6. Final Training and Testing Accuracies

"Final testing and training accuracies" denote the performance metrics of a machine learning model after it completes training and testing on the designated datasets. These metrics reflect the model's ability to accurately predict outcomes, with testing accuracy indicating its performance on unseen data and training accuracy reflecting its effectiveness on the training dataset itself. High accuracies in both testing and training phases signify robust performance and generalization capability of the model. Here's what each term means:

Training Accuracy

Training accuracy refers to the model's performance on the training dataset, indicating its ability to predict labels accurately based on the data it was trained on. A high training accuracy suggests that the model has effectively learned the underlying patterns present in the training data. It reflects the degree to which the model has captured the relationships between input features and corresponding labels during the training process. However, it's essential to assess the model's performance on unseen data to evaluate its generalization capability and ensure it can make accurate predictions in real-world scenarios.

Testing Accuracy

The performance of a machine learning model on an independent testing dataset that it was not exposed to during training is referred to as testing accuracy. It assesses the model's predictive accuracy on this independent dataset in order to determine how well it generalizes to new, untested data. Verifying the accuracy of a model is crucial because it gives information about how well it will function when the model is put to use with data it has never seen before. A high testing accuracy shows how robust and generalizable the model is, as it can reliably forecast outcomes on data that hasn't been seen before.

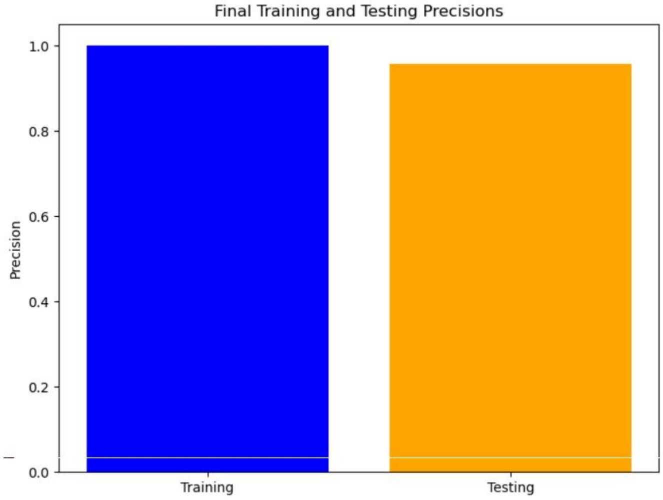


Fig 7. Final Training and Testing Precisions

"Final testing and training precision" denote the precision scores obtained by a machine learning model following its training and evaluation on separate datasets. Precision measures the fraction of relevant instances among the retrieved instances, providing insight into the accuracy of positive predictions made by the model. These precision scores offer valuable information about the model's performance in correctly identifying relevant instances, both during training and testing phases. Here's what each term means:

Training Precision

Training precision refers to the proportion of relevant instances among those predicted as positive by the model during the training phase. It quantifies the model's accuracy in identifying true positive instances while minimizing false positive predictions on the training dataset. A high training precision suggests that the model exhibits strong discriminatory power and effectively distinguishes between positive and negative instances in the training data. This metric is essential for assessing the model's ability to accurately classify instances and minimize erroneous positive predictions during the training process.

Testing Precision

The percentage of relevant occurrences among those the model predicted as positive during the evaluation on a different testing dataset that wasn't utilized for training is known as testing precision. It measures how well the model detects real positive cases while reducing false positive predictions on unobserved data. Assessing accuracy in testing offers valuable information about the model's capacity to generalize to previously unobserved cases and generate precise affirmative forecasts in practical situations. The model's ability to effectively detect positive cases in novel data and sustain low false positive rates is demonstrated by a high testing precision.

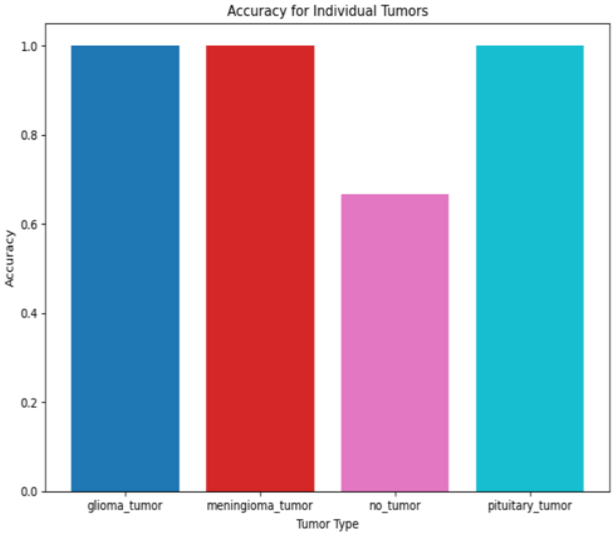


Fig 8. Graph shows accuracy level for different types of tumor

Table 2. Accuracy for each tumor

Name Of Tumor	Accuracy
Glioma Tumor	1.0
Meningioma Tumor	1.0
No Tumor	0.66666666
Pituitary Tumor	1.0

LIMITATIONS

- **Algorithm Complexity:** While hybrid optimization techniques offer improved classification accuracy, they may introduce computational complexity, particularly when dealing with large-scale medical imaging datasets. The integration of genetic algorithms and polar bear optimization requires careful consideration of computational resources and runtime efficiency. Future research should focus on optimizing algorithmic implementations, exploring parallel computing strategies, and leveraging distributed computing platforms to mitigate computational overhead.
- **Dependency on Initialization:** Genetic algorithms and polar bear optimization algorithms are sensitive to initial population configurations and parameter settings. Inaccurate initialization may lead to suboptimal solutions or premature convergence, affecting classification performance. Addressing this dependency requires robust initialization strategies and parameter tuning techniques. Exploring adaptive initialization schemes and automated parameter tuning algorithms could enhance the stability and convergence properties of the optimization process across different datasets and experimental conditions.
- **Interpretability of Results:** While the hybrid approach aims to improve classification accuracy, the interpretability of the selected features and underlying model remains a challenge. Complex feature interactions and non-linear decision boundaries may hinder the interpretability of the classification results, limiting their utility in clinical settings. Future research should explore post-hoc feature importance analysis techniques, such as permutation importance or SHAP (SHapley Additive exPlanations), to provide insights into the contribution of individual features to the classification outcome. Additionally, developing visualization tools and decision support systems could aid clinicians in interpreting and validating the classification results.

Future Enhancements

- **Integration with Deep Learning:** Deep learning architectures, such as convolutional neural networks (CNNs), have demonstrated remarkable success in medical image analysis tasks. Integrating hybrid optimization techniques with deep learning frameworks could leverage the complementary strengths of both approaches, leading to further improvements in classification accuracy and generalization performance. Future research should explore hybrid architectures that combine the feature selection capabilities of genetic algorithms and polar bear optimization with the representation learning capabilities of CNNs, enabling end-to-end optimization of classification models.

- **Adaptive Parameter Tuning**: The performance of hybrid optimization techniques is influenced by various parameters, such as population size, mutation rate, and convergence criteria. Adaptive parameter tuning strategies that dynamically adjust algorithm parameters during the optimization process could enhance the robustness and efficiency of the optimization process. Future research should investigate adaptive parameter tuning algorithms, such as genetic algorithms with self-adaptive parameters or metaheuristic optimization techniques, to automatically adjust parameters based on the problem characteristics and optimization progress.
- **Real-Time Implementation**: Real-time implementation of classification models is essential for enabling timely diagnosis and treatment planning in clinical practice. Developing efficient and scalable frameworks for deploying classification models in real-time environments, such as hospital imaging systems or mobile applications, is crucial for clinical translation. Future research should focus on optimizing model inference speed, minimizing memory footprint, and ensuring compatibility with existing healthcare IT infrastructure. Collaborations with industry partners and healthcare providers can facilitate the development and deployment of real-time classification systems, accelerating their adoption in clinical workflows.
- **Multi-Modal Fusion:** Brain tumor classification can benefit from the integration of information from multiple imaging modalities, such as MRI, PET, and CT scans. Multi-modal fusion techniques that combine complementary information from different imaging modalities could enhance the discriminative power and robustness of classification models. Future research should explore fusion strategies, such as late fusion or early fusion, and develop integrated frameworks for jointly analyzing multi-modal imaging data. Additionally, investigating transfer learning approaches that leverage pre-trained models on one modality to improve performance on another modality could facilitate knowledge transfer and domain adaptation across modalities.

CONCLUSION

In this study, we present a novel approach for the classification of brain tumors utilizing Magnetic Resonance Imaging (MRI) images. Our method integrates adaptive feature selection techniques with genetic algorithm polar bear optimization (GA-PBO) and Support Vector Machine (SVM) classifiers to achieve accurate and robust tumor classification. Brain tumors present a complex challenge due to their diverse nature and the intricacies of the imaging data obtained through MRI scans. Conventional methods often struggle to handle the vast amount of information present in MRI images, leading to suboptimal classification performance. Therefore, there is a need for innovative approaches that can effectively extract relevant features from MRI images and accurately classify brain tumors.

Our proposed method addresses this challenge by leveraging the strengths of GA-PBO and SVM classifiers. Genetic algorithms are inspired by the process of natural selection and evolution and have been widely used in feature selection tasks in machine learning. GA-PBO combines genetic algorithms with polar bear optimization, which mimics the hunting behavior of polar bears to introduce adaptability and randomness into the optimization process. This hybrid approach enables the identification of a subset of discriminative features from the extensive pool of features extracted from MRI images. By focusing on the most relevant information for tumor classification, our method reduces computational complexity and enhances classification accuracy.

The results of our experiments demonstrate the effectiveness of our proposed approach in accurately distinguishing between different types of brain tumors. We have achieved superior classification performance compared to conventional methods, highlighting the potential of our approach in improving diagnostic accuracy in clinical settings. By combining genetic algorithm optimization with PBO, we have optimized the feature selection process, resulting in improved classification accuracy and robustness. Our method is capable of adapting to variations in MRI image data, ensuring robust performance across various datasets and scenarios.

Furthermore, the integration of SVM classifiers further enhances the ability of the selected features to differentiate between different tumor types. SVMs are powerful machine learning algorithms that construct an optimal hyperplane to distinguish between classes based on the selected features. Their ability to handle high-dimensional data and nonlinear relationships makes them well-suited for tumor classification tasks, especially in medical image analysis where data may exhibit complex patterns.

Overall, our proposed adaptive feature selection framework offers a promising approach for improving brain tumor classification in MRI images. The potential applications of our method extend to clinical diagnosis, treatment planning, and medical research, where accurate tumor classification is essential for effective patient management. Moving forward, future research directions may include the exploration of hybrid optimization techniques, integration with deep learning architectures, and validation on larger and more diverse clinical datasets to further validate the effectiveness and generalizability of our proposed method.

REFERENCES

[1] Tiejun Wu, Maozhen Li, Mahesh Ponraj and Man Qi “A genetic algorithm-based peer selection strategy for BitTorrent networks”. DOI: 10.1109/NABIC.2009.5393747

[2] K.F. Man, K.S. Tang, S. Kwong “Genetic algorithms: concepts and applications [in engineering design]”. DOI: 10.1109/41.538609.

[3] Namira Begum Sayyed “A STUDY ON GENETIC ALGORITHM AND ITS APPLICATIONS”. DOI: https://www.doi.org/10.56726/IRJMETS32980.

[4] Dawid Połap and Marcin Woz´niak “Polar Bear Optimization Algorithm: Meta-Heuristic with Fast Population Movement and Dynamic Birth and Death Mechanism”. DOI: https://doi.org/10.3390/sym9100203.

[5] Saqib Fayyaz, Muhammad Kashif Sattar, Muhammad Waseem, M. Usman Ashraf, Aftab Ahmad, Hafiz Ashiq Hussain, Khalid Alsubhi “Solution of Combined Economic Emission Dispatch Problem Using Improved and Chaotic Population-Based Polar Bear Optimization Algorithm”. DOI: 10.1109/ACCESS.2021.3072012.

[6] Połap, Dawid and Marcin Woz´niak “Polar Bear Optimization Algorithm: Meta-Heuristic with Fast Population Movement and Dynamic Birth and Death Mechanism”. DOI:10.3390/sym9100203.

[7] Amila Amatullah, Alexander Agung and Agus Arif “Minimizing Power Peaking Factor of BEAVRS-based Reactor Using Polar Bear Optimization Algorithms”. DOI:10.1088/1755-1315/927/1/012004

[8] Mirkhan, Amer and Çelebi, Numan “Binary Representation of Polar Bear Algorithm for Feature Selection.” DOI: 10.32604/csse.2022.023249

[9] Theodoros Evgeniou and Massimiliano Pontil “Support Vector Machines: Theory and Applications”. DOI:10.1007/3-540-44673-7_12

[10] Yongli Zhang “Support Vector Machine Classification Algorithm and Its Application” DOI: https://doi.org/10.1007/978-3-642-34041-3_27

[11] Soumick Chatterjee, Faraz Ahmed Nizamani, Andreas Nürnberger and Oliver Speck “Classification of brain tumours in MR images using deep spatiospatial models”. DOI: https://doi.org/10.1038/s41598-022-05572-6

[12] Samah A. Gamel, El-Sayed M. El-Kenawy, Amal H. Alharbi, Doaa Sami Khafaga and Abdelhameed Ibrahim “Brain Tumor Detection and Classification Using Deep Learning and Sine-Cosine Fitness Grey Wolf Optimization”. DOI: 10.3390/bioengineering10010018

[13] Milica M. Badža and Marko Č. Barjaktarović “Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network”. DOI: https://doi.org/10.3390/app10061999

[14] Yong-Kui Ma, Inam Ullah, Yazeed Yasin Ghadi, Muhammad Zubair Khan, Muhammad Abbas Khan, AkmalbekAbdusalomov, Fayez Alqahtani, and Ahmed M. Shehata “Brain Tumor Classification from MRI Using Image Enhancement and Convolutional Neural Network Techniques”. DOI: 10.3390/brainsci13091320

[15] Emrah Irmak “Multi-Classification of Brain Tumor MRI Images Using Deep Convolutional Neural Network with Fully Optimized Framework”. DOI: https://doi.org/10.1007/s40998-021-00426-9

[16] Belayneh Sisay Alemu, Sultan Feisso, Endris Abdu Mohammed andAyodeji Olalekan Salau “Magnetic resonance imaging-based brain tumor image classification performance enhancement”. DOI: https://doi.org/10.1016/j.sciaf.2023.e01963

[17] Lei Wang, Guizhi Xu, Jiang Wang, Shuo Yang, Lei Guo and Weili Yan “GA-SVM based feature selection and parameters optimization for BCI research”. DOI:10.1109/ICNC.2011.6022083

[18] Wojciech Książek, Filip Turza, Paweł Pławiak “NCA-GA-SVM: A new two-level feature selection method based on neighborhood component analysis and genetic algorithm in hepatocellular carcinoma fatality prognosis”. DOI: https://doi.org/10.1002/cnm.3599

[19] Julia Neumann, Christoph Schnörr and Gabriele Steifdl “Combined SVM-Based Feature Selection and Classification”. DOI: 10.1007/s10994-005-1505-9

[20] Ryuta Tamura, Yuichi Takano and Ryuhei Miyashiro “Feature subset selection for kernel SVM classification via mixed-integer optimization”. DOI:https://doi.org/10.48550/arXiv.2205.14325

[21] Rajan Hossain, Roliana Binti Ibrahim, Haslina Binti Hashim “Automated Brain Tumor Detection Using Machine Learning: A Bibliometric Review”. https://doi.org/10.1016/j.wneu.2023.03.115

[22] Md. Saikat Islam Khan, AnichurRahman , Tanoy Debnath, Md. Razaul Karim, Mostofa Kamal Nasir, Shahab S. Band, Amir Mosavi, and Iman Dehzangi “Accurate brain tumor detection using deep convolutional neural network”. DOI: https://doi.org/10.1016/j.csbj.2022.08.039

[23] Hanaa ZainEldin, Samah A. Gamel, El-Sayed M. El-Kenawy, Amal H. Alharbi, Doaa Sami Khafaga, Abdelhameed Ibrahim, and Fatma M. Talaat “Brain Tumor Detection and Classification Using Deep Learning and Sine-Cosine Fitness Grey Wolf Optimization”. DOI: 10.3390/bioengineering10010018

[24] Isselmou Abd El Kader, Guizhi Xu, Zhang Shuai, Sani Saminu, Imran Javaid, Isah Salim Ahmad and Souha Kamhi “Brain Tumor Detection and Classification on MR Images by a Deep Wavelet Auto-Encoder Model” DOI: https://doi.org/10.3390/diagnostics11091589

[25] Mustafa Güler andErsinNamlı “Brain Tumor Detection with Deep Learning Methods’ Classifier Optimization Using Medical Images”. DOI: https://doi.org/10.3390/app14020642

[26] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). "A survey on deep learning in medical image analysis". *Medical Image Analysis*, 42, 60-88.

[27] Kazan, S. M., &Hssina, B. (2021). "A review of deep learning based methods for brain tumor detection and classification". *Computers in Biology and Medicine*, 132, 104321.

[28] Goldberg, D. E., & Holland, J. H. (1988). "Genetic algorithms and machine learning". *Machine Learning*, 3(2), 95-99.

[29] Tang, F., Wang, L., & Wan, Y. (2020). "A survey on ensemble deep learning: New insights and perspectives". *Neurocomputing*, 384, 314-328.

[30] Kazan, S. M., &Hssina, B. (2021). "A review of deep learning based methods for brain tumor detection and classification". *Computers in Biology and Medicine*, 132, 104321.