Automated Brain Image Classification Using Nature-Inspired Optimization-Based Machine Learning Algorithm

Gurumukh Das[1], Gurdeep Singh[2], Divya Zindani[3], G. Shanmugasundar[4], Jasgurpreet Singh Chohan[5], and Kanak Kalita[6,7](✉)

[1] Department of Mechanical Engineering, Faculty of Engineering, Dayalbagh Educational Institute (Deemed to be University), Agra 282 005, India
gurumukh@dei.ac.in
[2] Phullanwal, Punjab, Ludhiana 141013, India
[3] Department of Mechanical Engineering, Sri Sivasubramaniya Nadar (SSN) College of Engineering, Kalavakkam, India
[4] Department of Mechanical Engineering, Sri Sai Ram Institute of Technology, Chennai 600044, India
shanmugasundar.mech@sairamit.edu.in
[5] Department of Mechanical Engineering and University Center for Researdh & Development, Chandigarh University, Mohali 140413, India
jasgurpreet.me@cumail.in
[6] Department of Mechanical Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi 600062, India
kanakkalita02@gmail.com, drkanakkalita@veltech.edu.in
[7] Jadara University Research Center, Jadara University, Irbid 21110, Jordan

Abstract. The field of medicine greatly depends on medical imaging. In modern brain imaging, image classification is used to separate abnormal tissues from healthy tissue. Using various categorization algorithms, the region of the brain tumor as well as the brain tumor size is recognized in the MRI images, allowing the brain tumor to be diagnosed. Brain Magnetic resonance imaging (MRI) can quickly and effectively identify tumors, aiding neurologists in their diagnosis. Cancer, the most prevalent or leading cause of death globally, may be made more likely by tumors. Currently, efficient automation of tumor detection is crucial to finding brain tumors. In this research we suggested a strategy that combines the strength of machine learning algorithms with nature-inspired optimization approaches to classify brain image automatically. The study proposed a new classifier called Self-Adaptive Deer Hunting Optimized k-Nearest Neighbors (S-ADHO-k-NN). To start, the input image is first preprocessed to eliminate any noise using the median filtering (MF) method. The image was classified using a suggested S-ADHO-k-NN. The experimental results showed that the S-ADHO-k-NN model significantly outperformed the comparative methods with regard to of effectiveness.

Keywords: Brain Tumor · Optimization · k-Nearest Neighbors · Median Filtering · Watershed Algorithm · Principal Component Analysis

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
P. Vasant et al. (Eds.): ICO 2023, LNNS 1167, pp. 146–156, 2024.
https://doi.org/10.1007/978-3-031-73318-5_15

1 Introduction

An expansion of brain cells or cells near the brain is referred to as a cancer of the brain. Cancers of the brain may form in the head's cells. Also, potential are brain lesions close to brain tissue. For disease diagnosis and tissue segmentation, which are now issues of primary concern in clinical oncology, better visualization of brain images is necessary. This has been substantially accomplished with the use of machines, and it has led to an enhancement in the method that radiologists use to examine patients. The amount of MR scan images or slices is often enormous, making the manual examination of such large quantities laborious and time-consuming. Huge amounts of scanned images that must be manually processed and evaluated for accuracy run the risk of human mistake [1]. The recorded number of injury fatalities is about 30%. Following that, intracranial hemorrhages (ICH), extra-axial damages tumors, and traumatic brain injury (TBI) may occur. The primary cause of mortality for people of all ages globally is the ICH condition. The illness first starts in the brain as a consequence of a blood vessel leak, removing the route of interactions (following the brain function and instruction correspondingly), inside organ, and active bodily functions including memory loss, loss of vision, loss of voice, and other symptoms [2]. Deep Learning (DL) has recently been popular because of its outstanding NIA outcomes. The term "deep learning," or "DL," is used to describe machine learning (ML) techniques that place more emphasis on high-level data representation learning than task-specific learning. Due to improvements in machines and the creation of computationally quick and effective optimization techniques, DL has been widely used. DL may also be used to assess and exploit fascinating Big Data models [3]. One of the unusual cell developments in the brain is a tumor. Few tumors are brought on by cancerous or malignant cells, and the majority are benign. Primary tumors in the brain are cancers of the brain that start there. The term Atumor in the brain is an expansion of "the brain metastasis" or "secondary brain cancers." describes cancer that has progressed to the brain from another area of the body. Signs of a malignant brain tumor might include discomfort speech problems, movement problems, behavioral changes, vomit visual problems, seizures, and seizures depending on area of the disease and the section of the brain affected. Brain tumors may show clinically in a variety of ways depending on the kind, location, size, and growth rate of the tumor, making diagnosis challenging [4]. The processing of medical images involves the use of computer-aided algorithms to extract anatomical organs and analyze abnormalities like tumors and cysts, among other things. Restoration, enhancement, segmentation, classification, and compression are some of the numerous phases in image processing [5]. Because of qualities including enhanced It is used as a useful tool in the medical and surgery fields because of its soft tissue distinction, excellent resolution of space, and intensity. Additionally, it plays a significant role in the early diagnosis of brain tumors through diagnostic imaging. In order for radiologists to access patients for diagnosis and treatment, an MRI brain imaging is crucial [6]. Brain tissues may be categorized using MRI. The use of computer-based medical decision support systems is currently pervasive in fields of medicine such brain tumor treatment, heart disease, and cancer research. Due to the vast amount of image data, manual analysis (interpretation) of MRI images is tiresome and expensive (in terms of both time and money). This necessitates the creation of software for computer-aided diagnostics. In other words, there is now a motivation for automated human brain categorization. In clinical investigations, automatic categorization of the Human Brain (HB) as normal or sick using MRI images becomes crucial [7]. The brain's tumor cells multiply erratically. Tumors or lumps are formed when cancer cells proliferate and divide uncontrollably to produce a mass of tissue in any region of the body, which disrupts the body's overall metabolism. A clinical specialist contributes to the general public's improved health. A tumor is essentially an excessive development of malignant brain cells. While the A cancer tumor with a mixed phenotype does not contain active cancer cells, the benign brain tumor does not contain any [8]. Analysis is highly challenging due to the structured complexity of the human brain. Additionally, the analysis and viewing expertise are quite limited in comparison to a large number of MR images. The manual technique of examining these photos has a number of disadvantages, one of which is that it requires time. Additionally, it takes energy to maintain a high level of focus throughout categorization, which raises the accuracy of faulty identification. Because of this, MR image analysis requires an automated method with excellent accuracy. Computer assisted diagnosis (CAD) may provide the solution [9]. The study proposed a new classifier called Self-Adaptive Deer Hunting Optimized k-Nearest Neighbors (S-ADHO-k-NN). To start, the input image is first preprocessed to eliminate any noise using the median filtering (MF) method.

2 Related Work

Cellular abnormality leads to brain tumor formation. It is one of the main reasons why adult mortality worldwide. Early detection of brain tumors may save millions of fatalities. MRI earlier brain tumor identification may improve patient survival. The tumor is more clearly seen in an MRI, which aids in the process of further therapy [10]. The research [11] offers methods for brain MRI-based automated tumor identification and categorization. The foundation of these methods is hybrid-optimized. Brain MRI categorization methods separate healthy, benign, and malignant conditions. The study [12] develops a system that can automatically catalogtumors using MRI. Pre-processing is referred to as the first step in normalizing intensity. Here, min-max normalization is used for pre-processing. The research [13], presentsa brand-new brain MRI categorization system is put forward. First, they used a MobileNetV2 that had been trained on the Image net dataset to extract features from the input brain images. They just compute the output of a certain layer of the deep network to produce the feature vector rather than training it. The study [14] suggests a thorough procedure for identifying the malignant spot in MRI images. Here, optimum image segmentation based on the support vector neural network approach is used after image noise reduction. In study [15, 18], they offer a unique framework called Tumor Bagging, whose goal is to combine several Multilayer Perceptron (MLP)-based segmentation techniques in order to improve the performance of brain tumor segmentation. The paper [16, 19] proposes to provide an automated hybrid technique employing interval type-2 fuzzy based clustering and Cuckoo Based Search (CBS) to segment MR brain images effectively. A radiologist may quickly evaluate and easily analyze difficult tumor areas of ambiguous gray level regions with little user interface thanks to an automated MR brain imagesegmentation. One of the key issues in medical applications is brain tumors. It is clear that the majority of malignancies are brain-related and pose substantial issues for medical applications. So much effort is put into segmenting malignant areas, yet relatively little is done to classify malignancies and determine if a patient has cancer or not [17, 20].

3 Proposed Methodology

A novel classifier named Self-Adaptive Deer Hunting Optimized k-Nearest Neighbors (S-ADHO-k-NN) was suggested by the research. The MF approach is initially used to preprocess the input image in order to remove any noise. Images are segmented using the markers-based Watershed method. This section must discuss the suggested approach.

3.1 Data Set

The OASIS dataset for brain images (http://www.Oasis-brains.org/), Harvard Medical School (http://med.harvard.edu/AANLIB), and the AANLIB dataset for brain scans are all online resources. Were used to produce MRI images that were T2-weighted with 256256 radial phase directly sensitivity. The T2 model was selected because T2 images were more contrasted and clearer than those produced by the T1 and PET modalities. It is just one form of typical neuron. MRI and seven types of aberrant brain MRIs in the dataset, hence 160 patient MRIs were selected for the experiment, 20 normal and 140 abnormal.

3.2 Image Pre-processing Using Median Filter

It is a useful tool for partially discriminating between legitimate visual characteristics like edges and lines and isolated out-of-range noise. Instead of replacing a pixel with the average of all the pixels in a neighbourhood, the median filter substitutes a pixel. The median filter works by replacing a pixel's value with the average value of the pixels in the little window next to it. The following illustration shows how the median filtering procedure is depicted for the *mm* pixel window:

$$Med(j(o, r)) = Median\{j(o+q, r+s)\}|q \tag{1}$$

$$s \in \left[\frac{-m-1}{2}, \ldots, 0, \ldots, \frac{m-1}{2}\right] \tag{2}$$

where $j\ (o, r)$ is the pixel value at the position (j, , q), and j, q is a s. A fundamental and crucial method for pixel-based segmentation is the threshold. In its simplest form, it is used to convert grayscale photos into images that are binary. An appropriate threshold value is chosen for segmentation. Intensity values greater than the maximum value and equal to converted to white, while intensity values less the threshold value are greater or equal to altered to black pixels. The threshold value is ascertained and transformed from a grayscale to a digital image using Otsu's method.

$$b_a^2(sx) = a_0(sx)b_0^2(sx) + z_1(sx)b_1^2(sx) \tag{3}$$

where a_0 and a_1 are probabilities for two classes that are divided by a threshold value. sx Where the two classes of variances, b_a^2 and b_0^2, are used to determine the highest value of $b_1^2\ (sx)$.

$$a_0(sx) = \sum_{i=0}^{sx-1} v(i) \tag{4}$$

$$a_1(sx) = \sum_{i=sx}^{K-1} v(i) \tag{5}$$

3.3 Image Classification Using Self-Adaptive Deer Hunting Optimized k-Nearest Neighbors (S-ADHO-K-NN)

Self-Adaptive Deer Hunting Optimization

A novel metaheuristic SADHO approach has been created in this study to aid in the process of hyperparameter tuning, which was inspired by a group of hunters going deer hunting. When hunting deer, the hunter surrounds the animal and employs various tactics to get closer to it.This tactic involves carefully considering a number of factors, including the location of the deer, the wind's direction, and others. Hunting is particularly efficient when hunters cooperate with one another, which is another important criterion. The following illustration depicts the model's goal function:

$$e(x) = \max(accuracy) \tag{6}$$

According to the description of weight optimization using the SADHO approach, deer might readily evade hunters because of their special talents. A vector of the fictitious hunter population starts the procedure. The following equation describes it:

$$V = \left\{V_{1,} V_{2,\ldots,} V_m\right\} 1 < i \le m \tag{7}$$

where m stands for the population (weight) of predators, and the overall weight used for efficiency is indicated as follows. Then, important variables like position angle (weight) and wind angle are used. The whole search area is thought of as a circle, therefore the wind angle may be thought of as the circle's radius.

$$\theta_i = 2\pi a, \tag{8}$$

where a stands for the arbitrary value within and i stands for the current iteration. Now, implies the wind's direction. The location propagate for optimizing is then displayed together with the leader position (V_i) and successor location ($Z._O$). While the leader location designates the hunter's primary location, the successor location designates the exact spot of following weights. The position update method begins by simulating the around conduct" as follows. After establishing the ideal location, all the weights in the population attempt to reach the optimal place.

$$V_{i+1} = V_J - Z._O.|K \times V_J - V_i| \tag{9}$$

The Y and K coefficient vectors are used in this procedure. Assume that V_J represents the position at the current iteration, and V_{i+1} represents the location at the next iteration. The equation to estimate the Y and K coefficient vectors is provided below. An arbitrary value, that is, one presented by taking the wind speed into account, is designated asd, and it includes values from 0 to 2.

$$Y = \frac{1}{4} log\left(i + \frac{1}{i_{max}}\right)a, \ \mathrm{L} = 2.\mathrm{d}, \tag{10}$$

where the maximum iteration is indicated by i max. The value of the variable b ranges from -1 to 1, while the values of the other variables fall between. The hunter's initial location is indicated by (Y, , L), which is upgraded based on the location of the prey. To get to the ideal spot (Ya, L a), the Y and K coefficient vectors are both adjusted. The position updating method operates when d1, which suggests that the hunter might travel freely without taking the angles position into account. The searching space is thought to increase with angle location updates. It is essential to explain the angle location of As a way to improve the effectiveness of the predator's hunting technique. It may be carried out by

$$V_{i+1} = V_I - O.|\cos(x) \times V_J - V_i|, \tag{11}$$

The ideal location may be represented as V_{i+1}, V_t, and d, where d stands for arbitrarily chosen values. Because each individual position is opposite the angle location, the prey is not aware of the hunter's presence. The exploration presents the vector K as part of the encircling action. When using the arbitrarily searching method, K values are initially assumed to be less than 1.

$$V_{i+1} = V_t - Y.\ \mathrm{o.}\ |L \times V_t - V_i| \tag{12}$$

k-nearest neighbour method

It gives membership to the sample vector rather than giving it a clear-cut value of 0 or 1. The K-nearest neighbours of the sample vector and their class memberships in the potential classes are used as the foundation for the KNN algorithm's member assignment process.

$$w_j(v) = \frac{\sum_{i=1}^{l} wji\left(\frac{1}{\|v-v_i\|^{\frac{2}{n-1}}}\right)}{\sum_{i=1}^{l}\left(\frac{1}{\|v-v_i\|^{\frac{2}{n-1}}}\right)} \tag{13}$$

This m stands for determined by the flexible intensity option, which controls exactly highly a distance is weighted when calculating the worth that each neighbour adds to the group, where $j = 1,2,\ldots\ldots$, k and $i = 1,2\ldots\ldots.. v$, with k as the quantity of neighbours closest to you and C as the number of classes. The distance in Euclidean terms between x and its i th closest neighbor, w_j, is given by vw_j. Indicates the level of membership of the training set's pattern w_j to class I. There are several ways that wji may be defined.

Each training pattern is given a clear membership, which means that it is a full member of its specific class and is not a member of any other classes. The alternative method involves restricted fuzzy membership, in which the membership of w_j in each class is determined by finding the K closest neighbors of each training sample vector:

$$W_{ji}(v_l) = \left\{ \begin{array}{c} 0.51 + \frac{m_i}{L} \times 0.49, \textit{if } i = 1 \\ \frac{m_i}{L} \times 0.49, \textit{if } i \neq 1 \end{array} \right\} \tag{14}$$

This nj in Eq. (9) refers to the total amount of neighbors in the i th class. The membership calculated using Eq. (9) must meet the following equations

$$\sum_{j=1}^{2} W_{ji} = 1, i = 1{,}2 \ldots\ldots, n \tag{15}$$

$$w_{ji} \in [0{,}1] \tag{16}$$

In our first testing, we found that the fuzzy classifier with crisp initialization produced better classification accuracy than the limited initialization method. The membership value w_j for each query sample is calculated for each class taken into account, and then it is assigned to the class with the greatest membership, i.e.,

$$d(v) = arg\left[max\left(w_j(v)\right)\right]_{j=1}^{2} \tag{17}$$

4 Result and Discussion

To determine whether product reviews were positive or negative, we developed a system using the K-NN machine learning classification technique. Compared to the currently used procedures [2, 21, 22], the suggested strategy is extremely likely to provide outcomes that are more successful. Some of the results of the experiment include the following: Accuracy, Precision, F1score, and computation time as shown below.

Accuracy in statistics and machine learning refers to the proportion of accurate predictions a model or algorithm makes. It is derived by dividing the total number of forecasts made by the number of predictions that were accurate. For instance, a model's accuracy would be 90% if it successfully predicted 90 out of 100 instances. Higher accuracy is generally preferred in most applications, although it may not always be the most crucial statistic, depending on the particular use case. Figure 1a demonstrates the outcomes of accuracy for proposed and existing methods. The proposed S-ADHO-k-NN is better than existing methods.

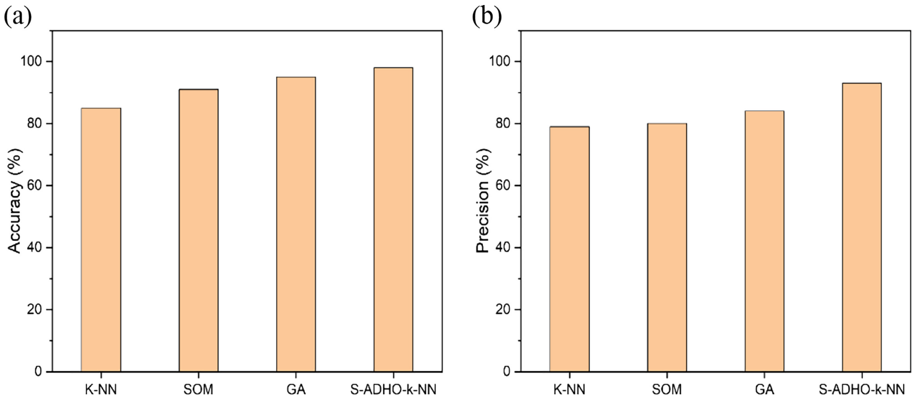

Fig. 1. Comparison of (a) accuracy (b) precision

Precision, or the percentage of true positive predictions among all positive predictions made by the model, is a measurement of how accurately a model or algorithm makes positive predictions. To put it another way, precision is the percentage of real positives out of all the occurrences the model correctly identified as positive. Figure 1b demonstrates the outcomes of precision for proposed and existing method. The suggested (S-ADHO-k-NN) technique outperforms the current one.

It is also referred to as sensitivity and represents the percentage of relevant documents that were found compared to all relevant instances. It is the percentage of pertinent instances that have been located out of all pertinent cases and documentation. 99.98% of the recall was successful. Figure 2a demonstrates the outcomes of recall for proposed and existing methods. It is concluded that the proposed S-ADHO-k-NN features a 2% increase in efficiency for Recall value.

F-measure or F1 score includes both recall and precision. The estimated harmonic average of Precision and Recall is the F1 score. Figure 2b demonstrates the outcomes of f1 score for proposed and existing methods. The suggested S-ADHO-k-NN is show to have a 2% boost in F1 score efficiency.

Profiling tools that keep track of the time spent on each operation or function call may be used to calculate the computation time of a program or algorithm. This data may be used to identify performance bottlenecks and improve the code. In general, cutting down on computation time may have a big impact on productivity and efficiency, particularly when working with lots of data or complicated computations. Figure 3 demonstrates the outcomes of computation time for proposed and existing methods.

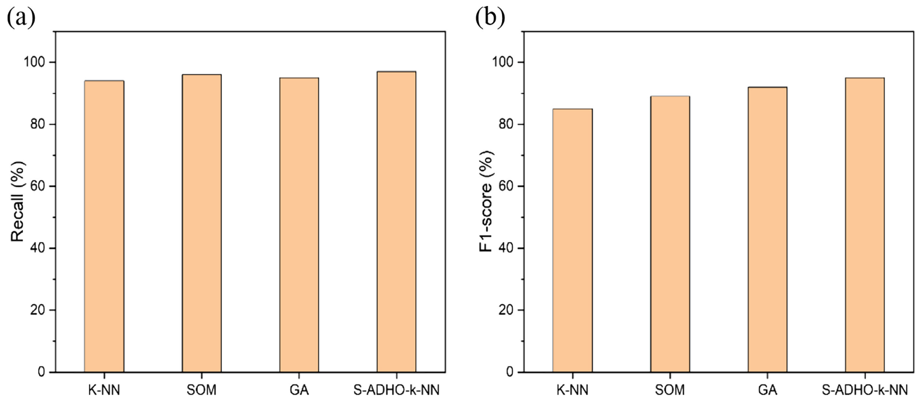

Fig. 2. Comparison of (a) recall (b) F1 score

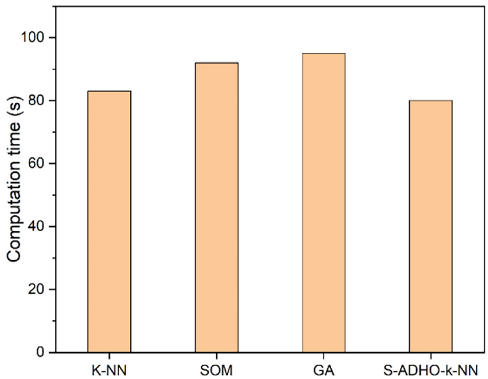

Fig. 3. Comparison of Computation time

5 Conclusion

This study uses S-ADHO-k-NN to propose effective robotic approach for classifying and segmenting brain cancers. Pre-processing, extracting features and separation with selection, and classification for brain tumours make up the main phases of this study. The MF and skull stripping are used to extract the regions of interest (ROI). The PCA is used to identify relevant features after feature extraction using texture data based on the DWT. The categorization of brain tumour forms utilizing good and bad the S-ADHO-k-NN algorithm. With 98% accuracy, 93% precision, and 97% recall, the suggested framework for categorising benign or cancerous tissues using brain MRI images is effective. The results highlight how important it is to compare the suggested approach's accuracy, precision, and recall to cutting-edge methods. By combining several classifiers and feature selection techniques, we want to assess the classifier's selecting scheme in future work.

References

1. Vishnuvarthanan A, Rajasekaran MP, Govindaraj V, Zhang Y, Thiyagarajan A (2017) An automated hybrid approach using clustering and nature inspired optimization technique for improved tumor and tissue segmentation in magnetic resonance brain images. Appl Soft Comput 57:399–426
2. Alfaer NM, Aljohani HM, Abdel-Khalek S, Alghamdi AS, Mansour RF (2022) Fusion-based deep learning with nature-inspired algorithm for intracerebral haemorrhage diagnosis. J Healthcare Eng
3. Bharti V, Biswas B, Shukla KK (2020) Recent trends in nature inspired computation with applications to deep learning. In 2020 10th international conference on cloud computing, data science & engineering (confluence). IEEE, pp 294–299. (2020 January)
4. Kurdi SZ, Ali MH, Jaber MM, Saba T, Rehman A, Damaševičius R (2023) Brain tumor classification using meta-heuristic optimized convolutional neural networks. J Pers Med 13(2):181
5. Kumar SN, Lenin Fred A, Ajay Kumar H, Sebastin Varghese P (2019) Firefly optimization based improved fuzzy clustering for CT/MR image segmentation. Nature inspired optimization techniques for image processing applications, pp1–28
6. Jothi G (2016) Hybrid tolerance rough set-firefly based supervised feature selection for MRI brain tumor image classification. Appl Soft Comput 46:639–651
7. Abdulameer AT (2018) An improvement of MRI brain images classification using dragonfly algorithm as trainer of artificial neural network. Ibn AL-Haitham J Pure Appl Sci 31(1):268–276
8. Kanmani P, Marikkannu P (2018) MRI brain images classification: a multi-level threshold based region optimization technique. J Med Syst 42:1–12
9. Amin J, Sharif M, Raza M, Saba T, Anjum MA (2019) Brain tumor detection using statistical and machine learning method. Comput Methods Programs Biomed 177:69–79
10. Elkorany AS, Elsharkawy ZF (2020) Automated optimized classification techniques for magnetic resonance brain images. Multimed Tools Appl 79(37–38):27791–27814
11. Neelima G, Chigurukota DR, Maram B, Girirajan B (2022) Optimal DeepMRSeg based tumor segmentation with GAN for brain tumor classification. Biomed Signal Process Control 74:103537
12. Lu SY, Wang SH, Zhang YD (2020) A classification method for brain MRI via MobileNet and feedforward network with random weights. Pattern Recogn Lett 140:252–260
13. Jayachandran A, Jegatheesan A, SreekeshNamboodiri T (2021) MRI brain image classification system using super pixel color contrast and support vector neural network. In Intelligence in big data technologies—beyond the hype: proceedings of ICBDCC 2019. Springer, Singapore, pp 353–360
14. Shivhare SN, Kumar N (2021) Tumor bagging: a novel framework for brain tumor segmentation using metaheuristic optimization algorithms. Multimed Tools Appl 80(17):26969–26995
15. Sidhu BK (2019) An optimize learning approach for brain tumor classification in medical applications. Think India J 22(14):858–866
16. Dhal KG, Rai R, Das A, Ghosh TK (2023) Hybridization of sine-cosine algorithm with k-means for pathology image clustering. In Artificial intelligence: first international symposium, ISAI 2022, Haldia, India, 17–22 Feb 2022, Revised Selected Papers. Springer Nature Switzerland, Cham, pp 76–86 (2023, January)
17. Kapoor A, Agarwal R (2021) Enhanced brain tumour MRI segmentation using K-means with machine learning based PSO and firefly algorithm. EAI Endorsed Trans Pervasive Health Technol 7(26):e2–e2

18. Malik A et al (2022) Deep learning versus gradient boosting machine for pan evaporation prediction. Eng Appl Comput Fluid Mech 16(1):570–587. https://doi.org/10.1080/19942060.2022.2027273
19. Jeyajothi ES, Anitha J, Rani S, Tiwari B (2022) A comprehensive review: computational models for obstructive sleep apnea detection in biomedical applications. BioMed Res Int 2022; 2022: Article No.: 7242667. https://doi.org/10.1155/2022/7242667
20. Khalaf OI et al (2022) Blinder Oaxaca and WilkNeutrosophic fuzzy set-based IoT sensor communication for remote healthcare analysis. IEEE Access. https://doi.org/10.1109/ACCESS.2022.3207751
21. Kaur T, Saini BS, Gupta S (2019) An adaptive fuzzy K-nearest neighbor approach for MR brain tumor image classification using parameter free bat optimization algorithm. Multimed Tools Appl 78:21853–21890
22. Mandle AK, Sahu SP, Gupta G (2022) Brain tumor segmentation and classification in MRI using clustering and kernel-based SVM. Biomed Pharmacol J 15(2):699–716