# Kernel Nearest Neigh-Bour Based Genetic Algorithm And Modified Kernel-Based Fuzzy C-Means Based MRI Image Brain Tumor Segmentation And Classification

**Dr.B.Balakumar. Dr. P. RAVIRAJ**

**Abstract:** The recognition of a mind cancer and its grouping from present day image modalities is an essential worry, yet a tedious and dreary work was achieved by radiologists or medical bosses. The precision of identification and tumor stages grouping done by radiologists is relied upon their knowledge just, so the PC supported innovation is imperative to help with the conclusion exactness. In this research, a tumor portion is segmented from the brain image by using Adjusted Kernel-Based Fuzzy C-Means (MKFCM) algorithm. The input images are resized like 256×256 in the pre - processing stage. The pre-processed MRI image segmented by MKFCM, which is a stretchy top ML system to locate the object in a difficult pattern. Next, Hybrid Feature Extraction (HFE) performed on the segmented image to increase the feature subsets. The feature selection (FS) process was performed by Kernel Nearest Neighbour (KNN) based Genetic Algorithm (GA) in order to acquire the best feature values. The best feature values given to the Naive Bayes (NB) classifier as an input, which is classified into Meningioma, Glioma and Pituitary regions in the MRI images. The performance of proposed KNN-GA-FS-NB method is validated by T1-WCEMRI dataset.

**Index Terms**: Modified Kernel-Based Fuzzy C-Means (MKFCM), Hybrid Feature Extraction (HFE), Kernel Nearest Neighbour (KNN) based Genetic Algorithm (GA) and Naive Bayes (NB).

## 1. INTRODUCTION

Brain downpour is the furthermost composite part in the group of individual, the unbounded of cells inbuilt in cerebrum. A consideration tumor rise when there is unreasonable parcel of cells demonstrating a strange social affair of cells intently the cerebrum. That social event of cells can impact the conventional value of the cerebrum activity and beat the strong cells [1-2]. Mind tumors assembled to merciful or below average and destructive tumors or high-grade. Liberal tumors are non-dynamic so saw as less strong, they started in the brain and grows steadily; in like manner it can't degree to wherever else in the body. Regardless, perilous tumors are destructive and grow rapidly with ill defined breaking points. They can be begun in the cerebrum that one which called basic undermining tumor or to be started anyplace else in the physical make-up and degree to the mind which called helper risky malignancy [3-5]. Best method to see the picture is the mind MRI that scientists depended on for recognizing the insight developments and displaying of the disease movement in together the revealing and the dealing with stages.


- [1]Assistant Professor,
- Centre for Information Technology and Engineering,
- Manonmaniam Sundaranar University,
- Tirunelveli – 627 012.

- [2]Professor, Department of CSE,
- GSSS Institute of Engineering and KRS Road, Metagalli,
- Mysore, Karnataka-570016.


X-beam pictures have a significant impact in the modified remedial picture examination field for its ability to give a huge amount of records about the cerebrum development and varieties from the standard inside the mind tissues since of the significant standards of the photos [6-8]. Believe it or not, Researchers showed differing mechanized methodologies for mind tumors revelation and type request utilizing cerebrum MRI films since it got possible to yield and load remedial pictures to the framework. In any case, SVM and NN are the comprehensively used procedures for their extraordinary execution over the span of the best late join of periods [9]. Regardless, starting late, (DL) models set a stimulating example in AI as the profound plan can proficiently address troublesome associations without requiring uncountable center points like in the shallow courses of action for instance SVM and K-nearest neighbor (KNN). Subsequently, they grew rapidly to undercover top tier in different prosperity informatics domains, for instance, bioinformatics, therapeutic informatics and remedial picture assessment [10].

## 2 LITERATURE SURVEY

Different investigates are endorsed via analysts as a pinnacle priority tumor (BT) discovery. On this circumstance, the brief exam of some focal obligations to the existing insightful fills in as advanced. Li, Haiyang et al. [11] provided a framework for MIS, referred to as Dynamic PSO and K-construes grouping structure (DPSOK). (DPSO) and K-prescribes bundling machine became the base structure of the DPSOK. They made DPSOK an everyday streamlined envisioning by using fresh the preparing technique for that one inaction importance and statistics elements Research consequences validated that DPSOK figuring can beneficially build up the K-assembles structure's general inquiring for limit. DPSOK assessment surpassed on excellent yields in creating photo department really worth and ability confirmed up differently in affiliation with fashionable PSO K-concludes game plan. Ronghua Shang et al. [12] delivered an unmatched FCM shape replica piece three-dimensional FCM (CKS-FCM). An ensured clone device was utilized to streamline the fundamental p.C. work environments, which permits the mixture as a rule introduction. The spatial bits of facts are fused within the goal highlight and CKS-FCM applied a non-Euclidean separation on an extremely fundamental level dependent on bit spotlight to restore the Euclidean parcel. The fundamental repression of the shape is Low parcel exactness and Less liberality. Y.T. Chen et al. [15] have proposed Independent part assessment primarily based segment zed padded by making use of MIS. They have examined the execution partition of 6 strategies (k-proposes, FCM, KFCM, ICFCM, KWFLICM, and ICKFCM) for the imitated MRI movies in calm case, disturbance case, and proper medicinal pictures. The pivotal weight of the framework is much less precise. The above territories gives an apparent point of view on the techniques that have been planned remarkably to secure the department location of interest, multiple frameworks for removing capabilities and a few to get equipped and check using the classifiers for portrayal in a way of speaking. Much convincing department with the combined segment extraction couldn't be driven, and just particularly few features had been removed which executed low precision in tumor singular insusceptible and revelation. The classifiers used to installation the capabilities also are subsequent to no convincing.

## 3 KNN-GA-FS-NB SYSTEM

The major impression of the KNN-GA-FS-NB method is to increase system's accuracy in head tumor recognition by using KNN-GA-FS-NB techniques. The proposed brain cancer subdivision methodology has five steps, those are 1.preprocessing, 2.segmentation, 3.feature extraction, 4.feature selection, and classification. The KNN-GA-FS-NB method's block diagram is represented in Fig. 1. A brief description of KNN-GA-FS-NB technique is determined below.

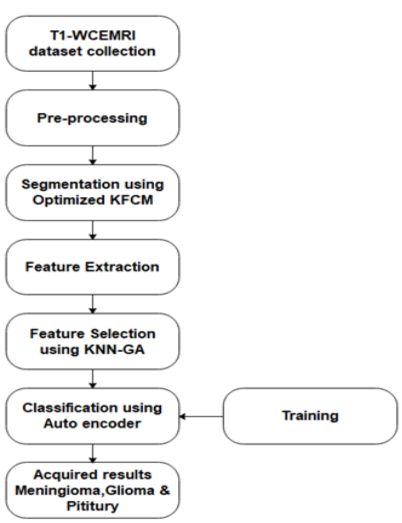


***Fig.1**. KNN-GA-FS-NB flow diagram.*

### 3.1. T1-WCEMRI data set

From 2005 to 2010, the cerebrum T1-gauged CEMRI dataset was gained from Guangzhou, Nanfang Clinic, China, Tianjin Medical University, and General Hospital, China. This dataset made out of 3064 cuts from 233 injured persons, containing, 1426-Gliomas, 708-meningiomas and 930-pituitary tumors. The pictures have an enplane goals of 512×512 with 0.49×0.49 mm2 pixel size. The size of cut has 6 mm thickness and the cut hole size is 1 mm. The tumor outskirt was physically clarified by 3-worked radiologists. The T1-WCEMRI dataset have three modules of cerebrum pictures, those are Meningioma, Glioma and Pituitary tumor. A portion of the T1-gauged CEMRI dataset pictures are uncovered in Fig.2.

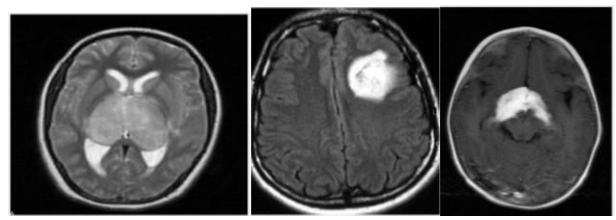

***Fig.2**. T1-weighed CEMRI dataset images.*

### 3.2. Pre-processing and Segmentation

In this research, the T1-W CEMRI dataset is cast-off to calculate the efficiency of the KNN-GA-FS-NB method. Input images are resized like the 512×512. The PP imageries are utilized to the segmentation process. Tumour images generally consist of normal and abnormal segments. Identification of normal and abnormal is critical to accurately extract the features. The brain tumor (BT) can be identified using segmentation techniques. The KNN-GA-FS-NB system consider Hybrid kernel-based FCM framework is represented as HKFCM. In this system first, we compute the Flexible parameter〖 φ〗_a related with all pixel to control the relative statistics by (1). The impartial meaning is defined.

$$JMKFCM = 2[\sum^{i} a=1 \sum^{c} b=1 u^{m} ab(1-K(m_a,v_b)) + \sum^{i} a=1 \sum^{c} b=1 \phi_a u^{m}{}_{ab}(1-K(m_a,v_b \tag{1}$$

The minimization of $J_{HKFCM}\ (m,v)$ can be considered through an alternative optimization technique using

$$u_{ij} = \frac{((1-K(m_a,v_b)) + \varphi_a(1-K(\bar{m}_a,v_j))^{-1/\ (m-1)}}{\sum_{a=1}^{N} u_{a,b}^{m}(K(m_a,v_j) + \varphi_a(1-K(\bar{m}_a,v_j))^{-1/\ (m-1)}} \tag{2}$$

$$v_a = \frac{\sum_{a=1}^{N} u_{a,b}^{m}(K(m_a,v_b)m_a + \varphi_a K(m_a,v_j)m_a)}{\sum_{a=1}^{N} u_{a,b}^{m}(K(m_a,v_b)m_a + \varphi_a K(\bar{m}_a,v_j))} \tag{3}$$

When $\bar{x}$ is substituted with the monochromic of the normal filter of the idea image, the set of code is represented as HKFCM. When $\bar{x}_i$ is changed with the weight image $\bar{\varepsilon}_i$ defined. The algorithm is represented as $\mathrm{HKFCM}_{\omega}$. The main step for the HKFCM is mentioned below:

HKFCM algorithm

(1) Initialized threshold $\varepsilon = 0.001, m = 2,$ loop counter t $t = 0, v \ and \ u^{(0)}$

(2) Calculate the Flexible regularization parameter $\varphi_i$.

(3) Calculate $\bar{x}_i$ for HKFCM

(4) Calculate cluster centres $v_j^{(t)}$ using $u^{(t)}$ as in (5).

(5) Calculate the membership function $u^{(t+1)}$ within (4)

(6) If max $||u^{(t+1)} - u^t| < \varepsilon \ or \ t > 100$ then stop; otherwise, update $t = t + 1$ and go to step (4).

Some of the T1-weighed CEMRI dataset images clustering and segmentation images are exposed in the below figure 4. And 5.

***Fig.3**. T1-weighed CEMRI dataset cluster Image.*

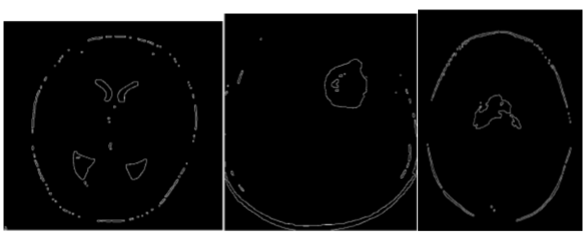

***Fig.4**. T1-weighed CEMRI dataset HKFCM segmented Image.*

### 3.3. Hybrid feature extraction

In KNN-GA-FS-NB method, the HFE performed on the transformed BT image. The feature withdrawal is the act of recording the spitting copy from image space to feature space. The GLCM is the well-known texture investigation method, which computes the image features linked to second directive measurements of the idea image and DWT process image. This research considered two types of GLCM texture features called homogeneity and energy. The homogeneity procedures the nearness of distribution geographies of the gray-level matrix. To quantitatively describe the same texture portions for comparison, the Limited Latitudinal Statistics (LSS) of the texture are computed using orientation and scale selective of Gabor. Energy is calculated grounded on the amount of recurrent couples and also the similarity is measured founded on the regularized pixel couple distributions. This energy supports to return the depth and softness of the BT image texture construction. In this research, the best article is particular by using LBP descriptor. Here, the charge of the significant pixel is employed as a threshold. The normalized histogram can be computed and employed as the features. The histogram provides the information related to edges and spread of the discontinuities of the duplicate. The LBP is most popular because it eases to calculate features and employed in many acknowledgment presentations. Additionally, the HOG descriptor is smeared to the segmented image processing and computer version is charity for mining the optimum feature standards. These ideals are specified to the FS block as an input for selecting the best feature, which is designated in the following section.

### 3.4. Feature extraction (Hybrid feature extraction)

After segmentation process, highlights pulling out is a significant advance in any grouping issue. Highlights contain applicable data required to recognize various classes. Surface properties of a picture container can be used for order reason. Surface contains data about the auxiliary disposed strategy of surfaces in a depiction. In this work, the underlying phase of the component extraction is wavelet highlights are extricated from each sectioned pictures. With the assistance of wavelet highlights (GLCM) is applied and the element esteems are extricated. Highlights register's such us, Autocorrelation, Contrast, Inverse distinction, Transformation etc. [24].

### 3.5. FS using KNN-GA

The FS method selects the greatest geographies from the HFE. The classification involves consideration of the dataset already transferring the data to a classifier. It is recommended to consider only an important feature for selection. Hence, it is useful for selecting the important and related features in this BT detection. Furthermore, the FS is used for the period of the arrangement to invention the significant article that reductions the classifier tasks and enhances the grouping correctness. In this research, KNN founded GA minimizes the severance within input voxels and defines more relevance between input /output voxels. Then, the qualification meaning is calculated for input voxel subsets by using KNN that maximizes the mutual data between the voxel. Lastly, the crossover and mutation operators are utilized to detect the voxels by reducing the redundancy-based (FF). The GA chooses the subsection of structures as the chromosomes and all chromosome is directed to the KNN for calculating fitness value. The KNN employs every single chromosome as a mask for catching the features. The KNN describes a suitability value of every single chromosome and GA uses these suitability tenets for the (DNA or RNA) computation method. At the end, the GA search a best subsection of the feature.

### 3.6. Classification using Naive Bayes

NB forecasts provisionally liberated class after a presumed period tags are specified as problematic of cases and then adjusts the article direction keen on feature standards. Now, giving a set of stranger data tuples, in that all are display a $n-$dimensional vector, $X = \{x_1, x_2, ..., x_n\}$ showing $n$ dimensions are complete to $n$ attributes of the tuple. Set of $m$ classes, $C_1, C_2, ....C_m$.Using Bayes theorem, the NB computes the subsequent possibility of class, acclimatized on $X$ and after it allocates the lesson tag on each class with extreme posterior probability. Therefore, probability is assumed by eq (4),

$$P(C_x \mid X) > P(C_y \mid X) \quad for \quad 1 <= y <= m, y \neq x \tag{4}$$

An Extreme posterior hypothesis can be consumed to attain an opinion approximation of an unnoticed magnitude on the beginning of empirical data. Maximize the worth of $P(X \mid C_x)P(C_x)$ with constant $P(X)$ lot of attributes; it is computationally risk to approximation the prospect of predicting $P(X \mid C_x)$. Maximum posterior Probability of equation is assumed by eq.(5),

$$P(C_x \mid X) = \frac{P(X \mid C_x)P(C_x)}{P(X)} \tag{5}$$

Where, $P(C_x)/X$ denotes class posterior probability, $P(C_x)$ is characterized as the class prior probability, $P(X \mid C_x)$ is characterized as the probability of predictor class, $P(X)$ is characterized as the prior probability of guessing.

$$P(X_k \mid C_x) \approx \prod_{k=1}^{n} P(X_k \mid C_x) \tag{6}$$

$$G(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{7}$$

If continuous-valued is $A_k$, at that point normally accept that the qualities have a Gaussian dissemination with a uncaring $\mu$ & STD $\sigma$ is assumed by Eq (8),

$$P(x_k \mid C_x) = g(x_k, \mu_{C_x}, \sigma_{C_x}) \tag{8}$$

Calculate the values of $\mu C_x$ ad $\sigma C_x$, which are measured as the mean and SD standards of quality $A_k$ for teaching examples of class $C_x$. Applying NB has display an important correctness and big database.

## 4. Results and Discussion

The KNN-GA-FS-NB technique was fulfilled in the MATLAB stimulator software version 2018b. The entire work implemented by using I3 system with 4 GB RAM. The performance criterion conceded out by using some T1WMRI scans. The T1 WMRI scans were selected from the T1-WCEMRI dataset. The enactment of the KNN-GA-FS-NB technique was instigated in relations of specificity, accurateness, sensitivity, and F-measure. To calculate the cataloguing efficiency of the KNN-GA-FS-NB method, the presentation of the KNN-GA-FS-NB technique is associated with conservative approaches with the same reputed dataset: T1-WCEMRI.

### A. Performance analysis

The challenge evaluation metrics is recycled for assessing the both splitting up and sorting performance of our method. For the segmentation, the evaluation criteria include (SE), specificity (SP), (AC), Jaccard index (JSI) and Dice coefficient (DSC). The performance standards are definite is as:

$$SE = \frac{tp}{tp + fn}$$

$$AC = \frac{tp + tn}{tp + fp + tn + fn}$$

$$P = \frac{tp + tn}{tp + tn + fp + fn}$$

$$SP = \frac{tn}{tn + fp}$$

$$DSC = \frac{2TP}{(2TP+FP+FN)}$$

$$JSI = \frac{TP}{FP+FN+TP}$$

There are four evaluation criteria, including sensitivity (SE), specificity (SP), accuracy (AC) Recall (R) and Precision (P).

**B**. Segmentation Performance

In this division, The KNN-GA-FS-NB system T1WMRI dataset images segmentation enactments are appraised some of the evaluated results are tabularized in the below table.1. The KNN-GA-FS-NB segmentation result are calculated in relations of Jaccard index (JSI) and Dice coefficient (DSC).

***Fig.5**. (a) Image-1.*

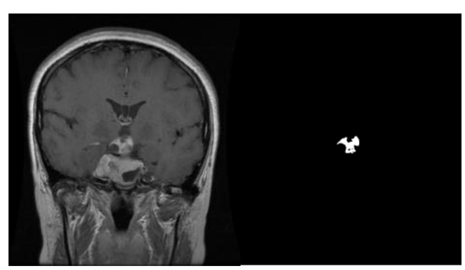

***Fig.5**. (b) Image-2.*

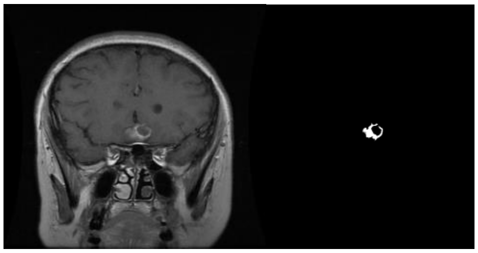

***Fig.5**. (c) Image-3.*

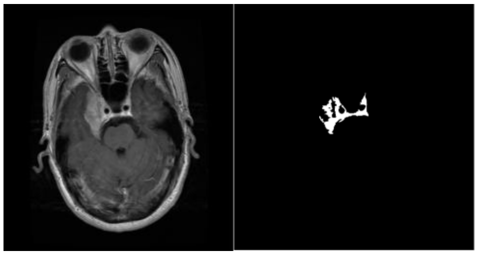

***Fig.5**. (d) Image-4.*

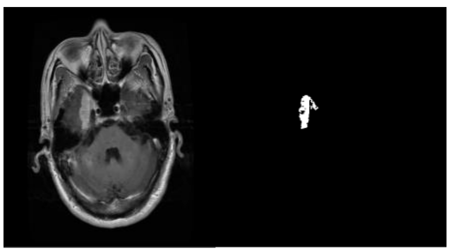

***Fig.5**. (e) Image-5.*

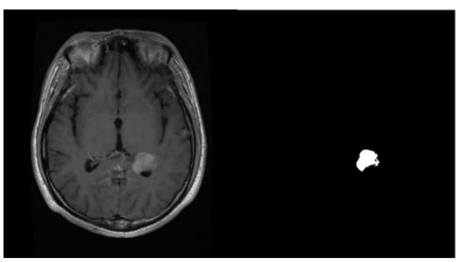

***Fig.5**. (f) Image-6.*

***Tab.1. T1WMRI scans dataset images Segmentation performances.***

| Title | DSC | JSI |
|---|---|---|
| Image-1 | 0.6312 | 0.4611 |
| Image-2 | 0.2741 | 0.1588 |
| Image-3 | 0.5915 | 0.5915 |
| Image-4 | 0.4590 | 0.2978 |
| Image-5 | 0.7181 | 0.7181 |
| Image-6 | 0.8606 | 0.7553 |

### C. Classification Performance

In this section, the KNN-GA-FS-NB system has been likened with the various classifiers and various optimization techniques. Which are tabulated in the Tabe.2. The KNN-GA-FS-NB system classification presentations are appraised in relations of SE, SP and AC.

***Tab.2**. Classification Performance.*

| Method | SE (%) | SP (%) | AC (%) |
|---|---|---|---|
| Without FS SVM | 72.25 | 83.24 | 83.24 |
| PSO FS with SVM | 100 | 94.65 | 83.55 |
| KNN-GA FS SVM | 92.11 | 100 | 92.00 |
| Without FS NB | 76.38 | 86.25 | 84.34 |
| PSO FS NB | 89.45 | 96.56 | 89.98 |
| KNN-GA-FS-NB | 96.11 | 100 | 97.44 |

### D. Comparative analysis

The KNN-GA-FS-NB system has compared with two different existing systems which are explained in below. And the relative examination is tabulated in table.3. N. Abiwinanda, et.al [16] proposed an awareness tumor order utilizing CNN. They ordered the cerebrum tumors into three sorts: Glioma, Meningioma, and Pituitary by utilizing CNN. They actualized the most straightforward conceivable engineering of the CNN (for instance, one every one of complication, max pooling and leveling covers pursued by whole association here single concealed layer). The CNN was prepared on a cerebrum tumor dataset comprising of 3064 T1 weighted CEMRI pictures. By and large, the routine of the precision is significant in mind tumor discovery. Be that as it may, the precision of the future method isn't assessed. J. Cheng, et.al [17] planned a technique to expand the exhibition of tumor arrangement. At first, they utilized enlarged growth area by means of picture widening as the ROI rather than the first tumor divide since tumor encompassing tissue can likewise give huge insights of a tumor type. In the subsequent stage, enlarged tumor segment split into various fine ring structure sub partitions. They assessed the productivity of this strategy dependent on an enormous database with 3-highlight extraction strategies: GLCM, power histogram and BoW. Be that as it may, SVM classifier doesn't accomplish the proficient exactness as a result of the inappropriate determination of the portion.

***Tab.**3.Comparative analysis.*

| Authors | Database | Tumour Classification Accuracy (%) |
|---|---|---|
| N. Abiwinanda, et.al [16] | T1-W CEMRI | 84.16 |
| J. Cheng, et.al [17] | T1-W CEMRI | 91.14 |
| KNN-GA-FS-NB | T1-W CEMRI | 97.44 |

The MRI brain image segmentation was performed using MRKFCM. In instruction to discovery the classification accuracy, the various texture and gradient orientation features mined by GLCM, HOG, and LBP techniques. The NB is utilized to categorize the abnormal brain image into Meningioma, Glioma and Pituitary. It can be concluded from the investigational outcomes that the KNN-GA-FS-NB method is more suitable for abnormal BT discovery and arrangement equated to present methods in rapports of accuracy.

## 5 CONCLUSION

Image segmentation process theatres a significant part in remedial image dispensation. In this work, the MRKFCM segmentation technique was employed to section the BT share of the mind image. The HFEs (GLCM, HOG, and LBP) extracted the optimal feature values from segmented region. Then, the feature range technique was applied to the information abstraction data for best feature selection. Finally, the suspicious portions were classified by using NB classifier based on selected features. The KNN-GA-FS-NB method delivered the accuracy rate is 97.44%, which is much promising to recognize the abnormal BT classification from patient MRI image, and the KNN-GA-FS-NB method can be recycled to categorize the various types of tumour according through medical diagnosis system.

## REFERENCES

[1] Bauer, Stefan, Lutz-P. Nolte, and Mauricio Reyes. "Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization." international conference on medical image computing and computer-assisted intervention. Springer, Berlin, Heidelberg, 2011.

[2] Logeswari, T., and M. Karnan. "An improved implementation of brain tumor detection using segmentation based on hierarchical self-organizing map." International Journal of Computer Theory and Engineering 2.4 (2010): 591.

[3] Verma, Ragini, et al. "Multiparametric tissue characterization of brain neoplasms and their recurrence using pattern classification of MR images." Academic radiology 15.8 (2008): 966-977.

[4] Hou, Zujun. "A review on MR image intensity inhomogeneity correction." International journal of biomedical imaging 2006 (2006).

[5] El-Dahshan, El-Sayed A., et al. "Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm." Expert systems with Applications 41.11 (2014): 5526-5545.

[6] Zacharaki, Evangelia I., Vasileios G. Kanas, and Christos Davatzikos. "Investigating machine learning techniques for MRI-based classification of brain neoplasms." International journal of computer assisted radiology and surgery 6.6 (2011): 821-828.

[7] Viergever, Max A., et al. "A survey of medical image registration–under review." (2016): 140-144.

[8] Singh, Lavneet, and Girija Chetty. "Detecting The Brain Abnormalities From MRI Structural Images Using Machine Learning And Pattern Recognition Tools." IJREISS 2.11 (2012): 15-30.

[9] Singh, Lavneet, and Girija Chetty. "Detecting The Brain Abnormalities From MRI Structural Images Using Machine Learning And Pattern Recognition Tools." IJREISS 2.11 (2012): 15-30.

[10] Ravì, Daniele, et al. "Deep learning for health informatics." IEEE journal of biomedical and health informatics 21.1 (2016): 4-21.

[11] Li, Haiyang, Hongzhou He, and Yongge Wen. "Dynamic particle swarm optimization and K-means clustering algorithm for image segmentation." Optik 126.24 (2015): 4817-4822.

[12] Shang, Ronghua, et al. "A spatial fuzzy clustering algorithm with kernel metric based on immune clone for SAR image segmentation." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 9.4 (2016): 1640-1652.

[13] Elazab, Ahmed, et al. "Segmentation of brain tissues from magnetic resonance images using adaptively regularized kernel-based fuzzy-means clustering." Computational and mathematical methods in medicine 2015 (2015).

[14] Liu, Li, et al. "Robust dataset classification approach based on neighbor searching and kernel fuzzy c-means." IEEE/CAA Journal of Automatica Sinica 2.3 (2015): 235-247.

[15] Chen, Yao-Tien. "Medical image segmentation using independent component analysis-based kernelized fuzzy-means clustering." Mathematical Problems in Engineering 2017 (2017).

[16] Kumar, Vemula Vinay, Kodali Sai Krishna, and Sarukonda Kusumavathi. "Genetic Algorithm Based Feature Selection Brain Tumour Segmentation and Classification.

[17] Cheng, Jun, et al. "Correction: Enhanced Performance of Brain Tumor Classification via Tumor Region Augmentation and Partition." PloS one 10.12 (2015): e0144479.