

Genetic Algorithm Based Fuzzy Local Information-C-Means (Gaflicm) Clustering Algorithm And Hybrid Kernel Convolution Neural Network (Hkcnn) With Distributed Processing Framework For Brain Mri Images

¹Vigneshwari K, ²Dr.Kalaiselvi K,

Department Of Computer Applications,
Vels Institute Of Science, Technology And Advanced Studies
Chennai, India, Email-Id: Vigneshwari.Raman10@Gmail.Com
Department Of Computer Applications, Vels Institute Of Science, Technology And Advanced Studies,
Chennai, India, Email-Id: Kalairaghu.Scs@Velsuniv.Ac.In

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

ABSTRACT: Cancer is the large important origin of casualty in today world. Among many cancers, brain cancer has been becomes one of the lowest living rate. It is formed based on the brain tumor. However brain tumors are able to have diverse categories based on their shape, texture, and position. Appropriate identification and retrieval of the tumor types create potential the doctor to compose the right cure option and assist keep the patient's life. Image processing has gained wide attention in medical analysis and health in recently. In general image processing methods, brain Magnetic Resonance Imaging (MRI) image collections cannot be processed efficiently on one computer due to large collection sizes and high computational costs. Hence, parallel computing and distributed system has been performed increasingly for brain MRI images in recently. In this paper, a novel Medical Image Cloud Processing (MICP) based distributed processing framework is proposed for brain MRI images by lesser computational time. In this work, image preprocessing is done by using the Adaptive Median Filtering (AMF) and image enhancement by Histogram Equalization (HE). The proposed MICP framework includes of Static Medical Image Cloud Processing (SMICP) and Dynamic Medical Image Cloud Processing (DMICP). In MICP framework, SMICP consists of two methods called Pure-Image and Big-Image. These methods are integrated to Genetic Algorithm based Fuzzy Local Information C-Means-MapReduce (GAFLICM-MR) algorithm to attain more optimized design and higher effectiveness. The core design of GAFLICM-MR framework is to make use of the rich computing resources given by means of the distributed system consequently as to apply efficient parallel processing. GAFLICM algorithm is also used as brain tumor segmentation in MRI images. In MICP framework, DMICP is developed via a parallel processing process of the distributed system. For retrieval and detection of brain Magnetic MRIimages into normal and tumor, Hybrid Kernel Convolution Neural Network (HKCNN) is developed in this work. Finally the results of the HKCNN classifier are compared to other previous works like precision, recall, f-measure, accuracy, time and memory.

INDEX TERMS: Brain tumor, Magnetic Resonance Imaging (MRI), Classification, Big data, Image processing, MapReduce, Distributed system, Image Cloud Processing (ICP), Genetic Algorithm based Fuzzy Local Information C-Means (GAFLICM), Hybrid Kernel Convolution Neural Network (HKCNN).

1.INTRODUCTION

Digital Image Processing (DIP) is a well-developed part in the biological sciences which involves classification and detection of tumour, classification and detection of cancer, and testing and clear examination of defected areas of the human body. In medical Science, Automatic brain tumor diagnosis is an important phase. [1]. The human body consists of various cells in which the brain (known as the body's processor) plays a vital role [2]. The brain's important part is the nervous system which is considered as the most important part of the human central nervous organism. A tumor is inlaid within the brain, malfunctioning of parts of the body. It is tedious process to identify its location and treat and arrest its capability of spreading [3-4]. Generally it may be either benign or malignant. The diagnosis carried out in initial phase and the brain tumour recognition is very important. Brain tumor detection is performed by Computer-Aided Diagnosis (CAD) systems [6]. The unnatural growth of tissues is called a brain tumor or the central spine which be able to disrupt the suitable functioning of the brain [7].The diagnostic process of a tumor should be carried out rapidly and are generated by means of MR image segmentation from complicated medical images. Suitable medical imaging system is MRI technology which is used to extract high standard images of human body, which plays a significant role in treating patients at the correct stage for tumor- affected patients.

Various techniques have been formulated for the purpose of classification of brain tumours that are extracted through MRI images, which involve Fuzzy Clustering Means (FCM), Support Vector Machine (SVM), Artificial Neural Network (ANN), and Expectation-Maximization (EM) algorithm are few noticeable techniques for segmentation based on region in order to gather important data on medical imaging modalities.

With these algorithms, the data storage and efficiency of processing have been raised as a critical issue. All these efforts are made on focussing on the optimization of image processing algorithms, whereas obviating the intrinsic insufficiency of the single node processing process. Hence in the previous works [8-10] many enhancements are to be made so as to minimize the complexities that are image processing are prone to, and its show is generally restricted towards a low level because of its ineffective processing carried on a single device. Though there may exist few issues like consumption of huge time, various experts segmenting MRI image distinctly. The result of tumour may also differ on various situations when consulted with the same physician; on the other hand there are some issues those are considering an increased computation time, and segmentation of MRI image with varied proficient might differ significantly. Furthermore, the outcome of tumor detection might differ in diverse conditions with the similar physician, and also the contrast and brightness of the image brain huge difference in the segmentation results. From the abovementioned issues, the handling of increased time plays a vital factor in this regard.

To manage these issues, Yan et al [11] constructed a Hadoop framework with the help of OpenCV library for processing the image; it is also modified as the speed is high for huge-sized files. Sweeney et al [12] did implementation on the HIPI library. At the same time as Hadoop Archive (HAR) files are able to be utilized as a file archive in HIPI library, which produces slow performance as the technique is used for getting the files with the purpose of lie inside. On the other side, HIPI is liable for serial reading process with the same speed as that of sequence files. Sozyki et al [13] developed MapReduce Image Processing framework (MIPr). It was proved that HIPI produced good results, when considering the time factor so as to carry out the task. Li et al [14] proved that the efficiency of Hadoop on greater amount of little range files which is lower than the fewer amount of big size files. These methods prove that the recent work considers greater efficiency in parallel processing, in which the researchers tries to formulate many algorithms on image processing with the intention be able to be developed in parallel manner with the steps like image classification [15], feature extraction and matching [16] be able to act as representative instances. Many of these methods are able to support several nodes run parallel and therefore high time efficiency is possible.

In spite of having all of these frameworks, they are successfully implemented in text data and stream data, but they are not widely implemented in medical image processing area. It motivates to employ a parallel processing framework via MapReduce (MR) for huge medical image data. In the proposed work, a framework for a novel effective Medical Image Cloud Processing (MICP) distributed system is formulated to an efficient model for brain tumor detection as either benign and malignant tumors accurately. MRI brain images are preprocessed via Adaptive Median Filtering (AMF) and Histogram Equalization (HE) for image enhancement. In MICP framework, Static Medical Image Cloud Processing (SMICP) is performed with two representations such as Pure-Image and Big-Image are implemented via Genetic Algorithm based Fuzzy Local Information C-Means -MapReduce (GAFLICM-MR) algorithm for image segmentation which achieves higher efficiency. In MICP, Dynamic Medical Image Cloud Processing (DMICP) is developed via a parallel processing process of the distributed system. For retrieval and detection of brain tumors, Hybrid Kernel Convolution Neural Network (HKCNN) has shown an optimistic performance in Computer-Aided Diagnosis (CAD) system. The results are also illustrated that benign tumors have more irregular appearance than malignant tumors and gives higher results.

2.LITERATURE REVIEW

Alam et al [17] proposed a Template-based K-means and improved Fuzzy C Means (TKFCM) for brain tumor detection in MRI image. At first, TKM is developed for segmentation considerably during the great choice of a template depending on gray-intensity of image. Further, new membership is computed via Euclidean distances between cluster center and samples from FCM clustering. Lastly TKFCM is used in the detection of position of tumor via TKM and FCM. The outputs of the simulation prove that the formulated algorithm gains a better diagnosis of usual and irregular tissues in the brain with the lesser detachment of intensity in the gray – level. Future work determination analyzes for decreasing the needed computational time and changes the complexity.

Xie [18] proposed a modified K-Nearest Neighbor (KNN) for brain tumor detection. The removal of the skull tissue is considered as the initial stage of the work which is done by pre-processing, that is carried out after the process of extraction and segmentation process. Minkowski distance which is a generalized form of Euclidean distance has been utilised in the proposed work and the segmentation results indicate an increased accuracy of nearly 1.5% over FCM and reduction in computation time as result of reduced iterations. The experimental result proves with the purpose of the proposed methodology works improved when compared with the other conventional methodologies.

Mathur et al [19] proposed a fuzzy-based K-means clustering for detection of brain tumor in a MRI image. Proposed approach for accurate detection of brain tumor of edges of MRI images. The segmentation process of the MRI image of the brain adds up the issue of identification of anatomical ROI, that aid the radiologists to collect the structural features like shape, appearance for the detection of disease or the evaluation

of the treatment. The brain image segmentation is tumor is located. The method of local threshold is utilized in order to identify the threshold of K-means segmentation.

Rehman et al [20] proposed a deep Convolutional Neural Network (CNN) for automatic brain tumor detection in MRI image and features are extracted automatically from the input space. In the formulated work, the 3 studies are conducted using 3 architectures in CNNs which includes AlexNet, GoogLeNet, and VGGNet, which are used in the classification of meningioma, glioma, and pituitary among the brain tumour. The techniques of data augmentation are implemented in the MRI slices for results that are to be generalized by maximizing the samples in the dataset and thereby minimizing the over-fitting possibilities. Deep CNN and transfer learning approaches has improves higher performance than the CNN which is left as scope of this work. Since deep CNN architectures for brain tumor classification gives lesser time complexity.

Wang et al [21] formulated a cascade of CNNs towards segment brain tumors by means of hierarchical subregions from multi-modal MRI, and 2.5D network to be a substitution among memory consumption, time complexity and open access. Added to it, the augmentation of employ test-time in order to gain enhanced accuracy in segmentation provides voxel-wise and structure-wise data that is unforeseen as a result of segmentation. An experiment with Brain Tumor Segmentation (BRATS 2017) dataset showed that cascade of CNNs with 2.5D CNNs performs better than the usual CNN methods.

Siar and Teshnehlab [22] proposed a CNN for brain tumor detection of MRI images. Images are feed into to the CNN classifier. From the results it concludes that the proposed CNN classifier with softmax function shows better results in terms of accuracy. Here CNN is used for both feature extraction and tumor detection from brain MRI images. The results are measured with respect to sensitivity, specificity, precision and accuracy.

Patil and Pradeepini [23] focused on real time detection of tumor levels in three dimensional MRI images. 3D MRI image in Digital Imaging and Communications in Medicine (DICOM) format is converted into Joint Photographic Experts Group (JPEG) format then denoising, grey scale and blurring techniques were applied on it to remove noise. Feature extraction algorithms were applied on each layer of 3D image to generate feature vectors and save these images on Hadoop Distributed File System (HDFS). Clustering techniques is used to identify region of interest in each and every layer of 3D image. Then classification algorithm is used on learning trained datasets to predict accurate level of brain tumor. To speed up diagnosis time and improve efficiency is achieved by using a parallel processing Hadoop's MapReduce framework.

Amin et al [24] proposed a Deep Neural Network (DNN) for brain tumor detection in a MRI image. Totally it consists of 7 layers those are 3 convolutional layers, 3 Rectified Linear Units (ReLU) layers and a softmax layer. Firstly input image samples are categorized into multiple patches and subsequently the center pixel value of every patch is complete towards the DNN. DNN assign labels related to center pixels and segmentation is carryout. Various research are carried out with the help of 8 huge sized bench mark datasets that contain BRATS 2012&2013 with benchmark and synthetic images, 2014, 2015 and Ischemic Stroke Lesion Segmentation (ISLES) 2015 & 2017. Proposed model works better with respect to accuracy and average processing time. It is used for brain lesion discovery at an early stage.

Zhang et al [25] proposed a 6-layer Dense Convolutional Network (DCN) for segmentation of brain tumor images. It connects all layers to each consequent layer in a feed-forward way. This architecture of exact connectivity confirms the greater flow of data among the layers in the network and DCN assures the highest data flow among layers in the DCN and preserves the feature propagation between each and every one layers. This arrangement maximizes the efficiency at the time of training and the accuracy obtained from the results. The method is experimented and validated via Multimodal BRATS 2017 dataset.

Imai et al [26] introduced a new method for 3D medical image segmentation which makes use of data-swapping, which extracts intermediate information from Graphics Processing Unit (GPU) memory towards Central Processing Unit (CPU) memory towards increase the efficient GPU memory size. Parameters are tuned in this data-swapping in order to get improved segmentation result for 3D U-Net based Deep Neural Network (DNN). 3D U-Net is trained by means of fullsize images of $192 \times 192 \times 192$ voxels in dataset.

3. PROPOSED METHODOLOGY

In this work, novel Medical Image Cloud Processing (MICP) is proposed for brain tumor detection. The pipeline of the proposed distributed processing framework is illustrated in Figure 1 which is mentioned in detail below. The proposed MICP framework is able to distinguish benign and malignant tumors precisely. It consists of several stages: Database collection, Data Acquisition, Data Preprocessing, Data augmentation, MICP, Model creation, Image retrieval and detection, Evaluation.

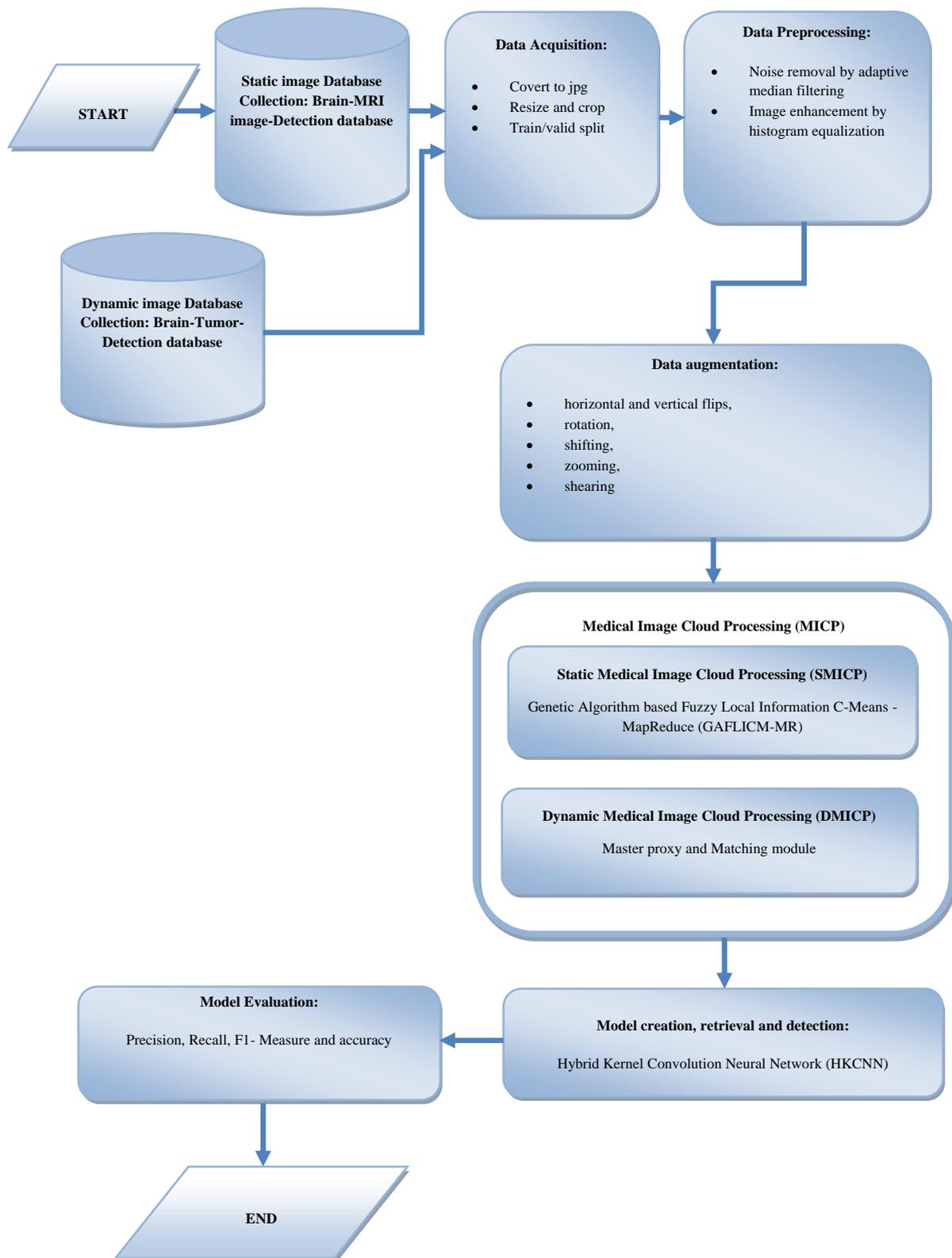


FIGURE 1. PIPELINE OF PROPOSED DISTRIBUTED PROCESSING FRAMEWORK

3.1. Database Collection

The dataset contains 253 Brain Magnetic Resonance Imaging (MRI) Images with 2 folders: Yes and No. Yes folder consists of 155 Brain MRI Images which are yes class and no folder consists of 98 Brain MRI Images which are considered to be no class. The dataset includes of 155 positive and 98 negative examples, before carrying out the process of data augmentation, resulted in 253 example images. Subsequent to the process

of data augmentation, the dataset resulted in 2065 example images with 1085 positive and 980 examples. These 2065 images enclose moreover the 253 original images with the folder name of augmented data. These samples are collected from the <https://github.com/MohamedAliHabib/Brain-Tumor-Detection>.

3.2. Data Acquisition

Data describes the task and plays a significant role in the performance model. The following preprocessing steps were applied to MRI image samples:

- Crop the element of the MRI image by 450×450 with the purpose of enclose simply the brain.
- Resize the MRI image towards contain a shape of (240, 240, 3)=(image_width, image_height, number of channels) in diverse sizes. Consequently, each and every one MRI images must contain the like shape towards feed it as an input to the classifier.
- Normalization is proposed in order to level pixel values to the range 0-1.
- Samples with training (70%), validation (15%) and testing (15%) samples are used for implementation.

3.3. Data Preprocessing

This work, brain MRI images are pre-processed via Adaptive Median Filtering (AMF) and brain image enhancement via Histogram Equalization (HE). Adaptive Median Filter (AMF) is considered as pre-processing stage to achieve the efficiency of image processing for image retrieval on brain MRI image. AMF algorithm is proposed to reduce noises in the brain MRI images [27] which increase the quality of the image. HE is a method in image processing of image enhancement using the image's histogram [28]. Once the images are pre-processed then it is feed into Medical Image Cloud Processing (MICP) framework.

3.3.1. Medical Image Cloud Processing (MICP)

Medical Image Cloud Processing (MICP) framework is proposed with two methods, i.e., Static Medical Image Cloud Processing (SMICP) and Dynamic Medical Image Cloud Processing (DMICP). To accomplish Static Medical Image Cloud Processing (SMICP), two forms of Pure-Image(P-Image) and Big-Image are introduced to Genetic Algorithm based Fuzzy Local Information C-Means -MapReduce (GAFLICM-MR). GAFLICM-MR algorithm is used to attain more optimized formation and higher effectiveness. Here GAFLICM is also used for brain tumor segmentation. Subsequent To the division of Big-Image, each Map Node should handle by means of its equivalent cluster group CGP[k] (the kth clusters of P-Images in parallel toward increase pixel values and achieve feature extraction. Mapping function i.e. $Ma(.)$ is named which takes CGP[k] as input. The ultimately increased features are able to be described in the equation (1)

$$Fe_{Im}(k) = Ma\{CGP[k]\} \tag{1}$$

where $Fe_{Im}(k)$ characterizes the total features of each $CGP[k]$ after the $Ma(.)$ operation. The input to the reducing function is gained FI (k). i.e. $Re(.)$, in which α_k is a new input coefficient. The final output ‘Op’ is expressed as equation (2),

$$Op = Re \left\{ \sum_{k=1}^{NumMapTask} \alpha_k Fe_{Im}(k) \right\} \tag{2}$$

In the reducing operation, the features $Fe_{Im}(k)$ extracted from each $Ma(.)$ would be used for the following steps such as classification, retrieval, and detection, etc.. In equation (2), α_k be set as 1 for image classification and image retrieval or 0 for image detection related toward the accurate response. In the equation (2) clustering has been performed by using GAFLICM. GAFLICM based on MapReduce is introduced in this work for segmentation. The map function performs the procedure of assigning each MRI brain images to the closest center while the reduce function performs the procedure of updating the new centers. Finding the closet center to the map becomes very difficult task, so GAFLICM clustering is introduced in this work. This clustering algorithm is also used as the tumor segmentation process.

Segmenting the tumor area in the brain MRI images are the most important for the identification of brain tumor exactly. Clustering is introduced of all MRI image pixels in a feature space constructed by the brain image gradients in the image. Fuzzy C Means (FCM) clustering with new fuzzy factor known as Fuzzy Local Information C-Means (FLICM) clustering algorithm [29]. Its objective function is described in the following equation (3),

$$J_m = \sum_{i=1}^N \sum_{k=1}^c [u_{ki}^m ||I_i - v_k||^2 + G_{ki}] \tag{3}$$

Novel fuzzy factor G_{ki} defined as in the equation (4)

$$G_{ki} = \sum_{j \in N_i, i \neq j} \frac{1}{ds_{ij} + 1} (1 - u_{kj})^m ||I_j - v_k||^2 + \alpha (1 - u_{kj})^m * I_j \tag{4}$$

where $I = \{I_1, \dots, I_n\} \subseteq \mathbb{R}^m$ is the MRI brain images in the m-dimensional vector space, N is the total number of MRI images, is the amount of clusters with $2 \leq c \leq N$, u_{ji} , is the degree of membership of in the cluster, m is the weighting exponent on each fuzzy membership, v_j is the cluster center prototype, ds_{ij} is a distance measure between MRI image and cluster center. α is a weighting coefficient which controls the influence of neighborhood gray information of MRI brain images. In equation (4), Euclidean distance measure is used in the FLICM objective function, which is computationally expensive [29]. In the equation (4) both fuzzy weighting exponent and α weighting coefficient is controlled via the use of the gray information of MRI brain images. So the computation of these weighting factors becomes very important task.

In this work, Genetic Algorithm (GA) is introduced for tuning of these parameters [30]. Initially, the population is usually included of a randomly created set of weighting factors. Known a fitness function F(P) for objective function of the clustering algorithm, which measures the fitness of each weighting factors in the population, the objective is to make best use of this value. The steps involved in the GA process are briefly described as follows:

1) Build initial population is randomly created, though it might be depending on a recognized solution set.

2) Evaluate fitness - every individual solution in a weighting factors are assessed and consequently assigned a measure of fitness in the clustering algorithm. Fitness function for evaluating the generalized solutions are stated as in the equation (5),

$$F(P) = \frac{1}{J_m} \tag{5}$$

3) New weighting factors are developed and considered this is a population, for the subsequently creation - the type of genetic process is probabilistically selected, choosing a particular weighting factor or two weighting factors for reproduction;

Selection - Weighting of the present population are chosen as appropriate topic for improvement of the next creation depending on their fitness function F(P).

Crossover - Two selected weighting factor of the clustering are combined by using a crossover point to create two new weighting factors.

Mutation - Known a small mutation probability factor, a new weighting factor might be probabilistically modified in the direction of a little degree.

4) Termination criteria - if the maximum clustering results are achieved with fitness improvement then stop, else go to step 2.

The two needed conditions for J_m to be at its local minimal extreme with respect in the direction of u_{ki} and v_k is computed as follows:

$$u_{ki} = \frac{1}{\sum_{j=1}^c \left(\frac{ds_{ik} + G_{ki}}{ds_{ij} + G_{ji}} \right)^{\frac{1}{m-1}}} \tag{6}$$

$$v_k = \frac{\sum_{i=1}^N u_{ki}^{m_{new}} I_i * \alpha_{new}}{\sum_{i=1}^N u_{ki}^{m_{new}} * \alpha_{new}} \tag{7}$$

Once the convergence of the image is carried out, the defuzzification is done in order to convert the fuzzy partition matrix U to a well-organized partition. The large quantity of membership procedure is considered to be the significant technique so as to de-fuzzify the partition matrix. This process allocates the pixel towards the class with the maximum membership as edge else it is not an edge of the brain image.

$$C_i = \arg_k \{ \max \{ u_{ki} \} \}, k = 1, \dots, c \tag{8}$$

ALGORITHM 1: GAFLICM clustering algorithm

INPUT: Brain MRI Images $I = (I_1, \dots, I_M)$, fuzzification parameter m, α is a weighting coefficient, $c, 2 \leq c < n; m, 1 \leq m < \infty$.

OUTPUT: Segmented input image

STEP 1: Let us consider that the total amount of cluster prototypes v , fuzzification parameter m and the ending criteria ϵ .

STEP 2: Set randomly the fuzzy partition matrix $U^{(0)}$.

STEP 3: Let us consider the loop counter $b=0$.

STEP 3.1. Calculate the cluster prototypes v_k using equation (7) for each image I.

STEP 3.2. Compute membership values u_{ki} using equation (6) for each image I.

STEP 3.3. Compute the new weighting factors (m & α) via the genetic algorithm

STEP 3.4. Update the cluster distance and find the objective function J_m for MRI image segmentation

STEP 3.5. Similar pixels are grouped as cluster c and then non-similar pixels are removed from the segmented area

STEP 4: If $\max \{U^{(b)} - U^{(b+1)}\} < \epsilon$ subsequently stop, or else, set $b=b+1$ and stop with step 3.

SMICP with GAFLICM algorithm assures an established processing process even when the MRI brain images scale reaches an enormous amount, which mostly features to the increased scalability with the purpose of the cloud computing technology. After each map task, concern a combiner to join the intermediate data of the similar map task. In reduce function, can sum each and every one the samples and calculate the entire amount of samples assigned to the similar cluster. Dynamic Medical Image Cloud Processing (DMICP) is developed via a parallel processing procedure. In DMICP mechanism, Master Proxy and a Matching Module are designed depending on conventional parallel file system, which assist to give elevated constancy and load resistance.

3.4. Data augmentation

Data augmentation is ways of expanding the dataset. Deep learning models are most suitable methods for performing this task. It is also measured towards be a type of regularization [31] carryout on the dataset stage, as it permit decreasing overfitting and improving classifier results by increasing the training images without changes of classifier model. Diverse augmentations methods [32] are used to exploit training images on runtime, allowing the model towards attaining higher results. Here the training images were augmented via steps like horizontal and vertical flips, rotation, changing, zooming, cut off, and brightness managing.

3.5. Creation Model and retrieval

For retrieval and detection of brain tumors, Hybrid Kernel Convolution Neural Network (HKCNN) has shown an optimistic performance in disease diagnosis system. HKCNN is an algorithm that accepts segmented Brain MRI image X as input and predicts outcomes Y as either true or false. CNN is a prominent architecture of deep learning. As an instance, LeNet-5 consists of three layers such as convolution, pooling, and fully connected layers. The first layer aims to study feature representations of the brain MRI images. As illustrated in figure 2, convolution layer is collected of numerous convolution kernels are used to calculate diverse feature maps. The new feature map be able to be extracted by first transforming the input by means of a learned kernel and subsequently developing on the converted values an element-wise non-linear activation function. Remember that in order to produce each function map, all spatial positions of the input must share the kernel. Through numerous different kernels the complete feature maps are produced.

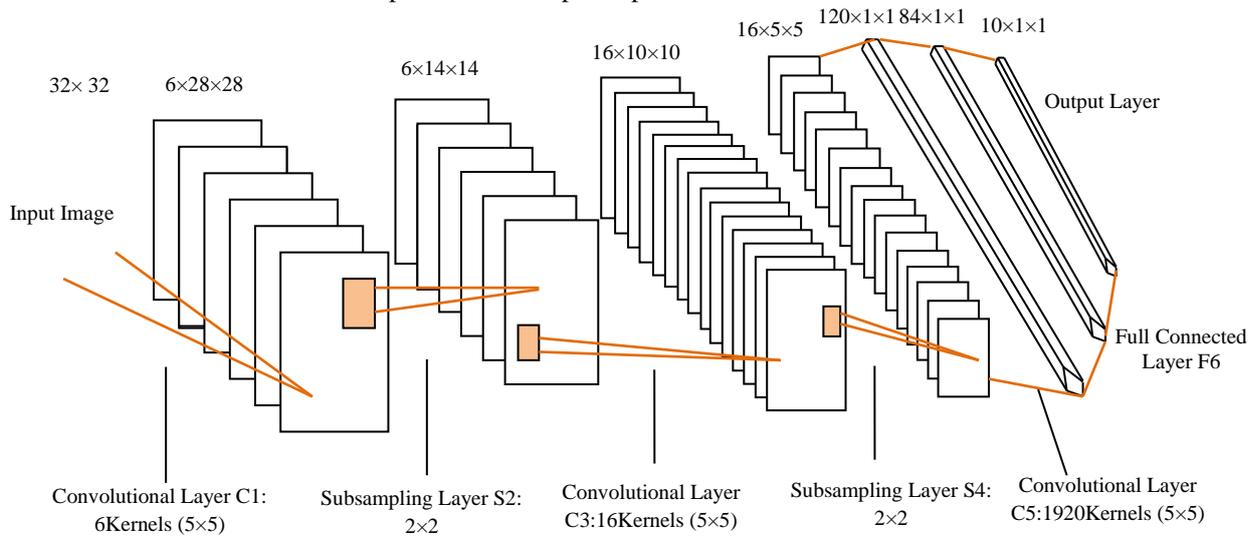


FIGURE 2. LENET-5 NETWORK ARCHITECTURE ON CLASSIFICATION TASK

Mathematically, the feature value of segmented MRI images at position (i,j) in the kth feature map of lth layer, $z_{i,j,k}^l$, is determined by equation (9):

$$z_{i,j,k}^l = w_k^{lT} X_{i,j}^l + b_k^l \tag{9}$$

where w_k^l and b_k^l is denoted as the weight vector and bias vector of the kth filter of the lth layer correspondingly, and $X_{i,j}^l$ is the input patch centered at position (i,j) of the lth layer. Weight kernel w_k^l with the purpose of creates the feature map $z_{i,j,k}^l$ is shared. The weight kernel w_k^l has major significant role in the classifier. Thus, weight kernel value is computed via the use of two different kernels are combined as multiplication operator. The results of two different kernels combined and named as the hybrid kernel. Consider $a(\cdot)$ is described as the nonlinear activation function. The activation value $a_{i,j,k}^l$ of convolutional feature $z_{i,j,k}^l$ is able to be calculated as in the equation (10):

$$a_{i,j,k}^l = a(z_{i,j,k}^l) \tag{10}$$

Common activations functions are sigmoid, tanh [33] and ReLU [34]. The pooling layer is used towards obtain shift-invariance of the feature maps for segmented MRI images. It is generally positioned among two

convolutional layers. Every feature map for segmented MRI images of a pooling layer is related towards its parallel feature map for segmented MRI images of the previous convolutional layer. Let us consider that the $pool(\cdot)$ is pooling function, for every feature map $a_{i,j,k}^l$ have in the equation (11):

$$y_{i,j,k}^l = pool(a_{m,n,k}^l), \forall (m, n) \in R_{i,j} \tag{11}$$

where R_{ij} is location's local neighbourhood (i, j) . Max pooling based function [35-36] is used for classification of image as Tumor and non-tumor. Then fully-connected layers which aims to perform high-level reasoning [37-38]. Final layer of CNNs is an output layer [39]. Let θ denote each and every one the parameters of a CNN. The optimum parameters for a specific task are able to be computed by reducing a suitable loss function described on with the purpose of task. Let us assume that there are N desired input-output pair $\{(x^{(n)}, y^{(n)})\}; n \in [1, \dots, N]$, where $x^{(n)}$ is the n^{th} input data, $y^{(n)}$ is its related target label and $o^{(n)}$ is the output of CNN. The loss of CNN classifier is computed as follows:

$$L = \frac{1}{N} \sum_{n=1}^N \ell(\theta; y^{(n)}, o^{(n)}) \tag{11}$$

Stochastic gradient descent is a traditional method for tuning the parameters of the CNN network [40-41]. In the equation (9), w_k^l is computed via the use of the kernel functions as sigmoid and polynomial.

$$\text{Sigmoid} = K_1(x_n, x_i) = \tanh(\gamma(x_n, x_i) + r) \tag{12}$$

$$\text{Polynomial} = K_2(x_n, x_i) = (\gamma(x_n, x_i) + r)^d \tag{13}$$

The results of these two kernels are computed via the function $w_k = K_1 * K_2$. Explanation, C: cost; γ : gamma; r: coefficient; d: degree is created via the search method. Let us consider the upper and lower bounds of these parameters are follows $C (2^{-15}, \dots, 2^2)$, $\gamma (2^{-10}, \dots, 2^2)$, $r (2^{-10}, \dots, 2^2)$, and $d (0, 1, 2)$ via grid-search method. This result is updated to the equation (9).

3.6. Training model

Training model is performed to tune its parameters and update weights values of the classifier.

3.7. Evaluation model

Proposed model is evaluated via the metrics like precision, recall, F-Measure and Accuracy.

4. RESULTS AND DISCUSSION

This section shows the experimentation results of proposed SSFCM-MR framework: 1) to confirm the results of Corel-10k dataset over the usual small image files; 2) to validate the time value of SICP when processing huge-size static image samples; 3) to show the constancy and complexity of DICP when processing the dynamic images as input. Corel-10k dataset has been taken from <http://www.ci.gxnu.edu.cn/cbir/Dataset.aspx> which includes of 100 categories with 10,000 images from various elements like sunset, beach, flower, building, car, horses, mountains, fish, food, door, etc. All categories consist of 100 images with JPEG format with size of 192x128 or 128x192. From these five categories such as buildings, rose, car, mountains, and dragon are used in this implementation. For these five categories 50 images are considered for each class. Totally 250 images are used in experimental analysis with MATLAB environment. In addition the proposed work is also implemented to another dataset. Brain MRI images dataset includes of two folders: yes and no with total of 253 Brain MRI Images. The first folder consists of 155 Brain MRI Images which belongs to positive class and the second folder consists of 98 Brain MRI Images which belongs to negative class. These collection of samples are done from <https://github.com/MohamedAliHabib/Brain-Tumor-Detection>

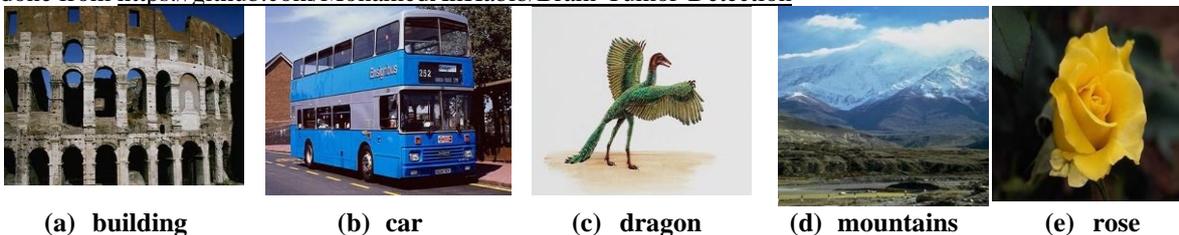


FIGURE 3. IMAGE SAMPLES FROM COREL-10K DATASET FOR FIVE CLASSES

In the figure 3 shows the images of Corel-10k dataset for five classes such as (a) Buildings, (b) Car, (c) Dragon, (d) Mountains, and (e) Rose.

Results are used to estimate retrieval methods are precision, recall, F-measure, accuracy, time and memory. Precision and recall metrics are described as follows,

$$\text{Precision} = (\text{Number of relevant images retrieved} / \text{Total number of images retrieved}) * 100 \tag{14}$$

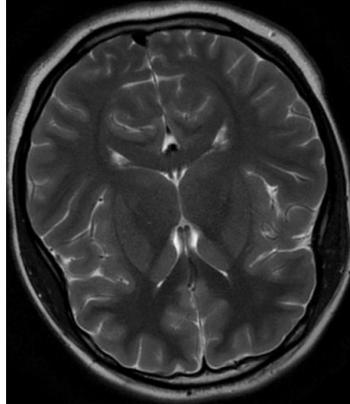
$$\text{Recall} = (\text{Number of relevant images retrieved} / \text{Number of relevant images in database}) * 100 \tag{15}$$

F-measure is defined as the weighted harmonic mean of the precision and recall of the test.

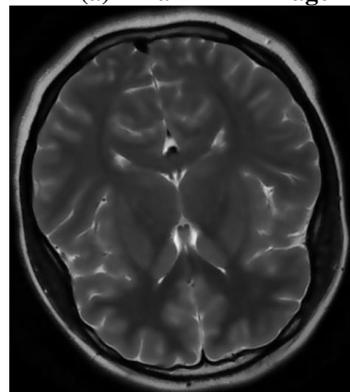
$F\text{-measure} = 2 * ((\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})) * 100$ (16)

Accuracy is defined as the number of relevant images divided by total number of retrieval images.

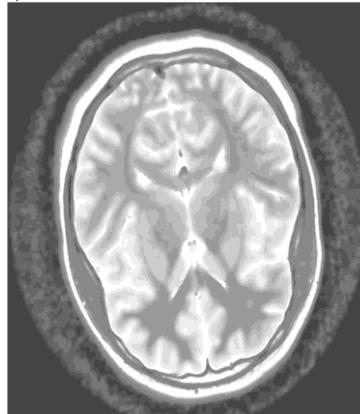
$\text{Accuracy} = (\text{No. of relevant images} / \text{total of retrieval images}) * 100$ (17)



(a) Brain MRI image



(b) AMF results for brain MRI image



(c) Histogram Equalization (HE) results for brain MRI image



(d) Segmented image results for brain MRI image
FIGURE 4. BRAIN MRI IMAGE RESULTS

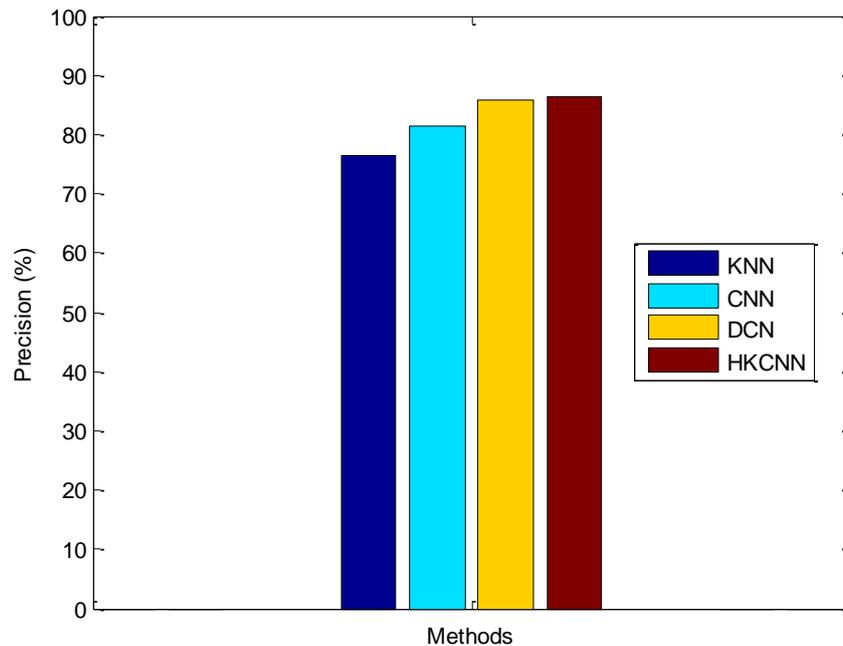
The steps included in the proposed work with brain MRI image sample are shown in the figure 4. In figure 4(a) shows the input image sample of brain MRI image, figure 4(b) shows the AMF results for brain MRI image and figure 4(c) shows the HE results for brain MRI image. Figure 4(d) shows the segmented tumor area results for brain MRI image.

TABLE 1. PERFORMANCE COMPARISON RESULTS OF IMAGE PROCESSING METHODS WITH COREL 10 K DATASET (DRAGON CLASS)

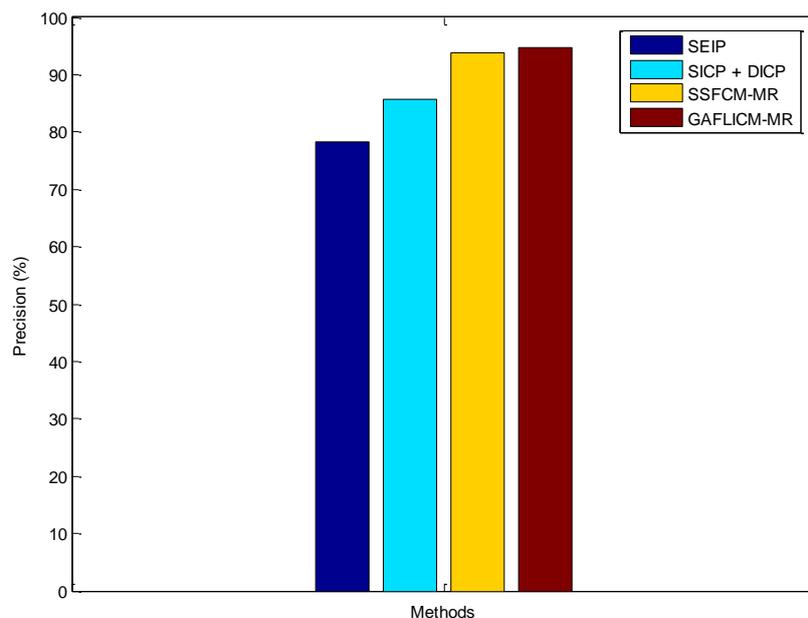
Metrics	SEIP	SICP+DICP	SSFCM-MR	GAFLICM-MR
Precision(%)	78.31	85.75	93.91	94.6800
Recall(%)	76.71	82.03	90.32	92.3500
F-measure(%)	77.50	83.85	92.08	94.8500
Accuracy(%)	80.00	84.00	92.00	94.00
Time(Seconds)	288.0046	112.8114	84.0681	62.5846
Memory(bytes)	7204	4908	3084	2642

TABLE 2. RESULTS COMPARISON OF BRAIN TUMOR RETREIVAL METHODS

Metrics	KNN	Convolutional Neural Network (CNN)	Dense Convolutional Network (DCN)	HKCNN
Precision(%)	76.3658	81.3654	85.6947	86.5487
Recall(%)	79.2500	84.2546	86.3254	89.1232
F-measure(%)	77.7812	82.7848	86.0089	87.8171
Accuracy(%)	80.7692	83.4615	88.4615	92.3077
Time(Seconds)	208.3521	172.4558	142.1039	113.6036
Memory(bytes)	40004	31564	30308	20256



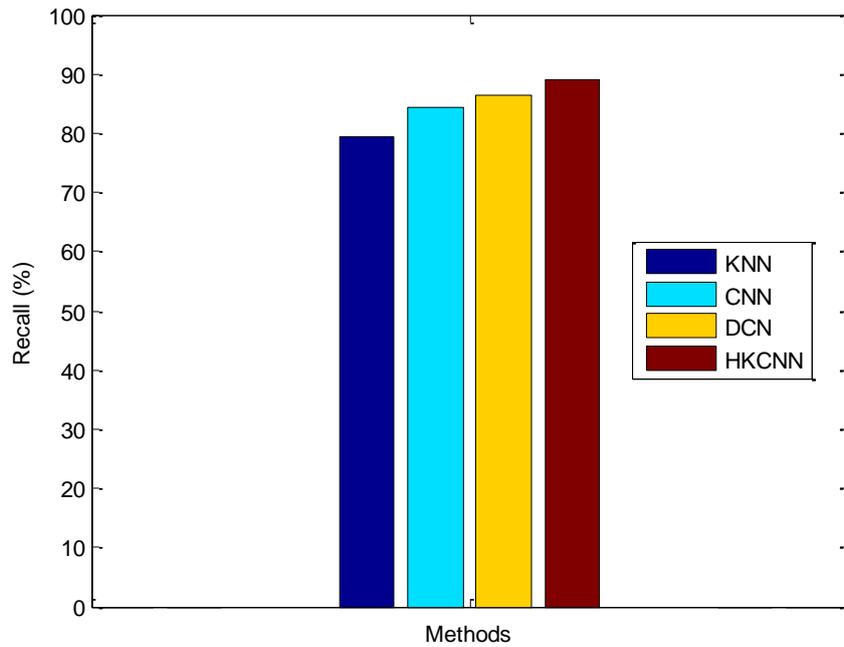
(A) PRECISION EVALUATION VS. CLASSIFICATION METHODS (BRAIN MRI IMAGES DATASET)



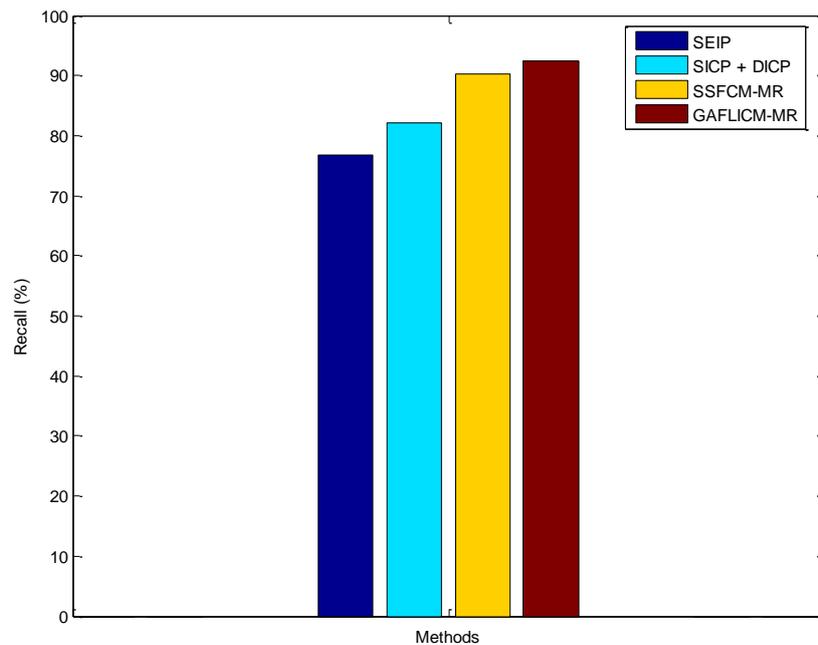
(B) PRECISION EVALUATION VS. RETRIEVAL METHODS(COREL 10 K DATASET (DRAGON CLASS))

FIGURE 5. PRECISION RESULTS ANALYSIS

The precision results comparison of the classification methods with respect to the brain MRI images are discussed in the figure 5(a).The precision results comparison of the retrieval methods for dragon class is illustrated in the figure 5(b). The performance measures in terms of precision via four classifiers such as KNN, CNN,DCN and HKCNN. The proposed HKCNN classifier gives higher precision results of 86.5487%, whereas other methods such as KNN, CNN, and DCN give only 76.3658%, 81.3654%, and 85.6947%for brain MRI images is illustrated in the figure 5(a) (Table 2). Similarly the proposed HKCNN classifier gives improved precision results of 94.6800%for dragon class which is shown in the figure 5(b)(Table 1).



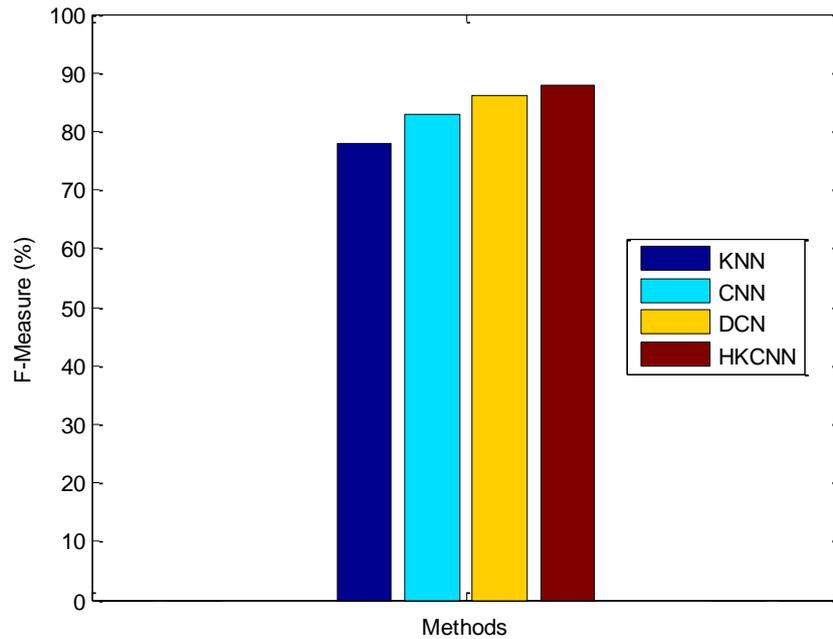
(A) RECALL EVALUATION VS. CLASSIFICATION METHODS (BRAIN MRI IMAGES DATASET)



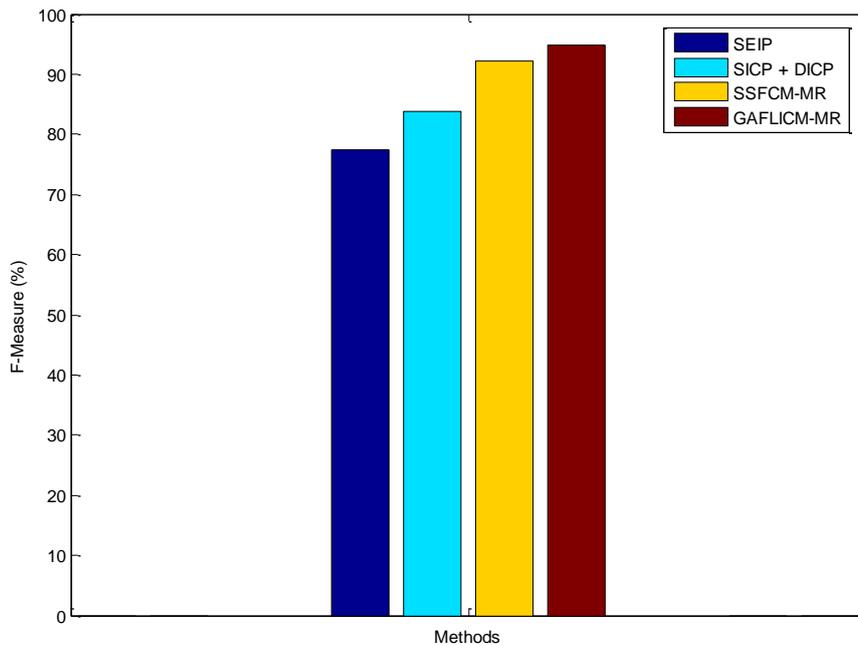
(B) RECALL EVALUATION VS. RETRIEVAL METHODS(COREL 10 K DATASET (DRAGON CLASS))

FIGURE 6. RECALL RESULTS ANALYSIS

Retrieval and classification methods with respect to the dragon class and the MRI images are discussed in the figure (6)(b) & 6(a) of recall results . Recall results of four classification methods such as KNN, CNN,DCN and HKCNN. The proposed HKCNN classifier gives higher recall results of 89.1232%, whereas other methods such as KNN, CNN, and DCN give only 79.2500%, 84.2546%, and 86.3254%for brain MRI images are shown in the figure 6(a)(Table 2). Similarly the proposed work gives higher recall results of 92.3500% in Corel 10K dataset which is shown in the figure 6(b)(Table 1).



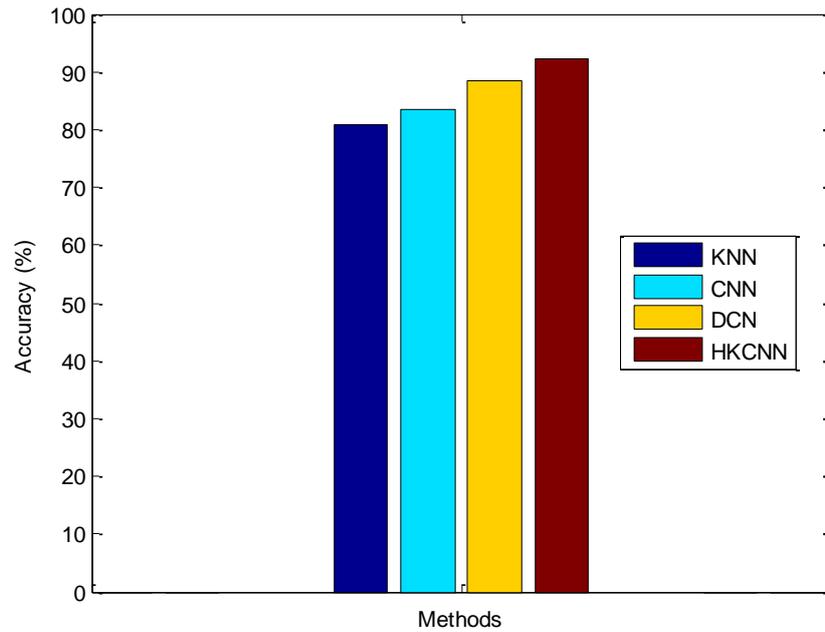
(A) F-MEASURE EVALUATION VS. CLASSIFICATION METHODS (BRAIN MRI IMAGES DATASET)



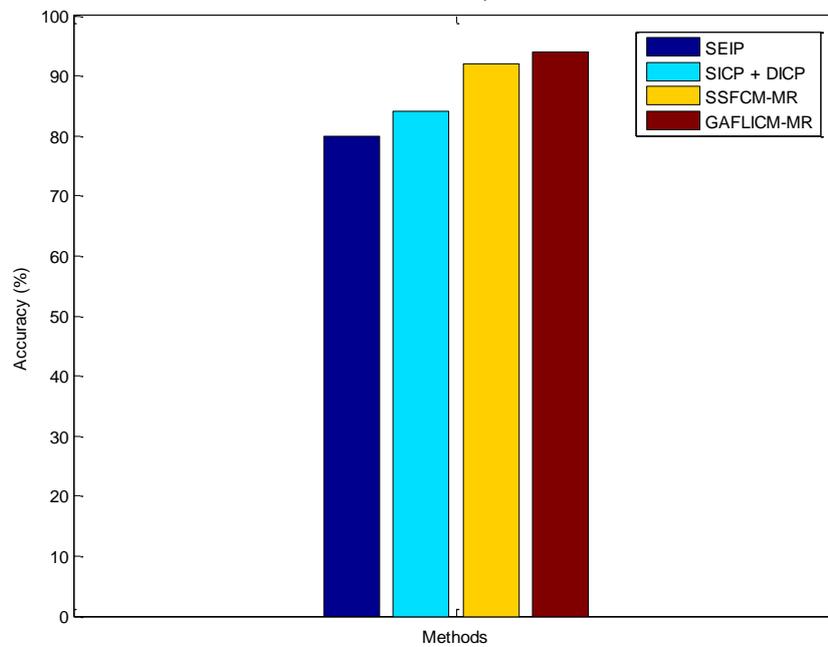
(B) F-MEASURE EVALUATION VS. RETRIEVAL METHODS(COREL 10 K DATASET (DRAGON CLASS))

FIGURE 7. F-MEASURE RESULTS ANALYSIS

Results evaluation of f-measure under retrieval and classification methods of dragon class and the brain MRI images are illustrated in the figure (7)(b) & 7(a). The methods like KNN, CNN, DCN and HKCNN are shown in the figure 7(a). The proposed HKCNN classifier gives higher f-measure results of 87.8171%, whereas other methods such as KNN, CNN, and DCN give only 77.7812%, 82.7848%, and 86.0089% for brain MRI images are discussed in the figure 7(a) (Table 2). Similarly the proposed work gives higher f-measure results of 94.8500% with dragon class which is shown in the figure 7(b) (Table 1).



(A) ACCURACY EVALUATION VS. CLASSIFICATION METHODS (BRAIN MRI IMAGES DATASET)

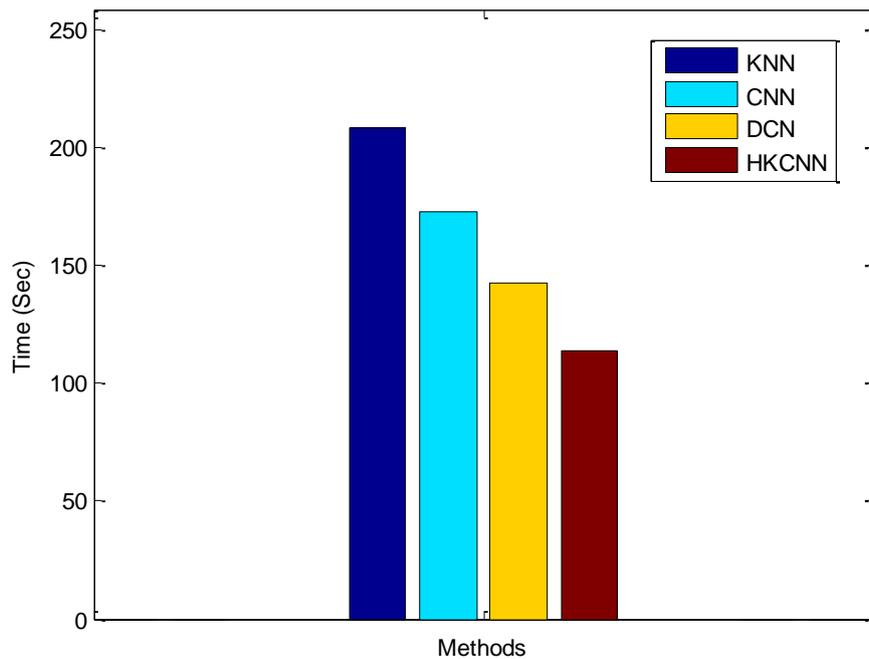


(B) ACCURACY EVALUATION VS. RETRIEVAL METHODS(COREL 10 K DATASET (DRAGON CLASS))

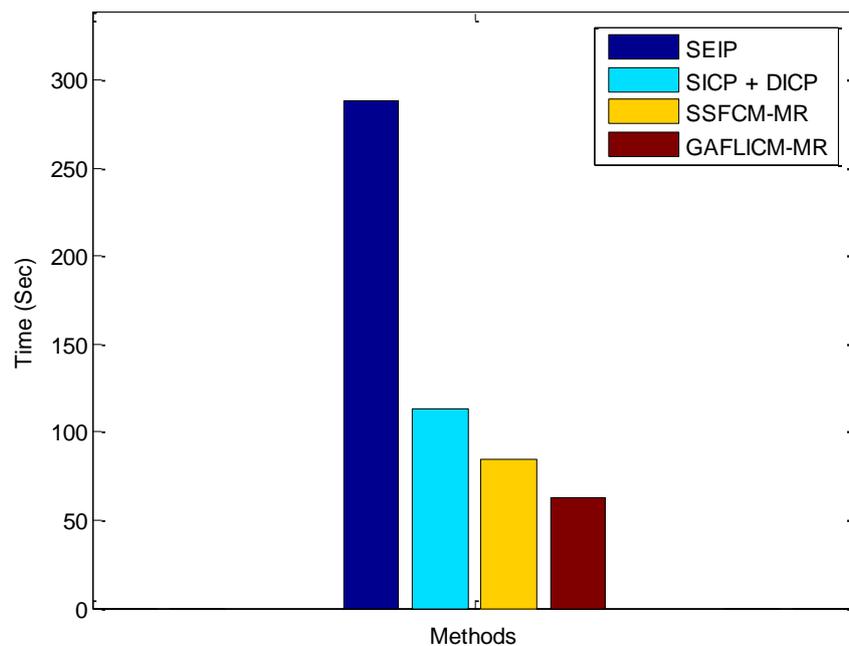
(C)

FIGURE 8. ACCURACY RESULTS ANALYSIS

Overall accuracy results with respect to classification and retrieval methods are shown in the figure 8(a) and figure 8(b). The proposed HKCNN classifier gives higher accuracy results of 92.3077%, whereas other methods such as KNN, CNN, and DCN give only 80.7692%, 83.4615%, and 88.4615% for brain MRI images in the figure 8(a) (Table 2). Similarly the proposed HKCNN classifier gives improved accuracy results of 94.00% with dragon class which is shown in the figure 8(b) (Table 1).



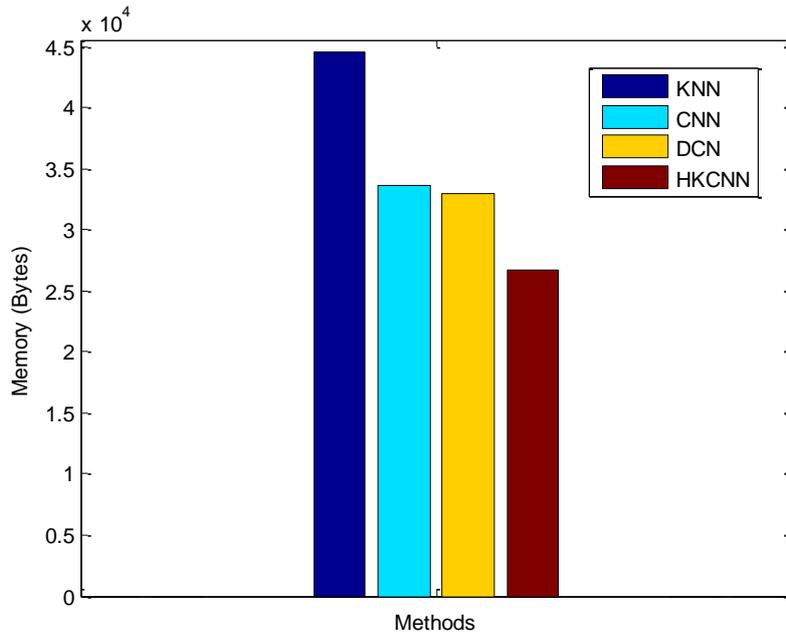
(A) TIME COMPARISON VS. CLASSIFICATION METHODS (BRAIN MRI IMAGES DATASET)



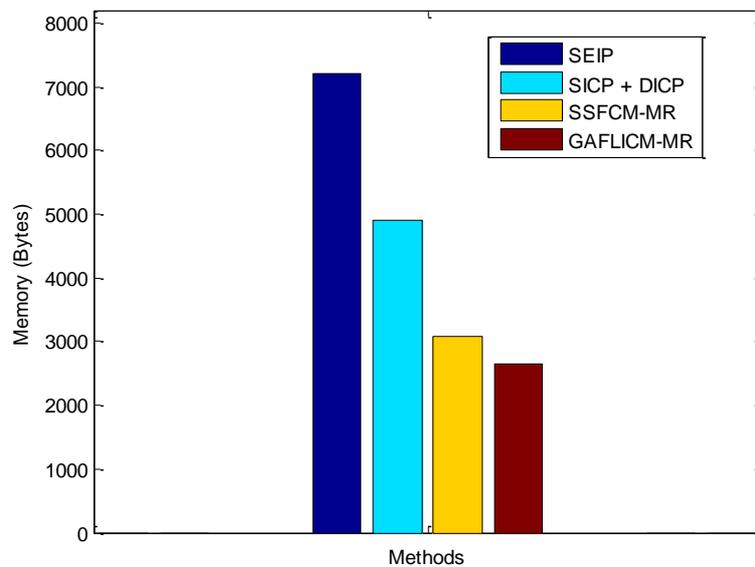
(B) TIME COMPARISON VS. RETRIEVAL METHODS(COREL 10 K DATASET (DRAGON CLASS))

FIGURE 9. TIME COMPARISON

Figure 9 shows the time comparison results of brain MRI images and Corel 10 K dataset. Figure 9(a) shows the results of classification methods like KNN, CNN, DCN and proposed HKCNN classifier with respect to time in seconds. From the figure 9(a) it shows with the purpose of the proposed HKCNN classifier has takes lesser time of 113.6036 seconds, other classifiers such as KNN, CNN and DCN takes more time of 208.3521 seconds, 172.4558seconds and 142.0000seconds for brain MRI image. From the figure 9(b) it shows with the purpose of the proposed GAFLICM-MR classifier has takes lesser time of 62.5846 seconds, other classifiers such as SEIP, SICP+DICP and SSFCM-MR takes more time of 288.0046 seconds, 112.8114 seconds and 84.0681 seconds for dragon class.



(A) MEMORY USAGE COMPARISON VS. CLASSIFICATION METHODS (BRAIN MRI IMAGES DATASET)



(B) MEMORY USAGE COMPARISON VS. RETRIEVAL METHODS (COREL 10 K DATASET (DRAGON CLASS))

FIGURE 10. MEMORY USAGE COMPARISON

The proposed HKCNN classifier has lesser memory usage of 26688bytes, whereas other methods such as KNN, CNN, and DCN has consumes higher memory usage of 44552bytes, 33600bytes, and 33004 bytes for brain MRI images in the figure 10(a) (Table 2). Similarly the proposed work has lesser memory usage of 2642 bytes for dragon class which is shown in the figure 10(b) (Discussed in table 1).

5. CONCLUSION AND FUTURE WORK

In this work, novel Medical Image Cloud Processing (MICP) based effective distributed framework is proposed for brain tumor detection in MRI images. This work, Adaptive Median Filtering (AMF) is applied for image pre-processing and Histogram Equalization (HE) for image enhancement. In MICP, Static Medical Image Cloud Processing (SMICP) is accomplished with two image representations such as Pure-Image and Big-Image. These representations are preformed based on the Genetic Algorithm based Fuzzy Local Information C-Means - MapReduce (GAFLICM-MR) algorithm and it is also used for image segmentation which achieves higher efficiency. In MICP, Dynamic Medical Image Cloud Processing (DMICP) is developed via a parallel processing procedure of the distributed system. For retrieval and detection of brain MRI images, Hybrid Kernel Convolution Neural Network (HKCNN) is proposed in this work. HKCNN classifier, weight kernel is computed via the use of the kernel functions as sigmoid and polynomial. From the results it shows with the purpose of the

proposed HKCNN classifier gives improved detection results when compared to other classifiers with respect to precision, recall, F-measure, accuracy, time and memory usage. The proposed HKCNN classifier gives higher accuracy results of 92.3077%, whereas other methods such as KNN, CNN, and DCN give only 80.7692%, 83.4615%, and 88.4615%. This work will be extended to experimentation through larger datasets and more tumor types. Future work, morphological operations and denoising methods are applied as post processing, which further improves the accuracy of segmentation.

REFERENCES

1. Selkar, R.G.; Thakare, M. Brain tumor detection and segmentation by using thresholding and watershed algorithm. *Int. J. Adv. Inf. Commun. Technol.* 2014, 1, 321–324.
2. Borole, V.Y.; Nimbhore, S.S.; Kawthekar, D.S.S. Image Processing Techniques for Brain Tumor Detection: A Review. *Int. J. Emerg. Trends Technol. Comput. Sci.* 2015, 4, 28–32.
3. Mustaqeem, A.; Javed, A.; Fatima, T. An efficient brain tumor detection algorithm using watershed & thresholding based segmentation. *Int. J. Image Graph. Signal Process.* 2012, 4, 34–39.
4. Kaur, H.; Mittal, M. Region Based Image Segmentation for Brain Tumor Detection. *Int. J. Eng. Manag. Res.* 2016, 6, 31–34.
5. Bahadure, N.B.; Ray, A.K.; Thethi, H.P. Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. *Int. J. Biomed. Imaging* 2017, 2017, 9749108.
6. Yazdani, S.; Yusof, R.; Karimian, A.; Pashna, M.; Hematian, A. Image segmentation methods and applications in MRI brain images. *IETE Tech. Rev.* 2015, 32, 413–427
7. Rana, R.; Singh, P. Brain Tumor Detection through MR Images: A Review of Literature. *IOSR J. Comput. Eng.* 2015, 17, 7–18
8. WuF., Z. Wang, Z. Zhang, Y. Yang, J. Luo, "Weakly Semi-Supervised Deep Learning for Multi-Label Image Annotation," *IEEE Transactions on Big Data*, vol.1, no.3, pp.109-122, 2015.
9. YangY., F. Shen, H. T. Shen, H. Li, X. Li, "Robust Discrete Spectral Hashing for Large-Scale Image Semantic Indexing," *IEEE Transactions on Big Data*, vol.1, no.4, pp.162-171, 2015.
10. WangY. C., C. C. Han, C. T. Hsieh, et al, "Biased Discriminant Analysis With Feature Line Embedding for Relevance Feedback-Based Image Retrieval," *IEEE Transactions on Multimedia*, vol.17, no.12, pp.2245- 2258, 2015.
11. Yuzhong Yan, and Lei Huang, "Largescale image processing research cloud" IARIA The Fifth International Conference on Cloud Computing, GRIDs, and Virtualization. May 25 - 29, 2014 - Venice, Italy.
12. Sweeney, C., Liu, L., Arietta, S. and Lawrence, J., 2011. HIPI: a Hadoop image processing interface for image-based mapreduce tasks. *Chris. university of Virginia*, 2(1), pp.1-5..
13. Sozykin, A. and Epanchintsev, T., 2015, MIPr-a framework for distributed image processing using Hadoop. In 2015 9th International Conference on Application of Information and Communication Technologies (AICT) ,pp. 35-39.
14. Jia Li, Kunhui Lin, and Jingjin Wang V. Vazhayil, "Map-Reduce based framework for instrument detection in large-scale surgical videos" *IEEE International Conference on Control Communication & Computing India (ICCC)*. 2015, Trivandrum, 2015, pp. 606-611.
15. LinY., F. Lv, S. H. Zhu, and M. Yanget al, "Large-scale image classification: fast feature extraction and SVM training," *CVPR*, pp.1689-1696, 2011.
16. SuJ., L. Dong, P. Ren, and E. R. Hancock, "Hypergraph matching based on marginalized constrained compatibility," *ICPR*, pp.2922-2925, 2012.
17. Alam M.S., M.M. Rahman, M.A. Hossain, M.K. Islam, K.M. Ahmed, K.T. Ahmed, B.C. Singh and M.S. Miah, Automatic Human Brain Tumor Detection in MRI Image Using Template-Based K Means and Improved Fuzzy C Means Clustering Algorithm, *Big Data and Cognitive Computing*, Vol.3, No.2, pp.1-18, 2019.
18. Xie X., A K-Nearest Neighbor Technique for Brain Tumor Segmentation Using Minkowski Distance, *Journal of Medical Imaging and Health Informatics*, Vol.8, No.2, pp.180-185, 2018.
19. Mathur N., Y.K. Meena, S. Mathur and D. Mathur, Detection of brain tumor in MRI image through fuzzy-based approach, *High-Resolution Neuroimaging-Basic Physical Principles and Clinical Applications*, pp.1-174, 2018.
20. Rehman, A., Naz, S., Razzak, M.I., Akram, F. and Imran, M., 2020. A deep learning-based framework for automatic brain tumors classification using transfer learning. *Circuits, Systems, and Signal Processing*, 39(2), pp.757-775.
21. Wang, G., Li, W., Vercauteren, T. and Ourselin, S., 2019. Automatic brain tumor segmentation based on cascaded convolutional neural networks with uncertainty estimation. *Frontiers in computational neuroscience*, 13, pp.1-13.

22. Siar, M. and Teshnehlab, M., 2019, Brain Tumor Detection Using Deep Neural Network and Machine Learning Algorithm. In 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE), pp. 363-368.
23. Patil, J.S. and Pradeepini, G., 2019. Brain Tumor Levels Detection in Three Dimensional MRI using Machine Learning and MapReduce. *Indian Journal of Public Health Research & Development*, 10(6), pp.1465-1471.
24. Amin, J., Sharif, M., Yasmin, M. and Fernandes, S.L., 2018. Big data analysis for brain tumor detection: Deep convolutional neural networks. *Future Generation Computer Systems*, 87, pp.290-297.
25. Zhang, Z., Odaibo, D., Skidmore, F.M. and Tanik, M.M., 2017. A big data analytics approach in medical imaging segmentation using deep convolutional neural networks. In *Big Data and Visual Analytics* (pp. 181-189). Springer, Cham.
26. Imai, H., Matzek, S., Le, T.D., Negishi, Y. and Kawachiya, K., 2018. Fast and accurate 3d medical image segmentation with data-swapping method. *arXiv preprint arXiv:1812.07816*, pp.1-13.
27. Kalaiselvi, K., 2019. Adaptive Median Filter Based Noise Removal Algorithm for Big Image Data. *International Journal of Advanced Studies of Scientific Research*, Vol. 3, No. 10, 2018, pp.154-159.
28. Abdullah-Al-Wadud, M., Kabir, M.H., Dewan, M.A.A. and Chae, O., 2007. A dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics*, 53(2), pp.593-600.
29. Krinidis, S. and Chatzis, V., 2010. A robust fuzzy local information C-means clustering algorithm. *IEEE transactions on image processing*, 19(5), pp.1328-1337.
30. Rashid, T. and Abdullah, S., 2018. A Hybrid of Artificial Bee Colony, Genetic Algorithm, and Neural Network for Diabetic Mellitus Diagnosing. *ARO-The Scientific Journal of Koya University*, 6(1), pp.55-64.
31. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. MIT Press; 2016. <http://www.deeplearningbook.org>.
32. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J Big Data* 2019.
33. LeCun Y. A., L. Bottou, G. B. Orr, K.-R. Müller, Efficient backprop, in: *Neural Networks: Tricks of the Trade - Second Edition*, 2012, pp. 9–48.
34. Nair V., G. E. Hinton, Rectified linear units improve restricted boltzmann machines, in: *Proceedings of the International Conference on Machine Learning (ICML)*, 2010, pp. 807–814.
35. Wang T., D. J. Wu, A. Coates, A. Y. Ng, End-to-end text recognition with convolutional neural networks, in: *Proceedings of the International Conference on Pattern Recognition (ICPR)*, 2012, pp. 3304–3308.
36. Boureau Y., J. Ponce, Y. LeCun, A theoretical analysis of feature pooling in visual recognition, in: *Proceedings of the International Conference on Machine Learning (ICML)*, 2010, pp. 111–118.
37. Hinton G. E., N. Srivastava, A. Krizhevsky, I. Sutskever, R. R. Salakhutdinov, Improving neural networks by preventing co-adaptation of feature detectors, *CoRR abs/1207.0580*.
38. Lin M., Q. Chen, S. Yan, Network in network, in: *Proceedings of the International Conference on Learning Representations (ICLR)*, 2014.
39. Tang Y., Deep learning using linear support vector machines, in: *Proceedings of the International Conference on Machine Learning (ICML) Workshops*, 2013.
40. Wijnhoven R. G. J., P. H. N. de With, Fast training of object detection using stochastic gradient descent, in: *International Conference on Pattern Recognition (ICPR)*, 2010, pp. 424–427.
41. Zinkevich M., M. Weimer, L. Li, A. J. Smola, Parallelized stochastic gradient descent, in: *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*, 2010, pp. 2595–2603.