Optimize KNN Algorithm for Cerebrospinal Fluid Cell Diseases

Soobia Saeed[1†], Afnizanfaizal Abdullah[2††], NZ Jhanjhi[3†††]

Faculty of Engineering, Department of Software Engineering, Universiti Teknologi Malaysia, Johor Bhrau Malaysia[1, 2]
School of Computer Science and Engineering, SCE, Taylor's University, Malaysia[3]

Abstract
Medical imaginings assume a important part in the analysis of tumors and cerebrospinal fluid (CSF) leak. Magnetic resonance imaging (MRI) is an image segmentation technology, which shows an angular sectional perspective of the body which provides convenience to medical specialists to examine the patients. The images generated by MRI are detailed, which enable medical specialists to identify affected areas to help them diagnose disease. MRI imaging is usually a basic part of diagnostic and treatment. In this research, we propose new techniques using the 4D-MRI image segmentation process to detect the brain tumor in the skull. We identify the issues related to the quality of cerebrum disease images or CSF leakage (discover fluid inside the brain). The aim of this research is to construct a framework that can identify cancer-damaged areas to be isolated from non-tumor. We use 4D image light field segmentation, which is followed by MATLAB modeling techniques, and measure the size of brain-damaged cells deep inside CSF. Data is usually collected from the support vector machine (SVM) tool using MATLAB's included K-Nearest Neighbor (KNN) algorithm. We propose a 4D light field tool (LFT) modulation method that can be used for the light editing field application. Depending on the input of the user, an objective evaluation of each ray is evaluated using the KNN to maintain the 4D frequency (redundancy). These light fields' approaches can help increase the efficiency of device segmentation and light field composite pipeline editing, as they minimize boundary artefacts.

Keywords:
Brain Tumor, MRI, Image Segmentation, CSF, KNN

1. Introduction

The cerebrospinal fluid (CSF) is the fluid that travels through the brain's ventricles and around the surface of the brain and spine. CSF is one of the most challenging pre-neurosurgical complications. CSF leakage is a condition that happens when the CSF leaks through deformity in the dura or head and exits through the nose or ear. CSF leakage is the aftereffect of a gap or tear of the dura that is the outermost layer of meningitis. The purpose behind this research is to identify the hole or tear due to brain injury which can damage the head (brain) or spinal cord. CSF leakage can be detected in the lower backside, also called spinal cord or spinal anesthesia. The brain cells that make up these interfaces are also sites of mechanisms of exchange (transporter) that regulate the brain's cell entry and exit to a broad spectrum of molecules.
An important mechanism for organizing what is unique

Manuscript received February 5, 2024
Manuscript revised February 20, 2024
https://doi.org/10.22937/IJCSNS.2024.24.2.5

The formation of interstitial fluid in the brain is the secretion of cerebrospinal fluid by the choroid plexuses that flow through the ventricular system and are exchanged between the cerebrospinal fluid and the brain. Understanding the complexity of blood-brain-barrier (BBB) mechanisms is necessary to assess the effects of inflammatory conditions on the brain, both in children and adults during growth. The exhaustion of the cerebrospinal liquid may happen by leakage, a shunt, insufficient generation, or exceptionally fast retention of leakage of fluid. There are additionally some comparable disorders where there is high intracranial consistency, causing comparative side effects when the cerebrum contracts on standing in the upper backside. CSF has been widely focused for the recognition of neuron cells for malignant growth identification. This investigation analyses the use of machine learning (ML) for the biochemical components that have been accounted for brain cancer and CSF leakage.

K- Nearest Neighbors, also known as K-NN, is part of the supervised learning algorithms family, which uses a set of tagged data to predict the new category of data points. The K-NN algorithm is a powerful technique that is often used as a point of reference for more complex techniques such as artificial neural networks (ANN) or the SVM. The K-NN algorithm can be easily understood and implemented in machine learning algorithms. The K-NN algorithm is read through a complete data set to find the closest neighbors to classify the new data point. Clustering is used as a research model and is a very effective technique used to identify similarity among different clusters or groups. There are many clustering algorithms in which the k-means algorithm is used to identify hidden patterns from the data. In this research, the k-means algorithm explores the unseen information by taking the attribute of the missing values of the KNN algorithm in brain cancer MRI images. It is an unsupervised clustering count that makes a specific number of disjoint level (non-dynamic) assemblies. The strategy takes after a clear and basic way to deal with the request for a given data set through a particular number of groups. The number of groups (expect k- mean) is developed to find the distance in MRI images.

K-means estimations randomly pick k objects and address the k beginning gathering the data in the images. Further, in this research we find the distance for each step of applying the K-NN algorithm in the given data sets to find the nearest center of object (distance point) as shown in Figure 1.

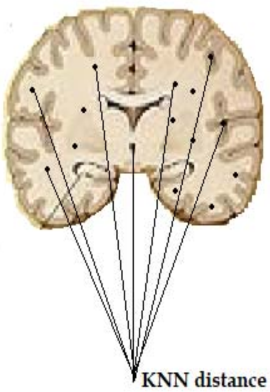


Fig.1 KNN long distance values

We use the Euclidean division and recalculate new *k*-group centers. The strategy is repeated until there is no conformity in *k* collection centers. This estimation goes for minimizing an objective limit known as squared values of missing information of KNN. We also use 4D images of MRI to investigate the stage of cancer and how to recover the hidden information of damaged-cells of brain cancer and CSF with K-NN algorithm. The selection of 4D image segmentation is to identify cancer with efficiency and accuracy. Cancer exists in the brain with different locations in the skull. We use the latest technology of 4D imaging using LFT tools for finding the high accuracy of the images. The algorithm for the 4D light field decides the significant areas in the 4D light field. In the 4D light fields, by means of a 2D input device, the user selects part of the areas in the 4D light field data. However, the auto- selection of a complete region in 4D space remains difficult for users because the functional user interfaces (UI) are different for each region. In the 4D light field, the UI must provide pointers to determining a locale of interest by entering a name in a segment of the area. One can see that the segmentation results of the 4D light fields can be obtained by applying to each point of the two-dimensional image segmentation method; however, there is no guarantee that the 4D light field frequency (redundancy) will be preserved.

Section 1 provides an overview of brain cancer and details of the CSF. Section 2 presents the "problem statement" of the hybrid KNN algorithm for brain cancer and CSF issues. Section 3 explains the previous related work done so far in this domain. Section 4 presents the methodology based on the detection of brain cancer through the MRI-4D interface of image segmentation to identify the missing values of KNN algorithms with statistical results and pseudo codes of algorithm. Section 5 provides details of the results and discussions while mentioning the findings with the new achievements. Finally, Section 6 presents the conclusion, describes the contributions made in this study and outlines the directions for the future.

2. Problem Statement

The problem associated with the brain cell having an abnormality is due to an increased number of cancer cells with low levels of glucose and high levels of protein [1]. This is a potentially serious condition that can cause infection in the CSF (meningitis) or the brain itself (brain abscess) [2]. In this research, we address the above-mentioned issues associated with the damaged cancer cell due to the leakage of CSF. We also focus on the *k*-NN algorithm which is highly sensitive to outliers because it simply selects neighbors based on distance criteria. We use 4D MRI images for finding the distance of cancer cells in the brain to overcome the missing values of the *k*-NN datasets.

3. Related Work

S. Saeed and A. Abdullah [3] proposed a method of image segmentation for classifying MRI images of brain tumor applying a statistical model for tumor detection and classification. They performed comparative analysis of images and applied histogram representation to generate images using binary transformation to predict tumor and non-neoplastic MRI scans based on a support vector machine (SVM).Their approach detected the right tumor region that a radiologist can easily identify and highlight.
S. Saeed and A .Abdullah [4] suggested a method to recognize the causes of brain tumor which is created by abnormalities of brain cells. This research focuses on tumor detection methods using MRI image segmentation to identify the tumor and non-tumor area by image classification of predictive methods of binary transformation and image pre-processing techniques.

A. H. Sin et al.[5] suggested a tool for the segmentation of brain tumor image to classify CSF leakage of brain tumor images using classification to improve the accuracy of surgical patient results. Their analysis provides for a prospective comparison of image evidence from CSF leakage in patients during surgery.

A. Al-Badarneh et al [6] proposed a model based on enhanced prevalence K and medium template C (TKFCM) for identifying brain tumors on MRI. In the proposed algorithm, the template-based K-mean algorithm initializes

the hash by choosing a template based on image strength (gray level) (gray level). Additionally, using the C-Means Blurry Algorithm (FCM), they measure distances between the cluster midpoint and the cluster data points. Finally, an optimized FCM clustering algorithm was used to detect the tumor site based on the different characteristics of the tumor image. The simulation results showed that the proposed algorithm detected normal and abnormal brain tissue within 0.9412 seconds compared to minutes for other algorithms.
W. Wu and colleagues [7] proposed a technique in which multimodal MRI images are divided into super-pixels to help represent samples correctly and reduce sampling problem. They removed super pixel properties using multi-level Gabor wave filters. The support vector (SVM) model and the affinity metric model for tumors dependent on characteristics have been learned to address the shortcomings of previous generative models.

N. Samina et al.[8] have studied four cluster-based image segmentation algorithms from contemporary literature. Both these methods have modified the objective function of the traditional FCM and used spatial information in the objective function of the standard FCM to solve the problems associated with the conventional FCM.

Venkatesh and M.Judith Leo [9] suggested. They included these traits in the k-Nearest Neighbor classifier (k-NN) and used them to classify neoplasms either malignant or benign. They mentioned that the proposed algorithm is as accurate as 85%.

S. Saeed and A. Abdullah [10] proposed a segmented-based machine learning algorithm to detect brain tumors in MRI images datasets. They focused on image segmentation techniques and histograms to achieve better results. S. Chowdhary et al.[11] suggest that the CSF leakage can be created due to abnormalities in the brain tumor that can also be validated in patient cohorts.

L. A. V.D. Kleij et al.[12] suggested an update to their predecessors where they built a short series to focus on CSF based on CSF's long T2. Data obtained with the CSF MRI series were analyzed to automatically obtain the brain parenchymal volume (BPV) and the intracranial volume (ICV). They proposed a change to test the accuracy of BPV and ICV assessments of LCR MRI sequences and to validate LCR MRI sequences by combining it with T1-based 3D image segmentation techniques. They concluded that the short computation time for rapid CSF MRI sequencing is preferable to 3D T1 sequencing as the image segmentation is performed by conventional methods.

K. Usman and K. Rajpoot [13] suggest the prospective improvement in the present technique. They anticipated that their research would provide the reader with a better knowledge of comparative study with the K-NN algorithm on individual establishment methods such as “k-mean” and “standard deviation” and modality direction. They proposed a training group with relevant information on different dimensions of research. In each group, the above the results are compared and techniques are applied. It was concluded that better the mean values than standard deviation.

Junejo et al. [14] used a marker controlled watershed segmentation method using input from 3D MR image to segment brain tumors as 2D MRI provides a confined view of the brain. The framework proposed in their research gives a thorough view of the brain with segmented areas showing the tumor. Their results show higher accuracy as compared with other methods.

H. Mihara et al.[15] suggested a controlled four-dimensional (4D) light field image segmentation technique using a graph-cut algorithm. 4D light field data contains redundancy of intrinsic depth information that varies from 4D hyper-volume. They identified two adjacent types of rays (spatial and angular) in light field data to ensure redundancy. They also designed a learning-based probability of using presence and difference indices to obtain better segmentation accuracy. They demonstrated the feasibility of their approach using numerical estimation and some light field editing software applications through both real-world and synthetic light fields.

4. Method

Here is a following instruction of the given proposed work 4.1 Light Field Segmentation of 4-Dimensional with Spatial and Angular Component

In [15], a lumigraphy method has been adopted to illustrate rays in three-dimensional space. The ray is defined by the two points of intersection with "u-v" and "x-y" in 3D coordinates. The ray can be described as a point in the 4D scale as p = (u, v, x, y), and p is the intensity parameter. A representation of lumigraph can be transformed into *x* and *y* planes and also *u* and *v* planes converted into images. H. Mihara et al. [15] describe an easy-to-understand multi-component representation technique in which *u-v* and *x-y* combine together with multi-point representation.

4.2 Tools and Platform

In this section, we discuss the tools of proposed method given below:

Manuscript received February 5, 2024
Manuscript revised February 20, 2024
https://doi.org/10.22937/IJCSNS.2024.24.2.5

4.2.1 Light Field Tool (LFT) Segmentation in 4D

There are several ways to represent the four-dimensional light field segmentation process. We have adopted the lumigraph method to illustrate rays in three-dimensional space. Two intersection points with "*u-v*" and "*x-y*" in the three-dimensional coordinates define the beam (ray). The radius can be expressed as a point at a distance of 4D, such as p = (u, v, x, y), and the strength of p is represented by lp where "l" denotes the representation of the lumigraph. The representation of the lumigraph can be transformed into several representations containing the pixels of the image. In a multi-point view, the u-v and x-y parameters are the same as the view and image planes. There are several ways in which the four-dimensional light field segmentation process can be interpreted. Here we explain the form of multi-component representation that is simple to understand.

4.2.2 Description of 4D Light Field Segmentation with Spatial and Angular Component

The four dimensional light field segmentation techniques use the diagram-cutting algorithm (grab cut). In our method, the 4D light field information contains the given data and frequency (range), in which we apply the contour method to partition a 4D hyper-volume. We recognize the two neighboring rays (spatial and angular) in light field segmentation. For segmentation goals, additionally a method of light field tools is considered a function of an image that utilizes indications of appearance and variety. To demonstrate the efficiency of our proposed method, we explain the calculation of image size and resolution by using the graph cutting technique which determines the size of selected image (The MRI data of brain images has been obtained from the Web Brain Database of the National Cancer Research Database-UK). In this research, we deal with tumor images and use the malignant cancer segments in MATLAB to calculate the size of CSF leakage in the brain. These images essentially represent the selected areas of tumor with CSF leakage after computing the size of cancer with the help of MATLAB programming. We screened 200 brain samples using the MATLAB programming that identified the place of brain cancer and also implemented lumigraph and multi-view representation

Table 1: CSF leakage identification

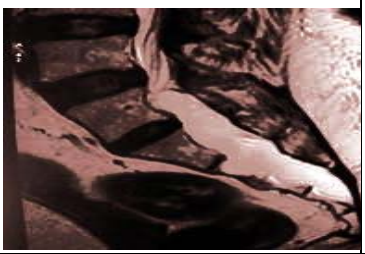
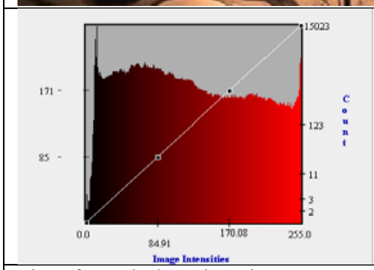
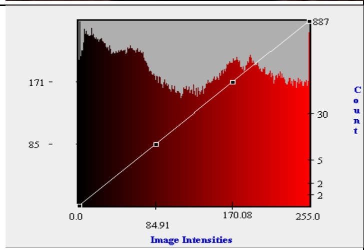
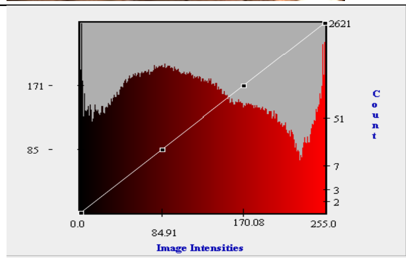
IMAGE 1	**IMAGE 2**	**IMAGE 3**
		
		
		
Size of CSF leakage in Spine	Size of CSF leakage in Spine	Size of CSF leakage in Spine
max_c =15023	max_c =887	max_c =2621

Table.1 demonstrates the initial 3D pictures of the spinal cord CSF leakage, which are transformed into 4-dimensional pictures. We generate a histogram through MATLAB to determine the precision and size of the spinal cord CSF leakage. The graphical representation shows image 1 which indicates the size of CSF leak, which is more prominent than the prior picture in the tumor region. After implementing the 4D image process, the image converts into grayscale and it clearly appears that the size of the CSF tumor is 0. Here is another value of maximum intensity (max_c) representing the intensity of the CSF-tumor which is 15023, showing the appearance of CSF leakage. Image 2 shows the size of the CSF-tumor which is still 0, but the value of the statistical graph is different which is 2550, showing the situation of CSF leakage which is slowly increasing in the brain and damaging the brain cells; the max_c of image 2 is 887. Image 3 also shows the size of the CSF-tumor is 0 and the statistical graph shows the size of tumor value is 2550 but the max_c is 2621, which represents the tumor shell in hard form after the deposits of CSF in the brain. This is basically the leakage from initial to final stage of the tumor in the spinal cord which is shown in Table 1.

4.3.2 KNN Algorithm (K-Means Clustering)

We know that K-Nearest neighbor, also known as *k*-NN, is a part of the supervised learning algorithms group and is a robust tool for creating more complex techniques for example artificial neural networks or support vector machines.

Table 2: k-NN Testing Result

SNO.	No. of Structures	Model Category	Evaluation Criteria	Acc	ROC	Sen	Spec
1	Decision_ Tree	Stated: Satisfactory Max_ splits: 100 Splits_ criterion: Multiplicity _Directory Substitute_ decision splits: off	Training Time Testing Time Prediction Speed	3.533Sec 72.7% 600Obs/Sec	0.83	93%	28%
2	Decision1_ Fine-Tree	Stated: Satisfactory _Tee Max _ splits: 100 Splits _criterion: Multiplicity_ Directory Substitute _decision splits: turn off	Training Time Testing Time Prediction Speed	333.53 Sec 90.9% 59Obs/Sec	0.83	93%	28%
3	Decision _Medium-Tree	Stated: Medium_ Tree Max _ splits: 20 Splits_ criterion: Multiplicity_ Directory Substitute_ decision splits: turn off	Training Time Testing Time Prediction Speed	260.4Sec 90.9% 29Obs/Sec	0.83	85%	60%
4	Coarse_ Tree	Stated: Coarse _Tree Max_ splits: 4 Splits_ criterion: Multiplicity_ Directory Substitute_ decision splits: turn off	Training Time Testing Time Prediction Speed	510.54Sec 90.9% 1600Obs/Sec	0.83	93%	28%
5	Linear _Classify	Stated: Linear Classify Covariance Configuration: Full	Training Time Testing Time Prediction Speed	507.42Sec 86.4% 150Obs/Sec	0.91	87%	70%
6	Quadratic_Discriminant	Stated: Quadratic_ Discriminant Covariance_ Configuration: Full	Training Time Testing Time Prediction Speed	510.28Sec 90.9% 180Obs/Sec	0.91	73%	70%
7	Logistic_regression	Stated: Logistic_ regression	Training Time Testing Time Prediction Speed	502.7Sec 86.4% 190Obs/Sec	0.89	73%	70%
8	SVM_ Linear	Stated: Linear _SVM Kernel _Function: Linear Kernel _Sale: Spontaneous Box _Limitation_ Level: 01 Multiclass _Scheme: One –to-One Regular _Records: True	Training Time Testing Time Prediction Speed	459.65Sec 93.2% 210Obs/Sec	0.94	92%	60%
9	SVM_ Coarse	Stated: Coarse_ SVM Kernel_ Function: Coarse Kernel_ Sale: Spontaneous Box_ Limitation_ Level: 01 Multiclass _ Scheme: One –to-One Regular_ Records: True	Training Time Testing Time Prediction Speed	182.97Sec 86.4% 690Obs/Sec		85%	60%

Manuscript received February 5, 2024
Manuscript revised February 20, 2024
https://doi.org/10.22937/IJCSNS.2024.24.2.5

10	SVM_ Quadratic	Stated: Quadratic _SVM Kernel _Function: Quadratic Kernel_ Sale: Spontaneous Box _Constraint _Level: 01 Multiclass _Scheme: One –to-One Regular _Records: True	Training Time Testing Time Prediction Speed	458.11Sec 90.9% 710Obs/Sec	0.97	92%	60%
11	SVM_ Cubic	Stated: Cubic _SVM Kernel _Function: Cubic Kernel _Sale: Spontaneous Box _Limitation _Level: 01 Multiclass _ Scheme: One –to-One Regular_ Records: True	Training Time Testing Time Prediction Speed	457.32Sec 93.2% 710Obs/Sec	0.97	10%	90%
12	SVM_ Fine _Gaussian	Stated: Gaussian _SVM Kernel_ Function: Gaussian Kernel _Sale: 0.79 Box _Limitation _Level: 01 Multiclass _Scheme: One –to-One Regular_ Records: True	Training Time Testing Time Prediction Speed	15.624Sec 93.2% 150Obs/Sec	0.91	92%	60%
13	SVM_ Medium _Gaussian	Stated: Medium Gaussian _SVM Kernel_ Function: Gaussian Kernel_ Scale: 3.2 Box _Limitation _Level: 01 Multiclass _Scheme: One –to-One Regular _Records: True	Training Time Testing Time Prediction Speed	455.89Sec 93.2% 720Obs/Sec	0.96	92%	60%
14	SVM_ Coarse_ Gaussian	Stated: Coarse Gaussian _SVM Kernel_ Function: Gaussian Kernel_ Sale: 13 Box _Limitation _Level: 01 Multiclass_ Scheme: One –to-One Regular_ Records: True	Training Time Testing Time Prediction Speed	455.728Sec 93.2% 1100Obs/Sec	0.91	85%	60%
15	KNN _Fine K=1	Stated: Fine_ KNN No._ of_ Neighbor: 1 Distance _Metric: Euclidean Distance_ Weight: Equal Regular_ Data: True	Training Time Testing Time Prediction Speed	449.36Sec 90.9% 250Obs/Sec	0.89	92%	60%
16	KNN _Medium (Euclidean) K=10	Stated: Medium _KNN No._ of_ Neighbor: 10 Distance _Metric: Euclidean Distance_ Weight: Equal Regular _Records: True	Training Time Testing Time Prediction Speed	447.88Sec 93.2% 550Obs/Sec	0.94	0%	30%
17	KNN _Coarse K =100	Stated: Coarse _KNN No._ of _Neighbors: 100 Distance_ Metric: Euclidean Distance_ Weight: Equal Regular _Records: True	Training Time Testing Time Prediction Speed	447.48Sec 70.5% 730Obs/Sec	0.44	92%	60%
18	KNN_ Cosine K=10	Stated: Cosine _KNN No._ of _Neighbour:10 Distance _Metric: Euclidean	Training Time Testing Time Prediction Speed	447.02Sec 93.2% 220Obs/Sec	0.93	92%	60%

		Distance _Weight: Equal Regular _Records: True					
19	KNN_ Cubic K=10	Stated: Cubic_ KNN No. _of _Neighbor: 10 Distance _Metric: Minkowski(cubic) Distance_ Weight: Equal Regular_ Records: True	Training Time Testing Time Prediction Speed	446.18Sec 93.2% 310Obs/Sec	0.94	92%	60%
20	KNN _Weighted K=10	Stated: Weighted_ KNN No._ of Neighbor: 10 Distance _Metric: Euclidean Distance_ Weight: Square_ Inverse Regular_ Records: True	Training Time Testing Time Prediction Speed	445.64Sec 93.2% 570Obs/Sec	0.97	0%	30%
21	Ensemble Boosted Tree (ADA Boost)	Stated: Boosted _Tree Ensemble Scheme: ADA _Boost Beginner Category: Decision _Tree Max _splits: 20 No. of_ Beginner: 30 Learning_ Rate: 0.1	Training Time Testing Time Prediction Speed	444.81Sec 70.5% 240Obs/Sec	0.96	92%	60%
22	Ensemble Bagged Tree Ensemble method (Decision Tree)	Stated :Bagged _Tree Ensemble_ Scheme: Bagged Beginner _Category: Decision Tree No. of Beginner: 30	Training Time Testing Time Prediction Speed	441.52Sec 93.2% 80Obs/Sec	0.96	92%	60%
23	Ensemble subspace (Discriminate)	Stated: Subspace _Discriminant _Ensemble Scheme: Subspace_ Beginner _Type: Classify No. _of _Beginner: 30 Subspace _Measurement: 5	Training Time Testing Time Prediction Speed	438.77 Sec 90.9% 31Obs/Sec	0.91	85%	60%
24	Ensemble Subspace (KNN)	Stated: Subspace_ KNN Ensemble_ Scheme: Subspace Beginner _Category: Nearest Neighbor Decision Tree No._ of_ Beginner: 30 Subspace_ Measurement: 5	Training Time Testing Time Prediction Speed	434.28Sec 79.5% 33Obs/Sec	0.91	65%	0%
25	RUS Boost Tree Ensemble method (Decision Tree)	Stated: RUS_ Boost _Tree Ensemble _Scheme: RUS_ Boost Beginner _Category: Decision _Tree Max _ splits: 20 No._ of _Beginner: 30 Erudition_ Rate: 0.1	Training Time Testing Time Prediction Speed	429.53Sec 90.9% 250Obs/Sec	0.96	85%	60%

Manuscript received February 5, 2024
Manuscript revised February 20, 2024
https://doi.org/10.22937/IJCSNS.2024.24.2.5

5. Results and Discussion

In this section, we discuss the classification results of the *k*-NN algorithm which is based on semi-supervised machine learning. We worked on four segmentation procedures of brain cancer and CSF leakage using 4D images. The findings of the tests are conducted by a dataset called "Malignant Brain Cancer with CSF Leakage" comprising nine functions with 25 characteristics. The initial information set with 4-dimensional information is then randomly divided into four segmentation methods and datasets, which conclude the accuracy of 96.9% as shown in the results. Using the training dataset, we trained several supervised machine-learning models namely Decision Tree, Decision Medium-Tree, Coarse Tree, Linear Discriminate, Quadratic Discriminant, Logistic regression, SVM (Linear), SVM (Coarse), SVM (Quadratic), SVM (Cubic), SVM (Fine Gaussian), SVM (Medium Gaussian), SVM (Coarse Gaussian), KNN Fine (K=1), KNN Medium (Euclidean)(K=10), KNN Coarse (K =100), KNN Cosine (K=10), KNN Cubic (K=10), KNN Weighted (K=10), Ensemble Boosted Tree (ADA Boost), Ensemble Bagged Tree, Ensemble subspace(Discriminate), Ensemble Subspace(KNN) and RUS Boost Tree. Once the models are trained, the new data from the testing dataset are predicted. Measuring four evaluation metrics namely Sensitivity (Sen), Specificity (Spe) ROC, and precision (Acc), we evaluate the efficiency of the proposed techniques. The experimental results are mentioned in Table 2. Among twenty-four machine-learning methods, nine models: SVM (Linear), SVM (Cubic), SVM (Fine Gaussian), SVM (Medium Gaussian), SVM (Coarse Gaussian), KNN Medium (Euclidean)(K=10), KNN Cosine (K=10), KNN Cubic (K=10), KNN Weighted (K=10) and Ensemble Bagged Tree achieved the accuracy of 93.2%. Each classifier's computational time is also calculated to assess their complexity and Table 2 shows each classifier's processing time.

The *k*-NN Coarse is comparatively small in computational time, but the practical use satisfies their sensitivity and specificity. The other remaining models: Decision Tree, Decision Medium-Tree, Coarse Tree, Linear Discriminate, Quadratic Discriminant, Ensemble subspace (Discriminate), and RUS Boost Tree achieved 90.9% accuracy. SVM (Coarse), Logistic Regression and Ensemble Subspace (KNN) are among three techniques that produce 86.40-79.5% precision. We also compare the results obtained with the results reported in the literature by means of related work. The comparison shows that the k-NN models (K = 1-100) in our study outperformed the previous studies, giving an accuracy of 96.9%. Our methods have improvements over previously reported works in the literature due to its shorter calculation time and taking the far-distant values. However, it does not meet the 100% accuracy for all methods but in 4 dimensional, the *k*-NN gives better results as compared to previous work.

6. Conclusion

We have proposed four new approaches to classify MRI images (4D) using hybrid *k*-NN algorithm techniques at the beginning of tumors or CSF development. Our research addressed the issues of accuracy and the efficiency of missing *k*-NN hybrid values. *K-NN* missing values are one of hottest research issues especially in 4D frequency. The proposed 4D method aims at improving the sensitivity of the segmentation process of an image and using *k*-NN algorithm to improve the efficiency, as well as to increase the segmentation range of the tumor area as compared to other reported methods. *k*-NN significantly enhances the performance of distant values mainly from two key factors which is brain cancer and CSF leakage. The first improvement resides in the fact that the *k*-NN algorithm constructs an Ensemble of adjacent values with parameters of neighborhood size (K=1-100). These values show the efficiency of our method using classification of *k*-NN algorithms. The performance our research methods is to focus the measured four evaluation metrics, i.e., Sensitivity (Sen), Specificity (Spe) ROC, and accuracy (Acc). The second improvement is that the *k*-NN algorithm provides an efficient weighting for all far values which makes the projections much less sensitive to the neighborhood size parameter as compared to the previous works. One of the other focuses of this research is to increase the accuracy of tumor-dependent images and apply a segmented area to the tumor for the cause of the CSF leakage. These images mainly represent the size of the tumor and need to calculate the frequency, intensity, and tumor mass with high accuracy using the latest 4D segmentation LFT tools.

We take the longer values of *k*-NN (K=1-100) and show their performance for the new proposed 4D model. We increase the accuracy of missing values of *k*-NN algorithms. Therefore, the proposed *k*-NN algorithm method that can be suitable for the latest segmentation of 4D images in various tasks, such as image processing, increasing the sensitivity within the region of curve and various deep learning tasks. In future work, we plan to extend our proposed improved methods using SVM classification in hybrid models.

Reference

[1]I. Altaf, AH. Vohra and S. Shams, “Management of cerebrospinal fluid leak following posterior cranial fossa surgery,” *Pakistan Journal of Medical Sciences*, vol.32, no.6, pp.1439-144, 2016.

[2]L.G. Alexander, A. Axel , J. Batiller, S. Eljamel , J. Gauld . P. Jones et al., “A multi-center, prospective, randomized

controlled study to evaluate the use of a fibrin sealant as an adjunct to sutured dural repair," *British Journal of Neurosurgery*, vol.29, no.1, pp.11-17, 2015.
[3]Saeed, Soobia, and Afnizanfaizal Abdullah, "Recognition of brain cancer and cerebrospinal fluid due to the usage of different MRI image by utilizing support vector machine," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no.2, pp.619-625, 2020.
[4]S. Saeed and A. Abdullah, "Investigation of a Brain Cancer with Interfacing of 3-Dimensional Image Processing," *In Proc. International Conference on Information Science and Communication Technology (ICISCT)*, Karachi, Pakistan,PP.1-12, 2019.
[5]A. H. Sin , G. Caldito, D. Smith, M. Rashidi, B. Willis and A. Nanda, "Predictive factors for dural tear and cerebrospinal fluid leakage in patients undergoing lumbar surgery," *German Cancer Research Center*,vol.5, no.1, pp.224-227, 2019.
[6]A. Al-Badarneh, H. Najadat and A. M. Alraziqi, "A classifier to detect tumor disease in brain MRI brain images," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 6, no.12, pp.2264-2269, 2013.
[7]W. Wei, A.YC Chen, L. Zhao, and J. J. Corso, "Brain tumor detection and segmentation in a CRF (conditional random fields) framework with pixel-pairwise affinity and superpixel-level features," *International journal of computer assisted radiology and surgery*, vol.9, no. 2, pp.241-253, 2014.
[8]S. Naz, H. Majeed, and H. Irshad, "Image segmentation using fuzzy clustering: A survey," *In Proc. 6th international conference on emerging technologies (ICET)*, IEEE, Islamabad, Pakistan, pp. 181-186, 2010.
[9]Venkatesh and M.Judith Leo. "MRI Brain Image Segmentation and Detection Using K-NN Classification," *In Proc. International Conference on Physics and Photonics Processes in Nano Sciences*, India, vol. 1362, no. 1, pp. 1-6, 2019.
[10] S. Saeed, A. Abdullah and NZ Jhanjhi," Investigation of a Brain Cancer with Interfacing of 3-Dimensional Image Processing," *Indian Journal of Science & Technology*, vol.12, no.32, pp.1-6, 2019.
[11] S. Chowdhary, S. Damlo and M. C. Chamberlain, "Cerebrospinal Fluid Dissemination and Neoplastic Meningitis in Primary Brain Tumors," *Journal of Mofitt Cancer Center,* vol. 24, no.1, pp.1-16, 2017.
[12] L. A. V.D. Kleij , J. D. Bresser, J. Hendrikse, J. C. W. Siero, E. T. Petersen et al., "Fast CSF MRI for brain segmentation; Cross-validation by comparison with 3D T-based brain segmentation methods, " *Plos One*, vol. 13, no.4, pp.1-14, 2018.
[13] K. Usman and K. Rajpoot, "Brain tumor classification from multi-modality mri using wavelets and machine learning," *Pattern Analysis and Application*, vol.20, no.1, pp.871–881, 2017.
[14] Junejo, A. Zahid, S. A. Memon, I. Z. Memon, and S. Talpur, "Brain Tumor Segmentation Using 3D Magnetic Resonance Imaging Scans." *In Proc. 2018 1st International Conference on Advanced Research in Engineering Sciences (ARES)*, IEEE, Dubai, pp. 1-6, 2018.
[15] H. Mihara, T. Funatomi, K. Tanaka, H. Kubo, Y. Mukaigawa et al., "4D light field segmentation with spatial and angular consistencies," *In Proc.2016 International Conference on Computational Photography (ICCP)*, IEEE, Evanston, pp. 1-8, 2016.

Manuscript received February 5, 2024
Manuscript revised February 20, 2024
https://doi.org/10.22937/IJCSNS.2024.24.2.5

Soobia Saeed is working as an Assistant Professor, Head of publication Department, and Coordinator of Seminars and Training at Institute of Business & Technology-IBT, Karachi, Pakistan. Currently, she is a Ph.D. Scholar in software engineering, from University Teknologi Malaysia-UTM, Malaysia She did MS in Software Engineering from Institute of Business & Technology- IBT, Karachi, Pakistan, and Masters in Computer Science fromInstitute of Business & Technology-IBT, Karachi, Pakistan and Bachelors in Mathematical Science from Federal Urdu University of Art, Science & Technology (FUUAST), and Karachi, Pakistan. She is a farmer research Analytic from University Teknologi Malaysia and supervises ICT & R and D funded Final Year Project (FYP).

Afnizanfaizal Abdullah is a senior lecturer at the School of Computing, with a PhD. in Computer Science, specializing in artificial intelligence techniques for analyzing biological data. My research interests are in the designing of machine learning algorithms for healthcare applications in the cloud environments. In 2015, I have co-founded Synthetic Biology Research Group to drive innovation in research and development of healthcare, biotechnology, and environment areas through computing and engineering. I am also active in engaging with industrial partners and professional communities to contribute the knowledge and skills for the public.

Noor Zaman has completed his PhD. in IT from University Technology Petronas (UTP) Malaysia. He has 19 years of teaching and administrative experience internationally. He has an intensive background of academic quality accreditation in higher education besides scientific research activities, he had worked for academic accreditation for more than a decade and earned ABET accreditation twice for three programs at College of computer sciences and IT, King Faisal University Saudi Arabia. He also worked for National Commission for Academic Accreditation and Assessment (NCAAA), Education Evaluation Commission Higher Education Sector (EECHES) formerly NCAAA Saudi Arabia, for institutional level accreditation. He also worked for National Computing Education Accreditation Council (NCEAC) Pakistan. He has experienced in teaching advanced era technological courses including, Mobile Programming (Android), Mobile Computing and .Net Framework programming besides other processing, and software development methodologies and models. Undergraduate and postgraduate courses, graduation projects and thesis supervision.

Noor Zaman has authored several research papers in ISI indexed and impact factor research journals\international conferences, edited 07 international reputed Computer Science area books, focused on research students, has many journal, IEEE conferences and book chapter publications to his credit. He has successfully completed more than 18 international funded research grants. He is Associate Editor, Regional Editor, Editorial board member, PC member, reviewer, Keynote speaker for several reputed international journals and conferences around the globe. He also chaired international conference sessions and presented session talks internationally. He has strong analytical, problem solving, interpersonal and communication skills. His areas of interest include Wireless Sensor Network (WSN), Internet of Things IoT, Security, Mobile Application Development, Ad hoc Networks, Cloud Computing, Big Data, Mobile Computing, and Software Engineering.