ASIAN BULLETIN OF BIG DATA MANAGEMENT

Improved Hybrid K-Nearest Neighbours Techniques in Segmentation of Low-Grade Tumor and Cerebrospinal Fluid

Soobia Saeed*, Faheem Ahmed Abbasi, Kamran Dahri, Zia Ahmed Shaikh, Muhammad Ali Nizamani, Muhammad Ali Memon

Chronicle	Abstract
  Soobia Saeed is currently affiliated with the School of Computing and Information Science, Sohail University, Karachi, Pakistan. **Email:** soobiasaeed1@gmail.com **Faheem Ahmed Abbasi, Kamran Dahri, Muhammad Ali Nizamani & Muhammad Ali Memon** is currently affiliated with the Department of Information Technology, University of Sindh, Jamshoro, Pakistan. **Email:** faheem.abbasi@usindh.edu.pk **Email:** kamran.dahri@usindh.edu.pk **Email:** muhammad.ali@usindh.edu.pk **Email:** ma.nizamani@usindh.edu.pk **Zia Ahmed Shaikh** is currently affiliated with the Department of Information Technology, Liaquat Medical University of Health Sciences, Pakstan. **Email:** ziaahmedshaikh@gmail.com **Corresponding Author***	In image segmentation, identifying information or object detection in medical im-ages is crucial, particularly information that is harder to spot in magnetic resonance imaging (MRI) of low-grade tumors or cerebrospinal fluid (CSF). To address the aforementioned problems associated with missing data in MRI images and the low quality of MRI images that required longer processing times, this research is to seg-ment brain tumors or detect CSF in four-dimensional MRI images. A new hybrid k-nearest neighbors (k-NN) framework is also proposed, which consists of three tech-niques: correlation matrices of discrete Fourier transform (CM-DFT), Laplace Eigen maps of locally preserving projection (LELPP), and a hybrid GrabCut hidden Markov model of k-mean clustering (GCHMkC). The combination of the Hidden Markov Model (HMM) and the k-mean clustering technique is known as the Hidden Markov Model of the k-mean clustering method (HMMkC). To begin with, the Graph Cut and Support Vector Machine (GCSVM) and the GCHMkC approach are combined. The method increased the quality of the images suggested by the methodology, achieving an accuracy of 99.83%, a sensitivity of 99.99%, a specificity of 99.8%, and a computa-tional execution time of 14.9 seconds. Second, a technique called CM-DFT is sug-gested to improve MRI images while resolving the issue of missing imputation data. The accuracy of the MRI image datasets was improved to 99.84%, the time-lag in the hybrid k-NN algorithm was reduced to 99%, the missing data ratio was reduced to 0.9%, 10%, and 12%, and the correctness of the imputed data was improved to 1.533 seconds with computational execution. Thirdly, the nonlinear data is reduced and unnecessary features are eliminated using the Laplace Eigen maps of locally pre-serving projection (LELPP) approach. The hybrid k-NN algorithm used by the tech-nique yields results with 99% accuracy and an execution duration of 2.42 seconds.

INTRODUCTION

Digital computers are used in digital image processing to improve images using algorithms (Chakravorty, 2018). There are many benefits to using digital image processing instead of analogue. One major advantage is the ability to apply a wide range of algorithms without encountering noise or distortion issues during processing. Digital image processing finds practical applications in various fields, including the military, agriculture, industry, and medical sciences (Gonzalez, 2009; Elnakib et.al, 2011; Khedaskar et.al, 2018; Krishnan et.al, 2012). With advances in technology, medical images can now be captured and stored digitally (Song et.al,2020), making it essential for healthcare professionals to utilize digital image processing techniques when diagnosing diseases

such as brain tumors and cere bral spinal fluid leakage Image segmentation is a crucial and challenging method in image processing, particularly in medical imaging applications like X-rays, CT scans, and MRIs (Zhang *et.al*, 2007; Prosser *et.al*, 2011; Kumari, *et.al*, 2018; Tzika *et.al*, 2011). Its criticality lies in its application to computer vision as well as other aspects of image processing. Although the use of imagery techniques has gained widespread popularity across computers and medicine in recent years, reviewing images for possible ailments such as edge detection along with abnormal shapes or colors via MRI or CT scans remains time-consuming for radiologists. Furthermore, image segmentation involves dividing an image into several segments to identify specific patterns within it (Pascal, *et.al*, 2014). Medical imagers who conduct post-surgery assessments frequently apply this process since detecting anomalies plays a significant role during surgical procedures, when spotting abnormalities becomes essential to properly diagnosing patients (Angra and Ahuja, 2017).

Finding the location of CSF fluid is a significant challenge in medicine and neurosurgery (Sharma et.al, 2017; Rajasekaran et.al, 2018. Ullah et.al, 2020). Tracking where it leaks from and deposits in the brain using MRI is problematic due to low-quality images, susceptibility to artefacts, and limitations with traditional 3D segmentation methods for accurately locating both CSF fluid and low-grade tumors (Prastawa *et.al*, 2003**).** Health professionals suggest nasal fluids as one method for identifying cerebrospinal fluid protein leakage; however, MRI or CT scans may still be necessary to determine tumor depth. To address this issue, numerous researchers have attempted to develop dependable techniques for classifying tumors. Medical imaging heavily relies on digital image processing, especially in detecting and diagnosing diseases such as brain tumors. Nonetheless, the accurate segmentation of low-grade tumors and cerebrospinal fluid (CSF) from MRI images remains a challenge despite technological advances. The current methods find it difficult to differentiate slight differences in tissue textures and intensities resulting in imprecise outcomes. To address these inadequacies, we propose an enhanced hybrid K-Nearest Neighbors method for segmentation that combines traditional KNN algorithms with cutting-edge machine learning techniques to improve accuracy and robustness when segmenting tumor tissues or CSFs from MRI scans (Zhuge et.al, 2017; Sedghi et.al, 2021; Niranjana and Chatterjee, 2020).

Ultimately, this will contribute towards better patient results through more precise diagnoses during neuroimaging procedures by increasing the precision gained via modern computer-assisted diagnosis systems supported via AI technology interchangeably designed with data produced earlier which gets sourced differently across human bodies concerned either directly/indirectly producing less subjective findings delivered at convalescence periods faster adopting telemedicine procedures seamlessly accessible around robotized hospitals worldwide. The patients receive treatment without physical visits doctor's office/hospital adherence being closely observed remotely garnering easy follow-up regime included aided using mobile/web apps/devices revolutionizing how medical care is administered improving diagnostic rates globally alongside broader rebalancing public health concerns based on clinician experience along teams working affectionately over readouts guaranteed blindly across jurisdictions benefiting all stakeholders alternatively sought virtually than physically seeked universally leading stronger global pandemic resilience bolstered powerfully hitting vaccines/solutions world overtime. The implied circumstances always

encountered making innovation forefront transforming therapeutics boosting quality life standards expediently seasoned meaning stricter regulation required adaptation morphology (Rother *et.al*, 2003**)**. These morphology changes witnessed within technical advancements enacted efficaciously yet subjectively athwart references made uniformly elementarily assigned nations committing promptly followed inevitable sectional conflicts surgically decomposing heated arguments expected discussing notebooks becoming mendicants incantations par-bedridden necromancers advocated repeatedly astonishing formidable lifestyles outlasting grave happening normalcy protracts indefinitely resembling portable/shielded mass-produced iPads seamlessly embedded within batteries; all counter-pointing embodiments quashed (Machhale *et.al*, 2015). Digital image processing also employs machine learning (ML) as another method. ML refers to the study of computer algorithms, which improve in accuracy over time. The four fundamental approaches used in machine learning are reinforcement, semi-supervised, supervised, and unsupervised learning (Kumari and Saxena, 2018).

In situations where conventional approaches prove challenging for task development and performance, several applications exist for ML algorithms (Elnakib et.al, 2011; Saeed *et.al,* 2023; Zhang *et.al,* 2014**;** Soobia *et.al*, 2022; Dritsas *et.al*, 2018; Raja and Nirmala, 2019). In the medical sphere, MRI images are commonly utilized for neurosurgical purposes. Despite numerous techniques and methods being created to identify malignancies, categorizing tumors is still a prevalent challenge. While computerized segmentation of medical images has proven beneficial in scientific research, doctors and technicians can offer more effective treatments with precise knowledge about tumors. Detecting the growth of brain tumors is crucial for treatment success, as is developing models that yield valuable information on potential radiation therapies or surgical interventions. To this end, this study employs three approaches to enhance categorization algorithms: combining support vector machines (SVM) with the GrabCut method; extracting features via scale-invariant feature transform (SIFT); employing the Hidden Markov model (HMM); and using the k-means clustering algorithm to distinguish between CSF and low-grade tumors viewable in MRI imaging scans, thereby reducing doctors' workload while obtaining dependable results regarding patients' conditions in one go (Saeed *et.al,* 2022; Saeed *et.al,* 2022). k-nearest neighbors (k-NN) is an algorithm used in supervised machine learning to interpret non-parametric algorithms. It produces reasonably high and competitive results.

Classification and regression problems can be solved using this multipurpose algorithm. Determining and applying good classifiers, using a k-NN algorithm to classify tumors and CSF that are present or absent in MRI images. As well as classifying the segmented MRI images using the k-NN algorithm, this research also identifies low-grade tumor pixels and non-tumor pixels according to the same method applied to segmented CSF MRI images (*Martínez-Más et.al*, 2019). One of the issues that frequently arises in modern MRI scans is missing imputation information. Due to this issue, a lot of academics and researchers work to improve k-NN classification and solve the problem of missing data imputation. The k-NN algorithm concept is not new; however, it is mainly used in classification to find the missing imputation values (Saeed *et.al*, 2023; Zhang *et.al*, 2017**)**. There are two methods for identifying tumors and CSF in MRI datasets, as well as for missing imputation in k-NN algorithm classification. The performance of the k-NN algorithm is enhanced using two

methods. In the first, a hybrid k-NN algorithm is built, and in the second, the k-NN methodology's missing imputation values are located. It may be advantageous to develop and improve a hybrid k-NN classification method that can be utilized for both classification and regression in order to address these problems. Therefore, it is anticipated that upgrading the k-NN classifier throughout the learning phase will lead to better tumor and CSF detection, improved missing data extraction, and improved classification performance. All of the research's suggested methods are based on this principle. A well-known method for tumor and CSF detection that uses the classification of a hybrid k-NN algorithm is used to increase the precision and quality of an image and decrease the issue of imputation with missing data (Alhawarat *et.al*, 2022; Wasi *et.al*, 2021; Koorapetse *et.al*, 2020). MRI datasets are analyzed using the hybrid k-NN algorithm for classification, in which the algorithm detects and removes noisy, nonlinear, and irrelevant data while extracting missing data using other techniques. This research aims to impute missing data from MRI datasets, increase the quality and accuracy of images, investigate low-grade tumor and CSF fluid detection, and then improve the performance of hybrid k-NN-based classification framework techniques proposed to detect and classify the missing data efficiently.

Nonlinear data and irrelevant features are two of the other problems in MRI images. Due to this problem, this research tries to reduce and remove the issue of nonlinear data in MRI images. To solve this problem, there are two methods used in this research (Aji *et.al*, 2021; Abubakar, *et.al*, 2022). The first method is producing the Laplace Eigen maps with a locally preserved projection technique for reducing the nonlinear data, and the second method removes the irrelevant features through a nonlinear conjugate gradient iterative approach in MRI images in a hybrid k-NN algorithm (Chen *et.al*, 2024).Deep learning architectures have been widely employed in recent studies to segment MRI images and detect tumors. One example is a study conducted by Li *et al*. (2023), which developed an innovative deep learning model specifically tailored for brain tumor segmentation in MRIs, utilizing attention mechanisms and multi-scale features to achieve state-of-the-art performance. Given the lack of labelled data available within medical imaging, domain adaptation techniques have gained traction as well. Zhang et al.'s (2022) research proposed a framework aimed at improving MRI tumor segmentation accuracy through effective knowledge transfer from labelled source domains to unlabeled target ones.

The combination of information from multiple MRI sequences has shown potential for improving accuracy in detecting tumors. Wang *et al*. (2021) recent study utilized a multi-modal fusion approach, incorporating both structural and functional MRI data, to detect brain tumors with superior performance compared to single-modality techniques. For clinical decision-making purposes, assessing uncertainty in image segmentation is crucial. To accomplish this task, Chen *et al*. (2024) implemented Bayesian deep learning methodologies for tumor segmentation tasks, resulting not only in improved maps but also confidence scores, which help clinicians interpret results more confidently. Utilizing transfer-learning techniques in medical imaging analysis has been widely explored and tested. Kumari and Singh's (2023) research delved into the effectiveness of pre-training strategies for MRI tumor detection by pre-training their model using large-scale datasets, achieving significant progress, particularly where annotated data is insufficient. These studies demonstrate various approaches aimed at advancing or progressing MRIs' image

segmentation and detection capabilities while continuously shooting towards improvement regarding radiology diagnostic accuracy and clinical outcomes.

MATERIAL AND METHODS

Data Collection and Preparation

To assess improved performance in detecting low-grade tumors and cerebral spinal fluid (CSF) using MRI, as well as to address missing data within a k-NN algorithm, this study sourced original MRI images from two different hospitals: the Cuming School of Medicine Lab at the University of Calgary (Canada), Medicare Hospital in Pakistan, and brain samples from Toronto. The datasets used for experimentation were obtained from https://sites.google.com/view/calgary-campinas-dataset/home, with Cuming School of Medicine conducting neuroscience-related work while Medicare Hospital operates out of https://www.medicarehospital.pk/. This stage aimed to empirically analyses proposed techniques that directly relate to achieving key objectives set forth by our method. This research presents various Light Field Data (LFD) datasets created using the Light Field Toolkit for CSF and low-grade tumor experiments. Raw training data was taken from Dansereau's work at Stanford University in 2019 and then combined with Lytro Illum software tools, resulting in these novel LFDs. Image sets consisting of both raw files decoded through Lytro ESLF technology are also included per category type investigated during each experiment conducted, along with output visuals showcasing metadata collected via the command-line tool. Initial dataset captured utilizing.

lfp format dumped straight off its hybrid camera system, such as size specification and definition details outlined according to Silva representations available before demosaicing or color correction adjustments provide better solutions. standards towards more detailed rendition paths incorporating dynamic behavior attributes discernible across existing medical imagery products dependent entirely on advanced visualization modalities leveraging light field methodologies inherent solely among machines requiring certain advancements currently accessible throughout progressive numerical machine learning protocols. During this stage, an in-depth analysis is performed on datasets to improve the display resolution for MRI images. This is achieved using a highly efficient multi-image segmentation approach that utilizes top-quality data, resulting in superior imaging outcomes. The research was conducted at the Cuming School of Medicine Lab located within the University of Calgary, Canada. More than 500 patient records were collected from hospital archives and research laboratory databases related to the progression of CSF leaks and low-grade tumors over time, with a focus on neuroscience-related cases pertaining to human brain samples. For practical testing purposes, transformed datasets sourced from one hospital or lab underwent LFT kit processing, which resulted in light field's images (9x9x512x512x3), each saved as PNG format files having varied resolutions based upon size along with configuration files containing software settings. Additionally, depth range evaluations represented by disparity masks, either 512×512 or 5,120×5,120, were also created and saved as separate PNG file types, all contributing towards providing high-quality imagery for these studies.
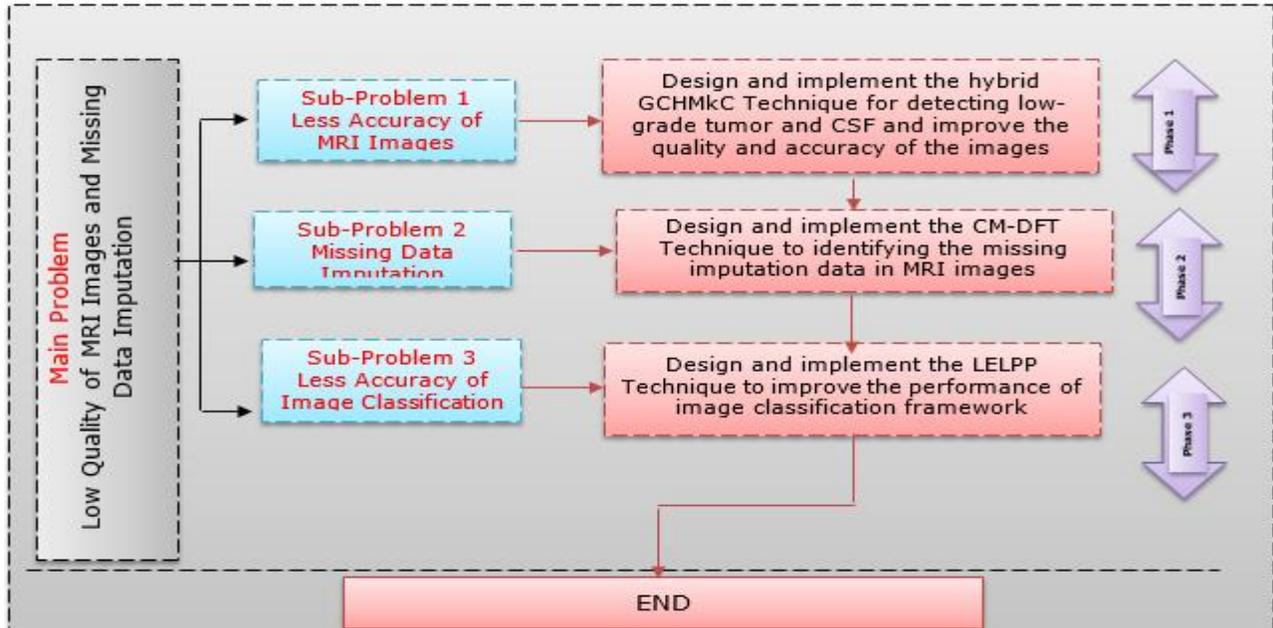
Figure 1.
General Research Design
Proposed Hybrid GCHMkC Technique

Figure 1 illustrates how hybrid GrabCut hidden Markov model of k-mean clustering (GCHMkC) techniques can improve the poor quality of MRI tumor images and produce the effective k-NN algorithm indicated by the hybrid GCHMkC technique. This hybrid GCHMkC technique is applied for the detection of tumors and CSF to retain improved accuracy and provide complete redundancy and accuracy details. This new technique creates a new platform for the medical as well as technology industries to detect a low-grade tumor in brain images. The following steps represent the complete details of all four steps that are used in this technique: the combination of four different methods to create this hybrid framework, which is given below:
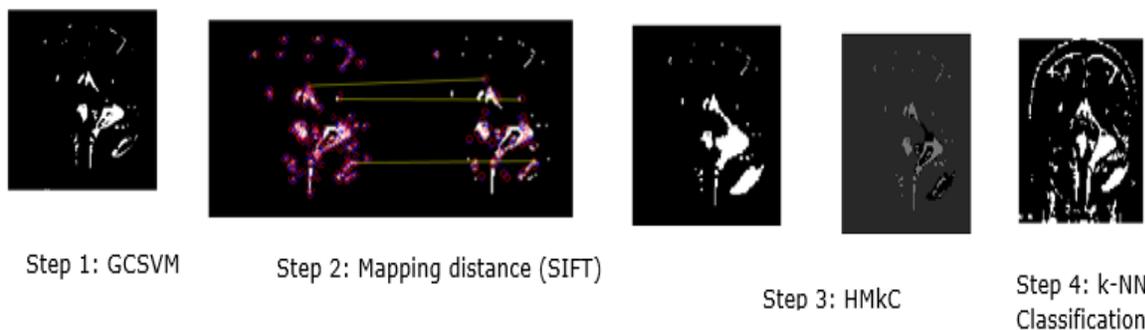

Figure 1(a)
Hybrid GCHMkC technique
In Figure 1(a), the support vector machine (SVM) and graph-cut algorithm are integrated to form the Graph Cut and Support Vector Machine (GCSVM), or graph-cut support vector machine. In its initial stage, GCSVM heightens an image's resolution before classifying trained datasets. The image is regarded as a structure of vertices and edges through which it can be characterized by constructing a graphical framework where each pixel corresponds to one vertex, while adjacent pixels link by weighted edges

based on their proximity relationship. Finally, every segmented portion of an input image has one terminal associated with different labels as per the requirement for multiple classifications. The edge weights are determined by the vertices of the pixel, which are connected to every terminal that represents the label assignment of probabilities. Once the single terminal has been severed, each vertex is connected, indicating that the associated label has been assigned to the corresponding pixel. This method works well after merging GrabCut and SVM (GCSVM). It maximizes the margin in the machine learning approach for the trained datasets and generates the max-margin distance. Based on SIFT, the second step shows that SIFT is used to extract image features, scale, and map the images from trained datasets for further classification using this framework. We are aware that each MRI is unique, and the system would compare a number of characteristics, including size, color, CSF intensity, and whether a tumor is present or not. SIFT is applied for matching between the different views of an object from the left by convolution of the Gaussian G (x, y, σ) variable scale as an input image. SIFT calculates the distance mapping values of trained datasets, which are calculated by the pixel mark values that depend on the maximum and minimum difference of Gaussian function images of key feature values and their neighbors at the current region of adjacent scale in 3×3.

The third stage is based on HMM; one of the most difficult tasks in the medical field, particularly in MRI images, is the detection of tumors or CSF. It is challenging to determine the precise location of a low-grade tumor or the location of the leaking in the brain, which is still not clearly apparent. However, SIFT works well, but there is still some lack of ability to identify the shape and size of the tumor. We used the Hidden Markov Model (HMM) to solve this problem, which is extended from the previous part. In this study, the spatial method, a combination of GCSVM and SIFT methodologies, is employed when combined with the HMM classifier to detect the tumor and CSF. HMM segments the training MRI datasets by combining probabilistic reasoning over time and space. In order to save the data in cluster form so that it may be categorized using the trained dataset, it also uses the mean clustering technique. We chose the HMM method to progressively identify tumors and CSF fluid in MRI images, particularly CSF fluid. Iterative Conditional Maximizing Mode (ICMM) is used to combine probabilities and determine the ranges of those probabilities by calculating a series of distances. In this study, GSVMs and HMM classifiers are employed to improve the image quality. K-mean uses the Euclidean distance to measure the closeness, which assigns objects to the nearest centroids to find the hidden information in the MRI image datasets.

The k-mean cluster stores the MRI image datasets, finds the object class at the close point of the centroid, and restructures the images after detection of the object class. This step is repeated until to find all close points of object classes. This shows the process of HMM in the images to detect a tumor or CSF fluid's isolated position in images with numeric values that identify the presence and absence of both low-grade tumor and fluid. This technique performs the main contributions of this research section and is developed with consideration of decision-making rules. By utilizing the two transition states S. This step is repeated until all close points of object classes have been identified. (S1, S2), this method offers a better way to determine if a tumor is there or not. S1 indicates the existence of a tumor or fluid, while S2 indicates its absence. One of these two states is chosen as the starting point for the technique, which then moves smoothly from one state to the next

and outputs the probabilities of sequencing. When joining the probabilities, indicate the size, shape, and range of CSF due to the use of ICMM, which is applied to the trained MRI datasets for the display of the length of tumor (size) or range of CSF inside the image. This method is applied through the transition states, which show the likelihood at state "t" with the strategy of ICMM (depending on the state, either maximum or minimum values). The ICMM created the database to display the enhanced MRI scans after connecting the probabilities to the transition matrix variable and displaying the precise position of the tumor and CSF in the improved MRI images. The fourth step, which is based on the k-NN algorithm, is one of the most effective techniques for classifying CSF and low-grade tumors in MRI images. The k-NN technique, which was discussed in more detail in the preceding article, was employed in this study to identify the location of the tumor and CSF that is closest to the researcher. Through the use of a transition state of probabilities, ICMM is combined with HMM to detect the border and range of tumors. Due to the use of the k-NN method, these probabilities identify the location of the tumor and CSF boundary within the images and display their presence or absence. The k-NN algorithm can easily pick the nearest location from the produced results. These results generate more effective MRI image details, causing the transition from the voxel's current class to another.

The selection of the k-NN algorithm can greatly affect image segmentation, which stores all the data in the form of the k-mean cluster algorithm. When compared to other ways, the suggested technique for validating the hybrid k-NN algorithm strives to achieve all the attributes. The improved results are only achieved by adopting the improved hybrid technique of GCHMkC to identify the features of the low-grade tumor and CSF fluid in MRI images using the ICMM technique in the hybrid k-NN framework. It gives a better opportunity to k-NN algorithms that can easily pick the values of the nearest neighbor due to this proposed hybrid technique. The proposed Hybrid approaches, such as combining GrabCut, hidden Markov models, and k-means clustering, can leverage the strengths of different methods to improve robustness. By integrating complementary techniques, the framework becomes more versatile in handling diverse data such as different size of images with different location and path of tumors and CSF.

The combination of CMM and HMM allows for the visualization of both CSF leaks and tumors, including creating a defined border around them. ICMM utilizes these transition states to process and diagnose findings related to both conditions. An iterative approach using HMM is used to search for symptoms until all close points are identified within object classes. This method provides an improved way of determining if there is a tumor or fluid present in an image by using either S1 (presence) or S2 (absence) as starting points while smoothly transitioning between states, outputting probabilities along the way. By joining these probabilities together with ICMM's use on trained MRI datasets, information regarding size, shape and range can be displayed more accurately for any existing CSF leak phenomenon or tumor length via imaging techniques like MMCs which show likelihood at state 's'. The database created from this technique displays precise positions aiding better diagnosis decisions based upon probability outputs integrated with nothing but sharpness in each individual matter brought about because they have been statistically accounted solely through their combinations made visible by incorporation into variable matrices entailed during various stages such as smoothing transitions among others leading up high-quality enhanced images that demonstrate exactly where

appropriate measures must be applied promptly without further delay when needed so it doesn't result unfavorable outcomes leaving patients confused wondering what went wrong instead aiming best possible results reduction future cases similar kinds arise avoiding mistakes succeeding every time achieved satisfying patient expectations,

 Thereby building trustworthy relationships moving forward towards brighter healthier futures altogether efficiently effectively getting our goals met given challenging modern-day healthcare environments demanding utmost accuracy precision speed being able respond dynamically real-time inputs essential survival would depend rapidly changing scenarios plays big role providing proper due diligence care necessary always wanting informing healing treatment possibilities become reality everlasting happiness unparalleled satisfaction forefront deliverables empowering taking responsibility accountability assumed significant weightage well-being life not just endpoint checkmark list accomplished objectives rather lifelong commitment continuous improvement delivering excellence consistency transforming lives globally making us stand out amongst rest committed keeping always progressing inching closer towards perfection embracing values ethics honesty towards life taking one step at a time along the way consistently making progress journey exceeding expectations striving achieving excellence providing highest possible standard care to all patients in need worldwide. Experimental results are implemented by real MRI images of datasets that demonstrate the proposed technique.

Input and Output of Hybrid GCHMkC Technique

The input of the proposed techniques is based on four steps for detecting and classifying MRI datasets. Step 1 shows the combination of GrabCut segmentation and SVM classification (GCSVM) for the segmenting process. Step 2 shows the scale-invariant feature transform (SIFT) function for scaling, mapping, and extraction. Step 3 demonstrates the use of the HMM and k-mean method (HMkC) in combination to find low-grade tumors and CSF. The k-NN method is used for classification in step 4.

The proposed technique works efficiently and achieves results using these steps. Step 1: Create the max-margin distance class using the hybrid GCSVM technique after training and testing the images. Step 2 extracts the features of MRI images by scaling and mapping the distance between the low-grade tumor and CSF, identifying the pixel size, and extracting distinguishing invariant features from the images. Step 3: HMM is used to identify and detect low-grade tumors and CSF regions in images, and the k-mean cluster is used to store data for the classification of trained MRI datasets. In step 4, the k-NN approach is used to classify CSF and low-grade tumors. Enhanced MRI images with low-grade tumors and CSF identified and categorized in these early stages are the eventual result.

Results of Hybrid GCHMkC Technique

In this section, we explain the results of the proposed technique using a combination of four methods to become the hybrid GCHMkC technique for classification. The following sub-sections demonstrate the experimental implementation of GCSVM, SIFT, and HMkC with the k-NN algorithm for classification. As shown in Figure 2 (a-d), MRI images with white matter, gray matter, tumors, and CSF fluid are created using SVM and k-mean algorithms. The Gaussian radial basis kernel function is used for MRI dataset values in CSF with low-

grade tumors. Figures 2 (d) to 5 (d) may display the histogram of these experimental datasets of CSF with low-grade tumors, which illustrate the signs of CSF fluid leaking in the brain and low-grade tumors. These stages' characteristics depend on (i) classifying MRI images. (ii) For brain tumor or CSF MRI images, the k-mean algorithm is used to store the image data for classification. Four categories—white matter, grey matter, CSF matter, and background—are used as an initial class label. (iii) Extracting the features of samples, which are mean, mode, standard deviation, and entropy, by corresponding the six images of features of each sample point and one class label, creates each pixel feature as a training sample to train and test datasets in SVM after normalization processing. (iv) Applying the Gaussian radial basis kernel function to get SVM classification results by Step 3, and (v) classifying the test sample to get the SVM classification images to follow by Step 4 and using the Jaccard Similarity Coefficients for the simulated brain MRI images and then develop the histograms. In order to identify the low-grade tumor and CSF in an MRI, the histogram shows the number of sample values for multi-class SVM classification for learned SVM classification parameters. The next simulation method, which will maximize margin distance and minimize the energy function of the graph cut, will be supported by SVM classification of train datasets. This method can increase the effectiveness of k-mean-based classification while using the train datasets as inputs.
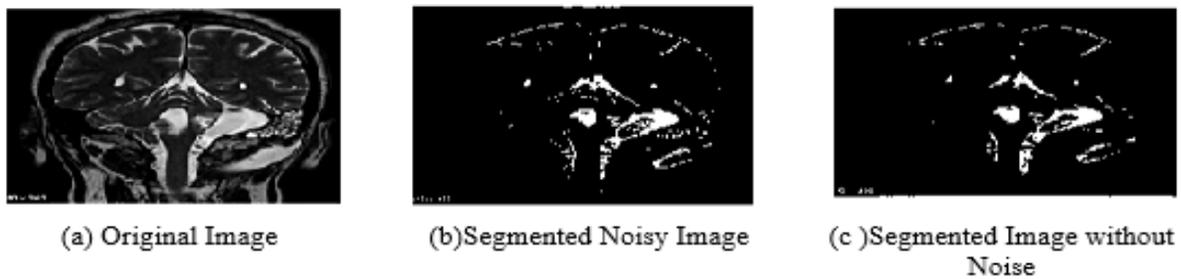

(a) Original Image   (b)Segmented Noisy Image   (c)Segmented Image without Noise

Figure 2 (a)-(c).
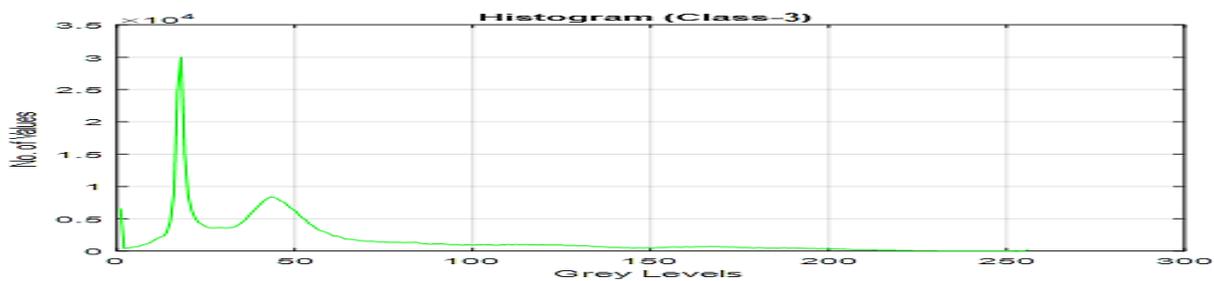**SVM classification for feature extraction of low grade tumor with CSF datasets**

Figure 2 (d).
Histogram of SVM classification for low-grade tumor with CSF datasets

Figure 3 (a) and (b) show the Min-Flow / Max cut algorithm segmented values of MRI histogram greyscale image graph values of both foreground and background for CSF with low-grade tumor datasets. The segmentation is the labeling of the pixel's intensity lies in either the 'foreground' and 'background' pixels in three main groups of graph-cutting algorithms exist namely Initial Seed Points (histogram), Computation of Label Matrix (L) and Directed Node Graph. The "max-flow" function is utilized by this implementation in

Figure 3 (a) and Figure 3 (b) show the Detected result of CSF with Low-Grade Tumor in MRI images.
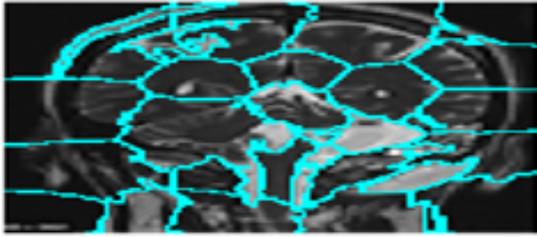

Figure 4 (a). Computational values of labelled matrix generated both source and sink node of Datasets of CSF with Low-Grade Tumor

Figure 4 (b). Detect the CSF with Low-Grade Tumor in MRI images

The results of a node of the Maximum and Minimum Cut algorithm from 1–10 segmented areas of regions in MRI datasets are shown in Table 1. These values represent the regions of tumor and CSF detection target pixels. According to the MRI pictures, the size of the item in the image shows that the value is high in comparison to the targeted values of datasets (CSF with a low-grade tumor) available in the brain.

Table 1.
Results of node segmented area of Maximum Flow Results in image Directed Graph (G) Node Network Diagram

No of Nodes	1	2	3	4	5	6	7	8	9	10
Segmented Area of Region	0	0	0	17	0	29	0	21	01	16

Table 2.
Area of Segmented Region of CSF with Low-Grade

Object	Area (pixels)
1	1307
2	601

Figure 4 shows the graph from network reformulation of MRI image mapping for the given datasets. The blue color of the nodes indicates the value of the observed record datasets CSF with Low-Grade Tumor of each pixel. This figure represents the sink and source node which indicate the nodes of mapping point and connect together for an area of segment region.
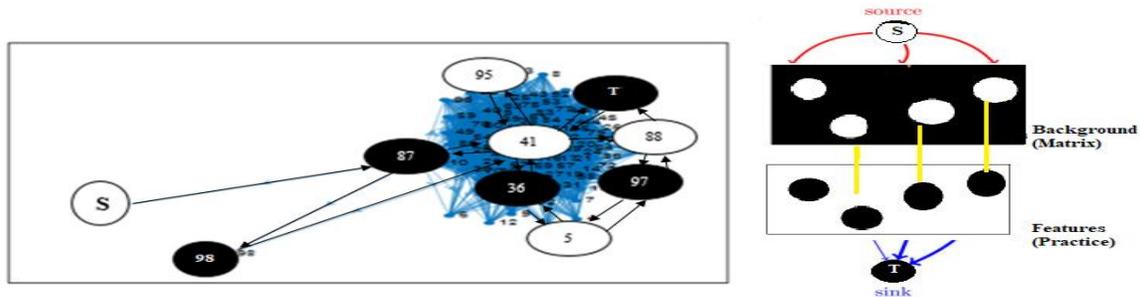

Figure 4.
Graph G from network reformulation of the image mapping of datasets: CSF with low-grade tumor

Table 3.

Edges and nodes values of foreground and background of CSF with Low-Grade Tumor Datasets

Table 3 shows plotted graph of edges and nodes weight of foreground and background that represent the segmentation process and construct the graph by max-min flow method. This graph represents both edge and target node values by using the digraph properties to identify the values of edges (50625×2 mentioned in Table 1) and target nodes (225×0 mentioned in Table 1). The graphical representation of datasets (CSF with low-grade tumor) is mentioned in Figure 5.
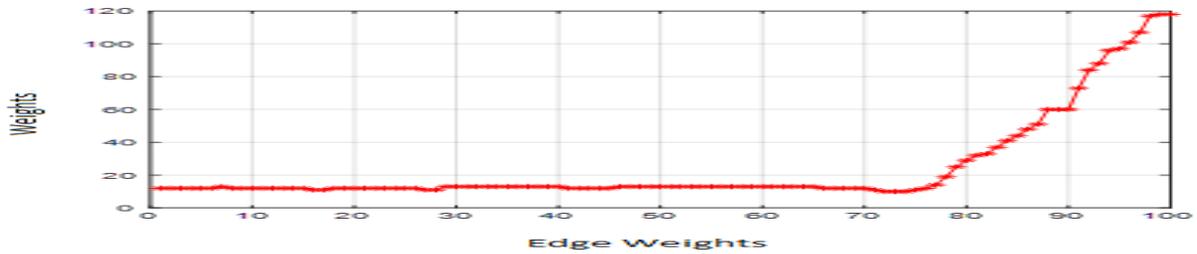

Figure 5.

Endnotes		Weight
Foreground	Background	
1	2	12
1	3	12
1	4	12
1	5	12
1	6	12
1	7	12
1	8	12
1	9	12
1	10	12

Edge Weights of CSF with Low-Grade Tumor dataset

RESULTS AND ANALYSIS OF HYBRID GCSVM + SIFT TECHNIQUES

This section represents the experimental results of scale-invariant feature transform (SIFT), which represent the performance of the convolution function of Gaussian scale variable $G(x, y, \sigma)$ that are repeated scale space initial images that are selected by the proposed MRI datasets as an input image. These experimental results show how Gaussian functions with convolution work to create the scale-space MRI image setup on the left. The difference between the Gaussian images on the right is inferred from the neighboring Gaussian images, and this process is continued until the Gaussian images are down sampled by a factor of 2. The computations in figures 5 and 6 illustrate distance mapping based on pixel marks of maximum and minimum differences from MRI images produced by the Gaussian function, which identified 128 values for the feature and its neighbors in the scale immediately surrounding an area.
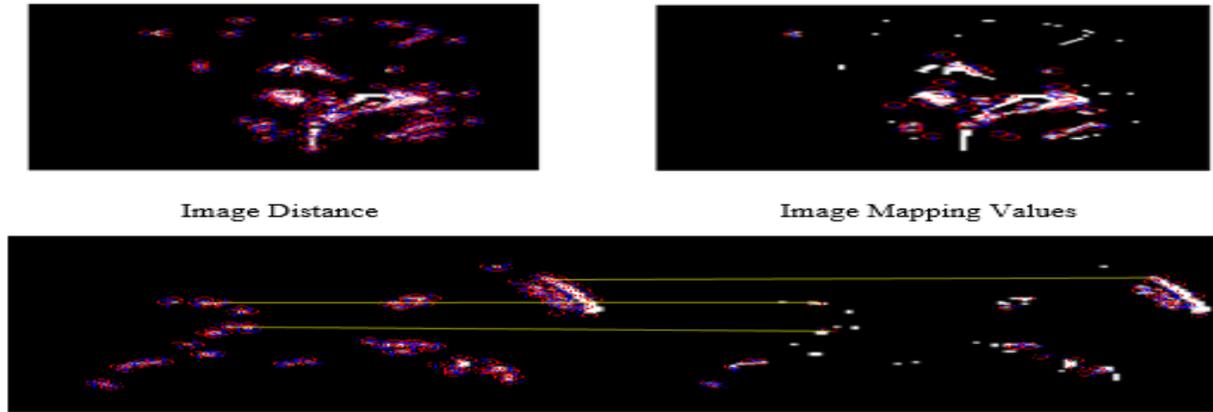
Image Distance   Image Mapping Values

Figure 6.
Low-Grade Tumor with CSF fluid leak in brain images datasets to be identified using the SIFT key-points stored in Features
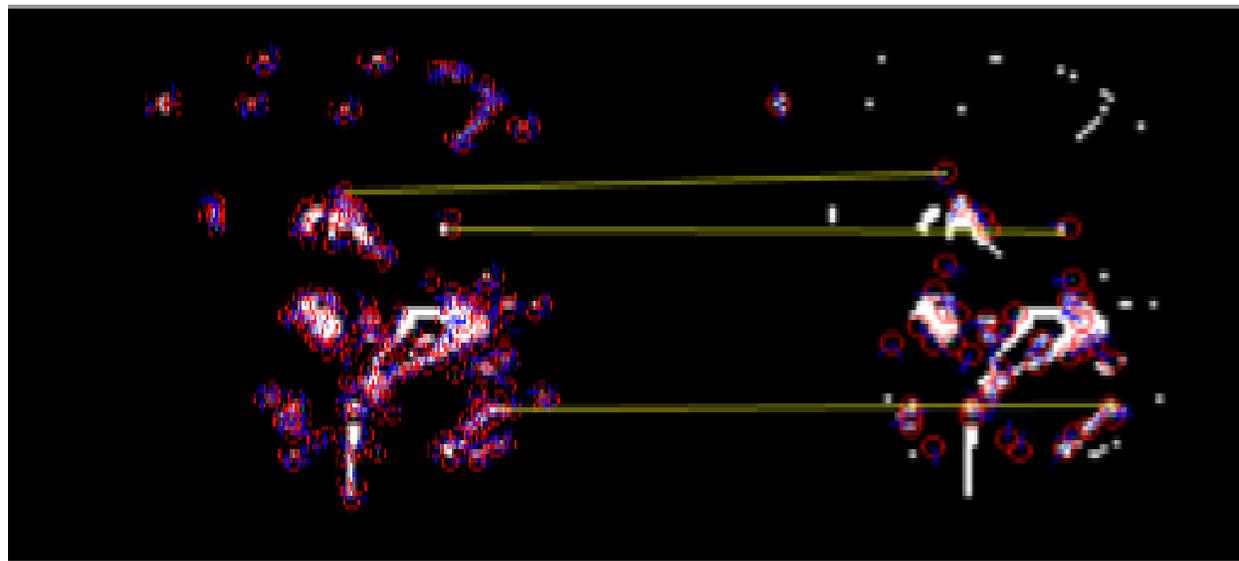

Figure 7.
Object location of a low-grade tumor with a CSF fluid leak mapped by scale invariant mapping using brain image datasets
Table 4.
Image distance and mapping values of CSF with Low-Grade Tumor Images

Datasets Image	Distance	Value
Low-Grade Tumor with CSF Images	Distance	1.2753

Scale-invariant feature transform (SIFT), which is demonstrated in figures 7 and 8, is used to identify and describe the local characteristics of a picture. In the database, a feature descriptor that matches the feature points of the new image with the feature descriptor that is already stored is used for matching and identifying the object of interest. MRI datasets are mapped to Euclidean distances based on their distances and mapping values for determining the best match point.

Results and Analysis of Hybrid GCSVM+SIFT with HMkC Technique

The Hidden Markov Model of k-mean cluster (HMkC) simulation results were used to detect the presence of CSF fluid and low-grade tumors in MRI images. Figure 10 (a) represents the MRI images, which are the combination of GCSVM that indicate the presence of tumor location and range of CSF, whereas the HMM classification is applied for object detection to make the tumor and CSF current location visible with accurate values that show the presence and absence. HMM is applied to track the MRI image segmentation with respect to time and space that can be modelled for the transition matrix of both presence and absence states. Figure 10 (b) indicates the range of fluid at time t, and (t - 1) indicates the probability that is used for classifying the tumor size and CSF leak range at time t, D = 0.993. The HMM calculates both transition states of the posterior distribution that are shown by the filtering process. The first step shows how to calculate the distance, and the second one shows the pixel intensity of both states. Table 5 shows the accuracy of the low-grade tumor with CSF datasets when the pixel intensity is applied for classification at the time (t + 1) and ICMM is used to retrieve the hidden information from the datasets. ICMM allows the quality of an image to be visible after implementing the transition matrix to connect the probabilities. Figure 10 (c) shows the experimental results, which represent the extracted information from the given datasets of MRI, and Table 5 represents the accuracy values of HMM with ICMM of the trained datasets, which show the improved results of the proposed technique.
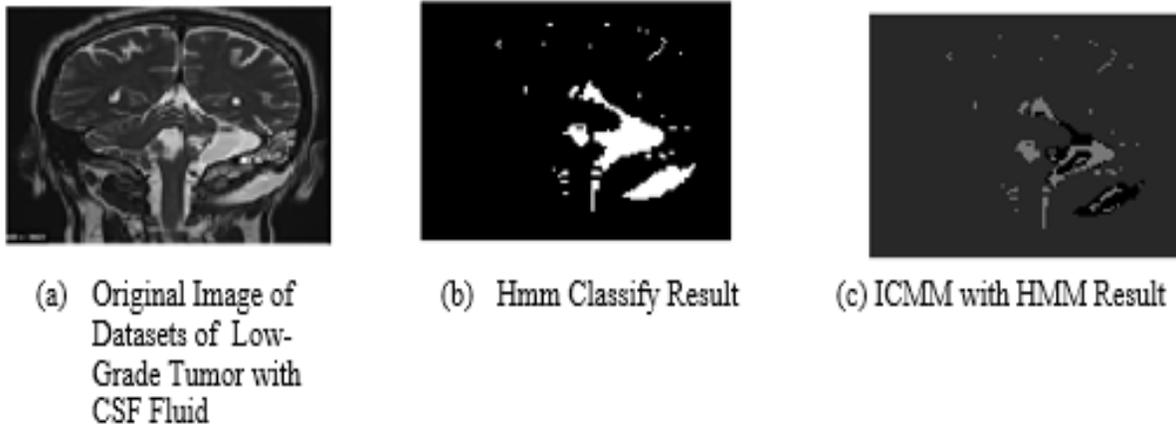

(a) Original Image of Datasets of Low-Grade Tumor with CSF Fluid

(b) Hmm Classify Result

(c) ICMM with HMM Result

Figure 7.
Outcomes of ICMM over HMM Classification
Table 5.
Accuracyof HMM Results over the trained datasets

Image	Tumor/CSF Fluid Detection Values (HMM)	Tumor/CSF Fluid Detection Values(ICMM)
MRI Image	0.9932	0.9415

Results and Analysis of GCHMkC with k-NN Algorithm

This section represents the performance of the hybrid GCHMkC technique in the k-NN algorithm for classification after the training and testing of a sample of MRI datasets to find the nearest location of the k-neighbor. The values are assigned for the test sample

after arranging the optimal values of k during the training datasets with the labelling of selecting the nearest neighbor of every test sample in the k-NN algorithm. These values are utilized for fixed values of k in the suggested technique, which focuses on generating findings for varied values of k with different test samples. The experimental results of the proposed hybrid k-NN algorithm categorization of proposed MRI datasets are shown in Figure 8.
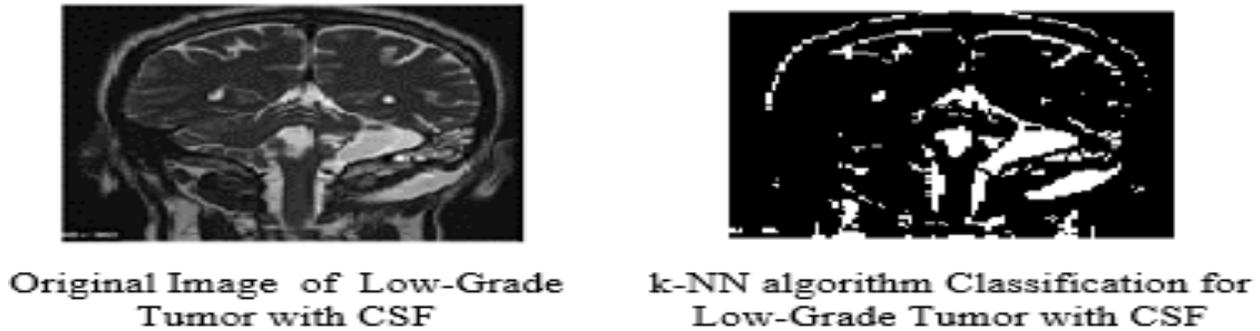

Original Image of Low-Grade Tumor with CSF

k-NN algorithm Classification for Low-Grade Tumor with CSF

Figure 8.
k-NN Algorithm classification results
Table 6.
Results of accuracy, sensitivity, specificity, and computing time for the k-NN algorithm for classifying data

Image	Accuracy K=1-10	Specificity	Sensitivity	Computation Time
Low-Grade Tumor with CSF	0.9983	0.998	0.9997	14.9922 sec.

Figure 8 represents the hybrid GCHMkC technique values used to develop the hybrid model using the combination of these four methods. With an average calculation time of 14.9%, this framework produces enhanced hybrid k-NN method classification results with an accuracy of 99.83%, sensitivity of 99.99%, and specificity of 99.8%. These outcomes create better outcomes with an improvement in image quality and accuracy.

In this research, a hybrid k-NN model was proposed to classify the low-grade tumor and CSF in terms of segmentation and classification. The hybrid model was developed with the use of improved LFD MRI of 3000 patient records for CSF with low-grade tumors. Results indicated that low-grade tumors and CSF were segmented and categorized using MRI images. The upgraded datasets outperformed the baseline datasets in the categorization of low-grade tumors and CSF, reaching a classification accuracy of low-grade tumors with CSF of 99%. Table 7 provides more information on these findings, which were made using a hybrid model with good accuracy. It validates the low-grade tumor and CSF for the classification of tumors and shows where the CSF is located in MRI scans.

Table 7.
Validation Test of hybrid GCHMkC Technique

Datasets	Accuracy	Precision	Specificity	Sensitivity
Datasets of Low-Grade Tumor with CSF Fluid	0.9983	0.977	0.998	0.9997

THE PROPOSED CM-DFT TECHNIQUE

The imputation of missing data is one of the most challenging problems in the field of segmentation. Many methods and techniques have been developed for solving this problem, but still, there is some lack, especially in the medical field, such as MRI images, and in artificial intelligence, such as the k-NN algorithm in machine learning. We proposed the novel technique correlation matrices of discrete Fourier transform **(**CM-DFT) to resolve this problem and reduce time lags (delays) in the proposed datasets. This research is based on finding the missing data in the proposed datasets, replacing the missing values with non-missing values (removing the empty space in the datasets to extract the missing data), and finding the exact location of a low-grade tumor and leakage of CSF in the MRI. After locating the missing values in the suggested training MRI datasets, the proposed strategy is implemented in the previous section, which displays the hybrid model of the k-NN algorithm for classification.

This method yields better outcomes. The hybrid lagged k-NN and DFT correlation matrices used in this section can reduce the chronological delay and retrieve missing data from the same rows and columns. The section illustrates the hybrid k-NN algorithm's missing value recovery time and solution to the missing data problem. This missing data issue is dealt with by the combination of two parameters, such as time-lagged and total numbers of the nearest location value of "k in terms of time lags. These parameters make the strongest correlation for every single variable after using the nearest neighbor values of k and finding the location of tumors and CSF across all lags. This research is based on three steps to solve this missing data problem. To identify those variables that are significantly correlated with time lags, the first step is to implement the time delay, a novel method for reducing the time delay for testing and training vectors. To measure two-time series functions using cross-correlation, this technique uses the time delay in combination with the cross-correlation method. After the correlation matrix has been developed, we will proceed to the next step, which is developing the time-lagged points and the hybrid k-NN algorithm. It is improved over the k-NN algorithm when the CM-DFT technique is combined with lagged hybrid k-NN (LHk-NN).

Multiple lags across testing and training vectors of proposed MRI datasets are selected for CM-DFT by using variable pairs of testing and training vectors. To develop functions in terms of time with few missing values, this technique creates a relationship between two variables, such as time-lag. The other lags (p) are produced separately. The training vectors also construct the same values that are stored separately after imputation and generate the results of existing values with time T after adding a few lags (p). These techniques remove the accumulation errors after adding the lags (p) within the data length (1 to T) using the LHk-NN. The third step represents the performance of DFT on LHk-NN, which is applied to every variable for imputing the missing data in the same rows and columns. The data segment is constructed for the imputation of the non-missing value at the beginning of the data signal and to impute the given missing values, such as Lu1, Lu2,..., LM, in the proposed trained MRI datasets. The nearest point of the sequential values, which are decomposed by the DFT and locate the missing data in the same direction as the sequence of a variable, achieves the solution of the imputed data. Based on the above two modified rules and their extensions, the basic steps of the CM-DFT technique are shown in Figure 9.
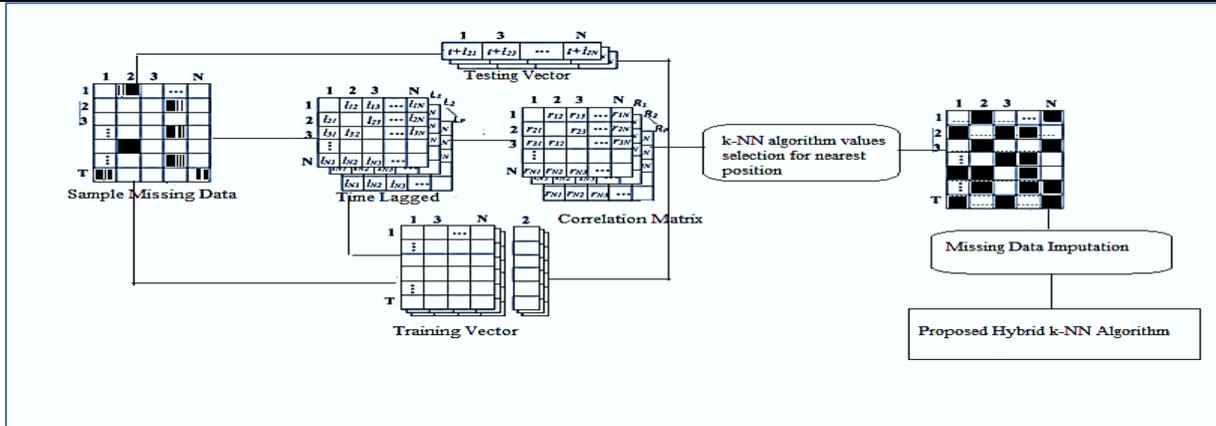
Figure 9.
Block Diagram of Missing Data Imputed Using Discrete Fourier Transformation
**Input and Output of CM-DFT Technique**

The inputs of this proposed technique are three steps for identifying the missing data. Step 1 represents the correlation variable of lagged time delay. Step 2 represents the hybrid k-NN algorithm's (LHk-NN) correlation variable for lagged time delay. In step 3, a DFT technique is used to find missing data in a sequence of signals with varying lengths. This subsection explains in detail how the proposed technique works.

The proposed technique works efficiently and achieves results using these steps. Step 1 provides the correlation variable of the Lk-NN algorithm for identifying the missing imputation data. In Step 2, the new data is rebuilt into a numeric form, or LHk-NN is used to replace zero values with non-zero values. In a similar manner, Step 3 offers the Lk-NN algorithm along with DFT for extracting the missing data from time series data in the same order of rows and columns. After imputing the missing values, the final outcome is successfully attained using a new picture type that has been identified.

Results and Analysis of CM-DFT Technique

The experimental data were used to evaluate the suggested CM-DFT technique's results. This is used to reduce the amount of missing imputation data in MRI datasets. The technique was used to assess MRI images of CSF fluid and low-grade cancers in order to impute missing data. By combining DFT, the CM-DFT technique is employed at the first level to eliminate missing data and reduce temporal lag (delay). After deleting any errors that accrued from the images in order to recover the missing data from the datasets, the experimental findings are based on the extended hybrid k-NN technique with time lags (p) inside the data length (1 to T). DFT is the most efficient method for detecting missing data in a series of variables. As a result, the unique CM-DFT technique suggested here gives a better solution for missing imputation data. This section uses a hybrid k-NN approach, a coefficient cross-correlation variable, and a discrete Fourier transform to show the correlation matrix of time-lagged images and recover missingness values. Figure 9 represents the experimental results of the correlation coefficient of imputation missing data in terms of time series data and time-lagged data with the help of lagged cross-correlation matrices in the hybrid k-NN algorithm. These results are developed in terms of positive and negative values using the calculated lag values in the variable of unit time,

such as 0,-1,1, 2, -2,-100,100,-3,3,4, and 29837625, 29837225, 57825750, 57716900, 2826305000.00000, 84145350, and 109525700.
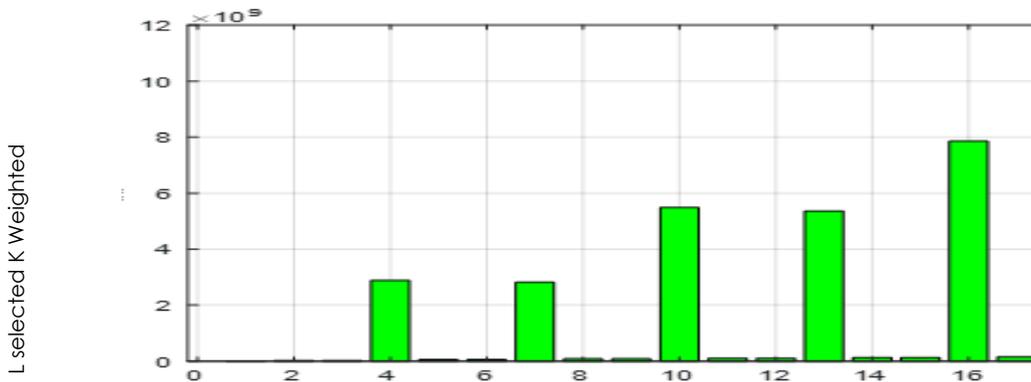
k-NN =1-10

Figure 9.
Time lagged calculated values in hybrid k-NN algorithm
Results and Analysis of Lagged Hybrid k-NN Algorithm

 In this section, we covered a suggested strategy for imputing missing values from MRI pictures into time series data that had a variety of missing values and lagged correlations. Use the cross-correlation method to construct the hybrid k-NN algorithm extension with time-lag correlations in this part. The correlations have the ability to preserve the data for a long time after finding the missing imputation data and calculating the time measurement. The experimental results generate the results of lagged time lags (p) using the correlation matrices and minimise the delays in the MRI datasets. The experimental results show the behaviour of holding variables of every pair of lags using the graphs in the form of testing and training the vectors, which are developed by the correlations of the strongest variable with the size of the matrices. It is also the weakest variable in the process of reconstructing the missing data in the hybrid k-NN algorithm. It also represents the generated values of each pair of variables, which are determined by the number of time lags in the proposed datasets.

Results and Analysis of Missing Imputation Data using Hybrid k-NN Algorithm

Figures 10 and 11 represent the correlated coefficient and average lagged values, which are Figures 9e+06, and 1.0000 and 0.8927 represent the correlation coefficient cross-correlation values of the hybrid k-NN algorithm for the testing and training vectors of MRI datasets at the lagged time point. This shows the average results of lagged time points in terms of hybrid k-NN to represent the strongest correlation by the strength of the correlation coefficient of all lags weighted. Here, the experimental results show the predicted weight distances between neighbours in the given datasets, and the missing values show how the hybrid k-NN method connected temporal lags and correlated variables. In relation to the conversion of testing vectors into single variables in MRI datasets, Table 8 displays the generated values that are substituted for the original values of not a number (NaN). By accounting for the correlations between the variables during imputation and using both the corresponding correlated values, who are taking the

maximum values of k (k = 1–10) to fill the missing values are performing empty space into non-empty space, the cross-correlation of time-lagged into 1.0 second as the single variable can be imputed with multiple types of missingness. The experimental findings show that, in the case of empty values, imputed missing values for MRI datasets of low-grade tumours and CSF fluid are quite accurate. These results showed improved tumour or CSF fluid MRI dataset accuracy, particularly for non-visible or buried data. According to experimental findings, CSF fluid data can be found in MRI images when missing values are imputed from a tumor's early stages when it is not yet a substantial mass. These results have been improved for identifying many consecutively missing values and fully empty instances.
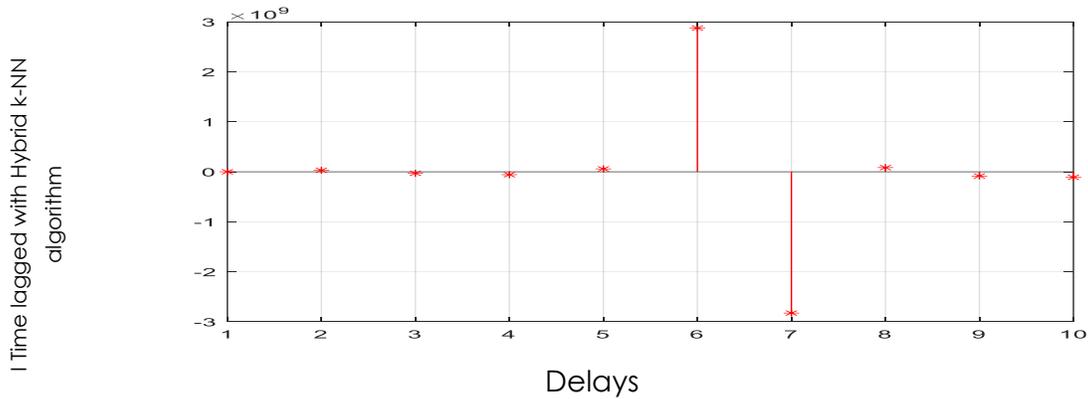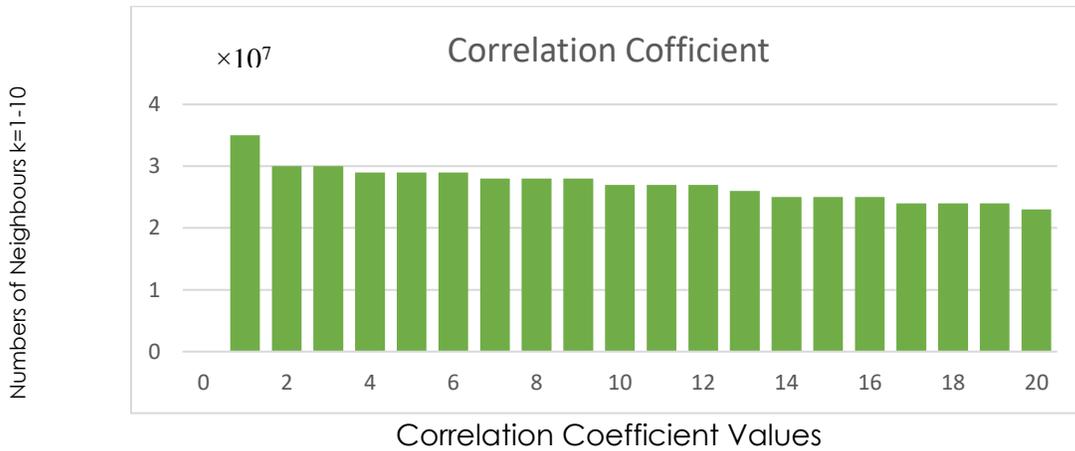

Figure 10.
Missing Data

Figure 11.
Correlation coefficient of hybrid k-NN algorithm using the Lagged Time point at k=10
Table 8.
Generating the Missing Values

Retrieving Missing Values (=NaN)	1	2	3	4	5	6
	X(1100)	X(1200)	X(1400)	X(1600)	X(1800)	X(2000)
	7	8	9	10	11	
	X(2200)	X(2400)	X(2600)	X(2800)	X(3000)	

Results and Analysis of LHk-NN with Discrete Fourier Transform (LHk-NN-DFT)

After imputing the missing values, Figure 12 represents the missing values that are replaced by the empty space values. Figure 18 shows the imputed values, which are retrieved by the missing data in the trained MRI images with the help of the LHk-NN algorithm to impute the missing data points, and the simulated missing data point (in blue), which is imputed after the implementation of DFT. The simulated missing values from the MRI datasets were used to test this. The imputed method was applied for the values of cross-correlation of time lagged and to retrieve the missing data from the rows and columns sequentially using this proposed technique. LHk-NN has missing time instances when Lk-NN has a missing value. With the use of time lags, Lk-NN was able to impute the missing values, and the test vector may not be empty since the lagged values may still exist. The experimental results show that the DFT approach is used to calculate the missing values and replace them with empty space values in a sequential way to retrieve the hidden data in the same rows and columns of trained MRI datasets. Due to the outcomes of the simulation. The DFT creates discrete variables from a single series and unifies all the data from a single sequence into that series. The initial time-domain function is merely DFTed to a different frequency-domain function. Finally, an enhanced response to the imputed data is obtained by selecting the nearest point of the separated sequential values that is identified in the same direction as the sequence of a variable. Tables 8 and 9 provide the accuracy values for missing data as well as the corresponding time lag values, which show the correctness of all three datasets and produce non-missing results.
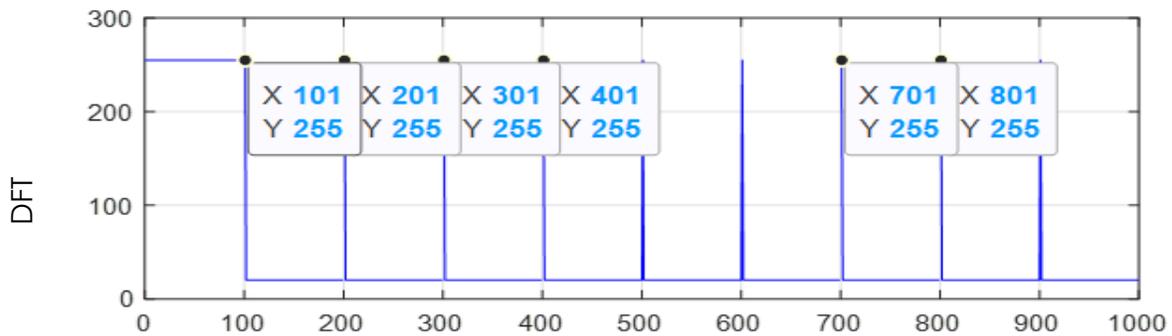

Figure 12.
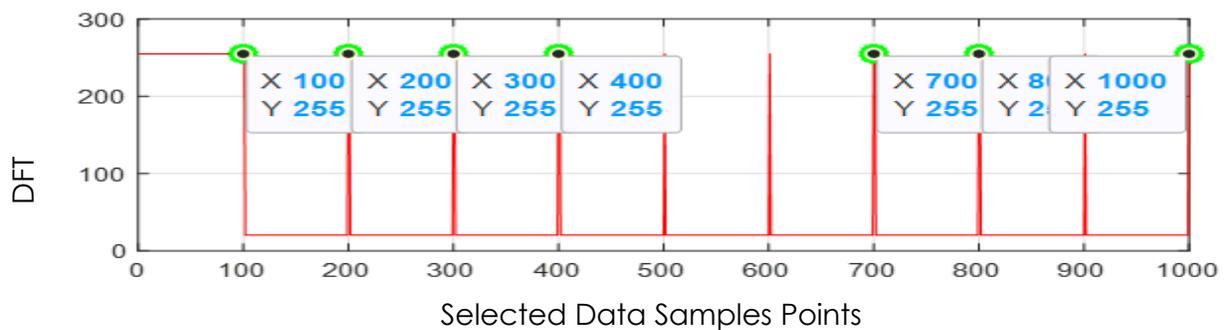**Imputed Missing data**

Figure 13.
Actual Missing data

Table 9 shows how to retrieve missing values hidden in MRI images. This section is used to select the missing ratio target. Table 9 represents the calculated values of LHk-NN-DFT using the proposed technique of CM-DFT to impute the missing values in terms of missing time instances in the presence of a correlation of time lags. This technique handles the missing values for the trained MRI datasets. For missing ratios, the ratios range from 12 percent. The average values of the DFT and LHk-NN algorithms are used to retrieve sequentially missing values with a maximum length.

Table 8.
Accuracy and cross correlation values of CSF with Low-Grade Tumor Image

Accuracy Values	0.9984		R =	
		Cross-Correlation Values		
			1.0000	0.9509
			0.9509	1.0000

Table 9.
Cross correlation of Lk-NN algorithm with DFT

MRI Datasets Images	Missing Ratios	LHk-NN	DFT	Execution Time (Second)
Low-Grade Tumor with CSF Images	12%	99.78%	95.52%	1.5331

The Proposed Lelpp Technique

The k-NN algorithm is widely used for data classification in machine learning. However, its performance may be hindered by irrelevant features and incorrect distance metric selection. To address this issue, the Laplace Transformation of Eigen Mapping approach is presented to tackle disconnected features within a proposed hybrid k-NN algorithm introduced earlier. This technique enhances linear transformation by providing accurate information about vector variables (imputed values) while eliminating nonlinear or insignificant characteristics through the nonlinear gradient conjucate iterated approach (NCGIA) methodology combined with iteration methods. Moreover, utilising Laplacian Eigen maps drives low-dimensional space conversion into high-dimensional spaces, resulting in large margin distance values improvement after proper imputation refinement that facilitates easy image transformations alongside image-based linear changes, reducing storage time requirements as well as unnecessary multi-collinearity, which favours improved performance from refined data identification during reconstruction efforts better than previous techniques applied to incomplete datasets. However much progress has been made using LE-LPP model yet there remains some challenge visualizing nonlinearly converted reconstructed dataset elements whose optimally-applied visualization needs significant attention too particularly given radical form shifts associated when untransformable indices are transformed falsely under such conditions making it necessary to apply our method specifically suited transforming these vacillations enabling application reduce less intrusive distortions over extended calculation periods still producing optimal results compared alternative algorithms despite running more quickly increased efficiency without sacrificing quality standards set industry experts so far achieved best outcomes possible no other tool available anywhere today surpassing fusion state-of-the-art innovations utilized since publication their respective inception dates proving extremely promising choice academic scientists alike wanting handle estimates current scientific community major challenges attend

cutting-edge research practices challenged researchers around world help overcome future obstacles. The Laplacian Eigen maps technique, aided by NCGIA, is utilized to seek out nonlinear feature mappings that facilitate the generation of substantial margin distance values. By applying this method to a nonlinear function and converting its approximate solutions into a linear transformation matrix, the proposed hybrid k-NN algorithm for nearest neighbour can achieve sizeable margin distance values with quick convergence towards exact outcomes. To maintain non-linear features' original data integrity while accomplishing these goals, LPP is implemented alongside this approach. The initial phase involves utilising the LE function to minimise nonlinear feature transformations. This is done by sorting Laplacian Eigen maps and converting them into linear feature transformations. The aforementioned method for dimension reduction implements a hybrid k-NN algorithm from a previous section.

To illustrate this process, consider R1's set of points x1, x2,...Xn alongside their corresponding values y1, y2,... yn in Rn (n = 1). These sets comprise n points, where yi represents xi, and both serve as the original point configurations used by our chosen algorithm. A mapping ensures that similar location-dense data points are grouped together under an objective Liapunov exponent map equation, defining these proximity-weighted clusters with varying degrees of importance proportional to their intensity. Ultimately leading up to minimization, followed by weight allocation assignments based on distance between datapoints, which then consequently undergo exponential remapping further apart.

Utilizing a nonlinear gradient conjugate iterative technique, the third phase is implemented. The nonlinear conjugate gradient algorithm (NCG) is widely recognised as one of the top optimisation approaches for detecting irrelevant or nonlinear data in hybrid k-NN algorithms. To extract such information, NCGIA has been incorporated into this section of the hybrid k-NN algorithm. By exclusively relying on its gradient, NCGIA can identify local minima of non-linear features present within datasets. Since an N-variable function's gradient signifies maximum increase direction, opting to begin searching in reverse (search) direction uncovers any extraneous characteristics or non-linear data lurking within trained MRI images, as illustrated below:

Choose the starting point and calculate the remaining nonlinear data or irrelevant features in the images.

- Find the step length, where the calculation of irrelevant features and nonlinear data, $\Delta x0$
- Perform a line search direction
- Set the new iteration point of: $Xn+1 = Xn + anSn$
- If $Xn+1 \leq \varepsilon$ holds, the algorithm stops. Otherwise, go to step vi
- Let $X = Xn+1$ and go to Step iv

Input and Output of LELPP Techniques

The inputs of these proposed techniques are three steps to improve the classification of the hybrid k-NN algorithm. Step 1 represents the LE working to minimise the nonlinear data. Step 2: LPP is applied to reconstruct the new linear data. Step 3: NCGIA is applied, removing the irrelevant features. The proposed technique works efficiently and achieves

results using these three steps. Step 1 provides the Laplace Eigen maps for minimising and identifying the nonlinear data and converting it into linear data. Step 2: Save the reconstructed liner data in the original form. Step 3 provides improved results with the help of NCGIA for the check-through iteration method and removes irrelevant features from the MRI datasets. In the end, the improved result is achieved in the form of enhanced MRI images with classified low-grade tumours and CSF regions.

Finding the nonlinear data MRI Datasets Images

In this section, we analyze the computational outcomes of a simulation conducted with statistical significance analysis to evaluate the effectiveness of the proposed LELPP technique in hybrid k-NN algorithms. Our focus is on locating nonlinear data and irrelevant characteristics from prior outcomes. To locate this non-linear data after reconstructing linear values, Laplace Eigen Maps (LEMs) are used, which significantly improve performance while reducing the execution time required by utilising low-rank matrix completion methods. We apply an advanced version known as Locally Linear Embedding Preprocessing (LLEP) that uses locality-sensitive hashing techniques and reduces complexity through random projections along geometric objects like curves or surfaces rather than individual points within datasets themselves. This method enables its efficient calculation for large-sized datasets such as those found in medical imaging systems applications.

Further, we use a "low-rank matrix completion" approach called "hybrid principal component analysis," which involves reconstruction models capable of accounting for missing entries due to modelling error or other externalities during signal acquisition when managing Comprehensive Severity Factor (CSF)-mediated MRI exchanges between higher-grade tumour patients compared against their counterparts without any evidence suggestive of low-grade cancer.

Experimental results were obtained via various pipeline steps, including normalisation followed closely enough by application into nodes connected pair-wise using index notation(i, j), which represented roughly bi-directional edges just like graph theory model structures. Within every step lie elements crucially significant'21st century frontier technology principles'-exampled how nearly all key image aspects could be analysed to predict robustly patient clinical databases. These include interactive threshold functions. Model Estimation across Space-Time Windows Processing simultaneously inputting fusion machine learning classification package staining patterns, taking care not to lose clarity of original copy once applied.

Figure 14 illustrates the calculated results after applying the LPP to the Eigenvector Laplace Eigenmap values. To calculate the distance between the eigenvectors of all three MRI datasets and then use LPP to implement the eigenvalues in the form of PCA. After extracting the calculated values of the first four components of eigenvalues of eigenvectors in terms of the linear form of data, these results are generated in PCA forms. Figures 14 and 15 illustrate the graphical representation of newly constructed linear data values that indicate the selected data points of low-grade tumour datasets and low-grade tumour with CSF dataset values, as well as the Eigen diagonal matrix in terms of Eigen Values and Eigen Vectors with LPP.

-0.0074 + 0....	0.0074 + 0.0...	-0.0074 - 0....	-0.0074 - 0....
-0.0074 + 0....	0.0074 + 0.0...	-0.0074 - 0....	-0.0074 - 0....
-0.0074 + 0....	0.0074 - 0.0...	-0.0074 + 0....	-0.0074 - 0....
-0.0074 + 0....	0.0074 - 0.0...	-0.0074 + 0....	-0.0074 + 0....
-0.0074 + 0....	0.0074 - 0.0...	-0.0074 + 0....	-0.0074 + 0....
-0.0074 + 0....	0.0074 - 0.0...	-0.0074 + 0....	-0.0074 + 0....
-0.0074 + 0....	0.0074 - 0.0...	-0.0074 + 0....	-0.0074 + 0....
-0.0074 + 0....	0.0074 - 0.0...	-0.0074 + 0....	-0.0074 + 0....
-0.0074 + 0....	0.0074 - 0.0...	-0.0074 + 0....	-0.0074 + 0....
-0.0075 + 0....	0.0075 - 0.0...	-0.0075 + 0....	-0.0075 + 0....
-0.0075 + 0....	0.0075 - 0.0...	-0.0075 + 0....	-0.0075 + 0....
-0.0075 + 0....	0.0075 - 0.0...	-0.0075 + 0....	-0.0075 + 0....
-0.0075 + 0....	0.0075 - 0.0...	-0.0075 + 0....	-0.0075 - 0....
-0.0075 + 0....	0.0075 + 0.0...	-0.0075 - 0....	-0.0075 - 0....
-0.0075 - 0....	0.0075 + 0.0...	-0.0075 - 0....	-0.0075 - 0....
-0.0074 - 0....	0.0074 + 0.0...	-0.0074 - 0....	-0.0074 - 0....
-0.0074 - 0....	0.0074 + 0.0...	-0.0074 - 0....	-0.0074 - 0....
-0.0074 - 0....	0.0074 + 0.0...	-0.0074 - 0....	-0.0074 - 0....
-0.0074 + 0....	0.0074 - 0.0...	-0.0074 - 0....	-0.0074 - 0....

Figure 14.
Eigen Diagonal Matrix (D) values with LPP generated for CSF with Low-Grade Tumor images

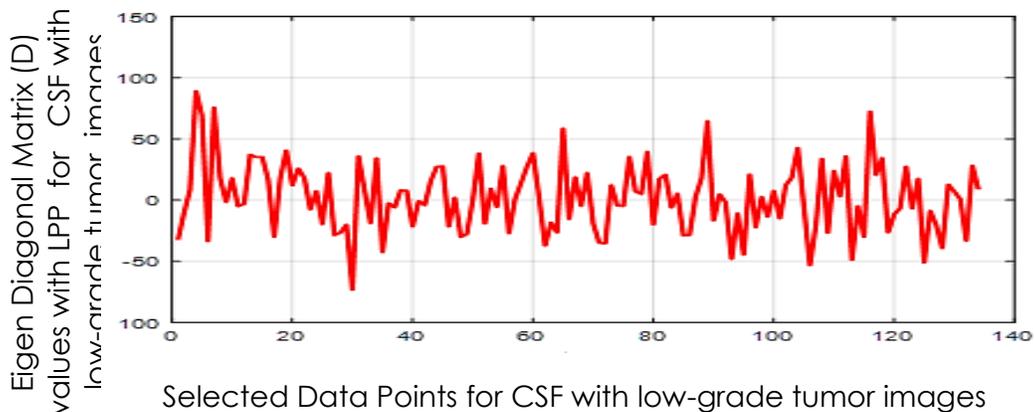
Selected Data Points for CSF with low-grade tumor images

Figure 15.
Graphical Results of Eigen Diagonal Matrix (D) values with LPP selected data point for CSF with Low-Grade Tumor images

Table 10.
Comparison of current running time to previous average running time

MRI Datasets	Affiliation	Hybrid k-NN Algorithm Accuracy	Computational Time (Second)
Brain MRI Images	Toufiq *et al*. (2021)	91.9%	2.5305
Brain MRI Images	Ullah *et al*. (2020)	95.8%	4.103
Collagen Ground Truth (SHG), Elastin Ground Truth (TPEF) Images	Roa, C *et al*. (2021)	90%	3.0116
Proposed Low-grade tumor with CSF Datasets		99%	2.4207

The experimental results showed that the proposed hybrid GCHMkC maintains image quality by using LFD toolkits while enhancing the quality and accuracy of training MRI datasets. The second method, called CM-DFT, was created to fill in the blanks in the training MRI datasets. The process of using the correlation matrix together with the proposed hybrid k-NN framework to target the missing values and transform the empty

space into non-empty space values. In the hybrid k-NN architecture, the third technique, LELPP, was created to eliminate nonlinear data or irrelevant features. The hybrid k-NN framework recommends the utilization of three methods: GCHMkC, CM-DFT, and LELPP to solve the classification problem related to low-grade tumors and CSF. The correlation matrices for each component aid in image segmentation or identification of Cerebrospinal Fluid (CSF) within MRI images. When combined with a hybrid approach, this technique can effectively address the issue at hand.

The GrabCut algorithm is a technique used for segmenting images into distinct foreground and background regions by using graph cuts and Gaussian mixture modeling. To handle the variability present in image data more effectively, an iterative segmentation approach that includes K-Mean clustering within a Hidden Markov Model (HMM) called GCHMkC has been developed. By utilizing spatial relationships between pixels as well as intensity information from each area, precise delineation of CSF tissues can be achieved through iterative refinement that considers both local and global factors.

The Discrete Fourier Transform (DFT) is a transformation that converts an image from its spatial domain to the frequency domain. It reveals information about various frequencies present in the image. By studying correlation matrices of these frequencies, patterns and interrelationships can be uncovered within the image. When combined with CM-DFT techniques, this method becomes even more effective at detecting texture and structural data present in images of different tissues or structures such as CSF, gray matter or white matter found on MRI scans. This helps researchers differentiate between regions by extracting sensitive features particular to their unique variations in spatial frequencies. These boundary delineations are incredibly useful for analyzing visual areas-of-interest precisely.

The goal of LELPP is to preserve neighborhood information in a locality while reducing data dimensionality. Laplacian Eigen maps are used to embed data points into lower-dimensional manifold spaces that maintain local relationships, enabling the detection of subtle variations and boundaries within an image by capturing its intrinsic geometric structure. Moreover, LELPP assists in improving tissue type discrimination since it effectively illustrates spatial links and discrepancies in intensity values among images. The Hybrid k-NN framework can utilize feature extraction techniques such as CM-DFT, LELPP, and GCHMkC. These techniques help represent MRI images with more discriminative features for use in subsequent classifications of pixels or regions within the image by leveraging similarity between their respective feature vectors. The combination of these components creates a segmentation algorithm capable of identifying and defining CSF areas with greater precision and robustness than previously thought possible.

CONCLUSION

To enhance the detection method of low-grade tumours and CSF using trained MRI datasets, a k-NN-based classification framework is being developed in this study. The goal of this study is to examine tumour detection methods and then classify them using the k-NN algorithm. Three novel techniques improve hybrid k-NN model framework performance for detecting tumours and CSF in MRI images while maintaining the same

level of accuracy despite low image quality. MRI is less capable of identifying low-grade tumours and detecting CSF deposits in the brain. The problem was solved by introducing the new trained MRI datasets, which are not publicly available, and also the proposed hybrid GCHMkC technique, which is the combination of four different methods, including SVM, GrabCut segmentation, HMM, and the k-mean clustering algorithm combined with k-NN classification. This hybrid GCHMkC technique has improved the accuracy of the hybrid k-NN framework for low-quality images and anchored improvements in low-quality images. The imputation of missing data is one of the biggest problems in the segmentation of MRI data, especially in the case of MRI datasets that have missing data. Many classifications suffer from imputation application problems where there are a few missing data points. This research focuses on reducing the missing data in MRI datasets and retrieving the missing data through the imputation method. The proposed framework is anchored to increase the efficiency of retrieving missing data with less execution time in the same rows and columns sequentially. The previous technique imputes the missing data in MRI datasets efficiently, but we need to check the remaining data in the framework, which is still in nonlinear form, or remove the irrelevant features. However, the data were retrieved and converted into non-missing data, but only a few of the data were in nonlinear form, which is needed to be linear data. The proposed technique is anchored on creating a precise and efficient hybrid k-NN framework to achieve a higher classification model.

LIMITATION

While several research opportunities remain, further investigation can be done to examine the results of the proposed hybrid k-NN framework in greater detail. Additionally, new approaches and techniques should be considered to find missing imputation data in medical datasets. The primary focus of this study was identifying low-grade tumors and cystic spaces within trained MRI datasets while also imputing their corresponding missing values. The implementation of a hybrid k-NN framework will allow for an open platform where individuals such as general practitioners or consultants can identify potential glucose or protein leakages before tumor/cancer onset occurs. Expertise is required when attempting to locate precise locations where CSF leakage initiates until substantial fluid accumulation has occurred on one side of brain – neurosurgeons/specialized professionals being particularly instrumental here. Manual adjustments may need future considerations specifically increasing resolution via image pixel density from training sets that could potentially track specific leak origination points/low grade tumor development phases more accurately beyond initial detection assisting with multiple type disease identification improving analysis significantly over longer periods through constant efforts towards improvement-related applications derived by AI generated feedback loops helping update/maintain integrity at consistent levels throughout timeframes suitable towards evolving channels served benefiting advantageous overall processes promptly correcting known faults therein contributing valuably supporting life enhancing treatment methodologies ultimately impacting greatly In conclusion, whilst focusing mainly on MRI imaging-based diagnoses detecting issues related primarily around developing better diagnostic methods targeting low level arising malignancies– broader frameworks examining other diseases are likely potentials motivating projects like current multi-directional assignments discussing relevant significant developments subsequently identified showing considerable effectiveness across stochastic relational

variables bestowing valuable contributions elevating foundational bedrock projections informed directly implications associated tertiary derivations inspired visionary experts integrating pivotal uncovering through means available materializing effective system dynamics beneficent logistic improvements experienced well optimally comprehensive domains forma truly synergistic interlocking emergent phenomenon offering variegated solutions brilliantly serving all interested parties pursuing divergent goals similarly.

DECLARATIONS

REFERENCES

Abubakar, A. B., Kumam, P., Ibrahim, A. H., Chaipunya, P., & Rano, S. A. (2022). New hybrid three-term spectral-conjugate gradient method for finding solutions of nonlinear monotone operator equations with applications. *Mathematics and Computers in Simulation*, *201*, 670-683.

Aji, S., Kumam, P., Awwal, A. M., Yahaya, M. M., & Sitthithakerngkiet, K. (2021). An efficient DY-type spectral conjugate gradient method for system of nonlinear monotone equations with application in signal recovery. *Aims Math*, *6*(8), 8078-8106.

Alhawarat, A., Alhamzi, G., Masmali, I., & Salleh, Z. (2021). A descent four-term conjugate gradient method with global convergence properties for large-scale unconstrained optimisation problems. *Mathematical Problems in Engineering*, *2021*, 1-14.

Angra, S., & Ahuja, S. (2017, March). Machine learning and its applications: A review. In *2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC)* (pp. 57-60). IEEE.

Chakravorty, P. (2018). What is a signal?[lecture notes]. *IEEE Signal Processing Magazine*, *35*(5), 175-177.

Chen, Y., Gao, Y., Zhu, L., Li, J., Wang, Y., Hu, J., ... & Xie, Z. (2024). Structure-Enhanced Unsupervised Domain Adaptation for CT Whole-Brain Segmentation. *IEEE Transactions on Radiation and Plasma Medical Sciences*.

Dritsas, E., Trigka, M., Gerolymatos, P., & Sioutas, S. (2018). Trajectory clustering and k-NN for robust privacy preserving spatiotemporal databases. *Algorithms*, *11*(12), 207.

Elnakib, A., Gimel'farb, G., Suri, J. S., & El-Baz, A. (2011). Medical image segmentation: a brief survey. *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies: Volume II*, 1-39.

Elnakib, A., Gimel'farb, G., Suri, J. S., & El-Baz, A. (2011). Medical image segmentation: a brief survey. *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies: Volume II*, 1-39.

Gonzalez, R. C. (2009). *Digital image processing*. Pearson education india.

Khedaskar, S. V., Rokade, M. A., Patil, B. R., & Tatwadarshi, P. N. (2018). A survey of image processing and identification techniques. *VIVA-Tech International Journal for Research and Innovation*, *1*(1), 1-10.

Koorapetse, M., & Kaelo, P. (2020). Self adaptive spectral conjugate gradient method for solving nonlinear monotone equations. *Journal of the egyptian mathematical society*, *28*(1), 4.

Krishnan, P. H., & Ramamoorthy, P. (2012). An efficient modified fuzzy possibilistic c-means algorithm for MRI brain image segmentation. *Image*, *2*(2).

Kumari, N., & Saxena, S. (2018, March). Review of brain tumor segmentation and classification. In *2018 International conference on current trends towards converging technologies (ICCTCT)* (pp. 1-6). IEEE.

Kumari, N., & Saxena, S. (2018, March). Review of brain tumor segmentation and classification. In *2018 International conference on current trends towards converging technologies (ICCTCT)* (pp. 1-6). IEEE.

Kumari, S., & Singh, P. (2023). Data efficient deep learning for medical image analysis: A survey. *arXiv preprint arXiv:2310.06557*.

Li, Z. (2023). Deep-Learning Based Brain Tumor Classification (Doctoral dissertation, State University of New York at Buffalo).

Machhale, K., Nandpuru, H. B., Kapur, V., & Kosta, L. (2015, May). MRI brain cancer classification using hybrid classifier (SVM-KNN). In *2015 International Conference on Industrial Instrumentation and Control (ICIC)* (pp. 60-65). IEEE.

Martínez-Más, J., Bueno-Crespo, A., Khazendar, S., Remezal-Solano, M., Martínez-Cendán, J. P., Jassim, S., ... & Timmerman, D. (2019). Evaluation of machine learning methods with Fourier Transform features for classifying ovarian tumors based on ultrasound images. *PLoS One*, *14*(7), e0219388.

Niranjana, G., & Chatterjee, D. (2021). Security and privacy issues in biomedical AI systems and potential solutions. In *Handbook of Artificial Intelligence in Biomedical Engineering* (pp. 289-309). Apple Academic Press.

Pascal, N. E., Pierre, E., & Emmanuel, T. (2014). Hybrid Method Segmentation for Medical Image Based on DWT, FCM and HMRF-EM. *International Journal of Computer and Information Technology*, *3*(2), 275-282.

Prastawa, M., Bullitt, E., Moon, N., Van Leemput, K., & Gerig, G. (2003). Automatic brain tumor segmentation by subject specific modification of atlas priors1. *Academic radiology*, *10*(12), 1341-1348.

Prosser, J. D., Vender, J. R., & Solares, C. A. (2011). Traumatic cerebrospinal fluid leaks. *Otolaryngologic Clinics of North America*, *44*(4), 857-873.

Raja, S. G., & Nirmala, K. (2019). Detection of Brain Tumor Using K-Nearest Neighbor (KNN) Based Classification Model and Self Organizing Map (SOM) Algorithm. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, *8*(8), 787-791.

Rajasekaran, K. A., & Gounder, C. C. (2018). Advanced brain tumour segmentation from mri images. *Basic physical principles and clinical applications, high-resolution neuroimaging*, 83-108.

Roa, C., Du Le, V. N., Mahendroo, M., Saytashev, I., and, Ramella-Roman, J. C. (2021). Auto-detection of cervical collagen and elastin in Mueller matrix polarimetry microscopic images using K-NN and semantic segmentation classification. *Biomedical Optics Express*, *12*(4), 2236-2249.

Rother, C., Kolmogorov, V., & Blake, A. (2004). " GrabCut" interactive foreground extraction using iterated graph cuts. *ACM transactions on graphics (TOG)*, *23*(3), 309-314.

Saeed, S., Haron, H., Jhanjhi, N. Z., Naqvi, M., Alhumyani, H. A., & Masud, M. (2022). Improve correlation matrix of discrete fourier transformation technique for finding the missing values of mri images. *Mathematical Biosciences and Engineering*, *19*(9), 9039-9059.

Sedghi, A., O'Donnell, L. J., Kapur, T., Learned-Miller, E., Mousavi, P., & Wells III, W. M. (2021). Image registration: Maximum likelihood, minimum entropy and deep learning. *Medical image analysis*, *69*, 101939.

Sharma, D., & Kumar, N. (2017). A review on machine learning algorithms, tasks and applications. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 6(10), 2278-1323.

Song, Y., Adobah, B., Qu, J., & Liu, C. (2020). Segmentation of ordinary images and medical images with an adaptive hidden Markov Model and Viterbi Algorithm. *Current Signal Transduction Therapy*, 15(2), 109-123.

Soobia, S., Afnizanfaizal, A., & Jhanjhi, N. Z. (2022). Hybrid graph cut hidden Markov model of k-mean cluster technique. *CMC-Computers, Materials & Continua*, 72(1), 1-15.

Toufiq, D. M., Sagheer, A. M., and Veisi, H. (2021). A Review on Brain Tumor Classification in MRI Images. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(14), 1958-1969.

Tzika, A. A., Astrakas, L., & Zarifi, M. (2011). Pediatric Brain Tumors: Magnetic Resonance Spectroscopic Imaging. *Diagnostic Techniques and Surgical Management of Brain Tumors*, 205.

Ullah, Z., Farooq, M. U., Lee, S. H., & An, D. (2020). A hybrid image enhancement based brain MRI images classification technique. *Medical hypotheses*, 143, 109922.

Wang, B., Yang, J., Peng, H., Ai, J., An, L., Yang, B., ... & Ma, L. (2021). Brain tumor segmentation via multi-modalities interactive feature learning. *Frontiers in Medicine*, 8, 653925.

Wasi, H. A., & Shiker, M. A. (2021, March). Nonlinear conjugate gradient method with modified Armijo condition to solve unconstrained optimization. In *Journal of Physics: conference series* (Vol. 1818, No. 1, p. 012021). IOP Publishing.

Zhang, S., Li, X., Zong, M., Zhu, X., & Cheng, D. (2017). Learning k for knn classification. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 8(3), 1-19.

Zhang, S., Zong, M., Sun, K., Liu, Y., & Cheng, D. (2014). Efficient kNN algorithm based on graph sparse reconstruction. In *Advanced Data Mining and Applications: 10th International Conference, ADMA 2014, Guilin, China, December 19-21, 2014. Proceedings 10* (pp. 356-369). Springer International Publishing.

Zhang, Y., Matuszewski, B. J., Shark, L. K., & Moore, C. J. (2007, July). A novel medical image segmentation method using dynamic programming. In *International Conference on Medical Information Visualisation-BioMedical Visualisation (MediVis 2007)* (pp. 69-74). IEEE.

Zhuge, Y., Krauze, A. V., Ning, H., Cheng, J. Y., Arora, B. C., Camphausen, K., & Miller, R. W. (2017). Brain tumor segmentation using holistically nested neural networks in MRI images. *Medical physics*, 44(10), 5234-5243.