International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.1 (2025) pp. 124-134
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Optimizing 3D Brain Tumor Detection with Hybrid Mean Clustering and Ensemble Classifiers

Vijayadeep GUMMADI[1,2]*, Naga Malleswara Rao NALLAMOTHU[3]

[1]Research Scholar, Department of Computer Science & Engineering, Aacharya Nagarjuna University, Guntur, Andhra Pradesh, India-522510

[2]Lecturer and Incharge, Department of Computer Science, SRR & CVR Govt. Degree College(A), Vijayawada, India

* **Corresponding Author Email:** vijayadeep.gummadi@gmail.com - **ORCID:** 0009-0005-8020-3936

[3]Professor and Head of the Department, Department of CSE-IoT, RVR & JC College of Engineering (A), Guntur, India

Email: nnmrao@rvrjc.ac.in - **ORCID:** 0000-0002-1360-1150

Article Info:

DOI: 10.22399/ijcesen.719
Received : 02 October 2024
Accepted : 05 December 2024

Keywords :

Brain stroke,
feature extraction,
stroke classification,
machine learning models.

Abstract:

Magnetic Resonance Imaging (MRI) scans are extensively used in the medical field for identifying brain cancers. However, a major drawback is that accurate diagnosis of brain tumours can be difficult. This happens due to several reasons such as noise oversegmented regions with a big feature space and high false-positive error rate. This paper presents a novel hybrid algorithm for brain tumour detection. The method exploits state of the art data-preprocessing, feature selection, clustering and classification techniques. The proposed method uses ensemble feature selection to select the most relevant features and ultimately increase the predictive accuracy of the detection system. A hybrid means clustering approach with modulus function clusters the similar data points, filter out noise and identify patterns and anomalies related to brain tumours. The ensemble classifier further explores the accuracy of the proposed scheme and helps in improving the predictive performance of brad tumours. Use of feature selection can highly boost the detection and finally achieve accurate diagnosis and prediction. Eventually, brad tumors can be treated at an earlier stage without being fully developed. The experimental outcomes are promising to the extent that the proposed algorithm can be further incorporated into the treatment of brad tumor.

1. Introduction

According to the recent research statistics, it has been indicated by the increasing number of the neurologists in the United States of America. As per the more popular prediction from the American Brain Tumor Association, more than eighty thousand brain tumour cases are detected in the USA on an annual basis with a great percentage of cases are malignant. That is, the approximate maximum detection of brain tumour cases identified and detected as malignant is 32 percent. The neuro-physicians competent in handling this type of brain tumour detection and analysis by performing a histopathology either on the brain or on brain tumour to remove the brain tumour to further reduce the growth of the brain tumour. The outcome of the brain tumour operation may lead to the many side-effects like the weakness, loss of coordination, confusion, dizziness, speech problem. So, the brain tumour early detection and diagnosis exhibit the increased possibility of the successful medical treatment leading to drop the short-and long-term medical outcomes [1]. Nowadays, the requirement of the interactive computer aided detection and diagnosis of the brain tumour is highly increasing with more population in day to day life. As the increasing human population automatically increases the prone of the patient which is also a dual edge sword in terms of the data acquiring.

The MRI medical data will increase the exponential form and may hit the neuro physicians to produce the more precise medical analysis and diagnosis. Motivating the research work toward to devise the brain tumour segmentation and classification to increase disease severity and diagnostic ability of the neuro physicians to increases the medical accuracy utilizing the machine learning based computer aided diagnosis technique.

The brain tumour segmentation is the recent research area in medical imaging which can segment the tumour into many types AD, ET, malignant and benign tumour to extract the relevant biological information like the region of interest. The existing research study implemented the various segmentation approaches like threshold based, region-or edge based, contour based, atlas based approach. After implementing the feature extraction, the classification based segmentation approaches for skin tumour such as SVM(support vector machine), k-NN (k nearest neighbour), random forest (RF), Naive bayes (NB), decision tree (DT) classifier were done to classify the region whether it is cancer/malignant tumour or non-cancer/benign tumour.

New medical imaging modality and medical image processing techniques is developing day by day which makes medical and health care industries to be an evergreen industries. Because no of patients is increasing tremendously due to hybrid or artificial foods, refrigerated and frozen foods, by more fertilized agriculture. In most of the cases, the incoming number of patients is more than the houshy surgoon (medical expert) per day and in daily routine no of expert (radologists) is always lower than the incoming number of patients. So, the medical imaging modality is required for the purpose of detection and classification of the tumor with its severity. None of the medical imagining can recoganise the tumor manually and it is weaken to get more collapses in the clasification system if any mistakes done overlly oncoming no of patients. Especially, if the tumor is classified manually then it requires more time, it will cost more and it is also weak in accuracy.In medicine, computer assisted diagnosis and image anoylysis have an important role in making a detected disorder to be diagnosed precisely. Medical Imaging plays an important role in the medical field. Since the early 1980s, there are certain techniques by which the information about inside of the human body can be retrieved without cutting or injuring the body. Medical-image analysis uses computer algorithms to transform medical images (or more generally, images in a medical context) into meaningful information for clinicians and scientists. It employs images as a new window for rapidly exploring core biological processes and disorders. In early days, some of the reseachers used conventional methods for segmenting brain tumors and validating segmented part manually. Some of the researches have involved with segmenting and classifying brain tumors, but all of them took more time, high cost and accuracy is also less. There were some of the earlier methods which have involved with features extraction and classification. Depending on the tumors, the extracted texture, shape, and some statistical features cannot present full information about the tumor.

Medical Image Processing is an interdisciplinary field for processing Human 3D Image Datasets obtained from Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). These models are obtained by Radiologists and clinicians to diagnose the human disease. The MRI images yields a clear interior structure with no injecting dyes. So, it is widely used to detect the brain anatomy. However it is., but also other human body parts like. Moreover, there are other medical imaging modality such as xray, CT, ultrasonic, . These modalities aids the doctors in diagonizing several other body parts like brain, breast, bone marrow, lungs etc via MRI. The image is converted to energies, then to signals by the fast processors. The signals generated indicates various tissues present in the body. The computer can perform this task in two way. The computer can have the capabilities in Medical Image Processing. There are different techniques in which bunch of images are stored, processed and transmitted.

A set of digital image processing approaches used to edit or diagnose medical images is called Medical Image Processing (MIP), which begins with read medical images from its input generating digital images. Hence the process of MIP is call Digital Image Processing as well for the process of medical images. After that the preprocess from images, e.g., image registration, denoising, color conversion, resizing and more. After preprocessing stage, the segmentation from images, which learn internal and exterior determinant from images feature identifying and feature classification at the same time.There are three types of cell in the produced the tumor . Those are Benign, Pre-malignant, and Malignant . Benign are Non-Cancer cells. Pre-malignant is actually the inappropriate cells. And Malignant are cancerous cells . Brain tumor are classified into two categories, namely primary, and secondary . Primary brain tumor they appear on the brain or around the tissues adjacent to the brain . For primary brain tumors , it makes it differnt types . Some types include Glioonmas or Astrocytomas , Meningioma , Acoustic neuromas also name as schwannomas , Pituitary adenomas , Medulloblastomas , Germ cell tumors, and Craniopharyngiomas . The acoustic neuroma were a schwannomas . The secondary brain tumour which it start at any parts in body and it spread to the brain . It are common affected to adults . The examples of common cancer include colorectal cancer , Renal cell carcinoma , Lung cancer , Breast cancer , Uterine , Prostate , Melanoma skin cancer and Leukaemia . The risk of brain tumor include Radiation exposure and the ancestry of studding brain tumor . The brain tumors happened in glial

cells . Gliomas the glia cell grow in the brain which it includes nearly all in all brain tumor . This cell primarily to offer nutrition and they give supports to central nervous system .

MRI gives complete information about the brain tumor by using magnetic fields. A contrast medium back to the veins like a special dye, helps to take clear picture of the brain tumor. MRI scans or Magnetic resonance imaging of various types are better that any other scans to decide the brain tumor. Neurological also takes his/her conclusion about brain tumour. MRI can be done either at brain or a spinal cord.MRI are of 6 types. They are as follows; Intravenous (IV) gadolinium, diffuse- weighted imaging, perfusion imaging, spinal MRI, functional (fMRI), Magnetic resonance spectroscopy (MRS). Intravenous (IV) gadolinium. In that, it enhances the image of brain tumor. That is done at the back of the vein, this particular medication is called as gadolinium, which is a contrast medium. Diffusion-weighted imaging. That shows clearly the cell structure of brain. Perfusion imaging. That shows the amount of blood flowing to the brain tumor. Spinal MRI. It is used to detect the tumor from spine. Functional (fMRI), That produces detailed informations about the brain. This kind of fMRI has some tasks like playing musical notes or moving hands, legs, etc.., all these task done by the patient during the fMRI process makes the clear images while reflecting on fMRI. This is more usefull in identifying the brain portion which affects during operation to remove the brain tumor.

By means of the chemical compositions extracted from brain.Patient's brain using MRS (Magnetic resonance spectroscopy). here information is generated how new and dead cells placed in the brain tumor.It also detect the cancer type of the tumors in human brain.On the other hand using MRI night images which are shift into Gray Scale imagesafter from using FCM clustering algorithm the objective is to use spatial domain using BY (boxy) filter for image enhancement. After that discarding the component using the spatial thresholding and watershedding which been used to isolate the noise and change the image at desired component.Above all the watershed you may get the thresholding value that used to measure the stenotic and non-stenotic lesions. After the morphological operators used to segment the threshold and watedsedding method than the tumour(Brain) detected from involved images The research focuses on finding the area in fMRI brain to accept human physical activity and guide to find how much brain able to work.

Brain tumor cells can be seen as Cancerous and Non-Cancerous cells. Cancerous cells or the malignant cells contains those types of cells, which kills the one who possesses them and spreads it to the environment. Non-Cancerous cells are just benign cells, which is not really causing any types of severe problems. those are called as Malignant Tumor and they were divided by World Health Organization(WHO) by Grades from 1 to 4. (P. Kleihues et al., 1993) They have classification such as

1) Pilocytic Astrocytoma,
2) Low-Grade Astrocytoma,
3) Anaplastic Astrocytoma 4) Glioblastoma The first two tumors called they are semi - malignant ones, and later tumors called they are malignant one and causing the death if does not treat before the certain threshold period.

Initially, segmentation is performed on the images acquired. Segmentation is a fundamental process in image processing algorithms. The segmentation perform the output of Image processing at a higher rate. An Image having appropriately segmented will increase the processing performance of image processing techniques. The brains tumor image should be diagnose at the prior of proceeding to treatment of the brain tumor. It is necessary to treat the tumor region that is most affected. By diagnosing at that area the zones will be heeled instantaneously. Also, diagnoses of the result of the treatment will present prognostic data. Brains tumor image segmentation is done in both automatic and semi-automatic, but the automatically segmentation distributes over a larger area. Meanwhile, the result of the segmentation will decrease in two way either the segmentation took more time for the accurate segmentation or the result of the segmentation have more errors. To decrease this downside the semiautomated image segmentation will produce fewer errors and it will reduce the time in segments. The brains tumor affect the MRI (Magnetic Resnance Image) that it have multiple sequence for segmentation of the tumor image. Like T1 (T1-weighted), T1c (T1-weighted contrast-enhanced), and T2-weighted and FLAIR techniques (Fluid Attenuated Inversion Recovery.

The Discrete Wavelet Transform (DWT) is a widely used technique in image processing, particularly in the area of image denoising. This technique is applied in various fields of image processing, such as texture analysis, object recognition, and image segmentation. Recently, new nonlinear and adaptive filters have been proposed, including the Non-Local Means (NLM) algorithm that uses PCA for improved precision.

The process of removing noise from an image can be divided into two main approaches: image thresholding-based denoising and image segmentation-based denoising. The method chosen depends on the characteristics extracted from the image and the segmented regions. Figure 1

illustrates the basic steps of image reconstruction, which involve applying image filtering techniques to remove noise and enhance the brightness of high-resolution images. The thresholding-based approach involves selecting a threshold value that separates the image into two categories: desired and undesired values.

2. Literature Survey

Brain tissue that proliferates and seems undisturbed by normal cellular processes is known as Brain Tumor in medicine. A tumor is a mass of abnormal cells in the brain. The skull serves as a barrier to keep the brain secure. Any expansion in such a small area might be an issue. Brain tumors may be both cancerous and noncancerous (benign). Whether benign or malignant, a tumor may cause a buildup of pressure in the skull. Severe brain tumors create brain damage and sudden deaths. There are two types of brain tumors: primary and secondary. A primary brain tumor occurs when a tumor develops in the brain. In many cases, early brain tumors are harmless. Metastatic brain tumors, or secondary brain tumors, originate when cancer cells move from another organ, such as the lungs or breasts, to the brain, resulting in tumor growth. Caucasians are more likely than other races to have a brain tumor. There are various reasons which cause brain tumors; Age: As with any cancer, brain tumors are more common in youngsters and the elderly, but they may occur at any age. For example, 60% of brain tumor patients are aged between 30 to 50yrs old. Gender: Brain tumors are seen in males three times more than in females. Among brain tumors, meningioma is more common in women than in men. From the overall comparison, the number of brain tumor patients are Male only. Occupational And Household Exposures: The chance of getting a brain tumor may be increased by soil, herbicides, rubber, and vinyl chloride. Even though these assertions may be accurate, there is no scientific evidence to support them. Family History: 5% to 10% of all brain tumors are believed to be caused by inherited conditions such tuberous sclerosis, von Hippel-Lindau disease, neurofibromatosis, and Turcot syndrome. Brain tumors have been observed in families with no known genetic link. Researchers are studying these clusters to discover what is causing them.A primary brain tumor originates in the brain and then spreads to other parts of the body. In both benign and malignant primary tumors, glial cells are the cell type of genesis. Brain tumors that originate elsewhere in the body and travel to the brain through the circulatory system are metastatic. A tumor must have spread to other parts of the patient's body in order to be considered malignant. About one-quarter of all cancer patients are affected by brain metastases, impacting about 150,000 individuals each year. There is a 40 percent probability that lung cancer patients with metastatic brain tumors would acquire the condition. For patients diagnosed with these tumors in the past, the typical survival time was only a few weeks. Improved diagnoses, better post-diagnosis outcomes, and surgical and radiation techniques have led to a large increase in survival rates in recent years. One of the famous and effective treatments for brain tumors is proton beam therapy. It is a worldwide benchmark treatment applied to treat pediatric brain tumors. Brain tumors have been linked to severe head trauma for many years. Many studies have shown a relationship between head trauma and meningioma; however, glioma is not. As a result, it is not clear whether seizures raise a person's risk of brain tumors, if seizures themselves are the cause of tumors, or if anti-seizure medication is a contributing factor. Types of Benign Brain Tumors: A brain tumor is classified into various types and is explained here. Chordomas: People in the age group of 50 and 60 are more likely to develop one of these benign tumors. Most people have them around the base of their skulls and near their spines. Even though these tumors are benign, they may pierce neighboring bone and exert pressure on the brain's underlying tissue. Fewer than 0.2 percent of all brain tumors are primary brain tumors. Craniopharyngiomas: Craniopharyngiomas are difficult to remove because of their proximity to crucial brain locations, even when benign. Usually, the growth of these tumors is linked to the pituitary gland, which controls a wide variety of hormones in the body. Gangliocytomas: These uncommon tumors, composed of well differentiated neoplastic nerve cells, may occur in young people. Magnetic resonance imaging (MRI) is another therapy that uses a magnetic field and not x-rays, so we get a very descriptive picture of the body by using this technique.we can get a MRI to evaluate how big is the tumor. Before having the CT scan, the patient needs to receive an injection of a contrast medium.A MRI and CT scan might be useful to diagnose a brain tumor.If the tumor is the one that has a tendency to spread to another parts of the brain or spinal cord, it is necessary to realize a MRI of the brain and of the spinal cord. There are many different MRI machines available. Evaluation made by an internist or neurologist will contribute to decide which type of MRI is the better for the patient.The "IV-gadolinium-enhanced MRI" is a very useful tool to better examine brain tumour, specially when we want to evaluate how serious is this tumour. Gadolinium contrast media is administered through a vein of the hand, after having done a standard MRI, and a second scan will be taken after having

administered the dye. It's so important to see clearly the tumour with this technique, that every care is needed for it. Monte Carlo simulations of 1 mT fields in the human brain and eye A special MRI approach called "diffusionweighted imaging" allows to observe more of the cells of the brain.A "perfusion imaging technique" is performed to find out how many blood arteries feed the tumour[2-4].

Perhaps they will help doctors estimate how well the therapy will function. A tumour in or near the spine can be identified with a spinal MRI Scans called functional magnetic resonance imaging (fMRI) can detect regions in the brain that are typically active in controlling muscular activity or speech. Any brain activity associated with acting, showing or receiving pain, speaking (or singing) and the like can be detected during the scan. Is there any part of the brain affected by the tumour that is crucial for the person's continued ability to function? This test is done before the tumours are removed . Diagnoses usually come after the biopsy of the tumour cell. There are several ways to diagnose a brain tumor, but most physicians use a biopsy. The fMRI images are samples brought back to the lab where pathologists interprest them. A pathologist is a doctor who interprets laboratory data through inspection of cells, tissues and organs. After the tumor is removed from the brain, a biopsy could be done on the cells. If the tumor was not able to be removed either because of the location of the tumor or the batter health of the patient then surgery could be done as a seperate treatment. Biopsy. A CAT scan. Contrary to popular belief, CAT scans are NOT realizing x-rays to see visble into the body. CT scans get multiple x-ray images at different angles then the computer puts them together to make 3D scans of patient. The images are useful for spotting tumors and other abnormalities[3-5].

One of the types of fMRI imaging is the PET (Positron Emission Tomography) that uses a substance that emits radiation known as a 'radiotracer' to measure fluctuations in the metabolic processes and depict them. It can also help obtain physical activity from blood flow, attraction and chemical structure. If a patient has a tumour and gets treatment for it to remove it, the PET scan can help to understand the tumour. If a tumour forms after the treatment, the tumour can be taken out. One of the most popular PET scans is the PET-CT scan, which combines a PET scan and a CT scan. When a doctor asks for a PET scan, they mean this technique. There are a number of chemicals, such as sugars and proteins, by which images of organs and tissues from the inside are obtained. There is a radioactive material given to patients injected into the body. The distribution of the material is more in growing cells. While the radiotracer substance goes in through cell division, it goes deep into the body of the cell. It is more likely that a tumorous cell is in proliferation, so it takes more radioactive matters in than healthy tissue. However, it is unlikely the amount of radiation for the body will be a problem. When scanning the inside of the body, pictures of what is happening inside. Sometimes, it is known as a cerebral angiography or an arteriogram of the cerebral arteries. A cerebral arteriogram is an x-ray or a sequence of x-rays of the artery systems in the brain. It is accomplished by injecting a type of dye, called a contrast medium, in the patient's major arteries in the skulls. In order to seek for tumor cells, blood or tumor markers in the patient's body, a lumbar puncture is needed. It is to inject an intravenous needle into the patient's lumbar region located at the bottom of the backbone. The procedure is called a lumbar puncture. Tumour patients may have an unusual high level of a biomarker in the patient's blood or other body fluids. Before the treatment, a local anesthetic is injected into the center of the backbone [6].

Myelogram: A doctor may order a myelogram to determine whether the tumour has spread to the spinal fluid, brain or spine. You can perform a myelogram by injecting the patient's spinal fluid with a tracing dye. Biomarkers: Your patient's doctor may do tumour-specific gene, protein or other variable testing. We typically refer to tumour testing as tissue characterisation in this scenario. Biomarkers can be used to predict the prognosis of your patient. Biomarker research is being developed to detect brain tumours before they cause symptoms, which are also referred to as clinically silent stages. These tests could influence the choices of therapies that are available to your patient. a hybrid clustering method combining k-means and fuzzy c-means algorithms for brain tumor segmentation, leveraging MRI scan data and demonstrating improved accuracy on both synthetic and real-time datasets. A new deep learning method, Dolphin-SCA based Deep CNN, which uses fuzzy deformable fusion segmentation, Dolphin Echolocation based Sine Cosine Algorithm, power LDP and statistical features to improve brain tumor classification accuracy, achieving a maximum accuracy of 0.963 on BRATS and SimBRATS datasets. [7-12]. Some other important works done on this topic and reported [13-19].

3. Proposed Model

The given flowchart illustrates a hybrid algorithm for image tumor detection in the brain area, combined several stages to select robust features, and group and classify them.

Figure 1. *The basic steps of image reconstruction*

Below its an explanation for each stage in the context of brain tumor detection:

Step-by-Step Explanation:

1.Input Datasets:

— The first step taken by the algorithm is importing the brain tumor datasets. These datasets usually consist of medical images (MRI scans etc) and patients' data.

2.Data Filtering Algorithm:

This step involves cleaning up the dataset (ie, preprocessing the data) by removing noise and redundancies, or normalising and augmenting it so that it is improved. This could be done by removing any artefacts.

3.Ensemble Feature Selection:

An ensemble approach is utilised with several different feature selection methods to obtain the most robust and informative features to detect tumours. This ensemble approach allows maximising the robustness and accuracy of the features obtained through a cross-validation, in which separate training datasets are used to minimise the possibility of overfitting and maximise generalisation of the trained model.

4.Hybrid Mean Clustering Algorithm:

It clusters the selected features with the help of a hybrid mean clustering algorithm, a typical aggregation approach that helps demand modellers to group together similar data points and recognise patterns relevant to brain tumours.

5.Hybrid Probabilistic Weighted Measure:

– A hybrid probabilistic weighted measure is used on the grouped data, in order to better refine the clusters by probabilistic (meaning probability) distributions of the data points. Now here it becomes easier to model the underlying data.

6.K-Representative Centers:

· Clusters are represented centrally, by finding the centre of each. So, for example, we can pick one item from each cluster and say: this one is typical of Cluster 1, this one of Cluster 2, etc. These items can be regarded as relevant 'centres' (or 'prototypes', or 'exemplars') for each cluster.

7.Mean K-Clustered Training Data:

– the training data is averaged within clusters to generate training samples that represent those clusters (this reduces the dimensionality of the database and its noise).

8.Data Training Module:

– The K-clustered training data is calculated through the mean from these brain tumour-ridden samples – this mean is fed into a data training module that uses the processed data to train machine learning models that ultimately learn to detect patterns related to brain tumours.

9.Ensemble Classifiers:

Several classifiers are used 'in ensemble' to increase its predictive performance. As the name indicates, the method consists of combining results from several classifiers, leading to higher accuracy and robustness of the detection.

10.Test Data:

With the trained model in hand, we can feed it new, unseen data (the test data) that will help us assess how well it generalises to reality.

11.Test Class Prediction:

The ensemble classifiers classify the test data according to this equation: The predictions have been generated by combining the patterns learned from the training set.

Algorithm1: Data filtering algorithm

Algorithm for Brain Tumor Detection Filtering

Input:

- Dataset: $\mathcal{D}_{\mathrm{BT}}$ (brain tumor dataset)
- Feature space: Σ

Procedure:

1. Read input dataset $\mathcal{D}_{\mathrm{BT}}$
   $\mathcal{D}_{\mathrm{BT}} \leftarrow \mathrm{ReadDataset}(\mathcal{D}_{\mathrm{BT}})$
2. Read each image from the input dataset $\mathcal{D}_{\mathrm{BT}}$
   For each image $\mathcal{I} \in \mathcal{D}_{\mathrm{BT}}$:
3. Find orientation using sine and cosine measures
$$\mathcal{D}_{\sigma} = \begin{cases} v\cos(\gamma) & \text{if } i \neq j \\ v\sin(\gamma) & \text{if } i = j \end{cases}$$
4. For each value in the block
5. Do
6. Compute probability noise as

$$P_v(\mathcal{D}_\sigma(y)) = \begin{cases} \dfrac{2\max\{\mathcal{D}_\sigma(y)\}e^{-\frac{(\mathcal{D}_\sigma(y)^2+\mu_\sigma)}{\sigma_{\mathcal{D}_\sigma(y)}}}}{v\cos(\gamma)\sigma_{\mathcal{D}_\sigma(y)}} & \text{if } i \neq j \\ \dfrac{2\min\{\mathcal{D}_\sigma(y)\}e^{-\frac{(\mathcal{D}_\sigma(y)^2+\mu_\sigma)}{\sigma_{\mathcal{D}_\sigma(y)}}}}{v\sin(\gamma)\sigma_{\mathcal{D}_\sigma(y)}} & \text{if } i = j \end{cases}$$

7. Done

Mathematical Notations:

- $\mathcal{D}_{\text{BT}}$: Brain tumor dataset
- Σ: Sensitive noise features
- σ: Individual feature
- γ: Angle used for sine and cosine transformations
- $\mathcal{D}_\sigma$: Transformed feature
- v: Noise factor
- μ_σ: Mean of the feature
- $\sigma_{\mathcal{D}_\sigma(y)}$: Standard deviation of the feature

Proposed Feature extraction measures

Log Inverse Differential Moment (LIDM)

LIDM is employed to determine the homogeneity of the image structure for image classification.

Max Correlation Inertia (MCI)

MCI measures the maximal correlation between the grey level linear dependence among pixels at given positions.

Probabilistic Gray-Level Co-occurrence Matrix (PGLCM) Algorithm

The PGLCM algorithm computes probabilistic measurements of image texture based on the spatial relationships between pixels of the same gray level.

Steps for the GLCM Algorithm

1. Convert the image to grayscale
   If the image is not already in grayscale, convert it to a grayscale image.
2. Define distance and direction parameters
   The distance parameter determines the number of pixels away from a given pixel to calculate the co-occurrence probability. The direction parameter determines the direction for the calculation.
3. Initialize the GLCM matrix with zeros
   The matrix size is $L \times L$, where L is the number of gray levels in the image.
4. Iterate through each pixel in the image
   For each pixel, calculate the co-occurrence probability with the pixel that is the specified distance and direction away. The co-occurrence probability is given by:
   $$P(i,j) = \frac{n(i,j)}{M \times N}$$
   where $n(i,j)$ is the number of occurrences of the pair of gray levels i and j, and M and N are the number of rows and columns in the image, respectively.
5. Increment the corresponding element in the GLCM matrix
   Update the GLCM matrix by the co-occurrence probability calculated in the previous step.
6. Normalize the GLCM matrix
   After iterating through all the pixels, normalize the GLCM matrix by dividing each element by the sum of all elements in the matrix.
7. Calculate probabilistic measurements
   Use the normalized GLCM matrix to calculate probabilistic measurements such as energy, entropy, contrast, and homogeneity using the following equations:
   $$\text{Energy} = \sum P(i,j)^2$$
   $$\text{Entropy} = -\sum P(i,j)\log P(i,j)$$
   $$\text{Contrast} = \sum (i-j)^2 P(i,j)$$
   $$\text{Homogeneity} = \sum \frac{P(i,j)}{1+(i-j)^2}$$
8. Return the probabilistic measurements

Proposed Bayesian feature selection

Estimation of Bayesian Probabilistic Measure

The Bayesian probabilistic measure is estimated using the following function:

Compressive Sensing Reconstruction Measure

The best compressive sensing reconstruction measure is computed using the following equation:

$$I_r = (1+\lambda)\frac{\sigma_E^2}{\sqrt{\max(\sigma_0^2 - \sigma_E^2, 0)}}$$

- B_r: Unique column values
  $$B_r = \text{uniCV}(\mathcal{D})$$
- HB: Histogram of $\mathcal{D}$
  $$HB = \text{Histobin}[] = \text{histogram}(\mathcal{D})$$
- Gaussian Kernel:
  $$\text{GK}(\phi,\theta) = e^{-\frac{\phi^2}{2\log(\theta)}}$$
- ψ:
  $$\psi = v \cdot \text{GK}\left(\sum HB_r, \frac{1}{2\sum B_r}\right)$$
- Exponential Gaussian Probability:
  $$\text{KP}(\mathcal{D}) = HB_r \cdot \left(\sum \log(y) \cdot HB_r\right)$$
- Poly Diffusion:
  $$\text{PD} = \text{KP}(\mathcal{D}) \cdot \frac{1}{2}\left(\int (\mathcal{I}_{\text{original}}(i,j) - \phi_{\text{noise}}(i,j))^2\, dx\, dy\right) + m \cdot \frac{2T_{\text{original}}(i,j)}{\alpha_x^2}$$
- Minimization function:
  $$\min(R(x)) = \text{PD} - \frac{1}{2}(\sigma_0^2 - \sigma_E^2) \parallel f \parallel^2 + L \cdot \phi(x)$$

Parameters and Variables

- σ_E^2: Error variance
- σ_0^2: Initial variance

- λ: Regularization parameter
- $\mathcal{D}$: Dataset
- ϕ: Parameter for Gaussian Kernel
- θ: Scale parameter
- ν: Scaling factor
- $\mathcal{I}_{\text{original}}$: Original image
- ϕ_{noise}: Noise component
- T_{original}: Original transformation
- α_x: Scaling parameter for x
- L: Regularization parameter
- f: Feature vector
- $R(x)$: Regularization function
- $\phi(x)$: Non-linear, non-smooth regularizer

Overall, the Bayesian feature selection algorithm aims to provide an estimate of the probabilistic measure that we can fit to a compressive sensing reconstruction function as an objective function of a L_1-norm compressive sensing algorithm for feature selection. The I_r measure incorporates a regularisation term λ and the error variance σ_E^2 normalised by the difference $\sigma_0^2 - \sigma_E^2$.

First, the list of unique values for the column (Br in mathematical terms) in the datasetD is calculated, followed by the histogram value of the dataset (HB). Finally, the Gaussian Kernel function is applied to the two values by using the transformation: w = GK(φ,θ)= (φ/θ)*exp(−wˆ2/θ2)Here above, φ is the histogramHB in the input, and θ is given by: θ = σ/√2 ln 2where σ is the scale factor derived from the list of unique values Br within a feature.

Finally, an exponential Gaussian probability is fitted based on HB and the logarithm of the histogram values. Now the probabilistic distribution is arranged in a way consistent with the data structure.

Additionally, a poly diffusion term comprised of the difference image $\mathcal{I}_{\text{distorted}} - \mathcal{I}_{\text{original}}$ is now averaged over the dimensions of the image, weighted by the same factor of the original transformation T_{original}, and by a parameter α_x.

In the final stage, the algorithm optimises the function $R(x)$, which is composed of the poly diffusion term, the variance difference and a regularisation term whose regulariser is $\phi(x)$, a non-linear, non-smooth regulariser. The iterative method is this way.

It preserves the sensitivity of the attributes that can cause errors, while optimising the utility of the dataset by maximising the PSNR and minimising the error rate. So the work can better filter brain tumor data.

Algorithm 2: hybrid filtering mean clustering algorithm

Characteristic Linearly-Weighted Sum Image

The characteristic linearly-weighted sum image ϵ_k is formed from the original image and its local neighbor average image, given by:

$$\epsilon_k = \frac{1}{1 + \alpha\left(\frac{\alpha}{N_r}\sum_{j\in N_r} x_j\right)\left(\frac{1}{k}\sum_{i=1}^{k} x(i)\cdot\beta\right)}$$

where $x(i)$ represents the edge points in the local neighborhood of x with β as a constant ranging from 0.1 to 0.9. ϵ_k is the gray value of the k-th pixel, and x_j represents the neighborhood of x_k. The parameter α is used to control the effect of the neighbor's term. The objective function for effectively segmenting the newly-generated image ϵ_k is defined as:

$$J_s = \sum_{i=1}^{p}\sum_{j=1}^{q} \gamma_j u_{ij}(\epsilon_j - v_i)^2$$

where v_i denotes the prototype of the i-th cluster, u_{ij} is the fuzzy membership of gray value j with respect to cluster i, and γ_j is the number of pixels with gray value equal to j.

Membership and Cluster Center Update

For J_s to be at its local minima, the membership values and cluster centers are updated using:

$$u_{ij} = \frac{(\epsilon_j - v_i)^{\frac{-2}{m-1}}}{\sum_{k=1}^{q}(\epsilon_j - v_k)^{\frac{-2}{m-1}}}$$

$$v_i = \frac{\sum_{j=1}^{q}\gamma_j u_{ij}^m \epsilon_j}{\sum_{j=1}^{q}\gamma_j u_{ij}^m}$$

Steps:

1. Set initial cluster center (c) ranges from 2 to $c_{\max}$, fix a certain value c for the work, and assign initial class prototypes and set $p \approx 0$ to a very small value.
2. Compute the new image ϵ using the equation:$\epsilon_k = \frac{1}{1+\alpha\left(\frac{\alpha}{N_r}\sum_{j\in N_r} x_j\right)\left(\frac{1}{k}\sum_{i=1}^{k} x(i)\cdot\beta\right)}$
3. Update the partition matrix using the membership update equation:$u_{ij} = \frac{(\epsilon_j - v_i)^{\frac{-2}{m-1}}}{\sum_{k=1}^{q}(\epsilon_j - v_k)^{\frac{-2}{m-1}}}$
4. Update the prototypes using the cluster center update equation:$v_i = \frac{\sum_{j=1}^{q}\gamma_j u_{ij}^m \epsilon_j}{\sum_{j=1}^{q}\gamma_j u_{ij}^m}$
5. Repeat steps 3-4 until $|v_n - v_{n-1}| < \rho$.

Apply CNN and Nonlinear SVM for class prediction.

4. Experimental results

The MRI images in the training dataset are applied to train the classifier. Besides, the second sub dataset holds a collection of ten HGT images with their individual ground truth image. A set of two classes

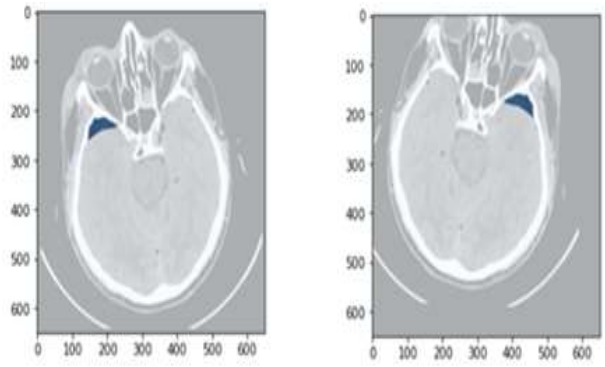

Figure 2. *Sample stroke image*

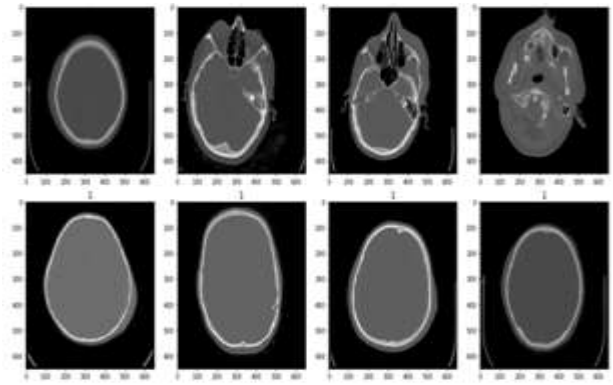

Figure 3. *Comparative analysis of proposed model.*

of images exist namely malignant and benign types.https://www.smir.ch/BRATS/Start2015RMS stands for "Root Mean Square" and is a measure of the difference between the predicted and actual values. In the context of BT segmentation data, RMS (in mm) indicates the average distance between the boundaries of the predicted segmentation and the actual segmentation, where a lower RMS indicates better accuracy. The Proposed Model has the lowest RMS (2 mm) among all the models, indicating the highest accuracy in BT segmentation. The MC-FCNN model has the second-lowest RMS (3.5 mm), followed by FCNN (4 mm), and U-Net (4.3 mm).

While figure 2 shows sample stroke image, the figure 3 shows comparative analysis of proposed model to the traditional techniques in terms of average PSNR ratio for all images with different levels of noise levels. On the other hand figure 4 is the comparative analysis of proposed segmentation-based classification model RMS to the conventional models RMS on Brain tumor dataset and figure 5 is the comparative analysis of proposed segmentation based classification model RMS to the conventional models VOE(%) on Brain tumor dataset. The VOE, or volume overlap error, is a measure of the disagreement between the ground truth and predicted volumes. A lower VOE indicates better performance.In this case, the Proposed Model has the lowest VOE value of 3%, which suggests that it has the highest level of agreement between the ground truth and predicted tumor volumes. The

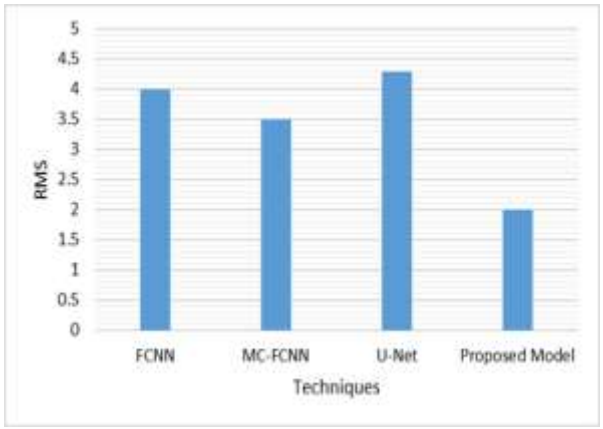

Figure 4. *Comparative analysis of proposed segmentation-based classification model.*

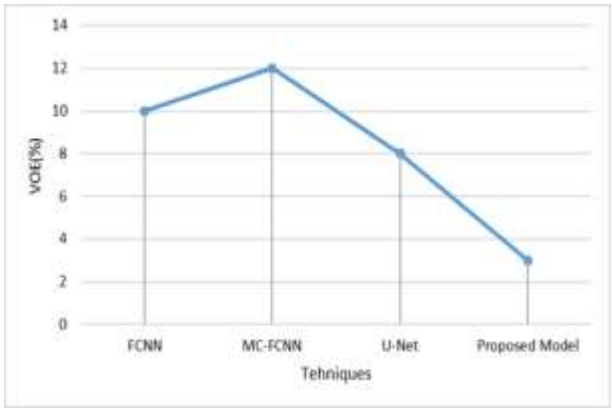

Figure 5. *Proposed model RMS to the conventional models VOE(%) on Brain tumor dataset*

FCNN and MC-FCNN have slightly higher VOE values of 10% and 12%, respectively, indicating slightly worse performance. The U-Net has the highest VOE value of 8%, indicating poorer performance than the Proposed Model but better than the FCNN and MC-FCNN.

5. Conclusion

The proposed hybrid algorithm for the detection of brain tumor from MRI scans is a complete ensemble learning algorithm for incorporation of robust data pre processing, advanced feature selection and clustering techniques before finnaly applying classifier. In order to select the most relevant features, ensemble feature selection methods are applied, which will add more weights to the informative features and subsequently, the accuracy of the detection system will increase in order to give better results. The proposed clustering algorithm is a hybrid mean algorithm of clustering which will form clusters of similar grouped points. An ensuing hybrid probabilistic weighted measure improves the clustering by finding the right balance between merging the data points of similar distributions and preserving dissimilar ones, while modelling the

underlying distributions of the data. Grouping data points into K-representative centres continues the process of simplifying the data, but with the inherent effect of retaining the information necessary for training CNN learning models. The filtering of the input data into all the different shapes and sizes using a multi-step algorithm guarantees consistently precise and reliable analysis. The hybrid approach is also successful in dealing with inherent variability and complex patterns arising from initial medical imaging data, making it a powerful and useful approach in the area of medical diagnosis and treatment planning.We also plan to examine the real-time implementation versions of the proposed system, as the use of deep learning techniques could potentially yield even better performance. In conclusion, concerted efforts should be made to integrate the multimodality of neuroscience data into healthcare as much as possible, as it appears to be increasingly significant in predicting and combining physical discrepancies or potential cognitive derailment, as captured by deep-learning models.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] M. O. Khairandish, M. Sharma, V. Jain, J. M. Chatterjee, and N. Z. Jhanjhi, (2022). A Hybrid CNN-SVM Threshold Segmentation Approach for Tumor Detection and Classification of MRI Brain Images. *IRBM*, 43(4);290–299, doi: 10.1016/j.irbm.2021.06.003.

[2] R. Sindhiya Devi, B. Perumal, and M. Pallikonda Rajasekaran, (2022). A hybrid deep learning based brain tumor classification and segmentation by stationary wavelet packet transform and adaptive kernel fuzzy c means clustering. *Advances in Engineering Software*, 170;103146, doi: 10.1016/j.advengsoft.2022.103146.

[3] O. Alpar, (2023). A mathematical fuzzy fusion framework for whole tumor segmentation in multimodal MRI using Nakagami imaging, *Expert Systems with Applications*, 216;119462, doi: 10.1016/j.eswa.2022.119462.

[4] E. Sert, F. Özyurt, and A. Doğantekin, (2019). A new approach for brain tumor diagnosis system: Single image super resolution based maximum fuzzy entropy segmentation and convolutional neural network, *Medical Hypotheses*, 133;109413, doi: 10.1016/j.mehy.2019.109413.

[5] Md. F. Ahamed et al., (2023). A review on brain tumor segmentation based on deep learning methods with federated learning techniques, *Computerized Medical Imaging and Graphics*, 110;102313, doi: 10.1016/j.compmedimag.2023.102313.

[6] K. Lamba, S. Rani, M. Anand, and L. P. Maguluri, (2024). An integrated deep learning and supervised learning approach for early detection of brain tumor using magnetic resonance imaging, *Healthcare Analytics*, 5;100336, doi: 10.1016/j.health.2024.100336.

[7] C.-J. Tseng and C. Tang, (2023). An optimized XGBoost technique for accurate brain tumor detection using feature selection and image segmentation, *Healthcare Analytics*, 4;100217, doi: 10.1016/j.health.2023.100217.

[8] S. Mohammadi, S. Ghaderi, K. Ghaderi, M. Mohammadi, and M. H. Pourasl, (2023). Automated segmentation of meningioma from contrast-enhanced T1-weighted MRI images in a case series using a marker-controlled watershed segmentation and fuzzy C-means clustering machine learning algorithm, *International Journal of Surgery Case Reports*, 111;108818, doi: 10.1016/j.ijscr.2023.108818.

[9] A. Batool and Y.-C. Byun, (2024). Brain tumor detection with integrating traditional and computational intelligence approaches across diverse imaging modalities - Challenges and future directions, *Computers in Biology and Medicine*, 175;108412, doi: 10.1016/j.compbiomed.2024.108412.

[10] S. P. Jakhar, A. Nandal, A. Dhaka, A. Alhudhaif, and K. Polat, (2024). Brain tumor detection with multi-scale fractal feature network and fractal residual learning, *Applied Soft Computing*, 153;111284, doi: 10.1016/j.asoc.2024.111284.

[11] K. Maheswari, A. Balamurugan, P. Malathi, S. Ramkumar, (2021). Hybrid clustering algorithm for an efficient brain tumor segmentation, *Materials Today: Proceedings*, 37(2);3002-3006, https://doi.org/10.1016/j.matpr.2020.08.718.

[12] Sharan Kumar, Dattatreya P. Mankame, (2020). Optimization driven Deep Convolution Neural Network for brain tumor classification, *Biocybernetics and Biomedical Engineering*, 40(3);1190-1204, https://doi.org/10.1016/j.bbe.2020.05.009.

[13] J. Prakash, R. Swathiramya, G. Balambigai, R. Menaha, & J.S. Abhirami. (2024). AI-Driven Real-Time Feedback System for Enhanced Student Support: Leveraging Sentiment Analysis and Machine Learning Algorithms. *International Journal of Computational and Experimental Science and Engineering*, 10(4);1567-1574. https://doi.org/10.22399/ijcesen.780

[14] Özlen, M. S., Cuma, A. B., Yazıcı, S. D., Yeğin, N., Demir, Özge, Aksoy, H., … Günay, O. (2024). Determination of Radiation Dose Level Exposed to Thyroid in C-Arm Scopy. *International Journal of Applied Sciences and Radiation Research*, 1(1);21-31. Retrieved from https://ijasrar.com/index.php/ijasrar/article/view/13

[15] Kosaraju Chaitanya, & Gnanasekaran Dhanabalan. (2024). Precise Node Authentication using Dynamic Session Key Set and Node Pattern Analysis for Malicious Node Detection in Wireless Sensor Networks. *International Journal of Computational and Experimental Science and Engineering*, 10(4);1462-1474. https://doi.org/10.22399/ijcesen.613

[16] Karpuz, N., & Serdar Ogel. (2024). Investigation of the Quality of Life of People in Need with the Effect of Social Assistance and Solidarity Foundation. *International Journal of Applied Sciences and Radiation Research*, 1(1);39-43. Retrieved from https://ijasrar.com/index.php/ijasrar/article/view/16

[17] Venkatraman Umbalacheri Ramasamy. (2024). Overview of Anomaly Detection Techniques across Different Domains: A Systematic Review. *International Journal of Computational and Experimental Science and Engineering*, 10(4);898-910. https://doi.org/10.22399/ijcesen.522

[18] Abuş, F., Gürçalar , A., Günay, O., Tunçman , D., Kesmezacar, F. F., & Demir, M. (2024). Quantification of Lens Radiation Exposure in Scopy Imaging: A Dose Level Analysis. *International Journal of Applied Sciences and Radiation Research*, 1(1);32-38. Retrieved from https://ijasrar.com/index.php/ijasrar/article/view/11

[19]Trivikrama Rao BATTULA, Narayana GARLAPATI, Srinivasa Rao CHOPPARAPU, Narasimha Swamy LAVUDIYA, & Prasad GUNDE. (2024). Real-Time E-commerce Insights with Mean Shift Clustering: A Dynamic Approach to Customer Understanding. *International Journal of Computational and Experimental Science and Engineering*, 10(4);1344-1350. https://doi.org/10.22399/ijcesen.607