Brain Tumor Detection Using GLCM and Machine learning Techniques

Yara Tarek
Department of Artificial intelligence, College of Information Technology, Misr University for Science & Technology.
Yara.tarek@must.edu.eg

Rania Elgohary
Department of Artificial intelligence, College of Information Technology, Misr University for Science & Technology.
Rania.elgohary@must.edu.eg

Mohaned Deif
Department of Artificial intelligence, College of Information Technology, Misr University for Science & Technology.
Mohanad.deif@must.edu.eg

***Abstract*— automated recognition of medical images poses a significant challenge in the field of medical image processing. These images are obtained from various modalities such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), etc., and are crucial for diagnosis purposes. In the medical field, brain tumor classification is very important phase for the further treatment. Human interpretation of large number of MRI slices (Normal or Abnormal) may leads to misclassification hence there is need of such a automated recognition system, which can classify the type of the brain tumor. The aim of this study is to detect brain tumor so we identify various features within an image. We extract the feature data from an image Using GLCM , LBP and other filters like Gaussian Filter, Sobel Filter, Laplace Filter, Gabor Filter, Hessian, Prewitt and create a data frame that can be fed into binary classification algorithms like** Logistic **Regression, KNN and decision tree**. **The accuracy achieved by Logistic Regression was 72%, KNN was 65% and decision tree was 80%.**

I. INTRODUCTION

Brain tumors indeed pose a significant health challenge, affecting individuals of all ages, from children to the elderly. The brain, being a highly complex organ, consists of billions of cells that work together to perform various functions crucial for human life. Brain tumors occurs when abnormal cells in the brain or its surrounding tissues multiply uncontrollably, forming a mass or growth. This uncontrolled cell growth can interfere with normal brain functions and cause damage to surrounding healthy tissue. The effects of brain tumors can vary widely depending on their location, size, type, and rate of growth. Common symptoms of brain tumors may include headaches, seizures, cognitive impairments, changes in behavior or personality, difficulties with balance or coordination, and sensory or motor disturbances. However, symptoms can vary greatly from person to person, and some tumors may remain asymptomatic until they reach a certain size or location. The treatment options for brain tumors depend on various factors, including the type and location of the tumor, its size and growth rate, and the overall health and preferences of the patient. Treatment may include surgery, radiation therapy, chemotherapy, targeted therapy, or a combination of these approaches. Early detection and diagnosis of brain tumors are crucial for improving outcomes and quality of life for patients. Advances in medical imaging techniques, such as magnetic resonance imaging (MRI) and computed tomography (CT) scans, have greatly facilitated the detection and characterization of brain tumors, allowing for more accurate diagnosis and treatment planning.

Features extracted from the MR image using methods like GLCM and LBP, then classifying the retrieved features with logistic regression, KNN, decision tree, are examples of efficient feature extraction and classification approaches.

The purpose of this research is to evaluate logistic regression, KNN, Decision tree classification algorithm for their ability to identify and categorize MR images of brain tumor. The results demonstrated that logistic regression classification was **72%** accurate, KNN accuracy was **65%** and decision tree accuracy was **80%.**

II. RELATED WORK

In [6], the authors presented a method for detecting brain tumors in MRI scans using a Support Vector Machine (SVM) classifier. They applied Grey-Level Co-occurrence Matrix (GLCM) and Discrete Wavelet Transform (DWT) as feature extraction techniques. By classifying test images into normal and tumorous brain MRIs, they achieved 93% accuracy with GLCM and 97% accuracy with DWT over the dataset.

In [7], the authors combined texture features from Grey-Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), and Gray-Level Run Length (GLRL). They employed various machine learning classification algorithms for tumor detection, categorizing MR images as normal or abnormal. Their proposed model achieved an average accuracy of 97.13%.In [8], the authors introduced a novel approach for brain tumor detection utilizing a convolutional neural network (CNN) with a transfer learning strategy alongside dimensionality reduction techniques. They employed EfficientNetB7 models for transfer learning to extract features and utilized Principal Component Analysis (PCA) to reduce the feature dimensionality. By combining features from PCA and the CNN EfficientNet model, they achieved an accuracy of 80%.In reference [9], the authors proposed a hybrid model that combines a Convolutional Neural Network (CNN) with a Support Vector Machine (SVM) to identify brain tumors in MRI images. They also implemented a pre-processing method on the MRI images, leading to a significant improvement in accuracy. The hybrid CNN-SVM model achieved an overall accuracy of 98.495%.

In reference [10], the authors classified brain MRIs into two categories: normal and diseased. They utilized Grey-Level Co-occurrence Matrix (GLCM) to extract features. For classification, they employed a probabilistic neural network (PNN) algorithm, achieving an impressive accuracy of 95%.

In reference [11], the authors introduced a model for tumor detection utilizing image processing techniques and machine learning algorithms. They employed Grey-Level Co-occurrence Matrix (GLCM) for feature extraction and the AdaBoost algorithm for tumor classification. Their approach achieved an accuracy of 89.90% in tumor classification.

In reference [12], the authors proposed an advanced approach for detecting brain tumors, incorporating preprocessing, segmentation, feature extraction, optimization, and detection processes. They utilized Grey-Level Co-occurrence Matrix (GLCM) for feature extraction and Convolutional Neural Network (CNN) classifiers for tumor detection. The system achieved an impressive accuracy of 98.9%.

Cheng et al[13]. presented a system for brain tumor analysis utilizing a dataset where the tumor area was delineated and subdivided into sub-areas using an adaptive spatial algorithm. Features were extracted using methods including Grey-Level Co-occurrence Matrix (GLCM), bag of words (BoW), and density graph. The system achieved commendable results, with GLCM features yielding an accuracy of 89.72%, BoW features achieving 91.28%, and density graph features reaching 87.54%.

Hao et al [14]. introduced a Deep Convolutional Neural Network (DCNN) approach that integrates symmetry into tumor region segmentation. They extended the network by incorporating symmetrical masks into certain layers. This method achieved good results, with an average Dice Similarity Coefficient (DSC) of 85.2%.Widhiarso et al[15]. employed Grey-Level Co-occurrence Matrix (GLCM) to extract features, which were then classified using Convolutional Neural Network (CNN). They combined GLCM features with contrast features, resulting in a system accuracy of 82%.Kumar and Vijayakumar [16]. Proposed utilizing Principal Component Analysis (PCA) in conjunction with a radial basis function kernel within Support Vector Machines (SVM) for the classification and segmentation of brain tumors. Their approach yielded a notable success rate of 94%.Solani et al [17]. conducted a study focusing on the challenges associated with diagnosing brain tumors, emphasizing the potential of MR imaging. They utilized statistical and machine learning techniques to detect brain tumors within a specific dataset.

III. METHODOLGY

3.1) Background

3.1.1) Preprocessing

Pre-processing plays a pivotal role in image processing techniques. Its primary function is to eliminate contaminants or unwanted noise from images to simplify processing while enhancing picture quality. The brain MRI images consist of a dataset of 253 images 98 for normal brain and 155 for abnormal brain . To prepare them for further analysis the images were resized to a standard size of 224x224 pixels. since various types of noise can deteriorate MRI brain images, employing different types of filters becomes essential to enhance MRI quality so we use some filters like Gaussian filter, Sobel filter, Gabor filter, hessian filter , Prewitt filter to enhance MRI quality.

3.1.2) GLCM feature extraction

GLCM is a statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM), also known as the gray-level spatial dependence matrix. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. (The texture filter functions, described in Calculate Statistical Measures of Texture cannot provide information about shape, that is, the spatial relationships of pixels in an image.).

After the creation of GLCMs using graycomatrix a several statistics can be derived from them using graycoprops these statistics provide information about the texture of an image.

greycomatrix - calculate the grey-level co-occurrence matrix , **greycoprops** - calculate texture properties of a GLCM , **Texture properties can be**

ontrast : Measures the local variations in the gray-level co-occurrence matrix.

$$\textbf{CON} = \sum_{i,j=0}^{n-1} (i-j)^2 \, \mathrm{p(i,j)} \qquad \textbf{(1)}$$

Where
CON: contrast

I &J : are intensity levels in the image, ranging from 0 to

$n-1n-1$.

(*i,j*): is the value from the co-occurrence matrix, representing the probability of a pixel with intensity i occurring adjacent to a pixel with intensity j.

2) Correlation: Measures the joint probability occurrence of the specified pixel pairs.

$$\textbf{CORT}=\sum_{i,j=0}^{n-1} \frac{(\mathrm{i}-\mu\mathrm{i})(\mathrm{j}-\mu\mathrm{j})p(i,j)}{\sigma\mathrm{i}\,\sigma\mathrm{j}} \qquad \textbf{(2)}$$

Where

CORT: Correlation

I &J : are intensity levels in the image, ranging from 0 to $n-1n-1$.

(*i,j*): is the value from the co-occurrence matrix, representing the probability of a pixel with intensity i occurring adjacent to a pixel with intensity j.

μi: is the mean of row sums of the GLCM

μj: is the mean of column sums of the GLCM

𝝈i: is the standard deviation of row sums of the GLCM.

𝝈j: is the standard deviation of column sums of the GLCM.

3) Homogeneity: Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

$$\textbf{HOM}=\sum_{i,j=0}^{n-1} \frac{p(i,j)}{1+(i-j)2} \qquad \textbf{(3)}$$

Where

HOM: Homogeneity

I &J : are intensity levels in the image, ranging from 0 to $n-1n-1$.

(i,j): is the value from the co-occurrence matrix, representing the probability of a pixel with intensity i occurring adjacent to a pixel with intensity j.

3.1.3) LBP features

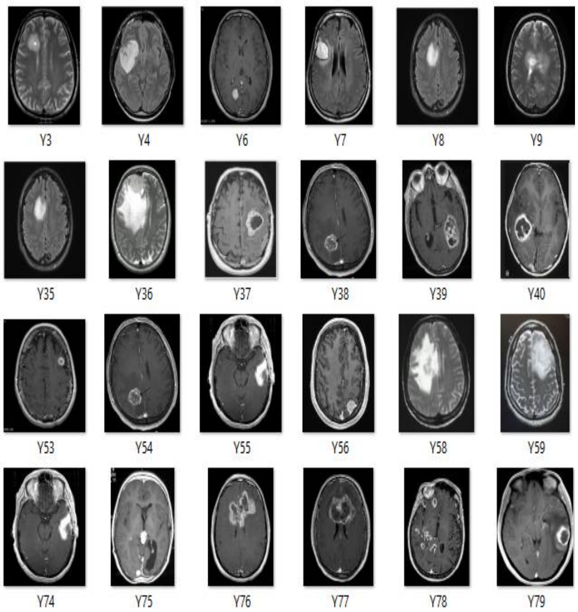
Local Binary Pattern (LBP) is a widely used texture descriptor in image processing and computer vision. LBP is particularly effective for texture analysis because it captures the local patterns in an image. This technique has been extensively applied in various domains, such as object recognition, face detection, and image segmentation. The step wise LBP Feature Extraction is, each pixel is compared with its neighboring pixels. Consider the Gray value of the center pixel as g_c, and the gray value of the neighboring pixel as g_p. Comparison is carried out with using Equation

$$\mathbf{S(g_p\,,\,g_c)=if\ \ g_p >= g_c\ \ else\ 0} \qquad \textbf{(4)}$$

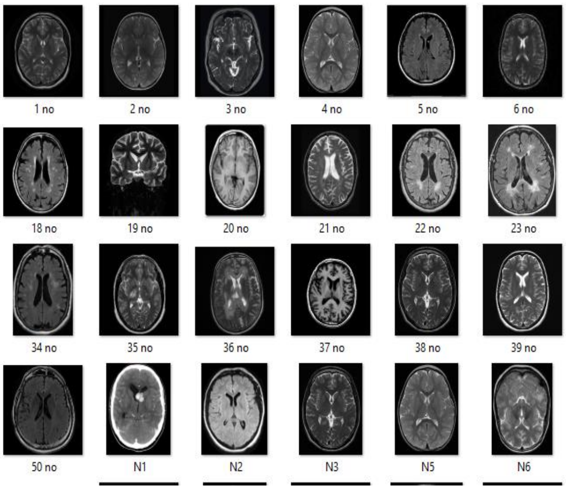
A circle of radius R is centered around the center pixel and the Function S is applied to P evenly spaced pixels.

3.2) Materials

The brain MR images dataset consist of 253 images 98 of them showed brain tumor and 155 of them showed healthy brain tissue.



Fig(1) Samples of images for Abnormal Brian



Fig(2) Samples of images for Normal Brian

3.3) Methods

Figure 3 is a schematic diagram of the overall methods used in this study. The approach is divided into three primary stages, including (1) brain MRI image pre-processing, (2) feature extraction, and (3) classification. Each step is explained below.

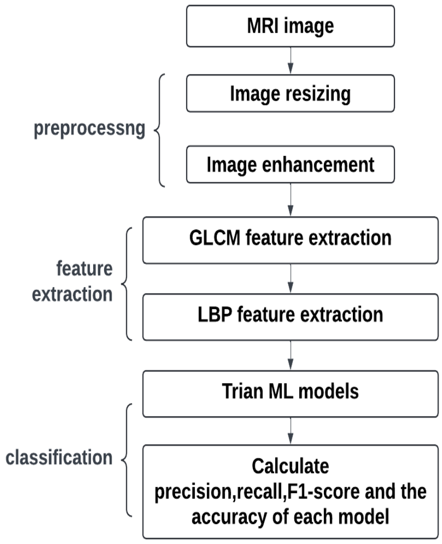


Fig (3): Proposed methodology

1) Preprocessing

To improve the classification results, some pre-processing steps were performed. The tasks completed during pre-processing are described below.

1.1 Image resizing

the images were resized to a standard size of 224x224 pixels. Resizing is crucial to standardize input dimensions for deep learning models, ensuring compatibility and consistent processing. The prevalent use of 224x224 pixel dimensions in computer vision is well-suited for automated computational techniques. To further enhance the quality and suitability of these resized images for analysis. Subsequently a crucial step was taken by Converting the images from the RGB (Red, Green, Blue) color space to grayscale. This conversion simplifies the images by representing each pixel with a single intensity value, which is especially beneficial for specific image analysis tasks.

1.2 Image enhancement

Filters play a crucial role in image processing, providing various functionalities that enhance, transform, and extract information from images. The filters used in this work was **gaussian filter** used to smooth out images and reduce random noise, improving the overall quality, **Sobel & Prewitt** used to identify and highlight edges within an image, **Gabor filters** are used to detect textures and patterns within an image, aiding in texture analysis and recognition, **Laplace filter** used in image processing to highlight areas of rapid intensity change, thus accentuating edges, **hessian filter** used for detecting complex structures like ridges, valleys, blobs, and other features in an image.

2) Feature extraction

In image processing and computer vision, feature extraction typically refers to the process of identifying and extracting meaningful patterns or structures from images. These patterns, known as features, can represent various characteristics of the image, such as edges, corners, textures, shapes, or colors.

3) Classification algorithms

3.1 logistic regressions

Logistic regression is a supervised machine learning algorithm used for classification tasks where the goal is to predict the probability that an instance belongs to a given class or not. Logistic regression is a statistical algorithm which analyzes the relationship between two data factors. The logistic regression model transforms the linear regression function continuous value output into categorical value output using a sigmoid function, which maps any real-valued set of independent variables input into a value between 0 and 1. This function is known as the logistic function.

3.2 KNN

KNN is one of the most basic yet essential classification algorithms in machine learning. It belongs to the supervised learning domain and finds intense application in pattern recognition, data mining, and intrusion detection.It is widely disposable in real-life scenarios since it is non-parametric, meaning it does not make any underlying assumptions about the distribution of data (as opposed to other algorithms such as GMM, which assume a Gaussian distribution of the given data). We are given some prior data (also called training data), which classifies coordinates into groups identified by an attribute.

3.3 Decision tree

A decision tree is one of the most powerful tools of supervised learning algorithms used for both classification and regression tasks. It builds a flowchart-like tree structure where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label. It is constructed by recursively splitting the training data into subsets based on the values of the attributes until a stopping criterion is met, such as the maximum depth of the tree or the minimum number of samples required to split a node.

During training, the Decision Tree algorithm selects the best attribute to split the data based on a metric such as entropy or Gini impurity, which measures the level of impurity or randomness in the subsets. The goal is to find the attribute that maximizes the information gain or the reduction in impurity after the split.

IV. RESULTS AND DISCUSION

On a dataset of MR images of brain tumors, the suggested approach for effective feature extraction and classification was evaluated using GLCM and LBP feature extraction methods and logistic regression, KNN and decision tree classification algorithms. Texture characteristics were extracted from the MR images using the GLCM, LBP methods. These features included contrast, energy, homogeneity, and correlation. Before feature extraction, the images pre-processed to eliminate artifacts and improve image quality. The brain MR images dataset consist of 253 images 98 of them showed brain tumors and 155 of them showed healthy brain tissue. The dataset is split into training and testing set with a ratio of 80:20. The classification model accuracy , precision , recall ,f1-score is evaluated using the following equations .

$$\textbf{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$\textbf{Precision} = \frac{TP}{TP+FP} \quad (6)$$

$$\textbf{Recall} = \frac{TP}{TP+FN} \quad (7)$$

$$\textbf{F1-score} = 2 * \frac{precision * recall}{precision + recall} \quad (8)$$

Where TP stands for true positive, TN stands for true negative, FP stands for false positive, and FN stands for false negative. According to the findings, the decision tree classifier has the highest accuracy.

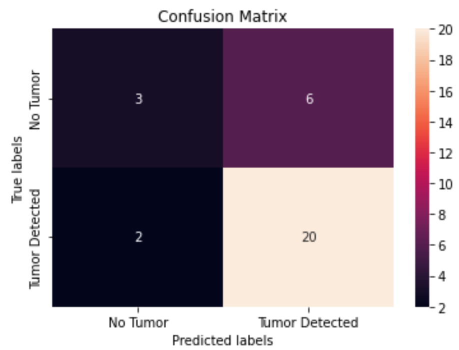


Fig (4) logistic regression confusion matrix

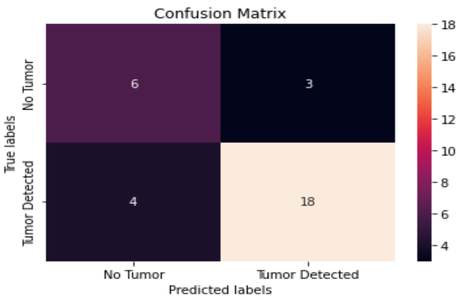


Fig (5) KNN confusion matrix

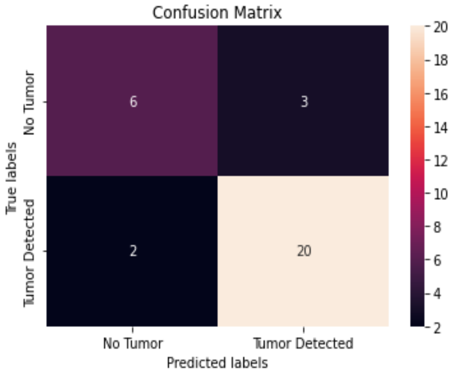


Fig (6) Decision tree confusion matrix

Table (1) comparison of performance metrics

Model	Precision		Recall		F1-score		Accuracy
	Normal	Abnormal	Normal	Abnormal	Normal	Abnormal	
Logistic regression	60%	77%	33%	91%	43%	83%	72%
KNN	60%	86%	67%	82%	63%	84%	65%
Decision tree	50%	84%	67%	73%	57%	78%	81%

CONCLUSION

In conclusion, our research focused on the vital task of brain tumor detection, employing GLCM and LBP feature extraction methods alongside logistic regression, KNN, and decision tree classification algorithms. Through rigorous evaluation on our dataset, we have gained valuable insights into the effectiveness of these techniques in accurately identifying brain tumors from MRI images. Our findings reveal that both GLCM and LBP feature extraction methods exhibit promising capabilities in capturing relevant information from the images, facilitating the discrimination between tumor and non-tumor regions. Moreover, by employing logistic regression, KNN, and decision tree classifiers, we were able to effectively classify these features, further enhancing the accuracy of our detection system. The comprehensive evaluation of our dataset demonstrated the robustness and versatility of the proposed approach. While logistic regression offers simplicity and interpretability, KNN leverages the local similarity of data points, and decision trees provide insight into the decision-making process. Each algorithm contributes unique strengths to the detection task, ultimately leading to a more comprehensive and reliable brain tumor detection system.

REFRENCES

[1]B.A.; Alazmi, M.; Alayba, A.M.; Alreshidi, A.; Alshahrani, M. Diagnosis of Histopathological Images to Distinguish Types of Malignant Lymphomas Using Hybrid Techniques Based on Fusion Features. Electronics 2022, 11, 2865.

[2] Koklu, M.; Unlersen, M.F.; Ozkan, I.A.; Aslan, M.F.; Sabanci, K. A CNN-SVM study based on selected deep features for grapevine leaves classification. Measurement 2022, 188, 110425.

[3] Mohammed, B.A.; Senan, E.M.; Al-Mekhlafi, Z.G.; Rassem, T.H.; Makbol, N.M.; Alanazi, A.A.; Almurayziq, T.S.; Ghaleb, F.A.; Sallam, A.A. Multi-Method Diagnosis of CT Images for Rapid Detection of Intracranial Hemorrhages Based on Deep and Hybrid Learning. Electronics 2022, 11, 2460

[4] Alanazi, M.F.; Ali, M.U.; Hussain, S.J.; Zafar, A.; Mohatram, M.; Irfan, M.; AlRuwaili, R.; Alruwaili, M.; Ali, N.H.; Albarrak, A.M. Brain Tumor/Mass Classification Framework Using Magnetic-Resonance-Imaging-Based Isolated and Developed Transfer Deep-Learning Model. Sensors 2022, 22, 372.

[5] Khan, A.H.; Abbas, S.; Khan, M.A.; Farooq, U.; Khan, W.A.; Siddiqui, S.Y.; Ahmad, A. Intelligent Model for Brain Tumor Identification Using Deep Learning. Appl. Comput. Intell. Soft Comput. 2022, 2022, 8104054.

[6] H. Mohsen, E.-S. A. El-Dahshan, E.-S. M. El-Horbaty, and A.-B. M. Salem, "Intelligent methodology for brain tumors classification in magnetic resonance images," 2017.

[7] N. Gupta, P. Bhatele, and P. Khanna, "Glioma detection on brain MRIs using texture and morphological features with ensemble learning," Biomedical Signal Processing and Control, vol. 47, pp. 115–125, 2019.

[8] P. Modiya and S. Vahora, "Brain Tumor Detection Using Transfer Learning with Dimensionality Reduction Method," International Journal of Intelligent Systems and Applications in Engineering, vol. 10, no. 2, pp. 201–206, 2022.

[9] M. O. Khairandish, M. Sharma, V. Jain, J. M. Chatterjee, and N. Jhanjhi, "A hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images," Irbm, vol. 43, no. 4, pp. 290–299, 2022.

[10] N. Varuna Shree and T. Kumar, "Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network," Brain informatics, vol. 5, no. 1, pp. 23–30, 2018.

[11] A. Minz and C. Mahobiya, "MR image classification using adaboost for brain tumor type," in 2017 IEEE 7th International Advance Computing Conference (IACC), 2017, pp. 701–705

[12] P. K. Ramtekkar, A. Pandey, and M. K. Pawar,

“Innovative brain tumor detection using optimized deep learning techniques,” International Journal of System Assurance Engineering and Management, pp. 1–15, 2023.

[13] Cheng, J.; Huang, W.; Cao, S.; Yang, R.; Yang, W.; Yun, Z.; Wang, Z.; Feng, Q. Enhanced performance of brain tumor classification via tumor region augmentation and partition. PloS One 2015, 10, e0140381.

[14] Chen, H.; Qin, Z.; Ding, Y.; Tian, L.; Qin, Z. Brain tumor segmentation with deep convolutional symmetric neural network. Neurocomputing 2020, 392, 305–313.

[15] Widhiarso, W.; Yohannes, Y.; Prakarsah, C. Brain tumor classification using gray level co-occurrence matrix and convolutional neural network. IJEIS (Indones. J. Electron. Instrum. Syst.) 2018, 8, 179–190.

[16] Kumar P, Vijayakumar B. Brain tumour mr image segmentation and classification using by PCA and RBF kernel based support vector machine. MiddleEast J Sci Res (2015) 23(9):2106–16.

[17] Solanki S, Singh UP, Chouhan SS, Jain S. Brain tumor detection and classification using intelligence techniques: an overview. IEEE Access(2023)11:12870–86.