Mutual Information-Based Optimum Metrics Identification in Symmetry-Based Brain Abnormality Detection

Mohammad A. N. Al-Azawi
Oman College of Management and Technology, Muscat, Sultanate of Oman
Email: mohd.alazawi@omancollege.edu.om

Abstract.

In our previous studies, we showed that brain abnormalities can be detected by comparing the features extracted from the two lobes with each other. Based on this, many metrics, such as those extracted from colour or texture features, have been extracted and used. The large number of extracted metrics posed a challenge in terms of how important each metric is. In this research, we use the mutual information content to measure the importance of the metrics and their influence on the classification process as it gives an indication of how the output and each input are related to each other. The algorithm was applied to 366 images, from which eleven metrics were extracted and studied. Random forest classifier was used as it was proven that it gives the highest accuracy. The obtained results showed that 30% of the features can be eliminated without a significant effect on the accuracy.

1. Introduction

Magnetic resonance imaging (MRI) is one of the most well-known methods to detect and identify brain abnormality due to its availability and the quality of the images it produces [1]. MR image processing methods go through five stages which are respectively: (i) preprocessing, (ii) feature extraction, (iii) feature selection, (iv) classification, and (v) segmentation [2]. In this research, our interest is focused on the second and third stages, which are the extraction and selection of features. In this study, the method of using the symmetry between the two lobes of the brain is relied upon to detect any defects in it as tumours cause deformation of the shape of one of the lobes and distort this symmetry. To perform this comparison, we used eleven metrics extracted from each lobe. These metrics are compared with each other to decide whether the brain is normal or has some kind of abnormalities. To study the importance of each metric separately, we needed to use a feature selection method, which is the main interest of this study.

The main contribution of this research is studying the various metrics that are extracted from the colour and texture features and using the mutual information measure to identify the relevant metrics that improve the performance of the classification process.

Based on previous studies, the best classification accuracy was obtained when using random forest classifier (RFC), hence, we adopted this classifier in our study. This is because of the advantages of this classifier in overcoming the overfitting and local optima problems.

The proposed algorithm was implemented using Python programming with JupyterLab version 2.2.6 on a PC with intel core i7 and 16 GB RAM. The algorithm was tested and applied to a dataset containing 366 various brain images constructed by selecting subsets from the “Brain MRI Images for Brain Tumor Detection” [3] and “Brain Tumor Classification (MRI)” [4] datasets. The selected images are axial images and with types of T1W, T2W, and Flair. More information about the types...
of the images is available in ref. [5]. The images are single-band with 256 grey levels and a size of 200 × 200.

The remainder of this paper is structured as follows. Section 2 presents the necessary theoretical background and reviews some of the state-of-the-art work. The proposed approach is presented in Section 3. In section 4, a thorough discussion and evaluation of the obtained results is presented. Finally, the conclusions are derived in section 5.

2. Background and Review

Brain tumour identification has been the focus of many studies in the past few decades and as a result of the large amount of published work, many reviews are available listing the methods and techniques used. Brain MR image processing approaches are divided primarily into two types: non-AI-based [6], [7], [8], [9] and [5], and artificial intelligence-based approaches [10], [11], [12], and [13]. Non-AI-based approaches use traditional image processing techniques, such as thresholding, clustering, segmentation, and edge detection to identify and segment the abnormality of the regions. They are classified in literature into pixel-based, edge-based, and region-based. Pixel-based approaches use pixel features such as intensity, colour-band value, and location in the analysis process. Thresholding and clustering are the most commonly used pixel-based methods where the features of the pixels are used to split the image into normal and abnormal regions [14]. Split/merge and region growing techniques are examples of region-based approaches which deal with regions rather than pixels [15].

On the other hand, AI-based approaches utilise the principles of artificial intelligence and machine learning, such as neural nets, Fuzzy, support vector machines, decision trees, deep learning, and others, to detect and identify the abnormality region in the brain. AI techniques such as classification and regression can be used in the various processing stages of the detection, identification, and segmentation processes. According to the scope of our research, we will focus on AI-based approaches which will be discussed in the next section.

2.1. AI-based algorithms

In this section, we shed light on some of the state-of-the-art related work. As mentioned earlier, although many studies have been published, we will only consider relevant AI-based methods in our discussion.

El-Dahshan et al. suggested an approach that uses a feed-forward back-propagation artificial neural network and k-nearest neighbour classifiers to classify whether the brain is normal or abnormal [16]. Zöllner et al. suggested an algorithm that utilises support vector machine (SVM) to perform the classification [17]. Zanaty proposed a hybrid approach, combining Fuzzy C-means, seed region growing, and Jaccard similarity coefficient algorithm for segmentation [18]. Cui et al. presented an approach that uses a localized fuzzy clustering with spatial information for clustering the images [19]. Sachdeva et al. used artificial neural networks and principal components analysis artificial neural networks to classify, segment, and extract features from MR images [20]. El-Dahshan et al. used Pulse Coupled Neural Network for segmenting the images [21]. Kumar and Vijayakumar introduced a method that uses principal component analysis, radial basis function and kernel-based SVM for segmentation and classification [22]. Damodharan and Raghavan presented a neural network-based method for classification [23]. Anitha and Murugavalli suggested a method that uses K-means, self-organising map neural network and K-nearest neighbour, and discrete wavelet transform features for segmenting and isolating the tumour from the rest of the brain tissues [24]. Bahadure et al. used Berkeley wavelet transformation, SVM, adaptive neuro-fuzzy inference system, backpropagation artificial neural network and K-nearest neighbour in the classification process [25]. Gilanie et al. used Gabor texture features and SVM to classify the images into normal and abnormal [26]. Bahadure et al. used genetic algorithms to classify and segment the tumour in brain MR images [13]. Özyurt et al. used convolutional neural networks with maximum fuzzy sure entropy for classification [27].
2.2. **Mutual information gain**

According to Battiti, mutual information (MI) can be used in selecting the optimum feature for any classification problem [28]. It was assumed that relevant features have a major impact on the output in any classification problem as they contain important information about the output, while some features may be irrelevant or have a small impact. Such features are known as irrelevant features and removing them may not affect the accuracy of the classification process much. MI was used in literature to evaluate and select the features such as in references [28], [29], [30] and [31] since it can find the relation between each feature and the output and then assign importance to that feature.

The Entropy from Shannon’s information theory (1949), which is given in equation (1), was widely used as a measure of the information contents.

\[
H(X) = - \sum_{i=0}^{n} P(x_i) \log_2 P(x_i)
\]  

(1)

where \(X\) is a discrete random variable expressed as \(X = \{x_0, x_1, ..., x_n\}\) and \(P(x_i)\) is the probability of occurrence of the variable \(x_i\).

To calculate the conditional entropy, we need to define another discrete random variable, be it \(Y = \{y_0, y_1, ..., y_m\}\). The conditional entropy of the variable \(Y\), which is the amount of uncertainty left in the variable \(Y\) after the variable \(X\) is introduced, can be calculated using the formula given in equation (2).

\[
H(Y|X) = - \sum_{x_i \in X} \sum_{y_j \in Y} P(y_j, x_i) \log_2 P(y_j|x_i)
\]  

(2)

The joint entropy of \(X\) and \(Y\), which is denoted as \(H(Y, X)\), is defined as the uncertainty that occurs simultaneously with two variables and can be calculated as given in equation (3).

\[
H(Y, X) = - \sum_{x_i \in X} \sum_{y_j \in Y} P(y_j, x_i) \log_2 P(y_j|x_i)
\]  

(3)

\[
H(Y, X) = H(X) + H(Y|X)
\]

Finally, the mutual information between variable \(Y\) and variable \(X\) can be defined as shown in equation (4).

\[
I(Y;X) = \sum_{x \in X} \sum_{y \in Y} P(y, x) \log_2 \left( \frac{P(y, x)}{P(x)P(y)} \right)
\]  

(4)

\[
I(X; Y) = H(X) + H(Y) - H(X, Y)
\]

Since the mutual information is defined as the amount of information that variable \(X\) contains about variable \(Y\), it can be used to indicate the level of shared information between the random variables. For the variables to be related to each other, the value of \(I(X; Y)\) should be high [32].

2.3. **Features and metrics**

In this section, we present the selected features and the metrics extracted from them that were used in our approach. Two main features were considered namely, colour feature, and texture feature. From colour feature, different metrics can be derived such as statistical metrics and colour distribution metrics. In our experiments, we adopted some statistical metrics extracted from the
intensity of each lobe such as mean ($\mu_h$) and standard deviation ($\sigma_h$), the information about the equations and the calculation is available at [33].

Due to the nature of the brain tissues, texture features can be very useful in describing the contents of its MR images. Grey Level Co-occurrence Matrix (GLCM) is one of the important features that can be used to describe the texture of the image. The metrics that were used here are Energy ($E_{GLCM}$), Dissimilarity ($\delta_{GLCM}$), Homogeneity ($h_{GLCM}$), Contrast ($C_{GLCM}$) and Correlation ($\rho_{GLCM}$). The details and the calculation of the metrics can be found in references such as [34].

Gabor filter is another approach that can provide important metrics which can be used to compare the tissues of the two lobes. From Gabor filters, the Energy ($E_{GAB}$) and the Entropy ($\epsilon_{GAB}$) were extracted and used. The formulas and the theoretical concepts of Gabor filters can be found in references such as [26].

The third set of metrics was derived from the local binary pattern (LBP) which is one of the texture spectrum descriptors. The metrics used are Entropy ($\epsilon_{LBP}$) and Energy ($E_{LBP}$). For more information about LBP, the reader can refer to [35], [36], and [37].

2.4. Feature selection
To avoid any confusion, in literature, the word features refer to the independent variables or the inputs of the classifier, while in this research we use features to refer to some properties of the images that metrics can be extracted from such as colour and texture features. While the inputs of the classifier are the metrics.

Two types of methods are used in feature selection, filter-based methods and wrapper-based methods. Each one of these types has its merits and demerits. Filter-based methods are faster than wrapper-based ones because they estimate the importance of each feature individually. While wrapper methods are more accurate as the importance of feature subsets is measured using a classification algorithm [38] [32]. In our research, we adopted a filter-based method to study the features individually.

3. Proposed Approach
The flowchart of the proposed algorithm is shown in Figure 1, where each image in the original dataset undergoes a preprocessing process that includes some basic preparations such as noise removal and registration [39]. Although the pre-processing stage is important to the proposed algorithm as it relies on the symmetry around the vertical axis, it is beyond the interest of this research as most of the available datasets include registered images and many approaches are available in the literature.

The next stage is to divide the image into two images, left and right, where the metrics are extracted from the features of these images and the distances between the corresponding measures are calculated and normalised to create the dataset used in the classification process. The resulting dataset is divided into two subsets, one is used to train the model and the other is used to test the model. Finally, if the accuracy obtained is acceptable, the model can be used for prediction. The following algorithm provides further clarification of the proposed approach.

1- Read the Dataset $\mathbb{D}_o$, which includes $N$ images.
2- For each image $I_i$, $0 < i \leq N$ in the dataset do the following:
   a. Divide the image around the vertical axis into two halves $L_i$ and $R_i$.
   b. Extract the metrics for each half $ML_i$ and $MR_i$.
   c. Calculate the distance vector $d_i = \{\delta_{i,1}, \delta_{i,2}, ..., \delta_{i,N}\}$ for each image.
3- Normalise the metrics vector dataset, $\hat{d}_i = \{\hat{\delta}_{i,1}, \hat{\delta}_{i,2}, ..., \hat{\delta}_{i,N}\}$.
4- Construct the dataset $\mathbb{D}_c = \{(d_1, c_1), (d_2, c_2), ..., (d_N, c_N)\}$.
5- For each normalised metrics vector $\hat{\delta}_k = \{\hat{\delta}_{1,k}, \hat{\delta}_{2,k}, ..., \hat{\delta}_{N,k}\}$, calculate the MI with the class $\hat{C}$. 

4
6- Sort the MI vector and select the appropriate percentile.

Figure 1. The proposed approach flowchart.

4. Results and discussion
To simplify the presentation of the results, we will refer to images of brains with abnormalities as (abnormal) and images of normal brains as (normal). The proposed algorithm was applied to a dataset containing 366 images, including 158 images of normal brains and 208 images of abnormal brains. Eleven metrics were extracted from the lobes of the same image, and the distances between the corresponding metrics were calculated and normalised.

Table 1 shows the normalised average distances of the various metrics used in our tests. From the table, it is clear that the normalised average distances of the abnormal brains are higher than those for normal brains which proves the hypothesis that abnormalities cause deformation in one of the lobes and distort the symmetry between them.

Table 1. Normalised average distances for the metrics derived from the features of the normal and abnormal brains.

| Brain Type | $\mu_1$ | $\sigma_1$ | $E_{LBP}$ | $E_{GGL}$ | $E_{GLCM}$ | $h_{GLCM}$ | $\rho_{GLCM}$ | $E_{GRB}$ | $E_{GAR}$ | $E_{GAR}$ |
|------------|---------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Normal     | 0.02085 | 0.01584   | 0.00918   | 0.01172   | 0.01456   | 0.02716   | 0.02304   | 0.01830   | 0.11892   | 0.02078   |
| Abnormal   | 0.12203 | 0.06678   | 0.09008   | 0.10021   | 0.09087   | 0.12233   | 0.09826   | 0.12069   | 0.16648   | 0.15801   |

To evaluate the results obtained from the proposed algorithm, we shall adopt the precision/recall/accuracy measures which are widely used in literature. To calculate these measures, we first need to provide some definitions as follows:

P: Positive samples, which is the number of abnormal images.  
N: Negative samples, which is the number of normal images.  
TP: True positive, which is the number of correctly identified images from the positive samples.  
TN: True negative, which is the number of correctly identified images from the negative samples.  
FP: False positive which is the number of incorrectly identified images from the positive samples.  
FN: False negative, which is the number of incorrectly identified images from the negative samples.
The precision (Pr), recall (Rec), F-Scale (FS), and accuracy (Acc) are then can be calculated using the formulas given in equation (5).

\[
\begin{align*}
Pr & = \frac{TP}{TP + FP} \\
Rec & = \frac{TP}{TP + FN} \\
FS & = \frac{2 \times Pr \times Rec}{TP + FN} \\
Acc & = \frac{Pr + Rec}{TP + TN + FP + FN}
\end{align*}
\]

Table 2 shows a sample of the results obtained, in this table, we studied the precision (P), recall (R), F-scale (FS), and Accuracy (Acc) for each metric separately. The mentioned metrics were measured to both classes, normal (0) and abnormal (1). Furthermore, the macro average (avg) and the weighted average (wa) were calculated for all metrics.

Table 2. Training time, Precision, Recall, FS, and Accuracy for each metric.

| Metrics | Time (ms) | P (1) | R (1) | FS (1) | P (0) | R (0) | FS (0) | P avg | R avg | FS avg | Pr avg | Rec avg | FS avg | Acc avg |
|---------|-----------|-------|-------|--------|-------|-------|--------|-------|-------|--------|--------|--------|--------|---------|
| H_k     | 403       | 0.77  | 0.73  | 0.75   | 0.66  | 0.70  | 0.68   | 0.71  | 0.72  | 0.71   | 0.72   | 0.72   | 0.72   | 71.8    |
| e_LBP   | 288       | 0.80  | 0.86  | 0.81   | 0.78  | 0.66  | 0.71   | 0.77  | 0.76  | 0.76   | 0.77   | 0.77   | 0.77   | 77.3    |
| E_GLCM  | 255       | 0.86  | 0.87  | 0.87   | 0.83  | 0.81  | 0.82   | 0.84  | 0.84  | 0.84   | 0.85   | 0.85   | 0.85   | 84.5    |
| δ_GLCM  | 291       | 0.83  | 0.79  | 0.81   | 0.74  | 0.79  | 0.76   | 0.79  | 0.79  | 0.79   | 0.79   | 0.79   | 0.79   | 79.1    |
| C_GLCM  | 282       | 0.79  | 0.83  | 0.81   | 0.75  | 0.70  | 0.73   | 0.77  | 0.77  | 0.77   | 0.77   | 0.77   | 0.77   | 77.3    |
| p_GLCM  | 275       | 0.85  | 0.73  | 0.79   | 0.70  | 0.83  | 0.76   | 0.78  | 0.78  | 0.78   | 0.79   | 0.79   | 0.79   | 77.3    |
| E_GAB   | 283       | 0.79  | 0.84  | 0.82   | 0.77  | 0.70  | 0.73   | 0.78  | 0.77  | 0.77   | 0.78   | 0.78   | 0.78   | 78.2    |
| e_GAB   | 279       | 0.81  | 0.90  | 0.86   | 0.85  | 0.72  | 0.78   | 0.83  | 0.81  | 0.83   | 0.83   | 0.83   | 0.83   | 82.7    |

Table 3 shows the relationship between the mutual information (MI) of each metric with the output and the accuracy it produces when used alone. From the table, it is clear that the metrics that produced higher MI value also produced higher accuracy and vice versa. This is evident from the correlation value between them which is 83%.

Table 3. Accuracy vs. mutual information contents MIC for each metric.

| Name  | ETGB | ETLBP | EGLBP | EGGGB | MEAN | EGGL | HGL | DGL | STDV | CGL | CRGL |
|-------|------|-------|-------|-------|------|------|-----|-----|------|-----|------|
| Metric | e_GAB | e_LBP | E_LBP | E_GAB | μ_k | E_GLCM | δ_GLCM | σ_k | E_GLCM | p_GLCM |
| Accuracy | 82.7 | 84.5 | 86.4 | 78.2 | 71.8 | 84.5 | 77.3 | 79.1 | 77.3 | 59.1 |
| MIC | 0.4464 | 0.3925 | 0.3651 | 0.3461 | 0.3077 | 0.28182 | 0.268546 | 0.260788 | 0.254211 | 0.233489 | 0.039644 |

Table 4 shows the evaluation measures Pr, Rec, FS, and Acc in addition to the time for each percentile selected.

Table 4. Evaluation measures vs. percentile.

| Percentile | Time (ms) | Pr | Rec | FS | Acc |
|------------|-----------|----|-----|----|-----|
| 10         | 279       | 0.83 | 0.83 | 0.82 | 82.7 |
| 20         | 293       | 0.91 | 0.91 | 0.91 | 90.9 |
| 30         | 381       | 0.94 | 0.94 | 0.94 | 93.6 |
| 40         | 299       | 0.94 | 0.93 | 0.93 | 92.7 |
| 50         | 342       | 0.96 | 0.95 | 0.95 | 95.45 |
| 60         | 377       | 0.96 | 0.95 | 0.95 | 95.45 |
| 70         | 380       | 0.97 | 0.96 | 0.96 | 96.36 |
| 80         | 380       | 0.97 | 0.97 | 0.97 | 97.27 |
The results are given also in Figure 2, which shows various graphs to evaluate the performance of the algorithm. The percentile here refers to the percentage of selected features considering those with the highest MI values first. In this figure, (a) shows the accuracy against the percentile, (b) gives the precision, recall and FS against percentile and (c) shows the learning curve of the classifier. From the curves in (a) and (b), it is clear that the evaluation measures are not improved always with the increase of percentile i.e. the number of features used in the training process. For example, consider the case when 30% of the features were selected which showed a performance reduction. Although the highest accuracy was obtained when 100% of the features were selected, the accepted accuracy can be obtained when 70% of the features is used.

5. Conclusions

In this study, we used mutual information contents to study the significance of each of the metrics used to classify brains into normal and abnormal. The method of the symmetry between the two lobes of the brain was relied upon to determine whether the brain has any kind of tumours. This method was proven in previous studies as accurate and gave excellent results and high accuracy. In order to apply the algorithms that depend on the symmetry between the two lobes of the brain, we need to extract some measures from each lobe and compare these measures and judge whether the brain contains some distortions. If the difference between the measures is large, the probability that the brain is abnormal is large and vice versa. Due to the fact that different metrics have different and different influence on the classification process, there was a need to study the importance of each metric and the extent of its impact on the accuracy of classification. Therefore, in this work, eleven metrics were studied and their importance and extent of their impact on the accuracy of the classification process based on the extent to which each measure linked with the output. The results obtained show that 30% of the features can be eliminated and this will not affect the results significantly. Further work can be done by considering other metrics that can be extracted from the images and by using wrapper-based methods to select the optimum features.

References

[1] M. S. H. Al-Tamimi and G. Sulong, “Tumor brain detection through MR images: A review of literature,” J. Theor. Appl. Inf. Technol., vol. 62, no. 2, pp. 387–403, 2014.

[2] D. M. Kumar, D. Satyanarayana, and M. N. G. Prasad, “An improved Gabor wavelet transform and rough K-means clustering algorithm for MRI brain tumor image segmentation,” Multimed. Tools Appl., Oct. 2020, doi: 10.1007/s11042-020-09635-6.

[3] N. Chakrabarty, “Brain MRI Images for Brain Tumor Detection,” 2018.
https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection (accessed Oct. 01, 2020).

[4] Sartaj, “Brain Tumor Classification (MRI),” 2020. https://www.kaggle.com/sartajbhuva/brain-tumor-classification-mri (accessed Aug. 01, 2020).

[5] M. Al-azawi, “Saliency-Based Brain Abnormality Identification in Magnetic Resonance Images,” Sci. Vis., vol. 13, no. 2, pp. 24–49, 2021, doi: 10.26583/sv. 13.2.03.

[6] S. R. Rajput and M. S. Raval, “A Review on End-To-End Methods for Brain Tumor Segmentation and Overall Survival Prediction,” arXiv, 2020, doi: 10.32010/26166127.

[7] S. M. Kulkarni and G. Sundari, “A Review on Image Segmentation for Brain Tumor Detection,” in Proceedings of the 2nd International Conference on Electronics, Communication and Aerospace Technology, ICECA 2018, 2018, vol. 3, pp. 552–555, doi: 10.1109/ICECA.2018.8474893.

[8] A. Wadhwa, A. Bhardwaj, and V. Singh Verma, “A review on brain tumor segmentation of MRI images,” Magnetic Resonance Imaging, vol. 61, pp. 247–259, 2019, doi: 10.1016/j.mri.2019.05.043.

[9] S. Tripathi, R. S. Anand, and E. Fernandez, “A Review of Brain MR Image Segmentation Techniques,” Int. J. Res. Anal. Rev., vol. 5, no. 2, p. 1295, 2018, [Online]. Available: http://ijrar.com/.

[10] A. Işin, C. Direkoğlu, and M. Şah, “Review of MRI-based Brain Tumor Image Segmentation Using Deep Learning Methods,” in Procedia Computer Science, 2016, vol. 102, pp. 317–324, doi: 10.1016/j.procs.2016.09.407.

[11] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, “Medical Image Analysis using Convolutional Neural Networks: A Review,” J. Med. Syst., vol. 42, no. 11, 2018, doi: 10.1007/s10916-018-1088-1.

[12] L. Pei, L. Vidyaratne, M. M. Rahman, and K. M. Iftekharuddin, “Context aware deep learning for brain tumor segmentation, subtype classification, and survival prediction using radiology images,” Sci. Rep., vol. 10, no. 1, pp. 1–11, 2020, doi: 10.1038/s41598-020-74419-9.

[13] N. B. Bahadure, A. K. Ray, and H. P. Thethi, “Comparative Approach of MRI-Based Brain Tumor Segmentation and Classification Using Genetic Algorithm,” J. Digit. Imaging, vol. 31, no. 4, pp. 477–489, Aug. 2018, doi: 10.1007/s10278-018-0050-6.

[14] M. A. N. Al-azawi, “Image Thresholding using Histogram Fuzzy Approximation,” Int. J. Comput. Appl., vol. 83, no. 9, pp. 36–40, 2013, doi: 10.5120/14480-2781.

[15] I. Zabir, S. Paul, M. A. Rayhan, T. Sarker, S. A. Fattah, and C. Shahnaz, “Automatic brain tumor detection and segmentation from multi-modal MRI images based on region growing and level set evolution,” in 2015 IEEE International WIE Conference on Electrical and Computer Engineering, WIECON-ECE 2015, Mar. 2016, pp. 503–506, doi: 10.1109/WIECON-ECE.2015.7443979.

[16] E. S. A. El-Dahshan, T. Hosny, and A. B. M. Salem, “Hybrid intelligent techniques for MRI brain images classification,” Digit. Signal Process. A Rev. J., vol. 20, no. 2, pp. 433–441, 2010, doi: 10.1016/j.dsp.2009.07.002.

[17] F. G. Zöllner, K. E. Emblem, and L. R. Schad, “SVM-based glioma grading: Optimization by feature reduction analysis,” Z. Med. Phys., vol. 22, no. 3, pp. 205–214, Sep. 2012, doi: 10.1016/j.zemedi.2012.03.007.

[18] E. Zanaty, “Determination of Gray Matter (GM) and White Matter (WM) Volume in Brain Magnetic Resonance Images (MRI),” Int. J. Comput. Appl., vol. 45, no. 3, pp. 16–22, 2012, [Online]. Available: https://pdfs.semanticscholar.org/1757/d8a0e6292835eb598bc9c40faca7b71079822.pdf.

[19] W. Cui, Y. Wang, Y. Fan, Y. Feng, and T. Lei, “Localized FCM clustering with spatial information for medical image segmentation and bias field estimation,” Int. J. Biomed. Imaging, vol. 2013, 2013, doi: 10.1155/2013/930301.
J. Suchdeva, V. Kumar, I. Gupta, N. Khandelwal, and C. K. Ahuja, “Segmentation, feature extraction, and multiclass brain tumor classification,” *J. Digit. Imaging*, vol. 26, no. 6, pp. 1141–1150, 2013, doi: 10.1007/s10278-013-9600-0.

E. A. S. El-Dahshan, H. M. Mohsen, K. Revett, and A. B. M. Salem, “Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm,” *Expert Syst. Appl.*, vol. 41, no. 11, pp. 5526–5545, 2014, doi: 10.1016/j.eswa.2014.01.021.

P. Kumar and B. Vijayakumar, “Brain Tumour Mr Image Segmentation and Classification Using by PCA and RBF Kernel Based Support Vector Machine,” *Middle-East J. Sci. Res.*, vol. 23, no. 9, pp. 2106–2116, 2015, doi: 10.5829/idosi.mejer.2015.23.09.22458.

S. Damodharan and D. Raghavan, “Combining tissue segmentation and neural network for brain tumor detection,” *Int. Arab J. Inf. Technol.*, vol. 12, no. 1, pp. 42–52, 2015.

V. Anitha and S. Murugavalli, “Brain tumour classification using two-tier classifier with adaptive segmentation technique,” *IET Comput. Vis.*, vol. 10, no. 1, pp. 9–17, 2016, doi: 10.1049/iet-cvi.2014.0193.

N. B. Bahadure, A. K. Ray, and H. P. Thethi, “Image Analysis for MRI Based Brain Tumor Detection and Feature Extraction Using Biologically Inspired BWT and SVM,” *Int. J. Biomed. Imaging*, vol. 2017, 2017, doi: 10.1155/2017/9749108.

G. Gilanie, U. I. Bajwa, M. M. Waraich, Z. Habib, H. Ullah, and M. Nasir, “Classification of normal and abnormal brain MRI slices using Gabor texture and support vector machines,” *Signal, Image Video Process.*, vol. 12, no. 3, pp. 479–487, 2018, doi: 10.1007/s11760-017-1182-8.

F. Özyurt, E. Sert, E. Avcı, and E. Dogantekin, “Brain tumor detection based on Convolutional Neural Network with neutrosophic expert maximum fuzzy sure entropy,” *Meas. J. Int. Meas. Confed.*, vol. 147, 2019, doi: 10.1016/j.measurement.2019.07.058.

A. I., M. Abu, and I. El-Henawy, “A Feature Selection Algorithm based on Mutual Information using Local Non-uniformity Correction Estimator,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 6, 2017, doi: 10.14569/ijacs.2017.080656.

X. Wang, B. Guo, Y. Shen, C. Zhou, and X. Duan, “Input Feature Selection Method Based on Feature Set Equivalence and Mutual Information Gain Maximization,” *IEEE Access*, vol. 7, pp. 15125–151538, 2019, doi: 10.1109/ACCESS.2019.2948095.

A. El Akadi, A. El Ouardighi, and D. Aboutajdine, “A Powerful Feature Selection approach based on Mutual Information,” *Int. J. Comput. Sci. Netw. Secur.*, vol. 8, no. 4, pp. 116–121, 2008, [Online]. Available: http://paper.ijcsns.org/07_book/200804/20080417.pdf.

S. Verron, T. Tiplica, and A. Kobi, “Fault detection and identification with a new feature selection based on mutual information,” *J. Process Control*, vol. 18, no. 5, pp. 479–490, 2008, doi: 10.1016/j.jprocont.2007.08.003.

J. Huang, Y. Cai, and X. Xu, “A hybrid genetic algorithm for feature selection wrapper based on mutual information,” *Pattern Recognit. Lett.*, vol. 28, no. 13, pp. 1825–1844, 2007, doi: 10.1016/j.patrec.2007.05.011.

M. Al-Azawi, “Saliency-Based Image Retrieval using Colour Histogram Feature,” *IEEE Int. Conf. Electr. Electron. Comput. Mech. Comput. EECCMC*, 2018.

M. Al-Azawi, “Computational Intelligence-based Semantic Image Background Identification using Colour-Texture Feature,” *Int. J. Comput. Appl.*, vol. 180, no. 10, pp. 27–31, Jan. 2018, doi: 10.5120/ijca2018916165.

M. S. Nixon and A. S. Aguado, “Introduction to texture description, segmentation, and classification,” *Feat. Extr. Image Process. Comput. Vis.*, pp. 399–434, 2012, doi: 10.1016/B978-0-12-396549-3.00008-2.
patterns and histogram orientation gradient,” *Neurocomputing*, vol. 219, pp. 526–535, 2017, doi: 10.1016/j.neucom.2016.09.051.

[38] R. N. Khushaba, A. Al-Ani, A. Alsukker, and A. Al-Jumaily, “A combined ant colony and differential evolution feature selection algorithm,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5217 LNCS, pp. 1–12, 2008, doi: 10.1007/978-3-540-87527-7_1.

[39] P. J. Kostelec and S. Periaswamy, “Image Registration for MRI,” *Mod. Signal Process.*., vol. 46, pp. 161–184, 2003.