Automated Glioma Grading based on an Efficient Ensemble Design of a Multiple Classifier System using Deep Iteration Neural Networks Matrix

[1,2]**Ahmed Al-Zurfi**, [1]**Farid Meziane and** [1]**Rob Aspin**

[1]School of Computing, Science and Engineering, University of Salford, UK

[2] Department of Electronic and Communications, Faculty of Engineering, University of Kufa, Iraq

a.n.alzurfi@edu.salford.ac.uk, f.meziane@salford.ac.uk, r.aspin@salford.ac.uk

Abstract— the preoperative diagnosis of brain Glioma grades is crucial for therapeutic planning as it impacts on the tumour's prognosis. The development of machine learning methods that can accurately evaluate Glioma grades is of great interest since it is a repeatable and reliable diagnosis procedure. Moreover, the classification accuracy of a single classifier can be further improved by using the ensemble of different classifiers. In this paper, a new strategy has been developed, which uses a deep neural network incorporating an extensive iteration matrix based on the combination of eleven different machine learning algorithms. The classification system is evaluated using a cross-validation technique, to add more generalization to the results of the classification system's reliability in unseen cases. Experimental results indicate that, when compared to both the single classification model, and the majority vote scheme, the grading accuracy has significantly improved using our proposed approach. The obtained sensitivity, specificity and accuracy are 100%, 90% and 93.3% respectively. The proposed approach has improved upon the highest accuracy of the single classification model by 13.3%. The proposed classification system presents an efficient method to evaluate the malignancy level of Glioma with more reliable and accurate clinical outcomes.

Keywords: Magnetic Resonance Imaging; Multiple Classifier System; Deep Neural Network; Glioma Grading, Ensemble Design

I. INTRODUCTION

Brain cancer is life-threatening, with Glioma being the most common type accounting for 70% of primary brain tumours in the UK adult population [1]. According to the World Health Organization (WHO) classification of brain tumour, Gliomas can be sub-classified based on the degree of malignancy into four grades (I, II, III, and IV). Grade IV, is extremely malignant, has a poor prognosis and a high mortality rate. Therefore, it is essential to discriminate grade IV from the lower Glioma grades preoperatively, as this impacts on the prognosis and treatment of the patients [2].

The ability of medical imaging techniques to create visual images of internal organs and tissues of the human body is used effectively to support clinicians in faster and more robust diagnosis and in informing treatment planning. Specifically, the diagnosis of brain tumours is determined by the aid of medical imaging techniques. Computer-aided diagnosis using image features extracted from brain tumours, which has been

presented as a complementary approach to improve upon radiological diagnosis performance [3].

Recent studies for the classification of Glioma grades used a combination of conventional MRI and/or advanced imaging modalities [4]. However, advanced MRI techniques as opposed to the conventional methods, require more expensive equipment, more experience, and relatively more time to extract the tumour's attributes. It also has limited availability in MRI clinics. Consequently, our work is based on only one conventional MRI modality, that is, T2-weighted MRI, which is a non-invasive imaging technique and the tumour is more hyperintense compared to the other brain tissue using this modality.

Although many studies have developed classifications of brain tumours using a single classification model, the fusion of different classifiers would lead to superior classification accuracy [5]. Therefore, to achieve further improvement in the classification accuracy for Glioma grades we investigated the use of a multiple classifier system that combines several machine learning algorithms in one grading system.

Glioma tumours have different levels of heterogeneity. Image texture analysis has been proven to be an efficient and objective method to measure the tumour heterogeneity [6]. Furthermore, image texture analysis plays a significant role in the recognition and identification of morphological characteristics of brain tumours [7]. Among different texture feature methods, Grey Level Co-occurrence Matrix (GLCM) has been commonly used in various applications [8, 9]. Specifically, GLCM has shown remarkable results in the evaluation of the malignancy level of brain tumours [10].

The rapid development of machine learning has played a vital part in the classification and prediction of many cancers, specifically brain tumours. Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Linear and Quadratic Discriminant Analysis (LDA and QDA respectively) and Decision Tree (DT) are common machine learning algorithms that are used for the classification of brain tumours. For instance, a comparative study was conducted on different classification methods for Glioma grading based on the single classification approach [11]. However, an approach that takes advantage of the combination of multiple machine learning algorithms would lead to an improvement in the classification performance. Furthermore, in the literature, different methods have been proposed to integrate multi-classification models to enhance the accuracy of the classification. For example, the bagging approach [12] where the output decisions that come

from several decision trees are combined and the outcome is determined by majority votes. Boosting, on the other hand, adjusts the input data distribution based on the outcomes from the previously trained classifiers, then a weighted majority vote is used to generate the final decision [13]. The majority voting has been applied widely to fuse multi-classification models [14], used with Multiple Classifier Systems (MCS). However, it has a limited ability to sense the complex relationships of information among different classifiers.

Using a learning strategy in the fusion stage of MCS is a much more powerful method. Neural Network (NN) is an efficient approach and is of great interest due to its ability to automatically uncover the nonlinear relationship of different data distributions. Deep learning can be designed based on Deep Neural Networks (DNN) with multi hidden layers of nonlinear information processing that can learn complex data pattern [15]. Hence, to obtain a better representation of the relationship among many classifiers, we developed DNN based MCS to stack multiple machine learning algorithms.

It is necessary with the application of backpropagation neural network to find the optimal divergence point which maximizes its accuracy. Indeed, at present, no such method gives a general solution to overcome this issue. Therefore, we developed a systematic deep iteration matrix of NN (DINN) to improve the classification accuracy of MCS for Glioma grades further.

In this paper, we proposed a novel strategy for developing a multiple classifier system to improve the classification accuracy of Glioma grades. This is based on the development of two stages of the classification system. The first stage is determined using eleven classifiers, and these are, DT, LDA, QDA, SVM (with three different kernels), and KNN (with five different designs). These classifiers are trained individually based on the co-occurrence textural features extracted from T2-weighted MRI modality to discriminate Glioma grade IV (Glioblastoma Multiforme) from the lowest Glioma grades (grade II and III). Then, in the second stage, an efficient method of combining all these classifiers is designed. The fusion stage is developed based on deep neural network incorporating an extensive iteration matrix. The difference in performance between the single classification model and the proposed MCS is analyzed in terms of classification accuracy, sensitivity, specificity. The performance is also investigated as compared to other traditional MCS based on the majority voting. The experimental outputs indicate that the proposed DINN-MCS is better than any other existing method in term of the performance measures. The dataset used for this work is downloaded from the cancer archive collection website [16]. The T2- weighted MRI of brain Glioma tumours of thirty patients, are used for this work. It is publicly available and confirmed with a histology diagnosis. This collection contains ten patients with grade IV (Glioblastomas) and twenty patients with lower Glioma grades (III and II). Each patient has a various number of slices ranging from 20 to 120, with varying post imaging parameters such as different gap space and slice thickness ranging from (2 to 7.5 mm).

II. OVERVIEW OF THE CLASSIFICATION MODELS

A brief description for each classifier used in this work is discussed starting with Linear and Quadratic Discriminant Analysis. These classifiers are parametric classifiers, and they are affected by the data distribution, thereby performing poorly when there is much overlap between classes of data [17]. The Decision Tree (DT) algorithm is the second common classifier, defined as tree-like structures, in which the neurons represent the input features and each branch indicates a different output decision [18]. DT offers many advantages. It is a non-parametric approach and is not dependent on the input data distribution. Also, it can deal with non-linear relationships between class labels and features. The main limitation of DT is that it can be subject to underfitting or overfitting, especially when using a small dataset [19]. The other popular classification algorithm is Support Vector Machine; it produces promising results in many applications [20]. It identifies a hyperplane that separates input data into two classes. The input data is mapped into a new features space with higher dimensionality. The major limitation of SVM is associated with the selection of the kernel function [21]. Finally, K-Nearest Neighbour is also used widely in different applications and showed superior results. KNN was designed based on searching for the close similarity of k-samples between a test sample and other instances that have similar behavior [22]. This searching process is guided by a distance function such as the Euclidean and the Camberra function. KNN algorithm is highly sensitive to noise as well as to the choice of similarity metric that is used to compare instances [18]. To conclude, there is no single classification model that can outperform all other classifiers as each one has a limitation. Hence, the integration of multiple classifiers can overcome the drawback of weak classifiers and lead to a more complementary approach. As a result, this would lead to greater accuracy than using single classification models.

III. PREPARATION OF MRI TEXTURAL FEATURES

The first step in this work is the segmentation of the regions of interest of the brain tumour using T2-weighted MR slices. The segmentation step is conducted based on an automatic segmentation system [23]. Then a cropping process is applied. After that, normalization is determined for the MRI grey levels based on mapping the original intensities of the MRI into the range (0-255). Textural features are extracted based on the grey level co-occurrence matrix using four angles ($\theta = 0\circ$, $45\circ$, $90\circ$, $135\circ$), and distance equal to one. Eighteen textural predicators incorporated with the GLCM are extracted as follows: autocorrelation, correlation, cluster prominence, contrast, cluster shade, dissimilarity, energy, entropy, maximum probability, homogeneity, sum of squares, sum average, sum variance, sum entropy, information measure of correlation 1, information measure of correlation 2, inverse difference normalized and inverse difference moment normalized [24]. Four directions with eighteen predictors are used. Hence, each patient is represented by seventy-two textural features. These features are used later in the classification stage to discriminate Glioma grades.

IV. FIRST STAGE: SINGLE CLASSIFIER SYSTEM

The first stage of the proposed MCS is providing the prepared MR image-textural features into different classification models. Eleven popular classifiers are evaluated to discriminate between grade IV and the lowest Glioma grades. We conducted DT, LDA, QDA, SVM with three kernels (linear, quadratic, and medium Gaussian), and KNN with five different design as presented in more detail in Table 1. The rationale behind applying different kernels and parameters with SVM and KNN is to increase the diversity of the output decisions produced at this stage. Leave-one-out cross-validation technique (LOO) is employed to avoid the overfitting problem, and to add more generalization to the classification system. All samples (N samples) are divided into training (N-1) and testing sample. The last step of this stage is the evaluation of the output results using different matrices. These metrics are sensitivity, specificity, and classification accuracy [25] which are used to evaluate the classifier accuracy for every single classifier. The parameters setup for different classifiers that are undertaken in this work are selected experimentally during the learning phase and are illustrated in Table 1.

Table 1 Experimental Parameters for the classifiers used in this work

Classifier Name	Parameters
DT	Maximum number of split is 4. Split criterion is Gini's Diversity Index. Maximum surrogate per node is 10.
LDA	Regularization is Diagonal covariance
QDA	Regularization is Diagonal covariance
SVML	Kernel function is linear
SVMQ	Kernel function is quadratic
SVMMG	Kernel function is medium Gaussian. kernel scale is 8.5
KNNF	Number of k-neighbours is 1 Distance metric is Euclidean Distance weight is identical
KNNM	Number of k-neighbours is 10 Distance metric is Euclidean Distance weight is identical
KNNCOS	Number of k-neighbours is 10 Distance metric is cosine Distance weight is identical
KNNCUB	Number of k-neighbours is 10 Distance metric is cubic Distance weight is identical
KNNW	Number of k-neighbours is 10 Distance metric is Euclidean Distance weight is squared inverse

V. SECOND STAGE: MULTIPLE CLASSIFIER SYSTEM

The second stage of the proposed MCS is the ensemble of the classifiers trained previously in the first stage. A binary classification to the output decisions vector is conducted. This is performed by supplying the output decision vector produced from each single classification model to the DNN, where all samples in this vector are passed through the training and testing phase using the LOO cross-validation technique.

VI. PROPOSED ENSEMBLE DESIGN

The proposed fusion stage is built based on the integration of multiple classification models using the back-propagation neural network incorporating a deep iteration matrix. All classifiers are integrated into one MCS system. The hyperbolic tangent sigmoid transfer function (tansig) is commonly used due to its high accuracy and faster divergence response. Therefore, it is selected to be the activation function for all neurons in the NN design. The formula for this function is defined by Equation (1) [26].

$$\tanh(z) = \frac{1-e^{-2z}}{1+e^{-2z}} \tag{1}$$

A backpropagation strategy is used to optimize the NN performance. Scaled conjugate gradient backpropagation (trainscg) is one of the most commonly applied functions with NN and has produced efficient results [27]. Therefore, this technique is selected to be the learning function of the design. We used the DINN matrix to evaluate the most significant design that reveals the maximum accuracy. In this matrix, we have examined a wide range of different numbers of neurons, and hidden layers, in addition to using a different structure of backpropagation neural network (BPNN) in each iteration based on testing a wide range of different values of weights and biases of Neural Networks.

Let I_i represent iterations of NN, where $i = 1, 2, …, E$; and E is a selected endpoint of iterations, R_j represent rounds of NN, where $j = 1, 2, …, N$; and N is the total number of samples. Let n is the number of neurons per layer L, and P is a probability which used to measure the output of BPNN for each cell of the deep iteration neural network matrix A as shown in Equation (2).

$$A(i,j) = P((i,j) \mid n, L, R_j, I_i) \tag{2}$$

Here, I_i, R_j represent columns and rows of matrix A respectively.

There are thirty rounds using the LOO technique to cover all N samples given to the classification system. In each round, the dataset is separated into different sets of training, validation and testing sets. The significant difference between each structure of BPNN in this matrix is of using different validation set which selected randomly from the original dataset and not including neither the training phase nor the testing phase. Each design of the DINN matrix should be self-optimized using the learning function. In this work, twenty-five patients are selected to be in the training phase and four samples in the validation set and one sample in the testing phase. This arrangement is re-rounded in each cell of the matrix to examine the behavior of each unique design of BPNN. When implementing this matrix, it is necessary to ensure that the cross-validation is fully controlled by applying a completely different dataset in each of the training, validation, and testing phases. In this work, both n, and L are in the range of $(1, 2, …, 30)$. All testing designs for the same iteration I_i should have the same network structure. After the calculation of DINN matrix is completed, the confusion matrix is measured by comparing the output results of each

column of the DINN matrix with the true class label. The results are then ranked to select the best model that shows the highest accuracy. The total number of iterations of DINN matrix can be calculated by $R_j \times I_i \times n \, per \, L$.

VII. EXPERIMENTAL RESULTS

The results of evaluating one-layer BPNN which shows the classification accuracy with respect to each number of neurons per layer is given in Figure 1. It can be seen in this Figure that the highest classification accuracy is achieved at 90%, by using 23 neurons in the layer. The results of sensitivity and specificity are 100% and 85% respectively. The confusion metrics are as follows: True positive (TP) = 10, False negative (FN) = 0, True negative (TN) =17, False positive (FP) = 3. Figure 2 reveals the evaluation results of classification accuracy with respect to different iterations sequence number using 23 neurons in the one-layer. Furthermore, it is found that the highest accuracy is produced by the second iteration.

Further investigation is conducted by evaluating DINN matrix using two-layer BPNN. Figure 3 presents the results of classification accuracy with respect to using a different number of neurons in the first and second layer. The results of this experiment reveal that the classification accuracy is improved at 93.33% using 26 neurons, and 16 neurons in the first and second layer respectively as shown in Figure 3. The results of sensitivity and specificity are also improved to 100%, 90% respectively. The confusion metrics are as follows TP = 10, FN=0, TN=18, FP=2. In this experiment, it is noticed that the accuracy using the second iteration (I_2) of DINN matrix is outperforming the other designs.

Ultimately, it is found that the results of evaluating of three-layer BPNN incorporating a DINN matrix, for all tested network designs, has shown less accuracy as compared to the two-layer BPNN performance. Therefore, we stopped further investigation at this stage.

The overall results comparison in terms of classification accuracy, sensitivity, specificity are shown in Figures 4, 5, and 6 respectively. These results show different evaluation methods for the classification of Glioma grades using both approaches: a single classification model, and the multiple classifier system. These results also present a comparison of the classification performance of the traditional MCS based on the majority vote and that of our proposed approach. The classification results illustrate that our proposed system significantly outperforms the other traditional algorithms undertaken in this work and raising the accuracy to 93.3%.

VIII. DISCUSSION

A novel approach has been proposed for the development of a multiple classifier system for Glioma grading. Three contributions have been addressed in this article, as follows:

First: a significant improvement in the results is achieved using MCS based on the proposed approach as compared to using only a single classification system. The classification accuracy is boosted up to 93.3% as shown in Figures 4, 5, and 6.
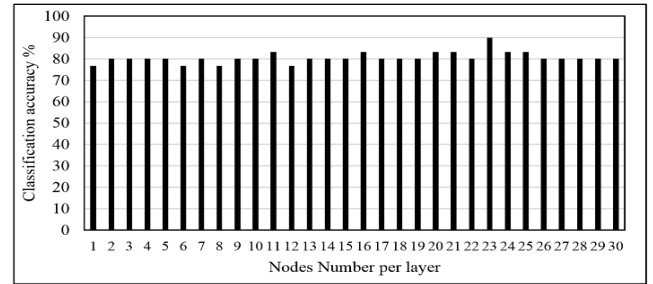

Figure 1 Classification accuracy results using one-layer NN corresponding to the number of nodes per the layer.
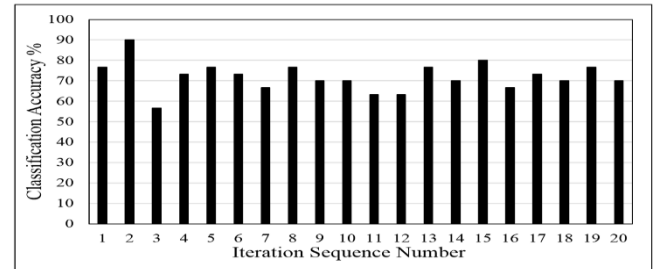

Figure 2 Classification accuracy results of the one-layer NN corresponding to different iterations using the 23 nodes per the layer.
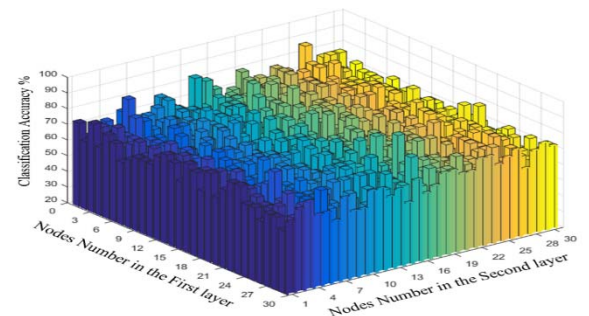

Figure 3 Classification accuracy results using two-layer BPNN corresponding to number of nodes per the first and second layer in the left and right axis respectively.
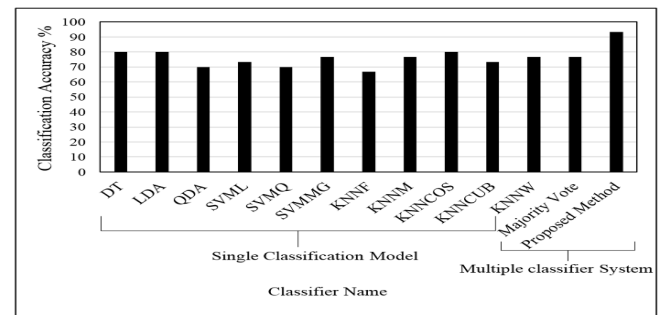

Figure 4 Comparative results in terms of the classification accuracy of the proposed method against all other single classification models and the majority vote.

Second: A comparison between the proposed system and the existing MCS using the majority vote has been investigated. These results have proved that the proposed approach demonstrates superior accuracy compared to the other classification system as shown in Figure 4, 5, and 6.

Third: the results confirmed that the deep iteration neural network matrix has the highest impact on DNN and enable it

to achieve its highest accuracy. Unlike the existing approaches which have used only a few trials, randomly selected. Moreover, the proposed DINN matrix is a systematic method to investigate all possible results and enable the highest accuracy in the MCS.

It has been shown that using DNN based on the fusion of multiple classifiers leads to improved classification accuracy for brain Glioma grades. A wide range of deferent designs of BPNN is investigated for the proposed MCS. It is found that adding more hidden layers would lead to enhanced classification accuracy. However, adding more than two hidden layers, reduced the classification accuracy. The additional layer increased the complexity of the network that raised the probability of overfitting which leads to degraded classification accuracy. Therefore, we ceased further evaluation for testing other hidden layers. There are different points of divergence that comes from testing different validation sets and initial weights, examined within various iterations of DNN. Hence, different network weights are produced in each iteration. Therefore, each iteration provides different results. The purpose of determining a considerable number of deferent designs of DNN with varying sets of validation is to find the optimal divergent point that is closer to the global minimum value, which represents, in other words, the best possible design of DNN that would reveal the highest classification accuracy. However, this will require a longer computation time in the learning phase.

The dataset set used in this work is relatively small, and this could have degraded the accuracy of the neural network. However, using deep BPNN incorporating the proposed iteration matrix has been proven in this work to overcome this limitation and shows remarkable results. Indeed, at present, it is a significant challenge to acquire a large image dataset of Glioma grades with the approval of histopathology test.
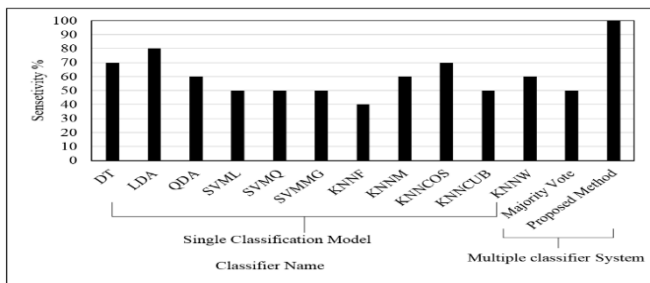

Figure 5 Comparative results in terms of the sensitivity of the proposed method against all other single classification models and the majority vote.
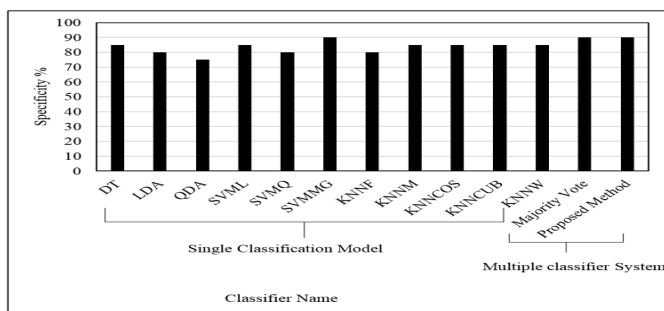

Figure 6 Comparative results in terms of the specificity of the proposed method against all other single classification models and the majority vote.

This approval is vital to have a solid confirmation of the malignancy grade of Glioma, thereby providing real validity of the evaluation and validation process to any classification or grading system. Consequently, the proposed classification system is the more suitable approach for a small dataset, achieved a promising result to improve the classification accuracy of Glioma grades. Also, it can be considered critically that the proposed system is an alternative method to the traditional approach of deep learning which usually requires an enormous dataset and greater intensive computation time. A significant improvement is obtained in the results of classification of MR images of Glioma grades using the proposed DINN-MCS. Moreover, the proposed system has the advantage of outperforming both the single classification method and the existing MCS based on the majority vote. However, it requires a high computation time, advanced computing hardware, and parallel computing where it is possible to reduce the execution time.

IX.   CONCLUSION

We proposed two stages of a novel multiple classification system for automated Glioma grading. The classification system aimed to discriminate grade IV from the lower Glioma grades. In the first stage, we tested eleven classification models individually. Each of these classification models is trained based on the co-occurrence of textural features extracted from T2-weighted MRI modality. In the second stage, DNN received all decision vectors as input, and then they are divided into training, validation, and testing sets. The performance is evaluated in terms of classification accuracy, sensitivity, and specificity. We investigated different approaches: a single classification method, majority vote based MCS, and our proposed method in terms of these performance measures. The Leave-One-Out cross-validation technique is adopted in all stages to add more generalization for the classification system. The results illustrated that the proposed system achieved a superior classification accuracy at 93.3% as compared to the existing methods. The experimental output revealed that the deep neural network incorporating the DINN matrix is an efficient approach to stack a multi-classification model. The proposed method further improved the final classification accuracy. Hence, enabling the MCS to achieve better discrimination accuracy for Glioma grades. As for future trends, experiments can be repeated with different dataset for possible enhancement in the classification accuracy.

Reference

[1]   B. Stefan, W. Roland, P. N. Lutz, and R. Mauricio, "A survey of MRI-based medical image analysis for brain tumor studies," *Physics in Medicine and Biology,* vol. 58, p. R97, 2013.

[2]   B. J. Theeler and M. D. Groves, "High-Grade Gliomas," *Current Treatment Options in Neurology,* vol. 13, pp. 386-399, 2011.

[3]   E.-S. A. 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 Systems with Applications,* vol. 41, pp. 5526-5545, 9/1/ 2014.

[4] E. I. Zacharaki, S. Wang, S. Chawla, D. Soo Yoo, R. Wolf, E. R. Melhem*, et al.*, "Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme," *Magnetic Resonance in Medicine,* vol. 62, pp. 1609-1618, Dec 2009.

[5] M. Woźniak, M. Graña, and E. Corchado, "A survey of multiple classifier systems as hybrid systems," *Information Fusion,* vol. 16, pp. 3-17, 2014.

[6] Y. J. Ryu, S. H. Choi, S. J. Park, T. J. Yun, J.-H. Kim, and C.-H. Sohn, "Glioma: Application of Whole-Tumor Texture Analysis of Diffusion-Weighted Imaging for the Evaluation of Tumor Heterogeneity," 2014.

[7] S. Roy, S. Nag, I. K. Maitra, and S. K. Bandyopadhyay, "A Review on Automated Brain Tumor Detection and Segmentation from MRI of Brain," *arXiv preprint arXiv:1312.6150,* 2013.

[8] M. Yazdi, Z. Adelpour, B. Bahraini, and Y. K. Jahromi, "Novel ridge orientation based approach for fingerprint identification using co-occurrence matrix," in *Proc. of World Academy of Science, Engineering and Technology*, 2007, pp. 371-375.

[9] H. Ahmed, S. Mashohor, and M. I. Saripan, "A texture-based approach for content based image retrieval system for plant leaves images," in *2011 IEEE 7th International Colloquium on Signal Processing and its Applications*, 2011, pp. 11-14.

[10] A. Larroza, V. Bodí, and D. Moratal, "Texture Analysis in Magnetic Resonance Imaging: Review and Considerations for Future Applications," in *Assessment of Cellular and Organ Function and Dysfunction using Direct and Derived MRI Methodologies*, ed: InTech, 2016.

[11] X. Zhang, L.-F. Yan, Y.-C. Hu, G. Li, Y. Yang, Y. Han*, et al.*, "Optimizing a machine learning based glioma grading system using multi-parametric MRI histogram and texture features," *Oncotarget,* vol. 8, pp. 47816-47830, 2017.

[12] L. Breiman, "Bagging predictors," *Machine Learning,* vol. 24, pp. 123-140, 1996/08/01 1996.

[13] Y. Freund, "Boosting a Weak Learning Algorithm by Majority," *Information and Computation,* vol. 121, pp. 256-285, 1995/09/01/ 1995.

[14] L. Xu, A. Krzyzak, and C. Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," *IEEE Transactions on Systems, Man, and Cybernetics,* vol. 22, pp. 418-435, 1992.

[15] P. Mamoshina, A. Vieira, E. Putin, and A. Zhavoronkov, "Applications of Deep Learning in Biomedicine," *Molecular Pharmaceutics,* vol. 13, pp. 1445-1454, 2016/05/02 2016.

[16] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel*, et al.*, "The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository," *Journal of Digital Imaging,* vol. 26, pp. 1045-1057, 2013/12/01 2013.

[17] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Regularization studies of linear discriminant analysis in small sample size scenarios with application to face recognition," *Pattern Recognition Letters,* vol. 26, pp. 181-191, 2005/01/15/ 2005.

[18] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," in *Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in EHealth, HCI, Information Retrieval and Pervasive Technologie*, ed, 2007.

[19] Y.-y. Song and Y. Lu, "Decision tree methods: applications for classification and prediction," *Shanghai Archives of Psychiatry,* vol. 27, pp. 130-135, 2015.

[20] W. Dubitzky, M. Granzow, and D. P. Berrar, *Fundamentals of data mining in genomics and proteomics*: Springer Science & Business Media, 2007.

[21] G. L. Prajapati and A. Patle, "On Performing Classification Using SVM with Radial Basis and Polynomial Kernel Functions," in *2010 3rd International Conference on Emerging Trends in Engineering and Technology*, 2010, pp. 512-515.

[22] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory,* vol. 13, pp. 21-27, 1967.

[23] A. Hasan, F. Meziane, R. Aspin, and H. Jalab, "Segmentation of Brain Tumors in MRI Images Using Three-Dimensional Active Contour without Edge," *Symmetry,* vol. 8, p. 132, 2016.

[24] W. Gómez, W. Pereira, and A. F. C. Infantosi, "Analysis of co-occurrence texture statistics as a function of gray-level quantization for classifying breast ultrasound," *IEEE transactions on medical imaging,* vol. 31, pp. 1889-1899, 2012.

[25] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical machine learning tools and techniques*: Morgan Kaufmann, 2016.

[26] G. Daniel, *Principles of artificial neural networks* vol. 7: World Scientific, 2013.

[27] F. D. Baptista, S. Rodrigues, and F. Morgado-Dias, "Performance comparison of ANN training algorithms for classification," in *Intelligent Signal Processing (WISP), 2013 IEEE 8th International Symposium on*, 2013, pp. 115-120.