Classification of brain tumors using the multilayer perceptron artificial neural network

Raid Adnan Omar, Jassim Mohammed Najim, Imad H. Abood

Department of Physics, College of Science, Anbar University

E-mail: raiedadnan_m@yahoo.com

Abstract

Information from 54 Magnetic Resonance Imaging (MRI) brain tumor images (27 benign and 27 malignant) were collected and subjected to multilayer perceptron artificial neural network available on the well known software of IBM SPSS 17 (Statistical Package for the Social Sciences). After many attempts, automatic architecture was decided to be adopted in this research work. Thirteen shape and statistical characteristics of images were considered. The neural network revealed an 89.1% of correct classification for the training sample and 100% of correct classification for the test sample. The normalized importance of the considered characteristics showed that kurtosis accounted for 100% which means that this variable has a substantial effect on how the network perform when predicting cases of brain tumor, contrast accounted for 64.3%, correlation accounted for 56.7%, and entropy accounted for 54.8%. All remaining characteristics accounted for 21.3-46.8% of normalized importance. The output of the neural networks showed that sensitivity and specificity were scored remarkably high level of probability as it approached % 96.

Key words

Brain Tumor, IBM SPSS Ver. 17 & 20, Artificial Neural Networks (ANN).

Article info.

Received: Aug. 2017
Accepted: Oct. 2017
Published: Mar. 2018

Tصنيف أورام الدماغ باستخدام الشبكة العصبية الاصطناعية لمستقبلات متعذدة الطبقات

زهاب عدنان عمر، جاسم محمد نجم، عيد هجول عبود

قسم الفيزياء، كلية العلوم، جامعة الإبتار

الخلاصة

جمع معلومات من 54 صورة لورم الدماغ من صور جهاز جيتز المغناطيسي (27 صورة لورم الدماغ الحميد و 27 صورة لورم الدماغ الخبيث). و عرضت هذه المعلومات من شبكة عصبية متعددة العقاقير متعددة الطبقات متوفرة على البرنامج الاصطناعي IBM SPSS 17. بعد محاولات عدة تم اختيار معمارية شبكة الاصناب الاصطناعية الدورية لعلوم هذا البحث. ثلاثة عشر معلومة عن العروض الاصنابية و الشكلية للصورة تم اعتمادها في هذا البحث. لقد أظهرت شبكة الاصناب المستخدمة نسب تصنيف صحيح في مجموعة التدريب بلغت 89.1% بينما بلغت 100% في مجموعة الاختبار. لقد اظهرت الاصناب الاصنابية المطلوبة ان التفرط كان المعلومات الأكثر أهمية (kurtosis) في التعريف بين نوعي ورم الدماغ. و هذا يؤكّر الاصناب الاصنابية (Entropy) المجهول لهذه المعلومة في التنبؤ بنوع الورم للحالات الجديدة من خلال معلومات صورة الورم. لقد كان لتباين (Contrast) الاصناب الاصنابية طبيعة 64.3% و للاقتراب (Correlation) الاصناب الاصنابية طبيعة 56.7% و للاقتراب الاصنابية طبيعة 54.8%. و الذي هو عبارة عن مؤشرات احصائية تتهم تقييم نسب الاصناب. لقد تراوحت الاصناب الاصنابية مداه بين 21.3% و 46.8% بنسب التنبؤ المطلوبة. و حالة الورم الخبيث (sensitivity) و حالة الورم الحميد (specificity) تصل إلى نفس المستوى.
Introduction

Brain tumor is a medical issue that harvest according to the American brain tumor association [1] thousands of lives every year. In the United States of America the number of new cases with brain tumor is expected to be 27110 in this 2017. It is also expected that 17000 will lose their lives during this year. In this context one can expect how significant is this issue when bearing in mind that thousands of people all over the world will be the victims of this disease. Two main types of brain tumors can be identified, these are; benign and malignant [2]. Benign tumor can be defined as a mass that is very slowly growing with well-defined borders and rarely spread or invade neighbor cells [2]. Malignant tumor can be defined as a mass that is invading neighbor cells and growing fast and is life threatening [2]. Epidemiological studies in different parts of the world showed that brain and central nervous system tumors are relatively low compared to other cancer types [3]. In Austria [4] the country that has developed many studies on brain tumors, the incidence rate of brain tumors is 18.1 per 100000 person/year. Brain tumors issue was and still the main concern of many researchers as well as many health organizations. The field of image processing witnesses many development in both segmentation and classification of brain tumors. Rajesh and Bhalchanda [5], used MATLAB to extract brain tumors from MRI images. Suhag and Mohan [6], used Support Vector Machines (SVM) classifier to detect and classify brain tumors from MRI images. Hassan and Aboshgifa [7], designed MATLAB GUI for the detection of brain tumors from MRI images. Saini and Singh [8] Singh used MATLAB image processing for the detection of brain tumors.

Artificial Neural Networks (ANN)

Artificial Neural Networks (ANN) was widely used in the process of detecting and classification of brain tumors. Monica Subashini and Sarat Kumar used ANN in the detection of brain tumors from MRI images [9]. Pulse coupled neural network was used to enhance MRI images before segmentation. The classification was to detect either normal or abnormal images depending on the performance of the network. The abnormal images are not classified to what this abnormal mass refers to, rather the issue left to further investigation under the aid of special medical support. In addition to the investigation of the MRI images, shape and texture characteristics of the images were also subjected to different statistical techniques in order to provide valuable information about characteristics that best describe the type of tumor. In this context Neelam Marshkole et al. [10], used texture and shape features of MRI images to classify brain tumors to either malignant or benign using linear vector quantization technique. They found these features very effective in the process of classification. In this paper the features that previously used by Zhang et al. [11] in an image processing program to classify two states of brain tumors (benign and malignant) were used here in order to see how good is the performance of ANN to enhance classification.

Preparing methods

The stage of introducing MRI brain images into the algorithm of the work of Zhang et al. [11] build it in MATLAB Ver. 2014a program which used the principal component analysis method (PCA) to reduce the output dimensions and using (K SVM) kernel and used the GRB kernel achieves to do classification, then the matrix output units produced by this
algorithm, which including Mean, Standard Deviation, Entropy, Root Means Square (RMS), Variance, Smoothness, Kurtosis, Skewness, Inverse Difference Movement (IDM), Contrast, Correlation, Energy and Homogeneity, as shown in Table 1.

Table 1: Statistical characteristics of images [11].

| Name       | Mean | Standard Dev. | Entropy | RMS | Variance | Smoothness | Kurtosis | Skewness | IDM | Correlation | Energy | Homogeneity | FBF Ac/L | Polys. Quad. | Type Tumor | Diagnose name          |
|------------|------|---------------|---------|-----|----------|------------|----------|----------|-----|-------------|--------|-------------|---------|--------------|-----------|------------------------|
|            |      |               |         |     |          |            |          |          |     |             |        |             |         |              |           |                        |
|            | 0.0028764 | 0.0087795 | 3.9215  | 0.001360 | 0.000151 | 0.037671 | 11.3126  | 1.1047   | -0.21141 | 0.072296 | 0.132051 | 0.77241 | 0.153138 | 0.9071   | 0.9071      | 0.9071   | BENIGN            |
|            | 0.0029496 | 0.0089482 | 2.3120  | 0.001518 | 0.057377 | 11.3126  | 1.1047   | -0.21141 | 0.072296 | 0.132051 | 0.77241 | 0.153138 | 0.9071   | 0.9071      | 0.9071   | MALIGANT          |
|            | 0.0032090 | 0.0096096 | 2.70277 | 0.001515 | 0.929970 | 10.3748  | 1.0095   | -0.74081 | 0.123112 | 0.113301 | 0.77241 | 0.153138 | 0.9071   | 0.9071      | 0.9071   | MALIGANT          |
|            | 0.0031741 | 0.0097498 | 2.40197 | 0.001525 | 0.503049 | 17.7999  | 1.52956  | -0.61584 | 0.042989 | 0.090831 | 0.77241 | 0.153138 | 0.9071   | 0.9071      | 0.9071   | MALIGANT          |
|            | 0.0040472 | 0.0080760 | 2.92325 | 0.001555 | 0.429544 | 23.7297  | 1.46087  | -0.02155 | 0.045055 | 0.122749 | 0.209322 | 0.792171 | 0.153138 | 0.9071   | 0.9071      | 0.9071   | MALIGANT          |
|            | 0.004131 | 0.0070828 | 2.54656 | 0.001555 | 1.95391 | 17.7999  | 1.52956  | -0.61584 | 0.042989 | 0.090831 | 0.77241 | 0.153138 | 0.9071   | 0.9071      | 0.9071   | MALIGANT          |
|            | 0.0044689 | 0.0084943 | 0.64316 | 0.001466 | 0.458337 | 15.7329  | 1.67276  | 0.169828 | 0.269258 | 0.208243 | 0.294137 | 0.146841 | 0.9071   | 0.9071      | 0.9071   | MALIGANT          |
|            | 0.0054690 | 0.0085759 | 2.39285 | 0.001565 | 0.506722 | 15.2545  | 3.41133  | 2.412884 | 0.254179 | 0.123181 | 0.637404 | 0.594154 | 0.77241 | 0.153138 | 0.9071   | 0.9071      | 0.9071   | MALIGANT          |
|            | 0.0033995 | 0.0075384 | 2.93509 | 0.001472 | 0.521901 | 15.621   | 1.10532  | -0.940037 | 0.352687 | 0.130951 | 0.604153 | 0.451318 | 0.77241 | 0.153138 | 0.9071   | 0.9071      | 0.9071   | MALIGANT          |
|            | 0.003529 | 0.0087502 | 2.94822 | 0.001576 | 0.506722 | 15.2545  | 3.41133  | 2.412884 | 0.254179 | 0.123181 | 0.637404 | 0.594154 | 0.77241 | 0.153138 | 0.9071   | 0.9071      | 0.9071   | MALIGANT          |
|            | 0.0043141 | 0.0089272 | 2.94800 | 0.001595 | 0.506722 | 15.2545  | 3.41133  | 2.412884 | 0.254179 | 0.123181 | 0.637404 | 0.594154 | 0.77241 | 0.153138 | 0.9071   | 0.9071      | 0.9071   | MALIGANT          |
| Patients and methods | MRI images from 54 patients attending Baghdad medical city, Baghdad-Iraq, for brain tumor investigations were collected during the period March to June 2016. The most common cancer tumors in Iraq that have taken samples of MRI images from hospitals in the Baghdad medical city are Meningioma for benign and Glioblastoma for malignant diseases as shown in Fig. 1. | Fig. 1: The brain images of the MRI of two types of diseases from Baghdad medical city. | 192
All images were subjected to the program of Zhang et al. [11] in order to collect information about the statistical and shape features considered in the program. These features are listed in Table 2. The ability of this program to detect benign tumors from the 54 images was not as good as expected and the best performance was no more than 78% of correct classification. As a matter of fact such a percentage cannot be considered well enough and many alterations were made to the program in an attempt to increase the percentage of correct classification, but unfortunately the results were always disappointed. For this reason the information of the thirteen features listed in Table 2 were subjected to a multilayer perceptron artificial neural network. The output of the image processing software designed by Zhang et al. [11] were used as an input to the ANN multilayer perceptron discriminant function. It is worthwhile mentioning that formulas and detailed explanation of the meaning of all functions listed in Table 2 are available in MATLAB13 and later versions.

ANN have been developed as generalizations of mathematical models of biological nervous systems [12]. The basic processing elements of neural networks are called artificial neurons, or simply neurons or nodes. In a simplified mathematical model of the neuron, the effects of the synapses are represented by connection weights that modulate the effect of the associated input signals, and the nonlinear characteristic exhibited by neurons is represented by a transfer function. The neuron impulse is then computed as the weighted sum of the input signals, transformed by the transfer function. The learning capability of an artificial neuron is achieved by adjusting the weights in accordance to the chosen learning algorithm.

### Table 2: Variables adopted in this study as described in the paper mentioned previously in this paper.

| Variable name       | Symbol |
|---------------------|--------|
| Mean                | $x_1$  |
| Standard deviation  | $x_2$  |
| Entropy             | $x_3$  |
| RMS                 | $x_4$  |
| Variance            | $x_5$  |
| Smoothness          | $x_6$  |
| Kurtosis            | $x_7$  |
| Skewness            | $x_8$  |
| IDM                 | $x_9$  |
| Contrast            | $x_{10}$|
| Correlation         | $x_{11}$|
| Energy              | $x_{12}$|
| Homogeneity         | $x_{13}$|

Artificial neural networks are characterized by their architecture, activation function and learning paradigm. Multilayer Perceptron (MLP) is one of the mostly used ANNs and about 80% of ANNs researches focused on [13]. It consists of a series of fully interconnected layers of nodes where there are only connections between adjacent layers. General structure is showed in Fig. 2.

If there is several inputs such as $x_1, x_2, ..., x_n$, then these inputs will represent the first layer in the design of the MLP NN, the other layer which will be serve as a hidden layer is the weights (also called synapses) that corresponds to each input $w_1, w_2, ..., w_n$. In addition there is a bias parameter which refers to $w_0$, and can be interpreted as synapse that is associated with artificial input $x_0 = -1$. The output neuron $y$ will be the sum of the products of the input vector $x_0, x_1, x_2, ..., x_n$ by the vector $w_0, w_1, w_2, ..., w_n$, that is:

$$x \cdot w = \sum_{i=0}^{n} x_i w_i$$  \hspace{1cm} (1)
The output neuron can then be calculated by the means of activation function
\[ y = f_{net}(x, y) \] (2)

where a hyperbolic tangent function is usually adopted defined for a generic value \( a \); however it is common to use other activation function in certain situations:
\[ f(a) = \frac{1 - e^{-a}}{1 + e^{-a}} \] (3)

\[ \text{Fig. 2: Multilayer perceptron neural network [9].} \]

One can easily realize that the architect of the multilayer perceptron involves:

- Input layer
- Hidden layer(s)
- Output layer

The software was used with an updated version IBM SPSS ver. 20 which provides a neural network tools that have two main options; radial base and multilayer perceptron. In this paper multilayer perceptron was used with automatic architect option.

**Results**

Table 3 shows the summary of the number of cases used in the procedure of the MLP ANN. In this table we can see the total number of cases used (valid cases) was 54 and that partitioned as 46 cases for the training group and 8 cases for the test group.

\[ \text{Table 3: Summary table.} \]

| Case Processing Summary | N   | Percent |
|-------------------------|-----|---------|
| Training                | 46  | 85.2%   |
| Testing                 | 8   | 14.8%   |
| Valid                   | 54  | 100.0%  |
| Excluded                | 0   |         |
| Total                   | 54  |         |

Fig. 3 shows the structure of the MLP ANN as produced by the SPSS ver. 20. It is clear from this figure the addition of the bias parameter which mentioned previously. There is no need to add it, rather it will be added automatically. It is also clear that the output layer contained two neurons namely Tumor Benign and Tumor Malignant. And these are connected to the input layer through a hidden layer and synapses. The classification process is then to retain the image to an output category that is closest to it.

Percentages of correct classification will calculated with respect to the classification rule derived by MLP ANN. The sensitivity and specificity are directly affected by the classification rule.
Fig. 3: Structure of the MLP ANN.
Table 4 shows the cross classification according to performance of the MLP ANN. The percent of correct classification for the training set is found to be 89.1 whereas for the testing group it found to be 100 \%. This result is actually promising and is much better obtained by the image processing program mentioned previously.

Fig. 4 shows the boxplot of the predicted pseudo probabilities with respect to the benign and malignant categories of the dependent variable tumor. The blue box in the category of benign represent the predicted probability of having benign tumor for the cases with benign tumor. The portion above 0.5 on the y-axis showing the correct classification listed in Table 3. The portion below 0.5 represents incorrect prediction. It is clear that malignant tumor cases have more chance to correctly classify according to the distribution of the pseudo probabilities of Fig. 4.

| Sample  | Observed | Predicted | Percent Correct |
|---------|----------|-----------|----------------|
|         |          | Benign    | Malignant      |                  |
| Training|          | 19        | 3              | 86.4\%           |
|         |          | 2         | 22             | 91.7\%           |
|         |          | 45.7\%    | 54.3\%         | 89.1\%           |
| Testing |          | 5         | 0              | 100.0\%          |
|         |          | 0         | 3              | 100.0\%          |
|         |          | 62.5\%    | 37.5\%         | 100.0\%          |

Fig. 5 shows the Receiver Operating Characteristic (ROC) curve which gives a visual display of the sensitivity and specificity. Two curves one for each category of the dependent variable tumor were showed in the figure. The two curves are about to be similar to each other which indicates the performance of the MLP ANN is very good in detecting cases from the two categories of tumor. Fig. 6 shows the normalized importance of the input variables sorted in an ascending manner. It is clear that x7 (kurtosis) accounted for the highest normalized importance which indicates its role in the judgment of the cases to either category of the dependent variable tumor. The variables contrast, correlation, entropy and energy have a remarkable normalized importance but less than that of the kurtosis.
Fig. 5: ROC curve for the categories of the tumor.

Fig. 6: The normalized importance for the input variables used to produce the MLP ANN.

Discussion

The performance of the ANN depends on the available data and how the set of independent variables prescribe the variation of the dependent variable. Sometimes, the independent variables are not well selected and have no significant effect in explaining any amount of the total variability of the dependent variable. In such situations neither the ANN nor the classic statistical models can help predicting the dependent variable. The size of the data has also an effect on the performance of the ANN, in the case of this research work the number of brain tumor images collected from the Baghdad medical city was not really enough to accomplish more reliable results. It is strongly
recommended that the number of these images should be increased by adding new cases in order to repeat the analysis at every time to observe potential changes.

With respect to the variables with high normalized importance such as kurtosis, contrast, entropy, energy, correlation and all other variables are not really reflecting a threshold of classification to categories of brain tumor categories by their own. This is because another factors needs to be involved and added to the set of independent variables. Laboratorial information may add substantial information about the type of brain tumor. Observation of the brain tumor images suggests that tumors can be characterized by three factors; these are: site, size and shape. Information about these factors may also helped producing more reliable criteria of classification.

References
[1] A. B. Lassman, M. R. Rossi, J. R. Razier, L. E. Abrey, F. S. Lieberman, C. N. Grefe, R. Wilson, Clinical Cancer Research, 01-03 and 00-01, 11, 21 (2005) 7841-7850.
[2] S. K. Singh, I. D. Clarke, M. Terasaki, V. E. Bonn, C. Hawkins, J. Squire, P. B. Dirks, Cancer Research, 63, 18 (2003) 5821-5828.
[3] MJ. Strong, J. Garces, JC. Vera, M. Mathkour, N. Emerson, Brain Tumors Neurooncol, 1, 102 (2015) 2475-3203.
[4] Adelheid Wohrer. Brain Tumour Epidemiology in Austria and the Austrian Brain Tumour Registry. Ph.D Thesis, University of Vienna, (2012), page 89-92.
[5] C. Rajesh Patil and A. S. Bhalchanda, International Journal of Electronics, Communication & Soft Computing Science and Engineering, 2, 1 (2012) 2277-9477.
[6] S. Suhag and L. Mohan Saini., Proceedings of ISER 2nd International Conference, Singapore, (2015).
[7] Esmail Hassan and Abobakr Aboshgifa, International Journal of Computer Science & Engineering Survey (IJCSES), 6 (2015) 67-77.
[8] P. Kr. Saini and M. Singh, International Research Journal of Engineering and Technology (IRJET), 02, 02 (2015) 191-196.
[9] Monica Subashini. M. and Sarat Kumar Sahoo, International Journal of Engineering and Technology (IJET), 5, 2 (2013) 87-93.
[10] Neelam Marshkole, Bikesh Kumar Singh, A.S. Thoke, International Journal of Computer Applications,30,11 (2011) 0975- 8887.
[11] Y. Zhang, L. Wu, S. Wang, Progress in Electromagnetics Research, 130 (2012) 369-388.
[12] H. Peter Sydenham and Thorn Richard. Handbook of Measuring System Design. 2005, John Wiley & Sons, Ltd.
[13] Kais Ncibi, Tarek Sadraoui, Mili Faycel, Amor Djenina, International Journal of Econometrics and Financial Management, 5, 1 (2017) 12-21.