ISSN:2320-0790

An International Journal of Advanced Computer Technology

An Improved Technique for Identification and Classification of Brain Disorder from MRI Brain Image

Finitha Joseph [1], Prof. S. Raja [2]

[1]ME Applied Electronics, Christ the King Engineering College, Karamadai, Coimbatore
[2]Head of the Department (ECE), Christ the King Engineering College, Karamadai, Coimbatore

Abstract: Medical image processing is developing recently due to its wide applications. An efficient MRI image segmentation is needed at present. In this paper, MRI brain segmentation is done by Semi supervised learning which does not require pathology modelling and, thus, allows high degree of automation. In abnormality detection, a vector is characterized as anomalous if it does not comply with the probability distribution obtained from normal data. The estimation of the probability density function, however, is usually not feasible due to large data dimensionality. In order to overcome this challenge, we treat every image as a network of locally coherent image partitions (overlapping blocks). We formulate and maximize a strictly concave likelihood function estimating abnormality for each partition and fuse the local estimates into a globally optimal estimate that satisfies the consistency constraints, based on a distributed estimation algorithm. After this features are extracted by Gray-Level Co-occurrence Matrices (GLCM) algorithm and those features are given to Particle Spam Optimization (PSO) and finally classification is done by using Library Support Vector Machine (LIBSVM).Thus results are evaluated and proved its efficiency using accuracy.

Keywords: Abnormality detection, Gray-Level Co-occurrence Matrices, Image Segmentation, Particle Spam Optimization, Support Vector Machine.

I. INTRODUCTION

Image segmentation is a significant process in image processing. They are used in various applications like biomedicine, remote sensing, control of quality and many others. The main aim of segmentation of image is to extract information from the images to make out different objects of significance. The segmented image separates abnormal area and normal area or differentiates the objects etc.

In medical image segmentation, brain, retina, breast, kidney and liver based image segmentations are the active area of research based on image processing.

The anatomy of the brain is complex due its complicate structure and function [1]. The brain is the part of the central nervous system. It is the centre to control the mental processes and physical action of a human being. Brain abnormality is a symptom where motor impairment and neuropsychological problems affect the central nervous system. It is an abnormal growth of cells within the brain, which can be cancerous or non-cancerous [2]. To date, numerous researches of brain abnormality detection had been conducted due to its important roles in identifying anatomical areas of interest for diagnosis, treatment, or surgery planning paradigms [3].

Magnetic Resonance Imaging (MRI) is a primary medical imaging modality that commonly uses to visualize the structure and the function of human body [4]. It provides rich information for excellent soft tissue contrast which is especially useful in neurological studies [5]. In previous years, MRI is observed to play an important role in brain abnormalities research in determining size and location of affected tissues [6].

Image segmentation refers to a process of assigning labels to set of pixels or multiple regions [7]. It plays a major role in the field of biomedical applications as it is widely used by the radiologists to segment the medical images input into meaningful regions. Thus, various segmentation techniques in medical imaging depending on the region of interest had been proposed [8].

The first MRF theory was introduced into the ground of statistical image analysis in the mid-1980s, Geman and Geman[9] and Besag [10] functional MRFs to image restoration, which can be viewed as a generalization of segmentation. Similar to the work of Geman and Geman [9], Geiger and Girosi [11] also added a second MRF (line process) to the original MRF for surface reconstruction. Likewise, in the work of Jeng and Woods [12] and Molina et al. [13], line process (edge MRF) was incorporated into the intensity process (label MRF). In general, adopting two or more MRFs in one task is a way to solve two or more different problems. For example, Sun et al. [14] integrated three MRFs, disparity, line process and occlusion, for stereo problems because these three factors are all critical to stereo

matching. Similarly, Arduini et al. [15] solved two problems, restoration of SAR images and extraction of intensity discontinuities, by using two distinct MRFs. Held et al. [16] used one added MRF, i.e., the bias field, to sweep the obstacle of MRI brain segmentation but they did not couple the two MRFs compactly because the two fields are assumed independent.

This work makes two fundamental contributions in discovering abnormality. First, an objective function is defined that evaluates probability of the test data according to a statistical model of normal data in a lower dimensional space, and also exploits similarity with the model representation as well as similarity with the original data. The objective function minimization is formulated as a quadratic optimization problem. Second, the curse of dimensionality is tackled by proposing a scheme where an image is partitioned into a set of overlapping blocks at various locations, similarly to [17]. The objective function is optimized for each local subspace and then the local subspace estimates are fused into a globally optimal estimate that satisfies coupling constraints. Data fusion is performed by applying a distributed estimation algorithm based on dual decomposition decomposition [18] and developed for solving large-scale problems. The proposed approach is comprehensively evaluated using receiver operating characteristic (ROC) analysis.

The paper is organized as follows. Section II gives the relation work. In Section III, presents the proposed work. Then results are presented in Section IV followed by conclusion in Section V.

II. RELATED WORK

Atlas-guided [19] approaches are an effective tool for medical image segmentation when a standard atlas or template is available. The atlas is generated by compiling information on the anatomy that requires segmenting. This atlas is then used as a reference frame for segmenting new images. It first finds a one-to-one transformation that maps a pre-segmented atlas image to the target image that requires segmenting. This process is often referred to as atlas warping [19].

An automatic image segmentation method using thresholding technique [20]. This is based on the assumption that adjacent pixels whose value (grey level, color value, texture, etc) lies within a certain range belong to the same class and thus, good segmentation of images that include only two opposite components can be obtained. Threshold based image segmentation are Global Thresholding, Local Thresholding, and Adaptive Thresholding. The key parameter in image segmentation using thresholding technique is the choice of selecting threshold value T.

There are two types Segmentation [21] -Soft Segmentation and Hard Segmentation. Segmentations that allow regions or classes to overlap are called soft segmentations. Soft segmentations are important in medical imaging because of partial volume effects, where multiple tissues contribute to a single pixel or voxel resulting in a blurring of intensity across boundaries [21]. A hard segmentation forces a decision of whether a pixel is inside or outside the object or class. Soft segmentations based on membership functions can be easily converted to hard segmentations by assigning a pixel to its class with the highest membership value. Automated segmentation and delineation of detailed structures remains a difficult task in MRI segmentation.

Clustering algorithms essentially perform the same function as classifier methods without the use of training data. Thus, they are termed unsupervised methods. Two commonly used clustering algorithms are the k -means [23], the fuzzy c-means algorithm. The K-means clustering algorithm clusters data by iteratively computing a mean intensity for each class and segmenting the image by classifying each pixel in the class with the closest mean and fuzzy c-mean [22]has membership function based on membership values it divides pixel into different classes which is also iterative based method.

III. PROPOSED WORK

The methodology for abnormality segmentation here uses 1) a set of pathology-free images in order to calculate an objective function measuring similarity to a healthy brain and 2) a test image which contains both normal and abnormalities for which the objective function is maximized. All images are co registered and the mean image is calculated and subtracted from them. The solution is based on partitioning the spatial domain into overlapping, equally sized blocks in random locations. The algorithmic steps are the following. First, the test image is scanned and a random block is selected (among the not already scanned locations).

By concatenating the image intensities in the block, the test vector $x_0 \in R^d$ is constructed, where d is the number of dimensions (e.g., number of voxels in the block). The same block is then extracted from all pathology-free images forming the training vectors $V_{n\times d}$, where n is the number of subjects. The training set V is used to calculate an objective function $l(x)$ the optimization of which gives a new vector the optimization of which gives a new vector $\hat{x} \in R^d$ that is "less abnormal" and also as similar as possible to the original vectorx_0. However, since the blocks are overlapping, the solutions cannot be independently calculated for each block. After merging the solutions of all blocks, a spatial abnormality score map is calculated for the whole image by subtracting the reconstructed image from the original one.

A. Formulation of the Objective Function

Since anomalies are defined as points with low probability density, it is expected to estimate $\hat{x}$ by maximizing the pdf obtained for the normal data. However, if the vector is high dimensional, the estimation of the pdf is not feasible. Therefore, we will maximize the pdf in a lower dimensional space $p(u)$ where u is the representation of x in a basis W:

$$u = W^T x$$

Here, x is a column vector assumed to be centered at the origin and T denotes the matrix transpose. If the Karhunen–Loeve (KL) transform (or PCA) is applied, the basis W is formed by the $(d \times d)$ matrix of the eigenvectors of the covariance matrix C of the training set V, i.e, $C = \left(\frac{1}{n} - 1\right) V^T V$. The KL transform can be inverted as follows: $x = Wu$.

Assuming that x follows a multivariate Gaussian distribution, the density of u is the multivariate Gaussian density:

$$p(u) = \frac{1}{(2\pi)^{\frac{k}{2}} |D|^{\frac{1}{2}}} e^{-\left(\frac{1}{2}\right) u^T D^{-1} u}$$

Where $D = W^T C W = \mathrm{diag}(\lambda_1, \lambda_2, \ldots\ldots, \lambda_d)$ is a $(d \times d)$ diagonal matrix of eigenvalues, assumed to be sorted in descending order. Typically, the number of samples is significantly smaller than the dimensionality in which case the eigen values λ_t, with $t \geq n$, are zero and the corresponding eigenvectors in W are ignored. If all other eigenvectors are retained $u \in R^{n-1}$.

According to previous equation, $p(u)$ is maximized when $\left(\frac{1}{2}\right) u^T D^{-1} u$ is minimized. Based on maximization of the density in respect to x then is equivalent to minimizing the following term:

$$E_1(x) = \frac{1}{2}(x^T (W D^{-1} W^T) x)$$

Since u is lower dimensional than x, there exist an infinite number of data points $x \in R^d$ with the same function cost value in above equation. In order to reduce the solution space, we use an additional term that constraints the solution to remain close to the subspace spanned by W. If the test vector is x_0 , then its projection on W is

$$x_{0W} = Wu = WW^T x_0$$

The second energy term expresses the distance to the projected point x_{0W}:

$$E_2(x; x_0) = \|x - x_{0W}\|^2 = (x - WW^T x_0)^T (x - WW^T x_0)$$

Where $\|.\|$ denotes the L_2-norm. If $= x_0$, this term expresses the reconstruction error or residual. Since x_0 does not necessarily lie within the subspace spanned by W, this term is larger than zero in this setting. This happens mainly because the abnormal vector x_0 is inconsistent with the normal data building the basisW. Generally by minimizingE_2, we infer that x becomes sufficiently linearly dependent on the current dictionary (normal data), and represents normal behavior.

The first two terms statistically model normality and are used to make the image look like if abnormality were removed. The final term is used to constrain the reconstructed image to be as similar as possible to the original image x_0 based on the assumption that the majority of the voxels in the test image are normal. If all voxels are equally possible to be abnormal, then the distance from can be used as dissimilarity criterion:

$$E_3(x; x_0) = \|x - x_{0W}\|^2 = \sum_{j=1}^{d} (x(j) - x_0(j))^2$$

Where j indicates the voxels in the image.

If prior knowledge exists on spatial locations of possible abnormality, then weights can be incorporated to penalize less the dissimilarity in those locations. Since this method is unsupervised for the abnormal class and aims to generalize for any kind of abnormality, we do not incorporate a prior for the abnormal areas. However, we focus on the normal class and introduce a confidence measure on the estimation ability of the calculated statistical model. Regions with large variability are much more difficult to model than uniform areas. A confidence map or vector shows the degree of certainty we have on the reconstruction of each parameter $x(j)$. Parameters with high uncertainty in estimation should not deviate significantly from their original values$x_0(j)$. This is achieved by penalizing any change on those parameters more than on other parameters. By incorporating an uncertainty vector $a \in R^d$ the third term becomes

$$E_3(x; x_0) = (x - x_0)^T A (x - x_0)$$

Where A is a $(d \times d)$ diagonal matrix with normalized elements $\frac{a(j)}{\sum_{j=1}^{d} a(j)}$ on the main diagonal. The uncertainty vector a is calculated as the average reconstruction error at each location over all training images obtained by leave-one-out cross validation:

$$a = \frac{1}{n} \sum_{t=1}^{n} \sqrt{x_t^T} (I - W_t W_t^T)^2 x_t$$

Where W_t is the basis formed without using training image t.

The previous three terms are combined into a single objective function, $l(x)$ by using different weights, shown as follows:

$$x = \arg\min l(x),$$

Where, $l(x) = w_1 E_1(x) + w_2 E_2(x; x_0) + w_2 E_2(x; x_0)$ And $0 \leq w_1, w_2, w_3 \leq 1$ and $w_1 + w_2 + w_3 = 1$.

According to the values of the weights, we balance between the model term (including E_1 and E_2), controlling the similarity with the training set consisting of normal data, and the data term (E_3), controlling the similarity with the original vector. The weights depend on the confidence we have on the statistical model, as well as on the dominance of novelty or anomaly over the data. The larger the anomaly, the smaller should be the contribution of the data term. The model term on the other hand should always contribute significantly to the solution since it guides toward normality. The weights can be empirically determined by maximizing segmentation accuracy through cross validation on labeled data.

Once the optimization problem is solved, the final reconstructed image is created by recentering to the original space, i.e., by adding the mean image to the result.

B. Optimization of the Objective Function

The objective function can be written in the form of a quadratic programming problem without any linear (equality or inequality) constraints

$$x = \arg\min_x l(x) = \arg\min_x \left(\frac{1}{2} x^T H_x + f^T x\right)$$

Subject to $b_l \leq x \leq b_u$

Where b_l, b_u lower and upper bounds on x, H are is a $(d \times d)$ positive semi definite symmetric matrix, and f is a d-element column vector.

$$H = 2(1 - w_1 - w_2)A + 2w_2 I + w_1 W_r D_r^{-1} W_r^T$$
$$f = -2\{(1 - w_1 - w_2)A + w_2 W_r W_r^T\} x_0$$

Where I is the $(d \times d)$ identity matrix, $D_r^{-1} = \text{diag}\left(\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots\right)$ is the inverse diagonal matrix of the largest eigenvalues retained and W_r the matrix of the corresponding retained eigenvectors.

C. Distributed Estimation

The maximum likelihood estimation problem in a distributed setting is solved using dual decomposition based on the algorithm. Let us assume that k blocks (partitions) are extracted from an image and that the k blocks are coupled through n_c consistency constraints that require the image intensities in overlapping voxels to be equal. The variables that are constraint to be equal across different blocks are denoted as public variables. The variables that are local to each block and are not common in other blocks are denoted as private variables.

Assume that $s_i \in R^{qi}$ and $y_i \in R^{pi}$ are the unknown private and public variables (image intensities) of blocki, respectively. If we concatenate s_i and y_i. we get the vector$x_i = \begin{pmatrix} s_i \\ y_i \end{pmatrix}$, indicating all variables (private and public) in block i. For each block a local (strictly) concave log-likelihood function is calculated by $l_i(x_i)$ or $l_i(s_i, y_i)$.

The public variables for all blocks are collected together into one vector variable $y = (y_1, \ldots\ldots, y_k) \in R^p$, where $p = P_1 + \ldots\ldots + P_k$, is the total number of public variables. A vector $z \in R^{n_c}$ is introduced to give the common values of the public variables in each consistency constraint. The constraints are expressed as

$y = E_z$

where $E \in R^{p \times n_c}$ specifies the set of coupling constraints for the given block interaction

$$E_{ij} = \begin{cases} 1 & \text{if } (y)_i \text{ is in constraint } j \\ 0 & \text{otherwise} \end{cases}$$

Lagrange multipliers $v \in R^p$ are introduced for the coupling constraints and a projected sub gradient method is used to solve the dual master problem. Using these dual variables, optimization is independently performed in each block, and later on, the net variables are updated using the optimal values of the public variables of the blocks adjacent to that net. The dual variables are then updated, in a way that brings the local copies of public variables into consistency.

A measure of the inconsistency of the current values of the public variables (consistency constraint residual) is given by the norm of the vector computed in the last step, $|E\hat{z} - y^*|$.

D. Implementation

The optimization for each block can be as a quadratic programming problem in respect to x_i, where the log likelihood function is given by the negative objective function. In order to extract y_i from x_i, the matrix $M = \begin{bmatrix} O_{qi \times pi} \\ I_{pi \times pi} \end{bmatrix}$, where O is composed of zeros and I is the identity matrix, is the identity matrix, such that $y_i = M^T x_i$.

Then the equation becomes,

$$s_i^*, y_i^* = \arg\min_{s_i y_i}(-l_i(s_i, y_i) + v_i^T y_i)$$
$$\Longrightarrow x_i^* = \arg\min_{x_i}(-l_i(x_i) + v_i^T M^T x_i)$$
$$\Longrightarrow x_i^* = \arg\min_{x_i}\left(\frac{1}{2} x_i^T H_i x_i + f_i^T x_i\right)$$

Where $H_i = H$ and is calculated for block i, and $f_i = f + M v_i$.

E. Independent Component Analysis(ICA)

The ICA segmentation is efficient segmentation which is used before feature extraction here.

ICA of a random vector x consists of estimating the following generative model for the data:

$$x = As$$

where the latent variable (components) s_i in the vector $s = (s_1, \ldots\ldots s_n)^T$ are assumed independent. The matrix A is a constant $m \times n$ 'mixing' matrix.

This is the simplest and widest used definition in most research on ICA. There are also other ICA definitions, which can be found in the literature [24,25].

To maximize by stochastic gradient ascent the joint entropy $H(g(y))$ of the linear transform squashed by a sigmoidal function g. The updating formula for W is:

$$\Delta W = (I + g(y)y^T)W$$

Where $y = Wx$ and $g(y) = 1 - \frac{2}{(1 + e^{-y})}$ is calculated for each component of y. Before the learning procedure, x is sphered by subtracting the mean $m_x$and multiplying by a whitening filter:

$$X = [(x - m_x)(x - m_x)^T]^{-1/2}(x - m_x)$$

This gives the segmented image from which features are extracted.

F. Feature Extraction:

The features are important for every classification algorithms. Here texture features of images are extracted.

The GLCMs features are stored in a $i \times j \times n$ matrix, where n is the number of GLCMs calculated usually due to the different orientation and displacements used in the algorithm. Usually the values i and j are equal to 'NumLevels' parameter of the GLCM computing function. Note that matlab quantization values belong to the set $\{1, \ldots, \text{NumLevels}\}$ and not from $\{0, \ldots, (\text{NumLevels} - 1)\}$ as provided.

The following GLCM features are extracted:

- Autocorrelation
- Contrast
- Correlation
- Correlation
- Cluster Prominence
- Cluster Shade
- Dissimilarity
- Energy
- Entropy
- Homogeneity
- Homogeneity
- Maximum probability
- Sum of squares
- Sum average
- Sum variance
- Sum entropy
- Difference variance
- Difference entropy
- Information measure of correlation1 and 2
- Inverse difference (INV)
- Inverse difference normalized (INN)
- Inverse difference moment.

G. Classification:

The classification of abnormality and normality is improved here by using PSO with LSVM technique.

I. Particle swarm optimization

Particle swarm optimization (PSO) is a population-based optimization algorithm modeled after the simulation of social behavior of birds in a flock [27]. The algorithm of PSO is initialized with a group of random particles and then searches for optima by updating generations. Each particle is flown through the search space having its position adjusted based on its distance from its own personal best position and the distance from the best particle of the swarm. The performance of each particle, i.e. how close the particle is from the global optimum, is measured using a fitness function which depends on the optimization problem.

Each particle i flies through an n-dimensional search space, R^n, and maintains the following:

x_i, the current position of ith particle (x-vector)

p_i, the personal best position of ith particle (p-vector), and

v_i, the current velocity of ith particle (v-vector).

The personal best position associated with a particle, i, is the best position that the particle has visited so far. If f denotes the fitness function, then the personal best of i at a time step t is updated as:

$$p_i(t+1) = \begin{cases} p_i(t) \text{ if } f(x_i(t+1)) \geq f(p_i(t)) \\ x_i(t+1) \text{ if } f(x_i(t+1)) < f(p_i(t)) \end{cases}$$

If the position of the global best particle is denoted by gbest , then :

$$\text{gbest} \in \{p_1(t), p_1(t), \ldots\ldots, p_m(t)\} = \min\{f(p_1(t)), f(p_2(t)), \ldots\ldots f(p_m(t))\}$$

The velocity updates are calculated as a linear combination of position and velocity vectors. Thus, the velocity of particle i is updated and the position of particle i is updated by the following equations.

$$v_i(t+1) = w.v_i(t) + c_1 r_1(p_i(t) - x_i(t)) + c_1 r_1(\text{gbest} - x_i(t))$$

$$x_i(t+1) = x_i(t) + v_i(t+1)$$

In the formula, w is the inertia weight [26], c_1 and c_2 are the acceleration constants, r_1 and r_2 are random numbers in the range [0,1] and V_{i1} must be in the range $[-V_{max}, V_{max}]$, where V_{max} is the maximum velocity.

II. *Library Support Vector Machine:*

LIBSVM is a library for Support Vector Machines (SVMs). The goal is to easily apply SVM to their applications. LIBSVM has gained wide popularity in machine learning and many other areas. In this work, we present implementation of LIBSVM. Issues such as solving SVM optimization problems theoretical convergence multiclass classification probability estimates and parameter selection.

A typical use of LIBSVM involves two steps: first, training a data set to obtain a model and second, using the model to predict information of a testing data set. For SVC and SVR, LIBSVM can also output probability estimates.

This is same as SVM technique, where in training SVM the An m by 1 vector of training labels (type must be double) is taken.

Then parameters for gamma in LIBSVM are taken from PSO algorithm. And the Cost parameter is set as the parameter C of C-SVC is taken.

Kernels:

Kernel methods in general have gained increased due to the grown of popularity of the Support Vector Machines. Support Vector Machines are linear classifiers and regressors that, through the Kernel trick, operate in reproducing Kernel Hilbert spaces and are thus able to perform non-linear classification and regression in their input space.

Here Radial Bias Function Kernel is used and it is expressed as

$$\text{RBF} = \exp\left(\frac{1}{2\sigma^2 \|x - x_i\|^2}\right)$$

And detailed discribtion is given in [28]. This is given as kernel in LIBSVM technique. By this the classification of abnormal and normal MRI brain images is performed and their affected disease is identified.

IV. EXPERIMENTAL RESULTS

The experiment is done to evaluate the performance of this proposed work. The MATLAB environment is chosen and MRI images are collected from various Scan centers. Here three major diseases type of Astrocytoma, Giloma and Metastasis and abnormality of dyaplasia and Brain infraction is identified.

The input image is preprocessed and features are extracted from it. The extracted features are given for the

PSO algorithm which helps in finding the parameters for the LIBSVM classification.

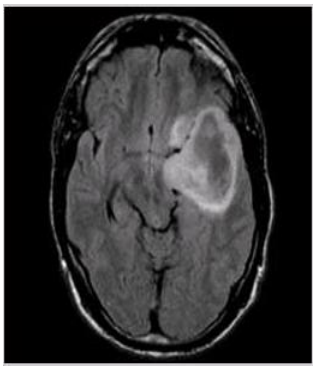

Figure 3.1: Input Image-1

Figure 3.2: Segmented Output

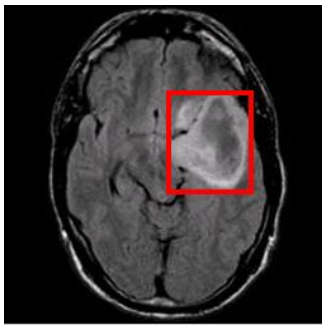

Figure3.3: Identified of Affected Area

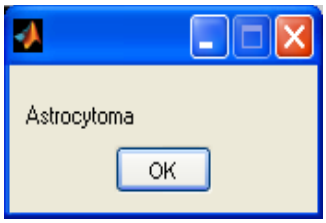

Figure 3.4: Output Result

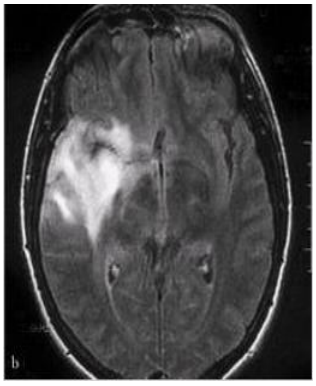

Figure 3.5: Input Image-2

Figure 3.6: Segmented Output

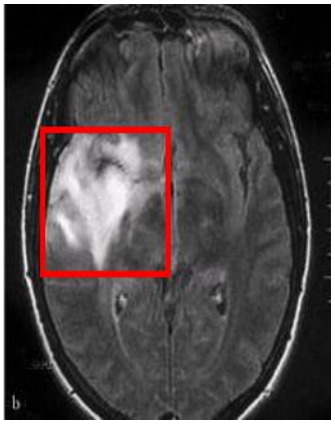

Figure3.7: Identified of Affected Area

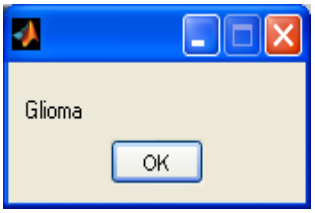

Figure 3.8: Output Result

Likewise for other diseases and abnormalities are identified by using this technique.

A. Performance Evaluation:

The classification accuracy is measured here by the MRI brain image database. The input database consists of MRI brain images containing diseases of Astrocytoma, Giloma and Metastasis. And fracture of Brain infraction and dyaplasia.

Table 1: Analysis of Classification Results For PSO with LIBSVM

Techniques	*Input Given*	*Correctly Classified*	*Wrongly Classified*	*Accuracy (%)*
Astrocytoma	34	34	0	100
Giloma	34	31	3	97
Metastasis	22	22	0	100
Brain infraction	6	6	0	100
dyaplasia	4	4	0	100
Total	100	97	3	97

The table 1 shows that proposed algorithm of PSO with LIBSVM produces best accuracy for the Metastasis, Brain infraction and dyaplasia. And better results for Giloma for the input images.

V. CONCLUSION

The identification and detection of abnormality and diseases in brain is carried here by MRI brain images. The proposed work is done by extracting GLCM features and given to the PSO. The PSO is efficient technique for segmentation which finds the parameters for the classification technique. The classification technique of Library SVM is used which is improved technique than traditional SVM in which RBF kernel is used to boost the classifier. The performance of this work is measured by accuracy calculation for three brain diseases and two tractors which proves the nearly maximum results. In future large database can be taken for evaluation with more number of brain images.

VI. REFERENCES

[1] L. G. Qurat-Ul-Ain, S. B. Kazmi, M. A. Jaffar, and A. M Mirza, "Classification and Segmentation of Brain Tumor using Texture Analysis", Recent Advances in Artificial Intelligence, Knowledge Engineering and Data Bases, ISBN: 978-960-474-154-0, pp. 147-155, 2010.

[2] L. M. DeAngelis, "Brain Tumors", The New England Journal of Medicine, Vol. 344, No. 2, pp 114-123, 2001.

[3] A. M. Riad, A. Atwan, H. M. El-Bakry, R. R. Mostafa, H. K. Elminir, and N. Mastorakis, "A New Approach for Segmentation of Brain MR Image", Proceedings of the WSEAS International Conference on Environment, Medicine and Health Sciences, ISSN: 1790- 5125, ISBN: 978-960-474-170-0, pp. 74-83, 2010.

[4] H. B. Kekre, and S. Gharge, "Direct Variance on MRI Images for Tumor Detection", Journal of Science, Engineering and Technology Management, Vol. 2, No. 1, January 2010, pp. 3-11.

[5] P. Vasuda, and S. Satheesh, "Improved Fuzzy C-Means Algorithm for MR Brain Image Segmentation", International Journal on Computer Science and Engineering (IJCSE), Vol. 02, No. 05, pp. 1713-1715, 2010.

[6] P. Anbeek, K. L. Vincken, and M. A.Viergever, "Automated MS-Lesion Segmentation by K-Nearest Neighbor Classification", 2008.

[7] R. Ganesan, and S. Radhakrisham, "Segmentation of Computed Tomography Brains Images Using Genetic Algorithm", International Journal of Soft Computing 4(4): pp. 157-161, 2009.

[8] J. Roerdink, and A. Meijster, "The Watershed Transform: Definitions, Algorithms and Parallelization Strategies", Fundamenta Informaticae, pp. 187-228, IOS Press, 2001.

[9] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images," IEEE Trans. Pattern Anal. Mach. Intell., vol. 6, no. 6, pp. 721–741, Jun. 1984.

[10] J. Besag, "On the statistical analysis of dirty pictures," J. Roy. Statist. Soc. B, vol. 48, pp. 259–302, 1986.

[11] D. Geiger and F. Girosi, "Parallel and deterministic algorithm from MRF's: Surface reconstruction," IEEE Trans. Pattern Anal. Mach. Intell., vol. 13, no. 5, pp. 401–412, May 1991.

[12] F. Jeng and J. Woods, "Compound Gauss-Markov random fields for image estimation," IEEE Trans. Signal Process., vol. 39, no. 3, pp. 683–697, Mar. 1991.

[13] R. Molina, J. Mateos, A. Katsaggelos, and M. Vega, "Bayesian multichannel image restoration using compound Gauss-Markov random fields," IEEE Trans. Image Process., vol. 12, no. 12, pp. 1642–1654, Dec. 2003.

[14] J. Sun, N. Zheng, and H. Shum, "Stereo matching using belief propagation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 7, pp. 787–800, Jul. 2003.

[15] F. Arduini, C. Dambra, and C. S. Regazzoni, "A coupled MRF model for sar image restoration and edge-extraction," in Proc. Int. Geosci. Remote Sensing Symp., 1992, vol. 2, pp. 1120–1122.

[16] K. Held, E. Kops, J. Krause, W. Wells, R. Kikinis, and H. Muller-Gartner, "Markov random field segmentation of brain MR images," IEEE Trans. Med. Imag., vol. 16, no. 6, pp. 878–886, Jun. 1997.

[17] G. Erus, E. I. Zacharaki, R. N. Bryan, and C. Davatzikos, "Learning high-dimensional image statistics for abnormality detection on medical images," in Proc. IEEE Comput. Soc. Workshop Math. Methods Biomed. Image Anal., 2010, pp. 139–145.

[18] S. Samar, S. Boyd, and D. Gorinevsky, "Distributed estimation via dual decomposition," in Proc. Eur. Control Conf., 2007, pp. 1511–1516.

[19] P.K. Sahoo, S. Soltani, and A.K.C.Wong. A survey of thresholding techniques. Comput. Vis. Graph. Im. Proc., 41:233–260, 1988.

[20] G. Evelin Sujji, Y.V.S. Lakshmi, G. Wiselin Jiji," MRI Brain Image Segmentation based on Thresholding ",International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume-3 Number-1 Issue-8, March-2013, pp.97-101.

[21] Dzung L. Phamy, Chenyang Xu, Jerry L. Prince," A Survey Of Current Methods In Medical Image Segmentation" Annual Review of Biomedical Engineering January 19, 1998.

[22] Xin-Bo Zhang And Li Jiang "An image Segmentation algorithm Based On Fuzzy C-Means Clustering", International Conference On Digital Image Processing,March 2009 ,pp 22-26.

[23] G.B. Coleman and H.C. Andrews. Image segmentation by clustering. Proc. IEEE, 5:773–785, 1979

[24] P. Comon, "Independent component analysis - a new concept", Signal Processing, 36, 1994, pp: 287-314

[25] C. Jutten, J. Herault, "Blind separation of sources", Signal Processing, Part I: An adaptive algorithm based on neuromimetic architecture. 24, 1991, pp: 1-10.

[26] Y. Shi, and R. Eberhart, "A modified particle swarm optimizer," in the Proceedings of the IEEE International Conference on Evolutionary Computation, Piscataway, NJ, pp. 69-73, 1998.

[27] J. Kennedy, and R. Eberhart, Swarm Intelligence, San Francisco: Morgan Kaufmann Publishers, 2001.

[28] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin," A Practical Guide to Support Vector Classification", April 15, 2010.