



## Analysis of Classification Algorithms for brain tumor detection

---

Sai Sasank Varthakavi, Devisetty Rohith Prasanna Babu,  
M B Amsitha and Sasi Jyothirmai Bonu

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

November 2, 2019

# Analysis of Classification Algorithms for Brain Tumor Detection

Sai Sasank Varthakavi<sup>1</sup>, Devisetty Rohith Prasanna Babu<sup>2</sup>, Amsitha.MB<sup>3</sup>, Sasi Jyothirmai Bonu<sup>4</sup>  
Department of Electronics and Communication Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India.  
saisasank86@gmail.com<sup>1</sup>, rp babu 461@gmail.com<sup>2</sup>, amsitha.mb@gmail.com<sup>3</sup>, sasibonu@gmail.com<sup>4</sup>

**Abstract:** Image classification of brain tumor MRI data is very critical and decisive work. In this paper, we implemented an algorithm to classify images of brain tumors from the data of those who do not have tumors using two efficient well-known classification techniques namely KNN and SVM. The pre-processing techniques include Gaussian Filtering and Adaptive median filtering were implemented for both classifiers, by extracting the HoG features. When an impulse noise is dominant in such an image, a normal adaptive median filter is well known way to remove the impulse noise. For both classifiers the accuracy has been determined, the results vividly show that SVM performs far better than K-NN with an accuracy of 81.1%, whereas the accuracy obtained for K-NN was 57.64%. The entire analysis is done using MATLAB 2019a software.

**Keywords:** MRI, Preprocessing filter, Gaussian, Median filter, FeatureExtraction, HOG, SVM, K-NN.

## I . Introduction

Brain tumor, in clinical treatment, is generally diagnosed by computer tomography and magnetic resonance imaging(MRI). Abnormal and uncommon development of brain cells can cause brain tumors. As the skull is a firm and volume-limited body part, any unanticipated development may affect the functioning according to the entangled part of the brain and it might spread to other organs of the body. Classification of brain tumors can be done in many ways, like primary or secondary tumors. Around 70% of all brain tumors are primary, whereas secondary tumors are the remaining 30%. According to its origin, this classification is determined. All the tumors initiating in the brain are called primary tumors, whereas the tumors whose origins can be traced in other parts of the body and then extended to the brain are said to be secondary tumors, and unfortunately many of them are said to be malignant.

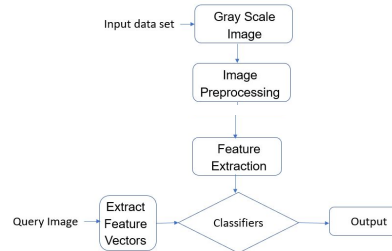


Fig.1. Block diagram representing Brain tumor detection.

Classification of brain tumors can be done by using many imaging techniques such as magnetic resonance imaging. MRI's prevalence arises from the fact that it does not use ionizing radiation while scanning, the predominant soft-tissue resolution and the ability to pertain different images using various imaging notations or by adopting contrast-enhanced agents. Here we are going to classify the benign tumors images from malignant tumor images using different algorithms like KNN (K nearest neighbors) and SVM (support vector machine) as shown in Fig.1. The classification here is done by finding similar data points in the training data the extension of support vector classifiers is the support vector machine that has non-linear class boundaries which are caused by making the feature space bigger using kernels. The more emphasis on KNN and SVM classifications and their results with respective accuracies are given in the further sections below.

## II . Preprocessing

**Gaussian Filtering:** Generally noise is random and there are several types of noises are there in present-day life but here our main focus is to deal with Image noise, which is a random variation of intensities over spatial domain i.e, pixel by pixel of an Image. Ideally preprocessing of an image will enhance the information hidden in an image i.e, filtering noise from an image. The main sources of image noise can be electrical, due to high temperature, to filter out the noise we have to opt one effective noise filtering technique An important concern is how to implement an effective noise filtering technique.

So we considered the Additive White Gaussian Noise (AWGN) as the primary noise in an image because like the white colour it has all frequencies of noise components over its bandwidth and if we have to filter out this noise the function used for filtering should be a gaussian. The main advantage is whether the filter is in the time domain or frequency domain the function is a Gaussian and the inbuilt command *imgaussfilt* in Matlab 2019a used for filtering will perform the filter operation in the time domain or frequency domain based on internal heuristics. To remove this noise we used Gaussian function [1] which is of the form (1)

$$F(x,y) = \frac{1}{2 \times \pi \times \sigma^2} \times e^{-\frac{(x^2+y^2)}{2 \times \sigma^2}} \quad (1)$$

Where (x,y) represents the location of each pixel of a 2D image.

$\sigma$  represents standard deviation.

When we apply Gaussian function to an image each pixel intensity value will be replaced by the average of all neighborhood pixel's, which shows that the pixels which are closer to the origin will have high-intensity value because of gaussian and the neighborhood pixels will have low-intensity values [2] when their distance from the origin increases, results in vanishing of edges in image and causes image blur, which is commonly known as Gaussian blur [3] as shown in Fig.2. Mathematically it is the convolution sum of a given image with the Gaussian function.

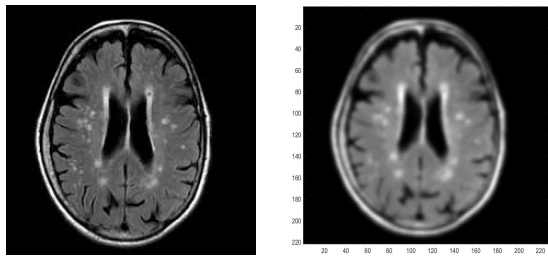


Fig.2. Normal Image and Gaussian Blur

**Adaptive Median Filter:** The window size of the adaptive median filter is not constant and keeps on changing over its period, this is the main reason it has been chosen over the standard median filter. The algorithm is described as follows:

The image is divided into rectangular blocks  $R_{xy}$  for the adaptive median filter to operate on. During the filtering operation, the size of  $R_{xy}$  can be amended

using adaptive median filtering[5] which varies with some specific constraints as mentioned below. The output of the operation shows a monotonous value that replaces the pixel value at the specified position[6]. notations have been used are :

- Hmin =the minimum pixel value in  $R_{xy}$
- Hmax =the maximum pixel value in  $R_{xy}$
- Hmed =the median pixel value in  $R_{xy}$
- $H_{xy}$  = coordinate at location (x,y)
- Rmax = approved maximum size

Now we look over the adaptive median filter. Hmed is not an impulse noise, Since  $H_{min} < H_{med} < H_{max}$ . If that is the case, the algorithm jumps to Stage B. The gist of Stage B is:  $H_{xy}$  is evaluated to see if it is an impulse noise or not. If  $H_{xy}$  is indeed an impulse noise, it will be replaced by Hmed. If that is indeed the case,  $H_{xy}$  is considered not an impulse noise. Now, if and only if the pixel is an impulse, the filtered image's pixel value remains the same as shown in Fig.3.

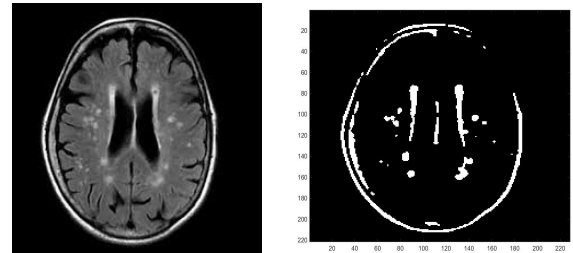


Fig.3. Normal Image and Filtered Image.

This filter carries out the operation of removing the impulse noise and decreases the distortion from the image. Images corrupted with the noise greater than 0.2 probability can be allowed and managed as well. The result so obtained is much better than the results obtained from the standard median filter.

### III. Feature Extraction

Feature extraction is one of the important steps before feeding it to any classifier. In this paper, we discussed the HOG feature extractor which can be used for improving the accuracy of object detection.

**Gradient determination:** Gradient calculation is the primary step in deriving the hog features. In the first

step, horizontal and vertical derivatives are applied as kernels. Generally 1D derivative mask of dimension 1\*3 and 3\*1 is applied in horizontal and vertical Direction.

Dx -Horizontal mask in the x-direction.

Dy- Vertical mask in the y-direction.

Dx =

-1	0	1
----	---	---

Dy =

-1
0
1

#### Orientation segregation:

In this process based on the gradient determination magnitude and orientation matrices are generated and based on the orientation i.e either from 0 to 360 degrees or from 0 to 180 degree unsigned bins are generated. The orientation value is mapped to the corresponding magnitude and the weighted voting process will be carried out in such a way that the magnitude value will be distributed into the corresponding bin in which its orientation belongs to and the magnitude will be constantly added. Another alternative is computing square of the gradient. Dalal et.al proposed 9 bins and 0-180 degree unsigned combinations give the best results for face detection and the same has been followed for the image classification.

#### Feature descriptor:

Block normalization is implemented to reduce the bad intensity levels in the pixel values. This process is implemented by concatenating the cells of the image and concatenating cells helps in the overlapping of the blocks at least once which helps in the contribution to form the feature descriptor accurately[7]. Blocks can be formed in two ways either rectangular or circular block. Rectangular hog block method is implemented by forming square grid and rectangular grids shown in Fig.4 and it depends on how many number of cells in each block and the number of pixels in each cell and also the number of channels in a histogram[8]. Based on Dalal et al exploration there are four methods[9] for block

normalisation which contributes to the effective formation of feature descriptors.

$$\text{L2-normalisation } f = u / \sqrt{\|u_2\|^2 + e^2} \quad (2)$$

Similarly L2-hys can be formed by clipping (limiting the maximum values of v to 0.2) and renormalisation), as in equation (3) and (4)

$$\text{L1-normalisation: } f = u / \sqrt{\|u_1\| + e} \quad (3)$$

$$\text{L1-square root: } f = \sqrt{u / \|u_1\| + e} \quad (4)$$

It has been proved that L2-hys, L2-normalisation (2) and L1-square root (4) has equal performance and L1-norm (3) has less reliability in performance over the other processes but an improved performance over the non-normalised method.

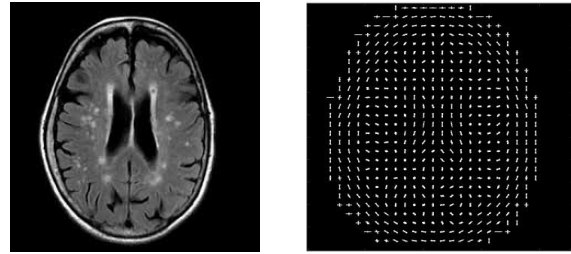


Fig. 4. Normal Image and Feature Vector Image.

## IV. Classification

### A. Support Vector Machine(SVM):

Support vector machine (SVMs) SVM finds the best isolating (maximalede) hyperplane between dataset and testset in the component space, which prompts maximal speculation. Albeit and few authors have exhibited the hypothetical foundation of the spatial properties of SVMs, uncovered completely [10], the standard of illuminating strategies originates from the arithmetical field (basically decay). A standout Optimisation amongst other delegate mathematical operations with regard to speed and simplicity of usage, by introducing some high standard adaptable properties, namely Sequential Negligible Optimization (SMO) [11]. The most important problem is to deal with the spatial properties of learning and lucidity of SVMs in the vector space, which results in double portrayal i.e, the projection of each and every class in the dataset and to calculate the hyperplanes from every individual points in vector space that the maximal edge for the detachable instance [12] and we have another instance i.e, non separable points in the component space, in such a case the optimisation used is Reduced Convex

Structure (RCH) [13]. All things considered, the geometric calculations displayed as of recently ([14]) are appropriate for unraveling legitimately the distinct case and in a roundabout way the non-divisible case through the stunt proposed in [15]. Be that as it may, the previously mentioned (misleadingly expanding the vector space by including different points from different vector space ) case which is identical to violation of the given constraints multiplied resulting parameters, due to the expansion of multifaceted nature by falsy expanding of elements in vector space resulting in the spatial properties of the subsequent SVMs can be poor[16].

This hyperplane is unique for every model and it is constructed in a way that it reduces the maximum error and maximize the geometric margin. Testing data label is predicted by calculating the distance between training data features and testing data features and the label which has got the minimal distance will be assigned to the testing dataset.

### B. K-Nearest Neighbours(K-NN):

It is one of the easiest algorithms in terms of application for classification and regression predictive problems. Also, it is easy to predict the output and minimize the calculation time.

Since we have noise filtered images of both benign(Class 1) and malignant (Class 2) classes, we intend to categorize the data into these two classes. All the features extracted through HoG lie on the axes. Each feature gets each axis. Since we have extracted 36 HoG features, we will have 36 different axes.

The “K” in the name is for the nearest neighbors we wish to take votes from. A circle is then drawn with the test input as the center and the size just enough to capture “k” points inside it[17]. Using the ‘k’ closest points to that input data point, we predict the output. The input is categorized as the class that gets the maximum number of votes.

The choice of parameter K plays a vital role in this algorithm. In general, a big K value is considered to be precise as it optimizes the total noise but there is no 100% assurance. The other way of determining the perfect K value is by using a cross-validation method [18]. These methods help in providing better results than neural networks. K value boundaries can be made for each class and these boundaries will segregate benign from malignant images. Consequently, the value of K will be shown on the class boundaries. The smoothening of boundaries is

directly proportional to the increase in k value. As k tends to reach infinity, every data point finally becomes one of either class which depends on the total majority. Different K values give different training error rate and validation error rate and these need to be accessed[19]. The error rate comes to be zero whenever the value of k is one as the closest point to any training point is the point itself. And so the prediction is always accurate with k=1.

After the segregation of the training data and it’s validation from the given dataset, the optimal value of k can be obtained by plotting the validation error curve. Now, this k is almost perfect and can be used for all the predictions.

The similarity measure between the input data points and the existing data points can be taken using distances between them. Our algorithm used the Euclidean distances between the data points.

$$d(p,q) = \sqrt{\sum \text{square}(p_i - q_i)}$$

## V. Results and evaluation

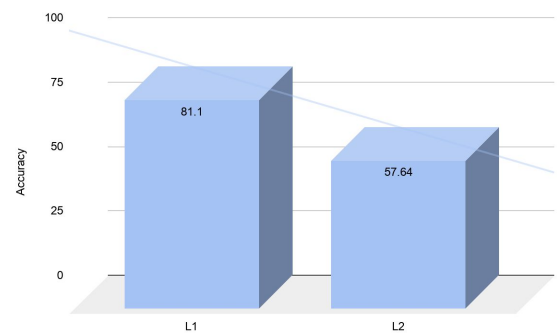


Fig.4. Graph showing Accuracies of best fitted models

L1 represents classification with SVM.

L2 represents classification with KNN.

**Conclusion:** The problem addressed here is the classification of image data for MRI classification as benign or malignant using two classification algorithms namely SVM and KNN in order to assess the effectiveness of these algorithms. Having used both SVM and KNN and analyzing all the facts and working with the same number of images and extracting the same features from the images, it can be safely and confidently said that SVM is way better and far more precise than KNN in terms of classification accuracy. As can be determined by the confusion matrix, the accuracy of SVM is 81% whereas the accuracy of KNN is 60%, which is very poor comparatively. Also the computation time

required for SVM was less when compared to KNN making it even more effective. SVM is better at traditional pattern recognition approach also. As a final remark, it can be said that parametric classifier like SVM should be preferred over non parametric ones like KNN.

## REFERENCES

- [1] Sandeep Chand Kumain, Maheep Singh, Navjot Singh, Krishan Kumar, "An efficient Gaussian Noise Reduction Technique For Noisy Images using optimized filter approach", *Secure Cyber Computing and Communication (ICSCCC) 2018 First International Conference on*, pp. 243-248, 2018.
- [2] A.K. Jain, Fundamentals of Image Processing, Prentice Hall, 1989.
- [3] Somkheart Kraisin, Natsuda Kaothanthong, "ACCURACY IMPROVEMENT OF A PROVINCE NAME RECOGNITION ON THAI LICENSE PLATE", *Artificial Intelligence and Natural Language Processing (iSAI-NLP) 2018 International Joint Symposium on*, pp. 1-6, 2018.
- [4] Stefano Di Carlo, Paolo Prinetto, Daniele Rolfo, Pascal Trotta, "AIDI: An adaptive image denoising FPGA-based IP-core for real-time applications", Adaptive Hardware and Systems (AHS) 2013 NASA/ESA Conference on, pp. 99-106, 2013.
- [5] Guoliang Fan, Wai-Kuen Cham, "Model-based edge reconstruction for low bit-rate wavelet-compressed images", *Circuits and Systems for Video Technology IEEE Transactions on*, vol. 10, no. 1, pp. 120-132, 2000.
- [6] Ajai, A & Nagaraj, Nithin. (2012). A Novel Methodology for Memory Reduction in Distributed Arithmetic Based Discrete Wavelet Transform. *Procedia Engineering*. 30. 226-233. 10.1016/j.proeng.2012.01.855.
- [7] Mariana-Eugenia Ilas, "Parameter selection for efficient HOG-based car detection", *Industrial Electronics (ISIE) 2017 IEEE 26th International Symposium on*, pp. 1439-1444, 2017.
- [8] Natthariya Laopracha, Khamron Sunat, Sirapat Chiewchanwattana, "A Novel Feature Selection in Vehicle Detection Through the Selection of Dominant Patterns of Histograms of Oriented Gradients (DPHOG)", *Access IEEE*, vol. 7, pp. 20894-20919, 2019.
- [9] Puji Lestari, Hans-Peter Schade, *Advances in Signal Processing and Intelligent Recognition Systems*, vol. 678, pp. 285, 2018.
- [10] K. P. Bennett, E. J. Bredensteiner "Geometry in Learning" in C. Gorini, E. Hart, W. Meyer and T. Phillips, editors, *Geometry at Work*, Mathematical Association of America, 1998.
- [11] K. P. Bennett, E. J. Bredensteiner "Duality and Geometry in SVM classifiers" in Pat Langley, editor, *Proc. 17th International Conference on Machine Learning*, Morgan Kaufmann, 2000, pp. 57-64.
- [12] D. J. Crisp, C. J. C. Burges "A geometric interpretation of v-SVM classifiers" *NIPS 12*, 2000, pp. 244-250.
- [13] T. T. Friess, R. Harisson "Support vector neural networks: the kernel adatron with bias and soft margin" Technical Report ACSE-TR-752, University of Sheffield, Department of ACSE, 1998.
- [14] S. S. Keerthi, S. K. Shevade, C. Bhattacharyya, K. R. K. Murthy "A fast iterative nearest point algorithm for support vector machine classifier design", Technical Report No. TRISL-99-03, Department of CSA, IISc, Bangalore, India, 1999.
- [15] J. Platt "Fast training of support vector machines using sequential minimal optimization" in B. Schölkopf, C. Burges and A. Smola, editors, *Advances in Kernel Methods – Support Vector Learning*, pp. 185-208. MIT Press, 1999.
- [16] Praveen, K & Poornachandran, Prabaharan & Ajai, A. (2012). Implementation of DES using pipelining concept with skew core key scheduling in secure transmission of images. *ACM International Conference Proceeding Series*. 637-641. 10.1145/2393216.2393322.
- [17] Che[n Lu, Dong Liang, Shan Wang, Lili Zeng, Yilin Zhao, "Pre-cut kNN Algorithm Based on Threshold of Distance", *Network Infrastructure and Digital Content (IC-NIDC) 2018 International Conference on*, pp. 309-314, 2018.
- [18] C.P. Samjith Raj, R. Shreeja, "Automatic brain tumor tissue detection in T-1 weighted MRI", *Innovations in Information Embedded and Communication Systems (ICIIECS) 2017 International Conference on*, pp. 1-4, 2017.
- [19] Iyer Aurobind Venkatkumar, Sanatkumar Jayantibhai Kondhol Shardaben, "Comparative study of data mining clustering algorithms", *Data Science and Engineering (ICDSE) 2016 International Conference on*, pp. 1-7, 2016.