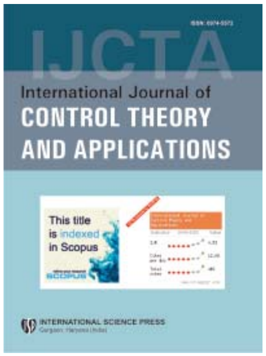

**International Journal of Control Theory and Applications**

ISSN : 0974–5572



**Volume 9 • Number 40 • 2016**

# An Outlier Detection of a Medical Image using SVM Technique

**P. Saibalaji[a] and N. Arivazhagan[b]**

*[a]SRM University, Department of information technology Chennai*

*E-mail: pakkarakarisaibalaji@gmail.com*

*[b]Department of information technology, Assistant professor (SL.G) Chennai*

*E-mail: n.azhagu@gmail.com*

***Abstract:*** Outliers When upon a period viewed Similarly as loud information over statistics, need turned out will make a critical issue which is, no doubt explored over different fields about Examine Furthermore provision domains. In this paper we recommend a Gaussian bilateral filter which may be An non-linear, Edge-preserving and noise reducing smoothing filter for images that need with be improved. . Here the Intensity values of the image at each pixel were replaced by average of intensity values nearby pixels. In the recommended system, we are taking MRI brain image for outlier detection and we use K-Nearest Neighbor(KNN) algorithm to segment the tumor affected part of the brain image and we utilize SVM classifier to classify the which type of tumor affected which analyses the data needed for classification.

***Keywords:*** *Bi lateral filter, outlier, k nearest neighbor, SVM classifier.*

## 1. INTRODUCTION

The set of data points that are considerably different than the remainder of the data are called outliers.

In many applications, data are generated by processes which may reflect either the activity of the system or observation of objects in the system.[1] Outliers may appear in the data due to some reasons like Mechanical fault, Changes in System behavior, Fraudulent Behavior, Human Error, Instrument error, Natural deviations in populations, changes of environment etc

Outlier can be noise or interesting item. There is no clear distinction between outlier and noise. It depends on interest of the analyst of the system. [1] In both the cases outlier detection is crucial because most of the statistical methods cannot work well in the presence of outliers. [2]

Global outliers: Global outliers are also called point outliers. A data point is called global outlier if it is different or far from the whole dataset. It is very simple form of outlier[3]. Most of the techniques are designed for finding this type of outliers

Contextual outliers: It is also called conditional Outlier. A data point is called contextual outlier if it is different or far from the other data points in the specific context. Contextual outlier detection techniques provide flexibility for detecting[3] outliers in different contexts.

In contextual outlier detection data attributes are divided into two groups Contextual Attributes: Attributes with respect to which a data point is considered an outlier are contextual attributes. It defines context of the object .For example time, spatial attribute (Longitude and latitude), network location etc.[2] Behavioral attributes: They are non-contextual attributes and evaluated to find outlierness of data point in context in which it belongs.[4] For example Temperature, humidity, pressure, rain fall etc. [2]

Collective outliers: A collection of data point as a whole is different from the entire data set is called collective outlier. An individual data point in a collection may not be outlier. Usually data points are related in collection. Finding subsequence as anomaly in time series data set, finding sub regions as anomaly in spatial imaginary data set or finding sub graph as anomaly in graph data set are examples of collective outliers [2]

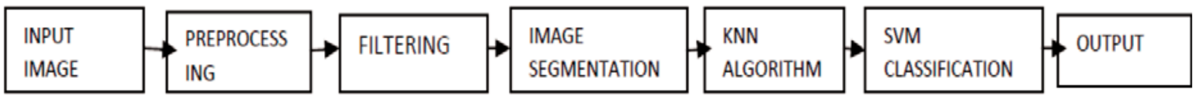


**Figure 1**

## 1.1. Proposed system

In this proposed system first we are going to filter the image by using Gaussian bilateral filter algorithm .In this we will add Gaussian noise to it so that by adding Gaussian noise the image will not become more blur .if we have directly applied filtering it will become more blur. Then we have MSE and PSNR for the image MSE(mean square error) and PSNR(peak signal noise ratio) are two error metrics used to compare various image compression quality .then we have used K nearest neighbor algorithm is used for segmentation of image and to detect outlier from that. Then we have used SVM(support vector machine) is used for the classification of the image based on the tumor[5]

## 1.2. About proposed system

**Gaussian bilateral filter:** Gaussian bilateral filter is used for the filtering of image .filtering is one of the important factor in image processing Gaussian bilateral filter is non linear edge preserving and noise reducing smoothing filter for images[6]. In this the intensity values of each pixel is replaced by weighted average of intensity values from nearby pixels

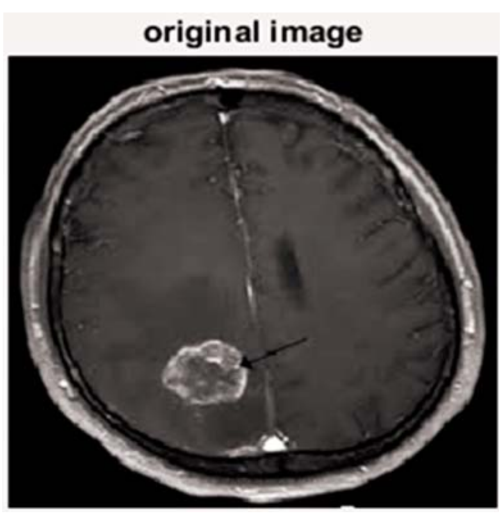


**Figure 2**

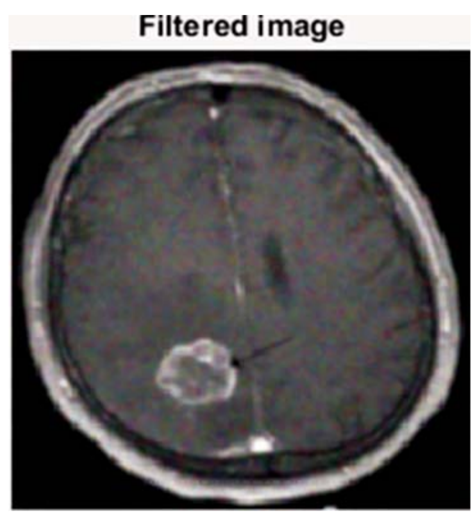


**Figure 2**

## 1.3. K nearest neighbor (Knn)

K nearest neighbor is considered as lazy learning algorithm that used for classification of datasets based on their similarity K stands for number data sets items used for the classification. In a simple way if you want to explain Knn algorithm is it will store all available cases and classifies new cases based on their similarity[6]

In this by KNN algorithm we will detect outliers from that by using nearest neighbor distance[7]

**Algorithm 1 Outlier Detection Algorithm**

1. Begin
2. Grayscale image we have taken
3. In this grayscale image we have to select one pixel In which K is constant K = any value
4. Select nearest neighbor pixel to it CC1, CC2, CC3, CC4
5. CC1, CC2, CC3, CC4 will select nearest neighbor pixel to it
6. Nearest pixel is not considered as outlier and pixel which is far is considered as outlier
7. End


**Figure 4**

## 1.4. SVM Classifier

Support vector machine is classifier which is used to classify data on their types.

In this paper we have used Svm classifier for the classification of images based on their tumor.[8] In this we have two tumor benign and malignant by using Svm classifier we will classify images based on their tumor whether it is benign or malignant tumor[9]

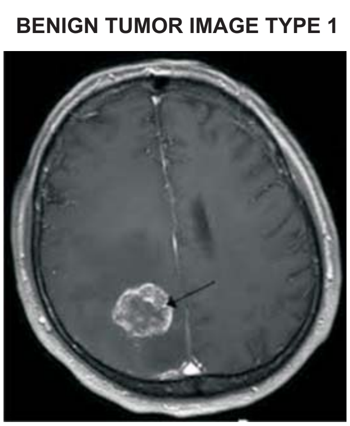


**Figure 5**

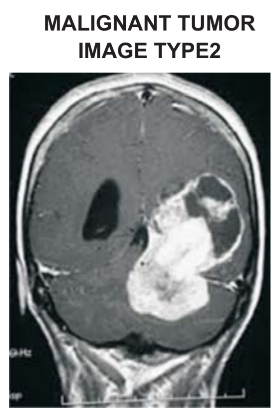


**Figure 6**

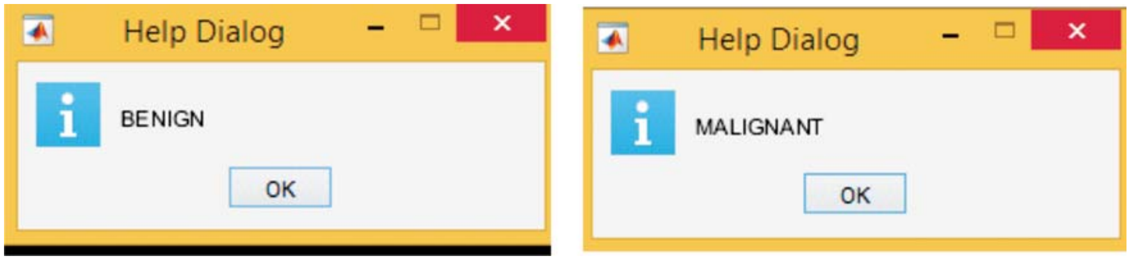


**Figure 7**

## 2. RESULT

**Table 1**

| *Image* | *Contrast* | *Correlation* | *Energy* | *Homogeneity* | *Mean* | *Standard Deviation* |
|---|---|---|---|---|---|---|
| 1 | 0.2333 | 0.1284 | 0.7491 | 0.9308 | 0.0019 | 0.0898 |
| 2 | 0.2717 | 0.0931 | 0.7686 | 0.9338 | 0.0024 | 0.0898 |
| 3 | 0.2272 | 0.1326 | 0.7439 | 0.929 | 0.0043 | 0.0897 |
| 4 | 0.2442 | 0.1007 | 0.7409 | 0.9263 | 0.0032 | 0.0898 |
| 5 | 0.2033 | 0.1126 | 0.7554 | 0.9331 | 0.0019 | 0.0898 |
| 6 | 0.2558 | 0.0895 | 0.7557 | 0.9314 | 0.0025 | 0.0898 |
| 7 | 0.2155 | 0.0951 | 0.7378 | 0.9274 | 0.0028 | 0.0898 |
| 8 | 0.2925 | 0.1584 | 0.7588 | 0.933 | 0.0057 | 0.0896 |
| 9 | 0.2341 | 0.1321 | 0.753 | 0.9315 | 0.0035 | 0.0897 |
| 10 | 0.2689 | 0.0977 | 0.7861 | 0.941 | 0.000687 | 0.0898 |
| 11 | 0.2433 | 0.1294 | 0.7606 | 0.9344 | 0.0034 | 0.0897 |
| 12 | 0.2272 | 0.1326 | 0.7439 | 0.929 | 0.0043 | 0.0897 |
| 13 | 0.275 | 0.118 | 0.7688 | 0.9346 | 0.0046 | 0.0897 |
| 14 | 0.2272 | 0.0908 | 0.7522 | 0.9308 | 0.0034 | 0.0897 |
| 15 | 0.2517 | 0.0734 | 0.7402 | 0.9267 | 0.0035 | 0.0897 |
| 16 | 0.2439 | 0.1072 | 0.731 | 0.9246 | 0.0046 | 0.0897 |
| 17 | 0.2925 | 0.1584 | 0.7588 | 0.933 | 0.0057 | 0.0896 |
| 18 | 0.2745 | 0.1095 | 0.7549 | 0.9308 | 0.0054 | 0.0897 |
| 19 | 0.2161 | 0.1382 | 0.7548 | 0.9325 | 0.0025 | 0.0898 |
| 20 | 0.2786 | 0.1427 | 0.7604 | 0.9321 | 0.0053 | 0.0897 |

**Table 2**

| *Image* | *Entropy* | *RMS* | *Variance* | *Smoothness* | *Kurtosis* | *Skewness* | *IDM* |
|---|---|---|---|---|---|---|---|
| 1 | 2.6632 | 0.0898 | 0.0081 | 0.8778 | 7.2707 | 0.6117 | 0.0366 |
| 2 | 3.2698 | 0.0898 | 0.0081 | 0.8974 | 7.9567 | 0.8862 | 0.4926 |
| 3 | 3.6046 | 0.0898 | 0.008 | 0.9406 | 5.9972 | 0.5218 | 0.37 |
| 4 | 3.5797 | 0.0898 | 0.008 | 0.9234 | 6.2735 | 0.6332 | 0.5257 |

| *Image* | *Entropy* | *RMS* | *Variance* | *Smoothness* | *Kurtosis* | *Skewness* | *IDM* |
|---|---|---|---|---|---|---|---|
| 5 | 3.6549 | 0.0898 | 0.008 | 0.8783 | 5.8117 | 0.3408 | 1.001 |
| 6 | 3.0756 | 0.0898 | 0.0081 | 0.904 | 7.7971 | 0.5774 | 0.2601 |
| 7 | 3.6283 | 0.0898 | 0.008 | 0.9132 | 5.3238 | 0.323 | 1.0419 |
| 8 | 2.6622 | 0.0898 | 0.008 | 0.9551 | 13.0402 | 1.3124 | 1.2778 |
| 9 | 3.1562 | 0.0898 | 0.008 | 0.9291 | 7.4848 | 0.5212 | 1.0392 |
| 10 | 2.7465 | 0.0898 | 0.0081 | 0.7186 | 10.9703 | 0.7365 | 0.119 |
| 11 | 2.9949 | 0.0898 | 0.0081 | 0.927 | 7.6801 | 0.6318 | 0.3816 |
| 12 | 3.6046 | 0.0898 | 0.008 | 0.9406 | 5.9972 | 0.5218 | 0.37 |
| 13 | 3.029 | 0.0898 | 0.0081 | 0.9453 | 13.1839 | 1.0085 | 0.2863 |
| 14 | 3.6783 | 0.0898 | 0.008 | 0.927 | 5.5966 | 0.4004 | 1.0469 |
| 15 | 3.5239 | 0.0898 | 0.008 | 0.9284 | 6.522 | 0.4979 | 1.6524 |
| 16 | 3.5484 | 0.0898 | 0.0081 | 0.9446 | 6.5235 | 0.6204 | 0.503 |
| 17 | 2.6622 | 0.0898 | 0.008 | 0.9551 | 13.0402 | 1.3124 | 1.2778 |
| 18 | 3.1085 | 0.0898 | 0.008 | 0.9523 | 11.1148 | 1.0231 | 0.6151 |
| 19 | 3.3156 | 0.0898 | 0.0081 | 0.9032 | 6.232 | 0.3121 | 0.5631 |
| 20 | 3.1943 | 0.0898 | 0.0081 | 0.9516 | 9.7318 | 0.9914 | 1.8546 |

## 3. CONCLUSION

In this conclusion we have concluded that image has been filtered by using Gaussian bilateral filter then MSE and psnr has been used for the efficiency of the image then Knn( K nearest neighbor) algorithm has been used for detection of outlier then feature extraction has been done and value of each feature extraction value has been stored then we have used SVM(support vector machine ) classifier for classification of the image based on their tumor.

## REFERENCES

[1] Charu C. Aggarwal, “Outlier Analysis”, Springer, 2013.
[2] Sreevidya S S, “A Survey on Outlier Detection Methods”, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (6), 2014.
[3] Karanjit Singh and Dr. Shuchita Upadhyaya, “Outlier Detection: Applications And Techniques”, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 3, January 2012.
[4] Jiawei Han, Micheline Kamber and Jian Pei, “Data mining Concepts and Techniques”, Third Edition, Morgan Kaufmann Series in Data management Systems.
[5] Shiblee Sadik and Le Gruenwald, “Research Issues in Outlier Detection for Data Streams”, SIGKDD Explorations, Volume 15, Issue 1.
[6] Rajendra Pamula, Jatindra Kumar Deka, Sukumar Nandi “An Outlier Detection Method based on Clustering” Second International Conference on Emerging Applications of Information Technology
[7] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander. Lof: identifying density-based local outliers. SIGMOD Rec., 29(2):93–104, 2000.
[8] M. Ester, H.-P. Kriegel, and X. Xu. A database interface for clustering in large spatial databases. In Proceedings of 1st International Conference on Knowledge Discovery and DataMining (KDD-95), 1995.
[9] A.Ghoting, S. Parthasarathy, and M. Otey. Fast mining of distance-based outliers in high-dimensional datasets. Data Mining and Knowledge Discovery, 16(3):349–364, June 2008.
[10] S. Guha, R. Rastogi, and K. Shim. CURE: An efficient clustering algorithm for large databases. SIGMOD Rec.,27(2):73–84, 1998.