A Novel technique for the Detection of Mixed Noise in Medical Images using Datamining

Er. Renuka Gujral[1] and Er. Nishi[2]

[1,2]Computer Science & Engineering Department,
[1,2]DAV University, Jalandhar, India
[1]renukagujral@gmail.com, [2]nishi.02.bti@gmail.com

Abstract

In Digital image processing; many researches have been done on image denoising so far. Nowadays, the noise detection from an image is the most challenging task. Though, the various algorithms introduced for the detection of noise type from a noisy image, but these algorithms work only for detection of single type of noise. To overcome the limitation of the previous built algorithms, we investigate the data mining technique called Support Vector Machine. The SVM is a powerful supervised learning method which is to be used for the detection of mixed noise models. Broadly, this technique detects the different types of noise from a mixed noise image; noise can be either single or mixed type of noise. The different parameters have combined to describe the properties of these different noise models so as to perform the detection. The detecting algorithm has been achieved by applying the SVM on the training dataset of different medical images and further extensive tests are performed on the test dataset for detection of each noise type model. This detection technique clearly outperforms various techniques with the high accuracy of results for different proposed noise models.

Keywords: Detection of noise type, Mixed Noise Models, Datamining, Support Vector Machine, Training dataset, Test dataset, Multiclass SVM

1. Introduction

Datamining aims to find a model which describes and distinguishes data classes or concepts for the purpose of predicting the class of objects whose class label is unknown and which yields to the detection of such particular data class [2]. Several studies have reported that SVM (Support Vector Machine) is one of the most prominent algorithm in the data mining area, delivers higher performance in terms of classification and object detection as well [1, 9]. However, here the SVM acts as a detecting algorithm which is used for solving the purpose of detection of noise type from an image. The Support Vector Machine was first proposed by Vapnik. It is a powerful machine method developed from statistical learning and has made significant achievements in different fields [7] such as Digital Image processing. Digital images are used in various scientific researches, medical sciences, astronomy, satellite television, *etc.* These digital images are transmitted through signals where the chances of noise occurrence arise. Thus, the images often get corrupted with the different types of noises while transmission, data acquisition, and storage systems [10]. This noise actually decreases the quality of noise. The two different types of noise occur mostly are Additive noise (Salt and pepper or Gaussian noise) and Multiplicative noise (Speckle noise.)[11] In image processing, the image restoration method is being used to remove the noise from an image so that the quality of an image is sustained [12].There are certain conventional filtering techniques to remove only one type of noise present in an image [20, 17]. However, there may a case arises when both the two types of noises are present in an image [18]. There is no such approach to detect and remove both types of noise simultaneously. So, it is important to develop an

approach which removes both the type of noise from an image simultaneously. Generally, Image denoising method follows two steps: noise detection and noise removal [4]. To build such denoising algorithms firstly there is a need to detect the mixture of noise present. The overall research of this paper is to first detect the mixed noise type present in an image. In 2007, Yixin Chen *et. al.,* has proposed the single pattern classification based on noise Identification. The basic idea for classification is to take some noise samples of each noise type, and extracts some properties using statistical features for noise type identification [21]. In this paper the noise is identified with the least number of parameters and needs deeper analysis in the classification of pattern. Many of the detection has been done through the Support Vector machine, For example, Xiaofu *et. al.,* 2007 proposed the Fake Iris detection technique based on statistical texture analysis using SVM [19]. This technique analyzes the properties of the image pixels using different parameters based on GLCM (gray level co-occurrence matrix). The SVM is used to characterize the class for good classification performance in high dimensional space. Bernd Heisele, 2010 have used the SVM approach for the face detection [6]. This technique address first by first locating the facial components, extract and combine them into a single feature vector machine called SVM. P.F.Felzenszwalb, 2010 has proposed an object detection system which is based on mixtures of multiscale deformable part models [16]. Jan Ruts, 2010 this paper gives the description of SVM technique which relates to the pattern classification. It describes how to deal with the binary classification problem as well as multiclass problem. It defines the practical or the steps involved during the implementation of SVM in different fields. Recently few researches have been done on noise diagnosis Wu, G-C. *et. al.,* 2014 develops a datamining approach for noise type diagnosis and a fuzzy filter is designed which improves the quality of noise corrupted images[22]. In this paper he demonstrates the proposed algorithm which is used to detect the two different noises, *i.e.,* salt and pepper and Gaussian noise from an image with the mixed noise present. Overall, a novel noise classifier using data mining techniques and fuzzy median-mean filter for removing complex noise from corrupted images was proposed. This research work has been done only on one type of additive noise diagnosis[17]. So, to advance this research, this paper introduces the novel technique which simultaneously detects the different noise models with different noise types additive as well as multiplicative noise. The noise identification is desirable in mixed noise for the further research in generating superior denoising algorithms. The better denoising technique can only be proved if we have better noise identification techniques.

2. Proposed Noise Models

The different noise models have been modeled with the different types of noise. We have considered three different types of noises are:

- Salt and pepper Noise
- Gaussian Noise
- Speckle Noise

The additive noises are Salt and pepper and Gaussian noise and the multiplicative noise is Speckle noise. In the research work, the different noise models where each type of model is a mixture of two types of noises. The different models which are modeled as in the following diagram:
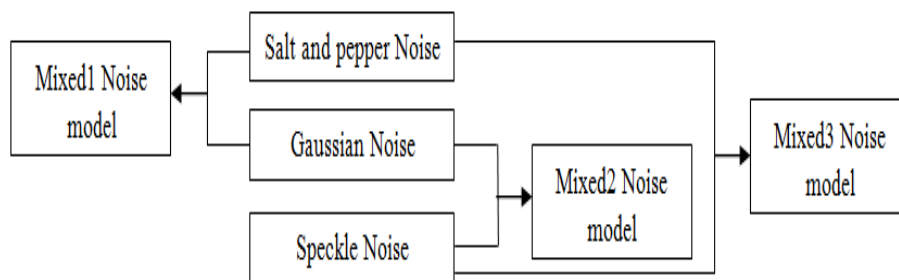
Figure 1. The Different Noise Models are Proposed with Different Types of Noises

These noise models are applied on an M×N image $I(i,j)$. The original image is degraded first, with the Mixed1 noise model. In the Mixed1 noise model, firstly the salt and pepper noise. This model is prescribed in following equations:

$$D_1(i,j) = I(i,j) + Sp(i,j), 1 \le i \le M, 1 \le j \le M \tag{1}$$

This degraded image is further degraded by the other type of additive noise that is Gaussian noise

$$D''_1(i,j) = D_1(i,j) + G(i,j), \ 1 \le i \le M, 1 \le j \le M \tag{2}$$

Second, the Mixed2 noise model is generated by contaminating the original image $I(i,j)$ with then Gaussian noise $G(i,j)$.

$$D_2(i,j) = I(i,j) + G(i,j), \ 1 \le i \le M, 1 \le j \le M \tag{3}$$

$$D''_2(i,j) = D_2(i,j) + S(i,j), \ 1 \le i \le M, 1 \le j \le M \tag{4}$$

Here, a second equation considers the multiplicative noise (Speckle noise) which is multiplicative in nature, thus is simply multiplied to an original value. The similar equations have been generated for the Mixed3 model.

$$D_3(i,j) = I(i,j) + Sp(i,j), \ 1 \le i \le M, 1 \le j \le M \tag{5}$$

$$D''_3(i,j) = D_3(i,j) + S(i,j), 1 \le i \le M, 1 \le j \le M \tag{6}$$

These all equations are involved in this technique to accomplish the system of proposed technique.

3. Support Vector Machines

The main aim of SVM is to separate the two different classes with optimal separating distance between hyperplanes on the basis of data points that are placed at the edge of class descriptors. These data points are called support vectors [3]. The data points other than the support vectors are discarded. Thus, it gives the optimal hyperplane with small training sets and high accuracy results. The SVM solves both the linear and nonlinear problems [9]. Basically, the SVM maps the original data from the input space into higher dimensional feature space so that it can easily separate the two classes. Consider a supervised binary classification where the training data are represented as $\{x_i \ y_i\}, i = 1, 2 \dots N$ and $y_i \in \{-1, +1\}$, where N is the number of training samples $y_i = +1$ for classes ω_1 and $y_i = -1$ for classes ω_2. Suppose two classes are linearly separable. This means it is separated by at least one hyperplane which is defined by a vector ω with a bias ω_o, where there is no error [13]. The equation is as:

$$f(x)\omega . x + \omega_0 = 0 \tag{5}$$

To find the hyperplane, ω and ω_0 should be estimated as:

$y_i(\omega.x_i + \omega_0) \geq +1$ for class $\omega_1 (y_i = +1)$ and $y_i(\omega.x_i + \omega_0) \leq -1$ for classes $\omega_2 (y_i = -1)$

These two equations combine and give the one equation as:

$y_i(\omega.x_i + \omega_0) -1 \geq 0$ (6)

Though there are many hyperplanes that easily separates the two classes, but there is only one hyperplane which is optimal. Now the goal is to find that optimal hyperplane which gives the maximum margin between the classes. To find this hyperplane, the support vectors must be defined [5]. The support vectors lie on the two hyperplane which are parallel to the optimal and written as:

$\omega.x_i + \omega_0 = \pm 1$ (7)

The margin between the hyperplane ω and ω_0 is given by $\frac{2}{||\omega||}$. The optimal hyperplane can be solved by solving the optimization problem:

Minimize $= \frac{1}{2}||\omega||^2$ (8)

Implies to $y_i(\omega.x_i + \omega_0) - 1 \geq 0$ where $i = 0, 1 \dots N$.

By Lagrangian formulation the above equation can be represented as:

Maximize $\sum_{i=1}^{N} \lambda_i - \frac{1}{2}\sum_{i,j=1}^{N} \lambda_i\lambda_j \, y_i y_j (x_i \cdot x_j)$ (9)

Implies to $\sum_{i=1}^{N} \lambda_i \, y_i = 0$ and $\lambda_i \geq 0$, i=1, 2…N where λ_i are the Lagrangian multipliers. The optimal hyperplane function becomes:

$f(x) = \sum_{i=s}^{N} \lambda_i \, y_i (x_i x) + \omega_0$ (10)

where S is the subset if training samples that correspond to non-zero Lagrangian multipliers. Here, the training vectors are the support vectors.

There are such cases where the classes are nonlinear separable where the equation (6) does not satisfy. So, to solve such cases, a cost function is to be formulated so as to combine the formulation of margin and minimization of error criteria, using the set of variables called slack variables (ξ). This cost function is defined as:

Minimize $J(\omega, \omega_0, \xi) = \frac{1}{2}||\omega||^2 + C\sum_{i=1}^{N}\xi_i$ (11)

Implies to $y_i(\omega.x_i + \omega_0) \geq 1 - \xi_i$

This implies a generalization of this method to the nonlinear discriminant function. The mapping of the input space to the high dimensional spaces increases the complexity to the problem. According to the Mercer's theorem, the inner product of the vectors in the mapping space is expressed as the function of the inner product of the corresponding vectors in the original space.
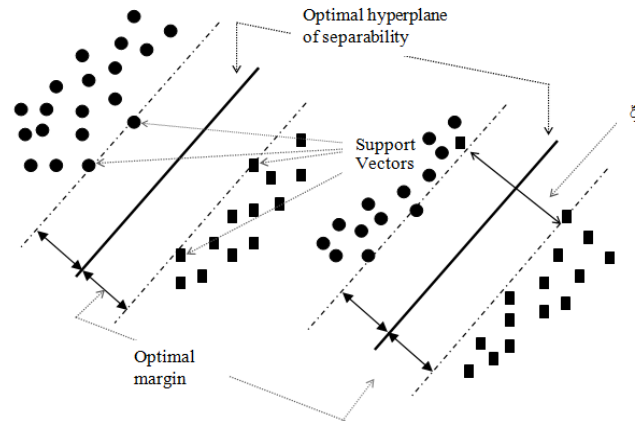
Figure 2. Left: the case of linear separable class. Right: the case of nonlinear separable class. ξ_i Measures the Error of Hyperplane Fitted

The inner product can be expressed as:

$$\phi(x)\phi(y) = K(x,z) \tag{12}$$

where the $K(x,z)$ is called a Kernel function. The kernel function is used in training without knowing the explicit form of ϕ.

The dual optimization problem is now represented as:

$$\text{Maximize } \sum_{i=1}^{N} \lambda_i - \frac{1}{2}\sum_{i,j=1}^{N} \lambda_i \lambda_j \, y_i y_j K(x_i \cdot x_j) \tag{13}$$

Implies to $\sum_{i=1}^{N} \lambda_i \, y_i = 0$ and $\lambda_i \geq 0$, i=1, 2… N

The final result of the SVM classifier would be:

$$f(x) = \sum_{i=s}^{N} \lambda_i \, y_i \mathrm{K}\,(x_i x_j) + \omega_0 \tag{14}$$

4. Multiclass Support Vector Machine

The SVM method was designed for two class problems. SVM can also be applied for multiclass problem. There are two such approaches for multi-class problem. The basic idea for multiclass problem is to reduce the multi-class into the set of binary set of variables so that the SVM can be applied. The first approach is "one against all". In this, a set of binary classifiers is trained and select each class separately from all others. Then each data object is classified in the class for which the largest decision value is determined. With this method N SVMs is trained where N is the number of classes and N decision functions are made. The second approach is "one against one". In this, a series of classifiers is applied to each pair of classes; with the most commonly computed class kept in each object. The max-win operator is used to determine to which class the object will be finally assigned. The application of this method requires N (N-1) /2 machines to be applied. Thus, this approach is applied according to the multi-class problem [8].

5. Methodology Of Proposed Technique

The proposed technique involves such method to process this whole system model for the noise diagnosis. The main issues regarding this technique are discussed in the flowchart. The following steps are involved as in the given figure below:
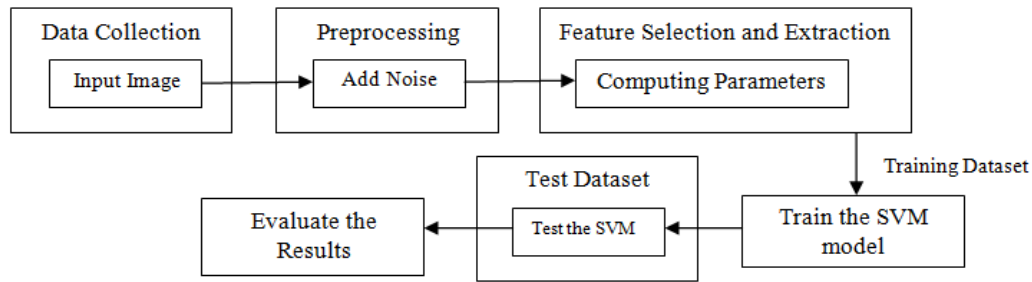
Figure 3. This Flowchart of the Proposed Technique

This flowchart addresses the each step in order to come to the final result[8]. The description of each step is explained as:

5.1. Data Collection

To define the dataset taken from the training, the different medical images have taken as a dataset.

5.2. Preprocessing

The preprocessing step involves the addition of noise into an image.

5.3. Feature Selection and Extraction

The main aim of the feature selection is to find the subset of variables which results in more accurate and compact models.

5.4. Train the SVM

To employ the SVM model according to the binary or multiclass problem for the further processing.

5.5. Test the SVM

Test the SVM on the test dataset for the evaluation of results.

5.6. Evaluate

The evaluation of the results and performance of the proposed technique and the confusion matrix is recorded according to the accuracy.

6. Experimental Analysis

In this section, the experimental dataset of different medical images has been used. The structure for experimental analysis involves the training and testing, which have been performed on a large set of medical images. During the training process, the 13 different types of parameters are selected for the extraction of feature of each class to improve the accuracy of the model for detection of that particular class to which it belongs [14-15]. The selected parameters through which detection of each class is analyzed are Contrast, Correlation, Energy, Homogeneity, Mean, Standard Deviation, Entropy, RMS, Variance, Smoothness, Kurtosis, Skewness and IDM.
The following selected parameters can be described as:

6.1. Contrast

It is the difference in luminance or gray level values in an image. It creates a new gray color map that defines the equal intensity distribution. It measures the joint probability occurrence of the specified pixel pairs.

$$C = \sum_{i,j=1}^{N-1} P_{i,j}(i-j)^2 \tag{15}$$

where i and j are equal and $(i-j) = 0$. The pixels with these values are entirely similar to their neighboring pixels, so they are given a weight of 0. Contrast has the strong, resolving power and detectability of an image.

6.2. Correlation

In 2-D digital correlation, displacement is directly detected from digital images of the surface of an object. It measures the linear dependency of the pixel gray level I in relation to the neighbor of gray level j.

$$\text{Corr} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_{i,j} \left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{\sigma_i^2 \sigma_j^2}} \right] \tag{16}$$

6.3. Energy

Energy is defined as the sum of the squared elements in the gray level co-occurrence matrix. It is an Angular Second Moment and also called as uniformity.

The ASM equation is written as:

$$\text{ASM} = \sum_{i,j=1}^{N-1} P_{i,j}{}^2 \tag{17}$$

Energy is represented as:

$$\text{Energy} = \sqrt{ASM} \tag{18}$$

6.4. Homogeneity

It measures the homogeneity as it assumes the larger values for smaller gray tone differences in each pair of distribution of elements in the gray level co-occurrence matrix.

$$\text{Homogeneity} = \sum_{i,j=1}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \tag{19}$$

6.5. Mean

The gray level co-occurrence matrix mean measures the mean of all the neighboring pixels of the reference pixel with the gray level i. This is calculated as:

$$\mu_i = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} i(P_{i,j}), \mu_j = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} j(P_{i,j}) \tag{20}$$

6.6. Variance

The GLCM variance measures the dispersion and the mean of combinations of reference pixels and the neighborhood pixel. The variance is calculated as:

$$\sigma_i^2 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (P_{i,j})(i-\mu_i)^2, \sigma_j^2 = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (P_{i,j})(i-\mu_j)^2 \tag{21}$$

6.7. Standard deviation

The gray level co-occurrence matrix can be calculated from the variance equation, which is calculated as:

$$\sigma_i = \sqrt{\sigma_i^2}, \ \sigma_j = \sqrt{\sigma_j^2} \tag{22}$$

6.8. Entropy

It is the statistical measure of randomness which is used to determine the texture of the input image. The entropy H of an image can be calculated as:

$$H = -\sum_{k=0}^{M-2} p_k \, log_2(p_k) \tag{23}$$

where M is the number of gray levels and p_k is the probability associated with gray level k.

6.9. Smoothness

The measure of smoothness is calculated as:

$$R = 1 - \frac{1}{1+\sigma^2} \tag{24}$$

where the bounded measure $0 \le \sigma^2 \le 1$.

6.10. Kurtosis

It is the measure of 'peakedness' or outlier prone distribution of the data around the sample mean. The kurtosis formulae as:

$$k = \sum_{i=0}^{N-1}(i-\mu)^4 \, P(i) \tag{25}$$

6.11. Skewness

It is the measure of the lack of symmetry of data around the sample mean. The formulae to calculate the skewness as:

$$s = \sum_{i=0}^{N-1}(i-\mu)^3 \, P(i) \tag{26}$$

6.12. IDM (Inverse Difference Moment)

It is influenced by the homogeneity of an image. The IDM is calculated as:

$$IDM = \sum_{i,j=1}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \tag{27}$$

These selected parameters are used to extract some information from an image. Further, the testing is done on an each image using the SVM which yields to successful results.

7. Results and Discussions

The following are the results shown which are performed during the testing phase of the proposed technique are shown clearly:

| (a) | (b) | (c) |
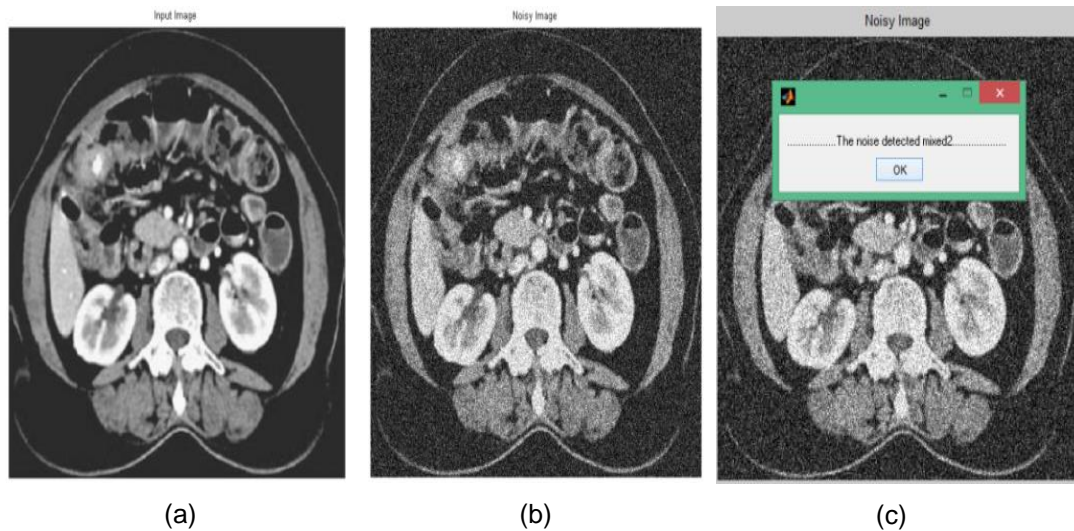
Figure 4. The Results for Detection of Mixed Noise Model: a) Input Image. b) The Noisy Image (Added Mixture of Two Unknown Types of Noises). c) Mixed Noise Model is Detected which is Present in an Image.

During the testing phase, In Figure 4, the input image is taken during the testing phase where the mixed type of noise is added first. Suppose Mixed2 Noise model is added manually to test and further testing is performed through implementing a technique for detecting this unknown type of noise. This proposed technique successfully detects that type of noise model present in an image.

Table 1. Confusion Matrix of the Detection of Different Noise Type from a Noisy Image

Noise Type	Training Images	Tested Images	Accuracy
Salt and pepper	100	100	100%
Gaussian	100	100	100%
Speckle	100	100	100%
Mixed1	100	100	100%
Mixed2	100	100	100%
Mixed3	100	95	95%

The Table 1 shows the confusion matrix of the detection of Different noise type models from mixed noisy image. This table shows the detection accuracy for the different types 100 medical images taken during the training and the testing phases. This proposed technique yields better performance with high accuracy for the detection of mixed noise from an image.

8. Conclusion

With SVM prominent properties, the proposed method shows better results than the other existing methods in real world data mining problems. The mixed noise diagnosis has been performed through SVM with higher accuracy. The SVM has a greater computational efficiency during training time. SVM also has the great feature where only a small training set is used to produce very good results, as the support vectors plays the main role during training. The future work can be done on this research work by increasing the types of input dataset and by increasing the corpus of experimental data.

References

[1] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", DataMining and Knowledge Discovery, vol. 2, no. 2, **(1998)**, pp. 121-167.

[2] S. Abe, "Support Vector for Pattern Classification", Springer Science and Business Media, United States of America, **(2005)**.

[3] T. Joachims, "Learning to Classify Text Using Support Vector Machines: Methods, Theory and Algorithms", Kluwer Academic Publishers Norwell , MA, USA, **(2002)**.

[4] I. Guyon, J. Weston, S. Barnhill and V. Vapnik, "Gene Selection for Cancer Classification using Support Vector Machines", Machine Learning, Springer., vol. 46, no. 1, **(2002)**, pp. 389-422.

[5] K. Kondo, M. Haseyama and H. Kitajima, "Efficient fixed-valued and random-valued impulse detection for accurate image restoration", Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis, Rome, Italy, **(2003)** September 18-20.

[6] K. G. Karibasappa, S. Hiremath and K. Karibasappa, "Neural Network Based Noise Identification in Digital Images", ACEEE International Journal on Network Security, vol. 2, no. 3, **(2011)**, pp. 28-31.

[7] K. He, X. C. Luan, C. H. Li and R. Liu, "Gaussian Noise Removal of Image on the Local Feature", Proceedings of Second International Symposium of Intelligent Information Technology Application, Shanghai, China, **(2008)** December 20-22.

[8] X. Han, J. Li , Y. Li and X. Xu, "A selective and adaptive image filtering approach based on impulse noise detection", Proceedings of Fifth World Congress on Intelligent Control and Automation, HangZhou, China, **(2004)** June 15-19.

[9] S. Deivalakshmi, S. Sarath and P. Palanisamy, "Detection and removal of Salt and Pepper noise in images by improved median filter", Proceedings of IEEE Conference on Recent Advances in Intelligent Computational Systems (RAICS), Trivandrum, **(2011)** September 22-14.

[10] S. Schulte, M. Nachtegael, V. D. Witte, D. V. D. Weken and E. E. Kerre, "A fuzzy impulse noise detection and reduction method", IEEE Transactions on Image Processing., vol. 15, no. 5, **(2006)**, pp. 1153-1162.

[11] C. Liu, W. T. Freeman, R. Szeliski and S. B. Kang, "Noise Estimation from a Single Image", Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, **(2006)** June 17-22.

[12] Y. Chen and M. Das, "An Automated Technique for Image Noise Identification using a Simple Pattern Classification approach", Proceedings of 50th Midwest Symposium on Circuits and Systems , Montreal, Que, **(2007)** August 5-8.

[13] X. He, S. An and P. Shi, "Statistical Texture Analysis-Based Approach for Fake Iris Detection Using Support Vector Machine", Advances in Biometrics, Springer Berlin Heidelberg, **(2007)**, pp. 540-546.

[14] B. Heisele, P. Ho and T. Poggio, "Face Recognitions with Support Vector Mechanics: global versus component based approach", Proceedings of the Eight IEEE International Conference on Computer Vision, Vancouver, Canada, **(2001)** July 09-12.

[15] P. F. Felzenszwalb, R. B. Girshick, D. McAllester and D. Ramanan, "Object Detection with Discriminatively Trained Part-Based Models", IEEE Transactions on Pattern Analysis and Machine Intelligence., vol. 32 , no. 9, **(2010)**, pp. 1627-1645.

[16] Y. Wang, G. Wua, G. S. Chen and T. Chai, "Data Mining based noise diagnosis and fuzzy filter design for image processing", Elsevier., vol. 40, no. 7, **(2014)**, pp. 2038-2049.

[17] C. Cortes and V. Vapnik, "Support-Vector Network", Machine Learning, vol. 20, no. 3, **(1995)**, pp. 273-297.

[18] A. Tozotsos and D. Argialas, "Support Vector Machine Classification for Object-Based Image Analysis", Edited T. Blaschke, S. Lang and G. J. Hay, Springer Berlin Heidelberg, **(2008)**, pp. 663-677.

[19] C. C. Chang and C. J. Lin, "LIBSVM: A Library for Support Vector Machines", ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 3, **(2011)**, pp. 1-27.

[20] J. Lutsa, F. Ojedaa, R. V. D. Plasa, B. D. Moora, S. V. Huffela and J. A.K. Suykensa, "A Tutorial on Support Vector Machine-based methods for classification problems in Chemometrics", Analytica Chimica Acta, Elsevier., vol. 665, no. 2, **(2010)**, pp. 129-145.

[21] O. Chapelle, V. Vapnik, O. Bousquet and S. Mukherjee, "Choosing Multiple Parameters for Support Vector Machines", Machine Learning, vol. 46, no. 1, **(2002)**, pp. 131-159.

[22] P. Mohanaiah, P. Sathyanarayana and L. GuruKumar, "Image Texture Feature Extraction Using GLCM approach", International Journal of Scientific and Research Publications., vol. 3, no. 5, **(2013)**, pp. 1-5.

Authors

Er. Renuka Gujral, she is a student of M.Tech in Department of Computer Science, DAV University, Jalandhar, and Punjab (INDIA). Completed B.Tech from RBCENT from Hoshiarpur, PTU, and Jalandhar. Have worked on many JAVA Projects. Several research papers have also been published in International journals.

Er. Nishi, she is working as assistant professor in the Department of Computer Science and Engineering at DAV University, Jalandhar, Punjab (INDIA). She has done her M.Tech in Computer Science and Engineering from Punjab Technical University, Jalandhar (INDIA). She has served NIT Jalandhar as an assistant professor in CSE Department and was a member of an antiragging squad. Her research areas include Digital Image Processing and Data Mining. She has various research publications on various international journals and has also presented several papers in various national and international conferences.