https://dx.doi.org/10.24001/eec.2.3.5 ISSN: 2456-2319

Detection and Classification of Fabric Defects in Textile using Image Mining and Association Rule Miner

Aswini Kumar Mohanty[1], Amalendu Bag[2]

[1]KMBB College of engg & CET, Bhubaneswar, Orissa, India
[2]Department Of Computer Science, KMBB College of engg & CET, Bhubaneswar – 752 054, Orissa, India

Abstract*—Image mining is concerned with knowledge discovery in image databases. It is the extension of data mining algorithms to image processing domain. Image mining plays a vital role in extracting useful information from images. In computer aided plant identification and classification system the image mining will take a crucial role for the fabric defect classification. Quality inspection is an important aspect of modern industrial manufacturing. In textile industry production, automate fabric inspection is important for maintain the fabric quality. In modern textile industry, Tissue online Automatic Inspection (TAI) is becoming an attractive alternative to Human Vision Inspection (HVI). HVI needs a high level of attention nevertheless leading to low performance in terms of tissue inspection. Automatic fabric inspection is valuable for maintenance of fabric quality. Defect inspection of fabric is a process which accomplished with human visual look-over using semi-automated way but it is labor prone and costly. Many sewing, knitting and dyeing units involve both manual and automated processes. Detecting faults in fabric manually, by human visual inspection is a tedious task. Its accuracy depends upon the skill of human operator and varies from person to person to address this difficulty a method is proposed for textile defect identification and classification based on image mining. The detection of local fabric defects is one of the most intriguing problems in computer vision. Texture analysis plays an important role in the automated visual inspection of texture images to detect their defects. Various approaches for fabric defect detection have been proposed in past and in this paper; we proposed a method based on texture analysis, association rule classifier, threshold segmentation and histogram to identify the textile defects. A feature extractor is designed based on Gray Level Co-occurrence Matrix (GLCM). An association rule miner is used as a classifier to identify the textile defects.*

Keywords— GLCM, HVI, TAI, Image Mining.

I. INTRODUCTION

Fabric defect detection is an important part of quality control in the textile industry. Usual methods of fabric inspection on the production line is done essentially by the worker on the circular knitting machine by introducing a light source in the middle of the circular product which enables the worker to detect the produced defects, and then stop the machine immediately. Stress and fatigue happens to the worker due to inspection in case of higher and quicker productivity. However, the method has been both time consuming and has lower accuracy of detection. Defect detection or inspection is a process identifying and locating defects. A fabric defect is a result of the manufacturing process. The textile industry is very concerned with quality. It is desirable to produce the highest quality goods in the shortest period of time possible [1]. Texture is also a main visual feature that refers to natural and fundamental surface properties of an object and their relationship with the surrounding atmosphere. The texture can be regarded as the visual look of a surface or material. Typically, textures and the analysis techniques associated with them are divided into two major categories with dissimilar computational approaches: the stochastic and the structural methods. Texture analysis is necessary for many computer image analysis applications such as classification, detection, or segmentation of images. In the other hand, defect detection is an important problem in fabric quality control process. At present, the texture quality identification is manually performed. Therefore, Tissue online Automatic Inspection (TAI) increases the efficiency of production lines and improve the quality of the products as well. Many attempts have been made based on three different approaches: statistical, spectral, and model based [2]. Texture is characterized not only by the gray value at a given pixel, but also by the gray value pattern in a neighborhood near the pixel. Generally, the term texture refers to repetition of texture primitives or fundamental texture elements called texels. A texel contains a number of image pixels, whose placement may be periodic, quasi-periodic, or random. The category of the texture determined by the repetitiveness of the Texel's and the texture analysis approach is decided [3]. In this research paper, we investigate the potential of the Gray Level Co-occurrence Matrix (GLCM) and association rule that used as a classifier to identify the textile defects. GLCM is a widely used texture descriptor]. The statistical features of GLCM are based on gray level intensities of the image. Such features of the GLCM are useful in texture recognition], image segmentation], image retrieval], color image analysis, image classification [4] [5], object recognition and texture analysis methods [6] [7] etc. The statistical features are extracted from GLCM of the textile digital image. GLCM is used as a technique for extracting texture features. The neural networks are used as a classifier to detect the presence of defects in textiles fabric products.

https://dx.doi.org/10.24001/eec.2.3.5 ISSN: 2456-2319

II. FABRIC DEFECTS

Fabric texture refers to the feel of the fabric. It is rough, velvety, smooth, soft, silky, and lustrous etc. The different textures of the fabric depend upon the types of weaves used. Textures are given to all types of fabrics, cotton, silk, wool, leather, and also to linen. Textile Fabric materials are used to prepare different categories and types of Fabric products in the textile industry. Natural fabric and synthetic fabric are the two different classifications of textile fabric. Synthetic fabrics are fairly new and have evolved with the continuous growth in textile industry [1, 8]. In a fabric, defects can occur due to:

A. Machine faults
B. Hole
C. Color bleeding
D. Yarn problems
E. Scratch
F. Poor finishing

III. IMAGE PROCESSING AND MINING

Image mining is the process of searching and discovering valuable information and knowledge in large volumes of data. Image mining draws basic principles from concepts in databases, machine learning, statistics, pattern recognition and 'soft' computing. Using data mining techniques enables a more efficient use of data banks of earth observation data. Image Mining is an extended branch of data mining that is concerned with the process of knowledge discovery concerning images. Image Mining deals with the extraction of image patterns from a large collection of images. In Image Mining, the goal is the discovery of image patterns that are significant in a given collection of images [8]. Image Mining deals with extraction of knowledge, image data relationship and other required patterns and uses ideas from image processing, image retrieval and machine learning, databases [9].

Image Processing represents an image which may be defined as a two-dimensional function, f(x,y), where x and y are spatial (plane) coordinates, and the amplitude of f at any pair of coordinates (x, y) is called the intensity or gray level of the image at that point. When x, y, and the amplitude values of f are all finite, discrete quantities, we call the image a digital image. The field of digital image processing refers to processing digital images by means of a digital computer. Note that a digital image is composed of a finite number of elements, each of which has a particular location and value. These elements are referred to as picture elements, image elements, pels, or pixels. Pixel is the term most widely used to denote the elements of a digital image [10].

IV. METHODOLOGY

The system of digital image mining may be presented schematically as shown in below Figure 1. The following operations are carried out during image quality improvement:

A. Image Acquisition
B. RGB to Gray Color Conversion
C. Image Enhancement (Thresholding, Segmentation)
D. Feature Extraction
E. Classification
F. Defect Identification

Block Diagram of the Proposed System:

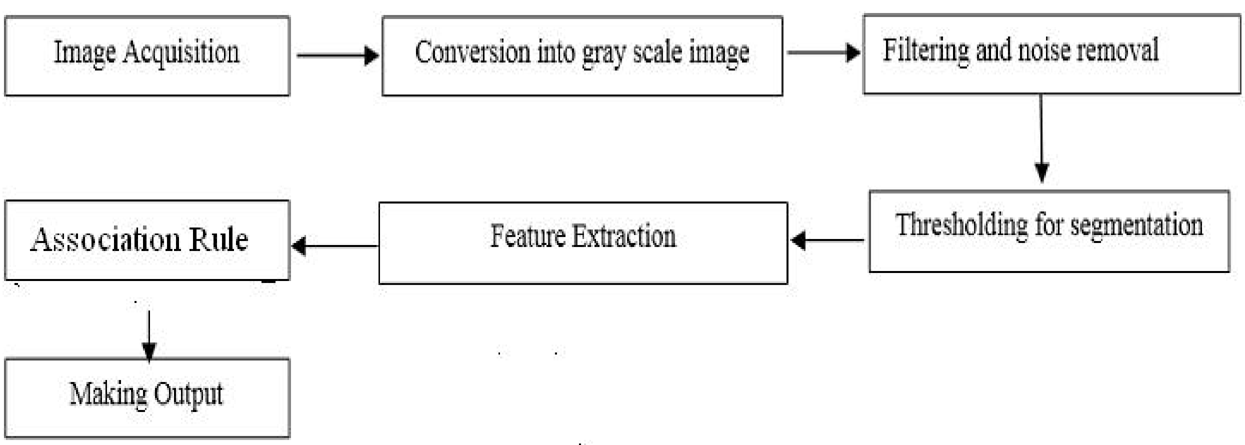


Fig.1: Block diagram of the textile defect inspection method

4. (a) Image Acquisition

Textile/fabric surface image is acquired by using the CCD camera from top of the surface from a distance adjusted so as to get the best possible view of the surface. Figures 2 show the quality of the acquired fabric images. The textile images under test are of size 256x256 (64KB). For proper imaging, uniform lighting system is to be maintained to avoid any illusive defect by virtue of light reflection properties falling on surface. Different Fabric/Textile Images Originally, the images are acquired at RGB color scale. The images then are converted to gray scale.

https://dx.doi.org/10.24001/eec.2.3.5 ISSN: 2456-2319

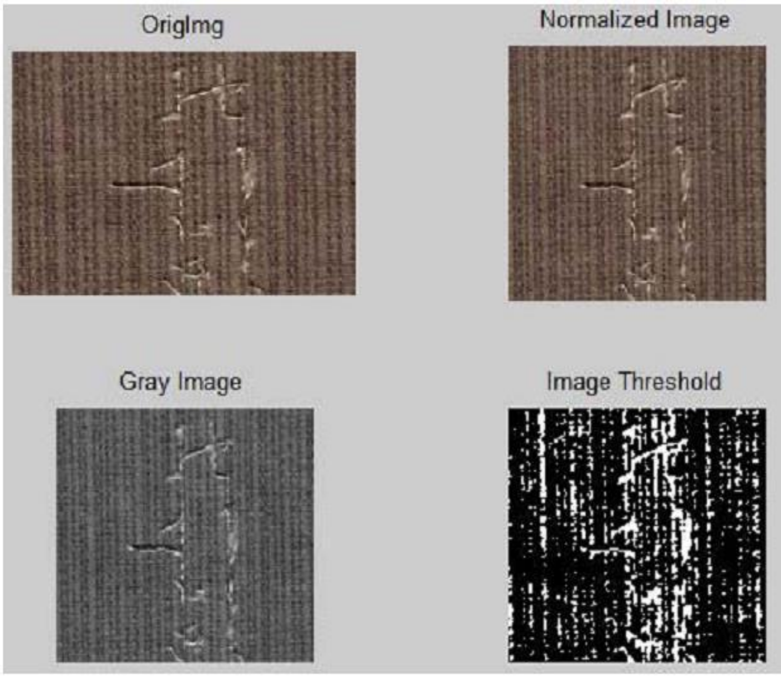


Fig.2: Preprocessing showing different images after being processed

4. (b) **RGB to Gray Color Conversion**

Weighted method or luminosity method

Weighted method has a solution to that problem. Since red color has more wavelengths of all the three colors, and green is the color that has not only less wavelength then red color but also green is the color that gives more soothing effect to the eyes.

It means that we have to decrease the contribution of red color, and increase the contribution of the green color, and put blue color contribution in between these two.

So the new equation that forms is:

New grayscale image = ((0.3 * R) + (0.59 * G) + (0.11 * B)).

According to this equation, Red has contributed 30%, Green has contributed 59% which is greater in all three colors and Blue has contributed 11%. A gray scale image (0 –255gray shades) is obtained.

4. (c) **Image Filtration and noise removal**

Since the leaf image for this study is taken by photographic camera and is converted from RGB to gray scale, it may contain some noise which could do difficulty to interpret. Therefore preprocessing is necessary to improve the quality of image and make the feature extraction phase as an easier and reliable one. A pre-processing; usually noise-reducing step [12, 13] is applied to improve image and contrast. Histogram equalization is a method in image processing of contrast adjustment using the image's histogram [16]. Through this adjustment, the intensities can be better distributed on the histogram. This allows for areas of lower local contrast to get better contrast. Histogram equalization accomplishes this by efficiently spreading out the most frequent intensity values. The method is useful in images with backgrounds and foregrounds that are both bright or both dark. In particular, the method can lead to better views of bone structure in x-ray images, and to better detail in photographs that are over or under-exposed. In mammogram images Histogram equalization is used to make contrast adjustment so that the image abnormalities will be better visible In this work the efficient filter (CLAHE) was applied..Contrast limited adaptive histogram equalization (CLAHE) method seeks to reduce the noise produced in homogeneous areas and was originally developed for medical imaging [14]. This method has been used for enhancement to remove the noise in the pre-processing of digital images [15]. CLAHE operates on small regions in the image called tiles rather than the entire image. Each tile’s contrast is enhanced, so that the histogram of the output region approximately matches the uniform distribution or Rayleigh distribution or exponential distribution. Distribution is the desired histogram shape for the image tiles. The neighboring tiles are then combined using bilinear interpolation to eliminate artificially induced boundaries. The contrast, especially in homogeneous areas, can be limited to avoid amplifying any noise that might be present in the image.

4. (d) **Image Thresholding**: The image after noise removal is brought under image into two colors i.e. we get a binary image with white as back ground and black as the object of interest. Segmentation algorithm is applied over the binary image to get the segmented patterns.

Basic Global Thresholding and segementation:

1) Select an initial estimate for T

2) Segment the image using T. This will produce two groups of pixels. G1 consisting of all pixels with gray level values >T and G2 consisting of pixels with values <=T.

3) Compute the average gray level values mean1 and mean2 for the pixels in regions G1 and G2.

4) Compute a new threshold value

T= (1/2)(mean1 +mean2)

5) Repeat steps 2 through 4 until difference in T in successive iterations is smaller than a predefined parameter T0.

4.(e) **Feature selection**

Features, characteristics of the objects of interest, if selected carefully are representative of the maximum relevant

https://dx.doi.org/10.24001/eec.2.3.5 ISSN: 2456-2319

information that the image has to offer for a complete characterization a lesion [17, 18]. Feature extraction methodologies analyze objects and images to extract the most prominent features that are representative of the various classes of objects. Features are used as inputs to classifiers that assign them to the class that they represent. In this Work Gray Level Co-Occurrence Matrix (GLCM) features are extracted.

Gray-Level Co-Occurrence Matrix (GLCM)

One of the simplest approaches for describing the texture is using a statistical moment of the histogram of the intensity of an image or region [19]. Using a statistical method such as co-occurrence matrix is important to get valuable information about the relative position of neighboring pixels of an image. Either the histogram calculation give only the measures of texture that carry only information about the intensity distribution, but not on the relative position of pixels with respect to each other in that the texture.

Given an image I, N × N, the co-occurrence, and the matrix P defined as:

$$P(i,j)=\sum_{x=1}^{N}\sum_{y=1}^{N}\begin{cases}1 & \text{if } I(x,y)=i \text{ and } I(x+\Delta_x, y+\Delta_y)=j\\ 0 & \text{otherwise}\end{cases}. \quad (1)$$

In the following, we present and review some features of a digital image by using GLCM. Those are Energy, Contrast, Correlation, and Homogeneity (features vector). The energy known as uniformity of ASM (angular second moment) calculated as:

$$\text{Energy: } \sum_{i=1}^{N}\sum_{j=1}^{N}P(i,j)^2 . \quad (2)$$

Contrast measurements of texture or gross variance, of the gray level. The difference is expected to be high in a coarse texture if the gray scale contrast is significant local variation of the gray level. Mathematically, this feature is calculated as:

$$\text{Contrast: } \frac{1}{(N)^2}\sum_{i=1}^{N}\sum_{j=1}^{N}(i-j)^2 p(i,j). \quad (3)$$

Texture correlation measures the linear dependence of gray levels on those of neighboring pixels (1). This feature computed as:

$$\text{Correlation: } \frac{\sum_{m=1}^{N}\sum_{n=1}^{N} mnp(m,n)-\mu_x\mu_y}{\sigma x \sigma y} \quad (4)$$

$$\text{where } \mu_x=\sum_{m=1}^{N}M\sum_{n=1}^{N}p(m,n), \quad \mu_y=\sum_{m=1}^{N}N\sum_{n=1}^{N}P(m,n)$$

$$\sigma_x=\sum_{m=1}^{N}(m-\mu_x)^2\sum_{n=1}^{N}p(m,n)$$

$$\sigma_y=\sum_{n=1}^{N}(n-\mu_y)^2\sum_{m=1}^{N}p(m,n).$$

The homogeneity measures the local correlation a pair of pixels. The homogeneity should be high if the gray level of each pixel pair is similar. This calculated function as follows:

$$\text{Homogeneity: } \sum_{m=1}^{N}\sum_{n=1}^{N}\frac{P(m,n)}{(1+|m-n|)} . \quad (5)$$

4.(f) **Classification**

Data classification [20] is a two step process. In the first step, a model is built describing a predetermined set of data classes. The model is constructed by analyzing database tuples described by attributes. Each tuple is assumed to belong to a predefined class, as determined by one of the attributes, called the class label attribute. In the context of classification, data tuples are also referred to as samples or examples. The data tuples analyzed to build the model collectively form the training dataset. The individual tuples making up the training set are referred to as training examples and are randomly selected form the training dataset[21]. Typically, the learned model is represented in the form of classification rules, decision trees, neural networks or mathematical formulas. In the second step the model is used for classification. But first the predictive accuracy of the model is estimated on a separate test dataset. If the accuracy of the model were estimated based on the training dataset, the estimate could be optimistic since the learned model tends to overfit the data. Therefore, a separate test dataset is used. If the accuracy of the model is considered acceptable, the model can be used to classify future data tuples for which the class level is not known.

Level Adaptive Classifier

Basic Idea:

Almost all classifiers use Apriori Rule mining Algorithm to mine the rule set. Apriori is an algorithm that generates frequent rule items on a level wise manner. That is, at first, all rules with one antecedent are mined, and then all rules with two antecedents and so on. This can generate a large number of rules and the number of rules in each new level can grow in an exponential manner. There are datasets with more than 20 attributes, so a rule of 20 antecedents can only be generated if all the subsets of that rule, 2^20 are generated previously. So this can be intractable. Moreover, rules of larger number of antecedents are generally over fitting rules and do not yield good performance on test data. But there is no way to ascertain when to stop generating rules in any state-of-the-art classifier. So this classifier was an attempt to make the “max length of a rule” parameter adaptive. The target was to achieve efficiency without sacrificing accuracy.

Description:

The idea was to generate all rules of level 1 first just as in A Priori and construct a classifier using these rules and calculate the number of errors made by this classifier on the validation set. Then level 2 rules are generated and a new classifier is constructed using both level1 and level2 rules. Again, this classifier's performance on validation set is noted by recording the number of errors committed on validation set. The number of errors [23] in this case is compared with the number of errors with level1 rules. If the number of errors does not decrease, we can make the assumption that level 2 rules don't help too much. In this way, classifiers are constructed using rules of level 1, then level 1 and 2, then level 1 ,2 ,3 etc. When two or more classifiers of higher length rules perform worse than previous classifiers, we can convince ourselves that the new long rules are over fitting and we can safely discard them and stop rule generation at that phase(do not generate any more higher length rules).

Algorithm

L1=find_frequent_1_itemset(D);
Construct a classifier
Find number of errors on validation set
For (k=2; Lk-1!=empty ;k++)
{
Generate Lk = frequent_k_itemset(Lk-1);
Construct a classifier using rules generated so far
Find number of errors on validation set and compare with previous
number of errors
If number of errors increases
Break;
}
This classifier is used to classify future test instances

V. EXPEIMENTAL RESULT

In this paper we used association rule mining level adaptive classifier using image contents for the classification of defects in fabric. The average accuracy is 92 %. We have used the precision and recall measures as the evaluation metric for fabric defect classification. Precision is the fraction of the number of true positive predictions divided by the total number of true positives in the set. Recall is the total number of predictions divided by the total number of true positives in the set. The testing result using the selected features is given in table 5.1. The selected features are used for classification. For classification of samples, we have employed the freely available Machine Learning package, WEKA [24]. Out of 100 images in the dataset, 70 were used for training and the remaining 30 for testing purposes.

Table.5.1: Results obtained by proposed method

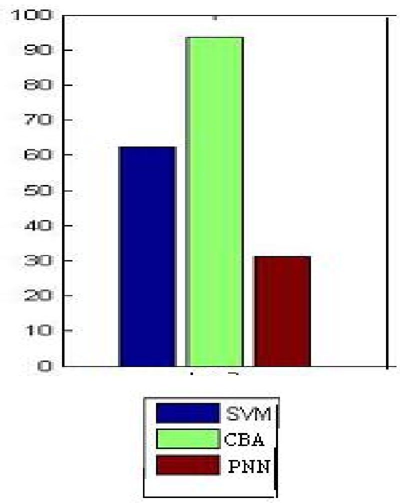
Non-Defected	**90%**
Defected	**94%**

The confusion matrix has been obtained from the testing part. In this case for example out of 50 actual defected images 3 images was classified as non-defected. In case of non-defected images 45 images are correctly classified and 5 images are classified as defected mages. The confusion matrix is given in Table 5.2.

Table 5.2: Confusion matrix

Actual	**Predicted class**	
	Defected	**Non-Defected**
Defected	**47**	**3**
Non-Defected	**5**	**45**

The following graph shows the comparative analysis of our method and various other methods.



VI. CONCLUSION

We studied and developed an efficient method of textile defects identification based on GLCM and Association rule miner which is a comprehensive methods in a uniform terminology, to define general properties and requirements of local techniques, to enable the readers to select the efficient method that is optimal for the specific application in detection of fabric defects images. The proposed method is tested with the best quality fabric images and the analysis classifications and detection has been sought. The descriptor of the textile image based on statistical features of GLCM is used as input to ARM classifier for recognition and classification defects of raw textile. Experimental results showed that the proposed method is efficient, and the recognition rate is 92% for overall testing including training. This study can take part in developing a computer-aided decision (CAD) system for Tissue online Automatic Inspection (TAI). In future work, various effective features will be extracted from the textile image used with other classifiers such as Reverse Association Rule. The algorithm uses simple statistical techniques in collaboration to develop a novel feature selection technique for defected fabric image analysis. The value of this technique is that it not only tackles the measurement problem but also provides a visualization of the relation among features. In addition to ease of use, this approach effectively addresses the feature redundancy problem. The method proposed has been proven that it is easier and it requires less computing time than existing methods.

https://dx.doi.org/10.24001/eec.2.3.5 ISSN: 2456-2319

REFERENCES

[1] Priya S, Kumar TA, Paul V. A novel approach to fabricdefect detection using digital image processing, Signal Processing, Communication, Computing and Networking Technologies (ICSCCN), 2011 internationalConference on IEEE, 2011, 228-232.

[2] Abdel Azim, G. and Nasir, S. (2013) Textile Defects Identification Based on Neural Networks and Mutual Information.International Conference on Computer Applications Technology (ICCAT), Sousse Tunisia, 20-22 January 2013

[3] E.M.Srinivasan, K.Ramar and A.Suruliandi,"Rotation Invariant Texture Classification using Fuzzy Local Texture Patterns", InternationalJournal of Computer Science and Technology, vol.3, Iss. 1, 2012.

[4] De Almeida, C.W.D., de Souza, R.M.C.R. and Candeias, A.L.B. (2010) Texture Classification Based on a Co-Occurrence Matrix and Self-Organizing Map. IEEE International Conference on Systems Man & Cybernetics,University of Pernambuco, Recife, 2010.

[5] Haralick, R.M., Shanmugam, S. and Dinstein, I. (1973) Textural Features for Image Classification. IEEE Transactionson Systems, Man, and Cybernetics, 3, 610-621.

[6] Srinivasan, G.N. and Shobha, G. (2008) Segmentation Techniques for ATDR. Proceedings of the World Academy ofScience, Engineering, and Technology, 36, 2070-3740.

[7] Tuceryan, M. (1994) Moment Based Texture Segmentation. Pattern Recognition Letters, 15, 659-667.

[8] Thilepa R, Thanikachalam M. A Paper on Automatic Fabric Fault Processing Using Image Processing Technique in Matlab, Signal & Image Processing: an International Journal (SIPIJ), 2010, 1(2).

[9] C. Ordonez and E. Omiecinski,"Discovering Association Rules Based on Image Content",01Proceedings of the IEEE Advances in Digital Libraries Conference (ADL'99), 1999.

[10] J. Salton, and M. J. McGill,"Introduction to Modern Information Retrieval",McGraw-Hill Book Company,1983.

[11] Anitha S, Radha V, Krishna PMHM, Enhanced Switching Median Filter For Denoising In 2D patterned Textile Image, International conference on modeling optimization and computing, Elsevier 2012; 38:3362–3372.

[12] Pisano ED, Gatsonis C, Hendrick E et al. "Diagnostic performance of digital versus film mammography for breast-cancer screening". NEngl J Med 2005; 353(17):1773-83.

[13] Wanga X, Wong BS, Guan TC. 'Image enhancement for radiography inspection". International Conference on Experimental Mechanics. 2004: 462-8.

[14] D.Brazokovic and M.Nescovic, "Mammogram screening using multisolution based image segmentation", International journal of pattern recognition and Artificial Intelligence, 7(6): pp.1437-1460, 1993

[15] Dougherty J, Kohavi R, Sahami M. "Supervised and unsupervised discretization of continuous features". In: Proceedings of the 12th international conference on machine learning.San Francisco:Morgan Kaufmann; pp 194–202, 1995.

[16] Yvan Saeys, Thomas Abeel, Yves Van de Peer "Towards robust feature selection techniques", www.bioinformatics.psb.ugent

[17] Gianluca Bontempi, Benjamin Haibe-Kains "Feature selection methods for mining bioinformatics data", http://www.ulb.ac.be/di/mlg

[18] Gonzalez, R.C. and Woods, R.E. (2008) Digital Image Processing. 3rd Edition, Prentice Hall, India.

[19] J. Han, J. Pei and Y. Yin. "Mining frequent patterns without candidate generation," In Proc. 2000 ACM-SIGMOD Int. Conf. on Management of Data (SIGMOD' 00), pages 1-12, Dallas, TX, May 2000.

[20] X. Yin and J. Han, "CPAR: Classification based on Predictive Association Rules," Proc. Of SIAM Int. Conf. on Data Mining (SDM'03), pp. 331-335, San Francisco, CA, 2003.

[21] E.Baralis, P. Gazza, "A lazy approach to pruning classification rules," Proc. IEEE Int. Conf. on Data Mining (ICDM'04), pages 35-42, 2002.

[22] Adriano Veloso, Wagner Meira Jr., Mohammed J. Zakib, "Lazy Associative Classification," Proc. IEEE Int. Conf. on Data Mining (ICDM'06)

[23] Holmes, G., Donkin, A., Witten, I.H.: WEKA: a machine learning workbench. In Proceedings Second Australia and New Zealand Conference on Intelligent Information Systems, Brisbane, Australia, pp. 357-361, 1994.