TONGUE IMAGE ANALYSIS FOR HEPATITIS DETECTION USING GA-SVM

Dr.N.Sandhya
CSE Department, VNRVJIET, Bachupally,
Hyderabad, Telangana 500090, India
sandhyanadela@gmail.com

Dr.M.Rajasekar
CSE Department, VNRVJIET, Bachupally,
Hyderabad, Telangana 500090, India
rajasekarm9@gmail.com

Abstract In this study, we proposed an automated methodology to analyze and detect and analyze hepatitis by using tongue images analysis based on genetic algorithm–support vector machine. There is a strong association in between the characteristics of tongue and human health diagnosis. GA-SVMs are used to establish a relationship between diseases like appendicitis, diabetes and characteristics of tongue. All the tongue samples are captured with the help of TDA-2 device. GA-SVMs are trained and tested with features like texture and chromatic information obtained from tongue image samples. GA-SVM classification accuracy of hepatitis is enhanced by obtaining unbiased features of tongue image samples. Proposed methodology achieved high accuracy by providing relevant data in training phase and removing irrelevant data. Tongue is very important organ in modern medical treatment as well as in Ayurveda medical treatment which is useful to detect and analyze diseases like appendicitis, pancreas, liver related diseases and diabetes. Experiments are performed on data set consisting of 1298 healthy and 2057 disease tongue samples. The accuracy of tongue images related to hepatitis is 93.74%.

Keywords: GA-SVM, appendicitis, diabetes, tongue and medical treatment.

1. Introduction

There is a strong relation in between texture, color of the tongue and lungs, heart, chest, neck etc. We can obtain significant information of human body from the analysis of tongue image as shown in Figure 1and each area in the tongue image represents the functioning of parts of human body. Neck, chest, lungs and heart functioning can be identified by tip of the tongue which is the region in between area9 and area10. Region in between AreaC and AreaD represents the diseases related to stomach and pancreas. EAA and EAD represent diseases related to liver and gallbladder as shown in Figure 1. Region in between Area1 and Area2 represents diseases related to abdominal organs as shown in Figure 1. This study detects and analyzes the hepatitis diseases with Genetic algorithm and support vector machine approach. Features like coating, shape, size and moisture plays vital role in the analysis of the disease and which are useful in diagnosis the working of major organs of human body. Ulceration or inflammations are identified by the red color of tongue image sample and weakness in the blood is diagnosed by the white color of the tongue image sample.

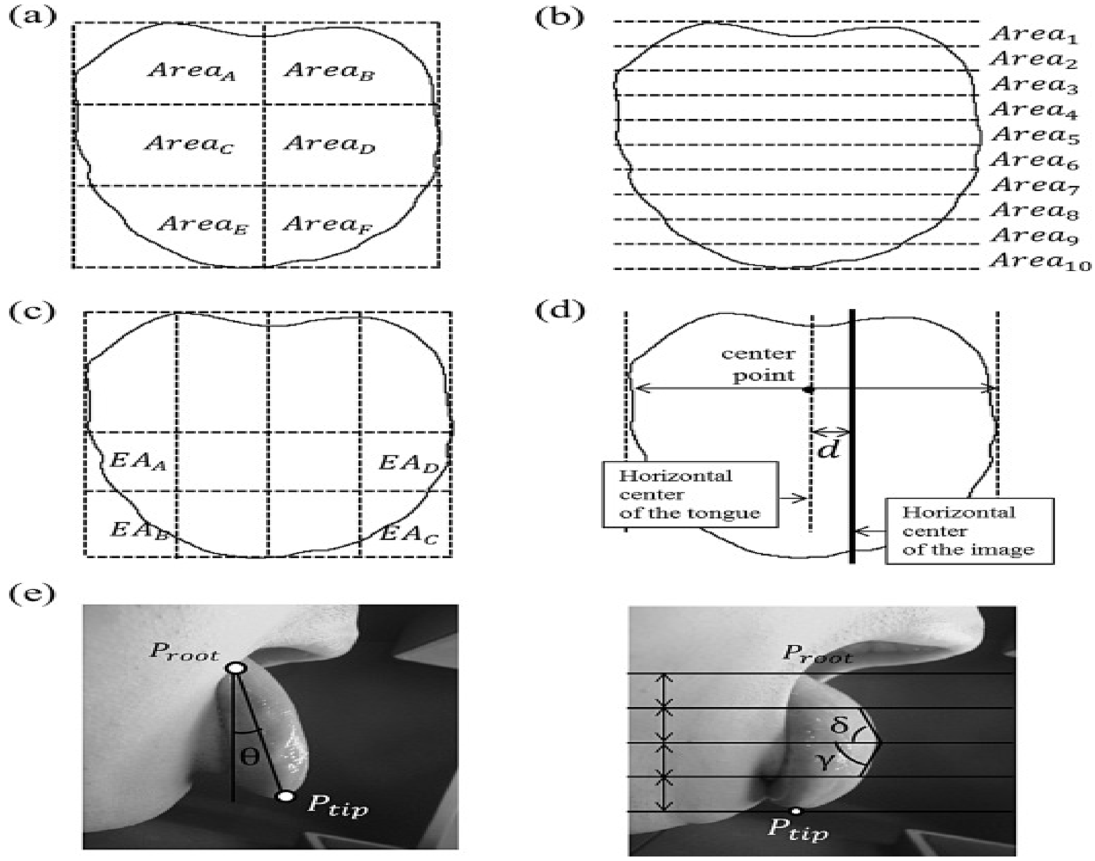


Figure1. Different areas of tongue

Diseases related to gallbladder and liver can be detected and analyzed on the basis of the yellow color of the tongue and blue color of the tongue represents diseases related to blood circulation. Robust and nice color represents the healthy tongue which means all the organs of the human body functioning properly. The main significance of tongue image analysis is it simplifies the pathological process in the cases of full of complex and contradictory diseases. This analysis not only plays an important role in diagnosis of diseases but also plays very important role in the treatment. Tongue image analysis based treatments are greatly influenced by illumination conditions and light sources. This study is useful in understanding syndromes which are not properly analyzed by modern medical scientist and western medical scientists. In this paper we proposed a model known as Tongue Computational Model (TCOM) which addresses the problems faced by hepatitis patients.

2. Literature Survey

Chiu et al. [1] identified several attributes of tongue to detect and analyze the diseases in which color of tongue image played major role. Xu, Su, Wang & Hu (2011) [2] used not only tongue color but also some qualitative and quantitative attributes to examine and analyze lung cancer. Li N.M proposed a methodology to detect and analyze the pancreatic cancer with the help of the shape of tongue image samples. Medical scientists from Western countries and East Asia used the tongue image based methodologies to diagnose and cure the diseases related liver and lungs [3]. NanJing and Jingui extensively used tongue image analysis in Chinese medical treatment methodology to detect and analyze ulcer related problems [4]. T. Han formulated a methodology in Japanese medical literature based on unified attributes of tongue image characteristics [5].

Odagu et al. [6] used various medical findings related to tongue sample image analysis in Kampo medical treatment for diabetes patient’s treatment. Zasla designed a reliable methodology to treat appendicitis patients based on intra and inter characteristics of tongue image samples [7]. Pang and Zing formulated an automatic frame work based on support vector machine and Bayes theorem to detect breast cancer in mammogram images [8]. C. C Chang et al. [9] designed a novel computerized medical system that detects and analyses disorders related thyroid problems [10].

C.H. Li et al. [11] developed a procedure to identify and classify cancer related tongue image samples based on color content. N. M Li et al. designed a system known as CICT to detect contemporary diseases which are not accurately diagnosed by western pathological procedures [11]. Zhu et al. [12, 13, 14] designed an automated system to identify and cure several diseases based on image processing methodologies. Chiu et al. [15, 16, 17] designed a methodology to detect lung cancer based on pattern recognition methods. Weng et al. [18, 19, 20] proposed method to detect brain tumor based on tongue Furring methodology.

3. GA SVM Methodology

The classification accuracy is improved by the combination of support vector machines (SVMs) and genetic algorithms. Feature selection can be performed by genetic algorithms and classification is performed by SVMs. The concept of fitness function of genetic algorithm is useful for extracting features of tongue image sample. The attributes of sample tongue images can be optimized by the objective function of the genetic algorithm. Prediction of various diseases is always being an issue in tongue image analysis research. Genetic algorithm is used to select optimal combination of parameters ε, c and kernel related variables γ and d.

The objective function of GA SVM Model is

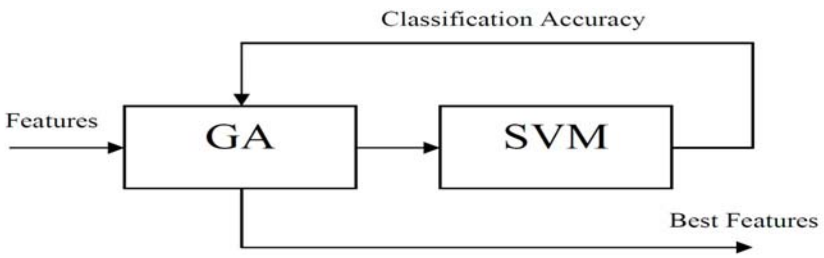
$$MinR(c) = \frac{1}{2}\|w\|^2 + c\frac{1}{N}M_{\varepsilon}(e_i, x_j)(1)$$

Hyper plane can be constructed by the following formula

$$W^* = \sum_{i=1}^{N}(\beta_i - \beta_i)k(y, y_i)(2)$$

Fitness function of GA is calculated by th following formula

$$\frac{1}{-\psi^2 N}\sum_{j=1}^{N}(e_j - x_j)^2(3)$$



$$f(x) = w.x + b = \sum_{j=1}^{k}(b_j - b_j^{\,*})k(y_i - y_j) + b(4)$$

Fig 2.Combination of GA and SVM to achieve the best features

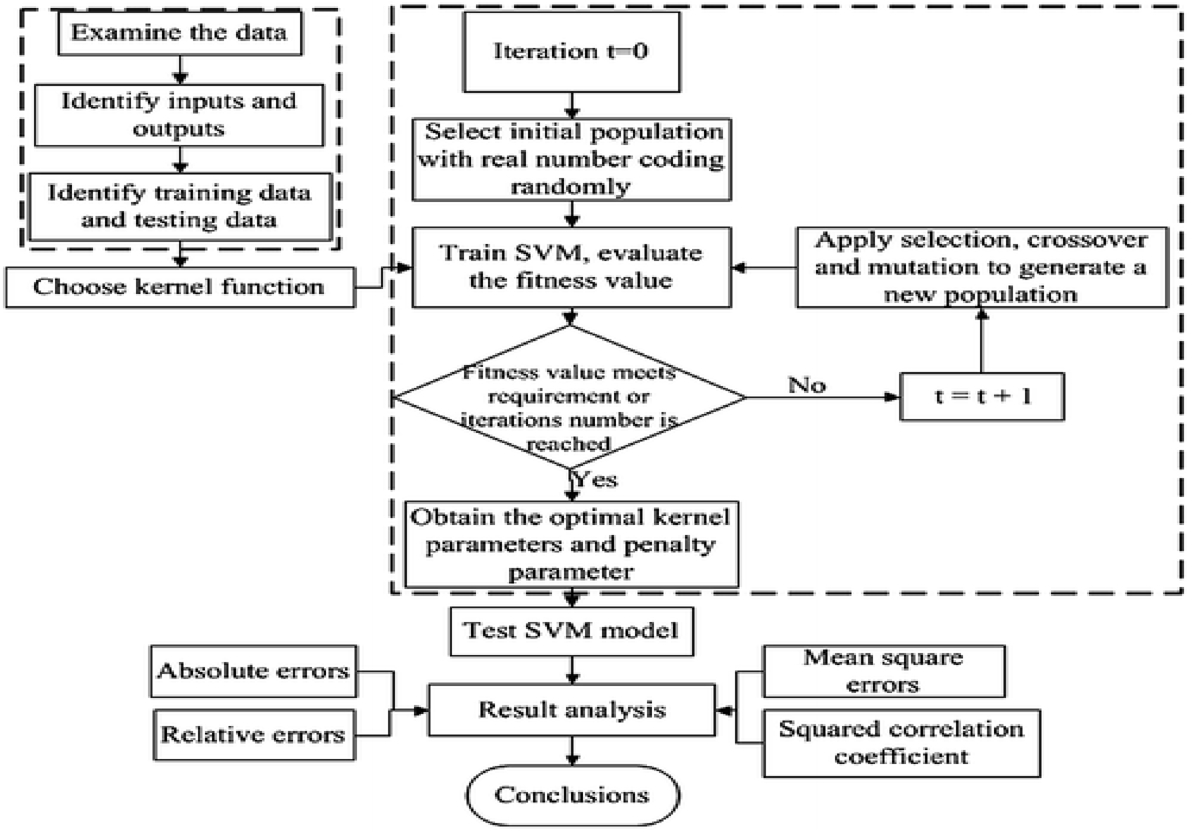


Fig 3. Flowchart depicting the GA-SVM model application process

3.1 Proposed GA SVM Methodology

Step 1. In the phase of initialization, an initial population of size 3355 tongue related images are selected in which 2057 images are unhealthy tongue image samples and 1298 healthy tongue image samples and selection of appropriate kernel and fitness parameters is performed.

Step 2. Calculate the fitness of each tongue image sample of hepatitis patient from constructed database

Step 3. New population is constructed by the following steps till it converges

i. Selection: Roulette wheel theorem is used to in the process of selection of chromosome related tongue images of hepatitis patients.

ii. Crossover: crossover probability is used to create offspring tongue image samples. In this process tongue image samples are paired by amortized process in which probability of each tongue image calculated.

Children = Par1 ± α(Par2 – Par1) where α is scaling factor which decides the crossover rate.

iii. Mutation: Local optimum problem is dealt with mutation function in which incremental technique is applied. Conditional probabilities are used to determine crossover and mutation rate of tongue image samples. In this study, crossover probability is 0.789 and mutation probability is 0.49.

iv. Accepting: Acceptability is decided based on the fitness value of the tongue image samples. This is useful to place new off springs in new population.

Step 4. In this old off spring tongue image samples are replaced by new offspring tongue image samples.

Step 5. This step indicates the stopping criteria. If equation 3 is satisfied, then stop the training process otherwise go to step 2 until it satisfy equation 2.

4. Applied Methodology and Results

In this study we designed database of 12,000 clinical tongue images collected from different hospitals in Hyderabad. All the tongue images of the database are collected by three charge coupled camera with D65 lights.

The tongue image sample of size 3355 is divided into two groups in which first group consists of 2057 unhealthy tongue images and second group consists of 1298 healthy tongue image samples. All the tongue images samples are classified into two 2 sets known as first one is training set and second one is testing set.

We can extract the features from tongue image by separating tongue body from its coating as part of analysis method. In this process both texture and color of tongue image and its coating is analyzed and detected. The main difference between tongue coating and tongue is function color information which may be in the form of grey, black, red, purple and white.

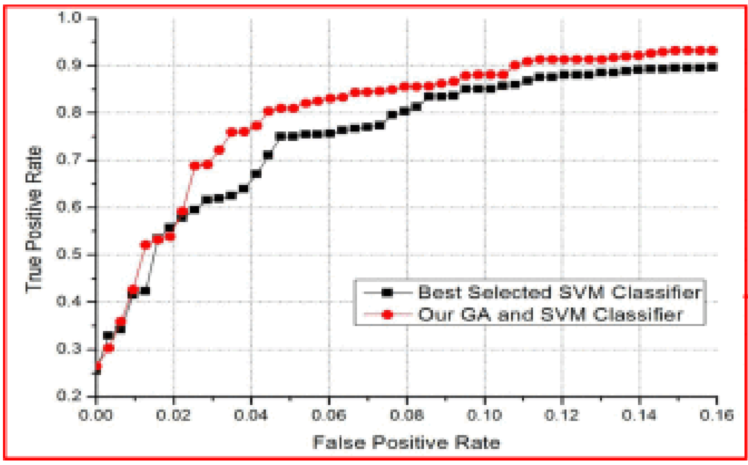
Proposed methodology obtained good results as it uses color information with Genetic algorithms and support vector machines. Tongue body and its coating are separated with merging algorithm and chromatic based threshold algorithms. Pixel average intensity values of tongue coating and its body are calculated based on red, blue and green. There are several types of tongues such as rough coating, tender tongue and greasy tongue in which tender tongue and greasy tongue has smooth textures. Tongue body and its coating are also well described by grey level scaling.

The texture feature and color feature are calculated from different groups of tongue image samples of hepatitis patients. The resultant tongue image features are normalized by using scaling transformation in the range between -1 and +1. Classification accuracy of tongue images with GA-SVM are significantly improved by reducing the number of dimensions by Principal component analysis. Genetic algorithm is to optimize the kernel variable g and penalty variable c.

Support vector machine classifier1 makes use of kernel of linear type, support vector machine classifier2 makes use of kernel of polynomial type and support vector machine classifier3 makes use of kernel type of radial basis function. ROC curve is plotted by taking false positive on X-axis and true positive rate on Y-axis. In ROC curve four methods are compared such as SVM Classifier1, SVM Classifier2, SVM Classifier3 and GA-SVM in which GA-SVM shows higher performance compare to other classifiers. This ROC curve is plotted for TTT=16.0 with various kernel parameters and penalty variable g. ROC curves describes that GA-SVM classifier has achieved a good true positive rate and lower false positive rate as shown in the Fig. 3. ROC curve in the Fig. 4 describes the trade-off in between specificity and sensitivity. Sensitivity is defined as ratio between true positive tongue image and all those with disease. Specificity is defined as the ratio between true negative tongue image and all those tongue image samples with diseases.

SVM Classifier1	SVM Classifier2	SVM Classifier3	Our Classifier
71.79%	76.21%	80.42%	86.74%

When TTT=16.0, the GA and SVM based classifier is better than three SVM classifiers (with random chose state combinations & classifier parameters)



ROC curves show that our classifier has achieved a better true positive rate and a lower false positive rate compared with the best selected SVM classifier

Fig 4. Comparison of GA-SVM with other classifier.

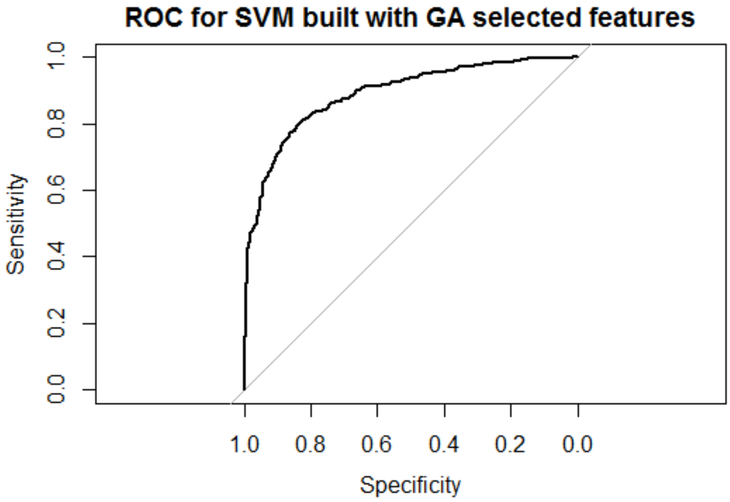


Fig 5. GA-SVM selected features

Fig. 5 describes the performance of PSO-SVM, SVM and GA-SVM in which it clearly describes that performance of GA-SVM is very significant. Error rate in the case of SVM is high compare to error rate of PSO-SVM. Error rate of PSO-SVM is high compare to GA-SVM. Fig 5. clearly describes the diagnosis accuracy rate of hepatitis tongue images with GA-SVM is very good. Table 1 indicates the result of SVM classifier1 with clusters of tongue image samples. Table indicates the result of SVM classifier2 with clusters of tongue image samples. Table 1 indicates the result of SVM classifier3 with clusters of tongue image samples. All tongue image samples of hepatitis patients are divided into five groups and each group is with different SVM classifiers. In all these classifiers GA-SVM exhibits higher performance in terms diseases diagnosis. In Table4 all the tongue related image are divided as H1-Sample, H2-Sample and H4-Sample to compare GA-SVM with CV-SVM and DC-SVM.

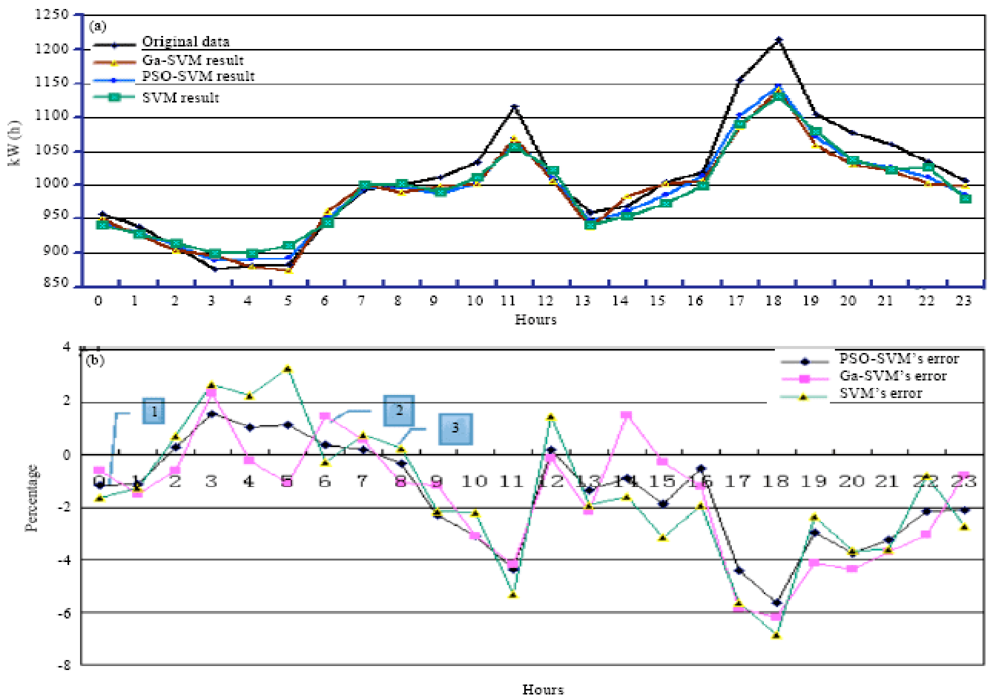


Fig 6. Comparison of PSO-SVM error, SVM error and GA-SVM error

Table 1: Accuracy results of SVM CLASSIFIER1

	SVM CLASSIFIER1					
TEST	**1**	**2**	**3**	**4**	**5**	**6**
1	**1228**	**340**	**0**	**0**	**0**	**84%**
2	**562**	**1000**	**0**	**0**	**1**	**72**
3	**0**	**0**	**1200**	**170**	**0**	**64%**
4	**0**	**0**	**25**	**1000**	**300**	**66%**
5	**11**	**16**	**0**	**69**	**100**	**66%**

Table 2:Accuracy results of SVM CLASSIFIER2

	SVM CLASSIFIER2					
TEST	**1**	**2**	**3**	**4**	**5**	**6**
1	**352**	**340**	**0**	**0**	**0**	**56%**
2	**402**	**104**	**0**	**0**	**1**	**52.1%**
3	**0**	**0**	**100**	**102**	**0**	**53.45%**
4	**0**	**0**	**35**	**1101**	**300**	**74.83%**
5	**210**	**250**	**0**	**109**	**1300**	**86.24%**

Table 3:Accuracy results of SVM CLASSIFIER3

	GA-SVM CLASSIFIER					
TEST	**1**	**2**	**3**	**4**	**5**	**6**
1	**1128**	**240**	**0**	**0**	**0**	**94%**
2	**432**	**932**	**0**	**0**	**1**	**82%**
3	**0**	**0**	**1300**	**80**	**0**	**94%**
4	**0**	**0**	**25**	**1000**	**300**	**96%**
5	**10**	**15**	**0**	**89**	**1300**	**86%**

Here 1 represents tongue image data related to group1

2 represents tongue image data related to group 2

3 represents tongue image data related to group3

1 represents tongue image data related to group4

2 represents tongue image data related to group 5

Table 4:Accuracy results GA-SVM with Hepatitis Samples

GA-SVM		**H1-Sample**	**H2-Sample**	**H3-Sample**
	SVM	**0.73**	**0.45**	**0.654**
	SVM+RIM	**0.89**	**0.92**	**0.89**
GA-SVM		**H1-Sample**	**H2-Sample**	**H3-Sample**
	SVM-1	**0.415**	**0.538**	**0.5317**
	SVM-2	**0.335**	**0.3223**	**0.473**
	RIM+SVM Linear	**0.92**	**0.92**	**0.97**
	RIM+SVM RBF	**0.93**	**0.94**	**0.98**
	RIM+SVM Poly	**0.92**	**0.93**	**0.93**
DC-SVM		**H1-Sample**	**H2-Sample**	**H3-Sample**
	SVM-1	**0.5285**	**0.538**	**0.531**
	RIM+SVM	**0.92**	**0.91**	**0.92**
CV-SVM		**H1-Sample**	**H2-Sample**	**H3-Sample**
	SVM-1	**0.527**	**0.5376**	**0.5337**
	RIM+SVM	**0.475**	**0.473**	**0.5872**
GA-SVM		**H1-Sample**	**H2-Sample**	**H3-Sample**
	SVM-1	**0.93**	**0.86**	**0.91**
	RIM+SVM	**0.96**	**0.98**	**0.99**

5. Conclusion

In this study, hepatitis tongue image is classified more accurately by the hybrid model derived from the combination of genetic algorithms and support vector machines. Results shows that classification accuracy of GA-SVM is significantly improved by dimensionality reduction and preprocessing methodology. The characteristics of GA-SVM helped to speed up the training and testing of tongue image sample obtained from hepatitis patients, this leads to avoid over fitting. The results obtained in this study is to diagnose the hepatitis depending on chromatic and texture information. Experimental results proven that gist and color information is more essential than color and texture. Our studies proved that GA-SVM is robust and showed superior performance compare to other methods like other conventional methods. We obtained significant results as we used the hybrid model GA-SVM. The results proved that classification accuracy of hepatitis tongue image is 93.74%.

References

[1] Abdel-Mottaleb,M., Elgammal, A., “Face Detection in complex environments from color images,’’ IEEE ICIP, pp. 622- 626, Oct. 2016.

[2] Aizerman, M., Braverman, E., Rozonoer, L.. "Theoretical foundations of the potential function method in pattern recognition learning". Automation and Remote Control25: 821–837, 2014.

[3] Antonio J. Colmenarez and Thomas S. Huang. Face detection with information-based maximum discrimination. In Computer Vision and Pattern Recognition, pp 782–787, 2015.

[4] Bellaire, G., Schluns, K., Oppermann, K., Schimke, W., "Matching Using Object Models Generated from Photometric Stereo Images", Proc. Machine Vision Applications in Industrial Inspection IV, SPIE Vol. 2665, San Jose, Califomia, pp. 70-81, 2015.

[5] M.Raja Sekar, “Region classification using SVMs”, Journal of Geomatics, pp 87-89, 2007.

[6] M.Raja Sekar, “Automatic Vehicle Identification” Journal of Advanced Research in Computer Engineering, pp 0974-4320, 2015.

[7] M.Raja Sekar “FER from Image sequence using SVMs “, Journal of Data Engineering and computer science, pp 80-89, 2016.

[8] Burges, C. J. C., A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery, 2(2):pp 121–167, 2015.

[9] Behrooz Kamgar-Parsi, Behzad Kamgar-Parsi, Jain, A., Dayhoff, J., “Aircraft Detection: A Case Study in Using Human Similarity Measure”, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 23, No. 12, 2001, pp. 1404-1414,2016.

[10] Beymer, D. J., Face recognition under varying pose. AI Memo 1461, Center for Biological and Computational Learning, M.I.T., Cambridge, MA, 2016.

[11] Brunelli, R., Poggio, T., Face recognition: Features versus templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(10):1042–1052, 2015.

[12] CHAPELLE, O., HAFFNE,R, P., AND VAPNIK, V. N. Support vector machines for histogram-based image classification. IEEE Transactions on Neural Networks 10, 5, 1055–1064, 2016.

[13] Cohen, I., Sebe, N., Garg, A., Lew, M.S., Huang, T.S., Facial Expression Recogni- tion From Video Sequences. In IEEE Intl Conf. on Multimedia and Expo (ICME), volume 2, pages 121 – 124, 2014.

[14] DORK ´O, G., AND SCHMID, Selection of scale-invariant parts for object class recognition Vol3, pp23-25, 2015.

[15] M.Raja Sekar, N. Sandhya, “Mammogram Images Detection Using Support Vector Machines” Vol2, pp 24-30, 217.

[16] Rowley, H. A., Baluja, S., Kanade, T., Neural networkbased face detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(1):23–38, 2015.

[17] Rui, Y., Huang, T. S., Ortega, M., Mehrotra, S., “Relevance Feedback: A Power Tool in Interactive Content-Based Image Retrieval”, IEEE Tran on Circuits and Systems for Video Technology, Vol 8, No. 5, Sept., 644-655, 2016.

[18] Sanderson, C. Paliwal, K., Fast Features for Face Authentication Under Illumination Direction Changes. Pattern Recognition Letters, vol. 24, no. 14, pages 2409–2419, 2015.

[19] Sung, K.-K., Learning and Example Selection for Object and Pattern Recognition. PhD thesis, MIT, AI Lab, Cambridge, MA, 2015.

[20] Turk, M. A., Pentland, A. P., Eigenfaces for recognition.Journal of Cognitive Neuroscience, 3(1):71–86, 2016.