iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IX **Month of publication:** September 2023

DOI: https://doi.org/10.22214/ijraset.2023.55735

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IX Sep 2023- Available at www.ijraset.com

Comparison of Classifiers with Application on DCE-MRI and Preterm Birth

G. Hema
Arunachala College of Engineering for Women

Abstract: ***Cervical cancer and Preterm birth in a pregnant woman is the most serious issues. The aim of our study was to compare different classifiers with SVM and its application. Various statistical methods are proposed to predict special factors. In this probability distribution is used to predict the desired outcome. However, most of the times enough information about the probability distribution is not available. In such situation, we need well predictors with minimum assumptions. Support vector machine (SVM) is a good statistical method for prediction. First order statistical features of the Brix parameters were used. First and second order features were used as explanatory variables for support vector machine (SVM) classification. Second order GLCM features could significantly predict treatment outcome with more accuracies. The performance of the classifiers is measured in terms of accuracy, specificity, and sensitivity. Finally, the result indicates the SVM classifier generates highest level accuracy, which is higher than other classifiers.***

Index Terms: Cervix, Preterm Birth, Magnetic Resonance Imaging (MRI), Features Extraction and Classification.

I. INTRODUCTION

Magnetic resonance imaging (MRI) is one of the important technique for planning and monitoring cancer treatment. Here, tissue distribution is imaged as a function of space and time [1]. Unlike a conventional MRI [2] which is linked to tumor oxygenation [3]. DCE-MRI is useful tool for prediction of treatment outcome and planning [4].

Brix parameters can be used to capture the properties of a tumor related treatment outcome. Texture analysis is used to capture important characteristics of an image [5], [6]. Texture features can be used for tumor classification [7]. Gray level co-occurrence is one approach for texture analysis, here GLCM is constructed, from which second order statistical features of texture can be calculated [5]. GLCM texture analysis is computationally efficient [10]. GLCM based features from DCE-MRI can also used for different types of cancer [11]. Preterm birth is one of the most problems to cause death and complications in pregnancy. A delivery before 37 weeks of pregnancy is called preterm birth. The preterm birth in developed countries is 5% to 7% and higher in developing countries. The bacterial infection can cause preterm labor. Preterm birth can be predicted by predicting factors affecting this [12]. The aim of the current study was to compare different classifiers by its accuracy and performance analysis and its applications on cervical cancer (identify the patient to be normal or critical stage) and preterm birth by using SVM classifier via Brix parameters estimation and GLCM. This study shows that the performance is more than other classifiers because of features vector based classification. This paper is organized as follows. Section II includes some notations and preliminaries of proposed method. In Section III, the proposed system and its motivations are introduced. We then develop algorithm for solving the proposed model. Section IV presents some experimental results. Finally, we conclude this paper with some discussions on future research in Section V.

II. MATERIALS AND METHODS

The data in this study consisted of dynamic contrast enhanced magnetic resonance images patients with cervical cancer and pregnancy. The size of tumor is different for different patient. The tumors are in stage I, stage II, stage III and stage IV. The three-dimensional parameter maps were transformed into two-dimensional images by appending tumor slices. Brix parameters can be used to predict the treatment outcome for a particular patient.

Preterm or premature birth is defined as birth before 37 weeks or less than 259 days in pregnancy. It can be classified based on gestational age. GA can be calculated from the first day of the last menstrual period (LMP). The four classification of PTB are,

- Extreme PTB (under 28 Weeks): birth before 28 weeks of pregnancy.
- Very PTB (28 to 32 Weeks): birth between 28 and 32 weeks of pregnancy.
- Moderate PTB (32 to 34 Weeks): birth between 32 and 34 weeks of pregnancy.
- Late PTB (34 to 37 Weeks): birth between 34 and 37 weeks of pregnancy.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IX Sep 2023- Available at www.ijraset.com

The relative signal increase (RSI), in the tumor can be calculated by comparing the signal at time *t* denoted as s(t) to the signal before injection of the contrast agent denoted as s(0). It is represented by,

$$RSI(t) = \frac{S(t)-S(0)}{S(0)} \quad \text{(II.1)}$$

The pharmacokinetic Brix model can be defined as,

$$RSI(t) = A\frac{k_{ep}}{k_{el}-k_{ep}}(e^{-k_{ep}t} - e^{-k_{el}t}) \quad \text{(II.2)}$$

Where, *A* is the amplitude, *kep* is the transfer rate of contrast agent from the tumor tissue to the blood stream and *kel* is the washout rate of contrast agent from the blood plasma.

In pregnancy cervical length, OS closed or open and water level are important factors for preterm birth. In most of the pregnancy cervical length <3cm are preterm birth and cervical length >4cm are post birth. A machine learning classifier is an algorithm that automatically classifies data. That can save time and make more efficient. It has both supervised and unsupervised classifiers. Unsupervised machine learning classifiers based on structures and anomalies of data and supervised classifiers based on training data, based on this classified. Machine Learning Classifiers are performed faster and high accuracy. Five types of classification Algorithms are,

- Decision Tree
- Naive Bayes Classifier
- K-Nearest Neighbors
- Support Vector Machines
- Artificial Neural Networks

A. Decision Tree

A decision tree is one of a supervised machine learning classification algorithm. This algorithm works like a structure of a tree. It classifies data like tree trunk to branches to leaves. It used to categorize a data to desired class.

B. Naive Bayes Classifier

Naive Bayes is a probabilistic Algorithm that works like Yes/No. That will calculate the possibility of given data. Moving from tag to tag, this calculates the probability of data within a particular category or not.

C. K-Nearest Neighbors

K-nearest neighbors (k-NN) is a pattern recognition algorithm. That will calculates how the data correspond to other data in n-dimensional space. k-NN categorized by calculating its nearest neighbor 'k'. If k = 1, it would be categorize into the class nearest 1.

D. Support Vector Machines

SVM algorithms classify a data by training dataset and creating a 3-dimensional classification model. SVM algorithms is excellent classifier because more accurate prediction and output as 3-dimensional. SVM algorithm is accurate machine learning models because it is multidimensional.

E. Artificial Neural Networks

Artificial neural networks are a collection of algorithms that work together to solve it. Artificial neural networks need lot of training data because the process is high. There are a different type of artificial neural networks depends on the problem we want to solve.

.

III. PROPOSED METHOD

The proposed system uses classification using SVM technique. Patient images are taken as input to proposed method. In training phase set of images to be trained and in testing particular new sick patient image to be tested. The input 3D image is transformed into 2D images by Brix parameters maps. SVM classification is a two step process,

- Testing
- Training

The testing and training stage consist of,

- Brix Parameters Estimation
- GLCM

Preterm birth problem is more in developing countries than developed countries. Using a support vector machine algorithm, we can classify whether a patient with cervical cancer to be normal or critical stage and classify factors that affecting premature birth.

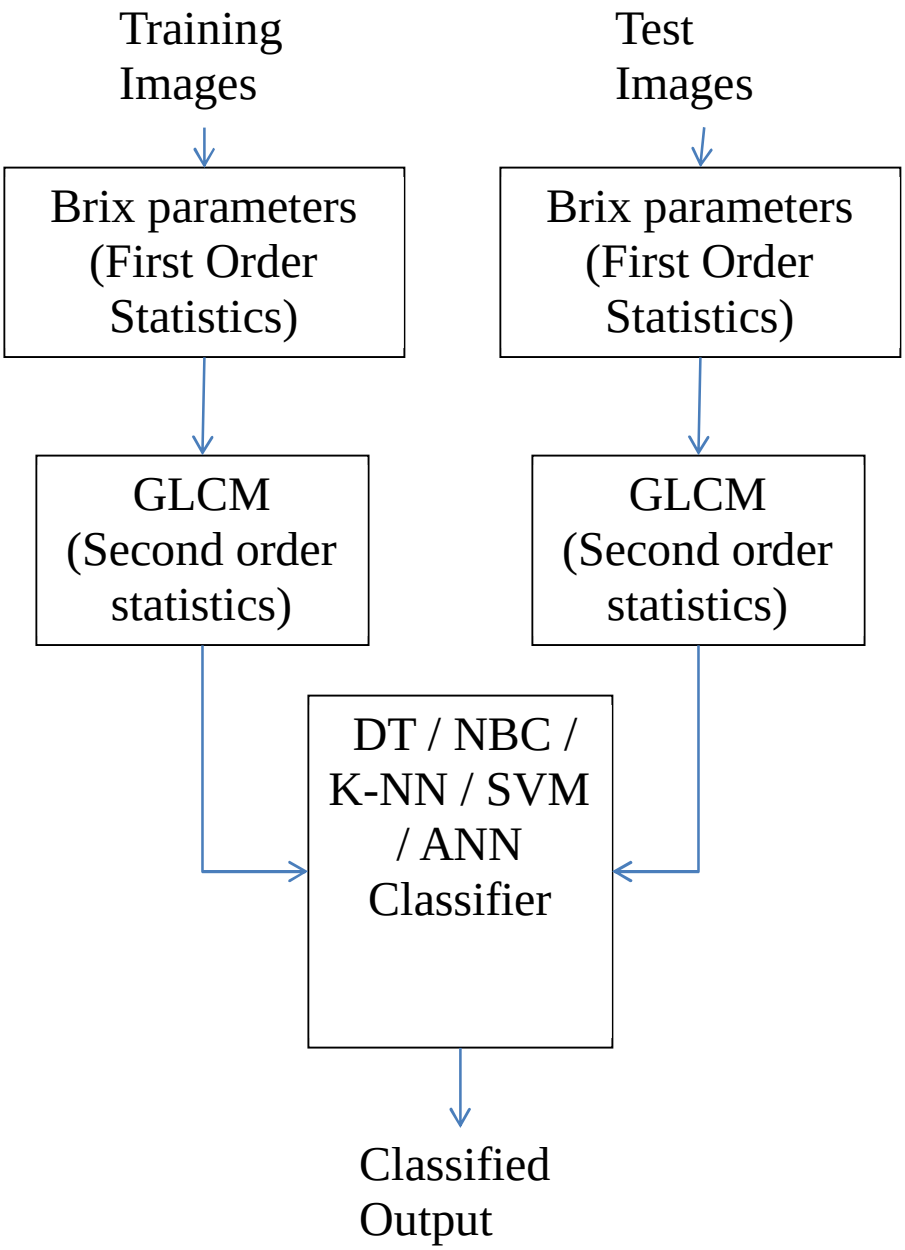


Fig. 1. Block diagram of proposed method

Fig. 1 shows that block diagram of proposed method for classification using SVM. First features are extracted from training images by Brix parameter estimation and GLCM then image trained and classes to class label. New patient's images are given to test image and features are extracted. Based on features SVM classified pregnancy patient as preterm birth or not and cervical cancer patient as normal or critical stage.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IX Sep 2023- Available at www.ijraset.com

A. First Order Statistics

Brix parameters can be used to captures the properties of outcome. The mean, median, mode, standard deviation, maximum and minimum value, skewness, and kurtosis of each Brix parameter (*A*, *kep*, and *kel*) were calculated. Brix parameters are based on variance. Where A is amplitude, *kep* is transfer rate and *kel* is washout rate. In addition, percentile values from 10% to 90% with 10% increments, and the percentile widths from 25% to 75% and from 10% to 90% were included to obtain a good representation. Totally, 21 first order statistical features were calculated. This can be used to predict particular patient condition.

B. Gray Level Co-occurrence

A Gray Level Co-occurrence Matrix can be constructed by square matrix where the number of rows and columns equals the number of gray levels in the original image. The GLCMs can be obtained by dividing each element with the sum of all elements in the matrix. Before constructing GLCM neighborhood to be choose. Here, 4th neighborhood method is chosen for easiest process. Then normalized GLCM is calculated. In normalized matrix, the sum of all elements to be one.

Second order statistical features can be calculated from the gray level co-occurrence matrix. The four features calculated from GLCM are Contrast, Correlation, Energy, and Homogeneity.

First measures the contrast in gray level from one pixel to its neighbor as defined as,

$$K = \sum_{(i,j)} \quad ((i-j)^2 p(i,j)) \qquad \text{(III.1)}$$

Second measures the correlation between intensities in neighboring pixels as defined as,

$$R = \sum_{(i,j)} \quad \left(\frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j}\right) \qquad \text{(III.2)}$$

Third measures the sum of square of all elements in GLCM as defined as Energy. It can be expressed as,

$$E = \sum_{(i,j)} \quad ((p(i,j))^2) \qquad \text{(III.3)}$$

Fourth measures the closeness of elements in GLCM to the diagonal as defined as Homogeneity. It can be expressed as,

$$H = \sum_{(i,j)} \quad \left(\frac{p(i,j)}{1+|i-j|}\right) \qquad \text{(III.4)}$$

The four features calculated from the matrix can be used to classification process.

C. Classification

Image classification is the process of classifying an image according to features of individual image. Two major steps in an image classification are,

- Features Extraction
- Classification

Different classifiers with merits and demerits as follows:

1) *Decision tree* is one of the commonly used supervised machine learning algorithms. That will divide the data into sub-data. This method split a data and continues until there is no possible of making any data splits. In this method both classification and regression are used. The outcome shows tree-like graph so it is called as a decision tree. In decision tree trunk, branches, and leaves are three parts. Based on this decision can be taken for result. DT method required less effort, not required normalization and scaling of data. But small change can cause large changes in outcome, higher time and some calculations increase complexity.
2) *Naive Bayes Classifier*, here variables are binary 0's/1's or Yes/No. In this dataset classified to Preterm birth as '0' and Term birth as '1'. Method classified a data by calculating probability of data. Equation for Naive Bayes is,

$$p(i/j) = p(j/i)\, p(i)/p(j) \qquad \text{(III.5)}$$

This classifier is simplest algorithm, faster and efficiency but rarely happening in real time.

3) *K-Nearest Neighbors* is supervised learning algorithm that used to classify a group of data. It uses either regression or classification method. Classification is mostly used with some assumption for patient outcome. Classification uses binary values but regression uses ordinary values so classification is best one for identification of patient. 'K' value defines how many neighbors are checked to be classification. Euclidean Distance can be calculated by,

$$d(x, y) = \sqrt{\sum_{j=1}^{m}(y_j - x_j)^2} \qquad \text{(III.6)}$$

If K=1, that are grouped to same class. This algorithm is simple and accuracy but if dataset increases it goes inefficiency.

4) *Support Vector Machine* is supervised learning algorithm used to classification and regression based on the concept of decision plane that defines the decision boundaries. SVM can be worked by mapping a data. It creates 2D as line or multidimensional as hyper plane which separate the data into classes. Supervised image classification can be divided in to two phases such as training data set and testing data set. In training phase, features are extracted based on energy, color and edges, etc. After features extraction images can be classified to corresponding image label. In testing phase also features extracted from testing image. After this process, the feature vector of testing image to be compared with feature vector of image label. Once the match is found with any of the class, an image is stored and labeled with same class. In cervical cancer patient, Images can be classified to level 1, level 2, level 3 and level 4. Level 1 is labeled as normal stage and level 4 are classified as critical stage. In pregnancy patient, images can be classified to term or preterm birth. In SVM clear classification between outcome dataset, comparison to other classifiers SVM has better computational complexity, more effective in multi dimensional, memory efficient, better accuracy.

5) *Artificial Neural Networks* predicting preterm birth at early stage by using advanced machine learning techniques and electrohysterography signal processing. EHG can be used to predict before labour. Six different ANN used to detect term and preterm births. First filtering the data then from the filtered data features are extracted that including root mean square, median frequency, peak frequency and sample entropy. Then we want to compare the oversampling results with results from original data for patient outcome. Artificial Neural Network is powerful machine learning technique but have some disadvantages. It is very flexible but required lot of computational power, complexity, hard to explain and training required lot of data.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed classification technique is implemented in the working platform of MATLAB with machine configuration. Magnetic Resonance Imaging (MRI) is a medical imaging technique (non-invasive method) that can be used in radiology to generate image of organs in the body. Cervical cancer is a type of cancer that occurs in cervix (lower part of uterus that connects to vagina). Preterm birth is birth that occurs before 37 weeks of pregnancy. In our proposed method, patient with cervical cancer images / pregnancy patient images are given to an image classification process by using the classifiers techniques.

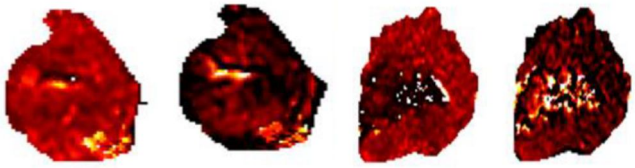


Fig. 2. Cervical cancer images with different stages

Fig. 2 shows that, different patient with different stage of cervical cancer. These images are taken as input image of proposed system. From an image features are extracted and based on features images are automatically classified by SVM.

Fig. 3. Term birth images

Fig. 3 shows that, full-term birth image of 20 weeks 5days pregnancy patient with liquor volume is normal, single vertical pocket measures 4.7 cm, FHR as 151 bpm and no placenta previa / retroplacental collection. In 12 weeks of pregnancy that patient measured adequate liquid, cervical length ~30mm and OS is closed.

TABLE II
Pregnancy Patient Report

Patient	Fetal parameters	Measurements/ Weeks	Amniotic fluid	EFW	Cervical length	Internal OS
1	BPD	4.3 cm/19w 2d	Adequate	289 gms ± 42gms	4.2 cm	Closed
	HC	15.9 cm/18w 6d				
	AC	13.7 cm/19w 1d				
	FL	2.8 cm/18w 6d				
2	BPD	4.88 cm/20w 5d	Adequate	370 gms ± 54 gms	3.8 cm	Closed
	HC	18.19 cm/20w 4d				
	AC	15.84 cm/21w 0d				
	FL	3.31 cm/20w 2d				
3	BPD	63 mm/25w 5d	Adequate	775 gms	3.3 cm	Closed
	HC	232 mm/25w 2d				
	AC	199 mm/24w 4d				
	FL	47 mm/25w 5d				

Table II shows that different term pregnancy patient in different weeks and their report. Here the important parameters are fetal parameters, liquid volume, cervical length and internal OS. Based on above factors, pregnancy patient are classified as whether term or preterm births.

TABLE III
Performance Comparison

Noise level to be added 0.0006					
Parameter	Decision Tree	Naive Bayes Classifier	K Nearest Neighbor	Support Vector Machine	Artificial Neural Networks
Total Checked	16	16	16	16	16
Correctly Detected	9	10	11	15	13
Incorrectly Detected	7	6	5	1	3
Accuracy	50	67	75	90	86

Table III shows that performance comparison of different classifiers. There may be noise can be added in input data. If noise is added to an input data, correctly prediction of data and an accuracy of support vector machine as higher than other classification methods.

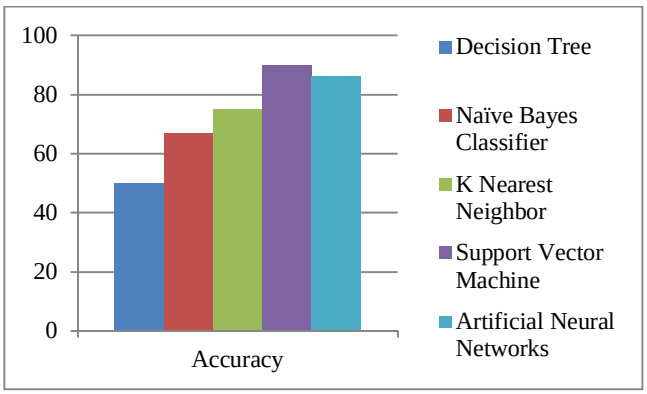


Fig. 4. Performance Comparison Chart

Fig. 4 shows that performance comparison of classifiers. In above figure each color denotes corresponding classifiers and an accuracy of SVM is higher than other classifiers. When noise level 0.0006 added to an input data then an accuracy of Decision Tree is 50%, Naive Bayes Classifier is 67%, K Nearest Neighbor is 75%, Support Vector Machine is 90% and Artificial Neural Networks is 86%.

In performance analysis we have measured an accuracy using the formula,

$$Accuracy = \frac{total\ number\ of\ images\ correctly\ classified}{total\ number\ of\ images\ tested} x100$$

Accuracy is based on four parameters such as True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN). Here, TP - Sick people correctly identified as sick, FP - Healthy people incorrectly identified as sick, TN - Healthy people correctly identified as healthy and FN - Sick people incorrectly identified as healthy. Accuracy can be defined as total number of image correctly classified divided by total number of image tested. When the test image is corrupted by noise, accuracy is more for support vector machine classification. The high accuracy value is 1(100%) and low value is 0 (0%).

TABLE IV
Performance Analysis

Noise intensity	DT	NBC	KNN	ANN	SVM
0	100	100	100	100	100
0.1	56	63	77	86	94
0.2	32	48	62	70	84
0.3	18	37	43	58	70
0.4	12	23	38	44	58
0.5	4	7	11	12	38

Table IV shows the performance analysis of the proposed system for image classification technique. The performance analysis shows when noise is added to an image, the performance analysis of different classifiers. When the noise intensity increases, the performance analysis of SVM is best than DT, NBC, KNN, ANN.

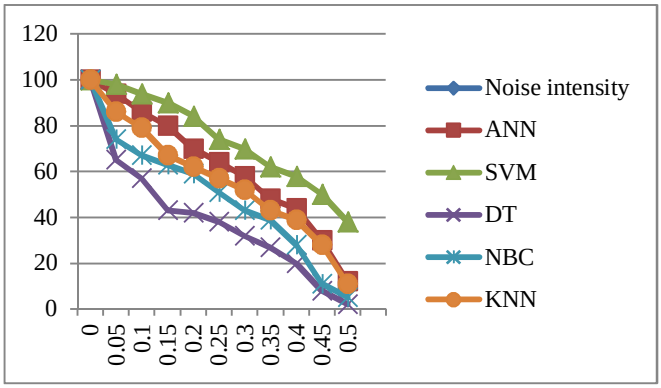


Fig. 5. Performance Analysis chart

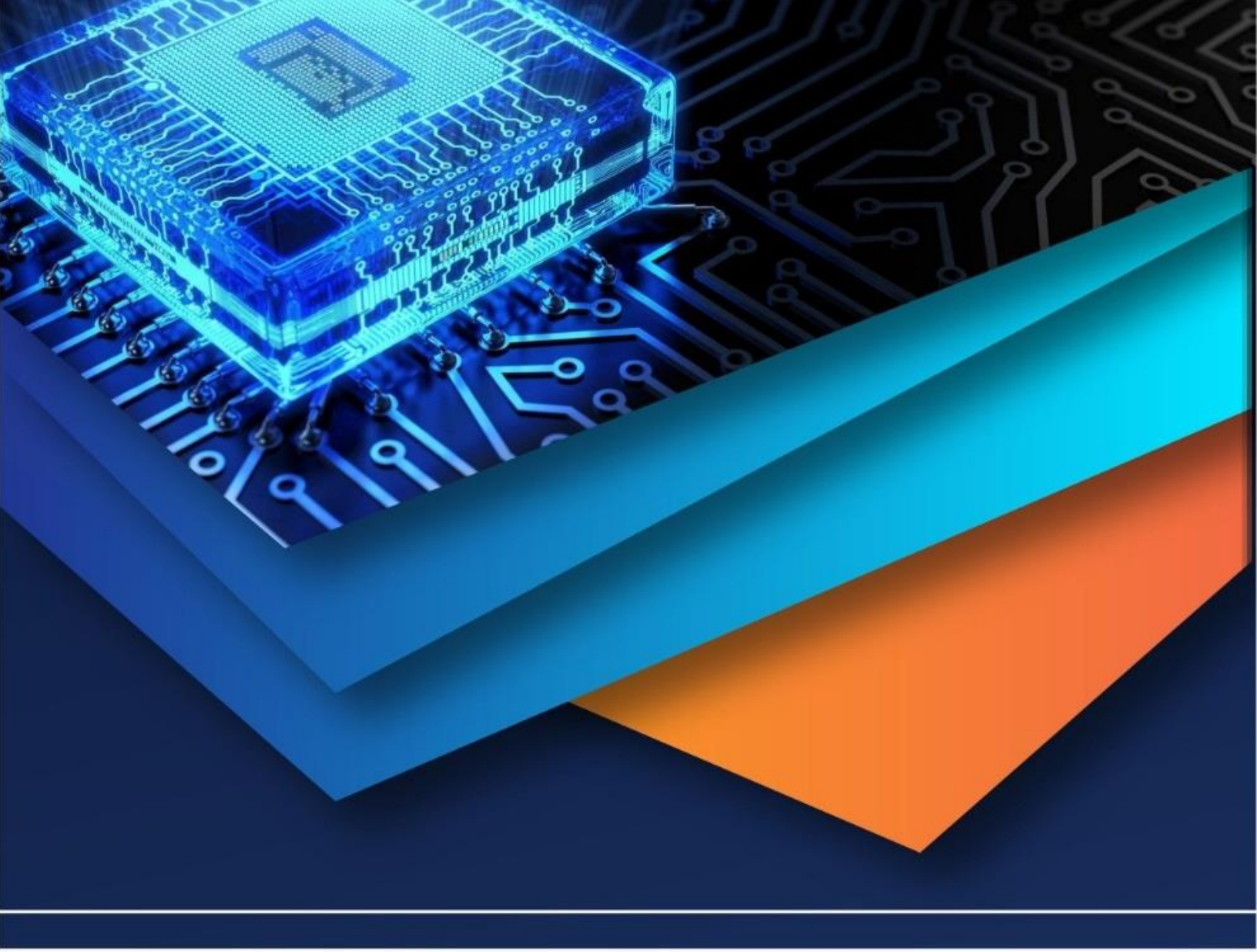
Fig. 5 shows that performance analysis of five classifiers. All classifiers are well classified when there is no noise intensity but for all classifier performance analysis is low when noise increases. Comparing the performance analysis of five classifiers SVM is better than DT, NBC, KNN and ANN when level of noise increases. From this project work, it can be easily observe that support vector machine classifier yield more satisfying results and can applied to different type of images. In future this method can be applied to different medical imaging process and also uses different classification methods to classifying images. The most important application of SVM is to recognize hand-written character.

V. CONCLUSION

In this paper, we demonstrated that image classification based on features extraction. Based on this context, we proposed a GLCM and DT/NBC/KNN/SVM/ANN for classifying cervical cancer images or premature birth. Features are extracted from images using Brix Parameters Calculation and GLCM methods. Based on features SVM automatically classify cancer patient as normal or critical stage and pregnancy patient as term or preterm birth. The Performance analysis of the GLCM shows that more accuracy. The results shows that the complexity of the SVM classification process reduces highly because the system classifies an image to class label automatically based on vector features. The accuracy of five classification method reduces as the noise intensity increases. But the accuracy of SVM is more than other classification method for the every noise intensity. Naive Bayes Classifier is better as compare to decision tree classification model in the preterm birth data. ANN method classified like SVM but increases computational power. We concluded that the SVM classifier has a better performance to predict the factors affecting preterm delivery than DT/NBC/KNN/ANN model.

REFERENCES

[1] N. Harry, F. J. Gilhert, and D. E. Parkin, “Predicting the Response of Advanced Cervical and Ovarian Tumors to Therapy,” Obstetrical & Gynecological Survey, vol. 64, no. 8, pp. 548-560, Aug, 2009.
[2] L. Alic, M. van Vliet, C. F. van Dijke, A. M. M. Eggermont, J. F. Veenland, and W. J. Niessen, “Heterogeneity in DCE-MRI parametric maps: a biomarker for treatment response?,” Physics in Medicine and Biology, vol. 56, no. 6, pp. 1601-1616, Mar, 2011.
[3] H. Lyng, A. O. Vorren, K. Sundfor, I. Taksdal, H. H. Lien, A. Kaalhus, and E. K. Rofstad, “Assessment of tumor oxygenation in human cervical carcinoma by use of dynamic Gd-DTPA-enhanced MR imaging,” Journal of Magnetic Resonance Imaging, vol. 14, no. 6, pp. 750-756, Dec, 2001.
[4] M. A. Zahra, K. G. Hollingsworth, E. Sala, D. J. Lomas, and L. T. Tan, “Dynamic contrast-enhanced MRI as a predictor of tumour response to radiotherapy,” Lancet Oncology, vol. 8, no. 1, pp. 63-74, Jan, 2007.
[5] R. C. Gonzalez, and R. E. Woods, Digital image processing, Upper Saddle River, N.J.: Pearson/Prentice Hall, 2008.
[6] M. Mirmehdi, X. Xie, and J. Suri, Handbook of texture analysis, London: Imperial College Press, 2008.
[7] X. Y. Yang, and M. V. Knopp. (2011, Feb). "Quantifying Tumor Vascular Heterogeneity with Dynamic Contrast-Enhanced Magnetic Resonance Imaging: A Review," Journal of Biomedicine and Biotechnology [Online], 2011, 12 pages. Avaliable: doi:10.1155/2011/732848
[8] G. Castellano, L. Bonilha, L. M. Li, and F. Cendes, “Texture analysis of medical images,” ClinicalRadiology, vol. 59, no. 12, pp. 1061-1069, Dec, 2004.
[9] L. Alic, “Quantification of Tumour Heterogeneity in MRI,” Ph.D. dissertation, Dept. of Medical Informatics, Erasmus University, Rotterdam, the Netherlands, 2013.
[10] Batoul Ahadi1, Hamid Alavi Majd, Soheila Khodakarim, Forough Rahimi, Nourossadat Kariman, Mahieh Khalili1, Nastaran Safavi,“Using support vector machines in predicting and classifying factors affecting preterm delivery,”.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY