

FEATURE SELECTION FOR GABOR FILTER BASED ON LEVEL MEASUREMENT USING NON - INTERACTING TANKS LEVEL IMAGES

B Kalaiselvi¹, Dr. Karthik B² Dr.A.Kumaravel³

Research Scholar, Dept of ECE, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India
Associate Professor, Dept of ECE, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu,
India

Dean, School of Computing, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India
kalaigopal1973@gmail.com¹, karthikguru33@gmail.com², drkumaravel@gmail.com³

Abstract

Level measurement models using image-based classifiers (pixel-based datasets) are used for estimation purposes. Preprocessing is thought-provoking in proceeding out the image filter technique and classifying the level. The level scenario of a two non-interacting tank system plays a vital role in predicting the level. Level monitoring is done using the supervised learning method using instance-based filters (Gabor Filter) and selected base classifiers for level measurements. The main scope of this case study is to improve the level measurements from the two non-interacting tank scenarios using Artificial Intelligent algorithms. The suggested article includes the finest feature selection process to increase the accuracy performance attained by the designated classifiers like IBK Instance base classifier for different neighborhood values and Tree category algorithm like Random Forest. The performance accuracy in level prediction obtained is 81.356%, the weighted Average of Receiver operator characteristics of (ROC) 0.931 are obtained by Random Forest Tree Category Classifier.

Key words: Level Monitoring, Gabor filter, machine learning, KNN, Random Forest, ROC

I Introduction

The Level prediction is carried out in 3 unique classes and provides the corresponding output with the given input image. In this Level Scenario is the enter given to the modeled machine in which its degree prediction the output system gaining knowledge of algorithms. However, in the enterprise, there can be a wonderful call for measuring the extent in order that the controlling venture is carried out in a completely unique way making use of this novel approach.

This article level image is the input given to the modeled system where in the level prediction is the output using machine learning algorithms. Nevertheless in industries there is a great demand in measuring the level so that the controlling task can be done in a faster rate by applying this novel approach. The applicable framework of studies in this article goes a long in the following steps first the novel Metrics has been used to analyse the measurement of level and its accuracy is checked. Level measurement performance is evaluated over a wide range of scenarios. The best possible acceptable performance within the dataset threshold is set for an optimal model. Data acquired is pre-process using image filter called Gabor in Weka tool. Model training is performed training is taking place over the data set using some of the selected classifiers. The result is being tabulated for its performance characteristics and the area under the curve that measures the weighted average of ROC is evaluated. The Characteristics is checked for better performance and attainment of an optimum model building is done. The model has the capability to give three different levels of the fluid by just giving an image as input.

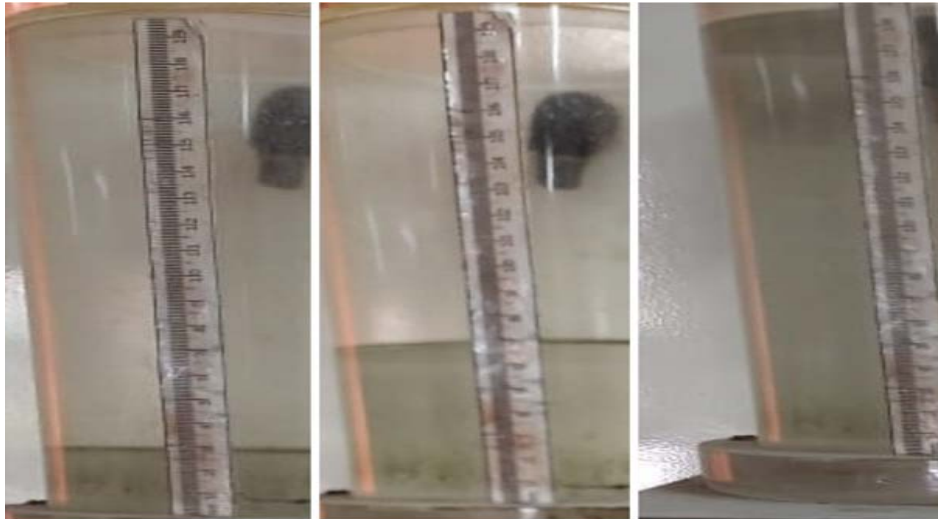


Figure 1 the staff gauge image of the non-interacting tank

The proposed dataset as shown in figure 1 establishes the accuracy of levels through experiments-primarily based on total characteristic choice with a new release on rating thresholds and Gabor with multiple iterations on elegance levels. However, thinking about the extent images for studying and predicting the extent through constructing a supervised version. The device used to construct the version is Weka. This device is to be had with Java implementation and it's miles used for image preprocessing.

The proposed system of research on this paper is going alongside the subsequent steps: [1] Firstly numerous robustness metrics were used to research the level prediction and its accuracy is checked. Expected overall performance is evaluated over a huge variety of eventualities, a couple of eventualities, and the nice viable proper overall performance in the set threshold is sought for procuring the very last most desirable model. Secondly [2] The image data statistics set received is preprocessed with the Weka tool and the set attribute is applied to the image dataset. [3] Training set acquired after attribute setting by removing the Gabor feature in ascending order of 3 is applied and training progresses. [4] The classification of level measurements is done and the resultant performances are evaluated. [5] The attribute setting using the ranker module is again applied for a group of 3 until getting a good threshold performance evaluation.[6] The optimum model is attained after training using the selected machine learning algorithms.

The goal of this project is to anticipate the three different fluid levels using several well-known machine learning methods. It differs from prior works in that it proposes yet another innovative framework. To obtain equivalent results with deep learning models utilising the features retrieved from the two tank non-interacting image scenario, the suggested method using shallow learning takes less computer resources and time.

For feature extraction and attribute selection, the proposed work employs a specialised filter on a specific image collection, and selected characteristics are provided for both J48 and KNN. Multi-classifiers for many degrees of classification like below the required set point of level as low class, above the setpoint as high class and then $\pm 5\%$ around the setpoint is classified as medium level.

This article is structured by four different stages they are the first section comprises data collection and converting the image to its corresponding numeric 2nd part covers apparent descriptions towards a required methodology for completion. The 3rd part comprises preparing the data set for further experimental analysis like selecting attributes and then classifying, and the concluding part of this article is experimental observations, report generation, and future inferences.

II Literature Survey

Predictive algorithms in Machine learning like Naïve Bayes, Instance base classifiers KNN, Decision Tree, and Random Forest are used for the early detection of Diabetes. Ensemble hybrid model increased the accuracy and performance time. Mohammed M Mazid, A. B. et al (2013) [18]. C4.5 Algorithm for Rule-Based Classification has been improved. Artificial Intelligence, Knowledge Engineering, and Data Bases: Recent Advances.C4.5 is a popular algorithm a rule-based classifier with very high accuracy and performance time. Nagaparameshwara chary, S. D. et al (2017) [19]. A Survey of Decision Tree Algorithms in Data Mining and Their Comparative Analysis. Innovative Engineering and Information Technology Applications International Conference. A innovative comparative study on ML algorithms like decision tree algorithms like CART, ID3, C4.5, CHAID and their performance and accuracy evaluation is carried out. Jian Huangl et al (2020) [22] proposed a novel automatic

analog instrument reading system using computer vision and an inspection robot has been presented. Paaranan Sivasothy, et al (2018) [23] proposed Machine learning-based filling level estimation for bulk solid silos and classify different filling levels. Changhyun Choi et al (2019) [24], Development of Water Level Prediction Models Using Machine Learning in Wetlands using various machine learning models such as artificial neural network (ANN), rules based category like decision tree (DT), trees based category like random forest (RF), and functions based category support vector machine (SVM). F. N. M. Ariff, et al (2020)[26] "A Character Segmentation for Automatic Vehicle License Plate Recognition Based on Fast K-Means Clustering" proposed to reveal and become aware of the automobile with the aid of using analyzing the automobile registration code numbers and discover the differences of the plate character. The received photo has segmented the use of Fast K means clustering and Fuzzy 'K' means algorithms with an accuracy of an extra the 80%. Facts set of one hundred photos are used for schooling and trying out purposes. The technique concerned is segmentation with clustering processes implemented. Fuzzy-like set of rules and in comparison for FCM display a barely better common percent of accuracy. V. V. Mainkar, et al (2020)[27] Involved preprocessing, segmentation, function extraction, and Post-processing on OCR to understand the person. An easy manner that may be applied is through simply taking pictures the photo of the handwritten information into an electric amount mild modulated via which the person is recognized. The proposed device makes use of an android mobile smartphone to seize the photo and optical person recognition. The hard challenge is the performance of the mild sign to understand one-of-a-kind handwriting styles. The degree of accuracy in overall performance is ready 90% for the handwritten files examined information set. Y. Peng et al (2020)[28] , recognition on analyzing the water meter man or woman automatically, which is additionally a unique approach for automated identity approach the usage of Deep Neural Network. This additionally proposes automated man or woman analyzing from the received underwater photograph and led to excessive accuracy. Self-made water meter photograph dataset is optimized the usage of region-primarily based totally absolutely convolution networks for goal detections with numerous iterations to as it should be study the water meter characters. The vital kind algorithms are implemented for identity. The consequences recommend that the popularity fee is excessive that could meet the needs of wide-ranging companies.

III METHODOLOGY

The existing resources help to make the suggested framework more understandable. This section also discusses the motivation behind classifiers as well as their representation models.

3.1 Gabor filter

Gabor filter is used for feature extraction. This filter is specially designed for statistical information of character structures. Gabor filter gives outputs to achieve better performance on low pixel quality images. An adaptive sigmoid function is applied to the outputs of Gabor filters to achieve better performance on low-quality images. To enhance the discriminability of the extracted features, the positive and the negative real parts of the outputs from the Gabor filters are used separately to construct histogram features. Experiments show that the proposed method has excellent performance on low-quality machine-printed character recognition and cursive handwritten character recognition^[1].

The transfer function $G(k)$ of a Gabor filter (Fourier transform of the impulse response) is given by:

$$G_{mn}(k) = e^{-1/2(k-k_{0mn})(A_{mn}^{-1})^T(k-k_{0mn})} \quad (1)$$

where $k = [k_1 \ k_2]^T$ is the spatial frequency. In order to establish a multi-resolution strategy, the image can be filtered with a set of N Gabor filters with different bandwidths and modulation frequencies. If the modulation frequencies are given by

$$k_{0n} = \frac{\pi}{2^{n+1}}, \quad n \in [0, \dots, N-1] \quad (2)$$

and the relative bandwidth is chosen to be constant for all filters the image is decomposed into octaves.

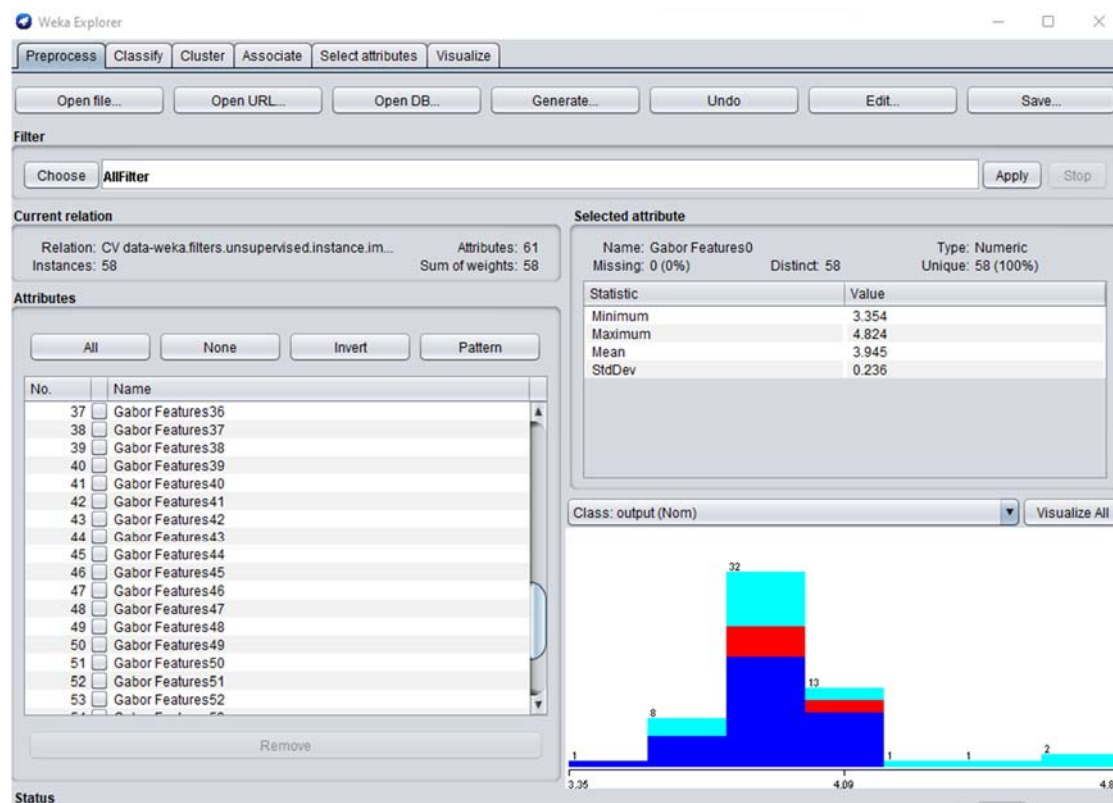


Figure 2 Weka Gabor Filter based data feature extraction

3.1 Gabor Feature filter in the combination of selected Attributes

The collected data is first converted into numeric utilizing the image filter like Gabor, as shown in figure 2 is the procedure in the Weka tool (Using this image filter, the image data set is converted from pixels to statistical features). This ARFF (Attribute Relation File Format) file is used for further classifying process, some of the selected classifiers are used to predict the level and the accuracy performance is evaluated and compared to converge into getting an optimum model.

3.1.1 Parameter Training

Parameter training is for refining the predictions done by the selected classifiers using machine learning algorithms. Here IBk classifier with changing number of nearest neighbors (k) values forecasts the accuracy level and produces the Receiver Operator Characteristics values

3.1.2 Attribute selection in Weka

In Weka, there are 3 options for performing attribute selection

- The built-in approach, using the attribute selection tab
- Openly using a meta category type classifiers
- Starting with filter approaches

This is a novel method of evaluation feature subset. This means that it tries every possible combination of the variables and returns the best-performing subset. Perform well in improve accuracy and less misleading data, reduce training time and reduce overfitting by removing the irrelevant data and redundant data. To create an optimal Machine learning model, it is critical to remove unnecessary attributes from the database. Some automatic feature selection tools can be found in Weka. We choose its attribute selection methods like InfoGain Attribute Evaluator /Ranker Method available in this tool to analyse attribute ranking order since we regard Waikato Environment for Knowledge Analysis (WEKA) to be an open source tool. To improve the performance of the accuracy percentage, we constantly delete attributes in lower order. To keep the number of iterations to a minimum, we repeat the list of sorted lists of attributes in relation to the selected threshold. The attribute evaluator applied for the selected database is examined in such a way it correctly classifies the output class

3.2 Designated Base Machine Learning Algorithms

Selected two classifiers are that play a vital role in building a model to predict the level parameter using the supervised learning algorithms to fit our modeling procedure and their knowledge and modest implementation foremost to easier clarifications.

3.2.1 IBk Instance base classifiers

One of the traditional algorithms in the machine learning concept that is the lazy based categorized machine learning algorithm like IBk the nearest neighborhood where the data are distributed in a hyper dimensionally and the manipulation is done depending on the “k” values nearest neighbors by just imitating the tags of the weight without any specific effects on prediction. Although it is very complex calculations, it is very unpretentious for execution.

Lazy category classifier, which Imitates ‘k’ nearest neighbor, An instance-based classifier IBk with k neighbors is implemented without taking care of any specific model, though it is computationally expensive as it is based on Euclidean distance between any two instances in the data set. Hence the laziness of this classifier for fitting the model turns out to be an advantage for escaping from any representational specific classifier. KNN classifier is a nonparametric algorithm which means it does not make any assumptions on underlying data. Because it just learns from the training set, it stores the data set, at the time of classification it performs an action on the data set. During the training phase, it stores the data set and when it gets new data then it classifies that data into a category that is much similar to the new data. This algorithm is explained on the basis of the below-mentioned steps:

- Step 1: Select the number of K, the neighbors
- Step 2: Euclidean distance of K number of neighbor
- Step 3: Take just neighbors as per the calculated Euclidean distance
- Step 4: For undefined among these k neighbors count the number of data points in each category
- Step 5: Assume the new data points to the category for which the number of neighbors maximum
- Step 6: The model is built using this algorithm.

Figure 3 specifies KNN Classifier in the parameter tuning option panel provides the option for changing the neighborhood values, the batch size can be varied, and a search algorithm can be varied. It has the capability to handle Binary class, Date class, Missing class values, Nominal class, Numeric class.

It uses nearest neighborhood search algorithms like ball tree, filter neighborhood search, KD tree, and by default linear NN Search applied. The output number in the module is two by default. The maximum number of instances allowed in the training pool is zero. The batch size is 100 by default.

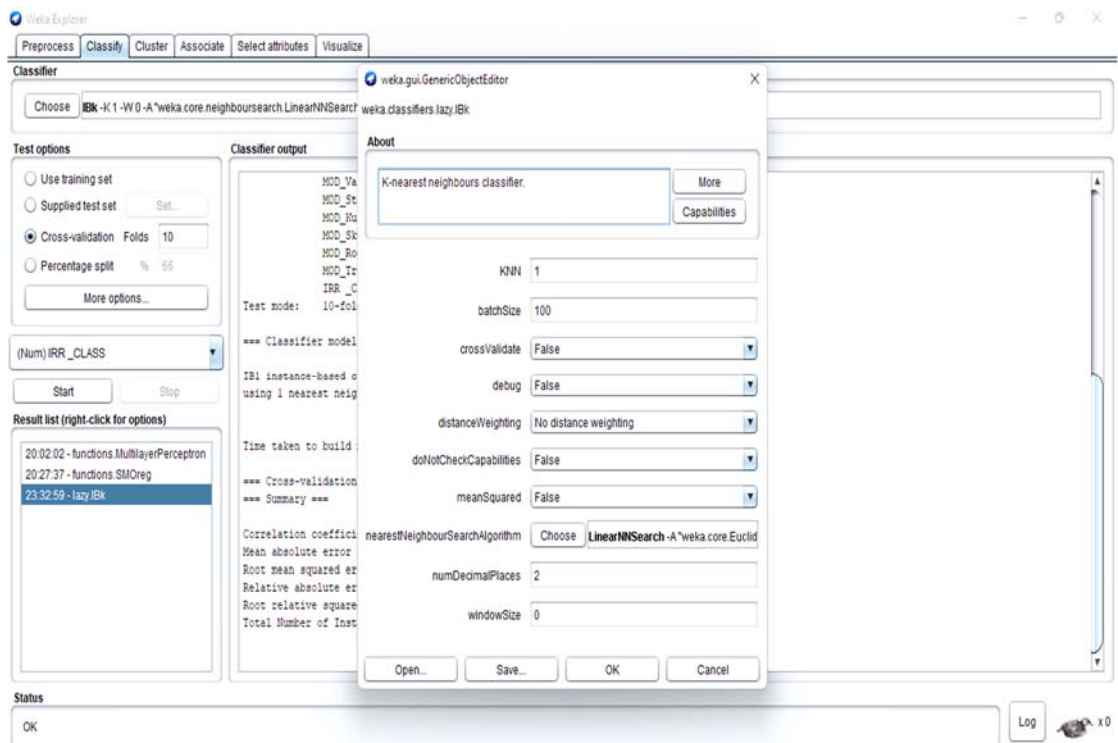


Figure 3 KNN Classifier in parameter tuning option panel

3.2.2 Random Forest

This is also a renowned type of tree category machine learning algorithm. The ensemble machine learning algorithm in weka. create multiple decision trees and merge them to get a more accurate and stable prediction. These decision tree-based predictors are best known for their computational power and scalability. However, in the case of highly unbalanced training data, as is often seen in data from medical studies with large control groups, the training algorithm or sampling method should be changed to improve the prediction quality for the classes. minority. In this work, a balanced random forest approach is proposed for WEKA.

It is an ensemble type of classifier. Mainly the bagging with base tree classifier. The main advantage is being an ensemble faster and more accurate than the base tree classifiers. One of the most powerful algorithms is Random Forest whose accuracy is high and training time is low. Moreover multiclass object detection can be accommodated.

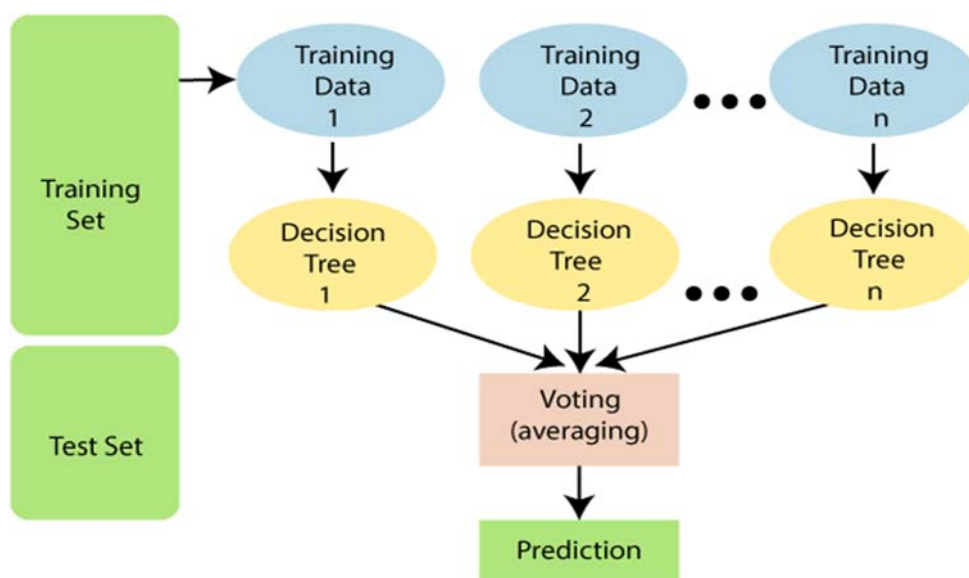


Figure 4 Splitting of the dataset in Random Forest classifier

Below are some points that explain why you should use the Random Forest algorithm as shown in figure 4

- Training time is reduced compared to other algorithms.
- Predict output with high accuracy, even for large datasets that run efficiently.
- Maintain accuracy even if a large amount of data is missing.

The work process can be described in the following steps and diagrams.

Step 1: Select a random K data point from the training set.

Step 2: Build a decision tree associated with the selected data points (subset).

Step 3: Select the number N of decision trees to build.

Step 4: Repeat steps 1 and 2.

3.3 Evaluating the Performance Criteria

The performance measures are derived from the prediction patterns provided by the classifiers in use, which are represented in the confusion matrices' available inputs. High Level (above the setpoint), Low Level (below the setpoint), and Medium Level (in between the setpoints) are the three levels (around 5 percent plus or minus the setpoint). The Receiver Operating Characteristic (ROC) area, a performance metric, is calculated at the end.

IV DATA DEMONSTRATION AND REPORT GENERATION

The main materials come from the existing laboratories, which contain a system of 2 tanks that do not interact, and the photos of the multi-level images are taken with a conventional smartphone and the level is measured with the lever indicator attached to the tanks.

Table 1: Experimental dataset for level prediction

S.No	Level type	Instances count	Class description of Level
1	High level	148	Above the set point
2	Medium level	112	Around +-5% with set point
3	Low level	94	Below the set point

This is the first processing step. The final level to achieve the optimal model is the main objective to measure the level.

Table 1 demonstrates how the three-class values are distributed with their Instances count. The classes High Level (above the setpoint), Low level (below the setpoint), and Medium (around 5% plus or minus the setpoint).

Flow chart for Level Monitoring describes the flow chart for the proposed system is shown and the experimental setup to predict the level is done by collecting the data from non-interacting tank scenarios from a normal camera image in JPEG format and the level is predicted using the conventional method. The collected data is pre-processed in WEKA software for the implementation of the machine learning algorithm with the Gabor filter (feature extraction). Data is classified using selected classifiers such as IBk and Random Forest. The next step in the progression is to check for better precision when the precision level is lower. We go to the preprocessing step called Info gain attribute evaluator and ranker for increasing the accuracy performance percentage and the classification steps are performed with the same base classifiers and the level of precision is checked. If the threshold is reached, the model is the optimal model.

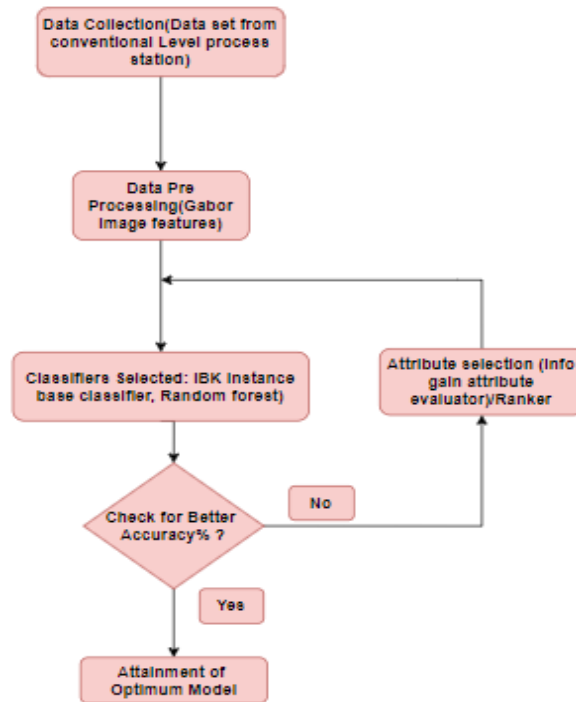


Figure 5 Flow diagram for level Monitoring

Figure 5 shows the experimental set up for predicting the Control valve stem position in the following steps:

- Data collection as raw images that has to be converted into an accessible format like Atriff or comma-delimited value files CSV
- Data preprocessing using Gabor filter and the image to numeric values is processed for further classification and building the model using selected classifiers
- Classification and building the model for training and obtaining the examination performance results
- Observations of the results for better accuracy percentage and area under the curve called weighted receiver operator characteristics(ROC) is better or not, for repeating the above process by varying the 'K' neighborhood values from 1 to 3 values to improve the overall performance of the built model.
- Attainment of the optimum model to predict the control valve stem position using machine-learning algorithms.

V EXPERIMENT AND RESULTS

In this section, we'll go over the recommended experimental setup. Figure 5 depicts the suggested framework methodology. Table 2 shows the results of applying the Java version of the Random Forest decision tree classifier to selected characteristics. With selected properties and training/parameter fitting to neighborhood size, the IBk (KNN) classifier was developed. "k" is shown in Table 3. The experimental statistical analysis of aspect level versus precision and attribute level prediction to the area under the curve called weighted average receiver operator characteristics are shown in Figures 2 and 3, correspondingly.

5.1 Proposed Methodology

The data enhancing techniques like data cropping and assigning the ground truth by a skilled person who already did the traditional level predictions. The categorized images are not evenly disseminated among the nominated classes and images with labels as LOW, MEDIUM, and HIGH. Because this implies a highly skewed distribution of images, training this image data set will skew the results and impair classification accuracy. To avoid this bias, data replication techniques were used to boost the count of photos in the remaining classes. [25] This pre-processed data is subjected to a feature extraction method with the help of image filters (Gabor) and then the features have been extracted, the characteristics. Data training activities such as Data preprocessing, Parameter Training, and Attribute Selection were performed on both Random and K-Nearest Neighbor classifiers in the process of obtaining the best model. For various setups, the performance of both classifiers is compared, and a better model based on accuracy is obtained. The above framework and its components are graphically explained in Figure 1

5.2 Accuracy Performance Examinations

The Random and IBk classifiers were used to test performance outcomes on feature selection and parameter training/tuning. Reiteration is a typical strategy for enhancing the performance of machine learning models and reducing training time. The information gain Attribute Eval/Ranker in WEKA is used to select the attributes. The IBk classifier modifies the nearby neighborhood to tune the parameters (k).

A quantity of various characteristic choice strategies as benefit the information from the associated works choice of method of the statistics is taken into consideration. Additionally the movement of various forms of a subset of statistics set mode the use of processing and classifying the use of in which is set of rules silly and after modern.

The reason is to pick out an excellent tool mastering method and to try and construct a version for every of the statistics. The approach of assessment of consequences is performed earlier than making use of that it's far an top of the line order the statistics set to the satisfactory overall performance furthermore this greater used on maximum issues for mastering algorithms.

The characteristic statistics characteristic choice is partitioned into parts: Attribute Evaluator and seek technique having more than one strategies in every direction from which possible pick out.

The characteristic evaluator is the technique with the aid of using which the characteristic of a characteristic with inside the statistics set is evaluated similar to every, the significance of the classes, etc. Techniques and navigation methodologies are combos of capabilities with inside the statistics set with the intention to attain the most accuracy in overall performance or the most fulfilling length of a listing of decided on capabilities. Attribute evaluator and seek technique strategies may be configured via the talk packing containers as special with inside the characteristic evaluator. This technique with the aid of using which every characteristic or statistics characteristic with inside the statistics set is evaluated similar to the magnificence characteristic of characteristic instance low, high, and medium. The simple looking method and of a distinctive aggregate of capabilities with inside the statistics set are completed for accomplishing the most fulfilling version with the chosen capabilities. The assessment method does now no longer paintings for all varieties of statistics sets, it varies with the request from a particular seek methodology. Implementation of correlation characteristic can most effective be carried out with a particular kind or seek technique this is primarily based totally at the valuation of every characteristic. Mind choosing distinctive characteristic evaluators it's far the requirement to alternate the compatibility of seek strategies like grasping set of rules for the breadth-first set of rules with selected method. Data characteristic choice primarily based totally on correlation is a well-known technique for choosing the maximum applicable attributes with inside the statistics set is figuring out the correlation in statistics. It is a greater widespread connection with an green correlation Coefficient in which we are able to calculate the dependencies with the aid of using the notation of correlation among every characteristic and the magnificence of variables. Based in this calculation, the choice technique follows most effective the ones attributes which have a pleasing common to superb or poor correlation and coffee correlation value. Correlation characteristic assessment method that calls for using ranker seek technique in Weka as shown in figure 6.

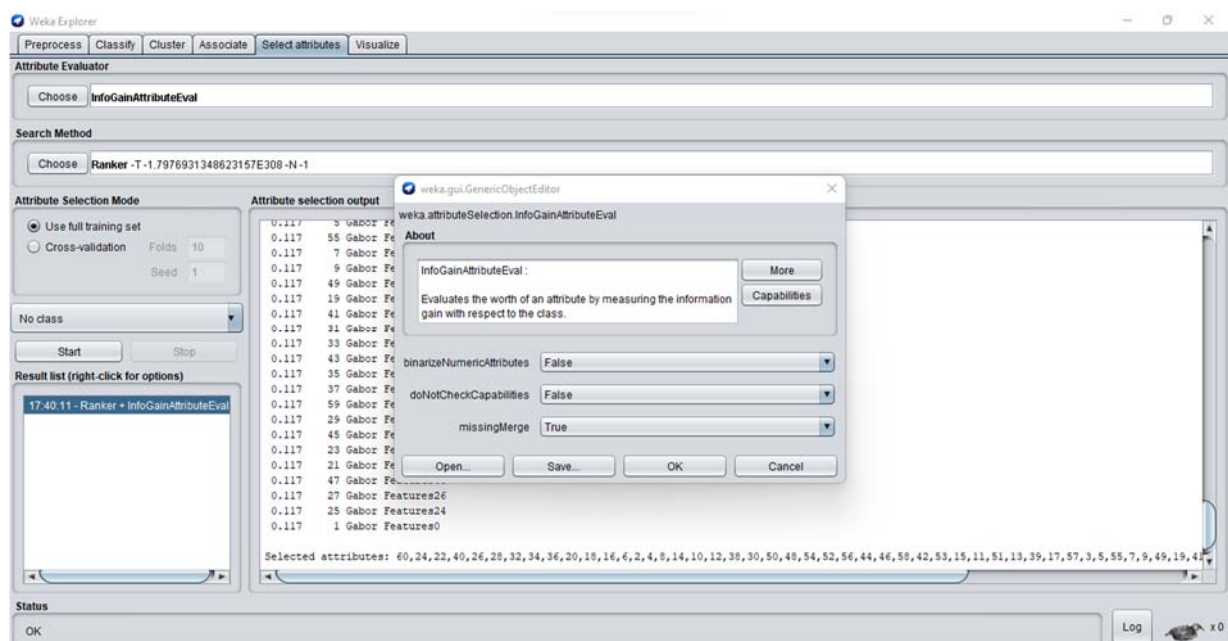


Figure 6 is the option panel for select attribute /Ranker in Weka

In Weka attribute selection info gain attribute evaluator can evaluate the worth of an attribute by measuring the information gain with respect to the class. It has the capability like handling binary class, missing class values, Nominal attributes, Numeric attributes, and Unary attributes. Attribute can be of the form of binary attributes, date attributes, empty nominal attributes and missing values nominal attributes numeric attributes and unary attributes.

The selected attribute starts from the range of 60, 24, 22, 40, 26, 28, 32, 34, 36, 20, 18, 16, 6, 2, 4, 8, 14, 10, 12, 38, 30, 50, 48, 54, 52, 56, 44, 46, 58, 42, 53, 15, 11, 51, 13, 39, 17, 57, 3, 5, 55, 7, 9, 49, 19, 41, 31, 33, 43, 35, 37, 59, 29, 45, 23, 21, 47, 27, 25, 1 : 60}. As shown in figure 7. It uses full dataset for training and 10 fold cross validation method without any class specification is done.

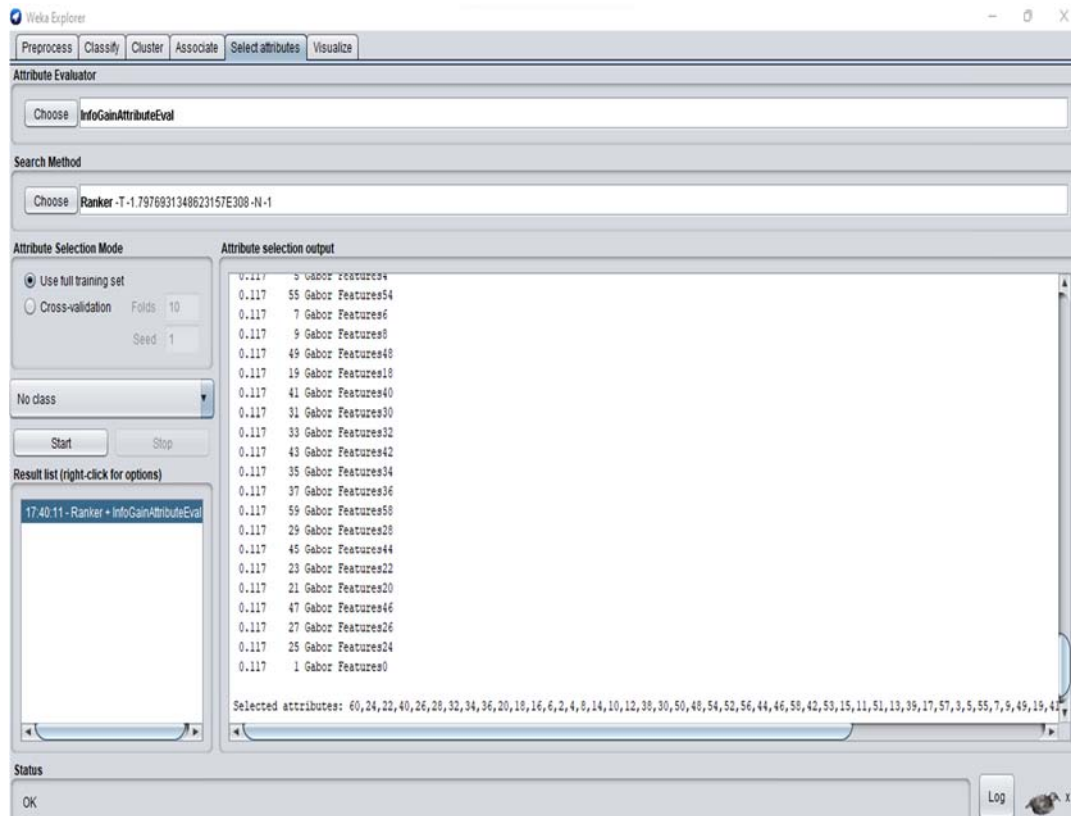


Figure 7 is the option panel for select infogain /Ranker in Weka

The run information of infogain attribute value it is as follows: The number of instances is 140 and 61 attributes. The ranks of all the features according to the information present in the features after evaluating the entire dataset. Attribute selection on all input data and the search method adopted is an attribute ranking and it is of information gained ranking filter that gives an output rank for all the 60 features.

The performance characteristics of the data set is then applied to the classifier tab where two different classifiers random forest and instance based classifiers are been applied to the data set with the retained attributes, number of attributes for various category the accuracy and the weighted-average receiver operator characteristics are tabulated in table 2 and 3.

The performance of random forest with feature selection module is done for both retaining all the attributes and removing the attributes in step by step with an account of 10 at a time. The result of the random forest classifier is such a way that accuracy is increased from 80.791% to 81.3559%. The average area under the curve or the weighted-average receiver operator Characteristics is 0.931 and 0.932.

Table 2: Performance of Random Forest with feature selection module

S.No	Uninvolved attribute list	Attribute selection session (Attribute assessor/Search method)	Tree based RF - Random Forest Classifier accuracy (%)	Tree based RF - Random Forest Classifier Area under the Curve [ROC]
1	Retaining all attributes (61 attributes are retained)	InfoGainAttributeEval / Ranker	80.791	0.931
2	Gobor Features removed - 16,02,18,20,24,40,14,46,12,48 and retained 51 attributes	InfoGainAttributeEval / Ranker	81.3559	0.932
3	Gobor Features removed - 58,38,30,52,32,26,44,50,4,6,22 and retained 40 attributes	InfoGainAttributeEval / Ranker	80.791	0.930
4	Gobor Features removed - 10,36,54,8,42,56,28,34,3 and retained 31 attributes	InfoGainAttributeEval / Ranker	74.011	0.875
5	Gobor Features removed - 51,53,17,13,59,55,57,11,23,31,49 and retained 20 attributes	InfoGainAttributeEval / Ranker	75.988	0.895
6	Gobor Features removed - 37,45,27,49,9,7,1 and retained 13 attributes	InfoGainAttributeEval / Ranker	75.706	0.907

The performance of instant based classifier IBK @1,2,3 selection module shows the performance characteristics of this classifier whose output accuracy percentage for retaining all the attributes is 75.141% with a weighted average ROC of point 0.848 as shown in table 3

Table 3: Presentation of IBK @ (K = 1,2,3) with data feature selection segment

S.No	Uninvolved attribute list	Attribute selection session (Attribute assessor/Se arch method)	IBk (k@1) Accuracy%	weighte d Avg. ROC (k@1)	IBk (k@2) Accuracy%	weighted Avg. ROC (k@2)	IBk (k@3) Accuracy%	weighted Avg. ROC (k@3)
1	Retaining all attributes (61 attributes retained)	InfoGainAttributeEval / Ranker	69.492	0.777	68.079	0.836	75.141	0.848
2	Gobor Features removed - 16,02,18,20,24,40,14,46,12,48 and retained 51 attributes	InfoGainAttributeEval / Ranker	68.644	0.772	66.949	0.828	75.424	0.851
3	Gobor Features removed - 58,38,30,52,32,26,44,50,4,6,22 and retained 40 attributes	InfoGainAttributeEval / Ranker	66.384	0.758	66.667	0.827	72.881	0.837
4	Gobor Features removed - 10,36,54,8,42,56,28,34,3and retained 31 attributes	InfoGainAttributeEval / Ranker	62.429	0.720	59.322	0.770	62.994	0.789
5	Gobor Features removed - 51,53,17,13,59,55,57,11,23,31,49 and retained 20 attributes	InfoGainAttributeEval / Ranker	63.842	0.734	62.7119	0.799	65.538	0.810
6	Gobor Features removed - 37,45,27,49,9,7,1 and retained 13 attributes	InfoGainAttributeEval / Ranker	66.667	0.766	63.277	0.871	64.972	0.811

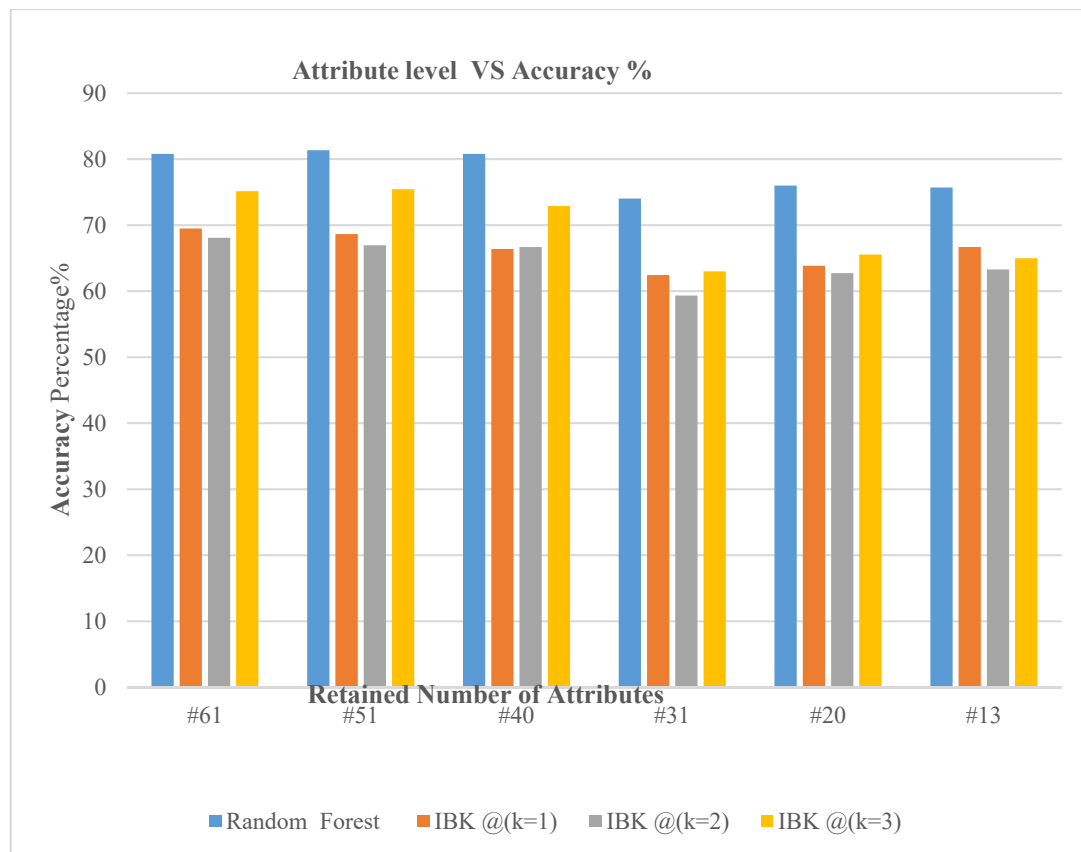


Figure 2 Presentation of RF - Random Forest & IBK For k=1,2,3 Selected attribute level vs accuracy

Performance characteristics of Random Forest, IBK Classifiers versus accuracy percentage for different attribute levels. On observing the results that the Random Forest Depending on the approach used to choose attributes, the accuracy ranges from 74.011 percent to 81.356 percent. When comparing 31 and 51 features, the Random Forest classification algorithm has the greatest accuracy rate of 81.356 percent.

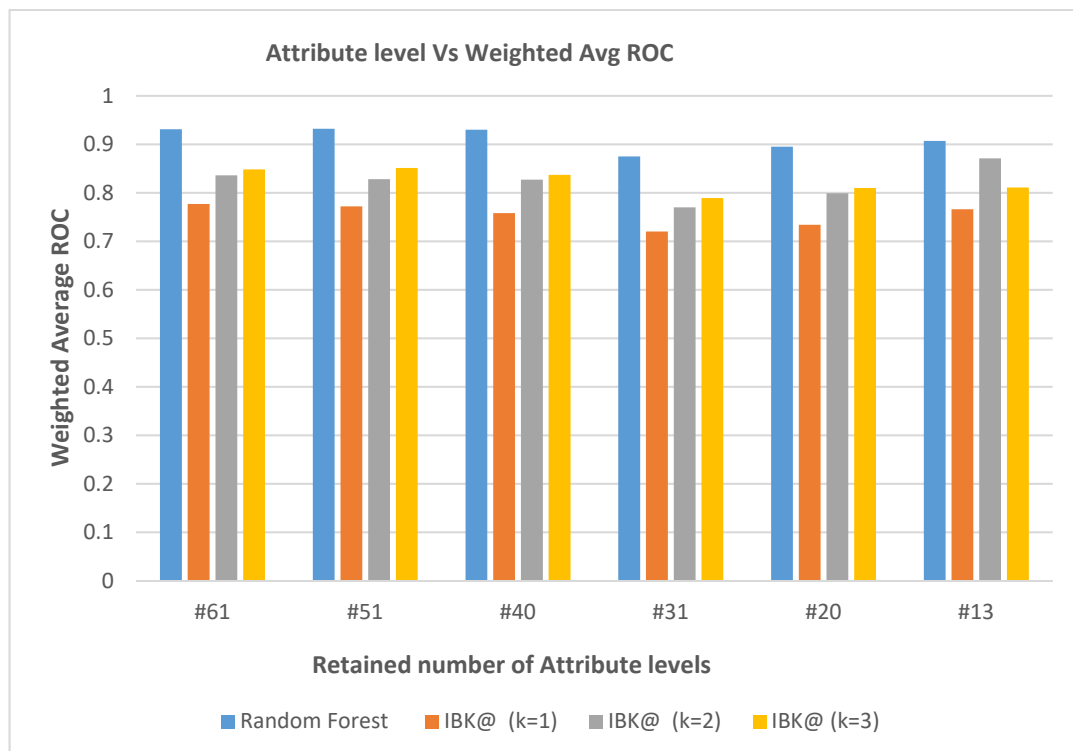


Figure 3 Performance of Random Forest & IBK (K = 1,2,3) attribute level vs Weighted Average ROC

The comparison of the Random & IBK algorithms with multi-k values is depicted in the graph above. Accuracy from a neighborhood of size $k = 1, 2, 3$ is represented by bars. With a $k = 3$ value, IBK achieves 75.424 percent accuracy. Based on the feature selection approach, we find that the Weight Avg. ROC achieved ranges from 0.851 to 0.720. Table 3 shows that when trained with 61, 51, and 40 selected features, the Random Forest classification algorithm yielded 0.932 high weighted average receiver operator characteristics (ROC) values. With varying k values, the IBK method is iterated. In the WEKA tool, the yellow line represents the standard k value, the blue line represents the nearest neighbour $k = 2$, and the red line represents the standard k value.

V. Deduction

The proposed system is made up of two primary modules with filters and classifications: attribute selection among filtered features in the first component, and iterations with instance-based closest neighbor models in the second. In the first component, the ideal accuracy with maximum performance was discovered to be 74.59 percent, and in the second component, it was 81.99 percent. These findings are the first of their type in this framework for level prediction utilizing digital image processing that is based on the Gabor image filter.

Table4 Conclusion Summary of research article inferences:

S.No	Interpretation of results	Performance Evaluating Criteria	Alterations in the examinations	Final output result
1	Feature Selection for Gabor filter based on level measurement using Non - interacting Tanks Level images	Accuracy: 81.356%, and Weighted average ROC: 0.931	Selected attributes with information gain and ranker, Increased K value in Instance base Classifier	Diverse approaches and obtained results in a structure arising new emergence of measurement.

FUTURE POSSIBILITIES

1. The Gabor image filter, in particular, allows us to identify the image set of level scenarios in a more efficient manner. These filters and expanded classifiers should be investigated further and evaluated for improved performance and accuracy
2. Trying other image-based parameter measurements and monitoring can be implemented.
3. Alterations are done on many more selections of ensembles other than Random Forest adopted in this study and framework.

AGREEMENT WITH MORAL STANDARDS

Ethical approval: Any of the authors' investigations with human participants or animals are not included in this article.

References

- [1] T. Hagglund. A friction compensator for pneumatic control " valves. Journal of Process Control, 2002; vol. 12, no. 8, pp. 897–904.
- [2] M. A. De Souza L. Cuadros, C. J. Munaro, and S. Munareto. Improved stiction compensation in pneumatic control valves. Computers and Chemical Engineering, 2012; vol. 38, pp. 106–114.
- [3] R. P. Champagne and S. J. Boyle. Optimizing valve actuator parameters to enhance control valve performance. ISA Transactions, 1996; vol. 35, no. 3, pp. 217–223.
- [4] K. Ogata. Modern Control Engineering. Prentice-Hall, Upper Saddle River, NJ, USA, 5th edition, 2010.
- [5] Basak, D, Pal, S, Patranabis, D.C. Support vector regression. Neural Inf. Process. -Lett. Rev. 2007; 11, 203–224.
- [6] S. Zerkaoui, F. Druaux, E. Leclercq, and D. Lefebvre. Stable adaptive control with recurrent neural networks for square MIMO non-linear systems. Engineering Applications of Artificial Intelligence. 2008; vol. 22, no. 4-5, p
- [7] M.A Cuadros, C.J. Munaro and S. Munareto. Improved stiction compensation in pneumatic control valves. Computers & Chemical Engineering. 2012; vol. 38, pp. 106-114.
- [8] Puneet Mishra, Vineet Kumar, and K. P. S. Rana. An online tuned novel nonlinear PI controller for stiction compensation in pneumatic control valves. ISA Transactions. 2015; vol. 58, pp. 434-445.
- [9] Pardeep Rohilla, Vineet Kumar. Fuzzy I+PD controller for stiction compensation in pneumatic control valve. International Journal of Applied Engineering Research 2017; Volume 12, no.13, pp. 3566-3575, ISSN 0973-4562.
- [10] Srinivasan Arumugam and R.C. Panda. Identification of stiction nonlinearity for pneumatic control valve using and method. International Journal of Engineering & Technology. 2014; vol. 6, no. 2, pp. 570-578.
- [11] Fei Li, You Jun Fan, Hus Tian Zhao. The Research of a Piezoelectric Valve Positioner Control Algorithm[J]. Advanced Materials Research, 2014; 3470
- [12] Wang J, Wang D J D, Moore P R. Modelling Study, Analysis and Robust Servo Control of Pneumatic Cylinder Actuator System[J]. Control Theory and Applications, 2014; 148(1): 35-41.
- [13] N.Saravanan, V Gayathiri. Performance and Classification Evaluation of J48 Algorithm and Kendall's Based J48 Algorithm (KNJ48). International Journal of Computational Intelligence and Informatics. 2018; Vol. 7: No. 4.
- [14] P. Wijerathna and L. Ranathunga. Rice Category Identification using Heuristic Feature Guided Machine Vision Approach. 2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS), 2018; pp. 185-190, doi: 10.1109/ICIINFS.2018.8721396.
- [15] O.D. Trier, A.K. Jain. Goal-directed evaluation of binarization methods. IEEE Trans. 1995; PAMI 16 (12) 1191–1201.
- [16] W. Niblack. An Introduction to Digital Image Processing. Prentice-Hall, Englewood Cliffs, NJ, 1986;

- [17] Minyechil Alehegn, R. J. Analysis and prediction of diabetes diseases using machine learning algorithm: Ensemble approach. International Research Journal of Engineering and Technology (IRJET). 2017.
- [18] Mohammed M Mazid, A. B. Improved C4.5 Algorithm for Rule-Based Classification. Recent Advances in Artificial Intelligence, Knowledge Engineering and Data Bases. 2013.
- [19] Nagaparameshwara chary, S. D. A Survey on Comparative Analysis of Decision Tree Algorithms in Data Mining. International Conference on Innovative Applications in Engineering and Information Technology (ICIAEIT-2017).
- [20] Dr. Anjali B Raut, A. A. Students Performance Prediction Using Decision Tree Technique. International Journal of Computational Intelligence Research, 2017; 1735-1741.
- [21] Gonzalez, Digital Image Processing, Electronic Industry Press.2005.
- [22] J.R. Parker. Image Processing and Computer Vision Algorithms and Applications, Tsinghua University Press.2006.
- [23] Paaranan Sivasothy, Jens Wagener and Jörg Seewig. Machine learning based filling level estimation for bulk solid silos Institute for Measurement and Sensor-Technology, Technische Universität Kaiserslautern, Kaiserslautern, Rheinlandpfalz, GERMANY, 2018; Vol. 35, 055002
- [24] Choi, Changhyun, Jungwook Kim, Heechan Han, Daegun Han, and Hung S. Kim. Development of Water Level Prediction Models Using Machine Learning in Wetlands: A Case Study of Upo Wetland in South Korea. 2020; Water 12, no. 1: 93. <https://doi.org/10.3390/w12010093>
- [25] Vijayan T , Sangeetha M , A. Kumaravel, Karthik B , Gabor Filter and Machine learning Based Diabetic Retinopathy Analysis and Detection, Microprocessors and Microsystem, 2020; doi: <https://doi.org/10.1016/j.micpro.2020.103353>
- [26] F. N. M. Ariff, A. S. A. Nasir, H. Jaafar and A. N. Zulkifli, "Character Segmentation for Automatic Vehicle License Plate Recognition Based on Fast K-Means Clustering," 2020 IEEE 10th International Conference on System Engineering and Technology (ICSET), 2020, pp. 228-233, doi: 10.1109/ICSET51301.2020.9265387.
- [27] V. V. Mainkar, J. A. Katkar, A. B. Upade and P. R. Pednekar, "Handwritten Character Recognition to Obtain Editable Text," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 599-602, DOI: 10.1109/ICESC48915.2020.9155786.
- [28] Y. Peng and Z. Chen, "Application of Deep Residual Neural Network to Water Meter Reading Recognition," 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), 2020, pp. 774-777, doi: 10.1109/ICAICA50127.2020.9182460.

Authors Profile



Mrs B. Kalaiselvi working as Assistant Professor and Part Time Research Scholar in the department of Electronics and Communication Engineering, Bharath Institute of Higher Education and Research, Chennai. She Received her M. Tech degree in Electronics and Instrumentation Engineering form Bharath University, Chennai, in 2014. She has received her B.E Electronics and Communication Engineering from Satyabhama University, Chennai in 2011. Her current area of interest is Measurements and Instrumentation, Digital image Processing Artificial Intelligence and Machine Learning. Email:kalaigopal1973@gmail.com



Dr. B. Karthik is an Associate professor in the Department of Electronics and Communication Engineering, Bharath Institute of Higher Education and Research, Chennai, India. He received his PhD (ECE) – Image Processing in May 2017. He received his Master of Engineering - Applied Electronics at Thiruvalluvar College of Engineering & Technology in 2011. He received his Bachelor of Technology in Electronics and Communication Engineering at Bharath Institute of Higher Education and Research, Chennai, India in 2007. His area of interests includes Image & Video Processing, Cryptography, and Network Security System Techniques. He is a member of IEEE, ITEEA, IAIP, CSTA and IAENG. Email: :karthikguru33@gmail.com



Dr A. Kumaravel is working as a Professor and Dean, School of Computing, Bharath Institute of Higher Education and Research, Chennai. His research interest includes Functional Paradigm, Soft Computing, Cloud Computing, Machine Learning, Computational Intelligence and Knowledge Engineering. He is a life Member of ISTE and IET. He is a reviewer for IEEE transactions and Elsevier journals. Email:drkumaravel@gmail.com