iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** II **Month of publication:** February 2025

DOI: https://doi.org/10.22214/ijraset.2025.67133

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue II Feb 2025- Available at www.ijraset.com

An Efficient Optimization Approach for Steel Surface Flaw Classification using Machine Learning

Dr. Edison Prabhu K[1], Godlin Debby K J[2]
[1]Assistant Professor, Department of Electrical and Electronics Engineering, Nehru Institute of Engineering and Technology, Coimbatore, India
[2]Assistant Professor, Department of Information Technology, Nehru Institute of Technology, Coimbatore, India

Abstract: Automatic steel surface defect classification is a well-known problem and is being considered for more than two decades. On the other hand, manual methods for defect classification in steel products are time consuming, labor intensive, and are prone to error. As a result, a real-time surface defect classification method using genetic algorithm is proposed to classify four kinds of surface defects in steel surfaces such as Roller Mark, Rust Spot, Emulsion Spot and Scrape. The proposed inspection system consists of visual dataset of 500 steel defect images for processing. The overview of the classification model has online and offline process. In offline, the input image is pre-processed and the following visual features such as geometric features, shape features and texture features are extracted from the defect image. In order to optimize the extracted visual features and to improve the accuracy of the classification model, genetic algorithm is used to find an optimal solution. For optimization, an initial population is generated randomly and the fitness function is calculated for every chromosome. Then the genetic operations such as selection, crossover and mutation are done to generate new offspring chromosome and evaluated to pick optimal chromosome or optimal feature. Finally, in online, the selected features are given as input to the Optimized SVM classifier, in order to predict the class of defect to which the input image belongs. Accuracy, sensitivity, specificity, f-measure and precision are used as the performance metrics to evaluate the system.
Keywords: Deep Learning, Genetic Algorithms, SVM Classifier, Convolutional Neural Network

I. INTRODUCTION

In this paper Image processing is a method to convert an image into digital form and perform some operations on it, to get an enhanced image or to extract some useful information from it. The term digital image processing generally refers to processing of a two-dimensional picture by a digital computer. A digital image is an array of real numbers represented by a finite number of elements arranged in rows and columns. These elements are called a pixel (picture elements) which has a particular value and location [2]. An image feature is a piece of information of a particular image which is relevant for solving the computational task related to a certain application. It is a primitive characteristic or attribute of an image. Features may be specific structures in the image such as points, edges or objects.

To capture the descriptive feature information from an image, the geometric, the shape and texture features are extracted.

1) *Geometric Features:* Geometric feature methods extract distinctive geometric features from images. Geometric features are features of objects constructed by a set of geometric elements like points, lines, curves or surfaces.
2) *Shape Features:* The shape of an object is an important and basic visual feature for describing image content. It provides vital information about the shape of an object.
3) *Texture Features:* Image texture is a set of metrics calculated in image processing designed to quantify the perceived texture of an image. Image texture gives us information about the spatial arrangement of color or intensities in an image or selected region of an image.

A. *Image Classification*

Image has been processed to put each pixel into a category. The goal of image classification is to predict the categories of the input image using its features. Depending on the interaction between the analyst and the computer during classification, it can be broadly categorized into two types namely i) *Supervised classification* and ii) *Unsupervised classification*

1) *Supervised Classification:* Supervised classification can be defined as the process of samples of known identity to classify pixels of unknown identity. Pixels located within these areas term the training samples used to guide the classification algorithm to assigning specific values to appropriate informational class.
2) *Unsupervised Classification:* Unsupervised classification does not use training data as the basis for classification. Rather, this family of classifiers involves algorithms that examine the unknown pixels in an image and aggregate them into a number of classes based on the natural groupings or clusters present in the image values.

II. OVERVIEW OF GENETIC ALGORITHM

Genetic algorithm is a part of evolutionary computing, a rapidly growing area of artificial intelligence and used to solve optimization problems.

A. *Biological Background*

It deals with the mechanisms responsible for similarities and differences in species called Genetics. The concepts of Genetic Algorithms are directly derived from natural evolution

B. *Fitness Function*

The fitness of an individual in a genetic algorithm is the value of an objective function. For calculating fitness, the chromosome has to be first decoded and the objective function has to be evaluated. The fitness not only indicates how good the solution is, but also corresponds to how close the chromosome is to the optimal one.

C. *Encoding*

Encoding is a process of representing individual genes. Each chromosome encodes a binary (bit) string. Each bit in the string can represent some characteristics of the solution. The process can be performed using bits, numbers, trees, arrays, lists or any other objects.

D. *Genetic Operators*

The breeding process is the heart of the genetic algorithm. Chromosomes are selected from the initial population to be parents for reproduction. The three genetic operators used to generate new offspring are selection, crossover and mutation.

1) *Selection:* Selection is the process of choosing two parents from the population for crossing. After encoding, the next step is to decide how to perform selection i.e., how to choose individuals in the population that will create offspring for the next generation and how many offspring each will create**.**
2) *Crossover:* Cross over is the process of taking two parent solutions and producing from them a child. After the selection process, the population is enriched with better individuals. The traditional genetic algorithm uses single point crossover, where the two mating chromosomes are cut once at corresponding points and the sections after the cuts exchanged. Here, a crossover point is selected randomly along the length of the mated strings and bits next to the cross-sites are exchanged.
3) *Mutation*: After crossover, the strings are subjected to mutation. There are different forms of mutation for different kinds of representation. For binary representation, a simple mutation can consist in inverting the value of each gene with a small probability. The probability is usually taken about 1/L, where L is the length of the chromosome.

E. *Termination*

The stopping criteria or search termination depends on maximum generation, elapsed time stall generations and stall time limit.

F. *Need for the Study*

In steel firms, the use of surface defect detection and classification is necessary to improve the quality of product, accuracy as well as to eliminate the need of a human intervention in a hazardous environment.

G. *Corresponding Work*

Various Procedures for Image Classification Reza Entezari Maleki *et al.* (2013), have made a comparative study of various classifiers based on the effect of parameters including the size of dataset and kind of independent attributes and the number of discrete attributes.

The SVM classifier has higher efficiency when compared to other classifiers. Arun Irtaza *et al* (2010), have presented a new method for extracting the features using Wavelet transform, Curvelet transform and Gabor filters. The features extracted by using these transform and filters are mean, variance, entropy and eigen value. Binary classifier is used to classify the images into relevant and non-relevant images. To classify the image into relevant and non- relevant fitness function is calculated using genetic algorithm. And this method can easily avoid the irrelevant images with high precision. Someyh Mousavi *et al* (2012), have proposed a new method to classify the image semantically using Genetic Algorithm. Zernike Moments are calculated using radial polynomials in both RGB and HSV Color Space. Then to transform the crisp input into degree of match, fuzzification is done. By using the fuzzification value the fuzzy rule base is built with Mamdani fuzzy inference system. Genetic Algorithm compares both Fuzzy Inference System and Zernike Moments and obtained an optimum MF for further classification. The dataset used here is COREL and it classified into five classes with low complexity and the approach overcomes the semantic concept gap accurately. Mei-Yun Chen *et al.* (2012), have presented non-linear SVM, which requires high computational complexity as the training data grows. It evaluates the performance of feature descriptors using linear SVM model. It improves the performance of original SIFT and LBP feature descriptors. Aarti Kaushik *et al*. (2013), have identified that Support Vector Machine is kernel-type sensitive and hence, Data Miner Analyst must ensure the choice of correct kernel parameter for particular data set. Support Vector Machines have shown their great promise in many multitudinous areas and in few cases, they have surpassed other methods. The important disadvantage of SVM is computational appeal because it includes quadratic optimization problem. The SVMs has solved various realistic problems varying from economics to genetics.

Md. Hafizur Rahman *et al.* (2013), have built a model for extracting two sets of data for both male and female and separate them accurately is a challenging job. Among all kinds of recognition algorithms, Support Vector Machine (SVM) has provides a sound theoretic basis for constructing classification models with high generalization ability. So SVM classifier is then used to recognize the facial features. It is proved that SVM can provide superior performance. Different kernel functions have been useful in cases where the data are not linearly separable. These kernel functions transform data to higher dimensional space where they can be separated easily.

III. PROPOSED METHOD

All the works reported under various procedures for Surface Defect Classification in literature survey prove that the Support Vector Machine (SVM) is the best method for the classification of defects. It is very difficult to predict the class of defect with more accuracy. Genetic Algorithm may provide better result for feature selection. Large collection of samples will be required for predicting the defects accurately.

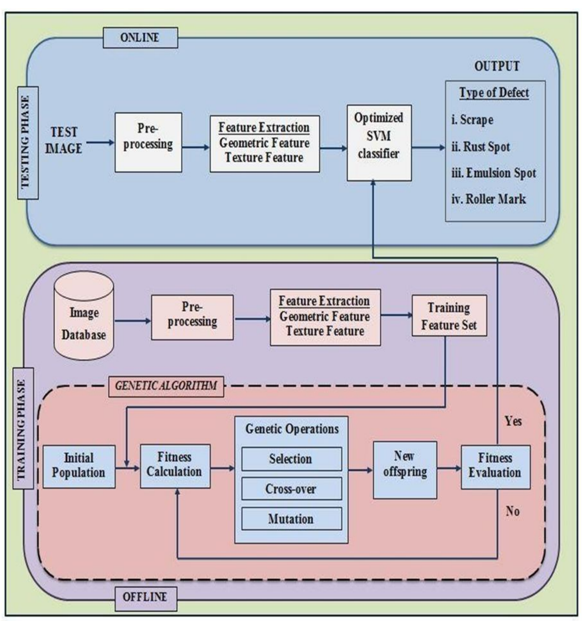


Figure 1: Architecture Diagram

The architecture diagram of the proposed system is shown in Fig 1. It consists of two parts: (i) online and (ii) offline. In offline the steel image is given as input to the system and its noise will be removed by median and average filters. Then eliminate uneven illumination and enhance the image by removing blur in the image. For further processing the input image is binarized using gray thresholding. The geometric features such as perimeter, centroid is computed. The texture feature can be extracted by using local binary pattern, gray-level co- occurance matrix (GLCM) and tamura texture features. From that, the statistical features such as mean, variance, entropy, homogeneity, skewness, energy, contrast are computed. The shape features such as rectangularity, density is extracted. Then the classification is done by using Multi-class support vector machine. In order to optimize the extracted visual features and to improve the accuracy of the classification model, genetic algorithm is used to identify optimal features. For optimization, an initial population is generated randomly and the fitness function is calculated for every chromosome. Then, generate offspring chromosomes by applying the genetic operators such as selection, crossover and mutation and compute fitness evaluation for offspring chromosome. If fitness evaluation is satisfied, the offspring is placed in the new population and given as input to the optimized SVM classifier or else generate new generation again. Finally, in online, the selected features are given as input to the Optimized SVM classifier, in order to predict the class of defect to which the input image belongs to.

A. Modules

The various modules are

1. Image pre-processing 2. Feature extraction
3. Feature selection 4. Classification

IV. METHODOLOGY

A. Image Pre-processing

The steel image is given as input and it is converted to gray-scale image for further processing. Then add noise by using imnoise function and removed by using median and average filters. Then resize the image and improve the contrast of the image. For further processing the gray image is binarized by using gray thresholding. Feature extraction to capture the descriptive feature information, the geometric, the shape and the texture features are extracted from the input image.

a) Geometric features

Geometric images provide vital information about the shape of objects in an image. Totally 3 shape features are extracted such as

(a) Area

Area is the size of a surface or the amount of space inside the boundary of a 2-d object. It is calculated by counting the number of non-zero pixels in the region. (b) Centroid.

Centroid is the center of mass of a geometric object of uniform density.

b) Shape features

Shape feature provides information about the shape of the objects in an image. Totally 2 features are extracted from the input image such as

(a) Rectangularity means the property of being like a rectangle.

(b)Density is a measurement of the pixel density (resolution) of an image.

c) Texture features

Texture gives information about structural arrangement of surfaces and objects in the image. It depends on the distribution of intensity over the image and it is not defined for single pixel. Here, texture features are extracted by using following methods such as Local binary pattern (LBP), Gray-level co-occurrence matrix (GLCM) and Tamura texture features.

1) Local Binary Pattern

To extract the texture features by using local binary pattern, the gray image is given as input to find local patterns in a texture as binary codes. Then compute local binary Pattern histogram and the statistical features of the histogram are calculated.

Mean

The mean is simply the arithmetic average of the values in the set, obtained by summing the values and dividing by the number of values. It is given by Equation

$$\text{Mean} = \sum^{L-1} (b) \qquad (1)$$

Where p(b) is the probability value of row and column in the image.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue II Feb 2025- Available at www.ijraset.com

Variance

The variance is the arithmetic average of the squared differences between the values and the mean. It is given by Equation

Variance = $\sqrt{\sum_{b=0}^{L-1}(b-b)^2 p(b)}$ (2)

Where p(b) is the probability value of row and column in the image and $(b-b)^2$ is the squared difference of the mean.

Skewness

Skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution is symmetric if it looks the same to the left and right of the center point. It is given by Equation

Skewness = $1/variance^3 \sqrt{\sum_{b=0}^{L-1}(b-b)^3 p(b)}$ (3)

Where p(b) is the probability value of row and column in the image and $(b-b)^3$ is the cubic difference of the mean and cubic variance

Kurtosis

Kurtosis is a measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution. It is given by Equation.

Kurtosis = $1/variance^4 \sqrt{\sum_{b=0}^{L-1}(b-b)^4 p(b)}$ (4)

Where p(b) is the probability value of row and column in the image and $(b-b)^4$ is the difference of the mean.

2) Gray-level co-occurrence matrix

A statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM). The features such as homogeneity, contrast, energy and entropy are extracted using GLCM.

a) Homogeneity

Homogeneity returns a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. It measures the number of local changes in the image texture. It is given by Equation.

Homogeneity = $\sum_i \sum_j p_{ij}/1 + |i-j|$ (5)

Where i is the number of rows in the image, j is the number of columns in the image and p(i,j) is the gray level value at the co-ordinate.

b) Contrast

Contrast returns the measure of intensity contrast between a pixel and its neighbour over the whole image. It is given by Equation.

Contrast = $\sum_i \sum_j (i-j)^2 p_{ij}$ (6)

Where i is the number of rows in the image, j is the number of columns in the image and p(i,j) is the gray level value at the co-ordinate.

c) Correlation

Correlation returns a measure of how correlated a pixel is to its neighbour over the whole image. It is given by Equation.

Correlation = $\sum_i \sum_j (i-mc)(j-md)p_{ij}/\sigma c\sigma d$ (7)

Where m_c and m_d are the mean values of row and column in the image, σc and σd the standard deviation of row and column in an image.

d) Energy

Energy is the sum of squared elements in the GLCM. It is also known as angular second moment. It is given by Equation

Energy = $\sum x \sum y\, p(x, y)^2$ (8)

Where x is the number of rows in the image, y is the number of columns in the image and p(x,y) is the gray level value at the co-ordinate.

e) Entropy

Entropy is a statistical measure of randomness and it can be used to characterize the texture of the input image. It is given by Equation.

Entropy $= \sum i \sum j\ pij\,log2\ pij$ (9)

3) Tamura Features

Tamura feature is an approach that explores texture representation from a different angle since it is motivated by the psychological studies on human visual perception of textures.

a) Tamura coarseness

Coarseness relates to distances of notable spatial variations of grey levels, that is, implicitly, to the size of the primitive elements (texels) forming the texture. The proposed computational procedure accounts for differences between the average signals for the non-overlapping windows of different size:

1. At each pixel compute six averages for the windows of size 2^k around the pixel.
2. At each pixel, compute absolute differences between the pairs of non-overlapping averages in the horizontal and vertical directions.
3. At each pixel, find the value of k that maximizes the difference in both direction and set the best size.
4. Compute the coarseness feature by averaging over the entire image.

b) Tamura contrast

Tamura contrast measures how grey levels vary in the image and to what extent their distribution is biased to black or white.

c) Tamura directionality

Tamura directionality is measured using the frequency distribution of oriented local edges against their directional angles. The edge strength and the directional angle are computed using the Sobel edge detector approximating the pixel-wise x-derivatives and y-derivatives of the image. A histogram of quantized direction values is constructed by counting numbers of the edge pixels with the corresponding directional angles and the edge strength greater than a predefined threshold.

d) Tamura roughness

The roughness feature is described by simply summing the coarseness and contrast measures. It is given by Equation

Roughness $= F_{coarseness}+F_{contrast}$ (10)

B. Feature selection

The feature selection is done by genetic algorithm. Here, a population of candidate solutions to an optimization problem is evolved toward better solutions. Each candidate solution has a set of properties which can be mutated and altered; traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. The evolution usually starts from a population of randomly generated individuals, and is an iterative process, with the population in each iteration called a generation. In each generation, the fitness of every individual in the population is evaluated; the fitness is usually the value of the objective function in the optimization problem being solved. The more fit individuals are stochastically selected from the current population, and each individual's genome is modified to form a new generation. The new generation of candidate solutions is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population. The flow of a Genetic Algorithm is shown in Figure

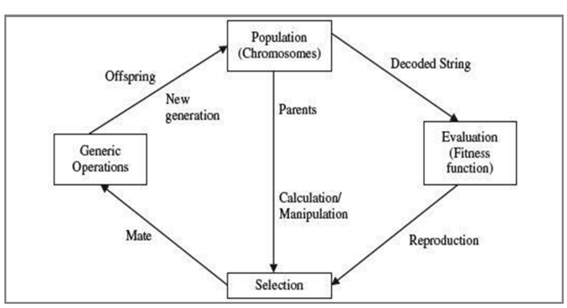


Figure 2: Flow of a Genetic Algorithm

Classification

A Support Vector Machine (SVM) constructs a hyper-plane or set of hyper-planes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyper-plane that has the largest distance to the nearest training- data point of any class. The hyper-planes in the higher-dimensional space are defined as the set of points whose dot product with a vector. The vectors defining the hyper-planes can be chosen to be linear parameters of images of feature vectors.

C. Implementation Aspects

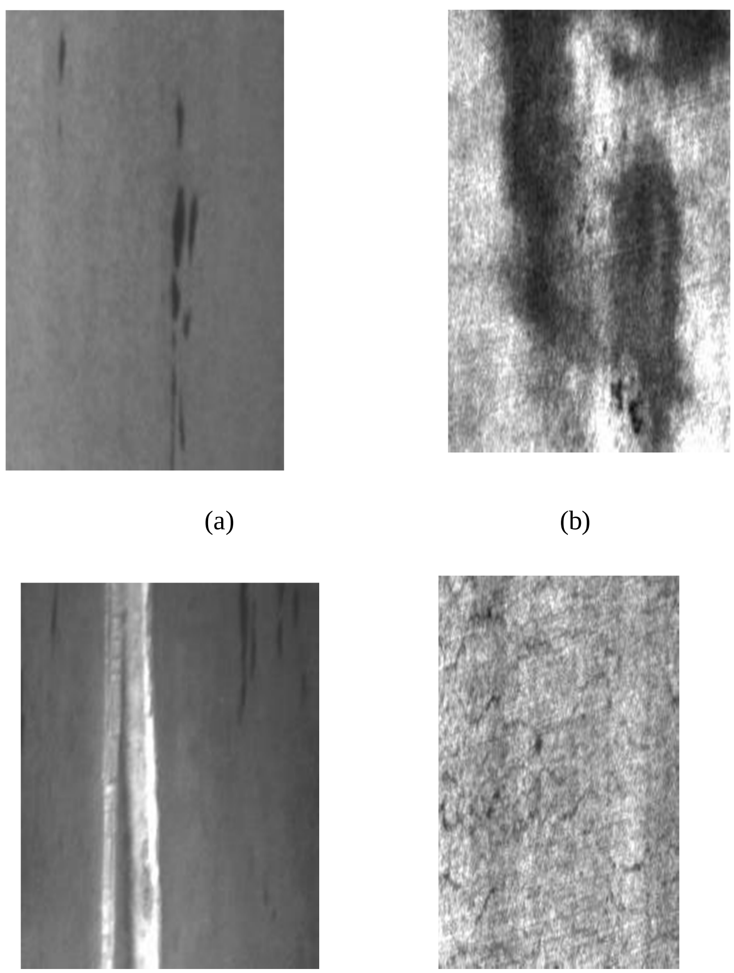
In the proposed system, MATLAB R2022 is used for implementation. The database of proposed system consists of 500 steel image samples. For evaluating the method 100 samples are used.

1) Modules

The various modules are

1. Image pre-processing
2. Feature extraction
3. Feature selection
4. Classification

The steel images acquired from camera corresponding to various surface defects like scrape, rust spot, emulsion spot and roller mark are shown in Fig 3 (a), (b), (c), (d).



(a) (b)

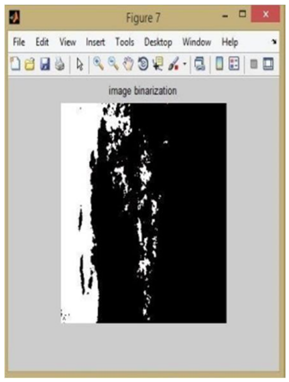
(c) (d)

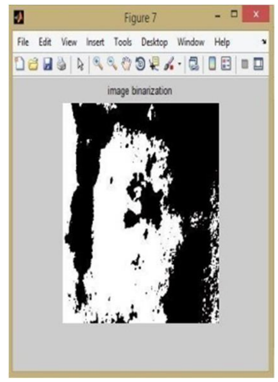
Figure 3: Various input steel defect images

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue II Feb 2025- Available at www.ijraset.com

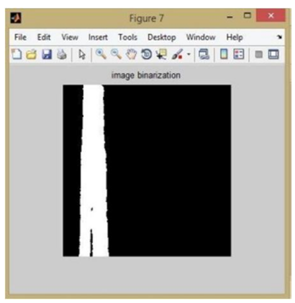
D. Image pre-processing

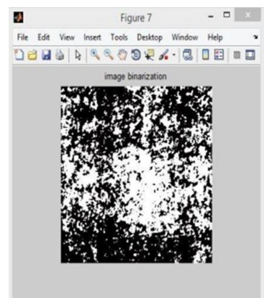
The image samples are given as input for pre-processing to remove noise present in the image and to binarize the image for further processing. The binarized images of four kinds of steel defect images such as scrape, rust spot, emulsion spot and roller mark are shown in Fig 4 (a),(b),(c) and (d).





(a) (b)





(c) (d)

Figure 4: Image pre-processing

E. Feature Extraction

To capture the descriptive feature information, the geometric, and the texture features are discussed in section 4.2. Totally 16 features are extracted from the four kinds of steel defect images for defect classification. Table 4.1 describes various geometric features, Table 4.2 describes texture features such as mean, variance, skewness and kurtosis extracted by using Local binary pattern (LBP) method, Table 4.3 describes texture features such as homogeneity, contrast, entropy and energy extracted by using Gray-level concurrence matrix (GLCM), and Table 4.4 describes texture features such as tamura oarseness, tamura_contrast, tamura_directionality and tamura_roughness extracted by using tamura texture feature extraction method.

Table 4.1 Geometric features-Area, Centroid

Sample images	Area	Centroid
Scrape	49193.38	0.802207
Rust Spot	30665.63	0.609381
Emulsion Spot	7090.25	0.028223
Roller Mark	30915.13	0.538459

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue II Feb 2025- Available at www.ijraset.com

Table 4.2 LBP texture features-Mean, Variance, Kurtosis, Skewness

Sample images	Mean	Variance	Kurtosis	Skewness
Scrape	0.312785	0.156169	0.395182	0.4195
Rust Spot	0.318809	0.159223	0.399027	0.100343
Emulsion Spot	0.316799	0.158684	0.398351	0.326195
Roller Mark	0.321827	0.15966	0.399575	0.204277

Table 4.3 GLCM texture features-Homogeneity, Contrast, Energy, Entropy

Sample images	Homogeneity	Contrast	Correlation	Energy
Scrape	0.753265	1.864847	0.987232	0.025536
Rust Spot	0.800447	1.930945	0.886223	0.232292
Emulsion Spot	0.800126	1.954515	0.955766	0.091238
Roller Mark	0.749844	1.870094	0.856424	0.29185

Table 4.4 Tamura texture features- Tamura_coarseness, Tamura_contrast, Tamura_directionality, Tamura_roughness

Sample Images	Tamura Coarseness	Tamura Contrast	Tamura directionality	Tamura roughness
Scrape	50.58673	8.297473	0.438622	58.8842
Rust Spot	47.79825	64.38797	0.208282	112.1862
Emulsion Spot	46.82217	23.28077	0.264233	70.10294
Roller Mark	46.14563	21.03417	0	67.1798

4.4 and 4.5 describes population setting options, genetic operation options such as selection, crossover and mutation and output for feature selection.

F. *Feature Selection*

The feature selection is discussed in the above section. It is done by using an evolutionary algorithm known as genetic algorithm. To do this, an optimtool is used for optimization. In that tool, feed the input as per the requirements. Fig .5,5.1,5.2

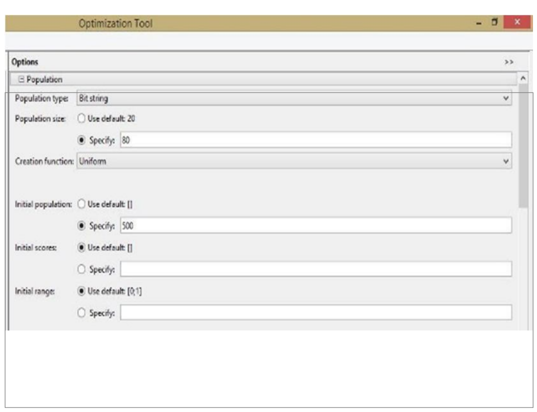


Figure 5 GA population setting

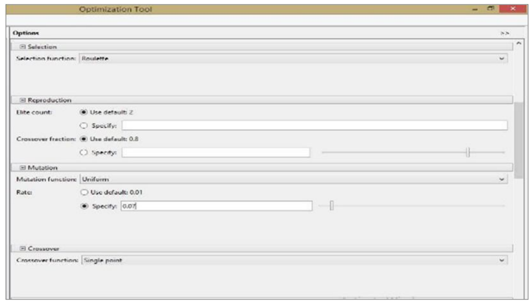


Figure 5.1 GA operators options

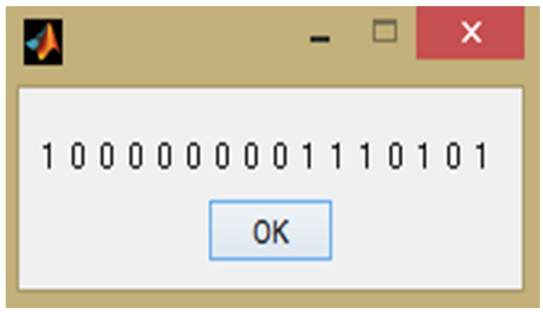


Figure 5.2 GA output for feature selection

The selected 6 features such as homogeneity, contrast, correlation, tamura_coarseness, tamura_contrast and tamura_roughness are tabulated in Table 4.5 & Table 4.6.

Table 4.5 Selected features- Homogeneity, Contrast, Correlation

Sample images	Homogeneity	Contrast	Correlation
Scrape	0.753265	1.864847	0.987232
Rust Spot	0.800447	1.930945	0.886223
Emulsion Spot	0.800126	1.954515	0.955766
Roller Mark	0.749844	1.870094	0.856424

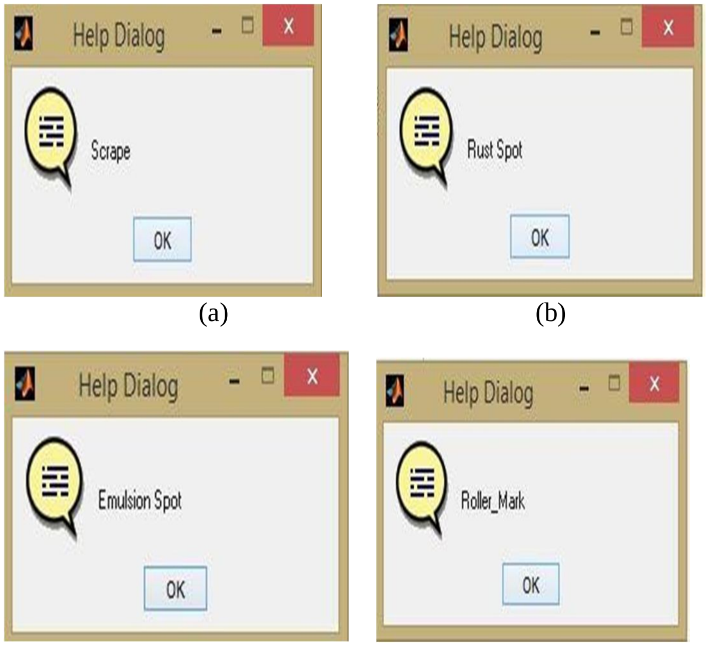
Table 4.6 Selected features- Tamura_coarseness, Tamura_contrast and Tamura_roughness

Sample images	Tamura coarseness	Tamura contrast	Tamura roughness
Scrape	50.58673	50.58673	50.58673
Rust Spot	8.297473	8.297473	8.297473
Emulsion Spot	47.79825	47.79825	47.79825
Roller Mark	64.38797	64.38797	64.38797

Classification

Selected features are shown in above Tables 4.4 and Table 4.5 are given as input for Support Vector Machine (SVM). Fig 6 (a), (b), (c) and (d) describes four types of different classes.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue II Feb 2025- Available at www.ijraset.com



(a) (b) (c) (d)

Figure 6 Different classes of defects

V. RESULTS AND DISCUSSION

The metrics used for the evaluation of proposed method are (i) Sensitivity, (ii) Specificity, (iii) Accuracy, (iv) Precision, (v) F-measure.

A. Sensitivity

Sensitivity is also called the true positive rate, or the recall. It measures the proportion of positives that are correctly identified. The sensitivity is given by Equation

$$\textbf{Sensitivity} = \frac{TP}{TP+FN}$$

Where TP is True Positive (TP), FN is False Negative (FN).

B. Specificity

Specificity is also called the true negative rate. It measures the proportion of negatives that are correctly identified. The specificity is given by Equation

$$\textbf{Specificity} = \frac{T\ N}{TN+FP}$$

Where TN is True Negative (TN), FP is False Positive (FP)

C. Accuracy

Accuracy is the percentage of test set correctly classified by classifier. It is given by Equation

$$\textbf{Accuracy} = \frac{TP^{+}TN}{TP+TN+FP+FN}$$

Where TP is True Positive (TP), TN is True Negative (TN), FP is False Positive (FP) and FN is False Negative (FN).

D. Precision

Precision is referred to as positive predictive value. It is given by Equation

$$\textbf{Precision} = \frac{TP}{TP+FP}$$

Where TP is True Positive (TP) and FP is False Positive (FP)

E. F-measure

It is the harmonic mean of precision and recall. It is given by Equation

$$\textbf{F-measure} = \frac{2*Precision*Recall}{Precision+Recall}$$

Table 4.7 show result analysis with SVM. Overall accuracy of the proposed system yielded by SVM is reported as 72%

Table 4.7 Result Analysis with SVM

Target Class	Accuracy	Sensitivity	Specificity	F-measure	Precision
Scrape	0.7500	0.8824	0.4286	0.8333	0.7895
Rust Spot	0.6000	0.7143	0.4545	0.6667	0.6250
Emulsion Spot	0.7200	0.8750	0.4444	0.8000	0.7368
Roller Mark	0.6800	0.7500	0.6154	0.6923	0.6429

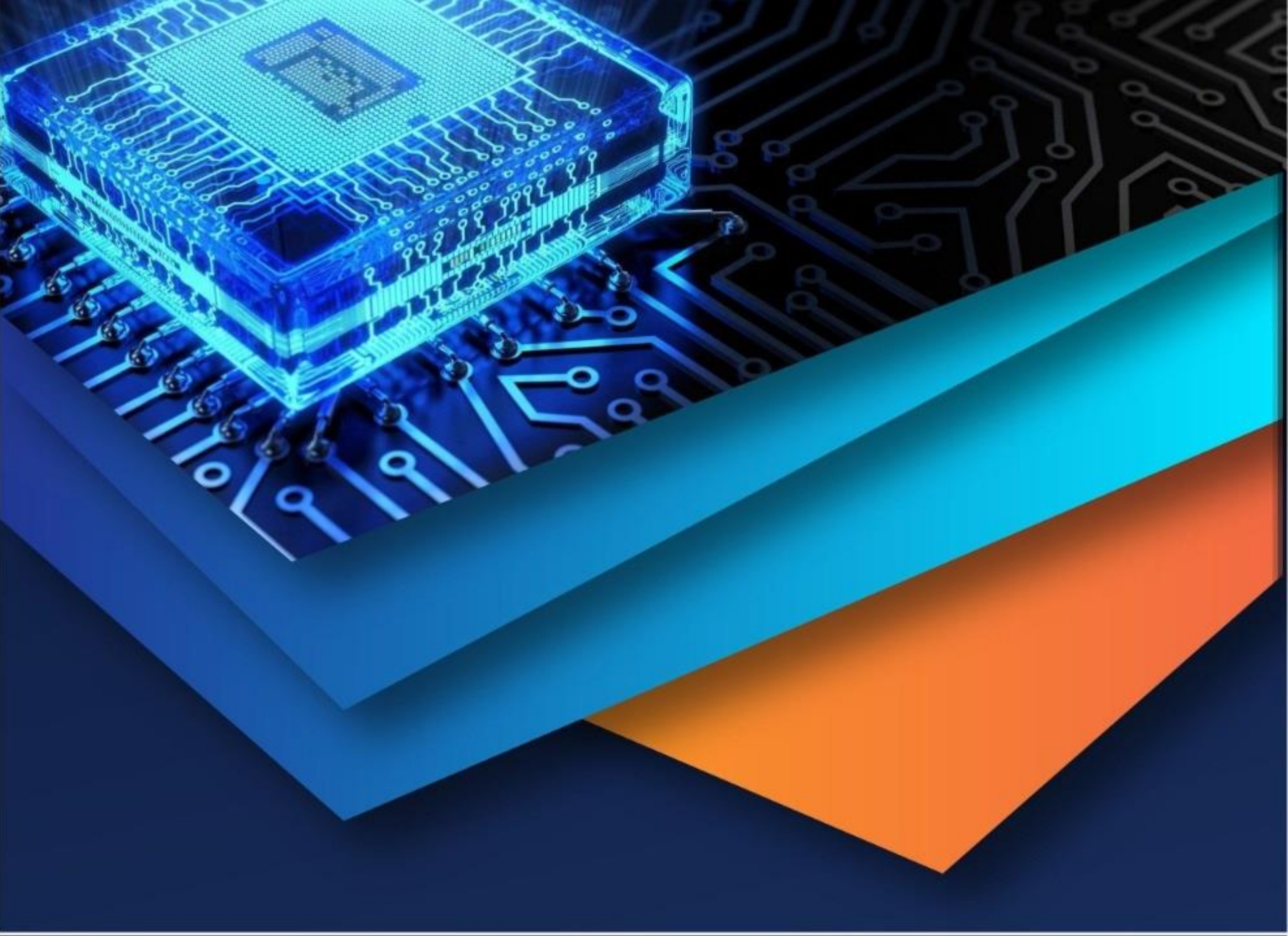
VI. CONCLUSION

The proposed system is to classify four kinds of surface defects in steel surfaces such as Roller Mark, Rust Spot, Emulsion Spot and Scrape. In offline, the input image is pre-processed and the following visual features such as geometric features, shape features and texture features are extracted from the defect image. In order to optimize the extracted visual features and to improve the accuracy of the classification model, genetic algorithm is used to find an optimal solution. For optimization, an initial population is generated randomly and the fitness function is calculated for every chromosome. Then the genetic operations such as selection, crossover and mutation are done to generate new offspring chromosome and evaluated to pick optimal chromosome or optimal feature. Finally, in online, the selected features are given as input to the Optimized SVM classifier, in order to predict the class of defect to which the input image belongs. Accuracy, sensitivity, specificity, f-measure and precision are used as the performance metrics to evaluate the system. The accuracy of the proposed system is reported by SVM as 78%. Thus, the proposed system provides an automatic surface inspection of steel surfaces and it can be employed in steel firms. As a future work, the research will focus on other shapes for more effective defect prediction and to improve accuracy.

REFERENCES

[1] Yun J. P, Jeon Y. J Choi, D C Kim. (2012), 'Real-time defect detection of steel wire rods using wavelet filters optimized by univariate dynamicencoding algorithm for searches'. JOSA A, Vol 29, No 5, pp 797-807.
[2] Alimohamadi H, Ahmadyfard A. (2009), 'Detecting skin defect of fruits using optimal Gabor wavelet filter'. In Digital Image Processing, 2009 International Conference on IEEE, pp 62-64.
[3] Rahman, M. H, Chowdhury S, Bashar M. A. (2013), 'An Automatic Face Detection and Gender Classification from Color Images using Support Vector Machine, Journal of Emerging Trends in Computing andInformation Sciences, Vol 4, No 1, pp 92-101.
[4] Moazenzadeh R, Mohammadi B, Shamshirband S, Kw Chau (2018) Coupling a firefly algorithm with support vector regression to predict evaporation in northern iran. Eng Appl Comput Fluid Mech 12(1):584–597.
[5] Najafi B, Faizollahzadeh Ardabili S, Shamshirband S, Kw Chau, Rabczuk T (2018) Application of ANNs, ANFIS and RSM to estimating and optimizing the parameters that affect the yield and cost of biodiesel production. Eng Appl Comput Fluid Mech 12(1):611–624.
[6] Houssein, E.H.; Gad, A.G.; Wazery, Y.M. Jaya algorithm and applications: A comprehensive review. Metaheuristics Optim. Comput. Electr. Eng. 2021, 696, 3–24.
[7] Ogundokun, R.O.; Awotunde, J.B.; Sadiku, P.; Adeniyi, E.A.; Abiodun, M.; Dauda, O.I. An enhanced intrusion detection system using particle swarm optimization feature extraction technique. Procedia Comput. Sci. 2021, 193, 504–512.
[8] Zitar, R.A.; Al-Betar, M.A.; Awadallah, M.A.; Doush, I.A.; Assaleh, K. An intensive and comprehensive overview of JAYA algorithm, its versions and applications. Arch. Comput. Methods Eng. 2022, 29, 763–792.
[9] Mohammadi, M.; Rashid, T.A.; Karim, S.H.T.; Aldalwie, A.H.M.; Tho, Q.T.; Bidaki, M.; Rahmani, A.M.; Hosseinzadeh, M. A comprehensive survey and taxonomy of the SVM-based intrusion detection systems. J. Netw. Comput. Appl. 2021, 178, 102983.
[10] Chen, M. Y, Yang, Y. H, Ho, C. J, Wang, S. H, Liu, S. M, Chang, E, Ouhyoung, M. (2012), 'Automatic chinese food identification and quantity estimation', In Siggraph Asia, Technical Briefs pp. 29, ACM.
[11] Kaushik, A, Singh G, Bhatia A. (2013), 'SVM Classification in Multiclass Letter Recognition System', Global Journal of Computer Science and Technology, Vol 13, No 9, pp 39-48.

[12] Rahman, M. H, Chowdhury S, Bashar M. A. (2013), 'An Automatic Face Detection and Gender Classification from Color Images using Support Vector Machine, Journal of Emerging Trends in Computing and Information Sciences, Vol 4, No 1, pp 92-101.
[13] Al-Al-Janabi, M.; Ismail, M.A. Improved intrusion detection algorithm based on TLBO and GA algorithms. Int. Arab J. Inf. Technol. 2021, 18, 170–179.
[14] Aljanabi, M.; Ismail, M.A.; Ali, A.H. Intrusion detection systems, issues, challenges, and needs. Int. J. Comput. Intell. Syst. 2021, 14, 560–571.
[15] Kavak, H.; Padilla, J.J.; Vernon-Bido, D.; Diallo, S.Y.; Gore, R.; Shetty, S. Simulation for cyber security: State of the art and future directions. J. Cyber Secur. 2021, 7, tyab005.
[16] Al-Qatf, M.; Lasheng, Y.; Al-Habib, M.; Al-Sabahi, K. Deep learning approach combining sparse autoencoder with SVM for network intrusion detection. IEEE Access 2018, 6, 52843–52856.
[17] A.R.B. Abad, H. Ghorbani, N. Mohamadian, S. Davoodi, M. Mehrad, S.K.ye Aghdam, H.R. Nasriani, Robust hybrid machine learning algorithms for gas flow rates prediction through wellhead chokes in gas condensate fields, Fuel 308 (2022).
[18] Al-Mahasneh, A. J., Anavatti, S. G., & Garratt, M. A. (2018a). Review of Applications of Generalized Regression Neural Networks in Identification and Control of Dynamic Systems.
[19] Azam, A., Bardhan, A., Kaloop, M. R., Samui, P., Alanazi, F., Alzara, M., & Yosri, A. M. (2022). Modeling resilient modulus of subgrade soils using LSSVM optimized with swarm intelligence algorithms. Scientific Reports 2022 12:1, 12(1), 1–20.
[20] A.R. Behesht Abad, S. Mousavi, N. Mohamadian, D.A. Wood, H. Ghorbani, S. Davoodi, …., K. Shahbazi, Hybrid machine learning algorithms to predict condensate viscosity in the near wellbore regions of gas condensate reservoirs, Journal of Natural Gas Science and Engineering 95 (2021).
[21] Q. Fan, T. Fan, A Hybrid Model of Extreme Learning Machine Based on Bat and Cuckoo Search Algorithm for Regression and Multiclass Classification, Journal of Mathematics 2021 (2021).
[22] J.C. Huang, K.M. Ko, M.H. Shu, B.M. Hsu, Application and comparison of several machine learning algorithms and their integration models in regression problems, Neural Computing and Applications 32 (10) (2020) 5461–5469.
[23] I. Izonin, R. Tkachenko, M. Gregus ml, K. Zub, P. Tkachenko, A GRNN-based approach towards prediction from small datasets in medical application, Procedia Computer Science 184 (2021) 242–249.
[24] N. Kardani, A. Bardhan, B. Roy, P. Samui, M. Nazem, D.J. Armaghani, A. Zhou, A novel improved Harris Hawks optimization algorithm coupled with ELM for predicting permeability of tight carbonates, Engineering with Computers 1 (2021) 1–24.
[25] H. Khodakhah, P. Aghelpour, Z. Hamedi, Comparing linear and non-linear data-driven approaches in monthly river flow prediction, based on the models SARIMA, LSSVM, ANFIS, and GMDH, Environmental Science and Pollution Research 29 (15) (2022) 21935–21954.
[26] Kui, L., Ke, Y., Xiaojun, X., Jianjun, Z., Yun, L., & Jinghong, W. (2022). Information Price Intelligent Prediction Model of Major Equipment in Power Grid Projects Based on the SSA-LSSVM Model. 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers, IPEC 2022, 549–553.
[27] W. Lu, R. Du, P. Niu, G. Xing, H. Luo, Y. Deng, L. Shu, Soybean Yield Preharvest Prediction Based on Bean Pods and Leaves Image Recognition Using Deep Learning Neural Network Combined With GRNN, Frontiers in Plant Science 12 (2022) 3160.
[28] L. Mancipe-Castro, R.E. Gutiérrez-Carvajal, Prediction of environment variables in precision agriculture using a sparse model as data fusion strategy, Information Processing in Agriculture 9 (2) (2022) 171–183.
[29] C.S.W. Ng, A. Jahanbani Ghahfarokhi, M. Nait Amar, Application of nature-inspired algorithms and artificial neural network in waterflooding well control optimization, Journal of Petroleum Exploration and Production 11 (7) (2021) 3103–3127.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY