An automatic system to surface defect classification of hot rolled steel

Adel Boudiaf[1*], Abdelkrim Moussaoui [2], Rachid Zaghdoudi[1,2], Zoheir Mentouri[2], Achour Saadoune[3]
[1]*Research Center in Industrial Technologies CRTI P.O.Box 64, cheraga 16014 Algiers, Algeria*
[2] Laboratory of Electrical Engineering of Guelma (LGEG), Université 08 Mai 1945, Guelma, 24000 –Algeria
[3]*Laboratoryof Metallic and Semiconducting Material (LMSM), University of Biskra, 07000 Biskra,Algeria.*
Email[*]: *a.boudiaf@crti.dz*

***Abstract*—The primary objective of this paper is to develop an automatic surface defect inspection system for hot-rolled flat steel The proposed technique consists of four major steps. The proposed technique consists of four major steps. The first step is image acquisition. The second step is features extraction of image by Histogram of oriented gradients (HOG). In the third step the principal component analysis (PCA) is applied on the HOG descriptor to reduce the dimensionality of the feature vector. In the final step the K- nearest neighbor classifier (KNN) is used to classify the different steel surface defects. The experimental results showed that the proposed steel inspection system which is based on KNN classifier provide a better results and recognition accuracy of 91.12%.**

.

Keywords*— *Industrial vision; the* K- nearest neighbor classifier *; classification of surface defects; Histogram of Oriented Gradients (HOG); Principal Component Analysis (PCA).

I. INTRODUCTION

Today, products of rolling represents approximately 90% of all materials produced in the sector of metalworking because these rolling products are widely used in different fields of industry such as food packaging, military, medical, auto industry, aerospace and other fields [1].

Many of these industries (aerospace and automotive) reject the rolling products which has surface defects because a minor defect in these materials might result in Human and material losses at a later stage [1]. Studies [2] show that, 70% of the failures of rotating machinery are due to the surface defects (See Fig.1). Therefore, these surface defects must be detected as early as possible to avoid loss of production and also preserve the safety of personnel.

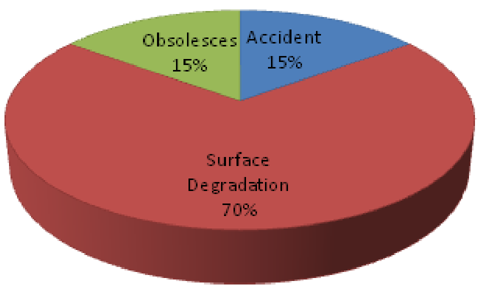

Fig. 1. Failures Sorces of rotating machinery

The detection of these surface defects by human vision remains tedious, fatiguing, bit fast, bit robust, sketchy, dangerous or impossible. These and other disadvantages led many researchers to focus their research on the study of the artificial intelligence (AI) approach to automatic the diagnosis procedures and reduce the errors caused by subjective human judgment ,like neural network [3], the sequential minimal optimization (SMO) algorithm [4], Bayesian classifier [5], fuzzy inference [6], , k-means clustering [7-9]. In this context, the primary objective of this paper is to develop an automatic surface defect inspection system for hot-rolled flat steel based on the principal component analysis (PCA), Histogram of oriented gradients (HOG) and K- nearest neighbor classifier.

The proposed technique consists of four major steps. The proposed technique consists of four major steps. The first step is image acquisition. The second step is features extraction of image by Histogram of oriented gradients (HOG). In the third step the principal component analysis (PCA) is applied on the HOG descriptor to reduce the dimensionality of the feature vector. In the final step the K-nearest neighbor classifier is used to classify the different steel surface defects. The developed flow chart of the surface defect inspection system is shown in Fig. 2

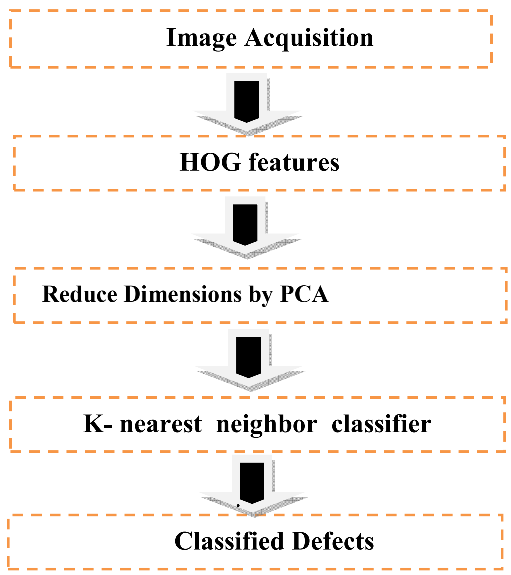

Fig 2. Flowchart of the proposed method

The paper is organized as follows: Section 2 presents Dataset of images used in this work. Section 3 presents the features extraction of defects images by Gray-level co-occurrence matrix (GLCM). Section 4 presents the K-nearest neighbor classifier is used for classification the surface defects.Section.5.Presents the experimental results of our work. Finally, our conclusions are provided in Section 6.

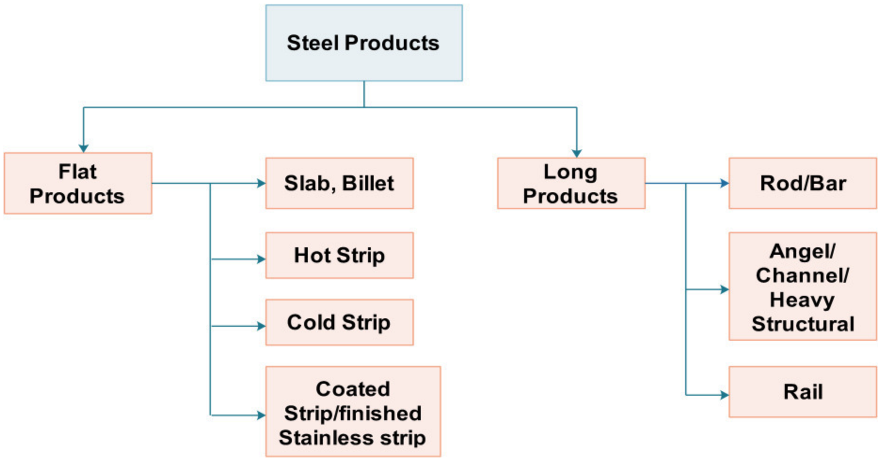

Fig. 3. Categories of steel products

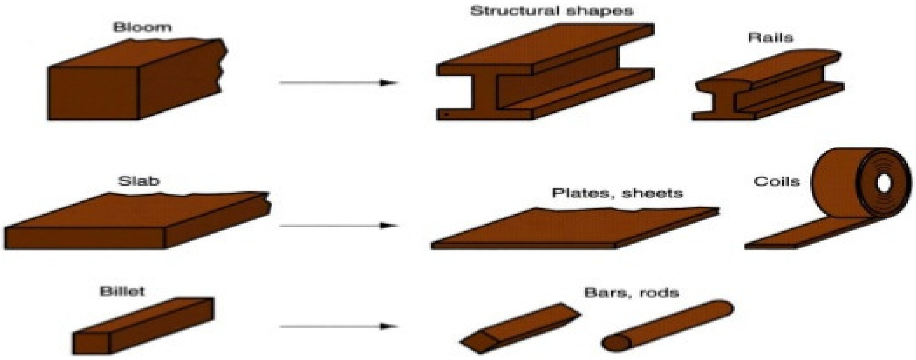

Fig. 4. Some products made from rolling

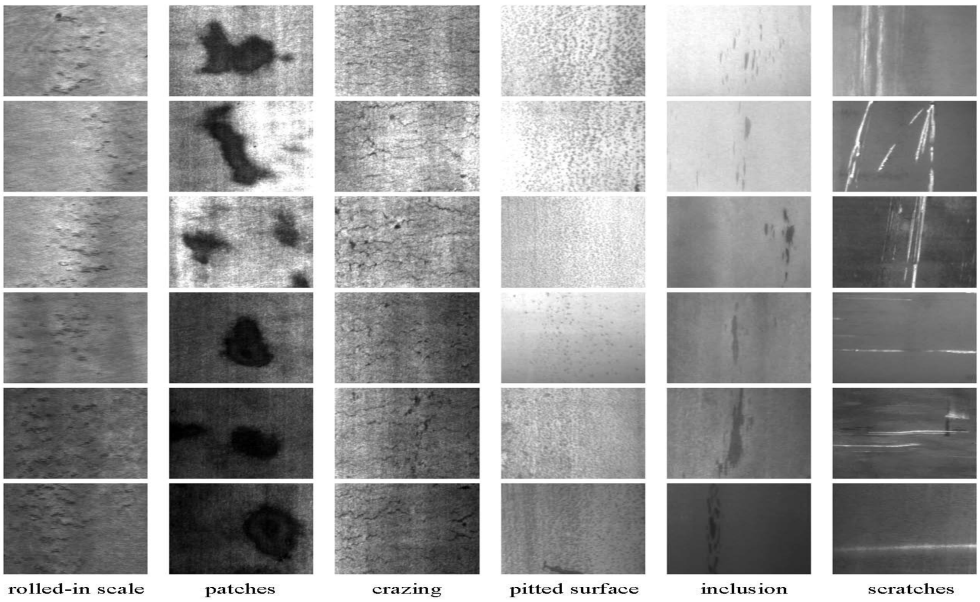

Fig. 5. Sample images of six kinds of typical surface defects

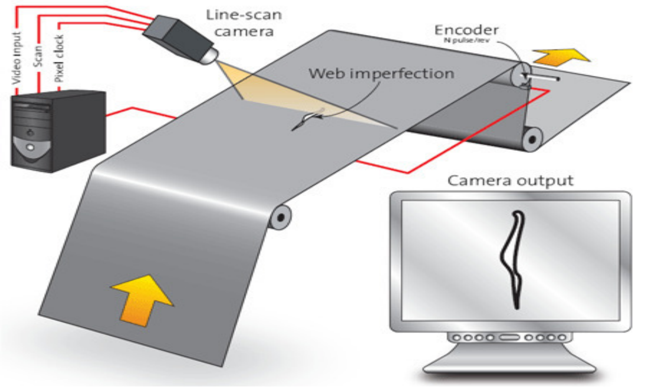

Fig. 6. Surface inspection system of steel strip based on machine vision

II. IMAGE ACQUISITIONS

Dataset used in this work are obtained from the surface defect database of Northeastern University (NEU) [10], six kinds of typical surface defects of the hot-rolled steel strip are collected which are rolled-in scale (RS), patches (Pa), crazing (Cr), pitted surface (PS), inclusion (In) and scratches (Sc) see fig.5. The database includes 1800 grayscale images: 300 samples each of six different kinds of typical surface defects. The original resolution of each image is 200×200 pixels.

III. FEATURES EXTRACTION

In this work we have used the Histogram of Oriented Gradients (HOG) method to extract the feature of image Due to its capabilities

A. Histogram of Oriented Gradients (HOG)

Histogram of Oriented Gradients (HOG) is a gradient based feature descriptor that was first proposed by Dalal and Triggs in the human detection framework [11]. HOG has been widely used in object detection and has shown great success in other various computer vision applications. The basic idea of this descriptor is to represent the appearance and shape of an object in an image by the way in which the intensity of the gradient or the direction edges is distributed. This is by dividing the image into cells and calculating for each cell a histogram of the directions of the gradient for the pixels within this cell. The concatenation of these histograms forms the Hog descriptor. In addition, a step of normalization is necessary to avoid the changes of illuminations and the shadows effects. The following steps outline the computational procedure of the HOG descriptor [12].

Step1: Calculates the horizontal and vertical gradient of the image by convolving the image with the respective gradient masks ([−1, 0, 1] and [−1, 0, 1]^T).

Step2: Uses (1) and (2) to compute the strength and orientation of the gradient.

$$SG = \sqrt{G_h(x, y)^2 + G_v(x, y)^2} \quad (1)$$

$$OG = \arctan \frac{G_h(x, y)}{G_v(x, y)} \quad (2)$$

Where: G_h and G_v denote the horizontal and vertical gradient; SG and OG represent the strength and orientation respectively at point (x, y) in the image.

Step3: Divides the image into N*N cells and computes the histogram of orientations for each cell. If the histogram is divided into *k* bins based on the orientation, the value of the i^{th} bin V_i for cell *C* is computed using (3).

$$V_i = \sum_{(x,y)\in C} SG(x, y)/OG(x, y) \in bin_i \quad (3)$$

Step4: The L2 Norm is used to normalize the histogram of each cell.

Step5: Forms the HOG descriptor by concatenating the histograms of all cells.

The computation of HOG features requires setting of two important parameters, the number of cell (N) and the number of orientation bins (K), which produces a descriptor of dimension N*N*K.

B. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) technique is one of the most excellent unsupervised dimensionality reduction techniques. The main goal of this technique is to transform the data or features from a higher dimensional space to a lower dimensional space [13-14].

In our work, principal component analysis (PCA) is applied on the HOG descriptor to reduce the dimensionality of the feature vector.

Algorithm: PCA

Begin
1: Given a data matrix ($X = [x_1, x_2, x_3 ... x_N]$), where N represents the total number of samples and x_i represents the i^{th} sample.
2: Compute the mean of all samples as follows:

3: Subtract the mean from all samples as follows:

4: Compute the covariance matrix as follows:

5: Compute the eigenvectors V and eigenvalues of the covariance matrix ;
6: Sort eigenvectors according to their corresponding eigenvalues ;
7: Select the eigenvectors that have the largest eigenvalues ;
9: All samples are projected on the lower dimensional space of PCA (W) as follows ;
Y = WT D.
End

For the classification the surface defects in products of hot rolling we have used in this work the K- nearest neighbor classifier.[15-16].

IV. K-NEAREST NEIGHBOR (K-NN) CLASSIFIER

The K- nearest neighbor classifier is a simple supervised classifier that has yield good performance . This classifier computes the distance from the unlabeled data to every training data and selects the K neighbor with shortest distance . No requirement for training process makes this classifiers implementation simple [15-16].

In Fig 6, the data instance marked by red coloured star is classified to Class B if k=3 and the same data instance is classified to Class A if k=6.

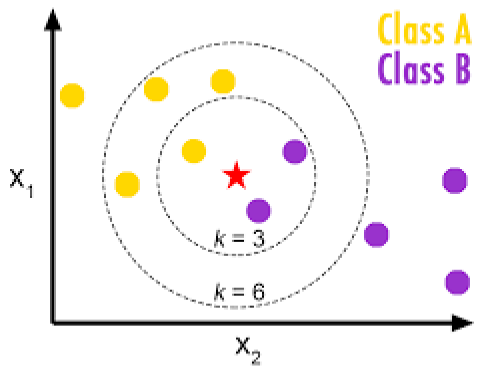

Fig. 6. Depicts the classification of KNN algorithm for k=3 and k=6

V. EXPERIMENTAL RESULT

Image dataset consist of 1800 images including six kinds of typical surface defects of the hot-rolled steel strip which are rolled-in scale (RS), patches (Pa), crazing (Cr), pitted surface (PS), inclusion (In) and scratches (Sc).

This 1800 grayscale images includes 300 samples each of six different classes of typical surface defects.

We are use 1200 images (200 samples each of six different classes of typical surface defects.) are used for training process, and 600 images(100 samples each of six different classes of typical surface defects) for test process .

The computation of HOG features requires setting of two important parameters, the number of cell (N) and the number of orientation bins (K), which produces a descriptor of dimension N*N*K. For experiments, we consider a number of configurations for the parameters (N, K). In addition, to reduce the dimensionality of the HOG descriptor, principal component analysis (PCA) is also applied. The number of principal components is determined empirically after many tests

Table I summarizes the recognition accuracy of the SVM classifier by computing the HOG features for different values of N, fixing the value of K to 4.

The impact of the number of orientation bins in the HOG features on the overall recognition rates is studied by fixing the cell size N to 6 and varying the number of bins K. The results of these evaluations are presented in Table II.

Table I and Table II summarize the recognition rates realized by the KNN classifier for different configuration of HOG descriptor. The highest recognition rates achieved stand at 91.12% for a number of cell N=6 and a number of orientation Bins k=5. It is worth noting that the use of PCA reduces the number of features less than half without lowering the performance.

TABLE I. RECOGNITION RATES AS A FUNCTION OF NUMBER OF CELLS (N).

Hog features (N*N*K)	HOG-PCA	KNN
3*3*4	36	61.53%
4*4*4	64	72.90%
5*5*4	60	77.00%
6*6*4	**80**	**85.30%**
7*7*4	100	84.66%
8*8*4	130	82.70%
9*9*4	160	76.33%
10*10*4	220	72.83%
11*11*4	240	70.16%
12*12*4	280	65.48%

TABLE II . RECOGNITION RATES AS A FUNCTION OF NUMBER OF ORIENTATION BINS (K).

Hog features (N*N*K)	HOG-PCA	KNN
6*6*3	70	80.66%
6*6*4	80	87.25%
6*6*5	**100**	**91.12%**
6*6*6	120	87.50%
6*6*7 6*6*8	130 140	85.66% 84.33%
6*6*9	150	82.16%
6*6*10	170	81.66%
6*6*12	200	79.16%
6*6*14	250	77.36%
6*6*16	270	74.00%

VI. CONCLUSION

In this research paper, we tried to classify the six kinds of typical surface defects of the hot-rolled steel strip are collected which are rolled-in scale (RS), patches (Pa), crazing (Cr), pitted surface (PS), inclusion (In) and scratches (Sc) . principal component analysis (PCA) is applied on the HOG descriptor to reduce the dimensionality of the feature vectorthes new vector (HOG+PCA) used to train the KNN classifier .The experimental results showed that our method is effective and efficient in classifying the surface defects .Moreover, the proposed steel inspection system which is based on KNN classifier provide a better results and recognition accuracy of 91.12%.

REFERENCES

[1] S. Kalpakjian and S. R. Schmid,," Manufacturing engineering and Technology ", 4th ed. Dorling Kindersley Pvt Ltd, , pp. 340-341. 2011.

[2] T. aarthi, M. karthi, M. abinesh "Detection and analysis of surface defects in metals using wavelet transform", International Journal of Scientific and Research Publications, Volume 3, Issue 6, June 2013

[3] G. Gatica, "Olive Fruits Recognition Using Neural Networks", Procedia Computer Science., Vol.17, no. 2, pp. 412–419, 2013

[4] V. Duraisamy , R. Durai , " Sequential Minimal Optimization Approach for Content Based Image Retrieval", International Journal of Information Sciences and Application , Vol.3, no. 2, pp. 93–97, 2011

[5] F Pernkopf, "Detection of surface defects on raw steel blocks using Bayesian network classifiers". Pattern Anal. Appl. 7, 333–342 (2004)

[6] A.Borselli, ""A fuzzy inference system applied to defect detection in flat steel production"". proceedings of the 2010, International Conference on Fuzzy Systems, , no.14, pp. 1–6, 2010.

[7] V. Pham · B. Ryong Lee, “An image segmentation approach for fruit defect detection using k-means clustering and graph-based algorithm”, Vietnam Journal of Computer Science, Vol.2, no.06, pp. 25-33, 2015.

[8] H. SIDQI · J. KAKBRA, “Image Classification Using K-mean Algorithm”, International Journal of Emerging Trends & Technology in Computer Science, Vol.3, no.01, pp. 38-41, 2014.

[9] N. Neogi, D.K. Mohanta, and P.K. Dutta, “Review of vision-based steel surface inspection systems”, Eurasip journal of image and video processing, 2014: 50. https://doi.org/10.1186/1687-5281-2014-50.

[10] Song. K., Yan Y., “A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects,” Appl. Surface Sci. Vol :285:,pp : 858–864, 2013.

[11] N. Dalal, B. Triggs, “Histograms of Oriented Gradients for Human Detection,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 886–893, 2005.

[12] S. Chen, C. Liu, “Precise Eye Detection Using Discriminating HOG Features,” Computer Analysis of Images and Patterns, Lecture Notes in Computer Science vol. 6854, pp 443-450, 2011.

[13] . B.-K. Bao, G. Liu.., C. Xu.., S. Yan., "Inductive robust principal component analysis," IEEE Transactions on Image Processing, vol. 21, no. 8, pp. 3794-3800, 2012.

[14] R. D., Santo. "Principal component analysis applied to digital image compression," Einstein, vol. 10, no. 2, pp. 135-139, 2012.

[15] J. S. Raikwal, C. Saxena, “Performance Evaluation of SVM and K-Nearest Neighbor Algorithm over Medical Data set”, International Journal of Computer Applications, vol. 50, no. 14, pp. 34–39, 2012.

[16] R. Rithesh.,“SVM-KNN: a novel approach to classification based on svm and knn,” *International Research Journal of Computer Science.*, vol. 4, no. 8, pp. 43–49, 2017.