Bullet Image Classification using Support Vector Machine (SVM)

Dwi Ratna S¹, Budi Setyono², Tyara Herdha³
¹²³Department of Mathematics, InstitutTeknologiSepuluh November Campus ITS, Sukolilo, Surabaya, 60111, Indonesia
E-mail: dwiratna@matematika.its.ac.id

Abstract. All Quality control of the bullet is usually done manually. Manual inspection process has some disadvantages for instance human error factor and the time inspection are relatively longer. In this study, defect detection and classification of bullets by using Support Vector Machine (SVM)has been done. The stages of classification of scheme bullets in this study include pre-processing, feature extraction with Principal Component Analysis (PCA) and classification of bullets by using SVM normalized. Performance of the proposed classification scheme was tested using 80 images of bullets in which 40 images free of defects and 40 images with defect. The results show that the scheme can classify images of bullets with a 90% accuracy rate with test data in the form of 10 images free-defect and 10 image defects.

1. Introduction

Sector of industry defense in Indonesia is increasingly developed, especially in the field of weaponry. One of the efforts in order to remain credible in producing weapons is to maintain and improve the quality and quantity of the product. For example, for product bullets, even with the slightest defect can be really dangerous when it is used. Therefore, the bullets quality inspection is very important to be done in the production of bullets before commercialized.

Quality inspection of ammunition are usually done by manual selection and this process certainly has its disadvantages including the existence human error and the inspection time is relatively long. Whereas at one time production could produce thousands bullets. Accordingly, the technology for automatically detecting defects bullets needed. Digital image processing can be used as a solution, because it has been used to detect defects on the fabric, ceramic and metal surfaces.

There has been some research on defect inspection on the bullet based on image processing. Identification of defects in the bullet with Learning Vector Quantization (LVQ) has been conducted [1], the results show that LVQ has identified defect bullet with an accuracy of 78%. Pusparini et al [2] have studied the defect detection bullets using line detection and fuzzy set with an accuracy of 70.58%.

Support Vector Machine (SVM) method is one of popular methods that as machine learning this past few years, because the method is performed analytically and also provide good generalization capabilities in its application. That's why it pushed many studies using this method, such as research on gender classification is based on the facial image [3] or research on the classification of the type of music based on audio features [4].
In this study, classification of bullets image using Support Vector Machine (SVM) has been performed. This classification will distinguish between image with free of defect and image with defect.

2. Methods

2.1. Support Vector Machine for Classification Bullets Image

Figure 1 shows the scheme of the classification process bullets image by using Support Vector Machine. Classification process consists of several stages: image reading stage, pre-processing stage, feature extraction stage and classification stage.

![Figure 1. Scheme of the classification system](image)

The input image is a color image of the bullets. At the stage of pre-processing the input image is processed by image processing operations that is segmentation, cropping, resize, Gray Scaling, and Conversion. Results of pre-processing is input to the feature extraction stage with Principal Component Analysis (PCA). Classification based on disability bullets image performed using SVM
2.2. Feature Extraction with Principal Component Analysis (PCA)

PCA has been successfully used in face recognition and object recognition. The standard algorithm of PCA [5]:

1. A set of $M$ image, each of them has a size $r \times c$. Each of this $M$ image will be represented as a column vector $Z_n$ with length $rc$. The average of object is defined by the following equation:

$$
\mu = \frac{1}{M} \sum_{n=1}^{M} z_n
$$

where:
- $\mu$ = The average object of image
- $M$ = Number of images
- $z_n$ = Column vector for each image

2. Calculate covariance matrix $C$, that has the following expansion

$$
C = \frac{1}{M} \sum_{n=1}^{M} (z_n - \mu)(z_n - \mu)^T
$$

where:
- $C$ = Covariance matrix
- $\mu$ = The average object of image
- $M$ = Number of images
- $z_n$ = Column vector for each image

3. Find the eigenvalues and eigenvectors. Then sort the eigenvectors in the matrix from the largest to smallest based on its eigenvalues. So the first column vector of the matrix will have eigenvalues greater than the eigenvalues from the next column vector.

4. The next step is to perform matrix reduction. Eigenvectors that already sorted according to its eigenvalues, the matrix will be reduced.

5. Create a feature vector from eigenvectors that has been sorted by this equation:

$$
F_n = E^T(z_n - \mu)
$$

where:
- $F_n$ = Feature vector
- $E^T$ = Transpose of eigenvectors
- $z_n$ = Column vector for each image
- $\mu$ = The average object of image

The weight of the projection $F_n (F_n \in \mathbb{R}^{M'-1})$ refer to the standard feature vector of each bulletsand $M'$ is called the feature length.

6. The last step is create feature matrix from feature vector which had been obtained previously.

2.3. Classification with Support Vector Machine (SVM)

Support Vector Machine (SVM) is a linear classification method with a feature set that has been predetermined. The SVM method will look for a linear hyperplane with the largest margin to separate the existing class. The largest margin, which is then referred to as Maximum Marginal hyperplane (MMH) will provide the greatest distance between classes. The distance between the hyperplane with a side of the margin is equal to the distance between the hyperplane with margin on the other side[6].
Before performing classification, data have to be normalized. Because SVM work more effectively with the data that have a small scale range. In this study, the method chosen to normalize the data is the min-max method which has the formula

\[ v'_i = \frac{v_i - \text{min}_A}{\text{max}_A - \text{min}_A} (\text{new}_\text{max}_A - \text{new}_\text{min}_A) + \text{new}_\text{min}_A \]  

(4)

where:
- \( v'_i \) = new data after normalization
- \( v_i \) = data input
- \( \text{min}_A \) = the minimum value from data input
- \( \text{max}_A \) = the maximum value from data input
- \( \text{new}_\text{min}_A \) = the minimum value from the desired outcome normalization data
- \( \text{new}_\text{max}_A \) = the maximum value from the desired outcome normalization data

In our case min-max normalization maps a value \( v_i \) to \( v'_i \) in the range \([0,1]\) so put \( \text{new}_\text{min}_A = 0 \) and \( \text{new}_\text{max}_A = 1 \) in the above equation (4), we get the simplified formula of min-max normalization

\[ v'_i = \frac{v_i - \text{min}_A}{\text{max}_A - \text{min}_A} \]  

(5)

In the real case, the data usually is non linearly separable data. So it is necessary to change the usual SVM method. To solve this problem, the SVM used "kernel trick". There are many types of kernel used in the process of SVM. In this study, there are three kernel applied that is a polynomial kernel, gaussian / RBF kernel and linear kernel.

Polynomial kernel has the form of the equation below:

\[ K(x, x_i) = ((x^T \cdot x_i) + 1)^n \]  

(6)

RBF kernel has the form of the equation below:

\[ K(x_i, x_j) = \exp\left(\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \]  

(7)

Linear kernel has the form of the equation below:

\[ K(x_i, x_j) = x^T \cdot x_i \]  

(8)

where:
- \( K \) = Kernel
- \( x_i \) = Data input to be trained
- \( x_i^T \) = Data input that has been transposed
- \( \sigma \) = Free parameter
The algorithm for obtaining the hyperplane classifier is [7]:

1. The data notated as $x \in \mathbb{R}$, while each label is denoted $y \in \{-1, +1\}$ for $i = 1, 2, \ldots, \ell$, where $\ell$ is the number of data. Assumed two class -1 and +1 can be separated completely by hyperplane dimension $d$, which is defined:

$$\vec{w}, x + b = 0 \quad (9)$$

A pattern $x$, which includes class -1 (negative sample) can be formulated as a pattern that satisfies equation:

$$\vec{w}, x + b = -1 \quad (10)$$

while pattern which includes class +1 (positive sample):

$$\vec{w}, x + b = +1 \quad (11)$$

2. The greatest margin can be found by maximizing the value of the distance between the hyperplane and the closest point $1/||w||$. It can be formulated as a Quadratic Programming (QP) problem. It finds a minimal point by solving the following optimization problem:

$$\min_{\vec{w}} \tau(w) = \frac{1}{2}||w||^2 \quad (12)$$

with the constraints:

$$y_i(x_i, w + b) - 1 \geq 0$$

where $x_i$ is the input data, $y_i$ is the output of the data $x_i$, $\vec{w}$, $b$ is the parameters that searched value.

3. The optimization problem above can be solved by some computation technique, one of computation technique that can be used is Lagrange Multipliers as shown in the following equation:

$$L(w, b, \alpha) = \frac{1}{2}||w||^2 - \sum_{i=1}^{\ell} \alpha_i \left( y_i \left( (x_i, w + b) - 1 \right) \right) \quad (13)$$

$\alpha_i$ are Lagrange Multipliers, which is zero or positive ($\alpha_i \geq 0$).

4. The optimal value of the above equation can be calculated by minimizing $L$ to $w$ and $b$, and maximize $L$ to $\alpha_i$. With notice the characteristic that the optimum gradient point $L = 0$. The equation can be modified as maximization problem which only contains only $\alpha_i$, as the following equation:

$$\max \sum_{i=1}^{\ell} \alpha_i \frac{1}{2} \sum_{i,j=1}^{\ell} \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (14)$$

with the constraints:

$$\alpha_i \geq 0 \quad (i = 1, 2, \ldots, \ell)$$

$$\sum_{i=1}^{\ell} \alpha_i y_i = 0$$

Dot product $x_i^T x_j$ often replaced by the symbol $K$, where $K$ is the matrix kernel. From the results of this calculation are mostly obtained $\alpha_i$ is positive. Data were correlated with $\alpha_i$ positive is called a support vector.

3. Experimental Result

In the experiments, we used the images of a bullet which consists of 40 images are free of defects and 40 images are with defects with image size is 1728x2592. The performances of the proposed
classification scheme are evaluated for three data set based on the number of training data. Table 1 shows the detailed information about the three data sets.

Table 1. The number of training Data

| Data Set  | Free-defect | With defect | Total Image |
|-----------|-------------|-------------|-------------|
| Data 1    | 10          | 10          | 20          |
| Data 2    | 20          | 20          | 40          |
| Data 3    | 30          | 30          | 60          |

Figure 3 shows an example of image reading of a bullet that is free of defects and with defects. The aim of the image reading stage is to prepare the image data that serves as the initial data of the proposed scheme.

The Otsu thresholding algorithm used in image segmentation step which aims to separate the object and background such as presented in Figure 4. Area background is given white colour to distinguish the object.

Table 2. Result of pre-processing step

|                   | Free Defect    |        |        | With Defect    |        |        |
|-------------------|----------------|--------|--------|----------------|--------|--------|
| Cropping          |                |        |        | Cropping       |        |        |
| Resize            |                |        |        | Resize        |        |        |
| Grayscaling       |                |        |        | Grayscaling    |        |        |

The results of pre-processing step for a bullet image shown in Table 2. Cropping process aims to minimize the image area that is by cutting area of the image containing only the bullet without a
The purpose of the resize process is to change the image size becomes smaller so that the computing time is reduced, in this experiment the size of the image was changed became 130x26. The grayscaling process is useful to convert color images become gray scale images. In the conversion process, the dimension of the gray scale image is transformed into a one-dimensional vector. This process will produce a column vector with size 3380x1 pixels that will be the input vector for classification with SVM.

The classification process using SVM aims to classify image bullets that were tested include defect-free images or images with defect. Tests performed on a different feature vector length, which 50, 100, 150 and 200 features. The test was also conducted on 3 different kernel: the kernel polynomial, gaussian / RBF kernel and linear kernel. However, the most satisfactory results given by linear kernel. Table 3 shows the accuracy provided by each kernels that have been tested. And the computing time of linear kernel presented in Table 4.

### Table 3. Accuracy of Each Kernel

| Data | Feature Vector Length | Poly  | Rbf  | Linear |
|------|-----------------------|-------|------|--------|
| 1    | 50                    | 58,33%| 58,33%| 55%    |
|      | 100                   | 55%   | 58,33%| 55%    |
|      | 150                   | 51,67%| 55%   | 55%    |
|      | 200                   | 51,67%| 53,33%| 55%    |
| 2    | 50                    | 62,5% | 52,5%| 67,5%  |
|      | 100                   | 65%   | 50%  | 67,5%  |
|      | 150                   | 72,5% | 75%  | 67,5%  |
|      | 200                   | 55%   | 70%  | 67,5%  |
| 3    | 50                    | 70%   | 55%  | 90%    |
|      | 100                   | 75%   | 50%  | 90%    |
|      | 150                   | 70%   | 55%  | 90%    |
|      | 200                   | 65%   | 80%  | 90%    |

### Table 4. Computing Time of Linear Kernel

| Data | Feature Length | Accuracy | Computing Time(second) |
|------|----------------|----------|------------------------|
|      |                |          | Training               | Testing                      |
|      |                |          |                        |                            |
| 1    | 50             | 55%      | 7,63 second            | 0,030 second                |
|      | 100            | 55%      | 7,68 second            | 0,031 second                |
|      | 150            | 55%      | 7,93 second            | 0,034 second                |
|      | 200            | 55%      | 7,91 second            | 0,032 second                |
| 2    | 50             | 67,5%    | 7,70 second            | 0,031 second                |
|      | 100            | 67,5%    | 7,69 second            | 0,033 second                |
|      | 150            | 67,5%    | 7,68 second            | 0,043 second                |
|      | 200            | 67,5%    | 7,74 second            | 0,040 second                |
| 3    | 50             | 90%      | 8,09 second            | 0,030 second                |
|      | 100            | 90%      | 7,40 second            | 0,035 second                |
|      | 150            | 90%      | 7,92 second            | 0,042 second                |
|      | 200            | 90%      | 8,04 second            | 0,041 second                |
4. Conclusion

In this study, support vector machine has been successfully applied to bullets image classification. Classification is done by several stages: image reading, pre-processing, feature extraction with principal component analysis and classification with support vector machine.

The results show that the system can recognize the image of a bullet with an accuracy rate of 90% at 10 free defect images and 10 defect images of testing data with 200 feature length and linear kernel parameter with the best computing time for training is 7.4 seconds. So the best result given by Linear kernel with accuracy 90% at the third training data.

References

[1] Setiyono, B, et (2014), “Pendeteksian Primer Peluru Cacat Menggunakan Learning Vector Quantization (LVQ)”, Prosiding Semnas Mat-PMat USD 2014, pp. 108 – 655.

[2] Pusparini, et al (2015), “Bullets Defect Detection Based On Digital Image Processing Using Line Detection And Fuzzy Sets”, Global Journal of Pure and Applied Mathematics.Volume 11, Number 4, pp. 2215-2222.

[3] Sinha, Chandrakamal. (2013). "Gender Classification from Facial Images using PCA and SVM". Thesis. Department of Biotechnology and Medical Engineering of National Institute of Technology Rourkela

[4] Augustianto, Ramadhika. (2009). “Music Genre Classification Based Audio Features Using Support Vector Machine”. Final Project, Faculty of Information Technology, Institut Teknologi Sepuluh Nopember, Surabaya.

[5] Lu, G., Zhang, D., Wang, K. (2003). “Palmprint recognition using eigenpalms features”. Pattern Recognition Letters 24, pp. 1463-1467.

[6] Abe, Shingo dan Takuya Inoue. 2002. “Fuzzy Support Vector Machine for Multiclass Problems”. European Symposium on Artificial Neural Networks, Bruges, Belgium.

[7] Santosa, Budi. (2007). “Data Mining, Teknik Pemanfaatan Data untuk Keperluan Bisnis”. Yogyakarta. Graha Ilmu.