Detection of Contraband in Milk Powder Cans by Using Stacked Auto-Encoders Combination with Support Vector Machine

Yuping Zhu[1], Lei Wang[2,*] and Wei Zhang[2]

[1]Library, Zaozhuang University, Shandong, China
[2]Information Science and Engineer Department, Zaozhuang University, Shandong, China

*Corresponding author e-mail: look_00@163.com

Abstract. The carrying of contraband has brought increasingly serious harm to people's lives. At present, detection devices used in important places such as customs, airports and railway stations can not automatically identify contraband, and the final identification is entirely done by hand. Therefore, all countries have devoted a great deal of manpower and material resources to studying and developing more effective contraband detection technologies. In this article, we propose a model based on the stacked auto-encoders (SAE) method to detect contraband in milk powder cans. Firstly, we construct a representative of the majority of the reality of the milk CT image data set, secondly, we use the SAE method to extract the features, and finally use the support vector machine (SVM) classifier to determine whether the contracted product is carried in the milk powder cans. In order to prevent the data from over fitting, in the experiment we used the 5-fold cross-validation method. In addition, we also use the grid method to adjust the parameters of SVM. The excellent experimental results show that the model we proposed has a good effect on the detection of carrying contraband in milk powder cans.

1. Introduction

In recent years, drug smuggling has always been the focus of the customs, aviation and transportation safety precautions. Due to the increasingly sophisticated means of criminals, the detection of contraband by the traditional means is becoming more and more difficult. How to use computer vision to improve the identification accuracy of contraband and detect the hidden contraband quickly and accurately in the crowd and cargo group of high-speed flow has become an urgent problem that people need to solve [1]. Contraband is usually sealed inside the container or package, so devices that detect contraband must have the ability to identify the characteristics of objects inside the container [2]. The X-ray detector has been proved to have the ability to reveal the characteristics of matter molecules or atoms, and this characteristic is very suitable for the detection of contraband. It has been developing rapidly in recent years, and has become the first line of defense for safety defense.

The X-ray detector can emit X-ray beams that will absorb some of the energy as they pass through the object being detected [3]. Since different kinds of materials have different absorptive capacity to X-rays, the energy of X-ray beams traveling through different materials reaches the detector differently [4]. After processing, the detector can give the gray scale image of the detected object, and then through the computer vision technology to determine whether the baggage carrying drugs [5]. Wang et al. proposed an illegitimate target classification X-ray detection system based on computer vision, which uses computer vision to automatically detect the X-ray images of baggage so as to realize the classification of illegal targets. The system combines Taruma features of Contourlet transform and histogram, and uses random forest classifier to detect illegal objects from baggage [6]. Gao et al. proposed a beam hardening correction method for liquid safety testing. The method uses a Full-liquid linearization and introduces look-up table techniques to achieve rapid liquid detection [7].

In this article, we automatically detect contraband in milk powder based on the deep learning algorithm. Firstly, we convert the computed tomography (CT) images of the collected milk powder cans into digital matrices that can be processed by depth learning algorithms; secondly, we extract the features of the image using a deep learning stacked auto encoder; finally, we classify them using the state-of-the-art support vector machines (SVM) classifier to detect contraband hidden in milk powder. In the experiment, we used the five-fold cross-validation method to detect contraband in milk powder. In addition, in order to improve the accuracy of SVM, we have also adjusted its parameters. The excellent experimental results show that the method proposed in this paper can effectively improve the detection accuracy of contraband in milk powder.

2. Materials and methods

2.1. Data Set

In real life, criminals usually carry contraband in a sealed package. These contrabands may be stored in milk cans, possibly outside the milk cans, and have a wide variety of contrabands, different shapes and densities, so no regularity can be found during safety testing. This entrapped contraband has great randomness, which increases the difficulty of automatically detecting the computer. In order to simulate the actual situation as much as possible to form such a small but random contraband example, we created a data set using CT images. In this data set, contraband shapes, sizes and locations are randomized and most of them are vague and illegible. In addition, we scan the milk powder can from different directions and angles to form different images by CT scanning. The data set takes into account different situations in real life, and these conditions are harsh enough to represent the vast majority of real-world cases. Figure 1 shows images that do not carry contraband and carry contraband of different situations in the data set.

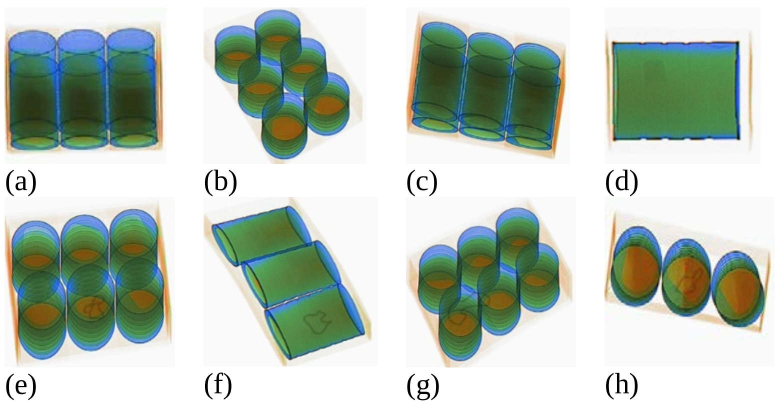

Figure 1. CT Images of Milk Powder Cans in Data Set. (a)- (d) for normal milk powder cans, (e)- (h) for carrying contraband milk powder cans.

2.2. Stacked Auto-Encoders

The deep learning stacked auto-encoders (SAE) use automatic encoders to create deep networks [8]. The auto-encoder (AE) attempts to output the same features as input, which has an input layer, a hidden layer, and an output layer, and its structure is shown in Figure 2. Given a training set X, The encoder converts the input into a representation of the hidden layer by mapping g_c.

$$Y = g_c(X) = F_c(W_1^T X + b_1) \tag{1}$$

where W_1^T is the weight, b_1 is the bias, $F_c(\cdot)$ is the activation function. The decoder then converts the hidden layer representation Y to output by mapping F_d.

$$Z = g_d(Y) = F_d(W_2^T Y + b_2) \tag{2}$$

where W_2^T is the weight, b_2 is the bias, $F_d(\cdot)$ is the activation function. Finally, the loss function $L(X, Z)$ is minimized by the back-propagation program.

$$L(X, Z) = L_r(X, Z) + 0.5\,\partial(\|W_1\|_2^2 + \|W_2\|_2^2) \tag{3}$$

where ∂ is the weight decay cost and $L_r(X, Z)$ is the reconstruction error.

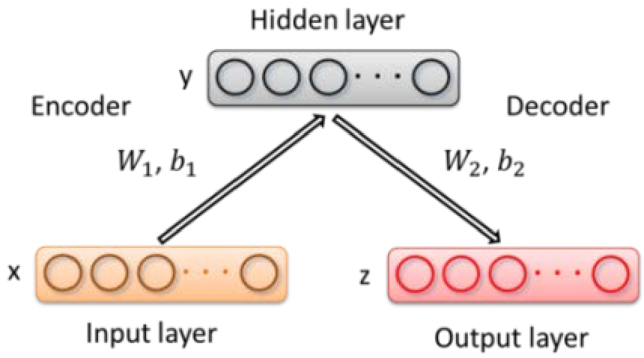

Figure 2. Structure of Auto-Encoder.

Stacked multiple Auto-Encoder can be formed Stacked Auto-Encoders, in which the output of the lower level as the input of the upper layer, extract the abstraction of the original data layer by layer until the highest level [9]. Stacked Auto-Encoders structure shown in Figure 3.

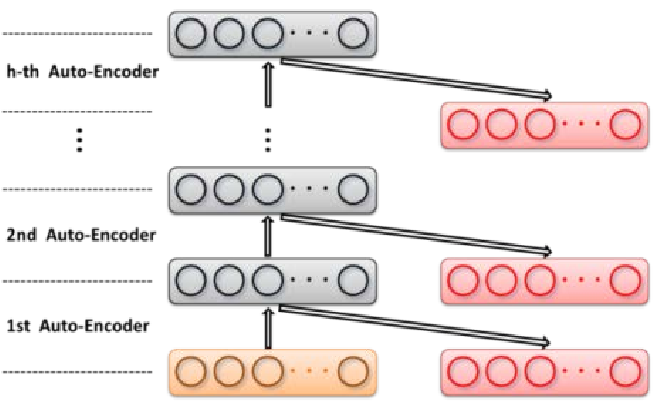

Figure 3. Structure of Stacked Auto-Encoders.

2.3. Support Vector Machines

Support vector machine (SVM) is a machine learning method based on the theory of VC dimension of statistical learning theory and the principle of structural risk minimization [10]. It exhibits many unique advantages in solving small sample, non-linear and high-dimensional pattern recognition problems and largely overcomes the problems of 'dimensionality disaster' and 'over-study' [11].

The mechanism of SVM is to find an optimal classification hyperplane that satisfies the classification requirements so that the hyperplane can maximize the classification accuracy while maximizing the free space on both sides of the hyperplane [12]. In theory, support vector machines can achieve the optimal classification of linear separable data.

The basic idea of support vector machines is: Firstly, the training dataset is nonlinearly mapped to a high dimensional feature space. The purpose of this nonlinear mapping is to transform linearly indivisible datasets in input space into high dimensional feature space into linearly separable datasets. Then establish a maximum isolation distance of optimal separating hyperplane in the feature space, which is equivalent to the input space to generate an optimal nonlinear decision boundary.

3. Results and Discussion

3.1. Evaluation Criteria

The unified evaluation criteria can better measure the advantages and disadvantages of the model. Therefore, in this paper, we choose the commonly used evaluation criteria to evaluate our model. These criteria include: accuracy (Accu.), sensitivity (Sen.), precision (Prec.), and Matthews correlation coefficient (MCC). Their formulas are as follows:

$$\mathrm{Accu.} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}} \tag{4}$$

$$\mathrm{Sen.} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{5}$$

$$\mathrm{Prec.} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{6}$$

$$\mathrm{MCC} = \frac{\mathrm{TP}\times\mathrm{TN}-\mathrm{FP}\times\mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}} \tag{7}$$

where TP (true positive) indicates the number of images without the contraband properly detected, TN (true negative) indicates the number of images with the contraband correctly detected, FP (false positive) indicates the number of images without the contraband incorrectly detected, TN (true negative) indicates the number of images with the contraband incorrectly detected. The Receiver Operating Characteristic (ROC) curve is used to graphically demonstrate the performance of our model, with the true positive rate (sensitivity) as the ordinate and the false positive rate (1-specificity) as the abscissa. In addition, we also calculate the area under an ROC curve (AUC), which ranges from 0.5 to 1. The larger the AUC value, the better the performance of the model.

3.2. Parameter selection

In order to achieve the best performance of our model, we use the grid method to select the parameters of SVM. Since SVM classifier has two main parameters: c and g, so we adjust them separately. In the experiment, referring to the prior knowledge, we set the values of c to 0.01 to 100 and g to 1 to 500, respectively. According to the accuracy of these parameters, we draw a three-dimensional image to describe the visualization, as shown in Figure 4.

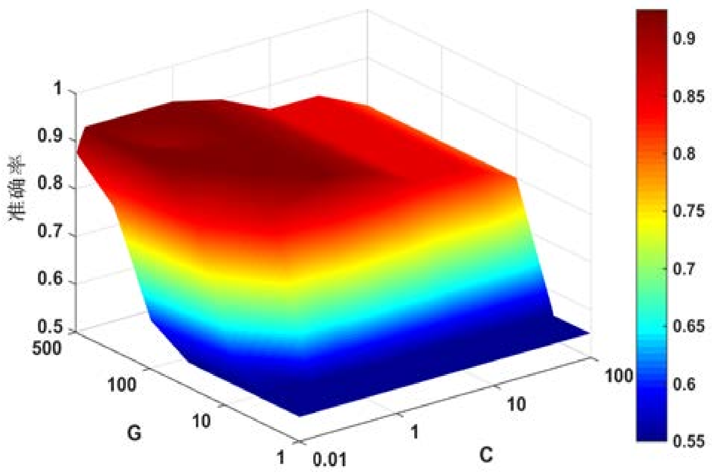

Figure 4. Accuracy of 3D images Generated by Grid Method.

In order to facilitate the observation, we set different colors for different accuracy in the image. Accuracy from low to high set in blue to deep red. It can be seen from the figure, when the value of g is too small, the accuracy does not change much and low with the increase of C. When the value of G is greater than 10, the rate of accuracy increases with the increase of G, but the value is mainly between 0.001 and 10 in C. Therefore, considering the cost of computing, we finally choose the values of C and G to be 1 and 10, respectively.

3.3. Model performance

In order to make the experimental results persuasive, we use the 5-fold cross-validation method to experiment. Specifically, we first upset the data set and randomly divide them into roughly five equal parts. Secondly, we take one of them as a test set, and the remaining four as a training set to conduct the first set of experiments. And then take a different as a test set, the remaining as a training set for the second set of experiments, in order until the five have done the test set, so we get five sets of results. Finally, we take the average of these five sets of results as the final experimental result, and calculate the standard deviation. The results of the 5-fold cross-validation of our model on the milk powder data set are shown in Table 1.

Table 1. The 5-fold cross-validation results value (%) were generated on the milk powder dataset.

Test set	Accu.	Sen.	Prec.	MCC	AUC
1	90.00	100	84.00	81.43	98.25
2	92.50	88.24	93.75	84.65	98.72
3	87.50	73.68	100	77.15	96.49
4	92.50	90.91	95.24	85.03	98.99
5	92.50	85.71	100	86.04	98.75
Average	91.00	87.71	94.60	82.86	98.24
Deviation	2.24	9.52	6.55	3.63	1.01

We can see from the table, our model achieved an accuracy of 91% with a standard deviation of only 2.24%. The values of other evaluation criteria such as accuracy, sensitivity, precision and Matthews correlation coefficient were 87.71%, 94.60%, 82.86% and 98.24% respectively, and their standard deviations were 9.52%, 6.55%, 3.63% and 1.01% respectively. From the five sets of results, we achieved the lowest accuracy rate of 87.50%, up to 92.50%. The ROC curve obtained by our method is shown in Figure 5.

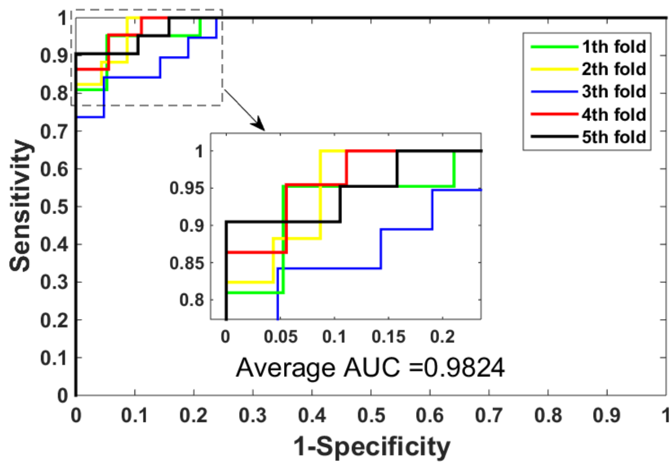

Figure 5. The ROC curves were generated on the milk powder data set by using the proposed method.

4. Conclusions

In this article we proposed a model that combines stacked auto-encoders and support vector machine to detect contraband in milk powder. The model first extracts features of CT images of milk powder using stacked auto-encoders and then classifies them using a support vector machine classifier to detect whether the contraband is carried in the milk powder. In the experiment, we used the 5-fold cross-validation method to prevent data over-fitting. In addition, in order to improve the accuracy of the test, we also tested the effect of different parameters on the test results. The excellent experimental results show that our proposed model can effectively identify the contraband with different shapes, sizes and densities entrained in the milk powder cans. In future research, we will improve the method of feature extraction so as to expect to achieve better results.

Acknowledgments

This work is supported in by the National Science Foundation of China, under Grant 61702444. The authors would like to thank all anonymous reviewers for their constructive advices.

References

[1] Mouton A, Breckon TP. A review of automated image understanding within 3D baggage computed tomography security screening [J]. Journal of X-ray science and technology, 2015, 23(5): 531.
[2] Han Y, Han Y, Li R, et al. Application of X-ray digital radiography to online automated inspection of interior assembly structures of complex products [J]. Nuclear Instruments & Methods in Physics Research, 2009, 604(3): 760-764.
[3] Albert RD. Multiple image scanning X-ray method and apparatus. In: US, 1993.
[4] Champness JN, Bennett MS, Wien F, et al. Exploring the active site of herpes simplex virus type-1 thymidine kinase by X-ray crystallography of complexes with aciclovir and other ligands [J]. Proteins-structure Function & Bioinformatics, 2015, 32(3): 350-361.
[5] Ikonen AEJ, Manninen HI, Vainio P, et al. Three-dimensional respiratory-gated coronary MR angiography with reference to X-ray coronary angiography [J]. Acta Radiologica, 2015,44(6): 583-589.
[6] Wang Y, Yang X, Wu W, et al. An X-ray inspection system for illegal object classification based on computer vision [J]. International Journal of Security and Its Applications, 2016,

10(10): 155-168.
[7] Gao H, Zhang L, Chen Z, et al. Application of X-ray CT to liquid security inspection: System analysis and beam hardening correction [J]. Nuclear Inst & Methods in Physics Research A, 2007, 579(1): 395-399.
[8] Wang W, Ooi BC, Yang X, et al. Effective multi-modal retrieval based on stacked auto-encoders [J]. Proceedings of the Vldb Endowment, 2014, 7(8): 649-660.
[9] Suk HI, Lee SW, Shen D. Latent feature representation with stacked auto-encoder for AD/MCI diagnosis [J]. Brain Structure & Function, 2015, 220(2): 841-859.
[10] Sun L, Liu H, Zhang L, et al. lncRScan-SVM: A Tool for Predicting Long Non-Coding RNAs Using Support Vector Machine [J]. Plos One, 2015, 10(10): e0139654.
[11] Harris T. Credit scoring using the clustered support vector machine [J]. Expert Systems with Applications, 2015, 42(2): 741-750.
[12] Nasiri JA, Charkari NM, Jalili S. Least squares twin multi-class classification support vector machine [J]. Pattern Recognition, 2015, 48(3): 984-992.