IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 26, 2019, accepted June 14, 2019, date of publication June 24, 2019, date of current version July 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2924445

A Machine Learning-Based Approach for Counting Blister Cards Within Drug Packages

MAHDI BAHAGHIGHAT[1], LEILA AKBARI[1], AND QIN XIN[2]
[1]Department of Electrical Engineering, Raja University, Qazvin 95834, Iran
[2]Faculty of Science and Technology, University of the Faroe Islands, Tórshavn FO 100, Faroe Islands

Corresponding author: Mahdi Bahaghighat (m.bahaghighat@raja.ac.ir)

The work of Q. Xin was supported in part by the EU COST grants from European Network for Cost Containment and Improved Quality of Health Care under Grant CA15222.

ABSTRACT Nowadays with the rapid development of technologies, machine vision has been used widely in various industries. The main applications of machine vision in industrial product lines are quality control (QC) and quality assurance (QA). The intelligent defects and anomalies recognition throughout the supply chain have come to be an integral part of quality control systems, in particular, in the food and pharmaceutical industries. In these industries, it is a legal requirement in manufacturing processes which can lead to minimizing the total number of defected products as well as maximizing the performance. In this paper, the machine vision has been utilized to monitor and control the proper packaging of drugs in pharmaceutical product lines. The main goal is counting the number of blister cards within a drug package. To tackle this problem, a new model based on object detection, feature extraction, and classification is proposed. Thanks to several strong approaches, such as the Haar cascade, HOG, ORG, Gabor wavelet, Radon transform, KNN, and SVM, and the accuracy over 88% is achieved in our experiments.

INDEX TERMS Machine vision, quality control, object detection, object counting, Blister counting, SVM, KNN, Radon transform.

I. INTRODUCTION

Quality control (QC) is a set of operations such as measurement or testing which is applied to a product to determine whether its specifications comply with technical standards or not. One of the main duties of a quality control system is to supervise production and in the case of any problem, it should be capable to cease production line and prevent the production of low quality or damaged products. The quality of pharmaceutical products has great significance for healthcare services providers, assurance companies, standardization organizations, and patients. As a result, its inspection has become more crucial and urgent with increasing attention [4], [61]. In addition, in order to achieve zero defect QC systems, machine vision systems have emerged as intelligent technologies for visual inspection of different attributes in wide range industries [4], [61]. For examples, in [21] an algorithm has been proposed for describing an automated inspection system in order to test slate slabs by a 3D color camera and study slate slab characteristics through using some algorithms in computer vision. In their model, texture analysis had been used to extract required features before applying a neural network classifier [21]. In [44], authors proposed an algorithm of computer vision in order to measure and inspect O-rings. In their design edge detection was followed by two main modules. The first module for detecting a defect of O-rings and the second one for measuring its dimension and thickness [44]. In [57], a plane for automatic illusion regarding inspection system for robot vision was proposed. In that system, through using of intensity histogram, features were extracted and for amplifying the difference in primary features, a submodule using a neural network was deployed before sending it to the classifier. Then, they applied the stacked auto-encoder as an amplifier to get the evolved features just because those are more discriminative [57]. Afterward, assessment of brightness quality of the image was interpreted to the problem of classification. The Softmax classifier was used for identification of image brightness quality via final features that had been evolved [57]. In [61], an automated machine vision system (MVS) for inspection of liquid medical injection was designed. In this algorithm, a captured raw sequence of images was gained and then an algorithm for segmentation was applied according to the difference in image sequence by using a fuzzy cellular

The associate editor coordinating the review of this manuscript and approving it for publication was Feng Shao.

neural network (FCNN). Finally, particles and bubbles could be distinguished from the image sequence via the method of classification and moving tracking. In [59], a system for automated optical inspection was developed. It contained integrated linear motors [48], lighting device and CCD cameras for detection of the area along with holes location on a PCB. Drillings defects were classified into missing and multi-holes as well as aperture error through the use of recognition and matching approaches. In [12], at the first, removing noise from a binary image was done using some preprocessing methods then fabric defects could be detected from a binary image through the labeling algorithm along with size filtering. In the case of successive appearance of defects, defects were saved in a memory queue. Extraction of defects could be done when the storing comes to end.

A vision-based quality control system for pharmaceutical factories is modeled in this work. To the best of our knowledge, there is a few works concentrating on counting blister cards within a drug package while moving in a conveyor belt. So we targeted designing and developing such a system based on computer vision techniques. In this paper, drug packages are detected using some object detection algorithms such as Haar cascades and template matching (TM) algorithms. Afterward, features are extracted using HOG, ORB, Radon transform and Gabor wavelet. In the end, our algorithm would be finalized with SVM and KNN classifiers to make a decision about the number of blisters within a package. The rest of this paper is constructed as follows: In Section II, we discuss related works. In Section III, the methodology and proposed model are introduced. Simulation results are then presented in Section IV and finally, Section V contains conclusions.

II. RELATED WORKS

There are a lot of applications for object detection algorithms. In [28], a novel method which is able to detect the face in analog cameras was proposed [28]. First, an efficient method for calculation of integral image via using of cumulative line sum was implemented then, in order to avoid division, an alternative approach found on compact integral image generator was suggested. One of the most degrading factors in object detection problems is illumination variations. An upgraded feature descriptor for better detection of objects under different illumination condition, named Haar contrast feature was proposed in [12]. This feature uses exactly the Haar-like feature prototype and by using the normalization factor, the feature was able to compute contrast. The normalization factor was designed to show the average intensity in a feature region. A new model for object detection was evaluated in [28]. In that method, several innovative ideas for Adaboost learning was presented. Firstly, non-adjacent Haar-like features with all topological properties and secondly a new Adaboost algorithm based on topology were suggested [19]. Thirdly, a scheme for tuning the parameters of cascade for Adaboost cascade learning was presented. Finally, the introduction of estimation of Gaussian kernel density to Adaboost was done [19]. In [43], a method for estimation of the Rotation, Scaling, and Translation (RST) transforms, relevant to unknown images from their projection was developed. To fulfill this goal, the Radon transform was defined to encounter the problem of rotation using an angle between 0 and 2π. The application of the algorithm can be extended on spatial image sonogram. Moreover, the algorithm is useful in raw data in projections. In [51], authors tried to understand whether some pre-filters such as a curve pre-filter and pre-filters for edge detection/smoothing were able in object detection accuracy improvement of ORB or not (without having a noticeable effect on the runtime). Moreover, a new survey on co-variance ellipses was done in order to evaluate the performance of the filter and determining the threshold in Hamming distance [51]. In the suggested technique of denoising in [47], the noisy image was divided into a great number of patches which were overlapped. Also by use of Scale Invariant Feature Transform (SIFT), local features were extracted from each patch. Afterward, for classification of the patches into two classes like flat patches and texture patches, the Support Vector Machine (SVM) was used. The texture patches proceeded through gradient histogram preservation (GHP) and reconstruction of the flat patches was done by sparse based denoising method analysis K-means Singular Value Decomposition (K-SVD) [1], [47]. Lastly, through merging the final results of the denoising tasks, the reconstructed image was gained. In [41], the goal of the paper was to take a model based on clustering by suggesting a different technique. In this method, a faster k nearest neighbor (KNN) technique was developed to replace the routine mean shift method. Moreover, for robustness, ideas from ensemble consensus clustering were incorporated considering hyperparameters. [42] presented two famous and pioneer feature extraction methods that are being used. One of these methods was Gabor wavelet based feature extraction. This method is used in capturing orientation, frequency, and locality. It also provides information about multi-resolutions texture, both in spatial and frequency domains. The second method was a statistical feature extraction method. This uses texture-based feature extraction methods like the linear binary pattern (LBP), gray level co-occurrence matrix (GLCM), the histogram of oriented gradient (HOG) and gray level run length matrix (GLRLM). These methods show the connection between the intensity of pixels in two images or pixel group. Moreover, estimation of statistics of first- and second-order related to the image properties was done using this method [42]. Apart from segmentation and tumor detection, with respect to this application, a research about how effective and complex feature extraction methods are, was done.

III. METHODOLOGY

The main task of a quality control system is to supervise production procedure to prevent any defect and damage in final goods. In this paper, a new model for blister counting in the pharmaceutical production line is presented. The proposed algorithm is divided into three main steps:

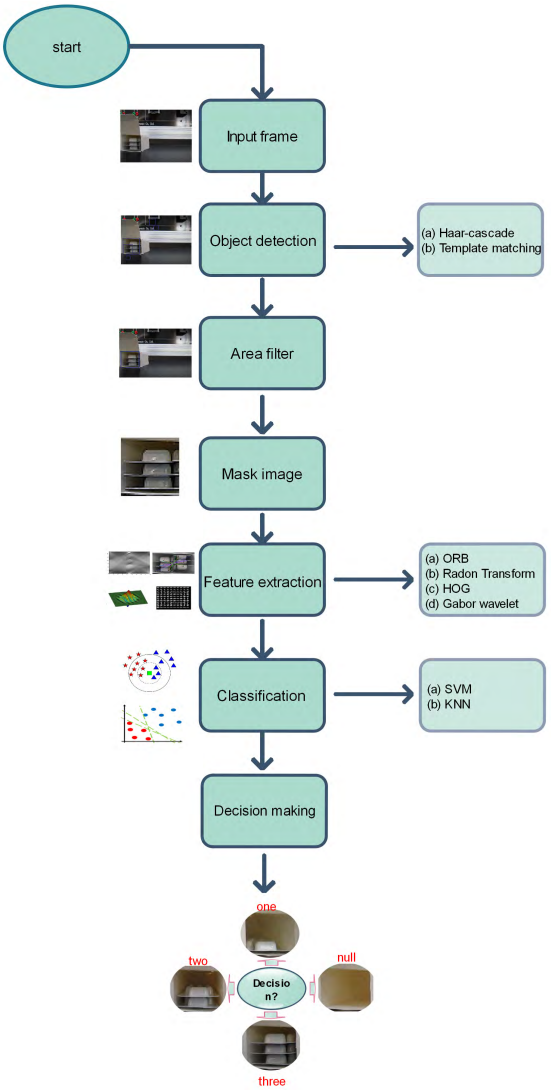

FIGURE 1. Block diagram of our proposed algorithm for counting blister cards within drug packages.

object detection, feature extraction, and classification. Object detection is implemented using Haar cascade and template matching (TM) methods while feature extraction is done by HOG, ORB, Gabor Wavelet [4], [5], [26], and Radon transform [4], [38]. At the final step, features are fed into SVM and KNN classifiers to make a decision among four defined classes **Null, One, Two,** and **Three** which are the number of blisters cards may be found in a drug package. Fig. 1 shows the block diagram of our proposed algorithm.

A. DATASET

For designing an efficient machine vision system, every aspect of the procedure from the dataset to the implementation phase should be considered accurate. In this work, to generate a proper dataset, several important terms were taken into consideration: intensity and color of the light source, angle of view, distance, and moving object's speed. Since

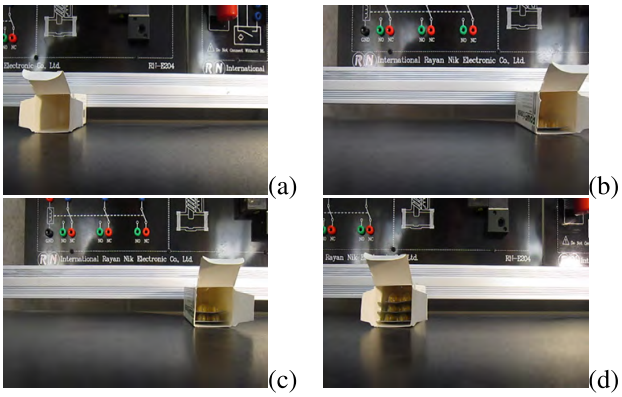

FIGURE 2. Some selected samples from our database: DB1. (a) Null, (b) one, (c) two, and (d) three.

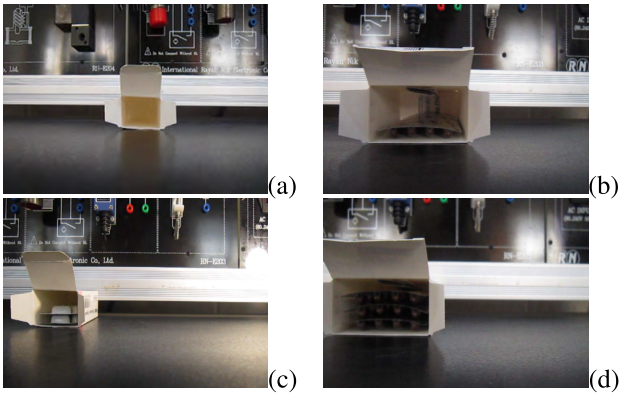

FIGURE 3. Some selected samples from our database: DB2. (a) Null-Type A, (b) one-type B, (c) two-type A, and (d) three-type B.

the rate of emitted light varies depended on environmental conditions, gathering samples in different light intensities and colors was done. In addition, angles between the camera and a drug package and their distances have great significance. Due to the fact that a drug package is carried by a conveyor belt, when a camera is located in front side with a fixed angle it can capture images in different angles linearly. In order to emphasize scalability, we also changed the distance between the camera and a package in database creation steps. As a result, two databases named DB1 and DB2 were created. In Figs. 2 and 3, some samples are depicted. Fig. 4 also indicates more information about data distribution for all classes.

For object detection problem DB1 and DB2 are deployed while for the classification problem, the bounding box Ground truth information of drug packages are required for all images in DB1 and DB2. This information was created using powerful annotation software: the LabelImg [55]. As a result, two databases named DB3 and DB4 were also developed. These new databases include all patches of drug package images according to DB1 and DB2, respectively. We use DB3 and DB4 in the learning phase of our proposed classifiers.

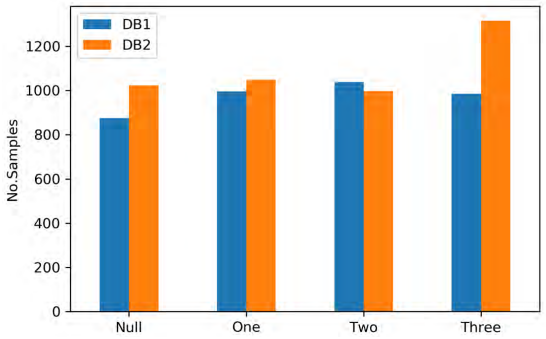

FIGURE 4. Distribution of samples in each class in our dataset (DB1 and DB2).

TABLE 1. Template matching (TM) list with descriptions.

TM index	TM method [22]	Description [22]
TM 1	CV_TM_SQDIFF	Square difference matching that tries to minimize the square difference among the pattern and the image area.
TM 2	CV_TM_SQDIFF_NORMED	Normalized version of square difference matching.
TM 3	CV_TM_CCORR	Correlation matching method multiplicatively matching a template against the image and then maximizing the matched area.
TM 4	CV_TM_CCORR_NORMED	Normalized version of correlation matching method.
TM 5	CV_TM_CCOEFF	Correlation coefficient matching method that matches a template against the image relative to theirs means and generates a matching score ranging from -1 to 1.
TM 6	CV_TM_CCOEFF_NORMED	Normalized version of the correlation coefficient matching method.

B. OBJECT DETECTION

The main step of our model which presented before is the object detection. In this paper, some important template matching algorithms and Haar cascade are used in order to compare detection accuracy. Some different types of template matching algorithms are described in the next section. Afterward, the Haar cascade algorithms will be explained in detail.

1) TEMPLATE MATCHING

Template matching (TM) is a technique to find a small object from an image which matches to a template object [20]. It compares a template image with sliding a window throughout an input image to probe the most similar window [20]. In TM, the procedure usually is run with zeros padding operation in the reference image and exploring the center point of a template image [20]. Table 1 shows TM methods of OpenCV libraries [40] that are implemented and compared in our work.

2) HAAR CASCADE

A very successful method of Viola and Jones [56] is the typical cascade classifier [13]. The technique starts with Haar features phase in which an object detection is done using a high differentiator between a non-object and an object [13]. The procedure of object detection from this method consists of four main parts: integral image, Haar-like features, adaptive boosting or the Adaboost and cascade classifier combination [37].

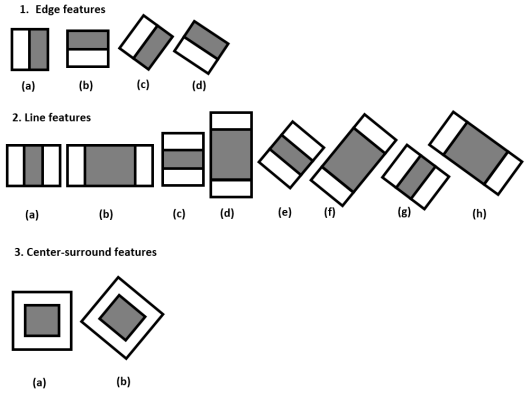

FIGURE 5. Haar-like features [13].

Fig. 5 shows the Haar-like characteristics by default. These can be used with all scales in the boosted classifier and computed so fast from an integral version of the image to be detected [13].

The difference between the sums of pixels of the dark area calculates this feature using below formula [13], [15]:

$$F(Haar) = \sum_{area} F_{White} - \sum_{area} F_{Black} \tag{1}$$

where $\sum_{area} F_{Black}$ is the sum of pixels of the dark area and $\sum_{area} F_{White}$ is the sum of pixels of the white area. As the Haar-like features get higher than a predefined threshold, this means that an object (or more) is within the area.

In order to filter a large number of objects within the image efficiently, the Integral image technique is used [15]. It is used for calculating the amount of the Haar like feature rapidly via modifying each pixel into a new image representation [16], [50]. In the input image, each pixel of the integral image has an amount which is built up from the top and left of the pixel position [16].

The Adaboost is a widely used classifier to train a large number of features. In the detection process, instead of only one kind of feature, different types of features with various scales and positions are used. That is why the calculation in the detection process slows down [63]. It is an algorithm which is used in the Viola and Jones [56] method in detection procedure in order to decrease the number of features which are not relevant. Usually, Adaboost is used to find features that have a high degree of differentiation [63]. This algorithm urges to find these features via evaluating every feature against the training data using the feature values. Most efficient features have the highest boundary between an object and non-object [16]. It is deployed to enhance the classification performance with simple training and combining a lot of weak classifiers into just one strong classifier [26]. A weak classifier has a correct answer but with a degree of truth which

is less accurate, comparing to a strong classifier which has a correct answer with an accurate degree of truth [26].

Viola and Jones [56] also introduced a classifier with cascade structure in order to speed up the detection step [37]. The Adaboost learning algorithm selects a group of Haar like features in each stage of cascade. While detecting almost all candidates in the first stages, the classifier is trained to decline most of the negative subwindows [8], [39]. This structure can accelerate the detection process greatly as most of negative images can be thrown away during the first two or three stages and the computation can be focused on just object like subwindows [37]. The stage classifier evaluates subwindows in sequences and in each stage the result of each Haar-like features is accumulated [19]. The accumulated value is compared with a stage threshold value when all the features in the stage are computed. This is done to see whether the current subwindow is an object like candidate or not. In the case that the result of the previous stage is positive, the next stage will be activated. In order to determine if the subwindow contains an object, the candidate should pass all the stages in the cascade. To implement Haar cascade classifier, we used OpenCV libraries [28].

C. FEATURE EXTRACTION

After detection of a drug package, proper features should be extracted from it. In this work, some strong methods such as HOG, ORB, Radon transform and Gabor wavelet are candidates to tackle this issue.

1) ORB FEATURES

Combination of the Feature from Accelerated Segment Test (FAST) key point detector and Binary Robust Independent Elementary Features (BRIEF) descriptor with some changes is called ORB [51]. Two main features of ORB which can be highlighted are being resistant to noise and rotation invariant [25]. Legacy experiments evidently demonstrate how ORB is more than two orders of magnitude faster than SIFT, while performing as well in many situations [25]. By modifying the intensity threshold between the center pixel and pixels in a circular shape ring around the center, a set of binary tests over a patch is implemented by the original FAST proposal [17]. To evaluate the corner intensity, the Harris corner measure has been used [25]. Computation of the moment m_{pq} of the patch I_p and centroid C can be done by Equations 2 and 3 [17], [51]:

$$m_{pq} = \sum_{x,y} x^p y^p I(x, y) \tag{2}$$

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}\right) \tag{3}$$

The relative orientation of the patch can be defined via the construction of a vector which starts from the patch center O and ends at the centroid C as below [17], [51]:

$$\omega = Atan2(m_{01}, m_{10}) \tag{4}$$

The patch description has been provided starting from the BRIEF operator, a bit string representation constructed from a set of binary intensity tests. Given a smoothed image patch I_p of an intensity image, a binary test can be performed as [17], [51]:

$$\tau(I_p, x, y) = \begin{cases} 1 & I_p(x) < I_p(y) \\ 0 & I_p(x) \geq I_p(y) \end{cases} \tag{5}$$

The final feature of image patch I_p, $b(I_p)$ is also defined as a weighted summation of n binary tests [17], [51]:

$$b(I_p) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(I_p, x_i, y_i) \tag{6}$$

The intra patch locations for the tests, therefore the choice of vector influences the quality of the descriptor itself [17]. Some sampling tests from 5×5 subwindow of the 31×31 patch window which have been chosen for the descriptor, running each test opposed to all training patches could be a solution for this. The result is a predefined set of 256 tests called Binary Robust Independent Elementary Features (BRIEF) [25], [51].

$$L = \begin{pmatrix} x_1, & \cdots & x_n \\ y_1, & \cdots & y_n \end{pmatrix} \tag{7}$$

Area Filtering: After running a TM algorithm, we have one object which can be either the true or false object but the output of Haar cascade could be multi-objects (which are detected by the algorithm) or nothing. After implementing the Haar cascade method on the input image area, a filtering process is done to reject an unwanted object which is detected. The area filtering is done to reject objects that are considered too small or very large to be the correct object. Assuming the width and height of the detected object is W_0 and H_0 respectively, the area is $A_0 = W_0 \times H_0$ and according to Equation 8, an appropriate object is held (H) while others will be removed ($\overline{H}$):

$$A_S = \begin{cases} \overline{H} & A < \alpha \\ H & \alpha \leq A \leq \beta \\ \overline{H} & A > \beta \end{cases} \tag{8}$$

where A_S is the selected area and α and β are pre-defined thresholds.

2) HOG FEATURES

For taking out the features in a video frame, Histogram Oriented Gradient (HOG) [15] is also considered in our work. The HOG is a feature descriptor that attracts a lot of attention in object recognition applications [30], [49]. The band gesture is distinguished and classified by HOG descriptors through a distribution of pixel gradients in the cell [30]. By using magnitude and direction of the gradient of a pixel, every single cell of the image has a histogram constructed [30]. The intensity $i(x, y)$ at every single cell of the input image

is used for calculating Gradient $m(x, y)$ and its orientation $\theta(x, y)$ [22], [24], [29], [53], [60]:

$$m(x, y) = \sqrt{f_x^2(x, y) + f_y^2(x, y)} \tag{9}$$

$$\theta(x, y) = tan^{-1}\frac{f_y(x, y)}{f_x(x, y)} \tag{10}$$

$$\begin{cases} f_x(x, y) = I(x + 1, y) - I(x - 1, y) \\ f_y(x, y) = I(x, y + 1) - I(x, y - 1) \end{cases} \tag{11}$$

According to the gradient $m(x, y)$ and its direction $\theta(x, y)$ in each cell of $p \times p$ pixels, generation of a weighted histogram of the gradient of intensity is done [22], [27]. At this stage direction of gradient $m(x, y)$ is quantized θ to N orientations and afterward generation of histogram of oriented gradients becomes complete [10], [45]. At the last stage, normalization of the histogram is done [27], [54]. Here each count of the histogram is divided by sum of HOG features in the cell $q \times q$ [22], [24], [29], [53], [60]:

$$v_i^n = \frac{v_i}{\sqrt{\sum_{k=1}^{q \times q \times N} v^2(k) + \epsilon^2}} \tag{12}$$

where v_i^n represents $n'th$ HOG features and ϵ is a near zero constant which is added to prevent the division by zero [22].

3) RADON TRANSFORM

There is an operative technique in the analysis of signal between the spatial domain and its projection space called Radon transform which is defined as an integral transform including the integral of the function over straight lines in each image [3], [32]. For a 2D image $f(x, y)$, Radon transform is described as [2], [4], [6], [7], [32], [33], [36]:

$$g(r, \theta) = \Re(f(x, y)) = \iint (f(x, y)\delta \times (r - xcos(\theta) - ysin(\theta))dxdy \tag{13}$$

In which the distance between the origin and the projection line is r. The projection angle is shown by θ and $\delta()$ denotes the pulse function [9], [43].

4) GABOR WAVELET

A very useful tool for analyzing texture such as image retrieval classification and segmentation is Gabor wavelet [34]. Its filter bank is one of the most practical tools for extracting directional textural features [34]. In the case that a Gabor function is used as the mother wavelet, the Gabor wavelet (GW) can be contemplated as a wavelet transform [46]. For 2D signals, a bidimensional Gabor wavelet is defined as below [5], [18], [42]:

$$g_{w,\theta}(x, y) = \frac{1}{\sqrt{\sigma_v}} \exp\left(-\frac{v^2}{2\sigma_v^2}\right)\frac{1}{\sqrt{\sigma_u}} \exp\left(-\frac{u^2}{2\sigma_u^2}\right) \tag{14}$$

where [2], [33], [36]:

$$u = xcos(\theta) + ysin(\theta) \tag{15}$$

$$v = -xsin(\theta) + ycos(\theta) \tag{16}$$

Here σ_v and σ_u are scale parameters in the direction of the wave and in orthogonal direction respectively [5]. At each point (x_0, y_0) a pixel neighborhood $I(x, y)$ of size $w \times w$ is considered [5]. $I(x, y)$ represents the image intensity at pixel $(x_0 + x, y_0 + y)$ [5]. At first, according to Equation 17, $I(x, y)$ should be normalized to a new image $I'(x, y)$ with constant energy [35]. In this formula, $m(x_0, y_0)$ denotes the mean value of $I(x, y)$ while $v(x_0, y_0)$ indicates its variance [2], [5], [33], [36]:

$$I'(x, y) = \frac{V}{v(x_0, y_0)}[I(x, y) - m(x_0, y_0)] \tag{17}$$

So after the normalization process, the mean value of $I'(x, y)$ is equal to zero and its norm is $V = 1$ independence of a point (x_0, y_0). Afterward, computation of the local projections of I' on each filter of the bank is done. The projection of I' on a Gabor wavelet of frequency $\overline{w}$ and direction α is a complex number $\alpha_{\overline{w},\alpha}$ [2], [5], [33], [36]:

$$\alpha_{\overline{w},\alpha} = \frac{\int_{-W/2}^{W/2}\int_{-W/2}^{W/2} I'(X, Y)g_{\overline{w},\alpha}(x, y)dxdy}{\|g_{\overline{w},\alpha}\|} \tag{18}$$

$$A_{\overline{w},\alpha} = \frac{1}{W^2}\text{Real}(\alpha_{\overline{w},\alpha}) \tag{19}$$

Now the following features can be obtained for a given point (x_0, y_0):

$$(w, \theta) = \text{Arg max}A_{\overline{w},\alpha} \tag{20}$$

where w and θ are the estimated frequency and orientation in the sub-image, respectively.

D. CLASSIFICATION

The last part of our blister counting algorithm is feeding extracted features into a proper classifier. In our work, KNN and SVM are selected for this goal. KNN includes simple implementation and its performance of classification is significant. In comparison, the transformation of a nonlinearly separable problem into a linear one can be done by the SVM. In the next sections, we describe the SVM and KNN models in more details.

1) SVM

A powerful machine learning method which was developed in the early 1990s is called Support Vector Machine or SVM [23], [58]. The method is based on the projection of data into the feature space and then probing the separating hyper plan [23], [58]. Furthermore, the transformation of a nonlinearly separable problem into a linearly separable problem can be done using SVM. The initial SVM which is designed for binary classification applications is not usable in problems with multi-classes classification [23], [58]. Currently, SVM is one of the greatest methods for different classification tasks including text, voice, image data and etc. By designing suitable kernel functions, it can be applied in complex data types. Below are some of the general kernels which are used in such data [14], [47]:

p^{th} Degree Polynomial

$$K(x_i, x_j) = (x_i, x_j + 1)^p \tag{21}$$

Radial bases (RBF) [14], [47]:

$$K(x_i, x_j) = e^{-\| \frac{x_i - x_j^2}{2\sigma^2} \|} \tag{22}$$

Multi-layers perceptron (MLP) [14], [47]:

$$K(x_i, x_j) = tanh(\beta x_i . x_j + \delta) \tag{23}$$

The objective function of the dual problem should be maximized. So the dual problem is defined as Equation 24 [14], [47]:

$$W(\alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_i \sum j \alpha_i \alpha_j y_i y_j K(x_i xj) \quad i, j \in 1, \cdots, n \tag{24}$$

where $0 \leq \alpha_i \leq C$, $\sum_i \alpha_i y_i = 0$, x_i is the region index, and x_j is the color index. After solving, the α matrix can be found [14]. The bias and y-classification are shown as Equations 25 and 26:

$$b = \frac{1}{\|s\|} \sum_i [y_i - \sum_j \alpha_j y_j K(x_j, x)] \tag{25}$$

$$y = sign(\sum_i \alpha_i y_i K(x_i, x) + b) \tag{26}$$

In which the nonzero value of α matrix is shown by s and x denotes the input matrix. Finally, making a decision about the class of input data can be achieved depending on the sign of the classifier.

2) KNN

The K-nearest neighbors or KNN is a very popular method in data mining and statistics [11], [41]. It is a method that calculates distances among all training samples and every test samples [11], [41]. Then we have the nearest neighbor of all test samples. The label of each sample is assigned using a majority rule. It is clear that KNN has simple implementation and results show that it has a great implementation performance [11], [41]. The KNN contains two major steps [11], [31], [41]: allocation of optimal samples [62] and allocation of different optimal values of k for different samples in test data [62]. To select the optimal k in KNN great efforts have been put [11], [31], [62]. For example in [31], authors indicated that the optimal value for K in all test samples could be $K = \sqrt{n}$ with $n > 100$ (n is the number of samples were involved in the training phase).

E. ACCURACY

The performance of a method could be evaluated using different evaluation metrics. Some of these metrics such as accuracy, specificity, sensitivity, precision, average accuracy, Jaccard index record, Dice coefficient, and F1 score are defined in Table 2 [52]. In these expressions, TP stands for True Positive which is the number of samples that are classified correctly that belongs to one of the classes. TN or True Negative is the number of samples which are classified correctly and does not belong to any of classes. False Positive (FP) is the number of samples which are misclassified to one of the classes or number of samples that are classified into the class that they do not belong to that class. Finally, FN which stands for False Negative is the number of samples that don't belong to any classes but classified as it belongs to one of the classes. In this work, Accuracy was used to make a comparison among different approaches.

TABLE 2. Evaluation metrics [52].

Metric	Expression
Accuracy(%)	$\frac{TP+TN}{TP+FP+TN+FN}$
Specificity(%)	$\frac{TN}{TN+FP}$
Sensitivity (Recall)(%)	$\frac{TP}{TP+FN}$
Precision(%)	$\frac{TP}{TP+FP}$
Jaccard-index score(%)	$\frac{TP}{TP+FN+FP}$
Dice coefficient(%)	$\frac{2\times TP}{2\times TP+FN+FP}$
F1_score(%)	$2 \times \frac{(Precision \times Recall)}{Precision+Recall)}$

IV. EXPERIMENT AND RESULTS

In this paper, template matching (TM) and Haar cascade were used for object detection. Also for feature extraction, ORB, Radon transform (RT), Gabor wavelet (GW), and HOG were applied. Finally, in the classification section, we used KNN and SVM. Results have been tested on DB1 to DB4 (See Section III-A) using Python and OpenCV environment.

A. DETECTING A DRUG PACKAGE

Drug package detection algorithm tries to find the location of a drug package in the input frames. Each detected object in frames has a different width and length. The number of detected objects may also vary. Frames which have been extracted from the video may contain a drug package with a background or background itself. In this paper, six TM methods named CV_TM_SQDIFF, CV_TM_SQDIFF_NORMED, CV_TM_CCORR, CV_TM_CCORR_NORMED, CV_TM_CCOEFF, CV_TM_CCOEFF_NORMED and Haar cascade were compared for detection of a drug package. In the next section, discussions and comparisons among all TM methods with achieved results of Haar cascade algorithm on DB1 and DB2 are presented.

1) DETECTING A DRUG PACKAGE: TEMPLATE MATCHING

In this section, a comparison among all TM methods mentioned in previous sections is drawn. In this regard, Fig. 6 demonstrates different results of TM algorithms. The statistical results of these algorithms are also shown in Table 3 and Table 4. Found on results of these tables, it is evident that TM algorithms are not capable enough to get accurate results.It can be viewed that the best accuracy about 51.74% has been achieved by TM_CCOEFF_NORMED for 500×700 images while other TM algorithms got lower accuracies in comparison.

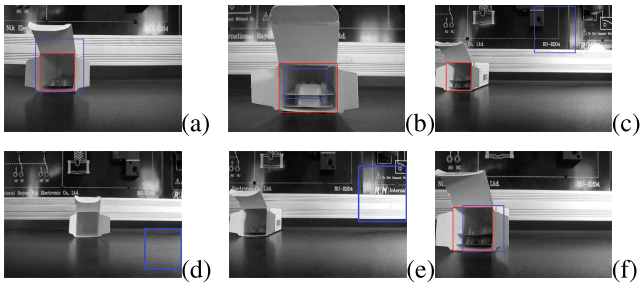

FIGURE 6. Detection of a drug package based on template matching. (a) TM_CCOEFF_NORMED in DB1 (b)TM_CCOEFF_NORMED in DB2 (c)TM_CCOFF in DB1 (d)TM_SQDIFF in DB1 (e)TM_SQDIFF_NORMED in DB1 (f)TM_CCOFF in DB1.

TABLE 3. TM algorithms comparison for a drug package detection in DB1.

Row	database	Input image	Run time(sec)	TM algorithm	Accuracy(%)
1	DB1	500 × 700	0.0209	TM 6	51.74
2	DB1	300 × 500	0.0308	TM 6	37.89
3	DB1	150 × 250	0.0276	TM 6	24.349
4	DB1	500 × 700	0.0221	TM 5	33.40
5	DB1	300 × 500	0.0199	TM 5	5.29
6	DB1	500 × 700	0.0173	TM 4	51.51
7	DB1	300 × 500	0.0171	TM 4	36.61
8	DB1	150 × 250	0.0162	TM 4	18.51
9	DB1	500 × 700	0.0206	TM 2	49.36
10	DB1	300 × 500	0.0263	TM 2	43.06
11	DB1	150 × 250	0.0222	TM 2	25.42
12	DB1	500 × 700	0.0203	TM 1	44.40
13	DB1	300 × 500	0.0206	TM 1	37.56
14	DB1	150 × 250	0.0175	TM 1	26.49

TABLE 4. TM algorithms comparison for a drug package detection in DB2.

Row	database	Input image	Run time(sec)	TM algorithm	Accuracy(%)
1	DB2	500 × 700	0.0315	TM 6	34.93
2	DB2	300 × 500	0.0304	TM 6	38.08
3	DB2	150 × 250	0.0306	TM 6	15.61
4	DB2	500 × 700	0.0256	TM 5	29.32
5	DB2	300 × 500	0.0301	TM 5	8.02
6	DB2	500 × 700	0.0217	TM 4	48.16
7	DB2	300 × 500	0.0177	TM 4	29.64
8	DB2	150 × 250	0.0128	TM 4	16.45
9	DB2	500 × 700	0.0235	TM 2	49.79
10	DB2	300 × 500	0.0300	TM 2	29.00
11	DB2	150 × 250	0.0249	TM 2	14.87
12	DB2	500 × 700	0.0233	TM 1	46.42
13	DB2	300 × 500	0.0205	TM 1	30.00
14	DB2	150 × 250	0.0295	TM 1	35.96

2) DETECTING A DRUG PACKAGE: HAAR CASCADE

An important point that should be considered is that despite having powerful methods for feature extraction and classification, we will not have a high accuracy counting system in the case of failure in the drug package detection step. For this purpose, in the object detection stage of the proposed algorithm, we have used Haar cascade method. In order to have a wider presentation of obtained results, not only Accuracy but also Specificity, Recall, Precision, Jaccard-index score, Dice coefficient, and F1- score are demonstrated in Table 5. Table 5 shows the fact that the accuracy of the smaller dataset (DB1) with single package type outweighs the larger dataset

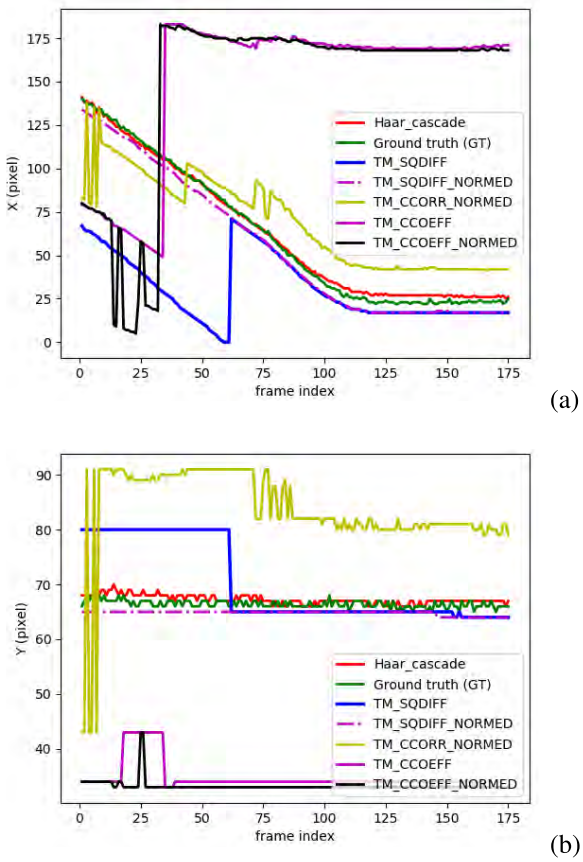

FIGURE 7. The drug package estimated position vs. video frame for a sample video (video 1 with image size 150×250). (a) x-Axis (horizontal variation) and (b) y-axis (vertical variation).

TABLE 5. Using Haar cascade for a drug package detection.

DB	Input image	Accuracy(%)	Specificity(%)	recall(%)	Precision(%)	Jaccar_index(%)	Dice coeeff(%)	F1_score(%)
DB1	500 × 700	90.07	63.50	98.16	89.83	88.35	93.81	93.81
DB1	300 × 500	81.08	33.95	100	79.04	79.04	88.29	88.29
DB1	150 × 250	85.11	42.63	100	83.25	83.25	90.86	90.86
DB2	500 × 700	84.60	62.53	98.29	80.87	79.75	88.73	88.73
DB2	300 × 500	58.75	21.28	97.84	54.37	53.73	69.90	69.90
DB2	150 × 250	67.19	39.32	94.95	61.11	59.18	74.36	74.36

with two types A and B. For example, it is about 90.07% for 500 × 700 images in DB1 comparing to 84.60% for the same sizes in DB2.

3) DETECTING A DRUG PACKAGE: COMPARING HAAR CASCADE AND TMS

In order to make a wider comparison between Haar cascade and TMs, Fig. 7 is depicted. In this new experiment, a sample video (Video 1) is used. It contains a moving object (a drug package) that is moving horizontally from the right to left and stop at its end. X and Y in Fig. 7 (a) and (b) are estimated positions of the drug package throughout the Video 1. Obtained results indicate that the proposed Haar cascade algorithm

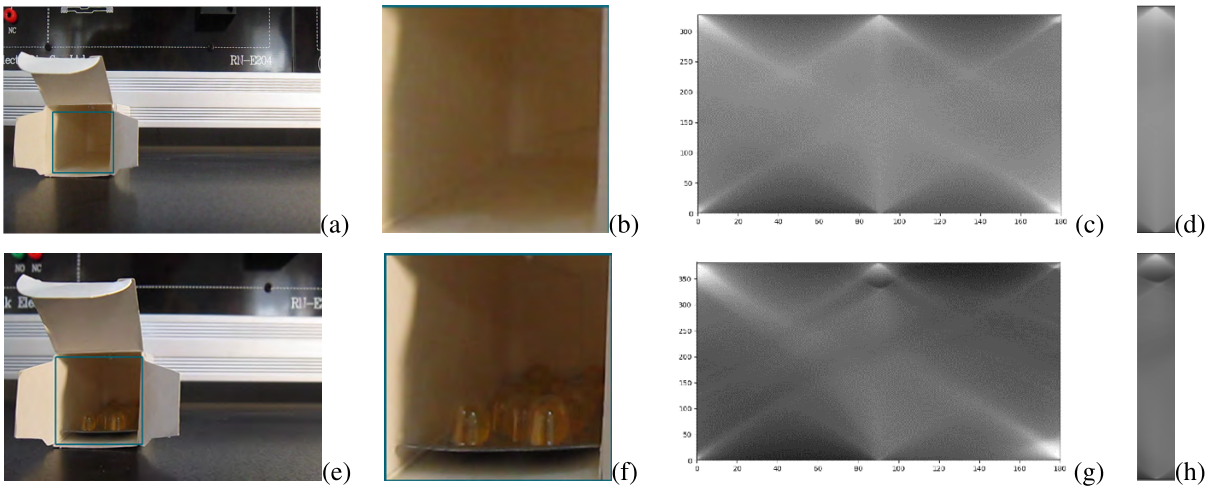

FIGURE 8. Feature extraction based on Radon transform for Null & One. (a) A detected drug package (Null). (b) Masked image (Null). (c) Radon transform (Null). (d) Masked RT (Null). (e) A detected drug package (One). (f) Masked image (one). (g) Radon transform (one). (h) Masked RT (one).

accurately tracks the object position and outperforms all TM methods strongly. Consequently, it would be a proper choice in object detection stage in our model.

B. COUNTING BLISTER CARDS: FEATURE EXTRACTION AND CLASSIFICATION

After object detection, its proper features should be extracted. For this purpose, Radon transform (RT), Gabor wavelet, ORB and HOG were implemented. Fig. 8 and Fig. 9 shows the generated output using RT. In this stage, Radon function is deployed for the masked image resulting from the Haar cascade algorithm. Due to the fact that blister cards within a package are located horizontally, the angle ranges around $\frac{\pi}{2}$ have more important information about the directional pattern and its variation. Therefore in order to reduce the size of the feature vector and also increase run time performance, a small patch which named Masked RT is extracted around the $\frac{\pi}{2}$ from RT sinogram. Fig. 8 and Fig. 9 show the results of applying RT on the masked input image.

It has been shown in these figures that Masked RT images follow special patterns for each class: Null, one, Two and Three. We found that the number of dark cloud points in Masked RT images are equal with the number of blister cards so outputs from RT potentially can lead to an acceptable distinction between all classes. As a result, we have used this feature of RT for counting blister cards.

Besides, in order to draw a wide comparison, simulation results for all proposed feature extraction methods including GW, ORB, HOG, and RT are presented in Table 6 and Table 7 for Ground truth drug package image datasets (DB3 and DB4). After applying a feature extraction algorithm, feature vectors should be fed into a classifier. In this work, two strong classifiers SVM and KNN are candidates to tackle decision making task among four defined classes. In our experiments, DB3 and DB4 were split in to train and test dataset by 0.8 split ratio. Based on achieved results, our proposed approach based on RT outperforms all other methods with the accuracy of about 100%. Regarding classifiers, better results were evidently obtained by KNN in comparison to SVM. It implicitly shows that our data set would not be easily separable using the SVM decision planes. In fact, the SVM deploys linear hyperplanes to discriminate all classes while the KNN is capable to produce a highly convoluted decision boundary as it is driven by the raw training data itself.

TABLE 6. Classification results (in percent) for DB3 (Ground truth (GT) drug packages).

Image size	Classifier	Train	Test	GW	ORB	HOG	RT
500 × 700	SVM	☑	☒	80.65	71.44	87.09	100
500 × 700	KNN	☑	☒	85.84	65.71	98.56	100
300 × 500	SVM	☑	☒	79.25	69.93	81.36	100
300 × 500	KNN	☑	☒	85.11	62.52	98.079	99.92
500 × 700	SVM	☒	☑	56.85	57.02	80.89	92.11
500 × 700	KNN	☒	☑	61.35	54.43	85.71	93.43
300 × 500	SVM	☒	☑	54.30	45.88	71.64	95.54
300 × 500	KNN	☒	☑	50.05	48.67	81.07	98.37

TABLE 7. Classification results (in percent) for DB4 (ground truth (GT) drug packages).

Image size	Classifier	Train	Test	GW	ORB	HOG	RT
500 × 700	SVM	☑	☒	89.47	71.19	99.03	99.51
500 × 700	KNN	☑	☒	85.43	66.12	99.23	99.73
300 × 500	SVM	☑	☒	85.65	70.09	93.91	99.11
300 × 500	KNN	☑	☒	71.51	64.23	95.13	99.02
500 × 700	SVM	☒	☑	53.91	45.71	82.52	84.94
500 × 700	KNN	☒	☑	56.86	57.14	78.07	82.21
300 × 500	SVM	☒	☑	50.84	55.13	73.61	78.90
300 × 500	KNN	☒	☑	43.88	52.10	75.30	75.40

In addition, Fig. 10 demonstrates a deeper comparison between two superior feature extraction methods: RT and HOG via depicting confusion matrix results. It can be observed that HOG experienced more challenge to discriminate among Null & One and also One & Two classes. This ultimately decreases total accuracy for HOG comparing to RT.

At the last stage, simulation results have been presented for all steps from the object detection to the fea-

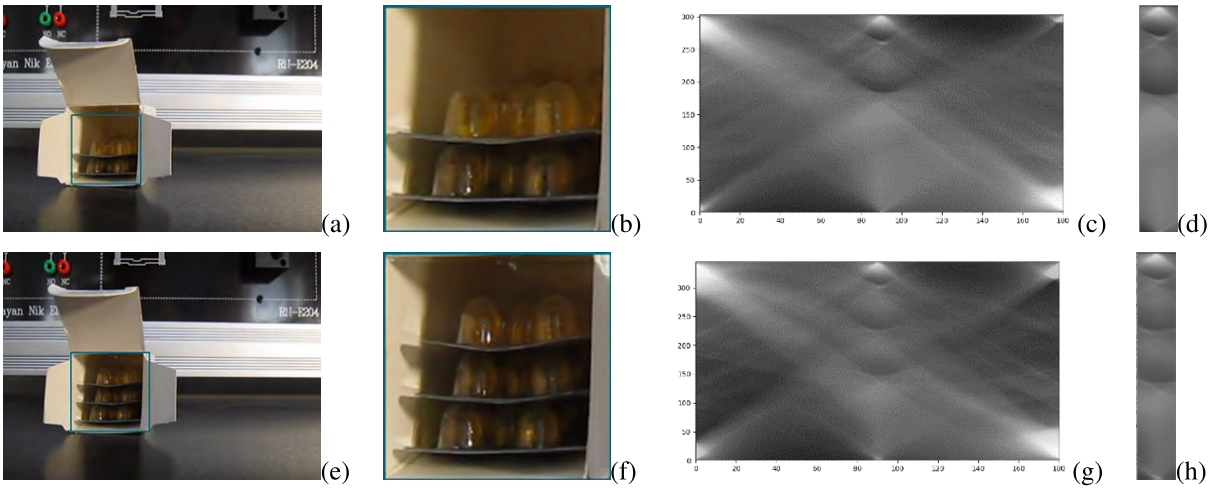

FIGURE 9. Feature extraction based on Radon transform for Two & Three. (a) A detected drug package (two). (b) Masked image (two). (c) Radon transform (two). (d) Masked RT (two). (e) A detected drug package (three). (f) Masked image (three). (g) Radon transform (three). (h) Masked RT (three).

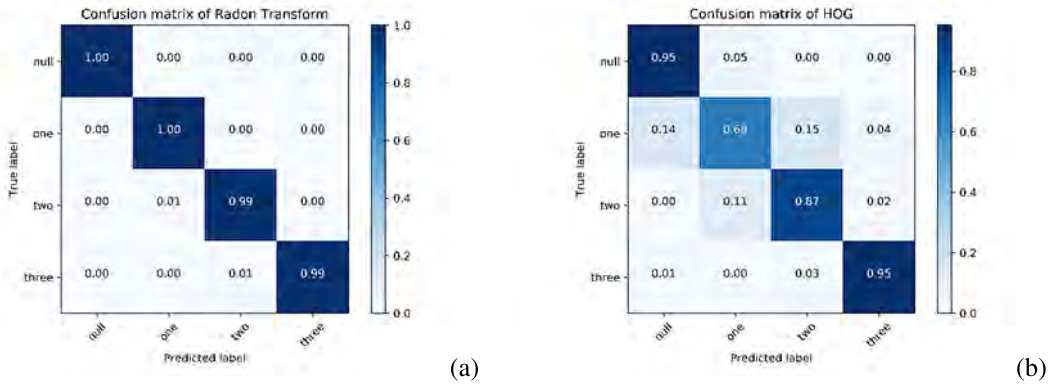

FIGURE 10. Confusion matrix for Top-2 feature extraction approaches: RT and HOG. (a) Confusion matrix of Radon transform and (b) confusion matrix of HOG.

TABLE 8. The prediction accuracy (in percent) comparison for our proposed blister counting system.

DB	Image size	Object detector	Classifier	Feature extraction	Prediction accuracy (%)
DB1	500 × 700	Haar cascade	SVM	RT	86.29
DB1	500 × 700	Haar cascade	KNN	RT	88.33
DB1	300 × 500	Haar cascade	SVM	RT	83.04
DB1	300 × 500	Haar cascade	KNN	RT	83.69
DB2	500 × 700	TM2	SVM	RT	25.33
DB2	500 × 700	TM2	KNN	RT	29.25
DB2	500 × 700	TM1	SVM	RT	52.50
DB1	300 × 500	TM2	KNN	RT	26.90
DB2	500 × 700	Haar cascade	SVM	RT	68.63

ture extraction (RT) and classifications in Table 8. The scenario was implemented for original datasets DB1 and DB2 rather than DB3 and DB4. Due to the degrading effect of object detecting step in our cascade system, achieved accuracies in this stage is lower than previous experiments in Tables 6 and 7.

V. CONCLUSION

The inspection of different elements in packaging systems for a pharmaceutical product is aimed at ensuring that medicines will hand over safely to end users and patients. To the best of our knowledge, there is no similar work in the literature that fully concentrated on blister counting in product lines. Therefore, the innovative and new model for blister counting in drug production lines was presented in this paper. Our model included the object detection, feature extractions and classifications methods to solve a challenging object counting problem. The object detection is a very important step in our proposed system. For object detection, we have compared template matching (TM) and Haar cascade algorithms. It has been observed that the TM algorithms could not result in a proper accuracy and in the best scenario just reached about 50% while Haar cascade led to acceptable accuracy around 90%. Afterward, in the feature extraction step, several strong methods such as Radon transform, Gabor wavelet, HOG and ORB algorithms have applied. Obtained results indicated that Radon transform outperformed others while HOG was the second best. In the classification stage, the SVM and KNN have been used. In comparison to SVM, better results were relatively obtained with KNN. Ultimately, the total accuracy of more than 88% was obtained based on our best model using Haar cascade, Radon transform, and KNN.

Todays, deep learning methods are achieving state-of-the-art results in a lot of scientific problems. Specially, Convolutional Neural Networks (CNNs) have been used for many challenges in computer vision. CNN obtained outstanding performance on different tasks, such as visual object recognition, image classification, handwritten character recognition and more. CNNs are fully data-driven and can retrieve hierarchical features automatically by building high-level features from low-level ones, thus obviating the need to manually customize hand-crafted features. In the future works, we are going to deploy deep learning algorithms to improve the accuracy of our object counting system as much as possible.

REFERENCES

[1] M. Aharon, M. Elad, and A. Bruckstein., "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, p. 4311, Nov. 2006.

[2] G. Ambartsoumian, "Inversion of the V-line radon transform in a disc and its applications in imaging," *Comput. Math. Appl.*, vol. 64, no. 3, pp. 260–265, 2012.

[3] R. Aramyan, "To local reconstruction from the spherical mean radon transform," *J. Math. Anal. Appl.*, vol. 470, no. 1, pp. 102–117, 2019.

[4] M. Bahaghighat, M. Mirfattahi, L. Akbari, and M. Babaie, "Designing quality control system based on vision inspection in pharmaceutical product lines," in *Proc. Int. Conf. Comput., Math. Eng. Technol. (iCoMET)*, Mar. 2018, pp. 1–4.

[5] M. K. Bahaghighat and R. Akbari, "Fingerprint image enhancement using GWT and DMF," in *Proc. 2nd Int. Conf. Signal Process. Syst.*, vol. 1, 2010, p. V1-253.

[6] O. Bimber, "An image sensor based on optical radon transform," *Comput. Graph.*, vol. 53, pp. 37–43, Dec. 2015.

[7] F. Biondi, "Low-rank plus sparse decomposition and localized radon transform for ship-wake detection in synthetic aperture radar images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 1, pp. 117–121, Jan. 2018.

[8] A. Bria, C. Marrocco, M. Molinara, and F. Tortorella, "An effective learning strategy for cascaded object detection," *Inf. Sci.*, vol. 340, pp. 17–26, May 2016.

[9] S. Chelbi and A. Mekhmoukh, "Features based image registration using cross correlation and radon transform," *Alexandria Eng. J.*, vol. 57, no. 4, pp. 2313–2318, 2017.

[10] P.-Y. Chen, C.-C. Huang, C.-Y. Lien, and Y.-H. Tsai, "An efficient hardware implementation of HOG feature extraction for human detection," in *Proc. IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 2, pp. 656–662, Oct. 2013.

[11] S.-B. Chen, Y.-L. Xu, C. H. Q. Ding, and B. Luo, "A nonnegative locally linear knn model for image recognition," *Pattern Recognit.*, vol. 83, pp. 78–90, Nov. 2018.

[12] C.-S. Cho, B.-M. Chung, and M.-J. Park, "Development of real-time vision-based fabric inspection system," *IEEE Trans. Ind. Electron.*, vol. 52, no. 4, pp. 1073–1079, Aug. 2005.

[13] L. Cuimei, Q. Zhiliang, J. Nan, and W. Jianhua, "Human face detection algorithm via Haar cascade classifier combined with three additional classifiers," in *Proc. 13th IEEE Int. Conf. Electron. Meas. Instrum. (ICEMI)*, Oct. 2017, pp. 483–487.

[14] M. Ebrahimi, M. Khoshtaghaza, S. Minaei, and B. Jamshidi, "Vision-based pest detection based on SVM classification method," *Comput. Electron. Agricult.*, vol. 137, pp. 52–58, May 2017.

[15] N. L. Fitriyani, C.-K. Yang, and M. Syafrudin, "Real-time eye state detection system using Haar cascade classifier and circular Hough transform," in *Proc. IEEE 5th Global Conf. Consum. Electron.*, Oct. 2016, pp. 1–3.

[16] P. Goel and S. Agarwal, "Hybrid approach of Haar cascade classifiers and geometrical properties of facial features applied to illumination invariant gender classification system," in *Proc. Int. Conf. Comput. Sci.*, Sep. 2012, pp. 132–136.

[17] C. Grana, D. Borghesani, M. Manfredi, and R. Cucchiara, "A fast approach for integrating ORB descriptors in the bag of words model," *Proc. SPIE*, vol. 8667, Mar. 2013, Art. no. 866709.

[18] H. Hadizadeh, "Multi-resolution local Gabor wavelets binary patterns for gray-scale texture description," *Pattern Recognit. Lett.*, vol. 65, pp. 163–169, Nov. 2015.

[19] Z. He, T. Tan, and Z. Sun, "Topology modeling for AdaBoost-cascade based object detection," *Pattern Recognit. Lett.*, vol. 31, no. 9, pp. 912–919, 2010.

[20] M. Hisham, S. N. Yaakob, R. A. Raof, A. A. Nazren, and N. W. Embedded, "Template matching using sum of squared difference and normalized cross correlation," in *Proc. IEEE Student Conf. Res. Develop.(SCOReD)*, Dec. 2015, pp. 100–104.

[21] C. Iglesias, J. Martínez, and J. Taboada, "Automated vision system for quality inspection of slate slabs," *Comput. Ind.*, vol. 99, pp. 119–129, Aug. 2018.

[22] Y. Iwahori, A. Hattori, Y. Adachi, M. K. Bhuyan, R. J. Woodham, and K. Kasugai, "Automatic detection of polyp using Hessian filter and HOG features," *Procedia Comput. Sci.*, vol. 60, pp. 730–739, Jan. 2015.

[23] D. K. Jain, S. B. Dubey, R. K. Choubey, A. Sinhal, S. K. Arjaria, A. Jain, and H. Wang, "An approach for hyperspectral image classification by optimizing SVM using self organizing map," *J. Comput. Sci.*, vol. 25, pp. 252–259, Mar. 2018.

[24] H. G. Jung, "Analysis of reduced-set construction using image reconstruction from a HOG feature vector," *IET Comput. Vis.*, vol. 11, no. 8, pp. 725–732, 2017.

[25] E. Karami, S. Prasad, and M. Shehata, "Image matching using SIFT, SURF, BRIEF and ORB: Performance comparison for distorted images," 2017, *arXiv:1710.02726*. [Online]. Available: https://arxiv.org/abs/1710.02726

[26] N. Karimimehr and A. A. B. Shirazi, "Fingerprint image enhancement using Gabor wavelet transform," in *Proc. 18th Iranian Conf. Elect. Eng.*, May 2010, pp. 316–320.

[27] J. Kim, J. Baek, and E. Kim, "A novel on-road vehicle detection method using π HOG," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 6, pp. 3414–3429, Dec. 2015.

[28] M. Kim, D. Lee, and K.-Y. Kim, "System architecture for real-time face detection on analog video camera," *Int. J. Distrib. Sensor Netw.*, vol. 11, no. 5, 2015, Art. no. 251386.

[29] D. Konstantinidis, T. Stathaki, V. Argyriou, and N. Grammalidis, "Building detection using enhanced HOG–LBP features and region refinement processes," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 3, pp. 888–905, Mar. 2017.

[30] H. Lahiani and M. Neji, "Hand gesture recognition method based on HOG-LBP features for mobile devices," *Procedia Comput. Sci.*, vol. 126, pp. 254–263, Jan. 2018.

[31] U. Lall and A. Sharma, "A nearest neighbor bootstrap for resampling hydrologic time series," *Water Resour. Res.*, vol. 32, no. 3, pp. 679–693, 1996.

[32] Y. Lei, Y. Wang, and J. Huang, "Robust image hash in Radon transform domain for authentication," *Signal Process., Image Commun.*, vol. 26, no. 6, pp. 280–288, 2011.

[33] Y. Lei, L. Zheng, and J. Huang, "Geometric invariant features in the radon transform domain for near-duplicate image detection," *Pattern Recognit.*, vol. 47, no. 11, pp. 3630–3640, 2014.

[34] C. Li, G. Duan, and F. Zhong, "Rotation invariant texture retrieval considering the scale dependence of Gabor wavelet," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2344–2354, Aug. 2015.

[35] C. Li, Y. Huang, and L. Zhu, "Color texture image retrieval based on Gaussian copula models of Gabor wavelets," *Pattern Recognit.*, vol. 64, pp. 118–129, Apr. 2017.

[36] D. P. K. Lun, T. C. L. Chan, T.-C. Hsung, D. Feng, and Y.-H. Chan, "Efficient blind image restoration using discrete periodic Radon transform," *IEEE Trans. Image Process.*, vol. 13, no. 2, pp. 188–200, Feb. 2004.

[37] T. Mantoro and M. A. Ayu, "Multi-faces recognition process using Haar cascades and eigenface methods," in *Proc. 6th Int. Conf. Multimedia Comput. Syst. (ICMCS)*, May 2018, pp. 1–5.

[38] J. Mohammadi and R. Akbari, "Vehicle speed estimation based on the image motion blur using radon transform," in *Proc. 2nd Int. Conf. Signal Process. Syst.*, vol. 1, Jul. 2010, pp. V1–V243.

[39] M. A. Mohammed, M. K. A. Ghani, N. Arunkumar, R. I. Hamed, M. K. Abdullah, and M. Burhanuddin, "A real time computer aided object detection of nasopharyngeal carcinoma using genetic algorithm and artificial neural network based on Haar feature fear," *Future Gener. Comput. Syst.*, vol. 89, pp. 539–547, Dec. 2018.

[40] M. Mozgovoy and E. Pyshkin, "Using image recognition for testing hand-drawn graphic user interfaces," in *Proc. 11th Int. Conf. Mobile Ubiquitous Comput., Syst., Services Technol. (UBICOMM)*, 2017, pp. 1–4.
[41] J. N. Myhre, K. O. Mikalsen, S. Løkse, and R. Jenssen, "Robust clustering using a knn mode seeking ensemble," *Pattern Recognit.*, vol. 76, pp. 491–505, Apr. 2018.
[42] N. Nabizadeh and M. Kubat, "Brain tumors detection and segmentation in MR images: Gabor wavelet vs. statistical features," *J. Comput. Elect. Eng.*, vol. 45, pp. 286–301, Jul. 2015.
[43] N. Nacereddine, S. Tabbone, and D. Ziou, "Similarity transformation parameters recovery based on radon transform. Application in image registration and object recognition," *Pattern Recognit.*, vol. 48, no. 7, pp. 2227–2240, 2015.
[44] G. Peng, Z. Zhang, and W. Li, "Computer vision algorithm for measurement and inspection of O-rings," *Measurement*, vol. 94, pp. 828–836, Dec. 2016.
[45] A. Radman, N. Zainal, and S. A. Suandi, "Automated segmentation of iris images acquired in an unconstrained environment using HOG-SVM and GrowCut," *Digit. Signal Process.*, vol. 64, pp. 60–70, May 2017.
[46] U. Raghavendra, U. R. Acharya, H. Fujita, A. Gudigar, J. H. Tan, and S. Chokkadi, "Application of Gabor wavelet and locality sensitive discriminant analysis for automated identification of breast cancer using digitized mammogram images," *Appl. Soft Comput.*, vol. 46, pp. 151–161, Sep. 2016.
[47] S. Routray, A. K. Ray, C. Mishra, and G. Palai, "Efficient hybrid image denoising scheme based on SVM classification," *Optik*, vol. 157, pp. 503–511, Mar. 2018.
[48] M. S. S. Sajadi, M. Babaie, and M. Bahaghighat, "Design and implementation of fuzzy supervisor controller on optimized DC machine driver," in *Proc. 8th Conf. AI Robot. 10th RoboCup Iranopen Int. Symp. (IRANOPEN)*, Apr. 2018, pp. 26–31.
[49] K. Seemanthini and S. S. Manjunath, "Human detection and tracking using HOG for action recognition," *Procedia Comput. Sci.*, vol. 132, pp. 1317–1326, Jan. 2018.
[50] C. H. Setjo and B. Achmad, "Thermal image human detection using Haar-cascade classifier," in *Proc. 7th Int. Annu. Eng. Seminar (InAES)*, Aug. 2017, pp. 1–6.
[51] H. Sharif and M. Hölzel, "A comparison of prefilters in ORB-based object detection," *Pattern Recognit. Lett.*, vol. 93, pp. 154–161, Jul. 2017.
[52] A. Soudani and W. Barhoumi, "An image-based segmentation recommender using crowdsourcing and transfer learning for skin lesion extraction," *Expert Syst. Appl.*, vol. 118, pp. 400–410, Mar. 2019.
[53] H. Tan, Z. Ma, and B. Yang, "Face recognition based on the fusion of global and local HOG features of face images," *IET Comput. Vis.*, vol. 8, no. 3, pp. 224–234, Oct. 2013.
[54] S. Tuermer, F. Kurz, P. Reinartz, and U. Stilla, "Airborne vehicle detection in dense urban areas using HoG features and disparity maps," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 6, pp. 2327–2337, Dec. 2013.
[55] Tzutalin. (2015). *Labelimg is a Graphical Image Annotation Tool and Label Object Bounding Boxes in Images*. [Online]. Available: https://github.com/tzutalin/labelImg
[56] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004.
[57] H. Wang, J. Wang, W. Chen, and L. Xu, "Automatic illumination planning for robot vision inspection system," *Neurocomputing*, vol. 275, pp. 19–28, Jan. 2018.
[58] R. Wang, W. Li, R. Li, and L. Zhang, "Automatic blur type classification via ensemble SVM," *Signal Process., Image Commun.*, vol. 71, pp. 24–35, Feb. 2019.
[59] W.-C. Wang, S.-L. Chen, L.-B. Chen, and W.-J. Chang, "A machine vision based automatic optical inspection system for measuring drilling quality of printed circuit boards," *IEEE Access*, vol. 5, pp. 10817–10833, 2017.
[60] S. Yao, S. Pan, T. Wang, C. Zheng, W. Shen, and Y. Chong, "A new pedestrian detection method based on combined HOG and LSS features," *Neurocomputing*, vol. 151, pp. 1006–1014, Mar. 2015.
[61] H. Zhang, X. Li, H. Zhong, Y. Yang, Q. J. Wu, J. Ge, and Y. Wang, "Automated machine vision system for liquid particle inspection of pharmaceutical injection," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 6, pp. 1278–1297, Jun. 2018.
[62] S. Zhang, X. Li, M. Zong, X. Zhu, and R. Wang, "Efficient kNN classification with different numbers of nearest neighbors," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 5, pp. 1774–1785, May 2018.
[63] Y. Zhao, J. Gu, C. Liu, S. Han, Y. Gao, and Q. Hu, "License plate location based on Haar-like cascade classifiers and edges," in *Proc. 2nd WRI Global Congr. Intell. Syst.*, vol. 3, Dec. 2010, pp. 102–105.

MAHDI BAHAGHIGHAT received the Ph.D. from the Electrical Engineering Department, Amirkabir University of Technology (AUT), in 2017. He is currently an Assistant Professor and the Chair of the Electrical Engineering Group, Raja University. He is also the Head of the Artificial Intelligence in Science and Technology (AIST) Lab. His current research interests include signal, image and video processing, computer vision, artificial intelligence, machine learning, deep learning, sensor networks, and wireless multimedia transmission.

LEILA AKBARI received the M.Sc. degree in electrical engineering from Qazvin Azad University (QIAU). She is currently a Lecturer with Raja University, and she is also a Researcher with the Artificial Intelligence in Science and Technology (AIST) Lab. Her research interests include computer vision and image processing.

QIN XIN received the Ph.D. degree from the University of Liverpool, U.K., in 2004. He is currently a Full Professor of computer science with the Faculty of Science and Technology, University of the Faroe Islands (UoFI), Faroe Islands. Prior to joining UoFI, he held various research positions in world leading universities and research laboratories including a Senior Research Fellow of the Université Catholique de Louvain, Belgium, a Research Scientist/Postdoctoral Research Fellow of the Simula Research Laboratory, Norway, and the University of Bergen, Norway. Moreover, he also investigates the combinatorial optimization problems with applications in bioinformatics, data mining, and space research. He is currently serving on the Management Committee Board of Denmark for several EU ICT projects. He has published over 70 peer-reviewed scientific papers. His works have been published in leading international conferences such as ICALP, ACM PODC, SWAT, the IEEE MASS, ISAAC, SIROCCO, and the IEEE ICC. His works have also been published in international journals including *Algorithmica*, *Theoretical Computer Science*, *Distributed Computing*, the IEEE TRANSACTIONS ON COMPUTERS, the *Journal of Parallel and Distributed Computing*, the IEEE TRANSACTIONS ON DIELECTRICS AND ELECTRICAL INSULATION, and *Advances in Space Research*. His main research interests include the design and the analysis of sequential, parallel, and distributed algorithms for various communication and optimization problems in wireless communication networks, and cryptography and digital currencies including quantum money. He has been very actively involved in the services for the community in terms of acting (or acted) on various positions, including the session chair, a member of the technical program committee, the symposium organizer, and the local organization co-chair, for numerous international leading conferences in the fields of distributed computing, wireless communications, and ubiquitous intelligence and computing, including the IEEE MASS, the IEEE LCN, ACM SAC, the IEEE ICC, the IEEE GLOBECOM, the IEEE WCNC, the IEEE VTC, IFIP NPC, and the IEEE Sarnoff. He is the Organizing Committee Chair of the 17th Scandinavian Symposium and Workshops on Algorithm Theory (SWAT 2020, Torshavn, Faroe Islands). He currently serves on the editorial board for more than ten international journals.