Arabian Journal for Science and Engineering (2021) 46:4049–4058
https://doi.org/10.1007/s13369-021-05401-4

RESEARCH ARTICLE-COMPUTER ENGINEERING AND COMPUTER SCIENCE

Adaptive Technique for Brightness Enhancement of Automated Knife Detection in Surveillance Video with Deep Learning

Mai K. Galab[1] · Ahmed Taha[1] · Hala H. Zayed[1]

Received: 3 August 2020 / Accepted: 20 January 2021 / Published online: 8 February 2021

Abstract
Detecting knives in surveillance videos are very urgent for public safety. In general, the research in identifying dangerous weapons is relatively new. Knife detection is a very challenging task because knives vary in size and shape. Besides, it easily reflects lights that reduce the visibility of knives in a video sequence. The reflection of light on the surface of the knife and the brightness on its surface makes the detection process extremely difficult, even impossible. This paper presents an adaptive technique for brightness enhancement of knife detection in surveillance systems. This technique overcomes the brightness problem that faces the steel weapons and improves the knife detection process. It suggests an automatic threshold to assess the level of frame brightness. Depending on this threshold, the proposed technique determines if the frame needs to enhance its brightness or not. Experimental results verify the efficiency of the proposed technique in detecting knives using the deep transfer learning approach. Moreover, the most four famous models of deep convolutional neural networks are tested to select the best in detecting knives. Finally, a comparison is made with the-state-of-the-art techniques, and the proposed technique proved its superiority.

Keywords Knife detection · Smart video surveillance · Deep neural network · CNN · Weapon detection

1 Introduction

Public safety is one of the significant burdens of society today. Nowadays, a lot of researchers paid considerable attention to smart video surveillance system (SVSS) as it empowers governments to trace all possible threats for maintaining public safety. SVSS is the use of automatic video analytics to enhance the effectiveness of traditional surveillance systems and avoid their drawbacks. The weapon in human hands in public places forms a severe threat to the safety and the security of people in their daily life. Recently, the world has witnessed a notable increase in acts of violence. The presence of efficient ways to avoid such behavior is a critical necessity. Automatic detection of weapons in images taken with surveillance cameras is one of these effective ways. The research in the area of detecting dangerous weapons is comparatively recent. Detecting a person holding weapons like firearms or any sharp object like knives in public or restricted areas is a robust indicator of potentially dangerous action. Such incidents have led to a need for automatic weapons detection systems. However, until now, the number of techniques presented to detect weapons from surveillance cameras is not widely available.

The principle of analysis and detection of dangerous situations by automated closed-circuit television (CCTV) has been raised in several studies and many domains. The research in the field of weaponry detection can be branched into the detection of firearms, and the detection of knives, as shown in Fig. 1.

The first category is more fortunate. Many techniques are developed using traditional image feature extraction techniques. These non-deep learning algorithms are based primarily on shapes, color segmentation, interesting point, and edge detectors. Some related work includes: using the canny edge detection technique in detecting weapons [1], employing SURF in gun detection [2], adopting a combination of SIFT and Harris interest point detector in detecting guns

✉ Mai K. Galab
may.gallab@fci.bu.edu.eg

Ahmed Taha
ahmed.taha@fci.bu.edu.eg

Hala H. Zayed
hala.zayed@fci.bu.edu.eg

1 Faculty of Computers & Artificial Intelligence, Benha University, Benha, Egypt

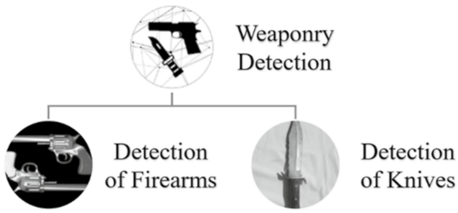

Fig. 1 Weapons detection types

Fig. 2 Checking the existence of a hidden weapon. Left: The weapon detected on a person`s luggage. Right: The weapon detected on a person`s body

[3], using the background subtraction technique [4], and exploiting the ontological nodes [5]. In many places such as airports, malls, stadiums, and schools, other hardware solutions are designed to detect weapons concealed in jackets, pockets, shoes, luggage, … etc. Concealed weapon detection (CWD) technologies are based on an electro-magnetic scan, millimetric wave, and X-ray as in [6–8], and [9] (see Fig. 2).

Gun detection is the pivot point of the most recent researches. They produced reasonably satisfactory results. However, the research in knife detection is rather new and little-studied [10, 11] and [12]. The wide variations in the size and shape of knives affect the detection results. Furthermore, the light reflection and brightness that manifests on the surface of the knife is a very challenging problem. It profoundly disturbs the detection process. The metal surface of the knife causes critical issues during the detection process. Changing the lighting on the metal surface and the brightness on the knife surface destroys the knife shape in the image. The small size of the knife and the distance from the camera make the detection difficult. Hence, the detection of knives from CCTV images needs more attention from researchers.

In this paper, an adaptive technique for eliminating the brightness problem in knife detection is proposed. To evaluate the brightness level, an automatic threshold is first calculated for each frame. Accordingly, the proposed technique employs gamma correction (GC) to enhance those frames with a brightness level greater than the generated threshold. A knife detection system is built using a deep learning (DL) approach. The system utilizes DL as a feature extractor as well as a classifier and enhances the results by applying transfer learning (TL). The proposed threshold succeeded in filtering the frames having a brightness problem. Also, we analyzed four different models of convolutional neural network (CNN), including AlexNet, GoogLeNet, VGGNet, and ResNet pre-trained networks, to determine the best one with the knife detection process. AlexNet proved its superiority.

In Sect. 2, the related work is reviewed. Section 3 presents the proposed technique. Section 4 draws a detailed picture of the experimental results. Finally, Sect. 5 offers conclusions and future directions.

2 Related Work

Automatic knife detection in images is rather new and little-studied compared to other weapon detection techniques. In the literature, the work of weapons detections can be classified as deep learning techniques and non-deep learning techniques.

2.1 Non-Deep Learning Techniques

Zywicki et al. [10] present a technique to detect the knife using Haar cascades. Kmieć and Glowacz in [11] use the histograms of oriented gradients (HOG) as a feature descriptor. This method evaluates the performance using the HOG-based SVM classifier. Moreover, Kmieć et al. [12] used a technique that detects a knife using the active appearance model (AAMs). His system depends mainly on the appearance and shape of the knife. It initializes the knife position in the first frame. Then, it locates the object on the next frames using optical flow.

However, these techniques suffer from some drawbacks. They need intensive processing. Therefore, they are not suitable for real-life situations. Besides, these techniques require the knife object to be entirely clear in the scene. Also, these techniques do not achieve satisfactory results unless the images used are of high quality. Some attempts are made to become more suitable for real-time situations and to work in open areas. One of these attempts is what is presented by Maksimove [13]. He presents a formal model of knife object representation by a pair of second-degree curves. The knife is detected depending on the angle between the two estimated curves. The basic characteristic of the geometric form of a knife is the angle between the blade and the butt of the knife. The fuzzy and possibilistic shell clustering methods are applied in this paper to generate the contours of objects in the image. Nevertheless, this method is highly sensitive to lighting conditions, equipment quality, and camera resolution. Also, giving a good result requires high complexity. Glowacz et al., [14] also present an approach that uses AAMs to detect knives in images. In their work, the knife blade has a particular interest point. This method trained 24 AAMs to cover all the variations in a knife orientation. His method used forty positive test set and forty negative test set for detector evaluation. However, this approach is only suited to applications where the knife tip is visible, such as in luggage-scanning systems.

In general, methods presented by Maksimova [13] and Glowacz [14] process the image pixel by pixel, which is unsuitable for real-time applications. The researchers in [15] dissolve the problem of knife detection in images by selecting MPEG-7 descriptors to find feature vectors. The underlying rationale behind their approach is testing the fuzzy classification with the generated feature vectors. The described method is experimentally validated on a dataset with over 12 thousand images. About 15% of the obtained results are miss labeled. It is a good result for describing the problem, but for a real system, this result needs to be improved. Kmieć combines the DED detector and the HOG detector in [16]. He assumes that the knife is centered vertically in the detection window, and he finds the dominant edges. Kmieć creates a feature vector dependent on the intensities of the dominant edges and the straight-line segments. Then, SVM is utilized to classify the image as a knife.

Furthermore, the authors in [17] design an algorithm for detecting the knife based on machine learning and visual descriptors. Their algorithm applies a sliding window technique as the first step to crop the interesting part. Then, it works only with knives close to the human silhouette. Finally, it determines the part of the image in which a hand holds the knife. In [18], the background subtraction technique and HOG are selected as a feature extractor. A neural network is used in the classification step. The performance of this neural network to detect a knife was calculated using tenfold cross-validation on the complete dataset. The work in [19] presents an object detection algorithm to detect a knife from video data. This approach is tolerant to the rotation, and change in scale and pose. It has three stages: foreground segmentation, Features from Accelerated Segment Test (FAST) based salient feature detection for image localization, and Multi-Resolution Analysis (MRA) for classification, and confirmation of the target.

2.2 Deep Learning Techniques

Deep learning (DL) is one of the best techniques that automatically extract features from raw data [20]. There are many architectures of DL as shown in Fig. 3 [21]. It has demonstrated its robust ability to learn and represent the features when compared with traditional object detection techniques [22].

With DL, it is smoother than ever to discover objects [21]. In the presented work, CNN is selected as a DL model. CNN has proved its tremendous success in image

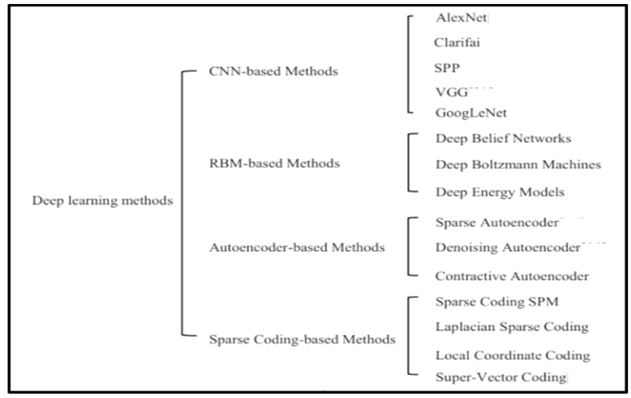

Fig. 3 Categorization of the deep learning methods and their representative works [21]

recognition. It can learn to represent the features from raw data automatically. It is a specific type of multi-layered network, designed to recognize features directly from images with minimal pre-processing. The CNN architecture is built of a single input layer, a single output layer, and multiple hidden layers. [23].

Pre-trained CNN models achieved a massive breakthrough and succeeded in winning the annual challenges of ImageNet a Large-Scale Visual Recognition Competition (ILSVRC). These models are shown in Table 1.

Using the DL approach in knife detection is still new. Here is a list of the algorithms that use DL in detecting weapons. Olmos et al. [24] used DL in detecting firearms. Their work is considered the leading DL solution in gun detection. This method uses DL as a classifier on their datasets within a sliding window and region proposals detection-based methods. This method achieved recall equal to 100% and precision equal to 84.21%. Verma et al. [25] used CNN to detect weapons especially the handheld guns from the cluttered scene automatically. They used a deep convolutional network (DCN). In the detection process, the faster region-based CNN model (Faster R-CNN) uses through TL. This gun detection approach was evaluated via the IMFDB benchmark weapons database. It achieved 93% accuracy.

Al-Shoukry [26] presented an automated hybrid approach to detect concealed weapons. This approach obtained fused imagery using a traditional discrete wavelet transform (DWT) with hybrid bag-of-words (BOF). Then, he utilized CNN with a pre-trained CNN model to train a multiclass SVM classifier. This approach employed the DL in CWD Applications. It works well with X-ray images.

Moreover, Castillo et al. claimed that his work in [27] is the first one that uses DL to detect knives. He develops an automatic model for detecting cold steel weapons for video surveillance using CNN. He tries to dissolve the problem of light conditions by applying a brightness guided

Table 1 The architectures of the pre-trained CNN models [23]

Year	CNN	Developed By	Top-5 error rate	No. of parameters
1998	LeNet	Yann LeCun et al.,	–	60 thousand
2012	AlexNet	Alex Krizhevsky et al.,	15.3%	60 million
2013	ZFNet	Matthew Zeiller and Rob Fergus	14.8%	–
2014	GoogLeNet	Google	6.67%	4 million
2014	VGGNet	Simoryan, Zisseman	7.3%	138 million
2015	ResNet	Kaiming He	3.6%	–

pre-processing procedure called Darkening and Contrast at Learning and Test stages (DaCoLT).

Finally, Noever in [28] introduced methods for solving the knife threat identification as both a classification and detection problem using DL network. This method depends on the TL. It achieves accuracy greater than 95%.

Generally, CNN has many benefits [29]. It is easier to train. CNN is weight sharing. It gives a beneficial and accurate result with fewer parameters compared to other networks with an equal number of hidden layers. Using CNN leads to time-saving. It preserves the memory requirements and computation complexity requirements. Furthermore, CNN provides better performance for applications where the inputs have a local correlation (e.g., image). CNN not only gives superior accuracy results compared to other detection algorithms but also, it even exceeds humans in some cases.

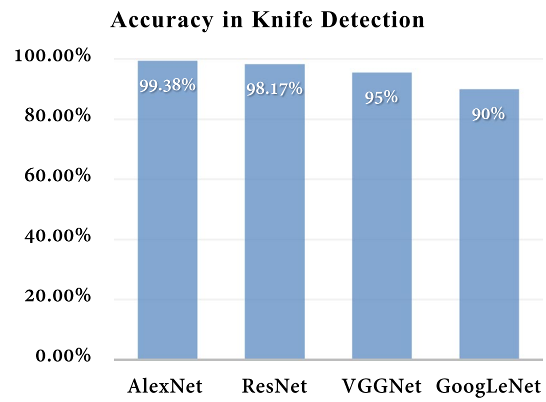

Fig. 4 Accuracy of knife detection with the most popular pre-trained CNN models

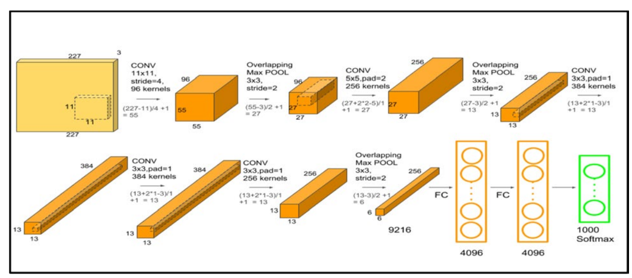

Fig. 5 AlexNet Architecture [30]

3 The Proposed Technique

One of the most severe and vital issues that significantly affect the performance of detection in surveillance systems is the change in the brightness of the captured images. To reduce the effect of reflecting light on the knife's surface, we introduce a technique that significantly enhances the visual features of the knife image in the scene.

The proposed technique makes a comparative evaluation of the most popular pre-trained networks commonly used for recognition problems. The four deep CNN models are, AlexNet, GoogLeNet, VGGNet, and ResNet to pick the pre-trained model with the highest accuracy in detecting knives. We tested these four CNN models with a large knife dataset images gathering positive and negative samples. After a series of experiments with these CNN models, AlexNet proved its superiority (see Fig. 4). So, AlexNet has chosen to build the proposed knife detection technique because it achieves the highest accuracy.

AlexNet architecture is employed (see Fig. 5). It consists of eight layers: five convolutional layers followed by three fully connected (FC) layers. FC layers are the final layers of CNN that determine the result of the intended output. It sums the weights of the previous feature layers, indicating the right mix of "ingredients". The final FC generates 4096 neurons. The final step in CNN reduces the features vector from 4096 to 1000, for the nonlinear part.

3.1 Knife Detection Technique

Figure 6 shows the block diagram of the presented knife detection technique. It consists of six steps. Step 1 begins by applying pre-processing to the frames captured from the CCTV camera to make the frame suitable for the selected CNN. The images collected from the raw input data are resized to 227 × 227 to be suitable to work with AlexNet. In step 2, TL is applied after dividing the dataset into two groups: the first group is the learning set and the second group is the test set:

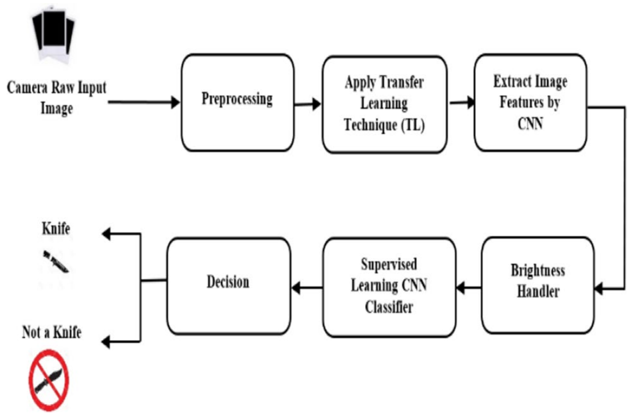

Fig. 6 Block diagram of the proposed technique

1. Learn group: used for creating the model in the learning mode. The Learn group contains 70% of the dataset (training images).
2. Test group: used for estimating the classification quality in the test mode. The Test group contains 30% of the dataset (testing images).

Transfer learning (TL) is applied in the proposed technique. To accelerate the training process and improve the training results, TL is used to avoid DL defects. TL is the task of using the knowledge provided by a pre-trained network to learn new patterns in new data [31]. The pre-trained network with TL is typically much more comfortable and faster than that trained from scratch. Using pre-trained deep networks enables the system to quickly learn new tasks without defining and training a new network, having millions of images, or need long training times. The advantage of TL is that the pre-trained network, which has already been learned before, is being used in similar work [32].

The original pre-trained network is trained on a subgroup of ImageNet with 1000 objects. The proposed technique transfers the learned knowledge to solve the problem with the small private dataset (classifying the object as a knife or a not-knife). The length of the output layer in AlexNet is not equal to the number of classes in the presented work. So, we modify the structure of AlexNet to make the final FC layer classify only our two classes (knife or not-knife) instead of a thousand classes.

Step 3 finds significant features using CNN. It applies a sequence of the convolutional layers and pooling layers. The classifier is composed of FC layers.

Step 4 tests whether the frame exceeds the brightness level range or not. If the frame brightness exceeds this level, it must be improved by applying our "Brightness Handler Procedure" (BHp). Otherwise, go directly to the next step. This step will be explained in detail in the next subsection (Sect. 3.2). Steps 5 and 6 classify the images and decide whether it is a knife or not.

3.2 Brightness Handler Procedure (BHp)

We proposed Brightness Handler Procedure (BHp) to improve the capability of the proposed technique to overcome the brightness problems in images. The presented Brightness Handler Procedure (BHp) determines the brightness value of the video frames based on an Automatic Brightness Threshold (ABT). According to the generated value of the ABT threshold, we decide if the frame has a *brightness and needs to be enhanced* or not. ABT rule expressed as:

$$\mathrm{ABT} = \frac{1}{3}(\beta + Avg_{Otsu} + Avg_{Brightness)} \tag{1}$$

Where

$$\beta = \frac{\alpha}{3}. \tag{2}$$

$$\alpha = \frac{-0 \cdot 3}{\log_{10}\left(Avg_{Brightness}\right)}. \tag{3}$$

$$Avg_{Brightness} = \frac{1}{N}\sum_{i=1}^{N} frame_{Brightnesvalue}(i). \tag{4}$$

$$Avg_{Otsu} = \frac{1}{N}\sum_{i=1}^{N} frame_{Otsuvalue}(i). \tag{5}$$

Algorithm 1: Procedure for Correcting Image Brightness

Input: ƑRGB ϵ V, ƑRGB is RGB image & V is a video with consecutive frames.

Output: The Brightness-Enhanced Image.

1: Convert the RGB frame "ƑRGB" into a grayscale image "ƑGry",

2: Calculate the value of the presented Automatic Brightness Threshold "ABT" for all "ƑGry",

3: Determine the brightness level "BL" of "ƑGry"

4: IF BL > ABT THEN

Apply Enhancement on "ƑGry" by using the GC technique

ELSE

Continue

ENDIF

The proposed BHp consists of three stages. The first stage converts the colored video frames into grayscale images. The second stage generates the value of the proposed Automatic Brightness Threshold (ABT) of all frames. The final stage checks the brightness level (BL) value of each video frame with the proposed generated threshold. If the BL of the frame is smaller than the ABT value, it moves to the next frame. Otherwise, it enhances the contrast of the acquired frame by using the GC technique (see Algorithm 1). If the ABT value is lower than the brightness level of the frame, the frame is darkened by applying GC [33] (Eq. 6).

$$P_{out} = CP_{in}^{\gamma}. \tag{6}$$

where C is a constant, usually equal to 1 [34], P_{in} is the old pixel, P_{out} is the new pixel, and γ is a gamma value. The value of gamma must be greater than 1 to darken the image [34]. In our experiments, the gamma is empirically equal to 6.

There are many darkening techniques like histogram equalization (HE) and its related methods such as contrast stretching, and contrast limited adaptive histogram equalization (CLAHE). By experiments, the GC technique achieved the best result [33].

HE redistributes the pixel intensity values evenly by using a cumulative (sum) histogram. HE is effective when both the background and the objects are dark or bright. However, the contrast of the images with both under-exposed and over-exposed regions cannot be improved effectively. To overcome the shortcomings of the standard HE method, CLAHE was developed [35–37]. CLAHE is quite complex and costly. It also works only in a small region. GC is also a popular pixel-domain contrast enhancement (CE) method. It is a transform-based histogram modification technique that processes a varying parameter γ (Gamma). It is a cost-effective and great way to handle frames that are bright and dim.

4 Experimental Results

Performance evaluation is a very crucial task after the development process. It is significant to show that the final proposed technique achieves an acceptable level of performance and that it represents a considerable improvement over the existing techniques or not. The proposed technique has been implemented by a MATLAB 2017b. This implementation is running under a 64-bit Windows 10 operating system on a computer with the Intel Core i7 processor and 16 GB RAM. It executes on the CPU.

4.1 Dataset

In this paper, four different datasets are used to test the performance of the proposed technique. Matiolanski, et al., and Grega, et al. Dataset [38] is considered to be a benchmark knife detection dataset. It was published in 2016. It is widely used to evaluate this kind of algorithms, such as [13, 15, 17, 18, 28]. The dataset is obtained from CCTV recordings. The images are cropped from the original frames and are set to 100 × 100 pixels. It is often of poor quality, blurred, and the object is rather small. Figure 7 shows some samples of this dataset. The dataset contains two picture groups (positive examples and negative examples). Positive samples are photographs of a human hand holding a knife, while negative samples are images without knives. The whole dataset consists of 12,899 colored images divided into 9340 Negative Examples Images (NEI) and 3559 Positive Examples Images (PEI). Some of these images were captured indoors, while others were captured in the street through a car.

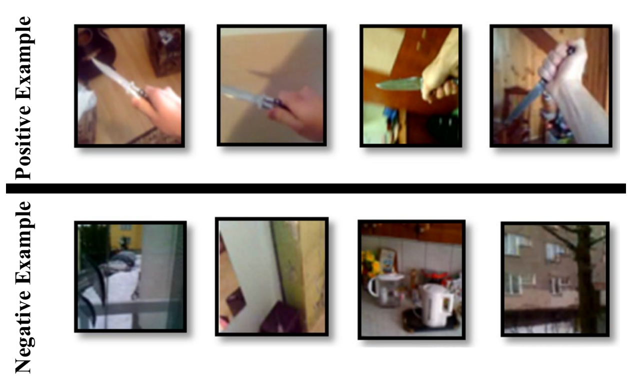

Fig. 7 Example images of the dataset, The Top Row: represents positive examples, and The Bottom Row: represents negative examples

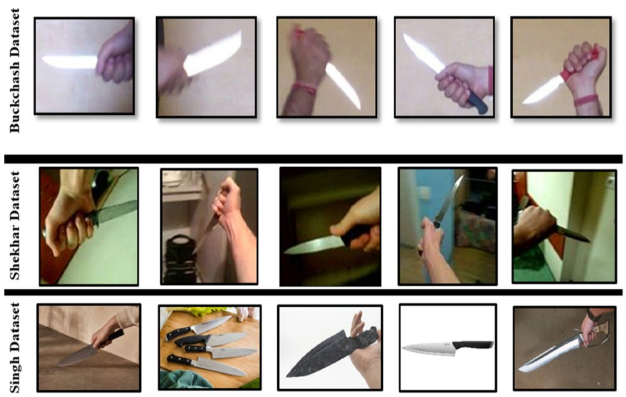

Fig. 8 Example of images of the newly published datasets with different brightness conditions. The Top Row: Buckchash's dataset, 2017. The Middle Row: Shekhar's dataset, 2020. The Bottom Row: Singh's dataset, 2020

Recently, some authors have released further knife datasets. Buckchash dataset [39] was created in 2017. This dataset contains 164 images of hand knives for positive samples, 123 images of hand-only, for negative samples, and videos with different brightness conditions (low, high, and normal). All of these videos present the same person in the same place. He repeats the same action but with different positions of the knife. Moreover, the Shekhar dataset [40] was published in 2020. This dataset contains 400 training images and 100 testing images of knives. Besides, the Singh dataset [41] was also created in 2020. Shekhar dataset and Singh dataset have images less than one thousand examples. The last three datasets images have different brightness conditions. Some samples of these last three datasets are shown in Fig. 8.

4.2 Experiments

For evaluating the final proposed technique, we use the metrics that are widely used in the evaluation process of this kind of application. These metrics [42] include true negative rate (TNR), true positive rate (TPR), positive predictive rate (PPV), false omission rate (FOR), recall (sensitivity), specificity, F1-score, and accuracy. All these metrics mainly depend on calculating the values of true positive (TP), false positive (FP), false negative (FN), and true negative (TN).

A series of experiments are implemented in the presented work to precisely evaluate the performance of the proposed technique. Two types of experiments are conducted. The first set of experiments is performed without applying the proposed technique to correct the frame brightness (BHp). This means the knife detection is performed without any enhancement to the input frames. The second type of experiment is conducted by applying our proposed brightness adaptive technique (BHp).

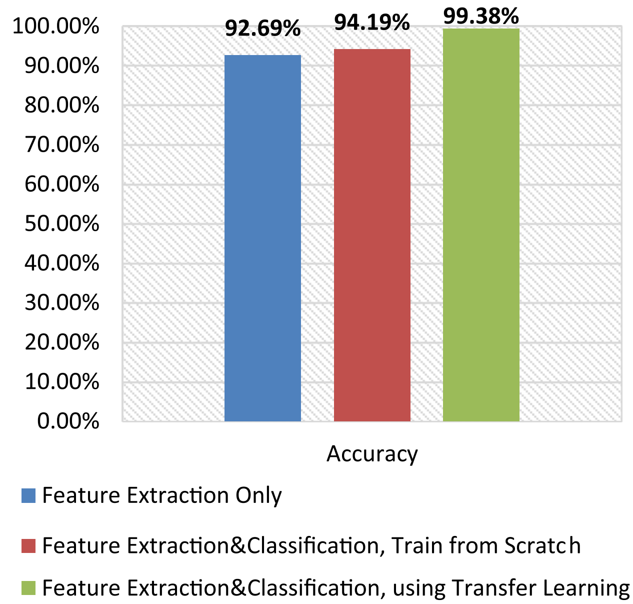

Fig. 9 Accuracy of the first set of experiments

4.2.1 Deep Learning Evaluations

In the first set of experiments, we employ DL once as a feature extractor, again as a classifier, and a third time both. The achieved accuracy was 92.69%, 94.19%, and 99.38% for each experiment, respectively. As can be seen, using DL as a feature extractor and, at the same time, as a classifier, achieves the highest accuracy in detecting knives. It uses fine-tuning to train the last fully connected layer of the network. The use of a TL was essential in enhancing accuracy, as shown in Fig. 9. Detailed results of the final experiment are shown in detail in Table 2.

These results are achieved with the normal cases of light in images. The brightness problem in images greatly affects the efficiency of knife detection techniques. Although we use deep learning in the detection process, it did not give robust results with the problems of light variation in images. Using deep learning alone is not enough. So, we develop the proposed Brightness Handler Procedure (BHp) with DL to treat the problem of the knife brightness.

Table 2 The proposed technique results with the evaluation metrics

Measurement	Result (%)
TPR	99.57
TNR	98.87
PPV	99.57
FOR	98.87
Recall	99.5
Specificity	98.87
Accuracy	99.38

Table 3 Summarized the accuracy of the proposed BHp technique on the different previous datasets

Dataset	Accuracy (%)
Buckchash 2017 [39]	96.95
Shekhar 2020 [40]	98.46
Singh 2020 [41]	90.59

4.2.2 Brightness Handler Procedure (BHp) Evaluation

In the second set of experiments, the proposed Brightness Handler Procedure (BHp) is evaluated. It is developed to handle the problem of the knife brightness. The results with the normal light case give accuracy equal to 99.38%. With high brightness, the accuracy is decreased to 84.9% with DL without applying the proposed Brightness Handler Procedure (BHp). This is because the high brightness destroys the image. However, when the proposed BHp is used to enhance the frame brightness, the accuracy is increased to 96.95%, the F1-score value is improved from 90.3% to 98.42%, and the FN value is reduced from 15 to 3%.

To confirm the efficiency of the Brightness Handler Procedure (BHp), the proposed technique is tested with the knife datasets that suffer from high brightness conditions. Table 3 reports the accuracy of the proposed technique on these datasets. As the results show, the proposed technique with BHp proves its ability to detect knife images and achieves great accuracy results.

4.2.3 Comparative Results:

To verify the performance of the proposed technique with the proposed BHp, it is compared with some related work for knife detection, including [13, 15, 17], and [18]. All the methods are tested using the same dataset (Matiolanski, et al., and Grega, et al., Dataset). Table 4 summarizes the overall comparative results. As noticed from the table, the methods presented in [13, 15, 17], and [18] achieve accuracy equal to 79%, 77%, 91%, and 84.6% respectively. However, the proposed technique with the proposed BHp achieves an accuracy that is superior to all of them. This proves that the proposed technique is superior to its peers.

Furthermore, we compared the proposed technique with recent deep learning-based methods [27] and [28]. In [28], Noever applies DL to detect the knife. His method achieves accuracy greater than 95% with the test data. However, his technique does not address the brightness problem. Table 5 shows the comparative results of the proposed technique with the proposed BHp and Noever's technique. As the results indicate, the proposed technique is superior to Noever's method when tested on the same dataset that was published in 2016 by Matiolanski, et al. and Grega, et al. [38] and that published in 2020 by Shekhar,2020 [40].

On the other hand, Castillo [27] addresses the subject of brightness and its effect on the detection process. He used his procedure (DaCoLT) to solve the brightness problem. DaCoLT is mainly based on CLAHE, which limits contrast, and enhances the low contrast images successfully. It is quite complex and costly. It also works only in a small region. Castillo method achieves a recall equal to 78.55%, a precision equal to 100%, and an F1 score equal to 87.74%. In the proposed technique, we used our Brightness Handler Procedure (BHp) which is based on GC rather than CLAHE. GC is a well-known contrast enhancement tool for the pixel domain because it is cost-effective and handles bright and blurred images. It produced better results during the test. Table 6 reports the comparative results.

From the reported results in Table 5 and Table 6, it is observed that the accuracy of the proposed technique with

Table 5 Comparing the experimental results of the proposed technique and other methods that use CNN at the same Datasets

Dataset	Accuracy	
	The proposed technique	Noever 2020
Matiolanski, et al. and Grega, et al. Dataset [38]	99.38%	95%
Shekhar, Dataset [40]	98.46%	98%

Table 6 Comparing the experimental results of the proposed technique and other methods that use CNN

Methods	Accuracy	Precision	Recall	F1-score
Castillo 2019[27]	*N/A*	100%	78.55%	87.74%
The proposed technique with a brightness procedure	96.95%	100%	96.8%	98.42%

Table 4 Comparative Results of The Proposed Technique with The Traditional Methods that used The Same Data Set in [40]

	The proposed technique (%)	Vajhala et al., (2016) [18] (%)	Grega et al., (2016) [17](%)	Maksimove et al., (2014) [15]	Maksimove, (2013) [13]
Specificity	**98.87**	85.7	94.9	N/A	N/A
Sensitivity	**99.5**	80.4	81.8	N/A	N/A
Accuracy	**99.38**	84.6	91	77%	79%

the proposed BHp overcomes other similar methods that use DL. The proposed technique with the proposed BHp is based on AlexNet, while the other methods use other pre-trained CNN models. In fact, AlexNet is smaller and faster than VGG networks. The VGG network is more challenging to train. Training the VGG network requires much more memory and computation time than AlexNet. VGG also has a higher number of parameters compared to AlexNet and GoogLeNet, which makes it more computationally expensive. ResNet is very deeper than AlexNet so, it is harder to train than AlexNet.

The presented work attempts to avoid the drawbacks of the previous methods. Some of these previous methods [13] and [14] work with images pixel by pixel, which is inefficient in many cases. The results shown in other publications [12, 16], and [18] use only simple test examples when the knife is clearly observed in the image. Other publications [12] work with a few numbers of images (not exceed 40 images). Other publications suit to applications in which the knife tip is clearly visible, such as in baggage-scanning systems. The majority of these approaches are not suitable for real-life situations. On the contrary, DL does not need to extract features from the image manually. It extracts features automatically in convolution layers.

Furthermore, the detection process using CNN is rugged to distortions in the image (such as different lighting conditions, different poses, and presence of partial occlusions). CNN is simpler and better in training than other traditional neural networks. It also needs low memory requirements.

The most four popular CNN models are tested in the proposed approach to determine the most suitable pre-trained network to detect knives. AlexNet is characterized by an architecture having a small number of layers. This feature reduces complexity and saves both time and memory.

The superiority of the proposed technique is due to our Brightness Handler Procedure (BHp). It makes the method more adaptive by determining precisely the frames that have a brightness problem using the proposed ABT. It provides very accurate results with various brightness conditions.

5 Conclusions and Future Work

The paper addresses a very hot topic, especially with the increase in violence related to the use of knives in many countries. Unfortunately, the research does not reflect this need with few works about knives compared to other weapons. Also, there are still unsolved problems about knife detection. One of these problems is the reflection of light addressed in this paper. The reflection of light on the knife blade hinders its detection in smart video surveillance systems. This paper presents an adaptive technique for brightness enhancement of knife detection in surveillance systems. This technique overcomes the brightness problem that faces the steel weapons and improves the knife detection process. The novelty of the proposed technique is that it suggests an automatic threshold (ABT) to assess the level of frame brightness. It also proposes a technique (BHp) to selectively apply brightness correction only on the required frames which improves the performance. DL is used in the presented work to extract features immediately from raw images. Applying DL as a feature extractor and a classifier makes the proposed technique very suitable for real-life situations. We analyze the CNN pre-trained networks to select the best network in the detection process for the knives.

Experimental results demonstrate that the proposed technique achieves a promising performance for knife detection. It achieves an accuracy, specificity, sensitivity, and F1-score equal to 99.38%, 99.1%, 98.3%, and 96.1%, respectively. Due to the proposed Brightness Handler Procedure (BHp), these results were achieved. These results are better than state-of-the-art methods. As future work, further research will be intended to concentrate on providing a complete solution for detecting other dangerous objects like firearms and postils in outdoor scenarios.

References

1. Grega, M. et al.: Automated recognition of firearms in surveillance. In: IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), San Diego (2013)
2. Kumar, R.T.; Verma, G.K.: A computer vision based framework for visual gun detection using SURF. In: Proceedings of International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO,2015) (2015)
3. Kumar, RT; Verma, G.K.: A computer vision based framework for visual gun detection using Harris Interest Point Detector. In: Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015) in ElSevier, vol. 54, pp. 703–712 (2015)
4. Dhulekar, P.A., et al.: Motion estimation for human activity surveillance. In: International Conference Emerging Trends & Innovation in ICT (ICEI,2017), Pune, India (2017)
5. Arslan, A.N.; Hempelmann, C.F., et al.: Threat assessment using visual hierarchy and conceptual firearms ontology. Opt. Eng. **54**(5), 105–109 (2015)
6. Darker, l. et al.: Can CCTV reliably detect gun crime?. In: Proceedings of IEEE, United States of America, United States of America (2007)
7. Blechko, A. et al.: Skills in detecting gun carrying from CCTV. In: Proceedings of IEEE, Prague, Czech Republic (2009)
8. Darker, I.T.; Gale, A.G.; Blechko, A.: CCTV as an automated sensor for firearms detection: Human-derived performance as a precursor to automatic recognition. In: Proceedings of the International Society for Optical Engineering, (ISOE,2008), Cardiff, Wales, United Kingdom (2008)
9. Xiao, Z.; Lu, X.; Yan, J.; Wu, L.; Luyao: Automatic detection of concealed pistols using passive millimeter wave imaging. In: 2015

IEEE International Conference on Imaging Systems and Techniques (IST), 1–4, IEEE, 2015
10. Zywicki, M. et al.: Knife detection as a subset of object detection approach based on Haar cascades. In: Proceedings of 11th International Conference on Pattern Recognition and Information Processing, Minsk, Belarus (2011)
11. Kmiec, M.; Glowacz, A.: An approach to Robust Visual Knife Detection (MG&V,2011). Mach. Graphics Vis. **20**(2), 215–227 (2011)
12. Kmieć, M. et al.: Towards robust visual knife detection in images: active appearance models initialised with shape-specific interest points. In: Proceedings of International Conference on Multimedia Communications, Services and Security., Krakow, Poland (2012)
13. Maksimova, A.: Knife detection scheme based on possibilistic shell clustering. In: International Conference on Multimedia Communications, Services and Security, Springer, Berlin, Heidelberg (2013)
14. Glowacz, A., et al.: Visual detection of knives in security applications using active appearance models. Multimed. Tools Appl. Springer **74**(12), 4253–4267 (2015)
15. Maksimova, A., et al.: Fuzzy classification method for knife detection problem. Int. Conf. Multimed. Commun., Serv. Secu., Springer **429**, 159–169 (2014)
16. Kmieć, M.; Glowacz, A.: Object detection in security applications using dominant edge directions. Pattern Recogn. Lett., in Elsevier **52**, 72–79 (2015)
17. Grega, M., et al.: Automated detection of firearms and knives in a CCTV image. Sensors **16**(1), 47–63 (2016)
18. Vajhala, R. et al.: Weapon detection in surveillance camera images, Department of Applied Signal Processing Blekinge Institute of Technology, Karlskrona, Sweden (2016)
19. Buckchash, H.; Raman, B.: A robust object detector: Application to detection of visual knives. In: Proceedings of the IEEE International Conference on Multimedia and Expo Workshops (ICMEW) 2017 (2017)
20. LeCun, Y.; Kavukcuoglu, K.; Farabet, C.: Convolutional networks and applications in vision. In: Proceedings of 2010 IEEE International Symposium on Circuits and Systems (2010)
21. Yanming Guo, Yu.; Liu, A.O.; Lao, S.: Deep learning for visual understanding: A review. Neurocomputing **187**, 27–48 (2016)
22. Liua, W.; Wanga, Z.; Liua, X.; Ze, N.: A survey of deep neural network architectures and their applications. Neurocomputing **234**, 11–26 (2017)
23. Haridas, R.; Jyothi, R.L.: Convolutional neural networks: A comprehensive survey. Int. J. Appl. Eng. Res. **14**(3), 780–789 (2019)
24. Olmos, R.; Tabik, S.; Herrera, F.: Automatic handgun detection alarm in videos using deep learning. Neurocomputing **275**, 66–72 (2018)
25. Verma, G.K.; Dhillon, A.: A handheld gun detection using Faster R-CNN deep learning. In: Proceedings of the 7th International Conference on Computer and Communication Technology, pp. 84–88 (2017)
26. Al-Shoukry, S.: An automatic hybrid approach to detect concealed weapons using deep learning. ARPN J. Eng. Appl. Sci. **12**(16), 4736–4741 (2017)
27. Castillo, A.; Tabik, S.; Pérez, F.: Brightness guided preprocessing for automatic cold steel weapon detection in surveillance videos with deep learning. Neurocomputing **330**, 151–161 (2019)
28. Noever, DA.; Noever, SEM.: "Knife and Threat Detectors," white paper (2020)
29. Zhong-Qiu, Z.; Peng, Z.; Shou-tao, X.; Xindo: Object detection with deep learning: A review. IEEE Trans. Neural Netw. Learning Syst. **30**(11), 3212–3232 (2019)
30. Alom, M.Z.; Taha, T.M.; Yakopcic, C.; Westberg, S.: "The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches,", pp. 1–39 (2018)
31. Torrey, L.; Shavlik, J.: Transfer learning. In: Appears in the Handbook of Research on Machine Learning Applications, published by IGI Global (2009)
32. Pan, S.J.; Yang, Q.: A survey on transfer learning. IEEE Trans. Knowl. Data Eng. **22**(10), 1345–1359 (2009)
33. Nithya, M.D.K.: Comparison of contrast enhancement technique with partitioned iterated function system. Int. J. Sci. Eng. Res. **5**(5), 167–172 (2014)
34. Rahman, S.; Rahman, M.M.; Abdullah-Al-Wadud, M.; Al-Quaderi, G.D.; Shoyaib, M.: An adaptive gamma correction for image enhancement. EURASIP J. Image Video Process. SpringerOpen **2016**(1), 1–13 (2016)
35. Pizer, S.M.; Johnston, R.E.; Eri, J.P.: Contrast-limited adaptive histogram equalization: speed and effectiveness. In: Proceedings of the First Conference on Visualization in Biomedical Computing, Atlanta, Georgia (1990).
36. Zuiderveld, K.: Contrast Limited Adaptive Histogram Equalization, p. 474–485. Academic Press Professional Inc, Singapore (1994)
37. Reza, A.M.: Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. J. VLSI Signal Process. Syst. Signal, Image Video Technol. Springer **38**(1), 35–44 (2004)
38. [Online]. Available: http://kt.agh.edu.pl/~matiolanski/KnivesImagesDatabase/
39. [Online]. Available: https://llwww.sites.google.com/site/kdsdataset/
40. "Knife dataset for detection," [Online]. Available: https://www.kaggle.com/shank885/knife-dataset?
41. "Knife detection," [Online]. Available: https://www.kaggle.com/vijaysingh888/knife-detection
42. Tom, F.: An introduction to ROC analysis. Pattern Recogn. Lett. **27**(8), 861–874 (2006)