Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 196 (2017) 895 – 902

Procedia Engineering

www.elsevier.com/locate/procedia

Creative Construction Conference 2017, CCC 2017, 19-22 June 2017, Primosten, Croatia

# Automated Hardhat Detection for Construction Safety Applications

Bahaa Eddine Mneymneh[a], Mohamad Abbas[a], Hiam Khoury[a,*]

[a]*American University of Beirut, Department of Civil and Environmental Engineering, Beirut, 1107 2020, Lebanon*

---

**Abstract**

Despite various safety inspections carried out over the years to ensure compliance with regulations and maintain acceptable and safe working conditions, construction is still among the most dangerous industries responsible for a large portion of the total worker fatalities. A construction worker has a chance of 1-in-200 of dying on the job during a 45-year career, mainly due to fires, falls, and being struck or caught-in/between objects. This in part can be attributed to how monitoring the presence and proper use of personal protective equipment (PPE) by safety officers becomes inefficient when surveying large areas and a considerable number of workers. Therefore, this paper takes the initial steps and aims at evaluating existing computer vision techniques, namely object detection methods, in rapidly detecting whether workers are wearing hardhats from images captured on many indoor jobsites. Experiments have been conducted and results highlighted the potential of cascade classifiers, in particular, in accurately and precisely detecting hardhats under various scenarios and for repetitive runs.


Peer-review under responsibility of the scientific committee of the Creative Construction Conference 2017

*Keywords:* computer vision; construction; hardhat; personal protective equipment; safety

---

## 1. Introduction and related work

It has been long stated that construction is unarguably one of the most perilous industries whereby many hazardous tasks and conditions exist, which may pose injuries, risks and fatalities to the workers. For instance, fires, falls, and being struck or caught-in/between objects contribute to over 50% of the total casualties in the sector. Hence, proper use and adoption of safety equipment such as personal protective equipment (PPE), in particular hardhats, was deemed necessary on jobsites [1-3]. However, the current monitoring of hardhat-wearing remains manual, tedious and time-consuming. Therefore, there is a great need to automate this process in a cost-effective manner with highly reduced turnaround times in order to mitigate the risks of injuries and fatalities and allow a safer environment.

* Corresponding author. Tel.: +961-1-350000 Ext: 3428; fax: +961-1-744462.
*E-mail address:* hiam.khoury@aub.edu.lb


Peer-review under responsibility of the scientific committee of the Creative Construction Conference 2017
doi:10.1016/j.proeng.2017.08.022

Among several information technology (IT) and computer-based tools widely adopted in the construction field to automate various processes [4-7], computer vision techniques proved to be efficient in rapidly and conveniently retrieving relevant data from construction sites such as the detection and tracking of workers and equipment [8]. For instance, recent research efforts used computer vision for PPE detection [1]. An algorithm was developed to detect workers using standard face detection and hardhats by applying edge detection on the region directly above the worker's head. In this case, the system detected a hardhat if its outline was found to be a semicircle and its color identified as red. However, their algorithm required a set of high resolution CCTV cameras to be installed on site and was only able to detect hardhats when applied on images captured from the front. Additionally, their system was not assessed on an actual construction site. Many other research studies have targeted similar applications [9-13]. However, research in this field remains in its infancy. The existing systems were either victims of over-prediction and false identification of unwanted objects as hardhats, or were never tested against increasing levels of challenge due to variations in orientation, color, background contrast, image resolution and on-site lighting conditions. In addition, none of the previous studies were concerned with the time efficiency of the detection method for adoption in real-time scenarios. After all, the computational speed of the algorithm is as important as its accuracy, especially in safety applications. In order to address the aforementioned limitations, this paper aims at automating part of construction safety inspections by evaluating existing computer vision techniques in rapidly and efficiently detecting hardhats in standard resolution images captured from any perspective on actual indoor construction sites.

## 2. Research methodology

This section describes and evaluates existing computer vision techniques deemed useful for detecting hardhats. Among these techniques, object detection/recognition methods proved promising in particular: (1) Feature detection, extraction and matching, (2) template matching, and (3) cascade classifiers models. The usefulness of each visual object recognition method varies according to many factors included but are not limited to color, orientation, shape, scale, etc. of the target object. The components of the algorithms were implemented using Matlab 2016a.

### *2.1. Feature detection, extraction, and matching*

In this study, gradient-based features such as the Speeded-Up Robust Features (SURF) or binary features including Binary Robust Invariant Scalable Keypoints (BRISK) and Features from Accelerated Segment Test (FAST) [14], are used to detect point correspondences between the input image ( and a reference image containing a hardhat. In Matlab, SURF features are detected from a grayscale image using the following code:

*Surf_Features = detectSURFFeatures(rgb2gray(Image));*

On the other hand, feature extraction locates the detected features within each image, while feature matching identifies similarities between the reference and input images, using the following code snippet based on SURF features:

*[feats1, validpts1] = extractFeatures(rgb2gray(Reference), Surf_Features_Reference);*

*[feats2, validpts2] = extractFeatures(rgb2gray(Input), Surf_Features_Input);*

*Index_Matched_Features = matchFeatures(feats1, feats2);*

In this case, the hardhat having the best match with the reference image is detected. The number of hardhats is then computed by hiding identified ones from the target image so that the next best match hardhat can be detected in the next iteration of the algorithm. This counting iterative process halts when no more hardhats can be detected in the target image. It is worth noting that this method works best for objects displaying non-repeating texture patterns to allow unique and numerous feature matches.

### *2.2. Template matching*

Template matching often refers to a series of operations aiming at detecting and identifying a certain form or pattern in an input image, by comparison with a reference or template [15]. The template is positioned over the input image at every possible position and a similarity coefficient is calculated. Possible metrics to determine the similarity include the sum of absolute differences (SAD), the sum of squared differences (SSD), and the maximum absolute difference (MaxAD) [16]. Other methods searching for the minimum difference between two images consist of either an Exhaustive search (ES) or a Three-Step search (TSS). The former is more accurate but more computationally expensive, while the latter is quicker, but may not always find the optimal solution. In Matlab, a template matcher is typically based on SAD unless otherwise stated (e.g. Three-Step) as shown in the code snippet below:

*Detector = vision.TemplateMatcher('SearchMethod','Three-step');*

A hardhat is detected when the difference computed between a template image containing a hardhat and the input image is less than a required threshold. Generally, template matching algorithms are limited by the available computation power due to a required high detection accuracy that necessitates lengthy iteration processes.

### *2.3. Cascade classifier*

In this study, different cascade classifiers based on Histogram of Oriented Gradients (HOG), Haar-like, and Local Binary Pattern (LBP) features are assessed. This requires a training process, using two sets of positive and negative instances. Positive instances contain images of the relevant object while negative instances are images that do not contain the relevant object. A sample of 75 positive and 164 negative images was collected from construction environments to train the three cascade object detectors. The training process requires as well a set of input parameters, including the number of cascade stages, the true positive rate and the false alarm rate (FAR). Experimenting with those parameters yields different results, allowing for the creation of a more effective detector. For example, training a cascade object detector based on HOG features and with the required parameters is performed using the following Matlab code:

*trainCascadeObjectDetector('Hog_7_10.XML', positiveInstances, negativeFolder, 'FalseAlarmRate', 0.10, 'NumCascadeStages', 7, 'FeatureType', 'HOG')*

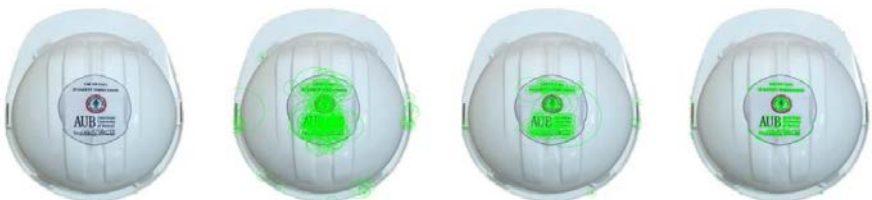
## 3. Preliminary results and analysis

### *3.1. Performance of the feature detection, extraction, and matching algorithm*

Due to the uniform shape and color of a hardhat, the number of detected features was found to be low (see Fig. 1). One suggested solution to the problem was to add a customized sticker to the hardhat (see Fig. 2a). This then greatly increased the number of extracted features (see Fig. 2b, 2c, 2d).

(a) (b) (c) (d)

Fig. 1. White Hardhat: (a) Original Image; (b) Detected SURF features (4); (c) Detected BRISK features (0); (d) Detected FAST features (0).



(a) (b) (c) (d)

Fig. 2. Hardhat with sticker: (a) Original Image; (b) Detected SURF features (285); (c) Detected BRISK features (172); (d) Detected FAST features (424).

In order to further assess the applicability of this algorithm in detecting hardhats in indoor construction environments, experiments were conducted on close-up images (see Fig. 3) showing clearly the customized stickers on both hardhats. The algorithm is independent from any type of feature used, but given that a minimum number of features needs to be extracted with the least computational power, the choice landed on SURF features. For example, in Fig.3a, 63 matching features were found between the reference and the target image in the first iteration of the algorithm. The first detected hardhat was then hidden from the target image and in the second iteration, 44 matching features were found between the reference and the new target image (see Fig. 3b). The iteration process halts once the second hardhat is hidden and the program returns the final number of detected hardhats.

As such, due to the lack of pertinent features on the hardhat, the algorithm searches for the customized sticker and identifies its target irrespective of the color or shape. However, further testing revealed some deficiencies in the system. The method is actually susceptible to misclassifying any object carrying the sticker. Moreover, in a three-dimensional dynamic construction environment, a clear view of the sticker cannot be always guaranteed. As a matter of fact, in another sets of experiments, a hardhat could not be detected because either the size or resolution of the sticker was low, or the sticker was not visible due to the orientation of the hardhat. Additionally, the feature extraction and filtering together with the iteration processes required a relatively high calculation cost.

### 3.2. Performance of the template matching algorithm

Experiments conducted in an indoor construction environment showed that the algorithm wrongly predicted the location of a blue hardhat when using a template with a slightly different rotation (see Fig. 4). As such, a unified template is not sufficient to detect all instances and a classic template matching is relatively inaccurate when dealing with any form of difference in scale and rotation. Furthermore, the lengthy calculation process of classic template matching eliminates any usefulness of such an algorithm in a real-time application. In fact, scanning full resolution images from construction sites required hours of processing and a lot of computational power.

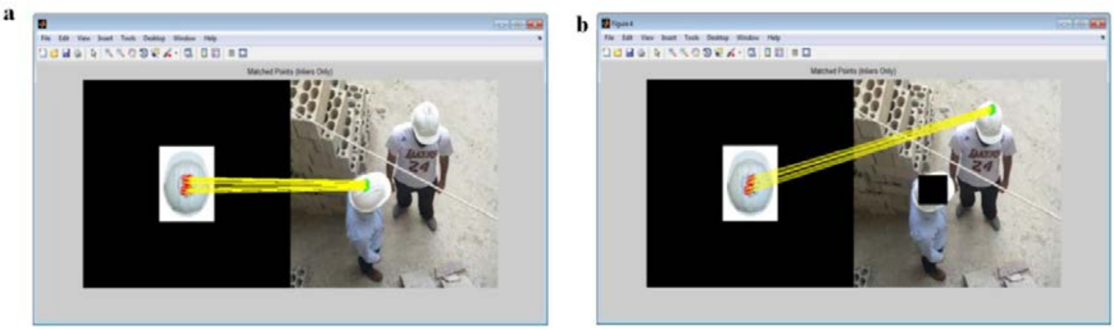



Fig. 3. Matching features for the: (a) first hardhat; (b) second hardhat.


Fig. 4. Wrong detection using template matching.

### *3.3. Performance of the cascade classifier*

Objects detectors are often sensitive to out of plane transformation. However, this should not be a problem in the case of hardhat detection since its semi-circular shape remains unchanged regardless of the viewing angle. Cascade detectors based on Haar and LBP features yielded high rates of wrong detection in all testing images (see Fig. 5). On the other hand, detectors based on HOG features could accurately describe the circular shape of the hardhat irrespective of its color (see Fig. 6). The ability of the detector to correctly identify hardhats from different viewpoints was verified as well using a set of 3 testing images containing front, side and back views of a blue hardhat (see Fig. 7). The classifier was also capable of recognizing two objects simultaneously. Color variations had also no effect and the computational speed was acceptable. Furthermore, the capacity to experiment with training parameters to obtain different results is another main advantage of cascade classifiers.

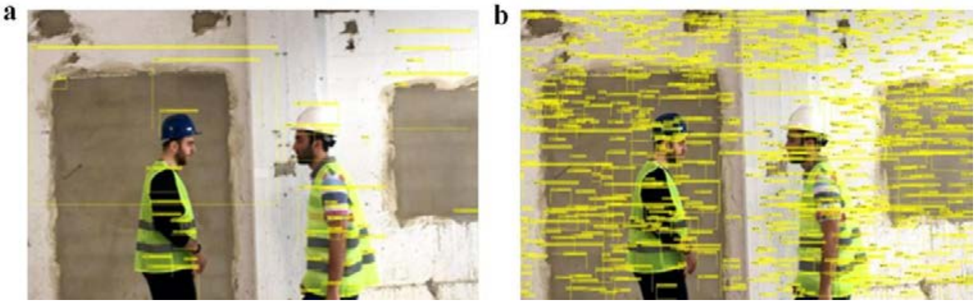


Fig. 5. High rate of incorrect detections using: (a) Haar features; (b) LBP features.

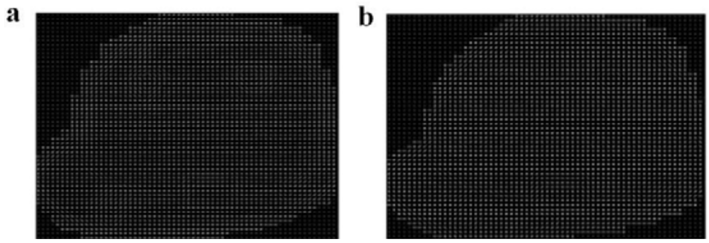


Fig. 6. Extracted HOG features: (a) blue hardhat; (b) white hardhat.

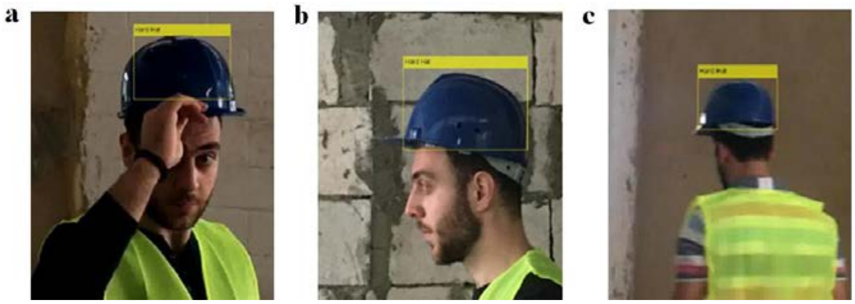


Fig. 7. Detected hardhat: (a) front view; (b) side view; (c) back view.

### 3.4. Comparison and selection

Based on preliminary experiments and results featured above, each technique is assessed according to the following criteria summarized in Table 1. The cascade object detector clearly outperforms the other object detection techniques and can be potentially adopted in real-time safety applications.

Table 1. Comparative summary of object detection techniques

| | Feature Detection, Extraction & Matching | Template matching | Cascade classifier |
|---|---|---|---|
| Computational Duration | Medium | Very High | Low |
| Color invariance | Yes | No | Yes |
| Orientation Invariance | No | No | Yes |
| Practicality | No | No | Yes |
| Customizability | Yes | No | Yes |
| Training database/fingerprinting | No | No | Yes |

## 4. Experimental analysis of the cascade classifier

In this section, further assessment of the HOG-based cascade object detector is carried out. Two seven-stage cascade object detectors were trained using the same image datasets of 75 positive and 164 negative images and given two different values of false alarm rates (FAR) set to 0.05 and 0.1. In theory, a larger false alarm rate should yield more false positive results and the detector should be less likely to miss a desired object. The performance of the cascade classifier was then analyzed against variations in orientation and color, background contrast, image resolution, and lighting conditions. As such, three scenarios were devised and respective results are depicted in Table 2, including computational durations (t). A fourth scenario (i.e. low luminosity) was also considered. However, its results were not reported because the variation in image luminosity did not considerably affect the performance of the HOG-based detector as it is capable of describing the shape of the object irrespective of its color.

Table 2. Performance of cascade object detectors in scenarios 1, 2, and 3

| **Scenario 1- High contrast against background, variable colors and orientations** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Image ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| True number of hardhats | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | |
| Detected- FAR = 0.05 | 1 | 1 | 0 | 0 | 1 | 2 | 1 | 2 | 1 | 3 | **t= 2.171s** |
| Detected- FAR = 0.1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 3 | **t= 2.231s** |
| **Scenario 2- Low contrast against background, variable colors and orientations** | | | | | | | | | | | |
| Image ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| True number of hardhats | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | |
| Detected- FAR = 0.05 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | **t=2.191s** |
| Detected- FAR = 0.1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 2 | 2 | 1 | **t=2.196s** |
| **Scenario 3- Different image resolutions** | | | | | | | | | | | |
| Image ID | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| True number of hardhats | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 2 | 2 | |
| Detected- FAR = 0.05 | 1 | 1 | 0 | 0 | 1 | 2 | 0 | 1 | 1 | 3 | **t= 0.541s** |
| Detected- FAR = 0.1 | 1 | 1 | 1 | 1 | 1 | 2 | 0 | 1 | 2 | 3 | **t=0.568s** |

In scenario 1, the level of challenge was relatively low and all hardhats could be easily discerned from their respective backgrounds. The two detectors were tested on 10 images with 13 hardhats in total and the one with an FAR set to 0.05 missed 3 while the detector with an FAR set to 0.1 did not miss any hardhat. Nevertheless, both classifiers were subject to wrong identification and for instance, in image 6, a mobile worker's head was mistakenly classified as hardhat (see Fig.8). The computational durations of both detectors were similar with an average processing time per image of around 2 seconds.

As for scenario 2, the level of challenge was greatly increased. The low contrast between the hardhat and its background (e.g. white hard hat in front of white wall) may reduce the significance of the detected HOG features which, in turn, can reduce the efficiency of the detector. In fact, the performance of the cascade object detector dropped compared to scenario 1. More specifically, for the higher FAR of 0.1, the detector was able to identify 10 out of 13 hardhats. The computational durations of both detectors did not significantly vary with a similar average processing time per image of around 2 seconds.

In order to assess the effect of changing the resolution of test images on detection results independently of other factors, the third experiment (scenario 3) was carried out using the same images of scenario 1 but cropped or resized to obtain an image resolution of (1920 × 1080) pixels. Images 1 to 5 were cropped while images 6 to 10 were resized. Cropped images (1-5) yielded results identical to scenario 1. On the other hand, resizing the image can possibly decrease the size of hardhats below the trained size and accordingly does not allow hardhat detection as is the case of images 7 and 8. Therefore, training the detector using images of the same resolution and taken from a similar distance as expected test images can positively impact detection results. Compared to scenarios 1 and 2, these scenario's computational durations improved significantly from around 2 seconds to around 0.5 seconds.

## 5. Conclusion and future work

This paper evaluated existing computer vision algorithms, in particular object detection methods, in efficiently and rapidly detecting hardhats on indoor construction sites. Several experiments were conducted and results revealed that a well-trained cascade classifier was found to be robust under various scenarios and conditions. Additionally, it was proven to be relatively time-efficient and in a real-time application, it is capable of scanning for violations every 2 seconds. The process can even be expedited by reducing the resolution of the training and test images.

Further work will look into improving the accuracy of the hardhat detection process and eliminating false detections by combining the cascade classifier with image and color segmentation techniques. Further testing is also required to evaluate the accuracy of other object detection algorithms as well as to explore the potential use of heat cameras besides digital imagery. Future studies will aim as well to integrate this hardhat detection process into a complete safety inspection framework capable of rapidly, efficiently, and conveniently issuing safety warnings when a safety violation is detected and alerting nearby workers of hazards.

Fig. 8. Wrong classification of head region in image number 6 – Scenario 1

## Acknowledgements

The presented work has been supported by AUB's University Research Board (URB). The authors gratefully acknowledge URB support. Any opinions, findings, conclusions, and recommendations expressed by the authors in this paper do not necessarily reflect the views of URB.

## References

[1] K. Shrestha, P. P. Sherestha, D. Bajracharya and E. A. Yfantis, Hard-Hat Detection for Construction Safety Visualization, J. of Const. Eng. 2015 1-8.
[2] Health and Safety Executive, Health and safety in construction sector in Great Britain, 2014/15. Available at: http://www.hse.gov.uk/statistics/industry/construction/construction.pdf
[3] Occupational Safety and Health Administration, Module 13—Personal protective equipment. Available at: https://www.osha.gov/dte/
[4] Y. S. Kim, J. H. Lee, H. S. Yoo, J. B. Lee and U. S. Jung, A performance evaluation of a Stewart platform based Hume concrete pipe manipulator, Auto. in Const. 18 (2009) 665-676.
[5] T. Yoshida and S. Chae, Application of RFID technology to prevention of collision accident with heavy equipment, Auto. in Const. 19 (2010) 368-374.
[6] L. Ding, H. L. Yu, C. Zhou, X. Wu and M. H. Yu, Safety risk identification system for metro construction on the basis of construction drawings, Auto. in Const. 27 (2012) 120-137.
[7] M. J. Skibniewski, Information technology applications in construction safety assurance, J. of Civ. Eng. and Manag. 20 (2014) 788-794.
[8] M. Memarzadeh, M. Golparvar-Fard and J. C. Niebles, Automated 2D detection of construction equipment and workers from site video streams using histograms of oriented gradients and colors, Auto. in Const. 32 (2013) 24-37.
[9] S. Du, M. Shehata and W. Badawy, Hardhat Detection in Video Sequences based on Face Features, Motion and Color Information, 3rd Int. Conf. on Comp. Res. and Dev. 2011.
[10] G. Gualdi, A. Prati, and R. Cucchiara, Contextual Information and covariance descriptors for people surveillance: an application for safety of construction workers, EURASIP J. on Im. and Vid. Proc. 9 (2011) 1-16.
[11] D. Bajracharya. Real time pattern recognition in digital video with application to safety in construction sites. University of Nevada, Las vegas. 2013.
[12] M.W. Park, N. Elsafty and Z. Zhu, Hardhat-Wearing Detection for Enhancing On-Site Safety of Construction Workers, J. of Const. Eng. and Manag. 141 (2015).
[13] A. H. M. Rubaiyat, T. T. Toma, M. Kalantari-Khandani, S. A.Rahman, L. Chen, Y. Ye, C.S. Pan. Automatic detection of helmet uses for construction safety. 2016 IEEE/WIC/ACM Int.l Conf. on Web Intel. Work. Nebraska. 2016. 135-142.
[14] K. Mikolajczyk and T. Tuytelaars, Local Image Features, in Encyclo-pedia of Biometrics, S. Li and A. Jain, Eds. Springer US, 2009, pp. 939-943.
[15] R. Brunelli, Template Matching Techniques in Computer Vision: Theory and Practice. Wiley 2009.
[16] J. Yu, J. Amores, N. Sebe, Q. Tian, A New Study on Distance Metrics as Similarity Measurement, IEEE Int. Conf. on Mult. and Expo, 2006.