Journal of Network Intelligence ©2022 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **7**, Number **2**, May **2022**

Safety Helmet Wearing Detection Based on Contour and Color Features

Yi-Jia Zhang, Fu-Su Xiao

School of Information Science and Technology
Zhejiang Sci-Tech University
Hangzhou 310018, P. R. China
waiting@zstu.edu.cn

Zhe-Ming Lu

School of Aeronautics and Astronautics
Zhejiang University
No. 38, Zheda Road, Hangzhou 310027, P. R. China
Corresponding author: zheminglu@zju.edu.cn

Received November 2021; revised February 2022

Abstract. *In many construction sites and motorcycle driving scenes, whether a safety helmet is worn or not is directly related to the life safety of workers and drivers. Therefore, the detection of helmet wearing is the key technology to assist site safety and road safety. However, the existing manual feature extraction methods typically use person location to further determine the wearing of the helmet, which is cumbersome and has a high false detection rate. In this paper, a fast and practical safety helmet-wearing detection framework is proposed, which detects the head with a helmet directly by contour and color features. Firstly, reduce image noise through image preprocessing. After that, based on the result of image preprocessing, the secondary region of interest(ROI) for the helmet is extracted by face detection, skin color detection and helmet contour detection. Finally, color space conversion and color feature recognition are performed on the two-stage ROI to realize the detection of helmet wearing. Extensive experimental results based on various photos of construction sites, in the dataset SHWD, demonstrated the effectiveness of the proposed framework.*

Keywords: Region of interest, Helmet wearing detection, HSV color space

1. **Introduction.** Nowadays, traffic and construction safety has been paid more and more attention and importance. Drivers riding electric bicycles and motorcycles, as well as workers on construction sites, need to wear helmets. However, tragedies caused by driving a non-motorized vehicle without a helmet still occur from time to time, and accidents on the construction site caused by working without a helmet emerge in endlessly. Therefore, it is very practical and necessary to use the helmet-wearing detection algorithm to assist in supervising the wearing of the helmet on various occasions.

Safety helmet wearing detection belongs to the target detection and recognition of digital images. The main research idea is to detect and locate the target(human usually) through features or state extraction in the image to be detected, and further recognize the feature(such as color, contour, texture, etc.) of the target, and finally complete the detection. In an earlier study, DALAL N and TRIGGS B proposed a human detection algorithm based on the Histograms of Oriented Gradients (HOG) feature[1], which is mostly used for pedestrian extraction in helmet-wearing detection algorithm[2-3]. After extracting features, Support Vector Machine(SVM) is usually used for classification, Wu.Q proposed an improved SVM to make it have higher speed and better performance[4]. R Waranusast et al. designed a method[5] to extract and classify moving objects by using K-Nearest Neighbor (KNN) classifier. In [6], a method of combining face and helmet detection in construction sites was proposed. In recent research, Li. K, Zhao. X and other scholars achieved helmet-wearing status recognition by human head detection and color features[7], but the head detection of this method treats the top 1/5 of the human detection frame as the head area, which is not universal because it limits the diversity of detection objects. For example, when the image to be detected is a photo of a bust, it loses its practicality. Chen. C and Hu. S combined image feature extraction and random forest algorithm to realize intelligent detection of helmet wearing, this method achieved good results in power equipment construction sites.

In this paper, we design a practical helmet-wearing detection algorithm based on contour and color features for workers on construction sites. In order to reduce image noise, the image preprocessing including image smoothing and image enhancement is adopted. Based on the result of image preprocessing, the two-stage ROI for the helmet is extracted by face detection, skin color detection and helmet contour detection. According to the result of the region of interest extraction, the safety helmet wearing detection is implemented by color space conversion and color feature recognition to detect whether a helmet is worn or not for workers.

The remainder of this paper is arranged as follows. Section 2 shows the overall framework of helmet wearing detection system. Section 3 gives the details of our proposed method including image preprocessing, the two-stage region of interest extraction and color feature recognition. Section 4 presents the results and evaluation of the experiments. Finally, Section 5 concludes the whole paper and arranges the future work.

2. **Helmet wearing detection system framework.** This paper proposes a method for detecting helmet wearing based on contour and color features, which mainly consists of three steps: image preprocessing, ROI extraction, and helmet-wearing detection. The framework of the helmet-wearing detection scheme in this paper is shown in Fig.1.

FIGURE 1. Framework of the helmet wearing detection scheme

The purpose of image preprocessing is that reducing image noise as much as possible by image filtering and gamma transformation. Firstly, performing a two-dimensional Gaussian noise process on the image to reduce most of the image noise and preserve image edge at the same time. After that, gamma transformation is used to standardize the color space of the image to adjust the contrast of the image and reduce the impact of partial light changes in the image.

The ROI extraction is decisive for the practicability of the whole algorithm because it directly affects the judgment of helmet wearing. The extraction of ROI in this paper is mainly achieved by the combination of face detection, skin color detection and helmet contour detection. The main function of skin color detection is to assist in judging the presence of human faces. According to the judgment of the skin color detection, the faces in the image are recognized by the haar feature, and then selects the region of the eyes up to a certain distance as the primary ROI for the helmet wearing detection. Performing binary and morphological processing on the image of the primary ROI to obtain a low-noise enhanced image. The last step of ROI extraction is to determine the contour in the enhanced image and treat the contour area with more than 10 sides as the secondary ROI. Fig.2 shows the process of two-stage ROI extraction.

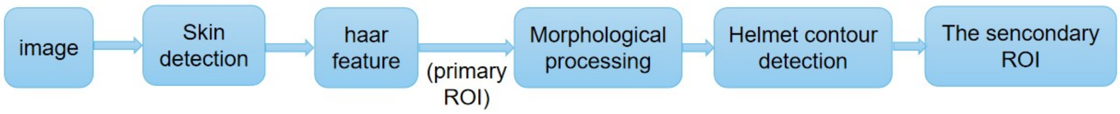

FIGURE 2. The process of two-stage ROI extraction

Due to the color of the helmet is more prominent in HSV color space, whether a helmet is worn or not is mainly determined by the HSV color space. In this paper, the extracted ROI of the original image would be converted to HSV color space and filtered by the range of helmets of different colors to determine whether the person in the image is wearing a helmet or not. The final result is that the helmet-wearing situation is framed out in the original image and output. Fig.3 shows the ROI in HSV color space and the result of helmet-wearing detection.

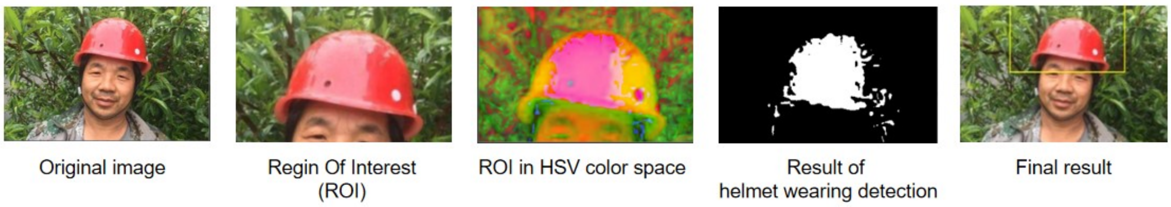

FIGURE 3. The ROI in HSV color space and the result of helmet detection

3. Method Principles of System Framework.

This section mainly introduces some methods involved in our proposed algorithm, such as image filtering, gamma transformation, skin color detection, face detection, contour detection and color feature recognition.

3.1. Image Preprocessing.

Image preprocessing, as the first step of helmet wearing algorithm in this paper, is committed to reducing noise and enhancing key information before ROI extraction, which can improve the performance of the whole algorithm. We designed image filtering and gamma transformation to pre-process the image.

3.1.1. *Image Filtering.*

The original image inevitably has some noise points, and the image can be filtered and denoised by Gaussian blur, moreover, bilateral Gaussian blur can achieve the dual effects of noise removal and image edge preservation.

Suppose the Gaussian convolution operator gaussKernel with width W and height H is constructed, where W and H are odd numbers, and the anchor point is at $(\frac{H-1}{2},\frac{W-1}{2})$. At first, calculate the Gaussian matrix according to Eq. (1).

$$gaussMatrix = [guass(r, c, \sigma)], 0 \le r \le H-1, 0 \le c \le W-1, r, c \in N \quad (1)$$

Where guass(r,c,σ) is expressed as Eq.(2)

$$guass(r,c,\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(r-\frac{H-1}{2})^2+(c-\frac{W-1}{2})^2}{2\pi\sigma^2}} \tag{2}$$

After calculating the Gaussian matrix, the key step is to normalize and get the Gaussian convolution operator gaussKernel by Eq.(3)

$$gaussKernel = \frac{gaussMatrix}{sum(gaussMatrix)} \tag{3}$$

The Gaussian convolution operator is the same as itself after being flipped by 180°, and the Gaussian convolution kernel is a separable convolution kernel. Therefore, when performing Gaussian smoothing on an image, we can first perform Gaussian smoothing in the one-dimensional horizontal direction on the image, and then perform Gaussian smoothing in the one-dimensional vertical direction.

3.1.2. *Gamma Transformation.* Gamma transformation is often used to adjust the contrast of overexposed or underexposed images. Specifically, it is through nonlinear transformation to brighten the darker area, and darken the over-bright areas in the image at the same time. After gamma transformation, the overall details of the image will be enhanced[8-10].

Assume the input image is I, and its width and height are W and H respectively. First, normalize its gray value to the range of [0,1]. I (r, c) represents the gray value of the r-th row and c-th column after normalization, the output image is marked as O, and the gamma transformation is defined as follows.

$$O(r,c) = I(r,c)^\gamma, 0 \le r < H, 0 \le c < W \tag{4}$$

Where O (r,c) is the output value of the r-th row and c-th column after gamma transformation. γ is the gamma factor's magnitude, which governs the zoom degree of transformation. When γ is less than 1, it will stretch the lower gray-level area in the image while compressing the higher gray-level part.

The γ value set in this paper is 0.5. After debugging and analysis, we conclude that gamma transformation has obviously improved the detection effect of scenes that are affected by illumination. For example, an image whose original detected order is 90, and the detected order becomes 103 after gamma transformation, The local accuracy rate is increased by 14.4 percent, and the overall accuracy rate is increased by about 8 percent. Furthermore, Fig.4 shows the contrast with or without gamma transformation in skin color detection, through which we can clearly see the effect of gamma transformation.

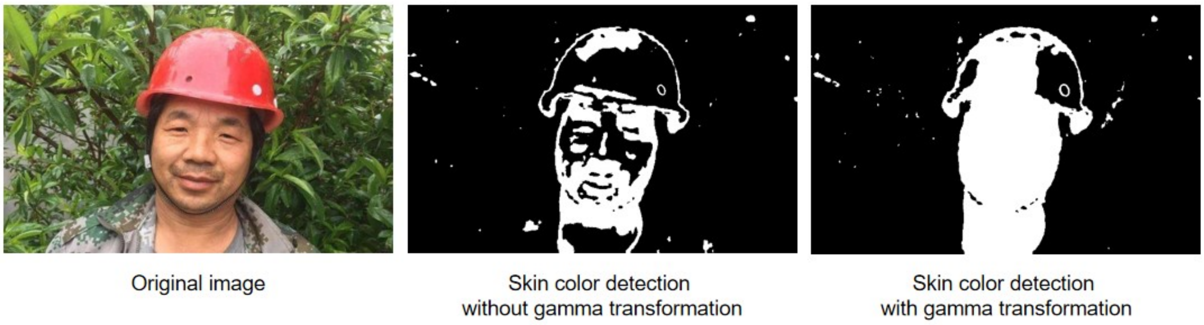

Figure 4. Contrast with or without gamma transformation in Skin color detection

3.2. **Two-Stage ROI Extraction.** Based on the result of image preprocessing, the two-stage ROI for the helmet in this paper is extracted by the combination of face detection, skin color detection and helmet contour detection.

3.2.1. *Skin Color Detection.* The YCrCb color space of skin color is a commonly used color model for skin color detection. The skin information is mapped to the YCrCb space, and the skin pixels in the CrCb two-dimensional space will be approximately distributed as an ellipse[11-12], as shown in Fig.5.

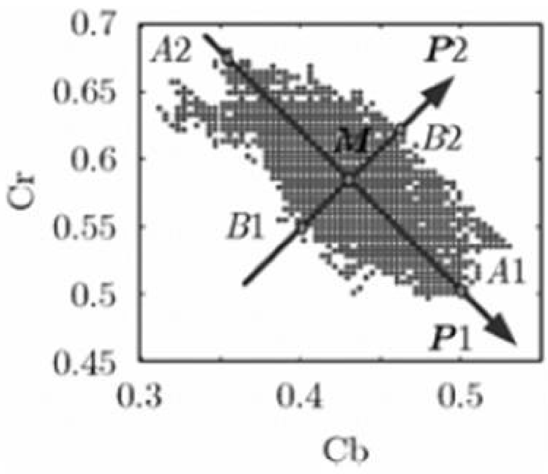

Figure 5. Ellipse Skin color Model in YCrCb Color Space

Therefore, if we get a CrCb ellipse, convert the image to YCrCb color space and set a two-dimensional all-zero matrix of the same size as the input image, and then traverse the pixel values of the entire image. For the coordinates (Cr, Cb) of the pixels, we only need to judge whether it is inside the ellipse. If it is, then it can be judged as a skin pixel, and the corresponding matrix value is set to 1; otherwise, it is a non-skin pixel, and the corresponding matrix value remains zero.

3.2.2. *Face detection based on Haar Feature.* The Haar algorithm that is used to detect faces includes the following steps: Haar classifier training, Haar feature selection, Adaboost algorithm and Cascade classifiers.

The Haar classifier is trained by a large number of images. Haar feature selection uses a sliding window of simple rectangular blocks to traverse the image and collect Haar features from it. As shown in Fig.6, Haar features can be divided into three categories: edge features, linear features, and center features.

When recognizing a certain part of the face, not all features are relevant. For example, features suitable for the nose are not suitable for the mouth. At this time, we can use an algorithm called Adaboost which uses a weak classifier to construct a strong classifier by assigning a higher weighted penalty to incorrect classification[13-14].

3.2.3. *Contour Feature.* Canny edge detection, which includes edge gradient calculating, non maximum gradient point suppression and edges generating, is optimal for step edges affected by white noise[15-16].

The edge gradient calculating use the Sobel operator to calculate the gradient value of the image in the x and y directions, and finally calculate the gradient value and gradient

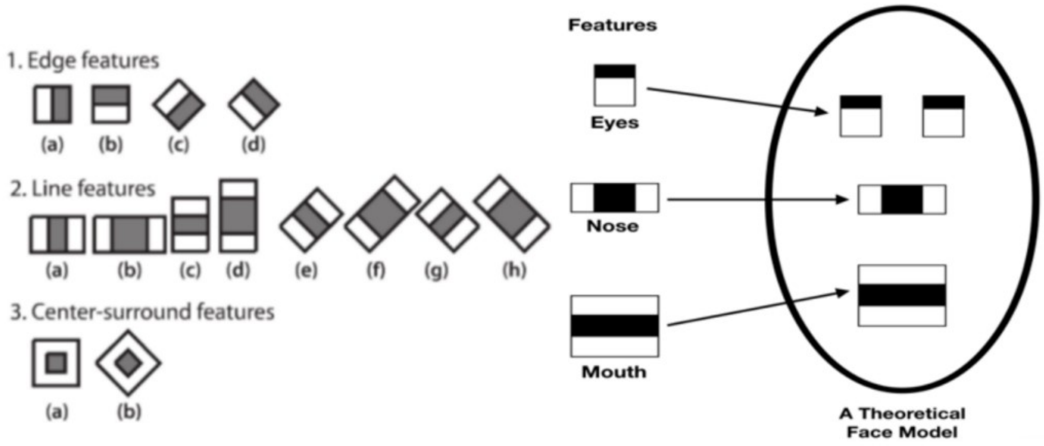

FIGURE 6. Haar Features

angle of each point of the image. The calculated gradient angle needs to be approximated, approximately four values -45 (or 135), 0, 45, 90.

Non-maximum gradient value point suppression traverses each point and does the following operations:

(1) If the gradient angle of the point (x,y) is 0, when its gradient value is greater than the gradient value of (x-1,y) and (x+1,y), it is considered (x,y) Point is an edge point, otherwise its value is suppressed, and its gradient value is the set background value;

(2) If the gradient angle of the point (x,y) is 90, when its gradient value is greater than (x,y-1) and (x,y+1) have large gradient values, then point (x,y) is considered an edge point, otherwise, its value is suppressed, and its gradient value is the set background value;

(3) If the gradient angle of the point (x,y) is 135 (or -45), when its gradient value is greater than the (x-1,y+1) and (x+1, if the gradient value of y-1) is large, the point (x,y) is considered to be an edge point, otherwise, its value is suppressed, and its gradient value is the set background value;

(4) If the gradient angle of the point (x,y) is 45, when its gradient value is larger than the gradient value of (x-1,y-1) and southeast (x+1,y+1), the point (x,y) is considered An edge point, otherwise suppress its value, and its gradient value is the set background value.

Based on the results of the edge gradient calculating, we use the edge value and background value to divide the two image regions to extract the edge features of the image. At last, we treat the contour area with more than 10 sides as the ROI.

3.3. **Color Feature Recognition.** Since the image is more adapted to color segmentation in HSV color space, the extracted ROI of the original image would be converted to HSV color space and filtered by the range of different color helmets to determine whether the person in the image is wearing a helmet or not.

4. **Results and Discussions.** This section begins with the introduction of the dataset, followed by showing the ROI extraction results and helmet-wearing detection results in single or multiple person images. All experiments were conducted on a laptop equipped with Core-i5 and 16-GB RAM, and the implementation and experiment of the algorithm were carried out using Python 3.8 and Pycharm 2019.3.3 x64 version. Finally, we discussed the significance and limitations of the proposed method.

4.1. **Dataset.** In order to effectively evaluate the proposed framework of the helmet-wearing detection, the public dataset SHWD was selected as the test images in this paper, which contains 7581 images, including 3241 positive images (containing frontal single person images, multi-pose single person images, multi-person images) and 4340 negative images. 1999 of these negative images are the frames of classroom video, and the remaining 2341 are portrait images of other scenes on campus.

It is worth mentioning that the sources of the dataset are extensive, such as web pictures, screenshots, and the photos taken from construction sites and campuses, which makes the background of these images different and complex, and improves the difficulty of helmet wearing detection. Fig.7 shows part of the image of dataset SHWD. The good news is that the complex background makes the experimental results more convincing and universal.

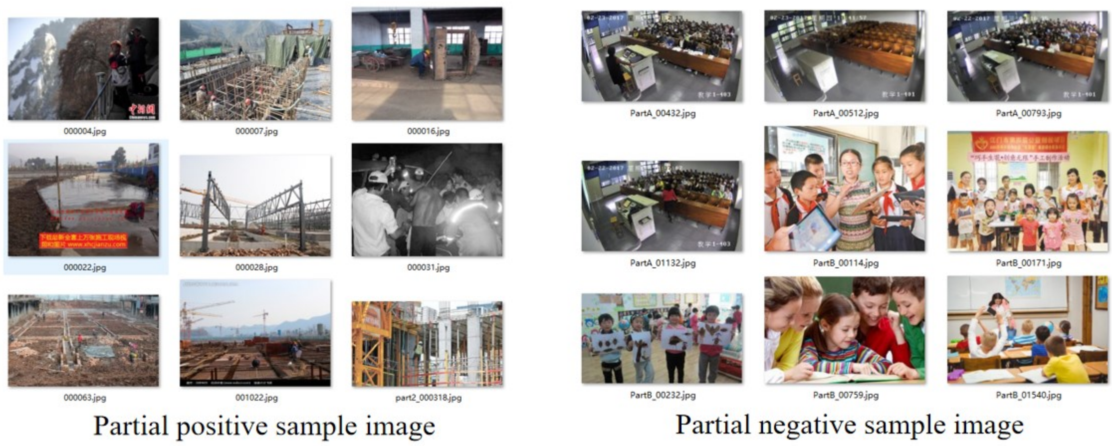

FIGURE 7. Part of the image of dataset SHWD

4.2. **Experimental Results and Analyses.** ROI extraction is the critical step of the whole algorithm. As is shown in Fig.8, the ROI of the original image including single person and multiple person are extracted respectively. It needs to be pointed out that the ROI in the image containing multiple persons are extracted one by one. It can be seen that the ROI extraction in this paper achieved good results on various images with complex background.

Based on the results of ROI extraction, the performance of helmet-wearing detection is shown in Fig.9, which shows that the helmet wearing of single person can be detected pretty well. However, the wearing of helmets in the images containing multiple persons is not detected sometimes. After analysis, false detection mainly occurs in ROI extraction. We carefully checked the pictures where the error occurred and found that they are either far away from the target or the background is too complicated. In addition, the color of the helmet is similar to the background sometimes, which can also cause false detection. Therefore, the method proposed in this paper still has much space for improvement.

Above is the qualitative analysis of the proposed method in this paper, and then we use the Accuracy, Precision and Recall, which are respectively defined in Eq.(5), Eq.(6) and Eq.(7) to evaluate the performance of the framework. The Accuracy is defined as the ratio of correct judgments to total detected images, and the Precision is defined as the ratio of correct detected helmet wearing images to total detected helmet wearing images; and the Recall means the ratio of correct detected helmet wearing images to the total number of positive samples.

FIGURE 8. The ROI extraction results

$$Accuracy = \frac{T}{T+F} \tag{5}$$

Where T is the number of right detections, which includes judging the positive samples as wearing helmets and judging the negative samples as not wearing helmets, F denotes the number of false detections, and T+F means the total number of detected images.

$$Precision = \frac{TP}{TP+FP} \tag{6}$$

$$Recall = \frac{TP}{TP+FN} \tag{7}$$

Where TP denotes the number of correct detection results of positive samples, FP represents the number of false detection results of negative samples which treat the image of no-helmet-wearing as worn; and FN denotes the number of incorrect detection results which treat the image of helmet wearing as not wearing; and TP +FP means the number of images with positive predictions, and TP+FN means the number of positive images.

In addition to the above mentioned, we calculate the F1 score using Eq.(8), to synthesize the Precision and Recall as a new evaluating indicator.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{8}$$

(a) Single person helmet wearing detection

(b) Multiple persons helmet wearing detection

FIGURE 9. Helmet wearing detection results

we tabulated the evaluation based on SHWD in Table. 1. From the table, we can conclude that helmet wearing detection in conjunction with the two-stage ROI extraction and color feature recognition can provide Satisfactory results.

TABLE 1. Performance evaluation based on SHWD

Accuracy	76.709%
Precision	70.409%
Recall	64.596%
F1	67.377%

4.3. **Discussion.** The work of this paper is meaningful because helmet wearing is very crucial for the life safety of construction workers. A large number of experiments show that the proposed framework is effective for this target.

However, the method proposed in this paper also faces some challenges. For example, the background of the detected image is intricate or similar to the targets, the detection of helmet wearing may fail. In addition, the helmet-wearing detection of persons in complex poses needs to be strengthened.

5. **Conclusion and Future Works.** In many scenarios, helmet-wearing plays a vital role, thus the detection of helmet wearing has certain practical significance for assisting construction safety. The algorithm in this paper mainly realizes the detection of wearing helmets in batches of images, which can meet some detection requirements, but for some real-time detection, this paper has not covered. The idea of future research is to realize the

real-time detection of helmet wearing and set the non-wearing alarm prompt, to complete a safety helmet wearing assistance system which can be practically applied.

Acknowledgment. This work is partially supported by Science Foundation of Zhejiang Sci-Tech University(ZSTU). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] N. DALAL, B. TRIGGS, Histograms of oriented gradients for human detection, *Computer Vision and Pattern Recognition*, [S.l.]:IEEE, pp. 886-893, 2005.

[2] L. Jie, H. Liu, T. Wang, et al. Safety helmet wearing detection based on image processing and machine learning, *2017 Ninth International Conference on Advanced Computational Intelligence (ICACI)*, 2017.

[3] M. Jin et al, Safety Helmet Detection Algorithm based on Color and HOG Features, *2020 IEEE 19th International Conference on Cognitive Informatics* & *Cognitive Computing (ICCI*CC)*, pp. 215-219, 2020.

[4] Q. Wu, B. Zang, Y. Zhang, Z. Qi, Wavelet Kernel Twin Support Vector Machine, *Journal of Information Hiding and Multimedia Signal Processing*, Vol. 12, No. 2, pp. 93-101, June 2021.

[5] R. Waranusast, N. Bundon, V. Timtong, et al. Machine vision techniques for motorcycle safety helmet detection, *IEEE*, pp. 35-40, 2013.

[6] H. Wang, Z. Hu, Y. Guo, et al. A Combined Method for Face and Helmet Detection in Intelligent Construction Site Application, *Recent Featured Applications of Artificial Intelligence Methods*, LSMS 2020 and ICSEE 2020 Workshops, 2020.

[7] L. Kang, X. Zhao, B. Jiang, et al. Automatic Safety Helmet Wearing Detection, *The State Key Laboratory of Management and Control for Complex System, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Institute of Automation, Beijing, China;The State Key Laboratory of Management and Control for Complex System*, Un, 2018.

[8] C. Gang, Z. Yao, R. Ni, Forensic estimation of gamma correction in digital images, *IEEE International Conference on Image Processing*, IEEE, 2010.

[9] M.F. Doustar, H. Hassanpour, A locally-adaptive approach for image gamma correction, *International Conference on Information Sciences Signal Processing* & *Their Applications*, IEEE, 2010.

[10] W. Ping, F. Liu, C. Yang, et al. Blind Forensics of Image Gamma Transformation and Its Application in Splicing Detection, *Journal of Visual Communication* & *Image Representation*, 55:80-90, 2018.

[11] S.L. Phung, A. Bouzerdoum, D. Chai, A novel skin color model in YCBCR color space and its application to human face detection, *Image Processing*, 2002. Proceedings. 2002 International Conference on. 2002.

[12] A.Y. Da Wod, J. Abdullah, M.J. Alam, Adaptive skin color model for hand segmentation, *International Conference on Computer Applications* & *Industrial Electronics*, 2011.

[13] R. Lienhart, J. Maydt, An extended set of Haar-like features for rapid object detection, *Proceedings. International Conference on Image Processing*, IEEE, 2002.

[14] P. Viola, Rapid object detection using a boosted cascade of simple features, *Computer Vision and Pattern Recognition*, 2001.

[15] Z. Xu, B. Xu, G. Wu, Canny edge detection based on Open CV, *2017 13th IEEE International Conference on Electronic Measurement* & *Instruments (ICEMI)*, IEEE, 2017.

[16] L. Xuan, Z. Hong, An improved canny edge detection algorithm, *2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, pp. 275-278, 2017.