Safety helmet wearing status detection based on improved boosted random ferns

Shiqin Yue [1,3] · Qian Zhang [1,2] · Dingqin Shao [1,3] · Yu Fan [1,3] · Jinhua Bai [1,3]

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

The safety helmet wearing of workers is extremely important to their safety in construction scenarios, and it is very meaningful for computer vision, pattern recognition and artificial intelligence. This paper proposes a new Improved Boosted Random Ferns algorithm (IBRFs) for safety helmet wearing status detection. IBRFs originates from the Boosted Random Ferns algorithm (BRFs) and introduces the weighted coefficient to improve. In IBRFs, firstly, the feature is extracted by Histogram of Oriented Gradient (HOG) to form the feature domain space of the image. Secondly, the random binary test method is used to construct random ferns in the feature domain space. Then, a weak (acceptable) classifier is constructed by random ferns. Finally, an improved Real AdaBoost algorithm is used to select the most discriminative ones to construct IBRFs. Experimental evaluation on an enlarged public Safety Helmet Wearing-datasets (GZMU-SHWD) shows that the result of IBRFs outperforms those of the existing advanced detection algorithms, including SSD, YOLOv3 and Faster R-CNN, which further demonstrates the effectiveness of IBRFs for safety helmet wearing status detection.

1 Introduction

The safety helmet, as the basics safety protection equipment for constructers, is mainly used to protect workers' head and prevent them from being hit and collided by sharp objects [18]. It is

✉ Qian Zhang
  gzmuzq@gzmu.edu.cn

1 School of Data Science and Information Engineering, Guizhou Minzu University, Guiyang 550025, China

2 Academic Affairs Office, Guizhou Minzu University, Guiyang 550025, China

3 Key Laboratory of Pattern Recognition and Intelligent Systems of Guizhou Province, Guiyang 550025, China

of great significance to the personal safety of constructers. Therefore, it has become a hot topic of artificial intelligence to detect the wearing status of safety helmet, in different construction scenarios, based on computer vision technology, especially in pattern recognition community.

The task of object detection is to find out the interested objects in the image and determine their classes and locations. We form the safety helmet wearing status detection as two class object detection problems. In pattern recognition community, there are two main methods, traditional statistical pattern recognition and deep learning, to detect the safety helmet wearing status.

In traditional object detection, firstly, some candidate regions are selected from a given image. Then, feature extraction is carried out for these regions, the commonly used features are Histogram of Oriented Gradient (HOG) [4], Local Binary Patterns (LBP) [12], Scale-Invariant Feature Transform (SIFT) [10] etc. Finally, the pretrained classifier is used for classification. Limei Cai et al. constructed standard images of helmets, extracted four directional features, modeled the distribution of these features using a Gaussian function and separated local images of frames into helmet and non-helmet classes [2]. Bahaa et al. Detect the helmet in different scenarios [1]. Park et al. proposed a background difference method for detection [13]. After extracted HOG features from the input image, SVM classifier is used to detect person and safety helmet.

There are two problems in the existing safety helmet wearing detection based on traditional methods. One is the region selection strategy based on sliding window lack of pertinence, with high time complexity and redundant window. Another is the hand-designed features have poor robustness to the changes of diversity.

In deep learning method, after multiple convolutional and pooling operations, the extracted features are put into the classifier for detection. There are two distinct ideas of deep learning detection. One is using semantic segmentation [11, 27] for reference, which mainly includes You Only Look Once (YOLO) [14], YOLOv3 [17] and Single Shot MutiBox Detector (SSD) [19], predicting the classes and locations of different objects with convolutional neural network. This kind of method can achieve higher accuracy while slower detection speed. Another is the R-CNN framework based on candidate region, which mainly includes R-CNN [8], Fast R-CNN, and Faster R-CNN [15]. Within this framework, the object candidate box is generated for classification and regression. This framework has strong generalization ability and has been widely used. Cheng Raoet al. proposed a helmet wearing detection algorithm based on SAS-YOLOv3-Tiny [3]. Fang Ming et al. proposed a fast detection algorithm based on YOLOv2 [5].Added dense blocks, and compressed the network by adopting the lightweight network structure, to realize the rapid detection of safety helmet. Wen P et al. proposed a method based on the YOLOv3 algorithm to detect the wearing of safety helmet [24]. By using multi-scale training and increasing the anchor points to enhance the robustness of network for different sizes object detection, the Fast R-CNN is optimized.

In general, deep learning can achieve high accuracy, however a lot of samples, high hardware, and long time is need for training. Moreover, there is no public large-scale safety helmet datasets. In this way, the unbiased estimation based on deep learning is in efficient on small samples.

In order to realize the accurate detection of helmet wearing status, this paper proposes an new IBRFs. Based on HOG feature to construct random ferns, then weak classifiers are constructed. Selected the most discriminative ones to build a strong classifier to detect the wearing status of the safety helmet. The main contirbutions of our work are summarized as follows:

- Feature extraction of image based on HOG domain solves the inaccurate feature description.
- Weak classifiers are constructed based on BRFs, by selecting the position points and parameters with strong discriminative ability to improve their performance.
- An improved ensemble learning algorithm is proposed to enhance the discriminative ability of strong classifiers.

The rest of this paper is composed of 4 sections. Section 2 briefly describes the theory of random ferns. Section 3 presents the proposed IBRFs. Section 4 submits experimental validation of IBRFs on GZMU-SHWD. And section 5 concludes the work and points out the future work.

2 RandomFerns Algorithm

Random Ferns algorithm (RFs) is an ensemble learning method that performs well in classification and regression tasks in machine learning [21, 25]. As shown in Fig. 1, RFs takes a particular decision tree as the basic meta-model, and there is only one judgment criterion in each layer of fern. In the random ferns, the training time increases linearly as the number of ferns increases, so the calculating feasibility can be confirmed.

2.1 Random ferns cluster based on HOG domain

In order to improve the robustness of the random ferns classifier for illumination and intra-class variation, we extract image HOG features referencing to the feature representation based on the intensity domain in the RFs [20]. Given an image window X, HOG is calculated in the center position U of the sub-window S. Each local gradient in the sub-window S contributes to the construction of the HOG, and is positioned in the corresponding box through weighting processing. In the HOG, two boxes are randomly selected for binary test:

$$f(x; u, \theta) = I\left(HOG(x; u, b) > HOG\left(x; u, b'\right)\right) \tag{1}$$
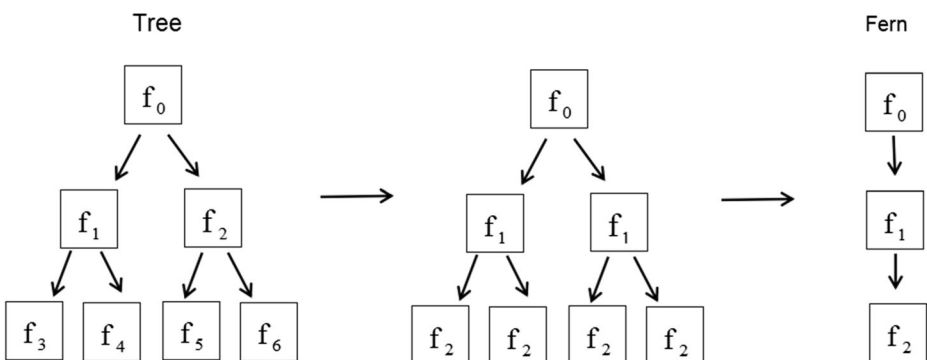

Fig. 1 Tree classifier and Fern classifier

Where $I(a)$ is the index function, if a is true, then $I(a) = 1$, otherwise $I(a) = 0$. $\theta = \{b, b'\}$ is two randomly selected d indexes in the interval $[1, B]$, and B is the total number of cases in the histogram. A binary feature vector of fern is synthesized by aggregating M local features to represent S belong to one class. Therefore a single random fern $f(x; u, \theta)$ based on HOG domain is represented as follows:

$$f(x; u, \theta) = [f(x; u, \theta_1), , f(x; u, \theta_M)] \qquad (2)$$

In Fig. 2, the right one is random ferns based on HOG, each output of the random ferns is a M dimensional binary feature vector Z, $Z \in 0, …, 2^M - 1$. In the image window X, When $M = 3$, the output feature of random fern s is $Z = \{011\} = 3$, $Z = \{0, …, 255\}$. The output of the random ferns is determined by u and $\theta = \{\theta_1, …, \theta_M\}$.

As shown in Fig. 3, when training N samples in the same class, random ferns will output a probability distribution $P(F| C)$. For training samples with multiple classes, random ferns will output probability distribution $P(F| C_k)$ of each class, C_k denotes the probability of the input sample belongs to class k. When testing a sample with an unknown class, using a fixed-size rectangular sliding window to scan it, its feature is calculated by random ferns. Semi-naive Bayes is used to get the final classification.

$$Class(f) = \text{argmax}_k P(C_k) \prod_{n=1}^{N} p(F_n | C_k) \qquad (3)$$

2.2 Boosted random ferns

In order to improve the performance of the classifier, BRFs [22] adopts the Real AdaBoost strategy [7] to select T weak classifiers $h_t(x)$, $t = 1, …, T$ with strong discriminative ability to build a strong classifier $H(x)$:

$$H(x) = sign\left(\sum_{t=1}^{T} h_t(x) - \beta\right) \qquad (4)$$

Where β is a threshold that determines the classifier tolerance, each weak classifier returns a confidence score:
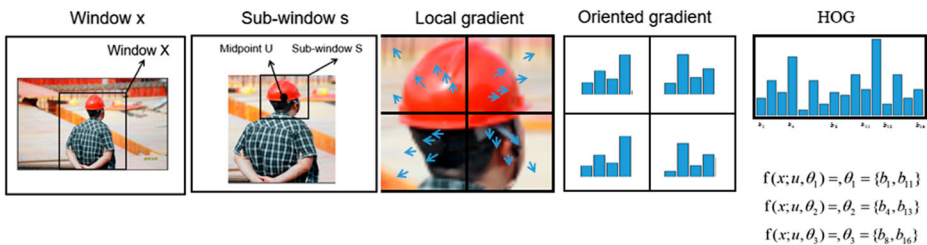

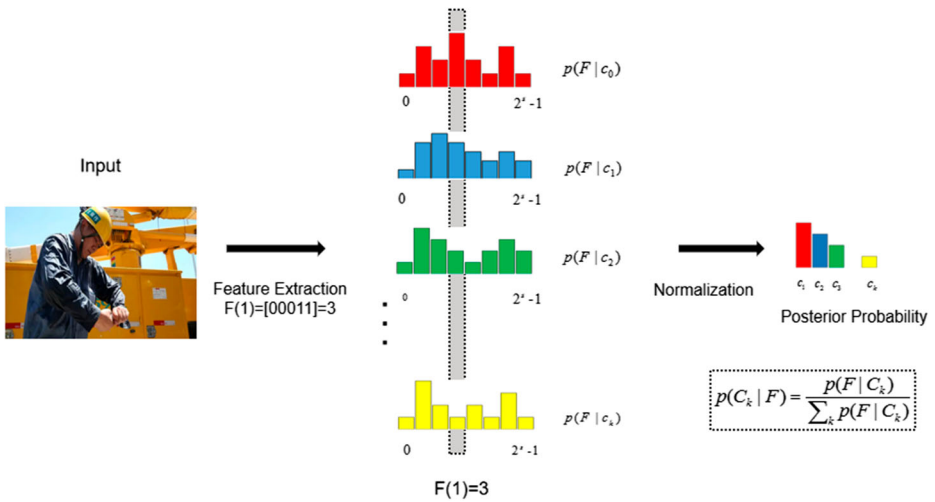
Fig. 2 Random Ferns based on HOG

Fig. 3 Random Ferns classifier

$$h_t(x) = \frac{1}{2}\log(o(x; u_l, \theta_r)) \tag{5}$$

$$o(x; u_l, \theta_r) = \frac{P(z|O) + \varepsilon}{P(z|B) + \varepsilon} \tag{6}$$

Where ε is a small positive integer to avoid zero division. Higher score weak classifier $h_t(x)$ denotes there is a significant difference between object and background. It means the random ferns can produce a weak classifier with recognition ability on the location of u and parameter selection θ.

In order to get the most suitable u and θ for each weak classifier, using N samples $S = \{(x_1, y_1), \ldots, (x_N, y_N)\}$ to train the classifier, where $y_i \in \{-1, +1\}$, $y_i = -1$ represents the image window x_i belongs to the background class, and $y_i = +1$ represents x_i belongs to the object. Every training sample also has an associated weight $\omega_t(i)$, it is initialized as $\omega_1(i) = \frac{1}{N}$. The main steps of BRFs as follows:

Specifically, we first define a set $\{u_1, \ldots, u_L\}$ of all possible 2D pixel coordinates within a window X and a pool $\{\theta_1, \ldots, \theta_R\}$ of R different sets of random pair of histogram bin indices.

Then, a weak classifier is constructed. Each iteration of a possible random pair (u_l, θ_r) generates a random fern $f(x; u_l, \theta_r)$. Each random fern is evaluated according to the whole training datasets, and HOG of object and background was constructed.

$$P(z|O) = \sum_{i:f(x_i;u_l,\theta_r)=z^\wedge y_i=+1} \omega_t(i) \tag{7}$$

$$P(z|B) = \sum_{i:f(x_i;u_l,\theta_r)=z^\wedge y_i=-1} \omega_t(i) \tag{8}$$

Where $P(z|\ O)$ stands for $P(f(x_i;\ u_l,\ \theta_r)|\ O)$, the output value z of random ferns under the parameters u_t and θ_t represents the probability that the image window x belongs to the object. $P(z|\ B)$ stands for $P(f(x_i;\ u_l,\ \theta_r)|\ B)$. It is the probability that the image window belongs to the background.

The discriminative ability of each weak classifier $h_t(x)$ is determined by the Bhattacharyya coefficient [16] Q between $P(z|\ O)$ and $P(z|\ B)$.

$$Q\ (u_l, \theta_r) = 2 \sum_{z=0}^{2^M-1} \sqrt{P(z|O)P(z|B)} \tag{9}$$

The value of Q is lower, the discriminative ability of weak classifier is stronger. A weak classifier $h_t(x)$ is constructed by retaining the parameters (u_{l^*}, θ_{r^*}) that minimize Q.

$$(u_{l^*}, \theta_{r^*}) = \mathrm{argmin}_{l,r} Q\ (u_l, \theta_r) \tag{10}$$

$$h_t(x) = \frac{1}{2} \log o(x; u_{l^*}, \theta_{r^*}) \tag{11}$$

Finally, sample weight is updated:

$$\omega_{t+1}(i) = \frac{\omega_t(i)\exp(-y_i h_t(x_i))}{\sum_{i=1}^{N} \omega_t(i)\exp(-y_i h_t(x_i))} \tag{12}$$

3 Algorithm design

3.1 Improved boosted random ferns algorithm

In order to enhance the detection accuracy, the Real AdaBoost method [7] is introduced in BRFs [22]. It selects the weak classifier $h_t(x)$ with strong discriminative ability, and obtains the classifier $H(x)$ by accumulating sum.

This method simply combines multiple weak classifiers, and pays less attention to weak classifiers with low weight. In order to achieve optimal combination of classifiers, this paper enhances the BRFs by Real AdaBoost with the help of a new definition method [23]. Random variable is defined as:

$$D_t = \begin{cases} h_t(x), & if\ y = +1,\ x \in s \\ -h_t(x), & if\ y = -1,\ x \in s \end{cases} \tag{13}$$

Let $D = \sum_{t=1}^{T} \gamma_t D_t$, then weighted combination is used to construct strong classifier $H(x)$:

$$H(x) = sign\left(\sum_{t=1}^{T} \gamma_t h_t(x) - \beta\right) \tag{14}$$

$$\gamma_t = \frac{\mu_t}{\delta_t^2} \tag{15}$$

Where γ_t is the ratio of mean μ_t and variance δ_t^2 of random variable D, μ_t and δ_t^2 are expressed as:

$$\mu_t = \frac{1}{2} \sum_{i=1}^{N} \left((P(z|O) - P(z|B)) \log\left(\frac{P(z|O)}{P(z|B)}\right) \right) \tag{16}$$

$$\delta_t^2 = \sum_{i=1}^{N} \left(P(z|O) \left(\frac{1}{2} \log\left(\frac{P(z|O)}{P(z|B)}\right) - \mu_t \right)^2 + P(z|B) \left(\frac{1}{2} \log\left(\frac{P(z|O)}{P(z|B)}\right) + \mu_t \right)^2 \right) \tag{17}$$

After introducing the weighted coefficient, the sample weight is updated as follows:

$$\omega_{t+1}(i) = \frac{\omega_t(i)\exp(-y_i\gamma_t h_t(x_i))}{\sum_{i=1}^{N} \omega_t(i)\exp(-y_i h_t(x_i))} \tag{18}$$

IBRFs is show as algorithm 1:

Input:N sample $(x_1, y_1), \dots, (x_N, y_N)$

Output: IBRFs

---

1 Initialize sample weight $\omega_1(i) = \frac{1}{N}, i = 1, \dots, N$

2 Create a 2D pixel position $u_1, \dots, u_L$ in the sample window x

3 Create a set of histogram random index pairs $\theta_1, \dots, \theta_R$

4 For $t = 1, \dots, T$

5 ......For $l = 1, \dots, L$

6 ....For $r = 1, \dots, R$

7 ....Calculate cumulative sample weight $P(z|O), P(z|B)$

8 ....Calculate Bhattacharyya coefficient $Q(u_l, \theta_r)$

9 ....Calculate weighted coefficient γ_t

10 ......Select the optimal parameters (u_{l^*}, θ_{r^*}) to construct classifier $h_t(x)$

11 ......Save probability histogram $P(z|O), P(z|B)$

12 ......The classification probability is evaluated on the image window and the sample weight is updated $\omega_{t+1}(i)$

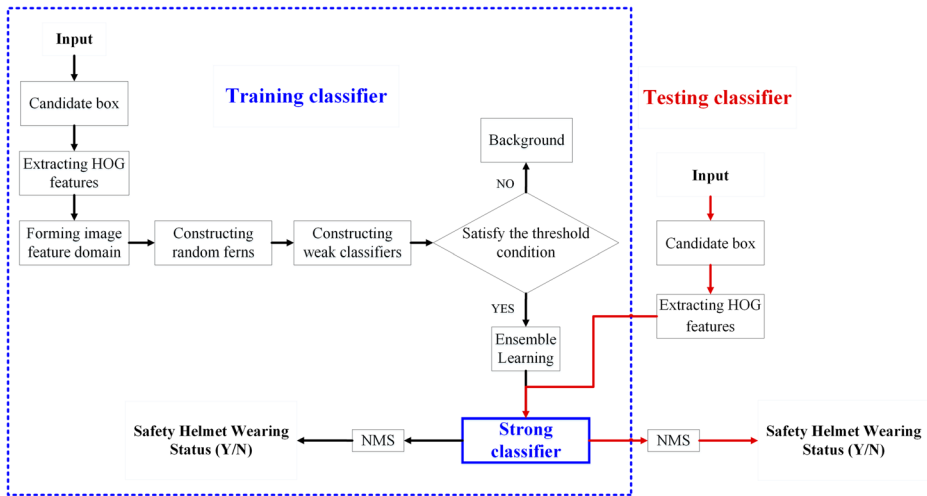13 Construct strong classifier $H(x) = sign(\sum_{t=1}^{T} \gamma_t h_t(x) - \beta)$

Fig. 4 Pipeline framework of the IBRFs

3.2 Algorithm implementation

As shown in Fig. 4, the pipeline framework of IBRFs is given as follows.

At the training stage, for an input image, firstly, the feature is extracted by HOG with candidate box to form the feature domain space. After that, the random binary test is used to construct the random ferns. Next, a weak classifier is built by the random ferns. At last, the improved Real AdaBoost algorithm is used to select the most discriminative ones to construct IBRFs. At the testing stage, the HOG for input image with candidate box is obtained, and estimated by the IBRFs to complete the safety helmet wearing status detection.

4 Experimental results and analysis

4.1 Datasets and experimental platform

The experimental environment of this paper is as follows: window10, 64-bit operating system, Intel(R) Core(TM) i5-8300H CPU @ 2.30GHz, 8G RAM.

In this experiment, an enlarged Safety Helmet Wearing-datasets (SHWD[1]) called GZMU-SHWD is taken to evaluate the IBRFs. The GZMU-SHWD originates from Internet crawler data (Google-Net crawler) and SHWD. It consists of 7589 images, including 5000 training samples, 2481 test samples and 108 verification samples. Some data are shown in Fig. 5.

The image comes from the different scenarios, different lighting conditions and different angles, and the sample distribution is uneven. Therefore, image preprocessing is performed. The image is augmented by horizontal flipping and clipping, and normalized to 100*120.

---

[1] The SWHD database: [Online]. Available: https://github.com/njvisionpower/Safety-Helmet-Wearing-Dataset#model, October 11, 2019.

Fig. 5 Some images fromGZMU-SHWD

4.2 Experimental parameter setting

In the training phase, some parameters of the model are fine-tuned with reference to BRFs [22]. The experimental parameters are shown in Table 1.

4.3 Experimental results

In order to evaluate the effectiveness of the proposed IBRFs, precision and recall are used in the measuring the effectiveness of the proposed method for safety helmet wearing status detection. The experimental evaluation formula is given as:

$$precision = \frac{TP}{TP + FP} \tag{19}$$

$$recall = \frac{TP}{TP + FN} \tag{20}$$

Table 1 Parameters

Parameter name	Parameter value
Epoch	50
Oriented box	4
Feature number (M)	8
Feature sharing	10
Number of weak classifiers (T)	300
Overlap threshold	0.3
Detection threshold	0.1020

Table 2 Comparative analysis of different ensemble algorithms

	UIUC Car		GZMU-SUWD		INRIA horse	
	BRFs	IBRFs	BRFs	IBRFs	BRFs	IBRFs
FP	4	2	20	11	12	9
FN	5	1	19	9	10	10
TP	134	138	105	115	35	35
Recall/%	96.40	99.28	84.00	91.27	77.78	77.78
Precision/%	97.10	98.57	84.68	92.74	74.47	79.55
EER	96.75	98.93	76.36	84.38	92.00	78.66

Where True Positive (TP) is the positive sample predicted to be positive, False Positive (FP) is the negative sample predicted to be positive, False Negative (FN) is the negative sample predicted to be negative.

4.3.1 Comparative analysis of different ensemble algorithms

The weighted coefficient is used to improve the Real AdaBoost algorithm to get the best combination coefficient. In order to verify the superiority of improved Real AdaBoost algorithm for classification, the algorithms are tested on Ionoshpere and Sonar datasets [23] and compared with other algorithms. The results are shown in Table 2.

As show in Table 2, the detection error rate and variance of Improved Real AdaBoost are lowest. On Ionoshpere datasets, the detection error rate of Improve Real AdaBoost is 0.0939, which is 0.0129 lower than Gentle AdaBoost, and 0.0111 lower than Real AdaBoost. This verifies the effectiveness of the Improve Real AdaBoost algorithm to help the IBRFs algorithm in the classification problem.

4.3.2 Analysis of BRFs and IBRFs

In order to verify the effectiveness and generalization ability of IBRFs.In the same experimental environment, the BRFs and IBRFs are trained on public datasets include UIUC cars [26], INRIA horses [6], and GZMU-SHWD. There are 108 images in UIUC cars datasets, including 139 detection objects. There are 35 images in the INRIA horses datasets, including 45 detection objects. There are 108 images in GZMU-SHWD datasets, including 124 detection objects. The detection results of the two models are shown in Table 3.

Table 3 Compares the performance of BRFs and IBRFs on different datasets, and the detection accuracy of the IBRFs on INRIA horses datasets, UIUC cars datasets and GZMU-

Table 3 Comparative analysis of different datasets

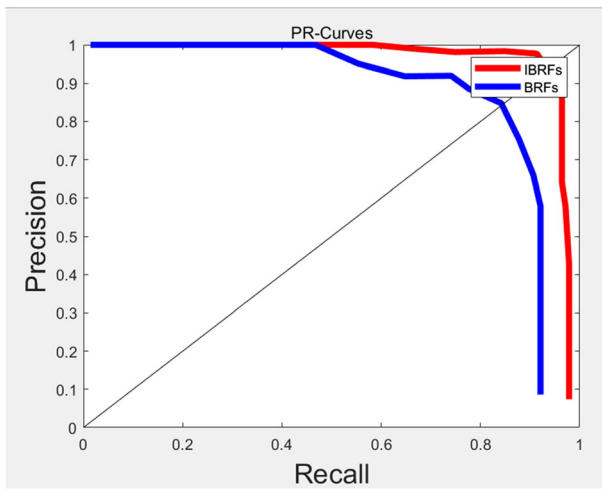|	Ionoshpere		Sonar	
	Error	Variance	Error	Variance
Real AdaBoost	0.1068	0.0237	0.2346	0.0392
Improved Real AdaBoost (Method of this paper)	**0.0939**	**0.0174**	**0.2300**	**0.0412**
Gentle AdaBoost	0.1050	0.0210	0.2337	0.0372

Fig. 6 Precise and Recall curves

SHWD datasets is improved compared with the BRFs, especially on GZMU-SHWD, the detection accuracy reaches 92.74%, which is about 9.5% higher than the BRFs, The results show that the IBRFs is effective to detect the wearing status of safety helmet.

Figure 6 is the PR (Precision and Recall) curves of GZMU-SHWD, it can be seen that the IBRFs is better than BRFs in detection accuracy and recall rate. The results show that the weighted coefficient introduced in this paper can improve the detection accuracy of the model.

The IBRFs can distinguish and select random ferns during the detection process, so that the classifier can focus on the object parts that are more relevant to the classification. In Fig. 7, we observe the results of feature selection for the three different object classifiers. The first row depict the spatial layout of the ferns used to ensemble the classifiers. The colored contour boxes indicate the positional density and weight of individual ferns that give rise to the strong classifier. Red contours indicate higher density of ferns. We can see how ferns concentrate on those semantically relevant regions, such as the wheels of cars, or the neck and head on horses.
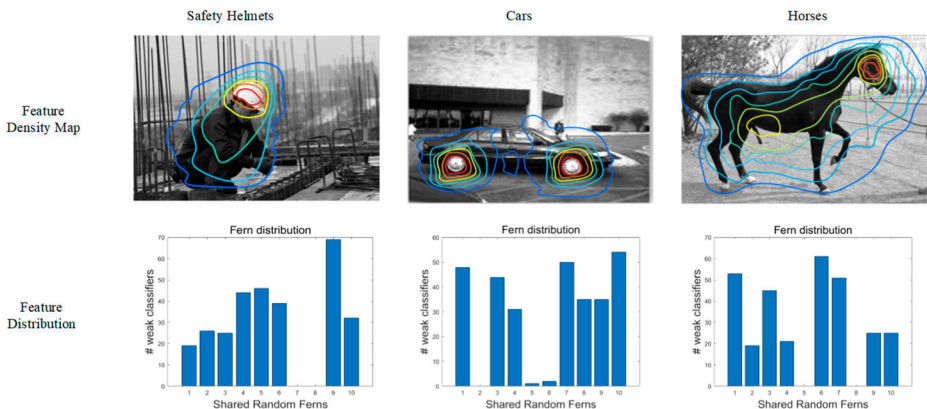

Fig. 7 The first row denotes spatial layout of ferns for different object categories.The second row denotes the distribution of weak classifiers for each class

Table 4 Comparison of different algorithms

Method	Recall/%	Precision/%	frame(f.s⁻¹)
HOG+SVM	71.95	79.83	#
SSD	87.89	88.23	38.0
Faster R-CNN	90.45	91.42	0.6
YOLOv3	92.56	94.13	48
BRFs	84.0	84.68	7
IBRFs	**91.26**	**92.74**	**9**

The second row contains shared ferns distributions map for each strong classifier, the height of the i column indicates the number of weak classifiers that use the parameters θ_i .

4.3.3 Analysis of experimental results

In order to verify the superiority of our algorithm, compared it with other algorithms in the same hardware environment. IBRFs, BRFs, SVM + HOG [9], SSD [19], Faster R-CNN [15] and YOLOv3 [17] are tested on GZMU-SHWD. The detailed detection results are shown in Table 4.

From Table 4, in the safety helmet wearing status detection task, the detection accuracy of IBRFs is 92.74%, which is higher than SSD, Fast R-CNN, BRFs, especially HOG + SVM. Compared with BRFs, the detection accuracy is improved by 9.5%, which verifies the effectiveness of IBRFs. In CPU environment, the detection speed of IBRFs is about 15 times faster than Fast R-CNN, compared with BRFs, the detection speed of IBRFs is also improved. It verifies IBRFs meets the requirements of real-time detection. Above all, in the helmet
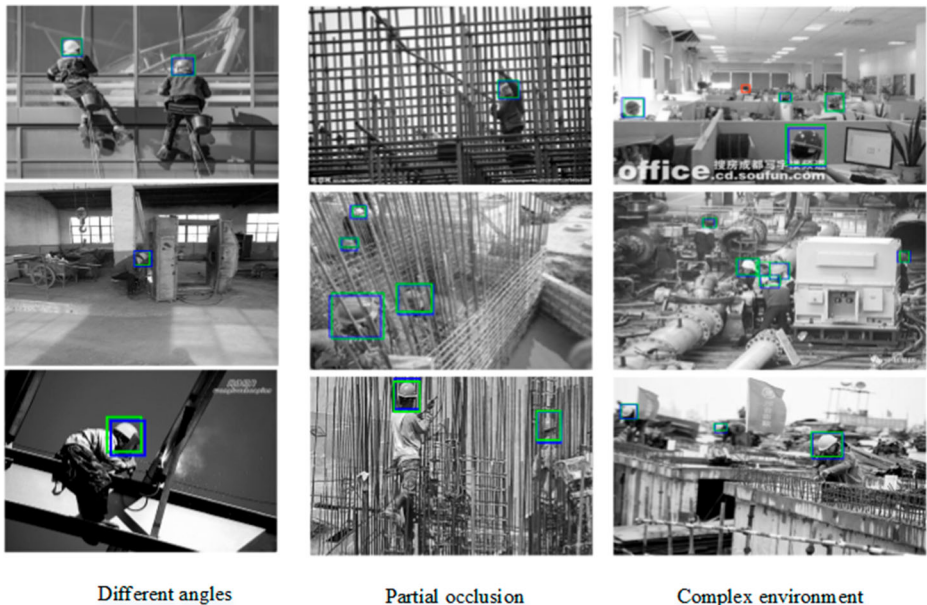

Different angles   Partial occlusion   Complex environment

Fig. 8 Detection results of IBRFs. The blue box represents Ground Truth (GT), the green box represents the detected person with helmet, and the red box represents the situation of mistaking the background as the object

wearing status detection task, the IBRFs proposed in this paper can achieve faster detection speed and higher detection accuracy.

Figure 8 shows the detection results of IBRFs in the safety helmet detection task in different angles, partial occlusion and complex environment. The reason for the success of IBRFs is that it adopts a crowd voting strategy based on prior probability, takes the differences between various features into full consideration, and gains strong adaptability in generalization.

5 Conclusions

A novel object detection method, known as Improved Boosted Random Ferns algorithm (IBRFs) is proposed in this work for safety helmet wearing status detection. The IBRFs makes use of the advantage of basic BRFs and the improved Real AdaBoost algorithm for voting. In the proposed method, it adopts a crowd voting strategy based on prior probability, takes the differences between various features into full consideration, and gains strong adaptability in generalization for safety helmet wearing status detection. The proposed IBRFs shows more efficient outcomes compared to some deep object detectors. Furthermore, two other public datasets, INRIA horses datasets and UIUC cars datasets, are tested and compared with the basic BRFs. The results aslo verified the competitiveness of IBRFs.

Due to the effectiveness of the proposed IBRFs, it is suitable for safety helmet wearing status detection under various scenarios. Next, some recent object detection methods, such as FCOS and DETR, will be taken into account to enhance our IBRFs for object detection task so as to complete complex and various applications.

References

1. Bahaa E, Mneymneh et al (2017) Automated hardhat detection for construction safety applications[J]. Procedia Eng 196:895–902
2. Cai L, Qian J (2011) A method for detecting miners based on helmets detection in underground coal mine videos[J]. Min Sci Technol (China) 21(4):553–556
3. Cheng R, He X, Zheng Z, Wang Z (2021) Multi-scale safety helmet detection based on SAS-YOLOv3-tiny [J]. Appl Sci 11(8):3652–3652
4. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection[C]. Comput Vis Pattern Recognit 33(9):886–893
5. Fang M, Sun TT, Sao Z (2019) Fast helmet wearing detection based on improved yolov2 [J]. Opt Precis Eng 27(05):1196–1205
6. Ferrari V, Fevrier L, Jurie F, Schmid C (2008) Groups of adjacent contour segments for object detection. IEEE Trans Pattern Anal Mach Intell 30(1):36–51. https://doi.org/10.1109/TPAMI.2007.1144
7. Freund Y (1995) Boosting a weak learning algorithm by majority [J]. Inf Comput 121(2):256–285
8. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp 580–587. https://doi.org/10.1109/CVPR.2014.81
9. Jebril NA, Al-Zoubi HR, Abu Al-Haija Q (2018) Recognition of handwritten Arabic characters using histograms of oriented gradient (HOG). Pattern Recognit Image Anal 28:321–345. https://doi.org/10.1134/S1054661818020141

10. Lowe DG (2004) Distinctive image features from scale-invariant keypoints[J]. Int J Comput Vis 60(2):91–110

11. Martinez M, Yang K, Constantinescu A, Stiefelhagen R (2020) Helping the blind to get through covid-19: social distancing assistant using real-time semantic segmentation on rgb-d video. Sensors (Switzerland) 20:1–17. https://doi.org/10.3390/s20185202

12. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Trans Pattern Anal Mach Intell 24(7):971–987

13. Park MW, Elsafty N, Zhu ZH (2015) Hardhat-wearing detection for enhancing on-site safety of construction workers [J]. J Constr Eng Manag 141(9):1–16

14. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 779–788. https://doi.org/10.1109/CVPR.2016.91

15. Ren S, He K, Girshick R, Sun J (2017) Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Trans Pattern Anal Mach Intell 39(6):1137–1149

16. Schapire RE, Singer Y (1999) Improved boosting algorithms using confidence-rated predictions [J]. Mach Learn 37(3):297–336. https://doi.org/10.1023/A:1007614523901

17. Shi T, Niu Y, Liu M, Yang Y, Wang C, Huang Y (2019) Underwater dense targets detection and classification based on YOLOv3. 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp 2595–2600. https://doi.org/10.1109/ROBIO49542.2019.8961615

18. Siebert FW, Lin H (2020) Detecting motorcycle helmet use with deep learning. Accid Anal Prev 134:105319. https://doi.org/10.1016/j.aap.2019.105319

19. Thakar V, Saini H, Ahmed W, Soltani MM, Aly A, Yu JY (2018) Efficient single-shot multibox detector for construction site monitoring. 2018 IEEE International Smart Cities Conference (ISC2), pp 1–6. https://doi.org/10.1109/ISC2.2018.8656929

20. Villamizar M, Moreno-Noguer F, Andrade-Cetto J, Sanfeliu A (2010) Shared random ferns for efficient detection of multiple categories. 2010 20th International Conference on Pattern Recognition, pp 388–391. https://doi.org/10.1109/ICPR.2010.103

21. Villamizar M, Andrade-Cetto J, Sanfeliu A, Moreno-Noguer F (2012) Bootstrapping boosted random ferns for discriminative and efficient object classification [J]. Pattern Recogn 45(9):3141–3153

22. Villamizar M, Andrade-Cetto J, Sanfeliu A, Moreno-Noguer F (2018) Boosted random ferns for object detection. IEEE Trans Pattern Anal Mach Intell 40(2):272–288. https://doi.org/10.1109/TPAMI.2017.2676778

23. Wang B, Chen F, Deng P (2014) Online object tracking based on sparse subspace representation. The 26th Chinese Control and Decision Conference (2014 CCDC), pp 3975–3980. https://doi.org/10.1109/CCDC.2014.6852876

24. Wu F, Jin G, Gao M, He Z, Yang Y Helmet detection based on improved YOLO V3 deep model. 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC), pp 363–368. https://doi.org/10.1109/ICNSC.2019.8743246

25. Xudong C, Yichen W, Fang W et al (2014) Face alignment by explicit shape regression [J]. Int J Comput Vis 107(2):177–190

26. Yokoya M, Iwasaki A (2015) Generalized-hough-transform object detection using class-specific sparse representation for local-feature detection. 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp 2852–2855. https://doi.org/10.1109/IGARSS.2015.7326409

27. Zhang J, Yang K, Stiefelhagen R (2020) ISSAFE: improving semantic segmentation in accidents by fusing event-based data. arXiv preprint arXiv:2008.08974