candidate results are sorted to determine the optimal grabbing position, so as to speed up the detection speed; stage 2 Calculate the grab angle for the local position image output by Angle-Net in the previous stage. Compared with the method in [9], the directly calculated grasp angle error is smaller and the grasp detection accuracy is improved.

(3) The transfer learning [60] mechanism is used in the model training process, which ensures the dynamic expansion of the data set and shortens the training period.

(4) Verify the effectiveness of the crawl detection method in public data sets and actual scenarios. Experiments show that the crawl detection method is fast (17.5 frames/second), with high accuracy and strong robustness.

The paper is organized as follows. The enhancement of real-time grasp detection by the detector consisting of cascaded R-FCN [61] and Angle-Net are presented in Section 2. Section 3 analyzes the experimental setup and the results. Conclusions are provided in Section 4.

2. Detector Consisting of Cascaded R-FCN and Angle-Net

The grasp detection task includes two stages: grasping point determination and grasping pose estimation. In a coarse-to-fine way, the corresponding convolutional neural network is designed for each part of the task and the network is cascaded into the final detection model. The model structure is shown in Figure 2. The first stage can be regarded as a positioning and classification problem, based on R-FCN to achieve grasping positioning and rough estimation of the grasping angle; the second stage is converted into a regression problem by constructing Angle-Net model realizes fine estimation of grasp angle.

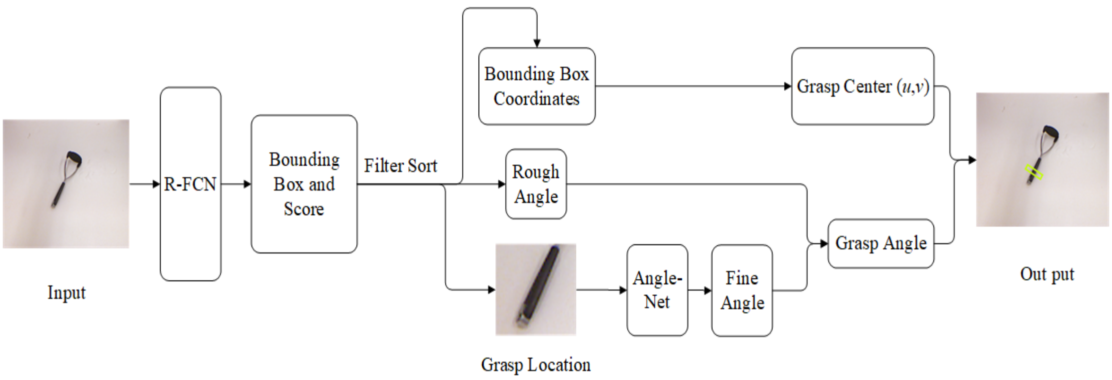


Figure 1. Structure of real-time grasp detection model.

2.1 Grasp Point Location and Pose Estimation

In this paper, R-FCN is used to extract the candidate grasp positions in the image. The candidates are marked by the bounding box on the image, and the grasping point is the center point of the bounding box. In order to achieve a rough estimation of the grasp angle, the grasp angle θ is used as a classification label, and there are 4 categories in total: 0^0, 45^0, 90^0, 135^0. In order to improve the detection speed and minimize the impact on the detection results, the candidate frame for the grasp position is set at 300. The input of the R-FCN model is a scene image containing the target object of any size, and the output is the candidate frame and its corresponding reliability score. The screening position of the highest score in the working area is determined through screening and sorting. R-FCN is based on the R-CNN framework, that is, to make regional recommendations first and then to classify the regions. In order to make the detection respond accurately to the translation of the target, a full convolutional network (FCN) is used to construct a position-sensitive score map with a special convolutional layer. Each spatially sensitive map encodes the relative spatial position information of RoI, and adds a position sensitive RoI pooling layer on the FCN to supervise these score maps. The structure of R-FCN is shown in Figure 3, which is composed of four parts: basic convolutional network, region suggestion network, position-sensitive score graph, and decision layer after RoI pooling.

Assuming that there are *c* categories to be detected and *c*=4 in the robot grab detection model. The basic convolutional network in the R-FCN structure is based on the ResNet, which uses the first 100 layers of ResNet and a 1×1×1024 full convolutional layer at the end. The basic convolutional network is used for feature extraction and output feature map. The region proposal network follows the RPN network in R-CNN, which generates multiple RoIs, that is, grabbing position candidate regions, and each RoI is divided into $k \times k$ blocks. The k^2 position-sensitive score map is used as the last convolutional layer in R-FCN. Its function is to output the results for classification. R-FCN performs position-sensitive pooling operation on RoI's (i, j) block $(0 \le i, j \le k-1)$ defined as Eq.1:

$$r_c\left(i,j\middle|\Theta\right)=\sum_{(x,y)\in(i,j)} Z_{i,j,c}(x+x_0,y+y_0|\Theta)/n \tag{1}$$

Where $r_c\left(i,j\middle|\Theta\right)$ represents the pooling response of the (i, j) block to category C, $Z_{i,j,c}$ is one of the $k^2(4+1)$ score graphs, (x_0, y_0) the upper left corner of RoI, n represents the pixel value in each block, and Θ is the parameter to be learned. After the pooling operation, k^2 position-sensitive score maps are output, and the final scores of each category are obtained using Eq.2 and Eq.3, which are used to calculate the loss.

In order to achieve bounding box regression, the output of the basic convolutional network is followed by a convolutional layer of $4k^2$ channels, and a position-sensitive RoI pooling operation is performed on the $4k^2$ map. Each RoI corresponds to a $4k^2$ vector, and then A 4-dimensional vector is generated through the average vote operation, which is the parameter of the bounding box $t = (t_x, t_y, t_w, t_h)$.

$$r_c\left(\Theta\right)=\sum_{i,j} r_c(i,j|\Theta) \tag{2}$$

$$s_c\left(\Theta\right)=e^{r_c(\Theta)}/\sum_{c`=0} e^{r_{c`}(\Theta)} \tag{3}$$

2.2 Fine Estimation of Grasp Pose

The direct output of the model angle value instead of the angle classification label value can achieve a more accurate grasping pose estimation, so the angle-Net fine estimation model is constructed, and the structure is shown in Figure 3.

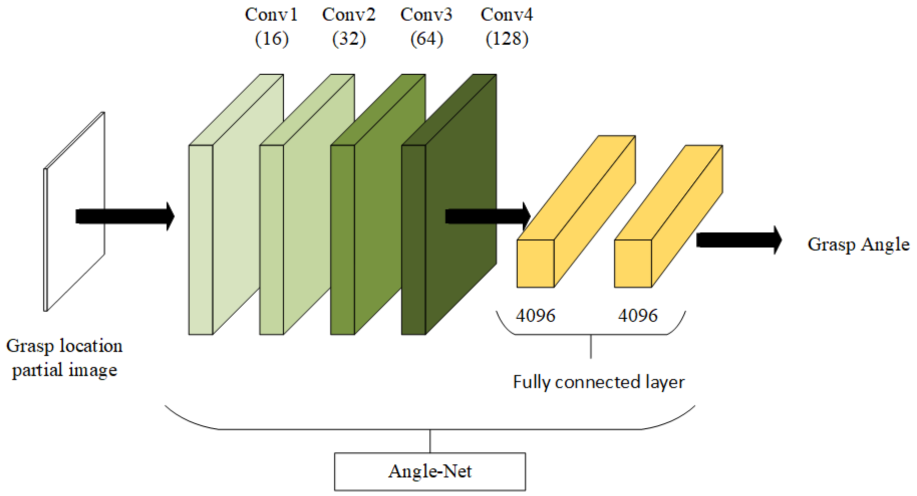


Figure 2. Structure of Angle-Net.

Angle-Net consists of 4 convolutional layers and 2 fully connected layers. The number of convolution kernels of the convolutional layer is 16, 32, 64, 128, and the number of neurons of the fully connected layer is 4096. Loss function (loss function), as a function of estimating the difference between the predicted value and the true value of the model, determines the convergence speed and final effect of model training. The loss function of Angle-Net uses L1 norm function. Among them, θ is the desired grasp angle, λ is the regularization term, and ω is the model weight parameter.

$$L = \frac{1}{N}\left(\left|\theta' - \theta_0\right| + \sum_{i}^{n} \lambda \omega_i^{\,2} \right) \tag{4}$$

The input of Angle-Net is the local image of the grabbing position output by the previous level, and the output is the grasp angle θ’ accurate to 10 under the image plane. The angle θ’ output by Angle-Net is the precision of the coarse estimation angle in stage 1. In view of the stability of R-FCN, the coarse estimation angle θ_{label} in stage 1 can supervise the results of Angle-Net calculation. In order to enhance the fault tolerance of the model and further improve the detection rate of the gripping detection and the success rate of the actual gripping, the final grasp angle fed back to the robot is determined by Eq.5.

$$\theta = \begin{cases} \theta', \left|\theta' - \theta_{label}\right| \le 45^0 \, or \left|\theta' - \theta_{label}\right| \ge 135^0 \\ \theta_{label} \end{cases} \tag{5}$$

3. Experimental Evaluation and Discussion

3.1 Model training

In order to prove the real-time grasp detection capability, we performed experiments on the Cornell Grasping Dataset [62]. Cornell obtained a total of 7,365 rectangular samples. The data set consists of 2 parts, one part is the background image without grasp in the image, a total of 11 images, and the other part is the image with grabs in the image, a total of 885 images, and these 885 images contain 240 A variety of objects are marked with 8019 grasp frames. A total of 5110 frames in these grasp frames can be used for grasping. Tests were performed on each data set were tested and compared. In testing, the combined model predicts both the best grasp and object type at the same time. Our learning rate on all layers is 0.0005 and the initial weight is 0.1, the offset is 0, and the moving average model attenuation rate is 0.99. We used TensorFlow object detection package running on NVIDIA GTX 1050Ti GPU to train and test our model.

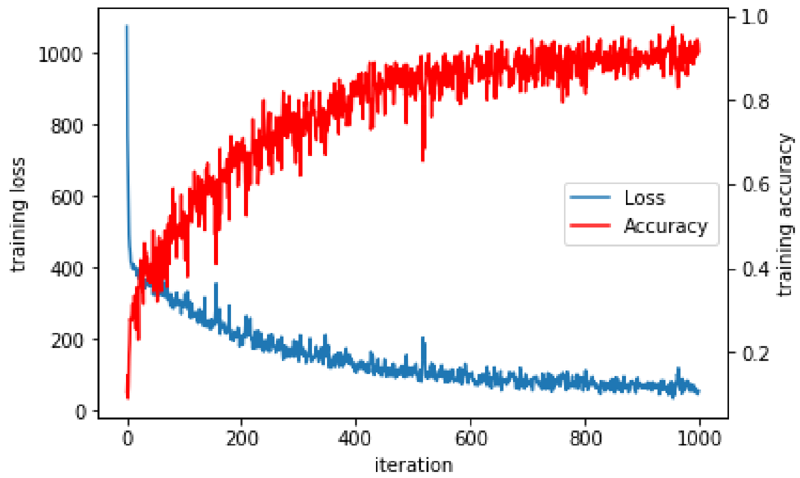
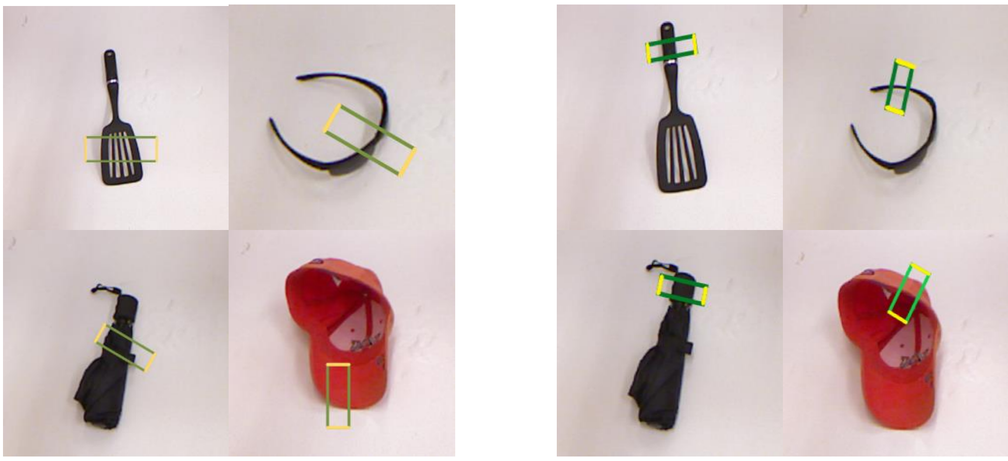


Figure 3. Grasp detection training accuracy (red) and Grasp detection training loss (blue).

3.2 Verification and Comparison of Detection Methods

The detection effect of the grasp detection model is verified on the Cornell Dataset. For comparison, only the vector rectangular frame is used here to represent the grabbing position of objects on the image. As shown in Figure 4, the target object has different geometric shapes and different placement directions. This method can accurately locate the grab position and give a reasonable grasp angle. Compared with the method of Lenz et al. [50], the pose results of the method in this paper are more stable when actually captured.

Extract the same amount of two types of data from the Cornell Dataset to check the model recognition rate. Type 1 represents an image of an object similar to the shape of the training data but with a different placement angle. Type 2 represents an image of an object that is completely different from the shape of the object in the training data. It can be seen from Table 1 that compared with the other four methods, the proposed method has the highest recognition rate in the two types of data tests. In terms of model structure, the model uses a convolutional neural network.



(**a**) (**b**)

Figure 4. Verification effect on the Cornell Dataset: (a) Lenz et al.; (b) Cascade convolutional neural network

Compared with the method in [7], the grasping pose feature is changed from artificial setting to autonomous extraction. Compared with the paper [50,56,8], this model has more network layers in the extraction part of the grasping pose feature, so the extracted features are more diverse and conducive to grasping pose judgment. Regarding the grab detection strategy, other methods first obtain window images through a sliding window search or random sampling at certain intervals, and then judge each window image. It is easy to miss the better grasp poses and the detection result of the same image determine. The method first extracts features from the image as a whole, and then selects the candidate frame for the grabbing position according to the features. The multiple detection results of the same image are consistent. Table 1 the accuracy results in Type 1 indicate that the Angle-Net model has strong generalization ability, and the high detection rate in Type 2 also indicates that the cascaded convolutional neural network grab detection model is very robust, that is, the robot is facing. When a new grab target is used, it can also achieve grasp detection.

Table 1. The results of different grasp pose detection methods.

Methods	**Detection accuracy/%**	
	Type 1	**Type 2**
Reference[7]	60.5%	58.3%
Reference[51]	73.9%	75.6%
Reference[8]	88.0%	87.1%
Reference[60]	93.2%	89.1%
Ours	94.2%	91.3%

The detection speed of this method can reach 17.5 frames/second (Table 2), significantly leading the experimental results of the other two published detection methods on the same hardware platform 0.02 frames/second and 0.62 frames/second and the calculation time of the grasp positioning method is much shorter than the sliding window method [7] and the random sampling method under large samples [9].

Table 2. The detection speeds of different methods.

Methods	**Speed**/(frames/second)
Reference[7]	0.02
Reference[9]	0.62
Ours	17.5

3.3 Robotic Experiment

The actual robot grab detection result is expressed by the "dot line method [7]". It can be seen from Figure 5 that using a rough estimation angle to grasp a long-handled object may be successful, but the angle is not optimal. If the rough angle is used, when the robot grabs objects of other shapes (such as a ring), it is likely that the grasp will fail because the force is not vertical or the contact point is incorrect. The precise angle output by Angle-Net is more reliable in actual crawling. The grabbing positions of the objects in Figure 5 are not at their geometric centers, which cannot be solved by traditional detection methods, and this method can accurately locate the grasp points.

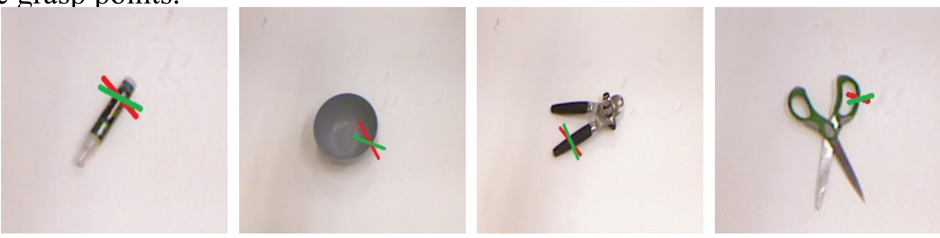
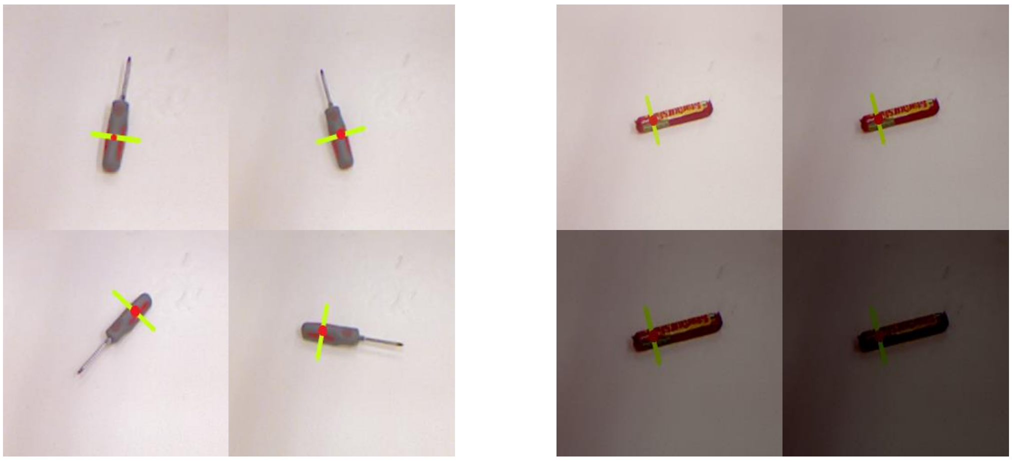


Figure 5. The real results of grasp pose detection (red represents the result of coarse estimation, and green represents the result of fine estimation).



(**a**) (**b**)

Figure 6. Grasp pose detection results under artificially changed environment: (a) The grasp pose detection results of the object with different poses; (b) The grasp pose detection results influenced by illumination.

Compared with traditional methods, the robotic grasp detection method proposed in this paper can be applied. The types of objects are more varied, and it has better adaptability to the background color change. The grasp pose detection is performed on the same object in different postures. The results are shown in Figure 6(a). Although the grab points and postures detected in any posture are different, they are all reasonable. We artificially change the brightness of the same scene to verify the effect of illumination on the detection results. As shown in Figure 6(b), the light intensity in the four images decreases in sequence. There is obvious noise in the image due to insufficient light, but the grab posture detected by each image is almost the same, indicating that the method is suitable for images with uneven illumination or noise.

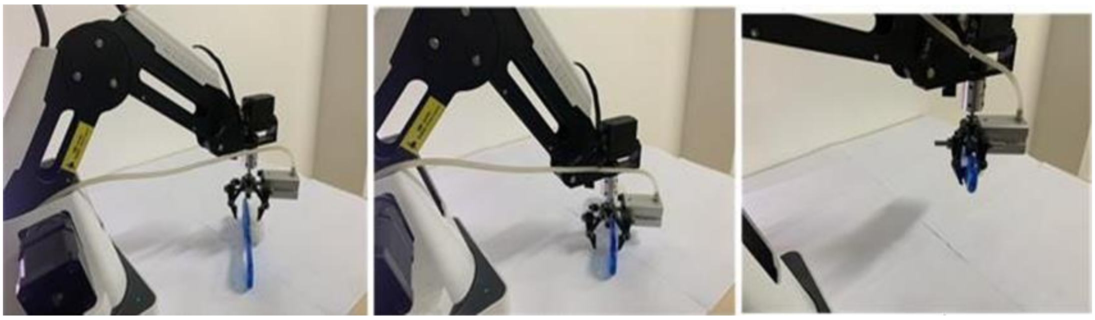


Figure 7. Robotic grasp for objects

Table 3. Results of robotic grasp for different objects with different poses.

Category	Grasp times	Successful times	Success rate
Milk carton	20	20	100
Ballpoint pen	20	18	90
Plate	20	17	85
Tape	20	20	100
Screwdriver	20	18	90
Pliers	20	18	90
Total	120	111	92.5

A total of 120 robot grasp experiments were performed on 6 types of objects in different postures. The results of some successful gripping are shown in Figure 7. The left and right halves of each figure represent the end effector closure and the object was caught. Since there are errors in the calibration of the vision system and the execution of the robot, the actual gripping point is consistent with the detection results through error compensation. The statistical results of the crawling experiment are shown in Table 3. From the limited representative experiments, the average success rate of robotic grasp is 92.5%. This experiment proves the practicality of the grasp detection method proposed in this paper.

4. Conclusion

We proposed a fast and accurate camera-based grasp pose detection system. Our model improves on the latest technology and runs faster than the former methods. We propose a fast detection method for plane grab posture based on cascaded convolutional neural network. In this paper, based on the RGB information of the capture scene, the self-built small-scale data set training model is used to autonomously extract the capture pose characteristics, and the R-FCN model is used to extract the candidate frame for screening to achieve the capture positioning and rough angle estimation. The Angle-Net model performs fine estimation of the grasping angle to realize grasping detection. The use of R-FCN to extract a small number of reliable candidate grabbing positions significantly improves the speed of grabbing detection. The proposed Angle-Net effectively improves the detection accuracy. Experiments show that this method can quickly and accurately calculate the position and attitude angle of the robot's grasping posture that is in line with human grasping habits and maps to the robot's end paws for scene images with diverse backgrounds, uneven lighting, and noise. Compared with the previous methods, not only improves the timeliness of the algorithm because of avoiding traversal search, but also improves the accuracy of attitude angle detection because of the addition of cascading Angle-Net.

Acknowledgements

This work was supported by grants of National Natural Science Foundation of China (Grant Nos.51575407, 51505349, 61733011, 41906177); the Grants of Hubei Provincial Department of Education (D20191105); the Grants of National Defense Pre-Research Foundation of Wuhan University of Science and Technology (GF201705) and Open Fund of the Key Laboratory for Metallurgical Equipment and Control of Ministry of Education in Wuhan University of Science and Technology (2018B07). **References**

[1] Jiabing Hu, Ying Sun, Gongfa Li, Guozhang Jiang, Bo Tao. Probability analysis for grasp planning facing the field of medical robotics. Measurement, 2019, 141, 227-234.

[2] Zhihua Cui, Xianghua Xu, Fei Xue, Xingjuan Cai, Yang Cao, Wensheng Zhang, Jinjun Chen. Personalized Recommendation System based on Collaborative Filtering for IoT Scenarios. IEEE Transactions on Services Computing, 2020, DOI:10.1109/TSC.2020.2964552.

[3] Wei Miao, Gongfa Li, Guozhang Jiang, Yinfeng Fang, Zhaojie Ju, Honghai Liu. Optimal grasp planning of multi-fingered robotic hands: a review. Applied and Computational Mathematics, 2015, 14(3): 238-247.

[4] Zhihua Cui, Fei Xue, Shiqiang Zhang, Xingjuan Cai, Yang Cao, Wensheng Zhang, Jinjun Chen. A Hybrid BlockChain-Based Identity Authentication Scheme for Multi-WSN. IEEE Transactions on Services Computing, 2020, 13(2): 241-251.

[5] Zhihua Cui, Jiangjiang Zhang, Di Wu, Xingjuan Cai, Hui Wang, Wensheng Zhang, Jinjun Chen. Hybrid Many-Objective Particle Swarm Optimization Algorithm for Green Coal Production Problem. Information Sciences, 2020, 518: 256-271.

[6] Xingjuan Cai, Penghong Wang, Lei Du, Zhihua Cui, Wensheng Zhang and Jinjun Chen. Multi-Objective Three-Dimensional DV-Hop Localization Algorithm With NSGA-II. IEEE Sensors Journal, 2019, 19(21): 10003-10015.

[7] Yun Jiang, S. Moseson, A. Saxena. Efficient grasping from RGB-D images: Learning using a new rectangle representation. In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, Piscataway, USA, IEEE, 2011, pp.3304-3311

[8] J. Redmon, A. Angelova. Real-time grasp detection using convolutional neural networks. In Proceedings of the 2015 IEEE International Conference on Robotics

and Automation, Piscataway, USA, IEEE, 2015, pp.1316-1322.
[9] L. Pinto, A. Gupta. Supersizing self-supervision: Learning to grasp from 50k tries and 700 robot hours. In Proceedings of the 2016 IEEE International Conference on Robotics and Automation. Piscataway, USA, IEEE, 2016, pp.3406-3413.
[10] Xingjuan Cai, Yun Niu, Shaojin Geng, Jiangjiang Zhang, Zhihua Cui, Jianwei Li and Jinjun Chen. An under-sampled software defect prediction method based on hybrid multi-objective cuckoo search. Concurrency and Computation: Practice and Experience, 2020, 32(5): e5478.
[11] S. Levine, P. Pastor, A. Krizhevsky, J. Ibarz, D. Quillen. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. International Journal of Robotics Research, 2017, 37(4/5): 421–436.
[12] Penghong Wang, Jianrou Huang, Zhihua Cui, Liping Xie and Jinjun Chen. A Gaussian Error Correction Multi-Objective Positioning Model with NSGA-II. Concurrency and Computation Practice and Experience. 2020, 32(5): e5464.
[13] Zhihua Cui, Lei Du, Penghong Wang, Xingjuan Cai and Wensheng Zhang. Malicious code detection based on CNNs and multi-objective algorithm. Journal of Parallel and Distributed Computing, 2019, 129: 50-58.
[14] Gongfa Li, Jiang Du, Yanling Zhou, Guozhang Jiang, Jianyi Kong. Manogaran, G. Human lesion detection method based on image information and brain signal. IEEE Access, 2019, 7: 11533-11542.
[15] V. Babin, C. Gosselin. Picking, grasping, or scooping small objects lying on flat surfaces: A design approach. International Journal of Robotics Research. 2018, 37, 1484–1499.
[16] Tinggui Chen, Shiwen Wu, Jianjun Yang, Guodong Cong. Risk Propagation Model and Its Simulation of Emergency Logistics Network Based on Material Reliability. International Journal of Environmental Research and Public Health. 2019, 16(23): 4677, DOI: 10.3390/ijerph16234677.
[17] Tinggui Chen, Qianqian Li, Jianjun Yang, Guodong Cong, Gongfa Li. Modeling of the Public Opinion Polarization Process with the Considerations of Individual Heterogeneity and Dynamic Conformity. Mathematics, 2019, 7(10): 917, DOI: 10.3390/math7100917.
[18] Tinggui Chen, Shiwen Wu, Jianjun Yang, Guodong Cong, Gongfa Li. Modeling of Emergency Supply Scheduling Problem Based on Reliability and Its Solution Algorithm under Variable Road Network after Sudden-Onset Disasters. Complexity, 2020, DOI: 10.1155/2020/7501891.
[19] Tinggui Chen, Qianqian Li, Peihua Fu, Jianjun, Chonghuan Xu, Guodong Cong, Gongfa Li. Public Opinion Polarization by Individual Revenue from the Social Preference Theory. International Journal of Environmental Research and Public Health, 2020, 17(3): 946, DOI: 10.3390/ijerph17030946
[20] Ruyi Ma, Leilei Zhang, Gongfa Li, Du Jiang, Shuang Xu, Disi Chen. Grasping force prediction based on sEMG signals. Alexandria Engineering Journal, 2020, 59(3): 1135-1147.
[21] Bei Li, Ying Sun, Gongfa Li, Jianyi Kong, Guozhang Jiang, Du Jiang, Bo Tao, Shuang Xu, Honghai Liu. Gesture recognition based on modified adaptive

orthogonal matching pursuit algorithm. Cluster Computing, 2019, 22 (Supplement 1): 503-512, DOI: 10.1007/s10586-017-1231-7.
[22] Ying Sun, Jinrong Tian, Du Jiang, Bo Tao, Ying Liu, Juntong Yun, Disi Chen. Numerical simulation of thermal insulation and longevity performance in new lightweight ladle. Concurrency and Computation: Practice and Experience, 2020, DOI: 10.1002/CPE.5830.
[23] Disi Chen, Gongfa Li, Ying Sun, Jianyi Kong, Guozhang Jiang, Heng Tang, Zhaojie Ju, Hui Yu, Honghai Liu. An interactive image segmentation method in hand gesture recognition. Sensors, 2017, 17(2): 253.
[24] Jinxian Qi, Guozhang Jiang, Gongfa Li, Ying Sun, Bo Tao. Intelligent human-computer interaction based on surface EMG gesture recognition. IEEE Access, 2019, 7: 61378-61387.
[25] Jinglin Li, Jing Xiao. Progressive planning of continuum grasping in cluttered space. IEEE Transactions on Robotics, 2016, 32(3): 707-716.
[26] Yang He, Gongfa Li, Yajie Liao, Ying Sun, Jianyi Kong, Guozhang Jiang, Du Jiang, Honghai Liu. Gesture recognition based on an improved local sparse representation classification algorithm. Cluster Computing, 2019, 22(Supplement 5): 10935-10946, DOI: 10.1007/s10586-017-1237-1.
[27] Du Jiang, Zujia Zheng, Gongfa Li, Ying Sun, Jianyi Kong, Guozhang Jiang, Hegen Xiong, Bo Tao, Shuang Xu, Honghai Liu, Zhaojie Ju. Gesture recognition based on binocular vision. Cluster Computing, 2019, 22(Supplement 6): 13261-13271, DOI: 10.1007/s10586-018-1844-5
[28] Ying Sun, Chao Xu, Gongfa Li, Wanfen Xu, Jianyi Kong, Du Jiang, Bo Tao, Disi Chen. Intelligent human computer interaction based on non redundant EMG signal. Alexandria Engineering Journal, 2020, 59(3): 1149-1157.
[29] Jing Li, Yang Mi, Gongfa Li, Zhaojie Ju. CNN-based facial expression recognition from annotated RGB-D images for human–robot interaction. International Journal of Humanoid Robotics, 2019, 16(4): 1941002.
[30] Ying Sun, Cuiqiao Li, Gongfa Li, Guozhang Jiang, Du Jiang, Honghai Liu, Zhigao Zheng, Wanneng Shu. Gesture recognition based on Kinect and sEMG signal fusion. Mobile Networks and Applications, 2018, 23(4): 797-805.
[31] Yajie Liao, Ying Sun, Gongfa Li, Jianyi Kong, Guozhang Jiang, Du Jiang, Haibin Cai, Zhaojie Ju, Hui Yu, Honghai Liu. Simultaneous calibration: a joint optimization approach for multiple Kinect and external cameras. Sensors, 2017, 17(7): 1491.
[32] U. Asif, M. Bennamoun, F.A. Sohel. RGB-D Object Recognition and Grasp Detection Using Hierarchical Cascaded Forests. IEEE Transactions on Robotics. 2017, 33: 547–564.
[33] Gongfa Li, Du Jiang, Ying Sun, Guozhang Jiang, Bo Tao. Life prediction mechanism of ladle composite structure body based on simulation technology. Archives of Metallurgy and Materials, 2019, 64(4): 1555-1562.
[34] Gongfa Li, Leilei Zhang, Ying Sun, Jianyi Kong. Towards the sEMG hand: internet of things sensors and haptic feedback application. Multimedia Tools and Applications, 2019, 78(21): 29765-29782.

[35] Li Huang, Meiling He, Chong Tan, Du Jiang, Gongfa Li ,Hui Yu. Jointly Network Image Processing: Multi-task Image Semantic Segmentation of Indoor Scene Based on CNN. IET Image Processing, 2020, DOI: 10.1049/iet-ipr.2020.0088.
[36] Du Jiang, Gongfa Li, Ying Sun, Jianyi Kong, Bo Tao. Gesture recognition based on skeletonization algorithm and CNN with ASL database. Multimedia Tools and Applications, 2019, 78(21), 29953-29970.
[37] Gongfa Li, Heng Tang, Ying Sun, Jianyi Kong, Guozhang Jiang, Du Jiang, Bo Tao, Shuang Xu, Honghai Liu. Hand gesture recognition based on convolution neural network. Cluster Computing, 2019, 22(Supplement 2): 2719-2729, DOI: 10.1007/s10586-017-1435-x.
[38] W. Eric Shaw-Cortez, D. Oetomo, C. Manzie, et al. Tactile-based blind grasping: a discrete-time object manipulation controller for robotic hands. IEEE Robotics and Automation Letters, 2018, 3(2): 1064-1071.
[39] Gongfa Li, Jiahan Li, Zhaojie Ju, Ying Sun, Jianyi Kong. A novel feature extraction method for machine learning based on surface electromyography from healthy brain. Neural Computing and Applications, 2019, 31(12): 9013-9022.
[40] Gongfa Li, Hao Wu, Guozhang Jiang, Shuang Xu， Honghai Liu. Dynamic gesture recognition in the internet of things. IEEE Access, 2019, 7: 23713-23724.
[41] Ying Sun, Jiabing Hu, Gongfa Li, Guozhang Jiang, Hegen Xiong, Bo Tao, Zujia Zheng, Du Jiang. Gear reducer optimal design based on computer multimedia simulation. The Journal of Supercomputing, 2020, 76(6): 4132-4148.
[42] Chengcheng Li, Gongfa Li, Guozhang Jiang, Disi Chen, Honghai Liu. Surface EMG data aggregation processing for intelligent prosthetic action recognition. Neural Computing and Applications, 2018, DOI: 10.1007/s00521-018-3909-z.
[43] Chong Tan, Ying Sun, Gongfa Li, Guozhang Jiang, Disi Chen, Honghai Liu. Research on gesture recognition of smart data fusion features in the IoT. Neural Computing and Applications, 2019, DOI: 10.1007/s00521-019-04023-0.
[44] Jinxian Qi, Guozhang Jiang, Gongfa Li, Ying Sun, Bo Tao. Surface EMG hand gesture recognition system based on PCA and GRNN. Neural Computing and Applications, 2020, 32(10): 6343-6351.
[45] Bowen Luo, Ying Sun, Gongfa Li, Disi Chen, Zhaojie Ju. Decomposition algorithm for depth image of human health posture based on brain health. Neural Computing and Applications, 2020, 32(10): 6327-6342.
[46] N. Correll, K.E. Bekris, D. Berenson, et al. Analysis and observations from the firs amazon picking challenge. IEEE Transactions on Automation Science and Engineering. 2018, 15(1): 172-188.
[47] Wentao Cheng, Ying Sun, Gongfa Li, Guozhang Jiang, Honghai Liu. Jointly network: a network based on CNN and RBM for gesture recognition. Neural Computing and Applications, 2019, 31(Supplement 1): 309-323, DOI: 10.1007/s00521-018-3775-8.
[48] Du Jiang, Gongfa Li, Ying Sun, Jianyi Kong, Bo Tao, Disi Chen. Grip strength forecast and rehabilitative guidance based on adaptive neural fuzzy inference system using sEMG. Personal and Ubiquitous Computing, 2019, DOI: 10.1007/s00779-019-01268-3.

[49] Mingchao Yu, Gongfa Li, Du Jiang, Guozhang Jiang, Bo Tao, Disi Chen. Hand medical monitoring system based on machine learning and optimal EMG feature set. Personal and Ubiquitous Computing, 2019, DOI: 10.1007/s00779-019-01285-2.
[50] Ying Sun, Yaoqing Weng, Bowen Luo, Gongfa Li, Bo Tao, Du Jiang, Disi Chen. Gesture recognition algorithm based on multi-scale feature fusion in RGB-D images. IET Image Processing, 2020, DOI: 10.1049/iet-ipr.2020.0148.
[51] Lenz I, Lee H, Saxena A. Deep Learning for Detecting Robotic Grasps. International Journal of Robotics Research, 2015, 34(4/5): 705–724.
[52] Mingchao Yu, Gongfa Li, Du Jiang, Guozhang Jiang, Fei Zeng, Haoyi Zhao, Disi Chen. Application of PSO-RBF neural network in gesture recognition of continuous surface EMG signals. Journal of Intelligent & Fuzzy Systems, 2020, 38(3): 2469-2480, DOI: 10.3233/jifs-179535.
[53] Yu-Chi Lin, S.-T. Wei, L.-C. Fu. Grasping unknown objects using depth gradient feature with eye-in-hand RGB-D sensor. In Proceedings of the 2014 IEEE International Conference on Automation Science and Engineering, 2014: pp.1258-1263.
[54] M. Hassan, M. Rehmani, J. Chen. Differential Privacy Techniques for Cyber Physical Systems: A Survey, IEEE Communications Surveys and Tutorials, 2020, 22(1): 746-789.
[55] J. Feng, L.T. Yang, R. Zhang, W. Qiang, J. Chen. Privacy Preserving High-Order Bi-Lanczos in Cloud-Fog Computing for Industrial Applications, IEEE Transactions on Industrial Informatics, 2020, DOI: 10.1109/TII.2020.2998086.
[56] Yangwei Cheng, Gongfa Li, Jiahan Li, Ying Sun, Guozhang Jiang, Fei Zeng, Haoyi Zhao, Disi Chen. Visualization of activated muscle area based on sEMG. Journal of Intelligent & Fuzzy Systems, 2020, 38(3): 2623-2634, DOI: 10.3233/jifs-179549.
[57] Jinrong Tian, Wentao Cheng, Ying Sun, Gongfa Li, Du Jiang, Guozhang Jiang, Bo Tao, Haoyi Zhao, Disi Chen. Gesture recognition based on multilevel multimodal feature fusion. Journal of Intelligent & Fuzzy Systems, 2020, 38(3): 2539-2550, DOI: 10.3233/JIFS-179541.
[58] Shangchun Liao, Gongfa Li, Jiahan Li, Du Jiang, Guozhang Jiang, Ying Sun, Bo Tao, Haoyi Zhao, Disi Chen. Multi-object intergroup gesture recognition combined with fusion feature and KNN algorithm. Journal of Intelligent & Fuzzy Systems, 2020, 38(3): 2725-2735, DOI: 10.3233/JIFS-179558.
[59] Li Huang, Qiaobo Fu, Gongfa Li, Bowen Luo, Disi Chen, Hui Yu. Improvement of maximum variance weight partitioning particle filter in urban computing and intelligence. IEEE Access, 2019, 7: 106527-106535.
[60] Di Guo, Fuchun Sun, Huaping Liu, et al. A hybrid deep architecture for robotic grasp detection. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation, Piscataway, USA, IEEE, 2017: pp.1609-1614.
[61] Jifeng Dai, Yi Li, Kaiming He, Jian Sun. R-FCN: Object detection via region-based fully convolutional networks. In Proceedings of Advances in Neural Information Processing Systems 29. Cambridge, USA: MIT Press, 2016: pp.379-387.

[62] "Cornell grasping dataset," http://pr.cs.cornell.edu/grasping/rect_data/data.php. (Accessed: 2019-09-01).