Jurnal Integrasi
Vol. 13 No. 1, April 2021, 15-25
e-ISSN: 2548-9828

Article History
Received January 2021
Accepted April 2021

Industry 4.0: Hand Recognition on Assembly Supervision Process

Riska Analia*, Andika Putra Pratama*, Susanto*

*Batam Polytechnics
Department of Electrical Engineering
Parkway Street, Batam Centre, Batam 29461, Indonesia
E-mail: riskaanalia@polibatam.ac.id, andikaputrapratama17@gmail.com, susanto@polibatam.ac.id

Abstrak

Pada industri perakitan, proses merakit komponen merupakan hal yang sangat penting guna menghasilkan produk yang berkualitas. Perakitan komponen hendaklah dilakukan secara urut berdasarkan standar yang telah ditentukan oleh perusahaan. Bagi perusahaan yang masih menggunakan proses perakitan secara manual yakni dengan menggunakan tenaga manusia, terkadang terjadi kesalahan dalam proses perakitan, sehingga dapat mempengaruhi kualitas produksi. Agar proses perakitan dapat dilakukan sesuai prosedur, maka diperlukan sebuah sistem yang dapat mendeteksi tangan karyawan ketika melakukan proses perakitan secara otomatis. Penelitian ini mengusulkan sistem pendeteksian tangan karyawan secara real-time berbasis kecerdasan buatan. Sistem ini akan menjadi dasar untuk pengembangan proses perakitan produk industri secara otomatis untuk menyambut industri 4.0. Untuk memverifikasi kinerja sistem, beberapa percobaan dilakukan yaitu mendeteksi tangan kanan dan kiri karyawan serta mendeteksi tangan ketika menggunakan aksesoris atau tidak. Dari hasil percobaan dapat disimpulkan bahwa sistem mampu mendeteksi tangan kanan dan kiri karyawan dengan baik dengan rata-rata FPS yang dihasilkan adalah 15.4.

Kata kunci: Industry 4.0, real-time pendeteksian tangan, real-time *supervision assembly process*

Abstract

In the assembly industry, the process of assembling components is very important in order to produce a quality product. Assembly of components should be carried out sequentially based on the standards set by the company. For companies that still operate the assembly process manually by employee, sometimes errors occur in the assembly process, which can affect the quality of production. In order to be carried out the assembly process according to the procedure, a system is needed that can detect employee hands when carrying out the assembly process automatically. This study proposes an artificial intelligence-based real-time employee hand detection system. This system will be the basis for the development of an automatic industrial product assembly process to welcome the Industry 4.0. To verify system performance, several experiments were carried out, such as; detecting the right and left hands of employees and detecting hands when using accessories or not. From the experimental results it can be concluded that the system is able to detect the right and left hands of employees well with the resulting FPS average of 15.4.

Keywords: Industry 4.0, real-time hand detection, real-time supervision assembly process

1. Introduction

Nowadays, the world is entering the era of the 4.0 revolution or the fourth industry where technology has become the basis of human life. Everything becomes unlimited due to the development of the internet and digital technology. In industrial companies, for example, the usage of technology for the assembly process has been grown rapidly. Therefore, they are always able to compete fiercely to produce better products compared to previous years in terms of quality and quantity.

The assembly process in industrial companies mostly combines two or more components mechanically into an item or product, where this is called permanent process, i.e., welding and soldering, and not a permanent process, such as joining the bolts, nuts, and screws. All of these jobs need an industrial employee who is able to help the company to produce a product in the assembly process.

However, in the current assembly process, there are still some employees who are less efficient and optimal in carrying out each step or process and often made some mistakes when taking the equipment to be used for their work. The human error in assembly the products can cause customer complaint and gave a bad impact to the company, i.e., more expensive cost lost costumer and also affect the development of the company. To overcome human error, the supervisor in charge needs to monitor the employee for mistake prevention. However, the supervisor cannot always monitor the one by one because each of the companies has a lot of employees who should be monitored. Thus, the hand detection, motion, and coordinate need to be developed automatically to help the supervisor supervise each of employee.

The study of hand detection or hand gesture detection which has been reviewed in [1] can be implemented in some areas such as in education [2,3], robotics [4,5], healthcare application [6-8], interactive projector scree [9], human computer interaction [10,11] and banking [12]. In detecting hands most of researches used the image of hand with various different method. As presented in [13], they implement the FIFO and path overlap in order to detect the hand motion. Another, the hand detection by generated hand ROIs and verified using Histogram of Oriented (HOG) feature and Support Vector Machine (SVM) [14] and detected the hands using the face skin color in the ROI [15]. The hand gesture detection also can be recognized from hand contour and applied the spatial homomorphic filter and variational level set as done by Zheng, et.alt [16]. The hand gesture detection also can be detected using a new 24GHz transmitting/receiving antenna module which consisted of 16 microstrip elements for transmitting, 4 microstrip for receiving, and 8 elements to support multichannel signal receiving [17]. Moreover, [18] developed the hand tracking algorithm by using the Lucas Kanade Optical flow and combined with color blob detection and used the convex hull algorithm for recognize hand gesture. The other work, used the depth camera image [19,20] for hand and wrist visual tracking. By using the camera, the hand gesture can be recognized by labeling the pixels and make it as a data set [21], while in [22] utilized the taxonomy texture and color-texture feature from the 2D hand gesture image and performed the classification process using Naïve bayes, Real AdaBoost, Gentle AdaBoost, and modest for analyzing the notable information in the image feature.

Mostly, in recognizing the image in this work was hand gesture or hand motion, the high computation is needed to avoid some noise or distortion while detecting the image. Therefore, in [10] presented a distance image based approach using CPU-GPU, heterogeneous computing. The other approached which need a high computation process by using the R-FCN [3] and CNN [23]. Since the study of object detection has been grown rapidly and the computer technology is sufficient for heavy computation, Redmon, et.al [24] introduced the new object detection method called You Only Look Once (YOLO) which the structure consists of some convolutional layer and this method is able to predict the bounding box image and object coordinate precisely. This method is the open source, thus, make some researchers are able to implemented this method on their study i.e., for robotics in detecting ball and goals [25], improved the traffic signs [26], detected human face [27,28], detected pedestrian [29], and hand gesture detection [30]. Two years later, Redmon, et.al, introduced the YOLO V3 which has the architecture bigger than last version but three times faster in detecting and has the accuracy the same as SSD algorithm [31]. This method has already implemented to detect hand [6] and detect the user whom not wearing and wearing the helmet [32]. Based on the state of arts of object detection and the results has already proven by some researcher, therefore in this work we developed the hand detection by using the YOLO V3. In contrast with [6], the hand detection in this work will be implemented as welcoming the industry 4.0 at the industrial in real-time application. Thus, it is hoped that the presence of Industry 4.0 very helpful in order to make the connections between people, processes, and products of a company faster, more flexible, and more efficient.

2. Hand Detection Algorithm

In industry 4.0 area, an assembly process supervision needs to be developed to avoid human error while constructing the device. As for manual supervision in assembly process took long time and sometimes produce some error from the hand step of the assembly process. Therefore, in this section will present the hand detection algorithm for applying in real-time application. The YOLO tiny V3 has been chosen as a proposed method in order to detect the hands on the table. The prototype of this system can be seen on **Figure 1**, where it consists of a webcam camera, camera stand, and computer which is equipped with GPU. The area of hand detection will be recognized by 1280 x 720 pixel.

Figure 1. The prototype of hand detection system.

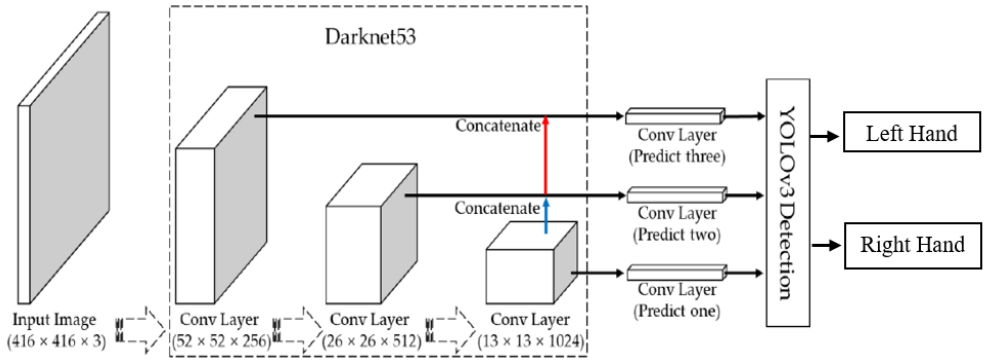


Figure 2. Hand detection YOLO V3 darknet53 convolution architecture.

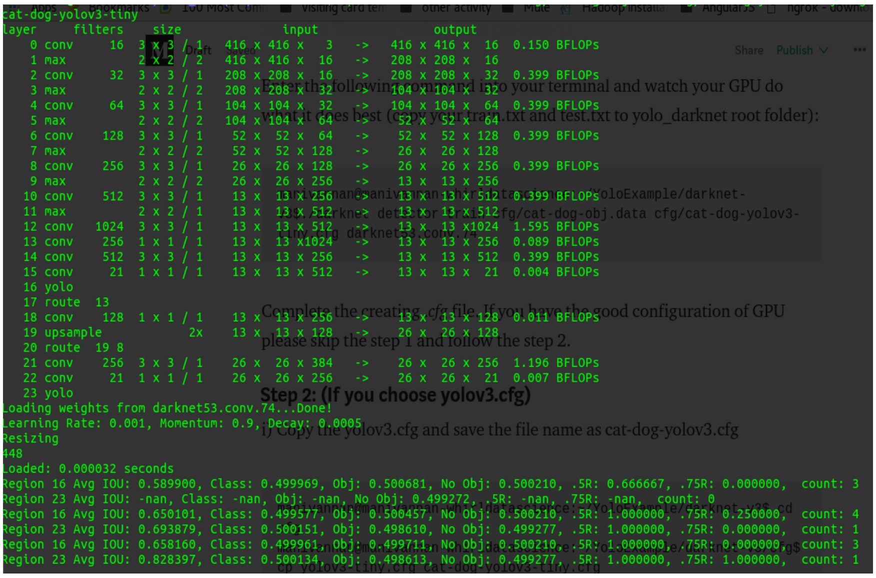


Figure 3. The hand detection convolution layer trained processs.

The hand detection architecture presents on Error! Reference source not found., the system will collect the raw image from the camera, then compress the raw image pixel into 416x416. Then, the system will perform the convolution layer using the darknet53 algorithm to reduce the pixel into 13x13. Moreover, the darknet53 algorithm gave three predictions convolution layer as the input for the YOLO detection method. The YOLO tiny V3 will do prediction of the sample image and classify it into two class in this work which is left and right hand.

In training process, this system used 64 train images in every step, then divided the batch into 16 subdivisions to reduce the VRAM requirement from the GPU. Moreover, the input pixel image will be reduced into 416 x 416 with 3 layers. The learning rate which is implemented in this work was 0.001 with total filter was 21. Each of layer information can be seen on **Figure 3** which is will be appeared while the training process. After doing the training process, the next step was to evaluate the model which has been trained before by using the mean average precision. After doing the calculation, the average precision of detection in this work is around 55.3%. In order to run the system in windows, this work used the Cmake-gui which already installed on Visual Studio and used CUDA 10.0. The OpenCV also implemented in this work for processing the image or video from the YOLO detection results which it can be done in real-time application.

3. Experimental Results

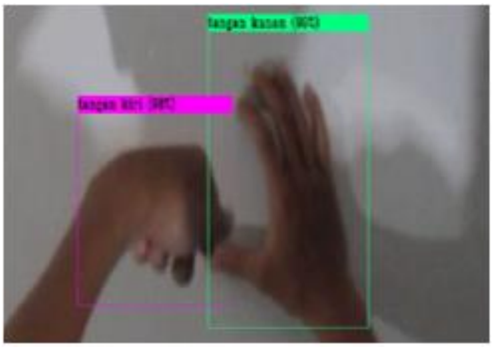
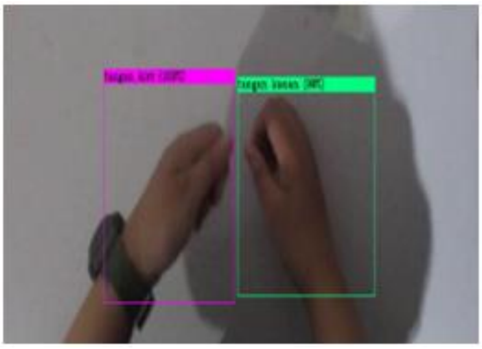
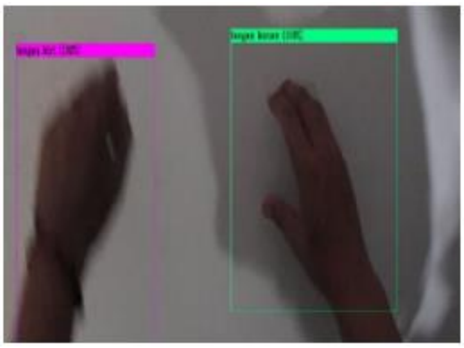
For verifying the hand detection system, the real-time application has been carried out in this work which is implemented on NVIDIA GTX 1050. The GPU usage when detection process was activated around 16% to 20%. This method used the IoU loss = mse (2) and IoU norm= 0.75, the BFLOPS total was 65.312 and the average output 516922.

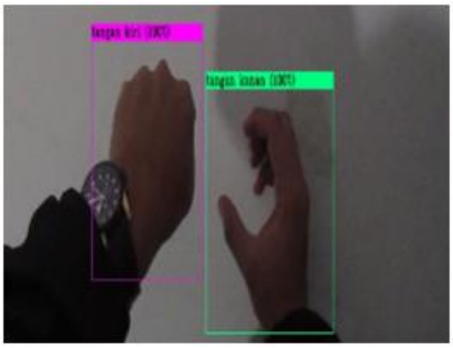
In this experiment, we testified our system with 8 different users to detect their left and right hand. As seen in TABLE I, the results presented that our system is able to detect the left and right hand precisely with total average FPS about 15.4. In order to distinguish left and right hand, the bounding box for each object has been developed. We also shown the confidently percentage of hand detection while the system detects the object. As presented in TABLE I, the magenta box denoted to the left hand and the green box for the right hand. The results in TABLE I, denoted that the system was able to handle every situation in order to detect the left or right hand. It is show in TABLE I no. 7, when the right hand was covered by the sleeve the system did not detected it as a hand and only detected the left hand. The proposed method also detected the hands even the users wearing any accessories such as watch and bracelet, see TABLE I no. 2, 3, 4, 6. From TABLE I, the system always detects the left and right hand with confidently percentage more than 95%, it is denoted that the system able to detect hands correctly and precisely even the average precision detection from the system only 55.3%.

Another, we also collected the coordinate of each hands in real-time as seen in TABLE II from 8 samples. In this experiment, system will record the coordinate regarding the hand movement which is still moved on the camera frame. From this experiment, camera still detected the hands and record the hand position even when it moved closer to each other, see TABLE II no. 7 and the light generated the shadow from the user body to the hand detection. Hence, when the other hand tried to entry the camera frame, the system also detected it as a right hand and recorded the coordinate as well, see TABLE II no. 8.

TABLE I

HAND DETECTION RESULTS

No	Result detection	Average FPS	Confidently Percentages	
			Left hand (%)	Right hand (%)
1		15.8	98	99
2		16.2	100	99
3		11.0	100	100

No	Result detection	Average FPS	Confidently Percentages	
			Left hand (%)	Right hand (%)
4		16.2	100	100
5		16.2	99	100
6		16.1	100	100
7		16.0	95	-

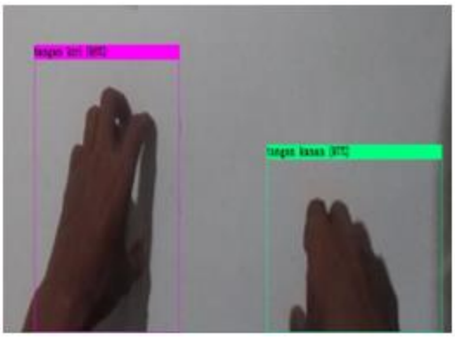
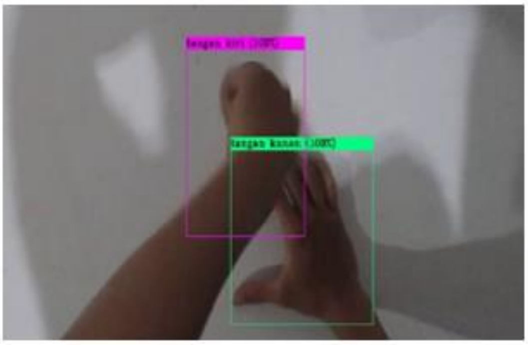
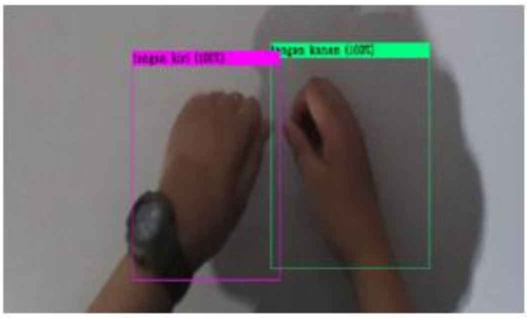
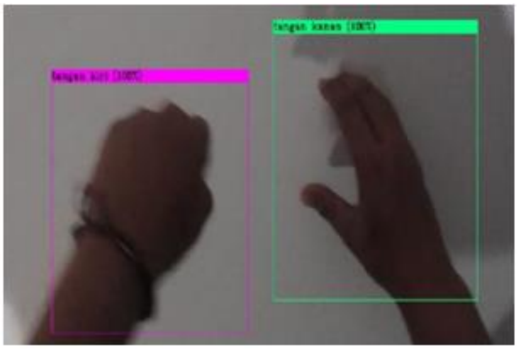
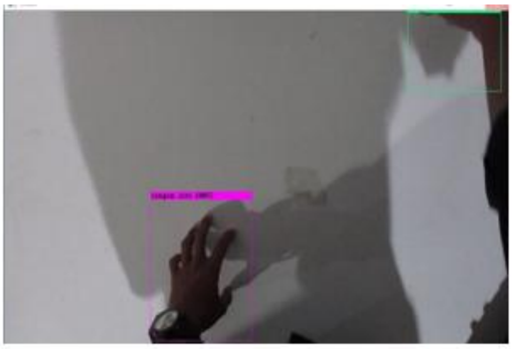
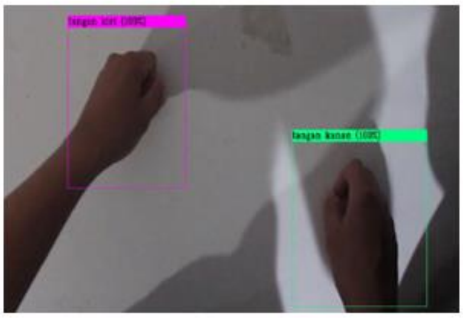
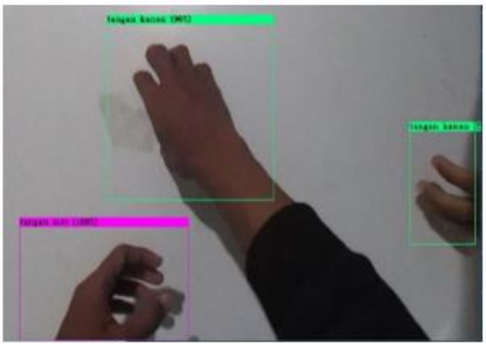
No	Result detection	Average FPS	Confidently Percentages	
			Left hand (%)	Right hand (%)
8		16.2	96	97

TABLE II

HAND DETECTION COORDINATE

No	Result detection	FPS	Confidently Percentages		Coordinate	
			Left hand (%)	Right hand (%)	Left hand	Right hand
1		15.8	100	100	X:547 Y:107 Width: 213 height: 241	X:627 Y:237 Width: 257 height: 225
2		16.0	100	100	X:347 Y:399 Width: 267 height: 241	X:599 Y:390 Width: 290 height: 237

No	Result detection	FPS	Confidently Percentages		Coordinate	
			Left hand (%)	Right hand (%)	Left hand	Right hand
3		16.2	100	100	X:226 Y:326 Width: 364 height: 315	X:637 Y:264 Width: 378 height: 335
4		9.6	98	81	X:371 Y:406 Width: 256 height: 304	X:1022 Y:4 Width: 237 height: 170
5		16.3	100	100	X:275 Y:310 Width: 251 height: 228	X:755 Y:472 Width: 288 height: 235
6		16.2	98	100	X:8 Y:79 Width: 174 height: 216	X:644 Y:420 Width: 323 height: 255

No	Result detection	FPS	Confidently Percentages		Coordinate	
			Left hand (%)	Right hand (%)	Left hand	Right hand
7		16.1	100	100	X:348 Y:308 Width: 256 height: 255	X:549 Y:344 Width: 289 height: 232
8		15.9	100	*99 #58	X:293 Y:519 Width: 361 height: 205	*X:480 Y:163 Width: 356 height: 308 #X:1128 Y:351 Width: 140 height: 199

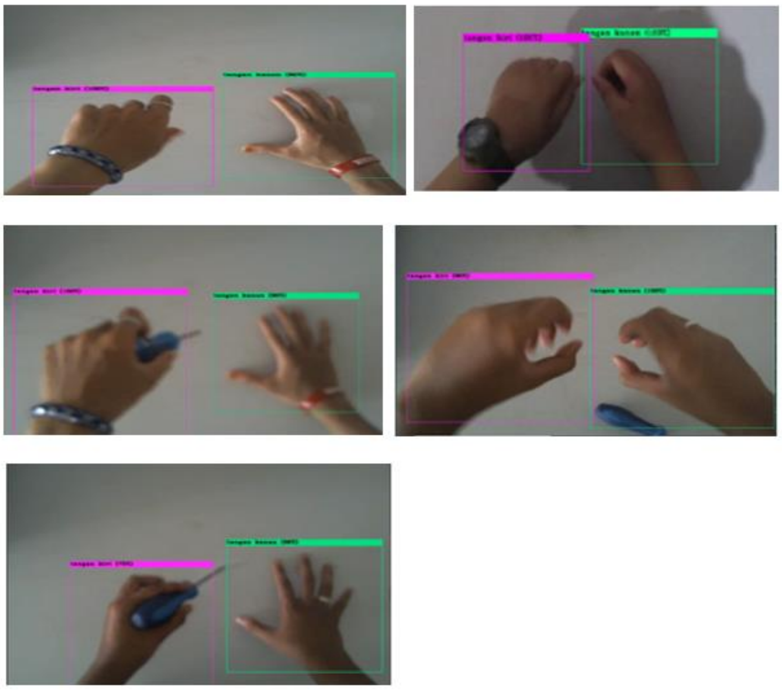


Figure 4. Hand detection when wearing and holding things.

We also testify our system to detect the hands when wearing something such as a watch, a ring, or a bracelet, and holding a screwdriver. This experiment, order the system to detect the hands when it is wearing things on the hand, holding a good, or even moving the hands. From the experiment show on **Figure 4**, the system still able to detect hands correctly with different hand condition.

4. Conclusions

The aim of this work was detecting the hands when it moved on the table in real-time application. Hand detection on the other hand, was developed to support the assembly industrial which need the automatic supervision for look after the employee when assembly the product. In order to developed the hand detection system, we implemented the YOLO tiny V3 in this work. From the experiment results, the system is able to detect the hands properly even the hands are holding or wearing some accessories. This system is able to detect the hands with the average FPS generated by the system is around 15.4. In the future, this system is hopped to be implement in industry to replace the human in working on the assembly station, as one of the implementations of the industry 4.0.

References

[1] S. Gulati and R. K. Bhogal, "Comprehensive Review of various Hand Detection Approaches," 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET), Kottayam, India, 2018, pp. 1-5, doi: 10.1109/ICCSDET.2018.8821238.

[2] M. Bousaaid, T. Ayaou, K. Afdel and P. Estraillier, "Hand gesture detection and recognition in cyber presence interactive system for E-learning," 2014 International Conference on Multimedia Computing and Systems (ICMCS), Marrakech, 2014, pp. 444-447, doi: 10.1109/ICMCS.2014.6911197.

[3] J. Lin, F. Jiang and R. Shen, "Hand-Raising Gesture Detection in Real Classroom," 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, 2018, pp. 6453-6457, doi: 10.1109/ICASSP.2018.8461733.

[4] Q. Gao, J. Liu, Z. Ju and X. Zhang, "Dual-Hand Detection for Human–Robot Interaction by a Parallel Network Based on Hand Detection and Body Pose Estimation," in IEEE Transactions on Industrial Electronics, vol. 66, no. 12, pp. 9663-9672, Dec. 2019, doi: 10.1109/TIE.2019.2898624.

[5] T. Grzejszczak, A. Łegowski and M. Niezabitowski, "Application of hand detection algorithm in robot control," 2016 17th International Carpathian Control Conference (ICCC), Tatranska Lomnica, 2016, pp. 222-225, doi: 10.1109/CarpathianCC.2016.7501098.

[6] X. Yu, L. Zhu and L. Jia, "Detection and recognition of hand abnormal state based on deep learning algorithm," 2020 IEEE International Conference on Artificial Intelligence and Information Systems (ICAIIS), Dalian, China, 2020, pp. 190-193, doi: 10.1109/ICAIIS49377.2020.9194919.

[7] M. -F. Tsai, R. H. Wang and J. Zariffa, "Generalizability of Hand-Object Interaction Detection in Egocentric Video across Populations with Hand Impairment," 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 2020, pp. 3228-3231, doi: 10.1109/EMBC44109.2020.9176154.

[8] F. Naim, R. Jaafar, N. W. Arshad, R. Hamid and M. N. Razali, "Unclean hand detection machine using vision sensor," 2013 Saudi International Electronics, Communications and Photonics Conference, Fira, 2013, pp. 1-4, doi: 10.1109/SIECPC.2013.6550750.

[9] R. Sharma, R. Shikher, N. V. Bansode and P. R. Rajarapollu, "Interactive projector screen with hand detection using gestures," 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), Pune, 2016, pp. 574-577, doi: 10.1109/ICACDOT.2016.7877650.

[10] Z. Yi, X. Hu, B. Jang and K. K. Kim, "A robust and parallel-friendly distance image based hand detection," 2015 International SoC Design Conference (ISOCC), Gyungju, 2015, pp. 33-34, doi: 10.1109/ISOCC.2015.7401651.

[11] M. Ç. Uysal, T. Karapinar, B. Benligiray and C. Topal, "Dataset augmentation for accurate object detection," 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp. 1-4, doi: 10.1109/SIU.2018.8404807.

[12] A. Gupta, Y. Kumar and S. Malhotra, "Banking security system using hand gesture recognition," 2015 International Conference on Recent Developments in Control, Automation and Power Engineering (RDCAPE), Noida, 2015, pp. 243-246, doi: 10.1109/RDCAPE.2015.7281403.

[13] Y. Wang, W. Lin and L. Yang, "A fast hand motion detection based on FIFO," 2017 International Conference on Machine Learning

and Cybernetics (ICMLC), Ningbo, 2017, pp. 599-604, doi: 10.1109/ICMLC.2017.8108978.

[14] J. Kim, J. Baek and E. Kim, "A part-based rotational invariant hand detection," 2013 International Conference on Fuzzy Theory and Its Applications (iFUZZY), Taipei, 2013, pp. 127-129, doi: 10.1109/iFuzzy.2013.6825422.

[15] S. Han, S. Park, M. Kim and J. Paik, "Hand detection using facial information," 2016 IEEE 6th International Conference on Consumer Electronics - Berlin (ICCE-Berlin), Berlin, 2016, pp. 167-168, doi: 10.1109/ICCE-Berlin.2016.7684746.

[16] Y. Zheng and P. Zheng, "Hand Contour Detection Using Spatial Homomorphic Filtering and Variational Level Set," 2015 International Conference on Computer Science and Applications (CSA), Wuhan, 2015, pp. 172-176, doi: 10.1109/CSA.2015.15.

[17] K. Yao, S. Lan, H. Tang, Z. He and C. Yang, "A 24GHz Micropatch Antenna Array for Human Hand Gestures Detection," 2018 USNC-URSI Radio Science Meeting (Joint with AP-S Symposium), Boston, MA, 2018, pp. 95-96, doi: 10.1109/USNC-URSI.2018.8602899.

[18] P. N. Crisnapati, M. Setiawan, I. G. N. Wikranta Arsa, P. Devi Novayanti, M. S. Wibawa and K. G. Oka Ciptahadi, "Real-Time Hand Palm Detection and Tracking Augmented Reality Game Using Lucas Kanade Optical Flow Combined with Color Blob Detection," 2019 1st International Conference on Cybernetics and Intelligent System (ICORIS), Denpasar, Bali, Indonesia, 2019, pp. 263-268, doi: 10.1109/ICORIS.2019.8874892.

[19] J. Y. Oh, J. Lee, J. H. Lee and J. H. Park, "A hand and wrist detection method for unobtrusive hand gesture interactions using HMD," 2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia), Seoul, 2016, pp. 1-4, doi: 10.1109/ICCE-Asia.2016.7804805.

[20] Y. Sun, X. Liang, H. Fan, M. Imran and H. Heidari, "Visual Hand Tracking on Depth Image using 2-D Matched Filter," 2019 UK/ China Emerging Technologies (UCET), Glasgow, United Kingdom, 2019, pp. 1-4, doi: 10.1109/UCET.2019.8881866.

[21] C. Li and K. M. Kitani, "Pixel-Level Hand Detection in Ego-centric Videos," 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, 2013, pp. 3570-3577, doi: 10.1109/CVPR.2013.458.

[22] S. Misra and R. H. Laskar, "Taxonomy of Texture and Color-Texture Features for Developing Hand Detection System under Non-Ideal Conditions," 2017 14th IEEE India Council International Conference (INDICON), Roorkee, 2017, pp. 1-6, doi: 10.1109/INDICON.2017.8487569.

[23] R. Adiguna and Y. E. Soelistio, "CNN Based Posture-Free Hand Detection," 2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE), Kuta, 2018, pp. 276-279, doi: 10.1109/ICITEED.2018.8534743.

[24] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, “You Only Look Once: Unifield, Real-Time Object Detection,” unpublished, 2016.

[25] Susanto, E. Rudiawan, R. Analia, P. Daniel Sutopo and H. Soebakti, "The deep learning development for real-time ball and goal detection of barelang-FC," 2017 International Electronics Symposium on Engineering Technology and Applications (IES-ETA), Surabaya, 2017, pp. 146-151, doi: 10.1109/ELECSYM.2017.8240393.

[26] C. Liu, Y. Tao, J. Liang, K. Li and Y. Chen, "Object Detection Based on YOLO Network," 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 2018, pp. 799-803, doi: 10.1109/ITOEC.2018.8740604.

[27] W. Yang and Z. Jiachun, "Real-time face detection based on YOLO," 2018 1st IEEE International Conference on Knowledge Innovation and Invention (ICKII), Jeju, 2018, pp. 221-224, doi: 10.1109/ICKII.2018.8569109.

[28] D. Garg, P. Goel, S. Pandya, A. Ganatra and K. Kotecha, "A Deep Learning Approach for Face Detection using YOLO," 2018 IEEE Punecon, Pune, India, 2018, pp. 1-4, doi: 10.1109/PUNECON.2018.8745376.

[29] C. Zhao and B. Chen, "Real-Time Pedestrian Detection Based on Improved YOLO Model," 2019 11th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Hangzhou, China, 2019, pp. 25-28, doi: 10.1109/IHMSC.2019.10101.

[30] T. Le, D. Jaw, I. Lin, H. Liu and S. Huang, "An efficient hand detection method based on convolutional neural network," 2018 7th International Symposium on Next Generation Electronics (ISNE), Taipei, 2018, pp. 1-2, doi: 10.1109/ISNE.2018.8394651.

[31] Joseph Redmon, & Ali Farhadi. (2018). YOLOv3: An Incremental Improvement.

[32] J. Hu, X. Gao, H. Wu and S. Gao, "Detection of Workers Without the Helments in Videos Based

on YOLO V3," 2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Suzhou, China, 2019, pp. 1-4, doi: 10.1109/CISP-BMEI48845.2019.8966045.