APEM journal
Advances in Production Engineering & Management
Volume 21 | Number 2 | June 2026 | pp 163–177
https://doi.org/10.14743/apem2026.2.567

ISSN 1854-6250
Journal home: apem-journal.org
Original scientific paper

Industrial multi-object detection for automotive battery manufacturing using a CBAM-enhanced YOLOv5 framework

Li, Y.X.[a,*], Chen, H.[a], Zhou, L.[b]

[a]School of Intelligence Technology, Geely University of China, Chengdu, P.R. China
[b]Huayou Cobalt Co., LTD, Chengdu, P.R. China

ABSTRACT

The mixing stage of automotive battery production requires reliable monitoring of raw material types, personnel actions, and correct tool use. However, accurate multi-scale object detection in complex scenes remains challenging because of background interference and the requirements of embedded deployment and real-time operation. This study proposes an enhanced YOLOv5 framework for industrial multi-object detection. The method adopts a dual-stage feature-enhancement strategy designed to improve robustness while limiting parameter count and computational complexity. First, the Convolutional Block Attention Module (CBAM) is embedded in the Backbone and Neck of YOLOv5 to provide multi-granularity feature enhancement and improve the detection of small objects, such as tools. Second, the conventional CIOU loss is replaced with the Focal-EIOU loss function to optimize bounding-box regression, reduce false detections across multiple target scales, and accelerate model convergence. Finally, K-means clustering is applied to target geometric features to generate specialized anchor-box parameters better suited to industrial scenarios. Experimental results on an automotive battery production-site dataset show that the improved model's mAP@0.5 increased by 2.4 % compared with the original model, reaching 95.2 %, while mAP@0.5:0.95 improved by 3.6 %, reaching 78.5 %. The proposed framework provides a lightweight and reliable solution for multi-object detection in complex industrial environments and supports the development of intelligent visual monitoring systems for smart manufacturing.

ARTICLE INFO

Keywords:
Automotive battery manufacturing;
Industrial computer vision;
Multi-object detection;
Real-time object detection;
Deep learning;
YOLOv5;
Convolutional Block Attention Module (CBAM);
Dual-stage Feature Enhancement;
Focal-EIOU loss

**Corresponding author:*
liyunxue@bgu.edu.cn
(Li, Y.X.)

Article history:
Received 26 June 2025
Revised 22 June 2026
Accepted 25 June 2026

References

[1] Xu, J., Sun, Q., Han, Q.-L., Tang, Y. (2025). When embodied AI meets Industry 5.0: Human-centered smart manufacturing, *IEEE/CAA Journal of Automatica Sinica,* Vol. 12, No. 3, 485-501, doi: 10.1109/JAS.2025.125327.

[2] Zhou, M.C. (2022). Editorial: Evolution from AI, IoT and big data analytics to metaverse, *IEEE/CAA Journal of Automatica Sinica,* Vol. 9, No. 12, 2041-2042, doi: 10.1109/JAS.2022.106100.

[3] Duan, J., Yu, S., Tan, H.L., Zhu, H., Tan, C. (2022). A survey of embodied AI: From simulators to research tasks, *IEEE Transactions on Emerging Topics in Computational Intelligence,* Vol. 6, No. 2, 230-244, doi: 10.1109/TETCI.2022.3141105.

[4] Song, Y., Sun, P., Liu, H., Li, Z., Song, W., Xiao, Y., Zhou, X. (2024). Scene-driven multimodal knowledge graph construction for embodied AI, *IEEE Transactions on Knowledge and Data Engineering,* Vol. 36, No. 11, 6962-6976, doi: 10.1109/TKDE.2024.3399746.

[5] Ren, S., He, K., Girshick, R., Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 39, No. 6, 1137-1149, doi: 10.1109/tpami.2016.2577031.

[6] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A.C. (2016). SSD: Single shot multibox detector, In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.), *Computer vision – ECCV 2016. ECCV 2016. Lecture notes in computer science,* Vol. 9905, Springer, Cham, Switzerland, 21-37, doi: 10.1007/978-3-319-46448-0_2.

[7] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). You only look once: Unified, real-time object detection, In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Las Vegas, Nevada, USA, 779-788, doi: 10.1109/CVPR.2016.91.

[8] Yin, Q., Yang, W., Ran, M., Wang, S. (2021). FD-SSD: An improved SSD object detection algorithm based on feature fusion and dilated convolution, *Signal Processing: Image Communication*, Vol. 98, Article No. 116402, doi: 10.1016/j.image.2021.116402.

[9] Moon, J., Son, M., Oh, B., Jin, J., Shin, Y. (2024). Automatic voltage stabilization system for substation using deep learning, *Computer Science and Information Systems*, Vol. 21, No. 2, 437-452, doi: 10.2298/CSIS220509050M.

[10] Pang, J.L. (2023). Adaptive fault prediction and maintenance in production lines using deep learning, *International Journal of Simulation Modelling,* Vol. 22, No. 4, 734-745, doi: 10.2507/IJSIMM22-4-CO20.

[11] Jian, Z.-J., Su, M.-H. (2024). Surface defect detection in industrial parts using the YOLO-based method with enhanced spatial features and attention mechanism, In: *Proceedings of 16th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, Takamatsu, Japan, 511-515, doi: 10.1109/IIAI-AAI63651.2024.00098.

[12] Hsu, W.-Y., Lin, W.-Y. (2021). Ratio-and-scale-aware YOLO for pedestrian detection, *IEEE Transactions on Image Processing*, Vol. 30, 934-947, doi: 10.1109/tip.2020.3039574.

[13] Badmos, O., Kopp, A., Bernthaler, T., Schneider, G. (2020). Image-based defect detection in lithium-ion battery electrode using convolutional neural networks, *Journal of Intelligent Manufacturing*, Vol. 31, No. 4, 885-897, doi: 10.1007/s10845-019-01484-x.

[14] Lee, J., Hwang, K. (2022). YOLO with adaptive frame control for real-time object detection applications, *Multimedia Tools and Applications*, Vol. 81, No. 25, 36375-36396, doi: 10.1007/s11042-021-11480-0.

[15] Marco-Detchart, C., Rincon, J.A., Carrascosa, C., Julian, V. (2024). Evaluation of deep learning techniques for plant disease detection, *Computer Science and Information Systems*, Vol. 21, No. 1, 223-243, doi: 10.2298/CSIS221222073M.

[16] Zhao, Q., Fang, S., Li, Y., Shang, H., Zhang, H., Shen, T. (2025). Track defect detection based on improved YOLOv5s, *IEEE Transactions on Industrial Informatics*, Vol. 21, No. 4, 3346-3355, doi: 10.1109/TII.2024.3523593.

[17] Liu, J., Zhou, G. (2024). Learning discriminative representations through an attention mechanism for image-based person re-identification, *Computer Science and Information Systems*, Vol. 21, No. 4, 1483-1498, doi: 10.2298/CSIS230829044L.

[18] Zhu, G., Liu, Y., Wang, J. (2024). Multi-frame network feature fusion model and self-attention mechanism for vehicle lane line detection, *Computer Science and Information Systems*, Vol. 21, No. 4, 1699-1723, doi: 10.2298/CSIS240314054Z.

[19] Lv, Q.H., Chen, J., Chen, P., Xun, Q.F., Gao, L. (2024). Flexible Job-shop Scheduling Problem with parallel operations using Reinforcement Learning: An approach based on Heterogeneous Graph Attention Networks, *Advances in Production Engineering & Management*, Vol. 19, No. 2, 157-181, doi: 10.14743/apem2024.2.499.

[20] Yang, J., Zhang, Z., Li, W., Wang, X., Yang, S., Yang, C. (2025). Underwater acoustic target classification using auditory fusion features and efficient convolutional attention network, *IEEE Sensors Letters*, Vol. 9, No. 3, Article No. 7001304, 1-4, doi: 10.1109/LSENS.2025.3541593.

[21] Zhou, K., Haimudula, A., Tang, W. (2024). Dual-branch convolution network with efficient channel attention for EEG-based motor imagery classification, *IEEE Access*, Vol. 12, 74930-74943, doi: 10.1109/ACCESS.2024.3404634.

[22] Wei, S., Shen, S., Liu, D., Song, Y., Gao, R., Wang, C. (2024). Coordinate attention enhanced adaptive spatiotemporal convolutional networks for traffic flow forecasting, *IEEE Access*, Vol. 12, 140611-140627, doi: 10.1109/ACCESS.2024.3460405.

[23] Zhang, Y.-P., Zhang, Q., Kang, L., Luo, Y., Zhang, L. (2022). End-to-end recognition of similar space cone–cylinder targets based on complex-valued coordinate attention networks, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 60, 1-14, Article No. 5106214, doi: 10.1109/TGRS.2021.3115624.

[24] Woo, S., Park, J., Lee, J.-Y., Kweon, I.S. (2018). CBAM: Convolutional block attention module, In: *Proceedings of the European Conference on Computer Vision (ECCV)*, 3-19, doi: 10.1007/978-3-030-01234-2_1.

[25] Zong, H., Shuai, B. (2025). Manufacturing process quality prediction via temporal knowledge graph reasoning with adaptive multi-scale temporal path fusion and self-attention mechanism, *Advances in Production Engineering & Management*, Vol. 20, No. 3, 380-390, doi: 10.14743/apem2025.3.547.

[26] Ding, L., Wang, L., Wang, Y., Yu, S., Xiao, J. (2024). A text recognition algorithm based on a dual-attention mechanism in complex driving environment, *Tehnički Vjesnik – Technical Gazette,* Vol. 31, No. 1, 247-253, doi: 10.17559/TV-20231023001052.

[27] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N. (2021). An image is worth 16×16 words: Transformers for image recognition at scale, *arXiv*, doi: 10.48550/arXiv.2010.11929.

[28] Hao, H., Yu, X. (2024). YOLOv8-UCB: Visual detection of pouch battery using improved YOLOv8, *IEEE Access*, Vol. 12, 194899-194910, doi: 10.1109/ACCESS.2024.3520181.

[29] Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R., Ren, D. (2020). Distance-IoU loss: Faster and better learning for bounding box regression, In: *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34, No. 7, 12993-13000, doi: 10.1609/aaai.v34i07.6999.

[30] Zheng, H., Gao, X., Yang, X., Jing, G., Yang, M., Liu, Y. (2025). A multi-objective feature selection and self-paced ensemble framework for semiconductor defect detection, *Advances in Production Engineering & Management*, Vol. 20, No. 4, 458-474, doi: 10.14743/apem2025.4.552.

[31] Wang, S.-Y., Qu, Z., Gao, L.-Y. (2024). Multi-spatial pyramid feature and optimizing focal loss function for object detection, *IEEE Transactions on Intelligent Vehicles*, Vol. 9, No. 1, 1054-1065, doi: 10.1109/TIV.2023.3282996.