# REFERENCES

[1] L. Zhou, L. Zhang, and N. Konz, "Computer vision techniques in manufacturing," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 53, no. 1, pp. 105–117, 2022.

[2] F. M. Talaat and H. ZainEldin, "An improved fire detection approach based on yolo-v8 for smart cities," Neural Computing and Applications, vol. 35, no. 28, pp. 20 939–20 954, 2023.

[3] X. Liu, J. Yang, C. Zou, Q. Chen, X. Yan, Y. Chen, and C. Cai, "Collaborative edge computing with fpga-based cnn accelerators for energy-efficient and time-aware face tracking system," IEEE Transactions on Computational Social Systems, vol. 9, no. 1, pp. 252–266, 2021.

[4] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom, "nuscenes: A multimodal dataset for autonomous driving," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 11 621–11 631.

[5] Y. Wei, S. Tran, S. Xu, B. Kang, and M. Springer, "Deep learning for retail product recognition: Challenges and techniques," Computational intelligence and neuroscience, vol. 2020, 2020.

[6] B. Santra, A. K. Shaw, and D. P. Mukherjee, "Part-based annotation-free fine-grained classification of images of retail products," Pattern Recognition, vol. 121, p. 108257, 2022.

[7] W. Wang, Y. Cui, G. Li, C. Jiang, and S. Deng, "A self-attention-based destruction and construction learning fine-grained image classification method for retail product recognition," Neural Computing and Applications, vol. 32, pp. 14 613–14 622, 2020.

[8] R. Ghosh, "Product identification in retail stores by combining faster r-cnn and recurrent neural network," Multimedia Tools and Applications, vol. 83, no. 3, pp. 7135–7158, 2024.

[9] H. Sun, J. Zhang, and T. Akashi, "Templatefree: product detection on retail store shelves," IEEJ Transactions on Electrical and Electronic Engineering, vol. 15, no. 2, pp. 242–251, 2020.

[10] B. B. Wantania, S. R. Sompie, and F. D. Kambey, "Penerapan pendeteksian manusia dan objek dalam keranjang belanja pada antrian di kasir," Jurnal Teknik Informatika, vol. 15, no. 2, pp. 101–108, 2020.

[11] Q. Song, S. Li, Q. Bai, J. Yang, X. Zhang, Z. Li, and Z. Duan, "Object detection method for grasping robot based on improved yolov5," Micromachines, vol. 12, no. 11, p. 1273, 2021.

[12] W. Lestari and S. Sumarlinda, "Implementation of k-nearest neighbor (knn) and suport vector machine (svm) for clasification cardiovascular disease," INTERNATIONAL JOURNAL OF MULTI SCIENCE, vol. 2, no. 10, pp. 30–36, 2022.

[13] G. Jocher, A. Chaurasia, and J. Qiu, "Ultralytics yolov8," 2023. [Online]. Available: https://github.com/ultralytics/ultralytics

[14] G. Jocher, "Ultralytics yolov5," 2020. [Online]. Available: https://github.com/ultralytics/yolov5

[15] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2023, pp. 7464–7475.

[16] J. Terven and D. Cordova-Esparza, "A comprehensive review of yolo: From yolov1 to yolov8 and beyond. arxiv 2023," arXiv preprint arXiv:2304.00501.

[17] H.-K. Jung and G.-S. Choi, "Improved yolov5: Efficient object detection using drone images under various conditions," Applied Sciences, vol. 12, no. 14, p. 7255, 2022.

[18] D. Reis, J. Kupec, J. Hong, and A. Daoudi, "Real-time flying object detection with yolov8," arXiv preprint arXiv:2305.09972, 2023.

[19] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in International conference on machine learning. PMLR, 2020, pp. 1597–1607.

[20] D. Torpey and R. Klein, "Deepset simclr: Self-supervised deep sets for improved pathology representation learning," arXiv preprint arXiv:2402.15598, 2024.

[21] K. Nguyen, Y. Nguyen, and B. Le, "Semi-supervising learning, transfer learning, and knowledge distillation with simclr," arXiv preprint arXiv:2108.00587, 2021.

[22] V. Margapuri and M. Neilsen, "Classification of seeds using domain randomization on self-supervised learning frameworks," in 2021 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, 2021, pp. 01–08.

[23] J. Ma, H. Duan, C. Yang, X. Wang, Y. Niu, and Y. Han, "Simclr-unet: An ecg feature wave segmentation algorithm based on a self-supervised learning strategy," in Proceedings of the 2022 4th International Conference on Robotics, Intelligent Control and Artificial Intelligence, 2022, pp. 1354–1359.

[24] T.-Y. Lin, A. RoyChowdhury, and S. Maji, "Bilinear cnn models for fine-grained visual recognition," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 1449–1457.

[25] D. S. Rao, R. B. Ch, V. S. Kiran, N. Rajasekhar, K. Srinivas, P. S. Akshay, G. S. Mohan, and B. L. Bharadwaj, "Plant disease classification using deep bilinear cnn," Intell. Autom. Soft Comput, vol. 31, no. 1, pp. 161–176, 2022.

[26] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.

[28] Z. Geng, "Improve unsupervised machine learning model on fruits by using vgg-16," Highlights in Science, Engineering and Technology, vol. 85, pp. 1135–1140, 2024.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[30] M. Klasson, C. Zhang, and H. Kjellstrom, "A hierarchical grocery store image ¨ dataset with visual and semantic labels," in 2019 IEEE winter conference on applications of computer vision (WACV). IEEE, 2019, pp. 491–500.

[31] M. Klasson, C. Zhang, and H. Kjellstrom, "Using variational multi-view learning for classification of grocery items," Patterns, vol. 1, no. 8, 2020.

[32] M. Ahmed, K. A. Hashmi, A. Pagani, M. Liwicki, D. Stricker, and M. Z. Afzal, "Survey and performance analysis of object detection in challenging environments," 2021

[33] Carranza-Garcia, M., et al. "On the performance of one-stage and two-stage object detectors in autonomous vehicles using camera data," in *Remote Sensing*, vol. 13, no. 1, pp. 89, 2020.

[34] H. Zhang, R. Cloutier. "Review on one-stage object detection based on deep learning," in *EAI Endorsed Transactions on e-Learning*, vol. 7, no. 23, pp. e5–e5, 2021.

[35] Tian, Z., Shen, C., Chen, H., & He, T. (2019). FCOS: Fully convolutional one-stage object detection. arXiv 2019. *arXiv preprint arXiv:1904.01355*.

[36] L. Du, R. Zhang, X. Wang, "Overview of two-stage object detection algorithms," in *Journal of Physics: Conference Series*, 2020, pp. 012033.

[37] Krishna, M., et al. "Image classification using deep learning," in *International Journal of Engineering & Technology*, vol. 7, no. 2.7, pp. 614–617, 2018.

[38] Chaudhary, Amit. 2020. "The Illustrated SimCLR Framework." March 4, 2020. https://amitness.com/posts/simclr.html.

[39] Cho, T., et al. "Deep-ACTINet: End-to-end deep learning architecture for automatic sleep-wake detection using wrist actigraphy," in *Electronics*, vol. 8, no. 12, pp. 1461, 2019.