International Journal of Innovative Research in Computer Science and Technology (IJIRCST)
ISSN(Online): 2347-5552, Volume-12, Issue-2, March 2024
https://doi.org/10.55524/ijircst.2024.12.2.13
Article ID IJIRD-1278, Pages 75-80
www.ijircst.org

Advancements in AI for Oncology: Developing an Enhanced YOLOv5-based Cancer Cell Detection System

Xin Chen[1], Yuxiang Hu[2], Ting Xu[3], Haowei Yang[4], and Tong Wu [5]

[1] Department of Computer Science, Cornell University, Ithaca, USA
[2] Whiting School of Engineering, Johns Hopkins University, Baltimore, USA
[3] Department of Computer Science, University of Massachusetts Boston, Boston, USA
[4] Cullen College Of Engineering, University of Houston, Houston, USA
[5] School of Engineering and Technology, University of Washington, Seattle, USA

Correspondence should be addressed to Xin Chen; xc235@cornell.edu

Received 5 March 2024; Revised 18 March 2024; Accepted 28 March 2024

ABSTRACT- As artificial intelligence (AI) theory becomes more sophisticated and its utilization spreads across daily life, education, and professional settings, the adoption of AI for medical diagnostic and service purposes stands as a logical progression in the evolution of medical technologies. This document outlines a novel approach to detecting cancer cell targets using a deep learning-based system, marking a critical step towards integrating AI into cancer diagnostics. The process of detecting cancer cell targets entails the localization of cell types within images of cells. By capitalizing on the strengths of the YOLOv5 model—a deep learning-driven, end-to-end, real-time object detection framework known for its efficiency, superior performance, adaptability, and user-friendly PyTorch integration—this research presents an enhanced YOLOv5 model incorporating both a feature pyramid network and the original YOLOv5 architecture. The ultimate aim is to facilitate precise detection of targets in cancer cell images. The experimental data demonstrate the system's negligible error rate in detection, swift processing capabilities, and exceptional reliability.

KEYWORDS- Deep Learning, Target Detection, Neural Network**,** Cancer Cell Detection, YOLOv5.

I. INTRODUCTION

Cancer poses a significant global health challenge, with early detection and diagnosis being crucial for patient survival rates and treatment outcomes. In the field of medical imaging, cancer cell detection is a key task that involves accurately identifying potential cancerous regions from complex cell images, allowing for timely therapeutic interventions.

Traditional methods for cancer cell detection often rely on subjective judgment and experience of physicians, leading to limitations such as time-consuming manual operations and susceptibility to subjective bias. However, with the advancement of deep learning technology, particularly in the field of object detection [1][2][3][4], automated cancer cell detection methods based on deep learning are becoming an increasingly prominent trend[5][6].

However, conventional object detectors cannot be applied directly to cancer cell detection due to various challenges, one of which arises from the limited inter-class variation observed between the malignant and the benign cancer cells [7]. Apart from algorithmic challenges, automated cancer diagnosis also suffers from difficulties in real-world deployments.

Medical devices for cancer diagnosis are deployed in vast numbers in challenging environments and have a high demand in terms of robustness. Such an elevated complexity of cancer diagnosis requires the awareness of platform and environmental constraints and optimized utilization of computational resources [8] [9].

Acknowledging the challenges inherent in cancer cell detection, we propose an improved lightweight YOLOv5 object detection network tailored for cancer cell images[10]. The objective is to strike a balance between energy consumption, accuracy, and speed. The contributions of this work are delineated as follows:

Our method fully leverages YOLOv5's fast convergence rate on datasets and strong model customizability[11]. Through careful design and optimization, it significantly reduces the number of model parameters, lowers storage overhead, and notably accelerates the inference, while maintaining detection accuracy. Our method shows good generalizability, not only holding significant implications for the field of cancer cell detection but also providing valuable insights and references for automated detection tasks in other medical imaging domains[12][13][14].

II. CANCER CELL TARGET DETECTION SCHEME

The cancer cell detection framework consists of several key components: a module for augmenting images of cancer cells to create an improved collection of images[15]; a preprocessing module that assigns labels to images of both cancerous and non-cancerous cells in this enhanced collection, tagging them with both category and spatial labels to prepare a processed dataset; a module designed to determine anchor aspect ratios for this dataset[16]; a module for predicting bounding boxes and identifying cancer cells simultaneously by extracting image features through a Feature Pyramid Network (FPN) [17] consisting of Cross Stage Partial Network (CSPNet) [18] and Perceptual

Adversarial Network (PAN) [19]; and a loss function value (Generalized Intersection over Union, GIoU_Loss) for assessing the detection system's effectiveness by computing the difference between the predicted bounding boxes and the ground truth and then updating the network's parameters via back-propagation.

The enhanced YOLOv5 model, an advancement from YOLOv4, presents a significant improvement due to its dynamic scaling capability across various channels [20]. This advanced cancer cell detection framework utilizes the YOLOv5 model as its core, leveraging multi-scale features for enhanced detection. Initially, the system processes images of cancer cells for data augmentation[21]. Subsequently, it labels images of cancer cells and other types of cells within the augmented dataset with category and spatial information, organizing them into distinct sets for training, cross-validation, and performance evaluation[22]. Following this organization, the system calculates target anchor boxes for cancer cells, conducts object detection through an FPN with multiple CSP blocks, and ultimately assesses the detection performance using the GIoU_Loss value[23]. This comprehensive process ensures the generation of reliable cancer cell detection outcomes, as depicted in Figure 1.

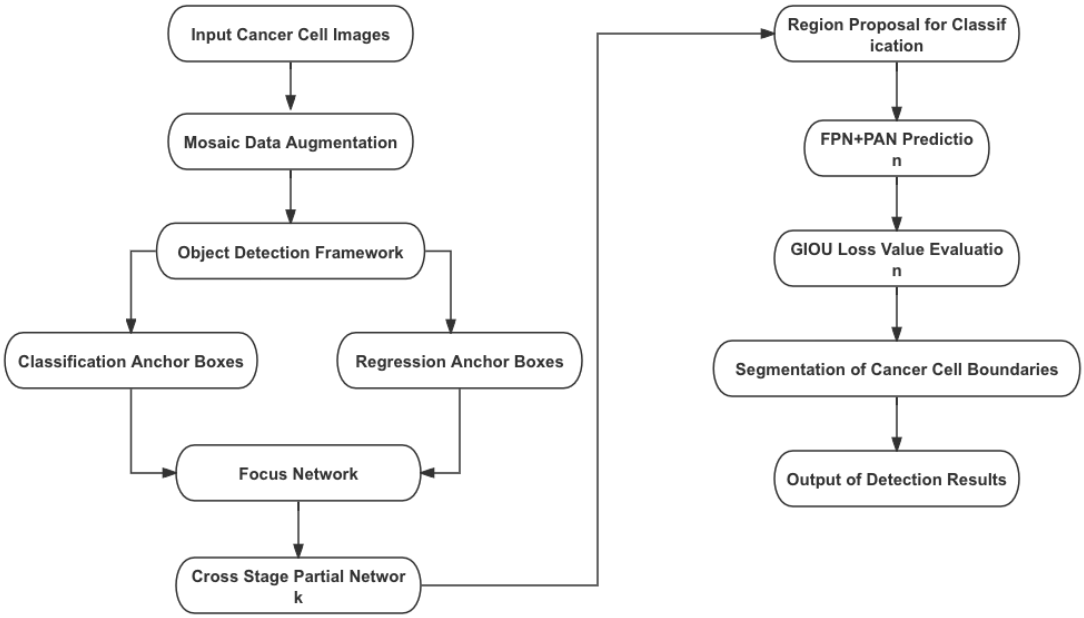

Figure 1: Schematic Diagram of the Cancer Cell Target Detection System

A. Enhanced YOLOv5 Network Model

Utilizing the strengths of residual structures from the Deep Residual Network (ResNet), the YOLOv5 enhances its network depth while preventing convergence difficulties linked to gradient vanishment. This improvement is made possible by adding a shortcut path between the input and the output. This shortcut path omits the pooling and fully connected layers in the middle of the block, instead employing adjustments in convolutional kernel strides for modifying tensor sizes[24]. Echoing the design principles of its predecessors, it scales down the feature output to one thirty-second of the input size, necessitating that the input image's resolution be divisible by 32. The model also applies methods like resizing and merging to increase tensor sizes, facilitating more comprehensive information extraction[25]. This merging process occurs in the intermediate and later stages of Darknet-53, followed by an upsampling process[26]. The numeral 53 in Darknet-53 denotes the presence of 53 convolutional layers out of the first 74, with the subsequent 22 layers serving as residual connections; the layers from 75 to 105 are dedicated to feature fusion layers, implementing multiscale detection capable of producing outputs at three different scales - 76x76, 38x38, and 19x19 - for identifying objects of varying sizes, with each scale predicting three anchor boxes.

Outlined in Figure 2, the refined YOLOv5 network structure developed in this research is segmented into four principal sections: the input segment, the central backbone network, the Neck portion, and the output stage for predictions [27][28]. At the input phase, techniques such as Mosaic data augmentation and adaptive anchor box calculations are employed to visualize anticipated positive sample anchor boxes[29]. The backbone network is enhanced with Focus and CSP architectures, supporting the creation of models of different sizes and complexities. The Neck section makes use of an integrated FPN+PAN architecture and the prediction stage operates through the GIoU_Loss mechanism.

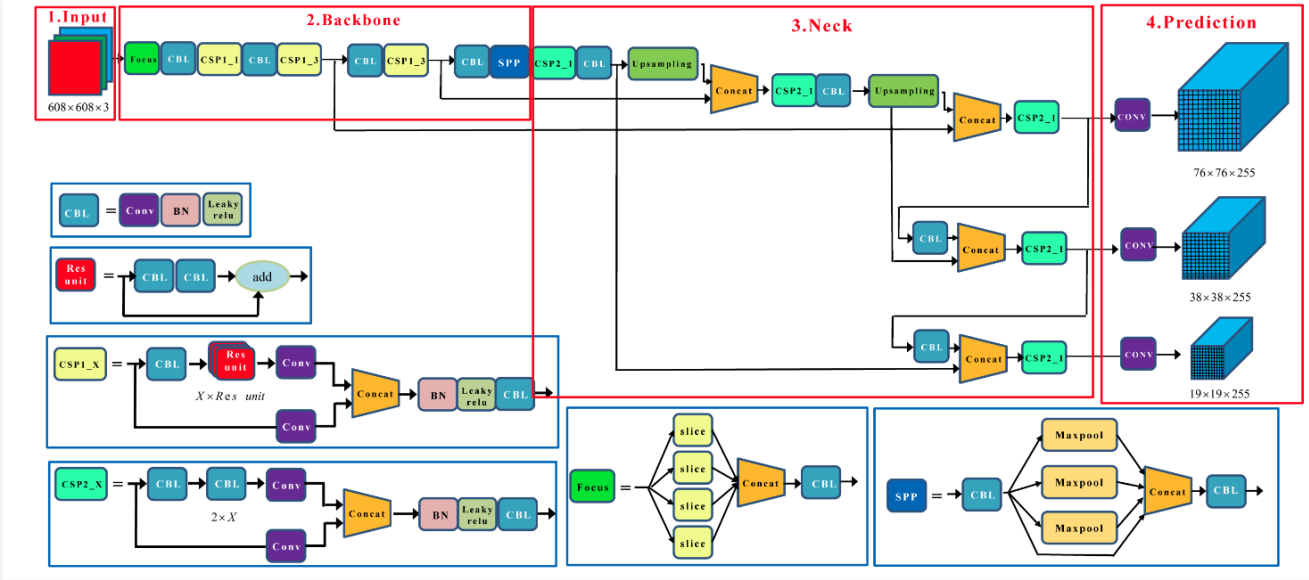

Figure 2: Network Structure Model of YOLOv5

This upgraded YOLOv5 model presents several key improvements:

- It incorporates a strategic approach to matching positive sample anchor boxes in proximity, which boosts the count of positive samples.
- It supports the development of models with diverse levels of complexity via adjustable configuration parameters.
- It improves the overall system performance with inbuilt strategies for hyperparameter tuning.
- It leverages Mosaic data augmentation to enhance the detection accuracy for smaller objects.

B. Mosaic Data Enhancement

Enhancing the model's ability to generalize and mitigating overfitting can be significantly achieved through data augmentation[30][31]. Considering the diminutive nature of cancer cells and the limited availability of training datasets for their detection, which are further compounded by the high costs of labeling new data, data augmentation serves as a critical tool. It artificially expands the training dataset through various transformations or modifications[32]. Enhancing the diversity of input images in this way, the targeted detection model becomes more adept at handling images captured in different settings. The YOLOv5 input stage capitalizes on Mosaic data enhancement's benefits by diversifying the dataset via random scaling, placement, and cropping for amalgamation, especially aiding in the enhancement of detection contexts for small entities and improving the detection rates for diminutive targets such as cancer cells [33].

C. Dynamic Anchor Box Optimization

The YOLOv5 framework allows for the presetting of anchor box dimensions tailored to specific datasets. During the training phase, the model uses these pre-set anchor boxes to generate predicted boxes, then compares these to the actual boxes to identify discrepancies. Adjustments to the network's parameters are made iteratively based on these discrepancies[34][35]. YOLOv5 incorporates the initial anchor box dimensions directly into its code, dynamically deriving the most suitable anchor box dimensions for a dataset with each training iteration [36]. This procedure involves: Applying random data enhancements and transformations to all bounding boxes in the set; Calculating optimal recall rates for bounding boxes using the initial anchor box dimensions, through a process that computes the aspect ratio values for n bounding boxes and compares these against 9 anchor box dimensions, selecting the closest matches and calculating a matching ratio. This process aims to achieve the highest possible recall rate; Should the recall rate meet or exceed 0.98, further optimization is deemed unnecessary. If it falls below this threshold, anchor box dimensions are refined using a combination of genetic algorithms and k-means clustering, with the aim of achieving the highest recall rate.

D. Multilevel Prediction via FPN

The Feature Pyramid Network (FPN) employs a dual approach, integrating bottom-up and top-down methodologies to merge lower-level detail features with higher-level abstract features, thus enriching the feature map's detail and resolution[37]. In the context of detecting cancer cell targets, this approach surpasses the capabilities of the SSD (single shot multibox detector) algorithm by allowing each layer within the FPN to operate independently, as illustrated in Figure 3. FPN enhances predictions by conducting separate analyses on each feature layer, blending upsampling with features from preceding layers to extract deeper insights, and applying a 3x3 convolution to each combined output to mitigate the effects of upsampling aliasing[38]. The FPN execution unfolds in three stages: Generating variably-sized features in a bottom-up fashion; Refining these features in a top-down manner; Enabling the extracted features to be linked with the varying dimensions output by the convolutional neural network (CNN).

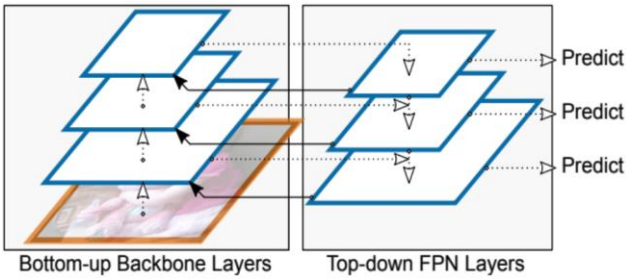

Figure 3: Implementation of FPN in the Cancer Cell Detection Model

E. Enhancements through Perceptual Adversarial Network (PAN)

Building on the FPN, the introduction of the Perceptual Adversarial Network (PAN) introduces a new layer to the pyramid, aimed at overcoming FPN's limitations in conveying location information, which primarily boosts semantic detail[39]. By channeling lower-layer detail features upward, PAN creates a composite structure enriched with both semantic and locational insights [40]. PAN differentiates between the model's outputs and actual images, refining detection capabilities through continuous feedback. Perceptual loss evaluates these differences by extracting output and real image features using a pre-trained image classification network, focusing on high-level attributes like content and texture[41]. Minimizing the disparities between these features, the model learns to transform the input image into an output that mirrors the actual image's high-level characteristics.

F. Optimization through GIoU_Loss

The Generalized Intersection Over Union (GIoU_Loss) loss function introduces a solution to the gradient vanishing dilemma in Intersection Over Union (IoU) calculations[42][43]. It provides a methodology for assessing bounding box prediction loss, especially in scenarios where candidate boxes do not intersect, ensuring the loss does not default to 0. Loss functions play a vital role in evaluating the performance of deep learning prediction models by quantifying the difference between predicted and actual outcomes[44]. In object detection, comparing the predicted bounding boxes against the actual annotated boxes is essential for computing loss and assessing the detection model's efficacy.

III. EXPERIMENTAL DESIGN

The detection model initiates its process with a feature extraction network, which is tasked with generating feature maps at varying scales. Utilizing these maps, a multi-layered predictive network identifies the positions and classifications of targets across different scales [45]. The endeavor to detect smaller targets encounters specific technical hurdles:

- The scarcity of usable features, as smaller targets comprise a smaller fraction of the dataset. Due to their diminutive size, these targets embody less information and suffer from reduced resolution, complicating the task of distilling unique features. Furthermore, they pose annotation challenges and are particularly vulnerable to errors.
- The need for precise localization. Small targets are highly susceptible to environmental changes, making accurate localization critical. A minor shift, even by one pixel, in prediction can substantially affect the outcome for small targets.
- The problem of sample imbalance and target clustering. The practice of setting thresholds for classifying anchor boxes as positive samples can cause imbalances across various target sizes. Small targets tend to cluster, leading to potential omissions of these targets by the predictive network as a result of non-maximum suppression eliminating correct predictions or due to the closeness of the boxes, which complicates model convergence.

To address the intricacies of intelligent detection of cancer cell targets, the feature extraction network is designed to embody several key features:

- It boasts potent capabilities for extracting nonlinear features.
- The network architecture is crafted to minimize computational complexity, facilitating rapid operation.
- It demonstrates enhanced robustness, mitigating the risks of training issues such as gradient vanishing or exploding.

For this study, the dataset is constituted of 2,000 breast cancer images, 2,000 cervical cancer images accessible via official platforms, 2,000 images of healthy cells from a hospital's pathology department, and 1,000 images depicting a mixture of healthy and cancerous cells, amounting to a total of 7,000 cell images. Following data augmentation, the dataset expanded to 10,000 images. A split of 80% of these images is utilized for training purposes, with the remaining 20% earmarked for validation. Performance evaluation employs standard precision *(P)* and recall *(R)* metrics to determine the Average Precision *(AP)*, alongside the F1-score, which serves as an additional critical performance metric [46]. The methodology behind these metrics is detailed as follows:

$$R = \frac{T}{T + F} \quad (1)$$

$$R = \frac{T}{T + \mathrm{X}} \quad (2)$$

$$\text{F1-score} = \frac{2 \times R \times P}{R + P} \quad (3)$$

In the formula, *T* represents true positives, *F* denotes false negatives, and *X* stands for false positives. Table 1 illustrates the detection result metrics for the cancer cell target detection model.

Table 1: Target Detection Result Metrics

	Recall	AP	Precision	F1-score
Result	0.807	0.77	0.911	0.867

The effectiveness of detecting cell images containing varying numbers of cancer cells is shown in Figure 4. These images include normal cell images, cell images with a small number of cancer cells, and cell images with a mix of several types of cancer cells among normal cells.

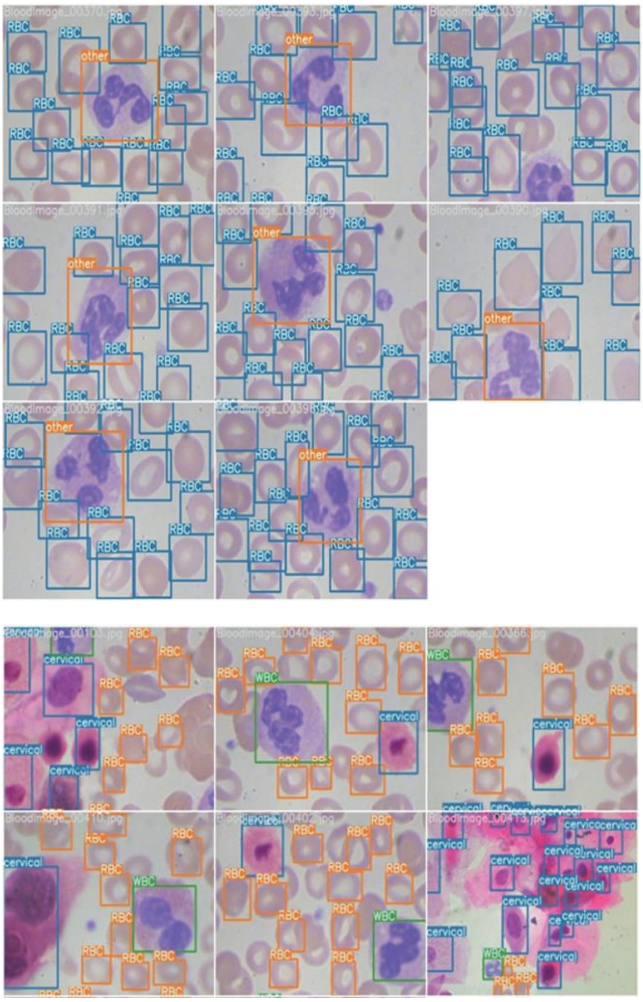

Figure 4: Cancer Cell Target Detection Results

IV. CONCLUSION

This study stands at the forefront of the intersection between deep learning and the medical field, introducing an improved lightweight YOLOv5 target detection network that successfully addresses the key technical challenges in implementing a cancer cell detection system. Our approach not only accurately identifies the presence of cancer cells in complex cell images but also provides detailed information on their location and size, thereby offering crucial decision support for subsequent cancer diagnosis and treatment.

By fully leveraging the rapid convergence and high customizability of YOLOv5, this project has effectively resolved a range of issues traditionally encountered by target detection networks in practical medical applications, such as excessive model weight parameters, high storage consumption, and the difficult balance between speed and accuracy. Moreover, the improved lightweight YOLOv5 network significantly reduces model complexity and operational time while maintaining high detection accuracy, substantially enhancing the practicality and acceptability of cancer cell detection.

Future work will focus on further optimizing the model structure to lower computational costs while maintaining or even improving detection accuracy. Additionally, considering the diversity of cancer types, expanding the model's adaptability to cover more types of cancer diagnoses is our next major research direction. We also plan to collaborate with clinical medical institutions to more broadly apply the outcomes of this study in actual cancer detection and diagnosis, contributing to the early detection and treatment of cancer patients.

Through this research, we have demonstrated the immense potential of artificial intelligence technology in the healthcare field, especially in the crucial area of cancer diagnosis. We believe that as technology continues to advance and more deeply integrates with medical practice, deep learning-based cancer cell detection systems will play an increasingly important role in the future, bringing hope and light to cancer patients worldwide.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] S. Diao et al., "Deep multi-magnification similarity learning for histopathological image classification," IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 3, pp. 1535-1545, 2023.

[2] W. Huang, A. Rezvani, H. Chen, Y. Ni, S. Yun, S. Jeong, and M. Imani, "A Plug-in Tiny AI Module for Intelligent and Selective Sensor Data Transmission," arXiv preprint arXiv:2402.02043, 2024.

[3] S. Yun, H. Chen, R. Masukawa, H. E. Barkam, A. Ding, W. Huang, et al., "HyperSense: Accelerating Hyper-Dimensional Computing for Intelligent Sensor Data Processing," arXiv preprint arXiv:2401.10267, 2024.

[4] H. Chen, W. Huang, Y. Ni, S. Yun, F. Wen, H. Latapie, and M. Imani, "TaskCLIP: Extend Large Vision-Language Model for Task Oriented Object Detection," arXiv preprint arXiv:2403.08108, 2024.

[5] Z. Deng, Y. Xin, X. Qiu, and Y. Chen, "Weakly and semi-supervised deep level set network for automated skin lesion segmentation," in Innovation in Medicine and Healthcare: Proceedings of 8th KES-InMed 2020, pp. 145-155, Springer Singapore, 2020.

[6] S. Diao, Y. Tian, W. Hu, J. Hou, R. Lambo, Z. Zhang, et al., "Weakly supervised framework for cancer region detection of hepatocellular carcinoma in whole-slide pathologic images based on multiscale attention convolutional neural network," The American journal of pathology, vol. 192, no. 3, pp. 553-563, 2022.

[7] Z. Qi, J. Wang, X. Wu, and H. Zhao, "OCBEV: Object-Centric BEV Transformer for Multi-View 3D Object Detection," arXiv preprint arXiv:2306.01738, 2023.

[8] Z. Li, H. Zhu, H. Liu, J. Song, and Q. Cheng, "Comprehensive evaluation of Mal-API-2019 dataset by machine learning in malware detection," arXiv preprint arXiv:2403.02232, 2024.

[9] S. Li, P. Kou, M. Ma, H. Yang, S. Huang, and Z. Yang, "Application of Semi-supervised Learning in Image Classification: Research on Fusion of Labeled and Unlabeled Data," IEEE Access, 2024.

[10] Y. Xin, S. Luo, P. Jin, Y. Du, and C. Wang, "Self-Training with Label-Feature-Consistency for Domain Adaptation," in International Conference on Database Systems for Advanced Applications, pp. 84-99, Cham: Springer Nature Switzerland, 2023, April.

[11] H. Xu and J. A. Colmenares, "Admission Control with Response Time Objectives for Low-latency Online Data Systems," arXiv preprint arXiv:2312.15123, 2023.

[12] X. Yan, M. Xiao, W. Wang, Y. Li, and F. Zhang, "A Self-Guided Deep Learning Technique for MRI Image Noise Reduction," Journal of Theory and Practice of Engineering Science, vol. 4, no. 01, pp. 109–117, 2024. [Online]. Available: https://doi.org/10.53469/jtpes.2024.04(01).15

[13] L. Li, "CPSeg: Finer-grained Image Semantic Segmentation via Chain-of-Thought Language Prompting," in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 513-522, 2024.
[14] W. Huang, M. Bulut, R. van Lieshout, and K. Dellimore, "Exploration of using a pressure sensitive mat for respiration rate and heart rate estimation," in 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 298-301, IEEE, November 2021.
[15] Y. Xin, J. Du, Q. Wang, Z. Lin, and K. Yan, "VMT-Adapter: Parameter-Efficient Transfer Learning for Multi-Task Dense Scene Understanding," arXiv preprint arXiv:2312.08733, 2023.
[16] J. Chen, X. Chen, R. Wang, C. Le, A. Khalilian-Gourtani, E. Jensen, et al., "Subject-Agnostic Transformer-Based Neural Speech Decoding from Surface and Depth Electrode Signals," bioRxiv, 2024-03.
[17] H. Xu, S. Wen, A. Gimenez, T. Gamblin, and X. Liu, "DR-BW: identifying bandwidth contention in NUMA architectures with supervised learning," in 2017 IEEE International Parallel and Distributed Processing Symposium (IPDPS), pp. 367-376, IEEE, May 2017.
[18] Y. Liu, H. Yang, and C. Wu, "Unveiling patterns: A study on semi-supervised classification of strip surface defects," IEEE Access, vol. 11, pp. 119933-119946, 2023.
[19] Q. He, H. Su, J. Zhang, and X. Hou, "Frustratingly easy regularization on representation can boost deep reinforcement learning," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 20215-20225, 2023.
[20] K. He et al., "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, June 2016, pp. 770-778. [Online]. Available: https://doi.org/10.1109/CVPR.2016.90
[21] X. Chen, R. Wang, A. Khalilian-Gourtani, L. Yu, P. Dugan, D. Friedman, et al., "A Neural Speech Decoding Framework Leveraging Deep Learning and Speech Synthesis," bioRxiv, 2023-09.
[22] C. Liu, J. Cui, R. Shang, Y. Xiao, Q. Jia, and E. Gehringer, "Improving Problem Detection in Peer Assessment through Pseudo-Labeling Using Semi-Supervised Learning," International Educational Data Mining Society, 2022.
[23] J. Y. Shi et al., "Combining incremental conductance and firefly algorithm for tracking the global MPP of PV arrays," Journal of Renewable and Sustainable Energy, vol. 9, no. 2, 2017.
[24] S. Chen, L. Lu, Q. Zhang, and M. Li, "Optimal binomial reliability demonstration tests design under acceptance decision uncertainty," Quality Engineering, vol. 32, no. 3, pp. 492-508, 2020.
[25] Z. Qi, J. Wang, X. Wu, and H. Zhao, "OCBEV: Object-Centric BEV Transformer for Multi-View 3D Object Detection," arXiv preprint arXiv:2306.01738, 2023.
[26] Y. Xiao et al., "Modeling review helpfulness with augmented transformer neural networks," in 2022 IEEE 16th International Conference on Semantic Computing (ICSC), pp. 83-90, January 2022.
[27] R. Wang et al., "Distributed feedforward and feedback cortical processing supports human speech production," Proceedings of the National Academy of Sciences, vol. 120, no. 42, article e2300255120, 2023.
[28] D. Ma et al., "Implementation of computer vision technology based on artificial intelligence for medical image analysis," International Journal of Computer Science and Information Technology, vol. 1, no. 1, pp. 69-76, 2023.
[29] H. Zang, "Precision Calibration of Industrial 3D Scanners: An AI-Enhanced Approach for Improved Measurement Accuracy," Global Academic Frontiers, vol. 2, no. 1, pp. 27-37, 2024. [Online]. Available: https://gafj.org/journal/article/view/30
[30] W. Wang et al., "Enhancing Liver Segmentation: A Deep Learning Approach with EAS Feature Extraction and Multi-Scale Fusion," International Journal of Innovative Research in Computer Science & Technology, vol. 12, no. 1, pp. 26–34, 2024. [Online]. Available: https://ijircst.irpublications.org/index.php/ijircst/article/view/21
[31] R. Wang et al., "Stimulus speech decoding from human cortex with generative adversarial network transfer learning," in 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), pp. 390-394, April 2020.
[32] J. Wu et al., "SwitchTab: Switched Autoencoders Are Effective Tabular Learners," arXiv preprint arXiv:2401.02013, 2024.
[33] W. Dai et al., "Addressing Unintended Bias in Toxicity Detection: An LSTM and Attention-Based Approach," in 2023 5th International Conference on Artificial Intelligence and Computer Applications (ICAICA), pp. 375-379, November 2023.
[34] R. Yuan, S. Ding, and C. Wang, "Comparative Analysis of Machine-Learning Models for Recognizing Lane-Change Intention Using Vehicle Trajectory Data," Infrastructures, vol. 8, no. 11, p. 156, 2023.
[35] Y. Xin et al., "Parameter-Efficient Fine-Tuning for Pre-Trained Vision Models: A Survey," arXiv preprint arXiv:2402.02242, 2024.
[36] Q. Jia et al., "All-in-one: Multi-task learning BERT models for evaluating peer assessments," arXiv, 2021.
[37] H. Xu, S. Song, and Z. Mao, "Characterizing the Performance of Emerging Deep Learning, Graph, and High-Performance Computing Workloads Under Interference," arXiv:2303.15763, 2023.
[38] F. Liu, H. Wang, W. Chen, H. Sun, and Y. Duan, "Make-Your-3D: Fast and Consistent Subject-Driven 3D Content Generation," arXiv preprint arXiv:2403.09625, 2024.
[39] S. Chen, J. Wu, N. Hovakimyan, and H. Yao, "Recontab: Regularized contrastive representation learning for tabular data," arXiv preprint arXiv:2310.18541, 2023.
[40] Y. Xin, J. Du, Q. Wang, K. Yan, and S. Ding, "MmAP: Multi-modal Alignment Prompt for Cross-domain Multi-task Learning," arXiv preprint arXiv:2312.08636, 2023.
[41] T. Zhang, L. Li, C. Igel, S. Oehmcke, F. Gieseke, and Z. Peng, "LR-CSNet: low-rank deep unfolding network for image compressive sensing," in 2022 IEEE 8th International Conference on Computer and Communications (ICCC), pp. 1951-1957, IEEE, Dec. 2022.
[42] X. S. Wang, J. D. Turner, and B. P. Mann, "Constrained attractor selection using deep reinforcement learning," Journal of Vibration and Control, vol. 27, no. 5-6, pp. 502-514, 2021.
[43] N. S. Lesmana, H. Su, and C. S. Pun, "Reinventing Policy Iteration under Time Inconsistency," Transactions on Machine Learning Research, 2022.
[44] F. Ni, H. Zang, and Y. Qiao, "Smartfix: Leveraging machine learning for proactive equipment maintenance in industry 4.0," in The 2nd International scientific and practical conference “Innovations in education: prospects and challenges of today”(January 16-19, 2024), Sofia, Bulgaria, International Science Group, 2024, pp. 313.
[45] L. Li, "Hierarchical Edge Aware Learning for 3D Point Cloud," in Computer Graphics International Conference, pp. 81-92, Cham: Springer Nature Switzerland, Aug. 2023.
[46] X. S. Wang and B. P. Mann, "Attractor Selection in Nonlinear Energy Harvesting Using Deep Reinforcement Learning," arXiv preprint arXiv:2010.01255, 2020.