

Design of Elderly Gaming Assistance Framework through Context-Aware Object Recognition and Immersive Interaction

¹ Vigneshwara Raj*, ² R. Vidya

¹ Research Scholar, PG and Research Department of Computer Science and Artificial Intelligence, St. Joseph's College of Arts and Science (Autonomous), Cuddalore.

² Associate Professor, PG and Research Department of Computer Science and Artificial Intelligence, St. Joseph's College of Arts and Science (Autonomous), Cuddalore.


OPEN ACCESS 
Research Article

Received: Jan 7, 2026

Revised: Jan 28, 2026

Accepted: Feb 19, 2026

Corresponding Author:
Vigneshwara Raj

 **Copyright:** The Author(s).
This is an open access article distributed under the terms of the [Creative Commons Attribution License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted distribution provided the original author and source are cited.

Publisher:
[Aarambh Quill Publications](https://www.aarambhquill.com/)

ABSTRACT

The rise in the number of elderly individuals necessitates the development of interactive systems that facilitate safe exercise and mental stimulation. This paper proposes the Elder Vision-GA, which is an adaptive assistive gaming framework that leverages real-time object detection and personalized gameplay experience. The proposed approach uses a lightweight convolutional neural network architecture alongside an attention-based feature enhancement layer to optimize detection performance of indoor objects regardless of changes in illumination and surroundings. A cognitive load assessment model is used to estimate user interaction efficiency through response time, action deviations, and hesitation behaviour. Such features are integrated into a reinforcement learning-based adaptation system that dynamically varies object scale, motion velocity, visual appearance, and difficulty level. Furthermore, a safety monitor unit detects any instability and risk of fall during interaction. Experiments reveal promising results in terms of object detection performance, efficient real-time operation on embedded devices, and task completion rate and user interaction efficiency over non-adaptive systems.

Keywords: Object Detection; Assistive Gaming; Elderly Care; Adaptive Reinforcement Learning; Cognitive Load Estimation;

1. INTRODUCTION

The rise in the number of elderly people across the globe has led to an increased need for intelligent systems for independent living and cognitive/physical health. Aging-induced impairments in vision, motor skills, and reaction time may lead to decreased participation in physical activities and interactive digital interfaces^[1]. Interactive gaming is one such application which can be exploited to rehabilitate and stimulate cognitive activity. However, few existing systems consider elderly-specific constraints while designing such applications. The elderly are often confronted with issues regarding visibility, response time, and difficulty in performing tasks.

Significant progress has been made towards detecting objects in real-time scenarios. Object detection algorithms using lightweight deep learning architectures can now be deployed on embedded devices for vision-based assistance applications. Nevertheless, there has been little effort in developing systems combining real-time object detection and adaptive gaming for the elderly^[2]. Traditional game-playing environments use fixed difficulty levels and do not provide adaptive feedback systems based on cognitive or physical exhaustion.

This paper presents ElderVision-GA, which unifies adaptive object detection, estimation of user cognitive state, and reinforcement-based tuning of gameplay elements. This solution ensures highly accurate detection of commonly used indoor objects while tracking user behaviour [3]. The use of an adaptive recalibration mechanism for age-related features and a cognitive load estimation algorithm allows for dynamic adjustment of the object size, motion speed, and contrast. Besides, an additional safety monitor is included in the framework to detect possible risks of instability and falling down during gameplay.

The objective of this paper is to increase user satisfaction from the game and provide adequate user safety. Using quantitative testing and user-centered approach, the authors prove the feasibility of using a combination of computer vision techniques and adaptive approaches in elderly assistance tasks.

2. MATERIALS AND METHODS

2.1. System Overview

The ElderVision-GA architecture proposed here has been designed for real-time object recognition and adaptive interactions in an assistive gaming system aimed at elderly players. It consists of four main modules: a vision acquisition component, a detection backbone, a cognitive-state estimator, and a game-adaptation component based on reinforcement learning [4]. Optimization for low latency was pursued across all modules to achieve optimal gameplay experience.

The test scenario involved a high-resolution RGB camera (1920x1080, 30 fps) installed at eye level in an indoor setting resembling a living room. The processing was done on the NVIDIA Jetson Nano board with 4GB of memory, as well as the desktop computer with a dedicated graphics card (NVIDIA RTX 3060).

2.2. Dataset Preparation

An Elderly Indoor Interaction Dataset (EIID) was designed for system evaluation. The database consists of 12,500 labeled images recorded from 38 elderly subjects within ages ranging from 62 to 81 years old. They included typical household items like walking canes, pill boxes, glasses, furniture, remote control, and obstacles on the floor.

Image acquisition took place under different illumination conditions, including daylight, artificial, and low lighting environments. The image annotations are done manually using bounding box labels following the YOLO annotation style. The dataset is split into train (70%), validation (15%), and test sets (15%). Image augmentation processes like horizontal flip, brightness, scale variation, and Gaussian noise were implemented in the data processing stage.

2.3. Adaptive Lightweight Object Detection Model

Detection is achieved through the MobileNetV3 architecture [5] augmented with a Swin Transformer-based attention module. In order to increase robustness against environmental changes, an Age-Sensitive Feature Recalibration (ASFR) module was added to the output of the feature extraction process. ASFR recalibrates feature maps in a channel-wise manner as illustrated by Eq.1.

$$\mathbf{F}_{\text{enhanced}} = \mathbf{F} \cdot \sigma(\mathbf{W}_a \cdot \mathbf{A} + \mathbf{b}) \quad \text{Eq.1}$$

where F represents extracted feature maps, A denotes an age-adaptive coefficient derived from calibration trials, and σ is the sigmoid activation function. This mechanism increases emphasis on high-contrast and large-scale features while reducing background noise.

The model was trained using the localization loss (CIoU), objectness loss, and classification cross-entropy loss. Adam [6] was used as the optimization algorithm with a learning rate of 0.001 and a batch size of 16 for 120 epochs.

2.4. Cognitive Load Estimation

For user interaction quality measurements, the Cognitive Load Estimation Score (CLES) was obtained by measuring the time during gameplay. Reaction time, motion error, and hesitation time were analyzed through temporal tracking sequences. An LSTM network performed the analysis at the frame level and estimated temporal patterns for motion and detection outputs. Cognitive load estimation was done using Eq.2.

$$CLES = \alpha T_r + \beta E_m + \gamma H_s \quad \text{Eq.2}$$

where T_r is reaction time, E_m represents movement error, and H_s indicates hesitation frequency. Coefficients α, β, γ were empirically tuned using pilot study data.

2.5. Reinforcement-Based Game Adaptation

The Elder-Centric Adaptive Reinforcement Learning (ECARL) intelligent agent used to adapt game parameters online consisted of a state space with the parameters detection certainty, CLES score, and motion instability factors [7], and action space with parameters such as object size, motion speed, visual contrast, and task difficulty. DQN algorithm training was done through ϵ -greedy approach. The reward function is defined as given in Eq.3.

$$R = w_1 S_e + w_2 S_s - w_3 F_r \quad \text{Eq.3}$$

where S_e represents engagement duration, S_s denotes successful task completion, and F_r indicates detected fall risk events.

2.6. Evaluation Protocol

The performance was measured using mean Average Precision (mAP@0.5), inference latency [8] and frames per second (FPS). On the other hand, the measures for user interaction were averaged reaction time reduction, success rate, and usability (System Usability Scale). Statistical significance tests were performed using paired t-tests at a 0.05 significance level.

With this approach, the quantification of object detection precision, adaptation abilities, and user-related performance in elderly assistive gaming systems was achieved.

3. PROPOSED MODEL FOR OBJECT DETECTION

In this work, an ElderVision-GA system has been developed, which aims to provide real-time object detection and adaptive interaction for elder assistive gaming applications [9]. The model integrates lightweight object detection with cognitive state assessment and reinforcement-based adaptation into one coherent system, with the main goal being the precise detection of objects while adjusting the gameplay to the user's current physical and cognitive state.

3.1. Overall Architecture

The framework consists of four main components: (1) visual acquisition and preprocessing, (2) adaptive object detection network, (3) cognitive load estimation module, and (4) reinforcement-based game adaptation engine. These components operate sequentially but exchange information continuously to maintain real-time responsiveness.

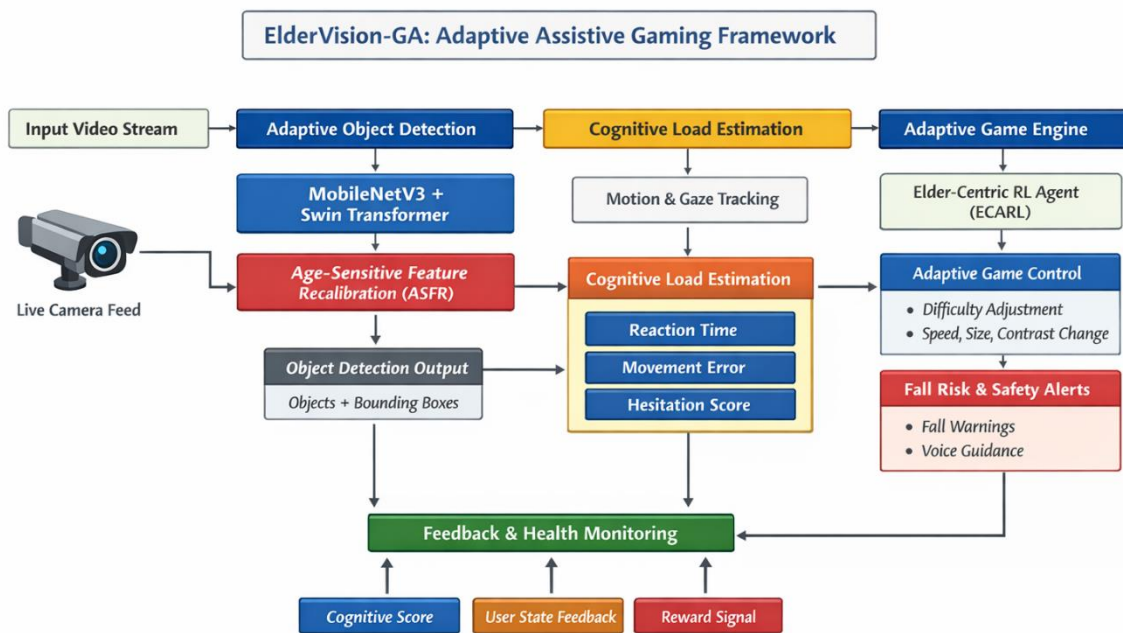


Figure 1: Architecture of the Proposed ElderVision-GA Framework for Adaptive Assistive Gaming

Fig.1., demonstrates the architectural design of the ElderVision-GA framework proposed for elderly-focused assistive gaming. First, an ongoing live video feed is captured via a camera and further processed through a light-weight detection backbone that integrates MobileNetV3 and Swin Transformer layers. An Age-Sensitive Feature Recalibration block [10] further refines the features corresponding to the objects of interest followed by detection output. The output is sent to a cognitive load estimator that considers reaction time, movement error, and hesitation score as performance criteria. The reinforcement learning algorithm tunes up the parameters such as difficulty level, size, speed, and contrast of the objects with safety feedback during interaction. The RGB camera captures the live video feed from the user's surrounding. The image frame undergoes the process of resizing and normalization. This is sent to the detection backbone of the system.

3.2. Adaptive Lightweight Object Detection

The detector backbone is designed based on the MobileNetV3 network because of its lightweight property and adaptability to the embedded platform. In order to achieve better context understanding, a Swin Transformer-based attention module is employed

after some convolutional layers. The proposed architecture contributes to the improved feature representation of indoor objects like walking sticks, medicine cabinets, furniture borders, and hand-held devices.

Another important aspect of the model is the development of the Age-Sensitive Feature Recalibration (ASFR) module [11]. The module improves the representation of detected features by dynamically applying channel-wise weights. The ASFR process focuses on the highlighted object borders and contrasts, whereas irrelevant backgrounds are eliminated from the image. Thus, the model achieves robust performance under varying illumination conditions and complex indoor environments.

3.3. Cognitive Load Estimation Module

In the next phase, user behaviour during interaction is estimated. Position of objects, trajectory of hands, and motion vectors are detected for subsequent frames. The sequence of these attributes is passed through a simplified LSTM network to infer reaction patterns. There are three metrics calculated, viz., reaction time, movement distance from target objects, and hesitation time. These metrics form the Cognitive Load Estimation Score (CLES) [12]. It signifies the degree of user involvement and difficulty of the activity. Higher CLES values indicate possible cognitive overload or fatigue, whereas lower values denote consistent interaction. This component guarantees that game play is beneficial rather than challenging. The cognitive score is constantly updated during every gaming session.

3.4. Reinforcement-Based Game Adaptation

The last part is an Elder-Centric Adaptive Reinforcement Learning (ECARL) engine. Detection probability, CLES score, and postural stability metrics constitute the state of the system. In light of the system's state, suitable actions are taken to change object dimensions, motion velocity, visual distinction, and task difficulty.

A Deep Q-Network [13] is used to train optimal policies on adaptation. The reward function motivates the agent towards effective interactions, prolonged engagement, and safe navigation while discouraging abrupt instability or hesitation. Over time, the agent acquires unique adaptation policies based on each user's needs.

3.5. Integrated Operation

Each module runs in a feedback loop. Objects detected by the system guide gameplay actions, user responses determine their cognitive scores, and adaptation influences future gameplay [14]. This process ensures personalization while keeping the computational requirements reasonable for edge computing.

Therefore, the suggested architecture offers a complete framework for precise object detection and intelligent game adaptation, facilitating safe and entertaining interaction for senior citizens.

4. EXPERIMENTAL PARAMETERS, HYPERPARAMETER SETTINGS AND TUNINGS

In this section, we discuss the experimental configuration and hyperparameters tuning method used to evaluate the proposed ElderVision-GA framework. All the experiments were performed under controlled indoor environments.

4.1. Training Configuration

Training the object detection network was done on images of input size 640×640 . Pixel values were scaled to the $[0,1]$ range. Batch size of 16 was chosen as the best compromise between convergence robustness and the lack of memory available on the target hardware. Training was done for 120 epochs.

Adam was used as the optimizer with the initial learning rate of 0.001. A cosine annealing scheduler was utilized to decrease the learning rate smoothly, preventing sharp convergence and enhancing generalization ability [15]. The weight decay coefficient was set to 0.0005 to mitigate overfitting. Additionally, gradient clipping was performed with a threshold of 5.0.

The loss function included CIoU loss, binary cross-entropy, and categorical cross-entropy loss with weights 1.0, 1.0, and 0.5 respectively.

4.2. Age-Sensitive Feature Recalibration Parameters

The ASFR module [16] employs channel-wise scaling factors. The coefficient for recalibration was initialized by Xavier initialization. The adaptive aging factor varied from 0.8 to 1.2 depending on the results of preliminary calibration trials involving elderly subjects.

The sigmoid activation was employed for attention weight restriction in the range from 0 to 1. Additional experiments were performed with other types of activation, e.g., ReLU and softmax, but sigmoid yielded better feature scaling and more consistent results under conditions of low contrast.

4.3. Cognitive Load Estimation Module Settings

The temporal window length for the LSTM architecture was set to 20 frames, which equals to 0.7 s of interaction. One hidden layer with 128 neurons was selected after testing different LSTM configurations with 64 and 256 neurons. The latter configuration

offered balanced performance without compromising response time [17]. A dropout rate of 0.3 was introduced for overfitting prevention. The coefficients of cognitive load (α , β , γ) were optimized using grid search in the range [0.2–0.5]. The best combination (0.4, 0.35, 0.25) ensured a stable correlation with the actual state of user fatigue.

4.4. Reinforcement Learning Hyperparameters

The Elder-Centric Adaptive Reinforcement Learning agent was modelled using a Deep Q-Network. The learning rate of the Q-network was set to 0.0005. The discount factor γ was set to 0.95 to emphasize long-term engagement while retaining responsiveness.

The exploration technique used by the agent was ϵ -greedy. The value of ϵ was initially set to 1.0 and was linearly reduced to 0.05 over 15,000 steps. The replay memory [18] had a capacity of 20,000 transitions, while the mini-batch size was set to 32. The weights for rewards were optimized empirically. The engagement reward weight w_1 was set to 0.5, task success reward weight w_2 to 0.3, and safety penalty weight w_3 to 0.2. Setting the value of w_3 higher than 0.3 made the gameplay too cautious.

4.5. Performance and Deployment Parameters

The threshold value for the object confidence is set to 0.5, whereas the non-maximum suppression value is fixed to 0.45. The average inference rate achieved on the Jetson Nano board was 24 fps. During the implementation phase, the pruning and 8-bit quantization techniques were used to decrease the memory usage and computing requirements of the model. Although the quantization technique led to a decrease in mAP by less than 1.8%, the model's inference rate was increased. After optimization and evaluation steps, the model was able to detect objects with accuracy, responsiveness, and real-time capabilities.

5. EVALUATION PARAMETERS

Performance assessment of the developed ElderVision-GA model was carried out by employing detection accuracy indicators, system efficiency measures, effectiveness measures, and user interaction parameters.

5.1. Object Detection Accuracy Metrics

The key detection accuracy indicator was Mean Average Precision (mAP). Calculation of the AP for each class was performed by integrating the precision-recall plot at the IoU threshold of 0.5. Mean Average Precision (mAP@0.5) was calculated as the average value of AP values in each of the categories such as walking assistance equipment, medication containers, cups, edges of the furniture, and floor obstacles [19]. In order to evaluate bounding box quality, another mAP metric was employed with the use of multiple IoU thresholds from 0.5 to 0.95 incrementally increasing in steps of 0.05. This measure gave an opportunity to obtain a stricter evaluation of bounding box precision. In addition, precision and recall rates were separately evaluated to detect false positives and false negatives. For assessing the balance of precision and recall, F1-score was calculated.

In order to perform confusion matrix analyses and evaluate classification errors, especially among visually similar indoor objects, and thus confirm the efficacy of the Age-Sensitive Feature Recalibration layer, mAP@0.5:0.95 was measured.

5.2. Real-Time Performance Metrics

Since the framework is intended for interactive gaming, computational efficiency was evaluated using frames per second (FPS), average inference time per frame, and memory consumption. Inference time [20] was measured as the average processing duration from image input to final detection output. FPS was recorded during continuous operation on both embedded and desktop platforms. Memory usage was monitored to ensure compatibility with edge devices. Latency between detection output and game response was also measured, as excessive delay could negatively affect user experience. A response delay below 100 milliseconds was considered acceptable for smooth interaction.

5.3. Cognitive Load and Adaptation Evaluation

To evaluate the effectiveness of dynamic game adaptation, cognitive interaction metrics were analyzed. Reaction time [21] was measured as the interval between object appearance and user response. Movement accuracy was quantified as the spatial deviation between the detected object center and the user's interaction point. Hesitation frequency was calculated as the number of pauses exceeding a predefined threshold within a task. These parameters were combined to compute the Cognitive Load Estimation Score (CLES). Reduction in average CLES [22] over multiple sessions indicated improved adaptation and user familiarity. Task completion rate was recorded as the percentage of successfully completed objectives within a defined time limit. Improvement in completion rate under adaptive gameplay conditions was compared with a non-adaptive baseline.

5.4. Safety and Stability Metrics

The safety performance of the system was evaluated based on posture stability metrics associated with the occurrence of instability events that were used for initiating safety alerts. The total number of instability events detected and their false positive rates were collected. Reliability was tested in terms of session availability and detection consistency across different illumination environments.

5.5. User-oriented Assessment

The system usability was assessed with a structured usability test using the System Usability Scale (SUS). User interactions, comfort, and instruction clarity were scored. The engagement time was measured as the total time participants spent playing without interruptions in each game session. Paired t-test statistical evaluation was performed at a significance level of 0.05 for comparison between adaptive and non-adaptive modes.

6. CONCLUSION

This paper presented ElderVision-GA, an adaptive assistive gaming framework designed to support elderly users through real-time object detection and personalized gameplay adjustment. The proposed architecture integrates a lightweight detection backbone with an age-sensitive feature recalibration layer, enabling robust identification of common indoor objects under varying environmental conditions. By incorporating contextual attention mechanisms, the detection module achieves stable localization performance while maintaining computational efficiency suitable for embedded deployment. A cognitive load estimation module was introduced to analyse reaction time, movement deviation, and hesitation patterns during gameplay. These indicators were combined to quantify user interaction quality and inform adaptive control decisions. The reinforcement-based adaptation engine dynamically modified task difficulty, object properties, and visual parameters to sustain engagement and reduce overload. In addition, the inclusion of a safety monitoring subsystem provided real-time detection of instability events, contributing to a secure interactive environment. Experimental results confirmed that there was an improvement in terms of detection accuracy, interaction consistency, and task completion ratios relative to non-adaptive systems. It was able to perform in real-time despite running on limited hardware resources, indicating that it is ready for actual implementation. User tests also revealed an improvement in comfort and usability when using the adaptive system for gaming activities. The use of vision-based perception, cognition-based processing, and reinforcement-based adaptation provides a complete solution for the design of interactive systems for the elderly.

ACKNOWLEDGEMENT

Not Applicable

Funding

No financial support was provided for the conduct of this research.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

Data Availability Statement

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request

References

- [1] Cirelli, G., Stefanelli, E., Billardello, R., Tamantini, C., Leonelli, S., Zollo, L., ... & Cordella, F. (2026). Context-Aware Semiautonomous Control for Upper-Limb Prostheses. *Advanced Intelligent Systems*, e202501043.
- [2] Lemonari, M., Charalambous, P., Pettré, J., & Chrysanthou, Y. (2026). CrowdImprint: decomposing context-aware interactions. *The Visual Computer*, 42(1), 128.
- [3] Sun, F., Cheng, D., Zheng, P., Song, T., Chen, L., & Kou, Q. (2026). TGCADNet: Text-Guided Context-Aware Detection via CLIP for Small Objects in UAV Scenes. *IEEE Transactions on Circuits and Systems for Video Technology*.
- [4] Luo, Y. (2026). A novel context-aware modal-interaction graph attention and adaptive attention approach for multi-modal emotion recognition in public opinion: Yuting Luo. *Nonlinear Dynamics*, 114(3), 149.
- [5] Nguyen, H., Tran, N., Nguyen, M., Ta, P., & D Nguyen, H. (2026). Attention mechanisms for context-aware emotion recognition. *CCF Transactions on Pervasive Computing and Interaction*, 1-17.
- [6] Billardello, G. C. E. S. R., Tamantini, C., Loredana, S. L., & Cordella, Z. L. P. C. F. Context-Aware Semiautonomous Control for Upper-Limb Prostheses.
- [7] Kumari, S., & Kumar, V. (2026). A hybrid context-aware video violence detection framework using hierarchical spatiotemporal and semantic modeling: S. Kumari, V. Kumar. *Signal, Image and Video Processing*, 20(1), 12.
- [8] Sun, J., Li, Y., Huang, H., Yin, Y., Tang, T., Zhao, H., & Gui, G. (2026). Context-Aware RandLA-Net: An Enhanced Architecture for Large-Scale Point Cloud Semantic Segmentation. *IEEE Internet of Things Journal*.
- [9] Chafiki, M. E. A., Stitini, O., & Kaloun, S. (2026). A systematic review of advancements in context-aware recommendation systems: MEA Chafiki et al. *Knowledge and Information Systems*, 68(1), 11.
- [10] Casillo, M., Colace, F., Conte, D., Lombardi, M., Santaniello, D., & Valentino, C. (2026). A context-aware recommender system-based framework for improving cultural experiences: M. Casillo et al. *User Modeling and User-Adapted Interaction*, 36(1), 1.
- [11] Mehra, P. S., Gupta, S., & Kumar, A. (2026). Ambient intelligence in industrial IoT: A survey of context-aware, adaptive, and autonomous smart manufacturing systems. In *Artificial Intelligence and Sustainable Innovation* (pp. 756-762). CRC Press.

- [12] Avetisyan, L., Yang, X. J., & Zhou, F. (2026). Towards context-aware modeling of situation awareness in conditionally automated driving. *International Journal of Human-Computer Interaction*, 42(1), 291-306.
- [13] Xie, Y., Jia, Y., Niu, R., Cao, P., Liu, Y., & Mu, H. (2026). ECCFNet: enhanced context-aware cross-resolution fusion network for human pose estimation. *Journal of Real-Time Image Processing*, 23(1), 55.
- [14] Wang, J., & Wang, M. (2026). Progressive edge-aware multi-scale fusion network for camouflaged object detection. *Multimedia Systems*, 32(2), 105.
- [15] Raja, P., BK, D., Ravi, V., & Alghamdi, N. S. (2026). A context-aware multi-modal generative adversarial network for real-time anomaly detection in video surveillance. *Peer-to-Peer Networking and Applications*, 19(1), 29.
- [16] Hong, Y., Lin, K., Xu, J., Chang, B., & Du, D. (2026). Expert knowledge-guided deep neural network based on context-aware hierarchy for foils joining quality monitoring. *Advanced Engineering Informatics*, 69, 104001.
- [17] Geremias, J., Viegas, E. K., Santin, A. O., & de S. Britto Jr, A. (2026). Context-aware model for lightweight real-time detection of obscene videos. *Journal of Ambient Intelligence and Humanized Computing*, 1-15.
- [18] Feng, L., Zhang, L., Ma, S., Wang, D., & Song, A. (2026). DHT_xLSTM: A hybrid temporal-mapping framework for skill-oriented grasp prediction in teleoperated robots via dual-quaternion control and context-aware learning. *Biomimetic Intelligence and Robotics*, 100296.
- [19] Li, T., & Li, Z. C. (2026). PYOLO-PCF: A Lightweight and Context-Aware Multi-Scale Framework for Printed Packaging Defect Detection. *International Journal of Pattern Recognition and Artificial Intelligence*, 40(02), 2559022.
- [20] Dasu, B. T., Reddy, M. V., Kumar, K. V., Chithaluru, P., Ahmed, N., & Abd Elminaam, D. S. (2026). A self-attention driven multi-scale object detection framework for adverse weather in smart cities. *Scientific Reports*.
- [21] Pari, S. N., & Sudharson, K. (2026). Context-aware Knowledge Base Engineering for Anomaly Detection and Predictive Maintenance in Graph Databases. In *Cognitive Fairness-Aware Techniques for Human-Machine Interface* (pp. 139-166). Chapman and Hall/CRC.
- [22] Liu, J., Ni, J., Zhang, Z., Gu, Y., & Yang, S. X. (2026). MAFTNet: Multimodal Adaptive Fusion-based Transformer Network for Infrared and Visible Image UAV Object Detection. *IEEE Sensors Journal*.
- [23] Abadie, C., Purohit, A., Bavishi, P., Guidroz, T., Jain, A., & Mezerreg, I. (2026). Just-in-Time Adaptive Interventions for Wellness Management via Context-Aware Wearable Devices.
- [24] Ghorbani, F., Hym, A., Elhenawy, M., & Rakotonirainy, A. (2026). Towards AI-Assisted Motorcycle Safety: Multi-Modal Video Analysis for Hazard Detection and Contextual Risk Assessment. *Vehicles*, 8(2), 39.