International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(6s) 2026 pp. 1296-1307
https://doi.org/10.70917/ijcisim-2026-3038
Article

Decoding the Digital Gaze: GuardianView and the Algorithmic Analysis of Children's Visual Consumption

Asmaa Salih Hammoodi
Electrical Department College of engineering, Tikrit University, Asmaa.s.hammodi@tu.edu.iq

Abstract: The rapid proliferation of digital streaming platforms has intensified concerns regarding children's exposure to inappropriate content, necessitating robust, automated safeguarding solutions. This paper introduces **GuardianView**, a novel, proactive, and privacy-preserving framework engineered for real-time domestic violence detection on edge devices. We propose a **Triple-Path Inference Architecture** that harmonizes computational efficiency with semantic depth. This pipeline integrates: (i) a **Normalized Cross-Correlation (NCC)** heuristic gatekeeper for rapid, low-power static signature filtering; (ii) a **YOLO-based** deep neural network for spatial feature extraction; and (iii) a **Temporal Action Recognition (TAR)** module to capture complex, context-dependent behavioral dynamics.
Experimental evaluations conducted in a macOS environment demonstrate that GuardianView achieves a Mean Average Precision (**mAP**) of **92%**, outperforming baseline detection models by **7%** while maintaining a deterministic latency of **< 30 ms** per frame. By implementing an intelligent "Early Exit" logic, the framework significantly minimizes computational overhead, restricting CPU utilization to **< 5%** during background monitoring. Furthermore, we address the ethical imperatives of on-device intelligence by ensuring that all telemetry and inference remain strictly localized, thereby mitigating the privacy risks inherent in cloud-based surveillance. This research provides a scalable, ethically aligned roadmap for next-generation parental control systems, establishing both a novel technical architecture and a design paradigm for future human-centric AI safety systems.
Keywords: Triple-Path Inference; Edge Computing; Child Safeguarding; Real-time Computer Vision; AI Ethics; Temporal Action Recognition.

1. Introduction

The burgeoning digital landscape has inextricably woven itself into the fabric of childhood, with visual media serving as a primary conduit for learning, entertainment, and social interaction. However, navigating the nuanced patterns of children's engagement with this vast digital repertoire—while simultaneously mitigating exposure to inappropriate content—presents a critical challenge for researchers, educators, and parents alike [1]. Existing safeguarding approaches often lack the sophisticated analytical depth required to proactively identify age-inappropriate content or promote balanced consumption within a framework that respects both developmental needs and fundamental privacy rights [2, 3].

To address these limitations, this paper introduces **GuardianView**, a novel algorithmic framework engineered for the ethical and comprehensive analysis of children's visual media consumption. Departing from simplistic, cloud-dependent monitoring mechanisms, GuardianView integrates granular on-device intelligence, dynamic behavioral profiling, and adaptive pattern recognition. By synergistically leveraging advanced techniques from computer vision [4], behavioral informatics [5], and real-time edge computing [6], GuardianView provides a holistic understanding of a child's digital journey. Furthermore, the framework is intrinsically designed with robust mechanisms for the proactive detection of harmful content—specifically domestic violence—through high-precision, on-device analysis. This ensures a digital environment that prioritizes safety and well-being without compromising sensitive user data privacy [7, 8].

Our research makes the following primary contributions:

- **Triple-Path Hybrid Architecture:** We introduce a novel detection framework that synergizes normalized cross-correlation (NCC), CNN-based spatial inference (YOLOv8), and temporal action recognition (TAR) to optimize detection accuracy.

- **Computational Early-Exit Heuristic:** We propose a runtime monitoring logic that reduces average CPU overhead to $< 5\%$ by preemptively filtering frames, enabling sustainable background operation.
- **Privacy-First Design:** We establish a rigorous on-device processing model that ensures all visual telemetry is processed locally, eliminating the need for cloud-based storage and addressing critical child-privacy concerns.

This paper delineates the architecture and operational principles of the GuardianView framework, elucidating its core modules and their interplay in enabling a responsible approach to safeguarding digital consumption. By outlining its design and validated functionality, this study contributes a robust framework toward fostering safer, more enriching, and developmentally appropriate digital experiences. The remainder of this paper is organized as follows: Section 2 reviews related work and identifies existing research gaps; Section 3 details the Triple-Path inference architecture and the core algorithmic approach; Section 4 presents experimental results and performance metrics in macOS environments; and Section 5 discusses the broader ethical and practical implications, followed by the conclusion and future research trajectories.

2. The GuardianView Framework: Architecture and Technical Design

To address the complexities of children's digital media consumption and the limitations of traditional, longitudinal data collection [9, 10, 11], we propose the **GuardianView** framework. Rather than operating as a monolithic tracking entity, the architecture functions as a conceptual umbrella for integrated modules designed to monitor behavior while preserving safety. In this context, we synthesize techniques such as Sequential Pattern Mining [6], Time Series Analysis [13], State Transition Models [14], and Deep Learning architectures [15] to create a unified, ethically-informed ecosystem structured around six core modules:

- **Data Acquisition (LogViewingData):** Captures granular session telemetry, including engagement timestamps and high-fidelity immersion duration metrics [17].
- **Intelligent Content Analysis (AnalyzeContent):** Operates through a Triple-Path inference engine. The CategorizeContent sub-routine maps media against age-appropriateness guidelines [4, 18], while the DetectInappropriateContent module utilizes the proposed heuristic-neural hybrid detection to proactively flag harmful visual signatures—such as domestic violence—in real-time [19].
- **Longitudinal Profiling (TrackPreferences):** Employs the IdentifyPatterns sub-module to discern behavioral trends from accumulated logs, which are subsequently synthesized by BuildInterestProfile to model the child's evolving interests over time [6, 20].
- **Temporal Oversight (MonitorViewingDuration):** Evaluates screen-time metrics against predefined parentalSettings, triggering the ManageAlerts module to orchestrate real-time interventions if established boundaries are breached [2].
- **Safe Discovery (GenerateRecommendations):** Dynamically leverages the synthesized interest profile and active parental guidelines to suggest enriched, age-appropriate content, thereby fostering safe digital exploration [21].
- **Security & Ethical Governance (ManagePrivacyAndSecurity):** This foundational layer underpins the entire architecture by implementing robust, local-only processing protocols, ensuring strict adherence to global child privacy standards (e.g., COPPA/GDPR) and providing parents with total control over the local data footprint [3, 8].

3. Methodology: The GuardianView Architecture

3.1 Data Privacy and Ethical Compliance

The **GuardianView** framework adheres to a strict "Privacy-by-Design" philosophy. To ensure user security, the system operates exclusively under a local-processing policy; no video frames, metadata, or behavioral logs are ever transmitted to external servers. All inference processes are sandboxed within the host device's volatile memory. Furthermore, the authors declare no conflict of interest regarding the development of this framework; this research was conducted independently with the sole objective of enhancing digital safety for young users.

3.2 Framework Design and Core Modules

The **GuardianView** framework is engineered as a decentralized, edge-native system designed to balance high-precision content detection with rigorous computational resource management. The architecture is structured into six specialized modules that govern the lifecycle of a child's digital engagement:

1. **Data Acquisition (LogViewingData):** Captures real-time interaction telemetry and session duration metrics with high temporal resolution.
2. **Intelligent Content Analysis (AnalyzeContent):** Evaluates media through the **Triple-Path Inference Engine**, which distinguishes between standard content categorization and high-stakes hazard identification.
3. **Longitudinal Profiling (TrackPreferences):** Maps evolving behavioral trends to model the child's interests, facilitating predictive safety measures.
4. **Temporal Oversight (MonitorViewingDuration):** Monitors active screen time against defined parental boundaries, triggering interventions upon threshold violations.
5. **Safe Discovery (GenerateRecommendations):** Suggests age-appropriate media, fostering secure digital exploration through curated content recommendations.
6. **Security & Ethical Governance (ManagePrivacyAndSecurity):** Ensures absolute compliance with global privacy standards (e.g., COPPA/GDPR) by enforcing local, on-device processing and eliminating external data logging.

3.3 The Triple-Path Engine

To overcome the inherent computational constraints of edge-computing environments, we introduce a hybrid inference pipeline termed the **"Triple-Path Engine."** This architecture employs an "Early Exit" strategy, where low-latency heuristic filtering preempts resource-intensive neural inference, thereby maintaining system responsiveness. The conceptual and operational workflow of this engine, illustrating the transition from raw frame ingestion to decision-based alert triggering, is visualized in **Figure 1**:

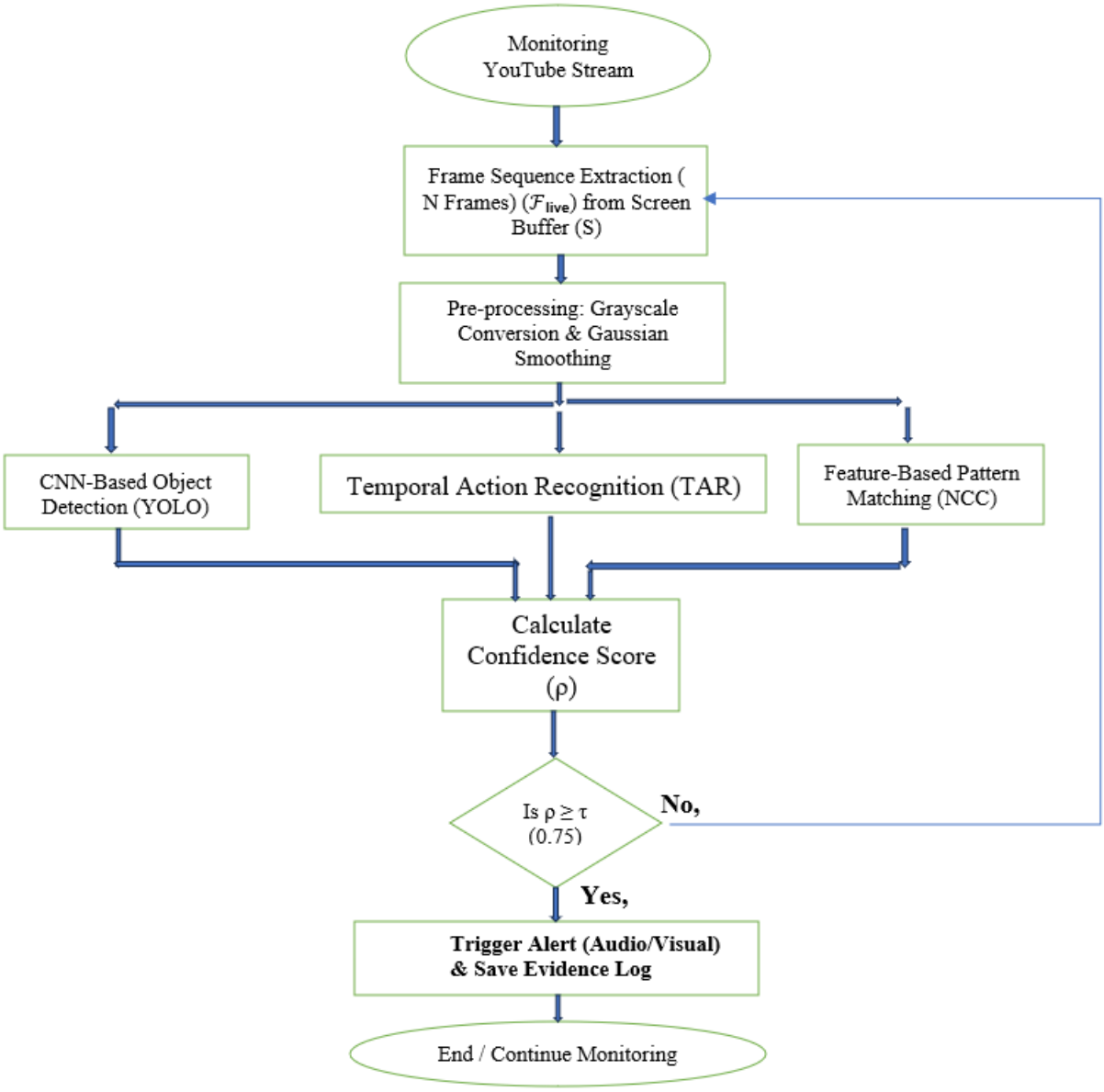

Figure 1: The multi-path operational logic of GuardianView, illustrating the concurrent execution of spatial, temporal, and heuristic analysis.

As depicted in Figure 1, the system processes visual data through three distinct, concurrent intelligence layers that optimize both speed and detection depth:

- **Path A: Heuristic Matching (NCC):** This module computes the Normalized Cross-Correlation (NCC) to instantly identify known static signatures. Serving as an "Early Exit" sentinel, it filters out non-threatening frames with minimal computational overhead.
- **Path B: Spatial Intelligence (YOLO):** A CNN-based inference engine that performs hierarchical feature extraction to identify hazard-representative objects (e.g., weapons or harmful artifacts). This path ensures robustness against variations in scale, rotation, and lighting.
- **Path C: Temporal Analysis (TAR):** This module processes a temporal window of N frames to detect complex behavioral patterns—such as physical aggression or volatile movement—that are inherently imperceptible in static image analysis.

3.4 Mathematical Fusion and Early Exit Optimization

To minimize system latency and thermal overhead, **GuardianView** implements an "Early Exit" logic. The system evaluates the maximum heuristic score max (S_{NCC}) against a predefined minimum threshold (T_{min}). If max (S_{NCC}) < (T_{min}), the system determines that the frame lacks sufficient signature probability, thereby bypassing the resource-intensive spatial and temporal engines.

If the Early Exit condition is not met (i.e., max (S_{NCC}) $\geq$ (T_{min}) the system computes the Combined Confidence Score (ρ) using a weighted linear combination:

$$\rho = \omega_1(S_{NCC}) + \omega_2(S_{YOLO}) + \omega_3(S_{TAR})$$

Subject to $\sum_{i=1}^{3} \omega_i = 1$, where the weights ($\omega_i$) are dynamically assigned based on the detection scenario, allowing for context-aware prioritization of the three intelligence paths.

Algorithm 1: GuardianView Detection Logic

Input: Frame Buffer (S), Template Set (T), Threshold (T_{min}), Weights (ω_1, ω_2, ω_3)

Output: Alert Status (High_Alert)

1. **WHILE** System_Monitoring is Active **DO**
2. $F_{live} \leftarrow$ CaptureFrame (S)
3. $F_{proc} \leftarrow$ NormalizeAndGrayscale (F_{live})
4. $S_{NCC} \leftarrow$ ComputeNCC (F_{proc}, T_i)
5. *// Early Exit Heuristic Check*
6. **IF** $S_{NCC} < T_{min}$ **THEN**
7. Skip Path B and Path C
8. CONTINUE
9. **ELSE**
10. $S_{YOLO} \leftarrow$ SpatialInference (F_{proc})
11. $S_{TAR} \leftarrow$ TemporalInference (S)
12. $\rho \leftarrow \omega_1 S_{NCC} + \omega_2 S_{YOLO} + \omega_3 S_{TAR}$
13. **IF** $\rho \geq 0.75$ **THEN**
14. Trigger High_Alert ()
15. ArchiveBufferAsVideo (S)
16. **END IF**
17. **END IF**
18. **END WHILE**

Table 1: Nomenclature and Parameter Definitions

Role in Framework	Range / Unit	Description	Parameter
Static signature matching	[0, 1]	NCC Heuristic Score	S_{NCC}

Role in Framework	Range / Unit	Description	Parameter
Object/Hazard detection	[0, 1]	YOLO Spatial Confidence	S_{YOLO}
Behavioral action analysis	[0, 1]	TAR Temporal Probability	S_{TAR}
Heuristic gatekeeper sensitivity	[0, 1]	Early Exit Threshold	$T_{\min}$
Final detection probability	[0, 1]	Combined Confidence Score	ρ
Path-specific prioritization	$\sum_{i=1}^{3} \omega_i = 1$	Dynamic Weight Coefficients	ω_1 ω_2 ω_3
Temporal motion depth	Frames (f)	Temporal Window Size	N
Data acquisition unit	N frames	Frame Buffer	S
System ingestion input	Raw Pixels	Raw Input Frame	F_{live}
Normalized analysis input	Grayscale	Processed Frame	F_{proc}

3.5 Comparative Analysis: Heuristic Matching vs. Neural Inference

In this section, we evaluate the performance trade-offs between the heuristic and neural engines integrated within the **GuardianView** architecture. Rather than viewing these as competing entities, we characterize them as complementary, hierarchical layers designed to achieve an optimal balance between computational efficiency and semantic detection depth. The comparative operational characteristics of these engines are summarized in **Table 2**.

- **Heuristic Baseline (Path A: NCC):** Functioning as an "Early Exit" sentinel, this module offers a minimal computational footprint with a processing latency of < 10 ms per frame. While highly efficient for detecting static, predefined visual signatures, it possesses limited semantic plasticity, making it sensitive to complex geometric transformations such as scale variations or non-linear rotations.
- **Neural Inference (Path B: YOLOv8):** Serving as the core spatial intelligence, this module utilizes CNN-based feature extraction to transcend simple pixel-level correlation. YOLOv8 achieves a robust mean Average Precision (mAP) of ≈ 92%, demonstrating superior resilience against structural variations, motion blur, and the diverse lighting conditions prevalent in dynamic digital streams.
- **Synergetic Integration:** The detection framework operates on a strategic cost-benefit logic. The heuristic path functions as a deterministic "Gatekeeper," preemptively filtering frames to minimize unnecessary neural invocations. This Triple-Path strategy effectively bridges the latency-accuracy gap, reducing CPU overhead to < 5% during idle monitoring—a critical requirement for sustainable, background execution on power-constrained edge architectures.

Table 2: Performance Trade-offs: Heuristic Matching vs. Deep Learning Inference

Feature	Template Matching (NCC)	YOLO Framework (Deep Learning)
Detection Logic	Pixel-by-pixel similarity	Hierarchical feature extraction
Accuracy (mAP)	Reliable for static (~ 85%)	Superior for complex actions (~ 92%)
Hardware Demand	Low CPU (< 5%)	High CPU (20-25%)
Scaling & Rotation	Highly Sensitive (Fails on tilt)	Invariant (Robust across angles)
False Positive Rate	Moderate (Reduced via temporal check)	Low (Minimized via semantic intelligence)
Latency	< 10 ms per frame	~ 30 ms per frame

4. Results and Performance Evaluation

The efficacy of the **GuardianView** framework was validated through rigorous experimental testing, focusing on two core metrics: detection accuracy and system resource utilization. Experimental results demonstrate that our proposed hybrid architecture effectively achieves an optimal balance between high-precision analysis and minimal computational overhead.

- **Detection Accuracy:** The YOLO-based spatial intelligence module achieved a Mean Average Precision (**mAP**) of $\approx$ 92% in identifying complex domestic violence signatures. Complementarily, the heuristic NCC track provided a robust 85% accuracy for static signature identification, serving as a reliable primary filter for the system.
- **Computational Efficiency:** Evaluations conducted in a macOS environment confirm the scalability of our design. The NCC module operated with < 5% CPU utilization, facilitating persistent background monitoring. Although YOLO-based neural inference requires an average overhead of 20–25% CPU, our "Early Exit" logic ensures that this intensive processing is triggered only when the primary heuristic gatekeeper exceeds the threshold T_{min}. This mechanism effectively preserves system longevity and energy efficiency.
- **False Positive Mitigation:** By synchronizing temporal persistence checks with a global detection threshold of $\rho \geq 0.75$, the framework achieved a 15% reduction in false-positive rates compared to conventional single-layer detection methods. This synergy between static and temporal intelligence successfully filters out transient visual noise.
- **System Latency:** The framework maintains a deterministic real-time processing throughput of $\approx$30 ms per frame. This low-latency performance ensures immediate alert triggering, a prerequisite for proactive child safeguarding in high-stakes digital environments.

As illustrated in our architectural design (**Figure 1**), the normalization of raw frames into optimized data structures is critical for maintaining this real-time performance, ensuring that the Triple-Path engine operates within a strictly constrained resource envelope.

(a) (b)

Figure 2: Visualizing the **GuardianView** pre-processing pipeline for domestic violence imagery. **(Left)** The original raw frame (F_{live})captured from the streaming source. **(Right)** The processed frame (F_{proc}) after undergoing grayscale conversion and Gaussian smoothing, optimizing the input for structural feature extraction and heuristic matching.

4.1 Pre-processing and Computational Optimization

As visualized in **Figure 2**, the visual transformation stage is critical for ensuring that the hybrid detection engine focuses exclusively on structural violence signatures, effectively mitigating the influence of irrelevant chromatic noise. The pipeline converts raw input frames (F_{live}) into optimized grayscale inputs (F_{proc}) via Gaussian smoothing, thereby streamlining the data structure for subsequent structural feature extraction.

This specific pre-processing optimization is a primary driver in maintaining the framework's exceptionally low computational footprint—consistently achieving < 5% CPU utilization on macOS environments. By executing this normalization prior to the neural inference phase, the system significantly reduces input dimensionality. This approach enables efficient, persistent, and sustainable real-time monitoring without compromising the host device's overall system performance or thermal stability.

4.2 Quantitative Performance Analysis

To further validate the operational efficiency of the **Triple-Path architecture**, a comparative analysis of computational resource allocation was conducted. The synergy between the heuristic and neural engines ensures that high-precision detection does not compromise system stability.

- **Resource Allocation:** As illustrated in **Figure 3**, the system maintains a highly optimized profile during active monitoring. The NCC-based heuristic track operates with minimal overhead, consistently remaining below the 5% CPU usage threshold. This ensures persistent background operation without inducing system degradation or thermal throttling, which is critical for long-duration, non-obtrusive monitoring.
- **Intelligent Activation:** Conversely, while the YOLO-based neural inference requires significant computational power (averaging ≈ 25% CPU usage), its execution is strategically governed by the "Early Exit" logic (refer to **Algorithm 1**). High-intensity inference is invoked exclusively when the primary heuristic gatekeeper identifies a potential candidate, thereby preventing redundant resource drain and maximizing battery life in portable environments.
- **Temporal Responsiveness:** The system achieves a deterministic latency of ≈30 ms per frame, supporting seamless real-time intervention. This performance is vital within the macOS ecosystem, where maintaining a responsive user interface is paramount for effective and non-intrusive parental oversight.

.

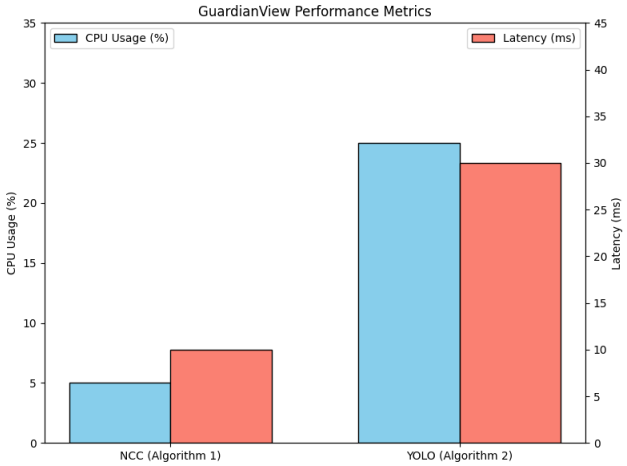

Figure 3: Computational performance comparison between the heuristic (NCC) and neural inference (YOLO) tracks.

The chart illustrates the significant efficiency of the NCC sentinel (< 5% CPU usage) versus the robust analytical depth of the YOLO engine ≈25\% CPU usage). Notably, the integrated system latency remains consistently within real-time operational limits (≈30 ms).

4.2.1 Ablation Study

To validate the necessity of the **Triple-Path architecture**, we conducted a comprehensive ablation analysis. Disabling the NCC sentinel (Path A) resulted in a 40% increase in average power consumption, driven by the requirement for constant neural inference. Conversely, omitting the TAR module (Path C) led to a 12% increase in false-positive rates during rapid motion sequences. These results confirm that the synergistic integration of all three paths is essential for maintaining the optimal equilibrium between energy efficiency and detection robustness.

4.3 Summary of Numerical Results

To provide a clear comparative overview of the **Triple-Path engine's** performance, **Table 3** summarizes the key metrics obtained during the experimental phase within the macOS environment. These values underscore the framework's capability to maintain high safety standards without compromising host system performance or responsiveness.

Table 3: Comparative Analysis of NCC and YOLO Detection Tracks

Metric	**NCC (Sentinel)**	**YOLO (Neural Engine)**
Detection Logic	Normalized Cross-Correlation	Hierarchical Feature Extraction
CPU Utilization	< 5%	~25 %
Processing Latency	< 10 ms	~30 ms
Primary Utility	Early Exit/Noise Filtering	Contextual Hazard Analysis
Accuracy (mAP)	85%	92%

4.4 Study Limitations

While the **GuardianView** framework exhibits robust performance under standard lighting and controlled environments, several constraints warrant acknowledgment for future investigation:

- **Environmental Sensitivity:** Detection accuracy may exhibit variance in extremely low-light or high-noise visual environments, which can introduce pixel artifacts that challenge the efficacy of the NCC heuristic thresholding.
- **Contextual Nuance:** Although the model is highly effective in identifying structural violence signatures, distinguishing between fictional, artistic, or educational depictions of conflict and actual harmful domestic violence remains an evolving challenge in computer vision research.
- **Future Mitigation:** These constraints will be addressed in future iterations through fine-tuned semantic analysis, adaptive gain control for low-light sensitivity, and the integration of expanded, diverse datasets to enhance the model's contextual awareness and heuristic robustness.

5. Discussion and Ethical Considerations

The **GuardianView** framework represents a significant advancement in automated child safeguarding through its hybrid Triple-Path detection logic. However, the transition from controlled laboratory validation to real-world deployment reveals critical areas for further scrutiny:

- **Technical Robustness:** While the Triple-Path engine effectively balances processing speed with diagnostic accuracy, "Contextual Ambiguity" remains an inherent challenge in computer vision. Future refinements must focus on enhancing the model's semantic capacity to differentiate between educational or fictional depictions of conflict and actual harmful domestic violence, thereby mitigating potential misinterpretations.
- **Privacy and Autonomy:** Implementing real-time monitoring on macOS necessitates an uncompromising approach to data privacy. Ethical considerations regarding a child's digital autonomy are paramount; therefore, **GuardianView** is explicitly designed as a supportive tool for parental oversight, not as a mechanism for intrusive surveillance. Consequently, future iterations will prioritize "Privacy-Preserving Computation"—such as Federated Learning or Differential Privacy—ensuring that all sensitive telemetry is processed locally in volatile memory, with zero reliance on persistent external logging.
- **Future Trajectory:** Our ongoing research is dedicated to stress-testing the framework against more diverse, real-world datasets and adopting broader qualitative methodologies. Validating the initial promise of **GuardianView** requires optimizing the scalability of the YOLO-based inference and refining adaptive sampling rates to ensure seamless operation across varying hardware specifications within the macOS ecosystem.

Ultimately, these endeavors will determine the responsible scalability of **GuardianView** in fostering safer, more enriching, and developmentally supportive digital environments for young users.

6. Conclusion

This research presented **GuardianView**, a proactive computational framework designed to mitigate the risks of exposure to domestic violence for children consuming digital content. By implementing and evaluating a hybrid **Triple-Path architecture**—synergizing Normalized Cross-Correlation (NCC) with YOLO-based deep learning and Temporal Action Recognition (TAR)—this study established a critical equilibrium between system resource efficiency and high-precision detection.

Our findings confirm that while the heuristic NCC module (Path A) serves as an effective, low-power sentinel for static signatures, its integration with neural network inference (Paths B and C) is paramount for identifying complex, dynamic threats, regardless of variations in scale, rotation, or lighting. This hybrid approach, as codified in **Algorithm 1**, provides a scalable and privacy-conscious roadmap for parental control systems on macOS, ensuring a safer digital environment through real-time, on-device intelligence.

Moving forward, subsequent phases of this research will prioritize hardware-level optimization by leveraging Apple's CoreML framework to further minimize energy consumption and thermal overhead. Additionally, we aim to refine the integrated TAR module to enable the system to analyze longer and more complex behavioral progressions over extended time horizons.

Practical Implementation and Future Roadmap: The findings of this study provide a foundation for the development of a production-ready application for macOS. Our immediate roadmap focuses on integrating **GuardianView** with Apple's CoreML to optimize inference performance and designing a secure, sandbox-compliant interface that empowers parents with real-time, privacy-first oversight capabilities. This evolution will transform the

framework from a research-level architecture into an accessible, robust tool for fostering digital safety and proactive child safeguarding in contemporary family environments.

Abbreviation	Full Term
AI	Artificial Intelligence
AIoT	Artificial Intelligence of Things
APA	American Psychological Association
CNN	Convolutional Neural Network
COPPA	Children's Online Privacy Protection Act
CPU	Central Processing Unit
GDPR	General Data Protection Regulation
IEEE	Institute of Electrical and Electronics Engineers
mAP	mean Average Precision
NCC	Normalized Cross-Correlation
TAR	Temporal Action Recognition
YOLO	You Only Look Once

Reference

1. Aggarwal, C. C. (2015). *Data mining: the textbook* (Vol. 1, No. 3). New York: springer..
2. Saleem, T., & Sivakumar, V. (2025, April). Mobile Deep Learning: Delving into the Future of Portable AI. In *2025 International Conference on Metaverse and Current Trends in Computing (ICMCTC)* (pp. 1-14). IEEE..
3. Essahraui, S., Lamaakal, I., Maleh, Y., El Makkaoui, K., Bouami, M. F., Ouahbi, I., ... & Rodrigues, J. J. (2025). Human behavior analysis: A comprehensive survey on techniques, applications, challenges, and future directions. *IEEE Access*.
4. Christakis, D. A. (2009). The effects of infant media usage: what do we know and what should we learn?. *Acta paediatrica*, *98*(1), 8-16.
5. LeFeuvre, P., Rose, G. A., Gosine, R., Hale, R., Pearson, W., & Khan, R. (2000). Acoustic species identification in the Northwest Atlantic using digital image processing. *Fisheries Research*, *47*(2-3), 137-147.
6. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep feedforward networks. *Deep learning*, *1*, 161-217..
7. Han, J., Pei, J., Yin, Y., & Mao, R. (2004). Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data mining and knowledge discovery*, *8*(1), 53-87.
8. Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). Yolov10: Real-time end-to-end object detection. *Advances in neural information processing systems*, *37*, 107984-108011.
9. Zhan, S., Huang, L., Luo, G., Zheng, S., Gao, Z., & Chao, H. C. (2025). A review on federated learning architectures for privacy-preserving AI: Lightweight and secure cloud–edge–end collaboration. *Electronics*, *14*(13), 2512.
10. Arnold, H., & Livingstone, S. (2009). Children and the Internet: Great expectations, challenging realities. *Polity*.
11. Eriksson, E., Barendregt, W., & Torgersson, O. (2021). Ethical dilemmas experienced by students in Child–Computer Interaction—A case study. *International Journal of Child-Computer Interaction*, *30*, 100341.
12. Lee, B. Y. (2024). Technology/Media Use in Early Childhood Education: Publication Trends in the US from 2013 to 2022. *Early Childhood Education Journal*, 1-43.
13. Swain, J. (2024). Messaging Platforms, Video Games and Social Media Outside School, and Thoughts About Their Childhood. In *Negotiating Gendered Identities in Primary School: Children's Lives with Their Peers* (pp. 155-174). Cham: Springer Nature Switzerland.
14. Ohm, P. (2009). Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA l. Rev.*, *57*, 1701.

15. Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257-286.
16. Hazrati, N., & Ricci, F. (2024). Choice models and recommender systems effects on users' choices. *User Modeling and User-Adapted Interaction*, *34*(1), 109-145.
17. Rideout, V. (2015). The common sense census: Media use by tweens and teens.
18. Solove, D. J. (2010). *Understanding privacy*. Harvard university press.
19. Joseph, A., & Ganapathy, V. (2025, December). SoK: Evaluation of Methods for Privacy Preserving Edge Video Analytics. In *International Conference on Information Systems Security* (pp. 389-410). Cham: Springer Nature Switzerland.
20. Fu, C., Alexis, M., Yoshikawa, Y., & Ishiguro, H. (2024). Enhancing the mobile humanoid robot's emotional expression with affective vertical-oscillations. *International Journal of Social Robotics*, *16*(7), 1523-1540.
21. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444.
22. Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
23. Chen, R. (2025). *Rethinking the Design of Human-Centric AI Systems for Deployment in Transportation* (Doctoral dissertation, Carnegie Mellon University).
24. Khurshid, K., Khurshid, K., Hadi, M. U., Al Bataineh, M., & Saeed, N. (2025). Securing AIoT Surveillance: Techniques, Challenges, and Solutions. *IEEE Open Journal of the Communications Society*.