Visual Computing for Industry, Biomedicine, and Art

Predictive Safety Intelligence: NANI — A DeepLab-Based Framework for Infant and Baby Posture and Fall Risk Prevention

--Manuscript Draft--

Manuscript Number:	VICO-D-25-00446
Full Title:	Predictive Safety Intelligence: NANI — A DeepLab-Based Framework for Infant and Baby Posture and Fall Risk Prevention
Article Type:	Original Article
Funding Information:	
Abstract:	Infant and toddler safety remains one of the most critical challenges in smart home monitoring. Most existing vision-based systems focus on post-incident fall detection, providing alerts only after accidents occur. In this paper, we introduce NANI (Neural Assistant for Nurturing Infants), a predictive safety intelligence framework that prevents both falls and posture-related risks in infants and young children up to five years old. The proposed system integrates a DeepLab (ResNet-101) segmentation model, fine-tuned on a custom bed dataset (IoU = 88%), with Google ML Kit (MediaPipe) for pose estimation. This combination enables real-time extraction of safe zone boundaries and continuous monitoring of key infant body points such as the head, hips, and limbs. By analyzing the spatial proximity between body keypoints and the edges of the segmented safe zone, NANI predicts potential falls before they happen. Additionally, it evaluates posture angles to identify unhealthy positions, preventing the development of bad habits such as prolonged face-down or "W-sitting." The system was deployed and tested in real-world environments using smartphone cameras, demonstrating reliable inference on edge devices without recording any video, ensuring full privacy compliance.
Corresponding Author:	Kareem Yasser Khalf, BSc Egyptian E-Learning University SOHAG, NORTH AFRICA EGYPT
Corresponding Author E-Mail:	kareemyasser1054@gmail.com
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Egyptian E-Learning University
Corresponding Author's Secondary Institution:	
First Author:	Kareem Yasser Khalf, BSc
First Author Secondary Information:	
Order of Authors:	Kareem Yasser Khalf, BSc
	Rodaina Abdelsalam, PhD
Order of Authors Secondary Information:	
Suggested Reviewers:	
Opposed Reviewers:	
Additional Information:	
Question	**Response**

Predictive Safety Intelligence: NANI — A DeepLab-Based Framework for Infant and Baby Posture and Fall Risk Prevention

November 2025

By:

Kareem Yasser, BSc

Rodaina Abdelsalam, PhD

Egyptian E-Learning University

Abstract

Infant and toddler safety remains one of the most critical challenges in smart home monitoring. Most existing vision-based systems focus on post-incident fall detection, providing alerts only after accidents occur. In this paper, we introduce NANI (Neural Assistant for Nurturing Infants), a predictive safety intelligence framework that prevents both falls and posture-related risks in infants and young children up to five years old. The proposed system integrates a DeepLab (ResNet-101) segmentation model, fine-tuned on a custom bed dataset (IoU = 88%), with Google ML Kit (MediaPipe) for pose estimation. This combination enables real-time extraction of safe zone boundaries and continuous monitoring of key infant body points such as the head, hips, and limbs. By analyzing the spatial proximity between body keypoints and the edges of the segmented safe zone, NANI predicts potential falls before they happen. Additionally, it evaluates posture angles to identify unhealthy positions, preventing the development of bad habits such as prolonged face-down or "W-sitting." The system was deployed and tested in real-world environments using smartphone cameras, demonstrating reliable inference on edge devices without recording any video, ensuring full privacy compliance.

Keywords: Infant safety, Posture analysis, Fall prevention, DeepLab, MediaPipe, Edge AI, Predictive intelligence

---

1. Introduction

Unintentional falls and unsafe postures are among the most frequent causes of infant injuries during early development. Traditional fall detection systems rely on post-event recognition, alerting caregivers only after a child has fallen. While such approaches are useful, they lack predictive intelligence capable of preventing incidents before they occur.

Recent advances in computer vision and deep learning have enabled real-time scene understanding, human pose estimation, and environmental segmentation. However, applying these technologies to infants presents unique challenges due to smaller body size, variable poses, and loose clothing that obscure key body joints. Furthermore, ethical considerations prohibit continuous video recording or intrusive wearable sensors, making vision-based edge intelligence a promising alternative.

In this context, we propose NANI, a preventive safety system that merges semantic segmentation and human pose estimation to provide proactive alerts in infant environments. Unlike previous fall detection methods, NANI identifies both environmental risks (proximity to unsafe edges) and postural risks (harmful body configurations). The framework is designed for deployment on standard mobile devices, leveraging on-device models to preserve privacy and ensure low latency. Inspired by hybrid approaches in computer vision and ergonomic assessment, NANI prioritizes non-intrusive CV for infant privacy.

2. Related Work

Vision-based infant and toddler monitoring systems have evolved from general fall detection in adults and the elderly to proactive, child-specific safety applications. Recent research increasingly leverages deep learning, semantic segmentation, and pose estimation to enable early risk detection in domestic environments, particularly in cribs and play areas. Real-time fall detection for toddlers has been a primary focus. Yang et al. [1] enhanced the YOLOv8 architecture with GELAN backbone and Generalized Hough Transform to accurately detect toddler falling events and compute safe zones, achieving 96.33% accuracy on a custom dataset. Hazard detection inside cribs has also advanced significantly. Zhu et al. [2] introduced CribNet, a vision-based framework that uses Detectron2 for segmenting blankets and toys, combined with a fine-tuned infant pose estimation model (FiDIP), to identify dangerous situations such as facial occlusion and prone sleeping, reaching 80% mAP on the CribHD dataset. Accurate infant body part segmentation is crucial for such systems.

Voss et al. [3]. proposed a multi-modal deep learning approach using RGB and depth data to segment infant body parts even under partial occlusion and varying illumination. Several works have validated deep-learning-based pose estimation for infants and children. Ienaga et al. [4] verified the clinical applicability of pose estimators such as OpenPose and HRNet for assessing postural control in children. Latreche et al. [5] rigorously evaluated the reliability and validity of MediaPipe for human motion tracking in rehabilitation contexts.

Tu et al. [6] integrated MediaPipe pose landmarks with machine learning classifiers to detect motor developmental delays through abnormal posture analysis. Comprehensive monitoring systems have been proposed by multiple research groups. Suresh et al. [7]. presented InfantZA, a deep-learning-enabled surveillance platform that detects hazards around infants with real-time alerts.

Kavitha et al. [8]. developed Baby Vision Pro, combining YOLO-based crying and movement detection with IoT sensors for live streaming and parental notification. Wang et al. [9] proposed a multi-branch CNN architecture for infant face detection, crying recognition, and behavioral monitoring. Zainab et al. [10] introduced a computer-vision-based smart home monitoring system dedicated to early childhood safety protecting them from risks like fire and pools. Khan and Dey [11] utilized deep learning to classify safe versus unsafe objects in children's surroundings. IoT-integrated smart cradles represent another active direction. Venu et al. [12] designed a smart infant care system using ESP32 with multiple sensors for cry detection, bedwetting alerts, and automated rocking. Chandnani et al. [13] presented an advanced smart cradle incorporating Raspberry Pi, CNN, and FFT analysis for high-accuracy cry detection and environmental control.

A systematic review by Khan et al. [14] highlighted persistent challenges in infant sleep and movement monitoring systems (0–4 years), including falls from beds, suffocation risks, privacy concerns, and the need for truly proactive rather than reactive solutions. From the regulatory standpoint, the UK Medicines and Healthcare products Regulatory Agency (MHRA) [15] classifies baby breathing and movement monitors as medical devices when they claim to monitor vital signs or prevent health risks.

Despite these advances, most existing systems remain predominantly reactive and lack predictive fusion of real-time pose dynamics, environmental semantics, and historical patterns to anticipate falls or hazardous postures habits this gap motivates the current work

Approach	**Segmentation**	**Pose Estimation**	**Predictive Risk**	**Edge/ Privacy**	**Regulatory Focus**	**Gap or Addition Addressed by NANI**
Yang et al. [1]	No	Bounding boxes	Post-event	N/A	No	General Proactive fall predictiona nd bad posture habits prevention

Zhu et al. [2]	Detectron2	FiDIP/MediaPipe	Hazard occlusion	Partial	No	Predictive fall/posture
Suresh A et al. [7]	Object detection	No	Reactive	Partial	No	Proactive segmentation and pose estimation
Kavitha et al. [8]	YOLO/CNN	Limited	Cry/baby detection	Partial	No	Edge segmentation and pose detection
Venu et al. [12]	No	Sensors	Cry/environment	Yes	Partial	Vision-based prediction and posture monitoring
Chandnani et al. [13]	No	CNN/FFT	Cry/rocking	Partial	Partial	Full posture / fall monitoring
Khan et al. [14]	Partial (IoT)	Sensors/ML	Sleep risks	Low	Partial	Proactive Predictive home edge and pose alerts

In summary, while existing research advances individual components from ergonomic systems to systematic infant reviews and crib hazards, they often overlook the integration of environmental segmentation with precise pose analysis for predictive infant safety, compounded by regulatory challenges for non-intrusive devices. NANI addresses this by fusing fine-tuned DeepLab (ResNet-101) for safe zone delineation with MediaPipe for keypoint-based proximity and angle evaluation, enabling proactive alerts in privacy-compliant, edge-deployed home environments and filling gaps in proactivity, hardware-free monitoring and bad posture detection exclusively.

---

3. Methodology

The proposed NANI framework is structured as a dual-stage predictive safety architecture that combines semantic segmentation and human pose estimation to proactively identify both environmental and postural risks in infants. The system operates in real time on mobile devices, ensuring privacy, efficiency, and robustness under natural home conditions.

3.1 System Overview

NANI's architecture consists of two major components:

1. **Safe-Zone Segmentation Module** implemented using a fine-tuned DeepLab model with a ResNet-101 backbone, responsible for isolating the infant's physical environment.
2. **Pose Estimation and Risk Analysis Module** powered by Google ML Kit (MediaPipe), responsible for extracting key infant body landmarks and assessing posture and fall risks.

These modules operate sequentially and share information through geometric mappings between segmented safe zones and detected keypoints.

3.2 Safe-Zone Segmentation

In the first stage, the DeepLab model performs semantic segmentation to delineate the safe area-typically the bed or-where the infant is located. The model is fine-tuned on a custom dataset containing labeled images of infant environments, focusing on accurately separating safe surfaces from surrounding hazards such as bed edges or floor boundaries.

optimization is carried out using a cross-entropy loss function with data augmentation for lighting and texture variations. The segmentation model achieves a mean Intersection over Union (mIoU) of 88%, confirming its reliability in varied domestic settings.

The generated segmentation masks serve as the spatial foundation for evaluating the infant's position and movement relative to safe boundaries.

3.3 Pose Estimation and Keypoint Detection

The second stage uses Google's ML Kit with MediaPipe BlazePose for detecting and tracking infant body joints, including the head, shoulders, elbows, hips, knees, and ankles. MediaPipe's lightweight and efficient topology enables real-time inference (~25 FPS) directly on mobile hardware without external processing.

Detected keypoints are projected onto the segmented environment to determine their spatial relationship with the safe zone. The system tracks these keypoints across frames to detect motion patterns indicating of potential falls or unsafe postures.

3.4 Predictive Risk Assessment

The Predictive Risk Analysis module fuses segmentation and poses information to infer safety conditions:

- **Fall Risk Prediction:** The Euclidean distance between the infant's head or hip centroid and the nearest safe-zone boundary is continuously monitored. When this distance drops below a dynamic threshold adaptively scaled to the bed dimensions and previous frame velocity trends the system triggers an early warning for possible fall risk.
- **Postural Risk Detection:** Using joint-angle analysis, the system evaluates key posture configurations to identify unsafe positions such as “face-down sleeping”, prolonged asymmetric lying, or “W-sitting” or “back bending” Critical joint angles (e.g., neck-shoulder, hip-knee, ankle-hip alignment) are compared to medically accepted ranges derived from pediatric and WHO guidelines to ensure physiologically meaningful evaluation.

3.5 Design Scalability and Future Enhancements

The modular design of NANI allows seamless integration of additional sensing modalities and advanced AI features. Future iterations may incorporate:

- Temporal modeling using LSTM or 3D CNNs for motion trend prediction.
- Integration with IOT cameras for easier use .
- EMG or IMU sensor fusion for load estimation and fine-grained motion tracking.
- Advanced segmentation models such as Detectron2 or the Segment Anything Model (SAM) for multi-surface scene understanding.

This modular methodology ensures that NANI remains scalable, privacy-compliant, and medically interpretable, advancing the field of predictive infant safety intelligence.

---

4. Experiments and Results

Hardware: Android smartphone with on-device inference (edge mode).

Models: DeepLab (ResNet-101) fine-tuned on custom bed dataset; MediaPipe for pose detection.

Testing Conditions: Real infants in natural home settings (with parental consent), without recording video or storing images.

Segmentation Accuracy: IoU: 88%

Pose Detection: Real-time inference at 25 FPS (MediaPipe default).

Average Latency: 40 ms per frame.

False Alert Rate: Low; primarily triggered by occluded limbs or extreme lighting.

The following figures represent a sample of the experiments showing some of the poses detected after exceeding allowed time sending alert to prevent these bad habits creation.

Fig (1) Safe area detected using deeplab V3

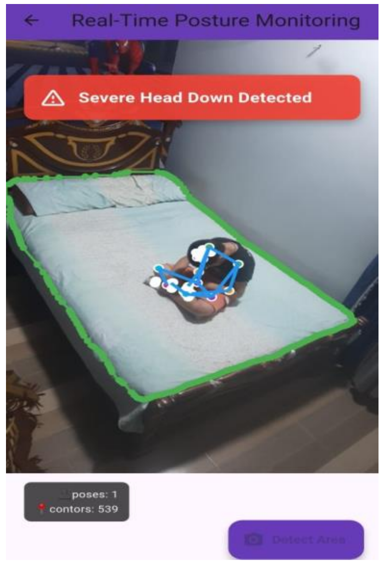
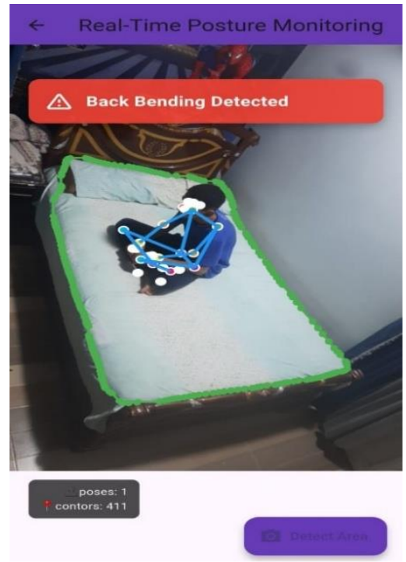


Fig (2) Serve head down alert triggered after passing the allowed time



Fig(3) Back bent alert triggered after passing the allowed time

Fig(4) Head down alert triggered after passing the allowed time

Fig(5) Kid sleeps near the edges of the bed any wrong move puts him in danger

5. Discussion

NANI bridges the gap between reactive fall detection and proactive infant safety intelligence. The fusion of segmentation and pose estimation enables **spatial reasoning** not achievable through bounding-box methods like YOLO.

While the system achieved reliable results, several challenges persist: lighting variation, occlusions, and limited generalization. As highlighted in Khan et al. [14] and Yang et al. [1], current systems overlook 0–6-month coverage and dynamic falls; NANI's flexible pose estimation addresses this across up to 5 years, with predictive edge alerts. Nevertheless, NANI's privacy-first, real-time design and edge implementation make it suitable for deployment in homes and childcare facilities without additional sensors.

Also, the posture detection to correct some bad posture habits is a great addition, As it ensures that kids grow with good body postures and healthy joints.

---

6. Conclusion and Future Work

This paper presented NANI, a predictive safety intelligence framework for infant and baby monitoring. By combining **DeepLab-based safe zone segmentation** with **ML Kit pose estimation**, the system detects potential falls and harmful postures before they occur. The model achieved 88% mIoU and reliable real-time performance on mobile edge devices, proving the feasibility of preventive AI for child safety. Future work includes integration with IOT cameras for more easy use, extending segmentation to additional furniture types, incorporating temporal modeling for motion prediction, and integrating multi-modal sensing (e.g., cry detection and thermal cues) for comprehensive infant wellbeing analysis.

---

7. Ethical Statement

All experiments were conducted with full parental consent, without storing or transmitting any visual data unless by the approval of the parents. The system processes all frames locally on the device, ensuring compliance with child privacy and data protection standards (e.g., GDPR). NANI aligns with MHRA guidelines [15] and reviews like Khan et al. [14] by avoiding vital sign monitoring claims, focusing on preventive posture/fall alerts as a non-intrusive, edge-based tool rather than a diagnostic device.

Competing Interests:

The authors declare no competing interests.

Data Availability Statement:

The original dataset used in this study is publicly available from Roboflow. The dataset was further modified by the authors for bed segmentation annotations. The modified dataset is available on :
[https://app.roboflow.com/kaiten/kaiten55/5]

Authors and Affiliations

Kareem Yasser[1]*, Rodaina Abdelsalam[2]

Egyptian E-Learning University, Egypt [1],[2]

Corresponding Author
Kareem Yasser
Egyptian E-Learning University
Email: kareemyasser1054@gmail.com.

Declaration Sections:

Funding

Not applicable.

Authors' Contributions

KY conceived and designed the study, prepared the data ,performed the experiments and analysis, and prepared the manuscript draft.
RA contributed to study revision, guidance, and manuscript revision.
All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

[1] Yang Z, Tsui B, Ning J, Wu Z. Falling Detection of Toddlers Based on Improved YOLOv8 Models. Sensors. 2024;24(19):6451.

[2] Zhu S, Mathew A, Hatamimajoumerd E, et al. CribNet: Enhancing Infant Safety in Cribs Through Vision-Based Hazard Detection. In: Proc. 18th IEEE Int. Conf. Automatic Face and Gesture Recognition (FG); 2024. p. 1–8.

[3] Voss F, Brechmann N, Lyra S, et al. Multi-modal body part segmentation of infants using deep learning. BioMedical Engineering OnLine. 2023;22:28.

[4] Ienaga N, Takahata S, Terayama K, et al. Development and Verification of Postural Control Assessment Using Deep-Learning-Based Pose Estimators: Towards Clinical Applications.

[5] Latreche A, Kelaiaia R, Chemori A, Kerboua A. Reliability and validity analysis of MediaPipe-based measurement system for some human rehabilitation motions. Measurement. 2023.

[6] Tu Y-F, Lin L-Y, Tsai M-H, Sung Y-S, Liu Y-S, Chen M-Y. Motor delay image recognition based on deep learning and human skeleton model. Applied Soft Computing. 2024.

[7] Suresh A, Sai Jitta A, Anil J, Sai Jayanth Immadi-setty, Bharathi R. InfantZA: Computer Vision and Deep Learning Enabled Infant Surveillance System. Int J Eng Res Technol (IJERT). 2024;13(06).

[8] Kavitha K, Ajay Krishna M, Suresh R, Karthikeyan P, Akshay Kumar. Baby Vision Pro–Smart Baby Monitoring System in Domestic Environment. Int J Multidisciplinary Res. 2024;6(3).

[9] Wang P, Jia J. Deep Learning Based Infant and Child Monitoring System. In: Proc. Int. Conf. Artificial Life and Robotics (ICAROB); 2025.

[10] Zainab Z, Akbarzi J, Qasim M, et al. Smart Home Monitoring System for Early Childhood Using Computer Vision Technology. Int J Innovations Sci Techno. 2025.

[11] Khan FA, Dey A. Towards Enhancing Child Safety: A Deep Learning Approach to Detect Child Safe and Unsafe Objects. In: Proc. 2024 IEEE Int. Women in Engineering Conf. Electrical and Computer Engineering (WIECON-ECE); 2024.

[12] Venu PS, Rajashekar KS, Pasare SB, Gowda BN, Anandaraju MB, ManojKumar SB. Smart Infant Care and Monitoring System Using IoT. Int J Innovative Res Eng. 2025;6(3):52–56.

[13] Chandnani K, Tripathy S, Parbhakar AK, et al. A novel smart baby cradle system utilizing IoT sensors and machine learning for optimized parental care. Scientific Reports.

[14] Khan N, Al-Hamid DZ, Chong PHJ, Mirza F, Kumar A. Systematic Review of Sleep Monitoring Systems for Babies: Research Issues, Current Status, and Future Challenges. Health & Social Care in the Community.

[15] MHRA. Decision Document on Baby Breathing/Movement Monitors. Medicines and Healthcare products Regulatory Agency; February 2021.