Sarcouncil Journal of Medical Sciences

ISSN(Online): 2945-3526

Volume- 05| Issue- 01| 2026

Research Article

Received: 01-12-2025 | Accepted: 25-12-2025 | Published: 10-01-2026

Lightweight Pose-Aware CNN for Real-Time Detection of Health Emergencies on Edge Devices

Kartik Venkataraman

Texas A&M University, College Station, TX

Abstract: Lightweight pose-aware Convolutional Neural Networks (CNNs), integrated with edge computing technologies, have created new opportunities in the field of real-time health emergency detection. Such systems use visual information about human body posture and behavior to detect critical medical events such as falls or immobility and to provide immediate and localized responses without relying on cloud infrastructure. This review examines the state-of-the-art in pose-aware CNNs optimized for deployment on edge devices, emphasizing their design, optimization techniques, deployment challenges, and application use cases. It draws insights from adjacent domains, including education and animal health monitoring, and illustrates how these cross-disciplinary advancements can inform healthcare solutions. The discussion covers an in-depth examination of hardware specifications, environmental limitations, energy efficiency, data privacy, and scalability of the system. By analyzing the current architectures and future opportunities, this review shows the potential of lightweight pose-aware CNNs to transform health monitoring to be both proactive and accessible.

Keywords: Pose-aware CNNs, Edge computing, Health emergency detection, Real-time monitoring.

INTRODUCTION

The introduction of artificial intelligence (AI) and edge computing has sparked considerable improvements in real-time health monitoring systems to the extent that the implementation of highly responsive, intelligent, and lightweight systems that can work on edge devices is possible. A potentially attractive field of application is pose-conscious Convolutional Neural Network (CNN) for monitoring health crises in the form of falls, loss of consciousness, or unusual physiological activities in real-time. The dire need of such systems is supported by the growing number of chronic diseases and the aging population of the world, which also needs constant supervision. Furthermore, traditional centralized healthcare systems often fail to deliver prompt responses due to network delays and privacy concerns, making edge-based solutions not only more effective but increasingly necessary.

Pose estimation has been an asset to the field in various ways, such as smart education, intelligent surveillance, and accurate animal monitoring. One area that is gaining more and more attention is its usage in healthcare, especially by lightweight and efficient CNNs running on edge devices. These models can be used to detect in real-time without using cloud infrastructure and thereby maintain user privacy and reduce delays in transmitting data. This review is a critical analysis of the state of the art, architectures, methodologies, and issues surrounding the implementation of pose-aware CNNs to detect health emergencies in real-time on edge devices.

BACKGROUND AND MOTIVATION

With the increase in the number of smart devices that are fitted with cameras and sensors, human pose estimation has become a practical area in which real-time applications can be performed. Pose-wise CNNs identify and track important body positions and use this information to make inferences about human movement and posture and, therefore, are useful in detecting falls or suspicious actions that may signal health crisis scenarios. Pose estimation is already being harnessed in related fields such as education, where facial features and body posture are used to assess student attentiveness [Singh, R. *et al.*, 2025]. These advances underscore the greater generalizability of visual cues to real-time inference and a conceptual basis for applying such techniques to healthcare.

One of the primary reasons to use lightweight pose-aware CNNs on edge devices is the necessity to provide a rapid reaction and efficiency. Latency is one critical aspect, unlike in traditional cloud-based solutions where data transfer introduces delay, which is minimized in edge computing due to local data processing. It is particularly important when it comes to emergency health situations in which each second counts. Moreover, pose-aware CNNs are efficient to optimize and, hence, can be deployed to resource-constrained devices such as smartphones, wearables, or microcontrollers.

Visual data interpretation to make real-time decisions has also been studied in education

***Corresponding Author:** Kartik Venkataraman

DOI- https://doi.org/10.5281/zenodo.18204445

technology, with deep learning models used to analyze the engagement of students through facial expressions and posture [Aly, M. 2025]. This idea may be applied to healthcare, where it is possible to fine-tune similar designs and identify the anomalies in human behavior. Such lightweight systems can be deployed in various environments, such as homes, assisted living facilities, and hospitals, providing 24/7 non-intrusive surveillance.

Pose-Aware CNN Architectures for Edge Devices

Pose-aware CNNs would detect important locations on the human body, including joints, limbs, and facial features. The architecture of these models must be lightweight enough to run on edge devices while being accurate enough to distinguish between normal and emergency behaviors. The trade-off between compute efficiency and complexity of the model is also another important factor in the context of real-time applications.

Deep learning models combined with social network analysis have shown that non-invasive behavioral monitoring can be used in animal health monitoring [Parivendan, S. C. *et al*., 2025]. The methods can be scaled to human applications, and CNNs are trained on annotated datasets to learn human movement patterns. As an example, such behaviors as falling, shakiness, or inability to move may be considered the signs of health emergencies.

The potential of the CNN-based systems implemented on edge devices has been confirmed in real-time license plate recognition [Sonnara, F. *et al*., 2025]. Such realizations show that the computationally intensive models can be optimized to work effectively on low-power devices. Model pruning, quantization, and knowledge distillation are some of the techniques used to decrease the model size and inference time without a serious accuracy drop.

The features CNNs that are pose-aware usually use the two-step scheme: feature extraction and pose estimation. In the former, convolutional layers are used to identify the visual features, and in the latter, the regression method identifies the important body points. Real-time detection can be made highly reliable by training the model to identify poses that can be associated with health risks, e.g., when a person is lying motionlessly or showing signs of distress. Moreover, edge-based methods of personalization, including few-shot learning and online adaptation, have also increased the accuracy of the model by adjusting to the specific behavioral patterns of individuals [Davalos, E. *et al*., 2025].

Another level of detection accuracy is provided by the integration of facial micro-expression analysis into pose-aware CNNs. Facial expression and pose estimation have been applied in video conferencing to identify stress and fatigue in real-time, which proves the topicality of the given technology in the non-verbal analysis of behavior [Ranjan, A., & Ravinder, M. 2025]. The abilities are particularly relevant to a healthcare environment where visual clues might be the only indicators of a health crisis.

Edge Computing for Real-Time Health Monitoring

Edge computing is a term that is used to refer to the process of processing information close to the origin of the data, instead of using centralized cloud servers. Edge computing enables real-time analysis and response in health emergency detection, which is vital for time-sensitive emergencies. Architectures of edge computing are especially beneficial in minimizing the latency, saving bandwidth, and maximizing data privacy.

A hybrid fog-edge architecture has been demonstrated to be optimal on Internet of Medical Things (IoMT) systems, both in terms of latency and threat resilience [Islam, U. *et al*., 2025]. This architecture performs in-place data processing but coordinates with the fog nodes to perform intermediate calculations, forming a highly scalable architecture. These architectures have the potential of assisting pose-sensitive CNNs by offering localized computational support and safe data processing services.

Figure 1 illustrates a conceptual diagram of a hybrid edge-fog architecture used in real-time health monitoring systems.

Publisher: SARC Publisher

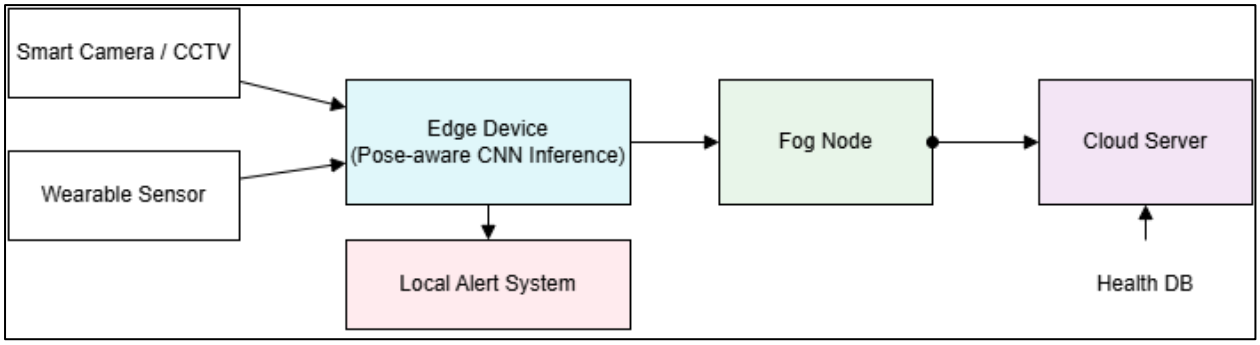

Figure 1: Hybrid Fog-Edge Architecture for Real-Time Health Monitoring
***Source**: Adapted from* [Islam, U. *et al*., 2025]

Edge computing is also being applied in association with AI to increase monitoring of the wellness of the general population and issue early alerts of disease outbreaks and patient decline [Konda, R. 2021]. Using lightweight AI models, such systems will be capable of analyzing real-time streams of wearables, video cameras, or biosensors and determining what patterns require immediate focus. This is achieved by combining pose-aware CNN and edge AI, which provides an all-encompassing approach to detect subtle behavior or posture variations that can lead to a medical emergency.

Personalized models that still maintain privacy have also been enabled through federated learning used in edge AI. An example is that real-time ICU alarm reduction has been met by using federated anomaly models on wearable data streams [Chadha, K. S. 2025]. These structures enable health monitoring systems to train on the distributed sources of data without violating patient privacy, which is a critical factor in a health setting.

Figure 2 presents a performance graph comparing latency between edge and cloud-based health monitoring systems, demonstrating the superior responsiveness of edge computing.

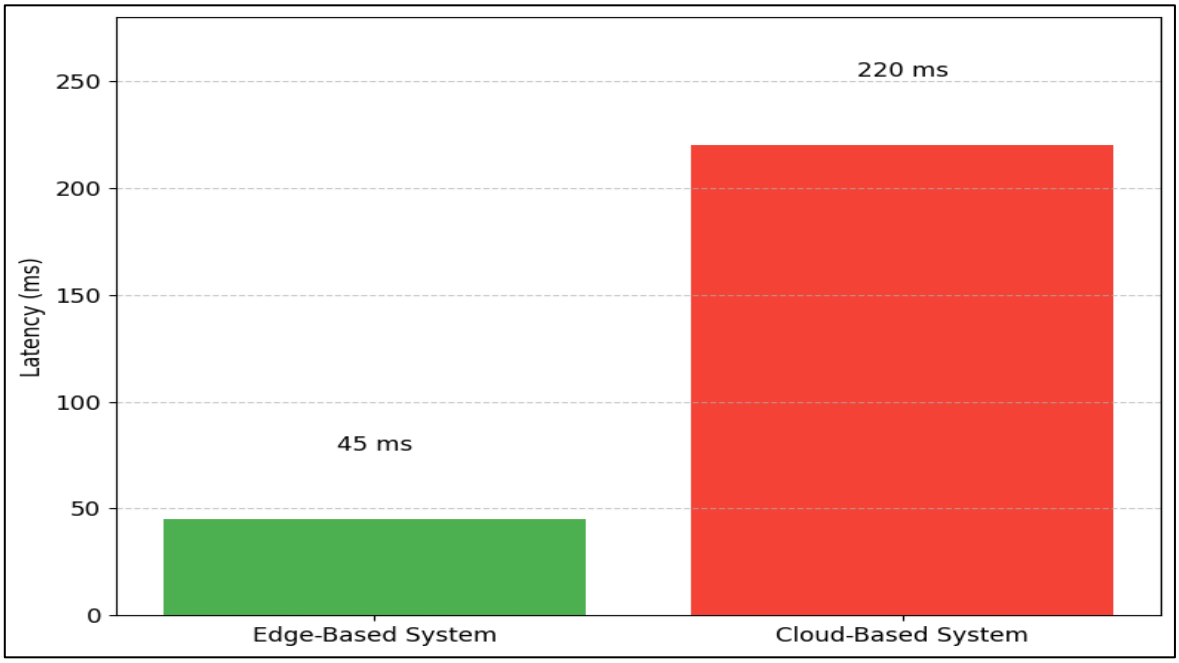

Figure 2: Latency Comparison Between Edge and Cloud-Based Monitoring Systems
***Source**: Created using data from* [Islam, U. *et al*., 2025; Chadha, K. S. 2025]

Optimization Techniques for Pose-Aware Models

Applications of CNNs on edge devices will take serious model optimization to achieve efficiency without compromising accuracy. Model pruning is also one of the main methods that are used, and it entails the elimination of unnecessary weights in the neural network. This minimizes the use of memory and increases the time taken to infer. Quantization also reduces model size by lowering

Publisher: SARC Publisher

the precision of weights, often using 8-bit integers instead of 32-bit floating-point numbers, which is sufficient for many applications.

The optimization strategies based on AI have greatly improved the real-time analytics in smart devices, especially when paired with edge computing [Bargavi, M. *et al*., 2025]. The strategies can be used to effectively deploy the models even on devices with low computational resources, including wearable sensors or mobile phones. Knowledge distillation also enables smaller models (students) to be trained by the bigger, pre-trained models (teachers), which can still be deployed as lightweight models but with high accuracy.

Another important consideration is energy efficiency, which is considered to be particularly important to battery-powered devices. Algorithms should be programmed in such a way as to reduce the amount of computation and maximize the accuracy of the output. The IoT health monitoring systems that use edges have been concerned with maximizing responsiveness, data privacy, and power consumption, making them viable in real-life applications in the long run [Swathi, B. *et al*., 2025]. These considerations are critical when deploying pose-aware CNNs, which continuously process video or sensor streams to detect anomalies.

Table 1: Comparison of Optimization Techniques for Edge Deployment

Optimization Technique	**Description**	**Benefits**	**Application**
Pruning	Removing unnecessary weights	Reduced size and faster inference	Lightweight CNNs
Quantization	Lowering numerical precision	Improved energy efficiency	Embedded systems
Knowledge Distillation	Student-teacher model training	Maintains accuracy in small models	Edge AI
Few-Shot Learning	Training with minimal data	Personalized inference	Health monitoring
Federated Learning	Distributed training on-device	Enhances privacy	Wearable devices

***Source**: Compiled from* [Chadha, K. S. 2025; Bargavi, M. *et al*., 2025; Swathi, B. *et al*., 2025]

Challenges in Deployment and Scalability

Although the use of pose-aware CNNs is beneficial in real-time health emergency detection with significant benefits of lightweight methods, a number of challenges continue to obstruct their extensive application. These are the computational constraints of edge devices, environmental variability, and quality of data. Even CNN models optimized need a minimum of hardware capability to be effectively used. Devices with limited RAM, low processing power, or outdated GPUs may be insufficient to support real-time inference workloads, leading to delayed or inaccurate responses.

Vision-based health monitoring systems face a great challenge posed by environmental variability. Pose estimation can be extremely inaccurate depending on lighting conditions, occlusions, camera angles, and background noise. The real-world setting usually creates unpredictability as opposed to a controlled setting like that of a classroom or a hospital bed. Ideally trained models can perform poorly with these kinds of variations, so robust data augmentation strategies and the use of real-world databases during training need to be developed.

A second issue is that of generalization of the trained models. A pose-trained CNN based on one demographic or type of body structure may not be as effective with another sample because of the physical features, clothing, or behavior of the model. This is added to by the fact that in health emergency detection, normal postures and movements of people can differ considerably to the extent that more personalized or adaptive modeling methods are needed. The models of eye-tracking, such as those employed in WEBEYETRACK, proved that it is possible to achieve the few-shot personalization directly on the devices and enable the systems to adjust to the user-specific patterns in a short time [Davalos, E. *et al*., 2025].

Scalability of edge-based health monitoring systems has become particularly important as it allows the deployment of edge-based health monitoring systems across a wide range of

Publisher: SARC Publisher

hardware configurations without exhaustive reconfigurability or retraining. This requires a design that is modular and interoperable and can be readily modified to support different operating systems, types of sensors, and communication protocols. Additionally, firmware updates and over-the-air retraining will have to be included in the system architecture to sustain performance in the long run.

Other issues are also critical in health monitoring systems: data privacy and ethical concerns. Despite edge computing aiming to reduce the possibility of data leakage through local processing, vulnerabilities are present. Poor device security, insecure firmware, or vulnerable communication protocols can all be exploited to access sensitive health information. Therefore, systems should have effective encryption, secure boot procedures, and data access control tools as the building blocks in the design.

APPLICATION USE CASES AND REAL-WORLD IMPLEMENTATIONS

Pose-aware CNNs that can be used to detect health emergencies in real-time can be used in various applications, especially in elder care, assisted living facilities, and remote patient monitoring. Falls among the elderly are a leading cause of serious injury, and the ability to detect such incidents accurately and promptly can be the difference between life and death. Home automation systems can include lightweight models that can detect such events and send an instant alert to the caregivers or the emergency services.

Pose-aware monitoring can be used to complement conventional patient monitoring systems in critical care units. These systems offer an extra safety measure to patients that are unable to express their distress due to physical distress signals, including convulsions, abnormal posturing, or sudden movement cessation. Federated anomaly detection has been implemented in ICUs to eliminate alarm fatigue, with edge AI used to filter false alarms and focus on the high-risk alerts [Chadha, K. S. 2025]. The same applies to the visual monitoring systems used to enhance the accuracy and decrease caregiver load.

Real-time health analytics based on AI models have also been introduced in wearable technologies. Pose estimation can be applied to smartwatches, fitness bands, wearable cameras, and other devices to identify an unusual behavior. These devices are able to detect falls, collapse, or unusual movement patterns in high-risk areas like construction, firefighting, or security to activate automated safety procedures.

Edge-based pose-aware systems are a scalable and affordable solution in rural and resource-constrained environments. And these environments do not always have high-speed internet or cloud backends, and decentralized solutions are not practical. With the help of the locally processed models, the healthcare delivery is able to reach the underserved groups, minimizing the differences in access and outcomes. The advantage of such implementations is that the architectures are energy-consuming and can run on the solar-powered or battery-operated devices, which makes them more sustainable.

Visual AI systems have already been applied to the educational sector to engage in surveillance. The attentiveness and participation of students during real-time have been monitored through facial expression and body posture [Singh, R. *et al.*, 2025; Aly, M. 2025]. Such principles can be applied to pediatric health monitoring, where behavioral indicators can be used in the initial stages of illness or children whose ailment or discomfort cannot be easily communicated.

Future Directions and Research Opportunities

The development of lightweight pose-aware CNNs that can run on the edge will gain a lot of utility due to the continued progress in hardware, AI optimization algorithms, and multi-modes. The integration of pose estimation with other modalities (e.g., audio analysis, physiological signal processing (e.g., ECG, PPG), and natural language processing) is one of these promising directions. This would enable the use of more thorough health-related assessments taking into account various aspects of patient data.

More studies on self-supervising and semi-supervising learning could be used to solve the issue of the unavailability of labeled data, especially those of rare emergency events. By using these techniques of learning, models are able to acquire helpful features of large amounts of unlabeled data, thereby becoming more suitable to real-life applications. Contrastive learning or temporal coherence modeling are some of the techniques that can help to make pose estimations more robust in situations where they are not.

The integration of hardware accelerators for deep learning, such as Tensor Processing Units (TPUs)

Publisher: SARC Publisher

or Neural Processing Units (NPUs), into consumer-grade devices can significantly enhance the real-time performance of AI models. With the spread of such specialized chips, pose-aware CNNs can grow more complicated without reducing the inference time or energy use. TensorRT and ONNX Runtime optimization frameworks will be critical towards closing the gap between model design and hardware capability.

Interdisciplinary collaboration will be essential for the future development of this field. To make sure that these systems, in addition to being technically solid, are ethically accountable, clinically proven, and universally available, healthcare professionals, AI researchers, hardware engineers, and policymakers should collaborate to guarantee that such systems are responsible and accountable. The regulatory frameworks should be revised to reflect the peculiarities of risks and opportunities of edge AI in healthcare, with the focus on safety, transparency, and accountability.

Lastly, pose-aware health monitoring systems must be assessed using large-scale longitudinal studies to determine their long-term efficacy in different populations and settings. Such studies ought to evaluate clinical outcomes as well as patient satisfaction, workload of the caregiver, cost-effectiveness, and the ability of the system to endure over time.

CONCLUSION

Lightweight pose-aware CNNs are a revolutionary concept of real-time health emergency detection, especially in cases where edge devices are involved. This is because they are able to process images in their immediate area, thus responding quickly, increasing privacy, and minimizing reliance on cloud technology. The technology is based on the progress in related areas like education and animal tracking and is augmented by the developments in edge computing, AI optimization, and federated learning.

Though the outcomes are encouraging, there are still critical issues in the context of the model generalization, the scalability of deployment, and the environmental flexibility. Yet, as further studies and interdisciplinary cooperation are conducted, pose-conscious CNNs can become a part of the modern healthcare system. Their use in home care, ICU monitoring, rural health care, and occupational safety highlights their ability and significance in the worldwide initiative to make healthcare more active, personalized, and accessible.

REFERENCES

1. Singh, R., E, R., & M, N. B. "MMSAD—A multi-modal student attentiveness detection in smart education using facial features and landmarks." *Journal of Ambient Intelligence and Smart Environments* (2025): 18761364251315239.
2. Aly, M. "Revolutionizing online education: Advanced facial expression recognition for real-time student progress tracking via deep learning model." *Multimedia Tools and Applications* 84.13 (2025): 12575-12614.
3. Parivendan, S. C., Sailunaz, K., & Neethirajan, S. "Socializing AI: Integrating Social Network Analysis and Deep Learning for Precision Dairy Cow Monitoring—A Critical Review." *Animals* 15.13 (2025): 1835.
4. Sonnara, F., Chihaoui, H., & Filali, F. "Efficient real-time license plate recognition using deep learning on edge devices." *Journal of Real-Time Image Processing* 22.4 (2025): 159.
5. Davalos, E., Zhang, Y., Srivastava, N., Thatigotla, Y., Salas, J. A., McFadden, S., ... & Biswas, G. "WEBEYETRACK: Scalable Eye-Tracking for the Browser via On-Device Few-Shot Personalization." *arXiv preprint arXiv:2508.19544* (2025).
6. Ranjan, A., & Ravinder, M. "ViConNet: Temporal-Spatial Probe Distilled (TSPD) Network for Real-Time Video Deblurring and Enhancement with Facial Micro-expression and Pose Estimation for Real-Time Video Conferencing." *Circuits, Systems, and Signal Processing* (2025): 1-43.
7. Islam, U., Alatawi, M. N., Alqazzaz, A., Alamro, S., Shah, B., & Moreira, F. "A hybrid fog-edge computing architecture for real-time health monitoring in IoMT systems with optimized latency and threat resilience." *Scientific Reports* 15.1 (2025): 25655.
8. Konda, R. "Edge computing and AI integration for enhancing real-time public health monitoring systems." *International Journal of Multidisciplinary Research and Growth Evaluation* 2.3 (2021): 579-583.
9. Chadha, K. S. "Edge AI for real-time ICU alarm fatigue reduction: Federated anomaly detection on wearable streams." *Utilitas Mathematica* 122.2 (2025): 291-308.

Publisher: SARC Publisher

10. Bargavi, M., Muhammed, H., Harish, P., & Dhanush, D. "Edge Computing and AI for Real-time Analytics in Smart Devices." *Asian Journal of Basic Science & Research* 7.2 (2025): 01-09.
11. Swathi, B., Veerabomma, S., Archana, M., Bhadru, D., Somu, N. L., & Bhavsingh, M. "Edge-Centric IoT Health Monitoring: Optimizing Real-Time Responsiveness, Data Privacy, and Energy Efficiency." *2025 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)*. IEEE, (2025).

Source of support: Nil; **Conflict of interest:** Nil.

Cite this article as:
Venkataraman, K. “Lightweight Pose-Aware CNN for Real-Time Detection of Health Emergencies on Edge Devices.” *Sarcouncil journal of Medical sciences* 5.1 (2026): pp 14-20.

Publisher: SARC Publisher