Received: 2 April 2023 | Revised: 28 June 2023 | Accepted: 28 June 2023

IET Wireless Sensor Systems

DOI: 10.1049/wss2.12064

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

A conditioned reflex action embedded associative context learning-based energy efficient paradigm in smart city milieu

Majid Hussain[1] | Ahmad Bilal[2] | Muhammad Faheem[3] | Muhammad Anwar[4] | Muhammad Sultan Zia[5]

[1]Department of Computer Science, The University of Faisalabad, Faisalabad, Pakistan

[2]Department of Software Engineering, The University of Lahore, Lahore, Pakistan

[3]School of Technology and Innovations, University of Vaasa, Vaasa, Finland

[4]Division of Science and Technology, Department of Information Sciences, University of Education, Lahore, Pakistan

[5]Department of Computer Science, NFC-IEFR, Faisalabad, Pakistan

Correspondence
Muhammad Faheem.
Email: muhammad.faheem@uwasa.fi

Funding information
University of Vaasa; Academy of Finland

Abstract
An intelligent video surveillance system is crucial to enhance public safety, crime prevention, traffic, and crowd management in a smart city milieu. Situational awareness is an essential aspect of these surveillance systems and it is inferred through underlying context aware frameworks. However, these systems may not possess the ability to proactively disseminate the real-time context among its sensor nodes. Moreover, in the specific conditions of occurrence of related or repeated events, these systems may also perform inefficiently through afresh context processing and disseminate cycles, without learning from the relevant context that has already been occurred and processed by the system. It leads to deteriorated performance, especially delay in reaction, overwhelmed processing, and energy expenditures. Therefore, to counter such issues, this research work proposes an energy efficient situational aware framework deployed in visual sensors network that is incorporated with context associative learning. System observes currently occurring context at each instance of an event. Overtime, context is refined and stored in context database. Such mechanism empowers the system to learn from previous experiences and develop relationship among the subsequent events that is embedded through this associative (adaptive) learning. Eventually, each event is processed through intelligent resource allocation, supported through mechanism of context learning that further illustrates the independent functions of reduced processing and improved (rapid) decision making resulting in evolution of energy efficient computing paradigm. Ultimately, the capability of learned reflex-action is induced through introspectively evolved context of the system in entirety and against specific condition of recurred situation depicting minimum energy expenditure.

KEYWORDS
energy conservation, object detection, smart cities, wireless sensor networks

1 | INTRODUCTION

From last few decades Wireless Sensor Networks (WSNs) have been subjected to the intensive research for developing flexible and reliable infrastructures for industrial, commercial, civil, and individual applications [1–3]. WSNs are deployed to acquire, process, and transmit data from an environment optimally without compromising resource efficiency. Visual Sensor Networks (VSNs) are special featured WSNs in which the sensor nodes are equipped with visual devices (cameras), onboard memory and processing capabilities for collaborative image analysis. In smart cities where real-time video surveillance is crucial aspect for ensuring public safety such VSNs can be utilised for automatic event and object detection and immediate decision making [2]. Public safety and security, particularly related to vehicular environment demands smart visual surveillance for mission critical situations that is acquired through real-time visual processing triggered by automatic event detection, for instance, kidnaps, police chases, identifying stolen vehicles and un-authorised intrusions [4]. Those applications of VSNs require capabilities, such as smart use of on-board resources, autonomous decision making in distributed manner, multi-camera synchronisation, and optimal transmission of data [5]. Such intelligent resource utilisation in VSNs is attainable through visual context awareness [6]. Context awareness is the ability of a computer system, artefact, or service to be aware of real-time characteristics of its environment and respond intelligently and proactively against a requisite situation through utilising a sense-decide-actuate cycle and adopt cycle [7].

Context awareness is physically realised through Context Aware Systems (CAS). Although, as contemporary CAS performs proactively and intelligently, it may still be lacking in the ability of constructing association among subsequent events and learn the pertinent pattern of occurrences. Such systems may pose overwhelmed energy expenditure due to reason of over processing caused by inefficient utilisation of system resources when it deals with similar situations [7]. Against each occurrence of an event, a traditional CAS always responds by executing complete sense-decide-actuate cycle to carry out the requisite operation. In a case where an event is reiterated the same sense-decide-actuate cycle is executed in entirety once again. In other words, a CAS may not possess the ability to learn from previous experiences. Consequently, it responds to each instance of event through similar context semantics, situational calculus, perception, decision, actuation and adaptation. Furthermore, limitations of contemporary CASs may become worsen when it is deployed in VSN environment because of the reason that visual processing demands higher computing resources and energy extent [5, 8]. Therefore, a system in which each event is stimulus to entirely engaging all system resources to carry out requisite task results in degraded performance. Due to such limitation the accumulated processing extent becomes significantly higher and system poses overwhelmed energy extent when it operates against replicated events.

However, the limitation of inefficient resource utilisation and its consequent elevated energy extent of the system can be optimised through establishing of relationship and learning the pattern of occurrences among subsequent events. This research work proposes a Conditioned Reflex Action Embedded Associative Context Learning-based Energy Efficient Paradigm in Smart City Milieu (CRIEEM) for detection, tracking and identification of intruding Mobile Object (MO) in urban environment, composed of human inspired memory architecture. As system is exposed to multiple events overtime, its internal database is refined with processed retrospectively occurring context that eventually implants it with the capability of associative context learning. Context database is utilised to provide meta-context about upcoming situations to the system so that it can perform against them with maximum resource efficacy that eventually leads towards the optimal energy expenditure. In other words, system evolves internally and gains the ability to make decisions based on refined context in database rather than exploiting all resources at entirety. A scenario where system responds through entirely self-evolved decision making/actuation in specific circumstances of situational/event repetition imitates as learned reflex-action. The mechanism of context and associative learning among subsequent events transform the system into autonomous, intelligent and proactive computing paradigm that depicts the optimal resource utilisation and corresponding efficient energy extent against each instance of an event.

The rest of the manuscript is divided as follows. Section 2 presents literature review. Section 3 explains the components of the proposed architecture. Whereas, memory and analytical models are discussed in Sections 4 and 5. Performance aspects of the proposed methodology are described in the Section 6. Moreover, Section 7 discusses the limitations and Section 8 concludes this research work.

2 | LITERATURE REVIEW

Context and situation, in contemporary research, are often used as substitutes of each other yet there exists a clear difference between such counterparts. Context is perceived as information about current time, location, nearby objects, people and changes pertinent to them. Schilit et al. [9] presented three important aspects of context those are user location, environment, and nearby resources. The authors in refs. [10, 11] described context as concerned time, location, environment, nearby objects, and devices. Abowd et al. [12] referred context as the information that defines a situation of an entity, that entity may be an object, user and environment that is relevant to interactions between user and application. We perceive context as more than location, time and environmental aspects, it contains any bit of information that has the association (event naïve) with user and application including user's emotional state and focus of attention. Contextual information leads towards the extraction of eventual situation.

However, situation can be described at higher level of abstraction than context. Zimmermann et al. [13] portrayed that situation can be perceived as a snapshot captured through camera that represents the instantaneous and structured settings of contexts at a certain time and space that can be described through specific name. Baker et al. [14] defined situation as “the complete state of universe at an instance of time”, they further stated that situation may be an accumulation of infinite variety of contextual information. Situational recognition with its pertinent reasoning based on related situational calculus and context models are the prime attributes that lay the foundation of context awareness.

Context awareness is the ability of a computer system, application, artefact or service to be aware of its current operating context, understand the accumulated situation and respond intelligently and proactively against any change occurring in this situation/environment based on such contextual awareness [14]. Jaouadi et al. [15] described that context aware applications possess the ability to capture, compute and actuate through adopting the real time context. Rezaee et al. [16] stated that context aware computing has ability to sense, measure and process the multiple situations in the environment. The authors in refs. [12, 17] defined CASs those provide the user with such information and services relevant to the user task. Such capabilities of context awareness, situational consciousness and adaptation imbued by real-time context acquisition is utilised to realise a pragmatic approach to administer the dynamic and heterogeneous computing environments of smart cities.

Implementation of smart cities are primarily based on WSNs and Internet of Things (IoT) [2]. In ref. [18], the authors argued that the IoT is a dynamic environment whose entities change continuously. Furthermore, conventional approaches provide static security measures those may prove to be insufficient when an attack occurs in IoT environment. They presented an adaptable and reconfigurable context sharing mechanism to prevent such security threats. Liu et al. [19] presented a framework for intelligent urban traffic management through maintaining a context repository with contextual information generated by conjunction of 5 G wireless, vehicle ad-hoc software defined networks and mobile edge computing. Muchtar et al. [20] came up with a novel unified framework for detection of moving objects from video captured through surveillance cameras deployed in smart city environment. The authors utilised a block-based texture and spatial-temporal information with the combination of deep-learning based image classification to detect the moving objects rapidly and precisely. Alshammari et al. [21] presented automatic human target detection approach for smart cities to detect, identify, and track targets though utilising multiple wall mounted cameras. Furthermore, the proposed framework intelligently analyses and investigates whether detected target is a threat or not without any human intervention.

Moreover, a research work [22] suggested that there is an immense need for efficient data transmission and collection methods in WSNs due to limited battery power in nodes. The authors propose a cluster formation technique using Particle Swarm Optimisation and a Fuzzy-based Energy Efficient Routing Protocol (E-FEERP) to transmit data optimally from cluster heads to the Base Station. The proposed E-FEERP demonstrates improved network performance in terms of packet delivery ratio, Residual Energy, throughput, energy consumption, load balancing ratio and network lifetime. In ref. [23], the authors highlighted the significance of energy-conserving in WSNs for smart city applications and also proposed an energy efficient cluster-based routing methodology achieved through Zone-Stable Election Protocol deployed in the sensing layer of smart city. The authors in ref. [24] presented a novel authentication mechanism incorporated with blockchain technology and game theory principles for securing Internet of Vehicles (IoV). The proposed solution is incorporated with three-layer multi-trusted authorisation approach, Physical Unclonable Functions, dual gaming for vehicle authentication and Proof-of-Work (dPoW) consensus mechanism. Furthermore, the proposed framework is also practically implemented and showcases its ability to achieve authentication while minimising the computational overhead.

The authors in refs. [25–27] converged the context awareness with internet of things for efficient energy management for smart spaces and buildings. In ref. [28], the authors presented that in recent years E-Healthcare services gained significant attention from the researchers and such services require optimal resource utilisation to achieve efficient energy expenditure. Hence, a context aware model based on combination of three different levels of algorithm is proposed to counter resource wastage and optimise energy expense. Whereas, it is also claimed that system performs extremely well for the E-healthcare services without degrading the time complexity. In ref. [29], the authors proposed an energy efficient asynchronous data transmission strategy under time constraints in the domain of mobile edge computing. The feature of energy saving is realised through binding the computing and channel resources. Furthermore, a game theory based strategy is also suggested to reduce the energy consumption of the mobile edge computing by using a deep reinforcement learning that manages and allocates channel resources. Authors have also argued that overhead cost of network synchronisation is eliminated and various transmission and computation latencies are also reduced. An energy efficient (wireless) flying ad hoc network (FANET) is proposed in ref. [30] to detect forest fires. Clustering approach is utilised to counter the routing and energy challenges that eventually improves the lifetime of an unmanned aerial vehicle. Features such as disaster region and other ground factors are detected through locally installed sensors in forest termed as distinct clusters. These nodes are further connected to the base station via FANET, whereas, interconnection between these sensors and the base station is carried out via optima resource utilisation to realise optimal energy expenditure. Moreover, authors claimed that such system may perform extremely better compared to other similar purposed systems in traditional literature.

In ref. [14] the authors conceptualised the context life cycle (context gathering, reasoning, actuation and adoption) through the context based situation. Authors further presented the implications of context awareness on smart societies. The authors in refs. [6, 31] presented an event-based CAS for MO detection, classification, tracking and identification implemented through camera based networks and distributed image analyses. In addition to that, as argued by ref. [7] systems those can adapt from the context can be plausible solution to efficiently utilise resources and processing. Authors also presented a context adaptive paradigm in WSNs to improve the resource allocation of a surveillance system for processing efficiency. A sink privacy preservation scheme presented by ref. [32] in which authors created a camouflage of cooperating nodes around the sink to prevent the location of the sink from being revealed. In ref. [33] the authors presented a congestion control model to reduce the packet loss in health care networks through developing the new limits of cwnd size in slow start phase. Kane et al. [34] came up with "reflex-tree" a reflex action inspired framework for gas pipeline maintenance in urban environment. Hussain et al. presented A Conditioned Reflex Action Inspired Adaptive Model (CRAM), in which author mimics the learned reflex-behaviour of the system in specific condition of repetition of a situation. Although, context awareness provides significant performance enhancement over context un-aware systems but there still may exist certain performance limitations posed by conventional CASs when they deal with replicated events and situations. An urban environment is susceptible to occurrences of similar or repeated events particularly related to appearances of vehicular MOs.

A video surveillance system for such domain composed of conventional CAS may lack the capability to establish distinction between afresh or reiteration of events due to the fact that system does not construct any internal association among subsequent events over time. Therefore, each instance of an event is considered as arrival of independent situation that may be dealt by system through entirely replicated sense-decide-actuate (processing) cycle leading the system towards degraded performance in terms of overwhelmed energy expenditure. In this research work we present a self-learning and internally evolving processing paradigm to overcome such limitations through optimally retaining and employing current operating context of the system. Overtime, system learns from previous experiences and improves its introspective context. Retained and refined context is engaged to provide meta-context to next actuations of the system that empowers it to utilise its resources optimally when it deals with an instance of afresh event. It leads the system towards the introspectively enhanced (internal) actuations that results in efficient energy expenditure. In Smart Cities, particularly video surveillance, often requires handling of huge amounts of network traffic while maintaining strict quality of service that may not be achievable all the time due to the reason of non-optimal resource utilisation [35]. However, CRIEEM explicitly prevents such limitations by utilising its resources intelligently and delivers proactive situational awareness that can potentially enhance the Smart Cities in terms of optimal extents of energy consumption.

3 | PROPOSED ARCHITECTURE

Underlying framework of CRIEEM is based on CRAM [6] that is developed through tenon-mortise cross-layered modular architecture. Proposed CRIEEM potentially improves the processing layer through altering the certain algorithmic aspects that leads the system towards enhanced energy efficiency however, physical layer does not pose any change. Therefore, in this section we only present processing layer and pertinent modules in perspective to the proposed alterations.

CRIEEM utilises a distributed context learning approach for MO detection, identification, and tracking. Behaviour of all conceptual layers improves introspectively as system starts memorising previous actuations and refines its context repository overtime. When a MO is detected by the system, first node carries out processing at entirety. As MO proceeds towards subsequent node in its trajectory the context processing starts to reduce because of the reason that actuation that is once performed by predecessor node provides the successor node with meta-context based on its own actuation. Since system learns through previous experiences and refines its introspective context, such capability of event based associative learning aids the system to evolve internally and reduces its external actuation. Finally, system reaches a state where context processing reduces to minimum and system truly realises the behaviour of context learning system. Furthermore, at maximum internal (introspective context based) actuation, system actuates impulsively against recurred situation called reflex-action. Figure 1 demonstrates the behaviour of different layers of CRIEEM with the increase in internal actuation the decision, processing, and sensing extents decrease respectively. Although, system sensing time remains the constant but system acquires capability to optimise the quality of image being captured through *instantaneous image acquisition optimisation scheme* that further augments the system in terms of energy efficiency.

3.1 | Processing layer

Processing layer administers logical and algorithmic aspects of the CRIEEM. At each node, in sequential process flow system detects the intruding MO, captures the image, extracts the silhouette, compares it with stored reference silhouettes and calculates the surety of that comparison to determine the type and class of the MO.

Image capturing module manages the acquisition of instantaneous image. This module utilises an *instantaneous image acquisition optimisation scheme* to enhance the resource efficiency at very first phase of processing layer.

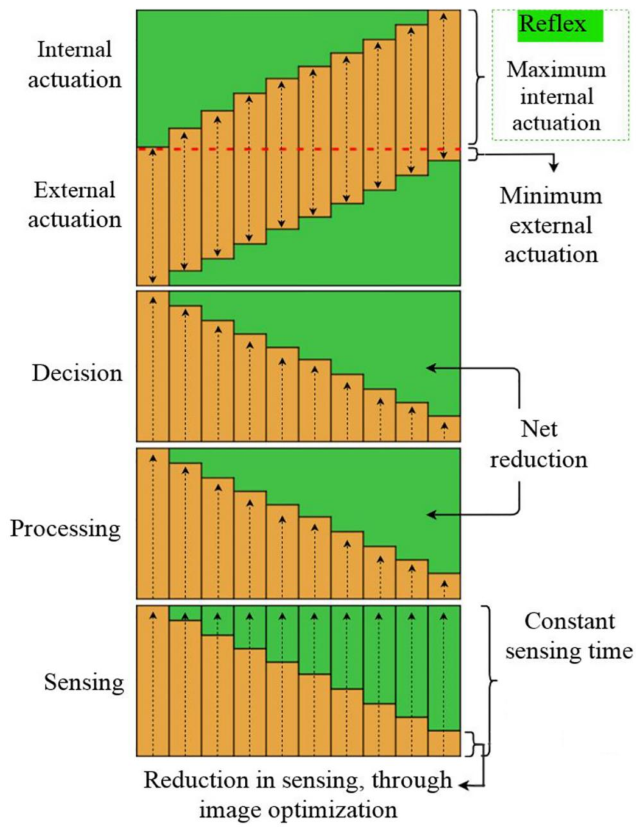

FIGURE 1 Conceptualised behaviour of layers of CRIEEM.

System optimises the quality of instantaneous image being captured based on current surety level of the detected MO as demonstrated in Table 1. Since image processing is the main resource demanding task in proposed framework hence, such optimisation of instantaneous image aids in reducing further processing load that eventually improves the overall energy expenditure of the system. CRIEEM declares its result when it reaches 90 percent surety. At each node surety is calculated through following Equation (1), where k represents the total number of silhouettes deployed at a specific node.

$$surety = \frac{1}{K} \times 100 \tag{1}$$

Instantaneous image is then processed by image change detection module that utilises Gaussian Mixture Model (GMM) to detect intruding MO.

Subsequent to image change detection module there exists silhouette subtraction module that subtracts the silhouette of the detected MO from the image (single frame) yielded by GMM. This module further optimises the processing through Don't Care operation [6]. Table 2 demonstrates the silhouette extraction under different instantaneous image quality levels. It is evident from the extracted silhouettes that distinct shapes are identifiable even when at very low resolutions.

Finally, silhouette comparator module computes the similarity between two (extracted and reference) silhouettes through Correlation Coefficient Equation (2). This approach is often adopted by the traditional literature [36, 37] for feature and image comparison. In our proposed approach, I and S are denoted as two separate input silhouettes (2d vectors) subsequently, $\overline{I}$ and $\overline{S}$ represent the mean values respectively. Yielded compression results through this module have been exemplified in Table 3. Extracted silhouette of intruding MO is identical to stored reference silhouette at second comparison hence, algorithm outputs maximum similarity 1.

$$similarity = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(I_{mn}-\overline{I}\right)\left(S_{mn}-\overline{S}\right)}{\sqrt{\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\left(I_{mn}-\overline{I}\right)^{2}\right)\left(\sum_{i=1}^{m}\sum_{j=1}^{n}\left(S_{mn}-\overline{S}\right)^{2}\right)}} \tag{2}$$

Previous study [6] utilises feature-dependent silhouette segmentation (FRILL) that is an aspect ratio-based operation. It requires feature segmentation of both extracted and stored silhouettes before initiating any comparison processing. It may pose certain processing overhead as segmentation process takes place and then for each segment to be loaded into memory. Therefore, a system that utilises FRILL may perform inefficiently in terms of processing where large number of

TABLE 1 Surety based instantaneous image optimisation.

Surety level (%)	Image quality level	Image quality (%)
0–25	Level 4	100
26–50	Level 3	75
51–75	Level 2	50
76–100	Level 1	25

TABLE 2 Silhouette subtraction under different instantaneous image quality levels.

Video quality	Video resolution	Extracted silhouette
100%	1080 × 720	
75%	810 × 456	
50%	540 × 304	
25%	270 × 152	

TABLE 3 Similarity of silhouettes through Correlation Coefficient.

Input silhouette	Stored silhouette	Similarity	Recognised?
		0.40	✗
		1.00	✓
		0.22	✗

silhouettes are deployed. However, operation of Correlation Coefficient is performed through pixel by pixel comparison but it potentially performs efficiently when compared to the FRILL. Because of the fact that extracted silhouettes have low pixel density moreover, there are minimum loading delay because it does not require segmentation of either of extracted or reference silhouettes. Table 4 demonstrate the different scenarios of silhouette comparisons through both FRILL and Correlation Coefficient.

4 | MEMORY ARCHITECTURE

CRIEEM utilises a human inspired memory architecture as illustrated in Figure 2. It is comprised of three types of sensory, short-term and long-term memories. Sonars operate as sensory receptors to detect the intruding MO and short-term memory is utilised as runtime working memory. Whereas, the long-term memory is subdivided into two sub-categories, declarative (explicit), and non-declarative (implicit) memories. Declarative memory is responsible for storing permanent data such as topology and silhouette tables with respective identifications, camera positions and orientations, views, octets, angles and sureties. Simultaneously, the non-declarative memory is inducted in CRIEEM as priming memory that is deployed by only Edge Nodes (EN) in outer stratum and it is the core aspect for realising reflex-action.

Studies [38, 39] provide evidence about the existence of two—instinctive and conditional reflexive behaviours in humans and animals. Instinctive reflexes are triggered to perform required functions without acquiring underlying foundations of association among preceding experiences and those are never learned consciously or unconsciously by the subject. On the contrary, conditional reflexes carry out their pertinent operations based on learned information through repetition (priming) of events in the non-declarative memory acquired through prior experiences.

In a scenario, when a MO is detected at EN, its silhouette is compared with previously primed MO silhouette. A case when such comparison yields that both silhouettes are identical then system disseminates a recognition message to sink without initiating any further processing and activation of the downstream nodes. Otherwise, system stores newly extracted silhouette in the priming memory and discards previously retained silhouette (if there is any). Such mechanism where first node responds instantly through referring to its priming memory gives the system the capability to imitate the learned reflex-action in specific scenario of recurred situation (Discuss Learned/Unreeled Reflex action in literature review).

TABLE 4 Frill versus Correlation coefficient based comparison.

Operation	Silhouettes × segments	Lapsed time (sec)
Frill	1 × 4	0.041
	3 × 4	0.128
	18 × 4	0.638
Correlation coefficient	1	0.007
	3	0.023
	8	0.151

5 | ANALYTICAL MODEL FOR SILHOUETTE COMPARISON

In this section we present the analytical treatment based on hierarchal classification for comparison scenarios between extracted and stored reference silhouettes. Suppose a trajectory of n nodes (Figure 3) triggered by intruding MO as represented by set $N = \{N_1, N_2, N_3, \ldots, N_n\}$, where each node N_i possesses one of following sets compromised of reference silhouette information as depicted in Figure 4. Extracted and reference silhouettes are compared linearly.

$$P_i = \{1, 2, 3, \ldots, t_i\}, C_i = \{1, 2, 3, \ldots, v_i\} \; and \; J_i = \{1, 2, 3, \ldots, u_i\}.$$

Where $P_i \& C_i$ are independent sets and $J_i \subset C_i$, $u_i \leq v_i$.

There also exists a set x that represents the skipped elements from C_i due to context awareness.

At very first node N_1, system only expects front view of target MO. In this simulation, all Parent-Classes (PC) are

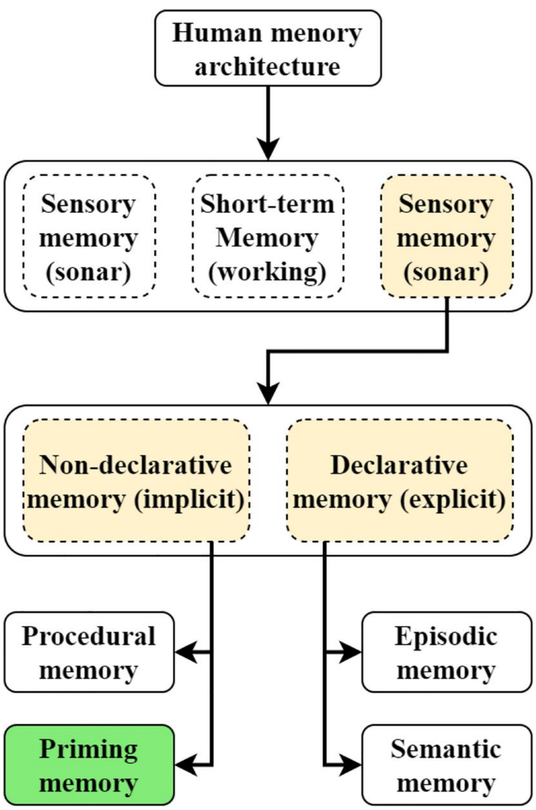

FIGURE 2 Human inspired memory architecture.

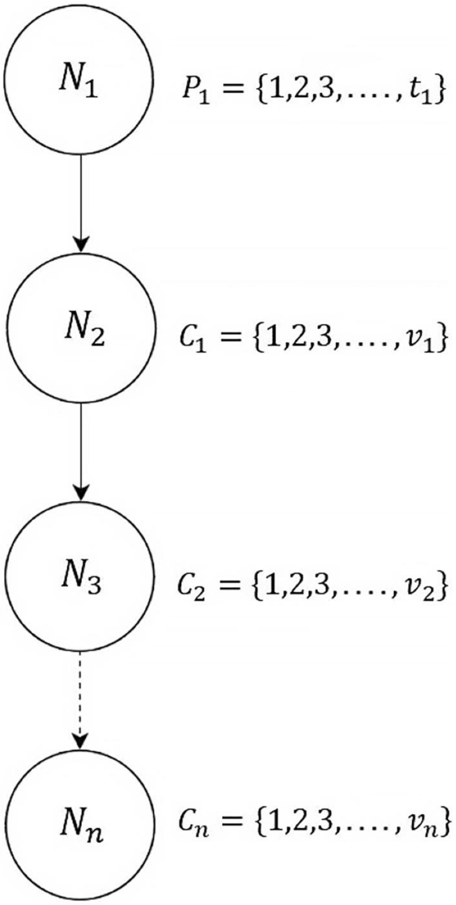

FIGURE 3 Trajectory of intruding Mobile Object.

comprised of two Child-Classes (CC). Each CC has resemblance to its sibling CC that can be utilised to reduce comparison processing through deploying silhouettes in P_1 optimally. For example, front views of Car and SUV (sibling CCs) share similarities with each other therefore, system only deploys one of them in P_1 to determine the Parent-class of the MO as demonstrated in Table 5.

Worst case: If intruding MO's silhouette matches in P_i at very last comparison t_1.

$$\sum_{t=1}^{t_1} 1 = t_1. \tag{3}$$

Average case: If intruding MO's silhouette matches in P_i in-between $1_t h$ and t_1.

$$\frac{1}{t_1}\sum_{i=1}^{t_1} i = \frac{t_1+1}{2}. \tag{4}$$

Best case: It is the depiction of reflexive behaviour of the system, where only 1 silhouette comparison is carried out through utilisation of priming memory as illustrated in Figure 5.

Actuation of N_1 is retained and utilised as input context at N_2. In other words, previous experience of the system provides awareness about upcoming event to next node in MO's trajectory. Therefore, at N_2 system only traverses the silhouettes elements in C_1 those are related to PC of detected MO at N_1.

Worst case: If intruding MO's silhouette matches in c_1 at very last comparison v_1.

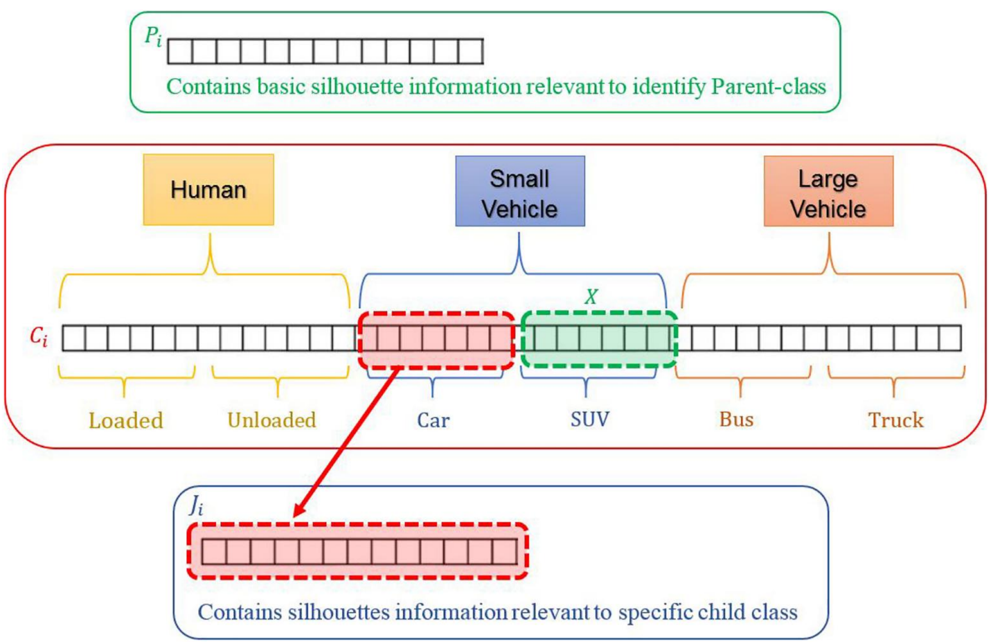

FIGURE 4 Conceptualised Arrangement of Silhouette sets.

TABLE 5 Sibling child-classes with resembling features.

Extracted silhouette	Extracted silhouette description	Stored silhouette	Stored silhouette description	Similarity	Recognition?
	Car front		Human-loaded front	0.3849	✗
			SUV front	0.5453	✓
			Truck front	0.3584	✗

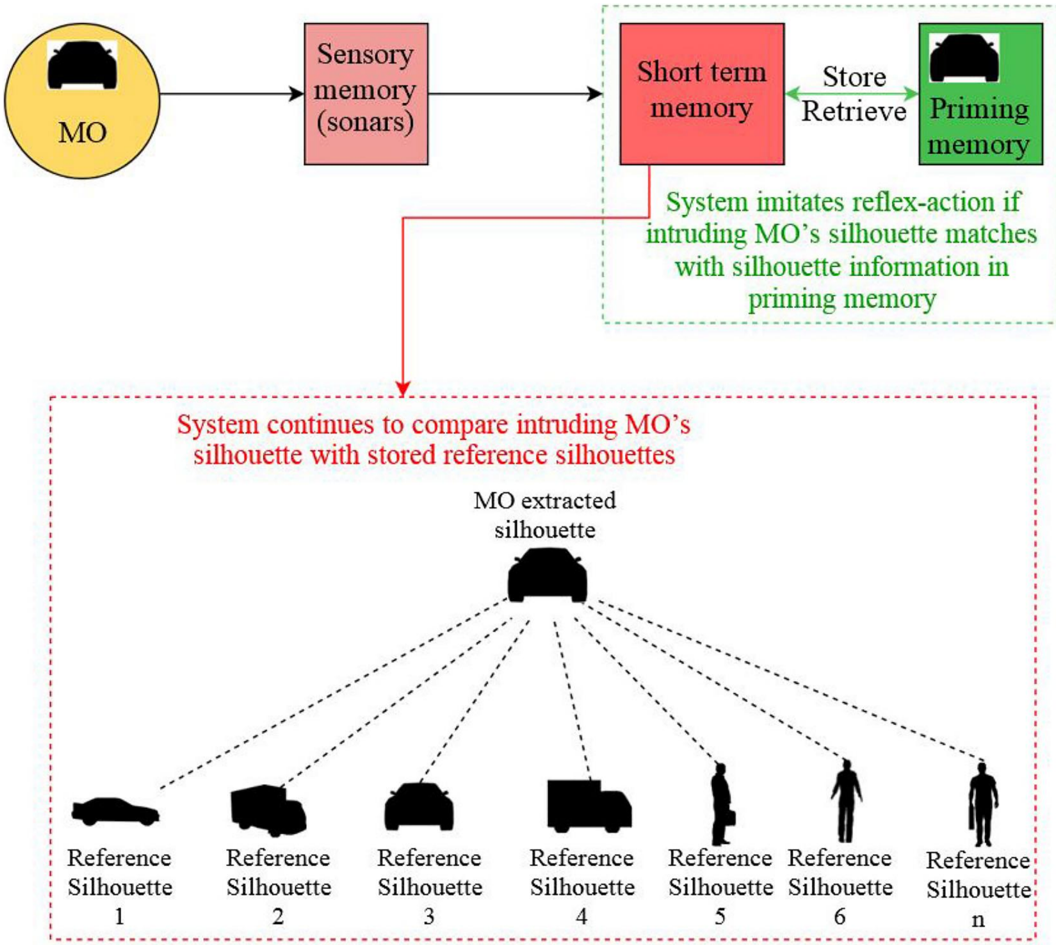

FIGURE 5 CRIEEM reflex-action through priming memory and normal function through episodic memory.

$$\left(\sum_{i=1}^{v_1} 1\right) - x = v_1 - x \qquad (5)$$

Where x represents non-relevant silhouette information (skipped elements) from set C_1.

Average case: If intruding MO's silhouette matches in C_i in between 1_th and v_1.

$$\left(\frac{1}{v_1}\sum_{i=1}^{v_1} i\right) - x = \frac{v_1 + 1}{2} - x. \qquad (6)$$

Best case: Only 1 silhouette comparison is carried out at N_2, if target's silhouette matches in C_1 at $1_t h$ compression.

A scenario when wrong object detection message is received at N_2 from N_1 then comparison operation yields incorrect result. Consequently, at N_2 system has to traverse the entire silhouette information in C_1.

$$\begin{aligned}&\sum_{i=1}^{t_1} 1 + \left(\sum_{i=1}^{v_1} 1\right) - x + \sum_{i=1}^{v_1} \\ &= t_1 + 2v_1 - x.\end{aligned} \tag{7}$$

We suppose, at this point system has already recognised MO PC at N_1 and CC at N_2. As system learns from previous actuations therefore, at N_3 it will only traverse a set j_2 that is sub set of C_1 and comprises of such silhouette information that is relevant to the exact CC of target MO.

Worst case: When all elements in J_2 are traversed.

$$\sum_{i=1}^{u_2} 1 = u_2. \tag{8}$$

Average case: If intruding MO's silhouette matches in J_2 in-between $1_t\ h$ and u_2.

$$\frac{1}{u_2}\sum_{i=1}^{u_2} i = \frac{u_2+1}{2}. \tag{9}$$

Best case: Only 1 silhouette comparison is carried out at N_3.
By adding all silhouette compared till N_3.
Worst case:

$$\begin{aligned}&\sum_{i=1}^{t_1} 1 + \left(\sum_{i=1}^{v_1} 1\right) - x + \sum_{i=1}^{u_2} 1 \\ &= t_1 + u_2 + v_1 - x.\end{aligned} \tag{10}$$

Average case:

$$\begin{aligned}&\frac{1}{t_1}\sum_{i=1}^{t_1} i + \frac{1}{v_1}\left(\sum_{i=1}^{v_1} i\right) - x + \frac{1}{u_2}\sum_{i=1}^{u_2} i \\ &= \frac{t_1+1}{2} + \frac{v_1+1}{2} - x + \frac{u_2+1}{2} \\ &= \frac{t_1+v_1+u_2+3}{2} - x\end{aligned} \tag{11}$$

Best case: Total 3 silhouettes are compared.

In a case when wrong object detection message is sent to N_3 from N_2. Consequently, at N_3 system has to traverse the entire silhouette information in set C_2 once again.

$$\begin{aligned}&\left(\sum_{i=1}^{v_1} 1\right) - x + \sum_{i=1}^{u_2} 1 + \sum_{i=1}^{v_2} 1 \\ &= v_1 - x + u_2 + v_2.\end{aligned} \tag{12}$$

By adding all silhouette compared till N_n.
Worst case:

$$\begin{aligned}&\sum_{i=1}^{t_1} 1 + \left(\sum_{i=1}^{v_1} 1\right) - x + \sum_{i=1}^{u_2} 1 +, \ldots.., + \sum_{i=1}^{u_n} 1 \\ &= t_1 + v_1 - x + u_2 +, \ldots.., + u_n \\ &= t_1 + v_1 - x + \sum_{i=2}^{n} u_i.\end{aligned} \tag{13}$$

Average case:

$$\begin{aligned}&\frac{1}{t_1}\sum_{i=1}^{t_1} i + \frac{1}{v_1}\left(\sum_{i=1}^{v_1} i\right) - x + \frac{1}{u_2}\sum_{i=1}^{u_2} i +, \ldots.., + \frac{1}{u_n}\sum_{i=1}^{u_n} i \\ &= \frac{t_1+1}{2} + \frac{v_1+1}{2} - x + \frac{u_2+1}{2} +, \ldots.., + \frac{u_n+1}{2} \\ &= \frac{1}{2}\left(t_1 + v_1 - x + n + \sum_{i=2}^{n} u_i\right).\end{aligned} \tag{14}$$

Best case: n number of silhouettes comparison are carried out.

A scenario, where comparison operation yields wrong results at each node and incorrect object detection messages are sent to next hop-nodes.

$$\begin{aligned}&\sum_{i=1}^{t_1} 1 + \left(\sum_{i=1}^{v_1} 1\right) - x + \sum_{i=1}^{v_1} 1 + \sum_{i=1}^{u_2} 1 + \sum_{i=1}^{v_2} 1 + \sum_{i=1}^{u_3} 1 \\ &\quad + \sum_{i=1}^{v_3} 1, \ldots.., \sum_{i=1}^{u_n} 1 + \sum_{i=1}^{v_n} 1 \\ &= t_1 + (v_1 - x + v_1 + u_2 + v_2 + u_3 + v_3, \ldots.., + u_n + v_n) \\ &= t_1 + v_1 + \left(\sum_{i=2}^{n} u_i + \sum_{i=1}^{n} v_i\right) - x\end{aligned} \tag{15}$$

6 | CRIEEM PERFORMANCE

Proposed model utilises associative context learning to optimise the resource utilisation of the system against subsequent events that eventually leads towards efficient energy expenditure. Table 6 depicts the performance of CRIEEM with conventional

TABLE 6 System specifications for performance evaluation.

Features	Specifications
Image processing	MATLAB R2017b
Mobile object animations	Unity 2018.1.4f1
Processor	Core i5, 2.2 GHz

TABLE 7 CRIEEM performance with conventional context aware behaviour.

Nodes	IQ	SC	DC	Surety	T (sec)	E_T (J)
1	4	3	4	16.66	3.88	46.56
2	4	6	4	33.33	3.80	57
3	4	18	4	38.88	3.24	87.48
4	4	6	4	55.55	3.23	48.45
5	4	18	4	61.11	3.24	87.48
6	4	6	4	77.77	3.23	48.45
7	4	18	4	83.33	3.24	87.48
8	4	6	4	100	3.25	48.75

TABLE 8 CRIEEM performance with conventional context aware behaviour.

Nodes	IQ	SC	DC	Surety	T (sec)	E_T (J)
1	4	3	4	33.33	3.44	41.28
2	3	2	4	50	1.63	16.3
3	3	3	4	55.54	1.52	16.72
4	2	2	4	72.20	0.72	6.48
5	2	3	2	77.75	0.72	5.76
6	1	2	2	94.41	0.39	2.34

Context Aware (C_A) behaviour when it is simulated to detect, identify and track the intruding MO of type Car. Such system lacks the capability to establish association among successive events. Therefore, each event is considered afresh and dealt with replicated processing cycle that eventually results in static (non-optimal) energy expenditure. For example, image quality remains the maximum (at level 4) and number of silhouettes compared at each node is also higher.

Whereas, in terms of context associative learning, Table 7 demonstrates the performance of the CRIEEM when it is deployed with event based associative Context Learning (C_L) behaviour. It significantly enhances the energy efficiency of the system through intelligent and proactive resource utilisation achieved by introspectively refined context. When the system reaches at maximum internal actuation (introspective context based), it responds against repetition of a situation by utilising priming memory that provides the mechanism to associate with previously occurred event. For instance, at each node, system learns from previous context, optimises image quality and reduces its processing and energy expenditure. Therefore, Table 7 illustrates that image quality is changing (optimising) and number of silhouettes compared at each node is also reducing.

System declares the result at very first node without initiating any further processing termed as learned-reflex action that depicts minimum energy extent (Table 8). When the system performs through node by node operation (reflex-action not initiated), each node consumes certain Time (T) and corresponding Energy (E) to carry out requisite processing. Therefore, in that case, total elapsed E and T expenditures of the system are determined through the accumulation of energy and time system takes at all triggered nodes individually. However, in case of reflexive-behaviour of the system it responds through activating only a single node (edge node) therefore, total time and energy extents of the system is determined by only that single node performance.

Energy extent of computing devices is directly based on CPU utilisation (amount of processing carried out for certain time period) when performing an operation. A conceptualised model based on relation $Energy = Power \times Time$ can be utilised to calculate the energy extent of a system at each specific node.

$$E_T = [(GMM \times IQ) + (SS \times DC) + SC + C] \times T \quad (16)$$

where,

E_T = Total Energy Consumption at a specific node
GMM = Gaussian Mixture Model
IQ = Image Quality
SS = Silhouette Subtraction
DC = Don't Care operation [6]
SX = Silhouettes Compared at specific node
T = Time taken by processing at specific node
C = Sensing Constant (Sonar + Image Sensor)
W = Watt

Let's Assume, following parameters consumes certain power in Watts.

GMM = 1 W
SS = 1 W
DC, 4 Segments = 4 W, 3 Segments = 3 W, 2 Segments = 2 W, 1 Segments = 1 W
IQ, 100% = 4 W, 75% = 3 W, 50% = 2 W, 25% = 1 W
C = 1 W

C_L behaves energy efficient at each subsequent node. Based on simulated results of CRIEEM we can formulate following conclusions, where E and T represent Energy and Time expenditures respectively (Table 9).

$$E_m < E_{(m-1)} < E_{(m-2)} < \ldots.. < E_1 \quad (17)$$

$$T_m < T_{(m-1)} < T_{(m-2)} < \ldots.. < T_1 \quad (18)$$

$$E_{C_L} < E_{C_A} \;\&\; T_{C_L} < T_{C_A} \quad (19)$$

TABLE 9 CRIEEM performance with imitation of reflex-action.

Case	Captured image	MO type	T (sec)	E_T (J)
1		Car front	3.24	32.4
2		Bus front	3.27	32.7

7 | LIMITATIONS

This proposed system is presented as a testbed for context association and learning (intelligent context acquisition, processing, storage, and dissemination) in the paradigm of VSNs. With results this research work has established that the system achieves significant improvements in energy conservation compared to traditional CAS. There does exist certain limitations, for example, natural environment is dynamic and continuously changing, especially the light conditions. Proposed system captures and processes the instantaneous images to detect and classify the MO in it. Light conditions may cause inconsistent and false calculations because of the reason that captured image may contain glares and shadows. Moreover, if MO image is captured from an angle that has no reference in the system, then it may pose false results. Such limitations can be eliminated through increasing the reference silhouette angles stored in system database and number of cameras covering different expected angles of MOs. Machine learning algorithms can also be utilised to eliminate impurities in the captured image and improve the system performance in terms of precision and accuracy. However, such integrations into the system may make it slower and more resource consuming.

8 | CONCLUSIONS

The proposed research work presents a situational aware video surveillance architecture for MO detection, identification, and tracking in smart cities. System achieves optimal energy extent through intelligent and proactive resource utilisation against subsequent visual events through associative context leaning and sharing among its nodes. Conceptual, architectural, prototyped deployment, and analytical aspects of this context learning system are proposed through establishing an analogy with human inspired memory architecture that have shown optimal operational flow. Furthermore, it has also been proved that context learning and conditional reflex-action can be incorporated into VSNs. Through simulated implantation, we have successfully established that visual context learning system delivers improved performance through cumulative resource efficiency and minimum energy expense when compared to conventional visual context aware system. Visual sensor networks comprised of associative context learning have a great potential to provide flexible and reliable video surveillance infrastructures in urban environments, eventually leading towards the initiative of safe and secure *smart Cities with optimal resource utilisation*.

Whereas, future work aims to incorporate the deep leaning models with such context learning systems to further enhance the accuracy and energy efficiency. Moreover, such integration could also make the system scalable to more classes of objects and scenarios.

AUTHOR CONTRIBUTIONS

Majid Hussain and Ahmad Bilal contributed in conceptualisation, methodology, data collection, data analysis, writing – original draft. Muhammad Faheem contributed in supervision, software, testing, validation, writing – review & editing. Muhammad Anwar and Muhammad Sultan Zia contributed in formal analysis, writing – review & editing.

ACKNOWLEDGEMENTS

The authors are highly grateful to their affiliated universities for providing research funding and facilities.

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no conflicts of interest as defined by this specific Journal/Publisher, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

DATA AVAILABILITY STATEMENT

No dataset was used in this research work. All data and information presented in this study are derived from publicly available sources. The silhouettes are extracted from publically available images of objects. Therefore, there is no dataset available for sharing.

ORCID

Ahmad Bilal https://orcid.org/0000-0002-0568-3467
Muhammad Faheem https://orcid.org/0000-0003-4628-4486

REFERENCES

1. Kocakulak, M., Butun, I.: An overview of wireless sensor networks towards internet of things. In: 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC), pp. 1–6. IEEE (2017)
2. Rashid, B., Rehmani, M.H.: Applications of wireless sensor networks for urban areas: a survey. J. Netw. Comput. Appl. 60, 192–219 (2016). https://doi.org/10.1016/j.jnca.2015.09.008
3. Modieginyane, K.M., et al.: Software defined wireless sensor networks application opportunities for efficient network management: a survey. Comput. Electr. Eng. 66, 274–287 (2018). https://doi.org/10.1016/j.compeleceng.2017.02.026
4. Hu, L., Ni, Q.: IoT-driven automated object detection algorithm for urban surveillance systems in smart cities. IEEE Internet Things J. 5(2), 747–754 (2017). https://doi.org/10.1109/jiot.2017.2705560

5. Soro, S., Heinzelman, W.: A survey of visual sensor networks. Adv. Multimed. 6, 1–21 (2009). https://doi.org/10.1155/2009/640386
6. Hussain, M., et al.: CRAM: a conditioned reflex action inspired adaptive model for context addition in wireless sensor networks. J. Sens. 2016, 1–24 (2016). https://doi.org/10.1155/2016/6319830
7. Bilal, A., Hussain, M.: LRAP: A learned reflex action embedded associative context learning based processing efficient paradigm in visual sensor networks. In: Proceedings of the 3rd International Conference on Future Networks and Distributed Systems. p. 42. ACM (2019)
8. Charfi, Y., et al.: Challenging issues in visual sensor networks. IEEE Wireless Commun. 16(2), 44–49 (2009). https://doi.org/10.1109/mwc.2009.4907559
9. Schilit, B.N., et al.: Context-aware Computing Applications. Xerox Corporation, Palo Alto Research Center (1994)
10. Baldauf, M., Dustdar, S., Rosenberg, F.: A survey on context-aware systems. Int. J. Ad Hoc Ubiquitous Comput. 2(4), 263–277 (2007). https://doi.org/10.1504/ijahuc.2007.014070
11. Ryan, N.S., Pascoe, J., Morse, D.R.: Enhanced reality fieldwork: the context-aware archaeological assistant. In: Computer Applications in Archaeology. Tempus Reparatum (1998)
12. Abowd, G.D., et al.: Towards a better understanding of context and context-awareness. In: International Symposium on Handheld and Ubiquitous Computing, pp. 304–307. Springer (1999)
13. Zimmermann, A., Specht, M., Lorenz, A.: Personalization and context management. User modeling and user-adapted interaction. User Model. User-Adapted Interact. 15(3-4), 275–302 (2005). https://doi.org/10.1007/s11257-005-1092-2
14. Baker, N., et al.: Context-aware systems and implications for future internet. In: Future Internet Assembly, pp. 335–344 (2009)
15. Jaouadi, I., Djemaa, R.B., Abdallah, H.B.: Approach to model-based development of context-aware application. J. Comput. Commun. 3(05), 212–219 (2015). https://doi.org/10.4236/jcc.2015.35027
16. Rezaee, Z., Malek, M.R.: A context-aware approach for the spatial data infrastructure portal. Int. J. Geosci. 6(01), 79–87 (2015). https://doi.org/10.4236/ijg.2015.61006
17. Casillo, M., et al.: Context-aware computing for improving the touristic experience: a pervasive app for the Amalfi coast. In: 2017 IEEE International Workshop on Measurement and Networking (M&N), pp. 1–6. IEEE (2017)
18. de Matos, E., et al.: Providing context-aware security for IoT environments through context sharing feature. In: 2018 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE), pp. 1711–1715. IEEE (2018)
19. Liu, J., et al.: High-efficiency urban-traffic management in context-aware computing and 5G communication. IEEE Commun. Mag. 55(1), 34–40 (2017). https://doi.org/10.1109/mcom.2017.1600371cm
20. Muchtar, K., et al.: A unified smart surveillance system incorporating adaptive foreground extraction and deep learning-based classification. In: 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), pp. 302–305. IEEE (2019)
21. Alshammari, A., Rawat, D.B.: Intelligent multi-camera video surveillance system for smart city applications. In: 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), pp. 0317–0323. IEEE (2019)
22. Narayan, V., Daniel, A., Chaturvedi, P.: E-FEERP: enhanced fuzzy based energy efficient routing Protocol for wireless sensor network. Wireless Pers. Commun. 131, 1–28 (2023). https://doi.org/10.1007/s11277-023-10434-z
23. Dogra, R., et al.: Essence of scalability in wireless sensor network for smart city applications. In: IOP Conference Series: Materials Science and Engineering, vol. 1022. IOP Publishing (2021).012094
24. Gupta, M., et al.: Game theory-based authentication framework to secure internet of vehicles with blockchain. Sensors 22(14), 5119 (2022). https://doi.org/10.3390/s22145119
25. Fakeeh, K.A.: An IoT based smart power mangement system for technical university. Int. J. Comput. Appl. 975, 8887 (2016)
26. Kamienski, C., et al.: Context-aware energy efficiency management for smart buildings. In: 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT), pp. 699–704. IEEE (2015)
27. Mahapatra, C., Moharana, A., Leung, V.: Energy management in smart cities based on Internet of Things: peak demand reduction and energy savings. Sensors 17(12), 2812 (2017). https://doi.org/10.3390/s17122812
28. Sharma, A., Kumar, R.: A constrained framework for context-aware remote E-healthcare (CARE) services. Trans. Emerg. Telecommun. Technol. 33(8), e3649 (2022). https://doi.org/10.1002/ett.3649
29. Chen, M., et al.: A game-based deep reinforcement learning approach for energy-efficient computation in MEC systems. Knowl. Base Syst. 235, 107660 (2022). https://doi.org/10.1016/j.knosys.2021.107660
30. Bharany, S., et al.: Wildfire monitoring based on energy efficient clustering approach for FANETS. Drones 6(8), 193 (2022). https://doi.org/10.3390/drones6080193
31. Jabbar, S., et al.: VISTA: achieving cumulative VIsion through energy efficient Silhouette recognition of mobile Targets through collAboration of visual sensor nodes. EURASIP J. Image Video Process. 2014(1), 32 (2014). https://doi.org/10.1186/1687-5281-2014-32
32. Ubaid, S., et al.: SCOUT: a sink camouflage and concealed data delivery paradigm for circumvention of sink-targeted cyber threats in wireless sensor networks. J. Supercomput. 74(10), 5022–5040 (2018). https://doi.org/10.1007/s11227-018-2346-1
33. Ahmad, M., et al.: End-to-end loss based TCP congestion control mechanism as a secured communication technology for smart healthcare enterprises. IEEE Access 6, 11641–11656 (2018). https://doi.org/10.1109/access.2018.2802841
34. Kane, J., et al.: Reflex-Tree: A Biologically Inspired Parallel Architecture for Future Smart Cities. In: 44th International Conference on Parallel Processing, pp. 360–369 IEEE (2015)
35. Kunst, R., et al.: Improving network resources allocation in smart cities video surveillance. Comput. Network. 134, 228–244 (2018). https://doi.org/10.1016/j.comnet.2018.01.042
36. Zhou, H., Wang, X., Zhu, R.: Feature selection based on mutual information with correlation coefficient. Appl. Intell. 52(5), 1–18 (2022). https://doi.org/10.1007/s10489-021-02524-x
37. Setiawan, A.W.: Image segmentation metrics in skin lesion: accuracy, sensitivity, specificity, dice coefficient, Jaccard index, and Matthews correlation coefficient. In: 2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM), pp. 97–102. IEEE (2020)
38. Anokhin, P.K.: Biology and Neurophysiology of the Conditioned Reflex and its Role in Adaptive Behavior: International Series of Monographs in Cerebrovisceral and Behavioral Physiology and Conditioned Reflexes, vol. 3. Elsevier (2016)
39. Iles, J., Stokes, M., Young, A.: Reflex actions of knee joint afferents during contraction of the human quadriceps. Clin. Physiol. 10(5), 489–500 (1990). https://doi.org/10.1111/j.1475-097x.1990.tb00828.x

How to cite this article: Hussain, M., et al.: A conditioned reflex action embedded associative context learning-based energy efficient paradigm in smart city milieu. IET Wirel. Sens. Syst. 13(4), 151–162 (2023). https://doi.org/10.1049/wss2.12064