Introduction: AI's Role in Safeguarding Critical Infrastructure

Critical infrastructure spanning energy grids, transportation, water systems, and urban assets underpins society but faces growing threats from climate change, cyberattacks, faults, degradation, and physical intrusions. Artificial intelligence (AI) offers proactive, efficient defenses, outperforming traditional approaches. This seminar examines AI applications in these areas, highlighting innovations, benefits, and challenges.

First, to counter degrading infrastructure, AIoT-CitySense equips waste trucks with cameras and AI for roadside monitoring in Melbourne, detecting dumped rubbish (97% F1-score), vandalized shelters (96%), and damaged signs (71%) 85% more issues than citizen reports, at low cost (~AUD $200/day) [1].

To protect against cyberattacks in energy sectors, AI excels: Deep learning detects ransomware with 98% accuracy by monitoring networks. Algorithms like Random Forest and SVM spot hardware anomalies, cutting undetected incidents by 90%. Predictive analytics forecast failures with 95% precision, reducing response times by 70% and false positives [2].

For fault protection, LSTM networks in wastewater plants analyze sensor data over time, achieving 96.5% accuracy in detecting ammonia faults surpassing variance analysis (93.3%) and PCA-SVM (93%) minimising false alarms and environmental risks [3].

Against physical intrusions, reconfigurable CPS for surveillance uses edge CNNs (e.g., MobileNetV2) for real-time human detection and cloud-based tracking with Kalman filters and facial recognition, switching modes for threats while addressing privacy and bias concerns [4].

Finally, AI aids climate adaptation with Deltares models simulating sea-level rise effects on infrastructure, suggesting protections like sea walls. FloodNet uses deep learning on images and water data for real-time flood predictions, enabling cities to manage power and water to prevent outages [5].

These cases illustrate AI's shift to predictive protection, with ethical issues like data privacy and explainability. Future work includes enhancing resilience and scalability for broader impact.

AIoT-CitySense: AI and IoT-Driven City-Scale Sensing for Roadside Infrastructure Maintenance

Motivation

Traditional roadside infrastructure maintenance relies on citizen reports or manual inspections by council staff. This approach is reactive, slow, and costly, often resulting in delays, inaccurate data, and unreported issues. With cities expanding and infrastructure budgets under pressure, there is a strong need for proactive, scalable, and cost-efficient solutions. Emerging smart city technologies, particularly AI, IoT, and 5G, offer the potential to transform infrastructure monitoring into a real-time, automated process that can improve safety, efficiency, and citizen satisfaction [1].

Solution and Key Findings

The paper [1] introduces AIoT-CitySense, a flexible framework that integrates IoT-enabled mobile assets (such as waste collection trucks) with AI-based analytics for city-scale sensing. A tailored adaptation, called Mobile IoT-Roadbot, was piloted in Melbourne, Australia, using 11 waste trucks fitted with stereo cameras, GNSS, and 5G connectivity. These vehicles acted as roaming sensor nodes, collecting visual and location data during routine operations [1]. Data was processed via AWS cloud services and custom deep learning (DL) models, which detected damaged road signs, dumped rubbish, and vandalized bus shelters. The pilot demonstrated:

1. 85% more roadside issues detected compared to citizen reporting, often 3-4 days earlier.
2. High accuracy: DL models achieved 97% F1-score for dumped rubbish, 96% for bus shelters, and 71% for damaged road signs [1].
3. Scalability: Trucks processed ~5,000 images/day each, at a modest cloud cost (~AUD $200/day across all trucks). These results validate the framework's effectiveness and real-world feasibility for proactive maintenance

Implications / Ethical Considerations

The deployment of AIoT-CitySense highlights how smart city technologies can reduce reliance on citizen labor, speed up maintenance, and improve urban liveability. It supports a shift from reactive to proactive governance, increasing public trust in municipal services. However, ethical issues arise around data privacy and surveillance. Since roadside cameras inevitably capture images of people, vehicles, or private property, anonymisation and filtering are critical. Transparency in how data are collected, stored, and used to ensure public trust. Additionally, councils must balance automation with equitable access, ensuring that benefits extend beyond well-resourced urban areas to rural and marginalised communities.

Limitations / Future Works

While promising, the system faces challenges:
Data limitations: Detection of damaged road signs performed less accurately due to limited training datasets (~10% damaged sign images). Expanding datasets will improve reliability.

1. Technical constraints: Network coverage and weather conditions caused 10-20% daily data loss. Edge deployment of DL models is suggested to reduce reliance on cloud bandwidth and lower costs.
2. Scope: Current focus is on three roadside issues. Future work should expand to pothole detection, road surface monitoring, and line marking analysis.
3. Scalability: Although effective in one municipality, broader deployment requires consideration of infrastructure costs, interoperability with different vehicle fleets, and standardization across regions.

Transforming Cybersecurity into Critical Energy Infrastructure: A Study on the effectiveness of Artificial Intelligence

Motivation: This paper [2] explores the applications of AI to defend critical energy infrastructure against cyber threats. Cyber attacks against the energy sector are becoming increasingly sophisticated. This threatens the continuity of energy services, risking significant economic and social instability.

Solutions/key findings: The paper [2] evaluates three AI solutions against traditional signature-based cybersecurity systems. Collectively, the AI models achieved 94.7% threat detection accuracy, versus 74.9% for traditional systems. The AI models also reduced incident response times by over 70% and reduced the false positive rate from 12% to 4%. The three AI solutions that were evaluated were:

1. Deep learning for ransomware detections: A deep learning model was applied to energy control systems. By monitoring real-time network operations the model could identify activity patterns indicating a ransomware attack. This solution achieved a 98% detection rate of ransomware threats.
2. AI algorithms to secure operational hardware: Machine learning algorithms (Random Forest and Support Vector Machine) were used to analyse behavioural patterns within the operational hardware that manages energy facilities. The algorithms performed anomaly detection, leading to a 90% reduction in undetected security incidents compared to traditional methods.
3. Predictive analytics for failure prevention: AI was used to analyse large volumes of operational data from the power distribution network to predict disruptions before they occurred. By identifying patterns indicating potential system failures or cyber attacks, the model achieved 95% precision in failure prediction and reduced system outages by 70%.

Implications: These results confirm that integrating AI can help to shift from a reactive to proactive cybersecurity approach in critical energy infrastructure. All three solutions use AI to detect anomalies or suspicious patterns at the beginning of an attack, allowing security operators to intervene earlier in the threat cycle. Operationally, the reduction in false positives also addresses the critical issue of “alert fatigue”, allowing human security teams to allocate their time and financial resources efficiently. A 70% reduction in incident response time shows that AI can reduce the time between detection and mitigation. Further supporting this, the paper explored the potential for threat mitigation using AI, such as automated system isolation, account deactivation and security patching.

Limitations/Future work:

1. **Adversarial Attacks:** The paper does not address adversarial attacks that specifically target AI models. Developing AI systems that are resilient to attack is an important area for future research.
2. **Explainability:** The paper discussed the issue of explainability, as many NN operate as “black box” algorithms. This lack of transparency limits security operators’ ability to understand why a threat was flagged, reducing trust in AI-driven decisions.
3. **Automation:** Although automation offers clear benefits, the use of AI to initiate mitigations in critical energy infrastructure raises concerns about fail-safe mechanisms. Errors could create safety hazards or economic disruptions.
4. **Deployment:** There are significant implementation costs and integration challenges when deploying AI in legacy energy infrastructure. Addressing these concerns is important to ensure any AI solutions proposed are feasible.

Monitoring and detecting faults in wastewater treatment plants using deep learning

Motivation:
In this paper the researchers developed an automated system to detect faults in wastewater treatment plant sensors, focusing on ammonia sensors in oxidation tanks. Traditional methods struggle to identify complex, time-based sensor failures, which can cause environmental harm and energy waste. To address this, the team used Long Short-Term Memory (LSTM) networks, a type of deep learning model that captures temporal patterns by retaining relevant past information while ignoring irrelevant data. This is to catch Sensor faults, which can lead to untreated discharge, strain the electrical grid, and disrupt 24/7 monitoring.

Solution
Data Collection Data from 12 sensors over a year (438,181 readings each) at a Italian treatment plant [3], including chemical (ammonia, oxygen, nitrates) and operational measurements (blower frequencies, temperatures).
LSTM Architecture: They built a multi-layered neural network with four LSTM layers, each containing 60 processing units [3]. The network processes 60-minute windows of data, automatically learning which temporal patterns indicate normal versus faulty behaviour. Each LSTM layer includes gates that control what information to remember, forget, or pass forward.
Data Preprocessing: Missing values were filled with the last known readings, statistical features from ammonia (mean, variance, etc.) were engineered, normalised for training.
Window-Based Analysis: Rather than analysing individual sensor readings, the system examines sequences of data over hour periods to detect "collective faults" - problematic patterns that emerge across multiple readings rather than single anomalous points.

Key findings
Baseline Comparisons: The LSTM approach was tested against three established methods:
- **Variance Analysis:** Simple statistical threshold detection
- **ARIMA Time Series:** Classical forecasting model that failed to detect collective faults
- **PCA-SVM:** Traditional machine learning combining feature reduction with classification

Performance Superiority: The LSTM method achieved 96.5% accuracy compared to 93% for traditional approaches [3].

Method / Metric	Variance Analysis	PCA-SVM	LSTM
Accuracy (%)	93.3	93.0	96.5
Recall (%)	80.9	94.1	93.9
Precision (%)	70.7	65.9	81.7
F1-Score (%)	75.4	77.5	87.4

The 11 percentage point improvement in precision from LSTM means less false alarms, reducing unnecessary maintenance costs and operator fatigue from alert overload.

Implications / Ethical Considerations:
- Environmental Protection: Prevents untreated discharge contaminating water bodies
- Public Health: Reduces waterborne disease risk from undetected failures
- Operational Equity: Reduces false alarms
- Data Privacy: Infrastructure monitoring raises surveillance concerns
- Technology Dependence: Over-reliance may reduce human expertise

Limitations / Future Work:
- Single Plant Validation: Only tested on one facility - generalizability unknown
- Limited Fault Types: Focused on ammonia sensors only
- Real-time Implementation: 60-minute windows could be too slow for rapid faults
- Future Directions: Multi-plant, faster prediction, additional sensors, lightweight models

Reconfigurable cyber-physical system for critical infrastructure protection in smart cities via smart video-surveillance

Motivation
Traditional CCTV surveillance depends on human operators, leading to high costs, errors, and limited scalability. Smart video-surveillance with distributed CPS and edge devices overcomes these issues by performing local analytics, reducing traffic, latency, and costs, while enhancing critical infrastructure protection in smart cities [4].

Solution and key findings
1. Edge Node Processing:
Local processing is performed on high-performance, low-power System-on-Chip (SoC) devices, assisting in achieving real-time performance, enabling efficient video stream management.
The edge nodes are connected to a surveillance device and uses subtraction (Mixture of Gaussians) to identify moving objects. These regions are then analysed by an efficient Convolutional Neural Network (CNN), specifically an optimised version of MobileNetV2, to detect humans in real-time [4].
2. Cloud Server Processing:
The cloud server aggregates processed data and video streams from multiple edge nodes to perform advanced analytics. It tracks individuals across different camera views using Kalman filters and appearance features, identifies faces through a pipeline with MTCNN and MobileFaceNet, and monitors secure perimeters by mapping all camera feeds to a common real-world coordinate system [4].

The System performance and efficiency
The system operates in real-time at 32-98 fps [4]. Its adaptive bandwidth management reduces resource consumption by 76% by defaulting to low resolution and switching to high resolution only during detections.

Human detection
Achieved an F1-score of 0.9646, nearly matching the best model while being 36x smaller (0.88MB), enabling highly efficient performance [4].

Facial recognition
Reached 99.5% accuracy on a standard dataset, with results showing image resolution is critical for classification [4].

Object tracking
Performed effectively in multi-camera setups, distributing tasks across edge nodes while maintaining real-time performance [4].

Implications/ethical considerations
Although the paper does not address ethics, key concerns arise. Automated tracking in public spaces threatens privacy and anonymity. Facial recognition may show demographic bias, risking false positives/negatives in critical infrastructure contexts. Finally, as a powerful surveillance tool, the system risks function creep beyond its intended purpose.

Limitations/future work
Environmental robustness
The paper evaluates metrics such as frame rate and bandwidth reduction, but real-world performance can decline due to untested environmental factors such as poor lighting, weather, or dirty camera lenses.
Scalability of re-identification
The system tracks individuals within a single camera view and automates identification, but the paper lacks evaluation of long-term tracking and complex re-identification scenarios.
Vulnerability to adversarial attacks
The system relies heavily on deep learning models such as CNNs, which are vulnerable to adversarial attacks - small changes to clothing or accessories can cause missed detections or misidentifications.

AI-enabled strategies for climate change adaptation: protecting communities, infrastructure and businesses from the impacts of climate change.

Motivation

As the effects of climate change continue to worsen, it is becoming increasingly common for infrastructure to be badly impacted by the effects of climate related weather events. In 2021, snowstorms in Texas crippled the electricity grid leaving 4.5million homes and businesses without power. Closer to home, the effects of Cyclone Gabrielle in 2023 saw critical road and transport infrastructure destroyed and strained agricultural food supply. [5] discusses how the increasing rate of these events necessitates the need for protections against the consequences of these events on critical national infrastructure and the role AI can play in these protections.

Solutions

The paper highlights two key areas where AI systems can be applied to protect critical infrastructure from the effects of climate change: predictive simulations of climate effects and early warning systems. Predictive simulations can be used to analyse how the effects of climate change, such as sea level rise, are likely to impact infrastructure within specific regions. The paper discusses an AI model, called Deltares, which is used to simulate the impacts of sea level rise in the Netherlands and the potential impact on infrastructure. The model ingests data about the nearby infrastructure, along with climate data to model the region and predict how critical infrastructure, such as drainage systems and roads, could be affected by sea level rise. The model is also capable of recommending different protections against these risks (e.g., sea walls) and modelling how this can reduce the impact of climate events.

AI can also be used to predict both the likelihood and severity of flooding in real time. The paper discusses a model, named FloodNet, as an example of this application. The model uses deep learning to analyse images of relevant bodies of water along with metrics on factors such as water level and flow rate to predict the likelihood and severity of flooding taking place. These same techniques can also be applied to other events, such as wildfires, to extend the application of these tools. Having these early warning systems allows cities to prepare and better manage infrastructure (such as power and water) to avoid widespread outages.

Limitations

Both models are limited by the quality of available data. For the models to perform optimally, high-quality region-specific data is needed but often may not be readily available. Equally, these models aren't yet widely used so it is unclear how the outcomes of these models will compare to infrastructure protection work done by humans.

Ethical Considerations

For both applications, model explainability is a key ethical concern. The risks of false negatives can have detrimental consequences, including loss of life, so it's crucial that humans provide oversight into the models' application and can understand how outcomes are determined. Another ethical concern is the energy usage and carbon footprint of these AI systems. While the AI may be helping protect against the possible impacts of climate change, it is simultaneously contributing to the underlying issue, especially when models are powered by 'dirty' energy sources such as coal and gas.

[1]
M. Forkan et al., “AIoT-CitySense: AI and IoT-Driven City-Scale Sensing for Roadside Infrastructure Maintenance,” Data Science and Engineering, Dec. 2023, doi: https://doi.org/10.1007/s41019-023-00236-5.

[2]
J. Govea, W. Gaibor-Naranjo, and W. Villegas-Ch, “Transforming Cybersecurity into Critical Energy Infrastructure: A Study on the Effectiveness of Artificial Intelligence,” Systems, vol. 12, no. 5, p. 165, 2024. doi: 10.3390/systems12050165.

[3]
B. Mamandipoor, M. Majd, S. Sheikhalishahi, C. Modena, and V. Osmani, "Monitoring and detecting faults in wastewater treatment plants using deep learning," Environmental Monitoring and Assessment, vol. 192, no. 148, Jan. 2020, doi: 10.1007/s10661-020-8064-1.

[4]
J. Isern, F. Barranco, D. Deniz, J. Lesonen, J. Hannuksela, and R. R. Carrillo, “Reconfigurable cyber-physical system for critical infrastructure protection in smart cities via smart video-surveillance,” Pattern Recognition Letters, vol. 140, pp. 303–309, Dec. 2020, doi: https://doi.org/10.1016/j.patrec.2020.11.004.

[5]
H. Jain, R. Dhupper, A. Shrivastava, et al., "AI-enabled strategies for climate change adaptation: protecting communities, infrastructure, and businesses from the impacts of climate change," Comput. Urban Sci., vol. 3, p. 25, 2023, doi: 10.1007/s43762-023-00100-2.