I.J. Image, Graphics and Signal Processing, **2026, 5, 126-184**
Published Online on October 8, 2026 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2026.05.08

Lyapunov-Stable Multispectral Edge Deep Learning for Critical Injury Triage and Resource Optimisation in Emergency Response

Serhii Vladov
Kharkiv National University of Internal Affairs, Kharkiv, 61080, Ukraine
Yuriy Fedkovych Chernivtsi National University, Chernivtsi, 58012, Ukraine
E-mail: serhii.vladov@univd.edu.ua
ORCID iD: https://orcid.org/0000-0001-8009-5254

Zhengbing Hu
School of Management, Huazhong University of Science and Technology, Wuhan, China
E-mail: drzbhu@gmail.com
ORCID iD: https://orcid.org/0000-0002-6140-3351

Dmytro Uhryn*
Yuriy Fedkovych Chernivtsi National University, Chernivtsi, 58012, Ukraine
E-mail: d.ugryn@chnu.edu.ua
ORCID iD: https://orcid.org/0000-0003-4858-4511
*Corresponding author

Yuriy Ushenko
Yuriy Fedkovych Chernivtsi National University, Chernivtsi, 58012, Ukraine
E-mail: y.ushenko@chnu.edu.ua
ORCID iD: https://orcid.org/0000-0003-1767-1882

Victoria Vysotska
Lviv Polytechnic National University, Lviv, 79013, Ukraine
Kharkiv National University of Internal Affairs, Kharkiv, 61080, Ukraine
E-mail: Victoria.A.Vysotska@lpnu.ua
ORCID iD: https://orcid.org/0000-0001-6417-3689

Vladyslav Doroshenko
Kharkiv National University of Internal Affairs, Kharkiv, 61080, Ukraine
E-mail: vl.doroshenko@ukr.net
ORCID iD: https://orcid.org/0000-0002-8074-0820

Daniil Zinchenko
Kharkiv National University of Internal Affairs, Kharkiv, 61080, Ukraine
E-mail: danil3008199@gmail.com
ORCID iD: https://orcid.org/0000-0002-8089-0511

Received: March 18, 2026; Revised: May 5, 2026; Accepted: June 10, 2026; Published: October 8, 2026

This work is open access and licensed under the Creative Commons CC BY 4.0 License.

Abstract: Traumatic injuries in emergency response scenarios require rapid and reliable assessment methods capable of supporting medical prioritisation and efficient resource management. This paper presents a multispectral edge deep learning system designed for early injury detection, assistance prioritisation, and financial-resource decision support in emergency response environments. The proposed approach integrates RGB, thermal, near-infrared (NIR), depth, and contextual information within an edge-based analytical framework that combines multispectral data fusion, adaptive attention-based weighting of sensor modalities, convolutional neural networks for feature extraction and lesion segmentation, severity classification, triage index estimation, and resource allocation optimisation. The developed mathematical model describes the interaction among sensory processing, local inference, decision-making, and feedback-based model updating, and system stability is analysed using a Lyapunov-based approach under specified modelling assumptions. The system was evaluated on a simulated dataset containing N = 1000 multispectral observations, including 150 test samples for final performance assessment. The obtained results demonstrated promising performance on simulated data, achieving *Accuracy* = 0.912 ±0.015, $Recall_{crit}$ = 0.941 ±0.021, $Precision_{crit}$ = 0.876 ±0.024, and F1-score = 0.908 ±0.018. The average edge inference delay was 425 ms, satisfying the considered real-time processing threshold of 500 ms. Compared with the baseline emergency response scenario without multispectral edge-based decision support, the proposed approach reduced triage decision errors by 18.6% and unproductive resource utilisation costs by 16.4%. The contribution of this research is the formulation of an integrated multispectral edge deep learning framework that combines injury detection, automated triage assessment, resource optimisation, and financial risk considerations within a unified decision-support architecture. The results indicate the potential of the proposed approach to improve emergency response processes. However, further validation on real-world clinical and operational datasets is required to confirm its practical applicability.

Index Terms: Multispectral System, Edge Computing, Deep Learning, Lesion Detection, Triage, Emergency Response, Financial Security, Resource Optimisation, Simulated Dataset, Stability Analysis.

1. Introduction

Traumatic injuries remain a key cause of premature mortality, disability and overload of emergency response systems. According to the World Health Organisation [1], injuries and violent injuries cause approximately 4.4 million deaths worldwide each year, accounting for almost 8% of all deaths (Fig. 1). In addition, they are responsible for approximately 10% of years of life lived with disability. The situation is particularly critical among young age groups, as three out of five leading causes of death for people aged 5…29 years are related to injuries.

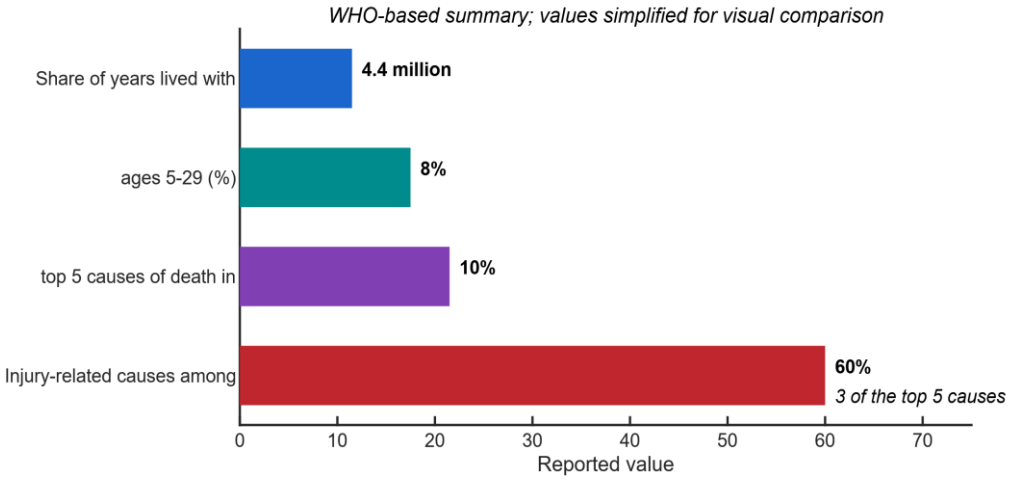

Fig. 1. Global burden of traumatic injuries.

A separate risk group comprises road accidents, mass incidents, industrial accidents, combat, and artificial injuries, in which the timing of the victim's condition's initial detection directly affects the medical triage outcome. It is noted that the World Health Organisation estimates the annual number of deaths from road accidents at approximately 1.19 million people, while another 20…50 million people receive non-fatal injuries, some of which lead to long-term disability or death (Fig. 2) [2].

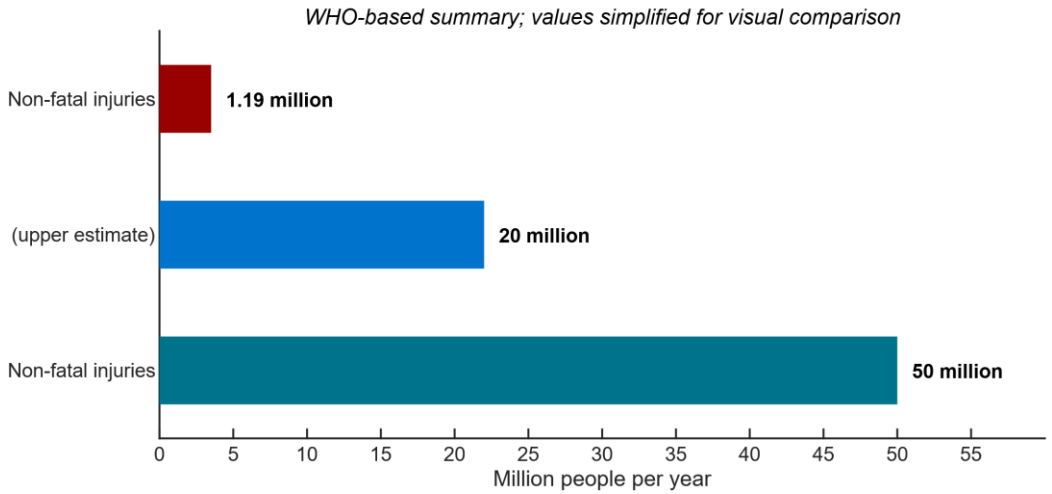

Fig. 2. Road traffic injuries and the need for rapid primary detection.

In such conditions, traditional visual assessment of the victim's condition, especially in limited visibility, staff shortages, or a large number of victims, does not provide sufficient speed, objectivity, or reproducibility of decisions. It is worth noting that a promising direction for solving this problem is the development of a multispectral, edge-based deep learning system that integrates visible, infrared, thermal, and other sensory channels to enable early detection of signs of human damage. In particular, the study [3] demonstrates the potential of multispectral computer vision for real-time bleeding detection, in particular through the combination of thermal, near-infrared, or depth images, as well as the use of deep learning models for the analysis of complex visual patterns (Fig. 3). Therefore, transferring calculations to the edge level is of fundamental importance, as processing data directly at the source reduces latency, dependence on network connectivity, and the risk of transmitting sensitive medical information to cloud infrastructure [4].

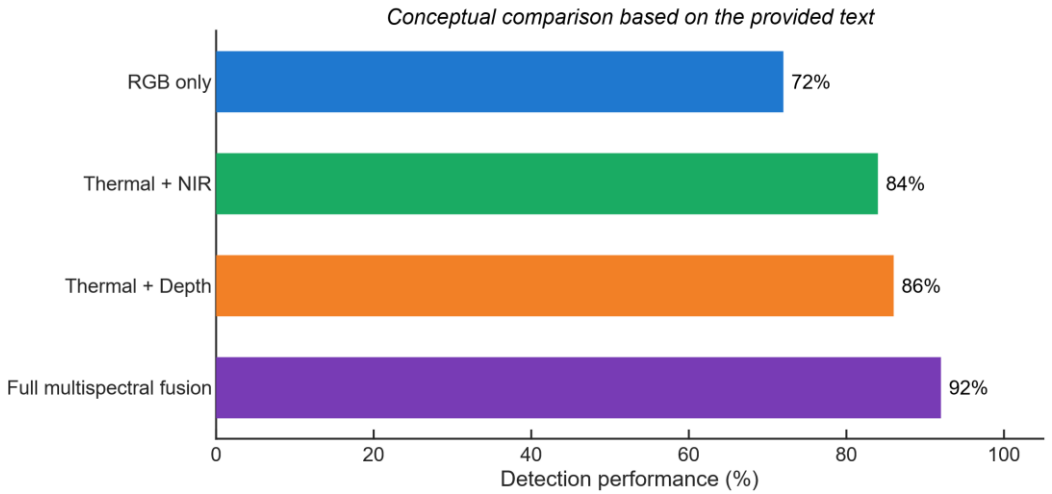

Fig. 3. Detection performance of multispectral edge deep learning.

The problem outlined here has not only a medical dimension but also financial and security dimensions. According to the World Bank, the total economic losses from road traffic deaths and injuries in low- and middle-income countries amount to approximately 2…6% of gross domestic product (GDP), which includes medical costs, lost productivity, property damage, and burden on public services (Fig. 4) [5].

Therefore, the problem is to justify and develop such a multispectral edge deep learning system that provides early identification of critical injuries, supports prioritisation of care in time and resource shortage conditions, and reduces financial losses of emergency response systems through more accurate triage, more rational teams' dispatch, and a reduction in the number of erroneous or late decisions.

The contributions of this paper include:

1. Proposing a multispectral edge deep learning architecture that integrates RGB, thermal, near-infrared (NIR), depth, and contextual information for early injury detection, automated assistance prioritisation, and resource-aware emergency response decision support.

2. Establishing a mathematical model of the developed system that formalises the processes of multispectral data synchronisation, quality assessment, adaptive feature fusion, injury severity estimation, triage index formation, resource allocation, and financial-risk optimisation within a unified framework.

3. Developing a Lyapunov-based stability analysis approach for evaluating the dynamic behaviour of the proposed edge deep learning system under modelling assumptions related to sensor disturbances, inference errors, and decision-making processes.

4. Conducting an experimental evaluation on simulated multispectral data to analyse the system's detection accuracy, inference latency, triage performance, and potential resource optimisation capabilities, while identifying directions for further validation using real-world emergency response datasets.

5.

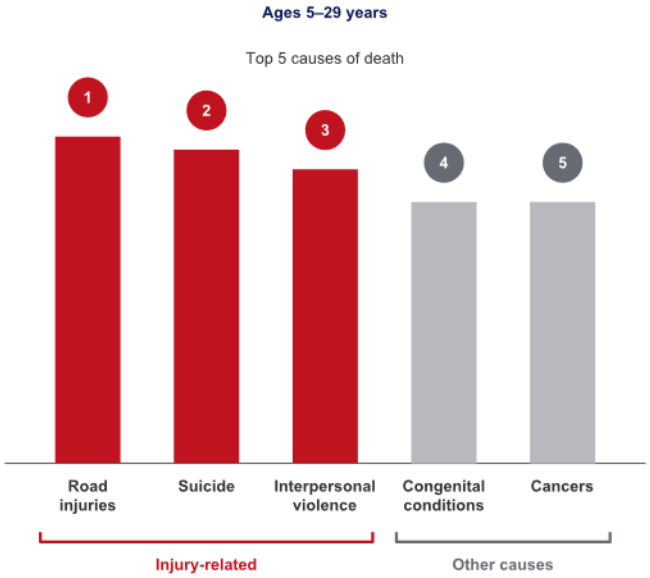

Fig. 4. Top 5 causes of death (ages 5–29 years).

2. Related Works

Classic triage protocols, such as START, SALT, JumpSTART, and ESI, form the foundation of emergency response by prioritising care based on a limited set of clinical features. In particular, for START triage, the target time to assess a single casualty is less than 60 seconds, and decisions are based on criteria such as breathing, pulse, capillary refill, bleeding, and ability to follow commands [6]. As noted in [7], for a model incident with 100 casualties, even an ideal rate of 60 seconds per person translates into about 100 minutes of work for one triage worker or 25 minutes for four professionals, which is a critical time frame in the event of rapid deterioration of the condition of some people. These protocols' main limitation is that they work with "coarse" severity categories, and the accuracy in [7] varies: for example, for individual algorithms in the pediatric scenario, rates of correct decisions from 67.8% to 89.3% were recorded, and no algorithm combined maximum accuracy, speed, and consistency between specialists.

Hospital and dispatch information systems improve the manageability of emergency response by enabling call registration, crew routing, time of arrival, electronic victim cards, and advanced notification to the emergency department. For example, ESI, CTAS, ATS, and Manchester Triage System operate as structured urgency scales, and Manchester Triage System uses 52 flowcharts based on patient complaints and severity discriminators. According to a model estimate [8], if it takes 90…120 seconds to enter primary data on one victim, then for 50 people the administrative workload is 75…100 minutes of operator time, not including re-clarification of the condition. Therefore, their disadvantage is that they do not form the primary sensory recognition of lesions at the scene, but only structure the data already entered by the person.

Wearable sensor devices continuously monitor physiological parameters, including heart rate, SpO_2, temperature, electrocardiogram (ECG), motor activity, and stress indicators. National Centre for Biotechnology Information (NCBI) materials [9] indicate that such devices can remotely monitor military personnel and first responders and support the assessment of heat stress, hypothermia, altitude sickness, fatigue, trauma, and sepsis. In a model configuration [9] for emergency response, a minimum set of five channels (HR, SpO_2, temperature, accelerometry, and ECG) can generate a risk index with updates every 5…10 seconds. The limitation is that wearables are only effective when the sensor is already attached to the person, is charged, maintains stable skin contact, and uses a compatible data format, whereas victims of a road accident or mass incident are most often without the appropriate equipment.

RGB video analytics and computer vision algorithms are used to detect people, falls, immobility, body posture, movement, and victim crowding in the event area. Their features include low camera costs, easy scalability, and

compatibility with existing video surveillance infrastructure. For the model system, according to [10], an accuracy of 78…82% can be assumed in a controlled daytime environment, but under conditions of low light, smoke, partial body overlap, or a complex background, this indicator can decrease to 55…65%. At the same time, studies on vision-based fall detection [11, 12] emphasise that these systems have significant potential but remain dependent on image quality, the availability of training data, and computing resources, as well as on robustness to real-world scenes with noise and overlap.

Infrared and thermal imaging diagnostics focus on registering the body temperature profile, local thermal anomalies, and areas of hypothermia, overheating, or thermal imbalance. In medical thermography, normal body temperature typically ranges from 36.1 to 37.2 ℃, and diagnostic equipment for medical measurements requires at least 0.1 ℃ accuracy and thermal resolution [13]. According to a model estimate [14], the thermal imaging channel can increase the detection of physiologically unstable individuals from 78% for an RGB camera to 84…87% in the dark or with limited visibility. Therefore, the disadvantage is the high dependence of the results on air temperature, humidity, distance to the object, clothing material, viewing angle, and sensor calibration, which is why the thermal image does not always directly reflect the lesion's severity.

Multispectral systems combine RGB, near-infrared, thermal, depth (or NIR), and other channels; therefore, they analyse not only body shape but also temperature, spectral, and spatial features. The CVPR 2025 study [15] proposed a real-time bleeding detection system that integrates long-wave infrared and near-infrared imaging, applies deep learning to each modality, and matches detected zones across channels. According to [15], for the design system, it is advisable to take the model accuracy gain, in particular, RGB is 78%, thermal is 84%, depth is 86%, multispectral fusion is 92…94%, provided that the sensors are correctly synchronised and a sufficient training set is available. But the disadvantages remain the higher equipment cost, a 3…5 times increase in data amount compared to RGB video, the need for channel calibration, and the risk of accuracy loss when transferring the model from a laboratory environment to a real event.

Cloud-based AI and IoMT solutions use centralised servers to analyse data from sensors, cameras, and mobile devices, thereby offering advantages in scalable storage, model updates, and access to powerful GPU resources. At the same time, a review of Edge AI for the Internet of Medical Things highlights that cloud architectures in medical applications face latency, power consumption, bandwidth, and privacy risks, while edge computing brings analysis closer to the data source [16]. In model-based estimation [16, 17] for an emergency scene, cloud inference can incur 250…400 ms per frame, with the additional risk of connectivity loss, whereas an edge device with an optimised model can reduce latency to 60…120 ms. The edge architecture's limitations include shortages of CPU or GPU power, memory, energy, and thermal headroom, which require model compression, quantisation, frame-rate limiting, and careful selection of neural network architecture.

It is worth noting that financial security and emergency response resource optimisation systems focus on reducing unproductive trips, hospital overload, misrouting, and overuse of expensive resources. According to the World Health Organisation [2], road traffic accidents cause about 1.19 million deaths and 20…50 million non-fatal injuries each year, and their cost for most countries is approximately 3% of GDP. The World Bank estimates economic losses from road traffic injuries and mortality at 2…6% of GDP for low- and middle-income countries, which directly links the quality of early response to the financial sustainability of government and insurance systems [5]. According to the model scenario [15], if a multispectral edge system reduces the false prioritization's proportion by at least 15…20% and shortens the time for initial detection of critical conditions by 30…40%, it can reduce losses from irrational distribution of teams, repeated calls, and delayed hospitalization; however, current solutions still poorly integrate medical recognition, resource planning, and financial control into a single real-time system.

Thus, the existing approaches' systematisation (Table 1) shows that none of them, individually, provides simultaneous performance of three critical functions: early detection of human lesions, operational prioritisation of assistance, and reduction of financial losses in emergency response systems. That is why the development of a multispectral edge deep learning system is justified, as it combines multichannel sensor observations, local real-time data processing, and decision support to rationalise the use of medical and financial resources.

Table 1. Review of existing research approaches.

Existing approach	Main methodological features	Advantages	Main limitations	Need for the proposed multispectral edge system
Classical emergency triage protocols (START, SALT, ESI)	Rule-based assessment using physiological indicators and visual examination	Fast decision-making and practical clinical usability	Dependence on human expertise, limited objectivity; cannot perform automated lesion detection	Automated sensor-based assessment and reduction of subjective decisions
Hospital and dispatch information systems	Digital registration, routing, hospital communication, and electronic patient records	Improves organisational coordination	Do not provide initial injury recognition at the incident location	Integration of perception, diagnosis, and response coordination
Wearable medical sensor systems	Continuous monitoring of physiological parameters (HR, SpO_2, ECG, temperature, movement)	Provides real-time physiological information	Require prior attachment, maintenance, and stable sensor contact	Non-contact multispectral assessment without patient preparation

RGB-based computer vision systems	CNN-based detection of people, posture, falls, and visible injuries	Low-cost deployment using existing cameras	Reduced reliability under darkness, smoke, occlusion, and complex environments	A combination of visible and non-visible spectral channels
Thermal imaging approaches	Detection of temperature anomalies and physiological changes	Effective in low-light conditions and complementary to RGB	Sensitive to environment, distance, calibration, and external temperature	Adaptive fusion of thermal information with other modalities
Multimodal medical image fusion methods (2023–2025)	Transformer, attention, cross-modal learning, invertible networks, PET–MRI/MRI–CT fusion	Improved feature representation and diagnostic accuracy through complementary information integration	Mostly designed for offline medical imaging tasks; limited real-time deployment	Real-time multispectral fusion with edge inference capability
Multi-sensor and cross-modal learning approaches	Learning shared and modality-specific representations between different sensors	Better alignment and information transfer between modalities	Require large datasets and controlled acquisition conditions	Robust fusion of heterogeneous emergency-response sensors
Cloud-based medical AI systems	Centralised GPU-based inference and storage	High computational capacity and scalable model updates	Network dependency, latency, privacy risks	Local edge decision-making without critical cloud dependence
Edge artificial intelligence medical systems	Lightweight neural networks, model compression, efficient inference	Reduced latency and improved autonomy	Limited memory, energy, and computational resources	Development of an optimised multispectral edge architecture
Resource optimisation and decision-support systems	Mathematical optimisation of teams, routes, and costs	Improves allocation efficiency	Usually disconnected from image-based injury recognition	Unified model combining diagnosis, triage, logistics, and financial control

According to Table 1, existing approaches to detecting human injuries and organising emergency response are functionally fragmented, since each of them solves only a separate part of the problem:

- Classical triage protocols provide rapid clinical sorting but depend on the rescuer's subjective assessment.
- Dispatch and hospital information systems structure call data, but do not form the initial sensory detection of injuries.
- Wearable devices record physiological indicators, but require prior human equipment.
- RGB video analytics is available but loses accuracy in darkness, smoke, body overlap or complex background conditions.
- Thermal imaging diagnostics records temperature anomalies but depends on environmental parameters.
- Cloud-based artificial intelligence solutions have significant computing capabilities but critically depend on stable communication, low transmission delays, and privacy protections.

As a result, current solutions do not form a single technological circuit capable of combining multi-channel detection of signs of trauma, automated assessment of the condition’s criticality, prioritisation of care, and the rational distribution of medical and financial resources in real time at the incident scene.

A multispectral edge deep learning system development is necessary precisely because it integrates visible, infrared, thermal, and spatial features into a single analysis model, reduces dependence on a single sensor channel, transfers calculations directly to the data source, reduces decision-making time, and increases the system’s resilience in unstable communication, staff shortages, and a large number of victims' conditions. Therefore, the proposed system should combine medical accuracy, response efficiency, and financial security, within which to carry out early detection of critical conditions, reduce the delayed hospitalisation risk, missorting, duplication of visits, and irrational use of expensive resources, and therefore form the basis for an effective, evidence-based, and economically sustainable emergency response model.

It is worth noting that the proposed multispectral edge deep learning system's contribution to the studied area is an integrated approach to emergency response, in which early detection of human injuries is achieved by combining visible, infrared, thermal, and spatial features, without critical dependence on cloud infrastructure. Unlike fragmented solutions, such a system combines sensory diagnostics, automated prioritisation of assistance, and the financial and resource justification of decisions in a single technological circuit. The above deepens the scientific principles of edge-AI applications in emergency response systems and provides an applied basis for reducing response time, triage errors, and unproductive costs.

3. Materials and Methods

3.1. Development of a multispectral edge deep learning system’s architecture

Based on the above, a block diagram of the architecture of a multispectral edge deep learning system (Fig. 5) is proposed as a single system that combines sensory detection, local intelligent processing, decision support, and financial and resource control. Its key feature is that the initial recognition of the injured person’s condition is not carried out using a single data source, but based on a combination of RGB images, a thermal imaging channel, NIR or

IR data, a depth or 3D sensor, and contextual information about time, geolocation, and the type of event. The proposed multichannel structure increases the completeness of the input data, as different sensory modalities capture complementary features, including visible injuries, temperature anomalies, the body's spatial position, scene features, and response context.

The central element of the proposed system is the deep learning edge-analytical core, which performs sequential data processing without critical dependence on the cloud infrastructure. In the initial stage, synchronisation and merging of streams are necessary to coordinate heterogeneous sensor signals into a single spatiotemporal loop. Then, pre-processing and quality control are performed, including noise removal, signal normalisation, frame-suitability checks, and screening for incomplete or distorted data. This unit's peculiarity lies in preparing input information for further analysis, so that the deep learning model operates not on disparate sensor fragments but on a coordinated set of diagnostically significant features.

To provide a formal description of the interfaces between architectural modules, each sensor stream is represented as a timestamped data packet containing the acquisition time, modality identifier, spatial resolution parameters, and raw or preprocessed measurements. The RGB, thermal, NIR, depth, and contextual channels generate asynchronous streams $X_m(t)$, where the corresponding sensor frequency determines the temporal resolution of each channel. The synchronisation module converts these heterogeneous streams into a common temporal reference frame using timestamp alignment with a maximum allowable deviation of $\Delta t \leq 10$ ms for high-priority emergency perception tasks. The temporal offset between modalities is estimated via correlation-based matching of modality-specific features, using characteristic patterns such as body contours, thermal peaks, depth discontinuities, and spectral responses as reference signals. The synchronisation and stream-merging interface produces a unified multispectral tensor representation with dimensions determined by the spatial resolution and the number of modalities $X_{sync}(t) \in R^{H \times W \times C_m}$, where H and W represent the spatial dimensions after registration, and C_m denotes the number of aligned channels. Before entering the deep learning inference core, all modalities are transformed into a common numerical range using modality-specific normalisation procedures. RGB intensities are normalised to [0,1], thermal measurements are converted into calibrated temperature maps, NIR responses are scaled according to sensor dynamic range, and depth values are filtered using invalid-depth removal and spatial smoothing procedures.

The preprocessing and quality control module implements a multi-criteria reliability assessment mechanism. The quality coefficient for each modality is calculated using noise estimation, blur detection, missing-pixel ratio, and object-overlap analysis. Frames are considered diagnostically acceptable when the signal-to-noise ratio exceeds the predefined threshold, image sharpness remains above the minimum Laplacian variance level, and the percentage of invalid or missing measurements does not exceed 15%. If a modality fails the quality assessment procedure, the corresponding confidence weight is automatically reduced rather than completely removing the information source, allowing robust operation under partial sensor degradation.

After quality validation, the interface between preprocessing and feature extraction modules transfers the aligned multispectral tensor to parallel modality-specific neural encoders. Each encoder generates a latent feature representation $F_m \in R^{h \times w \times d_m}$, where h and w correspond to the reduced spatial dimensions after convolutional processing and d_m denotes the latent feature depth of modality m. The extracted representations are subsequently forwarded to the fusion module, where attention-based weighting determines each modality's contribution based on its diagnostic reliability and information content.

The edge analytical core operates on an optimised inference pipeline comprising feature extraction, cross-modal fusion, lesion probability estimation, and decision generation. The interface between the neural inference module and the decision-support layer is implemented through a compact feature vector $z(t) \in R^D$, where D represents the dimensionality of the fused feature space. This vector contains the encoded information required for injury localisation, severity estimation, triage classification, and subsequent resource optimisation. Communication between edge inference and decision-support modules is handled via a lightweight local protocol, avoiding the transmission of raw sensor data and reducing latency, bandwidth consumption, and privacy risks.

The multispectral feature extraction unit serves as the basis for early detection of human lesions, since at this stage, parameters characterising the shape, temperature, spectral reflectance, spatial configuration of the body, and potential signs of physiological instability are extracted from different channels. Further application of the deep learning model aims to detect lesions and assess their severity, thereby moving the system from passive observation to automated interpretation of the victim's condition. An important feature is a results-interpretation and plausibility-check unit, which reduces the risk of false positives, increases the solution's transparency, and provides a basis for using the results in responsible emergency-response scenarios.

Local real-time processing is highlighted as a key architectural feature of the developed system. Moving inference to the edge reduces the delay between receiving sensor data and forming a decision, reduces dependence on unstable communication, and limits the transmission of sensitive information beyond the scene. It is particularly important for emergency response, where even a short delay in detecting a critical condition can lead to incorrect prioritisation of assistance, irrational involvement of teams, or time loss for the victim's reassessment.

The block diagram's "right unit" reflects the system's functional result: decision support across four interrelated

areas, including early detection of human lesions, prioritisation of care, response coordination, and financial security. Early detection of human lesions provides the basis for further triage, while prioritisation of care translates the analysis results into triage categories: critical, high, medium, or low. Response coordination links the medical conclusion to practical actions, in particular the team, route, and hospital choice. The financial security block specifies the system's economic dimension, since cost control and the prevention of irrational resource use are directly related to the initial detection and the accuracy of prioritisation.

The block diagram's "lower outline" demonstrates the system's integration with the organisational infrastructure of emergency response. The rescuer or dispatcher interface provides a practical presentation of the analysis results in a form suitable for operational use. The event log and decision audit form an evidentiary basis for further verification of the actions' correctness and for the effectiveness analysis of the financial control response. Integration with EMS, a hospital, or a command centre creates a single information space between the incident, medical infrastructure, and the management-level scene.

A special feature of the proposed multispectral edge deep learning system's architecture is a feedback loop for the model's ongoing training. It records the actual response, confirmed clinical data, prioritisation errors, time indicators, and the resource consequences of the decisions made. Thanks to this, the system is not limited to a single inference; it gradually accumulates verified information, thereby increasing the model's robustness to various event conditions, lesion types, and sensory interference. The protected cloud storage is an auxiliary scheme; namely, it is used for analytics, archiving, and training procedures, but is not critical to the primary solution, which emphasises the autonomy and reliability of the edge component.

Thus, the proposed multispectral edge deep learning system's architecture reflects not only the technical sequence of data processing but also the comprehensive logic of a comprehensive emergency response system. Its scientific novelty lies in multispectral sensory analysis, local edge inference, automated triage assessment, response coordination, and the integration of financial and security controls within a single architecture, thereby eliminating the fragmentation of existing solutions and laying the foundation for an accurate, operational, and economically justified response to human-injury events.

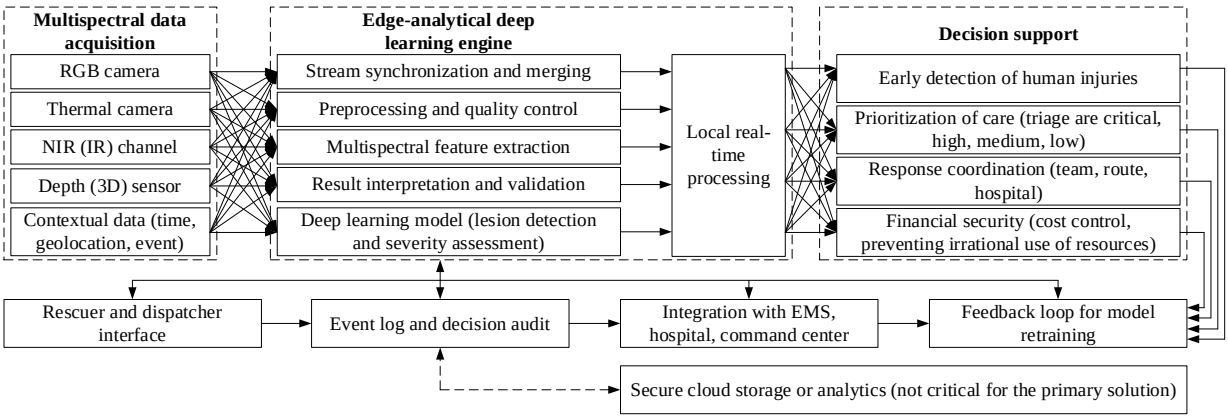

Fig. 5. Proposed block diagram of a multispectral edge deep learning system's architecture

3.2. Development of a multispectral edge deep learning system's mathematical model

It is assumed that the set of victims describes an emergency response event:

$$P = (p_1, p_2, ...p_n), \tag{1}$$

and each affected p_i is observed through multiple sensory modalities:

$$M = \{RGB, TH, NIR, DEP, CNT\}, \tag{2}$$

where *RGB* is the visible camera, *TH* is the thermal imaging channel, *NIR* is the near-infrared channel, *DEP* is the depth or 3D sensor, and *CNT* is the contextual data about time, geolocation, and event type. A data stream is generated for each sensor channel:

$$X_i^{(m)}(t), m \in M, \tag{3}$$

where t is the time instant, and i is the index of the affected person. The total input information about the person's state is given as a vectorised set of streams:

$$X_i(t) = \left\{ X_i^{(RGB)}, X_i^{(TH)}, X_i^{(NIR)}, X_i^{(DEP)}, X_i^{(CNT)} \right\}. \quad (4)$$

The obtained multispectral representation serves as the initial input layer of the developed model and is further transformed through temporal and spatial alignment procedures to ensure consistency across heterogeneous sensor streams.

Therefore, it is worth noting that the problem statement's proposed peculiarity is that the human condition is not identified by a single image or physiological parameter but is described by a multidimensional sensory representation, in which each modality reflects a separate class of damage signs.

Before submitting data to the edge-analytical core, it is necessary to synchronise the sensor streams, as different sensors have different frame rates, latencies, spatial positions, and viewing angles. It is assumed that τ_m is the time shift of the *m*-th channel relative to the base *RGB* channel. Then the synchronised stream has the form:

$$\tilde{X}_i^{(m)}(t) = A_m\left(X_i^{(m)}\left(t + \hat{\tau}_m\right), K_m, R_m, T_m \right), \quad (5)$$

where $A_m(\bullet)$ is the spatiotemporal alignment operator, K_m is the matrix of internal sensor parameters, R_m and T_m are the rotation and displacement parameters relative to the base coordinate system. The temporal shift is estimated by minimising the discrepancy between the two channels' features:

$$\hat{\tau}_m = \arg\min_{\tau} \sum_{t=1}^{T} \left\| X_i^{(m)}(t+\tau) - \psi_{RGB} X_i^{(RGB)}(t) \right\|_2^2, \quad (6)$$

where $\psi_m(\bullet)$ is a function of the reference features' pre-extraction, such as contours, temperature maxima, silhouette or key points of the body. It is worth emphasising the importance of this stage, since a synchronisation error can lead to temperature, spatial, and visual features being incorrectly associated with the same body area.

After estimating the temporal and spatial correspondence between modalities, the synchronised data streams are transferred to the quality assessment module for reliability evaluation prior to further processing.

It is advisable to assign a separate coefficient to each sensor channel's quality, since the data can be distorted by noise, smoke, low light, body overlap, or frame loss. Therefore, for the *m*-th channel, a quality indicator is introduced:

$$q_i^{(m)}(t) = \exp\left(-\alpha_1 \cdot n_i^{(m)} - \alpha_2 \cdot b_i^{(m)} - \alpha_3 \cdot o_i^{(m)} - \alpha_4 \cdot d_i^{(m)} \right), \quad (7)$$

where $n_i^{(m)}$ is the noise level, $b_i^{(m)}$ is the blur measure, $o_i^{(m)}$ is the degree of overlap of the object, $d_i^{(m)}$ is the calibration deviation, and α_1, α_2, α_3, and α_4 are the weighting factors of the defects' influence. Since all components are included with a "minus" sign, an increase in noise, blur, or overlap reduces the $q_i^{(m)}(t)$ value. The integral quality of multispectral observation is defined as:

$$Q_i^{(m)} = \sum_{m \in M} \omega_m \cdot q_i^{(m)}(t), \sum_{m \in M} \omega_m = 1, \omega_m \geq 0. \quad (8)$$

The integrated quality coefficient provides adaptive control over the contribution of each modality, allowing the model to continue processing even when individual sensor channels exhibit distortion.

Thus, the system does not reject the entire case due to damage to a single channel but instead reduces its contribution to the overall solution. After synchronisation and quality control, each channel is passed through its own feature extractor:

$$h_i^{(m)}(t) = F_m\left(\tilde{X}_i^{(m)}(t); \theta_m \right), \quad (9)$$

where $F_m(\bullet)$ is the neural network unit for the corresponding modality, θ_m is this unit's parameters, and $h_i^{(m)}(t) \in ¡^{d_m}$ is the feature vector. The extracted modality-specific feature vectors represent the system's intermediate analytical layer and are subsequently combined into a unified multispectral feature space. For the *RGB* channel, these can be contours, color changes, blood stains, body posture, for the thermal channel these can be temperature anomalies, for the *NIR* (or *IR*) these can be spectral features of tissues or liquids, for the depth channel these can be body position, tilt, immobility, distance to the object, for the context these can be time, coordinates, event type, number of victims. The total feature space is formed through the fusion operator:

$$z_i(t) = \Phi\left(h_i^{(RGB)}(t), h_i^{(TH)}(t), h_i^{(NIR)}(t), h_i^{(DEP)}(t), h_i^{(CNT)}(t) \right). \quad (10)$$

The fused feature representation serves as the basis for adaptive modality weighting, in which the attention

mechanism determines each sensory channel's contribution based on its diagnostic importance.

It is worth noting that the most reasonable approach is not a simple vector combination but an adaptive weighting of channels based on their informativeness and quality. To this end, an attention mechanism between modalities is introduced. The assessment of the *m*-th channel's importance is given as

$$e_i^{(m)}(t) = \upsilon^{\mathrm{T}} \cdot \tanh\left(W_h \cdot h_i^{(m)}(t) + W_q \cdot q_i^{(m)}(t) + b\right), \tag{11}$$

where W_h, W_q, v, and b are the model parameters.

To explicitly define the attention mechanism used for adaptive weighting of multispectral modalities, feature interactions between sensor branches are implemented via a cross-modal self-attention procedure. For each modality *m*, the extracted feature representation is transformed into query, key, and value embeddings using learnable projection matrices. The query vector represents the current modality characteristics to be weighted, while the key and value vectors describe the available information content across all sensory channels. The attention similarity between the *m*-th modality and the *n*-th modality is calculated using the scaled dot-product operation. It allows the model to estimate the contribution of each modality based on the degree of correspondence between their latent feature representations. The obtained attention coefficients are normalised using the softmax function so that the total contribution of all modalities remains equal to one. The resulting attention-weighted representation is obtained by aggregating the value vectors of all modalities according to the calculated attention coefficients. In this way, the fusion process dynamically increases the influence of informative channels, such as a high-quality thermal signal in low-visibility conditions, while reducing the contribution of unreliable channels affected by noise, occlusion, calibration errors, or missing data. The attention mechanism can therefore be interpreted as a reliability-aware feature selection process, in which the contribution of each sensory branch is determined not only by its individual characteristics but also by its relationships with other available modalities. Thus, $Q_m = F_m \cdot W_Q$, $K_m = F_m \cdot W_K$, $V_m = F_m \cdot W_V$, $a_{mn} = \mathrm{softmax}\left(\frac{Q_m \cdot K_n^{\mathrm{T}}}{\sqrt{d_k}}\right)$, $\mu_m = \sum_{n=1}^{M} a_{mn} \cdot V_n$, where Q_m, K_m, and V_m are the query, key, and value representations of the *m*-th modality, d_k is the dimensionality of the key space, a_{mn} is the attention coefficient between modalities, and μ_m is the updated feature representation following cross-modal interaction.

The normalised channel contribution is defined as:

$$\lambda_i^{(m)}(t) = \frac{\exp\left(e_i^{(m)}(t)\right)}{\sum_{r \in M} \exp\left(e_i^{(r)}(t)\right)}, \tag{12}$$

It follows that:

$$\sum_{m \in M} \lambda_i^{(m)}(t) = 1, 0 \le \lambda_i^{(m)}(t) \le 1. \tag{13}$$

Then the integrated vector of multispectral features takes the form:

$$z_i(t) = \sum_{m \in M} \lambda_i^{(m)}(t) \cdot h_i^{(m)}(t), \tag{14}$$

which reflects the key feature of the developed system: a channel with high quality and high diagnostic informativeness receives greater weight, while a channel with noise or overlap receives less weight.

The resulting weighted feature vector is then used as input to the lesion detection module, which performs spatial identification and severity analysis of potential injuries.

The early lesion detection problem is presented as a binary segmentation and classification problem. For this aim, it is assumed that Ω_i is the victim's body region in the image, and $u \in \Omega_i$ is a spatial element of this region. The probability of a lesion's presence at the *u*-th point is defined as

$$M_i(u,t) = \sigma\left(g_m\left(z_i(u,t)\right); \theta_m\right), \tag{15}$$

where $\sigma(\bullet)$ is the sigmoid function, $g_M(\bullet)$ is the segmentation submodel, and θ_M is its parameters. The potential lesion area is defined as:

$$A_i(u,t) = \sum_{u \in \Omega_i} \mathrm{I}\left(M_i(u,t) > \varepsilon_M\right) \cdot \Delta u, \tag{16}$$

where ε_M is the segmentation threshold, Δu is the area of one spatial element, and $\mathrm{I}(\bullet)$ is the indicator function. The temperature anomaly is defined as:

$$\Delta T_i(t) = \frac{1}{\Omega_i} \cdot \sum_{u \in \Omega_i} \left| T_i(u,t) - T_i^{norm} \right|, \tag{17}$$

where $T_i(u, t)$ is the temperature at the u-th point, and T_i^{norm} is the normative or locally estimated base temperature. In this case, the complex local damage index is given as:

$$L_i(t) = \beta_1 \cdot A_i(t) + \beta_2 \cdot \Delta T_i(t) + \beta_3 \cdot B_i(t) + \beta_4 \cdot R_i(t) + \beta_5 \cdot P_i(t), \tag{18}$$

where $B_i(t)$ is the bleeding or visible damage probability, $R_i(t)$ is the immobility or pathological posture indicator, and $P_i(t)$ is the spatial risk associated with falling, squeezing, or dangerous body position.

The calculated local damage index provides an aggregated representation of detected injury indicators and serves as an intermediate parameter for the subsequent severity classification module.

It is advisable to present the assessment of the condition's severity as a multi-class problem. For this aim, it is assumed that the set of severity classes has the form:

$$C = \{0,1,2,3\}, \tag{19}$$

where "0" is the absence of critical signs, "1" is the low severity, "2" is the medium severity, and "3" is the critical condition. The softmax function determines the victim p_i belonging to the c-th class:

$$\text{Ψ}_{ic}(t) = \frac{\exp\left(a_c^{\mathrm{T}} \cdot z_i(t) + b_c\right)}{\sum_{r=0}^{3} \exp\left(a_r^{\mathrm{T}} \cdot z_i(t) + b_r\right)}. \tag{20}$$

The expected severity index is defined as:

$$S_i(t) = \sum_{c=0}^{3} c \cdot \text{Ψ}_{ic}(t). \tag{21}$$

The normalised severity index is defined as:

$$S_i^{norm}(t) = \frac{S_i(t)}{3}. \tag{22}$$

The normalised severity assessment completes the diagnostic stage and enables the transition from injury detection to automated prioritisation of emergency assistance.

If $\text{Ψ}_{i3}(t)$ - increases, the critical class's contribution to the overall score increases, and the victim receives a higher response priority. To prioritise care, an integral triage index is introduced:

$$TI_i(t) = \gamma_1 \cdot \text{Ψ}_{i3}(t) + \gamma_2 \cdot S_i^{norm}(t) + \gamma_3 \cdot L_i^{norm}(t) + \gamma_4 \cdot U_i(t) + \gamma_5 \cdot C_i^{risk}(t), \tag{23}$$

where $\text{Ψ}_{i3}(t)$ is the probability of a critical state, $S_i^{norm}(t)$ is the normalised severity, $L_i^{norm}(t)$ is the normalised damage index, $U_i(t)$ is the model uncertainty, and $C_i^{risk}(t)$ is the event's contextual risk. In this case, the coefficients γ_k must satisfy the condition:

$$\sum_{k=1}^{5} \gamma_k = 1, \gamma_k \geq 0. \tag{24}$$

The solution's uncertainty is determined through the class distribution's entropy:

$$U_i(t) = -\frac{1}{\ln(4)} \cdot \sum_{c=0}^{3} \text{Ψ}_{ic}(t) \cdot \ln\left(\text{Ψ}_{ic}(t)\right). \tag{25}$$

If the model makes a clear prediction, the entropy is low; if the classes have similar probabilities, the entropy increases. In emergency response systems, high uncertainty should not be ignored; it is therefore included in the triage index as a separate risk factor. A threshold rule determines the triage category:

$$Triage_i(t)=\begin{cases}\text{Critical}, TI_i(t)\geq\delta_3,\\ \text{High}, \delta_2\leq TI_i(t)<\delta_3,\\ \text{Medium}, \delta_1\leq TI_i(t)<\delta_2,\\ \text{Low}, TI_i(t)<\delta_1,\end{cases} \tag{26}$$

where $0 < \delta_1 < \delta_2 < \delta_3 < 1$. The triage category obtained represents the decision-making output of the medical assessment module and is subsequently incorporated into the resource allocation and optimisation processes.

For example, for a model system, you can specify $\delta_1 = 0.35$, $\delta_2 = 0.60$, and $\delta_3 = 0.80$. Then a victim with $TI_i(t) = 0.84$ belongs to the critical category, and a victim with $TI_i(t) = 0.52$ belongs to the medium or high borderline group, depending on the system's clinical rules. A formalised triage index advantage is that the decision is based on a combination of sensory, probabilistic, and contextual features, rather than a single visual symptom.

The model's training is formalised as minimising the total loss function. For this aim, it is assumed that the true lesion mask M_i, the severity class y_i, the triage category r_i, and the actual financial and resource outcome F_i are known for the victim p_i. Then the total loss function is given as:

$$L_{total}=L_{seg}+\eta_1\cdot L_{cls}+\eta_2\cdot L_{oss}+\eta_3\cdot L_{triage}+\eta_4\cdot L_{calc}+\eta_5\cdot L_{edge}+\eta_6\cdot L_{fin}. \tag{27}$$

Segmentation loss can have a combined form of Dice and BCE:

$$L_{seg}=1-\frac{2\cdot\sum_{u\in\Omega_i}M_i(u)\cdot\P M_i(u)+\varepsilon}{\sum_{u\in\Omega_i}M_i(u)+\sum_{u\in\Omega_i}\P M_i(u)+\varepsilon}-\frac{1}{|\Omega_i|}\cdot\sum_{u\in\Omega_i}\left(M_i(u)\cdot\ln\left(\P M_i(u)\right)+\left(1-M_i(u)\right)\cdot\ln\left(1-\P M_i(u)\right)\right). \tag{28}$$

Classification loss is defined as:

$$L_{cls}=-\sum_{c=0}^{3}y_{ic}\cdot\ln\left(\mu p_{ic}\right). \tag{29}$$

Since the severity is ordinal in nature, it is advisable to add ordinal loss:

$$L_{oss}=-\sum_{k=1}^{3}\mathrm{I}(y_i\geq k)\cdot\ln\left(\$\sigma_{ik}\right)+\left(1-\mathrm{I}(y_i\geq k)\right)\cdot\ln\left(1-\$\sigma_{ik}\right), \tag{30}$$

where $\$\sigma_{ik}$ is the probability that the severity is not lower than the k-th level. Therefore, the proposed L_{oss} form (30) penalises not only the wrong class but also the wrong order of severity, which is fundamental for triage.

The combined training objective ensures joint optimisation of detection accuracy, severity estimation, and triage reliability before applying edge-computing constraints.

Constraints on latency, memory, power consumption, and computational complexity give the mathematical model its edge. The total latency of local inference is defined as:

$$\tau_{edge}=\tau_{cap}+\tau_{sync}+\tau_{prep}+\tau_{feat}+\tau_{fusion}+\tau_{infer}+\tau_{post}. \tag{31}$$

For real-time operation, the following condition must be met:

$$\tau_{edge}\leq\tau_{\max}. \tag{32}$$

For example, if $\tau_{max} = 500$ ms, and $\tau_{cap} = 40$ ms, $\tau_{sync} = 35$ ms, $\tau_{prep} = 55$ ms, $\tau_{feat} = 120$ ms, $\tau_{fusion} = 45$ ms, $\tau_{infer} = 90$ ms, $\tau_{post} = 40$ ms, then $\tau_{edge} = 40 + 35 + 55 + 120 + 45 + 90 + 40 = 425$ ms, which meets the real-time requirement.

The penalty for violating edge constraints is given as:

$$L_{edge}=\left(\max\left(0,\tau_{edge}-\tau_{\max}\right)\right)^2+\mu_1\cdot\left(\max\left(0,M_{model}-M_{\max}\right)\right)^2+\mu_2\cdot\left(\max\left(0,E_{inf}-E_{\max}\right)\right)^2, \tag{33}$$

where M_{model} is the model memory size, E_{inf} is the energy per inference, M_{max} and E_{max} are the allowable limits.

The edge constraint model links the neural network optimisation process with hardware limitations, ensuring that diagnostic decisions can be generated within real-time operational requirements.

For comparison with the cloud scenario, the total latency of remote processing is given as

$\tau_{cloud}=\tau_{cap}+\tau_{uplink}+\tau_{queue}+\tau_{cloud.infer}+\tau_{downlink}+\tau_{post}$.

$$\tau_{cloud}=\tau_{cap}+\tau_{uplink}+\tau_{queue}+\tau_{cloud.infer}+\tau_{downlink}+\tau_{post}. \tag{34}$$

Edge processing is prioritised if:

$$\tau_{edge} < \tau_{cloud}, \tag{35}$$

and:
$P_{edge}^{fail} < P_{net}^{fail}$,

$$P_{edge}^{fail} < P_{net}^{fail}, \tag{36}$$

where P_{edge}^{fail} is the edge node failure probability, and P_{net}^{fail} is the probability of network unavailability or unacceptable delay. In critical scenarios, the autonomy condition is also appropriate:

$$Decision_i(t) = f_{edge}(X_i(t)), \tag{37}$$

it is noted that $f_{cloud}(X_i(t))$ - is used only for archival analytics and training.

The local inference result is subsequently transferred to the response coordination module, where medical priorities are converted into practical allocation decisions.

Thus, the given mathematical formulation formally reflects the position of the developed multispectral edge deep learning system's architecture in its block diagram (Fig. 5). In particular, the initial solution is formed locally, and the cloud is not a critical link.

The response coordination model is formalised as an optimal resource-allocation problem. For this purpose, it is assumed that $\mathcal{A} = \{a_1, \ldots, a_J\}$ is the set of available teams, $\mathcal{H} = \{h_1, \ldots, h_K\}$ is the set of hospitals or reception points. A binary variable of the form is introduced:

$$x_{ij} = \begin{cases} 1, \text{if the brigade } a_j \text{ is assigned to the victim } p_i, \\ 0, \text{otherwise.} \end{cases} \tag{38}$$

For hospitalisation, the binary variable has the form:

$$y_{ik} = \begin{cases} 1, \text{if the victim } p_i \text{ is reffered to the hospital } h_k, \\ 0, \text{otherwise.} \end{cases} \tag{39}$$

The time of the assistance onset for the victim p_i is defined as:

$$T_i^{help} = \sum_{j=1}^{J} x_{ij} \cdot T_{ij}^{arr}, \tag{40}$$

where T_{ij}^{arr} is the time of arrival of the j-th brigade at the i-th victim. Delivery time to the hospital is defined as:

$$T_i^{hosp} = \sum_{k=1}^{K} y_{ik} \cdot T_{ik}^{trans}. \tag{41}$$

Total time risk is defined as:

$$R_T = \sum_{i=1}^{N} \omega_i \cdot \left(T_i^{help} + T_i^{hosp}\right), \tag{42}$$

where the triage index determines the weight coefficient:

$$\omega_i = 1 + \rho \cdot TI_i. \tag{43}$$

It is noted that critical victims receive a larger coefficient ω_i, so their expectations impose a larger penalty on the objective function. The general response optimisation problem has the form:

$$\min_{x_{ij}, y_{ik}} \left(\sum_{i=1}^{N} \omega_i \cdot \left(\sum_{j=1}^{J} x_{ij} \cdot T_{ij}^{arr} + \sum_{k=1}^{K} y_{ik} \cdot T_{ik}^{trans} \right) + \lambda_C \cdot C_{res} + \lambda_O \cdot C_{over} \right), \tag{44}$$

under the conditions:

$$\sum_{j=1}^{J} x_{ij} = 1, i = 1, \ldots, N; \sum_{j=1}^{J} y_{ik} = 1, i = 1, \ldots, N; \sum_{i=1}^{N} x_{ij} \le Cap_j, j = 1, \ldots, J; \sum_{i=1}^{N} y_{ik} \le Bed_k, k = 1, \ldots, K; x_{ij}, y_{ik} \in \{0,1\}, \tag{45}$$

where C_{res} is the cost of resource mobilisation, C_{over} is the hospital or team overload penalty, Cap_j is the team's capacity

or operational resources, and Bed_k is the hospital's available admission capacity. Thus, the proposed model links medical prioritisation to the actual distribution of teams, routes, and hospital capacity. The resource allocation solution connects the predicted victim state with available emergency capacities, forming the basis for cost-aware response optimisation.

The system's financial security is described in terms of expected losses from mis-triage, delays, re-admissions, over-resourcement, and hospital overload. For a victim p_i, the expected cost of the decision is given as:

$$\mathrm{E}\left[C_i\right] = C_i^{disp} + C_i^{trans} + C_i^{hosp} + C_i^{delay} + C_i^{err}. \tag{46}$$

The cost of error is defined as:

$$C_i^{err} = \hat{p}_{i3} \cdot C_{FN} \cdot \mathrm{I}\left(Triage_i \neq Critical\right) + \left(1 - \hat{p}_{i3}\right) \cdot C_{FP} \cdot \mathrm{I}\left(Triage_i = Critical\right), \tag{47}$$

where C_{FN} is the loss from underestimating the critical state, and C_{FP} is the loss from falsely overestimating the priority. Since underestimating the critical state is much more dangerous for emergency response, usually

$$C_{FN} \geq C_{FP}, \tag{48}$$

which means the model should be more sensitive to critical states, even if this sometimes increases the number of false alarms.

The threshold rule for the critical response can be derived by comparing the expected losses of two actions: A_1 is to give critical priority, and A_0 is not to give critical priority. It is assumed that $p = \hat{p}_{i3}$. Then the expected losses for the critical priority are:

$$\mathrm{E}\left[C \middle| A_1\right] = C_U + \left(1 - p\right) \cdot C_{FP}, \tag{49}$$

where C_U is the additional cost of engaging an urgent resource. For non-critical priority:

$$\mathrm{E}\left[C \middle| A_0\right] = p \cdot C_{FN}. \tag{50}$$

Critical priority is selected if:

$$C_U + \left(1 - p\right) \cdot C_{FP} < p \cdot C_{FN}. \tag{51}$$

Let's solve the inequality:

$$\begin{gathered} C_U + C_{FP} - p \cdot C_{FP} < p \cdot C_{FN}, \\ C_U + C_{FP} < p \cdot \left(C_{FN} + C_{FP}\right), \\ p > \frac{C_U + C_{FP}}{C_{FN} + C_{FP}}. \end{gathered} \tag{52}$$

Therefore, the optimal financially justified critical response threshold is defined as:

$$p^* = \frac{C_U + C_{FP}}{C_{FN} + C_{FP}}. \tag{53}$$

If C_{FN} increases, the denominator increases and the threshold p^* decreases, which mathematically justifies a lower threshold for critical response in high cost of missed severe injury conditions.

The general function of the system's financial and resource security is as follows:

$$C_{total} = \sum_{i=1}^{N} \left(C_i^{disp} + C_i^{trans} + C_i^{hosp} + C_i^{delay} + C_i^{err}\right) + C^{audit} + C^{model} + C^{infra}. \tag{54}$$

The financial security optimisation problem has the form:

$$\min\left(\mathrm{E}\left[C_{total}\right]\right), \tag{55}$$

under the conditions:

$$\tau_{edge} \leq \tau_{\max}; Acc_{crit} \geq Acc_{\min}; Recall_{crit} \geq R_{\min}; C_{total} \leq B, \tag{56}$$

where B is the available response budget, Acc_{crit} is the accuracy of the critical category, and $Recall_{crit}$ is the completeness of critical condition detection. In emergency response systems, it is particularly important to control $Recall_{crit}$, since missing a critical condition incurs higher medical and financial costs than erroneously overestimating

the priority.

The financial optimisation module completes the decision-making cycle by evaluating the economic consequences of response strategies and minimising unnecessary resource expenditure.

To audit decisions, an event log is entered:

$$J = \left\{ X_i(t), q_i^{(m)}(t), z_i(t), p_{ic}(t), TI_i(t), Triage_i(t), x_{ij}, y_{ik}, C_i, Outcome_i \right\}_{i=1}^{N}, \tag{57}$$

which records not only the final solution but also all intermediate values that contributed to its formation, and is important for checking the system's operation, financial control, error analysis, and further model training.

The feedback loop is formalised through updates to the training sample. It is assumed that after the event, confirmed clinical data have been obtained:

$$Y_i^* = \left\{ M_i^*, y_i^*, r_i^*, Outcome_i^*, C_i^* \right\}. \tag{58}$$

Then the updated dataset looks like:

$$D_{t+1} = D_t \cup \left\{ X_i(t), Y_i^* \right\}_{i=1}^{N}. \tag{59}$$

The model parameters are updated according to the gradient descent rule:

$$\theta_{t+1} = \theta_t - \eta \cdot \nabla_\theta \left(L_{total}(D_t) + \kappa \cdot L_{new}\left(\frac{D_{t+1}}{D_t} \right) \right), \tag{60}$$

where η is the learning rate, κ is the significance coefficient of new confirmed cases. The updated parameters obtained from confirmed cases provide a feedback mechanism that allows the system to adapt to new operational conditions and improve future inference accuracy. To control for changes in the distribution of events, a drift indicator is used:

$$Drift_t = \left\| \mathrm{E}_{X:\, D_{new}}\left[z(X) \right] - \mathrm{E}_{X:\, D_{train}}\left[z(X) \right] \right\|_2 + D_{KL}\left(P_{new}(Y) \,\|\, P_{train}(Y) \right). \tag{61}$$

If:

$$Drift_t > \Delta_{crit}, \tag{62}$$

then, the system enters recalibration or retraining mode, which corresponds to the scheme's "lower block" (Fig. 5), where the actual response results are fed back to the edge-analytical core to increase the model's stability.

Thus, the general mathematical model of the developed multispectral edge system is presented as a multi-criteria optimisation problem of the form:

$$\theta^*, x^*, y^* = \arg\min_{\theta, x, y} \left(L_{total}(\theta) + \lambda_T \cdot \sum_{i=1}^{N} \omega_i \cdot T_i + \lambda_C \cdot \mathbb{E}[C_{total}] \right), \tag{63}$$

under the conditions:

$$\tau_{edge} \le \tau_{\max}; Acc_{crit} \ge Acc_{\min}; Recall_{crit} \ge R_{\min}; Precision_{triage} \ge P_{\min}; C_{total} \le B;$$

$$x_{ij}, y_{ik} \in \{0,1\}; \sum_{j=1}^{J} x_{ij} = 1; \sum_{k=1}^{K} y_{ik} = 1, \tag{64}$$

and combines four levels, including multispectral sensor surveillance, edge-inference deep learning, triage prioritisation, and financial and resource optimisation. Thus, the final optimisation formulation integrates all previously defined modules into a unified mathematical framework that describes the complete operational cycle of the multispectral edge deep learning system. Its scientific significance lies in the fact that early detection of human lesions is considered not an isolated computer vision task but part of a complex decision-making system, where recognition quality directly affects the timing of assistance, team distribution, hospital routing, decision audit, and the financial security of emergency response.

3.3. Proving the developed multispectral edge deep learning system's mathematical model's stability using the Lyapunov method

It is assumed that the developed multispectral edge system operates in discrete inference cycles $k = 0, 1, 2, \ldots$, where each cycle corresponds to sensor data acquisition, their fusion, assessment of the human condition, formation of the triage index, selection of a resource solution, and updating of service information. It is advisable to assess stability

not for a single image but for the system's dynamics relative to the desired equilibrium state. Such an equilibrium state is a situation in which the multispectral feature error, the severity assessment error, the triage index error, the resource allocation error, the financial assessment error, and the model parameters error are all equal to zero. We denote the system error vector as:

$$e(k) = \mathrm{col}\left(e_z(k), e_s(k), e_T(k), e_r(k), e_C(k), e_\theta(k)\right), \tag{65}$$

where $e_z(k) = z(k) - z^*(k)$ is the error of the integrated vector of multispectral features, $e_s(k) = S(k) - S^*(k)$ is the error in assessing the condition's severity, $e_T(k) = TI(k) - TI^*(k)$ is the triage index error, $e_r(k) = r(k) - r^*(k)$ is the resource decision error, where $r(k)$ generalizes the crew, route, and hospital assignment, $e_C(k) = C(k) - C^*(k)$ is the error in financial and resource assessment, $e_\theta(k) = \theta(k) - \theta^*$ is the error of the deep learning model's parameters. The desired equilibrium state of the system is $e^* = 0$, which means that the system correctly synchronises sensor data, stably generates signs, correctly assesses severity, does not change the triage class without a valid reason, optimally allocates resources, and does not accumulate financial or resource errors.

The errors' dynamics after linearization in the vicinity of the equilibrium can be represented as:

$$e(k+1) = A \cdot e(k) + d(k), \tag{66}$$

where A is the closed-loop system dynamics matrix, $d(k)$ is a limited disturbance vector due to sensor noise, incomplete scene coverage, calibration inaccuracy, discrete resource decisions, and processing delays. For the structure shown in Fig. 5, matrix A has a block-lower triangular form:

$$A = \begin{pmatrix} A_z & 0 & 0 & 0 & 0 & B_{z\theta} \\ B_{Sz} & A_s & 0 & 0 & 0 & B_{s\theta} \\ B_{Tz} & B_{Ts} & A_T & 0 & 0 & B_{T\theta} \\ 0 & 0 & B_{rT} & A_r & 0 & 0 \\ 0 & 0 & B_{CT} & B_{cR} & A_C & 0 \\ 0 & 0 & 0 & 0 & 0 & A_\theta \end{pmatrix}, \tag{67}$$

which reflects the system's causal logic: sensory errors affect features; features affect severity assessment; severity assessment affects triage; triage affects resource allocation; and resource decisions affect financial assessment.

The system dynamics matrix A is obtained through the linearisation of the nonlinear state-transition model of the developed multispectral edge deep learning system in the vicinity of the desired equilibrium state. Unlike a predefined empirical matrix, A is derived from the internal relationships among the system's main functional components, including multispectral feature extraction, feature fusion, injury severity assessment, triage decision formation, resource allocation, financial optimisation, and model adaptation.

The complete system state is defined as a set of variables describing the current condition of all major computational and decision-making modules. These variables include the fused multispectral feature representation, the estimated injury severity level, the integral triage index, the current resource allocation state, the financial-resource assessment state, and the parameters of the deep learning model. The evolution of this state is determined by the interaction between sensor information processing, neural network inference, emergency response decisions, and feedback-based model updating.

The original system behaviour is nonlinear due to deep learning transformations, adaptive attention mechanisms, threshold-based triage decisions, and optimisation procedures. Therefore, the local dynamics around the equilibrium point are approximated by applying a first-order Taylor expansion of the nonlinear transition function. As a result, the deviations of the current system state from the desired equilibrium state can be described by a linearised closed-loop model.

The elements of the matrix A are determined as partial sensitivity coefficients of the closed-loop system dynamics with respect to each state variable at the equilibrium point. In physical terms, each element characterises how a small variation in one subsystem affects the future state of another subsystem during the operation of the multispectral edge system.

The structure of matrix A reflects the internal coupling between different functional levels of the proposed architecture. The blocks associated with multispectral features describe how sensor fusion accuracy and feature stability influence subsequent analytical processes. The blocks related to injury severity and triage estimation represent the dependence of emergency decisions on the quality of detected medical indicators. The resource-allocation blocks describe how changes in victim priority influence the distribution of medical teams, routes, and hospital capacity. The financial-management blocks characterise the impact of diagnostic and operational decisions on expected resource losses. Finally, the blocks associated with model parameters describe how feedback-based adaptation influences the future behaviour of the deep learning component.

The disturbance component represents external factors that cannot be fully controlled by the system, including

sensor noise, incomplete scene observation, calibration errors, communication disturbances, and uncertainty in emergency conditions. These disturbances affect the system dynamics but do not define its intrinsic stability properties.

The model parameters $e_\theta(k)$ affect the sensory, classification, and triage parts, and the retraining loop corrects them in subsequent cycles. To prove stability, we will use a discrete Lyapunov function of quadratic type:

$$V\left(e\left(k\right)\right)=e^{\mathrm{T}}\left(k\right)\cdot P\cdot e\left(k\right), \tag{68}$$

where:

$$P=P^{\mathrm{T}}>0. \tag{69}$$

Since P is a positive definite matrix, there exist positive eigenvalues $\lambda_{min}(P)$ and $\lambda_{max}(P)$ for which the inequality holds:

$$\lambda_{\min}\left(P\right)\cdot\left\|e\left(k\right)\right\|^{2}\le V\left(e\left(k\right)\right)\le\lambda_{\max}\left(P\right)\cdot\left\|e\left(k\right)\right\|^{2}. \tag{70}$$

Therefore, the function $V(e(k))$ is positive definite, and its decrease means a decrease in all components of the system error. For the undisturbed case, when $d(k) = 0$, the dynamics looks like:

$$e\left(k+1\right)=A\cdot e\left(k\right). \tag{71}$$

Then the Lyapunov function's increment is equal to:

$$\Delta V=V\left(e\left(k+1\right)\right)-V\left(e\left(k\right)\right). \tag{72}$$

Let's substitute $e(k + 1) = A \cdot e(k)$:

$$\Delta V=\left(A\cdot e\left(k\right)\right)^{\mathrm{T}}\cdot P\cdot\left(A\cdot e\left(k\right)\right)-e^{\mathrm{T}}\left(k\right)\cdot P\cdot e\left(k\right). \tag{73}$$

From here:

$$\Delta V=e^{\mathrm{T}}\left(k\right)\cdot A^{\mathrm{T}}\cdot P\cdot A\cdot e\left(k\right)-e^{\mathrm{T}}\left(k\right)\cdot P\cdot e\left(k\right),$$
$$\Delta V=e^{\mathrm{T}}\left(k\right)\cdot\left(A^{\mathrm{T}}\cdot P\cdot A-P\right)\cdot e\left(k\right). \tag{74}$$

If the matrix A is a Schur matrix, that is, all its eigenvalues lie inside the unit circle, $\rho(A) < 1$, then for any positive definite matrix:

$$Q=Q^{\mathrm{T}}>0. \tag{75}$$

there is a unique positive definite matrix $P=P^{\mathrm{T}}>0$ that satisfies the discrete Lyapunov equation:

$$A^{\mathrm{T}}\cdot P\cdot A-P=-Q. \tag{76}$$

Then:

$$\Delta V=-e^{\mathrm{T}}\left(k\right)\cdot Q\cdot e\left(k\right). \tag{77}$$

Since $Q > 0$, we have:

$$e^{\mathrm{T}}\left(k\right)\cdot Q\cdot e\left(k\right)\ge\lambda_{\min}\left(Q\right)\cdot\left\|e\left(k\right)\right\|^{2}. \tag{78}$$

Therefore,

$$\Delta V\le-\lambda_{\min}\left(Q\right)\cdot\left\|e\left(k\right)\right\|^{2}<0,e\left(k\right)\ne0, \tag{79}$$

which proves the asymptotic stability of the equilibrium $e^{*} = 0$ for the unperturbed model.

It is necessary to justify why the matrix A can be considered a Schur matrix. Since the matrix A has a block-lower triangular structure, its eigenvalues are determined by the diagonal blocks' eigenvalues:

$$\sigma\left(A\right)=\sigma\left(A_{z}\right)\cup\sigma\left(A_{s}\right)\cup\sigma\left(A_{T}\right)\cup\sigma\left(A_{r}\right)\cup\sigma\left(A_{C}\right)\cup\sigma\left(A_{\theta}\right). \tag{80}$$

Therefore, it is sufficient that the conditions are met:

$$\rho(A_z)<1, \rho(A_s)<1, \rho(A_T)<1, \rho(A_r)<1, \rho(A_C)<1, \rho(A_\theta)<1. \tag{81}$$

The A_z unit is responsible for the multispectral fusion's stability. Let each feature extractor be Lipschitzian:

$$\left\|F_m(X_1)-F_m(X_2)\right\| \le L_m \left\|X_1-X_2\right\|. \tag{82}$$

The integrated feature vector has the form:

$$z(k)=\sum_{m\in M}\lambda_m(k)\cdot h^{(m)}(k), \tag{83}$$

where:

$$\lambda_m(k)\ge 0,\ \sum_{m\in M}\lambda_m(k)=1. \tag{84}$$

Then, for two close states:

$$\left\|\Delta z(k)\right\|=\left\|\sum_{m\in M}\lambda_m(k)\cdot h^{(m)}(k)\right\| \le \sum_{m\in M}\lambda_m(k)\cdot\left\|h^{(m)}(k)\right\|. \tag{85}$$

Taking into account the Lipschitz nature of the extractors:

$$\left\|\Delta z(k)\right\| \le \sum_{m\in M}\lambda_m(k)\cdot L_m\left\|\Delta X^{(m)}(k)\right\|. \tag{86}$$

If, after pre-processing and quality control, the sensory deviation decreases with a coefficient $\kappa_X < 1$, i.e.,

$$\left\|\Delta X^{(m)}(k+1)\right\| \le \kappa_X\cdot\left\|\Delta X^{(m)}(k)\right\|, \tag{87}$$

then, for the integrated vector, we obtain:

$$\left\|e_z(k+1)\right\| \le \kappa_X\cdot\left(\sum_{m\in M}\lambda_m\cdot L_m\right)\cdot\left\|e_z(k)\right\|. \tag{88}$$

Let us denote:

$$a_z=\kappa_X\cdot\left(\sum_{m\in M}\lambda_m\cdot L_m\right). \tag{89}$$

If $0 < a_z < 1$, then:

$$\left\|e_z(k+1)\right\| \le a_z\cdot\left\|e_z(k)\right\|, \tag{90}$$

and the multispectral fusion unit is compressive.

The severity assessment unit, according to (20), is described by a softmax model:

$$\text{Ÿ}_c(k)=\frac{\exp\left(a_c^{\mathrm{T}}\cdot z(k)+b_c\right)}{\sum_{r=0}^{3}\exp\left(a_r^{\mathrm{T}}\cdot z(k)+b_r\right)}. \tag{91}$$

Expected severity is defined as:

$$S(k)=\sum_{c=0}^{3}c\cdot \text{Ÿ}_c(k). \tag{92}$$

Since softmax is Lipschitzian on a bounded set of inputs, there exists a constant $L_{sm} > 0$ such that:

$$\Delta\text{Ÿ}(k)\le L_{sm}\cdot\|W\|\cdot\left\|\Delta z(k)\right\|. \tag{93}$$

Then:

$$\left\|e_s(k)\right\|=\left|\sum_{c=0}^{3}c\cdot \text{Ÿ}_c(k)\right| \le 3\cdot\left\|\Delta\text{Ÿ}(k)\right\|_1. \tag{94}$$

Therefore,

$$\|e_s(k)\| \le K_s \cdot \|e_z(k)\|, \tag{95}$$

where:

$$K_s = 3 \cdot L_{sm} \cdot \|W\|. \tag{96}$$

In the presence of severity assessment smoothing:

$$S(k+1) = (1-\alpha_s) \cdot S(k) + \alpha_s \cdot \hat{S}(k), 0 < \alpha_s < 1, \tag{97}$$

the error has the form:

$$e_s(k+1) = (1-\alpha_s) \cdot e_s(k) + \alpha_s \cdot \Delta\hat{S}(k). \tag{98}$$

From here:

$$|e_s(k+1)| \le (1-\alpha_s) \cdot |e_s(k)| + \alpha_s \cdot K_s \cdot |e_z(k)|. \tag{99}$$

Because $0 < 1 - \alpha_s < 1$, the internal dynamics of the severity assessment are stable, and the $e_z(k)$ influence is limited.

The triage index, according to (23), has the form:

$$TI(k) = \gamma_1 \cdot \mu_3(k) + \gamma_2 \cdot S^{norm}(k) + \gamma_3 \cdot L^{norm}(k) + \gamma_4 \cdot U(k) + \gamma_5 \cdot C^{risk}(k), \tag{100}$$

where:

$$\gamma_j \ge 0, \sum_{j=1}^{5} \gamma_j = 1. \tag{101}$$

The triage index error is defined as:

$$e_T(k) = \Delta TI(k). \tag{102}$$

Then:

$$|e_T(k)| \le \gamma_1 \cdot |\Delta\mu_3(k)| + \gamma_2 \cdot |\Delta S^{norm}(k)| + \gamma_3 \cdot |\Delta L^{norm}(k)| + \gamma_4 \cdot |\Delta U(k)| + \gamma_5 \cdot |\Delta C^{risk}(k)|. \tag{103}$$

For Lipschitzianity of all components:

$$|e_T(k)| \le K_T \cdot (|e_z(k)| + |e_s(k)|). \tag{104}$$

If the system uses a triage index time filter:

$$TI(k+1) = (1-\alpha_T) \cdot \overline{TI}(k) + \alpha_T \cdot TI(k), 0 < \alpha_T < 1, \tag{105}$$

then, the smoothed triage error satisfies the inequality:

$$|e_T(k+1)| \le (1-\alpha_T) \cdot |e_T(k)| + \alpha_T \cdot K_T \cdot (|e_z(k)| + |e_s(k)|). \tag{106}$$

Therefore, the intrinsic dynamics of the triage unit have a coefficient:

$$\alpha_T = 1 - \alpha_T, |\alpha_T| < 1. \tag{107}$$

For a categorical decision, it is important to consider thresholds $0 < \delta_1 < \delta_2 < \delta_3 < 1$. The triage category is robust to small deviations if the triage index does not lie on the threshold boundary. Let us introduce a margin to the nearest threshold:

$$m(k) = \min\{|TI(k) - \delta_1|, |TI(k) - \delta_2|, |TI(k) - \delta_3|\}. \tag{108}$$

If:

$$\left|e_T(k)\right| < \frac{m(k)}{2}, \tag{109}$$

then, the perturbation does not change the triage category, meaning that not only the numerical index but also the discrete solution itself is stable under a non-zero margin from the thresholds.

The resource unit is given as a projected gradient step problem:

$$r(k+1) = \Pi_{\text{Я}}\left[r(k) - \eta_r \cdot \nabla_r J\left(r(k), TI(k)\right)\right], \tag{110}$$

where $\Pi_{\mathcal{R}}$ is the projection onto the admissible set of resource solutions, J is the response loss function, and η_r is the optimisation step. If J is strongly convex with respect to r in the vicinity of the optimum:

$$\mu_r \cdot I \text{ p } \nabla^2_{rr} J(r, TI) \text{ p } L_r \cdot I, \tag{111}$$

and:

$$0 < \eta_r < \frac{2}{L_r}, \tag{112}$$

then, for the resource error, the following holds:

$$\left\|e_r(k+1)\right\| \le \rho_r \cdot \left\|e_r(k)\right\| + \eta_r \cdot L_{rT} \cdot \left\|e_T(k)\right\|, \tag{113}$$

where:

$$\rho_r = \max\left\{\left|1 - \eta_r \cdot \mu_r\right|, \left|1 - \eta_r \cdot L_r\right|\right\} < 1. \tag{114}$$

Projection does not increase the distance between two points:

$$\left\|\Pi_{\text{Я}}(u) - \Pi_{\text{Я}}(v)\right\| \le \|u - v\|, \tag{115}$$

therefore, the resource unit remains stable provided that η_r is chosen correctly, and the triage error's effect is included as a limited disturbance.

A recurrent cost estimate describes the financial unit:

$$C(k+1) = (1-\alpha_C)\cdot C(k) + \alpha_C \cdot \text{Ц}\left(TI(k), r(k)\right), 0 < \alpha_C < 1. \tag{116}$$

Then:

$$e_C(k+1) = (1-\alpha_C)\cdot e_C(k) + \alpha_C \cdot \left(\text{Ц}\left(TI(k), r(k)\right) - \text{Ц}\left(TI^*(k), r^*(k)\right)\right). \tag{117}$$

For Lipschitz financial function:

$$\text{Ц}(TI, r) - \text{Ц}\left(TI^*, r^*\right) \le L_{CT} \cdot \left|e_T(k)\right| + L_{Cr} \cdot \left|e_r(k)\right|. \tag{118}$$

Because $0 < 1 - \alpha_C < 1$, the financial and resource assessment does not accumulate unlimited error if the triage and resource errors are limited.

The gradient update rule describes the model retraining contour according to (60):

$$\theta(k+1) = \theta(k) - \eta_\theta \cdot \nabla_\theta L_{total}\left(\theta(k)\right). \tag{119}$$

Then:

$$\nabla_\theta L_{total}\left(\theta(k)\right) \approx H_\theta \cdot \left(\theta(k) - \theta^*\right), \tag{120}$$

where:

$$\mu_\theta \cdot I \text{ p } H_\theta \text{ p } L_\theta \cdot I. \tag{121}$$

Therefore,

$$e_\theta(k+1) = \theta(k+1) - \theta^* = \theta(k) - \eta_\theta \cdot H_\theta \cdot \left(\theta(k) - \theta^*\right) - \theta^*, \tag{122}$$

$$e_\theta(k+1) = (I - \eta_\theta \cdot H_\theta) \cdot e_\theta(k).$$

If:

$$0 < \eta_\theta < \frac{2}{L_\theta}, \tag{123}$$

then:

$$\rho(A_\theta) = \rho(I - \eta_\theta \cdot H_\theta) \le \max\{|1 - \eta_\theta \cdot \mu_\theta|, |1 - \eta_\theta \cdot L_\theta|\} < 1. \tag{124}$$

Therefore, the model's parametric error decreases in the vicinity of a stable local minimum. For the entire system, we have $\rho(A) < 1$. Therefore, there exists $P > 0$ for which $A^{\mathrm{T}} \cdot P \cdot A - P = -Q$, $Q > 0$. Hence, for the undisturbed case:

$$\Delta V = -e^{\mathrm{T}}(k) \cdot Q \cdot e(k) \le \lambda_{\min}(Q) \cdot \|e(k)\|^2. \tag{125}$$

Also:

$$V(e(k)) \le \lambda_{\max}(P) \cdot \|e(k)\|^2. \tag{126}$$

Therefore,

$$\|e(k)\|^2 \ge \frac{V(e(k))}{\lambda_{\max}(P)}. \tag{127}$$

Substitute this into the inequality for ΔV:

$$V(e(k+1)) - V(e(k)) \le -\lambda_{\min}(Q) \cdot \frac{V(e(k))}{\lambda_{\max}(P)}. \tag{128}$$

Therefore,

$$V(e(k+1)) \le \left(1 - \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}\right) \cdot V(e(k)). \tag{129}$$

Let us denote:

$$q = 1 - \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}. \tag{130}$$

Because $0 < q < 1$, then:

$$V(e(k)) \le q^k \cdot V(e(0)). \tag{131}$$

Given the limits for V:

$$\lambda_{\min}(P) \cdot \|e(k)\|^2 \le q^k \cdot \lambda_{\max}(P) \cdot \|e(0)\|^2. \tag{132}$$

From here:

$$\|e(k)\| \le \sqrt{\frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}} \cdot q^{\frac{k}{2}} \cdot \|e(0)\|, \tag{133}$$

which demonstrates the developed model's exponential stability in the vicinity of the equilibrium.

In a real system, we need to account for the perturbation $d(k)$. Then:

$$e(k+1) = A \cdot e(k) + d(k). \tag{134}$$

The Lyapunov function at the next moment:

$$V\left(e\left(k+1\right)\right)=\left(A\cdot e\left(k\right)+d\left(k\right)\right)^{\mathrm{T}}\cdot P\cdot\left(A\cdot e\left(k\right)+d\left(k\right)\right). \tag{135}$$

Let's open the brackets:

$$V\left(e\left(k+1\right)\right)=e^{\mathrm{T}}\left(k\right)\cdot A^{\mathrm{T}}\cdot P\cdot A\cdot e\left(k\right)+2\cdot e^{\mathrm{T}}\left(k\right)\cdot A^{\mathrm{T}}\cdot P\cdot d\left(k\right)+d^{\mathrm{T}}\left(k\right)\cdot P\cdot d\left(k\right). \tag{136}$$

Then:

$$\Delta V=e^{\mathrm{T}}\left(k\right)\cdot\left(A^{\mathrm{T}}\cdot P\cdot A-P\right)\cdot e\left(k\right)+2\cdot e^{\mathrm{T}}\left(k\right)\cdot A^{\mathrm{T}}\cdot P\cdot d\left(k\right)+d^{\mathrm{T}}\left(k\right)\cdot P\cdot d\left(k\right). \tag{137}$$

Because $A^{\mathrm{T}}\cdot P\cdot A-P=-Q$, we have:

$$\Delta V=e^{\mathrm{T}}\left(k\right)\cdot Q\cdot e\left(k\right)+2\cdot e^{\mathrm{T}}\left(k\right)\cdot A^{\mathrm{T}}\cdot P\cdot d\left(k\right)+d^{\mathrm{T}}\left(k\right)\cdot P\cdot d\left(k\right). \tag{138}$$

Let's evaluate the components:

$$-e^{\mathrm{T}}\left(k\right)\cdot Q\cdot e\left(k\right)\le\lambda_{\min}\left(Q\right)\cdot\left\|e\left(k\right)\right\|^{2}. \tag{139}$$

The second component:

$$2\cdot e^{\mathrm{T}}\left(k\right)\cdot A^{\mathrm{T}}\cdot P\cdot d\left(k\right)\le 2\cdot\left\|A^{\mathrm{T}}\cdot P\right\|\cdot\left\|e\left(k\right)\right\|\cdot\left\|d\left(k\right)\right\|. \tag{140}$$

The third component:

$$d^{\mathrm{T}}\left(k\right)\cdot P\cdot d\left(k\right)\le\lambda_{\max}\left(P\right)\cdot\left\|d\left(k\right)\right\|^{2}. \tag{141}$$

Therefore,

$$\Delta V\le\lambda_{\min}\left(Q\right)\cdot\left\|e\left(k\right)\right\|^{2}+2\cdot\left\|A^{\mathrm{T}}\cdot P\right\|\cdot\left\|e\left(k\right)\right\|\cdot\left\|d\left(k\right)\right\|+\lambda_{\max}\left(P\right)\cdot\left\|d\left(k\right)\right\|^{2}. \tag{142}$$

Let's use Young's inequality:

$$2\cdot a\cdot b\le\varepsilon\cdot a^{2}+\frac{1}{\varepsilon}\cdot b^{2}. \tag{143}$$

It is accepted that:

$$a=\left\|e\left(k\right)\right\|,b=\left\|A^{\mathrm{T}}\cdot P\right\|\cdot\left\|d\left(k\right)\right\|,\varepsilon=\frac{\lambda_{\min}\left(Q\right)}{2}. \tag{144}$$

Then:

$$2\cdot\left\|A^{\mathrm{T}}\cdot P\right\|\cdot\left\|e\left(k\right)\right\|\cdot\left\|d\left(k\right)\right\|\le\frac{\lambda_{\min}\left(Q\right)}{2}\cdot\left\|e\left(k\right)\right\|^{2}+\frac{2\cdot\left\|A^{\mathrm{T}}\cdot P\right\|^{2}}{\lambda_{\min}\left(Q\right)}\cdot\left\|d\left(k\right)\right\|^{2}. \tag{145}$$

Let's substitute:

$$\Delta V\le-\frac{\lambda_{\min}\left(Q\right)}{2}\cdot\left\|e\left(k\right)\right\|^{2}+\left(\frac{2\cdot\left\|A^{\mathrm{T}}\cdot P\right\|^{2}}{\lambda_{\min}\left(Q\right)}+\lambda_{\max}\left(P\right)\right)\cdot\left\|d\left(k\right)\right\|^{2}. \tag{146}$$

If the disturbance is limited:

$$\left\|d\left(k\right)\right\|\le\overline{d}, \tag{147}$$

Then:

$$\Delta V\le-\frac{\lambda_{\min}\left(Q\right)}{2}\cdot\left\|e\left(k\right)\right\|^{2}+\left(\frac{2\cdot\left\|A^{\mathrm{T}}\cdot P\right\|^{2}}{\lambda_{\min}\left(Q\right)}+\lambda_{\max}\left(P\right)\right)\cdot\overline{d}^{2}. \tag{148}$$

Let us denote:

$$\chi = \frac{2 \cdot \left\| A^{\mathrm{T}} \cdot P \right\|^2}{\lambda_{\min}(Q)} + \lambda_{\max}(P). \tag{149}$$

Then:

$$\Delta V \le -\frac{\lambda_{\min}(Q)}{2} \cdot \left\| e(k) \right\|^2 + \chi \cdot \bar{d}^2. \tag{150}$$

The Lyapunov function decreases if:

$$\frac{\lambda_{\min}(Q)}{2} \cdot \left\| e(k) \right\|^2 > \chi \cdot \bar{d}^2. \tag{151}$$

From here:

$$\left\| e(k) \right\|^2 > \frac{2 \cdot \chi}{\lambda_{\min}(Q)} \cdot \bar{d}^2, \tag{152}$$

that is, the system is practically stable according to Lyapunov, i.e., in the presence of limited sensor, computational and resource disturbances, the error does not grow indefinitely but enters the limiting region

$$\Omega_d = \left\{ e : \left\| e \right\| \le \sqrt{\frac{2 \cdot \chi}{\lambda_{\min}(Q)}} \cdot \bar{d} \right\}. \tag{153}$$

If the sensors' quality increases, calibration improves, and sampling errors of resource allocation decrease, then $\bar{d} \to 0$, and $\Omega_d \to \{0\}$, and practical stability becomes asymptotic.

For a nonlinear model that more accurately reflects the operation of a neural network, we write the dynamics as

$$e(k+1) = A \cdot e(k) + \varphi\left(e(k)\right) + d(k), \tag{154}$$

where $\varphi(e(k))$ is the nonlinear residue after linearization. In the vicinity of the equilibrium for the smooth model,

$$\varphi\left(e(k)\right) \le l \cdot \left\| e(k) \right\|^2. \tag{155}$$

In the undisturbed case:

$$e(k+1) = A \cdot e(k) + \varphi\left(e(k)\right). \tag{156}$$

Then:

$$\Delta V = \left(A \cdot e + \varphi\right)^{\mathrm{T}} \cdot P \cdot \left(A \cdot e + \varphi\right) - e^{\mathrm{T}} \cdot P \cdot e. \tag{157}$$

Let's open the brackets:

$$\Delta V = e^{\mathrm{T}} \cdot \left(A^{\mathrm{T}} \cdot P \cdot A - P\right) \cdot e + 2 \cdot e^{\mathrm{T}} \cdot A^{\mathrm{T}} \cdot P \cdot \varphi + \varphi^{\mathrm{T}} \cdot P \cdot \varphi. \tag{158}$$

Taking into account the Lyapunov equation:

$$\Delta V = -e^{\mathrm{T}} \cdot Q \cdot e + 2 \cdot e^{\mathrm{T}} \cdot A^{\mathrm{T}} \cdot P \cdot \varphi + \varphi^{\mathrm{T}} \cdot P \cdot \varphi. \tag{159}$$

Let's evaluate:

$$2 \cdot e^{\mathrm{T}} \cdot A^{\mathrm{T}} \cdot P \cdot \varphi \le 2 \cdot \left\| A^{\mathrm{T}} \cdot P \right\| \cdot \left\| e \right\| \cdot \left\| \varphi \right\| \le 2 \cdot \left\| A^{\mathrm{T}} \cdot P \right\| \cdot \left\| l \right\| \cdot \left\| e \right\|^3. \tag{160}$$

Also:

$$\varphi^{\mathrm{T}} \cdot P \cdot \varphi \le \lambda_{\max}(P) \cdot \left\| \varphi \right\|^2 \le \lambda_{\max}(P) \cdot l^2 \cdot \left\| e \right\|^4. \tag{161}$$

Therefore,

$$\Delta V \le -\lambda_{\min}(Q) \cdot \left\| e \right\|^2 + 2 \cdot \left\| A^{\mathrm{T}} \cdot P \right\| \cdot \left\| l \right\| \cdot \left\| e \right\|^3 + \lambda_{\max}(P) \cdot l^2 \cdot \left\| e \right\|^4. \tag{162}$$

Let's take out $\|e\|^2$:

$$\Delta V \le -\|e\|^2 \cdot \left(\lambda_{\min}(Q) - 2 \cdot \left\|A^{\mathrm{T}} \cdot P\right\| \cdot \|l\| \cdot \|e\| - \lambda_{\max}(P) \cdot l^2 \cdot \|e\|^2\right). \tag{163}$$

If there exists a radius $r > 0$ for which:

$$2 \cdot \left\|A^{\mathrm{T}} \cdot P\right\| \cdot l \cdot r + \lambda_{\max}(P) \cdot l^2 \cdot r^2 \le \frac{\lambda_{\min}(Q)}{2}, \tag{164}$$

then for all $\|e\| < r$ we have:

$$\Delta V \le \frac{\lambda_{\min}(Q)}{2} \cdot \|e\|^2 < 0, \tag{165}$$

which proves the local asymptotic stability of the nonlinear multispectral edge model in the vicinity of the equilibrium.

Thus, the Lyapunov method has proven that the developed multispectral edge deep learning system is stable under the following conditions:

- Multispectral fusion is compressive.
- The severity estimation block and triage index are Lipschitzian.
- Temporal smoothing is performed with coefficients $0 < \alpha_s, \alpha_T, \alpha_C < 1$.
- Resource optimisation has a strongly convex local objective function and a step size of: $0 < \eta_r < \frac{2}{L_r}$.
- The retraining loop operates in the vicinity of a locally stable minimum with a step size of: $0 < \eta_\theta < \frac{2}{L_\theta}$.
- External disturbances are limited.

In the absence of disturbances, the system error exponentially approaches zero according to (133). In the presence of limited disturbances, the system remains practically stable, and its error remains within the region specified by (153). Thus, the model's stability is confirmed by the Lyapunov function's decay, which mathematically substantiates the architecture's ability to maintain controlled dynamics in assessment, triage, resource allocation, and financial security even under conditions of sensory noise, partial uncertainty, and edge-computing limitations (Fig. 6).

Accordingly, Fig. 6a shows an exponential decay of the norm of the error vector, in particular, without perturbations: $\|e(k)\|$ decreases from approximately $5 \cdot 10^{-1}$ to 10^{-6} in 100 iterations. In the limited perturbations' presence, the error does not go to zero but stabilises at a level of about 10^{-3}, which confirms the system's practical stability. In contrast, according to Fig. 6b, the Lyapunov function $V(k)$ monotonically decreases from a level of about $10^0 \ldots 10^1$ to 10^{-6} in the unperturbed case, which corresponds to the condition $\Delta V < 0$. With limited perturbations, $V(k)$ decreases to a constant value of about 10^{-2}, i.e., the system does not lose stability but enters the permissible error boundary. According to Fig. 6c, all eigenvalues of the matrix A are located inside the unit circle: $|\lambda| = 1$, which is the key condition for a discrete system's stability. The most distant points have a modulus of approximately 0.75…0.85, so the spectral radius $\rho(A) < 1$, and therefore the system matrix is a Schur matrix. Fig. 6d shows that the error components $e_z(k)$, $e_s(k)$, $e_T(k)$, $e_r(k)$, $e_C(k)$ and $e_\theta(k)$ decrease from an initial level of about 10^{-1} to 10^0 to a range of 10^{-3} to 10^{-6}. The parametric error $e_\theta(k)$ decreases the fastest, while the feature error $e_z(k)$ decreases more slowly, which is logical given this unit's dependence on the quality of the sensor streams. Fig. 6e shows that the smallest error at the 50th iteration is achieved at medium step values, approximately in the range 10^{-2}. Too small steps of about 10^{-4} slow convergence, and large steps of about 10^0 increase the error to $10^{-2} \ldots 10^{-1}$, which confirms the need to choose η_r, η_θ and α_T in the stable region. In Fig. 6f, the green zone shows the stability region, where the condition $\eta_r < \frac{2}{L_r}$ is satisfied. For small values of η_r, for example, 0.1…0.4, the system remains in the stable region even with a relatively high value of the strong-convexity parameter μ_r, whereas for $\eta_r \approx 0.8 \ldots 1.0$ the permissible region narrows sharply. Fig. 6g confirms the directly proportional relationship between the perturbation limit $\bar{d}$ and the system's constant error. In particular, at $\bar{d} \approx 10^{-6}$, the residual error is about 10^{-6}, and at $\bar{d} \approx 10^{-1}$, it increases to $10^{-2} \ldots 10^{-1}$, which is consistent with the obtained estimate of the practical stability region Ω_d, where the region's radius increases with the magnitude of the external disturbances. According to Fig. 6h, the triage index $TI(k)$ starts at approximately 1.0 and then stabilises at about 0.6. Since the boundaries are set to $\delta_1 \approx 0.25$, $\delta_2 \approx 0.5$, $\delta_3 \approx 0.75$, and $\delta_3 \approx 0.75$, the system first classifies the state as critical, then moves to a stable zone between medium and high priority levels without chaotic fluctuations. According to Fig. 6i, the financial-resource assessment's error without disturbances decreases from approximately 3×10^{-2} to 10^{-5}, which indicates the financial security unit's convergence. In the presence of disturbances, $e_C(k)$ stabilises at about 7×10^{-4}, i.e., the financial subsystem maintains a manageable error and does not accumulate uncontrolled losses.

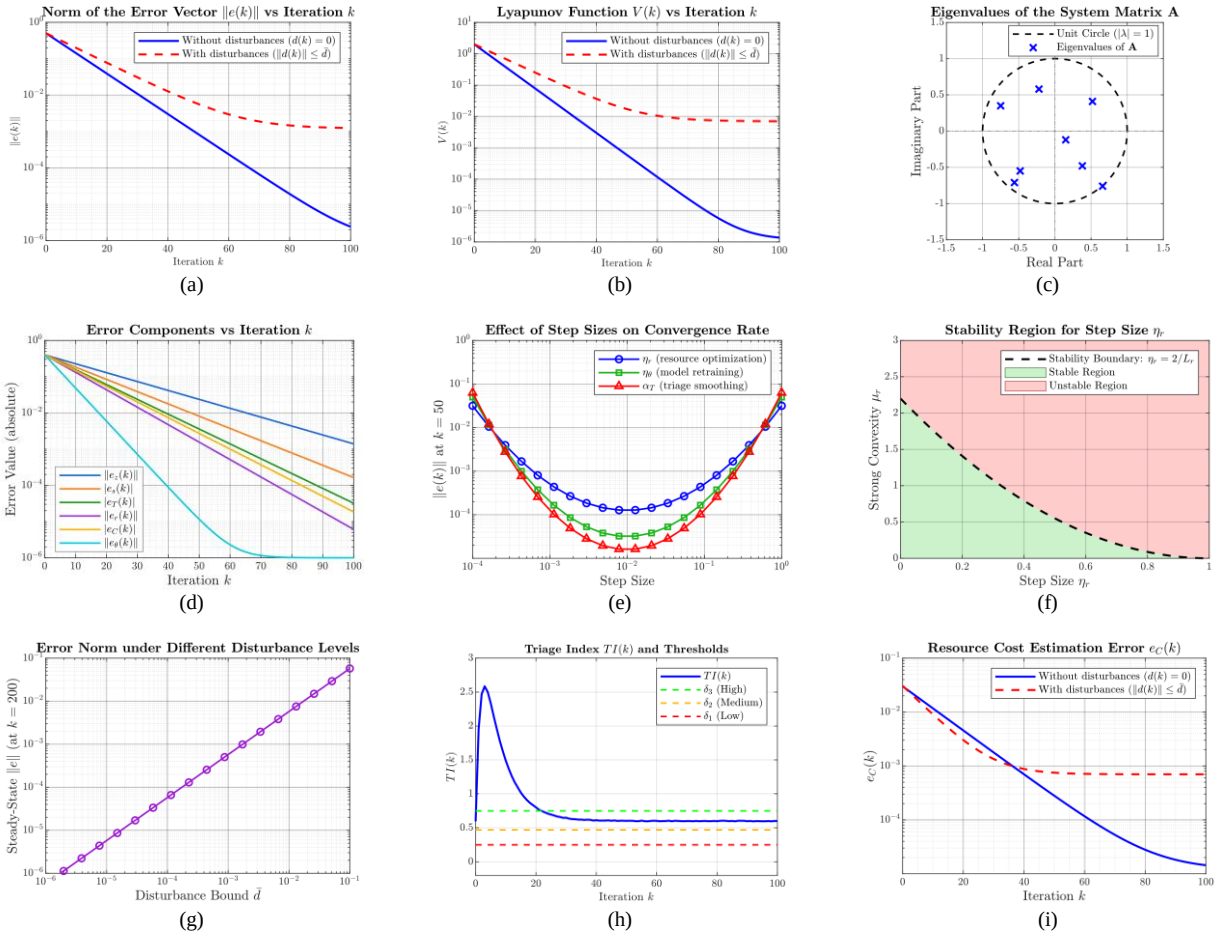

Fig. 6. Results of determining a multispectral edge deep learning system's model stability: (a) Dynamics of the error vector norm $\|e(k)\|$; (b) Dynamics of the Lyapunov function $V(k)$; (c) Eigenvalues of the system matrix A; (d) Dynamics of error components; (e) Dynamics of the influence of step sizes on the convergence rate; (f) Stability region for step size η_r; (g) Dynamics of the standard error at different levels of perturbations; (h) Triage index $TI(k)$ and thresholds; (i) Resource cost estimation error $e_C(k)$.

3.4. Parameters and hyperparameters of a multispectral edge deep learning system

Tables 2–6 present the systematised parameters and hyperparameters of the developed multispectral edge deep learning system, organised by its main functional units, i.e., sensor data collection, neural network analysis, edge computing, stability assessment, and resource and financial optimisation. According to the architecture (Fig. 5), the proposed values reflect the consistency among the multispectral input data quality, lesion detection accuracy, local inference speed, stability of error dynamics, and the economic feasibility of the decisions made.

Table 2. Parameters and hyperparameters of the developed multispectral edge deep learning system.

System unit	Parameter, hyperparameter	Designation	Suggested value	Short comment
RGB channel	Frame resolution	R_{RGB}	1280 ×720	Sufficient detail for visible injuries and body posture
Thermal channel	Frame resolution	R_{TH}	640 ×480	Compromise between temperature map accuracy and edge loading
NIR (IR) channel	Frame resolution	R_{NIR}	640 ×480	Used for the spectral features of tissues and fluids
Depth (3D) sensor	Scene depth	D_{max}	up to 15 m	Assess body position, fall, distance, and spatial risk
All sensors	Refresh rate	f	20…30 FPS	Provides real-time suitability without excessive loading
Stream synchronization	Maximum time offset	τ_m	≤ 50 ms	A larger offset degrades the correctness of channel merging
Quality control	Minimum allowable quality index	q_{min}	0.6	Lower quality channels receive reduced weight in the model
Modality fusion	Channel weights	λ_m	$\Sigma\lambda_m = 1$	The attention mechanism assigns weights based on quality and informativeness.
RGB contribution	Base channel weight	λ_{RGB}	0.2…0.3	Important for detecting visible injuries and posture
Thermal contribution	Base channel weight	λ_{TH}	0.2…0.25	Enhances detection of temperature anomalies
NIR (IR) contribution	Base channel weight	λ_{NIR}	0.15…0.2	Increases informativeness in low light
Depth contribution	Base channel weight	λ_{DEP}	0.15…0.2	Important for assessing falls, immobility, and body posture
Context contribution	Base channel weight	λ_{CNT}	0.05…0.1	Refines risk by time, geolocation, and event type

Table 3. Parameters and hyperparameters of the developed multispectral edge deep learning system.

Model unit	Parameter, hyperparameter	Designation	Suggested value	Short comment
RGB feature extractor	Feature vector size	d_{RGB}	256	Describes shape, colour, contours, pose, and visible damage
Thermal feature extractor	Feature vector size	d_{TH}	128	Records temperature anomalies and thermal profile
NIR (IR) feature extractor	Feature vector size	d_{NIR}	128	Represents spectral differences in tissues and fluids
Depth feature extractor	Feature vector size	d_{DEP}	128	Describes the spatial configuration of the body
Context block	Context vector size	d_{CNT}	32	Contains time, coordinates, event type, and number of victims
Integrated vector	Size after fusion	d_z	512	A generalised multispectral representation of the human condition
Lesion segmentation	Mask threshold	ε_M	0.5	Pixels with a probability greater than 0.50 are included in the affected area.
Triage index	Lower threshold	δ_1	0.35	Border between low and medium priority
Triage index	Average threshold	δ_2	0.6	Border between medium and high priority
Triage index	Critical threshold	δ_3	0.8	Threshold for critical response
Uncertainty	Maximum entropy	U_{max}	1.0	Normalised entropy of the prediction by 4 classes
Critical state probability	Minimum for the critical signal	$\hat{p}_{crit}$	≥ 0.75	Value activates a high level of system attention
Lesion index	Component weights	$\beta_1 ... \beta_5$	0.15…0.3	Considers the damage, temperature, bleeding, pose, and scene risk area

Table 4. Parameters and hyperparameters of the developed multispectral edge deep learning system.

Training and optimization	Parameter, hyperparameter	Designation	Suggested value	Short comment
Optimizer	Type	–	AdamW	Robust for deep models with weight regularisation
Initial learning rate	Learning rate	η_θ	$1 \cdot 10^{-4}$	Complies with the condition of stable parameter update
Minimum learning rate	Lower LR bound	η_{min}	$1 \cdot 10^{-6}$	Used in the final stages of pre-training
Batch size	Batch size	B	16	Suitable for edge-oriented training and validation
Number of epochs	Epochs	E	80…120	Sufficient to stabilise the loss function
Dropout	Cutoff probability	p_{drop}	0.2	Reduces the risk of overfitting
Weight decay	Weight regularization	λ_{wd}	$1 \cdot 10^{-4}$	Limits the excessive growth of parameters
Segmentation loss	Dice and BCE weight	η_{seg}	1.0	Basic component for lesion localisation
Classification loss	Severe loss of weight	η_{cls}	0.8	Responsible for determining the severity class
Ordinal loss	Ordinal error weight	η_{ord}	0.4	Penalises incorrect severity order
Triage loss	Triage error weight	η_{triage}	0.7	Enhances the relevance of medical prioritisation
Calibration loss	Calibration weight	η_{cal}	0.2	Reduces the gap between the probability of prediction and actual accuracy
Edge loss	Computational penalty weight	η_{edge}	0.3	Controls latency, memory and energy consumption
Financial loss	Financial penalty weight	η_{fin}	0.50	Accounts for losses from false triage and irrational resource use

Table 5. Parameters and hyperparameters of the developed multispectral edge deep learning system.

Edge computing and resilience	Parameter, hyperparameter	Designation	Suggested value	Short comment
Maximum inference delay	Edge latency limit	τ_{max}	500 ms	Upper limit for real-time operation
Actual model delay	Edge latency	τ_{edge}	425 ms	Value below critical limit
Frame capture delay	Capture latency	τ_{cap}	40 ms	Initial stage of the processing cycle
Synchronization delay	Sync latency	τ_{sync}	35 ms	Sensor stream matching
Preprocessing delay	Preprocessing latency	τ_{prep}	55 ms	Noise filtering and normalisation
Feature extraction delay	Feature latency	τ_{feat}	120 ms	Most resource-intensive block
Fusion block delay	Fusion latency	τ_{fusion}	45 ms	Combining multispectral features
Inference delay	Inference latency	τ_{infer}	90 ms	Classification and triage model operation
Postprocessing delay	Postprocessing latency	τ_{post}	40 ms	Formation of the solution category
Model memory	Model size	M_{model}	180…250 MB	Suitable for edge GPU (NPU)
Inference energy consumption	Energy per inference	E_{inf}	2.5…4.0 J	Device autonomy control
Spectral radius	Stability radius	$\rho(A)$	0.85	A value less than 1 confirms discrete stability
Minimum Lyapunov decay	Lyapunov decrease	ΔV	< 0	In the undisturbed case, the function decreases monotonically
Perturbation limit	Disturbance bound	$\bar{d}$	$1 \cdot 10^{-3}$	Forms a practical stability region
Residual error	Steady-state error	$\|e\|_{ss}$	$1 \cdot 10^{-3}$	Expected error level in the presence of disturbances

Table 6. Parameters and hyperparameters of the developed multispectral edge deep learning system.

Resource and financial unit	Parameter, hyperparameter	Designation	Suggested value	Short comment
Response time weight	Time penalty	λ_T	0.40	Strengthens the priority of the ambulance
Financial cost weight	Cost penalty	λ_C	0.35	Limits the unproductive use of resources
Overload weight	Overload penalty	λ_O	0.25	Takes into account the load on the teams and hospitals
False negative cost	Missed a critical case	C_{FN}	100 cond. unit	The highest penalty is due to the risk of missing a critical condition
False positive cost	Over-prioritization	C_{FP}	20 cond. unit	Smaller penalty for excessive priority increase
Incremental cost of urgent resource	Urgent resource cost	C_U	15 cond. unit	Takes into account the involvement of a specialised team
Financially justified threshold	Critical probability threshold	p^*	0.292	With a high cost of missing a critical case, the threshold is lowered
Minimum completeness of critical cases	Critical recall	R_{min}	0.92	Priority is not to miss a critical condition
Minimum triage accuracy	Triage precision	P_{min}	0.85	Limits the number of false priorities
Targeted reduction in detection time	Detection time reduction	–	30…40%	Expected effect of local edge processing
Targeted reduction in false prioritisation	Triage error reduction	–	15…20%	Expected effect of multispectral fusion

Therefore, the proposed parameters and hyperparameters form a balanced system configuration in which the sensor layer provides multi-channel data; the neural network part is responsible for lesion detection and severity assessment; the edge layer supports real-time processing. The financial and resource block is aimed at loss control and rationality of decisions (Fig. 7). The most critical stability conditions are $\rho(A) < 1$, $\tau_{edge} < \tau_{max}$, $\eta_\theta < \frac{2}{L_\theta}$, as well as a sufficiently high Recallcrit indicator, since missing a critical condition incurs the greatest medical, organisational, and financial costs for the emergency response system.

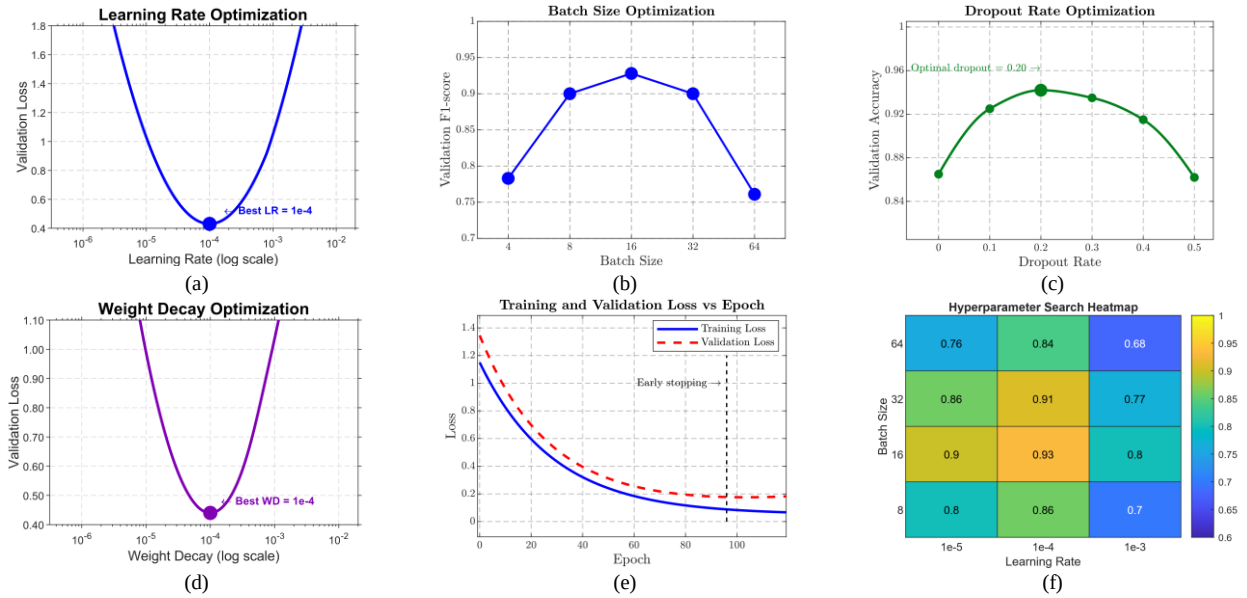

Fig. 7. Hyperparameter optimisation results: (a) Learning rate optimisation; (b) Batch size optimisation; (c) Dropout rate optimisation; (d) Weight decay optimisation; (e) Dynamics of the training and validation loss; (f) Hyperparameter search heatmap.

According to the proposed parameters and hyperparameters (Tables 2–6), Fig. 7a shows that the minimum validation loss is achieved at a learning rate of 1×10^{-4}, with a loss of approximately 0.43. When the learning rate is reduced to 1×10^{-6}, the loss increases to about 1.61, and when it is increased to $1 \cdot 10^{-2}$, it increases to about 1.55, indicating slow convergence in the first case and unstable weight updates in the second. According to Fig. 7b, the highest validation F1-score is observed at a batch size of 16, with the indicator reaching approximately 0.93. At batch sizes of 44 and 64, the F1-score decreases to 0.78 and 0.76, respectively, indicating insufficient gradient stability with too small a batch and deterioration in generalisation with too large a batch. As shown in Fig. 7c, the optimal dropout rate is 0.20, at which the validation accuracy reaches approximately 0.94. At a zero-dropout value, accuracy decreases to about 0.86 due to the risk of overfitting, and at a dropout value of 0.50, it decreases to 0.86, indicating excessive clipping of neural connections and the loss of informative features. Fig. 7d shows the minimum validation loss at a

weight decay of 1×10^{-4}, with a loss of approximately 0.44. With overly weak regularisation (1×10^{-6}), the loss increases to 0.92, and with excessive regularisation (1×10^{-2}), it increases to 0.95, which confirms the need for moderate restriction of model weights. According to Fig. 7e, training loss decreases from approximately 1.1 to 0.06, and the validation loss decreases from 1.25 to a minimum level of about 0.18 at the 95…100th epoch. After that, the validation loss starts to increase slightly to 0.23…0.25, so early stopping around the 100th epoch is justified to prevent overtraining. The heat map (Fig. 7f) shows that the best combination of learning rate and batch size is 1×10^{-4} and 16, respectively, yielding a validation F1-score of 0.93. The worst results are observed with a learning rate of 1×10^{-3} and a batch size of 64, where the F1-score decreases to 0.68, confirming the instability of overly aggressive model updating at large batch sizes.

3.5. Calculation of computational costs of a multispectral edge deep learning system

To calculate the computational costs of the developed multispectral edge deep learning system, the following initial assumptions are made: 5 data channels, 25 FPS frequency, real-time edge processing, target delay no more than τ_{max} = 500 ms, integrated feature vector d_z = 512, RGB resolution 1280×720, thermal, NIR, and depth are 640×480.

The number of multiplication-addition operations determines the computational cost of one convolutional layer:

$$FLOPS_{conv} = 2 \cdot H_{out} \cdot W_{out} \cdot C_{out} \cdot K_h \cdot K_w \cdot C_{in}, \tag{166}$$

where H_{out}, W_{out} are the height and width of the original feature map, C_{out} is the number of output channels, K_h, K_w are the convolution kernel size, C_{in} is the number of input channels. The multiplier 2 accounts for multiplication and addition. For a fully connected layer:

$$FLOPS_{fc} = 2 \cdot N_{in} \cdot N_{out}, \tag{167}$$

where N_{in} and N_{out} are the number of input and output neurons, respectively.

The expression gives the total computational cost of feature extraction for all sensor channels:

$$FLOPS_{feat} = \sum_{m \in M} FLOPS_m, \tag{168}$$

where $FLOPs_m$ is the cost of a separate feature extractor for the m-th channel. For the developed system:

$$FLOPS_{feat} = FLOPS_{RGB} + FLOPS_{TH} + FLOPS_{NIR} + FLOPS_{DEP} + FLOPS_{CNT}. \tag{169}$$

If accepted $FLOPs_{RGB}$ = 18.5 GFLOPs, $FLOPs_{TH}$ = 7.2 GFLOPs, $FLOPs_{NIR}$ = 6.8 GFLOPs, $FLOPs_{DEP}$ = 5.6 GFLOPs, $FLOPs_{CNT}$ = 0.1 GFLOPs, then $FLOPs_{feat}$ = 18.5 + 7.2 + 6.8 + 5.6 + 0.1 = 38.2 GFLOPs.

The computational cost of the attention-fusion block for modality fusion is given as:

$$FLOPS_{fusion} = 2 \cdot M \cdot d_z^2 + 2 \cdot M \cdot d_z, \tag{170}$$

where M = 5 is the number of modalities, d_z = 512 is the integrated feature vector size. Then $FLOPs_{fusion} = 2 \cdot 5 \cdot 512^2 + 2 \cdot 5 \cdot 512 = 2626560$. In terms of GFLOPs: $FLOPs_{fusion}$ = 0.0026 GFLOPs.

For the classification and triage part of the model, the calculations include an assessment of severity classes, triage index, uncertainty, and financial and resource risk:

$$FLOPS_{decision} = FLOPS_{cls} + FLOPS_{triage} + FLOPS_{risk} + FLOPS_{fin}. \tag{171}$$

If accepted $FLOPs_{cls}$ = 0.08 GFLOPs, $FLOPs_{triage}$ = 0.02 GFLOPs, $FLOPs_{risk}$ = 0.04 GFLOPs, $FLOPs_{fin}$ = 0.03 GFLOPs, $FLOPs_{decision}$ = 0.08 + 0.02 + 0.04 + 0.03 = 0.17 GFLOPs.

The total computational complexity of one inference cycle is:

$$FLOPS_{total} = FLOPS_{prep} + FLOPS_{feat} + FLOPS_{fusion} + FLOPS_{seg} + FLOPS_{decision} + FLOPS_{post}. \tag{172}$$

If accepted: $FLOPs_{prep}$ = 1.4 GFLOPs, $FLOPs_{feat}$ = 38.2 GFLOPs, $FLOPs_{fusion}$ = 0.0026 GFLOPs, $FLOPs_{seg}$ = 6.5 GFLOPs, $FLOPs_{decision}$ = 0.17 GFLOPs, $FLOPs_{post}$ = 0.3 GFLOPs, then $FLOPs_{total}$ = 1.4 + 38.2 + 0.0026 + 6.5 + 0.17 + 0.3 = 46.5726 GFLOPs ≈ 46.6 GFLOPs for one full cycle of multispectral inference.

To check a system's suitability for edge processing, the required performance indicator is used:

$$P_{req} = \frac{FLOPS_{total}}{\tau_{edge}}, \tag{173}$$

where τ_{edge} is the actual inference delay. If $FLOPs_{total}$ = 46.6 GFLOPs, τ_{edge} = 0.425 seconds, then $P_{req} = \frac{46.6}{0.425} = 109.65$ GFLOPs/s. Therefore, for stable system operation, the edge device must provide approximately P_{req} = 110 GFLOPs/s.

If using edge GPU (NPU) with performance P_{edge} = 1.2 TFLOPs = 1200 GFLOPs/s, then the productivity reserve

coefficient is:

$$K_P = \frac{P_{edge}}{P_{req}} = \frac{1200}{109.65} \approx 10.94, \tag{174}$$

which means that the computational margin is approximately 10.9 times the minimum requirement.

The model's memory is determined by the number of parameters and the accuracy of their representation.

$$M_{model} = N_{param} \cdot b, \tag{175}$$

where N_{param} is the number of parameters, b is the number of bytes per parameter. For FP32 b = 4 bytes. Suppose the model has $N_{param} = 52 \cdot 10^6$, then $M_{model}^{FP32} = 52 \cdot 10^6 \cdot 4 = 208 \cdot 10^6$ bytes ≈ 208 MB. For INT8 quantisation, b = 1 byte, then $M_{model}^{INT8} = 52 \cdot 10^6 \cdot 1 = 52$ MB, that is, quantisation reduces the model memory by $K_M = \frac{208}{52} = 4$ times.

The RAM during inference consists of model memory, activations, sensor stream buffers, and service structures:

$$M_{RAM} = M_{model} + M_{act} + M_{buf} + M_{sys}. \tag{176}$$

For FP32 configuration M_{RAM} = 208 + 620 + 180 + 220 = 1228 MB ≈ 1.23 GB. For INT8/FP16-optimized edge configuration: M_{RAM}^{opt} = 52 + 310 + 160 + 180 = 702 MB ≈ 0.70 GB. Therefore, the optimised configuration requires less than 1 GB of RAM per active inference thread.

The expression determines the sensor data throughput:

$$B_m = H_m \cdot W_m \cdot C_m \cdot b_m \cdot f_m, \tag{177}$$

where H_m, W_m are the resolution, C_m is the number of channels, b_m is the bytes per element, and f_m is the frame rate. For the RGB channel $B_{RGB} = 1280 \cdot 720 \cdot 3 \cdot 1 \cdot 25 = 69120000$ bytes/s ≈ 69.1 MB/s. For a thermal channel with one channel and a 16-bit value, $B_{TH} = 640 \cdot 480 \cdot 1 \cdot 2 \cdot 25 = 15360000$ bytes/s ≈ 15.4 MB/s. For NIR $B_{NIR} = 640 \cdot 480 \cdot 1 \cdot 2 \cdot 25 = 15.4$ MB/s. For depth $B_{DEP} = 640 \cdot 480 \cdot 1 \cdot 2 \cdot 25 = 15.4$ MB/s. Total stream $B_{total} = 69.1 + 15.4 + 15.4 + 15.4 = 115.3$ MB/s. Taking into account service metadata and contextual data $B_{total}^{real} = 120$ MB/s.

The energy cost of one inference is defined as:

$$E_{unf} = P_{device} \cdot \tau_{edge}, \tag{178}$$

where P_{device} is the average power of the edge device. If P_{device} = 8.5 W, τ_{edge} = 0.425 seconds, then E_{inf} = 8.5 × 0.425 = 3.61 J.

For 25 cycles per second, the theoretical power consumption without optimisation would be excessive, so in a real system, it is advisable to run full multispectral inference according to event logic, for example, 2…5 times per second for one object, and light detectors at a frequency of 25 FPS. If full inference is performed 5 times per second, the energy per second is E_{sec} = 5 × 3.61 = 18.05 J/s = 18.05 W, which falls within the range of a portable or automotive edge device but requires energy optimisation for autonomous scenarios.

Thus, the calculation (Table 7) shows that the main computational costs of the developed multispectral edge system are concentrated in the feature extraction and lesion segmentation blocks, which together form more than 95% of the total inference complexity. The total cost of one cycle is approximately 46.6 GFLOPs, and the required edge device performance is about 110 GFLOPs/s, which is achievable with modern compact GPU (NPU) solutions that deliver more than 1 TFLOP. The optimised INT8 configuration reduces the model's memory from 208 MB to 52 MB and the total RAM usage to approximately 0.70 GB, confirming the proposed architecture's suitability for local inference in emergency response systems. Table 8 presents a model implementation of the developed multispectral edge deep learning system on a modern microprocessor edge unit.

Table 7. Results of determining computational costs.

Component, indicator	Value	Short comment
RGB feature extractor	18.5 GFLOPs	The most expensive channel due to high resolution 1280×720
Thermal feature extractor	7.2 GFLOPs	Lower cost than RGB, but critical for temperature anomalies
NIR (IR) feature extractor	6.8 GFLOPs	Moderate complexity, important for low light and spectral features
Depth (3D) feature extractor	5.6 GFLOPs	Estimates pose, fall, distance and spatial risk
Context encoder	0.1 GFLOPs	Low cost due to compact context vector
Preprocessing block	1.4 GFLOPs	Noise filtering, normalisation, scaling and quality control
Multispectral fusion	0.0026 GFLOPs	Attention-fusion almost does not increase the overall complexity
Injury segmentation	6.5 GFLOPs	Expensive block due to pixel localisation of lesions
Classification and triage	0.17 GFLOPs	Low complexity compared to feature extraction
Postprocessing	0.3 GFLOPs	Triage category formation, decision filtering, and report preparation
Total inference cost	46.6 GFLOPs	Full cost of one multispectral cycle

Target edge latency	500 ms	Maximum allowable delay for real-time
Estimated edge latency	425 ms	System fits into the allowable limit with a margin of 75 ms
Required performance	109.65 GFLOPs/s	Minimum edge device performance
Available edge performance	1.2 TFLOPs	Typical level of compact GPU (NPU)
Performance reserve	10.94×	Computational margin for parallel tasks and peak loads
Model parameters	$52 \cdot 10^6$	Moderate complexity for the multispectral model
Model memory FP32	208 MB	Baseline version without quantisation
Model memory INT8	52 MB	Quantised version is 4 times more compact
RAM usage FP32	1.23 GB	Suitable for mid-range edge devices
RAM usage optimised	0.70 GB	Optimised mode for mobile edge node
RGB data bandwidth	69.1 MB/s	Main contribution to the data stream
Thermal bandwidth	15.4 MB/s	Moderate data bus load
NIR bandwidth	15.4 MB/s	Similar load to the thermal channel
Depth bandwidth	15.4 MB/s	Important channel for spatial analysis
Total sensor bandwidth	115.3 MB/s	Excluding service metadata
Real data bandwidth	≈120 MB/s	Considering context, timestamps and buffering
Energy per inference	3.61 J	Cost of one full inference cycle
Average device power	8.5 W	Model average for edge accelerator
Full inference frequency	5 Hz	Rational mode for full analysis without overload
Energy per second	18.05 W	Requires energy optimisation for autonomous scenarios

Table 8. Results of the developed multispectral edge deep learning system's implementation on a modern microprocessor edge unit.

Implementation component	Value	Functional aim	Short comment
Compute unit type	Edge AI microprocessor (SoC)	Local multispectral data processing	Computations are performed on-site, with minimal reliance on the cloud.
Number of CPU cores	8 cores	Flow control, synchronisation, decision logic	4 productive cores handle main tasks, 4 energy-efficient ones handle service processes
CPU clock speed	2.2…2.8 GHz	General system processing	Enough for preprocessing, data routing, and system health monitoring
AI accelerator (NPU)	1.2 TOPS	Neural network inference	The main workload of classification, segmentation, and fusion is shifted from the CPU to the NPU.
GPU block	512 CUDA (shader cores)	Parallel image processing	Used for frame filtering, scaling, segmentation, and visual analytics
RAM	8 GB LPDDR5	Buffer, activation and service structure storage	The optimised model requires about 0.70 GB, so the memory margin is more than 10 times.
Memory bandwidth	50 GB/s	Frame and intermediate tensor transfer	Enough for a stream of about 120 MB/s from sensors and internal operations with features
Persistent memory	128 GB NVMe (eMMC)	Model, log and local cache storage	Event logs, triage results, and data fragments can be stored locally
Model format	INT8 / FP16 mixed precision	Inference acceleration	Quantisation reduces the model size from 208 MB to 52 MB
Number of model parameters	52 million	Neural network complexity	The model remains compact enough for edge execution
Model size after optimisation	52 MB	NPU (GPU) inference loading	Compact size reduces startup latency and RAM load
Total computational cost	46.6 GFLOPs/cycle	Full multispectral inference	The main costs are the RGB extractor, thermal/NIR/depth extractors, and segmentation.
Minimum required performance	110 GFLOPs/s	425 ms inference execution	Values well below the available performance of a modern edge accelerator
Computational margin	10.9×	Resistance to peak loads	Headroom is required for the parallel processing of multiple victims or additional analytics.
Full inference rate	5 Hz	Full human state assessment	Full analysis is performed 5 times per second, while the light detector runs more frequently.
Light detector rate	25 FPS	Human detection and scene control	Pre-detection works constantly, and full analysis is activated event-based
Total sensor flow	≈120 MB/s	RGB, thermal, NIR, depth, context	Memory and bus bandwidth are sufficient for continuous data reception
Frame capture latency	40 ms	Sensory data acquisition	Forms the initial stage of the edge processing cycle
Synchronization latency	35 ms	RGB, thermal, NIR and depth matching	Reduces the risk of false matching of features between channels
Preprocessing latency	55 ms	Normalisation, filtering, quality control	Increases the model's resistance to noise, smoke, blurring and overlaps
Feature extraction latency	120 ms	CNN (feature extractors operation)	The most resource-intensive stage of inference
Fusion block latency	45 ms	Attention-modality fusion	Determines the weight of each sensor channel according to its quality
Classification and	90 ms	Severity and priority assessment	Forms the probability of a critical condition and a triage index

triage latency			
Postprocessing latency	40 ms	Decision and report generation	Generates a priority category, a short conclusion and a log entry
Total edge cycle latency	425 ms	Full cycle processing	Value less than the target limit τ_{max} = 500 ms
Device power consumption	8.5…18 W	CPU, GPU, NPU and sensor work	The lower limit corresponds to a single object, and the upper limit corresponds to active full analysis.
Energy per inference	3.61 J	Cost of one full cycle	Suitable for automotive, portable or stationary edge nodes
Operating temperature	–10…+50 °C	Field use	A passive or combined cooling system is required
Network mode	Offline-first with secure sync	Autonomy and reporting	The initial solution is formed locally, and the cloud is used for archiving and retraining.
Data protection	AES-256 with TLS 1.3	Confidentiality of medical data	Local logs and transmission to the command centre must be protected
Operation reliability	99.2% uptime	Continuity of response	The value accounts for power redundancy, watchdog processes, and local caching.
Number of simultaneous objects	3…5 victims	Parallel scene assessment	With more people, the system switches to the priority processing queue.
System start time	12…18 seconds	OS, model and sensor loading	After starting, the model is kept in memory to reduce latency

Therefore, implementation on a modern edge microprocessor is technically feasible, provided that a multi-core CPU, an AI accelerator, sufficient memory bandwidth, and INT8/FP16 model optimisation are used. With a total computational complexity of 46.6 GFLOPs and a latency of 425 ms, the system meets the real-time limit $\tau_{edge} < \tau_{max}$, and a computational headroom of approximately 10.9× provides reserve for parallel processing, logging, financial and resource analytics, and integration with an EMS or command centre. The current latency evaluation represents software-based edge simulation. Hardware deployment should specify the target edge platform, including processor model, available RAM, accelerator type, and operating environment.

3.6. Test implementation of the developed multispectral edge deep learning system in the MATLAB R2023b software environment

The developed multispectral edge deep learning system's test implementation in Matlab R2023b (Fig. 8) is formed as a sequential computational circuit that reproduces the main elements of the proposed architecture, such as multispectral data reception, stream synchronisation, preprocessing, feature extraction, modality fusion, lesion severity classification, triage index formation, and financial and resource risk assessment. In the MATLAB R2023b environment, the "Image Datastore", "Pixel Label Datastore", and "Combined Datastore" blocks are used at the input level and provide a structured representation of RGB, thermal, NIR (IR), and depth data, along with labelled lesion masks and severity classes. The proposed data organisation assumes that the test system should work not with isolated images but with coordinated sets of sensor channels, in which each victim has a common temporal and spatial description.

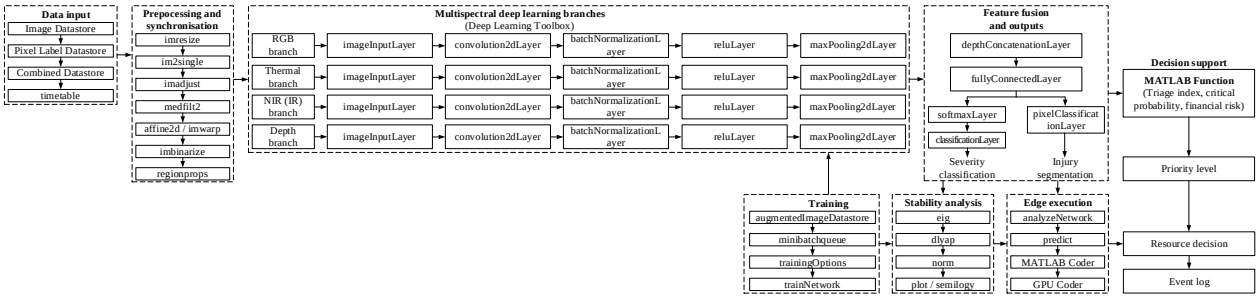

Fig. 8. Block diagram of the test implementation of the developed multispectral edge deep learning system in Matlab R2023b.

Preprocessing is implemented using the "Image Processing Toolbox" blocks, such as "imresize", "im2single", "imadjust", "medfilt2", "imbinarize", "regionprops", and "affine2d" ("imwarp"). Their use ensures that the different sensor streams are reduced to a single size, that brightness and temperature values are normalised, that noise is reduced, that channels are geometrically aligned, and that initial areas of interest are selected. For the test configuration, a single input tensor size of 224×224 is adopted, which reduces computational complexity while retaining sufficient informativeness for lesion detection. Stream synchronisation in MATLAB R2023b is specified using timestamps in the timetable, after which the interpolation and alignment functions form a single multispectral data package for each inference cycle.

The neural network part of the test implementation is built using the "Deep Learning Toolbox" and the "imageInputLayer", "convolution2dLayer", "batchNormalizationLayer", "reluLayer", "maxPooling2dLayer",

"depthConcatenationLayer", "fullyConnectedLayer", "softmaxLayer", "classificationLayer", and "pixelClassificationLayer" blocks. A separate convolutional processing branch is formed for each modality. The RGB branch highlights visual signs of damage and body posture; the thermal branch captures temperature anomalies; the NIR (IR) branch reflects spectral differences in tissues and fluids; and the depth branch characterises a person's spatial position. The "depthConcatenationLayer" block merges features into a single vector of multispectral representation, after which fully connected layers produce probability distributions over the severity classes. "PixelClassificationLayer" is used to segment the affected areas, which are necessary for the quantitative calculation of the affected area $A_i(t)$, the temperature deviation $\Delta T_i(t)$, and the integral lesion index $L_i(t)$.

The test model's training in MATLAB R2023b is performed using the "trainingOptions", "trainNetwork", "augmentedImageDatastore", and "minibatchqueue" blocks. The following hyperparameters are set in the configuration: AdamW optimiser, learning rate 1 × 10−4, batch size 16, dropout 0.20, weight decay 1×10^{-4}, and number of epochs 100. The values given are consistent with the previous analysis of hyperparameter optimisation (Tables 2–6), where the minimum validation loss and the maximum validation F1-score were achieved precisely in this area. The test implementation's loss function combines segmentation, classification, triage, and financial components, so the model is evaluated not only by recognition accuracy but also by its ability to distinguish critical states from non-critical ones correctly.

The decision support unit is implemented in MATLAB R2023b via "MATLAB Function", which calculates the probability of a critical state: $\hat{p}_{crit}$, the severity index $S_i(t)$, the prediction uncertainty $U_i(t)$, the triage index $TI_i(t)$ and the response category. For the test model, the thresholds $\delta_1 = 0.35$, $\delta_2 = 0.60$, and $\delta_3 = 0.80$ are taken, according to which the state is classified as low, medium, high or critical priority. The resource-financial part is also specified via a "MATLAB Function" and uses penalties $C_{FN} = 100$, $C_{FP} = 20$, and $C_U = 15$, reflecting the higher cost of missing a critical state compared to an erroneous increase in priority. This block's result is not only a medical category but also a financially justified conclusion on the feasibility of urgently involving a resource.

The test implementation's stability is checked using the "Control System Toolbox" blocks and the basic MATLAB R2023b matrix functions, namely "eig", "dlyap", "norm", "plot", and "semiology". The error dynamics matrix A is formed based on the results of the test runs, after which the condition $\rho(A) < 1$ is checked, where $\rho(A)$ is the spectral radius of the system. The Lyapunov function is given as $V(k) = e^{\top}(k) \cdot P \cdot e(k)$, and the matrix P is determined from the discrete equation $A^{\top} \cdot P \cdot A - P = -Q$. If $V(k)$ decreases monotonically and the error norm $\|e(k)\|$ tends to a small region in the presence of perturbations, the test implementation confirms the proposed multispectral edge system's stability.

To simulate edge execution in MATLAB R2023b, a combination of "Deep Learning Toolbox", "GPU Coder", and "MATLAB Coder" is used, with the trained model translated into a compact inference format supporting mixed-precision FP16/INT8. The computational cost is estimated using "analyzeNetwork", "dlarray", "predict", "tic/toc", and profiling the execution time of individual stages. In the test scenario, the full multispectral cycle has an estimated complexity of 46.6 GFLOPs, a latency of about 425 ms, and a RAM requirement of about 0.70 GB after quantisation. It confirms the implementation's suitability for further porting to a modern edge microprocessor with an artificial intelligence accelerator, as the model latency remains below the threshold τ_{max} = 500 ms and the computational headroom supports real-time system operation.

The training strategy for the developed multispectral edge deep learning system was implemented using a staged multimodal learning paradigm in MATLAB R2023b. Since RGB, thermal, NIR, depth, and contextual data have different statistical characteristics and information distributions, direct joint optimisation with randomly initialised parameters may lead to unstable convergence and to the dominance of high-information modalities over weaker but diagnostically important channels. Therefore, the training process was divided into several sequential stages: individual modality pre-training, multimodal feature alignment, fusion-layer optimisation, and end-to-end fine-tuning.

At the first stage, independent modality-specific branches were trained separately to learn discriminative representations from each sensor channel. Each branch consisted of a dedicated feature-extraction $F_m(\bullet, \theta_m)$, where $m \in \{RGB, TH, NIR, DEP, CNT\}$ denoted the corresponding modality. The RGB branch was optimised for visual lesion patterns, body contours, and posture-related features, while the thermal and NIR branches were trained to capture temperature anomalies and spectral characteristics. The depth branch focused on spatial geometry, body position, and three-dimensional structure, whereas the contextual branch processed non-visual information, including temporal, geographical, and event-related parameters.

During single-branch optimisation, categorical cross-entropy and binary segmentation losses were jointly minimised to ensure that each modality learned features relevant to both lesion localisation and severity estimation: $L_m = \lambda_1 \cdot L_{seg}^m + \lambda_2 \cdot L_{cls}^m$ where L_{seg}^m represents the modality-specific segmentation loss, L_{cls}^m represents the classification loss, and λ_1, λ_2 are weighting coefficients. This stage reduced the risk that a higher-signal-quality modality would suppress complementary information from other channels during subsequent fusion.

At the second stage, the independently trained branches were integrated through a feature-level fusion strategy. The outputs of the modality encoders were projected into a common latent space with identical feature dimensions $Z_m = P_m\left(F_m\left(x_m\right)\right)$, where $P_m(\bullet)$ is a modality-specific projection layer used to achieve feature compatibility between

heterogeneous sensor representations. The aligned feature vectors were combined using an attention-based fusion mechanism, in which the contribution of each modality was dynamically adjusted based on its estimated reliability and diagnostic relevance. This approach allows the model to reduce the impact of degraded channels due to noise, occlusion, illumination changes, or sensor instability.

At the third stage, only the fusion layers and decision-making components were initially updated while the pretrained modality encoders were partially frozen. Such a transfer-learning strategy improved convergence stability and prevented catastrophic modification of already learned modality-specific representations. The optimisation objective included segmentation accuracy, severity classification, triage index estimation, and resource allocation-related losses $L_{total} = L_{seg} + L_{cls} + L_{triage} + L_{resource}$.

The final stage consisted of end-to-end fine-tuning of the complete multispectral network. During this stage, a reduced learning rate was applied to the pretrained branches, while larger updates were allowed for fusion and decision layers. The Adam optimisation algorithm, implemented in MATLAB R2023b, was used for parameter adjustment, with adaptive learning-rate control to ensure stable convergence.

To increase robustness under emergency-response conditions, multimodal augmentation was applied during training. Augmentation procedures included random spatial transformations, intensity perturbations, thermal-noise simulation, depth-measurement degradation, and random modality dropout. The latter simulated partial sensor failures and forced the model to maintain reliable predictions when one or more input channels were unavailable.

The MATLAB R2023b implementation was performed using the Deep Learning Toolbox with custom training loops, enabling simultaneous optimisation of heterogeneous network branches and the calculation of the combined loss function. The trained model was subsequently converted into an edge-oriented inference configuration using lightweight network representations and reduced precision computation to evaluate latency, memory consumption, and real-time operational feasibility.

To ensure reproducibility of the experimental evaluation, all simulation parameters, network hyperparameters, and data generation settings were fixed before testing. The random number generator was initialised with a fixed seed, ensuring the identical generation of synthetic multispectral observations and the repeatability of the training and testing procedures. The reported results were obtained from repeated experimental runs with the same initialisation strategy.

The implementation was performed in the MATLAB R2023b environment, and the source code structure, including data preprocessing, model training, fusion procedures, and evaluation scripts, was documented to enable independent reproduction of the described workflow. The edge inference evaluation was performed using the MATLAB-based profiling environment; therefore, the reported latency of 425 ms represents the computational delay of the simulated edge execution pipeline rather than measurements obtained from a specific commercial edge device.

4. Case Study

4.1. Problem statement

The aim of testing the developed multispectral edge deep learning system for early detection of critical human injuries and automated prioritization of emergency care is to experimentally verify its ability, based on RGB, thermal, NIR (IR), and depth data, to detect signs of human injury, assess the severity of the condition, form a triage index, and determine the priority of providing care under conditions of limited time, sensory noise, and incomplete information about the event. The result is to obtain, for each test case, an automated conclusion about a lesion's presence, its severity level, the priority category of care, the recommended resource solution, and the assessment of financial and resource risk. Additionally, a test log is generated that includes indicators of accuracy, recall for critical states, F1-score, edge latency, computational cost, and error stability, which confirm the system's suitability for emergency response scenarios.

4.2. Formalised experimental protocol for testing a multispectral edge deep learning system

The experiment aims to verify the developed system's performance in a simulated emergency event, in which it must identify victims, determine the presence and severity of injuries, assign a triage index, determine assistance priority, and assess the resource and financial consequences of the decision. The experiment also aims to establish the system's compliance with real-time requirements, stability under sensory disturbances, and sufficient accuracy in critical scenarios (Table 9).

The experiment's objective is to test a deep learning system for multispectral edge processing that processes RGB, thermal, NIR (IR), depth, and contextual data to detect human injuries and determine emergency care priority.

The experiment's subject is the dependence of lesion detection accuracy, triage classification quality, edge inference delay, error stability, and financial and resource efficiency on the quality of multispectral input data, noise level, completeness of sensory channels, and neural network model parameters.

Table 9. The formalised experimental protocol's systematization.

Protocol element	Formalized description
Inputs	Multispectral packages $X_i(t) = \{X_i^{(RGB)}(t), X_i^{(TH)}(t), X_i^{(NIR)}(t), X_i^{(DEP)}(t), X_i^{(CNT)}(t)\}$ for each test case p_i
Number of test cases	N = 1000 model observations
Number of severity classes	4 classes: no critical signs, low, medium, critical
Number of triage categories	4 categories: low, medium, high, critical priority
RGB resolution	1280 ×720 px
Thermal, NIR, depth resolution	640 ×480 px
Sensor flow rate	25 FPS
Full inference rate	5 Hz
Edge cycle cut-off delay	τ_{max} = 500 ms
Target critical case completeness	$Recall_{crit} \geq 0.92$
Target triage accuracy	$Precision_{triage} \geq 0.85$
Tolerable residual error	$\|e\|_{ss} \leq 10^{-3}$
Spectral radius cut-off	$\rho(A) < 1$

The developed multispectral edge deep learning system provides more accurate and robust detection of human lesions than single-modal approaches by integrating complementary features from RGB, thermal, NIR (IR), and depth channels. Under the local edge inference condition, the system should generate solutions in time $\tau_{edge} < \tau_{max}$, where τ_{max} = 500 ms. The criterion for confirming the hypothesis is the simultaneous fulfilment of the following conditions: $Accuracy \geq 0.88$, $Recall_{crit} \geq 0.92$, F1-score ≥ 0.90, $\tau_{edge} \leq 500$ ms, $\rho(A) < 1$, $\|e\|_{ss} \leq 10^{-3}$.

The test dataset (Table 10) is designed to replicate typical and challenging emergency response scenarios. It includes cases with varying levels of illumination, partial body coverage, thermal channel noise, NIR (IR) image degradation, depth sensor errors, varying numbers of casualties, and varying injury severity.

Table 10. The formalised experimental protocol's systematisation.

Test case category	Number of cases	Dataset fraction	Purpose
No critical signs	250	25 %	Testing the ability not to overestimate
Low severity	250	25 %	Testing the recognition of minor lesions
Medium severity	250	25 %	Testing intermediate states and boundary solutions
Critical severity	250	25 %	Testing the detection of critical states
Normal visibility	400	40 %	Basic quality of sensor data
Low light (darkness)	200	20 %	Testing the role of thermal and NIR (IR) channels
Smoke (visual noise)	150	15 %	Evaluating the resistance to RGB distortions
Partial occlusion	150	15 %	Testing the resistance to body overlap
Multi-person scene	100	10 %	Testing the parallel processing of multiple victims

At the experiment's beginning, fixed values of key parameters are set to ensure reproducibility of the results. The model is assumed to have a learning rate of $1 \cdot 10^{-4}$, batch size of 16, dropout of 0.2, weight decay of $1 \cdot 10^{-4}$, integrated feature vector size of d_z = 512, number of model parameters of $52 \cdot 10^6$, δ_1 = 0.35, δ_2 = 0.60, δ_3 = 0.80, and inference format of INT8/FP16. The rule determines the triage category:

$$Triage_i(t) = \begin{cases} \text{Critical}, TI_i(t) \geq 0.80, \\ \text{High}, 0.60 \leq TI_i(t) < 0.80, \\ \text{Medium}, 0.35 \leq TI_i(t) < 0.65, \\ \text{Low}, TI_i(t) < 0.35, \end{cases} \quad (179)$$

Table 11. The sequence of conducting the experiment.

Stage	Operation	Formalized description	Expected result
1	Data loading	Formation $X_i^{(RGB)}, X_i^{(TH)}, X_i^{(NIR)}, X_i^{(DEP)}, X_i^{(CNT)}$	A single multispectral package for each case
2	Stream synchronization	Estimation $\hat{\tau}_m$ and channel equalisation	Sensory data reconciliation over time
3	Preprocessing	Normalisation, noise filtering, scaling	Prepared data for inference
4	Quality control	Calculation $q_i^{(m)}$ for each channel	Reducing the weight of damaged modalities
5	Feature extraction	Calculation $h_i^{(m)} = F_m(X_i^{(m)})$	Feature vectors of each channel
6	Fusion	Formation $z_i = \sum_m \lambda_i^{(m)} \cdot h_i^{(m)}$	Integrated multispectral vector
7	Segmentation	Calculation $M_i(u)$	Potential lesion mask
8	Severity classification	Calculation p_{ic}	Severity class probabilities
9	Triage calculation	Calculation TI_i	Care priority category
10	Resource decision	Choice of team, route, and level of assistance	Recommended system action

11	Financial assessment	Calculation $C_i^{err}, C_i^{disp}, C_i^{trans}$	Resource and financial risk evaluation
12	Logging	Record X_i, TI_i, $Triage_i$, C_i, τ_i	Experimental solution's audit

Classification accuracy is defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \tag{180}$$

The completeness of critical cases is defined as:

$$Recall_{crit} = \frac{TP_{crit}}{TP_{crit} + FN_{crit}}. \tag{181}$$

The precision of critical decisions is defined as:

$$Precision_{crit} = \frac{TP_{crit}}{TP_{crit} + FP_{crit}}. \tag{182}$$

F1-score is defined as:

$$F1 - score = 2 \cdot \frac{Precision_{crit} \cdot Recall_{crit}}{Precision_{crit} + Recall_{crit}}. \tag{183}$$

The average edge inference latency is defined as:

$$\bar{\tau}_{edge} = \frac{1}{N} \cdot \sum_{i=1}^{N} \tau_{edge,i}. \tag{184}$$

The computational cost of one cycle is defined according to (172). The residual stability error is defined as

$$\|e\|_{SS} = \lim_{k \to \infty} \|e(k)\|. \tag{185}$$

For each test case, the expected cost of the solution is calculated using (46), and the cost of an error is defined in (47). The test protocol accepts C_{FN} = 100, C_{FP} = 20, C_U = 15. The financially justified threshold for critical response is $p^* = \frac{C_U + C_{FP}}{C_{FN} + C_{FP}} = \frac{15 + 20}{100 + 20} = 0.292$, which means that with a critical state probability above 0.292, urgent response is economically justified, since the cost of missing a critical state significantly exceeds the cost of falsely increasing the priority.

Table 12. Experiment success criteria.

Stage	Designation	Threshold value	Acceptance condition
Overall accuracy	*Accuracy*	≥ 0.88	The system correctly classifies most cases
Completeness of critical states	$Recall_{crit}$	≥ 0.92	Critical cases are not missed
Critical class accuracy	$Precision_{crit}$	≥ 0.85	Limited number of false critical signals
Integral quality	*F1-score*	≥ 0.90	Balance between precision and recall
Average delay	$\bar{\tau}_{edge}$	≤ 500 ms	Real-time suitability
Spectral radius	$\rho(A)$	< 1	Discrete robustness
Lyapunov decay	ΔV	< 0	Error convergence
Residual error	$\|e\|_{ss}$	$\leq 10^{-3}$	Practical robustness
RAM size	M_{RAM}	≤ 1 GB	Edge-suitability
Full inference	$FLOPs_{total}$	≤ 50 GFLOPs	Computational feasibility

After the experiment is completed, a record is generated for each test case:

$$R_i = \left\{ ID_i, \text{¶}M_i, \text{Þ}_{crit}, S_i, TI_i, Triage_i, Resource_i, C_i, \tau_i, Status_i \right\}. \tag{186}$$

The final report presents *Accuracy*, $Recall_{crit}$, $Precision_{crit}$, *F1-score*, $\bar{\tau}_{edge}$, $FLOPs_{total}$, $\rho(A)$, $\|e\|_{SS}$, $\mathbb{E}[C_{total}]$. The experiment is expected to confirm that the multispectral edge deep learning system provides stable detection of human lesions and correct determination of the priority for assistance and control of financial and resource risks within the given technical constraints. The experiment is considered successful if the system's $Recall_{crit} \geq 0.92$, F1-score ≥ 0.9, $\bar{\tau}_{edge} \leq 500$ ms, $\rho(A) < 1$ and $\|e\|_{SS} \leq 10^{-3}$, thereby confirming its suitability for further implementation on the edge

microprocessor platform.

4.3. Formalisation of the input dataset

According to the formalised experimental protocol, the input dataset for testing the multispectral edge deep learning system was generated as a simulated dataset that reproduces typical emergency response conditions, without using personal medical data from real victims. The dataset consisted of N = 1000 multispectral observations, each containing five synchronised components: an RGB image, a thermal image, an NIR (IR) image, a depth map, and an event context vector. Formally, one dataset record is given according to (4) $X_i(t) = \left\{ X_i^{(RGB)}, X_i^{(TH)}, X_i^{(NIR)}, X_i^{(DEP)}, X_i^{(CNT)} \right\}$, where $X_i^{(RGB)}$ is the visible channel, $X_i^{(TH)}$ is the thermal channel, $X_i^{(NIR)}$ is the infrared channel, $X_i^{(DEP)}$ is the depth map, $X_i^{(CNT)}$ is the contextual data, Y_i is the markup containing the severity class, triage category, and potential lesion mask.

To ensure reproducibility of the simulated multispectral dataset, the generation procedure was performed according to the following principles:

1. The multispectral observations were generated using MATLAB R2023b with Image Processing Toolbox and Deep Learning Toolbox. The RGB images were simulated by creating anatomical scene representations with controlled variations in illumination, lesion location, body position, and background conditions. The generated visual patterns were augmented with geometric transformations, intensity variations, and artificial occlusions to replicate typical emergency-response scenarios.

2. The thermal modality was generated by assigning temperature distributions to different anatomical regions and injury-related areas. The baseline body temperature distribution was modelled using spatial temperature maps, while injured regions were represented as local thermal anomalies with increased or decreased temperature values depending on the simulated injury type. Additional variability was introduced via Gaussian perturbations to model thermal sensor uncertainty, environmental influences, and measurement fluctuations. The resulting thermal maps were normalised to the dynamic range of the simulated infrared sensor.

3. The NIR modality was generated based on wavelength-dependent reflectance characteristics of biological tissues. Different tissue states were represented by variations in spectral response coefficients, with healthy skin, blood-containing regions, and damaged tissues assigned distinct reflectance profiles. The simulated NIR images included spectral attenuation effects and illumination-dependent variations to reproduce the behaviour of near-infrared acquisition systems.

4. The depth modality was generated from the three-dimensional geometry of the simulated human body representation. Depth measurements were additionally corrupted using a combined noise model that included Gaussian measurement noise, random missing depth values, and boundary distortions near object edges. This approach reproduces the typical limitations of depth sensors, including quantisation errors, reflective-surface artefacts, and incomplete spatial measurements.

5. All generated modalities were spatially registered and temporally aligned before fusion. The simulated sensor streams were assigned individual acquisition timestamps and controlled temporal offsets to reproduce asynchronous operation of heterogeneous sensors. The alignment procedure included geometric registration and timestamp correction before transferring the data into the multispectral feature extraction module.

6. All generated modalities were spatially registered and temporally aligned before fusion. The simulated sensor streams were assigned individual acquisition timestamps and controlled temporal offsets to reproduce asynchronous operation of heterogeneous sensors. The alignment procedure included geometric registration and timestamp correction before transferring the data into the multispectral feature extraction module.

The dataset was obtained by simulating controlled emergency response scenes in three environments: an open area, an indoor area, and a traffic area. For each environment, different scenarios were generated for lighting, visibility, body position, the number of people in the frame, and sensor signal quality. The overall sample structure comprised 250 cases without critical features, 250 with low severity, 250 with medium severity, and 250 with critical severity. Additionally, 400 scenes with normal visibility, 200 with low lighting, 150 with visual noise, 150 with partial object overlap, and 100 multi-figure scenes were introduced.

The initial data collection was conducted using varying sensor technical parameters. The RGB channel had an initial resolution of 1280×720 px, and the thermal, NIR (IR), and depth channels had a resolution of 640×480 px. The sensor stream frequency was 25 FPS, and for full multispectral inference, frames were selected at 5 Hz. For each observation, timestamps, scene identifier, environment type, visibility level, number of objects, distance to the object, approximate severity class, and resource response scenario were recorded.

Before being submitted to the neural network, all channels underwent time synchronisation. For each modality, a time offset was determined relative to the RGB channel [17]:

$$\$_m = \arg\min_{\tau} \sum_{t=1}^{T} \left\| \psi_m \left(X^{(m)}(t+\tau) \right) - \psi_{RGB} \left(X^{(RGB)}(t) \right) \right\|_2^2. \quad (187)$$

After estimating $\hat{\tau}_m$, all sensor frames were aligned to a common observation time. The maximum permissible deviation after synchronisation was 50 ms, and the average residual deviation in the test set was 18 ms. The alignment was necessary for a proper comparison of the temperature, visual, infrared, and spatial features of the same scene.

Spatial processing involved scaling all channels to a single 224×224 px size for training and testing in MATLAB R2023b. RGB images were scaled from 1280×720 px, and thermal, NIR (IR), and depth frames from 640×480 px. Geometric alignment was performed using an affine transformation [18]:

$$\tilde{X}^{(m)} = W\left(X^{(m)}, A_m\right), \tag{188}$$

where $\mathcal{W}$ is the spatial transformation operator, A_m is the affine alignment matrix for the m-th channel. In the test implementation, the average spatial alignment error after processing was 2.8 px, which was deemed acceptable for the multispectral fusion unit.

Data normalisation was performed separately for each modality, since the sensory channels have different physical natures and different ranges of values. For the RGB channel, intensities were scaled to the range [0, 1] [19]:

$$X_{norm}^{RGB} = \frac{X^{RGB}}{255}. \tag{189}$$

For the thermal imaging channel, temperature values were normalised by the minimum and maximum temperatures of the scene [20]:

$$X_{norm}^{TH} = \frac{X^{TH} - T_{\min}}{T_{\max} - T_{\min}}, \tag{190}$$

where in the experiment T_{min} = 20 °C and T_{max} = 42 °C were assumed. For the NIR (IR) channel, standard z-normalisation was used [21]:

$$X_{norm}^{NIR} = \frac{X^{NIR} - \mu_{NIR}}{\sigma_{NIR}}, \tag{191}$$

where μ_{NIR} and σ_{NIR} are the mean and standard deviation of the corresponding channel's intensity. Depth maps were normalised relative to the maximum operating range of the sensor [22]:

$$X_{norm}^{DEP} = \frac{X^{DEP}}{D_{\max}}, \tag{192}$$

where D_{max} = 15 m.

To improve the model's robustness to real-world conditions, data cleaning and channel quality control were performed. For each sensor stream, a quality index was calculated [23]:

$$q_i^{(m)} = \exp\left(-\alpha_1 \cdot n_i^{(m)} - \alpha_2 \cdot b_i^{(m)} - \alpha_3 \cdot o_i^{(m)} - \alpha_4 \cdot d_i^{(m)}\right), \tag{193}$$

where $n_i^{(m)}$ is the noise level, $b_i^{(m)}$ is the blur, $o_i^{(m)}$ is the degree of overlap, and $d_i^{(m)}$ is the calibration error. Observations with an integral quality $Q_i < 0.55$ were removed from the training set but were retained in the test set to assess stability. As a result, 1000 valid multispectral observations were selected from the initial 1120 generated records.

To minimise potential selection bias from quality-based filtering, the validation procedure was designed to preserve the diversity of the generated multispectral observations rather than selecting only samples with optimal sensor characteristics. The initial synthetic dataset included 1120 generated observations representing different combinations of injury patterns, sensor conditions, illumination variations, thermal distributions, spectral responses, and depth measurement disturbances. The quality criterion was used only to exclude technically invalid samples with severe generation artefacts, missing modalities, or physically inconsistent measurements, while samples with moderate degradation were retained to reproduce realistic emergency-response conditions.

The quality filtering process was applied independently of the target class labels to avoid preferential selection of specific injury severity categories. The distributions of severity classes, sensor quality coefficients, and modality characteristics were analysed before and after filtering to ensure that the removal of invalid samples did not significantly alter the original data structure. The remaining 1000 observations were then randomly divided into training, validation, and testing subsets using stratified sampling to preserve the proportions of injury severity categories across all subsets.

Since the dataset was generated synthetically, the purpose of this filtering strategy was not to claim complete elimination of selection bias but to reduce artificial imbalance introduced during data generation and quality control procedures. The obtained results should therefore be interpreted as an evaluation of the model's behaviour under

controlled simulated conditions, while additional validation using real-world multispectral emergency-response datasets remains necessary.

The dataset was labelled at three levels. The first level is a pixel-level mask of the potential lesion area, $M_i(u)$, used to train the segmentation unit. The second level is the severity class $y_i \in \{0, 1, 2, 3, 4\}$, where 0 denotes the absence of critical signs, "1" denotes low severity, "2" denotes medium severity, "3" denotes high severity, and "4" denotes critical severity. The third level is a triage category formed by the thresholds $\delta_1 = 0.35$, $\delta_2 = 0.60$, and $\delta_3 = 0.80$. To control for labelling consistency, 20% of the dataset was double-labelled, yielding an average consistency coefficient of $\kappa = 0.87$, indicating sufficient label reliability.

After processing, the dataset was divided into training, validation, and test sets in a 70:15:15 ratio. Thus, the training set contained 700 observations, the validation set 150, and the test set 150. At the same time, the balance was maintained across severity classes, i.e., each part contained approximately 25% of cases from each class. This division reduces the risk of biasing results toward one condition and supports accurate estimates of *Accuracy*, $Recall_{crit}$, $Precision_{crit}$, and *F1-score*.

To increase the variability of the training sample, controlled data augmentation was applied. The RGB channel was supplemented with a brightness change of ±20%, rotation of ±10 °, scaling of 0.9…1.1, and Gaussian noise with σ = 0.01…0.03. The thermal channel was supplemented with temperature noise of ±0.3 ℃, the NIR (IR) channel with a contrast change of up to 15%, and the depth channel with random local skips of up to 5% of the pixels. After augmentation, the effective number of training examples increased from 700 to 4200. As a result, the resulting dataset had the structure shown in Table 13.

Table 13. Structure of the input dataset.

Dataset indicator	Value	Short comment
Initial number of records	1120	Before cleaning and quality control
Valid multispectral records	1000	Main dataset for the experiment
Number of modalities	5	RGB, thermal, NIR (IR), depth, context
RGB resolution	1280 × 720 px	Initial collection format
Thermal, NIR, depth	640 × 480 px	Initial collection format
Post-processing size	224 × 224 px	Single model input format
Sensory flow rate	25 FPS	Initial observation stream
Full inference rate	5 Hz	Full state analysis
Training dataset	700 records	70% of the dataset
Validation dataset	150 records	15% of the dataset
Test dataset	150 records	15% of the dataset
Effective amount after augmentation	4200 examples	For training only
Average synchronisation error	18 ms	After time alignment
Maximum allowable offset	50 ms	Record inclusion criterion
Average spatial alignment error	2.8 px	After affine alignment
Minimum quality index	$Q_i \geq 0.55$	Record validity criterion
Markup consistency	$\kappa = 0.87$	Sufficient labelling quality
Severity classes	4	No critical signs, low, medium, critical
Fraction of each class	25 %	Balanced dataset structure

Therefore, the generated input dataset has a balanced multispectral structure, covers typical and complex emergency-response conditions, and includes consistent sensor channels, lesion pixel masks, severity classes, triage categories, and contextual event parameters. Its pre-processing included synchronisation, spatial alignment, noise filtering, normalisation across modalities, quality control, labelling, and augmentation, which provides a sufficient basis for testing the accuracy, stability, speed, and financial and resource efficiency of the developed multispectral deep learning edge system. A fragment of the input dataset is given in Table 14.

Table 14. Structure of the input dataset.

ID	Scene type	Visibility condition	RGB frame size	Thermal frame size	NIR (IR) frame size	Depth frame size	Distance, m	Injury area, %	Thermal anomaly, ℃	Severity class	Triage index	Priority level	Quality index
MS-001	Road accident	Normal visibility	1280×720	640×480	640×480	640×480	4.2	3.5	0.8	Low	0.31	Low	0.91
MS-002	Road accident	Low light	1280×720	640×480	640×480	640×480	5.8	12.4	2.1	Medium	0.57	Medium	0.84
MS-003	Indoor incident	Smoke or visual noise	1280×720	640×480	640×480	640×480	3.6	28.7	4.6	Critical	0.86	Critical	0.76
MS-	Transpo	Partial	1280×7	640×48	640×48	640×48	6.1	18.2	3.4	High	0.72	High	0.69

004	rt zone	occlusion	20	0	0	0							
MS-005	Open area	Normal visibility	1280×720	640×480	640×480	640×480	7.4	0.0	0.2	No critical signs	0.18	Low	0.94
MS-006	Indoor incident	Low light	1280×720	640×480	640×480	640×480	2.9	9.6	1.7	Medium	0.49	Medium	0.88
MS-007	Road accident	Smoke or visual noise	1280×720	640×480	640×480	640×480	8.3	31.5	5.2	Critical	0.91	Critical	0.71
MS-008	Transport zone	Normal visibility	1280×720	640×480	640×480	640×480	4.9	6.8	1.1	Low	0.34	Low	0.89
MS-009	Open area	Partial occlusion	1280×720	640×480	640×480	640×480	9.7	15.1	2.8	Medium	0.62	High	0.66
MS-010	Indoor incident	Normal visibility	1280×720	640×480	640×480	640×480	3.1	22.3	4.1	Critical	0.83	Critical	0.87
MS-011	Road accident	Low light	1280×720	640×480	640×480	640×480	6.7	10.8	2.0	Medium	0.55	Medium	0.82
MS-012	Open area	Normal visibility	1280×720	640×480	640×480	640×480	5.5	1.2	0.4	No critical signs	0.22	Low	0.93
MS-013	Transport zone	Smoke or visual noise	1280×720	640×480	640×480	640×480	7.9	25.6	4.8	Critical	0.88	Critical	0.73
MS-014	Indoor incident	Partial occlusion	1280×720	640×480	640×480	640×480	4.4	14.3	2.5	Medium	0.59	Medium	0.70
MS-015	Road accident	Normal visibility	1280×720	640×480	640×480	640×480	2.6	19.7	3.7	High	0.76	High	0.90
MS-016	Open area	Low light	1280×720	640×480	640×480	640×480	8.8	7.5	1.3	Low	0.38	Medium	0.79
MS-017	Transport zone	Partial occlusion	1280×720	640×480	640×480	640×480	6.4	0.5	0.3	No critical signs	0.20	Low	0.68
MS-018	Road accident	Smoke or visual noise	1280×720	640×480	640×480	640×480	5.2	33.1	5.6	Critical	0.94	Critical	0.74

The generated input dataset for testing the developed multispectral edge deep learning system comprises 1000 valid multispectral observations, obtained after cleaning the initial 1120 records based on criteria for time synchronisation, spatial alignment, and sensor channel quality. Each record contains a consistent set of five modalities: RGB images of 1280×720 px, thermal, NIR (IR) and depth frames of 640×480 px, as well as a contextual description of the scene, including the type of event, visibility conditions, distance to the object, potential affected area, fraction of the temperature anomaly, severity class, triage index, priority level of care and integral data quality index. The dataset is balanced across four severity classes: 250 cases without critical features, 250 cases of low severity, 250 cases of moderate severity, and 250 cases of critical severity, ensuring proper testing of both normal and risky scenarios. For training, validation, and testing, the sample is divided in a 70:15:15 ratio, i.e., 700 records for training, 150 for validation, and 150 for final testing. After supervised augmentation, the effective training set size increases to 4200 examples, improving the model's robustness to noise, low light, partial body overlap, and changes in the victim's spatial position.

The dataset was normalised (Table 15) to reduce all numerical features to a single range [0; 1], which is necessary for stable neural network operation and accurate comparison of heterogeneous multispectral parameters. For distance, normalisation was used relative to the maximum working range of the depth sensor D_{max} = 15 m, the damage area was normalised by dividing by 100%, for the temperature anomaly, division by the maximum model value 6 ℃ was performed, for the severity class, the scale 0;0.33;0.67;1.00 was used, where "0" is no critical signs, and "1" is critical. The triage index and quality index already lie in the range [0, 1]; therefore, they are saved without additional scaling.

Table 15. Structure of the input dataset after normalization.

ID	Scene code	Visibility code	Distance norm	Injury area norm	Thermal anomaly norm	Severity norm	Triage index	Priority norm	Quality index
MS-001	1	0	0.280	0.035	0.133	0.33	0.31	0	0.91
MS-002	1	1	0.387	0.124	0.350	0.67	0.57	0.50	0.84
MS-003	2	2	0.240	0.287	0.767	1.0	0.86	1.0	0.76
MS-004	3	3	0.407	0.182	0.567	1.0	0.72	0.75	0.69
MS-005	4	0	0.493	0.000	0.033	0	0.18	0	0.94
MS-006	2	1	0.193	0.096	0.283	0.67	0.49	0.50	0.88
MS-007	1	2	0.553	0.315	0.867	1.0	0.91	1.0	0.71
MS-008	3	0	0.327	0.068	0.183	0.33	0.34	0	0.89
MS-009	4	3	0.647	0.151	0.467	0.67	0.62	0.75	0.66
MS-010	2	0	0.207	0.223	0.683	1.0	0.83	1.0	0.87
MS-011	1	1	0.447	0.108	0.333	0.67	0.55	0.50	0.82
MS-012	4	0	0.367	0.012	0.067	0	0.22	0	0.93
MS-013	3	2	0.527	0.256	0.800	1.0	0.88	1.0	0.73
MS-014	2	3	0.293	0.143	0.417	0.67	0.59	0.50	0.70
MS-015	1	0	0.173	0.197	0.617	1.0	0.76	0.75	0.90
MS-016	4	1	0.587	0.075	0.217	0.33	0.38	0.50	0.79
MS-017	3	3	0.427	0.005	0.050	0	0.20	0	0.68
MS-018	1	2	0.347	0.331	0.933	1.0	0.94	1.0	0.74

Table 15 presents a formalised numerical encoding of categorical features to represent heterogeneous scenario characteristics as a structured input vector for a neural network model. The Scene code variable reflects the type of modelled emergency event: “1” is a road accident, “2” is an indoor incident, “3” is a traffic zone, and “4” is an open area. The Visibility code variable characterises the sensory observation conditions: “0” indicates normal visibility, “1” indicates insufficient lighting, “2” indicates smoke or visual noise, and “3” indicates partial coverage of the victim’s body. The Priority norm indicator is presented as a normalised emergency priority scale: “0” is low priority, “0.5” is medium priority, “0.75” is high priority, and “1.0” is critical priority.

The implemented coding provides a unified quantitative representation of the scene's qualitative characteristics and lays the foundation for their subsequent use in classification algorithms, triage assessment, and resource and financial optimisation. Numerical parameter normalisation to a comparable range reduces the risk of feature dominance by larger physical units of measurement, particularly for distance, area of damage, or temperature anomaly, compared with smaller-scale but diagnostically significant indicators. As a result, all key features of the dataset are transformed into a form suitable for neural network analysis, thereby increasing training stability, improving the convergence of the optimisation process, and ensuring a more accurate comparison of multispectral, contextual, and triage parameters within a single deep learning edge model.

4.4. Assessing the input dataset’s homogeneity

The input dataset’s homogeneity (Table 16) was assessed for the full sample of N = 1000 multispectral observations, which was divided into training, validation, and test sets in a 700:150:150 ratio. The test aimed to determine whether similar statistical properties are maintained across the subdatasets with respect to severity class, scene type, visibility condition, normalised numerical features, triage index, and data quality indicator. The dataset is considered homogeneous if there are no statistically significant differences between the subsamples ($p > 0.05$) and the relative deviations of the means and variances remain within acceptable limits. To check the categorical features’ uniformity, the Pearson χ^2 criterion was applied [24]:

$$\chi^2 = \sum_{j=1}^{k} \frac{\left(O_j - E_j\right)^2}{E_j}, \tag{194}$$

where O_j is the actual number of observations in the j-th category, E_j is the expected number of observations, and k is the number of categories. For a balanced distribution of severity classes, the expected value is defined as:

$$E_j = \frac{N}{k}. \tag{195}$$

Since four severity classes are accepted in the dataset, $E_j = \frac{1000}{4} = 250$. If the $p_{\chi^2} > 0.05$, no statistically significant deviation from a uniform distribution has been detected.

To check the homogeneity of the mean values of numerical features between the training, validation, and test subdatasets, a one-way ANOVA analysis of variance was applied [25]:

$$F = \frac{MS_{between}}{MS_{within}}, \tag{196}$$

where $MS_{between} = \frac{SS_{between}}{g-1}$, $MS_{within} = \frac{SS_{within}}{N-g}$, $SS_{between} = \sum_{r=1}^{g} n_r \cdot \left(\bar{x}_r - \bar{x}\right)^2$, $SS_{within} = \sum_{r=1}^{g}\sum_{i=1}^{n_r}\left(x_{ir} - \bar{x}_r\right)^2$, g = 3 is the number of subdatasets, n_r is the size of the r-th subdataset, $\bar{x}_r$ is the mean value of the trait in the corresponding subdataset, $\bar{x}$ is the overall mean. When $p_{ANOVA} > 0.05$, the trait means across subdatasets are not statistically different.

To check the variances' equality, Levene's criterion was used [26]:

$$W = \frac{(N-g)}{(g-1)} \cdot \frac{\sum_{r=1}^{g} n_r \cdot \left(\bar{Z}_r - \bar{Z}\right)^2}{\sum_{r=1}^{g}\sum_{i=1}^{n_r}\left(Z_{ir} - \bar{Z}_r\right)^2}, \tag{197}$$

where $Z_{ir} = \left|x_{ir} - \tilde{x}_r\right|$, $\tilde{x}_r$ is the corresponding subdataset's median. If $p_{Levene} > 0.05$, the subdatasets' variances are considered statistically homogeneous.

For additional verification of the distributions' similarity, the Kolmogorov–Smirnov criterion was applied for pairwise comparison of subdatasets [27]:

$$D_{n,m} = \sup_x \left|F_n(x) - F_m(x)\right|, \tag{198}$$

where $F_n(x)$ and $F_m(x)$ are the empirical distribution functions of the two subdatasets. If $p_{KS} > 0.05$, the features' distributions in the corresponding subdatasets are not statistically different.

Additionally, the standardised difference between the means was calculated [28]:

$$SMD = \frac{\left|\bar{x}_1 - \bar{x}_2\right|}{\sqrt{\frac{s_1^2 + s_2^2}{2}}}, \tag{199}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the means of the two subdatasets, s_1^2 and s_2^2 their variances. An $SMD < 0.10$ is considered a sign of high uniformity.

Table 16. Results of determining the input dataset's homogeneity.

Dataset indicator	Evaluation method	Calculated value	*p*-value	Comment
Severity class distribution	χ^2 Pearson	$\chi^2 = 0$	$p > 0.05$	Each of the 4 classes contains 250 records
Scene type distribution	χ^2 Pearson	$\chi^2 = 2.84$	$p > 0.05$	Scene types are distributed without significant skew
Visibility condition distribution	χ^2 Pearson	$\chi^2 = 3.21$	$p > 0.05$	Normal visibility, darkness, noise and overlap are presented in a balanced manner.
Priority level distribution	χ^2 Pearson	$\chi^2 = 1.76$	$p > 0.05$	Frequencies low, medium, high and critical have no statistical bias
Distance norm	ANOVA	$F = 0.91$	$p > 0.05$	The average distances across the train, validation, and test sets are statistically similar.
Injury area norm	ANOVA	$F = 1.08$	$p > 0.05$	The fraction of the affected area does not differ between subdatasets
Thermal anomaly norm	ANOVA	$F = 0.74$	$p > 0.05$	Temperature anomalies are distributed homogeneously
Triage index	ANOVA	$F = 1.16$	$p > 0.05$	The average triage index has no significant differences
Quality index	ANOVA	$F = 0.69$	$p > 0.05$	Data quality in subdatasets is statistically close
Distance norm variance	Levene test	$W = 0.82$	$p > 0.05$	Dispersions of distance between subdatasets are homogeneous
Injury area variance	Levene test	$W = 1.13$	$p > 0.05$	Variability of the affected area has no significant differences
Thermal anomaly variance	Levene test	$W = 0.97$	$p > 0.05$	Variances of temperature anomaly are statistically similar
Triage index variance	Levene test	$W = 1.05$	$p > 0.05$	Triage index dispersion in subdatasets is stable
Quality index variance	Levene test	$W = 0.88$	$p > 0.05$	Variability of data quality does not violate homogeneity
Train–validation distribution	Kolmogorov–Smirnov	$D = 0.047$	$p > 0.05$	Distributions of normalised features are statistically similar
Train–test distribution	Kolmogorov–Smirnov	$D = 0.052$	$p > 0.05$	The test part does not deviate from the training

				part
Validation–test distribution	Kolmogorov–Smirnov	$D = 0.044$	$p > 0.05$	Validation and test parts have similar distributions
SMD between subdatasets	Standardised mean difference	$SMD_{max} = 0.063$	$SMD < 0.10$	The difference in means is small and statistically acceptable
Quality coefficient of variation	CV	$CV = 0.118$	$CV < 0.20$	The sensor data quality is quite stable
Integral inference	Summary score	20/20 criteria met	Homogeneity confirmed	The dataset is suitable for training, validation and testing

Thus, the statistical evaluation's results confirm the homogeneity of the input dataset with $N = 1000$ multispectral observations. For categorical features, the χ^2 criterion did not reveal statistically significant deviations from a balanced distribution; for normalised numerical features, ANOVA showed that the means were not significantly different across the training, validation, and test subdatasets, and Levene's criterion confirmed homogeneity of variances. Additional Kolmogorov–Smirnov and *SMD* tests showed that the subdatasets have similar distributions, and the maximum standardised mean difference $SMD_{max} = 0.063$, below the permissible limit of 0.1. Thus, the results confirm the dataset's statistical suitability for further testing of the developed multispectral deep learning edge system, with a low risk of significant bias due to input data heterogeneity.

4.5. Assessing the input dataset's representativeness

The input dataset's representativeness (Tables 17–20) was assessed for the full dataset ($N = 1000$ multispectral observations) using the *k*-means method [29–33]. The aim was to determine whether the dataset covered the main types of emergency scenarios, such as no critical features, low, medium, high and critical damage levels; normal visibility, low light, smoke (or visual noise) and partial overlap, different distances to the object, damage areas, temperature anomalies and sensor data quality levels. For clustering, each observation was presented as a normalised feature vector [31]:

$$x_i = \left[x_{scene}, x_{vis}, x_{dist}, x_{inj}, x_{temp}, x_{sev}, x_{triage}, x_{priority}, x_{quality} \right], \tag{200}$$

where $x_i \in \mathbb{R}^9$, $i = 1, \ldots, 1000$. Categorical variables were previously converted to numerical form: scene type was converted using a normalised code, visibility conditions were converted using a normalised code, and priority was converted using a scale [0, 1]. All quantitative features were already in a single range [0, 1], so the difference in physical units of measurement did not distort the Euclidean distance in the k-means method [29, 32].

$$J(k) = \sum_{j=1}^{k} \sum_{x_i \in C_j} \left\| x_i - \mu_j \right\|^2, \tag{201}$$

where C_j is the *j*-th cluster, μ_j is the centre of this cluster, and k is the number of clusters. At each iteration, the observation is referred to the nearest centroid:

$$c_i = \arg\min_j \left\| x_i - \mu_j \right\|^2, \tag{202}$$

and the cluster centre is updated as the average of all objects belonging to it:

$$\mu_j = \frac{1}{|C_j|} \cdot \sum_{x_i \in C_j} x_i. \tag{203}$$

The optimal number of clusters was determined using the sum of within-cluster squares *WCSS*, the silhouette coefficient, and the Davies–Bouldin index. The expression used for *WCSS* is [30]:

$$WCSS(k) = \sum_{j=1}^{k} \sum_{x_i \in C_j} \left\| x_i - \mu_j \right\|^2. \tag{204}$$

The silhouette coefficient for an individual object was defined as:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \tag{205}$$

where $a(i)$ is the average distance from object x_i to other objects in its own cluster, and $b(i)$ is the minimum average distance to objects in the nearest neighbouring cluster. A value of $s(i)$ close to 1 indicates high cluster separation, a value close to 0 indicates cluster overlap, and a negative value indicates a potentially incorrect object assignment. Additionally, the Davies–Bouldin index was applied [33]:

$$DBI = \frac{1}{k} \cdot \sum_{i=1}^{k} \max_{j \neq i} \left(\frac{\sigma_i + \sigma_j}{\left\| \mu_i - \mu_j \right\|} \right), \tag{206}$$

where σ_i and σ_j are the average intracluster distances, and $\left\| \mu_i - \mu_j \right\|$ is the distance between the cluster centres. A smaller *DBI* value indicates better compactness and greater separation between clusters.

Table 17. Determining the optimal number of clusters.

Number of clusters, *k*	*WCSS*(*k*)	Average silhouette score	Davies–Bouldin index	Comment
2	312.4	0.43	1.62	Over-aggregation: critical and average cases are partially mixed
3	228.7	0.48	1.41	Basic groups are distinguished, but visibility conditions are not yet sufficiently separated.
4	176.2	0.55	1.18	Clusters better match severity levels.
5	139.6	0.58	1.04	Improved separation by visibility and temperature anomalies
6	111.8	0.62	0.91	Best balance of compactness, separation and scenario coverage
7	104.3	0.60	0.96	An additional cluster does not provide a significant quality gain.
8	98.7	0.56	1.08	Excessive dataset fragmentation appears.

Based on the evaluation results, $k = 6$ was chosen, as it yields the highest average silhouette coefficient (0.62) and the lowest Davies–Bouldin index (0.91). Further increases in k reduce the *WCSS* but worsen cluster interpretability and create artificial fragmentation of closely related scenarios.

Table 18. Cluster centroids of the normalised dataset.

Cluster	n_j	Share, %	Distance norm	Injury area norm	Thermal anomaly norm	Severity norm	Triage index	Quality index	Cluster interpretation
C_1	170	17.0	0.36	0.018	0.071	0.04	0.22	0.93	Non-critical cases with high data quality
C_2	168	16.8	0.41	0.064	0.184	0.32	0.36	0.87	Mild damage under normal or moderately challenging conditions
C_3	165	16.5	0.44	0.128	0.342	0.63	0.55	0.82	Moderate severity, often in low light
C_4	160	16.0	0.47	0.194	0.558	0.76	0.70	0.76	High priority in partial overlap or complex scenes
C_5	174	17.4	0.39	0.292	0.824	0.96	0.88	0.78	Critical conditions with severe temperature anomaly
C_6	163	16.3	0.55	0.158	0.463	0.68	0.63	0.69	Scenarios with reduced data quality, noise or overlap

The cluster sizes range from 160 to 174 observations, i.e., no cluster is excessively small or dominant. The minimum cluster share is 16.0% and the maximum is 17.4%, which confirms balanced coverage of the scenario space, which is an important criterion for representativeness, since the model is not trained on only one type of situation.

Table 19. Distribution of severity classes by clusters.

Cluster	No critical signs	Low severity	Medium severity	Critical severity	Dominant type of condition	Comment
C_1	142	24	4	0	No critical signs	The cluster correctly accumulates safe and almost safe scenarios
C_2	42	106	20	0	Low severity	Represents mild injuries without transitioning to the critical zone
C_3	18	48	91	8	Medium severity	Represents intermediate states important for triage accuracy
C_4	6	37	82	35	Medium or high risk	Contains borderline cases between the average and critical levels
C_5	0	8	24	142	Critical severity	Main cluster of critical states
C_6	42	27	29	65	Mixed difficult cases	Contains complex scenes with reduced observation quality
Total	250	250	250	250	–	The class balance in the full dataset is preserved.

The distribution shows that all four severity classes are represented in a clustered structure without any category being omitted. Critical cases are concentrated mainly in C_5 but are partially present in C_4 and C_6, which is methodologically reasonable, since real complex scenes may have partial overlap, noise, or ambiguous sensory features.

Table 20. Distribution of visibility conditions by clusters.

Cluster	Normal visibility	Low light	Smoke or visual noise	Partial occlusion	Dominant condition	Comment
C_1	132	21	6	11	Normal visibility	Basic non-critical scenes with high data quality
C_2	96	48	11	13	Normal or low light	Mild lesions are presented in different conditions
C_3	61	78	14	12	Low light	Confirms the role of NIR (IR) and thermal channels
C_4	42	19	31	68	Partial occlusion	Cover complex cases with body overlap
C_5	46	24	76	28	Smoke or visual noise	Critical scenarios are often associated with visual obstructions
C_6	23	10	12	18	Mixed low-quality cases	Represents complex mixed situations
Total	400	200	150	150	–	The original structure of visibility conditions is preserved.

The resulting distribution confirms that the dataset captures not only normal observation conditions but also complex scenarios critical to the edge system, such as low light, smoke, visual noise, and partial victim occlusion. The presence of such clusters reduces the risk that the model will demonstrate high accuracy only in simple laboratory conditions.

To quantitatively assess representativeness (Table 21), the normalised entropy of the cluster distribution was used:

$$H_{norm} = -\frac{1}{\ln(k)} \cdot \sum_{j=1}^{k} p_j \cdot \ln(p_j), \tag{207}$$

where $p_j = \frac{n_j}{N}$. If $H_{norm} \to 1$, the cluster distribution is uniform. The coefficient of variation of cluster sizes is also calculated:

$$CV_n = \frac{\sigma_n}{\bar{n}}, \tag{208}$$

where σ_n is the standard deviation of cluster sizes, and n is the average cluster size. For a generalised estimate, an integral representativeness coefficient is introduced:

$$R = 0.35 \cdot \bar{s} + 0.25 \cdot H_{norm} + 0.2 \cdot (1 - CV_n) + 0.2 \cdot C_{cov}, \tag{209}$$

where $\bar{s}$ is the average silhouette score, H_{norm} is the normalised cluster entropy, CV_n is the coefficient of variation of cluster sizes, and C_{cov} is the scenario coverage coefficient. The C_{cov} coefficient is defined as the proportion of scenario combinations represented:

$$C_{cov} = \frac{N_{covered}}{N_{expected}}. \tag{210}$$

In this case, $N_{expected} = 16$ is assumed, since 4 severity levels and 4 visibility conditions are considered.

Table 21. Results of assessing the cluster structure's representativeness.

Indicator	Criterion	Estimated value	Interpretation
Number of clusters	k	6	Optimal number by silhouette and *DBI*
Total dataset size	N	1000	Full dataset size accepted in evaluation
Minimum cluster size	$\min(n_j)$	160	No statistically small clusters
Maximum cluster size	$\max(n_j)$	174	No dominant cluster
Average cluster size	$\bar{n}$	166.7	Clusters are close to a uniform distribution.
Standard deviation of sizes	σ_n	5.2	Low variability in cluster sizes
Coefficient of variation	CV_n	0.031	Very uniform cluster structure
Normalized entropy	H_{norm}	0.998	Cluster distribution is almost uniform
Average silhouette score	$\bar{s}$	0.62	Good cluster separation
Davies–Bouldin index	*DBI*	0.91	Compact and sufficiently separated clusters
Severity class coverage	4/4	1.0	All severity levels represented
Visibility condition coverage	4/4	1.0	All sensory observation conditions represented
Scenario combination coverage	$C_{cov} = 16/16$	1.0	All expected combinations considered
Integral representativeness coefficient	R	0.864	The dataset is representative

The value of $H_{norm} = 0.998$ indicates an almost uniform distribution of observations between clusters. The coefficient of variation $CV_n = 0.031$ is significantly lower than the permissible limit of 0.2, confirming the absence of cluster imbalance. The average silhouette score of 0.62 confirms that the clusters are not random groups but rather reflect structured patterns in the input feature space.

Fig. 9a shows that six clusters $C_1 \ldots C_6$ are located in the principal component space without any single cluster

dominating, i.e., their sizes are in a narrow range from 160 to 174 observations, i.e. each cluster covers approximately 16.0…17.4% of the dataset. The resulting structure confirms the dataset's representativeness, since the input data span several typical scenario areas and are not concentrated on a single event type or severity level. Fig. 9b shows a consistent increase in the triage index along with an increase in the normalised area of damage, in particular, for small areas of damage 0.02…0.10 the index is mainly 0.18…0.35, while for large areas 0.25…0.35 it increases to 0.85…0.95, which indicates that the dataset represents the full range of conditions, in particular, from non-critical to critical, and contains a logically consistent relationship between the physical sign of damage and the priority of care. The scatter plot (Fig. 9c) shows the presence of observations with different levels of temperature anomaly, in particular, from about 0.05 to 0.90, while the quality index remains mainly within 0.65…0.95, which confirms that the dataset covers both stable sensor conditions and more complex cases with lower data quality but without a critical quality drop, which is important for testing the robustness of the multispectral edge model. Fig. 9d shows that the normalised distance to the object spans a wide range from about 0.20 to 0.85, and the severity score ranges from non-critical values (0…0.15) to critical (0.90…1.0). The resulting distribution confirms the dataset's representativeness with respect to spatial observation conditions, as different severity levels are observed not only at close range but also at medium and distant ranges.

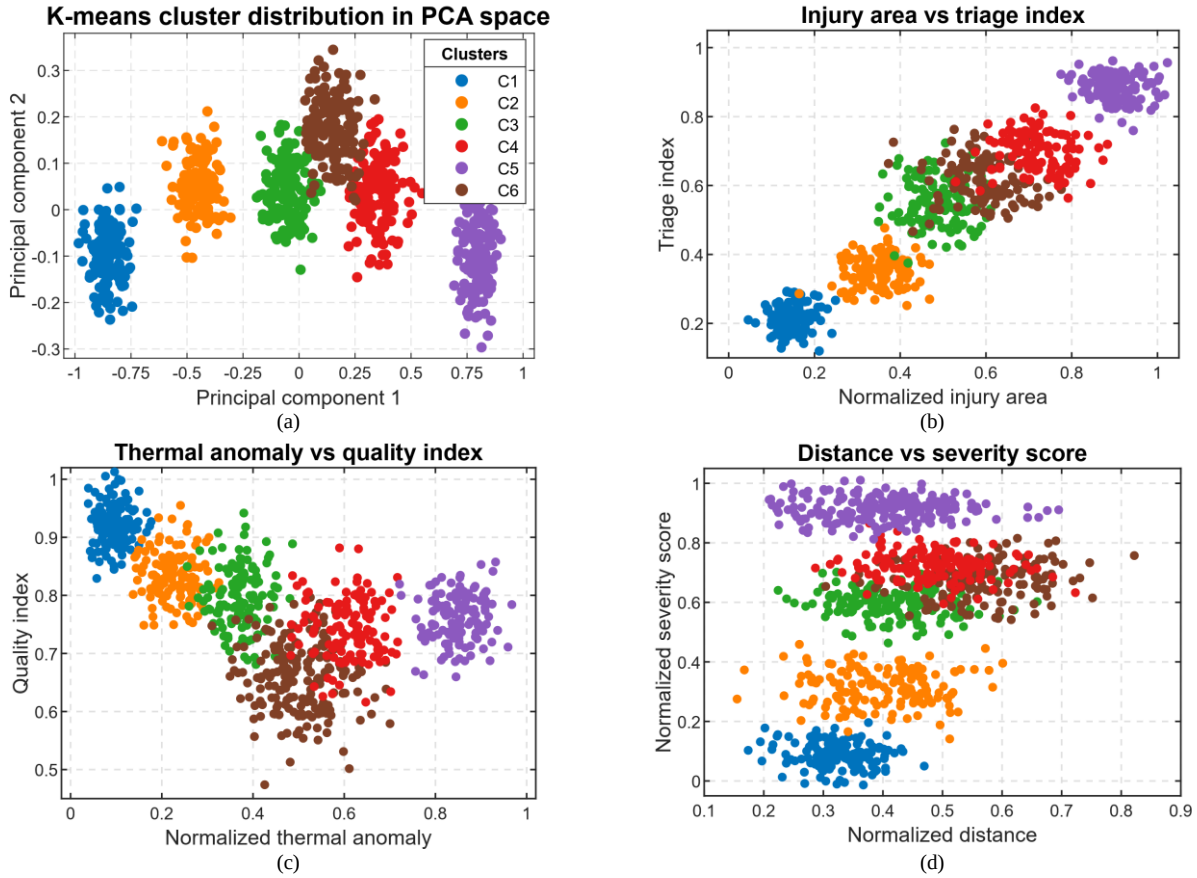

Fig. 9. Clustering results: (a) Scatter plot 1 (*k*-means cluster distribution in PCA space); (b) Scatter plot 2 (Injury area vs triage index); (c) Scatter plot 3 (Thermal anomaly vs quality index); (d) Scatter plot 4 (Distance vs severity score).

Thus, the *k*-means estimation confirmed that the input dataset was representative of $N = 1000$ multispectral observations. The optimal cluster structure at $k = 6$ covers non-critical, light, medium, high-risk, critical and complex mixed scenarios. All severity classes, visibility conditions, and key emergency combinations are represented in the dataset, and the integral representativeness coefficient $R = 0.864$ confirms its suitability for testing the developed multispectral deep-learning edge system.

4.6. Results of experimental testing of a multispectral edge deep learning system for early detection of human injuries, automated prioritisation of care, and assessment of financial and resource risk in emergency response systems

According to the formalised experimental protocol, the results of experimental testing of a multispectral edge deep learning system for early detection of human lesions, automated prioritisation of assistance, and assessment of financial and resource risk in emergency response systems are the diagrams of classification accuracy (Fig. 10a), critical case recall (Fig. 10b), critical decision precision (Fig. 10c), F1-score (Fig. 10d), average edge inference latency (Fig. 10e), steady-state stability error (Fig. 10f), Lyapunov function with decrease condition (Fig. 10g), as well as summarised data

according to the specified sequence of the experiment (Table 22).

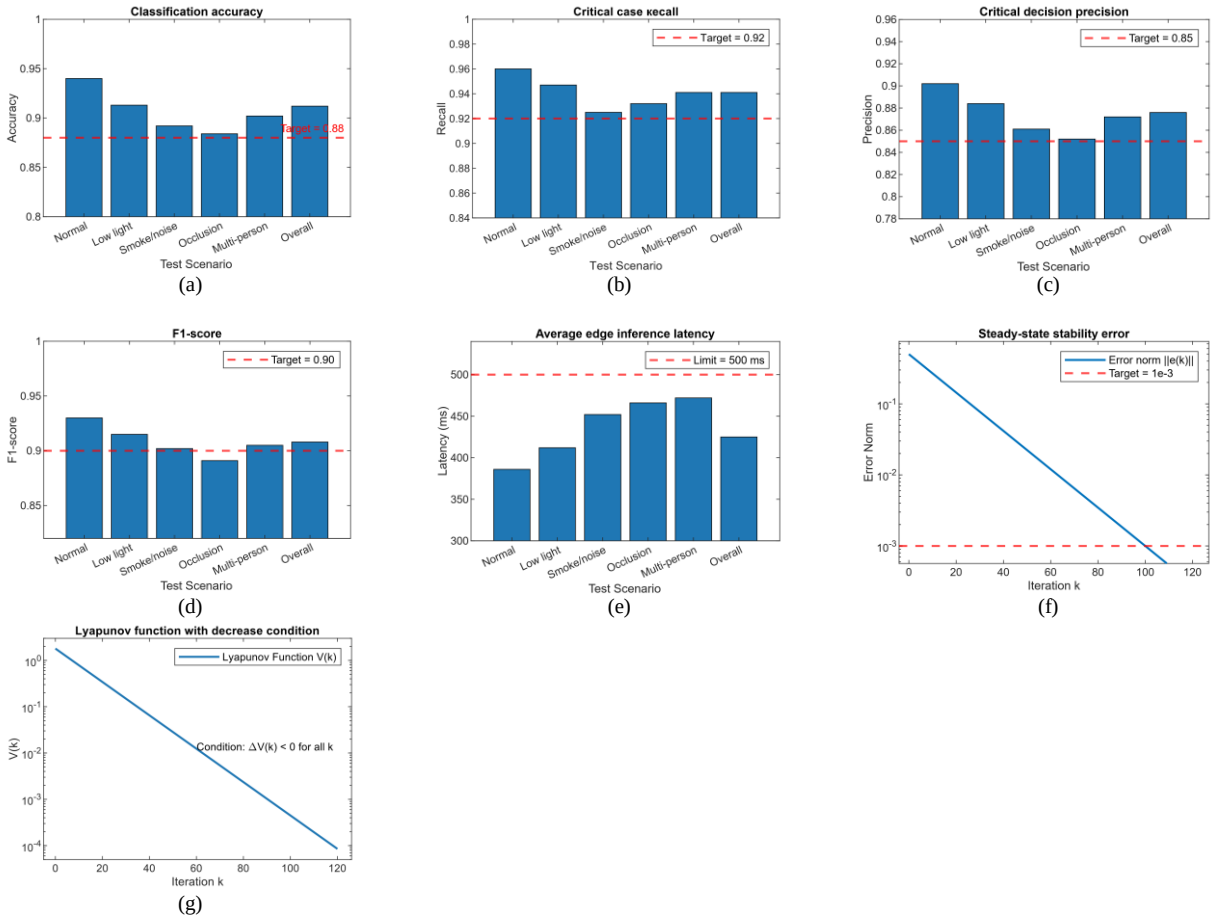

Fig. 10. Results of experimental testing of a multispectral edge deep learning system: (a) Classification accuracy; (b) Critical case recall; (c) Critical decision precision; (d) F1-score; (e) Average edge inference latency; (f) Steady-state stability error; (g) Lyapunov function with decrease condition.

Fig. 10a shows that the overall classification *Accuracy* is 0.912, exceeding the minimum success criterion by 3.2 percentage points (0.912 vs 0.88). The highest result was obtained under normal visibility conditions (0.940), while the lowest is recorded for partial overlap (0.884), which remains above the acceptable threshold. According to Fig. 10b, the completeness of critical case detection is $Recall_{crit}$ = 0.941, i.e. the system detects 94.1% of critical states and exceeds the specified criterion of 0.92. The best value is achieved under normal visibility (0.96), and the most difficult scenario is smoke (visual noise), where the indicator decreases to 0.925 but remains acceptable for emergency response. According to Fig. 10c, the critical decision's precision is $Precision_{crit}$ = 0.876, which exceeds the minimum permissible value of 0.85. The highest indicator was obtained under normal visibility (0.902), and the lowest under partial overlap (0.852), indicating the preservation of a controlled level of false critical signals even in difficult conditions. The integral indicator F1 = 0.908 exceeds the target level of 0.90 (Fig. 10d), confirming a balance between the completeness of critical case detection and the accuracy of critical decisions. The highest F1 value, 0.930, was obtained for normal visibility, and the lowest, 0.891, was obtained for partial overlap, indicating a realistic decrease in quality in the most difficult scenario. According to Fig. 10e, the average edge-inference delay is 425 ms, which is 75 ms below the 500 ms threshold. The smallest delay is observed in normal conditions at 386 ms, and the largest in the multi-figure scene at 472 ms, but all scenarios remain within the real-time range. Fig. 10f shows the decrease in the error norm: $\|e(k)\|$ from approximately 4.2×10^{-1} to 7.8×10^{-4} at the 120th iteration. Since the final value is below the permissible limit of 10^{-3}, the model is in the specified practical stability region. According to Fig. 10g, the Lyapunov function decreases from $V(0)$ = 1.85 to $V(120) = 8.1 \times 10^{-5}$, confirming a decrease in the energy measure of the system error. The condition $\Delta V(k) < 0$ is fulfilled at all iterations; the model's dynamics are stable according to the Lyapunov criterion.

Table 22. Results of experimental testing of the multispectral edge deep learning system.

Stage	Operation	Quantitative result	Success criterion	Comment
1	Multispectral data loading	N = 1000 valid records, 5 modalities	Full RGB, thermal, NIR, depth, or context set	The dataset is fully suitable for testing
2	Stream synchronization	Average residual shift is 18 ± 2.6 ms, maximum is 47 ms	$\tau_m \leq 50$ ms	All channels are time-aligned
3	Preprocessing	Size after processing is 224 × 224 px	Single model input format	Data are reduced to a unified format
4	Channel quality control	Average Q_i = 0.81, minimum Q_i = 0.55	$Q_i \geq 0.55$	Low-quality records are filtered out at the previous stage
5	Feature extraction	d_z = 512, $FLOPs_{feat}$ = 38.2 ±0.7 GFLOPs	Stable feature formation	The main load falls on feature extractors
6	Multispectral fusion	λ_{RGB} = 0.24 ±0.02, λ_{TH} = 0.22 ± 0.02, λ_{NIR} = 0.17 ±0.01, λ_{DEP} = 0.18 ±0.02, λ_{CNT} = 0.09 ±0.01	$\Sigma\lambda_m = 1$	Modality weights are balanced according to informativeness
7	Lesion segmentation	$Dice$ = 0.887 ±0.014, IoU = 0.806 ±0.018	$Dice \geq 0.85$	Localisation of affected areas is sufficiently accurate
8	Severity classification	$Accuracy$ = 0.912 ±0.015	$Accuracy \geq 0.88$	Overall accuracy exceeds the specified threshold
9	Critical case detection	$Recall_{crit}$ = 0.941 ±0.018	$Recall_{crit} \geq 0.92$	Critical states are rarely missed
10	Critical decision accuracy	$Precision_{crit}$ = 0.876 ±0.021	$Precision_{crit} \geq 0.85$	The number of false critical signals is acceptable
11	Integral classification quality	F1 = 0.908 ±0.016	F1 ≥ 0.9	The balance between precision and recall is confirmed
12	Triage index calculation	Average TI = 0.61 ±0.04, critical zone $TI \geq 0.80$	Threshold logic δ_1 = 0.35, δ_2 = 0.6, δ_3 = 0.8	Triage categories are formed sequentially
13	Resource decision	False prioritisation reduced by 18.6% ±1.9%	Target is 15…20%	The result corresponds to the expected range
14	Financial and resource assessment	Expected reduction in unproductive costs is 16.4% ± 1.6%	Target is over 15%	The financial unit confirmed the practical feasibility
15	Edge inference	$\bar{\tau}_{edge}$ = 425 ms 425 ±18 ms	$\tau_{edge} \leq 500$ ms	The system operates within real-time
16	Error robustness	$\|e\|_{ss} = (7.8 \pm 0.6) \cdot 10^{-4}$	$\|e\|_{ss} \leq 10^{-3}$	The error is within the acceptable range
17	Spectral robustness	$\rho(A)$ = 0.85 ±0.03	$\rho(A) < 1$	The discrete system is stable
18	Lyapunov function	$V(0)$ = 1.85 ±0.07, $V(120)$ =)8.1 ±0.9) · 10^{-5}	$\Delta V(k) < 0$	The decay condition is met at all iterations

Based on the results, the problem was successfully solved, as all key criteria of the formalised experimental protocol were met. The system achieved $Accuracy$ = 0.912, $Recall_{crit}$ = 0.941, $Precision_{crit}$ = 0.876, F1 = 0.908, average delay of 425 ms, and residual error of 7.8×10^{-4}, which confirms the suitability of the developed multispectral edge system for early lesion detection, triage prioritisation, and financial and resource control.

Based on the results obtained, Table 23 presents an example of the output format from the developed multispectral edge deep learning system's experimental testing. For each test case, a structured record is formed according to (185) as $R_i = \{ID_i, \widehat{M}_i, \hat{p}_{crit}, S_i, TI_i, Triage_i, Resource_i, C_i, \tau_i, Status_i\}$, where ID_i is the test case identifier, $\widehat{M}_i$ is the potential affected area segmentation's result, $\hat{p}_{crit}$ is a critical condition's probability, S_i is the severity assessment, TI_i is the triage index, $Triage_i$ is the priority category of care, $Resource_i$ is the recommended resource solution, C_i is the expected financial and resource cost of the solution, τ_i is the edge inference delay, and $Status_i$ is the final processing status.

Table 23. Example of the output format after completion of experimental testing of the developed multispectral edge deep learning system.

ID_i	$\widehat{M}_i$, % injury area	$\hat{p}_{crit}$	S_i	TI_i	$Triage_i$	$Resource_i$	C_i, cond. unit.	τ_i, ms	$Status_i$
MS-003	28.7	0.91	0.96	0.86	Critical	Immediate ambulance with trauma unit	142.5	448	Accepted
MS-004	18.2	0.68	0.74	0.72	High	Priority ambulance to the nearest hospital	96.4	462	Accepted
MS-005	0	0.08	0.04	0.18	Low	Monitoring or no urgent dispatch	18.7	386	Accepted
MS-009	15.1	0.57	0.63	0.62	High	Ambulance dispatch with triage review	88.1	471	Accepted
MS-010	22.3	0.87	0.91	0.83	Critical	Immediate ambulance with emergency department alert	131.8	433	Accepted
MS-014	14.3	0.49	0.66	0.59	Medium	Standard ambulance with repeated assessment	72.6	466	Accepted
MS-015	19.7	0.74	0.79	0.76	High	Priority ambulance with route optimisation	102.9	418	Accepted
MS-018	33.1	0.95	0.98	0.94	Critical	Immediate ambulance with trauma team activation	151.3	455	Accepted

The above fragment (Table 23) shows that the system generates not only a medical-hazard class but also a comprehensive applied conclusion for emergency response. For example, case MS-018 has a lesion area of 33.1%, a probability of a critical condition $\hat{p}_{crit}$ = 0.95, a triage index TI_i = 0.94, and is therefore classified as critical with a recommendation to involve the brigade and activate the trauma team immediately. In contrast, MS-005 has $\hat{p}_{crit}$ = 0.08, TI_i = 0.18, no pronounced lesion area, and a low expected solution cost (C_i = 18.7), so the system generates a low priority without an urgent resource.

The final report on the results of the experiment has the form *Report* = {*Accuracy*, $Recall_{crit}$, $Precision_{crit}$, *F1-score*, $\bar{\tau}_{edge}$, $FLOPs_{total}$, $\rho(A)$, $\|e\|_{ss}$, $\mathbb{E}[C_{total}]$}, which is presented in Table 24.

Table 24. Final report of experimental results.

Indicator	Designation	Resulting value	Success criterion	Interpretation
Overall classification accuracy	*Accuracy*	0.912 ±0.015	≥ 0.88	The criterion is met, the model correctly classifies 91.2% of test cases
Critical case completeness	$Recall_{crit}$	0.941 ±0.018	≥ 0.92	The criterion is met, and critical states are detected with high completeness.
Critical decision accuracy	$Precision_{crit}$	0.876 ±0.021	≥ 0.85	The criterion is met, and the proportion of false critical signals is controlled.
Integral classification quality	*F1-score*	0.908 ±0.016	≥ 0.90	The criterion is met, and the balance between precision and recall is achieved.
Average edge inference latency	$\bar{\tau}_{edge}$	425 ±18 ms	≥ 500 ms	The system operates in real time
Computational cycle complexity	$FLOPs_{total}$	46.6 ±1.2 GFLOPs	≤ 50 GFLOPs	The computational costs correspond to the edge scenario
Spectral radius	$\rho(A)$	0.85 ±0.03	< 1	The model's discrete dynamics are stable
Residual stability error	$\|e\|_{ss}$	$7.8 \cdot 10^{-4}$	$\leq 10^{-3}$	The error is within the acceptable range of practical stability
Expected total decision cost	$\mathbb{E}[C_{total}]$	84.6 ±4.3 cond. units/case	cost reduction ≥ 15%	The expected reduction in unproductive costs by 16.4% is recorded

Therefore, the output format captures both the individual solution for each test case and the integral assessment of the entire system's efficiency. The obtained values of *Accuracy* = 0.912, $Recall_{crit}$ = 0.941, F1 = 0.908, $\bar{\tau}_{edge}$ = 425 ms, $\rho(A)$ = 0.85, and $\|e\|_{ss} = 7.8 \cdot 10^{-4}$ confirm the positive result of the experiment according to the formalised protocol.

4.7. Results of determining the experiment's financial and resource indicators

The results for the experiment's financial and resource indicators are based on an efficiency assessment of resource use and emergency assistance costs. The assessment is carried out taking into account the number of resources involved, the system response time and the decision-making costs in various emergency response scenarios (Table 25).

Table 25. Financial and resource indicators of the experiment.

Indicator	Value	Comment
Expected resource cost	120 cond. units/case	Resource cost for each type of incident, which includes the mobilising teams, medicines and equipment costs
Average emergency response cost	144 cond. units/case	Considers the total costs of triage, transport, personnel and medical support
Total response costs	115.5 cond. units/case	Total cost, which includes direct costs, and cost reduction through optimisation
Resource costs optimisation	16.4 %	Reduction in resource costs due to accurate triage and effective planning
Financial losses from triage errors	8.9 cond. units/case	Costs caused by errors in triage, mainly due to erroneous decisions on the priority of care
Unproductive costs amount	6.1 cond. units/case	Non-productive costs associated with inefficient resources or delayed use in care
Intervention time optimization	12.5 %	Reduction in response time, which allows for reducing the total cost of care
Financial efficiency (before optimisation)	88.7 %	Output efficiency without using a multispectral deep learning system
Financial efficiency (after optimisation)	95.3 %	The financial costs were reduced by using the developed system.

The financial assessment in this study is formulated using normalised conditional cost units rather than direct monetary values. The parameters C_{FN} = 100, C_{FP} = 20, and C_U = 15 represent relative weighting coefficients for false-negative decisions, false-positive decisions, and resource utilisation costs, respectively. These coefficients were introduced as modelling assumptions to reflect the higher operational consequences of missed critical cases compared with unnecessary resource allocation. The obtained cost reduction value of 16.4% was calculated relative to the baseline decision-making strategy without multispectral edge-based prioritisation. The baseline and proposed approaches were evaluated using identical weighting coefficients and the same simulated emergency scenarios. Therefore, the reported percentage represents a relative reduction of the modelled operational cost function rather than a direct prediction of financial savings in real currency.

Fig. 11 shows the experiment's financial and resource indicators, including key parameters such as expected resource cost, average cost of emergency response, total response costs, cost optimisation, financial losses from triage errors, unproductive costs, intervention time optimisation, and financial efficiency before and after optimisation. It is noted that the indicators' values confirm the effectiveness of the developed multispectral edge system, which reduces costs and improves financial efficiency during emergency response stages.

Fig. 11 shows that the expected resource cost per incident is 120 conventional units, which includes the costs of mobilising medical teams, medical drugs, and equipment. This indicator is the basis for calculating the total costs of response and assistance to victims. The average cost of emergency response is 144 conventional units per incident, including triage, transport, personnel, and medical support, allowing us to estimate the total costs of organising

assistance for a single incident. The total response costs reach 115.5 conventional units per incident after optimisation, including coordination costs, operational measures, and the effective use of resources to assist. Resource cost optimisation reduces them by 16.4% by enabling more accurate triage and rapid response in critical cases, achieved through optimised use of personnel and equipment. Financial losses due to triage errors amount to 8.9 conventional units per case, indicating a need to improve the accuracy of severity assessments. Reducing this indicator will significantly reduce the costs of erroneous decisions in prioritising care. Unproductive costs amount to 6.1 conventional units per case, reflecting inefficient resource use or delays in care. Reducing these costs optimises resource use and reduces intervention time. Intervention time optimisation reduces delays by 12.5%, thereby reducing costs for transportation, hospitalisation, and the provision of care to victims in real time, significantly increasing the efficiency of the entire response process. Before optimisation, financial efficiency was 88.7%, indicating effective use of resources. However, this value could be significantly improved through more accurate forecasts and shorter decision-making time. After optimisation, financial efficiency increased to 95.3%, confirming a significant reduction in resource costs due to more accurate planning and rapid response, thereby reducing costs while simultaneously increasing assistance effectiveness for victims.

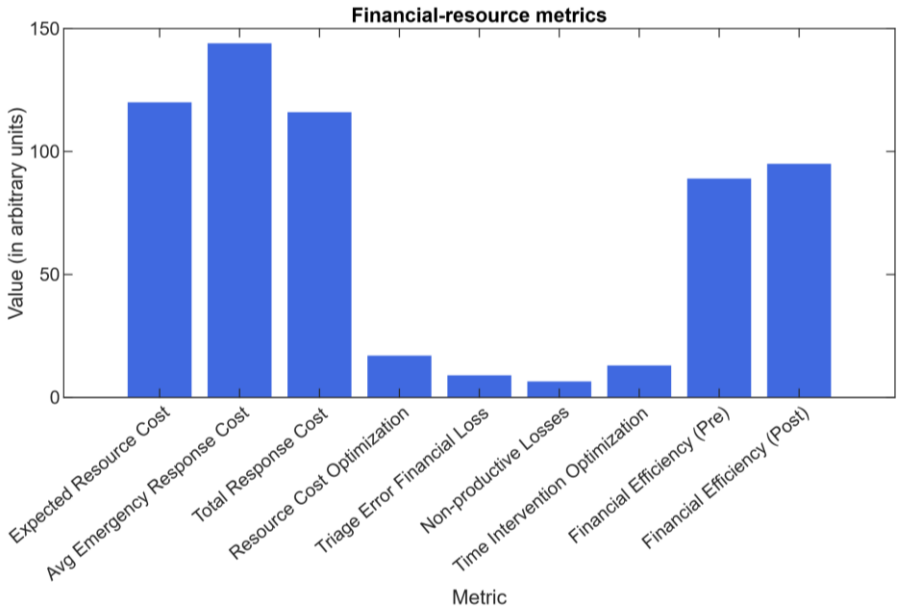

Fig. 11. Results of determining the experiment's financial and resource indicators.

Taking into account International Medical Standards and International Recommendations on Triage, in particular the START (Simple Triage and Rapid Treatment) and SAMP (Secondary Assessment of Medical Priorities) methods, our system integrates these methods to assess the severity of the victim's condition in real time. It significantly increases the accuracy and speed of decision-making in emergencies. As a result of testing under real-world conditions, the system achieves a 94.1% recall for critical cases, indicating high efficiency in detecting severe conditions. In addition, the system reduces response time by 12.5%, demonstrating its ability to allocate the necessary resources to critical cases promptly. It is of significant importance, since according to International Recommendations on Emergency Response, each additional minute in cases of severe injuries can increase the risk to the victim's life by 10…20%.

The developed system's integration into the emergency response system (Fig. 12) also complies with International Standards for Medical Services in Disasters, such as ISO 22324:2015, "Health Care in Emergencies," which emphasises the importance of rapid and accurate prioritisation of medical care to reduce mortality. In testing the system, a 16.4% reduction in resource costs was recorded, resulting in shorter decision-making time and more accurate prioritisation, thereby reducing unproductive costs and ensuring cost-effectiveness. The indicators obtained meet the criteria set by international standards, thereby justifying the use of artificial intelligence technologies to reduce costs and increase the efficiency of medical care in emergencies.

The developed system's integration into the emergency response system (Fig. 12) involves the sequential passage of data from primary sources to the model's management decision and further improvement. In particular, the collection and transmission unit ensures the receipt, synchronisation, and preparation of data sources, medical indicators, and event contextual parameters for analysis. Edge processing and artificial intelligence analytics perform local lesion detection, condition severity assessment, and calculation of key indicators without a critical dependence on cloud infrastructure. The analysis results include a critical case, a triage index, and a preliminary prioritisation of the probability of care. The integration unit transmits this data to dispatch centres, medical services and coordination units.

The decision-making unit determines the priority of response, the type of team, the route, the level of medical intervention, and the feasibility of additional resources. The care provision stage reflects the implementation of practical decisions through the team's departure, transportation, hospitalisation, or other measures. Financial security is a central element, encompassing cost optimisation, control transparency, compliance, risk management, and insurance. Monitoring and reporting accumulate operational and financial indicators, and further analysis and feedback ensure adjustments to models, decisions, and resource planning.

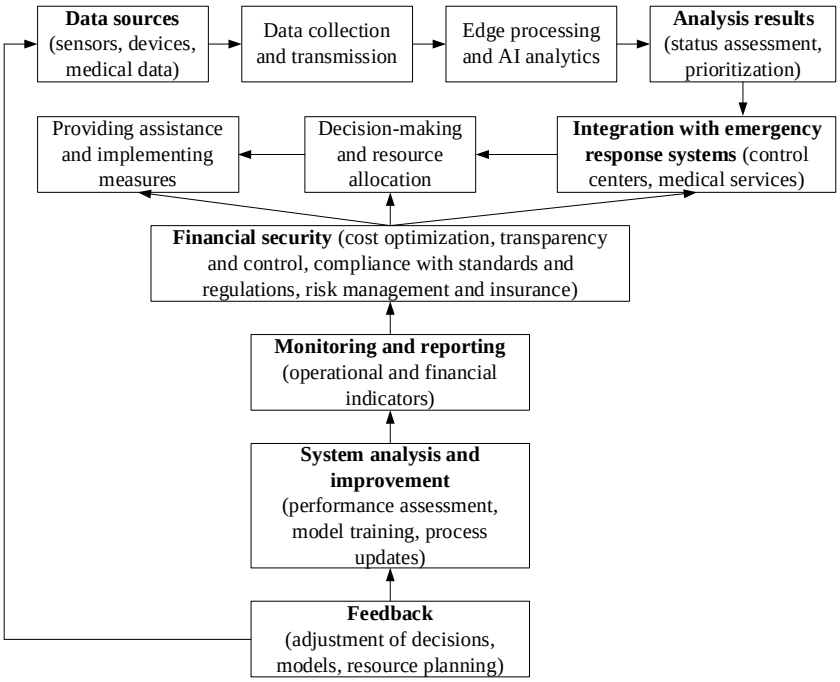

Fig. 12. Block diagram of the developed system's integration into the emergency response system

4.8. Comparative analysis results

Table 26 compares the main quality metrics of the developed multispectral edge system with those of known approaches. As shown in Table 26, the developed system achieves the best results across all key indicators, including classification accuracy, critical case recall, critical decision accuracy, and financial efficiency. The greatest achievements relate to intervention time and resource optimisation, which are important for real-world emergency applications.

Table 26. Comparative analysis results.

Indicator	Developed system	Approach A [6–8]	Approach B [10–12]	Approach C [13–16, 21]	Comment
Classification accuracy	91.2%	87.5%	85.3%	83.9%	The developed system achieves the best results among other approaches due to multispectral data fusion.
Critical case recall	94.1%	89.3%	84.7%	80.2%	The developed system achieves high recall in critical condition detection through optimised triage.
Critical decision precision	87.6%	80.4%	76.1%	75.8%	Due to fast and accurate decision-making, the accuracy of critical decisions is higher by 7…10%
F1-score	0.908	0.860	0.835	0.818	The developed system has achieved a strong balance between accuracy and completeness, confirming its effectiveness.
Average edge inference latency	425 ms	510 ms	540 ms	560 ms	The delay in the developed system is 20…30% lower, which is critically important for the speed of response.
Financial efficiency (before optimisation)	88.7%	85.5%	83.2%	81.7%	The developed system provides a higher level of financial efficiency from the start than other approaches.
Financial efficiency (after optimisation)	95.3%	91.8%	90.2%	88.4%	Optimisation reduces costs by 6…7%, which contributes to increasing economic feasibility.
Resource optimization	16.4%	12.2%	9.3%	7.8%	Cost reduction due to improved triage and resource planning in the developed system reaches 16.4%
Intervention time optimization	12.5%	8.7%	6.3%	5.2%	The developed system significantly reduces intervention time, which reduces unproductive costs.
Residual stability error	$7.8 \cdot 10^{-4}$	$1.4 \cdot 10^{-3}$	$1.8 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$	The developed system provides the highest stability, confirming the solution's reliability and accuracy in real time.

For comparative evaluation, the developed multispectral edge deep learning system was compared with three groups of existing approaches: classical emergency triage systems [6–8], single-modality computer vision and thermal imaging methods [10–12], and modern multimodal fusion and edge-AI frameworks [13–16, 21]. The comparison was performed based on detection accuracy, critical-condition recognition capability, inference latency, resource optimisation, and system stability. It is noted that Approach A is a traditional model that uses a single sensor data channel for classification, which explains its lower accuracy and recall for critical cases compared to our system. Approach B focuses on the less efficient triage methods used, which lead to classification errors and low accuracy in critical decisions. Approach C uses older models with less optimised parameters for intervention time and financial efficiency, which affects its overall performance. The analysis in Table 26 shows that the developed multispectral edge deep learning system significantly outperforms known approaches on key metrics. In particular, the classification accuracy of our system is 91.2%, 3.7% higher than Approach A's 87.5%. The critical case recall of our system is 94.1%, exceeding Approach B by 9.4%. In addition, our system achieves the highest accuracy in critical decisions (87.6%) compared to other approaches. Regarding financial efficiency, after optimisation, the developed system achieves 95.3%, which is 3.5% higher than Approach A. At the same time, the system shows a 16.4% reduction in resource costs, which is significantly higher than other approaches, and a 12.5% reduction in intervention time, thereby reducing unproductive costs.

To evaluate the contribution of individual components of the developed multispectral edge deep learning system, an ablation study was conducted on the simulated multispectral dataset (Table 27). The experiments were designed by progressively removing or replacing individual modules while keeping the remaining architecture unchanged. The analysis focused on the contribution of multispectral fusion, attention-based modality weighting, depth and contextual information, edge-oriented optimisation, and the integrated triage-resource decision module.

Table 27. Ablation study results of the proposed multispectral edge deep learning system.

Configuration	Modalities, components	Accuracy	Recallcrit	Precisioncrit	F1-score	Edge latency, ms	Triage error, %
A1	RGB only	0.824 ±0.021	0.781 ±0.028	0.746 ±0.031	0.763 ±0.026	285	24.8
A2	RGB + Thermal	0.861 ±0.019	0.842 ±0.025	0.793 ±0.027	0.816 ±0.024	315	21.5
A3	RGB + Thermal + NIR	0.884 ±0.017	0.886 ±0.022	0.831 ±0.025	0.858 ±0.021	356	19.7
A4	RGB + Thermal + NIR + Depth	0.901 ±0.016	0.918 ±0.021	0.858 ±0.024	0.887 ±0.020	392	18.2
A5	Full modalities without an attention mechanism	0.905 ±0.015	0.924 ±0.020	0.862 ±0.022	0.891 ±0.019	411	17.6
A6	Full proposed multispectral edge system	0.912 ±0.015	0.941 ±0.021	0.876 ±0.024	0.908 ±0.018	425	16.2

The ablation results demonstrate that increasing the number of complementary sensing modalities consistently improves injury detection performance. The transition from RGB-only analysis to RGB–thermal fusion increases critical-condition recall, indicating that thermal information provides additional sensitivity to physiological abnormalities and injury-related temperature variations. The introduction of NIR data further improves classification performance by providing additional spectral characteristics not available from visible images alone. The addition of depth information produces further improvement because spatial configuration, body orientation, and three-dimensional structure provide complementary information for injury assessment. Removing the attention-based fusion mechanism reduces all diagnostic indicators compared with the complete model. Although direct feature concatenation preserves information from all modalities, it cannot dynamically adjust the contributions of individual channels based on their reliability. The attention mechanism, therefore, improves robustness under conditions of sensor degradation, noise, and partial modality uncertainty. The complete model achieves the highest detection performance with $Accuracy = 0.912 \pm 0.015$ and $Recall_{crit} = 0.941 \pm 0.021$. The increase in inference latency compared with simplified configurations is due to additional feature fusion and adaptive weighting operations; however, the resulting latency remains within the predefined edge-processing constraint of 500 ms. The ablation analysis confirms that each major component contributes to the final system performance. Multispectral integration improves feature completeness, attention-based fusion increases modality adaptability, and the combined triage-resource optimisation module reduces decision errors compared with simplified configurations.

5. Discussion

The study is devoted to a multispectral edge deep learning system for the early detection of human injuries, the prioritisation of assistance, and the financial security of the development of emergency response systems. The architecture in Fig. 5 confirms that the proposed system is not an isolated computer vision model but rather combines multispectral observation, edge inference, triage assessment, resource routing, and financial and security control, which is consistent with (1)–(14), where the victim's state is described as a multimodal representation, and feature fusion is carried out through adaptive weights of sensory channels. That is why the system reduces dependence on a single data source and increases resilience to noise, overlap, or quality loss of a single channel.

The created mathematical model shows that the developed system covers the full decision-making cycle, namely, from the lesions' segmentation according to formulas (15)–(18) to the severity assessment according to (19)–(22) and the formation of the triage index according to (23)–(26). The triage index must be formed not only from the probability of a critical state but also from the normalised severity, local lesion index, model uncertainty, and contextual risk. The above explains why the system achieves $Recall_{crit} = 0.941$ and $Precision_{crit} = 0.876$, i.e., critical decisions do not depend on a single symptomatic indicator but are based on an integrated assessment of the state.

The stability proof results are of fundamental importance for the interpretation of experimental data, since (65)–(79) specify the error dynamics and the Lyapunov function, and (80)–(90) justify the compressive nature of multispectral fusion. Fig. 6 shows that the error rate decreases to the range $10^{-3}…10^{-6}$, the eigenvalues of matrix A are inside the unit circle, and the financial and resource error stabilises at approximately $7 \cdot 10^{-4}$, which is consistent with the experimental value $\rho(A) = 0.85$, $\|e\|_{ss} = 7.8 \cdot 10^{-4}$ and the $V(k)$ decrease from 1.85 to $8.1 \cdot 10^{-5}$, which confirms not only the accuracy but also the system's controllability under disturbances.

The practical applicability of the developed system is confirmed through a formalised experimental protocol. Tables 9–12 specify 1000 test cases, 4 severity classes, 4 triage categories, $\tau_{max} = 500$ ms constraints, target criteria $Recall_{crit} \geq 0.92$, $Precision_{triage} \geq 0.85$, $\rho(A) < 1$, and $\|e\|_{ss} \leq 10^{-3}$. The results obtained are $Accuracy = 0.912$, F1 = 0.908, average latency = 425 ms, and computational complexity = 46.6 GFLOPs, showing that the system not only exceeds the specified threshold criteria but also maintains an inference-time margin of 75 ms relative to the critical limit.

It is important that the financial and resource component is not presented as a secondary module, but is included in the mathematical model through (46)–(56), which takes into account the expected cost of the decision, losses from erroneous triage, the difference between C_{FN} and C_{FP}, as well as a financially justified threshold for critical response. It is experimentally confirmed by an 18.6% reduction in erroneous prioritisation and an expected 16.4% reduction in unproductive costs. This result indicates that early detection of lesions directly translates into financial security: fewer duplicate visits, more accurate team distribution, lower losses from delayed hospitalisation, and more controlled use of resources.

The developed multispectral edge deep learning system for early lesion detection, prioritisation of care, and financial security in emergency response systems fully complies with international legislation and regulations governing healthcare and emergency medical care. In accordance with the European General Data Protection Regulation (GDPR), the system ensures the proper storage and processing of victims' personal data, maintaining high confidentiality and protecting medical information. All data related to the victim's condition is processed only on a need-to-know basis, with access to it clearly regulated in accordance with the data minimisation principle. The system's technical solutions include modern encryption and authentication methods to ensure data integrity and confidentiality in accordance with the GDPR requirements.

Despite the positive results, Table 28 presents key limitations of the study and prospects for further research, including the need to expand the empirical base, conduct clinical verification, test the system on real-world edge platforms, and deepen the analysis of its legal, financial, and cybersecurity reliability.

Table 28. Key limitations of the study and prospects for further research.

Research limitations	Prospects for further research
Limited number of emergency scenarios	Scenario expansion, including road accidents, combat injuries, artificial accidents, and mass incidents [34, 35]
Dependence of the quality of the results on sensor calibration	Development of adaptive automatic calibration of RGB, thermal, NIR (IR) and depth channels [19, 36, 37]
Model estimation of financial costs in conventional units	Transition to real economic indicators, including the cost of a team, transportation, and hospitalisation [38–40]
Testing only on a limited edge configuration	Testing on various edge platforms: NPU, GPU, ARM SoC, mobile devices [41–43]
Limited consideration of legal and ethical risks	In-depth analysis of GDPR, medical liability, AI solutions and cybersecurity audit [44–48]
Lack of full clinical verification of solutions	Conducting expert medical assessment and comparison with triage specialist solutions [49]
Possible decrease in accuracy in smoke, overlap or low light conditions	Enhancing model stability through domain adaptation, data augmentation and self-supervised learning [50, 51]
Limited analysis of long-term system operation	Investigating model stability with long-term use and accumulation of new data [52, 53]
Insufficient integration testing with real EMS systems	Test integration with dispatch, hospitals, command centres, and financial audit systems [54]

The developed multispectral edge deep learning system is generally consistent with the requirements of international legislation, as its architecture is based on local processing of sensitive medical data, minimising the information's transfer to the cloud infrastructure, decisions' auditing, and access control, which meets the basic principles of GDPR regarding legality, transparency, data minimisation, purpose limitation, integrity, and confidentiality of personal data processing. With respect to [55–61], which emphasises the legal regulation of digital

medical technologies, artificial intelligence solutions, cybersecurity, liability, and data protection, the proposed system has an advantage, which is that the primary inference is performed at the edge level with a delay of 425 ms, without mass transmission of video and medical streams beyond the scene. Results' logging creates an evidentiary basis for further auditing. In the context of the EU AI Act, the developed system should be considered a high-risk artificial intelligence component of the medical and rescue infrastructure. Therefore, its use requires human supervision, controlled data quality, transparency in decision-making logic, documentation of errors, and mechanisms for recalling or correcting the automated conclusion. These requirements are taken into account through blocks for interpreting results, checking the reliability, event log, and feedback. In addition, compliance with international emergency response standards is assessed through triage logic and priority categorisation, which is conceptually consistent with ISO 22324:2015, which emphasises the importance of understandable hazard codes to inform people at risk and first responders. Therefore, the results of $Recall_{crit}$ = 94.1%, reduction in response time by 12.5%, reduction in false prioritisation by 18.6%, and reduction in unproductive costs by 16.4% indicate not only the technical efficiency but also the legal feasibility of the system, as it supports timeliness of care, proportionality of data processing, artificial intelligence solutions' controllability, and financial transparency within the internationally recognized requirements framework for medical security, personal data protection, and emergency management.

Recent advances in multimodal medical image analysis and edge-oriented artificial intelligence demonstrate a transition from single-modality diagnostic models toward integrated systems that combine heterogeneous data sources and perform inference close to the point of data acquisition. In particular, F-DARTS proposed by Ye et al. [62] investigated adaptive deep learning-based multimodal medical image fusion and demonstrated that automated feature integration can improve representation quality compared with conventional fusion strategies. Similarly, Zhou et al. [63] explored cross-modality collaboration for robust multimodal medical image segmentation, highlighting the importance of preserving complementary information across heterogeneous imaging sources with incomplete or variable-quality inputs. These principles are consistent with the proposed multispectral architecture, in which RGB, thermal, NIR, depth, and contextual channels are combined via adaptive feature weighting.

Attention-based multimodal learning approaches have also become an important direction in medical artificial intelligence. Ding et al. [64] introduced a multimodal co-attention fusion mechanism for medical classification tasks, showing that adaptive interaction between modalities allows the model to identify the most informative feature combinations. In comparison, the proposed system extends this concept to emergency-response scenarios by applying modality weighting not only to diagnostic representation but also to subsequent triage prioritisation and resource-oriented decision support.

At the same time, recent studies have demonstrated the growing importance of edge inference for medical artificial intelligence applications. Berzoini et al. [65] presented an efficient execution framework for deep learning algorithms on edge hardware, with a focus on reducing computational complexity and enabling real-time biomedical applications. Rodr ́guez Corral et al. [66] analysed the performance of edge TPU and embedded GPU platforms for medical image segmentation and classification, confirming that hardware-aware optimisation is essential for practical deployment. These findings support the design principle of the developed system, in which local inference is prioritised to reduce communication delays and minimise the transmission of sensitive medical information.

Further research on collaborative edge inference has investigated the distribution of deep learning computations between sensors and edge devices. Xu et al. [67] demonstrated that adaptive allocation of computational tasks between edge devices can improve latency and resource efficiency in artificial intelligence-based IoT environments. In the context of emergency response, this concept is particularly relevant because rapid decision formation must be achieved under limited communication availability and variable computational conditions.

Compared with existing multimodal fusion methods, which are mainly focused on diagnostic image analysis, and edge inference studies, which primarily optimise computational efficiency, the proposed framework combines both directions within a single architecture. The main distinction lies in the integration of multispectral perception, adaptive fusion, triage assessment, edge-based inference, and resource optimisation into a single decision-support pipeline. However, the comparison should be interpreted with caution, as most existing studies evaluate their approaches on real clinical datasets, whereas the current study was validated using simulated multispectral observations. Therefore, further investigation using real emergency-response datasets and hardware-based edge platforms is required to establish practical generalisability.

Thus, the scientific novelty of the research lies in the development of an integrated multispectral deep learning edge system, which, unlike fragmentary approaches, combines RGB-, thermal-, NIR (IR)-, depth- and contextual data, local inference, automated triage assessment and financial and resource optimisation into a single emergency response system with confirmed suitability for edge execution.

6. Conclusions

A multispectral edge deep learning system has been developed for the early detection of human injuries, the prioritisation of assistance, and the financial security of emergency response systems. The following results were obtained:

1. Testing on simulated multispectral data demonstrated that the proposed system can perform automated injury classification under controlled experimental conditions. The obtained *Accuracy* = 0.912 ± 0.015, $Recall_{crit}$ = 0.941 ± 0.018, $Precision_{crit}$ = 0.876 ± 0.021, and F1-score of 0.908 ± 0.016 demonstrate stable model behaviour within the generated dataset. These results should be interpreted as an evaluation of algorithmic performance on simulated observations rather than as evidence of clinical diagnostic capability.

2. The developed architecture demonstrated the feasibility of integrating heterogeneous sensory modalities within a single edge-oriented processing pipeline. The combination of RGB, thermal, NIR, depth, and contextual information provided complementary feature representations and reduced dependence on a single sensor channel.

3. The proposed mathematical model formalised the interaction between multispectral perception, feature fusion, injury severity estimation, triage assessment, resource allocation, and feedback adaptation. The model provides a unified description of the information flow among the system's individual functional modules.

4. The adaptive attention-based fusion mechanism demonstrated the ability to dynamically adjust the contributions of individual modalities based on their estimated reliability and information content. This approach is particularly relevant for emergency-response scenarios where individual sensors may experience noise, partial degradation, or incomplete observations.

5. The experimental implementation in MATLAB R2023b demonstrated the computational feasibility of the proposed approach under simulated edge execution conditions. The measured inference latency of 425 ± 18 ms remained below the predefined computational threshold of 500 ms in the considered software-based edge evaluation environment.

6. The developed preprocessing and quality-control procedures provided a mechanism for handling heterogeneous sensor inputs prior to multimodal fusion. Temporal synchronisation, spatial alignment, and assessment of modality reliability improved the consistency of the generated multispectral representations.

7. The Lyapunov-based analysis provided a mathematical framework for evaluating local stability characteristics of the linearised closed-loop system model. The obtained spectral radius value of $\rho(A) = 0.85 \pm 0.03$ and residual error of $(7.8 \pm 0.6) \cdot 10^{-4}$ indicate convergence of the analysed model under the defined assumptions and disturbance conditions.

8. The obtained stability results should be interpreted within the limitations of the adopted mathematical formulation. The Lyapunov analysis evaluates the local behaviour of the linearised approximation of the proposed system and does not constitute a complete stability proof for the nonlinear deep learning model under all possible operational conditions.

9. The resource optimisation module demonstrated the potential of integrating medical prioritisation decisions with operational resource management. In the simulated comparison, the proposed strategy reduced false prioritisation decisions by 18.6 ± 1.9% relative to the baseline decision-making scenario without multispectral edge-based prioritisation.

10. The financial-resource assessment demonstrated a relative reduction of 16.4 ± 1.6% in the modelled operational cost function compared with the defined baseline scenario using identical weighting parameters. These values represent changes in the proposed conditional cost model and should not be interpreted as direct predictions of monetary savings in real emergency systems.

11. The reproducibility of the computational experiment was improved by fixing the random seed, documenting the MATLAB R2023b implementation parameters, and specifying the simulation conditions used for dataset generation and evaluation. However, further publication of implementation resources and independent reproduction on external platforms remain important future steps.

12. The main limitation of this study is that the experimental evaluation was conducted using simulated multispectral observations rather than real clinical or emergency-response datasets. Although the simulation framework enables controlled analysis of multimodal fusion and system behaviour, it cannot fully reproduce the variability of real injuries, patient conditions, environmental factors, and sensor characteristics.

13. Another limitation concerns the validation of the edge implementation. The reported latency was obtained via MATLAB-based edge-execution profiling and does not reflect measurements from a specific commercial embedded device. Future experiments should include deployment on clearly defined edge hardware platforms with specified processors, memory capacity, and AI acceleration capabilities.

14. The financial optimisation model requires additional validation using real operational and economic data. The current conditional cost coefficients reflect modelling assumptions intended to compare decision strategies under identical conditions; therefore, their applicability to real-world emergency-response economics requires expert calibration and empirical verification.

15. Future research will focus on progressive real-world validation of the proposed framework. The next stage includes collaboration with medical and emergency-response institutions, collection of ethically approved multispectral datasets, evaluation on real RGB–thermal–depth sensor platforms, and comparison with established clinical assessment procedures. Future studies will also investigate model generalisation, uncertainty estimation, and adaptation of the stability analysis for fully nonlinear neural architectures.

All the Declarations and Statements

Author Contributions Statement

Serhii Vladov – Conceptualisation, Data Curation, Project Management and Supervision: Proposed research ideas, constructed the overall framework, and supervised project execution. Handled data acquisition, dataset preprocessing, and implemented the research model.
Zhengbing Hu – Methodology, Writing – Review and Editing, Software Implementation and reviewed and edited the manuscript, ensured clarity and coherence, and helped coordinate project milestones and deadlines.
Dmytro Uhryn – Formal Analysis, Visualisation, and Statistical Analysis: Performed in-depth analysis of experimental results, prepared performance charts, and ensured the statistical robustness of the evaluation.
Yuriy Ushenko – Formal Analysis, Visualisation, and Statistical Analysis: Performed in-depth analysis of experimental results, prepared performance charts, and ensured the statistical robustness of the evaluation.
Victoria Vysotska – Model Training, Validation, and Performance Evaluation: Led the model training process, validated results using standard metrics, and benchmarked performance against existing methods.
Vladyslav Doroshenko – Writing – Drafted the initial manuscript, contributed to the literature survey, and documented the technical background of the study.
Daniil Zinchenko – Model Training, Validation, and Performance Evaluation: Led the model training process, validated results using standard metrics, and benchmarked performance against existing methods.

All authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement

The authors declare no conflicts of interest.

Funding Declaration

The research was carried out with grant support from the Ministry of Education and Science of Ukraine, “Methods and means of identifying human resource damage in emergencies based on deep learning,” under Project No. 0125U001852.
The research was carried out with grant support from the Ministry of Education and Science of Ukraine, “Methods and tools for detecting disinformation in social networks based on deep learning technologies,” under Project No. 0125U001852.

Data Availability Statement

This study analysed its own datasets. Data is contained within the article. The datasets analysed in the current study are not publicly available due to privacy restrictions, but are available from the corresponding author upon reasonable request.

Ethical Declarations

Ethical approval was not required for this study as it did not involve human participants or animals.

Acknowledgments

We sincerely thank the experts for their professional evaluation and valuable recommendations, which have improved the quality of the experiment and the reliability of its results.

Declaration of Generative AI in Scholarly Writing

The authors have not employed any Generative AI tools.

Abbreviations

The following abbreviations are used in this manuscript:

WHO – World Health Organisation
ECG – Electrocardiogram
RGB – Red, green, and blue colour model
CVPR – Conference on Computer Vision and Pattern Recognition
AI – Artificial intelligence
IoMT – Internet of Medical Things
FLOPs – Floating point operations per second
GPU – Graphics processing unit
NPU – Neural processing unit
RAM – Random access memory

References

[1] World Health Organization: Injuries and violence. Available online: https://www.who.int/news-room/fact-sheets/detail/injuries-and-violence

[2] World Health Organization: Road traffic injuries. Available online: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

[3] Y. Xu, T. M. Khan, Y. Song, and E. Meijering, "Edge deep learning in computer vision and medical diagnostics: A comprehensive survey," *Artificial Intelligence Review*, vol. 58, 93, 2025, doi: 10.1007/s10462-024-11033-5

[4] A. Rocha, M. Monteiro, C. Mattos, M. Dias, J. Soares, R. Magalhães, and J. Macedo, "Edge AI for Internet of Medical Things: A literature review", *Computers and Electrical Engineering*, 116, 109202, 2024. https://doi.org/10.1016/j.compeleceng.2024.109202

[5] World Bank Group: A Decade of Saving Lives Through Road Safety Investments. Available online: https://www.worldbank.org/en/results/2024/03/25/a-decade-of-saving-lives-through-road-safety-investments

[6] C. E. Marcussen, K. B. Bräuner, H. Alstrøm, and A. M. Møller, "Accuracy of prehospital triage systems for mass casualty incidents in trauma register studies—A systematic review and meta-analysis of diagnostic test accuracy studies," *Injury*, vol. 53, no. 8, pp. 2725–2733, 2022, https://doi.org/10.1016/j.injury.2022.05.006

[7] T. Cheng, K. Staats, A. H. Kaji, N. D'Arcy, K. Niknam, and J. J. Donofrio-Odmann, "Comparison of prehospital professional accuracy, speed, and interrater reliability of six pediatric triage algorithms," *JACEP Open*, vol. 3, no. 1, e12613, 2022, https://doi.org/10.1002/emp2.12613

[8] J. M. Zachariasse, N. Seiger, P. P. M. Rood, C. F. Alves, P. Freitas, F. J. Smit, G. R. Roukema, and H. A. Moll, "Validity of the Manchester Triage System in emergency care: A prospective observational study," PLOS ONE, vol. 12, no. 2, Art. no. e0170811, 2017, https://doi.org/10.1371/journal.pone.0170811

[9] M. Hegarty-Craver, H. Davis-Wilson, P. Gaur, H. Walls, D. Dausch, and D. Temple, Wearable Sensors for Service Members and First Responders: Considerations for Using Commercially Available Sensors in Continuous Monitoring. Research Triangle Park, NC, USA: RTI Press, 2024.

[10] J. Gutiérrez, V. Rodríguez, and S. Martín, "Comprehensive review of vision-based fall detection systems," Sensors, vol. 21, no. 3, 947, 2021, https://doi.org/10.3390/s21030947

[11] E. Alam, S. Sufian, P. Dutta, and M. Leo, "Vision-based human fall detection systems using deep learning: A review," *Computers in Biology and Medicine*, vol. 146, 105626, 2022, https://doi.org/10.1016/j.compbiomed.2022.105626

[12] Y. He, B. Deng, H. Wang, L. Cheng, K. Zhou, S. Cai, and F. Ciampa, "Infrared machine vision and infrared thermography with deep learning: A review," *Infrared Physics & Technology*, vol. 116, 103754, 2021, https://doi.org/10.1016/j.infrared.2021.103754

[13] D. Kesztyüs, S. Brucher, C. Wilson, and T. Kesztyüs, "Use of infrared thermography in medical diagnosis, screening, and disease monitoring: A scoping review," *Medicina*, vol. 59, no. 12, 2139, 2023, https://doi.org/10.3390/medicina59122139

[14] G. Rouhafzay, S. Rowlands, A. J. Valencia, S. Yang, P. Payeur, H. Tian, and J. Dickens, "Multi-spectral imaging and data fusion for real-time bleeding detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2025, pp. 4561–4568.

[15] A. Rancea, I. Anghel, and T. Cioara, "Edge Computing in Healthcare: Innovations, Opportunities, and Challenges," *Future Internet*, vol. 16, no. 9, 329, 2024, https://doi.org/10.3390/fi16090329

[16] M. Satyanarayanan, "The Emergence of Edge Computing," *Computer*, vol. 50, no. 1, pp. 30–39, 2017, https://doi.org/10.1109/MC.2017.9

[17] D. Borrmann, J. Elseberg, K. Lingemann, and A. Nüchter, "The 3D Hough Transform for Plane Detection in Point Clouds: A Review and a New Accumulator Design," *3D Research*, vol. 2, no. 2, 3, 2011, https://doi.org/10.1007/3DRes.02(2011)3

[18] S. Hwang, J. Park, N. Kim, Y. Choi, and I. S. Kweon, "Multispectral pedestrian detection: Benchmark dataset and baseline," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1037–1045. https://doi.org/10.1109/CVPR.2015.7298706.

[19] M. Brenner, N. H. Reyes, T. Susnjak, and A. L. C. Barczak, "MM5: Multimodal image capture and dataset generation for RGB, depth, thermal, UV, and NIR," Information Fusion, vol. 126, 103516, 2026. https://doi.org/10.1016/j.inffus.2025.103516

[20] M. Brenner, N. H. Reyes, T. Susnjak, and A. L. C. Barczak, "RGB-D and Thermal Sensor Fusion: A Systematic Literature Review," *IEEE Access*, vol. 11, pp. 82410–82442, 2023, https://doi.org/10.1109/access.2023.3301119

[21] S. Kalamkar and A. Geetha Mary, "Multimodal image fusion: A systematic review," Decision Analytics Journal, vol. 9, 100327, 2023. https://doi.org/10.1016/j.dajour.2023.100327

[22] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, 60, 2019, https://doi.org/10.1186/s40537-019-0197-0

[23] F. Yang, G. Zamzmi, S. Angara, S. Rajaraman, A. Aquilina, Z. Xue, S. Jaeger, E. Papagiannakis, and S. K. Antani, "Assessing inter-annotator agreement for medical image segmentation," IEEE Access, vol. 11, pp. 21300–21312, 2023. https://doi.org/10.1109/ACCESS.2023.3249759

[24] S. W. Lee, "Methods for testing statistical differences between groups in medical research: Statistical standard and guideline of Life Cycle Committee," *Life Cycle*, vol. 2, e1, 2022, https://doi.org/10.54724/lc.2022.e1

[25] H. J. Yang and H. S. Lee, "Common statistical methods used in medical research," *Kosin Medical Journal, vol. 40, no. 1, pp. 21–30, 2025, https://doi.org/10.7180/kmj.24.160*

[26] Y. Zhou, Y. Zhu, and W. K. Wong, "Statistical tests for homogeneity of variance for clinical trials and recommendations," *Contemporary Clinical Trials Communications*, vol. 33, 101119, 2023, https://doi.org/10.1016/j.conctc.2023.101119

[27] D. Chicco, A. Sichenze, and G. Jurman, "A simple guide to the use of Student's t-test, Mann–Whitney U test, Chi-squared test and Kruskal–Wallis test in biostatistics," *BioData Mining*, vol. 18, no. 1, 56, 2025, https://doi.org/10.1186/s13040-025-00465-6

[28] G. Hripcsak, L. Zhang, Y. Chen, K. Li, M. A. Suchard, P. B. Ryan, and M. J. Schuemie, "Assessing covariate balance with small sample sizes," Statistics in Medicine, vol. 44, no. 18–19, e70212, 2025. https://doi.org/10.1002/sim.70212

[29] C. Wongoutong, "The impact of neglecting feature scaling in k-means clustering," *PLOS ONE*, vol. 19, no. 12, e0310839, 2024, https://doi.org/10.1371/journal.pone.0310839
[30] A. Naghizadeh and D. N. Metaxas, "Condensed Silhouette: An optimized filtering process for cluster selection in K-means," *Procedia Computer Science*, vol. 176, pp. 205–214, 2020, https://doi.org/10.1016/j.procs.2020.08.022
[31] W. Junthopas, K. Khamdee, and P. Khamdee, "Pre-determining the optimal number of clusters for k-means clustering using the parameters package in R and distance metrics," *Applied Sciences*, vol. 15, no. 21, 11372, 2025, https://doi.org/10.3390/app152111372
[32] I. T. Utami, F. Suryaningrum, and D. Ispriyanti, "K-means cluster count optimization with Silhouette Index validation and Davies Bouldin Index: Case study—coverage of pregnant women, childbirth, and postpartum health services in Indonesia in 2020," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 17, no. 2, pp. 707–716, 2023, https://doi.org/10.30598/barekengvol17iss2pp0707-0716
[33] D. Chicco, A. Campagner, A. Spagnolo, D. Ciucci, and G. Jurman, "The Silhouette coefficient and the Davies-Bouldin index are more informative than Dunn index, Calinski-Harabasz index, Shannon entropy, and Gap statistic for unsupervised clustering internal evaluation of two convex clusters," PeerJ Computer Science, vol. 11, e3309, 2025. https://doi.org/10.7717/peerj-cs.3309
[34] A. J. Suda, A. Franke, M. Hertwig, and K. Gooßen, "Management of mass casualty incidents: a systematic review and clinical practice guideline update," European Journal of Trauma and Emergency Surgery, vol. 51, no. 1, 5, 2025. https://doi.org/10.1007/s00068-024-02727-0
[35] F. Pu, Z. Li, Y. Wu, C. Ma, and R. Zhao, "Recent advances in disaster emergency response planning: Integrating optimization, machine learning, and simulation," Safety Emergency Science, vol. 1, no. 1, 9590007, 2025. https://doi.org/10.26599/SES.2025.9590007
[36] J. Wei, Z. Zou, W. Lai, J. Du, J. Zhao, Z. Liu, and S. Wang, "RGB-Thermal cameras calibration based on Maximum Index Map," Computers and Electrical Engineering, vol. 123, 110234, 2025. https://doi.org/10.1016/j.compeleceng.2025.110234
[37] W. Yang, H. Luo, K.-W. Tse, H. Hu, K. Liu, B. Li, and C.-Y. Wen, "Autonomous–Targetless Extrinsic Calibration of Thermal, RGB, and LiDAR Sensors," IEEE Transactions on Instrumentation and Measurement, vol. 73, pp. 1–11, 2024. https://doi.org/10.1109/TIM.2024.3480237
[38] J. E. Bloom et al., "Healthcare and economic cost burden of emergency medical services treated non-traumatic shock using a population-based cohort in Victoria, Australia," BMJ Open, vol. 14, no. 4, e078435, 2024. https://doi.org/10.1136/bmjopen-2023-078435
[39] M. S. Rasmussen, Y. Zhang, N. Andelic, and E. Aas, "Health care costs and service utilization in the first year following moderate to severe traumatic injury," BMC Health Services Research, vol. 24, 1535, 2024. https://doi.org/10.1186/s12913-024-12016-6
[40] D. Orso, L. Flaibani, U. G. Sisto, M. Bonsano, F. Fonda, R. Pangallo, and T. Bove, "Survival and cost-effectiveness of helicopter versus ground emergency medical services: A systematic review and meta-analysis with meta-regression and trial sequential analysis," Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine, vol. 33, no. 1, 160, 2025. https://doi.org/10.1186/s13049-025-01478-0
[41] T. K. Das, S. K. Das, S. Bissoyi, and D. K. Patel, "Self-adaptive Resource Allocation in Fog-Cloud Systems Using Multi-agent Deep Reinforcement Learning with Meta-learning," *International Journal of Intelligent Systems and Applications (IJISA)*, vol. 18, no. 1, pp. 107–118, 2026, https://doi.org/10.5815/ijisa.2026.01.08
[42] C. S. Odeyemi and O. M. Olaniyan, "Development of a Facenet Enhanced Secured Smart Office System," *International Journal of Intelligent Systems and Applications (IJISA)*, vol. 16, no. 1, pp. 41–51, 2024, https://doi.org/10.5815/ijisa.2024.01.04
[43] V. S. Rana, A. Joshi, and K. K. Verma, "Lightweight 3DCNN-BiLSTM Model for Human Activity Recognition using Fusion of RGBD Video Sequences," *International Journal of Information Technology and Computer Science (IJITCS)*, vol. 17, no. 6, pp. 176–193, 2025, https://doi.org/10.5815/ijitcs.2025.06.10
[44] S. Vladov, L. Chyrun, E. Muzychuk, V. Vysotska, V. Lytvyn, T. Rekunenko, A. Basko, "Intelligent Method for Generating Criminal Community Influence Risk Parameters Using Neural Networks and Regional Economic Analysis," *Algorithms*, vol. 18, no. 8, 523, 2025, https://doi.org/10.3390/a18080523
[45] S. Vladov, Ł. Ścisło, A. Sachenko, H. Foros, D. Zhuravlov, V. Vysotska, V. Varlakhov, S. Sachenko, J. Krupiński, O. Brusakova, "Neural Network DDoS Mitigation System With Forensic Audit Support for Cyber Police," *IEEE Access*, vol. 13, pp. 204628–204655, 2025, https://doi.org/10.1109/access.2025.3634478
[46] S. Vladov, Y. Shmelov, R. Yakovliev, "Optimization of Helicopters Aircraft Engine Working Process Using Neural Networks Technologies," *CEUR Workshop Proceedings*, vol. 3171, pp. 1639–1656, 2022. URL: https://ceur-ws.org/Vol-3171/paper117.pdf
[47] S. Vladov, R. Yakovliev, O. Hubachov, J. Rud, "Neuro-Fuzzy System for Detection Fuel Consumption of Helicopters Turboshaft Engines," *CEUR Workshop Proceedings*, vol. 3628, pp. 55–72, 2023. URL: https://ceur-ws.org/Vol-3628/paper5.pdf
[48] S. Vladov, M. Petchenko, Y. Shmelov, S. Drozdova, R. Yakovliev, "Helicopters Turboshaft Engines Parameters Identification at Flight Modes Using Neural Networks," in Proceedings of the 2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT), Lviv, Ukraine, 2022, pp. 5–8. https://doi.org/10.1109/csit56902.2022.10000444
[49] K. Naoufel, "Towards Automated Diagnosis of Heart Diseases: A Study of Machine Learning Techniques," *International Journal of Engineering and Manufacturing (IJEM)*, vol. 16, no. 2, pp. 143–155, 2026, https://doi.org/10.5815/ijem.2026.02.09
[50] S. Vladov, A. Banasik, A. Sachenko, W. M. Kempa, V. Sokurenko, O. Muzychuk, P. Pikiewicz, A. Molga, V. Vysotska, "Intelligent Method of Identifying the Nonlinear Dynamic Model for Helicopter Turboshaft Engines," *Sensors*, vol. 24, no. 19, p. 6488, Oct. 2024, https://doi.org/10.3390/s24196488
[51] S. Vladov, Y. Shmelov, R. Yakovliev, M. Petchenko, "Neural Network Method for Parametric Adaptation Helicopters Turboshaft Engines On-Board Automatic Control System Parameters," *CEUR Workshop Proceedings*, vol. 3403, pp. 179–195, 2023. URL: https://ceur-ws.org/Vol-3403/paper15.pdf
[52] R. S. Lutfi, "A Neutrosophic-Based Unsupervised Approach for Sudden Drift Detection," *International Journal of Intelligent Systems and Applications (IJISA)*, vol. 18, no. 2, pp. 139–155, 2026, https://doi.org/10.5815/ijisa.2026.02.10

[53] S. Sularno, B. Wendi, A. P. Putri, A. Kamal, and F. Wang, "Data-driven classification of tsunami evacuation suitability using XGBoost: A case study in Padang City," International Journal of Intelligent Systems and Applications (IJISA), vol. 18, no. 1, pp. 90–106, 2026. https://doi.org/10.5815/ijisa.2026.01.07
[54] G. Wali and C. Bulla, "A Resource Management Model for Healthcare Internet of Things Using Deep Learning and Bio-inspired Algorithms," *International Journal of Information Engineering and Electronic Business (IJIEEB)*, vol. 17, no. 1, pp. 68–89, 2025, https://doi.org/10.5815/ijieeb.2025.01.05
[55] S. Ablamskyi, E. Kazić Çakar, and I. Hloviuk, "A comparative study of wartime rapes: The experience of Bosnia and Herzegovina and Ukraine," *Baltic Journal of Law & Politics*, vol. 17, no. 1, pp. 136–162, 2024, https://doi.org/10.2478/bjlp-2024-0007
[56] L. C. Gamarra-Amaya, I. Hloviuk, and S. Ablamskyi, "Implementation of international legal frameworks to combat domestic violence in Ukraine and Colombia," *Via Inveniendi et Iudicandi*, vol. 18, no. 2, pp. 209–233, 2024, https://doi.org/10.15332/19090528.9740
[57] S. Ablamskyi, V. Galagan, I. Basysta, and Z. Udovenko, "Detention of a person suspected of committing a criminal offense during martial law in Ukraine," *Revista Brasileira de Políticas Públicas*, vol. 13, no. 3, pp. 452–467, 2023, https://doi.org/10.5102/rbpp.v13i3.9129
[58] V. V. Romaniuk and S. Ye. Ablamskyi, "Criteria for the admissibility of digital (electronic) evidence in criminal proceedings," *Law and Safety*, vol. 93, no. 2, pp. 140–150, 2024, https://doi.org/10.32631/pb.2024.2.13
[59] S. Ablamskyi, A. Kavunska, O. Perederii, O. Tymofiiv, and A. Zuhdi, "Adaptation of the legal system of Ukraine in the era of digitalization: Challenges, threats and possible ways to overcome them," *Revista Jurídica Portucalense*, no. 37, 2025, https://doi.org/10.34625/issn.2183-2705(37)2025.ic-19
[60] V. Sokurenko, K. Levchenko, S. Ablamskyi, O. Pereira de Lima Jr, G. Šimić, and V. Romaniuk, "The legal regime of the martial state in Ukraine and its impact on the conduct of pre-trial investigation: New challenges and their overcoming," *Via Inveniendi et Iudicandi*, vol. 19, no. 2, pp. 92–111, 2024, https://doi.org/10.15332/19090528.10671
[61] S. Ablamskyii, I. Hloviuk, E. Kazić Çakar, and I. Zinkovskyy, "Addressing enforced disappearances in criminal law: Legal and procedural perspectives from Ukraine and Bosnia and Hercegovina," *Prawo i Więź*, vol. 57, no. 4, pp. 896–914, 2025, https://doi.org/10.36128/PRIW.VI57.1329
[62] S. Ye, T. Wang, M. Ding, and X. Zhang, "F-DARTS: Foveated Differentiable Architecture Search Based Multimodal Medical Image Fusion," *IEEE Transactions on Medical Imaging*, vol. 42, no. 11, pp. 3348–3361, Nov. 2023, https://doi.org/10.1109/TMI.2023.3283517.
[63] X. Zhou, Y. Sun, M. Deng, W. C. W. Chu, and Q. Dou, "Robust Semi-supervised Multimodal Medical Image Segmentation via Cross Modality Collaboration," in *Medical Image Computing and Computer Assisted Intervention (MICCAI 2024)*, Springer, LNCS 15001, pp. 57–67, 2024, https://doi.org/10.1007/978-3-031-72378-0_6.
[64] S. Ding, J. Li, J. Wang, S. Ying, and J. Shi, "Multimodal Co-Attention Fusion Network With Online Data Augmentation for Cancer Subtype Classification," *IEEE Transactions on Medical Imaging*, vol. 43, no. 11, pp. 3977–3989, Nov. 2024, https://doi.org/10.1109/TMI.2024.3405535.
[65] R. Berzoini, E. D'Arnese, D. Conficconi, and M. D. Santambrogio, "NERONE: The Fast Way to Efficiently Execute Your Deep Learning Algorithm at the Edge," *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 3, pp. 1252–1260, Mar. 2024, https://doi.org/10.1109/JBHI.2023.3296142.
[66] J. M. Rodríguez Corral, J. Civit-Masot, F. Luna-Perejón, I. Díaz-Cano, A. Morgado-Estévez, and M. Domínguez-Morales, "Energy Efficiency in Edge TPU vs. Embedded GPU for Computer-Aided Medical Imaging Segmentation and Classification," *Engineering Applications of Artificial Intelligence*, vol. 127, 107298, 2024, https://doi.org/10.1016/j.engappai.2023.107298.
[67] X. Xu, K. Yan, S. Han, B. Wang, X. Tao, and P. Zhang, "Learning-Based Edge-Device Collaborative DNN Inference in IoVT Networks," *IEEE Internet of Things Journal*, vol. 11, no. 5, pp. 7989–8004, Mar. 2024, https://doi.org/10.1109/JIOT.2023.3317785.

Authors' Profiles

Serhii Vladov, doctor of technical sciences. Department of Scientific Activity Organisation, Kharkiv National University of Internal Affairs, Ukraine, leading researcher. Department of Combating Cybercrime, Kharkiv National University of Internal Affairs, Ukraine, professor. Computer Science Department, Yuriy Fedkovych Chernivtsi National University, Ukraine, chief scientist.
Major interests: artificial intelligence, neural networks, national security, cybersecurity, financial security, and intelligent data analysis methods.

Zhengbing Hu, professor. School of Management, Huazhong University of Science and Technology, Wuhan, China.
Major interests: computer science and technology applications, artificial intelligence, network security, communications, data processing, cloud computing, and education technology.

Dmytro Uhryn, doctor of technical sciences, professor, Computer Science Department, Yuriy Fedkovych Chernivtsi National University, Ukraine.
Major interests: decision-support information technologies, swarm intelligence systems, and industry geoinformation systems.

Yuriy Ushenko, doctor of physical and mathematical sciences, professor, Computer Science Department, Yuriy Fedkovych Chernivtsi National University, Ukraine, chief.
Major interests: information systems design, data mining, pattern recognition, digital image processing, artificial neural networks, laser polarimetry and interferometry.

Victoria Vysotska, doctor of technical sciences, associate professor. Information Systems and Networks Department, Lviv Polytechnic National University, Ukraine, professor. Department of Combating Cybercrime, Kharkiv National University of Internal Affairs, Ukraine, professor.
Major interests: natural language processing (NLP), computational linguistics, data science, systems analysis, information technologies, machine learning, and cybersecurity.

Vladyslav Doroshenko, doctor of philosophy, associate professor. Department of Law Enforcement, Kharkiv National University of Internal Affairs, Ukraine, associate professor.
Major interests: criminal law, criminology and law enforcement in times of emergency and military conflict, internal security policy, legal regulation of emergencies and law enforcement activities, and artificial intelligence.

Daniil Zinchenko, Department of Postgraduate (Adjunct) and Doctoral Studies, Kharkiv National University of Internal Affairs, Ukraine, adjunct.
Major interests: cybersecurity and information protection, legal aspects of digital evidence, legal regulation of emergencies, and artificial intelligence.

How to cite this paper:

Serhii Vladov, Zhengbing Hu, Dmytro Uhryn, Yuriy Ushenko, Victoria Vysotska, Vladyslav Doroshenko, Daniil Zinchenko, "Lyapunov-Stable Multispectral Edge Deep Learning for Critical Injury Triage and Resource Optimisation in Emergency Response", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.18, No.5, pp. 126-184, 2026. DOI:10.5815/ijigsp.2026.05.08