Journal of Power Sources 629 (2025) 235982

Contents lists available at ScienceDirect

Journal of Power Sources

journal homepage: www.elsevier.com/locate/jpowsour

AI-enabled thermal monitoring of commercial (PHEV) Li-ion pouch cells with Feature-Adapted Unsupervised Anomaly Detection

Abdelrahman Shabayek [a,**], Arunkumar Rathinam [a], Matthieu Ruthven [a], Djamila Aouada [a], Tazdin Amietszajew [b,*]

[a] Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg
[b] Centre for E-Mobility and Clean Growth, Coventry University, Coventry, United Kingdom

HIGHLIGHTS

- Unsupervised thermal anomaly detection.
- Li-ion cell thermal profiling.
- Deep learning using images as training data.

ARTICLE INFO

Keywords:
AI
Thermal
Battery
Simulation
Anomaly detection
Unsupervised learning

ABSTRACT

Distributed temperature profiling of lithium-ion batteries provides valuable insights, aiding thermal management and minimising risk of battery failures. Highlighted by Batteries Europe as crucial for battery safety, advances in thermal monitoring are imperative to continuous safe adoption of battery technology. Deep Learning techniques have recently emerged as powerful tools for anomaly detection (AD) in many thermal mapping applications. These data-driven methods can handle common challenges like data unavailability or environment variations. Our study devises a methodology to leverage Deep Learning with thermal data from commercially available pouch cells and an infrared camera. We explain the building blocks of FAUAD (Feature-Adapted Unsupervised Anomaly Detection), which models the normality of the input data and synthesizes anomalies in its feature space. The resulting model is benchmarked against some of the latest state-of-the-art methods and achieves high anomaly detection capability; Area Under the ROC Curve (AUROC) score of 0.971 on simulated data, 0.990 on contaminated real data, and a perfect score of 1.0 on real clean data. While maintaining a compact size of 15 MB. FAUAD offers a notable advancement in unsupervised anomaly detection for battery thermal monitoring. The proposed method is cell chemistry agnostic and open to usage scenarios beyond this works' scope.

1. Introduction

Lithium-ion cells have solidified themselves as the dominant energy storage solution within the realm of portable electronics [1], electric vehicles [2] and grid storage [3]. Amidst pressing global environmental concerns, the pursuit of a more sustainable future further accelerated the adoption of energy storage systems in all aspects of life. Nevertheless, the high energy density and thermal instability inherent to Li-ion batteries introduce significant challenges to ensuring safety, necessitating robust thermal management systems [4]. This is especially challenging in high performance systems [5], due to significant internal Joule heating [6]. Furthermore, the phenomenon of battery thermal runaway, a potential catastrophic failure mechanism, became one of the greatest challenges for battery safety often exacerbated through a cascading domino effect [7,8]. Recognised by Batteries Europe as one of the key areas of battery safety [9], advances in thermal monitoring of battery cells are imperative to continuous safe adoption of this technology across the wider market.

At present, to monitor cells operating parameters and estimate their State of Health, they are fitted with sensors that monitor voltage, current

* Corresponding author.
** Corresponding author.
E-mail addresses: abdelrahman.shabayek@uni.lu (A. Shabayek), taz.amietszajew@coventry.ac.uk (T. Amietszajew).

https://doi.org/10.1016/j.jpowsour.2024.235982
Received 2 September 2024; Received in revised form 12 November 2024; Accepted 28 November 2024
Available online 3 December 2024

flow and surface temperature. However, the common practice of single-point surface temperature measurements does not accurately represent a true condition due to its limited spatial resolution [10]. Consequently, this method overlooks critical occurrences like thermal hotspots, uneven heat distribution, or atypical events, which can lead to permanent damage to the electrochemical system in use. Additionally, various thermal-management strategies used, e.g. air flow cooling [11], can cause point-based instrumentation to produce inaccurate data. Distributed temperature profiling [12] produces significantly more insightful data, which could offer an insight into the events preceding cell failure, support the optimisation of thermal management and minimise battery degradation [13,14]. Certain diagnostic techniques, e.g. Differential Thermal Voltammetry [15] or recent Artificial Intelligence powered State-of-Charge models [16], require high fidelity thermal data for robustness and long-term operation [17].

Deep learning techniques have emerged as powerful tools for thermal anomaly detection (AD) [18,19]. These techniques are particularly important due to their ability to identify unusual patterns. Advanced deep learning models facilitate accurate and efficient fault detection, contributing to the reliability, longevity, and safety of battery cells. The use of these techniques is critical in improving performance and ensuring safety, as early detection and tracing of anomalous operations in batteries are essential [20]. Furthermore, these data-driven methods can handle common challenges like data unavailability, environment variations, and battery aging [19]. Therefore, the application of machine and deep learning techniques for AD in thermal images of batteries is not only a promising research direction but also a crucial step towards safer and more efficient energy systems.

In this work we aim to leverage Deep Learning combined with cell thermal data acquisition. Learning normal battery thermal propagation patterns, a system can be trained to spot anomalous thermal patterns that deviate from the learned normality. The methodology described results in a system capable of identifying thermal anomalies, while coping with limitations like computational constraints, limited training data, noise, and dynamic operating conditions. Our approach offers a way to increase battery safety through early anomaly detection, leveraging AI tools to analyse thermal maps, capturing the interplay between temperature measurements, rather than relying solely on localised readings.

2. Methodology

The main objective is to build a model capable of detecting anomalies in the thermal patterns of battery cells with high accuracy. Our approach firstly uses a pretrained deep neural network model to extract features [21]. Given that the extracted features are generic, it subsequently adapts these features to represent input thermal images of the batteries [22]. As the anomalies are very rare, we follow the ideology of One Class Classification (OCC) anomaly detection algorithms where a model is trained on what is considered "normal" data and then using that model to detect whether new data is normal or not [23]. Therefore, pseudo anomalies are generated to properly train the deep neural network model to recognize normal features and reject pseudo anomalies [24]. Finally, the model learns to discriminate the normality patterns from others (i.e. anomalies).

2.1. Proposed model

Herein we propose FAUAD (Feature-Adapted Unsupervised Anomaly Detection). It models the normality of the input data and synthesizes anomalies in its feature space. Fig. 1 a) shows the training process where only normal images are given as an input to the deep neural network, while Fig. 1 b) shows that the network is tested using normal and anomalous images (see Section 2.5) (see Fig. 2).

Following are the building blocks of FAUAD:

Features extractor: Similarly, as in Ref. [25], we leverage a light-weight pretrained resnet18 model in PyTorch [26] to obtain local features from the first three layers of the network to ensure compatibility with edge devices and real-time applications as required in Electric Vehicles.

Features adaptor: To bridge the gap between the captured thermal images and the extracted generic features, a feature adaptor is employed. This adaptor, consisting of a single fully connected layer without a bias term [22], transfers the training features to the batteries in the thermal domain.

Pseudo anomalies generator: To provide a tractable way to simulate real-world thermal anomalies, Poisson noise is added to normal features [27] to create pseudo anomalies. Gaussian noise has bee also tested akin to Ref. [22] but its performance is slightly behind FAUAD using Poisson noise. This is discussed in Section 4.2.

Features classifier: Finally, a two-layer multi-layer perceptron (MLP) classifier [28] is trained to distinguish between normal and anomalous features, acting as a normality scorer and outputting high values for genuine data and low values for anomalies.

The following figure gives a simplified explanation of these building blocks. For a detailed definition of the layers please refer to Ref. [29].

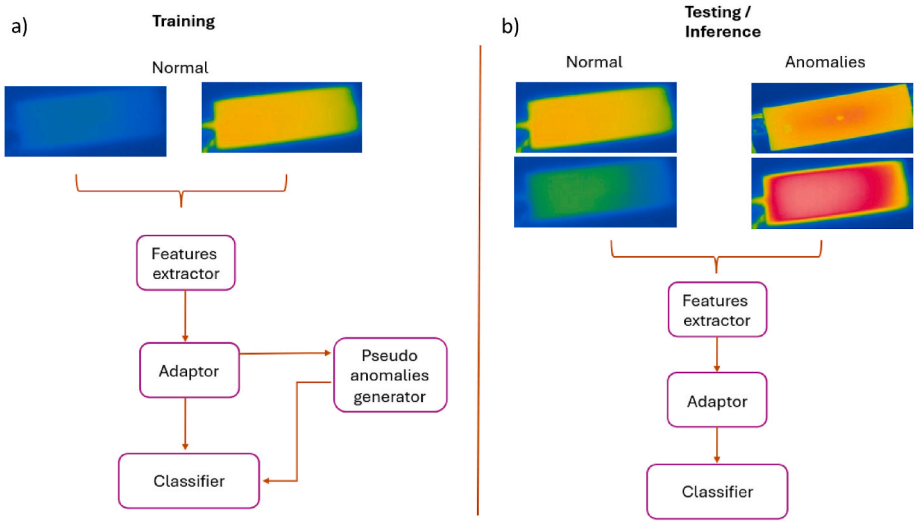

Fig. 1. A simplified illustration of FAUAD for unsupervised anomaly detection in thermal images of batteries, schematically representing a) training process and b) testing and verification steps.

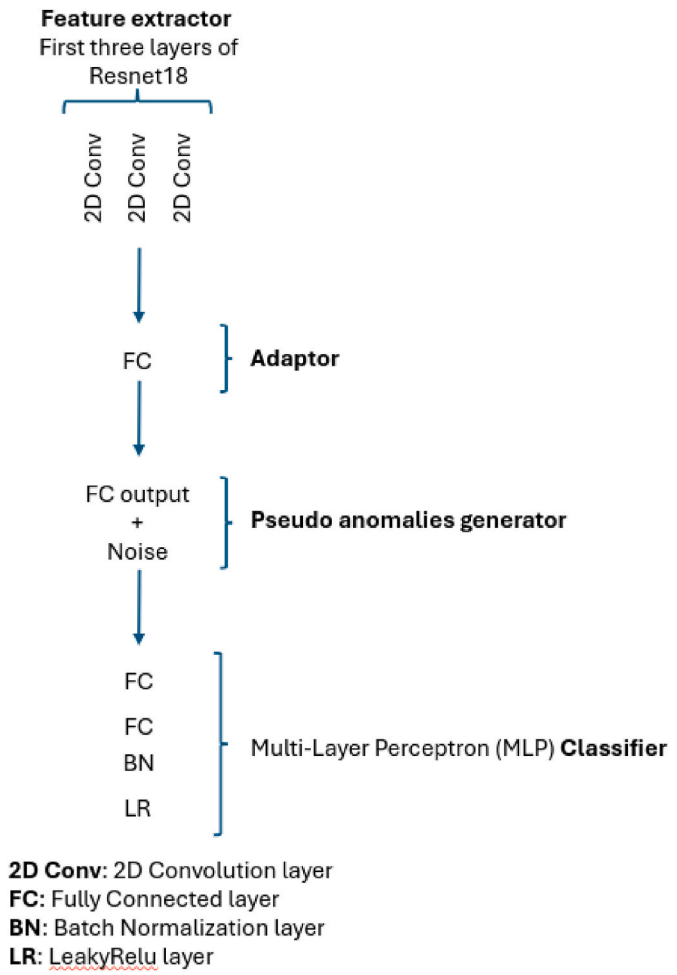

Fig. 2. A schematic illustration of the Feature-Adapted Unsupervised Anomaly Detection system building blocks.

2.2. Anomaly detection

Anomaly Detection (AD) can be either supervised, where anomalous examples are provided, or unsupervised, where only normal examples are available. Hence, unsupervised AD offers a compelling alternative to supervised methods for identifying anomalies [30]. This approach eliminates the significant effort required to collect and label anomalous samples, which are often rare in real-world applications. Furthermore, it bypasses the time-consuming and expensive process of labelling training data altogether. Finally, unsupervised methods are immune to labelling bias, a common issue in supervised learning where human error during annotation can skew the model. In essence, unsupervised AD leverages readily available normal data to learn the characteristics of healthy images, allowing it to effectively flag deviations from this norm as potential defects. This technique, also known as outlier detection or one-class classification [23], focuses on identifying individual anomalous images within a larger pool of normal ones.

The field of AD has been extensively explored, with numerous surveys available to delve deeper. This plethora of resources reflects the ongoing advancements and diverse approaches within the field. To provide a concise overview, we will highlight some of these surveys. While Ehret et al. [31] focused on classical approaches, omitting recent deep learning advancements, Chalapathy et al. [32] explored deep AD across supervised, semi-supervised, and unsupervised domains. Similarly, Tao et al. [30] offers a recent review on unsupervised AD specifically.

Bridging the gap between traditional and deep learning methods, Ruff et al. [33] provides a comprehensive analysis of their connections. Mohammadi et al. [34] dives deeper into image/video deep learning for AD, classifying methods into self-supervised learning, generative networks, and anomaly generation. Building upon this, Pang et al. [35] proposes a thorough taxonomy for deep AD, encompassing advancements in these three categories alongside other refinement categories.

Our approach aims to be simple yet effective. It is based on pipelining a series of simple steps starting from feature extraction, passing by domain adaptation and finally learning the normality patterns by exposing the discriminator to normal features and pseudo anomalies. Hence, the proposed model (FAUAD) has been compared with some of the latest state-of-the-art methods that follow similar strategies. These approaches can be broadly classified into Normalization Flows [36,37] based methods where features are mapped into a lower dimension and Deep feature embedding based methods where features are learned through knowledge distillation [38–44].

These methods excel at finding anomalies. They achieve this by first extracting features from the data, then condensing those features into a lower-dimensional space. This compression makes it simpler to identify anomalies, as they typically deviate significantly from the clusters formed by the compressed normal features. However, these techniques have some drawbacks. Real-world data can differ visually from the training data used for pre-training models; this domain mismatch may lead to inaccurate anomaly detection. Additionally, some methods involve computationally expensive steps like calculating covariance inverses or searching large memory banks, limiting their use on resource-constrained devices or for real-time applications. Furthermore, techniques that rely on normalizing flows can be memory-intensive due to the need for processing full-sized features and utilizing memory-hungry layers.

Our approach addresses the limitations of the above methods by tackling them through three key strategies. Firstly, it creates synthetic features within a specialized feature space for battery data, leading to more accurate anomaly representation. Secondly, it incorporates a "feature adapter" to bridge the gap between pre-trained models and battery images, improving effectiveness. Finally, a simplified yet efficient architecture facilitates faster training, inference, and deployment. These innovations unlock significant advantages. By leveraging unsupervised learning, the model can learn from normal data and identify deviations without needing pre-defined anomaly examples, making it adaptable to unforeseen scenarios. Furthermore, the approach can identify diverse and unexpected anomaly types, offering crucial benefits for early detection of potential battery failures and preventing catastrophic events.

2.3. Cell simulation

Prior to capturing real-world thermal images of lithium-ion cells, we conducted the first stage of our experiments using simulated pouch cell images [45] generated with COMSOL Multiphysics® software in combination with Matlab [46]. The major toolboxes used are "Battery Design Module" toolbox for battery thermal propagation and "LiveLink for MATLAB" for communication between MATLAB and COMSOL. This initial phase served a critical purpose: validating the effectiveness of our proposed algorithm for anomaly detection in thermal data.

The COMSOL-generated images encompassed three key scenarios:

- **Normal Thermal Patterns:** These images represented batteries functioning within their expected temperature ranges.
- **Overheating Thermal Patterns:** Simulated high-temperature events were depicted with reddish colour progressions, signifying potential overheating situations.
- **Non-Homogeneous Thermal Patterns**: Simulated abnormal heat propagation on the battery surface that created distinct thermal regions, characterized by bluish, gradual cold spots.

There were several compelling reasons for employing simulated experiments at this initial stage:

1. **Controlled Environment:** COMSOL allowed us to create a precisely controlled environment for testing. We could tailor the simulated images to include a wide range of thermal conditions, encompassing both normal and dangerous scenarios. This level of control would not be reasonably executed with real-world battery experiments.
2. **Cost-Effectiveness:** Simulations are significantly less resource and time-intensive than real-world tests. This initial validation phase allowed us to refine our algorithm efficiently before moving on to real-world testing.
3. **Safety Considerations:** Lithium-ion batteries pose a safety hazard during thermal anomaly events. By using simulated data, we could explore a broader range of dangerous scenarios without risking damage to the surrounding environment.

The simulated thermal images have been structured to extend the MVTec [38] industrial AD dataset for the classification task. The "simulated battery AD dataset" folder contains "train" and "test" folders. The train folder contains normal data only in a folder called "good". The "test" folder contains the "good" folder for normal data and remaining folders ("overheat" and "non-homogeneous"). The overheat folder contains simulations with reddish propagation patterns. The non-homogeneous folder contains simulated images with bluish gradual circular shaped regions that introduce non-homogeneous heat propagation patterns. We have manipulated the original model [45] to assign different thermal conductivity to random circular regions on the battery surface to produce the non-homogenous heat propagation anomalies. Fig. 3 shows examples of the simulated images introduced to FAUAD to learn and test normal patterns and examples for testing anomalies of overheat and non-homogenous distribution.

The simulated experiments provided a safe, cost-effective, and controlled environment to validate our anomaly detection algorithm. This initial validation stage ensured a higher likelihood of success when we transitioned to capturing real-world thermal images of lithium-ion batteries.

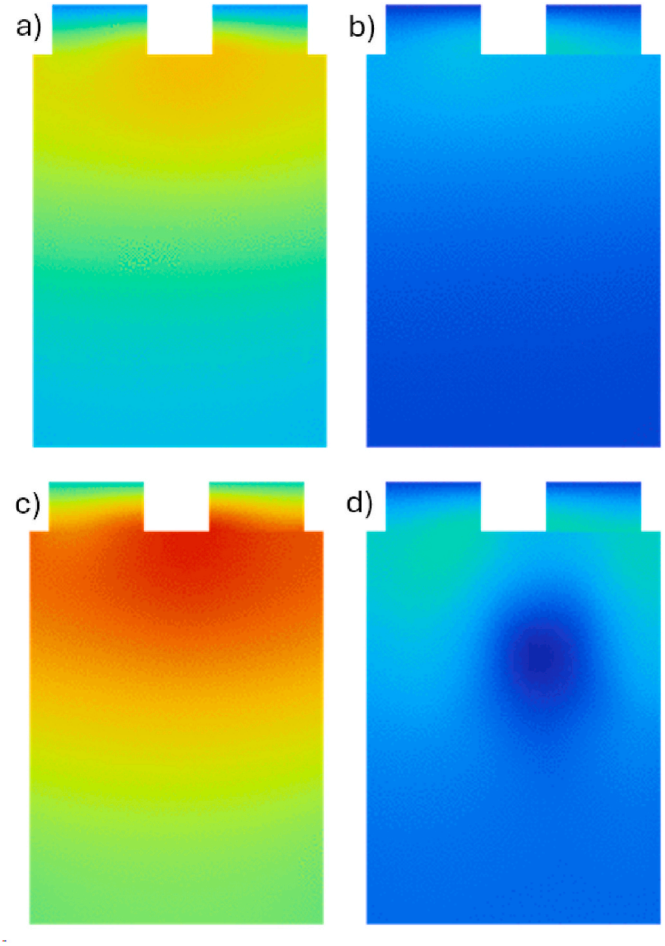

Fig. 3. Simulated images showing a) & b) normal behaviour, c) localised overheating and d) abnormal distribution.

2.4. Cell testing

Experiments were performed using commercially available SPIM11309102-GL40 Li-ion NMC pouch cells from MGL, intended for Plug-in Hybrid Electric Vehicle application, with a rated capacity of 40Ah. The cells are 309 mm × 102 mm with opposing tabs, with an average weight of 730g, resulting in 200 Wh/kg gravimetric energy density at nominal voltage of 3.65V. The cell schematics are shown in Fig. 4. The cells were connected via brass block terminals, mounted to an acrylic baseplate, to ensure mechanical stability and good connection to the high current test leads fastened via lug bolts.

Standard test cycles consisting of constant-current (CC) followed by constant-voltage (CV) charge and constant-current (CC) discharge were performed. Cells were cycled between 2.5V (0 % SoC) and 4.2V (100 % SoC) with a BTS9000 multi-channel potentiostat (NEWARE) capable of providing a maximum continuous current of 200A per cell. Specific cycling regimes are detailed in Table 1. All tests were conducted under 25 °C ambient temperature in a LabEvent environmental chamber (Weiss Technik GmbH).

2.5. Image acquisition

For AI thermal image analysis model training, real thermal images had to be obtained. This has been performed using FLIR T640 infrared camera (FLIR Systems). This system is rated at ± 2% of reading accuracy, capturing 640 x 480 pixel images at 24bit depth using sRGB colour representation. Focal length of 13 mm and exposure time of 1/46s are pre-set. Thermal Images were obtained in a time-lapse mode with 15s intervals, resulting in thousands of images per test. The camera was suspended above the cell being evaluated, as shown in Fig. 5 a). To avoid

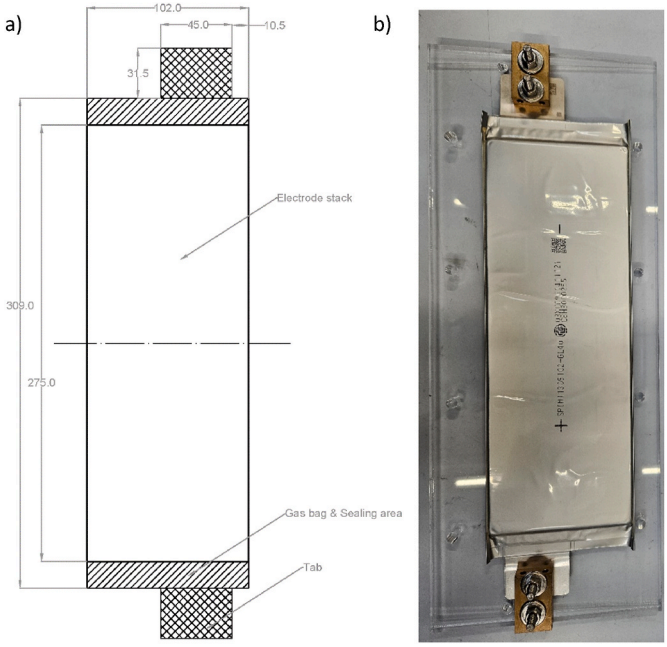

Fig. 4. Tested pouch cell a) dimensions and b) mounted in the test jig.

Table 1
Cell characterisation test profiles.

Test	Steps	Current/Voltage input	Limit
Static capacity	Constant Current (CC) Charge	40A (1C)	4.2V (100 % SoC)
	Constant Voltage (CV) Charge	4.2V/Variable current	2A (C/20)
	Rest	Open circuit	1h
	Constant Current (CC) Discharge	280 mA (C/5)	2.5V (0 % SoC)
	Rest	Open circuit	1h
	Repeat above (5 cycles)		
Waterfall test	Constant Current (CC) Charge	40A–180A (1C–4.5C)	4.2V (100 % SoC)
	Constant Voltage (CV) Charge	4.2V/Variable current	2A (C/20)
	Rest	Open circuit	2h
	Constant Current (CC) Discharge	280 mA (C/5)	2.5V (0 % SoC)
	Rest	Open circuit	2h
	Repeat above (5 cycles total) in 0.5C increments from 1C to 4.5C		

light reflections overlapping the thermal image, the surface of the cell was sprayed with a matte black paint (Ambersil) as advised by the thermal camera manufacturer, and the test chamber was blacked out for the duration of the test. Cell cycling performance data was collected alongside. The combined data obtained was used as training material for the AI model. To replicate non-homogeneity, while coating some of the cell with matte black paint, an area was masked, leaving a reflective patch, as shown in Fig. 5 b). That patch resulted in spatial anomaly (Fig. 5 c), leaving areas of sharp temperature gradient, which would indicate abnormal cell operation and areas of potential cell failure. Similar to the simulated dataset, the captured thermal images have been structured to extend the MVTec [38]. The "battery AD dataset" folder contains "train" and "test" folders. The train folder contains normal data only in a folder called "good". The "test" folder contains the "good" folder for normal data and remaining folders ("overheat" and "reflection") to represent anomalies. Examples of normal and anomalous data are shown in Section 2.6.

To ensure high data quality, a professional thermal image acquisition camera was employed. However, this approach poses a significant barrier to the widespread application of the solution. While high-quality training data is crucial for achieving high model accuracy, the deployment of such systems in compact devices remains impractical. Nevertheless, ongoing advancements in microelectronics have led to the development of micron-sized photodetectors in the infrared range, which present a potential solution to the challenges posed by constrained environments. This article focuses on the AI model proposition - future work includes focus on the miniaturization of the anomaly detection capable hardware.

2.6. Normal vs. abnormal propagation patterns

Thermal cameras typically offer two colormap scale settings: fixed and variable. Fixed scales present a predefined temperature range with consistent colour associations. Variable scales, on the other hand, dynamically adjust the colormap based on the minimum and maximum temperatures captured in the current image.

In the context of electric vehicles (EVs), using a fixed colormap scale is crucial for accurate anomaly detection. This ensures that heat propagation patterns are consistently represented, allowing the algorithm to effectively differentiate between normal thermal behaviour (represented by cooler colours) and dangerous situations (represented by hotter colours).

While simulating these thermal patterns with a fixed scale in COMSOL is straightforward, replicating overheating scenarios in real-world EV batteries poses a significant safety risk. Reaching such extreme temperatures can trigger thermal runaway, a potentially catastrophic event.

Therefore, for our algorithm validation, we can leverage real thermal images captured with a variable scale for testing, using images with reddish colour regions (anomalies) that typically represent high temperatures. These regions, even though displayed at a different absolute temperature value due to the variable scale, still exhibit the same fundamental thermal propagation patterns that would be observed at a fixed scale. Our anomaly detection algorithm doesn't rely on absolute temperature values, but rather learns the normal propagation patterns of heat across the image. Any deviations from this learned norm, which would be represented by colours at the high end of the variable scale, will be flagged as potential anomalies by the algorithm. This approach allows us to effectively test our algorithm using real-world data while maintaining safety during the validation stage.

Hence, under normal operation, a thermal image of a battery should display a consistent pattern of heat dissipation with cooler colours. This is shown in Fig. 6 a) & b). In case of overheat (hotter colours), while the thermal distribution appears uniform, the shades of red colour indicate the presence of dangerously high temperatures. This is shown in Fig. 6 c) & d).

Finally, to replicate non-homogenous heat propagation anomalies,

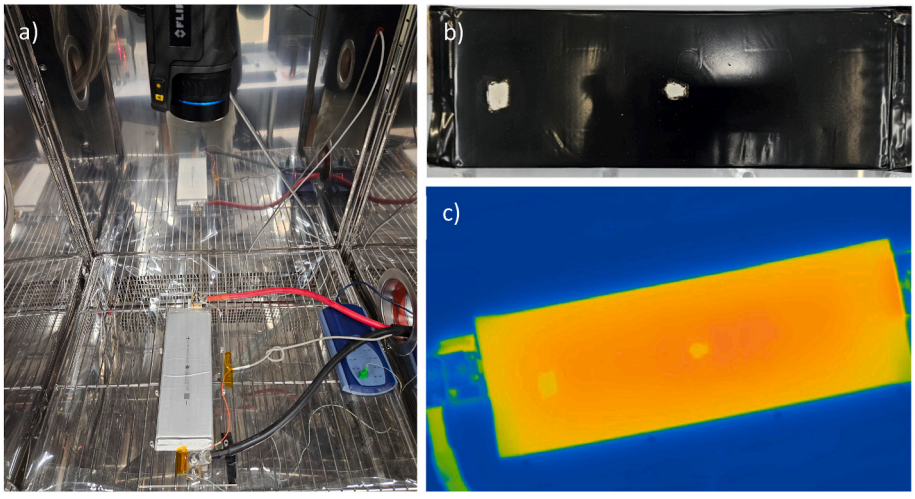

Fig. 5. a) Cell thermal camera set-up, b) painted cell and c) the resulting thermal image. Coating the cell with matte black paint, an area was masked and then uncovered, leaving a reflective patch resulting in anomalous data.

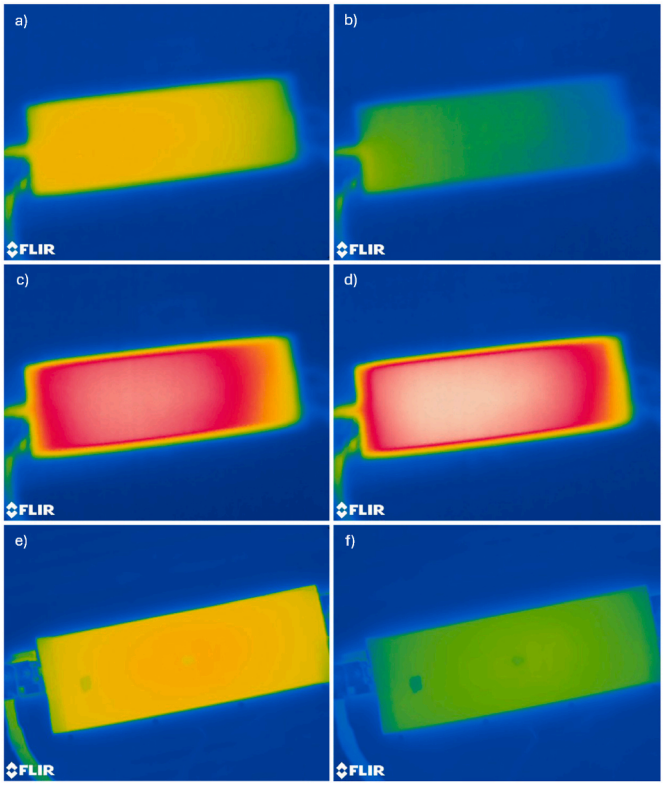

Fig. 6. Heat propagation pattern examples of a-b) 'normal' operation, c-d) 'overheat' scenario and e-f) 'non-homogeneous' propagation anomaly.

we can leverage reflections in thermal images to simulate non-uniform heat patterns. These reflections may cause thermal readings colours in the image to deviate from the expected homogenous distribution on the battery surface which resembles discontinuous propagation anomalies. This is shown in Fig. 6 e) & f).

3. Results & discussion

3.1. Cell testing

Prior to thermal testing and images acquisition, the cells were tested on receipt. At 1C cycling rate, the real capacity was measured at 42.5Ah, exceeding the minimum manufacturer's rating of 40Ah, as shown in Fig. 7, indicating the cells are in good health. Subsequent waterfall tests of increasing C-rate were conducted to assess the safe operating parameters of the cells under evaluation, with the results shown in Fig. 8.

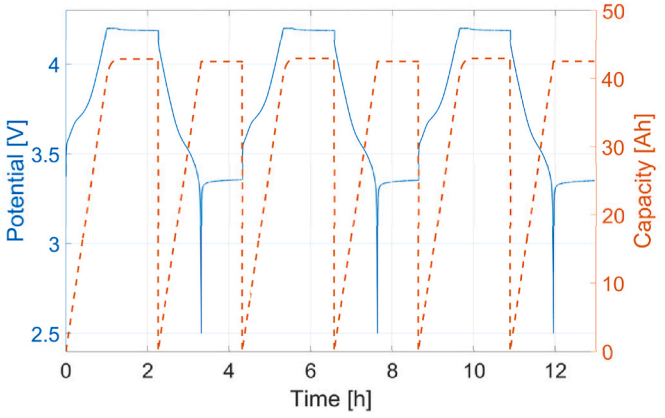

Fig. 7. Baseline static capacity (CC-CV) test at 1C cycling rate.

Stable thermal behaviour, with temperatures within manufacturer-specified operating window, were maintained under ambient cooling as specified in the methodology section (see Fig. 9).

3.2. Deep learning-based anomaly detection

To evaluate FAUAD's effectiveness, we initially conducted tests using the simulated pouch cell images generated with COMSOL, see Section 2.3. These simulations encompassed a wide range of scenarios, including both normal thermal behaviour and dangerous anomalous events (overheating and non-homogenous thermal patterns). Gratifyingly, the algorithm achieved a 0.971 AUROC in anomaly detection within the simulated data. This performance on controlled test data provided strong basis for us to proceed with empirical tests.

FAUAD had also been tested using Poisson [27] and Gaussian [22] noise. Although Poisson noise is simple, it is chosen to simulate rare (anomalous) events in thermal images of the pouch cell that are beyond its normal thermal distribution. It is important to note that this noise was added to the feature space rather than the image space. Given the thermal images homogeneous distribution pattern, Poisson noise was a better candidate to generate pseudo anomalies that are closer to the normal feature space. This in turn enabled the discriminator to learn a boundary that properly encloses the normal data features. An ablation study to compare Poisson and Gaussian noises is given in Table 2. It is noticed that using Poisson noise is slightly better than using Gaussian noise in the contaminated data scenario. Gaussian noise is expected to achieve better performance with more complex distributions possibly observing multiple cells within a battery pack, which is beyond this study. Gaussian noise is a better candidate when a larger diversity in the pseudo-anomalies' generation is required to handle more complex types of images like industrial images [22].

We compared FAUAD's performance on real captured images, see Sections 2.5 and 2.6, to recent state-of-the-art algorithms (detailed in Section 2.2) using two controlled scenarios designed to reflect real-world data acquisition challenges:

- **Simulating Real-World Anomalies (10 %):** This scenario injects a small percentage of anomalies (10 %) into the data, mimicking potential issues that might arise during data collection. While this controlled setting may not perfectly mirror real-world battery data (where environments are typically well-controlled for safety), it allows us to evaluate FAUAD's effectiveness in learning normal data patterns compared to existing methods.
- **Clean Data Training (Control):** The second experiment serves as a control, training all models on clean, anomaly-free data.

This two-part evaluation approach allows us to assess the robustness of all models under both realistic (potentially noisy data) and ideal (clean data) data collection conditions.

Performance comparison of various anomaly detection methods using the Area Under the ROC Curve (AUROC) metric is shown in Table 3, and visually represented in Fig. 9. AUROC has been used as it is the standard measure used for Anomaly Detection applications in the compared state of the art. It is the default measure used in the "anomalib" library [25] by Intel®. The evaluation considers both clean and contaminated data to assess the models' ability to distinguish anomalies under different conditions. It's important to note that all compared models leverage a pre-trained ResNet-18 network for feature extraction, following the protocol established by Akcay et al. [25]. Specifically, the first three layers of the pre-trained network are used to extract features for all methods in this comparison.

FAUAD seems to stand out for its anomaly detection capability. It achieves an impressive Area Under the ROC Curve (commonly called AUROC or AUC in literature) score of 0.990 on data that has been manipulated to include anomalies (contaminated data). This indicates a strong ability to distinguish anomalies from normal data in these

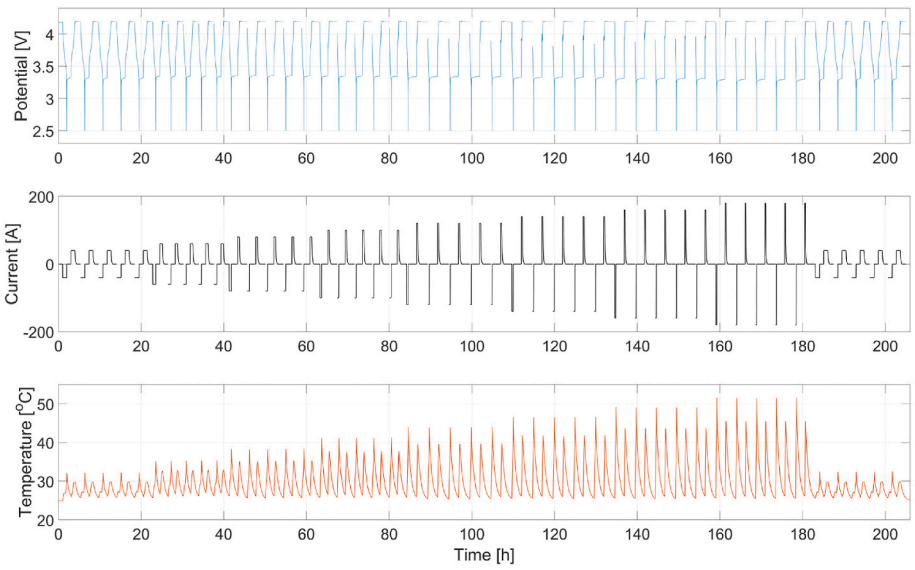

Fig. 8. Cycling rate test results of the evaluated 40Ah pouch cells. Within the tested range, the cells showed stable behaviour and maintained temperatures within specification.

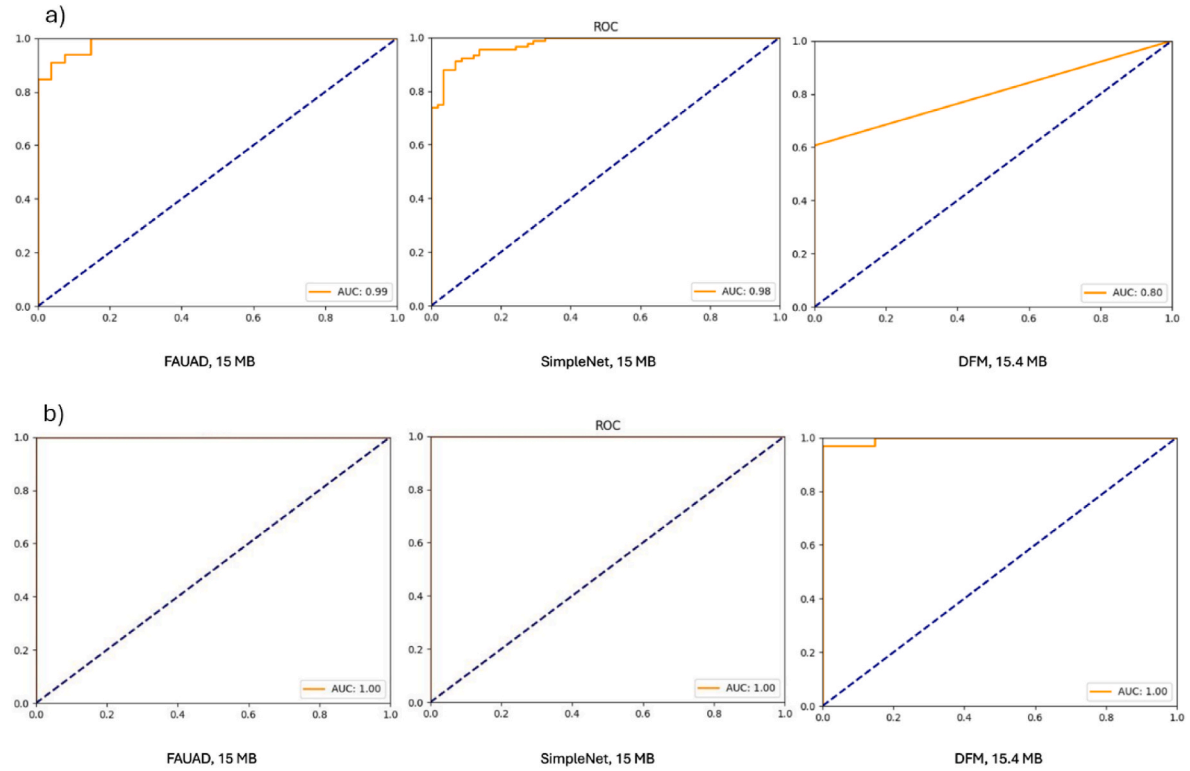

Fig. 9. a) AUROC curves resulting from testing the (<20 MB) models on normal data contaminated with 10 % anomalies, and b) resulting from testing the (<20 MB) models on clean normal data. X-axis denotes False Positive Rate and the Y-axis denotes The Positive Rate.

Table 2
Comparison of FAUAD performance using Poisson and Gaussian noises.

Technique	AUROC (normal data contaminated with 10 % anomalies)	AUROC (clean normal data)
FAUAD (Poisson noise)	**0.99**	**1**
FAUAD (Gaussian noise)	**0.98**	**1**

challenging scenarios. Furthermore, FAUAD achieves an AUROC score of 1.0 on clean data, signifying its flawless detection of anomalies when presented with uncorrupted data. At the same time, it is the smallest of the evaluated models, a modest 15 MB. This combination makes it a very attractive option for edge devices deployment.

We can observe that the performance on the real data (0.99 and 1.0 AUROC) is better than the simulated one (0.971 AUROC). This is probably due to having more challenging (gradient variation) non-homogenous thermal patterns in the simulated scenario, see Fig. 3 d), while having clearly distinguished regions in the real non-homogeneous

Table 3
Comparison of anomaly detection methods performance using the AUROC metric.

Technique	AUROC **(normal data contaminated with 10 % anomalies)**	AUROC **(clean normal data)**	Model head size in MB
FAUAD	**0.990**	**1**	**15**
SimpleNet [22]	0.977	**1**	**15**
DRAEM [39]	0.922	0.991	1100
CFA [40]	0.878	0.942	30.5
STFPM [41]	0.871	0.961	42.7
PaDiM [42]	0.863	0.996	175.1
EfficientAD [43]	0.810	**1**	81
DFM [44]	0.803	0.996	15.4
FastFlow [37]	0.793	**1**	64.7
PatchCore [21]	0.774	0.990	41.9
CFLOW-AD [36]	0.769	0.873	171.4

images acquired, see Fig. 6 e) & f).

While not quite reaching the same heights as FAUAD, SimpleNet still delivers a very good performance. Its AUROC score on contaminated data is 0.977, indicating a significant ability to differentiate anomalies from normal data even when the data has been tampered with. Like FAUAD, SimpleNet achieves a score (AUROC of 1.0) on clean data, showing its accuracy in detecting anomalies in uncompromised datasets. It also shares the same 15 MB model size as FAUAD.

Some techniques, such as EfficientAD (81 MB) and FastFlow (64.7 MB), prioritise perfect detection of normal data, achieving an AUROC score of 1.0 on clean data. This signifies their ability to flawlessly identify normal data points. However, their performance on contaminated data is lower. This suggests that while they excel at identifying normal data, they might struggle to distinguish anomalies effectively when anomalies are deliberately introduced into the data. Furthermore, their model size is larger than FAUAD and SimpleNet, which makes them less attractive for mobility applications that require deployment on edge devices.

4. Conclusions

The need for advanced thermal monitoring of energy storage systems is evident as it underpins both safety and performance aspects. The objective of this study was to propose advancements in this field by building and evaluating an automatic thermal anomaly detection model, using simulated images followed by testing commercially relevant cells. Leveraging recent developments in Deep Learning, a Feature-Adapted Unsupervised Anomaly Detection (FAUAD) model has been constructed. Building blocks of the model have been explained, enabling application of the proposed methodology by a wider audience. A key strength of FAUAD in battery health monitoring lies in its ability to leverage unsupervised learning. This approach is particularly well-suited due to the scarcity of labelled thermal image-based anomaly data.

By focusing on readily available normal data, FAUAD can effectively detect anomalies without the need for extensive and potentially hazardous anomalous data collection. The conducted experiments evaluating a pouch cell charge-discharge scenario demonstrate FAUAD's effectiveness; amongst the evaluated techniques, FAUAD and SimpleNet emerge as the frontrunners, while also the smallest at 15 MB, crucial aspect for embedded systems such as BMS. Both models achieve a AUROC score of 1.0 on clean data, indicating notable ability to identify normal battery behaviour. Furthermore, FAUAD boasts the highest AUROC (0.990) on normal data contaminated with few anomalies, highlighting its proficiency in anomaly detection even under less-than-ideal conditions.

Importantly, the proposed modelling methodology is cell chemistry and application agnostic. While the model was evaluated using Li-ion cells under specific scenario, batteries are currently in constant dynamic development, necessitating cross-chemistry solutions. As the model building blocks rely only on thermal images, such can be obtained from Na-ion, solid-state, etc. in a similar way, requiring little adaption to deploy utilizing the same methodology across various scenarios.

In summary, FAUAD presents a significant advancement in unsupervised anomaly detection for battery thermal mapping. Its ability to leverage available normal data, combined with its performance and efficient design, makes FAUAD a compelling option for ensuring safety and reliability of high-power batteries. Future work could explore incorporating labelled data, if such becomes available, to potentially improve FAUAD's ability to pinpoint the root cause of the detected anomalies.

CRediT authorship contribution statement

Abdelrahman Shabayek: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Arunkumar Rathinam:** Writing – original draft, Formal analysis. **Matthieu Ruthven:** Writing – review & editing, Supervision. **Djamila Aouada:** Writing – review & editing, Supervision. **Tazdin Amietszajew:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Tazdin Amietszajew reports financial support was provided by Horizon Europe. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The research work presented in this article is financially supported by the European Union's Horizon Europe project ENERGETIC (Grant No 101103667).

This work is also supported by the Luxembourg National Research Fund (FNR), under the project reference C21/IS/15965298/ELITE.

Data availability

Data will be made available on request.

References

[1] M. Bini, D. Capsoni, S. Ferrari, E. Quartarone, P. Mustarelli, Rechargeable Lithium Batteries, vol. 4, Elsevier Ltd., 2015, https://doi.org/10.1016/B978-1-78242-090-3.00001-8.
[2] L. Gaines, R. Cuenca, Costs of Lithium-Ion Batteries for Vehicles, Argonne, IL, Aug. 2000, https://doi.org/10.2172/761281.
[3] C. for E. Cooperation, Environmentally Sound Management of End-Of-Life Batteries from Electric-Drive Vehicles, North America, 2015.
[4] C. Roe, et al., Immersion Cooling for Lithium-Ion Batteries – A Review, Elsevier B. V., Mar. 30, 2022, https://doi.org/10.1016/j.jpowsour.2022.231094.
[5] A. Pesaran, Battery thermal management in EVs and HEVs : issues and solutions. Advanced Automotive Battery Conference, 2001, p. 10.
[6] P.V. Braun, J. Cho, J.H. Pikul, W.P. King, H. Zhang, High power rechargeable batteries, Curr. Opin. Solid State Mater. Sci. 16 (4) (2012) 186–198, https://doi.org/10.1016/j.cossms.2012.05.002.
[7] X. Feng, M. Ouyang, X. Liu, L. Lu, Y. Xia, X. He, Thermal runaway mechanism of lithium ion battery for electric vehicles: a review, Energy Storage Mater. 10 (Jan. 2018) 246–267, https://doi.org/10.1016/J.ENSM.2017.05.013.
[8] Q. Wang, P. Ping, X. Zhao, G. Chu, J. Sun, C. Chen, Thermal runaway caused fire and explosion of lithium ion battery, J. Power Sources 208 (Jun. 2012) 210–224, https://doi.org/10.1016/j.jpowsour.2012.02.038.

[9] A. Martins, E. Macchi, F. Di Persio, G. Thenaisie, "BEPA Safety Task Force Position Paper,", 2024.
[10] J. Fleming, T. Amietszajew, J. Charmet, A.J. Roberts, D. Greenwood, R. Bhagat, The design and impact of in-situ and operando thermal sensing for smart energy storage, J. Energy Storage 22 (October 2018) (2019) 36–43, https://doi.org/10.1016/j.est.2019.01.026.
[11] P. Cicconi, D. Landi, M. Germani, Thermal analysis and simulation of a Li-ion battery pack for a lightweight commercial EV, Appl. Energy 192 (2017) 159–177, https://doi.org/10.1016/j.apenergy.2017.02.008.
[12] T. Amietszajew, et al., Hybrid thermo-electrochemical in situ instrumentation for lithium-ion energy storage, Batter Supercaps 2 (11) (2019) 934–940, https://doi.org/10.1002/batt.201900109.
[13] Chester G. Motloch, et al., High-power battery testing procedures and analytical methodologies for HEV's, in: *Future Car Congress*, SAE International, Jun. 2002, https://doi.org/10.4271/2002-01-1950.
[14] B. Wu, V. Yufit, M. Marinescu, G.J. Offer, R.F. Martinez-Botas, N.P. Brandon, Coupled thermal–electrochemical modelling of uneven heat generation in lithium-ion battery packs, J. Power Sources 243 (Dec. 2013) 544–554, https://doi.org/10.1016/J.JPOWSOUR.2013.05.164.
[15] B. Wu, V. Yufit, Y. Merla, R.F. Martinez-Botas, N.P. Brandon, G.J. Offer, Differential thermal voltammetry for tracking of degradation in lithium-ion batteries, J. Power Sources 273 (2015) 495–501, https://doi.org/10.1016/j.jpowsour.2014.09.127.
[16] J. Yeregui, L. Oca, I. Lopetegi, E. Garayalde, M. Aizpurua, U. Iraola, State of charge estimation combining physics-based and artificial intelligence models for Lithium-ion batteries, J. Energy Storage 73 (Dec. 2023) 108883, https://doi.org/10.1016/J.EST.2023.108883.
[17] M.J. Koshkouei, N. Fereshteh Saniee, A. Barai, Thermocouple selection and its influence on temperature monitoring of lithium-ion cells, J. Energy Storage 92 (Jul) (2024), https://doi.org/10.1016/j.est.2024.112072.
[18] C. Lile, L. Yiqun, Anomaly detection in thermal images using deep neural networks, in: 2017 IEEE International Conference on Image Processing (ICIP), IEEE, Sep. 2017, pp. 2299–2303, https://doi.org/10.1109/ICIP.2017.8296692.
[19] X. Li, J. Li, A. Abdollahi, T. Jones, Data-driven thermal anomaly detection for batteries using unsupervised shape clustering, in: 2021 IEEE 30th International Symposium on Industrial Electronics (ISIE), IEEE, Jun. 2021, pp. 1–6, https://doi.org/10.1109/ISIE45552.2021.9576348.
[20] J. Zhang, et al., Realistic fault detection of li-ion battery via dynamical deep learning, Nat. Commun. 14 (1) (Sep. 2023) 5940, https://doi.org/10.1038/s41467-023-41226-5.
[21] K. Roth, L. Pemula, J. Zepeda, B. Schölkopf, T. Brox, P. Gehler, "Towards Total Recall in Industrial Anomaly Detection,", Jun. 2021.
[22] Z. Liu, Y. Zhou, Y. Xu, Z. Wang, "SimpleNet: A Simple Network for Image Anomaly Detection and Localization,", Mar. 2023.
[23] N. Seliya, A. Abdollah Zadeh, T.M. Khoshgoftaar, A literature review on one-class classification and its potential applications in big data, J Big Data 8 (1) (Sep. 2021) 122, https://doi.org/10.1186/s40537-021-00514-x.
[24] J. Cai, J. Fan, "Perturbation Learning Based Anomaly Detection,", Jun. 2022.
[25] S. Akcay, D. Ameln, A. Vaidya, B. Lakshmanan, N. Ahuja, U. Genc, Anomalib: A Deep Learning Library for Anomaly Detection, Feb. 2022.
[26] https://pytorch.org/vision/main/models/generated/torchvision.models.resnet18.html.".
[27] Y. Altmann, D. Yao, S. McLaughlin, M.E. Davies, Robust Linear Regression and Anomaly Detection in the Presence of Poisson Noise Using Expectation-Propagation, 2021, pp. 143–158, https://doi.org/10.1007/978-981-15-9199-0_14.
[28] Zhang et al., "https://d2l.ai/chapter_multilayer-perceptrons/mlp.html.".
[29] https://pytorch.org/docs/stable/index.html.".
[30] X. Tao, X. Gong, X. Zhang, S. Yan, C. Adak, Deep Learning for Unsupervised Anomaly Localization in Industrial Images: A Survey, Jul. 2022, https://doi.org/10.1109/TIM.2022.3196436.
[31] T. Ehret, A. Davy, J.-M. Morel, M. Delbracio, Image Anomalies: a Review and Synthesis of Detection Methods, Aug. 2018, https://doi.org/10.1007/s10851-019-00885-0.
[32] R. Chalapathy, S. Chawla, Deep Learning for Anomaly Detection: A Survey, Jan. 2019.
[33] L. Ruff, et al., A Unifying Review of Deep and Shallow Anomaly Detection, Sep. 2020, https://doi.org/10.1109/JPROC.2021.3052449.
[34] B. Mohammadi, M. Fathy, M. Sabokrou, Image/Video Deep Anomaly Detection, A Survey, Mar. 2021.
[35] G. Pang, C. Shen, L. Cao, A. van den Hengel, Deep learning for anomaly detection, A Review (Jul. 2020), https://doi.org/10.1145/3439950.
[36] D. Gudovskiy, S. Ishizaka, K. Kozuka, CFLOW-AD: Real-Time Unsupervised Anomaly Detection with Localization via Conditional Normalizing Flows, Jul. 2021.
[37] J. Yu, et al., FastFlow: Unsupervised Anomaly Detection and Localization via 2D Normalizing Flows, Nov. 2021.
[38] P. Bergmann, M. Fauser, D. Sattlegger, C. Steger, Uninformed Students: Student-Teacher Anomaly Detection with Discriminative Latent Embeddings, Nov. 2019, https://doi.org/10.1109/CVPR42600.2020.00424.
[39] V. Zavrtanik, M. Kristan, D. Skočaj, Reconstruction by inpainting for visual anomaly detection, Pattern Recogn. 112 (Apr. 2021) 107706, https://doi.org/10.1016/J.PATCOG.2020.107706.
[40] S. Lee, S. Lee, B.C. Song, CFA: coupled-hypersphere-based feature adaptation for target-oriented anomaly localization, IEEE Access 10 (2022) 78446–78454, https://doi.org/10.1109/ACCESS.2022.3193699.
[41] G. Wang, S. Han, E. Ding, D. Huang, Student-Teacher Feature Pyramid Matching for Unsupervised Anomaly Detection, Jun. 2021.
[42] T. Defard, A. Setkov, A. Loesch, R. Audigier, PaDiM: a Patch Distribution Modeling Framework for Anomaly Detection and Localization, Nov. 2020.
[43] K. Batzner, L. Heckler, R. König, EfficientAD: Accurate Visual Anomaly Detection at Millisecond-Level Latencies, Mar. 2023.
[44] N. Ahuja, I. Ndiour, T. Kalyanpur, O. Tickoo, Probabilistic Modeling of Deep Features for Out-Of-Distribution and Adversarial Detection, Jun. 2019.
[45] J. Lin, H.N. Chu, D.A. Howey, C.W. Monroe, Multiscale coupling of surface temperature with solid diffusion in large lithium-ion pouch cells, Commun. Eng. 1 (1) (May 2022) 1, https://doi.org/10.1038/s44172-022-00005-8.
[46] www.comsol.com," COMSOL Multiphysics® v. 6.1.