

THEORY, RESEARCH AND DISCUSSIONS IN ENGINEERING

Editor

Assist. Prof. Dr. Çağrı ARISOY



THEORY, RESEARCH AND DISCUSSIONS IN ENGINEERING

Editor

Assist. Prof. Dr. ÇAĞRI ARISOY*



Theory, Research and Discussions in Engineering

Editor

Assist. Prof. Dr. ÇAĞRI ARISOY

General Director: Berkan Balpetek

Cover and Page Design: Duvar Design

Printing : 30 September-2026

Publisher Certificate No: 49837

E-ISBN: 978-625-5471-08-6

© Duvar Yayınları

853 Sokak No:13 P.10 Kemeraltı-Konak/İzmir

Tel: 0 232 484 88 68

www.duvar yayinlari.com

duvarkitabevi@gmail.com

"The authors bear full responsibility for the sources, opinions, findings, results, tables, figures, images, and all other content presented in the chapters of this book. They are solely accountable for any financial or legal obligations that may arise in connection with national or international copyright regulations. The publisher and editors shall not be held liable under any circumstances"

TABLE OF CONTENTS

Chapter 11
AI-Based Detection and Classification of Power Quality Disturbances
Ahmet NUR

Chapter 222
Unit Disk Automorphisms for MNIST Data Augmentation:
An Empirical Study with a Shallow Neural Network
Bilal DEMİR, Özden KORUOĞLU

Chapter 336
Artificial Intelligence Applications in Battery Management Systems
Faruk KURKER

Chapter 1

AI-Based Detection and Classification of Power Quality Disturbances

Ahmet NUR¹

1. Introduction

Power Quality (PQ) in electrical power systems is a term used to describe how much voltage, current, frequency, and waveform characteristics stay within specified limits and how well these conditions help with the reliable operation of electrical devices for the entire Power System Generation Phase and Transmission Phase, Distribution Phase, and Utilization Phase. Any deviation from nominal conditions (i.e., change in voltage magnitude, frequency, waveform distortion, voltage imbalance, transients, interruptions, and other Electrical phenomena) are generally referred to as disturbances in Power Quality [1-3].

Over the last twenty years, the increasing use of renewable energy sources (predominantly solar and wind) added to the increasing number of power electronic converters, Flexible AC Transmission System (FACTS) devices, and nonlinear loads has increased the number and complexity of Power Quality disturbances by a wide margin [4,5]. The increasing number of disturbances, such as harmonics, have necessitated the creation of a new area of study dealing with the analysis of harmonics using Artificial Intelligence [6].

All disturbances such as voltage sags, voltage swells, harmonic distortion, transients, flicker, interruptions, notches, and various combinations can result in various adverse effects on sensitive electronic/microprocessor-based equipment, the unwanted opening of protective devices and circuit breakers, the shut down of production lines, and consequently major economic losses. For example, an estimate of the annual economic loss due to electrical outages to brokerage firms by the U.S. Department of Energy is approximately \$6.5 million/hour; to credit card companies it is approximately \$2.5 million/hour. In Europe alone, the annual estimated costs of Power Quality issues in industry and commerce are in excess of 10 billion euros [5,6]. The magnitude of impact has caused Automatic, Rapid, and Accurate Detection and Classification of power quality disturbances to become one of the most important research areas in today's Electric Power System Engineering.

¹ Dr. Öğr. Üyesi, Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Bitlis Eren University, Bitlis, 13100, Türkiye
ORCID: 0009-0006-5671-6923

Power Quality Problems classified according to electrical characteristics (amplitude, voltage magnitude, duration) and duration (i.e., microseconds, milliseconds) of the event are shown in Figure 1 below. Disturbances in Power Systems must be assessed in relation to voltage amplitude and the length of time and frequency components, phase relationship, and waveform. Transient events, which represent some of the shortest-duration power quality disturbances, are generally classified into two main categories: impulsive and oscillatory transients. Impulsive transients may occur over time scales ranging from nanoseconds to milliseconds, whereas oscillatory transients are commonly categorized according to their frequency content as low-, medium-, and high-frequency oscillations. Typical causes of transient events include lightning strikes, switching operations, and capacitor energization.

Voltage Changes are defined by two time periods: short-term and long-term. Events occurring for less than one minute are categorized as voltage sags, voltage swells, and interruptions. Events occurring for greater than one minute are categorized as Overs, Unders, and Sustained Interruptions. This categorization of disturbances in Power Quality emphasizes the importance of duration as an equally important aspect of Power Quality in addition to magnitude. Especially with sensitive electronic loads, motor drives, automation systems, and computer hardware/software, short-term voltage sags can result in enormous ramifications such as loss of production, activating of protective systems, or malfunction of equipment.

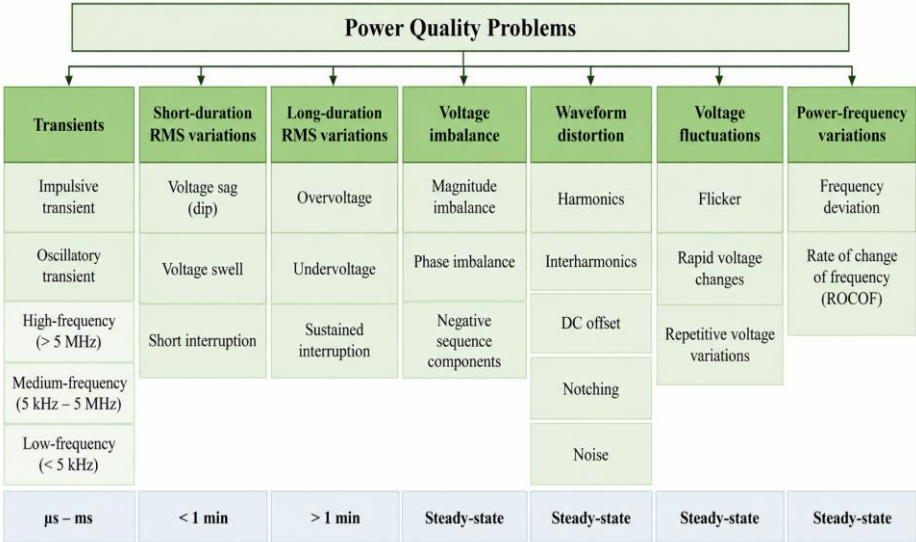


Figure 1. Classification of power quality problems (adapted from [5]).

Figure 1 illustrates that voltage imbalance is a unique power quality concern; it consists of two components: magnitude and phase. When there is a difference in the phase voltage amplitudes or phase angles of a three-phase system, this causes negative sequence components in the system. Negative sequence components can cause electric motor stators to experience increased losses (overheating) and reduced efficiency; hence voltages unbalance is an important parameter under steady-state conditions.

Waveform distortion is another type of power quality concern that includes components that distort the fundamental frequency sinusoidal voltage and current waveforms (harmonics, interharmonics, DC component, noise and notching). With the increased usage of power electronics, converters, rectifiers, variable speed drives and other nonlinear loads, harmonics and interharmonics are becoming increasingly important in power systems.

Harmonics and interharmonics distort the ideal fundamental frequency sinusoidal waveform. Waveform distortion problems include DC offset, notching, etc., which can be more apparent with converter applications and switched mode power supply applications. Voltage fluctuation problems are characterised by rapid and/or repetitive changes in voltage magnitude and can be perceived visually as flicker (primarily with lighting). The primary contributors to voltage fluctuations are loads requiring variable and rapidly changing power. Conversely, variations in power frequency will be evaluated separately from voltage fluctuation and waveform distortion problems.

In general, the findings of figure one demonstrate that the classification of power quality issues is multi-dimensional and that to evaluate power quality issues, the following are considered: effective voltage value, duration of the event, frequency spectrum, waveform, and phase balance. This classification provides a functional basis for identifying the existence of power quality issues, measuring them accurately, and determining the appropriate means of addressing those issues with the greatest fundamental accuracy.

At present, the use of artificial intelligence (AI) methods has become an integral part of power quality monitoring and detection systems because of the advantages associated with automatic feature extraction, nonlinear pattern recognition, and multi-class classification capabilities [7]. This section will cover the range of techniques from feature extraction techniques based on signal processing, to the application of traditional AI methods and reinforcement learning all designed for the purpose of detecting and classifying power quality disturbances [8,9].

2. Signal Processing-Based Feature Extraction for the Detection of Power Quality Disturbances

AI-based classifiers depend heavily on the quality of the features that are used to train them [9]. Due to this dependence on feature extraction techniques for AI classifiers, feature extraction from power quality disturbance signals has traditionally been considered an important preprocessing step. The most commonly employed signal processing techniques for feature extraction in the detection and classification of power quality disturbances are the Fourier Transform (FT) and its many derivative versions such as the Discrete Fourier Transform (DFT), Fast Fourier Transform (FFT), and the Short-Time Fourier Transform (STFT). Other signal processing techniques include the Wavelet Transform (WT), S-Transform (ST), Hilbert-Huang Transform (HHT), Kalman Filter (KF), and Mathematical Morphology (MM) [10].

Although the Fourier Transform can effectively extract the frequency components of stationary signals, its inability to retain temporal information limits its effectiveness for use in detecting sudden start/end points in power quality disturbance signals because of their inherent non-stationary nature [7]. The Short-Time Fourier Transform (STFT) addresses this limitation by dividing a signal into small stationary windows, which allows for time-frequency analysis, but it has the disadvantage of having a fixed window size and therefore represents a compromise between time resolution and frequency resolution.

The Wavelet Transform (WT) uses a short-duration waveform called the mother wavelet to scale and shift the wavelet and, as such, decomposes the signal into multiple levels of resolution while retaining both time and frequency information. Furthermore, when combined with the technique called Multi-Resolution Analysis (MRA), wavelet transforms are for the detection and classification of transient events [11]. Since their initial development, wavelet neural networks (WNNs) have garnered significant attention and research interest, including studies related to the use of wavelet networks for detecting and classifying transient events [12], wavelet neural networks for recognising power quality distortions [13], the development of the theoretical framework for the wavelet neural classifier [14] and practical applications of this neural network architecture [15].

The S-Transform (ST) is an extension of the continuous wavelet transform, which combines the benefits of both the STFT and WT. It has frequency-dependent resolution as well as the ability to retain absolute-reference local phase information. The S-transform has proven to be an effective detection technique, especially in noisy environments. Two early examples of the S-transform being used with neural networks for automatic classification include the S-transform

and neural network-based expert system approach [16] and a hybrid version with a radial basis function network [17]. The Hilbert-Huang Transform (HHT) works by decomposing the signal into intrinsic mode functions (IMFs) through an Empirical Mode Decomposition (EMD) technique, and then applying the Hilbert Transform to obtain the instantaneous amplitude, phase and frequency of the signal, making it effective for analysing nonlinear and nonstationary signals. Researchers have also developed intelligent sensor architectures based on the HHT for detecting, classifying and quantifying disturbances [18].

Another important tool for estimating the amplitude, phase angle, and frequency information of a noisy harmonic signal is the Kalman Filter (KF). The Extended Kalman Filter (EKF) is an example of a KF; it has been applied to classify short-term disturbances via S-Transform analysis. The Mathematical Morphology (MM) technique employs Dilation and Erosion Operators to provide low computational complexity and high detection efficiencies, but due to the absence of clearly defined guidelines within the existing literature on how to select the structuring element, specific applications of the MM technique require tuning to meet their requirements [4]. Therefore, no single technique of signal processing can adequately represent all forms of distortion by itself; thus, there has been a proliferation of Hybrid Techniques within the literature, including: combining Wavelet Transform with Support Vector Machines [19], combining Higher Order Statistics with Genetic Algorithms [20], and designing Real-Time Higher Order Statistics-based Smart Sensors on Field programmable Gate Arrays [21]. Additionally, Meta-Heuristic Optimization techniques have been used for feature selection and dimensionality reduction.

3. Standards

Standards for defining, measuring, and classifying power quality disturbances consistently are critical. Standards are developed by the Institute of Electrical and Electronics Engineers (IEEE), the International Electrotechnical Commission (IEC), and the European Committee for Electrotechnical Standardisation (CENELEC) and are accepted globally. The IEEE 1159 standard classifies power quality disturbances based on their duration and magnitude, while EN 50160 defines the characteristics of voltage supplied by public power grids within Europe. The IEEE, IEC, and EN standards used to monitor, assess, and improve electrical power quality can be found in Table 1.

Table 1. Key standards for power quality monitoring

Organization	Standard	Description
IEEE	519-2022	Guidelines for harmonics control of electrical power quality
IEEE	1159-2019	Guidelines for monitoring of electrical power quality issues
IEEE	1250-2018	Guidelines for Identifying and Improving Voltage Quality in Power Systems
IEEE	1366-2022	Guidelines for electric power utility indices
IEEE	1547-2018	IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces
IEC	61000-2-2	Guidelines for maintaining compatibility with low frequency disturbances in the utility network
IEC	61000-2-4	Guidelines for compatibility level in industrial plants for low frequency disturbances
IEC	61000-4-7	Guidelines for non-linear distortion in the utility network, harmonics/interharmonics measurement, and instrumentation
IEC	61000-4-15	Guidelines for measurement of voltage flicker
IEC	61000-4-30	Guidelines for power quality measurement methods
EN	50160:2022	Guidelines for voltage characteristics of the utility network

By reviewing the standards listed, it becomes clear that power quality is not only addressed concerning harmonic distortion but throughout other areas such as voltage fluctuations, supplying sensitive loads, grounding, flicker, low-frequency disturbances, and measurement accuracy.

3.1. Standards Relevant to AI-Enabled Power Quality Applications

Power quality management, using an AI and ML-based approach, involves compliance with multiple international standards as indicated in Table 2, including Scope, Technical Criteria and Relationship to AI/ML Applications for Power Quality Management.

Table 2. International standards relevant to AI applications in power quality management [5]

Standard	Scope	Criteria	Relevance to AI/ML
IEEE 1547-2018	Interconnection of DERs with Power Systems	Interoperability, voltage regulation, PQ standards	Ensures AI systems managing DERs maintain PQ by adhering to voltage and frequency limits
IEC 61850	Communication Networks in Substations	Interoperability, data exchange	Facilitates data exchange for real-time AI-based PQ monitoring and response
IEEE 519	Harmonics in Power Systems	Harmonic distortion limits	Sets guidelines for THD, which AI/ML algorithms must monitor and manage in high-frequency grids
IEEE/IEC 60255-118-1-2018	Synchrophasor Measurements for Power Systems	Synchrophasor data accuracy, communication requirements	Provides standardized high-resolution synchronized measurement data that can be used as input for AI/ML-based power quality monitoring, disturbance detection and dynamic system analysis
IEEE 2807.3-2022	Guide for Electric-Power-Oriented Knowledge Graph	Accuracy, robustness, reliability	Does not specifically target AI applications in power systems; focuses on knowledge graph frameworks useful for organizing information across sectors, including power systems

Collectively considering all of the standards shown in the table indicate that AI-based systems manage Power Quality more than just simply by using accurate algorithms; AI-based systems utilize a wide variety of conditions to be in compliance including but not limited to the following: quality of data, communication system(s) required to be interconnected, Interoperability, Reliability of Measurement Devices and Boundary Protections in the Electric Power Grid. Another major standard that provides a broad technical framework for AI-based control and decision support systems for all aspects of the integration of Distributed Energy Resources (DER) into the Electric Power Grid is the IEEE 1547-2018 standard. The IEEE 1547-2018 standard establishes

criteria for Interoperability (I.e. Voltages and Frequencies) as well as Technical Competence for Interconnecting DERS.

Real-time AI/ML applications depend on reliable data exchange. Thus, a major focus of IEC 61850 is to support standardized data communications within substations and allow for standardized data communication between electrical systems in power system automation. Another standard that plays a major role in supporting AI/ML applications is IEEE 519 which defines limits on harmonic distortion. By defining limits on harmonic distortion, IEEE 519 provides the algorithms used by AI to identify and classify harmonic distortion and to make decisions on how to improve the power quality of electrical systems. High-resolution synchrophasor measurements standardized by IEEE/IEC 60255-118-1-2018 provide synchronized voltage and current phasor, frequency, and rate-of-change-of-frequency (ROCOF) information. Such measurements can serve as valuable input data for AI/ML-based monitoring, anomaly detection, disturbance identification, and dynamic assessment of power systems. These documents create a more direct relationship between machine learning and power quality. They allow for the evaluation of the performance of the AI/ML models in relation to how effectively they use the data and how reliable the data is. AI/ML models that have been developed to support power quality must be assessed not only by the accuracy of their predictions and classifications, but also by the reliability, robustness, and compatibility of the models with current grid conditions. As a result, the standards identified in the table serve as an important complement to the reliable and practical deployment of AI-enabled power quality monitoring, prediction, classification, and control systems.

3.2. Taxonomy of Artificial Intelligence-Based Classifications

In Figure 2, the four primary categories of artificial intelligence techniques used in identifying and classifying the power quality events are shown. The first category contains all neural networks (ANN), the second category contains all support vector machines (SVM), the third category contains all fuzzy systems, and finally the last category contains classification methods. During the signal processing stage of all of these classification systems, features have been extracted that allow a computer to automatically identify the different types of power quality disturbances including voltage sags, voltage swells, interruptions, harmonic distortion, and transient events, based on the features derived from the processing of the signals.

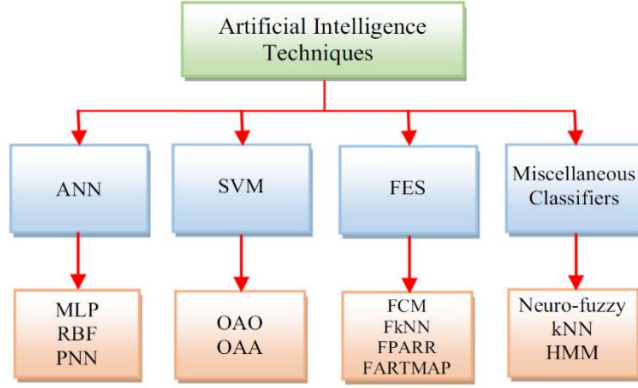


Figure 2. Classical AI classifier taxonomy [7]

Numerous ANNs, e.g., RBF, MLP, PNN and, etc., have been used for power quality classification due to their ability to classify complicated nonlinear events. On the other hand, in the case of SVM classifiers, both the OAA (One Against All) and OAO (One Against One) strategies have been applied when dealing with multiple classes of data in power quality classification. The strength of SVM classifiers is their capacity for clear separation between classes of data, in a multivariate feature vector space.

Fuzzy-based methods, such as FCM, FkNN, FPARR, and FARTMAP, have been used for power quality classification and prove advantageous if the class boundaries are not clear-cut and/or measurement data have high uncertainty (i.e., measurement data contain uncertainty). In addition to fuzzy methods, classifiers such as neuro-fuzzy, kNN, and HMM can also be applied based on the structure of the data being classified and application requirement. Overall, classifier selection in power quality classification is dependent on several factors including data size, feature vector structure, the nature of the relationships among features (i.e., linear vs. nonlinear), computational cost, and whether real-time operation is required. Therefore, classification accuracy will also depend on the AI algorithm used as well as appropriate signal processing and feature extraction methods prior to the application of the AI classifier.

Throughout this taxonomy, it is worth noting that each of the four primary categories can contain multiple variations derived from a core algorithm, e.g., ANN, SVM, FES, kNN/HMM. For instance, PNN (Probabilistic Neural Network), as an ANN, provides a probabilistic model that helps stabilize the classifications produced by PNNs in noisy environments; however, OAA and OAO SVM approaches allow for the solution of multi-class classification

problems through the reduction of multi-class classification problems to two binary classification sub-problems.

Furthermore, the presence of four variations under the Fuzzy Expert System category illustrates how important fuzzy logic was in the classification of power quality disturbance patterns that are uncertain and lack clear-class boundaries. More specifically, during this time, many additional classifiers were independently proposed, including fuzzy expert systems and Fourier linear combiners [22], neuro-fuzzy classifiers [23], rule-based expert systems [24], dynamic time-warping classifiers [25,26], ensemble models driven by multidimensional features [27], a combination of deep learning and curvelet transformation hybrid approaches [28], deep learning detection, automatic classification by deep learning [29,30], and a variety of new methods based on deep learning recognition of power quality disturbances [31].

3.3. Comparison of Artificial Intelligence Techniques

Various strengths and weaknesses of Machine Learning (ML) & AI are compared in Table 3. It is clear from the analysis that it is inappropriate to choose a method based only on how well it classifies data but also on other criteria such as the type of data used, how much computation is needed to run each algorithm, whether or not the algorithm is resistant to noise, whether the results must be produced within a particular amount of time, and how easy a model can be understood. In particular, SVMs present advantages for processing high-dimensional and complex datasets; however, it is difficult to understand what SVMs are doing, so SVM training times on large datasets are long. The ability of an SVM to classify data appropriately depends greatly on the amount of training examples available.

Table 3. Comparison of AI and soft-computing techniques [2]

Technique	Advantages	Disadvantages
SVM (Support Vector Machine)	Good for high dimensional data. It can work with unstructured data.	Understanding the final model is not easy. The classification is more accurate if training samples are more, so with low training samples the system is not accurate. Large data sets require long training time.
ANN (Artificial Neural Networks)	Inherently distributed nature of ANNs. Has parallel processing capability, is fault-tolerant if some part of the network is corrupted. Good for real-time applications and is flexible mathematically.	Low convergence speed, multiple local minima problem and over-fitting problems. Learning process can get stuck in case of multiple local minima. Chosen architecture, algorithm and noise are the major factors for convergence speed and robustness.
FL (Fuzzy Logic)	Uncertainty and ambiguity of the system are managed with the membership degrees assigned to the variables. Can handle incomplete data. Not so expensive.	New disturbances could not be adapted, so each new problem requires adjustment
GA (Genetic Algorithm)	Good for noisy data. Robust to local maxima, minima problems.	Computationally expensive. Difficult to develop.
ES (Expert System)	Large data not required, can work with limited data.	Slow, expensive and assumption power is low.
k-NN (k-Nearest Neighbors)	Classification accuracy is good.	Less accuracy under noisy conditions.

Benefits associated with artificial neural networks (ANNs) include their capacity to model nonlinear relationships, ease of use for real-time products and services, and their capacity to carry out parallel processing tasks. The slow convergence rate, possibility of falling into local minima, and the potential for overfitting all add to the complexity of creating and training ANN models. Furthermore, the architecture of a network, the choice of training algorithm, and the amount of noise present in the training data can directly impact the stability and generalisation capability of the model. Fuzzy logic (FL) provides an alternative methodology to solve problems where the boundaries cannot be definitively established. It offers an avenue to create interpretations of ambiguous or uncertain situations through membership degrees of uncertainty. However, as

there is no capability built-in to FL to react to previously defined situations or to learn, reconfiguration and/or retraining of the system may be necessary when confronted with unpredictable situations.

Genetic algorithms (GAs) have been identified as powerful optimisation tools due to their ability to manage noise in a training dataset and their reduced risk of being trapped by local optimum solution; however, the high cost of computation and the extensive complexity of designing an algorithm are both serious drawbacks. Expert systems (ES) are advantageous due to their ability to operate on a small dataset, as well as direct incorporation of expert knowledge. Nevertheless, their lack of flexibility is due to the amount of time and resources required for their construction and ongoing maintenance.

The k-nearest neighbour (k-NN) method is well known for its simple structure, and generally high accuracy of classification under the appropriate conditions; however, it has a significant limitation in that classification accuracy may degrade when noisy data is present.

In summary, all the evaluation criteria shown in Table 4 demonstrate that there is no one-size-fits-all algorithm for problems that involve multiple variables and potentially many sources of disturbance, such as power quality. SVM is more suitable for high-dimensional data, ANN is more appropriate for addressing complex relationships, fuzzy logic is better for modelling uncertainty, and genetic algorithms are more useful for optimising parameters. Therefore, algorithm selection should be performed by assessing data volume, level of noise, computational time, interpretability and real-time performance requirements collectively. Table 4 highlights the comparative strengths and weaknesses of various artificial intelligence methods against key attributes such as tolerance for uncertainty, ability to learn, ability to explain answers, degree of adaptability, ability to represent knowledge and degree of generalisation performance.

Table 4. Comparison of selected AI techniques based on key performance attributes [1]

Attribute	k-NN	ANN	SVM	FL	GA	ES
Uncertainty tolerance	★★★	★★★	★★★★	★★★	★★★★	★★
Imprecision tolerance	★★★	★★★	★★★★	★★★	★★★★	★
Knowledge representation	★	★★★★★	★	★★★	★	★★
Adaptability	★★★	★★★	★★★	★★	★★★★	★

Attribute	k-NN	ANN	SVM	FL	GA	ES
Maintainability	★★★	★★★	★★★	★★	★★	★
Learning ability	★★★	★★★★	★★★★	★	★★★★	★
Explanation ability	★	★★	★	★★★★	★	★★★★
Data mining	★★★	★★★	★★★	★★	★★	★
Generalized performance	★★★	★★★	★★★★★	★	★	★

Note: The star ratings represent qualitative assessments adapted from [1] and should not be interpreted as quantitative performance scores or absolute rankings. A higher number of stars indicates a comparatively stronger capability with respect to the corresponding attribute as reported in the source study. The ratings are intended only for general comparative interpretation, since actual performance may vary depending on the dataset, disturbance type, model configuration, and application conditions.

From the review of the table evaluations, it appears that no one method excels in all categories, and the selection of a method to use will depend upon the requirements of the specific application. SVM has received high ratings for its tolerance for uncertainty and ambiguity in data, ability to learn, and generalisation performance, resulting in many researchers promoting it as a strong candidate for data-driven classification tasks; however, it has limited structure regarding the representation and explanation of knowledge. ANNs perform exceptionally well when it comes to representing knowledge and learning capabilities; they model complex nonlinear relationships extremely well. However, the low explainability of ANNs is a significant drawback when the ability to interpret the decision-making process is a requirement of a given application. While the overall structure of fuzzy logic is reasonably balanced between the power of explanation and how knowledge is represented, fuzzy logic is less efficient at learning and generalizing than the other data-driven methods described in this chapter. Genetic Algorithms are adaptable and optimized, but General Algorithm, like fuzzy logic, does not have direct explainability or generalization capabilities. Expert Systems offer the advantage of interpretability but, as with Genetic Algorithms, do not perform well when developing adaptations and learning or perform poorly at data mining.

4. Comparison of Techniques Used for Power Quality Issues

A comparison of the frequency with which certain power quality issues have been addressed in prior research when evaluating the approaches taken by

artificial intelligence and machine learning on those same problems is indicated in Figure 3.

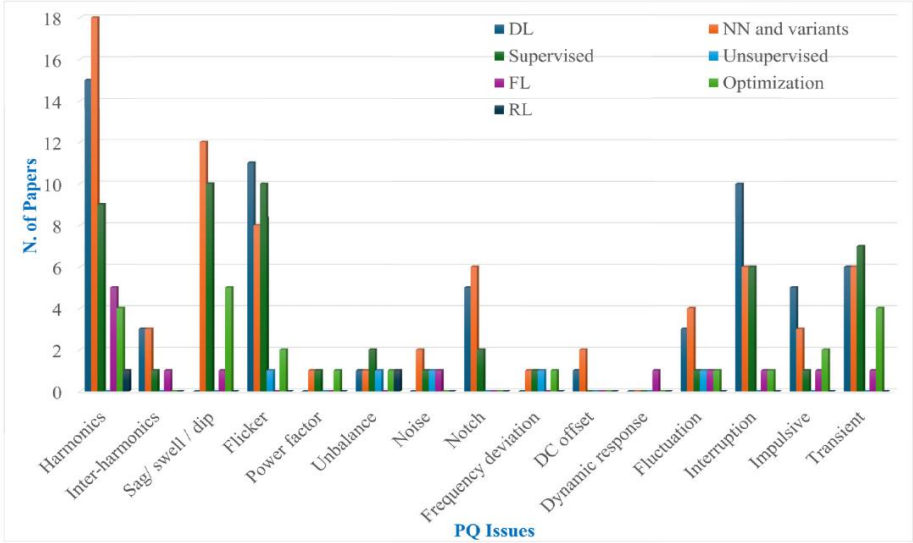


Figure 3. Comparison of the considered PQ issues and techniques used in the literature [5]

The general distribution shows that researchers have concentrated their efforts primarily on a smaller set of the identified power quality issues, namely: harmonics, voltage sags/swells, flickering, interruption, and transient events. Harmonical problems were also identified as the type of power quality issue for which neural network and related applications were most heavily utilised in almost 18 studies, while deep learning methods were also found to be the most commonly used in about 15 studies. For voltage fluctuations, researchers found neural network and supervised learning methods to be the most routinely used, whereas researchers preferred the use of deep learning and supervised learning methods for flicker evaluation.

The figure demonstrates conclusively the existence of strong methodological trends within the area of power quality disturbance research. In particular, NN-based power quality disturbance research predominates with respect to classic and common distortion types (e.g., harmonic distortion and voltage sag/swells), whereas DL methods dominate with respect to sudden and erratic occurrences (e.g., interruptions and impulsive distortions). This finding is influenced by the technical advantages of the ability of DL architectures (including CNNs and LSTMs) to learn and characterize transitions from raw time series data when

compared with hand-crafted features for sudden and irregular transitions. The superiority of this feature automatically learned via DL architectures is further demonstrated through the existence of research efforts that combine two DL architectures: for example, a hybrid of a bidirectional LSTM and CNN [32], a hybrid of a GRU and CNN [33], a method of combining continuous wavelet transforms with CNN [34], a hybrid CNN (LSTM) with fast S-transform [35], and the Lissajous-orbit-based DenseNet architecture [36]. Additionally, 2D DL CNNs are specifically designed to classify disturbances occurring in microgrids [37]; and the use of PCA with CNNs to classify disturbances occurring in wind-grid distribution systems [38] exemplifies the use of DL in practical applications to classify disturbances specific to renewable energy integration. Other examples of important contributions made by the research community include the end-to-end DL approach for voltage sag classification [39], the Bundle Extreme Learning Machine (BELM) based method for power quality analysis in transmission networks [40], and the combination of an Identity Feature Vector with an Extreme Learning Machine (ELM) for classification [41]. Furthermore, the development of a machine-learning based framework for the analysis of voltage quality disturbances [42] and real-time detection and classification systems [43] further demonstrates the current focus on developing practical solutions for real-time applications.

5. Conclusions

This chapter evaluates the use of various signal processing and artificial intelligence methods for the detection and classification of electric power-quality disturbances. The rise of renewable energy sources, power electronics converters, and non-linear loads have increased the complexity of disturbance events (e.g., voltage sags, swells, harmonics, interruptions/transients, flicker, composite disturbances). Thus, there is a demand for rapid and automatic adaptive analysis systems that supersede the traditional methods of using fixed thresholds based on expert evaluation. The Fourier, Short-Time Fourier Transform, Wavelet Transform, the S-Transform, and the Hilbert–Huang Transform, along with Kalman Filter methods, Mathematical Morphology, and more, are signal processing techniques that can be applied to obtaining distinguishing features from power quality signals.

Methods of classification of power quality disturbances include artificial neural networks (ANN), support vector machine (SVM), fuzzy logic, genetic algorithms, expert systems, and k-nearest neighbors (KNN); all have their strengths and weaknesses. For example, SVM is well suited for classifying high dimensional and sparse datasets, ANN can model non-linear relationships,

whereas fuzzy logic can process uncertain and incomplete data. On the other hand, the problem of overfitting in artificial neural networks (ANNs) and SVMs, the limited interpretability of artificial intelligence, and difficulties in adapting to new or previously unseen types of distortions are typical limitations of current methods. Selection of a method should take into account issues such as the size of the data set, the presence of noise, the characteristics of the distortion, computation costs, and real-time practical requirements.

The utilization of CNNs, LSTMs, GRUs, and their hybrid architectures is particularly advantageous for events involving fleeting, complex, and composite power quality disturbances; these architectures are capable of automatically learning attributes directly from time-series data. The following hybrid signal processing and Artificial Intelligence (AI) architectures demonstrate classification performance: CWT–CNN, CNN–LSTM, CNN–GRU, and image-based transfer learning. Nonetheless, these architectures struggle due to their requirement of extremely large and balanced data sets, high computational costs, and limited interpretability in real-time.

In summary, future a power-quality monitoring systems are predicted to rely on hybrid architectures that meet prescribed standards, are explainable, noise resistant, and able to operate in a real-time practical environment. Areas of proposed research include the application of unsupervised and reinforcement learning for less studied disturbance types; the development of publicly available data sets that contain actual grid data; and the application of lightweight deep learning models designed for a wide range of embedded systems. The integration of signal processing and AI will significantly enhance the ability to detect and monitor issues pertaining to power quality and to improve the reliability of electrical grid operations.

References

- [1] Mahela, O.P., Shaik, A.G., Gupta, N. (2015). A critical review of detection and classification of power quality events. *Renewable and Sustainable Energy Reviews*, 41, 495–505. <https://doi.org/10.1016/j.rser.2014.08.070>
- [2] Khetarpal, P., Tripathi, M.M. (2020). A critical and comprehensive review on power quality disturbance detection and classification. *Sustainable Computing: Informatics and Systems*, 28, 100417. <https://doi.org/10.1016/j.suscom.2020.100417>
- [3] Gargoom, A.M., Ertugrul, N., Soong, W.L. (2008). Automatic classification and characterization of power quality events. *IEEE Transactions on Power Delivery*, 23(4), 2417–2425. <https://doi.org/10.1109/TPWRD.2008.923998>
- [4] Chawda, G.S., Shaik, A.G., Shaik, M., Padmanaban, S., Holm-Nielsen, J.B., Mahela, O.P., Kaliannan, P. (2020). Comprehensive review on detection and classification of power quality disturbances in utility grid with renewable energy penetration. *IEEE Access*, 8, 146807–146830. <https://doi.org/10.1109/ACCESS.2020.3014732>
- [5] Nabian Dehaghani, M., Korotko, T., Rosin, A. (2025). AI applications for power quality issues in distribution systems: A systematic review. *IEEE Access*, 13, 18346–18365. <https://doi.org/10.1109/ACCESS.2025.3533702>
- [6] Eslami, A., Negnevitsky, M., Franklin, E., Lyden, S. (2022). Review of AI applications in harmonic analysis in power systems. *Renewable and Sustainable Energy Reviews*, 154, 111897. <https://doi.org/10.1016/j.rser.2021.111897>
- [7] Khokhar, S., Mohd Zin, A.A.B., Mokhtar, A.S.B., Pesaran, M. (2015). A comprehensive overview on signal processing and artificial intelligence techniques applications in classification of power quality disturbances. *Renewable and Sustainable Energy Reviews*, 51, 1650–1663. <https://doi.org/10.1016/j.rser.2015.07.068>
- [8] Wang, S., Chen, H. (2019). A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network. *Applied Energy*, 235, 1126–1140. <https://doi.org/10.1016/j.apenergy.2018.09.160>
- [9] Kheta, K., Giannetti, C. (2022). An image-based deep transfer learning approach to classify power quality disturbances. *Electric Power Systems Research*, 213, 108795. <https://doi.org/10.1016/j.eprsr.2022.108795>
- [10] Oubrahim, Z., Amirat, Y., Benbouzid, M., Ouassaid, M. (2023). Power quality disturbances characterization using signal processing and pattern

- recognition techniques. *Energies*, 16(6), 2685. <https://doi.org/10.3390/en16062685>
- [11] Bollen, M.H.J., Gu, I.Y.H., Axelberg, P.G.V., Styvaktakis, E. (2007). Classification of underlying causes of power quality disturbances: Deterministic versus statistical methods. *EURASIP Journal on Advances in Signal Processing*, 2007, 079747. <https://doi.org/10.1155/2007/79747>
- [12] Angrisani, L., Daponte, P., D'Apuzzo, M. (2001). Wavelet network-based detection and classification of transients. *IEEE Transactions on Instrumentation and Measurement*, 50(5), 1425–1435. <https://doi.org/10.1109/19.963220>
- [13] Gaing, Z.-L. (2004). Wavelet-based neural network for power disturbance recognition and classification. *IEEE Transactions on Power Delivery*, 19(4), 1560–1568. <https://doi.org/10.1109/TPWRD.2004.835281>
- [14] Santoso, S., Powers, E.J., Grady, W.M., Parsons, A.C. (2000). Power quality disturbance waveform recognition using wavelet-based neural classifier—Part 1: Theoretical foundation. *IEEE Transactions on Power Delivery*, 15(1), 222–228. <https://doi.org/10.1109/61.847255>
- [15] Santoso, S., Powers, E.J., Grady, W.M., Parsons, A.C. (2000). Power quality disturbance waveform recognition using wavelet-based neural classifier—Part 2: Application. *IEEE Transactions on Power Delivery*, 15(1), 229–235. <https://doi.org/10.1109/61.847256>
- [16] Uyar, M., Yildirim, S., Gencoglu, M.T. (2009). An expert system based on S-transform and neural network for automatic classification of power quality disturbances. *Expert Systems with Applications*, 36(3), 5962–5975. <https://doi.org/10.1016/j.eswa.2008.07.030>
- [17] Jayasree, T., Devaraj, D., Sukanesh, R. (2009). Power quality disturbance classification using S-transform and radial basis network. *Applied Artificial Intelligence*, 23, 680–694. <https://doi.org/10.1080/08839510903205563>
- [18] Granados-Lieberman, D., Valtierra-Rodriguez, M., Morales-Hernandez, L.A., Romero-Troncoso, R.J., Osornio-Rios, R.A. (2013). A Hilbert transform-based smart sensor for detection, classification, and quantification of power quality disturbances. *Sensors*, 13(5), 5507–5527. <https://doi.org/10.3390/s130505507>
- [19] Erişti, H., Demir, Y. (2010). A new algorithm for automatic classification of power quality events based on wavelet transform and SVM. *Expert Systems with Applications*, 37(6), 4094–4102. <https://doi.org/10.1016/j.eswa.2009.11.015>

- [20] Romero-Ramirez, L.A., Elvira-Ortiz, D.A., Jaen-Cuellar, A.Y., Morinigo-Sotelo, D., Osornio-Rios, R.A., Romero-Troncoso, R.D.J. (2020). Methodology based on higher-order statistics and genetic algorithms for the classification of power quality disturbances. *IET Generation, Transmission & Distribution*, 14(20), 4580–4592. <https://doi.org/10.1049/IET-GTD.2020.0366>
- [21] Martinez-Figueroa, G.D.J., Morinigo-Sotelo, D., Zorita-Lamadrid, A.L., Morales-Velazquez, L., Romero-Troncoso, R.D.J. (2017). FPGA-based smart sensor for detection and classification of power quality disturbances using higher order statistics. *IEEE Access*, 5, 14259–14274. <https://doi.org/10.1109/ACCESS.2017.2732726>
- [22] Dash, P.K., Mishra, S., Salama, M.M.A., Liew, A.C. (2000). Classification of power system disturbances using a fuzzy expert system and a Fourier Linear Combiner. *IEEE Transactions on Power Delivery*, 15(2), 472–477. <https://doi.org/10.1109/61.852971>
- [23] Huang, J., Negnevitsky, M., Nguyen, D.T. (2002). A neural-fuzzy classifier for recognition of power quality disturbances. *IEEE Transactions on Power Delivery*, 17(2), 609–616. <https://doi.org/10.1109/61.997947>
- [24] Styvaktakis, E., Bollen, M.H.J., Gu, I.Y.H. (2002). Expert system for classification and analysis of power system events. *IEEE Transactions on Power Delivery*, 17(2), 423–428. <https://doi.org/10.1109/61.997911>
- [25] Youssef, A.M., Abdel-Galil, T.K., El-Saadany, E.F., Salama, M.M.A. (2004). Disturbance classification utilizing dynamic time warping classifier. *IEEE Transactions on Power Delivery*, 19(1), 272–278. <https://doi.org/10.1109/TPWRD.2003.820178>
- [26] Chung, J., Powers, E.J., Grady, W.M., Bhatt, S.C. (2002). Power disturbance classifier using a rule-based method and wavelet packet-based hidden Markov model. *IEEE Transactions on Power Delivery*, 17(1), 233–241. <https://doi.org/10.1109/61.974212>
- [27] Liu, Y., Yuan, D., Fan, H., Jin, T., Mohamed, M.A. (2023). A multidimensional feature-driven ensemble model for accurate classification of complex power quality disturbance. *IEEE Transactions on Instrumentation and Measurement*, 72, Art no. 1501613. <https://doi.org/10.1109/TIM.2023.3265756>
- [28] Liu, H., Hussain, F., Shen, Y., Arif, S., Nazir, A., Abubakar, M. (2018). Complex power quality disturbances classification via curvelet transform and deep learning. *Electric Power Systems Research*, 163, 1–9. <https://doi.org/10.1016/j.epsr.2018.05.018>

- [29] Sekar, K., Kanagarathinam, K., Subramanian, S., Venugopal, E., Udayakumar, C., Samikannu, R. (2022). An improved power quality disturbance detection using deep learning approach. *Mathematical Problems in Engineering*, 2022, 7020979. <https://doi.org/10.1155/2022/7020979>
- [30] Albalooshi, F.A., Qader, M.R. (2025). Deep learning algorithm for automatic classification of power quality disturbances. *Applied Sciences*, 15(3), 1442. <https://doi.org/10.3390/app15031442>
- [31] Gonzalez-Abreu, A.-D., Delgado-Prieto, M., Osornio-Rios, R.-A., Saucedo-Dorantes, J.-J., Romero-Troncoso, R.-de-J. (2021). A novel deep learning-based diagnosis method applied to power quality disturbances. *Energies*, 14(10), 2839. <https://doi.org/10.3390/en14102839>
- [32] Kandasamy, P., Kumar, C., Lakshmanan, M., Jaisiva, S., Alexander Stonier, A., Peter, G., Ganji, V. (2023). Initial condition based real time classification of power quality disturbance using deep convolution neural network with bidirectional long short-term memory. *IET Generation, Transmission & Distribution*, 17(23), 5135–5154. <https://doi.org/10.1049/gtd2.13026>
- [33] Yiğit, E., Özkaya, U., Öztürk, Ş., Singh, D., Gritli, H. (2021). Automatic detection of power quality disturbance using convolutional neural network structure with gated recurrent unit. *Mobile Information Systems*, 2021, 7917500. <https://doi.org/10.1155/2021/7917500>
- [34] Perez-Anaya, E., Jaen-Cuellar, A.Y., Elvira-Ortiz, D.A., Romero-Troncoso, R.d.J., Saucedo-Dorantes, J.J. (2024). Methodology for the detection and classification of power quality disturbances using CWT and CNN. *Energies*, 17(4), 852. <https://doi.org/10.3390/en17040852>
- [35] Bai, H., Yao, R., Zhang, W., Zhong, Z., Zou, H. (2025). Power quality disturbance classification strategy based on fast S-transform and an improved CNN-LSTM hybrid model. *Processes*, 13(3), 743. <https://doi.org/10.3390/pr13030743>
- [36] Zhang, X., Zheng, J., Mei, F., Miao, H. (2025). Classification of complex power quality disturbances based on Lissajous trajectory and lightweight DenseNet. *Applied Sciences*, 15(14), 8021. <https://doi.org/10.3390/app15148021>
- [37] Chen, C.-I., Berutu, S.S., Chen, Y.-C., Yang, H.-C., Chen, C.-H. (2022). Regulated two-dimensional deep convolutional neural network-based power quality classifier for microgrid. *Energies*, 15(7), 2532. <https://doi.org/10.3390/en15072532>

- [38] Shen, Y., Abubakar, M., Liu, H., Hussain, F. (2019). Power quality disturbance monitoring and classification based on improved PCA and CNN for wind-grid distribution systems. *Energies*, 12(7), 1280. <https://doi.org/10.3390/en12071280>
- [39] Turović, R., Dragan, D., Gojić, G., Petrović, V.B., Gajić, D.B., Stanisavljević, A.M., Katić, V.A. (2022). An end-to-end deep learning method for voltage sag classification. *Energies*, 15(8), 2898. <https://doi.org/10.3390/en15082898>
- [40] Ucar, F., Cordova, J., Alcin, O.F., Dandil, B., Ata, F., Arghandeh, R. (2019). Bundle extreme learning machine for power quality analysis in transmission networks. *Energies*, 12(8), 1449. <https://doi.org/10.3390/en12081449>
- [41] Wei, S., Wenjuan, D., Xia, C. (2024). Power quality disturbances categorization using identity feature vector and extreme learning machine. *Intelligent Systems with Applications*, 24, 200446. <https://doi.org/10.1016/j.iswa.2024.200446>
- [42] Bagheri, A., de Oliveira, R.A., Bollen, M.H.J., Gu, I.Y.H. (2022). A framework based on machine learning for analytics of voltage quality disturbances. *Energies*, 15(4), 1283. <https://doi.org/10.3390/en15041283>
- [43] Mozaffari, M., Doshi, K., Yilmaz, Y. (2022). Real-time detection and classification of power quality disturbances. *Sensors*, 22(20), 7958. <https://doi.org/10.3390/s22207958>

Unit Disk Automorphisms for MNIST Data Augmentation: An Empirical Study with a Shallow Neural Network

Bilal DEMIR¹, Özden KORUOĞLU²

Abstract

This chapter examines unit disk automorphisms as a geometric data augmentation procedure for MNIST classification with a shallow multilayer perceptron. Building on previous Möbius and conformal augmentation studies, we describe a discrete implementation using inverse mapping, bilinear interpolation, and zero filling outside the source image. Disk preservation does not imply preservation of the square image boundary or lossless resampling. We compare no augmentation, affine augmentation, and Möbius augmentation using the same 50,000 training images, 10,000 validation images, architecture, and 15-epoch training budget across three random seeds. Models selected by minimum validation loss are evaluated on the separate 10,000-image MNIST test set. Mean test accuracies are 97.663%, 98.120%, and 98.200%, respectively. Both augmentation procedures improve performance relative to the unaugmented control. Möbius augmentation provides a small average advantage over the selected affine baseline, but its test accuracy is higher in only two of the three runs. Both augmented conditions continue to improve during later epochs. These findings support a limited empirical use of disk automorphisms for augmentation; they do not establish general superiority, complete prevention of overfitting, or learning of topological invariants.

Keywords: Data Augmentation, Möbius transformations, Deep Learning, multilayer perceptron, learning curves.

1. Introduction

Data augmentation introduces controlled variations into training examples to improve prediction on unseen data. In image classification, the usefulness of a transformation depends on its compatibility with the task and the preservation of the intended class label. A geometric transformation can be mathematically well

¹ Assoc. Prof. Dr. Balıkesir University Necatibey Faculty of Education, Department of Mathematics, Balıkesir, Türkiye. ORCID: 0000-0002-6638-6909, E-Mail: bdemir@balikesir.edu.tr

² Prof. Dr. Balıkesir University Necatibey Faculty of Education, Department of Mathematics, Balıkesir, Türkiye. ORCID: 0000-0002-9622-4847, E-Mail: ozdenk@balikesir.edu.tr

defined while still producing an ambiguous digit or moving relevant content outside the finite image window. Consequently, geometric properties and predictive performance require separate evaluation.

Affine operations such as translation, rotation, and scaling are common baselines, but image augmentation is not restricted to these operations. While computationally efficient, their effectiveness depends on the data, model, and parameter choices; affine augmentation does not inherently imply early learning stagnation.

This study examines the effect of Möbius-based geometric data augmentation on handwritten digit classification using the MNIST dataset and a shallow multilayer perceptron (MLP). Building on previous studies of Möbius and conformal image augmentation [11,13], we investigate how these transformations affect classification performance and learning curves in this setting. The study is designed to compare three training conditions: no augmentation, affine augmentation, and Möbius-based augmentation. The main research question is whether Möbius-based augmentation improves validation loss and test accuracy relative to these baselines. We also examine the learning curves over the selected training period. The contribution is a comparative empirical evaluation, with conclusions limited to the dataset, architecture, and experimental conditions considered.

Section 2 reviews related augmentation methods. Section 3 presents the mathematical background of unit disk automorphisms. Section 4 describes image sampling and the experimental protocol. Section 5 reports the results, and Section 6 discusses their interpretation and limitations.

2. Literature Review

In deep learning literature, data augmentation is a very common trick to stop models from overfitting. A broad survey by Shorten and Khoshgoftaar [1] survey augmentation techniques spanning geometric, image-level, feature-space, and generative methods. Mixup [2] forms convex combinations of input examples and their labels, whereas Manifold Mixup [3] interpolates hidden representations and labels. Cutout [4] masks regions of an input image and is therefore an image-level method, not a feature-space transformation. These approaches should not be treated as interchangeable forms of spatial warping or generative latent-space models like Variational Autoencoders [5]. Similarly, Sawhney et al.[6] applied hyperbolic interpolations inside the hidden representations. While these approaches help generalization, they manipulate abstract state vectors rather than the actual 2D image grid, meaning they do not provide direct geometric control over the input morphology. More recently, researchers started to explore non-

linear mappings and Möbius transformations directly on spatial data. However, most of these works focus on different geometries or specialized tasks. For example, some studies use Möbius math for spherical CNNs [7], modern architectures like pure Transformers also heavily rely on robust data augmentation for complex tasks like medical segmentation [8], 360-degree panoramic zoom problems [9], or Diffusion models [10]. When non-linear Möbius transformations are applied directly to standard 2D image datasets, they usually face a practical problem with the outer boundary. For instance, Zhou et al. [11] and Chhipa et al. [12] showed that non-linear mappings can improve model robustness against perspective distortions. However, their transformations do not fix the image frame. As a result, the warped images get curved edges, black empty backgrounds, or require artificial padding, which causes data loss around the corners. Zhou et al. [11] investigated Möbius transformations for image data augmentation. Rainio et al. [13] developed a related approach that maps square images conformally onto a disk, applies disk-preserving transformations, and maps the results back to the square. These studies establish prior work on conformal image augmentation.

The present chapter builds on these studies. It does not introduce Möbius transformations as a new augmentation method or establish that the selected MNIST experiment is the first of its kind. Its scope is the implementation and evaluation of a specified procedure with a shallow MLP. In particular, scaling a square into a disk is different from mapping the square onto the full disk and back as in Rainio et al. [13]. The simpler embedding used here does not preserve the square boundary.

3. Mathematical Preliminaries

To establish a mathematically rigorous foundation for our proposed data augmentation pipeline, we briefly review the theory of fractional linear transformations and the specific subgroup used to warp discrete image spaces.

3.1 Conformal Mappings and the Möbius Group

In complex analysis, a mapping is called conformal if it preserves local angles and orientations between intersecting curves [14]. The most fundamental class of biholomorphic automorphisms on the Riemann sphere $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ is the family of Möbius transformations, also known as fractional linear transformations [15]. A general Möbius mapping $f: \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$ is defined by the algebraic form:

$$f(z) = \frac{az + b}{cz + d}$$

where the complex coefficients satisfy the determinant condition $ad - bc \neq 0$ [14]. From a geometric perspective, these transformations are uniquely determined by their action on circles and straight lines. Under any Möbius mapping, generalized circles (where straight lines are treated as circles of infinite radius) are mapped to generalized circles [15]. These facts concern continuous geometry and do not imply lossless image interpolation

3.2 Automorphisms of the Unit Disk

While general Möbius transformations provide extensive degrees of geometric freedom, applying an arbitrary fractional linear mapping to an image domain often shifts central pixels toward infinity or distorts the exterior boundaries. We restrict our attention to conformal automorphisms of the open unit disk. These mappings preserve the disk as a set, but this property does not imply preservation of the boundary of a square image embedded within the disk. By standard Schwarz-Pick theory, every biholomorphic automorphism $f: D \rightarrow D$ that maps the unit disk bijectively onto itself can be parameterized as:

$$f(z) = e^{i\theta} \frac{z - \alpha}{1 - \bar{\alpha}z}$$

where $\alpha \in D$ is a complex constant representing the interior point that maps to the origin $f(\alpha) = 0$, $\bar{\alpha}$ denotes its complex conjugate, and $\theta \in [0, 2\pi)$ is a real angle representing a pure rotation around the origin [14], [16]. In the implementation considered here, we set ($\theta = 0$), so that the transformation takes the form:

$$f_\alpha(z) = \frac{z - \alpha}{1 - \bar{\alpha}z}, \quad |\alpha| < 1$$

3.3 Boundary Behaviour and Discrete Images

The mathematical power of the unit disk automorphism for data augmentation lies in its boundary behavior. For any point on the boundary circle where $|z| = 1$, the modulus of the transformed point evaluates to:

$$|f_\alpha(z)| = 1$$

This identity shows that the unit circle is preserved as a set; individual boundary points are generally displaced. Within the disk, the transformation produces smooth nonlinear geometric deformations and preserves local angles. These properties apply to the continuous mapping. They do not guarantee

preservation of a square image frame, lossless pixel resampling, or preservation of the semantic identity of a digit after discretization and cropping.

The continuous transformation is a homeomorphism of the disk and therefore does not change its topology. We use the terms geometric deformation and conformal warping for the image operation. The experiments do not measure whether a network learns topological invariants

4. Methodology

In Section 3, we established the mathematical properties of Unit Disk Automorphisms from a pure complex analysis perspective. We now explain how we translate these continuous Möbius operators into a computational pipeline for discrete image tensors and describe our experimental setup.

4.1 Discrete-to-Continuous Domain Mapping

A standard digital image is stored as a discrete grid of pixel intensities of size $H \times W$. Let (u, v) denote the row and column indices of a pixel, where $u \in \{0, 1, \dots, H-1\}$ and $v \in \{0, 1, \dots, W-1\}$. Because standard image coordinates start from the top-left corner, we first center the image at the origin and normalize the spatial coordinates to the range $[-1, 1]$:

$$x = \frac{2v - (W - 1)}{W - 1}, \quad y = \frac{(H - 1) - 2u}{H - 1}.$$

We represent each normalized spatial coordinate as $(q = x + iy)$, where $(x, y \in [-1, 1])$, and embed the square in the unit disk by setting $(z = rq)$. Since the maximum modulus over the normalized square is $(\sqrt{2})$, the condition $(0 < r < 1/\sqrt{2})$ places the entire square strictly inside the disk. We use $(r = 0.70)$. This embedding condition does not guarantee that the transformed image content remains within the square output window.

4.2 Inverse Warping and Bilinear Interpolation

We use inverse spatial mapping to assign a source sampling location to each output pixel. This avoids the unassigned output locations that can arise from forward mapping with coordinate rounding. For an output pixel with normalized coordinate $(q = x + iy)$, we first set $(w = rq)$. With $(\theta = 0)$, the inverse transformation is

$$z = f_{\alpha}^{-1}(w) = \frac{w + \alpha}{1 + \bar{\alpha}w}$$

The corresponding normalized source coordinate is ($s = z/r$). Its source row and column coordinates are

$$u_{\text{src}} = \frac{H-1}{2}(1 - \text{Im}(s)),$$

$$v_{\text{src}} = \frac{W-1}{2}(\text{Re}(s) + 1)$$

Pixel intensities are converted to float32 and divided by 255 before sampling. Bilinear interpolation is used at source locations inside the image. Source coordinates outside the square receive zero intensity. In the implementation, coordinates within 10^{-10} pixels of a boundary are snapped to that boundary to remove numerical round-off errors; genuinely out-of-range coordinates are not clipped to an edge. Inverse sampling avoids unassigned output locations caused by rounded forward mapping, but it does not eliminate cropping or interpolation-related information loss.

4.3 Data and Preliminary Parameter Inspection

MNIST contains ten digit classes, 0 – 9, represented by 28×28 grayscale images [17]. We used the standard 60,000-image training partition to construct a 50,000-image training set and a 10,000-image validation set. The first three examples of each class were used for an initial visual inspection and were explicitly retained in the training set. The remaining images were divided with stratification by class and split seed 42. The separate standard 10,000-image test partition was reserved for the final evaluation. Augmentation was applied only to training images.

An initial comparison on the 30 training examples showed visible clipping at $\alpha = 0.4 + 0.4i$. The smaller magnitude $|\alpha| = |0.05 + 0.05i|$ was then inspected in eight directions on a fixed random sample of 1,000 training images, with 100 images per class and sampling seed 2026. This was a geometric inspection rather than a performance-based hyperparameter search.

For this inspection, a foreground pixel centre was defined by normalized intensity greater than 0.20. We counted how many such centres were mapped outside the output square by the forward transformation. Across the eight directions, 24 of the 1,000 images had at least one foreground centre outside the output square in at least one direction. The largest per-image fraction was 3.614%, and the largest direction-specific mean fraction was 0.038%. These fractions are geometric diagnostics, not percentages of total information loss or

estimates of classification error. The nonzero counts also rule out a lossless-transformation claim.

We compared three conditions: *A*, no augmentation; *B*, affine augmentation; and *C*, Möbius augmentation. In *B* and *C*, each source image was independently selected for transformation with probability 0.5 at each epoch. The same selection mask was used for the two augmented conditions within each seed and epoch. Every source image appeared once per epoch, and a selected transformation was always applied to the original image rather than to an already transformed version.

In condition *B*, the angle was sampled uniformly from -10° to 10° . Horizontal and vertical translations were sampled independently and uniformly from -2 to 2 pixels. Rotation about the image centre was followed by translation, with a single inverse bilinear resampling. Positive horizontal and vertical translations refer to rightward and upward Cartesian directions, respectively.

In condition *C*, a direction index was sampled uniformly from the eight possible values for each selected image:

$$\alpha_k = \rho e^{ik\pi/4}, \quad k \in 0, 1, \dots, 7$$

$$\rho = |0.05 + 0.05i| = \sqrt{0.005} \approx 0.07071.$$

The direction of α varies, while the separate rotation parameter θ remains zero. Both augmentation procedures use bilinear interpolation and zero filling outside the source square. Their deformation magnitudes were not matched; the comparison concerns the selected protocols rather than an equal-strength comparison of transformation families.

All conditions used an MLP with a flattening layer, one dense hidden layer containing 128 ReLU units, and a 10-unit softmax output. Training used sparse categorical cross-entropy and Adam with learning rate 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 10^{-7}$. Batch size was 128, and every run lasted 15 epochs without early stopping. No dropout or weight decay was used. This shallow architecture defines the scope of the experiment; it does not isolate a general mechanism of memorization in deep networks.

The runs used seeds 42, 123, and 2026. Within each seed, all three conditions began with identical weights and used the same source-image order and batch boundaries. Random streams for ordering, augmentation selection, and transformation parameters were separated. The data split remained fixed across seeds. Epoch-end loss and accuracy were measured on the unaugmented training set and the common unaugmented validation set. These clean training metrics are

distinct from the online training loss measured on changing, possibly augmented batches.

Algorithm 1: Training and evaluation procedure

1. Split the standard MNIST training partition into disjoint training and validation sets; reserve the standard test partition.
2. For each seed, create and retain one set of initial MLP weights.
3. For each condition A, B, and C, restore those weights and initialize a fresh Adam optimizer.
4. For each of 15 epochs, generate the shared source-image order.
5. For B and C, select source images with the shared probability-0.5 mask and sample condition-specific transformation parameters.
6. Inverse-sample selected images with bilinear interpolation; assign zero outside the source square. Keep other images unchanged.
7. Train for one epoch, using each source example once.
8. Evaluate the resulting model on clean training and validation data; record metrics and save the lowest-validation-loss model.
9. After all training and selection are complete, evaluate each selected model on the unchanged test set without further training

We report the 15th-epoch validation metrics and the mean validation loss over epochs 11–15 separately. For the final test evaluation, the saved model with the smallest validation loss across epochs 1–15 was selected for each condition and seed. This selection rule was fixed before test evaluation. Test images received no augmentation, and test results were not used to select parameters, seeds, or checkpoints.

Training was performed in Google Colab with TensorFlow 2.20.0, NumPy 2.1.3, SciPy 1.16.3, and scikit-learn 1.6.1. Saved models were evaluated in a separate environment using TensorFlow 2.21.0. Before accessing the test arrays, all nine selected models reproduced their stored validation losses within 10^{-5} and accuracies within 10^{-6} . The largest observed validation-loss discrepancy was approximately 1.5×10^{-8} . Fixed seeds and deterministic TensorFlow operations were used, but bitwise equality across software and hardware environments is not assumed. Split indices, configuration files, epoch logs, and selected model files were retained.

5. Results and Discussion

5.1. Visual Inspection

Figure 1 illustrates the two augmentation procedures on the same training image of digit 4. The Möbius example uses $\alpha = 0.05 + 0.05i$, one of the eight

directions in the training protocol. The affine example uses a 10° rotation and translations of two pixels in both Cartesian directions. These examples illustrate the operations and do not establish equivalent deformation strengths or semantic preservation for every image.

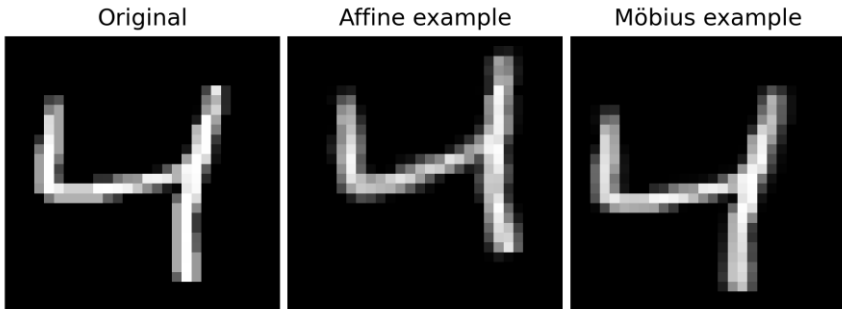


Figure 1: Original MNIST digit 4 (standard training index 2), an affine example, and a Möbius example with $r = 0.70$ and $\alpha = 0.05 + 0.05i$.

5.2. Learning Curves and Validation Results

Figure 2 shows the epoch-end clean training and validation losses averaged over three seeds. The unaugmented condition achieves the lowest clean training loss but the highest validation loss at the end of training. Both augmentation procedures reduce the validation loss relative to this control. Affine and Möbius curves follow similar trajectories and both continue to improve during the later epochs. The results therefore do not support an explanation in which affine augmentation necessarily stops learning while Möbius augmentation alone avoids a plateau.

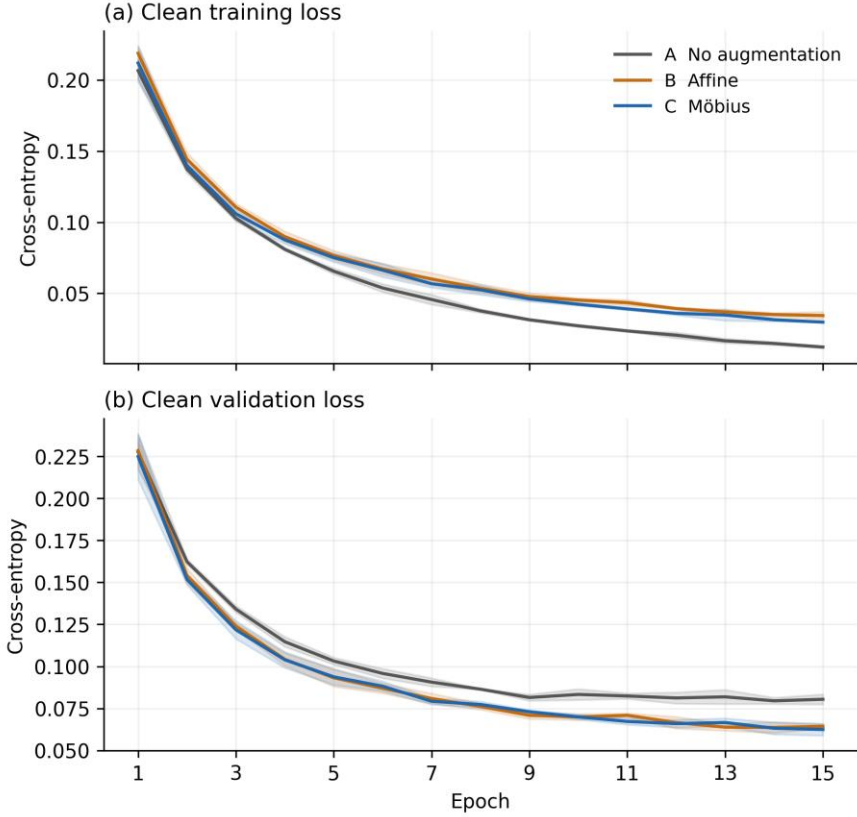


Figure 2: Clean training loss (a) and common clean validation loss (b). Lines show means across three seeds; shaded regions show ± 1 sample standard deviation, not confidence intervals. All conditions use the same fixed data split.

Table 1. Validation results. Epoch-15 values are means $\pm$ sample standard deviations across three seeds. The final column averages both the five specified epochs and the three runs.

Condition	Epoch 15 accuracy (%)	Epoch 15 loss	Epochs 11–15 mean loss
A No augmentation	97.640 ± 0.036	0.08042 ± 0.00299	0.08108
B Affine	97.980 ± 0.036	0.06433 ± 0.00136	0.06585
C Möbius	98.187 ± 0.100	0.06242 ± 0.00357	0.06508

At epoch 15, mean validation losses are 0.08042 for the control, 0.06433 for affine augmentation, and 0.06242 for Möbius augmentation. The average Möbius loss is lower than the affine loss, but the seed-wise comparison is mixed: Möbius has lower final validation loss only for seed 2026. Over epochs 11 – 15, the corresponding means are 0.08108, 0.06585, and 0.06508. These small

differences between the augmented conditions should not be interpreted as uniform superiority.

The final mean differences between clean validation and clean training losses are approximately 0.06818, 0.02995, and 0.03267 for A, B, and C, respectively. The smaller gaps under augmentation are consistent with a regularization effect. They do not demonstrate that overfitting has been eliminated, and the affine condition has the slightly smaller average gap. No representation analysis was performed to identify learned geometric or topological invariants.

5.3. Independent Test Evaluation

The models selected by minimum validation loss were evaluated on the separate 10,000-image test partition. Table 2 reports their performance. These results refer to selected checkpoints, not necessarily the models at epoch 15.

Table 2. Test performance of models selected by validation loss. Values are means $\pm$ sample standard deviations across three seeds.

Condition	Test accuracy (%)	Test loss
A No augmentation	97.663 $\pm$ 0.123	0.07635 $\pm$ 0.00242
B Affine	98.120 $\pm$ 0.082	0.05732 $\pm$ 0.00050
C Möbius	98.200 $\pm$ 0.056	0.05481 $\pm$ 0.00178

Both augmentation conditions outperform the unaugmented control in mean test accuracy and loss. Möbius augmentation attains 98.200% mean accuracy, compared with 98.120% for affine augmentation, a difference of 0.080 percentage points. On this test set, this corresponds to an average of eight fewer errors per run. Möbius test accuracy is higher in two seeds and lower in one, while its test loss is lower in all three seeds.

Table 3. Selected epoch / test accuracy (%) for each run. Epoch selection used validation loss only.

Seed	A control	B affine	C Möbius
42	15 / 97.63	14 / 98.14	15 / 98.19
123	12 / 97.56	14 / 98.19	14 / 98.15
2026	13 / 97.80	12 / 98.03	15 / 98.26

These results indicate a small average advantage for the specified Möbius procedure over the specified affine baseline. Three seeds on one fixed split do not establish statistically significant or broadly generalizable superiority. The

reported standard deviations describe variation between runs on the same test set; they are not confidence intervals for performance on other datasets.

6. Conclusion, Limitations and Future Directions

This study presented an empirical evaluation of unit disk automorphisms for MNIST augmentation with a shallow MLP. The implementation uses an embedded square, inverse mapping, bilinear interpolation, and zero filling. Its continuous conformal properties do not imply preservation of the square image boundary or lossless discretization. Explicit geometric checks and a common validation protocol are therefore necessary when interpreting the method.

Under the selected experimental conditions, both affine and Möbius augmentation improve test performance relative to training without augmentation. Möbius augmentation achieves a slightly higher mean test accuracy and a lower mean test loss than the affine baseline. However, the accuracy advantage is small and is not present in every run. Both augmented conditions continue learning in later epochs. The observed learning curves do not support an affine-specific early learning failure.

The study is limited to MNIST, one shallow architecture, one fixed data split, three seeds, and a 15-epoch training budget. The transformation strengths were not matched, and no comparison with elastic deformation or learned augmentation policies was conducted. The geometric inspection does not guarantee label preservation. Future work could investigate independent data splits, additional architectures and datasets, and controlled comparisons of transformation magnitude and computational cost. The present contribution is an explicit empirical implementation and comparison of established conformal augmentation ideas.

7. Declaration of Computational and AI Tools

The training experiments were implemented in Python using Google Colab, TensorFlow, and Keras. NumPy and SciPy were used for image transformations, scikit-learn for the stratified split, and Matplotlib for visualization. Independent test evaluation used the saved model files in a separate Python environment, as described in Section 4.5.

Google Gemini and OpenAI ChatGPT/Codex were used to assist with code development, debugging, mathematical and methodological checks, literature searches, manuscript editing, and execution of verification and saved-model evaluation scripts. Numerical results reported here were computed from the experimental data and model files. The authors retain responsibility for reviewing and approving the methods, interpretations, references, and final manuscript.

References

- [1] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data*, vol. 6, no. 1, p. 60, 2019.
- [2] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond Empirical Risk Minimization," in *International Conference on Learning Representations (ICLR)*, 2018.
- [3] V. Verma, A. Lamb, C. Beckham, A. Najafi, I. Mitliagkas, D. Lopez-Paz, and Y. Bengio, "Manifold Mixup: Better Representations by Interpolating Hidden States," in *Proceedings of the 36th International Conference on Machine Learning (ICML)*, vol. 97, pp. 6438-6447, 2019.
- [4] T. DeVries and G. W. Taylor, "Improved Regularization of Convolutional Neural Networks with Cutout," *arXiv preprint arXiv:1708.04552*, 2017.
- [5] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in *Proc. Int. Conf. Learn. Representations (ICLR)*, 2014.
- [6] R. Sawhney, M. Thakkar, S. Agarwal, D. Jin, D. Yang, and L. Flek, "HYPMIX: Hyperbolic interpolative data augmentation," in *Proc. Conf. Empirical Methods in Natural Language Processing (EMNLP)*, 2021, pp. 9858-9869.
- [7] T. W. Mitchel, N. Aigerman, V. G. Kim, and M. Kazhdan, "Möbius convolutions for spherical CNNs," in *Proc. SIGGRAPH Conf.*, 2022, Art. no. 3528233.3530724.
- [8] H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian, and M. Wang, "Swin-Unet: Unet-like pure transformer for medical image segmentation," in *European Conference on Computer Vision (ECCV 2022 Workshops)*, Springer, Cham, 2023.
- [9] Z. Cao, J. Zhu, W. Zhang, H. Ai, H. Bai, H. Zhao, and L. Wang, "PanDA: Towards panoramic depth anything with unlabeled panoramas and Möbius spatial augmentation," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, 2025.
- [10] S. Azizi, C. Kornblith, M. Saharia, M. Norouzi, and D. J. Fleet, "Synthetic Data from Diffusion Models Improves ImageNet Classification," *Transactions on Machine Learning Research*, 2023.
- [11] S. Zhou, J. Zhang, H. Jiang, T. Lundh, and A. Y. Ng, "Data augmentation with Mobius transformations," *Mach. Learn.: Sci. Technol.*, vol. 2, no. 2, Art. no. 025016, 2021.
- [12] M. S. Chippa, P. C. Chhipa, K. De, M. Liwicki, and R. Saini, "LCM: Log Conformal Maps for robust representation learning to mitigate perspective distortion," in *Proc. Asian Conf. Computer Vision (ACCV)*, 2024.

- [13] O. Rainio, M. M. S. Nasser, M. Vuorinen, and R. Klén, "Image augmentation with conformal mappings for a convolutional neural network," *Comput. Appl. Math.*, vol. 42, no. 8, Art. no. 361, 2023.
- [14] L. V. Ahlfors, *Complex Analysis*, 3rd ed. New York, NY, USA: McGraw-Hill, 1979.
- [15] T. Needham, *Visual Complex Analysis*. Oxford, U.K.: Oxford University Press, 1997.
- [16] E. M. Stein and R. Shakarchi, *Complex Analysis*. Princeton, NJ, USA: Princeton University Press, 2003.
- [17] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998. doi: 10.1109/5.726791

Artificial Intelligence Applications in Battery Management Systems

Faruk KURKER¹

1. Introduction

The increasing demand for energy globally has led to a growing number of applications using lithium-ion batteries, including electric vehicles (EVs) and renewable energy storage. Due to the many advantages of lithium-ion batteries, including their high energy density, low self-discharge rate, long cycle life, and high voltage per cell, they have become the preferred method of storing energy in many different types of applications, including portable electronic devices, electric vehicles, and grid applications [1,2]. However, the ability to operate a lithium-ion battery safely, reliably and efficiently relies on the accurate knowledge of its instantaneous internal state at all times. Therefore, battery management systems (BMSs) are essential; BMSs use voltage, current, and temperature data to help estimate the internal state variables of the battery that cannot be directly measured (for example, SOC, SOH, SOP, and SOE) [3,4]. SOC and SOH are two of the most important outputs produced by a BMS. The SOC is defined as the ratio of the currently available charge to the rated capacity of the battery, which can be compared to the fuel gauge found in traditional vehicles. The SOH is an indication of how much the battery has degraded when compared with a new battery; it measures the remaining performance capacity of the battery, taking into account both the capacity loss and the increase in internal resistance since it was manufactured [5,6]. Since SOC and SOH cannot be measured directly, they must be estimated using indirect methods based on the measurement of other quantities (voltage, current, and temperature). The challenge of estimating SOC and SOH is made even more difficult by the non-linear electrochemical nature of batteries and their dependence on temperature, age, charge/discharge rate, and usage history [7,8].

The methods that have historically been developed for estimating SOC and SOH can be divided into four categories: experimental methods, physics-based (model-based) methods, data-driven (artificial intelligence/machine learning)

¹ Dr. Öğr. Üyesi , Department of Electrical and Electronics Engineering, Faculty of Engineering, Adiyaman University, Adiyaman, 02040, Türkiye
ORCID: 0000-0003-1544-9743

methods, and hybrid methods combining the previous three categories of estimation methods [9]. While experimental methods yield the highest level of accuracy, they require extensive testing over a long period and expensive equipment found in laboratory environments and thus are not feasible for use in real-time in-vehicle applications. Physics-based methods attempt to model the internal dynamics of the battery using equivalent circuit models or electrochemical models; while these models have physical interpretability, they are difficult to set up, and as a result, their performance in real-time is limited. On the other hand, data-driven methods that employ artificial intelligence and machine learning approaches can directly learn the nonlinear relationships between the recorded input data and the desired output (SOC or SOH) from the actual data without having to explicitly model the internal mechanisms of the battery. Thus, AI-driven methodologies are becoming the fastest growing and heavily researched methodologies dealing with estimating battery states over the last decade [10,11].

It is important to understand the hierarchical relationship between Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) in order to appropriately classify the methodologies that have been used to estimate a battery's state. The most general of the three terms, Artificial Intelligence (AI) refers to the entire body of research that has been done with the goal of creating machines that can replicate human-like abilities to make decisions and solve problems. The general category of Artificial Intelligence contains the subset of Machine Learning (ML), which is defined as the method where the machine learns via observation of patterns in data rather than via explicit programming. Finally, Deep Learning (DL) falls into the subset of Machine Learning (ML) and refers to the methodology in implementing a multi-layered Artificial Neural Network to build abstract structural representations of data obtained from the raw data. The key components of these three interrelated areas and the historical progress in the development of these areas of expertise are illustrated in Figure 1; beginning with some of the first notable developments in Artificial Intelligence in the 1950s and progressing through the maturation of ML in the 1980s and 1990s to the excitement generated by the development of DL in the 2010s that has set off yet another wave of interest in AI.

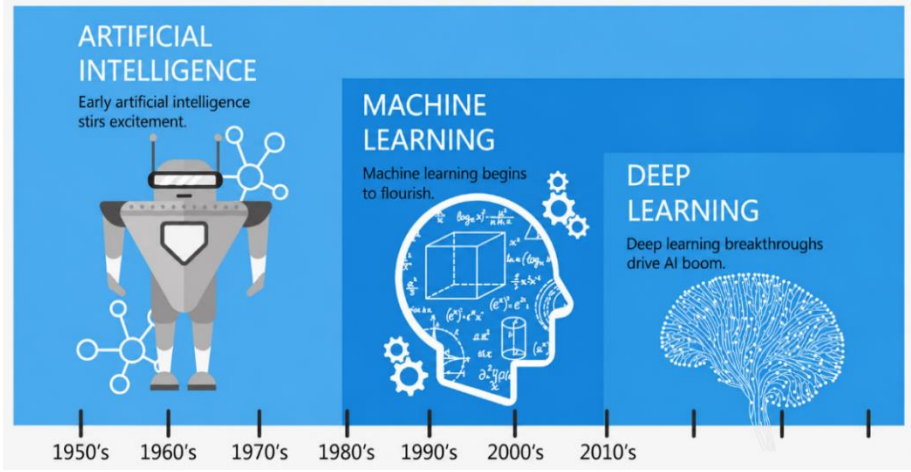


Figure 1. The relationship between artificial intelligence, machine learning, and deep learning [12]

Machine learning techniques provide some advantages over classical statistical methods regarding estimating battery state. Specifically, ML techniques provide enhanced ability to approximate complex, non-linear functions, allow for the modelling of changes in varying operating conditions, handle varying levels of noise, and assist with the capability to process complex patterns. Over the past few years, as a result of these advantages, both machine learning and deep learning have emerged as key techniques used to estimate SOC and SOH for challenging scenarios. Some examples of such challenging scenarios are dynamic driving conditions, widely varying temperature ranges, and various battery chemistries where standard methodology does not suffice [13]. The broad application of ML in the estimation of SOC and SOH has become commonplace due to the recent advancements in deep learning models incorporating convolutional, recurrent, and attention-mechanism architectures. With these advances have come new models based on transformer architectures designed to improve performance in the estimation of SOC and SOH in recent years [14,15].

2. Battery Management System Architecture

A battery management system (BMS) monitors the battery pack's (Rechargeable) status and parameters to ensure maximum safety and high efficiency in operating both the BMS and the battery pack. As shown by Figure 2, the BMS is divided into functional subsystems (sensing, controlling, monitoring) with components that transfer information between them via the communication interface. The data acquisition unit, which is placed centrally in the BMS structure, is responsible for continually measuring several critical

battery parameters (current, voltage, and temperature) and sending that data to other control units to process. Accurate knowledge of these battery parameters allows for reliable assessment of the operational status of a battery pack.

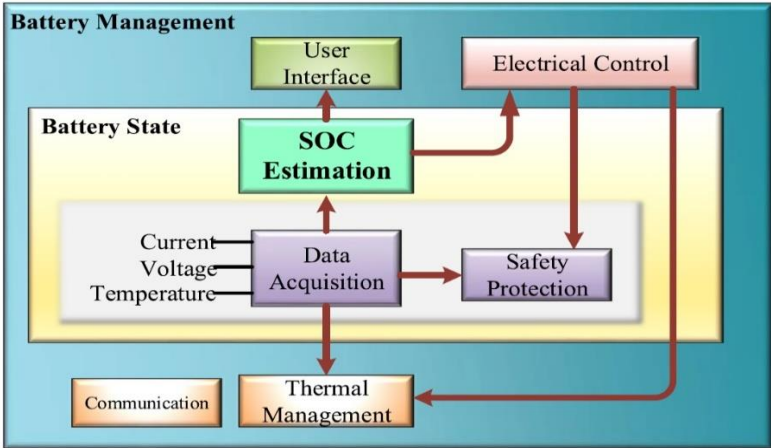


Figure 2. Block diagram of a battery management system [3]

The data collected from the Data Acquisition Unit (DAU) is used for state of charge (SOC) estimation, safety protection and thermal management. SOC Estimation provides users with an indication of how much ‘charge/capacity’ remains (available) in the battery. It supports the User Interface and Electrical Control Systems’ (ECS) decision-making processes. Safety protection protects against all (or several) adverse operational conditions, including overcurrent, overvoltage, undervoltage and exceeding temperature limits. Thermal management protects against overheating, especially when operating under extreme load conditions. The ECS uses the SOC, safety and thermal condition data collected to manage the charging and discharging of the battery. The User Interface (UI) and Communications Units offer an overall view of the BMS functions and enable communication with external systems. Figure 3 shows the physical structure of a (BMS) and demonstrates how information flows between the core functions of the battery management systems. Voltage, current and temperature data from cell packs are sent to the Measurement Unit (MU), which generates the BMS’s core decision-making data. The charge and state of charge (SOC) and state of health (SOH) are estimated from this data, and also includes thermal management (TM) functions, keeping the battery operational within a safe temperature range. This data is passed on to the Control Unit so that the battery system can be safely and effectively managed.

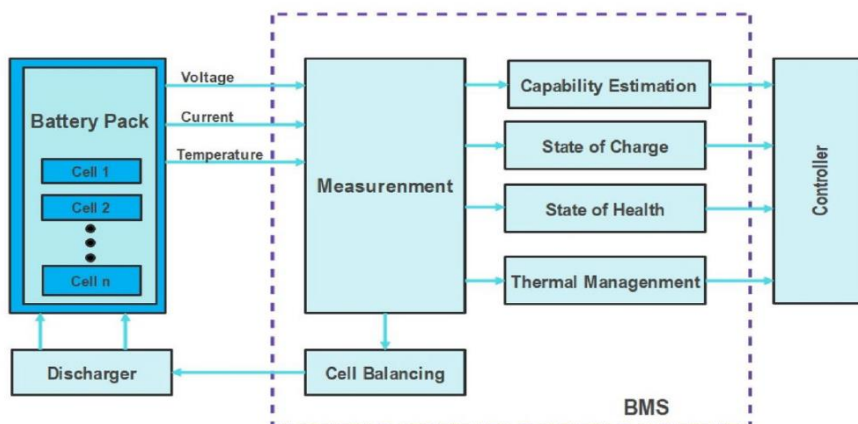


Figure 3. Architecture of a battery management system [4]

The cell balancing function of the BMS is seen to be a crucial aspect of the BMS as depicted in Figure 3. Cell balancing minimises voltage and capacity imbalances within series-connected cells, which result from factors such as manufacturing tolerances, differences in temperature and aging. Excess energy present in certain cells can be safely removed via the discharge unit when needed. In doing so, the BMS ensures that cell imbalances are limited, maximises the overall usable capacity of the battery pack and prevents cells from experiencing overcharging or over-discharging. Therefore, Figure 3 clearly indicates that, although based on measurements, the BMS is much more than a simple measurement device; it is the principal control unit that will manage all of the sub-systems of the BMS, including state estimation, thermal management, cell balancing and high-level control.

3. SOC Estimation Methods

Different methods used to estimate the State of Charge (SOC) of batteries through Battery Management Systems (BMS) can be divided into several general categories according to their basic principles, as shown in figure 4. SOC is an important calculation as it tells you how much capacity is left in your battery, but it isn't possible to measure SOC directly; therefore it must be estimated with many different mathematical, physical, and data-based approaches - which are grouped in figure 4 into 5 basic categories: Coulomb counting methods, lookup table algorithms, model-based methods, data-based methods and hybrid techniques.

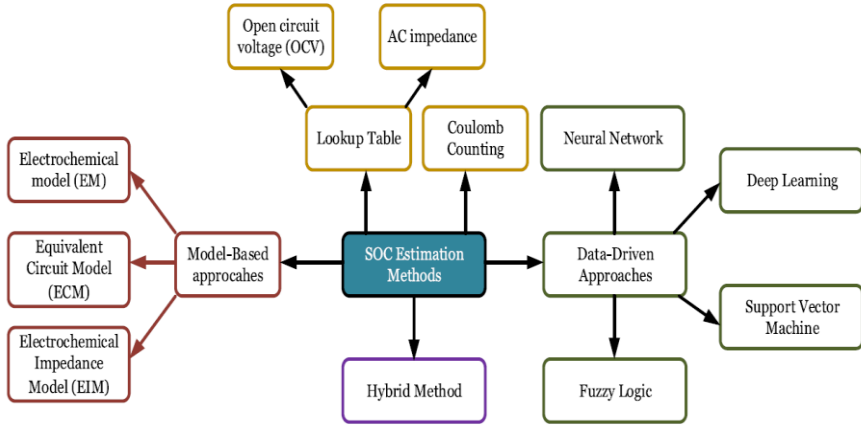


Figure 4. SOC estimation methods within the BMS [3]

The technique called Coulomb counting represents one of the most straightforward approaches to estimating the state of charge (SOC) for batteries and is based on measuring how much current flows into and out of the battery during a certain period. It is one of the most commonly used SOC estimation techniques because it can be implemented easily. However, with continued use, estimating SOC using Coulomb counting becomes increasingly dependent on the uncertainty associated with how long before initial SOC measurement is taken and the accumulation of measurement errors in the SOC value from time to time.

Lookup table-based methods estimate SOC using measured variables (for example, open-circuit voltage) and predefined relationships that exist between the measured variables and SOC. Even though lookup table methods are also simple to implement and inexpensive to use, they generally have a much larger degree of sensitivity to the impact of temperature, ageing, and other operating conditions than do Coulomb counting methods.

The third general method for estimating SOC in a lithium-ion battery system involves the use of model-based SOC estimation methodologies. Model-based methodologies are based on mathematical representations of the physical and electrochemical processes that occur in batteries. As depicted in the figure, model-based methodologies can be divided into three main categories—electrochemical models, equivalent circuit models, and electrochemical impedance models. Model-based methodologies are accurate at representing the dynamic behaviour of battery systems. However, model-based methodologies require the accurate measurement of all of the model parameters; this is particularly true with equivalent circuit models, which are often used in real-time

battery management systems because they provide a reasonable compromise between accuracy and computational complexity.

The fourth general method for SOC estimation is data-driven methodologies. Examples of data-driven methodologies are artificial neural networks, deep learning, support vector and fuzzy logic-based systems. Data-driven methodologies reduce the need for physical models of battery systems; they learn and understand non-linear relationships within structured measurement data such as the current, voltage, and temperature of a battery system. However, data-driven methodologies rely heavily on the volume, diversity, and quality of the datasets used to train the models.

Hybrid methodologies combine the strengths of both model-based and data-driven SOC estimation methodologies to improve prediction accuracy and robustness. As illustrated in the figure, the appropriate method for estimating SOC depends on multiple factors, including the level of accuracy required, computational load, the ability to operate in real-time, the complexity of the model used to represent SOC, and the ability of the model to adapt to the varying conditions of operation.

The method comparison presented in Table 1 helps to illustrate the advantages and disadvantages of the most common artificial intelligence (AI) and machine learning (ML) methodologies for estimating SOC and other similar battery-related applications. In this regard, Back Propagation Neural Networks (BPNN) have the strongest nonlinear mapping ability and self-learning and adaptive capabilities. However, BPNNs are often limited by their slow convergence rates, lengthy training times, uncertain parameter selection, and risk of being trapped in local minima.

Radial Basis Function (RBF) Neurons offer computational advantages over BPNNs due to their ability to provide local approximations, good interpolation and fast convergence rates. However, RBF training relies heavily on the choice of basis functions used, which can significantly impact the overall model performance and may result in overfitting of the model. Extreme Learning Machines (ELM) also garner interest due to their simple structure, relatively few model parameters, rapid convergence and high accuracy potential. However, the ELMs' high degree of variability in performance depending on the accuracy and quality of parameter selection is a significant disadvantage.

Wavelet Neural Networks (WNN) have notable advantages over traditional BPNN in terms of greater accuracy in approximating outputs and tolerance for error; however, WNN are also very difficult to apply practically because of the potential complexities associated with determining the parameter selection.

Table 1. Advantages and disadvantages of different models [13]

Model Name	Model Advantages	Model Disadvantages
BPNN	Mature technology with strong nonlinear mapping, self-learning, and adaptive capabilities. It also provides good generalization and fault-tolerance performance.	Slow convergence, difficulty in parameter selection, and long training time. It is also prone to becoming trapped in local minima.
RBFNN	Provides local approximation, good interpolation capability, simple computation, and fast convergence.	Performance is highly influenced by the choice of basis function and the model may be prone to overfitting.
ELM	Simple structure, few model parameters, fast convergence, and potentially high prediction accuracy.	Model performance is highly dependent on appropriate parameter selection.
WNN	Provides advantages over conventional BP networks in terms of design, approximation accuracy, sensitivity, and fault tolerance.	Model parameters are relatively difficult to determine and optimize.
CNN	Effectively handles complex nonlinear relationships and high-dimensional data, enabling high prediction accuracy.	Requires long training times and relatively high computational cost.
RNN	Capable of modeling temporal dependencies and short-term sequential information.	Susceptible to gradient explosion and related training instability.
LSTM	Provides long-term memory and controlled memory mechanisms, high prediction accuracy, and alleviates the vanishing-gradient problem encountered in conventional RNNs.	Has a relatively complex architecture, resulting in increased computational requirements and longer training times.
GRU	Simple architecture with fewer model parameters, easier parameter adjustment, high prediction accuracy, and reduced vanishing- and exploding-gradient problems.	Cannot completely eliminate the vanishing-gradient problem. Although it contains fewer parameters than LSTM, its modeling capability may vary depending on the application.
SVM	Effective in controlling underfitting and overfitting and is particularly suitable for high-dimensional, small-sample datasets.	Sensitive to parameter settings and noise, with high computational complexity; therefore, it may be less suitable for very large datasets.

Based on the evaluation of these machine-learning and deep-learning approaches, it appears that no single model is the best for every application, and that the selection of the appropriate model must be dependent upon many different factors, including (but not limited to) the amount of data being used, the temporal nature of the data, the accuracy of the results to be produced and how quickly a result needs to be produced. Specifically with regard to battery management systems where temporal changes of data need to be monitored in real time, neural network architectures that have been specifically designed to

process sequential data appear to be the most adaptable for battery management applications, while for smaller and higher-dimensional datasets or where speed at which results can be produced is of primary importance, methods such as SVMs or ELMs may be the best options to consider.

4. SOH and RUL Estimation Methods

4.1. SOH Estimation Methods

The State of Health (SOH) shows the amount of decline in a battery's ability to deliver energy or power based on age compared to when it was new. SOH is defined by taking the current capacity of the battery and comparing it to its rated capacity at the time of purchase, as well as looking at the internal resistance of the battery at this date and comparing it to when the battery is nearing the end of its useful life (end-of-life, or EOL). Battery aging occurs due to two primary processes: calendar and cycle aging. Calendar aging is caused solely by time and occurs during storage of the battery when it is stored at a high State of Charge (SOC) or at extremely high temperatures, whether or not the battery is being used. Cycle aging has additional influences from the charge-discharge current rate, depth of discharge and voltage range. Calendar and cycle aging degrade the battery through three primary mechanisms: loss of lithium inventory, loss of available active material and an increase in the internal resistance of the battery. In addition to providing a direct relationship between the SOH and capacity degradation, including aging mechanisms will allow batteries to be more accurately predicted to experience capacity degradation for a given amount of time and/or conditions [5,16,17].

Battery SOH was classified into four categories, which are presented in Figure 5, with a comparison of the advantages and disadvantages of these methods. Battery State of Health (SOH) estimation methods are divided into four main types: experimental-based, physics-based, data-driven and hybrid methods. Experimental-based methods use direct measurements of the battery's current state using methods such as open circuit voltage (OCV), capacity testing or electrochemical impedance spectroscopy (EIS). The advantages of the experimental-based methods are that they provide accurate and high-quality data, due to the physical nature of the experimental measurement techniques. However, disadvantages include a time consuming nature and high associated costs in addition to not being very practical for on-demand testing and identification of degradation. Physics-based methods provide a set of mathematical representations of the electrochemical models of the internal dynamics of the battery, including equivalent circuit models. While these methods have advantages in providing results that are interpretable in terms of the battery's

physical and electrochemical behaviour, they present difficulties in terms of determining the necessary parameters for the model to work correctly, resulting in a cumbersome structure that can lead to a higher computational burden than what was expected. Data-driven approaches to estimating SOH established direct correlations between the battery data collected through measurements of the battery, using either traditional machine learning or deep learning, on a historical basis, thus creating a method or approach that can be continually updated on a real-time basis and be highly practical with respect to battery management systems. Data-driven techniques have critical limitations. Model performance is highly dependent upon the quality and quantity of representative data used for training the model [18,19].

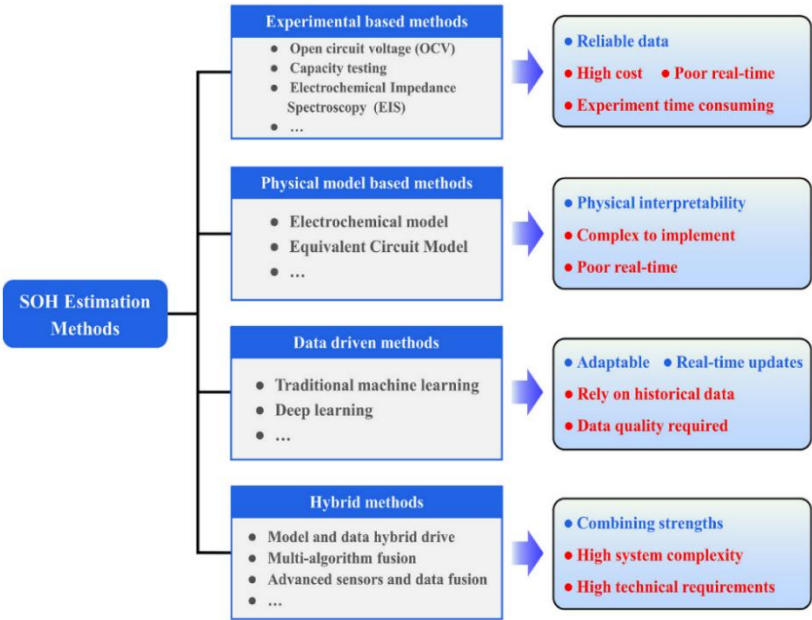


Figure 5. SOH estimation methods and their advantages/disadvantages [9]

Hybrid techniques that combine physical modelling, data-driven methods and/or multiple algorithms, and/or the techniques of advanced sensors and data fusion methods have the ability to take advantage of the strengths of the various approaches that are available. As such, hybrid techniques have tremendous potential for accuracy, robustness and generalisation. However, the increasing complexity of these systems and the growing computational and integrated requirements associated with hybrid techniques, as well as the additional need for highly technical expertise is a primary disadvantage of hybrid techniques. In

summary, the diagram illustrates that the method selected for estimating SOH will depend upon the trade-off between accuracy, real-time application, data requirements, model complexity and computational cost as presented in Table 2. A comparison is made between the analytical and machine-learning methods to estimate battery health; including both non-probabilistic and probabilistic methods regarding the advantages and disadvantages of each. Simpler methods, such as semi-empirical and AR models, have low computational burden and ease of use, but their long-term prediction capability and generaliseability may be limited. However, although ANN and SVM methods provide high prediction accuracy, they also face such issues as overfitting, hyperparameter selection, and computational cost. In contrast to ANN and SVM methods, GPR and RVM methods offer the benefit of incorporating uncertainty information into the model. However, as the size of the training data increases, GPR becomes increasingly computationally expensive, while RVM requires a significant amount of time and memory to train. Therefore, the evaluation of the method selected should occur by considering the combined factors of accuracy, computational cost, data volume, generalisability, and real-time application.

Table 2. Comparison of battery health prediction methods [5].

Method Category	Model	Advantages	Disadvantages
Analytical model with data fitting	Semi-empirical model	• Easy to construct and capable of producing rapid predictions. • Model parameters can be readily extracted. • Low computational burden. • Suitable for online implementation in BMS by monitoring cycling conditions, operating time, and/or cycle number.	• Requires extensive laboratory testing over the entire operating range, resulting in considerable time and cost. • Developing suitable ageing tests to characterize interactions among different degradation mechanisms is challenging. • Limited generalizability, as models are generally restricted to specific battery types and operating conditions.
Analytical model with data fitting	Empirical ageing model with filtering	• Requires only a relatively small amount of ageing data for model development. • Estimation errors can be continuously updated using actual measurements.	• Relatively high computational burden, which may complicate online implementation.
Machine learning (ML)	AR model	• Simple model structure. • Parameters can be easily identified. • Easy to implement.	• Susceptible to underfitting because of its linear regression structure. • Limited generalization capability.

			• Poor long-term prediction performance.
Machine learning (ML)	ANN	• Strong capability for modeling nonlinear relationships. • RNN-based architectures provide strong long-term RUL prediction capability through recurrent connections. • High prediction accuracy.	• Susceptible to overfitting. • Limited capability for uncertainty quantification. • Performance is highly dependent on the training process.
Machine learning (ML)	SVM	• High prediction accuracy. • Non-parametric structure. • Robust to outliers. • Short prediction time.	• High computational cost. • Limited uncertainty quantification capability. • Requires cross-validation for hyperparameter selection.
Machine learning (ML)	GPR	• Provides predictive covariance, enabling uncertainty quantification. • Non-parametric. • Flexible modeling capability.	• Performance is strongly influenced by the choice of kernel function. • High computational cost, particularly for large datasets.
Machine learning (ML)	RVM	• Can directly provide a probability distribution for predictions. • Non-parametric. • Produces highly sparse models. • Can reduce the need for extensive cross-validation.	• Requires a sufficient amount of data for reliable modeling. • Training may involve considerable time and memory consumption. • Susceptible to local optimization problems. • Potential risk of overfitting.

Supervised Learning Algorithms, including ANN, SVM, RVM, KNN and GPR, are popular for SOH prediction in the existing Literature. Convolutional Transformer Hybrid Architectures have been shown to improve the accuracy of SOH predictions by combining multiple views of the same dataset, while GMDH technique increase the robustness of SOH predictions and frameworks that do not require additional degradation testing serve to further reduce the cost of data collection [20-23]. Cycle based architectures and Single Use Strategies have been developed to generate and identify SOH trend predictions over the life of a Lithium-Ion Battery with limited cycle points [24-26]. Similar probabilistic GPR-based techniques have been applied not only to Lithium-Ion Batteries but also for predicting the SOH of other electro-chemical Energy Storage devices, such as Super Capacitors. Additionally, there is evidence to suggest that a number of deep

learning SOH estimation methods have been developed for other electro-chemical systems, including, PEM Fuel Cells [27].

4.2. RUL Estimation Methods

The literature lists two primary approaches for remaining useful life (RUL) estimation: analytical (semi-empirical life models and filtering-based empirical aging models) and machine learning based. Machine learning-based RUL prediction methods have been classified into two types: non-probabilistic (autoregressive models, ANN, SVM) and probabilistic (GPR, RVM). The primary advantage of the probabilistic methods is that in addition to providing a point estimate of RUL, they also provide confidence limits for that estimate; this feature is particularly important in battery applications where safety is of utmost importance [18]. Long Short-Term Memory (LSTM) networks have shown to be much more accurate at estimating RUL than classical methods, as they capture the long-term dependencies within the cycle-to-cycle time series data. Hybrid architectures that combine LSTM and Elman networks have also been shown to provide competitive accuracy for RUL estimation [28-31]. Gated Recurrent Unit (GRU)-based recurrent neural networks have also been applied successfully to estimate the lifespan of other energy storage components such as onboard supercapacitor systems. In addition, since the state of health (SOH) and RUL are not independent but are two interdependent values, techniques have been developed that enable both SOH and RUL estimation to be performed simultaneously (co-estimation) using optimally designed hybrid frameworks. Additionally, it has recently been shown that next-generation deep learning techniques, such as knowledge distillation-based methods for cross-modal learning and Graph Neural Networks (GNNs) that have adaptive loss functions, increase the accuracy of SOH predictions [32-35].

4.3. AI-Based Thermal Management and Fault Diagnosis

Artificial neural networks can accurately predict temperature profiles and heat generation rates in liquid-cooled air-cooled cooling systems and phase-change (PCM) based heat sink systems. Convolutional neural networks can be used as fast surrogate models to replace finite element based modelling of PCM microstructure. Long Short-Term Memory (LSTM) networks are capable of detecting abnormal heat generation before thermal runaways occur by processing temperature time-series data, thus allowing for early warning [36]. The fact that thermal runaway is a domino effect/prior cause-and-effect chain reaction and begins with the first stage of self-heating and micro-internal short circuit formation; progresses through the second stage of separator melting and the

appearance of internal short circuits; finally arriving at the third stage of cathode decomposition and electrolysis leading to the potential fire/explosion hazard demonstrates that early detection can enhance prevention and response strategies to thermal runaway. Hybrid algorithms that combine data-driven and model-driven techniques or data-driven models/methods with density-based clustering (DBSCAN) have been shown to provide reliable thermal runaway alerts at the earliest stages of thermal runaway [37–39]. Simultaneously, voltage failure diagnostics using entropy and Z-score and predicting voltage anomalies with LSTM networks are two additional complementary predictive techniques that can provide early warning of battery failures [40–42].

5. Conclusions

This section presents an analysis of various methods for estimating SOC, SOH, and RUL in batteries, including model-based, traditional, and AI-supported approaches. From the overall analysis, it is evident that today's battery management systems have progressed from basic measurement and safety functions to include intelligent capabilities for state estimation, cell balancing, thermal management, fault diagnosis, and lifespan forecasting. Data-driven approaches, including those from the fields of Machine Learning and Artificial Intelligence, take the opposite approach by building from measurement data (voltage, current and temperature) to provide direct learning of the complex, nonlinear coefficients that exist between these measurements and battery state variables. All of the AI models evaluated exhibit both unique strengths and limitations.

Accordingly, the key criteria for selecting a model should include prediction accuracy and practicality, data volume, computational capacity, and real-time applicability. Model-based methods offer a relatively fast and low-cost alternative for SOH/RUL prediction. However, they have limited generalizability across different battery types and operating conditions. Although Artificial Neural Networks and Support Vector Machines provide a high level of accuracy in predictions, neither may adequately represent uncertainty. On the other hand, probabilistic methods such as Gaussian Process Regression and Relevance Vector Machines offer a significant advantage in safety-critical applications because they provide confidence intervals for the predicted value and, due to their lower data requirements compared to model-based methods, are better suited for real-time use with BMS.

AI also provides strong capabilities for modelling batteries for temperature distribution, early detection of potential thermal runaway indicators, and identification of voltage anomalies, in addition to SOC and lifespan estimation.

The ultimate success of these AI architectures relies on the adequacy of training data sets to account for the entire range of operating temperatures, levels of degradation, all charging strategies, and driver behaviour patterns in the validation phase. Therefore, for the purposes of verification, validation with laboratory data alone is not sufficient; validation will require independent validation against real-world performance data. Thus, there is currently no definitive predictive methodology that is superior under every potential condition. Hybrid modelling techniques that incorporate physical models along with AI-based approaches are likely to yield greater accuracy, robustness and applicability in future studies. Models developed using AI methodologies that have the capacity for representing uncertainty in predictions, generating interpretable outputs, performing real-time updates and being deployed in low-cost embedded hardware will be critical in the development of improved, safer and more energy-efficient battery management systems.

References

- [1] Kurucan, M., Özbaltan, M., Yetgin, Z., Alkaya, A. (2024). Applications of Artificial Neural Network Based Battery Management Systems: A Literature Review. *Renewable and Sustainable Energy Reviews*, 192, 114262. <https://doi.org/10.1016/j.rser.2023.114262>
- [2] Sun, Z., Eskandari, M., Zheng, C., Li, M. (2022). Handling Computation Hardness of Battery Energy Storage Scheduling in Microgrids by Deep Reinforcement Learning. *Energies*, 16(1), 90. <https://doi.org/10.3390/en16010090>
- [3] How, D.N.T., Hannan, M.A., Hossain Lipu, M.S., Ker, P.J. (2019). State of Charge Estimation for Lithium-Ion Batteries Using Model-Based and Data-Driven Methods. *IEEE Access*, 7, 136116-136136. <https://doi.org/10.1109/ACCESS.2019.2942213>
- [4] Obuli Pranav, D., Preethem S. Babu, Indragandhi, V., Ashok, B., Vedhanayaki, S., Kavitha, C. (2024). Enhanced SOC Estimation of Lithium Ion Batteries with RealTime Data Using Machine Learning Algorithms. *Scientific Reports*, 14(1), 1-17. <https://doi.org/10.1038/s41598-024-66997-9>
- [5] Li, Yi, Liu, Kailong, Foley, Aoife M., Zülke, Alana, Berecibar, Maitane, Nanini-Maury, Elise, Van Mierlo, Joeri, Hoster, Harry E. (2019). Data-Driven Health Estimation and Lifetime Prediction of Lithium-Ion Batteries: A Review. *Renewable and Sustainable Energy Reviews*, 113, 109254. <https://doi.org/10.1016/j.rser.2019.109254>
- [6] Hossain Lipu, M.S., Hannan, M.A., Hussain, A., Hoque, M.M., Ker, P.J., Saad, M.H.M., Ayob, A. (2018). A Review of State of Health and Remaining Useful Life Estimation Methods for Lithium-Ion Battery in Electric Vehicles. *Journal of Cleaner Production*, 205, 115-133. <https://doi.org/10.1016/j.jclepro.2018.09.065>
- [7] Vidal, C., Malysz, P., Kollmeyer, P., Emadi, A. (2020). Machine Learning Applied to Electrified Vehicle Battery State of Charge and State of Health Estimation. *IEEE Access*, 8, 52796-52814. <https://doi.org/10.1109/ACCESS.2020.2980961>
- [8] Xiao, Y., Wen, J., Yao, L., Zheng, J., Fang, Z., Shen, Y. (2023). A Comprehensive Review of Lithium-Ion Battery State of Health Prognosis Methods. *Journal of Energy Storage*, 65, 107347. <https://doi.org/10.1016/j.est.2023.107347>
- [9] Wang, Yaxuan, Guo, Shilong, Cui, Yue, Deng, Liang, Zhao, Lei, Li, Junfu, Wang, Zhenbo (2025). A Comprehensive Review of Machine Learning-Based State of Health Estimation for Lithium-Ion Batteries. *Renewable*

- and Sustainable Energy Reviews, 224, 116125. <https://doi.org/10.1016/j.rser.2025.116125>
- [10] Massaoudi, M., Abu-Rub, H., Ghayeb, A. (2024). Advancing Lithium-Ion Battery Health Prognostics With Deep Learning. IEEE Open Journal of Industry Applications, 5, 43-62. <https://doi.org/10.1109/OJIA.2024.3354899>
- [11] Guirguis, J., Ahmed, R. (2024). Transformer-Based Deep Learning Models for SOC and SOH Estimation of Li-Ion Batteries: A Survey Study. Energies, 17(14), 3502. <https://doi.org/10.3390/en17143502>
- [12] Li, Ao, Weng, Jingwen, Yuen, Anthony Chun Yin, Wang, Wei, Liu, Hengrui, Lee, Eric Wai Ming, Wang, Jian, Kook, Sanghoon, Yeoh, Guan Heng (2023). Machine Learning Assisted Advanced Battery Thermal Management System: A State-of-the-Art Review. Journal of Energy Storage, 60, 106688. <https://doi.org/10.1016/j.est.2023.106688>
- [13] Zhao, Feng, Guo, Yun, Chen, Baoming (2024). A Review of Lithium-Ion Battery State of Charge Estimation Methods Based on Machine Learning. World Electric Vehicle Journal, 15(4), 131. <https://doi.org/10.3390/wevj15040131>
- [14] Guirguis, J., Abdulkmaksoud, A., Ismail, M., Kollmeyer, P.J., Ahmed, R. (2024). Transformer-Based Deep Learning Strategies for Lithium-Ion Batteries SOX Estimation. IEEE Access, 12, 167108-167119. <https://doi.org/10.1109/ACCESS.2024.3495560>
- [15] Hannan, M.A., How, D.N.T., Hossain Lipu, M.S., Mansor, M., Ker, P.J., Dong, Z.Y., Sahari, K.S.M., Tiong, S.K., Muttaqi, K.M., Indra Mahlia, T.M., Blaabjerg, F. (2021). Deep Learning Approach Towards Accurate State of Charge Estimation Using Self-Supervised Transformer Model. Scientific Reports, 11, 19541. <https://doi.org/10.1038/s41598-021-98915-8>
- [16] Meng, H., Li, Y.-F. (2019). A Review on Prognostics and Health Management (PHM) Methods of Lithium-Ion Batteries. Renewable and Sustainable Energy Reviews, 116, 109405. <https://doi.org/10.1016/j.rser.2019.109405>
- [17] Ren, J., Cai, J., Li, J. (2021). High Precision Implicit Function Learning for Forecasting Supercapacitor State of Health Based on Gaussian Process Regression. Scientific Reports, 11, 12081. <https://doi.org/10.1038/s41598-021-91241-z>
- [18] Thelen, A., Huan, X., Paulson, N., Onori, S., Hu, Z., Hu, C. (2024). Probabilistic Machine Learning for Battery Health Diagnostics and

- Prognostics. *npj Materials Sustainability*, 2(1), 14. <https://doi.org/10.1038/s44296-024-00011-1>
- [19] Li, S., He, H., Su, C., Zhao, P. (2020). Data Driven Battery Modeling and Management Method with Aging Phenomenon Considered. *Applied Energy*, 275, 115340. <https://doi.org/10.1016/j.apenergy.2020.115340>
- [20] Liu, P., Sun, Z., Wang, Z., Zhang, J. (2018). Entropy-Based Voltage Fault Diagnosis of Battery Systems for Electric Vehicles. *Energies*, 11(1), 136. <https://doi.org/10.3390/en11010136>
- [21] Bai, T., Wang, H. (2023). Convolutional Transformer-Based Multiview Information Perception Framework for Lithium-Ion Battery State of Health Estimation. *IEEE Transactions on Instrumentation and Measurement*, 72, 1-12. <https://doi.org/10.1109/TIM.2023.3300451>
- [22] Wei, Z., Sun, X., Wang, J., Liu, W., Rong, C., Liu, C. (2024). State of Health Estimation of Lithium-Ion Batteries Based on Multi-Health Features Fusion and Improved GMDH. *IEEE Transactions on Instrumentation and Measurement*, 73, 1-15. <https://doi.org/10.1109/TIM.2024.3417600>
- [23] Lu, J., Xiong, R., Tian, J., Wang, C., Sun, F. (2023). Deep Learning to Estimate Lithium-Ion Battery State of Health Without Additional Degradation Experiments. *Nature Communications*, 14(1), 2760. <https://doi.org/10.1038/s41467-023-38458-w>
- [24] Bairwa, B., Pareek, K., Jadoun, V.K. (2025). Cycle Based State of Health Estimation of Lithium Ion Cells Using Deep Learning Architectures. *Scientific Reports*, 15(1). <https://doi.org/10.1038/s41598-025-20995-7>
- [25] Li, W., Sengupta, N., Dechent, P., Howey, D., Annaswamy, A., Sauer, D.U. (2021). One-Shot Battery Degradation Trajectory Prediction with Deep Learning. *Journal of Power Sources*, 506, 230024. <https://doi.org/10.1016/j.jpowsour.2021.230024>
- [26] Khumprom, P., Yodo, N. (2019). A Data-Driven Predictive Prognostic Model for Lithium-Ion Batteries Based on a Deep Learning Algorithm. *Energies*, 12(4), 660. <https://doi.org/10.3390/en12040660>
- [27] Zhang, Y., Tang, X., Xu, S., Sun, C. (2024). Deep Learning-Based State of Health Estimation of PEM Fuel Cells Under Dynamic Operation Conditions. *Sensors*, 24(14), 4451. <https://doi.org/10.3390/s24144451>
- [28] Zhang, Y., Xiong, R., He, H., Pecht, M.G. (2018). Long Short-Term Memory Recurrent Neural Network for Remaining Useful Life Prediction of Lithium-Ion Batteries. *IEEE Transactions on Vehicular Technology*, 67, 5695-5705. <https://doi.org/10.1109/TVT.2018.2805189>
- [29] Li, X., Zhang, L., Wang, Z., Dong, P. (2019). Remaining Useful Life Prediction for Lithium-Ion Batteries Based on a Hybrid LSTM-Elman

- Model. *Journal of Energy Storage*, 21, 510-518. <https://doi.org/10.1016/j.est.2018.12.011>
- [30] Wei, L., Wang, Y., Lin, T., Huang, X., Yan, R. (2025). Life Prediction of On-Board Supercapacitor Energy Storage System Based on GRU Neural Network. *Applied Energy*, 379, 124917. <https://doi.org/10.1016/j.apenergy.2024.124917>
- [31] Ansari, S., Hannan, M.A., Ayob, A., Abdolrasol, M.G.M., Dar, M.A. (2025). Co-Estimation of State of Health and Remaining Useful Life for Lithium-Ion Batteries Using a Hybrid Optimized Framework. *Journal of Energy Storage*, 105, 114711. <https://doi.org/10.1016/j.est.2024.114711>
- [32] Xie, W., Zeng, Y. (2024). A Knowledge Distillation Based Cross-Modal Learning Framework for Lithium-Ion Battery State of Health Estimation. *Complex & Intelligent Systems*, 10(4), 5489-5511. <https://doi.org/10.1007/s40747-024-01458-4>
- [33] Kim, W.H., Chung, Y.M., Stinner, B. (2025). State of Health Prediction of Lithium-Ion Batteries with Adaptive Loss-Based Graph Neural Network. *Journal of Energy Storage*, 124, 116637. <https://doi.org/10.1016/j.est.2025.116637>
- [34] Han, X., Dai, Z., Ren, M., Cui, J., Shi, Y. (2024). One-Time Prediction of Battery Capacity Fade Curve Under Multiple Fast Charging Strategies. *Batteries*, 10(3), 74. <https://doi.org/10.3390/batteries10030074>
- [35] de la Vega, J., Riba, J.R., Ortega-Redondo, J.A. (2024). Real-Time Lithium Battery Aging Prediction Based on Capacity Estimation and Deep Learning Methods. *Batteries*, 10(1), 10. <https://doi.org/10.3390/batteries10010010>
- [36] Legala, A., Li, X. (2022). Hybrid Data-Based Modeling for the Prediction and Diagnostics of Li-Ion Battery Thermal Behaviors. *Energy and AI*, 10, 100194. <https://doi.org/10.1016/j.egyai.2022.100194>
- [37] Chen, Q., He, Y., Fang, N., Yu, G. (2024). A Combined Data-Driven and Model-Based Algorithm for Accurate Battery Thermal Runaway Warning. *Sensors*, 24(15), 4964. <https://doi.org/10.3390/s24154964>
- [38] Hong, J., Wang, Z., Liu, P. (2017). Big-Data-Based Thermal Runaway Prognosis of Battery Systems for Electric Vehicles. *Energies*, 10(7), 919. <https://doi.org/10.3390/en10070919>
- [39] Li, D., Zhang, Z., Liu, P., Wang, Z. (2019). DBSCAN-Based Thermal Runaway Diagnosis of Battery Systems for Electric Vehicles. *Energies*, 12(15), 2977. <https://doi.org/10.3390/en12152977>
- [40] Wang, Z., Hong, J., Liu, P., Zhang, L. (2017). Voltage Fault Diagnosis and Prognosis of Battery Systems Based on Entropy and Z-Score for Electric

Vehicles. *Applied Energy*, 196, 289-302.
<https://doi.org/10.1016/j.apenergy.2016.12.143>

- [41] Hong, J., Wang, Z., Yao, Y. (2019). Fault Prognosis of Battery System Based on Accurate Voltage Abnormity Prognosis Using LSTM. *Applied Energy*, 251, 113381. <https://doi.org/10.1016/j.apenergy.2019.113381>
- [42] Hong, J., Wang, Z., Chen, W., Yao, Y. (2019). Synchronous Multi-Parameter Prediction of Battery Systems Using LSTM Networks. *Applied Energy*, 254, 113648. <https://doi.org/10.1016/j.apenergy.2019.113648>