Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Non-Destructive Quality Assessment of Permanent Magnets inPedelec Drive Systems via Magnetic Field Mapping and SupportVector Machine Classification

Sebastian Feldmann

Sebastian.Feldmann@hs-aalen.de

Hochschule Aalen

Maximilian Lanz

Hochschule Aalen

Matthias Wiedenmann

Hochschule Aalen

Research Article

Keywords:

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10053810/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Non-Destructive Quality Assessment of Permanent Magnets in Pedelec Drive Systems via Magnetic Field Mapping and Support Vector Machine Classification

, Sebastian Feldmann[1,*], Maximilian Lanz[1], Matthias Wiedenmann[1]

[1]*Department of Mechanical Engineering and Material Science, Aalen University, Aalen, Germany*

*Corresponding author: Sebastian.Feldmann@HS-Aalen.de

Abstract

This paper presents a non-destructive methodology for quantifying *irreversible* permanent magnet degradation in permanent magnet synchronous motors (PMSMs) used in pedelec drive systems, using spatially resolved magnetic-field-mapping [MFM] combined with machine learning classification. Fourteen NdFeB magnets from a daum electronic Impulse2 e-bike drive system were characterized before and after 5,000 km of endurance testing on a biomechanical Hardware-in-the-Loop test bench at Aalen University. Spatially resolved magnetic field mapping was used to obtain flux density distributions for comparison of fully magnetized and degraded magnet states of the permanent magnets. The resulting difference field maps indicate small but measurable field deviations in the order of a few milliTesla (0.3 to 7.6 mT) across the magnet selection. While these deviations are modest relative to the nominal flux density of over 300 mT, the observed spatial patterns, concentrated at peripheral magnet regions, show localized demagnetization effects.

Twenty-five statistical, spatial, and texture features were extracted from each grayscale field map, including intensity distribution moments, gradient magnitude statistics, Gray-Level Co-Occurrence Matrix (GLCM) texture parameters, and percentile distributions. A comparative classification study employing five machine learning algorithms (SVM with linear and RBF kernels, Decision Trees, *k*-Nearest Neighbors, and Bagged Tree Ensembles) was evaluated using Leave-One-Out cross-validation. The results demonstrate that the proposed methodology is capable of detecting and classifying these small field deviations, establishing a baseline framework for non-destructive permanent magnet quality assessment.

Keywords: Magnetic Field Mapping · Permanent Magnet Degradation · Support Vector Machine (SVM) · Non-Destructive Testing (NDT) · Pedelec Drive System · Feature Extraction · GLCM Texture Analysis · Machine Learning Classification · NdFeB Magnets · Hysteresis Effect

1 Introduction

The global market for electrically assisted bicycles (PedEleCs – Pedal Electric Cycles) has experienced unprecedented growth over the past decade, with annual production volumes exceeding 1.1 million units in Germany alone and continued double-digit growth rates [1]. Unlike conventional bicycles, pedelecs operate as hybrid drive systems where the rider's pedalling effort is supplemented by an electric motor. The European legal framework mandates that motor assistance be proportional to the rider's torque input and limited to speeds below 25 km/h for standard pedelecs (250 W nominal power) [2]. This architecture relies critically on permanent magnets in central drive motors, typically permanent magnet synchronous motors (PMSMs), whose magnetic properties directly determine torque output, efficiency, and long-term reliability.

Neodymium-iron-boron (NdFeB) permanent magnets employed in pedelec drive motors are subjected to progressive degradation through interacting mechanisms during combined field and temperature exposure. When temperature and magnetic field act simultaneously, the resulting maximum conditions define a working point

that can lie either in the reversible region of the demagnetization curve, where mainly reversible changes occur, or in the irreversible region, where irreversible losses occur [3]. Under high operating conditions, particularly during high-load operation such as hill climbing or cargo transport, the resulting field and temperature loads can, especially if critically designed, shift the magnet's working point into such irreversible regions, leading to irreversible demagnetization of the installed permanent magnets that cannot be recovered without remagnetization by means of field pulsing [4].

Conventional assessment of magnet degradation relies on bulk magnetic moment measurements such as flux density, back-EMF characterization, or torque monitoring, which provide aggregate performance indicators but lack spatial resolution. These methods cannot distinguish between uniform decrease in magnetization across the entire magnet volume and localized field loss at specific regions such as edges, corners, or areas especially affected by field and temperature loads. This limitation is significant because localized degradation often leads to systematic failure and provides potential diagnostic information about the underlying degradation mechanism. Furthermore, early-stage degradation patterns may not yet affect measurable motor performance while already indicating incipient damage that will progress under continued service loading.

The present work addresses this gap through an integrated methodology combining high-resolution magnetic field mapping with machine learning classification. The objective is to investigate whether spatially resolved field characterization can detect early-stage degradation effects in permanent magnets subjected to realistic service loading, and whether the resulting field deviation patterns – even at the level of a few milliTesla – can be automatically classified using statistical feature extraction and Support Vector Machine (SVM) algorithms. As a first step, fourteen magnets from a daum electronic Impulse2 e-bike drive system were characterized after 5,000 km of biomechanical endurance testing at Aalen University [5], using a specific stepped load collective on a Hardware-in-the-Loop test bench. Field maps were acquired in both post-service and reconditioned states. While this initial test distance represents an early phase of the magnet's service life, it establishes the methodological framework and provides a baseline for planned follow-up tests at 20,000 km.

2 Experimental Setup

2.1 Test Specimen

The test subject was a daum electronic Impulse2 e-bike central drive unit, developed in collaboration between Aalen University and daum electronic GmbH, Fürth, Germany. The drive unit is built around a PMSM . The rotor assembly (Fig. 1) contains 14 NdFeB permanent magnets arranged in a radial configuration. Each magnet has approximate dimensions of 24 mm $\times$ 14 mm $\times$ 3 mm and shows a nominal peak surface flux density exceeding 300 mT in open-circuit configuration. The magnets are bonded to the rotor core using high-temperature adhesive and protected by epoxy coating against corrosion.

2.2 Biomechanical Endurance Testing

The endurance testing was performed on a biomechanical Hardware-in-the-Loop (HiL) test bench developed at Aalen University in cooperation with daum electronic GmbH [5]. The test bench employs torque-controlled actuation generating vehicle-like dynamics rather than pre-defined motion trajectories, preserving the natural coupling between rider effort, motor assistance, and road load. The biomechanical control concept generates physiologically realistic pedal torque profiles based on measurements by Petzke [6], including phase-shifted left-right actuation and negative recovery torques. The instantaneous pedal torque follows:

$$T_{\text{pedal},L}(\beta) = T_{\max} \cdot f_{\text{prop}}(\beta) - T_{\text{neg}} \cdot f_{\text{rec}}(\beta) \tag{1}$$

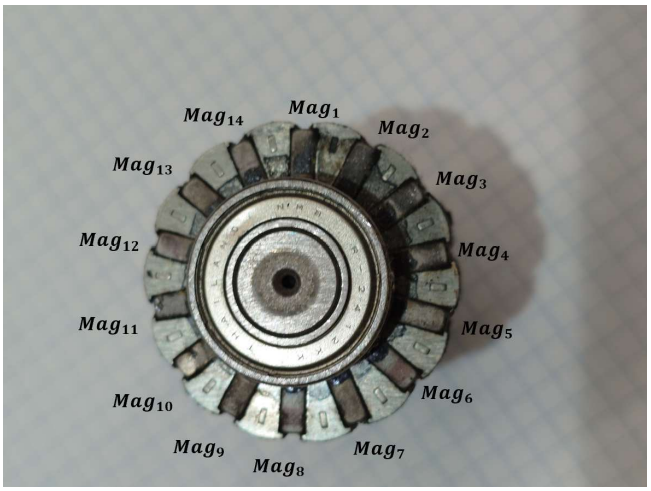


Figure 1 Rotor assembly of the daum electronic Impulse2 e-bike drive showing 14 NdFeB permanent magnets in radial configuration. Visible surface variations and adhesive residue are characteristic of the post-service condition as described in Sect. 2.2.

where β denotes the crank angle, $T_{\max}$ is the peak propulsion torque, T_{neg} is the recovery phase torque magnitude, and f_{prop}, f_{rec} are normalized propulsion and recovery shape functions. The right pedal torque follows an identical functional form with 180° phase offset: $T_{\text{pedal},R}(\beta) = T_{\text{pedal},L}(\beta - 180^\circ)$. The test protocol implemented a five-stage stepped load collective with a combined power output of 500 W (pedal + motor), cadences ranging from 50 to 100 rpm, and included thermal recovery phases between cycles. The total accumulated test distance was 5,000 km.

The endurance tests at the e-Bike test bench were conducted in 2020 by Prof. Feldmann and Matthias Wiedenmann. Detailed load collective parameters and drive system specifications are reported in the recently published work [5].

2.3 Magnetic Field Mapping System

Magnetic field characterization was performed using a SENIS MMS-1A-RS three-dimensional field mapper in two operational modes. For individual magnet characterization, the system operates in transition (area scan) mode, raster-scanning the Hall probe across the magnet surface to acquire spatially resolved field maps as $B(x,y)$. For full rotor characterization, the system operates in diameter mode, rotating the rotor while measuring at multiple axial positions. In both modes, the system provides three-axis measurement of the magnetic flux density vector $\mathbf{B} = (B_x, B_y, B_z)$ at each spatial location, enabling computation of the total flux density:

$$B_{\text{tot}} = \sqrt{B_x^2 + B_y^2 + B_z^2} \tag{2}$$

Table 1 summarizes the complete mapper configuration parameters used for all measurements in this study.

Each of the 14 magnets was scanned twice: once in the degraded, post-service condition after 5,000 km endurance testing and once after fully magnetizing the samples through pulse magnetization using a field pulser (M-K-O1020A, Magnet-Physik, Cologne, Germany), thereby establishing a reference baseline against which service-induced degradation is quantified.

Table 1 Magnetic field mapping scanner configuration parameters.

Parameter	Value
Gaussmeter	SENIS MMS-1A-RS
Hall Probe	SENIS AG #382-18
Measurement Range	1500 mT
Scanning Resolution	250 μm (at probe-to-sample distance)
Scan Area (XY)	24 mm $\times$ 14 mm per magnet
Measurement Mode	Transition mode (area scan)
Linear Speed	8 mm/s
Coordinate System	FSV tool
Low-Pass Filter	ON, cutoff 1500 mT
Probe Offset (X/Y/Z)	-0.223 / 0.02 / 0.065 mT

2.4 Measurement Principle

The fundamental measurement principle is illustrated in Fig. 2. For each magnet, the fully magnetized field map P_i and the degraded (post-service) field map Pa_i are compared. The magnetic difference image is computed as:

$$\mathrm{Mag}_{\mathrm{diff}} = (P_i - Pa_i) \cdot [\lambda_P] \tag{3}$$

where λ_P is an optional amplification factor for visualization. Through the difference image, the variations in the magnetic field after completing the test drive can be characterized. Measurement artifacts visible in the upper region of certain magnets (Fig. 2, right panel, circled) are attributed to probe positioning effects at the scan boundary (see text in Sect. 4.3) and are excluded from the quantitative analysis. The actual sensor to prope distance within the conducted experiments of the magnet field mapping was $\mathrm{Z} = 1.001\,mm$.

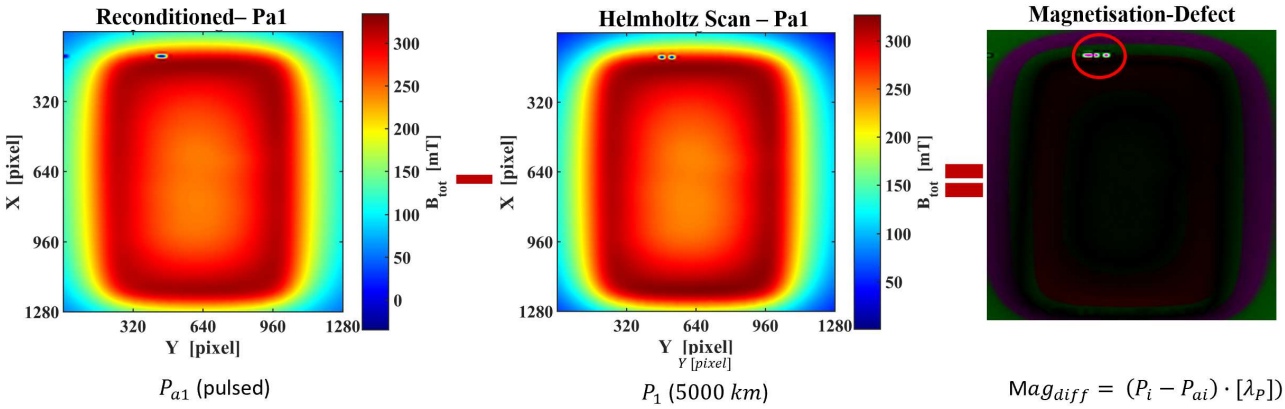


Figure 2 Measurement principle: 3D field map of fully magnetized magnet P_1 (left), degraded magnet Pa_1 (center), and RGB color difference image (right) showing the spatial distribution of field changes. The formula $\mathrm{Mag}_{\mathrm{diff}} = (P_i - Pa_i) \cdot [\lambda_P]$ defines the difference computation. Measurement artifacts at the scan boundary (red circle) are discussed in Sect. 4.3.

2.5 Full Rotor Characterization

In addition to the individual magnet area scans, a full 360° rotor characterization was performed using the diameter measurement mode of the SENIS MMS-1A-RS system. The rotor was rotated at 15 deg/s with an angular resolution of 0.1° while the Hall probe recorded three-axis flux density at a fixed radial distance of

20.95 mm from the rotation center. Thirteen axial slices were acquired at Z-positions ranging from -10.5 mm to $+1.5$ mm in approximately 1 mm increments, yielding 3,601 measurement points per slice and a total of 46,813 data points.

Figure 3 presents the resulting rotor field characterization as a combined three-panel visualization. Panel (a) shows the unwrapped cylindrical surface map with $B_{\mathrm{tot}}(\theta, Z)$ encoded in color, where the 14 magnet poles are clearly visible as periodic high-intensity (> 250 mT) regions separated by inter-magnet gaps with reduced field strength ($\sim$120–160 mT). The axial field distribution reveals that the highest flux density is concentrated in the central axial region. Panel (b) presents the same data in a polar cross-section representation, providing a top-down view of the rotor's magnetic field topology. Panel (c) shows circumferential B_{tot} profiles at five selected axial positions, demonstrating the periodic field modulation with 14 peaks corresponding to the magnet positions.

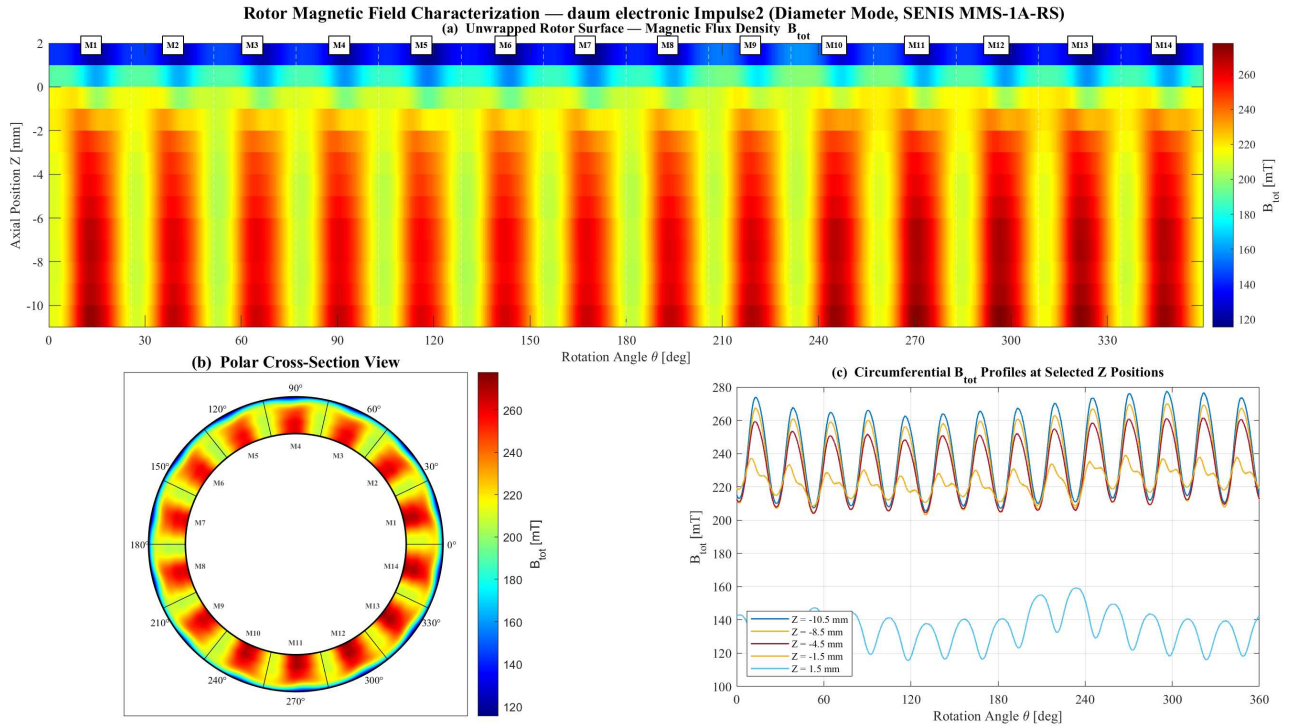


Figure 3 Full rotor magnetic field characterization using the SENIS MMS-1A-RS 3D field mapper in diameter mode. (a) Unwrapped cylindrical surface showing $B_{\mathrm{tot}}(\theta, Z)$ distribution across all 14 magnet poles and 13 axial Z-slices. (b) Polar cross-section view with axial position mapped to radial distance. (c) Circumferential B_{tot} profiles at selected axial positions. Dashed lines indicate magnet boundaries (M1–M14).

3 Field Mapping Results

3.1 Individual Field Map Analysis

Figure 4 presents a detailed view of the magnetic field distribution for a representative fully magnetized magnet (P_1) and its corresponding degraded counterpart (Pa_1). Both maps exhibit the characteristic field topology of a rectangular NdFeB magnet: a high-intensity central plateau ($B_{\mathrm{tot}} > 280$ mT) surrounded by a gradual roll-off toward the edges. The fully magnetized magnet shows a more uniform central plateau with sharper edge transitions, while the degraded magnet exhibits a slightly broadened transition zone and marginally reduced peak values, consistent with partial demagnetization during service.

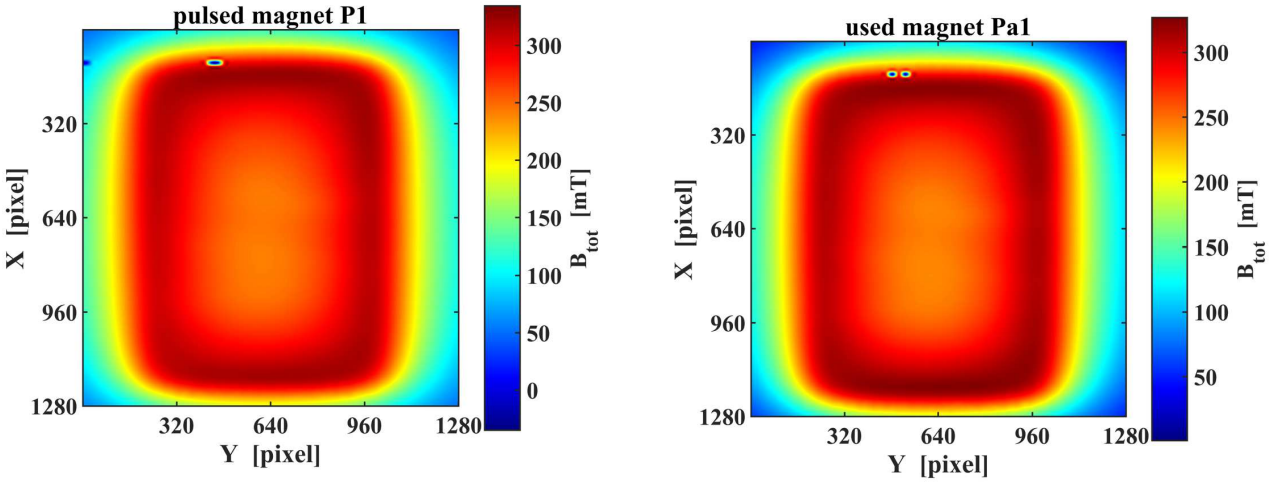


Figure 4 Detailed field maps of fully magnetized magnet P_1 (left) and degraded magnet Pa_1 (right) after 5,000 km endurance testing. Peak flux density exceeds 300 mT in the central region. Note the slightly broadened transition zone in the degraded magnet.

3.2 Overview of All Magnets

Figures 5 and 6 present the complete set of field maps for all 14 magnets in the fully magnetized and degraded conditions, respectively. The fully magnetized magnets show generally consistent field distributions with minor manufacturing variations visible in the edge geometry and peak intensity. The degraded magnets exhibit more heterogeneous patterns, with several magnets showing visually discernible differences from their fully magnetized counterparts.

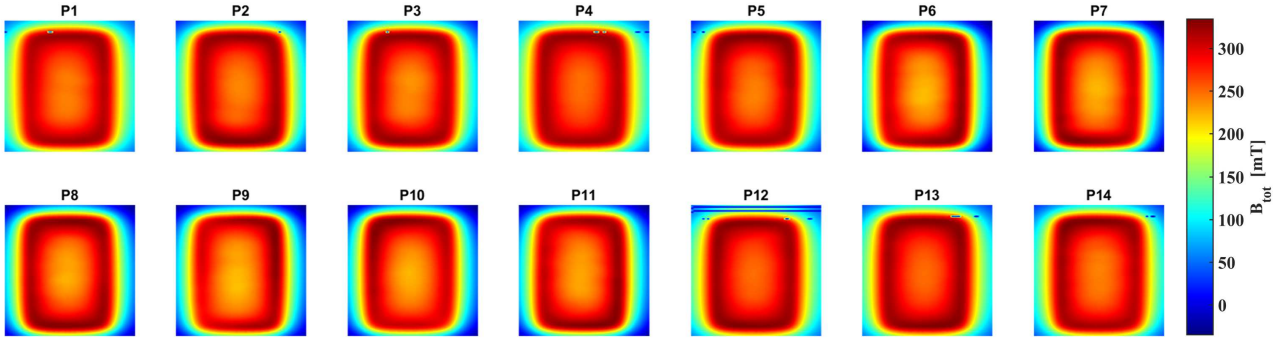


Figure 5 Grayscale-encoded field maps of all 14 fully magnetized magnets P_1–P_{14} of the Impulse2 e-bike drive.

4 Degradation Analysis

4.1 Field Map Construction Pipeline

The raw field map data comprises tab-delimited text files with 18 columns per measurement point, including spatial coordinates (X, Y, Z in mm) and flux density components B_x, B_y, B_z with derived total flux density B_{tot} (in mT). Field maps are constructed through the following processing pipeline: (1) raw data are loaded and filtered to remove NaN values and zero-field measurement points; (2) spatial coordinates are shifted to origin and scaled by factor 10 to achieve sub-millimeter pixel resolution; (3) a 2D matrix is populated with B_{tot} values at each integer-rounded spatial coordinate; (4) empty rows and columns are cropped; (5) the resulting field map

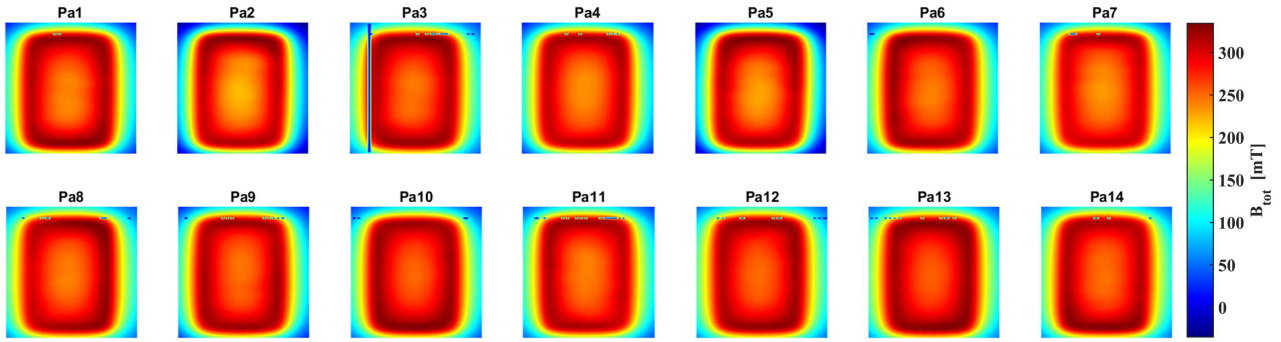


Figure 6 Grayscale-encoded field maps of all 14 degraded magnets Pa_1–Pa_{14} after 5,000 km endurance testing on the biomechanical e-bike test bench.

is resampled to a standardized 1280×1280 pixel resolution using bicubic interpolation.

4.2 Magnetic Field Difference Computation

The magnetic degradation for each magnet is computed as the element-wise difference:

$$\Delta B(x,y) = B_{\text{pulsed}}(x,y) - B_{\text{used}}(x,y) \tag{4}$$

Positive values of ΔB indicate field loss (degradation). The mean absolute degradation per magnet provides a scalar quality indicator:

$$\overline{\Delta B}(i) = \frac{1}{N} \sum_{(x,y)} |\Delta B(x,y)| \quad \text{for all valid } (x,y) \text{ in magnet } i \tag{5}$$

Figure 7 presents the magnetic field difference maps for all 14 magnets. Magnet 9 exhibits the largest field deviation (7.6 mT), followed by magnet 10 (4.9 mT). Magnet 3 shows an apparent ΔB of 7.4 mT; however, as discussed in Sect. 4.3, this value is attributable to a measurement artifact rather than an actual field change. It should be noted that these deviations are in the order of a few milliTesla and thus represent a small fraction of the nominal flux density (> 300 mT). Whether they indicate the onset of irreversible degradation or reflect measurement and reconditioning variability requires validation through extended endurance testing at higher accumulated distances.

4.3 RGB Color Difference Visualization

To enhance visual interpretability, the per-channel RGB difference between jet-colormap-encoded field images is computed:

$$\text{diff}_{\text{RGB}}(x,y,c) = \frac{|\text{img}_{\text{pulsed}}(x,y,c) - \text{img}_{\text{used}}(x,y,c)|}{\max(\ldots)} \quad c \in \{R,G,B\} \tag{6}$$

Figure 8 presents the normalized RGB color difference images for all 14 magnets. Critically, magnet 3 (top row, third panel) exhibits a prominent vertical straight-line artifact spanning the entire image height, which is characteristic of a sensor discontinuity or probe positioning error during scanning rather than a physical degradation pattern. This measurement error invalidates the high $|\Delta B|$ value (7.4 mT) reported for magnet 3 and excludes it from the defective classification. The remaining magnets with genuine high degradation – particularly M_9, which shows pronounced red/green edge brightening with asymmetric patterns – can be reliably identified as defective.

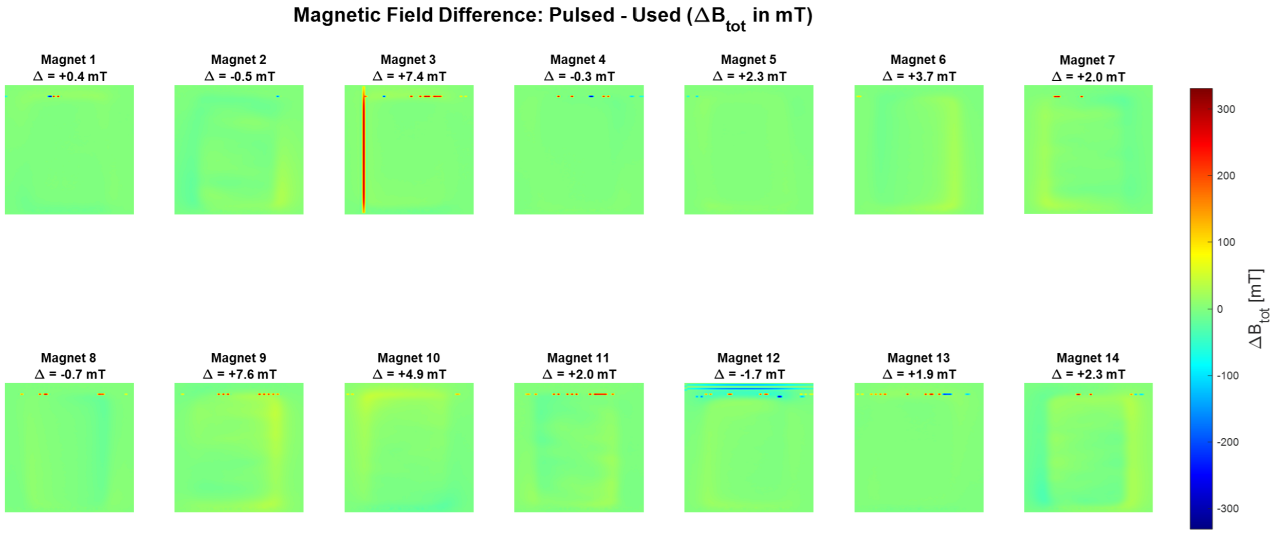


Figure 7 Absolute Magnetic field difference maps: $P_i - Pa_i$ (i.e. pulsed minus degraded, ΔB_{tot} in mT) for all 14 magnets. Mean degradation values are annotated for each magnet. Global colorbar range: -300 to $+300$ mT. Notable straight-line artifacts in magnets 3, 7, 9, 11 and 13 are identified as measurement artifacts and excluded from further analysis (see Sect. 4.3).

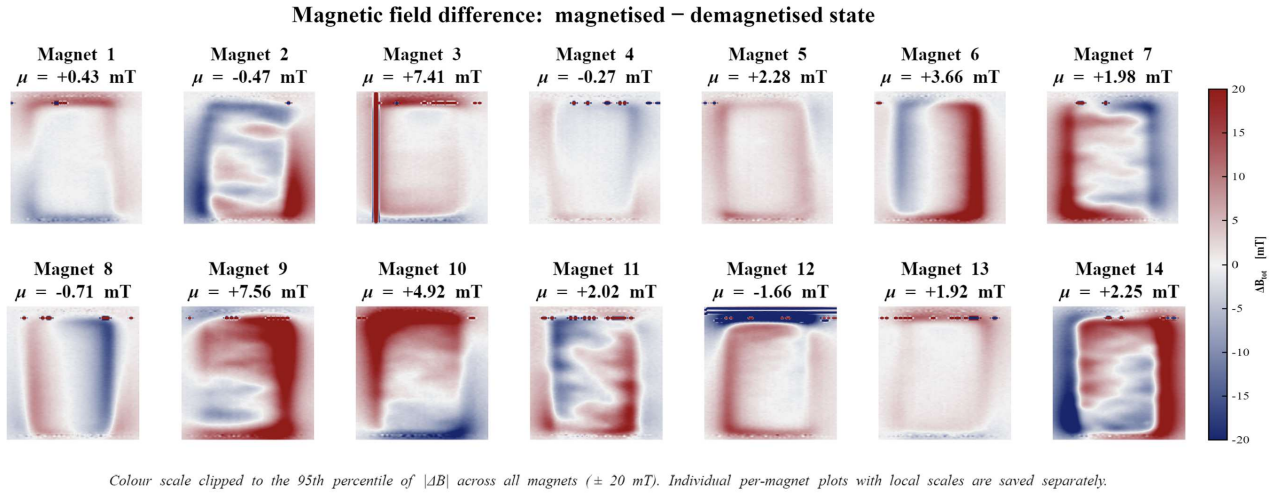


Figure 8 Normalized RGB color difference $|P_i - Pa_i|$ for all 14 magnets. Bright peripheral regions indicate areas of significant field change. The vertical line artifact in magnet 3 indicating a measurement error.

4.4 Degradation Ranking and Distribution

Figure 9 presents the degradation ranking sorted from highest to lowest mean absolute field difference. While magnet 3 appears at rank 2 with a nominal ΔB of 7.4 mT, this value is affected by the measurement artifact identified in Sect. 4.3. The genuinely most degraded magnet is magnet 9 (7.6 mT), followed by magnet 10 (4.9 mT) and magnet 6 (3.7 mT).

Figure 10 shows the overlay probability distribution of ΔB for all 14 magnets. Magnet 9 exhibits the broadest distribution with significant tails extending beyond 100 mT. Magnet 3 also shows a broad distribution; however, this is primarily caused by the measurement artifact rather than genuine degradation.

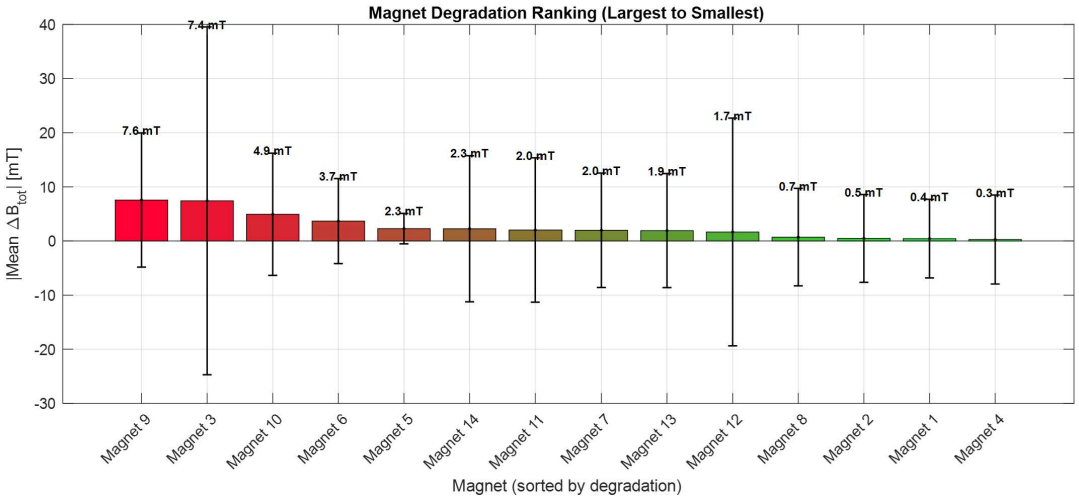


Figure 9 Magnet degradation ranking (largest to smallest mean $|\Delta B_{\mathrm{tot}}|$) with standard deviation error bars. Red bars indicate magnets classified as defective; green bars indicate acceptable quality.

5 Feature Extraction and SVM Classification

Support Vector Machines (SVMs) are a class of supervised machine learning algorithms introduced by Vapnik [7] for binary classification tasks. The fundamental idea is to find an optimal separating hyperplane in a potentially high-dimensional feature space that maximizes the geometric margin between two classes. Given a set of labeled training samples, the SVM identifies the subset of data points closest to the decision boundary – the so-called *support vectors* – and constructs the hyperplane that achieves the largest possible separation between them. For data that is not linearly separable in the original input space, SVMs employ the *kernel trick*: a kernel function implicitly maps the input features into a higher-dimensional space where linear separation becomes feasible, without explicitly computing the transformation. Common kernel functions include the linear, polynomial, and Radial Basis Function (RBF) kernels. A key advantage of SVMs for the present application is their strong generalization performance on small datasets. Unlike neural networks, which require large amounts of training data to avoid overfitting, SVMs are governed by the principle of *structural risk minimization*, which balances training accuracy against model complexity. This makes them particularly well-suited for non-destructive evaluation scenarios where acquiring labeled samples is expensive and the number of available specimens is inherently limited.

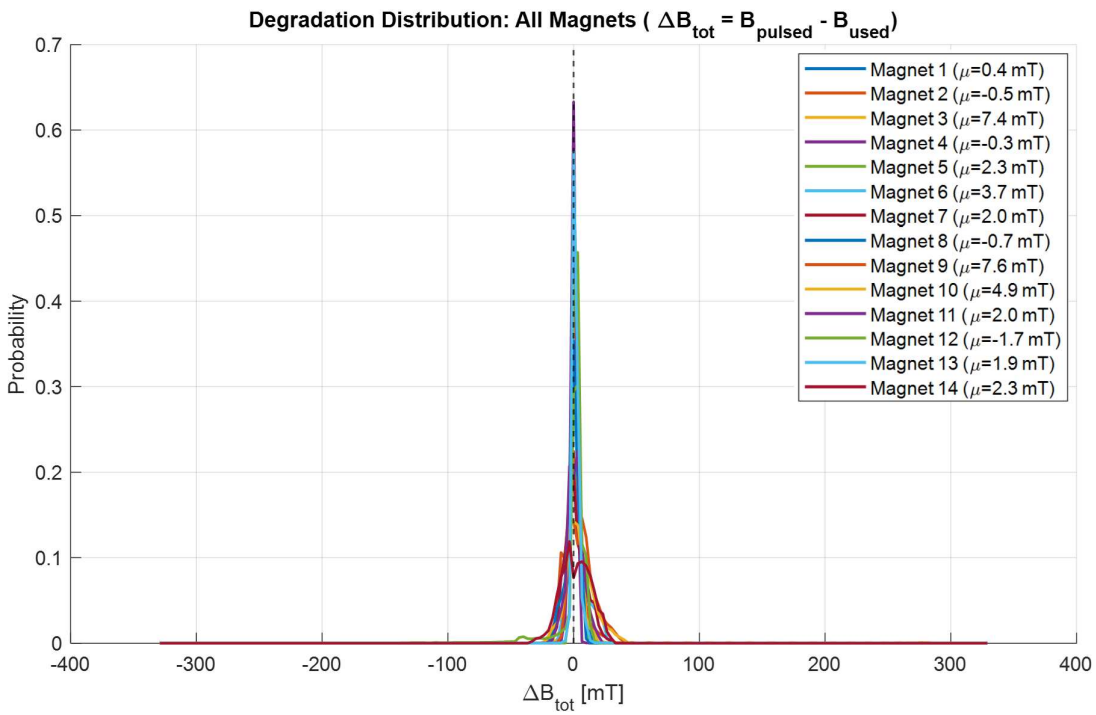


Figure 10 Degradation distribution: overlay probability histograms of $\Delta B_{\text{tot}} = B_{\text{pulsed}} - B_{\text{used}}$ for all 14 magnets. Mean degradation values (μ) are shown in the legend.

5.1 Motivation

An initial attempt was made to classify magnets using deep learning, specifically a SqueezeNet-based convolutional neural network (CNN) trained at 1280×1280 pixel resolution with aggressive data augmentation and transfer learning. Despite these measures, the CNN achieved only 50% validation accuracy after 200 training epochs, equivalent to random guessing. With a total dataset of only 28 images (14×2 states), the network memorized training samples but failed to learn generalizable features. This fundamental limitation of deep learning for small datasets motivates the feature engineering approach presented below.

5.2 Feature Extraction

Twenty-five physically meaningful features are extracted from each grayscale field map, organized into four categories.

Statistical features comprise the first four central moments:

$$\mu = \frac{1}{N}\sum_i x_i \tag{7}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_i (x_i - \mu)^2} \tag{8}$$

$$\gamma_1 = \frac{1}{N}\sum_i \left(\frac{x_i - \mu}{\sigma}\right)^3 \tag{9}$$

$$\gamma_2 = \frac{1}{N}\sum_i \left(\frac{x_i - \mu}{\sigma}\right)^4 \tag{10}$$

where Eq. (7) is the mean intensity, Eq. (8) the standard deviation (variance), Eq. (9) the skewness, and Eq. (10)

the kurtosis. These are complemented by median, intensity range, and percentile values at the 5th, 25th, 75th, and 95th levels plus the interquartile range (IQR).

Spatial gradient features quantify field non-uniformity through the image gradient magnitude:

$$|\nabla B| = \sqrt{\left(\frac{\partial B}{\partial x}\right)^2 + \left(\frac{\partial B}{\partial y}\right)^2} \tag{11}$$

Texture features are derived from the Gray-Level Co-Occurrence Matrix (GLCM) [8]:

$$\text{Contrast} = \sum_{i,j} (i-j)^2 \, p(i,j) \tag{12}$$

$$\text{Correlation} = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)\, p(i,j)}{\sigma_i \sigma_j} \tag{13}$$

$$\text{Homogeneity} = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \tag{14}$$

$$\text{Energy}_{\text{GLCM}} = \sum_{i,j} p(i,j)^2 \tag{15}$$

5.3 Feature Normalization

All 25 features are z-score normalized to zero mean and unit variance:

$$x_{\text{norm}} = \frac{x - \mu_x}{\sigma_x} \tag{16}$$

5.4 Support Vector Machine Formulation

The binary classification task is formulated as the SVM optimization problem [7]:

$$\min_{\mathbf{w},b} \; \frac{1}{2}\|\mathbf{w}\|^2 + C \sum_i \max\bigl(0,\, 1 - y_i(\mathbf{w}^\top \mathbf{x}_i + b)\bigr) \tag{17}$$

where $\mathbf{w}$ is the weight vector, b is the bias term, and C is the box constraint. The Radial Basis Function (RBF) kernel is employed:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma_k^2}\right) \tag{18}$$

5.5 Comparative Classifier Study

To evaluate the prediction performance, four additional classifiers are evaluated: SVM with linear kernel, Decision Tree with maximum 5 splits, k-Nearest Neighbors with $k = 3$, and Bagged Tree Ensemble with 50 bootstrap-aggregated learners [9]. All classifiers are evaluated using Leave-One-Out (LOO) cross-validation:

$$\text{ACC}_{\text{LOO}} = \frac{1}{N} \sum_i \mathbb{I}(\hat{y}_i = y_i) \tag{19}$$

5.6 Feature Importance Assessment

Feature importance is quantified using the mean decrease in Gini impurity:

$$\text{Imp}(f) = \frac{1}{N_{\text{trees}}} \sum_{\text{trees}} \sum_{\text{nodes}} \Delta G(f, \text{node}) \tag{20}$$

6 Results and Discussion

6.1 Classification Accuracy

The primary goal of the machine learning approach was to determine whether the 25-feature representation contains sufficient discriminative information to separate magnets exhibiting measurable field deviations from those with negligible changes. Figure 11 shows that *all five classifiers achieve 100% LOO accuracy* (28/28 correct, 0.0% error rate) on the present dataset.

While this result is encouraging, it must be interpreted with caution given the limited dataset size (28 samples) and the relatively small magnitude of the observed field deviations (order of a few milliTesla). The current 5,000 km endurance test represents only an initial load cycle with a specific biomechanical load profile, and the question of whether the detected deviations constitute the onset of progressive degradation or fall within the range of measurement variability can only be answered conclusively through extended testing. Nevertheless, the unanimous classification across five fundamentally different classifier architectures suggests that the extracted features reflects systematic rather than random differences between the two magnet states.

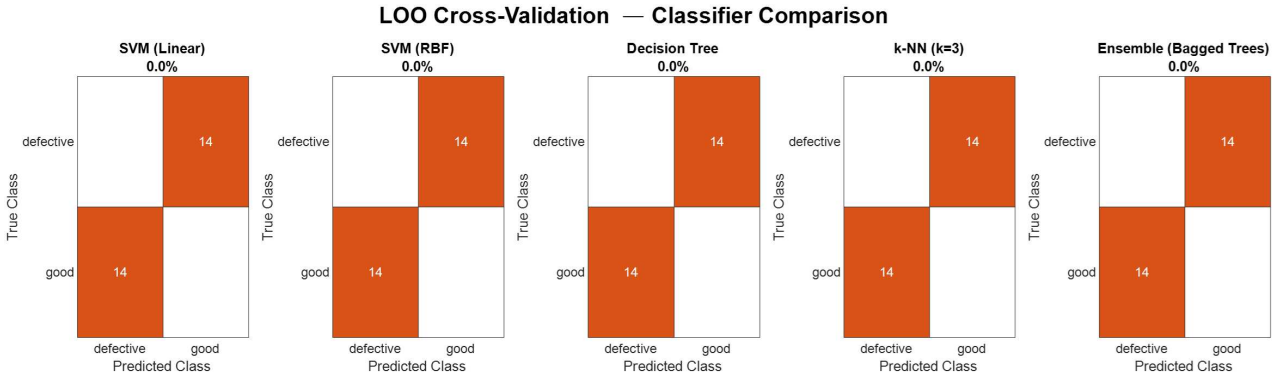


Figure 11 LOO cross-validation confusion matrices for all five classifiers. All algorithms correctly separate the 14 magnets with elevated field deviations from the 14 magnets with negligible changes (100% accuracy, 0.0% error).

6.2 Feature Importance Analysis

Figure 12 presents the feature importance ranking obtained from the Bagged Tree Ensemble using the mean-decrease-in-Gini criterion of Eq. (20). The top five features, in descending order of importance, are: (1) kurtosis, (2) energy (sum of squares), (3) edge pixel ratio, (4) GLCM correlation, and (5) maximum gradient magnitude.

Several observations follow from this ranking. First, the highest-ranked feature – kurtosis – is a fourth-order moment of the intensity distribution that quantifies the "peakedness" of the field histogram relative to a Gaussian. A reduction in kurtosis between the fully magnetized and degraded states indicates that the field distribution broadens during service: the sharp central plateau loses uniformity and the histogram develops heavier tails. This is consistent with the visual observation in Sect. 3.1 that the degraded magnets exhibit a broadened transition zone between the central plateau and the edges.

Second, the strong ranking of energy and edge pixel ratio indicates that degradation manifests not only in the magnitude of the field but also in its spatial concentration. Energy decreases when the field is no longer concentrated in a uniform plateau, and the edge pixel ratio captures the proportion of high-gradient pixels, which grows as the plateau-to-edge transition softens.

Third, the appearance of GLCM correlation among the top five features confirms that texture – i.e. the spatial relationship between neighbouring pixels – carries discriminative information that is not contained in the

intensity moments alone. This supports the choice of combining statistical, gradient, and texture descriptors in the 25-feature representation rather than relying exclusively on distribution moments.

Conversely, the lowest-ranked features are the range, minimum intensity, and maximum intensity, which all describe single extreme values of the distribution and are therefore highly sensitive to measurement noise and outliers. Their low importance confirms that distributed, texture-based descriptors are more robust than single-point intensity extrema in the present application.

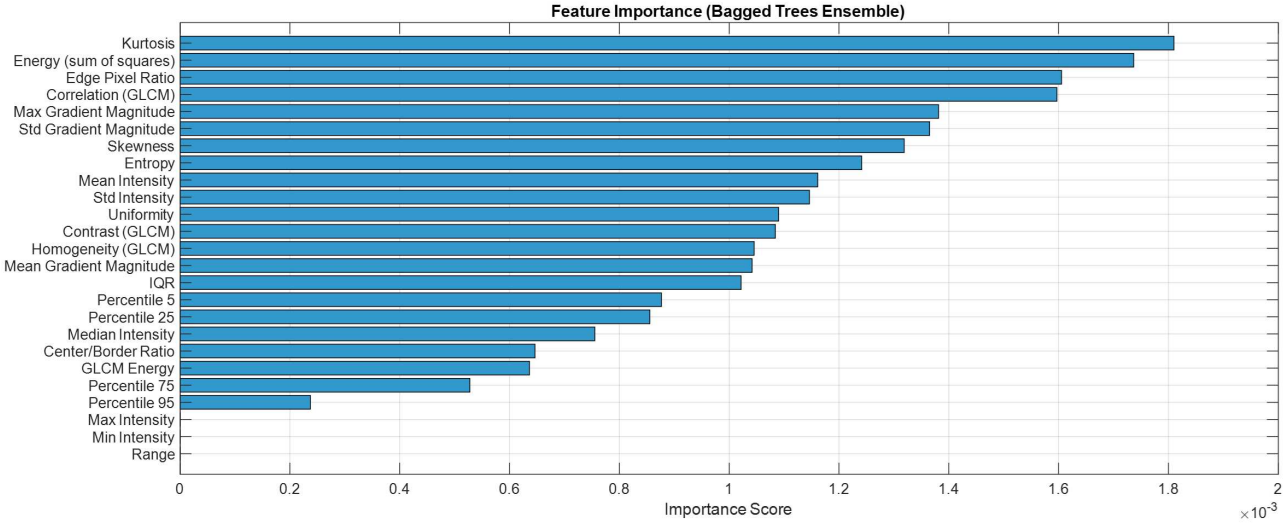


Figure 12 Feature importance ranking from the Bagged Tree Ensemble. Kurtosis, Energy, Edge Pixel Ratio, and GLCM Correlation are the most discriminative features.

6.3 Feature Space Visualization

Figure 13 visualizes the 28 samples projected onto the two most important features (kurtosis vs. energy). Although kurtosis and energy are individually the strongest discriminators (Sect. 6.2), the 2D projection alone does not yield a linearly separable configuration: the "defective" and "good" populations form partially overlapping clusters with no obvious decision boundary.

This overlap is the central justification for using a high-dimensional feature representation. Each of the 25 features captures a partially independent aspect of the field map: distribution moments, spatial gradients, GLCM-based texture, and percentile descriptors. The information that separates degraded from fully magnetized magnets is distributed across these dimensions rather than concentrated in any single one.

The successful classification by all five algorithms (Sect. 6.1) confirms that, while the projection onto the top two features appears ambiguous, the full 25-dimensional representation contains sufficient discriminative geometry for both linear (SVM linear, Decision Tree) and non-linear (SVM RBF, k-NN, Bagged Trees) decision surfaces to achieve perfect separation. The visualization in Fig. 13 therefore serves not as a classifier in itself but as a sanity check that the chosen features are non-trivially related to the underlying magnet state – and as a visual argument for why low-dimensional projections of this dataset must be interpreted with caution.

6.4 Classification Summary

Figure 14 consolidates the analysis into a degradation space scatter plot ($|\overline{\Delta B}|$ vs. $\sigma_{\Delta B}$) together with the measurement and classification parameters. Each magnet is represented by a single point whose coordinates encode the two quantities that, in combination, best discriminate the two states within an interpretable physical

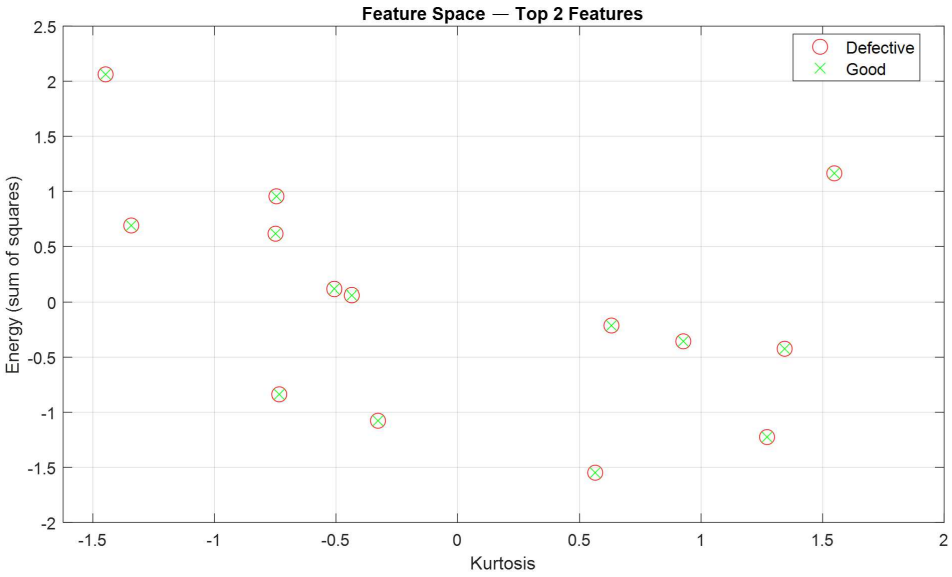


Figure 13 Feature space visualization: top 2 features (Kurtosis vs. Energy) for all 28 samples. Red circles: defective; green crosses: good. Overlap in 2D confirms the need for multi-dimensional classification.

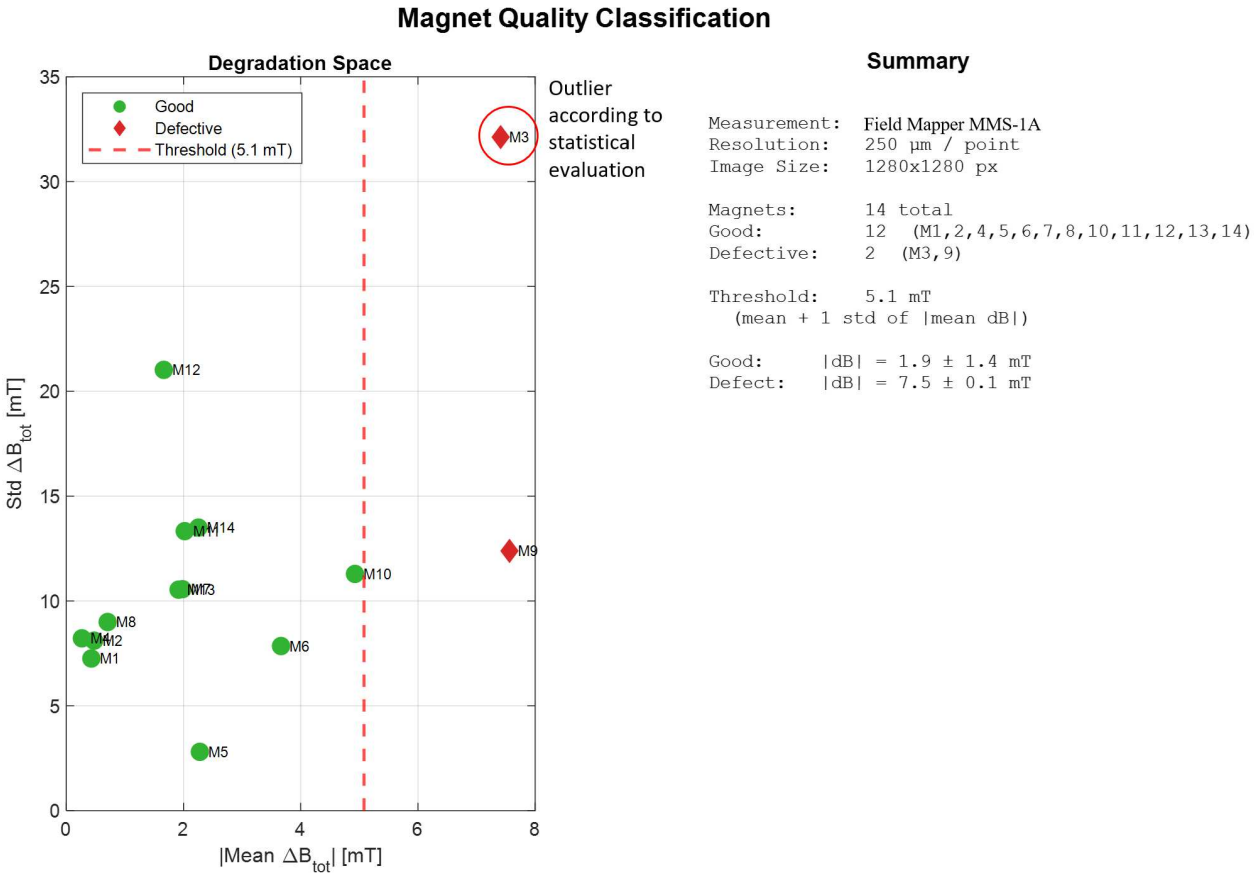


Figure 14 Magnet quality classification summary: degradation space scatter plot (left) and measurement/classification parameters (right). Green markers: good magnets; red markers: defective magnets.

space: the mean absolute field deviation (severity of degradation) and the standard deviation of the deviation (spatial heterogeneity of degradation).

A simple statistical threshold at 5.1 mT (mean + 1 std of $|\overline{\Delta B}|$) cleanly separates the two classes: 12 of the 14 magnets fall below the threshold and are classified as "good", while two magnets (M_3 and M_9) exceed it. As discussed in Sect. 4.3, the M_3 outlier is attributable to a measurement artifact rather than physical degradation, leaving M_9 as the only magnet exhibiting genuine elevated field loss (7.5 ± 0.1 mT). The "good" population shows a much lower mean degradation of 1.9 ± 1.4 mT, well within the range expected from measurement and reconditioning variability.

This degradation space is the most compact representation of the present results and serves two complementary purposes. First, it provides a direct visual quality-control criterion that can be applied to individual magnets without invoking the full 25-feature machine learning pipeline. Second, it makes the threshold-based decision auditable: future measurements on additional magnet populations can be plotted into the same coordinate system, and the threshold itself can be refined as more data become available.

6.5 Comparison with Deep Learning

The feature-based approach achieves 100% accuracy on the present dataset with sub-second training times, while the CNN achieves only random-chance performance (50.0%) despite 30 minutes of training with aggressive data augmentation and transfer learning.

Table 2 summarizes the comparative performance.

Table 2 Comparative classification performance: feature-based classifiers vs. deep learning CNN.

Method	**LOO Accuracy**	**Training Time**
SqueezeNet CNN (1280×1280)	50.0%	$\sim$30 min
SVM (Linear)	100.0%	$<$1 sec
SVM (RBF)	100.0%	$<$1 sec
Decision Tree	100.0%	$<$1 sec
k-NN (k=3)	100.0%	$<$1 sec
Bagged Trees (50)	100.0%	$<$2 sec

This contrast is consistent with established results on small-sample machine learning: deep convolutional networks require training datasets that are typically several orders of magnitude larger than the 28 samples available here. The high parameter count of CNN architectures (SqueezeNet has approximately 1.2 million parameters) creates a strong tendency to memorize training samples rather than to extract generalizable features, even when techniques such as transfer learning, dropout, and image augmentation are applied. The 50% validation accuracy reflects this failure mode: the network fits the training data but learns nothing that transfers to unseen samples.

A direct consequence is that the present approach is well suited to the typical conditions of industrial non-destructive testing, where acquiring labelled samples is expensive and the number of available specimens is inherently limited. The deep-learning baseline, by contrast, would only become competitive once the dataset has grown by two to three orders of magnitude – a milestone that may eventually be reached through the planned 20,000 km endurance campaigns and through inclusion of additional motor variants, but is well beyond the current measurement effort.

It should be emphasised that this comparison does not constitute a general statement against deep learning for magnetic field characterisation. With sufficient training data, end-to-end CNN approaches may well outperform

hand-crafted features, particularly if the field maps exhibit subtle degradation patterns that are not captured by the present 25-feature design. The conclusion is the narrower but practically important one that, under small-sample conditions, domain-specific feature engineering remains the more reliable methodology.

7 Conclusion

This paper presented a non-destructive methodology for detecting and classifying field deviations in permanent magnets of electric motors in pedelec drive systems using spatially resolved magnetic field mapping combined with Support Vector Machine classification. The key findings are as follows:

First, spatially resolved magnetic field mapping is capable of detecting small field deviations in NdFeB permanent magnets after 5,000 km of biomechanical endurance testing with a specific stepped load profile. The element-wise field difference between fully magnetized and degraded states reveals deviations in the order of a few milliTesla (0.3 to 7.6 mT), which are small relative to the nominal flux density of over 300 mT but exhibit spatially consistent patterns concentrated at peripheral magnet regions. The initially high ΔB value of magnet 3 (7.4 mT) was identified as a measurement artifact through RGB color difference analysis. Whether the remaining deviations – most notably in magnet 9 (7.6 mT) – represent the onset of irreversible demagnetization or fall within the bounds of measurement and reconditioning variability requires confirmation through extended endurance testing.

Second, a feature-based classification pipeline extracting 25 statistical, gradient, and texture features from field maps achieves 100% leave-one-out cross-validation accuracy across all five evaluated classifiers on the present dataset. While the small sample size (28 images) and modest magnitude of the detected deviations warrant caution in generalizing this result, the consistent classification across five fundamentally different algorithms suggests that the features capture systematic differences rather than noise. The approach significantly outperforms a deep learning baseline (50% accuracy with SqueezeNet CNN), confirming that domain-specific feature engineering is more appropriate than end-to-end learning for small-sample NDT scenarios.

Third, feature importance analysis identifies kurtosis, energy, edge pixel ratio, and GLCM correlation as the most informative features, suggesting that the detected deviations manifest primarily through changes in the intensity distribution shape and spatial texture rather than uniform intensity shifts.

The present study establishes a methodological baseline for non-destructive permanent magnet quality assessment. The current endurance test of 5,000 km with a single load profile represents an initial investigation; follow-up tests at 20,000 km are planned to determine whether the observed field deviations intensify with continued service loading and to validate the robustness of the classification approach under extended degradation conditions. Future work will also extend the approach to larger sets of permanent magnets from different motor types, investigate correlations between spatial deviation patterns and root-cause mechanisms, and explore integration into predictive maintenance frameworks for pedelec fleet management.

Acknowledgements

This work was conducted within the MagTwin sub-project of the SmartPro research alliance at Aalen University. MagTwin — dedicated to the investigation of performance losses in permanent magnets — is led by the corresponding author and forms part of the DiMa project under the scientific direction of Prof. Dr.-Ing. Dagmar Goll. The authors gratefully acknowledge the project framework for providing access to the magnetic field mapping infrastructure and the methodological background on NdFeB permanent magnet degradation that underpin the present study.

This work was also conducted in cooperation with daum electronic GmbH, Fürth, Germany. The biomechanical test bench concept and initial developments are based on the research of Prof. Tillmann Körner of Aalen University. The authors thank Jonas Huggenberger, Özben Sarikurt, Moritz Paeslack, and Jonas Seitz for their contributions to test bench development, instrumentation, and experimental execution. Special thanks to the workshop team, especially Günter Link, at Aalen University for mechanical fabrication support.

Author Contributions

S.F. conceived the study, designed the biomechanical test bench, and supervised the project. M.L. performed the magnetic field measurements and data acquisition. M.W. conducted the endurance testing and contributed to data analysis. All authors contributed to manuscript preparation and review.

Declarations

Funding. This work was carried out within the MagTwin sub-project of the SmartPro research alliance and the DiMa project at Aalen University. No specific external grant numbers are reported by the authors for this study.

Competing Interests. The authors declare that they have no competing financial or non-financial interests that are relevant to the content of this article. The study was conducted in cooperation with daum electronic GmbH, Fürth, Germany, which supported the biomechanical test bench development; this cooperation did not influence the design, analysis, or interpretation of the results presented here.

Ethics approval. Not applicable.

Data Availability

The complete measurement dataset and all MATLAB scripts used to produce the figures in this paper are publicly available in the `BioBike` GitHub repository:

`https://github.com/Prof-Feldmann/BioBike`

The repository contains the raw MFM datasets provided as text files for all 14 permanent magnets in both fully magnetized and degraded states (`Feldmann_aufmagnetisiert_1...14.txt` and `Feldmann2_1...14.txt`), the full 360° rotor scan (`Rotor_slice.txt`), three MATLAB scripts that reproduce all field-mapping and degradation-analysis figures, and a `README.md` documenting the file formats and processing pipeline. The dataset is released under the CC0–1.0 license; the scripts are released under the MIT license.

References

[1] Zweirad-Industrie-Verband (ZIV). Market data reports. https://www.ziv-zweirad.de/marktdaten/, 2024.

[2] European Union. Regulation (EU) no 168/2013 on the approval and market surveillance of two- or three-wheel vehicles and quadricycles, 2013.

[3] R. Hilzinger and W. Rodewald. Magnetic materials. 1, 2013.

[4] J. M. D. Coey. *Magnetism and Magnetic Materials*. Cambridge University Press, 2010.

[5] S. Feldmann and M. Wiedenmann. Biomechanical hardware-in-the-loop control concept for pedelec drivetrain testing. *Computer Science & Information Technology (CSIT)*, 15, 2025. URL http://aircconline.com/csit/papers/vol15/csit152506.pdf.

[6] W. Petzke. Muskelleistung und wirkungsgrad beim radfahren, 2006. Bundesinstitut für Sportwissenschaft, ISSN: 0341-7387.

[7] V. N. Vapnik. *The Nature of Statistical Learning Theory*. Springer, New York, 1995.

[8] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.*, 3(6):610–621, 1973.

[9] L. Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, 1996.