mathematics

MDPI

Article

Direct Parallel–Hierarchical Transformation Based on Evolutionary Development Models: Application to Laser Beam Spot Image Processing with a Prospective Extension to Railway Track Infrastructure Maintenance

Volodymyr Tverdomed [1,*], Natalia Kokriatska [1], Leonid Tymchenko [1], Yurii Didenko [1], Mykhailo Braikovskyi [1], Mariia Demchenko [1], Vaidas Lukoševičius [2,*], Ramūnas Skvireckas [2] and Rolandas Makaras [2]

[1] Educational and Research Kyiv Institute of Railway Transport, National Transport University, Mykhaila Omelianovycha—Pavlenka Str. 1, 01010 Kyiv, Ukraine; n.kokriatska@ntu.edu.ua (N.K.); l.tymchenko@ntu.edu.ua (L.T.); didenko_yuv@ntu.edu.ua (Y.D.); m.braikovskyi@ntu.edu.ua (M.B.); m.demchenko@ntu.edu.ua (M.D.)

[2] Department of Transport Engineering, Faculty of Mechanical Engineering and Design, Kaunas University of Technology, Studentų Str. 56, 51424 Kaunas, Lithuania; ramunas.skvireckas@ktu.lt (R.S.); rolandas.makaras@ktu.lt (R.M.)

* Correspondence: v.tverdomed@ntu.edu.ua (V.T.); vaidas.lukosevicius@ktu.lt (V.L.)

Abstract

Signal processing systems operating in noisy or dynamically changing environments require computational architectures capable of multilevel parallel information processing. Biological sensory systems have naturally evolved such architectures through hierarchical organization, providing inspiration for efficient computational models. This paper presents a mathematical model of the direct parallel–hierarchical transformation (PHT), its FPGA implementation, and an application to laser-beam spot processing. The original experiment yielded an average coordinate-determination error of 0.47%, a maximum error of 1.6%, and a standard deviation of 0.36%. The 269-frame tr0_269 dataset constitutes one sequentially acquired frame sequence; its 70/199 division is therefore treated as an internal functional division within that sequence rather than as a separation of independent laser tracks. For the subsequent 199-frame segment, the complete modified PHT chain gave a mean coordinate-determination error of 0.567 pixels; the repeated calculation gave 0.569 pixels, supporting reproducibility within this frame sequence. For an additional check, separate completed frame sequences were used without overlap: TRASSA00 and FRAME02-FRAME05 were used for implementation checks and parameter definition, whereas FRAME06-FRAME11 were not used for parameter selection and were used only for subsequent evaluation of the method. The FMT headers show that FRAME02-FRAME11 are separate completed acquisition sequences with individual start and end times; the gaps between successive sequences substantially exceed the 40 ms inter-frame period, excluding their interpretation as contiguous fragments of one continuously stored FMT file. No sequence and no frame contributed to both groups. The 38-section method produced a lower mean coordinate-determination error than the six-section method in all six evaluation sequences, with means across the six sequences of 0.815 and 1.321 pixels, respectively (Wilcoxon statistic 0; one-sided p-value 0.015625; 95% confidence interval from whole-sequence resampling, -1.178 to -0.152 pixels). In an additional evaluation of the complete PHT on FRAME06-FRAME11, the shape-coefficient limits were determined only from TRASSA00 and FRAME02-FRAME05 as the central 95% empirical interval [0.637, 0.862] and then applied unchanged to FRAME06-FRAME11. On the same evaluation frames, the mean same-frame error of the complete PHT was 0.961 pixels versus 0.822 pixels for the

Check for updates

Academic Editor: Renato Bruni

Received: 17 August 2026
Revised: 12 October 2026
Accepted: 19 October 2026
Published: 20 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

direct 38-section stage. The complete method produced a larger same-frame error in all six sequences because its acceptance logic can retain the preceding accepted coordinate. The FPGA implementation reached 7 ms/frame with register memory and 25 ms/frame with serial memory.

Keywords: parallel–hierarchical transformation; signal processing; laser beam spot classification; FPGA implementation; railway infrastructure maintenance

MSC: 68W10; 94A12; 68U10

1. Introduction

Signal processing systems must increasingly contend with large volumes of spatially and temporally structured data, demanding computational architectures that efficiently exploit the inherent hierarchical and parallel organization of such signals [1]. Biological sensory systems—auditory, visual, olfactory, and others—have evolved multilevel filtering mechanisms that perform convergence and divergence of signal features across hierarchical processing layers, achieving robust perception under noisy, high-dimensional conditions [2–4]. Formalizing these mechanisms as mathematical models provides a principled basis for developing artificial signal processing systems with similar properties.

The concept of parallel–hierarchical transformation (PHT) captures this biological organization within a computational framework: input signal fields are decomposed level by level, with each level forming sets of common and distinctive features that propagate to the next hierarchical stage. This structure corresponds closely to the modular cortical architecture described by Mountcastle [5] and the hierarchical sensory analysis documented across different sensory modalities [2–4]. While deep neural networks have demonstrated strong empirical performance on hierarchical signal processing tasks [6,7], their connection to the underlying biological principles of hierarchical organization remains largely implicit. Previous studies on PHT [8] established its theoretical foundations, whereas the present work extends these studies through a complete mathematical formalization and an FPGA-based implementation evaluated on a challenging real-world task.

Previously published laser-spot localization approaches include contour- or shape-based center estimation [9] and optical-image processing methods [8]. Their processing principles differ from the direct parallel–hierarchical transformation developed here; robustness under atmospheric disturbances must therefore be assessed experimentally rather than assumed. The proposed parallel–hierarchical transformation addresses these limitations through multilevel decomposition of spatial features and efficient parallel processing.

Recent studies have also demonstrated the effectiveness of diagnostic parameter extraction and mathematical modeling for improving the accuracy of technical condition assessment in engineering systems [10,11].

Beyond optical signal processing, parallel–hierarchical transformation has significant potential for processing heterogeneous diagnostic information in railway track infrastructure monitoring and maintenance. Modern railway inspection systems integrate laser scanning, machine vision [12,13], ultrasonic testing, eddy-current inspection, inertial measurement units, and track geometry measurement systems [14], generating large volumes of multidimensional data that require efficient hierarchical feature extraction and intelligent data fusion. However, the integration of heterogeneous diagnostic data into a unified computational framework capable of real-time hierarchical analysis remains a challenging problem for railway infrastructure maintenance [15]. Owing to its multilevel organization and computational efficiency, the proposed method could in principle be adapted

for analyzing multisource diagnostic information, in support of rail defect identification, infrastructure condition assessment, and condition-based maintenance [16] within digital railway asset management systems [17]. This application is outlined here as motivation only; it is not evaluated experimentally in the present study, which validates PHT on laser beam spot images.

The aim of this work is to develop and validate a direct parallel–hierarchical transformation method for the classification and coordinate determination of laser beam spot images. The principal contributions are: (I) a complete mathematical model of direct PHT incorporating matrix decomposition operators; (II) an FPGA-based implementation achieving processing speeds of up to 7 ms/frame; (III) an experimentally demonstrated reduction in coordinate-prediction error relative to the baseline contour-based method [8] used in the original comparison, for all five forecasting techniques tested, averaging 17.6% and ranging from 2.3% to 32.2%; (IV) evaluation on separate completed acquisition sequences, with TRASSA00 and FRAME02-FRAME05 used for implementation checks and parameter definition and FRAME06-FRAME11 used only for subsequent evaluation; (V) a separate evaluation of the complete PHT on FRAME06-FRAME11 using shape-coefficient limits determined only from TRASSA00 and FRAME02-FRAME05 and then applied unchanged; and (VI) comparative evaluation against two baseline coordinate-determination methods under a unified per-frame protocol.

Existing approaches to the classification and coordinate determination of laser beam spot images rely predominantly on contour-based processing, which treats the image as a single, flat feature space. Such approaches contain no explicit parallel–hierarchical structure and lose accuracy when the spot is deformed by atmospheric distortion, because turbulence-induced local variations are not separated from the stable spatial features carrying the position of the energy center. A transformation is therefore required that filters the input field level by level, progressively separating stable from variable components. The direct parallel–hierarchical transformation (PHT) developed in this work provides such a mechanism: it performs a level-by-level decomposition of the input numerical fields and forms the output results as tail elements. Its justification follows from the principles of hierarchy and parallelism observed in models of evolutionary development, above all in neurobiological sensory systems.

The research reported here aims to formalize this hierarchical and parallel organization as a mathematical model and to validate it on the classification and coordinate determination of laser beam spot images implemented on an FPGA platform. The accomplishment of the objective involved resolving the following tasks:

- Describe examples of hierarchy and parallelism in models of evolutionary development.
- Formulate a mathematical model of direct parallel–hierarchical transformation.
- Show the effectiveness of the model based on practical results.

2. Theoretical Framework of the Direct Parallel–Hierarchical Transformation

The mechanism of the evolutionary development model is directly observable in the formation of the auditory, visual, olfactory, somatosensory, gustatory, and vestibular representations by which man perceives the outside world [2–5,18], supplemented by signals from within the body reporting limb position, temperature, chemistry and blood volume [2,3]. All forms of sensation carry information about the state of the space-time environment (STE): where and when a stimulus appeared and how long it lasted. Comparing different sets of states, the perceptual system first determines the states common to them—formed from features such as brightness, color, timbre, pitch, taste, temperature and pressure—and then the distinctive states, as the segments that differ. Perception may

therefore be represented as an iterative process of parallel spatio-temporal comparison of signal parameters and selection from them of common and distinctive features–signals, each new set of distinctive features generating, through evolutionary change, a new set of common features; it proceeds continuously and simultaneously in all systems of the human senses. The common features–signals are conveyed to the cerebral cortex along independent pathways, where generalized filtering forms new sets, each of which characterizes a new hierarchical level of sensory processing. Images of the outside world are constructed at the later stages of sensory analysis by combining the most purified data passed through the filters of the individual sensory systems; synthesis proceeds in reverse order, a new set being formed at each level from the "reflections" of features–signals held in memory. The filtering is of a growing nature in some sensory systems and of a declining one in others, but in both cases convergence and divergence filter and amplify the information, which acts on the neurons of a single hierarchical level—a neural local area network retaining the information flow within that level. All parts of sensory systems include at least stimulus detectors, that is, specialized receptor neurons; a primary receiving center at which information from a group of detector blocks converges; and secondary and integrating centers fed by the primary centers. In more complex systems, the integrating centers, which distinguish the more general feature sets, are also interconnected, and their interaction creates thinking and consciousness. The formation and evolutionary change in the distinctive features–signals is the sign for their new selection at the next hierarchical level [2–5], and the speed of their formation in the cortex determines the human capacity for analysis. Thinking and awareness may thus be represented as parallel, interconnected processing of sensory information across sensory zones of increasing hierarchical order, the highest of which form the most common and the most distinctive feature sets.

In this regard, this approach coincides with the fundamental concept of the American physiologist W. Mountcastle [5]. According to Mountcastle's theory, the cerebral cortex consists of complex multicellular ensembles whose basic unit is formed by about a hundred vertically connected neurons of all cortical layers. Several such simple vertical ensembles can be combined by means of intercolumn connections into a larger information-processing unit—a module, or modular column, which acts as a unit that processes information [5]. Modules are grouped into larger aggregates, which create the primary visual or auditory cortex, or the cortex responsible for movement, depending on where the aggregate predominantly receives information from. Information processing is conducted in parallel channels: each aggregate performing one main function also contains smaller subgroups of vertical units associated with subgroups of other aggregates performing other functions, so that any subgroup of modules may belong to different distributed systems depending on the information currently received. The algorithm of such a network is universal: it consists in the parallel–hierarchical formation of sets of common and distinctive features of the state of STE, generalization between different types of sensory information occurring only after the network has been constructed for a given set of features.

The detector-based and spatial-frequency descriptions of signals provide complementary perspectives on hierarchical image processing. In optical applications, laser-beam images can be analyzed through explicitly structured processing stages; prior work on laser-beam image processing and modified parallel–hierarchical transformations provides the relevant technical background [8]. The present study applies these hierarchical processing principles specifically to laser-beam spot images.

Formally, such a model is a distributed generalized hierarchical network consisting of a finite set of hierarchical subnets. Each subnet reflects a development process of a strictly defined nature, the independent vertices of the subnets form independent branches of

the generalized network, and the whole set of subnets merges into a single hierarchical network, the structure of which is presented in Figures 1 and 2.

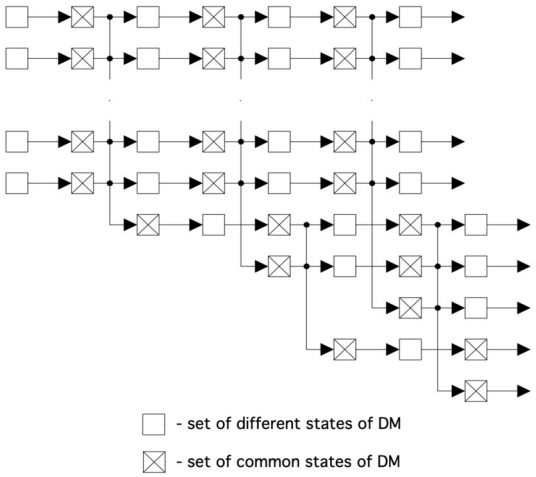


Figure 1. Structure of parallel–hierarchical network.

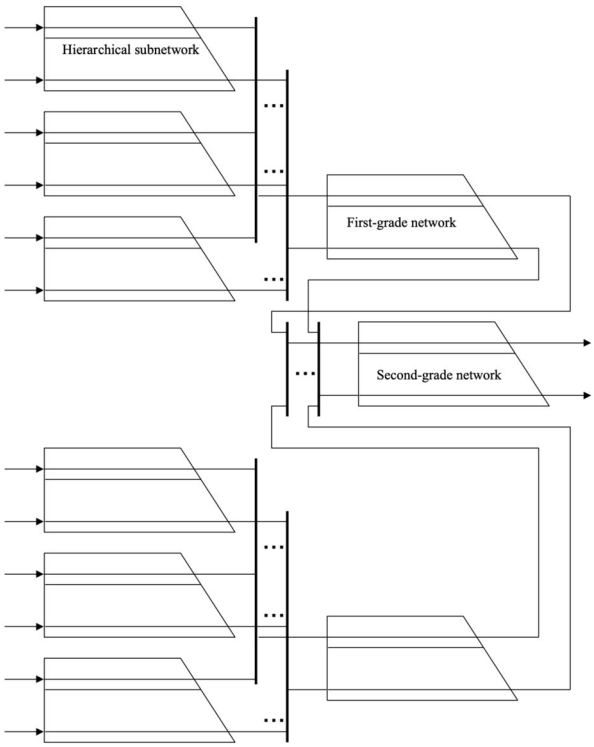


Figure 2. Formation of a multistage parallel–hierarchical network from individual parallel–hierarchical networks.

Figure 1 illustrates the mechanism of connection organization during the formation of an individual parallel–hierarchical network. The first three levels are shown, each containing parallel branches that perform the transformation. Figure 2 extends this principle to a set of hierarchical subnetworks that are successively combined into a single multistage parallel–hierarchical network. This multilevel interaction can be viewed in the broader context of interconnected complex systems [19], while the specific PHT operators are defined below. Within each individual network, information expands at the first level and is compressed at the subsequent levels; upon transition to the next network, expansion becomes possible again. Thus, the "expansion–compression" cycle is repeated as succes-

sive networks are involved in the processing. This alternation is the structural property formalized by the direct parallel–hierarchical transformation.

We use the following mathematical model of parallel set decomposition to formalize the direct PHT; its conceptual basis is described in earlier work on parallel–hierarchical transformations [8]: $\gamma = \{m_i\}, i = \overline{1,n}$

$$\gamma = \bigcup_{i=1}^{n} m_i = \bigcup_{j=1}^{R} \left(n - \sum_{k=0}^{j-1} n_k \right) \left(m^j - m^{j-1} \right) \tag{1}$$

The transformation defined by model (1) is treated here as a G-transformation operator. Its formal definition is given in Equation (2).

$$G(\gamma) = \gamma^1 \tag{2}$$

Applying the operator defined by Equation (2) to each source array gives the ordered decomposition presented in Equation (3).

$$\gamma_1^1 = \bigcup_{i=1}^{R_1^1} m_{1i}^1, \quad \gamma_2^1 = \bigcup_{i=1}^{R_2^1} m_{2i}^1, \quad \ldots, \quad \gamma_S^1 = \bigcup_{i=1}^{R_S^1} m_{Si}^1 \tag{3}$$

Equation (3) defines the source arrays of the first level and their elements. For an arbitrary hierarchical level, the level number and the number of elements in the corresponding array are specified separately. Combining the decomposition results of the source arrays forms the first-level matrix presented in Equation (4).

$$M_1 = \begin{pmatrix} m_{11}^1 & m_{12}^1 & \ldots & m_{1R_1^1}^1 & X & \ldots & X \\ m_{21}^1 & m_{22}^1 & \ldots & \ldots & \ldots & m_{2R_2^1}^1 & X \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ m_{S1}^1 & m_{S2}^1 & \ldots & m_{SR_S^1}^1 & X & \ldots & X \end{pmatrix} \tag{4}$$

or

$$M_1 = \bigcup_{s=1}^{S} \left\{ m_{si}^1 \right\}_{i=1}^{R_s^1} \tag{5}$$

The elements of the first-level matrix are then regrouped by columns. The resulting matrix form is presented in Equation (6).

$$M_1^T = \bigcup_{i=1}^{R^1} \left\{ m_{si}^1 \right\}_{s=1}^{S} \tag{6}$$

The transition to the transposed matrix is defined by the transposition operator in Equation (7).

$$T(M_1) = M_1^T \tag{7}$$

Because the source sets may contain different numbers of elements, the maximum set length determines the number of groups formed at the next stage of the transformation.

The elements of the transposed matrix pass to the next hierarchical level and provide the input for constructing the second-level matrix. Successive application of the G-transformation gives the relation presented in Equation (8).

$$\gamma_2 = \bigcup_{i=1}^{R^1} \gamma^2_{R^1_i} = \bigcup_{i=1}^{R^1} \left\{ m^2_{R^1_i i} \right\} \tag{8}$$

$$M_2 = \begin{pmatrix} m^2_{11} & m^2_{12} & \dots & m^2_{1R^2_1} & X & \dots & X \\ m^2_{21} & m^2_{22} & \dots & m^2_{2R^2_2} & X & \dots & X \\ m^2_{31} & m^2_{32} & \dots & \dots & \dots & m^2_{3R^2_3} & X \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ m^2_{R^1 1} & m^2_{R^1 2} & \dots & m^2_{R^1 R^2_{R^1}} & X & \dots & X \end{pmatrix} \tag{9}$$

Starting from the second level, matrix formation is not performed by transposition alone. The elements are also grouped diagonally, which provides the inter-level continuation of the transformation.

For construction of the next level, the second-level matrix is represented in the form shown in Equation (9).

$$M_2\prime = \begin{pmatrix} m^2_{11} & m^2_{12} & m^2_{13} & \dots & m^2_{2R^2_1} & X & \dots & X \\ X & m^2_{21} & m^2_{22} & m^2_{23} & \dots & m^2_{2R^2_2} & X & X \\ X & X & m^2_{31} & m^2_{32} & \dots & m^2_{3R^2_3} & X & X \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X & \dots & X & m^2_{R^1 1} & m^2_{R^1 2} & \dots & m^2_{R^1 R^2_{R^1}} & X \end{pmatrix} \tag{10}$$

After the structural transformation, the next matrix is formed from the preceding matrix without the first selected element; the result is given in Equation (10).

The first selected element is the first output result of the transformation. Starting from the second level, each subsequent level produces the next output element.

The transition between levels is performed by the structural row-shift operator. The rows are shifted by the required number of positions, after which the first column is excluded from further inter-level formation.

After the structural row shift, transposition is performed. The matrix of the next level is then obtained as shown in Equation (11).

$$M_2^T = \begin{pmatrix} m^2_{11} & m^2_{21} & X & \dots & \dots & X \\ m^2_{13} & m^2_{22} & m^2_{31} & X & \dots & X \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ m^2_{R^1 R^2_{R^1}} & X & \dots & \dots & \dots & X \end{pmatrix} \tag{11}$$

The resulting matrix provides the input for the next level. The same operations are repeated until only one resulting element remains; after that, no further non-empty matrix is formed. This sequence of operations defines the direct parallel–hierarchical transformation.

The notation in Equations (1)–(11) specifies the input numerical field, its elements, the number of input elements, the index and number of G-transformation steps, the number of elements at the current step, the selected value, the number of source arrays, the hierarchical-level index, the matrix elements and dimensions, transposition, and the structural shift. An empty padding position is not a data value.

Let us consider the theoretical justification of the PHT-output correlation threshold. For each newly processed frame, the output PHT vectors of the four most recent previously accepted frames are retained and averaged to form the current reference PHT vector. The

current-frame PHT vector and the averaged reference vector are normalized, after which the correlation coefficient P between them is calculated. A new coordinate is accepted only when P is not less than 0.95; otherwise, the previously accepted coordinate is retained. Equations (12)–(16) define vector averaging, correlation, the angle between the normalized vectors, and the corresponding Euclidean distance. The distance-based interpretation therefore means expressing the same correlation threshold of 0.95 as an allowable Euclidean distance between the two normalized PHT vectors.

$$T_{ref,n} = \frac{T_{k1} + T_{k2} + T_{k3} + T_{k4}}{4} \tag{12}$$

$$\rho_n \geq P \tag{13}$$

$$\rho_n = {u_n}^T v_n = \cos\theta_n \tag{14}$$

$${d_n}^2 = 2 - 2\rho_n \tag{15}$$

$$\cos\theta_{max} = 0.95; \theta_{max} = 18.19^\circ, d_{max} = \sqrt{0.10} = 0.316 \tag{16}$$

Equations (12)–(16) describe the four most recent accepted PHT vectors, their averaging, normalization of the current and reference vectors, their correlation coefficient, the angle between the normalized vectors, and the corresponding Euclidean distance. The acceptance threshold is 0.95.

Noise-model interpretation. Equations (17)–(20) represent the current and reference PHT vectors as a stable structural component plus noise and relate the correlation threshold of 0.95 to the corresponding noise-to-structure variance ratio. This model is used only to interpret the fixed correlation threshold; it does not replace the experimental reliability check.

$$T_n = S + \varepsilon_n, T_{ref,n} = S + \varepsilon_{ref,n} \tag{17}$$

$$q = \frac{{\sigma_\varepsilon}^2}{{\sigma_S}^2}, {\sigma_n}^2 = {\sigma_S}^2 + {\sigma_\varepsilon}^2, {\sigma_{ref}}^2 = {\sigma_S}^2 + \frac{{\sigma_\varepsilon}^2}{4} \tag{18}$$

$${\rho_{pop}}^{-2} \approx 1 + \frac{5}{4}q + \frac{1}{4}q^2 \tag{19}$$

$$P = 0.95; q \approx 0.08498, SNR = -10\log_{10}q \approx 10.71\ dB \tag{20}$$

Equations (17)–(20) represent the current and reference PHT vectors as the sum of a stable structural component and a noise component. The model relates noise variance, structural-component variance, correlation, and signal-to-noise ratio. It is used only for theoretical interpretation of the fixed correlation threshold and does not replace the experimental evaluation.

Thus, the correlation threshold of 0.95 has a direct geometric meaning: at the acceptance boundary, the normalized vectors differ by an angle of approximately 18.2°, and their normalized Euclidean distance is approximately 0.316. Under the stated additive-noise model, the same threshold corresponds to a noise-to-structure variance ratio of approximately 0.085.

3. Results

Let us consider one possible application of parallel–hierarchical transformation for coordinate calculation and classification of laser-beam spot images, examples of which are shown in Figure 3. The 269 BMP images in tr0_269 form one sequence of consecutively acquired frames. In the original processing scheme, the first 70 frames were used at the preliminary stage to form the initial processing characteristics required by the algorithm, whereas the following 199 frames were used for evaluation after those characteristics had

been formed. Therefore, in this article the 70/199 division is treated as an internal functional division of one frame sequence, not as a separation of independent laser tracks. Additional completed acquisition sequences are used separately for implementation checks and for comparison of the six-section and 38-section methods.

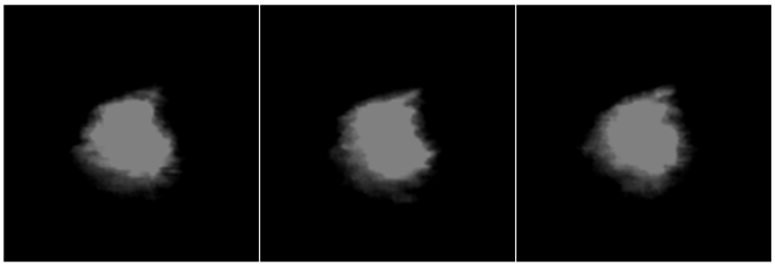


Figure 3. Images of 128 × 128 pixel laser path spots.

Origin of the sectioning and shape-coefficient criterion. The brightness-section stage follows the method of sections described in [8]. For each selected intensity section, its area and center are determined; the relation between areas at different intensity levels characterizes the spot shape. Equations (21) and (22) give the shape coefficients used in this procedure.

$$r_s = \frac{S_{0.75h}}{S_{0.50h}} \tag{21}$$

$$r_v \approx \frac{4r_s + 7}{12} \tag{22}$$

Coordinate calculation from brightness sections is performed as follows. For each selected brightness level, the pixels belonging to the corresponding cumulative section are identified, and the center of that section is calculated using Equation (23). The final coordinate of the energy center of the laser-beam spot is obtained by averaging the section-center coordinates according to Equation (24). Using six sections gives the original method; using 38 sections gives the modified direct method.

$$A_k x_k = \sum_{p=1}^{A_k} x_{k,p}, A_k y_k = \sum_{p=1}^{A_k} y_{k,p} \tag{23}$$

$$N x_N = \sum_{k=0}^{N-1} x_k, N y_N = \sum_{k=0}^{N-1} y_k \tag{24}$$

Equations (21)–(24) define the maximum frame intensity, the area of a cumulative brightness section at a selected threshold, the shape coefficients, the section index, the number and coordinates of pixels in a section, and the section-center coordinates. The number of sections is specified separately, and the final coordinate is obtained by averaging the centers of the selected sections. The six-section variant uses six successive brightness levels; the 38-section variant uses the same first six levels plus 32 additional successive levels.

Let us consider the construction of the computational reference and its limitations. For each frame, the reference coordinates are obtained by the same fixed image-processing procedure. First, the minimum and maximum intensities are determined, and the binary-segmentation threshold is calculated using Equation (25). Pixels satisfying the condition in Equation (26) form the binary region. Spatial moments of this region are then calculated, and Equation (27) gives the computational reference coordinates of the energy center of the laser-beam spot. These coordinates provide a common computational reference for the six-section, 38-section, and complete PHT methods, but they are not an external metrological measurement of the true physical beam position. The result may be affected by threshold choice, intensity quantization, saturation, background gradients, asymmetric or multimodal spot structure, and truncation at the image boundary. A sensitivity analysis

was therefore performed by decreasing and increasing the coefficient in Equation (25) by 5%. Relative to the nominal reference, the mean coordinate shifts for individual sequences were 0.114–0.401 pixels for the decreased coefficient and 0.120–0.415 pixels for the increased coefficient; the means across the six analyzed sequences were 0.178 and 0.181 pixels, respectively. The mean 38-section error across the six sequences was 0.851, 0.815, and 0.788 pixels for the decreased, nominal, and increased coefficients, respectively. Thus, the numerical values change slightly, but the conclusion that the 38-section method has a lower error than the six-section method is unchanged under this plus or minus 5% coefficient variation.

$$I_c = I_{min} + \frac{I_{max} - I_{min}}{\sqrt{e}} \quad (25)$$

$$b(x,y) = 1 \text{ if } I(x,y) > I_c;\ b(x,y) = 0 \text{ otherwise} \quad (26)$$

$$M_{pq} = \Sigma x^p y^q b(x,y), x_r = \frac{M_{10}}{M_{00}}, y_r = H - 1 - \frac{M_{01}}{M_{00}} \quad (27)$$

For each evaluated frame, the coordinate-determination error of the energy center of the laser-beam spot is calculated using Equation (28) as the Euclidean distance between the coordinate returned by the method and the computational reference coordinate for the same frame. The error is reported in pixels and as a percentage of the 128-pixel frame width. It therefore measures agreement with the common computational reference rather than the absolute metrological error of the physical laser-beam position.

$$e_i = \sqrt{{\Delta x_i}^2 + {\Delta y_i}^2}, E_i = \frac{100 e_i}{128} \quad (28)$$

The generalized scheme of the developed device is presented in Figure 4.

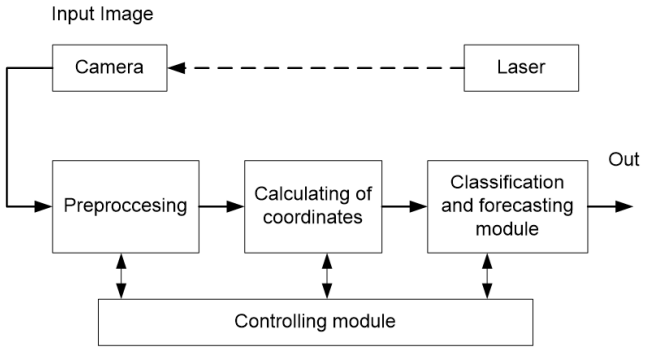


Figure 4. Generalized block diagram of image processing of laser beam spots.

Equations (25)–(28) use the minimum and maximum frame intensities, the binary-segmentation threshold, the intensity of an individual pixel, and the resulting binary mask. The summation in Equation (27) extends over all image pixels, and the zeroth-order spatial moment equals the number of pixels in the binary region, whereas the two first-order spatial moments determine the horizontal and vertical coordinates of its center. The image height is 128 rows. Equation (27) accounts for the adopted direction of the vertical axis. Equation (28) compares the computational reference coordinates with those returned by the method and gives the error for the current frame.

Figure 5 illustrates the algorithm used to calculate the coordinates and categorize laser beam spots within a single frame, employing the direct parallel–hierarchical transformation method [8].

The DE1-SoC board served as the foundation for conducting the simulation. The processing speed of the complete circuitry was evaluated using actual laser-beam spot images at a fundamental operating frequency of approximately 100 MHz:

- Up to 25 ms per frame when serial memory was used;
- Up to 7 ms per frame when register memory was used; an incomplete memory matrix was used because of the limited FPGA resources (Figure 6).

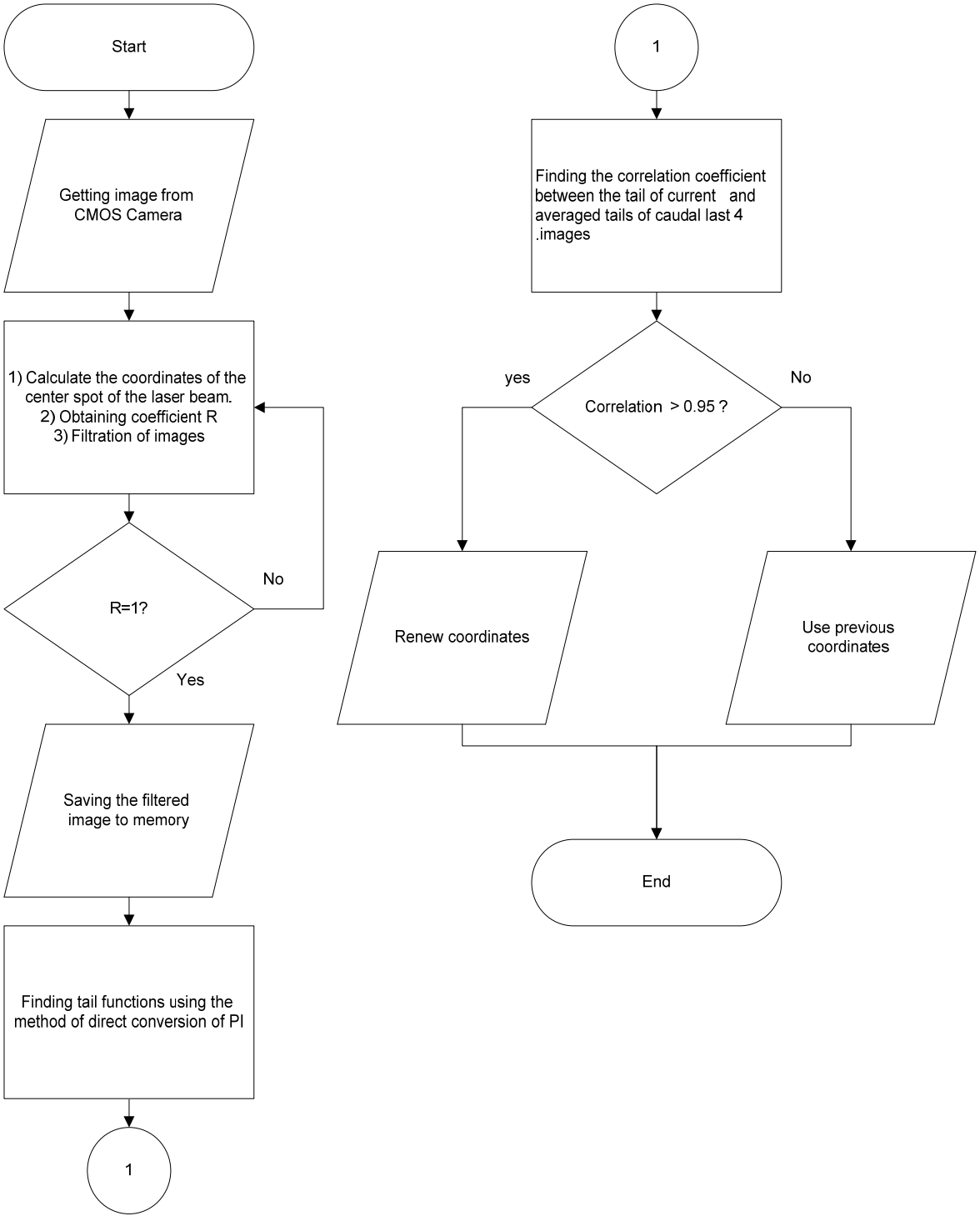
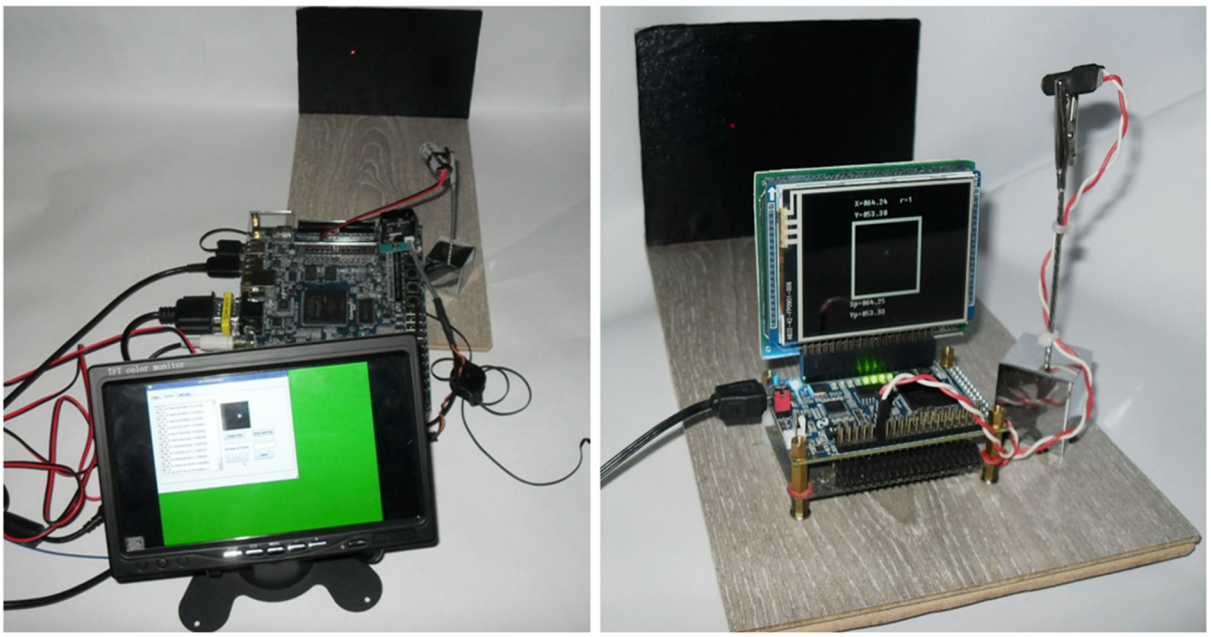


Figure 5. The combined approach algorithm for determining coordinates and classification of spots of laser beams using the parallel–hierarchical transformation method.

The selection of classification boundaries was based on minimizing the mean error relative to the center of the actual image. In order to identify the optimal boundaries, the mean error for different boundary distances was estimated depending on the selected correlation coefficient of the combined method.

The graph of the average error, depending on the choice of threshold for one of the tracks, is presented in Figure 7.



(**a**) (**b**)

Figure 6. Models of the computational complex of parallel–hierarchical processing of images of extended laser lines and predictive characteristics for optical communication systems based on FPGA usage: (**a**)—first modification and (**b**)—second modification.

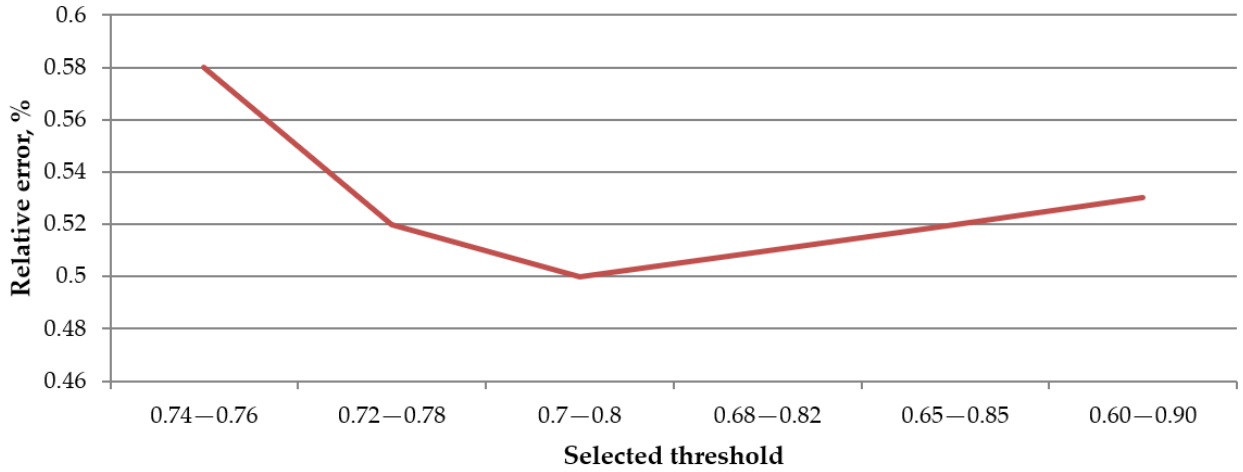


Figure 7. Average coordinate-determination error as a function of the selected threshold for one selected laser track.

Figure 7 shows the dependence of the average coordinate-determination error on the selected threshold for one selected laser track. The curve has a distinct minimum corresponding to the range in which the coordinate-determination error is smallest. Further widening of the acceptance range increases the error because the decision criterion becomes less selective. The position of the minimum therefore provides the experimental basis for choosing the working threshold for this track. Classification performance was also evaluated using the ROC method [20], as shown in Figure 8. The developed method produced a true-positive rate of 71% and a false-positive rate of 23%.

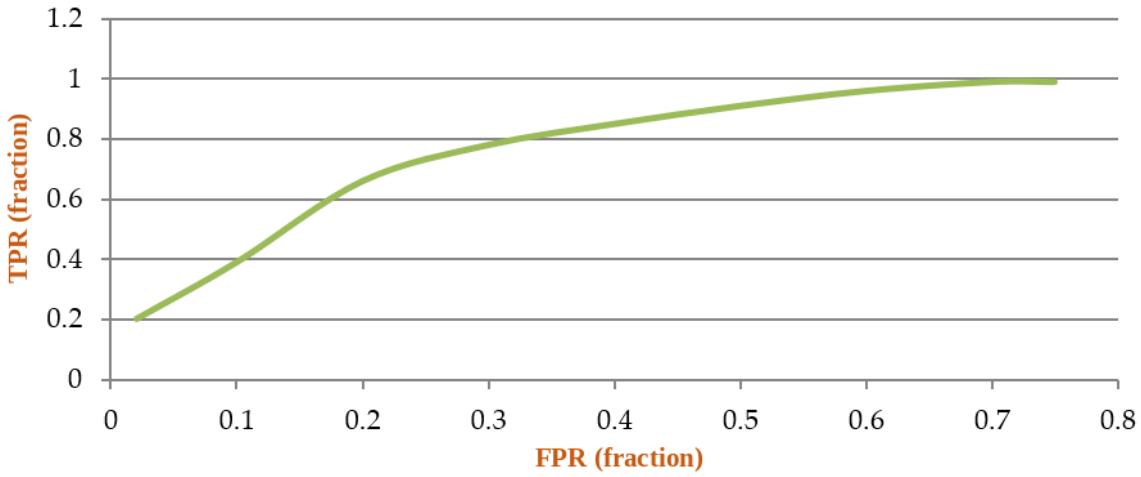


Figure 8. Receiver operating characteristic (ROC) curve of the developed method. Both axes show dimensionless fractions, not percentages.

The developed approach produced an average deviation of 0.47%, a maximum deviation of 1.6%, and a standard deviation of 0.36%. Figure 9 presents a comparative error graph for a route segment using the X coordinate, with the baseline contour-based method [8] used in the original comparison.

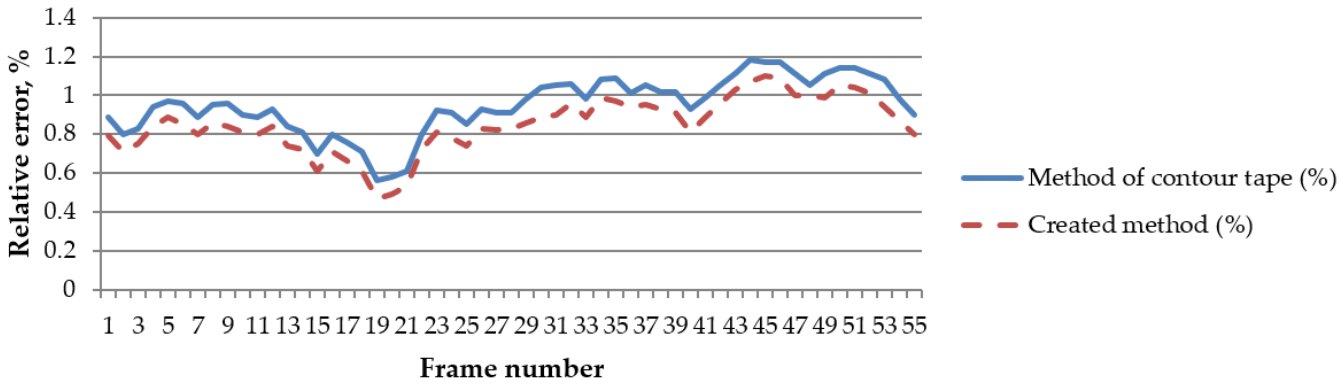


Figure 9. Comparative graph of relative errors.

The area under the ROC curve (AUC) was estimated as approximately 0.81 from the available ROC graph. This graph-derived estimate complements the reported operating point, for which the true-positive rate is 71% and the false-positive rate is 23%. Because frame-level ROC scores are unavailable, the AUC value should be interpreted as a descriptive estimate rather than as a directly computed frame-level AUC.

The graph shows that the combined method has a lower average error than the baseline contour-based method over the displayed route segment, corresponding to more accurate trajectory calculation for these data.

In laser-beam data transmission systems, it is important not only to determine the position of the energy center in each laser-path image but also to predict its subsequent position. The 199-frame reanalysis belongs to the same frame sequence and therefore characterizes that sequence only. Relative to the same computational reference, the intensity-weighted centroid produced a mean coordinate-determination error of 4.627 pixels (3.615% of image width), the six-section method produced 1.755 pixels (1.371%), and the complete modified PHT algorithm produced 0.567 pixels (0.443%). These values should not be interpreted as estimates of accuracy on other independent laser tracks.

The direct 38-section method produced the lower mean error of 0.368 pixels. This value characterizes direct coordinate determination for the current frame: a new coordinate of the energy center of the laser-beam spot is calculated for every frame and compared with the computational reference coordinate calculated for that same frame. The complete PHT method includes two additional sequential control stages after coordinate calculation. First, a laser-beam-spot shape coefficient is calculated as the mean normalized intensity of pixels belonging to the image region above one-half of the maximum frame intensity. For reproduction of the original SCR071-SCR269 calculation, the previously established acceptance interval 0.70–0.77 is retained. Second, the output PHT vector of the current frame is compared with an average reference PHT vector formed from the output PHT vectors of the four most recent previously accepted frames. Both vectors are normalized, and the correlation coefficient between them is calculated. The new coordinate is accepted only when the correlation coefficient is at least 0.95; otherwise, the newly calculated coordinate is not accepted and the previously accepted coordinate is retained. Consequently, the values 0.368 pixels for the direct 38-section method and 0.567 pixels for the complete PHT characterize different processing procedures and should not be interpreted as two measurements of exactly the same operation.

We next perform a reproducibility check of the complete processing chain. The complete PHT calculation was repeated for frames SCR071-SCR269. The laser-beam-spot shape coefficient was calculated as the mean normalized intensity of pixels in the region above one-half of the maximum frame intensity; for this check within the same frame sequence, the original interval 0.70–0.77 was retained. Four initial reference PHT vectors were formed from frames SCR067-SCR070. Of the subsequent 199 frames, 131 passed the primary shape-coefficient condition and 118 passed both the shape and PHT-vector correlation conditions. The repeated mean coordinate-determination error of the energy center of the laser-beam spot was 0.569 pixels; the standard deviation was 0.475 pixels, the median was 0.433 pixels, and the maximum was 3.318 pixels. The repeated mean differs from the previously reported value of 0.567 pixels by only 0.002 pixels. This confirms numerical reproducibility of the complete PHT for the SCR071-SCR269 sequence.

Let us examine the effect of the number of sections using separate completed frame sequences. The tr0_269 dataset contains 269 consecutively acquired BMP frames of the laser-beam spot image. In the original processing scheme, the first 70 frames were used at the preliminary stage to form the initial processing characteristics required by the algorithm, whereas the following 199 frames were used for evaluation after those characteristics had been formed. Thus, the boundary after frame 70 separates two processing stages within one frame sequence and does not represent a separation of independent laser tracks. For the additional numerical analysis, completed acquisition sequences were divided into two non-overlapping groups. TRASSA00 and FRAME02-FRAME05 were used for implementation checks and parameter definition. FRAME06-FRAME11 were not used when those parameters were selected and were used in their entirety only for subsequent evaluation. The FMT headers provide the frame count, image dimensions, inter-frame period, and individual acquisition start and end times. FRAME02-FRAME11 each contain 2044 frames of 128 $\times$ 128 pixels with a 40 ms inter-frame period and have separate start and end times. The intervals between the end of one acquisition sequence and the beginning of the next are substantially longer than one inter-frame period. FRAME02-FRAME11 are therefore treated as separate completed frame sequences rather than as contiguous fragments of one continuously stored FMT file. No sequence and no frame contributed to both groups. TRASSA00 contains 933 frames, and FRAME02-FRAME11 contain 2044 frames each, giving 21,373 readable frames in total. Both section-count variants use the same sequence of brightness levels starting from the maximum frame intensity.

The six-section variant uses the first six levels; the 38-section variant continues the same sequence with 32 additional levels. Thus, the first six sections are identical in both variants; the only difference is the additional 32 successive brightness sections. The center of each selected section is calculated, and the final coordinate is obtained from the centers of the selected sections.

For the same SCR071-SCR269 sequence, the standard deviation, median, and maximum coordinate-determination error of the energy center of the laser-beam spot were 0.463, 0.423, and 2.845 pixels for the complete modified PHT algorithm; 1.028, 1.629, and 5.690 pixels for the six-section method; and 1.157, 4.661, and 8.074 pixels for the intensity-weighted centroid method. These values characterize this single frame sequence and are reported descriptively.

Consecutive frames within one acquisition sequence are statistically dependent. Therefore, individual frames are not treated as independent statistical units in the comparison. For each of FRAME06, FRAME07, FRAME08, FRAME09, FRAME10, and FRAME11, the coordinate-determination error of the energy center of the laser-beam spot was first calculated for every frame using Equation (28) and then averaged over the complete sequence. Each sequence therefore contributes one paired result: a mean error for the 38-section method and a mean error for the six-section method, calculated from the same frames relative to the same computational reference. The mean errors for the 38-section and six-section methods, respectively, were: FRAME06, 0.885 and 1.084 pixels; FRAME07, 0.667 and 0.801; FRAME08, 0.705 and 0.887; FRAME09, 0.801 and 1.019; FRAME10, 0.638 and 0.769; and FRAME11, 1.196 and 3.367 pixels. FRAME11 has the same FMT format, dimensions, frame count, and frame period as FRAME06-FRAME10 and was therefore retained in the analysis. Its larger six-section error coincides with greater frame-to-frame intensity variability and is not used as a reason for exclusion. The 38-section method had the lower mean error in all six sequences. The means across the six sequences were 0.815 and 1.321 pixels, and the medians were 0.753 and 0.953 pixels for the 38-section and six-section methods, respectively. For each sequence, the difference between the mean 38-section error and the mean six-section error was calculated; all six differences were negative. A paired one-sided Wilcoxon signed-rank test gave a statistic of 0 and a p-value of 0.015625. A statistic of 0 means that none of the six sequences had a larger mean error for the 38-section method than for the six-section method. Resampling whole sequences rather than individual frames gave a 95% confidence interval from -1.178 to -0.152 pixels for the mean paired difference. Because the entire interval is below zero, this result further supports the lower mean error of the 38-section method. This analysis accounts for within-sequence frame dependence by using the mean of the entire sequence, rather than the individual frame, as the statistical unit of comparison.

We also perform an additional evaluation of the complete PHT on the separate completed frame sequences FRAME06-FRAME11. The original shape-coefficient interval 0.70–0.77 reproduced the calculation for SCR071-SCR269 but did not represent the shape-coefficient distribution in the other sequences. To avoid using FRAME06-FRAME11 to select new limits, the revised shape-coefficient interval was calculated only from TRASSA00 and FRAME02-FRAME05, the sequences used for implementation checks and parameter definition. The central 95% empirical interval of the shape coefficient gave fixed limits of 0.637–0.862. These limits, together with the unchanged PHT-vector correlation threshold of 0.95, were then applied to FRAME06-FRAME11 without sequence-specific adjustment. For each evaluation sequence, frames 67–70 were used only to initialize the four reference PHT vectors required to begin sequential processing; these four frames were excluded from the error calculation and were not used to determine or modify the shape-coefficient limits or the correlation threshold. Frames 71–2044 formed the evaluation segment, giving

1974 frames per sequence. The numbers of frames passing the shape-coefficient condition and both sequential conditions, respectively, were: FRAME06, 1912 and 1910; FRAME07, 1956 and 1950; FRAME08, 1932 and 1928; FRAME09, 1903 and 1901; FRAME10, 1935 and 1931; and FRAME11, 1750 and 1747. In total, 11,367 of 11,844 evaluated frames (95.97%) passed both conditions. The mean coordinate-determination errors of the complete PHT for FRAME06-FRAME11 were 0.903, 0.672, 0.763, 1.215, 0.650, and 1.562 pixels, respectively. On exactly the same frames, the direct 38-section method gave 0.894, 0.669, 0.710, 0.803, 0.642, and 1.216 pixels. Thus, the complete PHT had the larger same-frame error in all six evaluation sequences; the means across the six sequences were 0.961 and 0.822 pixels for the complete PHT and direct 38-section method, respectively. A paired one-sided Wilcoxon test for the six sequence-level pairs gave a statistic of 21 and a p-value of 0.015625 for the alternative that the complete-PHT error is greater. Resampling whole sequences gave a 95% confidence interval of 0.014–0.274 pixels for the mean increase. This additional evaluation supports the explanation of the 0.368-versus-0.567-pixel difference: the complete PHT is not simply another way of calculating the same coordinate. Its sequential acceptance logic can retain a previously accepted coordinate, so the same-frame error can increase when the spot position changes.

The forecasting results below were obtained using coordinates generated by the proposed PHT-based processing chain. Comparison of the average prediction error shows that the proposed preprocessing gave lower errors than the baseline contour-based method from the original comparison [8] for each of the five forecasting methods tested, although the magnitude of the improvement depended on the forecasting method. The corresponding relative errors and reductions were:

- Exponential smoothing with damping: 0.66% versus 0.69%; relative error reduction, 4.3%.
- Exponential smoothing with an additive model: 0.85% versus 0.87%; relative error reduction, 2.3%.
- AR model: 0.97% versus 1.43%; relative error reduction, 32.2%.
- Linear network S5: 0.37% versus 0.46%; relative error reduction, 19.6%.
- Multilayer perceptron MP5-2-1: 0.38% versus 0.47%; relative error reduction, 19.1%.

Across the five methods, the mean prediction error decreased from 0.78% to 0.65%, corresponding to an overall relative reduction of 17.6%. The reduction for individual methods ranged from 2.3% to 32.2%; the mean of the five relative reductions was 15.5%, with a standard deviation of 12.3 percentage points.

The results demonstrate that the proposed direct parallel–hierarchical transformation method achieves competitive accuracy in laser beam spot classification and coordinate determination. The average error of 0.47%, maximum error of 1.6%, and standard deviation of 0.36% represent a meaningful improvement over the baseline contour-based method [8]. The proposed method reduced the prediction error relative to the baseline contour-based method [8] for all five forecasting methods tested, which indicates that the gain originates from the PHT decomposition itself rather than from any particular downstream prediction technique; the magnitude of the reduction was, however, method-dependent, ranging from 2.3% to 32.2% relative to the corresponding baseline error.

The observed reduction in average prediction error (17.6% on average, though varying by method) aligns with the theoretical expectation that level-by-level decomposition progressively filters irrelevant spatial variation, leaving increasingly purified feature sets for downstream classification. This mirrors the convergence–divergence architecture of biological sensory pathways [2–4], where successive hierarchical levels filter common and distinctive signal features before integration in higher cortical areas. The comparison between the PHT network structure and Mountcastle's modular cortical columns [5] is a

structural analogy concerning hierarchical, partly parallel organization; it does not establish equivalence between the biological and computational mechanisms or independently prove their robustness.

The ROC operating point of a true-positive rate of 71% and a false-positive rate of 23% should be interpreted together with how it was obtained: it was not selected by optimizing an ROC-specific criterion such as the Youden index or a cost-weighted classification loss, but is the point reached on the curve when the classification boundary is fixed by the criterion described in Section 3—minimization of the mean coordinate error relative to the center of the actual image (Figure 7). The threshold was therefore tuned for coordinate accuracy, and the resulting classification behavior is a consequence of that choice rather than an independently optimized target; a localization criterion and a classification criterion need not share the same optimum.

Judged on its own terms as a classification result, this operating point is moderate rather than strong: it corresponds to a miss rate of 29%, a specificity of 77%, a Youden index of 0.48, where zero corresponds to chance-level separation and one to perfect separation, and a balanced accuracy of 74%. The area under the ROC curve (AUC) was estimated as approximately 0.81 from the available ROC graph. However, frame-level ROC scores, a second operating point, and a confidence interval for the AUC are not available, and no corresponding ROC curve was obtained for the baseline contour-based method. Therefore, the present data do not allow a direct classification-level comparison of PHT with the baseline contour-based method over the full range of decision thresholds. We regard this single-operating-point reporting, together with the absence of an explicit cost model justifying the chosen trade-off, as a limitation of the present classification evaluation rather than as evidence that a true-positive rate of 71% and a false-positive rate of 23% is optimal; a full frame-level ROC analysis with directly computed AUC and its confidence interval, together with a stated cost model for missed versus false detections, is identified as necessary future work. The well-defined minimum in the error-versus-threshold curve (Figure 7), with error rising again as the threshold widens, is nonetheless consistent with the theoretical properties of the PHT decomposition, in which tail elements at each level represent increasingly specific spatial features and over-broad thresholds aggregate elements belonging to distinct feature classes.

Processing speeds of up to 7 ms/frame and 25 ms/frame on the DE1-SoC platform confirm the practical viability of the method for real-time laser tracking in optical communication systems. These processing times correspond to throughputs of approximately 142.9 frames/s for the register-memory configuration and 40 frames/s for the serial-memory configuration. Thus, the register-memory configuration provides approximately 3.57 times the throughput of the serial-memory configuration. The incomplete memory matrix used in the register-memory configuration, necessitated by FPGA resource constraints, suggests that a full-resource implementation would yield further speed improvements.

The present study does not include a CNN- or RNN-based comparison implemented on the same laser beam spot dataset; a controlled, same-dataset benchmark is identified below as an important limitation and a priority for future work. To contextualize the magnitude of the reported improvement in the meantime, two recent deep-learning studies addressing the same physical task—laser-spot localization and prediction under atmospheric turbulence—are considered for comparison here, on the understanding that differences in dataset and error metric mean their figures cannot be compared directly with the results of this study. A deep-learning-enhanced spot center positioning and reconstruction method evaluated under atmospheric turbulence—the same distortion source considered in this study—improved Manhattan-distance positioning precision by more than 2.24 pixels and reconstruction PSNR by more than 3.25 dB relative to traditional

methods [21]. For the specific sub-task of predicting a future spot position under atmospheric turbulence, directly analogous to the forecasting comparison performed in Section 3, an RNN-based feature-matching prediction method reduced average prediction error by 44–62% relative to classical centroid/Hough/Gaussian-fit baselines and by a further 12% relative to a CNN-LSTM baseline [22]. These results provide only qualitative context for the 17.6% average improvement of PHT over the baseline contour-based method reported in Section 3, because differences in datasets, evaluation metrics, and experimental protocols do not permit a direct quantitative comparison with the cited deep-learning-based methods.

Beyond this contextual placement, the PHT method offers an interpretable hierarchical structure with explicitly defined decomposition levels, making it easier to analyze failure modes and tune classification boundaries than the implicit, high-dimensional feature representations learned by CNNs [7,23]—a relevant advantage in safety-critical optical applications. A direct, quantitative comparison of PHT against a CNN or RNN baseline trained and evaluated on the same dataset used in this study was not performed here; this is an important limitation of the present work and a priority for future research, since it would allow the accuracy-interpretability trade-off between PHT and deep learning methods to be assessed under identical experimental conditions rather than through cross-study literature comparison. Nevertheless, the scientific significance of the proposed approach lies not in exceeding deep-learning accuracy but in achieving improvements of a comparable order of magnitude through a deterministic, mathematically transparent PHT computational core that does not require iterative parameter learning or gradient-based optimization; the training subset is used only to form the reference characteristics required by the surrounding classification and reliability procedure, while every intermediate tail element remains traceable to an explicit decomposition level. Combined with the demonstrated real-time FPGA implementation at up to 7 ms/frame, this establishes PHT as a scientifically distinct and practically valuable alternative to deep learning for hierarchical signal processing in resource-constrained, safety-critical, and interpretability-sensitive applications, where a favorable accuracy-interpretability-efficiency trade-off is often more valuable than a marginal gain in raw accuracy.

Future work should also investigate the extension of PHT to three-dimensional and video signal contexts, building on earlier work on hierarchical optical-image processing [8]. Integration of adaptive threshold selection through online learning—rather than the static minimization used here—represents a further direction for reducing sensitivity to atmospheric variability. Extension to other modalities where hierarchical sensory structure is present, including medical imaging [8] and acoustic signal processing, warrants exploration.

The following discussion sets out an application domain for PHT that, unlike the laser beam spot results reported above, was not implemented or evaluated experimentally in the present study; it is included to motivate and scope future work. Railway infrastructure maintenance increasingly relies on automated inspection systems capable of early detection and classification of rail defects. Modern railway inspection vehicles integrate laser scanners, machine vision systems, ultrasonic testing, eddy-current sensors, and track geometry measurement equipment to acquire large volumes of heterogeneous diagnostic information. Efficient processing of these data requires advanced computational methods that support multilevel feature extraction, hierarchical data integration, and intelligent classification in real time.

The proposed parallel–hierarchical transformation offers a candidate unified computational framework for processing such multisource diagnostic information. In the envisaged scheme, which has not been implemented or evaluated on railway data here, lower hierarchical levels would extract local geometric, textural, and structural features from inspection

data, whereas higher hierarchical levels would integrate these features into generalized descriptors for defect identification and infrastructure condition assessment.

Taking the classification principles defined in UIC recommendations and EN 16729 [24] as the target taxonomy, the proposed approach could in principle be adapted for the identification of major categories of rail defects, such as rolling contact fatigue (head checking and gauge corner cracking), squats, corrugation, shelling and spalling, transverse and longitudinal cracks, welded joint defects, rail head and side wear, corrosion damage, plastic deformation, and other surface defects associated with cyclic loading and material degradation. No such adaptation was implemented or tested in this work; these categories indicate the intended scope of a future study rather than demonstrated detection capability.

The hierarchical representation is intended to allow simultaneous analysis of local defect characteristics and the overall condition of railway track components. Whether this would improve classification reliability or reduce false-positive detections relative to existing rail inspection pipelines has not been tested and cannot be inferred from the laser beam spot results reported above, for which classification performance is itself characterized only at a single ROC operating point. Integration of information obtained from laser scanning, machine vision, ultrasonic testing, eddy-current inspection, and track geometry monitoring is likewise proposed as a route towards comprehensive condition assessment in support of condition-based maintenance, rather than as a demonstrated result.

The frame rates reached by the FPGA implementation on laser beam spot images are of the order required by onboard railway inspection systems, although no railway sensor interface, data rate, or inspection payload was implemented or benchmarked here. Because the present study validates the proposed transformation only on laser beam spot images, the expectation that its computational principles carry over to railway infrastructure diagnostics rests on the application-independence of the underlying mathematical framework rather than on experimental evidence. The extracted hierarchical features are therefore put forward, but not yet shown, to be informative descriptors for rail defect identification, infrastructure condition assessment, and maintenance decision support.

Therefore, the proposed parallel–hierarchical transformation is put forward as a promising, but as yet experimentally unverified, computational framework for intelligent railway track infrastructure maintenance. Its hierarchical organization and efficient parallel implementation suggest opportunities for future integration into condition-based and predictive maintenance systems, Digital Twin technologies, and AI-based railway asset management platforms, with the aim of contributing to improved operational safety, infrastructure reliability, and lifecycle management. Realizing and quantifying these benefits would require validation on railway inspection data, which is beyond the scope of the present study.

4. Conclusions

This study formulated a complete mathematical model of the direct parallel–hierarchical transformation (PHT), extending the theoretical foundations established in [8,14]. The model was implemented on a DE1-SoC FPGA platform and demonstrated experimentally using classification and coordinate determination of laser-beam spot images distorted by atmospheric turbulence. The scope of statistical inference was revised after auditing the recovered acquisition archive.

In the original experiment, the developed method achieved an average coordinate-determination error of 0.47%, a maximum error of 1.6%, and a standard deviation of 0.36%, with an ROC operating point of a true-positive rate of 71% and a false-positive rate of 23%. In the descriptive evaluation of 199 frames from one frame sequence, the complete modified PHT gave a mean coordinate-determination error of the laser-beam spot energy center of

0.567 pixels (0.443% of the frame width), and the repeated calculation gave 0.569 pixels. The additional analysis of the section-count effect used FRAME06-FRAME11, which were not used for parameter selection. The 38-section method had a lower mean coordinate-determination error than the six-section method in all six sequences; the means across the six sequences were 0.815 and 1.321 pixels, respectively (Wilcoxon statistic 0; one-sided p-value 0.015625; 95% confidence interval for the paired difference from whole-sequence resampling, -1.178 to -0.152 pixels). A separate evaluation of the complete PHT on the same six sequences used the shape-coefficient interval 0.637–0.862 determined only from TRASSA00 and FRAME02-FRAME05 and the unchanged correlation threshold of 0.95. On frames 71–2044, 11,367 of 11,844 frames passed both sequential acceptance conditions. The mean same-frame error across the six sequences was 0.961 pixels for the complete PHT versus 0.822 pixels for the direct 38-section stage. The complete-PHT error was higher in all six sequences, quantitatively confirming that the two procedures serve different processing purposes: the complete chain applies sequential acceptance logic and may retain the preceding coordinate, whereas the direct 38-section method calculates a new coordinate for every frame. Relative to the baseline contour-based method, PHT preprocessing reduced the average prediction error across five forecasting methods by 17.6%, from 0.78% to 0.65%. The processing chain reached 7 ms/frame with register memory and 25 ms/frame with serial memory.

The archive review shows that the 269-image tr0_269 dataset used for the 70/199 division is one sequentially acquired frame sequence. Its frame-level error distribution is therefore reported as a descriptive characteristic of that sequence, and earlier frame-level significance calculations are not used as evidence of accuracy on other acquisition sequences. The additional analysis is based on six separate completed sequences, FRAME06-FRAME11, each used in its entirety and each having its own FMT header, start and end times, and 2044 frames. The intervals between successive sequences substantially exceed the 40 ms inter-frame period, so these sequences are not treated as contiguous fragments of one continuously stored FMT file. In the section-count comparison, the complete frame sequence, rather than the individual frame, is the statistical unit. FRAME06-FRAME11 were not used for implementation checks or parameter selection. The additional complete-PHT evaluation also shows that the original shape-coefficient interval 0.70–0.77 applies to the original SCR071-SCR269 sequence; for evaluation on other sequences, a fixed interval was therefore determined only from TRASSA00 and FRAME02-FRAME05 and then applied unchanged to FRAME06-FRAME11. The complete-PHT and direct 38-section results are intentionally distinguished: the direct stage calculates a coordinate for the current frame, whereas the complete chain additionally decides whether to accept that coordinate or retain the preceding one. Further limitations are the absence of a same-dataset CNN or RNN comparison and the availability of only a graph-derived AUC estimate of approximately 0.81, without frame-level ROC scores or an AUC confidence interval.

Because the transformation operates on abstract numerical fields, its computational principles are application-independent. A prospective extension, not evaluated in the present study, is the hierarchical fusion of heterogeneous railway diagnostic data—laser scanning, machine vision, ultrasonic and eddy-current testing, and track geometry measurement—in which low-level geometric and textural features would be aggregated into generalized descriptors supporting rail defect identification and condition-based maintenance decision support [12–17]. Experimental validation of PHT on railway inspection data remains a direction for future research.

Author Contributions: Conceptualization, V.T., N.K., L.T., Y.D., M.B. and M.D.; methodology, V.T., N.K., L.T., Y.D., M.B. and M.D.; software V.T., N.K., L.T., Y.D., M.B. and M.D.; validation, V.T., N.K., L.T., Y.D., M.B. and M.D.; formal analysis, V.T., N.K., L.T., Y.D., M.B. and M.D.; investigation, V.T., N.K., L.T., Y.D., M.B. and M.D.; resources, V.T., N.K., L.T., Y.D., M.B. and M.D.; data curation, V.T., N.K., L.T., Y.D., M.B. and M.D.; writing—original draft preparation, V.T., N.K., L.T., Y.D., M.B., M.D., V.L., R.S. and R.M.; writing—review and editing, V.T., N.K., L.T., Y.D., M.B., M.D., V.L., R.S. and R.M.; visualization, V.T., N.K., L.T., Y.D., M.B., M.D., V.L., R.S. and R.M.; supervision, V.T., N.K., L.T., Y.D., M.B., M.D., V.L., R.S. and R.M.; project administration, V.T. and V.L.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The experimental dataset of laser beam spot images used for validation in this study was originally recorded and reported in our earlier work [8]. The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PHT	Parallel–hierarchical transformation
STE	Space-time environment
DM	The model of development

References

1. Chen, K.C.; Peng, W.H.; Lee, C.G.G. Overview of Intelligent Signal Processing Systems. *APSIPA Trans. Signal Inf. Process.* **2023**, *12*, e36. [CrossRef]
2. Yufik, Y.M. The Understanding Capacity and Information Dynamics in the Human Brain. *Entropy* **2019**, *21*, 308. [CrossRef] [PubMed]
3. Fritzsch, B.; Elliott, K.L.; Yamoah, E.N. Neurosensory Development of the Four Brainstem-Projecting Sensory Systems and Their Integration in the Telencephalon. *Front. Neural Circuits* **2022**, *16*, 913480. [CrossRef] [PubMed]
4. Zhang, N.; Xu, N.–.L. Reshaping sensory representations by task-specific brain states: Toward cortical circuit mechanisms. *Curr. Opin. Neurobiol.* **2022**, *77*, 102628. [CrossRef] [PubMed]
5. Mountcastle, V.B. The columnar organization of the neocortex. *Brain* **1997**, *120*, 701–722. [CrossRef] [PubMed]
6. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
7. Xin, M.; Wang, Y. Research on image classification model based on deep convolution neural network. *EURASIP J. Image Video Process.* **2019**, *2019*, 40. [CrossRef]
8. Tymchenko, L.; Tverdomed, V.; Petrovsky, N.; Kokryatskaya, N.; Maistrenko, Y. Development of a Method of Processing Images of Laser Beam Bands with the Use of Parallel-Hierarchic Networks. *East.-Eur. J. Enterp. Technol.* **2019**, *6*, 21–27. [CrossRef]
9. Zhu, J.; Xu, Z.; Fu, D.; Hu, C. Laser Spot Center Detection and Comparison Test. *Photonic Sens.* **2019**, *9*, 49–52. [CrossRef]
10. Gubarevych, O.; Goolak, S.; Daki, O.; Tryshyn, V. Investigation of Turn-To-Turn Closures of Stator Windings to Improve the Diagnostics System for Induction Motors. *Probl. Energ. Reg.* **2021**, *2*, 10–24. [CrossRef]
11. Gubarevych, O.; Goolak, S.; Daki, O.; Yakusevych, Y. Determining an additional diagnostic parameter for improving the accuracy of assessment of the condition of stator windings in an induction motor. *East.-Eur. J. Enterp. Technol.* **2021**, *5*, 21–29. [CrossRef]
12. Kazemi, M.J.; Rashidi, M.; Kang, W.-H.; Siahkouhi, M. Toward Smart Railway Infrastructure Predictive and Optimised Maintenance Through Digital Twin (DT) System. *Sensors* **2026**, *26*, 2333. Correction in *Sensors* **2026**, *26*, 3430. https://doi.org/10.3390/s26113430. [CrossRef] [PubMed]
13. Alif, M.A.R.; Tucker, G.; Hussain, M. Advances in computer vision for comprehensive railway engineering: From track inspection to rolling stock and safety monitoring. *Railw. Eng. Sci.* **2026**. [CrossRef]
14. Lukoševičius, V.; Tverdomed, V.; Tymchenko, L.; Kokriatska, N.; Didenko, Y.; Demchenko, M.; Oliynyk, O. Analysis of G-Transformation Modes for Building Neuro-like Parallel–Hierarchical Network Identification of Rail Surface Defects. *Mathematics* **2025**, *13*, 966. [CrossRef]
15. Hu, Z.; Lau, A.; Tao, R.; Liu, Y.; Pan, Y. A comprehensive review of railway track condition monitoring: From sensor selection to data processing for defect-oriented, task-driven applications. *J. Railw. Sci. Technol.* **2026**, *2*, 1–28. [CrossRef]

16. Sresakoolchai, J.; Kaewunruen, S. Railway infrastructure maintenance efficiency improvement using deep reinforcement learning integrated with digital twin based on track geometry and component defects. *Sci. Rep.* **2023**, *13*, 2439. [CrossRef] [PubMed]
17. Sanz Bobi, J.d.D.; Garrido Martínez-Llop, P.; Rubio Marcos, P.; Solano Jiménez, Á.; Gómez Fernández, J. Prediction of Degraded Infrastructure Conditions for Railway Operation. *Sensors* **2024**, *24*, 2456. [CrossRef] [PubMed]
18. Grasso, P.A.; Gallina, J.; Bertini, C. Shaping the Visual System: Cortical and Subcortical Plasticity in the Intact and the Lesioned Brain. *Neuropsychologia* **2020**, *142*, 107464. [CrossRef] [PubMed]
19. Yang, K.; Li, J.; Liu, M.; Lei, T.; Xu, X.; Wu, H.; Cao, J.; Qi, G. Complex Systems and Network Science: A Survey. *J. Syst. Eng. Electron.* **2023**, *34*, 543–573. [CrossRef]
20. Saito, T.; Rehmsmeier, M. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLoS ONE* **2015**, *10*, e0118432. [CrossRef] [PubMed]
21. Wang, J.; Meng, X.; Zhou, S.; Wang, X.; Han, J.; Guo, Y.; Song, S.; Liu, W. SPRNet: Laser spot center position and reconstruction under atmospheric turbulence based on deep learning enhancement. *Opt. Lasers Eng.* **2025**, *186*, 108775. [CrossRef]
22. Wang, X.; Zhang, N.; Wang, Q.; Cui, L. Accurate prediction approach for the center position of a future light spot under atmospheric turbulence. *Opt. Express* **2025**, *33*, 36481–36504. [CrossRef] [PubMed]
23. Wang, J.; Zhu, H.; Wang, S.; Zhang, Y. A review of deep learning on medical image analysis. *Mob. Netw. Appl.* **2021**, *26*, 351–380. [CrossRef]
24. *EN 17397-1:2020*; Railway Applications—Rail Defects—Part 1: Rail Defect Management. European Committee for Standardization (CEN): Brussels, Belgium, 2020.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.