A survey of multimodal data fusion: applications and adverse conditions

Una revisión de fusión de datos multimodal: aplicaciones y condiciones adversas

DOI: https://dx.doi.org/10.17981/ingecuc.21.1.2025.08

Original Research.
Date Received: 16/08/2023, Date Accepted: 20/04/2024.

Daniel Fernando Quintero-Bernal
Universidad de Santiago de Chile. Santiago, (Chile)
daniel.quintero@usach.cl

Héctor Kaschel-Cárcamo
Universidad de Santiago de Chile. Santiago, (Chile)
hector.kaschel@usach.cl

John Kern-Molina
Universidad de Santiago de Chile. Santiago, (Chile)
john.kern@usach.cl

To cite this paper
D. Quintero-Bernal, H. Kaschel-Cárcamo & J. Kern-Molina "A survey of multimodal data fusion: applications and adverse conditions," *INGE CUC*, vol. 21, no. 1, 2025. DOI: https://dx.doi.org/10.17981/ingecuc.21.1.2025.08

Abstract
Multimodal data fusion is a research field that combines information from various sources, each with its own modalities. One of the prominent challenges in data fusion is addressing adverse conditions that arise when dealing with real-world implementations. This article presents a review of data fusion, conducting a comparative analysis through a literature search across various application domains that tackle the challenges of data heterogeneity and adverse conditions. The focus of this paper is centered on establishing a robust foundation to enable the development of future research in this field. This review reveals that almost half of the analyzed documents describe adverse conditions, but only a minority conduct analyses on how their techniques address noise or benefit from considering these conditions. The number of modalities used in the research is generally low, with most studies employing static data in 1D dimensions. Finally, it is crucial for researchers to continue working on an interdisciplinary vocabulary and consider applications in real-world environments with adverse conditions to advance in this field. It is also important to explore applications with higher-dimensional data and more modalities, which could provide valuable insights for addressing specific challenges in data fusion.

Keywords
Adverse conditions, data fusion, disturbances, heterogeneous data, multimodal, noise.

Resumen
La fusión de datos multimodal es un campo de investigación que combina información de diversas fuentes, cada una con sus propias modalidades. Uno de los desafíos más destacados de la fusión de datos es abordar las condiciones adversas que surgen al enfrentar implementaciones del mundo real. Este artículo presenta una revisión de la fusión de datos, realizando un análisis comparativo a través de una búsqueda bibliográfica en varios dominios de aplicaciones que abordan los desafíos de la heterogeneidad de datos y condiciones adversas. El enfoque de este documento se centra en establecer una base sólida que permita el desarrollo de futuras investigaciones en este campo. Esta revisión muestra que casi la mitad de los documentos analizados describen condiciones adversas, pero solo una minoría realiza análisis sobre cómo sus técnicas abordan el ruido o se benefician de considerar estas condiciones. La cantidad de modalidades utilizadas en las investigaciones es generalmente baja, y la mayoría de los estudios emplean datos estáticos en dimensiones 1D. Finalmente, es fundamental que los investigadores sigan trabajando en un vocabulario interdisciplinario y consideren aplicaciones en entornos del mundo real con condiciones adversas para avanzar en este campo. También es importante explorar aplicaciones con datos de mayor dimensionalidad y más modalidades, lo que podría proporcionar información valiosa para abordar desafíos específicos en la fusión de datos.

Palabras clave
Condiciones adversas, datos heterogéneos, fusión de datos, multimodal, perturbaciones, ruido.

I. INTRODUCTION

In recent years, data fusion (DF) techniques have significantly increased due to the growing availability of information across various fields—medicine, the Internet of Things (IoT), and robotics. DF methods leverage the combination of information to provide several advantages, including (1) enhanced accuracy, even in static systems relying on a single source; (2) reduction of redundancy in sensor networks, leading to decreased time and resource consumption in data transmission; and (3) transformation of raw information into a dataset with higher value and quality [1]. Specifically, multimodal data fusion (MMDF) refers to DF techniques that process diverse types of information —varying in dimensionality, temporality, resolution, etc.— from multiple data sources [2] to obtain a comprehensive representation of the observed phenomenon [1].

However, DF faces numerous challenges, including imprecision, inconsistency, privacy, and security issues [3]. Two challenges of particular relevance for scene understanding in real-world applications are heterogeneity and adverse environments addressed in this study. There is a strong relationship between the concepts of DF, heterogeneity, and complex environments since, in systems where noise, disturbances, or dynamic conditions are present, the MMDF approach becomes essential as it provides complementary perspectives of a scene of interest [4]. Many MMDF systems lack in-depth analysis focused on complex environments, which hinders their subsequent real-world implementation. Ma et al. [5] pointed out that traditional developments related to visible and infrared image fusion techniques have often ignored noise.

Furthermore, the researchers noted that these methods primarily focus on preserving valuable data information but do not typically address the requirements of a specific application. Blasch et al. [6] proposed that a future research direction for machine learning-based DF methods should involve scaling for detection in complex environments. Additionally, Cui & Niu [7] highlighted that selecting data from various modalities to meet the needs of specific tasks remains an active research topic, while Bokade et al. [2] emphasized that there is still no solid theoretical foundation for multimodal data fusion.

Various researchers have conducted reviews and surveys in the field of DF, as observed in [1], [3], [8]. Specifically, regarding MMDF, Y. Zhang et al. [9] conducted a study on models and techniques across various domains, while C. Zhang et al. [10] focused solely on deep learning techniques. Based on the literature reviewed, no existing DF study has primarily addressed applications considering the challenges of heterogeneity and complex environments.

This study highlights DF developments that have accounted for complex environments and heterogeneous datasets. It includes the areas of the applied techniques, the disturbances and noise analyzed, the purpose and level of fusion, the number of modalities, and the dimensionality and temporality of multimodality. The organization of this work is presented below (see Fig 1).

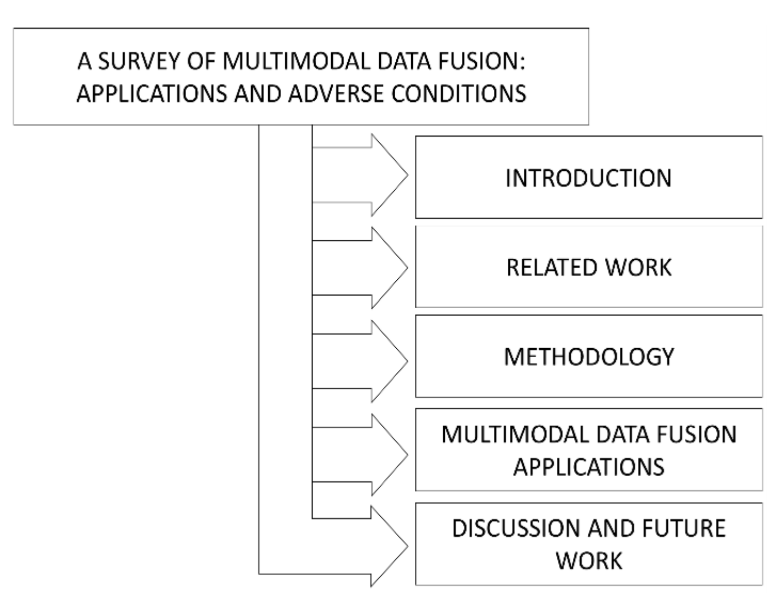

Fig 1. Structure of this document.

Section II presents the related work. Section III describes the methodology used. In Section IV, MMDF applications across various domains are examined. Finally, Section V discusses the obtained results and proposes possible directions for future research.

II. RELATED WORK

Table I presents the related works associated with this study. The reviews in references [3] and [8] provide a cross-sectional analysis of DF techniques across multiple domains, introducing various concepts from the field while also identifying open challenges and potential research directions. In [1], Meng et al. review machine learning methods. Y. Zhang et al. [9] surveyed available models and methods for multimodal fusion, while C. Zhang et al. [10] highlighted techniques centered on deep learning. Additionally, Bokade et al. [2] performed a comparative study across nine domains, considering various multimodal fusion methods.

Ullah & Youn [11] developed a review addressing data heterogeneity in the context of IoT, while Ghamisi et al. [12] presented a comprehensive study of multitemporal and multisource fusion techniques. References [13], [14] cover a broad range of topics—including definitions, classifications across different domains, evaluation metrics, open research questions, and potential future directions—specifically in data fusion algorithms incorporating images. In [4], H. Zhang et al. analyzed deep learning techniques for image fusion.

Since the lowest level of abstraction (data level) is the most common stage for multimodal data fusion (as indicated by Bokade et al. [2]), Li et al. [15] conducted a study on pixel-level image fusion methods. Additionally, [16] reviews deep learning-based methods in the same area (see Fig 2).

TABLE I. RELATED WORK

Reference	Year	Title	Covered Topics
[8]	2021	A survey on data fusion: what for? in what form? what is next?	A review on data fusion addressing various key questions: 1) Why is it necessary? 2) What is it used for? 3) What methods have been proposed? 4) What challenges remain open? and 5) What are potential research directions?
[3]	2019	Information Fusion for Multi-Source Material Data: Progress and Challenges	This document presents a review of techniques for material data fusion across multiple sources. It highlights: 1) Analysis of material data properties and multi-source fusion, 2) Recent achievements, and 3) Open challenges and future research directions.
[1]	2020	A survey on machine learning for data fusion	A review of machine learning techniques in data fusion. The document covers: 1) Introduction to background —definitions, models, techniques, and applications— of DF and machine learning, 2) Criteria for reviewing and evaluating algorithm performance in DF, and 3) Future research directions.
[9]	2021	Deep multimodal fusion for semantic image segmentation: A survey	The authors present an overview of MMDF methodologies, multimodal segmentation datasets, quantitative evaluations, and design choices. It also discusses current challenges and design options.
[10]	2020	Multimodal Intelligence: Representation Learning, Information Fusion, and Applications	This document reviews deep learning models and methods for multimodal fusion, focusing on vision and natural language through three aspects: 1) Multimodal representation learning, 2) Multimodal fusion at different levels, and 3) Applications.
[2]	2021	A cross-disciplinary comparison of multimodal data fusion approaches and applications: Accelerating learning through trans-disciplinary information sharing	A comparison across nine engineering domains of multimodal fusion techniques, quantifying the number and dimensions of modalities used in each case study. It also classifies algorithms and their application purposes.
[11]	2020	Intelligent Data Fusion for Smart IoT Environment: A Survey	A review of data fusion methods in IoT systems, highlighting trends, challenges, and research opportunities in DF.
[12]	2019	Multisource and Multitemporal Data Fusion in Remote Sensing	A compendium of multi-source and multi-temporal fusion, covering: 1) Point cloud fusion, 2) Hyperspectral and LiDAR data fusion, 3) Resolution and panoramic approaches, 4) Multitemporal data fusion, and 5) Big Data and social networks. It also outlines future challenges and research.

[13]	2021	Review of Various Image Fusion Algorithms and Image Fusion Performance Metric	A review of image fusion techniques, considering: 1) Spatial domain integration methodology, 2) Frequency domain and deep learning-based fusion techniques, and 3) Evaluation metrics.
[14]	2021	Image Fusion Techniques: A Survey	A review of strengths and weaknesses of different fusion techniques—spatial domain, frequency domain, and deep learning. Topics covered: 1) Evaluation metrics, 2) Applications across domains, and 3) Future directions.
[4]	2021	Image fusion meets deep learning: A survey and perspective	A review of deep learning techniques for image fusion, addressing: 1) Concepts, 2) Method classification, 3) State-of-the-art review across multiple applications, 4) Qualitative and quantitative evaluations of fusion methods, and 5) Challenges and future directions.
[15]	2017	Pixel-level image fusion: A survey of the state of the art	An in-depth study of pixel-level image fusion methods, discussing different evaluation metrics and concluding with future research directions.
[16]	2018	Deep learning for pixel-level image fusion: Recent advances and future prospects	A review of deep learning for pixel-level image fusion, covering: 1) Key difficulties in traditional DF, 2) Advantages and achievements of deep learning, 3) Various fusion techniques, and 4) Future research perspectives.

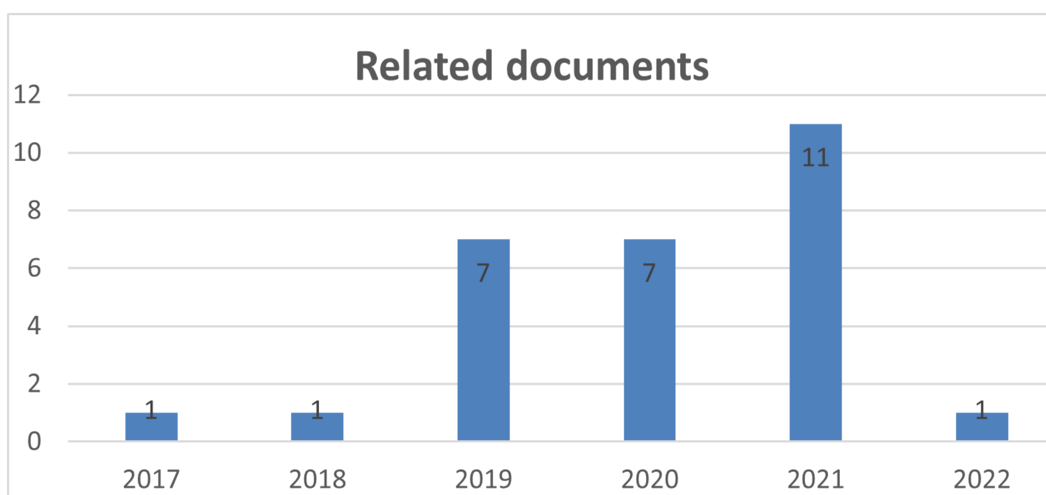

Fig 2. Histogram of related documentation.

Finally, some recent reviews and surveys on MMDF in specific domains are highlighted, including smart cities [17], activity recognition [18], agriculture [19], medical imaging [20], [21], [22], [23], [24], infrared and visible images [5], [25], hyperspectral and multispectral images [26], panoramic sharpening techniques for synthesizing panchromatic and multispectral images [27], pipeline monitoring [28], autonomous driving [29], [30], etc.

III. METHODOLOGY

The methodology section established the foundations to ensure a clear and precise understanding of the study presented. Initially, the specific terminologies used throughout the document were defined, which was crucial due to the diversity of terms employed in data fusion. Subsequently, the approach adopted for the identification, selection, and analysis of the literature was detailed, thereby ensuring the transparency and replicability of the review.

A. Key Concepts

The following presents definitions, classifications, and challenges associated with fusion, aiming to establish a conceptual framework for MMDF.

1. Data Fusion (DF)

DF is a process that enables the combination or integration of data through methods and models [7], which may originate from one or multiple sources [1], with the aim of obtaining more informative data [2] (Fig 3).

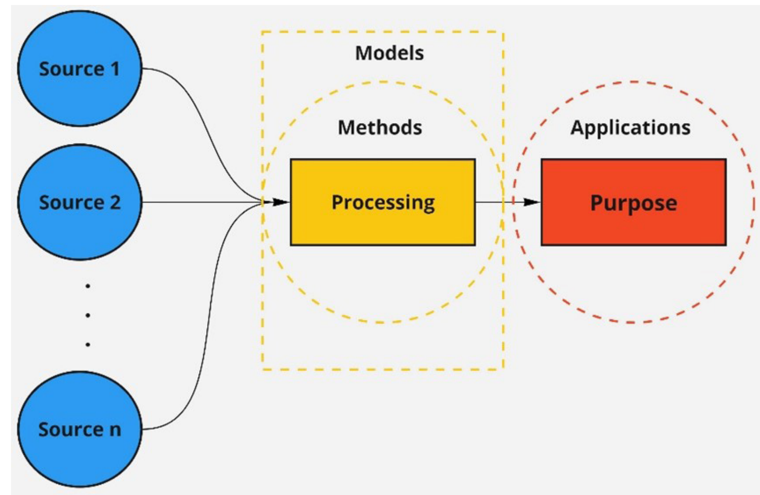

Fig 3. Data fusion.

From the previous concept, three key elements of DF can be identified: (1) Source, (2) Processing, and (3) Purpose.

- Data Source: Refers to elements such as databases, sensors, or prior system knowledge [1].
- Models and Methods: Models represent hierarchical transformations applied to data; these structures or architectures help unify terminologies to specify the characteristics and functionalities of DF systems. Researchers often use their architectures when developing applications, but fundamental models include (1) JDL (Joint Directors of Laboratories), (2) Luo & Kay, and (3) Dasarathy [1]. However, the fusion core lies not in the structures but in data processing methods [31].
- Purposes: DF methods are employed to achieve specific goals, such as dimensionality reduction, regression, classification, and clustering [2].

2. Multi-Modal Data Fusion (MMDF)

MMDF refers to integrating data from multiple modalities to generate more usable information. It also enables inferences in complex scenes that would not be possible using a single modality. Data obtained from different modalities exhibit specific characteristics, such as:

1. Dimensionality:
- 1D (one-dimensional information, i.e., data without spatial representation).
- 2D (two-dimensional information, such as images).
- 3D (three-dimensional information, such as point clouds).

2. Temporality:
- Static data (information without a temporal reference).
- Time series (e.g., video or samples acquired at specific time intervals) [2], [9].

A modality is a channel (mode) that captures information from a scene or object of study, representing how data is acquired [2]. Based on the findings of Y. Zhang et al. [9], some examples of modalities used in different applications include:

- Multimedia analysis: Video, audio, and text data.
- Medical imaging: Single-photon emission computed tomography, positron emission tomography, computed tomography, nuclear magnetic resonance, etc.
- Remote sensing: 3D point clouds, high-resolution optical data, and synthetic aperture radar.
- Other modalities mentioned include visible light cameras, infrared cameras, thermal cameras, and polarization cameras.

3. Classification of Data Fusion Methods

There are numerous classifications for addressing DF methods; below, multiple perspectives found in the literature are presented:

- Based on the mathematical methods used: Techniques based on Dempster-Shafer theory (also known as evidence theory or belief function theory), probability-based techniques, and artificial intelligence-based techniques [11], [31].

- Based on the relationship between information sources: Complementary, cooperative, and competitive [28].
- Based on the abstraction level of input and output: Data level (signal or pixel), feature level, and decision level [2], [3], [22], [27], [29], [32].

When the integration of information involves 2D data (images), various authors commonly use the following classifications:

- Based on the type of technique: Spatial domain techniques, frequency domain techniques, and deep learning-based techniques [13], [14], [23].
- Based on fusion categorization: Multi-view fusion, multi-focus fusion, multi-temporal fusion, and multimodal fusion [14], [21], [22].
- Finally, Y. Zhang et al. [9] categorized deep learning-based multimodal fusion methods into early, late, and hybrid fusion.

4. Data Fusion Challenges

Data fusion techniques enhance the quality of acquired information. According to Baroudi et al. [28], this improvement is reflected in at least four aspects: representation, accuracy, certainty, and completeness. However, due to the complexity of specific applications, fusion faces challenges such as those illustrated in Fig 4. Y. Zhang et al. [9] identified several challenges in multimodal fusion, including:

- Data diversity,
- Data quantity and quality,
- Data alignment and
- Dataset construction.

This study addresses two of the challenges highlighted in Fig. 4, which are also emphasized in research studies [1], [31]:

- Complex environments: DF is not a static process; data may only represent a specific period. This challenge arises due to dust particles, fog, lighting variations, and sensor-acquired noise. In this document, the terms (a) noise, (b) disturbances, and (c) dynamic environments are considered synonymous with complex environments.
- Heterogeneity: Information acquired from multiple sources may come from different sensors or modalities.

The focus of this study on these challenges stems from: (a) The need, as indicated by several researchers, for DF methods tailored to specific applications and (b) a strong relationship exists between heterogeneous data and adverse conditions in real-world implementations. Additionally, Y. Zhang et al. [9] state that high-quality, large-scale multimodal data are necessary for training deep learning methods. However, when these conditions are lacking, processing under noisy conditions must be considered.

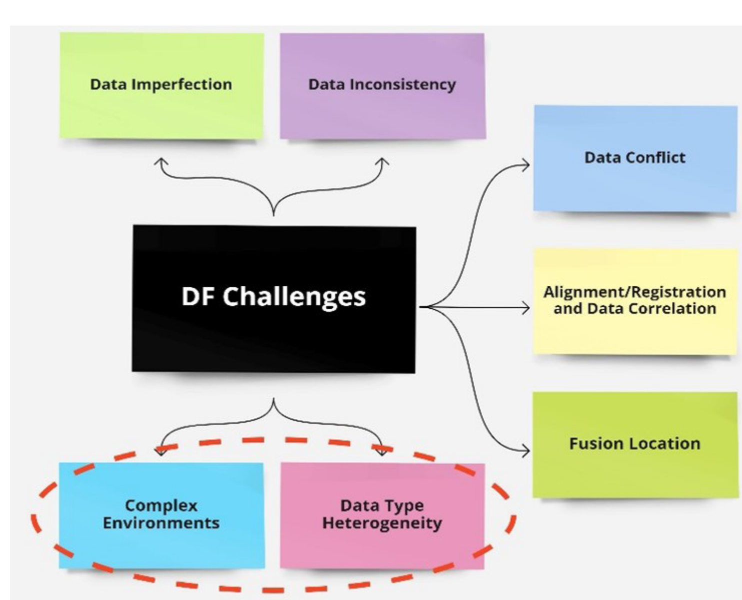

Fig 4. Data Fusion challenges.

B. Research Approach

This literature review was designed and conducted following a structured approach, aligned with the principles of PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), to ensure maximum transparency and replicability in the identification, selection, and analysis of relevant literature. The specific stages of the adopted methodological process are detailed below.

Literature Search:

- Databases used: A comprehensive search was conducted in Scite.ai and Web of Science, selected for their relevance and extensive coverage of scientific literature, which is essential for this study's focus.
- Search strategy: Carefully selected search terms, such as "multimodal data fusion" and "sensor fusion," among others, were employed. These terms were combined using the boolean operator OR —"data fusion" OR "multimodal data fusion" OR "multi-modal data fusion" OR "sensor fusion" OR "multisensor fusion" OR "multi-sensor fusion" OR "image fusion" OR "information fusion"— to encompass a broader range of studies.

Study Selection:

- Inclusion and exclusion criteria: Studies explicitly addressing data fusion and considering the heterogeneity of data sources in three specific domains—civil engineering, automation, and robotics—were included. Studies that did not align with these areas and lacked key terms related to adverse conditions (e.g., "noisy," "noise," "disturbance," "uncertainty," "adverse condition," "complex environment," "dynamic environment") were excluded.
- Selection process: The selection was conducted in two stages. Initially, articles were filtered based on their titles and abstracts to determine preliminary relevance. A full-text review was then performed to confirm inclusion based on the established criteria.

Data Analysis and Synthesis:

- Data extraction: Key information was extracted from each selected study, including descriptions of adverse conditions, DF level, modalities used, temporality and dimensionality, DF technique classification, and the specific purpose of fusion (see Fig 5).
- Information synthesis: A comparative analysis was conducted to identify trends, challenges, and research gaps in the existing literature.

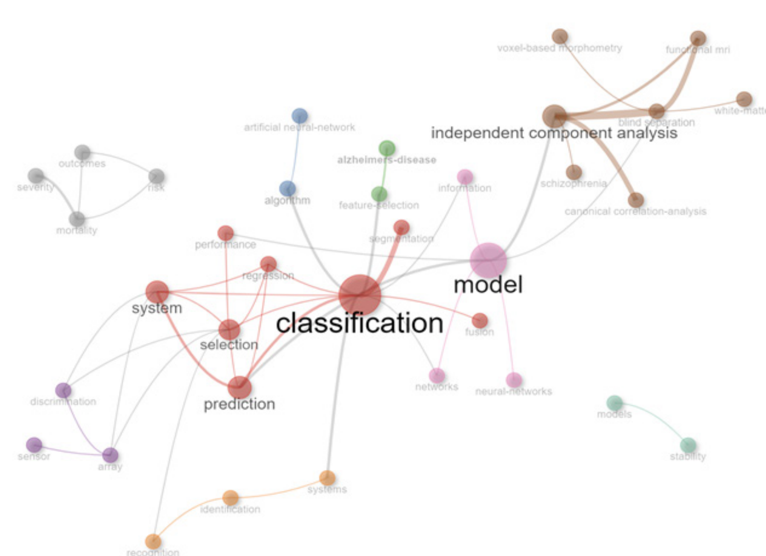

Fig 5. Bibliometric analysis

Fig 6 presents the connections between keywords in over 300 studies, allowing for a visualization of their relationships. These results highlight the interdisciplinary nature of data fusion, as it involves knowledge from various fields, which is reflected in the number and diversity of terms covered. However, it also reveals the lack of uniformity in the terminology used in this field.

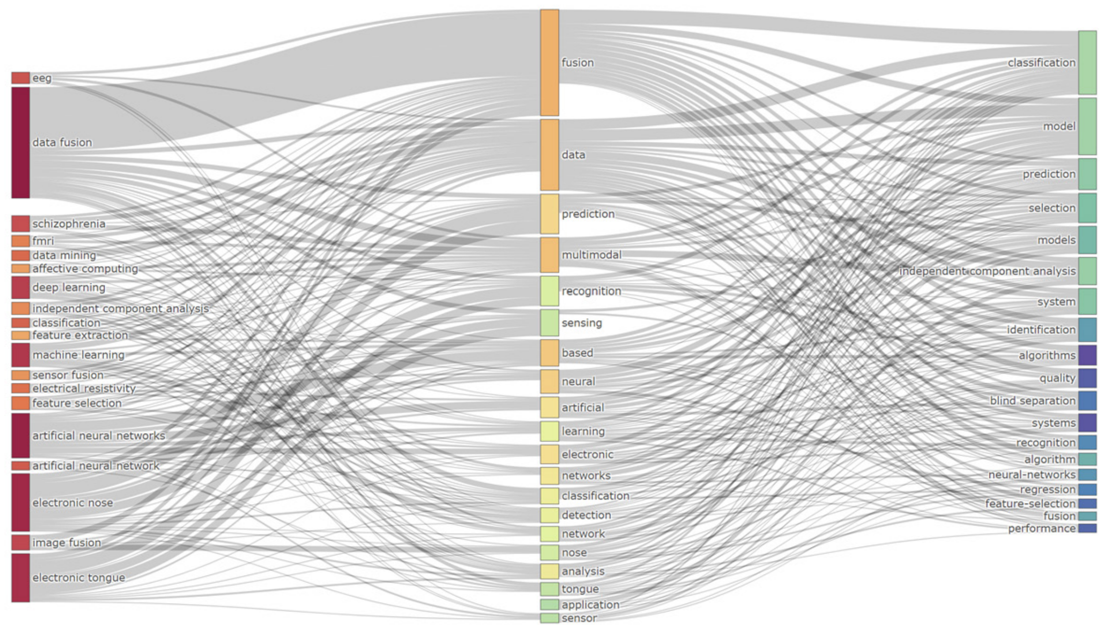

Fig 6. Word flow diagram.

IV. MULTIMODAL DATA FUSION APPLICATIONS

The following is a summary of studies in the field of data fusion across various domains that considered heterogeneous data (multimodality) and accounted for or mentioned complex conditions (dynamic environments, disturbances, or noise) within the studied system.

A. Civil Engineering

Gros et al. [33] conducted a non-destructive inspection of a carbon fiber-reinforced composite material using eddy currents and infrared thermography after subjecting it to low-energy impacts. The data were fused at the pixel level using probabilistic methods, including (1) maximum amplitude, (2) integration (AND operator), (3) averaging, (4) weighted averaging, (5) Bayesian analysis, and (6) Dempster-Shafer theory. These methods were evaluated qualitatively and quantitatively, with the authors concluding that the ensemble averaging technique is the most suitable for reducing noise in multiple images from the same source. Kurup and Griffin [34] developed a general regression neural network (GRNN) to predict soil composition. The dataset used included values obtained from cone penetration test (CPT) soundings —specifically, sleeve friction and cone resistance— as well as grain size distribution results from standard penetration test (SPT) boreholes. The researchers stated that site heterogeneity is expected to introduce noisy training data due to potential inconsistencies between CPT data from nearby soundings and soil samples recovered from boreholes, ultimately improving the model's performance. Cheng et al. [35] proposed a centralized data fusion method for monitoring workers' safe and unsafe ergonomic behaviors in construction sites. Their approach consists of two stages: (1) Posture estimation, combining data from physiological state monitoring (PSM) and video recordings. (2) Position estimation is done using ultra-wideband (UWB) devices. Due to PSM implementation in a real-world environment, posture angle and heart rate data were collected with noise. However, the researchers stated that error analysis of these measurements and their impact on activity classification were beyond the scope of their study. Additionally, UWB signals, which can contain outlier-related noise, were filtered using a Robust Kalman Filter.

Sbartaï et al. [36] conducted a study to predict concrete properties (such as water content and strength) through data fusion of non-destructive testing (NDT) measurements, including Ground Penetrating Radar (GPR), Ultrasonic Pulse Velocity, and Electrical Resistivity. The employed techniques were Artificial Neural Networks (ANN) and the Response Surface Method (RSM). Due to the limitations of using a single parameter —since samples are affected by noise caused by material variability or measurement precision— the authors indicated that by combining the three considered variables, data fusion techniques demonstrated their capability to evaluate in-situ concrete properties. Similarly, researchers in [37] developed a system for accurately assessing concrete properties. Two techniques —fuzzy neural networks (FNN) and possibility theory— fused information, as individual concrete properties influence different NDT methods used in the study. This fusion allowed for the prediction of the desired properties. Erzin et al. [38] compared the performance of three models (SOFM1, SOFM2, SOFM3) based on Self-Organizing Feature Maps (SOFM) to predict the critical safety factor (Fs) in an artificial slope subjected to seismic forces. The input parameters included internal friction angle (Φ), cohesion (c), bulk unit weight (γ), and seismic coefficient (k). According to the authors, all three models were developed using SOFM because of their ability to perform various operations, including noise rejection. Additionally, the study demonstrated the effectiveness of the SOFM1 model in estimating Fs. Ploix et al. [39] conducted a study on estimating durability indicators for undamaged concrete using a Possibility Theory and Fuzzy Set-based method. The input parameters included ultrasonic transmission velocity, Wenner resistivity, radar wave amplitude, quadripolar resistivity, the arrival time of a direct radar wave, and ultrasonic surface wave velocity. According to the authors, data fusion techniques were proposed due to the sensitivity of the measurements to variability, material heterogeneity, and experimental noise. The results demonstrated that the selected adaptive operator is suitable for this type of application.

Völker & Shokouhi [40] conducted a study on the detection of simulated panels in concrete through data fusion from three sensor modalities: (1) ground-penetrating radar data, (2) ultrasonic pulse-echo, and (3) impact-echo. The authors confirmed the superiority of the multisensor approach in complex environments compared to the best response obtained from a single sensor in the tests performed (ultrasonic sensor). The ROC curves of the Hadamard product and Dempster-Shafer techniques were more pronounced, indicating a noise reduction (inferred from the decrease in the false positive rate). Aryal et al. [41] conducted a study to prevent construction industry accidents by monitoring real-time physical fatigue. To achieve this, they trained a boosted tree classifier using 21 features obtained from (1) temperature sensors (9), (2) heart rate (4), (3) personal worker information (7), and (4) work duration (1). To eliminate large spikes in the sensor signals during the preprocessing stage, the researchers applied a Savitzky-Golay filter, a third-order one-dimensional median filter, and a moving average filter to smooth out (remove noise from) the signals. Finally, Völker & Shokouhi [42] studied identifying simulated honeycomb defects in a concrete slab. The employed data fusion techniques—DBSCAN (Density-Based Spatial Clustering of Applications with Noise), C-Means, and K-Means—enabled the clustering of features to classify results as "defective" or "non-defective." Due to noise reduction during the fusion process, the DBSCAN method achieved the best results in the tests conducted.

B. Automation

Cai B. et al. [43] proposed a fault diagnosis system for drive systems and power switch devices, specifically targeting inverters, as they are the weakest components in such systems. The diagnosis was performed using a data-driven Bayesian network and validated through simulation and experimental testing, considering sensor noise and diagnosis accuracy bias across different PCA dimensions. This methodology is suitable for various power switch-based electrical devices, including inverters and multilevel converters. Zhao et al. [44] introduced a deep fusion feature optimization method using deep neural networks to extract degradation features from bearings in centrifugal pumps within nuclear power plants, considering vibration, temperature, and seal conditions as adverse variables. The experiments in this study utilized the IEEE PHM 2012 database. Zhang et al. [45] proposed a new intelligent method for ball screw degradation recognition based on Deep Belief Networks (DBN) and implemented multisensor data fusion—using three accelerometers to capture five sets of vibration signals at multiple positions and directions. The recognition results indicate that the proposed DBN-based method can accurately classify ball screw degradation stages and is capable of adaptively extracting intrinsic degradation features. Ge Y. et al. [46] developed a gesture prediction system capable of interpreting hand movement intent and predicting the final gesture before the motion ends. The system features a glove with six flexible sensors to collect precise spatial information. A low-pass filter was applied to counteract motion instability, and the technique employed a combination of neural networks and support vector machines (SVM). Hsu Y. et al. [47] presented a portable 3D Simultaneous Localization and Mapping (SLAM) system for human navigation, adaptable to various environmental scenarios. The system integrates RGB-D SLAM and IMU/LiDAR SLAM, combining the strengths of both techniques through a sensor fusion-based SLAM algorithm. It fuses data from a Microsoft Kinect camera, a Hokuyo laser rangefinder (LRF), and inertial measurement unit (IMU) sensors to enable 3D positioning and mapping. The proposed sensor fusion system can adapt to different environments while considering perturbations such as velocity and vibration, with the LRF-based SLAM algorithm handling necessary corrections.

Kacprzynski G. et al. [48] conducted a study to identify the optimal combination of measured system data, data fusion algorithms, and associated architectures to achieve the highest overall prediction/detection confidence levels for a given application. These fusion techniques were applied to vibration features extracted during a transient failure test associated with an accelerated industrial gearbox and recorded using the Mechanical Diagnostic Test Bed (MDTB) at the ARL Laboratory of Penn State University. Ma M. et al. [49] developed a new deep multimodal fusion architecture, the Deep Coupling Autoencoder (DCAE) model, to identify joint features between vibration and acoustic signals for rotating machinery health state classification. Two experiments evaluated the proposed method's performance, demonstrating that DCAE efficiently utilizes multi-source sensor data for accurate fault diagnosis. Kumar M. and Garg D. [50] introduced specific strategies for fusing data from some of the most commonly used sensors in robotic work cells, namely vision sensors and proximity sensors. The proposed transformation and data fusion strategy was tested and validated in a robotic work cell using an ABB IRB140 six-axis articulated industrial robot equipped with a force/torque sensor, a proximity sensor, and an overhead camera. The data from these diverse but redundant sensors were fused using Bayesian inference to generate an occupancy grid model of the workspace. The Kalman filtering technique was also employed to estimate external forces acting on the robot's end effector, leveraging its underlying dynamics and the force/torque (F/T) sensor data mounted on the robot's wrist. Chen Y. et al. [51] designed a multimodal fusion method for face and fingerprint images using a block-based feature image matrix that extracts a semantic feature type from intermediate local features. They used the Variational Bayesian Extreme Learning Machine (VBELM) for recognition. This technique provides higher processing speed due to random input weights while offering superior stability and generalization by incorporating a fully non-informative Gaussian prior (noise filtering). The study utilized the FERET_FVC2002 public database containing face and fingerprint data.

Yuqin J. et al. [52] developed an intelligent inspection and monitoring model for Flexible Manufacturing Systems (FMS) based on Distributed Artificial Intelligence (DAI). The study analyzed methods for data preprocessing, feature extraction, and neural network construction. The case study focused on a textile machinery plant implementing FMS, which integrates multiple sensors to capture the condition signals of interrelated devices in the machining process. According to the authors, the oil cylinder pressure signal was periodically recorded, and noise interference was partially eliminated using a set of Rk rules, which apply a single-point fuzz method, product reasoning, and centroid method to obtain fuzzy system results, increasing the signal-to-noise ratio fivefold. The system can analyze, process, and inspect operations automatically while detecting abnormal and unforeseen failure conditions. Safizadeh M.S. and Latifi S.K. [53] proposed a fault diagnosis method for bearings using data fusion from two primary sensors: an accelerometer and a load cell. They developed a Condition-Based Monitoring (CBM) system consisting of six modules: detection, signal processing, feature extraction, classification, high-level fusion, and decision-making. The acceleration and load signals were acquired using a test bench, and the K-Nearest Neighbor (KNN) classifier was used to identify the bearing's condition based on vibration and load signals. Vibration measurements were conducted using a piezoelectric accelerometer (IMI Sensors 608A111) with a natural undamped frequency of 10 kHz. The test results demonstrated that the load cell is highly effective in distinguishing between healthy and defective ball bearings, while the accelerometer is particularly useful for pinpointing the fault location.

C. Robotics

Ahmad et al. [54] conducted a quantitative comparison of four segmentation methods for autonomous horizon line detection using Average Pixel Accuracy and Average Absolute Pixel Distance. Two of these methods belong to the deep neural network field, while the other two fall into the feature learning and classifier training by patches category. These techniques were trained and tested on a large dataset (GeoPose3K), which includes images with field-of-view parameters, camera position, and camera orientation while considering perturbations such as varying weather conditions and lighting. Tseng et al. [55] developed a system that enables a robot to infer social situations and determine whether it can approach and interact with a group of people. Since human detection in complex environments often results in false positives, the authors implemented data fusion through covariance intersection using two sensors: (a) Laser Range Finder (SICK LMS 100) and (b) RGB-D camera (ASUS Xtion Pro). Reina et al. [56] developed a multimodal system (stereo vision, LiDAR, radar, and thermography) for agricultural vehicle applications, considering the variability of field conditions such as terrain diversity and ambient lighting conditions. The researchers demonstrated that (1) stereo vision can help overcome LiDAR limitations, such as low acquisition frequency and data sparsity; (2) LiDAR helps mitigate vision limitations caused by reconstruction errors due to poor lighting; (3) radar measurements can define regions of interest in a stereo vision-generated point cloud; and (4) stereo vision and thermography measurements provide features for detecting living beings and obstacles (such as rocks and poles). Palumbo et al. [57] presented a system for daily activity recognition using decision tree-based fusion of implicit disturbances detected by wireless sensors placed in the environment and inertial sensor data from a smartphone. The researchers performed feature extraction in the time domain to compress the acquired time series, eliminate noise, and remove correlations.

Howard & Seraji [58] presented a data fusion system (fuzzy logic) for safe terrain classification using information from LiDAR, radar, and camera sensors aboard a spacecraft during descent. In the developed simulation—which accounted for spacecraft conditions, terrain type, and sensor parameters such as noise—the researchers concluded that the applied technique is ideal for this scenario, as it effectively handles imprecision and ambiguities in terrain data caused by spacecraft vibrations, among other factors. Y. Cheng & Zhang [59] developed an algorithm for obstacle avoidance in an unmanned marine vessel. The proposed system (CDRLOA) consists of (1) a data fusion module (convolutional neural network), (2) a decision-making module (deep reinforcement learning), and (3) an avoidance reward function. The authors conducted various experiments to validate the algorithm's effectiveness, including scenarios with unknown dynamic environments. Chilian et al. [60] presented a data fusion system (indirect feedback information filter) for the DLR Crawler six-legged robot. This algorithm integrates measurements from (1) an inertial measurement unit (IMU), (2) joint angle sensors in the robot's legs, and (3) a stereo camera. The authors accounted for odometry errors in the filtering process and successfully tested the proposed algorithm. Kubelka et al. [61] improved the accuracy and reliability of their multimodal data fusion system —combining (a) Xsens MTi-G IMU, (b) track velocity encoders, (c) a Point Grey Ladybug3 camera, and (d) a SICK LMS-151 sensor— using an extended Kalman filter for a mobile robot operating in complex environments. The researchers evaluated their system in both indoor and outdoor scenarios, considering conditions such as (a) slippery slopes, (b) unstable surfaces, (c) obstacles, and (d) artificial sensor limitations. Maatoug et al. [63] conducted a study on wheelchair localization. The simulation included ultrasonic sensors and incremental encoders, with sensor fusion performed via an extended Kalman filter, accounting for sensor noise. Qiu et al. [63] presented an autonomous landing control system for an unmanned aerial vehicle (UAV) during the docking process on an electrical tower. An extended Kalman filter was employed to estimate the vehicle's position by fusing camera, radar, and IMU data. The researchers indicated that this data fusion technique effectively reduces Pan-Tilt fluctuations and flight noise, improving landing stability.

Sumanarathna et al. [64] described the trajectory of a robot along a predefined path. Considering real-world implementation conditions, the researchers stated that relying on a single sensor for such applications is not feasible. Therefore, they implemented a data fusion technique based on Kalman filters to improve the robot's position estimation by combining information from (a) GPS, (b) encoders, and (c) an accelerometer. Novak et al. [65] conducted a study comparing the performance of various technologies aimed at predicting human reach movement targets. Their experiments combined different modalities using linear discriminant analysis, including (1) camera-based eye tracking, (2) electromyography, (3) electroencephalography, (4) electrooculography, (5) hand position, and (6) user preferences. Additionally, various analog filters were employed to eliminate sensor noise. Yang et al. [66] proposed a data fusion system integrating (a) a magnetic encoder and (b) an analog speed sensor for precision motion control of a linear motor. Position estimation weights were calculated using a radial basis function (RBF) neural network, which accounted for sensor velocity and noise. The researchers reported improved performance compared to single-sensor approaches and existing multimodal methods. Shamwell et al. [67] developed a study to estimate the state of a robot with size, weight, and power (SWaP) constraints using an unsupervised deep convolutional-deconvolutional fusion technique, Multi-Hypothesis DeepEfference 8 (MHDE). This approach intelligently integrates noisy sensor data (evaluated under four noise conditions) to generate probable pixel density hypotheses. The authors assessed their technique using the KITTI Odometry dataset. By comparing their results with DeepEfference and DeepMatching techniques, as well as average pixel error and execution time metrics, they demonstrated the superiority of their method, as reported in their findings.

Finally, metrics obtained from the reviewed documents that considered heterogeneity and adverse conditions are presented: (a) Description of adverse conditions (Fig 7), (b) DF level (Fig 8), (c) Modalities (Fig 9), (d) Temporality and dimensionality (Fig 10), (e) Classification of DF technique (Fig 11), and (f) Purpose (Fig 12).

Figure 7 provides a visual representation highlighting the percentage of documents addressing some form of adverse condition (49%) compared to those that do not (51%).

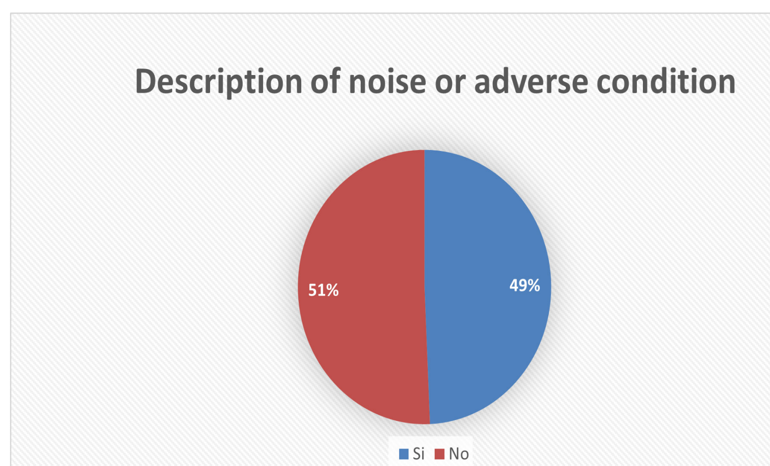

Fig 7. Description of noise or adverse condition.

Fig 8 shows the percentage of documents according to the classification of the fusion level: data-level fusion (26%), feature-level fusion (65%), and decision-level fusion (9%).

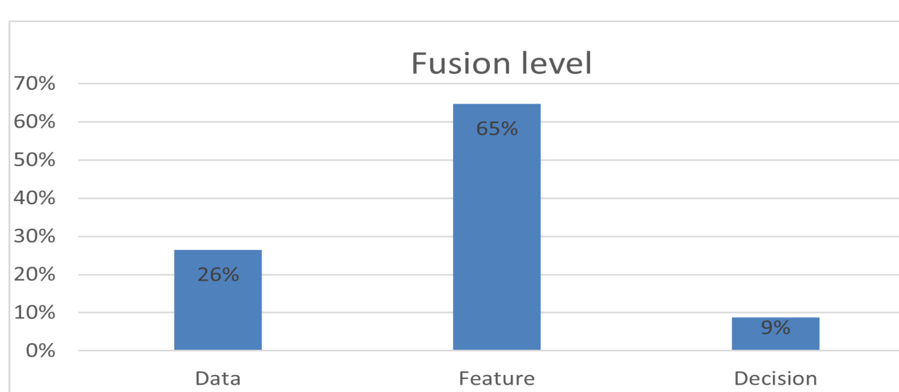

Fig 8. Fusion level.

Fig 9 represents the number of modalities used in the reviewed studies.

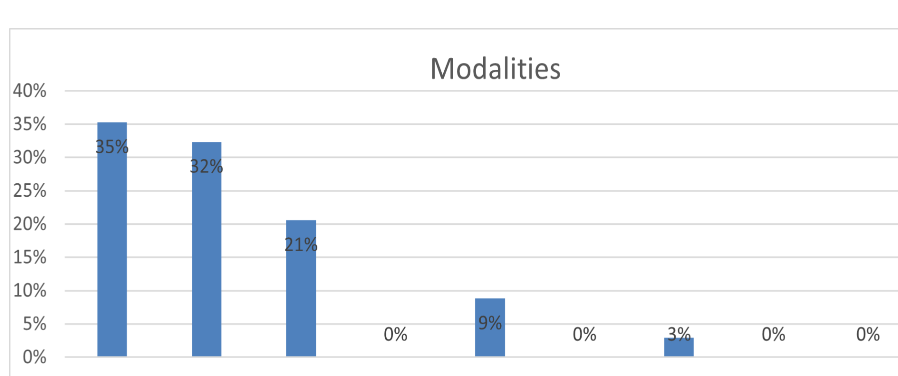

Fig 9. Modalities.

Fig 10 illustrates the percentage of documents that conducted DF with a specific dimensionality and temporality.

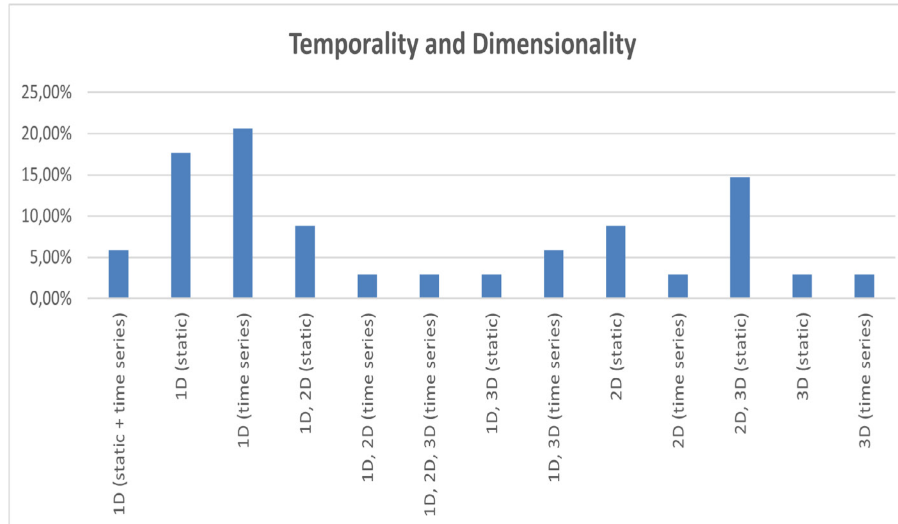

Fig 10. Temporality and Dimensionality.

Fig 11 provides a visual representation of the percentage of documents that conducted DF according to a classification into three groups, defined based on references [11] and [31] (see Table II).

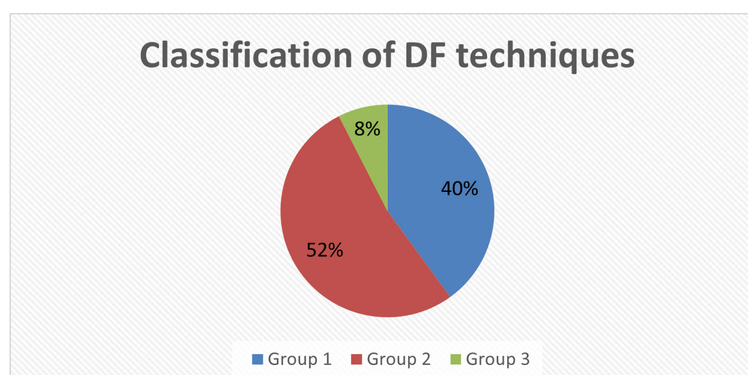

Fig 11. Classification of DF techniques.

TABLE II. DATA FUSION CLASSIFICATION

Reference	Group 1	Group 2	Group 3
[11]	Recursive operators based on probability and Bayesian analysis	Artificial intelligence based on neural networks and fuzzy logic	Dempster-Shafer theory based on evidence theory
[31]	Probability-based methods including Bayesian analysis, statistics, and recursive operators	Artificial intelligence-based techniques, including classical machine learning, fuzzy logic, artificial neural networks, and genetic evaluation	Data fusion methods based on evidence theory

Fig 12 illustrates the percentage of documents that conducted DF based on their purpose: regression/estimation (63%), classification (35%), clustering (1%), and dimensionality reduction (1%).

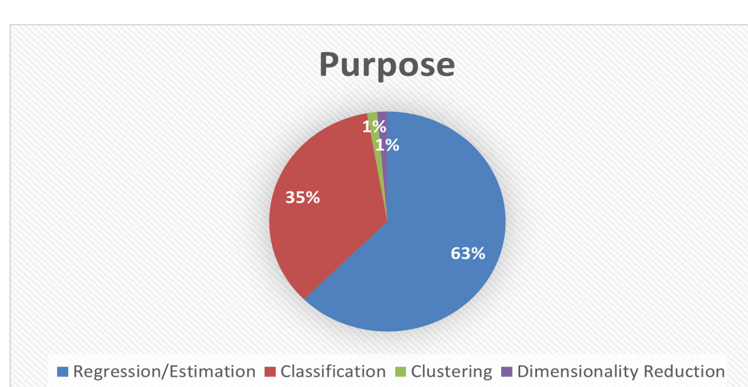

Fig 12. Purpose.

V. DISCUSSION AND FUTURE WORK

A. Discussion

First, according to the collected information, 49% of the reviewed documents described or defined some adverse condition for their system. However, only 28% conducted any analysis related to their techniques' ability to reject noise or the benefits of considering such conditions in their datasets. Among the reviewed areas, automation studies had the highest percentage of papers describing some adverse condition (78%). In comparison, robotics had the highest percentage of documents that performed analyses regarding these conditions (56%).

Second, the most commonly used fusion level (see Section III – Subsection C) is feature-level fusion (65%), which contradicts the findings of Bokade et al. [2]. This discrepancy may be due to either of the following reasons: (a) Bokade et al. considered a broader range of application domains (such as 2D MMDF in various applications). (b) The lack of theoretical development and integration across disciplines may result in a conceptual difference. In applications where the information does not originate from sensors, confusion may arise in classifying DF according to the fusion level.

Third, regarding data acquisition modes (see Section III – Subsection B): (a) The number of modalities used in the studies is generally low, with 88% of the documents utilizing between 2 and 4 modalities. (b) The most commonly used data dimensionality is 1D (68%). (c) Approximately 62% of the studies used static data. Notably, 39% of the authors considered multiple dimensionalities within a single study, while only about 6% of the studies incorporated both temporality conditions.

Fourth, based on references [11] and [31] (see Section III—Subsection C), which classify fusion methods according to mathematical techniques, Table II proposes a grouping of methods. 52% of the researchers employed techniques from Group 2 (AI-based methods, including classical machine learning, artificial neural networks, fuzzy logic, and genetic evaluation). 40% used techniques from Group 1 (probabilistic methods, including Bayesian analysis, recursive operators, and statistics). 8% used techniques from Group 3 (Dempster-Shafer theory). Regarding the purpose of data fusion (see Section III – Subsection A), regression (63%) and classification (35%) were the most common objectives in DF studies. However, clustering and dimensionality reduction accounted for only 1% each. Notably, classification objectives were dominant in automation-related studies (65%).

B. Future Work

First, it is crucial to continue developing an interdisciplinary vocabulary that facilitates both literature searches and understanding within the DF field. This will promote greater clarity and consistency in communication among researchers, ultimately fostering the advancement of knowledge in this rapidly evolving area.

Second, to ensure effective communication of developments in DF, authors should provide detailed information in their publications, including: The number of modalities (or sensors used), the dimensionality of the data (1D, 2D, or 3D), the temporality (static or time series) and the fusion level (data, feature, or decision). Additionally, other relevant DF classifications should be reported as appropriate.

Third, when considering real-world applications, researchers must clearly and comprehensively describe the type of adverse conditions affecting the studied system. These may include noise, disturbances, dust presence, variable lighting, fog, or other non-ideal environmental factors. Moreover, thorough analyses should be conducted based on the specific techniques employed in each study.

Fourth, to further advance the DF field, it is recommended to explore applications involving higher-dimensional data, such as 3D data, and to consider a broader range of modalities (more than 4), surpassing the traditionally used quantity. Additionally, researchers should explore variations in other aspects, such as temporality and the purpose of DF, to open new perspectives and expand the existing body of knowledge in this research domain.

VI. CONCLUSIONS

This article presents a literature review on DF, aiming to compare studies across various application domains —automation, robotics, and civil engineering— that considered information heterogeneity and adverse conditions. Additionally, key definitions and metrics extracted from the bibliographic search were provided.

According to the reviewed literature, most studies lack an analysis of the system's noise rejection capability (72%). The number of modalities used is relatively low (2 to 4 modalities), and the most commonly employed approaches include static temporality (62%), 1D data dimensionality (68%), feature-level fusion (65%), and regression/estimation as the primary fusion purpose (63%).

It is recommended that we continue advancing the development of an interdisciplinary vocabulary in the DF field and further explore the application of these techniques in real-world environments or scenarios involving adverse conditions, disturbances, or noise. Furthermore, investigating applications that involve higher-dimensional data and a broader range of modalities is encouraged, as this could provide valuable insights for addressing specific challenges in the field.

CRediT AUTHORSHIP CONTRIBUTION STATEMENT

D. Quintero-Bernal: Conceptualization, Methodology, Investigation, Writing – Original Draft. **H. Kaschel-Cárcamo:** Supervision, Funding Acquisition, Visualization, Review and Editing of the Manuscript. **J. Kern-Molina:** Supervision, Project Administration, Funding Acquisition, Visualization, Review and Editing of the Manuscript.

FUNDING

This work was supported by the DICYT project (Code 062113KC), the Agencia Nacional de Investigación y Desarrollo de Chile (ANID) through the IDeA I + D project ID21I10087, and the Vicerrectoría de Investigación, Innovación y Creación at the Universidad de Santiago de Chile (USACH).

ACKNOWLEDGMENTS

We acknowledge and appreciate the support received from ANID-Subdirección de Capital Humano/Doctorado Nacional/2022-21220739 for this research.

REFERENCES

[1] T. Meng, X. Jing, Z. Yan, and W. Pedrycz, “A survey on machine learning for data fusion,” *Information Fusion*, vol. 57, pp. 115–129, 2020, doi: https://doi.org/10.1016/j.inffus.2019.12.001.

[2] R. Bokade et al., “A cross-disciplinary comparison of multimodal data fusion approaches and applications: Accelerating learning through trans-disciplinary information sharing,” *Expert Syst Appl*, vol. 165, p. 113885, 2021, doi: https://doi.org/10.1016/j.eswa.2020.113885.

[3] J. Zhou, X. Hong, and P. Jin, “Information Fusion for Multi-Source Material Data: Progress and Challenges,” *Applied Sciences*, vol. 9, no. 17, 2019, doi: 10.3390/app9173473.

[4] H. Zhang, H. Xu, X. Tian, J. Jiang, and J. Ma, “Image fusion meets deep learning: A survey and perspective,” *Information Fusion*, vol. 76, pp. 323–336, 2021, doi: https://doi.org/10.1016/j.inffus.2021.06.008.

[5] J. Ma, Y. Ma, and C. Li, “Infrared and visible image fusion methods and applications: A survey,” *Information Fusion*, vol. 45, pp. 153–178, 2019, doi: https://doi.org/10.1016/j.inffus.2018.02.004.

[6] E. Blasch et al., “Machine Learning/Artificial Intelligence for Sensor Data Fusion–Opportunities and Challenges,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 36, no. 7, pp. 80–93, 2021, doi: 10.1109/MAES.2020.3049030.

[7] X. Cui and D. Niu, “A Study on Multimodal Data Fusion Based on Data Space for Power Equipment Identification,” in 2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2022, pp. 471–476. doi: 10.1109/CSCWD54268.2022.9776174.

[8] G. K. Canalle, A. C. Salgado, and B. F. Loscio, “A survey on data fusion: what for? in what form? what is next?,” *J Intell Inf Syst*, vol. 57, no. 1, pp. 25–50, 2021, doi: 10.1007/s10844-020-00627-4.

[9] Y. Zhang, D. Sidibé, O. Morel, and F. Mériaudeau, "Deep multimodal fusion for semantic image segmentation: A survey," *Image Vis Comput*, vol. 105, p. 104042, 2021, doi: https://doi.org/10.1016/j.imavis.2020.104042.

[10] C. Zhang, Z. Yang, X. He, and L. Deng, "Multimodal Intelligence: Representation Learning, Information Fusion, and Applications," *IEEE Journal on Selected Topics in Signal Processing*, vol. 14, no. 3, pp. 478–493, 2020, doi: 10.1109/JSTSP.2020.2987728.

[11] I. Ullah and H. Y. Youn, "Intelligent Data Fusion for Smart IoT Environment: A Survey," *Wirel Pers Commun*, vol. 114, no. 1, pp. 409–430, 2020, doi: 10.1007/s11277-020-07369-0.

[12] P. Ghamisi et al., "Multisource and Multitemporal Data Fusion in Remote Sensing: A Comprehensive Review of the State of the Art," *IEEE Geosci Remote Sens Mag*, vol. 7, no. 1, pp. 6–39, 2019, doi: 10.1109/MGRS.2018.2890023.

[13] S. Singh, N. Mittal, and H. Singh, "Review of Various Image Fusion Algorithms and Image Fusion Performance Metric," *Archives of Computational Methods in Engineering*, vol. 28, no. 5, pp. 3645–3659, 2021, doi: 10.1007/s11831-020-09518-x.

[14] H. Kaur, D. Koundal, and V. Kadyan, "Image Fusion Techniques: A Survey," *Archives of Computational Methods in Engineering*, vol. 28, no. 7, pp. 4425–4447, 2021, doi: 10.1007/s11831-021-09540-7.

[15] S. Li, X. Kang, L. Fang, J. Hu, and H. Yin, "Pixel-level image fusion: A survey of the state of the art," *Information Fusion*, vol. 33, pp. 100–112, 2017, doi: https://doi.org/10.1016/j.inffus.2016.05.004.

[16] Y. Liu, X. Chen, Z. Wang, Z. J. Wang, R. K. Ward, and X. Wang, "Deep learning for pixel-level image fusion: Recent advances and future prospects," *Information Fusion*, vol. 42, pp. 158–173, 2018, doi: https://doi.org/10.1016/j.inffus.2017.10.007.

[17] B. P. L. Lau et al., "A survey of data fusion in smart city applications," *Information Fusion*, vol. 52, pp. 357–374, 2019, doi: https://doi.org/10.1016/j.inffus.2019.05.004.

[18] A. A. Aguileta, R. F. Brena, O. Mayora, E. Molino-Minero-re, and L. A. Trejo, "Multi-sensor fusion for activity recognition—a survey," *Sensors (Switzerland)*, vol. 19, no. 17, pp. 1–41, 2019, doi: 10.3390/s19173808.

[19] D. Li, Z. Song, C. Quan, X. Xu, and C. Liu, "Recent advances in image fusion technology in agriculture," *Comput Electron Agric*, vol. 191, p. 106491, 2021, doi: https://doi.org/10.1016/j.compag.2021.106491.

[20] T. Zhou, S. Ruan, and S. Canu, "A review: Deep learning for medical image segmentation using multi-modality fusion," *Array*, vol. 3–4, p. 100004, 2019, doi: https://doi.org/10.1016/j.array.2019.100004.

[21] H. Hermessi, O. Mourali, and E. Zagrouba, "Multimodal medical image fusion review: Theoretical background and recent advances," *Signal Processing*, vol. 183, p. 108036, 2021, doi: https://doi.org/10.1016/j.sigpro.2021.108036.

[22] O. S. Faragallah et al., "A Comprehensive Survey Analysis for Present Solutions of Medical Image Fusion and Future Directions," *IEEE Access*, vol. 9, pp. 11358–11371, 2021, doi: 10.1109/ACCESS.2020.3048315.

[23] B. Huang, F. Yang, M. Yin, X. Mo, and C. Zhong, "A Review of Multimodal Medical Image Fusion Techniques," *Comput Math Methods Med*, vol. 2020, 2020, doi: 10.1155/2020/8279342.

[24] S. P. Yadav and S. Yadav, "Image fusion using hybrid methods in multimodality medical images," *Med Biol Eng Comput*, vol. 58, no. 4, pp. 669–687, 2020, doi: 10.1007/s11517-020-02136-6.

[25] C. Sun, C. Zhang, and N. Xiong, "Infrared and Visible Image Fusion Techniques Based on Deep Learning: A Review," *Electronics (Basel)*, vol. 9, no. 12, 2020, doi: 10.3390/electronics9122162.

[26] R. Dian, S. Li, B. Sun, and A. Guo, "Recent advances and new guidelines on hyperspectral and multispectral image fusion," *Information Fusion*, vol. 69, pp. 40–51, 2021, doi: https://doi.org/10.1016/j.inffus.2020.11.001.

[27] F. Dadrass Javan, F. Samadzadegan, S. Mehravar, A. Toosi, R. Khatami, and A. Stein, "A review of image fusion techniques for pan-sharpening of high-resolution satellite imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 171, no. October 2020, pp. 101–117, 2021, doi: 10.1016/j.isprsjprs.2020.11.001.

[28] U. Baroudi, A. A. Al-Roubaiey, and A. Devendiran, "Pipeline Leak Detection Systems and Data Fusion: A Survey," *IEEE Access*, vol. 7, pp. 97426–97439, 2019, doi: 10.1109/ACCESS.2019.2928487.

[29] Y. Cui et al., "Deep Learning for Image and Point Cloud Fusion in Autonomous Driving: A Review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 2, pp. 722–739, 2022, doi: 10.1109/TITS.2020.3023541.

[30] D. J. Yeong, G. Velasco-hernandez, J. Barry, and J. Walsh, "Sensor and sensor fusion technology in autonomous vehicles: A review," *Sensors*, vol. 21, no. 6, pp. 1–37, 2021, doi: 10.3390/s21062140.

[31] F. Alam, R. Mehmood, I. Katib, N. N. Albogami, and A. Albeshri, "Data Fusion and IoT for Smart Ubiquitous Environments: A Survey," IEEE Access, vol. 5, pp. 9533–9554, 2017, doi: 10.1109/ACCESS.2017.2697839.

[32] G. Xiao, D. P. Bavirisetti, G. Liu, and X. Zhang, Image Fusion. Singapore: Springer Singapore, 2020. doi: 10.1007/978-981-15-4867-3.

[33] X. E. Gros, J. Bousigue, and K. Takahashi, "NDT data fusion at pixel level," *NDT & E International*, vol. 32, no. 5, pp. 283–292, Jul. 1999, doi: 10.1016/S0963-8695(98)00056-5.

[34] P. U. Kurup and E. P. Griffin, "Prediction of Soil Composition from CPT Data Using General Regression Neural Network," *Journal of Computing in Civil Engineering*, vol. 20, no. 4, pp. 281–289, Jul. 2006, doi: 10.1061/(ASCE)0887-3801(2006)20:4(281).

[35] T. Cheng, G. C. Migliaccio, J. Teizer, and U. C. Gatti, "Data Fusion of Real-Time Location Sensing and Physiological Status Monitoring for Ergonomics Analysis of Construction Workers," *Journal of Computing in Civil Engineering*, vol. 27, no. 3, pp. 320–335, May 2013, doi: 10.1061/(ASCE)CP.1943-5487.0000222.

[36] Z. M. Sbartaï, S. Laurens, S. M. Elachachi, and C. Payan, "Concrete properties evaluation by statistical fusion of NDT techniques," *Constr Build Mater*, vol. 37, pp. 943–950, 2012, doi: https://doi.org/10.1016/j.conbuildmat.2012.09.064.

[37] Z. M. SBARTAÏ, V. GARNIER, and M. A. PLOIX, "DECISION MAKING FOR CONCRETE EVALUATION USING FUSION OF NDT TECHNIQUES," in Uncertainty Modeling in Knowledge Engineering and Decision Making, 2012, pp. 94–99. doi: 10.1142/9789814417747_0016.

[38] Y. Z. Erzin, M. Nikoo, M. R. Nikoo, and T. Cetin, "THE USE OF SELF-ORGANIZING FEATURE MAP NETWORKS FOR THE PREDICTION OF THE CRITICAL FACTOR OF SAFETY OF AN ARTIFICIAL SLOPE," *Neural Network World*, vol. 26, pp. 461–475, 2016, [Online]. Available: https://api.semanticscholar.org/CorpusID:41853327.

[39] M.-A. Ploix, V. Garnier, D. Breysse, and J. Moysan, "NDE data fusion to improve the evaluation of concrete structures," *NDT & E International*, vol. 44, no. 5, pp. 442–448, 2011, doi: https://doi.org/10.1016/j.ndteint.2011.04.006.

[40] C. Völker and P. Shokouhi, "Multi sensor data fusion approach for automatic honeycomb detection in concrete," *NDT & E International*, vol. 71, pp. 54–60, 2015, doi: https://doi.org/10.1016/j.ndteint.2015.01.003.

[41] A. Aryal, A. Ghahramani, and B. Becerik-Gerber, "Monitoring fatigue in construction workers using physiological measurements," *Autom Constr*, vol. 82, pp. 154–165, 2017, doi: https://doi.org/10.1016/j.autcon.2017.03.003.

[42] C. Völker and P. Shokouhi, "Clustering Based Multi Sensor Data Fusion for Honeycomb Detection in Concrete," *J Nondestr Eval*, vol. 34, no. 4, p. 32, 2015, doi: 10.1007/s10921-015-0307-7.

[43] B. Cai, Y. Zhao, H. Liu, and M. Xie, "A Data-Driven Fault Diagnosis Methodology in Three-Phase Inverters for PMSM Drive Systems," *IEEE Trans Power Electron*, vol. 32, no. 7, pp. 5590–5600, 2017, doi: 10.1109/TPEL.2016.2608842.

[44] L. Zhao and X. Wang, "A Deep Feature Optimization Fusion Method for Extracting Bearing Degradation Features," *IEEE Access*, vol. 6, pp. 19640–19653, 2018, doi: 10.1109/ACCESS.2018.2824352.

[45] L. Zhang, H. Gao, J. Wen, S. Li, and Q. Liu, "A deep learning-based recognition method for degradation monitoring of ball screw with multi-sensor data fusion," *Microelectronics Reliability*, vol. 75, pp. 215–222, 2017, doi: https://doi.org/10.1016/j.microrel.2017.03.038.

[46] Y. Ge, B. Li, W. Yan, and Y. Zhao, "A real-time gesture prediction system using neural networks and multimodal fusion based on data glove," in 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI), 2018, pp. 625–630. doi: 10.1109/ICACI.2018.8377532.

[47] Y.-W. Hsu, S.-S. Huang, and J.-W. Perng, "Application of multisensor fusion to develop a personal location and 3D mapping system," Optik (Stuttg), vol. 172, pp. 328–339, 2018, doi: https://doi.org/10.1016/j.ijleo.2018.07.029.

[48] M. J. Roemer, G. J. Kacprzynski, and R. F. Orsagh, "Assessment of data and knowledge fusion strategies for prognostics and health management," in 2001 IEEE Aerospace Conference Proceedings (Cat. No.01TH8542), 2001, pp. 2979–2988 vol.6. doi: 10.1109/AERO.2001.931318.

[49] M. Ma, C. Sun, and X. Chen, "Deep Coupling Autoencoder for Fault Diagnosis with Multimodal Sensory Data," *IEEE Trans Industr Inform*, vol. 14, no. 3, pp. 1137–1145, 2018, doi: 10.1109/TII.2018.2793246.

[50] M. Kumar and D. P. Garg, "Intelligent multi-sensor fusion techniques in flexible manufacturing workcells," in Proceedings of the 2004 American Control Conference, 2004, pp. 5375–5380 vol.6. doi: 10.23919/ACC.2004.1384707.

[51] Y. Chen, J. Yang, C. Wang, and N. Liu, "Multimodal biometrics recognition based on local fusion visual features and variational Bayesian extreme learning machine," *Expert Syst Appl*, vol. 64, pp. 93–103, 2016, doi: https://doi.org/10.1016/j.eswa.2016.07.009.

[52] J. Yuqin, W. Peixia, and L. Yue, "Study of manufacturing system based on neural network multi-sensor data fusion and its application," in IEEE International Conference on Robotics, Intelligent Systems and Signal Processing, 2003. Proceedings. 2003, 2003, pp. 1022–1026 vol.2. doi: 10.1109/RISSP.2003.1285729.

[53] M. S. Safizadeh and S. K. Latifi, "Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell," *Information Fusion*, vol. 18, pp. 1–8, 2014, doi: https://doi.org/10.1016/j.inffus.2013.10.002.

[54] T. Ahmad, P. Campr, M. Čadik, and G. Bebis, "Comparison of semantic segmentation approaches for horizon/sky line detection," in 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 4436–4443. doi: 10.1109/IJCNN.2017.7966418.

[55] S.-H. Tseng, Y. Chao, C. Lin, and L.-C. Fu, "Service robots: System design for tracking people through data fusion and initiating interaction with the human group by inferring social situations," *Rob Auton Syst*, vol. 83, pp. 188–202, 2016, doi: https://doi.org/10.1016/j.robot.2016.05.004.

[56] G. Reina, A. Milella, R. Rouveure, M. Nielsen, R. Worst, and M. R. Blas, "Ambient awareness for agricultural robotic vehicles," *Biosyst Eng*, vol. 146, pp. 114–132, 2016, doi: https://doi.org/10.1016/j.biosystemseng.2015.12.010.

[57] F. Palumbo, C. Gallicchio, R. Pucci, and A. Micheli, "Human activity recognition using multisensor data fusion based on Reservoir Computing," *J Ambient Intell Smart Environ*, vol. 8, pp. 87–107, 2016, doi: 10.3233/AIS-160372.

[58] A. Howard and H. Seraji, "Multi-sensor terrain classification for safe spacecraft landing," *IEEE Trans Aerosp Electron Syst*, vol. 40, no. 4, pp. 1122–1131, 2004, doi: 10.1109/TAES.2004.1386868.

[59] Y. Cheng and W. Zhang, "Concise deep reinforcement learning obstacle avoidance for underactuated unmanned marine vessels," *Neurocomputing*, vol. 272, pp. 63–73, 2018, doi: https://doi.org/10.1016/j.neucom.2017.06.066.

[60] A. Chilian, H. Hirschmüller, and M. Görner, "Multisensor data fusion for robust pose estimation of a six-legged walking robot," in 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2011, pp. 2497–2504. doi: 10.1109/IROS.2011.6094484.

[61] V. Kubelka, M. Reinstein, and T. Svoboda, "Improving multimodal data fusion for mobile robots by trajectory smoothing," *Rob Auton Syst*, vol. 84, pp. 88–96, 2016, doi: https://doi.org/10.1016/j.robot.2016.07.006.

[62] K. Maatoug, M. Njah, and M. Jallouli, "Multisensor data fusion for electrical wheelchair localization using extended Kalman Filter," in 2017 18th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), 2017, pp. 257–260. doi: 10.1109/STA.2017.8314970.

[63] R. Qiu, X. Miao, S. Zhuang, H. Jiang, and J. Chen, "Design and implementation of an autonomous landing control system of unmanned aerial vehicle for power line inspection," in 2017 Chinese Automation Congress (CAC), 2017, pp. 7428–7431. doi: 10.1109/CAC.2017.8244120.

[64] D. M. G. A. I. N. Sumanarathna, I. A. S. R. Senevirathna, K. L. U. Sirisena, H. G. N. Sandamali, M. B. Pillai, and A. M. H. S. Abeykoon, "Simulation of mobile robot navigation with sensor fusion on an uneven path," in 2014 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2014], 2014, pp. 388–393. doi: 10.1109/ICCPCT.2014.7054830.

[65] D. Novak, X. Omlin, R. Leins-Hess, and R. Riener, "Predicting targets of human reaching motions using different sensing technologies.," *IEEE Trans Biomed Eng*, vol. 60, no. 9, pp. 2645–2654, Sep. 2013, doi: 10.1109/TBME.2013.2262455.

[66] R. Yang, P. V. Er, Z. Wang, and K. K. Tan, "An RBF neural network approach towards precision motion system with selective sensor fusion," *Neurocomputing*, vol. 199, pp. 31–39, 2016, doi: https://doi.org/10.1016/j.neucom.2016.01.093.

[67] E. J. Shamwell, W. D. Nothwang, and D. Perlis, "A deep neural network approach to fusing vision and heteroscedastic motion estimates for low-SWaP robotic applications," in 2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), 2017, pp. 56–63. doi: 10.1109/MFI.2017.8170407.

Author 1 pursued studies in the Mechatronic Engineering program at the Technological University of Pereira, Colombia, from 2012 to 2017, where he received three Distinguished Student awards and two Honor Registrations. From 2017 to 2020, he completed postgraduate studies at the same university as a scholarship recipient (Jorge Roa Martínez Scholarship) in the Master's program in Physical Instrumentation. During his master's studies, he worked as a researcher. He is currently a scholarship recipient (BEE and ANID Scholarships) in the Ph.D. program in Engineering Sciences, specializing in Automation, at the University of Santiago, Chile. His research interests include instrumentation, automation, and control, with a particular focus on applications involving data from multiple sensors for specific objectives (data fusion). He holds a patent and is a co-author of five scientific articles. https://orcid.org/0000-0002-0745-6452

Author 2 received a Civil Electrical Engineering degree from the University of Santiago, Chile, and a Doctorate in Electrical Engineering (Dr.-Ing.) from the University of Paderborn, Germany. He is currently a Full Professor in the Department of Electrical Engineering at the University of Santiago, Chile. He has published more than 150 articles in national and international conferences and journals. His research interests include Industrial Communication Networks, Wireless Sensor Networks (WSN), Wireless Body Area Networks

(WBAN), Smart Cities, Smart Grids, Wireless Local Area Networks (WLAN), and Mobile Networks. https://orcid.org/0000-0002-0868-7834

Author 3 John Kern was born in Santiago, Chile. He received his M.Sc. Eng. and Ph.D. from the University of Santiago, Chile, in 2010 and 2013, respectively, and completed his postdoctoral studies at the same institution in 2014. His research focuses on Fault-Tolerant Systems. John Kern is currently a Professor in the Department of Electrical Engineering at the University of Santiago, Chile, specializing in automatic control and robotics. He has developed laboratories in control theory and systems, signal and communication systems, and electronics. Since 2016, he has served as the Head of the Process Control Laboratory. He is currently the Director of the Ph.D. Program in Engineering Sciences with a specialization in Automation at the University of Santiago, Chile. https://orcid.org/0000-0003-1958-7289