information

MDPI

Systematic Review

Multimodal Sensor Fusion for Indoor Localisation and Industrial Asset Tracking: A Systematic Review of Industrial and Transferable Evidence

Sidrah Shoaib and Mahsa Mohaghegh *

Department of Computer Sciences, Auckland University of Technology, Auckland 1010, New Zealand; sidrah.shoaib@aut.ac.nz
* Correspondence: mahsa.mohaghegh@aut.ac.nz

Abstract

Indoor localisation and asset tracking are central functions in industrial information systems, yet reliable positioning remains difficult in factories and warehouses where signal interference, non-line-of-sight propagation, dynamic layouts, and real-time constraints affect system performance. This systematic literature review examines multimodal sensor fusion architectures for indoor localisation and asset tracking, with emphasis on their relevance to industrial environments such as factories and warehouses. Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines, 61 peer-reviewed studies published between 2021 and 2026 were analysed, including 12 core industrial studies and 49 supporting studies from transferable indoor localisation contexts. Studies were identified through searches of ACM Digital Library, IEEE Xplore, Springer Link, Scopus, and ScienceDirect conducted in October 2026. The review synthesises evidence on radio-frequency identification (RFID), ultra-wideband (UWB), inertial measurement unit (IMU), light detection and ranging (LiDAR), visual, and hybrid sensing approaches, with particular attention to fusion architectures, information integration strategies, and algorithmic design. The findings show that fused systems generally outperform comparable standalone sensing approaches in representative studies under the evaluated experimental conditions, with reported accuracy improvements ranging from approximately 11.3% to 96% across the 61 included studies, and from 21.4% to 72.7% when restricted to the 12 core industrial studies, where baseline and fused-system results were directly comparable. This range is descriptive and condition-dependent rather than indicative of a central tendency. The strongest performance gains occur when complementary information sources are combined: radio-based systems provide absolute position cues, inertial and odometry-based methods provide continuous motion estimates, and vision or LiDAR provides spatial context. However, practical deployment remains limited by calibration effort, infrastructure cost, synchronisation complexity, map maintenance, scalability, privacy, and explainability. Future research should prioritise self-calibrating fusion architectures, edge-AI-based information processing, shared industrial benchmark datasets, explainable fusion models, and collaborative multi-asset tracking frameworks to support more reliable translation of indoor localisation methods into industrial asset-tracking applications.

Keywords: multimodal information fusion; sensor fusion; indoor localisation; asset tracking; real-time location systems; industrial IoT; NLOS mitigation; Kalman filter; data integration; systematic literature review

Check for updates

Academic Editor: Lin Chen

Received: 23 August 2026
Revised: 10 November 2026
Accepted: 10 November 2026
Published: 18 November 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.

1. Introduction

Indoor localisation and asset tracking are core functions in industrial information systems, particularly in Industry 4.0 factories and warehouses where operational decisions depend on reliable, real-time location data. These settings increasingly rely on connected sensors, data platforms, and decision-support systems to monitor assets, automate workflows, improve inventory visibility, and support safe human–machine interaction [1,2].

Real-Time Location Systems (RTLS) aim to provide this capability and have been associated with improved inventory visibility, reduced search time, and enhanced safety [3,4]. Achieving high accuracy and reliability indoors, however, remains difficult due to metallic infrastructure, signal interference, and dynamic operational conditions.

A wide range of sensing technologies has, therefore, been explored. RFID is low-cost and effective for identification but typically provides limited localisation accuracy. UWB can deliver high positioning accuracy, yet requires dedicated infrastructure and may degrade under non-line-of-sight (NLOS) conditions. Computer vision enables rich spatial understanding but depends on lighting and often requires substantial computation. IMUs and vibration sensors support motion tracking but suffer from cumulative drift over time [1,5–7].

Because no single modality is consistently robust across industrial conditions, single-sensor approaches are often insufficient. This motivates multimodal sensor fusion, where heterogeneous data streams are integrated to exploit complementarity and reduce uncertainty [7,8]. From an information-systems perspective, the challenge extends beyond sensor selection to the design of fusion architectures that can transform noisy, incomplete, and asynchronous data into reliable location intelligence.

Despite this potential, robust multimodal fusion in real-world factories and warehouses remains challenging. Signal interference, moving obstacles, and strict real-time constraints can degrade performance [2,9]. Many deployments still rely on single-sensor solutions such as RFID or UWB, limiting overall robustness and accuracy [3]. Meanwhile, the research landscape is fragmented, with limited comparative evaluation across approaches and insufficient discussion of deployment barriers such as system complexity, cost, and data integration [4].

In this review, RFID, UWB, computer vision, inertial measurement unit/pedestrian dead reckoning (IMU/PDR), vibration sensing, and light detection and ranging (LiDAR) are treated as primary modalities because they appear most frequently in industrial asset-tracking and indoor-localisation studies. Complementary modalities, including Wi-Fi, Bluetooth Low Energy (BLE), ultrasonic sensing, magnetic sensing, and odometry, are included where they provide transferable evidence on sensor complementarity, drift correction, NLOS mitigation, map matching, localisation continuity, or fusion architecture relevant to industrial asset tracking.

Existing surveys have advanced the literature on indoor positioning, RTLS, multisensor fusion, and smart manufacturing. Many, however, emphasise general indoor-localisation technologies, single sensor modalities, autonomous navigation, or smart-manufacturing data-fusion systems at a broad level. Fewer studies offer a targeted synthesis of multimodal information fusion for precise indoor industrial asset tracking that simultaneously compares sensing modalities, fusion architectures, algorithmic approaches, and real-world deployment constraints. Table 1 summarises representative surveys and positions the present review relative to this body of work.

Building on this gap, the novelty of this review lies in its focused analysis of multimodal information fusion for indoor industrial asset tracking. Rather than reviewing indoor-positioning technologies in isolation, the review examines (i) how sensing modalities are combined, (ii) how fusion architectures handle heterogeneous and asynchronous data,

(iii) which algorithmic strategies are used to improve robustness, and (iv) which practical barriers shape real-world deployment. Figure 1 illustrates this scope by linking sensing modalities, fusion approaches, algorithmic processing, and deployment outcomes.

Table 1. Comparison of representative existing surveys and the positioning of the present review.

Survey	Main Focus	Limitations Relative to This Review	Relevance to Present Review
Zafari et al. [1]	Indoor localisation systems and technologies	Provides a broad survey of indoor localisation technologies, but does not focus specifically on industrial asset tracking or multimodal fusion architectures	Establishes the broader technological background for indoor positioning systems
Zhang et al. [2]	Indoor positioning systems for warehouse environments	Focuses on warehouse positioning systems, but provides limited comparative analysis of multimodal fusion architectures for industrial asset tracking	Relevant to warehouse-specific deployment requirements and positioning challenges
Tsanousa et al. [4]	Multisensor data fusion in smart manufacturing	Addresses data-fusion solutions across smart manufacturing broadly, rather than precise indoor asset localisation and RTLS-focused tracking	Supports the role of multisensor data fusion in manufacturing and Industry 4.0 systems
Ušinskis et al. [7]	Sensor-fusion-based navigation for an autonomous mobile robot	Focuses on autonomous mobile robot navigation rather than industrial asset-tracking systems	Useful for understanding localisation, odometry integration, and mobile-platform fusion methods
Wang and Ahmad [10]	Sensor-fusion techniques for dynamic-target localisation in GPS-denied environments	Broadly reviews dynamic-target localisation in GPS-denied environments, but is not centred on indoor industrial RTLS, warehouse tracking, or asset-management applications	Relevant to sensor-fusion-based localisation where GNSS/GPS signals are unavailable
Present review	Multimodal information fusion for precise indoor industrial asset tracking	Focuses specifically on asset tracking, sensor combinations, fusion architectures, algorithms, and deployment barriers in industrial indoor environments	Provides a targeted synthesis for Industry 4.0 RTLS and asset-tracking applications

To address these gaps, this systematic literature review analyses research published between 2021 and 2026 on multimodal sensor fusion for indoor localisation and asset tracking. It considers both studies conducted in industrial environments and supporting evidence from related indoor-localisation contexts where findings are transferable.

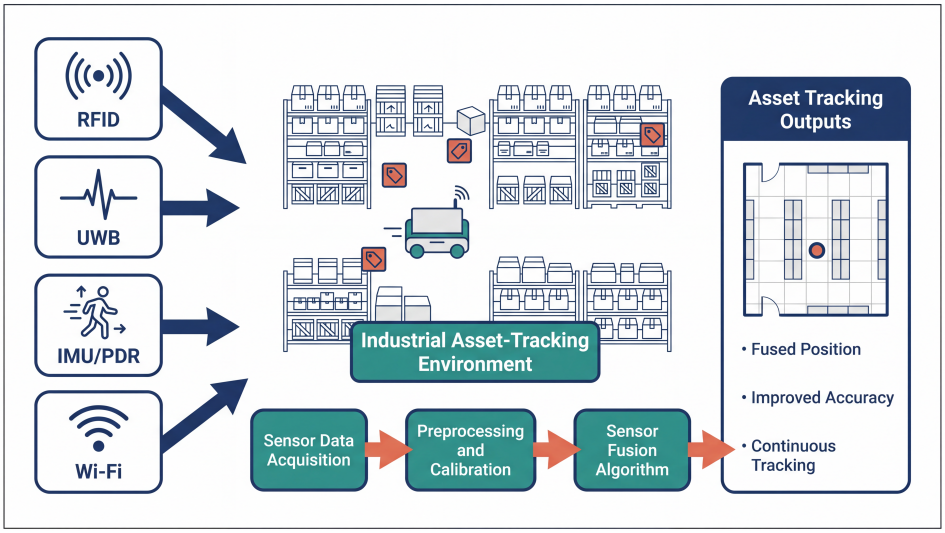


Figure 1. Multimodal information fusion for industrial localisation.

Here, multimodal information fusion is defined as the integration of heterogeneous sensing data (e.g., radio-frequency, inertial, visual, LiDAR, RFID, and odometry) to produce localisation estimates that are more accurate and robust than those obtained from any single modality. This definition highlights not only sensing hardware, but also the information-processing layer: data alignment, uncertainty handling, fusion algorithms, and the resulting decision-support value.

Accordingly, the review distinguishes between core industrial studies and transferable supporting studies, using the latter to interpret sensor complementarity, fusion design choices, algorithmic behaviour, and deployment implications.

The aim is to synthesise and critically evaluate multimodal fusion approaches for indoor localisation and asset tracking, with a particular focus on transferability to industrial environments. The review, therefore, examines sensor complementarity, fusion architectures, algorithmic strategies, reported performance improvements, operational benefits, and deployment challenges.

Research Questions

This systematic literature review is guided by the following research question:

RQ: How do multimodal sensor fusion architectures improve the accuracy, robustness, continuity, and reliability of indoor localisation and asset tracking, and how transferable are these approaches to industrial environments?

The following sub-research questions further structure the review:

Sub-RQ1: What information is provided by different sensing modalities, and what are their strengths, limitations, and complementary characteristics for indoor localisation and asset tracking?

Sub-RQ2: How do fusion architectures and algorithms combine heterogeneous sensing information to improve localisation performance compared with standalone sensing approaches?

Sub-RQ3: What operational benefits, deployment challenges, and future research directions affect the transfer of multimodal indoor localisation systems into industrial asset-tracking environments?

2. Materials and Methods

This review followed a systematic, multi-stage screening and quality-assessment protocol based on PRISMA 2020 [11] to support methodological rigour, transparency, and reproducibility. The process comprised three phases: (1) database-level filtering using predefined

inclusion/exclusion criteria, (2) title and abstract screening for thematic alignment, and (3) full-text eligibility assessment against the complete criteria set.

2.1. Review Protocol and Reporting Guideline

This systematic review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement [11]. The PRISMA 2020 checklist is provided as Supplementary Material, and the identification, screening, eligibility, and inclusion steps are summarised in the PRISMA flow diagram (Figure 2). The review was not registered in a systematic review registry, and no separate protocol was prepared or prospectively registered.

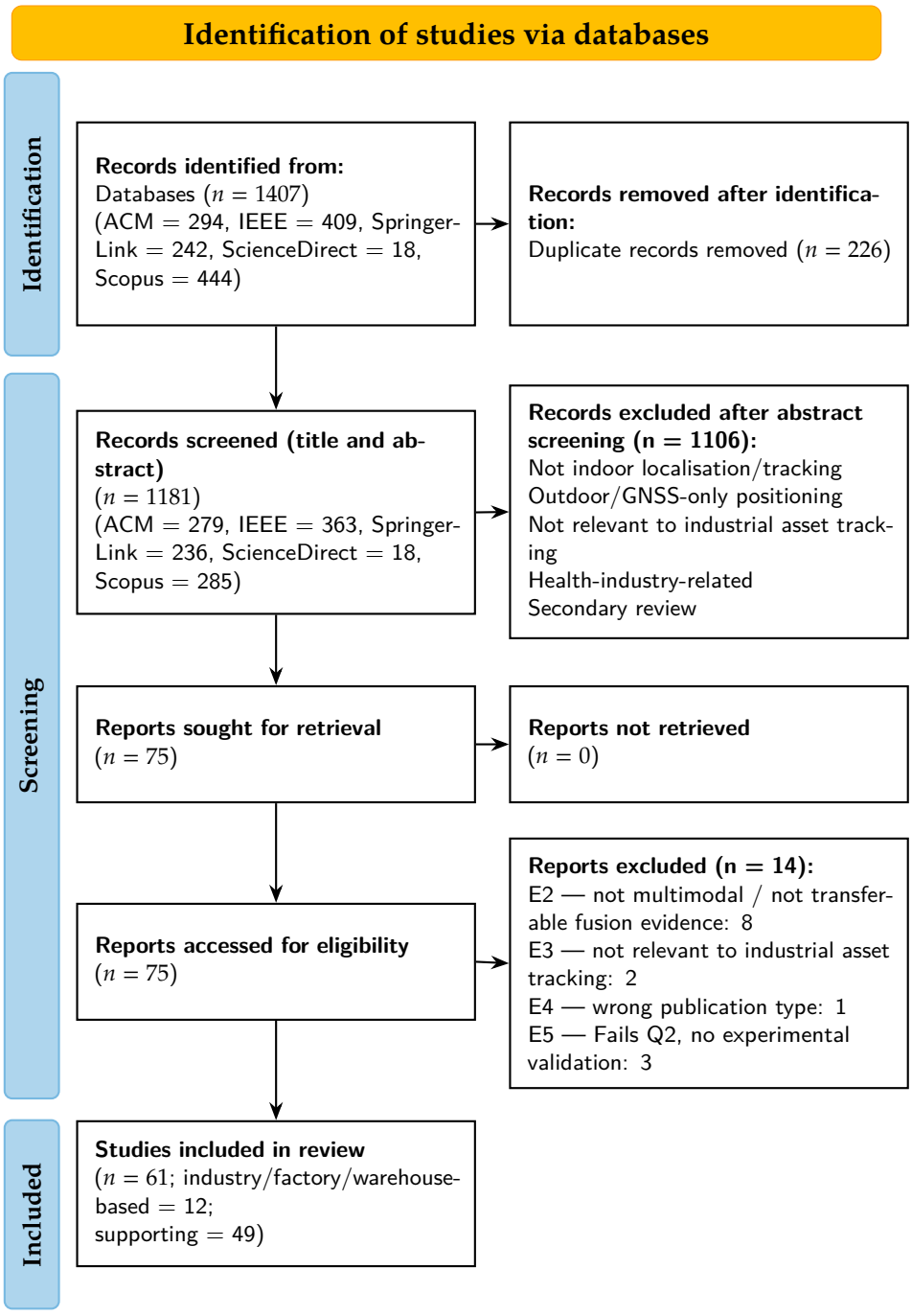


Figure 2. PRISMA 2020 flow diagram of study identification, screening, eligibility, and inclusion.

2.2. Search Strategy

A systematic database search was conducted across five peer-reviewed digital libraries: ACM Digital Library, IEEE Xplore, Springer Link, Scopus, and ScienceDirect. These databases were selected for their broad coverage of engineering, computer science, and IoT-related literature, providing representative coverage of the target research domain [12,13].

The database searches were conducted on 20 October 2026. The search strings used established terms related to sensor fusion, data fusion, indoor positioning, and asset tracking; however, the evidence was synthesised through the broader lens of multimodal information fusion because the reviewed systems integrate heterogeneous sensing data to support industrial localisation and decision-making. Table 2 reports the full search strings and the corresponding database links.

Table 2. Database search strings and result counts

Database	Search String	Records
ACM Digital Library	(“sensor fusion” OR “multi-sensor” OR “multimodal” OR “multi-modal”) AND (“asset tracking” OR “indoor tracking” OR “indoor positioning” OR RTLS) AND (factory OR warehouse OR “shop floor” OR “production line” OR “smart manufacturing” OR “industry 4.0” OR “industrial environment”) AND (RFID OR UWB OR “ultra wideband” OR vision OR camera OR IMU OR inertial OR vibration)	294
IEEE Xplore	(“sensor fusion” OR “multi-sensor” OR “multimodal” OR “multi-modal” OR “data fusion”) AND (“indoor positioning” OR “indoor localisation” OR “indoor localization” OR “indoor tracking” OR “asset tracking” OR “real-time location system” OR RTLS) AND (RFID OR “radio frequency identification” OR UWB OR “ultra wideband” OR “ultra-wideband” OR “computer vision” OR camera OR vision OR IMU OR “inertial measurement unit” OR inertial OR “vibration sensor”)	409
Springer Link	(“sensor fusion” OR “multi-sensor” OR “multimodal” OR “multi-modal” OR “data fusion”) AND (“indoor positioning” OR “indoor localization” OR “indoor localisation” OR “indoor tracking” OR “asset tracking” OR “real-time location system” OR RTLS) AND (RFID OR “radio frequency identification” OR UWB OR “ultra wideband” OR “ultra-wideband” OR “computer vision” OR camera OR vision OR IMU OR “inertial measurement unit” OR inertial OR “vibration sensor”)	242
Scopus	(“sensor fusion” OR “multi-sensor” OR “multimodal” OR “multi-modal” OR “data fusion”) AND (“indoor positioning” OR “indoor localization” OR “indoor localisation” OR “indoor tracking” OR “asset tracking” OR “real-time location system” OR RTLS) AND (RFID OR “radio frequency identification” OR UWB OR “ultra wideband” OR “ultra-wideband” OR “computer vision” OR camera OR vision OR IMU OR “inertial measurement unit” OR inertial OR “vibration sensor”)	444
ScienceDirect	(“sensor fusion” OR “multi-sensor” OR “multimodal” OR “data fusion”)AND (“indoor positioning” OR “indoor tracking” OR “asset tracking” OR “real-time location system”)	18
Total		**1407**

Search fields were adapted to each platform’s query syntax. For Scopus and ScienceDirect, searches were restricted to title, abstract, and keyword fields (using `TITLE-ABS-KEY` in Scopus and the platform’s default scope in ScienceDirect). IEEE Xplore, Springer Link, and the ACM Digital Library do not expose equivalent abstract/keyword-only filters in their standard interfaces; these searches were, therefore, executed against the full record (including full text). To mitigate spurious matches from full-record retrieval, all results were subsequently screened using titles and abstracts, ensuring consistent eligibility assessment across platforms.

2.3. Inclusion and Exclusion Criteria

An initial pool of 1407 records was retrieved from five databases. Table 3 lists the inclusion and exclusion criteria applied throughout screening.

The 2021–2026 publication window was chosen to capture recent developments in multimodal fusion architectures for industrial indoor localisation. This timeframe follows the release of IEEE 802.15.4z-2020, which enhanced UWB physical layers and ranging techniques for more accurate and reliable distance estimation [14], and coincides with increased use of edge computing, deep learning, graph-based optimisation, and real-time fusion in industrial and GPS-denied indoor environments. Earlier foundational studies were used only to support background context and were not included in the systematic evidence synthesis.

Table 3. Inclusion and exclusion criteria.

Criterion	Inclusion	Exclusion
Publication period	Studies published between 2021 and 2026.	Studies published before 2021.
Publication type	Peer-reviewed journal articles and peer-reviewed full-text research articles	Theses, editorials, grey literature, non-peer-reviewed reports, and papers without accessible full text.
Language	English-language publications.	Non-English publications.
Subject area	Engineering, computing, robotics, industrial IoT, and related technical domains.	Non-technical domains or studies outside the scope of sensor-based indoor localisation and tracking.
Topic focus	Studies addressing multimodal information fusion for indoor localisation, indoor positioning, RTLS, robot localisation, asset tracking, or transferable indoor localisation contexts	Single-sensor-only studies without fusion, outdoor-only positioning studies, or studies unrelated to localisation, tracking, or sensor fusion.
Sensor modalities	Studies involving at least one target modality, including RFID, UWB, Wi-Fi, computer vision, LiDAR, IMU/PDR, BLE, magnetic, ultrasonic, odometry, or vibration-based sensing.	Studies not involving any relevant sensing modality for indoor localisation.
Environment	Factory, warehouse, shop floor, production line, plant factory, industrial logistics setting, or transferable indoor environments such as corridors, offices, laboratories, and robotic testbeds	Outdoor-only environments or indoor contexts with no transferable relevance to industrial localisation.

2.3.1. Stage 1: Database-Level Filtering

In Stage 1, database-level filtering was applied using built-in platform filters: publication year (2021–2026), document type (peer-reviewed journal article), language (English), and subject area (computing and engineering). Because these filters were applied before export, removals at this stage are reflected in the 1407 identified records rather than reported as a separate screening step.

2.3.2. Stage 2: Deduplication

In Stage 2, duplicate records arising from overlap across databases were identified and removed using Mendeley's duplicate detection (matching title, author, and year). Detection was automated, and each flagged pair was manually confirmed prior to removal. This stage removed 226 duplicates.

2.3.3. Stage 3: Title and Abstract Screening

In Stage 3, titles and abstracts were screened by the first reviewer to assess alignment with the research questions. Screening was performed in Excel. Records were excluded using the following categories:

- Not indoor localisation or tracking
- Outdoor or GNSS-only positioning

- Not relevant to industrial asset tracking
- Health-industry-related
- Secondary review

Exclusion reasons were recorded as categories; per-category counts were not tracked at this stage. Of the 1181 records screened, 1106 were excluded, leaving 75 for full-text assessment (Figure 2).

2.3.4. Stage 4: Full-Text Eligibility Assessment

In Stage 4, the full text of each retained article was retrieved and assessed against the full inclusion and exclusion criteria, including the fusion architecture/algorithm, the deployment environment, and evidence of empirical validation. This was the only stage at which the complete criteria set was applied. Exclusion reasons were recorded per criterion as mentioned in Figure 2. Of the 75 reports assessed, 61 met the eligibility criteria and were included in the final synthesis. Table 4 summarises the record counts at each screening stage for each database.

All eligibility decisions, including borderline transferability judgements, were made by the first reviewer. To limit subjectivity, each borderline case was documented against the operational criteria (Section 2.5) and retained in the coding sheet. Where a record was ambiguous, it was provisionally coded and re-examined against the four mechanisms before a final decision. Independent second-reviewer screening of all records was not performed, and this is acknowledged as a limitation in Section 2.6.

Table 4. Record counts at each screening stage per database.

Database	Database Level Filtering	After Deduplication	After Title/Abstract Screening
ACM	294	279	2
IEEE Xplore	409	363	41
Springer Nature Link	242	236	9
Scopus	444	285	9
Science Direct	18	18	14
Total	**1407**	**1181**	**75**

Figure 2 presents the complete PRISMA 2020 flow diagram illustrating the flow of records through each phase of the review.

2.4. Data Extraction and Synthesis

For each study that passed full-text eligibility assessment, a structured data-extraction form was used to capture the following: (1) bibliographic metadata (2) sensing modalities and data sources (3) fusion architecture and algorithm (4) target environment and asset type (5) reported accuracy metrics (6) reported limitations or failure conditions and (7) findings relevant to the research questions [12].

2.5. Evidence Coding and Transferability Classification

Each included study was coded using six attributes: sensing modalities, fusion architecture, fusion algorithm, target platform, evaluation environment, and reported positioning performance. The evidence base was then divided into two categories: core industrial evidence and transferable supporting evidence, with a separate set of rubric fields recorded to document the basis of each decision.

Core industrial. A study was classified as core industrial only when its evaluation environment was itself industrial, such as a factory, warehouse, plant factory, industrial workshop, logistics space, or AGV operating scenario.

Transferable supporting. All other studies were classified as transferable supporting when they provided relevant technical evidence from related indoor-localisation contexts, such as robotic platforms, laboratories, offices, pedestrian localisation, smartphone-based tracking, QR-based localisation, or multi-object tracking. Inclusion required the study to demonstrate at least one of four mechanisms: (i) absolute position correction using an external reference; (ii) drift or uncertainty reduction through fusion; (iii) NLOS or multipath mitigation; or (iv) map construction, map matching, or scalable multi-asset localisation. Studies satisfying none of (i)–(iv) were excluded as non-transferable (E3).

Recorded rubric fields. For each study, the classification additionally recorded the evaluation environment, tracked target or asset, sensing infrastructure, whether NLOS or metallic conditions were tested, operational scale, and any reported real-time constraints. This separates industrial evaluation from laboratory-derived evidence and allows the two groups to be reported and compared independently.

Reviews and surveys were excluded during title/abstract screening and were used only for background framing, not as empirical evidence. The assigned category and the recorded rubric fields for every included study are reported in Supplementary Table S2.

2.6. Study Quality Assessment

To support the reliability of the included evidence, each study was assessed against four quality criteria: (1) clarity of the proposed methodology, (2) presence of experimental validation, (3) reporting of quantitative performance metrics, and (4) relevance to industrial localisation or transferable indoor-localisation contexts. These criteria were adapted from prior engineering systematic reviews and were used because no widely adopted risk-of-bias tool exists for multimodal localisation studies.

Studies meeting all four criteria were considered high quality. Studies lacking experimental validation or quantitative evaluation were excluded during full-text assessment. The quality assessment informed selection and interpretation rather than being used as a weighted scoring scheme. Reviews and surveys did not satisfy Q2 or Q3 and were excluded during title/abstract screening. Three reports were excluded at the full-text stage for failing Q2 because their evaluations were simulation-based rather than empirical.

Quality assessment was not applied solely as an inclusion threshold. Because controlled or simulated evaluations can overstate achievable accuracy, performance is reported separately for core industrial and transferable supporting studies, and conclusions about industrial deployment are drawn only from the 12 core industrial studies. A study-level quality (Q1–Q4 for every included study) is provided in Supplementary Table S2.

Full-text eligibility assessment was performed by the first reviewer, with the second reviewer checking final inclusion decisions. Independent dual screening was not conducted at the record-screening stage; this is acknowledged as a limitation of the review.

2.7. Study-Level Risk of Bias

Formal risk-of-bias tools are not directly applicable to many engineering studies of indoor localisation systems; nevertheless, potential sources of bias were considered during evidence synthesis. The most common risk was evaluation in controlled laboratory or simulated environments rather than in real industrial deployments. Only 12 of the 61 included studies evaluated their methods directly in factories, warehouses, plant factories, or industrial logistics settings; the remainder were conducted in transferable indoor environments such as offices, laboratories, corridors, or robotic testbeds.

A second source of bias arises from heterogeneous validation methodologies. Studies used different ground-truth systems (e.g., motion-capture systems, laser trackers, total stations, manually surveyed reference points, and vendor-specific localisation systems).

As a result, reported positioning errors are not strictly comparable across studies because measurement uncertainty varies between evaluation protocols.

Performance reporting also differed substantially. Some studies reported mean positioning error, whereas others used RMSE, MAE, cumulative distribution functions, trajectory accuracy, or qualitative robustness measures. Many studies also evaluated only a limited number of trajectories, assets, or environmental configurations, which may overestimate performance under controlled conditions.

These sources of heterogeneity and potential bias were considered during the narrative synthesis. Consequently, the review emphasises consistent patterns observed across multiple studies rather than direct numerical comparison of individual performance values. Where a source's headline improvement exceeded what its own tabulated baselines support, the reproducible figure was used and the original claim is recorded in Supplementary Table S2. For example, one study reports a 90% reduction in localisation error [15], whereas its own reported average errors (0.42 m for the proposed method against 1.16 m and 1.38 m for the filter-based baselines) correspond to reductions of 63.8% and 69.6%; the claimed figure is, therefore, not used. In another case, the fused system achieves the highest reported "enhanced accuracy" (96.32% against 80.49%) while simultaneously recording the largest average error (0.359 m) of the compared configurations [16]; because the two metrics move in opposite directions, no paired improvement can be derived and the study is cited only for the accuracy–error trade-off it documents. Both examples are transferable supporting studies rather than core industrial evidence, and their performance values are consequently excluded from the industrial synthesis and from the headline improvement range reported in Section 3.3.1.

Table 5 summarises the quality assessment criteria used to support study selection and evidence interpretation.

Table 5. Quality assessment criteria applied to the included studies.

Code	Criterion	Assessment Focus	Decision Use
Q1	Methodological clarity	The study clearly describes the proposed system, sensing setup, fusion method, and experimental procedure.	Included if clear
Q2	Experimental validation	The study reports empirical testing using real-world, laboratory, industrial, robotic, or transferable indoor localisation experiments.	Required
Q3	Quantitative performance reporting	The study reports measurable outcomes such as positioning error, RMSE, MAE, accuracy, latency, update rate, or robustness indicators.	Required
Q4	Relevance to industrial localisation	The study is either directly evaluated in an industrial context or provides transferable evidence for indoor tracking, sensor behaviour, fusion design, NLOS mitigation, drift correction, or real-time localisation.	Required

The completed PRISMA 2020 checklist is provided as Supplementary Table S1. The extracted evidence matrix, including each study's type (primary empirical) and its Q1–Q4 assessment outcome, is provided as Supplementary Table S2. No review protocol was registered or separately prepared for this systematic literature review.

3. Analysis and Synthesis

The 61 non-duplicate, peer-reviewed studies (2021–2026) retained after full screening were grouped into the following: core industrial studies ($n = 12$), evaluated in factory, warehouse, plant-factory, or industrial logistics environments and transferable supporting studies ($n = 49$), evaluated in generic indoor settings (e.g., offices, corridors, multi-floor

buildings) but reporting sensor behaviour, fusion methods, or performance findings that transfer to industrial asset-tracking contexts.

3.1. Overview of the Evidence Base

Across the evidence base, multimodal information fusion emerges less as a single-technology optimisation problem and more as a system-design problem. The reviewed studies span radio-frequency, inertial, visual, and LiDAR sensing; RFID-based identification; andhybrid software architectures. While modalities such as UWB, IMU/PDR, Wi-Fi/RSSI, LiDAR, vision, RFID, BLE, magnetic sensing, ultrasonic sensing, and odometry recur throughout the literature, they play different roles depending on the environment, infrastructure constraints, and required continuity.

Core industrial studies provide the strongest evidence for deployment relevance. For example, TrackInFactory demonstrates tightly coupled Wi-Fi and motion-sensor fusion for continuous industrial vehicle tracking without requiring an initial position [17]. Silva et al. further show that industrial vehicles can construct and maintain radio maps during normal operation [18]. The preciseSLAM study shows that LiDAR, IMU, and ultrasonic fusion can meet high-precision positioning requirements for plant-factory robots [19]. More recent work extends this trajectory toward deep-learning and graph-based fusion for industrial NLOS scenarios, including UWB–IMU fusion using cascaded DNNs and graph neural networks [20,21].

3.2. RQ1: Individual Sensor Strengths, Limitations, and Complementarity

Across the 61 included studies, sensing modalities were identified, grouped by sensor type, and analysed in terms of operating principle, measured performance, limitations, and the conditions under which each modality performs best. Figure 3 summarises the distribution of modalities across the evidence base, split by evidence group.

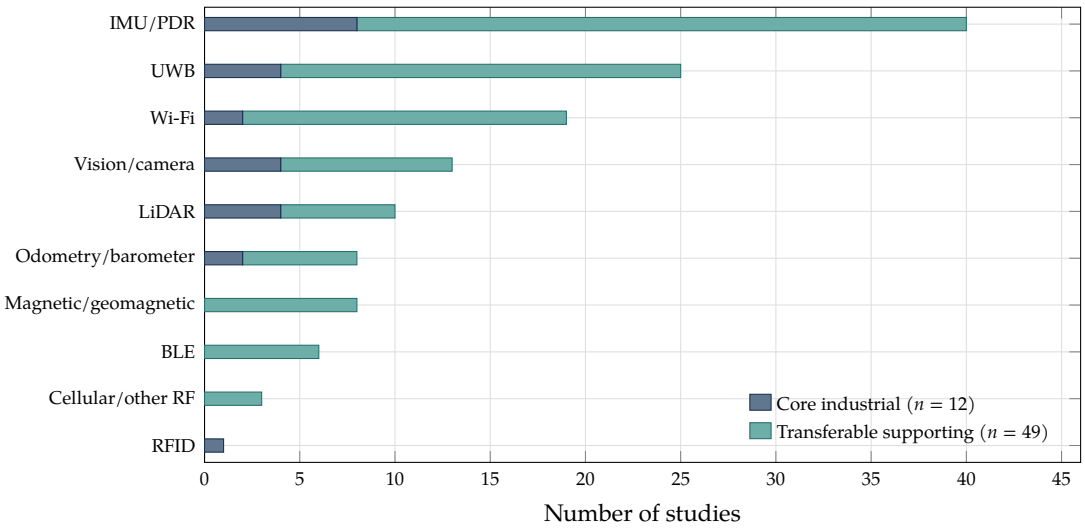


Figure 3. Frequency of sensing modalities, split by evidence group (core industrial vs. transferable supporting).

Table 6 summarises the comparative sensor-level analysis drawn from the 61 included studies. "Papers" counts studies in which the modality appears as a primary or supporting sensor; counts are non-exclusive and sum to more than 61. "Core" counts only core industrial studies in which the modality appears as a sensor. "Typical accuracy" values are study-specific and differ in metric, ground-truth system, trajectory length and environmental conditions, so values are not comparable across rows. Paired comparative accuracy evidence

is given in Tables 7 and 8. Core strength, key limitation and best use recorded for each modality are provided in Supplementary Table S2.

Table 6. Comparative sensor analysis across the 61 included studies (12 core industrial, 49 transferable supporting).

Modality	Papers	Core	Typical Accuracy	Best Use	Representative Studies
IMU/PDR	40	8	Mean error 0.97 m fused vs 3.55 m inertial-only [22]; MAE 0.053–0.076 m with UWB fusion [20]	Continuous tracking backbone in fused systems; low-infrastructure asset tracking	[16,17,19–43]
UWB	25	4	Mean error 0.12–0.14 m (GNN fusion) [21]; 0.4961 m fused [44]; MAE 0.053 m under deep-learning NLOS mitigation [20]	Factories and warehouses where anchor infrastructure is available	[20,21,29,32–34,36–38,42,44–58]
Wi-Fi	19	2	Mean error 1.24–1.42 m (auto-built radio map) [18]; RMSE 0.177 m [59]	Large facilities with existing Wi-Fi coverage where infrastructure reuse outweighs centimetre accuracy	[15–18,27,30,31,39–41,59–67]
Vision/camera	13	4	Mean error 0.38 m X/0.36 m Y (UWB + video) [54]; MAE 0.3–1.5 m (SoOp + photogrammetry) [63]	Structured layouts with adequate lighting; semantic scene understanding alongside RF sensing	[45,48–50,54,59,61–63,68–72]
LiDAR	10	4	RMSE 0.113 m simulation/0.287 m laboratory (EKF+SPF) [47]; 5.4 cm plant-factory accuracy [19]	Plant factories and structured warehouses requiring centimetre-level positioning	[19,28,47,51,67–70]
Odometry/barometer	8	2	RMSE 0.075 m fused vs. 0.835 m UWB-only [53]; 0.34–0.79 m across three cases [25]	Relative-motion component in fused systems; constrained or degraded environments	[18,25,30,35,40,52,53,58]
Magnetic/geomagnetic	8	0	Mean error 0.29 m fused vs. 1.90 m inertial-only and 7.21 m magnetic-only [26]	Multi-floor and infrastructure-free correction layer where metal interference is limited	[16,26,27,58,60,65,73]
BLE	6	0	Mean error 0.42 m fused [66]; 0.57–0.62 m (graph attention) [71]; 1.60–1.73 m (5G + BLE) [74]	Supplementary coverage in 5G or Wi-Fi deployments where cost dominates accuracy	[43,46,64,66,71,74]
Cellular/other RF	3	0	Mean error 0.21 m [73]; 1.73 m (5G+BLE HIPS) [74]	Emergency and wide-area monitoring; short-range static precision where line-of-sight holds	[73–75]
RFID	1	1	Mean error 0.97 m fused over a 10 m route vs. 3.55 m inertial-only [22]	Warehouse choke-point re-initialisation and identity verification alongside continuous tracking	[22]

Table 7. Fusion versus baseline accuracy on absolute positioning error.

Study	Fusion Combination	Metric (As Reported)	Baseline	Fused Result	Impr. (%)	Prov.
[22]	IMU/PDR + passive RFID	Mean location error	IMU/PDR alone 3.55 m	0.97 m (10 m route)	≈73	C
[21]	UWB + IMU, GNN (PosGNN)	Mean positioning error	EKF 0.25 m (Traj. 1)	0.12 m (Traj. 1)	52.0	C
[21]	UWB + IMU, GNN (PosGNN)	Mean positioning error	DNN 0.18 m (Traj. 1)	0.12 m (Traj. 1)	33.3	C
[17]	Wi-Fi RSSI + motion sensors, tight-coupling PF	Mean error	Loose-coupling PF 1.03 m	0.81 m	21.4	C
[54]	UWB RTLS + AI video, Kalman + EMA	Average localisation error (X)	No filter 0.50 m	0.38 m	24.0	C

Table 7. *Cont.*

Study	Fusion Combination	Metric (As Reported)	Baseline	Fused Result	Impr. (%)	Prov.
[44]	UWB + INS, KEF-based GPDAF	Average localisation error	MVG-GPDAF 0.7392 m	0.4961 m	32.9	S
[27]	Wi-Fi + magnetic + PDR (SF-PF)	Average error	PDR alone 7.33 m	1.05 m	85.7	S
[27]	Wi-Fi + magnetic + PDR (SF-PF)	Average error	PDR + Wi-Fi 4.58 m	1.05 m	77.1	S
[27]	Wi-Fi + magnetic + PDR (SF-PF)	Average error	WEKF 1.59 m	1.41 m	11.3	S
[60]	Wi-Fi + smartphone IMU + magnetic, PF	Mean error	Wi-Fi only (configuration)	<0.60 m	≈50	S
[59]	Wi-Fi + low-resolution vision, soft fusion	RMSE	Wi-Fi only 0.474 m	0.177 m (layer 2)	62.7	C
[23]	Smartphone IMU/PDR + conversational input	RMSE	Inertial-only 17.55 m	6.91 m	61.0	S
[28]	LiDAR + IMU, FGO coupled with AUKF	Positioning RMSE	AUKF + PL-ICP	≈0.27 m	37.2	S
[46]	UWB + BLE, improved Levenberg–Marquardt	Average positioning error	Single-BS model	Dual-BS model	27.8	S
[30]	Wi-Fi + PDR + barometer, factor graph (FGLA)	Average error	FG 0.37 m	0.31 m	16.2	C
[47]	UWB + LiDAR SLAM, EKF + SPF	RMSE (simulation)	EKF 0.143 m	0.113 m	21.0	S
[47]	UWB + LiDAR SLAM, UKF + SPF	RMSE (laboratory)	UKF 0.368 m	0.287 m	22	S
[31]	Wi-Fi RSSI + PDR (WBO filter)	Mean positioning error	PDR alone	≈0.68 m	42.0	S
[51]	UWB + LiDAR, ISSA-optimised PF	Positioning accuracy	UWB alone	ISSA-PF fusion	69.2	S
[51]	UWB + LiDAR, ISSA-optimised PF	Positioning accuracy	LiDAR alone	ISSA-PF fusion	59.6	S
[53]	UWB + wheel odometry, dynamic UKF	RMSE	UWB only 0.835 m	0.075 m	91.0	C
[53]	UWB + wheel odometry, dynamic UKF	RMSE	Odometry only 0.296 m	0.075 m	74.7	C
[25]	Smartphone IMU + barometer, KF + LR	RMSE	Existing literature values	0.51 m (mean)	19.5–78.6	S
[33]	UWB (Pozyx) + IMU, CNN–LSTM	RMSE (S-path)	CNN alone 0.238 m	0.142 m	40.3	C
[56]	UWB + INS, fuzzy clustering + multi-filter	RMSE	Two-Level EKF 2.978 m	0.781 m	73.7	S
[57]	UWB deep learning TCN–CAM	Average point error	LSTM-AM 13.9 cm	3.32 cm	76.1	S
[57]	UWB deep learning TCN–CAM	Average point error	TCN 4.12 cm	3.32 cm	19.4	S
[65]	Wi-Fi + magnetic + PDR (HV-DDTW + CPF)	Average error	APF 3.74 m (office)	1.1 m	70.0	S
[36]	UWB + IMU, UKF-based fusion	RMSE	LSQ/KF 0.85 m	0.43 m	49.4	S
[66]	Wi-Fi + BLE + IMU, Transformer + ESKF	MAE/mean error	CNN baseline 0.72 m	0.42 m	41.7	C
[71]	BLE + visible light + IMU (CrowdGraph)	Mean positioning error	KNN 1.10 m	0.62 m	43.6	C
[71]	BLE + visible light + IMU (CrowdGraph)	Mean positioning error	GAT 0.72 m	0.62 m	13.9	C

Table 7. *Cont.*

Study	Fusion Combination	Metric (As Reported)	Baseline	Fused Result	Impr. (%)	Prov.
[39]	Wi-Fi + IMU (GraphKANLoc)	ATE	AGNNConv (OWP-IMU)	0.073 m	53.3	S
[40]	CSI + IMU auxiliary branch (ConvNeXt)	Average localisation error	Without auxiliary branch 2.18 m	1.02 m	53.0	C
[74]	5G NR + BLE (HIPS)	Mean positioning error	RF baseline	1.73 m	38.2	S
[42]	UWB + IMU, cascaded DNN + AAFN	Mean positioning error	UWB-only trilateration	0.23 m	41.0	S
[42]	UWB + IMU, cascaded DNN + AAFN	Mean positioning error	IMU-only dead reckoning	0.23 m	67.1	S
[26]	PDR + geomagnetic, particle filter	Mean error	Inertial-only 1.90 m	0.29 m	85.0	S
[26]	PDR + geomagnetic, particle filter	Mean error	Magnetic-only 7.21 m	0.29 m	96.0	S
[67]	Wi-Fi RTT + LiDAR-inertial (RTT-LIO)	RMSE (Exp. 1)	Standalone LIO 0.98 m	0.76 m	22.4	S
[67]	Wi-Fi RTT + LiDAR-inertial (RTT-LIO)	RMSE (Exp. 1)	LS RTT-only 2.08 m	0.76 m	63.5	S
[67]	Wi-Fi RTT + LiDAR-inertial (RTT-LIO)	RMSE (Exp. 1)	Loose-coupling 0.87 m	0.76 m	12.6	S

Table 8. Fusion versus baseline accuracy on relative error and percentage-point metrics. These values are *not* comparable with those in Table 7: a relative-error baseline already normalises for path length, so reductions are smaller by construction.

Study	Fusion Combination	Metric (As Reported)	Baseline	Fused Result	Impr.	Prov.
[35]	Wheel odometry + IMU + LiDAR (EKF + AMCL)	Relative error (%)	Single wheel odometry	<1%	8.35	S
[35]	Wheel odometry + IMU + LiDAR (EKF + AMCL)	Relative error (%)	EKF-only fusion	<1%	4.0	S
[24]	Multi-sensor PDR (IPF–UKF + LM-SIF)	Relative positioning error (%)	Four traditional PDR baselines	0.36% (2D)/0.435% (3D)	0.76–1.17 pp (2D)	S

UWB and IMU/PDR are the most frequently reported modalities; however, frequency is not a proxy for performance. Studies differ substantially in metric choice, environment, and evaluation protocol. The discussion below is, therefore, structural—i.e., what each modality tends to provide, where it fails, and why it is typically paired with others—rather than a ranking.

IMU and PDR.

Inertial sensing is the only modality that provides continuous output without external infrastructure, but its error grows with time and distance. The evidence is consistent that even sparse absolute references can bound drift. For example, ref. [22] reports mean error decreasing from 3.55 m (inertial-only) to 0.97 m using passive RFID checkpoints, and ref. [26] reports 1.90 m (inertial-only) versus 0.29 m with geomagnetic correction (Table 7). The mechanism matters more than the headline magnitude: three tag reads along a 10 m corridor can be sufficient to reset drift [22], and sparse human-provided references reduce RMSE by 61% relative to an inertial-only baseline [23]. Two qualifications apply: dead-reckoning error depends on the tested trajectory length, and these improvements are measured against standalone dead reckoning, illustrating the value of correction rather than the superiority of a particular architecture.

UWB.

UWB provides the most accurate absolute ranging in the evidence base. Under line-of-sight (LOS) conditions, it is consistently sub-metre, and the lowest reported errors combine ranging with explicit NLOS mitigation. For example, ref. [21] reports mean error of 0.12 m and 0.14 m across two trajectories using a graph neural network, compared with 0.25–0.28 m for an EKF and 0.17–0.20 m for a DNN baseline on the same trajectories. Ref. [57] applies a temporal convolutional network with channel attention directly to UWB measurements, reaching mean error of 3.32 cm in field experiments and 10.87 cm under the most severe simulated NLOS condition tested. The governing failure mode remains discontinuous degradation under obstruction: when the LOS path is blocked, error can rise sharply. This explains why strong UWB results are typically obtained either under controlled LOS or with explicit NLOS handling, and why anchor infrastructure remains a key deployment cost.

LiDAR and vision.

LiDAR and vision provide geometric (and, for vision, semantic) context that radio modalities do not, and they achieve the highest standalone precision in the review. In an industrial plant-factory robot scenario, ref. [19] reports 5.4 cm positioning accuracy—meeting a 5 cm control requirement for ridge-raising—by tightly coupling LiDAR, inertial, and ultrasonic sensing; the same study notes that LiDAR–inertial SLAM without global correction accumulated drift over long runs. This precision is contingent on scene conditions: LiDAR degrades in geometrically repetitive or sparse environments, while vision degrades under poor illumination and occlusion. Ref. [70] addresses occlusion by removing dynamic obstacles so their points do not corrupt pose estimation. In fusion settings, ref. [54] reports mean localisation error of 0.38 m (X) and 0.36 m (Y) when UWB ranging is fused with video analytics, illustrating how RF measurements can compensate for visual occlusion rather than replace vision. Neither modality is infrastructure-free in practice: LiDAR typically requires a map or distinctive structure, and vision requires camera coverage (with attendant privacy concerns); both are also computationally heavier than RF or inertial sensing.

Wi-Fi.

Wi-Fi is the most infrastructure-efficient modality because it can reuse access points already present in many industrial buildings. Its standalone accuracy is modest and variable. Ref. [18] reports mean errors of 1.32 m, 1.24 m, and 1.42 m across complete, partial, and absent floor-plan configurations for a radio map built automatically during normal vehicle operation, with a 99th-percentile error of 5.54 m and a maximum error of 6.28 m in the no-floor-plan condition. The gap between mean and tail performance is central: Wi-Fi can be reasonably bounded under typical conditions but exhibits a heavy tail under adverse ones. As a result, it appears mainly as a complementary modality, paired with inertial sensing for continuity and with LiDAR or other absolute references for precision. Recurring weaknesses include multipath, signal fluctuation, fingerprint degradation after layout changes, and the ongoing effort of radio-map maintenance [17,18,59].

RFID.

RFID primarily provides identity rather than metric localisation accuracy: a tag read confirms that a specific asset is at a known location and can reset accumulated drift, making RFID a checkpoint mechanism rather than a continuous tracking method. Ref. [22] is the only included study where RFID acts as a primary modality, reporting a mean error of 0.97 m over a 10 m route when inertial trajectories are reoriented at tag reads. The single-study count is itself informative: although RFID is widely discussed in warehouse practice, this evidence base does not support modality-level performance generalisations. Structurally,

RFID provides discrete fixes (with no trajectory information between reads), and read range can degrade near metal—a common condition in industrial racking.

Why these sensors are complementary rather than interchangeable.

Each modality fails differently, and failure modes are often weakly correlated. RF modalities (UWB, Wi-Fi, RFID) provide absolute position or identity but degrade under obstruction and multipath. Inertial sensing provides continuity but drifts without bound. LiDAR and vision provide geometric/semantic context but depend on scene structure and illumination. Because these failure modes are largely orthogonal, what is unreliable for one sensor may still be informative for another. Ref. [47] illustrates this directly by fusing UWB (discontinuous NLOS errors) with LiDAR (gradual drift): EKF+SPF reaches RMSE of 0.113 m in simulation and 0.291 m in laboratory trials, compared with 0.160 m and 0.368 m for the single-modality baselines. Modality selection is, therefore, less about choosing a single "best" sensor and more about ensuring that at least one component remains reliable under each condition the deployment will encounter; the evidence does not support treating these modalities as interchangeable substitutes.

3.3. RQ2: Fusion Performance vs. Single-Sensor Approaches

This subsection addresses Sub-RQ2 by grouping findings into three outcome categories: (i) accuracy gains, (ii) robustness to NLOS/occlusion/drift, and (iii) real-time performance.

3.3.1. Accuracy Improvements

Table 7 reports quantitative accuracy results for studies that evaluate a fused system and a baseline on the same metric, trajectory, and experimental condition. Of the 61 included studies, 35 support this paired comparison. The remaining 26 are excluded from this analysis because they report no paired baseline, use a non-error metric, or provide a percentage improvement that cannot be reconstructed from the reported values. In Tables 7 and 8, "Prov." indicates whether each improvement percentage was reported by the source (S) or calculated by the review authors (C).

Across the verified comparisons in transferable supporting studies, fusion reduces positioning error for every architecture for which a paired comparison was available; however, the size of the reduction is strongly condition-dependent. Improvements on absolute positioning error span approximately 11.3%, where Wi-Fi, magnetic and PDR fusion is compared against a comparable filter baseline (WEKF) [27], to 96%, where PDR fused with geomagnetic matching is compared against magnetic-only positioning [26] (Figure 4). This range is descriptive rather than indicative of a central tendency, and the extremes reflect different situations. Large improvements typically occur when a single modality fails under the tested condition: in the same study, PDR fused with geomagnetic matching reduces mean error by 85% relative to inertial-only and by 96% relative to magnetic-only, the two figures coming from separate baseline comparisons within that study [26], while UWB fused with wheel odometry under severe NLOS reduces RMSE by 91% relative to UWB alone [53]. Smaller improvements often occur when baselines are already strong: the factor-graph fusion in [30] improves on a comparable factor graph by 16.2%, and UWB–BLE fusion in [46] improves on a single-base-station configuration by 27.8%. A single headline improvement figure would, therefore, be misleading: gains are largest where a modality is weakest, not necessarily where fusion is most complex.

Figure 5 shows the same comparison for RMSE-based studies and Figure 6 presents the same comparison for relative-error and percentage-point metrics.

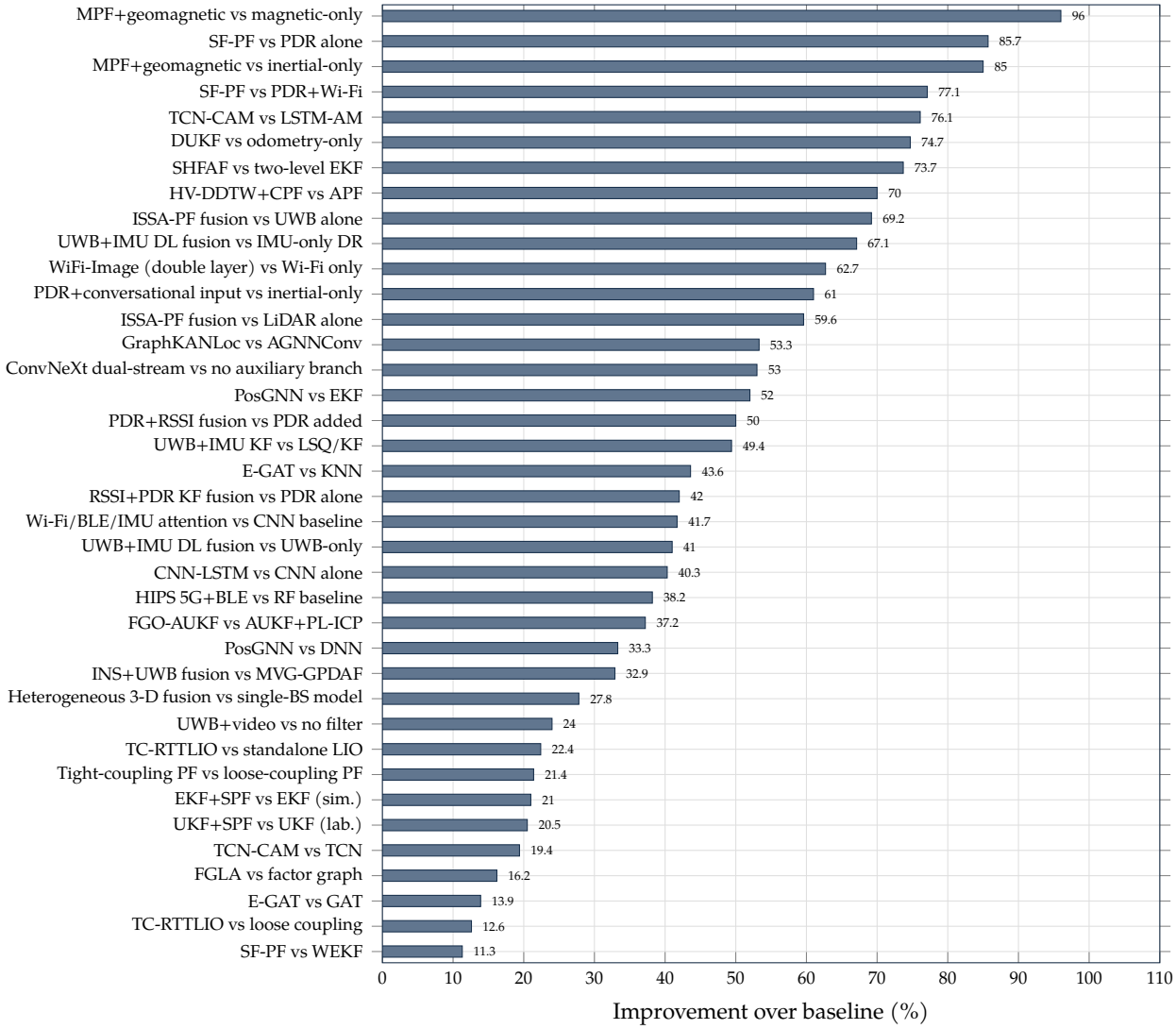


Figure 4. Comparisons on mean and average positioning error. Each bar is labelled with the study and the baseline against which the improvement was measured, so a study appears more than once where it reports comparisons against different baselines.Bars 1 and 3: Medjahed et al. [26]. Bars 2, 4 and 38: Guo et al. [27]. Bars 5 and 34: He et al. [57]. Bars 6 and 7: Liu et al. [53] and Cheng et al. [56]. Bar 8: Chen et al. [65]. Bars 9 and 13: Wang et al. [51]. Bars 10 and 22: Cao et al. [42]. Bar 11: Tahiri et al. [59]. Bar 12: Sheshadri et al. [23]. Bar 14: Zhang et al. [39]. Bar 15: Wang et al. [40]. Bars 16 and 26: Muthineni et al. [21]. Bar 17: Hu et al. [60]. Bar 18: AlOkby et al. [36]. Bars 19 and 36: Li et al. [71]. Bar 20: Mehrabian et al. [31]. Bar 21: Kien et al. [66]. Bar 23: Abhilasha et al. [33]. Bar 24: Wu et al. [74]. Bar 25: Zhou et al. [28]. Bar 27: Wang et al. [44]. Bar 28: Qin et al. [46]. Bar 29: D'Irienzio et al. [54]. Bar 30: Silva et al. [17]. Bars 32–33: Zhou et al. [47]. Bar 35: Li et al. [30]. Bar 37: Xu et al. [67].

Restricting attention to the 12 core industrial studies, where baseline and fused results were obtained in factory, warehouse or industrial-logistics settings, the verified improvement range is narrower. The lower end is 21.4% for tightly coupled Wi-Fi and motion-sensor fusion on an industrial tow tractor compared against a loose-coupling approach on the same sensors [17]. The upper end is 72.7% for passive-RFID checkpoint correction of inertial drift in a warehouse scenario compared against strap-down inertial positioning alone [22]. Intermediate values include 52.0% for UWB–IMU graph-neural-network fusion in an industrial NLOS cell with an AGV compared against an EKF baseline [21], and 24% per axis for UWB fusion compared against an unfiltered baseline [54].

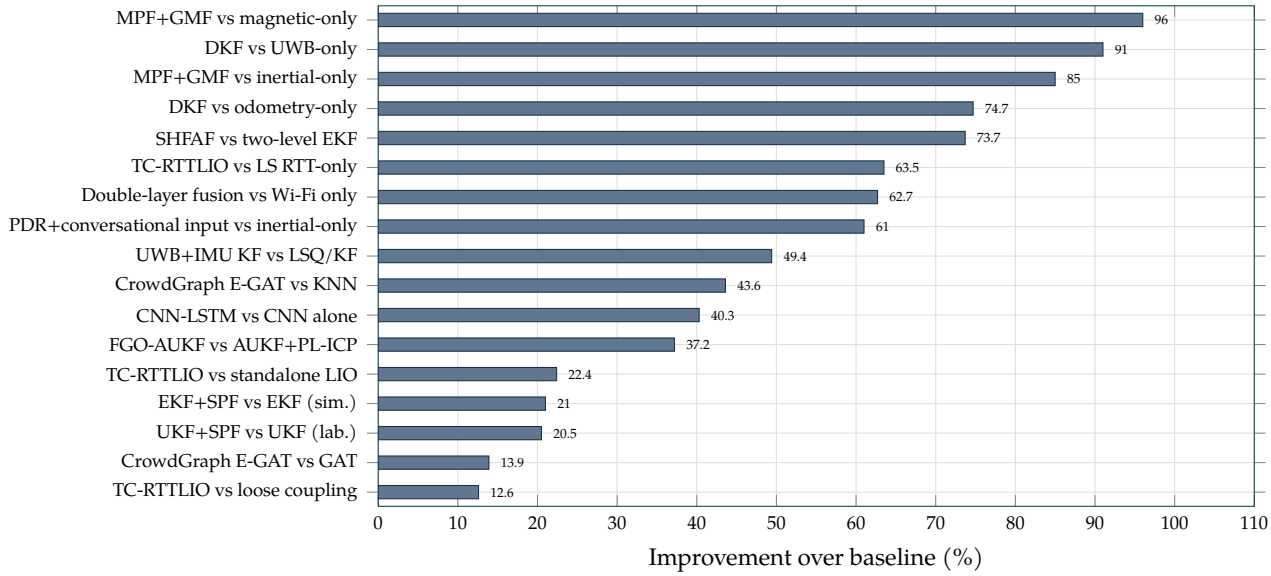


Figure 5. Comparisons on RMSE. Reported separately from mean and average positioning error because RMSE and mean error are not directly comparable. Bars 1 and 3: Medjahed et al. [26]. Bar 2 and 4: Liu et al. [53]. Bar 5: Cheng et al. [56]. Bars 6, 13 and 17: Xu et al. [67]. Bar 7: Tahiri et al. [59]. Bar 8: Sheshadri et al. [23]. Bar 9: AlOkby et al. [36]. Bars 10 and 16: Li et al. [71]. Bar 11: Abhilasha et al. [33]. Bar 12: Zhou et al. [28]. Bars 14–15: Zhou et al. [47].

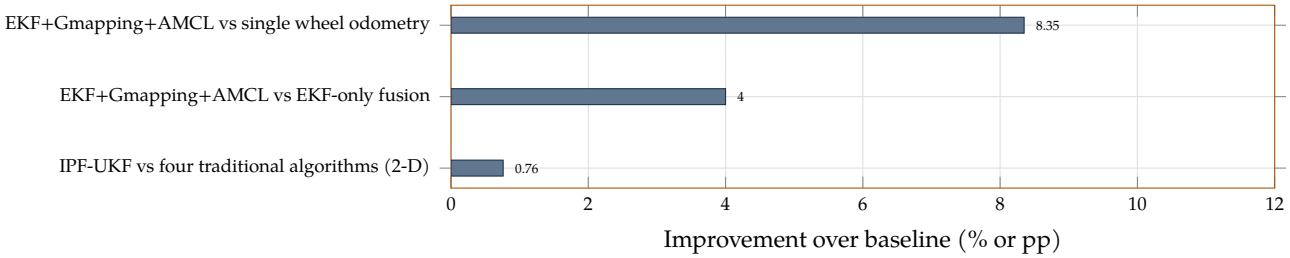


Figure 6. Comparisons on relative error and percentage-point metrics. These values are *not* comparable with Figures 4 and 5: a relative-error baseline already normalises for path length, so the reductions are smaller by construction. Bars 1–2: Teng et al. [35]. Bar 3: Wu et al. [24].

Each percentage reported above is drawn only from a study that reports both a baseline and a fused result on the same metric, trajectory, and experimental condition. Where a study reports several metrics or settings, the comparison is made within one metric and one condition only, and no cross-metric or cross-setting derivations are included. Table 7 records, for every comparison, the metric as reported, the baseline, the fused result, and whether the percentage was reported by the source (S) or calculated by the review authors (C). Comparisons that could not be reconstructed from a study's own reported values are excluded from the range.

A second pattern is baseline dependence: the same fused system can show very different improvements depending on what it is compared against. For instance, the Wi-Fi/magnetic/PDR fusion in [27] improves over standalone PDR by 85.7%, over PDR + Wi-Fi by 77.1%, but over a comparable filter by only 11.3%. Similarly, the UWB/INS fusion in [44] improves by 32.9% relative to its closest comparator but by 60.3% relative to a standard EKF on the same data. This is why Table 7 reports each improvement alongside its baseline, and why these percentages should not be interpreted as intrinsic properties of a fusion method.

Finally, absolute-error improvements (metres, Table 7) are not comparable with the relative-error and percentage-point results in Table 8 (4% to 8.35%). Relative-error metrics normalise for path length, leaving less accumulated drift for fusion to remove; the smaller range is, therefore, a consequence of the metric rather than weaker fusion performance.

3.3.2. Fusion Architecture and Algorithmic Patterns

Records were coded for coupling level, fusion family, and fusion stage. The patterns below, therefore, reflect counts across the evidence base rather than illustrative examples. Of the 61 included studies, 56 describe a multimodal fusion system; five were classified as not implementing sensor fusion and are excluded from the architecture counts.

Within the 56 fusion systems, three coupling levels dominate. Tightly coupled designs fuse raw or intermediate measurements prior to state estimation; these account for the majority and are common in UWB-based and factor-graph systems. Loosely coupled designs combine already-estimated positions; these remain common in Wi-Fi and RFID configurations where subsystems are often validated independently. A smaller set of studies use hierarchical designs, fusing at more than one level (e.g., a measurement-level front end feeding a decision-level classifier).

By fusion family, filter-based probabilistic methods remain the largest group. Kalman, extended Kalman, unscented Kalman, and particle filters underpin many UWB–IMU, Wi-Fi–PDR, RFID–IMU, and LiDAR–UWB systems [31,47,76,77]. Optimisation-based approaches—primarily factor graphs—form the second group and have increased sharply since 2022 [19,28,67,68]. Learning-based methods are the smallest but fastest-growing group, spanning LSTM, CNN–LSTM, cascaded DNNs, and graph neural networks [15,20,21,42]. Rule-based or heuristic aggregation (e.g., checkpoint correction and semantic rule engines) appears in a small number of studies, typically where identity or event detection—rather than continuous pose—is the main output.

Three qualifications are important when interpreting these counts. First, architecture family is not independent of modality: filter-based methods dominate whenever an inertial sensor is present because the prediction–correction structure directly addresses drift. Second, learning-based approaches are concentrated in the most recent publication years, so counts reflect a trend rather than a settled preference. Third, architectural labels are used inconsistently in the source literature; several papers describe a system as tightly coupled even when fusion occurs at the state or decision level. The coding applied here, therefore, records coupling based on where measurements are combined, not on the authors' terminology.

Overall, the evidence does not support a single "best" architecture. Classical filters are often appropriate when computation is constrained and sensor behaviour is approximately stationary; factor-graph approaches are well suited when loop closure or global consistency is required; learning-based methods can yield large gains when the error process is strongly nonlinear and representative training data are available. However, the reviewed studies do not provide controlled comparisons across these families on common datasets: no two studies evaluate different fusion families on identical inputs, conditions, and metrics. As a result, apparent performance differences are confounded by modality choice, environment, and evaluation protocol. This gap motivates reporting representative architectures (Table 9) rather than attempting a ranked performance comparison.

Table 9 summarises representative fusion architectures identified across the reviewed studies, grouped by architecture type.

Table 9. Representative fusion architectures identified across the reviewed studies.

Coupling/Family	Fusion Pattern	Representative Sensors	Typical Algorithms	Studies
Tightly coupled, filter-based	RF plus inertial fusion at measurement level	UWB, Wi-Fi, RFID, IMU, PDR	KF, EKF, UKF, PF	[20,31,76,77]
Tightly coupled, filter-based	UWB plus LiDAR or vision at measurement level	UWB, LiDAR, camera, IMU	PF, tightly coupled estimation	[47–49,53]
Loosely coupled, filter-based	Wi-Fi or RSSI assisted fusion at state level	Wi-Fi, RSSI, PDR, magnetic	PF, crowd-sensing fusion, LSTM	[15,16,18,27]
Tightly coupled, filter-based	Industrial vehicle tracking	Wi-Fi, IMU, wheel encoder	Tight-coupling particle filter	[17,18]
Tightly coupled, optimisation-based	LiDAR or vision plus inertial	LiDAR, camera, IMU, odometry	Factor graph, graph optimisation, SLAM	[19,28,67,68]
Tightly coupled, learning-based	AI-driven industrial fusion	UWB, IMU, vision, semantic features	DNN, GNN, deep NLOS mitigation	[20,21,57,70]
Rule-based aggregation	Sensor-agnostic multi-object tracking	Multiple heterogeneous sensors	Data correlation, clustering, prediction	[62]

3.3.3. Robustness to NLOS, Occlusion and Drift

Robustness is the most consistently reported benefit of fusion, even when headline accuracy gains are modest. Of the 61 included studies, 30 explicitly tested NLOS, metallic, or occlusion conditions and 17 explicitly did not. The remaining 14 do not report whether such conditions were tested; the count of NLOS-tested studies should, therefore, be interpreted as a lower bound.

NLOS mitigation. The clearest evidence comes from methods that detect and suppress NLOS measurements rather than implicitly averaging over them. Ref. [57] applies a temporal convolutional network with channel attention directly to UWB measurements, reducing average point error from 13.9 cm (LSTM baseline) and 4.12 cm (TCN baseline) to 3.32 cm; the gain is largest under the highest NLOS severity. Ref. [37] improves robustness through a different mechanism by modelling occlusion as an explicit Markov state. This allows the filter to respond to an occluded channel rather than treating it as generic noise, attaining particle-filter-equivalent accuracy at approximately 62% of the computational cost.

Drift bounding through complementarity. Absolute measurements need not be frequent or highly precise to bound inertial drift; they primarily need to be available. Ref. [22] demonstrates a limiting case: passive RFID tags read at three checkpoints along a 10 m corridor reduce mean error from 3.55 m to 0.97 m without continuous ranging. Ref. [26] generalises the principle using geomagnetic correction, reporting 0.29 m versus 1.90 m for inertial-only and 7.21 m for magnetic-only operation. Ref. [23] extends it to sparse human-provided references, reducing RMSE by 61% across 48 walking trials.

Robustness can also be driven by coupling strategy, not only by sensor choice. Ref. [17] reports that tight coupling reduced maximum error by more than 60% relative to loose coupling using the same sensors, while the corresponding mean-error improvement was 21.4% (only the mean-error value is comparable with Table 7). Similarly, ref. [27] shows a particle-migration strategy improving on a comparable filter by 11.3% but on standalone dead reckoning by 85.7%—illustrating again that the baseline strongly influences the reported gain.

Generalisability. Evidence for autonomous map construction is weaker than it is sometimes assumed to be. Ref. [18] reports mean errors of 1.32 m, 1.24 m, and 1.42 m across complete, partial, and absent floor-plan configurations, with the automatically built

map improving accuracy below the 99th percentile. However, this comparison varies floor-plan availability rather than directly comparing against a manually surveyed map; the no-floor-plan condition was also the weakest of the three.

Reporting form. Fewer than a third of studies that test NLOS conditions stratify results by severity; ref. [42] (low/medium/high NLOS reporting) is a notable exception. Aggregate figures can conceal whether a system degrades gracefully or fails at a threshold. For industrial deployment—where NLOS can be persistent rather than occasional—the degradation profile is often more decision-relevant than the average error.

3.3.4. Real-Time Performance

Real-time capability is the least well-evidenced outcome category. Of the 61 included studies, 17 report a measured latency or update rate, 29 claim real-time operation without measuring it, and 15 do not address real-time constraints. Only the 17 measured cases can, therefore, support evidence-based conclusions about deployability; in many other papers, "real time" is asserted as a design property rather than demonstrated as an observed outcome.

Where latency is measured, fused systems can meet practical update-rate requirements, but at measurable computational cost. Ref. [51] reports 0.09 m MAE at 45 ms per sample for Kalman–LSTM fusion in warehouse forklift tracking, and ref. [54] reports approximately 100 ms end-to-end latency for UWB–video fusion—both compatible with vehicle-tracking update rates. Ref. [62] reports a correlation engine that processes up to 80,000 measurements in linear time on commodity hardware, maintaining multi-object tracking accuracy near 92% across sensor-density scenarios.

The cost of fusion is most visible in particle-filter studies, where particle count is often selected as an explicit accuracy–efficiency trade-off. Ref. [51] reduces the required particle count to 25% of the original while preserving accuracy, and ref. [37] achieves particle-filter-equivalent accuracy at approximately 62% of the computational cost. Both findings come from studies that measured runtime rather than assuming it.

Cross-study comparison remains unreliable for two reasons: papers that report computation time rarely specify the hardware platform, and latency is seldom reported alongside accuracy at the same operating point. As a result, the accuracy–efficiency frontier of each method cannot be reconstructed from the published record. For deployments with a fixed compute budget, this omission can be more consequential than remaining accuracy gaps.

3.4. RQ3: Operational Benefits, Challenges, and Future Directions

3.4.1. Operational Benefits

Table 10 summarises the operational benefits and implementation challenges identified across all 61 included studies.

A commonly reported operational benefit across the industrial literature is improved asset visibility through continuous coverage. The TrackInFactory study [17] demonstrated that combining Wi-Fi fingerprinting with IMU-based motion tracking allowed an industrial tow tractor to be tracked continuously, even in areas where signals are weak or unavailable. The system achieved a mean error of 0.81 m and did not require prior knowledge of the starting position, addressing a key limitation of single-sensor systems that lose accuracy in signal "dead-zones".

A follow-up real-world deployment study by the same authors [18] further demonstrated that vehicles can automatically build and update their own radio maps during normal operation. These automatically generated maps outperformed manually created ones, while significantly reducing setup time, requiring only 53 min of vehicle operation for a 1000 m^2 industrial space, compared with 5.6 h of manual surveying.

Table 10. Operational benefits and implementation challenges of multimodal information fusion for industrial localisation.

Theme	Evidence from Literature	Studies
Operational Benefits		
Better asset visibility	Fused systems enabled more continuous tracking through dead zones, extending operational coverage by combining global RF and local inertial estimates	[17,18,78]
Lower drift in long-run operation	Correction from global sensors (UWB, RFID, ultrasonic) prevented cumulative drift in IMU/LiDAR odometry during extended trajectories	[17,19,51]
Higher resilience to environment	Tight coupling reduced maximum errors by >60% vs. loose coupling; DL-based NLOS mitigation reduced errors by up to 76% in dynamic environments	[17,27,57]
Real-time tracking capability	O(n) processing architectures and factor-graph optimisation enabled near real-time update cycles under industrial embedded compute constraints	[19,62]
Autonomous map maintenance	Vehicle-driven radio map construction eliminated manual site surveys and kept maps current with zero additional human effort	[18]
Safety and logistics improvement	Worker and asset visibility in mixed human-vehicle factory zones supported improved safety monitoring and logistics optimisation in industrial deployments	[18,78]
Implementation Challenges		
Calibration effort	Sensor extrinsic calibration, time synchronisation, and coordinate frame alignment require significant initial engineering effort and specialist knowledge	[19,51,62]
Infrastructure cost	UWB anchor installation (hardware, cabling, planning) and LiDAR per-unit cost create per-deployment capital expenditure barriers	[17,78]
Synchronisation Complexity	Tight timestamp alignment between sensors with different update rates (UWB 10 Hz, IMU 100 Hz, camera 30 Hz) introduces algorithmic and hardware overhead	[51,62]
Computational overhead	Tightly coupled factor graph and particle filter algorithms require significantly more compute than standalone methods; scaling to many assets is non-trivial	[19,27,62]
Fingerprint maintenance	Wi-Fi and magnetic radio maps degrade over time due to infrastructure changes, rearrangement of metallic assets, and AP additions or removals	[18,27,31]
Poor generalisation across environments	Systems optimised for one factory layout do not generalise well to other sites without re-calibration or retraining of learned components	[27,57,70]

The UWB logistics review [78] showed that using UWB-based RTLS in industrial environments such as smart manufacturing and warehouses can provide measurable operational benefits. These include faster localisation, more efficient vehicle routing, automated location-based tasks, and improved worker safety in environments where people and machines operate together.

For tracking both stationary and moving items on the production floor, UWB systems typically achieve accuracy in the range of 5–30 cm without extensive setup. This level of accuracy is sufficient for most applications, including inventory tracking and automated guided vehicle navigation.

3.4.2. Implementation Challenges

The most common implementation challenges identified in the 61 papers are the effort required for calibration and the cost of infrastructure. Gong et al. [19] reported that achieving the 5 cm positioning requirement for plant factory robots required careful confidence-zone calibration of ultrasonic sensors at 73 measurement points, along with careful tuning of multiple system parameters. This represents a significant engineering effort and would be difficult to apply in fast or large-scale deployments.

Wang et al. [51] noted that their particle filter approach was sensitive to how anchors were positioned in the environment. As a result, optimisation techniques were needed to adjust system parameters for each specific deployment, increasing setup complexity.

Maintaining fingerprint-based systems (such as Wi-Fi and magnetic mapping) is another key challenge in industrial environments. Silva et al. [18] showed that radio maps can quickly become outdated when indoor layouts change, which is common in factories where production lines are regularly reconfigured. Similarly, Mehrabian and Ravanmehr [31] found that signal models depend heavily on the specific environment and must be recalibrated whenever conditions change, requiring repeated setup at each new site.

The FITS study [62] also highlighted a broader challenge: the lack of standardisation across different sensor systems. Combining data from multiple sensors requires aligning timestamps, coordinate systems, and data formats, which add processing overhead and become more complex as more sensor types are included.

3.4.3. Emerging Future Directions

The main future research directions identified across the reviewed studies are discussed in detail in Section 4.3. To avoid repetition, this section does not restate them here.

4. Discussion

Figure 7 summarises the overall evidence synthesis developed from the reviewed studies. The framework illustrates how industrial tracking requirements drive sensor selection, which in turn influences fusion architecture and algorithm design. The resulting systems provide operational benefits such as improved accuracy, robustness, and continuity, while introducing practical deployment challenges. The identified research gaps and future directions emerge directly from these challenges and represent key priorities for the next generation of industrial localisation systems.

4.1. Interpretation of Findings

From an information systems perspective, the reviewed studies show that industrial localisation should be understood as a data integration problem rather than only a sensor deployment problem. Because only 12 of the 61 included studies were evaluated directly in industrial environments, the conclusions of this review should be interpreted as evidence-informed design guidance rather than definitive proof of performance across all industrial scenarios. The value of multimodal fusion lies in converting heterogeneous, noisy, asynchronous, and sometimes incomplete sensor observations into reliable location information that can support operational decisions. This shifts the design focus from selecting a single high-performing sensor to building information architectures that manage uncertainty, combine complementary data streams, and remain reliable under changing industrial conditions.

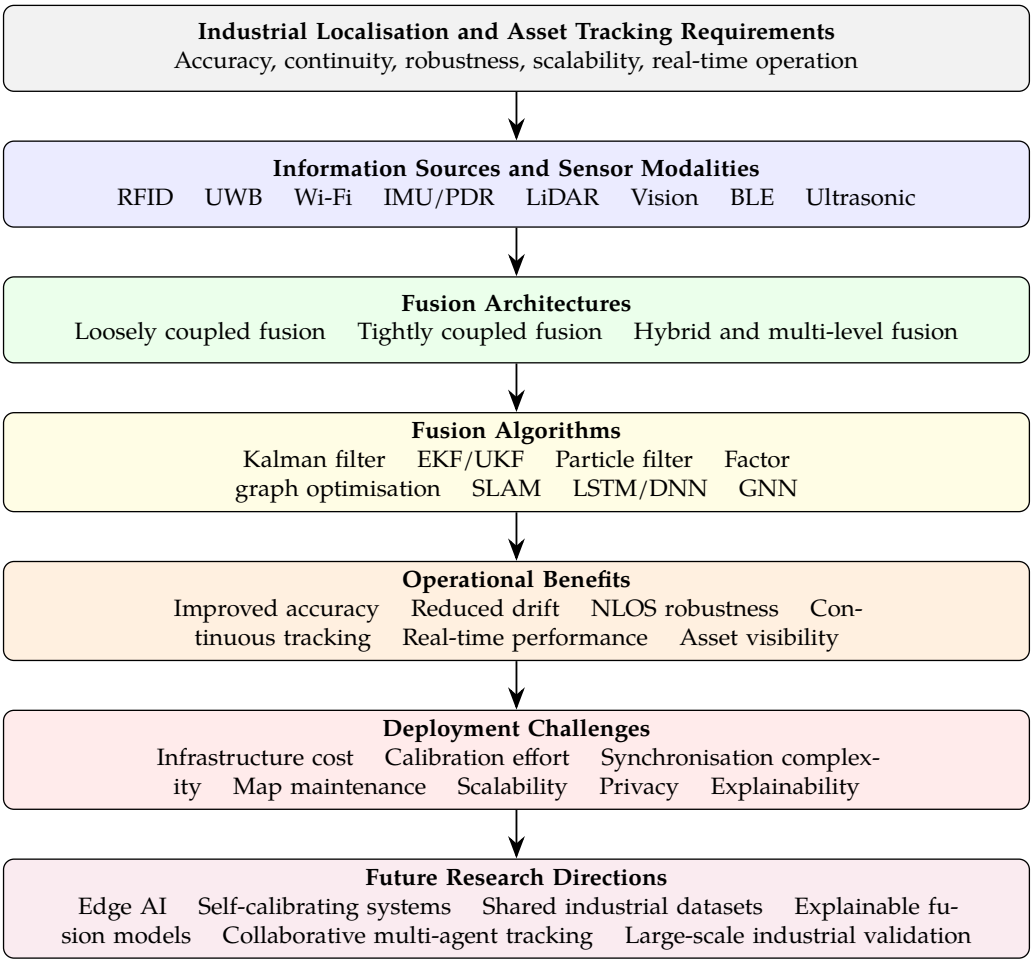


Figure 7. Conceptual synthesis framework for industrial multimodal information fusion.

The main finding of this review is that multimodal information fusion improves industrial indoor localisation not simply by increasing the number of sensors, but by combining sensors with structurally different failure modes. Across modalities, a key distinction is between absolute references, continuous motion estimation, and environment-derived context: RFID provides identity and checkpoint-level correction but cannot deliver continuous trajectories; UWB provides accurate absolute ranging but degrades under NLOS and multipath; IMU, PDR, and wheel odometry provide continuous motion information but suffer from cumulative drift and vision and LiDAR provide spatial and semantic context but are vulnerable to occlusion, feature degradation, lighting variation, and computational cost. Fusion is, therefore, effective because it enables one modality to compensate for the dominant failure mode of another.

The evidence also shows that fusion architecture is as important as sensor choice. Because tightly coupled approaches integrate measurements before final position estimation, they generally provide stronger correction under heterogeneous sensor error. In practice, this advantage appears in industrial vehicle tracking, where tightly coupled Wi-Fi and motion-sensor fusion reduced maximum error compared with loose coupling [17]. A similar pattern is reported in UWB–LiDAR fusion, where LiDAR compensates for abrupt UWB NLOS error while UWB constrains LiDAR drift [47], and in plant-factory robotics, where LiDAR, IMU, and ultrasonic fusion achieved high-precision positioning by introducing global ultrasonic constraints into a factor-graph SLAM framework [19].

A second major finding is that industrial deployment requires more than accuracy. In addition to positioning error, real-world usability depends on calibration effort, infras-

tructure cost, synchronisation complexity, computation, scalability, and maintainability, and these factors often trade off against one another. For example, UWB systems can deliver high positioning accuracy but require anchor planning and are sensitive to NLOS conditions [78]; Wi-Fi-based approaches can reuse existing access points but depend on radio maps that degrade as factory layouts change [18] and vision and LiDAR improve environmental awareness but can increase hardware and computation costs. These trade-offs explain why no single fusion architecture is universally optimal.

Beyond these deployment trade-offs, several reviewed implementations suggest potential financial benefits for industry. For example, the warehouse implementation combining UWB localisation, LSTM modelling, Kalman filtering, and QR-based product tracking identifies reduced manual inventory auditing, fewer losses from misplaced items, and improved forklift routing as potential sources of labour and operational savings [32]. Such benefits could arise when reliable location information is integrated into inventory and logistics workflows, rather than from improved positioning accuracy alone. Hardware selection may also influence affordability: UWB–wheel-odometry fusion offers a potentially lower-cost alternative to configurations requiring additional vision or LiDAR hardware, where it meets the application's positioning requirements [53]. However, these proposed advantages should not be interpreted as demonstrated monetary savings. The RFID–inertial asset-tracking study explicitly identifies industrial return on investment, particularly for small and medium-sized enterprises, as requiring further investigation [77]. Only a subset of the 12 core industrial studies report operational or economic outcomes; the economic implications discussed here are, therefore, indicative of reported deployment benefits rather than demonstrated monetary savings across the evidence base. Economic evaluation should, therefore, consider hardware, installation, integration, calibration, and maintenance costs alongside measured reductions in search time, inventory discrepancies, and vehicle idle time.

The reviewed studies further indicate a transition from hand-crafted filtering toward adaptive and learning-based fusion. Classical filters remain attractive for real-time implementation, but their assumptions about noise, motion, and sensor reliability may be insufficient in dynamic industrial environments. Recent DNN and GNN approaches show stronger potential for nonlinear UWB–IMU fusion in industrial NLOS scenarios [20,21]. This added modelling flexibility comes at a cost: learning-based fusion increases dependence on representative training data, making cross-site generalisation a major unresolved challenge.

Overall, the findings suggest that practical industrial localisation should be designed around the operational environment. In particular, fusion priorities differ by setting: in metal-dense factories, fusion should prioritise NLOS mitigation and absolute correction; in warehouses, it should prioritise scalable infrastructure and continuous coverage and in plant factories and structured robotic environments, it should prioritise drift suppression and real-time factor optimisation. Thus, future systems should move from sensor-centred design toward environment-aware fusion design.

4.2. Research Gaps

The purpose of identifying research gaps in this review is to explain why promising multimodal fusion methods have not yet translated into reliable, scalable, and widely adopted industrial asset-tracking systems. In relation to the review objective and RQ3, the gaps below are ranked according to their likely impact on implementation readiness. Together, they show that the main barrier is not the absence of accurate fusion algorithms, but the lack of validated, comparable, scalable, and maintainable systems for real industrial environments.

The highest-priority gap is the limited level of direct industrial validation. Although many reviewed studies report strong localisation performance, only a minority evaluate their systems in factories, warehouses, plant factories, industrial workshops, logistics settings, or automated guided vehicle scenarios. This limits the strength of conclusions that can be drawn about industrial deployment. Controlled laboratories, corridors, offices, and robotic testbeds are useful for understanding sensor behaviour and fusion mechanisms, but they do not fully reproduce the metallic infrastructure, dynamic layouts, vibration, occlusion, non-line-of-sight propagation, and operational constraints found in real factories and warehouses. Industrial studies such as TrackInFactory, real-world vehicle-based radio-map construction, preciseSLAM, and UWB–IMU fusion in industrial NLOS scenarios provide stronger deployment evidence, but such studies remain comparatively few [17–19,21]. This gap directly affects the review's central question because it limits how confidently results from transferable indoor localisation studies can be generalised to industrial asset tracking.

The second major gap is the absence of shared benchmarks and standardised evaluation protocols. This gap is closely connected to RQ2, because it prevents reliable comparison of fusion architectures and algorithms. Studies use different sensor combinations, target platforms, ground-truth systems, trajectories, motion speeds, update rates, and error metrics, including mean error, RMSE, MAE, percentile error, accuracy percentage, and qualitative robustness. As a result, reported improvements cannot be pooled quantitatively and must be interpreted as study-specific findings. Public resources such as UWB positioning and tracking datasets provide valuable range measurements, channel impulse response data, ground-truth positions, and tools for reproducible range-based positioning research, while architectures such as FITS support standardised data structures, synthetic data generation, and sensor-agnostic fusion experimentation. However, these resources do not yet constitute a common industrial benchmark for multimodal asset tracking across factories and warehouses [5,62]. Without standardised benchmarks, it remains difficult to determine whether one fusion method is genuinely more robust than another, or whether reported gains are mainly due to differences in environment, trajectory, or baseline selection.

The third priority is scalability from single-target or small-scale experiments to multi-asset industrial operation. Many studies evaluate one robot, one pedestrian, one vehicle, or a small number of tracked objects, whereas real warehouses and factories may require simultaneous tracking of hundreds of assets, forklifts, workers, tools, and automated guided vehicles. Scalability is not only a computational issue; it also involves data association, identity preservation, communication load, sensor density, update rate, and failure recovery. Architectures such as FITS and decentralised consensus-based IoT positioning begin to address multi-object correlation, high-workload tracking, and collaborative positioning; for example, FITS evaluates tracking performance under workloads up to 100,000 measurements and 1000 visitors, while consensus-based IoT positioning reduces reliance on centralised positioning infrastructure. However, large-scale validation in operational industrial environments remains limited [62,79]. This gap links RQ2 and RQ3 because an algorithm that is accurate for one target may not remain reliable when many moving assets compete for bandwidth, sensing coverage, and real-time processing.

The fourth gap concerns calibration, synchronisation, and map maintenance. These issues are often treated as implementation details, but they are central barriers to industrial adoption. High-performing fusion systems frequently depend on accurate extrinsic calibration, time alignment between sensors with different update rates, coordinate-frame registration, anchor placement, radio maps, visual maps, or LiDAR maps. In dynamic industrial environments, these assumptions can break when racks are moved, machinery is added, access points change, lighting varies, or metallic objects alter signal propagation. Vehicle-generated radio-map construction can reduce manual survey effort, while

tightly coupled SLAM and fusion approaches such as preciseSLAM, RSSI–PDR integration, UWB/LiDAR positioning, and enhanced particle-filter navigation show how global constraints, map-aware optimisation, and robust filtering can improve reliability under challenging indoor conditions. Nevertheless, most systems still require site-specific setup, calibration, infrastructure planning, or expert tuning [18,19,27,31,51]. This gap is important because it explains why technically accurate fusion systems may remain difficult to deploy and maintain in practice.

The fifth gap is the limited generalisability and explainability of learning-based fusion models. Deep learning, graph neural networks, and attention-based models show strong potential for modelling nonlinear sensor errors, especially in UWB–IMU fusion and NLOS mitigation [20,21,57]. However, these methods often depend on representative training data from the target environment. A model trained in one building, layout, or NLOS condition may not generalise to another factory without retraining or recalibration. In addition, the decision-making process of DNN- and GNN-based fusion remains difficult to interpret, which limits debugging, certification, trust, and adoption in safety-critical industrial environments. This gap connects directly to RQ2 because it affects whether learning-based fusion can be considered a robust alternative to classical filters such as Kalman filters, extended Kalman filters, and particle filters.

The sixth gap is privacy, governance, and human-centred deployment. This gap has lower technical visibility in the reviewed literature, but it is important for implementation. Industrial localisation systems may track workers, vehicles, tools, and production assets continuously. Vision-based and AI-assisted systems can improve safety monitoring and asset visibility, but they also raise concerns about worker privacy, surveillance, data ownership, and organisational trust [54,62]. Current studies acknowledge these issues only briefly and rarely propose technical or governance mechanisms such as privacy-preserving sensing, edge processing, access control, anonymisation, or transparent worker-facing policies. This gap is directly related to RQ3 because social acceptability and governance can determine whether technically effective systems are adopted in real workplaces.

These gaps are interdependent. Limited validation in industrial environments constrains the availability of realistic benchmark datasets, which in turn makes it difficult to assess generalisability or compare fusion architectures consistently. Even when systems perform well in laboratory studies, limited scalability and substantial calibration requirements can hinder their transition to operational use. Learning-based methods may improve robustness, but their effectiveness depends on access to representative industrial data and their behaviour needs to be explainable. Adoption also depends on addressing privacy and governance concerns, particularly for systems that continuously monitor workers and assets. Future research should, therefore, move beyond accuracy demonstrations in controlled environments and prioritise large-scale industrial validation, shared benchmarks, self-calibration, explainability, and privacy-aware system design.

4.3. Opportunities

The evidence synthesised in this review suggests that multimodal information fusion has progressed beyond demonstrating that combining multiple sensing modalities can improve localisation performance. The next stage of research should shift from improving localisation accuracy alone to intelligent localisation systems that are adaptive, scalable, interoperable, and capable of supporting industrial decision-making throughout the asset lifecycle. Based on the reviewed literature, five major research opportunities emerge.

The first opportunity is the development of context-aware adaptive fusion architectures. Many reviewed systems use fixed fusion strategies, where sensor weighting and fusion rules remain largely unchanged after deployment. However, industrial environ-

ments are dynamic, with changing layouts, moving machinery, variable radio propagation, intermittent non-line-of-sight conditions, and sensor failures. Evidence from UWB–IMU fusion, UWB–LiDAR fusion, and tightly coupled SLAM shows that localisation improves when complementary sensors compensate for each other's failure modes [19–21,51]. Future systems should, therefore, estimate sensor reliability in real time and automatically adjust the contribution of each modality according to environmental context. Important research questions include how localisation systems can detect sensor degradation, determine which sensing modality is most reliable under changing conditions, and continuously optimise fusion strategies without manual reconfiguration.

A second opportunity is the integration of localisation systems with industrial digital twins. Current localisation research primarily focuses on estimating position, whereas future Industry 4.0 environments require localisation data to become an active input to digital production systems. Continuous localisation information could support production planning, predictive maintenance, inventory management, autonomous logistics, and operational decision support. Studies such as TrackInFactory, vehicle-generated radio-map construction, UWB-based smart logistics, and industrial safety monitoring show that localisation information has operational value beyond position estimation, even though direct integration with digital twins remains underexplored [17,18,54,78]. Future research should examine how localisation data can continuously update digital representations of factory operations, how localisation uncertainty can be represented within digital models, and how multimodal fusion outputs can support predictive industrial analytics.

The third opportunity is the development of transferable learning models for multimodal sensor fusion. Recent studies increasingly employ deep neural networks, graph neural networks, attention mechanisms, and recurrent models for UWB positioning, UWB–IMU fusion, Wi-Fi–inertial fusion, and NLOS mitigation [15,20,21,57]. However, many learning-based approaches remain dependent on site-specific training data, sensor placement, environmental geometry, and local propagation conditions. Future research should investigate whether larger pre-trained, transfer-learning-based, or domain-adaptive multimodal models can learn representations that generalise across different factories, warehouses, and robotic environments with minimal recalibration. Key questions include what level of domain adaptation is required for a new industrial site, whether models trained in transferable indoor environments remain useful in industrial spaces, and how learning-based fusion compares with Kalman filtering, particle filtering, and factor-graph optimisation under unseen operating conditions.

The fourth opportunity is the development of explainable and human-centred localisation systems. Industrial localisation increasingly affects safety monitoring, autonomous vehicle coordination, worker tracking, and operational decision-making. As a result, localisation systems should not only provide accurate position estimates, but also communicate uncertainty, expose sensor disagreement, and support human trust. AI-based safety monitoring, semantic multisensor localisation, graph-based fusion, and multi-object tracking systems indicate the growing role of intelligent localisation in complex operational environments [21,62,70]. Future research should investigate how sensor contributions can be explained to operators, how localisation uncertainty can be visualised, how workers and supervisors can provide feedback to the system, and how accuracy can be balanced with privacy, transparency, and organisational trust.

The fifth opportunity is the evolution of self-managing localisation ecosystems. Current localisation systems often require manual calibration, infrastructure maintenance, map updates, and periodic reconfiguration. This limits long-term adoption in factories and warehouses where layouts, equipment, access points, anchor visibility, and signal propagation conditions change over time. Evidence from vehicle-generated radio maps, tightly coupled

plant-factory SLAM, robust particle filtering, and scalable data-tracking architectures suggests possible pathways toward more autonomous maintenance and recovery [18,19,27,62]. Future systems should continuously monitor sensor health, detect environmental change, recalibrate sensing infrastructure, update localisation models, and recover from sensor failures with minimal human intervention. Important questions include how such systems can identify outdated maps, diagnose faulty sensors, recover from missing modalities, and maintain performance over long-term industrial operation.

Collectively, these opportunities represent a transition from accurate positioning algorithms toward intelligent localisation ecosystems that function as core components of industrial information systems. Future research should, therefore, focus not only on improving localisation accuracy, but also on developing adaptive, explainable, transferable, scalable, and self-managing fusion architectures that enable reliable long-term deployment in real industrial environments.

4.4. Limitations of This Review

This review has several limitations. The relatively small number of studies conducted directly in industrial environments reflects the current maturity of the research field rather than the review methodology. First, the included studies are heterogeneous in terms of sensor combinations, target platforms, evaluation environments, and reported performance metrics. This prevented a quantitative meta-analysis and required a qualitative narrative synthesis. Second, only a subset of the included studies were evaluated directly in factories, warehouses, plant factories, or industrial logistics environments. Supporting studies from mobile robotics, pedestrian positioning, and smartphone-based localisation were included only where their findings were technically transferable to industrial localisation.

Third, reported accuracy values should be interpreted cautiously because studies used different ground-truth systems, trajectories, motion speeds, update rates, and environmental conditions. A method reporting centimetre-level accuracy in a controlled laboratory may not achieve the same performance in a metal-dense warehouse or a reconfigurable production line. Fourth, publication bias may favour studies reporting successful fusion performance, while failed deployments, maintenance burden, and negative industrial results are less visible in the literature. Fifth, independent dual screening was not performed for all records; therefore, inter-rater reliability statistics such as Cohen's kappa could not be reported. To reduce selection bias, predefined inclusion, exclusion, and quality assessment criteria were applied consistently, and the final set of included studies and extracted evidence was reviewed for consistency by the second author.

Despite these limitations, the review provides a structured synthesis of the available evidence and identifies consistent patterns across sensor modalities, fusion architectures, algorithms, and deployment challenges.

5. Conclusions

This systematic review synthesised 61 studies published between 2021 and 2026 on multimodal information fusion for indoor localisation and industrial asset tracking. The evidence base comprised 12 core industrial studies and 49 transferable supporting studies from related indoor-localisation contexts. Because only 12 of the 61 studies were conducted in industrial environments, much of the current evidence for industrial asset tracking remains indirect.

The findings show that no single sensing modality fully satisfies industrial requirements for accuracy, continuity, robustness, scalability, and deployability. RFID, UWB, vision, LiDAR, IMU/PDR, odometry, and vibration-based sensing each offer useful capabilities, but also suffer from limitations such as drift, NLOS sensitivity, occlusion, infrastructure

cost, calibration burden, or computational complexity. Multimodal fusion is, therefore, most valuable when it combines complementary modalities to reduce uncertainty and compensate for individual sensor weaknesses.

The main implication is that industrial localisation should be viewed as a data-integration and decision-support problem rather than only a sensor-selection problem. Reliable systems must fuse heterogeneous and asynchronous data streams, preserve asset identity, manage uncertainty, and deliver actionable location intelligence to industrial platforms. This highlights the importance of fusion architecture, data quality, synchronisation, and system integration.

This review contributes an evidence-informed synthesis of industrial and transferable studies, clarifying the roles of major sensing modalities, fusion architectures, reported benefits, and deployment challenges. However, the findings should not be interpreted as a pooled quantitative proof of universal superiority; instead, they should be read as indicative under the conditions of the reported study. Reported improvements are study-specific and depend on sensor combinations, environments, baselines, metrics, and validation protocols. Across all 61 included studies, verified improvements on absolute positioning error span approximately 11.3–96%; restricted to the 12 core industrial studies, they span approximately 21.4–72.7%, with all values recalculated from paired baseline and fused-system results on the same metric and condition.

Future research should prioritise full-scale industrial validation, shared benchmark datasets, standardised evaluation methods, scalable multi-asset experiments, self-calibrating fusion architectures, and explainable adaptive models. Overall, multimodal information fusion is a promising enabler of reliable Industry 4.0 asset tracking, but stronger real-world evidence is needed before its benefits can be generalised across industrial environments.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/info17111168/s1, Supplementary Table S1: PRISMA 2020 checklist; Supplementary Table S2: Evidence Matrix with Identity, Evidence classification, Modality coding, Fusion architecture, Performance Quality, and Risk of Bias of 61 included studies.

Author Contributions: Conceptualisation, S.S.; methodology, S.S.; formal analysis, S.S.; writing, original draft preparation, S.S.; writing, review and editing, S.S. and M.M.; supervision, M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new primary data were generated. The extracted evidence matrix used for this review is provided as Supplementary Table S2.

Acknowledgments: During the preparation of this manuscript, the authors used Claude (Anthropic) Instant (https://claude.com), web interface to support language editing, grammar correction, sentence restructuring, formatting consistency checks, image improvements, and organisation of review evidence into clearer narrative form. The tool was not used to select studies, extract data independently, perform analysis, generate results, or make conclusions. All AI-assisted text and suggestions were independently reviewed, fact-checked against the included studies, edited, and approved by the authors, who take full responsibility for the manuscript's content.

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

References

1. Zafari, F.; Gkelias, A.; Leung, K.K. A Survey of Indoor Localization Systems and Technologies. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 2568–2599. [CrossRef]
2. Zhang, X.D.; Tan, Y.C.; Tai, V.C.; Hao, Y.N. Indoor Positioning System for Warehouse Environment: A Scoping Review. *J. Mech. Eng. Sci.* **2024**, *18*, 10350–10381. [CrossRef]

3. Li, P.; Wu, W.; Zhao, Z.; Huang, G.Q. Indoor positioning systems in industry 4.0 applications: Current status, opportunities, and future trends. *Digit. Eng.* **2024**, *3*, 100020. [CrossRef]
4. Tsanousa, A.; Bektsis, E.; Kyriakou, T.; González, A.G.; Leturiondo, U.; Gialampoukidis, I.; Vrochidis, S.; Kompatsiaris, I. A Review of Multisensor Data Fusion Solutions in Smart Manufacturing: Systems and Trends. *Sensors* **2022**, *22*, 1734. [CrossRef]
5. Bregar, K. Indoor UWB Positioning and Position Tracking Data Set. *Sci. Data* **2023**, *10*, 733. [CrossRef]
6. Wang, X.; Sun, Z.; Chehri, A.; Jeon, G.; Song, Y. Deep learning and multi-modal fusion for real-time multi-object tracking: Algorithms, challenges, datasets, and comparative study. *Inf. Fusion* **2024**, *105*, 102247. [CrossRef]
7. Ušinskis, V.; Nowicki, M.; Dzedzickis, A.; Bučinskas, V. Sensor-Fusion Based Navigation for Autonomous Mobile Robot. *Sensors* **2025**, *25*, 1248. [CrossRef]
8. Karle, P.; Fent, F.; Huch, S.; Sauerbeck, F.; Lienkamp, M. Multi-Modal Sensor Fusion and Object Tracking for Autonomous Racing. *IEEE Trans. Intell. Veh.* **2023**, *8*, 3871–3883. [CrossRef]
9. Olejniczak, A.; Blaszkiewicz, O.; Cwalina, K.K.; Rajchowski, P.; Sadowski, J. Deep Learning Approach for LOS and NLOS Identification in the Indoor Environment. In Proceedings of the 2020 Baltic URSI Symposium (URSI), Warsaw, Poland, 5–8 October 2020; pp. 104–107. [CrossRef]
10. Wang, S.; Ahmad, N.S. A Comprehensive Review on Sensor Fusion Techniques for Localization of a Dynamic Target in GPS-Denied Environments. *IEEE Access* **2025**, *13*, 2252–2285. [CrossRef]
11. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ* **2021**, *372*, n71. [CrossRef]
12. Kitchenham, B.; Charters, S. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*; Technical Report; Keele University: Newcastle, UK; Durham University: Durham, UK, 2007.
13. Brereton, P.; Kitchenham, B.A.; Budgen, D.; Turner, M.; Khalil, M. Lessons from Applying the Systematic Literature Review Process within the Software Engineering Domain. *J. Syst. Softw.* **2007**, *80*, 571–583. [CrossRef]
14. *802.15.4z-2020*; IEEE Standard for Low-Rate Wireless Networks–Amendment 1: Enhanced Ultra Wideband (UWB) Physical Layers (PHYs) and Associated Ranging Techniques. IEEE: New York, NY, USA, 2020. [CrossRef]
15. Zhang, M.; Jia, J.; Chen, J.; Deng, Y.; Wang, X.; Aghvami, A.H. Indoor Localization Fusing WiFi With Smartphone Inertial Sensors Using LSTM Networks. *IEEE Internet Things J.* **2021**, *8*, 13608–13623. [CrossRef]
16. Abdellatif, A.G.; Salama, A.A.; Zied, H.S.; Elmahallawy, A.A.; Shawky, M.A. An Improved Indoor Positioning Based on Crowd-Sensing Data Fusion and Particle Filter. *Phys. Commun.* **2023**, *61*, 102225. [CrossRef]
17. Silva, I.; Pendão, C.; Torres-Sospedra, J.; Moreira, A. TrackInFactory: A Tight Coupling Particle Filter for Industrial Vehicle Tracking in Indoor Environments. *IEEE Trans. Syst. Man Cybern. Syst.* **2022**, *52*, 4151–4163. [CrossRef]
18. Silva, I.; Pendão, C.; Moreira, A. Real-World Deployment of Low-Cost Indoor Positioning Systems for Industrial Applications. *IEEE Sens. J.* **2022**, *22*, 5386–5397. [CrossRef]
19. Gong, L.; Gao, B.; Sun, Y.; Zhang, W.; Lin, G.; Zhang, Z.; Li, Y.; Liu, C. preciseSLAM: Robust, Real-Time, LiDAR–Inertial–Ultrasonic Tightly-Coupled SLAM With Ultraprecise Positioning for Plant Factories. *IEEE Trans. Ind. Inform.* **2024**, *20*, 8818–8827. [CrossRef]
20. Muthineni, K.; Artemenko, A.; Vidal, J.; Nájar, M.; Catalan, M.; Paradells, J. Deep Learning-Based UWB-IMU Data Fusion for Indoor Positioning in Industrial Scenario. *IEEE Open J. Veh. Technol.* **2025**, *6*, 1209–1221. [CrossRef]
21. Muthineni, K.; Artemenko, A.; Abode, D.; Vidal, J.; Nájar, M. PosGNN: A Graph Neural Network Based Multimodal Data Fusion for Indoor Positioning in Industrial Non-Line-of-Sight Scenarios. *IEEE Open J. Veh. Technol.* **2026**, *7*, 15–26. [CrossRef]
22. Ammanabrolu, M.L. Indoor Way-Finding Method Using IMU and Magnetic Tensor Sensor Measurements for Visually Impaired Users. *Int. J. Intell. Robot. Appl.* **2021**, *5*, 264–282. [CrossRef]
23. Sheshadri, S.; Hara, K. Enhancing Smartphone-based Inertial Indoor Tracking with Conversational User Input. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2025**, *9*, 1–37. [CrossRef]
24. Wu, L.; Guo, S.; Han, L.; Anil Baris, C. Indoor positioning method for pedestrian dead reckoning based on multi-source sensors. *Measurement* **2024**, *229*, 114416. [CrossRef]
25. Sharma, S.; Kaur, M.; Kumar, N. Enhanced 3D Indoor Position Estimation from Smartphone Inbuilt IMU and Pressure Sensor via Hybrid Approach. *Clust. Comput.* **2024**, *27*, 10953–10970. [CrossRef]
26. Medjahed, R.; Louali, R.; Messaoui, A.A.; Hebbache, N.; Nemra, A.; Azem, B.M.; Belounis, L.; Khellal, A.; Boulahmar, M. Particle Filter-Based Fusion of PDR and Geomagnetic Field for Indoor Localization. In Proceedings of the 2026 34th Mediterranean Conference on Control and Automation (MED), Ancona, Italy, 23–26 June 2026; pp. 61–67. [CrossRef]
27. Guo, K.; Zhang, J.; Chen, J. A Robust Indoor Navigation Method Using an Enhanced Particle Filter in Complex Architectural Environments. *Results Eng.* **2025**, *27*, 106525. [CrossRef]

28. Zhou, T.; Sun, J.; Shen, N.; Liu, J.; Chen, L. Factor Graph Optimization Coupled With Adaptive Unscented Kalman Filter for Real-Time Indoor Localization. *IEEE Internet Things J.* **2026**, *13*, 2880–2891. [CrossRef]
29. Zhou, X.; Chen, L.; Chen, Y.; Yin, H.; Chen, X.; Wang, W. Fusion of IMU and Probabilistic Model for Indoor Localization Based on Bayesian Framework. *IEEE Internet Things J.* **2025**, *12*, 17080–17094. [CrossRef]
30. Wang, M.; Shi, L.; Li, Y. Indoor Vehicle Positioning Based on Multi-Sensor Data Fusion. *J. Shanghai Jiaotong Univ. (Sci.)* **2023**, *28*, 77–85. [CrossRef]
31. Mehrabian, H.; Ravanmehr, R. Sensor Fusion for Indoor Positioning System through Improved RSSI and PDR Methods. *Future Gener. Comput. Syst.* **2023**, *138*, 254–269. [CrossRef]
32. Masoudi, B.; Rezazadeh, J. IoT-Enabled Indoor Real-Time Tracking Using UWB for Smart Warehouse Management. *Computers* **2025**, *14*, 510. [CrossRef]
33. Singh, A.; Karthikeyan, R.; Kalaichelvi, V. Machine learning-based multi-sensor fusion for warehouse robot in GPS-denied environment. *Multimed. Tools Appl.* **2023**, *83*, 56229–56246. [CrossRef]
34. Liu, C.; Kadja, T.; Chodavarapu, V.P. Experimental Evaluation of Sensor Fusion of Low-Cost UWB and IMU for Localization under Indoor Dynamic Testing Conditions. *Sensors* **2022**, *22*, 8156. [CrossRef]
35. Teng, X.; Shen, Z.; Huang, L.; Li, H.; Li, W. Multi-sensor fusion based wheeled robot research on indoor positioning method. *Results Eng.* **2024**, *22*, 102268. [CrossRef]
36. Ruzaij Al-Okby, M.F.; Roddelkopf, T.; Thurow, K. Sensor Fusion of UWB and IMU for Optimized Indoor Positioning Using Kalman Filtering. In Proceedings of the 2026 IEEE 30th International Conference on Intelligent Engineering Systems (INES), Budapest, Hungary, 2–4 July 2026; pp. 389–394. [CrossRef]
37. Cai, Q.; Huang, W.; Li, J.; Marelli, D.; Li, P.; Meng, W. Continuous Indoor Positioning Under Random UWB Channel Occlusions Using RBEKF for UWB/IMU Data Fusion. *IEEE Trans. Autom. Sci. Eng.* **2026**, *23*, 16011–16023. [CrossRef]
38. Zheng, B.; Liu, F.; Ping, Y.; Zhang, T. Factor Graph Optimization Based Data Fusion for Accurate 3-D Positioning Using UWB and IMU. In Proceedings of the 2026 IEEE International Conference on Communications Workshops (ICC Workshops), Glasgow, UK, 24–28 May 2026; pp. 1–6. [CrossRef]
39. Zhang, H.; Jiang, Y.; Deng, X.; Liu, K. GraphKANLoc: Physics-Informed Heterogeneous Graph Attention Networks With KAN Regression for Indoor Localization. *IEEE Trans. Mob. Comput.* **2026**, *25*, 14988–14999. [CrossRef]
40. Wang, J.; Duan, L.; Zhang, Z.; Xia, M.; Shi, C. High-Precision CSI-Based Continuous Indoor Localization With ConvNeXt Deep Representation and Main-Auxiliary Dual-Stream Fusion. *IEEE Sens. J.* **2026**, *26*, 15980–15992. [CrossRef]
41. Seo, J.; Choi, J.; Lee, J. Infra-Free Indoor Localization via On-Demand Wi-Fi Aware FTM with Extended Particle Filter. In Proceedings of the ICC 2026—IEEE International Conference on Communications, Glasgow, UK, 24–28 May 2026; pp. 1–6. [CrossRef]
42. Cao, E. Multi-Sensor Information Fusion for Indoor Positioning Based on UWB and IMU with Deep Learning Optimization. In Proceedings of the 2026 6th International Conference on Sensors and Information Technology (ICSI), Zhengzhou, China, 24–26 April 2026; pp. 439–442. [CrossRef]
43. Xie, P.; Lou, X.; Mao, D. Adaptive Sage-Husa Kalman Filter for 3D Indoor Positioning with Multi-Sensor Fusion. *Comput. Electr. Eng.* **2026**, *140*, 111530. [CrossRef]
44. Cheng, L.; Chen, Y.; Guan, J. INS/UWB multi-sensor integrated localization scheme in indoor LOS/NLOS environment based on MEF-GPDAF. *Comput. Electr. Eng.* **2024**, *120*, 109693. [CrossRef]
45. Liu, Y.A.; Zhang, L.; Fan, Y.; Yang, Z. EVIU-Link: Composite Collaborative Perception Framework for Enhanced Indoor Localization. *IEEE Trans. Instrum. Meas.* **2026**, *75*, 9505510. [CrossRef]
46. Qin, N.; Zhao, Y.; Yang, L. Indoor 3-D Localization Based on Heterogeneous Sensing Data Fusion. *IEEE Internet Things J.* **2025**, *12*, 33230–33242. [CrossRef]
47. Zhou, Z.; Lu, Y.; Xia, J. Research on the Fusion Positioning System of UWB/LiDAR Based on the Algorithm of SPF and KF. *Phys. Commun.* **2024**, *67*, 102529. [CrossRef]
48. Yang, B.; Li, J.; Zhang, H. Resilient Indoor Localization System Based on UWB and Visual–Inertial Sensors for Complex Environments. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 8504014. [CrossRef]
49. Hu, C.; Huang, P.; Wang, W. Tightly Coupled Visual-Inertial-UWB Indoor Localization System With Multiple Position-Unknown Anchors. *IEEE Robot. Autom. Lett.* **2024**, *9*, 351–358. [CrossRef]
50. Xu, Y.; Chen, Z.; Zhao, M.; Tang, F.; Li, Y.; Liu, J.; Kato, N. UVtrack: Multi-Modal Indoor Seamless Localization Using Ultra-Wideband Communication and Vision Sensors. *IEEE Open J. Comput. Soc.* **2025**, *6*, 272–281. [CrossRef]
51. Wang, X.; Gao, F.; Huang, J.; Xue, Y. UWB/LiDAR Tightly Coupled Positioning Algorithm Based on ISSA Optimized Particle Filter. *IEEE Sens. J.* **2024**, *24*, 11217–11228. [CrossRef]
52. Tran, Q.K.; Ryoo, Y.J. Multi-Sensor Fusion Framework for Reliable Localization and Trajectory Tracking of Mobile Robot by Integrating UWB, Odometry, and AHRS. *Biomimetics* **2025**, *10*, 478. [CrossRef]

53. Liu, A.; Wang, J.; Lin, S.; Kong, X. A Dynamic UKF-Based UWB/Wheel Odometry Tightly Coupled Approach for Indoor Positioning. *Electronics* **2024**, *13*, 1518. [CrossRef]
54. Di Rienzo, F.; Miglionico, G.C.; Ducange, P.; Marcelloni, F.; Salti, N.; Vallati, C. Improving Accuracy in Industrial Safety Monitoring: Combine UWB Localization and AI-Based Image Analysis. *J. Sens. Actuator Netw.* **2025**, *14*, 118. [CrossRef]
55. Cheng, L.; Zhao, F.; Zhao, W. An INS/UWB Joint Indoor Positioning Algorithm Based on Hypothesis Testing and Yaw Angle. *Wirel. Netw.* **2025**, *31*, 457–469. [CrossRef]
56. Cheng, L.; Cui, C.; Zhao, Z.; Shi, Y. An Improved Multi-Filter Fusion Indoor Localization Algorithm Based on INS and UWB. *Earth Sci. Inform.* **2024**, *17*, 2509–2521. [CrossRef]
57. He, L.; Lian, Z.; Núñez-Andrés, M.A.; Chen, H.; Zhao, C.; Tian, Y. Application of a Temporal Convolutional Network Algorithm Fused with Channel Attention Module for UWB Indoor Positioning. *Sci. Rep.* **2026**, *16*, 6305. [CrossRef]
58. Zheng, Y.; Lu, Z.; Chu, X.; Zhou, Y.; Luo, Z.; Hu, Z.; Guo, Y. Scalable Multimodal Localization for Underground Parking Lots Using Distributed Antenna Systems. *IEEE Internet Things J.* **2026**, *13*, 36600–36613. [CrossRef]
59. Boudlal, H.; Serrhini, M.; Tahiri, A. Towards a Low-Cost and Privacy-Preserving Indoor Activity Recognition System Using WiFi Channel State Information. *Multimed. Tools Appl.* **2025**, *84*, 35761–35792. [CrossRef]
60. Zhou, C.; Xiong, R.; Zeng, H.; Xiao, J.; Wang, Y.; Jia, P.; Ye, J.; Zhao, T.; Hu, K. Aerial Locating Method Design for Civil Aviation RFI: UAV Monitoring Platform and Ground Terminal System. *J. Intell. Robot. Syst.* **2021**, *103*, 29. [CrossRef]
61. Qin, P.; Hu, Q.; Yu, H. An Internet of Electronic-Visual Things Indoor Localization System Using Adaptive Kalman Filter. *IEEE Sens. J.* **2023**, *23*, 16058–16067. [CrossRef]
62. Rodrigues, B.; Scheid, E.J.; Willems, J.; Tornow, M.; Müller, K.O.E.; Stiller, B. FusIon Data Tracking System (FITS). *IEEE Sens. J.* **2022**, *22*, 19060–19072. [CrossRef]
63. Wei, Z.; Li, J.; Lian, B.; Wang, W.; Zhao, J. WiNAL: Joint Feature Fusion and Adversarial Training for Non-Line-of-Sight WiFi Localization. *Clust. Comput.* **2025**, *28*, 802. [CrossRef]
64. Li, Z.; Chamran, K.; Alobaedy, M.M.; Sheikh, M.A.; Siddiqui, T.; Ahad, A. An Optimization on 2D-SLAM Map Construction Algorithm Based on LiDAR. *J. Intell. Robot. Syst.* **2024**, *110*, 144. [CrossRef]
65. Chen, J.; Song, S.; Gong, Y.; Zhang, S. An Indoor Fusion Navigation Algorithm Using HV-Derivative Dynamic Time Warping and the Chicken Particle Filter. *Satell. Navig.* **2022**, *3*, 13. [CrossRef]
66. Kien, N.T.; Van Hung, N.; Nhat Hai, V.Q.; Xuan Son, N.; Chinh, H.D. An Improved Indoor Positioning System based on Wi-Fi, BLE, and IMU Fusion using Attention Mechanism. In Proceedings of the 2026 8th International Conference on Green Technology and Sustainable Development (GTSD), Ho Chi Minh City, Vietnam, 30–31 July 2026; pp. 614–619. [CrossRef]
67. Xu, R.; Liu, X.; Wang, X.; Wen, W.; Huang, Y. RTT-LIO: A Wi-Fi RTT-Aided LiDAR-Inertial Odometry via Tightly-Coupled Factor Graph Optimization in Complex Scenes. *IEEE Internet Things J.* **2026**, *13*, 22603–22617. [CrossRef]
68. Zhang, L.; Zhang, L.; Gao, R.; Pan, L.; Xu, C.; Cheng, K. Indoor Mobile Robot Localization Applying IMU/Stereo Camera/LiDAR and Graph-Based Optimization. *IEEE Sens. J.* **2024**, *24*, 21466–21478. [CrossRef]
69. Li, Y.; Tan, L.; Xu, X.; Zhang, Z. Room-Level Localization Method in Industrial Workshops Using LiDAR-Based Point Cloud Registration and Object Recognition. *Appl. Intell.* **2025**, *55*, 373–393. [CrossRef]
70. Li, Y.; Jiang, L.; Lei, B.; Tang, B.; Zhu, J.; Liu, H. SIMSL: A Semantic Instance-Based Multisensor Fusion Localization System for Challenging Dynamic and Degraded Environments. *IEEE Trans. Ind. Electron.* **2025**, *72*, 13647–13657. [CrossRef]
71. Li, Z.; Wang, C.; Zhao, X. CrowdGraph: Crowdsourcing Positioning Based on Multimodal Graph Attention Model. *IEEE Trans. Ind. Inform.* **2026**, *22*, 6558–6568. [CrossRef]
72. Belbachir, A.; Ortiz, A.M.; Hauge, E.T.; Belbachir, A.N.; Bonanno, G.; Ciccia, E.; Felline, G. Outdoor Warehouse Management: UAS-Driven Precision Tracking of Stacked Steel Bars. *SN Comput. Sci.* **2025**, *6*, 701. [CrossRef]
73. Yi, P.; Li, Z.; Braun, T. mmDif-DRL: Diffusion-Deep Reinforcement Learning for Collaborative mmWave Positioning. *IEEE Trans. Mob. Comput.* **2026**, *25*, 16897–16914. [CrossRef]
74. Wu, Z.; Hu, P.; Huang, Y.; Liu, N.; Niu, Q.; Liang, L.; Pang, T. HIPS: A Spatial-Temporal Attention-Guided Framework for Accurate Indoor Positioning via 5G/BLE Signal Fusion. *IEEE Trans. Veh. Technol.* **2026**, *75*, 14489–14505. [CrossRef]
75. Zhang, L.; Tao, Y.; Lin, J.; Zhang, F.; Fallon, M. Visual Localization in 3D Maps: Comparing Point Cloud, Mesh, and NeRF Representations. *Auton. Robots* **2026**, *50*, 14. [CrossRef]
76. Krishnaveni, B.V.; Reddy, K.S.; Reddy, P.R. Indoor Tracking by Adding IMU and UWB Using Unscented Kalman Filter. *Wirel. Pers. Commun.* **2022**, *123*, 3575–3596. [CrossRef]
77. Hayward, S.J.; Earps, J.; Sharpe, R.; van Lopik, K.; Tribe, J.; West, A.A. A Novel Inertial Positioning Update Method, Using Passive RFID Tags, for Indoor Asset Localisation. *CIRP J. Manuf. Sci. Technol.* **2021**, *35*, 968–982. [CrossRef]

78. Elsanhoury, M.; Mäkelä, P.; Koljonen, J.; Välisuo, P.; Shamsuzzoha, A.; Mantere, T.; Elmusrati, M.; Kuusniemi, H. Precision Positioning for Smart Logistics Using Ultra-Wideband Technology-Based Indoor Navigation: A Review. *IEEE Access* **2022**, *10*, 44413–44445. [CrossRef]
79. Chernov, A.; Erofeeva, V.; Granichin, O.; Moseiko, E.; Sudomir, A. SPSA-Based Consensus Algorithm for Mutual IoT Device Positioning. *IFAC-PapersOnLine* **2025**, *59*, 133–138. [CrossRef]

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.