Automation And Sample Classification Algorithms For Industrial Spectroscopy-based Systems

João Francisco Caramelo de Carvalho
Master in Engineering Physics
Department of Physics and Astronomy
Faculty of Sciences of the University of Porto
Faculty of Engineering of the University of Porto
2025

Automation And Sample Classification Algorithms For Industrial Spectroscopy-based Systems

João Francisco Caramelo de Carvalho

Project report carried out as part of the Master in Engineering Physics

Department of Physics and Astronomy

2025

Supervisor
Prof. Dr. Nuno Silva, Area Manager, INESC TEC

Co-supervisor
Prof. Dr. Pedro Jorge, Professor Auxiliar, Faculdade de Ciências da Universidade do Porto

Co-supervisor
Dra. Diana Guimarães, Investigadora Auxiliar, INESC TEC

Acknowledgments

Ao concluir este trabalho, gostaria de aproveitar esta oportunidade para, de certa forma, eternizar a minha porfunda gratidão a todas as pessoas que, de diferentes formas, contribuíram para a sua concretização.

Em primeiro lugar, quero agradecer aos meus orientadores. Ao Professor Nuno Silva, pela orientação atenta e pelo acompanhamento constante, mas sobretudo pela qualidade das sugestões e comentários construtivos que foram determinantes para o meu crescimento científico e para o desenvolvimento deste projeto. Ao Professor Pedro Jorge, agradeço a oportunidade de integrar um ambiente de investigação tão estimulante e, em conjunto com a Dra. Diana Guimarães, pelo apoio, incentivo e disponibilidade desmonstrados ao longo de todo este percurso.

Dirijo igualmente um sincero agradecimento a toda a equipa do Centro de Fotónica Aplicada do INESC TEC, pelo acolhimento e pela forma como contribuíram, direta ou indiretamente, para o progresso deste trabalho. Deixo uma menção especial à Diana e ao Rafael, pela ajuda incansável e inestimável, sem a qual este projeto não teria sido possível.

Por fim, mas com igual importância, agradeço profundamente aos meus pais, ao meu irmão, à minha família e aos meus amigos pelo apoio incondicional ao longo de todos estes anos, tanto dentro como fora do percurso académico. O vosso incentivo constante e a vossa presença foram fundamentais para tornar este caminho possível.

Resumo

Esta tese baseia-se na integração estratégica de duas tecnologias complementares, visão computacional e espetroscopia de emissão induzida por laser (LIBS), com o objetivo de desenvolver um ambiente híbrido de inspeção multimodal orientado para a deteção automática de contaminantes em materiais sólidos.

A crescente exigência de controlo de qualidade e segurança nas linhas de produção industriais reforça a necessidade de sistemas de inspeção capazes de identificar, de forma rápida e fiável, a presença de contaminantes. Contudo, metodologias tradicionais de monitorização, baseadas exclusivamente numa única tecnologia, revelam limitações e deixam claro o espaço existente para soluções inovadoras que aliem performance analítica com escalabilidade industrial.

Partindo dos avanços obtidos por cada tecnologia de forma independente, esta investigação visa superar as suas limitações intrínsecas quando aplicadas isoladamente, através da conceção e implementação de uma infraestrutura física e digital que integra a aquisição, sincronização e processamento de dados visuais e espectrais. O sistema resultante constitui uma solução robusta, automática e adaptável, capaz de conjugar características morfológicas com informações composicionais de modo mais completo e fiável, sendo aplicável tanto em cenários laboratoriais como em linhas industriais dinâmicas e de alta cadência.

Para explorar diferentes graus de integração e responder a requisitos operacionais distinto, foram desenvolvidos três modos de funcionamento: o “Modo Matriz”, orientado para caracterização espacial detalhada, o “Modo Singular” direcionado para a deteção seletiva em fluxo e o “Modo Contínuo” concebido para responder a cenários de elevada densidade e velocidade. Estes modos representam estratégias complementares de sinergia entre os módulos de visão computacional e LIBS, refletindo uma evolução progressiva desde abordagens estáticas e mais detalhadas até metodologias de análise contínua em tempo real.

Em conjunto, os resultados demonstram que a integração multimodal permitem ultrapassar as limitações estruturais das tecnologias individuais, ampliando a fiabilidade, a seletividade e escalabilidade do processo de deteção, e abrindo caminho para a evolução e transição de soluções laboratoriais em direção a sistemas industriais não apenas automatizados, mas também ambientalmente conscientes.

Palavras-chave: Fusão Multimodal, Visão Computacional, Espetroscopia LIBS, Inspeção em Linha, Deteção de Contaminantes

Abstract

This thesis focuses in integrating two complementary technologies - computer vision and laser-induced breakdown spectroscopy (LIBS) - with the aim of developing a hybrid multimodal inspection environment designed for the automatic detection of contaminants in solid materials.

The increasing demand for quality control and safety in industrial production lines reinforces the need for inspection systems capable of rapidly and reliably identifying the presence of contaminants. However, traditional monitoring methodologies, based exclusively on a single technology, reveal inherent limitations and clearly highlight the scope for innovative solutions that combine analytical precision with industrial scalability.

Building upon the advances achieved by each technology independently, this research seeks to overcome their intrinsic limitations when applied separately through the design and implementation of a physical and digital infrastructure that integrates the acquisition, synchronisation, and processing of both visual and spectral data. The resulting system constitutes a robust, automatic, and adaptable solution, capable of combining morphological characteristics with compositional information in a more comprehensive and reliable manner, with applicability ranging from laboratory scenarios to dynamic, high-throughput industrial environments.

To explore different levels of integration and address distinct operational requirements, three operating modes were developed: the “Matrix Mode”, aimed at detailed spatial characterization, the “Singular Mode” directed towards selective detection in flow and the “Continuous Mode” designed to respond to scenarios of high density and speed. These modes represent complementary strategies for achieving synergy between computer vision and LIBS modules, reflecting a progressive evolution from static, more detailed approaches towards continuous real-time analysis methodologies.

Taken together, the results demonstrate that multimodal integration enables the system to overcome the structural limitations of individual technologies, thereby enhancing reliability, selectivity, and scalability in contaminant detection, while paving the way for the evolution and transition of laboratory-based solutions towards industrial systems that are not only automated but also environmentally conscious.

Keywords: Multimodal Fusion, Computer Vision, LIBS Spectroscopy, Inline Inspection, Contaminant Detection

Table of Contents

List of Figures viii

List of Abbreviations xi

1. Introduction 1
   1.1. Motivation: Seeing Beyond the Visible Range 1
   1.2. Shaping an Analytical Synergy for Computer Vision and Laser-Induced Breakdown Spectroscopy 3
   1.3. Dissertation Objectives and Structure 5
   1.4. List of works and publications 7
2. Scientific Context and Challenges 8
   2.1. Industrial and Environmental Impact of Contaminants 8
   2.2. The Barriers of Conventional Analysis in Real-World Scenarios 10
   2.3. Two-Stroke Technology: The Discrete Power of Computer Vision and Laser Induced Breadown Spectroscopy 11
   2.4. Fusing Technology to Advance the State of the Art 16
   2.5. Concluding Remarks 17
3. Hardware Architecture and Acquisition Systems 19
   3.1. Functional Integration for Multimodal Inspection 19
   3.2. Computer Vision Module for Automated Analysis 21
   3.3. Laser Induced Breadown Spectroscopy (LIBS) Module for Precision Elemental Analysis 23
   3.4. Motion Module for Localised Chemical Analysis 25
   3.5. Module Integration Regarding Temporal Coordination and Spatial Alignment .26
   3.6. Concluding Remarks 28
4. Software Layer and Algorithms 30
   4.1. Computer Vision Model Architecture and Strategy 30
   4.2. Inferring Contaminants with LIBS Data 36
   4.3. Concluding Remarks 38
5. Multimodal Data Fusion and Integration Strategies 40
   5.1. Matrix Mode for High-Precision Mapping 41
   5.2. Singular Mode for Directed and Vision-Guided Analysis 43
   5.3. Continuous Mode for Continuous Acquisition and Full Autonomy 47

5.4. Concluding Remarks 50

6. Conclusion 51

6.1. Perspectives and Future Work 52

References 53

List of Figures

1.1: Comparative diagram illustrating the gap between human sensory limitations and technological extensions, illustrating how technology transcends biological constraints to expand perception and analytical capabilities..... 2
1.2: Conceptual scheme of multimodal fusion, where computer vision and LIBS spectroscopy are integrated into a unified framework, combining spatial context and compositional evidence for robust contaminant detection..... 4
2.1: Diagram of anthropogenic pressures and contaminant pathways, linking population growth and associated waste streams to the release of contaminants and their consequent impacts on water, soil and air..... 9
2.2: Analytical trade-off diagram mapping industrial throughput (speed, automation and robustness) against analytical confidence (precision, sensitivity and specificity), highlighting the design space for innovative solutions and positioning the proposed computer vision and LIBS fusion within that space..... 11
2.3: Example of computer vision deployment in environmental monitoring: aerial imagery used for automated detection and quantification of livestock, illustrating computer vision capacity for large-scale spatial surveillance..... 13
2.4: Field and planetary deployment of LIBS: rover-mounted LIBS analysis (Perseverance/SuperCam) demonstrating in-situ elemental interrogation of geological targets and the technique’s operational suitability for remote compositional analysis.. 15
2.5: Venn diagram of complementary strengths of computer vision and LIBS spectroscopy. The intersection (fusion) denotes spatially guided LIBS sampling, combined spatial-compositional analysis, and improved robustness in contaminant detection and decision-making..... 16
3.1: High-level functional architecture of the integrated multimodal inspection system, presenting sample intake and conveyance, the computer vision and LIBS paths, control and synchronisation, and the downstream fusion and actuation stages to provide an intuitive overview of data and control flow..... 20
3.2: Rear view of the system, the LIBS module, responsible for chemical analysis (left), while the system’s entry point (right) features the upstream Computer Vision subsystem for preliminary visual inspection..... 21
3.3: Comparative illustration of global-shutter and rolling-shutter acquisition under motion, emphasising motion-induced geometric distortion in rolling-shutter images and

the undistorted capture achieved with a global-shutter in dynamic inspection contexts.. . 23
3.4: LIBS subsystem: photograph of the assembled instrument paired with a technical block diagram of laser, spectrometer and DDG, including a representative emission spectrum with annotated elemental lines..... 24
3.5: Gantry motion geometry and implementation: left, 3D schematic showing X-Y-Z axes and principal degrees of freedom; right, photographic view of the physical gantry and its workspace..... 26
3.6: Temporal sequence diagram showing the ordered progression from computer vision assessment to LIBS spectral interrogation, culminating in integrated decision-making and actuation..... 27
3.7: Calibration and synchronisation workflow linking camera pixel coordinates to gantry millimetre coordinates via ArUco markers and transformation matrices, establishing metric alignment for spatially guided LIBS sampling..... 28
4.1: Comparative visualisation of segmentation paradigms (left: original image; centre: instance segmentation), highlighting instance segmentation as the selected strategy for its capacity to individually delineate and trace samples across the pipeline. right: semantic segmentation (top) and panoptic segmentation (bottom) for comparison..... 31
4.2: Dataset composition and partitioning: absolute sample counts per split and class-wise percentage distribution used for training, validation and testing. (numbers in parentheses denote annotation/mask counts)..... 33
4.3: Comparative performance of evaluated segmentation models for box (left) and mask (right) metrics, showing YOLOv11 in relation to Mask R-CNN, YOLOv8 and YOLOv12. (bottom) confusion matrix for the YOLOv11 model..... 34
4.4: Per-class mask-performance heatmap comparing precision, recall, mAP50 and mAP50-95 across models for the target classes (glass, rock, wood)..... 35
4.5: Schematic of LIBS plasma dynamics and spectral emission, illustrating the sequence from laser-induced ablation and plasma formation to the emission and collection of characteristic spectral lines..... 36
4.6: Processing pipeline for LIBS spectra: raw acquisition, pre-processing (baseline correction and normalisation), peak detection and matching against a contaminant reference database..... 37
4.7: Representative LIBS result for a painted wood sample: spectral peaks compared with the contaminant emission-line reference table (match Ti), illustrating targeted elemental identification..... 38

5.1: Matrix Mode demonstration: left, schematic of the gantry-driven laser raster scanning the halted sample on the conveyor; right, representative pictures illustrating execution on a tile..... 42
5.2: Matrix Mode spectral mapping results for a tile: single-wavelength maps, a vertical-layer composite, and a top-view contaminant mask that localises spectral signatures..... 43
5.3: Multi-channel spectral maps presented side-by-side to emphasise wavelength-dependent spatial contrast across the analysed tile..... 43
5.4: Focused operational pipeline for the Singular Mode, showing the ordered sequence from sample intake and visual analysis through candidate assignment and scheduling to fusion and actioning..... 45
5.5: Singular Mode demonstration with sparse samples: left, schematic illustrating the laser sweep/filter over the conveyor; right, selected frames from the demonstration video showing targeted shots on isolated samples..... 46
5.6: Dedicated user interface developed for the system, presenting the contamination class, captured image, and complementary information such as the type of contaminant detected..... 47
5.7: Continuous Mode demonstration under high sample density: left, schematic of oscillatory laser coverage across the moving conveyor belt; right, frames from the demonstration video showing continuous operation on densely packed samples..... 48
5.8: Focused operational pipeline for the Continuous Mode, showing the stamp-driven shot scheduling, gantry motion control and LIBS analysis feeding the fusion and actioning stage..... 49

List of Abbreviations

CV Computer Vision

Gantry Robotic System

LIBS Laser Induced Breadown Spectroscopy

YOLO You Only Look Once

1. Introduction

1.1. Motivation: Seeing Beyond the Visible Range

As observing beings, we possess characteristics that allow us to continuously explore and redefine the boundaries of knowledge. Our innate curiosity and human senses drive continuous discovery of the world. Yet, despite the fundamental importance of the human senses that we inherited through evolution, they still present constraints when confronted with the ambition to explore the unknown and to reach beyond our natural capabilities.

These limitations become particularly evident in challenging applications such as industrial contexts, where simple visual observation is not sufficient to detect critical risks and meet the required throughput. Materials that appear harmless, such as wood, glass, or rocks, may contain contaminants not only of organic and biological origin, but also radioactive and metallic, in concentrations high enough to compromise safety, the quality of the final product, and the sustainability of production processes. For instance, recent studies [1] indicate that a single kilogram of treated wood may contain the equivalent of a teaspoon of undesirable metallic contaminants, whether of natural or anthropogenic origin, which can affect the physical and chemical properties of materials, but more importantly, represent a sufficient load to generate significant environmental impacts and health risks, particularly when accumulated in large volumes.

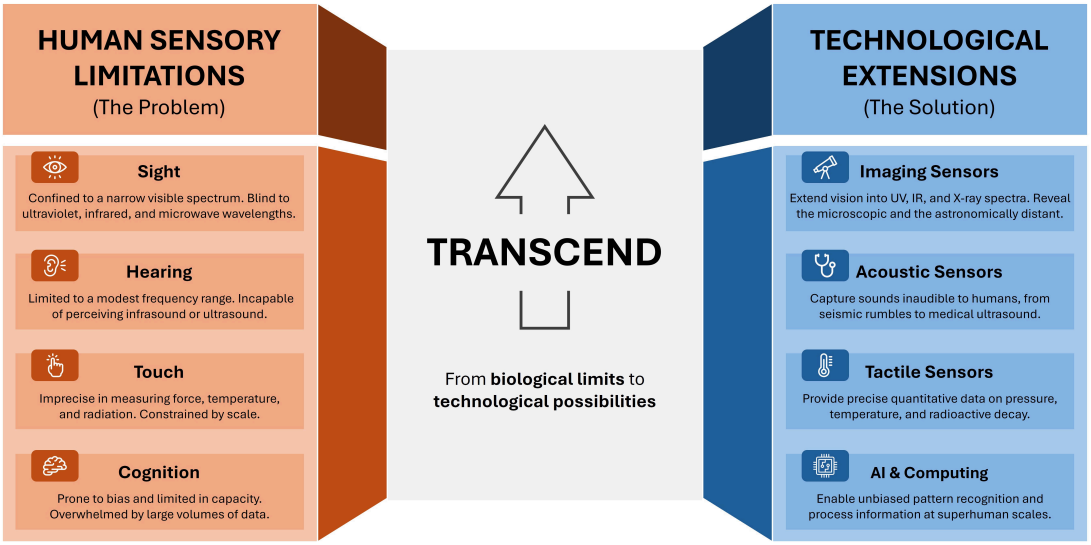


Figure 1.1: Comparative diagram illustrating the gap between human sensory limitations and technological extensions, illustrating how technology transcends biological constraints to expand perception and analytical capabilities.

In such contexts, where the challenges of perception far exceed human sensory capabilities, technologies emerge as indispensable extensions of our ability to observe, analyse, and make decisions. In particular, and just as the stethoscope amplifies hearing and the microscope expands vision beyond the visible, both Computer Vision (CV) and spectroscopy tools may enhance and augment human senses in demanding and specialised tasks (see Figure 1.1). Still, when employed individually, their intrinsic limitations may compromise their effectiveness in more complex real-world scenarios. For example, although Computer Vision has established itself as an essential and central technology across multiple domains, enabling systems to interpret, analyse, and extract information from large volumes of visual data in an automated way, it proves insufficient for detecting properties that do not manifest visually. On the other hand, spectroscopic techniques may provide data on atomic and molecular composition in an automated manner, but still face significant challenges related to the speed of operation, sample heterogeneity variation, and the complexity of interpreting the spectra generated, just to name a few.

These limitations clearly highlight the need for integrated approaches that combine the complementary capabilities of different technologies in order to overcome individual constraints and enhance the strengths of each one.

1.2. Shaping an Analytical Synergy for Computer Vision and Laser-Induced Breakdown Spectroscopy

In theory, one may conjecture that the integration of Computer Vision with spectral data presents a unique and interesting opportunity to create a hybrid data environment.

On one hand, Computer Vision replicates our vision and significantly extends it in terms of memory persistence, data throughput, and even machine-speed processing. Classical image processing and, more recently, data-driven models based on deep neural networks enable real-time detection, segmentation, tracking, and classification on modest hardware. Computer Vision captures spatial structure (shape, texture, colour, morphology), context (scene layout, co-occurrence), and temporal cues (motion, change), providing robust macroscopic descriptors of materials and their environment. However, colour and texture often lack clear information related to chemical information, making Computer Vision alone insufficient to support reliable decisions.

On the other hand, spectroscopy tools enable access to the richness of chemical and structural information. In particular, LIBS is a rapid, minimally invasive spectroscopic technique in which a focused laser pulse ablates a small volume of material, forming a micro-plasma whose emission lines encode elemental composition. LIBS offers broad elemental coverage (including light elements), sub-second acquisition, and may be suitable for in-line or in situ analysis for applications such as contaminant detection. Yet, the speed of operation (normally a few Hz) means that the data obtained is spatially sparse and even insufficient.

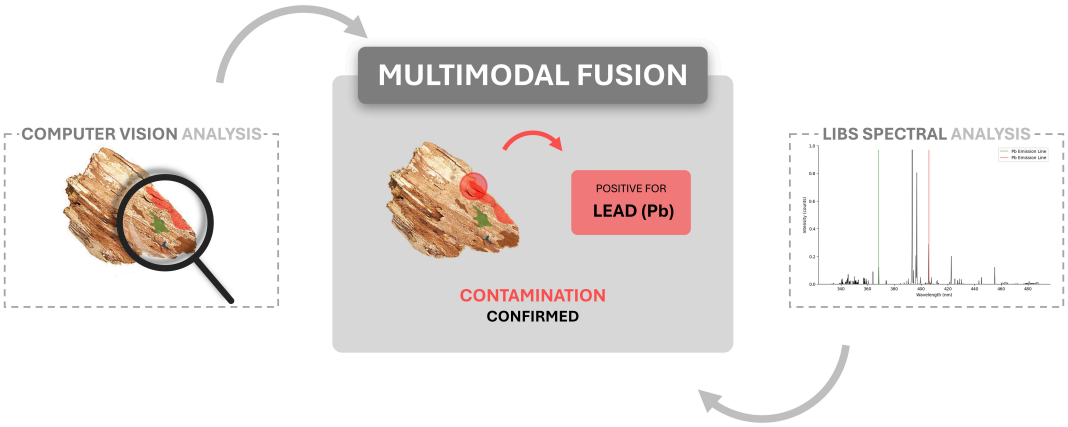


Figure 1.2: Conceptual scheme of multimodal fusion, where computer vision and LIBS spectroscopy are integrated into a unified framework, combining spatial context and compositional evidence for robust contaminant detection.

Thus, it is clear that each of the techniques presents individual strengths but also limiting weaknesses. Yet, as we introduced above, we may now conjecture that we can combine both technologies in a multimodal imaging system to try to minimise their individual limitations (see Figure 1.2). For example, in a possible solution, a Computer Vision module may guide attention to the right spatial regions via segmentation and supply additional context for interpretation, while LIBS may supply definitive compositional evidence that distinguishes visually similar classes. In this context, it becomes possible to explore both the spatial properties and spectral signatures of objects, enabling a more comprehensive, robust, and adaptable analysis system suited to a wide range of scenarios.

This collaborative approach not only increases the effectiveness of material detection and characterisation, broadening the range of potential applications, such as automated industrial environments, environmental monitoring scenarios and selective waste sorting. However, the synergistic fusion of Computer Vision with advanced spectroscopy comes with its own challenges. In summary, we may divide them into the two backbone components of these systems and their open questions:

- **Hardware Challenge**: What design should be used to make a modular solution that supports synchronised acquisition? How to make it robust for compatibility with industrial environments?
- **Software Challenge**: What algorithms and data pipelines are capable of seamless fusion of heterogeneous data streams? How to reconcile differences in temporal

and spatial resolution, and extract meaningful correlations across spatial and chemical domains?

Together, these challenges highlight that the creation of an effective multimodal system is not only a matter of combining sensors, but rather of engineering a coherent framework in which hardware and software are designed to complement one another. Addressing these issues establishes the foundation for the objectives of the present work.

1.3. Dissertation Objectives and Structure

The present dissertation is framed within this technological challenge and builds upon the emerging work developed in recent years by INESC TEC in the field of spectral imaging analysis, with a particular focus on Laser Induced Breadown Spectroscopy. In particular, the broader motivation is to develop a multimodal solution (LIBS and CV) capable of tackling industrial contexts that require significant processing speed and operational robustness, enabling enhanced traceability and quality control that does not hinder the overall performance of the production processes. For this, we establish three operational objectives:

- **Objective 1:** Design and deploy a multimodal hardware solution to incorporate Computer Vision and LIBS in a unique system, in a versatile setup close to industrial applications;
- **Objective 2:** Develop a software layer to integrate two distinct yet complementary data models - visual data from Computer Vision and chemical information obtained through LIBS spectroscopy;
- **Objective 3:** Demonstrate automated detection of contaminants in real-world samples under real industrial throughput conditions.

Moreover, and more on the strategic side, this research also aims to map the opportunities and challenges associated with integrating these two data acquisition modalities. In doing so, it not only aims to advance the state of the art in multimodal analysis, but also charts promising pathways for the development of technologically robust and environmentally responsible solutions, with tangible impact both in applied science and in sustainability-driven industrial sectors.

The present dissertation is structured into seven chapters, organised to reflect the logical progression of the development and integration of the proposed hybrid system, from theoretical and technological contextualisation to practical implementation and analysis of the results obtained.

Chapter 1 introduces the motivation underlying this dissertation, along with the central research problem that guides the investigation. It outlines the general and specific objectives of the study and provides a brief description of the methodologies adopted to achieve them. Finally, it highlights the scientific relevance of the work, as well as the main contributions it aims to deliver.

Chapter 2 offers a conceptual and technological overview of the challenges associated with contaminant detection in solid materials. It examines the limitations of human perception in demanding contexts and justifies the need for advanced technological approaches. Computer Vision and LIBS spectroscopy are presented as complementary techniques, including a discussion of their fundamental principles and the current state of the art, both in isolation and in integrated approaches.

Chapter 3 details the physical and digital infrastructure developed to support the implementation of the hybrid system. It describes the hardware components used, including the LIBS acquisition system, the camera supporting Computer Vision, and control mechanisms, as well as the software tools and techniques employed for data acquisition, processing, and synchronisation. The system architecture is presented with emphasis on its modularity and interoperability.

Chapter 4 describes the two core analytical subsystems, the model based on Computer Vision and the model based on spectral data acquired via LIBS. The first section addresses the selection, training, and validation of the Computer Vision model, with particular attention to dataset creation and annotation, segmentation techniques, and system calibration. The second section focuses on LIBS data processing and interpretation, including pre-processing, feature extraction, and classification strategies.

Chapter 5 presents the approach developed for the synergistic integration of the two models, with the aim of creating a hybrid analysis system. Different fusion strategies, at the levels of decision, features, or acquisition, are discussed, and the adopted architecture is justified. The implementation of the integrated pipeline is described in detail,

highlighting how Computer Vision can guide or constrain LIBS analysis, and vice versa, thus contributing to a more efficient and robust process.

Finally, Chapter 6 outlines the main conclusions drawn throughout the work, summarising the contributions of the dissertation and reflecting on the impact of the results obtained. Possible directions for future research are also discussed, focusing on system generalisation to new contexts, computational optimisation, and the expansion of data sources to be integrated.

1.4. List of works and publications

Journal Publications:

- Diana Capela; Rafael Cavaco; João Carvalho; Pedro A. S. Jorge; Nuno A. Silva; Diana Guimarães. “Spectral Imaging Applications for Industrial Settings: On-Site, At-Line, And In-Sight”. Manuscript in preparation for submission to Resources, Conservation & Recycling (Elsevier)

Oral Presentations:

- Diana Capela; Rafael Cavaco; João Carvalho; Pedro A. S. Jorge; Nuno A. Silva; Diana Guimarães “Laser-induced Breakdown Spectroscopy System for At-line Sorting of Recycled Wood”. 5th International Symposium on Advanced Measurement, Analysis and Control for Energy and Environment (5th AMACEE), Brno, August 2025
- Diana Guimarães, Diana Capela, Rafael Cavaco, João Carvalho, Joana Teixeira, Tomás Lopes, Pedro Jorge, Nuno A. Silva. “Spectral Imaging applications for industrial settings: On-Site, At-Line, and In-Sight”. XLIV Colloquium Spectroscopicum Internationale (CSI 2025), Ulm, Germany, July 2025. Invited Talk

2. Scientific Context and Challenges

The previous chapter introduced the motivation for this research, highlighting both the promise and the limitations of combining Computer Vision with spectral imaging in a multimodal system. While the combination between these two modalities offers a compelling pathway for more robust and versatile analysis, the associated hardware and software challenges also make a universal, “one-size-fits-all” solution for industrial applications unrealistic. Instead, the path forward requires fit-for-purpose approaches, designed and optimised for specific operational scenarios. In this context, this chapter defines our chosen focus - the detection of contaminants - introducing its importance as a critical challenge across industrial, environmental, and societal contexts. The background is framed by first outlining the risks posed by contaminants, followed by examining the shortcomings of conventional analysis, and finally discussing how Computer Vision and LIBS may work as promising complementary technologies. The chapter closes by positioning the fusion of these modalities as a largely unexplored but promising frontier for adaptive multimodal solutions.

2.1. Industrial and Environmental Impact of Contaminants

The rapid population growth, at times uncontrolled during various historical periods especially following the Industrial Revolution, has presented large-scale opportunities but also major critical challenges. Although global growth rates are currently slowing, projections still indicate continued population increase at least until the 2080s [2]. This growth has been sustained by major advances in areas such as public health, disease control, basic sanitation, and food security, which have largely contributed to increased life expectancy and sharply reduced mortality rates, particularly among children.

However, the civilisational advances observed in recent decades do not fully balance the side effects of continuous human expansion [3]. As population density rises and human activities intensify, substantial pressure is placed on natural resources, ecosystems, and productive infrastructure. These challenges manifest across multiple dimensions, from economy to society and environment (see Figure 2.1 ([4], [5])). Among the most visible effects are the significant increase in the generation of industrial and urban waste, the depletion of critical raw materials, and the degradation of water and air quality, to name just a few [4]. These dynamics undermine not only the regenerative capacity of natural systems but also the efficiency and safety of production systems.

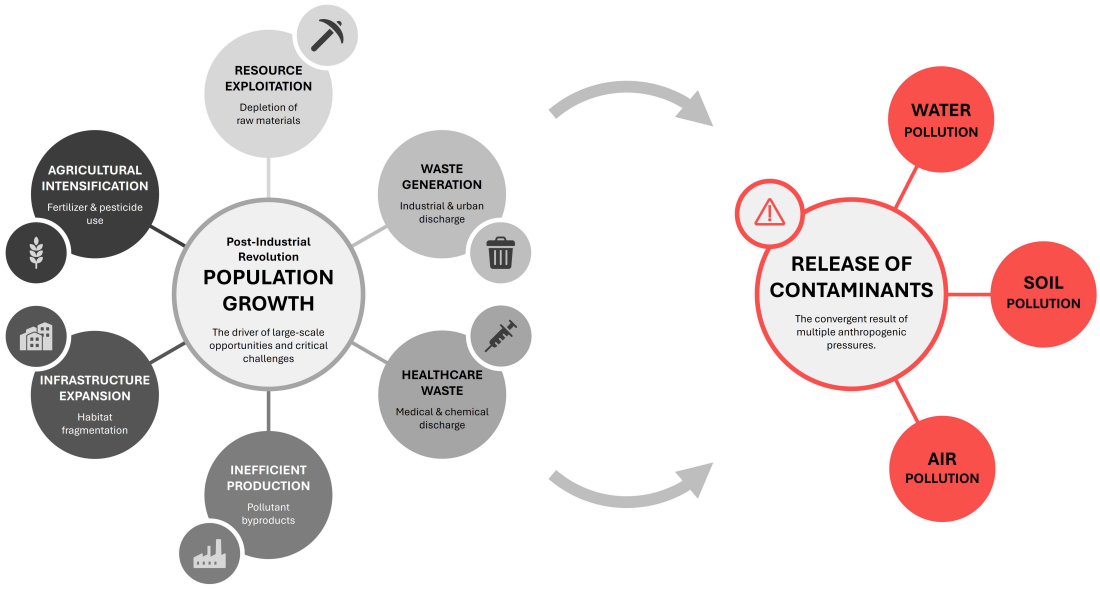


Figure 2.1: Diagram of anthropogenic pressures and contaminant pathways, linking population growth and associated waste streams to the release of contaminants and their consequent impacts on water, soil and air.

Alongside these visible consequences, a less apparent but equally critical issue has gained increasing attention: the production and dispersion of contaminants in various materials and environments. Extensively documented [6], [7], this event affects both industrial products, in which impurities can compromise quality, safety, or performance, and natural raw materials, which were once considered intrinsically pure and reliable. The detection of contaminants in food, groundwater, electronic components, or recyclable materials, for example, illustrates the pervasiveness and severity of this problem, and the consequences of this diffuse presence of contaminants vary depending on the domain in which they occur. For example, in environmental contexts, contaminants dispersed in soil, water, or even air can result in severe ecological imbalances, affecting biodiversity and posing risks to both human and animal health [5]. Another important context is the industrial and production sectors, where cross-contamination of materials can compromise product quality, increase production costs, and challenge compliance with increasingly stringent regulatory standards [8]. This issue becomes particularly critical in the incorporation of recycled materials into production chains, where heterogeneous inputs and variable purity levels increase the risk of impurities being introduced. Ensuring the reliable detection and control of contaminants in such scenarios is therefore not only a matter of operational efficiency, but also of environmental responsibility and regulatory

compliance. It is precisely within this context of controlling the reintegration of recycled materials into industrial processes that this work positions its focus.

2.2. The Barriers of Conventional Analysis in Real-World Scenarios

The selection of the most appropriate approach depends largely on the specific domain of application, as well as on the nature of the contaminants, whether of physical, chemical or biological origin. To name a few examples, conventional solutions include chemical analysis and separation techniques [9], microbiological methods for pathogen detection [10], [11], or a variety of instrumental approaches based on sensors, destructive testing, and visual inspections in industrial settings [12], [13], [14], [15].

Chemical contaminants are generally divided into inorganic and organic impurities, such as pentachlorophenol, benzopyrene, and phthalates, commonly used in wood finishes [1]. In this work, emphasis is given to inorganic chemical elements, as these are the species detectable by atomic analytical techniques such as X-ray Fluorescence (XRF) and LIBS.

Within the group of chemical analysis techniques are included, spectroscopy methods (in its various forms, e.g. mass spectrometry, atomic absorption or emission spectroscopy) are frequently employed to determine, with high precision, the elemental composition and concentration of contaminating substances [16], [17]. Due to their high sensitivity, specificity, and reproducibility, these spectroscopy methods are widely used in laboratory environments. However, in industrial contexts, control routines are more commonly based on systematic sampling and, in many cases, on human visual inspections, which often serve as a first-line screening method. This means that despite the technical advances enabled by these methodologies and their proven effectiveness, particularly in terms of analytical precision and regulatory compliance, they present significant obstacles that hinder their wider application outside controlled laboratory environments (see Figure 2.2).

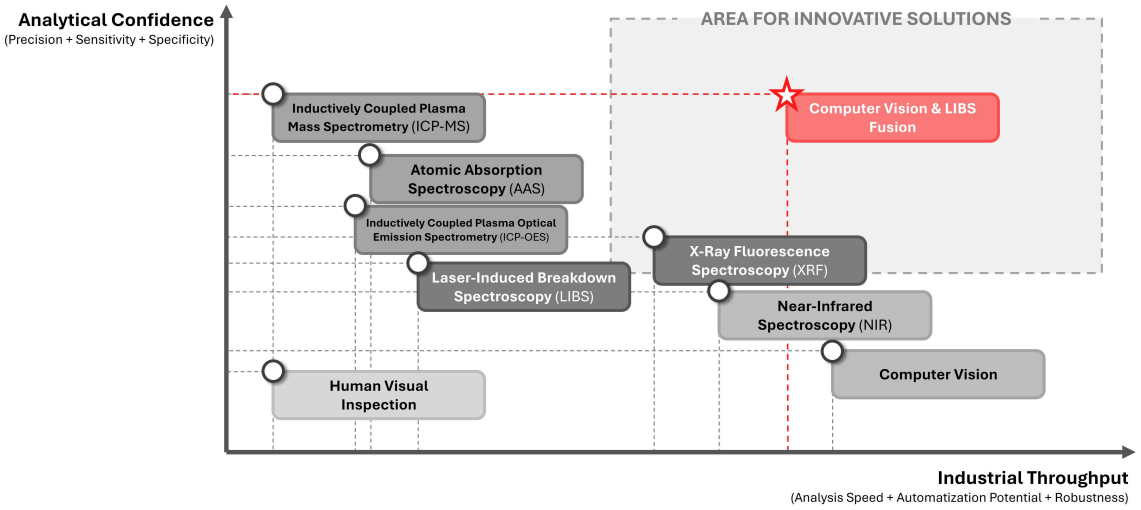


Figure 2.2: Analytical trade-off diagram mapping industrial throughput (speed, automation and robustness) against analytical confidence (precision, sensitivity and specificity), highlighting the design space for innovative solutions and positioning the proposed computer vision and LIBS fusion within that space.

Indeed, among the main constraints are the time-consuming nature of analytical procedures, the requirement for careful and rigorous sample preparation, the high acquisition and maintenance costs of equipment, and the need for highly trained personnel to ensure accurate interpretation of results [18], [19]. In addition to these operational limitations, many conventional techniques rely on point-based analyses conducted on representative samples from a batch or process. While effective in laboratory settings, this approach proves insufficient for detecting localised, sporadic, or heterogeneous contaminations, scenarios that demand a more continuous and dynamic analytical capacity.

Thus, it becomes increasingly important to develop solutions that can bridge analytical spectroscopy tools to industrial performance, by making them faster, minimal destructive, automatic, and suitable for on-line operation, without compromising analytical reliability and sensitivity.

2.3. Two-Stroke Technology: The Discrete Power of Computer Vision and Laser Induced Breadown Spectroscopy

In recent years, and due to distinct contexts as we will describe, two emerging approaches have gained particular relevance: i) Computer Vision, which enables the rapid and precise extraction of information from visual data; and ii) LIBS, recognised for its

ability to provide detailed compositional data at the atomic and molecular levels, in near real-time and without requiring prior sample preparation. Individually, both technologies have become increasingly consolidated across various domains, standing out due to their innovative characteristics and operational advantages compared to conventional techniques.

In the case of Computer Vision, it has established itself as an essential tool in numerous scientific and industrial fields, enabling a broad spectrum of applications across sectors such as medicine [20], [21], [22], agriculture [23], [24], [25], [26], astronomy, autonomous driving [27], [28], [29], [30], waste management, and industrial control [31], [32], [33], among many others. This has been particularly propelled by the convergence of several key developments:

- **Deep learning breakthroughs**: in recent years the advent of machine learning and in particular of convolutional neural networks (CNNs), vision transformers (ViTs), and more recently of hybrid attention-based architectures [34] have drastically improved accuracy and minimised development costs for tasks such as detection, segmentation, and classification, while also broadening applicability to multimodal contexts;
- **Specialised hardware acceleration**: the above breakthroughs cannot be dissociated from the exponential increase in computational power offered by graphics processing units (GPUs), tensor processing units (TPUs), and other dedicated accelerators that have reduced training and inference times. These devices are now crucial for performance applications, by enabling real-time deployment of complex models even on embedded and edge devices;
- **Data-driven methodologies**: also connected with the deep-learning breakthroughs, the availability of massive annotated datasets, together with new techniques for data augmentation, self-supervision, and transfer learning, has enabled models to learn robust and generalisable visual representations [35] far beyond handcrafted features;
- **Accessible and versatile AI tools**: finally, modern frameworks and libraries (such as TensorFlow, PyTorch, and ONNX) have also lowered the barrier to implementation, while cloud infrastructures and edge computing solutions have expanded scalability and deployment options.

Together, these factors have transformed Computer Vision into a mature and versatile technological domain with direct impact in applied sciences and industry (see Figure 2.3 (adapted from [24])). In the medical field, for instance, Computer Vision plays a decisive role in the early and accurate detection of diseases such as cancer, through the automated analysis of medical images, including computed tomography scans (CT), magnetic resonance imaging (MRIs), and X-rays [36]. These techniques directly impact clinical outcomes by increasing diagnostic accuracy and enabling the detection and classification of tumour subtypes, thereby supporting the development of personalised treatment plans [37]. Another example is astronomy, where the growing challenge of processing the massive volumes of data generated by telescopes and space probes, typical of the “big data era”, has driven the adoption of segmentation methods as an essential solution for detailed analysis of the structure, morphology, and evolution of galaxies across different cosmic epochs [38].



Figure 2.3: Example of computer vision deployment in environmental monitoring: aerial imagery used for automated detection and quantification of livestock, illustrating computer vision capacity for large-scale spatial surveillance.

Specifically in the context of contaminant detection, Computer Vision has shown increasing value in remote ecosystem monitoring, water and soil inspection, and automated sorting in recycling and waste management processes [39]. By identifying visual patterns associated with pollution, material degradation, or the presence of surface anomalies, such as visible impurities, colour shifts, or texture changes, this technology can indirectly infer the presence of unwanted substances, acting as an early indicator of potential contamination. However, in this context, its effectiveness is inherently limited to what can be visually observed. Subsurface pollutants, chemically indistinguishable materials, or contaminants with no visible manifestations often remain undetected. In addition, the

reliance on lighting conditions, surface characteristics, and the risk of false positives in visually similar classes highlight the need for complementary analytical techniques.

In the case of LIBS, the technique has likewise emerged as a highly relevant analytical method in recent years, steadily transitioning from laboratory research into real-world technological applications. At its core, LIBS operates by focusing a short, high-energy laser pulse onto the surface of a material. The energy ablates a microscopic portion of the sample and generates a micro-plasma plume. As this plasma cools, the excited atoms and ions emit light at characteristic wavelengths, which are then captured by a spectrometer. Each emission line corresponds to a specific element, enabling the rapid determination of the material's elemental composition [40]. This principle of operation makes LIBS inherently flexible and capable of probing a wide variety of matrices featuring a distinct set of advantages that position it as a versatile and powerful analytical tool:

- **Multi-elemental and composition detection capabilities**: in theory, LIBS can simultaneously detect all elements of the periodic table from a single plasma event, including light elements (H, Li, Be, B, C, N, O) that are often difficult to measure with other techniques such as X-Ray Fluorescence [41], [42], [43]. Besides, being so information-rich, LIBS spectra can be analysed with chemometrics, machine learning, or deep learning models, to support advanced quantitative analysis, sample classification, and automated interpretation [44], [45], [46];
- **Minimal destructiveness**: the ablation involves only micrograms of material, leaving negligible marks and rendering the method quasi-non-destructive, even for sensitive or valuable samples;
- **Rapid single point operation**: emission spectra can be generated in rates typically ranging from 1 Hz to 100 Hz, enabling measurements on sub-second timescales and making LIBS suitable for in-line and in situ monitoring;
- **No or minimal sample preparation**: unlike ICP-MS or AAS, LIBS can analyse solid, liquid, or gaseous samples directly, drastically reducing pre-treatment time and cost, and making it suitable for inline industrial operations;
- **Portability and field readiness**: the miniaturisation of solid-state lasers and spectrometers has led to portable and even handheld devices, expanding LIBS applications to fieldwork, environmental monitoring, and planetary exploration.

Together, these features explain the growing interest in LIBS as a disruptive analytical technology, now enabling application across an expanding array of contexts, from the metallurgical and automotive industries [47], to medicine and clinical diagnostics [48], [49], [50], environmental analysis and remediation [51], [52], [53], archaeometry and cultural heritage preservation [54], [55], food safety [56], [57], [58], and even the exploration of submerged or marine environments [59], [60]. For example, in the field of heritage conservation, there is a need for high-value methods and strategies that provide crucial information to guide preservation efforts. LIBS has been employed in stratigraphic analyses of ancient paintings, metals, and ceramics, allowing the identification of the elemental composition of sensitive surfaces without compromising the structural integrity of valuable artefacts. Another interesting example is planetary exploration, where LIBS has become increasingly prominent (see Figure 2.4 (adapted from [61], [62])). In addition to the many tools and sensors typically included in such missions, this robust technique, capable of withstanding adverse and extreme environmental conditions, and requiring no logistical infrastructure for sample collection or transport, has enabled detailed geochemical analyses directly on the surface of Mars, as demonstrated by instruments such as ChemCam and SuperCam on planetary rovers [63], [64].

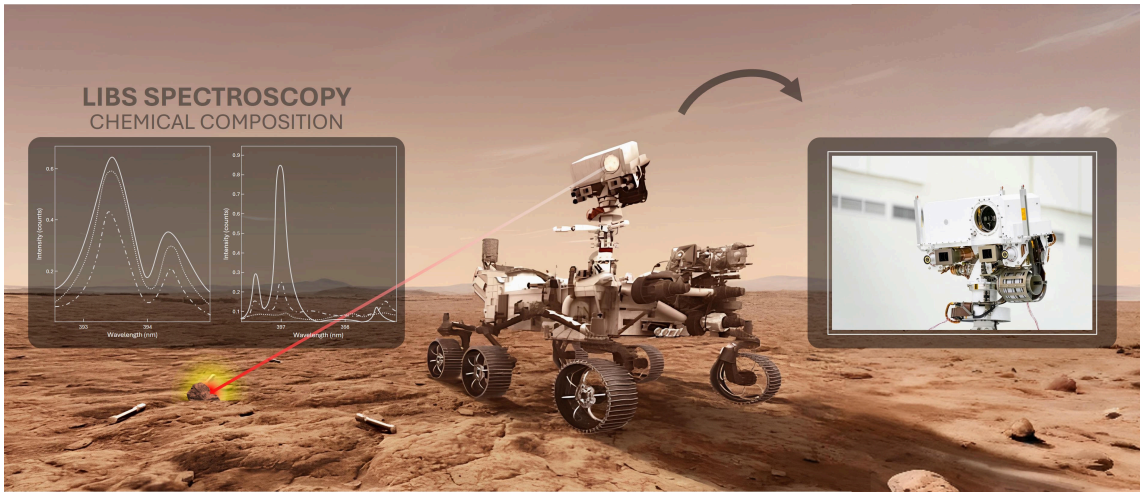


Figure 2.4: Field and planetary deployment of LIBS: rover-mounted LIBS analysis (Perseverance/SuperCam) demonstrating in-situ elemental interrogation of geological targets and the technique's operational suitability for remote compositional analysis.

More specifically regarding contaminant detection, LIBS continues to demonstrate its considerable potential, particularly through its ability to directly identify toxic chemical elements, heavy metals, and other undesirable substances [65], [66]. By analysing the emission spectra generated during the laser-sample interaction, it is possible to detect

characteristic spectral peaks for each element, allowing for accurate inference of their presence and, consequently, the assessment of contamination levels. However, the technique is not without limitations. The inherently sparse sampling - as each laser pulse interrogates only a microscopic spot - may limit the representativeness in heterogeneous or large-scale materials, where contaminants can be unevenly distributed. Issues such as the intrinsic surface-limited nature of the technique, which prevents probing the bulk of the sample, as well as variable surface roughness, particle size, and environmental factors (e.g., humidity, dust, ambient light) can also degrade spectral quality and consistency.

2.4. Fusing Technology to Advance the State of the Art

Despite the significant independent advancements recorded in both the field of Computer Vision and that of LIBS, integrated solutions that synergistically explore these two technologies remain largely underexplored. The vast majority of existing approaches focus on unimodal systems or, when multimodal, rely on alternative combinations that do not necessarily involve LIBS. This situation is particularly, and without exception, evident in the context of contaminant detection, where the potential of combining detailed visual characterisation with direct compositional analysis remains underutilised (see Figure 2.5).

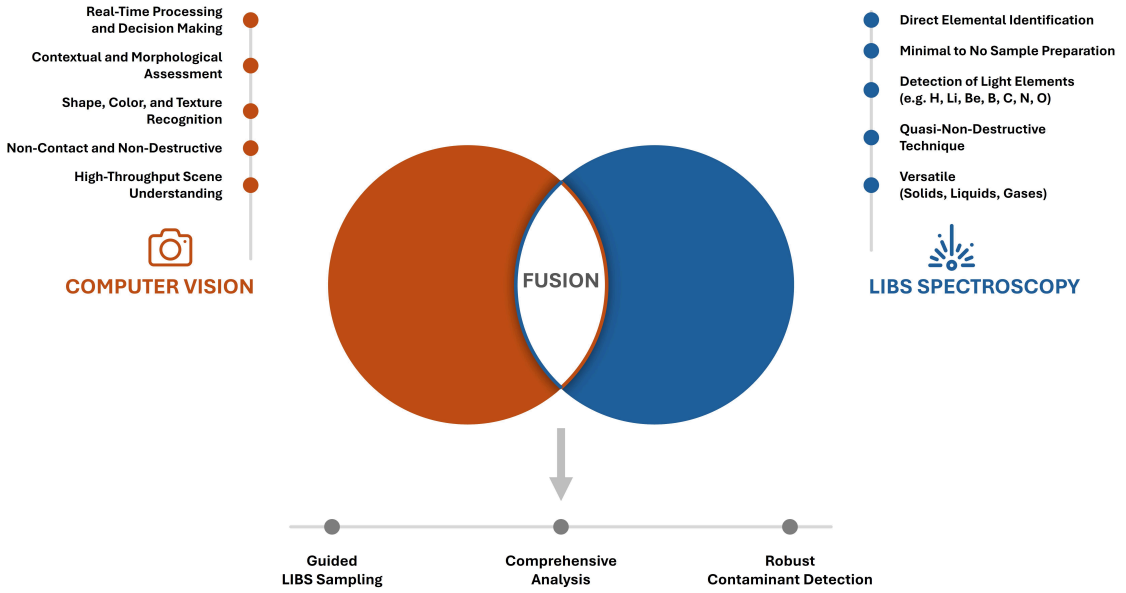


Figure 2.5: Venn diagram of complementary strengths of computer vision and LIBS spectroscopy. The intersection (fusion) denotes spatially guided LIBS sampling, combined spatial-compositional analysis, and improved robustness in contaminant detection and decision-making.

A few studies have begun to explore, at an early stage, the convergence between LIBS and Computer Vision, though with objectives distinct from those proposed here. For example, [67] developed a method in which LIBS-generated data on silicone insulators were converted into images and processed with the aim of assessing the degree of surface pollution. Although a form of technological fusion is present, the focus of the analysis is not on the direct detection of chemical contaminants, but rather on the extraction of visual patterns derived from the obtained spectra. In another example, within the food sector, [68] designed a continuous contaminant detection system for processed products based on X-ray imaging, incorporating advanced Computer Vision techniques.

These examples clearly illustrate that, although there are promising attempts at fusing visual and analytical modalities, the specific combination of LIBS and Computer Vision applied to contaminant detection, particularly in dynamic, heterogeneous, and continuous-flow environments, remains an unexplored frontier with high potential impact. Accordingly, the present work positions itself as an attempt to address this gap by proposing an innovative approach that effectively integrates the analytical power of LIBS with the interpretative capabilities of Computer Vision. The objective is not to advance the technical development of each technology independently, but rather to drive forward the state of the art in their joint and strategic application, demonstrating how such fusion can deliver more robust, versatile, and informed solutions for contaminant detection in complex scenarios.

2.5. Concluding Remarks

The contextual and technological exploration developed throughout this chapter highlights the multidimensional challenges associated with the often-invisible presence of contaminants in environmental and industrial contexts. Although conventional analytical methodologies have played a crucial role in the detection and control of contaminants, their operational limitations, particularly outside controlled laboratory settings, have increasingly exposed the need for alternative approaches that are faster, minimally destructive, automated, and adaptable to the real dynamics of productive and ecological systems.

Within this framework, Computer Vision and LIBS emerge as two complementary technologies, each addressing different dimensions of the analytical problem. While both technologies have demonstrated maturity and applicability across various domains, their joint application remains notably underexplored. Rather than pursuing unimodal

advancements, this work aims to design a hybrid architecture capable of coherently and functionally articulating the capabilities of both modalities.

3. Hardware Architecture and Acquisition Systems

Despite the significant progress achieved independently in the fields of Computer Vision and LIBS, it becomes essential to identify both hardware and software pathways to deploy a solution for a multimodal on-line contaminant screening. This section focuses on the hardware part, describing the technical realisation of the proposed integrated approach, based on the combination of optical, electronic, and mechanical components. Each module was carefully selected not only based on its specific functional requirements but also with the intent of ensuring robust interoperability, capable of supporting a continuous and efficient operational flow from the first to the final subsystem.

3.1. Functional Integration for Multimodal Inspection

Our target system is an adaptive multimodal inspection solution designed for semi-industrial environments, capable of detecting, segmenting, and analysing heterogeneous samples, such as rocks, glass, or wood, in an automated, localised, and non-destructive manner. Owing to the complementarity between Computer Vision and LIBS, the goal is to deliver a comprehensive analysis within a few seconds, with the potential to be scaled and integrated into continuous production lines or advanced sorting contexts.
Towards this global vision, the system was structured into three distinct functional modules, reflecting the division between the employed technologies and the infrastructure that connects them:

- **A Computer Vision module**: positioned upstream, this first module relies on Computer Vision equipment and processing techniques to perform image acquisition and applies identification, segmentation, and classification algorithms with the goal of identifying regions of interest that merit further investigation. These selected regions can be either fused with spectral data from the second module or serve as coordinates for targeted LIBS sampling;
- **A LIBS module**: the second module, mounted on a robotic support, performs precise localised spectral analyses through the emission of laser pulses and the collection of radiation emitted by the induced plasma. It provides detailed compositional information about the samples or regions identified by the vision subsystem;

- **A motion and spatial scanning module**: the third module is responsible for the positional coordination of the LIBS system relative to the conveyor belt and the samples under inspection. It ensures accurate spatial coverage, fine focal adjustment, and seamless integration between continuous material flow and targeted local analysis, thereby guaranteeing the synchronisation and reliability of the entire inspection workflow.

The three modules, which will be further described in the subsequent sections, operate not only sequentially, but also in a tightly coordinated manner, ensuring the overall fluency of the analytical process (see Figure 3.1). Following the analysis, the system outputs a reliable inference regarding the presence of one or more specific contaminants, grounded in the complementarity between visual, compositional, and positional information.

Besides, the samples of varying typologies and degrees of contamination move along a linear trajectory on a motorised conveyor belt, which approximates the context of an industrial application. Indeed, the whole combination not only enables the system to be scaled across different application contexts, due to its modular architecture, but also allows for the adaptation of detection criteria according to the material, the environment, and the production requirements.

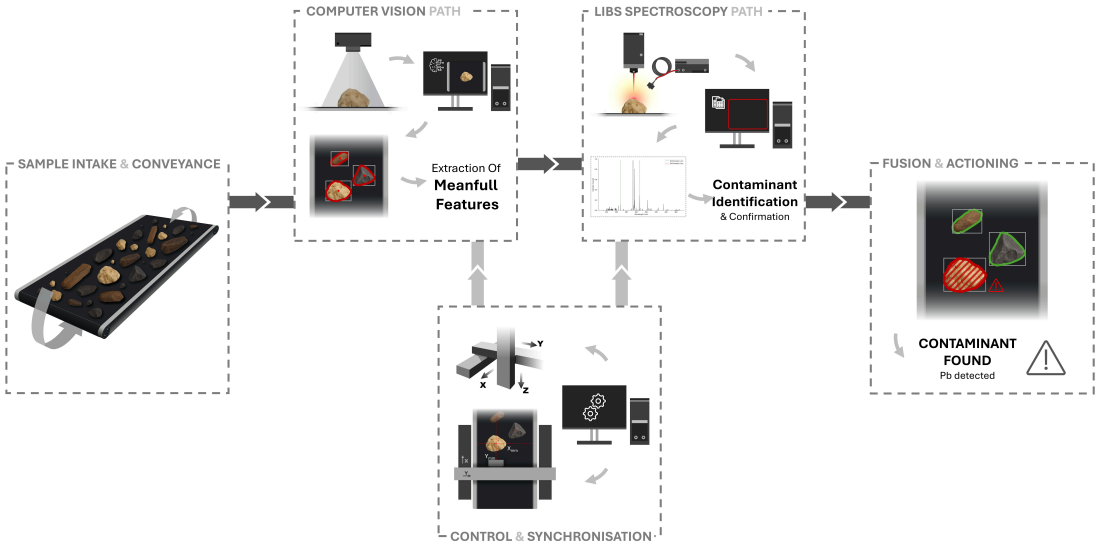


Figure 3.1: High-level functional architecture of the integrated multimodal inspection system, presenting sample intake and conveyance, the computer vision and LIBS paths, control and synchronisation, and the downstream fusion and actuation stages to provide an intuitive overview of data and control flow.

3.2. Computer Vision Module for Automated Analysis

The analytical flexibility extends, first and foremost, to the Computer Vision subsystem, which is positioned upstream (see Figure 3.2) in the process and plays a pivotal role in guiding the system's overall behaviour. By enabling the customisation of both the differential detection model and the decisions derived from it (such as which regions should be analysed or disregarded) this module serves as the operational starting point of the workflow, functioning as an intelligent filter that anticipates, organises, and rationalises the subsequent analytical efforts, thereby optimising the time and resources associated with the spectral stage.

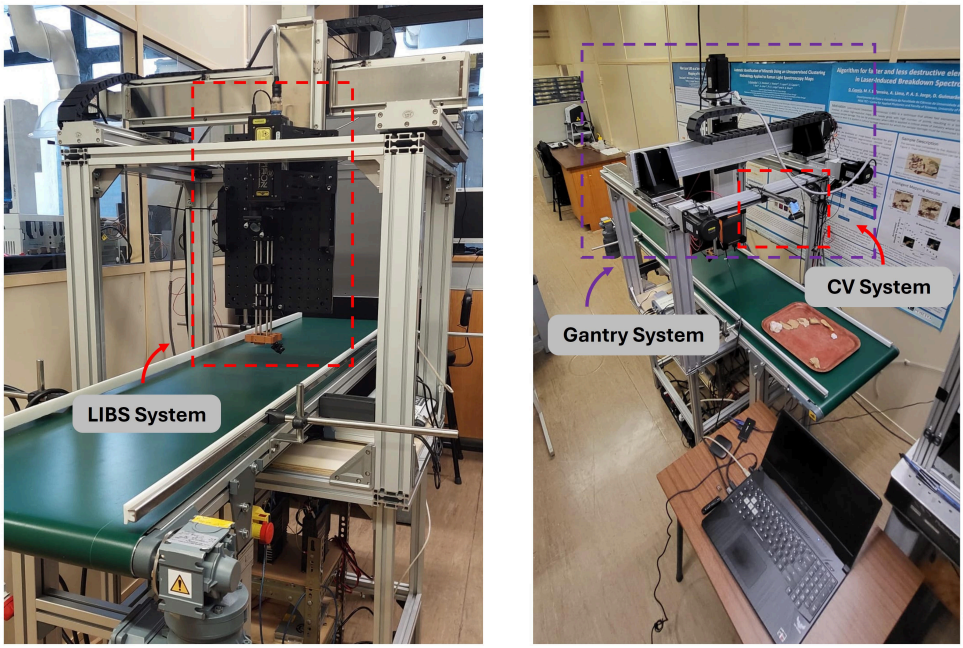


Figure 3.2: Rear view of the system, the LIBS module, responsible for chemical analysis (left), while the system's entry point (right) features the upstream Computer Vision subsystem for preliminary visual inspection.

In order for this visual decision-making process to be effective and consistent with the operational dynamics of the system, the selection of the image acquisition module required a thorough evaluation of multiple parameters, both operational and environment-based. The chosen camera had to meet a demanding set of conditions imposed by the physical constraints of the application, including the constant motion and morphological diversity of the samples conveyed on the belt, the variability of ambient lighting within the inspection environment, and, most importantly, the need to ensure stable, well-defined images with high spatial and chromatic fidelity. For this reason, one of the key design constraints imposed was that the projected pixel dimension in the object plane should

approximate the size of the laser spot produced during LIBS acquisition. This condition ensured that each visual unit could be meaningfully correlated with the physical region subjected to spectroscopy analysis, thereby maximising the consistency between the two modalities. In parallel, operational factors such as the conveyor speed, fixed at 17.5 cm/s, also played a critical role in defining the requirements for acquisition rate and data throughput. The camera therefore needed to sustain a frame rate high enough to capture non-blurred images with sufficient overlap between successive frames, while balancing resolution and memory usage.

To this end, and with the aim of ensuring precision in the detection, segmentation, and classification processes, a model was selected that offered a robust balance between resolution, acquisition speed, and sensor quality. The Basler ace 2 camera, equipped with the Sony IMX900 CMOS sensor, was found to meet these technical requirements, offering a resolution of 3.2 megapixels, a frame rate of up to 114 frames per second, and a 5 GigE interface capable of handling the high data throughput generated in real time.

In addition, one of the decisive factors in its selection was the integration of global-shutter technology, which enables all sensor pixels to be exposed simultaneously, an essential condition for preserving object geometry in motion and avoiding the typical distortions associated with rolling-shutter sensors [69] (see Figure 3.3). Furthermore, the fact that it is a colour camera introduces an additional chromatic dimension to the analysis, enhancing the system's discriminative capabilities without compromising the integrity and consistency of the acquired images. Collectively, these features contribute to continuous and stable image acquisition, fully synchronised with the process rhythm, thereby ensuring ideal conditions for the subsequent spectral analysis phase.

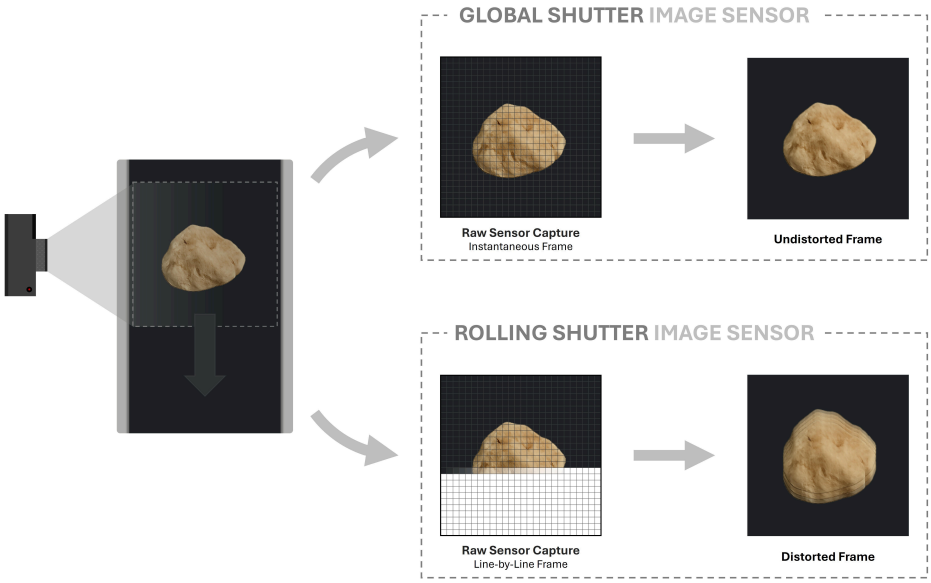


Figure 3.3: Comparative illustration of global-shutter and rolling-shutter acquisition under motion, emphasising motion-induced geometric distortion in rolling-shutter images and the undistorted capture achieved with a global-shutter in dynamic inspection contexts.

3.3. LIBS Module for Precision Elemental Analysis

The second module concerns a LIBS equipment and directly benefits from the information provided by the Computer Vision subsystem. This is particularly evident for the precise localisation of the regions of interest, which correspond to entire samples that have been previously segmented and classified, enabling the chemical analysis of the constituent elements of the selected samples to be conducted in a targeted, selective, and efficient manner.

To enable such performance, it is essential that the spectroscopic component operates in coordination with the rest of the system, both spatially and temporally. The LIBS module is based on an architecture composed of three main components, the laser system, the spectrometer, and the digital delay generator (DDG), which is responsible for the sequential and temporal control of events (see Figure 3.4). Indeed, the integration of the LIBS components is not limited to simply triggering a laser pulse, but it is rather designed to precisely orchestrate the timing of the material ablation and the capture of the optical signal emitted by the resulting plasma, thereby maximising the quality of the data acquired.

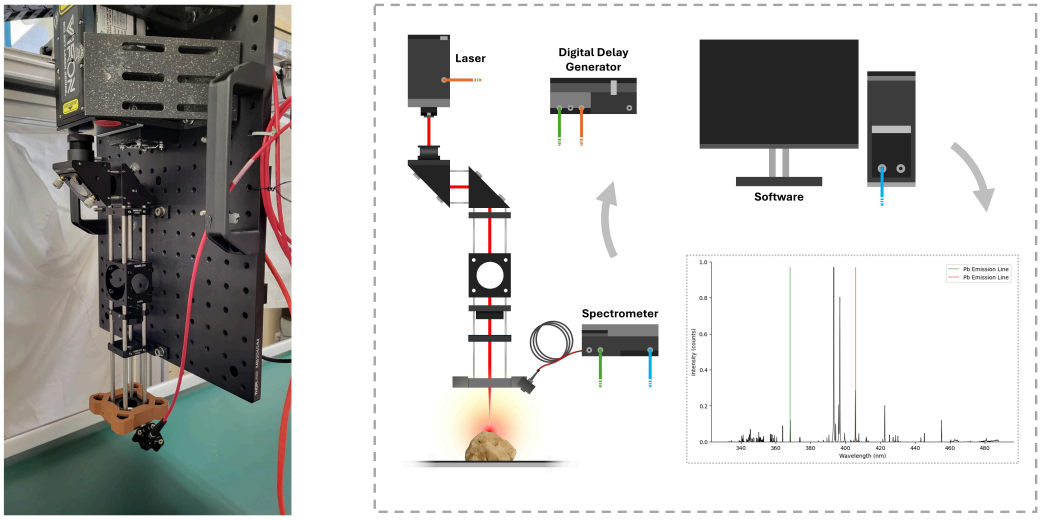


Figure 3.4: LIBS subsystem: photograph of the assembled instrument paired with a technical block diagram of laser, spectrometer and DDG, including a representative emission spectrum with annotated elemental lines.

Overall, the design of the LIBS module was the result of a technical process conducted with the support and expertise of the Center for Applied Photonics at INESC TEC, whose extensive experience with LIBS systems guided the definition of selection criteria and experimental validation. Drawing on this accumulated know-how, not only were the operational parameters for the laser and spectrometer specified, but also the need to ensure synchronisation between them, ensuring a robust response to the diverse demands of the application.

This specification process, coupled with the aim of ensuring the system's applicability to a wide range of contaminants and sample types, without compromising analytical speed or reliability, led to the selection of the VIRON laser model from Lumibird. This is a high-efficiency diode-pumped Nd:YAG laser, operating at a fundamental wavelength of 1064 nm, specifically designed for demanding applications that require operational robustness. This active Q-Switch laser provides firing rates of up to 20 Hz, corresponding to 20 pulses per second, with individual pulse of typical duration of 10 ns and energies up to 50 mJ, which are well-suited to effective plasma generation across various sample types. It is also worth highlighting its operational versatility, evident in its three available modes, single shot, continuous shot (multiple shot), and burst mode, which allows controlled emission of a pre-defined number of pulses. This flexibility grants the system significant control, not only during analysis but also in adjusting the response according to the specific characteristics of the samples.

The physical-chemical properties of the samples are detected via the electromagnetic spectrum that is emitted by the plasma during ablation. This information is captured by a multi-channel spectrometer integrated into the system. We have chosen a wide spectral range, from 200 to 1100 nm, enabling simultaneous detection of multiple emission lines associated with different chemical elements present in the samples. This performance is achieved without compromising high reading speed or dynamic response, both critical characteristics for this application, thanks to the use of a CMOS linear sensor, which offers technological advantages over traditional CCD sensors, particularly in high acquisition-rate contexts. Finally, the synchronisation between laser firing and optical signal acquisition by the spectrometer is ensured by a digital delay generator, a programmable control module that determines the activation timings of each component. This device emits triggering pulses with nanosecond-level resolution, enabling precise adjustment of the spectrometer's acquisition start time. This temporal coordination is essential, as it ensures that the detection window aligns with the optimal moment after the pulse, specifically, when the plasma begins to cool down and reaches its peak optical emission, thus optimising the signal-to-noise ratio and minimising undesirable background interference.

3.4. Motion Module for Localised Chemical Analysis

Beyond the necessary temporal synchronisation among the various components of the system, it is equally essential to ensure precise spatial coordination, particularly regarding the position of the LIBS module in relation to the samples under analysis. In this context, the ability to move the inspection system swiftly, accurately, and with high repeatability proves crucial to ensure adequate spatial coverage and the overall effectiveness of the spectral analysis process.

To this end, a Robotic System (Gantry) was implemented based on a FUYU three-dimensional motion system, FSL120 linear module, operated by an FMC4030 pulse-type controller. This architecture enables independent movements along the X, Y, and Z axes, as well as interpolated trajectories, either linear or circular, involving two or three axes simultaneously see Figure 3.5). With stroke ranges configurable between 50 mm and 1500 mm, the system offers sufficient flexibility to accommodate both small and large-scale samples. Its positioning accuracy of 0.02 mm ensures the fine spatial control required for LIBS acquisition, while a maximum speed of 500 mm/s allows the gantry to keep pace with the conveyor dynamics and the high-frequency demands of sequential laser firing. Such capability not only ensures smooth and precise navigation across the surface of the

samples but also guarantees the positional stability required at the LIBS system's focal region, an essential factor for maintaining consistency across spectral measurements. Additionally, vertical movement along the Z-axis allows for fine adjustments to the focal distance, adapting the system to variations in sample thickness and surface topography without compromising optical performance or ablation quality.

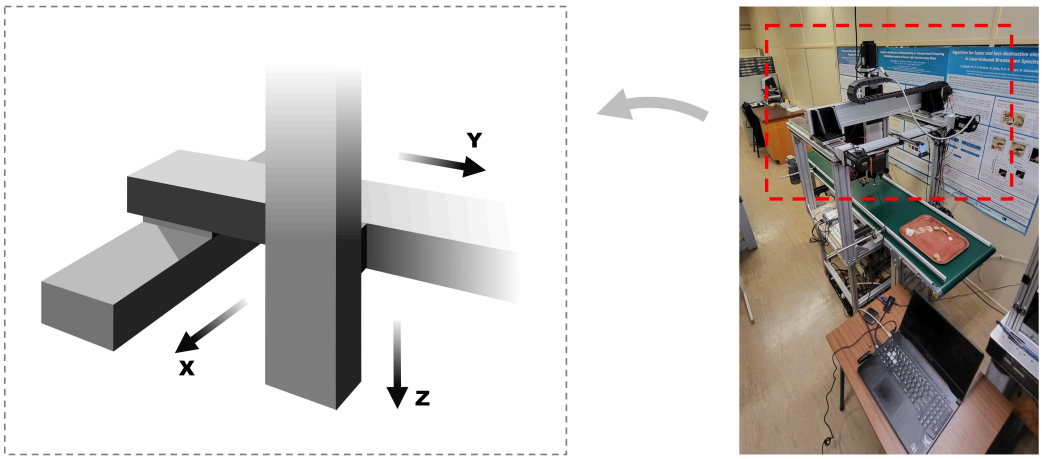


Figure 3.5: Gantry motion geometry and implementation: left, 3D schematic showing X-Y-Z axes and principal degrees of freedom; right, photographic view of the physical gantry and its workspace.

It is important to highlight that although this robotic system constitutes the primary mechanism for moving the LIBS module across the sample plane, its operation is carried out in coordination with a motorised conveyor belt, which is responsible for the continuous transport of materials through the inspection zone. This integrated configuration ultimately provides the system with a high degree of adaptability to industrial contexts, allowing for its installation on pre-existing production lines with minimal structural adjustments or disruption to operational workflows.

3.5. Module Integration Regarding Temporal Coordination and Spatial Alignment

The integration among the different modules, from the Computer Vision subsystem to the spectrometer, requires strict sequential control, both spatially and temporally, to ensure that the LIBS analysis is accurately targeted at regions previously identified as analytically relevant (see Figure 3.6). To achieve this level of precision and operational robustness,

it was necessary to implement a set of calibrations and specific adjustments tailored to the operational characteristics of each subsystem.



Figure 3.6: Temporal sequence diagram showing the ordered progression from computer vision assessment to LIBS spectral interrogation, culminating in integrated decision-making and actuation.

For the Computer Vision module, intrinsic camera calibrations were performed, correcting radial and pincushion optical distortions using specialised calibration patterns. Through the incorporation of ArUco markers, it was possible to accurately define the boundaries of the operational area, as well as to compute conversion factors between image units (pixels) and physical coordinates (millimetres), thereby establishing a reliable metric link with the Gantry positioning system (see Figure 3.7). Based on this data, spatial calibration lines were traced, enabling accurate transposition between the two systems.

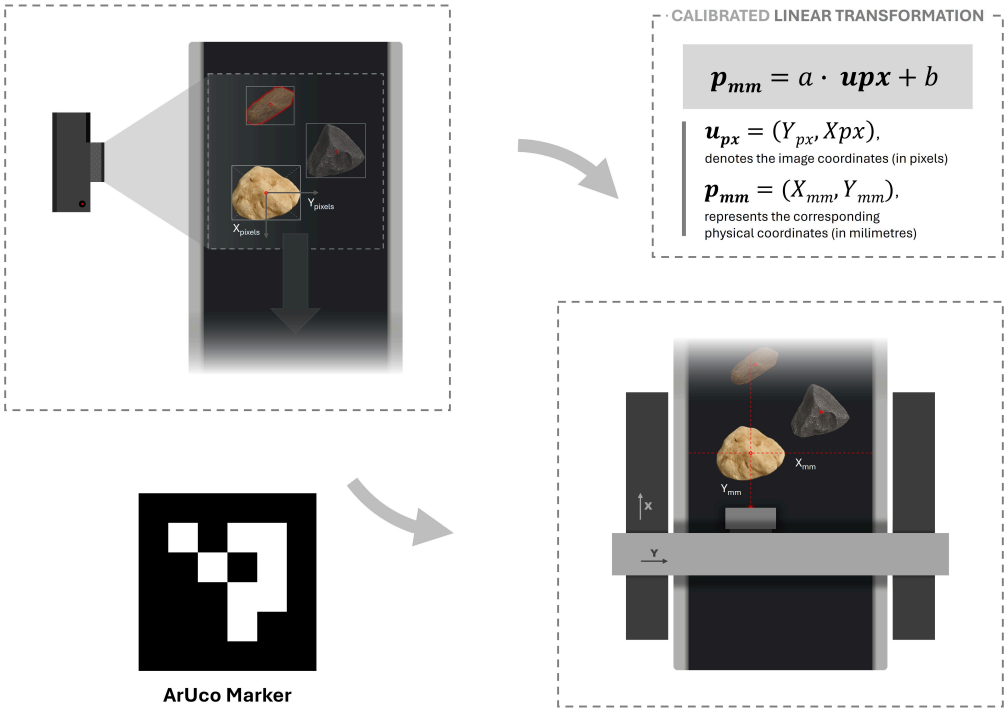


Figure 3.7: Calibration and synchronisation workflow linking camera pixel coordinates to gantry millimetre coordinates via ArUco markers and transformation matrices, establishing metric alignment for spatially guided LIBS sampling.

In parallel, the actual speed of the motorised conveyor belt was determined experimentally to be 17.5 cm/s and subsequently correlated with various acquisition time intervals relevant to the global synchronisation of events. In pratical terms, this corresponds to a transit time of approximately 1.75 s between the moment a sample leaves the field of view of the vision subsystem and its arrival at the LIBS interaction zone, thereby defining the effective temporal window available for detection, decision-making and laser positioning. Beyond this global synchronisation aspect, the same parameter also determines the spatial separation between consecutive laser shots during continuous operation modes. For example, at a repetition rate o 20 Hz, the conveyor speed corresponds to a displacement of roughly 8.75 mm between successive pulses.

3.6. Concluding Remarks

In this chapter we detailed the hardware architecture of our multimodal inspection system, purpose-built to address the complex challenges of automated contaminant screening in on-line industrial operation. The description of the system's three functional blocks, visual acquisition, compositional analysis, and robotic movement, has demonstrated that the

system’s effectiveness stems not only from the technical sophistication of its individual components, but above all from the way these components were designed to operate in a coordinated, precise, and interdependent manner.

The careful selection of optical acquisition modules, the robust implementation of the LIBS infrastructure, and the integration of a high-precision robotic platform, both spatially and temporally, converge into a technological solution that goes far beyond the mere juxtaposition of equipment. Rather, it represents the concrete realisation of an integrated approach, in which Computer Vision acts as an intelligent screening mechanism, selectively guiding the LIBS analysis, which in turn yields highly resolved compositional inferences on the identified regions of interest. Thus, the hardware described herein constitutes not only the physical foundation of the system, but also its technological enabler, upon which the subsequent layers of algorithmic intelligence and adaptive inference will be built.

4. Software Layer and Algorithms

The utility of the sophisticated hardware architecture only truly manifests when the acquired data are transformed into actionable and meaningful knowledge. Indeed, beyond the robustness of the physical components, the efficacy of the intended solution will reside in its ability to extract, structure, and interpret significant information from the collected signals of the analysed samples. This means that a functional software layer is also required to convert physical and chemical observations into well-founded analytical decisions and to allow autonomous actions. For this, the system should integrate multiple sources of information - high-resolution images and localised LIBS spectra - each with its own structure, granularity, and information content.

Considering this diversity, the solution proposed comprises two core specialised software models, each tailored to the specificities of the respective data type and oriented towards distinct analytical objectives. On one hand, we will have a Computer Vision software model that will be responsible for interpreting the visual content of the samples, outputting information such as sample location and classification. On the other hand, the spectral data software model will aim to utilise the LIBS technique to provide a compositional analysis and in particular, the detection of key contaminants.

In this chapter, we present the implementation of each software model, describing how they operate according to a specific technical pipeline, with distinct strategies for feature extraction, decision-making, and temporal application across different phases.

4.1. Computer Vision Model Architecture and Strategy

As previously indicated, the goal of the Computer Vision system is to provide an initial intelligent screening stage, capable of visually locating and interpreting the samples at the pipeline’s input. This visual component was conceived not merely as a passive detection mechanism but as a true “analytical trigger” enabling the system to perceive the type, quantity, and individuality of each sample in circulation, thereby defining how subsequent data should be acquired and interpreted. To systematise, it is essential that the data-processing model is suitable for the dynamic and demanding characteristics of industrial environments, namely:

1. automatically identify and spatially locate the samples by segmenting them on the conveyor belt;
2. classify, and distinguish different samples based on morphological, textural, and chromatic attributes;
3. deliver these results in real time.

Considering these requisites, our choice was to seek solutions based on the segmentation paradigm, a choice that directly influences not only the system's final performance but also the data annotation processes, training strategies, and the manner in which knowledge is subsequently operationalised.

Focusing on segmentation models, and reviewing the most established approaches, semantic segmentation stands out as the most common [70]. In semantic segmentation [71], each area of an image is assigned a specific class (such as wood, glass, or cork) which allows a rapid and coarse understanding of the scene's composition (see Figure 4.1). However, this approach does not distinguish between different instances of the same category, which constitutes a significant limitation when multiple fragments of the same class must be treated as independent entities.

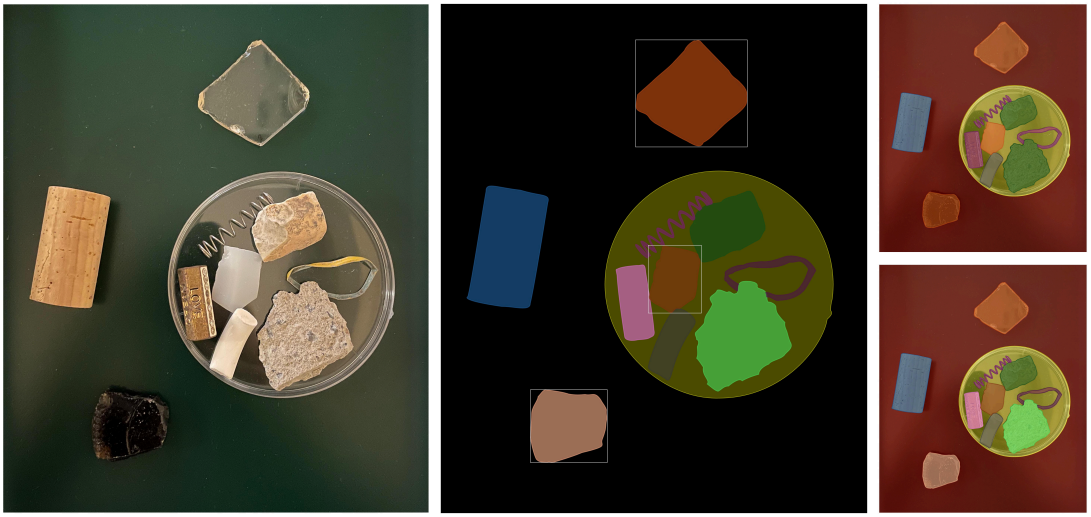


Figure 4.1: Comparative visualisation of segmentation paradigms (left: original image; centre: instance segmentation), highlighting instance segmentation as the selected strategy for its capacity to individually delineate and trace samples across the pipeline. right: semantic segmentation (top) and panoptic segmentation (bottom) for comparison.

To overcome this limitation, instance segmentation [72] offers an additional level of granularity by identifying and separating distinct objects, even when belonging to the same class. This feature renders it particularly suitable for contexts such as the present one, where the individualisation of samples is crucial to ensure traceability and integrity throughout the pipeline. Finally, panoptic segmentation [70] is a more recent approach that aims to simultaneously integrate the granularity of instance segmentation with the contextual coverage of semantic segmentation, providing a holistic representation of the image. Yet, in spite its conceptual merits, this approach presents considerable computational complexity. All things considered, we opted to utilise instance segmentation, which is well suited to the demands of the present project, establishing a promising balance between the required level of detail for sample individualisation and the technical and operational feasibility of real-time application.

Adopting this strategy imposed specific requirements in model development, particularly regarding dataset curation and annotation. Given the fundamental premise of distinguishing each sample individually, it became imperative to compile a robust dataset representative of the diversity and complexity observed under real operational conditions. To this end, controlled image acquisition was performed in scenarios closely resembling the target industrial environment, ensuring the inclusion of diverse fragments of rock, glass, and wood exhibiting significant variations in shape, colour, texture, and size. The dataset was constructed with high intra- and inter-class variability to foster the model's generalisation capacity and resilience to specific cases.

As a direct consequence of this morphological diversity, dataset annotation had to follow rigorous instance segmentation criteria, requiring the creation of individualised masks for each visible element, an inherently more demanding task compared to traditional approaches. To mitigate this complexity and enhance efficiency, the Roboflow platform was employed [73], whose assisted segmentation capabilities, based on pre-trained models, accelerated the initial annotation phase while ensuring minimum quality standards (see Figure 4.2). Nevertheless, in cases of overlapping objects or diffuse contours, manual intervention was necessary, proceeding with point-by-point annotation. Although more time-consuming, this stage was distributed among multiple team members, contributing not only to enhanced quality control but also to mitigating potential risks of subjective bias in visual interpretation.

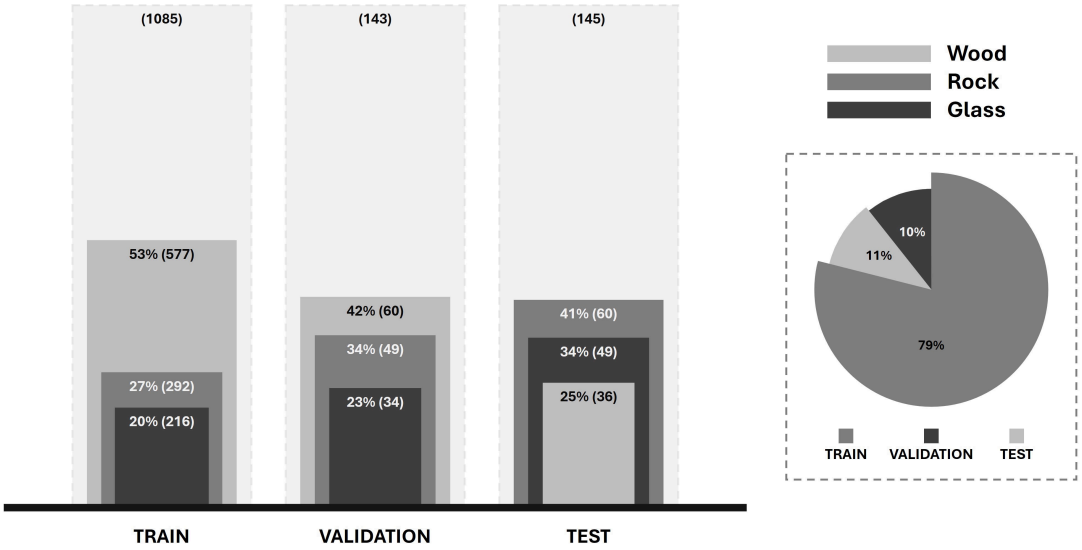


Figure 4.2: Dataset composition and partitioning: absolute sample counts per split and class-wise percentage distribution used for training, validation and testing. (numbers in parentheses denote annotation/mask counts)

Following curation and annotation, the dataset was augmented systematically to enhance visual diversity without compromising sample integrity. Operations such as random rotations, brightness and exposure adjustments, noise introduction, and selective cropping were applied, among others. Fundamentally, these augmentations expanded the range of simulated conditions, from lighting variations to spatial orientations, providing the model with increased robustness against natural fluctuations typical of real application contexts.

Considering this level of visual variability and the demand for generalisation, it was essential to adopt a segmentation architecture capable of combining performance, flexibility, and accuracy. Consequently, the choice fell on the eleventh version of the You Only Look Once (YOLO) architecture [74], [75], widely recognised for its high capability in real-time object detection and segmentation. This architecture has been successfully applied across multiple domains, from robotics to surveillance and biomedical analysis, standing out for its versatility, scalability, and efficiency [76], [77], [78]. Beyond object detection, the model supports additional functionalities such as tracking and even pose estimation, further reinforcing its suitability for scenarios with stringent operational constraints.

Structurally, the network is based on a unified architecture that processes the input image integrally and simultaneously, resulting in a highly efficient pipeline. This is achieved through a deep convolutional backbone that extracts hierarchical multiscale

feature maps. These semantically rich representations are utilised by specialised heads that generate spatial localisation of objects, their classification. In instance segmentation tasks, this can be performed either through bounding boxes, which provide a coarse but useful approximation of object extent, or through custom mask generation, which yields more precise and representative delineations of the samples. While bounding boxes do not fully capture the actual shape or composition of the objects, they can still serve as a valuable basis for extracting complementary information.

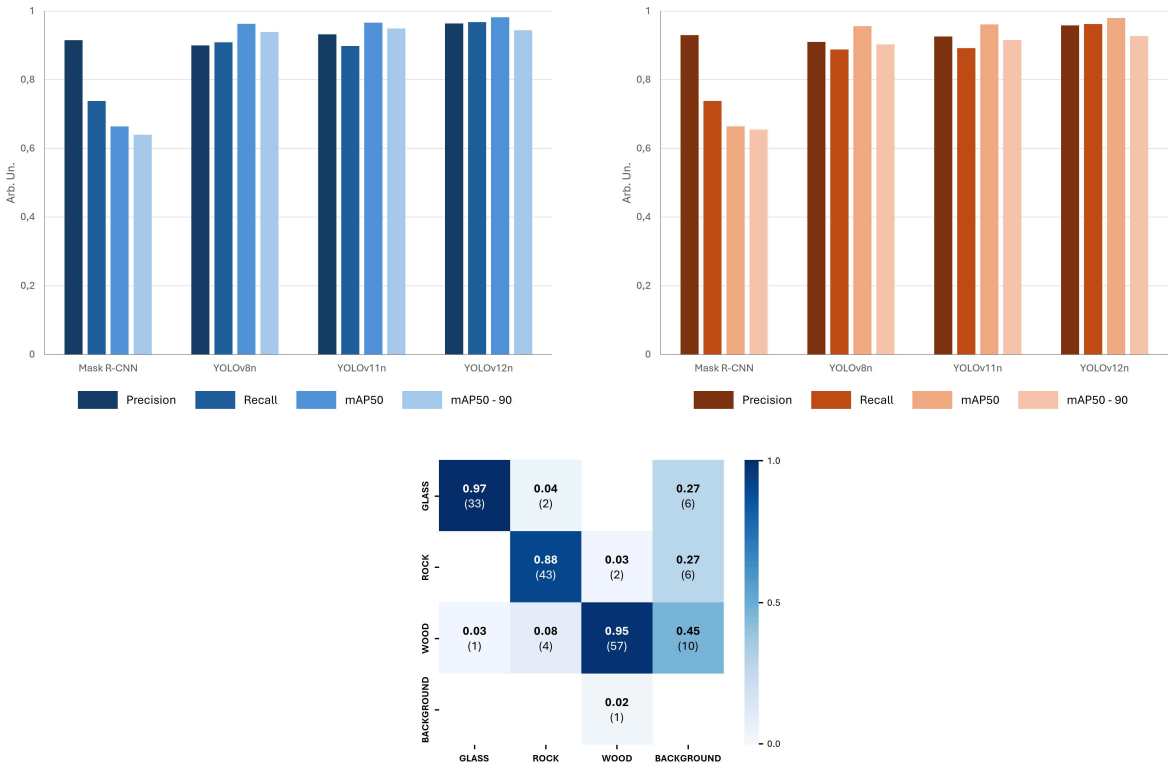


Figure 4.3: Comparative performance of evaluated segmentation models for box (left) and mask (right) metrics, showing YOLOv11 in relation to Mask R-CNN, YOLOv8 and YOLOv12. (bottom) confusion matrix for the YOLOv11 model.

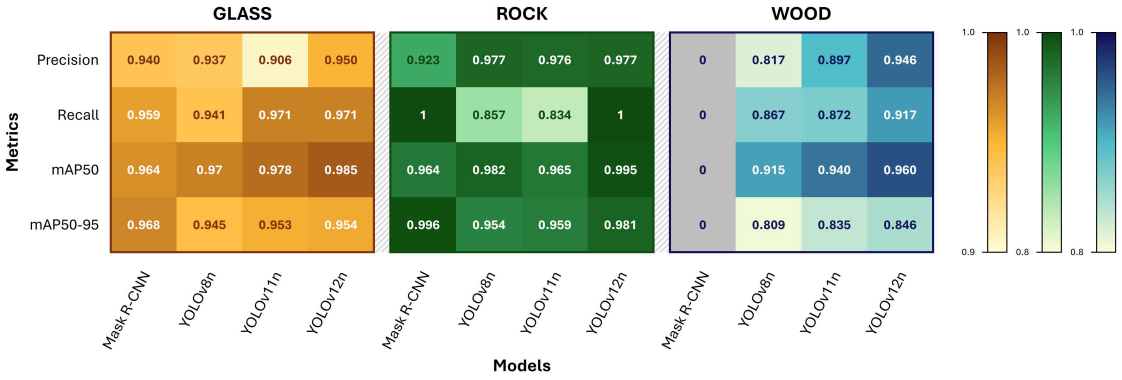


Figure 4.4: Per-class mask-performance heatmap comparing precision, recall, mAP50 and mAP50-95 across models for the target classes (glass, rock, wood).

Comparing different segmentation architectures allows not only to observe performance differences between them but, more importantly, to asses the positioning of the selected model relative to both traditional and state of the art alternatives.

Overall, the Mask R-CNN model, which represents a more traditional two-stage approach to instance segmentation, generally slower because of its two-stage processing [79], exhibits inferior performance compared to the others, achieving the lowest values indicating limitations in its ability to detect and segment the target samples (see Figure 4.3). When comparing YOLOv8 and YOLOv11, both models demonstrate similar precision metrics, yet YOLOv11n shows significantly reduced inference times on CPU ONNX (56.1 ms versus 80.4 ms for YOLOv8n [80], [81]), representing a relevant advantage for applications requiring real-time processing.

YOLOv12, on the other hand, while registering slightly higher metrics for both bounding boxes and masks, exhibits marginally slower inference times when executed on a T4 GPU with TensorRT 10 (1.64 ms versus 1.5 ms for YOLOv11n [81], [82]), emphasising the need to balance predictive performance and processing speed for this specific context.

Looking more closely at the per-class analysis, the same trend is confirmed. Each individual class (glass, rock, wood) shows that YOLOv11 maintains consistent performance, combining high precision and recall with reduced inference times (see Figure 4.4). Therefore, YOLOv11 was selected as the final software model, as it offers the best balance between accuracy and computational efficiency.

4.2. Inferring Contaminants with LIBS Data

The intelligence of the proposed system is not confined solely to the visual domain. It is precisely from this domain that a new analytical stage is triggered giving way to spectrum and chemical signatures. In this new data and decision domain, LIBS spectroscopy assumes the role of an indispensable complementary tool: as discussed it will allow to reveal properties invisible to morphological analysis by functioning as a compositional lens that decodes the elemental nature of the samples.

As previously explained, LIBS operates as a highly localised energetic probe capable of identifying the presence or absence of contaminants. The operation begins with the focusing of a high-energy laser pulse, serving as the excitation source, onto a specific region of the sample surface (see Figure 4.5). This short-duration pulse, typically lasting from a few nanoseconds to, in advanced cases, picoseconds or femtoseconds, deposits an extremely concentrated amount of energy at the impact point. Such excitation is sufficient to induce partial vapourisation of the material through rapid local temperature elevation, far exceeding its boiling point and culminating in the formation of an optically emitting plasma.

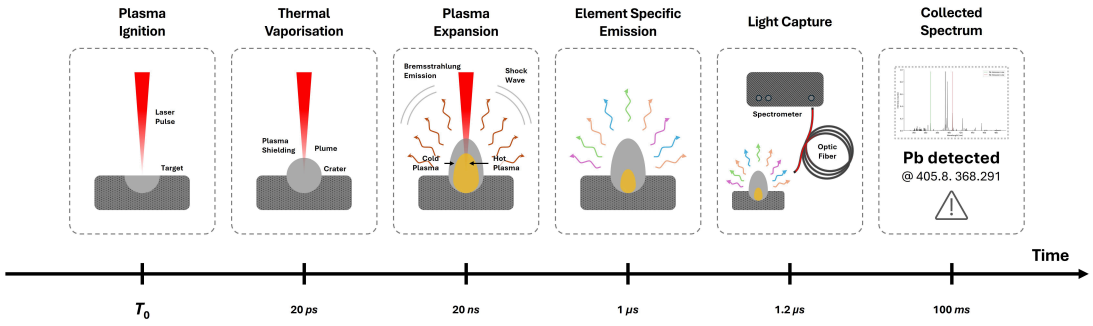


Figure 4.5: Schematic of LIBS plasma dynamics and spectral emission, illustrating the sequence from laser-induced ablation and plasma formation to the emission and collection of characteristic spectral lines.

This phenomenon results from the direct interaction between the electromagnetic field of the laser and the sample constituents, promoting gas breakdown and subsequent plasma formation, often accompanied by shock waves and acoustic emissions, reflecting the high energy intensity involved. The resulting plasma, initially unstable, dynamically evolves over a few microseconds, undergoing recombination and energetic relaxation processes. It is precisely during this cooling phase that spontaneous emission of charac-

teristic radiation occurs, producing a unique spectrum corresponding to the elemental composition.

The emission spectrum consists of multiple emission lines distributed across a spectral range extending from the ultraviolet to the infrared, and constitutes a true compositional signature of the elements present in the sample. However, direct reading and interpretation of this spectrum is far from trivial, given the high spectral density, line overlap, and dependence on experimental factors such as the ambient atmosphere, which introduce noise and complexity into the acquired signals. Therefore, rigorous and systematic computational methods become indispensable for extracting, processing, and interpreting the relevant information.

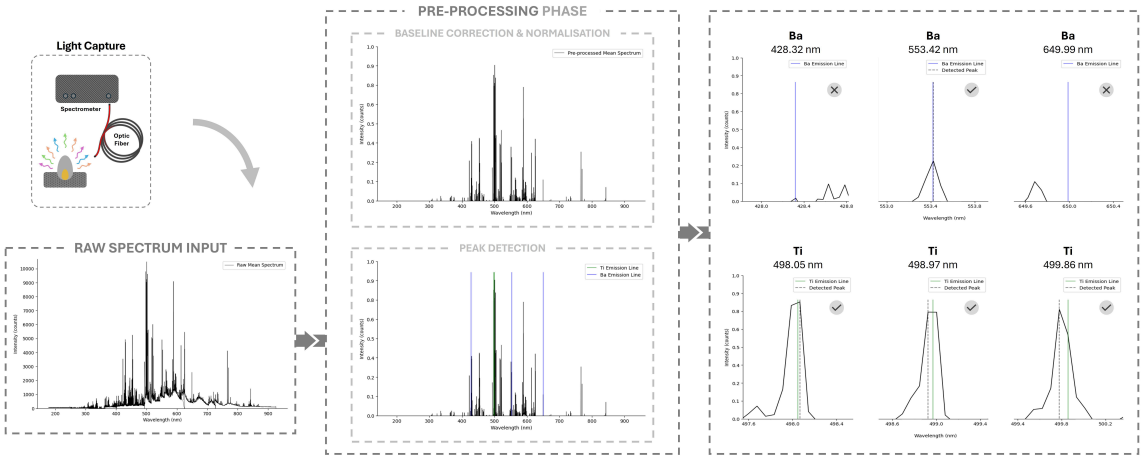


Figure 4.6: Processing pipeline for LIBS spectra: raw acquisition, pre-processing (baseline correction and normalisation), peak detection and matching against a contaminant reference database.

Just as a segmented image can be converted into categorised information through the visual model, the LIBS spectrum is transformed into chemical knowledge via a specialised analytical model developed from experimental data. The model developed essentially consists of a set of annotated spectral vectors that aggregate, for each contaminant of interest, the wavelengths corresponding to the most representative and recurrent emission lines within the spectrum. Initially, the acquired spectrum undergoes a series of pre-processing operations, including normalisation, baseline correction, and spectral peak detection (see Figure 4.6). These steps not only reduce noise and non-informative variability but also facilitate the isolation and reading of the most relevant spectral signatures [83].

After this first step, the identification of a given element's presence is then performed by direct comparison between observed peaks and characteristic wavelengths stored in the model, considering also an intensity threshold and an appropriate wavelength deviation tolerance. The presence of a contaminant is confirmed whenever at least two emission lines attributed to the same element are detected with significant intensity within defined wavelength and intensity threshold parameters (see Figure 4.7 (with reference emission lines obtained from [52])).

This process constitutes a deterministic inference method grounded in experimental knowledge, which, although not relying on statistical learning or machine learning models, provides a high degree of reliability and interpretability [52].

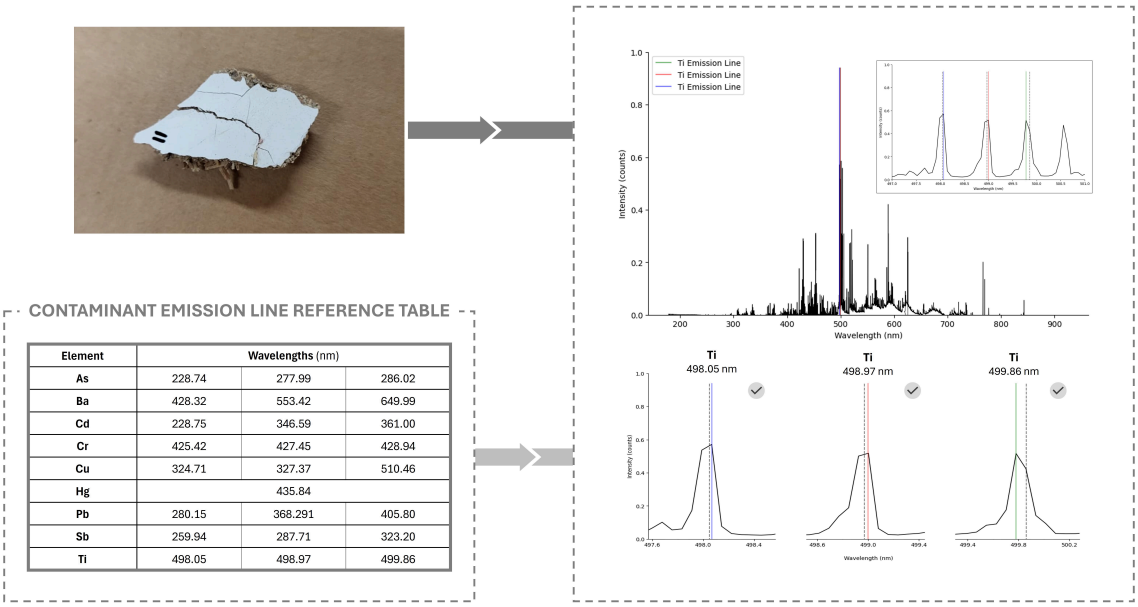


Figure 4.7: Representative LIBS result for a painted wood sample: spectral peaks compared with the contaminant emission-line reference table (match Ti), illustrating targeted elemental identification.

4.3. Concluding Remarks

This chapter has addressed the analytical foundations of modal interpretation, providing a detailed account of the computational architectures responsible for transforming sensory signals, both visual and spectral, into structured and operationalisable inferences.

Together, these two models not only enhance the system's level of analytical intelligence, but also make it adaptable to a wide range of operational scenarios. Nevertheless, it is essential to underscore that the true added value of the proposed architecture does

not lie in the isolated effectiveness of each component, but rather in the integration of both modalities, thus enabling synergies that surpass the inherent limitations of each one individually, as we shall see in the next chapter.

5. Multimodal Data Fusion and Integration Strategies

Although both Computer Vision and LIBS spectroscopy have individually demonstrated a high degree of analytical sophistication, the intelligence resulting from each modality remains constrained by structural and semantic limitations inherent to their respective domains. While an image, regardless of its resolution, does not reveal chemical composition, a spectrum lacks morphological context.

Rather than evolving as autonomous and self-contained modules, the models are now required to cooperate, whereby the individual knowledge accumulated within each modality is reused and fused into a new layer of joint inference, more informed, more contextualised, and, above all, more robust to the variabilities of the operational environment.

Multimodal fusion, here understood as the ability to synergistically combine visual interpretation and spectral interpretation, thus emerges as the foundational asset of the developed architecture. This integration forms the basis of a new layer of intelligence, enabling more robust decisions. Therefore, the true challenge, and simultaneously the potential, lies in designing an integration strategy that leverages the complementary strengths of each model.

This conceptual repositioning is concretely expressed through three distinct operational modes - “Matrix Mode”, “Singular Mode”, and “Continuous Mode” - designed to explore varying degrees of integration, autonomy, and technical sophistication. This modular organisation accommodates scenarios ranging from laboratory and academic contexts to demanding industrial environments with high processing throughput, representing a deliberate progression in terms of complexity and autonomy.

This chapter will discuss the conceptual and technical architecture of the adopted fusion strategy, with a particular focus on the mechanisms of articulation between the visual and compositional models. It will explore the concrete forms of coordination between the Computer Vision and LIBS spectroscopy subsystems, as well as the performance metrics associated with their joint operation. Finally, the chapter will present the different operational modes developed, each reflecting progressively higher levels of functional integration, tailored to scenarios with varying degrees of analytical demand and processing throughput.

5.1. Matrix Mode for High-Precision Mapping

The “Matrix Mode” offers a systematic and controlled approach, oriented toward detailed and fixed-sample acquisition, while the subsequent modes introduce increasingly adaptive and context-aware intelligence, culminating in real-time response capabilities.

While still a unimodal approach, the “Matrix Mode” based exclusively on LIBS spectroscopy, this mode serves as the foundation for the subsequent multimodal modalities. It was designed to replicate established methodologies of spectral mapping previously explored in other LIBS systems developed at the INESC TEC Centre for Applied Photonics, following a point-by-point acquisition principle.

Inspired by the traditional paradigm of two-dimensional scanning, the system performs a systematic sweep across the surface of the sample along the x and y axes, resulting in the generation of a spatial data matrix, also referred to as a spectral cube. At each two-dimensional coordinate, an individual spectrum is acquired and associated, enabling the reconstruction of elemental distribution maps with high spatial resolution, an approach particularly suited to the detailed characterisation of surface heterogeneities.

In conventional LIBS systems, this strategy typically requires not only the immobilisation of the optical system but also the controlled motion of the sample across automated platforms. Generally, the sample is adjusted along the z-axis to ensure proper laser focus and then moved with precision across the x and y planes while the acquisition components remain fixed.

In contrast, the system developed in this project inverts the logic of relative movement between components, while maintaining its underlying principles. In this configuration, the sample remains stationary on the industrial conveyor belt, which is halted during acquisition, while the Robotic System system (Gantry), coupled with the entire LIBS module, including the laser and optical collection devices, is responsible for executing the coordinated movement across the area of interest (see Figure 5.1 (showcase video [84])). This structural change introduces additional flexibility, enabling the analysis of larger or irregularly shaped samples, significantly expanding the range of materials that can be characterised. The primary constraints are the Gantry system’s effective scanning area and its minimum step resolution.

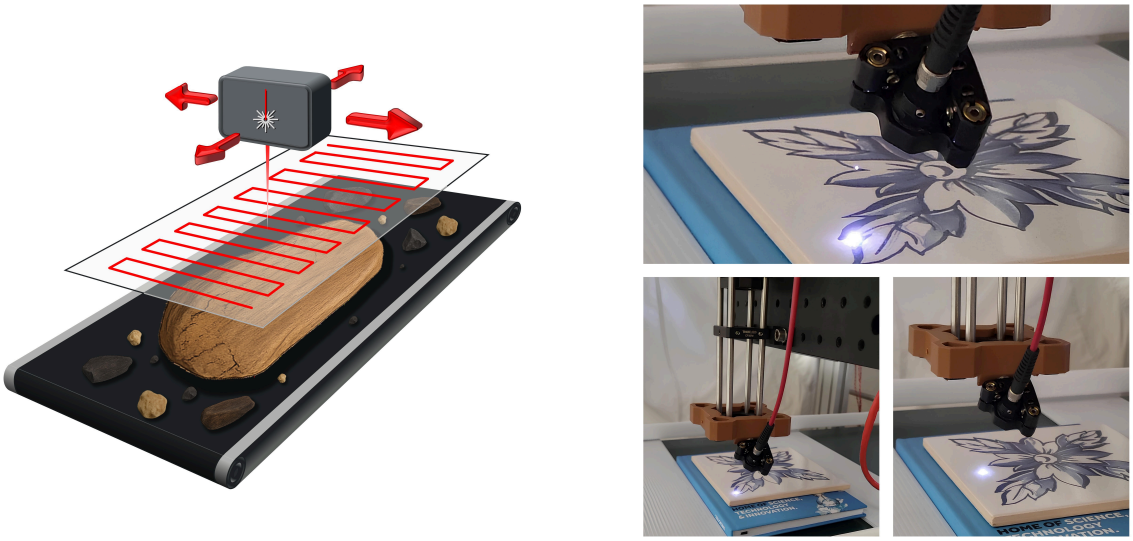


Figure 5.1: Matrix Mode demonstration: left, schematic of the gantry-driven laser raster scanning the halted sample on the conveyor; right, representative pictures illustrating execution on a tile.

The execution and initialization of the process follow a level of simplicity comparable to that of traditional LIBS systems, beginning with focus adjustment along the z-axis and sample alignment within the analysis area. Operational parameters are then defined, including the initial and final coordinates of the scan path, the number of matrix rows and columns (which determine the spatial density of acquisition), and the scanning speed. The resulting movement follows a zigzag trajectory, commonly referred to as a “whiskbroom” scanning pattern, systematically covering each line of the defined matrix (see Figure 5.1).

Although “Matrix Mode” provides extremely detailed chemical and morphological characterisation, its applicability in industrial settings is limited, not due to inherent constraints of the acquisition process itself, but due to its incompatibility with the continuous operational flow that characterises such environments. The requirement for the sample to remain stationary throughout acquisition, deeming for a complete halt in the production line for dedicated analysis, results in a lack of scalability and reduced processing throughput. Nevertheless, this mode remains a valuable tool in laboratory or experimental development contexts, where the priority lies in analytical precision and the detailed representation of the elemental composition of samples, as in the case of cultural heritage artifacts (see Figure 5.2 and Figure 5.3).

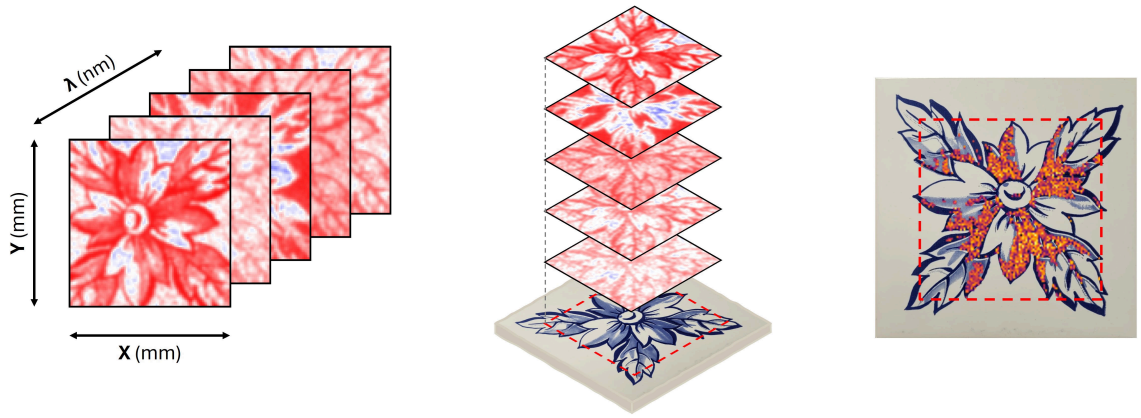


Figure 5.2: Matrix Mode spectral mapping results for a tile: single-wavelength maps, a vertical-layer composite, and a top-view contaminant mask that localises spectral signatures.

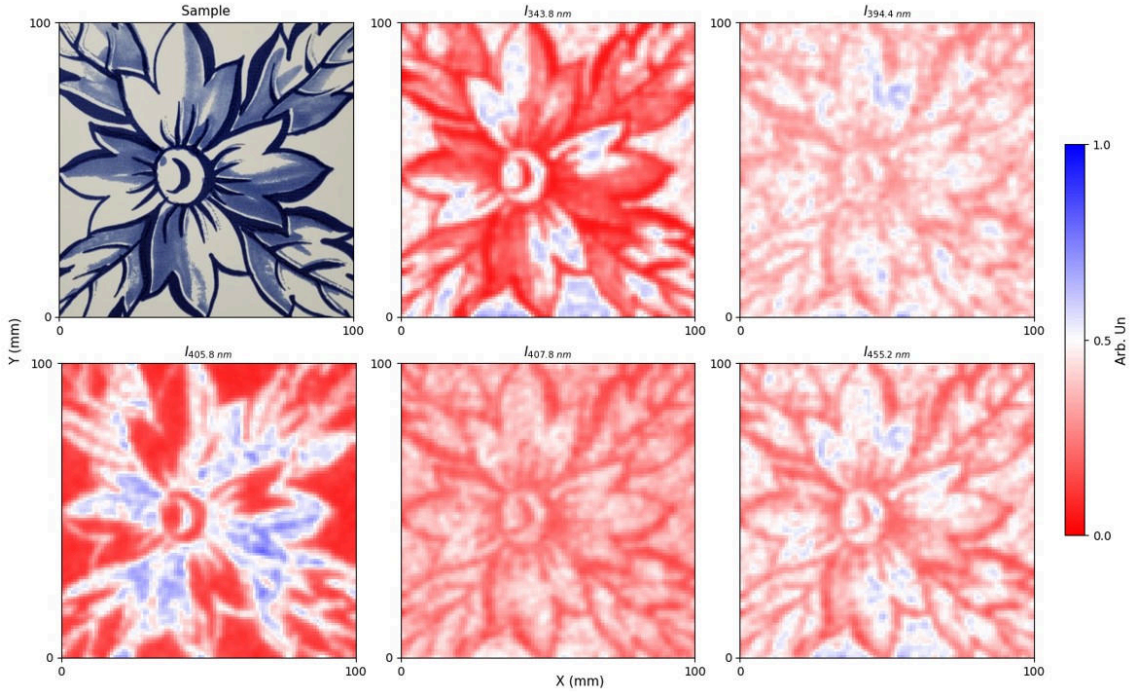


Figure 5.3: Multi-channel spectral maps presented side-by-side to emphasise wavelength-dependent spatial contrast across the analysed tile.

5.2. Singular Mode for Directed and Vision-Guided Analysis

In contrast with the meticulous but inherently static approach of the “Matrix Mode” the “Singular Mode”, an intermediate successor in the system’s strategic evolution, emerges as a step toward greater agility and adaptability, introducing a more dynamic and context-aware operational model. Here, chemical analysis is guided by real-time visual information, enabling each laser firing decision to be based on the morphological identification and position of the samples as they move along the conveyor line.

This articulation between Computer Vision and spectral analysis translates into a more refined operational capacity, in which the system not only detects objects of interest but also determines, based on spatial and temporal, which should be analysed. The paradigm of exhaustive and rigid coverage is thus replaced by a targeted and context-driven operation, representing a critical advancement for high-throughput semi-industrial or laboratory applications.

In this new hybrid mode, unlike the previous one, the samples move along the industrial conveyor belt, enabling a continuous and fluid analysis. This configuration proves particularly effective when the samples are sparsely distributed, being analysed individually through a single laser shot, perpendicular to the line flow. Each shot generates an immediate result, allowing for real-time flagging of samples that exhibit predefined critical contaminants, which could subsequently be sorted or removed automatically (see Figure 5.4).

The vision system assumes a central role in this configuration. Through continuous image acquisition and the application of the trained detection model, automatic segmentation and classification of the samples is performed. The information gathered extends beyond morphology or position and forms the basis for subsequent LIBS decisions. For each identified sample, the geometric centre is computed from the generated bounding box, which is employed as a simplified and computationally efficient representation of the object, thereby defining the representative point for the laser shot and preceding operations.

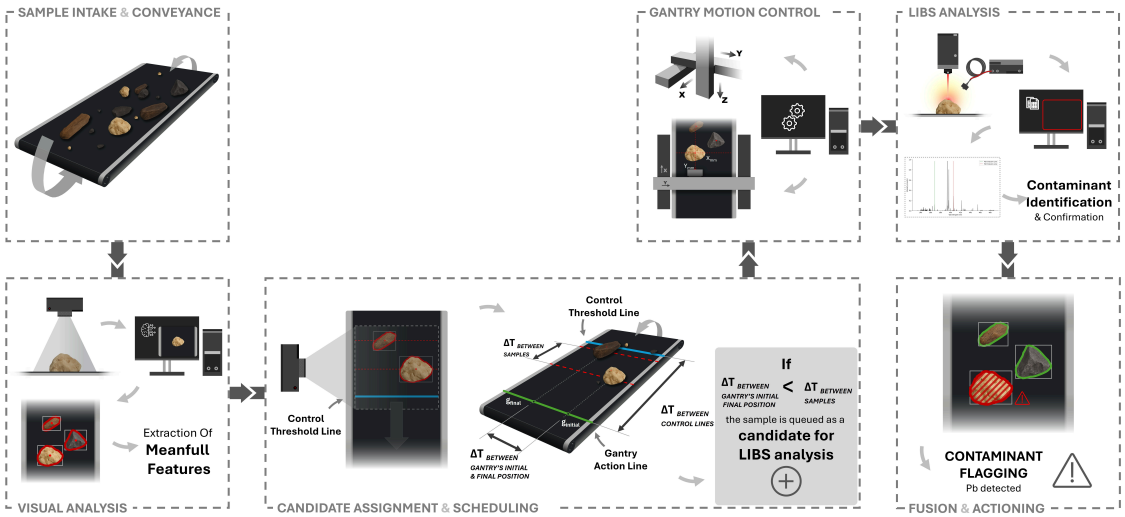


Figure 5.4: Focused operational pipeline for the Singular Mode, showing the ordered sequence from sample intake and visual analysis through candidate assignment and scheduling to fusion and actioning.

The effectiveness and reliability of the process are ensured by a dual-stage validation procedure, both spatial and temporal in nature. In the first stage, the complete exit of the sample from the field of view is verified, based on the intersection of its centre with a virtual line strategically placed at the lower boundary of the captured frame (blue line, shown in the schematic within the “candidate Assignment & Scheduling” block, see Figure 5.4). This seemingly simple mechanism avoids errors caused by partial segmentations or timing delays, which are common when the check is based solely on the image’s bottom edge.

Next, the Gantry system’s response time is simulated, taking into account its motion curve, to accurately estimate the time required to reach the predicted position of the sample at the exact moment of analysis. Several dynamic parameters are considered, including speeds, accelerations, and initial and final coordinates, allowing for comparison between this calculated time and the interval available before the sample reaches the laser’s operational zone. With this forecast, the system can anticipate whether a given sample can be analysed in time, thus avoiding inefficient Gantry movements or failed analyses. If a sample falls outside the viable window, due to, for instance, overlap with others or positioning near the edge of the field, the system opts to skip it and redirect resources toward subsequent samples within acceptable parameters.

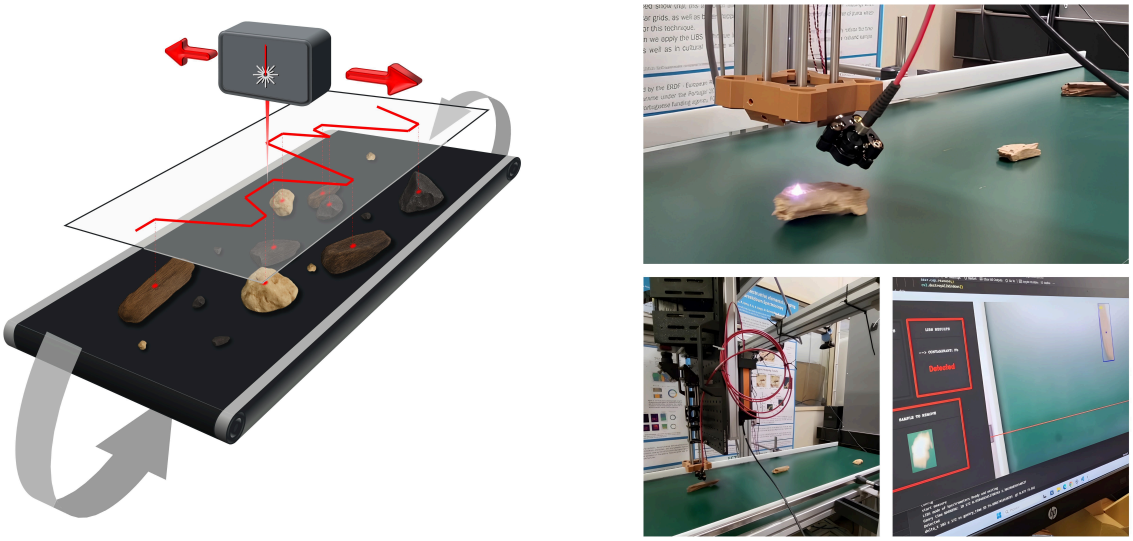


Figure 5.5: Singular Mode demonstration with sparse samples: left, schematic illustrating the laser sweep/filter over the conveyor; right, selected frames from the demonstration video showing targeted shots on isolated samples.

The careful management of samples along the conveyor is supported by a tracking system based on the SORT algorithm, which allows for the individual identification and monitoring of each detected sample. It assigns a unique identifier to every new object entering the field of view and maintains its trajectory throughout the process. This approach not only facilitates the matching between the captured image and the spectral analysis obtained, but also ensures robust traceability of contamination decisions, thereby enhancing system reliability. By combining visual records with unique numerical tags, ambiguity in the correspondence between detection and chemical analysis is eliminated.

Once the viability of analysing a candidate sample with a specific marker is confirmed, the LIBS analysis phase begins. The Gantry performs a fast unidirectional movement, positioning the laser over the predefined point, where a single shot is executed (see Figure 5.5 (showcase video [85])). While this localised approach does not allow for exhaustive characterisation, it provides a representative reading of the sample's overall elemental content. It is important to recognize, however, that the localised nature of the analysis may not detect contaminants that are distributed non-uniformly or within the bulk of the sample, constituting an inherent limitation of this approach.

The spectral processing sequence follows the same principles described for the "Matrix Mode", encompassing the steps of acquisition, pre-processing, and classification. The

results are then presented on a dedicated interface specifically developed for this purpose, providing clear and accessible indications of the contamination class, captured image, and other relevant information, such as the type of contaminant detected (see Figure 5.6). This tool thus plays an essential and central role in process supervision and in supporting informed real-time decision-making.

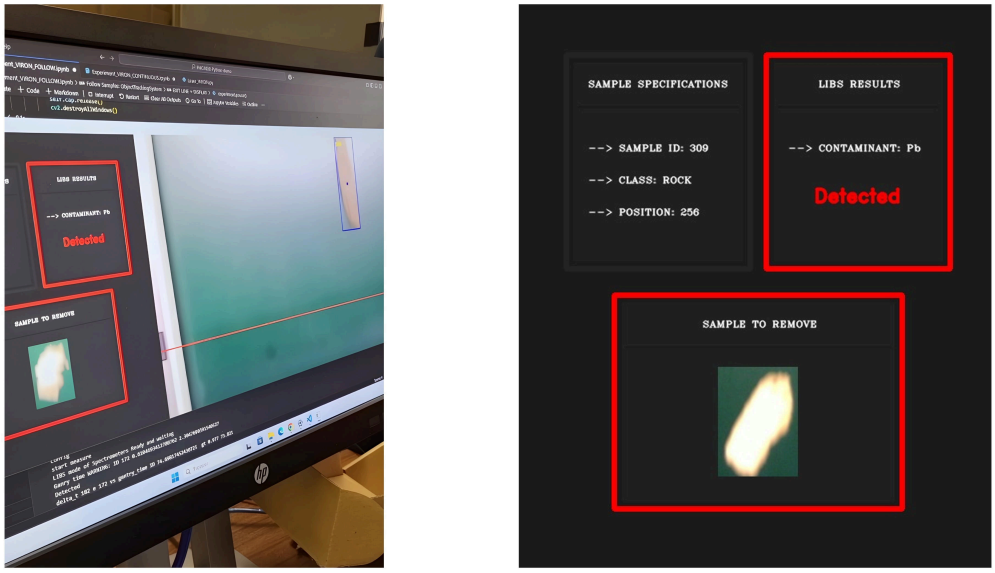


Figure 5.6: Dedicated user interface developed for the system, presenting the contamination class, captured image, and complementary information such as the type of contaminant detected.

5.3. Continuous Mode for Continuous Acquisition and Full Autonomy

Despite the optimizations introduced in the “Singular Mode” and its proven flexibility in laboratory and semi-industrial contexts, it remains constrained by minimum spacing requirements between samples and by the inherent physical and temporal limitations associated with repositioning the Gantry system. In real industrial environments, where processing rates are high and the density of transported materials challenges response times, these constraints become particularly detrimental, directly affecting the scalability and feasibility of the system in continuous operation.

To overcome these limitations and ensure an effective response to the demands of modern production lines, the “Continuous Mode” was developed as the final stage in the system’s strategic evolution. This new paradigm replaces the logic of discrete action with uninterrupted operation, in which the Gantry module performs a constant oscillatory movement along the transverse axis of the conveyor belt, synchronising with the motion

of the samples and enabling continuous acquisition of spectral data (see Figure 5.7 (showcase video [86])). As such, chemical analysis transitions from an episodic to a fluid and adaptive process, supported by a high and configurable pulse rate (up to 20 Hz), eliminating the need for physical separation between samples and maximising the density of acquired information.

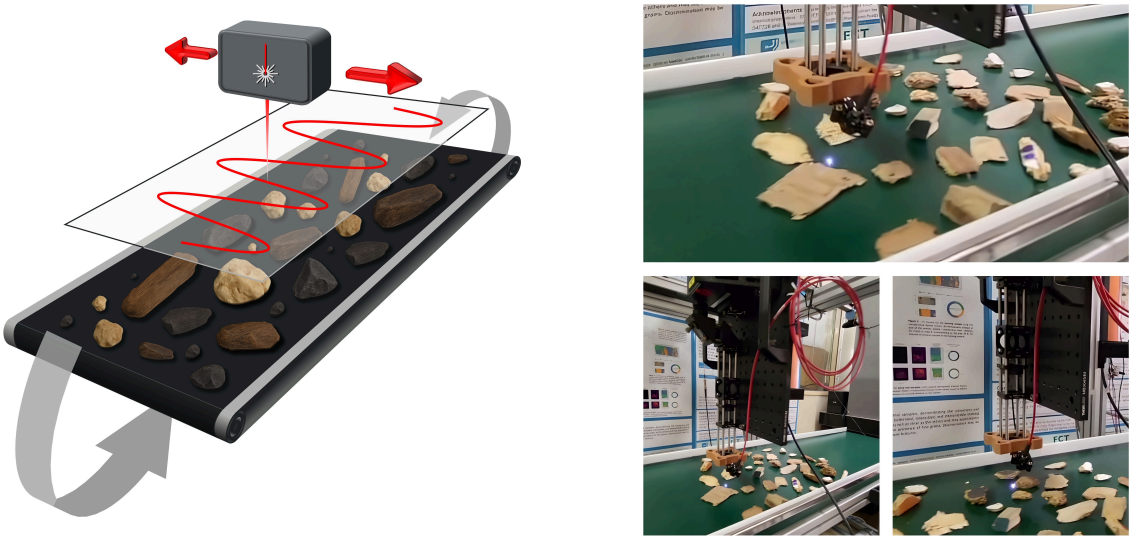


Figure 5.7: Continuous Mode demonstration under high sample density: left, schematic of oscillatory laser coverage across the moving conveyor belt; right, frames from the demonstration video showing continuous operation on densely packed samples.

This oscillatory operation introduces multiple advantages, not only in enabling the analysis of multiple points per sample, but also in facilitating the characterisation of continuous or adjacent areas. This expands the spatial representativeness of the readings and provides a more robust and reliable mapping of the transported material. As such, this approach effectively addresses the morphological and compositional heterogeneity of the samples, making it particularly relevant in scenarios characterised by dense, irregular flows with partially overlapping samples. Despite changes in acquisition dynamics, the system's core architecture maintains logical continuity with previous modes, particularly during the initial stage of visual detection, segmentation, and classification of samples. Each sample is identified upon entering the field of view and assigned a unique identifier along with a high-precision digital mask that defines the regions of interest eligible for LIBS analysis.

However, the absence of physical and mobile feedback sensors at the tip of the LIBS system, such as cameras or optical tracking sensors, introduced a new challenge, to accurately anticipate the laser's impact points on the physical plane of the samples, taking into account the Gantry's oscillatory motion. To meet this requirement, a predictive methodology was developed based on mathematical modelling of the Gantry's trajectory. By experimentally collecting multiple displacement points, it became possible to reconstruct the oscillation curve of the Gantry movement and, based on the configured laser pulse frequency, calculate the predicted positions for each laser shot (see Figure 5.8).

Initially, each of these predicted points was individually checked for intersection with valid regions of the segmented masks. However, this strategy proved too computationally demanding and incompatible with the low-latency requirements of real-time operation.

In response, a more efficient model was devised in the form of a compact digital structure, a "digital stamp", that aggregates the entire laser trajectory and its simulated shots across a complete oscillation cycle into a single unit. This stamp is divided into two halves, corresponding to the positive and negative directions of oscillation, and is periodically overlaid onto the global mask of segmented samples. This enables intersection checks to be carried out simultaneously and vectorially.

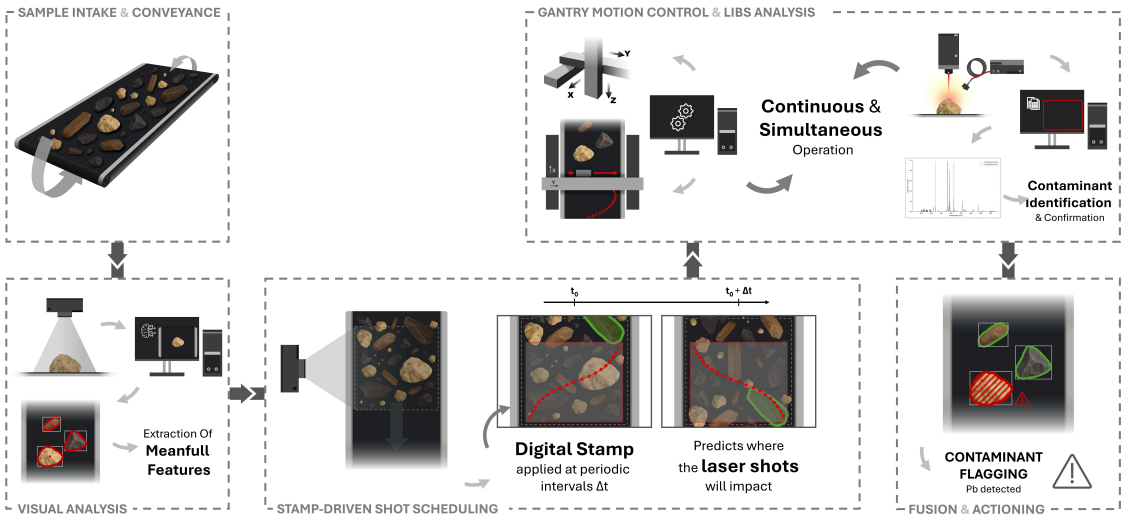


Figure 5.8: Focused operational pipeline for the Continuous Mode, showing the stamp-driven shot scheduling, gantry motion control and LIBS analysis feeding the fusion and actioning stage.

With this new operational model, the system becomes capable of truly continuous functioning, successively analysing different regions of the conveyor belt without the need for pauses, reconfigurations, or prior decisions based on discrete detections.

Finally, the results of the spectral analysis, following pre-processing and classification, are integrated into a real-time supervision panel, similar to those used in previous modes but now capable of handling multiple laser shots per sample. This panel clearly and accessibly displays the detected contaminations, their spatial distribution, and provides automatic identification of the areas to be removed or sorted.

5.4. Concluding Remarks

The evolution presented throughout this chapter reflects a deliberate and strategic trajectory toward greater integration, adaptability, and operational intelligence within the developed system. From the “Matrix Mode”, rooted in conventional laboratory practices, to the “Continuous Mode”, capable of operating in high-throughput and complex industrial settings, each configuration was designed to meet distinct technical and contextual requirements, ensuring broad and scalable functional coverage.

The modular structure of the three operational modes does not merely constitute a technical adaptation to the physical and temporal constraints of the environment, rather, it represents a progressive materialization of the concept of multimodal fusion. By synergistically combining the morphological acuity of Computer Vision with the compositional precision of LIBS spectroscopy, the system acquires the capacity to make more informed and context-aware decisions, thereby overcoming the inherent limitations of each modality in isolation.

Each mode should not be regarded as an end in itself, but rather as a step within an adaptive architecture, where resources are mobilized proportionally to the demands of the operational scenario.

6. Conclusion

The work developed in this dissertation focused on the exploration of the fusion and synergy between two distinct, yet intrinsically complementary technologies: computer vision and LIBS. The motivation was grounded in the conviction that the combination of heterogeneous data streams, morphological and compositional, could give rise to a more comprehensive, robust, and reliable inspection system, overcoming the limitations associated with the isolated use of each modality. During the course of this research, it also became evident that a gap exists in the state of the art regarding the effective integration of these technologies in dynamic, high-throughput industrial environments. This ultimately manifestsnin the absence of mature solutions, which reinforces the relevance and timeliness of the work, moving towards the development of a hybrid solution, composed of a physical layer (hardware), responsible for data acquisition and synchronisation, and a digital layer (software), dedicated to data processing, fusion, and decision-making.

Based on this structure, we designed and implemented different operational modes, tailored to address distinct scenarios and temporal requirements. Overall, the results enclosed in the thesis demonstrate that multimodal integration is not only technically feasible but also adds substantial value, resulting in significant improvements in system selectivity, reliability, and adaptability.

Nonetheless, as we put the work in a broader perspective, it is also important to acknowledge the limitations identified, which are inevitable in any innovative process. Focusing specifically on the operational modes with greater industrial orientation, the “Singular Mode” and “Continuous Mode” present operational constraints that limit their immediate applicability in real production lines. Starting with the “Singular Mode” it may prove useful in laboratory contexts, low-throughput production lines, or situations where selective analysis of individual samples is required. However, its effectiveness is reduced in high-throughput environments, as its operational logic relies on the analysis of sparsely positioned samples on the conveyor belt. Regarding the “Continuous Mode”, although the software demonstrates the capacity to cope with the high throughput of industrial environments, the available hardware revealed significant limitations. In particular, the oscillatory scanning of the LIBS system generates dead zones, preventing full coverage of the analysed surface. Some mitigating measures, such as increasing the speed of the robotic movement system (gantry) or reducing the speed of the motorised conveyor, may

alleviate the problem, but carry the risk of compromising the overall process balance. A more promising strategy, aligned with the multimodal philosophy, lies in adopting statistical and inferential approaches capable of extrapolating reliable information from effectively analysed regions, thereby mitigating the system's physical limitations.

These considerations highlight that, beyond the advancements already achieved, there remain clear opportunities for further development that can be explored in future work. Nevertheless, the results obtained demonstrate that the true strength of the system resides precisely in the articulation between modalities, rather than in their isolated exploitation, reinforcing the potential of the multimodal approach as a solid pathway for high-demand industrial applications.

6.1. Perspectives and Future Work

The final reflections of this work make it clear that, although the fusion of computer vision and LIBS already represents a significant advancement, the true potential of the multimodal approach is far from being fully realised. The limitations identified in the operational modes developed, as well as the opportunities opened by the concept of integrating heterogeneous information sources, highlight clear avenues for future development.

Some potential directions for future research include:

- **Expansion of the multimodal paradigm**: Beyond the combination explored in this dissertation, it is feasible to integrate additional diagnostic and characterisation techniques, such as Raman spectroscopy, hyperspectral imaging, or X-ray fluorescence. These resources are already available within the specialisations of the Applied Photonics research group at INESC TEC. Such an expansion would further enrich the fusion of heterogeneous data, enhancing the capacity for contaminant detection and characterisation across different materials.

References

[1] A. R. Hasan, L. Hu, H. M. Solo-Gabriele, L. Fieber, Y. Cai, and T. G. Townsend, “Field-scale leaching of arsenic, chromium and copper from weathered treated wood,” *Environmental Pollution*, vol. 158, no. 5, pp. 1479–1486, 2010, doi: https://doi.org/10.1016/j.envpol.2009.12.027. Available: https://www.sciencedirect.com/science/article/pii/S0269749109006356

[2] D. o. E. United Nations and P. D. Social Affairs, “World Population Prospects - Probabilistic Projections: Total Population,” 2024. Available: https://population.un.org/wpp/graphs?loc=900&type=Probabilistic%20Projections&category=Population&subcategory=1_Total%20Population

[3] H. Khan and I. Khan, “The effect of technological innovations, urbanization and economic growth on environmental quality: does governance matter?,” *Frontiers in Environmental Science*, 2023, doi: 10.3389/fenvs.2023.1239288. Available: https://www.frontiersin.org/journals/environmental-science/articles/10.3389/fenvs.2023.1239288

[4] V. K. Gaur, P. Sharma, R. Sirohi, M. K. Awasthi, C.-G. Dussap, and A. Pandey, “Assessing the impact of industrial waste on environment and mitigation strategies: A comprehensive review,” *Journal of Hazardous Materials*, vol. 398, p. 123019, 2020, doi: https://doi.org/10.1016/j.jhazmat.2020.123019. Available: https://www.sciencedirect.com/science/article/pii/S0304389420310086

[5] J. Zang, B. Luo, G. He, C. Yue, and Y. Yan, “The mini review of sustainable urban environment and emerging contaminants,” *Current Opinion in Environmental Science & Health*, vol. 47, p. 100648, 2025, doi: https://doi.org/10.1016/j.coesh.2025.100648. Available: https://www.sciencedirect.com/science/article/pii/S2468584425000571

[6] X. Li, X. Shen, W. Jiang, Y. Xi, and S. Li, “Comprehensive review of emerging contaminants: Detection technologies, environmental impact, and management strategies,” *Ecotoxicology and Environmental Safety*, vol. 278, p. 116420, 2024, doi: https://doi.org/10.1016/j.ecoenv.2024.116420. Available: https://www.sciencedirect.com/science/article/pii/S0147651324004962

[7] R. Khanna *et al.*, “Microplastics and Nanoplastics as Environmental Contaminants of Emerging Concern: Potential Hazards for Human Health,” *Sustainability*, vol. 16, p. 8704, 2024, doi: 10.3390/su16198704

[8] N. Chandimali, J. Bae, J. Y. Hwang, S. G. Bak, S. H. Cheong, and S. J. Lee, “Integrated Hygiene Control Strategies in Food Manufacturing: Technologies, Regulations, and Socioeconomic Impacts,” *Journal of Food Protection*, vol. 88, no. 7, p. 100547, 2025, doi: https://doi.org/10.1016/j.jfp.2025.100547. Available: https://www.sciencedirect.com/science/article/pii/S0362028X25000997

[9] I. Uddin, “Onsite visual detection of heavy metal contaminants using impregnated strip,” *Journal of Photochemistry and Photobiology A: Chemistry*, vol. 421, p. 113512, 2021, doi: https://doi.org/10.1016/j.jphotochem.2021.113512. Available: https://www.sciencedirect.com/science/article/pii/S1010603021003816

[10] Y. Xie *et al.*, “Combined detection of foodborne pathogens in irrigation water in Jiangsu, China,” *Food Microbiology*, vol. 133, p. 104885, 2026, doi: https://doi.org/10.1016/j.fm.2025.104885. Available: https://www.sciencedirect.com/science/article/pii/S0740002025001650

[11] X. Yan *et al.*, “Systematic review of the microbiological status of farmed and processed Tenebrio molitor: Insights on foodborne pathogens in food and feed applications,” *Food Control*, vol. 176, p. 111340, 2025, doi: https://doi.org/10.1016/j.foodcont.2025.111340. Available: https://www.sciencedirect.com/science/article/pii/S0956713525002099

[12] L. Zhang, A. Dehghani, Z. Su, T. King, B. Greenwood, and M. Levesley, “Real-time automated visual inspection system for contaminant removal from wool,” *Real-Time Imaging*, vol. 11, no. 4, pp. 257–269, 2005, doi: https://doi.org/10.1016/j.rti.2004.09.003. Available: https://www.sciencedirect.com/science/article/pii/S1077201404000889

[13] M. H. A. Bakar *et al.*, “Oil contaminant detection using long-range surface plasmon resonance-based sensors,” *Microchemical Journal*, vol. 212, p. 113576, 2025, doi: https://doi.org/10.1016/j.microc.2025.113576. Available: https://www.sciencedirect.com/science/article/pii/S0026265X25009300

[14] Y. Yuan, H. Jia, D. Xu, and J. Wang, “Novel method in emerging environmental contaminants detection: Fiber optic sensors based on microfluidic chips,” *Science*

of The Total Environment, vol. 857, p. 159563, 2023, doi: https://doi.org/10.1016/j.scitotenv.2022.159563. Available: https://www.sciencedirect.com/science/article/pii/S0048969722066621

[15] H. N. Hussain *et al.*, "Advances in the removal of Polyfluoroalkyl Substances (PFAS) from water using destructive and non-destructive methods," *Green Analytical Chemistry*, vol. 12, p. 100225, 2025, doi: https://doi.org/10.1016/j.greeac.2025.100225. Available: https://www.sciencedirect.com/science/article/pii/S2772577425000229

[16] H. Jiang, F. Meng, J. Deng, and J. Zhu, "Detecting fossil oil contaminants in peanut oil using NIR spectroscopy: Enhanced feature selection through mutual information and Shapley Additive Explanations," *Infrared Physics & Technology*, vol. 150, p. 105960, 2025, doi: https://doi.org/10.1016/j.infrared.2025.105960. Available: https://www.sciencedirect.com/science/article/pii/S1350449525002531

[17] A. Khumaeni, W. Setia Budi, K. Hendrik Kurniawan, K.-I. Fukumoto, K. Kurihara, and K. Kagawa, "Quantification of sodium contaminant on steel surfaces using pulse CO2 laser-induced breakdown spectroscopy," *Arabian Journal of Chemistry*, vol. 15, no. 1, p. 103474, 2022, doi: https://doi.org/10.1016/j.arabjc.2021.103474. Available: https://www.sciencedirect.com/science/article/pii/S1878535221004895

[18] S. Ma, Y. Li, and Y. Peng, "Spectroscopy and computer vision techniques for noninvasive analysis of legumes: A review," *Computers and Electronics in Agriculture*, vol. 206, p. 107695, 2023, doi: https://doi.org/10.1016/j.compag.2023.107695. Available: https://www.sciencedirect.com/science/article/pii/S0168169923000832

[19] Y.-Y. Pu, C. O'Donnell, J. T. Tobin, and N. O'Shea, "Review of near-infrared spectroscopy as a process analytical technology for real-time product monitoring in dairy processing," *International Dairy Journal*, vol. 103, p. 104623, 2020, doi: https://doi.org/10.1016/j.idairyj.2019.104623. Available: https://www.sciencedirect.com/science/article/pii/S0958694619302602

[20] H. M. Balaha, A. E.-S. Hassan, R. A. Ahmed, and M. H. Balaha, "Advancing eye disease detection: A comprehensive study on computer-aided diagnosis with vision transformers and SHAP explainability techniques," *Biocybernetics and Biomedical Engineering*, vol. 45, no. 1, pp. 23–33, 2025, doi: https://doi.org/10.1016/j.bbe.2024.11.005. Available: https://www.sciencedirect.com/science/article/pii/S0208521624000883

[21] W. Deng *et al.*, “Optical MRI imaging based on computer vision for extracting and analyzing morphological features of renal tumors,” *SLAS Technology*, vol. 29, no. 5, p. 100192, 2024, doi: https://doi.org/10.1016/j.slast.2024.100192. Available: https://www.sciencedirect.com/science/article/pii/S2472630324000748

[22] M. M. Kabir, A. Rahman, M. N. Hasan, and M. Mridha, “Computer vision algorithms in healthcare: Recent advancements and future challenges,” *Computers in Biology and Medicine*, vol. 185, p. 109531, 2025, doi: https://doi.org/10.1016/j.compbiomed.2024.109531. Available: https://www.sciencedirect.com/science/article/pii/S0010482524016160

[23] J. Ye, Z. Yu, J. Lin, H. Li, and L. Lin, “Vision foundation model for agricultural applications with efficient layer aggregation network,” *Expert Systems with Applications*, vol. 257, p. 124972, 2024, doi: https://doi.org/10.1016/j.eswa.2024.124972. Available: https://www.sciencedirect.com/science/article/pii/S0957417424018396

[24] D. Bárbulo Barrios, J. Valente, and F. van Langevelde, “Monitoring mammalian herbivores via convolutional neural networks implemented on thermal UAV imagery,” *Computers and Electronics in Agriculture*, vol. 218, p. 108713, 2024, doi: https://doi.org/10.1016/j.compag.2024.108713. Available: https://www.sciencedirect.com/science/article/pii/S0168169924001042

[25] S. Ghazal, A. Munir, and W. S. Qureshi, “Computer vision in smart agriculture and precision farming: Techniques and applications,” *Artificial Intelligence in Agriculture*, vol. 13, pp. 64–83, 2024, doi: https://doi.org/10.1016/j.aiia.2024.06.004. Available: https://www.sciencedirect.com/science/article/pii/S2589721724000266

[26] M. Mahmood Ur Rehman, J. Liu, A. Nijabat, M. Faheem, W. Wang, and S. Zhao, “Leveraging Convolutional Neural Networks for Disease Detection in Vegetables: A Comprehensive Review,” *Agronomy*, vol. 14, p. , 2024, doi: 10.3390/agronomy14102231

[27] Y. Bai, B. Zhang, N. Xu, J. Zhou, J. Shi, and Z. Diao, “Vision-based navigation and guidance for agricultural autonomous vehicles and robots: A review,” *Computers and Electronics in Agriculture*, vol. 205, p. 107584, 2023, doi: https://doi.org/10.1016/j.compag.2022.107584. Available: https://www.sciencedirect.com/science/article/pii/S0168169922008924

[28] H. Gajjar, S. Sanyal, and M. Shah, “A comprehensive study on lane detecting autonomous car using computer vision,” *Expert Systems with Applications*, vol. 233, p. 120929, 2023, doi: https://doi.org/10.1016/j.eswa.2023.120929. Available: https://www.sciencedirect.com/science/article/pii/S0957417423014318

[29] A. V. R. Katkuri, H. Madan, N. Khatri, A. S. H. Abdul-Qawy, and K. S. Patnaik, “Autonomous UAV navigation using deep learning-based computer vision frameworks: A systematic literature review,” *Array*, vol. 23, p. 100361, 2024, doi: https://doi.org/10.1016/j.array.2024.100361. Available: https://www.sciencedirect.com/science/article/pii/S2590005624000274

[30] M. A.-M. Khan *et al.*, “LLDNet: A Lightweight Lane Detection Approach for Autonomous Cars Using Deep Learning,” *Sensors*, vol. 22, p. 5595, 2022, doi: 10.3390/s22155595

[31] Q. Liao *et al.*, “Improving traceability and quality control in the red-meat industry through computer vision-driven physical meat feature tracking,” *Food Chemistry*, vol. 480, p. 143830, 2025, doi: https://doi.org/10.1016/j.foodchem.2025.143830. Available: https://www.sciencedirect.com/science/article/pii/S0308814625010817

[32] B. Amaral Pascarelli Ferreira, K. Soares Augusto, J. César Álvarez Iglesias, T. Dias Pinheiro Caldas, R. Bryan Magalhães Santos, and S. Paciornik, “Instance segmentation of quartz in iron ore optical microscopy images by deep learning,” *Minerals Engineering*, vol. 211, p. 108681, 2024, doi: https://doi.org/10.1016/j.mineng.2024.108681. Available: https://www.sciencedirect.com/science/article/pii/S0892687524001109

[33] E. Casas, L. T. Ramos, C. Romero, and F. Rivas-Echeverría, “A review of computer vision applications for asset inspection in the oil and gas industry,” *Journal of Pipeline Science and Engineering*, p. 100246, 2024, doi: https://doi.org/10.1016/j.jpse.2024.100246. Available: https://www.sciencedirect.com/science/article/pii/S2667143324000738

[34] Y. Haruna, S. Qin, A. H. Adama Chukkol, A. A. Yusuf, I. Bello, and A. Lawan, “Exploring the synergies of hybrid convolutional neural network and Vision Transformer architectures for computer vision: A survey,” *Engineering Applications of Artificial Intelligence*, vol. 144, p. 110057, 2025, doi: https://doi.org/10.1016/j.engappai.2025.110057. Available: https://www.sciencedirect.com/science/article/pii/S0952197625000570

[35] A. Mumuni and F. Mumuni, “Data augmentation: A comprehensive survey of modern approaches,” *Array*, vol. 16, p. 100258, 2022, doi: https://doi.org/10.1016/j.array.2022.100258. Available: https://www.sciencedirect.com/science/article/pii/S2590005622000911

[36] M. Javaid, A. Haleem, R. P. Singh, and M. Ahmed, “Computer vision to enhance healthcare domain: An overview of features, implementation, and opportunities,” *Intelligent Pharmacy*, vol. 2, no. 6, pp. 792–803, 2024, doi: https://doi.org/10.1016/j.ipha.2024.05.007. Available: https://www.sciencedirect.com/science/article/pii/S2949866X24000662

[37] Z. Diao and H. Jiang, “A multi-instance tumor subtype classification method for small PET datasets using RA-DL attention module guided deep feature extraction with radiomics features,” *Computers in Biology and Medicine*, vol. 174, p. 108461, 2024, doi: https://doi.org/10.1016/j.compbiomed.2024.108461. Available: https://www.sciencedirect.com/science/article/pii/S0010482524005456

[38] M. Gu, F. Wang, T. Hu, and S. Yu, “Localization and Segmentation of Galaxy Morphology Based on Mask R-CNN,” 2023, pp. 512–515. doi: 10.1109/ICCEA58433.2023.10135337

[39] D. Demetriou, P. Mavromatidis, M. F. Petrou, and D. Nicolaides, “CODD: A benchmark dataset for the automated sorting of construction and demolition waste,” *Waste Management*, vol. 178, pp. 35–45, 2024, doi: https://doi.org/10.1016/j.wasman.2024.02.017. Available: https://www.sciencedirect.com/science/article/pii/S0956053X24001028

[40] D. A. Cremers and L. J. Radziemski, *Handbook of laser-induced breakdown spectroscopy*. John Wiley & Sons, 2013.

[41] G. S. Senesi, R. S. Harmon, and R. R. Hark, “Field-portable and handheld laser-induced breakdown spectroscopy: Historical review, current status and future prospects,” *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 175, p. 106013, 2021, doi: https://doi.org/10.1016/j.sab.2020.106013. Available: https://www.sciencedirect.com/science/article/pii/S0584854720304523

[42] J. Rakovský, P. Čermák, O. Musset, and P. Veis, “A review of the development of portable laser induced breakdown spectroscopy and its applications,” *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 101, pp. 269–287, 2014, doi: https://doi.

org/10.1016/j.sab.2014.09.015. Available: https://www.sciencedirect.com/science/article/pii/S0584854714002663

[43] C. Duan, J. Li, Y. Zhang, K. Ding, X. Geng, and Y. Guan, “Portable instruments for on-site analysis of environmental samples,” *TrAC Trends in Analytical Chemistry*, vol. 154, p. 116653, 2022, doi: https://doi.org/10.1016/j.trac.2022.116653. Available: https://www.sciencedirect.com/science/article/pii/S0165993622001364

[44] X. Fang, H. Yu, Q. Huang, Z. Jiang, D. Pan, and W. Gui, “Full-spectrum LIBS quantitative analysis based on heterogeneous ensemble learning model,” *Chemometrics and Intelligent Laboratory Systems*, vol. 257, p. 105321, 2025, doi: https://doi.org/10.1016/j.chemolab.2025.105321. Available: https://www.sciencedirect.com/science/article/pii/S0169743925000061

[45] Z. Gajarska *et al.*, “Automated detection of element-specific features in LIBS spectra,” *Journal of Analytical Atomic Spectrometry*, vol. 39, p. , 2024, doi: 10.1039/D4JA00247D

[46] C. Noël *et al.*, “Automated line identification for atomic spectroscopy (ALIAS): Application to LIBS imaging data processing,” *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 231, p. 107255, 2025, doi: https://doi.org/10.1016/j.sab.2025.107255. Available: https://www.sciencedirect.com/science/article/pii/S0584854725001405

[47] D. Girón, T. Delgado, J. Ruiz, L. Cabalín, and J. Laserna, “In-situ monitoring and characterization of airborne solid particles in the hostile environment of a steel industry using stand-off LIBS,” *Measurement*, vol. 115, pp. 1–10, 2018, doi: https://doi.org/10.1016/j.measurement.2017.09.046. Available: https://www.sciencedirect.com/science/article/pii/S0263224117306188

[48] R. Zhang, S. Hu, C. Ma, T. Zhang, and H. Li, “Laser-induced breakdown spectroscopy (LIBS) in biomedical analysis,” *TrAC Trends in Analytical Chemistry*, vol. 181, p. 117992, 2024, doi: https://doi.org/10.1016/j.trac.2024.117992. Available: https://www.sciencedirect.com/science/article/pii/S0165993624004758

[49] F. Chen *et al.*, “Remote identification of microorganisms on various substrates using LIBS and machine learning integration,” *Optics and Lasers in Engineering*, vol. 193, p. 109051, 2025, doi: https://doi.org/10.1016/j.optlaseng.2025.109051. Available: https://www.sciencedirect.com/science/article/pii/S0143816625002374

[50] I. Cherni, R. Nouir, F. Daoud, S. Hamzaoui, and H. Ghalila, “Fast diagnostic of osteoporosis based on hair analysis using LIBS technique,” *Medical Engineering & Physics*, vol. 103, p. 103798, 2022, doi: https://doi.org/10.1016/j.medengphy.2022.103798. Available: https://www.sciencedirect.com/science/article/pii/S1350453322000479

[51] P. Vaisakh, U. Adarsh, K. Amrutha, A. K. Warrier, V. Kartha, and V. Unnikrishnan, “Integrated LIBS-Raman spectroscopy: A comprehensive approach to monitor microplastics and heavy metal contamination in water resources,” *Environmental Research*, vol. 231, p. 116198, 2023, doi: https://doi.org/10.1016/j.envres.2023.116198. Available: https://www.sciencedirect.com/science/article/pii/S0013935123009994

[52] D. Capela *et al.*, “From waste to resource: LIBS methodology development for rapid quality assessment of recycled wood,” *Journal of Hazardous Materials*, vol. 489, p. 137493, 2025, doi: https://doi.org/10.1016/j.jhazmat.2025.137493. Available: https://www.sciencedirect.com/science/article/pii/S0304389425004054

[53] M. López Ochoa *et al.*, “Methodological approach based on LIBS imaging for the identification of microplastics in the environment. A case study for water samples,” *Microchemical Journal*, vol. 216, p. 114791, 2025, doi: https://doi.org/10.1016/j.microc.2025.114791. Available: https://www.sciencedirect.com/science/article/pii/S0026265X25021393

[54] R. H. El-Saeid, M. Abdelhamid, M. F. Ali, and M. Abdel-Harith, “LIBS utilization for the elemental analysis of black resin and gold used by ancient Egyptians in embalming,” *Journal of Cultural Heritage*, vol. 67, pp. 101–110, 2024, doi: https://doi.org/10.1016/j.culher.2024.02.009. Available: https://www.sciencedirect.com/science/article/pii/S1296207424000463

[55] V. Detalle and X. Bai, “The assets of laser-induced breakdown spectroscopy (LIBS) for the future of heritage science,” *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 191, p. 106407, 2022, doi: https://doi.org/10.1016/j.sab.2022.106407. Available: https://www.sciencedirect.com/science/article/pii/S0584854722000519

[56] M. P. Casado-Gavalda *et al.*, “Quantification of copper content with laser induced breakdown spectroscopy as a potential indicator of offal adulteration in beef,” *Talanta*, vol. 169, pp. 123–129, 2017, doi: https://doi.org/10.1016/j.

talanta.2017.03.071. Available: https://www.sciencedirect.com/science/article/pii/S0039914017303685

[57] G. Bilge, B. Sezer, K. Eseller, H. Berberoglu, H. Köksel, and I. Boyacı, “Determination of Ca addition to the wheat flour by using laser-induced breakdown spectroscopy (LIBS),” *European Food Research and Technology*, vol. 242, p. , 2016, doi: 10.1007/s00217-016-2668-2

[58] G. Bilge, H. M. Velioglu, B. Sezer, K. E. Eseller, and I. H. Boyaci, “Identification of meat species by using laser-induced breakdown spectroscopy,” *Meat Science*, vol. 119, pp. 118–122, 2016, doi: https://doi.org/10.1016/j.meatsci.2016.04.035. Available: https://www.sciencedirect.com/science/article/pii/S0309174016301346

[59] M. Sui, Y. Fan, L. Jiang, Y. Xue, J. Zhou, and S. Zhong, “Online ultrasonic nebulizer assisted laser induced breakdown spectroscopy (OUN-LIBS): An online metal elements sensor for marine water analysis,” *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 180, p. 106201, 2021, doi: https://doi.org/10.1016/j.sab.2021.106201. Available: https://www.sciencedirect.com/science/article/pii/S0584854721001580

[60] M. Sui *et al.*, “An underwater microimage long-pulse LIBS system: Development and preliminary testing,” *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 230, p. 107234, 2025, doi: https://doi.org/10.1016/j.sab.2025.107234. Available: https://www.sciencedirect.com/science/article/pii/S0584854725001193

[61] NASA, “Mars 2020 Perseverance Rover Componentes.” 2024.

[62] NASA, “We Persevere.” 2020.

[63] S. Schröder *et al.*, “A Laser-Induced Breakdown Spectroscopy (LIBS) Instrument for In-Situ Exploration with the DLR Lightweight Rover Unit (LRU),” *Applied Sciences*, vol. 14, no. 6, p. 2467, 2024, doi: 10.3390/app14062467. Available: https://app.dimensions.ai/details/publication/pub.1169816396

[64] H. Saeidfirozeh *et al.*, “Laser-induced breakdown spectroscopy in space applications: Review and prospects,” *TrAC Trends in Analytical Chemistry*, vol. 181, p. 117991, 2024, doi: https://doi.org/10.1016/j.trac.2024.117991. Available: https://www.sciencedirect.com/science/article/pii/S0165993624004746

[65] J. Ren, Y. Zhao, and K. Yu, “LIBS in agriculture: A review focusing on revealing nutritional and toxic elements in soil, water, and crops,” *Computers and Elec-*

tronics in Agriculture, vol. 197, p. 106986, 2022, doi: https://doi.org/10.1016/j.compag.2022.106986. Available: https://www.sciencedirect.com/science/article/pii/S0168169922003039

[66] G. Kim *et al.*, “Rapid detection of soils contaminated with heavy metals and oils by laser induced breakdown spectroscopy (LIBS),” *Journal of Hazardous Materials*, vol. 263, pp. 754–760, 2013, doi: https://doi.org/10.1016/j.jhazmat.2013.10.041. Available: https://www.sciencedirect.com/science/article/pii/S0304389413007838

[67] A. Ravichandran, R. Sarathi, and A. Haddad, “Pollution Monitoring on Polymeric Insulators Adopting Laser-Induced Breakdown Spectroscopy, Computer Vision, and Machine Learning Techniques,” *IEEE Transactions on Plasma Science*, pp. 1–9, 2025, doi: 10.1109/TPS.2025.3539261

[68] D. Penman, “Automated x-ray detection of contaminants in continuous food streams,” *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 2908, p. , 1996, doi: 10.1117/12.257259

[69] Y. Zhou, M. Daakir, E. Rupnik, and M. Pierrot-Deseilligny, “A two-step approach for the correction of rolling shutter distortion in UAV photogrammetry,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 160, pp. 51–66, 2020, doi: https://doi.org/10.1016/j.isprsjprs.2019.11.020. Available: https://www.sciencedirect.com/science/article/pii/S0924271619302849

[70] A. Kirillov, K. He, R. Girshick, C. Rother, and P. Dollár, “Panoptic Segmentation,” 2019. Available: https://arxiv.org/abs/1801.00868

[71] Y. Mo, Y. Wu, X. Yang, F. Liu, and Y. Liao, “Review the state-of-the-art technologies of semantic segmentation based on deep learning,” *Neurocomputing*, vol. 493, pp. 626–646, 2022, doi: https://doi.org/10.1016/j.neucom.2022.01.005. Available: https://www.sciencedirect.com/science/article/pii/S0925231222000054

[72] A. M. Hafiz and G. M. Bhat, “A survey on instance segmentation: state of the art,” *International Journal of Multimedia Information Retrieval*, vol. 9, no. 3, pp. 171–189, Jul. 2020, doi: 10.1007/s13735-020-00195-x. Available: http://dx.doi.org/10.1007/s13735-020-00195-x

[73] Roboflow, “Annotate - Roboflow,” 2025. Available: https://roboflow.com/annotate

[74] G. Jocher and J. Qiu, “Ultralytics YOLO11,” 2024. Available: https://github.com/ultralytics/ultralytics

[75] R. Khanam and M. Hussain, “YOLOv11: An Overview of the Key Architectural Enhancements,” 2024. Available: https://arxiv.org/abs/2410.17725

[76] P. Nandal, N. Bohra, P. Mann, and N. N. Das, “YOLOv11 with transformer attention for real-time monitoring of ships: A federated learning approach for maritime surveillance,” *Results in Engineering*, vol. 27, p. 106297, 2025, doi: https://doi.org/10.1016/j.rineng.2025.106297. Available: https://www.sciencedirect.com/science/article/pii/S2590123025023692

[77] W. Wei *et al.*, “YOLOv11-based multi-task learning for enhanced bone fracture detection and classification in X-ray images,” *Journal of Radiation Research and Applied Sciences*, vol. 18, no. 1, p. 101309, 2025, doi: https://doi.org/10.1016/j.jrras.2025.101309. Available: https://www.sciencedirect.com/science/article/pii/S1687850725000214

[78] V. D. Cong and L. H. Phuong, “Improving robotic grasping accuracy through oriented bounding box detection with YOLOv11-OBB,” *Heliyon*, vol. 11, no. 12, p. e43512, 2025, doi: https://doi.org/10.1016/j.heliyon.2025.e43512. Available: https://www.sciencedirect.com/science/article/pii/S2405844025018985

[79] R. Kaur and S. Singh, “A comprehensive review of object detection with deep learning,” *Digital Signal Processing*, vol. 132, p. 103812, 2023, doi: https://doi.org/10.1016/j.dsp.2022.103812. Available: https://www.sciencedirect.com/science/article/pii/S1051200422004298

[80] Ultralytics, “YOLOv8 - Performance Metrics,” 2023. Available: https://docs.ultralytics.com/pt/models/yolov8/#performance-metrics

[81] Ultralytics, “YOLO11 - Performance Metrics,” 2024. Available: https://docs.ultralytics.com/pt/models/yolo11/#performance-metrics

[82] Ultralytics, “YOLO12 - Performance Metrics,” 2025. Available: https://docs.ultralytics.com/pt/models/yolo12/#performance-metrics

[83] X. Xu *et al.*, “Spectral preprocessing combined with feature selection improve model robustness for plastics samples classification by LIBS,” *Frontiers in Environmental Science*, 2023, doi: 10.3389/fenvs.2023.1175392. Available:

https://www.frontiersin.org/journals/environmental-science/articles/10.3389/fenvs.2023.1175392

[84] J. Carvalho, “Multimodal System - Matrix Mode (Showcase Video),” 2025. Available: https://youtube.com/shorts/tocXoXAfNlc?feature=share

[85] J. Carvalho, “Multimodal System - Singular Mode (Showcase Video),” 2025. Available: https://youtube.com/shorts/LXfIsZOYQJM?feature=share

[86] J. Carvalho, “Multimodal System - Continuous Mode (Showcase Video),” 2025. Available: https://youtube.com/shorts/3-OTTKK5FI8?feature=share