

Leveraging Large Language Models and Robotics for Sustainable Construction Project Management

Kwame Mensah

Department of Artificial Intelligence and
Machine Learning, Institute of Intelligent Systems, Ghana

Ama Boateng

Department of Robotics and
Autonomous Systems, Centre for AI Innovation, Ghana

VOLUME 03 ISSUE 08 (2026)

Published Date: 15 August 2026 // Page no.: - 16-21

ABSTRACT

The increasing complexity of construction projects has intensified the need for intelligent systems capable of supporting planning, monitoring, decision-making, and operational control while simultaneously addressing sustainability objectives. Large Language Models (LLMs) and robotics represent two complementary technological capabilities: LLMs can support information interpretation, reasoning, communication, and decision assistance, whereas robotic systems can translate digitally generated decisions into physical construction activities. This review-oriented research paper develops a conceptual framework for integrating these technologies into sustainable construction project management. The methodological foundation is derived exclusively from the supplied literature, which primarily examines neural-network-based classification, image processing, and machine-learning approaches for medical-image analysis. Although these references do not directly investigate construction management, their treatment of automated classification, feature interpretation, image-based recognition, and machine-learning decision processes provides transferable methodological principles for intelligent construction monitoring. The proposed framework therefore positions LLMs as a cognitive coordination layer and robotics as an execution and sensing layer within a sustainability-oriented project-management architecture. The analysis identifies potential benefits in automated progress interpretation, resource coordination, anomaly identification, safety-oriented monitoring, and sustainability decision support. It also identifies critical limitations, including domain-transfer uncertainty, data quality, model reliability, explainability, interoperability, and the danger of treating domain-specific machine-learning evidence as direct validation for construction applications. The study contributes a theoretically grounded integration model while emphasizing that empirical construction datasets and field experiments are required before the framework can be considered operationally validated.

Keywords: Large Language Models, Construction Robotics, Sustainable Construction, Project Management, Artificial Intelligence, Machine Learning, Digital Decision Support, Automated Monitoring, Resource Optimization, Intelligent Construction.

INTRODUCTION

Construction project management increasingly operates within an environment characterized by interconnected schedules, heterogeneous information, changing site conditions, resource constraints, quality requirements, and sustainability expectations. Traditional management approaches depend heavily on human interpretation of documents, visual observations, progress reports, and operational data. While such approaches remain essential, their effectiveness can be constrained when project information becomes too large, dynamic, or complex to process manually. An intelligent project-management architecture can address this challenge by combining computational interpretation with physical automation.

Large Language Models provide a potential cognitive interface for such an architecture. Their relevance to construction management lies not simply in text generation but in their ability to organize heterogeneous information into interpretable decision-support outputs. Project specifications, work descriptions, inspection observations, schedules, risk reports, and sustainability requirements can theoretically be connected through a language-based reasoning layer. Robotics, by contrast, provides a physical layer capable of sensing site conditions and executing repetitive or precision-oriented tasks. The combination creates an opportunity to connect digital reasoning with physical construction activity.

However, the conceptual integration of LLMs and robotics requires a rigorous theoretical foundation. The supplied literature does not directly address LLMs, construction

robotics, or sustainable construction. Instead, it focuses on machine-learning-based recognition and classification of leukemia cells through neural networks and image-processing techniques. Ahmed et al. (2019), Thanh et al. (2018), Harjoko et al. (2018), Jagadev and Virani (2017), Lim et al. (2016), Singh et al. (2016), Fatma and Sharma (2014), and Vincent et al. (2015) collectively demonstrate how computational systems can transform complex visual information into classification-oriented outputs. This principle is relevant to construction because construction sites also generate large quantities of visual and operational information that require automated interpretation.

Accordingly, this paper does not claim that the supplied studies empirically validate LLMs or robotics for construction. Instead, it investigates how their machine-learning principles can theoretically inform an integrated construction-management framework. The central proposition of Leveraging Large Language Models and Robotics for Sustainable Construction Project Management is that intelligent construction systems should be designed as interconnected decision and execution architectures rather than as isolated artificial-intelligence applications.

The objectives are therefore threefold. First, the paper synthesizes the supplied literature to identify transferable principles of automated classification and intelligent image interpretation. Second, it develops a conceptual framework connecting LLM-based decision support with robotic sensing and execution. Third, it critically evaluates the framework in terms of sustainability, project-management performance, technical limitations, and future research requirements.

LITERATURE REVIEW

2.1 Machine Learning as a Foundation for Automated Recognition

The supplied literature demonstrates the evolution of computational approaches from conventional neural networks toward more sophisticated image-based learning systems. Fatma and Sharma (2014) investigated the identification and classification of acute leukemia using neural networks, demonstrating the application of computational learning to complex classification problems. Lim et al. (2016) similarly employed a multilayer perceptron network for classification of bone-marrow acute leukemia cells. These studies establish a fundamental principle: complex visual patterns can be converted into computationally interpretable categories through trained learning architectures.

This principle has methodological relevance to construction monitoring. A construction site contains

visually distinguishable conditions such as material placement, equipment positioning, incomplete work, completed work, and deviations from expected site configurations. Although the biological and construction domains are fundamentally different, the underlying computational problem of extracting meaningful patterns from complex visual information has conceptual similarities. Consequently, the literature supports the feasibility of automated pattern-recognition logic, but not the direct transfer of medical-image models to construction environments.

2.2 Image Processing and Classification

Jagadev and Virani (2017) examined leukemia detection and classification through image processing and machine learning, while Singh et al. (2016) reviewed approaches for detecting leukemia in medical images. These works reinforce the importance of preprocessing, feature extraction, classification, and interpretation as interconnected stages rather than isolated computational activities.

For construction management, the corresponding conceptual sequence can be represented as sensing, preprocessing, feature interpretation, classification, decision generation, and action. A robotic system may acquire site images or sensor observations; an analytical model may identify relevant patterns; an LLM may interpret those outputs alongside project requirements; and a project-management system may translate the resulting interpretation into an operational recommendation. The critical distinction is that construction environments are dynamic and less controlled than many laboratory imaging conditions. Therefore, environmental variation, occlusion, lighting, equipment movement, and incomplete observations can substantially influence model reliability.

2.3 Neural Networks and Progressive Automation

Harjoko et al. (2018) combined segmentation and neural-network-based classification for leukemia subtype recognition. Vincent et al. (2015) examined a two-step neural-network classifier for acute lymphoid leukemia. These studies suggest that complex recognition problems may benefit from staged or hierarchical computational architectures.

This observation provides a theoretical basis for a layered construction-intelligence model. Rather than requiring one model to perform sensing, interpretation, reasoning, and physical control simultaneously, responsibilities can be distributed across specialized layers. Robotics can perform physical sensing and execution; machine-learning components can analyze structured or visual observations; and LLMs can provide higher-level contextual interpretation and communication. Such separation can improve modularity and facilitate human oversight.

Thanh et al. (2018) further demonstrated the application of convolutional neural networks to leukemia blood-cell image

classification. Ahmed et al. (2019) investigated convolutional-neural-network-based identification of leukemia subtypes from microscopic images. Collectively, these studies demonstrate the increasing role of learned representations in complex image-classification problems.

2.4 Research Gap and Theoretical Positioning

The most significant gap in the supplied literature is domain specificity. None of the eight references directly examines construction project management, sustainability, LLMs, or construction robotics. Therefore, a direct evidence-based claim that these references prove the effectiveness of LLM-robotics integration in construction would be methodologically inappropriate.

Instead, the literature provides an indirect methodological foundation. It demonstrates the progression from neural-network classification toward automated visual interpretation. This foundation can inform a conceptual architecture in which intelligent systems transform raw observations into structured information and subsequently into management decisions. The research gap consequently lies in the absence of empirical investigation connecting automated perception with language-based reasoning and robotic execution in sustainable construction.

METHODOLOGY

3.1 Research Design

This study adopts a conceptual research-and-review methodology. Because the supplied references concern machine-learning-based image classification rather than construction, the study avoids presenting their findings as direct construction evidence. Instead, recurring computational principles are extracted and theoretically mapped onto construction project-management functions.

The methodology consists of four stages: literature abstraction, functional mapping, framework development, and critical evaluation. During literature abstraction, the principal computational functions represented by the references are identified. These include image acquisition, preprocessing, segmentation, feature recognition, classification, and neural-network-based decision processes. The second stage maps these functions to analogous construction-management activities. The third stage introduces LLMs and robotics as complementary architectural components. The final stage evaluates the resulting framework according to sustainability, operational value, reliability, and limitations.

3.2 Proposed LLM-Robotics Architecture

The proposed framework contains five interconnected layers.

Layer 1: Physical Sensing and Robotic Observation.

Robotic platforms, cameras, autonomous equipment, and other sensing devices constitute the physical observation layer. Their function is to collect information about construction conditions. Examples include visual observations of work areas, material locations, equipment movement, and task progress. The conceptual relevance of the supplied literature emerges here because Ahmed et al. (2019) and Thanh et al. (2018) demonstrate that image-based systems can convert visual information into machine-readable classifications.

Layer 2: Perception and Machine-Learning Analysis.

Raw observations require computational processing before they can support project decisions. Image preprocessing, segmentation, feature extraction, and classification can be used to identify relevant site conditions. The staged classification approaches described by Vincent et al. (2015) and Harjoko et al. (2018) support the conceptual argument that complex recognition tasks can be decomposed into multiple analytical stages.

Layer 3: LLM-Based Contextual Interpretation.

The LLM functions as a contextual reasoning and communication layer. Instead of treating a classification output as an isolated result, the system can theoretically connect it with project requirements, schedules, task descriptions, sustainability objectives, and operational constraints. For example, a robotic perception system could identify an apparent material-storage anomaly, while an LLM-based layer could organize the observation into a project-management recommendation for human review.

Layer 4: Decision Support and Coordination.

The interpreted information can be connected to project-management functions such as scheduling, resource coordination, quality monitoring, risk prioritization, and sustainability assessment. This layer should not operate as an unrestricted autonomous authority. Decisions involving safety, contractual obligations, major resource changes, or regulatory compliance should remain subject to appropriate human verification.

Layer 5: Robotic Execution and Feedback.

Where automation is technically and operationally appropriate, decisions can be translated into robotic actions. The resulting physical state is then sensed again, creating a feedback loop. This produces an iterative cycle:

Sense → Analyze → Interpret → Decide → Act → Re-sense.

Such a cycle is theoretically more suitable for dynamic

construction environments than a one-directional automation model.

3.3 Sustainability Integration

Sustainability is positioned as a decision criterion rather than an independent technological feature. A construction-management system should evaluate whether proposed actions reduce material waste, unnecessary equipment operation, excessive transportation, idle time, or inefficient resource allocation.

For example, if robotic monitoring identifies that materials have accumulated in a location that obstructs workflow, the LLM-based coordination layer could interpret the observation in relation to planned activities and recommend relocation. The sustainability value would arise not from the LLM itself but from the reduction of avoidable movement and improved utilization of resources.

This distinction is important because artificial intelligence does not automatically produce sustainable outcomes. A technically efficient recommendation may still increase energy consumption, equipment utilization, or material waste. Sustainability therefore needs to be encoded as an explicit decision criterion.

3.4 Human-in-the-Loop Governance

The proposed architecture retains human supervision because machine-learning outputs are probabilistic and context-dependent. The supplied literature demonstrates the value of automated classification, but classification accuracy alone does not establish managerial correctness. A construction system must account for contractual requirements, site-specific constraints, worker coordination, and contextual information that may not be represented in sensor data.

Human oversight should therefore operate at critical decision points. The LLM can explain detected conditions, organize alternatives, and provide recommendations, while project managers retain authority over consequential decisions. This arrangement reduces the risk of interpreting automated outputs as unquestionable facts.

3.5 Hypothetical Operational Example

Consider a construction project in which an autonomous inspection robot periodically surveys a work zone. Its perception system identifies that a designated material area differs from the expected configuration. A classification model assigns the observation to an anomaly category. An LLM then receives the structured observation together with the relevant work-plan

description and generates a concise management recommendation. The project manager reviews the recommendation and determines whether relocation or another corrective action is appropriate. If approved, a robotic handling system may execute the physical adjustment, after which the inspection robot reassesses the area.

The example illustrates why the proposed framework is not merely an LLM application or robotics application. The value emerges from integration. Perception provides evidence, language-based reasoning provides contextual interpretation, human management provides authorization, and robotics provides physical execution.

4. RESULTS

The conceptual analysis produces four principal findings. First, machine-learning-based perception represents a foundational component of intelligent construction management. The supplied studies consistently demonstrate the transformation of complex visual information into classification outputs through neural-network and image-processing techniques (Fatma and Sharma, 2014; Lim et al., 2016; Jagadev and Virani, 2017). This supports the theoretical feasibility of automated perception, although construction-specific validation remains necessary.

Second, a layered architecture is more defensible than a monolithic AI model. The two-step classification approach discussed by Vincent et al. (2015) and the segmentation-plus-classification methodology of Harjoko et al. (2018) illustrate the value of dividing complex recognition processes into stages. In construction, similar modularity can separate sensing, perception, contextual interpretation, decision support, and physical execution.

Third, LLMs and robotics offer complementary rather than interchangeable capabilities. Robotics provides physical interaction and site-level sensing, whereas LLMs can organize complex information and support human communication and decision-making. The conceptual integration of these capabilities can create a closed operational feedback loop in which physical observations inform digital reasoning and approved decisions influence physical operations.

Fourth, the sustainability contribution is conditional on decision criteria. Neither automated classification nor language-based reasoning is inherently sustainable. Sustainability improvements occur only when resource consumption, waste, energy use, transportation, and operational efficiency are incorporated into decision processes. Therefore, the framework should evaluate technological decisions according to both operational and environmental criteria.

The analysis also identifies a major evidence limitation. The provided literature demonstrates machine-learning performance in medical-image classification, but it does not

establish performance in construction environments. Consequently, the proposed framework should be considered a theoretically derived research model rather than an empirically validated construction-management system.

5. DISCUSSION

The proposed framework contributes to the theoretical understanding of intelligent construction management by treating AI and robotics as components of a coordinated socio-technical system. Its principal theoretical implication is that perception, reasoning, and action should be architecturally separated while remaining connected through feedback. This proposition is consistent with the staged computational logic observable across the supplied literature, particularly in studies involving neural-network classification and image processing (Vincent et al., 2015; Harjoko et al., 2018).

From a practical perspective, the framework could support continuous monitoring rather than periodic manual inspection. Automated visual recognition may assist in identifying deviations, while an LLM-based layer could convert structured observations into project-management language. Such integration could potentially reduce information-processing delays and improve coordination between technical systems and human decision-makers. However, these benefits depend strongly on the quality and reliability of the underlying sensing and classification systems.

A central trade-off concerns automation versus accountability. Increasing automation may improve response speed, but it can also amplify erroneous classifications. Ahmed et al. (2019) and Thanh et al. (2018) demonstrate the use of convolutional neural networks for complex image classification, yet construction sites differ considerably from controlled medical-image contexts. Site images can contain changing illumination, dust, occlusion, moving equipment, incomplete structures, and irregular objects. Consequently, models trained in other domains cannot simply be assumed to generalize.

Another issue concerns interpretability. A project manager needs more than a classification label; the manager needs to understand what was detected, how confident the system is, why the observation matters, and what alternatives exist. The LLM layer may improve communication, but it does not automatically guarantee factual correctness. Therefore, language-based recommendations should remain connected to verifiable project data and clearly distinguish observations from generated interpretations.

The literature also reveals the importance of model

architecture. Singh et al. (2016) reviewed computational approaches to medical-image detection, while Jagadev and Virani (2017) connected image processing with machine learning. Their methodological progression suggests that successful intelligent systems depend on multiple preprocessing and analytical stages. Construction applications should therefore avoid conceptualizing an LLM as a replacement for perception systems. An LLM cannot compensate for poor sensing or unreliable underlying data.

The framework further has implications for sustainability research. Automated monitoring can theoretically identify inefficient resource use, while robotic systems may reduce unnecessary manual repetition in suitable tasks. Nevertheless, robotics itself consumes energy and requires equipment, maintenance, and digital infrastructure. Sustainable construction management should therefore evaluate net environmental performance rather than assuming that greater automation necessarily means greater sustainability.

The most significant limitation remains the mismatch between the supplied evidence base and the proposed application domain. The references concern leukemia detection and classification, not construction. Their value is methodological and theoretical rather than empirical. This limitation prevents claims of demonstrated construction performance and establishes a clear research agenda: future studies must develop construction-specific datasets, benchmark LLM-supported decision processes, evaluate robotic execution, measure sustainability outcomes, and test the proposed architecture under realistic site conditions.

6. CONCLUSION

This paper developed a conceptual framework for leveraging Large Language Models and robotics in sustainable construction project management. The framework is based on a critical synthesis of the supplied machine-learning literature, which demonstrates the use of neural networks, image processing, segmentation, and classification for complex visual-recognition problems (Ahmed et al., 2019; Thanh et al., 2018). Although those studies originate outside construction, their methodological principles provide a useful foundation for conceptualizing automated perception and staged intelligent processing.

The proposed architecture integrates robotic sensing, machine-learning perception, LLM-based contextual interpretation, decision support, human governance, and robotic execution. Its central operational logic is a continuous sense-analyze-interpret-decide-act-re-sense cycle. Such integration can potentially improve construction monitoring, resource coordination, operational responsiveness, and sustainability-oriented decision-making.

The paper also emphasizes that AI integration should not be equated with automatic sustainability or autonomous

decision-making. Sustainability must be explicitly incorporated into management objectives, while consequential decisions should remain subject to human oversight. Likewise, machine-learning results obtained in medical-image domains cannot be treated as direct evidence of construction performance.

Future research should therefore focus on construction-specific datasets and field experiments. Empirical studies should evaluate perception accuracy, LLM recommendation reliability, robotic execution performance, human acceptance, interoperability, and measurable sustainability outcomes. Comparative experiments should also determine whether integrated systems provide meaningful advantages over conventional project-management workflows. Ultimately, the contribution of intelligent construction technology will depend not merely on the sophistication of individual AI models but on the reliability, transparency, interoperability, and sustainability of the complete socio-technical system.

REFERENCES

1. N. Ahmed, A. Yigit, Z. Isik and A. Alpkocak, "Identification of leukemia subtypes from microscopic images using convolutional neural network," *Diagnostics*, vol. 9, no. 3, pp. 1–11, 2019.
2. M. Fatma and J. Sharma, "Identification and classification of acute leukemia using neural network," in *Proc. MedCom*, Greater Noida, India, pp. 142–145, 2014.
3. Harjoko, T. Ratnaningsih, E. Suryani, S. Palgunadi and N. P. T. Prakisy, "Classification of acute myeloid leukemia subtypes M1, M2 and M3 using active contour without edge segmentation and momentum backpropagation artificial neural network," in *Proc. MATEC Web of Conf.*, Yogyakarta, Indonesia, pp. 01041–01048, 2018.
4. P. Jagadev and H. G. Virani, "Detection of leukemia and its types using image processing and machine learning," in *Proc. ICEI*, Tirunelveli, India, pp. 522–526, 2017.
5. H. N. Lim, E. U. Francis, M. Y. Mashor and R. Hassan, "Classification of bone marrow acute leukemia cells using multilayer perceptron network," in *Proc. ICED*, Phuket, Thailand, pp. 486–490, 2016.
6. G. Singh, G. Bathla and S. Kaur, "A review to detect leukemia cancer in medical images," in *Proc. ICCCA*, Greater Noida, India, pp. 1043–1047, 2016.
7. T. T. P. Thanh, C. Vununu, S. Atoev, S. H. Lee and K. R. Kwon, "Leukemia blood cell image classification

using convolutional neural network," *International Journal of Computer Theory and Engineering*, vol. 10, no. 2, pp. 54–58, 2018.

8. Vincent, K. R. Kwon, S. H. Lee and K. S. Moon, "Acute lymphoid leukemia classification using two-step neural network classifier," in *Proc. FCV*, Mokpo, Korea, pp. 1–4, 2015.
9. Ramamurthy, K. (2023). *AI-Driven Test Automation Frameworks for the Modern Software Quality Engineering*. *International Journal of Emerging Trends in Computer Science and Information Technology*, 4(4), 257-269.
10. Geo Philip, Paulson, *Robotics-Enabled Sustainable Construction Management: A Socio-Technical Framework for Operational Efficiency, Digital Integration, and Sustainability Performance*. Available at SSRN: <https://ssrn.com/abstract=6845167> or <http://dx.doi.org/10.2139/ssrn.6845167>