DS Journal of Multidisciplinary
Volume 2 Issue 3, 1-14, Jul - Sep 2025
ISSN: 3048-4561/ https://doi.org/10.59232/DSM-V2I3P101

*Original Article*

# LC-IAF: Low-Code AI Framework for Image Analysis in the Context of Digital Learning Materials

**Ngan-Giang Ma[1], Kim-Quoc Nguyen[2], Son-Hai Tran[1*]**

[1]*Department of Information Technology, Ho Chi Minh University of Education (HCMUE), Ho Chi Minh City, Vietnam.*
[2]*Department of Information Technology, Nguyen Tat Thanh University (NTTU), Ho Chi Minh City, Vietnam.*

*haits@hcmue.edu.vn

Received: 06 June 2025; Revised: 05 July 2025; Accepted: 02 August 2025; Published: 15 August 2025

**Abstract -** This paper proposes a Low-Code Image Analysis Framework (LC-IAF) to address teachers' rules when analyzing images in learning materials. This work is time-consuming and subjective, while AI is too complex. LC-IAF bridges the gap between AI's image analysis capabilities and pedagogical needs. This framework analyzes deep learning and computer vision to extract attributes of images in digital learning materials, providing a low-code interface that allows teachers without programming expertise to analyze and develop image content in teaching materials. Based on visual learning theory, LC-IAF provides actionable insights, supports design guidance, personalized learning, and automated assessment, empowering teachers to improve visual teaching materials.

**Keywords -** Image analysis, Low-code, Educational technology, Digital learning materials, Personalized learning.

## 1. Introduction

Nowadays, diagrams, tables, and illustrations are increasingly popular and play an important role in digital learning materials (Foutsitzi, 2018). These visual elements help learners enhance their ability to understand lesson content and increase their retention, as previous studies have shown that converting abstract ideas into tangible representations through images increases learners' level of interaction with the lesson content (Trigka, M., & Dritsas, E., 2025). Therefore, exploiting and effectively analysing these visual elements is necessary to optimize the learning experience, improve the teaching process, and enhance learners' ability to absorb knowledge. Although Artificial Intelligence (AI) has powerful image analysis capabilities, existing tools and programming languages are often complex and require programming expertise (Trigka, M., & Dritsas, E., 2025). These have created a barrier for teachers who have innovative ideas but cannot implement them into separate software or support systems. Additionally, teachers currently do so manually when analyzing images for teaching materials. However, with many images, this is time-consuming and subjective when evaluating the images for pedagogical purposes, leading to inconsistencies and scalability issues (Johnson, A., Gabriel, O. K., Kazeem, A. O., & Kolawole, A. A., 2024), including the assessment of image quality, relevance, and potential cognitive load (Trigka, M., & Dritsas, E., 2025) (Abdulsahib, A. K., Mohammed, R., Ahmed, A. L., & Jaber, M. M., 2024). Problems that exist in converting rich visual data into pedagogical insights can hinder teachers from effectively applying research-based principles while also limiting learners' personalized and adaptive learning experiences (Abdrakhmanov, R., Tuimebayev, A., Zhussipbek, B., Utebayev, K., Nakhipova, V., Alchinbayeva, O., ... & Kazhybayev, O., 2024).

To address the issues – particularly the gap between AI's powerful image analysis capabilities and teachers' practical pedagogical needs, coupled with the complexity of existing tools and the subjectivity of manual analysis – this paper proposes a novel Low-Code Image Analysis Framework (LC-IAF) for digital learning materials. LC-IAF is specifically designed for the educational analysis of images in learning content, offering a groundbreaking solution that significantly differentiates it from current methods.

The novelty of LC-IAF lies in its ability to bridge the gap between AI's image analysis capabilities and real pedagogical needs. Unlike traditional computer vision tools such as OpenCV, which demand high programming proficiency and in-depth technical understanding, or commercial AI APIs that primarily focus on enterprise data processing with limited pedagogical understanding, LC-IAF is built with a low-code interface. This allows non-programmer teachers to create, modify, and run image analysis workflows efficiently. Contrary to current literature in computer vision and education, which mainly focuses on technical characteristics with a less immediate correspondence to instructional design principles, LC-IAF is informed by established concepts of visual learning theories, such as Paivio's Dual-Coding Theory, Sweller's Cognitive Load Theory, and Mayer's principles of multimedia learning. This integration allows LC-IAF to translate technical AI outputs into pedagogically meaningful insights.

This represents a core differentiation from previous works that primarily addressed the technical capabilities of CV in education without fully overcoming the expertise barrier or lacking direct pedagogical applicability. The key contributions of this work, highlighting its novelty and superiority over existing research, include:

- A novel LC-IAF framework for analyzing learning materials, with a particular emphasis on teacher and student usability – a factor often overlooked in complex AI solutions. This framework democratizes the use of AI in education, overcoming the limitations of high programming skill requirements associated with current tools.
- Extensive explanation of the tight integration of deep learning and state-of-the-art computer vision methodologies (e.g., semantic segmentation, object detection and recognition, and image classification) for the extraction of pedagogically relevant attributes from images. Unlike traditional CV tools, this is more than just treating content recognition problems, but enables fine-grained pedagogical analytics, essentially to tackle the extraneous cognitive load problem
- Demonstrate a novel low-code interface design that enables educators who are not programmers to easily configure and apply image analysis workflows without the need for programming. This is especially new in education – low-code AI platforms are so misunderstood. LC-IAF makes AI become a supplementary and easy-to-use tool for users.
- Introducing a low-code UI that allows users with no programming experience to easily create, modify, and deploy image analysis pipelines.
- An illustration of how LC-IAF facilitates advanced instructional design, personalized learning, and automated assessment by providing actionable insights from image content. This framework not only determines "what" is in an image but also "how" its visual attributes affect learning, offering a level of pedagogical understanding not provided by other technical image analysis tools.
- A sketch of various future research steps, such as multimodal document classification, integration with Open Educational Resources (OER), and the integration of Explainable AI (XAI) for pedagogical insights. Serving as guidance, these directions expand the potential of LC-IAF to a broader educational content intelligence system.

This paper is organized into five primary sections: The theoretical underpinnings and background are presented in Section 2. The architecture and fundamental methods of the suggested LC-IAF framework are explained in Section 3. The pedagogical implications and assessment techniques are covered in Section 4. Section 5 concludes and makes recommendations for future research challenges and directions.

## 2. Overview

To develop a practical and effective visual analysis framework for learning materials, a comprehensive understanding of existing works in the field of computer vision, pedagogical theories of visual learning, and the advancements of low-code/no-code platforms in educational technology is essential. This section provides a detailed overview of these areas, establishing the necessary context and identifying the research gap that this paper's proposal aims to address.

### *2.1. Computer Vision in Education and Current Limitations*

Computer Vision (CV) has been bringing about significant transformations across various domains, and education is no exception. CV applications in education range from operational support tasks like campus security monitoring and intelligent parking management to more direct applications aimed at enhancing the learning and teaching experience. Specifically, CV has been utilized to assess learners' emotional levels through facial expressions (Mothwa, 2018), evaluate learner interaction levels through eye gaze (Soares Jr, R. D. S., Pinheiro, E. D., Oku, A. Y. A., Rizzo, M. B., Vieira, C. D. N., & Sato, J. R., 2024), and support the creation of multimedia educational content (Abdulsahib, A. K., Mohammed, R., Ahmed, A. L., & Jaber, M. M., 2024).

Research suggests that computer vision has the capacity to observe, analyze, and enhance student engagement and interaction with digital content, thereby bridging the gap between static learning materials and the dynamic nature of learner involvement for instance, Real-time Automated STEM Engagement Detection Systems (RASEDS) have been developed to accurately distinguish student engagement levels using advanced object detection methods, particularly the YOLOR method (Abdrakhmanov, R., Tuimebayev, A., Zhussipbek, B., Utebayev, K., Nakhipova, V., Alchinbayeva, O., ... & Kazhybayev, O., 2024). Similarly, automated attendance systems employing facial recognition have achieved an average accuracy of 98%, saving time and improving classroom management effectiveness (Mothwa, 2018). Through image analysis, technology has also been used to identify student engagement in a classroom; research has shown that it can reliably determine whether a student is engaged or not (Pillai, 2022).

Nevertheless, a major drawback of the CV tools available today is that they frequently call for a high level of programming proficiency and in-depth technical understanding. These factors pose a serious obstacle for educators who lack the technical know-how to use or modify these tools for their own needs directly. It has been demonstrated that problems like a lack of domain knowledge, a lack of computational power, and data limitations can result in poor model architecture selections (Zviel-Girshin, 2024). Additionally, a lot of conventional CV tools concentrate on technical features like object detection and classification, with little connection to instructional design principles or the capacity to analyze images from an educational standpoint. Computer vision education studies have also shown that the lack of appropriate learning context materials and the separation between theoretical and practical courses reduce teaching effectiveness (Raiyn, 2016).

Despite these advancements, a significant drawback of currently available CV tools is that they frequently demand a high level of programming proficiency and in-depth technical understanding. This poses a serious obstacle for educators who lack the technical know-how to use or modify these tools for their specific needs directly. Problems such as a lack of domain knowledge, insufficient computational power, and data limitations can often result in poor model architecture selections, diminishing the effectiveness of AI applications. Furthermore, many conventional CV tools primarily focus on technical features like object detection and classification, with limited connection to instructional design principles or the capacity to analyze images from an educational standpoint. Computer vision education studies have also indicated that the lack of appropriate learning context materials and the separation between theoretical and practical courses reduce teaching effectiveness. This gap, between the powerful technical capabilities of AI and the practical pedagogical needs of teachers, represents the core problem this paper aims to address.

### *2.2. Pedagogical Theories of Visual Learning*

This study's framework is built upon established visual learning theories to ensure that image analysis is not merely a technical exercise but possesses profound pedagogical significance. A thorough understanding of these theories is crucial for translating technical image data into valuable educational insights.

Paivio's Dual-Coding Theory (Paivio, 2006): According to this theory, the human mind uses different channels to process verbal and non-verbal (visual) information and combining the two improves memory and comprehension. According to Paivio, two cognitive subsystems are one specialized in representing and processing non-verbal events or objects, such as images, and the other specialized in processing language.

Three types of processing are also identified by the theory:

- Representational processing, which activates linguistic or nonlinguistic representations directly.
- Referential processing activates the linguistic system through the nonlinguistic system or vice versa.
- Associative processing, which activates representations within the same linguistic or nonlinguistic system.

The use of imagery in this study gave students another way to retrieve information from memory, which improved their ability to do so. Specific images for shapes, sounds in the environment, actions, and the aural or visceral sensations connected to emotions are examples of nonlinguistic representations. Nonlinguistic representations have the ability to encode information simultaneously or in parallel, in contrast to sequential language processing.

Sweller's Cognitive Load Theory (CLT) (Paas, F., Renkl, A., & Sweller, J. , 2003): CLT distinguishes three types of cognitive load, including Intrinsic Cognitive Load (ICL), which is the inherent difficulty of the material; Extraneous Cognitive Load (ECL), which arises from poor instructional design; and Germane Cognitive Load (GCL), which is related to schema formation and deep learning. Effective instructional design aims to minimize ECL and optimize GCL through the use of dual processing. This technique typically involves presenting learners with an image that is processed by their visual-spatial sketchpad and talking about the image, which is processed by their phonological loop (Asma, H., & Dallel, S., 2021).

This study shows the importance of cognitive capacity in working memory for good learning results. Working memory has a limited capacity, so students need to be able to handle and connect a lot of different mental parts at the same time in order to understand information. These mental demands often overload working memory, so effective learning processes must work within its natural limits. Building upon the Cognitive Load Theory (CLT), (Mayer, 2005) has formulated the major tenets of effective multimedia design.

Stated in another manner, the Mayer model is based on three core assumptions: people process information via a dual-channel (auditory and visual) system (dual-channel assumption), people have limited capacity to process incoming information (limited capacity assumption), and people must be actively processing during learning (active processing assumption). Principles included the Multimedia Principle (using words and pictures instead of words alone), the Coherence Principle (excluding extraneous information), the Redundancy Principle (not presenting visually narrated information simultaneously in visual text), and the Temporal Contiguity Principle (presenting words and images concurrently).

The ultimate goal of this theory is to optimize the efficacy of cognitive processes in learning. Building an image analysis framework based on these pedagogical theories provides a solid scientific basis, enabling AI to analyse images technically and assess their pedagogical effectiveness. This ensures that the insights generated can be directly linked to researched learning principles, enhancing the validity and explainability of the system.

### *2.3. Low-Code/No-Code Platforms and Multimodal Learning Analytics*

The increasing adoption of Low-Code/No-Code (LC-NC) platforms is a recent and notable phenomenon in the IT landscape. These platforms empower non-technical users to create tailored applications to address specific concerns without the need for programming from scratch (Kandaurova, M., Skog, D. A., & Bosch-Sijtsema, P. M., 2024). With pre-made components and intuitive, easy-to-use interfaces, LC-NC makes it possible for non-developers to build valuable applications, fostering agility and creativity.

In the field of educational technology, LC-NC offers powerful benefits, including fast prototyping and deployment of personalized learning portals and adaptive learning experiences, while preserving the authority of educators and administrators to craft and direct their digital assets. Research shows that the use of low-code and no-code platforms can make developers more efficient by enabling agile approaches and reducing coding time, thereby slashing time to market. While misconceptions may limit the full potential of AI and low-code in education, the ability to democratize the use of AI-powered software tools could have a revolutionary impact on the creation of AI applications.

Meanwhile, Multimodal Learning Analytics (MMLA) has become a fruitful research field involving the interrelation of different data sources to provide a holistic view of learning behaviour and outcomes. According to Mangaroska et al. (Mangaroska, K., Sharma, K., Gaševic, D., & Giannakos, M., 2020), (VP, 2025), MMLA gathers and deals with a variety of such data forms, such as text, images, audio or even video. Another definition is that MMLA is a family of methods through which multichannel data, compiled from participating actors ‘interactions via video, activity logs, audio, and gestures, can be recorded at high frequency.

Its main purpose is to offer insight and to provide prompt feedback on learning in progress, whether learners are attending an online or face-to-face course. Studies have also revealed that learning to programme is a complex learning activity during which several cognitive processes and affective states work together. In this regard, Multimodal Learning Analytics (MMLA) is helpful in gathering data to measure psychological constructs that affect learning, such as cognitive load and confusion (Mangaroska, K., Sharma, K., Gaševic, D., & Giannakos, M., 2020).

Document Understanding Systems (DUS) are an emerging class of technologies that utilize the combination of Computer Vision, Natural Language Processing (NLP), and Deep Learning to classify, extract and structure data from multiple document types. In this context, the existing visual analytics system provides a fundamental basis for the development of a complete Educational Content Intelligence system.

This analytical model can be extended to describe how visual elements interact with text, layout, and other modalities of information in learning materials, not only by analyzing images isolated from text. This can be done both using MMLA and by itself, in the field of document understanding, to provide the learner with a deep semantic comprehension of educational documents. Several approaches have been suggested on MMLA system design, including the incorporation of technology, limitations and learning scenarios.

The Multimodal Learning Analytics Design Framework (MDF) has been developed in order to provide a systematic principle to MMLA design, emphasizing the significance of design at early stages to effectively and user-friendly MMLA systems. In light of the MMLA system design, the combination of technology considerations, constraints, and learning scenarios that have been offered by researchers (H. Ouhaichi, V. Bahtijar, and D. Spikol, 2024). To help the research community in the development of MMLA systems, the Multimodal Learning Analytics Design Framework (MDF) provides systematic guidance on the development process of MMLA. To ensure that the MMLA systems are efficient and user-friendly, design should be added during the development of an MMLA system.

The current visual analytics approach is an initial step toward a holistic educational content intelligence system. Leveraging the spirit of MMLA and the principles of document understanding, this framework could be extended to analyze the images alone and in the manner they are combined with text, layout and other modalities in learning materials to let users understand the semantics of educational materials.

## 3. Methods

The architecture, fundamental methods, and low-code design process of the suggested Low-Code Image Analysis Framework (LC-IAF) are described in detail in this section. LC-IAF seeks to close the gap between teachers' practical pedagogical needs and artificial intelligence's potent image analysis capabilities. This enables teachers to utilise visual content in instructional materials fully.

### *3.1. Framework Architecture and Data Flow*

LC-IAF is developed based on a modular architecture, ensuring outstanding flexibility and scalability. The main components of the framework and the data flow are meticulously designed to transform raw image data into highly applicable pedagogical insights, including (See Figure 1):

- Image Ingestion Module: The system begins with the process of receiving learning materials from the teacher or learner. This module supports a variety of input formats, including but not limited to scanned textbook pages, digital images, and PDF documents.
- Preprocessing Module: After the ingestion stage, the images are put into the preprocessing process. This process includes normalization, resizing, color channel conversion, and noise reduction steps, which aim to optimize image quality for subsequent AI analysis.
- Core AI Analysis Engine: This is the central component of the framework, integrating specialized Computer Vision and Deep Learning models. Its main function is to extract raw image features, such as edges, textures, objects, and semantic regions.
- Pedagogical Analysis Module: This is the core component of LC-IAF, which is responsible for interpreting raw outputs from AI models into pedagogically meaningful insights. This module converts technical findings, such as detected objects, into educational solutions (e.g., "the diagram contains five complex elements that are likely to cause high intrinsic load"). This pedagogical perspective is what turns a purely technical tool into an effective teaching aid.
- Low-Code Interface Layer (User Interface Layer): A user-oriented component that serves as the entry point for a teacher to use the system. Specifically, it provides a way to set up an analysis without programming knowledge.
- Output and Recommendation Module: Finally, the system outputs insights and recommendations. These insights are surfaced through a low-code interface, which enables teachers to review and act on their data easily.
- The flow of data is meant to produce an unbroken feedback loop for education reform. Educators apply the tool, obtain insights, revise the document, and reanalyze to create an iterative refinement of instructional design akin to adaptive learning principles.

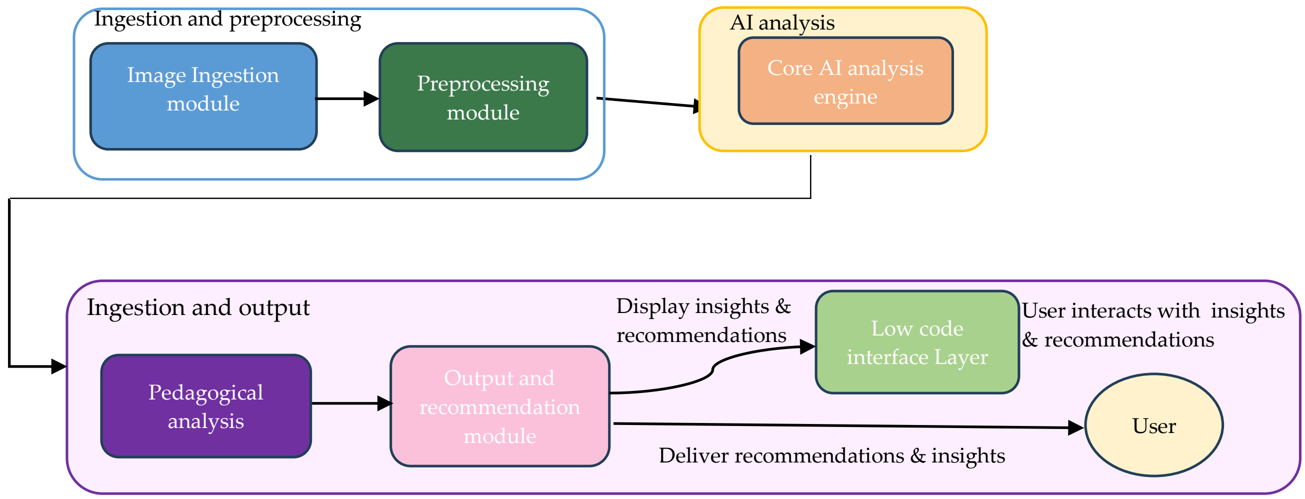


**Fig. 1 Educational image analysis system data flow**

### *3.2. Core Image Analysis Techniques for Educational Purposes*

This methodology utilizes advanced, pre-processed, and customized image processing techniques to extract features that are directly associated with the pedagogical effectiveness of images:

- Semantic Segmentation (Long, J., Shelhamer, E., & Darrell, T., 2015): Semantic segmentation aims to classify every pixel in the image into classes, which can be diagrams, photographs, text blocks, or decorative elements. The aim is to understand in detail what the image is made up of. This approach is particularly useful for the differentiation of the learning and decorative or distracting aspects and, as such, provides support on the basis of Mayer's principle of contiguity (Mayer, 2005).
- Object Detection and Recognition (Girshick, R., Donahue, J., Darrell, T., & Malik, J. , 2014): The framework estimates the relevance of the image with respect to the text content through the detection and localization of certain objects or concepts in the image.
- Image Classification (Krizhevsky, A., Sutskever, I., & Hinton, G. E., 207): This categorizes images according to their overall content or style (e.g., line-charts, bar-charts, pictorial illustrations, real-life photos, abstract images). This data may help guide decisions about when and how to use images consistent with Mayer's pedagogical principles to select the best type of image to communicate a concept (Mayer, 2005).
- Image feature extraction for educationally relevant attributes such as:
- Complexity appraisal: Examine image density, elements, and relations to estimate intrinsic cognitive load (ICL). This enables teachers to identify images that could be overwhelming for the learners.
- Text relevance: Models exist to measure the relevance of textual descriptions of images and discover problems of inconsistency and redundancy. This is the evidence for Mayer's redundancy and coherence principles.
- Layout component: Determine the spatial layout of the elements in the image (i.e., text boxes, arrows, labels, etc.) by checking their relevancy. This is consistent with the Spatial Coherence Principle (Mayer, 2005).

Its framework does not just determine "what" there is in an image, but also "how" its visual attributes affect learning. Applying those techniques together has enabled the framework to carry out a fine-grained pedagogical analysis of images even beyond that of simple content recognition, in an attempt to identify at the root level the instructional design characteristics of the images, thereby directly contributing to the extraneous cognitive load problem.

### *3.3. Low-Code Integrated Design*

In order to be practical and widely used by teachers, LC-IAF is implemented with a low-code approach where the complex AI capabilities are all designed to be intuitive tools for users. In particular, the following aspects of this methodology are illustrated as follows:

- Drag-and-Drop interface: No programming is required to build a custom analysis workflow of pre-built AI components, and it allows a teacher to develop a new idea or workflow visually.
- Configuration Panels: The system is equipped with customizable configuration interfaces that enable users to define their parameters without directly manipulating the source code.
- Ready Templates and Recipes: The platform includes pre-built templates for everyday learning tasks that streamline the creation and analysis of content.
- Live Feedback and Visualization: Analytics and pedagogical insights are delivered visually comprehensibly, allowing teachers to understand quickly and act faster.
- Integration with Current EdTech Ecosystems: The extent to which the item is compatible with Learning Management Systems (LMS) and other popular educational content creation tools is an important consideration for item publishers in order to maximize the seamless integration in teachers' existing practices.

The low-code architecture of LC-IAF is more than a technical aspect; it is a strategic choice to build users' access to cutting-edge AI. By hiding the complexity of coding, this design leads teachers to develop supportive tools and to make a proactive change in and re-creation of the teaching materials employing AI. This is important in fostering a perception of active technology use rather than passive submission to technology systems, which is crucial for the long-term success of embedding AI in classrooms.

As a demonstration of the low-code simplicity of LC-IAF, this paper introduces a practical microflow developed on the Mendix platform, which applies cheatsheet classification of educational resources using the Gemini Pro Vision API. This microflow (GeminiPro_Cheatsheet_Classification) is a great example of how a teacher or non-technical user can create a complex visual annotation workflow visually (see Figure 2).

#### *3.3.1. Flow Description*

- Encoding and API Call: The microflow starts by encoding an image file (cheatsheet) to Base64 format. It then calls a sub-microflow (SUB_GetProgrammingLanguageFromCheatsheet) to prepare the Gemini Pro Vision API request, including changing the Image property to ReqPrompt to send the image data. A REST call (POST) to the Gemini Pro Vision API is made to parse the image.
- Response Processing and Classification: The response from the API is logged. If the response is empty, the system logs an error and deletes the image. Otherwise, if there is a response, an action (ACT_AfterResponse) is performed, and the image is committed. Then, two more sub-microflows (SUB_AutoChooseProgrammingLanguage_FromRe... and SUB_AutoChooseSubjectFromResponse_For_Cheat...) are called to automatically extract the programming language and subject from the Gemini Pro response, based on the contents of the cheatsheet.
- Synthesis and display: After successful analysis, a confirmation message is displayed. Finally, a sub-microflow (SUB_GeminiPro_Cheatsheet_Summarization) is called to synthesize the information from the cheatsheet, possibly creating a summary or highlights. This information is then used to generate an email (ACT_CreateEmail, Message_For_Cheatsheet) and displayed on an overview page (Show page 'Cheatsheet_Overview_Lastest').

This example illustrates how LC-IAF (see Figure 3), through a low-code interface like Mendix, allows teachers to easily integrate advanced AI capabilities to classify and extract pedagogical information from visual materials without writing complex code. This directly supports the automation of tasks such as classifying learning materials and generating actionable insights.

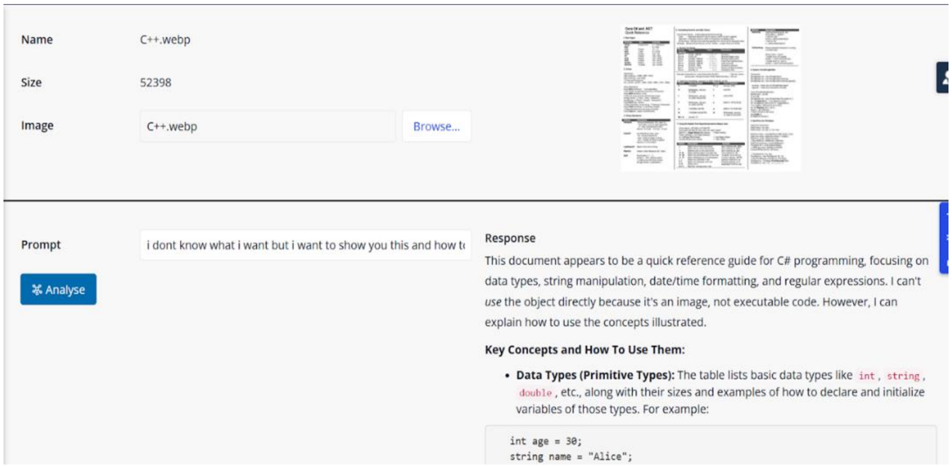


**Fig. 1 Cheatsheet analysis application interface on the Mendix platform.**

**Table 1. Key features of the proposed low-code image analysis framework**

| Module/Component | Main Function | Core Technology | Benefits |
|---|---|---|---|
| **Image Ingestion** | Supports various formats (PDF, Image, Digital Data Processing) | Data Processing | Reduced data processing time |
| **Data Preprocessing** | Transforms, normalizes, image processing, noise reduction | Image Processing | Optimizes and cleans AI analysis data |
| **Core AI Analysis Engine** | Object detection, CNN-based classification, deep learning, machine learning, and semantic segmentation | AI Algorithms | Provides insights into content |
| **Analysis & Interpretation Module** | Evaluates value in recognition, confidence levels related to text, accuracy, and customized algorithms | Multi-modal models, transfer learning | Converts data into meaningful insights |
| **Low-Code Interface Development** | Visual interface and dashboard, visual control of forms, and existing templates | Visual Programming, UX/UI | Empowers users |
| **Output & Recommendation Module** | Creates actionable recommendations, visualizes data, and visualizes results | Data Visualization, Recommendation Algorithms | Provides support and continuous improvement |

## 4. Impact on Education and Discussion

This section outlines how LC-IAF can improve instructional design and personalized learning, and outlines a comprehensive evaluation approach to validate the effectiveness of the framework.

### *4.1. Enhancing Instructional Design and Personalized Learning*

LC-IAF aims to facilitate teachers systematically using MM principles, transforming from subjective, intuition-based instructional design toward objective, data-driven pedagogy. Improving Instructional Design: This system automatically prioritizes irrelevant or distracting images, offering alternative images to reduce interference with extraneous cognitive load, analyzes the layout of images to propose appropriate placement of labels/highlights, and evaluates the proximity of images to relevant text that can be better integrated. It also returns formulae for quantitative or qualitative estimates of the image complexity, which is useful for teachers to choose the images needed by a level of students, limiting the cognitive load.

*4.1.1. Personalized Learning and Feedback*

This framework contributes to adaptive learning systems by automatically adjusting the difficulty or presentation of visual content based on students' individual needs or learning styles. For example, it can provide simplified diagrams for struggling learners or more complex images for advanced students, also recommend additional visual resources or alternative explanations based on student engagement and comprehension, and provide automated feedback on visual products created by students. By analyzing visual content at a granular level, this framework enables highly personalized visual learning paths, resulting in more accurate and effective adaptive learning experiences.

**Table 2. Mapping framework features to pedagogical principles and benefits for teachers**

| Framework Features | Pedagogical Principles for Resolution | Actions / Benefits for Teachers | Impact on Learner Learning |
|---|---|---|---|
| **Complexity Assessment** | Theories in Perception | Adjust the complexity of instructions | Reduces cognitive load |
| **Relevance Level of Context** | Cohesion Principle | Ensure appropriate image-text matching | Improves comprehension ability |
| **Low-code Interface** | Teacher's Ownership | Customizes the learning process | Enhances interaction |
| **Semantic Segmentation** | Multi-modal Principles | Distinguishes meaningful content | Optimizes attention and focus |
| **Report analysis** | Spatial cohesion principles | Arranges elements for effective visualization | Enhances information integration |

***4.2. Evaluation and Discussion***

A multi-faceted evaluation approach will be adopted to validate the effectiveness of LC-IAF, including technical validation, user usability, and pedagogical effectiveness.

*4.2.1. Technical Validation*

The core AI analytics tool will be evaluated using common measures such as Accuracy, F1-score, and mAP for image classification, object detection, and semantic segmentation. Moreover, the performance and reproducibility of the tool will be assessed. One major obstacle in this effort is access to large and representative datasets (the data to feed the framework, as well as well-annotated datasets that could be used as a reference for training and evaluating it), as well as to an accurate measure of the pedagogical accuracy of the framework's insights.

*4.2.2. Teacher Usability and Acceptance*

The usability of the low-code interface will be evaluated on a wide range of teachers by performing qualitative and quantitative user studies. The framework's ease of use and effectiveness will be assessed using metrics such as SUS (System Usability Scale) and task completion rates. Teacher feedback and satisfaction surveys will be used to evaluate their acceptability and perceived value. Also, gaining trust from users (teachers) and the AI system through explainability and transparency in the UX design is mandatory for long-term success.

*4.2.3. Pedagogical Effectiveness*

Different ways of measuring the effect of the framework on student learning will be presented. Pre-post tests will be used to compare learning gain between technology-enhanced AI materials and non-AI intervention groups. The human pose will also be used together with engagement indicators, such as monitoring students' interactions with visual content. Additionally, subjective or objective (i.e., eye tracking) measurements of cognitive load will be taken. Lastly, teacher efficiency, which saves time when preparing a superior lesson with better quality feedback for teachers, will be taken into account.

The framework's actual efficacy rests in its capacity to convert better instructional design (through AI) into quantifiable learning gains and less cognitive strain for students. This needs to be rigorously validated empirically and go beyond anecdotal evidence to show a causal relationship between improved learning outcomes and AI-powered visual analysis.

**Table 3. Comparison of existing image analysis tools against the proposed low-code approach**

| Tool/Method | Focus | Technical Expertise Requirement | Pedagogical Understanding Provided | Teacher Usability | Scalability |
|---|---|---|---|---|---|
| **OpenCV** | General Computer Vision | High (programming) | None | Low | High |
| **Commercial AI APIs** | Enterprise Data Processing | Medium (API usage) | Limited | Medium | High |
| **LC-IAF (proposed)** | Pedagogical Image Analysis | Low (low-code interface) | Comprehensive | High | High |

## 5. Conclusion and Proposal for Future Research

This final section concludes and outlines future research directions for LC-IAF, discusses challenges and ethical considerations associated with implementing AI in education, and draws overall conclusions about the potential of the framework.

### *5.1. Evaluation and Discussion Methods*

The introduced solution, Low-Code Image Analysis Framework (LC-IAF), bridges the advanced analytic capabilities of Artificial Intelligence (AI) with actual pedagogical needs among teachers. By making powerful image analysis more accessible with an intuitive low-code interface, LC-IAF enables educators to improve visual teaching materials in a rapid and systematic manner.

The strength of LC-IAF is that it converts raw AI-technical outputs into pedagogically useful insights by adopting well-known learning theories such as Cognitive Load Theory and Mayer's Multimedia Learning Principle. This not only enhances the quality of instructional design by minimizing cognitive load due to nonalignment but also encourages personalized cognitive and adaptive learning experiences targeting load and student engagement issues.

Implementing and validating LC-IAF is not easy, despite the previous point. Considerations of robust empirical evidence to estimate real pedagogical impact and barriers to obtaining the collection of accurately annotated data and the reliable quantification of “pedagogical accuracy” are the focus of investigation. In addition, ethical concerns regarding data privacy, algorithm bias, deployment cost, and the necessity of sustaining human-AI cooperation are important issues that should be dealt with in a careful and responsible manner.

### *5.2. Future Research*

The Future research and development directions for LC-IAF will focus on extending the framework’s capabilities to address more complex edge tracks of learning materials and deeper analytics into the educational ecosystem.

Advanced document classification: The framework will also be extended to carry out multi-method document classification, studying form images with textual content (via OCR and NLP), and the layout of the document. As a result, it will be able to support the automatic classification of different types of educational materials (e.g., textbooks, worksheets, presentations) according to their content and legal characteristics. Study deep architectures that can use various heterogeneous approaches to achieve better classification performance.

Integration with Open Educational Resources: This development direction will discuss how to facilitate OER's creation, discovery, and adaptation. Specifically, this includes the ability to automatically tag image content in OER repositories with professorial metadata (e.g., "cognitive load reduction schemes"), improve information accessibility through the generation of alternative text, and enable teachers to adapt OER images to their needs using low-code interfaces easily.

Explainable AI (XAI) for Pedagogical Insights: Future research will focus on incorporating XAI techniques to provide transparent solutions to professorial recommendations. XAI is not seen as just a convenience but also as a pedagogical requirement. By explaining the explanation behind certain AI recommendations, XAI can transform the framework from a mere tool to a virtual assistant, helping teachers use AI and better understand the underlying pedagogical principles.

Generate appropriate content applications: Use frameworks to automatically generate or modify content images in real-time for applied learning systems. This would include exploiting biologically generative AI models such as text-to-image models driven by human-informed attack vectors to generate new, optimized images according to the learning engine requirements.

### *5.3. Ethical Challenges and Considerations*

With the introduction of LC-IAF, as with any AI technology in education, there are several ethical issues and challenges that must be carefully considered. First, Data Privacy and Security are a concern, especially when considering the gathering, storage, and processing of sensitive visual educational data (e.g., student work, classroom observations). Compliance with regulations (e.g., GDPR, FERPA), strong data security measures, proper consent from students and parents, and clear data retention policies are all required to help manage these risks.

Second, Algorithmic Bias is a highly critical issue. If AI approaches learning from non-diverse or non-representative data, this can result in implicit biases and, thus, unbalanced or misleading pedagogical perceptions. They suggest addressing this by taking counteracting measures such as reducing bias in the training data, performing regular auditing, and using Explainable AI (XAI) for bias detection.

Third, the financial and technological barriers associated with the Implementation Costs and WI tools in education can significantly deter or support the adoption of AI tools. Adapting these solutions to various educational contexts is also difficult to scale.

Above all, promoting effective human-AI collaboration is key. AI needs to be an aid to a teacher, not a replacement. It is believed that evidence would arise wherein people question the absence of interactions with other humans and whether it may result in overreliance on AI-produced content and deterioration of critical thinking. There is a need to ensure teachers are maintained as independent and professional in all pedagogical matters. In short, LC-IAF promises to revolutionize the way we can create, analyze, and deliver visual learning content. While there are challenges to overcome, its potential to dramatically improve educational outcomes by empowering teachers and personalizing learning experiences is huge, positioning it as a significant step forward in the field of educational technology.

## Acknowledgments

This research is funded by Ho Chi Minh City University of Education Foundation of Science and Technology under grant number CS.2024.19.33.

## References

[1] Rustam Abdrakhmanov et al., "Applying Computer Vision and Machine Learning Techniques in STEM-Education Self-Study," *International Journal of Advanced Computer Science & Applications*, vol. 15, no. 1, pp. 1-9, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Asma Khazaal Abdulsahib et al., "Artificial Intelligence Based Computer Vision Analysis for Smart Education Interactive Visualization," *Fusion: Practice and Applications*, vol. 15, no. 2, pp. 245-260, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Houichi Asma, and Sarnou Dallel, "Cognitive Load Theory and its Relation to Instructional Design: Perspectives of Some Algerian University Teachers of English," *Arab World English Journal (AWEJ)*, vol. 11, no. 4, pp.110-127, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] Archontia Foutsitzi, "Images in Educational Textbooks and Educational Audiovisual Media," *European Journal of Language and Literature Studies*, vol. 4, no. 2, pp. 30-33, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[5] Ross Girshick et al., "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, pp. 580-587, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[6] Hamza Ouhaichi, Vogel Bahtijar, and Daniel Spikol, "Exploring Design Considerations for Multimodal Learning Analytics Systems: An Interview Study," *Frontiers in Education*, vol. 9, pp. 1-15, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Famuyide Alaba Johnson et al., "Assessing the Awareness and Usability of Artificial Intelligence Software Among Secondary School Teachers," *Educational Perspectives*, vol. 13, no. 1, pp. 1-10, 2024. [Google Scholar] [Publisher Link]

[8] Maria Kandaurova, Daniel A. Skog, and Petra M. Bosch-Sijtsema, "The Promise and Perils of Low-Code AI Platforms," *MIS Quarterly Executive*, vol. 23, no. 3, pp. 1-16, 2024. [Google Scholar] [Publisher Link]

[9] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84-90, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[10] Evan Shelhamer, Jonathan Long, and Trevor Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640 - 651, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[11] Katerina Mangaroska et al., "Multimodal Learning Analytics to Inform Learning Design: Lessons Learned from Computing Education," *Journal of Learning Analytics*, vol. 7, no. 3, pp. 79-97, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Richard E. Mayer, *The Cambridge Handbook of Multimedia Learning*, 2nd ed., Cambridge University Press, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[13] Mothwa Louis, "*Smart Attendance Monitoring System Using Computer Vision*," Doctoral Dissertation, University of KwaZulu-Natal, Westville, 2019. [Google Scholar] [Publisher Link]

[14] Fred Paas, Alexander Renkl, and John Sweller, "Cognitive Load Theory and Instructional Design: Recent Developments," *Educational Psychologist*, vol. 38, no. 1, pp. 1-4, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[15] A. Paivio, "*Dual Coding Theory and Education, Pathways to Literacy Achievement for High Poverty Children*," University of Michigan School of Education, 2006. [Google Scholar]

[16] Aravind Sasidharan Pillai, "Student Engagement Detection in Classrooms through Computer Vision and Deep Learning: A Novel Approach Using YOLOv4," *Sage Science Review of Educational Technology*, vol. 5, no. 1, pp. 87-97, 2022. [Google Scholar] [Publisher Link]

[17] Jamal Raiyn, "The Role of Visual Learning in Improving Students' High-Order Thinking Skills," *Journal of Education and Practice*, vol. 7, no. 24, pp. 115-121, 2016. [Google Scholar] [Publisher Link]

[18] Raimundo da Silva Soares et al., "Integrating Students' Real-Time Gaze in Teacher–Student Interactions: Case Studies on the Benefits and Challenges of Eye Tracking in Primary Education," *Applied Sciences*, vol. 14, no. 23, pp. 1-17, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] Maria Trigka, and Elias Dritsas, "A Comprehensive Survey of Deep Learning Approaches in Image Processing," *Sensors*, vol. 25, no. 2, pp. 1-46, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[20] V.P. Joshith, and S. Asha, "Multimodal Learning Analytics (MMLA) In Education-A Game Changer for Educators," *Indian Journal of Educational Technology*, vol. 7, no. 1, pp. 329-344, 2025. [Google Scholar] [Publisher Link]

[21] Rina Zviel-Girshin, "The Good and Bad of AI Tools in Novice Programming Education," *Education Sciences*, vol. 14, no. 10, pp. 1-17, 2024. [CrossRef] [Google Scholar] [Publisher Link]