Editor's Note

Recent advances in Artificial Intelligence (AI) applied to Human–Computer Interaction (HCI) have been characterized by the convergence of immersive technologies, semantic reasoning, and data-centric learning paradigms. In augmented reality (AR), AI enables real-time interaction by integrating computer vision and contextual understanding into dynamic environments, although it simultaneously introduces new security vulnerabilities that require robust AI-driven protection mechanisms. In parallel, AI-driven cybersecurity has evolved toward proactive and adaptive models that leverage machine learning for real-time threat detection, anomaly analysis, and predictive mitigation, significantly enhancing the resilience of digital systems. Ontologies further strengthen this landscape by providing structured, semantically rich representations of cybersecurity knowledge, facilitating interoperability, automated reasoning, and intelligent threat assessment. Additionally, advances in visual analytics and time series foundation models are enabling the processing of large-scale temporal and multimodal datasets, supporting more sophisticated data interpretation and decision-making processes in complex environments.

Modern computer vision systems leverage deep learning architectures such as convolutional neural networks and vision transformers to extract high-level features, enabling accurate object recognition, pattern detection, and scene understanding. Advances in multimodal AI systems—capable of integrating text, images, audio, and video—have become essential for achieving more comprehensive and context-aware intelligence, especially in domains such as medical imaging and interactive systems.

Based on these advances highlight, this regular issue presents the most notable research in this direction. In the field of convolutional networks Jin Wang et al. present a paper entitled "Demosaicking Algorithm Using Deep Residual Convolutional Network". The authors introduce a novel demosaicking framework grounded in a Deep Residual Convolutional Neural Network (DRCNN), designed to reconcile reconstruction fidelity with computational efficiency. The proposed methodology adopts a hybrid initialization strategy, wherein mosaicked inputs are preliminarily reconstructed using conventional demosaicking techniques, followed by channel-wise refinement through a learned residual mapping for the three color components. Architecturally, the DRCNN integrates complementary design paradigms to address the intrinsic challenges of ill-posed reconstruction in color filter array (CFA) data. Binary Convolution Units (BCUs) are incorporated to reduce computational overhead, while Efficient Layer Aggregation Networks (ELAN) facilitate multi-scale feature integration. In parallel, Dense Residual Blocks (DRBs) enhance gradient propagation and stabilize deep network training, reflecting the broader effectiveness of residual learning strategies in mitigating vanishing-gradient limitations in deep models. Extensive experimental validation demonstrates that the proposed framework achieves superior performance relative to state-of-the-art demosaicking approaches, as evidenced by improvements in peak signal-to-noise ratio (PSNR), reduced computational complexity, and enhanced perceptual image quality.

In the field of genetic algorithms, this issue presents the work proposal by Juan D. Gutierrez et al., entitled "Exploring the Limits of Foundation Models in Medical Image Segmentation: A Case Study With SAM and Genetic Algorithms". This study critically examines the operational limits of foundation models for medical image segmentation, with particular emphasis on the Segment Anything Model (SAM) proposed by Meta. Although prior work has demonstrated SAM's utility as a cost-efficient segmentation framework, the present contribution extends this paradigm by integrating prompt enhancement strategies with genetic algorithm–based optimization, aiming to reduce dependency on user-provided inputs while improving segmentation fidelity. As a proof of concept, the proposed methodology is evaluated across heterogeneous imaging modalities, including axial lung computed tomography (CT), frontal chest radiography, and spleen magnetic resonance imaging (MRI), using publicly available datasets. The empirical results indicate that genetic optimization of prompts yields consistent and statistically significant improvements in segmentation performance, approaching the accuracy levels of fully supervised, task-specific state-of-the-art models. However, the analysis also reveals an intrinsic performance ceiling, beyond which additional iterations of the genetic algorithm fail to yield further improvements. This observation suggests inherent limitations in prompt-based optimization when applied to pre-trained foundation models.

For pet lovers, especially dog lovers, the work proposal by Oluwakemi Akinwehinmi at al., entitled "Deep Learning for Detecting Abandoned Dogs", introduces a novel methodological framework that integrates an algorithmic model with a structured operational workflow for the detection of abandoned dogs in natural and public environments. The proposed approach is grounded in a temporal-contextual paradigm, combining temporal behavioral rules with context-aware object detection. Specifically, the methodology extends conventional bounding boxes of detected dogs to incorporate proximal visual cues indicative of abandonment, thereby enhancing semantic interpretation beyond isolated object recognition. The system leverages a combination of deep learning-based object detection, trajectory analysis, and image segmentation to discriminate effectively between abandoned and accompanied animals. In addressing the inherent challenges of this task—particularly data scarcity and the ambiguity associated with distinguishing ownership status—the study incorporates single-channel image augmentation strategies.

Next, in the field of visual analytics, in the proposal made by Inmaculada Santamaria-Valenzuela et al., entitled "Decoding Latent Spaces: Assessing the Interpretability of Time Series Foundation Models for Visual Analytics", authors investigate the interpretability of latent representations generated by time series foundation models, with a particular emphasis on their applicability to visual analytics tasks. Specifically, the analysis focuses on the MOMENT family of models, a suite of transformer-based, pre-trained architectures designed for multivariate time series applications, including imputation, forecasting, classification, and anomaly detection. The evaluation assesses the ability of these models, across five datasets, to capture intrinsic structural patterns within time series data through their latent space embeddings. Additionally, the study examines whether fine-tuning enhances the semantic clarity and separability of these embedding spaces. While fine-tuning yields notable improvements in model performance, as evidenced by a reduction in loss metrics, the corresponding gains in embedding interpretability remain limited according to visual inspection. These findings suggest that, despite the robustness and computational efficiency of foundation models such as MOMENT—particularly in reducing execution time and enabling interactive visual analytics—their latent spaces may not yet be readily interpretable without further methodological enhancements. Potential avenues for improvement include the adoption of alternative dimensionality reduction techniques, the design of interpretability-oriented loss functions, and more effective data preprocessing strategies.

DOI: 10.9781/ijimai.2026.8008

In the field of multimodal AI systems, we present the work proposal by Isabel Ferri-Molla et al., titled "Improving Aphasic Communication Using Multimodal AI Systems". This study aims to improve communication for people with aphasia (a condition that affects language after brain injury). Authors propose a system with two parts: a) A speech recognition model trained specifically to understand aphasic speech, which significantly reduces transcription errors; b) A multimodal component that uses visual context (e.g., images) to better understand what the speaker means, not just what they say. Results show that the adapted speech model performs much better than standard ones, and that combining audio with visual information helps capture the speaker's intended message more accurately than audio alone.

In the area of human computer interaction, the work proposal by Stevica Cvetković et al., entitled "Edge-Centric Augmented Reality Framework for Realtime Wristwatch Try-On", offers a well-founded and pragmatically oriented contribution to the advancement of augmented reality applications in e-commerce, particularly within the underexplored domain of small-accessory virtual try-on. The manuscript is situated within the expanding domain of e-commerce visualization technologies, emphasizing the growing demand for immersive and realistic product interaction. The authors correctly identify virtual try-on (VTO) systems as a strategic mechanism to enhance user confidence and mitigate return rates—two critical variables in online retail performance metrics. By prioritizing computational efficiency, real-time performance, and edge-device deployment, the proposed framework effectively addresses critical constraints related to latency, scalability, and data privacy. Although certain limitations persist—most notably in photorealistic rendering and robustness under diverse environmental conditions—the study convincingly demonstrates the technical feasibility and user acceptance of lightweight AR solutions in resource-constrained contexts. Overall, the manuscript constitutes a valuable step toward decentralized, high-quality AR experiences, with clear implications for future research in human-computer interaction and edge-based visual computing systems.

From visual analytics perspectives to computational learning field, the article written by Sofia Strukova et al., entitled "Classifying Professional Photographers on Instagram: Data Collection and Processing for Computational Learning", states that the widespread adoption of photo-sharing platforms provides a novel opportunity for data-driven inquiry. This work contributes with a comprehensive multimodal dataset comprising 29,679 Instagram posts derived from 1,042 user profiles, each annotated according to professional or non-professional photographer status. Leveraging this dataset, multiple machine learning (ML) approaches were systematically evaluated to determine their effectiveness in profile classification. Among the tested models, the Random Forest (RF) algorithm demonstrated superior performance, effectively capturing structural patterns characteristic of professional photography profiles. Further statistical analysis identified significant discriminative features between the two profile categories. Notably, the number of tagged users, the technical quality score of posts, and the variance in image height emerged as the most influential predictors of professional status. These findings not only highlight relevant indicators for expertise identification in visual social media contexts, but also establish a robust foundation for future research and practical applications in domains such as talent discovery, digital marketing, and social media analytics.

As an experimental study in the field of education, the work proposed by Joaquin González-Cabrera et al., entitled "Adaptation of the AIessphobia Scale in Education in a Sample of Spanish Adolescents", presents an important contribution to the emerging field of AI-related psychosocial research in education. By providing empirically supported evidence of its factorial structure, reliability, and convergent validity, the study establishes a foundational instrument for systematically examining anxiety associated with AI dependency among adolescents. The absence of significant differences across sex and educational levels further reinforces the transversal nature of this construct. Importantly, the identification of threshold scores enables early detection of potentially problematic patterns, supporting targeted educational and preventive interventions. As AI continues to permeate learning environments, future research should extend this line of inquiry toward longitudinal designs, cross-cultural validation, and the exploration of causal mechanisms, thereby informing evidence-based digital literacy policies and promoting a balanced, critical engagement with emerging technologies.

For its part, the proposal presented by Andrey Voynov et al., entitled "Typst: A Modern Typesetting Engine for Science", introduces Typst, a modern, markup-based document preparation system designed for producing professional documents with a strong emphasis on simplicity, performance, and powerful scripting for structured content and automated layout and styling. This paper presents a comprehensive review of the platform and its surrounding ecosystem through the systematic and comparative analysis, highlighting key features, some unique to this approach, that are relevant to both general purpose and niche use cases. It introduces the system and compares its output and functionality with LaTeX, outlines the historical development of document technologies and the core layout and composition algorithms they rely on, and explains the language through visual examples accompanied by corresponding source code. The paper also examines the official Typst compiler and web application, as well as adoption trends and ecosystem growth. It concludes with cross-domain case studies in Physics, Math, and Computer Science, intended for beginners and readers interested in modern document workflows and systems.

Finally, in the field of cybersecurity, this regular issue presents the manuscript proposal by Yixuan Wang et al. entitled "A Practical Cybersecurity Ontology Generator Based on Hierarchical Clustering and Multi-way Tree". In this study, the authors state that the construction of cybersecurity ontologies has traditionally relied on the expertise of cybersecurity specialists and ontology engineers, with prevailing approaches typically focusing on either unstructured textual sources or structured datasets in isolation. However, the simultaneous integration of both data types remains an underexplored challenge. In this context, the present study introduces Locust, a novel software tool designed to support the comprehensive generation of cybersecurity ontologies through the combined use of structured specifications and domain-specific corpora. Specifically, open-source cybersecurity standards are employed as structured inputs to establish the foundational ontology schema, while complementary domain corpora are utilized to enrich, refine, and complete the ontology content. Furthermore, the work proposes a systematic methodology for ontology filtering and simplification, leveraging hierarchical clustering techniques and multi-way tree representations to improve structural coherence and manage complexity. The reported experimental evaluation provides empirical evidence supporting the effectiveness of the proposed approach in constructing domain-specific cybersecurity ontologies from heterogeneous data sources.

We hope that this issue increases, supports and inspires future research in multidisciplinary fields.

Dr. Paulo Alonso Gaona-García
Universidad Distrital Francisco José de Caldas