Transforming Healthcare with AI: The Work of CINI-AIIS at Federico II University

Mariano Barone[1,†], Domenico Benfenati[1,†], Salvatore Capuozzo[1,†], Giovanni Maria De Filippis[1,†], Adriano De Simone[1,†], Simona Fioretto[1,†], Kyriakos Kristofer Georgiou[1,†], Michela Gravina[1,*,†], Julian Gortz[2,†], Lidia Marassi[1,†], Stefano Marrone[1,†], Enea Vincenzo Napolitano[1,†], Valeria Panzetta[2,3,†], Giuseppe Pontillo[4,5,†], Giuseppe Riccio[1,†], Antonio Romano[1,†], Cristiano Russo[1,†], Carmela Russo[6,†], Cristian Tommasino[1,†], Mariano Sirignano[3,†], Konstantinos Siettos[7,†], Elio Masciari[1,†], Antonio Maria Rinaldi[1,†], Paolo Antonio Netti[2,3,†], Vincenzo Moscato[1,†] and Carlo Sansone[1,*,†]

[1] *DIETI, University of Naples Federico II, Naples Italy*
[2] *CRIB, University of Naples Federico II*
[3] *DICMAPI, University of Naples Federico II, Naples Italy*
[4] *Department of Radiology and Nuclear Medicine, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands*
[5] *Department of Advanced Biomedical Sciences, University of Naples Federico II, Naples, Italy*
[6] *Istituto di Scienze e Tecnologia per l'Energia e la Mobilità Sostenibili (STEMS)-CNR, Naples Italy*
[7] *Department of Mathematics and Applications, University of Naples Federico II*

Abstract

Artificial Intelligence (AI) is playing an increasingly central role in the transformation of healthcare, offering powerful tools to support diagnostics, treatment planning, and the management of complex clinical data. In this paper, we provide a concise overview of several initiatives conducted at the University of Naples Federico II node of the CINI-AIIS Lab, highlighting their primary objectives and contributions. Our work spans a wide spectrum of AI methodologies, applied to diverse medical domains such as radiology, oncology, and personalized medicine. Collectively, these contributions underscore our commitment to developing robust, clinically relevant AI solutions that advance the state of the art in biomedical research and healthcare delivery.

Keywords

Artificial Intelligence, Healthcare, Deep Learning, Machine Learning

1. Introduction

In recent years, the integration of Artificial Intelligence (AI) into clinical and administrative processes holds significant promise for enhancing patient outcomes, optimizing resource allocation, and supporting medical professionals in complex diagnostic and therapeutic tasks. AI-driven systems are capable of processing large volumes of heterogeneous medical data, identifying subtle patterns, and delivering insights that can improve diagnostic accuracy, particularly in cases involving multifactorial conditions or extensive clinical histories. AI in healthcare comprises a diverse set of technologies rather than a single monolithic approach. Among these, Machine Learning (ML) plays a central role by enabling

---

Ital-IA 2025: 5th National Conference on Artificial Intelligence, organized by CINI, June 23-24, 2025, Trieste, Italy

[*] Corresponding author.

[†] These authors contributed equally.

michela.gravina@unina.it (M. Gravina); carlo.sansone@unina.it (C. Sansone)

0009-0004-0744-2386 (M. Barone); 0009-0008-5825-8043 (D. Benfenati); 0000-0002-2578-9349 (S. Capuozzo); 0009-0002-8395-0724 (G. De Filippis); 0009-0001-3740-416X (A. De Simone); 0009-0006-8700-8188 (S. Fioretto); 0000-0001-7549-2298 (K. K. Georgiou); 0000-0001-5033-9617 (M. Gravina); 0009-0006-8134-5466 (L. Marassi); 0000-0001-6852-0377 (S. Marrone); 0000-0002-6384-9891 (E. V. Napolitano); 0000-0001-6174-4116 (V. Panzetta); 0000-0001-5425-1890 (G. Pontillo); 0009-0002-8613-1126 (G. Riccio); 0009-0000-5377-5051 (A. Romano); 0000-0002-8732-1733 (C. Russo); 0000-0002-6891-4250 (C. Russo); 0000-0001-9763-8745 (C. Tommasino); 0000-0003-4363-7194 (M. Sirignano); 0000-0002-9568-3355 (K. Siettos); 0000-0002-1778-5321 (E. Masciari); 0000-0001-7003-4781 (A. Rinaldi); 0000-0002-2435-7181 (P. A. Netti); 0000-0002-0754-7696 (V. Moscato); 0000-0002-8176-6950 (C. Sansone)

systems to learn from data and improve performance over time without explicit programming. ML is widely employed in clinical decision support, prognostic modeling, and precision medicine, where it helps identify optimal treatment strategies based on patient-specific factors. A particularly impactful subfield is Deep Learning (DL), which uses artificial neural networks to model complex patterns in high-dimensional data. DL has revolutionized medical image analysis, facilitating tasks such as radiological diagnosis, image segmentation, and automated detection of pathological features. In parallel, the recent emergence of Large Language Models (LLMs) has opened new avenues for natural language understanding and generation in medicine. Trained on massive text corpora, LLMs are capable of interpreting and producing human-like language, making them valuable tools for clinical documentation, decision support, literature summarization, and patient communication.

This paper provides an overview of the research activities carried out at the University of Naples Federico II, operating within the national CINI-AIIS laboratory. We present a selection of projects that illustrate how advanced AI methodologies are being applied to address real-world medical challenges. Emphasis is placed on the scientific innovation, methodological rigor, and translational potential of each initiative, underscoring the lab's contribution to the advancement of AI in biomedical domains.

2. AI in Medicine: Opportunities and Risks

The application of AI in medicine is progressing rapidly, offering promising tools for enhancing diagnostic accuracy and predicting disease progression, particularly in complex and heterogeneous conditions such as neurodegenerative diseases. However, these advancements raise critical ethical, technical, and regulatory concerns, notably the risk of algorithmic bias. Bias can arise at multiple stages of AI development—from data collection to model deployment—and may lead to diagnostic inaccuracies or unequal care, especially for underrepresented populations [1].

To mitigate these risks, there is a growing emphasis on the adoption of certification protocols that ensure compliance with both technical standards and ethical principles. Recent efforts, such as the collaboration between the Italian accreditation body Accredia and the CINI consortium, have led to the development of structured validation frameworks inspired by ISO/IEC 24027 guidelines [2]. These frameworks incorporate bias detection, fairness evaluation, and mitigation strategies, and are designed to promote transparency and equitable performance across diverse patient subgroups. This is particularly relevant for AI applications in monitoring and stratifying neurodegenerative diseases like multiple sclerosis, where model generalizability and fairness are critical for clinical decision-making. Ultimately, while AI holds transformative potential in neurology and other medical fields, its deployment must be accompanied by rigorous oversight. Certification and validation protocols should not be viewed as bureaucratic obstacles, but rather as foundational tools for building trustworthy, interpretable, and ethically sound AI systems that can be confidently integrated into healthcare practice.

3. KFM System Enhancement on Trematodes Eggshells Detection

Building on the prototype by Cringoli et al. [3] and the AI methods of Capuozzo et al. [4], the Kubic FLOTAC Microscope (KFM) has been refined to better detect eggshells of *Fasciola hepatica* and *Calicophoron daubneyi*—trematodes causing fasciolosis and paramphistomosis in ruminants. These eggshells are difficult to distinguish visually due to their similar morphology and color, especially in debris-laden samples, and their emptied, deformed structure adds to the complexity.

To overcome these challenges, the KFM system was upgraded with an improved YOLO-based object detector and two image processing modules placed before and after inference. After scanning a 16×12 grid of images from the Mini-FLOTAC chamber, the data is sent to a server. A pre-inference color grading step then normalizes image color profiles using the PDF of the reference distribution derived from clean training samples, reducing misclassification caused by subtle color variations and enhancing detection accuracy.

Post-inference, a second module reduces redundancy caused by overlapping image edges regions, which are intentionally designed to compensate mechanical inaccuracies in scanning. This overlap can produce duplicate eggshell detections across adjacent frames, affecting egg counts. To address this, an overlap-aware algorithm uses SSIM and positional data to discard duplicates.

These enhancements led to a mAP50 of 94.2% for eggshell detection and a MAE of 8 in egg counts, marking significant progress toward validating and deploying a complete, portable, and cost-effective KFM system capable of detecting over 20 parasitic egg classes.

4. Synergy for COVID: a ML-based LOS Prediction System for Pneumonia Patients

The COVID-19 pandemic highlighted the urgent need for automated diagnostic tools to help clinicians quickly analyze test results and imaging, improving decision-making and hospital resource management during high patient influx. This study focuses on pneumonia patients using lung CT scans and clinical data. Most diagnostic systems rely on a single data modality limiting their robustness in real-world settings where missing data are common. In contrast, multimodal approaches provide flexibility for this issue and enhance the value of small datasets—an important benefit in the data-scarce biomedical field.

Developed with the University of Campania Luigi Vanvitelli and Cotugno Hospital, this work integrates multimodal data to build a system for predicting the length of stay (LOS) of pneumonia patients [5]. Clinical data were preprocessed using imputation, SMOTE for class imbalance, and feature selection methods like Backward Feature Elimination and Forward Feature Selection. CT images were processed with a modified ResNet, replacing the final layer with a fixed-size one to extract features for integration with tabular data. After testing various feature sets, data sources, and models, the best-performing models achieved 87.5% accuracy for outcome prediction and a mean absolute error (MAE) of 3.102 days for LOS. Training on a single data source further reduced the MAE to 1.43 days.

5. A Physiological-Informed Model for Lesion Classification

DL has significantly advanced medical image analysis, but its performance remains constrained by limited annotated data, particularly in tasks like breast lesion classification from Dynamic Contrast-Enhanced MRI (DCE-MRI). Traditional and generative data augmentation techniques often fall short in preserving the biological fidelity crucial for clinical deployment. To address this, the research activity described in [6] propose PBPK DAE-CNN, a physiologically-informed framework that augments small DCE-MRI datasets with biologically plausible synthetic images. The method combines a Physiologically-Based Pharmacokinetic (PBPK) model with a Deforming Autoencoder (DAE) to disentangle image components while preserving contrast agent dynamics through a pharmacokinetics-guided loss. Synthetic images are generated via latent feature recombination under class-consistent constraints, enhancing variability without compromising realism. A CNN classifier is trained using these samples within an iterative scheme that co-optimizes generative and discriminative models. Evaluated on three DCE-MRI datasets, the approach outperforms conventional and generative augmentation strategies in accuracy, generalization, and interpretability. This work highlights the value of integrating physiological priors into generative models for robust DL applications in medical imaging.

6. Carbon Nanoparticles in Breast Cell Imaging

The convergence of AI and fluorescence microscopy has enabled automated, high-precision analysis of cellular structures, while carbon dots (CDs)—fluorescent, biocompatible nanoparticles—have shown promise in enhancing cancer diagnostics. This research activity explores the effect of CD exposure on the performance of DL models applied to fluorescence microscopy images of breast cells. A modular pipeline is introduced, comprising an unsupervised cell segmentation module and a classification module based on EfficientNet-B0, applied to two breast cell lines (MCF10a and MCF7) under control and CD-exposed conditions. Explainable AI techniques are employed to enhance model interpretability. Results show that CD exposure improves classification accuracy in both intra- and cross-condition experiments and enhances the biological relevance of model attention, particularly around membranes and nuclei. These findings suggest that combining nanomaterial-enhanced imaging with AI can boost diagnostic accuracy and interpretability in early cancer detection.

7. LLM Applications in Nutrigenetics

Genetic variability plays a critical role in modulating individual responses to nutrition, with numerous studies linking metabolic pathway polymorphisms to diet-related disease susceptibility. While nutrigenetics promises personalized dietary interventions based on genetic profiles, its clinical translation remains constrained by fragmented data and poor standardization. Recent computational advancements, such as NLP and machine learning techniques, offer solutions to these challenges. In particular, De Filippis et al. (2023) used BioBERT for entity matching of GWAS phenotypes with food-related terms, creating genotype-specific dietary dataset [7]. Employing on this dataset, our work applies data-driven techniques using pretrained language models (PLMs), specifically fine-tuned BERT variants, to generate biomedical embeddings. We adopt the BERTopic framework—combining UMAP for dimensionality reduction and HDBSCAN for hierarchical clustering—to uncover latent semantic structures within a 37k-document corpus [8]. This process identified 50 topics, of which 21 were retained through cTF-IDF filtering to focus on genetic variants relevant to nutrition-related pathologies. There topics was used to build an LLM-based Q&A system, employing a RAG framework with vector databases to embed nutrigenetic knowledge. This approach improved accuracy and specificity over models like GPT-3.5 and Mistral-7B, aiding clinicians in prioritizing nutrigenetic variants and refining gene panels for personalized dietary strategies [9]. Our approach enhances the interpretability of genotype-phenotype-nutrition relationships and supports more refined, evidence-based gene panel selection for personalized nutrition strategies.

8. Self-Organizing Maps and Temporal Health Modeling for Corporate Wellness Programs

In collaboration with Antur SRL, we contributed to the design and evaluation of a data-driven protocol aimed at enhancing employee health and promoting organizational sustainability [10]. The study involved over 12,000 workers from 40 diverse companies and implemented a multidisciplinary intervention program, including personalized nutrition, postural education, music therapy, and stress management. Central to the approach was the use of unsupervised machine learning, specifically Self-Organizing Maps (SOMs), to analyze high-dimensional health data obtained from periodic Bioelectrical Impedance Analysis (BIA) and clinical records. The SOM model revealed ten distinct health clusters based on key anthropometric and metabolic features, enabling personalized intervention strategies tailored to each subgroup. Longitudinal trends in biometric parameters were modeled using linear regression, providing dynamic insights into intervention efficacy. The integrated system led to significant improvements in physiological health, reductions in absenteeism and workplace incidents, and measurable performance gains aligned with ESG goals.

9. AI-Powered Clinical Decision Support and Biomedical Fact-Checking

In this section, we present two systems developed to support clinical decision-making and biomedical fact verification through explainable and trustworthy AI. The first, PIE-Med [11], is a Clinical Decision Support System that integrates Graph Convolutional Networks (GCNs) and Large Language Models

(LLMs) to provide transparent and reliable medical recommendations. While LLMs offer contextual and natural language expressiveness, their limitations—such as hallucinations and lack of interpretability—are addressed by restricting their role to verbalizing the decisions inferred by the GCN, which is trained on structured patient data and domain-specific knowledge. The system, validated on the MIMIC-III [12] dataset, leverages interpretability methods like GNNExplainer and Integrated Gradients to trace the influence of input features, aligning algorithmic reasoning with clinical safety and explainability. The source code is available at https://github.com/PRAISELab-PicusLab/PIE-Med.

The second contribution, CER (Combining Evidence and Reasoning), is a framework for biomedical fact-checking that integrates retrieval-based evidence with LLM reasoning to overcome domain-specific limitations of generic fact-checkers. By grounding its outputs in peer-reviewed scientific literature (e.g., PubMed), CER minimizes hallucinations and enhances credibility, achieving strong performance across benchmarks such as HealthFC [13], BioASQ-7b [14], and SciFact [15]. The source code is available at https://github.com/PRAISELab-PicusLab/CER.

10. Multilingual and Multimodal Medical QA dataset

The advent of LLMs has highlighted the increasing necessity for medical QA systems to handle both informal and formal communications across various languages and modalities. This study presents two Italian datasets—one with 780,000 interactions from patient-doctor forums and another containing 26,000 licensing exam questions—to advance research in non-English clinical contexts. Findings confirm that domain-specific models, particularly those enhanced with the RAG technique, significantly outperform general-purpose LLMs. Additionally, a multilingual, multimodal benchmark, available at https://github.com/PRAISELab-PicusLab/MMMED, uses Spanish medical exam questions, images, and translations to assess vision-language models' clinical reasoning, supporting resilient multilingual medical AI.

11. Decoding Brain Networks with AI: the road to a Brain Digital Twin

The study of brain structure and function is central to advancing our understanding of neurodegenerative diseases such as Alzheimer's and Parkinson's. In this context, precision medicine is emerging as a transformative paradigm, increasingly driven by engineering approaches, including computational modeling, advanced data analysis, and AI. Our research activity focuses on developing AI-based methods for brain analysis to support personalized diagnostic and prognostic tools. Building on prior work in intelligent diagnosis [16], we now investigate the brain connectome using Granger causality to model temporal interactions among brain regions. This method offers promising potential for capturing functional dynamics and identifying predictive patterns linked to disease progression. Furthermore, the exploration of brain interactions contributes to uncovering computational principles of cognition, which may inspire novel AI architectures. We also review recent advances in brain and virtual brain digital twins, emphasizing the role of AI within these emerging systems. Overall, our work highlights AI's dual potential: to advance personalized approaches to neurodegenerative disease and to deepen our understanding of the brain's intrinsic complexity.

Acknowledgments

This work was supported in part by the Future Artificial Intelligence Research (FAIR) Project (Grant PE0000013-FAIR) and by the Production of sustainable carbon dots for cell diagnostics and targeting (PoLLINATE) Project (CUP E53D23017250001).

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] D. S. Char, N. H. Shah, D. Magnus, Implementing machine learning in health care—addressing ethical challenges, New England Journal of Medicine 378 (2018) 981–983.

[2] D. Leslie, Understanding artificial intelligence ethics and safety, arXiv preprint arXiv:1906.05684 (2019).

[3] G. Cringoli, A. Amadesi, M. P. Maurelli, B. Celano, G. Piantadosi, A. Bosco, L. Ciuca, M. Cesarelli, P. Bifulco, A. Montresor, et al., The Kubic FLOTAC microscope (KFM): a new compact digital microscope for helminth egg counts, Parasitology 148 (2021) 427–434.

[4] S. Capuozzo, S. Marrone, M. Gravina, G. Cringoli, L. Rinaldi, M. P. Maurelli, A. Bosco, G. Orrù, G. L. Marcialis, L. Ghiani, et al., Automating parasite egg detection: insights from the first AI-KFM challenge, Frontiers in Artificial Intelligence 7 (2024) 1325219.

[5] A. Annunziata, S. Cappabianca, S. Capuozzo, N. Coppola, C. Di Somma, L. Docimo, G. Fiorentino, M. Gravina, L. Marassi, S. Marrone, et al., A multimodal machine learning model in pneumonia patients hospital length of stay prediction, Big Data and Cognitive Computing 8 (2024) 178.

[6] M. Gravina, M. Maddaluno, S. Marrone, M. Sansone, R. Fusco, V. Granata, A. Petrillo, C. Sansone, A physiological-informed generative model for improving breast lesion classification in small dce-mri datasets, IEEE Journal of Biomedical and Health Informatics (2024).

[7] G. M. De Filippis, M. Monticelli, A. Pollice, T. Angrisano, B. H. Mele, V. Calabrò, Computational Strategies in Nutrigenetics: Constructing a Reference Dataset of Nutrition-Associated Genetic Polymorphisms, 2023. doi:10.1101/2023.08.04.23293659, iSSN: 2329-3659 Pages: 2023.08.04.23293659.

[8] G. M. De Filippis, A. M. Rinaldi, C. Russo, C. Tommasino, Advanced topic modeling in genomics: Towards personalized dietary recommendations through bertopic analysis, in: P. Delir Haghighi, M. Greguš, G. Kotsis, I. Khalil (Eds.), Information Integration and Web Intelligence, Springer Nature Switzerland, Cham, 2025, pp. 3–17.

[9] D. Benfenati, G. M. D. Filippis, A. M. Rinaldi, C. Russo, C. Tommasino, A Retrieval-augmented Generation application for Question-Answering in Nutrigenetics Domain, Procedia Computer Science 246 (2024) 586–595. URL: https://www.sciencedirect.com/science/article/pii/S1877050924025092. doi:https://doi.org/10.1016/j.procs.2024.09.467.

[10] M. L. Conza, S. Fioretto, E. Masciari, E. V. Napolitano, Enhancing employee health through an experimental diet: Insights from machine learning analysis, in: 2025 33rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP), IEEE, 2025, pp. 283–290.

[11] A. Romano, G. Riccio, M. Postiglione, V. Moscato, Pie-med: Predicting, interpreting and explaining medical recommendations, in: European Conference on Information Retrieval, ECIR, Springer, 2025, pp. 6–12.

[12] A. Johnson, T. Pollard, L. Shen, L. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. Celi, R. Mark, MIMIC-III, a freely accessible critical care database, 2016.

[13] J. Vladika, P. Schneider, F. Matthes, Healthfc: Verifying health claims with evidence-based medical fact-checking, in: Joint International Conference on Computational Linguistics, Language Resources and Evaluation, LREC/COLING, ELRA and ICCL, 2024, pp. 8095–8107.

[14] A. Nentidis, K. Bougiatiotis, A. Krithara, G. Paliouras, Results of the seventh edition of the bioasq challenge, in: Machine Learning and Knowledge Discovery in Databases - International Workshops of ECML PKDD, Springer, 2019, pp. 553–568.

[15] D. Wadden, S. Lin, K. Lo, L. L. Wang, M. van Zuylen, A. Cohan, H. Hajishirzi, Fact or fiction: Verifying scientific claims, in: Conference on Empirical Methods in Natural Language Processing, EMNLP, ACL, 2020, pp. 7534–7550.

[16] A. De Simone, M. Gravina, C. Sansone, Multimodality calibration in 3d multi input-multi output network for dementia diagnosis with incomplete acquisitions, in: Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR), Springer, 2024, pp. 92–101.