

# AI in Medicine: Activities of the CINI-AIIS Lab at University of Naples Federico II

Domenico Benfenati<sup>1</sup>, Salvatore Capuozzo<sup>1</sup>, Giovanni Maria De Filippis<sup>1</sup>, Adriano De Simone<sup>1</sup>, Michela Gravina<sup>1</sup>, Lidia Marassi<sup>1</sup>, Stefano Marrone<sup>1</sup>, Elio Masciari<sup>1</sup>, Enea Vincenzo Napolitano<sup>1</sup>, Giuseppe Pontillo<sup>1,4,5</sup>, Marco Postiglione<sup>1,3</sup>, Cristiano Russo<sup>1</sup>, Cristian Tommasino<sup>1,2</sup>, Antonio M. Rinaldi<sup>1</sup>, Vincenzo Moscato<sup>1</sup> and Carlo Sansone<sup>1,\*</sup>

<sup>1</sup>Department of Electrical Engineering and Information Technology (DIETI), University of Naples Federico II, Naples, Italy

<sup>2</sup>Interdepartmental Center for Research on Management and Innovation in Healthcare (CIRMIS), University of Naples Federico II, Naples, Italy

<sup>3</sup>Northwestern University, Department of Computer Science, McCormick School of Engineering and Applied Science, 2233 Tech Dr, Evanston, IL 60208, United States

<sup>4</sup>Department of Radiology and Nuclear Medicine, MS Center Amsterdam, Amsterdam Neuroscience, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

<sup>5</sup>Department of Advanced Biomedical Sciences, University of Naples Federico II, Naples, Italy

## Abstract

Artificial Intelligence (AI) encompasses a variety of methods and algorithms that have found applications across numerous domains over time. The increasing complexity and abundance of data in healthcare have spurred investigations into utilizing AI techniques within the medical field, leading to promising avenues for fostering innovation, facilitating early diagnosis, and aiding in treatments. In this study, we provide a concise overview of several initiatives conducted within this realm at the University of Naples Federico II node of the CINI-AIIS Lab, highlighting their primary objectives and contributions.

## Keywords

Artificial Intelligence, Healthcare, Deep Learning, Machine Learning

## 1. Introduction

Artificial intelligence (AI) mimics human intelligence in machines to carry out tasks involving abstraction and problem-solving. Among the various sectors influenced by AI, healthcare stands out as a highly promising field for its application. Indeed, the integration of AI in healthcare has the potential to aid both patients and healthcare professionals, revolutionizing patient care and administrative operations. Additionally, AI-driven platforms possess the capability to analyze patient data, highlight potential health issues, and enhance diagnostic precision for physicians, particularly in scenarios involving

intricate medical histories or multiple conditions.

AI applications encompass a variety of technologies rather than being confined to a single one. Among these, Machine Learning (ML) stands out as a subset of AI, comprising algorithms that enable systems to autonomously learn and enhance their performance through experience. In the realm of healthcare, traditional ML finds widespread use, notably in precision medicine, where it predicts optimal treatment protocols based on patient attributes and contextual factors.

Deep Learning (DL) is a class of ML algorithms characterized by the use of Artificial Neural Networks (ANNs) that simulate the structure of the human brain. DL approaches have gained popularity in pattern recognition tasks, particularly in image processing, improving medical image analysis. Moreover, in recent years, there has been a notable surge in interest surrounding the application of Large Language Models (LLMs) within the medical domain. LLMs are advanced natural language processing (NLP) models trained on massive amounts of text data, capable of understanding, generating, and processing human language with remarkable accuracy and fluency.

In this paper, we will illustrate some of the projects exploiting AI techniques in the medical field carried out at the University of Naples Federico II node of the CINI-AIIS Lab, highlighting their innovative aspects and contributions.

*Ital-IA 2024: 4th National Conference on Artificial Intelligence, organized by CINI, May 29-30, 2024, Naples, Italy*

\*Corresponding author.

<sup>†</sup>These authors contributed equally.

✉ carlo.sansone@unina.it (C. Sansone)

ORCID 0009-0008-5825-8043 (D. Benfenati); 0000-0002-2578-9349

(S. Capuozzo); 0009-0002-8395-0724 (G. M. D. Filippis);

0009-0001-3740-416X (A. D. Simone); 0000-0001-5033-9617

(M. Gravina); 0009-0006-8134-5466 (L. Marassi);

0000-0001-6852-0377 (S. Marrone); 0000-0002-1778-5321

(E. Masciari); 0000-0002-6384-9891 (E. V. Napolitano);

0000-0001-5425-1890 (G. Pontillo); 0000-0003-1470-8053

(M. Postiglione); 0000-0002-8732-1733 (C. Russo);

0000-0001-9763-8745 (C. Tommasino); 0000-0001-7003-4781

(A. M. Rinaldi); 0000-0002-0754-7696 (V. Moscato);

0000-0002-8176-6950 (C. Sansone)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



## 2. PIE: a novel paradigm for medical recommendations

Existing works in the biomedical domain employing LLMs often overlook the phenomenon of LLM hallucinations [1], which consists in the system generating responses that are either factually incorrect, nonsensical, or disconnected from the input prompt. To address this concern, we introduce a novel paradigm, denoted as Predict→Interpret→Explain (PIE). This paradigm entails employing a model trained on high-quality and controlled data for predictions, interpreting its internal mechanisms using state-of-the-art eXplainable Artificial Intelligence (XAI) techniques, and subsequently utilizing a pool of LLM-based agents as reasoning engines to convey medical recommendations and the interpretation to medical professionals. Below, we outline each step in the PIE pipeline, which we have experimented with in the context of a medication recommendation downstream task.

The Predict module exploits the wealth of information embedded within Electronic Health Records (EHRs) derived from the MIMIC-III dataset [2]. We systematically analyze both the structured and unstructured data present in these records to construct heterogeneous graphs, taking into account the semantic interrelations among various medical concepts and harnessing intrinsic correlations within EHR data. Subsequently, a suite of Graph Neural Networks (GNNs) has been trained for a link prediction task. This task entails discerning forthcoming associations between patients and medications.

The Interpret module is pivotal for elucidating the decision-making process of the GNN model. It consists in understanding the underlying rationales behind a decision. Interpretation serves as an indispensable tool bridging the intricate nature of AI models with the human requisites for transparency and trustworthiness in decision-making processes. In our study, we evaluated the efficacy of the Integrated Gradients [3] and GNNExplainer [4] methodologies, both furnishing node weights that delineate their significance within the predictive context.

In the Explain module, the outcomes derived from the previous phases undergo refinement through a pool of LLM-based collaborative agents, operating in accordance with a specific protocol. Initially, an *internist* physician enlists a panel of specialists (e.g., cardiologists, nephrologists), tailored to the individual patient's characteristics. Each specialist evaluates the patient's characteristics, the predictions from the GNN-based model, and the findings of the Interpret phase to generate a comprehensive assessment, elucidating the justification behind the model's prediction. Then, specialists review the reports of their peers and engage in a discussion that may lead to recon-

sideration of their initial perspectives, thus facilitating the generation of revised reports. Ultimately, the *internist* physician consolidates all reports to produce a unified summary for the human specialist utilizing the system.

## 3. AI in Patient Support: Opportunities and Risks

The capacity of LLMs to interpret and generate natural language with exceptional understanding and contextual awareness has paved the way for innovative approaches to enhancing patient support. One possible application of LLMs is represented by chatbots, virtual assistants capable of creating a welcoming communication environment for patients by providing both emotional support and informative responses, in a way that is similar to a human operator. The natural conversational element and the ability to understand and respond to patients' needs can contribute to making patients feel comfortable communicating with a virtual model, even when aware of its non-human nature. The potential of a model to emulate authentic human interaction had already been explored in the past, in the 1960s, when research in natural language processing led to the development of the ELIZA system, capable of emulating a Rogerian therapist [5]. The Naples' CINI AI-IS node had also pioneered this approach, designing and implementing a chatbot architecture intended to support patients in performing pre-screening procedures [6, 7].

With the advent of Large Language Models (LLMs), this emulation capability is coupled with a remarkable accuracy of responses. The fact that some studies are beginning to show a preference for responses provided by AI models like ChatGPT, in terms of accuracy, nuance, and even empathy, opens up many interesting possibilities for the future of medicine and healthcare. In a recent study, some responses from ChatGPT in the medical field were evaluated as significantly higher quality compared to those of doctors and more empathetic [8]. This latter data appears particularly interesting when considering the potential integration of AI models into healthcare systems to improve doctors' responses to patient inquiries and lighten their workload.

However, this progress also presents significant risks. Considering only the effect on patients, there is concern that the widespread use of AI to provide psychological support may contribute to exacerbating stigma towards certain categories of patients, especially in the psychiatric context, relegating them to "non-human" interactions. Additionally, there is a risk of eroding trust in traditional medicine and perceiving healthcare as impersonal, as well as potentially reducing the responsibility of human operators in expressing empathy and managing interactions with patients, prompting various reflections on the

areas of overlap between chatbots and human operators. If the new possibilities offered by AI in the medical field are manifold, it seems crucial to recognize its limitations and ensure that these technologies are used ethically and under the supervision of qualified professionals. Technology, as a tool and given its current state of development, should integrate rather than replace human interaction in the medical context, ensuring that values such as empathy, clinical judgment, and professional ethics remain at the centre of healthcare. While the combination of human expertise and technological advancements promises to significantly improve healthcare, its success will depend on the ability to effectively balance the skills of both.

#### 4. ASAD Project

In neuroimaging, Deep Learning (DL) has been widely used to model chronological age as a function of brain Magnetic Resonance Imaging (MRI) scans in healthy individuals with excellent results [9]. The extent to which a person deviates from healthy brain-ageing trajectories, expressed as the difference between predicted and chronological age (brain-predicted age difference, brain-PAD), has been proposed as an index of structural brain health, sensitive to brain pathology in a wide spectrum of neurological disorders [10]. However, the presence of subject-specific characteristics in different acquisitions necessitates capturing factors of variation within the target population. In this project, our aim is to propose an innovative DL-based model for age, shape, and appearance disentanglement (ASAD) in brain MRI. This model will enable a more precise quantification of the impact of pathologies on the brain, significantly enhancing the analysis of the brain-PAD. We will employ an Autoencoder architecture consisting of an Encoder to define three latent representations for age, shape, and appearance, respectively. Additionally, a Regressor will be utilized to compute the predicted age, along with a set of three decoders:  $D_a$ ,  $D_s$ , and  $D_t$ . Specifically,  $D_a$  and  $D_s$  will extract the texture (appearance) and the deformation field (shape component), respectively, while  $D_t$  will consider the age-specific latent representation to provide an age-specific template. Following a similar approach as proposed in [11], the ASAD architecture will be trained to reconstruct the input from the disentangled components, using only data from healthy controls with varying age ranges. This enables the network to model a healthy brain-ageing trajectory. Subsequently, the architecture will be applied to heterogeneous patient populations encompassing a wide spectrum of neurological disorders. This application aims to detect disease-specific characteristics within the disentangled components. The disentanglement of age, shape, and appearance may have several

clinical applications, providing deep insights into patient-specific brain characteristics. Indeed, the network's ability to predict the subject's age allows for the evaluation of brain-PAD as a biomarker for a range of neurological diseases, including Multiple Sclerosis, Alzheimer's disease, and schizophrenia. Moreover, the decomposition of brain MRI into its main components of shape and appearance enables the assessment of how a specific subject deviates from an age-specific standard template. When applied to a population of patients with neurological diseases, we expect that the deformation field will capture disease-specific characteristics such as atrophy, while the texture component may be useful for lesion detection. Finally, the ability to disentangle age-related features provides the architecture with generative capacity, allowing for longitudinal evaluation obtained by varying the age-specific latent representation

#### 5. Dementia Severity Assessment with Incomplete Multimodal Data

Alzheimer's disease (AD) is the most common cause of dementia, affecting millions of elderly people around the world. AD is a neurodegenerative disorder, and early detection is a key element to improve the quality of life of affected patients and their families. In clinical trials, Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) are mostly used for the early diagnosis of neurodegenerative disorders since they provide volumetric and metabolic function information of the brain, respectively.

There is the need of combining information from heterogeneous and complementary sources, such as MRI and PET, to evaluate the structural and metabolic characteristics of the brain. This makes the Multimodal Learning well suited in the case of dementia assessment. Techniques for multimodal data fusion can be categorized into early, intermediate, and late fusion. Early fusion integrates multiple sources of data into a single features vector, before being used as input to a learning model. Late fusion, also referred as decision-level fusion, combines according to a given rule the decisions from multiple classifiers, each trained on separate modalities. Intermediate fusion, also named as joint learning, exploits the deep neural networks to transform raw inputs into higher-level and shared representations, which are constructed, for instance, by merging into a single layer, units coming from multiple modality-specific paths. However, when working with a multimodal dataset in the medical field, it is not easy to have images of all the involved modalities, belonging to the same patient. For each subject, *paired acquisition* consists of images coming from all the differ-

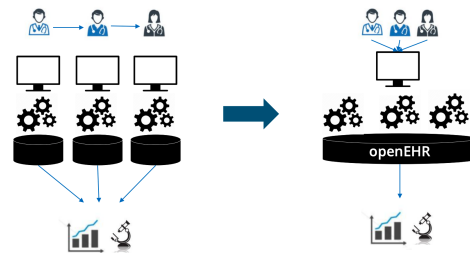
ent sources and collected at the same time or in a specific range. In a real scenario, patients may have *incomplete acquisitions*, in which some modalities are missed.

In the work proposed in [12], we conducted a systematic analysis of early, late and joint approaches in fusion for dementia severity assessment on the publicly available OASIS-3 dataset [13]. We focused on 3D Convolutional Neural Network (CNN) to exploit the volumetric features of the involved images, including in the training step strategies to handle a high imbalance and incomplete dataset. In particular, we analyzed the effects of the incomplete dataset in each multimodal fusion technique, and in the case of intermediate fusion, we proposed a Multiple Input - Multi Output 3D CNN whose training iteration changes according to the characteristics of the input volumes. To further assess the generalization ability of the implemented methodology, we are including the ADNI dataset [14], a study consisting of about 2500 subjects and focusing on the progression of mild cognitive impairment and early AD [15].

## 6. Data-Centric AI for Healthcare

In the age of digital transformation, healthcare is rapidly evolving into a data-driven ecosystem. Imagine a healthcare system where patient records are seamlessly interconnected, diagnosis is made with unprecedented precision, and treatments are tailored in real time. Data-centric architectures are the key to unlocking this visionary healthcare landscape. The transition from a model-centric approach to AI to a data-centric one signifies a shift in emphasis when it comes to creating and implementing AI systems. Model-centric AI aims at producing the best model for a given dataset, whereas data-centric AI aims at systematically and algorithmically producing the best dataset to feed a given ML model.

Current challenges and limitations in health data governance have demonstrated the need for a real digital transformation of healthcare, where decisions are made based on data, whether it is patient history, laboratory results, or imaging data. AI algorithms can assist in identifying high-risk patients, predicting treatment outcomes, and enabling personalized medicine. Nevertheless, ensuring patient data security, AI algorithm accountability, and transparency are crucial to address privacy, security, and bias concerns. Emerging technologies, such as Extended reality (XR) and blockchains, are being already used for improving patient care and guarantee the security and privacy of the data. In this context, the data-centric manifesto serves as a beacon for the healthcare community. Collaboration among clinicians, healthcare organizations, and technology vendors is indispensable in implementing a data-centric approach to coalesce around a common vision, and to ensure that all relevant data is considered



**Figure 1:** Use case: Data-centric approach for digital health transformation

when making decisions, leading to better outcomes for patients. This requires three fundamental steps: integrate, open, innovate: use interoperability standards to integrate existing systems and data. Storing data in an open, vendor-neutral format will then enable ecosystems of vendors to innovate. Real use cases of data-centric architectures for healthcare, such as the one proposed at Karolinska University Hospital, have been already developed (see Figure 1). The adoption of standards like OpenEHR, FHIR, HL7, and ontologies like Snomed CT represent the technical foundations upon which this vision can be realized to achieve semantic and structural interoperability in personal health data, that is to ensure high data quality.

## 7. UNet-based multi-class nuclei segmentation

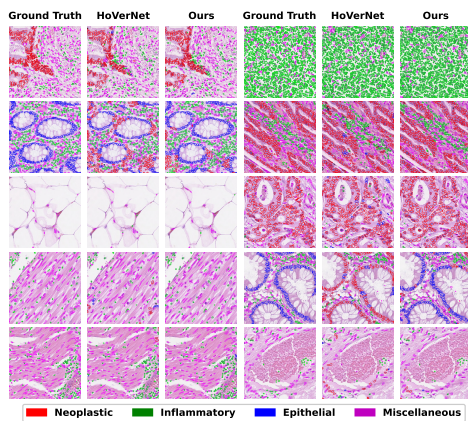
In recent years, the application of AI in Healthcare is increasingly stimulating researchers interests [16, 17]. Nuclei panoptic segmentation, i.e., the simultaneous detection, segmentation, and classification of nuclear instances, is at the core of the automation of several tasks in digital pathology, particularly in the analysis of routine Hematoxylin and Eosin (H&E) stained histology slides.

**Distillation Framework.** In our framework, we adopt an offline technique using HoVerNet as a pre-trained teacher network. Given that HoVerNet performs nuclei instance segmentation and classification through three branches, our distillation strategy is based on the idea of combining all output branches of HoVerNet into a single branch network. Note that we aim to train a student that can replace only the HoVerNet backbone, not its post-processing steps, which we left unvaried. We employ a single-branch UNet as our student model and join all HoVerNet branch outputs into a single branch with a number of output channels equal to the total number of HoVerNet’s branches. In particular, we used a Mix Vision Transformer (MixViT) as the backbone for UNet,

resulting in the best combination based on our experiments. Our loss is a linear combination of the *student loss* between the student and the ground truth and the *distillation loss* between the student and the teacher regulated by  $\alpha$  parameter.

In this work we used two datasets, namely PanNuke [18] for training HoVer-UNet and CoNSEP [19] for validating results on external data. In the case of PanNuke dataset, our solution achieved comparable performance to HoVerNet, demonstrating a significant advantage in terms of processing speed. When the CoNSEP dataset is considered, the results showed that our solution outperforms HoVerNet in terms of panoptic quality, though it falls short in terms of F-score detection. Regarding classification metrics, our solution outperforms HoVerNet across neoplastic and epithelial nuclei; it is practically equal for miscellaneous and worse for inflammatory. Lastly, the inference time is about three times lower.

Figure 2 shows visual examples of the results of HoVerNet and HoVer-UNet compared with the CoNSEP reference standard. Overall, the similarity between the results supports the practical effectiveness of our approach.



**Figure 2:** Nuclei segmentation and classification comparison between CoNSEP ground truth, HoVerNet, and our predictions.

## 8. Infantile Predictors of Functional Gastrointestinal Disorders

Functional Gastrointestinal Disorders (FGIDs) are a significant challenge in pediatric healthcare due to their prevalence and impact on infants. FGIDs refer to a range of conditions, including infant colic, regurgitation, functional diarrhea, and functional constipation, that are de-

finied by the absence of identifiable biochemical or structural anomalies. These conditions affect almost 50% of infants in their first year of life [20]

This study examines the considerable impact of FGIDs on children, their families and healthcare systems, and highlights the historic challenge of identifying children at risk due to unclear pathophysiology. The research aims to identify early-life risk factors for FGIDs [21], within the first year of life. Using a prospective observational cohort design, the study enrolled term and preterm infants from a tertiary care university hospital in Foggia, Italy, between 1 January 2020 and 31 December 2022, excluding infants with severe disease or major neonatal complications. By using conventional statistical methods and Machine Learning (ML), this study identified birth weight, cord blood pH, and maternal age as significant predictors for FGIDs. A logistic regression predictive model also established an inverse relationship between these variables and the occurrence of FGIDs. Using these findings, the study created a ML-based predictive model and a practical, user-friendly web interface for risk assessment. This enables clinicians to identify infants at a higher risk for FGIDs. The approach marks a pioneering step in FGID risk prediction.

## Acknowledgments

This work was supported in part by the Piano Nazionale Ripresa Resilienza (PNRR) Ministero dell’Università e della Ricerca (MUR) Project under Grant PE0000013-FAIR

## References

- [1] M. A. Ahmad, I. Yaramis, T. D. Roy, Creating trustworthy llms: Dealing with hallucinations in healthcare AI, CoRR abs/2311.01463 (2023). URL: <https://doi.org/10.48550/arXiv.2311.01463>. doi:10.48550/ARXIV.2311.01463. arXiv:2311.01463.
- [2] 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, Scientific Data 3 (2016).
- [3] M. Sundararajan, A. Taly, Q. Yan, Axiomatic Attribution for Deep Networks, CoRR abs/1703.01365 (2017). URL: <http://arxiv.org/abs/1703.01365>. arXiv:1703.01365.
- [4] Z. Ying, D. Bourgeois, J. You, M. Zitnik, J. Leskovec, GNNExplainer: Generating Explanations for Graph Neural Networks, in: H. M. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. B. Fox, R. Garnett (Eds.), Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019,

- Vancouver, BC, Canada, 2019, pp. 9240–9251. URL: <https://proceedings.neurips.cc/paper/2019/hash/d80b7040b773199015de6d3b4293c8ff-Abstract.html>.
- [5] R. Bowman, O. Cooney, J. W. Newbold, A. Thieme, L. Clark, G. Doherty, B. Cowan, Exploring how politeness impacts the user experience of chatbots for mental health support, *International Journal of Human-Computer Studies* 184 (2024) 103181.
  - [6] F. Amato, S. Marrone, V. Moscato, G. Piantadosi, A. Picariello, C. Sansone, et al., Chatbots meet ehealth: Automatizing healthcare., in: *WIAIAH@AI\* IA*, 2017, pp. 40–49.
  - [7] F. Amato, S. Marrone, V. Moscato, G. Piantadosi, A. Picariello, C. Sansone, Holmes: Ehealth in the big data and deep learning era, *Information* 10 (2019) 34.
  - [8] J. W. Ayers, A. Poliak, M. Dredze, E. C. Leas, Z. Zhu, J. B. Kelley, D. J. Faix, A. M. Goodman, C. A. Longhurst, M. Hogarth, et al., Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum, *JAMA internal medicine* 183 (2023) 589–596.
  - [9] M. Tanveer, M. Ganaie, I. Beheshti, T. Goel, N. Ahmad, K.-T. Lai, K. Huang, Y.-D. Zhang, J. Del Ser, C.-T. Lin, Deep learning for brain age estimation: A systematic review, *Information Fusion* (2023).
  - [10] J. H. Cole, R. P. Poudel, D. Tsagkrasoulis, M. W. Caan, C. Steves, T. D. Spector, G. Montana, Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker, *NeuroImage* 163 (2017) 115–124.
  - [11] Z. Shu, M. Sahasrabudhe, R. A. Guler, D. Samaras, N. Paragios, I. Kokkinos, Deforming autoencoders: Unsupervised disentangling of shape and appearance, in: *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 650–665.
  - [12] M. Gravina, A. García-Pedrero, C. Gonzalo-Martín, C. Sansone, P. Soda, Multi input–multi output 3d cnn for dementia severity assessment with incomplete multimodal data, *Artificial Intelligence in Medicine* 149 (2024) 102774.
  - [13] P. J. LaMontagne, T. L. Benzinger, J. C. Morris, S. Keefe, R. Hornbeck, C. Xiong, E. Grant, J. Hasenstab, K. Moulder, A. G. Vlassenko, et al., Oasis-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and alzheimer disease, *MedRxiv* (2019).
  - [14] C. R. Jack Jr, M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. J. Britson, J. L. Whitwell, C. Ward, et al., The alzheimer’s disease neuroimaging initiative (adni): Mri methods, *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* 27 (2008) 685–691.
  - [15] A. De Simone, C. Sansone, A multimodal deep learning based approach for alzheimer’s disease diagnosis, in: *International Conference on Image Analysis and Processing*, Springer, 2023, pp. 131–139.
  - [16] C. Tommasino, F. Merolla, C. Russo, S. Staibano, A. M. Rinaldi, Histopathological image deep feature representation for cbir in smart pacs, *Journal of Digital Imaging* 36 (2023) 2194–2209.
  - [17] A. M. Rinaldi, C. Russo, C. Tommasino, Effects of color stain normalization in histopathology image retrieval using deep learning, in: *2022 IEEE International Symposium on Multimedia (ISM)*, IEEE, 2022, pp. 26–33.
  - [18] J. Gamper, N. A. Koohbanani, K. Benes, S. Graham, M. Jahanifar, S. A. Khurram, A. Azam, K. Hewitt, N. Rajpoot, Pannuke dataset extension, insights and baselines, *arXiv preprint arXiv:2003.10778* (2020).
  - [19] S. Graham, Q. D. Vu, S. E. A. Raza, A. Azam, Y. W. Tsang, J. T. Kwak, N. Rajpoot, Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images, *Medical image analysis* 58 (2019) 101563.
  - [20] A. Chogle, C. A. Velasco-Benitez, I. J. Koppen, J. E. Moreno, C. R. R. Hernández, M. Saps, A population-based study on the epidemiology of functional gastrointestinal disorders in young children, *The Journal of pediatrics* 179 (2016) 139–143.
  - [21] D. Bi, H. Jiang, K. Yang, T. Guan, L. Hou, G. Shu, Neonatal risk factors for functional gastrointestinal disorders in preterm infants in the first year of life (2022).