Advancing e-health with AI: Insights from our research experience in neuroimaging, acoustic signals, and vital parameter monitoring

Gabriella **Casalino**, Giovanna **Castellano***, Gennaro **Vessio** and Gianluca **Zaza**

Department of Computer Science, University of Bari Aldo Moro, Italy

Abstract

This contribution briefly describes the research being carried out in the Computational Intelligence Laboratory of the Department of Computer Science, University of Bari Aldo Moro, in AI-based e-health. Our research encompasses a wide array of methodologies and applications aimed at leveraging the capability of AI to empower the diagnosis, monitoring, and treatment of various health conditions. Through multifaceted research that covers neuroimaging analysis, acoustic signal processing, and vital parameter monitoring, our goal is to shed light on the potential of AI in enhancing healthcare services.

Keywords

artificial intelligence, explainability, e-health, neuroimaging, acoustic signals, vital parameters

1. Introduction

Integrating Artificial Intelligence (AI) into healthcare is transforming how we diagnose, monitor, and care for patients. At the Computational Intelligence Laboratory (CILab) of the Department of Computer Science, University of Bari Aldo Moro, we are contributing to this transformation by applying AI to neuroimaging, acoustic signal analysis, and vital signs monitoring. Our work aims to address current healthcare challenges by developing innovative and practical AI solutions.

This paper presents an overview of our research efforts and achievements in these areas. By sharing our findings and methodologies, we aim to highlight AI's significant impact on improving healthcare services and patient outcomes. Our goal is to showcase our work and encourage ongoing innovation and dialogue in the rapidly evolving field of AI in healthcare.

2. Neuroimaging

Neuroimaging is a pivotal area within our research portfolio, where the application of AI-driven algorithms plays a crucial role in enabling the early and precise diagnosis of neurological disorders.

*Corresponding author.
✉ gabriella.casalino@uniba.it (G. Casalino);
giovanna.castellano@uniba.it (G. Castellano);
gennaro.vessio@uniba.it (G. Vessio); gianluca.zaza@uniba.it
(G. Zaza)
🆔 0000-0003-0713-2260 (G. Casalino); 0000-0002-6489-8628
(G. Castellano); 0000-0002-0883-2691 (G. Vessio);
0000-0003-3272-9739 (G. Zaza)

2.1. Alzheimer's disease detection

Dementia, with Alzheimer's disease (AD) being its most common form, poses a significant global health challenge, especially among the aging population. It currently affects around 55 million people worldwide, predominantly in low- and middle-income countries, and this number is expected to increase as the global population ages. Unfortunately, effective cures remain elusive, with available treatments focusing more on symptom management than addressing the underlying causes. This underscores the critical need for early and accurate diagnosis to improve patient care.

AI, mainly through advanced machine and deep learning techniques, is increasingly recognized for its potential to revolutionize the diagnosis of dementia, including AD. Neuroimaging techniques, such as Magnetic Resonance Imaging (MRI) and amyloid Positron Emission Tomography (PET) scans, have been identified as promising tools for early detection. MRI provides detailed images of the brain, enabling the identification of brain atrophy patterns characteristic of AD. At the same time, amyloid PET scans offer insights into the pathophysiology of the disease by detecting amyloid plaques in the brain, a hallmark of AD.

Our research has advanced the application of Convolutional Neural Network (CNN) models for the automated diagnosis of AD, using the strengths of both MRI and PET scans [1]. We examined these neuroimaging techniques' efficacy in uni-modal and multi-modal setups, underscoring the advantage of integrating data from diverse modalities to refine diagnostic precision. Additionally, we incorporated an explainable AI method to address the demand for transparency in medical AI applications, offering insights into the AI-driven diagnostic process and contributing to a deeper understanding of AD's un-
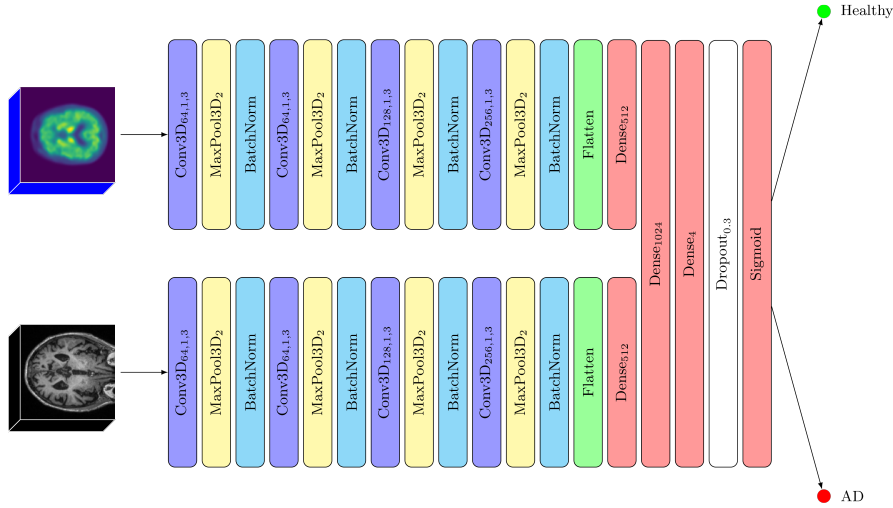
Figure 1: Diagram illustrating the proposed multi-modal CNN architecture designed for AD detection, which simultaneously processes 3D MRI and PET scan inputs for enhanced diagnostic accuracy.

derlying mechanisms.

Specifically, our investigation has yielded several key insights, highlighting the potential of multi-modal imaging strategies. By analyzing the classification results from various model configurations on the OASIS-3 benchmark dataset, we discovered that models utilizing 3D inputs consistently outperformed those using 2D inputs, likely due to the richer spatial information available in 3D scans. Moreover, regardless of whether in 2D or 3D, MRI scans significantly surpassed amyloid PET scans in diagnostic performance, emphasizing MRI's inherent value in AD detection within our study. However, multi-modal strategies, particularly our "fusion" model (shown in Fig. 1), demonstrated a clear advantage, achieving up to 95% accuracy. This underlines the complementary nature of MRI and PET scans in AD diagnosis.

Our adoption of the Grad-CAM technique further allowed us to pinpoint the brain regions most relevant for classification, offering valuable insights into the neuroanatomical underpinnings of AD. This supports the validity of our models and enhances our understanding of AD's neuropathology.

In another exploratory study [2], we delved into Diffusion Tensor Imaging (DTI), a sophisticated MRI technique that assesses the integrity of white matter fiber tracts in the brain. We explored fractional anisotropy (FA), a DTI metric that measures the uniformity of water diffusion directionality, which exhibits notable changes in AD patients, suggesting its utility as a diagnostic indicator.

We introduced a dual-stage deep learning approach, combining unsupervised and supervised techniques. Initially, a 3D convolutional autoencoder was employed to extract low-dimensional representations from FA images

in an unsupervised manner. These representations were then used to train a supervised 3D CNN for AD detection. This innovative strategy demonstrated encouraging outcomes on the OASIS-3 dataset and lessened the dependence on extensively annotated datasets, setting the stage for more autonomous and quantitative AD detection in clinical practice. Our future endeavors will aim to assess this method across broader and more varied datasets to affirm its diagnostic validity further.

2.2. Brain tumor segmentation

Brain tumor segmentation from MRI scans is crucial for accurate diagnosis, treatment planning, and patient monitoring. Recent strides in deep learning, particularly with CNNs, have significantly advanced the automation and precision of tumor segmentation. Nevertheless, these models' "black-box" nature raises challenges in explainability—a vital aspect of clinician trust and decision-making.

Graph Neural Networks (GNNs) have recently gained attention as a novel approach to medical image analysis, offering an alternative that might bridge the gap between accuracy and explainability. By conceptualizing brain images as graphs—where nodes represent voxels or regions of interest, and edges depict spatial relationships—GNNs use relational dependencies to achieve fine segmentation. This capability to capture both local and global contexts through message passing between nodes positions GNNs as a promising tool for achieving high precision in brain tumor segmentation and providing a pathway to model understandability.

In a recent study [3], we analyzed GNN models for seg-

menting brain tumors, focusing on their explainability. Using GNNExplainer, we aimed to improve the transparency of GNN models, making their decision-making processes accessible and understandable to clinicians.

Our exploration highlighted the effectiveness of GNNs in medical imaging. It also laid the foundation for future research, suggesting potential synergies between GNNs and CNNs, such as integrating GNNs with 3D U-Net architectures, to refine segmentation results further. In addition, collaboration with medical experts to examine critical features identified by GNNExplainer could further solidify the role of GNNs in clinical practice, combining accuracy and explainability in brain tumor management.

3. Acoustic signals

The analysis of vocal characteristics from speech samples is an effective approach to identifying conditions associated with mental diseases, notably bipolar disorder (BD). Our research focused on extracting acoustic features from patients with BD using a specialized mobile application, developed at the Department of Affective Disorders, Institute of Psychiatry and Neurology in Warsaw, Poland, under the project "Smartphone-based diagnostics of phase changes in the course of bipolar disorder". BD manifests through fluctuating mood states, including euthymia, depression, mixed states, and mania, traditionally diagnosed through regular consultations using standard psychiatric tools like the Hamilton Depression Rating Scale (HDRS) and the Young Mania Rating Scale (YMRS). These instruments allow healthcare providers to detect symptoms and evaluate the intensity of depressive and manic episodes, facilitating precise diagnoses and the development of customized treatment strategies.

Our research examined several critical dimensions of data related to BD, specifically focusing on the importance of continuous monitoring to track temporal fluctuations. We addressed the challenge of missing labels while also handling the uncertainty in labeling due to the inherent ambiguity and variability in data classification. Moreover, we worked on generating readily understandable explanations of BD state classifications leveraging the availability of multi-layered information.

3.1. Monitoring bipolar disorder states

The specialized application captures acoustic features daily, but patient assessments are less frequent, resulting in a scarcity of labeled data. This gap leaves many acoustic features without clear annotations of the patient's condition.

We explored semi-supervised learning algorithms to harness the geometric data properties and the predefined knowledge of patient states. These algorithms have shown promise in enhancing classification accuracy, even with limited labeled data [4].

We introduced a novel algorithm, the Dynamic Incremental Semi-Supervised Fuzzy C-Means (DISSFCM), designed to monitor BD states while considering the temporal acquisition of acoustic features. DISSFCM, an extension of the Semi-Supervised Fuzzy C-Means (SSFCM) algorithm, analyses data chunks sequentially in near real-time, maintaining historical data insights without extensive storage. It adapts to new information, refining the classification through an increased cluster count representing the patient's condition states. This method has proven effective in predicting episodes of health and illness with as little as 25% labeled data [5].

DISSFCM operates on labeled prototypes, summarizing data clusters for each segment. It generates membership matrices, clarifying each data point's cluster association and facilitating outcome explanation. Initially, we applied visual analytics for interpretation [6], advancing to natural language explanations or *linguistic summaries*, which translate complex data relations into understandable sentences [7]. For example, we could deduce that "Most calls in the state of hypomania have low loudness compared to the state of euthymia". This approach segments acoustic features into semantic categories—loudness, pitch, spectrum, and voice quality—guided by psychiatric expertise. Our experiments have demonstrated the practical application of linguistic summaries as informative granules for smartphone-based BD monitoring. They offer clear, insightful linguistic descriptions, making the complex data and sparse psychiatric evaluations comprehensible.

3.2. Explaining bipolar disorder states

We designed a versatile, multi-task neural network to leverage the detailed symptom information captured during patient assessments. This network is trained to generate several outputs, each aligning with the various levels of labels obtained from intermediate assessment stages. These intermediate outputs fulfill dual roles: they enhance the model's overall predictive accuracy and provide insights into classifying mid-level labels. Our architecture, designed to handle data with a hierarchical class structure, is a crucial component of PLENARY (exPlaining bLack-box modEls in Natural lAnguage thRough fuzzY linguistic summaries) [8]. PLENARY aims to categorize tabular data across different class levels and render the model's explanations into natural language, employing fuzzy linguistic summaries for clarity.

In collaboration with a neuropsychiatrist, we identified ten critical symptoms as intermediate labels, including anxiety, decreased activity, mood changes, disorganization, and sleep disorders, among others. The model's outcomes and explanations focus on the patient's state

and these specific symptoms. For instance, we found that "Among records that contribute positively to predicting mania, most of them have spectral-related features at low level" and "Among records that contribute against predicting decreased activity, most of them have quality-related features at low level".

Through rigorous experimental evaluation, we have demonstrated that augmenting model explanations with fuzzy linguistic summarization—especially those derived from SHAP analyses—significantly enhances understanding of the model's predictions. This approach effectively combines domain-specific knowledge with technical insight, providing a comprehensive and accessible explanation framework.

4. Vital parameters

Our research into vital parameter monitoring leverages AI to anticipate important diseases, equipping patients and physicians with critical insights for preemptive healthcare management. Herein, we detail our efforts in remote vital parameter estimation and creating eXplainable AI (XAI) models to support medical diagnosis. These models use vital signs data to aid medical professionals in the early detection of cardiovascular diseases and stress-related conditions.

4.1. Contact-less monitoring of vital parameters

Our endeavors in vital parameter monitoring have been concentrated on heart rate, breathing rate, blood oxygen saturation (SpO$_2$), and systolic and diastolic blood pressure—key indicators for cardiovascular health. Traditional methods, like ECG, require direct skin contact, often necessitating cumbersome wearable devices. To overcome the limitations and discomfort of contact-based monitoring, advancements have been made toward developing photoplethysmography (PPG) techniques that operate using camera-based systems. However, these can be expensive and not user-friendly for daily home use.

Addressing these challenges, our lab has developed an innovative, cost-effective approach for monitoring cardiovascular parameters that seamlessly integrates into everyday living environments [9]. This system employs a non-invasive, contactless device consisting of a transparent mirror equipped with a camera that identifies the user's face and uses remote photoplethysmography (rPPG) to analyze video frames. The prototype of the smart mirror is shown in 2(a). This method calculates vital parameters like blood oxygen saturation, heart rate, and breathing rate and includes a novel technique for automatic lip color detection through clustering-based color quantization. With this new method, we aim to relieve individuals from the discomfort of traditional contact-based monitoring, making it a more convenient and user-friendly option for daily home use.

Our methodological pipeline (shown in Fig. 2(b)) initiates with the detection of the subject's face, focusing specifically on the forehead as the region of interest (ROI) for signal extraction. The rPPG signal is then processed using Independent Component Analysis (ICA) and Fast Fourier Transform (FFT) to estimate HR and BR. At the same time, SpO$_2$ measurements are derived by applying the Beer-Lambert law. For lip color detection, the system identifies the lip ROI and determines the dominant color using clustering methods. Our contactless approach has not only demonstrated measurement accuracy within acceptable ranges for both stationary and minimally moving subjects, but it has also shown superior performance compared to traditional contact devices, instilling confidence in its reliability and accuracy.

Further enhancements included the addition of new ROIs and a face-tracking feature to accommodate head movements, improving usability on mobile devices [10, 11]. This comprehensive framework is adaptable to any camera-equipped device, leading to the creation of a smartphone application that facilitates easy, widespread monitoring of vital health parameters [12].

4.2. Cardiovascular risk assessment

While traditional machine learning algorithms have significantly aided physicians in diagnosing symptoms early to prevent disease progression, their often opaque nature presents a challenge. These "black box" models deliver accurate predictions but lack an intuitive explanation for their results. This makes them less practical in fields where end-users are non-technical professionals, notably in healthcare.

Our work concentrates on advancing XAI models, which are mainly aimed at supporting medical decisions in cardiovascular disease (CVD) assessment. CVDs are a primary global health concern, responsible for approximately 17.9 million deaths annually,[1] spanning conditions such as coronary heart disease and stroke. Given the multifactorial causes of CVDs, including lifestyle and genetic predispositions, early intervention and continuous monitoring of vital signs are crucial to prevention.

In our efforts, we have developed a fuzzy rule-based system to assist clinicians in evaluating cardiovascular risks with greater interpretability [13]. This system utilizes IF-THEN rules, a natural language format that simplifies understanding and application, incorporating patient data like heart rate and blood oxygen saturation to estimate CVD risk. Developed in collaboration with medical experts, this model prioritizes accuracy while
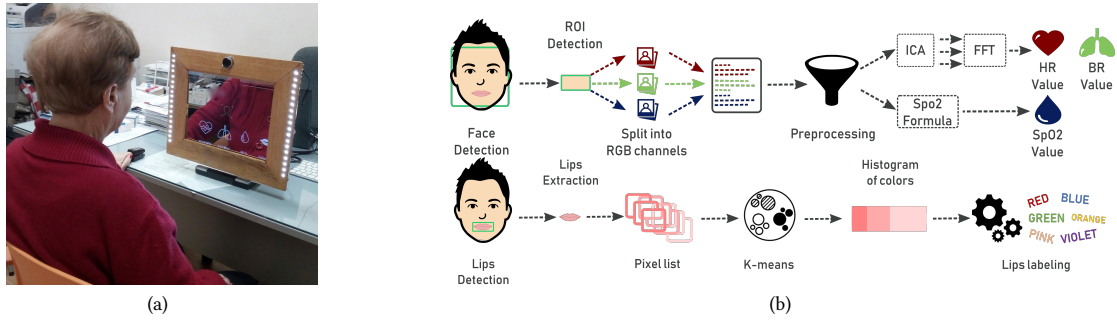
---

[1]https://www.who.int/health-topics/cardiovascular-diseases

Figure 2: (a) Prototype of the smart mirror developed in our lab and (b) methodological pipelines for vital sign measurement.

ensuring user-friendly interpretability, offering a slight trade-off in precision for much greater transparency.

To bridge the gap between data-driven precision and expert intuition, we explored neuro-fuzzy systems, which automate the generation of fuzzy rule-based models from data, streamlining the otherwise manual and labor-intensive process of rule formation. Our research demonstrates that models created through neuro-fuzzy systems maintain accuracy and significantly enhance interpretability, outperforming manually designed models in cardiovascular risk prediction [14].

Expanding beyond cardiovascular health, we have applied neuro-fuzzy systems to diagnose hypertension and stress, focusing on minimizing complexity for clearer understanding. We have balanced accuracy and interpretability by employing feature selection to refine the number of relevant indicators and fuzzy rules, making these models highly practical for real-world medical applications [15].

Acknowledgments

References

[1] G. Castellano, A. Esposito, E. Lella, G. Montanaro, G. Vessio, Automated detection of Alzheimer's disease: a multi-modal approach with 3D MRI and amyloid PET, Scientific Reports 14 (2024) 5210.

[2] G. Castellano, E. Lella, V. Longo, G. Placidi, M. Polsinelli, G. Vessio, Combining Unsupervised and Supervised Deep Learning for Alzheimer's Disease Detection by Fractional Anisotropy Imaging, in: 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS), IEEE, 2023, pp. 511–516.

[3] D. Amendola, A. Basile, G. Castellano, G. Vessio, G. Zaza, From Voxels to Insights: Exploring the Effectiveness and Transparency of Graph Neural Networks in Brain Tumor Segmentation, in: International Joint Conference on Neural Networks (IJCNN 2024), IEEE, to appear.

[4] G. Casalino, G. Castellano, O. Hryniewicz, D. Leite, K. Opara, W. Radziszewska, K. Kaczmarek-Majer, Semi–Supervised vs. Supervised Learning for Mental Health Monitoring: A Case Study on Bipolar Disorder, International Journal of Applied Mathematics and Computer Science 33 (2023) 419–428.

[5] G. Casalino, G. Castellano, F. Galetta, K. Kaczmarek-Majer, Dynamic incremental semi-supervised fuzzy clustering for bipolar disorder episode prediction, in: International Conference on Discovery Science, Springer, 2020, pp. 79–93.

[6] G. Casalino, G. Castellano, K. Kaczmarek-Majer, O. Hryniewicz, Intelligent analysis of data streams about phone calls for bipolar disorder monitoring, in: 2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), IEEE, 2021, pp. 1–6.

[7] K. Kaczmarek-Majer, G. Casalino, G. Castellano, O. Hryniewicz, M. Dominiak, Explaining smartphone-based acoustic data in bipolar disorder: Semi-supervised fuzzy clustering and relative linguistic summaries, Information Sciences 588 (2022) 174–195.

[8] K. Kaczmarek-Majer, G. Casalino, G. Castellano, M. Dominiak, O. Hryniewicz, O. Kamińska, G. Vessio, N. Díaz-Rodríguez, PLENARY: Explaining black-box models in natural language through fuzzy linguistic summaries, Information Sciences 614

(2022) 374–399.

[9] G. Casalino, G. Castellano, V. Pasquadibisceglie, G. Zaza, Contact-less real-time monitoring of cardiovascular risk using video imaging and fuzzy inference rules, Information 10 (2018) 9.

[10] G. Casalino, G. Castellano, G. Zaza, A mHealth solution for contact-less self-monitoring of blood oxygen saturation, in: 2020 IEEE Symposium on Computers and Communications (ISCC), IEEE, 2020, pp. 1–7.

[11] G. Casalino, G. Castellano, G. Zaza, Evaluating the robustness of a contact-less mHealth solution for personal and remote monitoring of blood oxygen saturation, Journal of Ambient Intelligence and Humanized Computing (2023) 1–10.

[12] G. Casalino, G. Castellano, A. Nisio, V. Pasquadibisceglie, G. Zaza, A mobile app for contactless measurement of vital signs through remote Photoplethysmography, in: 2022 IEEE international conference on systems, man, and cybernetics (SMC), IEEE, 2022, pp. 2675–2680.

[13] G. Casalino, G. Castellano, C. Castiello, V. Pasquadibisceglie, G. Zaza, A fuzzy rule-based decision support system for cardiovascular risk assessment, in: Fuzzy Logic and Applications: WILF 2018, Springer, 2019, pp. 97–108.

[14] G. Casalino, G. Castellano, U. Kaymak, G. Zaza, Balancing accuracy and interpretability through neuro-fuzzy models for cardiovascular risk assessment, in: 2021 IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, 2021, pp. 1–8.

[15] G. Casalino, G. Castellano, G. Zaza, Interpretable Neuro-Fuzzy Models for Stress Prediction, in: Conference of the European Society for Fuzzy Logic and Technology, Springer, 2023, pp. 630–641.