NEXT_AIM: descrizione ed obiettivi

[INFN-CSN5, 2022-2024]

Artificial Intelligence in Medicine (AIM): next *steps*
focus on no-so-big data and explainable techniques

Istituto Nazionale di Fisica Nucleare
Sezione di PISA

https://www.pi.infn.it/aim/

Involves 13 INFN groups

Resp. Nazionale: A. Retico

Resp. Locali:
- Bari (S. Tangaro)
- **Bologna (D. Remondini)**
- Cagliari (P. Oliva)
- Catania (M. Marrale)
- Ferrara (G. Paternò)
- Firenze (C. Talamonti)
- Genova (A. Chincarini)
- Lab. Naz. Sud (G. Russo)
- Milano (C. Lenardi)
- Napoli (G. Mettivier)
- Padova (A. Zucchetta) dal 2023
- Pavia (A. Lascialfari)
- Pisa (M.E. Fantacci)

*developing **robust** and **explainable** AI algorithms and validating them on realistic use cases in the **medical field***

The **next AIM** experiment aims to address the following specific challenges related to methodological aspects of the application of AI in Medicine (AIM):
1) manage limited datasets with AI techniques (**n**o-so-big dataset);
2) make solutions provided by AI models understandable by humans (explainable AI).

next_AIM - Artificial Intelligence in Medicine: *next steps*

https://www.pi.infn.it/aim/

AIM [2019-2021]
next_AIM [2022-2024]

Collaboration with Italian medical centers (hospitals and IRCCS) and with international consortia for data sharing

WP1 — Challenge I: no-so-big data

Strategies for efficient learning with limited data samples.

Evaluation of robustness and reliability of trained models.

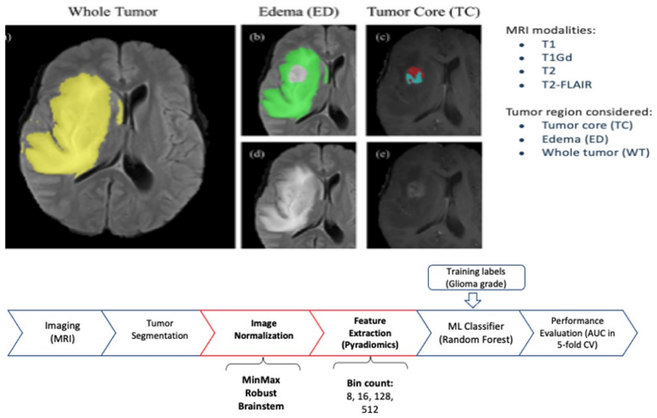
WP2 — Challenge II: explainable AI (XAI)

Make AI results understandable to humans.

Which image/data features are relevant to make a decision?

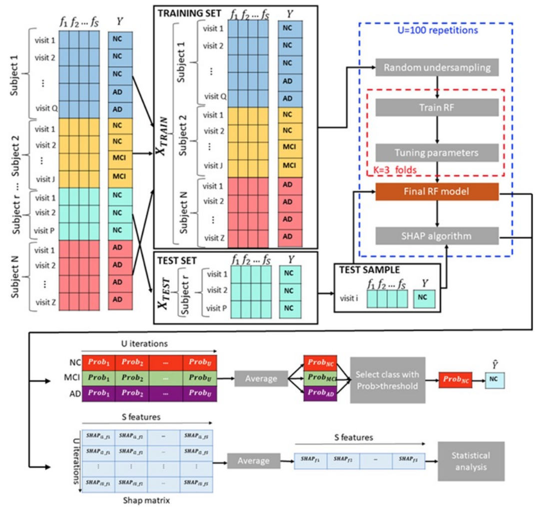
WP3 — Applications to real-world data samples include ….

Evaluation of the robustness of radiomic features in multiparametric MRI and its impact on predictive value of AI models



Ubaldi L, Saponaro S, Giuliano A, Talamonti C, Retico A. Deriving quantitative information from multiparametric MRI via Radiomics… *Phys Medica* 2023;107:102538 https://doi.org/10.1016/j.ejmp.2023.102538.

Robust implementation of explainable AI methods



Lombardi A, Diacono D, Amoroso N, Biecek P, Monaco A, Bellantuono L, … Tangaro S, Bellotti R. A robust framework to investigate the reliability and stability of explainable artificial intelligence …. *Brain Informatics* 2022; 9:17. https://doi.org/10.1186/s40708-022-00165-5.

WP4

Computing resources and SW repository organization

ReCaS, IBiSCo, INFN-Cloud + local HW resources

WP5

Exploitation of research results and communication

conferences, publications and outreach, collaboration with aifm

NEXT_AIM: stato di avanzamento generale delle milestones

Istituto Nazionale di Fisica Nucleare
Sezione di PISA

MILESTONES

2022		
31 Dec	M1.1	Identification of methodological pitfalls in case of small datasets
31 Dec	M2.1	Identification of explainability requirements for medical applications
31 Dec	M3.1	Identification of data samples for practical use cases and fist tests
30 Jun	M4.1a	Identification of available resources and usage instructions
31 Dec	M4.1b	SW package release instructions
31 Dec	M5.1	~~Workshop organization: "AI methods and applications in Medical Physics"~~
2023		
31 Dec	M1.2	Definition of robust pipelines for efficient model training on small datasets
31 Dec	M2.2	Customization of explainability pipelines to AI models for medical imaging
31 Dec	M3.2	Implementation of robust pipelines and explainability algorithms in at least three different use cases
31 Dec	M4.2	Integration of at least 1 application per site in **next**AIM SW package repository
31 Dec	M5.2	~~Workshop organization: "The right to explanation"~~
2024		
30 Jun	M2.3	Definition of optimal explainability methodology for medical problems
31 Dec	M3.3	Result evaluation for the practical use case and reporting
31 Dec	M4.3	Integration of all analysis pipelines trained for the use cases of WP3 in the **next**AIM SW package repository
31 Dec	M5.3	Submission of at least 1 scientific publication per use case

31 Dec M2.4 Workshop "The right to explanation"

Planned for **October 2024 in Bari**

WS AI@INFN, Bologna 2-3 Maggio 2022
https://agenda.infn.it/event/29907/

State of the art and challenges in AI
Speaker: Daniel Remondini

Deep Learning in Medical Image Analysis
Speaker: Francesca Lizzi

Trustworthy AI in medical applications
Speaker: Angela Lombardi

The Workshop "AI methods and applications in Medical Physics" October 2024.
2023: 3-day General Collaboration Meeting in Milan in February 2023
https://agenda.infn.it/event/34599/

List of tasks of WP3 Bologna

Sedi partecipanti													Task	Topic
	BO			FE			LNS		NA		PI		**T1**	Radiomics in **Digital Breast Tomosynthesis** (DBT)
	BO			FE					NA		PI		**T2**	**Super-Resolution in Medical Imaging**
	BO		CT										**T3**	Radiomics in prostate **cancer**
	BO		CT										**T4**	Radiomics and DL in tcMRgFUS
	BO				FI	GE	LNS						**T5**	**Nuclear Imaging Quantification and Radiomics**

eNeuro

Research Article: New Research
Novel Tools and Methods

Generalizing the Enhanced-Deep-Super-Resolution Neural Network to Brain MR Images: A Retrospective Study on the Cam-CAN Dataset

Cristiana Fiscone,[1*] Nico Curti,[2*] Mattia Ceccarelli,[3] Daniel Remondini,[2,4] Claudia Testa,[2,4] Raffaele Lodi,[1,5] Caterina Tonon,[1,5] David Neil Manners,[5,6**] and Gastone Castellani[7**]

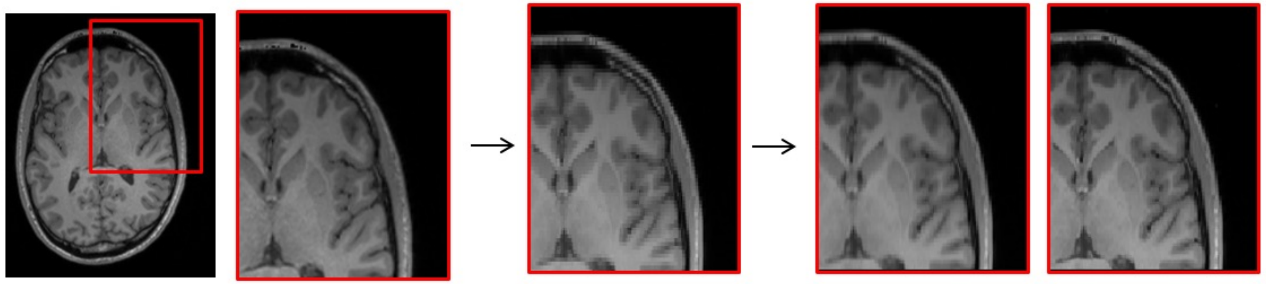
eNeuro 2024 paper: super-resolution for brain MR images

May 2024, 11(5). DOI: https://doi.org/10.1523/ENEURO.0458-22.2023. 1 of 12

CamCAN dataset
EDSR > bicubic

Background:
Super-resolution models are deep learning algorithms which enhance image spatial resolution. Their use on biomedical images has been explored, through *ad hoc* training stages [1]. EDSR (Enhanced Deep Super Resolution) [2] and WDSR (Wide activation for efficient and accurate Deep Super Resolution) [3] are convolutional neural networks, trained on general purpose images, which perform 2x- and 4x- upsampling. This work aims to validate their application to MR brain images, comparing the results with traditional upsampling methods.

Methods:
Data used in this work were provided by the Cambridge Centre for Ageing and Neuroscience (CamCAN) [4],[5].
3D sagittal high-resolution T1w and T2w images (3T Siemens Magnetom Trio, 1mm isotropic) of 70 subjects were convolved with a Gaussian filter and then down-sampled. EDSR and WDSR were used to up-sample low-resolution images. The reconstruction time for each image was ~ 1 min and ~ 15mins respectively (Ubuntu 20.04.2 LTS, 80 processors Intel(R) Xeon(R) Gold 6138 CPU, 20 cores each). Byron was used as custom library, released with MIT license and available on Github [6]. The results were compared to bicubic interpolation upsampling. The processing pipeline is shown in Fig. 1.
After the brain extraction, pixel-wise and whole-brain average analysis was performed. RMSE (Root Mean Square Error), pSNR (peak Signal-to-Noise Ratio), SSIM (Structural SIMilarity index) and HFEN (High Frequency Error Norm) were chosen as quantitative similarity parameters, and they were evaluated over the entire T1w and T2w images reconstructed by different upsampling techniques for all the subjects, using the original high-resolution images as ground truth. Since the two models work with 2D images, sagittal, coronal and axial directions were analyzed separately.



ORIGINAL HIGH RESOLUTION IMAGE — DOWNSAMPLED IMAGE – BICUBIC INTERPOLATION + GAUSSIAN FILTER — UPSAMPLED IMAGE – BICUBIC INTERPOLATION — UPSAMPLED IMAGE – EDSR – 2X

Results:
EDSR generally shows better performance than bicubic interpolation.
- T1w images: there is significant difference in favour of EDSR for all the considered criteria in sagittal, coronal and axial reconstructions, confirmed by both *p-value* and *Coehn's d*
- T2w images: there is significant difference in favour of EDSR for two out of four criteria (SSIM and HFEN) in sagittal, coronal and axial reconstructions confirmed by both *p-value* and *Coehn's d*. The trend of the other two criteria (RMSE and pSNR) is similar in the two upsampling methods leading to a not significant difference

No correlations were found between similarity parameters and subjects attributes (sex, age, handedness and total intracranial volume).
WDSR was not found to be suitable, since it enhances and creates line-like artifacts.

p-value		RMSE	pSNR	SSIM	HFEN
YZ – Sagittal	T1w	<.000 * (EDSR)	<.000 * (EDSR)	<.000 * (EDSR)	<.000 * (EDSR)
	T2w	0.711 (BC)	0.860 (BC)	<.000 * (EDSR)	<.000 * (EDSR)
XY - Axial	T1w	<.000 * (EDSR)	<.000 * (EDSR)	<.000 * (EDSR)	<.000 * (EDSR)
	T2w	.389 (EDSR)	.446 (EDSR)	<.000 * (EDSR)	<.000 * (EDSR)
XZ - Coronal	T1w	<.000 * (EDSR)	<.000* (EDSR)	<.000 * (EDSR)	<.000 * (EDSR)
	T2w	<.008 * (EDSR)	<.003* (EDSR)	<.000 * (EDSR)	<.000 * (EDSR)

Conclusions:
EDSR, that performs 2x-upsampling, outperforms the bicubic interpolation without fine-tuning, showing its ability of transfer learning. It is flexible with respect the analyzed MR sequence and subject characteristics.
In particular, in T1w images it shows significant better performance by all the considered metrics.

diagnostics

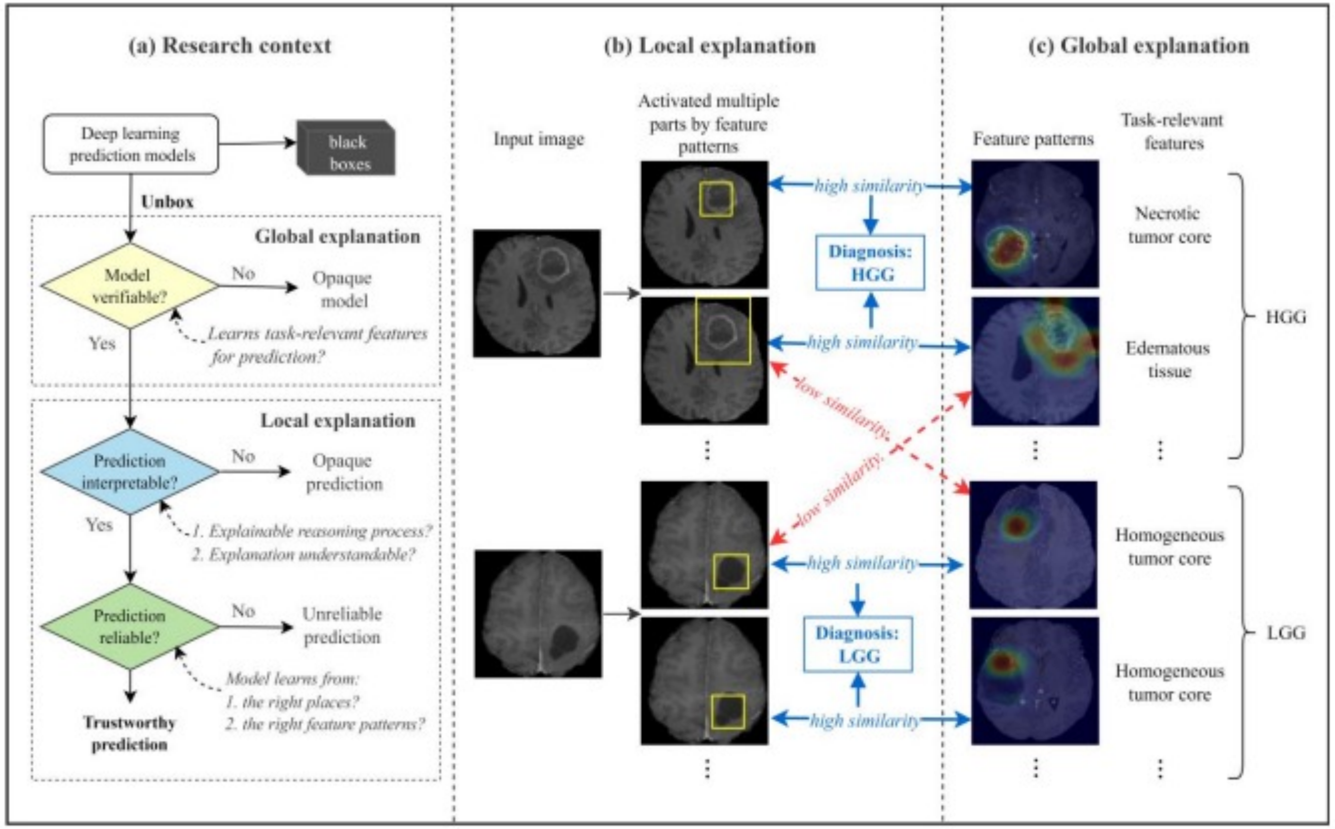
MDPI

Article

IMPA-Net: Interpretable Multi-Part Attention Network for Trustworthy Brain Tumor Classification from MRI

Yuting Xie [1,2], Fulvio Zaccagna [3,4], Leonardo Rundo [5], Claudia Testa [6,7], Ruifeng Zhu [8], Caterina Tonon [1,2], Raffaele Lodi [1,2] and David Neil Manners [2,9,*]

Diagnostics **2024**, *14*, 997. https://doi.org/10.3390/diagnostics14100997



Multi-part attention network (IMPA-Net) **for brain tumor classification**.

The proposed model **predicts the tumor grade** and provides a global explanation for the model interpretability and a local explanation as justification for the proffered prediction. Global explanation is represented as a group of **feature patterns** that the model learns to distinguish high-grade glioma (HGG) and low-grade glioma (LGG) classes. Local explanation interprets the reasoning process of an individual prediction by calculating the similarity between the prototypical parts of the image and a group of pre-learned task-related features.

Methods: BraTS2017 dataset demonstrate that IMPA-Net is a verifiable model for the classification task. A percentage of 86% of feature patterns were assessed by two radiologists to be valid for representing task-relevant medical features.

Results: the model shows a classification **accuracy** of 92.12%, of which 81.17% were evaluated as trustworthy based on local explanations. Our interpretable model is a trustworthy model that can be used for decision aids for glioma classification.

Bologna (main) activities:

ML & AI methods for discrimination tasks on biomedical images (carcinoma, COVID-19)

Other papers recently published - Bologna
Biondi R, Renzulli M, Golfieri R, **Curti** N, Carlini G, Sala C, **Remondini** D et al. Machine Learning Pipeline for the Automated Prediction of Microvascular Invasion in Hepatocellular Carcinomas. Appl Sci 2023;13. https://doi.org/10.3390/app13031371.
Carlini G, Gaudiano C, Golfieri R, **Curti** N, **Biondi** R, Bianchi L, **Remondini** D, et al. Effectiveness of Radiomic ZOT Features in the Automated Discrimination of Oncocytoma from Clear Cell Renal Cancer. J Pers Med 2023;13. https://doi.org/10.3390/jpm13030478.
Verzellesi L, Botti A, Bertolini M, Trojani V, Carlini G, Nitrosi A, et al. Machine and Deep Learning Algorithms for COVID-19 Mortality Prediction Using Clinical and Radiomic Features. Electron 2023;12:1–14. https://doi.org/10.3390/electronics12183878.
ongoing: generation of synthetic data in brain 3D PET context