

IJCAI-ECAI 2026 TUTORIAL 7

Data-Centric AI: Addressing Missing Data Imputation in Image and Tabular Contexts

Joana Cristo Santos¹, Ricardo Cardoso Pereira², Pedro Henriques Abreu²

¹ University of Lisbon, Portugal (LASIGE) ² University of Coimbra, Portugal (CISUC)

jcsantos@ciencias.ulisboa.pt, rdpereira@dei.uc.pt, pha@dei.uc.pt

Funded by the PRR Project NEURASPACE - AI Fights Space Debris, 02/C05-i01.01/2022.PC644877546-00000020.

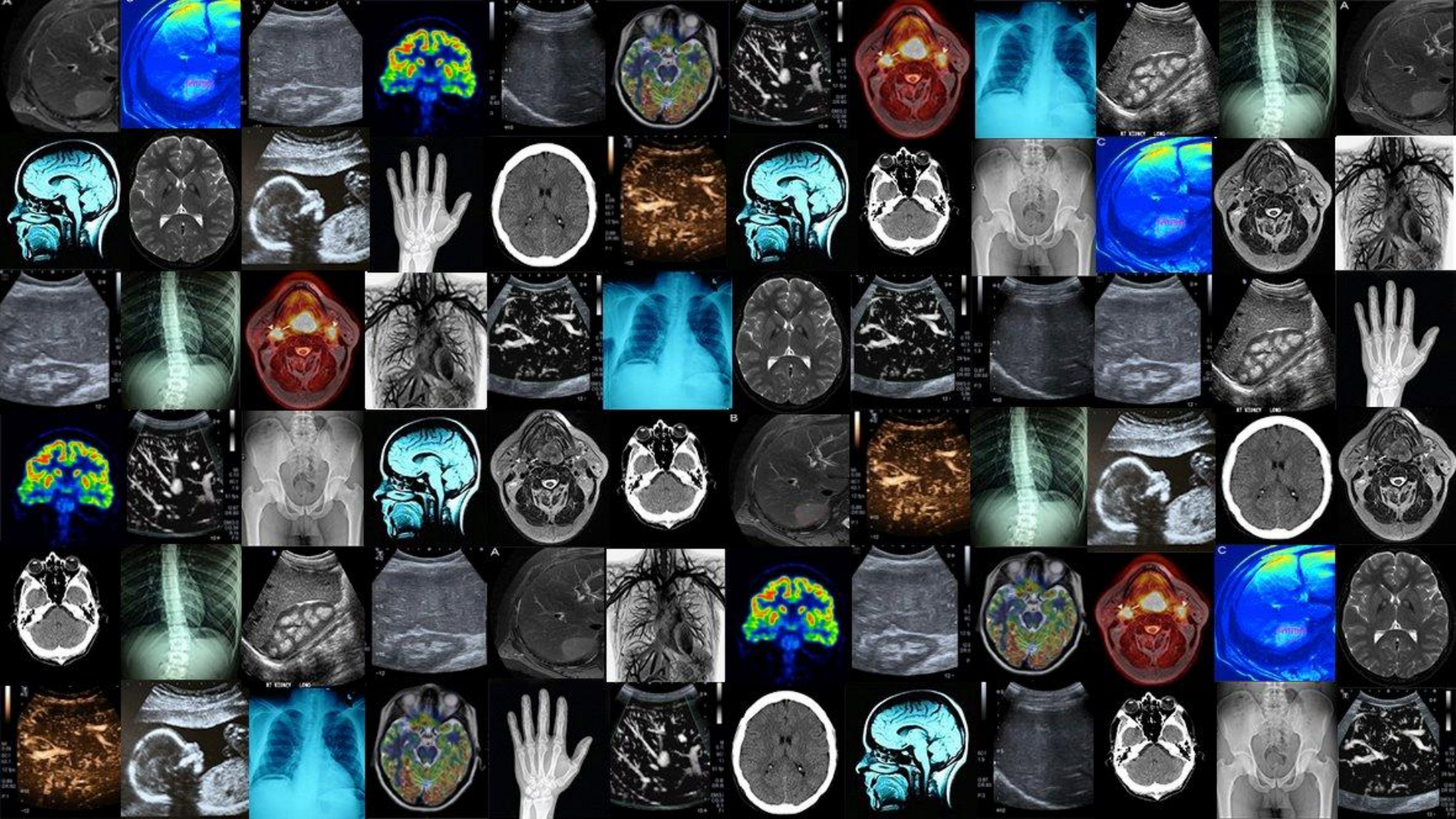
 **CISUC**

LASIGE driven by
excellence



The world's most valuable resource is no longer oil, but data.

— The Economist, Cover story (May 6, 2017)



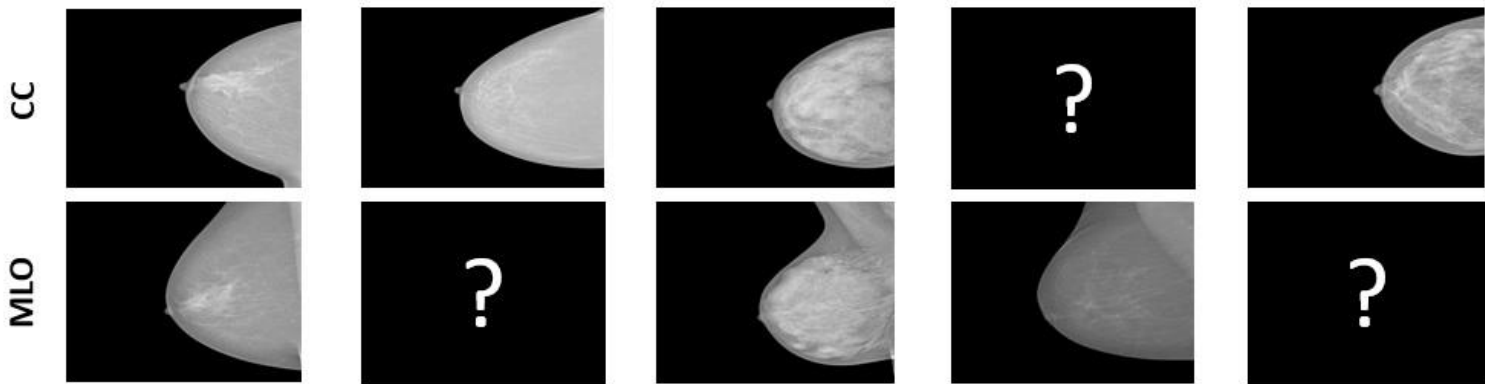
Missing Data in Imaging Datasets

Full Missing Images vs Missing Pixels

What Is Missing Data in Imaging Dataset?

Full Missing Images vs Missing Pixels

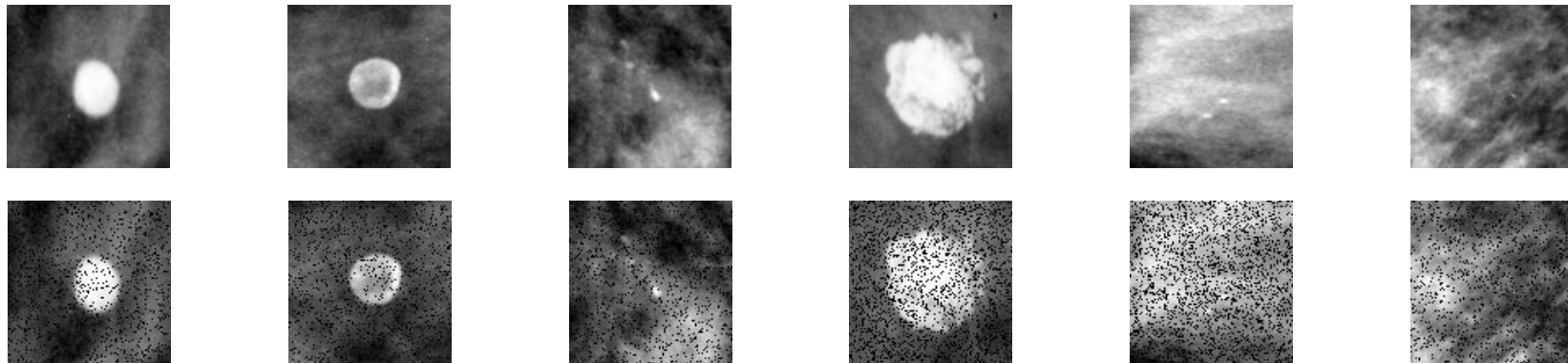
Full missing images is an **item-level missingness** = the whole image is absent.



What Is Missing Data in Imaging Dataset?

Full Missing Images vs Missing Pixels

Missing Pixels is a **value-level missingness** = pixels/regions within an image are absent or corrupted is an Item-level missingness.



Examples of Possible Situations

- **Protocol variability:** Different hospitals use different imaging protocols.
- **Emergency imaging:** Rapid acquisitions prioritize speed over comprehensive imaging.
- **Scanner or acquisition failures:** Equipment malfunctions or incomplete image acquisition.
- **Motion and metal artifacts:** Patient movement or implants degrade image quality.
- **Occlusions:** External devices, casts, or surgical instruments obscure anatomical structures.
- **Detector or sensor failures:** Faulty detector elements or coil failures result in missing image information.
- **Data anonymization:** Removal or masking of image regions to protect patient privacy.

FEATURED WORK 1

The role of deep learning in medical image inpainting: A systematic review

Joana Cristo Santos, Hugo Tomás Pereira, Miriam Seoane Santos, Pedro Henriques Abreu

ACM Transactions on Computing for Healthcare

2025

FEATURED WORK 2

Missing image data imputation using variational autoencoders with weighted loss

Ricardo Cardoso Pereira, Joana Cristo Santos, José Pereira Amorim, Pedro Pereira Rodrigues, Pedro Henriques Abreu

European Symposium on Artificial Neural Networks (ESANN), 475–480

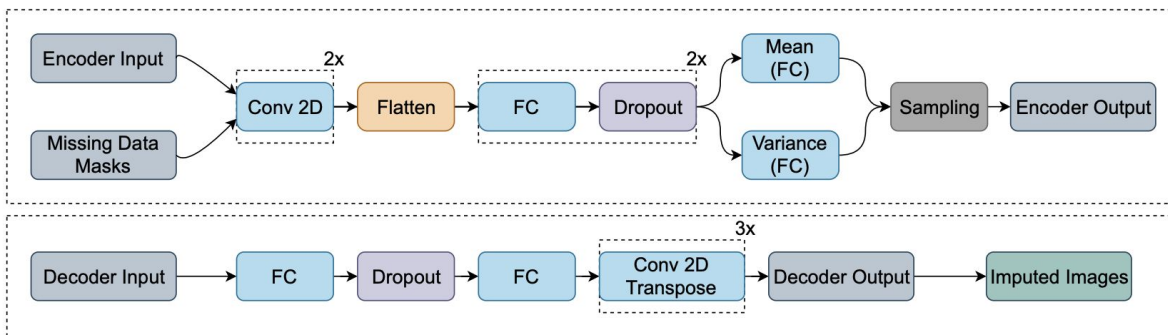
2020

Missing image data imputation using variational autoencoders with weighted loss

Ricardo Cardoso Pereira, Joana Cristo Santos, José Pereira Amorim, Pedro Pereira Rodrigues, Pedro Henriques Abreu

European Symposium on Artificial Neural Networks (ESANN), 475–480

2020



FEATURED WORK 3

The identification of cancer lesions in mammography images with missing pixels: analysis of morphology

Joana Cristo Santos, Miriam Seoane Santos, Pedro Henriques Abreu

IEEE International Conference on Data Science and Advanced Analytics (DSAA)

2022

FEATURED WORK 4

Diffusion Models in Medical Image Inpainting: Challenges, Solution Taxonomy, and Future Directions

Arthur Dantas Mangussi, Joana Cristo Santos, Ricardo Cardoso Pereira, Ana Carolina Lorena, Mário A. T. Figueiredo, Pedro Henriques Abreu

<https://arxiv.org/abs/2607.21904>

2026

Inpainting Pipeline

1. Identify or Generate Missing Regions

Detect or simulate missing pixels, regions, or artifacts to be inpainted.

2. Train Imputation/Inpainting Models

Fit classical statistical methods and deep generative models to learn how to reconstruct missing content from data.

3. Apply the Inpainting Methods

Restore missing image content using the trained models.

4. Evaluate Performance

Assess reconstruction quality using image quality metrics and evaluate the impact on downstream clinical tasks.

Thank You

Joana Cristo Santos

jcsantos@ciencias.ulisboa.pt

tinyurl.com/ijcaiecai7