

Received 14 October 2025, accepted 13 November 2025, date of publication 17 November 2025,
date of current version 21 November 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3633780



SynthVal: A Framework for Validating Synthetic Medical Images

DARIO GUIDOTTI¹, (Member, IEEE), **LAURA PANDOLFO¹**, (Member, IEEE),
ALBERTO GUTIERREZ-TORRE², **OMAR LOPEZ-RUBIO²**, AND **LUCA PULINA¹**, (Member, IEEE)

¹DUMAS, University of Sassari, 07100 Sassari, Italy

²Barcelona Supercomputing Center, Les Corts, 08034 Barcelona, Spain

Corresponding author: Dario Guidotti (dguidotti@uniss.it)

The work of Dario Guidotti, Laura Pandolfo, and Luca Pulina was supported by EU Project SECURED, “Scaling Up Secure Processing, Anonymization and Generation of Health Data for EU Cross-Border Collaborative Research and Innovation,” under Agreement 101095717. The work of Alberto Gutierrez-Torre and Omar Lopez-Rubio was supported by SECURED under Agreement 101095717; in part by Spanish Ministry of Science (MICINN), the State Research Agency (AEI), and European Regional Development Funds (ERDF/FEDER) under Grant PID2021-126248OB-I00; in part by MCIN/AEI/10.13039/501100011033/FEDER, UE; and in part by the Generalitat de Catalunya (AGAUR) under Grant 2021-SGR-00478.

• **ABSTRACT** Synthetic data is increasingly used in medical imaging to overcome data scarcity and privacy constraints. However, assessing the fidelity of synthetic images remains a critical challenge for ensuring their safe and effective use in clinical and AI applications. We present *SynthVal*, a Python-based framework for validating the quality of synthetic medical images through statistical comparisons in deep feature space. *SynthVal* extracts semantic image embeddings using transformer-based models and computes similarity metrics – including Fréchet Distance, Wasserstein Distance, and Kullback-Leibler Divergence – between real and synthetic data distributions. The framework is designed for modularity, scalability, and ease of integration into existing workflows via pip installation. We evaluate *SynthVal* using real images from the CSAW-CC dataset, and synthetic images produced by a model developed by the Barcelona Supercomputing Center. Our experiments systematically benchmark the influence of different feature extraction models and similarity metrics, providing practical insights for selecting validation strategies in medical image synthesis. *SynthVal* offers a reproducible and extensible solution for quality control in synthetic data pipelines.