Explainable Convolutional Neural Networks For Shm Based On Transmissibility Functions

Marc PARZIALE, Marco GIGLIO, Francesco CADINI

Damage diagnosis in the structural sector (mechanical, civil, aeronautical, etc.) has become a captivating debated and researched topic. Indeed, recent advances in sensor technologies have provided a massive amount of data that often embodies relevant diagnostic information exploitable for Structural Health Monitoring (SHM) purposes. Moreover, in order to accurately process the gathered signals, new deep learning-based algorithms have been developed, providing important predictions that would normally require many pre- and postprocessing steps, without even reaching the same accuracy. However, when the processing of big and complex data is required, high performance predictions are often achieved by increasing the algorithm complexity, especially when input data are perturbed by external factors (e.g. measurement noise, environmental conditions, etc.). As consequence, the considered models may turn into "black box" approaches, generating doubts regarding the way they work and come to decisions. In fact, this uncertainty has made deep learning algorithms to be hardly adopted in critical domains, though their role may be important, such as for the SHM. With this aim, this paper proposes to exploit methods developed in the Explainable Artificial Intelligence (XAI) field to explain and interpret the behavior of a convolutional neural network (CNN) that extracts subtle damage-related features from complex transmissibility functions (TFs) spectra. The interpretability of the vibration-based structural diagnosis algorithm is studied with reference to a simple, but realistic, numerical case study of a structural beam excited by an external load.