A comparison of data synthesis algorithms on small acoustic data sets for training unsupervised learning methods

Christian Scharr[1], Christoph Heinze[2], Konrad Jagusch[3], Laura Knitter[3], Michael Geist[2], Wilko Flügge[4]

[1]Measuring of Large Structures, IGP Fraunhofer, Germany, [1]Measuring of Large Structures , Fraunhofer IGP, Germany, [1]Production planning and control, Fraunhofer IGP, Germany, [1]Production engineering, University of Rostock, Germany

Engine failure has been the leading cause of marine accidents over the past decade. In addition to costly repairs and downtime, engine failure can lead to serious accidents. To ensure safe ship operation, predictive maintenance systems are used to anticipate failures, which requires a large and balanced set of machine data to train and test the learning algorithms. However, collecting sensor data while the ship is in operation is time-consuming and requires deep domain knowledge. In this paper, we compare and evaluate different approaches for generating synthetic data based on small sets of acoustic sensor data used for fault detection training. The audio synthesis methods examined include WaveGAN, WaveNet, and SpecGAN. These machine learning algorithms are used to enrich sensor data sets for a fault detection system for marine engines. The algorithms were evaluated using acoustic sensor data recorded from a small electric motor under various conditions in a laboratory environment. The evaluation criteria of the methods are efficiency, data representation, computational and data complexity. In addition, the quality of the synthetic sound files was validated by human subjects. Overall, the data synthesis algorithms have produced good results, but the applicability for error detection training needs to be demonstrated in further investigations.