SUPERVISED LEARNING FOR SELECTIVE MULTI-SPECIES QUANTIFICATION FROM NOISY INFRARED SPECTROSCOPY DATA

EMAD AL IBRAHIM, AAMIR FAROOQ, *Clean Combustion Research Center, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia*.

A supervised learning approach is implemented to extract information from noisy vibrational spectroscopy data. Our method tackles two of the main problems in any commercial sensing application: sensitivity and selectivity. First, an encoder takes in noisy spectra of complex mixtures and learns reduced representations referred to as embeddings. The learned embeddings are then used in the decoder to filter out noise and unwanted species. Embeddings are also simultaneously used as input to a regression network for the prediction of concentrations and baseline shift. The model was applied for gas sensing using Fourier-Transform Infrared spectroscopy (FTIR) data. We focus on identifying common volatile organic compounds (VOCs) in a realistic scenario. The multitask nature of the model gives better results compared to single task denoising followed by regression and classical techniques like non-negative linear regression. The denoising capability was also compared to other denoising methods like Savitzky-Golay filters (SVG) and wavelet transformations (WT).