We use a few feature extraction strategies listed below: We consider a 10-level decomposition of the learned wavelet (DeSpaWN) for all methods. Each level has a different number of coefficients, and they represent the same time window (with different sampling rates).

Grids2:. Implemented in notebook "Extract_features_DeSpaWN_Grids2", by function "extract_DeSpaWN_grids". Each level is separated into five regions, called grids. Each decomposition level has five grids. Each grid g_k^l comprises the average value of the coefficients that make up a time window $\frac{T}{5}$ of the level l, where T is the total period of the input temporal sample of the Learned Wavelet. Each page of the feature tensor $F \in \mathbb{R}^{n_e \times n_{grids} \times L}$ is given by an $F_n \in \mathbb{R}^{n_{grids} \times L}$ matrix, where n_{grids} is the number of grids and L is the number of decomposition levels.

For grids2, $F_n = \begin{bmatrix} g_{1,1} & \cdots & g_{1,L} \\ \vdots & \ddots & \vdots \\ g_{n_{grids},1} & \cdots & g_{n_{grids},L} \end{bmatrix}$, where $g_{g,l} = \frac{1}{n_c} \sum h_l^g[k]_{k=1}^{n_c}$

Grids3:. Implemented in notebook "Extract_features_DeSpaWN_Grids3", by function "extract_DeSpaWN_grids3". Each level is separated into five regions, called grids. Each decomposition level has five grids. Each grid g_k^l comprises the average value of the coefficients that make up a time window $\frac{T}{5}$ of the level l, where T is the total period of the input temporal sample of the Learned Wavelet. Each page of the feature tensor $F \in \mathbb{R}^{n_e \times n_{grids} \times L}$ is given by an $F_n \in \mathbb{R}^{n_{grids} \times L}$ matrix, where n_{grids} is the number of grids and L is the number of decomposition levels.

For grids3, $F_n = \begin{bmatrix} g_{2,1} - g_{1,1} & g_{2,2} - g_{1,2} & \cdots & g_{2,L} - g_{1,L} \\ g_{3,1} - g_{2,1} & g_{3,2} - g_{2,2} & \cdots & g_{3,L} - g_{2,L} \\ g_{4,1} - g_{3,1} & g_{4,2} - g_{3,2} & \cdots & g_{4,L} - g_{3,L} \\ g_{5,1} - g_{4,1} & g_{5,2} - g_{4,2} & \cdots & g_{5,L} - g_{4,L} \\ g_{5,1} & \cdots & g_{n_{grids},L-1} & g_{n_{grids},L} \end{bmatrix}$, where $g_{g,l} = \frac{1}{n_c} \sum h_l^g[k]_{k=1}^{n_c}$

Grids4:. Implemented in notebook "Extract_features_DeSpaWN_Grids4", by function "extract_DeSpaWN_grids4". Each level is separated into five regions, called grids. Each decomposition level has five grids. Each grid g_k^l comprises the average value of the coefficients that make up a time window $\frac{T}{5}$ of the level l, where T is the total period of the

input temporal sample of the Learned Wavelet. Each page of the feature tensor $F \in \mathbb{R}^{n_e \times n_{grids} \times (L+1)}$ is given by an $F_n \in \mathbb{R}^{n_{grids} \times (L+1)}$ matrix, where n_{grids} is the number of grids and L is the number of decomposition levels.

$$F_n = \begin{bmatrix} b_1 & g_{2,1} - g_{1,1} & g_{2,2} - g_{1,2} & \cdots & g_{2,L} - g_{1,L} \\ b_2 & g_{3,1} - g_{2,1} & g_{3,2} - g_{2,2} & \cdots & g_{3,L} - g_{2,L} \\ b_3 & g_{4,1} - g_{3,1} & g_{4,2} - g_{3,2} & \cdots & g_{4,L} - g_{3,L} \\ b_4 & g_{5,1} - g_{4,1} & g_{5,2} - g_{4,2} & \cdots & g_{5,L} - g_{4,L} \\ b_5 & g_{5,1} & \cdots & g_{n_{grids},L-1} & g_{n_{grids},L} \end{bmatrix}$$

For grids4, , where , $$a_0$$ are the approximation coefficients of the wavelet transform, $$n_e$$ is the number of elements in $$a_0$$, $$h_l^g$$ is the set of detailing coefficients of the level $$l $$ and in the grid $$g$$. $$g_{g,l} = \frac{1}{n_c} \sum h_l^g[k]_{k=1}^{n_c}$$ are the approximation coefficients of the wavelet transform, n_e is the number of elements in a_0, h_l^g is the set of detailing coefficients of the level l and in the grid g.