9. Supplementary Material

Detailed Information for the Experimental Setup

We define the *Similarity Trajectory* as a discrete time series $\{z_t\}_{t=T-1}^{1}$, where T is the total number of time steps in the sampling process. Each element z_t represents the similarity score between the denoised images in consecutive time steps t and $t-1$. To analyze fluctuations within this trajectory, we segment it either based on time steps or by projecting it onto different bases using the Haar Transform. We extract various statistical quantities to characterize each set which subsequently serve as inputs to a RF Classifier. Segmenting the Similarity Trajectory in the time domain allows us to capture variations at specific time steps, which convey varying information; notably, fluctuations during the middle section of sampling are critical for artifact detection. Additionally, the coefficients from different levels of the Haar Transform reveal fluctuations at various frequencies, enabling us to characterize the duration and magnitude of these changes.

Segmentation of the Time Series

We divide the entire *Similarity Trajectory* into three equal sets S_1, S_2, S_3 based on time steps, as well as considering the entire series as a single set, denoted as S_4. Segments in the time domain are formally defined as:

- **Segment 1** (S_1):
$$S_1 = \{z_t \mid t = 1, 2, \ldots, N_1\},$$
where
$$N_1 = \left\lfloor \frac{T-1}{3} \right\rfloor.$$
- **Segment 2** (S_2):
$$S_2 = \{z_t \mid t = N_1 + 1, N_1 + 2, \ldots, N_2\},$$
where
$$N_2 = \left\lfloor \frac{2(T-1)}{3} \right\rfloor.$$
- **Segment 3** (S_3):
$$S_3 = \{z_t \mid t = N_2 + 1, N_2 + 2, \ldots, T-1\}.$$
- **Segment 4** (S_4):
$$S_4 = \{z_t \mid t = 1, 2, \ldots, T-1\}.$$

Segmentation of Coefficients for Haar Transform

We apply the discrete Haar wavelet transform to the entire *Similarity Trajectories* $\{z_t\}_{t=T-1}^{1}$ because of its ability to detect sudden changes in time-series data [23]. In this section, we first introduce the Haar Transform, and then we explain how to process the *Similarity Trajectory* using the Haar Transform.

Haar Transformation

The Haar Transform decomposes the original time series into approximation and detail coefficients at various scales, capturing both global trends and local variations.

At the first level of decomposition, for $k = 1, 2, \ldots, \left\lfloor \frac{T-1}{2} \right\rfloor$, the approximation coefficients $a_1(k)$ and detail coefficients $d_1(k)$ are calculated as:

$$a_1(k) = \frac{z_{2k-1} + z_{2k}}{2}, \tag{10}$$

$$d_1(k) = \frac{z_{2k-1} - z_{2k}}{2}. \tag{11}$$

This process is recursively applied to the approximation coefficients to obtain higher-level coefficients. At level $j+1$, the coefficients are computed as:

$$a_{j+1}(k) = \frac{a_j(2k-1) + a_j(2k)}{2}, \tag{12}$$

$$d_{j+1}(k) = \frac{a_j(2k-1) - a_j(2k)}{2}, \tag{13}$$

where $j = 1, 2, \ldots, J$, and J is the maximum level of decomposition. From the transformation, we obtain a set of detail coefficients $\{d_j(k)\}$ corresponding to each basis function at various levels. Note that the detail coefficients capture the fluctuation information of the *Similarity Trajectory*, which is important for assessing the presence of artifacts in the image.

We segment the detail coefficients obtained from the Haar Transform of the entire *Similarity Trajectory* $\{z_t\}_{t=T-1}^{1}$ by grouping them according to their corresponding Haar basis functions. Each set S_j consists of all detail coefficients at decomposition level j:

$$S_j = \{d_j(k) \mid k = 1, 2, \ldots, N_j\},$$

where $N_j = \left\lceil \frac{T-1}{2^j} \right\rceil$ is the number of coefficients at level j. This segmentation aligns each set of detail coefficients with their respective scales in the time series, allowing us to analyze fluctuations captured by each Haar basis function effectively.

Feature Extraction

For all the sets S obtained—whether from the detail coefficients of the Haar Transform or in the time domain—we calculate ten statistical features for each set to perform the bag-of-statistics method. These statistical features are chosen to describe the dynamics of the *Similarity Trajectory* as we already established that the fluctuation in *Similarity Trajectory* is correlated to the presence of artifacts. They include:

1. **Mean (μ_S)**: The average value of the data in the set S.

$$\mu_S = \frac{1}{N_S} \sum_{s \in S} s, \quad (14)$$

where N_S is the number of elements in the set S.

2. **Standard Deviation (σ_S)**: Measures the dispersion of the data in the set S. This is related to fluctuation in the *Similarity Trajectory* for sets in the time domain.

$$\sigma_S = \sqrt{\frac{1}{N_S} \sum_{s \in S} (s - \mu_S)^2}. \quad (15)$$

3. **Percentile**: We extract the 5th, 25th, 50th, 75th, and 95th percentiles for each obtained set's values. The significance of percentiles lies in their relation to the fluctuation of the *Similarity Trajectory* for detail coefficients.

4. **Number of Mean Crossings ($C_{\mu,S}$)**: Counts how many times the data crosses its mean value in the set S. This describes how rapidly the *Similarity Trajectory* fluctuates in the time domain.

$$C_{\mu,S} = \sum_{i=1}^{N_S} \mathbb{I}\left[(S_{i+1} - \mu_s)(S_i - \mu_s) < 0\right], \quad (16)$$

where N_S is the total number of elements in set S. S_i is the i^{th} element in S and $\mathbb{I}[\cdot]$ is the indicator function.

5. **Number of Zero Crossings ($C_{0,S}$)**: Counts how many times the data crosses zero in set S. Note that in detail coefficients, this represents how many times the *Similarity Trajectory* changes direction, from monotonically increasing to monotonically decreasing or vice versa.

$$C_{0,S} = \sum_{i=1}^{N_S} \mathbb{I}\left[S_{i+1} S_i < 0\right]. \quad (17)$$

where N_S is the total number of elements in set S. S_i is the i^{th} element in S and $\mathbb{I}[\cdot]$ is the indicator function.

6. **Entropy (E_s)**: Measures how uniform of the data in set S is. Again, this is another metric characterizing the fluctuations of the set.

$$E_S = -\sum_i p_i^{(S)} \log_2 p_i^{(S)}, \quad (18)$$

where $p_i^{(S)}$ is the probability of the i-th bin in the histogram of the data in set S.

These features are computed for both the time-domain data and the detail coefficients from Haar Transform, resulting in a comprehensive feature set that captures both temporal and frequency-domain characteristics.

Using k-Nearest Neighbor Model Probabilities

We employed a k-Nearest Neighbor (k-NN) model trained directly on the time-domain *Similarity Trajectory*. This k-NN model estimates the probability of artifact presence by assessing the proportion of its nearest neighbors that are labeled as artifact or non-artifact images. These predicted probabilities are then incorporated as additional features into the RF Classifier.

Feature Vector Construction

For every set S, we formulate a feature vector $\mathbf{f}_S$ which includes ten statistical features. The comprehensive feature vector $\mathbf{F}$ for the trajectory is then assembled by concatenating all $\mathbf{f}_S$ vectors from every set along with the prediction probability from the k-NN model. The feature vector $\mathbf{F}$ serves as the input to the RF Classifier for both training and inference.