Homework 3: Dimensionality Reduction & Clustering

Goal: Perform dimensionality reduction and clustering on boiling images.

Data: The dataset consists of images from a boiling experiment within my lab. These can be found at this link: https://data.mendeley.com/datasets/5kjnphrbsz/1.

Assignment:

1. Run principal component analysis (PCA) and one of the following methods; single value decomposition (SVD), t-sne, u-map; on the boiling images to reduce the dimensionality of them.
2. Plot the percentage explained variance vs number of principal components (PC).
3. Pick a representative image, run PCA and plot the reconstructed images using a different number of PCs (e.g. using PC1, PCs 1-2, PCs 1-10, PCs 1-20, etc.)
4. Calculate the error of the reconstructed images relative to the original image and plot the error as a function of the number of PCs.
5. Run a clustering analysis of the boiling images using the PCs (the number of PCs to use is up to your choice) and evaluate the results of the clustering.