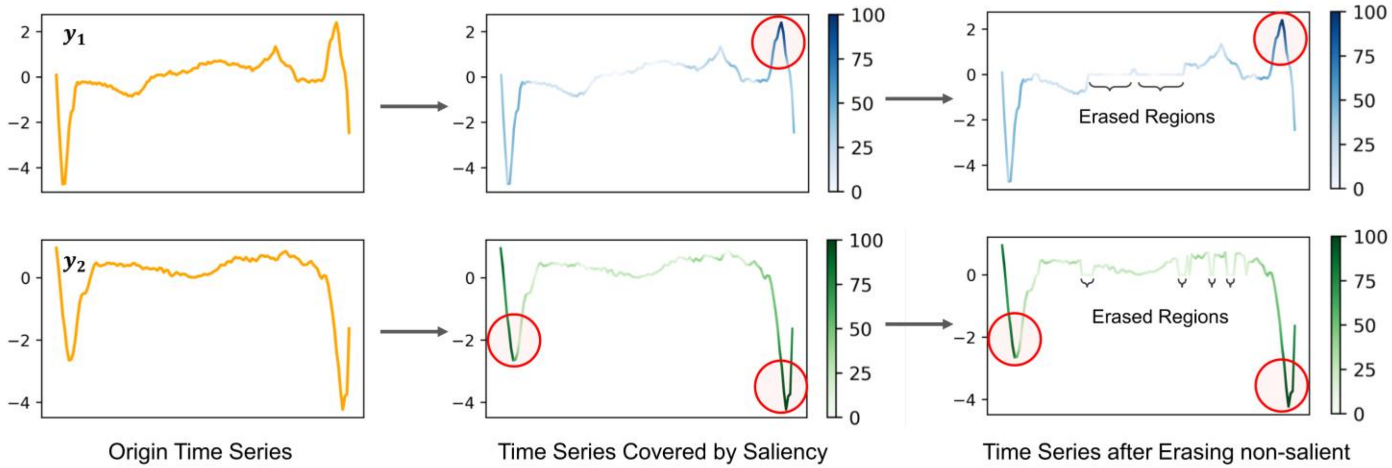
Non-salient region erasure for time series augmentation

Pin LIU, Xiaohui GUO, Bin SHI, Rui WANG, Tianyu WO, Xudong LIU

Frontiers of Computer Science, DOI: 10.1007/s11704-022-1765-6

Problems & Ideas

- The overfitting problem of the time series classification model:
  - The salient region is the core basis for distinguishing categories and has not received special attention.
  - Existing methods do not protect these salient regions, and lead to its inability to guide the model to improve classification performance.
- Ideas: An augmentation method that utilizes erasing non-salient regions (redundant information), which considers both the diversity and fidelity of augmented time series.



Core process based on erasing non-salient regions. Left: ECG original time series for two classes; Middle: Extracted saliency map of the original time series, where the red circles mark the salient regions; Right: Augmented time series after erasing non-salient regions, where the salient regions are always preserved.

Main Contributions

- Contributions:
  - A novel saliency-based time series augmentation method can not only increase the diversity of training samples but also improve the fidelity of augmented time series;
  - A lightweight method that does not require any additional parameter learning, reducing the burden of training time consumption for time series classification tasks;
  - An easy-to-integrate architecture that be implanted directly into 1D convolution-based neural networks to perform end-to-end training.

Table 1 Classification performance (Mean Acc.%) with different augmentation methods using FCN.

DataSet	None	wDBA	RGW-sD	SeaM
Adiac	84.78(0.38)	84.53(1.92)	75.06(0.38)	**86.13(0.83)**
ArrowHead	84.30(1.50)	88.00(2.00)	86.00(0.29)	**89.57(0.71)**
DiatomSR.	31.30(3.60)	76.14(0.00)	91.01(2.45)	**97.30(0.41)**
PowerCons	85.06(0.24)	86.39(2.50)	86.67(0.28)	**89.03(0.97)**
FordA	90.40(0.20)	90.00(1.03)	87.12(2.00)	**91.99(0.28)**
Worms	76.50(2.20)	80.58(3.50)	79.06(0.30)	**81.17(0.65)**
Lightning7	82.70(2.30)	89.65(5.00)	83.56(0.00)	**90.07(3.08)**
Earthquakes	72.70(1.70)	72.66(1.28)	74.10(3.33)	**74.10(1.44)**
ElectricD.	70.02(1.20)	71.66(1.04)	71.97(2.06)	**74.49(2.26)**
CROP	69.74(3.11)	71.51(1.52)	70.55(1.71)	**72.55(3.24)**
Average	74.75(1.64)	81.11(2.00)	80.51(1.28)	**84.64(1.39)**

Table 2 Classification performance (Mean Acc.%) with different augmentation methods using ResNet.

DataSet	None	wDBA	RGW-sD	SeaM
Adiac	82.90(0.60)	83.63(1.53)	75.70(0.77)	**84.14(1.02)**
ArrowHead	84.50(1.20)	86.00(0.86)	88.57(1.14)	**89.43(0.86)**
DiatomSR.	30.10(0.20)	80.07(0.33)	97.55(0.49)	**98.12(0.25)**
PowerCons	88.12(1.60)	86.67(1.67)	90.00(1.67)	**90.00(0.56)**
FordA	92.00(0.40)	93.45(0.05)	92.83(1.20)	**93.54(0.32)**
Worms	79.10(2.50)	78.06(6.75)	71.43(3.75)	**80.84(2.27)**
Lightning7	84.50(2.00)	86.47(2.00)	84.93(2.74)	**86.99(0.68)**
Earthquakes	71.20(2.00)	**74.82(0.25)**	73.38(0.10)	71.76(3.06)
ElectricD.	72.90(0.90)	73.56(0.62)	73.17(1.03)	**75.62(0.85)**
CROP	68.52(2.33)	69.15(2.15)	71.74(1.62)	**74.23(2.03)**
Average	75.38(1.37)	81.19(1.62)	81.93(1.45)	**84.47(1.19)**

Performance comparison of SeaM with state-of-the-art methods on typical datasets. Left: Accuracy and standard error using the FCN architecture; Right: Accuracy and standard error using the ResNet architecture.