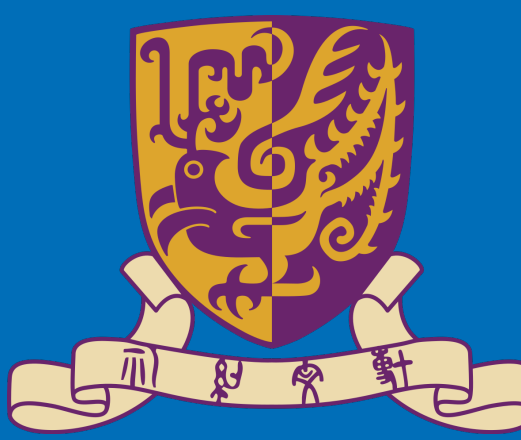
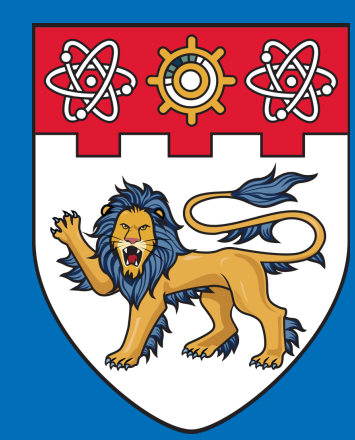




SHAPENET: A SHAPELET-NEURAL NETWORK APPROACH FOR MULTIVARIATE TIME SERIES CLASSIFICATION



Guozhong Li¹, Byron Choi¹, Jianliang Xu¹, Sourav S Bhowmick², Kwok Pan Chun³, and Grace L.H. Wong⁴

¹Computer Science, HKBU ²Computing Engineering, NTU ³Geography, HKBU ⁴Medicine, CUHK
csgzli@comp.hkbu.edu.hk

Introduction

Recent solutions for multivariate time series classification

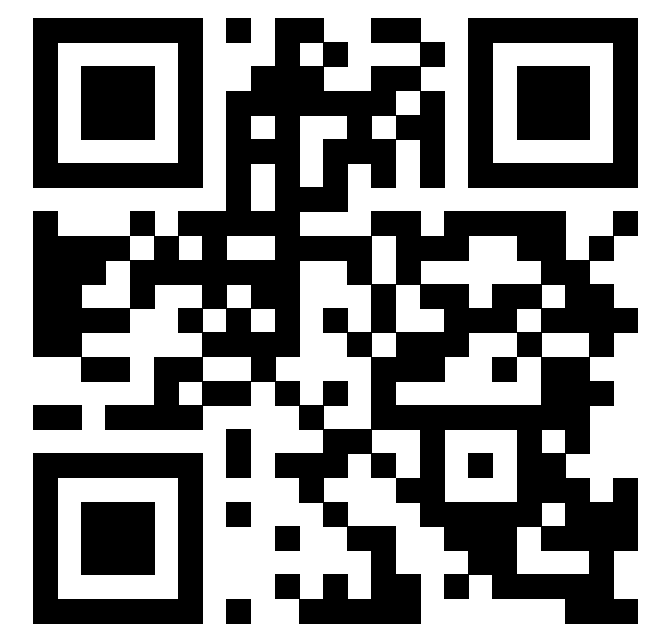
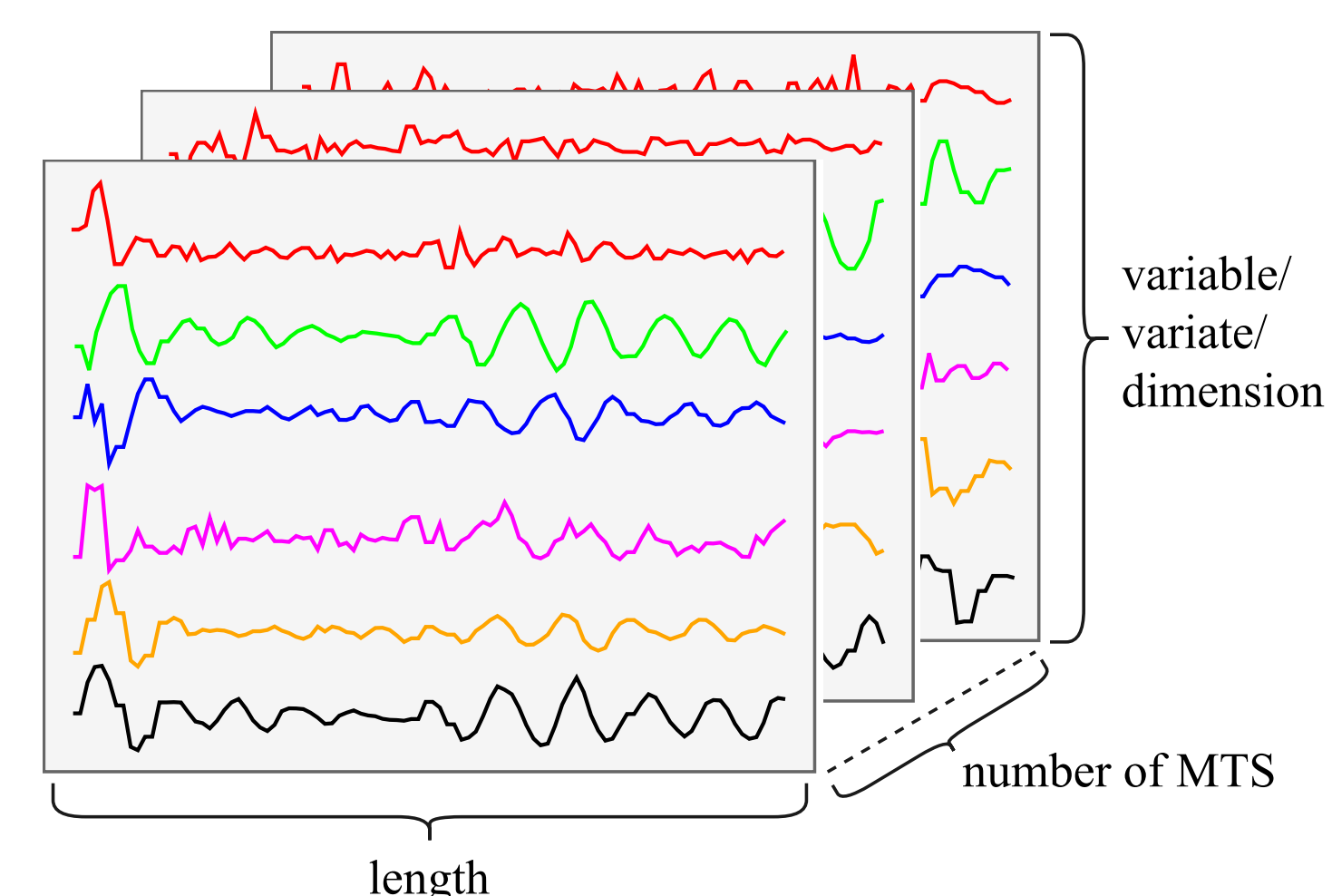
- ✓ **Model-based methods** A tree classifier based on a new symbolic representation to extract relationship;
- ✓ **Neural network-based methods:** An LSTM layer and stacked CNN layer to extract features for a softmax layer;

Problem statement

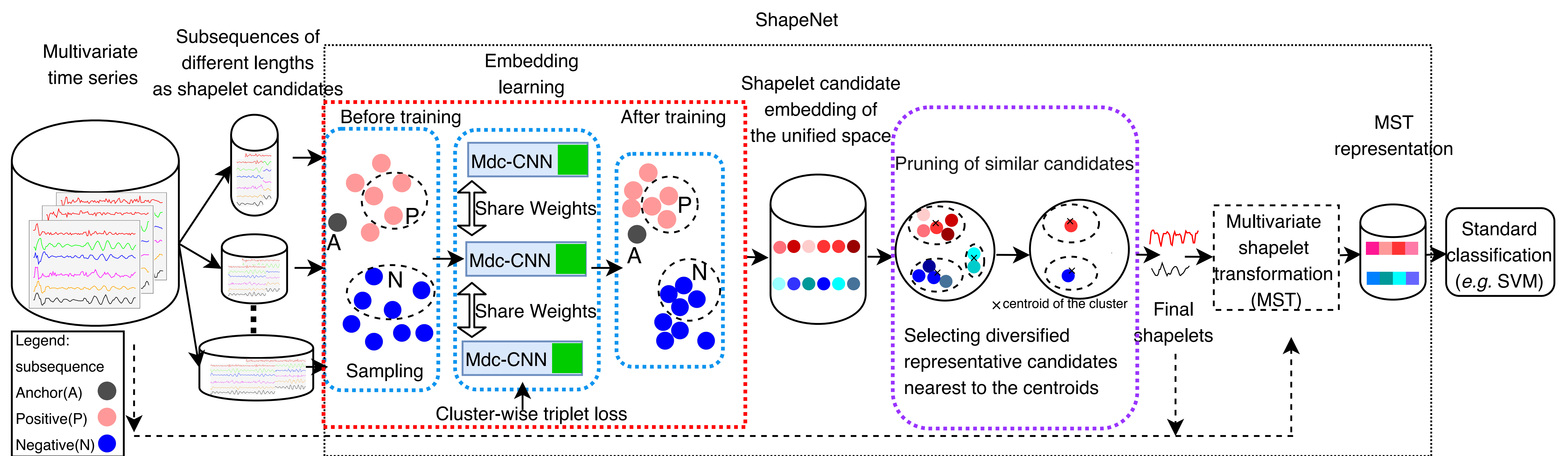
Given a multivariate time series dataset \mathbb{D} , consisting of M multivariate time series instances T_1, T_2, \dots, T_M with V variables, this paper investigates a shapelet-based classifier.

Challenges

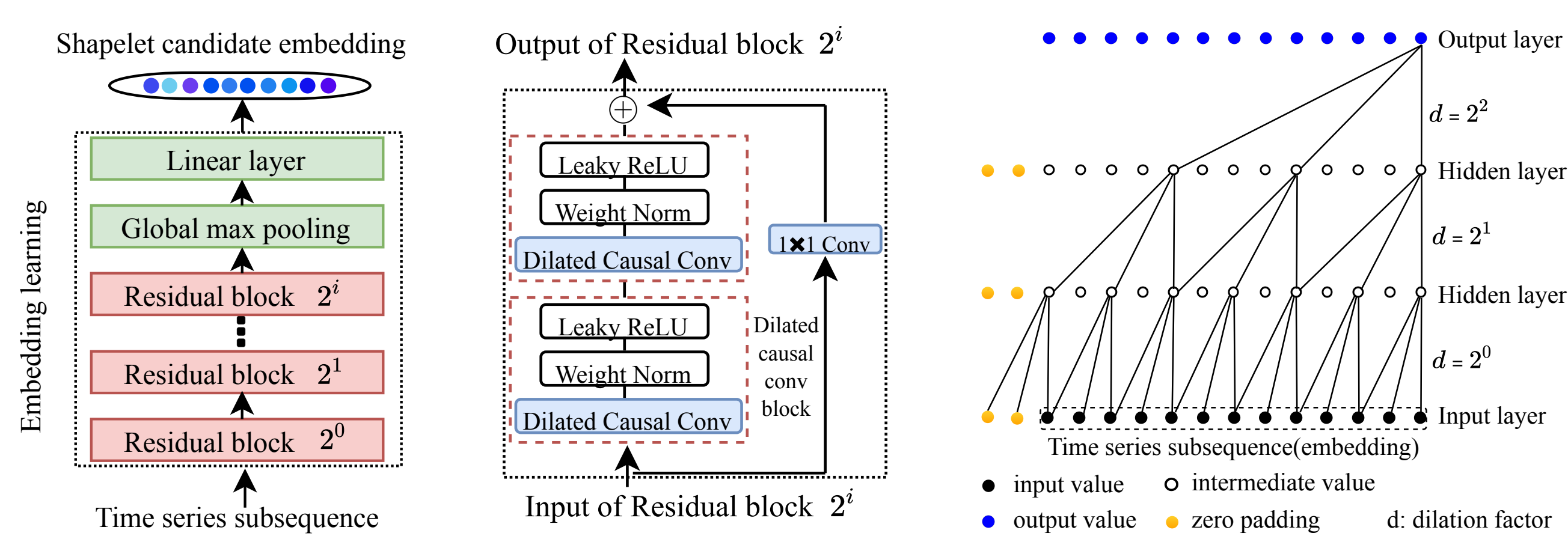
- ✓ multiple variables and different lengths, heterogeneous and voluminous candidates
- ✓ hard to compare candidates of different lengths
- ✓ original triplet loss, one positive (negative), may not converge and not stable
- ✓ triplets selection (negative samples), word2vec assumption not hold
- ✓ black-box nature of most existing studies



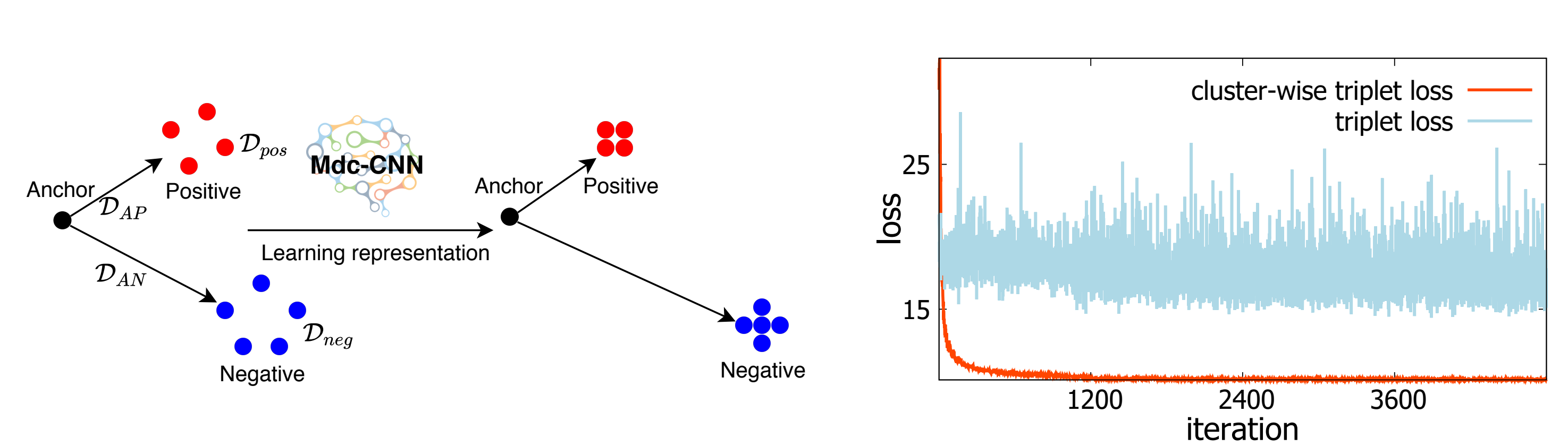
The overview of ShapeNet



Mdc-CNN details



Cluster-wise triplet loss - Learning time series representations



$$\mathcal{L}(f(x), f(x^+), f(x^-)) = \log \frac{D_{AP} + \mu}{D_{AN}} + \lambda D_{intra}$$

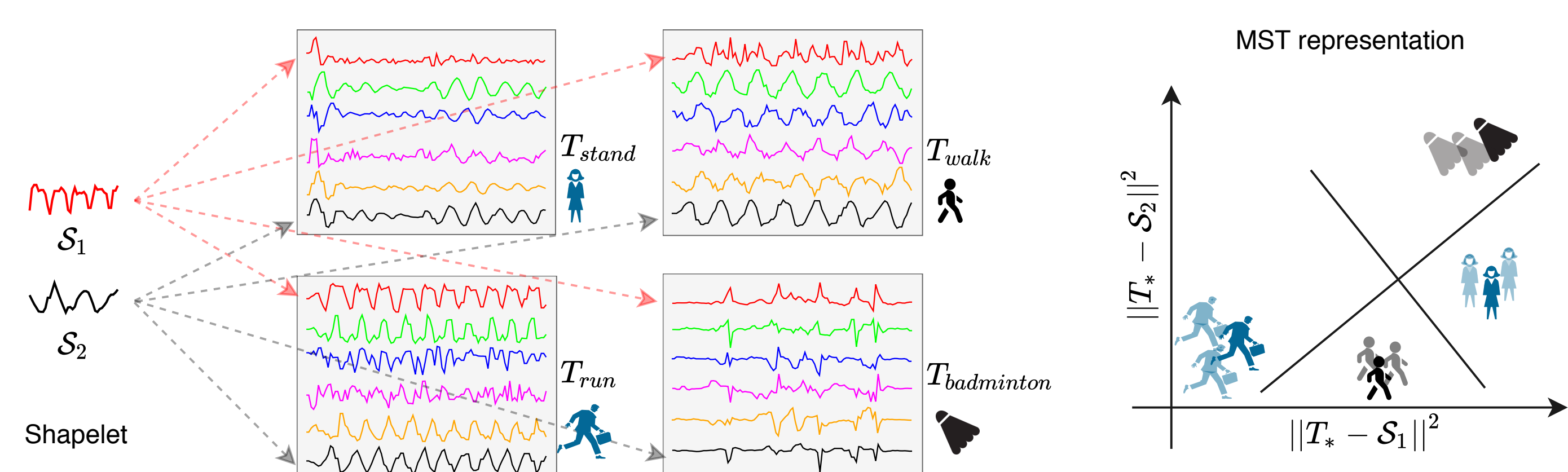
Selection of final shapelets

- A clustering method (kmeans) to yield Y clusters of the shapelet candidates
- A utility to rank the candidates that are nearest to the cluster centroids

$$\mathcal{U}(f(x_i)) = \beta \cdot \frac{\log(\text{size}(f(x_i)))}{\log(\max(\text{size}(f(x_i))))} + (1 - \beta) \frac{\log \sum_{j=1}^Y \|f(x_i) - f(x_j)\|_2^2}{\log(\max(\sum_{j=1}^Y \|f(x_i) - f(x_j)\|_2^2))}$$

- the size of the candidate's cluster
- the candidate's distance to other candidates in other classes

An interpretability result



An example of multivariate shapelet transformation on Basicmotions

Performance of overall accuracy

Accuracies on 30 UEA ARCHIVE datasets [1]

Dataset	EDI	DTWI	DTWD	MLSTM -FCNs	WEASEL +MUSE	NS	TapNet	ShapeNet
Total best acc	1	2	2	4	12	5	5	14
Ours 1-to-1-Wins	29	26	22	21	15	18	20	-
Ours 1-to-1-Draws	1	3	5	3	3	5	5	-
Ours 1-to-1-Losses	0	1	3	6	12	7	5	-
Rank Mean	6.2	5.43	4.77	4.6	3.47	3.67	3.23	2.23
Wilcoxon Test p-value	0.000	0.000	0.000	0.001	0.183	0.819	0.002	-

References

- A. Bagnall, H. A. Dau, J. Lines, M. Flynn, J. Large, A. Bostrom, P. Southam, and E. Keogh, "The uea multivariate time series classification archive, 2018," *arXiv preprint arXiv:1811.00075*, 2018.
- G. Li, B. Choi, J. Xu, S. S. Bhowmick, K. P. Chun, and G. Wong, "Shapenet: A shapelet-neural network approach for multivariate time series classification," *AAAI*, 2021.
- G. Li, B. Choi, J. Xu, S. S. Bhowmick, K.-P. Chun, and G. L. Wong, "Efficient shapelet discovery for time series classification," *TKDE*, 2020.
- G. Li, B. Choi, S. S. Bhowmick, G. L.-H. Wong, K.-P. Chun, and S. Li, "Visualet: Visualizing shapelets for time series classification," in *CIKM*, 2020, 3429–3432. <https://youtu.be/cAWyIzXDxik>.