TapNet: Multivariate Time Series Classification with Attentional Prototypical Network

Xuchao Zhang[1], [2]Yifeng Gao, [2]Jessica Lin, Chang-Tien Lu[1]

[1]Discovery Analytics Center, Virginia Tech, Falls Church, VA

[2]Department of Computer Science, George Mason University, Fairfax, VA

Full Paper

Source Code

Introduction

With the advance of sensor technologies, the Multivariate Time Series classification (MTSC) problem has continuously received a significant amount of attention in recent decades. Our paper focuses on:

- **Multivariate Time Series Classification**: Given a group of multivariate time series $\mathcal{X} = \{X_1, \ldots, X_n\} \in \mathcal{R}^{n \times m \times l}$, where n is the number of time series, and the corresponding labels $y = \{y_1, \ldots, y_n\} \in \mathcal{R}^n$ for each time series, the MTSC task is to train a classifier $f_X \mapsto y$ to predict a class label for a multivariate time series whose label is unknown.
- **Semi-supervised Multivariate Time Series Classification**: Assume model can utilize unlabeled data during the training process to help improve the overall classification performance.

Challenge

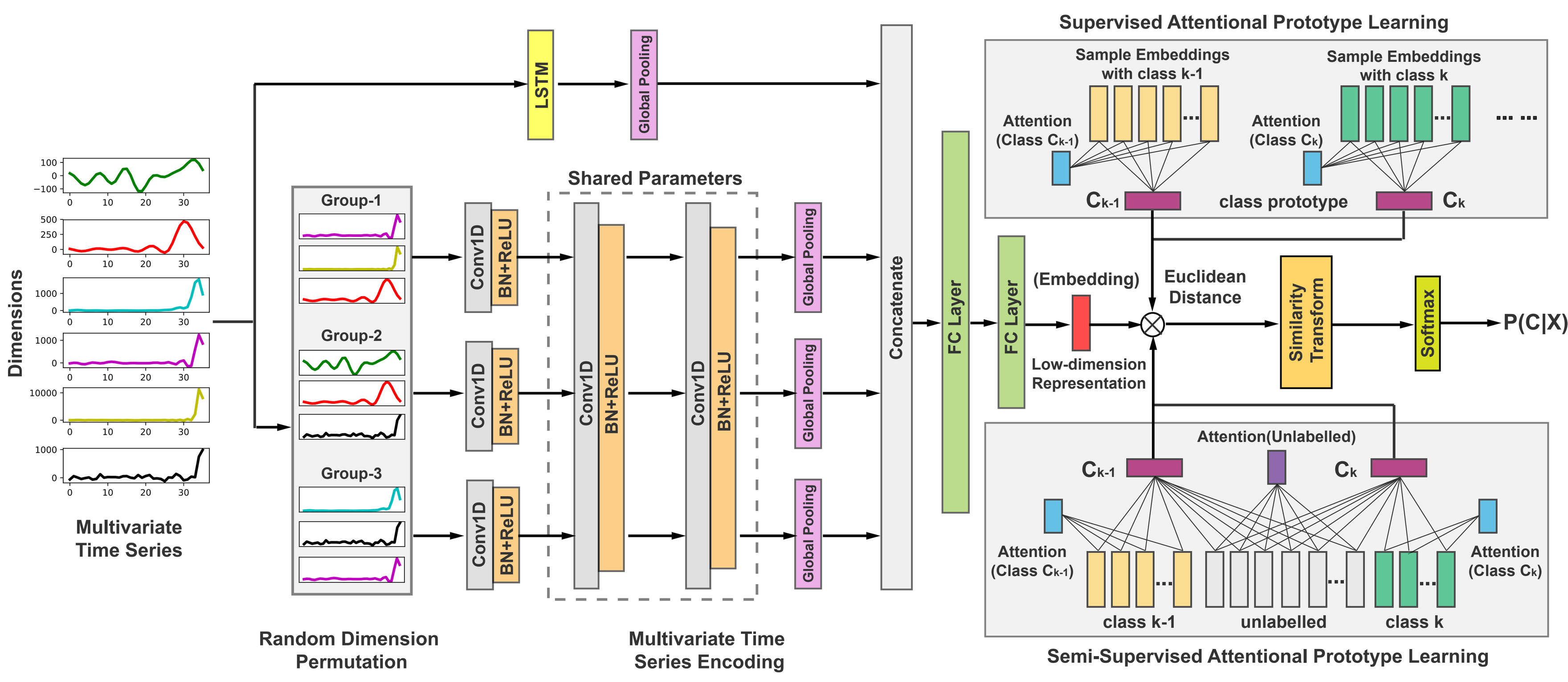
- Traditional time series classification approaches may generate huge amounts of feature candidates.
- Deep learning based methods suffer from a shortage of labelled data.

Contribution

We propose a novel attentional prototype network (TapNet):

- TapNet train the feature representation based on their distance to class prototypes with inadequate data labels.
- TapNet can be extended into its semi-supervised setting by utilizing the unlabeled data.

Attentional Prototypical Network



Let $H_k = [h_1, \ldots, h_{S_k}] \in \mathcal{R}^{S_k \times d}$ be a matrix of time series embeddings belonging to the class k, where S_k represents the set of indices for data samples with class label k.

$$c_k = \sum_i A_{k,i} \cdot H_{k,i}, \tag{1}$$

The attention weights $A_{k,i}$ for the k class can be computed by the following equation:

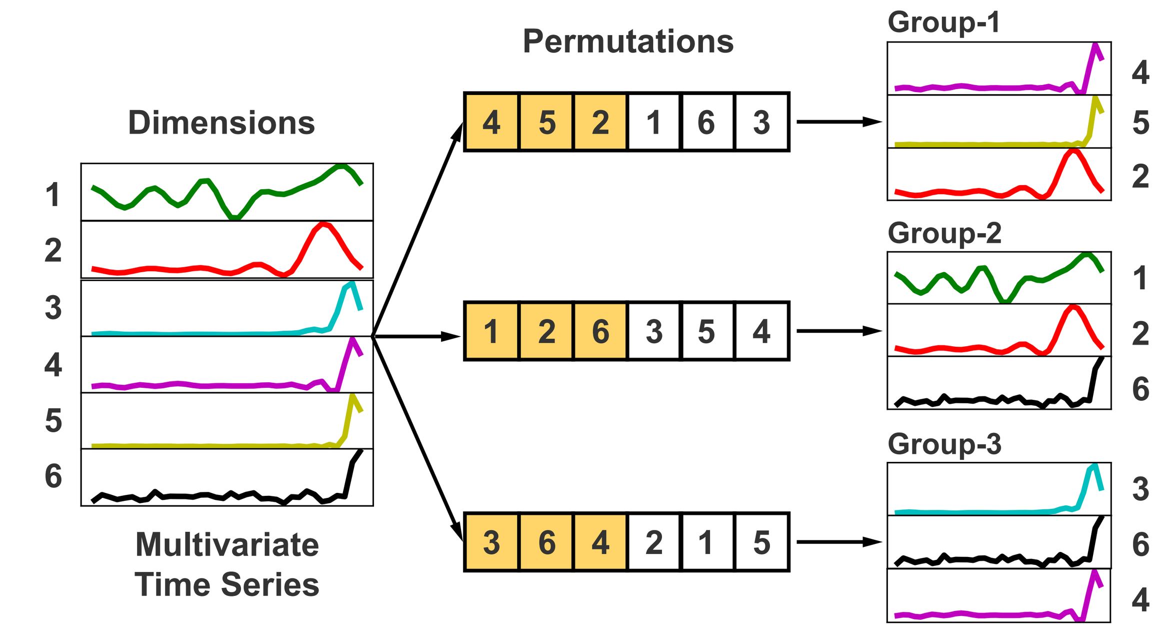
$$A_k = \Big(w_k^T \tanh \big(V_k H_k^T \big) \Big), \tag{2}$$

the distribution over classes for a given time series $x \in \mathcal{R}^d$ can be represented as a softmax over distances to the prototypes in the embedding space as follows:

$$p_\Theta(y = k|x) = \frac{\exp\big(-D(f_\Theta(x), c_k)\big)}{\sum_i \exp\big(-D(f_\Theta(x), c_i)\big)}, \tag{3}$$

where the function $D : \mathcal{R}^d \times \mathcal{R}^d \mapsto [0, +\infty)$ is the distance function to measure the distances between two embedding vectors. The distance function can be chosen from *regular Bregman divergences*Banerjee et al. (2005)

Random Dimension Permutation



Result

Multivariate Time Series Classification: We compare our proposed approach with eight different benchmark approaches, including the latest bag-of-patterns model based multivariate time series classification approach, deep learning framework, and common distance-based classifiers.

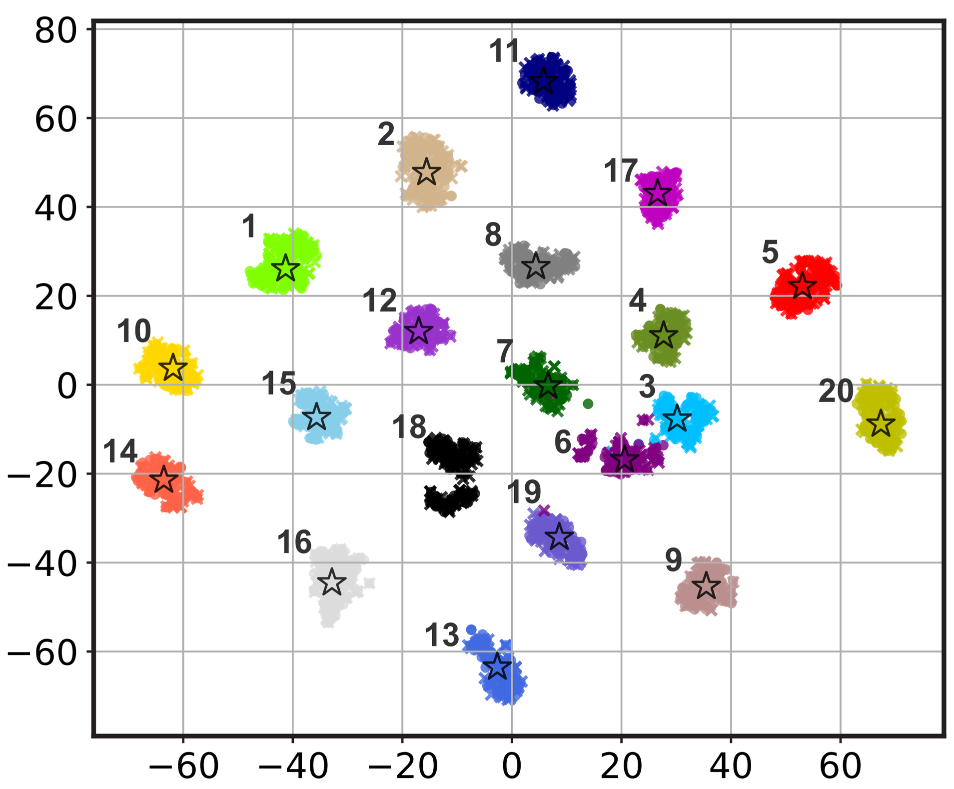
Dataset	TapNet	MLSTM -FCN	WEASEL +MUSE	ED-1NN	DTW-1NN-I	DTW-1NN-D	ED-1NN (norm)	DTW-1NN-I (norm)	DTW-1NN-D (norm)
ArticularyWordRecognition	0.987	0.973	**0.99**	0.97	0.98	0.987	0.97	0.98	0.987
AtrialFibrillation	**0.333**	0.267	**0.333**	0.267	0.267	0.2	0.267	0.267	0.22
BasicMotions	**1**	0.95	**1**	0.675	**1**	0.975	0.676	**1**	0.975
CharacterTrajectories	**0.997**	0.985	0.99	0.964	0.969	0.99	0.964	0.969	0.989
FaceDetection	**0.556**	0.545	0.545	0.519	0.513	0.529	0.519	0.5	0.529
HandMovementDirection	**0.378**	0.365	0.365	0.279	0.306	0.231	0.278	0.306	0.231
Heartbeat	**0.751**	0.663	0.727	0.62	0.659	0.717	0.619	0.658	0.717
MotorImagery	**0.59**	0.51	0.5	0.51	0.39	0.5	0.51	N/A	0.5
NATOPS	**0.939**	0.889	0.87	0.86	0.85	0.883	0.85	0.85	0.883
PEMS-SF	**0.751**	0.699	N/A	0.705	0.734	0.711	0.705	0.734	0.711
PenDigits	**0.98**	0.978	0.948	0.973	0.939	0.977	0.973	0.939	0.977
Phoneme	0.175	0.11	**0.19**	0.104	0.151	0.151	0.104	0.151	0.151
SelfRegulationSCP2	**0.55**	0.472	0.46	0.483	0.533	0.539	0.483	0.533	0.539
SpokenArabicDigits	0.983	**0.99**	0.982	0.967	0.96	0.963	0.967	0.959	0.963
StandWalkJump	**0.4**	0.067	0.333	0.2	0.333	0.2	0.2	0.333	0.2
Avg. Rank	**1.15**	4.23	3.23	5.76	5.15	4.46	6.15	5.38	4.7
Wins/Ties	**12**	1	4	0	0	1	0	1	0

Semi-supervised Multivariate Time Series Classification: We next evaluation the performance of our model on five datasets that have a significantly imbalanced training/test split.

Table 3: Performance of Semi-Supervised TapNet

Dataset (Training/Test)	TapNet	Semi-TapNet
Handwriting (150/850)	0.3565	**0.3882**
UWaveGestureLibrary (120/320)	0.894	**0.903**
ArticularyWordRecognition (275/300)	0.987	**0.993**
StandWalkJump (12/15)	**0.4**	**0.4**
JapaneseVowels (270/370)	0.965	**0.968**

Inspection of Class Prototype: Finally, we visualize the class prototypes and their corresponding time series embeddings:



References

Banerjee, A., Merugu, S., Dhillon, I. S., & Ghosh, J. (2005). Clustering with bregman divergences. *Journal of machine learning research*, *6*(Oct), 1705–1749.