Space-Time Domain Tensor Neural Networks: An Application on Human Activity Classification

L-Università ta' Malta
Institute of Digital Games

Konstantinos Makantasis[1], Athanasios Voulodimos[2], Anastasios Doulamis[3], Nikolaos Bakalos[3], Nikolaos Doulamis[3]

INTRODUCTION

Sensing technologies advances

Multimodal data

Spatiotemporal data

The current work has 2 **goals**:

1. Process and fuse multimodal spatiotemporal data
2. Create discriminative data representations for pattern recognition tasks

Application: Classification of human poses using 3D skeleton data

METHOD

Consider N sensors, J modalities and T time instances. Points from j-th modality are represented by a matrix $\boldsymbol{S}_j \in \mathbb{R}^{N \times T}$.

Steps:

1. Project $\boldsymbol{S}_j$ in $\mathbb{R}^{M \times T}$ using $\boldsymbol{W} \in \mathbb{R}^{N \times M}$
2. Compute feature

$$\boldsymbol{f}_j = \log\left[\frac{\mathrm{var}(\boldsymbol{W}^T\boldsymbol{S}_j^1)}{\sum_{i=1}^{M}\mathrm{var}(\boldsymbol{W}^T\boldsymbol{S}_j^i)} \cdots \frac{\mathrm{var}(\boldsymbol{W}^T\boldsymbol{S}_j^M)}{\sum_{i=1}^{M}\mathrm{var}(\boldsymbol{W}^T\boldsymbol{S}_j^i)}\right] \in \mathbb{R}^M$$

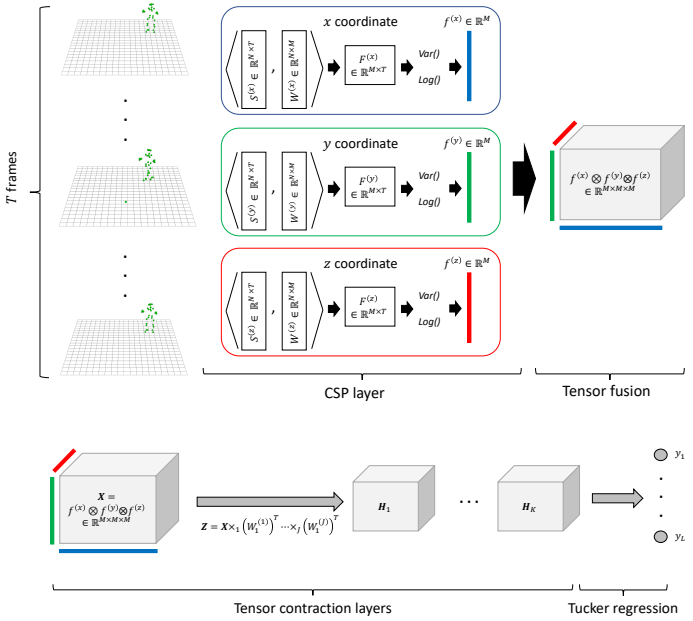
where $\boldsymbol{W}^T\boldsymbol{S}_j^i$ is the i-th row of $\boldsymbol{W}^T\boldsymbol{S}_j$

3. Fuse different modalities into a tensor

$$\boldsymbol{\mathcal{X}} = \boldsymbol{f}_1 \otimes \cdots \otimes \boldsymbol{f}_J \in \mathbb{R}^{M \times M \times \cdots \times M}$$

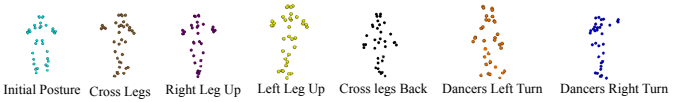
4. Classify $\boldsymbol{\mathcal{X}}$ using a tensor neural network

MODEL



RESULTS

Dataset:



- Kinect II 3D skeleton data
- 10 sessions, 3 dancers, 4 dances
- 10-fold cross validation

	Accuracy(%)	F1 Score (%)
LSTM	84.2%	82.0%
BOBi LSTM*	85.4%	80.7%
1D-CNN	91.1%	89.7%
Our approach	**91.6%**	**90.9%**

DISCUSSION

- The network is *end-to-end-trainable*
- State-of-the-art results with
  - 87 times less parameters than BOBi LSTM
  - 2 times less parameters than 1D-CNN

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Take a picture to download the full paper

*I. Rallis et al. "Learning choreographic primitives through a bayesian optimized bi-directional lstm model," in IEEE ICIP, 2019, pp. 1940–1944.

[1] University of Malta, Malta
[2] University of West Attica, Greece
[3] National Technical University of Athens, Greece