Cyclic-Permutation Invariant Networks for Modeling Periodic Sequential Data: Application to Variable Star Classification

@AstroKeming
arXiv: 2011.01243

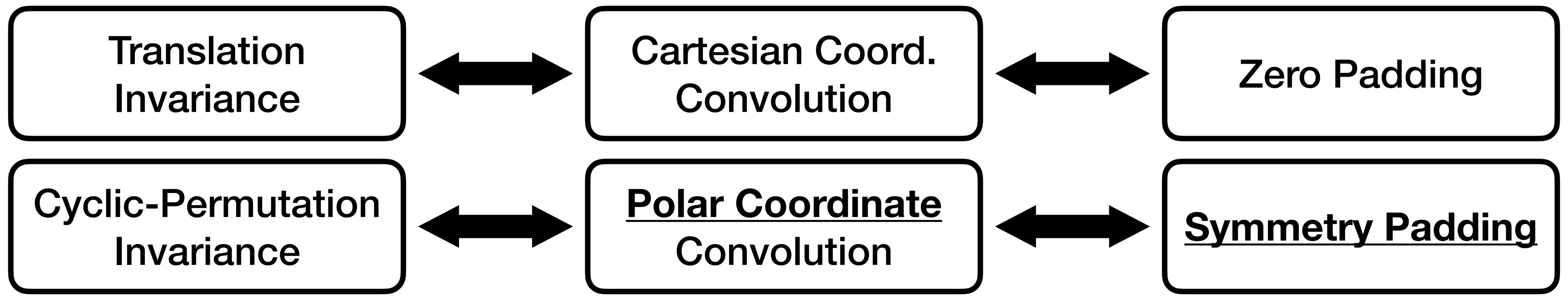
Keming Zhang (张可名)
University of California at Berkeley

Joshua S. Bloom
University of California at Berkeley

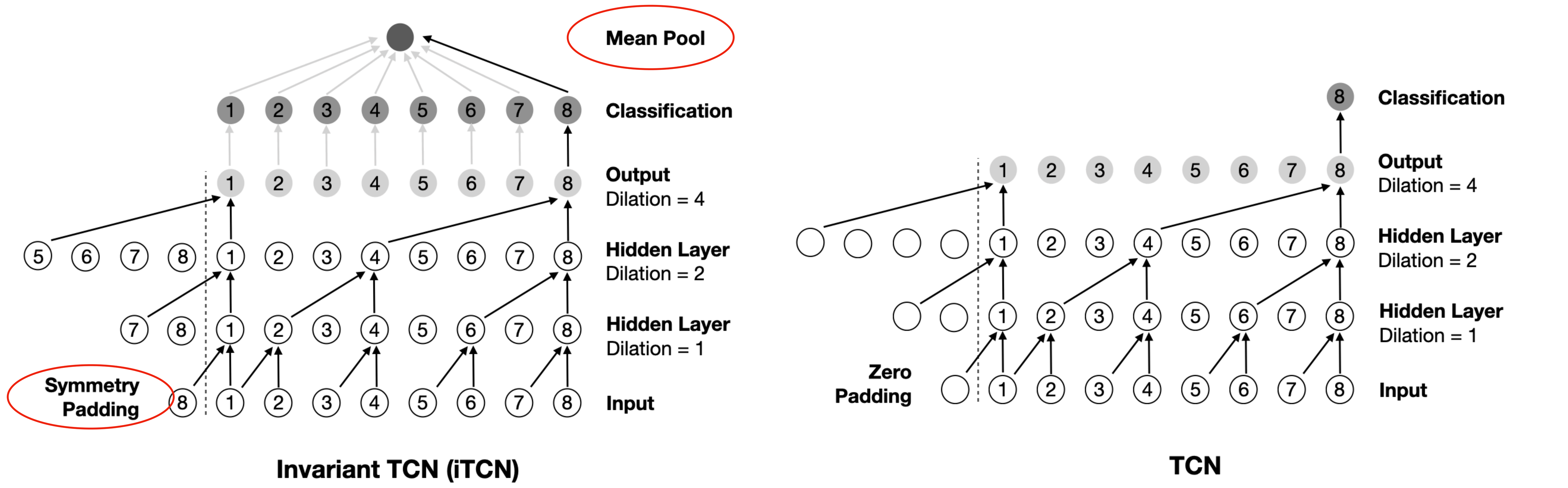
Berkeley
UNIVERSITY OF CALIFORNIA

Summary

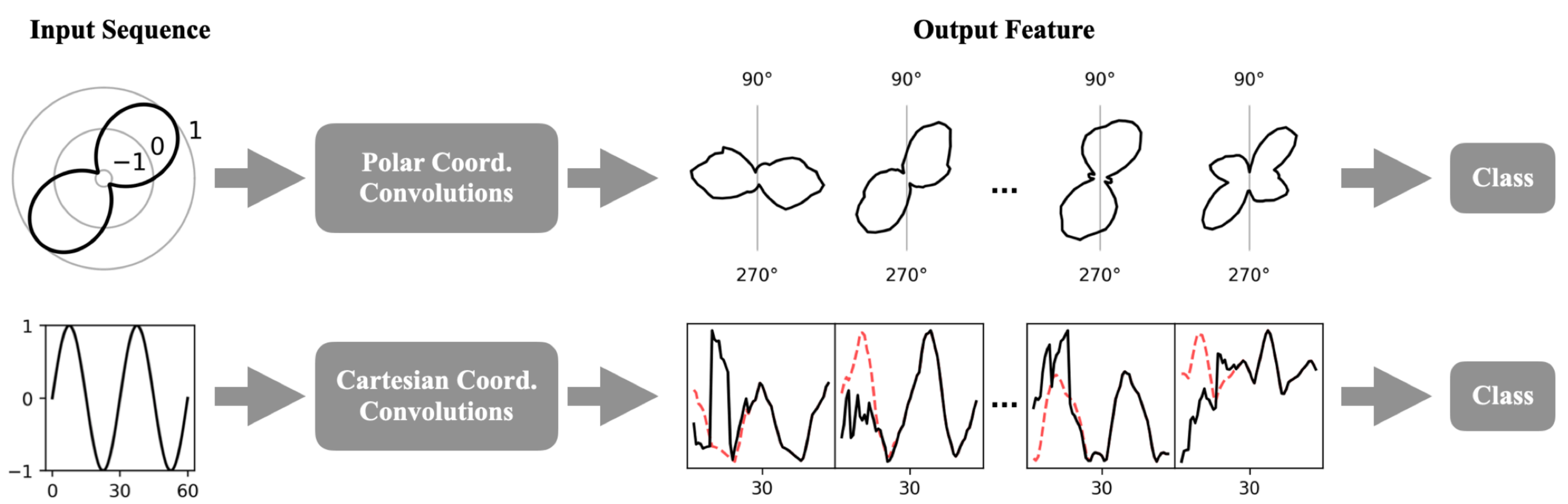
- **Periodic data is cyclic**: it can be wrapped in a closed ring after period folding (phase = t mod period)
- **Current RNN/CNN is acyclic**, and therefore not optimal for periodic data.
- **Cyclic neural network should be used to model cyclic data.**



Implementation: Symmetry Padding



(above figure) Simplified illustration of the **cyclic-permutation invariant Temporal Convolutional Network (iTCN)** and the usual TCN. Numbers refer to the ordering of the period-folded sequence. Dilated convolutions are represented by arrows where the dilation factor is indicated to the right of each layer. Gray arrows in the final two layers represent operations which are present only in the iTCN not the TCN. The classification layer consists of two convolutions of kernel size 1. **A comparison of the left and the right shows that cyclic-permutation is achieved by Symmetry padding and a final global mean pooling layer, which are circled in red.**



(above figure) Schematic illustration of the effect of polar coordinate convolutions in preserving cyclic-permutation invariance. The input and output sequences are shown in polar coordinates for iTCN, and in Cartesian coordinates for TCN. **The input sequence is a sine curve with two full oscillations** in both cases. In the upper diagram, 1-D **feature maps of the periodic input remains periodic; rotational symmetry is preserved**. These periodic feature maps are also shown in Cartesian coordinates of the lower plots in red dashed lines for comparison. As demonstrated by the discrepancy, **feature maps are distorted for the first full oscillation in the non-invariant network, which is shown in solid black lines.**

Results: Periodic Variable Star Classification

Model	MACHO	OGLE-III	ASAS-SN
iTCN	**92.7%** ± 0.43%	**93.7%** ± 0.09%	**94.5%** ± 0.14%
TCN	92.0% ± 0.35%	92.9% ± 0.11%	93.0% ± 0.20%
–[1]	$(-0.58\%^{+0.05\%}_{-0.17\%})$	$(-0.75\%^{+0.04\%}_{-0.02\%})$	$(-1.52\%^{+0.12\%}_{-0.09\%})$
iResNet*	92.6% ± 0.45%	**93.7%** ± 0.09%	93.9% ± 0.14%
ResNet	92.1% ± 0.34%	93.4% ± 0.11%	93.2% ± 0.22%
–[1]	$(-0.48\%^{+0.05\%}_{-0.13\%})$	$(-0.25\%^{+0.03\%}_{-0.02\%})$	$(-0.64\%^{+0.01\%}_{-0.07\%})$
GRU	92.3% ± 0.37%	92.8% ± 0.19%	93.6% ± 0.42%
–[2]	$(-0.34\%^{+0.05\%}_{-0.02\%})$	$(-0.86\%^{+0.17\%}_{-0.13\%})$	$(-0.71\%^{+0.07\%}_{-0.12\%})$
LSTM	91.7% ± 0.53%	92.6% ± 0.61%	93.5% ± 0.23%
–[2]	$(-0.93\%^{+0.45\%}_{-0.16\%})$	$(-0.85\%^{+0.17\%}_{-0.48\%})$	$(-1.00\%^{+0.11\%}_{-0.06\%})$

Table 1. Ablation study test accuracies demonstrating gains afforded by cyclic-permutation invariance. The network with the top accuracy for each dataset is shown in bold. Test accuracies are the mean values for 8 different data splits. Median test accuracy differences of the different data partitions are shown in parentheses with the uncertainty interval corresponding to 1-σ range of test accuracy differences calculated pair-wise for the same random partitions of data.

[1]Compared to the invariant version of the same network.

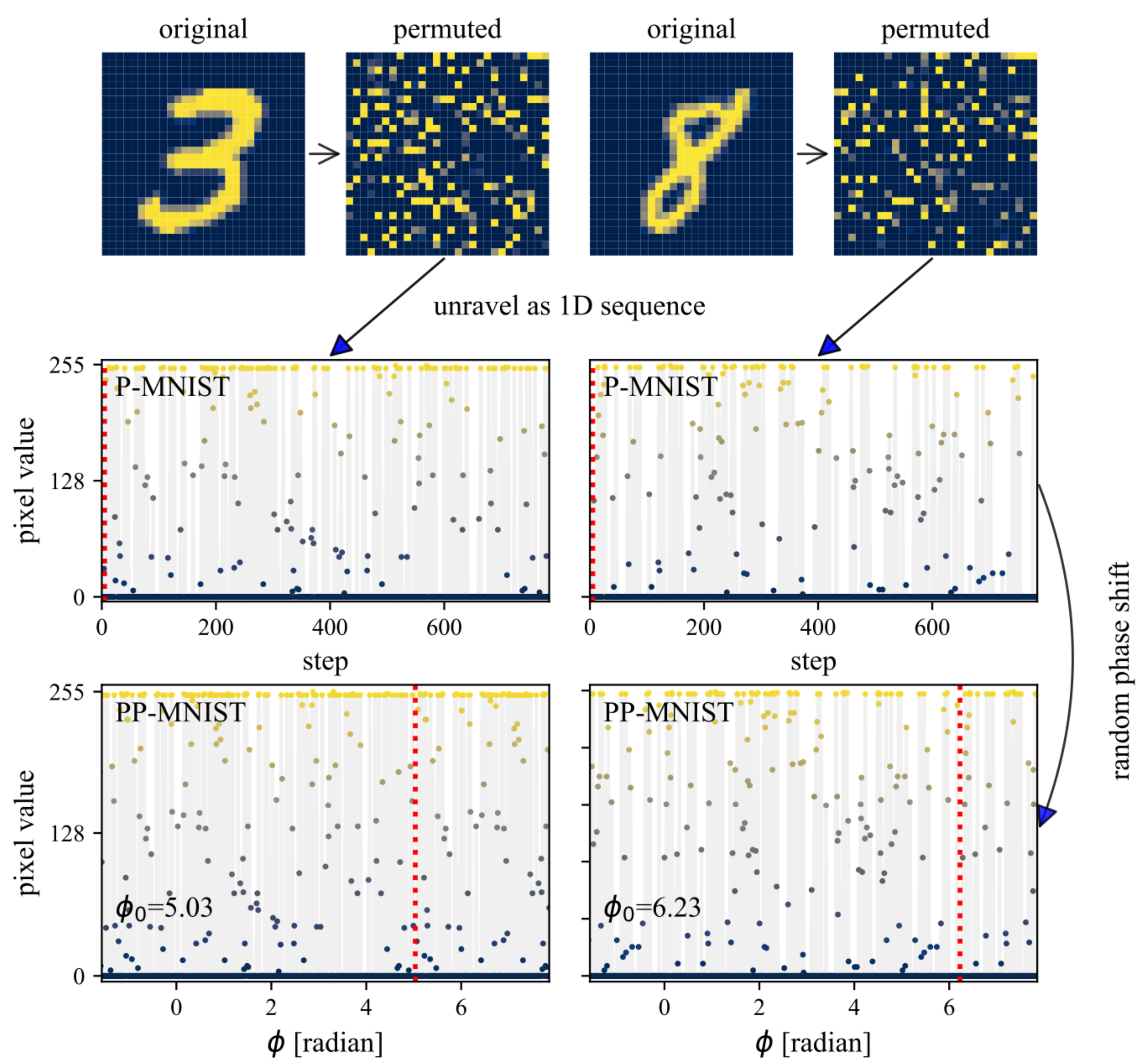
[2]Compared to the best performing network.

*Semi-invariant due to use of discrete max-pooling layers.

Astronomers routinely measure the change in brightness of various astronomical objects on a nightly basis with so time-domain surveys. Among these objects are **variable stars whose brightness vary periodically** with periods ranging from minutes to weeks. They are dozens of classes of variable stars, each varies it's brightness in its characteristic way and can be classified according to the shape of light-curves, or time series of brightness.

MACHO, OGLE-III, and ASAS-SN are three different datasets of variable star light curves. We segment each light curve into **L=200 equal length segments** in temporal order and **transform from time-space into phase-space, i.e. period-folding.**

Custom Task: Periodic Permuted MNIST



Periodic permuted MNIST (PP-MNIST) is our custom dataset which is a periodic variation of the sequential MNIST task.

For PP-MNIST, each MNIST image is first unrolled into a 1D sequence and permuted with a fixed order to remove the spatial/temporal structure. Then, a random cyclic-permutation is applied to each 1D sequence. This essentially makes the sequence periodic in that there's no unique "zero-index."

iResNet	**96.0%**	iTCN	**94.8%**
ResNet	95.1%	TCN	77.4%

(above table) Periodic permuted MNIST (PP-MNIST) classification accuracies.