NRITYANTAR: POSE OBLIVIOUS INDIAN CLASSICAL DANCE SEQUENCE CLASSIFICATION SYSTEM

Prerana Mukherjee*, Vinay Kaushik*, Brejesh Lall
Indian Institute of Technology, Delhi.

at 11th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP 2018)
December 18 - 22, Hyderabad, India

Abstract

In this paper, we attempt to advance the research work done in human action recognition to a rather specialized application namely Indian Classical Dance (ICD) classification. The variation in such dance forms in terms of hand and body postures, facial expressions or emotions and head orientation makes pose estimation an extremely challenging task. We construct a pose-oblivious shape signature which is fed to a sequence learning framework. We represent person-pose in first frame of a dance video using symmetric Spatial Transformer Networks (STN) to extract good person object proposals and CNN based parallel single person pose estimator (SPPE). The pose basis are converted to pose flows by assigning a similarity score between successive poses followed by non-maximal suppression. Instead of feeding a simple chain of joints in the sequence learner which generally hinders the network performance we constitute a feature vector of the normalized distance vectors, flow, angles between anchor joints which captures the adjacency configuration in the skeletal pattern. Thus, the kinematic relationship amongst the body joints across the frames using pose estimation helps in better establishing the spatio-temporal dependencies. We present an exhaustive empirical evaluation of state-of-the-art deep network based methods for dance classification on ICD dataset.

Fig. 1: Indian Classical Dance (ICD) Forms:

(a) Bharatnatyam
(b) Kuchipudi
(c) Manipuri
(d) Kathak
(e) Mohiniattam
(f) Odissi

Fig. 2: (a) Pose of a Kathak dancer (b) Visualization of the anchor joints

Methodology

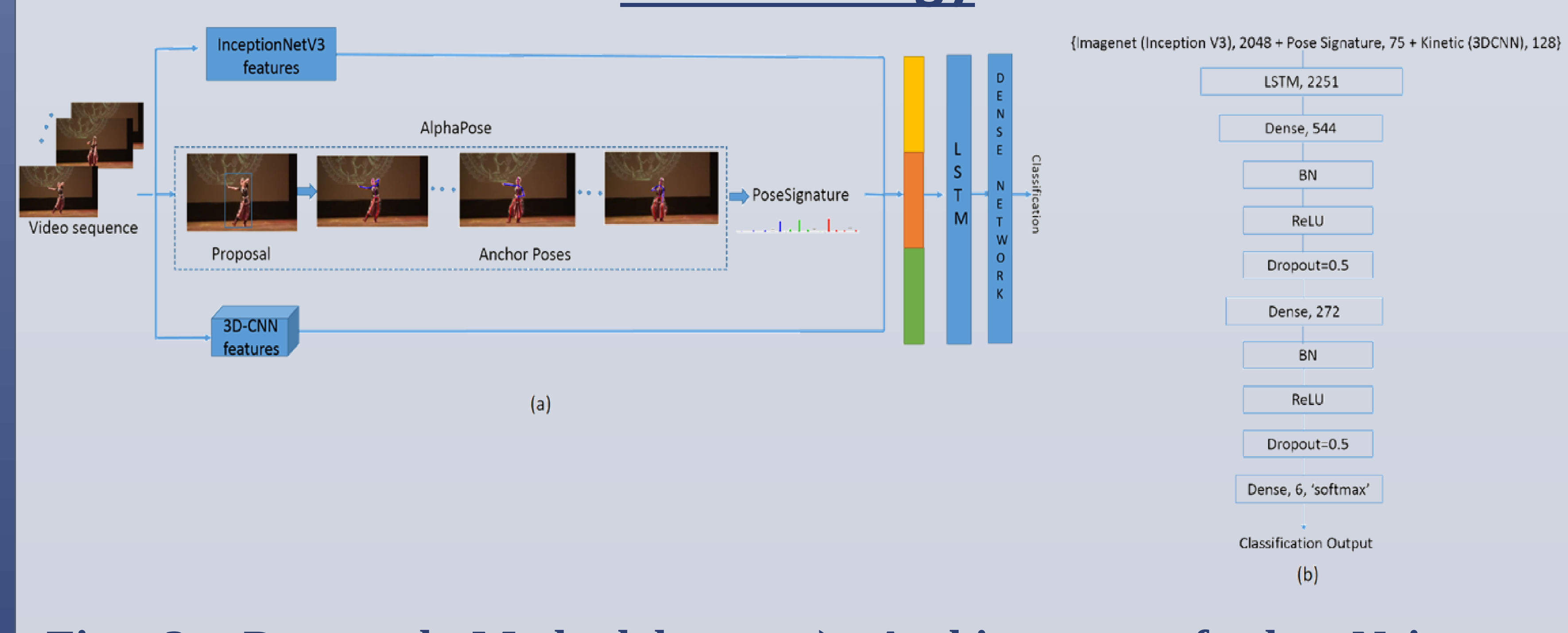

Fig. 3: Proposed Methodology: a) Architecture of the Nrityantar framework b) Block diagram of the Sequence learning framework in Nrityantar. The numbers in the layers indicate the number of features

Tables and Results

Optimizer	Adam
Loss	Categorical Cross-entropy
Learning rate	0.0001
Decay	0.000001
Feature length	2251
Output length	6
Sequence Length (One training sample)	48 frames (Uniformly Distributed)
Batch Size	32
Maximum Epoch	100

Table 1: Parameters used for training the LSTM

Method	Bharatnatyam	Kathak	Kuchipudi	Manipuri	Mohiniattam	Odissi	Average Accuracy
InceptionV3	80.48	62.71	28.90	30.64	98.41	53.62	59.1
Pose Signature	88.15	79.31	43.65	41.50	96.82	71.01	67.41
Kinetics	84.21	68.96	56.34	30.18	96.82	57.97	65.61
InceptionV3+ Pose Signature	51.31	67.24	65.87	67.92	93.65	73.91	68.98
InceptionV3+ Kinetics	85.52	68.96	63.49	24.53	96.82	57.97	67.19
InceptionV3+ Pose Signature+ Kinetics	86.84	87.93	70.63	24.52	93.65	63.76	72.35

Table 2: Comparison of various features and their combinations for Dance Classification

	Bharatnatyam	Kathak	Kuchipudi	Manipuri	Mohiniattam	Odissi
Bharatnatyam	66	0	9	0	0	1
Kathak	0	51	5	1	1	0
Kuchipudi	15	9	89	5	0	8
Manipuri	18	17	4	13	1	0
Mohiniattam	0	2	0	2	59	0
Odissi	24	0	1	0	0	44

Table 3: Confusion Matrix

Dance Class	Precision	Recall	F1-Score	Support	Class Accuracy
Bharatnatyam	0.54	0.87	0.66	76	86.84
Kathak	0.65	0.88	0.74	58	87.93
Kuchipudi	0.82	0.71	0.76	126	70.63
Manipuri	0.62	0.25	0.35	53	24.52
Mohiniattam	0.97	0.94	0.95	63	93.63
Odissi	0.83	0.64	0.72	69	63.76
Average	0.75	0.72	0.71	445	**72.35**

Table 4: Final Results

Conclusion

- We have presented a novel pose signature in a sequential learning framework for Indian Classical Dance (ICD) classification. We incorporated pose, flow and spatio-temporal dependencies in the pose signature to capture the adjacency relationship between anchor joints in the skeletal pattern of the dancer.
- We performed exhaustive experiments and demonstrated the effectiveness of the proposed methodology on dance classification.
- We showed that deep descriptors with handcrafted pose signature outperformed on ICD dataset. We also showed that due to high similarities between dance moves and dressing attires it is highly challenging to classify dance sequences.

References

- Chéron, G., Laptev, I. and Schmid, C., 2015. P-cnn: Pose-based cnn features for action recognition. In Proceedings of the IEEE international conference on computer vision (pp. 3218-3226).
- Samanta, S., Purkait, P. and Chanda, B., 2012, January. Indian Classical Dance classification by learning dance pose bases. In Applications of Computer Vision (WACV), 2012 IEEE Workshop on (pp. 265-270). IEEE.

Contact: Dr. Brejesh Lall (brejesh@ee.iitd.ac.in)