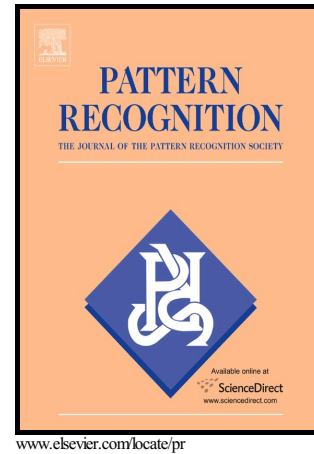


Author's Accepted Manuscript

Convolutional Neural Random Fields for Action Recognition

Caihua Liu, Jie Liu, Zhicheng He, Yujia Zhai, Qinghua Hu, Yalou Huang



PII: S0031-3203(16)30004-8
DOI: <http://dx.doi.org/10.1016/j.patcog.2016.03.019>
Reference: PR5680

To appear in: *Pattern Recognition*

Received date: 2 August 2015
Revised date: 13 March 2016
Accepted date: 14 March 2016

Cite this article as: Caihua Liu, Jie Liu, Zhicheng He, Yujia Zhai, Qinghua Hu and Yalou Huang, Convolutional Neural Random Fields for Action Recognition *Pattern Recognition*, <http://dx.doi.org/10.1016/j.patcog.2016.03.019>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and a review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain

Convolutional Neural Random Fields for Action Recognition

Caihua Liu^{a,b}, Jie Liu^{a,b,*}, Zhicheng He^{a,b}, Yujia Zhai^{a,b}, Qinghua Hu^{c,**}, Yalou Huang^{a,b}

^aCollege of Computer and Control Engineering, Nankai University, Tianjin, China

^bCollege of Software, Nankai University, Tianjin, China

^cSchool of Computer Science and Technology, Tianjin University, Tianjin, China

Abstract

A deep discriminative structured model, Convolutional Neural Random Fields (CNRF), is proposed for action recognition problem. In the proposed model, a spatio-temporal convolutional neural network (CNN) is developed for invariant feature learning from raw input frames, and the CNN is combined with Conditional Random Fields (CRF) for capturing the interdependencies between outputs. The parameters from both CRF and CNN are learned in a joint fashion which enables structured prediction and feature learning as well. We also explore different combinations of observation and transition feature functions based on the learned high level features from convolution part. The approach enjoys the advantages of both CNN and CRF, it has the invariant feature learning ability possessed by the former and structured prediction ability of the latter. The experimental results on both segmented and unsegmented human action recognition datasets show that CNRF boosts the performance over the comparison methods by a large margin.

Keywords: Action recognition; Sequence labeling; Conditional random fields; Spatio-temporal convolution;

*Corresponding author

**Joint corresponding author

Email addresses: jliu@nankai.edu.cn (Jie Liu), huqinghua@tju.edu.cn (Qinghua Hu)

Download English Version:

<https://daneshyari.com/en/article/4969949>

Download Persian Version:

<https://daneshyari.com/article/4969949>

[Daneshyari.com](https://daneshyari.com)