Extreme Learning Machine for Large-Scale Action Recognition

Gül Varol and Albert Ali Salah
Dept. of Computer Engineering, Boğaziçi University, İstanbul, Turkey

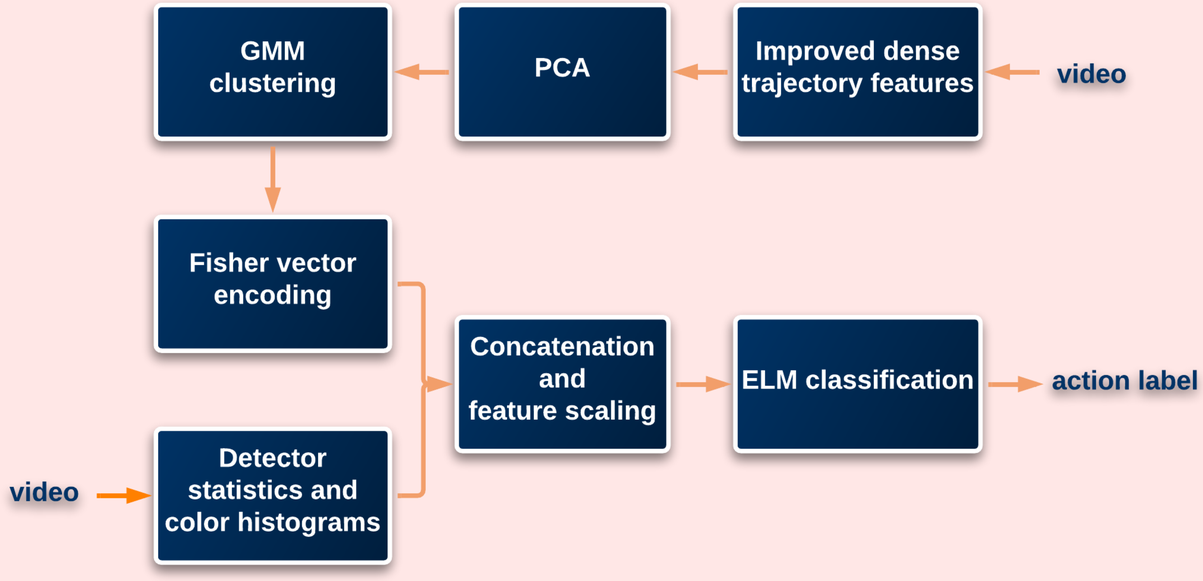
Overview

This work presents the method we applied for the action recognition task on the THUMOS 2014 challenge dataset. We study **human action recognition in RGB videos** through low-level features by focusing on **improved trajectory features** that are densely extracted from the spatio-temporal volume. We represent each video with **Fisher vector** encoding and additional mid-level features. Finally, we use **Extreme Learning Machines** for classification and achieve 62.27% mean average precision on the validation set.

[1]

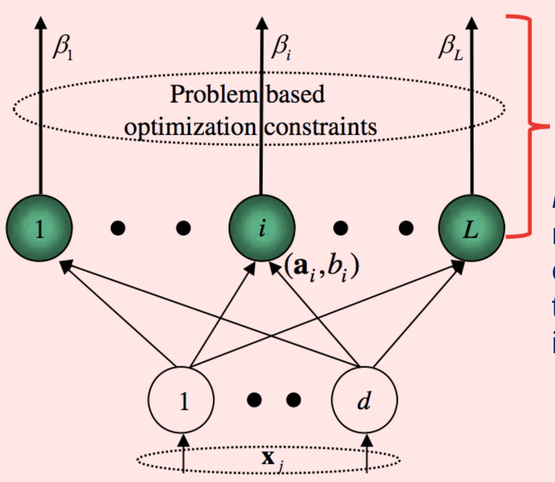
Feature Extraction

- Feature extraction, encoding and classification pipeline



Extreme Learning Machine

- Extremely fast alternative to other conventional popular learning algorithms
- Works for the generalized single-hidden-layer feed-forward networks
- No iterative tuning



Feature learning
Clustering
Regression
Classification

L Random Hidden Neurons (which need not be algebraic sum based) or other ELM feature mappings. Different type of output functions could be used in different neurons:

$$h_i(x) = G_i(\boldsymbol{a}_i, b_i, \boldsymbol{x})$$

[2]

d Input Nodes

Experimental Results

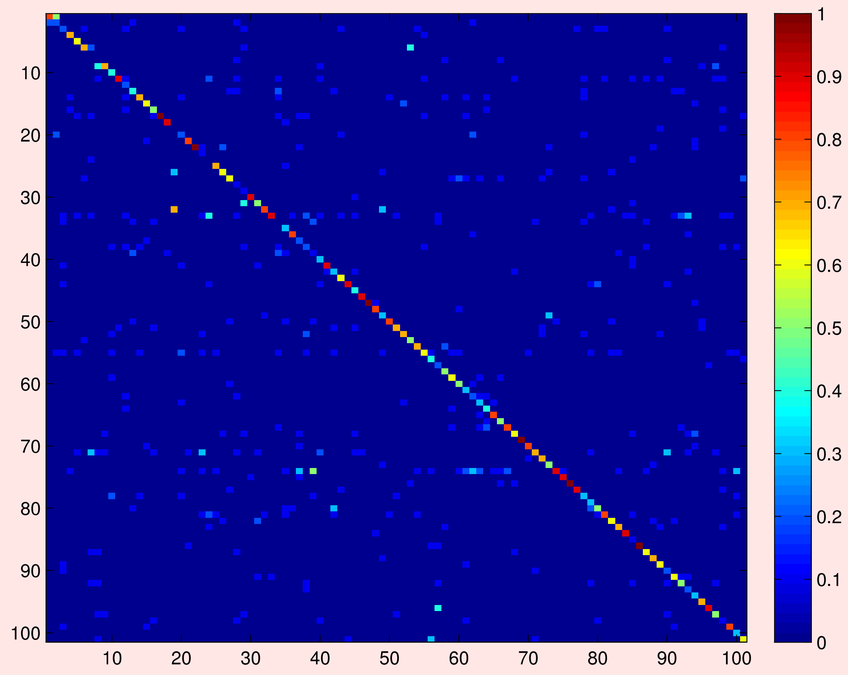
- Mean average precision values for various feature sets within improved dense trajectory features

Feature set	mAP (%)
HOG	50.18
HOF	52.77
MBH	55.72
MBH + HOG	56.78
HOG + HOF	59.26
MBH + HOG	60.61
MBH + HOG + HOF	61.02

- Bag of Features and Fisher Vector comparison for encoding MBH+HOG+HOF

Encoding	mAP (%)	Dimensionality
BOF	40.64	12000
FV	61.02	24576

- Confusion matrix on the validation set



- SVM and ELM comparison with the best feature combination (i.e. MBH+HOG+HOF+DS+RGBH+HSVH)

Algorithm	mAP (%)	Training time (sec)	Testing time (sec)
SVM	43.94	51920 (~14h)	885
ELM	62.27	92	11

References

1. Jiang, Y.G., Liu, J., Roshan Zamir, A., Toderici, G., Laptev, I., Shah, M., Sukthankar, R.: THUMOS challenge: Action recognition with a large number of classes. http://crcv.ucf.edu/THUMOS14/ (2014)

2. Huang, G. B.: An insight into extreme learning machines: Random neurons, random features and kernels. Cognitive Computation 6(3), 376–390 (2014).