

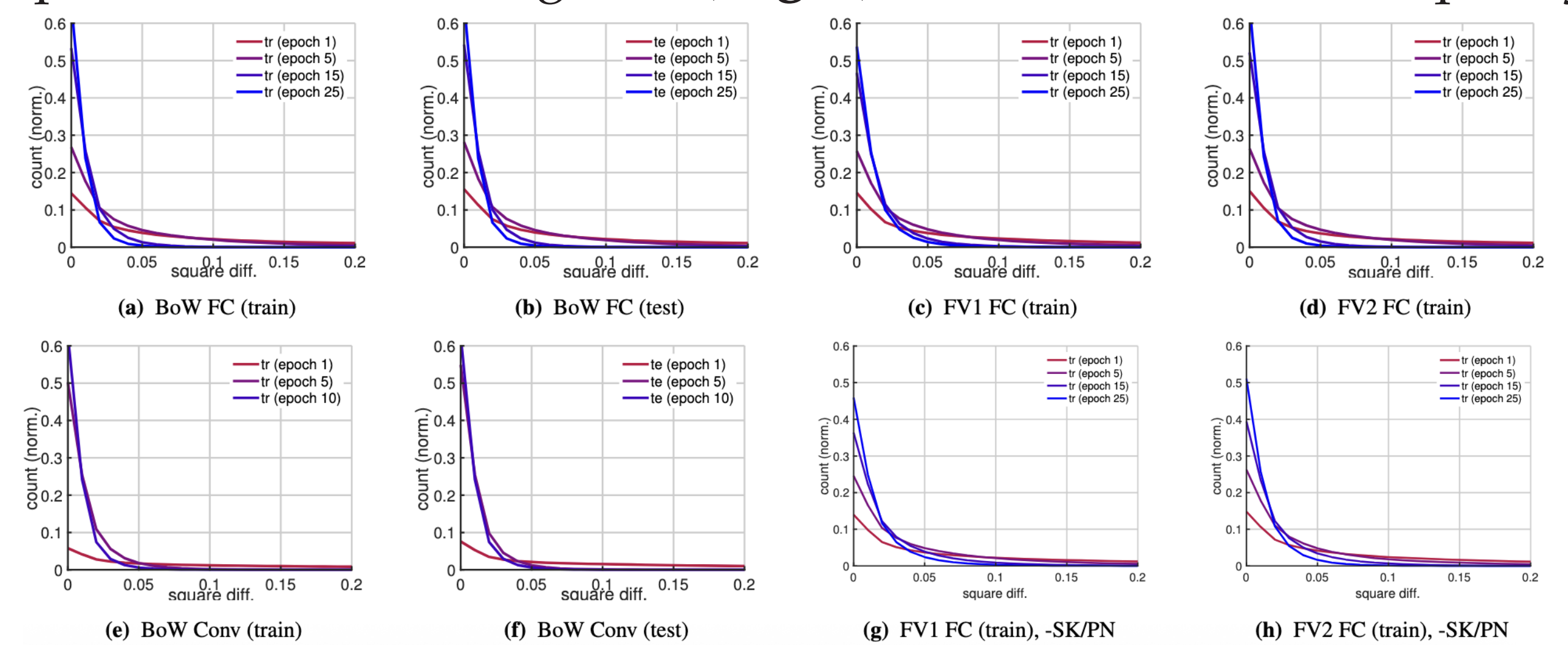
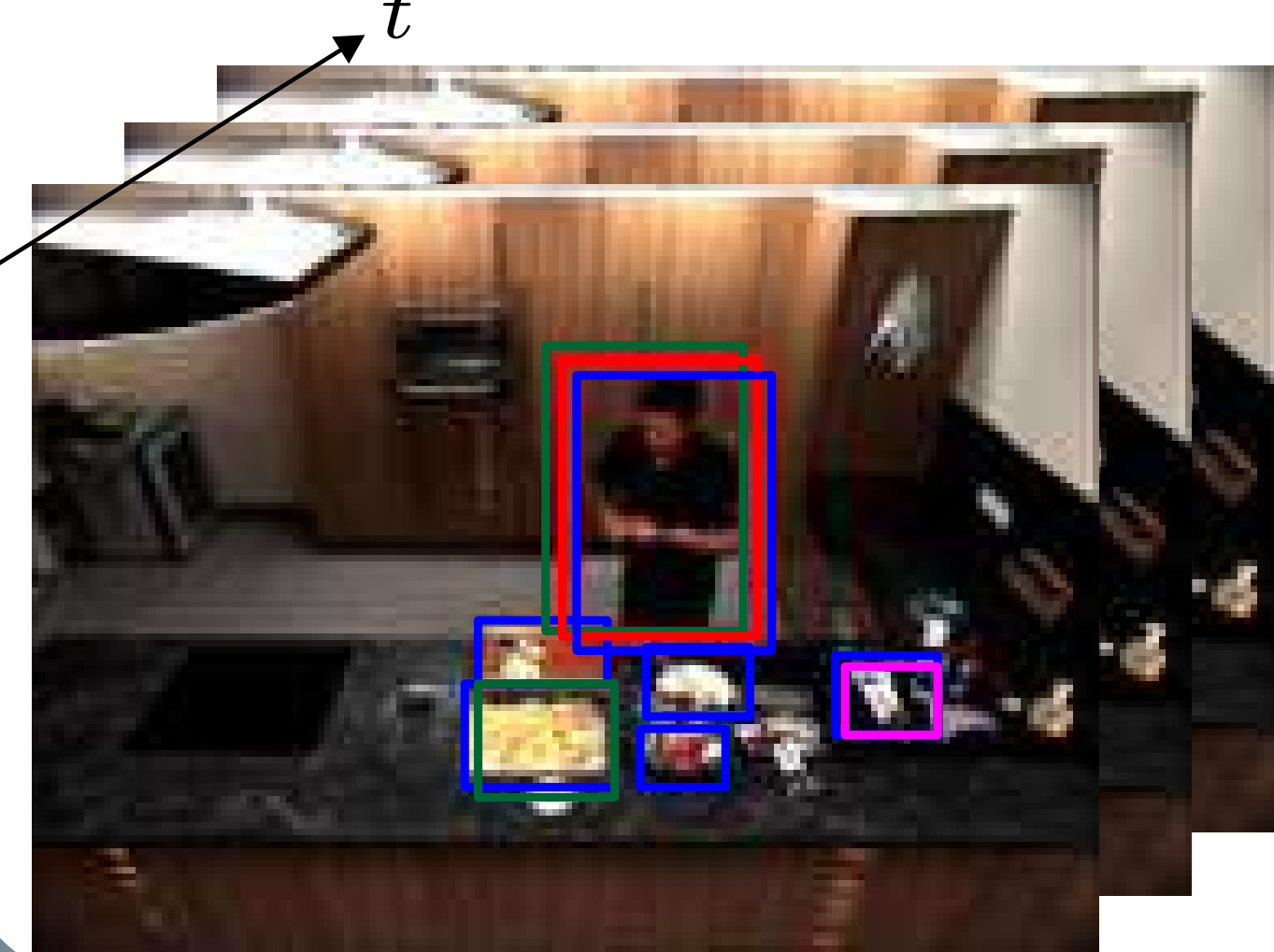
Feature Hallucination for Self-supervised Action Recognition

Lei Wang Piotr Koniusz

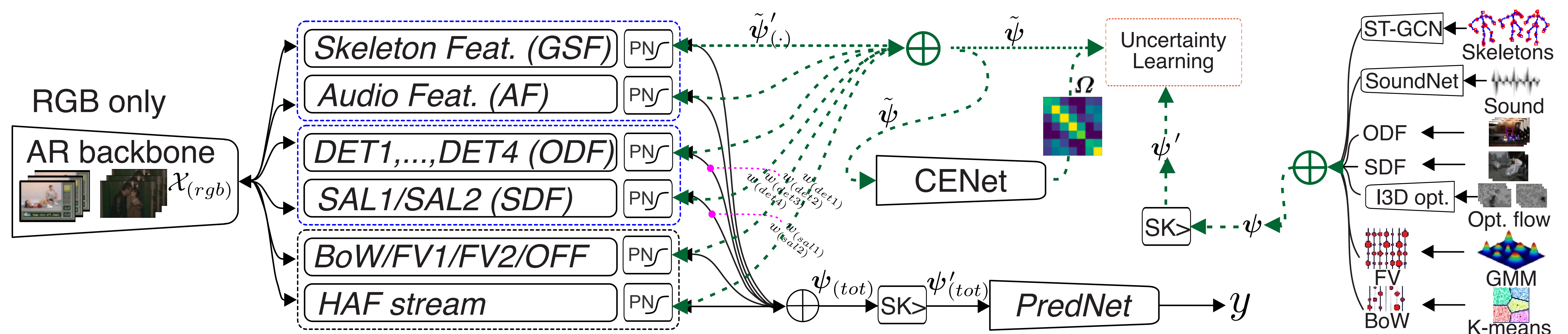


Motivation and key ideas

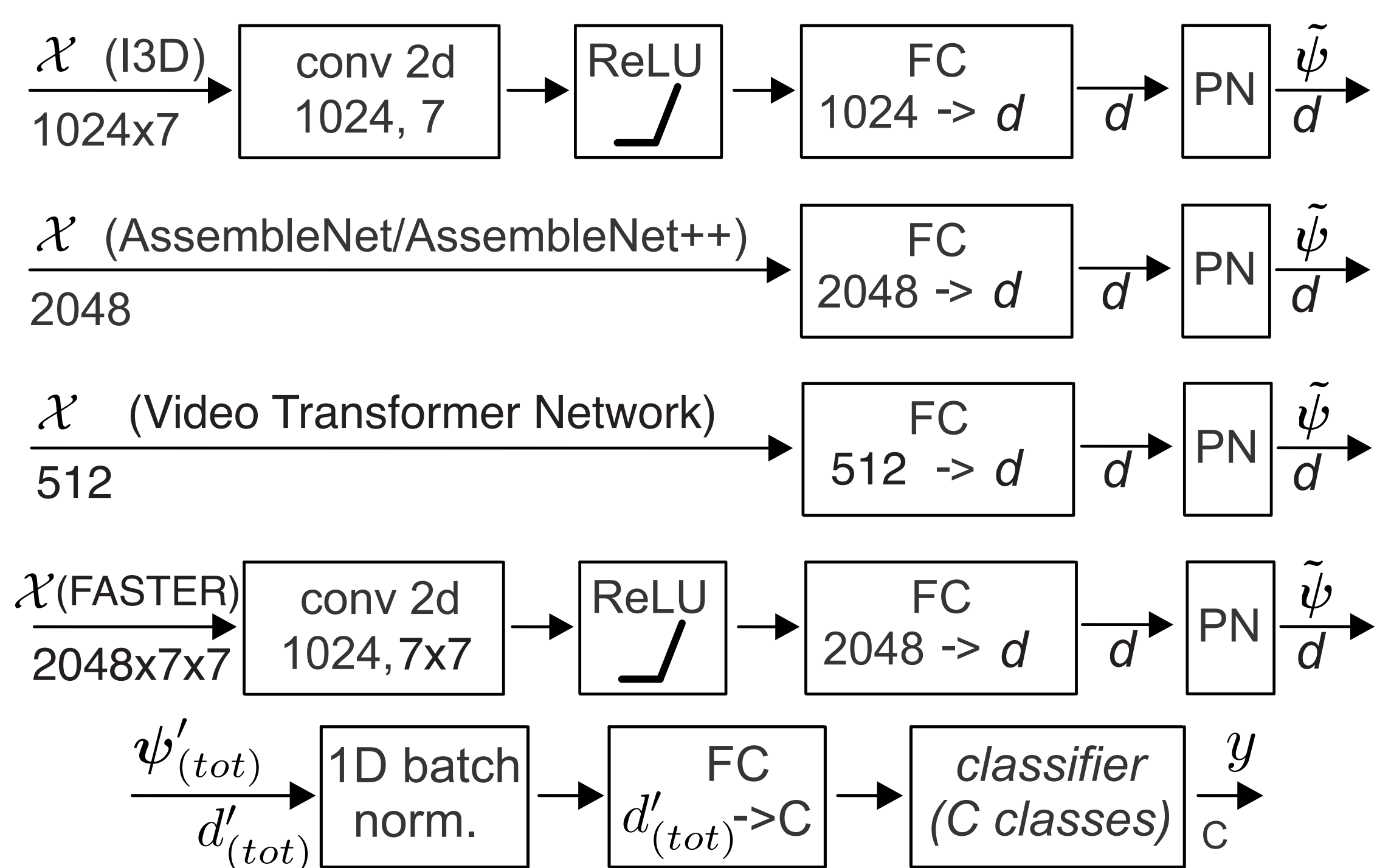
(Left) Bounding boxes from four detectors; (Middle) saliency detectors highlighting spatial and motion regions; (Right) feature hallucination quality.



The pipeline: further details

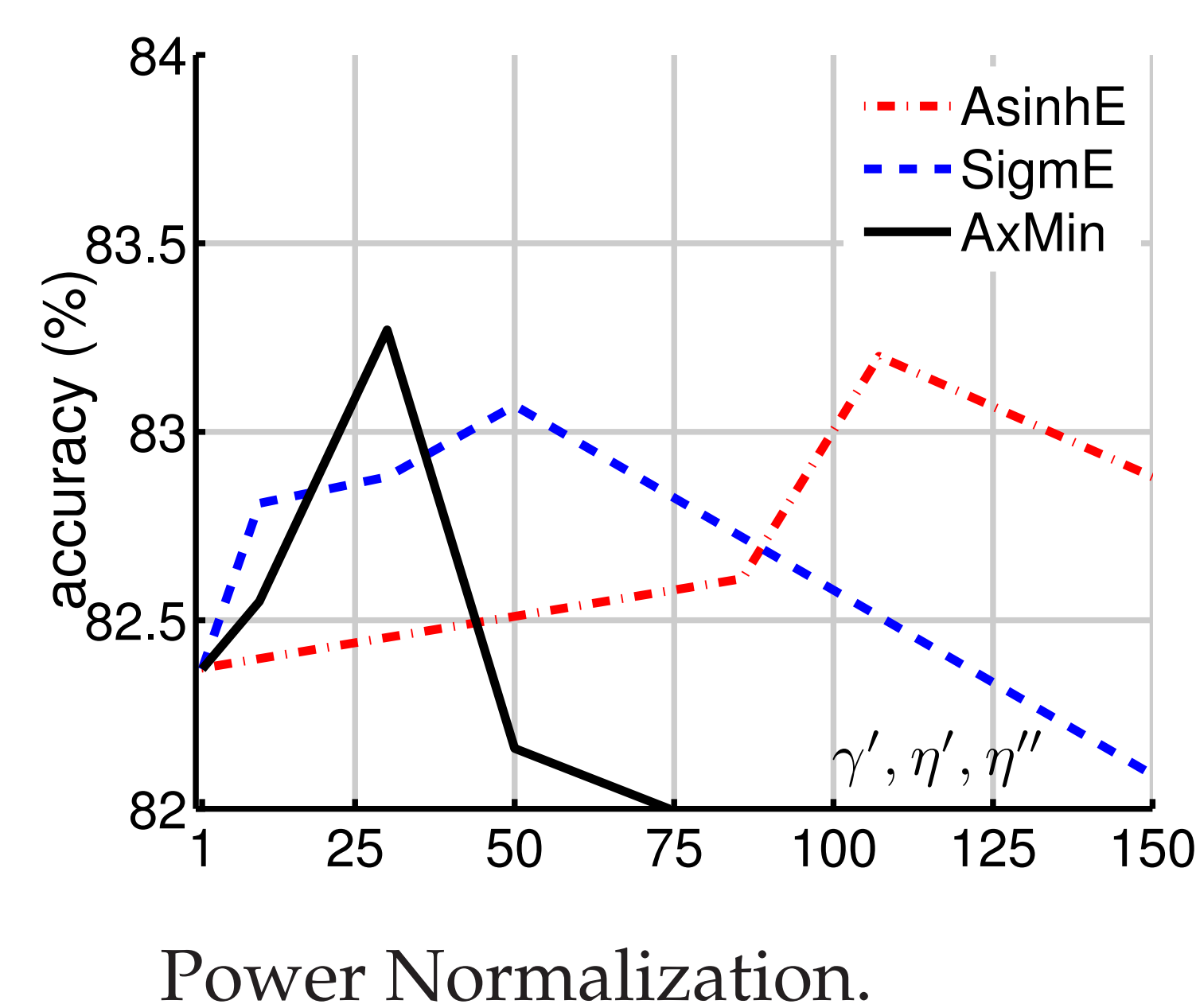


A conceptual overview of our multimodal self-supervised action recognition framework.

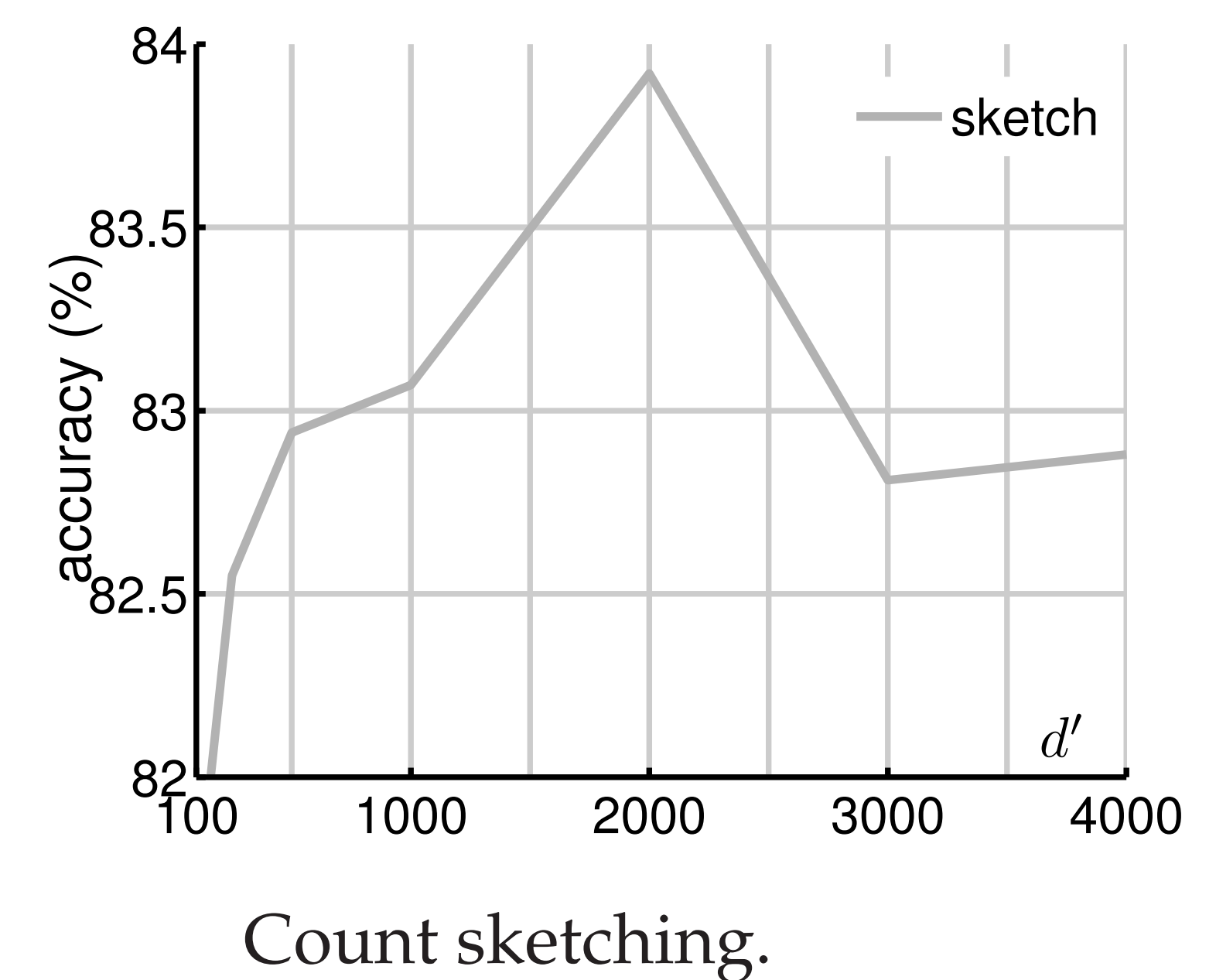


Stream details.

(Left) Top four subfigures show the architectures used for the BoW, FV, OFF, HAF, ODF, SDF, GSF, and AF streams with four different backbones; the bottom subfigure shows the PredNet architecture. (Right) Evaluation of Power Normalization and sketching on HMDB-51 (split 1).



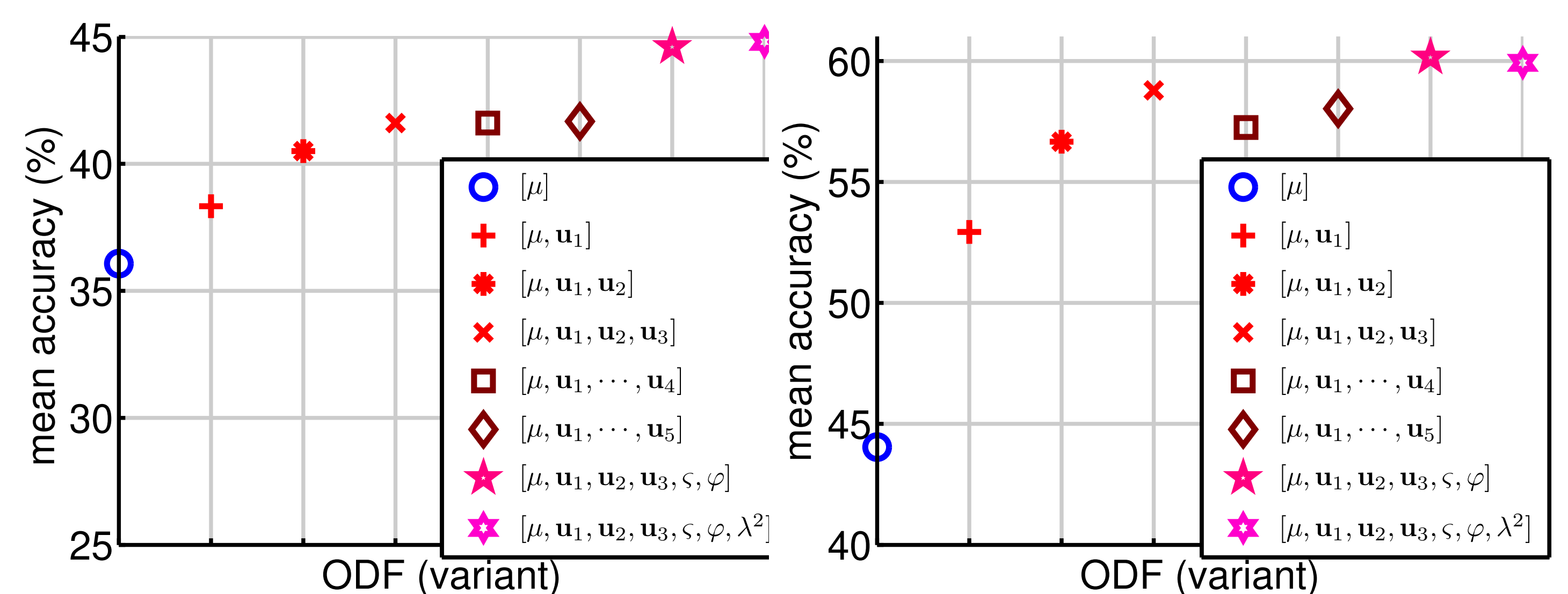
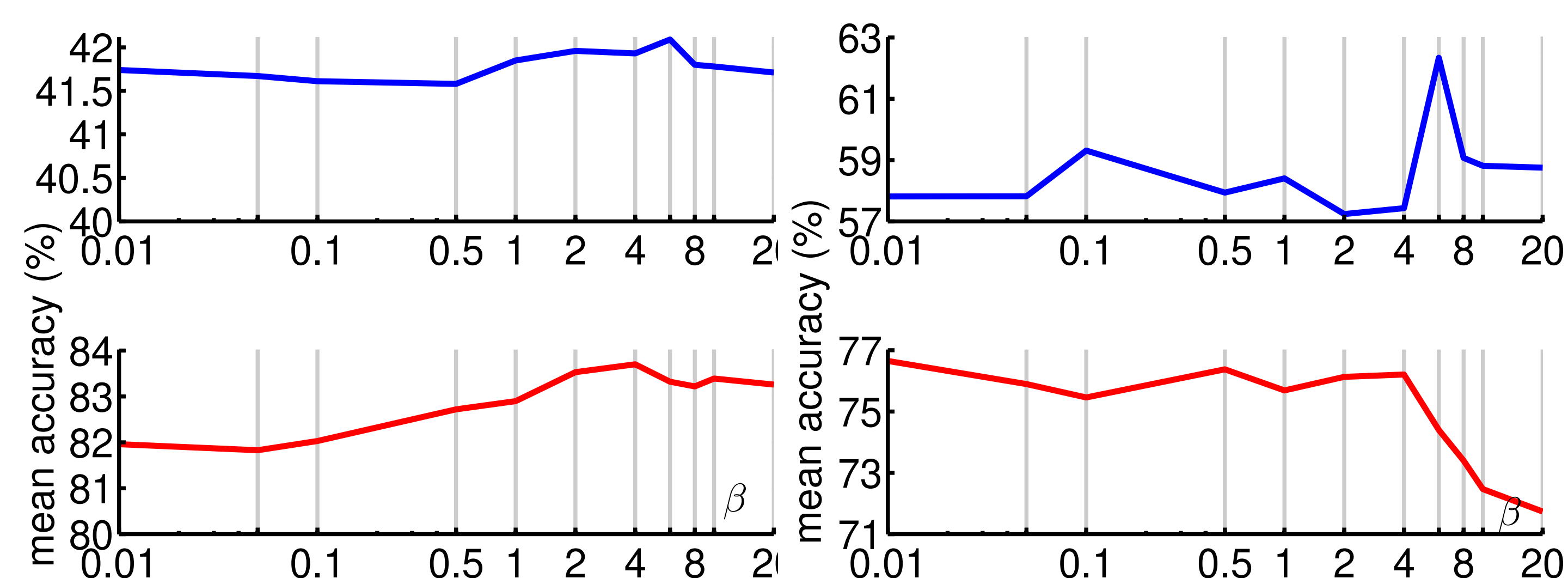
Power Normalization.



Count sketching.

Results

(Left) Impact of β in the weighted mean on classification performance. Results on HMDB-51 and YUP++: (top) four combined detectors + SVM; (bottom) DEEP-HAL with four combined detectors + SVM. (Right) Evaluation of ODF with SVM using the weighted mean across four detectors.



Visualization of the feature space (extracted from PredNet) for DEEP-HAL and DEEP-HAL+ODF on (Left) YUP++ and (Right) HMDB-51.

