Dataset clip frame distribution before sub-sampling, with red dotted line highlighting sampled 30 frames mark.

Error Bar Plot of the Mean and Standard Deviation (not standard error) of frames in clips ordered by class. This shows the significant variance between different clips.

State Farm Dataset [7]
Error Bar Plot of the Mean and Standard Deviation (not standard error) of frames in clips ordered by class. This shows the significant variance between different clips.

Dataset clip frame distribution before sub-sampling, with red dotted line highlighting sampled 30 frames mark.
B5 + Pose + Object + FC2

B5 + Pose + Object + FC2 Max Pooling

State Farm Dataset [7]
Transfer learning – state-of-the-art object detector (Faster R-CNN with Inception ResNet-V2 [17]) on the target dataset before filtering (frequency of 80 different objects).
Transfer learning – state-of-the-art object detector (Faster R-CNN with Inception ResNet-V2 [17]) on the target dataset after filtering (objects of interest).
B5 + Pose + Object

FC2 + Pose + Object

FC2 + B5 + Pose

FC2 + B5 + Object

State Farm Dataset [7]
B5 + Pose + Object + FC2

Distracted Drivers Dataset [1]
Progression (M-LSTM memory) example when adding contextual features to the appearance features (VGG16 B5, FC2) over frames. Accuracy (ACC) and Average Precision (AP)

Distracted Drivers Dataset [1]