

Recently, there has been a lot of interest in building compact models for video classification which have a small memory footprint ( $<1$  GB). While these models a