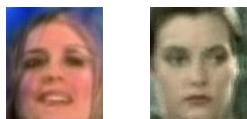


Aggregated Images



Aggregation
Network
 G



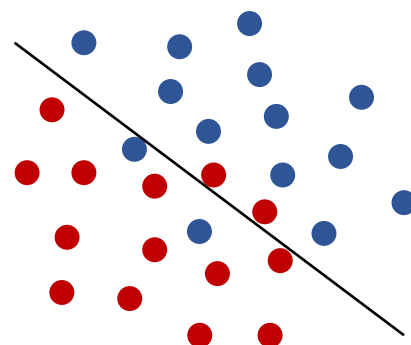
Input Videos

CNN F

CNN F

Image Space

Aggregated Features



Video Features

Feature Space