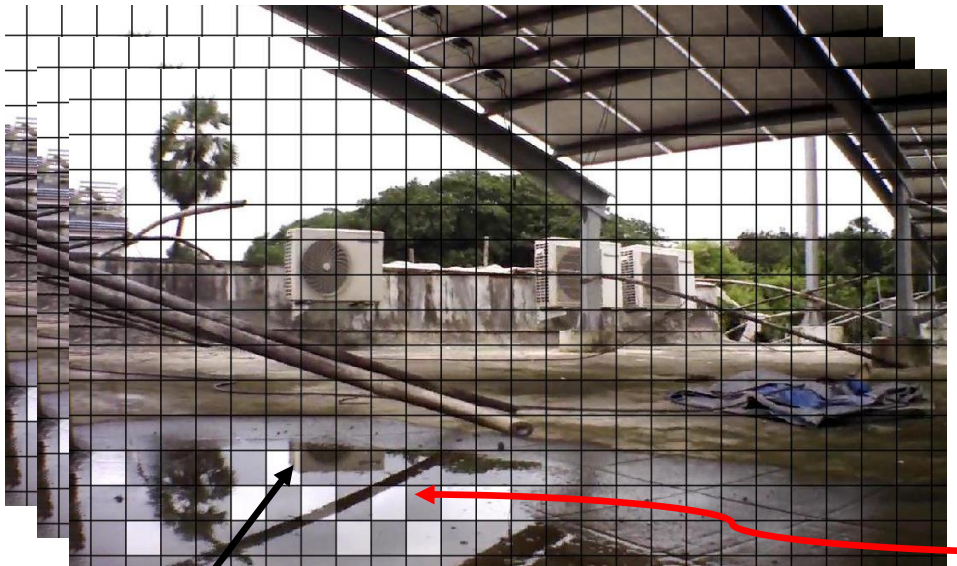
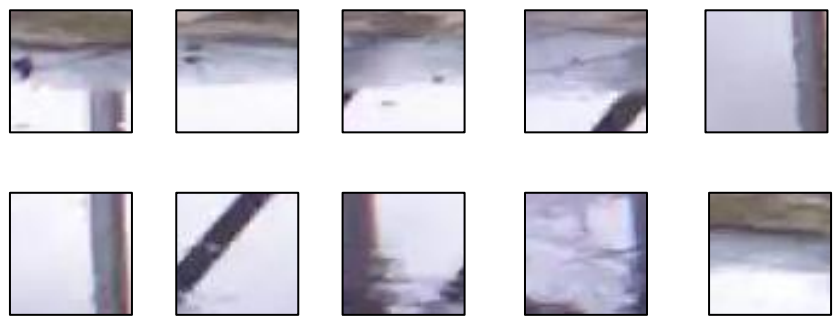
Input Frames from video



Area marked up by user

Blocks from non-puddle region



Blocks from puddle region