

**Video as a 4D array:  $N$  frames  $\times$  (height, width, 3)**

