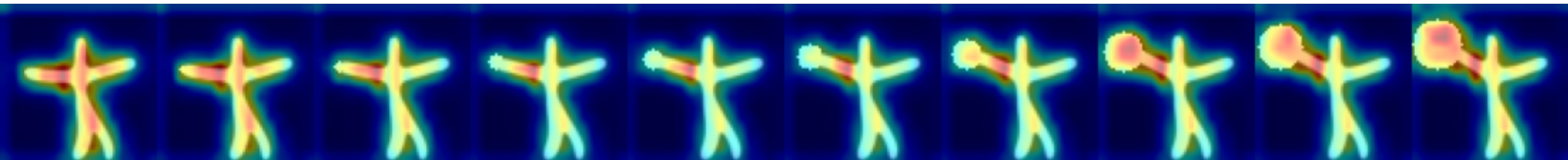
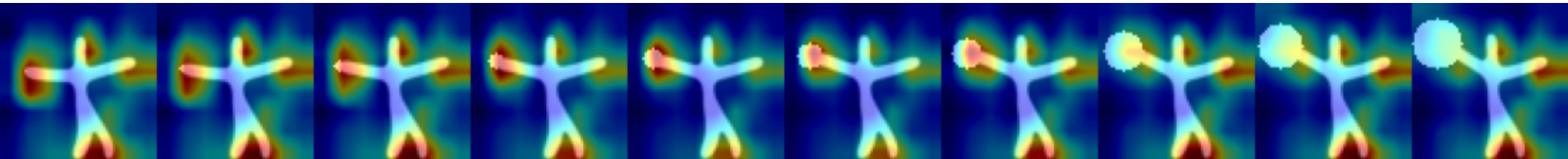


Feature map from layer1 semantic encoder (upscaled)  
mode: squeeze\_mean - feat\_size: torch.Size([256, 16, 16])



Feature map from layer2 semantic encoder (upscaled)  
mode: squeeze\_mean - feat\_size: torch.Size([512, 8, 8])



Feature map from layer3 semantic encoder (upscaled)  
mode: squeeze\_mean - feat\_size: torch.Size([1024, 4, 4])

