

A dominant paradigm for learning-based approaches in computer vision is training generic models, such as ResNet for image recognition, or I3D for video under