

usual
[9,13]

Videos

+

training
labels



train model



teenager shows
tricks in a park ...



church wedding
with orchestra ...



marching band
on the road ...



person builds a
wooden guitar ...



"play instrument"
model

**our
setting**

Videos

+

natural lang.
descriptions



identify
training labels



train model