

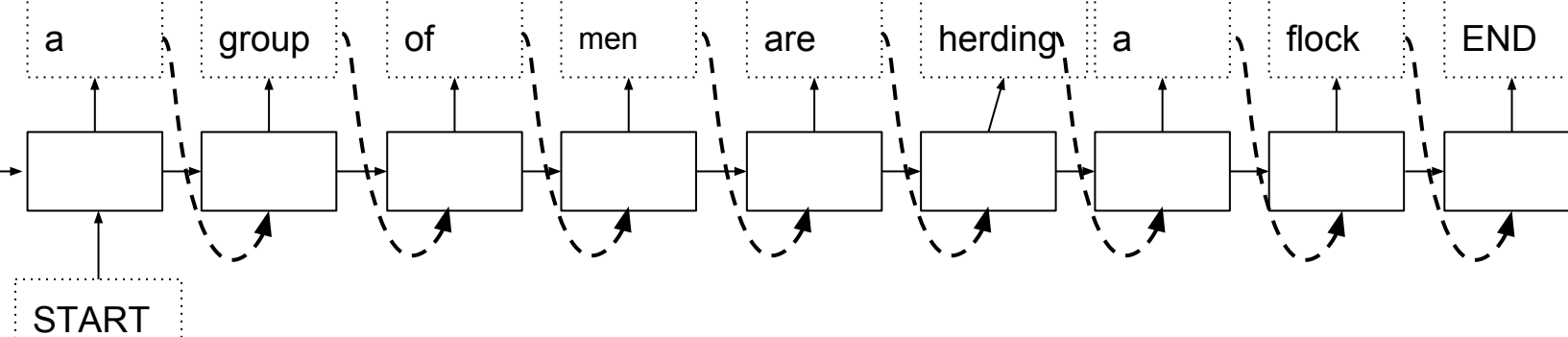
Input image

Resnet-50 pre-trained
on Imagenet 1000
classes

CNN-Encoder

Metadata:
Year: 1932
Location: Madison
County, Montana

Metadata-encoder
(see details in paper)



RNN-Decoder
(LSTM)