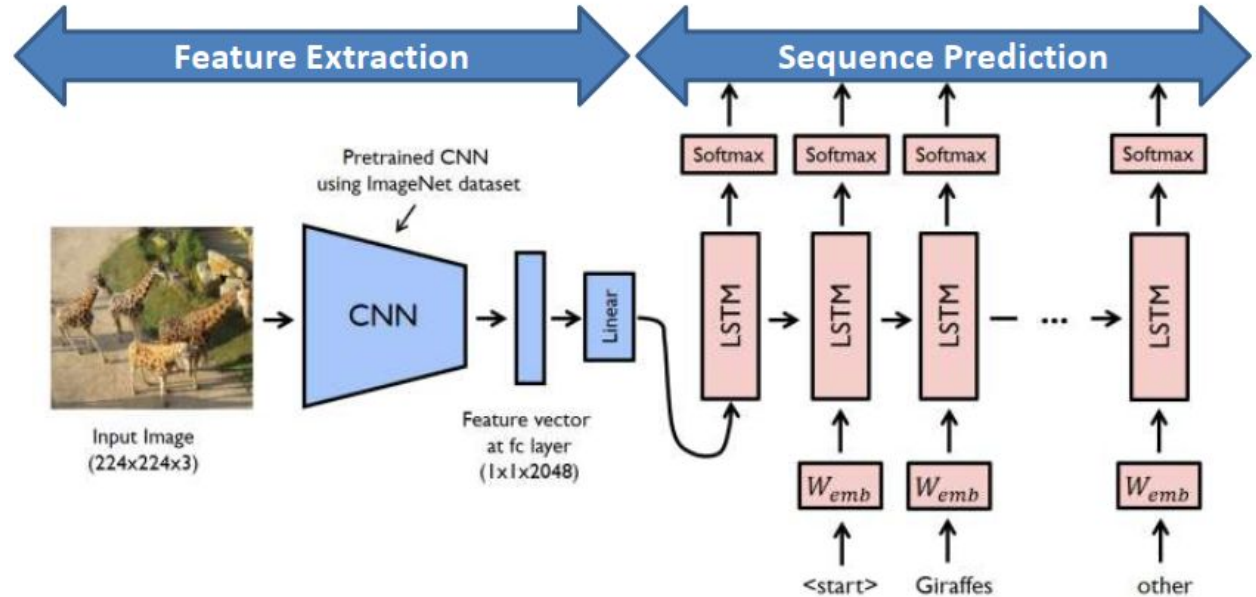


LSTM Variants

Nibir Mandal

CNN-LSTM

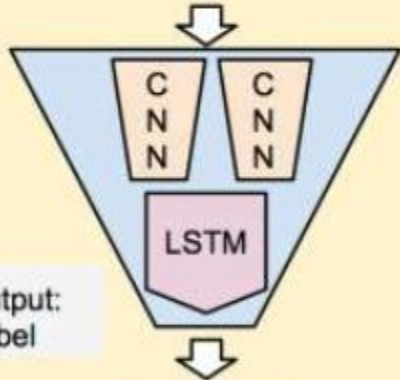
- Designed for sequence prediction problems with spatial inputs like images or videos



Application

Activity Recognition

Input:
Sequence
of Frames

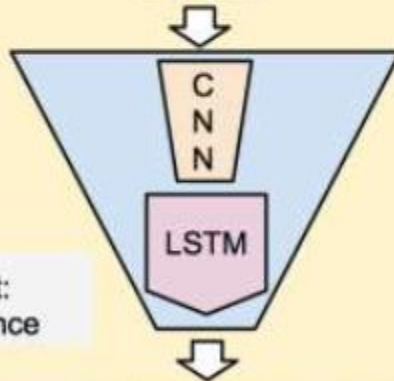


Output:
Label



Image Description

Input:
Image

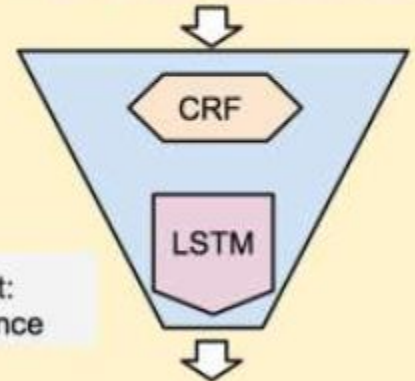


Output:
Sentence

A large building with a
clock on the front of it

Video Description

Input:
Video

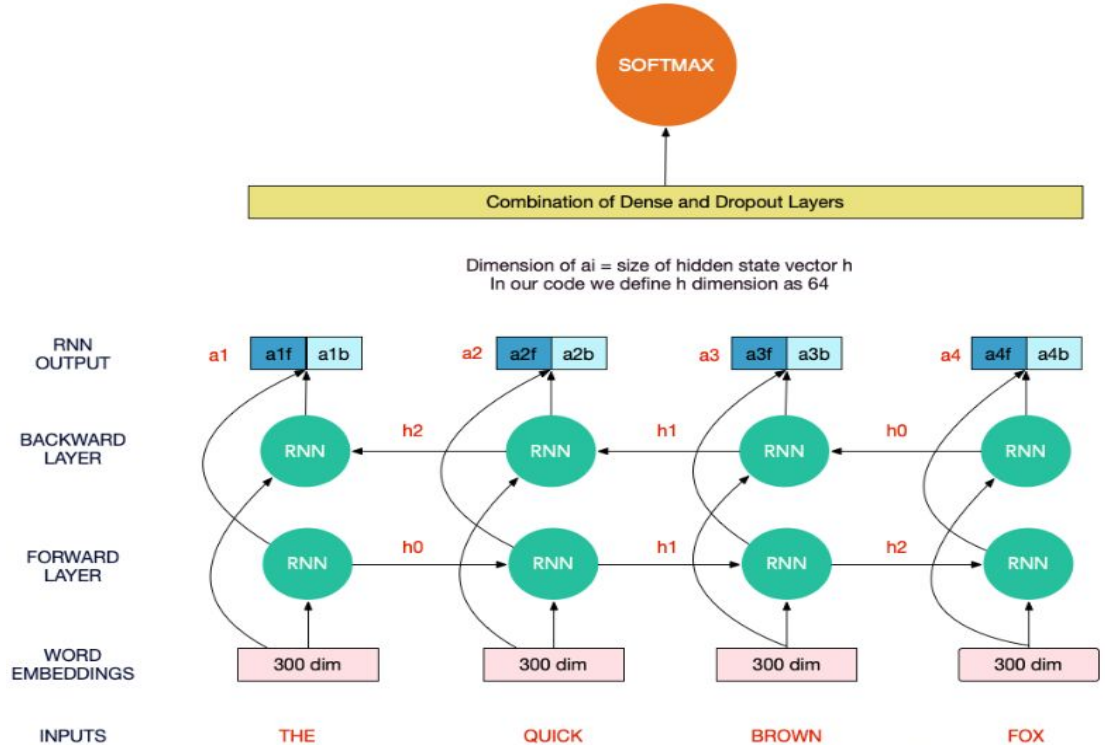


Output:
Sentence

A man juiced the orange

Bidirectional LSTM

- Forward & Backward training
- Capture complex contexts and dependencies



Bidirectional RNN has forward and backward RNNs. Source: MLWhiz 2018.* [📄](#)

Thank You!

Questions?