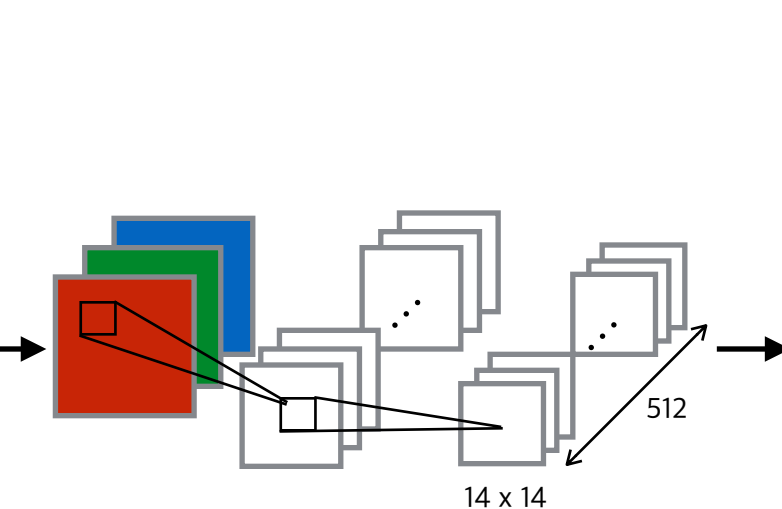
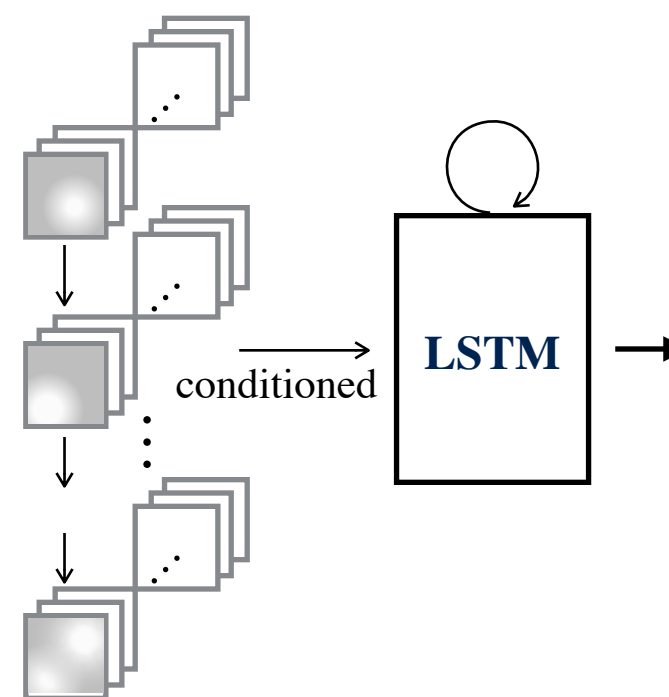


1. Input Images



2. Feature Extraction
(Oxford VGGnet pretrained)



3. Conditional LSTM
with Attention

That 's
very
nice
of
you

4. Word by Word
Generation