Attention - based - Models :

→ why do we need them ?

→ Intuition & Architecture

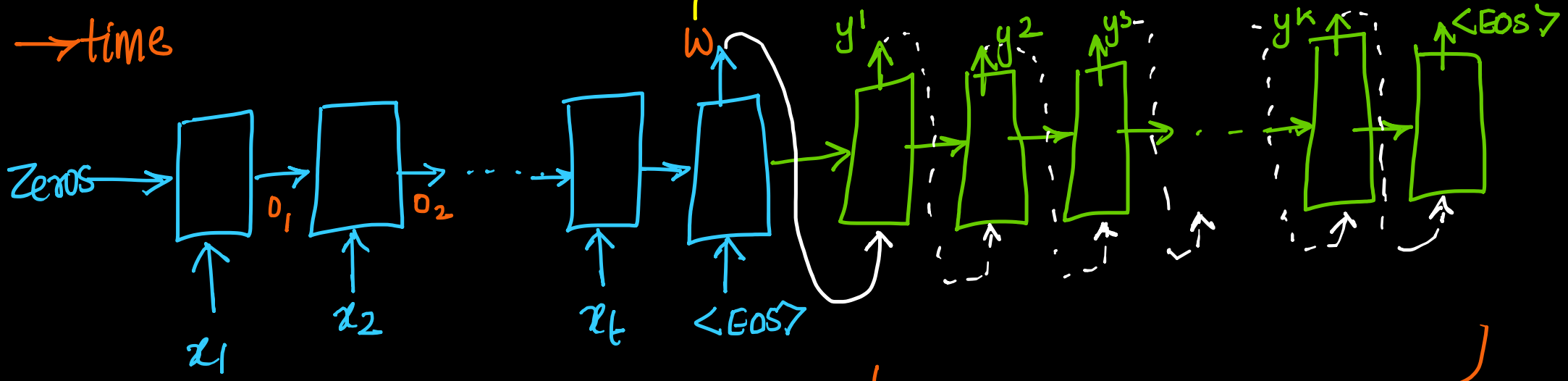
→ Attention on images

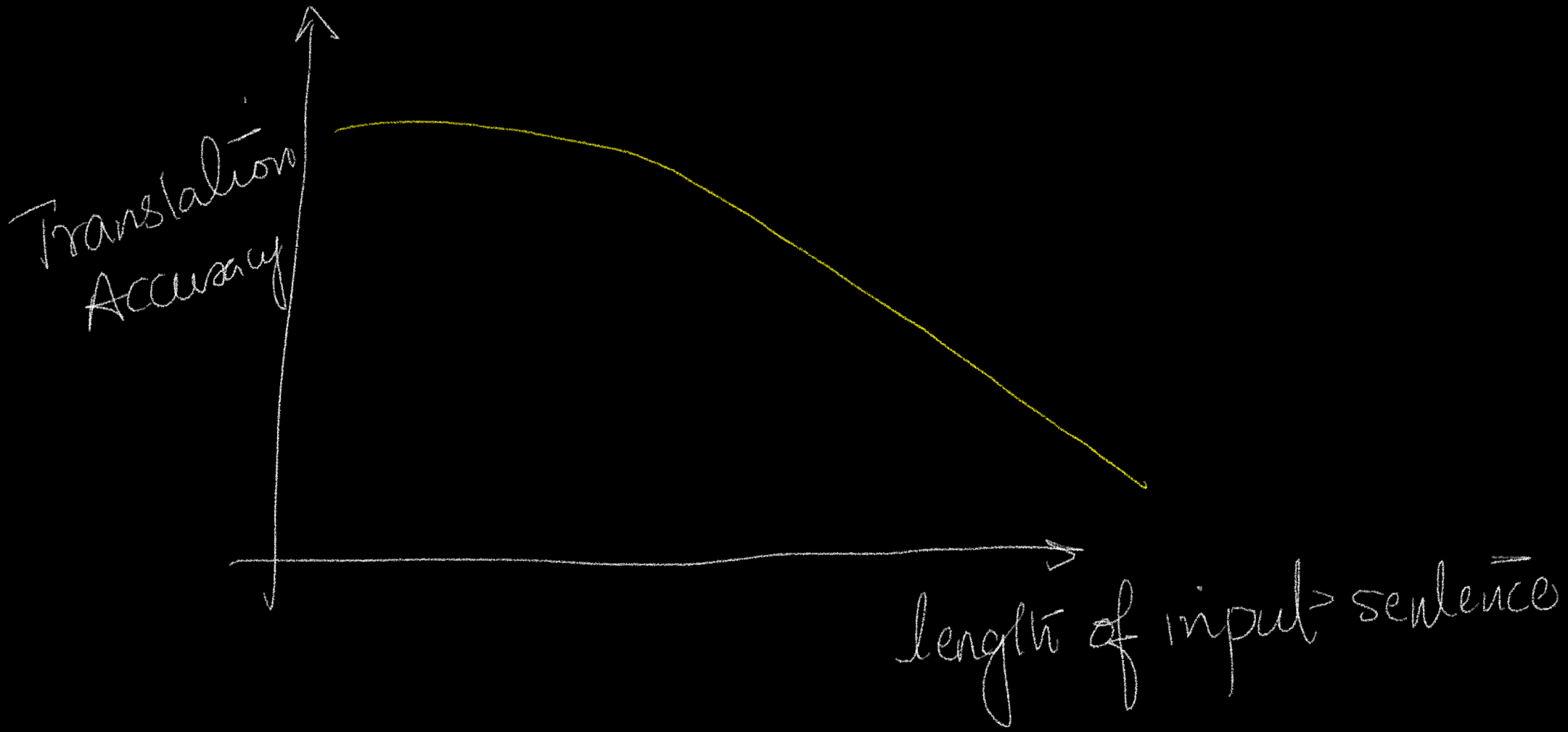
→ one-line code implementation .

Problems with simple seq2seq model: (previous video)

→ lengthy-selences

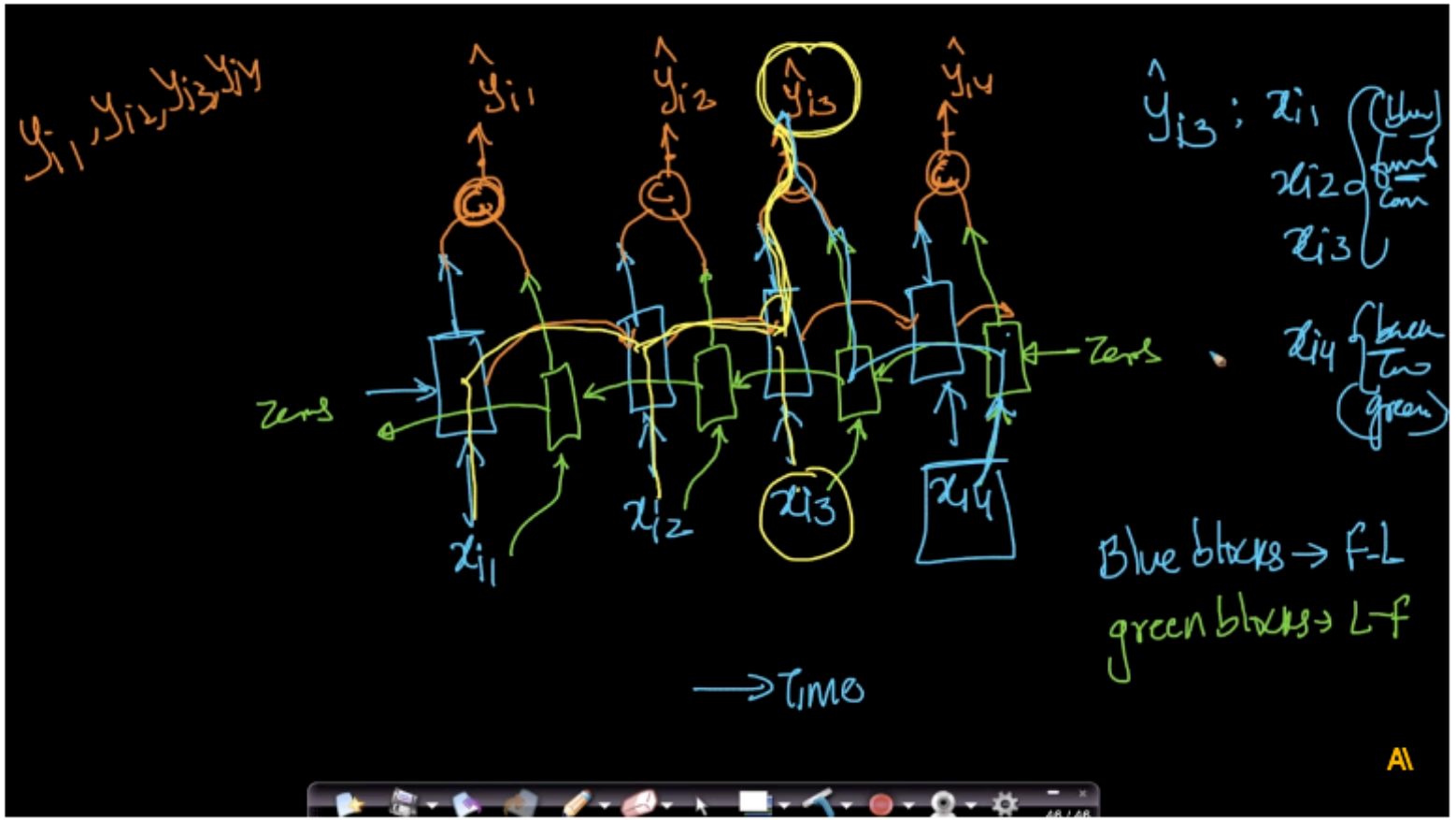
Not human-like
cannot capture the essence of a.

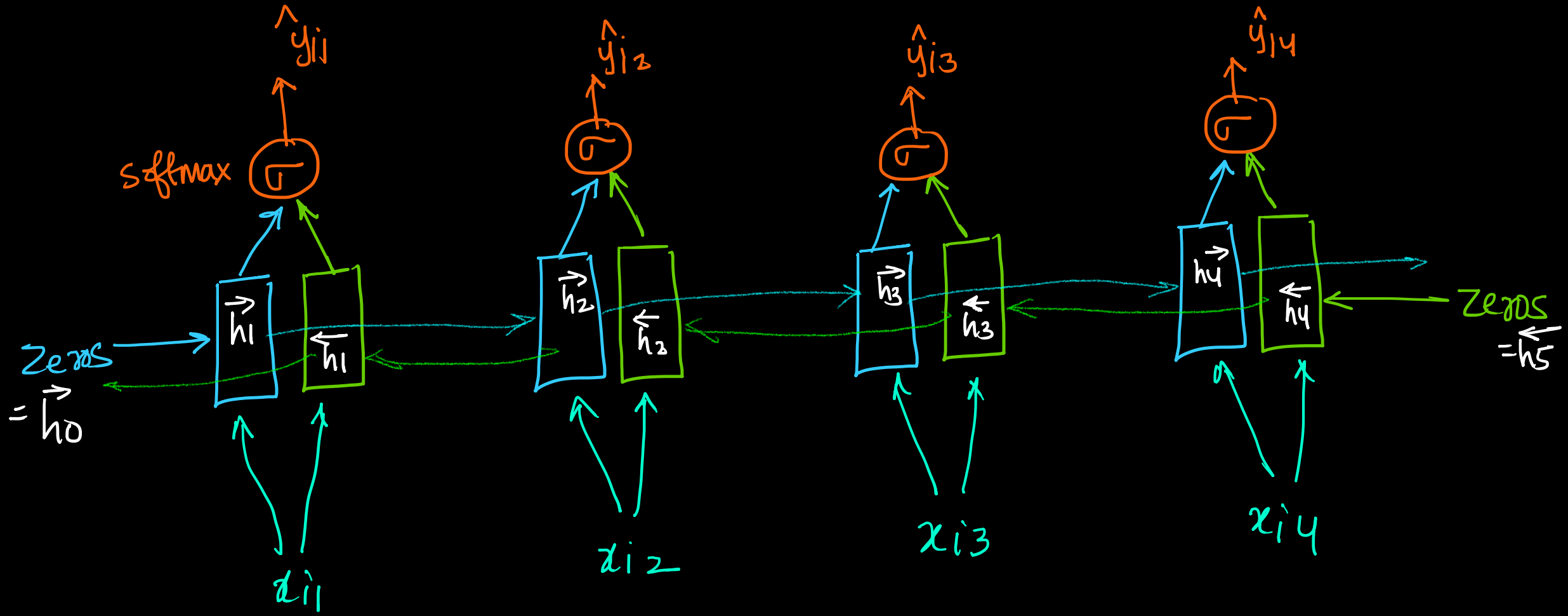




Bidirectional RNN.

Instructor: Applied AI Course **Duration:** 12 mins Full Screen





x_i : i^{th} input sentence

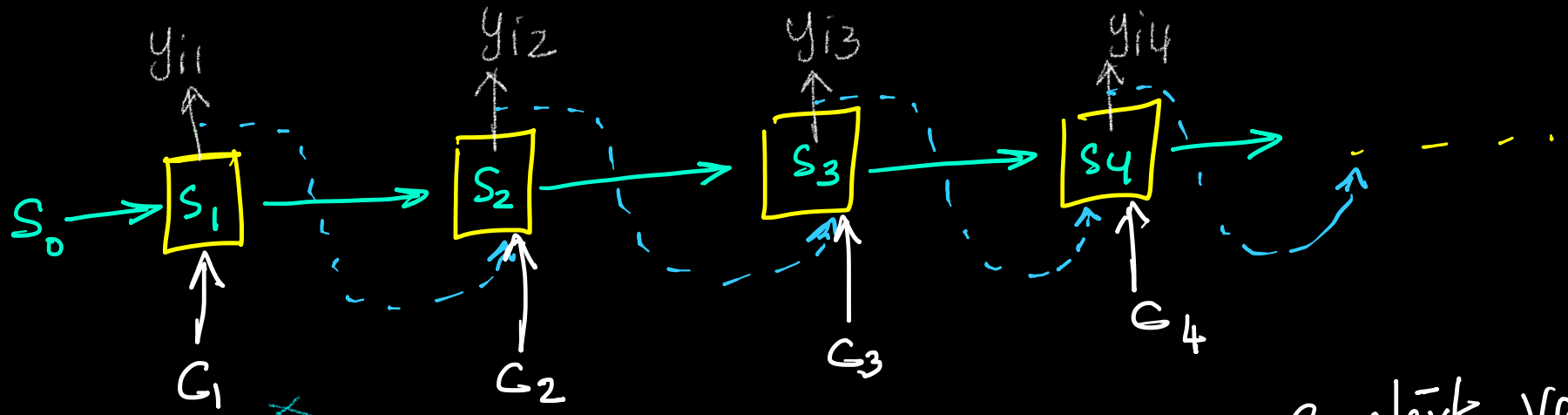
y_i : i^{th} output sentence

$$\mathcal{D} = \{x_i, y_i\}_{i=1}^{n}$$

$$h_t = \langle \vec{h}_t, \overleftarrow{h}_t \rangle$$

NEURAL MACHINE TRANSLATION BY JOINTLY LEARNING TO ALIGN AND TRANSLATE https://arxiv.org/pdf/1409.0473.pdf

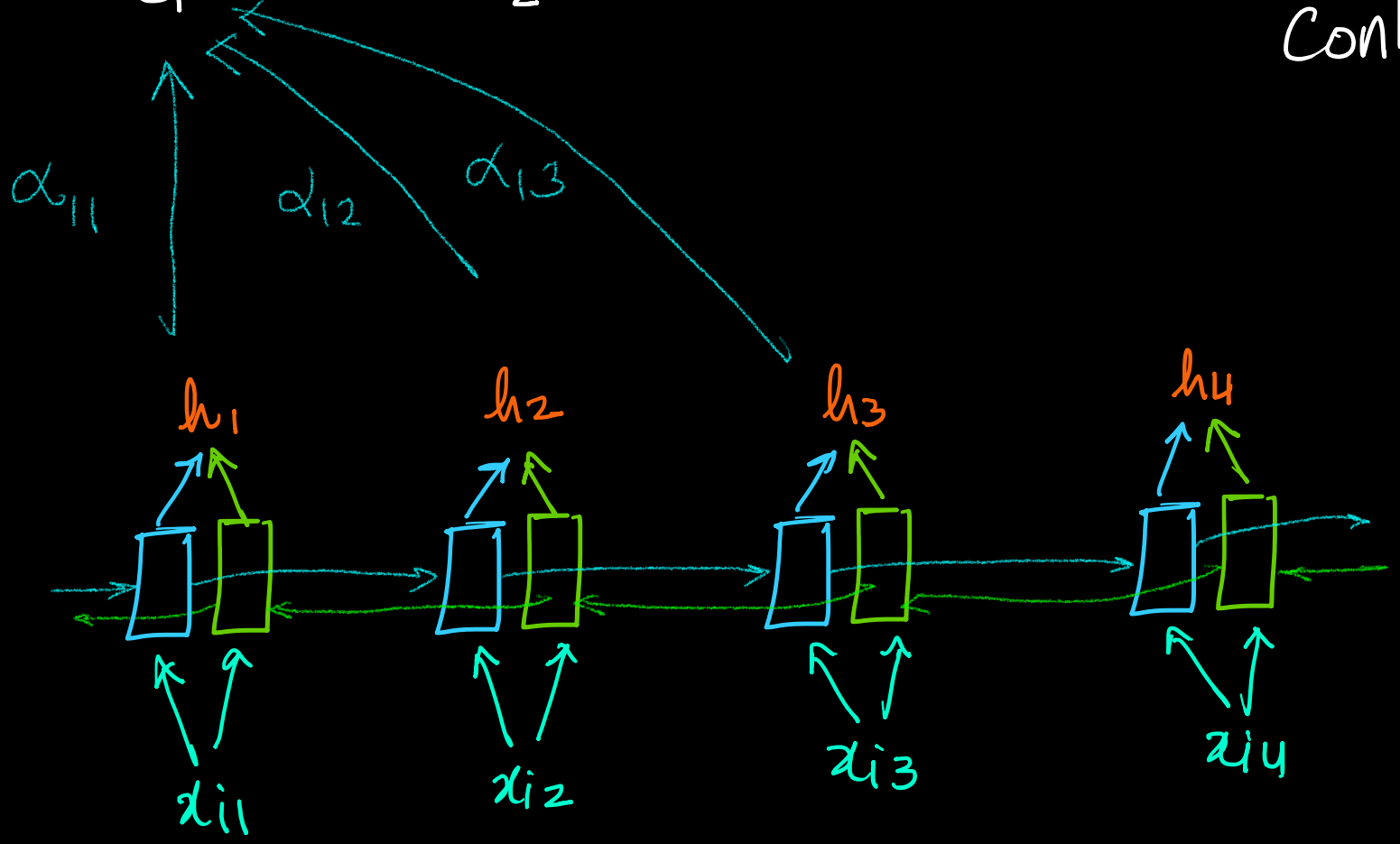
Dzmitry Bahdanau, KyungHyun Cho, Yoshua Bengio, ICLR 2015



decoder

$$\text{Context-vec} = C_i = \sum_{i=1}^{T_x} h_i \alpha_{ii}$$

$$\sum_{i=1}^{T_x} \alpha_{ii} = 1 \quad \& \quad \alpha_{ii} \geq 0$$



encoder

How to compute α_{ij}'s ?

$$c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j$$

(eqn 5 in the research paper)

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})}$$

≥ 0

$$\sum_j \alpha_{ij} = 1$$

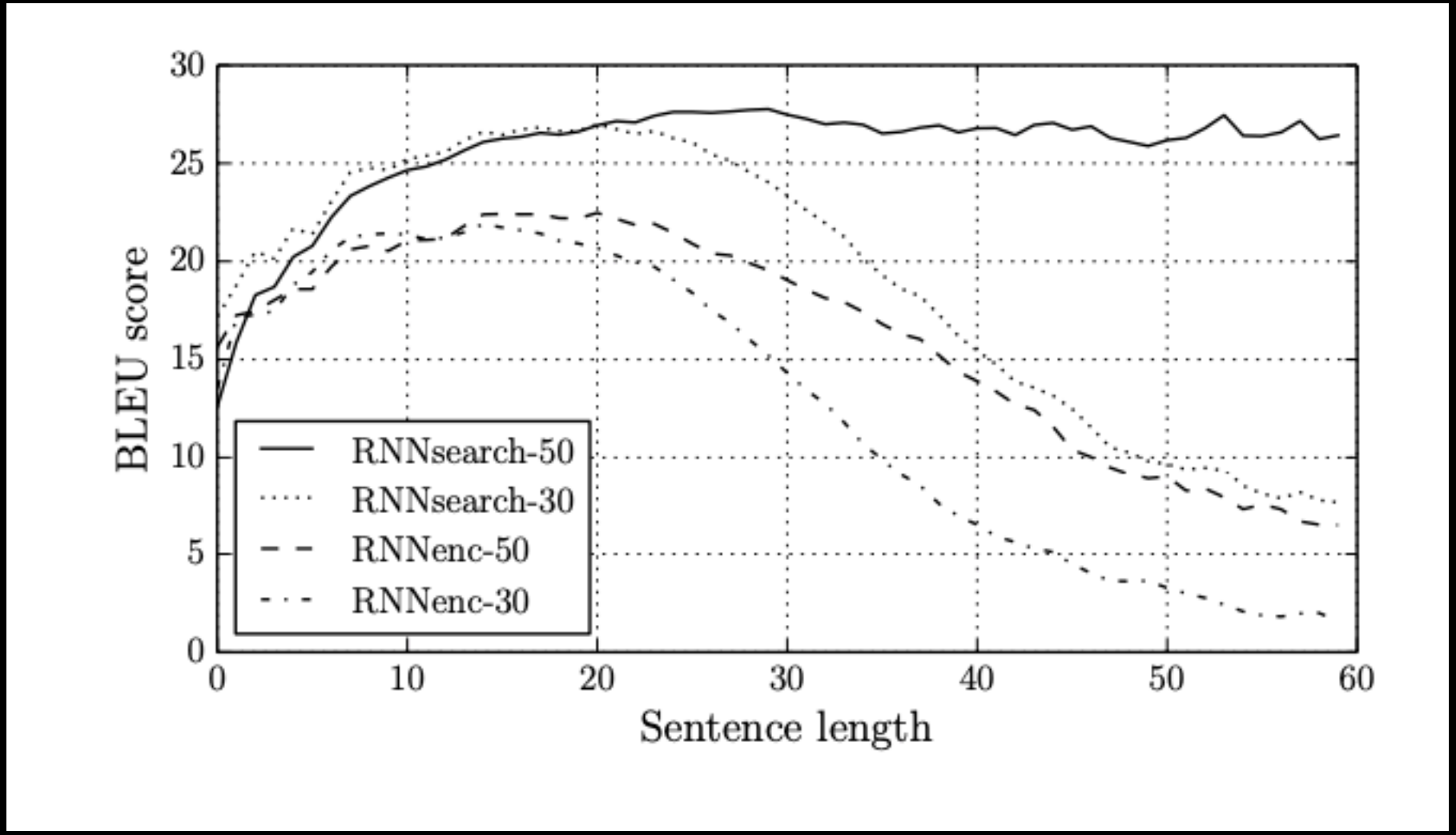
$$e_{ij} = a(s_{i-1}, h_j)$$

→ attention-model

↓

feed forward NN (small)

Train using BPTT & let it converge.



Drawback:

$$\text{Time-complx} = O(K_1 \cdot K_2)$$

K_2 → length of output sentence

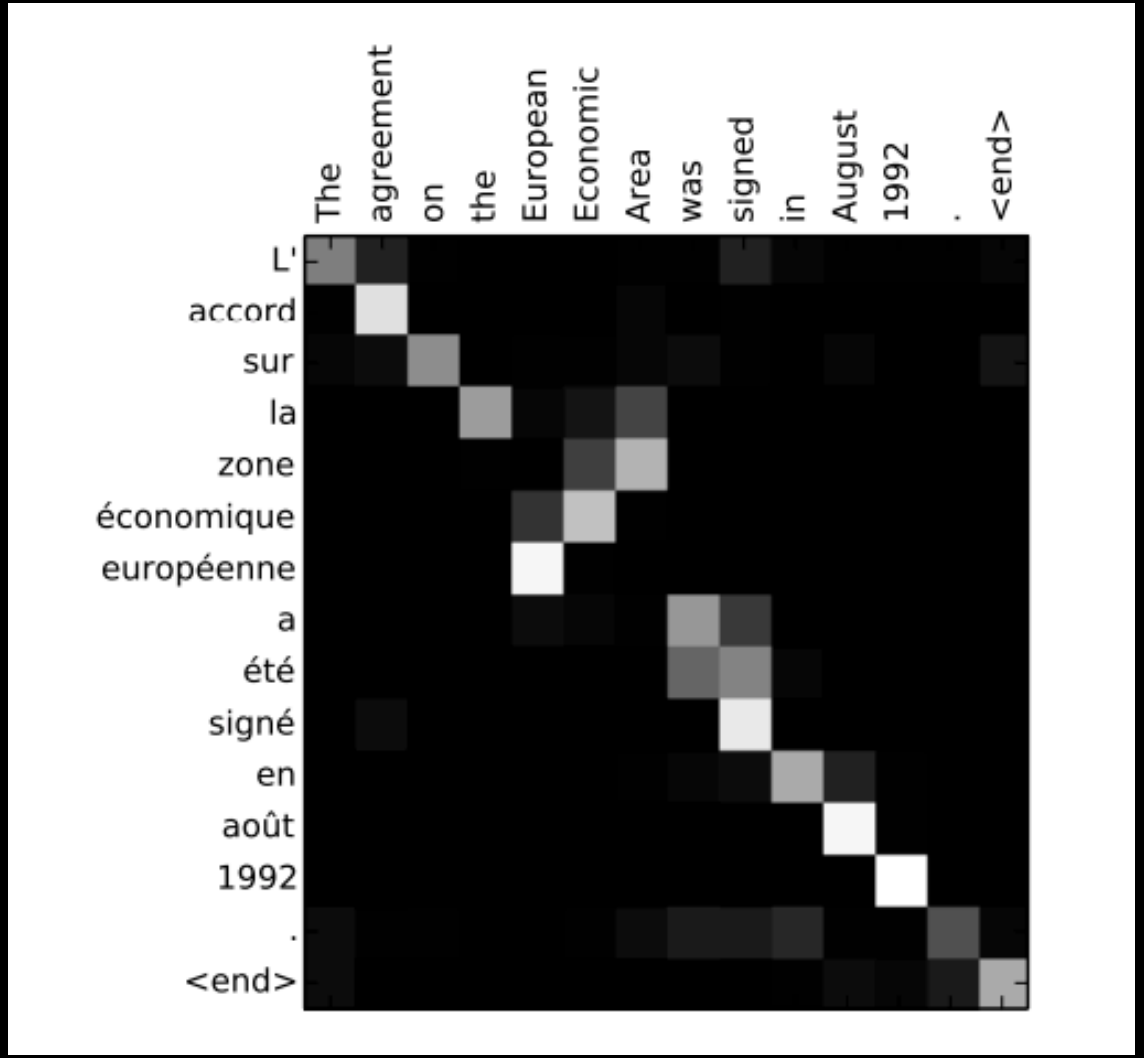
K_1 → length of input sentence

[Not a major drawback]

Visualize α_{ij}'s

making sense of what's happening

{Debugging}

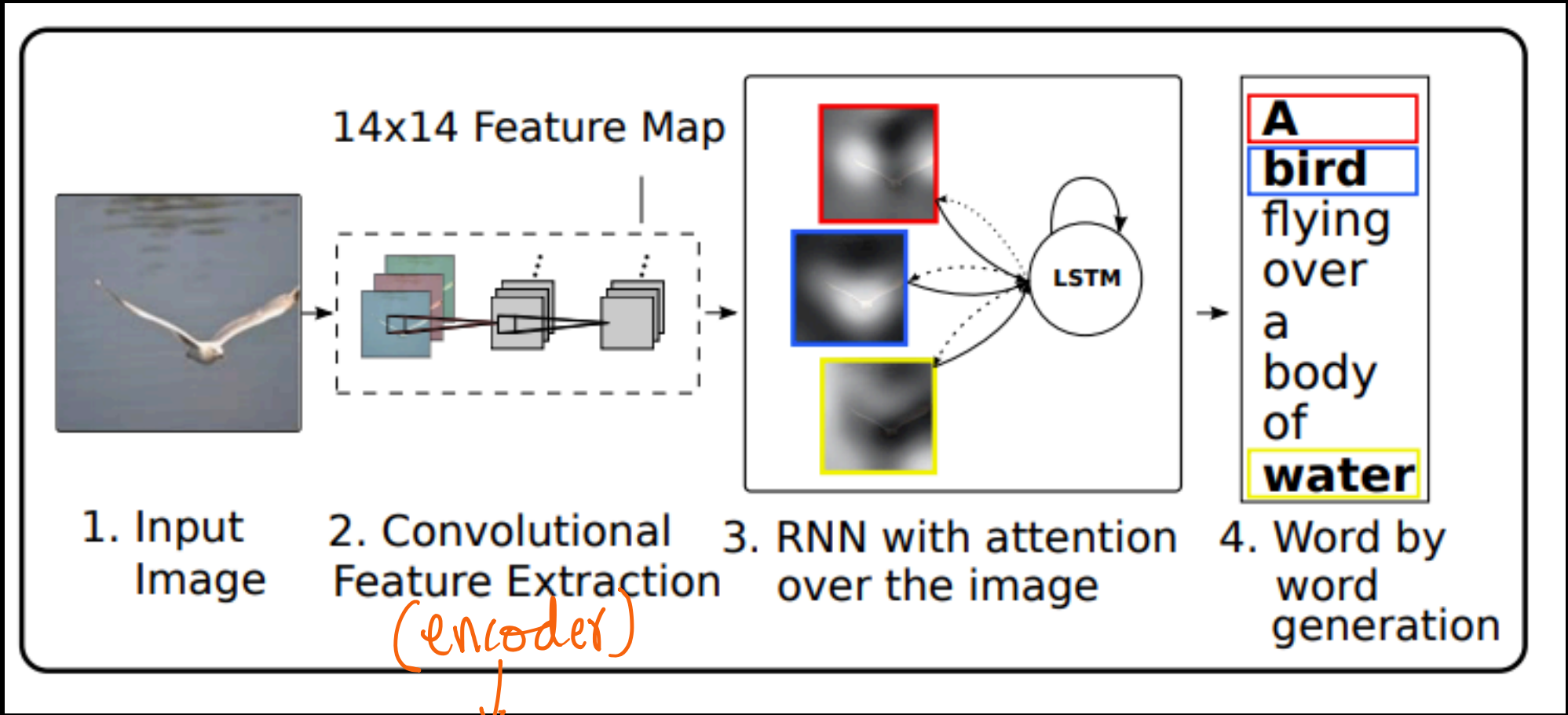


Show, Attend and Tell: Neural Image Caption Generation with Visual Attention

[2015]

Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhutdinov, Richard Zemel, Yoshua Bengio

https://arxiv.org/pdf/1502.03044.pdf



(encoder)

'L' d-dim vectors for each image-region

Visualizing dij's:

A woman is throwing a frisbee in a park.

A dog is standing on a hardwood floor.

A stop sign is on a road with a mountain in the background.

A little girl sitting on a bed with a teddy bear.

A group of people sitting on a boat in the water.

A giraffe standing in a forest with trees in the background.

One-line implementation:

https://www.tensorflow.org/api_docs/python/tf/keras/layers/Attention

https://towardsdatascience.com/light-on-math-ml-attention-with-keras-dc8dbc1fad39