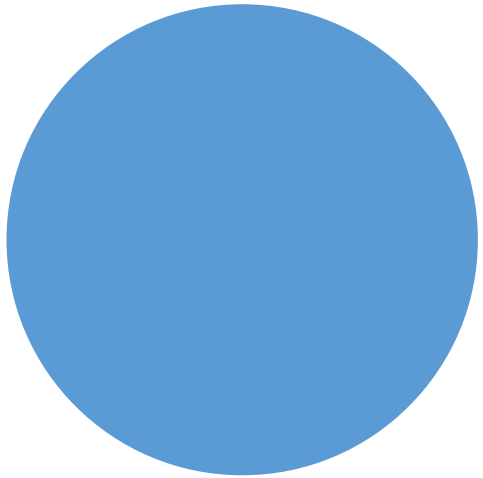
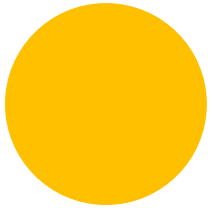



Attention Please!

Attention Mechanisms in Neural Networks

Patrick Michl
www.frootlab.org

PyData Heidelberg #4
2019-11-21

Milestones in Attention Research

Visual Attention in Human Perception

Visual attention in human perception is based on the dynamics between **Recognition** and **Selection**

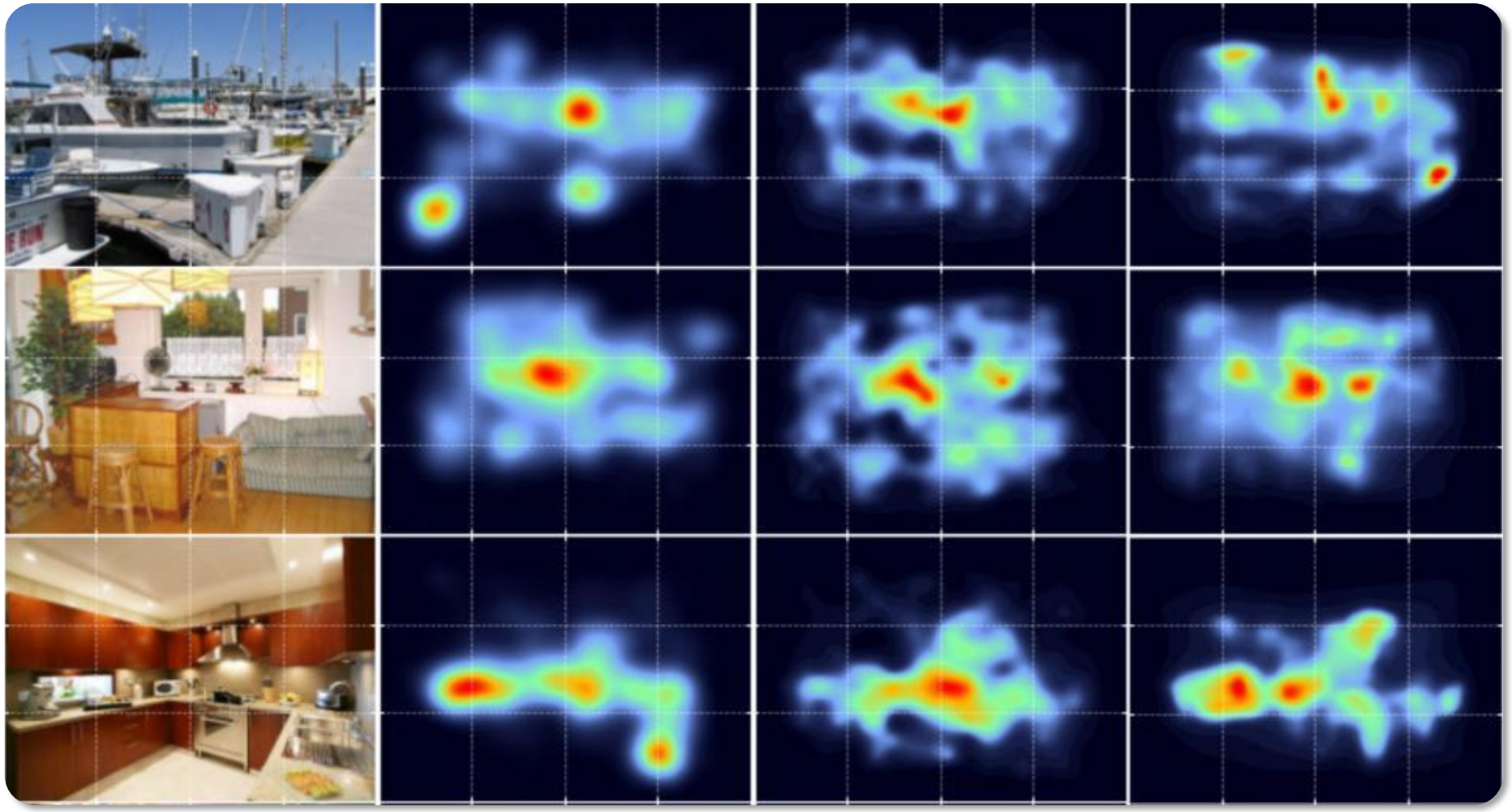
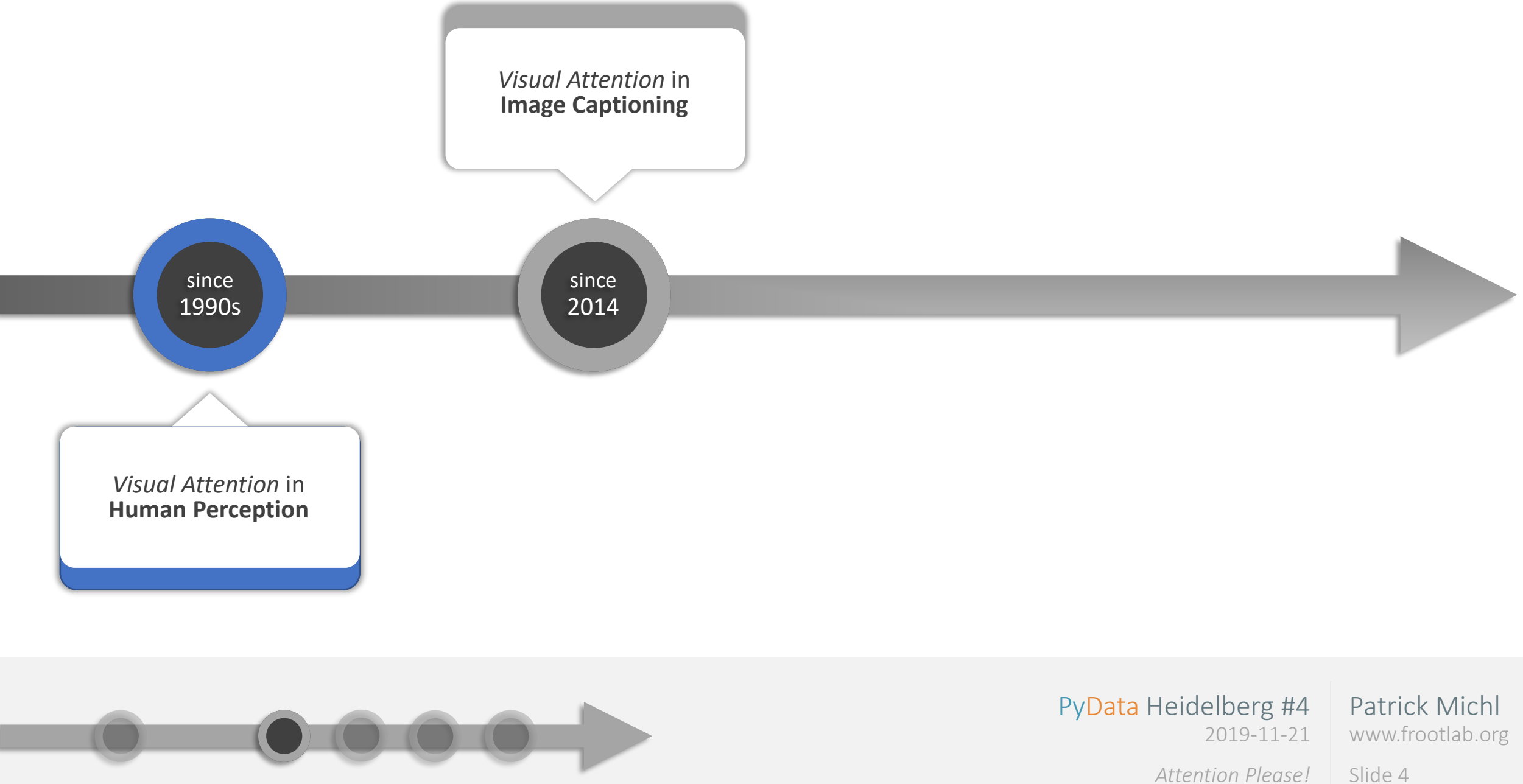


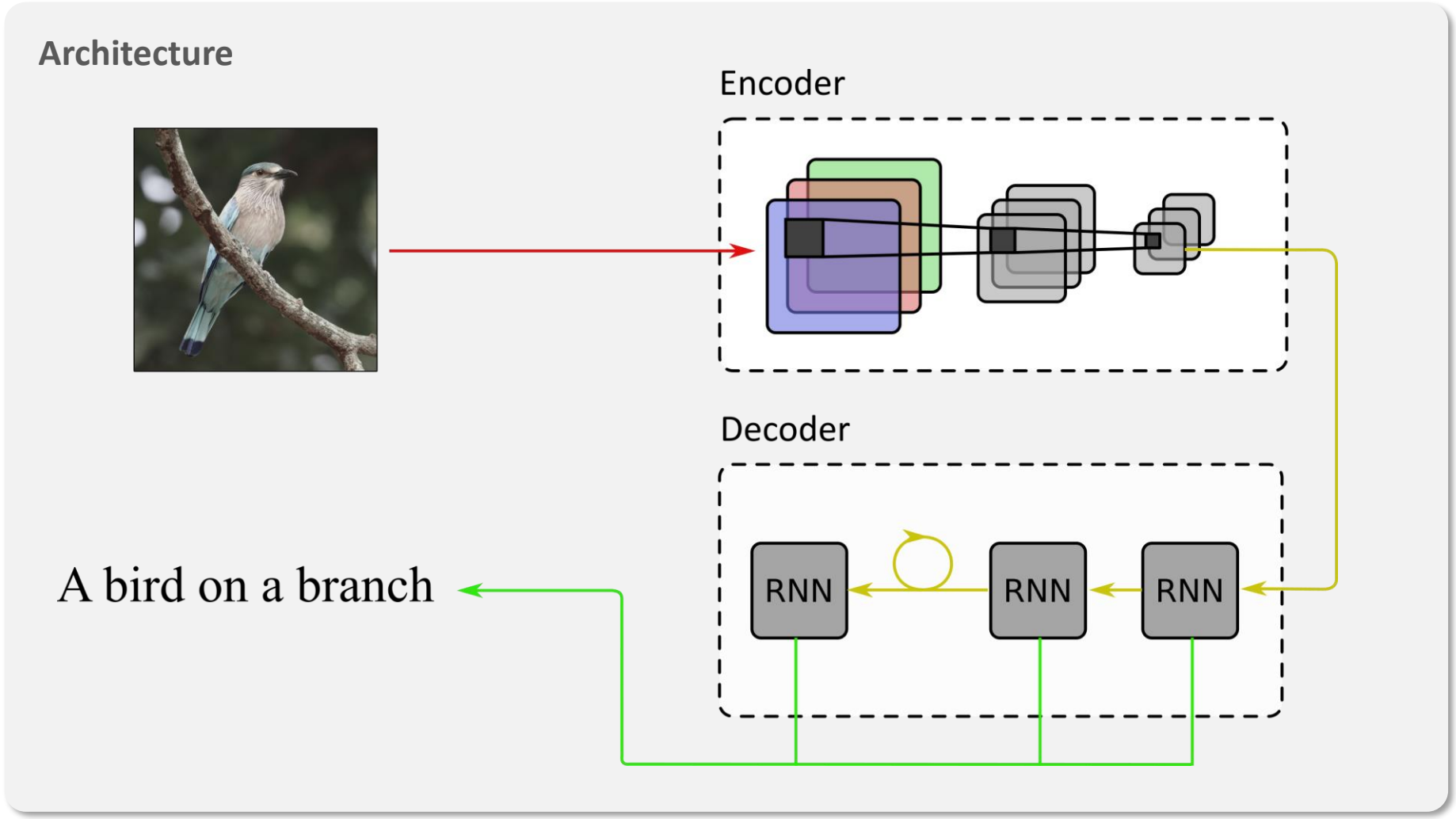
Image source*: John Henderson* and *Taylor Hayes*, UC Davis


Visual Attention in
Image Captioning
since
2014
since
1990s
Visual Attention in
Human Perception

Visual Attention in Image Captioning

Encoder: CNN

Decoder: RNN



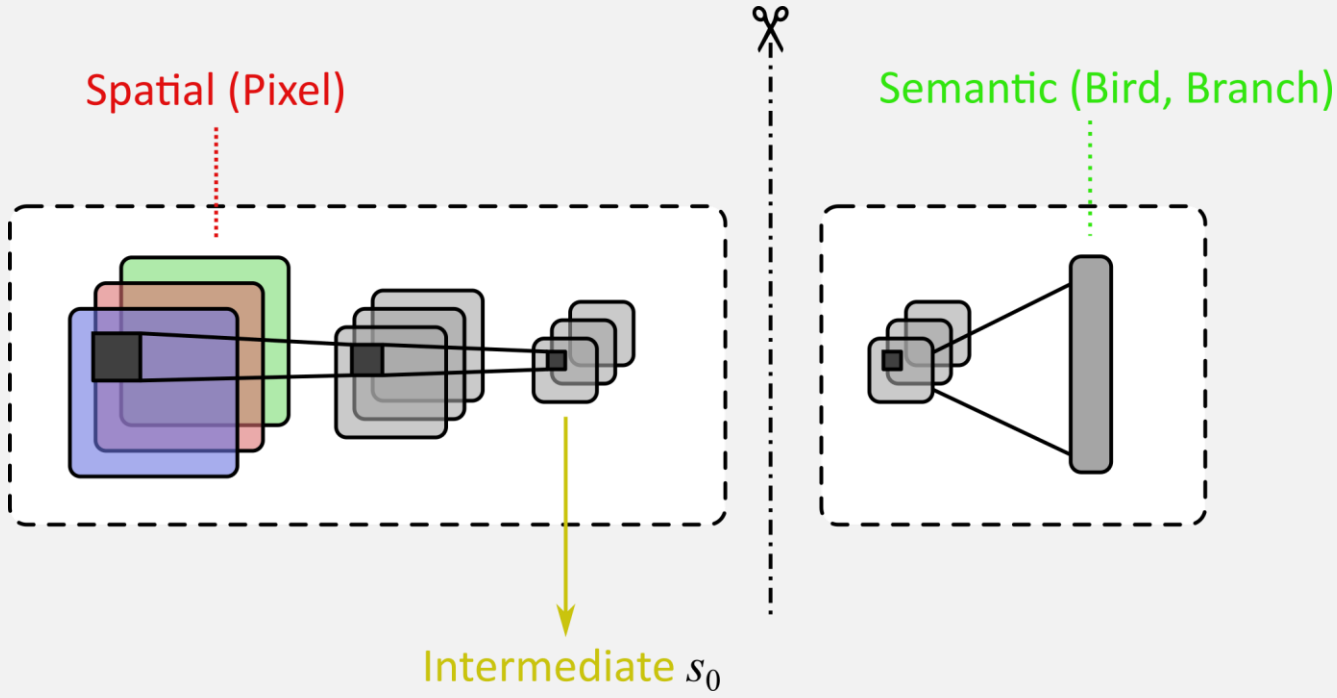
Visual Attention in Image Captioning

Encoder: CNN

Convolutional features entangle spatial with semantic information

Decoder: RNN

Encoder



Visual Attention in Image Captioning

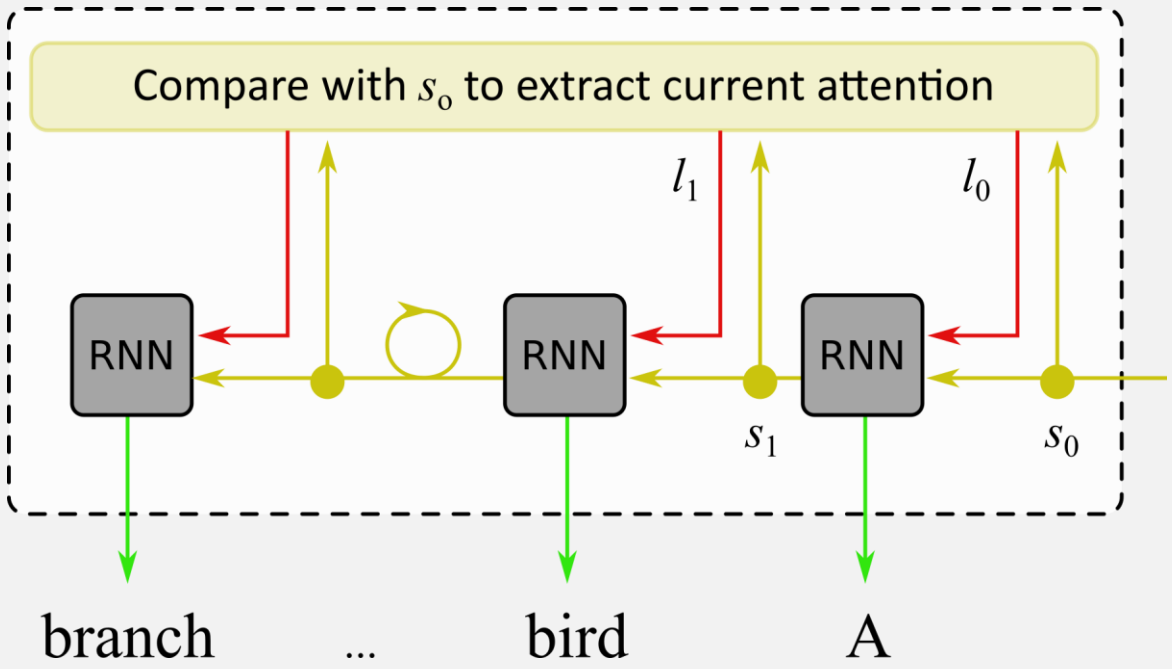
Encoder: CNN

Convolutional features
entangle spatial with semantic information

Decoder: RNN

Visual Attention
allows to recover spatial relations of current outputs

Decoder



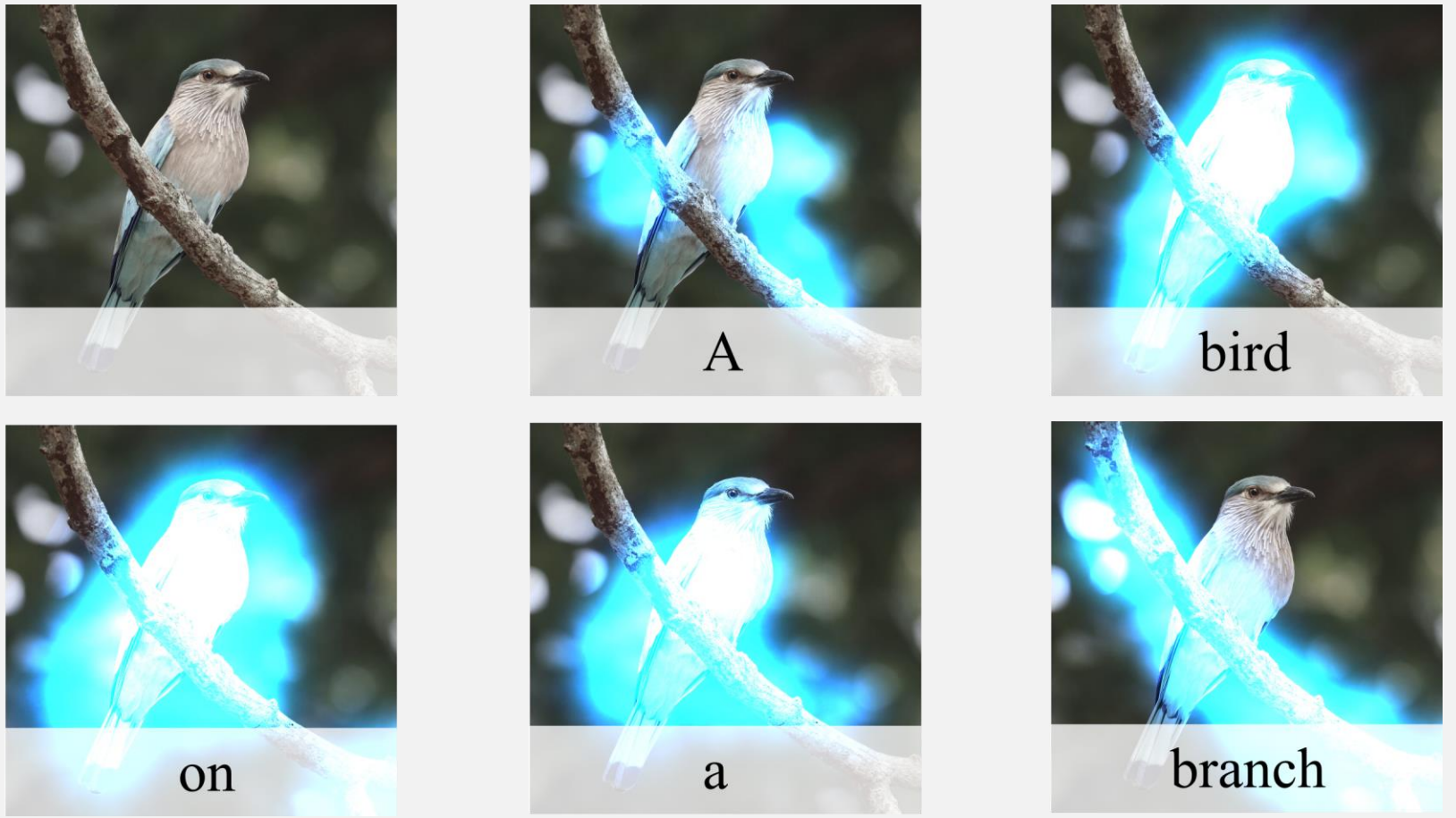
Visual Attention in Image Captioning

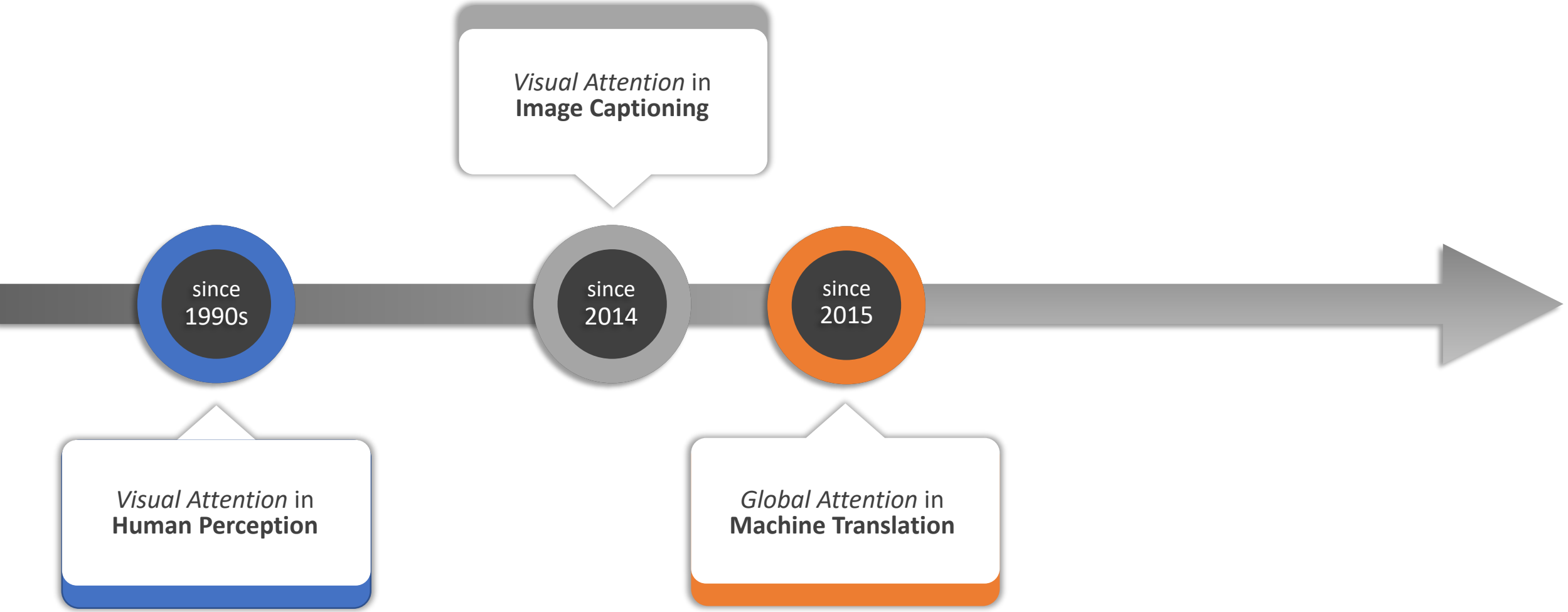
Encoder: CNN

Convolutional features
entangle spatial with
semantic information

Decoder: RNN

Visual Attention
allows to recover spatial
relations of current outputs



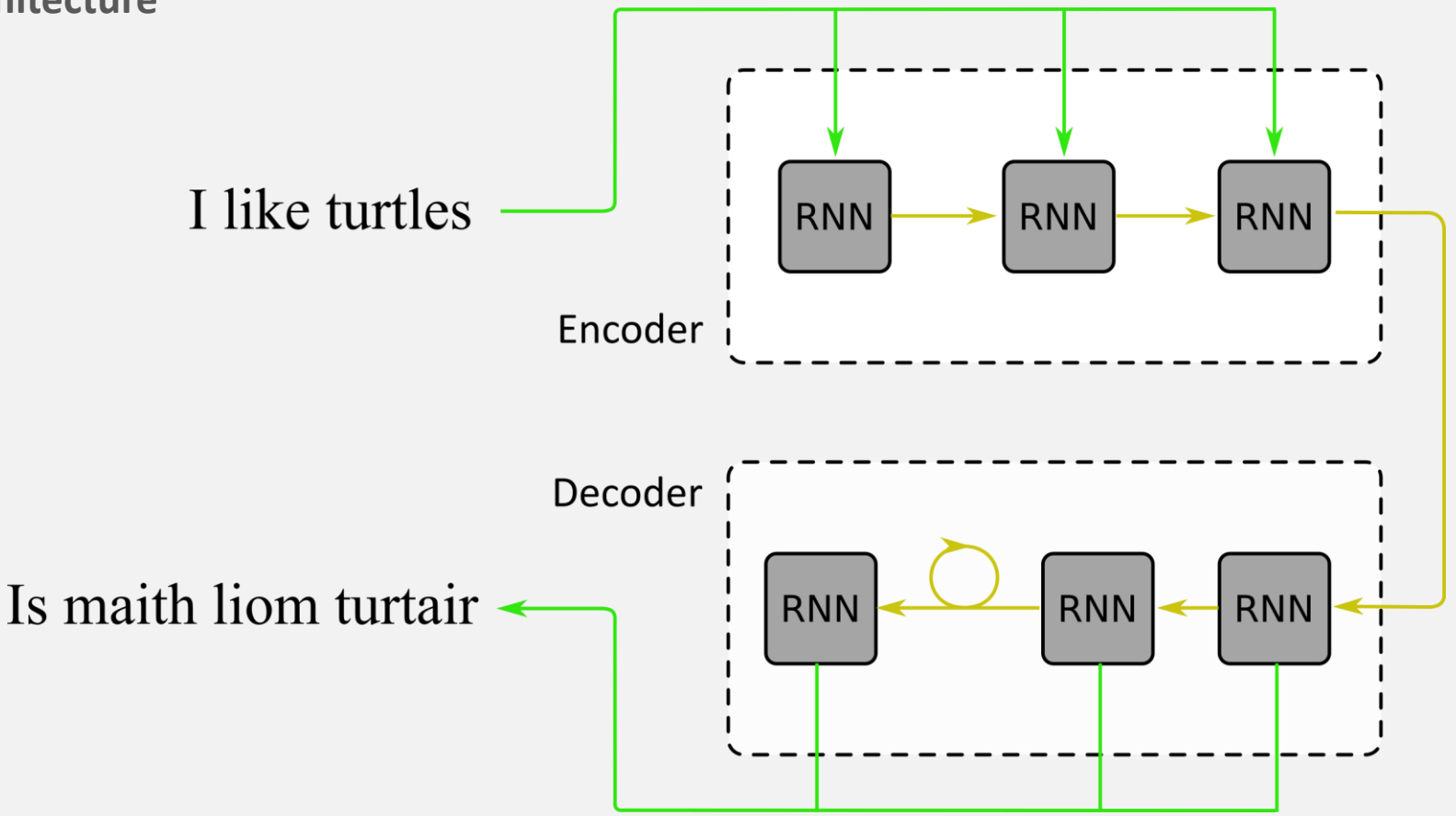

Visual Attention in
Image Captioning
since
1990s
since
2014
since
2015
Visual Attention in
Human Perception
Global Attention in
Machine Translation

Global Attention in Machine Translation

Encoder: RNN

Decoder: RNN

Architecture



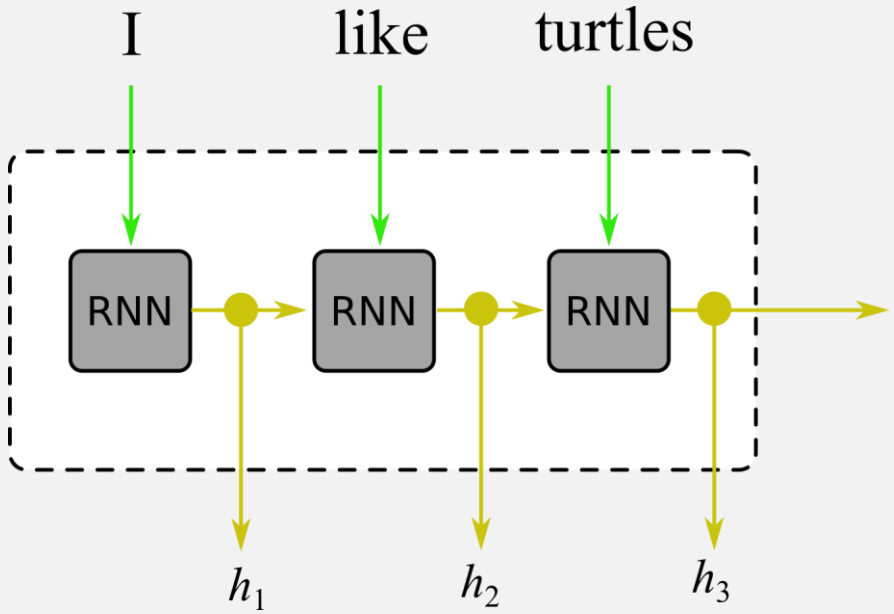
Global Attention in Machine Translation

Encoder: RNN

Sequence of hidden states encodes the context building stack

Decoder: RNN

Encoder



Global Attention in Machine Translation

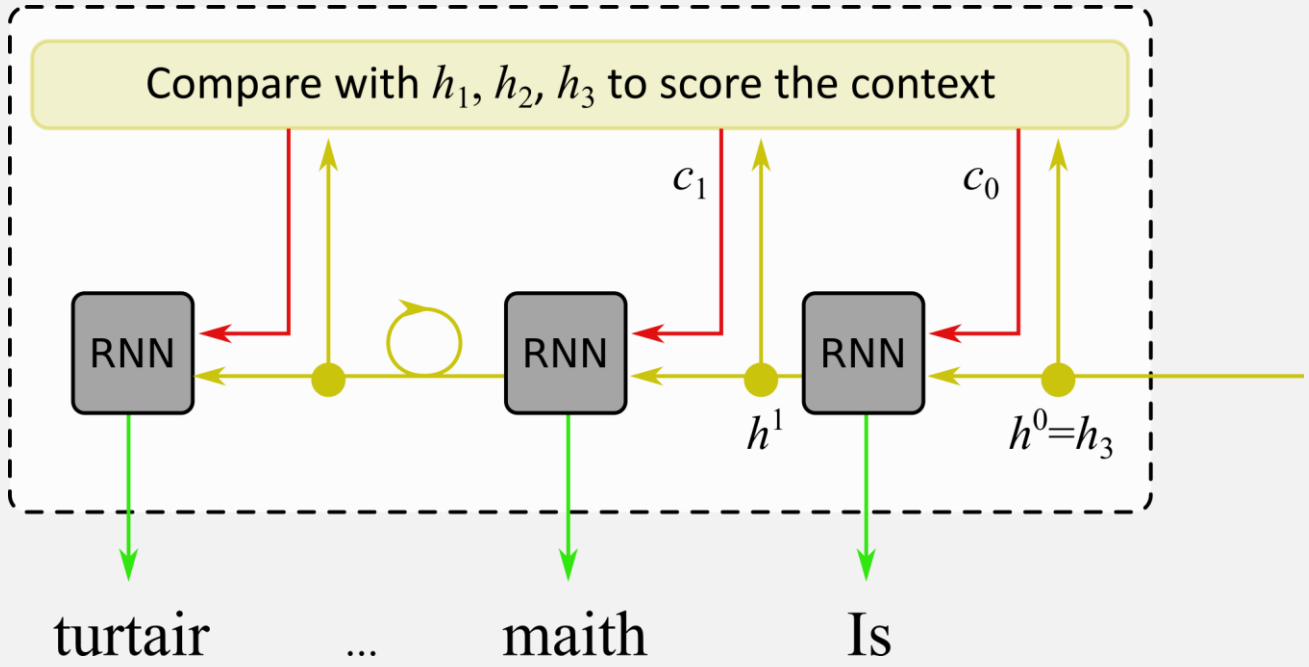
Encoder: RNN

Sequence of hidden states
encodes the context building stack

Decoder: RNN

Global Attention
allows to recover the context of current outputs

Decoder



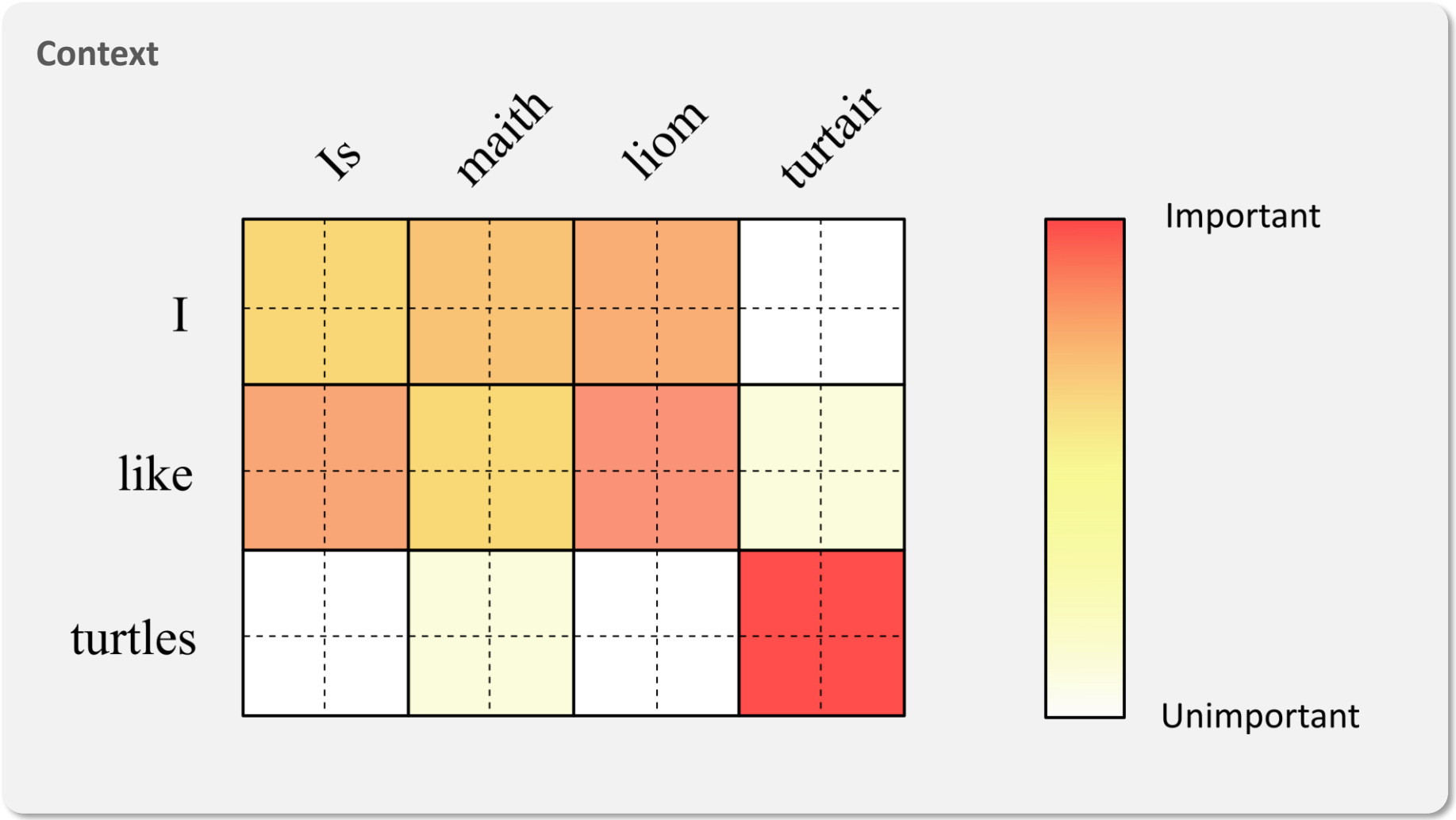
Global Attention in Machine Translation

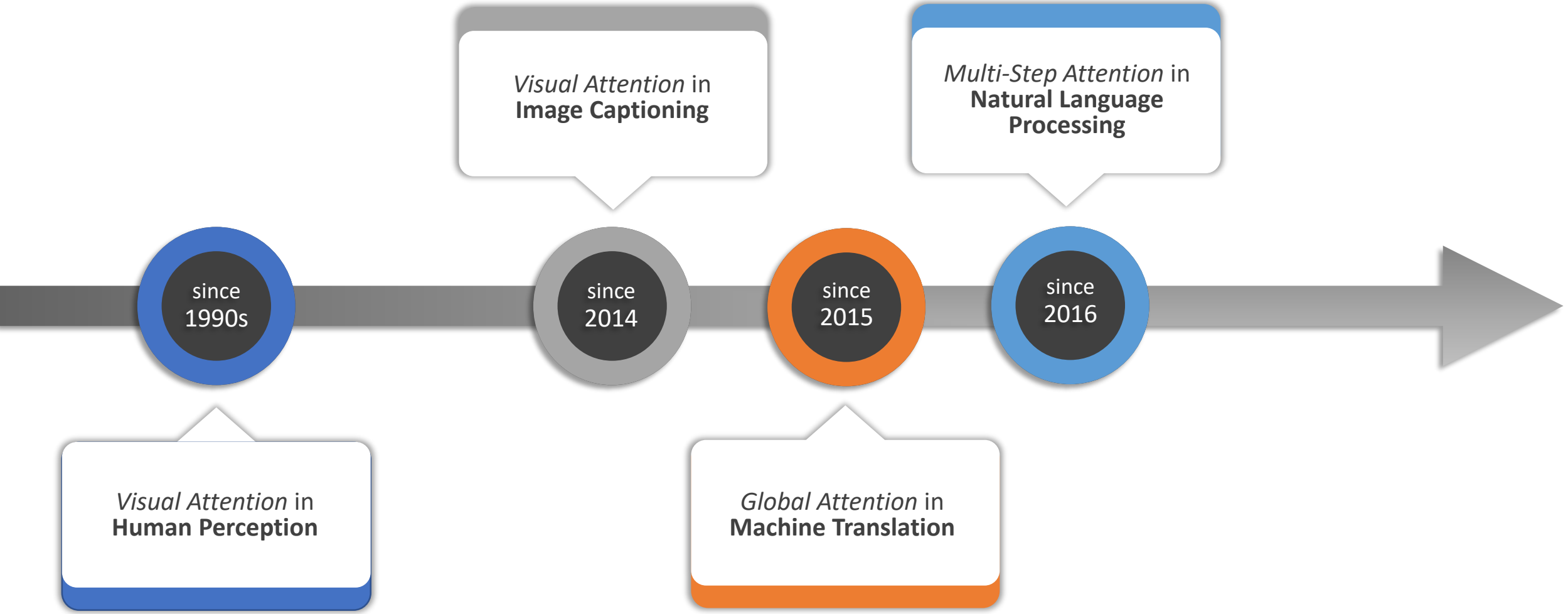
Encoder: RNN

Sequence of hidden states encodes the context building stack

Decoder: RNN

Global Attention allows to recover the context of current outputs



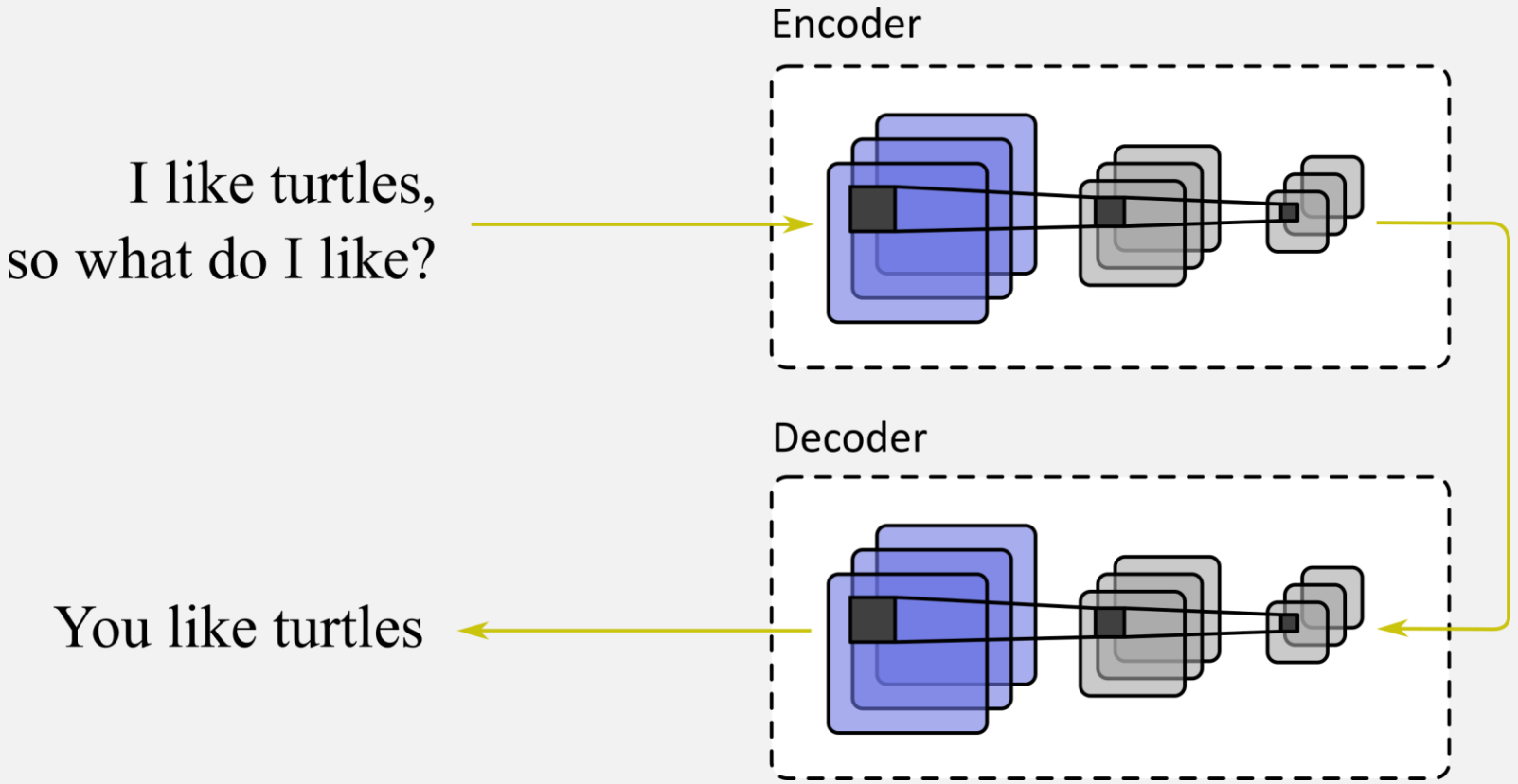

Visual Attention in
Image Captioning
Multi-Step Attention in
Natural Language Processing
since
1990s
since
2014
since
2015
since
2016
Visual Attention in
Human Perception
Global Attention in
Machine Translation

Multi-Step Attention
in **Natural Language Processing**

Encoder: CNN

Decoder: CNN

Architecture

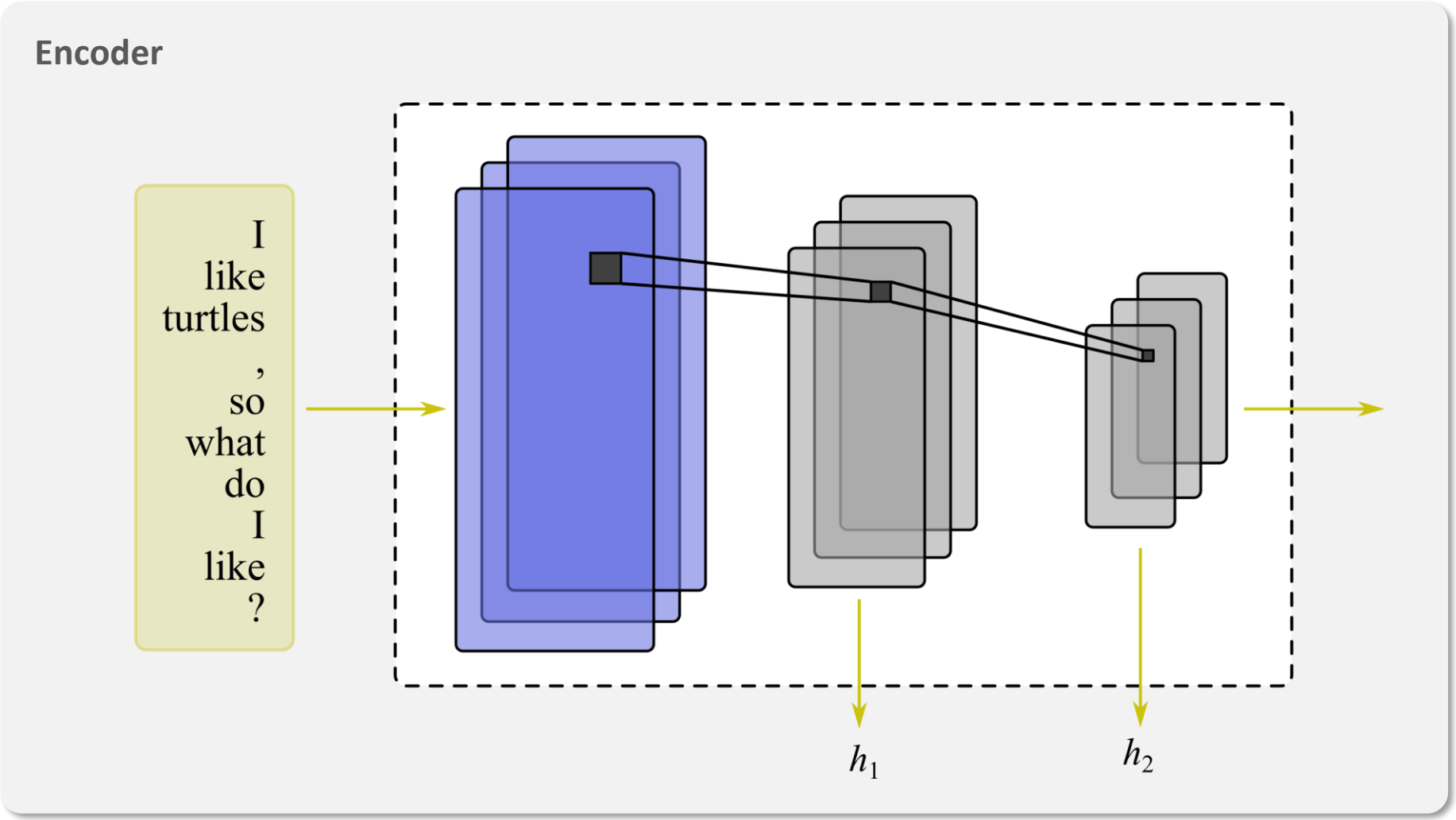


Multi-Step Attention
in **Natural Language Processing**

Encoder: CNN

Sequence of convolutional features encodes the context building stack

Decoder: CNN



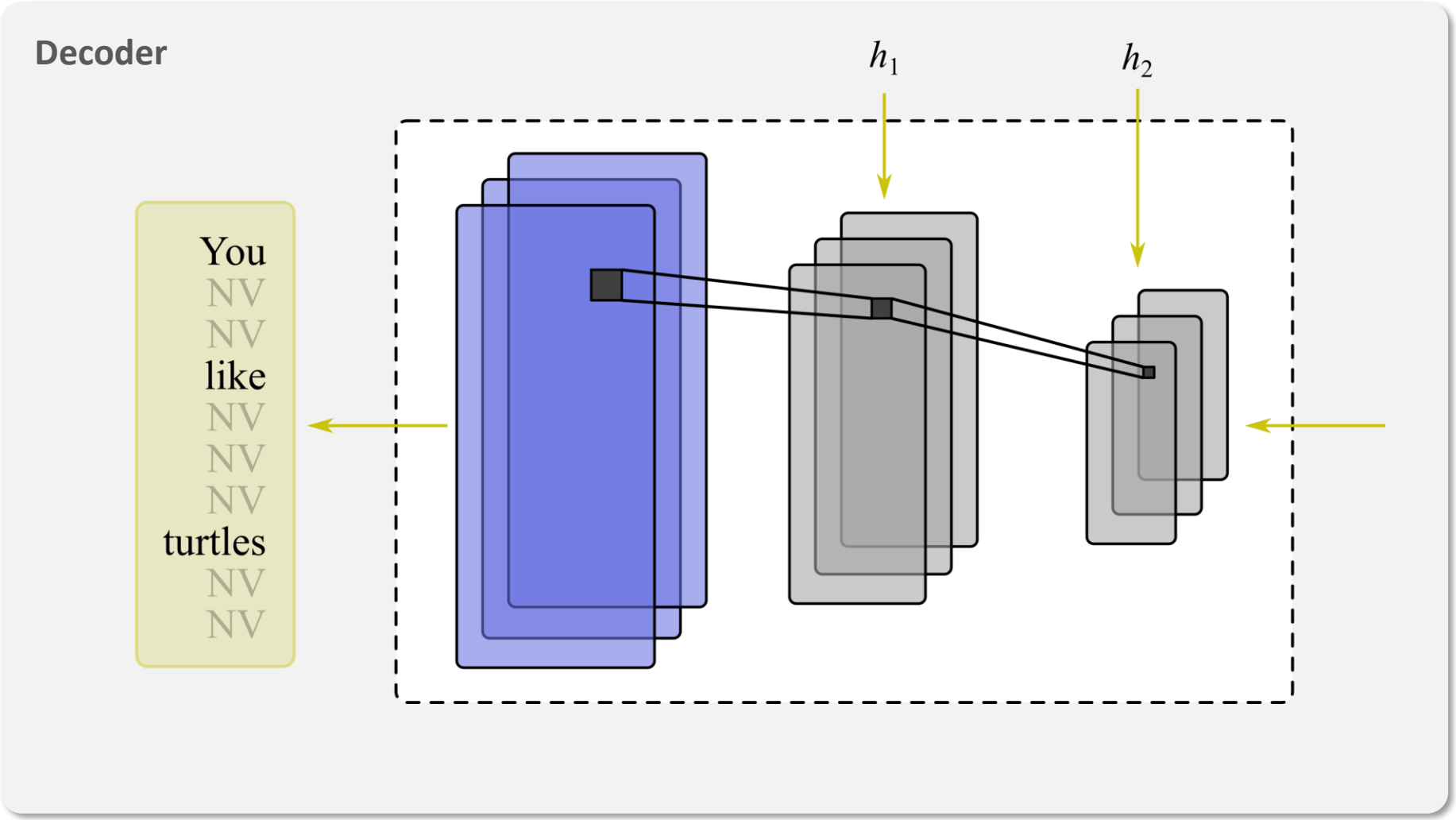
Multi-Step Attention in Natural Language Processing

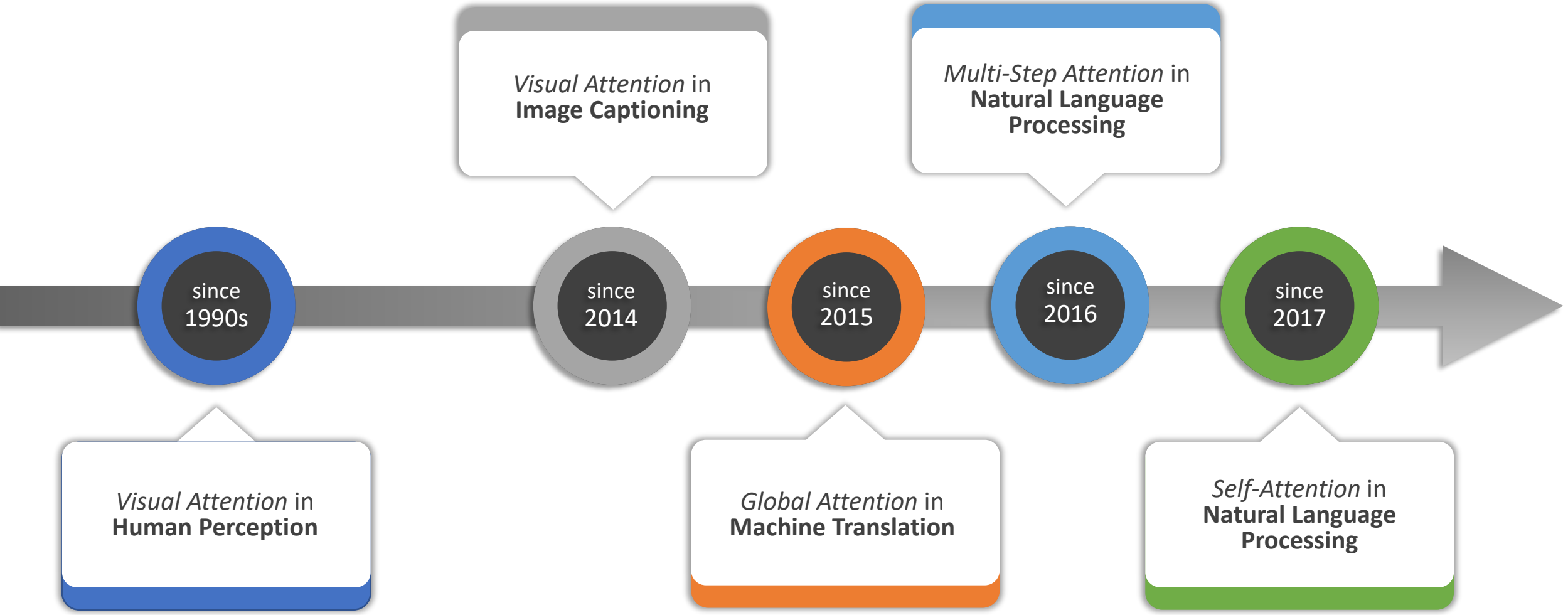
Encoder: CNN

Sequence of convolutional features encodes the context building stack

Decoder: CNN

Multi-Step Attention allows to recover the current context of current deconvolutional steps



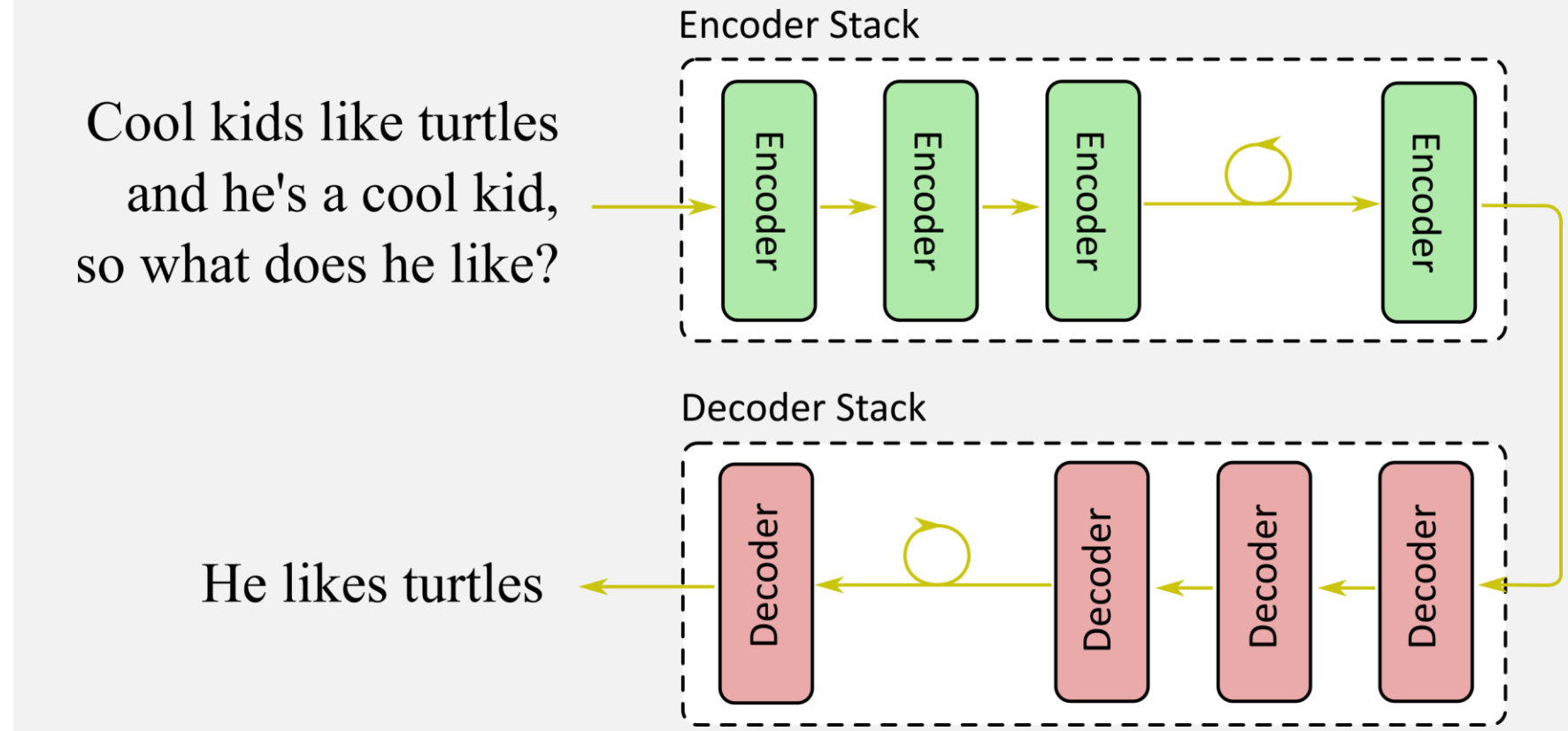

since 1990s
Visual Attention in Human Perception
since 2014
Visual Attention in Image Captioning
since 2015
Global Attention in Machine Translation
since 2016
Multi-Step Attention in Natural Language Processing
since 2017
Self-Attention in Natural Language Processing

Self-Attention in Natural Language Processing

Encoder: Stacked ANN

Decoder: Stacked ANN

Architecture

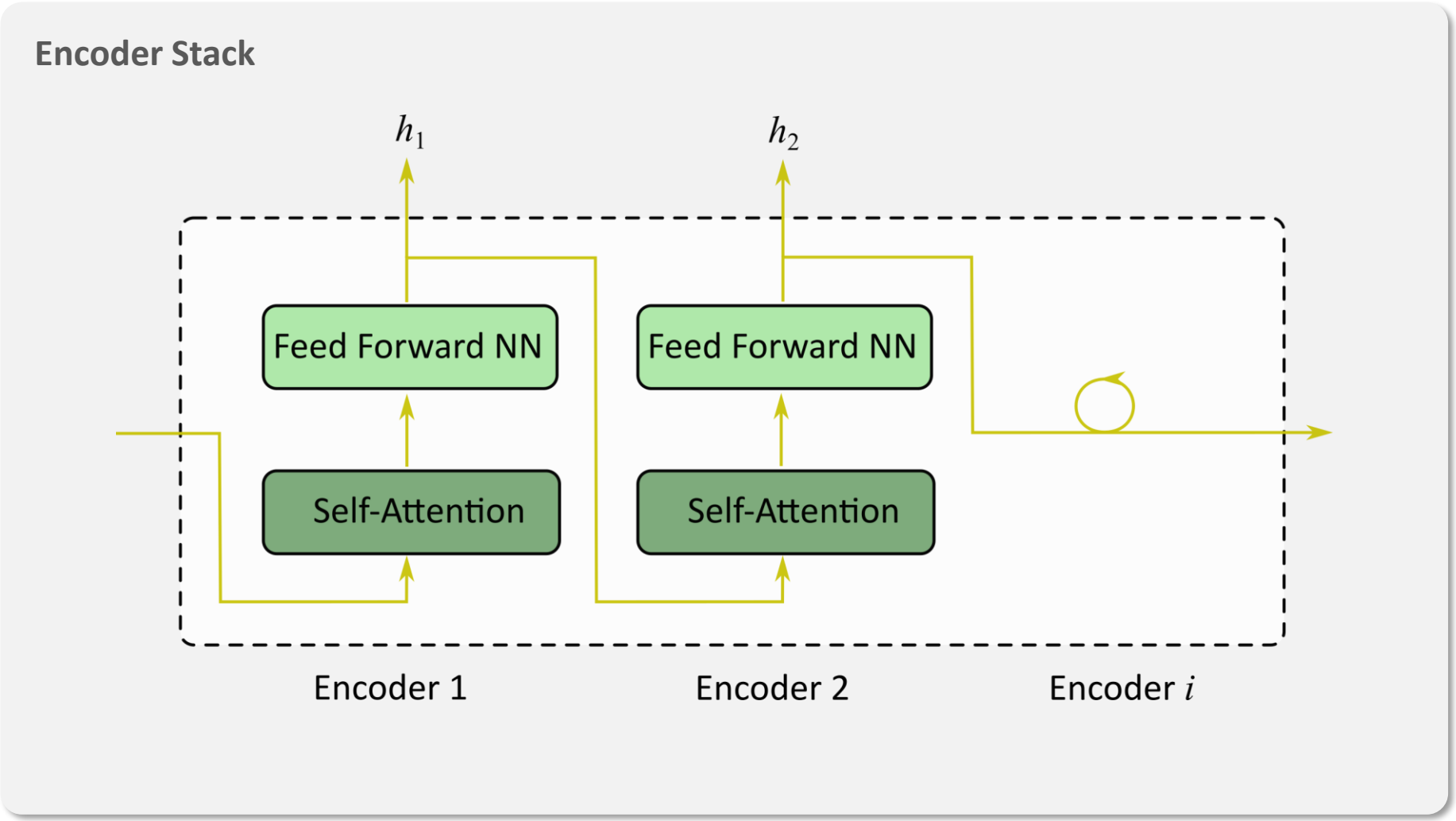


Self-Attention in Natural Language Processing

Encoder: Stacked ANN

Sequence of hidden features captures the encoder context building stack

Decoder: Stacked ANN



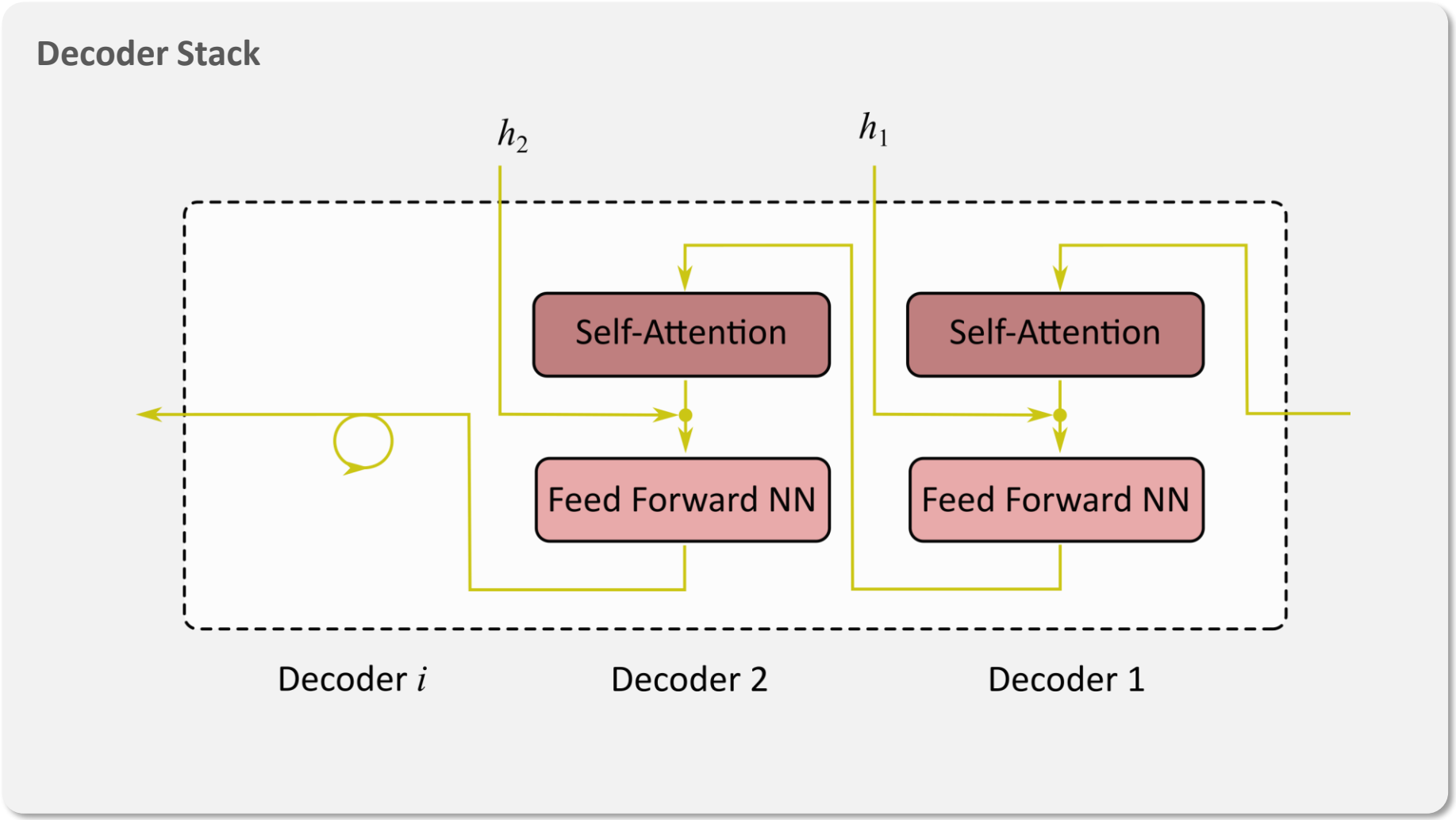
Self-Attention in Natural Language Processing

Encoder: Stacked ANN

Sequence of hidden features captures the encoder context building stack

Decoder: Stacked ANN

Multi-Step Attention unrolls the context stack to the decoder



Self-Attention in Natural Language Processing

Encoder: Stacked ANN

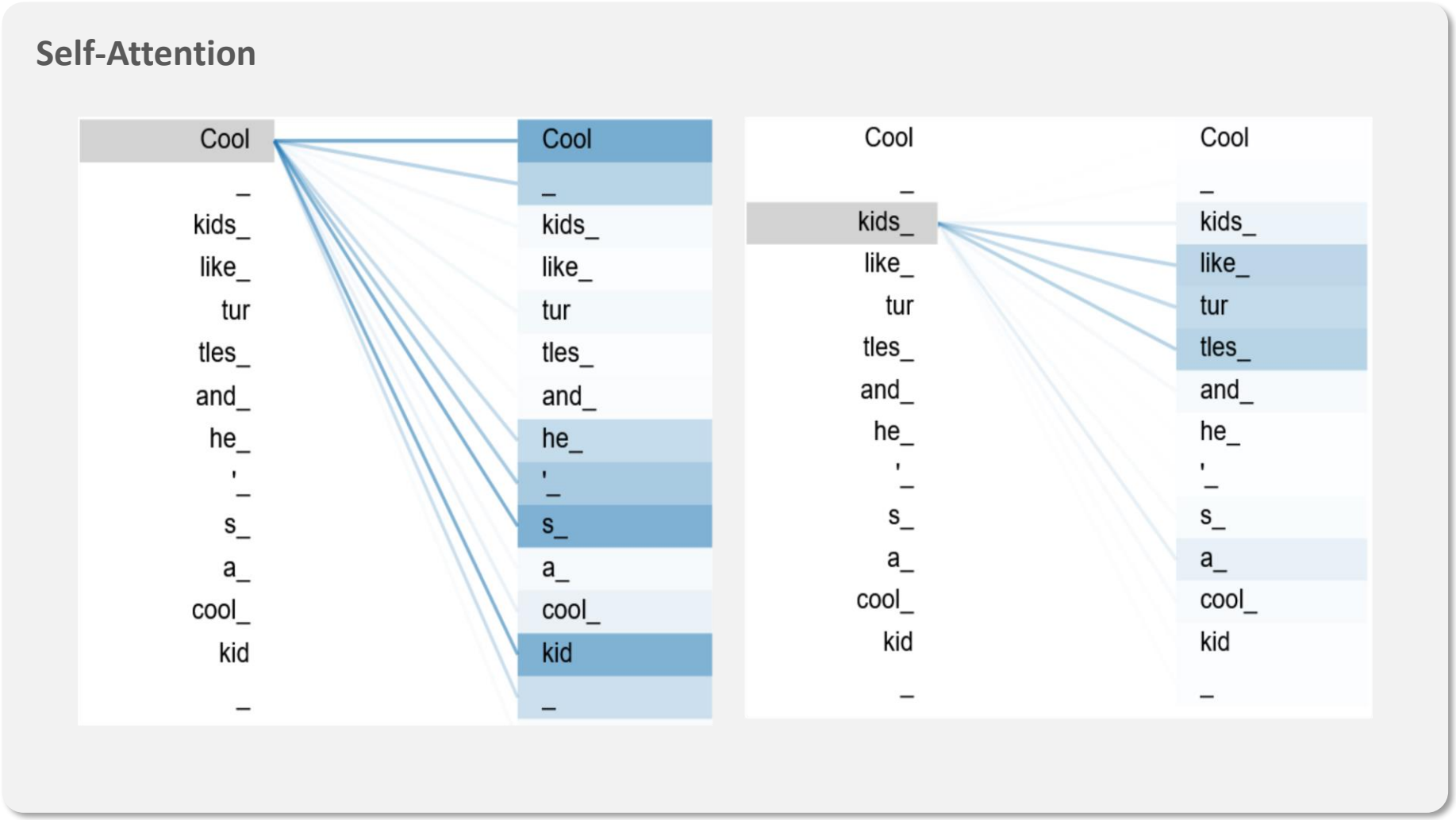
Sequence of hidden features captures the encoder context building stack

Decoder: Stacked ANN

Multi-Step Attention unrolls the context stack to the decoder

---

Self-attention aggregates context dependencies within the inputs



Summary

#1

Human Perception is based on the dynamics between ***selection*** and ***recognition***

#2

Attention machanisms immitate this behaviour by using intermediate states, that entangle *context* with *semantic* information

#3

The incorporation of context provides **dynamic features** that are context specific and therefore improve the model perfomance

#4

After all - attention mechanisms can also help to **understand the decisions** of deep networks

Summary

#1

Human Perception is based on the dynamics between ***selection*** and ***recognition***

#2

Attention machanisms immitate this behaviour by using intermediate states, that entangle ***context*** with ***semantic*** information

#3

The incorporation of context provides **dynamic features** that are context specific and therefore improve the model perfomance

#4

After all - attention mechanisms can also help to **understand the decisions** of deep networks

Summary

#1

Human Perception is based on the dynamics between ***selection*** and ***recognition***

#2

Attention machanisms immitate this behaviour by using intermediate states, that entangle ***context*** with ***semantic*** information

#3

The incorporation of context provides **dynamic features** that are context specific and therefore improve the model perfomance

#4

After all - attention mechanisms can also help to **understand the decisions** of deep networks

Summary

#1

Human Perception is based on the dynamics between ***selection*** and ***recognition***

#2

Attention machanisms immitate this behaviour by using intermediate states, that entangle ***context*** with ***semantic*** information

#3

The incorporation of context provides **dynamic features** that are context specific and therefore improve the model perfomance

#4

After all - attention mechanisms can also help to **understand the decisions** of deep networks

Thank you for your attention!