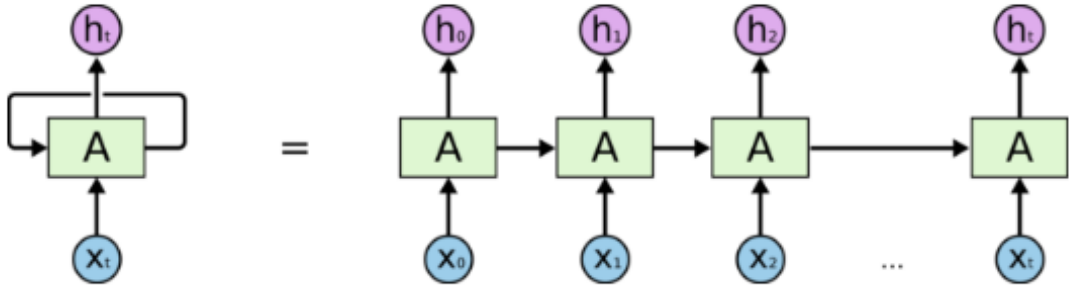
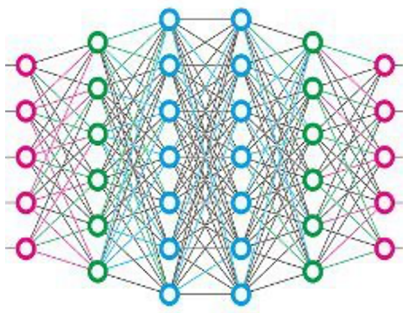
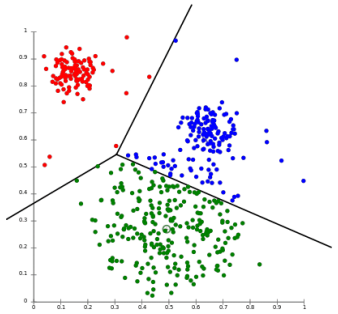
CS 445 Introduction to Machine Learning

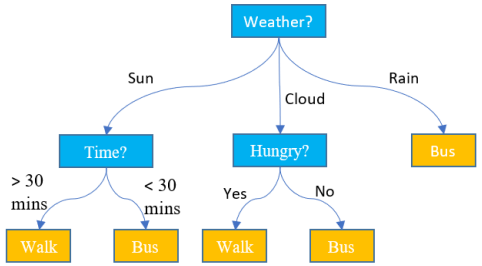


Recurrent Neural Networks (RNNs)

Instructor: Dr. Kevin Molloy







Announcements

PA 3

- Due this Friday, November 13th at 5:00 pm
- Should be starting your CNN models now

Learning Objectives

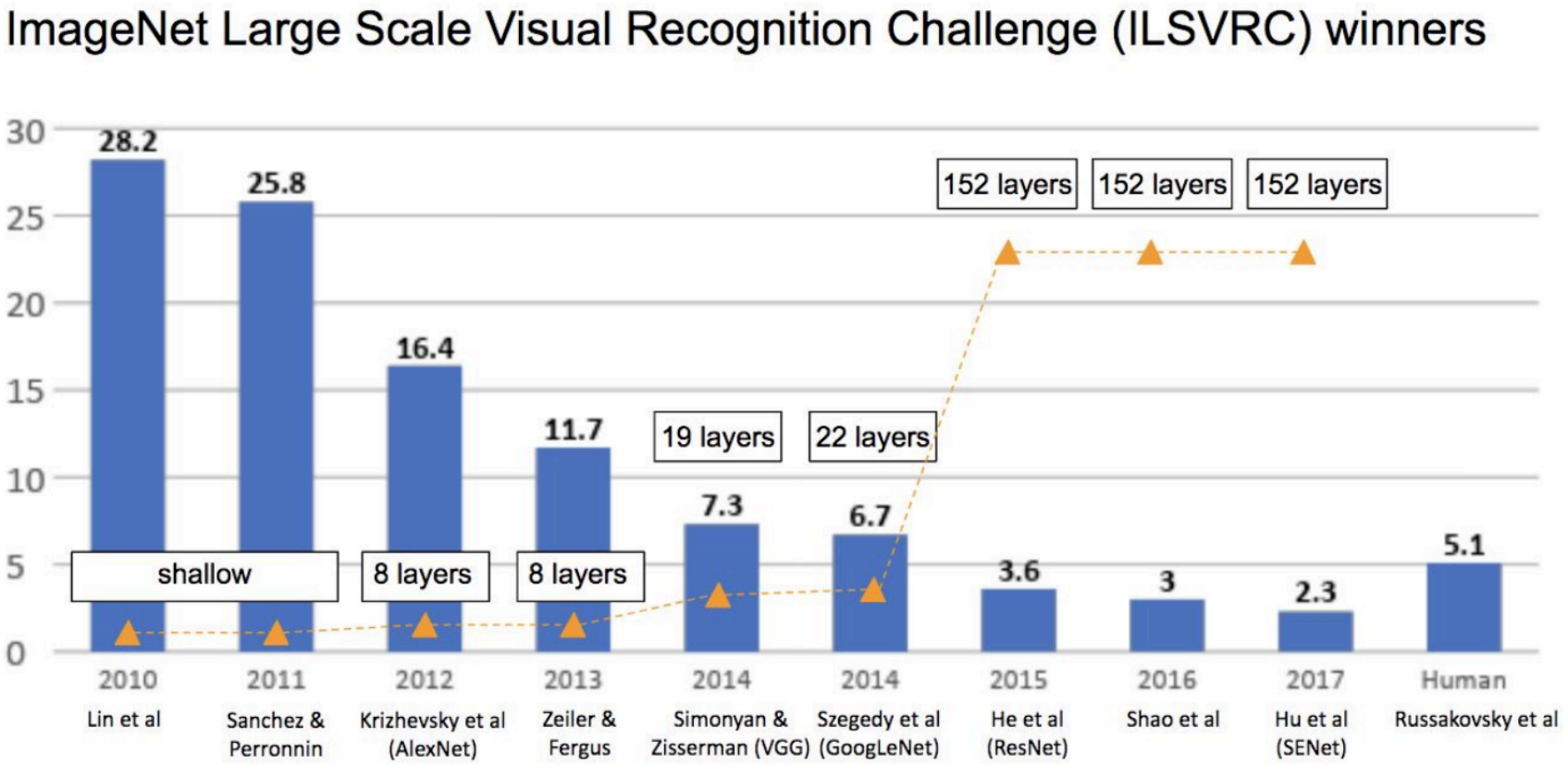
- Keras saving a model for autolab submission
- Review CNN architectures and present popular architectures
- Define recurrent neural networks (RNNs) and sequential type problems

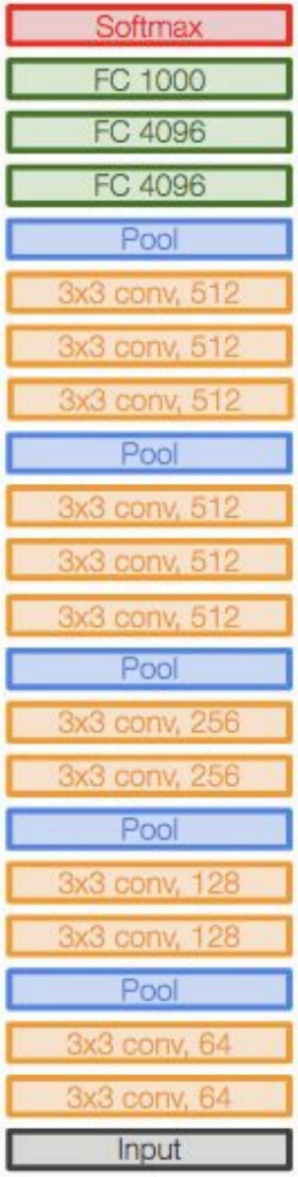
Saving Your Model in Keras

- Your model is saved so that Autolab does not need to train/fit
- Use the ***keras_test_digit_model.py*** to load your model and test it on the test data. If your files do not work with this script, they will not work within Autolab.

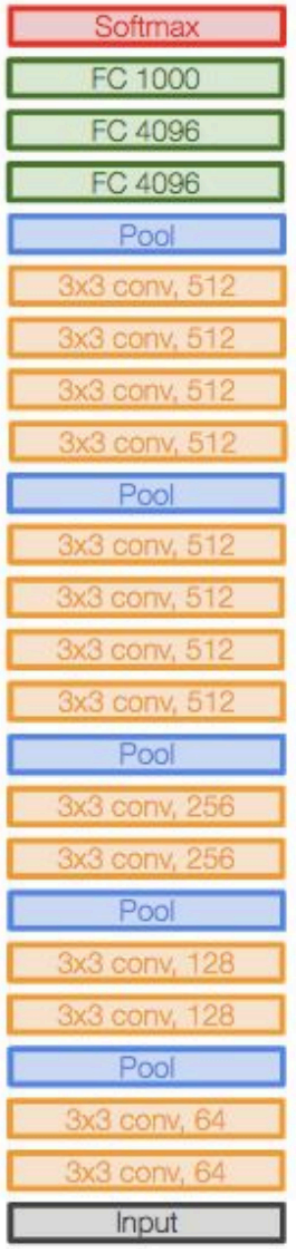
```
model = Sequential()
…
model.save('filename.h5')
```

Convolutional Progress



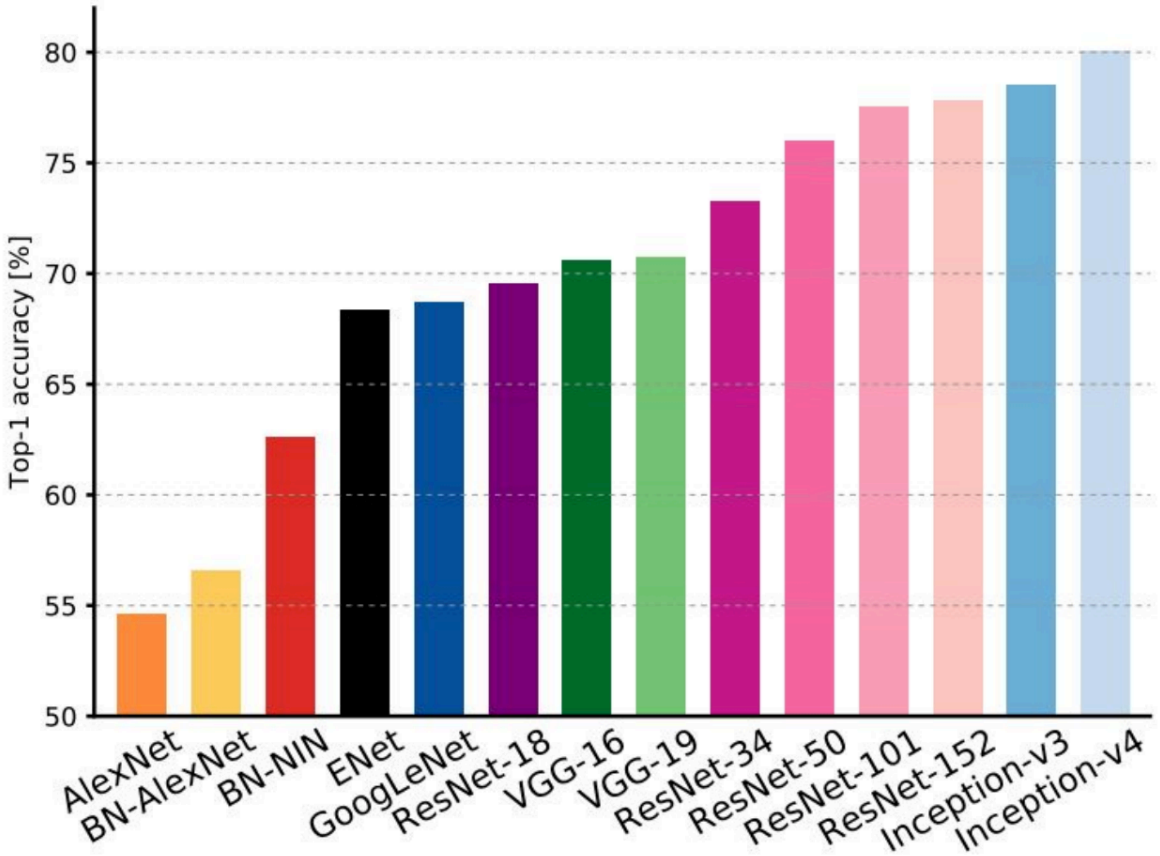


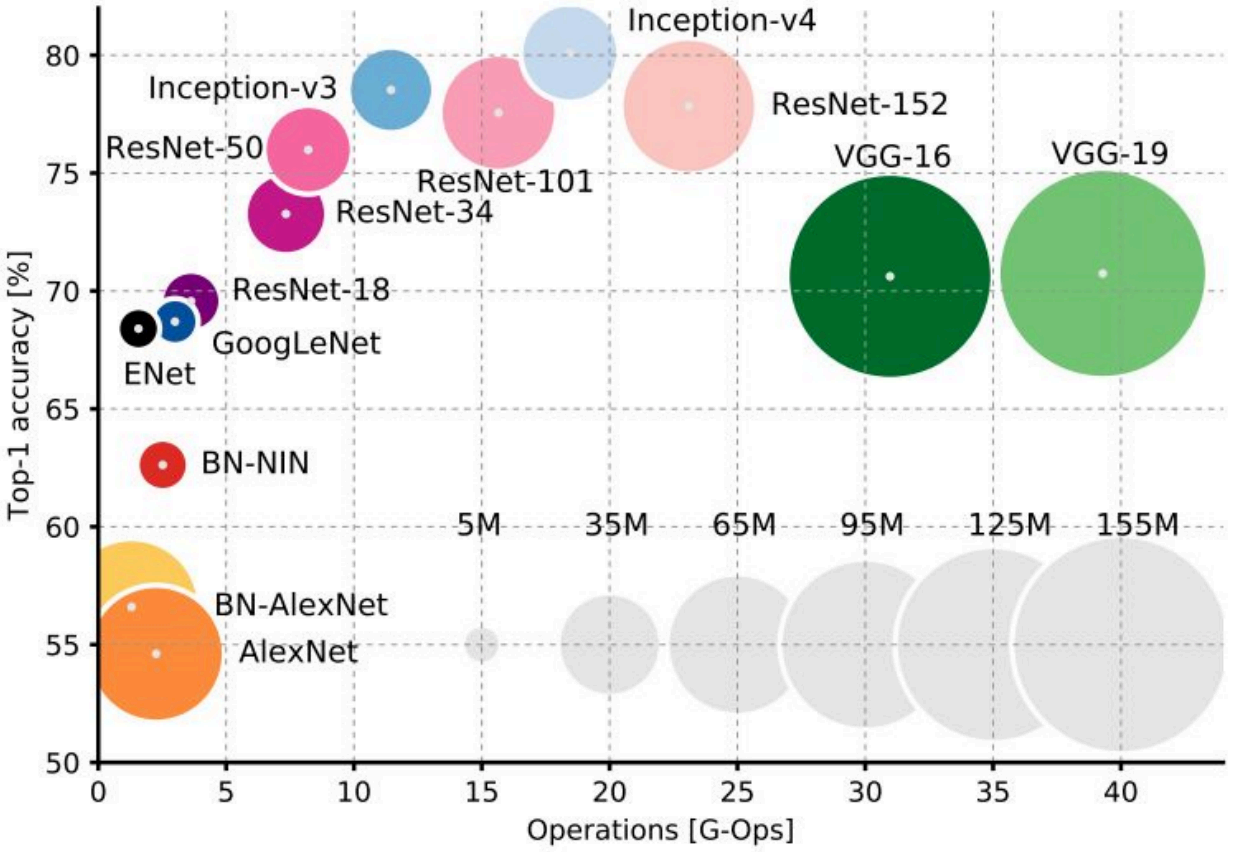
VGG16



VGG19

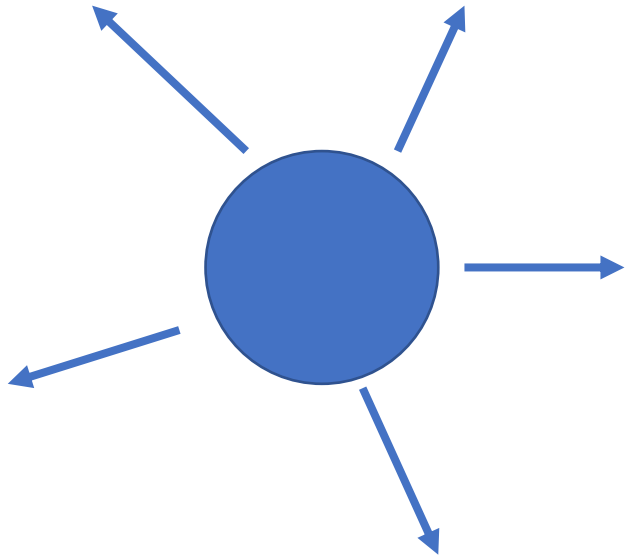
Convolutional Progress



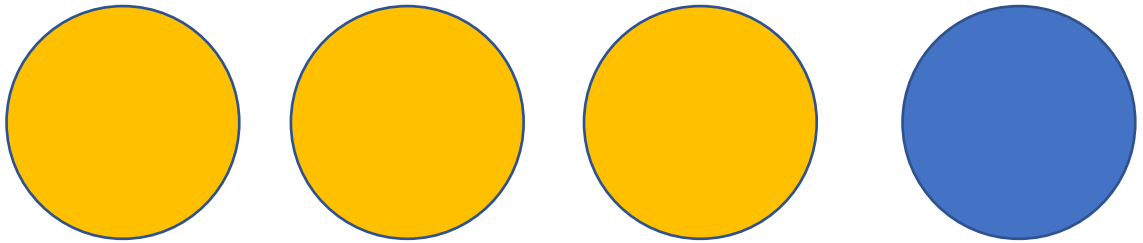


Figures from Canziani, Paszke and Culurciello, 2017

Making Sequential Predictions?



Making Sequential Predictions?

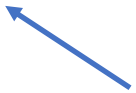


What Problems Have Sequential Data?

- Audio – sequence of soundwaves
- Text – sequence of words
- Medical signals (heartbeats, ekg)
- Amino acid sequences (in proteins)
- Stock market predictions

Example: Predicting the Next Word

I took my dog for a walk



Given these words

Predict this word

Challenges to do this with the networks we have reviewed so far?

Example: Predicting the Next Word

I took my dog for a walk

Given these words

Predict this word

Challenges to do this with the networks we have reviewed so far?

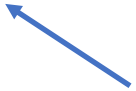
Feed forward networks have a **fixed input size** (you specify this in the first layer as the input_dim).

Potential fix – only look at a fixed number of words prior to the one we want to predict. Issues?

Example: Predicting the Next Word

I took my dog for a walk

Given these words



Predict this word

Challenges to do this with the networks we have reviewed so far?

Feed forward networks have a **fixed input size** (you specify this in the first layer as the input_dim).

Potential fix – only look at a fixed number of words prior to the one we want to predict (maybe one word). Issues?

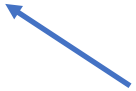
Imagine the sentence.

Toulouse, France is where I did my post doctoral studies. I speak fluent _____.

Example: Predicting the Next Word

I took my dog for a walk

Given these words



Predict this word

Challenges to do this with the networks we have reviewed so far?

Feed forward networks have a **fixed input size** (you specify this in the first layer as the input_dim).

Potential fix – only look at a fixed number of words prior to the one we want to predict (maybe one word). Issues?

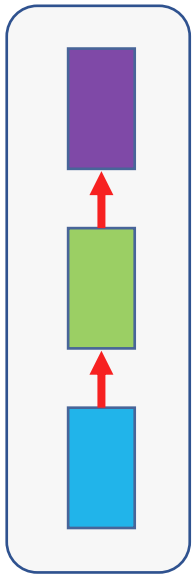
Imagine the sentence.

Toulouse, France is where I did my post doctoral studies. I speak fluent _____.

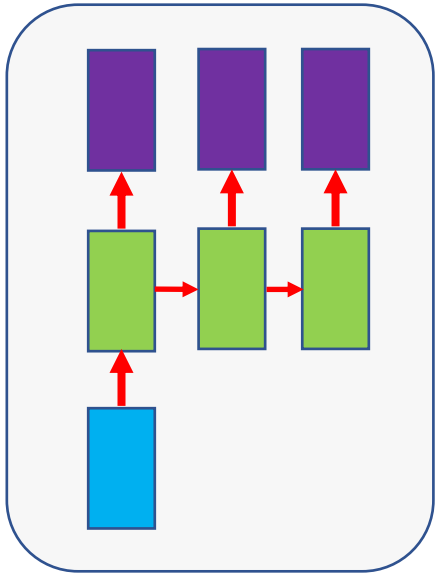
Highlights that we need information form the past in order to accurately predict the next word.

RNN: Process Sequences

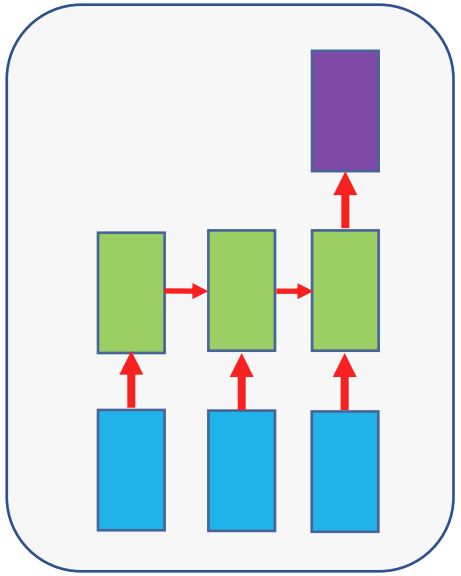
one to one



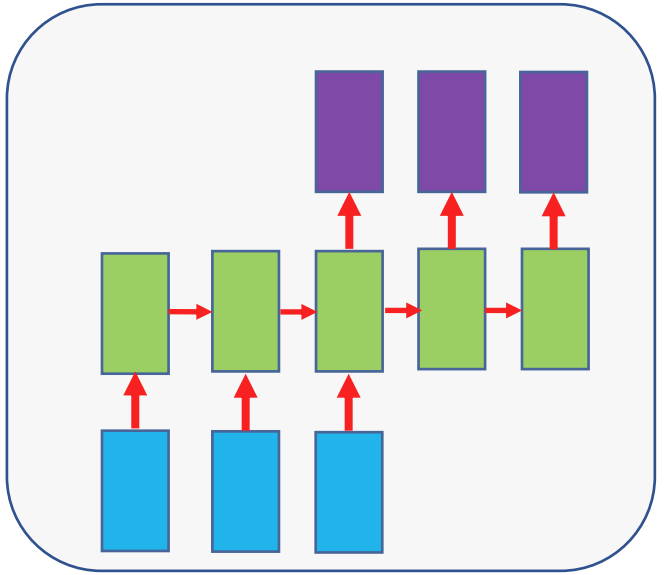
one to many



many to one



many to many

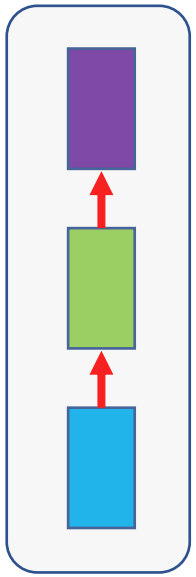


Vanilla Neural Networks

(Feed forward networks)

RNN: Process Sequences

one to one



one to many

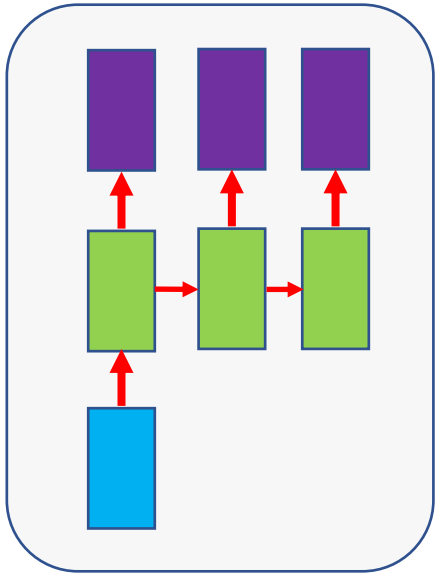
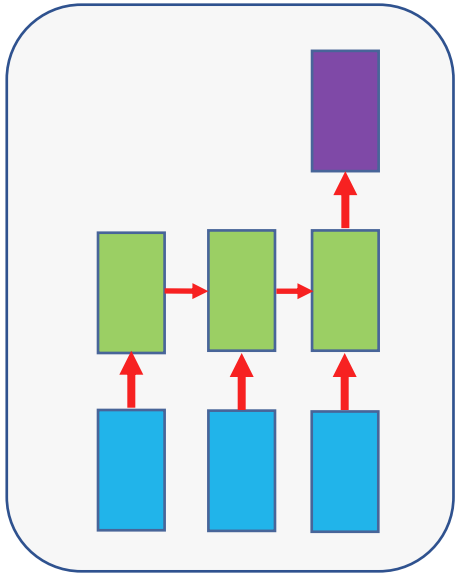


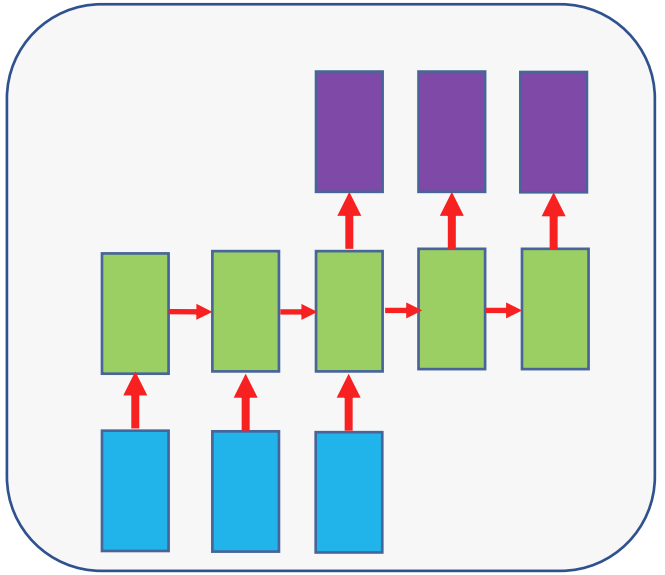
Image captioning

Image -> sequence of words

many to one

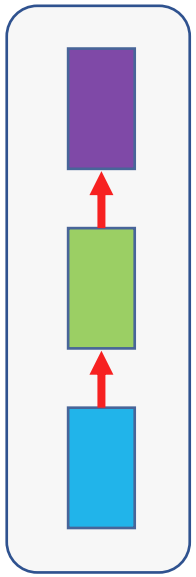


many to many

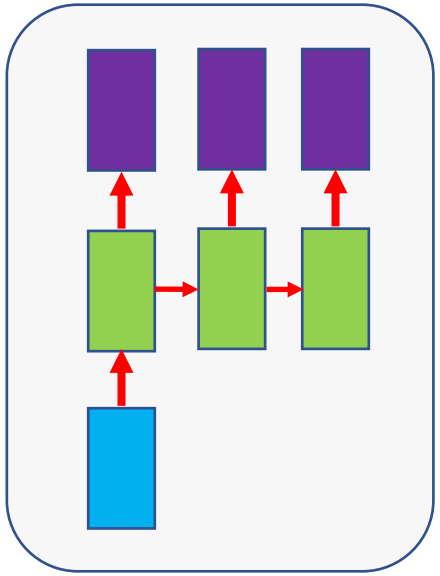


RNN: Process Sequences

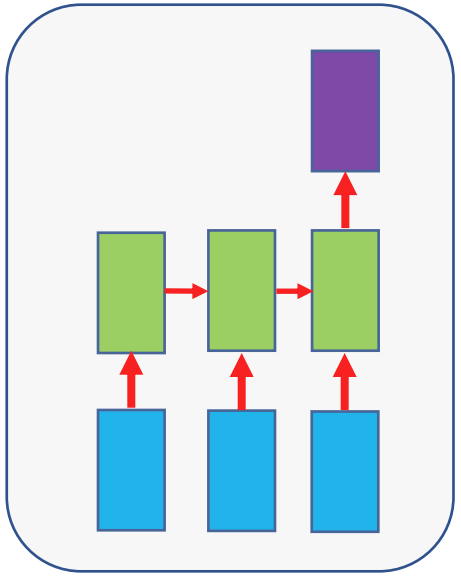
one to one



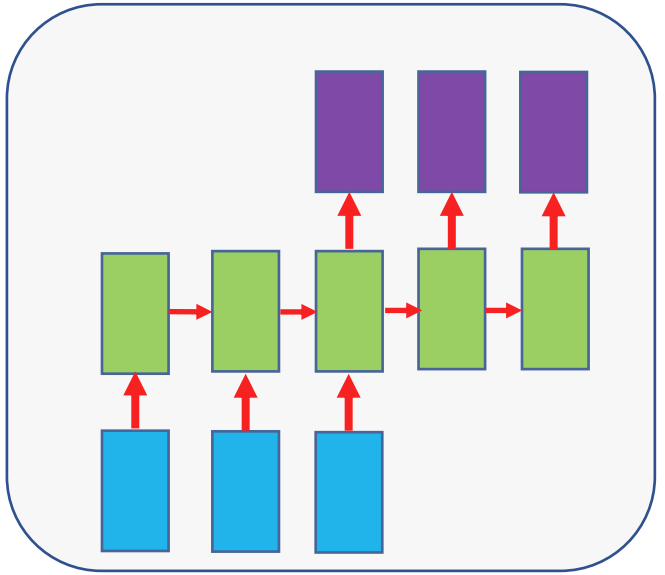
one to many



many to one



many to many

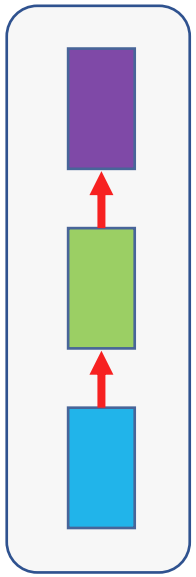


Action prediction

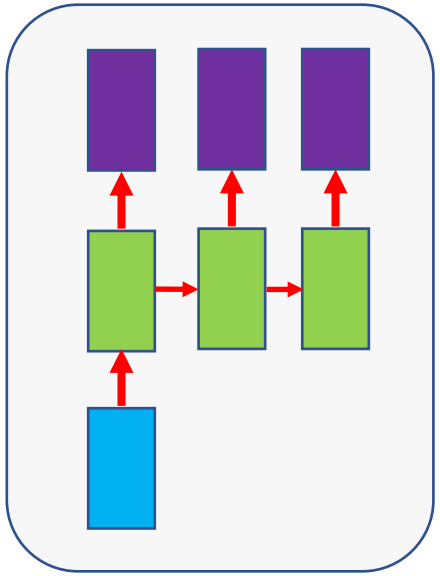
Sequence of video frames -> action class

RNN: Process Sequences

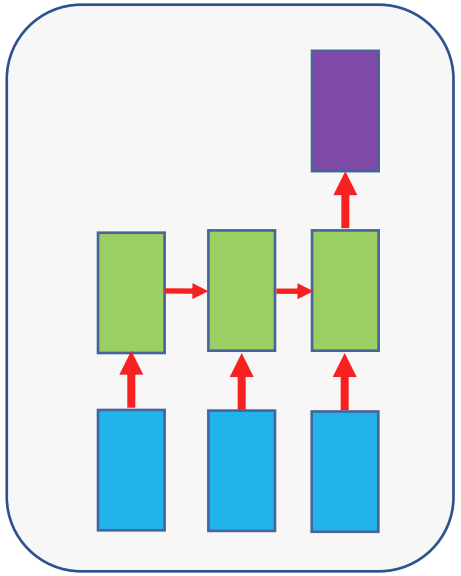
one to one



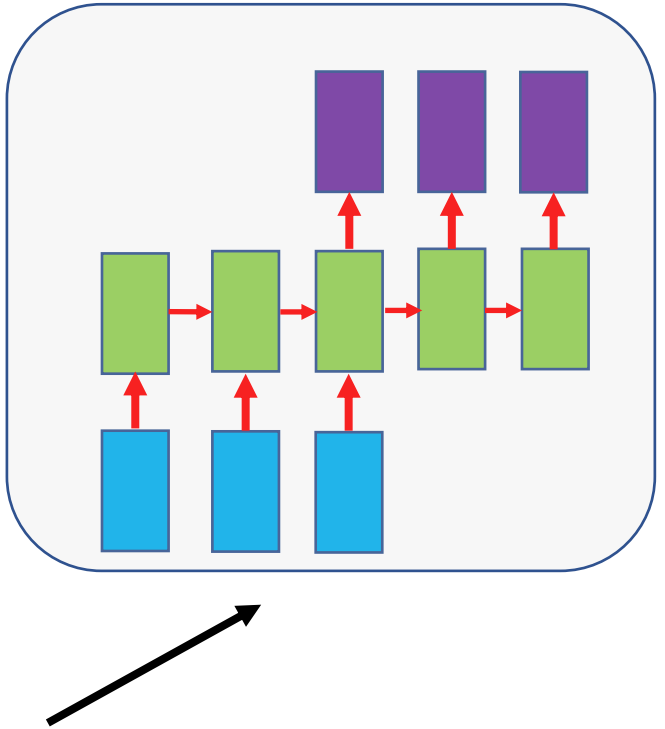
one to many



many to one



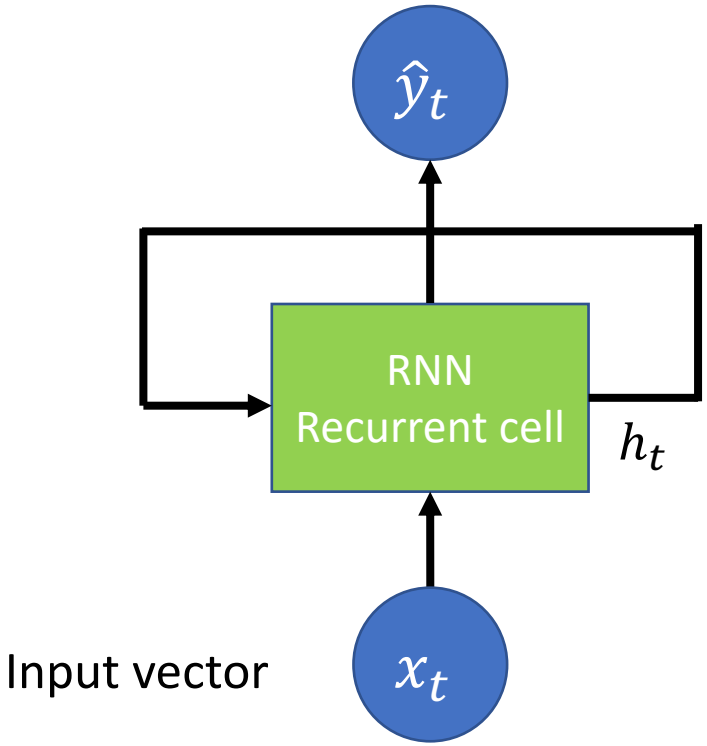
many to many



Video classification

Frame by frame

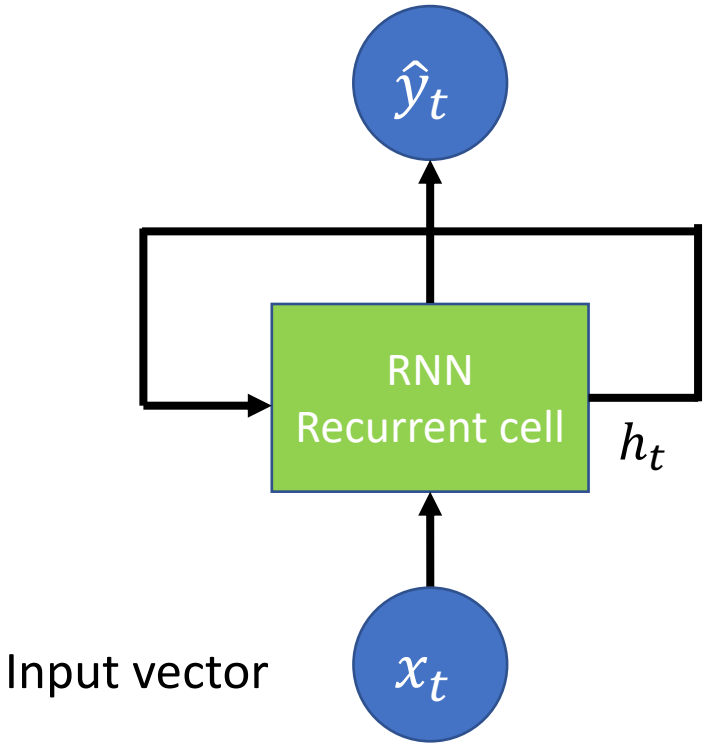
RNN: Process Sequences



Apply a recurrence relation at each time step

$$h_t = f_W\ (h_{t-1}, x_t)$$

RNN: Process Sequences



Apply a recurrence relation at each time step

$$h_t = f_W(h_{t-1}, x_t)$$

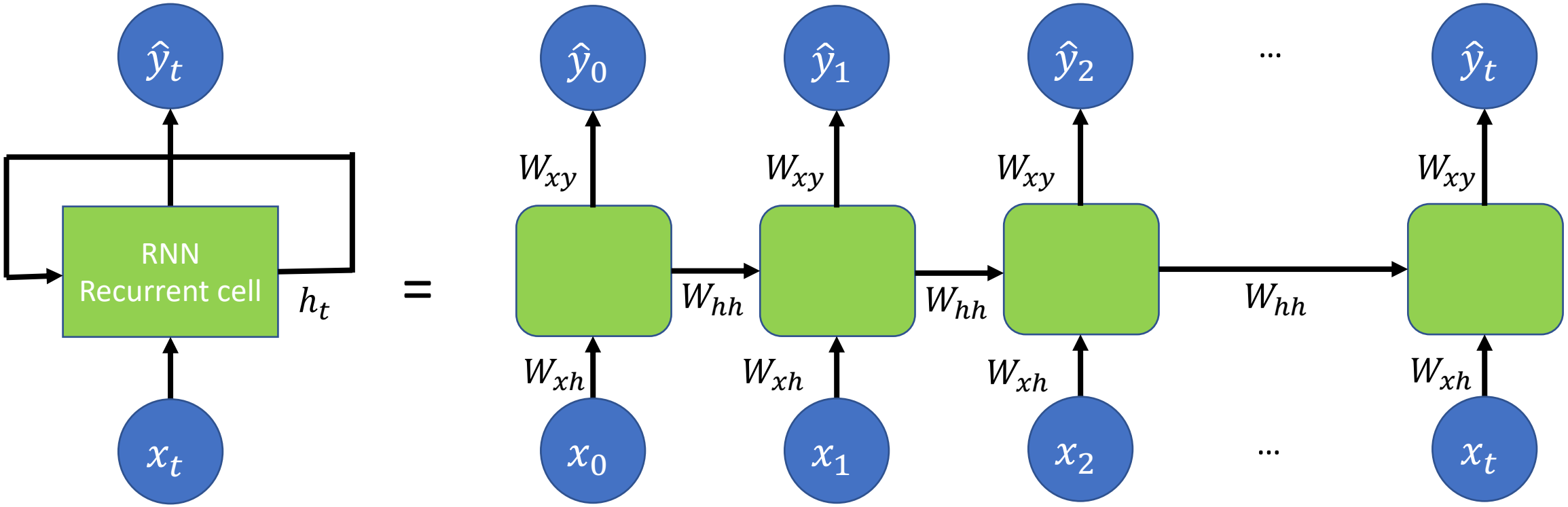
Cell state

Function (with weights W)

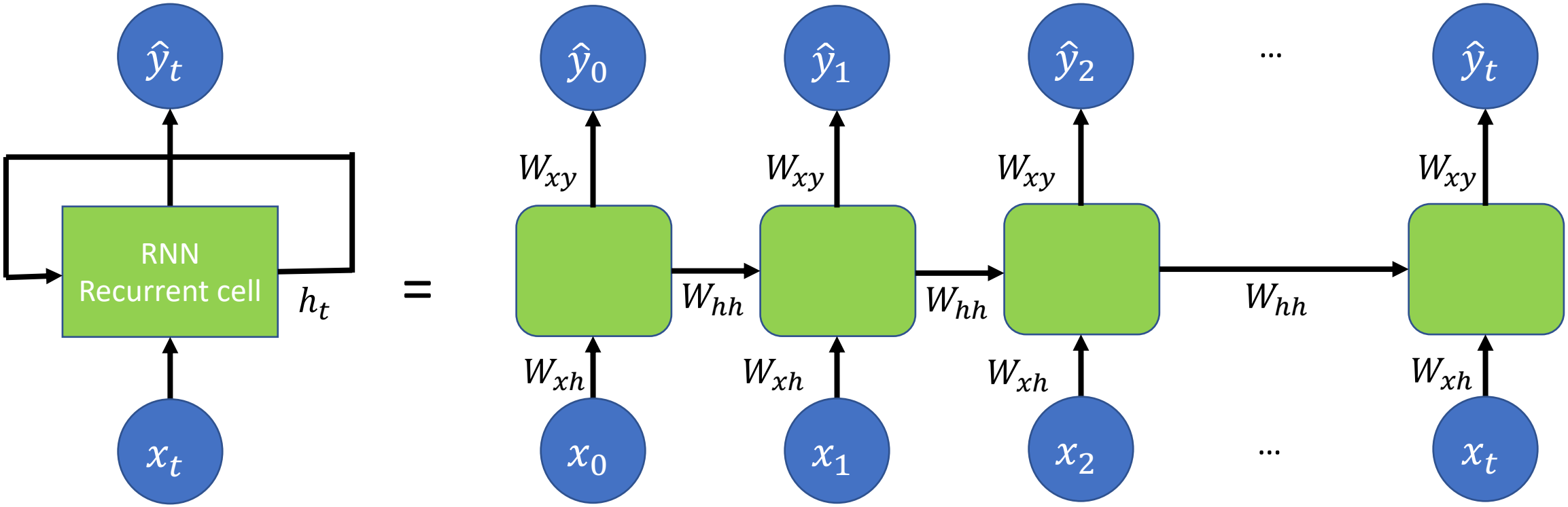
old state

Input at time t

RNN: Unrolling the Process



RNN: Unrolling the Process



Same weight matrices at every time step.

W_{xh}= input_dim (per step) x RNN_units = (1,1)

W_{hh}= RNN units x RNN_Units = (1,1)

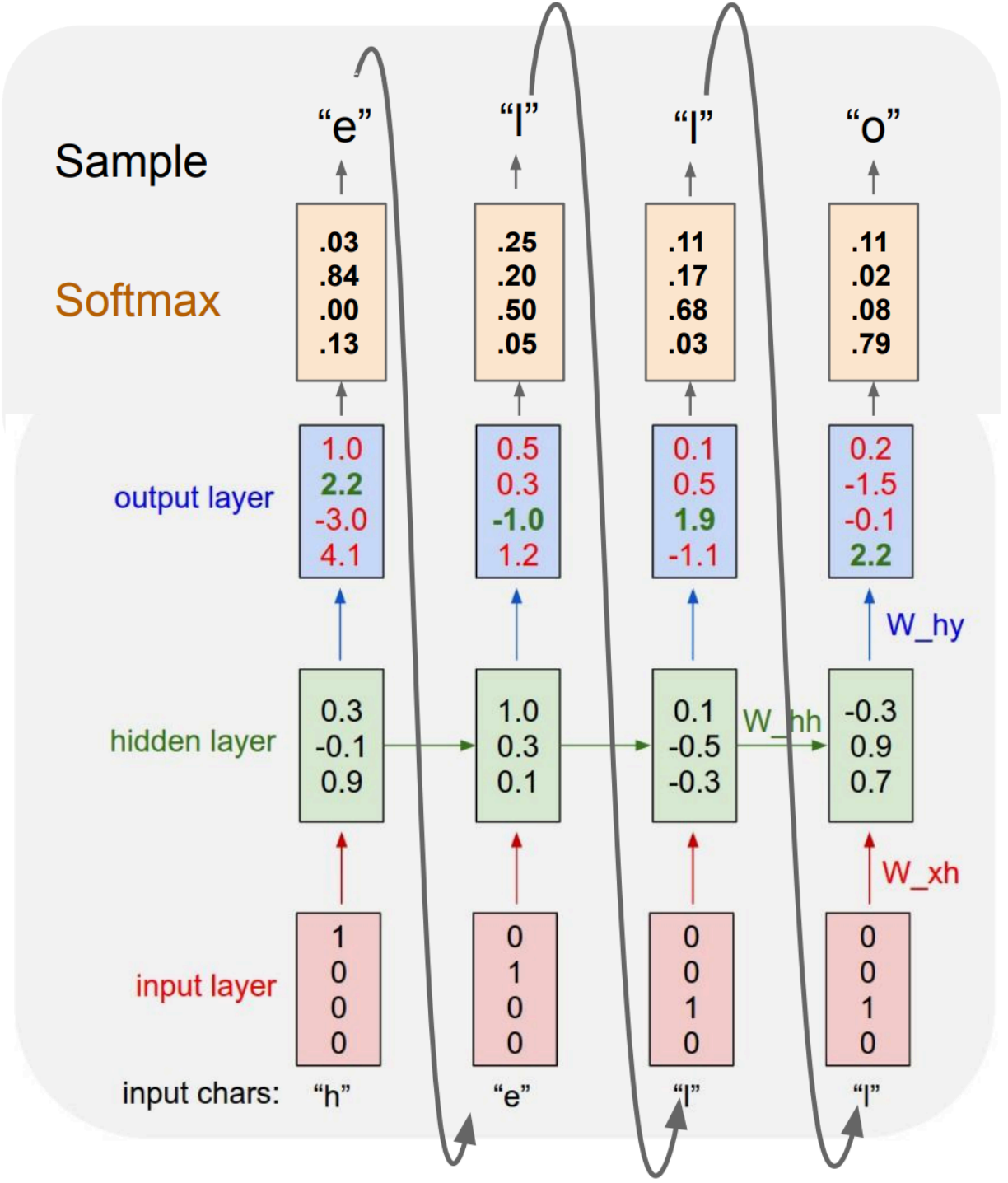
W_{xy}= Output_dim x RNN_Units = y times 1

RNN Examples

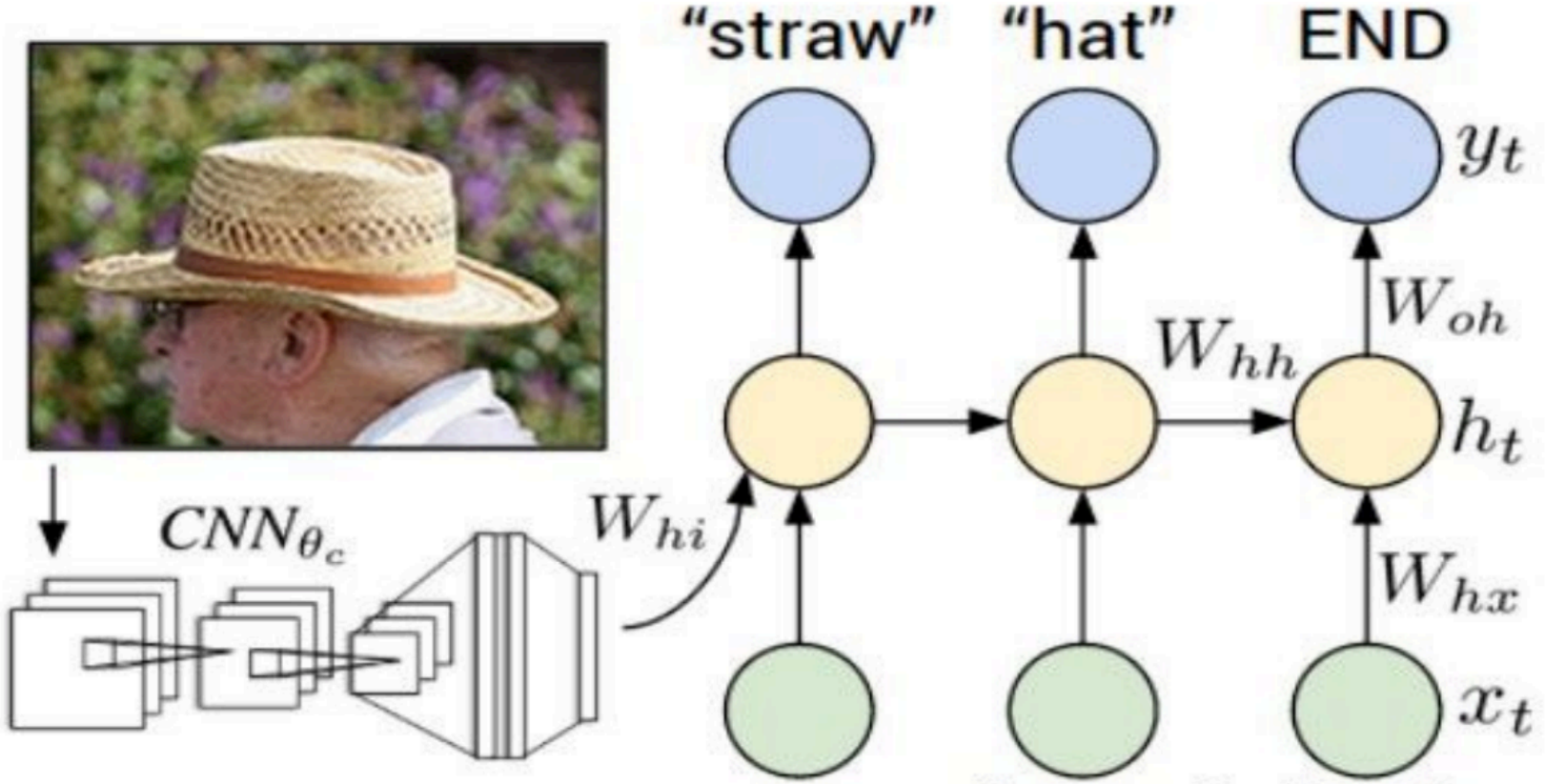
Example: Character-level Language model

Vocabulary: [h, e, l, o]

At test-time sample characters one at a time and feed back into model.



RNN Example – Image Captioning



Karpathy et al. "Deep Visual-Semantic Alignments for Generating Image Descriptions, 2015.

Image Captioning -- Results

A cat sitting on a suitcase on the floor

A cat is sitting on a tree branch

A dog is running in the grass with a frisbee

A white teddy bear sitting in the grass

Karpathy et al. "Deep Visual-Semantic Alignments for Generating Image Descriptions, 2015.

Image Captioning -- Results

A cat sitting on a suitcase on the floor

A cat is sitting on a tree branch

A dog is running in the grass with a frisbee

A white teddy bear sitting in the grass

A bird is perched on a tree branch

A man in a baseball uniform throwing a ball

Karpathy et al. "Deep Visual-Semantic Alignments for Generating Image Descriptions, 2015.

RNN Summary

Many variants of RNNs: **LSTM** (long short term memory)

Used everywhere (try texting on your cell phone and looking at word completion).

Material to know about RNNs:

- Define the recurrence relationship and be able to "unroll" an RNN
- Given a problem, decide what type of NN would be most appropriate (dense, CNN, RNN, combination) and why. And for RNN, is the network a (one to many, many to one, many to many)