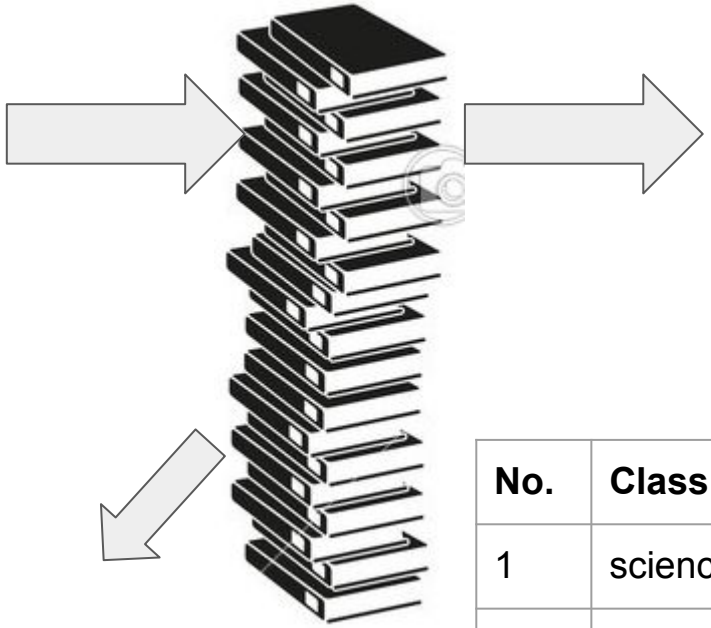
RNN for book classification

Mindlab Group, Ritual Group

Dataset construction

Project Gutenberg



goodreads Title

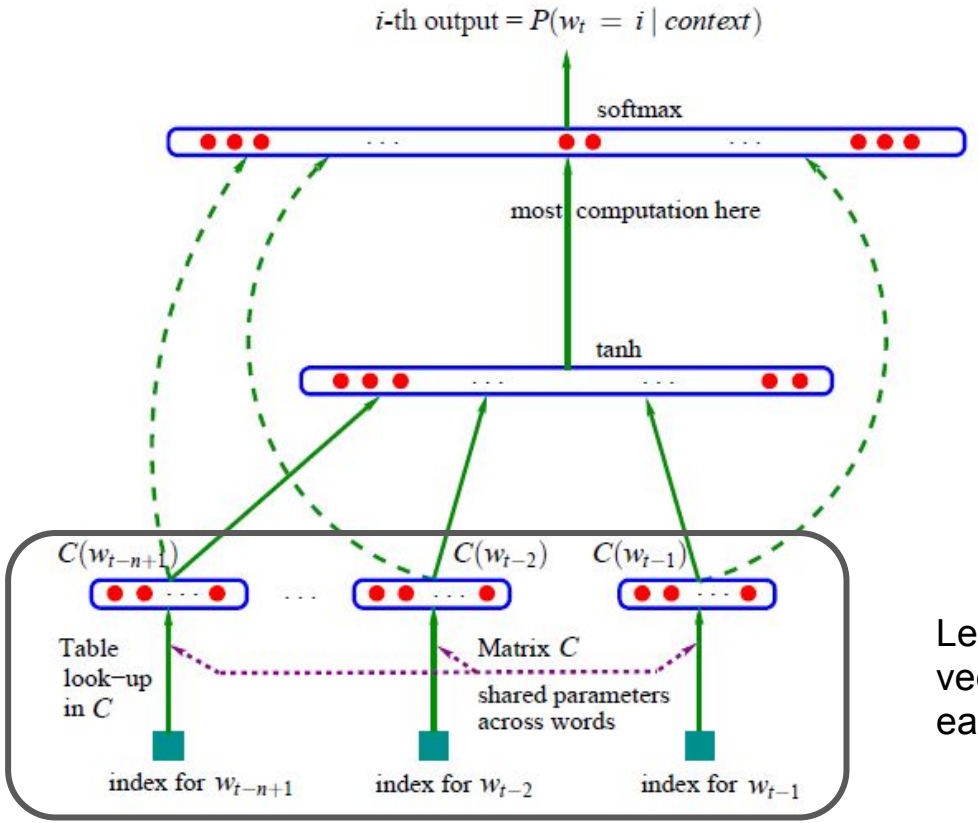
API

Get top tags from users

No.	Class	Tags
1	science_fiction	sci-fi, science-fiction
2	comedy	comedies, comedy, humor
...	...	...
9	religion	christian, religion, christianity,...

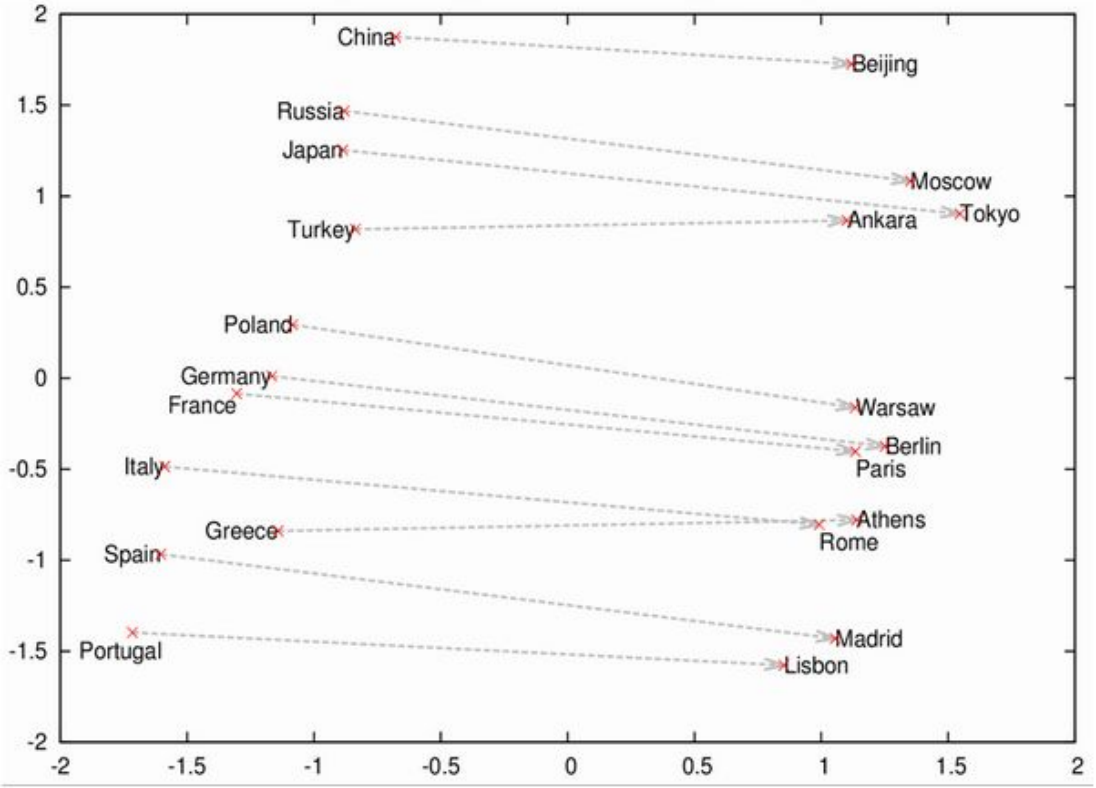
Annotated Dataset

Predict next word with NN (Language model)

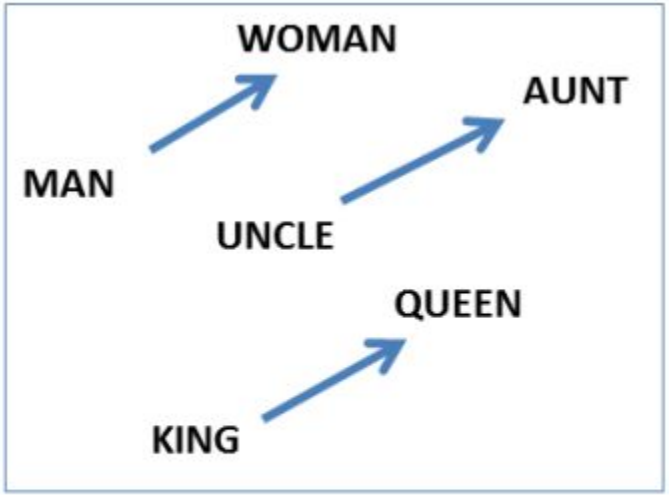


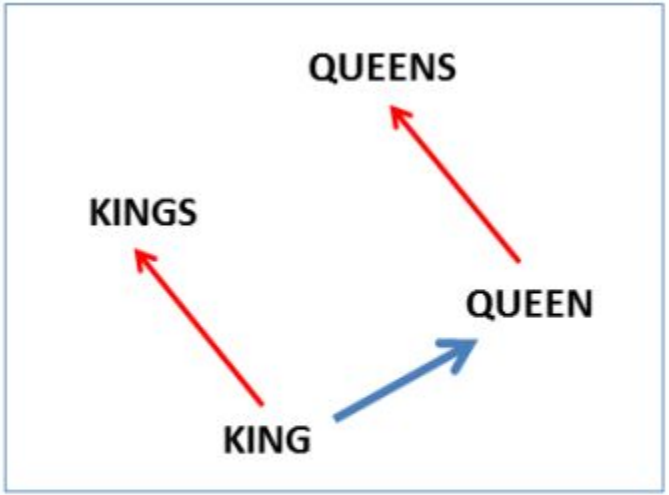
Learn one vector per each word

Word embeddings properties



Word embeddings properties





(Mikolov et al., NAACL HLT, 2013)

From words to book representation

Represent a sequence of N words, by representing each word using word2vec embedding space and average their word vectors.

Take M sequences of vectors as input for a RNN. Label all sequences with the genre of the source book.

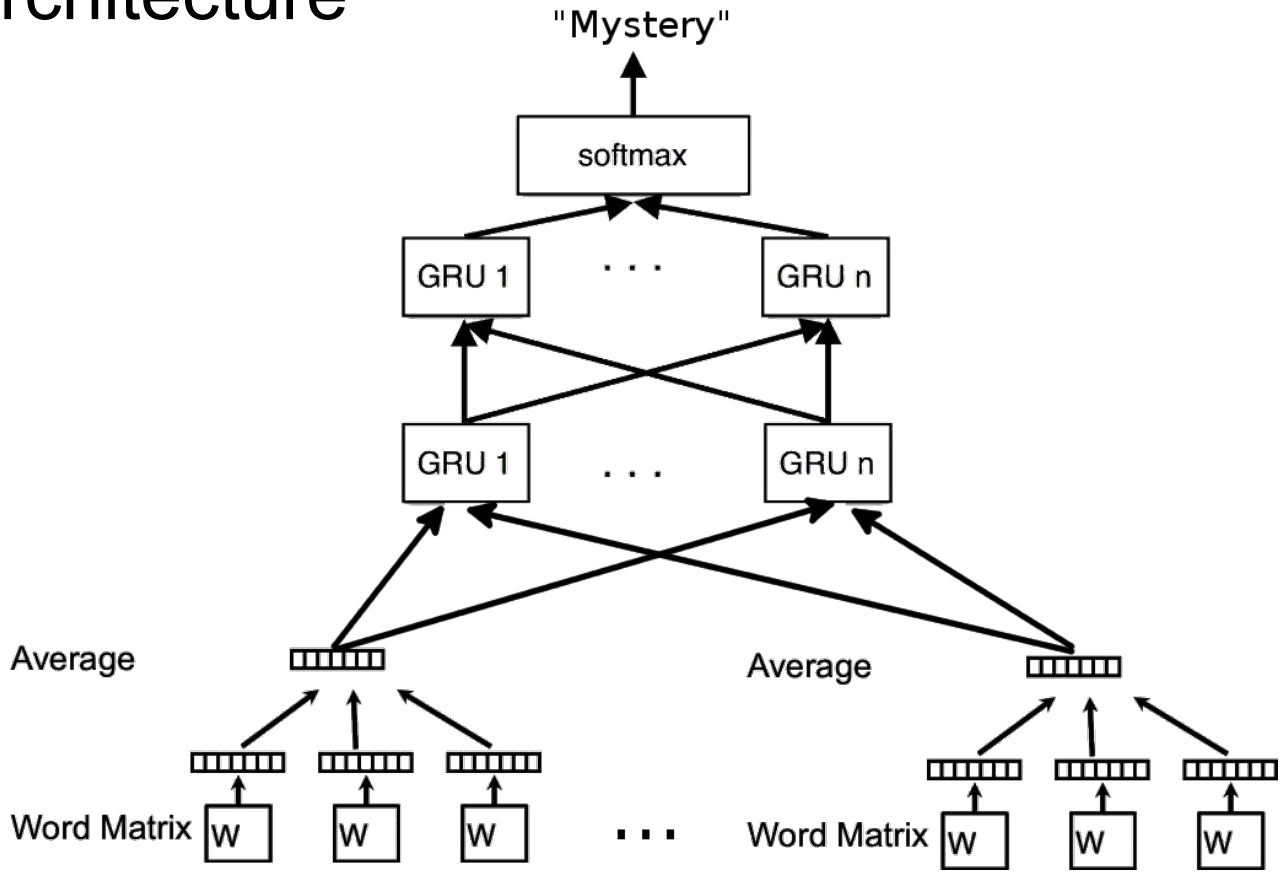
Total Books: 3629

Total Samples: ~68000

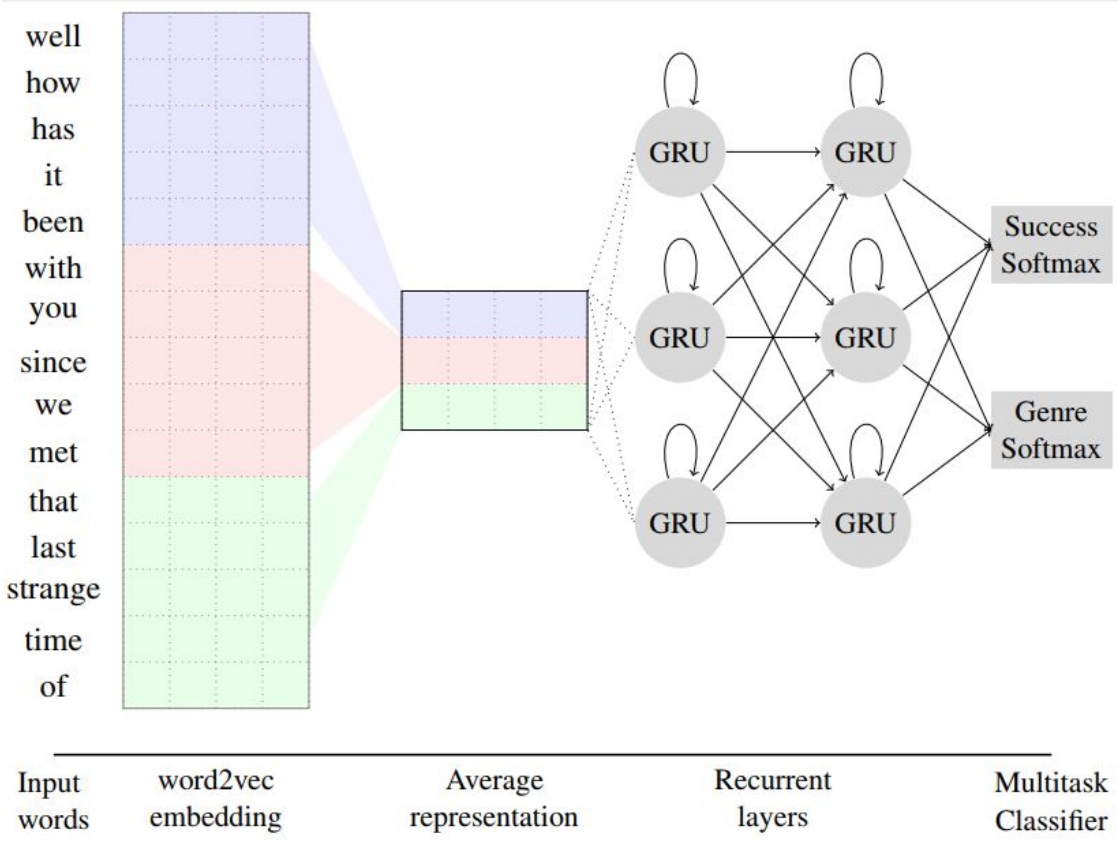
Dataset Distribution

Genre	Unsuccessful	Successful	Total
Detective Mystery	60	46	106
Drama	29	70	99
Fiction	30	81	111
Historical Fiction	16	65	81
Love Stories	20	60	80
Poetry	23	158	181
Science Fiction	48	39	87
Short Stories	123	135	258
Total	349	654	1,003

RNN Architecture



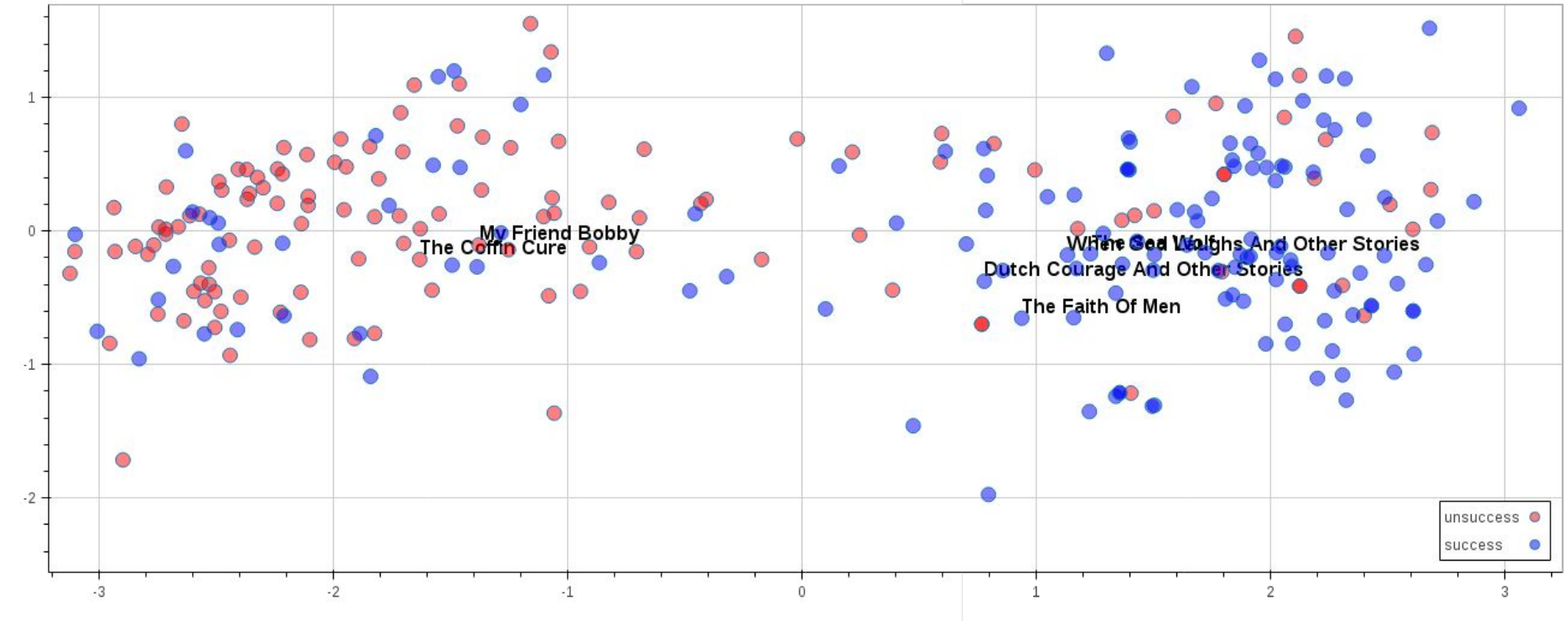
Proposed model



Results using RNN

Features	F1 score
Ashok (2013)	0.70
Char 3-5 grams	0.69
Typed prefix 3-gram	0.69
Writing Density (WR)	0.69
Readability	0.69
Sentic concepts & scores(SCS)	0.68
Book2Vec (DMM)	0.70
Book2Vec (DBoW+DMM)	0.70
Book10002Vec (DMC)	0.70
Book10002Vec (DBoW)	0.71
Book10002Vec (DBoW+DMC)	0.70
Unigram+Bigram	0.69
Book2Vec+SCS	0.71
Book2Vec+WR	0.70
Best Features except RNN	0.71
RNN	**0.79**

Representation of samples over 2D visualization



https://goo.gl/e9jO38

Image captioning

Next level of computer vision

Image Captioning

- A step beyond image classification or object detection.
- Requires the identification of complex relations between elements in the image
- Additionally requires a generative model to build meaningful sentences.
- A hard task to evaluate.
- Proposed methods focus on get higher BLEU scores, rather than solve the problem

Previous approaches

- Detect objects using complex features
- Identify actions, relations in the scene
- Train a language model
- Integrate…
- Sentence retrieval

Neural Image Caption generator

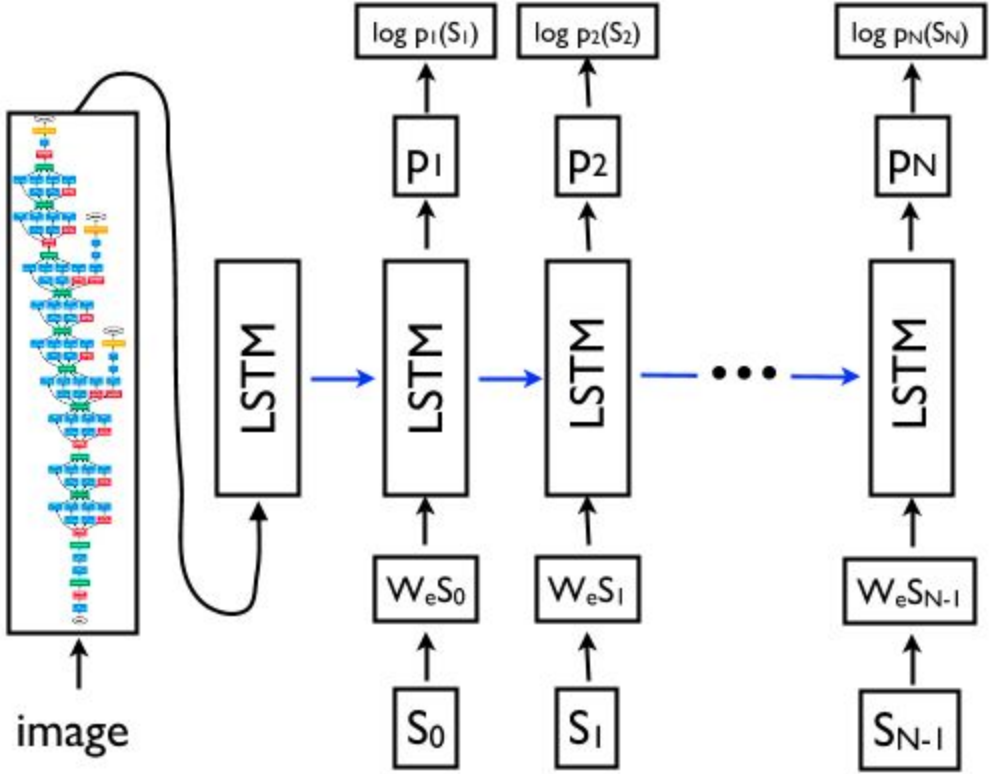
Model

- CNN for images
- RNN for language modeling
- Backpropagation for training

Data

Dataset name	size		
	train	valid.	test
Pascal VOC 2008 [6]	-	-	1000
Flickr8k [26]	6000	1000	1000
Flickr30k [33]	28000	1000	1000
MSCOCO [20]	82783	40504	40775
SBU [24]	1M	-	-

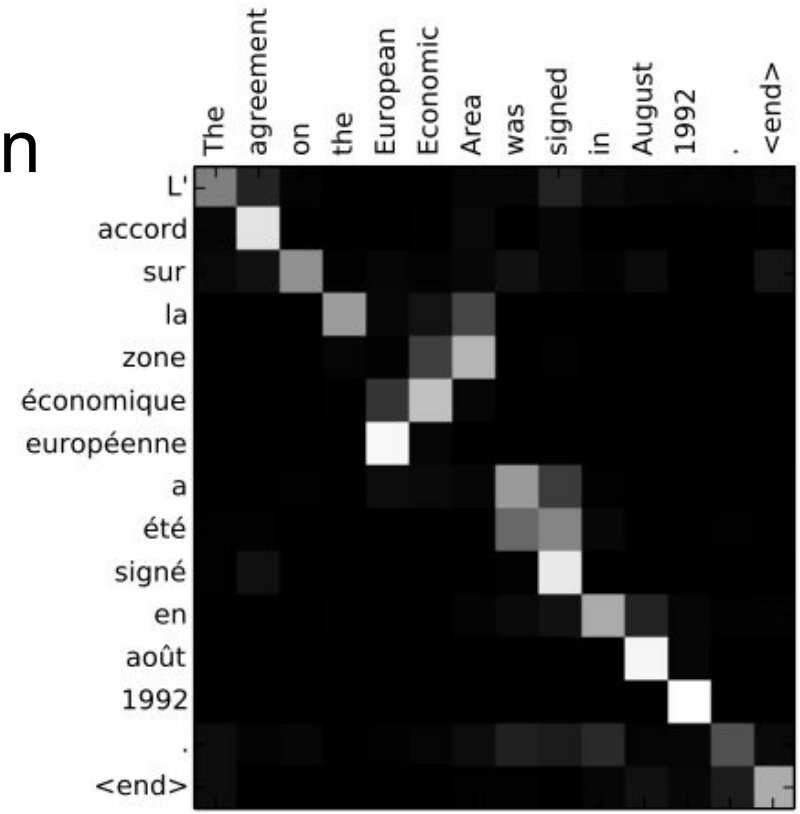
An End-to-End approach:



O Vinyals - 2015

Generated sentences

Attention models in Translation

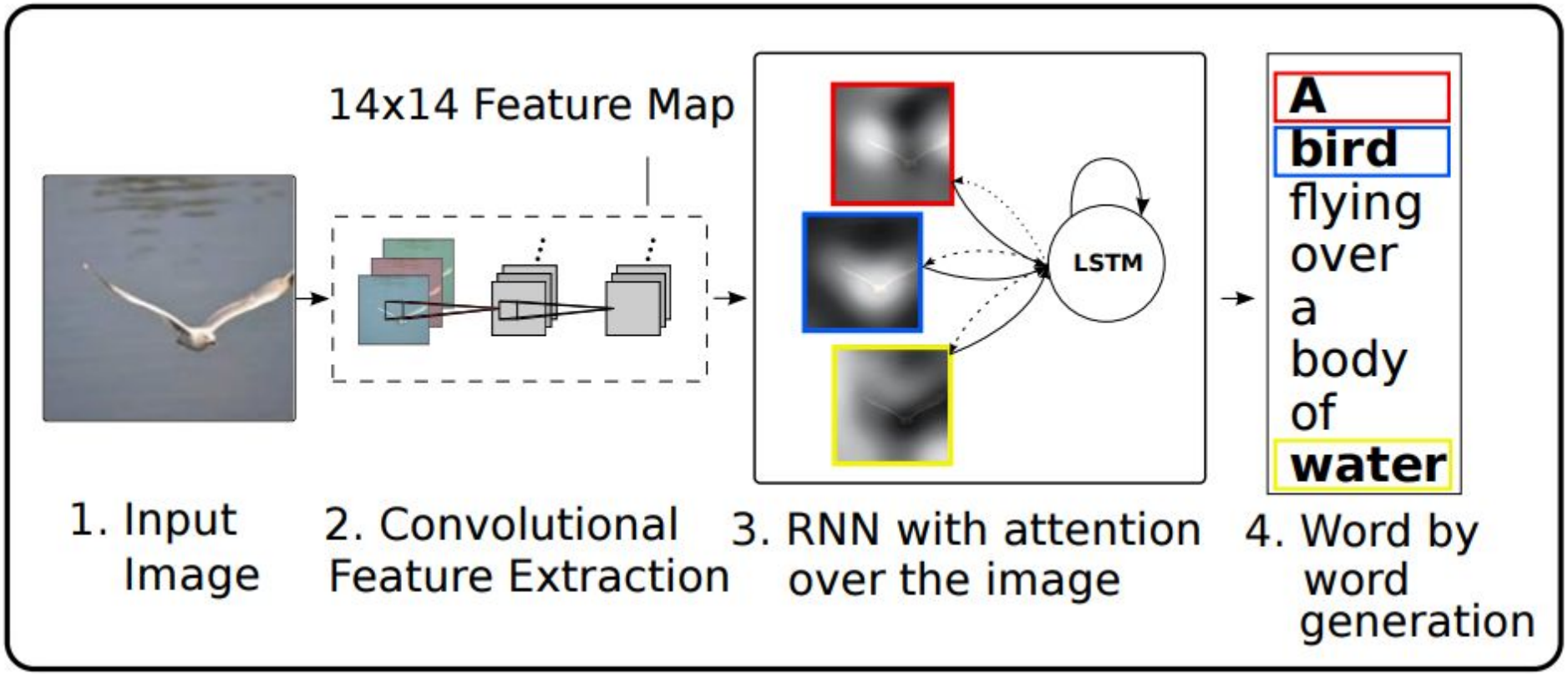


(a)

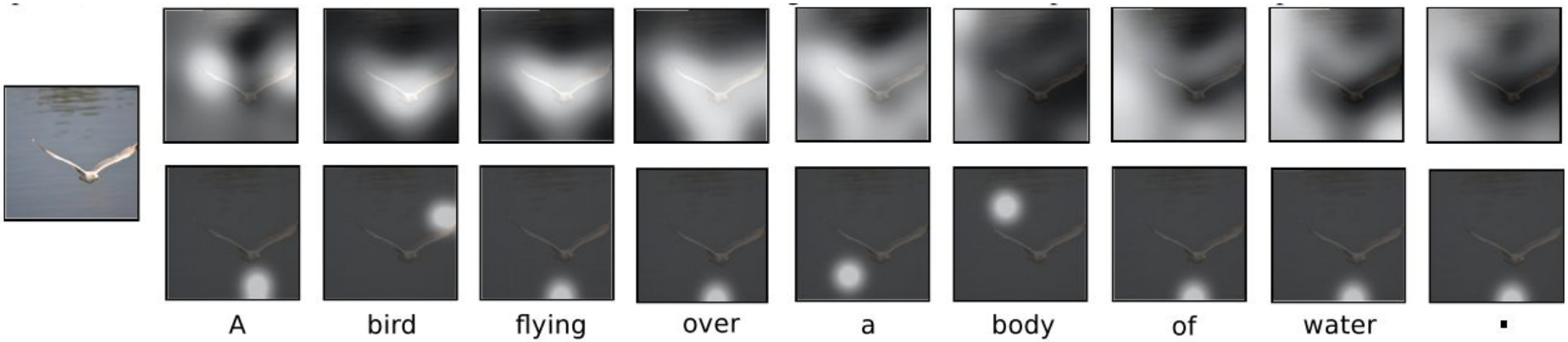
D. Bahdanau 2014

Attention models in Image Captioning

K. Xu, 2016



Visual Alignments



Generated phrases

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.