

Assignment #6

Attention(Image Captioning)

Due on Dec 28, 11:59 pm

Overview

- In the assignment #5, we implemented the image captioning tasks.
- In this assignment, you will add an attention module.
- You are free to use pre-trained models like ResNet or LSTM as your backbone structure.

Image Captioning

Image captioning is an interdisciplinary research problem that stands between computer vision and natural language processing.

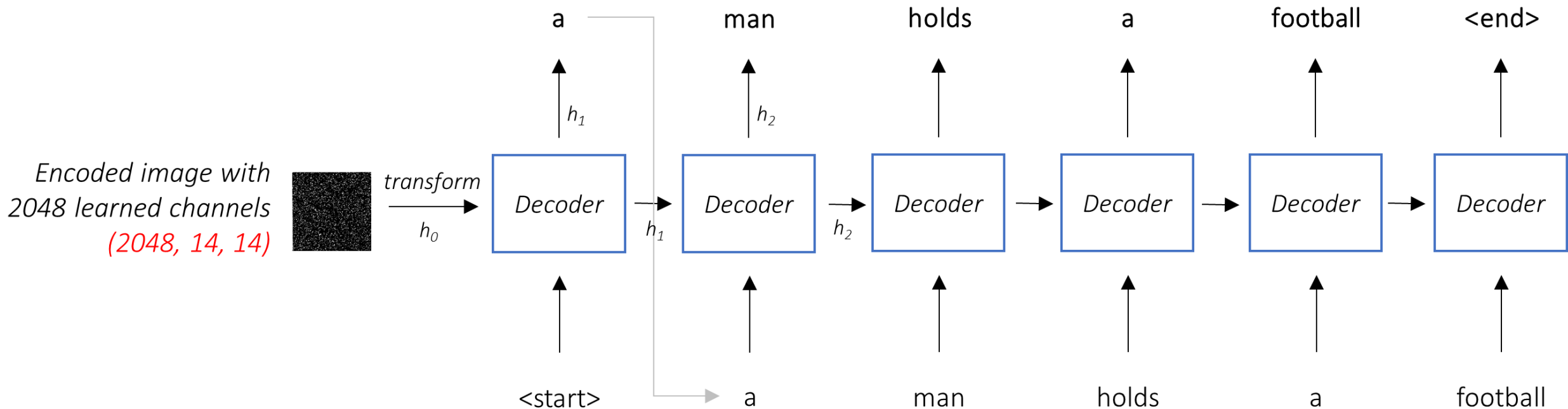
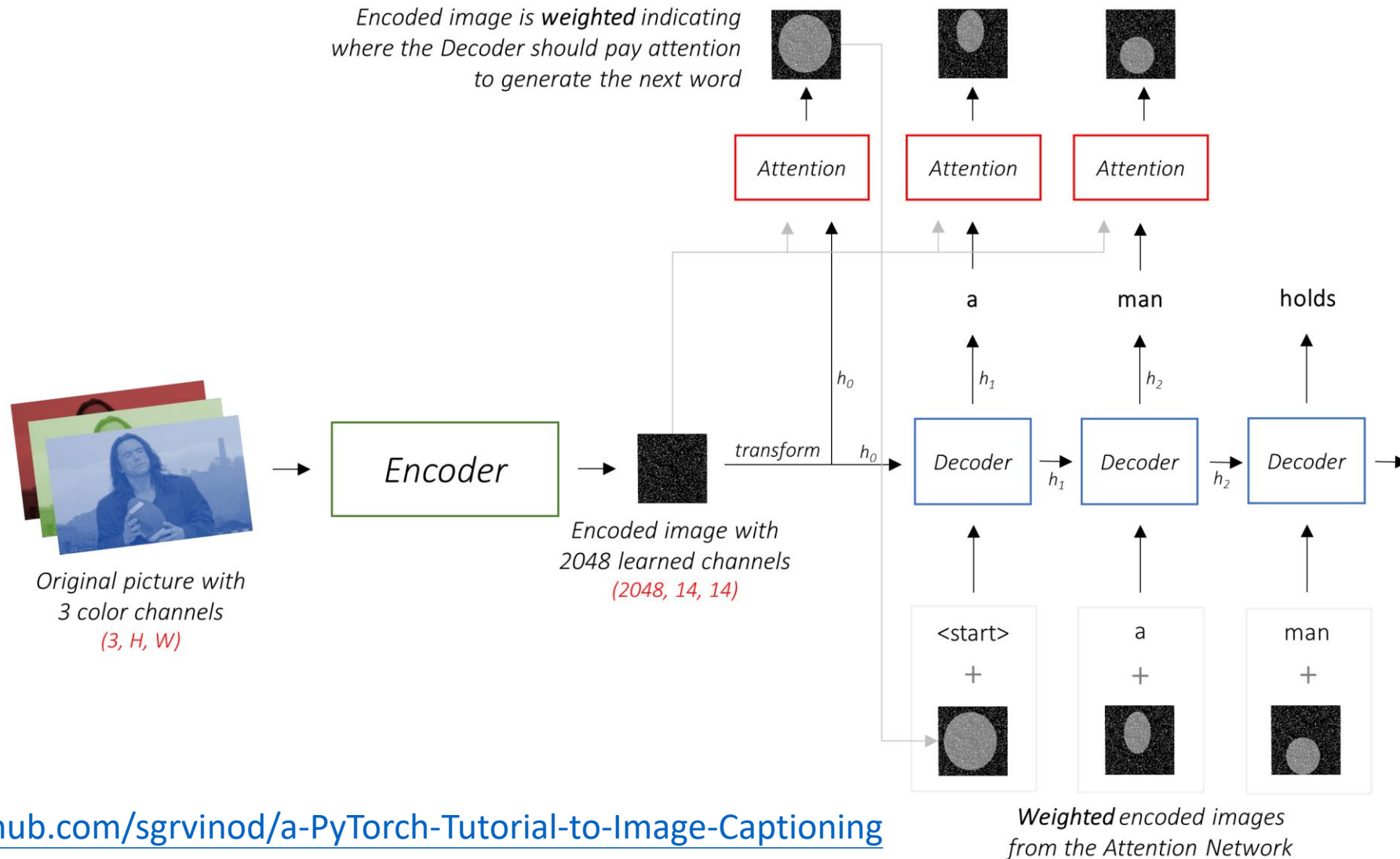


Image Captioning with Attention



ref: <https://github.com/sgrvinod/a-PyTorch-Tutorial-to-Image-Captioning>

Flickr8k Dataset

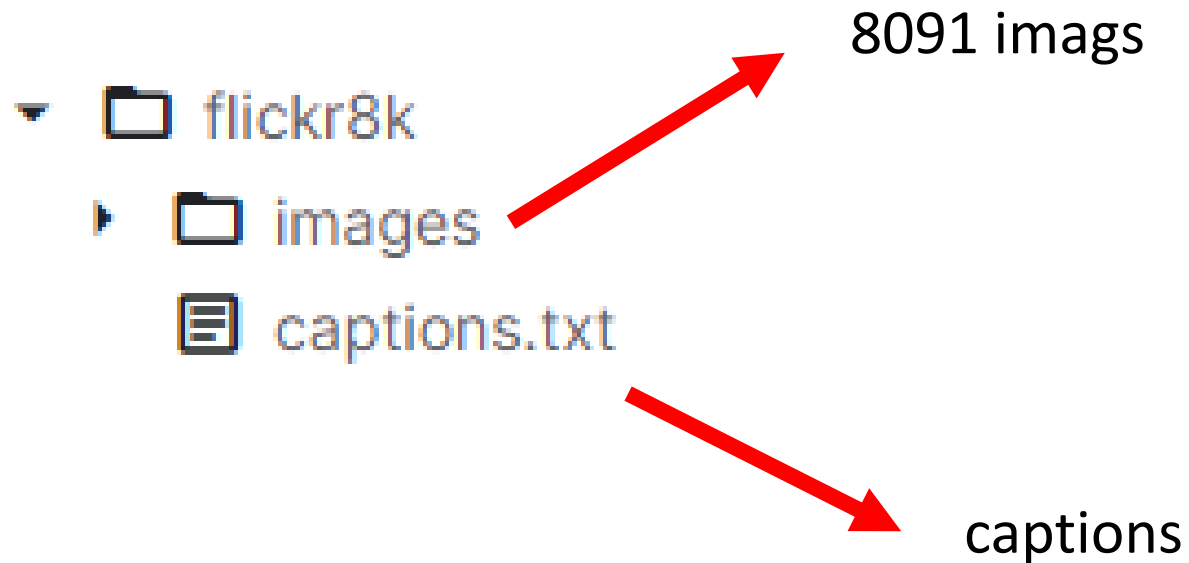
- [Flickr8k-Images-Captions](#)
- Collected by Alexander Mamaev.
- Sentence-based image description and search
- Consisting of 8,091 images that are each paired with five different captions



A child in a pink dress is climbing up a set of stairs in an entry way .

Assignment #6 Dataset

We use the same dataset as Assignment #5.



Your task

- We have code skeleton for you guys.
- https://colab.research.google.com/drive/15zLahwl2Ud8TAJlaUyh79SFee3hEb45_?usp=sharing
- Design an image captioning model with **attention module**.
- To get a high accuracy, you'll need to experiment with different filter sizes, different number of layers, and other design principles discussed in class to figure out a network architecture that works best.
- You'll also need to try data augmentation, dropout, batch normalization as well as different optimizers and other tricks to boost performance.

Things you cannot do

- You cannot copy trained models from others.
- Directly copy and paste the code from the internet(It won't work).

Any violation will result in no points!

Submission

Submit your code + report to the CU.

- The report include:
 - **Additional works.**
 - What you have experimented for performance improvements. This could include experiments that work and those that do not work.
 - Other findings if any
- Grading on report will mostly depend on the report quality. Please show your effort.

Grading

- Code: 85 points
- Report: 15 points report

Some Tips For you

- Image Captioning Tutorial (Assignment #5):
<https://www.youtube.com/watch?v=y2BaTt1fxJU>
- Image Captioning with Attention Tutorial (Assignment #6):
<https://github.com/sgrvinod/a-PyTorch-Tutorial-to-Image-Captioning>