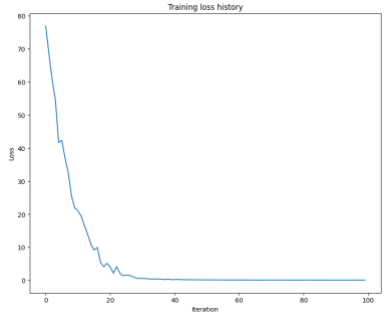
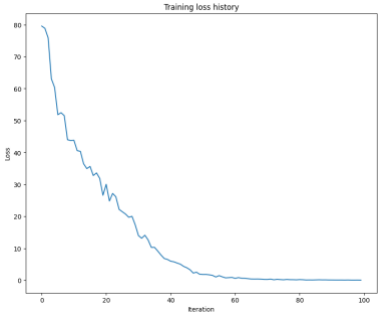
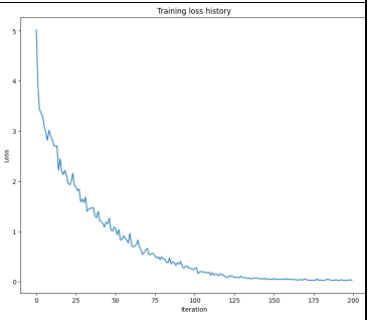
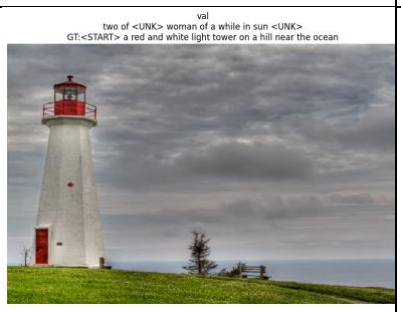
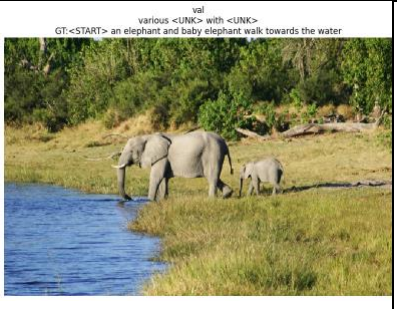
Computer Vision Homework#3

2022094093 Kim Dohoon, Dept. of Data Science

All codes are uploaded at https://github.com/kdh-yu/ComputerVision/tree/main/HW_3

1. Code implementation and inference

RNN	LSTM	Transformer
		
Final Loss 0.08207246780395508	Final Loss 0.08085013696488379	Final Loss 0.02125324
		
		
		

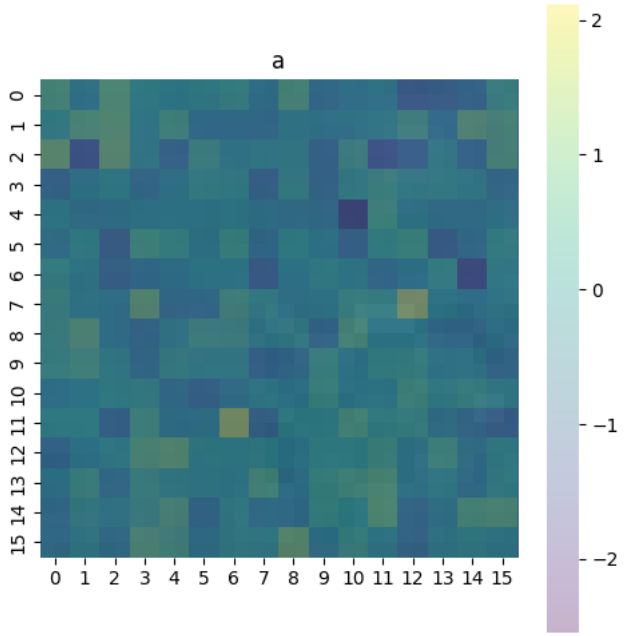
2. Implementation of evaluation matric (CIDEr score)

All scores are evaluated by models that are overfitted in small data, using valid data. Codes are implemented in "CIDEr.py"

RNN :	CIDEr score on valid data : 9.33
LSTM :	CIDEr score on valid data : 8.85
Transformer :	CIDEr score on valid data : 11.17

3. Visualize the attention map.

---- couldn't



Tried to implement, but all attention maps were exactly same. I think attention scores are not returned correctly. If possible in future, I want to fix it and implement the whole process again.

4. NICE data:

19	2022-—93	06-13 21:00:28	36.4	18.32	9.08	5	9.37	5.58

Submitted.