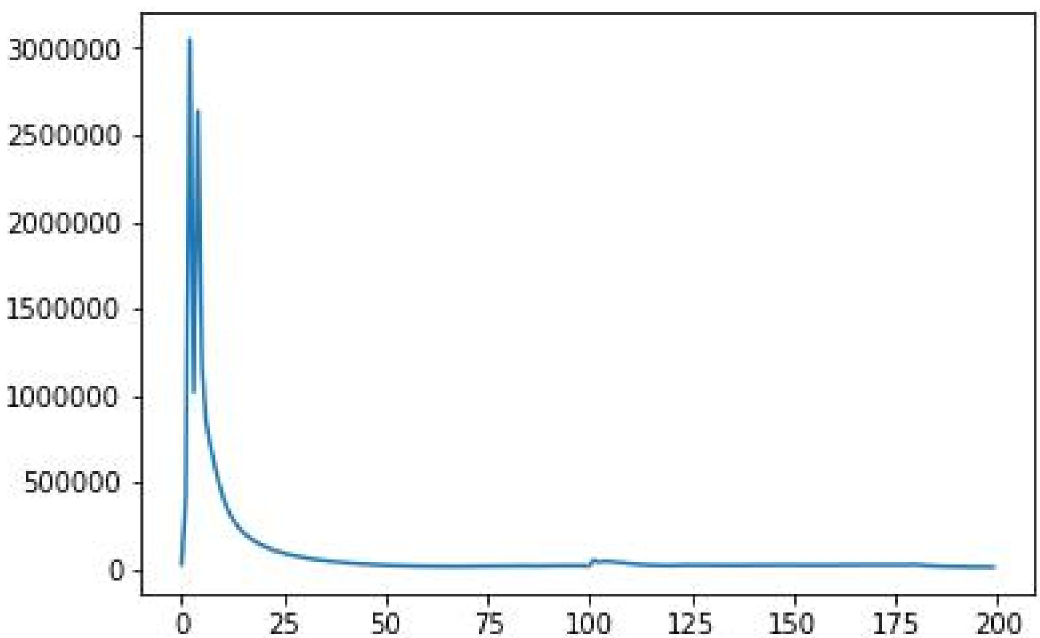
Yuanhang Luo

1.5. validation accuracy:

	Finetune	freeze
On pretrain	0.3529	0.3529
After training	0.9281	0.9477

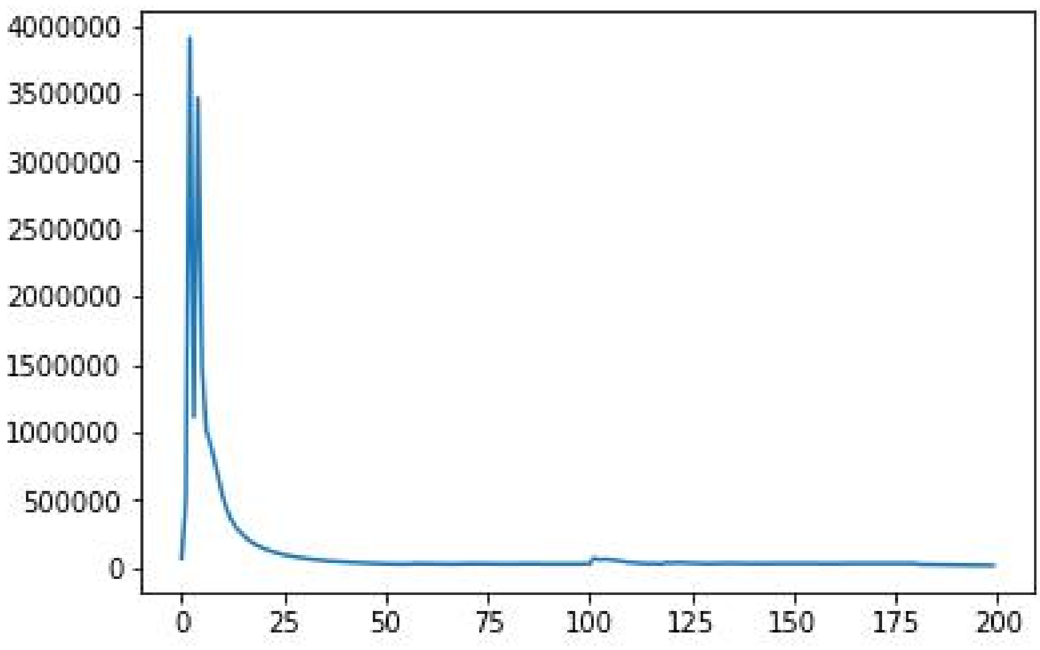
2.4

Composition VII + Tubingen:



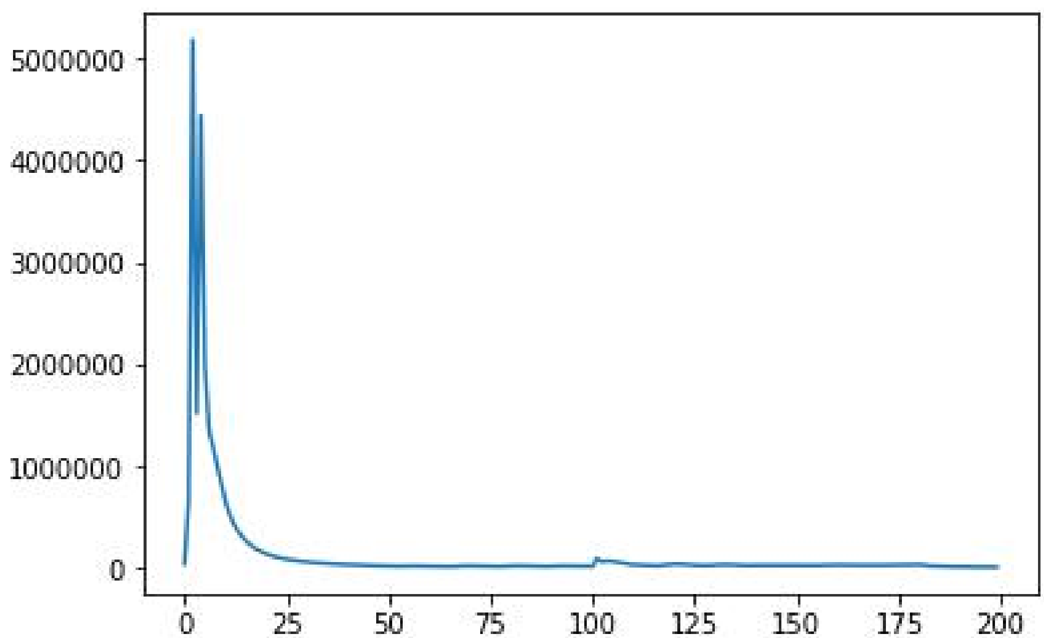
Yuanhang Luo

Scream + Tubingen:



Yuanhang Luo

Starry Night + Tubingen:



Yuanhang Luo

3.2 Let $[1-h_{t_i}^2]$ denote a vector with the i-th element equals to $1-h_{t_i}^2$.

* means element wise product.

$\frac{\partial L}{\partial b_i}=\frac{\partial L}{\partial h_{t_i}}\frac{\partial h_{t_i}}{\partial b_i}=\frac{\partial L}{\partial h_{t_i}}\left(1-h_{t_i}^2\right).$ So, $\frac{\partial L}{\partial b}=\frac{\partial L}{\partial h_t} * \ [1-h_{t_i}^2]$

$\frac{\partial L}{\partial x_{t_i}}=\frac{\partial L}{\partial h_{t_i}}\frac{\partial h_{t_i}}{\partial x_{t_i}}=\frac{\partial L}{\partial h_{t_i}}\left(1-h_{t_i}^2\right)W_{x_i}$ So, $\frac{\partial L}{\partial x_t}=W_x^T(\left[1-h_{t_i}^2\right] * \frac{\partial L}{\partial h_t})$

$\frac{\partial L}{\partial h_{t-1_i}}=\frac{\partial L}{\partial h_{t_i}}\frac{\partial h_{t_i}}{\partial h_{t-1_i}}=\frac{\partial L}{\partial h_{t_i}}\left(1-h_{t_i}^2\right)W_{h_i}$ So, $\frac{\partial L}{\partial h_{t-1}}=W_h^T(\left[1-h_{t_i}^2\right] * \frac{\partial L}{\partial h_t})$

$\frac{\partial L}{\partial W_{x_i}}=\frac{\partial L}{\partial h_{t_i}}\frac{\partial h_{t_i}}{\partial W_{x_{ij}}}=\frac{\partial L}{\partial h_{t_i}}\left[1-h_{t_i}^2\right]x_{t_j}$ So, $\frac{\partial L}{\partial W_x}=(\frac{\partial L}{\partial h_t} * \left[1-h_{t_i}^2\right])x_t^T$

$\frac{\partial L}{\partial W_{h_i}}=\frac{\partial L}{\partial h_{t_i}}\frac{\partial h_{t_i}}{\partial W_{h_{ij}}}=\frac{\partial L}{\partial h_{t_i}}\left[1-h_{t_i}^2\right]h_{t-1_j}$ So, $\frac{\partial L}{\partial W_h}=(\frac{\partial L}{\partial h_t} * \left[1-h_{t_i}^2\right])h_{t-1}^T$

3.4 For each timestep:

Recursively define: $\frac{\partial L}{\partial \tilde{h}_t}=\frac{\partial L}{\partial h_t}+\frac{\partial L}{\partial h_{t-1}}.$

$\frac{\partial L}{\partial h_{t-1}}=W_h^T\left(\left[1-h_{t_i}^2\right] * \frac{\partial L}{\partial \tilde{h}_t}\right).$

So at timestep t, $\frac{\partial L}{\partial \tilde{h}_t} = \sum_{a=0}^{T-t}((w_h^T)^{T-t-a}\frac{\partial L}{\partial h_{T-a}}\prod_{c=a}^{T-t-1}[1-h_{(T-c)_i}^2])$, Π is element wise product here.

So $\frac{\partial L}{\partial b}=\sum_{t=1}^{T}\frac{\partial L}{\partial \tilde{h}_t} * \ [1-h_{t_i}^2]$

$\frac{\partial L}{\partial W_h}=\sum_{t=1}^{T}(\frac{\partial L}{\partial \tilde{h}_t} * \left[1-h_{t_i}^2\right])h_{t-1}^T$

$\frac{\partial L}{\partial W_x}=\sum_{t=1}^{T}(\frac{\partial L}{\partial \tilde{h}_t} * \left[1-h_{t_i}^2\right])x_t^T$

$\frac{\partial L}{\partial h_0}=W_h^T(\left[1-h_{1_i}^2\right] * \frac{\partial L}{\partial \tilde{h}_1})$

$\frac{\partial L}{\partial x_t}=W_x^T(\left[1-h_{t_i}^2\right] * \frac{\partial L}{\partial \tilde{h}_t})$

Yuanhang Luo

4.2

$\frac{\partial L}{\partial c_t}$ *is the aggregate gradient for* c_t.

$$\frac{\partial L}{\partial c_{t-1}} = \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial c_{t-1}} = \frac{\partial L}{\partial c_t} * f_t$$

$$\frac{\partial L}{\partial b^o} = \frac{\partial L}{\partial h_t}\frac{\partial h_t}{\partial O_t}\frac{\partial O_t}{\partial b^o} = \frac{\partial L}{\partial h_t} * \tanh(C_t) * o_t * (1 - o_t)$$

$$\frac{\partial L}{\partial b^c} = \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial \widetilde{c_t}}\frac{\partial \widetilde{c_t}}{\partial b^c} = \frac{\partial L}{\partial c_t} * i_t * \left(1 - \widetilde{c_t^2}\right)$$

$$\frac{\partial L}{\partial b^i} = \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial i_t}\frac{\partial i_t}{\partial b^i} = \frac{\partial L}{\partial c_t} * \widetilde{c_t} * i_t * (1 - i_t)$$

$$\frac{\partial L}{\partial b^f} = \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial f_t}\frac{\partial f_t}{\partial b^f} = \frac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t)$$

$$\frac{\partial L}{\partial W_x^o} = \frac{\partial L}{\partial h_t}\frac{\partial h_t}{\partial O_t}\frac{\partial O_t}{\partial W_x^o} = (\frac{\partial L}{\partial h_t} * \tanh(C_t) * o_t * (1 - o_t))x_t^T$$

$$\frac{\partial L}{\partial W_x^c} = \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial \widetilde{c_t}}\frac{\partial \widetilde{c_t}}{\partial W_x^c} = (\frac{\partial L}{\partial c_t} * i_t * \left(1 - \widetilde{c_t^2}\right))x_t^T$$

$$\frac{\partial L}{\partial W_x^i} = \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial i_t}\frac{\partial i_t}{\partial W_x^i} = (\frac{\partial L}{\partial c_t} * \widetilde{c_t} * i_t * (1 - i_t))x_t^T$$

$$\frac{\partial L}{\partial W_x^f} = \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial f_t}\frac{\partial f_t}{\partial W_x^f} = (\frac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t))x_t^T$$

$$\frac{\partial L}{\partial W_h^o} = \frac{\partial L}{\partial h_t}\frac{\partial h_t}{\partial O_t}\frac{\partial O_t}{\partial W_h^o} = (\frac{\partial L}{\partial h_t} * \tanh(C_t) * o_t * (1 - o_t))h_{t-1}^T$$

$$\frac{\partial L}{\partial W_h^c} = \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial \widetilde{c_t}}\frac{\partial \widetilde{c_t}}{\partial W_h^c} = (\frac{\partial L}{\partial c_t} * i_t * \left(1 - \widetilde{c_t^2}\right))h_{t-1}^T$$

$$\frac{\partial L}{\partial W_h^i} = \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial i_t}\frac{\partial i_t}{\partial W_h^i} = (\frac{\partial L}{\partial c_t} * \widetilde{c_t} * i_t * (1 - i_t))h_{t-1}^T$$

$$\frac{\partial L}{\partial W_h^f} = \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial f_t}\frac{\partial f_t}{\partial W_h^f} = (\frac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t))h_{t-1}^T$$

$$\frac{\partial L}{\partial x_t} = \frac{\partial L}{\partial h_t}\frac{\partial h_t}{\partial O_t}\frac{\partial O_t}{\partial x_t} + \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial \widetilde{c_t}}\frac{\partial \widetilde{c_t}}{\partial x_t} + \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial i_t}\frac{\partial i_t}{\partial x_t} + \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial f_t}\frac{\partial f_t}{\partial x_t}$$

$$= (W_x^o)^T\left(\tfrac{\partial L}{\partial h_t} * \tanh(C_t) * o_t * (1 - o_t)\right) + (W_x^c)^T\left(\tfrac{\partial L}{\partial c_t} * i_t * \left(1 - \widetilde{c_t^2}\right)\right) +$$

$$\left(W_x^i\right)^T(\tfrac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t)) + \left(W_x^f\right)^T(\tfrac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t))$$

$$\frac{\partial L}{\partial h_{t-1}} = \frac{\partial L}{\partial h_t}\frac{\partial h_t}{\partial O_t}\frac{\partial O_t}{\partial h_{t-1}} + \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial \widetilde{c_t}}\frac{\partial \widetilde{c_t}}{\partial h_{t-1}} + \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial i_t}\frac{\partial i_t}{\partial h_{t-1}} + \frac{\partial L}{\partial c_t}\frac{\partial c_t}{\partial f_t}\frac{\partial f_t}{\partial h_{t-1}}$$

$$= \left(W_h^o\right)^T \left(\frac{\partial L}{\partial h_t} * \tanh(C_t) * o_t * (1 - o_t)\right) + \left(W_h^c\right)^T \left(\frac{\partial L}{\partial c_t} * i_t * \left(1 - \widetilde{c_t^2}\right)\right) +$$

$$\left(W_h^i\right)^T \left(\frac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t)\right) + \left(W_h^f\right)^T \left(\frac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t)\right)$$

4.4

For each timestep:

Recursively define: $\frac{\partial L}{\partial \tilde{h}_t} = \frac{\partial L}{\partial h_t} + \frac{\partial L}{\partial h_{t-1}}$

$$\frac{\partial L}{\partial h_{t-1}} = \left(W_h^o\right)^T \left(\frac{\partial L}{\partial \tilde{h}_t} * \tanh(C_t) * o_t * (1 - o_t)\right) + \left(W_h^c\right)^T \left(\frac{\partial L}{\partial c_t} * i_t * \left(1 - \widetilde{c_t^2}\right)\right) +$$

$$\left(W_h^i\right)^T \left(\frac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t)\right) + \left(W_h^f\right)^T \left(\frac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t)\right)$$

$\frac{\partial L}{\partial h_{t-1}} = W_h^T \left(\left[1 - h_{t_i}^2\right] * \frac{\partial L}{\partial \tilde{h}_t}\right).$

So at timestep t, $\quad \frac{\partial L}{\partial \tilde{h}_t} \; = \sum_{a=0}^{T-t} \left(\left(w_h^T\right)^{T-t-a} \frac{\partial L}{\partial h_{T-a}} \prod_{c=a}^{T-t-1} \left[1 - h_{(T-c)_i}^2\right]\right)$

So at timestep t, $\quad \frac{\partial L}{\partial \tilde{h}_t} \; = \sum_{a=0}^{T-t} \left(\left(\left(W_h^o\right)^T\right)^{T-t-a} \frac{\partial L}{\partial h_{T-a}} \prod_{k=a}^{T-t-1} \tanh(C_{T-k}) * o_{T-k} * (1 - o_{T-k})\right)$, Π is element wise product here.

$$\frac{\partial L}{\partial b^o} = \sum_{t=1}^{T} \frac{\partial L}{\partial \tilde{h}_t} * \tanh(C_t) * o_t * (1 - o_t)$$

$$\frac{\partial L}{\partial b^c} = \sum_{t=1}^{T} \frac{\partial L}{\partial c_t} * i_t * \left(1 - \widetilde{c_t^2}\right)$$

$$\frac{\partial L}{\partial b^i} = \sum_{t=1}^{T} \frac{\partial L}{\partial c_t} * \widetilde{c_t} * i_t * (1 - i_t)$$

Yuanhang Luo

$$\frac{\partial L}{\partial b^f} = \sum_{t=1}^{T} \frac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t)$$

$$\frac{\partial L}{\partial W_x^o} = \sum_{t=1}^{T} \left(\frac{\partial L}{\partial \tilde{h}_t} * \tanh(C_t) * o_t * (1 - o_t)\right) x_t^T$$

$$\frac{\partial L}{\partial W_x^c} = \sum_{t=1}^{T} \left(\frac{\partial L}{\partial c_t} * i_t * \left(1 - \widetilde{c_t^2}\right)\right) x_t^T$$

$$\frac{\partial L}{\partial W_x^i} = \sum_{t=1}^{T} \left(\frac{\partial L}{\partial c_t} * \widetilde{c_t} * i_t * (1 - i_t)\right) x_t^T$$

$$\frac{\partial L}{\partial W_x^f} = \sum_{t=1}^{T} \left(\frac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t)\right) x_t^T$$

$$\frac{\partial L}{\partial W_h^o} = \sum_{t=1}^{T} \left(\frac{\partial L}{\partial \tilde{h}_t} * \tanh(C_t) * o_t * (1 - o_t)\right) h_{t-1}^T$$

$$\frac{\partial L}{\partial W_h^c} = \sum_{t=1}^{T} \left(\frac{\partial L}{\partial c_t} * i_t * \left(1 - \widetilde{c_t^2}\right)\right) h_{t-1}^T$$

$$\frac{\partial L}{\partial W_h^i} = \sum_{t=1}^{T} \left(\frac{\partial L}{\partial c_t} * \widetilde{c_t} * i_t * (1 - i_t)\right) h_{t-1}^T$$

$$\frac{\partial L}{\partial W_h^f} = \sum_{t=1}^{T} \left(\frac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t)\right) h_{t-1}^T$$

$$\frac{\partial L}{\partial x_t} = (W_x^o)^T \left(\frac{\partial L}{\partial \tilde{h}_t} * \tanh(C_t) * o_t * (1 - o_t)\right) + (W_x^c)^T \left(\frac{\partial L}{\partial c_t} * i_t * \left(1 - \widetilde{c_t^2}\right)\right) +$$

$$\left(W_x^i\right)^T \left(\frac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t)\right) + \left(W_x^f\right)^T \left(\frac{\partial L}{\partial c_t} * c_{t-1} * f_t * (1 - f_t)\right)$$

$$\frac{\partial L}{\partial h_0} = \left(W_h^o\right)^T \left(\frac{\partial L}{\partial \tilde{h}_1} * \tanh(C_1) * o_1 * (1 - o_1)\right) + \left(W_h^c\right)^T \left(\frac{\partial L}{\partial c_1} * i_1 * \left(1 - \widetilde{c_1^2}\right)\right) +$$

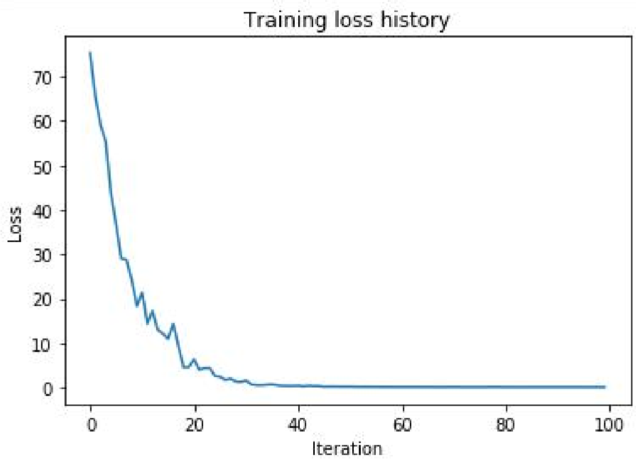
$$\left(W_h^i\right)^T \left(\frac{\partial L}{\partial c_1} * c_0 * f_1 * (1 - f_1)\right) + \left(W_h^f\right)^T \left(\frac{\partial L}{\partial c_1} * c_0 * f_1 * (1 - f_1)\right)$$

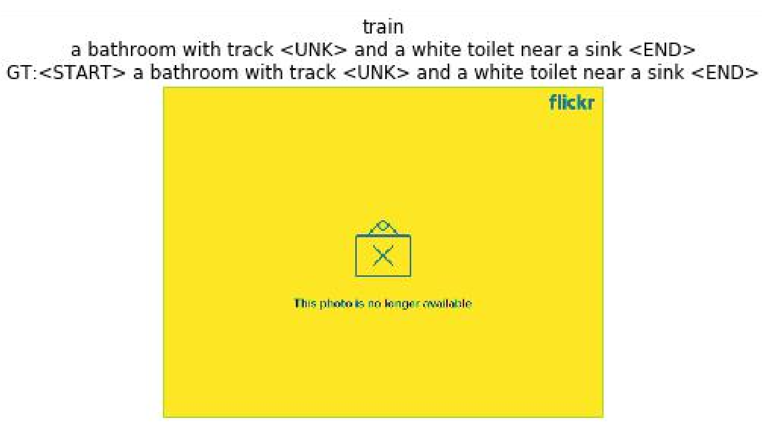
Yuanhang Luo

5.4.

In the prediction, I am ***not*** including <START> inside the predicted caption, because it is only initial state. And according to the code rnn.py, “The first element of captions should be the first sampled word, not the <START> token.”

RNN:





Yuanhang Luo

val
a red <UNK> with a <UNK> <UNK> in the the <END>
GT:<START> an open luggage bag near a laptop <END>

val
a number of the mirror <END>
GT:<START> the blue train engine <UNK> black smoke as it <UNK> down the track <END>

Yuanhang Luo

LSTM:

train
the <UNK> of an empty restaurant decorated in wood and leather <END>
GT:<START> the <UNK> of an empty restaurant decorated in wood and leather <END>

train
a kitchen is <UNK> lit by the light above the stove <END>
GT:<START> a kitchen is <UNK> lit by the light above the stove <END>

Yuanhang Luo

val
a female is <UNK> in a tree <END>
GT:<START> a lady with an apple <UNK> towards her <UNK> <END>

val
a city street with traffic and tall buildings <END>
GT:<START> a yellow school bus driving down a street with a red car <UNK> behind it <END>

5.5

I got a better model in image_captioning_better.py. With BLEU validation scores higher than 0.25 (in 5.4, the BLEU validation scores is around 0.15).

6. test performances:

    1. 92.14%
    2. 94.76%
    3. 94.84%
    4. 95.88%
    5. 92.98%