**Medical Image Report Generation**

**CS286-HW1**

1. **Introduction**

Radiologists routinely examine medical images such as XRay, CT, or MRI and write reports summarizing their descriptive findings and conclusive impressions. A computer-aided radiology report generation system can lighten the workload for radiologists considerably and assist them in decision making

1. **Requirements**

We aim to generate a realistic medical diagnostic report from given chest x-ray images. They are amount of existing works for image captioning in nature images. This time, we want to apply one of the most famous image captioning algorithms [1] into the domain specific image/text datasets, which is based on CNN/RNN framework. In this homework, you can realize how CNNs extract visual semantic feature and how RNN decode the visual image into medical report. There is a code example for nature images captioning. You can modify this code and adapt it into this chest x-ray datasets.

* 1. Related materials

Chest x-ray dataset, you can download from this url:

<http://10.19.124.26:8000/d/0f886e6deb3d460cb355/>

Code example, you can refer to: <https://www.tensorflow.org/tutorials/text/image_captioning>

* 1. Modify the dataloader to adapt this chest x-ray datasets, describe the standard image preprocessing techniques you have used in you code, explain why you use those techniques? (30 points)
  2. Running the whole code in this dataset and get the final performance. Eventually. we take CIDEr and BLUE metrics into consideration to evaluate the quality of generated reports. You need to show training loss curve, training logs and the highest performance you got. Besides, you need to replace the CNN backbone InceptionV3 with other famous backbones, like VGG-16, ResNet-50 and compare them with the default backbone. Show the training curve and accuracy curve (30 points)
  3. Describe what loss function you have used and explain the meaning of it. And explain how to calculate the CIDEr and BLUE metrics when given the generated report and the ground-truth report. (25 points)
  4. Have you met overfitting or underfitting during training? Explain what is overfitting and underfitting and describe how to avoid those. (15 points)

1. **Submission Instrument**
   1. The deadline is 2019.11.03 23:59, please submit your HW1 via git timely. Our late submission policy is that you will get 50% off the grade one day late.
   2. Our gitclassroom have strict submission deadline, you will not push the results after deadline. So please contact TAs in your tutorial session intermediately once you miss the deadline.

Bioinformatics and genomics: 宋柯 [songk@shanghaitech.edu.cn](mailto:songk@shanghaitech.edu.cn)

Imaging: 张晓峰[zhangxf2@shanghaitech.edu.cn](mailto:zhangxf2@shanghaitech.edu.cn)

Molecular Sciences 熊招平[xiongzhp@shanghaitech.edu.cn](mailto:xiongzhp@shanghaitech.edu.cn)

* 1. Please exclude all the datasets and checkpoints in the git. You just need submit the code and a document to present your results.
  2. You should write your name(chinese), student number and email in the jupyter code and documents.

Reference

[1] Xu, Kelvin, et al. "Show, attend and tell: Neural image caption generation with visual attention." *International conference on machine learning*. 2015.

[2] Amount of papers about medical image report generation, you can find here:

<https://github.com/wangleihitcs/Papers/tree/master/medical%20report%20generation>