ECE 228 Final Project:

Radiology Report Generation Using Deep Learning

Team 77: Ruochen Liu, Zixi Wang, Moira Feng

Background

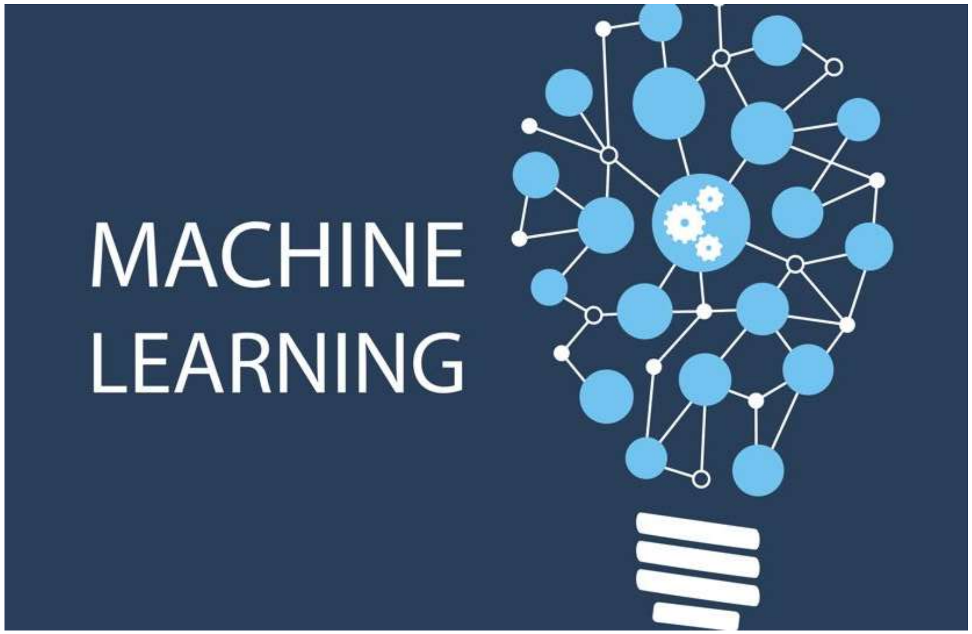
- In diagnostic radiology, diseases identified by findings on radiology images (CT/chest X-ray).
- Large-scale pneumonia outbreak, high-demand of chest X-ray tests.
- Need to accelerate diagnosis.

Background

Therefore, we want to

- Develope a radiology report generator using deep learning.
- Automatically generates descriptive text (findings/impressions) of a chest X-ray.
- Greatly expedite the workflow of radiologists.

How machine learning/deep learning can help solve this problem

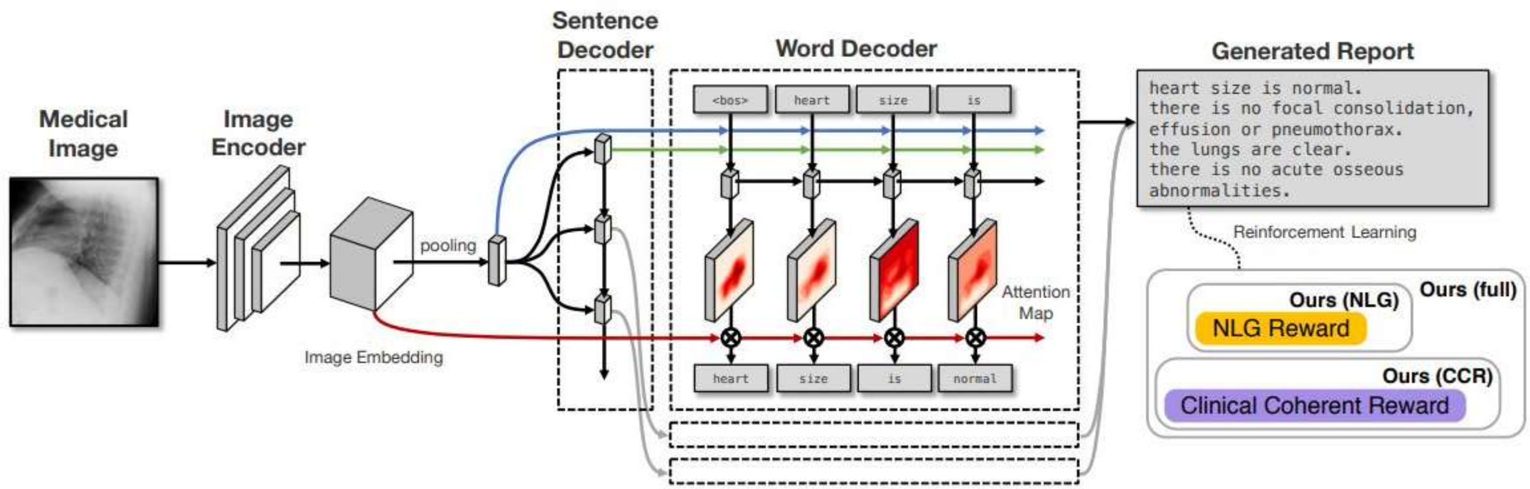


Literature survey

Following recent works presented possible approaches toward radiology report generation.

- Zhang et al. (2018): summarized radiology report's findings with extractive and abstractive techniques to generate impression sections.
- Han et al. (2018): MRI report generation with limited text training data using weak supervision for a Recurrent-GAN and template-based framework.
- Gale et al. (2018): used an RNN to generate template-generated descriptive texts of pelvic X-rays.
- Wang et al. (2018): used a CNN-RNN model with attention to generate descriptive texts on chest X-rays based on sequence decoder losses on the generated report.
- Li et al. (2018): used reinforcement learning to generate chest X-ray reports.

Literature survey



The CNN-RNN-RNN net in paper *Clinically Accurate Chest X-Ray Report Generation*

Dataset: Chest X-ray Images with Text Reports

- Open-I Indiana University Chest X-ray Collection from Indiana Network for Patient Care
- Contains 7466 frontal and lateral Chest X-ray images of 3,999 patients
- Diagnostic reports including “findings” and “impressions” provided by radiologists
- Contains reports for 3852 out of 3999 patients
    - Some “findings” or “impressions” items are empty
    - We currently work with “impressions”

Findings: no finding.

Impression: heart size is upper normal. No edema bandlike left base and lingular opacities. No scarring or atelectasis. No lobar consolidation pleural effusion or pneumothorax.

Model Schematic

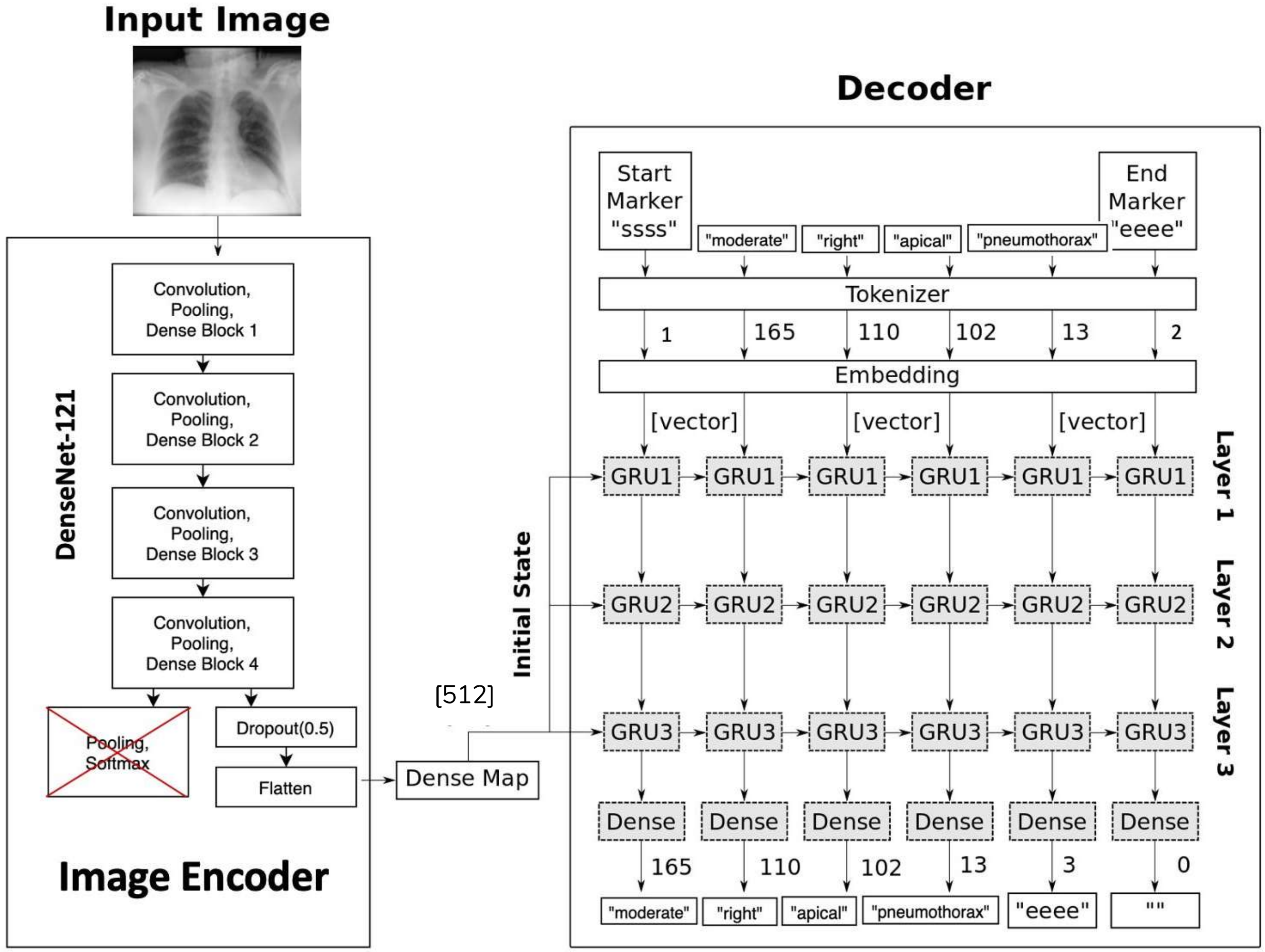
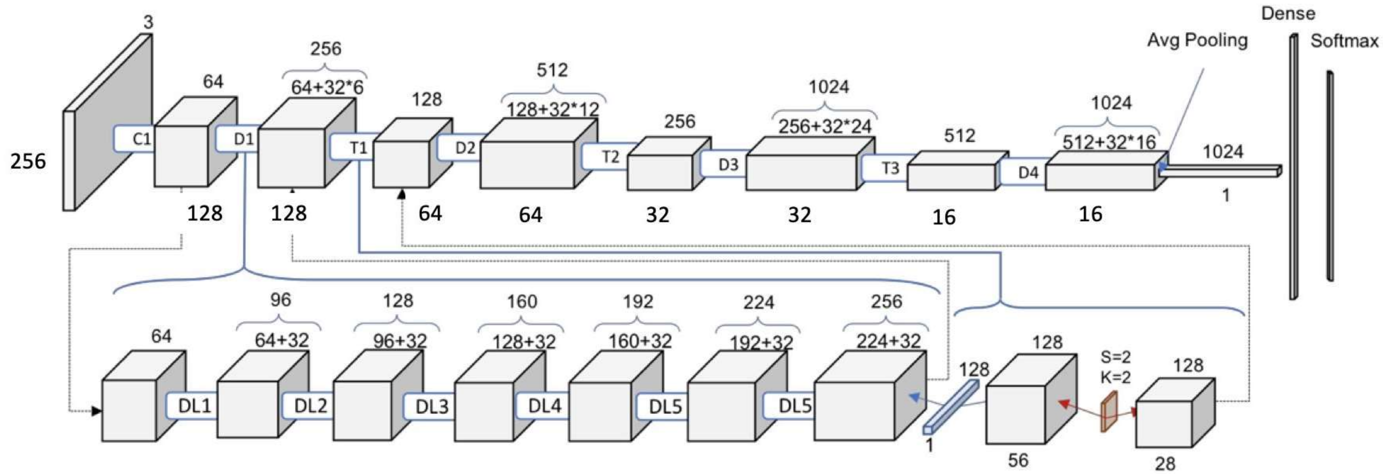


Image partially adapted from [TensorFlow Tutorial #22]

Model Architecture: DenseNet-121

DenseNet-121 Architecture:



Activation: ReLU + tanh

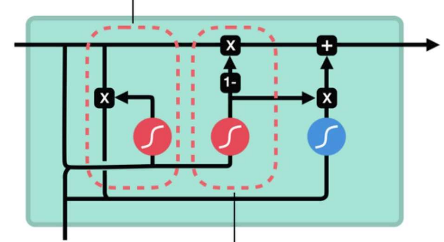
Total parameters: 7,037,504 (DenseNet) + 33,554,944 = 40,592,448

Image partially adapted from [Understanding & Visualizing DenseNet-121](#) by Pablo Ruiz

Model Architecture: RNN & GRU Layer

Recurrent Neural Network (RNN)
- Widely used for processing sequences
- “Information highway” passes hidden state

Gated Recurrent Units (GRU) Architecture:
- Newer generation of RNN
- More compacted version of LSTM cells
- Only update gate & reset gate



Total parameters: 6,060,984

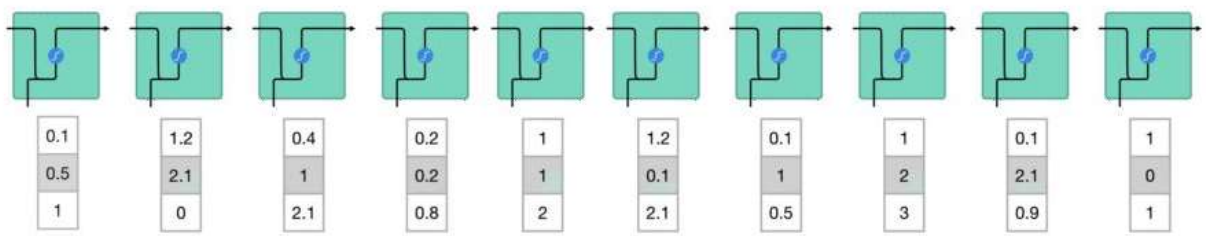
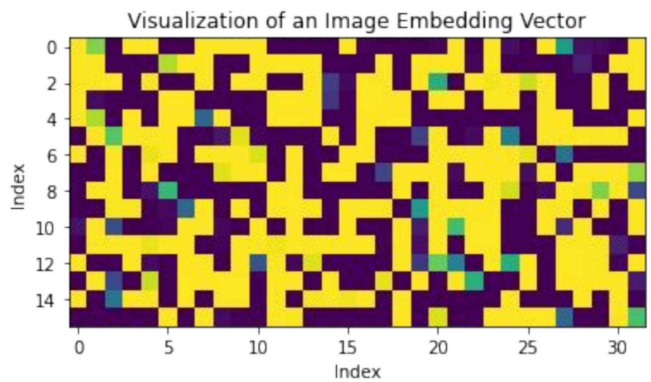


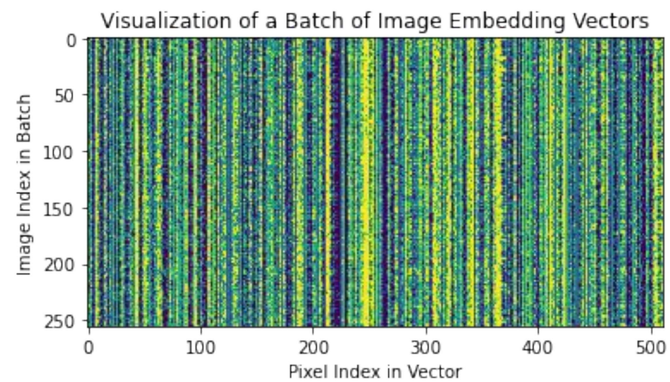
Image partially adapted from [Illustrated Guide to LSTM’s and GRU’s: A step by step explanation] by Michael Phi

Data Processing & Feature Extraction

Chest X-ray image feature extraction:

- Downsized & center-cropped [256x256x3] RGB images
- Encoded to an embedding vector of length 512 through CNN encoder model

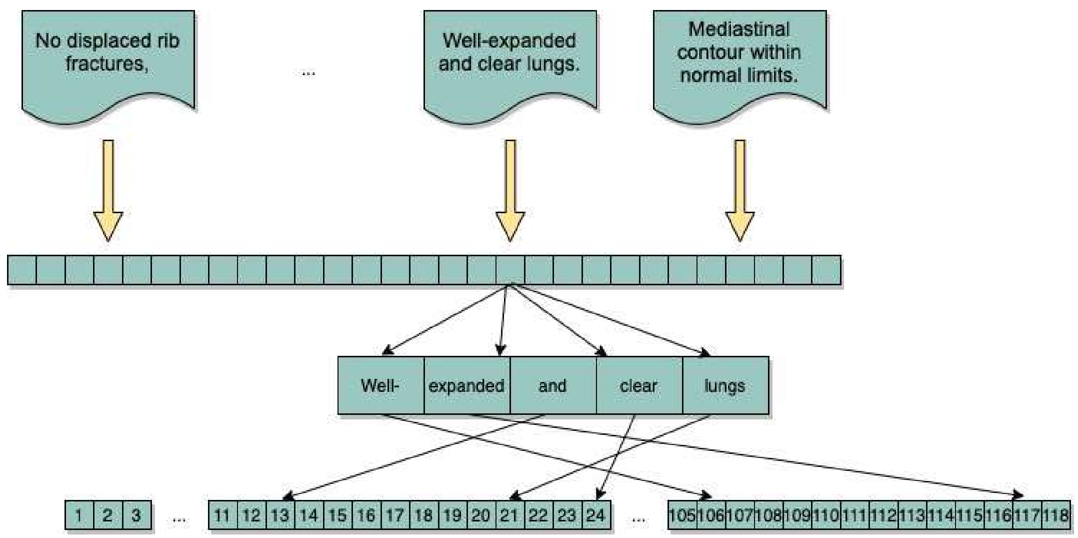




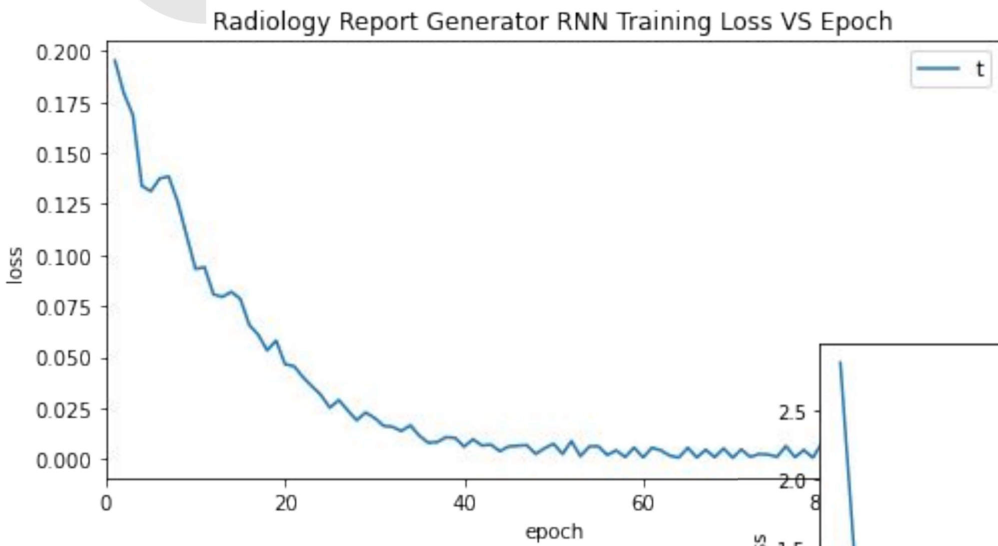
Data Processing & Feature Extraction

Radiology report text image feature extraction:

Mapped string impressions to integer tokens. Word frequency was embedded in the token string.

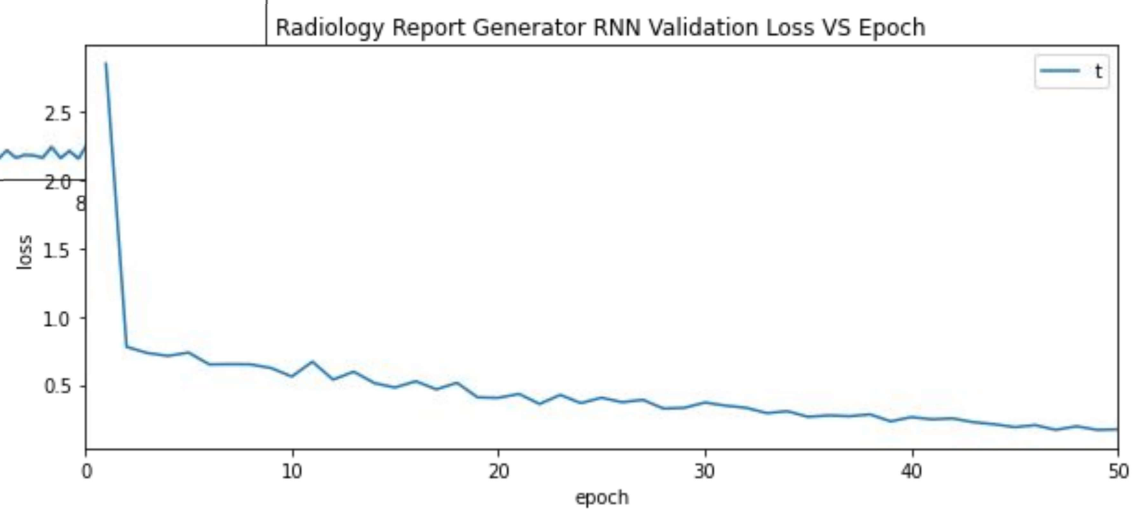


Results: Model Training & Loss Evaluation



Training Parameters

- Optimizer: RMSprop
- Learning Rate: 1E-4
- Loss Function: Sparse categorical cross entropy

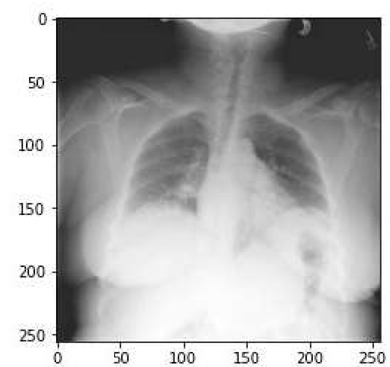


Loss at 50 epochs

- Training loss: ~0.015
- Validation loss: ~0.253

Results & Observations

Compare ground truth and model prediction: training set

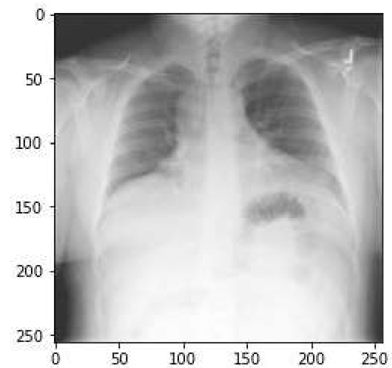


```
Image Index # 51
Patient ID:  53
Predicted caption:
 low lung volumes with minimal bibasilar atelectasis in the right lung base
Ground truth:
Low lung volumes with right basilar atelectasis. Otherwise, no acute cardiopulmonary disease.
```

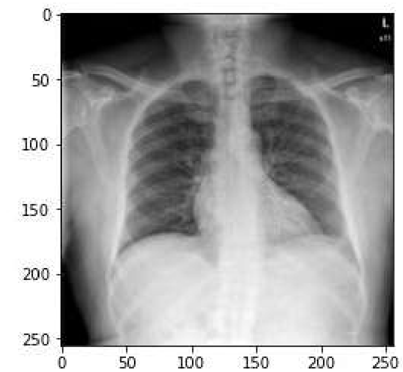
- The predicted caption of a training image should be rephrasing the ground truth
- Possible overfitting If the predicted caption is the same as the ground truth

Results & Observations

Compare ground truth and model prediction: validation set




```
Image Index # 3219
Patient ID:  3414
Predicted caption:
 1 no acute pulmonary abnormality 2 mild cardiomegaly without acute bony abnormality
Ground truth:
1. Low volume study without definite acute process. 2. Mild cardiomegaly.
```




```
Image Index # 3228
Patient ID:  3422
Predicted caption:
 no acute cardiopulmonary abnormality
Ground truth:
Normal chest
```

As the above two examples show, the predicted impressions have the same meaning as ground truths but in different words, just as expected.

Further items to be completed before final report submission

- Tokenize not only words, but also short phrases.
- Use both frontal and lateral images to train the model. (combine view position embedding with the image embedding as the input of the RNN)
- Experiment and compare other models(CNN:VGG16, RNN:LSTM, etc.)
- Experiment parameters(loss function, activation function, image & text feature size, training epochs, parameters for particular RNN models)
- Image pre-process(DCT, corp vs padding, resolution)
- There are two major problems:
  - It can describe the principal part part of the image (such as "there is a heart in the X-Ray image"), but tends to fail describing the details.
  - It takes too much effort to experiment with larger datasets.

References

Clinically Accurate Chest X-Ray Report Generation

Indiana University Chest X-ray Collection Kohli MD, Rosenman M - (2013)

Frequency and Distribution of Chest Radiographic Findings in COVID-19 Positive Patients

Zhang et al. (2018): summarized radiology report's findings with extractive and abstractive techniques to generate impression sections.

Han et al. (2018): MRI report generation with limited text training data using weak supervision for a Recurrent-GAN and template-based framework.

Gale et al. (2018): used an RNN to generate template-generated descriptive texts of pelvic X-rays.

Wang et al. (2018): used a CNN-RNN model with attention to generate descriptive texts on chest X-rays based on sequence decoder losses on the generated report.

Li et al. (2018): used reinforcement learning to generate chest X-ray reports.