Optimizing Sparse View Raw CT Data for Intracranial Hemorrhage Classification

Long Shen, Qiuwen Wei
Department of Biomedical Engineering
Duke University
Durham, NC 27705

Abstract

In this project, we build a machine learning model to classify intracranial hemorrhage based on sparse view raw CT data. We implement two separate trainable physical layers to simulate changing mean effective energy of X-ray at each CT scan and image filtering applied before back projection. We evaluate the performance of four models that consist different physical layers. To explore the effect of sparsity in sinogram on classification accuracy, we trained our final model with eight datasets containing different number of projections. Our results indicate that the Fourier mask layer brings more improvements to our model and a projection number greater than five will guarantee us a validation accuracy above 80%.

1 Introduction

Intracranial hemorrhage is an important and serious public health problem in the world. Statistically [1], it accounts for 10% of strokes happened in United States and strokes is the fifth leading cause of death. Intracranial hemorrhage often leads to irreversible neurologic damage and high risk of disability. Since its symptoms can be either acute or chronic, early detection and prompt diagnosis play an important role in its treatment and outcomes. Due to the advances in machine learning algorithm, there has been growing interests in applying machine learning technique in medical field to perform multiple tasks, such as diagnosis, detection and many others.

Computed tomography (CT) is the primary imaging system that used to evaluate the hemorrhage in the head. But one of the drawbacks of CT is the incomplete projections and loss of information in raw CT data happened a lot due to implementation conditions [2], such as patient movement and image reconstruction. For some background information, CT works by taking images around object from multiple angle. The projection at a given angle is a line of brighter and darker pixel depending on how much the object block the beam at that angle as shown in Figure 1(a). Thus, the raw CT data is simply a display of all different projections of an object stacked together, refers to sinogram shown in Figure 1(b). The x-axis of sinograms represents angle and y-axis represents projection. Due to incomplete projections and potential loss of information, we want to bypass the image reconstruction process and use sinogram to train machine learning model to do the intracranial hemorrhage detection task.

Another widely known problem of CT is the high doses of radiation applied to patient during the scanning, which may induce cancer or genetic mutation. Reducing projected number of CT scanners is one direct way to decrease the amount of radiation exposure. But with too little projections, the quality of final CT images will be affected. To explored more about this problem, we varied the projection number of sinogram, used sparse-view raw CT data to perform classification task, and evaluated the relationship between model accuracy and projection number. In order to analyze the importance of each projection column, we purpose to add projection weight and Fourier mask as physical layers in our machine learning model.

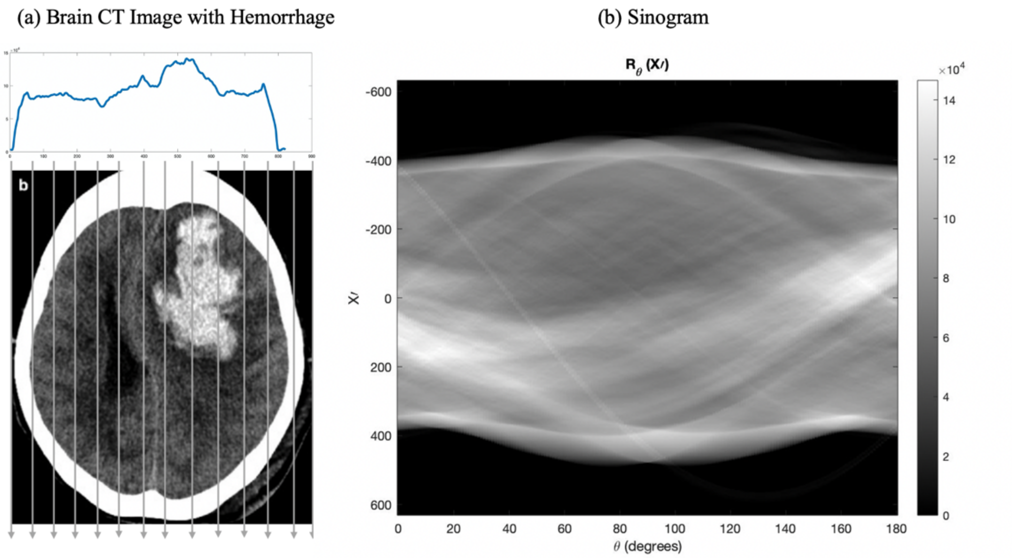

Figure 1: (a) An example of one line with all projections at one angle. The head CT scan contained intracranial hemorrhage. (b) A full sinogram of the image on the left.

In brief, our specific purpose for this project is: 1) to build a machine learning model using raw CT data to binarily classify whether there is intracranial hemorrhage; 2) to explore the relationship projection number and model accuracy; 3) to analyze the performance of Fourier mask and projection weight on classification performance. We demonstrated that to train machine learning model with sinogram containing 180 projections to detect intracranial hemorrhage could achieve above 85% on validation accuracy. The projection number of sinogram for training could be reduced to five without too much effect on classification accuracy.

2 Related Work

There are several researches about application of deep learning algorithms on raw CT data in sinogram-space and to perform on different tasks [2, 3, 5]. But most of researches were related to sparse view raw CT image reconstruction optimization. In Ref [2], they built a deep learning network to reconstruct high-quality CT images from incomplete projection sinogram. In Ref [3], they used deep learning method to interpolate the sparse view raw CT data in the sinogram and found the interpolation can be successful with 90% sparsity. These studies mentioned about incomplete projections happens quite often in CT and the image reconstruction involves recreation of images, which only can partially represent all raw data. In order to deal with this problem, for this project, we want to explore the intracranial hemorrhage detection accuracy with different projection number.

There is a study about developing a convolutional neural network (CNN) for interpreting sinograms, in Ref [2]. They proposed a CNN called SinoNet and compared its performance with normal CNN image-space-based model on body region identification task and also intracranial hemorrhage detection task. The SinoNet could achieve a better performance than a normal CNN training by reconstructed images on both tasks. In this project, in addition to classification accuracy and basic CNN structure, we would put more emphasis the model performance after implementing Fourier mask and projection weight.

3 Methods

3.1 Data Pre-processing

Our original data are head CT scans from Ref [4], which is a balanced dataset containing 100 normal head CT and 100 with hemorrhage in the head. Since the size of dataset is limited, we did data augmentation by rotating each image twelve times and each time by 30 degrees. Thus, we had 2400 images in total. Then, we converted head CT scans into sinogram using radon transform. The rotation of images was equivalent to horizontally shifting the sinogram (see in Figure 2). After transformed all images into sinogram, we split the training and validation data and the ratio was 9:1. Augmented data were shuffled by groups to make sure the training and validation dataset does not contain similar data augmented from same image,

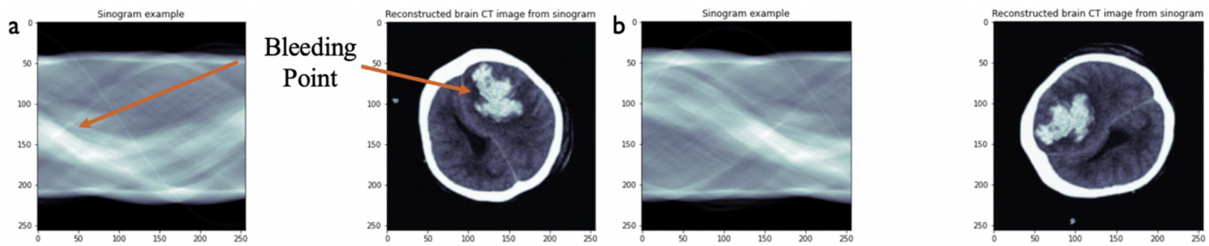

Figure 2: (a) A pair of sinogram and head CT scan with hemorrhage. The relative thicker and brighter lines in the sinogram was the bleeding point. (b) Figure 2(a) after rotated some degrees. It is equivalent to horizontally shift the sinogram in figure 2(a).

3.2 Physical layers and CNN model

The CNN model contains four convolution layers, each with 32 channels and 5x5 kernel, with max-pooling and batch normalization following by each two convolution layers. Then a dense layer, two dropouts with 0.5 rate to prevent overfitting and an output layer, which for binary classification. Since the size of our images are large, a complicated CNN will be computational expensive.

The two physical layers we implemented were projection weight and Fourier mask. The projection weight was used to simulate changing the mean effective energy of X-ray at each scan of CT. We created a trainable weight vector, which initial value was all ones, and multiplied it with a row vector filling with ones. The next step was taking transpose of the weight matrix and multiplied with the sinogram element-wisely, as demonstrated in Figure 3. By doing this, we assigned each column of sinogram a weight and evaluated the importance of each projection.

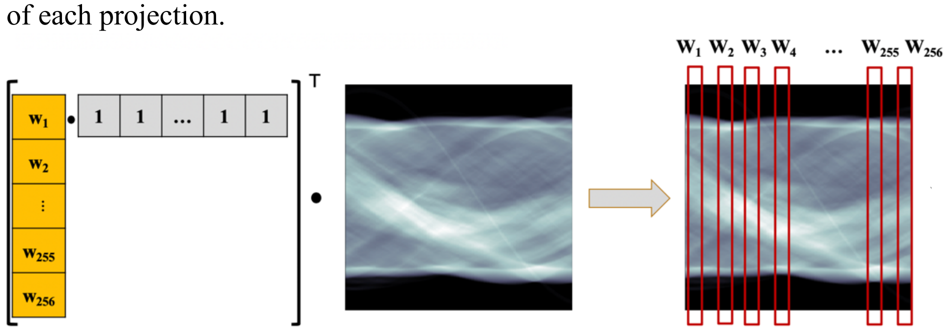

Figure 3: Demonstrated how projection weight physical layer worked.

The second physical layer we implemented was a Fourier mask, which is common process applied in filtered back projection [6]. We multiplied a trainable mask matrix all initialized as 0.01, with the sinogram in the frequency domain and then did inverse Fourier Transform to get a flited sinogram.

In order to compare the function of each physical layer and their performance to detect intracranial hemorrhage, we trained 180 projections sinogram with CNN model, CNN with projection weight layer, CNN with filter layer and CNN with both physical layers. For each model, we train it for 100 epochs and learning rate are set as 0.00001. The loss function used is binary cross-entropy and the optimizer is Adam.

3.3 Varied Projection Number

To explore the relationship between projection number and model accuracy, we tried to vary projection number to 2, 5, 10, 20, 30, 60, 90 and 180. Since the projection number represented by number of columns in the sinogram, directly applying radon transform will not give us square images. In order to generate a 256 x 256 sinogram, we did column-wise extension after radon transform. Figure 4 showed how the sinogram with 5 projections looks like after extension and its reconstructed CT image. After preprocessing sinogram, we trained the CNN model with both physical layers with different projection and compared their validation accuracies.

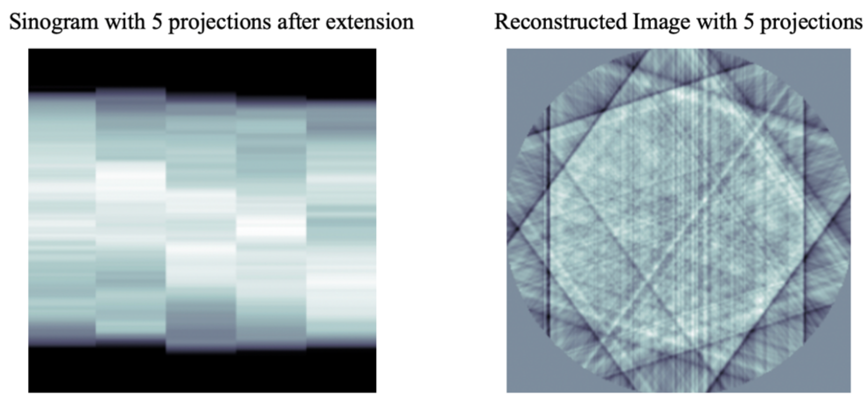

Figure 4: The sinogram with 5 projections after extension and its reconstructed image.

4 Results

4.1 Trained Physical Layers

For projection weight, the trained weight for each projection only changed slightly compared with their initial values. For example, the first five values for trained projection weights with 180 projections and CNN model with both physical layers are 1.0022097, 1.0017527, 1.0016527, 1.001285 and 1.0014513. This result meant that each projection has the same significance in our model.

For Fourier masks, after training on sinograms with eight different number of projections, we pulled out the mask and found that large values occurred around the center. As shown in Figure 5, the sinogram with 180 projection in Fourier space had a cross feature around the DC, which corresponds to the wave features in the image space and contains important information. The corresponding trained mask also expressed a cross characteristic. When trained on five projection sinogram, since the sinogram in the Fourier space does not have such cross feature, we cannot observe such feature in the trained mask.

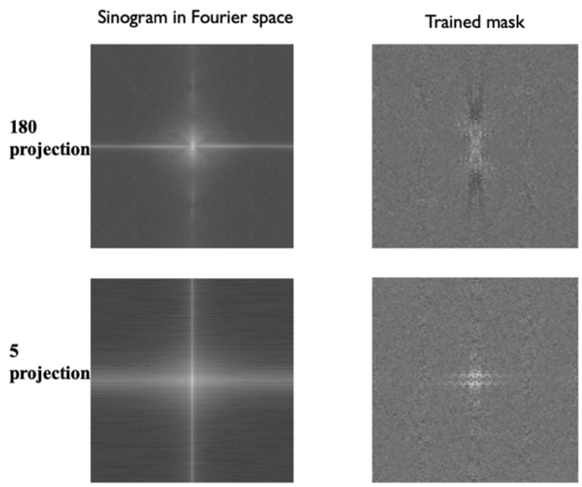

Figure 5: The shape of trained Fourier mask on sinograms with 180 and 5 projection.

4.2 Performance of Different Physical Layer

We compared the performance of four models containing different physical layers on the classification task. As shown in Figure 6, all four models reach a validation accuracy above 80% after 40 epochs. Models with filter layer and model with both physical layers have better performance than the other two, which achieve a validation accuracy of 85% after 40 epochs. For all these four models, training and validation accuracy converge within 100 epochs.

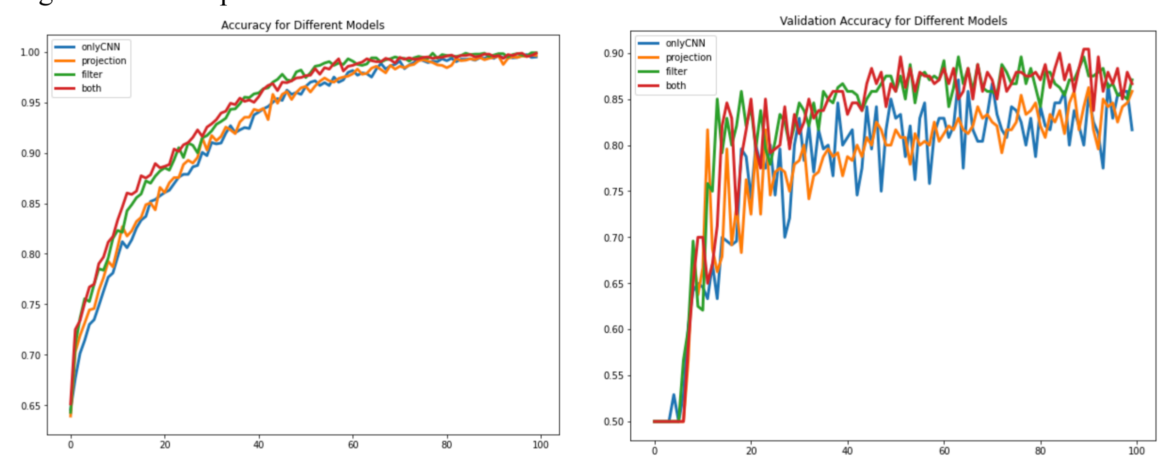

Figure 6: Traning and validation accuracy for four different models.

Table 1: The sensitivity and specificity values for four different models.

	CNN only	**Projection Weight**	**Filter**	**Both Layers**
Sensitivity	0.6413	0.6901	0.8035	**0.8268**
Specificity	0.9038	0.8980	**0.9056**	0.8930

In addition, we computed sensitivity and specificity values for each model, listed in Table 1, to evaluate their performance on this binary classification task. For sensitivity, the model with both physical layers had the highest value. For specificity, the model with filter had the highest value. These results match with the validation accuracy plot in Figure 5.

4.3 Performance with Different Projection Number

To explore the model performance on detecting intracranial hemorrhage with different projection number, we plotted the average validation accuracy for last 20 epochs versus projection number, as shown in Figure 7. While the projection number greater than 5, the average validation accuracy for last 20 epochs could achieve above 80%.

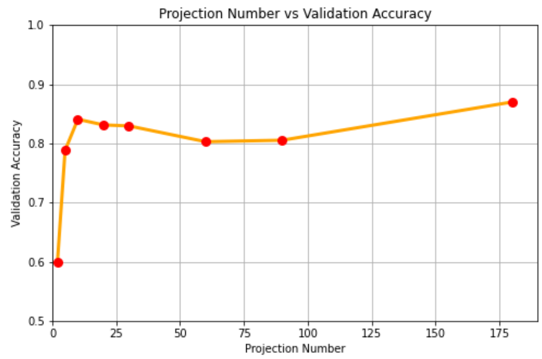

Figure 7: Validation accuracy at different projection number.

5 Discussion

Based on the results presented above, we were able to conclude that Fourier mask layer has more influence on model performance than projection weight layer. The model with filter mask achieved a higher validation accuracy than the model with projection weight. In general, all our model reached 85% accuracy on classifying intracranial hemorrhage, but in practical application on diagnosis, an 85% accuracy was far from enough. By looking at sensitivity and specificity, the pure CNN model and the model with projection weight only had 0.6413 and 0.6901 on sensitivity, which indicated a bad performance for detecting correctly with the samples with hemorrhage. Although, for these two models, they had high values on specificity, which were 0.9038 and 0.8980, we valued more on true positive rate, since we would rather identify a health people as sick than identify a sick people as health in real practice. For model with filter and with both physical layers, they had relatively higher values on both sensitivity and specificity. Thus, we concluded that filter had bigger effect on classification, which matched the conclusion drawn from validation accuracy plot.

Getting into details about our two physical layers. We found that the projection weights have little changes after training, our current interpretation is that each projection in the sinogram has same significance for our binary classification. If the classification task is more complicated, such as classifying hemorrhage types based on its location, our model will probably produce different trained projection weights. Customizing the learning rate for this single layer can also be a good attempt. For the Fourier mask physical layer, the trained filter is always magnifying the low frequency signal and the wave feature. About the varying projection number part, the model accuracy could reach above 80% when the projection number greater than 5, which indicated the strength of machine learning on detecting and analyzing images because the bleeding point was barely able to tell by using human eye from reconstructed image with only 5 projections. Out dataset has its limitation, all the hemorrhage presented in our data has relatively large area, this would make it easier for our model to do the classification. So, if we want to develop our model further, using a much larger dataset that contains more hemorrhage types is highly recommended. Another promising aspect to study is how row-wise 1D sum pooling the sinogram will affect the model classification performance, which is a good simulation of changing the X-ray detector size in CT.

References

[1] “RSNA Intracranial Hemorrhage Detection: Identify acute intracranial hemorrhage and its subtypes,” *Kaggle*, (2019).

[2] Dong, J., et al., “A deep learning reconstruction framework for X-ray computed tomography with incomplete data,” *PLOS ONE* 14(11): e0224426., (2019).

[3] Zhang, Z., et al., “Deep Learning Neural Network-based Sinogram Interpolation for Sparse-View CT Reconstruction,” *40th International Conference of the IEEE Engineering in Medicine and Biology Conference (EMBC),* (2018).

[4] Kitamura, F., “Head CT – hemorrhage: Balanced normal vs hemorrhage head CTs,” *Kaggle,* (2019).

[5] Lee, H., et al., “Machine Friendling Machine Learning: Interpretation of Computed Tomography Without Image Reconstruction,” *Scientific Reports* 15540, (2019).

[6] Prince, Jerry L., and Jonathan M. Links. *Medical Imaging Signals and Systems*. Pearson, 2015.