Iowa Initiative for Artificial Intelligence

Final Report

Project title:	Using Neural Networks to Derive Volume of Ischemic Core and Hypo-perfused Brain Tissue Using Computed Tomography Perfusion and Brain Datasets.
Principal Investigator:	Girish Bathla
Prepared by (IIAI):	Yanan Liu
Other investigators:	Honghai Zhang Colin P Derdeyn
Date:	

Were specific aims fulfilled:	Y
Readiness for extramural proposal?	Yes - Needs some work
If yes … Planned submission date	Jan 2021
Funding agency	ASNR
Grant mechanism	External, non-federal
If no … Why not? What went wrong?	

Brief summary of accomplished results:

We developed and validated a U-net model to automatically generate CBF and Tmax map from CT perfusion images and compared the volumes CBF<20% and Tmax>6 from prediction images with the software-calculated volumes. The averaged structural similarity index (SSI) between the AI-predicted CBF (Tmax) and commercial-software-derived CBF (Tmax) maps were 0.812±0.101 (0.825±0.101) for the validation group. The volumes of CBF<20%(Tmax>6) obtained from prediction image had a reasonable correlation of 0.632/0.687 (0.705/0.743) with the volume obtained using the commercially-available software (Siemens/Rapid).

Research report:

Aims (provided by PI):

Aim 1: To construct an artificial neural network using a combination of existing CTP datasets and subsequent MRI imaging to generate ischemic core and hypoperfusion volume maps through deep learning.

Aim 2: To validate the predicted outcomes with the results of the CTP and the final infarct volumes on the subsequent diffusion weighted imaging using additional, in-house, independent CTP and MRI datasets.

Aim 3: To explore the accuracy of the model using only 50 or 75% of the volumetric data of the existing dataset as well as with only the non-contrast CT (NCCT) Brain study. The NCCT will be from the same patients, acquired at the same time as the CTP, with the DWI images serving as the independent standard.

The aims were not changed, and the pilot project successfully completed all 3 aims as originally stated.

Data:

We have previously performed an internal data search and have 59 patient cases. Each case has one 4D CT perfusion image (512x512x9/11x28) and 3D CBF and Tmax maps (512x512x9/11).

AI/ML Approach:

In this study, U-net model was implemented for generating perfusion maps using Python. Training/validation split was 60/40% (36/23 patients) and was applied at the patient (not slice) level. Each slice of each case was treated separately due to small data set. Total of 352 and 217 slices were used for training and validation. The Structural Similarity Index (SSIM) [1] was used to quantify the agreement between prediction and ground truth. Ground truth was pre-processed perfusion maps (CBF/Tmax) from Siemens.

Experimental methods, validation approach:

Data preparation

Data preparation or pre-processing is an essential step in any machine learning study. In this project, we converted all our CT datasets in compressed Nifti format (nii.gz) and converted all the color maps to gray scale images using a data derived LUT. The input was downsampled by 4 in x and y axes to save computation time. After that, the downsampled CT data were normalized to [0,1].

Unet

In this project, Unet was implemented with Keras functional API, which makes it extremely easy to experiment with different interesting architectures. Input was the prepressed CT images. Output from the network is a 512x512 image which represents perfusion map that should be learned.

Volume analysis

After training the model, the volume calculated from prediction and ground truth were analyzed. For volume comparison, Pearson correlation coefficient was used. Pearson correlation coefficient is a statistic that measures linear correlation between two variables X and Y. It has a value between +1 and -1. A value of +1 is total positive linear correlation, 0 is no linear correlation, and -1 is total negative linear correlation.

Results:

Since CBF and Tmax maps are totally different, we trained two different models separately for CBF and Tmax.

Utilizing all 28 time-points data for training

After training, averaged SSIM achieved in the testing stage was 0.812±0.101 （0.825±0.101） 0.812±0.101 for CBF and 0.825± 0.101for Tmax in validation group (23 patients). Figures 1-4

showed examples of typical results obtained in patients using all 28 time-points data. There was not that much difference between prediction results and ground truth images.

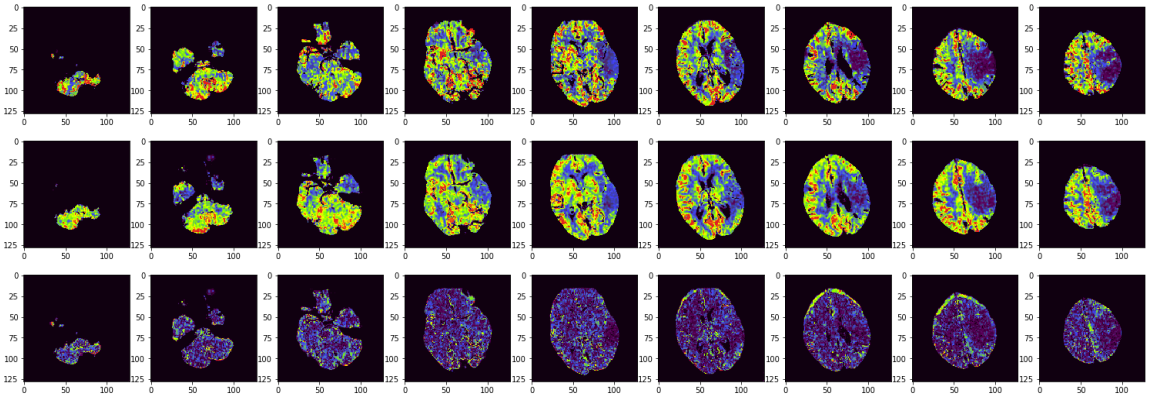


Figure 1. CBF prediction results from a typical example patient (SSIM=0.80). Top row: ground truth; middle row: prediction; third row: differences.

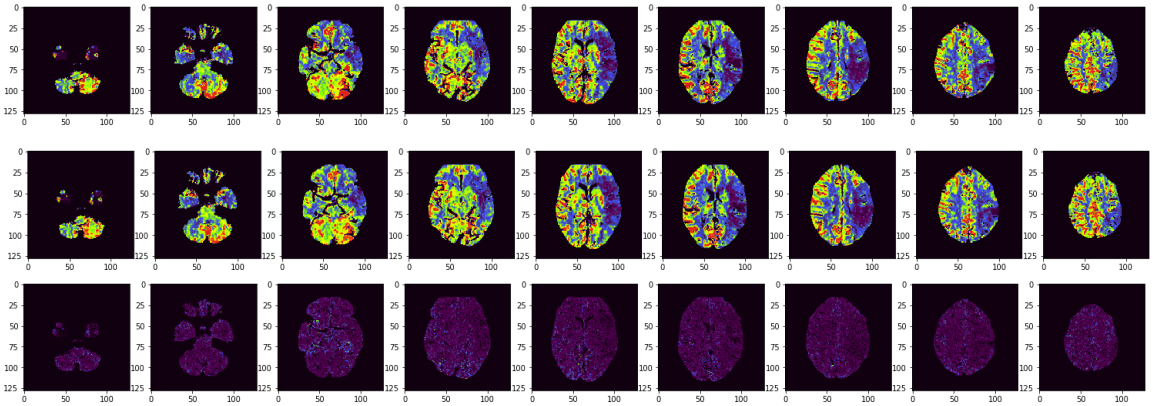


Figure 2. CBF prediction results from a typical example patient (SSIM=0.98). Top row: ground truth; middle row: prediction; third row: differences.

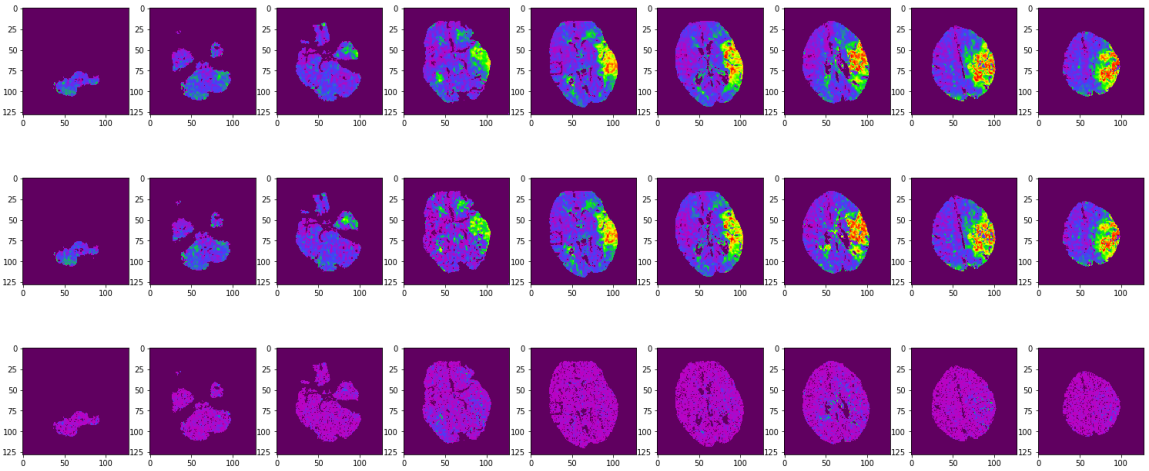


Figure 3. Tmax prediction results from a typical example patient (SSIM=0.80). Top row: ground truth; middle row: prediction; third row: differences.

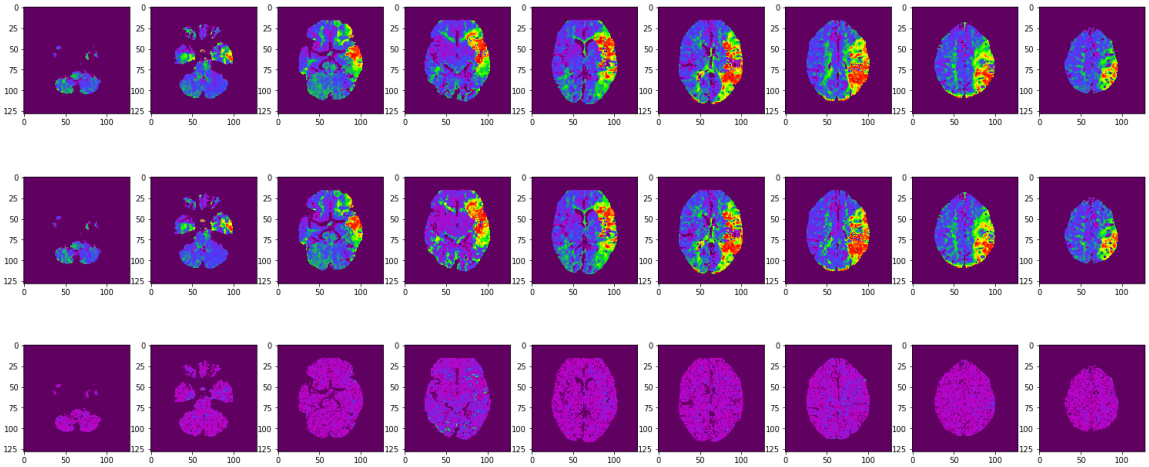


Figure 4. Tmax prediction results from a typical example patient (SSIM=0.98). Top row: ground truth; middle row: prediction; third row: differences

Utilizing less than all 28 time-points of image data

After we got the trained model, we explored the model with reduced time-points input. The reduced time-points input was achieved by selecting less time-points equally. Since our trained model still need 28 time-points input, the reduced time-points data was interpolated back to 28 time-points. Student t-test was employed to compare results obtained from all 28 time-points data and results obtained from reduced number of available data time-points. Tables 1 and 2 showed the t-test results. T-test is used to determine if the means of two sets of data are significantly different from each other. If the calculated *p*-value is below the threshold chosen for statistical significance (usually 0.05 level), then these two sets are significant different from each other. Figures 5-8 showed typical results obtained in patients from the validation group for which a reduced number of time-points was used.

Table 1: results for CBF (N=217)

Time-points	**SSIM**	**T-test (p-value)**
28	0.812±0.101	N/A
14	0.801±0.101	**0.246**
10	0.797±0.102	**0.115**
7	0.781±0.103	0.002

Table 2 results for Tmax

Time-points	**SSIM**	**t-test**

28	0.825±0.101	N/A
14	0.809±0.103	**0.107**
10	0.808±0.101	**0.086**
7	0.783±0.104	2.35e-05

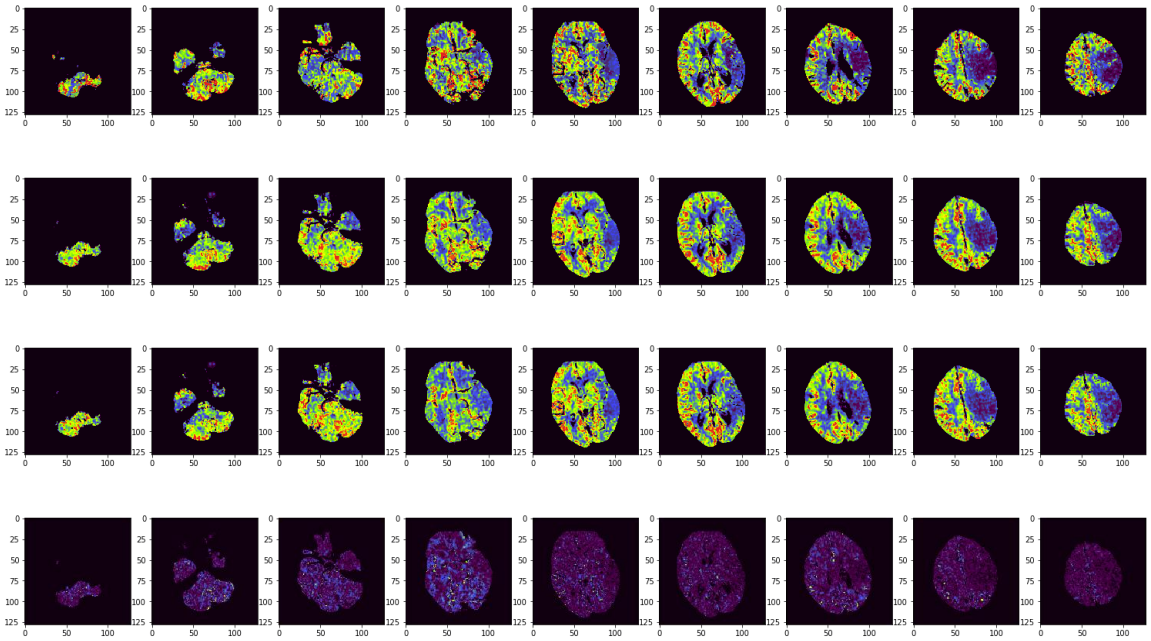
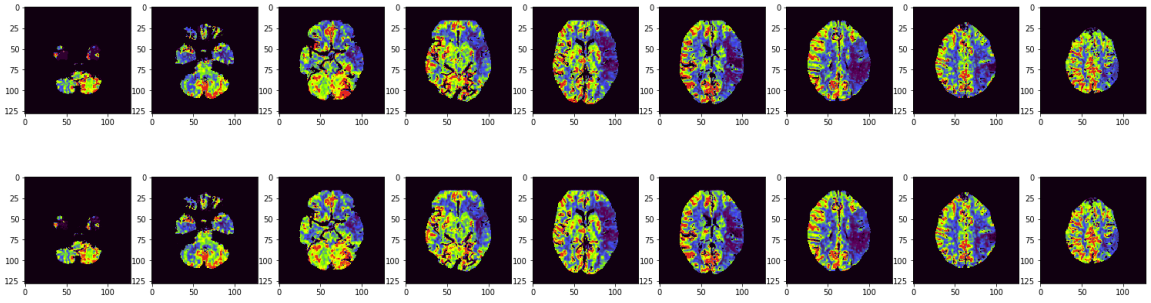


Figure 5. Typical patient (SSIM=0.80) CBF results. Top row: ground truth; middle row: prediction with 28 time-points; third row: prediction with 10 time-points; fourth row: differences between second and third row.



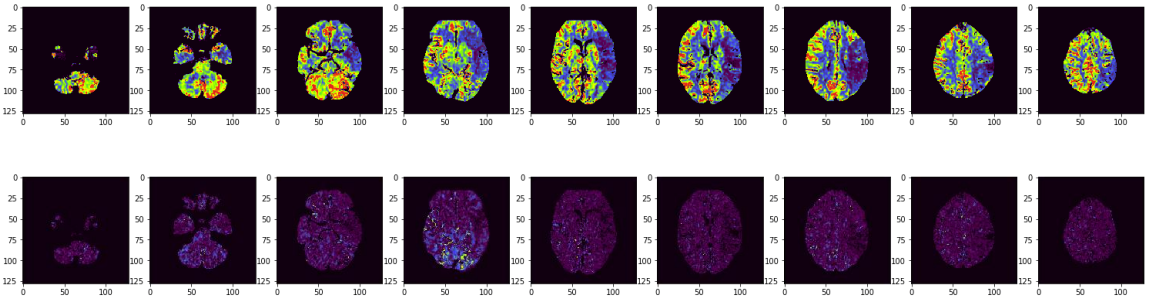


Figure 6. Typical patient (SSIM=0.98) CBF results. Top row: ground truth; middle row: prediction with 28 time-points; third row: prediction with 10 time-points; fourth row: differences between second and third row.

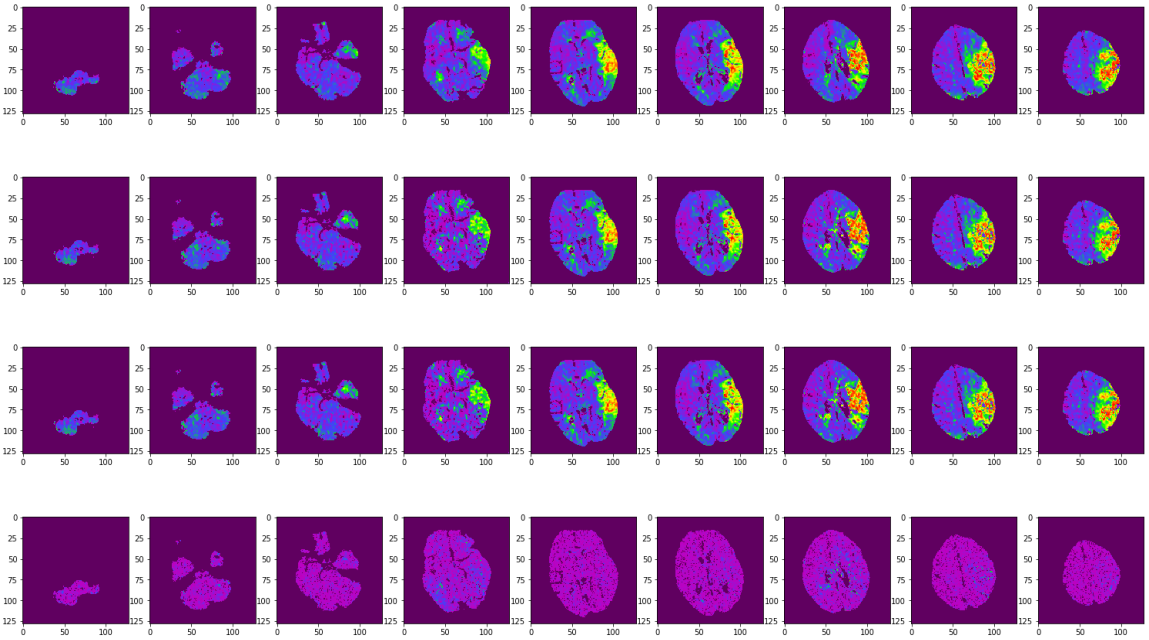
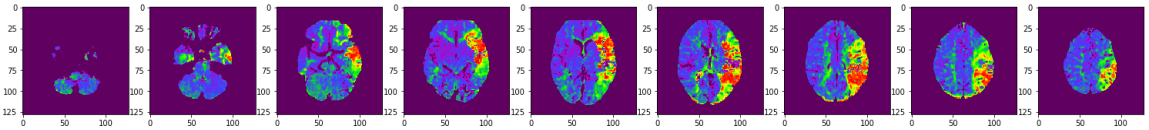


Figure 7. Selected patient (SSIM=0.80) Tmax results. Top row: ground truth; middle row: prediction with 28 time-points; third row: prediction with 10 time-points; fourth row: differences between second and third row.



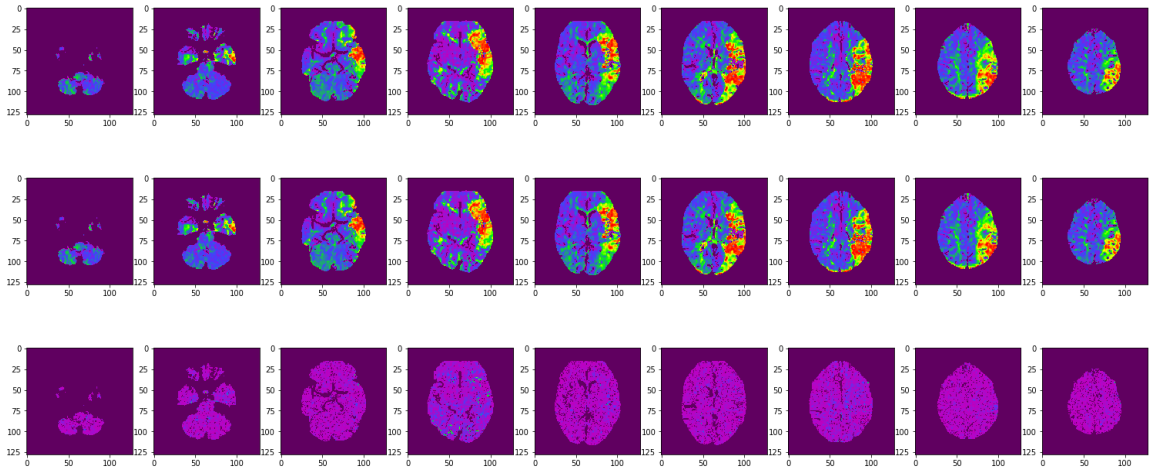


Figure 8. Typical patient (SSIM=0.98) Tmax results. Top row: ground truth; middle row: prediction with 28 time-points; third row: prediction with 10 time-points; fourth row: differences between second and third row.

Comparing volumes

We calculated volume in CBF <0.2 from ground truth and predicted CBF map, and compared them with software results (syngo, rapid) for patients in validation group (23 patients). Figures 9-12 show the scatterplot charts, regression equation and Pearson's correlation coefficients.

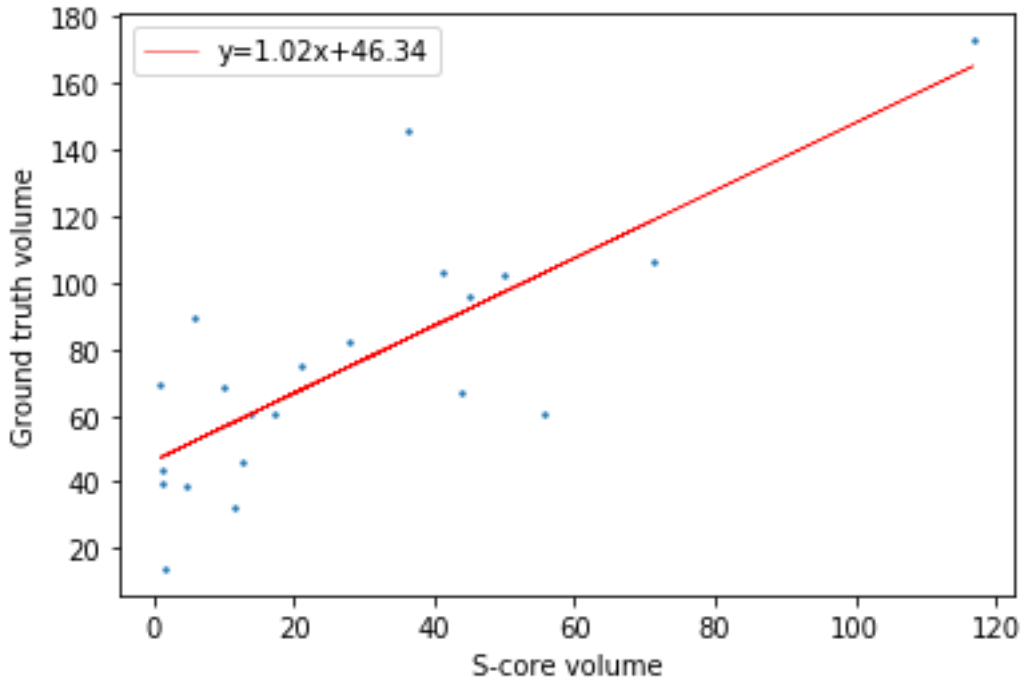


Figure 9. CBF Siemens software vs. ground truth volume scatter plot. X: software volume; Y: ground truth volume. Pearson's correlation coefficient=0.781.

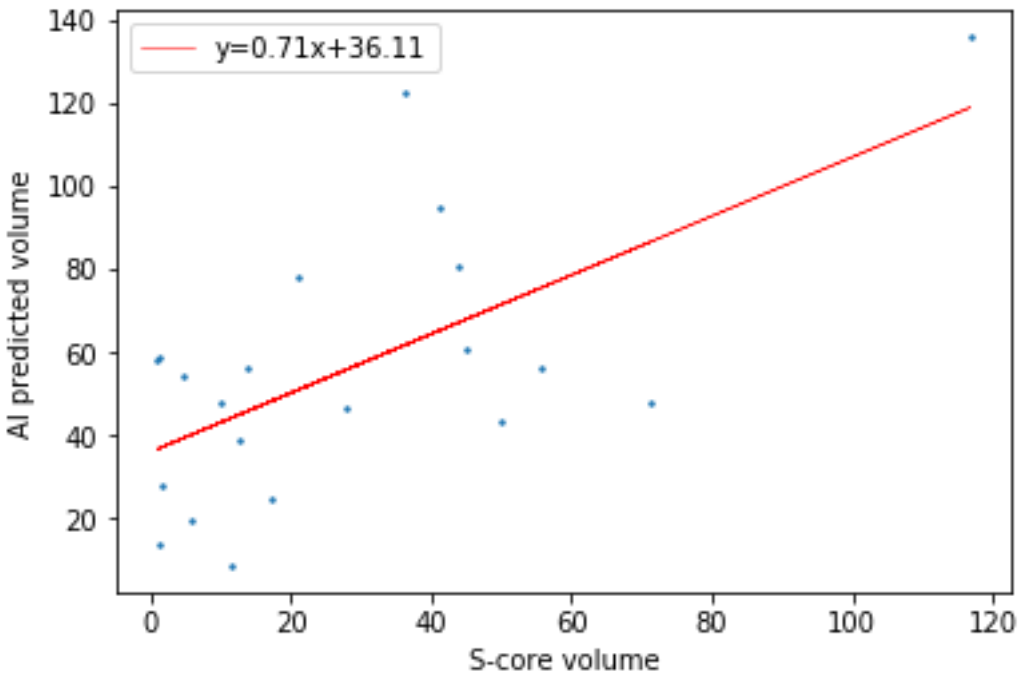


Figure 10. CBF Siemens software vs prediction volume scatter plot. X: software volume; Y: prediction. Pearson's correlation coefficient=0.632.

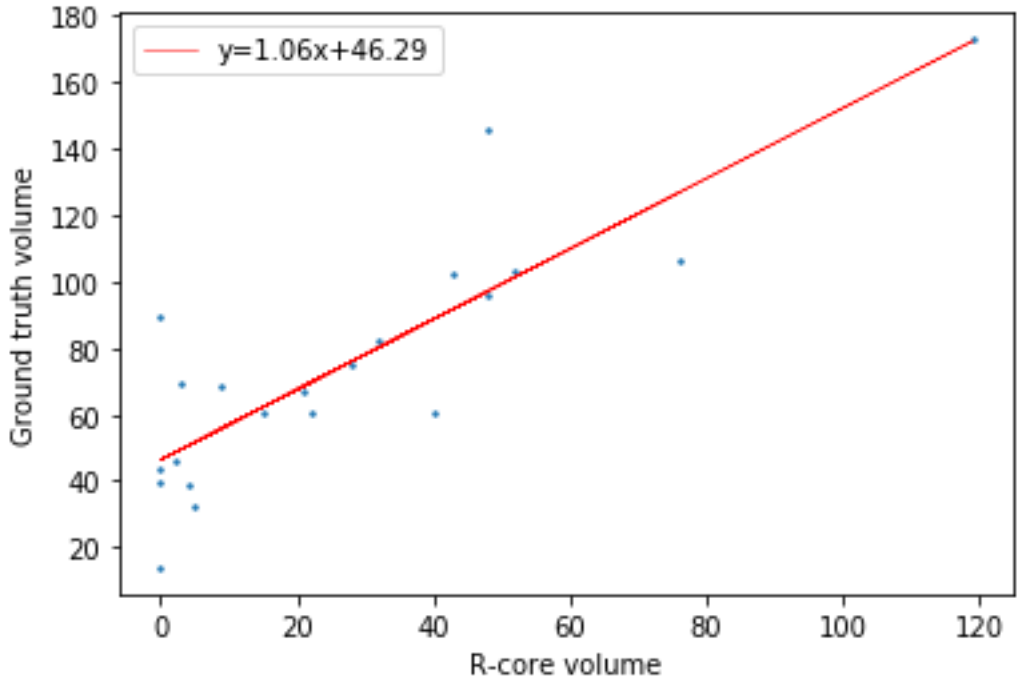


Figure 11. CBF rapid software vs ground truth volume scatter plot. X: software volume; Y: ground truth volume. Pearson's correlation coefficient=0.852.

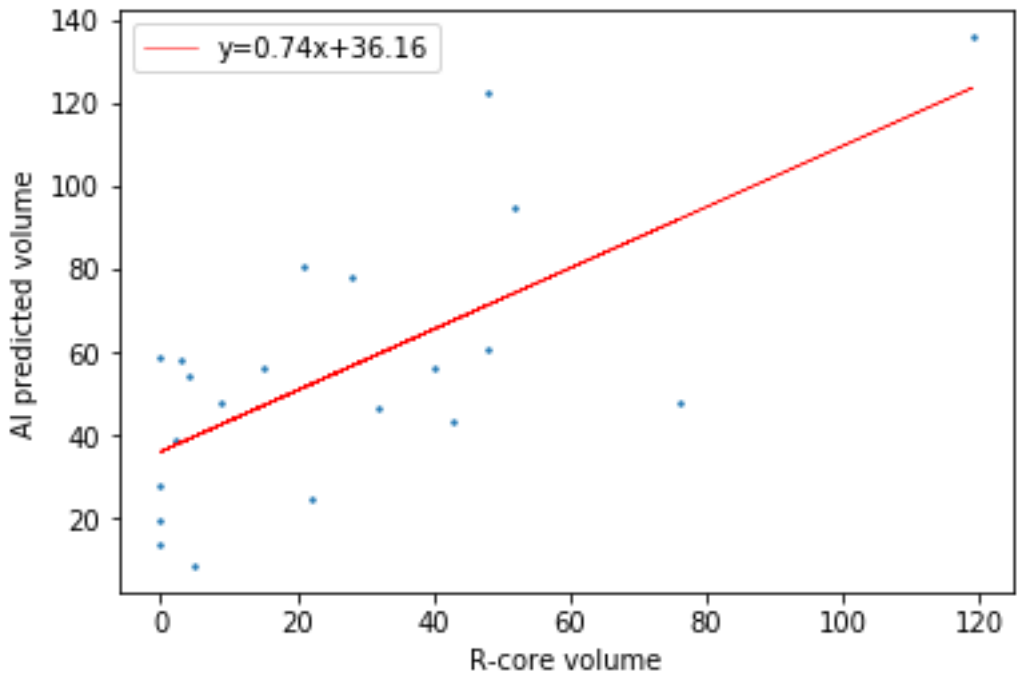


Figure 12. CBF rapid software vs prediction volume scatter plot. X: software volume; Y: prediction. Pearson's correlation coefficient=0.687.

For Tmax, we calculated volume in Tmax>6 from ground truth Tmax and predicted Tmax map and compared them with software results (syngo and rapid) for patients in validation group. Figure 13-16 showed the scatter plots, regression equation, and Pearson's correlation coefficients.

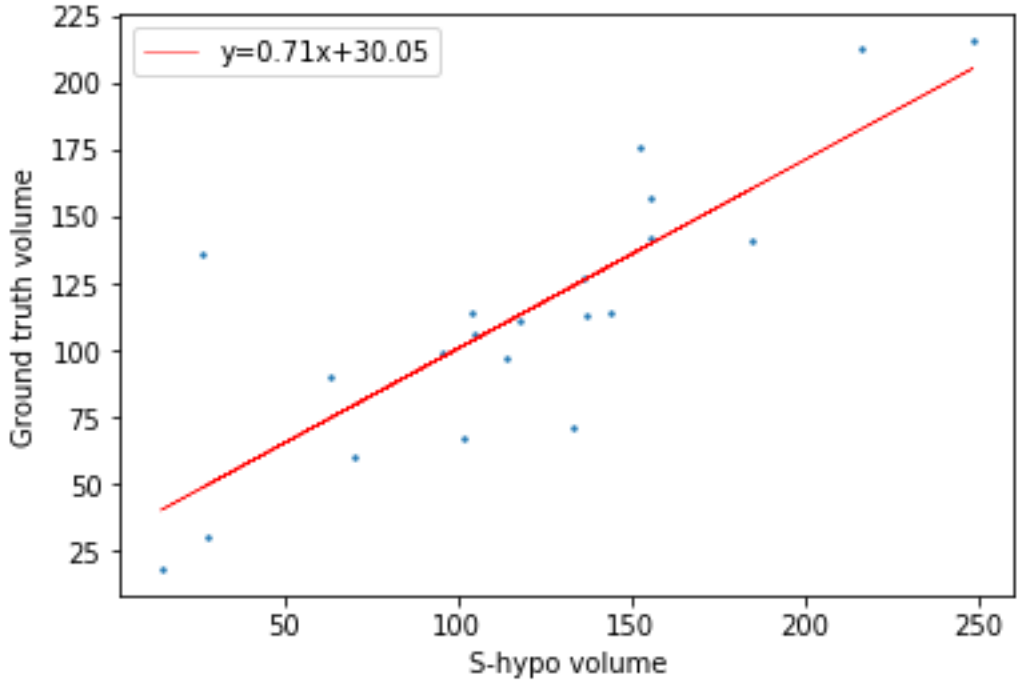


Figure 13. Tmax Siemens software vs ground truth volume scatter plot. X: software volume; Y: ground truth volume. Pearson's correlation coefficient=0.820.

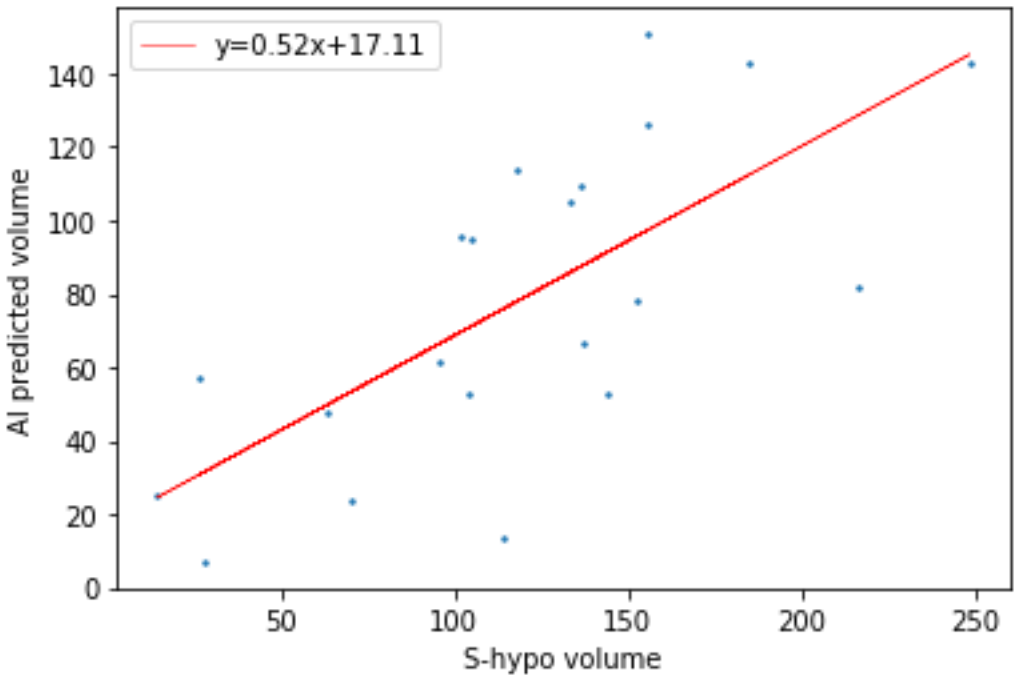


Figure 14. Tmax Siemens software vs prediction volume scatter plot. X: software volume; Y: prediction. Pearson's correlation coefficient=0.705.

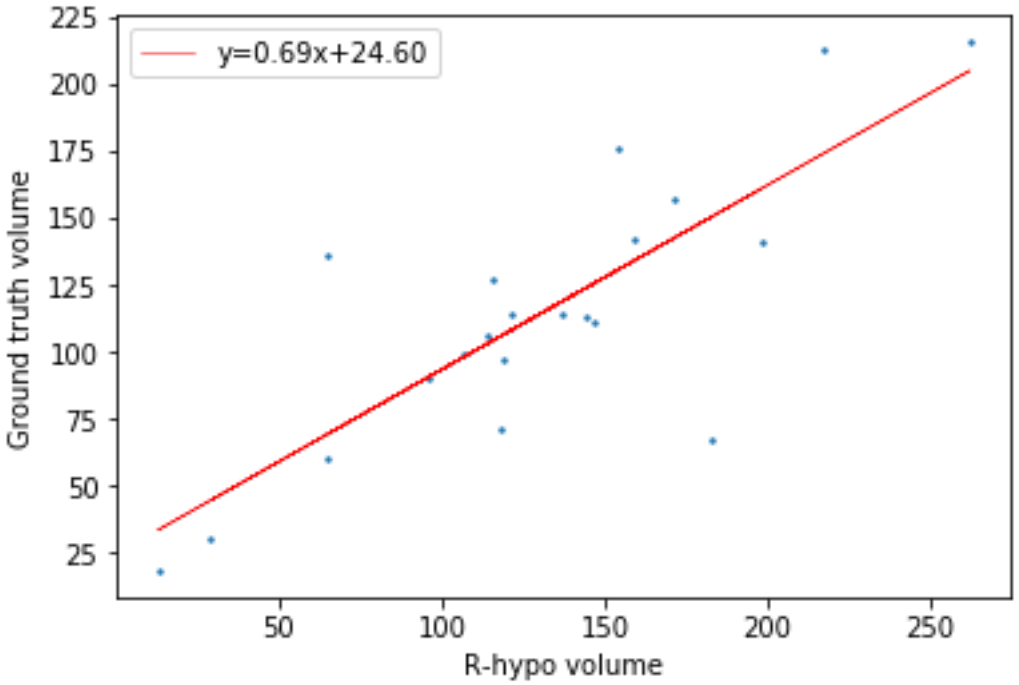


Figure 15. Tmax rapid software vs ground truth volume scatter plot. X: software volume; Y: ground truth volume. Pearson's correlation coefficient=0.803.

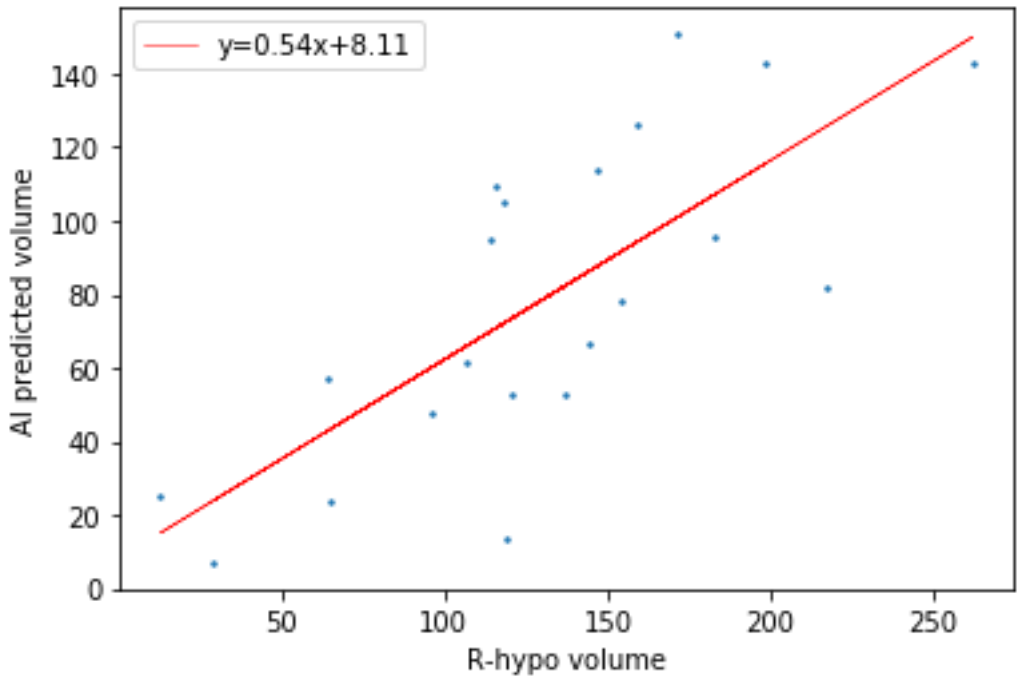


Figure 16. Tmax rapid software vs prediction volume scatter plot. X: software volume; Y: prediction. Pearson's correlation coefficient=0.743.

Exploring the effect of spatial downsampling

To understand the effect of pre-preprocessing step on volume, the volume was calculated from original CBF perfusion map without downsampling. In addition, volumes from different brain area (whole brain, left brain, right brain and affected brain) were calculated to understand the effect of brain area. Figure 17-20 showed scatter plots of volume obtained from non-downsampled image and downsampled image in different brain area. The scatter plot showed that the pre-processing step for machine learning did not affect the final volume.

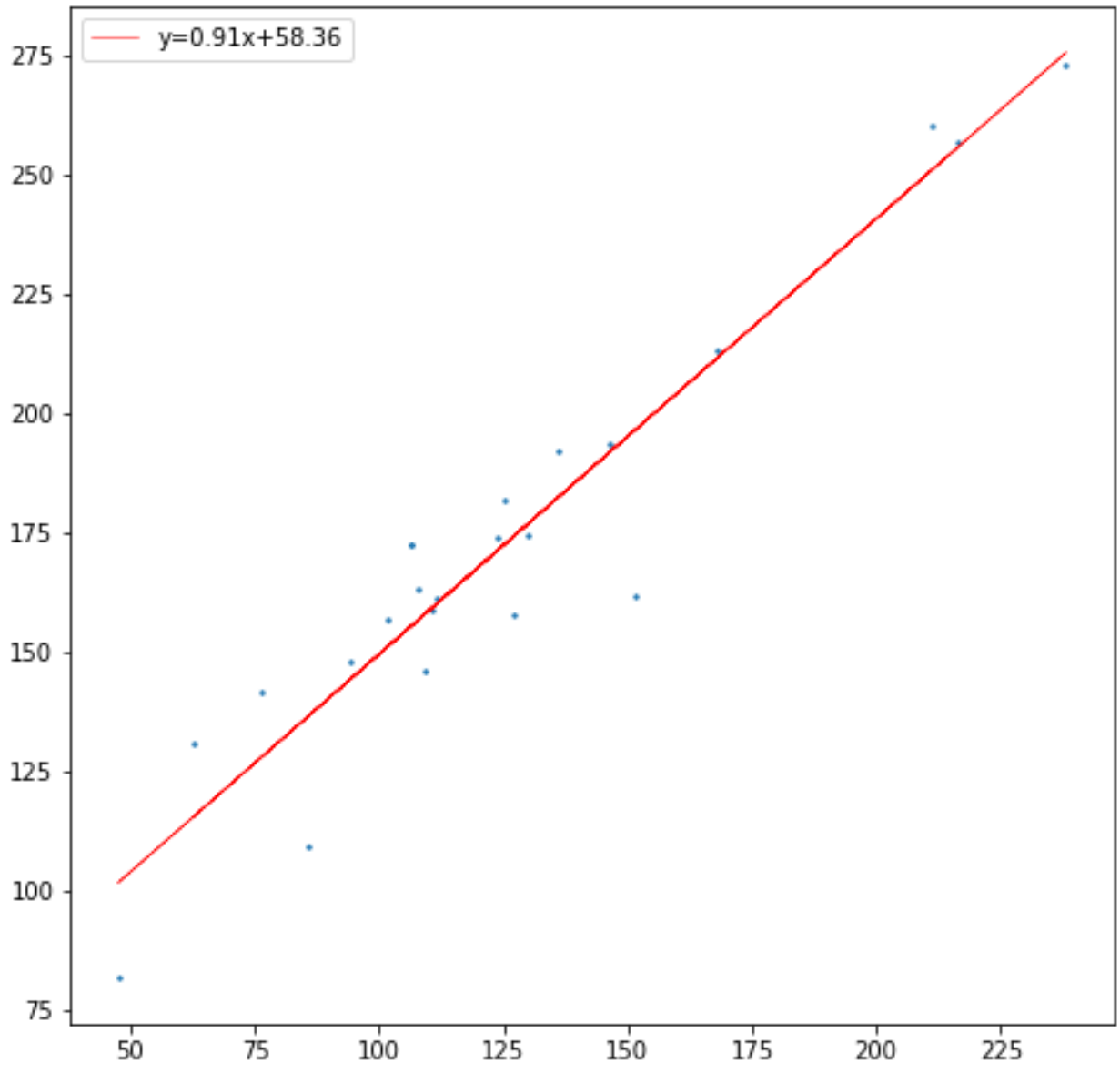


Figure 17. CBF non-downsampled vs downsampled volume of whole brain scatter plot. X: non-downsampled volume; Y: downsampled volume. Pearson's correlation coefficient=0.952

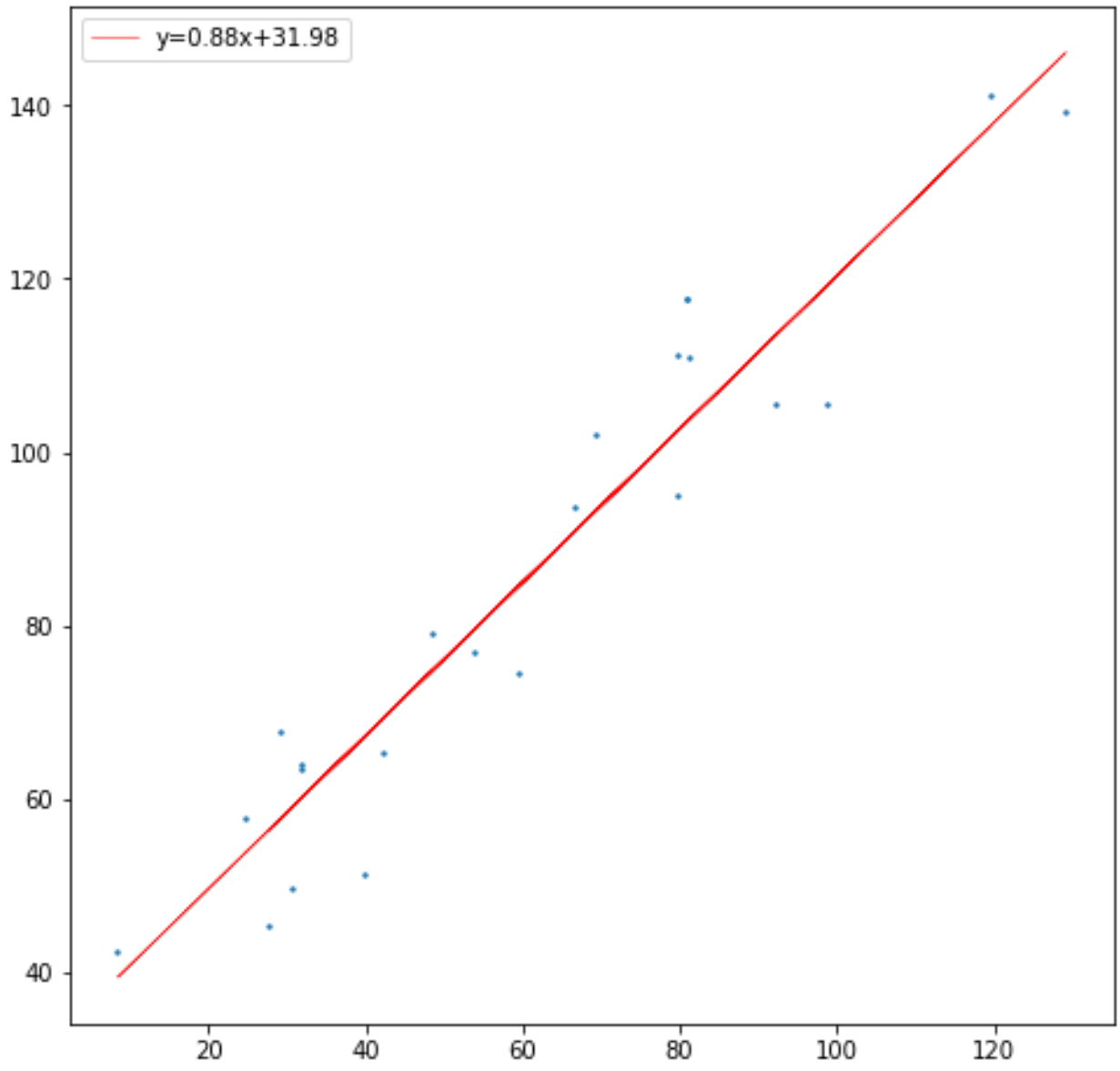


Figure 18. CBF non-downsampled vs downsampled volume of right brain scatter plot. X: non-downsampled volume; Y: downsampled volume. Pearson's correlation coefficient=0.954

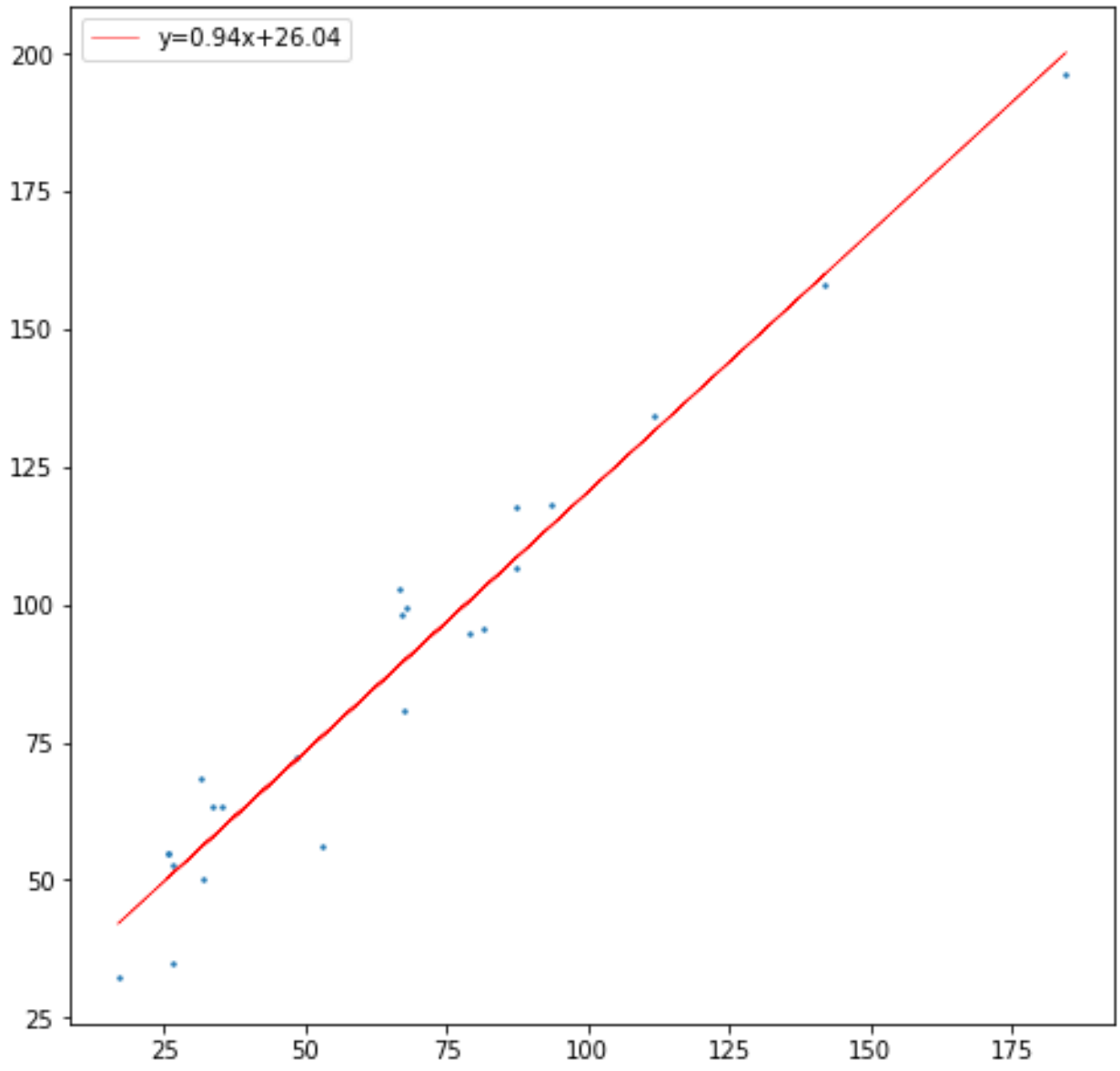


Figure 19. CBF non-downsampled vs downsampled volume of left-brain scatter plot. X: non-downsampled volume; Y: downsampled volume. Pearson's correlation coefficient=0.976

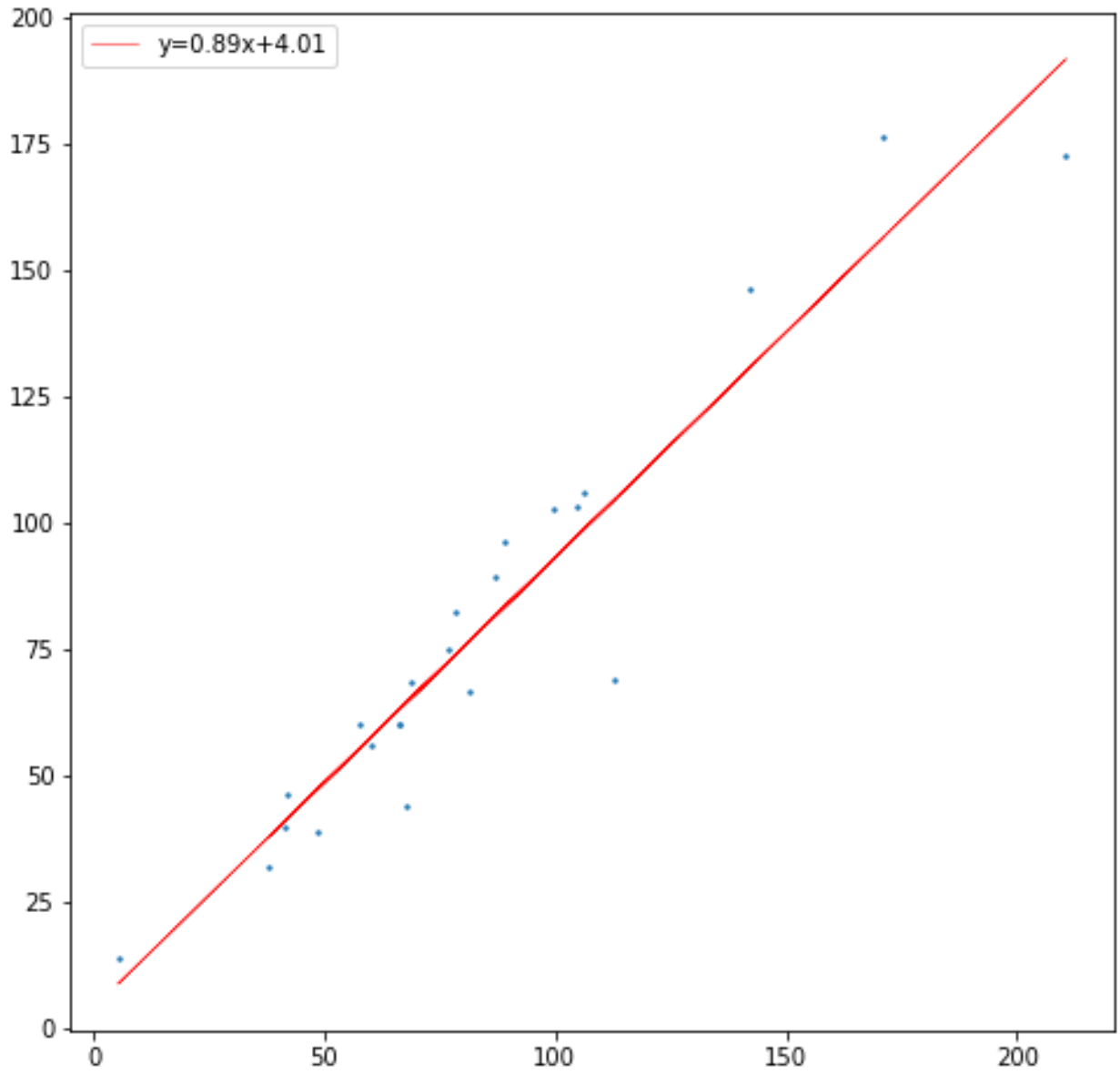


Figure 20. CBF non-downsampled vs downsampled volume of affected brain scatter plot. X: non-downsampled volume; Y: downsampled volume. Pearson's correlation coefficient=0.954

Calculating volumes from reduced numbers of time-point images

To explore the effect of reduced time-points results, volumes obtained from 28 time-points and reduced time-points prediction images were calculated. Scatter plots were showed in Figure 21-24.

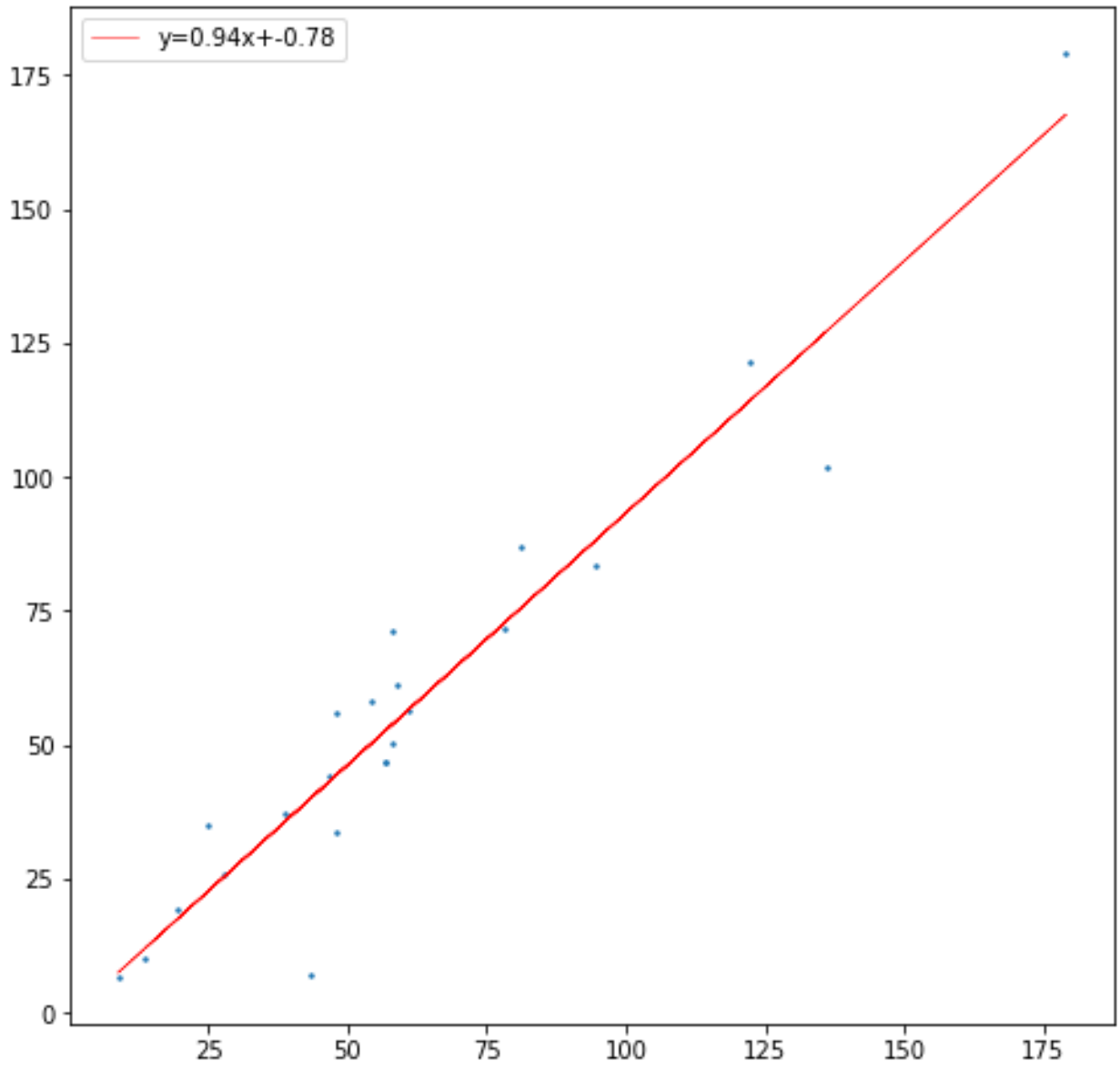


Figure 21. CBF 28-time-points vs 14-time-points volume of affected brain scatter plot. X:28-time-points volume; Y: 14-time-points volume. Pearson's correlation coefficient=0.957

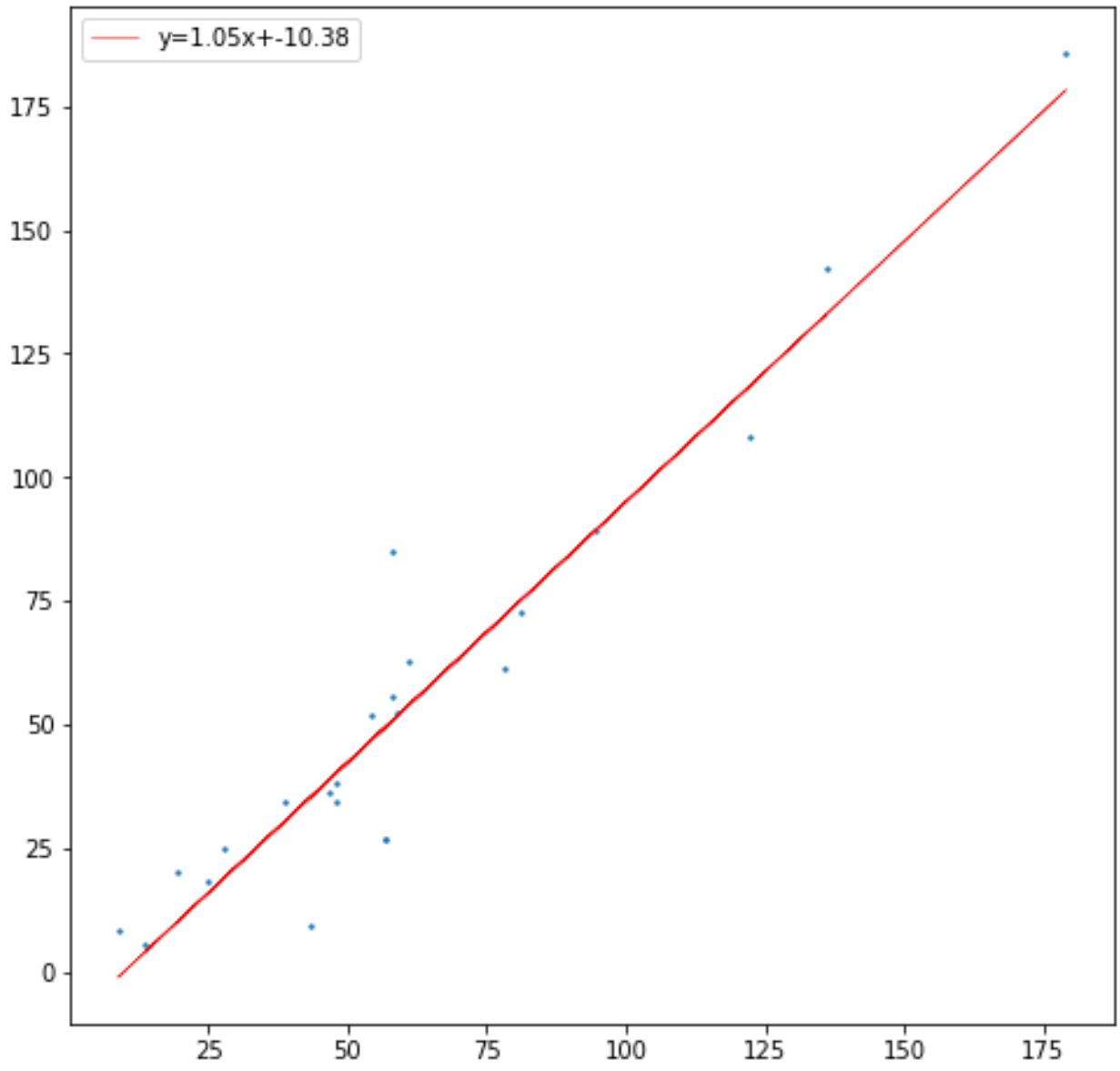


Figure 22. CBF 28-time-points vs 10-time-points volume of affected brain scatter plot. X:28-time-points volume; Y: 10-time-points volume. Pearson's correlation coefficient=0.957

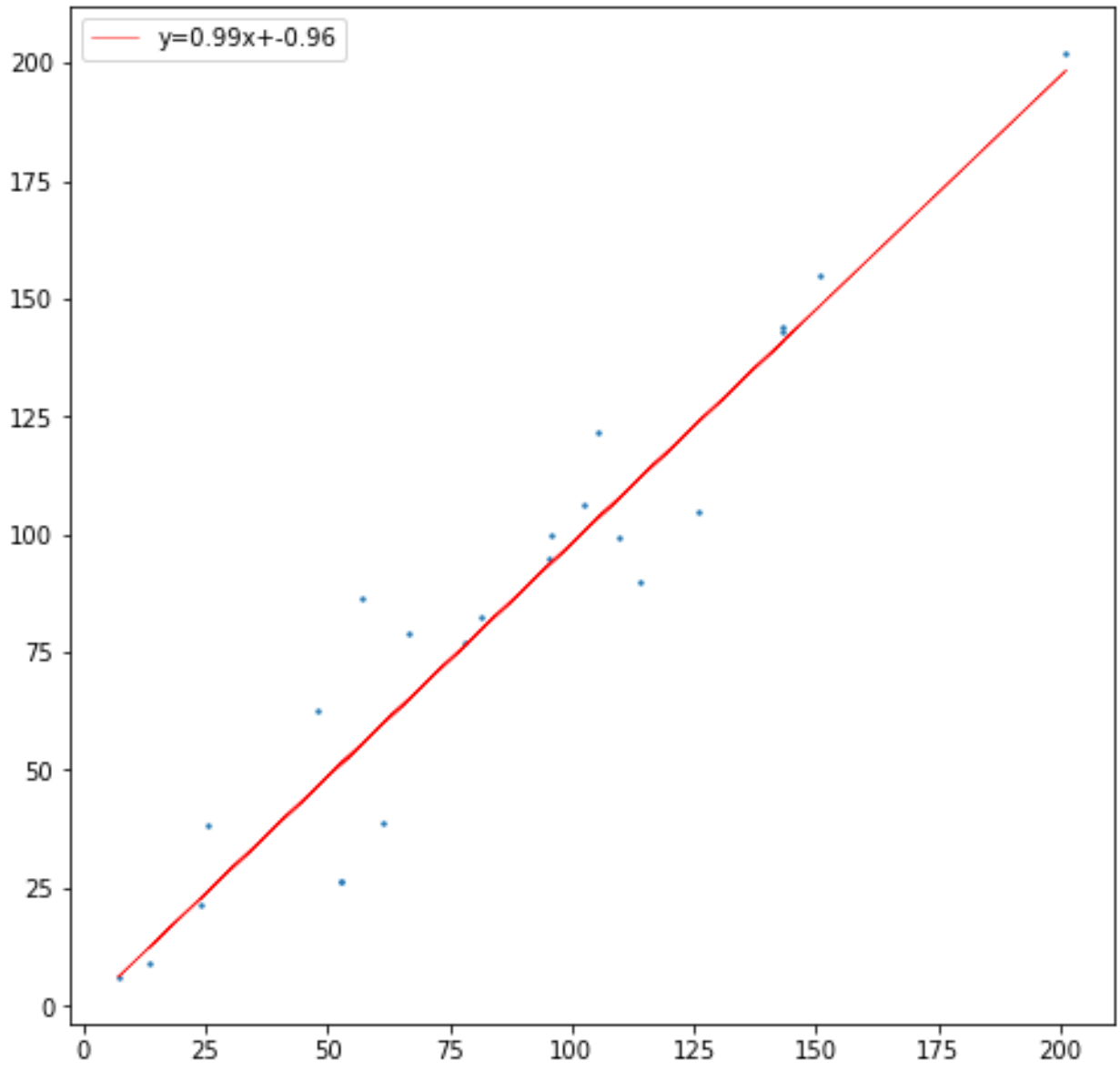


Figure 23. Tmax 28-time-points vs 14-time-points volume of affected brain scatter plot. X:28-time-points volume; Y: 14-time-points volume. Pearson's correlation coefficient=0.957

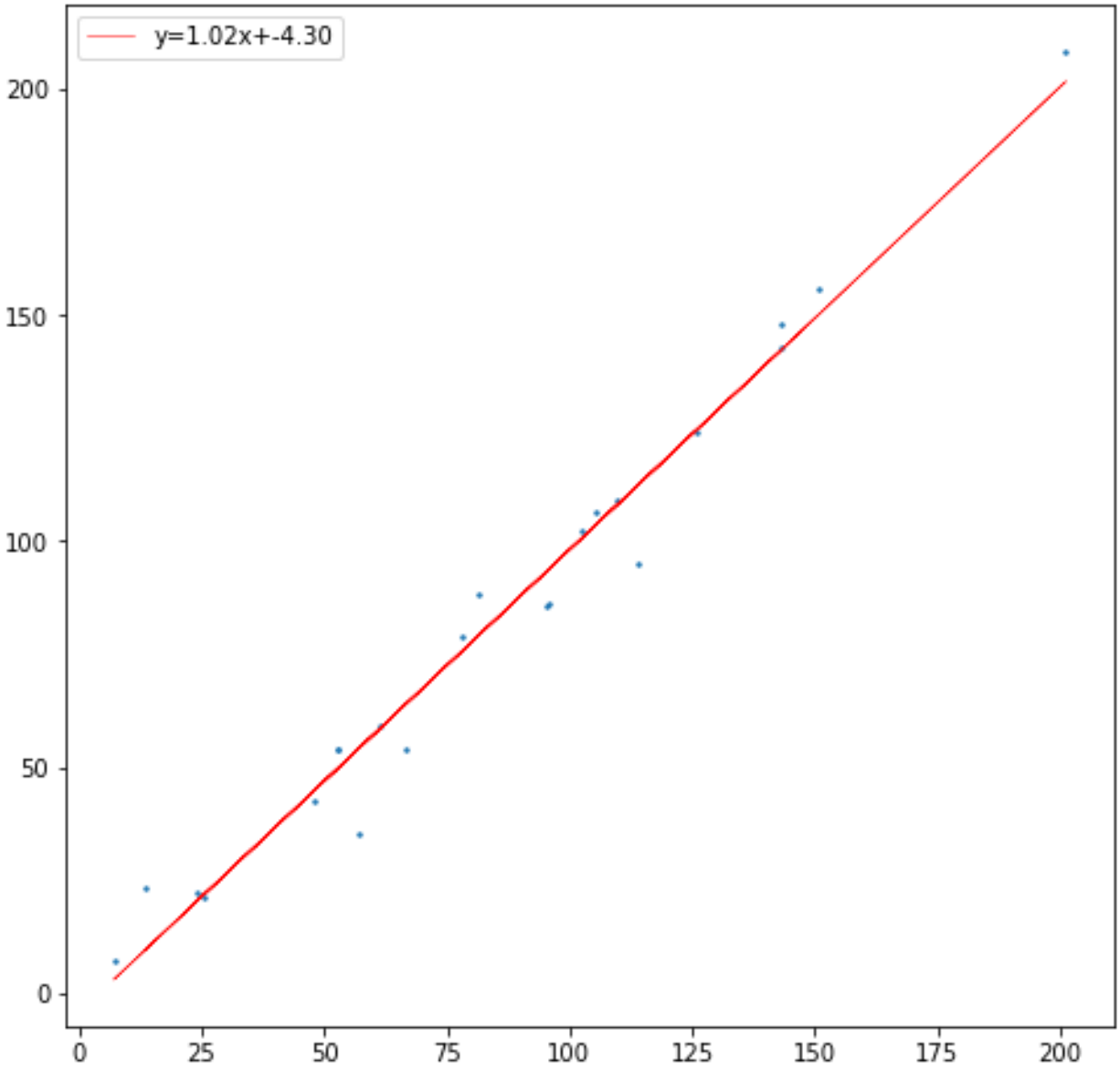


Figure 24. Tmax 28-time-points vs 10-time-points volume of affected brain scatter plot. X:28-time-points volume; Y: 10-time-points volume. Pearson's correlation coefficient=0.988

In this project, we clearly showed that deep learning method (Unet) can compute perfusion map for CT perfusion. Using only 1/3 time-points was not significantly different from using all time-points. Volume obtained from prediction perfusion map has a reasonable correlation with software results. Due to data limit for this pilot project, we only explored the deep learning model with 23 patients in validation group. The machine learning model and our approach can be further developed in larger dataset. Furthermore, the CT perfusion map used as ground truth to train machine learning in this project were prepared by Syngo (Siemens software). Since Rapid is used in clinal trial and is validated, we could explore this whole machine learning approach using Rapid prepared perfusion map to have a better correlation between predicted volume and software results.

Ideas/aims for future extramural project:

Among several ideas, the most promising seems to make separate models for Tmax and CBF and train/evaluate in leave-x-out fashion on Rapid maps, considering proper LUTs for the Rapid software data. Here, the training would specifically use only 14 time points rather than current 28 timepoints. Alternatively. Training models with only one third of temporal data may still be possible, but needs better data curation.

The current model using full data to train the prediction model has an advantage of being vendor independent and not requiring AIF selection, but such a design still requires training data with full radiation dose and thus is not preferred for the future.

Publications resulting from project:

Results are almost ready for publication but more subject level information is needed in tabular/graph format.

The following to be included:

- All patient ID's (both validation and training) with subjects clearly marked as whether they were used in the training or validation group.
- Individual SSIM scores
- Individual Rapid/Siemens volumes as well as model predicted volumes for both CBF and Tmax.

References

1. Wang, Zhou; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. (2004-04-01). "Image quality assessment: from error visibility to structural similarity". IEEE Transactions on Image Processing. **13** (4): 600–612