EasyChair Preprint

№ 6792

CT Perfusion Imaging of the Brain with Machine Learning

Kellen Cheng, Kunakorn Atchaneeyasakul, Zeid Barakat,
David Liebeskind and Fabien Scalzo

October 6, 2021

CT Perfusion Imaging of the Brain with Machine Learning

Kellen Cheng[1], Kunakorn Atchaneeyasakul[2], Zeid Barakat[3], David S. Liebeskind[1], and Fabien Scalzo[1,4]

[1] Department of Neurology, University of California, Los Angeles (UCLA), CA 90095
[2] School of Medicine, University of Pittsburgh, Pittsburgh, PA 15213
[3] Perfuse, inc
[4] Seaver College, Pepperdine University, Malibu, CA 90265

Abstract. Computed tomography (CT) perfusion imaging is a routinely used technique in the field of neurovascular imaging. The progression of a bolus of contrast agent through the neurovasculature is imaged in a series of CT scans. Relevant perfusion parameters, such as cerebral blood volume (CBV), flow (CBF) and delay (Tmax), can be computed by deconvolution of the contrast-time curves with the bolus shape measured at one of the feeding arteries. These parameters are crucial in the medical management of acute stroke patients, where they are used to identify the extent of likely salvageable tissue and irreversibly damaged infarct core. Deconvolution is normally achieved using singular value decomposition (SVD). However, studies have shown that such a technique is noise sensitive and easily influenced by artifacts in the source image, and may introduce further distortions in the output parameters. In this study, we present a machine learning approach to the estimation of perfusion parameters from CT imaging. Standard types of regression-based machine learning models were trained on the raw/native CT perfusion imaging data to reproduce the output of an FDA-approved commercial implementation of the SVD deconvolution algorithm. As part of our experiments, Kernel ridge regression and random forest models performed best (SSIM: 82%, 84%), trading off quick run time for better prediction accuracy while requiring a relatively low number of training examples.

1 Introduction

Clinical management of neurovascular diseases, and acute stroke in particular, requires perfusion imaging to quantify blood flow through the brain parenchyma. In computed tomography (CT) perfusion imaging [1, 2], or CT perfusion, images are obtained after a bolus of contrast agent, that attenuates the signal intensity on the X-ray receptor, is injected to flow through vessels and tissue while a series of consecutive scans is taken. The spatiotemporal signal attenuation resulting from the contrast agent can be inferred across the brain over the duration of the acquisition.

Due to the dimensionality of the resulting CT perfusion data (i.e. 4D), these concentration-time curves are not directly interpreted by clinicians. Instead, features are inferred from it and displayed as a set of 2D image slices. Such feature

maps include cerebral blood flow (CBF), cerebral blood volume (CBV), mean transit time (MTT), time-to-peak (TTP), and time-to-maximum (Tmax). The computation of these features typically amounts at deconvolving the data with the arterial input function (AIF) to obtain the residue function: a curve characterizing blood flow through a given volume element. These parametric estimations are widely used in assessing neurovascular conditions [3]. In acute stroke, treatment selection is performed by comparing the volume of the ischemic core with that of the hypoperfused tissue which is at risk, but which may still be salvaged. Parameters extracted from perfusion imaging are vital for identifying the tissue at risk. Various studies have shown correlations between the perfusion parameters and clinical outcome in terms of Rankin score and Barthel Index [4].

The problem of inferring the residue function, and thus the feature maps, is ill-posed, and deconvolution techniques such as singular value decomposition (SVD) are unstable in the presence of noise, causing biases [5] that could alter clinical decisions. Certain techniques have been developed to reduce this problem. A smoother residue function can be achieved through a Gaussian process for deconvolution (GPD) [6], Tikhonov Regularization [7], and a physiological model of microvasculature [8]. Attempts to provide better estimates of perfusion parameters have also used Expectation Maximization (EM) [9], and Bayesian estimation [10]. Other groups have focused on reducing the noise of CTP source images prior to the deconvolution using Deep Learning techniques [11] and restoring higher temporal resolution [12]. These novel algorithms have provided encouraging results to mitigate the risk of bias introduced by the deconvolution.

As part of the clinical workflow in acute stroke, perfusion maps are reviewed by visual inspection and by automated processing based on image segmentation. While these interpretations are complementary, both are susceptible to noise and bias in the reported maps. As an alternative to improving the quality of perfusion maps by post-processing, attempts have been made to move beyond thresholds and instead apply machine-learning to compute standard perfusion maps [13, 14]. Yu et al. presented a model predicting hemorrhagic transformation severity directly from MRI source perfusion imaging [15]. Similarly, recent models have used source MR [16] and angiographic [17] images to produce perfusion maps and have demonstrated greater robustness when compared to deconvolution-based approaches.

In this paper, we follow a similar strategy and propose to evaluate the use of regression-based machine learning models to compute perfusion maps from CT perfusion data without the use of deconvolution. The models are trained on a large number of concentration-time curves extracted locally. We demonstrate in our experiments the similarity of the maps obtained via machine learning to the ones obtained via deconvolution.

2 Methods

2.1 Overview

The study is based on imaging data taken from patients treated for acute ischemic stroke at the UCLA Ronald Reagan Hospital in Los Angeles. The deidentified medical imaging data composed the training set that was used to train various types of machine learning models, including but not limited to kernel ridge and random forest regression. The model utilizes regression analysis to predict the perfusion parameter scans of a patient based on their time series CT slices. This process was repeated for each patient and their 6 slice locations, outputting perfusion maps that were compared to the ground truth values. Model accuracy and correction were evaluated by the structural similarity index metric (SSIM) between the actual and predicted rCBV and rCBF perfusion maps.

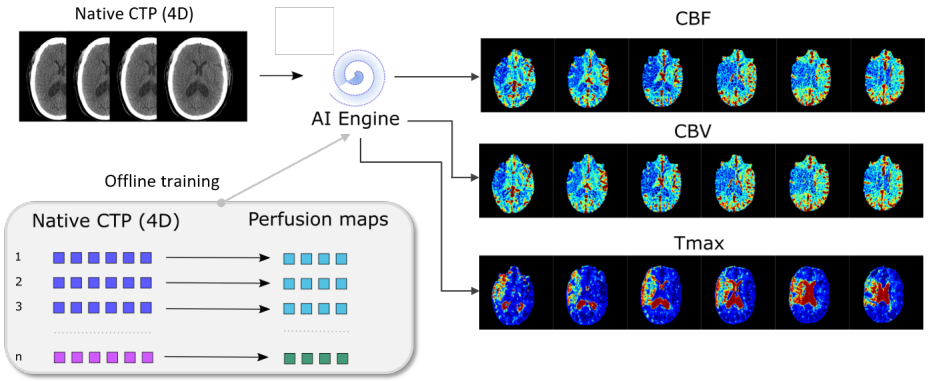


Fig. 1: Machine learning model pipeline for perfusion parameter prediction.

2.2 Dataset

The dataset is a retrospective collection of patients' CT scans gathered from perfusion imaging of the neurovascular artery as part of treatment for acute stroke. The usage of the data was approved by the internal review board (IRB) of our institution. This study uses a total of 38 patients, with the training set composed of 30 patients and the test set composed of 8 patients. Each patient has a total of 360 time series CT scans, split between 6 slice locations for 60 time points each. Additionally, each patient has a total of 24 perfusion maps, split between 6 slice locations and 4 perfusion parameters. For this study, the focus is on the results achieved by the models when predicting the perfusion parameters rCBV and rCBF.

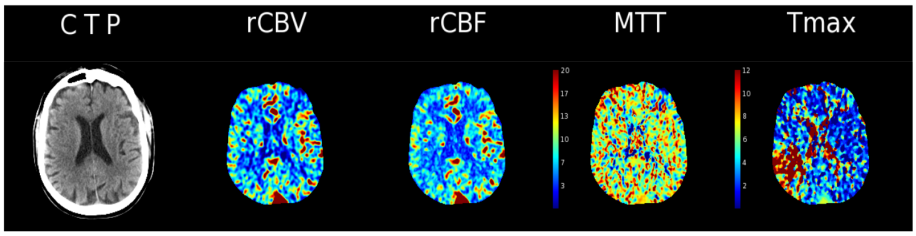


Fig. 2: Illustration of source CTP image (1st acquisition) with corresponding computed perfusion maps, including rCBV, rCBF, MTT, and Tmax using deconvolution.

2.3 Pre-processing

CT Processing All 360 time series CT scans for each patient are extracted as 2D floating point arrays from their DICOM files. The pixel intensity values are subsequently transformed to Hounsfield Units (HU), using linear rescaling

$$v = mx + b$$

where v is the resultant pixel intensity in HU, m is the rescale slope, x is the original pixel intensity, and b is the rescale intercept.

The timeseries CT scans are then median filtered with a size 3 kernel to reduce intracranial noise, and then windowed on an HU range from 0-600. This range was chosen as it encompasses the information contained from the bolus of a contrast agent. Each scan is normalized between 0-1 to remove variance that exists between different slice locations and patients. To reduce uncertainty generated via patient cranial movement, 3D volumetric rigid coregistration is applied on each patient's set of CT scans, where each volumetric slice is aligned, using an affine transformation.

Grayscale Conversion All 24 perfusion maps for each patient are converted to grayscale, as the time series CT scans are represented using grayscale intensities. Due to the nature of the JET color map, a standardized conversion formula from RGB to grayscale would fail to accurately retain the neurovascular information, as the maximum and minimum intensities would converge to the same grayscale intensity index. This problem is bypassed by implementing a manual lookup table (LUT), where each RGB pixel value is measured against the LUT. A grayscale index is selected from the LUT based off the entry with the minimum L2 distance to the RGB pixel value.

Sampling Process The training set is constructed by randomly sampling 25,000 pixels from the timeseries CT scans, spread evenly across all 30 train-

ing patients. Each point is sampled randomly, and without replacement, from a region in the scan containing only non-zero intensities, so as to limit the chance of selecting immaterial pixels. The set of selected points is additionally filtered by removing pixels that correspond to the background in the ground truth scan. For a single patient's slice location, the sampled pixel is not a single intensity value but an array containing the spatial location's intensity values for all time points. The data set is adequately represented by a 2D array of dimensions 25,000 by 60, as each pixel is an array of 60 time-varying intensities. The indices of valid points are saved, and then resampled from the perfusion maps to generate the ground truth, represented by a 1D array of 25,000 intensities.

The testing set is constructed by randomly sampling 10,000 pixels from the time series CT scans, spread evenly across all 8 testing patients. Each point is subject to the same requirements as to the training set, to maintain model and result consistency. The indices of valid points are saved in this model, to generate easy access when constructing the ground truth for the testing set. The ground truths are represented by a 1D array of 10,000 intensities spread across all 8 testing patients. For this study, the model used the training set to output prediction scans of the perfusion parameters rCBV and rCBF.

2.4 Training and Deployment

Overview This study utilizes four different types of regression based machine learning models to predict rCBV and rCBF, specifically kernel ridge regression, random forest regression, regularized linear regression (ridge regression), and linear regression. These models aim to identify general trends and varying observations from the training set in order to build an accurate prediction of the perfusion parameters.

Kernel Ridge Regression Model The model uses classic kernel ridge regression to predict the perfusion parameter intensities based on their temporal array values, using the equation below to determine the model parameters

$$\theta = (X^T X + \lambda I)^{-1} X^T y \tag{1}$$

where X is the input vector, y is the corresponding pixel output, and λ is the regularization. Note that unlike linear regression, kernel ridge regression aims to provide a more robust testing outlook by introducing both a regularization parameter to combat overfitting and by implementing a kernel transformation on the input, such that the model can predict nonlinear trends by transforming the data into the kernel space. The regularization parameter and kernel standard deviation used to train the model was 0.1 and 8.0, respectively.

Random Forest Regression Model The model uses an aggregation of singular decision trees, with each tree having its own weight on the influence of the model prediction. For this study, the model utilized 90 decision trees, each with

a maximum tree depth of 55. By introducing more trees as opposed to its singular component, the decision tree regression model, the random forest is able to combat severe overfitting on the training set and allows for weighted pixel prediction for the output pixel.

Ridge Regression Model Linear ridge regression aims to improve upon the basic linear regression model by introducing a regularization parameter, in an effort to weight the model parameters in the cost function analysis. The model cost function is determined by the equation below

$$J = \|y^{(i)} - f(x^{(i)})\|_2^2 + \lambda\|\theta\|_2^2$$

where J is the cost, $y^{(i)}$ is the ground truth value, $f(x^{(i)})$ is the predicted value, λ is the regularization parameter, and θ are the model weights. The regularization parameter of 0.1 was used to train this model.

Linear Regression Model The baseline linear regression model aims to predict pixel intensity values from the input intensities based on the following equation

$$y^{(i)} = \theta x^{(i)} + b$$

where $y^{(i)}$ represents the predicted output pixel, θ the model parameters, $x^{(i)}$ the input pixel intensity, and b the bias value.

3 Experiments

In order to assess the accuracy of the model, the predicted intensities for each pixel was compared to its corresponding intensity in the ground truth perfusion scan, using the SSIM as a measurement of each model's accuracy. The SSIM is typically used to determine the similarity in image structure and quality between a reference and processed image. The metric considers image degradation as a loss in visual information. We report the SSIM scores of various machine learning models and visualize their predicted rCBV and rCBF perfusion parameter maps compared to the ground truth maps.

Validity of Machine Learning Model Each machine learning model underwent hyperparameter testing to optimize results. The regression model was trained on a range of data points, from 5,000 to 25,000, with each point containing an array of 60 intensity values for each patient, and outputs the predicted perfusion map. The accuracy of these predictions was compared by their SSIM scores for the rCBV and rCBF perfusion parameters, which we display below. The most accurate model used kernel ridge regression paired with a Gaussian kernel. The kernel ridge regression model performs adequately in regards to both the training and testing set, with a training SSIM of 82.71% and testing SSIM of 82.53% at 25,000 training points.

Tables 1 and 2 illustrate the overall performance of the regression-based machine learning models for predicting rCBV and rCBF, respectively, both in the training set and the testing set. Tables 3 and 4 illustrate the results for the prediction of rCBV and rCBF for 8 representative patients. The effect of smoothing on the overall accuracy can be visualized from column 3 to 5 where an increasingly large Gaussian filter was applied.

Figures 3 and 4 provide a visualization of the accuracy (in terms of SSIM) with respect to the number of training data points. Figure 4 shows that while Kernel ridge and random forest regression can take advantage of a large number of training samples (best results are obtained with 22,500 samples), regularized and standard linear regression reach their best accuracy around 15,000 samples.

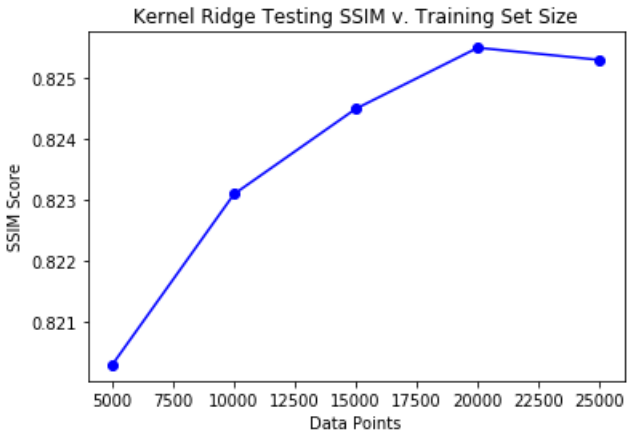


Fig. 3: Effect of training set size on kernel ridge SSIM scores, for regularization parameter 0.1 and gamma value 8.0.

Model Type	rCBV Training SSIM	rCBV Testing SSIM
Linear Regression	74.64%	77.01%
Ridge Regression	74.34%	76.89%
Kernel Ridge Regression	82.71%	82.53%
Random Forests Regression	81.14%	82.59%

Table 1: Average rCBV SSIM scores from batch model testing for all slice locations on the 8 testing patients.

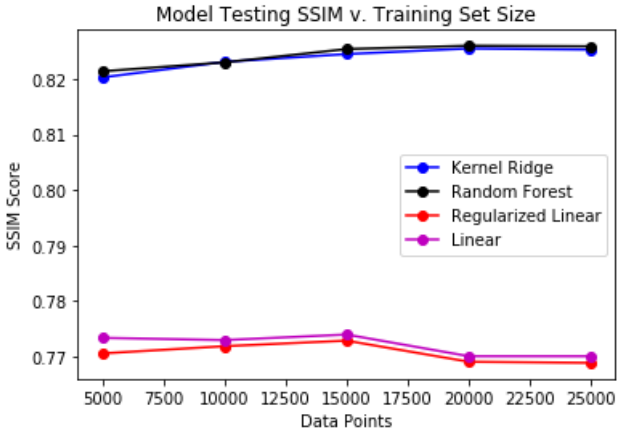


Fig. 4: Effect of training set size on all model testing SSIM scores.

Model Type	rCBF Training SSIM	rCBF Testing SSIM
Linear Regression	74.55%	77.46%
Ridge Regression	74.51%	77.67%
Kernel Ridge Regression	82.83%	83.25%
Random Forests Regression	82.12%	84.19%

Table 2: Average rCBF SSIM scores from batch model testing for all slice locations on the 8 testing patients.

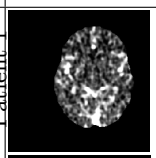
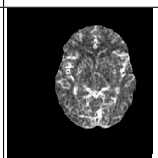
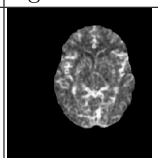
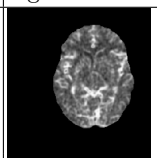
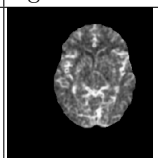
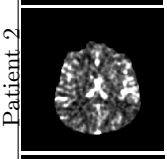
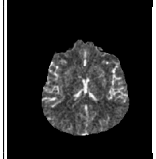
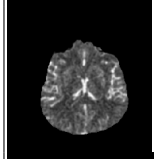
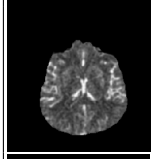
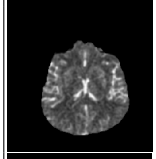
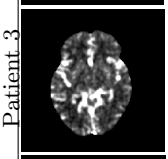
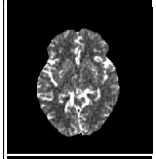
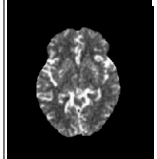
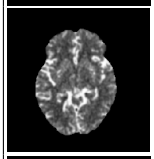
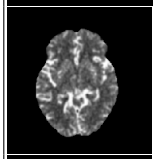
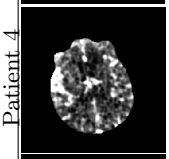
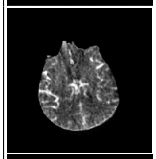
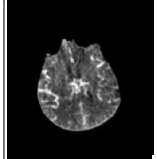
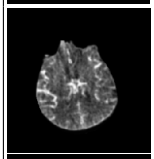
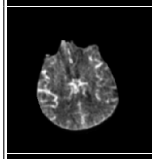
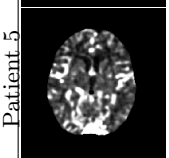
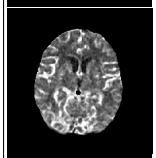
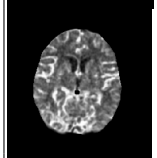
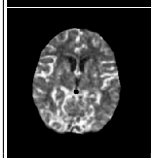
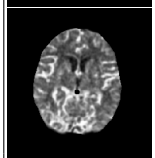
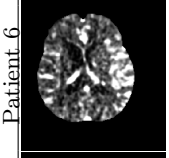
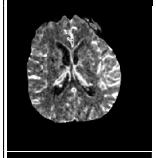
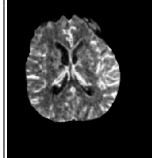
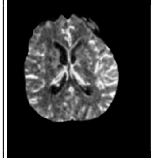
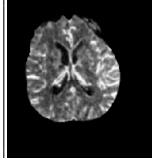
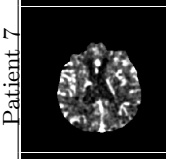
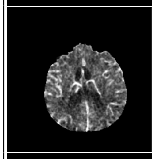
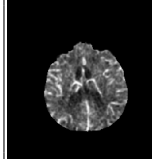
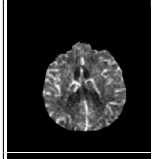
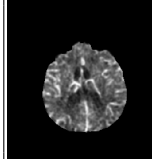
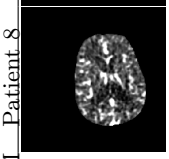
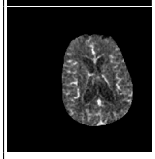
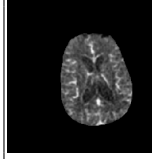
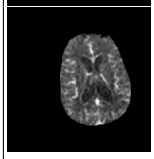
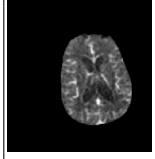
	Ground Truth	No Filter	Sigma= 0.9	Sigma= 1.0	Sigma= 1.1
Patient 1					
Patient 2					
Patient 3					
Patient 4					
Patient 5					
Patient 6					
Patient 7					
Patient 8					
Average SSIM	N/A	83.45%	84.79%	84.77%	84.70%

Table 3: Visualization of Gaussian filtering on kernel ridge regression rCBV perfusion predictions on slice location 3 for the 8 testing patients, trained on 25000 data points.

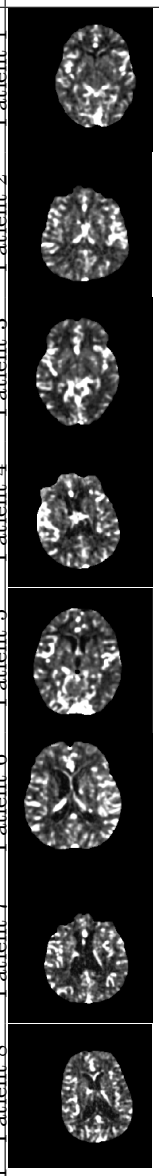
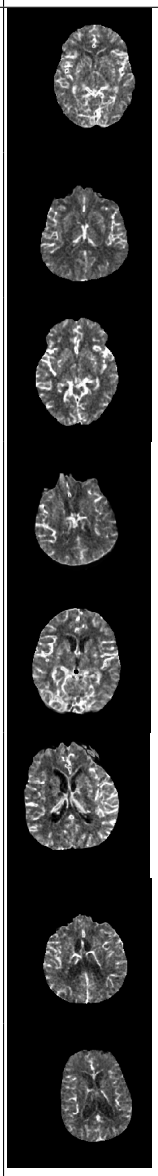
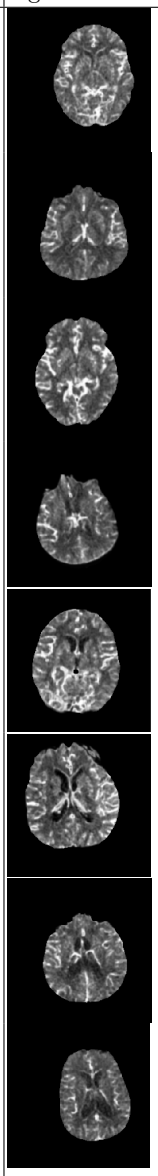
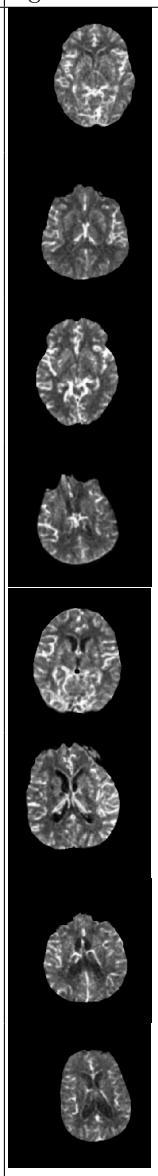
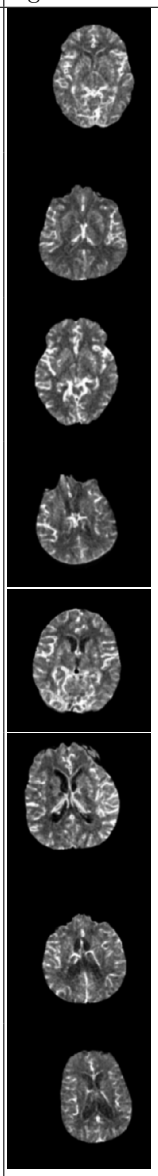
	Ground Truth	No Filter	Sigma= 0.9	Sigma= 1.0	Sigma= 1.1
Patient 1					
Patient 2					
Patient 3					
Patient 4					
Patient 5					
Patient 6					
Patient 7					
Patient 8					
Average SSIM	N/A	84.22%	86.00%	86.09%	83.25%

Table 4: Visualization of Gaussian filtering on kernel ridge regression rCBF perfusion predictions on slice location 3 for the 8 testing patients, trained on 25000 data points.

4 Discussion

There are numerous advantages to applying machine learning to CT imaging techniques. Machine learning models create consistent and reproducible results for analysis, far superior to conventional methods. Proper supervision of the training process entails successful treatment and diagnosis methods, which can become more widespread and reliable for medical use. Specifically, the prediction of perfusion parameters by machine learning models allows near instantaneous and convenient treatment solutions, without relying on expensive or inconvenient industry programs.

To accurately supervise the training process, the size and quality of the data set is the chief concern. CT image data was pruned to remove faulty and unreliable patients, such as those with excessive cranial movement or containing mismatched CT slice locations. This removal allowed the training process to occur only on patients with consistent scans and minimize variance between patients. The process to create the model was not trivial, and for 38 patients took less than 20 minutes. However, since the model only needs to be generated once, the time for construction can be disregarded. Processing of the model scales linearly with more training data, although this represents a trade-off between fitting the data and model runtime. Additionally, excessive data can give rise to the machine learning problem of overfitting, which occurs if the model fits the training data too well, at the cost of failing to capture and generalize trends.

For this study, the model is able to predict the spatio-temporal intensities of perfusion parameters associated with the intracranial artery. The framework of this study demonstrates that our model can learn these trends automatically, and predict the results with reasonable accuracy. The combination of this model with automatic imaging and training techniques allows for the possibility of a pipeline to process and predict perfusion parameters for any single patient.

Finally, deep learning methods such as neural networks and long short-term memory models could be considered to improve the results of the study. These models will, however, require a much larger data set and significantly more runtime to process.

5 Conclusion

We introduced a framework to utilize machine learning-based models for the computation of neurovascular perfusion parameters from patient CT scans. Overall, the use of machine learning models in automating the processing of patient data has been proven to be a viable alternative to current commercial methods based on deconvolution. This study concluded that machine learning models might improve accuracy in prediction and represent the potential for hospital cost cutting, which can significantly aid neurologists and neurosurgeons on a variety of neurological conditions, including acute stroke. Such results could monumentally facilitate the prediction of outcome based on raw CT perfusion.

References

1. Campbell, B.C., Yassi, N., Ma, H., Sharma, G., Salinas, S., Churilov, L., Meretoja, A., Parsons, M.W., Desmond, P.M., Lansberg, M.G., Donnan, G.A., Davis, S.M.: Imaging selection in ischemic stroke: Feasibility of automated ct-perfusion analysis. International Journal of Stroke **10** (2015) 51–54
2. Vagal, A., Wintermark, M., Nael, K., Bivard, A., Parsons, M., Grossman, A.W., Khatri, P.: Automated CT perfusion imaging for acute ischemic stroke: Pearls and pitfalls for real-world use. Neurology **93** (2019) 888–898
3. Tong, E., Sugrue, L., Wintermark, M.: Understanding the Neurophysiology and Quantification of Brain Perfusion. Top Magn Reson Imaging **26** (2017) 57–65
4. et al., T.D.F.: Use of magnetic resonance imaging to predict outcome after stroke: A review of experimental and clinical evidence. Journal Of Cerebral Blood Flow and Metabolism **30(4)** (2010) 703–717
5. Kudo, K., Sasaki, M., Yamada, K., Momoshima, S., Utsunomiya, H., Shirato, H., Ogasawara, K.: Differences in CT perfusion maps generated by different commercial software: quantitative analysis by using identical source data of acute stroke patients. Radiology **254** (2010) 200–209
6. Andersen, I.K., Szymkowiak, A., Rasmussen, C.E., Hanson, L.G., Marstrand, J.R., Larsson, H.B., Hansen, L.K.: Perfusion quantification using Gaussian process deconvolution. Magn Reson Med **48** (2002) 351–361
7. Calamante, F., Gadian, D.G., Connelly, A.: Quantification of bolus-tracking MRI: Improved characterization of the tissue residue function using Tikhonov regularization. Magn Reson Med **50** (2003) 1237–1247
8. Mouridsen, K., Friston, K., Hjort, N., Gyldensted, L., ?stergaard, L., Kiebel, S.: Bayesian estimation of cerebral perfusion using a physiological model of microvasculature. Neuroimage **33** (2006) 570–579
9. Vonken, E.P., Beekman, F.J., Bakker, C.J., Viergever, M.A.: Maximum likelihood estimation of cerebral blood flow in dynamic susceptibility contrast MRI. Magn Reson Med **41** (1999) 343–350
10. Boutelier, T., Kudo, K., Pautot, F., Sasaki, M.: Bayesian hemodynamic parameter estimation by bolus tracking perfusion weighted imaging. IEEE Trans Med Imaging **31** (2012) 1381–1395
11. Wu, D., Ren, H., Li, Q.: Self-supervised dynamic ct perfusion image denoising with deep neural networks (2020)
12. Zhu, H., Tong, D., Zhang, L., Wang, S., Wu, W., Tang, H., Chen, Y., Luo, L., Zhu, J., Li, B.: Temporally downsampled cerebral CT perfusion image restoration using deep residual learning. Int J Comput Assist Radiol Surg **15** (2020) 193–201
13. Stier, N., Vincent, N., Liebeskind, D., Scalzo, F.: Deep Learning of Tissue Fate Features in Acute Ischemic Stroke. Proceedings (IEEE Int Conf Bioinformatics Biomed) **2015** (2015) 1316–1321
14. Scalzo, F., Hao, Q., Alger, J.R., Hu, X., Liebeskind, D.S.: Regional prediction of tissue fate in acute ischemic stroke. Ann Biomed Eng **40** (2012) 2177–2187
15. Yu, Y., Guo, D., Lou, M., Liebeskind, D., Scalzo, F.: Prediction of hemorrhagic transformation severity in acute stroke from source perfusion mri. IEEE Transactions on Biomedical Engineering **65** (2018) 2058–2065
16. McKinley, R., Hung, F., Wiest, R., Liebeskind, D.S., Scalzo, F.: A Machine Learning Approach to Perfusion Imaging With Dynamic Susceptibility Contrast MR. Front Neurol **9** (2018) 717
17. Feghhi, E., Zhou, Y., Tran, J., Liebeskind, D., Scalzo, F.: Angio-ai: Cerebral perfusion angiography with machine learning. In: ISVC. (2019)