Random Forest and Gradient Boosted Trees for Patient Individualized Contrast Agent Dose Reduction in CT Angiography

René Pallenberg, Marja Fleitmann, Andreas Martin Stroth, Jan Gerlach, Alexander Fürschke, Jörg Birkhausen, Arpad Bischof, and Heinz Handsels

1 University of Lübeck, Institute of Signal Processing, Germany
2 University of Lübeck, Institute of Medical Informatics, Germany
3 German Research Center for Artificial Intelligence, Artificial Intelligence in Medical Imaging, Lübeck, Germany
4 University Medical Center Lübeck, Department of Radiology and Nuclear Medicine
5 IMAGE Information Systems Europe, Rostock, Germany

Abstract. This work aims to recognize the patient individual possibility of contrast dose reduction in CT angiography. This system should help to identify whether the dose of contrast agent in CT angiography can be reduced to avoid side effects. In a clinical study, 263 CT angiographies were performed and, in addition, 21 clinical parameters were recorded for each patient before contrast agent administration. The resulting images were labeled according to their contrast quality. It is assumed that the contrast dose could be reduced for CT angiography images with excessive contrast. These data was used to develop a model for predicting excessive contrast based on the clinical parameters using logistic regression, random forest, and gradient boosted trees. In addition, the minimization of clinical parameters required was investigated to reduce the overall effort. Therefore, models were tested with all subsets of clinical parameters and each parameter’s importance was examined. In predicting excessive contrast in CT angiography images covering the aortic region, a maximum accuracy of 0.84 was achieved by a random forest with 11 clinical parameters; for the leg-pelvis region data, an accuracy of 0.87 was achieved by a random forest with 7 parameters; and for the entire data set, an accuracy of 0.74 was achieved by gradient boosted trees with 9 parameters.

Keywords. Random Forest, Gradient Boosted Trees, Logistic Regression, CT angiography, Contrast

1 Corresponding Author: René Pallenberg, E-mail: r.pallenberg@uni-luebeck.de.
1. Introduction

Computed tomography angiography (CTA) is an important method for detecting diseases such as tumors, pulmonary embolism and vascular stenosis [1]. Administration of contrast agents is necessary for CTA. Many of these contrast agents contain iodine, which may harm the patient’s body. Observed side effects are allergic reactions, hyperthyroidism, and deterioration of renal function up to renal insufficiency [2].

One way to avoid side effects is to reduce the administered dose of contrast agents [3]. Currently, standard doses are usually administered in everyday clinical practice. The dose is often higher than required for sufficient contrast quality. This is to avoid generating an insufficiently contrasted image, which would result in repetition and thus in a renewed exposure of the patient.

In this paper, we present methods to recommend contrast dose reduction for patients on an individual basis. To do this, the system is trained to predict whether the contrast quality of a CTA image is “excessive” based on clinical parameters recorded before contrast agent administration [4]. It is assumed that if a CTA image had excessive contrast, the contrast dose could have been reduced [5]. It is important to achieve high precision to ensure that no CTA images are created with insufficient contrast. In addition, the influence of each clinical parameter in the prediction models was investigated.

2. Material and Methods

2.1. Data

A clinical study was conducted by the Department of Radiology and Nuclear Medicine at the University Hospital Schleswig-Holstein in Lübeck, Germany [4]. In this study, additional clinical parameters were recorded from patients who had a CTA performed as part of their treatment. They were recorded before the administration of the contrast agent.

These parameters were: ankle-brachial index (ABI), age (AGE), body mass index (BMI), blood pressures diastolic (BPD) and systolic at rest (BPS), 5 min (BPD5, BPS5) and directly (BPD0, BPS0) before administration, creatine (C), gender (G), glomerular filtration rate (GFR), γ-glutamyltransferase (GGT), height (H), hematocrit (HC), hemoglobin concentration (HB), oxygen saturation (OS), Pulse directly (P0) and 5 min (P5) before administration, weight (W) and waist circumference (WS). In total, the data set comprises 263 CTA images and the corresponding clinical parameters. The images were labeled as “excessive” based on the quality of their contrast [5-7]. Excessive means that the contrast is higher than necessary. To determine this, the mean HU values of selected regions of interest were determined by a radiologist [6]. The data set includes CTA images of eight different recording protocols covering different body regions.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Not Excessive</th>
<th>Excessive</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aorta (AORTA)</td>
<td>22</td>
<td>68</td>
<td>90</td>
</tr>
<tr>
<td>Leg and pelvic vessel (LPV)</td>
<td>48</td>
<td>22</td>
<td>70</td>
</tr>
<tr>
<td>All protocols combined (COMB)</td>
<td>107</td>
<td>156</td>
<td>263</td>
</tr>
</tbody>
</table>

Table 1. Distribution for the used data sets.
Three data sets were used for the experiments, one containing the aortic protocol (AORTA), one the leg and pelvic vessel protocol (LPV), and one all eight protocols combined (COMB), see Table 1.

2.2. Classification

To predict whether the contrast of a CTA image is “excessive” or not, three classification methods widely used in tabular data in medicine and biology were applied [8,9]: Random Forest, Gradient Boosted Trees, and Logistic Regression as established reference method [10-12]. The 21 clinical parameters of the patient belonging to the CTA served as input for the classifiers. Training and evaluation were performed using leave-one-out cross validation, for each of the three methods. A grid search was performed to optimize the hyperparameters of the classifiers. Accuracy was chosen as the optimization criterion. A preselection of input parameters was investigated to further improve the prediction and reduce the number of required clinical parameters. Therefore, the methods were tested for all possible subsets of the 21 clinical parameters. In addition, the importance of each clinical parameter for prediction was assessed by permutation and Gini importance [13]. The importance was predicted separately for the 50 models that achieved the highest accuracy on the respective dataset. For the overall importance, the results of these 50 models were averaged.

3. Results and Discussion

3.1. Classification Performance

The described evaluation procedure was done separately for the two subsets AORTA and LPV and also for all available data (COMB). Due to the unbalanced distribution of classes in the data sets, the Matthews correlation coefficient (MCC) was used as a measure of model performance. In Figure 1 the highest achieved MCC is related to the associated values for precision and the number of clinical parameters used.

This work should improve previous prediction methods of contrast quality. In [4] we show that a k-nearest neighbor method achieved an MCC of 0.34, an accuracy of 0.78 and a precision of 0.85 on the AORTA data set. By applying random forest, gradient boosted trees, and logistic regression and reducing the clinical parameters used, higher accuracy was achieved for the AORTA data set. The highest accuracy (0.84), MCC (0.54), and precision (0.86) were achieved by random forest using 11 clinical parameters (ABI, AGE, G, GFR, GGT, H, HB, P0, P5, W, WS). With the LPV data set, both gradient boosted trees with 7 (ABI, BMI, BPS, G, OS, P5, W) and random forest with 7 (ABI, AGE, BPD5, BPS5, GFR, HB, WS) clinical parameters achieved an MCC of 0.70 with a precision of 0.93 and an accuracy of 0.87. On the whole data set (COMB), the maximum MCC was 0.46 with a precision of 0.75 and an accuracy of 0.75, achieved by gradient boosted trees and 9 (ABI, AGE, BMI, BPD, BPD5, GFR, GGT, HC, OS, W) clinical parameters.

This shows that logistic regression cannot compete with the other two methods. On all data sets, the best results were achieved with 5-11 parameters. In addition to the reduced effort, the reduction of the 21 parameters also offers quality improvements. The differences in results between the data set LPV and AORTA can be explained by the distribution of the data. It is also seen that the results for COMB are much less.
3.2. Importance of the Clinical Parameters

To further investigate the relationship between the clinical parameters and the quality of contrast in the CTA images permutation and Gini importance was computed [13]. When looking at Figure 2, it can be seen that between 5 and 9 clinical parameters are rated as important and the others as hardly important. The results of the two methods differ only slightly in the evaluation of the importance of the parameters, which indicates the reliability of the results.

Depending on the data set, however, the important parameters differ. Also, the distance between the important and unimportant parameters is significantly greater in LPV and COMP than in AORTA. Therefore, generalization over different recording protocols is not feasible. On AORTA, AGE, BMI, BPS0, C, G, GFR, HC, P0, and WS were the most important parameters. On LPV, WS was by far the most important parameter W, BP5, and GFR were the next most important. On COMP, AGE, ABI, GFR, PD0 and W were the most important parameters.
4. Conclusion

This work aimed to develop improved recommendation models for contrast agent dose reduction. This is done by predicting excessive contrast of CTA images based on 21 clinical parameters of a patient. To build a predictive model three different methods gradient boosted trees, random forest, and logistic regression were tested. These methods were tested on three different data sets, AORTA, LPV, and COMB. Overall, a maximum MCC of 0.54 was achieved on the AORTA data set, with a precision of 0.86 and an accuracy of 0.84. This was achieved by a random forest using 11 clinical parameters. This represents an improvement to the results of our previous work, where a maximum MCC of 0.34 was achieved [4].

According to the experts, an accuracy of more than 90% and a precision of 95% should be achieved for clinical application. Even though this work has already improved the results, they still need to be increased. Our experiments show that generalization across multiple imaging protocols leads to poorer predictions. Therefore, the models should be built for each protocol and body region. Feature selection has shown that using about 5 to 11 clinical parameters as input leads to as good or better results than using all 21 clinical parameters. The results obtained may be used to reduce the number of clinical parameters to be included in the future.

References