1. Motivation

MR image quality evaluation is mainly performed by human observers (HO) to determine the underlying image quality with respect to a certain diagnostic question. HO evaluations are time-demanding and expensive. Furthermore, the lack of a reference image makes this task challenging. In order to support the HO and automate this process, we extend our previous no-reference MR image quality assessment system [1,2] which is based on a machine-learning model observer with an active learning (AL) loop to reduce the amount of needed labeled training data.

2. MR image quality assessment system

The proposed system including the active learning loop is shown in Fig. 1. 3D and 2D multislice MR images are considered as input.

3. Active Learning

The amount of needed labels can be reduced by selecting the most meaningful ones, without redundant information for classification. We implemented two query strategies based on pool-based uncertainty sampling [3], i.e. on how certain the classifier is in his decision. For $D_t$ training data $N_0 = |D_t|$ number of training samples $D_t$ initial training set $N_1 = |D_t|$ number of initial training samples $Q$, query set $N_1 = |Q|$ number of samples per query the goal is to keep $N_0$ as small as possible with an initial training of $D_t$. In each of the $N_1$ AL loops the HO is queried to label $N_1$ samples from the query set $Q$. Furthermore, since HOs label 3D images, but the sample selection takes place on 2D slices, the selection prioritizes samples belonging to the same 3D image.

3.1 probability-based selection

Probability estimates for multi-class SVM derived from pairwise coupling as described in [4] are selected to minimize the difference between the probabilities $P_1(x_n)$ and $P_2(x_n)$ of the 1st and 2nd most probable class [5].

$$L = \frac{1}{N_1} \sum_{i=1}^{N_1} \min \{ P_1(x_n), P_2(x_n) \}.$$ (1)

3.2 distance-based selection

Uncertainty of a sample to belong to class $y_n \in \{1, \ldots, 5\}$ is determined by the distance $d(x_n)$ of a feature vector $x_n$ to the hyperplane $f(x_n)$. The to be labeled set is thus composed by

$$L = \{ |d(x_n)| = d(x_n) \forall x_n \in y_n \} \setminus (Q \cup S)$$ (2)

with the ascending distances and $Q$ denoting a set of outliers and a group $S$ considering the slack variables (Fig. 2). Distances are determined by

$$d(x_n) = \|x_n\|^2 f(x_n) = |x_n| \sum_{i=1}^{50} n_{i}^{q} k(x_n, x_i) + b$$ (3)

where $\omega$ and $b$ denote the primal parameters learned by the SVM with an RBF kernel $k(x_n, x_i)$ and dual coefficients $\alpha_i$ of $N_0$ support vectors (SV).

Outlier correction $O$: Reject samples via distance $d(x_n)$ of class $y_n$.

$$O = \{ |d(x_n)| > |d(x_n)| \}$$

$$\forall x_n \in y_n, x_n \neq x_1.$$ (4)

Slack variable corrector $S$: Discard samples with minimum distance $\delta$ to the hyperplane

$$S = \{ |d(x_n)| < \delta \}$$ (5)

Figure 2: Exemplary 2D feature space to illustrate sample selection.

4. Materials and Methods

As reported in [1], the system is able to achieve an overall test accuracy of 91.2% with the whole training set $N_0 = 2038$. The aim of this study is to reduce the labeling cost, i.e. $N_0 = N_1 + N_2 + N_3$ with $N_1$ queries, while maintaining accuracy. Results are presented as mean of ten different runs.

5. Results

![Test accuracy for initial training set sizes $N_1$ with $N_2 = 40$ samples per query.](image)

(a) probability-based approach

(b) margin-based approach

Figure 5: Test accuracy for $N_1$ samples per query with initial training size $N_1 = 200$.

6. Conclusion and Outlook

On in vivo MR data both strategies reveal that training data can be reduced by roughly 50% while achieving comparable classification results to the previous system. Furthermore by selecting only the most meaningful 2D slices belonging to a few significant 3D images, the labeling effort is reduced tremendously.

References

contact information: thomas.kuestner@iss.uni-stuttgart.de