# Supervised pattern recognition for predicting the appearance of contrast-enhancement in high-grade gliomas using multimodality MRI and MRSI

M. C. Lee[1], S. M. Chang[2], S. J. Nelson[1]

[1]Surbeck Laboratory of Advanced Imaging, Department of Radiology, University of California, San Francisco, CA, United States, [2]Department of Neurological Surgery, University of California, San Francisco, CA, United States

## 1. Introduction

By manipulating contrast mechanisms in MRI, it is possible to gain insight into the physiological status of brain tumors and surrounding tissues. However, the complexity of multimodality imaging complicates visual inspection and leaves open the possibility that valuable but complex patterns may exist undiscovered in the data. In this work, we study the combined use of spectroscopic, diffusion, and perfusion imaging in predicting growth of contrast-enhancing gliomas over the course of radiotherapy. We apply a *k*-nearest neighbor pattern recognition algorithm and use a genetic algorithm for optimal feature selection and weighting. Similar techniques have previously been used for classifying spectra, predicting voxel fates in stroke[1], and identifying parameters useful in characterizing mammographic masses.[2] We hypothesize that the pre-radiation images contain information that can be used to predict where the contrast enhancing (CE) lesion will appear after radiotherapy. By applying feature selection and weighting, we are able to improve the accuracy of the classification and gain insight into the relevance of various MR measured indices.

## 2. Materials and Methods

Twenty-four patients (17 male, 7 female, age 27–75) with grade IV gliomas received MRI examinations on a 1.5 T GE Echospeed scanner within two weeks following surgery but before radiotherapy. A second MRI was performed within a month of the conclusion of radiotherapy. The MR examination included T1-SPGR with Gd-DTPA contrast, 3D $^1$H-MRSI (TR/TE = 1000/144 ms; 12 × 12 × 8; 1 cm$^3$), diffusion imaging ($b$ = 1000 s/mm$^2$; 1.5 × 1.5 × 2.2 mm$^3$), and dynamic susceptibility-contrast perfusion imaging (TR/TE = 1000/54 ms; 10 mmol/kg Gd-DTPA; 2 × 2 × 3–6 mm$^3$). The spectral amplitudes and linewidths of choline (Cho), creatine (Cr), *N*-acetyl-aspartate (NAA), lactate (Lac), and lipids (Lip) were quantified. Abnormality indices were derived for the Cho/NAA, Cr/NAA, and Cho/Cr ratios (CNI, CrNI, and CCrI, respectively).[3] The first-pass perfusion data were modeled as a gamma-variate, yielding CBV as well as curve shape parameters such as time-to-peak (TTP), peak width (FWHM), and recovery to baseline. The ADC and anisotropy were derived from the diffusion images. Data were aligned (non-rigidly, if needed) and resampled to the perfusion resolution. Data were expressed in three ways: normalized to normal-appearing white matter, normalized then standardized, and, if applicable, as a raw SNR, and in all cases, scaled to equal ranges. A total of 36 parameters were compiled for every voxel which were used to predict the post-therapy CE status of each voxel.

The classifier used in this study was a *k*-nearest neighbor (*k*-NN) algorithm. Briefly, the *k* closest training points (in a Euclidean space) vote on the classification of each unknown test point. The *k*-NN algorithm is sensitive to parameter scaling and the presence of irrelevant or noisy parameters. To improve classification accuracy, we use *feature selection*, in which only some of the parameters are used, and *feature scaling*, in which the parameters are scaled so that their numerical range better matches their relevance. Both approaches are applied and optimized simultaneously using a genetic algorithm.[4] Genetic algorithms 'evolve' solutions through a stochastic search, and are ideal for combinatorial problems. During each of the 500 generations, an exhaustive search of *k* values was performed for each of 500 individuals. Uniform crossover was used for masking, and continuous uniform crossover for the weights. The optimization was implemented in C and parallelized using the MPICH library on a grid-enabled cluster of 24 Intel Xeon processors running at 2.8 GHz.

Patients were randomly assigned to a design (13) and verification group (11). During the design phase, the classifier was applied to each design patient, using training data from all other design patients. For each patient, a receiver-operating characteristic (ROC) curve was estimated using 200 threshold values and the area under the curve (AUC) computed using trapezoidal integration. The goal of the genetic algorithm was to select a set of features and feature weights to minimize the mean AUC across the design patients. Once optimal feature weights were chosen, the algorithm was tested using the verification patients, with the 13 design patients comprising the training set. Five verification patients did not have enough new CE for meaningful ROC curves (less than ten voxels of new CE), and were not included in the analysis of new CE predictions.

TABLE I. AUC† OF PREDICTORS (VERIFICATION)

| Pre-existing CE | | Previously non-CE | |
|---|---|---|---|
| Pred. | AUC | Pred. | AUC |
| *k*-NN | 0.66 ± 0.12 | *k*-NN | 0.88 ± 0.04 |
| Lip | 0.61 ± 0.18 | CNI | 0.86 ± 0.10 |
| TTP | 0.56 ± 0.14 | CrNI | 0.78 ± 0.22 |
| Lac | 0.63 ± 0.13 | NAA | 0.79 ± 0.12 |

† Probability that a random pair of enhancing and non-enhancing voxels will be correctly ordered (mean ± SD across patients).

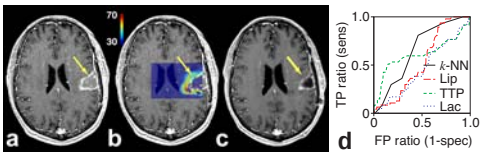


**Figure 1.** Example post-contrast T1 images (a) pre-radiation, (b) with prediction overlay, (c) post-radiation. (d) ROC curve for old CE only.

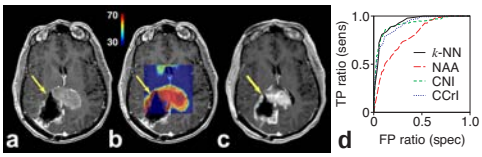


**Figure 2.** Same as fig. 1, for verification patient; ROC for new CE only.

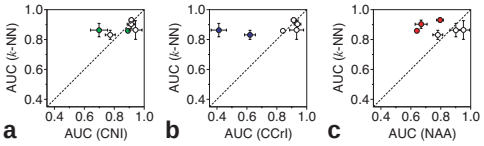


**Figure 3.** Comparison of *k*-NN and thresholding for predicting new CE: (a) Cho-NAA index, (b) Cho-Cr index, (c) NAA. Each point is a single verification patient. Filled circles are significant differences ($P < 0.01$).

## 3. Results and Discussion

The mean AUC of the verification patients was 0.875 ± 0.035 (range 0.859–0.930) when predicting new CE in previously non-enhancing regions, and 0.662 ± 0.118 (0.480–0.798) for predicting persistence and disappearance of CE in regions of pre-existing enhancement. Example patients are shown in figures 1 and 2. A region of disappearing CE is indicated by the arrow in figure 1, while a region of new CE growth is indicated by the arrow in figure 2. Using the same criteria of maximizing the mean AUC of the design patients, the three individual parameters most predictive of new CE were CNI, CCrI, and NAA. Similarly, for the fate of pre-existing CE, the most predictive features were Lac, TTP, and Lip. As shown in table I and figure 3, the optimized *k*-NN algorithm outperformed thresholding with these parameters in most patients. The mean AUC difference was smallest with CNI; however, the inter-patient variation is much smaller using the optimized *k*-NN than CNI. Note that in this study, the AUC can be less than 0.5, because the directionality of the thresholding is determined beforehand by the design patients.

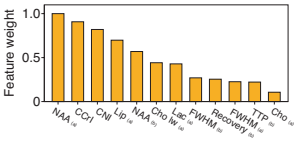


**Figure 4.** Weights of selected features found using the optimized *k*-NN. Features were (a) normalized or (b) standardized to NAWM.

Optimized weights and selected parameters are shown in figure 4. The selection of particular features can be interpreted as follows: voxels with similar values for these features before radiotherapy also tend to have similar CE status after therapy. Note that the same set of feature weights were used whether voxels were enhancing pre-therapy or not. In both cases, the optimized classifier exceeds the performance of single-variable thresholding; this suggests that the underlying biology of persistent and newly-appearing enhancement may be similar, and that the pre-therapy examination may simply be capturing the tissue at two stages of the same progressive time-course. However, the gap in performance between pre-existing and new CE suggests that training separate classifiers for each region may be useful.

## 4. Conclusions

Techniques of pattern recognition and evolutionary computing have previously been used in computer-aided diagnosis applications, such as in mammography and stroke imaging. The primary contribution of this work is to apply these methods to assess the importance of MRI derived parameters in predicting new contrast enhancement. Selecting features is useful in improving the robustness of feature weighting algorithms and improving the comprehensibility of the results. We believe that in addition to lending insight into parameter relevance, such algorithms can be applied to aid radiological interpretation of complex multimodality datasets.

[1] Gottrup C *et al. Artif Intell Med.* 33: 223–36 (2005).
[2] Sahiner B *et al. Med Phys.* 23: 1671–84 (1996).
[3] McKnight TR *et al. J Magn Reson Imaging.* 13: 167–77 (2001).
[4] Raymer ML *et al. IEEE Trans Evolut Comp.* 4: 164–71 (2000).

*This work was funded by NIH/NCI P50 CA97257 and F32 CA105944.*