



Published in final edited form as:

Comput Med Imaging Graph. 2013 June ; 37(4): 281–292. doi:10.1016/j.compmedimag.2013.04.003.

Statistical Learning Algorithm for In-situ and Invasive Breast Carcinoma Segmentation

Jagadeesan Jayender, Eva Gombos, Sona Chikarmane, Donnette Dabydeen, Ferenc A. Jolesz, and Kirby G. Vosburgh[†]

Department of Breast Radiology, Surgical Planning Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

Abstract

DCE-MRI has proven to be a highly sensitive imaging modality in diagnosing breast cancers. However, analyzing the DCE-MRI is time-consuming and prone to errors due to the large volume of data. Mathematical models to quantify contrast perfusion, such as the Black Box methods and Pharmacokinetic analysis, are inaccurate, sensitive to noise and depend on a large number of external factors such as imaging parameters, patient physiology, arterial input function, fitting algorithms etc., leading to inaccurate diagnosis. In this paper, we have developed a novel Statistical Learning Algorithm for Tumor Segmentation (SLATS) based on Hidden Markov Models to auto-segment regions of angiogenesis, corresponding to tumor. The SLATS algorithm has been trained to identify voxels belonging to the tumor class using the time-intensity curve, first and second derivatives of the intensity curves (“velocity” and “acceleration” respectively) and a composite vector consisting of a concatenation of the intensity, velocity and acceleration vectors. The results of SLATS trained for the four vectors has been shown for 22 Invasive Ductal Carcinoma (IDC) and 19 Ductal Carcinoma In Situ (DCIS) cases. The SLATS trained for the velocity tuple shows the best performance in delineating the tumors when compared with the segmentation performed by an expert radiologist and the output of a commercially available software, CADstream.

Keywords

Computer-aided diagnosis; Hidden Markov Models; DCE-MRI; Invasive Ductal Carcinoma; Ductal Carcinoma In Situ; Statistical Learning Algorithm

1 Introduction

Nearly 200,000 new cases of invasive breast cancer were expected to be detected in the US in 2011 (<http://www.breastcancer.org/>). Early diagnosis and treatment of breast cancer has proven to be beneficial in improving the survival rate. Conventional diagnostic tools involve palpation and imaging techniques such as mammography. Of late, DCE-MRI has proven to

© 2013 Elsevier Ltd. All rights reserved.

^{*}Corresponding author's jayender@bwh.harvard.edu.

[†]This project was supported by the National Center for Research Resources and the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health through Grant Numbers P41EB015898 and P41RR019703 (F.A.J) and Brigham Radiology Research and Education Fund (J.J.).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

be a highly sensitive technique in the diagnosis of breast cancer and staging of therapy [1]. To create a DCE-MRI data set, vascular contrast (generally Gd-DPTA) is injected intravenously, and a time series of volumetric images is made of the breast. Tumors have a greater density of blood vessels resulting in rapid wash-in of the contrast. However, due to the leaky nature of the blood vessels, the contrast tends to wash out of the tumor. Comparatively, the surrounding breast parenchyma has normal supply of blood resulting in a slower increase in contrast. As a result, voxels within a tumor in DCE-MRI show a rapid increase in signal intensity and a subsequent decrease over time, while voxels within the healthy parenchyma show a gradual increase in signal intensity.

Several investigators have developed mathematical techniques to represent the pharmacokinetics and derive quantitative information on the signal intensity changes as the contrast medium flows into and out of each voxel. These methods can be broadly divided into two groups : Black Box Methods and Tissue-Contrast Modeling [2].

Due to inherent differences in the shape of the intensity curves corresponding to tumor and healthy voxels, black box methods quantify the type of signal intensity change in terms of metrics such as maximal enhancement [3], initial rate of enhancement [4], time to peak [5], signal enhancement ratio [6], and washout slope [7]. Several commercial systems such as CADstream have also utilized the black-box methods to evaluate the tumor maps. However, these metrics utilize only a small portion of the curve to estimate the extent of angiogenesis. In addition, these metrics are highly dependent on external factors such as imaging parameters, timing of contrast injection and image acquisition, imaging noise, patient physiology and type of tumor. In addition, Ductal Carcinoma In Situ, a precursor to malignant cancer of the breast is low enhancing. Therefore, metrics such as maximum slope are not suitable for distinguishing regions of cancer from healthy tissue.

Pharmacokinetic models mathematically describe the underlying diffusion process of contrast from the blood vessels to the extracellular space making some key assumptions. Examples include the standard and extended Tofts model [8], two compartment model, shutter speed model [9] and Brix model [10]. A recent paper [11] mathematically elucidated the shortcoming of the Tofts and extended Tofts model by demonstrating the narrow range of applicability of the model to highly perfused or weakly vascularized tissue, both of which are not necessarily applicable to breast carcinomas. Luypaert et al. [12] found that for the two compartment model, the errors in estimating the pharmacokinetic models are extremely sensitive to the sampling time and increased rapidly with increased sampling time. This leads to unreasonable requirements on the sampling time. In addition, to allow for a 20% error in the estimation of the parameters, the limits imposed on the Contrast to Noise Ratio is well beyond the capability of modern day scanners. Litjens et al. [13] have shown that the Brix model performance is suboptimal over all sampling rates and does not adequately model the diffusion process. The accuracy of the pharmacokinetic models also greatly depend on the fitting algorithms, which are highly sensitive to imaging artifacts such as inhomogeneities, noise and limited temporal resolution. These fitting algorithms may not converge at all, or converge to a different local minima [14].

Several papers have also utilized clustering algorithms such as fuzzy c-means [15], artificial neural networks [16], PCA/ICA based methods [17] etc. to differentiate regions of tumor from healthy parenchyma. However, most clustering methods still rely on scalar heuristic parameters obtained from the intensity curves, thereby suffering from the same problems of the black-box methods. Noise and imaging inhomogeneities affect the accuracy of such clustering methods.

Hidden Markov Models (HMMs) have proven to be a useful tool in identifying and predicting patterns in noisy data. Some applications which use HMMs include speech recognition [18], detecting protein homologies [19], gesture recognition [20], predicting electricity prices [21], modeling cardiac arrhythmia [22], etc. A variation of the HMM has been used in [23] to segment prostate cancer from multispectral MR prostate datasets. Gong and Brady [24] utilized the Hidden Markov Random Measure Field to simultaneously register and segment breast carcinomas from DCE-MRI. However, in their framework they have integrated the texture information and a Markov Random Field model along with a pharmacokinetic model to segment the tumors. This framework would be sensitive to the same problems as the pharmacokinetic models. In a recent paper, probabilistic models based on HMMs have shown great promise in segmenting and estimating breast density from breast tomosynthesis images [25].

We have developed a novel Statistical Learning Algorithm for Tumor Segmentation (SLATS) based on the discrete form of the HMM. SLATS automatically segments regions of angiogenesis corresponding to tumor from healthy parenchyma without requiring the user to specify a training set for the HMMs. An initial version of this paper was presented at the IEEE International Conference on Biomedical Imaging (ISBI), 2012 [26]. In this paper, we have further trained the HMMs using the derivatives of the time-intensity curves and compared the results of SLATS for the different trajectories. The classification of the pixels into healthy and tumor classes is done in the frequency domain on short time intervals (using Short Time Fourier Transform), effectively suppressing frequencies outside the time window, making it less sensitive to noise and biases, thereby overcoming a major problem associated with the black-box method. Also, the entire data contained in the time-intensity curve is utilized in contrast to conventional methods of reducing the information to a single scalar metric. Since the method is independent of any underlying physiological model, it can be widely utilized to segment structures from 4D dynamic images and also avoids the limitations imposed by the pharmacokinetic models.

The paper has been organized as follows: Section II contains the description of SLATS, Sections III describes the methods used to load the DCE-MRI and perform the analysis, Section IV contains the results of the algorithm for 22 IDC and 19 DCIS cases along with results for DCE-MRI with greater time points and robustness analysis and finally, Section V has some concluding remarks.

2 Tumor Segmentation Algorithm

Key techniques used in SLATS include the Fast Fourier transform (FFT), Short-time Fourier Transform (STFT), the fuzzy c -means clustering (FCM) algorithm and the HMM. The flow chart of the algorithm is shown in Figure 1. Our algorithm is based on modeling the transition of each voxel in time and classifying the voxels into healthy and tumor classes. First, a region of interest (ROI) is delineated by a human operator on the dynamic image data. While the results of the processing are not particularly sensitive to the choice of ROI, a smaller ROI permits faster computation, which is particularly important for image guided surgery applications. The time-intensity profile of each voxel in the ROI is provided to the FFT algorithm, and the frequency components of the time-intensity profiles are computed. The voxels are approximately clustered based on the magnitude of the FFT into “tissue type” classes using the FCM clustering algorithm. A small sample of points from the two classes of voxels are then provided to train two discrete HMMs corresponding to healthy (normal parenchyma) (λ_1) and tumor (λ_2) classes. For the remaining pixels, the probability of observing the discrete sequence is computed from HMM λ_1 and HMM λ_2 . A similarity ratio χ is then evaluated to determine the similarity to the healthy or tumor class. The ratio

χ is then converted to a color scale and overlaid with the original images to determine the position of the tumor. We now go over each component of the pipeline in more depth.

2.1 Preparing data

The time-intensity profile of all the voxels within a mask is first logged. Since there are limited number of time points, the curves are interpolated to obtain greater samples within the given 5 data points. In this paper, we have trained the HMMs to learn four different trajectories corresponding to the time-intensity curves, derivative (“velocity”) and second derivative (“acceleration”) of the time-intensity curves and a composite tuple consisting of the concatenation of the position, velocity and acceleration tuples. At each voxel i , the interpolated intensity tuple can be written as $\bar{x}_i = [x_i^0, x_i^1, \dots, x_i^n]$, where the superscript represents the time index. Each of the intensity, velocity and acceleration tuples at a voxel is an n -tuple, while the composite tuple is a $3n$ -tuple.

Next, we create a prior for input to the HMMs. The discrete Fourier transform (DFT) for all the time-intensity curves are computed. Based on the magnitude of the DFTs, the voxels are approximately clustered into two classes, corresponding to the tumor and healthy class, using the FCM algorithm [27]. These are then input to the HMMs to further refine the classification of the voxels into tumor and healthy class. In order to train and predict using HMMs, we estimate the STFT features of the voxels within the ROI.

2.2 Short-time Fourier Transform (STFT)

Although the Fourier transform preserves the information in the signals and can be computed efficiently, they lack the temporal localization of the frequencies. Therefore, we use the STFT in short time periods and obtain a feature tuple corresponding to each time window [28]. Since the STFT is computed by multiplying the time-intensity curves with a small time window, the contribution of the signals outside this time window is effectively suppressed, thereby acting as a band pass filter. The STFT is computed as,

$$STFT_x^\gamma = \int_{\tau} [x(\tau) \gamma(\tau - t')] e^{-j2\pi\tau} d\tau \quad (1)$$

where $\gamma(\tau - t')$ is the sampling window of the trajectory. The Fourier transform in each sampling window is computed by the FFT algorithm. Information loss is minimized by overlapping the STFT windows. In each sampling window, the STFT consists of the magnitude of N discrete frequency contributions. Multiplication of the signal by the relatively short window effectively suppresses the signal outside the analysis time point, thereby acting as a bandpass filter. Considering that the STFT has been computed in p time intervals, the dimension of the STFT feature tuple is $N \times p$.

2.3 Vector Quantization

Since the HMM structure considered in this paper is discrete, we convert the N tuple STFT vector in each time interval into a single discrete observation symbol using the k -means clustering algorithm [29]. The k -means algorithm partitions the pN -tuples at different time intervals into $L = 8$ sets so as to minimize the within-cluster sum of squares. The discrete observation symbol of each tuple is the index of the codebook tuple closest to the given N tuple vector, i.e., the cluster in which the vector belongs.

2.4 Hidden Markov Model (HMM)

Our HMM analysis is based on the approach and notation of Rabiner [30]. The parameters of the HMM model are defined as follows:

- The HMM is assumed to have $N (= 2)$ states. These states correspond to the two possible classes the voxel can belong to. The transition probability between state i to j is given by

$$a_{ij}=P(q_{t+1}=S_j|q_t=S_i) \quad \text{and} \quad A=\{a_{ij}\} \quad (2)$$

- Each state also has $M (= 8)$ possible observation symbols O_t . M was chosen such that the HMM model could capture the variation in the time-intensity curves observed for different pixels while avoiding a computationally intensive model. The probability of observing a particular symbol O_t in state j is

$$b_j=P(O_t|q_t=S_j) \quad \text{and} \quad B=\{b_j\} \quad (3)$$

- Also a state prior π_j is defined, which is initial probability of beginning in S_j

In short the HMM can be represented as $\lambda = (A, B, \pi)$. The model λ is trained according to the time-intensity curves or its derivatives obtained from a sample of points from each class.

2.5 HMM Training

Having generated the discrete observation symbols, the observation sequence is provided to the HMM network to obtain the updated model $\hat{\lambda}$. For training the HMMs corresponding to the two classes, we use a sample of 20 points from each of the two classes of the soft priors generated from the FCM algorithm in Step A. The typical enhancement pattern for breast tissue comprises of three types of curves - Type I, II and III, as shown in Figure 2. It has been shown that voxels demonstrating a Type III curve are highly correlated to malignant tissue while Type I curves correspond to benign tissue [31]. The HMM corresponding to the tumor class is trained using voxels which demonstrate a Type III curve. The curves are chosen such that the initial enhancement ($= (y_i^1 - y_i^0)/y_i^0$) is greater than 0.5 while the washout ($= y_i^1 - y_i^4$) is greater than 0, where y_i^k is the uninterpolated signal intensity of voxel i at the k^{th} time point. The HMMs are trained using the interpolated time-intensity curves $\left(\bar{x}_i\right)$, derivative $\left(\dot{\bar{x}}_i\right)$ and second derivative $\left(\ddot{\bar{x}}_i\right)$ of the time-intensity curves and a composite tuple consisting of the three tuples $\left(\left(\bar{x}_i \quad \dot{\bar{x}}_i \quad \ddot{\bar{x}}_i\right)\right)$.

The parameters of the models are estimated by maximizing the auxiliary function

$$Q(\lambda, \bar{\lambda}) = \sum_Q P(Q|O, \lambda) \log \left[P(O, Q | \bar{\lambda}) \right] \quad (4)$$

This optimization problem is solved iteratively by the Baum-Welch method [30]. Two HMMs are trained corresponding to the ‘‘Healthy’’ (λ_1) and ‘‘Tumor’’ (λ_2) class.

2.6 HMM Prediction

Once the HMMs have been trained, the next step is to estimate the likelihood of each HMM observing a given sequence. The observation sequence consists of the STFT tuples generated from the intensity, velocity, acceleration or composite vectors of the voxels within the ROI. The probability of predicting the observation sequence given the HMM model is

computed inductively using the forward-backward algorithm [32] by summing the following joint probability over all possible state sequences:

$$P(O|\lambda) = \sum_{\text{all } Q} P(O|Q, \lambda) P(Q|\lambda) \quad (5)$$

The reader is referred to [30] for greater details.

2.7 Similarity Ratio

A similarity ratio χ_j is computed to evaluate if a particular point j belongs to the tumor or healthy class and is converted to a scalar color map. The hypothesis is that a voxel belonging to the tumor class will have a higher probability of being observed by the HMM trained for tumor voxel (λ_2) than the healthy HMM (λ_1). The measure is defined as

$$\chi_j = \log(P(O_j|\lambda_1)) / \log(P(O_j|\lambda_2)) \quad (6)$$

The similarity index χ_j is converted to a scalar image, representing the tumor map, and overlaid with the DCE-MRI.

3 Methods

In these retrospective studies, patient records were analyzed for cases in which biopsy confirmed breast cancer was detected. These included 22 breast examinations of women diagnosed with IDC and 19 examinations of women diagnosed with DCIS. The protocol was approved by the Brigham and Women's Hospital Institutional Review Board and compliant with Health Insurance Portability and Accountability Act (HIPAA). Five different MR scanners were used for imaging of the patients using either a 1.5-T or a 3-T closed magnet (Signa or HDX, GE Healthcare and Trio, Siemens Medical Solutions, Erlangen, Germany) and a dedicated breast-surface coil (InVivo 7-Channel Breast Biopsy Array, Invivo Research). Scheduling availability determined which magnet would be used. Scanning protocols included pre- and post- VIBRANT (Volume Imaging Breast Assessment) or VIBE (Volumetric Interpolated Breath hold Examination) fat saturated sequences in either the sagittal or axial planes. The dynamic contrast sequences were obtained at 90s, 180s, 270s and 360s following intravenous contrast administration. Patients received 0.2ml/kg of gadolinium (Magnevist, Bayer HealthCare) contrast infused at 2ml/s. The tumors diagnosed in the images were biopsy confirmed to have either IDC or DCIS.

The 4D DCE-MRI data set was loaded in 3D Slicer, an open-source image processing and navigation software (www.slicer.org). A region of interest was delineated on the DCE-MRI. The time-intensity curve for each voxel was logged and provided to the SLATS algorithm along with the computed velocity, acceleration and composite tuples. The resultant tumor maps were converted to DICOM images and imported into 3D Slicer and overlaid with the DCE-MRI. The workflow of a typical case is shown in Figure 3.

For each breast examination, an expert radiologist delineated the primary node of the tumor by overlaying the pre-contrast and first post-contrast image. The location of the tumor described in the biopsy reports was also made available to the radiologist. A visual inspection of brightly enhanced portions of the DCE-MRI was utilized for delineating the tumor boundaries. Although numerous tumor foci may be present, the radiologist only delineated the masses in the vicinity of the biopsy proven malignant tumor. In addition, for each breast exam the output of a commercial software, CADstream (Merge Healthcare Inc.) was obtained and registered to the DCE-MRI in 3D Slicer. The tumor map generated by CADstream was outlined. The performance of the algorithm was assessed with the

following measurables: Accuracy, Sensitivity and Dice Similarity Coefficient (DSC). Any overlap between the two tumor masks is considered as a true positive finding while no overlap between the two tumor masks is considered as a false positive or negative finding. The accuracy and sensitivity of detecting a tumor focus by the algorithm was calculated, considering the radiologists segmentation and the CADstream output as the reference validation independently. In addition, the amount of overlap of the tumor map generated by the SLATS and the radiologist's delineation and CADstream output was measured in terms of the DSC. A perfect overlap results in a DSC of 1 while no overlap results in DSC of 0.

4 Results

4.1 Algorithm results

Figure 4 shows the pre-contrast (a) and the four post contrast (b-e) T1-weighted images. A biopsy proven IDC is shown by a red arrow. A visual inspection of the images suggests an increase in the intensity of voxels corresponding to the tumor mass while the surrounding healthy parenchyma shows negligible increase in the signal intensity. Due to angiogenesis, tumors have large number of vessels supplying blood and other nutrients. As a result of the enhanced blood flow, the contrast rapidly washes into the tumor, thereby increasing the intensity of voxels corresponding to the tumor at the first time point. However, the tumor blood vessels are malformed and leaky, thereby resulting in washout of contrast. This can be seen over the next three time points, as shown in Figure 4(b)-(e). On the other hand, the surrounding healthy parenchyma has a lower density of blood vessels, resulting in minimal increase in signal intensity of the healthy parenchyma in the DCE-MRI. The time-intensity curve of a voxel within the tumor and healthy parenchyma is shown in Figure 5(a). The first and second derivatives of the time-intensity curves (hereafter denoted as the "velocity" and "acceleration" trajectories) are computed by the first and second order difference equations and are shown in Figure 5(b) and (c) respectively.

As detailed in Section 1, the time-intensity ("intensity"), velocity, acceleration and composite tuples are input to the HMMs to train the network to predict the occurrence of the discrete observation sequence at each voxel. The first step of the algorithm is to estimate the patient-specific priors, which are then utilized to refine the segmentation using the HMM. The magnitude of the Fourier transform of the observation tuples are input to the FCM algorithm, which approximately clusters the observation tuples from all the voxels into two classes corresponding to the healthy and tumor class. Figure 6 shows the patient specific priors generated from the intensity, velocity, acceleration and composite tuples. It can be seen in Figure 6 that the FCM algorithm detected the primary mass, which is clearly seen on the DCE-MRI, along with a smaller secondary mass that is difficult to identify using just the pre-contrast and post-contrast images. Having generated the priors, a small sample of points belonging to the tumor and healthy classes are provided to the two HMMs to train the HMMs corresponding to the healthy and tumor class. Thereafter, the observation tuples from all voxels under the mask are then provided to the two HMMs to evaluate the similarity metric for generating the final tumor mask. The output from the SLATS algorithm is shown in Figure 7. The tumor masks generated by the FCM algorithm is further refined by the HMM. The result of the four cases was compared with the radiologist's segmentation and CADstream output and quantified in terms of the accuracy and sensitivity of detecting the tumors and DSC for the amount of overlap. For 112×70 voxels ROI, the computation time for the FCM step was 6.76 seconds while the computation time for training the HMM and estimating the similarity ratio was 8.54 seconds. For a smaller ROI with size 56×32 voxels, the computation time for the FCM and HMM steps were 0.15 and 3.90 seconds respectively. The simulations were performed on a Dell Precision Workstation (Intel Xeon Quad-core, 32 GB RAM, Fedora 12, MATLAB 7.7.0).

4.2 Result for Invasive Ductal Carcinoma (IDC) cases

The algorithm was first tested for patients diagnosed with IDC that typically enhance faster than any other type of breast carcinoma. The result of SLATS for four patient cases is shown in Figure 8. The output of SLATS for the position tuple is shown in Figure 8(a), velocity tuple in Figure 8(b), acceleration tuple in Figure 8(c) and composite tuple in Figure 8(d). The tumor maps correspond closely with the CADstream output, shown in Figure 8(e) and the manual segmentations, shown in Figure 8(f). However, due to the inherent noise in the imaging modality, the tumor map generated using the acceleration tuple demonstrated numerous spurious tumor masses, as seen in the third case in Figure 8(c). For the 22 IDC cases analyzed in this paper, the results are summarized in Table I. The sensitivity of detecting the tumors for the four observation tuples is consistently higher than 92% with the composite observation tuple showing the highest sensitivity. The HMM trained with the composite tuple identified all the tumors delineated by the radiologist. In a few cases, the algorithm detected additional tumor masses as shown in Figure 9. The accuracy demonstrated by the four observation tuples is similar and ranges between 85.1% for the intensity tuple and 92.6% for the composite tuple.

In the absence of targeted biopsy results, the only way to characterize the additional mass for evaluating the accuracy of the algorithm is by observing the type of signal enhancement (see Figure 2). In order to quantify the accuracy of the tumor map, the enhancement curve for Regions of Interest (ROIs) chosen within the different tumor masks is shown in Figure 9. It can be seen that curves corresponding to ROI(a) and ROI(b) show a Type III curve, characterized by a rapid enhancement greater than 100% at the first time point and a subsequent washout, which is typical of a malignant tumor. ROI(c) which lies within the tumor map generated by all observation tuples except the acceleration tumor mask also shows a borderline Type III/Type II curve which could correspond to malignant tumor. However, enhancement curves corresponding to ROIs (d) and (e), which are within the tumor maps generated by the position and composite tuples but outside the velocity tumor map, correspond to a Type I persistent curve. The Type I curve is less likely to correspond to malignant tissue. This suggests that the tumor mask generated by the velocity tuple is more likely to be the most accurate in delineating the regions of angiogenesis. The secondary mass detected by the SLATS algorithm trained for the velocity tuple was also missed by the radiologist.

The accuracy of the tumor maps was also validated by evaluating the DSC, which measures the amount of overlap between SLATS detected tumor and manually outlined tumor. Considering the radiologist's segmentation as the gold standard, the DSC for the tumor map generated by the SLATS corresponding to the velocity tuple is higher than the DSC of the tumor maps generated from other observation tuples.

The result of the SLATS algorithm was also compared with the output of CADstream. The results are summarized in Table II. In 9% (2/22 cases) of the IDC cases, the commercial tumor map failed to detect a biopsy proven tumor due to slow enhancement of the primary mass. In these cases, the SLATS algorithm successfully detected the tumor mass with an average DSC of 0.71, 0.60, 0.39 and 0.69 respectively corresponding to the intensity, velocity, acceleration and composite tuples, demonstrating that the SLATS performs better than the commercially available software. In the remaining 20 cases, considering the CADstream output as the gold standard validation, the accuracy and sensitivity of detecting the tumor masses are similar for the four observation tuples, ranging in the 90% to 100% range. The tumor masks generated by the HMMs trained for the velocity and acceleration tuples demonstrated maximum overlap with the CADstream output with a DSC of 0.71 and 0.72 respectively.

4.3 Result for Ductal Carcinoma In Situ (DCIS) cases

Figure 10 shows the pre-contrast and four post-contrast images of a breast tumor confirmed to be DCIS. As with the IDCs, DCIS shows significant enhancement of contrast at the first time point and a subsequent decrease in contrast over the next three time points. An ROI is selected on the DCE-MRI and all the time-intensity tuples along with the computed velocity, acceleration and composite tuples are provided to the SLATS to generate the tumor mask. The result of the SLATS for the four observation tuples is shown in Figure 11.

The SLATS was implemented on 19 incisional or excisional biopsy proven DCIS cases. The result for four cases is shown in Figure 12. As with the IDC cases, the result of SLATS was compared with the radiologist's segmentation and the output of CADstream. The result of the SLATS algorithm compared with the manual segmentation is shown in Table III. The overall accuracy of detecting DCIS is lower compared to the result for the IDC cases, perhaps because DCIS is low enhancing tumor and can be difficult for the radiologist to detect using a visual inspection of the pre- and first post-contrast DCE-MRI. In order to evaluate the accuracy of the tumor maps generated from SLATS, we follow the same procedure as performed above for the IDC case shown in Figure 9. We evaluate the time-intensity curves of ROIs chosen within each tumor mask (see Figure 13). As seen in Figure 13 (left), ROI1 corresponds to the primary mass which is detected by the radiologist, CADstream and SLATS (indicated by a blue arrow), ROI2 to a mass detected by CADstream and SLATS (indicated by a red arrow) and ROI3 to a mass detected only by SLATS (indicated by a green arrow). The enhancement curves for ROI1 and ROI2 show a typical Type III curve with an initial enhancement greater than 100%, thereby suggesting that the two masses correspond to malignant tumors. ROI3 also shows a Type III curve. However, the initial enhancement is significantly lower than the two other masses but is greater than the 50% threshold. This type of low enhancement is typical of DCIS and therefore, the mass detected by the SLATS algorithm is more likely to correspond to a true positive.

The accuracy of the SLATS algorithm trained using the velocity tuple is the highest among the four observation tuples and has 100% sensitivity in detecting the DCIS tumors. The amount of overlap with the radiologists segmentation was also measured in terms of DSC and is shown in Table III for the four observation tuples. The tumor mask generated by the velocity tuple shows maximum overlap with the radiologist's segmentation with a DSC of 0.69.

The result of the SLATS algorithm for DCIS cases was also compared with the CADstream output. Since DCIS is a weakly enhancing tumor, CADstream did not detect significant contrast enhancement corresponding to the tumor in 26.3% (5/19) of the cases. In these cases, the SLATS algorithm detected DCIS with a DSC of 0.56, 0.66, 0.55 and 0.52 respectively for the four observation tuples when compared to the manual segmentations. In the remaining 14 cases, considering the CADstream output as the gold standard, the results of the SLATS algorithm are summarized in Table IV. The velocity tumor mask shows an accuracy of 90%, a sensitivity of 100% and a DSC of 0.58 in segmenting tumors which have been identified by CADstream.

4.4 Result for DCE-MRI with greater time points

In the previous subsections, the SLATS algorithm has been tested for DCE-MRI with four time points excluding the pre-contrast image. In this section, we will show that the SLATS algorithm can be applied to DCE-MRI with any number of time points. The algorithm has been tested on 10 datasets obtained from the Cancer Imaging Archive, an initiative funded by the National Cancer Institute (<https://cancerimagingarchive.net/>). The women were

imaged prone in a 1.5T Achieva MRI scanner from Philips Medical Systems. The pulse parameters were as follows: matrix size = 560×560; slice thickness = 4mm; flip angle = 12°. The DCE-MRI has 8 time points. The SLATS algorithm has been implemented on these datasets with no modifications. To evaluate the output of the algorithm, a radiologist manually segmented the tumor mass from the pre- and fourth post-contrast images with limited information obtained from the case study report.

The result of the SLATS algorithm compared to the manual segmentation mask is shown in Table V. The SLATS demonstrates 100% sensitivity in detecting biopsy proven tumors. The SLATS tumor mask corresponding to the acceleration tuple detected additional tumor masses, therefore resulting in an accuracy of 91.6%. The SLATS trained for other observation tuples detected the tumor masses with a 100% accuracy. The amount of overlap with the radiologist's segmentation is also significant with a DSC over 0.73. The results show that the SLATS algorithm can be applied to DCE-MRI with higher number of time points. In this preliminary result, it also appears that the SLATS performs better for DCE-MRI with greater time points.

4.5 Result for robustness analysis

In order to test the stability of the algorithm, the DCE-MRI and the associated tumor map T_t was translated by an arbitrary amount, which in this case was (-20, 37, 16) mm in the Right-Anterior-Superior coordinate frame. The ROI was chosen again and the time-intensity profiles were logged and provided to the SLATS algorithm trained for the intensity tuple based on which the tumor map T was obtained. The DSC computed between T_t and T was 0.99, thereby showing a robust stability to translational disturbances. In addition, random noise was also added to the DCE-MRI to evaluate the noise disturbance rejection capability. The DSC between the tumor map generated from the noisy DCE-MRI and original map was 0.995, showing excellent noise rejection capability. To test the reproducibility of the algorithm, we have further tested the SLATS algorithm for a single case for 10 trials with varying initial conditions. The mean DSC between the tumor map generated from the 10 trials and the original tumor map was computed to be 0.99934, showing excellent reproducibility of the results.

5 Conclusion

In this paper, we have described an algorithm based on HMMs to segment regions of angiogenesis corresponding to the tumor. The SLATS has been trained using STFT feature vectors obtained from the intensity, velocity, acceleration and composite tuples. The SLATS trained for the velocity tuple shows the best performance in terms of the accuracy and sensitivity of detecting tumors identified by the radiologist and CADstream. The tumor map generated from SLATS using the velocity tuple also shows maximum overlap with the radiologist's segmentation and CADstream output in terms of the DSC. In 9% of IDC cases and 26.3% of DCIS cases, CADstream failed to detect biopsy proven tumors. In these cases, the SLATS algorithm successfully delineated the tumor, thereby showing better performance than the commercially available software. The algorithm has also been tested for DCE-MRI with 8 time points. The results show the applicability of the algorithm for DCE-MRI with a higher number of time points. The algorithm is also robust to noise and translational disturbances. However, deformation of the breast during imaging needs to be addressed using non-linear deformable registration algorithms. Further validation of the algorithm using biopsy samples obtained under image-guidance is currently underway.

References

- [1]. Eliat PA, Dedieu V, Bertino C, Boute V, Lacroix J, Constans J, de Korvin B, Vincent C, Bailly C, Jo re F, de Certaines J, Vincensini D. Magnetic resonance imaging contrast-enhanced relaxometry of breast tumors: an MRI multicenter investigation concerning 100 patients. *Magnetic Resonance Imaging*. 2004; vol. 22(no. 4):475–481. [PubMed: 15120166]
- [2]. Radjenovic A, Dall BJ, Ridgway JP, Smith MA. Measurement of pharmacokinetic parameters in histologically graded invasive breast tumours using dynamic contrast-enhanced MRI. *Br J Radiol*. 2008; vol. 81(no. 962):120–128. [PubMed: 18070824]
- [3]. Pickles MD, Manton DJ, Lowry M, Turnbull LW. Prognostic value of pre-treatment DCE-MRI parameters in predicting disease free and overall survival for breast cancer patients undergoing neoadjuvant chemotherapy. *European Journal of Radiology*. 2009; vol. 71(no. 3):498–505. [PubMed: 18572340]
- [4]. Tuncbilek N, Tokatli F, Altaner S, Sezer A, Tre M, Omurlu IK, Temizoz O. Prognostic value DCE-MRI parameters in predicting factor disease free survival and overall survival for breast cancer patients. *European Journal of Radiology*. 2011; vol. 71(no. 3):498–505.
- [5]. Bhooshan N, Giger ML, Jansen SA, Li H, Lan L, Newstead GM. Cancerous breast lesions on dynamic contrast-enhanced MR images: Computerized characterization for image-based prognostic markers. *Radiology*. 2010; vol. 254(no. 3):680–690. [PubMed: 20123903]
- [6]. Hylton N. Dynamic contrast-enhanced magnetic resonance imaging as an imaging biomarker. *Journal of Clinical Oncology*. 2006; vol. 24(no. 20):3293–3298. [PubMed: 16829653]
- [7]. El Khouli, Riham H.; Macura, Katarzyna J.; Jacobs, Michael A.; Khalil, Tarek H.; Kamel, Ihab R.; Dwyer, Andrew; Bluemke, David A. Dynamic contrast-enhanced MRI of the breast: Quantitative method for kinetic curve type assessment. *American Journal of Roentgenology*. 2009; vol. 193(no. 4):W295–W300. [PubMed: 19770298]
- [8]. Tofts PS, Kermode AG. Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. fundamental concepts. *Mag Res in Med*. 1991; vol. 17(no. 2): 357–367.
- [9]. Li, Xin; Huang, Wei; Yankeelov, Thomas E.; Tudorica, Alina; Rooney, William D.; Springer, Charles S. Shutter-speed analysis of contrast reagent bolus-tracking data: Preliminary observations in benign and malignant breast disease. *Magnetic Resonance in Medicine*. 2005; vol. 53(no. 3):724–729. [PubMed: 15723402]
- [10]. Brix, Gunnar; Kiessling, Fabian; Lucht, Robert; Darai, Susanne; Wasser, Klaus; Delorme, Stefan; Griebel, Jorgen. Microcirculation and microvasculature in breast tumors: Pharmacokinetic analysis of dynamic mr image series. *Magnetic Resonance in Medicine*. 2004; vol. 52(no. 2): 420–429. [PubMed: 15282828]
- [11]. Sourbron SP, Buckley DL. On the scope and interpretation of the tofts models for DCE-MRI. *Magnetic Resonance in Medicine*. 2011; vol. 66(no. 3):735–745. [PubMed: 21384424]
- [12]. Luykaert R, Sourbron S, Makkat S, de Mey J. Error estimation for perfusion parameters obtained using the two-compartment exchange model in dynamic contrast-enhanced mri: a simulation study. *Physics in Medicine and Biology*. 2010; vol. 55:6431. [PubMed: 20952813]
- [13]. Litjens, GJS.; Heisen, M.; Buurman, J.; ter Haar Romeny, BM. Pharmacokinetic models in clinical practice: what model to use for dce-mri of the breast?. *Proceedings of the 2010 IEEE international conference on Biomedical imaging: from nano to Macro*; Piscataway, NJ, USA. IEEE Press; 2010. p. 185-188. ISBN: 10
- [14]. Buckley, David L. Uncertainty in the analysis of tracer kinetics using dynamic contrast-enhanced T1-weighted MRI. *Mag Res in Med*. 2002; vol. 47(no. 3):601–606.
- [15]. Chen W, Giger ML, Bick U. A Fuzzy C-Means (FCM)-based approach for computerized segmentation of breast lesions in dynamic contrast-enhanced MR images. *Academic Radiology*. 2006; vol. 13(no. 1):63–72. [PubMed: 16399033]
- [16]. Bhooshan, Neha; Giger, Maryellen; Edwards, Darrin; Yuan, Yading; Jansen, Sanaz; Li, Hui; Lan, Li; Sattar, Husain; Newstead, Gillian. Computerized three-class classification of mri-based prognostic markers for breast cancer. *Physics in Medicine and Biology*. 2011; vol. 56(no. 18): 5995. [PubMed: 21860079]

- [17]. Eyal, Erez; Badikhi, Daria; Furman-Haran, Edna; Kelcz, Fredrick; Kirshenbaum, Kevin J.; Degani, Hadassa. Principal component analysis of breast dce-mri adjusted with a model-based method. *Journal of Magnetic Resonance Imaging*. 2009; vol. 30(no. 5):989–998. [PubMed: 19856419]
- [18]. Juang BH, Rabiner LR. Hidden markov models for speech recognition. *Technometrics*. 1991; vol. 33:251–272.
- [19]. Karplus K, Barrett C, Hughey R. Hidden markov models for detecting remote protein homologies. *Bioinformatics*. 1998; vol. 14(no. 10):846–856. [PubMed: 9927713]
- [20]. Wilson AD, Bobick AF. Parametric hidden markov models for gesture recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1999; vol. 21:884–900.
- [21]. Gonzalez AM, Roque AMS, Garcia-Gonzalez J. Modeling and forecasting electricity prices with input/output hidden markov models. *IEEE Transactions on Power Systems*. 2005; vol. 20:13–24.
- [22]. Coast DA, Stern RM, Cano GG, Briller SA. An approach to cardiac arrhythmia analysis using hidden markov models. *Biomedical Engineering, IEEE Transactions on*. Sept.1990 vol. 37(no. 9):826–836.
- [23]. Liu, Xin; Langer, DL.; Haider, MA.; Yang, Y.; Wernick, MN.; Yetik, IS. Prostate cancer segmentation with simultaneous estimation of markov random field parameters and class. *Medical Imaging, IEEE Transactions on*. Jun; 2009 vol. 28(no. 6):906–915.
- [24]. Gong, Yang; Brady, Michael. Texture-based simultaneous registration and segmentation of breast dce-mri. In: Krupinski, Elizabeth, editor. *Digital Mammography*. Vol. vol. 5116. Springer; Berlin, Heidelberg; 2008. p. 174-180. *Lecture Notes in Computer Science*
- [25]. Shafer CM, Seewaldt VL, Lo JY. Validation of a 3D Hidden-Markov model for breast tissue segmentation and density estimation from MR and tomosynthesis images. *Biomedical Sciences and Engineering Conference (BSEC)*. 2011:1–4.
- [26]. Jayender, J.; Vosburgh, KG.; Gombos, E.; Ashraf, A.; Kontos, D.; Gavenonis, SC.; Jolesz, FA.; Pohl, K. *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*. IEEE; 2012. Automatic segmentation of breast carcinomas from dce-mri using a statistical learning algorithm; p. 122-125.
- [27]. Bezdek JC. FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*. 1984; vol. 10(no. 23):191–203.
- [28]. Hlawatsch F, Bordeaux-Bartles D. Linear and quadratic time-frequency signal representation. *IEEE Signal Process*. 1992; vol. 9:21–67.
- [29]. Likas A, Vlassisb N, Verbeekb JJ. The global k-means clustering algorithm. *Pattern Recognition*. 2003; vol. 36:451–461.
- [30]. Rabiner LR. A tutorial on Hidden Markov models and selected applications in speech recognition. *Proceeding of the IEEE*. 1989; vol. 77:257–286.
- [31]. Macura, Katarzyna J.; Ouwerkerk, Ronald; Jacobs, Michael A.; Bluemke, David A. Patterns of enhancement on breast mr images: Interpretation and imaging pitfalls1. *Radiographics*. Nov-Dec; 2006 vol. 26(no. 6):1719–1734. [PubMed: 17102046]
- [32]. Beyar R, Wenderow T, Lindner D, Kumar G, Shofti R. Concept, design and pre-clinical studies for remote control percutaneous coronary interventions. *Eurointervention*. 2005; vol. 1:340–345. [PubMed: 19758927]

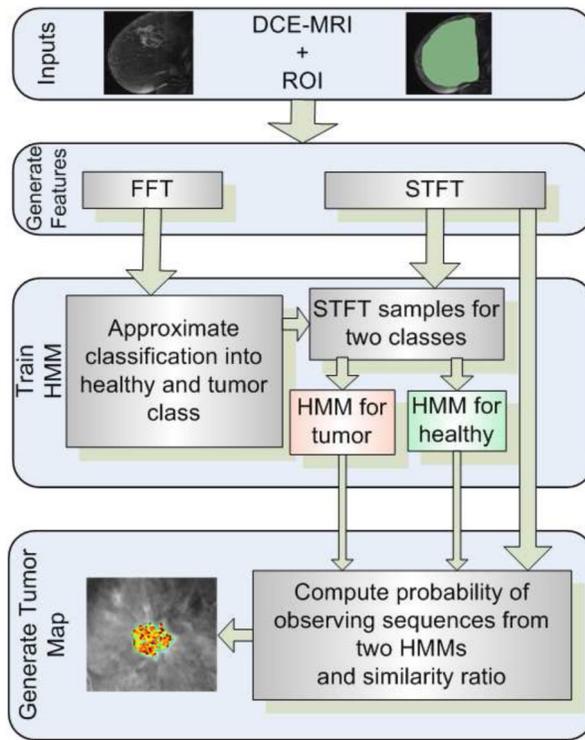


Figure 1.
Flowchart of SLATS

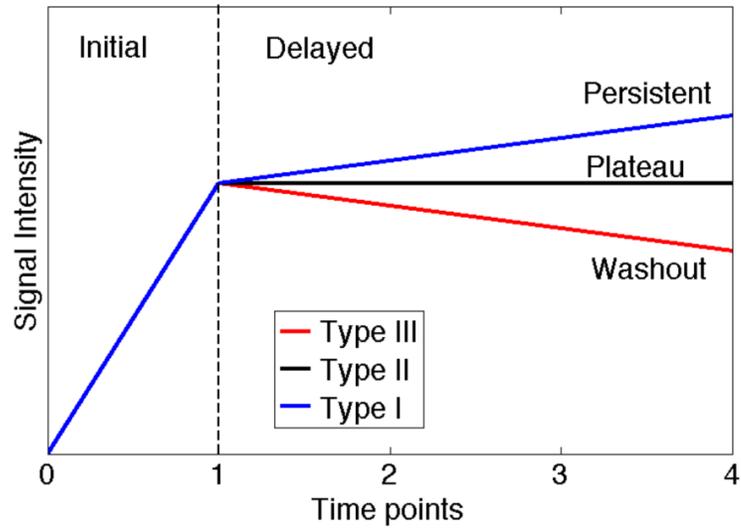


Figure 2.
Types of signal enhancement in DCE-MRI

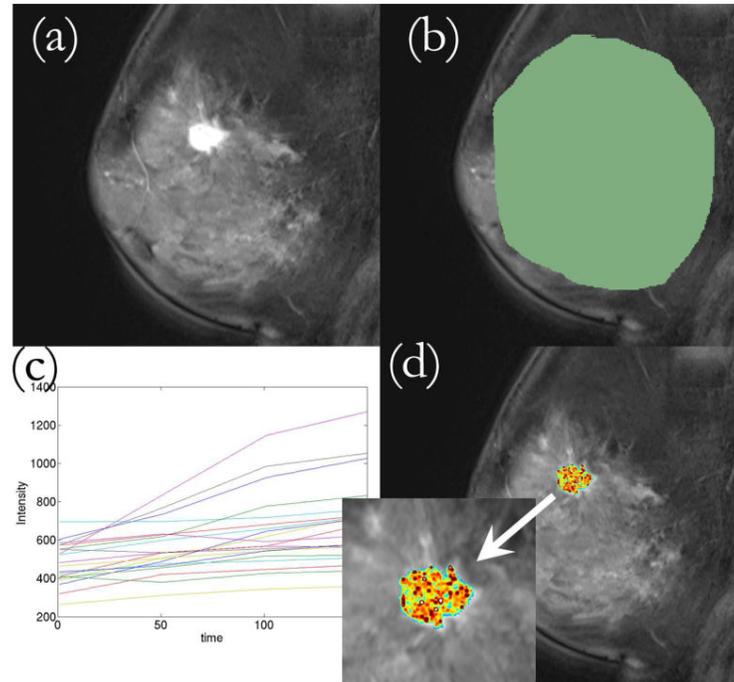


Figure 3. Workflow of the SLATS. (a) DCE-MRI loaded into 3D Slicer (b) ROI delineated (c) Time-intensity curves obtained from all voxels under ROI and provided to SLATS (d) Tumor map is generated

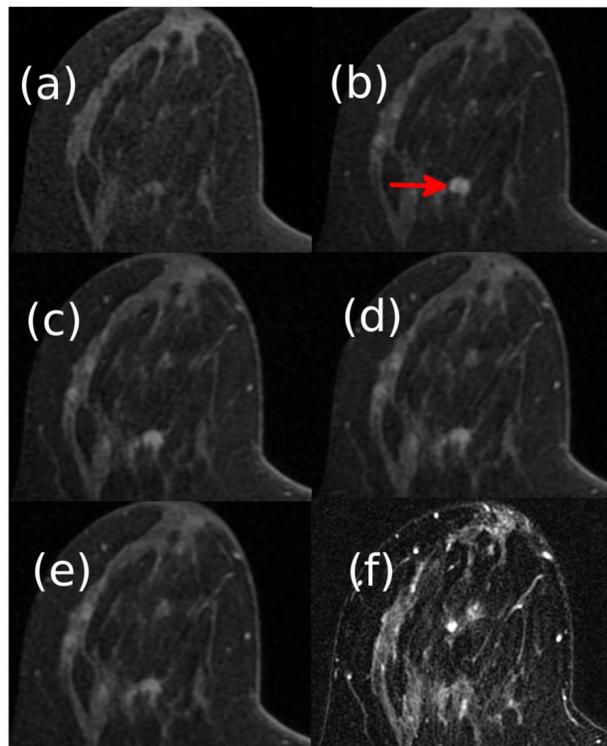


Figure 4. DCE-MRI of a patient with IDC. (a) Pre-contrast baseline image (b)-(e) Four post-contrast images (f) T2-weighted image

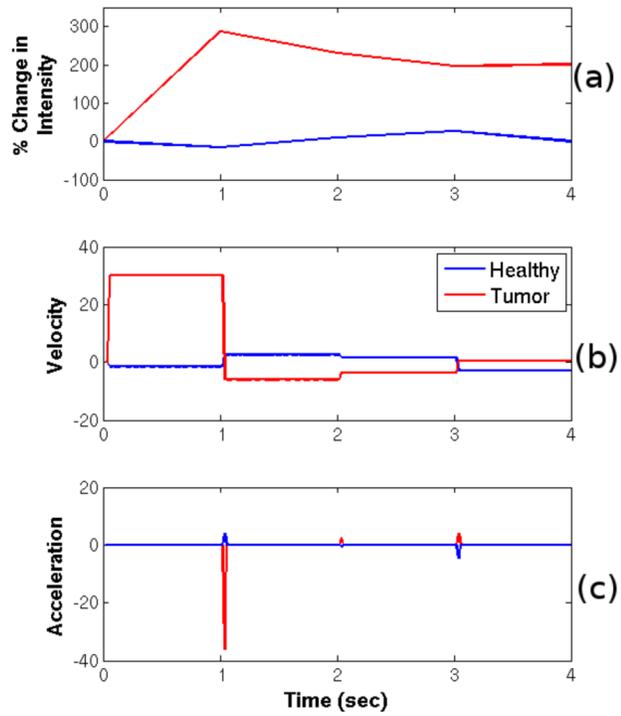


Figure 5. (a) Time intensity curves (b) Derivative of the time intensity curve (“velocity”) (c) Second derivative of the time intensity curve (“acceleration”)

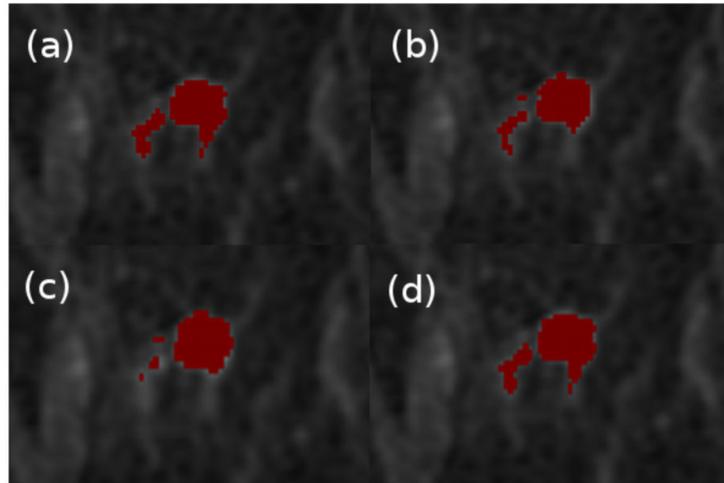


Figure 6. Output of the FCM algorithm for (a) Intensity (b) Velocity (c) Acceleration (d) Composite tuples

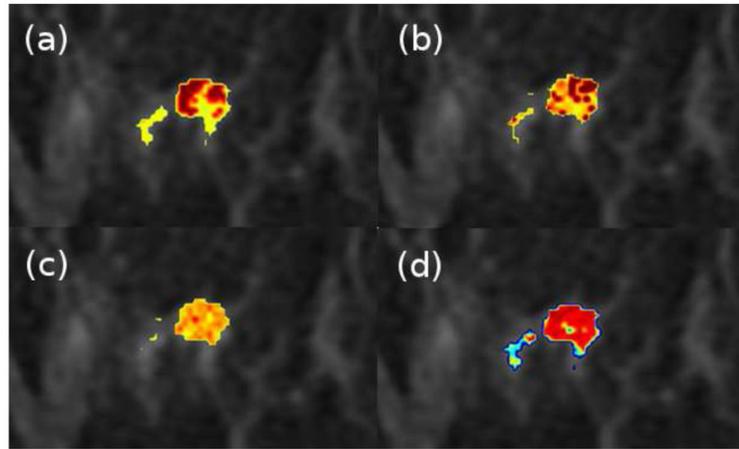


Figure 7. Result of the SLATS algorithm for (a) Intensity (b) Velocity (c) Acceleration (d) Composite tuples

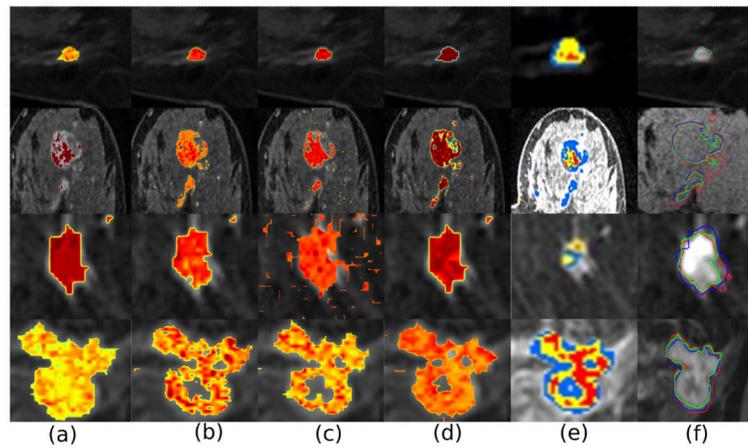


Figure 8. Results of the SLATS algorithm for four cases, each row is a case. (a) Position (b) Velocity (c) Acceleration (d) Composite tuples (e) CADstream output (f) Magnified view of the smoothed tumor outlines defined by velocity SLATS map (green), Radiologist (red) and CADstream output (blue).

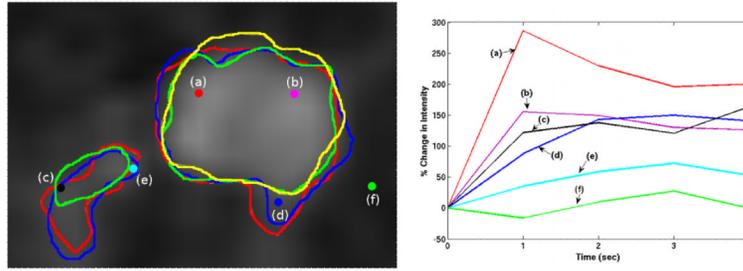


Figure 9. (left) Tumor map outlines generated by the position (red), velocity (green), acceleration (yellow) and composite (blue) tuples (right) Time-intensity curves of the ROIs

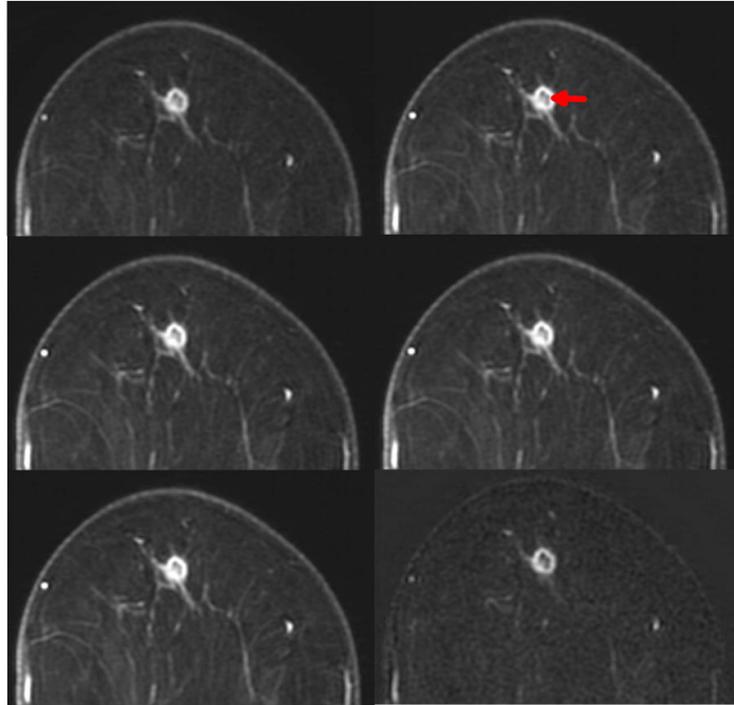


Figure 10. DCE-MRI of a patient with DCIS. (a) Pre-contrast baseline image (b)-(e) Four post-contrast images (f) Subtracted image (First post-contrast image minus baseline image)

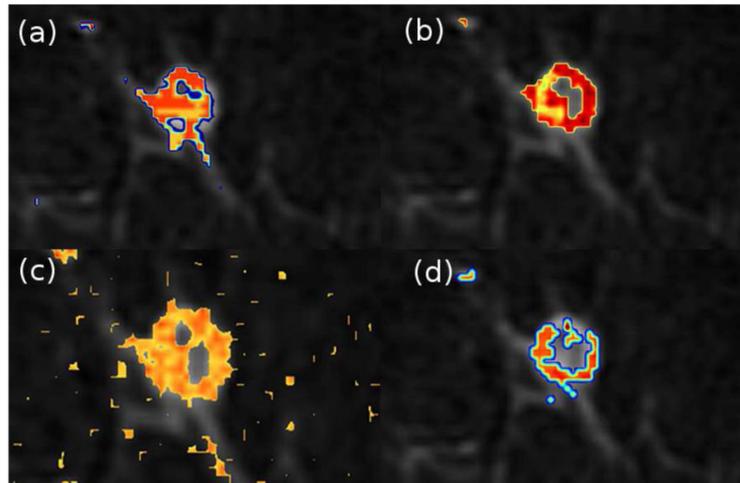


Figure 11.
Result of SLATS for (a) position (b) velocity (c) acceleration (d) composite tuples

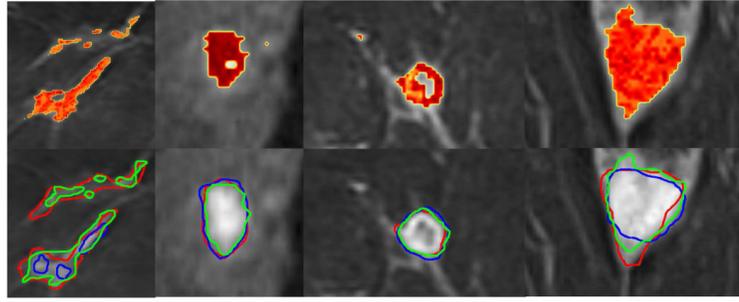


Figure 12.

Results of the SLATS algorithm for four DCIS cases, each column is a case. (Top) SLATS output generated from velocity tuple (bottom) Outlines of the velocity SLATS tumor map (green), radiologist's segmentation (red) and CADstream output (blue)

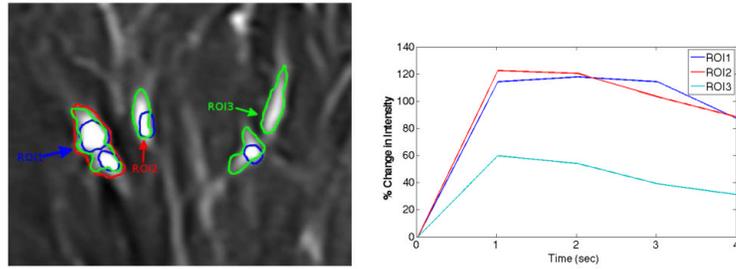


Figure 13. (left) Tumor map outlines generated by the SLATS velocity tumor map (green), radiologist (red) and CADstream output (blue) (right) Time-intensity curves of the ROIs

TABLE I

RESULTS OF THE SLATS COMPARED TO RADIOLOGIST'S DELINEATION FOR IDC CASES

	Intensity	Velocity	Acceleration	Composite
Accuracy	85.1%	88.9%	88.4%	92.6%
Sensitivity	92%	96%	92%	100%
DSC	0.63	0.75	0.71	0.72

TABLE IIRESULTS OF THE SLATS COMPARED TO CAD_{STREAM} OUTPUT FOR IDC CASES

	Intensity	Velocity	Acceleration	Composite
Accuracy	96%	92.3%	92%	88.8%
Sensitivity	100%	100%	95.8%	100%
DSC	0.59	0.71	0.72	0.68

TABLE III

RESULTS OF THE SLATS COMPARED TO RADIOLOGIST'S DELINEATION FOR DCIS CASES

	Intensity	Velocity	Acceleration	Composite
Accuracy	67.6%	79.3%	68.7%	62.1%
Sensitivity	100%	100%	95.6%	100%
DSC	0.58	0.69	0.56	0.60

TABLE IVRESULTS OF THE SLATS COMPARED TO CAD_{STREAM} OUTPUT FOR DCIS CASES

	Intensity	Velocity	Acceleration	Composite
Accuracy	76%	90.4%	75%	70.3%
Sensitivity	100%	100%	94.7%	100%
DSC	0.44	0.58	0.58	0.49

Table V

RESULTS OF THE SLATS COMPARED TO RADIOLOGIST'S DELINEATION FOR DCE-MRI WITH GREATER TIME POINTS

	Intensity	Velocity	Acceleration	Composite
Accuracy	100%	100%	91.6%	100%
Sensitivity	100%	100%	100%	100%
DSC	0.80	0.78	0.73	0.83