UNIVERSITY OF CALIFORNIA
RIVERSIDE

Computational Methods for Mild Traumatic Brain Injury

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Anthony Christopher Bianchi

December 2014

Dissertation Committee:
Dr. Bir Bhanu, Chairperson
Dr. Andre Obenaus
Dr. Hyle Park

The Dissertation of Anthony Christopher Bianchi is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgements

I thank my mentor and advisor Dr. Bir Bhanu for his guidance and the motivation he has instilled upon me to better myself and others. I would like to give my gratitude my co-advisor Dr. Andre Obenaus for his invaluable biological and clinically driven recommendations. Thank you to Dr. Hyle Park for being on my thesis defense committee. Also, I would like to thank Dr. Gerardo Beni, Dr. Subir Ghosh and Dr. Yingbo Hua for steering me in my oral examination. I am grateful to my colleagues in VISLAB for their thought provoking discussions and writing advice: Dr. Alberto C. Cruz, Dr. Ninad Thakoor, Dr. Hoang T. Nguyen, Dr.Yu Sun and Vincent On. I would like to give a special thanks to Dr. Virginia Donovan for her collaboration with data collection and analysis. I would like to thank the IGERT Videobioinformatics faculty and participants for their support and alternate points of view. Support of this work was provided by NSF IGERT: Video Bioinformatics grant DGE 0903667, NSF Learning Concepts in Morphological Image Databases DCMRP-DR080470, NSF: Learning Concepts in Morphological Image Databases 0641076. Portion of this dissertation appear in "Visual and Contextual Modeling for the Detection of Repeated Mild Traumatic Brain Injury" and "Dynamic Low-Level Context for the Detection of Mild Traumatic Brain Injury" by A. Bianchi et al. © 2014 the IEEE.

I dedicate this to my parents Alyce M. Castleman and David A. Castleman for their dedication to seeing me succeed in life.

ABSTRACT OF THE DISSERTATION

Computational Methods for Mild Traumatic Brain Injury

by

Anthony Christopher Bianchi

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, December 2014
Dr. Bir Bhanu, Chairperson

Awareness of mild traumatic brain injury and its potential long term effects has increased dramatically in recent years. Currently, there is a lack of computational methods for the evaluation of mild traumatic brain injury (mTBI) from magnetic resonance imaging (MRI). Further, the development of automated analyses has been hindered by the subtle nature of mTBI abnormalities, which appear as low contrast MR regions. In manual estimation a train operator must spend a copious amount of time locating and delineating the lesions. We have developed three innovations to aid in the analysis of mTBI by automatically detecting lesion in MR images. (1) A high level Bayesian network based contextual model, which is able to estimate the progression of lesion over time given basic information about the subject. The generative contextual model is fused with a discriminative visual model that estimates the lesion using 3D features extracted from T2maps. (2) Low level dynamic and static contextual features are developed that take advantage of sequential imaging to improve detection as an iterative

post processing step. After the initial estimate by a discriminative classifier, the context features are estimated from the posterior marginal to describe spatio-temporal information about the lesion to the classifier. (3) Multi-Modal features are developed to further improve the visual model. A temporal generative tissue model is developed to model tissue type probabilities within the brain over the course of the injury in the T2map modality. Maximum a posterior edge features are developed to take advantage of the susceptibility weighted imaging (SWI) channel. All of the proposed methods have been evaluated on a temporal mTBI dataset with manually evaluated ground truth and have been shown to outperform the current state-of-the-art approaches.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1 INTRODUCTION

Mild traumatic brain injury (mTBI) annually affects millions of people within the United States [1]. This is due to events, such as sports, military (blast), automobile accidents, assaults, and falls. There are many short and long term symptoms associated with mTBI including loss of memory, loss of reasoning, neuropsychiatric alterations including decrements in social interactions [2]. Recently, there has been increased focus on the potential effects of repeated mTBI as this may result in exacerbation of ongoing tissue pathology and neurological deficits. In extreme cases, second impact syndrome can occur in which a subsequent injury is sustained within hours to weeks of the first. This may result in catastrophic swelling of brain which may ultimately lead to death or the effects similar to a severe TBI [3].

Currently, neurological injuries are evaluated using the Glasgow Coma Scale (GCS) which evaluates a patient's consciousness level through his/her ability to respond to motor, verbal and visual stimuli. Other qualitative factors are used to assess the effects of TBI, including loss of consciousness, loss of memory, alteration in mental status, focal neurological deficits, and visual assessment of neuroimaging studies when performed. However, visual assessment of neuroimaging is subject to poor inter-rater reliability and may not be sensitive enough to detect mild injuries, which often show subtle or no symptoms. The clinical definitions for diagnosis of mTBI have been recently clarified [4]. If GCS abnormalities are observed, computed tomography (CT) or MR imaging is used, depending on the severity of presumed injury, to assist in diagnosis though quantitative measures are not routinely used in the clinic. However, clinical studies

assessing potential quantitative methods suggest that this type of analysis can improve detection and treatment of mild injuries.

In clinical studies and those using experimental models of TBI, manual quantitative analysis has been used to evaluate TBI in MRI [5][6][7]. These manual studies have sought to identify the location and size of a lesion from MRI with correlative histology and assessment of long term neurological effects. However, manual detection of lesions in mTBI is very difficult and often requiring, a) hours per scan for manual region-of-interest analysis, b) a trained operator to improve inter- and intra-rater reliability, c) large data sets for statistically sound analysis but this is not feasible as it becomes resource intensive, and d) multi-modal inputs as mild or subtle alterations on MR images are often difficult to identify (low contrast). Further consistency in manual detection is a problem with both inter and intra operator variability. This low contrast appearance (as shown in Figure 1) on MRI exacerbates the consistency in manual evaluation as individuals interpret presumed abnormalities differently, despite strict protocols, leading to large inter-operator variations. This variability is even present in higher contrast images, such as tumor size quantification [8]. Partial volume effects along with neighboring anatomical regions can affect adjacent voxels making them seem to be of a different class [9]. Operator fatigue also plays a large role since manual detection can take a long time to analyze. However, manual segmentation is still considered the "gold standard" and automated algorithms are routinely compared to this standard.

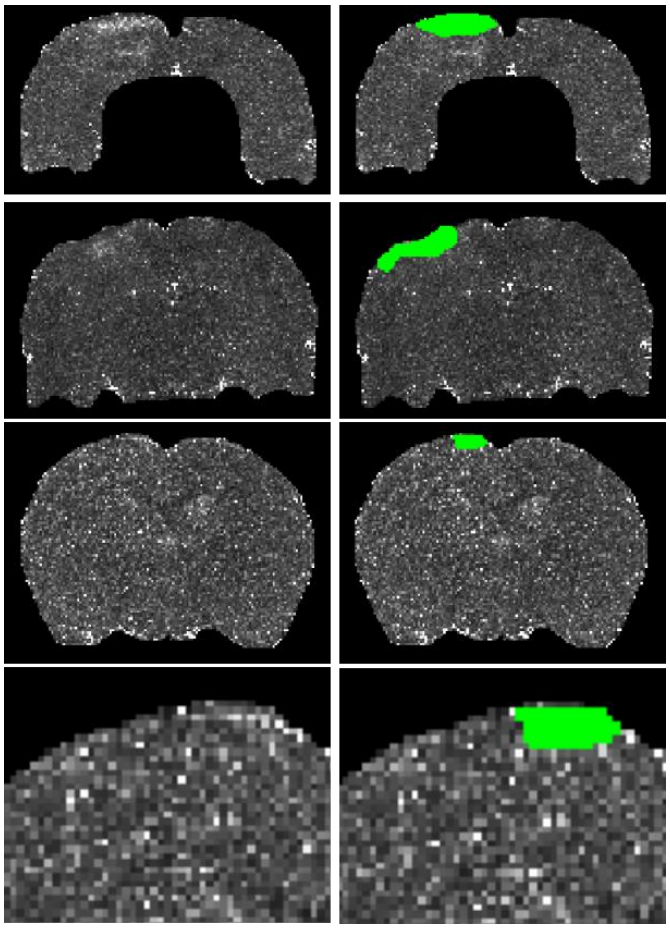

Figure 1: Examples of mTBI lesions in a rat (mildly controlled cortical impact model for an MRI T2 map). Left column - original coronal slices. Right column - manually drawn lesion regions. The last row is the expanded view of the row above it. Note the texture changes in the manually drawn regions.

On structural MR, gross tissue abnormalities, such as extravascular blood and edema (increased water content), can be identified in moderate to severe TBI; but these features are often not present (or little) on mTBI MR data. While CT is typically used in the initial assessment of head injuries and to identify extra vascular blood in the brain, new emerging technique, such as susceptibility weighted imaging (SWI), are even more sensitive to micro-hemorrhages [10]. T2-weighted imaging (T2WI) and the resultant T2 maps are routinely used to identify structural changes, edema and tissue relaxation

properties. In experimental models of TBI, T2 maps are used to identify and quantify regions of both edema (high T2 values) and extravascular blood (low T2 values). MRI T2 maps have been used for manual quantitative analysis of TBI [11]. However, edema, particularly in mTBI, can be subtle and transient further making identification of putative injury difficult. We have chosen to use T2 maps in this study due it its ability to find edema and blood in a single modality. Having a single modality decreases computation time and imaging time.

The key questions associated with detection of mTBI lesions are: single vs. multiple lesions and their locations; single/multiple lesions over time; focal vs. diffuse lesion; contralateral (both hemispheres) vs. ipsilateral (same hemisphere); coup vs. contra coup injury. For this paper we are concerned with modeling and analysis of single/multiple injuries (contralateral) over time where the injury is modeled by a mild control cortical impact (CCI) animal model, which only present a focal coup injury.

In Chapter 3 e explored the ability of computer vision and learning techniques to assess subtle alterations on quantitative T2 maps after induction of mTBI. The contrast-to-noise ratio (CNR) calculated from the manually segmented detected lesions to the normal appearing brain matter in our T2map dataset is 6 on average, where we estimated the noise variance of the lesion regions for the calculation of CNR [12].

In order to analyze such images in Chapter 3 we use a range of different models in our integrated approach that is proposed in this paper. These models include Bayesian networks (BN) for contextual modeling, support vector machines (SVM) for discrimination of lesion vs. non-lesion tissue based on visual appearance. In addition, we

used Gompertz function for effective representation of lesion progression and Fourier descriptor of brain perimeter for smoothing so as to reduce the effect low resolution and noise within MR images.

In Chapter 4 To estimate the mTBI lesion we use a discriminative approach comprised of voxel level classification utilizing visual and contextual features. Both static and dynamic contextual features are proposed. Static contextual features, proximity and directional, are calculated from local neighborhood information of the posterior marginal estimated by a preceding classifier. The dynamic context (posterior marginal edge distance) is calculated from the maximum posterior marginal of the cascade of classifier of a previous time point and then used to estimate the current lesion. Adding these contextual features overcomes the low contrast nature of the lesions.

Advanced MRI techniques examined for the analysis of TBI include diffusion tensor imaging (DTI) [13][14] and susceptibility weighted imaging (SWI) [15][16][17] DTI measures the "mobility" of water in the brain, and is used to identify white matter (WM) tracts. Significant WM tracts changes have been found in patients with mTBI, but it requires multiple imaging time points or a baseline image [18]. DTI based tractography becomes difficult to evaluate between patients. These limitations rule out the use of DTI for voxel based classification. SWI on the other hand has been shown to be three to six times more sensitive than traditional MRI methods in detecting microhemorrhages [15]. It utilizes the MRI phase images to estimate the magnetic susceptibility between tissues. In a recent study of mTBI patients [16] 19% were found to exhibit micro hemorrhages on SWI, while CT scans were negative. There have been few automated approaches for

finding lesions in SWI [17], however SWI tends to overestimate the physical mTBI lesions, both in size and in location. Therefore, in Chapter 5 we propose features that allow for co-operation between T2 and SWI modalities.

In Chapter 5 we propose a discriminative classification approach where features at every voxel are used to estimate its probability of being a lesion. The key contributions of this paper include: 1) development of temporally sensitive generative tissue model for estimating blood, normal appearing brain matter (NABM), edema, and noise utilizing the quantitative T2maps, and an estimate of the probability of each tissue type. 2) The use of new features based on the probability estimate of micro hemorrhage in SWI, and the use of texture features on T2 MRI echoes. 3) Also, we present experimental evaluations and comparison of results.

CHAPTER 2 RELATED WORK

A. Lesion Segmentation

To the best of our knowledge, there has not been another paper that has attempted to automatically detect mTBI abnormalities. Recently, a semi-automated approach [19] used an atlas based classification algorithm to detect lesions in severe TBI patients. The authors reported the use of a multi-modal MRI atlas of healthy human brains along with outlier detection to determine lesions which worked well in severe TBI due to the high contrast and large size of the lesions.

When comparing our proposed approach to previous works, other disease detections algorithms were considered. Some diseases that have been addressed by automated computer vision and image processing techniques are multiple sclerosis, tumors, and stroke. These diseases often have, on MRI, a high contrast appearance and a large injury size compared to mTBI, making them relatively less challenging to detect compared to mTBI. Table 1 has summarized some of the current methods utilized for lesion detection.

TABLE 1
OVERVIEW OF CURRENT LESION DETECTION METHODS FOR MRI DATA

Paper	Basic Idea	Disease	Notes
Kruggel et al., 2008 [23]	Co-occurrence features with PCA and the distance to the cluster centers is used to make a probability map of injury.	White Matter Lesions	+ Utilizes high dimensional texture features. - Registration and assumes single cluster per class (multiple appearances are possible).
Renske de Boer et al., 2009 [20]	K-nearest neighbor for initial tissue segmentation. Gray matter is reclassified to lesion or non-lesion using the FLAIR channel.	White Matter Lesions	+ Uses known tissue location priors for misclassification. - Registration, disease specific property.
Anbeek et al., 2004 [39]	Direct KNN classification with spatial information using previously known lesions in the training set.	White Matter Lesions	+ Incorporates voxel location. Allows for neighboring voxels to be considered. - Registration. Does not scale to large data.
Zhang et al., 2011 [40]	SVM classification followed by region growing.	Tumor	+ Prior segmentation knowledge and texture features. - Registration, highly depends on first segmentation.
Sun et al., 2009 [41]	Asymmetry detection followed by classification.	Stroke	+ No prior knowledge and symmetry aided. - Problems with symmetrical lesions and textured regions.
Seghier et al., 2008 [42]	Tissue segmentation with an extra class followed by outlier detection in the gray/white matter classes.	Various	+ Tissue classification with an iteratively learned abnormal class. - Registration, smoothing cause problems with small lesions.
This Chapter	SVM classification with texture features fused with a contextual disease model.	mTBI	+ Uses patient information, texture features, unimodal MRI - mTBI specific disease model.

Some approaches are atlas based, but the results vary since they are dependent upon the atlas being used [20]. Registration is used across many disease detection algorithms since it allows for multiple subjects to be put into a common registered space. This allows for statistical lesion extraction, through outlier detection, with a known atlas. However, brain anatomy is variable even among healthy brains, which causes a distribution of values to be learned at every voxel in the atlas. Registration can cause a distortion in the MR values due to interpolations in the calculation of the registration [21] [22]. These registration distortions make mTBI lesions from T2 maps nearly indistinguishable.

Texture features have been used to increase the discriminatory potential of MRIs [23] [24]. Many different types of texture features have been used such as neighborhood statistics, co-occurrence matrices, fractals, and more. Recently, [25] showed, using manually drawn ROIs, that mTBI causes texture changes in the brain with features based on region histogram statistics, co-occurrence matrix and gradient measures. Texture features are also used in this paper to increase the discrimination of our unimodal dataset. Most of the other approaches use multiple MRI modalities to detect the disease signature they we are trying to measure. Having multiple modalities allows us to measure multiple tissue properties, but leads to longer imaging and subsequent analysis time. Thus, analysis using a minimal number of modalities is essential for efficient results.

Context has been an active area of research in computer vision [26] [27]. It can be used to overcome low contrast detection problems. Context is defined [26] as, "any information that might be relevant to object detection, categorization and classification tasks, but not directly due to the physical appearance of the object, as perceived by the image acquisition system." Context types that have been described include: local pixels, 2D scene gist, 3D geometric, semantic, photogrammetric, illumination, weather, geographic, temporal, and cultural [28]. Object recognition in natural scenes has been one of the active areas for context [29][30][31][32][33]. Many of these approaches use context for position priming, which is the process of estimating the location of the object based on the context. Context can be split into two types local and global: a) Local context includes spatial relationships learned at the pixel/region level. It is used in conditional random fields and auto-context [34], and b) Global context can be thought of

as an estimation of the spatial location of an object. For instance, the sky usually occurs at the top of an image. Recently, Horsfield et al. [35] proposed using spatial priming in Multiple Sclerosis (MS) lesion detection. The spatial priors are learned directly from the data in a registered space. This approach works with MS since lesions occur regularly in white matter locations. However, using prior locations directly on mTBI data will not work since the injury can occur and progress at different sites within the brain. The proposed approach models the disease progression from any point of time in its progression.

B. Integration of Context

Context has been an active are of research in computer vision and can be used to overcome low contrast detection problems. Context is defined [36] as, "any information that might be relevant to object detection, categorization and classification tasks, but not directly due to the physical appearance of the object, as perceived by the image acquisition system." Context types that have been described include: local pixels, 2D scene gist, 3D geometric, semantic, photogrammetric, illumination, weather, geographic, temporal, and cultural [26]. Context can be split into two types local and global: a) Local context includes spatial relationships learned at the pixel/region level. It is used in conditional random fields [37] and auto-context [34], and b) Global context can be thought of as an estimation of the spatial location of an object. In this paper we exploit low (voxel) level context. Our previous approach [38] uses a contextually driven generative model to estimate the lesion location.

As a starting point algorithms for Multiple Sclerosis (MS) are examined due to similar low contrast lesions being present in MR images. The approaches for detecting MS lesions [39][23] use texture features and take advantage of the knowledge that MS occurs in a specific tissue type, which is consistently located in similar anatomical regions. This is a type of context (anatomical context) that can be exploited easily for MS, but mTBI lesions do not have this advantage since mTBI can occur in different tissue types.

Recently [34] proposed autocontext is a way to model context at the voxel level. The main premise is to estimate an object with a discriminative classifier and use a sampling of the estimated posterior probability as an additional feature to a subsequent classifier. It is able to take information from far away (across the brain) compared to other methods like conditional random fields (CRF) [37] which are local. The context features in this case are a sparse sampling of a distant neighborhood around every pixel. This can lead to overfitting due to the very specific locations and this can be seen in one of the examples in [34]. Other methods that exploit this type of context include [43], which has similar pitfalls as autocontext. These methods have been shown to segment rigidly positioned objects well, such as organs or body parts. In this paper we adopted the idea of cascading classifiers, but developed features that are more generalizable and integrate temporal information, which has not been addressed in [34].

C. Multi-modal Magnetic Resonance for Detection

Computed tomography (CT) is often the first and primary imaging method to aid in the diagnosis of TBI. The focus of the CT is to determine the presence of blood but

does not adequately reveal edema (excess fluid) in the injured brain. Additionally, CT requires the patient to be exposed to a high amount of radiation [25] which can be managed in adults but poses additional risks for children [44]. MRI imaging does not use radiation and thus allows repeated imaging to assess injury progression. Further, MRI provides multiple image modalities (e.g. T2, SWI) to assess brain pathology and/or function.

Past studies [45][5][6][7] have used MRI T2 maps to manually and automatically detect mTBI lesions in a rodent animal model. On T2 MRI mTBI lesions appear as small low contrast lesions, which are often a combination of blood and edema. Due to the low contrast and complex lesion composition, T2 values alone are unable to delineate the lesion. Within the T2 imaging modality texture has been used to increase its discriminative ability [25][38][46]. Significant texture changes have been shown in brain structures after mTBI [1]. We utilize texture in our paper using T2 echoes (T2e), which do not give physical relaxation values but assist in reducing noise relative values.

Advanced MRI techniques examined for the analysis of TBI include diffusion tensor imaging (DTI) [13][14] and susceptibility weighted imaging (SWI). DTI measures the "mobility" of water in the brain, and is used to identify white matter (WM) tracts. Significant WM tracts changes have been found in patients with mTBI, but it requires multiple imaging time points or a baseline image [18]. DTI based tractography becomes difficult to evaluate between patients. These limitations rule out the use of DTI for voxel based classification. SWI on the other hand has been shown to be three to six times more sensitive than traditional MRI methods in detecting microhemorrhages [15]. It utilizes the

MRI phase images to estimate the magnetic susceptibility between tissues. In a recent study of mTBI patients [16] 19% were found to exhibit micro hemorrhages on SWI, while CT scans were negative. There have been few automated approaches for finding lesions in SWI [17], however SWI tends to overestimate the physical mTBI lesions, both in size and in location. Therefore, we propose features that allow for co-operation between T2 and SWI modalities.

CHAPTER 3 VISUAL AND CONTEXTUAL MODELING

A. Contributions

The contributions of this Chapter are: 1) A contextual disease model based on Bayesian networks to estimate the spatial location of mTBI abnormalities. It allows for a range of contextual inputs, when the exact value (time or approximate location) is not known, which could be due to patient's memory loss. This contextual model helps to overcome the low contrast appearance of mTBI by focusing on the region of the search space that may have injury. 2) A visual model that is learned using texture features to build a probabilistic support vector machine (PSVM). The method does not require registration (it is spatially invariant), which can cause distortion. 3) A novel dataset of rat MR images is developed to follow mTBI to test the validity of the proposed model. The rat controlled cortical impact (CCI) mTBI model is used with multiple observations post-injury using T2 MRI maps to demonstrate the results. This paper is an extension of [36]. It extends the approach to repeated TBI and a more in-depth analysis of the contextual inputs is given.

B. Technical Approach

1. System Overview and Technical Rationale

The proposed system is a combination of a visual and contextual model. Figure 2 shows the system diagram. A database of known mTBI MRI brain and lesion volumes with manually detected lesions and the associated contexts is used to build the model. The visual model uses 3D texture features to build a Probabilistic Support Vector

Machine that describes the lesion/normal appearing brain matter (NABM) space. The output is a probability map that is combined with the contextual model, where higher probabilities signify higher chances for a particular voxel being part of the lesion, where the lesion is an estimate of both blood and edema. If the lesion splits into multiple parts, it will still be captured since a voxel based classifier is used. SVM is a discriminative model which performs well with a large amount of data and can describe a complex decision space. With a voxel based classifier, both these conditions are satisfied.

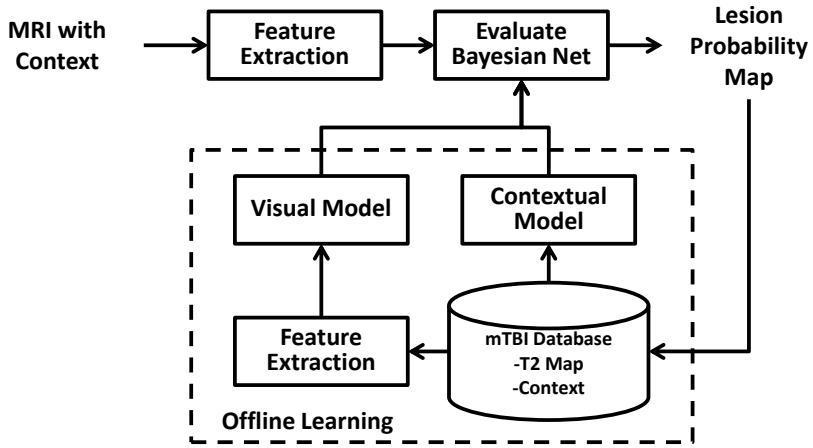

Figure 2: System flow diagram.

The contextual model uses a Bayesian network (BN) to estimate the locations of the mTBI lesions (Figure 3). This model uses knowledge about the subject, where both temporal and spatial information are used to describe the development of the disease. The BN is a generative model that is able to extrapolate based on little information when the underlying distribution assumptions hold. Generative models are used to simulate cause and effect relationships and processes, while discriminative models such as SVM do not

have this ability. Our system combines the advantages from both discriminative and generative models.

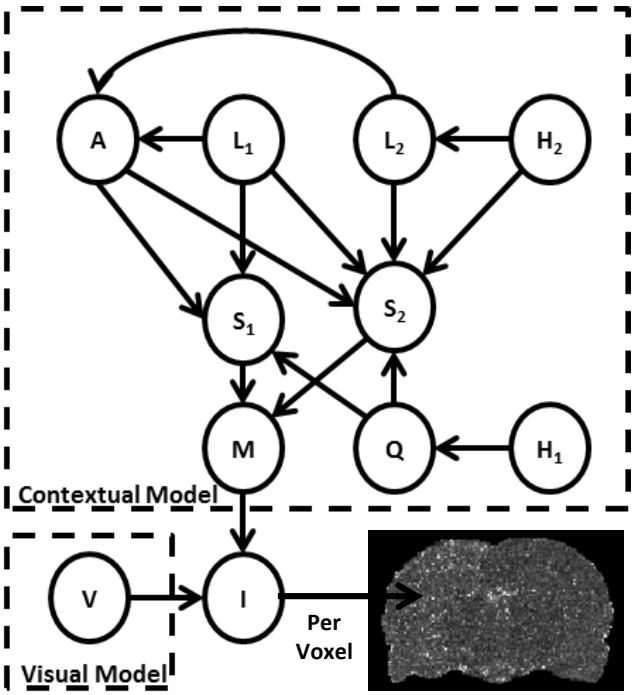

Figure 3: Graphical representation of the Bayesian network, showing the dependencies of random variables. Intuitive definition of distributions: A – anatomical constraints, L_1, L_2 – central location of injury for injury one and two, H_1 – time since the first event, H_2 – time between first and second injury, Q – volume (quantity) of injury with time, S_1, S_2 – spread for first and second injury, M – max operator, V – visual information, I – estimated injury. Where I = 1 is absolute certainty of injury and I = 0 is absolute certainty of NABM.

When a volume enters the system the texture features are extracted and then the Bayesian net is evaluated. Contextual information is also passed to the system which includes an estimated location of the focal points of impact (L_1,L_2), time since the first injury (H_1), and time between the two injuries (H_2). These inputs can be exact (e.g., 1 day since injury) or a range of values (e.g. 5-14 days since injury). Each of the major steps will be described in the subsequent subsections.

2. Bayesian Net for Contextual Modeling

Modeling a joint probability distribution with many variables that interact with each other is an extremely difficult task and requires a large number of training samples. At the other extreme, naïve Bayesian learning is simple requiring a reduced numbers of samples but often poor results are observed due to features not being independent. BNs allow for a middle ground, where dependencies among variables can be modeled in a graphical manner. Directed acyclic graphs are used to represent BNs, where nodes are random variables and edges represent dependencies among random variables [47]. A BN with nodes N_1, N_2… N_i results in the full joint distribution:

$$p(n_1, n_2, \dots, n_i) = p(n_1)p(n_2|n_1) \dots p(n_i|n_1, n_2, \dots, n_{i-1}) \quad (1)$$

$$= \prod_{q=1}^{i} p(n_q | n_1, n_2, \dots, n_{i-1}) \quad (2)$$

The properties of the BN allows for a node to only depend on its parents' values, so the joint distribution can be re-written as:

$$= \prod_{q=1}^{i} p\left(n_q | parents(N_i)\right) \quad (3)$$

A node in a BN is only conditional on its parents' values. This means that each probability distribution is conditionally independent of all its non-descendants in the graph given the values of its parent.

3. Description of Random Variables

mTBI, like most other injuries, evolves with time, so the contextual model (see Figure 3) has multiple random variables that evolve with time. H_1 is the distribution of the time since the first injury. This random variable can be learned directly from the data or it can

be modeled as an exponential distribution since it is most probable that a patient will arrive soon after the injury. Q is the random variable for the volume of injury over time. This distribution can also be learned from the data directly or modeled as a lognormal distribution. This distribution follows the natural progression of the injury where there is an initial peak in gross abnormality that resolves slowly overtime. As an example, Figure 9(f) shows that gross pathology (in our data set) is maximal during the acute phase of the injury, followed by a sharp drop during the sub-acute phase and it approaches a steady state in the chronic phase. H2 is the time between the first and second injury (i.e. repeated mTBI). After the first injury metabolic and cellular injury cascades occur which may lead to a window of vulnerability in which a second mTBI may worsen ongoing pathology [48]. This function can also be modeled directly from the data or using an exponential distribution from the time since first injury. The process of subsequent injuries can be thought of as a Poisson process, which has the time between injuries as an exponential distribution. H_1 and H_2 are variables that would be dependent on location and can be estimated through a regional epidemiological study.

4. Anatomical Constraints

Gross tissue level-pathology observed following mTBI may expand and contract throughout its duration, but tissue structures provide natural barriers within the brain that restrict the spread of injury. For the rat brain model, two strong barriers are observed. The first boundary is the midline which is a physical separation between the left and right hemispheres. If the injury occurs on the left hemisphere it will not progress to the right. However, this may not be the case with severe and moderate injuries as the injury will

extend throughout the brain. The second boundary that was observed is the major white matter tract (corpus callosum) that bounds the mild nature of the injury. To give an estimate of the location of the corpus callosum, the top 1/4th of the brain was used. These two constraints are represented in the form of a 3D binary mask.

5. Location

An estimate of the location of the focal impact point is essential to modeling the progression of the injury. In our current model we account for two possible injury sites modeled by distributions L_1 and L_2. Examples of these distributions, estimated from our dataset, are shown in Figure 9. Unlike other diseases that occur in specific locations/structures in the brain, mTBI progresses from an impact site which may occur at almost any point on the perimeter of the brain. To model the distribution of focal impact points, the perimeter distance in the coronal slices from the midline and the z location are used to make a 2D distribution. The values are binned to make histograms. A smaller number of bins give a better generalization of the distributions when the sample size is small. When the sample size is larger the number of bins can be increased while maintaining a generalized distribution.

6. Spread

Eq. (4) is used to describe the injury progression given all the contextual inputs. This function is a sigmoid like function, known as the Gompertz function [49], which has parameters (m, n) to control the shape and displacement. The Gompertz function has

more flexibility than the logistic function and less number of parameters than the generalized logistic function. The effects of these parameters are shown in Figure 4(b).

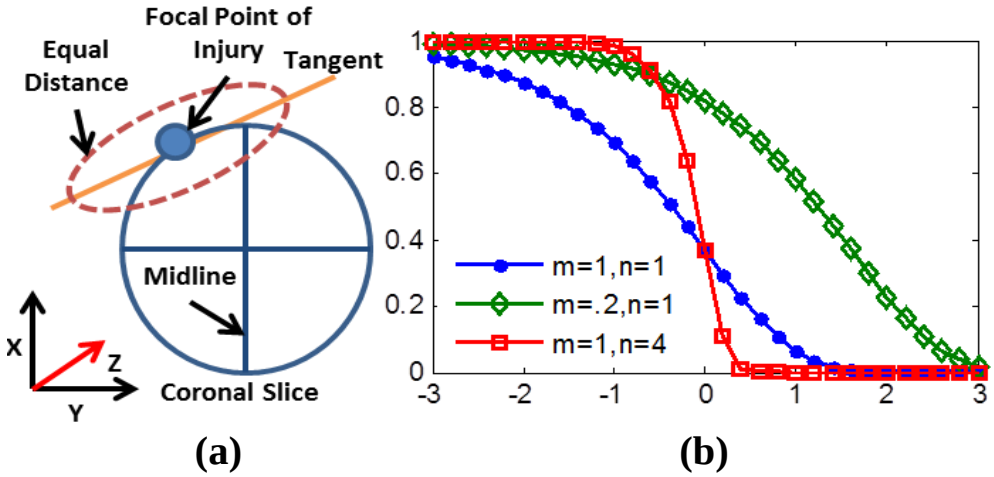

Figure 4: (a) Description of the space and distance function. The XY plane is the coronal plane and the YZ is the sagittal plane. (b) Effect of parameters (*m*,*n*) on the Gompertz function, with the other parameters being held constant.

Parameter m determines the displacement of the function and n determines the decay rate. The shape and displacement parameters are not the only variables that affect the shape of the curve; Q (quantity of lesion) also determines the shape. When Q is large, the function will shift to the right and have a more gradual slope. This represents a larger area for potential injury and more uncertainty in the location. When Q is small, the opposite is observed, the area is small and there is more certainty in the location. Q is estimated by taking the average lesion size at each time point for all the volumes in the database. There is a separate set of parameters for each H2 value. This is due to the shape of the injury being different when the repeated injuries are at distinct times from each other. When the first hit occurs, there are many cellular cascades that take place

throughout the progression of the injury potentially leaving the brain vulnerable to subsequent mTBIs [48]. This has been observed in animal models [48][50].

$$p(S|L,A,H,Q) = Ae^{\frac{-m_H}{Q}e^{\frac{n_H}{Q}d(x,L)}} \tag{4}$$

As described in eq. (5), d is a distance function weighted by Σ. It measures the distance from L to every other point in the 3D space. Σ weights the 3D space so it accounts for rotation and scaling. The parameters described in equations (7-9) describe the rotation and scaling in the coronal space, and σz describes the scaling in the z space. The parameters that need to be set in these functions are σx and σy, which control the scaling in the direction of the tangent and the direction of the normal, respectively. Finding the tangent angle θ at L is done by utilizing the Fourier descriptor [51] of the closed perimeter of the brain slice. When converting to Fourier space a low pass filter is applied that cuts off the upper twenty percent of the frequencies (Figure 5). The parameters, a, b and c control the rotation in the coronal plane using θ, to create the new "axis" for the distance function. Eq. (4) models the progression of mTBI, since the injury spreads along the perimeter of the brain more than into the center of the brain. Hence, the distance function is weighted more along the tangent of the perimeter.

$$d(x,L) = \sqrt{(x-L)^T \Sigma (x-L)} \tag{5}$$

$$\Sigma = \begin{bmatrix} a & b & 0 \\ b & c & 0 \\ 0 & 0 & \sigma_z \end{bmatrix} \tag{6}$$

$$a = \frac{\cos^2\theta}{2\sigma_x^2} + \frac{\sin^2\theta}{2\sigma_y^2} \tag{7}$$

$$b = \frac{\sin 2\theta}{4\sigma_x^2} + \frac{\sin 2\theta}{4\sigma_y^2} \tag{8}$$

$$c = \frac{\sin^2\theta}{2\sigma_x^2} + \frac{\cos^2\theta}{2\sigma_y^2} \tag{9}$$

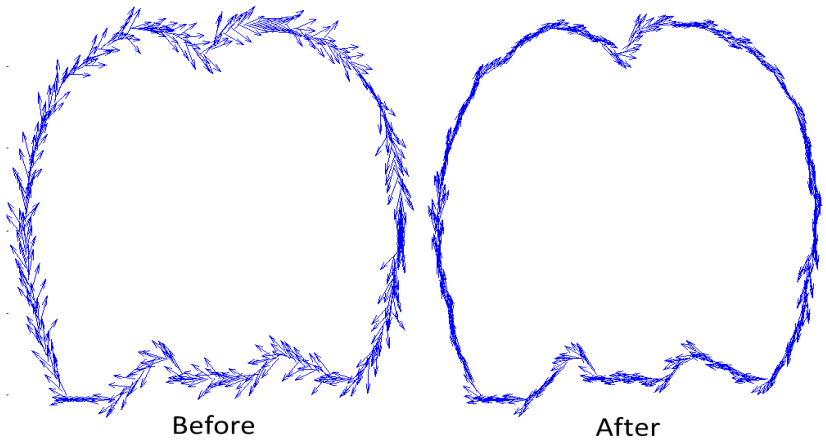

Figure 5: Example of tangents around the perimeter of the brain. The tangents are scattered due to the low resolution of the boundary. The high frequency noise is nullified after converting to the Fourier descriptor and filtering. The tangents are accurately estimated after filtering.

The contextual model is finalized in eq. (10). An example of the context over time is shown in Figure 6. This function takes the maximum of the two spread functions evaluated over each value of the range of each contextual input. The output of the contextual model is an estimate of the lesion extent. If a contextual input is known then the probability of it being that value is 1. For example, if H_1 is known to be 3 days then p(H1,3) = 1 and all other values in the distribution equal zero. Another example is when a ranged input is given. If H_1 is known to be between 5 and 14 days, the distribution becomes the priors known at those values normalized by the sum of the probabilities at those values (all other values are set to zero). If one of the contextual inputs is not known, all values in the prior distribution are considered.

$$p(M|S_1,S_2) = \sum_{\forall i|P(L_i)>0,\ \forall j|P(Q_j)>0,\ \forall k|P(H_{2k})>0} \sum \sum \max\left(p(L_i)p(Q_j)p(H_{2k})p(S_{1ijk}),\ p(L_i)p(Q_j)p(H_{2k})p(S_{2ijk})\right) \quad (10)$$

$$p(I|M,V) = p(M)p(V) \quad (11)$$

An estimation of the lesion is finally given by eq. (11). The contextual model and visual model are independent. This assumption has been made by other approaches using context [26]. The final output is a probability map.

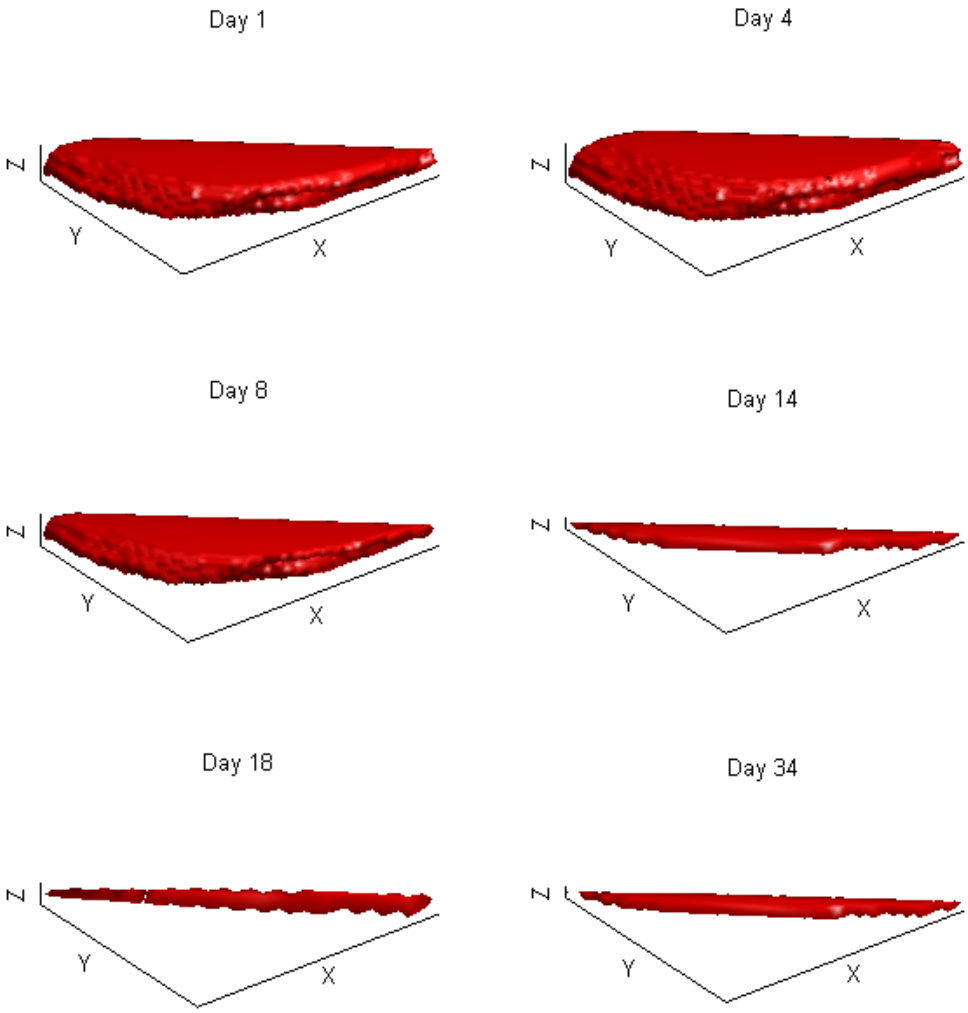

Figure 6: An example of the contextual model over time. The surface is the model manually thresholded at 0.2 to give a solid surface that illustrates the progression of the model over time with all other variables constant. The XY plane is the coronal plane. Note, the best threshold for lesion detection is given as the threshold that maximizes the dice curve. An example of this is shown in Figure 12.

7. Parameter Estimation

The parameters in the spread function (4) need to be learned. These parameters are learned from the known detected lesions in the database of mTBI volumes represented as a binary volume B, where 1 is lesion and 0 is NABM. r_{ij} is the set of variables (see eq. (12)) that need to be learned in the spread function (see eq. (4)) for each time of second impact (H_{2i}) and each anatomically constrained hemisphere A_j (left and right).

$$r_{ij} = [m_{ij}, n_{ij}, \sigma_{xij}, \sigma_{yij}, \sigma_{zij}] \tag{12}$$

We use weighted nonlinear least squares [52][53] to estimate the parameters. The cost function is given by eq. (13).

$$\widehat{r_{ij}} = \underset{r_{ij}}{\operatorname{argmin}} \sum_{\forall\, B \in H_{2i} \wedge A_j} \omega \left(p\left(S_{r_{ij}} \middle| L, A, H, Q\right) - B \right)^2$$
$$\text{Where} \begin{cases} \omega = 1, if\ B = 0 \\ \omega = \frac{\#\ NABM\ Voxels}{\#\ Lesion\ Voxeks}, if\ B = 1 \end{cases} \tag{13}$$

Each contextual input is known from the database for evaluating the spread function. Eq. (13) finds an estimate $\widehat{r_{ij}}$ that minimizes the sum of the differences between the spread function ($S_{r_{ij}}$) and the ground truth (B) at every voxel. The weight ω is used to give the lesion a higher value in the cost function since there are many more NABM voxels compared to lesion voxels.

8. Visible Model

The visible model is the probability of an injury given the visual appearance of a sample voxel in the MRI. The MRI modality that has been chosen for evaluation of mTBI

injury is T2WI and its quantitative T2 map. This modality allows for observation of both edema (increased signals) and extravascular blood (decreased signals), which are both potentially at the site of injury. Edema presents itself as high intensities and blood occurs as a low intensities. Note that we consider blood and edema as a single class. Since the edema and blood are believed to have local tissue effects, the texture around known lesions are used. The texture features that are used in this study are 3D local mean, variance, skewness, kurtosis, entropy, range, gradient x, gradient y, gradient z. An example of these features can be seen in Figure 7. These local statistics were chosen since they are able to describe small local neighborhoods at every voxel, which gives an estimate of the tissue. Note that partial voluming has less of an effect since the tissue is estimated and not a single voxel. The Bhattacharyya distance is a measure of the similarity of two distributions and can give a bound on the Bayesian error [54]. There are a total of ten features including the original T2 values. These features go through the feature selection process using the Bhattacharyya distance, to give the final feature set. The size of the filter is an important parameter and is found by finding the using the same feature selection process. The local neighborhood size found to be most discriminative is 5x5x5. For our current dataset the features that are selected are: T2 value, entropy, variance, skewness, gradient x, T2 mean and gradient y.

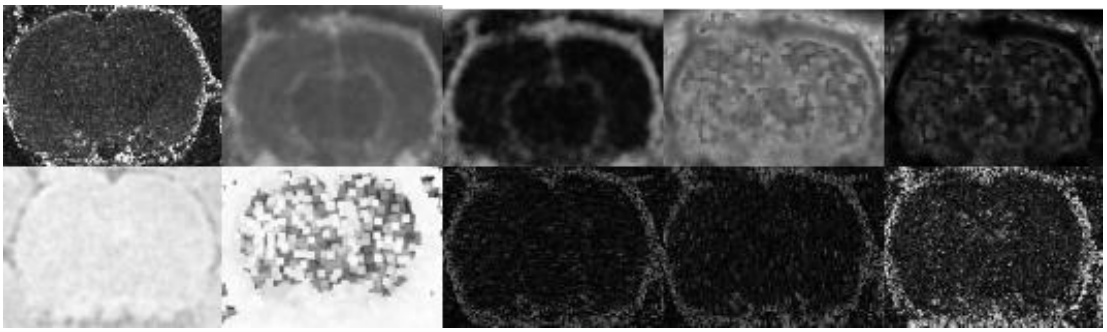

Figure 7: Example of features with a 2D coronal slice. Left to right, top to bottom: T2 map, mean, variance, skewness, kurtosis, entropy, range, gradient x, gradient y, gradient z.

Support Vector Machines (SVM) are one of the leading supervised classifiers in machine learning. It is a max-margin classifier, which means it gives a decision boundary that maximizes the distance between the two classes. Describing the boundary using a maximum margin has been shown to provide a high generalization. The kernel form of SVM is utilized to give non-linear boundaries. A radial basis function kernel is used to project the data to a high dimensional space making the separation linear. There are two parameters that need to be set when using the radial basis function, the soft margin parameter and the kernel shape parameter [55]. These two parameters are found through grid search that minimizes the error in a cross-validation test. SVM is generally used as a hard classifier with the output being one class or the other. However, Platt [56] proposed a framework for Probabilistic SVM (PSVM) which allows for estimation of the posterior probability p(class|features). A sigmoid (with two parameters) is fit to the SVM result giving an estimate of the posterior probability for each class. LIBSVMs [55] implementation of PSVM is used for testing. An example of the final output of the two models is shown in Figure 8.

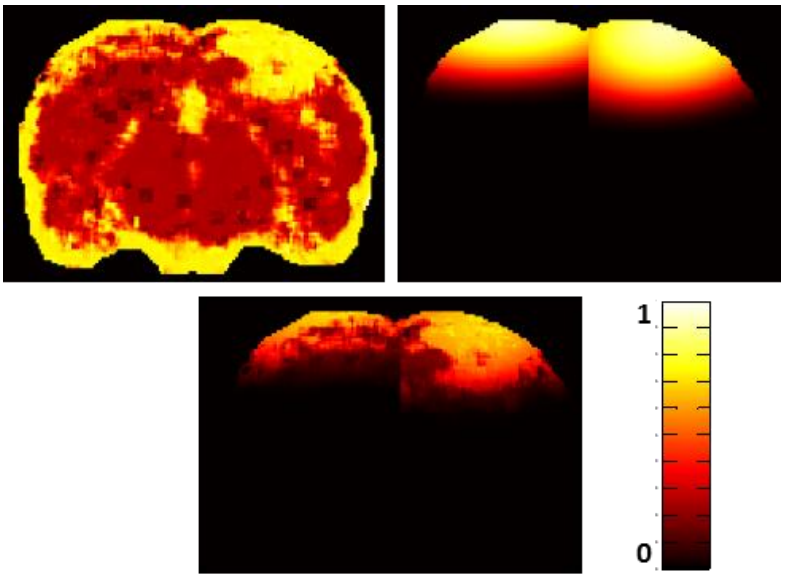

Figure 8: Example of fusion between visual model and contextual model on a contralaterally injured rat. Top left: Probability map after PSVM. Top right: Probability map from the contextual model. Bottom: Fusion of the contextual and visual model. Note this shows that the contextual model for repeated injuries progress at different rates.

C. Experimental Results

1. Dataset

Adult Sprague Dawley rats were randomly assigned to three experimental groups: Single mTBI, and repeated mTBI (rmTBI) induced 3 days or 7 days apart. A controlled cortical impact (CCI) was used to induce mild injury [6] [45]. The CCI model has a focal injury. Briefly, a craniectomy (5mm) was performed over the right hemisphere (3mm lateral, 3mm posterior to bregma) where a mild CCI was induced using an electromagnetically driven piston (4mm diameter, 0.5mm depth, 6.0 m/s). The incision made for the craniectomy is sutured after the injury is induced. At 3 or 7 days after the

initial injury, animals within the rmTBI groups (repeated injury) received a second craniotomy and mCCI was induced on the left hemisphere using identical parameters.

In vivo T2 weighted images (T2WI) were collected as multiple 2D slices [6]. Images were obtained at 1, 4, 8, 14, 21, 30 and 60 days post initial injury using a 4.7T Bruker Avance (Bruker Biospin, Billerica, MA). T2WI sequences (TR/TE=3453ms/20ms; 25x1mm slice) were collected with a field of view of 3cm (256x256 matrix). The resolution of each voxel is 0.12x0.12x1mm. On average 12 coronal slices were used to cover the cerebrum. T2 maps were computed using custom written software. Regions-of-interest (ROIs) were manually drawn using Cheshire imaging software (Haydan Image/Processing Group) [45]. ROIs included normal appearing brain matter (NABM) and cortical tissue containing T2 observable abnormalities using criteria from [45]. There were a total of 81 volumes in the dataset with time points as shown in Figure 9(e). Manual outlines of the brain were used to separate the brain from the skull and other tissues.

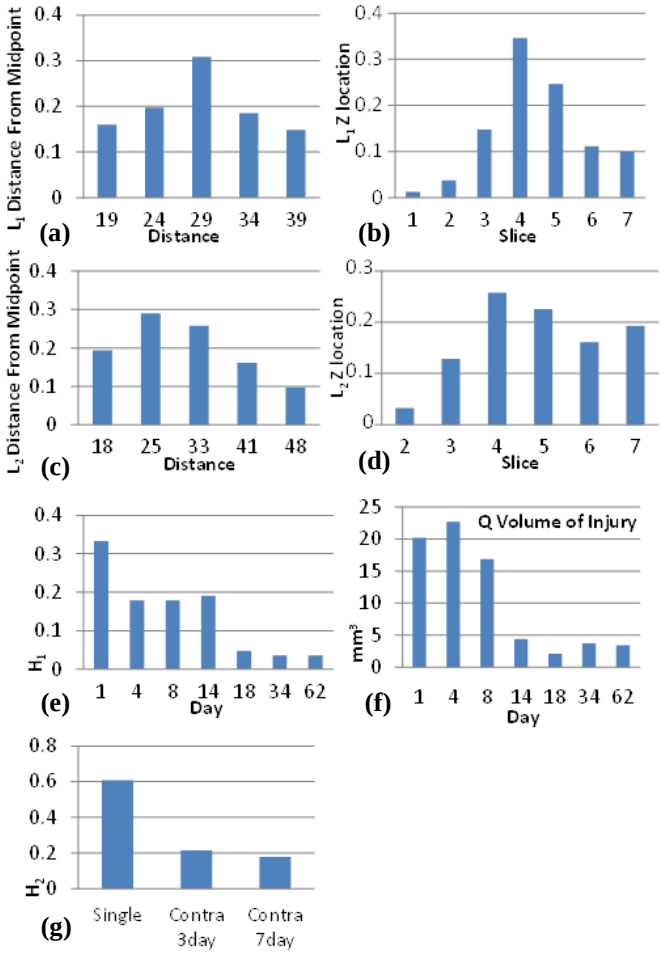

Figure 9: Distributions learned from the database for each of the nodes in the Bayesian network. Note the number of bins used in: (a) and (c) is 5, (d) is 6, and (b), (e), (f) is 7.

2. Performance Measures

To provide a sense of where errors are coming from and how the regions are correctly detected, the following measures were used. Sensitivity or true positive rate (TPR) eq. (14) measures the percentage of the target object that is correctly detected. Specificity or true negative rate eq. (15) is the percentage of the background that is correctly detected. 1-Specificity gives the false positive rate (FPR), which is the false

alarm rate. A receiver operating characteristic (ROC) curve is shown for each of the tests. A threshold is set on the probability map output from the algorithm to give a binary decision surface. Each threshold creates a point on the ROC curve and together they form a graph that represents the tradeoff between TPR and FPR. ROC has been a strong measure of performance in radiology and other medical disciplines [57]. Only FPR under 0.2 are considered for clinical relevance.

$$Sensitivity = \frac{TP}{TP + FN} \tag{13}$$

$$Specificity = \frac{TN}{TN + FP} \tag{15}$$

$$Dice = \frac{2TP}{(TP + FP) + (TP + FN)} \tag{16}$$

The Dice coefficient eq. (16) measures set agreement, which is the intersection of the known object and the detected object divided by the size of the known and detected objects. Dice gives a better idea of the intersection between the detected object and the actual object since it does not include true negatives in the calculation.

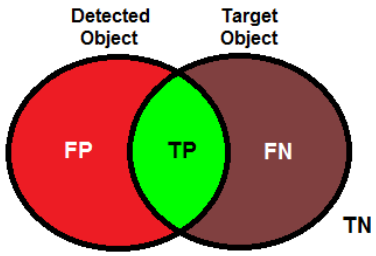

Figure 10: Visual meaning of measurements. FP – false positive, TP – true positive, FN – false negative, TN – true negative.

3. Varying Levels of Context

In these experiments each contextual input was tested to verify the effects they have on the model using our real data. Testing was conducted using the leave-one-out approach

for learning the parameters. The following cases were examined: all contexts known, L_1 and L_2 unknown position (unkPOS), H_1 unknown time (unkTIME), L_1 L_2 H_1 unknown (unkALL), and V alone (Probabilistic SVM). Figure 11 is the ROC plot for each of the aforementioned context tests after thresholding the output probability maps to get a hard classification. When all the context is known an equal error rate of 0.93 is achieved, and with missing context a decrease in performance is obtained.

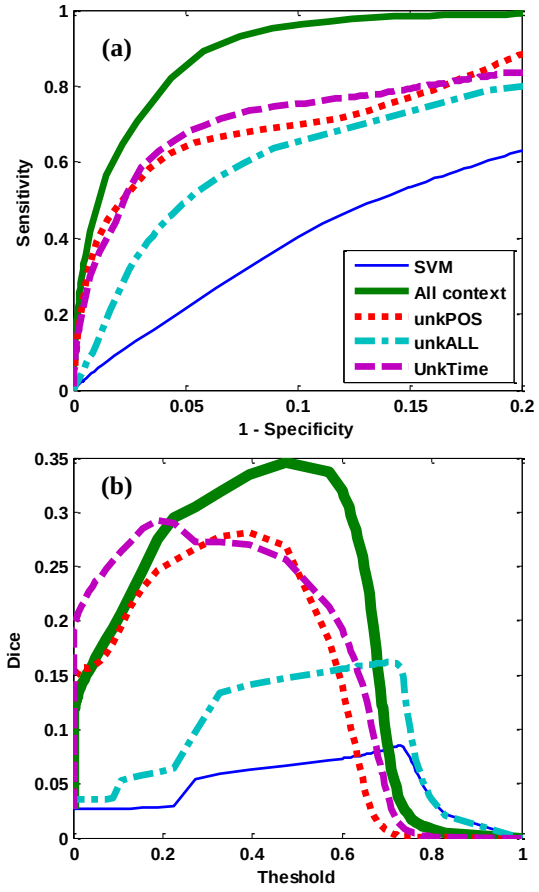

Figure 11:(a) ROC plot with multiple contextual inputs. Results were evaluated after thresholding the output probability map (from eq. (11)). (b) Dice plot with multiple contextual inputs. The peak of the Dice curve (maximum dice coefficient) is used to threshold the output probability map. Legends in (b) are the same as in (a).

The Dice plot in Figure 11 shows results corresponding to the results in the ROC plot. With unkPOS and unkTIME a similar performance is observed since each of these cases causes a smoothing effect in the output of the contextual model. This smoothing effect is clearly seen in Figure 12. Essentially, when the context is not known, a wider area has to be considered. In the case of unkPOS smoothing is experienced along the edges of a brain since the focal point is shifted along the perimeter of the brain. The unkTIME case has smoothing in all directions, which is due to the time points with a large amount of lesion having the largest effect on the output from the contextual model. In unkALL both of the smoothing effects occur.

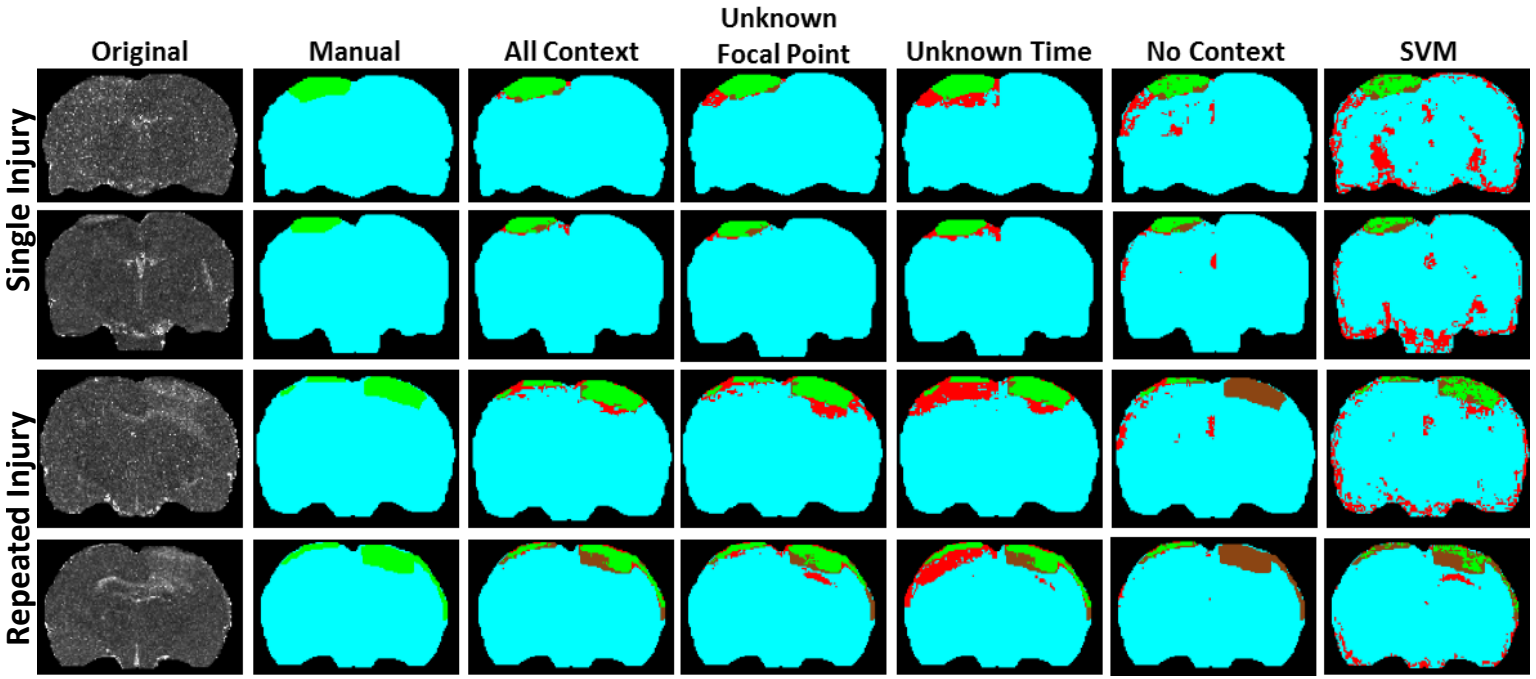

Figure 12: Example outputs from the thresholded probability maps (for single and repeated mTBI) with varying levels of contextual input. The thresholds are selected at the highest point on the dice curve for each respective level of context. Each row is a coronal slice, where the top two are from a single mTBI rat and the bottom two are from a repeated mTBI rat. The colors represent the following: Green: true positive, Teal: true negative, Red: false positive, Brown: false negative. (Note: This Figure should be viewed in color.)

4. Class Analysis

Evaluating the performance of each class (single, contra3, contra7) gives an insight into the potential errors of the model. From Figure 13 it is clear that the single mTBI class has the best performance. This is due to the single mTBI class being more consistent allowing the model to estimate the area of injury better. The distributions for L_2 (in Figure 9(C)) show that the second injury is less consistent than the first. Also, there is more variation in the shape and size of the second mTBI.

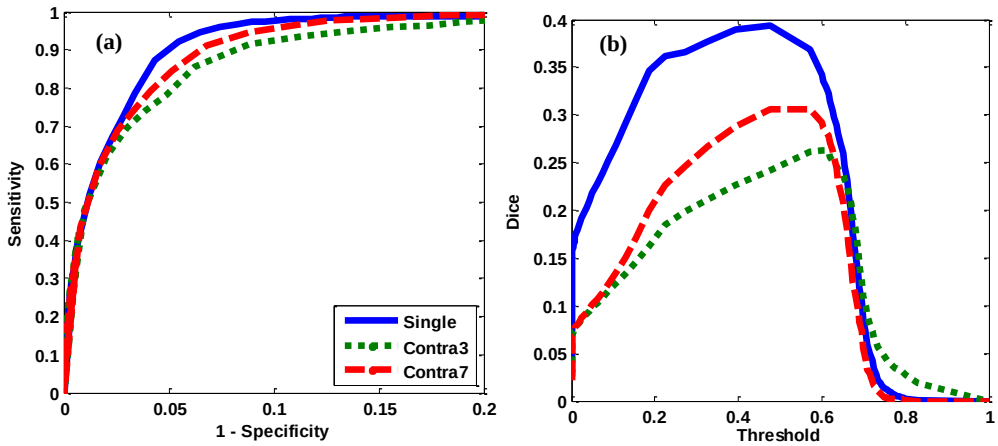

Figure 13: (a) ROC for each class. (b) Dice plot for each class. Definition of classes: Single-only a single injury, Contra3-contralateral injury 3 days after the first injury, Contra7-contralateral injury 7 days after the first injury. Legends of (a) and (b) are the same.

5. Ranged Context

In this test the effect of having ranged values for contextual input is examined. The known locations are first binned into the distribution shown in Figure 9(A) to 9(D). From the known location a range of values is obtained by taking a set number of bins on either side of the known location (Figure 14). For example, if L_1 is known to be 29 voxels from the midline and the range input is to be 2 steps away, the range input would be from

19-39 voxels from the midline (values from Figure 9(A)). The results from range testing are shown in Figure 15. From these plots it is evident that the performance increases as the position is closer to the known location. As the range gets wider the result approaches the performance of having the position unknown. This is due to the contextual model being able to focus when the values are known and hence follow the progression of the injury better. Ranged values are possible inputs to the system and give better performance than having completely unknown inputs.

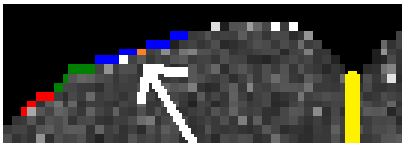

Figure 14: Example of the ranged inputs on H_1. The midline is shown in yellow. Orange (arrow) is the focal impact point, blue -1 away, green -2 away, red -3 away. Each higher range includes the previous. For example, 2 away includes 1 away. Increasing the away, increases the uncertainty in the position. It represents including the value in the prior histogram (Figure 9(e)) some step away (left and right) of the true point.

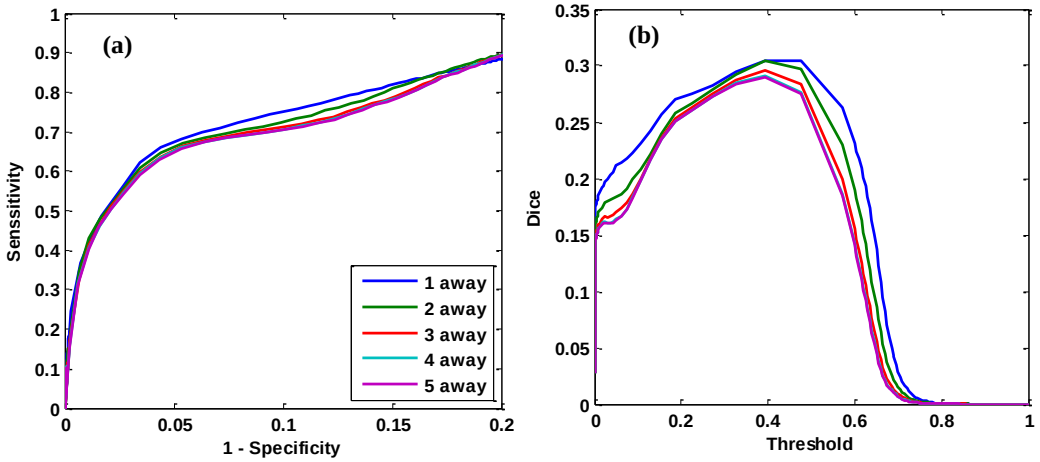

Figure 15: (a) ROC plot with varying distance input. (b) Dice for varying input distance. Legends of (a) and (b) are the same.

Comparison with State-of-the-Art Approaches

The proposed approach is compared to two other state-of-the-art approaches. The compared approaches are setup as described in the papers [39][41]. First, the approach of Sun et al. [41] is used since it is a purely data driven approach, which does not account for contextual or database information. This approach uses symmetry to find asymmetric lesions in the brain. This approach does not use any prior knowledge of the brain except that it should be symmetric. The user set parameters were optimized using three example T2 maps from the dataset, and then the parameters were fixed for testing of the entire dataset. The second approach that was chosen to compare against is that of Anbeek et al. [39]. Their approach is pure model-based approach, which takes advantage of visual information in the database. The volumes are registered, with rigid registration, and then a KNN model is built from the known lesions in the dataset. Since unimodal MRI data are used, the texture features are used in their approach as channels. The parameters were chosen as described in [39] and each feature was normalized.

The results from Table 2 demonstrate that the performance of our proposed approach outperforms the compared state-of-the-art approaches. All of the approaches tend to underestimate the lesion area [41], had difficulty in detecting the mTBI lesions due to the subtle T2 value differences and the lesions were mostly comprised of textured area. These two problems caused [41] to consider lesion within NABM. The symmetry approach [41] is best suited for stroke data where the lesions are more uniform and have higher contrast. The pure model-based approach [39] also had difficulties due to the low number of lesion voxels and the injury not occurring within the same locations

consistently. A large amount of data is required for a viable K-NN classifier and the lesions must occur in a similar location. MS is a disease where the model-based approach [39] excels since the lesions are primarily identified in the white matter locations. Also, a proper ratio between diseased brains and normal brains would have to be maintained to give a proper normal versus abnormal bias.

Run time is also a consideration when determining the effectiveness of an algorithm. All testing was carried out on the same computer system (PC with Intel Xeon (E5345) Quad CPU 2.33GHz and 2GB of RAM). The proposed contextual and visual approach has the fastest computation time. A majority of the time is spent in the PSVM calculation. The pure model-based approach [39] had increased computation time, due to each sample needing to be compared to every sample in the database. Some increases in speed could be made to the K-NN algorithm, but when the database becomes large computation time will become large.

TABLE 2
COMPARISON OF ALGORITHMS

Method	TPR	FPR	Dice	Time
Sun [24a]	0.06 (0.01-0.11)	0.003 (0.0-0.015)	0.02 (0.0-0.05)	~10min
Anbeek [22a]	0.07 (0.01-0.13)	0.002 (0.0-0.012)	0.13 (0.057-0.203)	~24hrs
PSVM (*)	0.73 (0.63-0.83)	0.27 (0.173-0.367)	0.07 (0.014-0.125)	~4mins
Proposed (*)	0.93 (0.87-0.99)	0.07 (0.014-0.126)	0.25 (0.156-0.344)	~5mins
Proposed (**)	0.57 (0.46-0.68)	0.015 (0.0-0.041)	0.35 (0.246-0.453)	~5mins

Time is in the online phase of the algorithm. The range (-) is the 95% confidence interval. (* – at the equal error rate, ** – at the highest dice). The time is for an entire brain.

7. Applicability of the Proposed Approach to Human Data

In this paper we showed single and multiple injuries over time where: (a) Our model accepts an input as a brain volume with contextual information at a single point in time as an input. Multiple volumes are not need to ensure an accurate estimate. (b) The location of the injury may be focal, diffuse or unknown. If the location is unknown then prior distributionsL_1 and L_2 (shown in Figure 9) are used as possible locations, where each location is tested and weighted by the distribution. If the injured region is diffuse then a ranged input can be given (as shown in section 4.5). (c) We have focused on focal lesions in this paper. Note that for diffuse lesions it is extremely difficult to get ground truth and one has to rely in histological data. This topic will be explored further in the future. (d) Multiple lesions can be handled by our approach (see Figure 12). Over time

the lesions may collapse into a single lesion or break into multiple lesions. These lesions can be detected using visual and contextual modeling. The model will need to be updated for cases where there are more than two injuries.

D. Conclusions

An approach was presented that incorporates visual and contextual knowledge to detect mTBI lesions using unimodal MRI, specifically T2 maps. Texture features are used with Probabilistic SVM to create a visual model of the mTBI appearance on the T2 MRI maps. The contextual model utilizes a Bayesian network which allows for exact or partial contextual inputs to be given to the system. Qualitative and quantitative results of the algorithm were shown using a novel rat mTBI dataset. The results show when more contexts are known, we obtain injury detections that are highly similar to the manually detected lesions by a trained observer. An equal error rate of 0.93 was achieved. The proposed algorithm outperformed other state-of-the-art-approaches, including computation time performance. This work will be extended to human data through the use of large mTBI databases and some modification. New animal models will be explored that include diffuse and contra coup injuries. Integration of susceptibility weighted imaging and diffusion tensor imaging is expected to improve the performance of the visual modeling and improve the contextual model through the estimation of WM tracts.

CHAPTER 4 DYNAMIC LOW LEVEL CONTEXT

A. Contributions

The contributions of this Chapter are: 1) development of three new contextual features to be used with a cascade of classifiers. These features include a proximity feature, a directional feature, and a maximum a posteriori edge distance feature. 2) Use of a temporal sequence of MR images to provide context from a previous time point, capturing the dynamics of the injury progression automatically. 3) Analysis of multiple contextual feature configurations on a rat controlled cortical impact (CCI) mTBI model dataset.

B. Technical Approach

1. System Overview

A discriminative approach is taken where classifiers are used to directly estimate the posterior probability of lesion and normal appearing brain matter (NABM) voxels (Note the ground-truth is obtained from segmentation of the mTBI lesion by a domain expert). The discriminative approach performs well when there is a large amount of training data and it can be used for a complex decision space. A voxel level classifier has a large amount of data considering the 3D nature of MR images. The appearance of lesions in MRIs can be very complex, which leads to a complex decision space.

Here a cascade of classifiers is used for estimating the detected lesion at each time point (shown in Figure 16), where information from a previous time point is propagated forward. The first classifier in the cascade estimates the lesion using only the visual

features. Then context features are formed from the posterior probability map estimated by the classifier. These features are recalculated for each iteration in the process, for a given number of classifiers shown as a column of classifier for a single time point in Figure 16. Spatial information is propagated by the contextual features leading to improved classification. This process was introduced by [34] for static images. The contextual features used by [34] were point samplings of the posterior estimates. Their features demonstrated good segmentation performance on objects that were rigid in shape, when that shape was distorted their performance faltered. We propose two new features that generalize the static contextual information so it will work with amorphous objects.

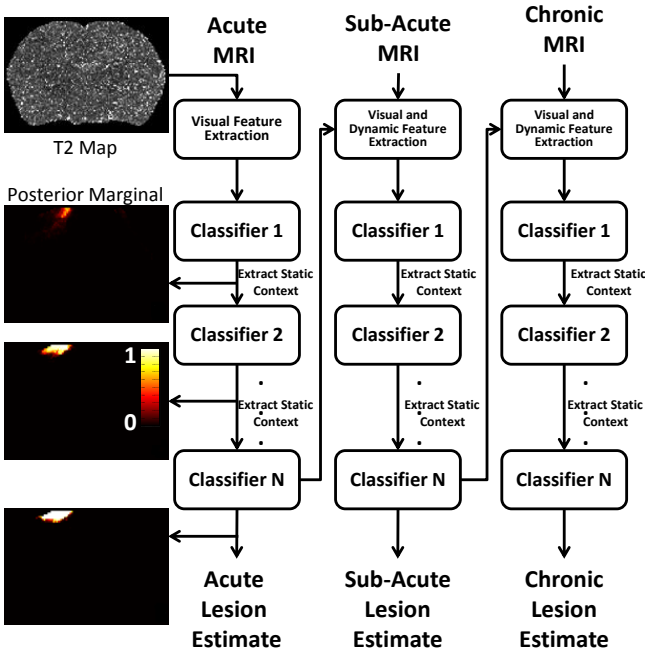

Figure 16: Overview of the proposed system. Context information is sent from one classifier to the next. After each classifier the static contextual features are extracted. The output of the N^{th} classifier is the final estimate of the lesion for the given time point. The dynamic contextual features are estimated from the final posterior marginal of the preceding time point.

Dynamic contextual features are calculated from the final classifier at a single time point. These features also in our approach are used by classifier at the subsequent time point, and they make use of spatial and lesion growth information. Note that [34] proposed using spatio-temporal volumes with their basic point features in the higher dimensional space. Their approach would require extreme amounts of data and the entire sequence to be known before segmentation. Our approach only considers pairs of brain volumes at a time, which allows for estimation at every time point.

2. Visual Features

Due to the low contrast nature of the unimodal MR images and the mild nature of the TBI, texture features are used to increase the discrimination in our approach. Four types of texture features are used: uniform local binary pattern (ULBP) local histograms [58] in the coronal plane (59 features), local statistics (mean, variance, skewness, kurtosis) of a Gabor filter bank [59] with 8 orientations and 4 scales in the coronal plane (128 features), basic histogram of oriented gradients in the coronal plane [60] (9 features), and local neighborhood statistical features (mean, variance, skewness, kurtosis, range, entropy, gradient magnitude xyz) (9 features). This gives a total of 205 visual features. This wide variety of features provides many different characteristics without being too specific (i.e., they will generalize well). The size of the local neighborhood used is 5x5 voxels [63], where the feature size is chosen to be the one that maximizes the Bhattacharyya distance between the two classes and gives a bound on the Bayesian error [54]. Examples of the visual features are shown in Figure 17.

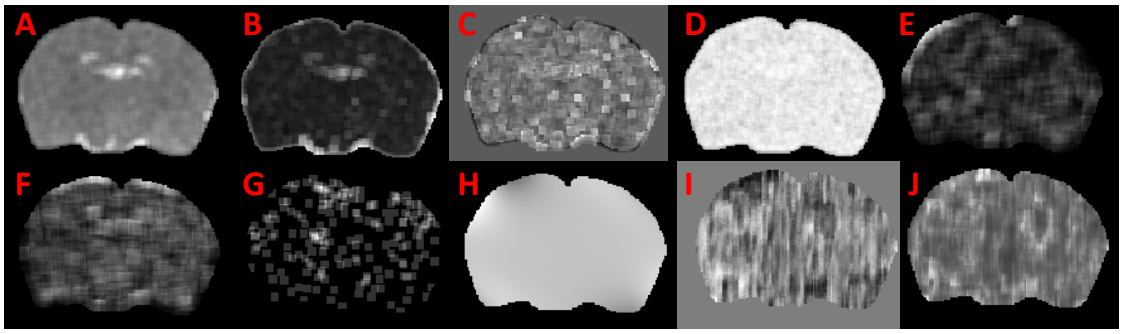

Figure 17: Examples of the visual features on coronal slices. The features from left to right, top to bottom: local mean (A), variance (B), skewness (C), kurtosis (D a bin from the histogram of oriented gradients (E-F), a bin from the LBP histogram (G), and example gabor features (H-J).

3. Static Contextual Features

The contextual features come from the posterior probability estimated by an already learned classifier. Previous approaches [45] have directly sampled a dense neighborhood around an observed voxel, making each location a potential feature. This method can lead to large feature sizes and can cause overfitting due to the specific locations that are learned. In this paper, two new static features are proposed to overcome this problem. One incorporates a sense of the surrounding without a known direction, while the other gives a general sense of direction.

The first feature, shown in Figure 18A, gives the average posterior probability at various distances around the observed voxel. This can be thought of as a proximity feature where: what is a close, medium and far away in distance are estimated. The distance function used here is the Manhattan distance allowing for a cuboidal region. These features are directionally invariant and can lead to better generalization since they describe a general area. By having a nesting of boxes the integral image can be utilized

for quick computation of the features. In 3D only 8 points are needed to find the average of a cuboidal region, using integral images [62]. Equation 17 provides these features, where f_xyz is the proximity feature, R_(1xyz),R_2xyz are square neighborhoods around the voxel at xyz and size(*) is the size of the bounding box.

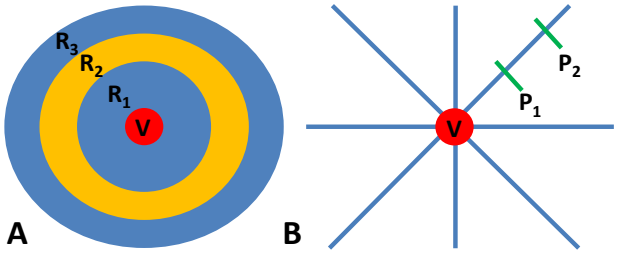

Figure 18: A) Illustration of the proximity feature. V is the observed voxel and the feature is the average probability of the regions (R_1, R_2, R_3). B) Illustration of the distance features. V is the observed voxel and an example feature is the average probability between P_1 and P_2 along the 45° ray.

$$f_{xyz} = \left(\sum R_{1xyz} - \sum R_{2xyz}\right)\left(\frac{1}{size(R_{1xyz}) - size(R_{2xyz})}\right) \quad (17)$$

Directional information is important for classification since the objects are rigidly registered to a naïve animal. The second contextual feature describes the posterior probability in various directions from the observed voxel. Rays are sampled at various distance ranges and angles from the observed voxel (see Figure 18B). From the distance ranges along the rays the mean is calculated. This gives a refined sense of the surrounding. An example would be what is close and above the observed voxel. The integral image is also used to calculate these features. Both features can be computed at coarse or fine distance bins without a significant increase in computational time. Equation (17) can also be used for the direction context features, the shape of the features

are changed to be that of Figure 18B. 3The features are rectangles with width one and the 45 degree features can be estimated using the rotated image.

4. Dynamic Contextual Feature

We propose the posterior marginal edge distance (PMED) feature is the distance a voxel is from the perimeter of objects of a class found by the maximum posterior marginal (MPM) estimate. To create this feature first the MPM at a voxel is obtained from the output of a classifier. This gives a binary image for each class. The distance transform is applied to the image and the inverse image and the feature is given by equation 18.

$$PMED = d(MPM) - d(\sim MPM) \tag{18}$$

$$MPM = \underset{c}{argmax}\, p(\omega = c|F) \tag{19}$$

Here d(*) is the Euclidean distance transform. This gives an image that is increasing as the voxels become farther away from the edge and smaller (more negative) as the voxels get further into the object. In eq. (19) ω is the estimated class, c is a specific class (lesion or normal brain in our case), and F is the feature set at a given voxel (see Figure 19).

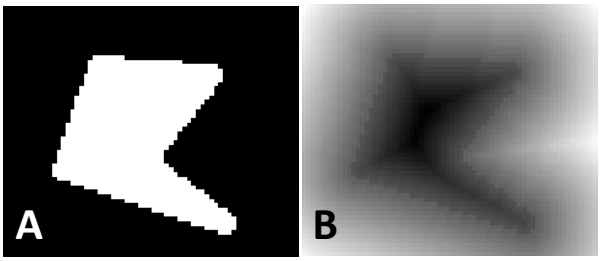

Figure 19: A) Example MPM estimate. B) Corresponding PMED feature. Note that the values become more negative towards the center of the object and more positive farther away from the object.

5. Classifier Cascade Selection

The base classifier in the cascade of classifiers has to be able to deal with a large number of features, a complex decision space and be able to give a posterior estimate. The primary classifier being used is adaboost [61] with small decision trees as base classifiers. Using small trees as a weak classifier (h(*)) allows for inherent feature selection, meaning erroneous features are disregarded. It is also not sensitive to feature normalization. In each iteration (t), the best classifier is selected and weighted with α. During the training process a cost matrix (4) is used where the top row represents the cost of normal appearing brain matter (NABM) and the bottom row the cost of Lesion voxels, such that the priors are offset to be even. This is done to account for the large disparity, in the number of training samples between the classes. Algorithm 1 describes the flow of adaboost.

ALGORITHM I
ADABOOST

1:	Input($\boldsymbol{x}$,$\boldsymbol{y}$) where $\boldsymbol{x}$ is the set of all training features and features and **y** is the class label.
2:	T = number of weak classifiers
3:	Initialize $\boldsymbol{D_1} = 1/m$ where $\boldsymbol{D}$ is the weight of each training example and m is the number of training examples.
4:	For t = 1 to T
5:	$h_t = argmax_{h_t \in H} \lvert 0.5 - \epsilon_t \rvert$ // Base Classifier
	where $\varepsilon_t = \sum D * \sum Cost\ Matrix * (y \neq h_t(x))$ //Error
6:	$\alpha_t = \frac{1}{2} ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right)$ // Classifier Weight
7:	$\boldsymbol{D}_{t+1} = \frac{\boldsymbol{D_t} * \exp(-\alpha_t y h_t(x))}{\boldsymbol{Z}}$ // Reweight the examples
	Where **Z** is a normalization term.
8:	Exit if $\lvert 0.5 - \epsilon_t \rvert < threshold$ //exit if error is small

$$Cost\ Matrix = \begin{bmatrix} 0 & 1 \\ \frac{\sum NABM\ voxels}{\sum Lesion\ voxels} & 0 \end{bmatrix} \quad (20)$$

It has been shown that the posterior marginal can be estimated using logistic regression [61] given by equation 21.

$$p(\omega_{xyz} = c | f_{xyz}) = \frac{e^{H_c(f_{xyz})}}{\sum_{c=1}^{C} e^{H_c(f_{xyz})}} \quad (21)$$

$$H_c(f_{xyz}) = \sum_{t=1}^{T} \alpha_t h_t(f_{xyz}) \quad (22)$$

Other classifiers such as the popular support vector machine (SVM) approach are also able to estimate the posterior marginal [56]. SVM is a strong classifier, but it does not have inherent feature selection and is sensitive to feature normalization. SVM tends to perform well with a small number of strong features. An example of the iterative training process for a single time point is shown in Figure 20. Notice that the estimate improves the greatest between classifier 1 (Figure 20B) and classifier 2 (Figure 20C). During training each successive classifier will have a performance greater than or equal to its predecessor, since the posterior marginal is one of the features.

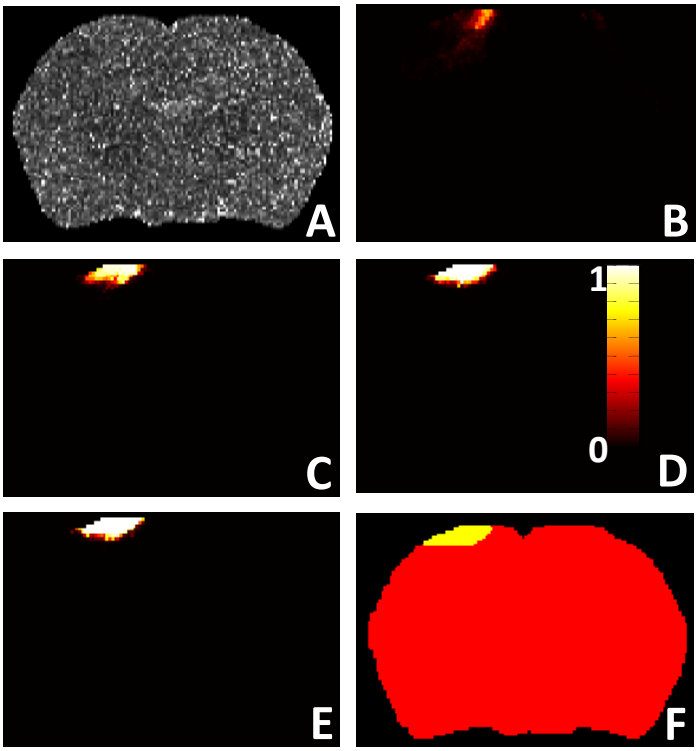

Figure 20: Example probability maps after each classifier. A) T2 map. B-E) Classifier output probability map for the training of classifier 1-4 respectively. (F) Manual Segmentation where yellow denotes the lesion. This shows the convergence of the algorithm to the manually segmented injury.

C. Experimental Results

1. Contribution of Features in a Temporal mTBI Dataset

Our temporal data set uses Sprague Dawley rats subjected to a single impact controlled cortical impact (CCI) as an animal model of mTBI [45][6]. The CCI model results in a focal injury. Briefly, a craniectomy (5mm) was performed over the right hemisphere (3mm lateral, 3mm posterior to bregma) where a mild CCI was induced using an electromagnetically driven piston (4mm diameter, 0.5mm depth, 6.0 m/s). The incision made for the craniectomy is sutured after the injury is induced.

The animals were imaged in vivo at 3 time points post injury: acute (1st day), sub-acute (8th day), and chronic (14th day). There are a total of 6 sequences, each with 3 time points. MRI data were acquired using a Bruker Advance 4.7T for T2 weighted images (T2WI; TR/TE/FA=3453 ms/20 ms/20°, 25x1 mm slices) with a 256x256 matrix and 3cm field of view. Each voxel has a dimension of 0.12x0.12x1mm. The T2 weighted images were converted to T2 maps. ROIs were manually segmented using Cheshire image processing software (Hayden Image/Processing Group, Waltham, MA) and included the right and left hemispheres and injured tissue volumes that were defined as abnormal (hyper/hypo-intense) signal intensities within the cortex with the remaining tissues designated as NABM. The injuries have been verified with histology in [6] at the end points of experiments. A direct comparison to the results of histology is not possible due to the deformations that occur in the histological staining process.

We examine the effect of the proposed features and effect of the dynamic information. For the training/testing split leave-one-out validation is used where a whole sequence is left out (resulting in 6 cross-folds). Three temporally consecutive volumes were left for testing and the rest were used for training. The parameters used were: 300 weak learners, learning rate 1, and 4 cascaded classifiers. Three approaches were tested: the original autocontext features [34], the proposed approach with only the proposed static features, and the proposed approach with the static and dynamic features.

To evaluate the segmentation results of the proposed approach the dice coefficient is used, given by equation (23),

$$Dice = \frac{2TP}{(TP + FP) + (TP + FN)} \tag{23}$$

where TP is true positive, FP is false positive, and FN is false negative. The dice coefficient gives the ratio of the intersection between the detected object and the target object to the size of the objects. The Dice gives a better idea of the intersection between the two objects because it does not consider the true negatives, which will be large with the unbalanced class size. The output of the classifier at each time point (acute, sub-acute, and chronic) shown in Figure 16, is thresholded to get a hard classification.

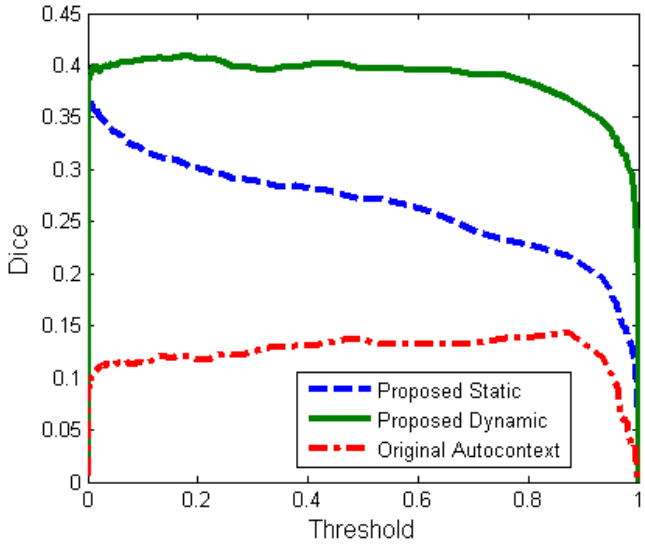

Figure 21: Dice coefficient after thresholding the posterior probability map at the end of each cascade of classifier (i.e. at each time point). This is the average of all the tests in the leave one out validation.

From Figure 21 it is clear that the proposed dynamic approach outperforms the other methods. This shows it is important to use the dynamic information. The original autocontext tends to over fit due to the specific locations the features represent. The same features locations proposed in [34] were used for the original autocontext testing. During the training phase it obtains a dice score above 0.9, but it does not generalize well to the testing data. From the testing, our proposed static features give a good generalization

compared to the original autocontext. The output from the original autocontext approach is patchy with many false negatives leading to an under-estimation of the injury. This is due to the fact that it looks at specific locations for the context features.

The proposed approach has a very flat dice curve, so it is not sensitive to a chosen threshold on the output probability map. This makes the selection of a threshold less critical. From the qualitative results (Figure 22) it is clear that the results of proposed approach work on small to medium size lesions. The qualitative results also show that false positives are only close to the majority of the lesion mass without having erroneous spatial outliers. It is also apparent that the approach has some difficulties at the edges of the lesions. This could be rectified by using shape constraint on the thresholded probability map.

In [38] a similar dataset was used for analysis except repeated contralateral injuries were also considered. The average dice score using a high level contextual model was 0.35 and 0.07 using probabilistic SVM alone. This is similar to our dice score using the proposed static features with an average dice score of 0.36 and it far exceeds the results of a probabilistic SVM. The dynamic features increase the dice up to 0.41. The method proposed here does not use any extra patient information as it has been done in [38].

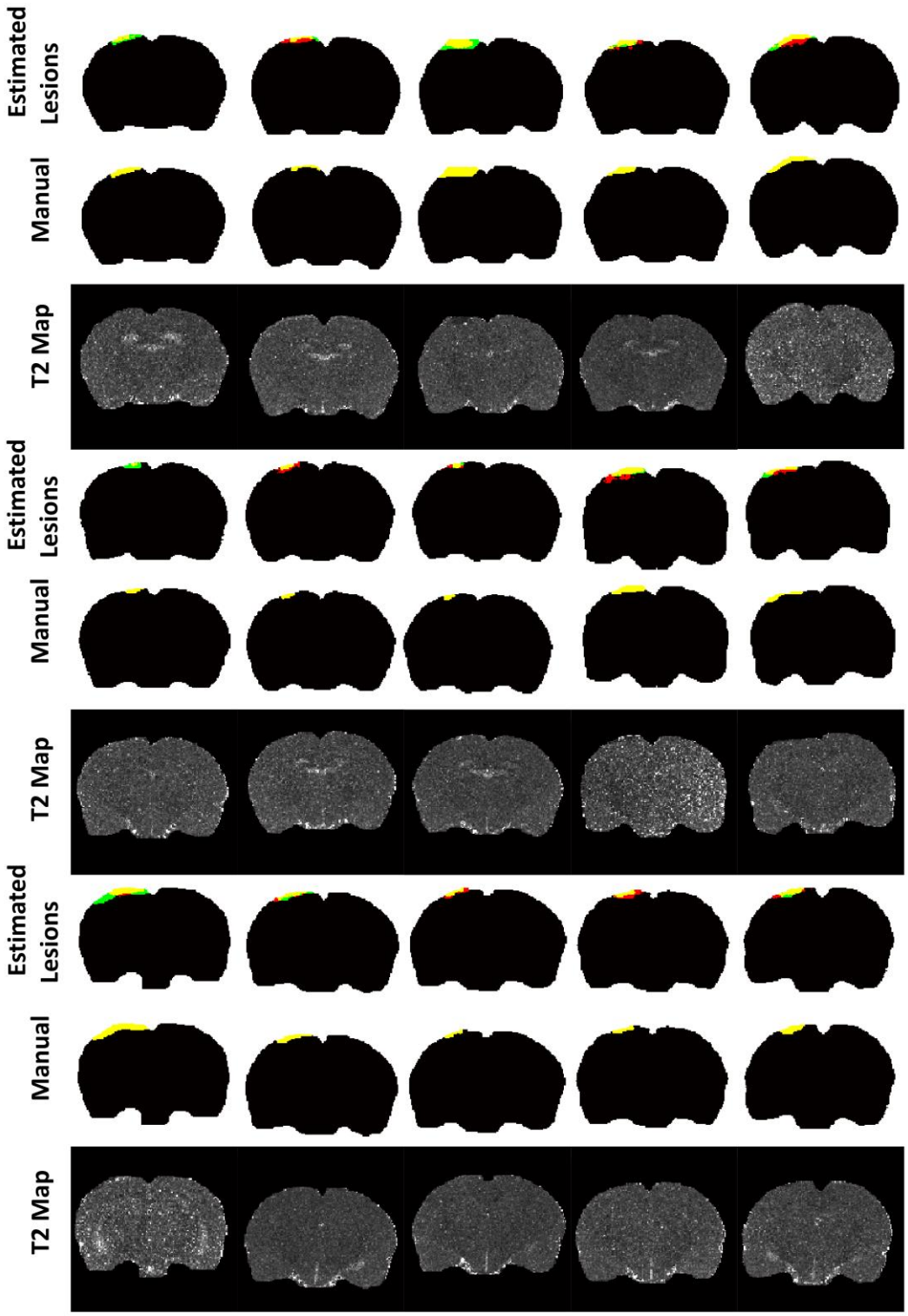

Figure 22: Qualitative results of the proposed approach using dynamic and static contextual features. Each coronal slice is from a separate volume. Color code: yellow = true positive, black = true negative, green = false negative, red = false positive.

2. Convergence Rate of the Features

We evaluate the strength of the individual features. The same dataset from section IV.A is used in this section. To determine the strength of the features the rate of convergence of the cost function during training is observed.

In Figure 23 the training cost after each tree during the training of adaboost is given. Both the autocontext and the proposed static features have the same results from the initial classifier utilizing only visual features. This is due to the proposed static features being able to generalize in the surrounding area better.

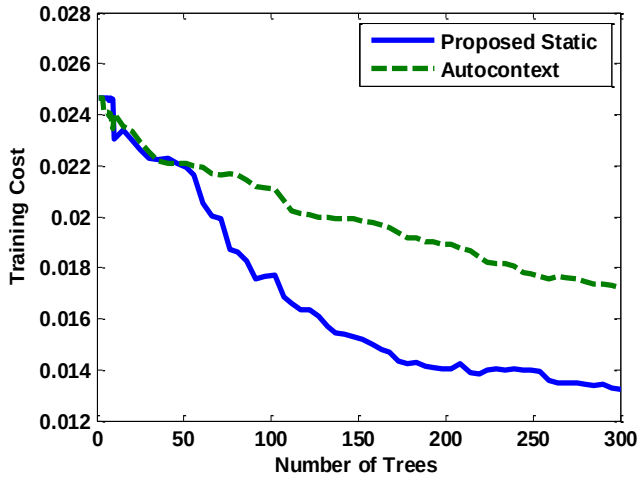

Figure 23: Training cost after each tree for the proposed static features versus the original autocontext. This is the training cost of the second classifier i.e. the input to both methods is the results after using visual features only. Lower cost is better.

To understand the convergence of each classifier in the cascade one must observe the output training cost after adding each tree during the training of adaboost. In Figure 24 the convergence of the training using only the static features is observed. Convergence using only the visual features takes a while compared to with the static contextual

features. After the second classifier, the improvement over the number of trees is less pronounced. Not much information is gained after the second classifier.

Also in Figure 24 we can observe the convergence of the classifier using the dynamic and static context features. It is apparent that the convergence is very fast for Classifier 1 when the dynamic context features are used, compared to the case when no dynamic information is used in the first iteration. Information is still gained by the static features but at a slower rate.

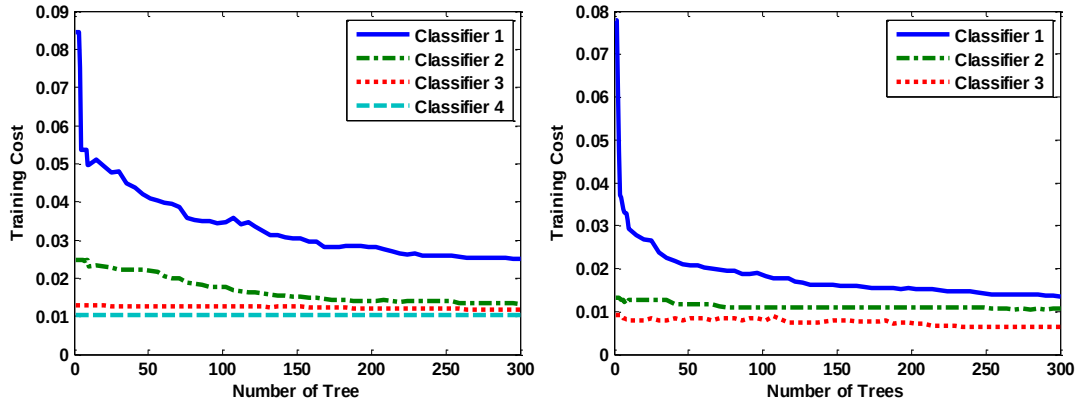

Figure 24: Left: Convergence of classifier training using only the static contextual features. Right: Convergence of classifier training using both dynamic and static contextual features. Lower cost is better.

3. Feature Importance Evaluation

An important question to ask when using adaboost is, "which features are being chosen and when?" Here those questions are answered by looking at the percent of features chosen throughout the adaboost training process. In the case of the static features alone, Figure 25 shows first 50 trees are almost exclusively context features. This shows that the contextual features hold the most information for discrimination. The visual

features start being used after the information from the contextual features is nearly all used. In the end 40-50% of all the features are contextual.

A similar pattern occurs when the dynamic features are used (shown in Figure 25). The contextual features in Classifier 1 are used early on in the learning process, and then the visual features almost exclusively take over. In Classifier 3 nearly 60% of the features are chosen as contextual features. This shows that the dynamic information plays a vital role in the classification.

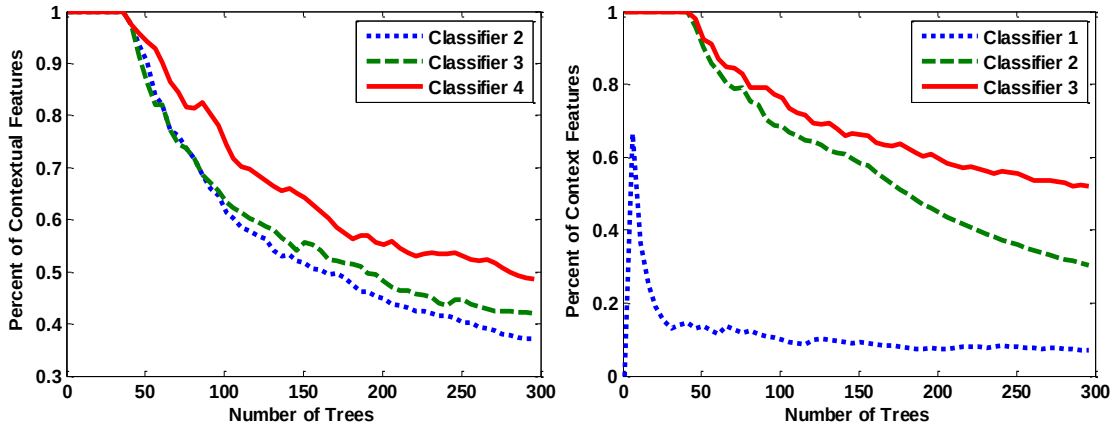

Figure 25: Top: The percentage of the static contextual features chosen versus the current number of trees. Bottom: The percentage of dynamic and static contextual features chosen up to a number of trees. Note in Classifier 1 only the dynamic contextual features are used.

4. Static Feature Performance on the BRATS Dataset

To show the versatility of the proposed approach the static features were tested on a brain tumor dataset. The dataset is from the MICCAI 2012 Multimodal brain tumor segmentation (BraTS) challenge. Here a subset of the data was used that included 15 High grade glioma patients. The volumes included the following MR modalities, T1, post-Gadolinium T1, T2, and FLAIR. Each modality has been co-registered, skull-

stripped and resampled to be 1mm isometric voxels. A mode shift was used to normalize the histogram across subjects. For this testing the edema and core tumor are considered to be a single class. The visual features used were the individual modalities, the same local neighborhood statistics as the mTBI dataset on the FLAIR channel, and LBP on the FLAIR channel. Since there is a single time point database only the static contextual features are used. Testing was done using 5 fold-cross validation.

For this testing an average dice score of 0.71 was obtained. The best scores reported during the BraTS challenge were between 0.7 - 0.8, for the high grade glioma cases. Note that the other approaches [63][64] used symmetry features. These features can be directly used by the proposed approach. This is another type of context that can be exploited. Another feature used in the challenge were long range context features [65]. These features sample the value around an observed voxel. The parameter space for these features is extremely large and requires random sampling of the space via decision forests [66]. The other features used by the winners of the challenge [65] were based on generative modeling of the tissue types. These features can also directly be used by the approach proposed here.

D. Conclusions

A fully automated method for detecting mTBI lesions that integrates low-level static and dynamic context is proposed. Three novel features are proposed that describe the posterior probability of classifier outputs in a cascade. These features were shown to have good qualitative and quantitative results on an mTBI dataset. The proposed approach outperformed the original autocontext [34], by being able to generalize and

integrate dynamic information. The context features were shown to be the most important features in the training process. Adding the dynamic context features lead to fast convergence during training leading a smaller number of trees being needed. The generality and flexibility of the proposed approach allows it to be applied to other brain lesion problems.

CHAPTER 5 MULTI-MODAL FEATURE

A. Contributions

We propose a discriminative classification approach where features at every voxel are used to estimate its probability of being a lesion. The key contributions of this paper include: 1) development of temporally sensitive generative tissue model for estimating blood, normal appearing brain matter (NABM), edema, and noise utilizing the quantitative T2maps, and an estimate of the probability of each tissue type. 2) The use of new features based on the probability estimate of micro hemorrhage in SWI, and the use of texture features on T2 MRI echoes. 3) Also, we present experimental evaluations and comparison of results.

B. Technical Rationale

To estimate the probability of lesions a discriminative approach was taken where the posterior probability is directly estimated using classifiers (Figure 26). We utilized manual segmentation as the ground-truth for training our classifiers, since it is the current gold standard. The chosen classifier was adaboost [61], which is an ensemble classifier that allows for a direct estimate of the posterior probability based on logistic regression. Features specific to each imaging modality were calculated to take advantage of the unique properties (Figure 26) of each modality. Entropy based multimodal registration was utilized to normalize the spatial locations of the brain among the modalities [67]. The T2 echoes have good anatomical definition but do not have physical meaning, therefore, this channel is used for calculation of texture. The T2 maps have physical meaning but

are inherently noisy, so a temporal generative tissue model was used in this channel. For SWI we take advantage of its sensitivity to blood by thresholding the image and relaxing the result to create a new feature.

Adaboost was chosen as the discriminative classifier due to its inherent feature selection and efficiency (discussed in Section 2.4). The features from each modality were fused at the feature level [68], which has been shown to do well in systems with adequate feature selection.

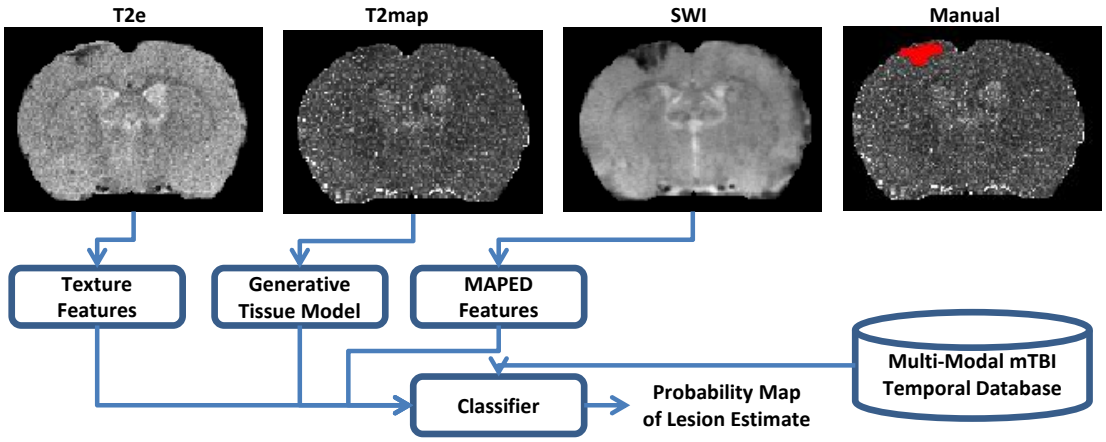

Figure 26: Overview of the online lesion estimation system. The output is a probability map where values close to one signify a high likelihood of lesion and values close to 0 signify NABM. The top row of images shows the multimodal input from a single subject. Each modality has its specific features extracted to maximize their discriminatory potential based on their specific characteristics. The manual image is an example of an expert segmentation.

1. Using Texture in T2 Echoes

MRI acquisition with multiple T2e images, when mathematically combined, is widely used clinically to obtain quantitative relaxation values of tissues in the brain [18]. Individual T2e images provide enhanced contrast for the estimation of many diseases with lesions, including tumors [20c][65] and multiple sclerosis (MS) [23]. The high

resolution and low noise in the T2e channel allows for texture to be used [25]. Texture has been shown to greatly improve the discriminatory value of T2e MRIs for detecting lesions caused by mTBI [25][38][46]. This is due to texture being able to estimate the underlying tissue characteristics of a region.

To describe the texture in local regions two feature types are used, 1) local neighborhood statistical features and 2) local statistics of a Gabor filter bank. The local neighborhood statistics include: mean, variance, skewness, kurtosis, range, and gradient magnitude xyz (8 features). The neighborhoods are only considered in the coronal plane due to the anisotropic nature of the voxels. For the Gabor filter bank [59] 8 orientations and 4 scales in the coronal plane are used. The Gabor filter bank is a group of bandpass filters at various frequencies and orientation. The filter bank was calculated such that the frequency space is maximally described [69]. To compute more informative features of the filtered images, features based on local statistics (mean, variance, skewness, kurtosis) are used. The size of the local neighborhoods used is 5x5 voxels, where the maximum Bhattacharyya distance between the two classes is used to choose the voxel size. The Bhattacharyya distance has been shown to give a bound of the Bayesian error [54]. Examples of these texture features are shown in Figure 27.

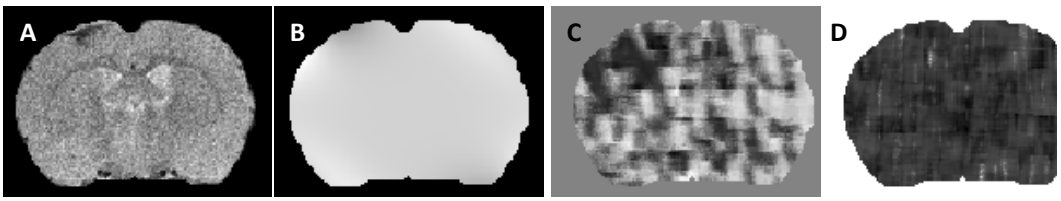

Figure 27: An example of texture features in the T2e modality. A) An original T2e coronal slice. B-D) examples of Gabor features from lower to higher frequency relative to the size of the image.

2. Temporal Generative Tissue Model Using T2 Maps

Manual analysis of T2 maps (T2map) has been used in the analysis of the underlying composition of lesions caused by mTBI [45]. T2maps give an estimate of the T2 relaxation of the brain tissues being imaged and is a value with physical meaning. The composition of a mTBI lesion is known to have four types: blood, NABM, edema and noise. Over time the injury progression results in altered lesion composition [45] which needs to be taken into account. In manual analysis the thresholds were based on T2 values derived by expert analysis.

We propose a temporal tissue model using a generative Gaussian mixture model (GMM) at the major time points: acute (1st day), sub-acute (3-8 days), and chronic (14-64 days). At each of these time points a GMM is estimated. GMM has been used for classification of the underlying brain tissue types (gray matter, white matter, and cerebral spinal fluid) [70][71], but it has not been previously used for estimating the tissues types we are investigating. An example of the mixture model is shown in Figure 28.

To estimate the mixture model, we used the manually found lesions to obtain an estimation of the mixture distribution. Then expectation maximization (EM) is used to fit the parameters of the distribution [72]. To initialize the parameter of the distribution, the thresholds found in [45] are used. This ensures that the estimate will not converge to local minima that are not close to the true distribution.

Each tissue distribution is evaluated to give a feature. The output of each distribution is smoothed using a Gaussian filter (5x5) in the coronal plane. Using a smoothing filter adds spatial information to the estimation and is done commonly when

estimating tissues [72]. An example of these tissue estimates are shown in the graph in Figure 28.

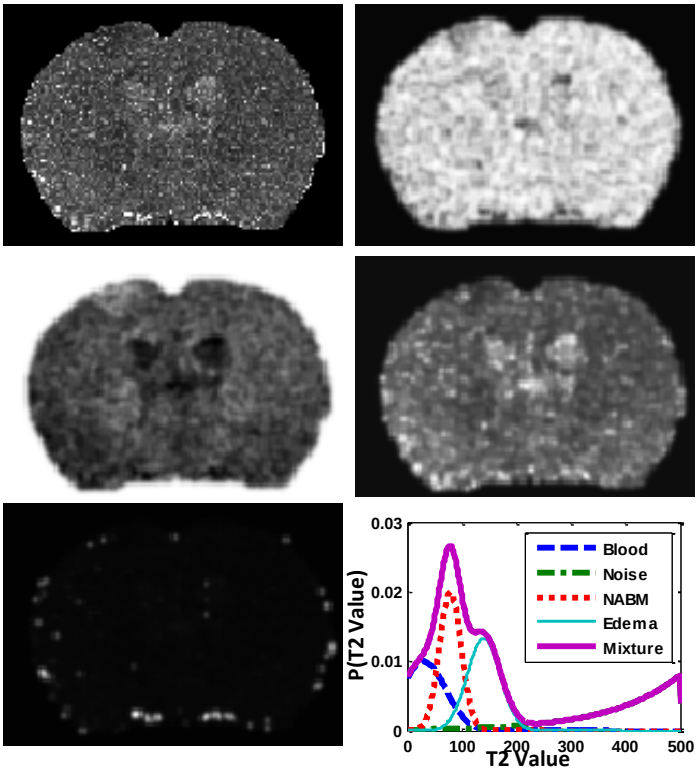

Figure 28: Example of the temporal generative tissue model features generated from the T2map. From top to bottom, left to right: Original T2map coronal slice, NABM estimate, Blood estimate, Edema estimate, Noise estimate. Edema estimate, Noise estimate, example of the generative tissue model at a single time point. Note brighter value of a pixel in an image means a higher probability.

To give further information from T2 maps a hard estimate of the tissue type is given using the threshold found in [45]. After the hard classification the connected components are found allowing for the size of the estimated tissue type to be used as a feature. When the connected components of the tissue is found to be small (one voxel) there is less probability of it being the tissue type it is initially estimated to be. Using

component size as a feature allows the classifier to automatically choose the component size that is important.

3. SWI Lesion Estimate

SWI uses the loss of signal intensity caused by the disturbance in a homogeneous magnetic field. The magnetic susceptibility is higher at the interface of two regions [15]. Extravascular blood is enhanced due to the local dephasing caused by its iron content. SWI takes advantage of both the magnitude and phase images to estimate the susceptibility. A phase mask is created that enhances the phase differences between paramagnetic materials and surrounding tissue [15]. These biophysical features can be exploited for imaging micro hemorrhages that occur following mTBI. The hemorrhages appear as dark spots on the SWI images (Figure 29). Other locations around the exterior surface of the brain may show false positives due to phase distortions at interface between air, bone and brain. Some approaches have been proposed for finding localized small depositions of blood in SWI images [17]. The hemorrhagic lesions shown in SWI tend to be overestimates as the region of magnetic homogeneity disruption tends to be larger than the actual lesions found on T2e and verified by histology [44]. Our proposed feature takes this into account.

The SWI image is first thresholded at multiple intensity threshold levels, since a single threshold may not contain all the spatial information needed. This gives an initial hard estimate of the lesion (Figure 29). To add spatial information the distance transform is utilized. Eq. (24) shows our feature. It is the distance transform applied to the lesion estimate and the inverse lesion estimate.

$$Rf = d(\bar{L}) - d(\sim\bar{L}) \quad (24)$$

$$where: L = I_{SWI} > th_{SWI} \quad (25)$$

I_SWI is the SWI image and 〖th〗_SWI is the threshold on that image. The threshold can be set by an expert or multiple thresholds can be given. The inherent feature selection of the classifier can select the best threshold automatically as is done in our testing. L is the thresholded SWI image, which can be thought of as the maximum posterior marginal using the thresholded image as a rough estimate of the lesion. d(*) is the Euclidean distance transform, which gives an image where voxels increase farther away from the estimated object. Rf is the maximum a posterior edge distance (MAPED) feature since the lesion estimate is being relaxed, where voxel going away from the edge of the estimate are increasing (more positive) and into the estimated lesion are decreasing (more negative).

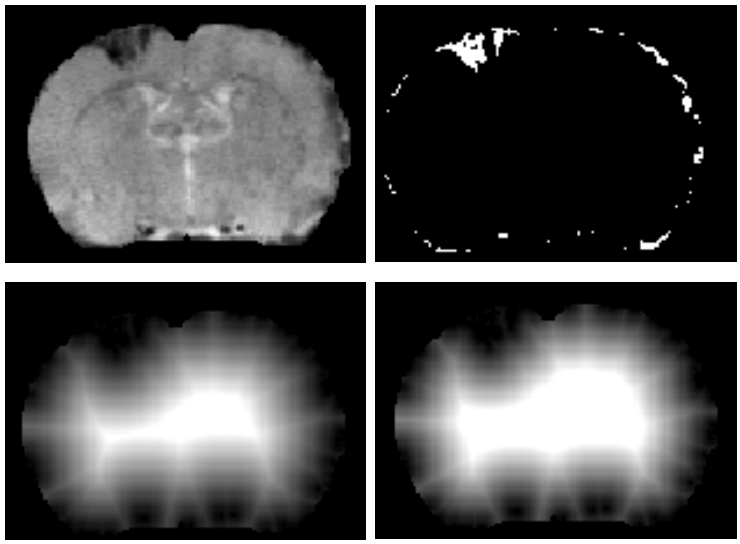

Figure 29: Examples of the SWI MAPED features. From top to bottom, left to right: Original SWI coronal slice, Thresholded SWI, 2D MAPED feature, 3D MAPED feature.

This is similar to the posterior marginal edge distance proposed in [46]. Except, in our application the overestimation of the lesion is being modeled but the growth of the lesion is not being estimated. As previously noted, our data is anisotropic, so the feature is estimated in the coronal plane. However, we have seen that for this particular feature 3D information is also important. Therefore, the 3D distance is also included as a feature. An example of these features is shown in Figure 29. This channel gives a total of (1+3*number of thresholds) features.

4. Classifier Selection

The voxel classifier is used to directly estimate the posterior probability of the lesion and NABM. There are a multitude of classifiers that try to accomplish the same goal: support vector machines (SVM) [73], decision forests (DF) [74], extreme learning machine (ELM) [75], K-th Nearest Neighbor (KNN), Neural Networks (NN), adaboost [61], and many more. Each of these classifiers has its merits and weaknesses, so we must first determine the problem attributes. The attributes of our problem include: a very large number of training examples, a complex decision space, non-normalized features from different domains, a probabilistic output is desired, a two class problem (Lesion vs. NABM) and we have a large disparity between the number of examples in each class. The classifier that satisfies all these properties is adaboost with small decision trees as its weak classifier. When using small decision trees as the weak classifier (h(*)) inherent feature selection, meaning superfluous features will be discarded. Using small decision trees also obviates the need for feature normalization. Feature normalization can become problematic in classifier like SVM.

For training, the best weak classifier (h(*)) is chosen at each iteration (t) and weighted by α. The training data is reweighted, assigning higher weights to misclassified samples. The error is evaluated using Cost Matrix in eq. (26) where the top row represents the cost of NABM and the bottom row the cost of the lesion voxels. A cost adjustment must be used due to the large imbalance between the number of training samples between the classes. The final classifier H(*) is an ensemble of the of the weak classifiers. Pseudo code for this algorithm is given by Algorithm II.

ALGORITHM II
ADABOOST

1:	Input(***x***,***y***) where ***x*** is the set of all training features and features and **y** is the class label.
2:	T = number of weak classifiers
3:	Initialize $\boldsymbol{D_1} = 1/m$ where ***D*** is the weight of each training example and m is the number of training examples.
4:	For t = 1 to T
5:	$h_t = argmax_{h_t \in H} \lvert 0.5 - \epsilon_t \rvert$ // Base Classifier
	where $\varepsilon_t = \sum D * \sum Cost\ Matrix * (y \neq h_t(x))$ //Error
6:	$\alpha_t = \frac{1}{2} ln\left(\frac{1-\varepsilon_t}{\varepsilon_t}\right)$ // Classifier Weight
7:	$\boldsymbol{D}_{t+1} = \frac{\boldsymbol{D}_t * \exp(-\alpha_t \boldsymbol{y} h_t(\boldsymbol{x}))}{\boldsymbol{Z}}$ // Reweight the examples
	where **Z** is a normalization term.
8:	Exit if $\lvert 0.5 - \epsilon_t \rvert < threshold$ //exit if error is small

$$Cost\ Matrix = \begin{bmatrix} 0 & 1 \\ \frac{\sum NABM\ voxels}{\sum Lesion\ voxels} & 0 \end{bmatrix} \quad (26)$$

We stipulated that a probabilistic output is desired, since it allows for post processing methods [38]. Adaboost does not inherently give a probabilistic output, but [61] showed that adaboost can give a posterior probability estimate by treating it as a logistic regression, given by eq. (27).

$$p(\omega_{xyz} = c | f_{xyz}) = \frac{e^{H_c(f_{xyz})}}{\sum_{c=1}^{C} e^{H_c(f_{xyz})}} \quad (27)$$

$$H_c(f_{xyz}) = \sum_{t=1}^{T} \alpha_t h_t(f_{xyz}) \quad (28)$$

where T is the maximum number of trees and f_{xyz} is the voxels feature vector.

C. Experimental Results

1. Dataset and Performance Measures

Our data set uses a Sprague Dawley rat as a model subjected to either a single or repeated bilateral mild controlled cortical impact (mCCI) [45][6]. In the CCI model a crainectomy (5mm) is performed over the right hemisphere (3mm lateral, 3mm posterior to bregma) and a mCCI is induced using an electromagnetically driven piston (4mm diameter, 0.5mm depth, 6.0 m/s). The incision made for the craniectomy is then sutured. For the repeated contra-lateral injury a second mCCI is induced at either 3 or 7 days post first injury to the identical location but on the contralateral cortex [45][6].

After the mTBI animals underwent multimodal MRI at at various time points between 1 and 64 days post injury. The imaging time points were classified into three groups: acute (1^{st} day), sub-acute (3-8 days), and chronic (14-64 days). These temporal groups were delineated based on the previously defined major epochs in the injury evolution. A total of 59 imaging time points were acquired: 15 3day-contra-lateral, 11 7day-contra-lateral, and 33 single volumes. The imaging time point distribution for our dataset were as follows: 19 acute, 21 sub-acute and 19 chronic. MRI data were acquired using a Bruker Advance 4.7T for T2 weighted images (T2WI; TR/TE/FA=3453 ms/20 ms/20°, 25x1 mm slices) and SWI (TR/TE/FA=39 ms/20 ms/20°, 48×0.8 mm slices) with a 256x256 matrix and 3cm field-of-view. Each voxel has a dimension of 0.12x0.12x1mm. The T2 weighted images were processed into T2 maps. ROIs were manually segmented on T2WI using Cheshire image processing software (Hayden Image/Processing Group, Waltham, MA) and included the right and left hemispheres and

injured tissue volumes that were defined as abnormal (hyper/hypo-intense) signal intensities within the cortex with the remaining tissues designated as NABM. These manual data provided the ground-truth for our computational approach. Note all images were normalized in intensity using a mode-shift normalization method.

The segmentation was evaluated using the Dice coefficient given by equation (29).

$$Dice = \frac{2TP}{(TP + FP) + (TP + FN)} \tag{29}$$

where TP is true positive, FP is false positive, FN is false negative. This metric describes the ratio of the intersection, between the true object and the estimated object, and the overall size of the two objects. Note that Dice does not consider the true negative, which can be very large in evaluation of voxel level segmentation.

We also use the *receiver operating characteristic* (ROC) curve to describe where the errors are occurring. It describes the results using the true positive rate (TPR) (Sensitivity) versus the false positive rate (FPR) (1-Specificity). Each point on the ROC represents a different threshold on the posterior probability estimated by the classifier. ROC is avidly used in radiology and other medical disciplines to determine the performance of predictive procedures [57]. The performance is considered to be the best when the ROC is a step function and is considered random when it is a line on the diagonal.

2. Feature Performance

In these experiments the performance of each feature is measured as well as the fusion of all the features. To show the performance of the features a 5-fold cross

validation is performed where one fifth of the data is held for testing and the other four fifths is used for training the classifier. The feature set used for testing are as follows: 1) *T2e texture* only where only the texture features are used on the T2e channel are used. 2) *T2map Generative* where only the temporal generative model on the T2map channel is used. 3) *SWI features* where only the SWI features are used. 4) *Intensity only* where only the intensities from SWI, T2e, and T2map are used. 5) *Proposed* where all the proposed features are fused. For the classifier 300 weak learners are used.

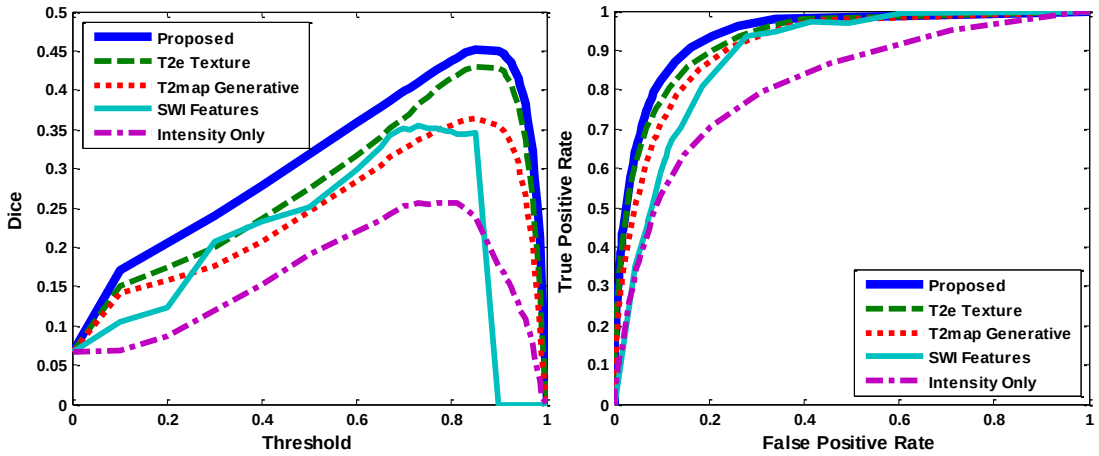

Figure 30: A: This plot shows the Dice score at different thresholds on the posterior probability map output by the classifier. We see the proposed method, which is the fusion of all the features from their modalities, and gives the best performance. B: The ROC plot, where each point on the plot represents a threshold on the posterior probability map output by the classifier.

From Figure 30 it is apparent that the *proposed* approach with all the features outperforms all the individual methods. The closest result, to the proposed, is using the *T2e Texture* which has been used in [46] as a uni-modal approach. An approximately ten percent increase over the other experiments is observed with the addition of the other channels and features. Adding the other channels and features contributes to the increase in performance that would not be possible in a uni-modal approach. The performance of

T2e Texture further solidifies the need for texture in mTBI analysis. The third best outcome in testing is the *T2map Generative* which performs below the *Proposed*, but it performs better than *Intensity Only* revealing that the information gain from theT2map features are significant. Note the Dice graph is skewed toward the higher thresholds. This shows that the classifier is overestimating the lesion voxels, suggesting that the posterior probability needs to be restricted to improve the results.

From the ROC plot, given by Figure 30, we see similar trends in performance as that given by the Dice plot, but there are some differences. The *SWI features* show better performance on the ROC plot by having a more convex shape than *Intensity Only* and *T2map Generative*. This demonstrates that the SWI features are better at distinguishing false alarms. When looking at the SWI images it becomes apparent why this is the case. The *proposed* approach still gives the best performance. The increase in performance of the *proposed* over *T2e Texture* is cluttered in the ROC due to the size of the true negatives, so the Dice plot gives a better view of the performance increase. Note on the qualitative results given by Figure 31.

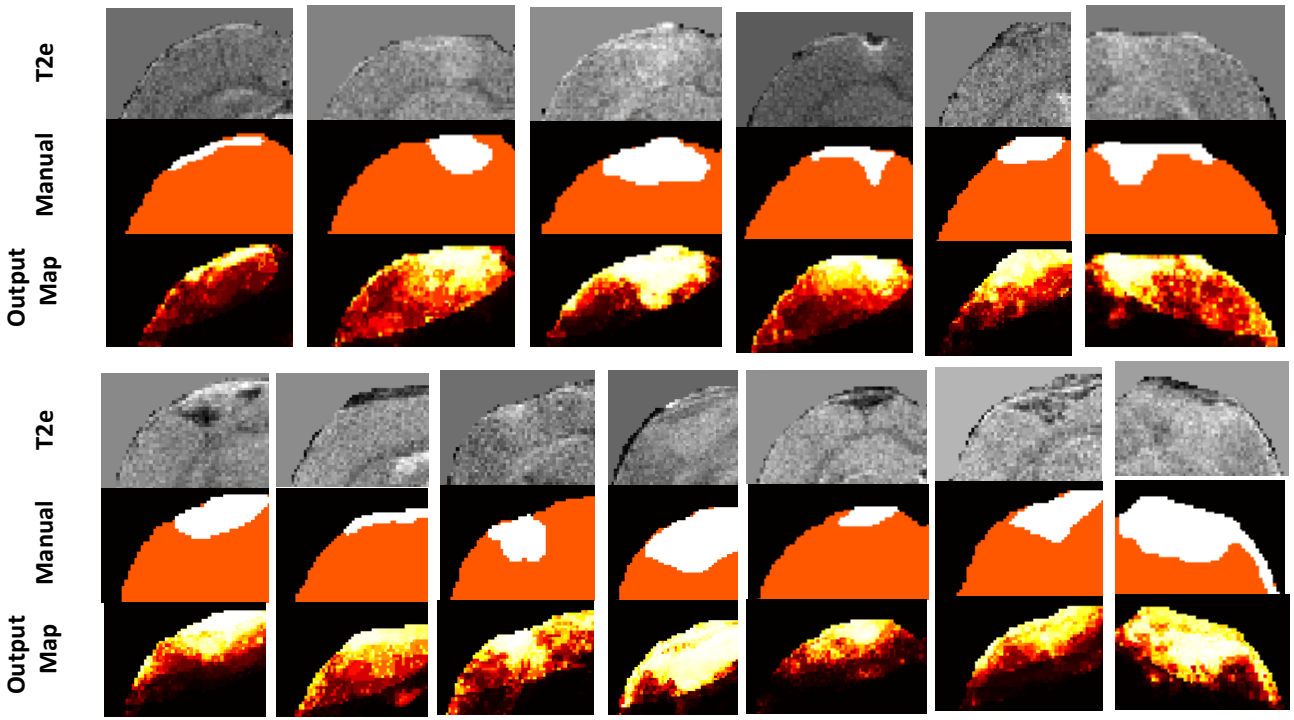

Figure 31: Qualitative outputs from the proposed approach are given. Note that these results are from various time points and injury types, but the output is obtained using the same classifier. T2e is the normalized visualization of the T2e channel. Manual is the ground-truth segmentation by an expert. Output map is the probabilistic output of the proposed approach. Note to get a hard classification the map is thresholded.

Our *proposed* approach is compared to other state-of-the-art algorithms in Table 1. The compared approaches are described previously in [57] and [39]. The Kth-nearest neighbor (KNN) approach proposed by Anbeek et al. [57] uses a database of volumes (our training set) to query whether each data point is lesion or NABM. The volumes are registered and the position of each voxel in the registered space is saved as a feature. For the other features the basic texture features on the T2e are used (not including the Gabor features) as well as the multi-modal intensities. In the probabilistic support vector machine (SVM) approach the same features are used. For the SVM the radial basis function is used as the kernel and the parameters are optimized using cross-validation. Note that both competing approaches do not possess inherent feature selection like our chosen classifier (adaboost). This can cause confusion within the classifier hence hindering the results.

TABLE 3
COMPARISON OF ALGORITHMS

Method	**Dice(std.)**	**TPR(std.)**	**FPR(std.)**
Proposed	**0.46(0.02)**	**0.88(0.05)**	0.12(0.002)
PSVM [36c]	0.08(0.1)	0.77(0.2)	0.31(0.1)
KNN[35c]	0.18(0.01)	0.1(0.02)	0.01(0.03)
T2e Texture	0.43(0.02)	0.85(0.06)	0.16(0.001)
Intensity Only	0.26(0.01)	0.71(0.03)	0.28(0.011)
Inter-Rater*	0.58(0.15)	0.92(0.07)	0.15(0.013)

The dice score is the highest dice achieved after thresholding the posterior probability map. The TPR and FPR are taken at the equal error rate (excluding KNN and Inter-Rater). *Note the inter-rater testing was carried out on a subset of the data: 3 single injury, 3 3day contra-lateral, and 3 7day contra-lateral.

To validate acceptable Dice scores, the inter-rater variability was estimated using a subset of the data. For this test a subset of the data included: 3 single injury, 3 3day contra-lateral, and 3 7day contra-lateral. A trained operator gave the estimates manually. From this we can see the difficulty in achieving a high Dice score on this type of data

particularly when the regions are small and are of low contrast. The Dice coefficient becomes more sensitive as the regions are reduced in size. Given that the regions are less than one percent of the total brain volume and encompass several hundred voxels at most, it is clear why the Dice score can be affected so dramatically.

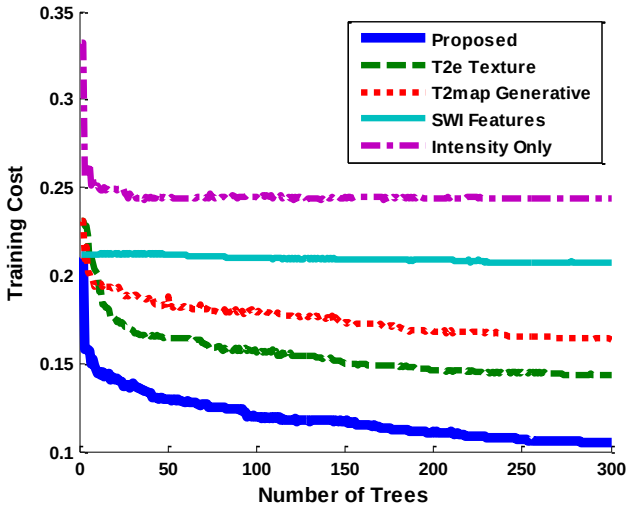

Figure 32: Training cost during training of the classifier (lower is better). This shows the convergence of the adaboost classifier, where a smaller number of trees means a simpler classifier and a lower number of selected features. The proposed method, which is the fusions of all the features, is shown to give the lowest training cost.

3. Feature Performance in the Classifier

To observe the importance of the features and how they are selected by the classifier, we look at the training cost versus the number of learned trees (Figure 32). The graph shows the average of the five cross folds from the test run in section 3.2. From this graph the training cost of the classifier shows the effect adding more trees to the ensemble. There are two reasons the number of trees can have an effect on the cost. 1) The decision space is complex in a single dimension requiring more trees to describe the space. 2) Different features are chosen by the new trees. Figure 32 shows the *proposed*

test gives the lowest cost but requires a larger number of trees to level off compared to the other tests, showing that information is gained by fusing all the features. This suggests that the features are working synergistically with each other to give the best performance. The *SWI Features* converge quickly and suggests that the decision space is less complex. *T2map Generative* has seven features, but takes longer to converge suggesting that the decision space is more complex. This test verifies that combination of features give an additive response and the decision space is not much more complex, leading to a classifier that requires less trees leading to less overfitting.

4. Feature Selection Evaluation

To give a better understanding of the importance and relationship of the features within the classifier training, we examined the order in which the features are selected by the classifier. In adaboost a feature is selected at each iteration, which maximizes the current ensemble classifier. Features that are selected early on will have the most discriminatory value. From Figure 33A it is apparent that the features from T2e and SWI are selected almost exclusively until iteration 60. This shows that the most important features are coming from these two channels. We expected T2e features to be used early on since the ground-truth derived from this channel.

We also expected SWI to be useful early but then not be used later as it does not contain as much texture information as the other channels. The texture features create a complex decision space leading to the need of more classifiers. This shows that SWI makes the classifier have a simpler feature space, which can lead to less overfitting and better generalizability.

Figure 33B gives another view of the feature selection. It shows the percentage of the features that are chosen at each iteration. This shows at the end of the training, most features are selected virtually evenly across all the channels.

Each of these classifiers is also weighted giving an importance to the selected feature. Figure 33C shows the weighted percentage of the features selected at each iteration. From this we see that the T2e features give the most information to the classifier, while the T2map and SWI features give about the same amount of information. However as was discussed before, the SWI features are shown to be more important early on in the classifier.

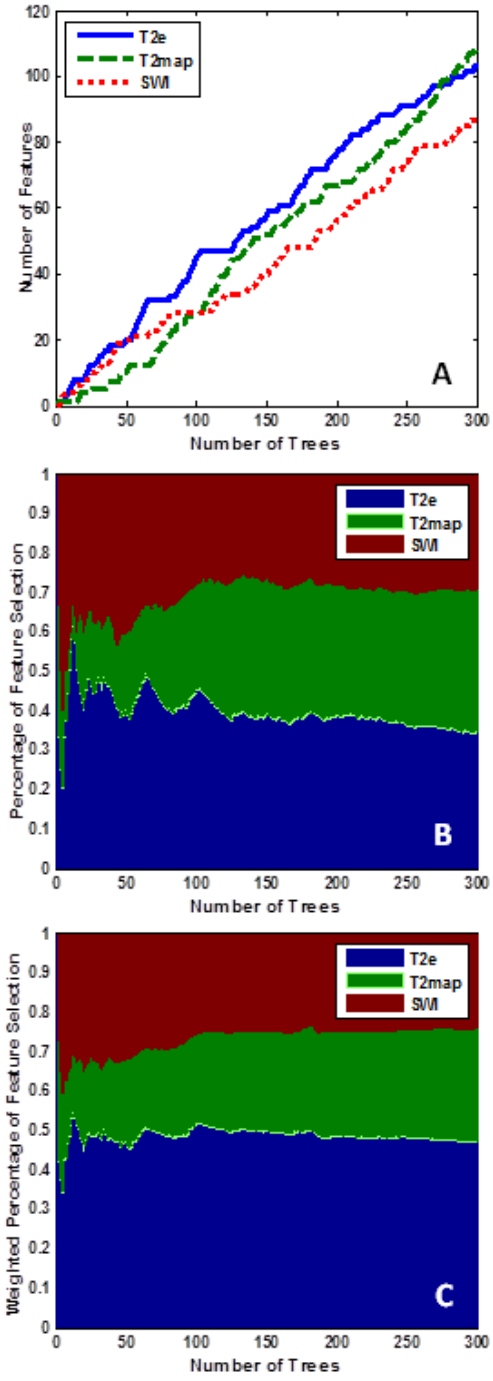

Figure 33:A: Number of features chosen from each channel at each iteration. B: Percentage of features chosen from each channel at each iteration. C: Weighted percentage of features chosen from each channel at each iteration. The legends are defined as follows: T2e features from the T2e channel only, T2map features from the T2map channel only, and SWI features from the SWI channel only.

5. Temporal Relevance of SWI Features

From looking at SWI images (Figure 26), initially we expect the performance to be extremely high due to its sensitivity to micro bleeds. However, we see in Figure 30 that SWI alone does not give the best performance. We investigated this further to identify the underlying reasons for this finding. The composition of the lesion changes over time [45]. At the early time points the lesion is comprised mostly of edema, and at later time point the lesion is almost exclusively blood. Based on this we grouped the test images by time into three time points: acute, sub-acute, and chronic. The resulting dice scores are shown in Figure 34. As expected the chronic time point shows a significant increase in performance over the other two time points, when the lesion is comprised of blood.

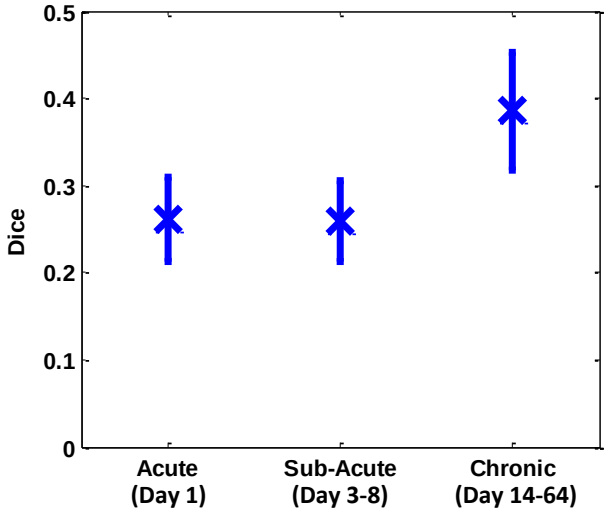

Figure 34: this depicts the dice scores at the major time points for the SWI features alone. The X shows the mean and the bars show one standard deviation. This shows that SWI performs best at the later time points where the lesion is made almost entirely of blood.

D. Conclusions

A multi-modal approach for mTBI lesion segmentation is presented, where features are chosen to take advantage of the specific attributes from each modality. In the T2map channel a temporal tissue model is proposed while in the SWI channel MAPED features are proposed. The combination of these features was shown to produce the best dice score when compared to other algorithms. Further analysis of the classifier training revealed that the SWI feature was important in obtaining generalizable results and allows for quick computation due to fewer weak classifiers being needed. SWI was also shown to be the most important feature in later time points, since the lesion is comprised mostly of blood. Future work includes post processing steps to take advantage of spatial arrangement of voxels.

CHAPTER 6 CONCLUSIONS

Mild traumatic brain injury is just starting to be understood. The progression of the disease and its underlying mechanisms are complex. Tools have been presented here to detect the lesions caused my mTBI. As with any biological process the use of tools for understanding the system is essential. Several technics have been presented that improve the detection of such lesions. When the lesion can be accurately and rapidly assessed, treatment routines can be accurately determined and lead to diminished long term effects.

Context has been shown to be an important factor when lesions are of a low contrast nature. Incorporating high level information helps to improve visual understanding. In Chapter 3 an approach was presented that incorporates visual and contextual knowledge to detect mTBI lesions using unimodal MRI, specifically T2 maps. Texture features are used with Probabilistic SVM to create a visual model of the mTBI appearance on the T2 MRI maps. The contextual model utilizes a Bayesian network which allows for exact or partial contextual inputs to be given to the system. Qualitative and quantitative results of the algorithm were shown using a novel rat mTBI dataset.

Further low level context has been shown to improve detection through the localization and spatial description of the lesions based on an initial estimate of the injury from the visual model. In Chapter 4 a fully automated method for detecting mTBI lesions that integrates low-level static and dynamic context is proposed. Three novel features are proposed that describe the posterior probability of classifier outputs in a cascade. These features were shown to have good qualitative and quantitative results on an mTBI dataset. The proposed approach outperformed the original autocontext, by being able to

generalize and integrate dynamic information. The context features were shown to be the most important features in the training process. Adding the dynamic context features lead to fast convergence during training leading a smaller number of trees being needed. The generality and flexibility of the proposed approach allows it to be applied to other brain lesion problems.

Finally, the use of multiple MRI modalities is explored where the advantages of each modality is evaluated to give fourth features that allow for cooperation between the modalities. In Chapter 5 A multi-modal approach for mTBI lesion segmentation is presented, where features are chosen to take advantage of the specific attributes from each modality. In the T2map channel a temporal tissue model is proposed while in the SWI channel MAPED features are proposed. The combination of these features was shown to produce the best dice score when compared to other algorithms. Further analysis of the classifier training revealed that the SWI feature was important in obtaining generalizable results and allows for quick computation due to fewer weak classifiers being needed. SWI was also shown to be the most important feature in later time points, since the lesion is comprised mostly of blood.

Future directions for this research would include the integration of DTI and DWI MRI modalities. This would lead to more information for the detection system and lead to a better description of the underlying tissue. Another avenue for research is adapting these models to humans. Many aspects of the proposed approaches will directly translate to human analysis, but there are some modeling issues that will have to be overcome. Improvement of evaluation methods should also be considered. Direct analysis of the

imaging detection with histology would be ideal. Other visual evaluation methods should also be considered.

REFERENCES

[1] T. Morris. “Traumatic Brain Injury.” in Handbook of Medical Neuropsychology, C. L. Armstrong and L. Morrow, Eds. Springer New York, 2010, ch. 2, pp. 17-32.

[2] (2002) Center for Disease Control and Prevention, Heads Up: facts for physicians about mild traumatic brain injury (MTBI) [Online]. Available: http://www.cdc.gov/ncipc/tbi/TBI.htm

[3] R. C. Cantu, “Second-Impact Syndrome,” Clinics in Sports Medicine, 17(1): 37-44, 1998.

[4] D. Cifu, et al. “VA/DoD Clincial practice guideline for management of concussion/mild traumatic brain injury.” Journal of Rehabilitation Research and Development, vol. 46, no. 6, pp. CP1-68, 2009.

[5] N. C. Colgan, M. M. Cronin, O. L. Gobbo, S. M. O'Mara, W. T. O'Connor, and M. D. Gilchrist. “Quantitative MRI Analysis of Brain Volume Changes due to Controlled Cortical Impact.” Neurotrauma, vol. 27, no. 7, pp.1265-1274, July 2010.

[6] A. Obenaus, M. Robbins, G. Blanco, N. R. Galloway, E. Snissarenko, E. Gillard, S. Lee, and M. Currás-Collazo. “Multi-Modal Magnetic Resonance Imaging Alterations in Two Rat Models of Mild Neurotrauma.” Neurotrauma, vol. 24, no.7, pp.1147-1160, July 2007.

[7] G. P. Mazzara, R. P. Velthuizen, J. L. Pearlman, H. M. Greenberg, and H. Wagner, “Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation.” International Journal of Radiation Oncology Biology Physics, vol. 59, no. 1, pp. 300-312, May 2004.

[8] T. Niimi, K. Imai, H. Maeda, and M. Ikeda. “Information loss in visual assessments of medical images.” European Journal of Radiology, vol. 61, no. 2, pp. 362-366, February 2007.

[9] Z. Metting, L. A. Rodiger, J. De Keyser, and J. Van Der Naalt. “Structural and functional neuroimaging in mild-to-moderate head injury.” The Lancet Neurology, vol. 6, no.8, pp.699-710, August 2007.

[10] K.A.Tong, “New MRI techniques for imaging of head trauma: DWI, MRS, SWI.” Applied Radiology, vol. 32, no. 7, pp. 29:34, July 2003.

[11] I. Kharatishvili, A. Sierra, R. J. Immonen, O. H. J. Gröhn, and A. Pitkänen, “Quantitative T2 mapping as a potential marker for the initial assessment of the severity of damage after traumatic brain injury in rat,” Experimental Neurology, vol. 217, no. 1, pp. 154-164, May 2009.

[12] J. Rajan, D. Poot, J. Juntu, and J. Sijbers, “Noise measurement from magnitude MRI using local estimates of variance and skewness,” Physics in Medicine and Biology, vol. 55, no. 16, pp N441, August 2010.

[13] S.N. Niogi, P. Mukherjee, “ Diffusion tensor imaging of mild traumatic brain injury.” J Head Trauma Rehabil., vol. 25, no. 4, pp. 241-255, July 2010.

[14] N. Ghosh, Y. Sun, B. Bhanu, S. Ashwal, A. Obenaus, “Automated detection of brain abnormalities in neonatal hypoxia ischemic injury from MR images,” Medical Image Analysis, 2014.

[15] E.M. Haacke, S. Mittal, Z. Wu, J. Neelavalli, Y.C.N. Cheng, “ Susceptibility-Weighted Imaging: Technical Aspects and Clinical Applications, Part 1.” American Journal of Neuroradiology, vol. 30, no. 1, pp. 19-30, January 2009.

[16] K.A. Tong , S. Ashwal , B.A. Holshouser BA, et al., “Diffuse axonal injury in children: clinical correlation with hemorrhagic lesions.” Ann Neurol, vol. 56, no. 1, pp. 36-50, July 2004.

[17] A. Fazlollahi, F. Meriaudeau, V.L. Villemagne, C.C. Rowe, P.M. Desmond, P.A. Yates, and O. Salvado, “ Automatic detection of small spherical lesions using multiscale approach in 3D medical images,” Image Processing (ICIP), 2013 20th IEEE International Conference on , pp.1158,1162, 15-18 Sept. 2013.

[18] E.D. Bigler, W.L. Maxwell, “ Neuropathology of mild traumatic brain injury: relationship to neuroimaging findings.” Brain Imaging Behav., vol. 6, no. 2, pp. 108-136, June 2012.

[19] A. Irimia, M. C. Chambers, J. R. Alger, M. Filippou, M. W. Prastawa, B. Wang, D. A. Hovda, G. Gerig, A. W. Toga, R. Kikinis, P. M. Vespa and J. D. Van Horn. “Comparison of Acute and Chronic Traumatic Brain Injury Using Semi-Automatic Multimodal Segmentation of MR Volumes.” Journal of Neurotrauma, vol. 28, no. 11, pp. 2287-2306, November 2011.

[20] R. de Boer, H.A. Vrooman, F. Van Der Lijn, M.W. Vernooij, M.A. Ikram, A. Van Der Lugt, M.M.B. Breteler and W.J. Niessen. “White matter lesion extension to automatic brain tissue segmentation on MRI.” NeuroImage, vol. 45, no. 4, pp. 1151-1161, May 2009.

[21] J. M. Fitzpatrick, J. B. West and J. Maurer C.R. “Predicting error in rigid-body point-based registration.” IEEE Trans. Med. Imag., vol. 17, no. 5, pp. 694-702, 1998.

[22] P. Thevenaz, T. Blu and M. Unser. “Interpolation revisited [medical images application.” IEEE Trans. Med. Imag., vol. 19, no. 7, pp. 739-758, July 2000.

[23] F. Kruggel, J. S. Paul, and H.-J. Gertz. “Texture-based segmentation of diffuse lesions of the brain’s white matter.” NeuroImage, vol. 39, no. 3, pp. 987-996, February 2008.

[24] S. Ahmed, K. M. Iftekharuddin and A. Vossough. “Efficacy of Texture, Shape, and Intensity Feature Fusion for Posterior-Fossa Tumor Segmentation in MRI.” IEEE Trans. Inf. Technol. Biomed., vol. 15, no. 2, pp. 206-213, March 2011.

[25] K. Holli, L. Harrison, P. Dastidar, M. Waljas, S. Liimatainen, T. Luukkaala, J. Ohman, S. Soimakallio and H. Eskola. “Texture analysis of MR images of patients with Mild Traumatic Brain Injury.” BioMed Central Medical Imaging, vol.10, no.8, May 2010.

[26] O. Marques, E. Barenholtz and V. Charvillat. “Context Modeling in Computer Vision: Techniques, Implications, and Applications.” Multimedia Tools and Applications, vol. 51, no. 1, pp. 303-339, January 2011.

[27] C. Galleguillos and S. Belongie. “Context based object categorization: A critical survey.” Computer Vision and Image Understanding, vol. 114, no. 6, pp. 712-722, Jun. 2010.

[28] S. K. Divvala, D. Hoiem, J. H. Hays, A. A. Efros and M. Hebert. “An empirical study of context in object detection,” in Proc. IEEE Cong. Computer Vision and Pattern Recognition, Miami FL, 2009, pp. 1271-1278.

[29] M. J. Choi, A. Torralba and A. S. Willsky. “A Tree-Based Context Model for Object Recognition.” IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 2, pp. 240-252, February 2012.

[30] A. Rabinovich, A. Vedaldi, C. Galleguillos, E. Wiewiora and S. Belongie. “Objects in Context.” Proc. IEEE Int’l Conf. Computer Vision, Rio de Janeiro, Brazil 2007.

[31] C. Galleguillos, A. Rabinovich, and S. Belongie. “Object categorization using co-occurrence, location and appearance.” Proc. IEEE Cong. Computer Vision and Pattern Recognition, Anchorage, Alaska, June 2008, pp. 1-8.

[32] R. Perko and A. Leonardis. “A framework for visual-context-aware object detection in still images.” Computer Vision and Image Understanding, vol. 114, no. 6, pp.700-711, June 2010.

[33] A. Torralba, K. P. Murphy and W. T. Freeman, “Using the forest to see the trees: exploiting context for visual object detection and localization.” Communications of the ACM, vol. 53, no. 3, pp.107-114, March 2010.

[34] Z. Tu and X. Bai, “Auto-Context and Its Application to High-Level Vision Tasks and 3D Brain Image Segmentation.” IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 10, pp. 1744-1757, October 2010.

[35] M. A. Horsfield, R. Bakshi, M. Rovaris, M. A. Rocca, V. S. R. Dandamudi, P. Valsasina, E. Judica, F. Lucchini, C. R. G. Guttmann, M. Pia Sormani and M. Filippi. “Incorporating Domain Knowledge Into the Fuzzy Connectedness Framework: Application to Brain Lesion Volume Estimation in Multiple Sclerosis.” IEEE Trans. Med. Imag., vol. 26, no. 12, pp. 1670-1680, December 2007.

[36] A. Bianchi, B. Bhanu, V. Donovan and A. Obenaus, “Contextual and Visual Modeling for Detection of Mild Traumatic Brain Injury in MRI.” IEEE International Conference on Image Processing, Orlando, FL, October 2012.

[37] J. Lafferty, A. McCallum, F. Pereira, “Conditional random fields: probabilistic models for segmenting and labeling sequence data,” Proc. 10th Int’l Conf. Machine Learning, pp. 282-289, 2001.

[38] A. Bianchi, B. Bhanu, V. Donovan, “Visual and contextual modeling for the detection of repeated mild traumatic brain injury,” IEEE Transactions on Medical Imaging, vol. 33, no. 1, pp.1-12, January 2014.

[39] P. Anbeek, K. L. Vincken, M. J. P. Van Osch, R. H. C. Bisschops and J. Van Der Grond. “Automatic segmentation of different-sized white matter lesions by voxel probability estimation.” Medical Image Analysis, vol. 8, no. 3, pp. 205-215, September 2004.

[40] N. Zhang, S. Ruan, S. Lebonvallet, Q. Liao and Y. Zhu. “Kernel feature selection to fuse multi-spectral MRI images for brain tumor segmentation.” Computer Vision and Image Understanding, vol. 115, no. 2, pp. 256-269, February 2011.

[41] Y. Sun, B. Bhanu and S. Bhanu. "Automatic symmetry-integrated brain injury detection in MRI sequences," IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis, June 20, 2009, held in conjunction with IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, June 20-25, 2009.

[42] M. L. Seghier, A. Ramlackhansingh, J. Crinion, A. P. Leff and C. J. Price. “Lesion identification using unified segmentation-normalisation models and fuzzy clustering.” NeuroImage, vol. 41, no. 4, pp. 1253-1266, July 2008.

[43] A. Montillo, J. Shotton, J. Winn, J. E. Iglesias, D. Metaxas, A. Criminisi, “Entangled decision forests and their application for semantic segmentation of CT images,” Information Processing in Medical Imaging, vol. 6801, pp. 184-196, 2011.

[44] K.A. Tong, S. Ashwal, A. Obenaus, J.P. Nickerson, D. Kido, E.M. Haacke, " Susceptibility-weighted MR imaging: a review of clinical applications in children," AJNR Am J Neuroradiol, vol. 29, no. 1, pp. 9-17, October 2007.

[45] V. Donovan, A. Bianchi, R. Hartman, B. Bhanu, M.J. Carson, and A. Obenaus. "Computational analysis reveals increased blood deposition following repeated mild traumatic brain injury." NeuroImage: Clinical, vol. 1, no. 1, pp.18–28, 2012.

[46] A. Bianchi, B. Bhanu, V. Donovan, A. Obenaus, "Detecting mild traumatic brain injury using dynamic low level context," IEEE International Conference on Image Processing, Melbourne, Australia, Sept. 15-18, 2013.

[47] M. Kafai and B. Bhanu. "Dynamic Bayesian Networks for Vehicle Classification in Video." IEEE Trans. Ind. Informat., vol. 8, no. 1, pp. 100-109, February 2012.

[48] Vagnozzi, R., Tavazzi, B., Signoretti, S., Amorini, A.M., Belli, A., Cimatti, M., Delfini, R., Di Pietro, V., Finocchiaro, A., Lazzarino, G. "Temporal window of metabolic brain vulnerability to concussions: mitochondrial-related impairment--part I." Neurosurgery, vol. 61, no. 2, pp.379-388, August 2007.

[49] A. K. Laird, "Dynamics of Tumor Growth," British Journal of Cancer, vol. 18, no. 3, pp. 490-502, September 1964.

[50] Longhi, L., Saatman, K.E., Fujimoto, S., Raghupathi, R., Meaney, D.F., Davis, J., McMillan, B.S.A., Conte, V., Laurer, H.L., Stein, S., Stocchetti, N., McIntosh, T.K. "Temporal window of vulnerability to repetitive experimental concussive brain injury." Neurosurgery vol. 56, no. 2, pp. 364-374, February 2005.

[51] L. F. Costa and R. M. Cesar, "Shape Analysis and Classification: Theory and Practice." Boca Raton, FL: CRC, 2001.

[52] T.F. Coleman and Y. Li, "An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds." Society for Industrial and Applied Mathematics Journal on Optimization, vol. 6, no. 2, pp.418–445, July 1996.

[53] T.F. Coleman and Y. Li. "On the Convergence of Reflective Newton Methods for Large-Scale Nonlinear Minimization Subject to Bounds." Mathematical Programming, vol. 67, no. 2, pp. 189-224, 1994.

[54] X. Guorong, C. Peiqi and W. Minhui. "Bhattacharyya distance feature selection," Proc. Int'l Conf. Pattern Recognition, Vienna, Austria 1996, pp.195 -199.

[55] C. C. Chang and C.J. Lin. "LIBSVM: A library for support vector machines." ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 3, pp. 1-39, 2011.

[56] J. C. Platt. “Probabilistic outputs for support vector machines and comparison to regularized likelihood methods.” In Advances in Large Margin Classifiers, Cambridge, MA: MIT Press, 2000, pp. 61-74.

[57] N. A. Obuchowski. “Receiver Operating Characteristic Curves and Their Use in Radiology.” Radiology, vol. 229, no. 1, pp. 3-8, 2003.

[58] T. Ojala, M. Pietikainen, T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with Local Binary Patterns,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp. 971-987, August 2002.

[59] B. S. Manjunath and W.Y. Ma, "Texture features for browsing and retrieval of image data", IEEE Trans. Pattern Anal. Mach. Intell., vol.18, no.8, pp.837-842, August 1996.

[60] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” Proc. IEEE Cong. Computer Vision and Pattern Recognition, San Diego, California, June 2005, pp.886-893.

[61] Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting.” J. of Computer and System Sciences, vol. 55, nol. 1, pp. 119-139, 1997.

[62] P. Viola and M. J. Jones, "Rapid object detection using a boosted cascade of simple features", Proc. IEEE Cong. Computer Vision and Pattern Recognition, Kauai, Hawaii, December 2001, pp.511–518.

[63] B. Menze, K. Van Leemput, D. Lashkari, M. Weber, N. Ayache, P. Golland, “Segmenting glioma in multi-modal images using a generative-discriminative model for brain lesion segmentation.”, Proc. MICCAI BraTS Challenge, Nice, France 2012.

[64] E. Geremia, B.H. Menze, N. Ayache, “Spatial decision forest for glioma segmentation in multi-channel MRI.” Proc.MICCAI BraTS Challenge, Nice, France 2012.

[65] D. Zikic, B. Glocker, E. Konukoglu, J. Shotton, A. Criminisi, D.H. Ye, C. Demiralp, O.M. Thomas, T. Das, R. Jena, S.J. Price, “Context-sensitive classification forests for segmentation of brain tumor tissues,” Proc. MICCAI, Nice, France, 2012.

[66] A. Bianchi, J.V. Miller, E.T. Tan, A. Montillo, “Brain tumor segmentation with symmetric texture and symmetric intensity-based decision forests,” IEEE International Symposium on Biomedical Imaging, San Francisco, California, April 2013.

[67] F. Maes, D. Vandermeulen, P. Suetens, " Medical image registration using mutual information," Proceeding of the IEEE, vol. 91, no. 10, pp. 1699-1722, October 2003.

[68] C.C. Lip, D.A. Ramli, "Comparative study on feature, score and decision level fusion schemes for robust multibiometric systems," in Frontiers in Computer Education. Springer Berlin Heidelberg, 2012, vol. 133, pp. 941-948.

[69] J.R. Movellan, " Tutorial on Gabor Filter", Retrieved 2014-2-10, http://mplab.ucsd.edu/tutorials/gabor.pdf

[70] H. Greenspan, A. Ruf, J. Goldberger, "Constrained Gaussian mixture model framework for automatic segmentation of MR brain images," IEEE Trans. Med. Img., vol. 25, no. 9, pp. 1233-1245, September 2006.

[71] S. Ruan, C. Jaggi, J. Xue, J. Fadili, D. Bloyet, " Brain tissue classification of magnetic resonance images using partial volume modeling." IEEE Trans. Med. Img., vol. 19, no. 12, pp. 1179-1187, August 2002.

[72] McLachlan, G., and D. Peel. Finite Mixture Models. Hoboken, NJ: John Wiley & Sons, Inc., 2000.

[73] C. Cortes and V. Vapnik. "Support-Vector Networks." Machine Learning, vol. 20, no. 3, pp. 273-297, September 1995.

[74] L. Breiman. "Random Forest." Machine Learning, vol. 45, no. 1, pp. 5-32, October 2001.

[75] G. Huang, Q. Zhu, and C. Siew. "Extreme Learning Machine: Theory and Applications." Neurocomputing, vol. 70, no. 1, pp. 489-501, December 2006.